Robust Forecasting Aided Power System State Estimation Considering State Correlations

Robust Forecasting Aided Power System State Estimation Considering State Correlations

Junbo Zhao, *Student Member, IEEE*, Gexiang Zhang, *Member, IEEE*,
Zhao Yang Dong, *Senior Member, IEEE*, Massimo La Scala, *Fellow, IEEE*

***Abstract*—With the increase of load fluctuations and the integration of stochastic distributed generations (DGs), there have been more and more research interests in forecasting-aided state estimation. In this paper, we propose a robust generalized maximum likelihood (GM)-estimator based power system forecasting-aided state estimation (GM-PSE), which integrates the statistical characteristics of both loads and DGs, i.e., spatial and temporal correlations. A first order vector auto-regressive model (VAR(1)) is developed to capture the statistical characteristics of load and DGs, facilitating short-term loads and DGs forecasting. These forecasted power injections are further combined with power balance equations to derive a new state transition model, where the relationship between forecasted state vector and predicted power injections is expressed explicitly. After that, a redundant batch regression model that simultaneously processes predicted state vector and received observations is derived, allowing the development of a robust estimator. To this end, we propose a robust GM-estimator that leverages modified projection statistics and a Huber convex score function, to bound the influence of observation outliers while maintaining its high statistical estimation efficiency. Finally, the iteratively reweighted least squares (IRLS) algorithm is adopted to solve the GM-estimator. Numerical comparisons on IEEE benchmark systems with DGs integration demonstrate the efficiency and robustness of the proposed method.**

***Index Terms*—State estimation, distributed generation, vector auto-regression, forecasting-aided state estimation, robust estimator, power systems.**

I. Introduction

REAL-time states play a vital role for reliable and secure power system operations. These states are usually obtained by the power system static state estimation (SE) using redundancy measurements from SCADA or PMUs. By the definition of static SE, the current state is determined exactly by present measurements regardless of the previous state. However, due to the continuous variations of loads and generators, a power system is slowly changeable with time rather than static. Once loads of a power system change, the generators have to keep up with the changes and consequently power flows and injections at all buses will change [1]. On the other hand, in recent years, with the integrations of more and more renewable-based distributed generation (DG) and microgrids, the power grid faces with new stochastic operating behaviors and dynamics [2]. For example, the uncertainty of loads and stochastic (intermittent) characteristics of DGs increases the probability of sudden changes in bus voltage phasor in a short time-frame [3]. These changes are mainly driven by the changes in active or reactive power injections from DGs or loads. Consequently, the estimation results from static state estimation (SSE) may not effectively reflect the present system operation states.

To address this issue, the forecasting-aided SE (some researchers also call it dynamic state estimation (DSE)) [4]– [13] was proposed using various Kalman filter (KF) like methods, such as extended Kalman filter (EKF) and unscented Kalman filter (UKF). The investigations in [6] extended the regression-based state forecasting method developed in [7] to consider fast sampling rates of voltage and phasor measurements from PMUs in DSE. The accuracy of the predicted state can be greatly improved and the trend in state variations can also be provided for power system operators. In [8], a PMU placement strategy in EKF-based DSE was developed to track states of a power grid. The number and locations of PMUs installed in the system to ensure a satisfactory state tracking performance were discussed. In [9], a method to extract the dynamic real-time model of an electric power system using PMU and SCADA data obtained in substations was proposed. The method is an extension of the standard SSE and is mainly based on data analysis techniques. The study in [10] presented a PMU-based DSE algorithm and discussed the impact of PMUs on a DSE technique. The effects of the number and locations of PMUs, and weighting factors on the accuracy of DSE were discussed. While in [11]–[13], DSE using different measurement weighting functions was proposed to handle outliers and system sudden changes. In [14], Gaussian mixtures models were adopted to account for the stochastic characteristic of power flow in SE process. But this method is a SSE and does not have forecasting ability. In [15], a short-term load forecasting method based static SE was proposed to consider the impacts of load variations on SE. However, the DG integration was not considered. In [1], the temporal correlations among loads and DGs were modeled and integrated into the EKF-based DSE method, resulting in improvement of SE accuracy. But it does not account for the

This work was partially supported by the National Natural Science Foundation (61672437, 61373047). *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 the School of Electrical Engineering, Southwest Jiaotong University, Chengdu, 610031 China (zhgxdylan@126.com).

Zhao Yang Dong is with the School of Electrical and Information Engineering, The University of Sydney, NSW 2006, Australia (e-mail: zydong@ieee.org).

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).

spatial correlations among loads and DGs [16], [17]. Besides, since KF like methods are not robust to outliers, bad data in measurements would produce large biases to SE results.

To solve these problems, we propose a robust generalized maximum likelihood (GM)-estimator based power system forecasting-aided state estimation (GM-PSE). This estimator provides an alternative way to integrate both spatial and temporal correlations of DGs and loads explicitly, and is also able to mitigate the negative impacts caused by bad measurements. To be specific, the first order vector auto-regressive model (VAR(1)), which is a good time-series analysis technique for modeling statistical characteristics of random variables, is developed to capture the spatial and temporal correlations of the loads and DGs. The forecasted system power injections (active and reactive power injections of the loads and DGs) are further used to derive a new dynamic state transition model through power balance equations, where the relationship between system state and the system power injections is presented explicitly. We then derive the redundant batch regression model that simultaneously processes the predicted states and the present observations. This redundancy regression model is very helpful for developing robust estimator to handle outliers occurred in the observed measurements. Thus, this paper extends the robust Kalman filter in [18] for a linear system to a nonlinear power system SE problem. The GM-estimator is developed to bound the influence of observation outliers while maintaining high estimation statistical efficiency. Finally, the iteratively reweighted least squares (IRLS) algorithm is used to solve the GM-estimator.

This paper is organized as follows: the spatial and temporal correlations of loads and DGs are modeled and analyzed in Section II. The state transition model and the batch regression model are derived in Section III. Section IV presents the proposed robust Schweppe-type Huber GM estimator for solving the regression problem. Numerical simulation results on IEEE test systems with DGs integration are provided in Section V, followed by conclusions and future work in Section VI.

II. Shot-Term Load and DGs Forecasting Integrating Spatial and Temporal Correlations

Similar to [1], [19], we consider only the randomness caused by loads and DGs in modeling and analysis for simplicity. In the literature, time series model based statistical methods [20], [21], i.e., auto-regressive (AR) process, auto-regressive moving average (ARMA) process, etc., for loads and DGs forecasting have been widely used and demonstrated to be effective. For example, a vector auto-regressive (VAR) model for solar power forecasting was proposed in [17]. Field data based tests were conducted to highlight its effectiveness and the improved performance compared with a simple auto-regressive (AR) model. [20] has evaluated both AR and AR with exogenous input (ARX) models for solar power forecasting, where the latter takes numerical weather predictions (NWPs) as input. The results suggested that AR is suitable for short-term forecasting while AR with exogenous input (ARX) achieves better performance in a long-term forecasting. In [21]–[23], AR-type time series models have been shown to be good statistical methods for short-term wind power forecasting. Since we are interested in short-term loads and DGs forecasting, AR-type time series model is adopted. On the other hand, except for temporal correlations among loads and DGs in the same geographic area due to weather, economic, social behavior, the changes of some loads and DGs also affect other generators and loads in the same geographic area, presenting cross-correlation, i.e., spatial correlations [24]. Therefore, to perform short-term predictions of loads and DGs considering their spatial and temporal correlations, the VAR model is advocated. The time series for active (P_g) and reactive (Q_g) power of DG and active (P_l) and reactive (Q_l) power of load by the order p and dimension D VAR model, i.e., VAR$_D(p)$, is expressed as,

$$\boldsymbol{X}_k = \boldsymbol{\varphi}_1 \boldsymbol{X}_{k-1} + \cdots \boldsymbol{\varphi}_p \boldsymbol{X}_{k-p} + \boldsymbol{\varepsilon}_k, \quad (1)$$

where $\boldsymbol{X} = [\boldsymbol{P}_l \, \boldsymbol{Q}_l \, \boldsymbol{P}_g \, \boldsymbol{Q}_g]^T$; $\boldsymbol{\varphi}_1, \cdots \boldsymbol{\varphi}_p \in \mathbb{R}^{D\times D}$ are coefficient matrices; $\boldsymbol{\varepsilon_k} \in \mathbb{R}^D$ is Gaussian noise and assumed to follow $\mathcal{N}(0, \boldsymbol{S})$, where the positive definite matrix $\boldsymbol{S}$ is not necessary diagonal. In this model, temporal correlations are contained in the diagonal elements of the estimated $\widehat{\boldsymbol{\varphi}}$ terms and the error covariance matrix $\widehat{\boldsymbol{S}}$, while spatial correlations are reflected in the non-diagonal elements of $\widehat{\boldsymbol{\varphi}}$ terms and $\widehat{\boldsymbol{S}}$. To estimate these auto-regression model parameters, the Yule-Walker estimator (YW) [25] can be used.

By multiplying by $\boldsymbol{X}_{k-j}^T, j = 1, \cdots, p$ on both sides of (1) and taking the expectations, we obtain

$$\boldsymbol{\Pi}(i) = \sum_{i=1}^{p} \boldsymbol{\Pi}(i-j), \; i = 1, \cdots, p\,, \quad (2)$$

where $\boldsymbol{\Pi}(0), \cdots, \boldsymbol{\Pi}(j)$ can be estimated by using the sample covariance matrices

$$\boldsymbol{\Pi}(\ell) = \frac{1}{M}\sum_{k=1}^{M-\ell} (\boldsymbol{X}_{k+\ell} - \boldsymbol{\mu})(\boldsymbol{X}_{k+\ell} - \boldsymbol{\mu})^T, \quad (3)$$

where $\ell \in [0, \; M-1]$, M is the number of considered history terms; $\boldsymbol{\mu}$ is the sample mean of the chosen historical data. In order not to violate the stationariness and to ensure acceptable forecasting accuracy, the value of $M = 20$ is chosen [12], [26]. Then, the forecasting error covariance matrix can be calculated through [26]

$$\widehat{\boldsymbol{S}} = \boldsymbol{\Pi}(0) - \sum_{i=1}^{p} \widehat{\varphi}_j \boldsymbol{\Pi}(j). \quad (4)$$

In this paper, following previous work [3], [24], the VAR(1) is adopted here, yielding the following forecasting model

$$\boldsymbol{X_k} = \widehat{\boldsymbol{\varphi}}_{\boldsymbol{1}} \boldsymbol{X_{k-1}} + \boldsymbol{\varepsilon_k}, \quad (5)$$

with the forecasting error covariance matrix $\widehat{\boldsymbol{S}}_k = \boldsymbol{\Pi}(0) - \boldsymbol{\Pi}(1)$, where $\widehat{\varphi}_1 = \boldsymbol{\Pi}(1)\cdot\boldsymbol{\Pi}(0)^{-1}$.

Remark 1: There might be a singularity problem of inverting the matrix $\boldsymbol{\Pi}(0)$ due to small variations in nodal loads and DGs. To mitigate this issue, the concept of "eligible buses" and its related determination procedures in [3] can be used.

III. Derivation of State Transition Model and Batch Regression Model

A. Derivation of the State Transition Model

As mentioned above, the system dynamic changes are driven by loads and DGs. This means the system state vector $\boldsymbol{X}$ should be the power injections of loads and DGs, while the conventional state vector $\boldsymbol{x}$ including nodal voltage magnitudes and angles should be regarded as the algebraic state vector. In this paper, the algebraic state vector $\boldsymbol{x}$ is augmented with $\boldsymbol{X}$, leading to the full system state vector $\boldsymbol{\Xi} = [\boldsymbol{X}^T \ \boldsymbol{x}^T]^T$. In fact, power flows at any observed node is a linear combination of real and reactive power of loads and generators. Besides, for any node of a power system, the power balance equation $\boldsymbol{f}$ should be satisfied,

$$\boldsymbol{f}(\boldsymbol{\Xi}) = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{X}) = \mathbf{0}. \tag{6}$$

The full derivation on the above equation indicates

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} d\boldsymbol{x} + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}} d\boldsymbol{X} = \boldsymbol{H} d\boldsymbol{x} + d\boldsymbol{X} = \mathbf{0}, \tag{7}$$

where $\boldsymbol{H} = \partial \boldsymbol{f}/\partial \boldsymbol{x} = \partial \boldsymbol{h}/\partial \boldsymbol{x}$ is the Jacobin matrix and $\boldsymbol{h}(\boldsymbol{x})$ is the power flow equation vector that relates the power injections of load and generator to the system algebraic state vector $\boldsymbol{x}$; $\boldsymbol{I} = \partial \boldsymbol{f}/\partial \boldsymbol{X}$ is an identity matrix; $d\boldsymbol{X}$ is the active and reactive power differences of loads or generators. Therefore, the commonly used algebraic state vector can be calculated as

$$d\boldsymbol{x} = -\boldsymbol{H}^{-1} d\boldsymbol{X}. \tag{8}$$

In a short time interval, a power system is assumed to operate under quasi-steady operation status, and loads and generators do not have large sudden changes. Thus, for two successfully adjacent system equilibrium operating points, i.e., $\boldsymbol{f}(\boldsymbol{x}_{k+1}, \boldsymbol{X}_{k+1}) = \mathbf{0}$, $\boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{X}_k) = \mathbf{0}$, $\boldsymbol{H}$ can be assumed to be constant. Therefore, we can derive from (8)

$$\boldsymbol{x}_{k+1} - \boldsymbol{x}_k = -\boldsymbol{H}^{-1} \left(\boldsymbol{X}_{k+1} - \boldsymbol{X}_k \right) + \boldsymbol{w}_k, \tag{9}$$

where $\boldsymbol{w}_k$ is the linearization error between two equilibrium operation points and is assumed to be normally distributed with zero mean. By taking expectations on both sides of (9), the new algebraic state prediction model is derived as

$$\boldsymbol{x}_{k+1|k} = \widehat{\boldsymbol{x}}_{k|k} - \boldsymbol{H}^{-1}(\boldsymbol{X}_{k+1|k} - \boldsymbol{X}_{k|k}) = \widehat{\boldsymbol{x}}_{k|k} - \boldsymbol{H}^{-1}\boldsymbol{X}_k^{\circ}, \tag{10}$$

where $\widehat{\boldsymbol{x}}_{k|k}$ is the filtered algebraic state at time k; $\boldsymbol{X}_{k+1|k}$ is the forecasted vector of loads and DGs; $\boldsymbol{X}_{k|k}$ represents the actual power injections of loads and DGs at time k; $\boldsymbol{X}_k^{\circ} = \boldsymbol{X}_{k+1|k} - \boldsymbol{X}_{k|k}$. It is clear that in the derived state forecasting model, the change of system state vector is driven by the power injections from the loads and generators.

Comment: If there is a singularity problem of inverting the Jacobian matrix $\boldsymbol{H}$, the voltage stability issue might occur [27], [28]. Thus, some preventive actions should be taken to avoid voltage collapse. In the mean time, we can perform only a robust static SE [29] to determine the current system states using received measurements while ignoring the forecasted state vector.

B. Derivation of the Batch Regression Model

To develop robust estimator, a regression model is the premise. In [18], the regression model for linear KF has been derived. However, the power system SE problem is nonlinear. Thus, a batch regression model for nonlinear SE should be developed. In this paper, when the state prediction from former time instant is available, the first-order Taylor series expansion is applied to linearize the nonlinear measurement function, i.e.,

$$\boldsymbol{z}_{k+1} = \boldsymbol{g}(\boldsymbol{x}_{k+1}) + \boldsymbol{\zeta}_{k+1}, \tag{11}$$

at the forecasted state $\boldsymbol{x}_{k+1|k}$, resulting in

$$\boldsymbol{z}_{k+1} \doteq \boldsymbol{g}\left(\boldsymbol{x}_{k+1|k}\right) + \boldsymbol{G}\left(\boldsymbol{x}_{k+1} - \boldsymbol{x}_{k+1|k}\right) + \boldsymbol{\zeta}_{k+1}, \tag{12}$$

where $\boldsymbol{z}_{k+1}$ is the observed measurement vector; $\boldsymbol{g}(\cdot)$ is the nonlinear measurement function; $\boldsymbol{G} = \boldsymbol{\partial g}/\boldsymbol{\partial x}$ is the measurement Jacobin matrix evaluated at $\boldsymbol{x}_{k+1|k}$; $\boldsymbol{\zeta}_{k+1} \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{R}_{k+1}\right)$ is the measurement error vector with the error covariance matrix $\boldsymbol{R}_{k+1}$. Combining predictions with observations $\boldsymbol{z}_{k+1}$ yields the following batch regression form

$$\begin{bmatrix} \boldsymbol{z}_{k+1} - \boldsymbol{g}\left(\boldsymbol{x}_{k+1|k}\right) + \boldsymbol{G}\boldsymbol{x}_{k+1|k} \\ \boldsymbol{x}_{k+1|k} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{I} \end{bmatrix} \boldsymbol{x}_{k+1} + \begin{bmatrix} \boldsymbol{\zeta}_{k+1} \\ \boldsymbol{\delta}_{k+1} \end{bmatrix}, \tag{13}$$

where $\boldsymbol{I}$ is an identity matrix; $\boldsymbol{\delta}_{k+1}$ is the error between true state vector $\boldsymbol{x}_{k+1}$ and its predicted state vector $\boldsymbol{x}_{k+1|k}$, and its covariance matrix is

$$\boldsymbol{\Sigma}_{k+1|k} = \mathrm{cov}\left(\boldsymbol{\delta}_{k+1}\right) = \boldsymbol{\Sigma}_{k|k} + \boldsymbol{H}^{-1}\widehat{\boldsymbol{S}}_k\left(\boldsymbol{H}^{-1}\right)^T, \tag{14}$$

where $\boldsymbol{\Sigma}_{k|k}$ and $\widehat{\boldsymbol{S}}_k$ are the algebraic state filtering error covariance matrix and the prediction error covariance matrix of loads and DGs at time sample k, respectively. (13) can be put in the following compact form

$$\mathbf{z}_{k+1} = \boldsymbol{\mathcal{H}}_{k+1}\boldsymbol{x}_{k+1} + \mathbf{e}_{k+1}, \tag{15}$$

where the covariance matrix of the error $\mathbf{e}_{k+1}$ is

$$\boldsymbol{\mathcal{C}}_{k+1} = \mathbb{E}\left[\mathbf{e}_{k+1}\mathbf{e}_{k+1}^T\right] = \begin{bmatrix} \boldsymbol{R}_{k+1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{k+1|k} \end{bmatrix} = \boldsymbol{L}_{k+1}\boldsymbol{L}_{k+1}^T, \tag{16}$$

where $\boldsymbol{L}_{k+1}$ is obtained by Cholesky decomposition. Before the prewhitening, we need to detect and downweight the outliers as they are adding negative impacts. This can be carried out by using the projection statistics (PS) algorithm. For example, [18] proposed to apply PS to the matrix $\boldsymbol{Z}$ with the column vectors $\mathbf{z}_{k+1}$ and $\mathbf{z}_k$. However, since the presence of the term $\boldsymbol{G}$ in (13) would induce smearing effect once bad data occurs in the predicted state vector, we propose to modify the PS and apply it to the matrix $\boldsymbol{Z}$ as

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{z}_{k+1} - \boldsymbol{g}(\boldsymbol{x}_{k+1|k}) & \boldsymbol{z}_k - \boldsymbol{g}(\boldsymbol{x}_{k|k-1}) \\ \boldsymbol{x}_{k+1|k} & \boldsymbol{x}_{k|k-1} \end{bmatrix}, \tag{17}$$

where $\boldsymbol{z}_{k+1} - \boldsymbol{g}(\boldsymbol{x}_{k+1|k})$ is the innovation vector. Then, the i-th computed projection statistic value, PS_i, is compared to a given threshold. The flagged outliers are then downweighted using the following weight function: $\varpi_i = \min\left(1, \ d^2/PS_i^2\right)$, where $d = 1.5$ is set to yield good statistical efficiency.

Then, we perform prewhitening by multiplying $\boldsymbol{L}_{k+1}^{-1}$ into both sides of (15),

$$\boldsymbol{L}_{k+1}^{-1}\mathbf{z}_{k+1} = \boldsymbol{L}_{k+1}^{-1}\boldsymbol{\mathcal{H}}_{k+1}\boldsymbol{x}_{k+1} + \boldsymbol{L}_{k+1}^{-1}\mathbf{e}_{k+1}, \tag{18}$$

yielding the following final batch regression model

$$\boldsymbol{y}_{k+1} = \boldsymbol{A}_{k+1}\boldsymbol{x}_{k+1} + \boldsymbol{\eta}_{k+1}, \tag{19}$$

where $\mathbb{E}[\boldsymbol{\eta}_{k+1}\boldsymbol{\eta}_{k+1}^T] = \boldsymbol{I}$ can be easily verified.

Remark 2: The batch regression model in (19) is a standard linear squares regression problem with identity weight matrix. By using the weighted least squares estimation, the state vector can be estimated as

$$\widehat{\boldsymbol{x}}_{k+1|k+1} = \left(\boldsymbol{A}_{k+1}^T \boldsymbol{A}_{k+1}\right)^{-1} \boldsymbol{A}_{k+1}^T \boldsymbol{y}_{k+1}, \tag{20}$$

with the estimation error covariance matrix

$$\boldsymbol{\Sigma}_{k+1|k+1} = \left(\boldsymbol{A}_{k+1}^T \boldsymbol{A}_{k+1}\right)^{-1}. \tag{21}$$

Remark 3: With the forecasted state vector $\boldsymbol{x}_{k+1|k}$ and the nonlinear measurement model in (11), extended Kalman filter (EKF) technique is widely used in the existing forecasting-aided SE methods (see [4], [6], [7] for example) to perform the state filtering. It is interesting that equations shown in (20) and (21) are equivalent to the state filtering and state estimation error covariance matrix updating in EKF. This can be easily proved using the Matrix Inversion Lemma and simple algebraic operations.

Remark 4: The weighted least squares estimation and EKF are not robust to any type of outliers since their influence functions are not bounded [18], [30]. Therefore, robust estimator should be proposed to mitigate the impacts of outliers.

IV. Proposed Robust Estimator

To estimate $\boldsymbol{x}_{k+1}$ in a robust way, this paper proposes to use the GM estimator that combines the modified PS and the Huber convex score function. The objective function of the proposed estimator is expressed as

$$J(\boldsymbol{x}) = \sum_{i=1}^{m} \varpi_i^2 \rho(r_S), \tag{22}$$

where $\rho(\cdot)$ is the Huber function and can be defined as:

$$\rho(r_S) = \begin{cases} r_S^2/2 & \text{for } |r_S| \le \beta \\ \beta|r_S| - \beta^2/2 & \text{for } |r_S| > \beta \end{cases} \tag{23}$$

associated with its first derivative with respective to r_S, $\psi(r_S) = \partial\rho(r_S)/\partial r_S$. The parameter β is set to 1.5 with high efficiency at Gaussian noise [18]; r_S is the standardized residual, i.e., $r_S = r_i/(s\varpi_i)$, where $r_i = y_i - \boldsymbol{a}_i^T\widehat{\boldsymbol{x}}$ is derived from (19) and $\boldsymbol{a}_i^T$ is the ith column vector of the matrix $\boldsymbol{A}_{k+1}^T$; $s = 1.4826\, b_m\, median_i\,|r_i|$ is the robust scale estimate and tends asymptotically to σ when the observations follow $\mathcal{N}(\mu, \sigma^2)$. In this paper, since the covariance of $\boldsymbol{\eta}_{k+1}$ is an identity matrix, s tends asymptotically to 1; ϖ_i has been determined by applying PS on (17). Note that, in the proposed robust estimator, the ϖ_i is used to bound the influence of the measurement and structural errors, while the Huber function is chosen to bound the influence of the residuals.

A. Solving Proposed Robust Estimator Using IRLS Algorithm

To solve the objective function, we calculate the first order derivation with respect to x

$$\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}} = \sum_{i=1}^{m} -\frac{\varpi_i \boldsymbol{a_i}}{s}\psi(r_{S_i}) = \boldsymbol{0}. \tag{24}$$

It is clear that this is a system of nonlinear equations and iterative method is required to solve it. Here, the iteratively re-weighted least square (IRLS) algorithm [18] is adopted. Multiplying and dividing function $\psi(r_{S_i})$ by r_{S_i} yields

$$\sum_{i=1}^{m} \boldsymbol{a_i}\frac{\psi(r_{S_i})}{r_{S_i}} \cdot \frac{\varpi_i}{s} \cdot r_{S_i} = \boldsymbol{0}, \tag{25}$$

which can be arranged as a matrix form

$$(\boldsymbol{A}_{k+1})^T \boldsymbol{Q}\left(\boldsymbol{y} - \boldsymbol{A}_{k+1}\widehat{\boldsymbol{x}}\right) = \boldsymbol{0}, \tag{26}$$

where $q(r_S) = \psi(r_S)/r_S$ and $\boldsymbol{Q} = diag(q(r_S))$.

Then, the estimated states can be obtained through the following iterations

$$\widehat{\boldsymbol{x}}_{k+1|k+1}^{(l+1)} = \left(\boldsymbol{A}_{k+1}^T \boldsymbol{Q}^{(l)} \boldsymbol{A}_{k+1}\right)^{-1} \boldsymbol{A}_{k+1}^T \boldsymbol{Q}^{(l)} \boldsymbol{y}_{k+1}, \tag{27}$$

where l is the iteration counter. The algorithm converges if

$$\left\| \widehat{\boldsymbol{x}}_{k+1|k+1}^{(l+1)} - \widehat{\boldsymbol{x}}_{k+1|k+1}^{(l)} \right\|_\infty \le tol,\ e.g., 10^{-2}, \tag{28}$$

where $\|\cdot\|_\infty$ is the infinity norm.

B. Estimation Error Covariance Matrix Updating

After the convergence of the IRLS algorithm, the estimation error covariance matrix $\boldsymbol{\Sigma}_{k|k}$ should be replaced and updated by $\boldsymbol{\Sigma}_{k+1|k+1}$ so that the state forecasting (see equations (10) and (14)) in the next time sample, i.e., $k+2$, can be performed. According to Hampel's proposal, the asymptotic variance matrix $\boldsymbol{V}$ of a linear regression model ($\boldsymbol{Y} = \boldsymbol{X\theta} + \boldsymbol{\varepsilon}$ for example) using M-estimator can be derived by the Influence Function (IF) and is expressed as [30]

$$\begin{aligned} \boldsymbol{V} &= \lim_{m\to\infty} \boldsymbol{V}\left(\sqrt{m}\widehat{\boldsymbol{\theta}}_m\right) = \mathbb{E}\left[\boldsymbol{IF}\cdot\boldsymbol{IF}^T\right] \\ &= \frac{\mathbb{E}\left[\psi^2(r)\right]}{\left(\mathbb{E}\left[\psi'(r)\right]\right)^2}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1} \end{aligned}, \tag{29}$$

where r is the residual. However, in this paper, the GM-estimator is used, yielding new $\boldsymbol{IF}$. Following the work in [18], [30], the new $\boldsymbol{IF}$ can be derived as (due to the space limitation, the derivation process is not shown here.)

$$\boldsymbol{IF}(\boldsymbol{r}, \Phi) = \frac{\psi(r_S)}{\left(\mathbb{E}_\Phi\left[\psi'(r_s)\right]\right)^2}\left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{a}\varpi, \tag{30}$$

where Φ is the cumulative probability distribution function of the residual $\boldsymbol{r}$. Using (29), the asymptotic covariance matrix of $\boldsymbol{\Sigma}_{k+1|k+1}$ is

$$\begin{aligned} \boldsymbol{\Sigma}_{k+1|k+1} &= \mathbb{E}\left[\boldsymbol{IF}\cdot\boldsymbol{IF}^{\boldsymbol{T}}\right] \\ &= \frac{\mathbb{E}_\Phi\left[\psi^2(r_S)\right]}{\left\{\mathbb{E}_\Phi\left[\psi'(r_S)\right]\right\}^2}\left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\left(\boldsymbol{A}^T\boldsymbol{Q}_\omega\boldsymbol{A}\right)\left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1} \end{aligned}, \tag{31}$$

where $\boldsymbol{A}$ refers to $\boldsymbol{A}_{k+1}$; $\boldsymbol{Q}_\omega = diag\left(\varpi_i^2\right)$. In this paper, since the threshold β for the Huber estimator is 1.5, the value of the first term in the right hand side of equation (31) is 1.0369. Please see the Appendix for the detailed calculation.

C. Algorithm Implementation

The implementation of the proposed robust SE integrating temporal and spatial correlations of loads and DGs can be summarized as follows:

Step 1: Estimate the correlation matrix $\widehat{\varphi}_1$ using the historical load and DG data and perform the short-term power injections of loads and DGs forecasting by equation (5).

Step 2: Perform the state forecasting using equation (10) and calculate its associated error covariance matrix through equation (14).

Step 3: Adopt the projection statistic (PS) algorithm to identify and downweight outliers, followed by the robust prewhitening using (18).

Step 4: Construct the batch regression model leveraging equation (19)

Step 5: Use the IRLS algorithm to solve the proposed robust GM-PSE in an iterative way using equation (27).

Step 6: Increase the loop counter to $l \leftarrow l+1$ and judge if the IRLS algorithm has converged for a given tolerance threshold, otherwise return to Step 5.

Step 7: Calculate and store the estimated injections of the loads and DGs and update the estimation error covariance matrix $\mathbf{\Sigma}_{k|k}$ with $\mathbf{\Sigma}_{k+1|k+1}$.

Step 8: Turn to Step 1 if measurements for the next time sample, i.e., $k+2$, and the historical load and DGs data at $k+1$, are available.

V. Numerical Results

Numerical simulations are conducted on the IEEE benchmark test systems with DGs integration for validating the performance of GM-PSE. This section first describes the test systems. Then, the evaluation indices are introduced. Finally, numerical results are presented and analyzed.

A. Description of Test Systems

To test the performance and robustness of GM-PSE, simulations on IEEE 14, 30, 57 and 118 bus test systems are used, whose data are available in [31]. The scaled 10-min load and wind data from BPA [32] are used for simulations. This 10-min interval is filled with 30 samples, which is similar to the simulations conducted in [3]. The generator outputs are changed according to the assignment of the participation factors. To simulate the variations of the systems, the smooth load changes over a period of time are obtained by successfully running load flows under different loading conditions. The outcome of the power flow are regarded as true states and true values of measurements, which include bus voltages, bus injections and line flows. The measurements are generated by adding random additive Gaussian noises with zero mean and standard deviations of 1% (voltages) or 2% (powers). The measurement configuration for the four test systems are shown in Table I, where NoS, NoPJ, NoPF and NoV represents the numbers of states, power injections, power flow measurements, and voltage measurements, respectively. The maximal iterations for the proposed GM-PSE is 20; the threshold for the convergence is 10^{-2}. All the tests are performed in MATLAB environment using Intel Core i5 2.5Hz CPU with 8 GB memory computer.

TABLE I
MEASUREMENT CONFIGURATION FOR THE FOUR TEST SYSTEMS

Test system	NoS	NoPJ	NoPF	NoV	Redundancy
IEEE 14-bus	27	10	34	1	1.67
IEEE 30-bus	59	16	76	2	1.61
IEEE 57-bus	113	36	168	2	1.82
IEEE 118-bus	235	44	352	4	1.70

B. Evaluation Indices

In this study, three methods are used to make comparisons under different simulation conditions:

- **Method 1:** Traditional EKF-based SE method (TSE) with the state space model based state forecasting, where the state transition matrix is identified by using the Holt's linear exponential smoothing method [4].
- **Method 2:** EKF-based SE method considering the *temporal correlation* (SETC). The state forecasting is performed using the first-order auto-regressive (AR(1)) model and the state accuracy-based weighting function is used to improve the robustness of the estimator, which is our previous work in [12].
- **Method 3:** Proposed robust power system SE method considering both *temporal and spatial correlations* (GM-PSE). The VAR(1) model is used to capture the correlations of loads and DGs, and the GM-estimator is proposed to robustify the SE.

In order to provide more generalized simulation results and higher statistical significance, N_{MC}=100 Monte-Carlo simulations are run for all simulations. The performances of the three methods are evaluated by the following indices [11]:

Prediction indices: the mean-absolute-error (MAE) of forecasted voltage magnitude $\widetilde{V}_{MAE}$ and voltage angle $\widetilde{\theta}_{MAE}$,

$$\widetilde{V}_{MAE} = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \frac{1}{N_s} \sum_{i=1}^{N_s} \left| \widetilde{V}_i - V \right| \tag{32}$$

$$\widetilde{\theta}_{MAE} = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \frac{1}{N_s} \sum_{i=1}^{N_s} \left| \widetilde{\theta}_i - \theta \right| \tag{33}$$

Filtering indices: the mean-absolute-error (MAE) of filtered voltage magnitude $\widehat{V}_{MAE}$ and voltage angle $\widehat{\theta}_{MAE}$,

$$\widehat{V}_{MAE} = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \frac{1}{N_s} \sum_{i=1}^{N_s} \left| \widehat{V}_i - V \right| \tag{34}$$

$$\widehat{\theta}_{MAE} = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \frac{1}{N_s} \sum_{i=1}^{N_s} \left| \widehat{\theta}_i - \theta \right| \tag{35}$$

where V, $\widetilde{V}$, $\widehat{V}$ represent the true, predicted and filtered voltage magnitude value, respectively; while θ, $\widetilde{\theta}$, $\widehat{\theta}$ represent the true, predicted and filtered voltage angle value, respectively; N_s is the number of states (voltage magnitudes or voltage angles).

C. Results and Discussion

The effectiveness and robustness of GM-PSE is investigated from three aspects: normal operation condition, occurrence of outliers and the occurrence of sudden load change.

TABLE II
COMPARISON OF FORECASTING ERROR FOR VOLTAGE MAGNITUDES IN DIFFERENT TEST SYSTEMS

Method	14-bus	30-bus	57-bus	118-bus
TSE (in pu)	0.0084	0.0611	0.113	0.145
SETC (in pu)	0.0072	0.0574	0.0847	0.104
GM-PSE (in pu)	0.0036	0.0268	0.0365	0.0618

TABLE III
COMPARISON OF FORECASTING ERROR FOR VOLTAGE ANGLES IN DIFFERENT TEST SYSTEMS

Method	14-bus	30-bus	57-bus	118-bus
TSE (in degree)	0.00938	0.1045	0.1181	0.1217
SETC (in degree)	0.00821	0.0921	0.102	0.1104
GM-PSE (in degree)	0.00345	0.0424	0.0631	0.0703

TABLE IV
COMPARISON OF FILTERING ERROR FOR VOLTAGE MAGNITUDES IN DIFFERENT TEST SYSTEMS

Method	14-bus	30-bus	57-bus	118-bus
TSE (in pu)	6.5×10^{-4}	0.0084	0.0095	0.012
SETC (in pu)	5.4×10^{-4}	0.0063	0.0091	0.0104
GM-PSE (in pu)	4.2×10^{-4}	0.0047	0.0065	0.0072

TABLE V
COMPARISON OF FILTERING ERROR FOR VOLTAGE ANGLES IN DIFFERENT TEST SYSTEMS

Method	14-bus	30-bus	57-bus	118-bus
TSE (in degree)	0.0036	0.005	0.0061	0.012
SETC (in degree)	0.0026	0.0036	0.0041	0.0068
GM-PSE (in degree)	0.0023	0.0033	0.0039	0.0047

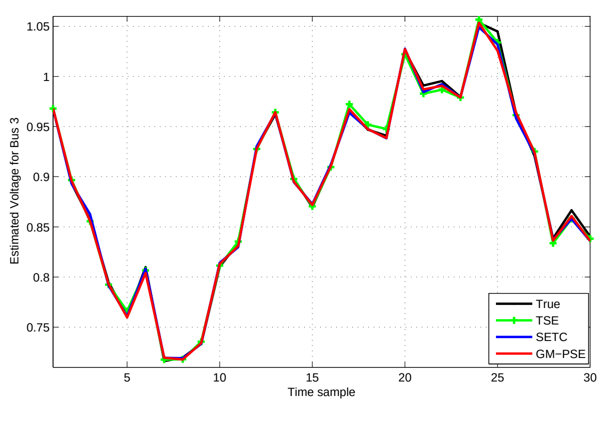

Fig. 1. Comparison results of estimating the voltage magnitude at bus 3 in the IEEE 118-bus system.

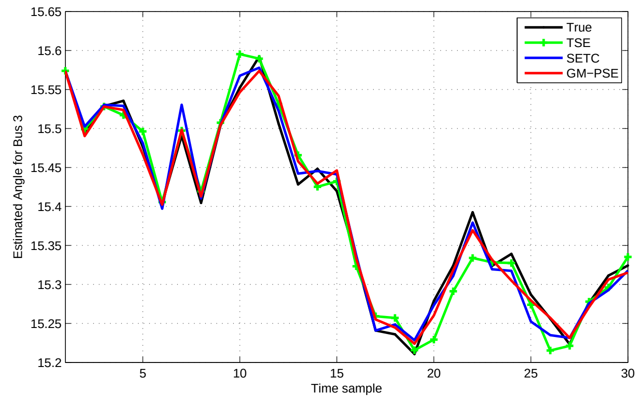

Fig. 2. Comparison results of estimating the voltage angle at bus 3 in the IEEE 118-bus system.

1) Normal Operation Condition: In this subsection, we consider the slowly changing loads, which are obtained by adding a random fluctuation with a linear trend (1%-3%) to the load curve, and the measurement set does not contain bad data. Tables II-V show the test results for all systems. As it can be seen from these tables, the GM-PSE that uses the VAR(1) model considering both temporal and spatial correlations is able to provide us more accurate forecasting results than the other two methods. The SETC, which only considers temporal correlation, performs a little better than the state space model based forecasting method. These tables also indicate that GM-PSE method obtains better filtering results than the other two methods due to the more accurate forecasting results and the robust filtering technique with high statistical efficiency. For example, Figs. 1 and 2 show the performance comparisons of the three approaches in terms of estimating the voltage magnitude and voltage angle at bus 3 in the IEEE 118-bus test system. It is clear that GM-PSE obtains the best results for tracking the trajectory of voltage changes, followed by the SETC method. Another observation from these results is that the estimation accuracy of the voltage magnitude of all three methods is quite similar while that of voltage magnitude is much different. This indicates the temporal and spatial correlations among voltage angles are more significant than those among voltage magnitude.

2) Occurrence of Outliers: To investigate the effects of outliers on the performance of three estimation methods, all measurements coming from bus 2 in the IEEE 30-bus system are contaminated with gross errors at time sample 25, i.e., all the measurements related to bus 2 have been increased to 1.8 times of their original values. The simulation results are presented in Figs. 3 and 4. It is obvious that the standard EKF based filtering technique, i.e., TSE method, is vulnerable to the effects of outliers. However, our previous robust SETC method and the proposed robust estimator, i.e., GM-PSE can bound the influence of outliers. Thanks to the accurate forecasting results and the high statistical efficiency of the GM-estimator, the GM-PSE performs better than the SETC method. The MAE for filtered voltage magnitude $\widehat{V}_{MAE}$ are 0.0034 and 0.0039 for GM-PSE and SETC, respectively. While the MAE for the filtered voltage angle $\widehat{\theta}_{MAE}$ of GM-PSE and SETC are

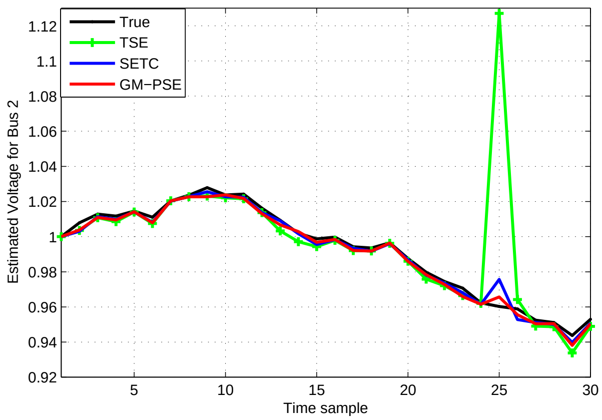

Fig. 3. Comparison results of estimating the voltage magnitude at bus 2 in the IEEE 30-bus system when outliers occur.

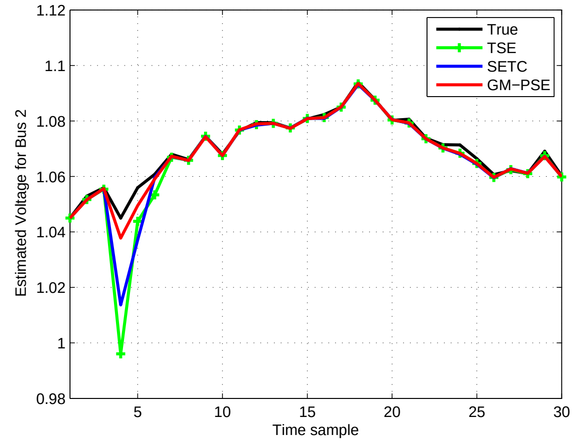

Fig. 5. Comparison results of estimating the voltage magnitude at bus 2 in the IEEE 14-bus system when sudden load change happens.

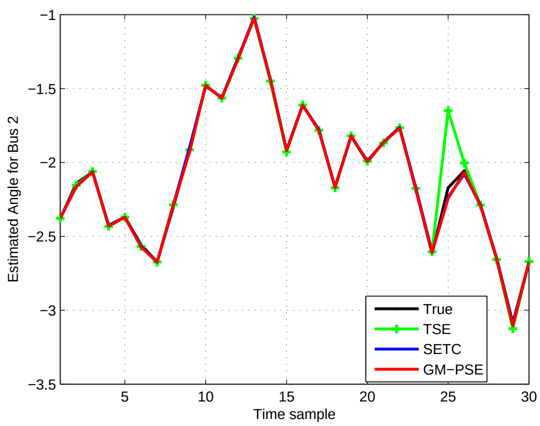

Fig. 4. Comparison results of estimating the voltage angle at bus 2 in the IEEE 30-bus system when outliers occur

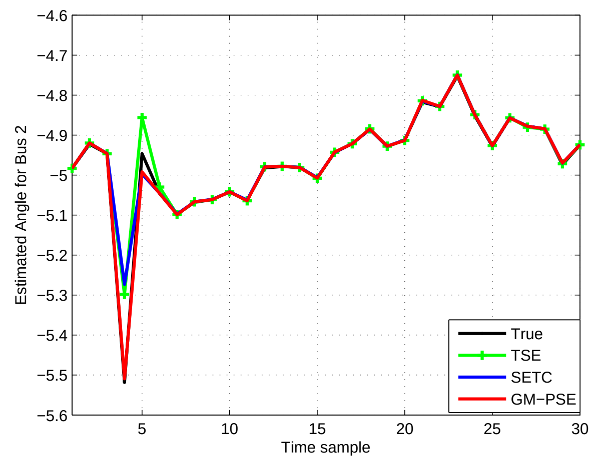

Fig. 6. Comparison results of estimating the voltage angle at bus 2 in the IEEE 14-bus system when sudden load change happens.

0.0048 and 0.0070, respectively. This indicates the efficiency improvement of estimating voltage magnitude and angle by using GM-PSE compared with SETC method.

To quantify how many outliers the proposed method can handle without giving unreliable estimation results, that is, the breakdown point of the proposed estimator, extensive simulations on the IEEE 30-test system are conducted. By replacing a varying number of data points by outliers in the observation vector z_{k+1}, it is observed that the GM-EKF can handle up to 25% of outliers among the data set no matter whether they are vertical or leverage outliers, the worst case being clustered ones. It is interesting to notice that this is consistent with the theoretical breakdown point of a GM-estimator in linear structured regression (e.g., in linearized power system state estimation model, which involves sparse Jacobian matrices) [33].

3) Sudden Load Change: To investigate the three methods in handling the system sudden changes, the estimation of states of bus 2 in the IEEE 14-bus test system is taken as an example, where the real power of the load at bus 2 is changed to 0.4 p.u. at the time sample k=4. The reason for change of test system from the IEEE 30 bus system to the IEEE 14 bus system is to test the scalability and adaptiveness of the three methods (Actually, we have also tested three methods in IEEE 30 bus system and other systems with sudden load change and the results are consistent with those in the IEEE 14 bus system here). Figs. 5 and 6 show the simulation results. Since the voltage magnitude and angle are closely related to the change of reactive and real power in the transmission system, respectively, the sudden change of real power at a load will cause larger voltage angle variations than that of the voltage magnitudes. This can be seen through Fig. 5, where the voltage at bus 2 changes more than that of voltage magnitude. It is obviously observed that with sudden load change, both TSE and SETC can not track the dynamic variations of the voltage magnitude and angle at bus 2. However, the GM-PSE can effectively track their dynamics by putting larger weights for observations while downweighting the unreliable predictions significantly. To be more specific, when unpredictable sudden change happens in power system, the predicted value will be far away from the actual system states. Unlike the proposed GM-PSE, the TSE and SETC can not obtain a good tradeoff between predictions and observations, resulting in unreliable estimation results.

VI. Conclusion and Future Work

The role of forecasting-aided state estimation is increasingly important for power system real-time modeling and control in modern energy management center. In this paper, a novel robust GM-estimator for power system forecasting-aided state estimation that integrates both temporal and spatial correlations of loads and DGs is proposed. The time series analysis technique, VAR(1), is used to model the temporal and spatial correlations. These correlations are further integrated into the state forecasting model through power flow analysis, resulting in more accurate state forecasting results. Finally, a GM-estimator is proposed to bound the influence of observation outliers while maintaining a high statistical efficiency. Simulation results on different IEEE test systems under various operating conditions demonstrate the efficiency and robustness of the proposed method.

Future work will apply this approach for large-scale realistic power systems with high penetration of renewable energy integration. Besides, the robust time series based forecasting technique, e.g., the Median-of-Ratios Estimator (MRE) [34] and the Phase-Phase Correlator (PPC) [35], which have high breakdown points, will be adopted and investigated. Finally, we are currently working on developing robust detector for system topology or parameter errors. It is well known that to handle system topology or parameter errors [36], [37]: (i) the topology or parameter errors detection problem needs to be reformulated, which might be different from the state estimation; (ii) a very high level of local measurement redundancy is required, which is difficult to achieve in practical power system due to the economic constraints; (iii) a robust detector with high breakdown point is required since the outliers caused by topology or parameter errors are strongly correlated. In our proposed method, because of the enhanced redundancy from forecasted measurements, and the good robustness of the projection statistics, measurements corrupted by the topology or parameter errors can be identified as outliers. However, how to correlate outliers with topology or parameter errors is still a challenging problem.

VII. Acknowledgements

The authors would like to thank Prof. Lamine Mili at Virginia Tech and Prof. Robson C. Pires at UNIFEI for valuable discussions and suggestions on robust GM-estimator and VAR(1) model for temporal and spatial correlations modeling.

Appendix

In this paper, since the covariance of $\boldsymbol{\eta}_{k+1}$ is an identity matrix, $\boldsymbol{\eta}_{k+1}$ has been standardized to the standard Gaussian distribution. Therefore, the probability distribution function of the residual $\boldsymbol{r}$ can be expressed as $\phi(r) = \frac{1}{\sqrt{2\pi}} e^{-\frac{r^2}{2}}$. On the other hand, from the Huber function with $\beta = 1.5$, we can calculate

$$\psi(r) = \begin{cases} r & \text{for } |r| \leq \beta \\ \beta sign(r) & \text{for } |r| > \beta \end{cases}, \tag{36}$$

$$\psi'(r) = \begin{cases} 1 & \text{for } |r| \leq \beta \\ 0 & \text{for } |r| > \beta \end{cases}. \tag{37}$$

Then, we can further obtain

$$\begin{aligned} \mathbb{E}\left[\psi'(r)\right] &= \int_{-\infty}^{\infty} \psi'(r)\phi(r)\,dr = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{r^2}{2}} dr \\ &= 2\Phi(\beta) - 1 = 0.8664, \end{aligned} \tag{38}$$

$$\begin{aligned} &\mathbb{E}\left[\psi^2(r)\right] = \int_{-\infty}^{\infty} \psi^2(r)\phi(r)\,dr \\ &= 2\frac{\beta^2}{\sqrt{2\pi}} \int_{-\infty}^{-\beta} e^{-\frac{r^2}{2}} dr + \frac{1}{\sqrt{2\pi}} \int_{-\beta}^{\beta} r^2 e^{-\frac{r^2}{2}} dr \\ &= \beta^2 \Phi(-\beta) - \frac{2\beta}{\sqrt{2\pi}} e^{-\frac{\beta^2}{2}} + 2\Phi(\beta) - 1 \\ &= 0.7784. \end{aligned} \tag{39}$$

Finally, we can calculate

$$\frac{\mathbb{E}\left[\psi^2(r)\right]}{\left(\mathbb{E}\left[\psi'(r)\right]\right)^2} = \frac{0.7784}{(0.8664)^2} = 1.0369. \tag{40}$$

References

[1] S. Deshmukh, N. Balasubramaniam, A. Pahwa, "State estimation and voltage/VAR control in distribution network with intermittent measurements," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 200–209, Jan. 2014.

[2] J. B. Zhao, G. X. Zhang, X. R. Wang, D. H. Zou, "Distribution state estimation with renewable sources based distributed generations using a modified quantum-inspired evolutionary algorithm," *J. Renewable Sustainable Energy* 5, 053133, Oct 2013.

[3] M. Hassanzadeh, C. Y. Evrenosoglu, L. Mili, "A short-term nodal voltage phasor forecasting method using temporal and spatial correlation," *IEEE Trans. Power Syst.,* DOI: 10.1109/TPWRS.2015.2487419, 2016.

[4] M. B. Do Coutto Filho, J. C. Stacchini de Souza, "Forecasting-aided state estimation–Part I: Panorama," *IEEE Trans. Power Syst.,* vol. 24, no. 4, pp. 1667–1678, 2009.

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

[6] M. Hassanzadeh, C. Y. Evrenosoglu, "Use of PMUs in regression-based power system dynamic state estimation," *in Proc. North American Power Symposium*. 2012, Champaign, IL, Sept. 9-11, 2012, pp. 1-5.

[7] M. Hassanzadeh, C. Y. Evrenosoglu, "A regression-based analysis based state transition model for power system dynamic state estimation," in *Proc. North American Power Symposium*. 2012, Champaign, IL, Sept. 9-11, 2012 , pp.1-5.

[8] Y. N. Sun, P. W. Du, Z. Y. Huang, K. Kalsi, R. S. Diao, K. K Anderson, Y. L. Li, "PMU placement for dynamic state tracking of power systems," in *Proc. North American Power Symposium*. 2011, Boston, MA, Aug. 4-6, 2011, pp.1-7.

[9] E. Farantatos, G.K. Stefopoulos, G. J. Cokkinides, A. P. Meliopoulos, "PMU-Based dynamic state estimation for electric power systems," in *Proc. IEEE Power Energy Society General Meeting*, 2009, Calgary, AB, Jul.26-30,2009, pp. 1-8.

[10] A. Jain, N. R. Shivakumar, "Impact of PMU in dynamic state estimation of power systems," in *Proc. North American Power Symposium*, 2008, Calgary, AB, Sep. 28-30, 2009, pp. 1-8.

[11] K. R. Shih, S. J. Huang, "Application of a robust algorithm for dynamic state estimation of a power system," *IEEE Trans. Power Syst*., vol. 17, no. 1, pp. 141–147, Feb. 2002.

[12] 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. pp. 26–30, 2015.

[13] 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, 2016.

[14] G. Valverde, A. T. Saric, V. Terzija, "Stochastic monitoring of distribution networks including correlated input variables," *IEEE Trans. Power Syst*., vol. 28, no. 1, pp. 246–255, Feb. 2013.

[15] E. Blood, B. Krogh, M. Ilic, "Electric power system static state estimation through Kalman filtering and load forecasting," in *Proc. IEEE Power Eng. Soc. General Meeting*, July. pp. 1–6, 2008.

[16] J. Dowell, P. Pinson, "Very-short-term probabilistic wind power forecasts by sparse vector autoregression," *IEEE Trans. Smart Grid.*, DOI: 10.1109/TSG.2015.2424078, 2015.

[17] 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.

[18] M. A. Gandhi, L. Mili, "Robust Kalman filter based on a generalized maximum-likelihood-type estimator," *IEEE Trans. Signal Process*., vol. 58, no. 5, pp. 2509–2520, May. 2010.
[19] A. Alimardani, F. Therrien, D. Atanackovic, J. Jatskevich, E. Vaahedi, "Distribution system state estimation based on nonsynchronized smart meters," *IEEE Trans. Smart Grid.*, Vol. 6. no. 6, pp. 2919–2928, 2015.
[20] P. Bacher, H. Madsen, H. A. Nielsen, "Online short-term solar power forecasting," *Solar Energy*, vol. 83, no. 10, pp. 1772–1783, 2009.
[21] P. Ailliot, V. Monbet, M. Prevosto, *An autoregressive model with time-varying coefficients for wind fields*. Hoboken, NJ, USA: Wiley, 2005.
[22] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, 2009.
[23] D. Hill, D. McMillan, K. R. W. Bell, D. Infield, "Application of auto-regressive models to UK wind speed data for power system impact studies," *IEEE Trans. Sustainable Energy*, vol. 3, no. 1, pp. 134–141, 2012.
[24] 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.
[25] J. D. Hamilton, *Time Series Analysis*. Princton, NJ: Princton University Press, 1994.
[26] P. J. Brockwell, R. A. Davis, *Introduction to Time Series and Forecasting*. Springer Texts in Statistics, 2nd. New York, NY, USA: Springer-Verlag, 2002.
[27] A. C. Z. de Souza, S. C. S. de Souza, A. M. L. da Silva, "On-line voltage stability monitoring," *IEEE Trans. Power Syst.*, vol. 15, no. 4, pp. 1300–1305, 2000.
[28] P. Kundur, J. Paserba, V. Ajjarapu V, et al, "Definition and classification of power system stability IEEE CIGRE joint task force on stability terms and definitions," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, 2004.
[29] L. Mili, M. Cheniae, N. Vichare, P. Rousseeuw, "Robust state estimation based on projection statistics," *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 1118–1127, May 1996.
[30] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*. New York: John Wiley & Sons, Inc., 1986.
[31] http://www.ee.washington.edu/research/pstca/
[32] Bonneville Power Administration, *Wind Generation & Total Load in the BPA Balancing Authority*, 2012 http://transmission.bpa.gov/business /operations/wind.
[33] L. Mili, C. W. Coakley, "Robust estimation in structured linear regression," *Annals of Statistics*, vol. 24, no. 6, p. 25932607, 1996.
[34] Y. Chakhchoukh, P. Panciatici, L. Mili, "Electric load forecasting based on statistical robust methods," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 982–991, 2011.
[35] P. Tamburello, L. Mili, "New robust estimators of correlation and weighted basis pursuit," *IEEE Trans. Signal Processing*, Vol. 63, no. 4, pp. 882–894, 2015.
[36] L. Mili, G. Steeno, F. Dobraca, D. French, "A robust estimation method for topology error identification," *IEEE Trans. Power Syst.*, vol. 14, no. 4, pp. 1469–1476, 1999.
[37] A. G. Exposito, A. V. J. de la A, "Reduced substation models for generalized state estimation," *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 839–846, 2000.

Junbo Zhao (S'13) has been pursuing the Ph.D. degree at the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University since 2015.

His research interests are in the theoretical and algorithmic studies in power system state estimation, power system operation and cyber security, robust statistics and signal processing.

Gexiang Zhang (M'03) received B.Sc. and M. Sc. degrees in 1997 and 2002 from Southwest University of Science and Technology, Mianyang, China, and 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).

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.

Zhao Yang Dong (M'99-SM'06) received the Ph.D. degree from the University of Sydney, Sydney, N.S.W., Australia, in 1999. He is currently a Professor and Head of School of Electrical and Information Engineering, University of Sydney, Australia. He was previously the Ausgrid Chair and Director of the Center for Intelligent Electricity Networks (CIEN), The University of Newcastle, Callaghan, N.S.W., Australia, and is now a Conjoint Professor there. He also held academic and industrial positions with the Hong Kong Polytechnic University, Hong Kong and Transend Networks (now TASNetworks), Tas., Australia. His research interests include smart grid, power system planning, power system security, load modeling, electricity market, and computational intelligence and its application in power engineering.

He is an Editor of the IEEE TRANSACTIONS ON SMART GRID, the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, the IEEE POWER ENGINEERING LETTERS, and IET Renewable Power Generation.

Massimo La Scala (M'88–SM'99–F'07) was born in Bari, Italy, in 1959. He received the B.S. and Ph.D. degrees in electrical engineering from the University of Bari, Bari, in 1984 and 1989, respectively.

In 1987, he joined ENEL. He is currently Professor of Power Systems at the Politecnico di Bari. His research interests are in the areas of power system analysis and control. In 2007, he was appointed Fellow Member of the IEEE PES for his contributions to computationally efficient power system dynamic performance simulation and control.