Enhanced Robustness of State Estimator to Bad Data Processing Through Multi-Innovation Analysis

Junbo Zhao, *Student Member, IEEE*, Gexiang Zhang, *Member, IEEE*, Massimo La Scala, *Fellow, IEEE*, Zhaoyu Wang, *Member, IEEE*

***Abstract*—To enhance the robustness of a power system state estimator to topology errors, bad critical measurements, multiple non-interacting or interacting bad data (BD), this paper presents a new robust detection method by exploiting the temporal correlation and the statistical consistency of measurements. Particularly, we propose three innovation matrices to capture the measurement correlation and statistical consistency by processing the forecasted states/measurements and the interpolated reliable information from phasor measurement units. The latter is achieved by using a robust generalized maximum-likelihood estimator. We then propose to apply the projection statistics (PS) to the proposed innovation matrices for BD detection. Extensive Monte Carlo simulations and QQ-plots are carried out to obtain an analytical threshold of the statistical test of the PS. Because of the robustness of PS and the enhanced measurement redundancy by the innovations, the proposed method is able to handle various types of BD in both PMU observable and PMU partially observable power systems. Moreover, the proposed method is suitable for parallel implementation, and can be integrated with online applications. Comparison results with existing methods under different BD conditions on IEEE 14-bus, 118-bus and Polish 2383-bus test systems demonstrate the effectiveness and robustness of the proposed method.**

***Index Terms*—Robust estimation, state estimation, bad data detection, phasor measurement unit, state forecasting, innovation vectors, correlation, statistical consistency.**

I. INTRODUCTION

POWER system state estimation (SE) is an important function in modern energy management systems (EMS). It computes the most likely states of the network based on redundant measurements provided by supervisory control and data acquisition (SCADA) system or phasor measurement units (PMUs). With the estimated states, various EMS functions can be performed, such as optimal power flow, contingency analysis and bad data (BD) detection/analysis, etc. [1]–[3].

In the field of power system SE, BD detection and processing is essential. To date, various approaches have been proposed. They include largest normalized measurement residual (LNR) test-based approach, the hypothesis test [4]-based approach, and the robust estimator-based approaches, such as least absolute value estimator and least median of squares estimator, [5] etc. Although these methods have relatively satisfactory performance with single and multiple BD, they may not be effective in the presence of bad critical measurement, smearing effect and the conforming and interacting BD. Thanks to the wide-area deployment of PMUs in recent years, enhanced performance of BD detection is achieved by exploring the redundant information from PMUs [6]–[11]. In [6], a PMU placement algorithm to transform the existing critical measurements into redundant ones was firstly performed, then an enhanced BD detection method by using both SCADA and PMU measurements was proposed. In [7], bad critical or pairs were addressed and subsequently converted to non-critical bad measurements by placing additional PMUs. Authors in [8] proposed a two-stage method with enhanced measurement redundancy by PMUs to detect bad critical measurements. In [9], a conventional normalized residual-based method was proposed to improve BD detection performance by processing PMU and SCADA measurements simultaneously. In [10], SCADA and PMU measurements were processed separately, yielding two independent estimators. Then the phase-aided normalized residual-based test was integrated with the measurement residual-based test to detect BD. In [11], a multistage phasor-aided method was proposed, where the difference between interpolated SCADA measurements by PMUs and raw SCADA measurements was defined as an innovation vector. Then statistical test was adopted to handle various types of BD.

Although the performance of normalized measurement residual-based test has been improved thanks to the enhanced measurement redundancy by PMUs, it is still prone to bad leverage points, the smearing effect, and the conforming and interacting BD [4], [5]. Note that the improved performance can be achieved only when BD occurs in PMU-observable areas and there is no simultaneous occurrence of BD in PMU and SCADA measurements. The latter may induce smearing effect as well as the conforming and interacting BD. To address these issues, the normalized innovation test is proposed by processing together the forecasted state/measurement information and

This work was supported by the National Natural Science Foundation of China (61672437, 61373047) and the Research Project of Key Laboratory of Fluid and Power Machinery (Xihua University), Ministry of Education, P.R. China (JYBFX-YQ-1). *Corresponding author*: Gexiang Zhang.

Junbo Zhao is with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu, 610031 China and the Bradley Department of Electrical Computer Engineering, Virginia Polytechnic Institute and State University, Northern Virginia Center, Falls Church, VA 22043, USA (e-mail: zjunbo@vt.edu).

Gexiang Zhang is with Robotics Research Center, Xihua University, Chengdu, 610039, China, and Key Laboratory of Fluid and Power Machinery (Xihua University), Ministry of Education, Chengdu, 610039, China, and School of Electrical Engineering, Southwest Jiaotong University, Chengdu, 610031 China (zhgxdylan@126.com).

Massimo La Scala is with the Department of Electrical Engineering and Computer Science, Politecnico di Bari, 70125 Bari, Italy (e-mail:massimo.lascala@poliba.it).

Zhaoyu Wang is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, USA (email: wzy@iastate.edu).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2016.2626782, IEEE Transactions on Industrial Informatics

received measurements. However, the detection threshold of the statistical test is difficult to tune in practice since it is system dependent [22], and sensitive to the measurement and process noise. Furthermore, this method is vulnerable to the clustered BD [12], [13].

In this paper, we relax the PMU observability assumption for BD detection enhancement, and propose an alternative robust detection method by exploiting the temporal correlation and the statistical consistency of measurements. Specifically, we propose three new innovation matrices to capture the correlation and measurement statistical consistency, where the innovation matrices are developed by processing the forecasted measurements and the interpolated information from cleaned PMU measurements. We then propose to apply the projection statistics (PS) to these matrices for BD detection with an analytical detection threshold. The latter is determined by extensive Monte Carlo simulations and QQ-plots. Thanks to the robustness of PS, we are able to handle clustered BD, multiple non-interacting or interacting BD as well as the smearing effect. In summary, the main contributions of this paper are:

1) We relax the PMU observability assumption for detecting bad critical measurements, topology errors and suppressing smearing/masking effect;
2) It presents a framework to effectively combine limited number of PMU measurements with forecasted states and measurements for BD detection;
3) Unlike the innovation-based statistical test using system dependent detection threshold, the proposed method has an analytical threshold;
4) Because of the good robustness of PS and the enhanced measurement redundancy by the innovations, the proposed method is able to handle various types of BD for both PMU observable and PMU partially observable power systems; they include topology errors, bad critical measurements and multiple non-interacting or interacting BD;
5) The proposed method is suitable for parallel implementation, and can be integrated with online applications.

The remainder of this paper is organized as follows: Section II presents the definitions of proposed innovation vectors. The robust BD detection framework is presented in Section III. Section IV analyzes the simulation results, and finally Section V concludes the paper.

II. Definitions of Proposed Innovation Vectors

The combination of SE with forecasted states and measurements, called the forecasting-aided state estimator (FASE), is considered as a promising way to enhance BD processing [14]–[16]. Using the forecasted information, the main difficulties in conventional BD processing can be overcome since the innovation-based analysis is an effective tool to detect multiple spurious measurements without being affected by the smearing effect. Also, the detected BD can be easily replaced by the forecasted measurements without causing any system observability problem. However, this innovation-based test has several disadvantages, such as the difficulties in determining an analytical threshold of the statistical test and its vulnerability to clustered BD (masking effect) [12], [13]. On the other hand, to our best knowledge, there are few papers investigating the benefits of combining the FASE with both PMU and SCADA measurements to enhance the BD detection. In this paper, we propose a framework to effectively combine limited number of PMU measurements with forecasted states and measurements for BD robust detection.

A. Short-term State Forecasting

Studies have shown that the temporal correlation exists among the loads in the same geographic area [17], [18]. Therefore, the states driven by loads should show temporal characteristics. To capture the temporal correlation between the states in different time instants, the stochastic vector auto-regressive time series model of a discrete time-variant power system is used

$$\boldsymbol{x}_{k+1} = \boldsymbol{\Phi}_k \boldsymbol{x}_k + \boldsymbol{w}_k, \tag{1}$$

where $\boldsymbol{x}_k$ is the state vector including the voltage magnitude and angle at every bus; k is the time sample; $\boldsymbol{\Phi}_k$ represents the state transition matrix and is updated recursively using the method in [19]; $\boldsymbol{w}_k$ represents the modeling uncertainties, which is usually assumed to follow a Gaussian distribution with zero mean and covariance matrix $\boldsymbol{Q}_k$, i.e., $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)$. When the estimated state transition matrix $\widehat{\boldsymbol{\Phi}}_k$ is available, the forecasted state vector $\widetilde{\boldsymbol{x}}_{k+1}$ with its covariance matrix $\boldsymbol{S}_{k+1}$ can be obtained through

$$\widetilde{\boldsymbol{x}}_{k+1} = \widehat{\boldsymbol{\Phi}}_k \widehat{\boldsymbol{x}}_k, \tag{2}$$

$$\boldsymbol{S}_{k+1} = \widehat{\boldsymbol{\Phi}}_k \boldsymbol{\Sigma}_k \widehat{\boldsymbol{\Phi}}_k^T + \boldsymbol{Q}_k, \tag{3}$$

where $\widehat{\boldsymbol{x}}_k$ and $\boldsymbol{\Sigma}_k$ are the estimated state vector and its corresponding covariance matrix at time sample k, respectively.

B. Proposed Innovation Vectors

To date, a given power grid is typically observable by SCADA measurements while partially observable through PMUs. In this paper, a hybrid topological/numerical method [20], [21] for power system observability analysis is used to decompose a given system into PMU observable areas and PMU unobservable but SCADA observable areas.

1) ***Innovation Vector for PMU Observable Area****:* When the measurement vector for a PMU observable area is available at time sample $k+1$, one innovation vector is defined as the difference between the received PMU measurement vector $\boldsymbol{z}_{k+1}^p \in \mathbb{R}^m$ and the forecasted measurement vector $\widetilde{\boldsymbol{z}}_{k+1}^p \in \mathbb{R}^m$, where $\boldsymbol{z}_{k+1}^p$ includes the bus voltage phasors and line current phasors that are adjacent to the PMU buses. In this paper, the rectangular coordinates expression is used for PMU measurements to obtain a linear measurement model. We may also adopt the polar expression suggested in [9]. Formally, we have

Definition 1. The difference between $\boldsymbol{z}_{k+1}^p(i)$ and $\widetilde{\boldsymbol{z}}_{k+1}^p(i)$ ($\widetilde{\boldsymbol{z}}_{k+1}^p = \boldsymbol{H}_{k+1}^p \widetilde{\boldsymbol{x}}_{k+1}^p$) is defined as the innovation for the ith PMU measurement, i.e.,

$$\boldsymbol{\nu}_{k+1}(i) = \boldsymbol{z}_{k+1}^p(i) - \widetilde{\boldsymbol{z}}_{k+1}^p(i), \tag{4}$$

where $\boldsymbol{\nu}_{k+1}(i)$ is the ith component of the innovation vector; $\boldsymbol{H}_{k+1}^p$ is the constant Jacobian matrix that relates linear relationship between the states and the PMU measurements; $\widetilde{\boldsymbol{x}}_{k+1}^{p}$ is the forecasted state vector of the PMU observable area.

Remark 1: It should be noted that when a sudden change takes place, which may be caused by loss of loads, unscheduled outages of generating units and/or the change of network configuration, the forecasted state/measurements are no longer valid. As a result, high innovations appear and consequently yield difficulties to the correct identification of which anomaly has occurred. To address this problem, that is, to distinguish the occurrence of sudden change, gross measurement error and network topology error, the skewness test and active line flow-based test have been proposed [22]–[24]. However, they did not step further to investigate the exact locations of BD and topology errors. In this paper, we assume the three anomalies have been distinguished by the approaches in [22]–[24], and our main scopes are to identify which measurement is bad and which topology is wrong.

Using the robust detection method proposed in the next section, the BD in PMU measurements can be detected and downweighted. After that, their influences are suppressed by the robust generalized maximum-likelihood (GM)-estimator, which solves the following objective function (the time instant is dropped for simplicity)

$$J(\boldsymbol{x}) = \sum_{i=1}^{m} \varpi_i^2 \rho\left(r_{S_i}\right), \tag{5}$$

where ϖ_i is calculated by (16) presented in the next section; $r_{S_i} = r_i/s\varpi_i$ is the standardized residual; $r_i = z_i^p - \boldsymbol{a}_i^T \hat{\boldsymbol{x}}^p$ and $\boldsymbol{a}_i^T$ is the ith column of the Jacobian matrix $\boldsymbol{H}^p$; $s = 1.4826 \cdot b_m \cdot \text{median}_i\, |r_i|$ is the robust scale estimate; b_m is a correction factor for unbiasedness at the Gaussian distribution; $\rho(\cdot)$ is the Huber cost function given by

$$\rho\left(r_{S_i}\right) = \begin{cases} \frac{1}{2} r_{S_i}^2, & \text{for } |r_{S_i}| < c \\ c\,|r_{S_i}| - c^2/2, & elsewhere \end{cases}, \tag{6}$$

where the parameter c is the breakpoint to balance the robustness and statistical efficiency at the desired distributions. Huber [25] has shown that if c is chosen between 1.5 and 3, the algorithm achieves at least 95% asymptotic efficiency under normal assumptions and performs well in most heavy tailed situations. In this paper, c is set to be 1.5, which is a commonly used value in literature.

To minimize (5), we take its partial derivative and set it to zero, yielding

$$\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}} = \sum_{i=1}^{m} -\frac{\varpi_i \boldsymbol{a}_i^T}{s} \psi\left(r_{S_i}\right) = \boldsymbol{0}, \tag{7}$$

where $\psi\left(r_{S_i}\right) = \partial\rho\left(r_{S_i}\right)/\partial r_{S_i}$. Then, we first divide and multiply the standardized residual r_{S_i} to both sides of (7) and rewrite it in a matrix form, yielding

$$(\boldsymbol{H}^p)^T \boldsymbol{Q}(\boldsymbol{z}^p - \boldsymbol{H}^p \hat{\boldsymbol{x}}^p) = \boldsymbol{0}, \tag{8}$$

where $\boldsymbol{Q}$ =diag$(q\left(r_{S_i}\right))$ and $q\left(r_{S_i}\right) = \psi\left(r_{S_i}\right)/r_{S_i}$.

By using the iterative reweighted least square algorithm, the state can be estimated and represented as

$$\widehat{\boldsymbol{x}}_{k+1}^p = \left(\boldsymbol{H}_{k+1}^{p}{}^T \boldsymbol{Q} \boldsymbol{H}_{k+1}^p\right)^{-1} \boldsymbol{H}_{k+1}^{p}{}^T \boldsymbol{Q} \boldsymbol{z}_{k+1}^p. \tag{9}$$

Therefore, the SCADA measurements can be interpolated/estimated by $\widehat{\boldsymbol{z}}_{k+1}^s = \boldsymbol{h}\left(\widehat{\boldsymbol{x}}_{k+1}^p\right)$, where $\boldsymbol{h}$ is the vector-valued nonlinear measurement function of the SCADA measurements. Finally, we have the innovation vector for SCADA measurements defined as

Definition 2. The difference between received raw SCADA measurement $\boldsymbol{z}_{k+1}^s(i)$ and its interpolated/estimated measurement $\widehat{\boldsymbol{z}}_{k+1}^s(i)$ is defined as the innovation of the ith SCADA measurement, i.e.,

$$\boldsymbol{\xi}_{k+1}(i) = \boldsymbol{z}_{k+1}^s(i) - \widehat{\boldsymbol{z}}_{k+1}^s(i). \tag{10}$$

where $\boldsymbol{\xi}_{k+1}(i)$ is the ith component of the innovation vector.

Remark 2: The idea of proposing this innovation vector is that two independent state estimations using PMU measurements and SCADA measurements, respectively, should produce consistent state estimates if no BD occurs. When the PMU measurement set is cleaned by the robust estimator, its estimation results can be used to estimate/interpolate the SCADA measurements. If there is no BD in SCADA measurements, the estimated SCADA measurements should be consistent with the received raw SCADA measurements, otherwise, bad measurements exist. Therefore, by checking the consistency of these two measurement sets, the BD in SCADA measurements can be detected. On the other hand, the state estimation model itself is an approximate model with parameter uncertainty, measurement bias, topology uncertainty, etc. Thus, a high measurement redundancy is needed to estimate the true system states. This motivates us to clean the SCADA measurements in PMU observable areas.

2) ***Innovation Vector for PMU Unobservable but SCADA Observable Area***: Similar to the definition of the innovation vector for PMU measurements, we define the innovation vector for PMU unobservable but SCADA observable area as

Definition 3. The difference between received raw measurement $\boldsymbol{z}_{k+1}^r(i)$ and its forecasted measurement $\widetilde{\boldsymbol{z}}_{k+1}^r(i)$ ($\widetilde{\boldsymbol{z}}_{k+1}^r = \boldsymbol{h}(\widetilde{\boldsymbol{x}}_{k+1}^r)$) is defined as the innovation for the ith SCADA measurement in PMU unobservable but SCADA observable area, i.e.,

$$\boldsymbol{\lambda}_{k+1}(i) = \boldsymbol{z}_{k+1}^r(i) - \widetilde{\boldsymbol{z}}_{k+1}^r(i), \tag{11}$$

where $\boldsymbol{\lambda}_{k+1}(i)$ is the ith component of the innovation vector.

Remark 3: The idea of proposing innovations 1 and 3 is that even when the system changes at each time instant, the loads and generators change slowly and show temporal correlations, which results in temporal correlations of the system states. Therefore, sampled measurements taken from the system should show temporal correlations with the calculated measurements through forecasted states. However, the correlations will be broken up by the BD. Thus, by checking the statistical correlations of the forecasted and received measurements, we can effectively detect the BD.

III. Proposed Robust Bad Data Detection and Processing Method

A. Motivation of Robust Detection

It has been shown in [15], [22] that the statistical test on normalized innovation of (11) can be used to detect BD

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2016.2626782, IEEE Transactions on Industrial Informatics

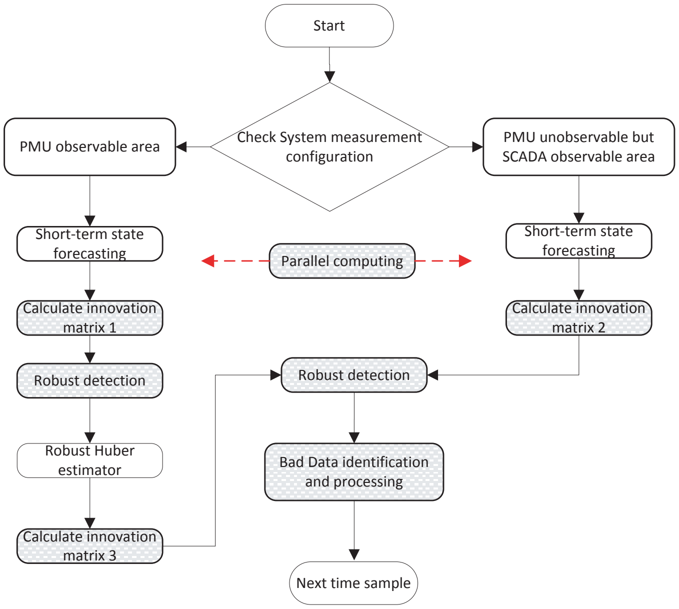


Fig. 1: The flowchart of the proposed method

assuming the detection threshold is chosen appropriately and no bad leverage point or clustered BD occurs. However, there are several disadvantages of this approach, such as the difficulties in determining an appropriate detection threshold (the detection threshold is difficult to tune in practice since it is system dependent [22] and sensitive to the measurement and process noise). In addition, the normalized innovations are in fact the well-known Mahalanobis distances that represent the surface of a multidimensional ellipsoid centered at the sample mean [12], [13]. With the assumption that the innovation vector follows a Gaussian distribution, the square of the Mahalanobis distances follow a χ^2 distribution. Then, the normalized innovation statistical test is applied to detect BD. However, this method is vulnerable to clustered BD because of masking effect [12], [13].

To handle these difficulties, this paper proposes an alternative robust BD detection method by exploiting the statistical properties of the developed innovation vectors. The detection threshold is system independent and can be determined in an analytical way. Because of the robustness of PS with asymptotic hight breakdown point, the masking effect is effectively suppressed. The flowchart of the proposed method is shown in Fig. 1. It can be observed that the proposed method has different strategies for PMU observable areas and PMU unobservable but SCADA observable areas. The left column of Fig. 1 is to detect and suppress BD in PMU observable areas while the right column is to detect suspicious measurements in PMU unobservable but SCADA observable areas. The central robust detection step is to perform the consistency check to identify whether suspicious measurements in PMU unobservable but SCADA observable areas are BD or not. In addition, it can be seen from the flowchart that the detection and suppression of BD in PMU observable areas as well as PMU unobservable but SCADA observable areas are suitable for parallel implementation, which is attractive for large-scale systems. Please note that the central robust detection

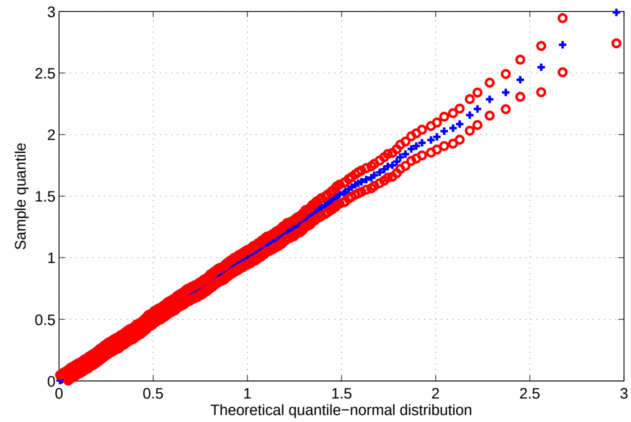


Fig. 2: QQ plots of the sample quantiles of the innovation vs. the corresponding quantiles of the normal distribution

is performed only for suspicious measurements and thus can be carried out very fast.

B. Robust Detection

The consistency of the interpolated measurements and received measurements is checked to detect the BD using the innovation vector in definition 2. To detect the BD using innovation vectors in Definitions 1 and 3, it is necessary to check the temporal correlations of the forecasted and received measurements. In this paper, we propose a modified version of PS [28] to perform the robust BD detection. PS uses the sample median and the median-absolute-deviation of the data point $\boldsymbol{\ell}$ in the direction of all possible vectors $\boldsymbol{l}$ for the robust location and scale estimation, i.e.,

$$PS_i = \max_{\|\boldsymbol{l}\|=1} \frac{\left|\boldsymbol{\ell}_i^T \boldsymbol{l} - med_j\left(\boldsymbol{\ell}_j^T \boldsymbol{l}\right)\right|}{1.4826\, med_k \left|\boldsymbol{\ell}_k^T \boldsymbol{l} - med_j\left(\boldsymbol{\ell}_j^T \boldsymbol{l}\right)\right|}. \tag{12}$$

The first proposal of PS for power engineering application is to develop a sparse version of the original PS [27] to identify the leverage points [28]. The latter is achieved by applying PS to the sparse Jacobian matrix. However, in the developed innovation vectors, no sparse Jacobian matrix is involved. On the other hand, the original PS by Gasko and Donoho [27] is actually a type of robust Mahalanobis distance that is robust to outliers no matter they are bad leverage points or vertical outliers. Therefore, the key point is how to design a matrix that captures the statistical characteristics of outliers. To do this, the following three innovation matrices are proposed.

- Matrix for the innovation vector 1:

$$\boldsymbol{Z}_1 = [\boldsymbol{\nu}_k \;\; \boldsymbol{\nu}_{k+1}] = [\boldsymbol{z}_k^p - \widetilde{\boldsymbol{z}}_k^p \;\;\; \boldsymbol{z}_{k+1}^p - \widetilde{\boldsymbol{z}}_{k+1}^p] \tag{13}$$

- Matrix for the innovation vector 2:

$$\boldsymbol{Z}_2 = [\boldsymbol{\xi}_k \; \boldsymbol{\xi}_{k+1}] = [\boldsymbol{z}_k^s - \widehat{\boldsymbol{z}}_k^s \;\; \boldsymbol{z}_{k+1}^s - \widehat{\boldsymbol{z}}_{k+1}^s] \tag{14}$$

- Matrix for the innovation vector 3:

$$\boldsymbol{Z}_3 = [\boldsymbol{\lambda}_k \;\; \boldsymbol{\lambda}_{k+1}] = [\boldsymbol{z}_k^r - \widetilde{\boldsymbol{z}}_k^r \;\;\; \boldsymbol{z}_{k+1}^r - \widetilde{\boldsymbol{z}}_{k+1}^r] \tag{15}$$

where the subscript k indicates the previous time instant.

Remark 4: The idea of developing the matrices for innovation vectors 1 and 3 is similar. Due to the temporal correlation

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2016.2626782, IEEE Transactions on Industrial Informatics

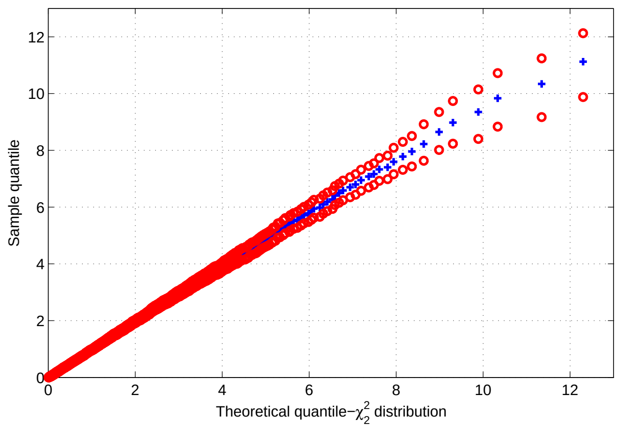


Fig. 3: QQ plots of the sample quantiles of the PS vs. the corresponding quantiles of the χ_2^2 distribution

between the received and the forecasted measurements, the innovation vectors at time instants k and $k+1$ should also show the correlation. Once BD occurs, the temporal correlation is broken up. As a consequence, some points are far away from the majority of the point cloud, yielding very large PS values. While the reason for constructing matrix 2 is as follows: if there is no BD, the received measurement vector is consistent with the interpolated measurement vector, then all of the measurement points would show the same statistical characteristics and fall into the same confidence ellipse; otherwise, the BD should be out of the confidence ellipse and detected by PS.

After computing the PS values for the elements of each matrix, they are compared to a threshold to identify outliers. The threshold is determined from the probability distribution of PS under the assumption that the majority of the data points in the innovation matrices follow a bivariate Gaussian probability distribution. To validate the Gaussian assumption for the proposed innovation vectors, extensive Monte Carlo simulations are conducted on both IEEE-14 and 118 bus systems. The test results of IEEE 118-bus system are used for illustration. A random fluctuation with a linear trend (1%-3%) is added to the load curve over a period of time, 100 time samples, then, at each time sample, the normalized innovation vectors are calculated and the procedure is repeated 100 times. The sample medians and the interquantile ranges of the empirical innovation quantiles are finally plotted versus the corresponding quantiles of the normal distribution. The QQ plots displayed in Fig. 2 provide evidence that the innovation vectors follow approximately the Gaussian distribution. To determine the probability distribution of the PS when the innovation vectors follow Gaussian distributions, we consider two random variables $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ that are independent and identically distributed according to $\mathcal{N}(\mathbf{0},\boldsymbol{I})$. Then we generate 500 realizations of these two random variables, apply the PS to $[\boldsymbol{v}_1\ \boldsymbol{v}_2]$ and repeat the procedure 100 times. The sample medians and the interquantile ranges of the empirical PS quantiles are plotted versus the corresponding quantiles of the chi-square distribution with 2 degrees of freedom. The QQ plots shown in Fig. 3 validate that PS follows chi-square distribution.

In this paper, we set the threshold of the statistical test to $\chi^2_{2,0.975}$ at a significance level of 97.5%. Thus, the measurements, whose associated PS values satisfy $PS_i > \chi^2_{2,0.975}$, are marked as BD. These BD are downweighted by the following weight function

$$\omega_i = \min\left(1, d^2/PS_i^2\right) \tag{16}$$

where d is the BD detection threshold and $d = \chi^2_{2,0.975}$.

Remark 5: In this paper, we propose to apply the PS to the three innovation matrices that contain multiple rows but only 2 columns. The calculation of PS for these matrices is very fast. For example, the calculation of PS for a matrix with 1000 rows and 2 columns (for simulating large-scale power systems) takes only around 1 second for a 2.50 GHz, 8GB of RAM, Intel Core i5 computer. If a more powerful computer is used, this computation time can be further reduced. In addition, as indicated by [14], [15] the computational effort related to the state forecasting step is negligible compared to the state filtering. In the proposed method, we are interested in preprocessing the SCADA and PMU measurements. Thus, only the state forecasting step is involved, which offers users the flexibility in choosing efficient state filtering methods, such as hierarchical or distributed estimation [29], [30], etc. Furthermore, FASE has been implemented and applied to the EMS of LIGHT Services of Electricity-the company responsible for Rio de Janeiro city energy supply [14], which demonstrates the applicability of the FASE for power system real-time monitoring. Last but not least, the proposed method is suitable for parallel implementation, which further enhances its suitability for online applications.

C. Robust Bad Data Identification and Processing

The PS is firstly applied to matrices $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_3$ simultaneously. The related BD are detected and downweighted. When the PMU measurements are cleaned by the robust GM-estimator, the matrix $\boldsymbol{Z}_2$ is constructed and PS is applied for detecting the BD in SCADA measurements in PMU observable areas.

1) BD in SCADA Measurements: Threshold violations in the tests of matrices $\boldsymbol{Z}_2$ and $\boldsymbol{Z}_3$ with negative results for matrix $\boldsymbol{Z}_1$ indicate the presence of BD in SCADA measurements, where $\boldsymbol{Z}_2$ and $\boldsymbol{Z}_3$ are associated with PMU observable, PMU unobservable but SCADA observable areas, respectively. Measurements larger than the detection threshold will be identified as BD and downweighted by (16).

2) BD in PMU Measurements: if the test results of matrix $\boldsymbol{Z}_1$ violate the detection threshold while the results of the other two matrices are negative, BD occurs in PMU measurements and are downweighted by (16).

3) BD in both SCADA and PMU Measurements: Threshold violations in the tests of all three matrices indicate the presence of BD in both SCADA and PMU measurements, and bad SCADA measurements occur in both areas. More specifically, threshold violations in the tests of matrices $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ with negative results for matrix $\boldsymbol{Z}_3$ indicate the presence of BD in both SCADA and PMU measurements, and the BD occurs only in PMU observable area. However, the SCADA

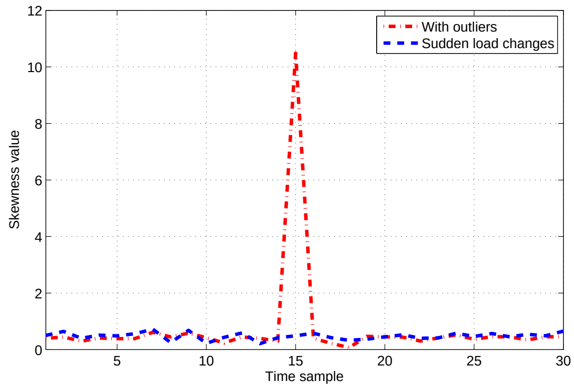


Fig. 4: The skewness test results

measurements in PMU unobservable but SCADA observable areas are free of BD. On the other hand, threshold violations in the tests of matrices $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_3$ with negative results for matrix $\boldsymbol{Z}_2$ indicate the presence of BD in PMU measurements (PMU observable area) and SCADA measurements (PMU unobservable but SCADA observable area), while the SCADA measurements in PMU observable area are free of BD. BD are then downweighted by (16).

4) Smearing Effect Elimination: It is well known that gross measurement errors impact filtering significantly through the error propagation, which makes many residuals related to valid measurements exceed the threshold (called the smearing effect). However, the proposed method is not prone to it because the forecasted measurements depend on the historical data, which are not affected by BD in the received measurements. Therefore, if BD exists, only the components associated with the spurious measurements break the temporal correlation and produce high innovations. PS is adopted to detect the BD all at once and bad measurements can be easily downweighted, which is more computationally attractive than conventional one-by-one BD elimination techniques.

5) Bad Critical Measurements Analysis: In the conventional BD detection and identification method, single or multiple critical measurements appear when measurement redundancy is low and reaches critical levels. The gross errors in these critical measurements are undetectable because the residuals of bad critical measurements are negligible. However, in the proposed method, enhanced measurement redundancy can be obtained from the innovations, thus making the critical measurements non-critical. Therefore, the bad critical measurement can be detected.

6) Topology Error Detection: To detect topology errors, a relative high level of local measurement redundancy and a robust detector with high breakdown point are required. This is because the outliers caused by topology errors are strongly correlated. In our proposed method, because of the enhanced measurement redundancy and the good robustness of PS, measurements corrupted by the topology errors can be identified as outliers, and subsequently topology errors can be detected.

Remark 6: Although the objective of this paper is to enhance the BD detection, a few remarks are made here for completeness of the whole procedure of the estimator. After the BD processing, the SCADA and PMU measurements are expected to be clean. The BD detection and processing is performed using the innovations and PMUs, while the existing static SE using SCADA measurements is not used yet. In other words, the proposed method can be easily integrated with existing static state estimators. To obtain the final SE results, there are several different ways, such as a weighted average of filtered states [31] by SCADA-based and PMU-based estimators.

IV. Numerical Results

In order to demonstrate the effectiveness and robustness of the proposed method, PMU full observable IEEE 14-bus system, PMU partial observable IEEE 14-bus system and 118-bus test system are considered. Several different BD conditions are analyzed, i.e., *single/multiple BD in SCADA or PMU measurements, bad critical measurements, BD in both SCADA and PMU measurements, topology errors*. The hybrid state estimator-based largest normalized measurement residual test (Hybrid) through the identification-by-elimination procedure [9], and the Phasor-aided method in [10], are used to make comparisons. In all tests, the errors in SCADA measurements are assumed to follow a normal distribution with zero mean and standard deviation of 10^{-2}, while standard deviations of PMU measurements are assumed to be 10^{-3}; All tests are implemented on Matlab R2012a and performed on a 2.50 GHz, 8GB of RAM, Intel Core i5 computer.

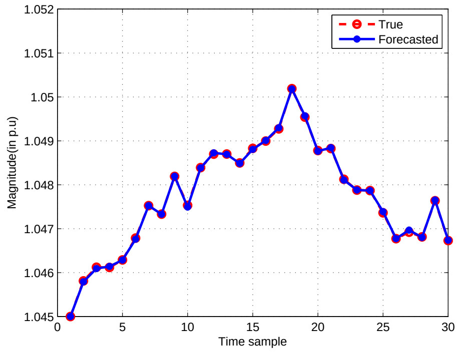
A. Unpredictable Sudden Changes

As emphasized before, [22]–[24] have already proposed effective method for distinguishing system anomalies and our work is to further identify which measurement is incorrect and which topology is wrong. To make this paper compete, we show a few cases to illustrate how the methods in [22]–[24] work. Two different scenarios on the IEEE 14-bus test system are considered at time sample k=15: (i) the real power of the load at bus 2 is suddenly changed to 0.4 p.u. and (ii) 20% error is added to real power flow P_{2-1} for simulating the BD. The skewness test [22] is applied here to distinguish the sudden system changes and BD. The test result is shown in Fig. 4. It can be observed from this figure that no abrupt response of skewness appears in presence of sudden load changes. However, in case (ii) the magnitude of the skewness value at time $t = 15$ has significantly increased, indicating the occurrence of BD. Therefore, the system sudden change and BD can be effectively distinguished. If a sudden change is identified, the state estimation is preinitialized and the BD processing will be only performed once the next time sample comes. Otherwise, the BD including topology error is detected and processed by proposed method.

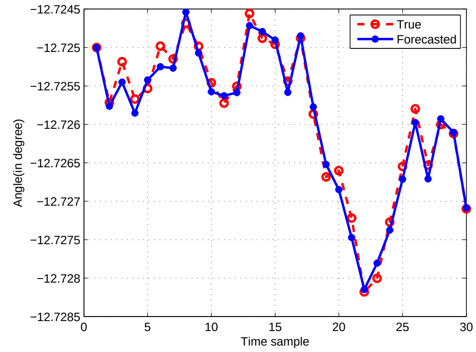
B. Test Results Under Normal System Operation Conditions

When there is no large disturbance, the system is operating under the steady state conditions. In this paper, a random fluctuation with a linear trend (1%-3%) is added to the load curve over a period of time, i.e. 30 time-sample intervals, then, at each time sample, the load flows are calculated and the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2016.2626782, IEEE Transactions on Industrial Informatics



(a) Voltage magnitude



(b) Voltage angle

Fig. 5: The evolution of the voltage magnitude and angle of bus 2 in normal operating conditions

TABLE I: Comparison Results under Different Bad Data Conditions When the System is PMU Observable

Case	Actual BD	Hybrid	Phase-aided	Proposed method
1	P_4	P_4	P_4	P_4
2	P_3, Q_3	$P_3, Q_3, I^i_{2-3}, I^r_{2-3}, I^i_{2-4}$	P_3, Q_3	P_3, Q_3
3	V_2 θ_2	$P_6, Q_3, P_{1-5}, P_{5-2}, V_2, \theta_2, \theta_6, \theta_7$ $I^r_{2-4}, I^r_{2-3}, I^r_{2-5}, I^r_{6-5}, I^r_{6-11}, I^r_{6-13}, I^r_{7-4}, I^r_{9-4}, I^i_{2-1}, I^i_{2-5}$	V_2 θ_2	V_2 θ_2
4	$V_7, \theta_7, I^r_{7-4}, I^r_{7-8}$ $I^r_{7-9}, I^i_{7-4}, I^i_{7-8}$	$P_{4-7}, Q_{7-8}, Q_{8-7}, P_9, Q_9, \theta_2, \theta_6, \theta_7, \theta_9$ $I^r_{9-7}, I^r_{9-10}, I^r_{9-14}, I^i_{2-4}, I^i_{7-4}, I^i_{7-8}, I^i_{7-9}$ $V_7, I^r_{2-4}, I^r_{7-4}, I^r_{7-9}, I^r_{9-4}, I^i_{9-4}, I^i_{9-7}, I^i_{9-10}$	$V_7, \theta_7, I^r_{7-4}, I^r_{7-8}$ $I^r_{7-9}, I^i_{7-4}, I^i_{7-8}$	$V_7, \theta_7, I^r_{7-4}, I^r_{7-8}$ $I^r_{7-9}, I^i_{7-4}, I^i_{7-8}$
5	P_4 V_2, θ_2	$V_2, \theta_2, V_2, P_{5-2}, P_4, \theta_6, \theta_7, P_{1-2}, P_{2-3}, Q_3, Q_6, I^r_{3-4}, I^r_{2-4}$ $I^r_{2-3}, I^r_{2-5}, I^i_{6-5}, I^r_{1-5}, I^r_{4-5}, I^r_{7-4}, I^i_{2-1}, I^i_{2-5}, P_{4-3}$	$V_2, \theta_2, P_{5-2}, P_3, I^i_{2-1}$ $P_4, P_{4-2}, I^r_{2-4}, I^r_{2-1}, I^r_{2-3}, I^i_{2-4}$	P_4 V_2, θ_2

time series of the system states are obtained. Fig. 5 presents the evolution of the voltage magnitude and angle of bus 2 in normal operating conditions as an example. It is observed that the forecasted state is close to the true operating state and can be used to help us process the BD.

1) PMU Observable System: To show the performance of the Hybrid method, the phasor-aided method and the proposed method, the PMU observable system is considered. The detailed configuration of the SCADA and PMU measurements can be found in [10]. The detection thresholds for all largesr normalized measurement residual-based test are set to 3, while the detection threshold for PS is $\chi^2_{2,0.975} = 7.38$

Case 1: A single injection active power with 30% error of the measured value at bus 4 in SCADA measurements;

Case 2: A critical measurement set (P_3, Q_3) with 30% error;

Case 3: A single PMU voltage vector of bus 2 with 10% error;

Case 4: Multiple bad PMU measurements, i.e., all PMU measurements related to bus 7 with 10% error;

Case 5: Single injection active power with 20% error of bus 4 and the magnitude and phase angle of voltage phasors at bus 2 with 15% error;

Table I presents the obtained results of the five cases, where I^r and I^i represent the real rectangular component and imaginary rectangular component of the current I, respectively. From the results, it is observed that conventional normalized measurement residual-based test by hybrid method can effectively address the single bad SCADA measurements. However, it cannot handle the smearing effect since the valid measurements are incorrectly identified as BD (see case 4 for example). On the other hand, when BD occurs in the critical measurement set, the hybrid method fails to detect them. The same problem does not happen in Phase-aided method and proposed method. However, when both bad SCADA and PMU measurements occur, the phasor-aided methods fail to mark the BD, while proposed method can still detect them accurately without being affected by the smearing effect.

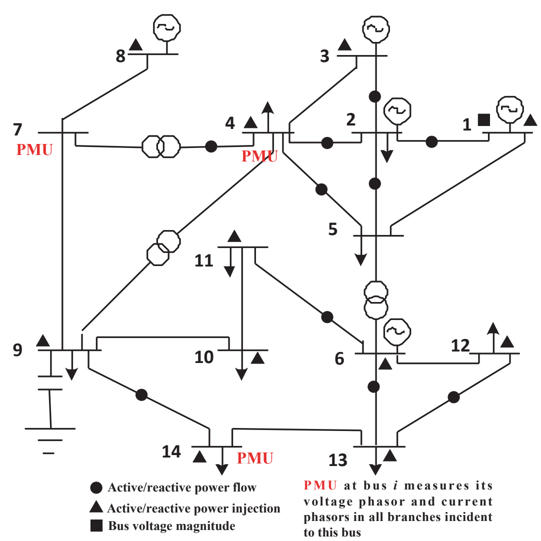


Fig. 6: Measurement configuration for PMU partial observable IEEE 14-bus system

2) PMU Partially Observable System: As shown in the former section, the Hybrid and phasor-aided method are unable to effectively detect all the BD even the PMU can observe the

TABLE II: Comparison Results for PMU Partial Observable System

Case	Actual BD	Hybrid	Phase-aided	Proposed method
6	P_4	P_4	P_4	P_4
7	Q_{12}, P_{12-13}	$Q_{12}, P_{12-13}, Q_{6-13}, Q_{12-13}$	$Q_{12}, P_{12-13}, Q_{6-13}, Q_{12-13}$	Q_{12}, P_{12-13}
8	V_7 P_4	$P_4, \theta_{14}, P_{14}, Q_{14}, P_{11}$ $P_{11-6}, Q_{11-6}, I^r_{4-5}, I^i_{4-5}$	V_7, P_4 $P_{14}, Q_{14}, P_{11-6}, Q_{11-6}$	V_7 P_4
9	V_7, θ_7 P_{12}	$V_7, V_4, \theta_4, \theta_7, I^r_{4-5}, I^i_{4-5}, P_4, Q_4, P_{14}, Q_4$ $P_{11-6}, Q_{11-6}, P_{4-7}, Q_{4-7}, P_{10}, Q_{10}, P_{11}$	$V_7, \theta_7, P_{11}, Q_{11}, P_{10}$ $Q_{10}, P_{11-6}, Q_{11-6}, Q_{11}$	V_7, θ_7 P_{12}

TABLE III: Comparison Results for IEEE 118-bus Test System

Case	Actual BD	Hybrid	Phase-aided	Proposed method
10	Q_{63-59}	Q_{63-59}, P_{63-59}	Q_{63-59}	Q_{63-59}
11	Q_{76-118}, Q_{28-27} $P_{28},\ P_{27}$	$Q_{76-118}, Q_{28-27}, P_{28-27}$ $P_{28},\ P_{27},\ P_{28-29},\ Q_{28-29}$	Q_{76-118}, Q_{28-27} $P_{28},\ P_{27},\ P_{28-27}$	Q_{76-118}, Q_{28-27} $P_{28},\ P_{27}$
12	I^r_{52-53}, I^i_{52-53}	I^r_{52-53}, I^i_{52-53}	I^r_{52-53}, I^i_{52-53}	I^r_{52-53}, I^i_{52-53}
13	$V_{114}, P_{28-27}, P_{27}$ P_{25-27}	$V_{114}, I^r_{27-114}, I^i_{27-114}, P_{27}$ $P_{28-27}, Q_{28-27}, P_{25-27}, Q_{25-27}$	$V_{114}, P_{28-27}, Q_{28-27}$ P_{25-27}, Q_{25-27}	$V_{114}, P_{28-27}, P_{27}$ P_{25-27}

whole system. In this section, we aim to show how all three methods behave when the system is partially PMU observable.

In this study, 3 PMUs are deployed on buses 4, 7 and 14 to ensure the system to be partially PMU observable, where busbar voltage magnitudes and phase angles are measured, and four branch currents are measured. On the other hand, 31 SCADA measurements are considered, which include 16 power flows, 14 power injections (all of them are in active/reactive pairs), and 1 bus voltage magnitude at bus 1. Note that, power injections at buses 4, 8, 10, 14 and power flow in lines 4–7 are critical measurements without PMU measurements. The detailed measurement configuration of the PMU partially observable system is depicted in Fig. 6. In the tests, BD are generated by adding large deviations into the original measurements. The following cases are considered:

Case 6: A critical measurement P_4 with 20% error;

Case 7: Multiple BD, i.e., reactive power injection at bus 12 and real power flow in bus 12 to bus 13 with 20% errors;

Case 8: A single PMU voltage vector at bus 7 with 20% error and P_4 with 20% error;

Case 9: Multiple BD, i.e., voltage and angle of bus 7 with 20% error and P_{12} with 20% error;

Simulation results for Cases 6–9 are shown in Table II. It can be seen from Case 9 that all of the three methods can effectively detect single BD. However, when multiple BD occurs in PMU unobservable area (see case 7), Hybrid and phasor-aided methods could not handle it due to the smearing effect. This does not happen in the proposed method. Furthermore, when BD in SCADA measurements and PMU measurements occurs simultaneously, the phasor-aided method is unable to detect them. However, the proposed method can still work, which can be shown by Cases 8 and Case 9. For example, in Case 8, when BD exists in V_7 and P_7, the Hybrid and phasor-aided methods incorrectly identify many valid measurements as BD, such as P_{14}, Q_{14}, etc. By contrast, the proposed method can still accurately detect them.

3) IEEE 118-bus System Test Results: This system is fully observable with 150 pairs of SCADA measurements including 39 pairs of injection measurements and 111 pairs of flow measurements. Note that, there are 5 critical power injections and 22 critical power flows. Only 19 PMUs are installed, allowing the system to be partially observable. The detailed system measurements placement and topology can be found in [8]. The following cases are considered:

TABLE IV: Comparison of Computation Time in IEEE 118-bus Test System under Different Scenarios

Scenario	Hybrid	Phase-aided	Proposed (Centralized)	Proposed (Parallel)
Without BD	0.679s	0.779s	1.13s	0.72s
Single BD	1.456s	1.642s	1.70s	0.82s
Two BD	2.346s	1.689s	1.76s	0.88s

Case 10: A critical measurement Q_{63-59} with 20% error of the measured value;

Case 11: Multiple BD, i.e., reactive power flow Q_{76-118}, reactive power flow Q_{28-27}, active power injections P_{28} and P_{27} are all with 20% errors (in this small area, two PMUs are installed at buses 27 and 114, where bus 27 is observable from PMU and bus 28 is not observable though the PMUs);

Case 12: A single PMU current in line 52-53 with 20% error;

Case 13: Multiple BD, i.e., voltage V_{114}, active power flows, P_{28-27} and P_{25-27} and active power injection P_{27} are all with 20% errors.

Test results for Cases 10–13 are shown in Table III. Similar to the results of IEEE 14-bus test system, the proposed method demonstrates better performance than existing methods in detecting BD.

4) Topology Error Detection: Since the Hybrid and Phasor-aided methods by their designs have assumed the correct topology for state estimation, only the results of the proposed method is shown here. Three cases are tested in *PMU partially observable* IEEE 14-bus and 118-bus systems, i.e.,

Case 14: A false report of the circuit breaker between buses 12 and 13 in IEEE 14-bus system;

TABLE V: Comparison of Computation Time in IEEE 2383-bus Test System under Different Scenarios

Scenario	Hybrid	Phase-aided	Proposed (Centralized)	Proposed (Parallel)
Without BD	12.96s	14.56s	20.82s	6.26s
Single BD	26.31s	19.27s	22.45s	6.93s
Two BD	39.84s	20.05s	23.05s	7.02s

Case 15: A false outage of the transformer line between bus 5 and 6 in IEEE 14-bus system;

Case 16: A false report of the circuit breaker between buses 110 and 112 in IEEE 118-bus system;

In case 14, the measurements, P_{12-13}, Q_{12-13}, P_{12}, Q_{12}, P_{13} and Q_{13} are flagged as outliers. From the results, we observe that all the measurements related to line 12-13 are corrupted, indicating the topology error between line 12-13. For case 15, measurements P_5, Q_5, P_6, Q_6, P_{5-6} and Q_{5-6} are detected as outliers, which means that the switch/ breaker on the transformer line 5-6 incorrectly reports the status of the line. Finally, in case 15, all measurements $P_{110-112}$, $Q_{110-112}$, P_{110}, Q_{110}, P_{112} and Q_{112} corresponding to line 110-112 are identified as outliers, the topology error between line 110-112 is therefore declared. It should be noted that in all three cases, only the measurements corresponding to incorrect topology are marked as outliers without being affected by smearing effect because of the good robustness of PS and the enhanced measurement redundancy by innovation vectors.

C. Assessment of Computational Efficiency

In this subsection, the computational efficiency of the proposed method (implemented in centralized or parallel manners) in the 118-bus test system is compared with the Hybrid and the Phasor-aided methods. The Cholesky factorization technique is adopted to increase the numerical stability. Table IV shows the computing time for four methods. It can be seen from this table that the conventional hybrid method has the fastest speed if there is no BD, followed by the parallel implementation of the proposed method. The centralized version of the proposed method is the slowest one as it needs additional time to perform centralized processing of all information. However, when multiple BD occurs, the conventional one-by-one BD elimination method is time consuming. This does not happen for the Phase-aided and the proposed parallel method since they can detect and eliminate the BD all at once. The computing time of the proposed method with the parallel implementation is very fast, demonstrating its ability for real-time implementation.

To further demonstrate the effectiveness of the proposed method for large-scale power systems, various tests have been conducted on the Polish 2383-bus system [32]. There are 1296 PMU and 4960 SCADA measurements of this test system, including 360 voltage phasors and 936 current phasors, 4760 power injections, 100 voltage magnitude measurements and 100 zero injections. According to the observability check, the system is divided into 18 areas, among which the largest one has 308 buses. Three scenarios are simulated and tested, i.e., without outliers, with single outlier and with two outliers randomly placed in the measurement set. The conclusions of BD detection are the same as IEEE 14 and 118 systems. To avoid repeating the similar BD detection results, only the computing time is presented. The test results are shown in Table. V. From this table, we find that the conventional one-by-one BD elimination strategy is the most time consuming method in presence of BD, followed by the proposed centralized and Phase-aided methods. With the parallel implementation of the proposed method, the computation time has been reduced significantly.

D. Discussions

State estimation aims to process a set of measurements that are assumed to be taken at the same snapshot in time to obtain a best estimate of the states of electric power systems. In this paper, the SCADA and PMU measurements are assumed to be taken at the same snapshot and further used for BD detection and state estimation. However, in some situations, due to the different sampling rates of SCADA and PMU, these two types of measurements may not arrive simultaneously. In this case, the detection and suppression of BD in PMU observable areas is not affected as only forecasted states and PMU measurements are used. But the detection of suspicious BD in PMU unobservable but SCADA observable areas will be affected as the PMU sampling rate is higher than SCADA. To mitigate this time skewness problem between SCADA and PMU measurements, the PMU optimal buffering strategy proposed in [17], [33], [34] can be used. The key idea is to buffer the PMU measurements between the two consecutive SCADA scans in an optimal way so that the statistical information of the PMU measurements is effectively extracted. Once new SCADA measurements arrive, the buffered information represented by its sample mean and sample covariance matrix is used together with SCADA measurements for state estimation. By using this strategy, SCADA and PMU measurements are coordinated to be at the same snapshot, thus the proposed method can be applied to detect BD.

V. Conclusion

A robust BD detection method is proposed by exploiting the temporal correlation and the statistical consistency of the measurements. Three innovation vectors are defined and used to develop the innovation matrices for BD detection by projection statistics. The proposed method can address various types of BD, such as the single/multiple BD, BD smearing effect, bad critical measurements, topology error and simultaneous occurrence of BD in SCADA and PMU measurements. Comparisons with existing methods on IEEE 14-bus, 118-bus and Polish 2383-bus test systems demonstrate its effectiveness and robustness. In the future work, we aim to investigate the performance of the proposed method for detecting false data injection attacks on the power system state estimator [35], [36].

References

[1] R. C. Pires, L. Mili, F. A. Becon Lemos, “Constrained robust estimation of power system state variables and transformer tap positions under erroneous zero-injections,” *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1144–1152, May 2014.

[2] W. Zhang, W. Liu, C. Zang, L. Liu, "Multi-agent system based integrated solution for topology identification and state estimation," *IEEE Trans. Industrial Informatics.*, DOI: 10.1109/TII.2016.2543200, 2016.

[3] J. B. Zhao, G. X. Zhang, K. Das, Ge. N. Korres, N. M. Manousakis, A. K. Sinha, Z. Y. He, "Power system real-time monitoring by using PMU-based robust state estimation method," *IEEE Trans. Smart Grid.*, vol. 7, no. 1, pp. 300–309, Jan 2016.

[4] L. Mili, T. V. Cutsem, M. Ribbens-Pavella, "Hypothesis testing identification: a new method for bad data analysis in power system state estimation," *IEEE Trans. Power Syst.*, vol. 103, no. 11, pp. 3239–3252, Nov. 1984.

[5] L. Mili, V. Phaniraj and P. Rousseeuw, "Least median of squares estimation in power systems," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 511–523, May 1991.

[6] 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, Nov. 2006.

[7] B. Gou, R. G. Kavasseri, "Unified PMU placement for observability and bad data detection in state estimation," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2573–2580, Nov. 2014.

[8] A. Tarali, A. Abur, "Bad data detection in two-stage state estimation using phasor measurements," in *Proceedings of the 3rd IEEE PES Innovative Smart Grid Technologies Conference*, 2012: 1–8.

[9] G. N. Korres and N. M. Manousakis, "State estimation and bad data processing for systems including PMU and SCADA measurements," *Elect. Power Syst. Res.,* vol. 81, no. 7, pp. 1514–1524, Jul. 2011.

[10] M. B. Do Coutto Filho, J. C. Stacchini de Souza, M . A. R. Guimaraens, "Enhanced bad data processing by phasor-aided state estimation," *IEEE Trans. Power Syst.,* no. 99, pp. 1–10, 2014.

[11] J. B. Zhao, G. X. Zhang, M. L. Scala, J. H. Zhang, "Multistage phasor-aided bad data detection and identification," in *Proc. IEEE Power Eng. Soc. General Meeting*, July. 26–30, 2015.

[12] P. J. Rousseeuw, B. C. van Zomeren, "Unmasking multivariate outliers and leverage points," *Journal of the American Statistical Association*, Vol. 85, No. 411, pp. 633–639, 1990.

[13] P. J. Rousseeuw, B. C. van Zomeren, "Robust distances: simulations and cutoff values," In Stahel, W. and Weisberg, S. (Eds.), Directions in Robust Statistics and Diagnostics, Part II, Vol. 34, The IMA Volumes in Mathematics and its Applications, Springer–Verlag, New York, 1991, pp. 195–203.

[14] M. B. Do Coutto Filho, J. C. Stacchini de Souza, "Forecasting-aided state estimation–Part II: Implementation," *IEEE Trans. Power Syst.,* vol. 24, no. 4, pp. 1678–1685, 2009.

[15] J. B. Zhao, G. X. Zhang, Z. Dong, M. La Scala, "Robust forecasting aided power system state estimation considering state correlations," *IEEE Trans. on Smart Grid*, DOI: 10.1109/TSG.2016.2615473, in press.

[16] M. Hassanzadeh, C. Y. Evrenosoglu, L. Mili, "A short-term nodal voltage phasor forecasting method using temporal and spatial correlation," *IEEE Trans. Power Syst.,* Vol. 31, no. 5, pp. 3881-3890, 2016.

[17] Y. Chakhchoukh, V. Vittal, G. T. Heydt, "PMU based state estimation by integrating correlation," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 617–626, Mar. 2014.

[18] R. J. Bessa, A. Trindade, V. Miranda, "Spatial-temporal solar power forecasting for smart grids," *IEEE Trans. Industrial Informatics.*, Vol. 11, no. 1, pp. 232–241, 2015.

[19] J. B. Zhao, G. X. Zhang, M. L. Scala, "PMU based robust dynamic state estimation method for power systems," in *Proc. IEEE Power Eng. Soc. General Meeting*, July. 26–30, 2015.

[20] G. N. Korres, N. M. Manousakis, "Observability analysis and restoration for state estimation using SCADA and PMU data," *IEEE Power and Energy Society General Meeting*, pp: 1–8, 2012.

[21] N. M. Manousakis, G. N. Korres, P. S. Georgilakis, "Taxonomy of PMU placement methodologies," *IEEE Trans. Power Syst.*, Vol. 27, no. 2, pp: 1070–1077, 2012.

[22] K. Nishiya, J. Hasegawa, T. Koike, "Dynamic state estimation including anomaly detection and identification for power systems," *IEE Proc. Gener., Transm. Distrib.*, vol. 129, no. 5, pp. 192–198, Sep. 1982.

[23] G. Valverde and V. Terzija, "Unscented Kalman filter for power system dynamic state estimation," *IET Gener. Transm. Distrib.*, vol. 5, no. 1, pp. 29–37, Jan. 2011.

[24] Z. Guo, S. Li, X. Wang, W. Heng, "Distributed point-based gaussian approximation filtering for forecasting-aided state estimation in power systems," *IEEE Trans. Power Syst.,* vol. 31, no. 4, pp. 2597–2608, Jul. 2016.

[25] P. J. Huber, *Robust Statistics*. New York: Wiley, 1973.

[26] M. Gandhi, L. Mili, "Robust Kalman filter based on a generalized maximum-likelihood-type estimator, *IEEE Trans. Signal Processing*, vol. 58, no. 5, pp. 2509–2520, 2010.

[27] M. Gasko, D. Donoho, "Influential observation in data analysis," American Statistical Association, Proceedings of the Business and Economic Statistics Section, 1982, pp. 104–110.

[28] L. Mili, M. Cheniae, N. Vichare, P. Rousseeuw, "Robust state estimation based on projection statistics," *IEEE Trans. Power Syst.*, vol. 11, no. 2, p. 1118–1127, May 1996.

[29] D. Mallieu, T. V. Cutsem, P. Rousseaux, and M. Ribbens-Pavella, "Dynamic multilevel filtering for real-time estimation of electric power system," Control Theory and Adv. Tech No. (C-TAT Special Issue on Large-Scale Systems), vol. 2, no. 2, pp. 255–272, 1986.

[30] P. Rousseaux, D. Mallieu, T. V. Cutsem, and M. Ribbens-Pavella, "Dynamic state prediction and hierarchical filtering for power system state estimation," *Automatica*, vol. 24, no. 5, pp. 595–618, Sep. 1988.

[31] B. P. Gibs, *Advanced Kalman Filtering, Least-Squares and Modeling: A Practical Handbook*. Hoboken, NJ, USA: Wiley, 2011.

[32] [Online]. Available: http://www.pserc.cornell.edu/matpower.

[33] Q. Zhang, Y. Chakhchoukh, V. Vittal, Gerald T. Heydt, N. Logic, S. Sturgill, "Impact of PMU measurement buffer length on state estimation and its optimization," *IEEE Trans. Power Syst.*, Vol. 28, no. 2, pp: 1657-1665, 2013.

[34] V. Murugesan, Y. Chakhchoukh, V. Vittal, Gerald T. Heydt, N. Logic, S. Sturgill, "PMU data buffering for power system state estimators," *IEEE Power and Energy Technology Systems Journal*, Vol. 28, no.2, pp: 94–102, 2015.

[35] J. B. Zhao, G. X. Zhang, M. La Scala, Z. Dong, C. Chen, J. Wang, "Short-term state forecasting-aided method for detection of smart grid general false data injection attacks," *IEEE Trans. Smart Grid*, in press.

[36] J. B. Zhao, G. X. Zhang, Z. Dong, K.P. Wong, "Forecasting-aided imperfect false data injection attacks against power system nonlinear state estimation," *IEEE Trans. Smart Grid*, Vol. 7, no. 1, pp. 6–8, 2016.

Junbo Zhao (S'13) is pursuing the Ph.D. degree at the Bradley Department of Electrical Computer Engineering, Virginia Polytechnic Institute and State University, USA. His research interests are in the theoretical and algorithmic studies in power system state estimation, power system operation and cyber security, signal processing. He is a member of IEEE PES and IEEE Communications Society.

Gexiang Zhang (M'03), received Ph.D. degree in 2005 from Southwest Jiaotong University, Chengdu, China. Since the year of 2005, he has been a Professor at the School of Electrical Engineering in Southwest Jiaotong University, where he leads the research group of Nature-Inspired Computation and Smart Grid (NICSG). He is also an Adjunct Professor in Xihua University.

His research interests include natural computing and smart grid. He has published over 100 scientific papers in international journals or conferences. He was selected as "New Century Excellent Talents in University" from Chinese Ministry of Education.

Massimo La Scala (M'88–SM'99–F'07) received the B.S. and Ph.D. degrees in electrical engineering from the University of Bari, Bari, in 1984 and 1989, respectively.

He is currently Professor of Electrical Energy Systems and Director of Lab ZERO, a lab for the development of sustainable technologies in smart cities, at the Politecnico di Bari, Italy. His research interests are in the area of power system analysis and control, smart grids, sustainable and smart cities.

Zhaoyu Wang (M'15) is an Assistant Professor at Iowa State Unviersity. He received Ph.D. degree in Electrical and Computer Engineering from Georgia Institute of Technology in 2015. He received B.S. degree and M.S. degree in Electrical Engineering from Shanghai Jiaotong University in 2009 and 2012, respectively, and the M.S. degree in Electrical and Computer Engineering from Georgia Institute of Technology in 2012.

His research interests include power distribution systems, microgrids, renewable integration, self-healing resilient power systems, and voltage/VAR control. He was a Research Aid in 2013 at Argonne National Laboratory and an Electrical Engineer at Corning Incorporated in 2014. He is an editor of IEEE TRANSACTIONS ON SMART GRID and IEEE POWER ENGINEERING LETTERS.