Modified Matrix Completion-based Detection of Stealthy Data Manipulation Attacks in Low Observable Distribution Systems

James Ranjith Kumar Rajasekaran, Balasubramaniam Natarajan and Anil Pahwa

***Abstract*—A composite detection technique against stealthy data manipulations is developed in this paper for distribution networks that are low observable. Attack detection strategies typically rely on state estimation which becomes challenging when limited measurements are available. In this paper, a modified matrix completion approach provides estimates of the system state and its error variances for the locations in the network where measurements are unavailable. Using the error statistics and their corresponding state estimates, bad data detection can be carried out using the chi-squared test. The proposed approach employs a moving target defence strategy (MTD) where the network parameters are perturbed through distributed flexible AC transmission system (D-FACTS) devices such that stealthy data manipulation attacks can be exposed in the form of bad data. Thus, the bad data detection approach developed in this paper can detect stealthy attacks using the MTD strategy. This technique is implemented on 37-bus and 123-bus three-phase unbalanced distribution networks to demonstrate the attack detection accuracy even for a low observable system.**

***Index Terms*—Bad Data detection, Distribution system, Matrix completion, Moving Target Defence, State Estimation.**

I. Introduction

Over the last decade, there have been growing concerns related to the cybersecurity of power networks. Even though conventional bad data detection techniques can detect inconsistencies in power system measurements, [1] shows how to evade such residual-based detection techniques. This circumvention process can be achieved as long as the manipulated measurements satisfy the network governing equations. Since conventional bad data detection schemes cannot detect such stealthy attacks, there is a need for more efficient, dedicated approaches that can better analyse the measurements for the presence of abnormalities. The research efforts related to cybersecurity for power distribution networks are summarised in [2]. The earliest among those studies [3] shows the impacts of manipulating the information about the overcurrent relay's position. Initial research attempts that are available in [4], [5] demonstrate how false data injection (FDI) attacks can be executed on three-phase balanced distribution systems. For observable unbalanced distribution networks, an approach is proposed in [6] to detect FDI attacks with system transients.

MTD has been a well-known approach used to identify the presence of stealthy data manipulations in electric networks [7]. With the MTD approach, the system parameters are perturbed intentionally by the operator such that the attacker is unaware of such changes in the system. Thus the attack vector injected by the attacker in the measurements will not comply with the system's governing equations with the perturbed network parameters. In a fully observable network, the verification of the measurements corresponding to the system parameters can be carried out using conventional bad data techniques. The MTD strategy proposed in [8] perturbs the nodal power injections at controllable points and since the manipulations made by the attacker do not come in line with these perturbations, they can be detected in the form of bad data.

Distributed flexible AC transmission systems (D-FACTS) are special devices used in power system networks because of their ability to change the line reactance from 80% to 120% of its nominal value [9]. D-FACTS devices are traditionally used for loss minimization, voltage control and optimal power dispatch in the power system networks [10]. The advantages of using D-FACTS devices include increased operational efficiency, reduced voltage fluctuations and better economical operation of the grid. In addition to the aforementioned applications of D-FACTS devices, it gained interest among researchers to employ the D-FACTS devices to induce MTD in a fully observable system and detect stealthy data manipulation attacks. The detection approaches given in [11], [12] are the initial attempts in the literature to employ the MTD strategy through D-FACTS devices to detect stealthy data manipulations in electric transmission systems. Following these works, various studies (which are listed in [13]) have been carried out to utilise the advantages of D-FACTS and detect stealthy attacks. In [14], a simplified AC model of transmission systems is used to develop a reactance perturbation algorithm so that it minimises the power loss in the transmission grid as well as reduces the losses in its lines. To account the effect of noisy measurements, a robust MTD is developed in [15] for detecting stealthy attacks in transmission systems. An MTD-based countermeasure is developed in [16] against attackers who have limited meter information in the given transmission system. In a recent study [17], it is shown that D-FACTS devices can be used to induce MTD in a fully observable system to detect data manipulation attacks in addition to the as well as carry out optimal power flow for the grid.

D-FACTS which were popularly used in transmission systems have been scaled down to distribution systems recently so that their advantages can be attained in such distribution

All authors are with Kansas State University, Manhattan, USA.

This material is based upon work partly supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Solar Energy Technologies Office, under Award # DE-EE0008767.

networks [18]. But, in the context of stealthy attack detection, only a handful of literature is available for D-FACTS-based MTD deployment in distribution networks. An implementation of the D-FACTS-based MTD strategy in distribution systems is presented in [19] where the line parameters are perturbed not only to expose the concealed data manipulation attacks but also to achieve optimal operation costs in the network. It is to be noted that [19] aims to compute the optimal set points for the D-FACTS devices so that attacks can be detected with an observable set of measurements. Thus, existing MTD-based stealthy attack detection approaches assume that the system is fully observable with the given set of measurements. Such an assumption may not be valid in distribution networks as the measurement devices are placed only at specific points of interest and typically such networks are not fully observable [20]. To implement MTD strategies, first, it is necessary to develop a technique to detect bad data in the measurements of the given unobservable distribution network such that it can work in conjunction with the MTD strategy to detect FDI attacks. Bad data detection techniques for observable networks use residual-based analysis with prior knowledge of probability distribution in the measurement noise. But for unobservable distribution networks, conventional bad data detection approaches cannot be carried out as a significant number of nodes in the network do not have measurements and hence the noise statistics of the system state are unavailable apriori.

To carry out the state estimation in unobservable distribution networks, traditionally pseudo measurements are used (see [21], [22]) with high error variances as compared to actual measurements. Such pseudo-measurements are typically obtained from forecasts or historical data. It is to be noted that, if pseudo measurements are used at points which have a prominent influence in the network, the estimation accuracy is significantly reduced [23]. The state estimation approach for low observable distribution networks[1] given in [24] utilizes a neural network approximation of the minimum mean squared error estimator. The deep neural network-based state estimator in [24] obtains its parameters from the offline training section which requires a large amount of data for the training process. If the network parameters are changed with devices like D-FACTS, it needs to be completely trained again with the information of the changed network which takes several hours for training. Thus, the approach given in [24] may not be applicable for detecting FDI attacks in streaming measurements. Recently, a matrix completion-based approach has been proposed to impute the missing data in unobservable networks by exploiting the low-rank property of the measurements. This matrix completion approach is used in [25] to recover the missing data in an unobservable set of PMU measurements. Techniques given in [20], [26] have adapted the low-rank matrix completion approach to estimate the system state of distribution networks that have a low observable set of measurements.

For observable networks, the closed-form solution of the weighted least squares (WLS) state estimator is used to compute the noise covariance in the estimates and with that bad data detection can be carried out directly. Such an approach is not applicable in unobservable systems because the WLS state estimator will not give a unique solution if the system is unobservable and perform poorly as shown in many prior works [27]. On the other hand, state estimators that are designed for unobservable systems such as matrix completion or compressed sensing approaches do not have a closed-form solution in the available literature. Hence the noise covariance terms are unavailable and thus the conventional bad data detection scheme cannot be applied.

To the best of our knowledge, this paper is the first to provide a closed-form solution for matrix completion-based state estimation for unobservable distribution networks through an analytical modification. With the closed-form solution, the proposed approach is not only able to provide the state estimates at a faster rate, but it also provides the values of the noise co-variance terms of the estimates. With the noise statistics of the estimates, a chi-square test can be carried out to detect the presence of any bad data in the measurements. To detect stealthy data manipulations, MTD is used due to its ability to change such stealthy attacks into bad data which can be detected with the proposed approach. This objective is attained with a multi-stage approach that is developed in this paper. The first stage involves the modification of the matrix completion technique from which the unknown states and their noise statistics are estimated. Based on these estimated values, a chi-squared test is used to detect the bad data in the measurements. To detect FDI attacks using the MTD strategy, D-FACTS devices are considered across various lines in the network. With such D-FACTS devices, the effective reactance of the line can be varied for MTD against stealthy FDI attacks. As the attacker is unaware of the network changes due to the line parameter perturbation by the MTD strategy, attacks are transformed into bad data injections which can be detected with the proposed modified matrix completion-based approach. The main contributions of this paper are listed as follows:

1) A linear relation is developed between the unavailable values and the measured values in an unobservable distribution network. This linear formulation is derived by modifying the matrix completion approach where a hybrid matrix is employed.
2) An estimation algorithm is developed based on the previously derived linear formulation that can impute the unavailable values as well as their error co-variance terms using the measurements from an unobservable distribution network and its noise variances as inputs.
3) With the estimated system state and its error co-variance terms implementing an MTD strategy where stealthy attacks are transformed into bad data.

The rest of the paper is organised as follows: In Section 2, a brief introduction to the FDI attack model and MTD-based detection strategy is given. The conventional matrix completion technique is described in Section 3. Section 4 details the proposed approach and its performance is demonstrated

[1]Throughout this paper, the term 'low observable' and 'unobservable' are used interchangeably

through simulation results in Section 5. In the final section, the concluding remarks are provided.

II. FDI Attack and MTD-based Detection

The proposed approach aims to detect stealthy or non-stealthy type data manipulation attacks regardless of the intention of the attacker. Among these types of attacks, stealthy data manipulations can be considered the most challenging as they can evade conventional bad data detection schemes and create operational damage as intended by the attacker. Therefore, this section discusses how a stealthy false data injection attack can be executed and how the moving target defence strategy can transform stealthy false data into bad data.

A. Attack Model

To explain the FDI attack model, let $\boldsymbol{z}$ be the vector of measurements and $\hat{\boldsymbol{x}}$ be the estimated state vector of the given network. Without loss of generality, let $\boldsymbol{H}$ be the function that linearly transforms the true system state with the additive noise to the measurement vector $\boldsymbol{z}$. With these definitions, the estimated residual can be written as:

$$\boldsymbol{r} = \boldsymbol{z} - \boldsymbol{H} \cdot \hat{\boldsymbol{x}} \tag{1}$$

Since the conventional bad data approach uses residuals for the detection process, the attacker can manipulate the measurements that do not affect the resulting residual value and the data manipulation stays undetected. Let $\tilde{\boldsymbol{x}}$ be the manipulated state vector intended by the attacker. The intention of such manipulation in the state vector could be either

1) to hide the overloading of one or more lines from the operators and impede them not to take any necessary corrective action; or
2) to disguise the nominal state in the system in a form of an emergency state such that the operators can be misled to take an inappropriate control action.

From (1), the necessary condition for the manipulated measurement $\tilde{\boldsymbol{z}}$ that does not affect the residual can be given as:

$$\boldsymbol{z} - \boldsymbol{H} \cdot \hat{\boldsymbol{x}} = \tilde{\boldsymbol{z}} - \boldsymbol{H} \cdot \tilde{\boldsymbol{x}} \tag{2}$$

To execute such an attack, it is assumed that the network model of the distribution system is available to the attacker. Such an assumption is reasonable as the configurations of the line conductors and the types of transformers that are used in the network are known to the distribution utility at the planning stage itself. Hence with that information, the network model of the distribution network can be obtained for the control centre operations. Moreover, this assumption is commonly found in the related literature [28]. Due to the vast geographic footprint of the network and the computational constraints, it is reasonable to assume that an attacker cannot hijack all the measurements in a given system. However, it is assumed that the attacker can hijack a subset of measurements in the network.

B. MTD detection strategy

In the previous subsection, it can be noticed that the attacker relies on the linear function $\boldsymbol{H}$ which is formed using the physical network information. The operator can leverage this fact to ensure that the stealthy attack gets detected. This underlying principle is used in the moving target defence (MTD) strategy where the network parameters are perturbed to a value known only to the operator from its nominal value. Let $\hat{\boldsymbol{H}}$ be the resulting linear function after perturbing the network parameters. The measurements will conform to the perturbed network parameters and the value of residual with the no-attack condition is given as follows:

$$\hat{\boldsymbol{r}} = \boldsymbol{z} - \hat{\boldsymbol{H}} \cdot \hat{\boldsymbol{x}} \tag{3}$$

During an attack scenario, let the attacker manipulates the measurement vector to $\tilde{\boldsymbol{z}}$ such that the state vector becomes $\tilde{\boldsymbol{x}}$. It is assumed that the attacker is unaware of perturbation made in the network parameters and uses the linear function $\boldsymbol{H}$ for generating the attack vector. Under such data manipulation, the residual computed at the control centre can be written as:

$$\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{z}} - \hat{\boldsymbol{H}} \cdot \tilde{\boldsymbol{x}} \tag{4}$$

Since the generated attack vector adheres to the nominal network parameters and hence to $\boldsymbol{H}$, the residual value given in (4) can be rewritten as:

$$\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{z}} - \boldsymbol{H} \cdot \tilde{\boldsymbol{x}} + \left(\boldsymbol{H} - \hat{\boldsymbol{H}}\right) \cdot \hat{\boldsymbol{x}} \tag{5}$$

$$= \hat{\boldsymbol{r}} + \left(\boldsymbol{H} - \hat{\boldsymbol{H}}\right) \cdot \hat{\boldsymbol{x}} \tag{6}$$

Thus by implementing the MTD strategy which involves perturbing the network parameters, the computed residual significantly changes during the stealthy attack scenario as compared to the no-attack scenario.

Network parameters could be perturbed in many ways. In this paper, we assume that D-FACTS devices employed across various lines in the distribution network can change the line reactance from 80% to 120% of the nominal value. By perturbing the set points of the D-FACTS devices, the linear mapping function $\boldsymbol{H}$ that represents the network information can be changed and through that, the MTD strategy can be implemented. But since distribution systems typically have an unobservable set of measurements, bad data detection approaches such as the chi-squared test cannot be directly implemented. In the next section, the conventional matrix completion approach is explained which will be used to develop the proposed composite state estimation and bad data detection for an unobservable system. Such an approach can be used along with the MTD strategy to detect stealthy FDI attacks.

III. Matrix Completion based State Estimation

In this section, the conventional matrix completion approach [25] is detailed. For ease in understanding, consider a complete measurement data matrix, $\boldsymbol{M} \in \mathbb{R}^{n_1 \times n_2}$ in which only the elements given by the set of indices $\Psi \subset [1, n_1] \times [1, n_2]$ are

available for the estimation process. With this definition, the observation matrix can be denoted as $\boldsymbol{M}_\Psi$ such that

$$(\boldsymbol{M}_\Psi)_{j,k} = \begin{cases} \boldsymbol{M}_{j,k} & \text{, for } (j,k) \in \Psi \\ 0 & \text{, for } (j,k) \notin \Psi \end{cases} \tag{7}$$

where the available measurements are assigned to their respective indices and unavailable terms are set to zero. The conventional matrix completion approach aims to obtain a reasonable estimate of these unavailable terms to complete the matrix. To achieve this objective, the low-rank property of the measurement data matrix is exploited. This low-rank property results due to the spatial correlation of the measurements. With this consideration, the matrix completion approach is formulated as a rank minimization problem:

$$\arg \min_{\boldsymbol{X} \in \mathbb{R}^{n_1 \times n_2}} \quad \operatorname{rank}(\boldsymbol{X}) \tag{8a}$$

$$\text{such that} \quad \boldsymbol{X}_\Psi = \boldsymbol{M}_\Psi \tag{8b}$$

where $\boldsymbol{X}$, the variable in the optimisation problem provides an estimate of the complete data matrix $\boldsymbol{M}$. It can be noticed that the rank function in the objective of (8) results in a non-convex optimization problem. Hence this problem is convexified by using the nuclear norm of the decision variable $\boldsymbol{X}$ as:

$$\arg \min_{\boldsymbol{X} \in \mathbb{R}^{n_1 \times n_2}} \quad \|\boldsymbol{X}\|_* \tag{9a}$$

$$\text{such that} \quad \boldsymbol{X}_\Psi = \boldsymbol{M}_\Psi \tag{9b}$$

The definition of the nuclear norm of the matrix $\boldsymbol{X}$ can be recalled as the sum of its singular values. That is,

$$\|\boldsymbol{X}\|_* = \sum_i \sigma_i(\boldsymbol{X}) \tag{10}$$

where $\sigma_i(\boldsymbol{X})$ is the i^{th} singular value of matrix $\boldsymbol{X}$. The resulting convex problem can be solved using methods like truncated nuclear norm regularization. This problem can be expanded to incorporate known relationships between the elements of the data matrix. Such a consideration is quite plausible as the physical system where the measurements are taken can be modelled using its governing equations. In addition to incorporating the system physics in the problem formulation, the noise content which is typically present in the measurements can also be accounted for. Under such considerations, the formulation given in (9) is extended as:

$$\arg \min_{\boldsymbol{X} \in \mathbb{R}^{n_1 \times n_2}} \quad \|\boldsymbol{X}\|_* \tag{11a}$$

$$\text{such that} \quad \|\boldsymbol{X}_\Psi - \boldsymbol{M}_\Psi\|_F \leq \alpha \tag{11b}$$

$$\|\boldsymbol{g}(\boldsymbol{X})\| \leq \beta \tag{11c}$$

where $\boldsymbol{g}(\cdot)$ is the vector of functions that model the physical system from which the measurements are taken. The parameters α and β are tuned to a non-negative value following the level of noise present in the measurements. Typically, the problem given in (11) is solved numerically since the analytical solution for the nuclear norm of the variable $\boldsymbol{X}$ is currently not attempted in the literature. Thus it is difficult to obtain the analytical relationship between the known values and the unknown terms that are imputed. Therefore, with the formulation given in (11), it is possible to quantify the amount of noise in the imputed unknown terms. Quantification of noise in the estimates is necessary to carry out residual-based bad data analysis which serves as a basis for MTD based FDI detection process.

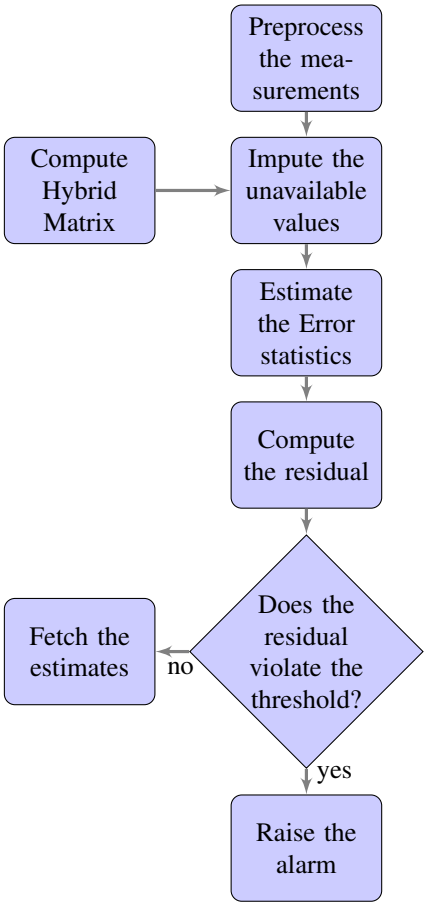


Fig. 1: Flow chart of the proposed approach

IV. Proposed Approach

This section develops a linear formulation for (11) from which the unknown terms are imputed. It also provides a framework to estimate the noise in the imputed terms. The flow chart of the proposed approach is illustrated in Fig. 1. This section is divided into four parts where the first three parts present the three stages of the proposed approach to detect the presence of bad data in an unobservable distribution system. The last part of this section deals with how the MTD strategy can translate stealthy FDI attacks into bad data which can be detected using the proposed approach.

A. Handing Power System Measurements

Traditionally, distribution networks take the values of voltage magnitudes and real power injections [20] as measurements at the nodes of interest across the network. These voltage magnitudes and real power injections naturally possess a non-linear relation with each other. The proposed approach

first transforms the values of voltage magnitudes and real power injections into equivalent voltage and current phasors so that its linear relation can be exploited. For transforming the voltage magnitudes into equivalent voltage phasors, it is assumed that in distribution systems, the voltage angle differences between the buses are negligible. This assumption is valid as the feeders in the distribution networks tend to have a low X/R ratio. With this assumption, the angular difference corresponding to its phase can be used with the voltage magnitude value $\left|\mathbf{v}_i^\phi\right|$ measured at phase ϕ of bus i to obtain its equivalent voltage phasor as:

$$\mathbf{v}_i^\phi = \left|\mathbf{v}_i^\phi\right| \angle \delta^\phi, \quad \phi \in \{a, b, c\} \tag{12}$$

where

$$\delta = \begin{cases} 0 & \text{if } \phi = a \\ -120 & \text{if } \phi = b \\ 120 & \text{if } \phi = c \end{cases} \tag{13}$$

The power factor of the loads in the distribution network tends to be consistent and hence it is assumed to be a constant with a value $\cos\boldsymbol{\theta}$. With this assumption, an approximate value of reactive power, $\mathbf{q}_i^\phi$, can be recovered using the real power measurement, $\mathbf{q}_i^\phi$, at phase ϕ of bus i as:

$$\mathbf{q}_i^\phi = \mathbf{p}_i^\phi \tan\boldsymbol{\theta} \tag{14}$$

With this recovered approximate voltage phasor, $\mathbf{v}_i^\phi$ along with the values of real power injection, $\mathbf{p}_i^\phi$, and reactive power injection, $\mathbf{q}_i^\phi$, the approximate value of current injection at phase ϕ of bus i can be obtained as:

$$\mathbf{i}_i^\phi = \left(\frac{\mathbf{p}_i^\phi + \mathrm{j}\mathbf{q}_i^\phi}{\mathbf{v}_i^\phi}\right)^* \tag{15}$$

B. Formation of Bus Hybrid Matrices

For a given distribution system, let $\mathcal{S}$ be the set of nodes in the reference bus corresponding to all its three phases and $\mathcal{B}$ be the set of all the remaining nodes in the network excluding the nodes corresponding to the reference bus. It is to be noted that to handle an unbalanced distribution network, each phase of a given bus is considered a node. Let $\mathbf{v} \in \mathbb{C}^{|\mathcal{B}|}$ and $\mathbf{i} \in \mathbb{C}^{|\mathcal{B}|}$ be the complex vectors of nodal voltages and current injections at all nodes given by set $\mathcal{B}$. Similarly, let $\mathbf{v}_\mathcal{S} \in \mathbb{C}^{|\mathcal{S}|}$ be the voltage at the nodes in the reference bus given by set $\mathcal{S}$. With these definitions, the relationship between the current injections and nodal voltages can be written as:

$$\mathbf{i} = \mathbf{Y}_\mathcal{S} \cdot \mathbf{v}_\mathcal{S} + \mathbf{Y} \cdot \mathbf{v} \tag{16}$$

where $\mathbf{Y}$ and $\mathbf{Y}_\mathcal{S}$ be the block forms of the nodal admittance matrix for the nodes given by set $\mathcal{B}$ and $\mathcal{S}$ respectively. The expression given in (16) can be written in linear form as:

$$\mathbf{Z} \cdot \mathbf{i} = \Delta\mathbf{v} \tag{17}$$

where,

$$\Delta\mathbf{v} = \mathbf{Z} \cdot \mathbf{Y}_\mathcal{B} \cdot \mathbf{v}_\mathcal{S} + \mathbf{v} \tag{18}$$

and consider that $\mathbf{Z} = \mathbf{Y}^{-1}$ is the nodal impedance matrix with $\mathcal{S}$ as the reference.

Consider a distribution system where the bus voltage phasor deviations, $\Delta\mathbf{v}$, and the current injection phasors, $\mathbf{i}$, are available at the buses given by set $\mathcal{K}$. Let $\mathcal{U}$ be the set of buses at which measurements are unavailable (complement of set $\mathcal{K}$). With this definition, the relation between voltage deviations and the current injections that are given in (17) can be rewritten as:

$$\begin{bmatrix} \mathbf{Z}_{\mathcal{K}\mathcal{K}} & \mathbf{Z}_{\mathcal{K}\mathcal{U}} \\ \mathbf{Z}_{\mathcal{U}\mathcal{K}} & \mathbf{Z}_{\mathcal{U}\mathcal{U}} \end{bmatrix} \begin{bmatrix} \mathbf{i}_\mathcal{K} \\ \mathbf{i}_\mathcal{U} \end{bmatrix} = \begin{bmatrix} \Delta\mathbf{v}_\mathcal{K} \\ \Delta\mathbf{v}_\mathcal{U} \end{bmatrix} \tag{19}$$

where

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{\mathcal{K}\mathcal{K}} & \mathbf{Z}_{\mathcal{K}\mathcal{U}} \\ \mathbf{Z}_{\mathcal{U}\mathcal{K}} & \mathbf{Z}_{\mathcal{U}\mathcal{U}} \end{bmatrix} \tag{20}$$

$$\Delta\mathbf{v}_\mathcal{U} = \mathbf{Z}_{\mathcal{U}\mathcal{K}}\mathbf{i}_\mathcal{K} + \mathbf{Z}_{\mathcal{U}\mathcal{U}}\mathbf{i}_\mathcal{U} \tag{21}$$

$$= \mathbf{Z}_{\mathcal{U}\mathcal{K}}\mathbf{Z}_{\mathcal{K}\mathcal{K}}^{-1}\left(\Delta\mathbf{v}_\mathcal{K} - \mathbf{Z}_{\mathcal{K}\mathcal{U}}\mathbf{i}_\mathcal{U}\right) + \mathbf{Z}_{\mathcal{U}\mathcal{U}}\mathbf{i}_\mathcal{U} \tag{22}$$

$$= \mathbf{Z}_{\mathcal{U}\mathcal{K}}\mathbf{Z}_{\mathcal{K}\mathcal{K}}^{-1}\Delta\mathbf{v}_\mathcal{K} - \left(\mathbf{Z}_{\mathcal{U}\mathcal{K}}\mathbf{Z}_{\mathcal{K}\mathcal{K}}^{-1}\mathbf{Z}_{\mathcal{K}\mathcal{U}} - \mathbf{Z}_{\mathcal{U}\mathcal{U}}\right)\mathbf{i}_\mathcal{U} \tag{23}$$

$$\Delta\mathbf{v}_\mathcal{U} + \left(\mathbf{Z}_{\mathcal{U}\mathcal{K}}\mathbf{Z}_{\mathcal{K}\mathcal{K}}^{-1}\mathbf{Z}_{\mathcal{K}\mathcal{U}} - \mathbf{Z}_{\mathcal{U}\mathcal{U}}\right)\mathbf{i}_\mathcal{U} = \mathbf{Z}_{\mathcal{U}\mathcal{K}}\mathbf{Z}_{\mathcal{K}\mathcal{K}}^{-1}\Delta\mathbf{v}_\mathcal{K} \tag{24}$$

The bus admittance matrix (excluding the slack bus) can be obtained from the bus impedance matrix as $\mathbf{Y} = \mathbf{Z}^{-1}$.

$$\begin{bmatrix} \mathbf{Y}_{\mathcal{K}\mathcal{K}} & \mathbf{Y}_{\mathcal{K}\mathcal{U}} \\ \mathbf{Y}_{\mathcal{U}\mathcal{K}} & \mathbf{Y}_{\mathcal{U}\mathcal{U}} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{v}_\mathcal{K} \\ \Delta\mathbf{v}_\mathcal{U} \end{bmatrix} = \begin{bmatrix} \mathbf{i}_\mathcal{K} \\ \mathbf{i}_\mathcal{U} \end{bmatrix} \tag{25}$$

where

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{\mathcal{K}\mathcal{K}} & \mathbf{Y}_{\mathcal{K}\mathcal{U}} \\ \mathbf{Y}_{\mathcal{U}\mathcal{K}} & \mathbf{Y}_{\mathcal{U}\mathcal{U}} \end{bmatrix} \tag{26}$$

$$\mathbf{i}_\mathcal{U} = \mathbf{Y}_{\mathcal{U}\mathcal{K}}\Delta\mathbf{v}_\mathcal{K} + \mathbf{Y}_{\mathcal{U}\mathcal{U}}\Delta\mathbf{v}_\mathcal{U} \tag{27}$$

$$= \mathbf{Y}_{\mathcal{U}\mathcal{K}}\mathbf{Y}_{\mathcal{K}\mathcal{K}}^{-1}\left(\mathbf{i}_\mathcal{K} - \mathbf{Y}_{\mathcal{K}\mathcal{U}}\Delta\mathbf{v}_\mathcal{U}\right) + \mathbf{Y}_{\mathcal{U}\mathcal{U}}\Delta\mathbf{v}_\mathcal{U} \tag{28}$$

$$= \mathbf{Y}_{\mathcal{U}\mathcal{K}}\mathbf{Y}_{\mathcal{K}\mathcal{K}}^{-1}\mathbf{i}_\mathcal{K} - \left(\mathbf{Y}_{\mathcal{U}\mathcal{K}}\mathbf{Y}_{\mathcal{K}\mathcal{K}}^{-1}\mathbf{Y}_{\mathcal{K}\mathcal{U}} - \mathbf{Y}_{\mathcal{U}\mathcal{U}}\right)\Delta\mathbf{v}_\mathcal{U} \tag{29}$$

$$\left(\mathbf{Y}_{\mathcal{U}\mathcal{K}}\mathbf{Y}_{\mathcal{K}\mathcal{K}}^{-1}\mathbf{Y}_{\mathcal{K}\mathcal{U}} - \mathbf{Y}_{\mathcal{U}\mathcal{U}}\right)\Delta\mathbf{v}_\mathcal{U} + \mathbf{i}_\mathcal{U} = \mathbf{Y}_{\mathcal{U}\mathcal{K}}\mathbf{Y}_{\mathcal{K}\mathcal{K}}^{-1}\mathbf{i}_\mathcal{K} \tag{30}$$

$$\mathbf{H}^\mathrm{v}\Delta\mathbf{v}_\mathcal{U} + \mathbf{H}^\mathrm{i}\mathbf{i}_\mathcal{U} = \begin{bmatrix} \mathbf{G}^\mathrm{v} \cdot \Delta\mathbf{v}_\mathcal{K} \\ \mathbf{G}^\mathrm{i} \cdot \mathbf{i}_\mathcal{K} \end{bmatrix} \tag{31}$$

where

$$\mathbf{H}^\mathrm{v} = \begin{bmatrix} \mathbf{I} \\ \mathbf{Y}_{\mathcal{U}\mathcal{K}}\mathbf{Y}_{\mathcal{K}\mathcal{K}}^{-1}\mathbf{Y}_{\mathcal{K}\mathcal{U}} - \mathbf{Y}_{\mathcal{U}\mathcal{U}} \end{bmatrix} \tag{32}$$

$$\mathbf{H}^\mathrm{i} = \begin{bmatrix} \mathbf{Z}_{\mathcal{U}\mathcal{K}}\mathbf{Z}_{\mathcal{K}\mathcal{K}}^{-1}\mathbf{Z}_{\mathcal{K}\mathcal{U}} - \mathbf{Z}_{\mathcal{U}\mathcal{U}} \\ \mathbf{I} \end{bmatrix} \tag{33}$$

$$\mathbf{G}^\mathrm{v} = \mathbf{Z}_{\mathcal{U}\mathcal{K}}\mathbf{Z}_{\mathcal{K}\mathcal{K}}^{-1} \tag{34}$$

$$\mathbf{G}^\mathrm{i} = \mathbf{Y}_{\mathcal{U}\mathcal{K}}\mathbf{Y}_{\mathcal{K}\mathcal{K}}^{-1} \tag{35}$$

The above expression provides the linear relationship between voltages and current terms taken at measured and unmeasured nodes. To provide a better explanation for why the bus hybrid matrix is employed instead of the traditionally used bus admittance matrix, let $|\mathcal{K}|$ and $|\mathcal{U}|$ be the total number of nodes with and without measurements, respectively. The number of constraints for the system model in the matrix completion optimisation problem with bus admittance matrix is $3|\mathcal{K}| + |\mathcal{U}|$ as seen in (11) ($|\mathcal{K}|$ constraints for (11b) and $|\mathcal{K}| + |\mathcal{U}|$ constraints for (11c)). Whereas with the bus hybrid matrix, the number of constraints is reduced to $2|\mathcal{K}|$ which

eventually reduces the number of Lagrange coefficients and effectively reduces the computational burden for the estimation process. It is to be noted that (31) is achieved through algebraic manipulations of (25) and not through linear approximations. This concept of the bus hybrid matrix is extended from the earlier work reported in [29].

C. Imputation of Unavailable Measurements

To further relax the problem, the linear model in (17), which is in the complex domain, is transformed into the real domain using the following definitions:

$$\Delta \mathsf{v}_i = \left[\begin{array}{cc} \Re\{\Delta \mathbf{v}_i\} & \Im\{\Delta \mathbf{v}_i\} \end{array} \right]^T \tag{36}$$

$$\mathsf{i}_i = \left[\begin{array}{cc} \Re\{\mathbf{i}_i\} & \Im\{\mathbf{i}_i\} \end{array} \right]^T \tag{37}$$

$$\mathsf{Z}_{ij} = \left[\begin{array}{cc} \Re\{\mathbf{Z}_{ij}\} & -\Im\{\mathbf{Z}_{ij}\} \\ \Im\{\mathbf{Z}_{ij}\} & \Re\{\mathbf{Z}_{ij}\} \end{array} \right] \tag{38}$$

$$\mathsf{H}^{\mathsf{v}}_{ij} = \left[\begin{array}{cc} \Re\left\{\mathbf{H}^{\mathsf{v}}_{ij}\right\} & -\Im\left\{\mathbf{H}^{\mathsf{v}}_{ij}\right\} \\ \Im\left\{\mathbf{H}^{\mathsf{v}}_{ij}\right\} & \Re\left\{\mathbf{H}^{\mathsf{v}}_{ij}\right\} \end{array} \right] \tag{39}$$

$$\mathsf{H}^{\mathsf{i}}_{ij} = \left[\begin{array}{cc} \Re\left\{\mathbf{H}^{\mathsf{i}}_{ij}\right\} & -\Im\left\{\mathbf{H}^{\mathsf{i}}_{ij}\right\} \\ \Im\left\{\mathbf{H}^{\mathsf{i}}_{ij}\right\} & \Re\left\{\mathbf{H}^{\mathsf{i}}_{ij}\right\} \end{array} \right] \tag{40}$$

$$\mathsf{G}^{\mathsf{v}}_{ij} = \left[\begin{array}{cc} \Re\left\{\mathbf{H}^{\mathsf{v}}_{ij}\right\} & -\Im\left\{\mathbf{H}^{\mathsf{v}}_{ij}\right\} \\ \Im\left\{\mathbf{H}^{\mathsf{v}}_{ij}\right\} & \Re\left\{\mathbf{H}^{\mathsf{v}}_{ij}\right\} \end{array} \right] \tag{41}$$

$$\mathsf{G}^{\mathsf{i}}_{ij} = \left[\begin{array}{cc} \Re\left\{\mathbf{G}^{\mathsf{v}}_{ij}\right\} & -\Im\left\{\mathbf{G}^{\mathsf{v}}_{ij}\right\} \\ \Im\left\{\mathbf{G}^{\mathsf{v}}_{ij}\right\} & \Re\left\{\mathbf{G}^{\mathsf{v}}_{ij}\right\} \end{array} \right] \tag{42}$$

The conventional matrix completion-based state estimator can now be restated as:

$$\arg \min_{\Delta \mathsf{v}_{\mathcal{U}}, \mathsf{i}_{\mathcal{U}}} \left\| \left[\begin{array}{cc} \Delta \mathsf{v}_{\mathcal{K}} & \mathsf{i}_{\mathcal{K}} \\ \Delta \mathsf{v}_{\mathcal{U}} & \mathsf{i}_{\mathcal{U}} \end{array} \right] \right\|_* \tag{43a}$$

$$\text{such that} \quad \left[\begin{array}{c} \Delta \mathsf{v}_{\mathcal{K}} \\ \Delta \mathsf{v}_{\mathcal{U}} \end{array} \right] = \left[\begin{array}{cc} \mathsf{Z}_{\mathcal{K}\mathcal{K}} & \mathsf{Z}_{\mathcal{K}\mathcal{U}} \\ \mathsf{Z}_{\mathcal{U}\mathcal{K}} & \mathsf{Z}_{\mathcal{U}\mathcal{U}} \end{array} \right] \left[\begin{array}{c} \mathsf{i}_{\mathcal{K}} \\ \mathsf{i}_{\mathcal{U}} \end{array} \right] \tag{43b}$$

Using a hybrid matrix (a combination of bus impedance and bus admittance matrix $\mathbf{Y}$), the equality constraint can be written in linear form and the matrix completion problem becomes:

$$\arg \min_{\Delta \mathsf{v}_{\mathcal{U}}, \mathsf{i}_{\mathcal{U}}} \left\| \left[\begin{array}{cc} \Delta \mathsf{v}_{\mathcal{K}} & \mathsf{i}_{\mathcal{K}} \\ \Delta \mathsf{v}_{\mathcal{U}} & \mathsf{i}_{\mathcal{U}} \end{array} \right] \right\|_* \tag{44a}$$

$$\text{such that} \quad \left[\begin{array}{c} \mathsf{G}^{\mathsf{v}} \cdot \Delta \mathsf{v}_{\mathcal{K}} \\ \mathsf{G}^{\mathsf{i}} \cdot \mathsf{i}_{\mathcal{K}} \end{array} \right] = \mathsf{H}^{\mathsf{v}} \Delta \mathsf{v}_{\mathcal{U}} + \mathsf{H}^{\mathsf{i}} \mathsf{i}_{\mathcal{U}} \tag{44b}$$

The singular values of the matrix $\mathbf{X}$ are the square roots of the eigenvalues of its projection matrix $\mathbf{X}^T\mathbf{X}$ which can be written as:

$$\sigma(\mathbf{X}) = \sqrt{\lambda\left(\mathbf{X}^T\mathbf{X}\right)} \tag{45}$$

Utilizing this definition of singular value given in (45) along with the definition of nuclear norm given in (10), the matrix completion problem can be rewritten as:

$$\min_{\Delta \mathsf{v}_{\mathcal{U}}, \mathsf{i}_{\mathcal{U}}} \sum_i \sqrt{\lambda_i\left(\mathbf{X}^T\mathbf{X}\right)} \tag{46a}$$

$$\text{such that} \quad \left[\begin{array}{c} \mathsf{G}^{\mathsf{v}} \cdot \Delta \mathsf{v}_{\mathcal{K}} \\ \mathsf{G}^{\mathsf{i}} \cdot \mathsf{i}_{\mathcal{K}} \end{array} \right] = \mathsf{H}^{\mathsf{v}} \Delta \mathsf{v}_{\mathcal{U}} + \mathsf{H}^{\mathsf{i}} \mathsf{i}_{\mathcal{U}} \tag{46b}$$

where

$$\mathbf{X} = \left[\begin{array}{cc} \Delta \mathsf{v}_{\mathcal{K}} & \mathsf{i}_{\mathcal{K}} \\ \Delta \mathsf{v}_{\mathcal{U}} & \mathsf{i}_{\mathcal{U}} \end{array} \right] \tag{47}$$

Note that the matrix completion problem given in (46) is the transformed version of the original problem given in (43) and hence both of these problems give the same solution. Further, since the original problem given in (43) is convex, the transformed version of the matrix completion problem given in (46) is also convex. Numerical techniques are typically used in convex optimization tools to solve such problems. But such solution approaches do not yield a direct analytical relation between the measured values of the indices given by set $\mathcal{K}$ and the imputed values using the matrix completion approach of the set of indices $\mathcal{U}$.

The objective of this paper is to obtain the relationship between the elements of the set of indices $\mathcal{U}$ and $\mathcal{K}$ so that bad data analysis can be carried out. To achieve it, a heuristic-based modification to the formulation in (46) is proposed as follows:

$$\min_{\Delta \mathsf{v}_{\mathcal{U}}, \mathsf{i}_{\mathcal{U}}} \sum_i \lambda_i\left(\mathbf{X}^T\mathbf{X}\right) \tag{48a}$$

$$\text{such that} \quad \left[\begin{array}{c} \mathsf{G}^{\mathsf{v}} \cdot \Delta \mathsf{v}_{\mathcal{K}} \\ \mathsf{G}^{\mathsf{i}} \cdot \mathsf{i}_{\mathcal{K}} \end{array} \right] = \mathsf{H}^{\mathsf{v}} \Delta \mathsf{v}_{\mathcal{U}} + \mathsf{H}^{\mathsf{i}} \mathsf{i}_{\mathcal{U}} \tag{48b}$$

In this heuristic, the objective function is modified which aims to minimise the sum of squares of singular values instead of the sum of singular values. By the definition given in (45), the objective function turns out to be the sum of the eigenvalues of $\mathbf{X}^T\mathbf{X}$ which is the same as the trace of $\mathbf{X}^T\mathbf{X}$. This can be expanded as:

$$\sum_i \lambda_i\left(\mathbf{X}^T\mathbf{X}\right) = \Delta \mathsf{v}_{\mathcal{K}}^T \Delta \mathsf{v}_{\mathcal{K}} + \Delta \mathsf{v}_{\mathcal{U}}^T \Delta \mathsf{v}_{\mathcal{U}} + \mathsf{i}_{\mathcal{K}}^T \mathsf{i}_{\mathcal{K}} + \mathsf{i}_{\mathcal{U}}^T \mathsf{i}_{\mathcal{U}} \tag{49}$$

Since the terms $\Delta \mathsf{v}_{\mathcal{K}}^T \Delta \mathsf{v}_{\mathcal{K}}$ and $\mathsf{i}_{\mathcal{K}}^T \mathsf{i}_{\mathcal{K}}$ are constants, they can be removed from the objective function and the simplified form of the proposed formulation is given as:

$$\arg \min_{\Delta \mathsf{v}_{\mathcal{U}}, \mathsf{i}_{\mathcal{U}}} \left(\Delta \mathsf{v}_{\mathcal{U}}^T \Delta \mathsf{v}_{\mathcal{U}} + \mathsf{i}_{\mathcal{U}}^T \mathsf{i}_{\mathcal{U}}\right) \tag{50a}$$

$$\text{such that} \quad \left[\begin{array}{c} \mathsf{G}^{\mathsf{v}} \cdot \Delta \mathsf{v}_{\mathcal{K}} \\ \mathsf{G}^{\mathsf{i}} \cdot \mathsf{i}_{\mathcal{K}} \end{array} \right] = \mathsf{H}^{\mathsf{v}} \Delta \mathsf{v}_{\mathcal{U}} + \mathsf{H}^{\mathsf{i}} \mathsf{i}_{\mathcal{U}} \tag{50b}$$

Using $\mathbf{w}$ as the Lagrange variable, it can be reformulated as an unconstrained problem:

$$\min_{\Delta \mathsf{v}_{\mathcal{U}}, \mathsf{i}_{\mathcal{U}}, \mathbf{w}} f\left(\Delta \mathsf{v}_{\mathcal{U}}, \mathsf{i}_{\mathcal{U}}, \mathbf{w}\right) \tag{51}$$

where

$$\begin{aligned} f = {} & \Delta \mathsf{v}_{\mathcal{U}}^T \Delta \mathsf{v}_{\mathcal{U}} + \mathsf{i}_{\mathcal{U}}^T \mathsf{i}_{\mathcal{U}} \\ & + \mathbf{w}^T \left(\mathsf{H}^{\mathsf{v}} \Delta \mathsf{v}_{\mathcal{U}} + \mathsf{H}^{\mathsf{i}} \mathsf{i}_{\mathcal{U}} - \left[\begin{array}{c} \mathsf{G}^{\mathsf{v}} \cdot \Delta \mathsf{v}_{\mathcal{K}} \\ \mathsf{G}^{\mathsf{i}} \cdot \mathsf{i}_{\mathcal{K}} \end{array} \right] \right) \end{aligned} \tag{52}$$

The derivatives of this new modified objective function are given as:

$$\nabla_{\Delta \mathsf{v}_\mathcal{U}} f = 2\Delta \mathsf{v}_\mathcal{U} + \mathsf{H}^{\mathsf{v}^T}\mathbf{w} \tag{53}$$

$$\nabla_{\mathsf{i}_\mathcal{U}} f = 2\mathsf{i}_\mathcal{U} + \mathsf{H}^{\mathsf{i}^T}\mathbf{w} \tag{54}$$

$$\nabla_{\mathbf{w}} f = \mathsf{H}^{\mathsf{v}}\Delta \mathsf{v}_\mathcal{U} + \mathsf{H}^{\mathsf{i}}\mathsf{i}_\mathcal{U} - \begin{bmatrix} \mathsf{G}^{\mathsf{v}} \cdot \Delta \mathsf{v}_\mathcal{K} \\ \mathsf{G}^{\mathsf{i}} \cdot \mathsf{i}_\mathcal{K} \end{bmatrix} \tag{55}$$

By applying the first order condition for optimality i.e., $\nabla f = 0$, the linear formulation for estimating the unknown values at the set of nodes given by $\mathcal{U}$ using the known set of nodes $\mathcal{K}$ can be given as:

$$\begin{bmatrix} 2\mathbf{I} & \mathbf{0} & \mathsf{H}^{\mathsf{v}^T} \\ \mathbf{0} & 2\mathbf{I} & \mathsf{H}^{\mathsf{i}^T} \\ \mathsf{H}^{\mathsf{v}} & \mathsf{H}^{\mathsf{i}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \mathsf{v}_\mathcal{U} \\ \mathsf{i}_\mathcal{U} \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathsf{G}^{\mathsf{v}} \cdot \Delta \mathsf{v}_\mathcal{K} \\ \mathsf{G}^{\mathsf{i}} \cdot \mathsf{i}_\mathcal{K} \end{bmatrix} \tag{56}$$

$$\begin{bmatrix} \Delta \mathsf{v}_\mathcal{U} \\ \mathsf{i}_\mathcal{U} \\ \mathbf{w} \end{bmatrix} = \mathbf{J} \cdot \begin{bmatrix} \mathbf{0} \\ \Delta \mathsf{v}_\mathcal{K} \\ \mathsf{i}_\mathcal{K} \end{bmatrix} \tag{57}$$

where

$$\mathbf{J} = \begin{bmatrix} 2\mathbf{I} & \mathbf{0} & \mathsf{H}^{\mathsf{v}^T} \\ \mathbf{0} & 2\mathbf{I} & \mathsf{H}^{\mathsf{i}^T} \\ \mathsf{H}^{\mathsf{v}} & \mathsf{H}^{\mathsf{i}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathsf{G}^{\mathsf{v}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathsf{G}^{\mathsf{i}} \end{bmatrix} \tag{58}$$

For simplification, $\mathbf{J}$ can be segmented as:

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\Delta \mathsf{v}_\mathcal{U}, \mathbf{0}} & \mathbf{J}_{\Delta \mathsf{v}_\mathcal{U}, \Delta \mathsf{v}_\mathcal{K}} & \mathbf{J}_{\Delta \mathsf{v}_\mathcal{U}, \mathsf{i}_\mathcal{K}} \\ \mathbf{J}_{\mathsf{i}_\mathcal{U}, \mathbf{0}} & \mathbf{J}_{\mathsf{i}_\mathcal{U}, \Delta \mathsf{v}_\mathcal{K}} & \mathbf{J}_{\mathsf{i}_\mathcal{U}, \mathsf{i}_\mathcal{K}} \\ \mathbf{J}_{\mathbf{w}, \mathbf{0}} & \mathbf{J}_{\mathbf{w}, \Delta \mathsf{v}_\mathcal{K}} & \mathbf{J}_{\mathbf{w}, \mathsf{i}_\mathcal{K}} \end{bmatrix} \tag{59}$$

The simplified formulation for the imputation is given as:

$$\begin{bmatrix} \Delta \mathsf{v}_\mathcal{U} \\ \mathsf{i}_\mathcal{U} \end{bmatrix} = \tilde{\mathbf{J}} \begin{bmatrix} \Delta \mathsf{v}_\mathcal{K} \\ \mathsf{i}_\mathcal{K} \end{bmatrix} \tag{60}$$

where

$$\tilde{\mathbf{J}} = \begin{bmatrix} \mathbf{J}_{\Delta \mathsf{v}_\mathcal{U}, \Delta \mathsf{v}_\mathcal{K}} & \mathbf{J}_{\Delta \mathsf{v}_\mathcal{U}, \mathsf{i}_\mathcal{K}} \\ \mathbf{J}_{\mathsf{i}_\mathcal{U}, \Delta \mathsf{v}_\mathcal{K}} & \mathbf{J}_{\mathsf{i}_\mathcal{U}, \mathsf{i}_\mathcal{K}} \end{bmatrix} \tag{61}$$

D. Bad Data Detection

With the diagonal covariance matrix $\Sigma_\mathcal{K}$ corresponding to the noise components in measured values, the complete noise co-variance matrix for both measured and imputed values are given as:

$$\Sigma = \begin{bmatrix} \Sigma_\mathcal{K} & \Sigma_\mathcal{K} \cdot \tilde{\mathbf{J}}^T \\ \tilde{\mathbf{J}} \cdot \Sigma_\mathcal{K} & \tilde{\mathbf{J}} \cdot \Sigma_\mathcal{K} \cdot \tilde{\mathbf{J}}^T \end{bmatrix} \tag{62}$$

To find the near positive definite symmetric matrix, $\tilde{\Sigma}$, for the obtained noise co-variance matrix, Σ, the algorithm given in [30] is employed. With this definition, it can be written as:

$$\begin{bmatrix} \Delta \mathsf{v}_\mathcal{K} \\ \mathsf{i}_\mathcal{K} \\ \Delta \mathsf{v}_\mathcal{U} \\ \mathsf{i}_\mathcal{U} \end{bmatrix} = \begin{bmatrix} \Delta \bar{\mathsf{v}}_\mathcal{K} \\ \bar{\mathsf{i}}_\mathcal{K} \\ \Delta \bar{\mathsf{v}}_\mathcal{U} \\ \bar{\mathsf{i}}_\mathcal{U} \end{bmatrix} + \eta, \quad \eta \sim \mathcal{N}\left(\mathbf{0}, \tilde{\Sigma}\right) \tag{63}$$

To calculate the residuals so that the chi-squared test can be carried out, the mean values $\Delta\bar{\mathbf{v}}$, $\bar{\mathbf{i}}$ is required at buses given by sets $\mathcal{U}$ and $\mathcal{K}$. To eliminate further estimation processes, the following condition is exploited:

$$\begin{bmatrix} \Delta \bar{\mathsf{v}}_\mathcal{K} \\ \Delta \bar{\mathsf{v}}_\mathcal{U} \end{bmatrix} = \begin{bmatrix} \mathsf{Z}_{\mathcal{K}\mathcal{K}} & \mathsf{Z}_{\mathcal{K}\mathcal{U}} \\ \mathsf{Z}_{\mathcal{U}\mathcal{K}} & \mathsf{Z}_{\mathcal{U}\mathcal{U}} \end{bmatrix} \begin{bmatrix} \bar{\mathsf{i}}_\mathcal{K} \\ \bar{\mathsf{i}}_\mathcal{U} \end{bmatrix} \tag{64}$$

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} = \mathbf{T} \begin{bmatrix} \Delta \bar{\mathsf{v}}_\mathcal{K} \\ \bar{\mathsf{i}}_\mathcal{K} \\ \Delta \bar{\mathsf{v}}_\mathcal{U} \\ \bar{\mathsf{i}}_\mathcal{U} \end{bmatrix} \tag{65}$$

where

$$\mathbf{T} = \begin{bmatrix} -\mathbf{I} & \mathsf{Z}_{\mathcal{K}\mathcal{K}} & \mathbf{0} & \mathsf{Z}_{\mathcal{K}\mathcal{U}} \\ \mathbf{0} & \mathsf{Z}_{\mathcal{U}\mathcal{K}} & -\mathbf{I} & \mathsf{Z}_{\mathcal{U}\mathcal{U}} \end{bmatrix} \tag{66}$$

With this definition of $\mathbf{T}$, the residuals correspond to:

$$\mathbf{r} = \mathbf{T} \begin{bmatrix} \Delta \mathsf{v}_\mathcal{K} \\ \mathsf{i}_\mathcal{K} \\ \Delta \mathsf{v}_\mathcal{U} \\ \mathsf{i}_\mathcal{U} \end{bmatrix}, \quad \mathbf{r} \sim \mathcal{N}\left(\mathbf{0}, \Omega\right) \tag{67}$$

where

$$\Omega = \mathbf{T} \cdot \tilde{\Sigma} \cdot \mathbf{T}^T \tag{68}$$

Since the residual $\mathbf{r}$ has normal distribution, the chi-square $\mathbf{r}^T\Omega^{-1}\mathbf{r}$ follows a chi-squared distribution. The presence of bad data in the measured values and the imputed values can be identified if the chi-squared test i.e., $\mathbf{r}^T\Omega^{-1}\mathbf{r} > \chi^2_{\nu,p}$ is not satisfied. The parameters of this test ν and p denote the degrees of freedom and detection probability respectively. As there are $2|\mathcal{K}|$ linearly independent states in this problem formulation, the degree of freedom is chosen as $\nu = 2|\mathcal{K}|$. The detection probability is typically chosen as $p = 0.95$.

E. Algorithmic Representation

The proposed approach is segmented into three separate algorithms based on their functionalities and their pseudo codes are written in Algorithms 1, 2 and 3. For a given network topology and meter placement information in a distribution system, $\tilde{\mathbf{J}}$ matrix is formed using Algorithm 1. The state estimation process is carried out with Algorithm 2 with the measurements taken at $\mathcal{K}$ set of buses. Algorithm 3 provides the process of bad data detection which is used to detect FDI attacks if the MTD strategy is employed.

Algorithm 1 Formation of $\tilde{\mathbf{J}}$ Matrix

Require: $\mathcal{K}, \mathcal{U}$
Obtain $\mathbf{Z}$ using the bus building algorithm
Compute $\mathbf{Y} = \mathbf{Z}^{-1}$
Compute $\mathbf{H}^{\mathsf{v}}$, $\mathbf{H}^{\mathsf{i}}$, $\mathbf{G}^{\mathsf{v}}$ and $\mathbf{G}^{\mathsf{i}}$ using (32)-(35), respectively
Compute H^{v}, H^{i}, G^{v} and G^{i} using (39)-(42), respectively
Compute $\mathbf{J}$ using (58)
Compute $\tilde{\mathbf{J}}$ using (61)

Algorithm 2 Estimation of $\Delta \mathrm{v}_{\mathcal{U}}$ and $\mathrm{i}_{\mathcal{U}}$

Require: $|\mathbf{v}_{\mathcal{K}}|$, $\mathbf{p}_{\mathcal{K}}$, $\tilde{\mathbf{J}}$
Assume $\cos\boldsymbol{\theta}$
Compute $\mathbf{v}_{\mathcal{K}}$ using (12)
Compute $\mathbf{i}_{\mathcal{K}}$ using (14) and (15)
Compute $\Delta\mathbf{v}_{\mathcal{K}}$ from (18)
Convert $\Delta\mathbf{v}_{\mathcal{K}}$ to real domain, $\Delta\mathrm{v}_{\mathcal{K}}$ using (36)
Convert $\mathbf{i}_{\mathcal{K}}$ to real domain, $\mathrm{i}_{\mathcal{K}}$ using (37)
With $\tilde{\mathbf{J}}$, estimate $\Delta\mathrm{v}_{\mathcal{U}}$ and $\mathrm{i}_{\mathcal{U}}$ using (60)

Algorithm 3 Bad Data Detection

Require: $\tilde{\mathbf{J}}$, $\Sigma_{\mathcal{K}}$, $\Delta\mathrm{v}_{\mathcal{K}}$, $\mathrm{i}_{\mathcal{K}}$, $\Delta\mathrm{v}_{\mathcal{U}}$ and $\mathrm{i}_{\mathcal{U}}$
With $\tilde{\mathbf{J}}$ and $\Sigma_{\mathcal{K}}$, Compute Σ using (62)
Compute $\tilde{\Sigma}$ by adopting [30] such that $\tilde{\Sigma} \approx \Sigma$ and $\tilde{\Sigma} \succ 0$
Compute $\mathbf{T}$ using (66)
With $\Sigma_{\mathcal{K}}$, $\Delta\mathrm{v}_{\mathcal{K}}$, $\mathrm{i}_{\mathcal{K}}$, $\Delta\mathrm{v}_{\mathcal{U}}$ and $\mathrm{i}_{\mathcal{U}}$, compute $\mathbf{r}$ using (67)
Compute Ω using (68)
if $\mathbf{r}^T\Omega^{-1}\mathbf{r} > \chi^2_{\nu,p}$ **then**
Raise Alarm
end if

V. Simulation Results

The performance of the proposed approach is demonstrated on the IEEE 37-bus and 123-bus unbalanced distribution systems with unobservable set of measurements which are subjected to both non-stealthy and stealthy FDI attacks. For brevity, in both test cases, the available node is defined as the node where the measured values of voltage magnitude and the real power injection are available. In both the test cases, the noise in the measured values is assumed to be Gaussian with zero mean and standard deviation of 1%. Such a consideration is quite practical for SCADA measurements [26]. To recover the reactive power component as explained in section IV-A, it is assumed that the overall value of the load power factor is constant and available to the operator. Hence, the load power factor in both the IEEE 37-bus and the 123-bus systems is considered to be 0.9 lagging.

A. Performance of Imputation Algorithm

The voltage estimates from the proposed approach are first validated by considering 15 three-phase available nodes in the 37-bus system and 60 three-phase available nodes in the 123-bus system. It is trivial to see that such systems are unobservable in the context of the least squares-based state estimation approach as the number of measurements is less than half of the number of all the nodes in the network. The estimated voltage magnitudes at phase 'a' for the 37-bus system using the proposed approach given in Algorithm 2 are plotted in Fig. 2. In this plot, the true values (TV) and the estimated values using the conventional matrix completion (CMC) approach (which is given in [20]) are plotted along with the results of the proposed method for comparison. It can be noticed that the proposed approach provides similar results as compared to the conventional matrix completion approach. Similar validation is carried out on the 123-bus system and the voltage magnitudes

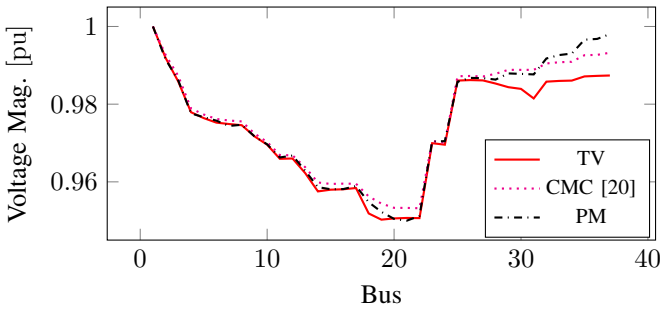


Fig. 2: Phase 'a' Voltage Magnitude for the 37-bus system with 15 available buses.

#1	No Attack with γ=0.0
#2	Non-stealthy Attack on single bus with γ=0.0
#3	Stealthy Attack on single target bus with γ=0.0
#4	Stealthy Attack on single target bus with γ=0.4
#5	Uncoordinated Attack on multiple target buses with γ=0.0
#6	Coordinated Attack on multiple target buses with γ=0.0
#7	Coordinated Attack on multiple target buses with γ=0.4

TABLE I: Testing conditions and its corresponding indications

at phase an estimated using the proposed method (PM) and conventional matrix completion (CMC) approach along with its true values (TV) are plotted in Fig. 3. These plots provide the conclusion that the proposed approach behaves similarly to the conventional matrix completion approach. In other words, the estimation accuracy is not much compromised with the proposed heuristic-based modification in the matrix completion approach.

In order to quantify the overall estimation accuracy of the proposed method (PM), a popular metric known as Mean Absolute Percentage Error (MAPE) [26] is used which can be defined as:

$$MAPE = \frac{1}{3}\sum_{\phi\in\{a,b,c\}}\left(\frac{1}{|\mathcal{U}|}\sum_{j\in\mathcal{U}}\frac{\left|\mathbf{v}_j^{\phi}\right| - \left|\tilde{\mathbf{v}}_j^{\phi}\right|}{\left|\mathbf{v}_j^{\phi}\right|}\right). \quad (69)$$

where $\left|\tilde{\mathbf{v}}_j^{\phi}\right|$ and $\left|\mathbf{v}_j^{\phi}\right|$ are the estimated and the true value of voltage magnitude at phase ϕ of bus j. The value of $MAPE$ is computed for both the proposed method (PM) and conventional matrix completion (CMC) approach using the estimates of voltage magnitudes by changing the number of measured buses from the lowest possible value to the maximum limit where the system reaches closer to observability. The calculated values of $MAPE$ are plotted in Figures 4 and 5 for 37-bus and 123-bus systems respectively. These plots indicate that the proposed method has a similar order of estimation accuracy as compared to the conventional matrix completion. Moreover, the proposed method relies on solving a linear system which takes less computation effort as compared to the conventional matrix completion as it relies on solving a convex optimisation problem numerically.

B. Performance of Detection Algorithm

This part of the paper focuses on validating the performance of the proposed approach for detecting anomalies in the

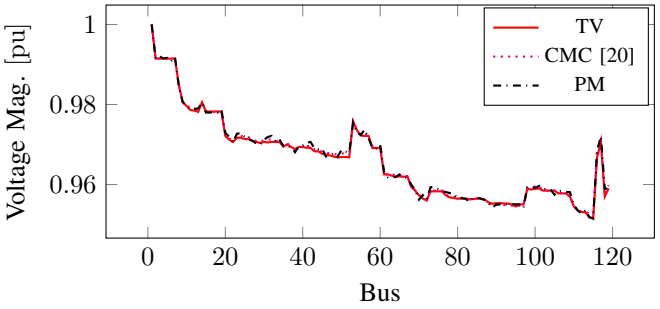


Fig. 3: Phase 'a' Voltage Magnitude for the 123-bus system with 60 available buses.

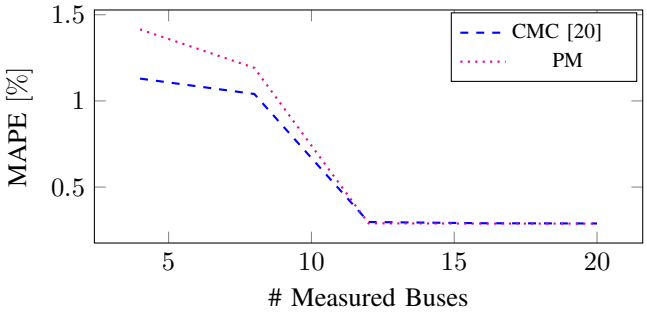


Fig. 4: Voltage magnitude estimation performance of the proposed method (PM) in 37-bus system compared to conventional matrix completion (CMC) [20].

measurements with the 37-bus and the 123-bus test systems. To detect the presence of stealthy FDI attacks, the MTD strategy given in [19] is implemented in both the test cases where D-FACTS devices are installed on 6 lines in the 37-bus system and 27 lines in the 123-bus system. In such a setup, only one-fifth of the overall number of lines have D-FACTS devices which can be achieved in a practical scenario. The amount of variation made through the D-FACTS devices from its initial set-point for detection is defined as MTD magnitude and denoted by γ. In such a setup, we first show how the computed chi-squared values using the proposed approach vary under six different conditions as tabulated in Table I.

For the 37-bus and the 123-bus systems, the chi-squared

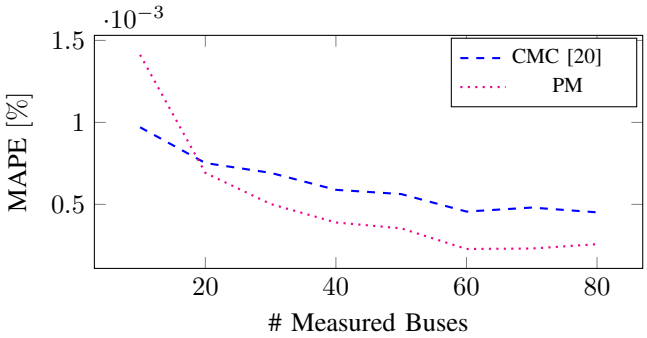


Fig. 5: Voltage magnitude estimation performance of the proposed method (PM) in 123 bus system compared to conventional matrix completion (CMC) [20].

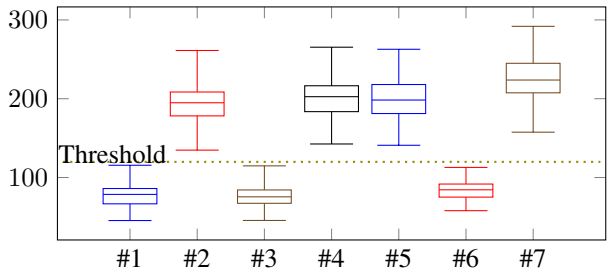


Fig. 6: Computed Residuals under multiple scenarios in 37-bus system

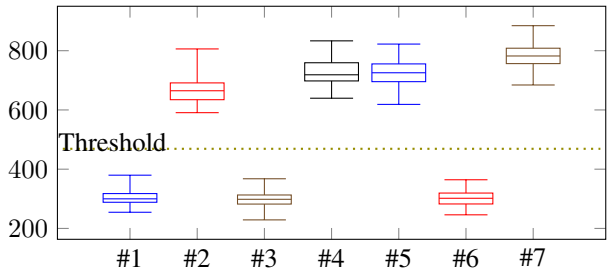
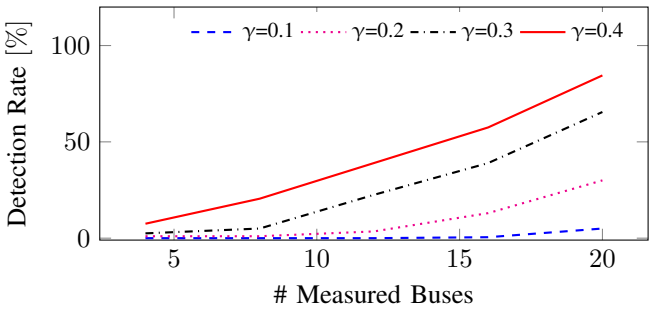


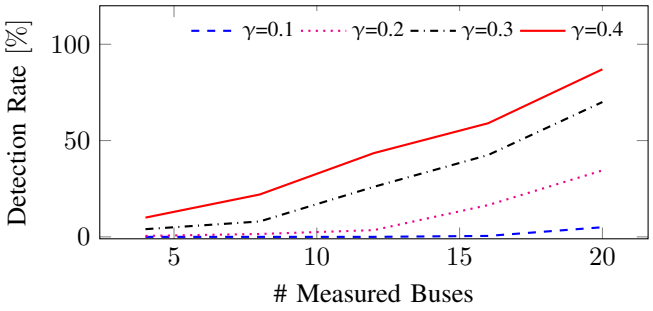
Fig. 7: Computed Residuals under multiple scenarios in 123-bus system

values are computed for each of the conditions listed in Table I and its box plots given in Fig. 6 and Fig. 7, respectively. In both test cases, the threshold is selected from the chi-squared distribution with a detection probability of 0.95. From both sets of box plots, it can be seen that chi-squared values in condition #1 stay below the threshold limit due to the absence of bad data in a no-attack condition. Data manipulations with a single bus and that do not conform to the network governing equations are carried out in condition #2 in both test systems. As seen from Fig. 6 and Fig. 7, the proposed approach can detect such non-stealthy anomalies as defined in condition #2 even when network parameter values are not perturbed with MTD strategy as γ=0. Conditions #3 and #4 consider stealthy attacks on a single target bus where the measurements in both such conditions satisfy the governing equations with the values of network parameters perceived by the attacker. If the operator does not implement any MTD strategy i.e., γ=0, as demonstrated in condition #3, then the proposed approach is not able to detect stealthy data manipulations in condition #3 as it does not generate any bad data in both 37-bus and 123-bus test systems.

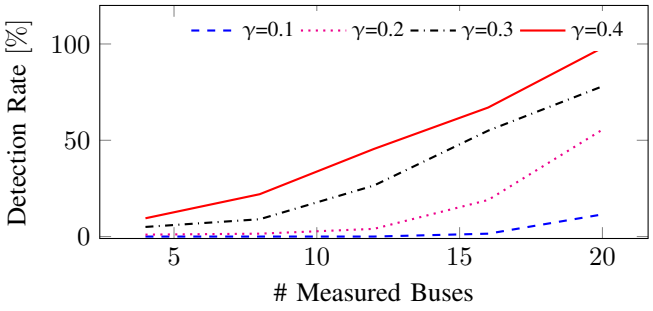
In condition #4, MTD magnitude, γ is set to 0.4 and it is considered that the attacker is unaware of the perturbations made in the network parameters by the operator to induce bad data in the measurements in the event of false data injection attacks. It is easy to see in Fig. 6 and Fig. 7 that after executing the MTD strategy as given in condition #4, the proposed approach can detect stealthy attacks as the chi-squared residuals violate the threshold limit under this condition. Condition #5 considers stealthy attacks on two target buses where both attacks are executed in an uncoordinated



(a) #D-FACTS=4

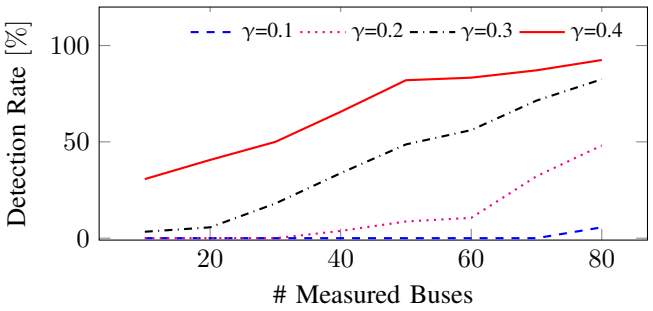


(b) #D-FACTS=6

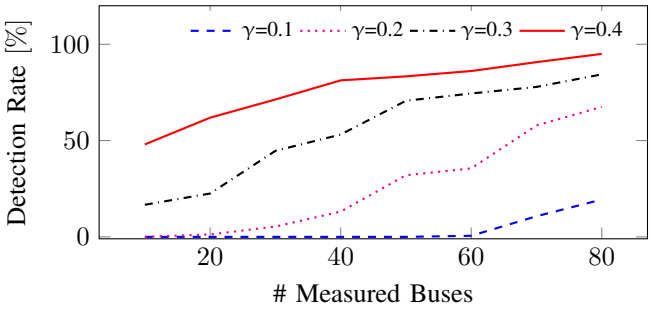


(c) #D-FACTS=9

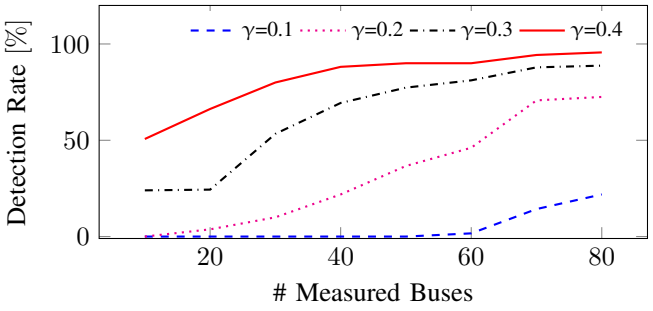
Fig. 8: Detection Rate for the 37-bus system.



(a) #D-FACTS=13



(b) #D-FACTS=27



(c) #D-FACTS=35

Fig. 9: Detection Rate for the 123-bus system.

manner. Even though both attacks individually can conform to the network governing equations but without any coordination between these attacks, the resulting manipulations will not satisfy the network model. Thus, as seen in Fig. 6 and Fig. 7 for 37-bus and 123-bus systems respectively, uncoordinated attacks naturally induce bad data and they violate the threshold with the proposed approach. In contrast to the previous test condition, conditions #6 and #7 consider a coordinated stealthy attack on two target buses. As indicated in Table I, condition #6 does not induce MTD perturbation i.e., γ=0 and therefore no bad data can be detected by the proposed approach as shown in Fig. 6 and Fig. 7 for 37-bus and 123-bus systems respectively. Since in condition #7, the system is perturbed with an MTD magnitude, γ of 0.4 for both the test systems, the proposed approach detects the coordinated stealthy attack as shown in Fig. 6 and Fig. 7.

To demonstrate the effectiveness of the proposed approach against stealthy attacks, an extensive analysis is carried out for 37-bus and 123-bus systems by varying different parameters namely the number of measured buses, number of D-FACTS devices and MTD magnitude. In each of the test systems, for each scenario, 1000 Monte-Carlo simulations are executed where measurement configurations and the possible set of target buses are randomly picked in every simulation so that a wide range of possible scenarios in a given test system are covered. In both the test systems, the detection rate of the proposed approach is analysed by varying three parameters namely the number of measured buses, the number of D-FACTS connected lines and MTD magnitude. The detection rate was computed by incrementing the number of measured buses from the least amount to a maximum value such that the system reaches closer to the observability limit. For the 37-bus system, the number of measured buses is varied from 4 to 20 with increments of 4 at every step. Similarly, in the 123-bus system, the number of measured buses is increased from 10 to 80 with increments of 10 at every step. In each of these analyses, we have considered three setups in both test systems, starting with the least number of D-FACTS connected lines to the maximum value that provides the detection rate closer to 100%. The first, second and third setups for the 37-bus system have 4, 6 and 9 D-FACTS connected lines respectively and for

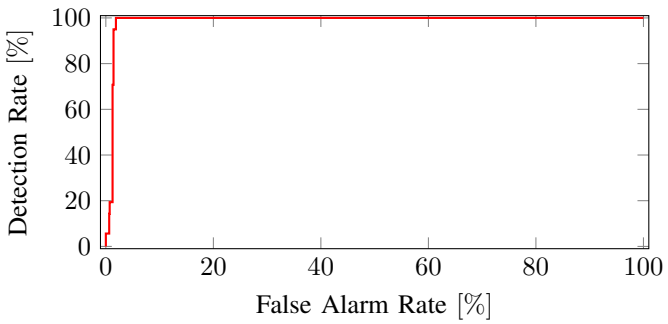


Fig. 10: ROC of the Proposed Detection Approach.

the 123-bus system have 13, 27 and 35 D-FACTS connected lines respectively. The detection rate in each of the three setups for 37-bus and 123-bus systems are plotted Fig. 8 and Fig. 9 for four levels of MTD magnitude i.e., γ=0.1, 0.2, 0.3 and 0.4.

It is intuitive to see in Fig.8 and Fig.9 that the increase in the number of measured buses eventually increases the level of visibility of the network and in turn increases the detection rate. In addition to that, an increase in the number of D-FACTS connected lines or the MTD magnitude, γ, will eventually increase the probability to expose stealthy data manipulations in the form of bad data and increase the detection rate of the proposed approach. Thus, in a distribution system with a modest number of measurements, the detection rate of the proposed approach will be high as long as either a considerable number of D-FACTS devices are used or MTD magnitude is high enough for the detection process. From this analysis, it is surmised that in 37-bus and 123-bus systems, stealthy data manipulations at any of its buses can be detected with a probability of 95% as long as the number of measured buses is at least half the amount of the total number of buses in the system and D-FACTS devices are installed in at least 20% of the lines in the network. Even though in this demonstration, MTD has been implemented through D-FACTS, it can be noted that the proposed approach is also compatible with other MTD schemes like perturbing the set points of distributed energy resources [8]. It is because the proposed approach uses the residuals and hence it can work with any MTD scheme that translates false data into bad data. The receiver operating characteristics (ROC) curve of the proposed approach is plotted in Fig. 10, which indicates the relationship between the detection rate and the false alarm rate. This ROC curve shows that it has characteristics closer to an ideal detector since the false alarm rate is almost negligible in the proposed approach.

VI. Conclusion

In this paper, a technique that can detect FDI attacks in an unobservable distribution network is presented. A modified matrix completion approach is developed such that it can provide the estimates of the system state and its error variances at the unobserved locations in the network. Using these estimates, the chi-squared test is employed to detect the presence of bad data. For detecting stealthy FDI attacks, D-FACTS are deployed to execute a moving target defence strategy and through that stealthy false data injection attacks can be revealed in the form of bad data which can be detected with the proposed approach. This methodology is tested on 37-bus and 123-bus distribution systems. The results show that the proposed approach provides around 90% detection rate for FDI attacks for distribution systems even if only 50% of its buses have measurements.

References

[1] 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, Jun. 2011.

[2] Z. Li, M. Shahidehpour, and F. Aminifar, "Cybersecurity in distributed power systems," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1367–1388, 2017.

[3] Q. Dai, L. Shi, and Y. Ni, "Risk assessment for cyberattack in active distribution systems considering the role of feeder automation," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3230–3240, 2019.

[4] R. Deng, P. Zhuang, and H. Liang, "False data injection attacks against state estimation in power distribution systems," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2871–2881, 2019.

[5] P. Zhuang, R. Deng, and H. Liang, "False data injection attacks against state estimation in multiphase and unbalanced smart distribution systems," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6000–6013, 2019.

[6] J. R. R. Kumar, B. Sikdar, and D. Kundur, "Electromagnetic transients based detection of data manipulation attacks in three phase radial distribution networks," *IEEE Trans. Ind. Appl.*, pp. 1–1, 2021.

[7] H. Zhang, B. Liu, and H. Wu, "Smart grid cyber-physical attack and defense: A review," *IEEE Access*, vol. 9, pp. 29 641–29 659, 2021.

[8] K. Jhala, P. Pradhan, and B. Natarajan, "Perturbation-based diagnosis of false data injection attack using distributed energy resources," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1589–1601, 2021.

[9] D. Divan and H. Johal, "Distributed facts—a new concept for realizing grid power flow control," *IEEE Transactions on Power Electronics*, vol. 22, no. 6, pp. 2253–2260, 2007.

[10] K. M. Rogers and T. J. Overbye, "Some applications of distributed flexible ac transmission system (d-facts) devices in power systems," in *2008 40th North American Power Symposium*, 2008, pp. 1–8.

[11] K. L. Morrow, E. Heine, K. M. Rogers, R. B. Bobba, and T. J. Overbye, "Topology perturbation for detecting malicious data injection," in *2012 45th Hawaii International Conference on System Sciences*, 2012, pp. 2104–2113.

[12] K. R. Davis, K. L. Morrow, R. Bobba, and E. Heine, "Power flow cyber attacks and perturbation-based defense," in *International Conference on Smart Grid Communications*, 2012, pp. 342–347.

[13] B. Li, G. Xiao, R. Lu, R. Deng, and H. Bao, "On feasibility and limitations of detecting false data injection attacks on power grid state estimation using d-facts devices," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 854–864, 2020.

[14] C. Liu, J. Wu, C. Long, and D. Kundur, "Reactance perturbation for detecting and identifying fdi attacks in power system state estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 763–776, 2018.

[15] J. Tian, R. Tan, X. Guan, and T. Liu, "Enhanced hidden moving target defense in smart grids," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2208–2223, 2019.

[16] R. Deng and H. Liang, "False data injection attacks with limited susceptance information and new countermeasures in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1619–1628, 2019.

[17] B. Liu and H. Wu, "Optimal planning and operation of hidden moving target defense for maximal detection effectiveness," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4447–4459, 2021.

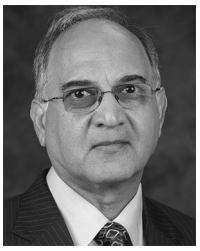
[18] A. R. Gupta and A. Kumar, "Deployment of distributed generation with d-facts in distribution system: A comprehensive analytical review," *IETE Journal of Research*, vol. 68, no. 2, pp. 1195–1212, 2022.

[19] B. Liu, H. Wu, A. Pahwa, F. Ding, E. Ibrahim, and T. Liu, "Hidden moving target defense against false data injection in distribution network reconfiguration," in *2018 IEEE Power Energy Society General Meeting (PESGM)*, 2018, pp. 1–5.

[20] P. L. Donti, Y. Liu, A. J. Schmitt, A. Bernstein, R. Yang, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2520–2530, 2020.
[21] R. Singh, B. C. Pal, and R. B. Vinter, "Measurement placement in distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 668–675, 2009.
[22] M. C. de Almeida and L. F. Ochoa, "An improved three-phase amb distribution system state estimator," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1463–1473, 2017.
[23] K. A. Clements, "The impact of pseudo-measurements on state estimator accuracy," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–4.
[24] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4910–4920, 2019.
[25] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1006–1013, 2016.
[26] S. Dahale, H. S. Karimi, K. Lai, and B. Natarajan, "Sparsity based approaches for distribution grid state estimation - a comparative study," *IEEE Access*, vol. 8, pp. 198 317–198 327, 2020.
[27] J. R. K. Rajasekaran, B. Natarajan, A. Pahwa, and H. Wu, "Detection of stealthy false data injection attacks in unobservable distribution networks," in *Innovative Smart Grid Technologies Conference*, 2022.
[28] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3875–3883, 2017.
[29] K. Nagappan and K. Mohanram, "Building algorithms for the formation of the bus hybrid matrix," *Electric Power Systems Research*, vol. 5, no. 3, pp. 219–229, 1982.
[30] N. Higham, "Matrix nearness problems and applications," *Applications of Matrix Theory*, vol. 22, 1989.

James Ranjith Kumar Rajasekaran James Ranjith Kumar Rajasekaran (Member, IEEE) received the B.E. degree (Dist.) in Electrical and Electronics engineering from Anna University, Chennai, in 2009, the M.E. degree in Electrical Engineering from the Indian Institute of Science, Bengaluru, in 2014, and the Ph.D. degree in Electrical and Computer Engineering from National University of Singapore, Singapore, in 2021. His areas of interest include power system analysis, transient analysis, and smart grid cyber-security.

Balasubramaniam Natarajan Balasubramaniam Natarajan (Senior Member, IEEE) received the B.E. degree (Hons.) in electrical and electronics engineering from the Birla Institute of Technology and Science, Pilani, India, in 1997, the Ph.D. degree in electrical engineering from Colorado State University, Fort Collins, CO, USA, in 2002, and the Ph.D. degree in statistics from Kansas State University, Manhattan, KS, USA, in 2018. He is currently a Clair N. Palmer and Sara M. Palmer Endowed Professor and the Director of the Cyber-Physical Systems and Wireless Innovations Research Group, Kansas State University. His research interests include statistical signal processing, stochastic modeling, optimization, and control theories. He has worked and published extensively on modeling, analysis and networked estimation, and control of smart distribution grids and cyber physical systems in general. He has published more than 200 refereed journals and conference papers. He has served on the editorial board for multiple IEEE journals, including IEEE Transactions on Wireless Communications.



Anil Pahwa Anil Pahwa (Fellow, IEEE) received the B.E. degree (Hons.) in electrical engineering from the Birla Institute of Technology and Science, Pilani, India, in 1975, the M.S. degree in electrical engineering from the University of Maine, Orono, ME, USA, in 1979, and the Ph.D. degree in electrical engineering from the Texas A&M University, College Station, TX, USA, in 1983. Since 1983, he has been with Kansas State University, Manhattan, KS, USA, where he is currently a University Distinguished Professor and holds the Logan Fetterhoof Chair with the Electrical and Computer Engineering Department. The National Academy selected him as a Jefferson Science Fellow in 2014 to serve as a Senior Science Advisor in the U.S. State Department for one year. He was with the East Asian and Pacific Affairs Bureau on international policies to facilitate higher deployment of renewable energy. His research interests include distribution automation, distribution planning, renewable energy integration into power systems, and intelligent computational methods for distribution system applications.