Structural-constrained Methods for the Identification of False Data Injection Attacks in Power Systems

Gal Morgenstern, *Student Member, IEEE* and Tirza Routtenberg, *Senior Member, IEEE*

Abstract—Power system functionality is determined on the basis of the power system state estimation (PSSE). Thus, corruption of the PSSE may lead to severe consequences, such as financial losses, maintenance damage, and disruptions in electricity distribution. Classical bad data detection (BDD) methods, developed to ensure PSSE reliability, are unable to detect well-designed attacks, named unobservable false data injection (FDI) attacks. In this paper, we develop novel structural-constrained methods for the detection of unobservable FDI attacks, the identification of the attacked buses' locations, and PSSE under the presence of such attacks. The proposed methods are based on formulating structural, sparse constraints on both the attack and the system loads. First, we exploit these constraints in order to compose an appropriate model selection problem. Then, we develop the associated generalized information criterion (GIC) for this problem. However, for large networks, the GIC method's computational complexity grows exponentially with the network size. Thus, based on the proposed structural and sparse constraints, we develop two novel low-complexity methods for unobservable FDI attack identification: 1) a modification of the state-of-the-art orthogonal matching pursuit (OMP); and 2) a method that utilizes the graph Markovian property in power systems, i.e. the second-neighbor relationship between the power data at the system's buses. The methods' performance is evaluated on a IEEE-30 bus test case system.

Index Terms—unobservable false data injection (FDI) attack, power system state estimation (PSSE), attack detection and identification, structural constraints, model selection, graph Markovian property

I. Introduction

MODERN electrical grids are monitored by energy management systems (EMSs). The EMS evaluates the power system state estimation (PSSE) for multiple monitoring purposes, including stability assessment, control, and security [1], [2]. To ensure the reliability of the measurements, residual-based bad data detection (BDD) methods have been integrated into the EMS [1]. However, BDD methods cannot detect well-designed attacks, named unobservable false data injection (FDI) attacks [3], [4]. Unobservable FDI attacks are achieved by manipulating measurements based on the power network topology [3], where the topology matrix is either known or can be estimated from historical data [5]–[9]. Unobservable FDI attacks may inflict severe damage by influencing the PSSE [10]–[14]. Therefore, appropriate tools for the detection, identification, and estimation of these attacks, that ensure the reliability of the PSSE, are essential for obtaining high power quality and maintaining stable power system operation.

The problem of detecting unobservable FDI attacks has been considered in the last decade. Methods that strategically protect a basic set of measurements were proposed in [15]–[19]. In the same context, the use of synchronized phasor measurement units (PMUs) has been suggested [19]–[21]. However, these approaches require integrating additional hardware into the power system, which results in high cost, a long installation period, and additional security vulnerabilities [22], [23]. Detection by the moving target defense technique, in which the system configuration is actively changed, has been proposed in [24]–[27]. However, in these methods, effective attack detection carries the cost of operating far from the optimal state and, thereby, incurs economic losses [28]. Detection methods based on machine learning and data mining have been suggested in [29], [30], but these methods require a large set of historic power system data, which is usually unavailable. Detection based on load forecasting has been suggested in [31] and [32], yet obtaining a reliable forecast is not always possible and may require extensive resources [33]. In conclusion, all the above methods require additional hardware, or a large amount of historical data, or downgrade the system operation, and, thus, cannot serve as practical solutions for the detection of unobservable FDI attacks in power systems. Moreover, these methods do not provide attack identification, i.e. localization of the attacked buses, and PSSE evaluation in the presence of this attack.

In this paper, we suggest a novel compressive sensing (CS) parametric approach, based on structural constraints, for the detection of unobservable FDI attacks, the identification of the attacked buses' locations, and PSSE under the presence of such attacks. Sparse recovery CS techniques have become a foremost research area in the last two decades (see e.g. [34]–[36] and references therein). In power systems, CS algorithms [37] have been proposed in the context of multi-line outage identification [38], [39], gross error identification [40], identification of imbalances [41], and BDD [42]. CS algorithms that exploit the sparsity of the unobservable FDI attack, as well as the temporal correlations of the power measurements, have been developed in [43]–[45] for detection of unobservable FDI attacks. In addition to attack detection, CS methods also enable the identification of the attacked buses. However, the methods in [43]–[45] require multi-time measurements and do not consider the difference between the structure of the unobservable FDI attack and the system measurements.

In this paper, we formulate structural and sparse constraints for both the attack and the change in the power system loads between two consecutive time samples. Accordingly, we formulate a model selection problem, where in each model the attack is assumed to be represented by a different set of dictionary elements, i.e. the topology matrix columns. Based on the model selection formulation and the generalized

information criterion (GIC) selection rule [46], we develop a structural-constrained method for the detection, identification, and estimation of unobservable FDI attacks. As a by-product, the PSSE can be evaluated on 'clean measurements', which are obtained by removing the estimated attack from the measurements. However, the proposed GIC-based method requires an exhaustive search, for which the required computer resources grow exponentially. Thus, based on the proposed structural and sparse constraints, we develop two novel low-complexity methods for unobservable FDI attack identification: 1) a modification of the state-of-the-art orthogonal matching pursuit (OMP) [47]; and 2) a method that is based on the graph Markovian property in power systems, i.e. second-neighbor relationship [48]. Finally, we show by numerical simulation that the proposed GIC, OMP, and Graph Markovian GIC (GM-GIC) methods outperform the detection and identification performance of existing methods. In particular, the proposed *parametric* methods that are based on information criteria outperform our previous *non-parametric* detection methods in [49], [50] that are heuristic in nature and are based on the Graph Fourier Transform (GFT).

Therefore, the main contributions of this paper are threefold: First, we present a new model that takes into account both the physical, statistical, and structural properties of unobservable FDI attacks and the typical load changes in power systems. Second, we propose a model-selection-based approach that enables the identification of unobservable FDI attacks by using information criteria and proving analytically that the nuisance parameters that stemmed from the typical load changes can be neglected. Finally, we propose an OMP based method and the novel low-complexity GM-GIC method that utilizes both the sparsity and the graphical properties of the problem in order to reduce complexity, while preserving high capabilities for the identification of unobservable FDI attacks. Moreover, GM-GIC is a theoretical method that can be used in the general context of sparse recovery of graph signals and is not limited to power system applications.

The remainder of this paper is organized as follows. In Section II we introduce the background on the model of the power flow equations, BDD, and unobservable FDI attacks. The proposed GIC-based approach for the detection and identification of unobservable FDI attacks is presented in Section III, while the two low-complexity methods are developed in Section IV. In Section V a simulation study is presented. Finally, the paper ends with conclusions in Section VI.

In this paper vectors are denoted by boldface lowercase letters and matrices are denoted by boldface uppercase letters. The operators $||\cdot||$ and $||\cdot||_0$ denote the Euclidean norm and the zero seminorm, respectively, where the latter specifies the number of non-zero elements in the vector. The operators $(\cdot)^T$ and $(\cdot)^{-1}$ denote the transpose and inverse operators, respectively. The projection and the orthogonal projection operators for a matrix $\mathbf{A} \in \mathbb{R}^{M \times K}$ with a full column rank are given by $\mathbf{P_A} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ and $\mathbf{P_A^\perp} = \mathbf{I} - \mathbf{P_A}$, respectively. The linear space spanned by the $\mathbf{A}$ matrix columns is denoted by col($\mathbf{A}$). For an index set, $\Lambda \subset \{1, \ldots, L\}$, $\boldsymbol{\theta}_\Lambda$ is the $|\Lambda|$-dimensional subvector of $\boldsymbol{\theta}$ containing the elements indexed by Λ, where $|\Lambda|$ denotes the set's cardinality. For any index sets, Λ_1 and Λ_2, $\mathbf{A}_{\Lambda_1,\Lambda_2}$ is the submatrix composed by the columns and rows of $\mathbf{A}$ associated with Λ_1 and Λ_2, respectively.

II. BACKGROUND

A. DC model

The power system network can be defined by a set of buses, $\mathcal{V} = \{0, 1, \ldots N\}$, and a set of transmission lines, $\mathcal{E}$, where the transmission line connecting buses $n \in \mathcal{V}$ and $k \in \mathcal{V}$ is denoted by the pair $(n, k) \in \mathcal{E}$. Furthermore, according to the π-model of transmission lines [1], each line, $(n, k) \in \mathcal{E}$, is characterized by an admittance, $Y_{n,k}$. Based on the above, the DC model, which is a linearized representation of the power measurements in the power system, is considered with the following assumptions:

a.1 Branches are without resistance losses. As a result, $Y_{n,k} = B_{n,k}$, where $B_{n,k}$ is the susceptance of the (n,k) transmission line.

a.2 The bus voltage magnitudes, V_n, $n = 1, 2, \ldots, N$, are approximated by 1 per unit (p.u.).

a.3 The differences in bus voltage angle between adjacent buses are assumed to be small such that

$$\sin(\theta_n - \theta_k) \approx \theta_n - \theta_k, \ \forall (n.k) \in \mathcal{E}.$$

Based on Assumptions **a.1-a.3**, the DC model of the active power injections in the buses is given by [1]:

$$P_n = \sum_{k:(n,k)\in\mathcal{E}} P_{n,k} = \sum_{k:(n,k)\in\mathcal{E}} B_{n,k}(\theta_n - \theta_k), \quad (1)$$

$\forall \, n \in \mathcal{V}$. In similar manner, the DC model of the active power flows in the transmission lines is given by

$$P_{n,k} = B_{n,k}(\theta_n - \theta_k), \quad \forall \, (n, k) \in \mathcal{E}. \quad (2)$$

In this paper we assume active power measurements, denoted by $\mathbf{z} = [z_1, \cdots, z_M]^T \in \mathbb{R}^M$, that include the power injections in (1) and the power flows in (2). In addition, $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]^T$ denotes the state vector and includes the bus voltage angles at all buses except the slack bus. Accordingly, the matrix-form noisy and attacked DC power flow model can be written as

$$\mathbf{z} = \mathbf{H}\boldsymbol{\theta} + \mathbf{a} + \mathbf{e}, \quad (3)$$

where the topology matrix, $\mathbf{H}$, is a constant $M \times N$ Jacobian matrix, which is composed of the susceptance elements, $B_{n,k}$ (as described, for example, in [51]). In addition, $\mathbf{e} \in \mathbb{R}^M$ is a zero-mean Gaussian additive noise vector with covariance matrix $\mathbf{R}$, $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$. The attack is denoted by $\mathbf{a} \in \mathbb{R}^M$.

B. Power system state estimation (PSSE)

The PSSE, which is used for multiple monitoring purposes, is implemented by using meter measurement data from (3). The PSSE is commonly computed using the weighted least squares (WLS) estimator [1]:

$$\hat{\boldsymbol{\theta}}^{\text{WLS}} = \operatorname*{argmin}_{\boldsymbol{\theta}\in\mathbb{R}^N} (\mathbf{z} - \mathbf{H}\boldsymbol{\theta})^T\mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\boldsymbol{\theta}) = \mathbf{Kz}, \quad (4)$$

where

$$\mathbf{K} = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}. \quad (5)$$

In order to robustify the PSSE against errors, classical BDD methods are implemented in the EMS [1]. These BDD methods are based on the residual error:

$$\mathbf{r} = \mathbf{z} - \mathbf{H}\hat{\boldsymbol{\theta}}^{\text{WLS}} = (\mathbf{I} - \mathbf{HK})\mathbf{z}, \quad (6)$$

where the last equality is obtained by substituting the estimated state vector as defined in (4). The residual error from (6) is used, for example, in the *Largest Normalized Residual* $\mathbf{r}_{max}^N$-*test* and the χ^2-test [1], also referred to as the $J(x)$-detector, which is given by

$$T^{\text{BDD}} = ||\mathbf{z} - \mathbf{H}\hat{\boldsymbol{\theta}}^{\text{WLS}}||^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma^{\text{BDD}}, \tag{7}$$

where $\mathcal{H}_1$ is the hypothesis that the measurements are corrupted by bad data, which may be a result of an attack, and $\mathcal{H}_0$ is the null hypothesis. The threshold, γ^{BDD}, is determined to obtain a desired false alarm probability. However, both the Largest Normalized Residual, which is an identification method, and the χ^2-test, which is a detection method, can neither detect nor identify the presence of unobservable FDI attacks, as described in the next subsection.

C. Unobservable FDI attacks

The unobservable FDI attack satisfies

$$\mathbf{a} \overset{\triangle}{=} \mathbf{Hc}, \tag{8}$$

where $\mathbf{c} \in \mathbb{R}^N$ is an arbitrary constant vector. By substituting the well-designed attack from (8) in the observation model from (3) we obtain

$$\mathbf{z} = \mathbf{H}(\boldsymbol{\theta} + \mathbf{c}) + \mathbf{e}. \tag{9}$$

From a comparison between the models in (3) and in (9), it can be verified that by observing $\mathbf{z}$, the state vector in (3), $\boldsymbol{\theta}$, cannot be distinguished from its corrupted (attacked) version in (9), $\boldsymbol{\theta} + \mathbf{c}$, since both $\boldsymbol{\theta}$ and $\mathbf{c}$ are unknown vectors. As a result, the residual error, obtained by substituting the WLS estimation (which is based on the unobservable FDI model from (9)) in (6) is

$$\mathbf{r} = (\mathbf{I} - \mathbf{HK})\mathbf{e}, \tag{10}$$

and therefore cannot be utilized in order to indicate the presence of an unobservable FDI attack. Consequently, all residual-based methods, as well as all the other methods that are based solely on the model in (9), are expected to fail in detecting unobservable FDI attacks. Therefore, new methods for the identification of manipulated measurements that integrate additional structural and physical constraints are required.

III. STRUCTURAL-BASED GIC APPROACH FOR THE IDENTIFICATION OF UNOBSERVABLE FDI ATTACKS

Considering the unobservable FDI attack formation and further sparsity restrictions, the additional power caused by the attack is constrained to a small subset of buses in the network. Hence, in this section a new framework that is based on observing power changes in small subnetworks of the system is proposed. First, a new model for power measurements in the presence of unobservable FDI attacks is developed in Subsection III-A. Next, in Subsection III-B, we develop a novel identification method for unobservable FDI attacks based on the GIC selection rule [46]. By using the identification solution, in Subsection III-C we develop a method for PSSE in the presence of unobservable FDI attacks. Finally, extensions are presented in Subsection III-D.

A. Structural-constrained model

The proposed framework, which facilitates structural constraints on the unobservable FDI attack and on the typical demand of the system's loads. is constructed by applying the following definitions and assumptions:

A.1 Difference-based model: Similar to the models in [31] and [52], two consecutive time samples are observed, where the first is considered free from malicious attacks and the second may contain an unobservable FDI attack.

A.2 Restricted measurements: Since generator buses are heavily secured and obtain direct communication to the control center [53], we assume that generator bus measurements could not be manipulated. Additionally, it is assumed that 'zero load' measurements cannot be manipulated. Thus, the set of buses that can be attacked only includes the load buses and is denoted by $\mathcal{L}$, where $\mathcal{L} \subset \mathcal{V}$.

A.3 State sparsity: The number of manipulated state variables is considered to be significantly smaller than the cardinality of the node set, $|\mathcal{V}|$, and is bounded by the sparsity term $K_{\mathbf{c}}$. Thus, we define the set

$$\mathcal{G}_{K_{\mathbf{c}}} \overset{\triangle}{=} \{\Lambda \subset \mathcal{V} : \ 1 \le |\Lambda| \le K_{\mathbf{c}}\}, \tag{11}$$

that includes all possible supports of the state attack $\mathbf{c}$. It is shown in [3], that this assumption (also used in [45]) stems directly from the commonly-used sparsity restriction on the number of manipulated meters, which states that the attack vector, $\mathbf{a}$, is sparse (see, e.g. [51]).

A.4 Typical load changes: Power systems under normal conditions are quasi-static systems that only change slightly over a short period of time [43], [54]. Therefore, a relatively small change in the load measurements is considered. In addition, the typical load changes, w.r.t. the actual buses' load measurements at a specific moment, are determined by the consumption demand [55]. Thus, representing the typical load changes in a vectorial fashion will output a non-sparse vector that is not related to the system topology matrix.

It can be seen that Assumptions **A.1**-**A.3** define structural constraints on the attack. In difference, Assumption **A.4**, which for the best of our knowledge has not yet been used in the field of power system cyber-security, defines a structural constraint on the typical load changes.

From Assumption **A.1**, the difference-based model is achieved by taking the difference between two consecutive time samples that satisfy the model in (9) with $\mathbf{c} = \mathbf{0}$ for the first time sample and an arbitrary $\mathbf{c}$ in the second time sample. Thus, the difference-based observation model is given by

$$\Delta\mathbf{z} = \mathbf{H}(\mathbf{c} + \Delta\boldsymbol{\theta}) + \Delta\mathbf{e}, \tag{12}$$

where $\Delta\boldsymbol{\theta}$ is the change in the state vector, $\boldsymbol{\theta}$, between the two time samples. Furthermore, by assuming that the noises of the two consecutive time samples are independent, we obtain that the difference-based measurement noise in (12) is distributed according to $\Delta\mathbf{e} \sim \mathcal{N}(\mathbf{0}, 2\mathbf{R})$, where, for the sake of simplicity, it is assumed that $2\mathbf{R} = \sigma_e^2\mathbf{I}$.

According to Assumption **A.2**, only the changes in the load buses' measurements are relevant for the attack. Thus, the relevant measurement model for the detection of attacks is

obtained by taking only the load measurements from (12), which results in

$$\Delta \mathbf{z}_{\mathcal{L}} = \mathbf{H}_{\mathcal{L},\mathcal{V}}(\mathbf{c} + \Delta \boldsymbol{\theta}) + \Delta \mathbf{e}_{\mathcal{L}}, \qquad (13)$$

where $\Delta \mathbf{z}_{\mathcal{L}}$ is the change in load measurements, $\mathbf{H}_{\mathcal{L},\mathcal{V}}$ is the associated submatrix of the topology matrix $\mathbf{H}$, and $\Delta \mathbf{e}_{\mathcal{L}}$ is the corresponding noise.

Assumption **A.3** implies that the state attack vector support, Λ, is in the set $\mathcal{G}_{K_{\mathbf{c}}}$. Therefore, there exists an element $\Lambda_i \in \mathcal{G}_{K_{\mathbf{c}}}$, where $i = 1, 2, \cdots, |\mathcal{G}_{K_{\mathbf{c}}}|$, such that the attack on the load measurements is given by

$$\mathbf{H}_{\mathcal{L},\mathcal{V}}\mathbf{c} = \mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i}, \qquad (14)$$

where $|\Lambda_i| \ll |\mathcal{V}|$. By substituting (14) in (13) we obtain

$$\Delta \mathbf{z}_{\mathcal{L}} = \mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i} + \mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta} + \Delta \mathbf{e}_{\mathcal{L}}. \qquad (15)$$

Furthermore, Assumption **A.4**, implies that the power of the unattacked load changes without noise, $\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta}$, is bounded. This assumption can be mathematically formulated as follows:

$$||\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta}||^2 < \eta, \qquad (16)$$

where η is a system tuning parameter that can be obtained from the power system statistics. In addition, from Assumption **A.4**, $\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta}$ cannot be represented as a sparse combination of the submatrix, $\mathbf{H}_{\mathcal{L},\Lambda_i}$ columns, for any $\Lambda_i \in \mathcal{G}_{K_{\mathbf{c}}}$. Therefore, it is assumed that the projection of $\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta}$ onto the column space of the submatrix, $\mathbf{H}_{\mathcal{L},\Lambda_i}$, results in a significant loss in power, as follows:

$$||\mathbf{P}_{\Lambda_i}\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta}||^2 \ll ||\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta}||^2, \qquad (17)$$

where

$$\mathbf{P}_{\Lambda_i} \triangleq \mathbf{H}_{\mathcal{L},\Lambda_i}(\mathbf{H}_{\mathcal{L},\Lambda_i}^T \mathbf{H}_{\mathcal{L},\Lambda_i})^{-1}\mathbf{H}_{\mathcal{L},\Lambda_i}^T \qquad (18)$$

is the projection matrix to the space $\mathrm{col}(\mathbf{H}_{\mathcal{L},\Lambda_i})$. The combination of the inequalities in (16) and (17) implies that

$$||\mathbf{P}_{\Lambda_i}\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta}||^2 \ll \eta. \qquad (19)$$

The inequality shows that the projection of $\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta}$ onto $\mathrm{col}(\mathbf{H}_{\mathcal{L},\Lambda_i})$ is small w.r.t. the network parameters. In contrast, by using the properties of projection matrices it can be verified that the projection of $\mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i}$ onto the space of the attack support remains unchanged, i.e.

$$\mathbf{P}_{\Lambda_i}\mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i} = \mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i}. \qquad (20)$$

Therefore (19) and (20) imply that

$$||\mathbf{P}_{\Lambda_i}\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta}||^2 \ll ||\mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i}||^2, \qquad (21)$$

i.e. when observing the measurement space $\mathrm{col}(\mathbf{H}_{\mathcal{L},\Lambda_i})$, the attack component is considered to be significantly larger than the projected loads component. Accordingly, the projection of the measurement vector in (15) onto $\mathrm{col}(\mathbf{H}_{\mathcal{L},\Lambda_i})$ by the projection matrix in (18) satisfies

$$\begin{aligned}
\mathbf{P}_{\Lambda_i}\Delta \mathbf{z}_{\mathcal{L}} &= \mathbf{P}_{\Lambda_i}(\mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i} + \mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta} + \Delta \mathbf{e}_{\mathcal{L}}) \\
&= \mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i} + \mathbf{P}_{\Lambda_i}(\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta} + \Delta \mathbf{e}_{\mathcal{L}}) \\
&\approx \mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i} + \mathbf{P}_{\Lambda_i}\Delta \mathbf{e}_{\mathcal{L}}, \quad (22)
\end{aligned}$$

where the second equality is obtained by substituting (20), and the last approximation is based on (21). It can be seen that in the new, approximated model in (22), the attack component,

$\mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i}$, can be detected and is no longer unobservable, as it was in (9). Moreover, the new approximated model in (22) can be obtained directly by enforcing

$$\mathbf{P}_{\Lambda_i}\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta} \approx \mathbf{0}, \quad \forall \Lambda_i \in \mathcal{G}_{K_{\mathbf{c}}}. \qquad (23)$$

While he approximation in (23) is not assumed in the first place, it can now be interpreted as a result of applying Assumptions **A.1**-**A.4** on the unobservable FDI attack model in (9). Thus, the Approximations introduced in (22)-(23) can be leveraged for the identification of unobservable FDI attacks and, accordingly, are used as basis for the proposed methods.

B. Identification of unobservable FDI attacks by the GIC approach

Based on the model from (15) and the result from (23), the task of identifying the subset of attacked buses, Λ_i, is performed by selecting from amongst the multiple hypotheses:

$$\begin{aligned}
\mathcal{H}_0 : & \quad \Delta \mathbf{z}_{\mathcal{L}} = \mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta} + \Delta \mathbf{e}_{\mathcal{L}} \\
\mathcal{H}_i : & \quad \Delta \mathbf{z}_{\mathcal{L}} = \mathbf{H}_{\mathcal{L},\Lambda_i}\mathbf{c}_{\Lambda_i} + \mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta} + \Delta \mathbf{e}_{\mathcal{L}} \quad (24) \\
& \quad \text{s.t. } \mathbf{P}_{\Lambda_i}\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta \boldsymbol{\theta} = \mathbf{0},
\end{aligned}$$

where $i = 1, \ldots, |\mathcal{G}_{K_{\mathbf{c}}}|$, and each hypothesis $\mathcal{H}_i$ in (24) assumes a different support $\Lambda_i \in \mathcal{G}_{K_{\mathbf{c}}}$ for the state attack vector $\mathbf{c}$. The null hypothesis $\mathcal{H}_0$ is obtained by substituting $\mathbf{c} = \mathbf{0}$ in (13).

In the following, we derive the identification of unobservable FDI attacks by selecting the most likely hypothesis in (24), i.e. the most suitable choice of attack support, Λ_i, from the set of candidate supports, $\{\mathcal{G}_{K_{\mathbf{c}}} \cup \emptyset\}$, given the measurement vector $\Delta \mathbf{z}_{\mathcal{L}}$. In order to solve (24), we implement the GIC selection rule [46], which is widely employed in signal and array processing. The GIC method chooses the hypothesis $\mathcal{H}_i$ which maximizes the sum of two terms: the likelihood term for data encoding, $L(\cdot)$, and a penalty function, $\tau(\cdot)$, that inhibits the number of free parameters of the model from becoming very large. For the considered model in (24) and a given difference-based state vector $\Delta \boldsymbol{\theta}$, the GIC statistic is given by

$$\mathrm{GIC}(\Lambda_i, \tau(|\Lambda_i|, |\mathcal{L}|)) \triangleq 2L(\Delta \mathbf{z}_{\mathcal{L}}; \hat{\mathbf{c}}_{\Lambda_i}^{\mathrm{ML}|i}, \Delta \boldsymbol{\theta}) - \tau(|\Lambda_i|, |\mathcal{L}|),$$
$$(25)$$

where $L(\Delta \mathbf{z}_{\mathcal{L}}; \hat{\mathbf{c}}_{\Lambda_i}^{\mathrm{ML}|i}, \Delta \boldsymbol{\theta})$ is the log-likelihood function of $\Delta \mathbf{z}$ under the ith hypothesis, which is associated with the support, Λ_i, and $\hat{\mathbf{c}}^{\mathrm{ML}|i}$ is the associated ML estimation of the state attack vector, $\mathbf{c}$. Therefore, in the general case the GIC statistic is a function of $\Delta \boldsymbol{\theta}$, which is an unknown deterministic vector. However, as presented in the following, the GIC selection rule in this case is independent of $\Delta \boldsymbol{\theta}$ and, thereby, is a valid rule. The term $\tau(|\Lambda_i|, |\mathcal{L}|)$ is a penalty function, which increases with the number of free unknown parameters that is determined in this case by the number of manipulated variables, $|\Lambda_i|$, and the number of load buses, $|\mathcal{L}|$. In particular, a special case of the GIC family is the Akaike information criterion (AIC), for which the penalty term is

$$\tau(|\Lambda_i|, |\mathcal{L}|) = 2|\Lambda_i|. \qquad (26)$$

Further discussion on the penalty term is provided in the simulation study in Section V.

Proposition 1. *The GIC statistic in* (25) *satisfies*

$$
\begin{aligned}
\mathrm{GIC}(\Lambda_i, \tau(|\Lambda_i|, |\mathcal{L}|)) &= \frac{1}{\sigma_\mathbf{e}^2} ||\mathbf{P}_{\Lambda_i} \Delta \mathbf{z}_\mathcal{L}||^2 - \tau(|\Lambda_i|, |\mathcal{L}|) \\
&\quad + const,
\end{aligned} \tag{27}
$$

where const is a term that does not depend on the hypothesis, $i = 0, 1, \ldots, |\mathcal{G}_{K_\mathbf{c}}|.$

Proof. Proof is provided in Appendix A. ∎

It should be noted that for the null hypothesis in Proposition 1, in which $\Lambda_0 = \emptyset$, we use the convention that $\mathbf{P}_{\Lambda_0} \Delta \mathbf{z}_\mathcal{L} = \mathbf{0}$. It can be seen from the GIC statistic in (27) that the selected hypothesis is the one that maximizes the sum of the two terms: 1) the projection of the load measurements onto the associated "attack subspace", $\mathrm{col}(\mathbf{H}_{\mathcal{L}, \Lambda_i})$, by computing $\mathbf{P}_{\Lambda_i} \Delta \mathbf{z}_\mathcal{L}$; and 2) a penalty function, $-\tau(|\Lambda_i|, |\mathcal{L}|)$, that keeps the attack sparse. As a result, the representation of the GIC selection rule from (25) by (27) shows that it is not a function of the unknown states, $\Delta \boldsymbol{\theta}$, and therefore, it is a valid selection rule.

The proposed structural-constrained GIC algorithm for the identification of unobservable FDI attacks is provided in Algorithm 1 and is based on the GIC selection rule in Proposition 1. Detection of unobservable FDI attacks is obtained as a by-product of the identification solution of Algorithm 1, s, where the proposed structural-constrained GIC-based detector decides that there is no attack if $s = 0$ and that an unobservable FDI attack is present, for $s \neq 0$.

Algorithm 1: Structural-constrained GIC

Input:
- difference-based measurements: $\Delta \mathbf{z}$
- network parameters: $\mathbf{H}$, $\mathcal{L}$
- set of candidate state attack supports: $\mathcal{G}_{K_\mathbf{c}}$
- GIC penalty function: $\tau(\cdot, \cdot)$

Output: selected hypothesis: s

1 **for** $\Lambda_i \in \mathcal{G}_{K_\mathbf{c}}$ **do**
2 $\quad$ evaluate $\mathrm{GIC}(\Lambda_i, \tau(|\Lambda_i|, |\mathcal{L}|))$ from (27)
3 **end**
4 **return** $s = \arg \max\limits_{i \in \{0, \ldots, |\mathcal{G}_{K_\mathbf{c}}|\}} \mathrm{GIC}(\Lambda_i, \tau(|\Lambda_i|, |\mathcal{L}|))$

The computational complexity of Algorithm 1 is dominated by step 2 in the **for** loop between steps 1-3, in which the GIC statistic in (27) is computed for each $\Lambda_i \in \mathcal{G}_{K_\mathbf{c}}$, where according to Assumption **A.3**

$$
|\mathcal{G}_{K_\mathbf{c}}| = \sum_{k=0}^{K_\mathbf{c}} \binom{|\mathcal{V}|}{k}, \tag{28}
$$

where $K_\mathbf{c}$ is the maximal sparsity level from (11). For any $k = 1, 2, \ldots, K_\mathbf{c}$, the binomial coefficient, $\binom{|\mathcal{V}|}{k}$, grows by $O(|\mathcal{V}|^k)$ (p. 36, in [56]). Thus, the number of times the GIC statistic is evaluated in Algorithm 1, which is the number of elements in (28), is in the order of $O(\sum_{k=0}^{K_\mathbf{c}} |\mathcal{V}|^k)$. The number of matrix-vector multiplications required for the computation of each GIC statistic in (27) is $O(|\mathcal{L}||\Lambda_i|^2 + (|\Lambda_i| + 1)|\mathcal{L}|^2 + |\mathcal{L}|)$. Thus, for large-scale power systems, where $|\mathcal{V}|$ and $|\mathcal{L}|$ are significantly large compared with $K_\mathbf{c}$, the complexity is in the order of $O(|\mathcal{V}|^{K_\mathbf{c}} (K_\mathbf{c} + 1)|\mathcal{L}|^2)$ and thus, it may be infeasible to use the structural-constrained GIC method. Therefore, in Section IV, we provide two low-complexity approximations.

C. PSSE in the presence of unobservable FDI attacks

The difference-based model in (12) is achieved by taking the difference between two consecutive time samples that satisfy the model in (9) with $\mathbf{c} = \mathbf{0}$ for the first time sample and an arbitrary $\mathbf{c}$ in the second time sample. Thus, according to (9), the power measurements at the two consecutive times t and $t + 1$ can be presented in a matrix form as follows:

$$
\begin{bmatrix} \mathbf{z}_t \\ \mathbf{z}_{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{H} & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{H} \end{bmatrix} \left(\begin{bmatrix} \mathbf{0}_N \\ \mathbf{c} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\theta}_t \\ \boldsymbol{\theta}_{t+1} \end{bmatrix} \right) + \begin{bmatrix} \mathbf{e}_t \\ \mathbf{e}_{t+1} \end{bmatrix}, \tag{29}
$$

where, for this case, the differential measurements in (12) are realized by subtracting $\mathbf{z}_t$ from $\mathbf{z}_{t+1}$ to obtain $\Delta \mathbf{z} = \mathbf{z}_{t+1} - \mathbf{z}_t$. Therefore, for the discussed case, the state attack vector, $\mathbf{c}$, in (29) and (12) is the same. Thus, the differential model in (12) is used in order to estimate the state attack vector in (29). Specifically, the state attack vector estimation, $\hat{\mathbf{c}}_{\Lambda_s}^{\mathrm{ML}|s}$, is obtained by applying Algorithm 1 on $\Delta \mathbf{z}$ and substituting the selected support, Λ_s, in (59) from Appendix A. Finally, the state estimation is computed by substituting $\mathbf{c} = \hat{\mathbf{c}}_{\Lambda_s}^{\mathrm{ML}|s}$ in (29) and performing standard WLS:

$$
\begin{bmatrix} \hat{\boldsymbol{\theta}}_t \\ \hat{\boldsymbol{\theta}}_{t+1}^{\hat{\mathbf{c}}_{\Lambda_s}^{\mathrm{ML}|s}} \end{bmatrix} = \begin{bmatrix} \mathbf{K} \mathbf{z}_t \\ \mathbf{K}(\mathbf{z}_{t+1} - \mathbf{H}\hat{\mathbf{c}}_{\Lambda_s}^{\mathrm{ML}|s}) \end{bmatrix}, \tag{30}
$$

where $\mathbf{K}$ is defined in (5).

D. Remarks

1) Remark 1: It should be noted that the proposed framework is designed specifically for unobservable FDI attacks with the associated measurement model from (9). In the general case, unobservable attacks can be combined with random attacks and bad data can be induced due to system faults. For this case, the measurements can be preprocessed by conventional BDD methods that remove conventional faulty measurements. Then, the reconstructed measurements at the output of the conventional PSSE can be used as the input measurement for our approach.

2) Remark 2: Assumption **A.1** can be relaxed by integrating the proposed method into an adaptive scheme in a similar manner to that outlined in [52], as illustrated in Fig. 1. This scheme takes the power measurements of any two sequential data collection slots into account and detects data injection attacks by constantly monitoring the measurement variations between the two time slots. In practice, in each time step $t + 1$ we apply the GIC approach from Algorithm 1 with the input $\Delta \mathbf{z} = \mathbf{z}_{t+1} - \mathbf{z}_t$. The adaptive scheme is terminated if an attack is detected and the locations of the infected measurements are provided to the operators in order to resecure the system. After the system is resecured, the adaptive scheme is reinitialized.

IV. LOW-COMPLEXITY IDENTIFICATION METHODS

As shown at the end of Subsection III-B, the computational complexity of the structural-constrained GIC method makes it impractical for large power system networks, where $|\mathcal{V}|$ and $|\mathcal{L}|$ are considered to be large. Therefore, in this section, we develop two low-complexity methods that rely on Assumptions **A.1**-**A.4**: 1) an OMP-based method [47]; and 2) a novel method that exploits the graph Markovian property of order two between the nodes (buses) in the graph representation of the power system [48].
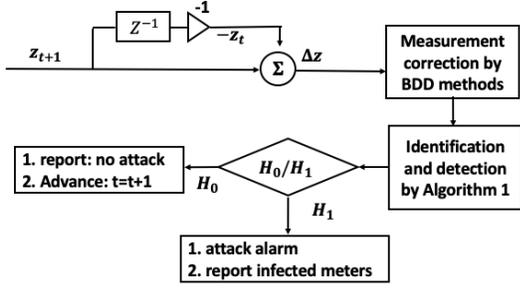
Fig. 1. Adaptive scheme for detection and identification of unobservable FDI attacks: In time step $t + 1$, the previous measurements at time step t, $\mathbf{z}_t$, are subtracted from the current measurements at time step $t + 1$, $\mathbf{z}_{t+1}$, to obtain the differential measurement vector: $\Delta\mathbf{z} = \mathbf{z}_{t+1} - \mathbf{z}_t$. In the first processing stage, $\Delta\mathbf{z}$ is corrected by a standard BDD method. In the second stage, Algorithm 1 is applied for the detection and identification of unobservable FDI attacks. If the null hypothesis is selected, then the adaptive scheme proceeds to the next time step. Otherwise, if an attack is detected, then the adaptive scheme is terminated and the operators are provided by the locations of the infected meters. After the system is resecured, the adaptive scheme is reinitialized.

A. OMP method

The OMP algorithm in [47] is an efficient method for the recovery of sparse signals. The basic principle behind the OMP algorithm is to iteratively find the support set of the sparse vector. The OMP method proceeds by finding the column of the CS matrix that correlates most strongly with the signal residual. The residual is constructed in each iteration by projecting the measurements onto the linear space spanned by the remaining columns that were not selected in previous iterations.

In this subsection, we apply the OMP algorithm for the sparse recovery of the state attack vector, $\mathbf{c}$, which is a sparse signal as described in Assumption **A.3**, from the measurements in (13). It should be noted that the measurement model in (13) contains a nuisance parameter vector, $\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta\boldsymbol{\theta}$, which is not a part of the conventional sparse recovery model. Nevertheless, similarly to in the derivations in Subsection III-A, it can be shown that under Assumptions **A.1**-**A.4** the nuisance parameter has a negligible effect on the OMP selection criterion. Thus, the conventional OMP method is valid for this setting.

The main iteration of the OMP algorithm, which is performed on the load measurements model from (13), is given as follows. Suppose $\Lambda^{(j)}$ is the estimated support set of $\mathbf{c}$ in the jth iterative step. In the $j + 1$th iteration, we compute

$$k = \arg\max_{\tilde{k} \in \mathcal{V}} ||\mathbf{P}_{\tilde{k}}\mathbf{r}^{\mathrm{OMP}|(j)}||^2, \qquad (31)$$

where $\mathbf{P}_{\tilde{k}}$ is obtained by replacing $\Lambda_i = \tilde{k}$ in (18), and

$$\mathbf{r}^{\mathrm{OMP}|(j)} = \mathbf{P}_{\Lambda^{(j)}}^{\perp}\Delta\mathbf{z}_{\mathcal{L}} \qquad (32)$$

is the signal residual at the jth iteration. By substituting (32) and $\mathbf{P}_{\Lambda^{(j)}}^{\perp} = \mathbf{I} - \mathbf{P}_{\Lambda^{(j)}}$ on the r.h.s. of (31) we obtain

$$||\mathbf{P}_{\{k\}}\mathbf{r}^{\mathrm{OMP}|(j)}||^2 = ||\mathbf{P}_{\{k\}}\Delta\mathbf{z}_{\mathcal{L}} - \mathbf{P}_{\{k\}}\mathbf{P}_{\Lambda^{(j)}}\Delta\mathbf{z}_{\mathcal{L}}||^2. \qquad (33)$$

By limiting the number of iterations to the maximal sparsity level, $K_{\mathbf{c}}$, we obtain that $\Lambda^{(j)} \in \mathcal{G}_{K_{\mathbf{c}}}$. In addition, any single bus, $k \in \mathcal{V}$, is an element in $\mathcal{G}_{K_{\mathbf{c}}}$, i.e. $k \in \mathcal{G}_{K_{\mathbf{c}}}$, as well. Therefore, under Assumptions **A.1**-**A.4**, we obtain from (22) and (23) that the nuisance parameter $\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta\boldsymbol{\theta}$ has a minor effect on both the terms $\mathbf{P}_{\{k\}}\Delta\mathbf{z}_{\mathcal{L}}$ and $\mathbf{P}_{\Lambda^{(j)}}\Delta\mathbf{z}_{\mathcal{L}}$. Consequently, the nuisance parameter has a minor effect on (33) and, thus, its influence on the OMP selection procedure in (31) can be neglected. The proposed structural-constrained OMP algorithm for the identification of unobservable FDI attacks is provided in Algorithm 2.

Algorithm 2: Structural-constrained OMP

Input:
- difference-based measurements: $\Delta\mathbf{z}$
- network parameters: $\mathbf{H}$, $\mathcal{V}$, $\mathcal{L}$
- maximal sparsity level: $K_{\mathbf{c}}$
- stopping rule threshold: γ^{OMP}

Output: estimated support: $\hat{\Lambda}^{\mathrm{OMP}}$

1 Initialize: $\Lambda^{(0)} = \emptyset$, $\mathbf{r}^{\mathrm{OMP}|(0)} = \Delta\mathbf{z}_{\mathcal{L}}$
2 **for** $j = 1 \ldots K_{\mathbf{c}}$ **do**
3 $\quad$ evaluate k from (31)
4 $\quad$ **if** $||\mathbf{P}_{\{k\}}\mathbf{r}^{\mathrm{OMP}|(j-1)}||^2 < \gamma^{\mathrm{OMP}}$ **then**
5 $\quad\quad$ **return** $\Lambda_s = \hat{\Lambda}^{(j-1)}$
6 $\quad$ **end**
7 $\quad$ update: $\Lambda^{(j)} = \{\Lambda^{(j-1)} \cup k\}$
8 $\quad$ update: $\mathbf{r}^{\mathrm{OMP}|(j)} = \mathbf{P}_{\Lambda^{(j)}}^{\perp}\Delta\mathbf{z}_{\mathcal{L}}$
9 **end**
10 **return** $\hat{\Lambda}^{\mathrm{OMP}} = \hat{\Lambda}^{(j)}$

The computational complexity of Algorithm 2 is dominated by Step 3 (computing (31)), which requires $O(2|\mathcal{L}|^2 + 3|\mathcal{L}|)$ matrix-vector multiplications for each $k = 1, 2, \ldots, |\mathcal{V}|$. The loop between Steps 2-8 is performed at most $K_{\mathbf{c}}$ times. As a result, the total complexity of Algorithm 2 is in the order of $O(|\mathcal{V}|K_c|\mathcal{L}|^2)$, which is significantly lower than the complexity of the GIC method (see discussion after (28)).

In general, the OMP method is used in a variety of applications due to its low computational complexity. However, the OMP method is a greedy algorithm with no optimal recovery guarantees, and usually requires an incoherent dictionary in order to provide high performance [35]. In our case, the CS matrix is the topology matrix $\mathbf{H}_{\mathcal{L},\mathcal{V}}$, which may be highly correlated, and thus, with large mutual coherence. Therefore, in the following subsection, we develop a novel low-complexity method that uses the power system graph representation Markovian properties.

B. Graph Markovian low-complexity GIC (GM-GIC)

In this section, we develop the low complexity GM-GIC method for the measurement model in (13) and under Assumptions **A.1**-**A.4**, which considers the graphical representation of the power system. Accordingly, in this subsection the system buses are referred to as the graph nodes. The power system graphical representation is utilized in Subsection IV-B1 to analyze the affect of an unobservable FDI attack on the measurements. Based on this analysis, we derive the GM-GIC

method, which consists of the following four stages: 1) low-scale pre-screening (Subsection IV-B2); 2) node partitioning (Subsection IV-B3); 3) local GIC, on the partitioned subsets; and 4) sparsity correction. The GM-GIC method, including Stages 3 and 4, is summarized in Subsection IV-B4.

1) Attack analysis: In the following we analyze the effect of an unobservable FDI attack on the system measurements. Specifically, we focus on how an attack, with a support Λ, affects the single-node measurement subspace, $\mathrm{col}(\mathbf{H}_{\mathcal{L},m})$, for any node $m \in \mathcal{V}$. Considering the measurement model in (13), the unobservable FDI attack can be linearly decomposed as follows:

$$\mathbf{Hc} = \mathbf{H}_{\mathcal{L},\Lambda}\mathbf{c}_\Lambda = \sum_{k \in \Lambda} \mathbf{H}_{\mathcal{L},k}c_k. \tag{34}$$

Based on the decomposition in (34), we can analyze the influence of an attack on a single-node measurement subspace by summing over the individual influences on each node, $k \in \Lambda$. The following proposition evaluates the effect of a single-node attack on a single-node measurement subspace.

Proposition 2. *The projection of a single-node attack on node $k \in \Lambda$, $\mathbf{H}_{\mathcal{L},k}c_k$, onto the measurement subspace associated with node $m \in \mathcal{V}$, $col(\mathbf{H}_{\mathcal{L},m})$, satisfies*

$$\mathbf{P}_{\{m\}}\mathbf{H}_{\mathcal{L},k}c_k = \mathbf{0}, \quad \forall k, m \in \mathcal{V}, \ d(k,m) > 2 \tag{35}$$

and

$$||\mathbf{P}_{\{m\}}\mathbf{H}_{\mathcal{L},k}c_k||^2 \le ||\mathbf{P}_{\{k\}}\mathbf{H}_{\mathcal{L},k}c_k||^2, \quad \forall k, m \in \mathcal{V}, \tag{36}$$

where $\mathbf{P}_{\{m\}}$ is obtained by substituting $\Lambda = \{m\}$ in (18).

Proof. The proof appears in Appendix B. ∎

This proposition demonstrates that the single-node attack obtains a 'local' effect. That is, a single-node attack on node k does not affect the single-node measurement subspace of node m if the geodesic distance (hop distance) [57] between nodes k and m, $d(k,m)$, is greater than two, $d(k,m) > 2$. By multiplying (34) by $\mathbf{P}_{\{m\}}$ and substituting (35), we obtain

$$\mathbf{P}_{\{m\}}\mathbf{H}_{\mathcal{L},\Lambda}c_\Lambda = \mathbf{0}, \quad \forall k \in \Lambda, \ d(k,m) > 2. \tag{37}$$

As a result of (37), the single-node measurement subspaces affected by the unobservable FDI attack are only those which are associated with nodes in the set

$$\mathcal{A} \triangleq \{m \in \mathcal{V} : \ \exists k \in \Lambda \text{ s.t. } 0 \le d(k,m) \le 2\}, \tag{38}$$

where, according to (38), the set $\mathcal{A}$ includes the attacked nodes, those in Λ, and first- or second-order neighbors of attacked nodes. This observation can be interpreted as a second-order graph Markov property [48].

In general, the power system network generates a *sparse* graph where nodes are connected to 2-5 neighbors [58]. Thus, the nodes in the power system have at most 5 first-order and 25 second-order neighbor nodes. As a result, the number of nodes affected by the attack is bounded as follows:

$$|\mathcal{A}| \le (1 + 5 + 25)K_{\mathbf{c}}, \tag{39}$$

where $K_{\mathbf{c}}$ is the sparsity term defined in Assumption **A.3** and $\mathcal{A}$ is defined in (38). Moreover, the power system network obtains a local behavior for the connectivity pattern, where

only substations geographically close are likely to be connected. Thus, the bound in (39) is not a tight bound and $|\mathcal{A}|$ is expected to be significantly lower than the r.h.s of (39). In conclusion, it is implied that for large networks $|\mathcal{A}| \ll |\mathcal{V}|$.

2) Pre-screening: The first stage of the GM-GIC method is a pre-screening stage. In order to evaluate if the dictionary matrix column associated with node $m \in \mathcal{V}$, $\mathbf{H}_{\mathcal{L},m}$, is correlated with the attack, $\mathbf{Hc}$, we can implement a binary hypothesis testing problem, which is obtained by substituting $\Lambda_i = \{m\}$ in (24):

$$\begin{aligned} \mathcal{H}_0: \quad & \Delta\mathbf{z}_{\mathcal{L}} = \mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta\boldsymbol{\theta} + \Delta\mathbf{e}_{\mathcal{L}} \\ \mathcal{H}_1: \quad & \Delta\mathbf{z}_{\mathcal{L}} = \mathbf{H}_{\mathcal{L},m}\mathbf{c}_m + \mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta\boldsymbol{\theta} + \Delta\mathbf{e}_{\mathcal{L}} \\ & \text{s.t. } \mathbf{P}_{\{m\}}\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta\boldsymbol{\theta} = \mathbf{0}. \end{aligned} \tag{40}$$

That is, instead of multiple hypothesis testing in (24), we solve here multiple problems of binary hypothesis testing in the form of (40). As a result, the linear dependencies between the dictionary matrix, $\mathbf{H}_{\mathcal{L},\mathcal{V}}$, columns are not taken into account.

In [59] it is shown that the GIC test provides the same decision rule as the GLRT for a binary hypothesis testing problem with a fixed threshold. Therefore, by using the GIC test in (27), the GLRT for the associated binary hypothesis problem in (39) is given by

$$||\mathbf{P}_{\{m\}}\Delta\mathbf{z}_{\mathcal{L}}||^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \rho, \tag{41}$$

where ρ is set to determine a desired false alarm rate. Applying the GLRT detector in (41) for any node in the system yields the following set of suspicious nodes:

$$\mathcal{S} \triangleq \{m \in \mathcal{V} : \ ||\mathbf{P}_{\{m\}}\Delta\mathbf{z}_{\mathcal{L}}||^2 > \rho\}. \tag{42}$$

Thus, the node m will be included in the set $\mathcal{S}$ if abnormal energy is detected in the single-node measurement subspace, $\mathrm{col}(\mathbf{H}_{:,m})$. From (23), the projection of the nuisance parameter onto this subspace, $\mathbf{P}_{\{m\}}\mathbf{H}_{\mathcal{L},\mathcal{V}}\Delta\boldsymbol{\theta}$, is negligible. Therefore, the inclusion of the node m in the set $\mathcal{S}$ indicates that the measurement subspace, $\mathrm{col}(\mathbf{H}_{:,m})$, is affected by the attack, i.e. that $\mathbf{P}_{\{m\}}\mathbf{H}_{\mathcal{L},\Lambda}\mathbf{c}_\Lambda \ne \mathbf{0}$. Thus, from (37), the set of suspicious nodes in (42), can be interpreted as an estimator of the set $\mathcal{A}$ in (38). From (36), it can be seen that the norm of the projected single-node attack on single-node measurement subspaces of attacked nodes is higher than the norm of the projection onto measurement subspaces associated with other nodes. Thus, considering the discussion after (34), the effect of the attack, $\mathbf{H}_{\mathcal{L},\Lambda}\mathbf{c}_\Lambda$, on attacked nodes is expected to be higher than the effect on their first- and second- order neighbor nodes. Therefore, even if the estimator $\mathcal{S}$ does not cover all the nodes in $\mathcal{A}$, the attacked nodes are still expected to be included.

3) Node partitioning: The second stage of the GM-GIC method is the partitioning of the suspicious set, $\mathcal{S}$, in (42), into disjoint subsets. This partitioning provides the condition that an attack on nodes located in one of the subsets does not affect the measurement subspaces associated with nodes in the other subsets. According to the proposed partitioning, a general set $\mathcal{S}$ is partitioned into Q disjoint subsets, $\{\mathcal{S}_q\}_{q=1}^{Q}$, if it is satisfied that

$$d(k,m) > 2, \ \forall k \in \mathcal{S}_q, m \in \mathcal{S}_p \tag{43}$$

for any two different subsets, S_q and S_p, $p \neq q$, selected from $\{S_q\}_{q=1}^Q$, where $1 \leq Q \leq K_c$. Thus, from (37) and (43), an attack on any node subset in S_q, $\Lambda_q \in S_q$, does not affect the column space associated with the nodes in any of the other subsets, e.g. $\text{col}(\mathbf{H}_{\mathcal{L},S_p})$, i.e.

$$\mathbf{P}_{S_p}\mathbf{H}_{\mathcal{L},\Lambda_q} = \mathbf{0}. \tag{44}$$

Therefore, a node partitioning that satisfies (43) enables identification of the attacked buses on each subset separately and, thus, reduces the problem dimension. In contrast, the GIC from Algorithm 1 considers the entire node set $\mathcal{V}$.

In practice, finding a node partitioning that satisfies (43) is performed as follows. First, based on the graph representation of the given power system, we generate the unweighted undirected graph $\tilde{\mathcal{G}} = (\mathcal{S}, \mathcal{E}_{\tilde{\mathcal{G}}})$, where $\mathcal{S}$ are the nodes and the set of edges is defined as

$$\mathcal{E}_{\tilde{\mathcal{G}}} \triangleq \{(k,m) : k,m \in \mathcal{S}, \ 1 \leq d(k,m) \leq 2\}, \tag{45}$$

in which $d(\cdot,\cdot)$ refers to the geodesic distance measured on the original graph that represents the power system network. Then, we find the connected components of $\tilde{\mathcal{G}}$ (e.g. by using the Matlab command conncomp). According to the definition in (45), we obtain that $d(k,m) > 2$ for any k and m that belong to different connected components of $\tilde{\mathcal{G}}$. Therefore, selecting the partition subsets $\{S\}_{q=1}^Q$ to be the node sets of the connected components of $\tilde{G}$ satisfies (43).

4) Summary: GM-GIC method: In this subsection we summarize the proposed GM-GIC method. In the first stage of this method, a reduced set composed of suspicious nodes, $\mathcal{S}$, is extracted from the node set $\mathcal{V}$, according to (42). In the second stage, the set $\mathcal{S}$ is partitioned into Q disjoint subsets, $\{S_q\}_{q=1}^Q$, as described in Subsection IV-B3. In the third stage, the GIC method is applied on each subset separately, by replacing the set of candidate state attack supports, $\mathcal{G}_{K_c}$, in Algorithm 1 with

$$\mathcal{G}_{K_c}^{S_q} \triangleq \{\Lambda \subset S_q : \ 1 \leq |\Lambda| \leq K_c\}. \tag{46}$$

As a result, for each subset S_q we obtain a partial estimation of the support set, which is denoted by $\hat{\Lambda}_q$. The total support set of the state attack vector is the union of all the partial estimates

$$\hat{\Lambda}_{\text{temp}}^{\text{GM-GIC}} = \bigcup_{q=1}^Q \hat{\Lambda}_q. \tag{47}$$

The estimated support in (47) may exceed the sparsity limit $|\hat{\Lambda}_{\text{temp}}^{\text{GM-GIC}}| \leq K_c$. Thus, $\hat{\Lambda}_{\text{temp}}^{\text{GM-GIC}}$ may include node-elements which are not attacked. It should be noted that in this stage, Assumption **A.2** is relaxed and the sparsity restriction is only imposed on the partial (separated) estimates, i.e. $\Lambda_q \leq K_c$ In order to reduce the identification errors that may be induced by this relaxation, in the fourth stage, the estimated the estimated support in (47) is corrected to satisfy $|\hat{\Lambda}^{\text{GM-GIC}}| \leq K_c$. This stage is performed by: 1) evaluating the state attack ML estimation, $\hat{\mathbf{c}}_{\hat{\Lambda}^{\text{GM-GIC}}}^{\text{ML}|i}$, by replacing Λ_i with (47) in (59) from Appendix A; 2) sorting $\hat{\mathbf{c}}_{\hat{\Lambda}_{\text{temp}}^{\text{GM-GIC}}}^{\text{ML}|i}$ in a descending order:

$$\mathcal{I} \triangleq \{i_1, i_2, \ldots, i_{|\hat{\Lambda}^{\text{GM-GIC}}|}\} = \text{sort}(\hat{\mathbf{c}}_{\hat{\Lambda}_{\text{temp}}^{\text{GM-GIC}}}^{\text{ML}|i}); \tag{48}$$

and 3) preserving only the K_c elements with the highest absolute values,

$$\hat{\Lambda}^{\text{GM-GIC}} = \{i_1, i_2, \ldots, i_{K_c}\}. \tag{49}$$

The proposed GM-GIC algorithm for the identification of unobservable FDI attacks is provided in Algorithm 3.

Algorithm 3: GM-GIC

Input:
- difference-based measurements: $\Delta\mathbf{z}$
- network parameters: $\mathbf{H}$, $\mathcal{V}$, $\mathcal{L}$
- maximal sparsity level: K_c
- energy threshold: ρ

Output: estimated support: $\hat{\Lambda}^{\text{GM-GIC}}$

1   Pre-screening: evaluate $\mathcal{S}$ by (42)
2   **if** $\mathcal{S} = \emptyset$ **then return** $\hat{\Lambda}^{\text{GM-GIC}} = \emptyset$
3   Node-partitioning: 1) generate $\tilde{\mathcal{G}} = (\mathcal{S}, \mathcal{E}_{\tilde{\mathcal{G}}})$ by computing (45); 2) partition $\tilde{\mathcal{G}}$ into its connected components (e.g. by conncomp in Matlab); and 3) set $\{S_q\}_{q=1}^Q$ to be these components
4   **for** $q = 1 \ldots Q$ **do**
5   generate the graph $\mathcal{G}_{K_c}^{S_q}$ by (46)
6   compute $\hat{\Lambda}_q$ by applying Algorithm 1 with the input set of candidate state attack supports $\mathcal{G}_{K_c}^{S_q}$
7   **end**
8   compute $\hat{\Lambda}^{\text{GM-GIC}}$ by (47)
9   **if** $|\hat{\Lambda}^{GM\text{-}GIC}| > K_c$ **then**
10   compute $\hat{\mathbf{c}}_{\hat{\Lambda}^{\text{GM-GIC}}}^{\text{ML}|i}$ by (59)
11   correct $\hat{\Lambda}^{\text{GM-GIC}}$ by (48) and (49).
12   **end**
13   **return** $\hat{\Lambda}^{\text{GM-GIC}}$

The computational complexity of the different stages of the GM-GIC method in Algorithm 3 are as follows: 1) the pre-screening stage in Step 1 requires $O(|\mathcal{V}|(2|\mathcal{L}|^2 + 3|\mathcal{L}|))$ matrix-vector multiplications; 2) the node partitioning stage in Step 3 is dominated by the graph partitioning, which is implemented by the conncomp($\cdot$) Matlab command and requires $O(|\mathcal{S}| + |\mathcal{S}|^2)$ computations [60]. Considering that the number of edges in $\tilde{\mathcal{G}}$ cannot be anticipated in advance, this analysis is based on the worst case in which $\tilde{\mathcal{G}}$ is a fully connected graph. 3) in Steps 4-7 the GIC method is applied on each of the subsets $\{S_q\}_{q=1}^Q$ and, thus, from Subsection III-B, the complexity is in the order of $O((\sum_{q=1}^Q |S_q|^{K_c})(K_c+1)|\mathcal{L}|^2)$; and 4) the sparsity correction stage, in steps 9-12, is dominated by step 10, which requires $O(K_c(K_c+1)|\mathcal{L}|)$ matrix-vector multiplications. In power systems, the number of load nodes (buses) $|\mathcal{L}|$ is, in general, in scale with the number of nodes $|\mathcal{V}|$. Thus, from Subsections IV-B1-IV-B2, we obtain that $|\mathcal{S}| \ll |\mathcal{L}|$. Thus, the total computational complexity of Algorithm 3 is $O((\sum_{q=1}^Q |S_q|^{K_c})(K_c+1)|\mathcal{L}|^2)$.

The computational complexity of the different methods is summarized in Table I. It can be seen that the computational complexity of the GM-GIC method is significantly lower than the complexity of the GIC method, but may be higher than the complexity of the OMP method. Finally, the worst complexity for the GM-GIC method is obtained when $\mathcal{S}$

GIC	$O(|\mathcal{V}|^{K_{\mathbf{c}}}(K_{\mathbf{c}}+1)|\mathcal{L}|^2)$
OMP	$O(|\mathcal{V}|K_c|\mathcal{L}|^2)$
GM-GIC	$O((\sum_{q=1}^{Q}|\mathcal{S}_q|^{K_{\mathbf{c}}})(K_{\mathbf{c}}+1)|\mathcal{L}|^2)$
GM-GIC (worst case)	$O(|\mathcal{S}|^{K_{\mathbf{c}}}(K_{\mathbf{c}}+1)|\mathcal{L}|^2)$

Table I: Computational complexity of the proposed methods.

cannot be partitioned. In this case, the complexity is in the order of $O(|\mathcal{S}|^{K_{\mathbf{c}}}(K_{\mathbf{c}}+1)|\mathcal{L}|^2)$, which is still significantly lower than the computational complexity of the GIC method.

The GM-GIC heuristic is designed to exploit the graphical properties of the power system networks that include: 1) graph Markovity; 2) graph sparsity; and 3) local graph connectivity behavior. These graphical properties are utilized as part of the attack analysis performed in Subsection IV-B1, which is the basis for the GM-GIC method. Thus, this method can be an effective method for sparse recovery in GSP frameworks in which these properties are satisfied. Otherwise, the low computational complexity of the GM-GIC method is not guaranteed, where the worst case complexity approaches the complexity of the GIC method.

V. SIMULATIONS

In this section, the performance of the proposed methods is demonstrated for the tasks of detection and identification of unobservable FDI attacks and for the task of PSSE in the presence of these attacks. The simulations are conducted on the IEEE-30 bus test case, where the topology matrix and measurement data are extracted using the Matpower toolbox for Matlab [61]. The simulation set-up is described in Subsection V-A, while the proposed method's performance is demonstrated in Subsection V-B.

A. Set-up

1) Measurements: The simulation study is conducted on the difference-based model in (12). For the sake of simplicity of implementation, we ensure Assumption **A.2** by defining the set $\tilde{\mathcal{V}}$ to include only state variables that are related to load measurements, and then restricting the support of the state attack vector to $\Lambda \subseteq \tilde{\mathcal{V}}$. In particular, for the IEEE-30 bus test case it can be verified that $\tilde{\mathcal{V}} = \{14, 16, 17, 18, 19, 20\}$. In each simulation, we set the cardinality of the state attack support vector to be $|\Lambda| = K_{\mathbf{a}}$ and draw the support vector, Λ, uniformly from the set $\{\Lambda \in \tilde{\mathcal{V}} : |\Lambda| = K_{\mathbf{a}}\}$. The attack values on the chosen support are randomly drawn from the uniform distribution over the interval $[-1, 1]$. Then, the values are scaled to obtain a desired value of attack norm, $||\mathbf{a}||$.

The differential state vector is obtained by first generating the state vectors at times t and $t+1$, $\boldsymbol{\theta}_t$ and $\boldsymbol{\theta}_{t+1}$, respectively, and then subtracting $\boldsymbol{\theta}_t$ from $\boldsymbol{\theta}_{t+1}$ to obtain $\Delta\boldsymbol{\theta} = \boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_t$. The state vector at time t, $\boldsymbol{\theta}_t$, is set to the values in the IEEE test case [61]. Based on $\boldsymbol{\theta}_t$ and considering Assumption **A.4**, the state vector at time $t+1$, $\boldsymbol{\theta}_{t+1}$, is simulated by first updating the load measurements with random scaling:

$$\phi\mathbf{H}_{\mathcal{L},\mathcal{V}}\boldsymbol{\theta}_t, \ \phi \sim \mathcal{N}(\mathbf{1}, \sigma_s^2\mathbf{I}), \tag{50}$$

and then computing $\boldsymbol{\theta}_{t+1}$ by the rundcpf($\cdot$) Matpower command in [61]. In addition, throughout the simulations, the noise is modeled as in (9) with $\sigma_e^2 = 0.01$.

2) Methods: The proposed methods include the structural-constrained GIC method from Section III, and the low-complexity methods in Section IV: the structural-constrained OMP and the GM-GIC method. The penalty function for the GIC and GM-GIC methods defined in (27) is set by

$$\tau(|\Lambda_i|, \mathcal{L}) = \zeta|\Lambda_i| + \gamma^{\text{GIC}}\delta[|\Lambda_i|], \tag{51}$$

where ζ and γ^{GIC} are user-defined regularization parameters and $\delta[\cdot]$ is the Kronecker delta function. The term $\zeta|\Lambda_i|$ is set to encourage sparse solutions for the identification problem, while the term $\gamma^{\text{GIC}}\delta(|\Lambda_i|)$ is set to maintain a desired false alarm rate for the detection problem. In particular, the probability of false alarm is set to $P_{FA} = 0.05$ unless stated otherwise. Furthermore, by selecting the AIC in (26), the GIC tuning parameter in (51) is set to $\zeta = 2$. The maximal sparsity rate is set to $K_{\mathbf{c}} = 6$.

The performance of the proposed methods is compared with the following methods that were all modified according to the difference-based model in (12):

M.1 The sparse optimization technique in [45], which considers an ℓ_2 relaxation for the attack sparsity restriction (denoted as ℓ_2).

M.2 The sparse optimization technique in [43], which considers an ℓ_1 relaxation for the attack sparsity restriction (denoted as ℓ_1).

M.3 The GFT-based detector in [50] (denoted as GFT).

M.4 The energy detector, which is obtained by comparing $T^{\text{ENG}} = ||\Delta\mathbf{z}||^2$ with a chosen threshold and was inspired by the detector in [52] (denoted as ENG).

M.5 The BDD detector in (7) (denoted as BDD).

All numerical results in this section were obtained using at least 500 Monte Carlo simulations. The detection thresholds were computed from simulated historic data obtained by 500 off-line simulations of (12) under the null hypothesis.

B. Simulations

1) Detection: In Fig. 2. the receiver operating characteristics (ROC) curves of the proposed methods: GIC, GM-GIC, and OMP, are presented and compared with those of the methods **M.1**-**M.5**. The simulations were conducted on the IEEE-30 bus system with an attack on four state variables, $K_a = 4$, where the attack norm is normalized to $||\mathbf{a}|| = 0.05K_a$. The loads scaling variation in (50) is set to $\sigma_s^2 = 0.05$. This figure shows that the detection performance of the proposed low-complexity methods, GM-GIC, and OMP, converges to the performance of the optimal GIC method. Furthermore, the proposed GIC, GM-GIC, and OMP methods provide a higher detection probability for any chosen false alarm rate, when compared to the previous methods. It should be noted that the sparse ℓ_2 and ℓ_1 methods were developed based on models with multiple time measurements and where the sparsity pattern of the attack is over both the time and the bus domains, while in the considered model only two time samples are provided. Finally, as expected from Subsection II-C, the BDD method cannot detect unobservable FDI attacks and is no better than flipping a coin.

2) Identification: The identification performance is measured by the capability to classify each state variable as
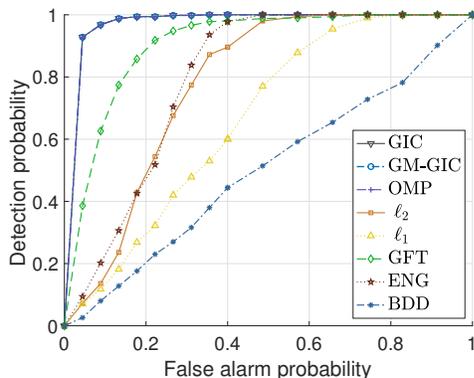
Fig. 2. Unobservable FDI attack detection: The ROC curve of the different methods for $K_{\mathbf{a}} = 4$, $\frac{||\mathbf{a}||}{K_{\mathbf{a}}} = 0.05$, and $\sigma_s^2 = 0.05$.

manipulated or not. Therefore, we evaluate the identification performance by the F-score classification metric [62]:

$$FS(\hat{\Lambda}, \Lambda) = \frac{2t_p}{2t_p + f_n + f_p}, \qquad (52)$$

where Λ is the true support of the state attack vector, $\mathbf{c}$, and $\hat{\Lambda}$ is the estimated support by a given method. The terms t_p, f_p, and f_n, are the true-positive, false-positive, and false-negative probabilities, respectively. The F-score metric takes values between 0 and 1, where 1 means perfect identification.

In Fig. 3, the identification performance, measured by the F-score metric, of the GIC, GM-GIC, OMP, ℓ_2, and ℓ_1 methods is examined w.r.t. to the attack characteristics. Methods **M.3**-**M.5** are not included in the comparison since they provide solely detection and do not identify the attacked buses' locations. In Fig. 3.a the F-score is presented versus the normalized attack norm, $\frac{||\mathbf{a}||}{K_{\mathbf{a}}}$, for $K_{\mathbf{a}} = 4$ and $\sigma_s^2 = 0.1$. In addition, in Fig. 3.b the F-score is presented versus the number of attacked elements, $K_{\mathbf{a}}$, for $||\mathbf{a}|| = 1.2$ and $\sigma_s^2 = 0.05$. It can be observed from both figures that the proposed methods have a significantly higher F-score than those of the previous sparsity-based methods, ℓ_2 and ℓ_1. In addition, it can be seen from Fig. 3.a that the performance of all the methods increases with the increase of the attack measured by $\frac{||\mathbf{a}||}{K_{\mathbf{a}}}$. Similarly, Fig. 3.b shows that the performance of all the methods decreases as $K_{\mathbf{a}}$ increases. Nonetheless, the F-score of the proposed methods is above 0.8 even when a considerable portion (> 0.2) of the system is attacked. Finally, it can be seen that both the GM-GIC and OMP methods show relatively close results to the the GIC method, where the GM-GIC method provides higher identification rates than the OMP method.

3) Identification under mismatch: The use of the GIC approach for identifying and detecting unobservable FDI attacks is based on the assumption that the nuisance parameter $\mathbf{H}\Delta\boldsymbol{\theta}$ is bounded and small, as described in (16). In Fig. 4, we study the robustness of the performance to this assumption by evaluating the influence of the nuisance parameter, $\mathbf{H}\Delta\boldsymbol{\theta}$, on the identification performance of the different methods. In particular, the F-score of the different methods is compared versus $||\mathbf{H}\Delta\boldsymbol{\theta}||^2$ for $K_{\mathbf{a}} = 4$ and $||\mathbf{a}|| = 0.05K_{\mathbf{a}}$. Furthermore, as an upper bound on the F-score performance we use the oracle GIC. This method is obtained by applying the GIC method in Algorithm 1 after replacing $\Delta\mathbf{z}$ with the true (unknown) value $\Delta\mathbf{z}_{\mathcal{L}} - \mathbf{H}\Delta\boldsymbol{\theta}$, i.e. by applying the GIC criterion with
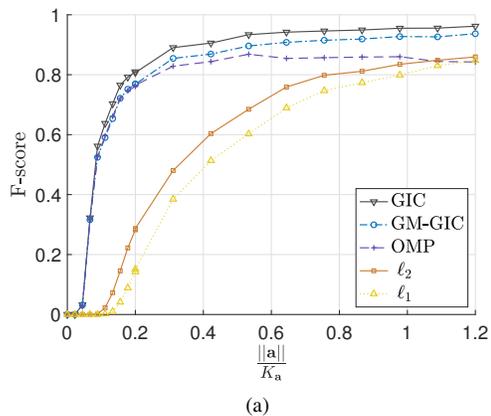
(a)
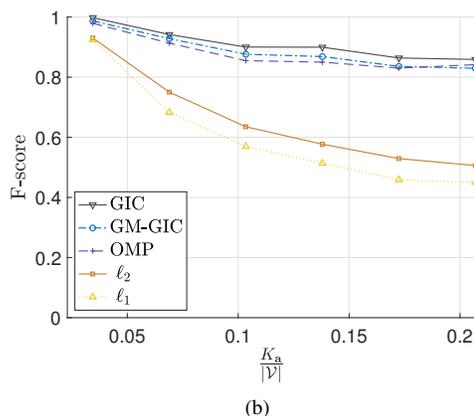
(b)

Fig. 3. Unobservable FDI attack identification: The F-score of the different methods versus the normalized attack, $\frac{||\mathbf{a}||}{K_{\mathbf{a}}}$, where $K_{\mathbf{a}} = 4$ and $\sigma_s^2 = 0.1$ (a); and versus the ratio of attacked elements, $\frac{K_{\mathbf{a}}}{\mathcal{V}}$, for $||\mathbf{a}|| = 1.2$ and $\sigma_s^2 = 0.05$ (b).

the following GIC statistic:

$$\begin{aligned}
\text{GIC}_{or}(\Lambda_i, \tau(|\Lambda_i|, |\mathcal{L}|)) &= \frac{1}{\sigma_{\mathbf{e}}^2}||\mathbf{P}_{\Lambda_i}(\Delta\mathbf{z}_{\mathcal{L}} - \mathbf{H}\Delta\boldsymbol{\theta})||^2 \\
&\quad - \tau(|\Lambda_i|, |\mathcal{L}|) + const. \qquad (53)
\end{aligned}$$

Fig. 4 shows that the F-score of the different methods, excluding the oracle GIC, decreases as $||\mathbf{H}\Delta\boldsymbol{\theta}||^2$ increases, due to the model mismatch. It can be seen that this degradation in the identification performance is significantly greater for the sparsity-based methods, ℓ_2 and ℓ_1, than for the proposed methods. Therefore, the proposed methods are more robust to a mismatched model in comparison to the existing methods. Finally, for lower values of $||\mathbf{H}\Delta\boldsymbol{\theta}||^2$ the proposed GIC method converges to the oracle GIC.

4) PSSE: In Fig. 5, the mean-squared-error (MSE) of the PSSE at time $t + 1$, $\hat{\boldsymbol{\theta}}_{t+1}$, of the different methods is presented versus the normalized attack norm, $\frac{||\mathbf{a}||}{K_{\mathbf{a}}}$, with the same parameters as in Fig. 3. The PSSE is implemented by (30) after substituting the selected support, Λ_s, that is obtained from the chosen identification method. In addition, the WLS estimator in (4), performed directly on the differential measurements in (12), is provided as an upper bound. Fig. 5 shows that the GIC method provides the lowest MSE for any scale of attack and that the GM-GIC method outperforms the remaining methods.
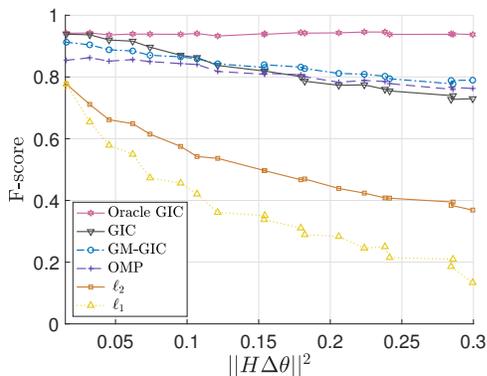
Fig. 4. Unobservable FDI attack identification: The F-score of the different methods with $K_{\mathbf{a}} = 4$ and $||\mathbf{a}|| = 0.05 K_{\mathbf{a}}$.

For lower degrees of attack, the OMP method is superior to the existing sparse techniques, while for a significant attack ($\frac{||a||}{K_{\mathbf{c}}} > 0.7$) the ℓ_2 and ℓ_1 methods obtain a lower MSE than the OMP method. Thus, in terms of state estimation for large-scale networks under a significant unobservable attack, the existing sparsity-based methods outperform the proposed OMP method. Nevertheless, our proposed GM-GIC method that incorporates the available graphical information achieves better estimation and identification performance.
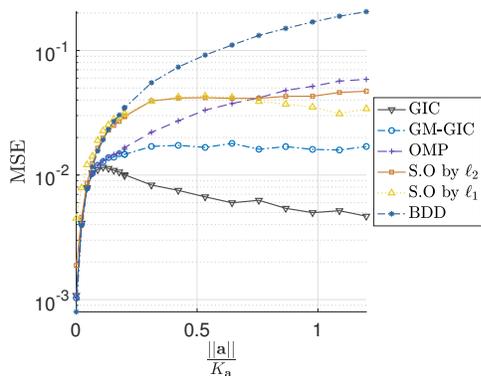

Fig. 5. PSSE in the presence of unobservable FDI attacks: The MSE of the different methods versus the normalized attack norm $\frac{||\mathbf{a}||}{K_{\mathbf{a}}}$ for $K_{\mathbf{a}} = 4$ and $\sigma_s^2 = 0.1$.

5) Run-time: In Fig. 6, the averaged run-time of the identification methods is presented versus the ratio of node elements attacked, $\frac{K_{\mathbf{a}}}{|\mathcal{V}|}$, for $\sigma_s^2 = 0.05$ and $||\mathbf{a}|| = 1.2$. It can be seen that the GIC method requires the highest run-time, which is significantly higher than those of the other methods. This makes the GIC method impractical for large networks. The CS methods [43], [45] have a higher averaged run-time than the proposed low-complexity methods, GM-GIC and OMP. As expected from the discussion in Subsection IV-B4, the GM-GIC method requires a significantly lower run-time than the GIC method, but a higher run-time than the OMP method.

VI. CONCLUSIONS

In this paper, we introduce novel methods for the identification of unobservable FDI attacks. The proposed methods are
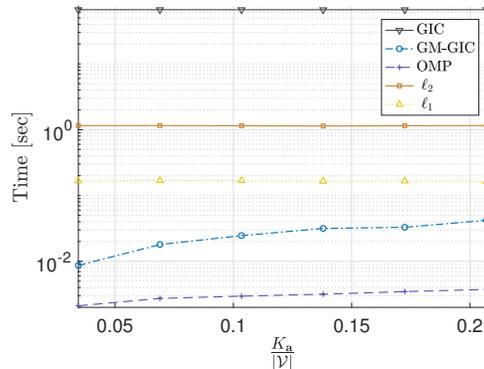

Fig. 6. Unobservable FDI attack identification: Run-time of the different methods versus the ratio of attacked elements, $\frac{K_{\mathbf{a}}}{\mathcal{V}}$, for $||\mathbf{a}|| = 1.2$ and $\sigma_s^2 = 0.05$.

based on defining structural constraints on both the attack and the typical load changes, and then formulating the identification problem as a model selection problem. We develop the GIC selection rule for the identification task, as well as two low-complexity methods: 1) the structural constrained OMP method, which is a modification to the standard OMP that accounts for the proposed structural and sparse constraints; and 2) the novel GM-GIC method that exploits the graph sparsity, graph Markovity, and the local behavior of graph connectivity in power systems. As by-products of the identification process, the proposed methods also enable the detection of unobservable FDI attacks and PSSE in the presence of these attacks. Our numerical simulations show that the proposed methods outperform existing methods for detection, identification, and PSSE, and are robust to the model assumptions. In addition, the PSSE conducted by the proposed methods significantly reduces the estimation error in comparison with the standard estimation. For both identification and PSSE, the GM-GIC method, that integrates the structural information regarding the underling graph behind the data, obtains better performance than that of the OMP method and the existing sparsity-based methods. The OMP and GM-GIC methods require shorter run-time than the GIC method.

APPENDIX A
PROOF OF PROPOSITION 1

Based on (24) and the noise statistics, $\Delta \mathbf{e}_{\mathcal{L}} \sim \mathcal{N}(0, \sigma_e^2 \mathbf{I})$, under each hypothesis, i, the log-likelihood function of the measurements, for a given $\Delta \boldsymbol{\theta}$, is

$$L(\Delta \mathbf{z}_{\mathcal{L}}; \mathbf{c}_{\Lambda_i}, \Delta \boldsymbol{\theta}) = -\frac{|\mathcal{L}|}{2} \ln 2\pi \sigma_e^2$$
$$-\frac{1}{2\sigma_{\mathbf{e}}^2} ||\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L}, \Lambda_i} \mathbf{c}_{\Lambda_i} - \mathbf{H}_{\mathcal{L}, \mathcal{V}} \Delta \boldsymbol{\theta}||^2. \quad (54)$$

Thus, the ML estimator of the state attack vector, $\mathbf{c}_{\Lambda_i}$, is obtained by maximizing (54) w.r.t. $\mathbf{c}_{\Lambda_i}$. Since (54) is a concave function w.r.t. $\mathbf{c}_{\Lambda_i}$, the ML estimator can be computed by

equating the derivative of (54) w.r.t. $\mathbf{c}_{\Lambda_i}$ to zero, as follows:

$$\left.\frac{\partial L(\Delta \mathbf{z}_{\mathcal{L}}; \mathbf{c}_{\Lambda_i})}{\partial \mathbf{c}_{\Lambda_i}}\right|_{\mathbf{c}_{\Lambda_i} = \hat{\mathbf{c}}_{\Lambda_i}^{\mathrm{ML}|i}}$$

$$= \frac{1}{2\sigma_{\mathbf{e}}^2} \mathbf{H}_{\mathcal{L},\Lambda_i}^T (\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\Lambda_i} \hat{\mathbf{c}}_{\Lambda_i}^{\mathrm{ML}|i} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta}) = \mathbf{0}, \quad (55)$$

which implies that the ML estimator is

$$\hat{\mathbf{c}}_{\Lambda_i}^{\mathrm{ML}|i} = (\mathbf{H}_{\mathcal{L},\Lambda_i}^T \mathbf{H}_{\mathcal{L},\Lambda_i})^{-1} \mathbf{H}_{\mathcal{L},\Lambda_i}^T (\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta}). \quad (56)$$

By using the projection matrix from (18), $\mathbf{P}_{\Lambda_i}$, and its orthogonal projection, $\mathbf{P}_{\Lambda_i}^{\perp}$, we can decompose $\mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta}$ as follows:

$$\mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta} = \mathbf{P}_{\Lambda_i} \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta} + \mathbf{P}_{\Lambda_i}^{\perp} \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta}. \quad (57)$$

By substituting the constraint from (24), $\mathbf{P}_{\Lambda_i} \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta} = \mathbf{0}$, in (57), one obtains

$$\mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta} = \mathbf{P}_{\Lambda_i}^{\perp} \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta}. \quad (58)$$

Substitution of (58) in (56) results in

$$\begin{aligned}\hat{\mathbf{c}}_{\Lambda_i}^{\mathrm{ML}|i} &= (\mathbf{H}_{\mathcal{L},\Lambda_i}^T \mathbf{H}_{\mathcal{L},\Lambda_i})^{-1} \mathbf{H}_{\mathcal{L},\Lambda_i}^T (\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{P}_{\Lambda_i}^{\perp} \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta}) \\ &= (\mathbf{H}_{\mathcal{L},\Lambda_i}^T \mathbf{H}_{\mathcal{L},\Lambda_i})^{-1} \mathbf{H}_{\mathcal{L},\Lambda_i}^T \Delta \mathbf{z}_{\mathcal{L}}, \end{aligned} \quad (59)$$

where the last equality is obtained from the properties of projection matrices. The generalized log likelihood is obtained by substituting (58) and (59) in (54):

$$L(\Delta \mathbf{z}_{\mathcal{L}}; \hat{\mathbf{c}}_{\Lambda_i}^{\mathrm{ML}|i}, \Delta \boldsymbol{\theta})$$

$$= -\frac{|\mathcal{L}|}{2} \ln 2\pi\sigma_e^2 - \frac{1}{2\sigma_{\mathbf{e}}^2} ||\mathbf{P}_{\Lambda_i}^{\perp} (\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta})||^2. \quad (60)$$

Additionally, from the properties of projection matrices, it can be verified that (see, e.g. p. 46 in [63])

$$||\mathbf{P}_{\Lambda_i}^{\perp} (\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta})||^2$$

$$= ||(\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta})||^2 - ||\mathbf{P}_{\Lambda_i} (\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta})||^2. \quad (61)$$

By substituting the constraint from (24), $\mathbf{P}_{\Lambda_i} \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta} = \mathbf{0}$, in (61) we obtain

$$||\mathbf{P}_{\Lambda_i}^{\perp} (\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta})||^2$$

$$= ||(\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta})||^2 - ||\mathbf{P}_{\Lambda_i} \Delta \mathbf{z}_{\mathcal{L}}||^2, \quad (62)$$

$\forall \Lambda_i \in \mathcal{G}_{K_{\mathbf{c}}}$. By substituting (62) in (60) we obtain

$$L(\Delta \mathbf{z}_{\mathcal{L}}; \hat{\mathbf{c}}_{\Lambda_i}^{\mathrm{ML}|i}, \Delta \boldsymbol{\theta}) = -\frac{|\mathcal{L}|}{2} \ln 2\pi\sigma_e^2$$

$$-\frac{1}{2\sigma_{\mathbf{e}}^2} ||(\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta})||^2 + \frac{1}{2\sigma_{\mathbf{e}}^2} ||\mathbf{P}_{\Lambda_i} \Delta \mathbf{z}_{\mathcal{L}}||^2, \quad (63)$$

$\forall \Lambda_i \in \mathcal{G}_{K_{\mathbf{c}}}$. By substituting (63) into (25), we obtain that the GIC statistic satisfies

$$\mathrm{GIC}(\Lambda_i, \tau(|\Lambda_i|, |\mathcal{L}|)$$

$$= -|\mathcal{L}| \ln 2\pi\sigma_e^2 - \frac{1}{\sigma_{\mathbf{e}}^2} ||(\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta})||^2$$

$$+ \frac{1}{\sigma_{\mathbf{e}}^2} ||\mathbf{P}_{\Lambda_i} \Delta \mathbf{z}_{\mathcal{L}}||^2 - \tau(|\Lambda_i|, |\mathcal{L}|), \quad (64)$$

where $i = 0, 1, \ldots, |\mathcal{G}_{K_{\mathbf{c}}}|$. Since the term

$$-|\mathcal{L}| \ln 2\pi\sigma_e^2 - \frac{1}{\sigma_{\mathbf{e}}^2} ||(\Delta \mathbf{z}_{\mathcal{L}} - \mathbf{H}_{\mathcal{L},\mathcal{V}} \Delta \boldsymbol{\theta})||^2$$

is independent of the hypothesis, i.e. is not a function of Λ_i, maximizing the GIC rule from (64), w.r.t. $i = 0, 1, \ldots, |\mathcal{G}_{K_{\mathbf{c}}}|$,

is equivalent to maximizing the r.h.s. of (27), w.r.t. the same candidate sets, $i = 0, 1, \ldots, |\mathcal{G}_{K_{\mathbf{c}}}|$.

APPENDIX B
PROOF OF PROPOSITION 2

In Proposition 1 in [48] it was shown that under the assumption that the injected powers in the nodes are Gaussian distributed and mutually independent, which holds under the model in (12), the columns of the topology matrix satisfy

$$\mathbf{H}_{\mathcal{V},m}^T \mathbf{H}_{\mathcal{V},k} = 0, \quad \forall k, m \in \mathcal{V}, \ d(k,m) > 2. \quad (65)$$

Similarly, the columns of the subtopology matrix associated with the load buses satisfy

$$\mathbf{H}_{\mathcal{L},m}^T \mathbf{H}_{\mathcal{L},k} = 0, \quad \forall k, m \in \mathcal{V}, \ d(k,m) > 2. \quad (66)$$

By substituting (18) with $\Lambda_i = \{m\}$ in the l.h.s. of (35) we obtain

$$\mathbf{P}_{\{m\}} \mathbf{H}_{\mathcal{L},k} c_k = \mathbf{H}_{\mathcal{L},m} (\mathbf{H}_{\mathcal{L},m}^T \mathbf{H}_{\mathcal{L},m})^{-1} \mathbf{H}_{\mathcal{L},m}^T \mathbf{H}_{\mathcal{L},k} c_k. \quad (67)$$

Thus, by substituting (66) in (67) we obtain (35). In addition, by using the properties of projection matrices it can be verified (see Theorem 2.22 in [63]) that

$$||\mathbf{P}_{\{m\}} \mathbf{H}_{\mathcal{L},k} \mathbf{c}_k|| \leq ||\mathbf{H}_{\mathcal{L},k} \mathbf{c}_k||, \quad \forall k, m \in \mathcal{V}. \quad (68)$$

Thus, by substituting (20), with $\Lambda_i = \{k\}$, in the l.h.s. of (68) we obtain (36).

REFERENCES

[1] A. Gomez-Exposito and A. Abur, *Power system state estimation: theory and implementation*. CRC press, 2004.

[2] G. B. Giannakis, V. Kekatos, N. Gatsis, S.-J. Kim, H. Zhu, and B. F. Wollenberg, "Monitoring and optimization for power grids: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 107–128, 2013.

[3] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, p. 13, 2011.

[4] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1630–1638, 2017.

[5] V. Kekatos, G. B. Giannakis, and R. Baldick, "Online energy price matrix factorization for power grid topology tracking," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1239–1248, 2015.

[6] X. Liu, Z. Bao, D. Lu, and Z. Li, "Modeling of local false data injection attacks with reduced network information," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1686–1696, 2015.

[7] Z.-H. Yu and W.-L. Chin, "Blind false data injection attack using pca approximation method in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1219–1226, 2015.

[8] S. Grotas, Y. Yakoby, I. Gera, and T. Routtenberg, "Power systems topology and state estimation by graph blind source separation," *IEEE Trans. Signal Process*, vol. 67, no. 8, pp. 2036–2051, 2019.

[9] J. Kim, L. Tong, and R. J. Thomas, "Subspace methods for data attack on state estimation: A data driven approach," *IEEE Trans. Signal Process*, vol. 63, no. 5, pp. 1102–1114, 2014.

[10] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets," in *SmartGridComm*, 2010, pp. 226–231.

[11] ——, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, 2011.

[12] L. Jia, J. Kim, R. J. Thomas, and L. Tong, "Impact of data quality on real-time locational marginal price," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 627–636, 2014.

[13] Y. Yuan, Z. Li, and K. Ren, "Quantitative analysis of load redistribution attacks in power systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1731–1738, 2012.

[14] J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao, "On false data injection attacks against distributed energy routing in smart grid," in *Int. Conf. Cyber-Physical. Syst*, 2012, pp. 183–192.

[15] R. B. Bobba, K. M. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T. J. Overbye, "Detecting false data injection attacks on dc state estimation," in *Preprints of the First Workshop on Secure Control Systems, CPSWEEK*, vol. 2010, 2010.

[16] S. Bi and Y. J. Zhang, "Graphical methods for defense against false-data injection attacks on power system state estimation," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1216–1227, 2014.

[17] ——, "Defending mechanisms against false-data injection attacks in the power system state estimation," in *GLOBECOM*, 2011, pp. 1162–1167.

[18] R. Deng, G. Xiao, and R. Lu, "Defending against false data injection attacks on power system state estimation," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 198–207, 2015.

[19] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 326–333, 2011.

[20] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1608–1615, 2006.

[21] J. Kim and L. Tong, "On phasor measurement unit placement against state and topology attacks," in *SmartGridComm*, 2013, pp. 396–401.

[22] S. Gong, Z. Zhang, H. Li, and A. D. Dimitrovski, "Time stamp attack in smart grid: Physical mechanism and damage analysis," *arXiv preprint arXiv:1201.2578*, 2012.

[23] A. Mohammadi and K. N. Plataniotis, "Noncircular attacks on phasor measurement units for state estimation in smart grid," *IEEE J. Sel. Topics Signal Process*, vol. 12, no. 4, pp. 777–789, 2018.

[24] K. R. Davis, K. L. Morrow, R. Bobba, and E. Heine, "Power flow cyber attacks and perturbation-based defense," in *SmartGridComm*, 2012, pp. 342–347.

[25] K. L. Morrow, E. Heine, K. M. Rogers, R. B. Bobba, and T. J. Overbye, "Topology perturbation for detecting malicious data injection," in *Hawaii International Conference on System Sciences*, 2012, pp. 2104–2113.

[26] M. A. Rahman, E. Al-Shaer, and R. B. Bobba, "Moving target defense for hardening the security of the power system state estimation," in *Proceedings of the First ACM Workshop on Moving Target Defense*, 2014, pp. 59–68.

[27] J. Tian, R. Tan, X. Guan, and T. Liu, "Hidden moving target defense in smart grids," in *Proceedings of the 2nd Workshop on Cyber-Physical Security and Resilience in Smart Grids*, 2017, pp. 21–26.

[28] S. Lakshminarayana and D. K. Yau, "Cost-benefit analysis of moving-target defense in power grids," in *International Conference on Dependable Systems and Networks (DSN)*, 2018, pp. 139–150.

[29] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1644–1652, 2017.

[30] Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2505–2516, 2017.

[31] R. Xu, R. Wang, Z. Guan, L. Wu, J. Wu, and X. Du, "Achieving efficient detection against false data injection attacks in smart grid," *IEEE Access*, vol. 5, pp. 13 787–13 798, 2017.

[32] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Online detection of stealthy false data injection attacks in power system state estimation," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1636–1646, 2018.

[33] M. A. Abu-El-Magd and N. K. Sinha, "Short-term load demand modeling and forecasting: A review," *IEEE Trans. Syst., Man, and Cybernetics*, vol. 12, no. 3, pp. 370–382, May 1982.

[34] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[35] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[36] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, ser. Compressed Sensing: Theory and Applications. Cambridge University Press, 2012.

[37] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*, 1st ed. New York, NY, USA: Cambridge University Press, 2015.

[38] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2215–2224, Nov. 2012.

[39] J. Chen, Y. Zhao, A. Goldsmith, and H. Poor, "Line outage detection in power transmission networks via message passing algorithms," in *Asilomar Conference*, Nov. 2014, pp. 350–354.

[40] S. Kibria, J. Kim, and R. Raich, "Sparse error correction with multiple measurement vectors," in *IEEE SSP Workshop*, June 2016, pp. 1–5.

[41] T. Routtenberg and Y. C. Eldar, "Centralized identification of imbalances in power networks with synchrophasor data," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1981–1992, 2017.

[42] W. Xu, M. Wang, J. F. Cai, and A. Tang, "Sparse error correction from nonlinear measurements with applications in bad data detection for power networks," *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6175–6187, Dec. 2013.

[43] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, "Detecting false data injection attacks on power grid by sparse optimization," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 612–621, 2014.

[44] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Chin, and Z. Fan, "Sparse malicious false data injection attacks and defense mechanisms in smart grids," *IEEE Trans. Ind. Informat.*, vol. 11, no. 5, pp. 1–12, 2015.

[45] P. Gao, M. Wang, J. H. Chow, S. G. Ghiocel, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky, "Identification of successive unobservable cyber data attacks in power systems through matrix decomposition," *IEEE Trans. Signal Process*, vol. 64, no. 21, pp. 5557–5570, 2016.

[46] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, 2004.

[47] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process*, vol. 41, no. 12, pp. 3397–3415, 1993.

[48] H. Sedghi and E. Jonckheere, "Statistical structure learning to ensure data integrity in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1924–1933, 2015.

[49] E. Drayer and T. Routtenberg, "Detection of false data injection attacks in smart grids based on graph signal processing," *IEEE Syst. J.*, 2019.

[50] ——, "Detection of false data injection attacks in power systems with graph fourier transform," in *Glob. Conf. Sig. and Info. Process. (GlobalSIP)*, 2018, pp. 890–894.

[51] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, 2011.

[52] J. Jiang and Y. Qian, "Defense mechanisms against data injection attacks in smart grid networks," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 76–82, 2017.

[53] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 382–390, 2011.

[54] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, 2000.

[55] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proc. IEEE*, vol. 75, no. 12, pp. 1558–1573, 1987.

[56] M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing*. Springer Science & Business Media, 2010.

[57] H. Sedghi and E. Jonckheere, "On the conditional mutual information in the gaussian–markov structured grids," in *Information and Control in Networks*. Springer, 2014, pp. 277–297.

[58] Z. Wang, A. Scaglione, and R. J. Thomas, "Generating statistically correct random topologies for testing smart grid communication and control networks," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 28–39, 2010.

[59] P. Stoica, Y. Selen, and J. Li, "On information criteria and the generalized likelihood ratio test of model order selection," *IEEE Signal Process. Lett.*, vol. 11, no. 10, pp. 794–797, Oct. 2004.

[60] MATLAB, *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc., 2010.

[61] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb 2011.

[62] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing and Management*, vol. 45, pp. 427–437, 07 2009.

[63] H. Yanai, K. Takeuchi, and Y. Takane, "Projection matrices," in *Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition*. Springer, 2011, pp. 25–54.