MMSE-based analytical estimator for uncertain power system with limited number of measurements

Hasnae Bilil, Member, IEEE, and Hamid Gharavi, Life Fellow, IEEE,
Advanced Network Technologies Division, National Institute of Standards and Technology, MD, USA

Abstract—The expected penetration of a large number of renewable distributed energy resources (DER’s) is driving next generation power systems toward uncertainties that can have a huge impact on the reliability and complexities of state estimation. Therefore, the stochastic power flow (SPF) and forecasting-aided state estimation of power systems integrating DER’s are becoming a major challenge for operation of the future grid. In this paper we propose a new state estimation method referred to as ‘mean squared estimator’ (MSE) to deal with the uncertain nature of the power system parameters. The estimator aims at achieving minimum mean squared error (MMSE) and benefits from the prior study of SPF, which involves the probability density functions (PDF’s) of the system parameters. The main advantage of this estimator is based on its ability to instantaneously incorporate the dynamics of the power system. Moreover, the analytical formula of MSE expresses the mean value of the estimated parameters corrected by an additional term that takes into account the measurement of the parameters. It is shown that the proposed MSE can provide an accurate state estimation with a limited number of measurements where WLS and UKF may lead to redundancies. The results have been compared to methods such as weighted least square (WLS), unscented Kalman filter (UKF) and compressive sensing-based UKF (CS-UKF). The numerical results show superior performances, especially under a limited number of measurements where WLS and UKF may lead to divergence.

Index Terms—Dynamic state estimation; minimum mean-squared error; Gaussian mixture model; analytic estimator; limited number of measurements.

I. INTRODUCTION

The deployment of renewable resources in distributed grid systems poses a set of new challenges mainly due to their variability and dependency on climate parameters, which can have a major impact on the system parameters that are needed to measure power flow and state estimation. The first aims at calculating the entire system parameters, based on prior knowledge (or prediction) of some of the parameters, while the latter, which is the backbone of energy management systems, estimates these parameters measured under noisy conditions. More precisely, a state estimator aims at providing estimated values of the system parameters roughly equal to the true values by eliminating measurement errors. In the presence of renewables, many studies in the past have incorporated the system uncertainties for power flow measurement and state estimation. Stochastic power flow (SPF) and forecast-aided state estimation, also called dynamic state estimation (DSE), take into consideration the uncertain behavior of power generations and loads. The main objective is to determine the probability density function (PDF) of the power system parameters to solve the SPF problem [1]–[13]. For instance, in [1], the authors propose to combine the concept of Cumulants and Gram-Charlier’s expansion theory, and [2] presents a comparative and analytic study of four different Hong’s point estimate schemes to solve the SPF problem. A practical method to tackle various random variables, such as renewable energy sources parameters that follow different types of probability distributions, has also been proposed in [4]. In the case of unknown input PDF’s, a method for non-parametric probabilistic load flow analysis is developed in [5]. Other studies mainly use a Gaussian mixture model (GMM) to approximate the PDF of loads or a mix of DER’s, which cannot be approximated with a known shape PDF [6]–[8]. Furthermore, there has been tremendous effort towards real-time DSE in support of next generation power systems [14]–[23]. In particular, with the emergence of the phasor measurement units (PMU’s) having the unique capability of providing synchronous measurements, real-time DSE has been receiving considerable attention [24]–[36]. Moreover, various versions of Kalman filtering have also been developed in order to improve DSE performances and its robustness while reducing the execution time [33], [34]. For instance, three versions of Kalman filtering, namely, extended Kalman filter (EKF), the unscented Kalman filter (UKF), and the Cubature Kalman filter (CKF) with or without PMU’s measurements have been evaluated in [35]. By utilizing the PMU’s measurements, the authors in [36] have developed an extended Kalman filter with unknown inputs (EKF-UI) that can estimate the states as well as the unknown inputs of the synchronous machine. On the other hand, the study in [22] describes a generalized maximum likelihood (GM)-estimator on power systems (GM-PSE). In this method, short-term forecasts of the distributed energy resources (DER’s) and loads are first calculated to predict the states. A redundant batch regression model is then used to process the predicted states and measured parameters. The method uses an iteratively re-weighted least squares (IRLS) algorithm to obtain the final estimated states. To track the system transients faster and with better reliability, the same authors present an iterated extended Kalman filter using the generalized maximum likelihood approach (termed GM-IEKF) [23]. We should point out, however, that all aforementioned methods are designed to predict and then estimate the states regardless of any prior knowledge of the states PDF’s that the SPF is supposed to determine.

U.S. Government work not protected by U.S. copyright.
In this paper, we present a new analytic-based state estimator, referred to as ‘Mean Square Estimator’ (MSE), that uses the outputs of the SPF by assuming a prior knowledge of the state parameters distribution regardless of their shapes. Specifically, our approach is based on the calculus of the minimum mean squared error (MMSE) that deals with generally non-Gaussian Random Variables (NGRV’s). Therefore, the GMM, which is an approximate presentation of the PDF’s of NGRV’s, is utilized in the proposed estimator. As a weighted sum of several Gaussian components, GMM is widely used for NGRVs probability distribution [6]–[8], [37], [38]. More specifically, in many studies, GMM has been suggested to model the PDFs of load power and renewable energy sources [39]–[43]. Since the proposed MMSE-based estimator is an analytical approach, it can provide estimated values instantaneously. In addition, its performance remains good even with a limited number of measurements, which is important in the case of a wide system. For instance, an extension of state estimation to medium and low voltage grids may require a large number of measurements that can have a considerable impact on state estimation complexities. In terms of performance, the proposed estimator is compared with the weighted least squares method (WLS) and UKF using various IEEE test systems.

The paper is organized as follows: Section II presents the fundamentals of GMM presentation of NGRV’s. Section III introduces the formulation of the proposed MSE as a conditional expectation of the estimated states for the given measurements. Section IV derives formulas of generating moments followed by the general formula of MSE when states and measurements are NGRV’s. The results of the proposed state estimation approach using the IEEE 14, 30 and 118 bus models are then presented in Section V. Finally, Section VII provides the conclusion and some perspectives.

II. GAUSSIAN MIXTURE MODEL

In this section we give a brief overview of the fundamentals of GMM for the presentation of any NGRV as a function of Gaussian components. Gaussian distribution, which is commonly used for modeling univariate data, can be extended to two or more variables. There are many studies in open literature that provide the formulation of any order of the moments of Gaussian random variables (GRVs) [41], [42]. Furthermore, Gaussian distribution can be considered to formulate the moments of NGRVs. Note that in the rest of the paper, the notations $\mathbb{E}(X)$, $X^T$ and $X^{-1}$ are used to express the expectation, the transposed and inverse matrices of $X$ respectively.

A. Gaussian Density

A multivariate random variable $\mathbf{x} = [x_1, x_2, \ldots, x_n]$ is said to be distributed as the multivariate Gaussian distribution with mean $\mu_\mathbf{x}$ and covariance matrix $\Sigma_\mathbf{x}$, $\mathbf{x} \sim \mathcal{N}(\mu_\mathbf{x}, \Sigma_\mathbf{x})$, if the density function of $\mathbf{x}$ is given by

$$g(\mathbf{x}|\mu_\mathbf{x}, \Sigma_\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_\mathbf{x}|}} \exp \left[-\frac{1}{2} ((\mathbf{x} - \mu_\mathbf{x})^T \Sigma_\mathbf{x}^{-1} (\mathbf{x} - \mu_\mathbf{x})) \right]$$

(1)

B. Gaussian Mixture Model

The Gaussian Mixture Model (GMM) is an approximate presentation of a non-Gaussian density function of a random variable $\mathbf{x}$ (it can be a multivariate variable) by a mixture (a summation) of $L_\mathbf{x}$ Gaussian distribution components [6]–[8], [37], [38]. It can be expressed as,

$$f_\mathbf{x}(\mathbf{x}) = \sum_{i=1}^{L_\mathbf{x}} w_i g(\mathbf{x}|\mu_{\mathbf{x}|i}, \Sigma_{\mathbf{x}|i}),$$

(2)

where $\mathbf{x}$ is an $n$-dimensional vector of continuous values, the vector of the estimation problem parameters, $w_i$, $i = 1, \ldots, L_\mathbf{x}$, are the mixture weights, and $g(\mathbf{x}|\mu_{\mathbf{x}|i}, \Sigma_{\mathbf{x}|i})$, $i = 1, \ldots, L_\mathbf{x}$, are the components Gaussian densities. Each component is an $n$-variate Gaussian function as defined precisely, with $\mu_{\mathbf{x}|i}$ and $\Sigma_{\mathbf{x}|i}$ being the mean vector and the covariance matrix of $\mathbf{x}$, respectively. The mixture weights satisfy the following equality,

$$\sum_{i=1}^{L_\mathbf{x}} w_{\mathbf{x}|i} = 1,$$

(3)

Therefore, the mean value vector and covariance matrix of $\mathbf{x}$ can be approximated in terms of the Gaussian components parameters as follows,

$$\mu_{\mathbf{x}} = \sum_{i=1}^{L_\mathbf{x}} w_{\mathbf{x}|i} \mu_{\mathbf{x}|i},$$

(4)

$$\Sigma_{\mathbf{x}} = \sum_{i=1}^{L_\mathbf{x}} w_{\mathbf{x}|i} \left( \Sigma_{\mathbf{x}|i} + (\mu_{\mathbf{x}|i} - \mu_{\mathbf{x}})(\mu_{\mathbf{x}|i} - \mu_{\mathbf{x}})^T \right),$$

(5)

III. MINIMUM MEAN SQUARE ERROR BASED POWER SYSTEM STATE ESTIMATION

A. Problem formulation

State estimation aims at determining an optimal estimation of a vectors’ parameters $\mathbf{x} = [x_1, \ldots, x_N]$ giving a vector of measurements $\mathbf{z} = [z_1, \ldots, z_M]$ and functions $\mathbf{h} = [h_1, \ldots, h_M]$ based on the following relationship,

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{v}$$

(6)

where $\mathbf{v}$ represents the noise. In the case of power system state estimation, the estimated parameters are voltage amplitudes and angles, e.g. $\mathbf{x} = [\theta, V]$. Moreover, using PMU’s or Supervisory Control And Data Acquisition (SCADA) measurements, the function $h_j(\mathbf{x})$ is either a linear function of voltage amplitude and angle measurements or a non-linear function of active and reactive power measurements. Notations $\mathcal{Z}_V, \mathcal{Z}_\theta, \mathcal{Z}_P, \mathcal{Z}_Q$ stand for the sets of buses equipped with
measurements meters of voltage amplitude and angles, active and reactive power injections respectively. \( Z_{P_j} \) and \( Z_{Q_j} \) present the sets of pairs of buses with measured active and reactive power flow respectively. Therefore,

\[
\begin{align*}
\tilde{h}_j(x) &= V_n 	ext{ for } z_j = z_{V_n}, \ n \in Z_V \\
\tilde{h}_j(x) &= \theta_n 	ext{ for } z_j = \theta_{n_s}, \ n \in Z_\theta \\
\tilde{h}_j(x) &= P_n 	ext{ for } z_j = z_{P_n}, \ n \in Z_P \\
\tilde{h}_j(x) &= Q_n 	ext{ for } z_j = z_{Q_n}, \ n \in Z_Q \\
\tilde{h}_j(x) &= P_{nk} \text{ for } z_j = z_{P_{nk}}, (n,k) \in Z_{P_j} \\
\tilde{h}_j(x) &= Q_{nk} \text{ for } z_j = z_{Q_{nk}}, (n,k) \in Z_{Q_j} \\
\end{align*}
\]  

(7)

In the above, \( z_{V_n}, \theta_{n_s} \) are the measurement values of voltage amplitude and angle, respectively at bus \( n \) and \( z_{P_n}, z_{Q_n} \) are the measurements of active and reactive power: \( P_n, Q_n, \) defined in \([8, 9]\) respectively, and \( z_{P_{nk}}, z_{Q_{nk}} \) correspond to the measurements of active and reactive power flow \( P_{nk}, Q_{nk} \) between the buses \( n \) and \( k, \) as defined in \([10, 11]\) respectively.\([45]\)

\[
P_n = V_n \sum_{k \in \mathcal{N}} V_k (G_{nk} \cos \theta_{nk} + B_{nk} \sin \theta_{nk}) \tag{8}
\]

\[
Q_n = V_n \sum_{k \in \mathcal{N}} V_k (G_{nk} \sin \theta_{nk} - B_{nk} \cos \theta_{nk}) \tag{9}
\]

\[
P_{nk} = -V_n^2 G_{nk} + V_n V_k (G_{nk} \cos \theta_{nk} + B_{nk} \sin \theta_{nk}) \tag{10}
\]

\[
Q_{nk} = -V_n^2 B_{nk} - V_n V_k (G_{nk} \sin \theta_{nk} - B_{nk} \cos \theta_{nk}) \tag{11}
\]

where the notation \( V_n \) stands for the voltage magnitude at the \( n^{th} \) bus, \( \theta_{nk} = \theta_n - \theta_k \) is the angular difference of voltage phases at the buses \( n \) and \( k, \) \( G_{nk} \) and \( B_{nk} \) are the real and imaginary part of admittance elements. \( \mathcal{N} \) is the set of the network buses.

**B. Minimum mean-squared error**

The objective function of an MMSE-based estimator is defined as follows \([44]\),

\[
\min_{\hat{x}_{MS}} \mathbb{E}(\tilde{h}_{MS}^2) = \mathbb{E}(h_{MS}^2) \tag{12}
\]

where \( \tilde{h}_{MS} = h - \hat{x}_{MS} \) is the estimation error of \( h \) and \( \hat{x}_{MS} \) is the estimated value of \( x \) using MSE. \( \tilde{x}_{MS}^T \) is the transposed vector of \( \tilde{x}_{MS}. \) The solution to this problem, as proven in \([44]\), is the conditional expectation of \( x \) knowing the measurements vector \( z, \)

\[
\hat{x}_{MS} = \mathbb{E}(x|z), \tag{13}
\]

This result is called the fundamental theorem of estimation theory and is true for both linear and non-linear systems \([44]\). Moreover, the minimum value of \( J(\tilde{x}_{MS}) \), which also presents the variance of the estimation error, is given as,

\[
J^*(\tilde{x}_{MS}) = \mathbb{E}((x^T x)|z) - \mathbb{E}(x^T |z) \mathbb{E}(x|z). \tag{14}
\]

where \( x^T \) is the transposed vector of \( x. \) This expression is the summation of the conditional variances of the elements of \( x \) knowing \( z. \)

**C. Mean squared estimator for Gaussian input parameters**

According to \([44]\), when \( x \) and \( z \) are jointly Gaussian, the estimator that minimizes the mean-squared error is

\[
\hat{x}_{MS} = \mu_x + \Sigma_{xz} \Sigma_z^{-1} [z - \mu_z]. \tag{15}
\]

Furthermore, the above estimator for \( x \) is NGRV’s can be expressed in two forms depending on \( z \) is Gaussian or non-Gaussian. Nonetheless, the estimator given in \([15]\) provides an optimal state estimation which would require determining the value of the covariance matrix \( \Sigma_{xz} \) of the estimated parameters and the measurements, as well as the covariance matrix \( \Sigma_z \) of the measurements. Elements \( \Sigma_{xz}, \Sigma_z \) for the elements \((i, j) \in \mathcal{N}_e \times \mathcal{N}_m \) \((\mathcal{N}_e, \mathcal{N}_m \) are the sets of estimated parameters and measurements respectively) in the matrix \( \Sigma_{xz} \) can be calculated using the following expression of two RV’s covariance,

\[
\Sigma_{xz} = \mathbb{E}(x_i z_j) - \mu_x \mu_{z_j}, \tag{16}
\]

From \([6]\),

\[
z_j = h_j(x) + v_j. \tag{17}
\]

Therefore,

\[
\mathbb{E}(x_i z_j) = \mathbb{E}(x_i h_j(x)) + \mathbb{E}(x_i v_j), \tag{18}
\]

Consider \( x_i \) and \( v_j \) uncorrelated, then \( \mathbb{E}(x_i v_j) = \mathbb{E}(x_i) \mathbb{E}(v_j) \). Assuming that the noise has zero mean \( \mathbb{E}(v_j) = 0 \), the following expression can be obtained,

\[
\mathbb{E}(x_i z_j) = \mathbb{E}(x_i h_j(x)). \tag{19}
\]

This clearly indicates that computation of matrix \( \Sigma_{xz} \) would require calculating the expectation values \( \mathbb{E}(x_i h_j(x)) \). This matrix can be expressed in terms of sub-matrices that correspond to covariances between the estimated state parameters and the measurements as given in \([7]\). Therefore, \( \Sigma_{xz} \) can be expressed as,

\[
\Sigma_{xz} = \begin{bmatrix}
\Sigma_{\theta, \theta} & \Sigma_{\theta, V} & \Sigma_{\theta, P} & \Sigma_{\theta, Q} \\
\Sigma_{V, \theta} & \Sigma_{V, V} & \Sigma_{V, P} & \Sigma_{V, Q} \\
\Sigma_{P, \theta} & \Sigma_{P, V} & \Sigma_{P, P} & \Sigma_{P, Q} \\
\Sigma_{Q, \theta} & \Sigma_{Q, V} & \Sigma_{Q, P} & \Sigma_{Q, Q}
\end{bmatrix} \tag{19}
\]

where \( \Sigma_{\theta, \theta}, \Sigma_{\theta, V}, \Sigma_{\theta, P}, \Sigma_{\theta, Q}, \Sigma_{P, \theta}, \Sigma_{P, V}, \Sigma_{P, P}, \Sigma_{P, Q}, \Sigma_{Q, \theta}, \Sigma_{Q, V}, \Sigma_{Q, P}, \Sigma_{Q, Q} \) are the covariance matrices between the voltage angle (estimated parameter) and the measurement parameters, i.e. voltage angle and amplitude, power injections and power flows, respectively. While \( \Sigma_{V, \theta}, \Sigma_{V, V}, \Sigma_{V, P}, \Sigma_{V, Q}, \Sigma_{P, \theta}, \Sigma_{P, V}, \Sigma_{P, P}, \Sigma_{P, Q} \) hold the covariance matrices between the voltage amplitude (estimated parameter) and the measurement parameters, i.e. voltage angle and amplitude, power injections and power flows, respectively.

The calculus of the mean values \( \mathbb{E}(x_i z_j) \) can be achieved using the PDF’s of voltage and measurement parameters, given by the SPF study. In the same way, the elements of the measurements covariance matrix \( \Sigma_z \) can be expressed as the difference between the mean value of the product of two measurements and the product of their mean values.
IV. GENERAL FORMULA OF MEAN SQUARED ESTIMATOR

Monitoring and state estimation of power system parameters requires the treatment of several NGRV’s. For theoretical reasons, assume a Gaussian distribution of uncertain variables can be most conveniently used for state estimation [44], [46]. The purpose behind this section is to present the general formula of MSE by developing the calculus of $E(x|z)$ in the case where states and measurements are NGRV’s. To achieve this, we develop the calculus of the moments of a generic set of NGRV’s, which are based on the GMM approximation of these variables.

A. On the moments of correlated non-Gaussian random variables

As mentioned before, the MMSE-based estimator is formulated in terms of the moments of voltage parameters, e.g., amplitudes and angles, and measurements. The PDF’s of these parameters are not generally Gaussian. Therefore, based on the GMM approximation of these parameters, an analytical solution of the moments of NGRV’s is developed in the following using formulas of the moments of GRV’s.

Proposition 1 (Moments of NGRV): Let $x$ be a RV and its GMM is as given in (2). Therefore,

$$E(x_1 x_2 \ldots x_n) = \sum_{i=1}^{L_x} \omega_{x|z} E(x_1 x_2 \ldots x_n | \mu_{x|i}, \Sigma_{x|i}),$$

(20)

where $E(x_1 x_2 \ldots x_n | \mu_{x|i}, \Sigma_{x|i})$ is the moments of $x_1, x_2, \ldots, x_n$ based on the $i^{th}$ component of $x$’s PDF.

Proof: Consider the characteristic function of $x$ as,

$$\varphi_x(u) = \int_{\mathbb{R}^n} \exp\{j(u, x)\} f_x(x) dx$$

$$= \sum_{i=1}^{L_x} \omega_{x|i} \int_{\mathbb{R}^n} \exp\{j(u, x)\} g(x | \mu_{x|i}, \Sigma_{x|i}) dx$$

$$= \sum_{i=1}^{L_x} \omega_{x|i} \varphi_{x|i}(u)$$

(21)

where $\varphi_{x|i}$ is the characteristic function of a jointly Gaussian random variable. $j$ is the complex number defined as $j^2 = -1$ and $(u, x)$ denotes the scalar product of $u$ and $x$, i.e., $(u, x) = u^T x$. Given that the moments of $n$ random variables can be obtained by differentiating the characteristic function,

$$E(x_1 x_2 \ldots x_n) = \frac{1}{n!} \left. \frac{\partial^n}{\partial u_1 \ldots \partial u_n} \varphi_x(u) \right|_{u=0}$$

(22)

Using the expression in (21), the moment of $n$ NGRV’s can be obtained as,

$$E(x_1 x_2 \ldots x_n) = \sum_{i=1}^{L_x} \omega_{x|i} \frac{1}{n!} \left. \frac{\partial^n}{\partial u_1 \ldots \partial u_n} \varphi_{x|i}(u) \right|_{u=0}$$

(23)

Therefore, the expression of the moment of NGRV’s can be given in terms of the moments of the Gaussian components in GMM of these RV’s, as expressed in (20).

B. General formula of MSE

The MMSE-based estimator presented in (15) is optimal when the states are GRV. Generally the voltage amplitude and angle are non-Gaussian. A new formula of the MSE estimator is presented in the following proposition for an NGRV for $x$, when GMM components of $x$ and $z$ are jointly Gaussian.

Proposition 2 (Conditional expectation of NGRV): Let $x$ be an NGRV whose GMM is given in (2) and $z$ being a GRV. If each of the GMM components of $x$ and $z$ are jointly then,

$$E(x|z) = \mu_x + \sum_{i=1}^{L_x} \omega_{x|i} \Sigma_{x|i,z}^{-1} [z - \mu_z].$$

(24)

with $\Sigma_{x|i,z}$ is the covariance matrix of $x$ (i.e., the $i^{th}$ components of its GMM) and $z$.

Proof: Applying Proposition 1 in the case of the first moment,

$$E(x|z) = \sum_{i=1}^{L_x} \omega_{x|i} E(x | z, \mu_{x|i}, \Sigma_{x|i}),$$

(25)

where $E(x | z, \mu_{x|i}, \Sigma_{x|i})$ is the expectation of $x$ following the $i^{th}$ Gaussian component of GMM of $x$’s PDF, conditioned by $z$. Thus, we can show,

$$E(x | z, \mu_{x|i}, \Sigma_{x|i}) = \mu_{x|i} + \Sigma_{x|i,z}^{-1} [z - \mu_z].$$

(26)

Giving that,

$$\sum_{i=1}^{L_x} \omega_{x|i} | \mu_{x|i} = \mu_x$$

This indicates that $E(x|z)$ can indeed be expressed as in (24). However, we should emphasize that proposition 2 provides the formula of conditional expectation when the general expected parameter (NGRV) is conditioned by a GRV. However, measurements in a power system, i.e., voltage parameters (amplitude and angle), power injection, and power flow can also be NGRV. Later, we present the calculus of the conditional expectation of NGRV conditioned by NGRV.

Proposition 3 (Expectation of NGRV conditioned by NGRV): Let $x$ and $z$ be NGRV’s whose PDF’s are approximated by GMM according to (2) and (27), respectively,

$$f_z(z) = \sum_{j=1}^{L_z} \omega_{z|j} g(z | \mu_{z|j}, \Sigma_{z|j}),$$

(27)

then

$$E(x|z) = \mu_x + \sum_{j=1}^{L_x} \sum_{j=1}^{L_z} \omega_{x|i} \omega_{z|j} A_{x|i,z|j} [z - \mu_z].$$

(28)

with

$$A_{x|i,z|j} = \frac{g(z | \mu_{z|j}, \Sigma_{z|j})}{f_z(z)} \Sigma_{x|i,z|j} \Sigma_{z|j}^{-1}.$$
where \( \Sigma_{x|i, x|j} \) is the covariance matrix of \( x \) and \( z \) based on the \( i \)th and \( j \)th components of their GMM’s.

**Proof** To prove this proposition, we use the conditional density function as defined in [44], [47], expressed as

\[
fx|x(z|x) = \frac{fx,z(x, z)}{fz(z)}.
\]

(30)

Now let \( y = \begin{bmatrix} x \\ z \end{bmatrix} \), \( \mu_{y|i, j} = \begin{bmatrix} \mu_{x|i} \\ \mu_{z|j} \end{bmatrix} \) and \( \Sigma_{y|i, j} = \begin{bmatrix} \Sigma_{x|i} & \Sigma_{x|i, x|j} \\ \Sigma_{x|i, x|j} & \Sigma_{z|j} \end{bmatrix} \).

Herein, the joint PDF of \( x \) and \( z \) is approximated as follows,

\[
f_{x, z}(x, z) = \sum_{i=1}^{L_x} \sum_{j=1}^{L_z} \omega_{x|i} \omega_{z|j} g(y|\mu_{y|i, j}, \Sigma_{y|i, j}),
\]

(31)

Therefore,

\[
f_{x, z}(x, z) = \sum_{j=1}^{L_z} \omega_{z|j} f_{x, z}(x, z|\mu_{z|j}, \Sigma_{z|j}),
\]

(32)

where \( f_{x, z}(x, z|\mu_{z|j}, \Sigma_{z|j}) \) is the joint PDF of \( x \) and \( z \) based on the \( j \)th GMM component of \( z \)’s PDF. Multiplying both the numerator and denominator in (30) by \( g(z|\mu_{z|j}, \Sigma_{z|j}) \), we can obtain the following formula,

\[
f_{x|z}(x|z) = \sum_{j=1}^{L_z} \omega_{z|j} \frac{f_{x, z}(x, z|\mu_{z|j}, \Sigma_{z|j})}{f_{z}(z)} \cdot g(z|\mu_{z|j}, \Sigma_{z|j}).
\]

(33)

Finally, integrating \( x f_{x|z}(x|z) \) with respect to \( x \) and using Proposition 2, the formula in (28) can be obtained. Therefore, \( \mathbb{E}(x|z) \) in (28) presents the general formula of the proposed MMSE-based estimator, \( \hat{x}_{MS} \), as originally expressed in (13).

V. PROPERTIES OF MEAN SQUARED ESTIMATOR

The purpose behind this section is to analyze the main properties of the proposed estimator, which are described below.

A. Unbiasedness and minimum variance

Property 1 (Unbiasedness): The estimator \( \hat{x}_{MS} \) as defined in (13) is unbiased.

**Proof** It is obvious that \( \hat{x}_{MS} = \mathbb{E}(x|z) \) is unbiased when \( z \) is GRV. By applying the mean function on both sides of the equation in (15), we can show,

\[
\mathbb{E}(\hat{x}_{MS}) = \mu_{x}.
\]

(34)

For \( z \) being an NGRV, the expected value of the expression in (28) can be shown as,

\[
\mathbb{E} \left( \frac{z - \mu_{z|i}}{f_{z}(z)} \right) = \int_{-\infty}^{\infty} \frac{g(z|\mu_{z|i}, \Sigma_{z|i})}{f_{z}(z)} [z - \mu_{z|i}] f_{z}(z) dz
\]

(35)

Therefore, applying the expectation function on all the elements of the summation in (28), the equation (34) is obtained which verifies that \( \hat{x}_{MS} \) is an unbiased estimator.

Property 2 (Minimum variance): an estimator is said to be a minimum variance estimator (MVE) if it has the smallest error variance [44].

**Proof** Since this is a minimum mean squared error estimator [44], it is therefore an MVE. Moreover, the value of the error variance is given in (14).

Moreover, the minimum error variance that minimize the function in (12) as expressed in (14) can be further developed as,

\[
J^*(\hat{x}_{MS}) = \mathbb{E} \left( \sum_{n \in N} x_{n}^2 \right) - \sum_{n \in N} (\mathbb{E}(x_{n}|z))^2
\]

(36)

or,

\[
\sigma^2(x_{n}|z_1, \ldots, z_M) \leq \sigma^2(x_{n}|z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_M),
\]

\[\forall n \in \{1, \ldots, N\}, \forall j \in \{1, \ldots, M\}\]

(37)

Where \( M \) is the maximum number of PMU measurements based on full network observability. For instance, since the error is a zero-mean, the impact of missing a single measurement, e.g., \( z_j \) in (37), can be quantified by a positive added value to the minimum error variance as,

\[
\Delta J^* = \sum_{n \in N} \left( \sigma^2(x_{n}|z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_M) - \sigma^2(x_{n}|z_1, \ldots, z_M) \right).
\]

(38)

Therefore, the maximum value of the minimum error variance can be reached in the absence of any measurements. This can be quantified as,

\[
J^*(\hat{x}_{MS}) \leq J^*_{\text{max}} = \sum_{n \in N} \sigma^2_{x_{n}}.
\]

(39)

It should be noted that the minimum required number of measurements depends on the statistics of the state parameters, as well as the selection of a suitable threshold value that
we have imposed on the error variance. The threshold value is selected in order to provide a tradeoff between the state estimation accuracy and the number of measurements. This will be further discussed in the next section (i.e., see also Eq. 47).

B. Bad measurement identification

Many factors can lead to bad or corrupt measurements such as device malfunction or malicious data injection [48], [49]. The presence of bad measurements can be quantified by [50],

$z' = z + o$

(40)

where $z$ is as defined in (6) and $z'$ is the infected measurements by an unknown vector $o$. An element $o_j$ in $o$ is non-zero only if $z_j$ is a bad datum. Therefore, using the general formula of MSE as given in (28), the impact of bad measurements on the estimated states can be expressed as,

$$\Delta \hat{x}_{MS} = Bo.$$  

(41)

with,

$$B = \sum_{i=1}^{L_x} \sum_{j=1}^{L_z} \omega_{xi} \omega_{zi} A_{zi,j}.$$  

(42)

On the other hand, from (14) and using $E(x|z) = \hat{x}_{MS}$, we obtain,

$$J^*(\hat{x}_{MS}) = E(x^T x|z) - \hat{x}_{MS}^T \hat{x}_{MS}.$$  

(43)

Therefore, based on an proposition [3], $J^*(\hat{x}_{MS})$ can be calculated immediately using the GMM presentation of the PDF of $x^T x$. Moreover, the mean value of $J^*(\hat{x}_{MS})$ can be evaluated using the following property of conditional variance,

$$\sigma^2_x = E(\sigma^2(x|z)) + \sigma^2(E(x|z)),$$  

(44)

In other expression,

$$\sigma^2(x|z) = \sigma^2_x - \sigma^2_{\hat{x}_{MS}},$$  

(45)

From (36) and (45),

$$E(J^*(\hat{x}_{MS})) = \sum_{n \in \Omega} (\sigma^2_x - \sigma^2_{\hat{x}_{MS}}),$$  

(46)

where $(\sigma^2_x - \sigma^2_{\hat{x}_{MS}})$ is the $n^{th}$ component of $\sigma^2_x - \sigma^2_{\hat{x}_{MS}}$. Knowing the mean value of the minimum error variance in the case of no bad data present in the measurements vector, a threshold of $J^*(\hat{x}_{MS})$ can be defined in terms of $E(J^*(\hat{x}_{MS}))$ as,

$$J^*(\hat{x}_{MS}) \leq \alpha E(J^*(\hat{x}_{MS})).$$  

(47)

This criteria allows to determine if the measurements vector is holding a bad datum. In the case of the minimum variance bigger than the threshold, it is proposed to eliminate one-by-one measurements and reevaluate until respect of the criteria.

VI. SIMULATION AND DISCUSSIONS

Simulations using benchmark IEEE test networks i.e., IEEE 14-bus, IEEE 30-bus and IEEE 118-bus have been conducted to validate the performances of the MMSE-based approach for power systems. In this section, a description of our simulation environment and test systems are presented followed by a brief description of a benchmark estimation method such as WLS, which is used for the sake of comparison. Finally, we analyze and discuss the results.

A. Simulations description

Performances of the proposed MMSE-based estimator for the three test models are evaluated using theIEEE test models. Bear in mind that the proposed approach requires a prior probability of the SPF in order to provide distribution for both estimated and measured parameters. For the sake of generalization, we use the data provided in [51], [52] for power generation while the load power is assumed to be random variables with mean values equal to the load, as in [51], [52]. We assume that all the loads have a ratio of the standard deviation over the mean value; $CV = 0.1$. To validate the performance of the proposed approach for a general case of non-Gaussian parameters, GMM parameters are required. For the sake of simplification and generalization, a data in the case of two Gaussian components is randomly generated given the mean and standard deviation of the power at each bus and verifying the GMM conditions in (3) and (5).

The active power at all the buses is assumed to be correlated. The correlation matrix for load powers is generated randomly in [0, 1] and is kept the same within the GMM components while the active and reactive powers are assumed to be fully correlated while keeping the power factor constant. Then the Monte Carlo simulations (MCS) are ran for $N_{mcs} = 10000$ simulations in order to solve the Newton Raphson based power flow using the package of MATLAB MATPOWER provided in [53]. Based on the MCS results, the fitting function of Matlab “gmdistribution.fit” is then used to provide the GMM components of the MCS outputs; mainly the estimation parameters and the measurements. This function gives maximum likelihood estimates of the parameters in a Gaussian mixture model with, for instance, $k = 2$ components. The analytical formula of the proposed MSE estimator expresses the estimated parameters as the mean value corrected by an additive term which is the product of the covariance matrices of estimated and measured parameters and the subtraction of current value and mean value of the measurements (see equation (28)). This process would allow the estimator to perform a highly accurate state estimation with a limited number of measurements. To validate this capability, four different cases, defined by the number redundancy of the measurements as shown in Table [4] have been investigated. Note that the redundancy is defined as the ratio of the measurements number over estimated parameters number [22]. The notations $|\mathcal{N}_i|$, $|\mathcal{Z}_{x,V_i}|$, $|\mathcal{Z}_{x,Q_i}|$ and $|\mathcal{Z}_{P_{f},Q_{f}}|$ are the sizes of the estimated parameters, the measured voltage, the measured power injection and measured power low sets respectively. We assume an Additive White
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/TPWRS.2018.2801121, IEEE Transactions on Power Systems

TABLE I: Measurements configuration for the 3 test networks

| Test system | Case | $|N_v|$ | $|Z_{0,v}|$ | $|Z_{f,v}|$ | $|Z_{F,v}|$ | $|Z_{P,v}|$ | Redundancy |
|-------------|------|--------|-------------|-------------|-------------|-------------|------------|
| IEEE        | 1    | 27     | 14          | 28          | 40          | 3.04        |
|             | 1-4  | 22     | 8           | 14          | 20          | 1.56        |
| 14-bus      | 3    | 27     | 4           | 10          | 14          | 1.04        |
|             | 3    | 27     | 4           | 8           | 10          | 0.81        |
| 30-bus      | 9    | 59     | 60          | 82          | 2.90        |
|             | 11   | 59     | 60          | 82          | 2.90        |
| 118-bus     | 2    | 235    | 118         | 236         | 372         | 3.04        |
|             | 4    | 235    | 118         | 236         | 372         | 3.04        |
| 30-bus      | 3    | 235    | 40          | 78          | 124         | 1.03        |
|             | 3    | 235    | 40          | 78          | 124         | 1.03        |
| 30-bus      | 4    | 235    | 30          | 60          | 94          | 0.78        |
|             | 4    | 235    | 30          | 60          | 94          | 0.78        |
| IEEE        | 1    | 77     | 40          | 80          | 110         | 2.99        |
|             | 11   | 77     | 20          | 40          | 56          | 1.51        |
| 30-bus      | 3    | 77     | 14          | 26          | 38          | 1.01        |
|             | 3    | 77     | 14          | 26          | 38          | 1.01        |
| 30-bus      | 4    | 77     | 10          | 20          | 28          | 0.75        |
|             | 4    | 77     | 10          | 20          | 28          | 0.75        |

Gaussian noise (AWGN) for the measurements with standard deviations of 1% for voltage (amplitude and angle) and 2% for power injections and flows. A sample size of $N_{spl} = 200$ has been considered for the MCS running of WLS, UKF, and MSE. In these experiments we consider $N_{spl}$ different measurements generated according to the obtained PDFs from the SPF study. For each experiment, WLS, UKF, and MSE have been used to calculate the estimated states, i.e. voltage amplitude and angle. These simulations have been run on an Intel i7-4600U processor at 2.7 GHz using a Matlab code.

### B. Comparison methods

In this study, WLS and UKF have been used as reference methods to evaluate the relative performance of the proposed MSE method, especially in the case of redundant measurements scenarios. Furthermore, for scenarios with a limited number of measurement where the is no guarantee that WLS and UKF can converge, we consider a combination of compressive sensing and UKS which will be referred to as CS-UKF. Such a combination is inspired by the SPF methods presented in [49], [50], [54], [55]. In this paper, a CS-UKF-based approach is developed as an additional reference method to aid our evaluations. This method is based on two successive steps. The first considers CS to reconstruct the measurements vector followed by UKF to estimate the system states. The CS process exploits the available statistics on the measurement parameters, as well as their correlations to reconstruct the complete vector of measurements. Bear in mind that the CS theory is based on reconstruction of the complete signal (more redundant measurement vector) from available compressed measurements (less redundant measurements vector). The availability of statistics on the state parameters makes the measurement vector a spare signal which can then be compressed and reconstructed. The recovery of a more redundant measurement vector $z^{\circ}$ from a limited measurement vector using the SPF information can be developed as,

$$z^{\circ} = \mathbb{E}(z^{\circ}|z^{\text{m}})$$  \hspace{1cm} (48)

where $z^{\circ}$ is the recovered measurements vector. Therefore, using Proposition 3 the conditional expectation can be evaluated from the measurement parameters statics. Note that $z^{\text{m}}$ contains $z^{\text{m}}$ elements besides other unavailable measurements, then $|z^{\circ}| \geq |z^{\text{m}}|$.  

### C. Performance comparisons

In these experiments we compare the performance of the proposed MSE to those obtained by WLS and UKF. The true values are considered as the state values obtained from the power flow. To assess MSE, UKF and WLS performances, the mean absolute error ($\hat{x}_{\text{MAE}}$) of both voltage amplitude and angle are calculated using the following expression,

$$\hat{x}_{\text{MAE}} = \frac{1}{N_x} \sum_{i=1}^{N_x} \mathbb{E}_{\text{spl}}(|\hat{x}_i - x_{i\text{true}}|),$$  \hspace{1cm} (49)

with

$$\mathbb{E}_{\text{spl}}(x) = \frac{1}{N_{\text{spl}}} \sum_{j=1}^{N_{\text{spl}}} x_j$$

where $x_j$ is the value of $x$ corresponding to the $j$th element in the considered sample. $x_{i\text{true}}$ and $\hat{x}_i$ are the true and estimated values of state $i$, while $N_x$ is the number of estimated states.

### D. Results and discussion

Table II displays the $\hat{V}_{\text{MAE}}$ and $\hat{\theta}_{\text{MAE}}$ results for the IEEE test models using 4 different numbers of measurements (i.e., case 1-4). For the three test system, WLS converges for cases 1 and 2. However, it diverges for cases 3 and 4 where the measurement redundancy is lower ($<\sim 1$). Comparing cases 1 and 2, the results show that MSE outperforms WLS and UKF in terms of MAE values (see equation 49). Fig. 1 illustrates the values of mean absolute error over the considered sample of each estimated state. Figs. 1a, 1b and 1c show the voltage amplitudes absolute error for IEEE-14, -30 and -118 bus, respectively, while Figs. 1d, 1e and 1f display their voltage angles absolute error. As can be observed for all three IEEE test models, the MSE performances are well below those of WLS and UKF for all buses verifying the superior performances of the proposed estimator. In addition, with an even smaller number of measurements (e.g., cases 3 and 4), the performance of MSE does not deteriorate considerably as indicated in Table II. Fig. 2 illustrates the obtained curves of the mean absolute error over the considered simulation sample of each estimated state. Indeed, Figs. 2a, 2b and 2c illustrate the voltage amplitudes absolute error for IEEE-14, -30 and -118 bus respectively. Moreover, Figs. 2d, 2e and 2f are illustrating voltage angles absolute error for IEEE-14, -30 and -118 bus respectively. Finally, for the sake of comparison in cases 3 and 4 where WLS and UKF do not converge, CS-UKF has been deployed and the results are illustrated in Fig. 3. It shows that MSE overcomes CS-UKF. Besides, comparing the execution time of the estimation methods which are given in Table II which shows that MSE can significantly
TABLE II: Mean values of the absolute error for different measurements configuration of IEEE benchmark systems

<table>
<thead>
<tr>
<th>Test system</th>
<th>Method</th>
<th>$V_{\text{MAE}} \times 10^3$</th>
<th>Execution time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test system</td>
<td>Cases</td>
<td>Method</td>
<td>$\theta_{\text{MAE}} \times 10^3$</td>
</tr>
<tr>
<td>IEEE 118-bus</td>
<td>WLS</td>
<td>0.0760</td>
<td>0.0856</td>
</tr>
<tr>
<td></td>
<td>UKF</td>
<td>0.7197</td>
<td>0.3534</td>
</tr>
<tr>
<td></td>
<td>CS-UKF</td>
<td>0.0719</td>
<td>0.3534</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0422</td>
<td>0.0614</td>
</tr>
<tr>
<td>IEEE 30-bus</td>
<td>WLS</td>
<td>0.2945</td>
<td>0.8930</td>
</tr>
<tr>
<td></td>
<td>UKF</td>
<td>0.0839</td>
<td>0.3755</td>
</tr>
<tr>
<td></td>
<td>CS-UKF</td>
<td>0.0839</td>
<td>0.1235</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0595</td>
<td>0.0905</td>
</tr>
<tr>
<td>IEEE 118-bus</td>
<td>WLS</td>
<td>0.1241</td>
<td>0.3632</td>
</tr>
<tr>
<td></td>
<td>UKF</td>
<td>0.0540</td>
<td>0.3575</td>
</tr>
<tr>
<td></td>
<td>CS-UKF</td>
<td>0.0540</td>
<td>0.0983</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>0.0147</td>
<td>0.0361</td>
</tr>
</tbody>
</table>

Fig. 1: Comparison of $E_{\text{ap}(|V_\hat{i} - V_{\text{true}}|)}$ for (a) IEEE-14 bus, (b) IEEE-30 bus and (c) IEEE-118 bus and of $E_{\text{ap}(|\theta_\hat{i} - \theta_{\text{true}}|)}$ for (d) IEEE-14 bus, (e) IEEE-30 bus and (f) IEEE-118 bus using WLS, UKF and MSE in cases 1 and 2 of measurements configuration.

outperform CS-UKF. In addition, as shown in Table II, CS-UKF, UKF and WLS, due to their iterative nature, would require a significantly higher execution time than MSE. This unique future is important for real-time grid monitoring and fast state estimation.

E. Case of bad measurements

The study of this case aims at showing the performances of the proposed estimator on identifying and eliminating bad measured data. To realize this purpose, a samples of 200 estimations is evaluated using MSE. In this case, arbitrary values have been added to the measurement vector at iterations over the range [50−59] to one parameter and over [100−109] to two parameters applied to IEEE-118 bus network. Fig. 4 illustrates the variation of the voltage angle estimation. It is shown in Fig. 4a that the error increases clearly when some measurements are corrupted. Moreover, as expected in our previous analysis, the error variance also increases in terms of number and value of the bad datum which allows to identify and then eliminate it. To decide whether the measurements are corrupted or not, a chosen value of $\alpha = 2$ in the criteria in (47).

F. Use case using EMTP-RV

The design of the proposed estimator is based on availability of the system statistics under various test environments. To gather such statistics we use an EMTP-RV software tool for real-time assessment of the proposed method [56]. EMTP-RV

1Certain commercial equipment, instruments, or materials are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.
is an Electro Magnetic Transients Program (EMTP) software tool where PMU devices can be placed at any desirable location in the grid network. EMTP-RV has been widely used as a time domain transient solution [24], [57]. It is able to provide almost all power system components, such as power plant, transformer, windfarm, different kinds of faults, overhead lines with line and ground wires and towers, as well as underground cables. In our experiments using EMTP-RV we considered the IEEE 39-bus transmission model where PMUs are placed on buses to collect and process samples of power signals. To assess the impact of uncertainties on the proposed state estimation, we added 4 wind parks, each of 100 wind
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/TPWRS.2018.2801121, IEEE Transactions on Power Systems

I E
I E
1
2
arrays of configurations given in Table I.

collected and then applied to evaluate the MSE performance.

Power and voltage signals generated by EMTP-RV have been error on the MSE results.

measurements. This work will be extended in order to present methods. It is also shown that the proposed state estimator can verify that the proposed approach outperforms the background compared to those of WLS, UKF and CS-UKF. The results

The performance of the proposed MSE approach has been

we use the GMM to approximate the NGRV’s PDF as a
general case where a state parameters and measurements

are non-Gaussian random variables (NGRV’s). In our analysis,

are non-Gaussian random variables (NGRV’s). In our analysis,

in correlated uncertain environment using unscented transformation.


VII. CONCLUSION

In this paper, a new analytical-based state estimator, referred to as Mean Squared Estimator (MSE), has been proposed for the general case where a state parameters and measurements

are non-Gaussian random variables (NGRV’s). In our analysis,

we use the GMM to approximate the NGRV’s PDF as a

weighted summation of Gaussian components to derive a general formula as a conditional expectation of the estimated states for a given set of measurements. Simulations have been carried out using three benchmark IEEE test models. The performance of the proposed MSE approach has been compared to those of WLS, UKF and CS-UKF. The results verify that the proposed approach outperforms the background methods. It is also shown that the proposed state estimator can perform highly accurate estimations with a limited number of measurements. This work will be extended in order to present more analysis on dynamic SPF and on the sensitivity of SPF error on the MSE results.

REFERENCES


Fig. 5: Comparison of (a) $E_{\text{app}}(|\hat{V}_i - V_i^{\text{true}}|)$ and of (b) $E_{\text{app}}(|\theta_i - \theta_i^{\text{true}}|)$ for IEEE-118 bus using MSE in cases 1, 2, 3 and 4 of measurements configuration.
Fig. 6: IEEE 39-bus with high penetration of distributed generation


[33] J. Milton Brown Do Coutto Filho, "Forecasting-Aided State Estimation-


