Adaptive Extended Kalman Filter with Correntropy Loss for Robust Power System State Estimation

Zhiyu Zhang, Jinzhe Qiu and **Wentao Ma ***

School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China; zhangzhiyu@xaut.edu.cn (Z.Z.); qiujinzhe@163.com (J.Q.)

* Correspondence: mawt@xaut.edu.cn

Abstract: Monitoring the current operation status of the power system plays an essential role in the enhancement of the power grid for future requirements. Therefore, the real-time state estimation (SE) of the power system has been of widely-held concern. The Kalman filter is an outstanding method for the SE, and the noise in the system is generally assumed to be Gaussian noise. In the actual power system however, these measurements are usually disturbed by non-Gaussian noises in practice. Furthermore, it is hard to get the statistics of the state noise and measurement noise. As a result, a novel adaptive extended Kalman filter with correntropy loss is proposed and applied for power system SE in this paper. Firstly, correntropy is used to improve the robustness of the EKF algorithm in the presence of non-Gaussian noises and outliers. In addition, an adaptive update mechanism of the covariance matrixes of the measurement and process noises is introduced into the EKF with correntropy loss to enhance the accuracy of the algorithm. Extensive simulations are carried out on IEEE 14-bus and IEEE 30-bus test systems to verify the feasibility and robustness of the proposed algorithm.

Keywords: correntropy loss; extended Kalman filter; adaptive update mechanism; power system robust state estimation; non-Gaussian noises

1. Introduction

The power system state estimation (SE) is the foundation and core of the energy management system, and it is indispensable for power system safety, reliability, quality and economic operation [1]. SE is usually divided into static state estimation (SSE) and forecasting-aided state estimation (FASE), FASE is also called dynamic state estimation in some studies [2]. SSE can grasp the real-time operating state of the power system, and FASE can predict the operating trend of the system. The power system is a quasi-steady state system and the state information of buses will change with the change of loads. The SSE cannot consider the impact of the change of system loads on system status information, while the FASE with the function of analyzing and predicting the operation trend of the system is required. Therefore, research on power system forecasting-aided state estimation (PSFASE) is of great significance [3–5].

At present, the achievements of FASE are mainly based on different Kalman filters and their modifications. The original Kalman filtering algorithm can only solve linear problems, and it loses its advantage when facing complex nonlinear systems like a power system. Subsequently, an extended Kalman filter (EKF) was proposed in [6], which can linearize the nonlinear measurement function by Taylor series and can be applied in the nonlinear system. In [7], EKF has been used to solve the PSFASE. Considering the adaptability of the algorithm is poor when the power system sudden load changes, an adaptive EKF was proposed in [8]. To enhance the robustness of the algorithm when bad data exists in the measurement data, the robust EKF was proposed in [9]. A new robust

generalized maximum-likelihood-type EKF is developed in [10] which can bound the influence of the disturbances. To avoid first-order approximation errors of the EKF, which may be large under strong nonlinearities of the model, the iterated EKF (IEKF) [11] has been proposed as an alternative method. The IEKF iteratively linearizes the functions of nonlinear system to compensate for the higher order terms. To improve the stability of algorithm when the system model involves uncertainty, the H∞ EKF was proposed in [12] which quotes the robust control theory. To summarize, the EKF and its modifications make a great breakthrough in the optimization of the nonlinear system and widely applied for the PSFASE due to its simplicity, high computational efficiency and superior performance in highly nonlinear systems.

Although EKF can solve the nonlinear problem, these methods ignore the influence of higher order terms in the process of linearization and increase the computational complexity because it needs to calculate Jacobian matrix. For resolving the above challenge, the unscented Kalman filter (UKF) was developed in [13] to approximate the nonlinear distribution by sampling method to solve the nonlinear problem. In [14], a UKF based power system dynamic state estimation was proposed. However, all these approaches mentioned above have suffered from several important defects, limiting them from being adopted for PSFASE. To be specific, they cannot handle: 1) the non-Gaussian process and observation noises of the system nonlinear dynamic models, and 2) the unknown noise covariance matrices. In the actual power grids, the non-Gaussian process and observation noises are ubiquitous and hence the inappropriate noise covariance matrices will lead to imprecise estimation results.

The traditional EKF and UKF methods are derived based on mean square error (MSE) which is optimal under Gaussian assumption. Therefore, the SE methods based on the original EKF and UKF will show un-robustness when the system suffers from the non-Gaussian noises, that is, the state cannot be estimated correctly. To overcome the influence of non-Gaussian noise, a novel maximum correntropy criteria (MCC) [15–19] has been developed in an information theoretic learning (ITL) methodology for non-Gaussian signal processing. At present, a novel robust EKF based on the MCC (called MCC-EKF) was developed in which the MSE was substituted by the MCC to solve estimation issues in non-Gaussian noise environments [20–22]. In this paper, the MCC-EKF is firstly used to design a robust SE method to suppress the interference of non-Gaussian noise. In addition, a novel adaptive update mechanism of the predicted error covariance matrix and measurement noise covariance matrix is introduced into the MCC-EKF, and an adaptive MCC-EKF (called AMCC-EKF) method is proposed to enhance the tracking ability of the original MCC-EKF. The performance of the proposed AMCC-EKF, MCC-EKF, UKF and EKF for PSFASE are tested in this paper. Experimental results illustrate that the proposed algorithm performs better than other algorithms with respect to the estimation accuracy under the non-Gaussian circumstances.

The remainder of this paper is organized as follows. In Section 2 we briefly review the correntropy and extended Kalman filter. In Section 3 we derived the proposed AMCC-EKF algorithm and in Section 4 the proposed algorithm is applied to PSFASE. Section 5 gives the simulation results. Finally, Section 6 concludes with a summary of the main findings of this paper.

2. Preliminaries

2.1. Maximum Correntropy Criteria

Given two random variables X and Y, the correntropy is defined as

$$V_\sigma(X, Y) = E[G_\sigma(X - Y)] = \iint\limits_{x,y} G_\sigma(x - y) dF_{X,Y}(x, y) \tag{1}$$

where $E[\cdot]$ denotes the expectation operator, $F_{X,Y}(x,y)$ is the joint distribution function of X and Y, and $G_\sigma(\cdot)$ represents a shift-invariant Mercer kernel, with bandwidth σ. In this paper, we use the following Gaussian kernel

$$G_\sigma(x-y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{2}$$

Correntropy can be used as a cost function to develop novel robust adaptive filtering algorithms [23–25]. In practice, only finite samples of the variables X and Y are given, and the joint probability density function is unknown in general. Hence the sample mean estimator of corrrentropy is used as

$$\hat{V}_\sigma(X,Y) = \frac{1}{N}\sum_{i=1}^{N} G_\sigma(x_i - y_i) \tag{3}$$

where N is the number of samples. And the performance surface of the maximum square mean (MSE) and correntropy are shown in Figure 1. It is can be seen that correntropy is local whereas MSE is global. By global, we mean that all the samples in the joint space will contribute appreciably to the value of the similarity measure while the locality of correntropy means that the value is primarily dictated by the kernel function along the $x = y$ line. Therefore, correntropy of the error can be used as a robust cost function for adaptive systems training, which will be called the MCC. MCC has the advantage that it is a local criterion of similarity and it should be very useful for cases when the measurement noise is nonzero mean, non-Gaussian, with large outliers [19].
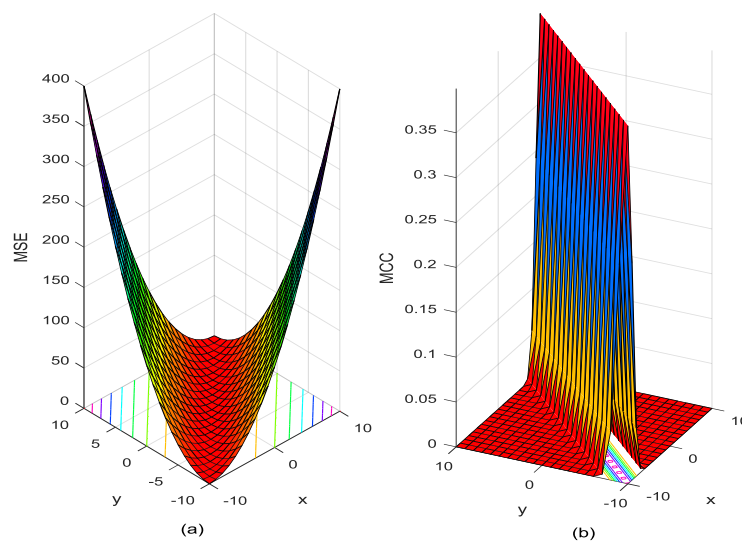

Figure 1. Performance surface of MSE and Correntropy: (**a**) MSE (**b**) Correntropy.

2.2. Review of Extended Kalman Filter

The extended Kalman filter, a derivation of Kalman filter, is developed to resolve the problem of nonlinear system by Taylor series. Consider a nonlinear system by the nonlinear state and linear measurement functions:

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \tag{4}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \tag{5}$$

where $\mathbf{x}_k \in R^n$ denotes a n-dimensional state vector at time step k, $\mathbf{y}_k \in R^m$ denotes a m-dimensional measurement vector at time step k, $f(\cdot)$ represents the vector-valued function and $\mathbf{H}_k$ stands for the observation matrix, $\mathbf{w}_k$ and $\mathbf{v}_k$ are process and measurement noises respectively, which are generally assumed to be uncorrelated with zero mean and covariance matrices. For EKF, the measurement noise

$\mathbf{v}_k$ is also assumed to be a zero mean Gaussian white noise sequence while the MCC-EKF avoids the Gaussian assumption of $\mathbf{v}_k$.

$$E\left[\mathbf{w}_{k-1}\mathbf{w}^T{}_{k-1}\right] = \mathbf{Q}_k \tag{6}$$

$$E\left[\mathbf{v}_k\mathbf{v}^T{}_k\right] = \mathbf{R}_k \tag{7}$$

In general, extended Kalman filter includes the following two steps:

(1) predict: the prior mean and covariance matrix are given by:

$$\hat{\mathbf{x}}_{k|k-1} = f\left(\hat{\mathbf{x}}_{k-1|k-1}\right) \tag{8}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}^T{}_{k-1} + \mathbf{Q}_{k-1} \tag{9}$$

where $\hat{\mathbf{x}}_{k|k-1}$ denotes the predicted state vector at time $(k-1)$. $\mathbf{F}_{k-1}$ denotes a Jacobian matrix of $f(\cdot)$, and is described as $\mathbf{F}_{k-1} = \frac{\partial f\left(\hat{\mathbf{x}}_{k-1|k-1}\right)}{\partial \mathbf{x}_k}$.

(2) update: the gain matrix of extended Kalman filter can be obtained as

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \tag{10}$$

the posterior state is equal to the prior state plus the innovation weighted by the gain matrix of extended Kalman filter

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}\right) \tag{11}$$

Additionally, the iterative formula for the posterior covariance is as follows

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \tag{12}$$

3. Adaptive Extended Kalman Filter With Correntropy Loss

3.1. Extended Kalman Filter with Correntropy Loss

The traditional EKF based on MSE loss and it is not robust when the system noise is non-Gaussian. To overcome the influence of non-Gaussian noise, a novel maximum correntropy criteria (MCC) has been developed in ITL. In [15], the MCC-EKF was developed by using the MCC to replace the MSE, which may perform much better in non-Gaussian noise environments. The main reason is that the correntropy contains second and higher order moments of the error, that is, the correntropy in ITL involves all even moments of the error which can be obtained by Taylor expansion [15].

For the nonlinear model given by (4) (5) and expression (8) (9), we can obtain

$$\begin{bmatrix} \hat{\mathbf{x}}_{k|k-1} \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{H}_k \end{bmatrix}\mathbf{x}_k + \mathbf{q}_k \tag{13}$$

We assumed that the state vector $\mathbf{x}_k$ and process noise $\mathbf{w}_k$ and measurement noise $\mathbf{v}_k$ are non-zero correlation respectively.

Where $\mathbf{I}$ is the $n \times n$ identity matrix, and $\mathbf{q}_k$ can express as

$$\mathbf{q}_k = \begin{bmatrix} -\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\right) \\ \mathbf{v}_k \end{bmatrix} \tag{14}$$

With

$$
\begin{aligned}
E\left[\mathbf{q}_k\mathbf{q}_k^T\right] &= \begin{bmatrix} \mathbf{P}_{k|k-1} & 0 \\ 0 & \mathbf{R}_k \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{B}_{p,k|k-1}\mathbf{B}_{p,k|k-1}^T & 0 \\ 0 & \mathbf{B}_{r,k}\mathbf{B}_{r,k}^T \end{bmatrix} \\
&= \mathbf{B}_k\mathbf{B}_k^T
\end{aligned}
\tag{15}
$$

where $\mathbf{B}_k$ can be computed by Cholesky decomposition of $E\left[\mathbf{q}_k\mathbf{q}_k^T\right]$. Left multiplying both sides of (13) by $\mathbf{B}^{-1}{}_k$, we have

$$
\mathbf{D}_k = \mathbf{W}_k\mathbf{x}_k + \mathbf{e}_k
\tag{16}
$$

where $\mathbf{D}_k = \mathbf{B}_k^{-1}\begin{bmatrix} \hat{\mathbf{x}}_{k|k-1} \\ \mathbf{y}_k \end{bmatrix}$, $\mathbf{w}_k = \mathbf{B}_k^{-1}\begin{bmatrix} \mathbf{I} \\ \mathbf{H}_k \end{bmatrix}$, $\mathbf{e}_k = \mathbf{B}_k^{-1}v_k$. Since $E\left[\mathbf{e}_k\mathbf{e}_k^T\right] = \mathbf{I}$, the residual error $\mathbf{e}_k$ is white.

Then, we define the following cost function $J_L(\mathbf{x}_k)$ based on MCC

$$
J_L(\mathbf{x}_k) = \frac{1}{L}\sum_{i=1}^{L} G_\sigma(d_{i,k} - \mathbf{w}_{i,k}\mathbf{x}_k)
\tag{17}
$$

where $d_{i,k}$ is the i-th element of $\mathbf{D}_k$, $\mathbf{w}_{i,k}$ is the i-th row of $\mathbf{W}_k$, and $L = m + n$ is the dimension of $\mathbf{D}_k$.

Under the MCC, the optimal estimate of $\mathbf{x}_k$ is

$$
\hat{\mathbf{x}}_k = \operatorname*{argmax}_{x_k} J_L(\mathbf{x}_k) = \operatorname*{argmax}_{x_k}\sum_{i=1}^{L} G_\sigma(e_{i,k})
\tag{18}
$$

where $e_{i,k}$ is the i-th element of e_k and

$$
e_{i,k} = d_{i,k} - \mathbf{w}_{i,k}\mathbf{x}_k
\tag{19}
$$

Hence, the optimal solution can be obtained by solving

$$
\frac{\partial J_L(\mathbf{x}_k)}{\partial \mathbf{x}_k} = 0
\tag{20}
$$

It follows easily that

$$
\mathbf{x}_k = \left(\sum_{i=1}^{L}\left[G_\sigma(e_{i,k})\mathbf{w}_{i,k}^T\mathbf{w}_{i,k}\right]\right)^{-1} \times \left(\sum_{i=1}^{L}\left[G_\sigma(e_{i,k})\mathbf{w}_{i,k}^T d_{i,k}\right]\right)
\tag{21}
$$

Since $e_{i,k} = d_{i,k} - \mathbf{w}_{i,k}\mathbf{x}_k$, the formula (21) is a of $\mathbf{x}_k$ and can be rewritten as

$$
\mathbf{x}_k = g(\mathbf{x}_k)
\tag{22}
$$

A fixed-point iterative algorithm can be expressed as

$$
\hat{\mathbf{x}}_{k,t+1} = g(\hat{\mathbf{x}}_{k,t})
\tag{23}
$$

where $\hat{\mathbf{x}}_{k,t}$ denotes the state $\hat{\mathbf{x}}_k$ at the fixed-point iteration t

The Equation (21) can also be expressed as

$$
\mathbf{x}_k = \left(\mathbf{W}_k^T\mathbf{C}_k\mathbf{W}_k\right)^{-1} \times \mathbf{W}_k^T\mathbf{C}_k\mathbf{D}_k
\tag{24}
$$

where $\mathbf{C}_k = \begin{bmatrix} \mathbf{C}_{x,k} & 0 \\ 0 & \mathbf{C}_{y,k} \end{bmatrix}$, with

$$\mathbf{C}_{x,k} = diag(G_\sigma(e_{1,k}), \ldots, G_\sigma(e_{n,k})) \tag{25}$$

$$\mathbf{C}_{y,k} = diag(G_\sigma(e_{n+1,k}), \ldots, G_\sigma(e_{n+m,k})) \tag{26}$$

Based on formula (24), we can obtain a recursive formula [20]

$$\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1} + \overline{\mathbf{K}}_k \left(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \right) \tag{27}$$

where

$$\overline{\mathbf{K}}_k = \overline{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{H}_k \overline{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T + \overline{\mathbf{R}}_k \right)^{-1} \tag{28}$$

$$\overline{\mathbf{P}}_{k|k-1} = \mathbf{B}_{p,k|k-1} \mathbf{C}_{x,k}^{-1} \mathbf{B}_{p,k|k-1}^T \tag{29}$$

$$\overline{\mathbf{R}}_k = \mathbf{B}_{r,k} \mathbf{C}_{y,k}^{-1} \mathbf{B}_{r,k}^T \tag{30}$$

With the above derivations, the optimal state variable of MCC-EKF algorithm can be obtained by the formula (25)–(30).

3.2. Adaptive Extended Kalman Filter with Correntropy Loss

The traditional EKF algorithm is usually used under the condition that the statistical characteristics of state noise and measurement noise of the system are known, but it is unknown in the actual situation. Therefore, the filtering divergence and inaccurate estimation results may be generated by the influence of the uncertain factors. Similarity to the EKF, The MCC-EKF algorithm will still suffer from this problem. In order to improve the accuracy of MCC-EKF algorithm, a covariance matrix adaptive mechanism is introduced into the MCC-EKF algorithm to continuously estimate and modify the filter noise statistical characteristics online, and we call the novel algorithm adaptive MCC-EKF algorithm which can improve the filtering accuracy while filtering by using the information of observation data. Thus, the optimal value of the estimated state is obtained [26,27]. Now, the covariance matrix adaptive update mechanism is given.

We define the new information d_k that is the error with the actual observation value $\mathbf{y}_k$ and the predicted observation value $\hat{\mathbf{y}}_{k|k-1}$ at time k as

$$d_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1} = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} = \mathbf{H}_k \widetilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_k \tag{31}$$

where $\widetilde{\mathbf{x}}_{k|k-1} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}$ denotes one-step prediction error.

Then, according to the windowing estimation method, the real-time estimation variance of the new information d_k is

$$\hat{\mathbf{C}}_{d_k} = \begin{cases} \frac{k-1}{k} \hat{\mathbf{C}}_{d_{k-1}} + \frac{1}{k} d_k d_k^T, & k \le W \\ \frac{1}{W} \sum\limits_{i=k-W+1}^{k} d_i d_i^T, & k > W \end{cases} \tag{32}$$

where W is the size of the moving window. Since $\mathbf{v}_k$ and $\hat{\mathbf{x}}_{k|k-1}$ are uncorrelated, we have

$$\hat{\mathbf{C}}_{d_k} = E\left(d_k d_k^T \right) = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \tag{33}$$

The residual r_k at time k is defined as the error with the actual observed value $\mathbf{y}_k$ and the estimated observed value $\hat{\mathbf{y}}_k$

$$
\begin{aligned}
r_k &= \mathbf{y}_k - \hat{\mathbf{y}}_k \\
&= \mathbf{y}_k - \mathbf{H}_k \left[\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \right) \right] \\
&= (\mathbf{I} - \mathbf{H}_k \mathbf{K}_k) d_k \\
&= \left(\mathbf{I} - \frac{\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T}{\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k} \right) d_k \\
&= \mathbf{R}_k C_{d_k}^{-1} d_k
\end{aligned}
\tag{34}
$$

The variance of the residual r_k is defined as

$$
C_{r_k} = E\left(r_k r_k^T \right) = \mathbf{R}_k C_{d_k}^{-1} C_{d_k} C_{d_k}^{-1} \mathbf{R}_k = \mathbf{R}_k C_{d_k}^{-1} \mathbf{R}_k
\tag{35}
$$

According to Kalman filtering principle [26–29], the filtering gain is

$$
\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T C_{d_k}^{-1} = \mathbf{P}_k \mathbf{H}_k^T \mathbf{R}_k^{-1}
\tag{36}
$$

Left multiplying both sides of Equation (36) by $\mathbf{H}_k$, and substitute into Equation (33), we can obtain

$$
\begin{aligned}
\mathbf{H}_k \mathbf{P}_k \mathbf{H}_k^T \mathbf{R}_k^{-1} &= \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T C_{d_k}^{-1} \\
&= \left(C_{d_k} - \mathbf{R}_k \right) C_{d_k}^{-1} \\
&= \mathbf{I} - \mathbf{R}_k C_{d_k}^{-1}
\end{aligned}
\tag{37}
$$

Right multiplying both sides of Equation (37) by $\mathbf{R}_k$, we have

$$
\mathbf{H}_k \mathbf{P}_k \mathbf{H}_k^T = \mathbf{R}_k - \mathbf{R}_k C_{d_k}^{-1} \mathbf{R}_k
\tag{38}
$$

Combining (35) and (38), we can get the $\hat{\mathbf{R}}_k$ as follows

$$
\hat{\mathbf{R}}_k = \hat{C}_{r_k} + \mathbf{H}_k \mathbf{P}_k \mathbf{H}_k^T
\tag{39}
$$

Similarly, the new information d_k can be used to estimate the variance matrix $\mathbf{Q}$ of the state noise, Right multiplying both sides of Equation (36) by $C_{d_k} \mathbf{K}_k^T$, we have

$$
\mathbf{K}_k C_{d_k} \mathbf{K}_k^T = \mathbf{P}_{k|k-1} \mathbf{H}_k^T C_{d_k}^{-1} C_{d_k} \mathbf{K}_k^T = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{K}_k^T
\tag{40}
$$

Transposing both sides of formula (22) and applying Kalman filtering principle, we can obtain

$$
\mathbf{K}_k C_{d_k} \mathbf{K}_k^T = \mathbf{P}_{k|k-1} - \mathbf{P}_k = \mathbf{F}_{k,k-1} \mathbf{P}_k \mathbf{F}_{k,k-1}^T + \mathbf{Q}_{k-1} - \mathbf{P}_k
\tag{41}
$$

To keep $\mathbf{Q}$ semi-positive, the approximate estimate of $\mathbf{Q}_{k-1}$ can be expressed as

$$
\hat{\mathbf{Q}}_{k-1} = \mathbf{K}_k \hat{C}_{d_k} \mathbf{K}_k^T
\tag{42}
$$

4. Adaptive Extended Kalman Filter With Correntropy Loss for PSDSE

4.1. Power System Dynamic Model

The power system is a nonlinear and complex system. For the FASE of power system, dynamic equation and measurement equation can be expressed in the following form

$$
\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1}
\tag{43}
$$

$$
\mathbf{y}_k = h(\mathbf{x}_k) + \mathbf{v}_k
\tag{44}
$$

where $\mathbf{x}_k$ denotes the state vector consisting of magnitudes and angles of nodal voltage. The measurement vector $\mathbf{y}_k$ comprises of voltage magnitude measurements, real power injection measurements, reactive power injection measurements, real power flow measurements, reactive power flow measurements. The noise $\mathbf{w}_k$ and $\mathbf{v}_k$ are usually assumed to be Gaussian noise and independent of each other. Specially, the noise $\mathbf{w}_k$ is the error of system and the noise $\mathbf{v}_k$ is the error of measurement. $f(\cdot)$ represents the function that relates $\mathbf{x}_{k-1}$ to $\mathbf{x}_k$ and $h(\cdot)$ stands for the measurement function that relates $\mathbf{x}_k$ to $\mathbf{y}_k$.

In order to determine the dynamic model of power system, it is necessary to identify the model parameters. The Holt's two-parameter linear exponential smoothing technique [30], also known as linear extrapolation method, is most commonly used to calculate the state transition matrix $\mathbf{F}_k$, and in this paper we use this method as the dynamic model. It can also be used as a simple short-term load forecasting method. It has the advantages of less storage variables and a faster computation speed, and it is suitable for online calculation. Employing this method, the state transition function $f(\cdot)$ is defined as

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{a}_k + \mathbf{b}_k \tag{45}$$

where

$$\mathbf{a}_k = \alpha_k \hat{\mathbf{x}}_{k|k} + (1 - \alpha_k)\hat{\mathbf{x}}_{k+1|k} \tag{46}$$

$$\mathbf{b}_k = \beta_k(\mathbf{a}_k - \mathbf{a}_{k-1}) + (1 - \beta_k)\mathbf{b}_{k-1} \tag{47}$$

where both α_k and β_k are parameters lying in the range from 0 to 1, and vectors $\mathbf{a}_k$ and $\mathbf{b}_k$ at time k are obtained as (43) and (44). Linearization of the nonlinear model in formula (43) above can be expressed as follows

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}_k\hat{\mathbf{x}}_{k|k} + \mathbf{u}_k + \mathbf{w}_k \tag{48}$$

where $\mathbf{F}_k = \frac{\partial f(\mathbf{x}_k)}{\partial \mathbf{x}_k}$, $\mathbf{u}_k$ is a nonrandom external actor in the expansion. Combining formulae (48) and (45), we have

$$\mathbf{F}_k = \alpha_k(1 + \beta_k) \tag{49}$$

$$\mathbf{u}_k = (1 + \beta_k)(1 - \alpha_k)\hat{\mathbf{x}}_{k+1|k} - \beta_k\mathbf{a}_{k-1} + (1 - \beta_k)\mathbf{b}_{k-1} \tag{50}$$

The composition of the measurement vector $\mathbf{y}_k$ changes with the measurement method of the power system. This paper studies the data measured by SCADA system, so $h(x)$ is a nonlinear function. For bus i, the relationship between the measurement and the state vector as follows [31]

$$P_i = \sum_{j=1}^{N}|V_i||V_j|\left(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}\right) \tag{51}$$

$$Q_i = \sum_{j=1}^{N}|V_i||V_j|\left(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}\right) \tag{52}$$

$$P_{ij} = V_i^2 G_{ij} - |V_i||V_j|\left(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}\right) \tag{53}$$

$$Q_{ij} = -V_i^2\left(B_{gi} + B_{ij}\right) - |V_i||V_j|\left(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}\right) \tag{54}$$

where P_i is the real power injection at bus i, Q_i is the reactive power injection at bus i, P_{ij} is the real power flow between buses i and j, Q_{ij} is the reactive power flow between buses i and j, V_i is the voltage magnitude at bus i, G_{ij} is the conductance of the line between buses i and j, B_{ij} is the susceptance of the line between buses i and j.

And the Jacobian matrix $\mathbf{H}_k$ can be expressed as

$$\mathbf{H}_k = \begin{bmatrix} \frac{\partial V_i}{\partial \theta} & \frac{\partial P_i}{\partial \theta} & \frac{\partial Q_i}{\partial \theta} & \frac{\partial P_{ij}}{\partial \theta} & \frac{\partial Q_{ij}}{\partial \theta} \\ \frac{\partial V_i}{\partial V} & \frac{\partial P_i}{\partial V} & \frac{\partial Q_i}{\partial V} & \frac{\partial P_{ij}}{\partial V} & \frac{\partial Q_{ij}}{\partial V} \end{bmatrix}^T \tag{55}$$

Combined with the power system network, the specific elements of the Jacobian matrix $\mathbf{H}_k$ are shown as

$$
\begin{cases}
\frac{\partial V_i}{\partial \theta_i} = 0 & \frac{\partial V_i}{\partial \theta_j} = 0 \\
\frac{\partial V_i}{\partial V_i} = 1 & \frac{\partial V_i}{\partial V_j} = 0
\end{cases}
\tag{56}
$$

$$
\begin{cases}
\frac{\partial P_i}{\partial \theta_i} = -V_i^2 B_{ii} - V_i \sum_{j=1}^{N} V_j \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right) \\
\frac{\partial P_i}{\partial \theta_j} = V_i V_j \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right) \\
\frac{\partial P_i}{\partial V_i} = V_i G_{ii} + \sum_{j=1}^{N} V_j \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right) \\
\frac{\partial P_i}{\partial V_j} = V_i \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right)
\end{cases}
\tag{57}
$$

$$
\begin{cases}
\frac{\partial Q_i}{\partial \theta_i} = -V_i^2 G_{ii} + V_i \sum_{j=1}^{N} V_j \left(G_{ij} \sin \theta_{ij} + B_{ij} \cos \theta_{ij} \right) \\
\frac{\partial Q_i}{\partial \theta_j} = -V_i V_j \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right) \\
\frac{\partial Q_i}{\partial V_i} = -V_i B_{ii} + \sum_{j=1}^{N} V_j \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right) \\
\frac{\partial Q_i}{\partial V_j} = V_i \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right)
\end{cases}
\tag{58}
$$

$$
\begin{cases}
\frac{\partial P_{ij}}{\partial \theta_i} = -V_i V_j \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right) \\
\frac{\partial P_{ij}}{\partial \theta_j} = V_i V_j \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right) \\
\frac{\partial P_{ij}}{\partial V_i} = 2V_i G_{ii} - V_j \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right) \\
\frac{\partial P_{ij}}{\partial V_j} = -V_i \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right)
\end{cases}
\tag{59}
$$

$$
\begin{cases}
\frac{\partial Q_{ij}}{\partial \theta_i} = V_i V_j \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right) \\
\frac{\partial Q_{ij}}{\partial \theta_j} = -V_i V_j \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right) \\
\frac{\partial Q_{ij}}{\partial V_i} = -2V_i \left(B_{gi} + B_{ij} \right) - V_j \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right) \\
\frac{\partial Q_{ij}}{\partial V_j} = -V_i \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right)
\end{cases}
\tag{60}
$$

4.2. Adaptive Extended Kalman Filter with Correntropy Loss for Power System Forecasting-Aided State Estimation

Now, we apply this method to PSFASE to solve the problem of non-Gaussian noise and bad data. For n-bus power system, there are $2n-1$ states (contains n voltage amplitudes and $n-1$ voltage phase angles) that needs to be estimated. The dynamic model reflects the change law of system state variables with time. The dynamic model of power system in this paper assumes that the changes in the system parameters, such as load variations, are very slow. Then, we give the detail procedure of MCC-EKF algorithm for PSDSE.

(1) Select the appropriate initial parameters: a proper kernel bandwidth σ and a small positive ε; Set an initial state value $\hat{\mathbf{x}}_{0|0}$ and corresponding covariance matrix $\mathbf{P}_{0|0}$; Let $k = 1$;

(2) Use Equations (8) and (9) to calculate the $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$, and obtain the $\mathbf{B}_{p,k|k-1}$ by Cholesky decomposition;

(3) Let $k = 1$ and $\hat{\mathbf{x}}_{k|k,0} = \hat{\mathbf{x}}_{k|k-1}$, where $\hat{\mathbf{x}}_{k|k,t}$ stands for the estimated state at the fixed-point iteration k;

(4) Calculate the state transition function using (45)–(47) and the Jacobian matrix $\mathbf{H}_k$ using (51)–(60);

(5) Get the estimates state $\hat{\mathbf{x}}_{k|k,t}$ by Equations (61)–(69);

$$
\hat{\mathbf{x}}_{k|k,t} = \hat{\mathbf{x}}_{k|k-1} + \widetilde{\mathbf{K}}_k \left(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \right)
\tag{61}
$$

With

$$\widetilde{\mathbf{K}}_k = \widetilde{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{H}_k \widetilde{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T + \widetilde{\mathbf{R}}_k \right)^{-1} \tag{62}$$

$$\widetilde{\mathbf{P}}_{k|k-1} = \mathbf{B}_{p,k|k-1} \widetilde{\mathbf{C}}_{x,k}^{-1} \mathbf{B}_{p,k|k-1}^T \tag{63}$$

$$\widetilde{\mathbf{R}}_k = \mathbf{B}_{r,k} \widetilde{\mathbf{C}}_{y,k}^{-1} \mathbf{B}_{r,k}^T \tag{64}$$

$$\widetilde{\mathbf{C}}_{x,k} = diag(G_\sigma(\widetilde{e}_{1,k}), \dots, G_\sigma(\widetilde{e}_{n,k})) \tag{65}$$

$$\widetilde{\mathbf{C}}_{y,k} = diag(G_\sigma(\widetilde{e}_{n+1,k}), \dots, G_\sigma(\widetilde{e}_{n+m,k})) \tag{66}$$

$$\widetilde{e}_{i,k} = d_{i,k} - \mathbf{w}_{i,k} \hat{\mathbf{x}}_{k|k,t-1} \tag{67}$$

$$\widetilde{\mathbf{R}}_k = \mathbf{R}_k \left(\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + R_k \right)^{-1} \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k \mathbf{H}_k^T \tag{68}$$

$$\widetilde{\mathbf{Q}}_k = \mathbf{K}_k \left(\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + R_k \right) \mathbf{K}_k^T \tag{69}$$

(6) Compare the estimation of the current step and the estimation of the last step. If (70) holds, let $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k,t}$ and continue to 7). Otherwise, $t + 1 \rightarrow t$, and go back to 5);

$$\frac{\left\| \hat{\mathbf{x}}_{k|k,t} - \hat{\mathbf{x}}_{k|k,t-1} \right\|}{\left\| \hat{\mathbf{x}}_{k|k,t-1} \right\|} \leq \varepsilon \tag{70}$$

(7) Moreover, the posterior matrix is updated as (71), $k + 1 \rightarrow k$ and go back to 2).

$$\mathbf{P}_{k|k} = \left(\mathbf{I} - \widetilde{\mathbf{K}}_k \mathbf{H}_k \right) \mathbf{P}_{k|k-1} \left(\mathbf{I} - \widetilde{\mathbf{K}}_k \mathbf{H}_k \right)^T + \widetilde{\mathbf{K}}_k \mathbf{R}_k \widetilde{\mathbf{K}}_k^T \tag{71}$$

5. Results

In this section, we perform experiments on the standard IEEE 14-bus and IEEE 30-bus test system to verify the effectiveness and superiority of the proposed algorithm compared with the EKF, UKF, A-EKF and MCC-EKF algorithms. We use the 50 time-sample intervals, which were obtained by running successful load flows under different loading conditions to simulate the slow dynamics of the power system. The variation of loads can be divided into linear and nonlinear variation, among which the linear variation means the whole observed time interval with 50 samples are changes follow a linear trend of 10% and the nonlinear variation denotes the whole observed time interval with 50 samples changes follow a random fluctuation of 5%. The convergence tolerance threshold of all algorithms is 0.001. In addition, all free parameters of the algorithms mentioned above are selected such that each algorithm can achieve its optimal performance.

The mean absolute error (MAE) and root mean square error (RMSE) are utilized to evaluate the performance of each method. In addition, the overall performance is one of the indices for evaluating the performance of the state estimation algorithms, and it can be defined by

$$J_i = \frac{\sum\limits_{i=1}^{N} \left| \hat{y}_k^i - \overline{y}_k^i \right|}{\sum\limits_{i=1}^{N} \left| y_k^i - \overline{y}_k^i \right|} \tag{72}$$

where $\hat{y}_k^i$ is the estimated measurement vector at time index k for bus i, $\overline{y}_k^i$ denotes the true measurements and y_k^i represents the measurement vector with noises. The MAE, RMSE and the over performance will be used to verify the available of the proposed method in the following experiments.

5.1. Case 1: Gaussian Measurement Noise Environment

In general, the measurement noise is assumed for Gaussian noise, and we first evaluate the proposed algorithm in the normal environment. Loads are changed according to the above specified linear trend. The average overall performance of all algorithms in standard IEEE 30-bus is given in Table 1. From the results, we know that (1) both EKF, UKF and A-EKF algorithms have good performance in this case; (2) the adaptive mechanism can improve the estimation accuracy; (3) the performance of MCC-EKF is slightly better than that of EKF and UKF and the proposed AMCC-EKF performs better than other algorithms which illustrate that on the one hand, the correntropy is robust for the noise, on the other hand, the adaptive mechanism The adaptive mechanism makes the algorithm have stronger tracking performance by updating the predicted error covariance matrix and measurement noise covariance matrix.

Table 1. The average overall performance of all algorithms in standard IEEE 30-bus.

	EKF	UKF	A-EKF	MCC-EKF	AMCC-EKF
Index J (p.u.)	0.39	0.29	0.31	0.25	0.16

5.2. Case 2: Gaussian Mixture Measurement Noise Environment

In this case, we evaluate the performance of the proposed algorithm under the non-Gaussian measurement noise circumstance. The measurement of noise is modeled by the mixed Gaussian distribution which is defined as

$$(1 - \theta)N(\mu_1, v_1^2) + \theta N(\mu_2, v_2^2) \tag{73}$$

where $N(\mu_i, v_i^2)(i = 1, 2)$ denotes the Gaussian distributions with mean values μ_i and variances σ_i^2, and the θ is the mixture coefficient. In this simulation, the mean values μ_1 and μ_2 both are set at zero, the variances σ_1^2 and σ_2^2 are set at 1 and 80 respectively, and the mixture coefficient θ is set at 0.25. The overall performance of all algorithms in standard IEEE 14-bus and IEEE 30-bus test system are shown in Figures 2 and 3, respectively. One can observe that the performance of EKF, UKF and A-EKF is significantly worse under the Gaussian mixture measurement noise environment, while the MCC-EKF has good performance because of it is insensitive to the non-Gaussian noise. Moreover, we see that the performance of the proposed AMCC-EKF is better than the original MCC-EKF algorithm. In addition, the average overall performance of all algorithms in standard IEEE 14-bus and 30-bus systems are given in Tables 2 and 3, respectively.
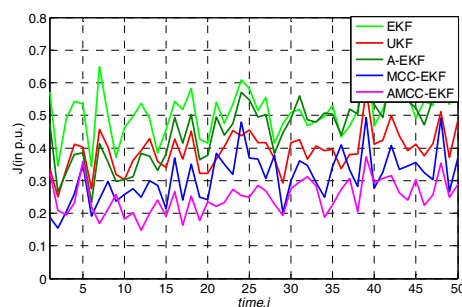

Figure 2. The overall performance of all algorithms in standard IEEE 14-bus under Gaussian mixture measurement noise environment.
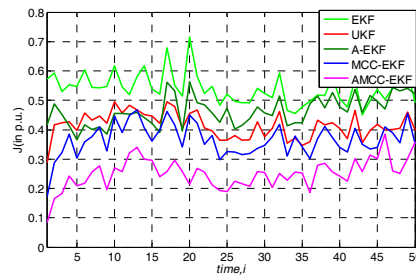
Figure 3. The overall performance of all algorithms in standard IEEE 30-bus under Gaussian mixture measurement noise environment.

Table 2. The average overall performance of all algorithms in standard IEEE 14-bus under Gaussian mixture measurement noise environment.

	EKF	**UKF**	**A-EKF**	**MCC-EKF**	**AMCC-EKF**
Index J (p.u.)	0.53	0.41	0.49	0.33	0.23

Table 3. The average overall performance of all algorithms in standard IEEE 30-bus under Gaussian mixture measurement noise environment.

	EKF	**UKF**	**A-EKF**	**MCC-EKF**	**AMCC-EKF**
Index J (p.u.)	0.55	0.41	0.48	0.32	0.24

In addition, we further analyze the tracking performance of the algorithm from the perspective of accurately estimating the voltage amplitude and phase angle of bus at each time. Specifically, loads are changed following a linear trend of 10%. The true voltage amplitude and voltage angle of no.3 bus in IEEE 30-bus test system and the estimated values of each algorithms are shown in Figures 4 and 5, respectively. We know that the estimated value of the proposed AMCC-EKF algorithm is close to the true value than other algorithms.
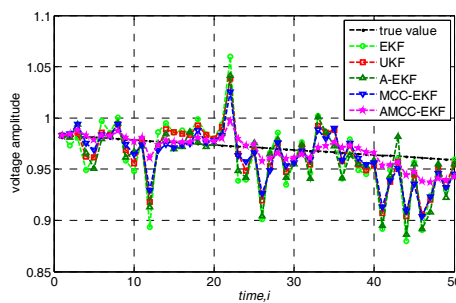
Figure 4. The true voltage amplitude of no.3 bus in IEEE 30-bus test system.
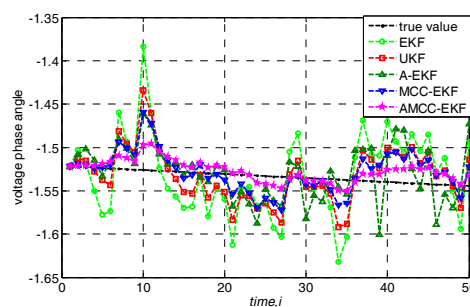
Figure 5. The voltage angle of no.3 bus in IEEE 30-bus test system.

This result proves that the AMCC-EKF has both a higher accuracy and filtering capacities than the corresponding EKF, UKF, A-EKF and MCC-EKF algorithm under the Gaussian mixture measurement noise environment.

5.3. Case 3: Laplace and Gaussian Mixture Measurement Noise Environment

In this case, we test the universality of the proposed algorithm under the mixed Gaussian and Laplace noises environment. The noise model can be represented as

$$v(n) = (1 - a(n))A(n) + a(n)B(n) \tag{74}$$

where $a(n)$ is an independent and identically distributed binary process with an occurrence probability $0 \leq c \leq 1$. In this simulation, c is set at 0.5, $A(n)$ is a noise obey the Laplace distribution with zero-mean and unit variance, and $B(n)$ denotes another noise process with Gaussian distribution with zero-mean and variance 0.55. The noise processes $A(n)$ and $B(n)$ are mutually independent and they are both independent of $a(n)$. The obtained overall performance of all algorithms in IEEE 14-bus and IEEE 30-bus test system are displayed in Figures 6 and 7, respectively. One can see that the EKF, UKF and A-EKF still show worse performance in presence of the Laplace and Gaussian mixture noise. In addition, we can obtain the same conclusion with the Section 5.2 that the MCC based methods (MCC-EKF and AMCC-EKF) are robustness in this case, and the proposed AMCC-EKF algorithm can achieve the best performance. The average overall performance of all algorithms in standard IEEE 14-bus and 30-bus system are shown in Tables 4 and 5, respectively. The results illustrate the outstanding properties of the AMCC-EKF for SE again.
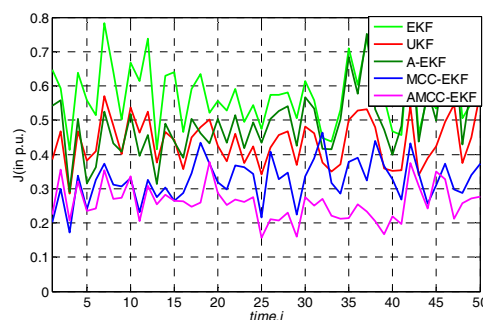

Figure 6. The overall performance of all algorithms in standard IEEE 14-bus under Laplace and Gaussian mixture measurement noise environment.
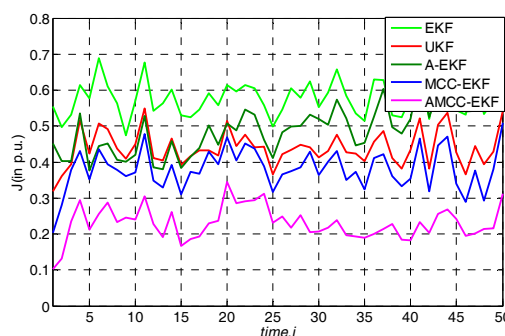

Figure 7. The overall performance of all algorithms in standard IEEE 30-bus under Laplace and Gaussian mixture measurement noise environment.

Table 4. The average overall performance of all algorithms in standard IEEE 30-bus under Laplace and Gaussian mixture measurement noise environment.

	EKF	UKF	A-EKF	MCC-EKF	AMCC-EKF
Index J (p.u.)	0.65	0.48	0.53	0.36	0.28

Table 5. The average overall performance of all algorithms in standard IEEE 30-bus under Laplace and Gaussian mixture measurement noise environment.

	EKF	UKF	A-EKF	MCC-EKF	AMCC-EKF
Index J (p.u.)	0.59	0.52	0.46	0.38	0.23

Furthermore, to test the effect of load change on the proposed algorithm, we changed the linear variation trend of load and consider 10%, 20%, 30%, respectively. The mean absolute error and root mean square error of voltage angle of no.3 bus in IEEE 30-bus in different variation trend of load are shown in Figures 8 and 9, respectively. It can be seen from the bar chart that although the error of the proposed algorithm increases with the linear variation trend of the load, the error of the proposed algorithm is minimal compared with other methods.
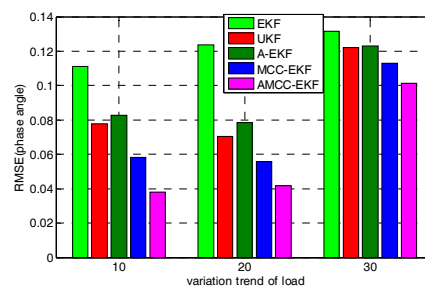
Figure 8. The mean absolute error of voltage angle of no.3 bus in IEEE 30-bus in different variation trend of load.
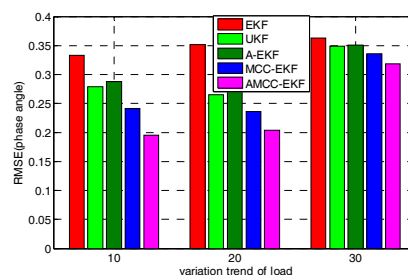

Figure 9. The root mean square error of voltage angle of no.3 bus in IEEE 30-bus in different variation trend of load.

5.4. Case 4: the Nonlinear Variation of Loads

Now, from the analysis above, we consider this case that the loads change follows a random fluctuation of 5%. The other settings are the same as those of Case 2. The true value of voltage amplitude and angle of no.3 bus in IEEE 30-bus test system and estimated value of other algorithms are shown in Figures 10 and 11, respectively. It can be seen from the figure that even in the case of nonlinear variation of loads, the estimated value of the proposed AMCC-EKF algorithm is closest to the true state value.
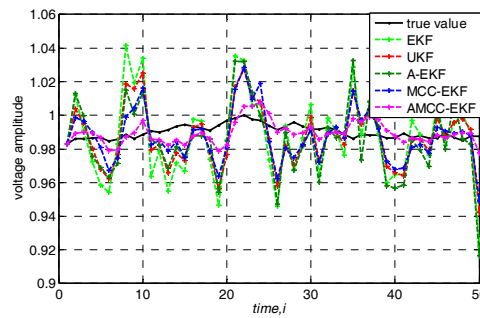
Figure 10. The true value of voltage amplitude of no.3 bus in IEEE 30-bus test system and estimated value of other algorithms when the loads change follows a random fluctuation.
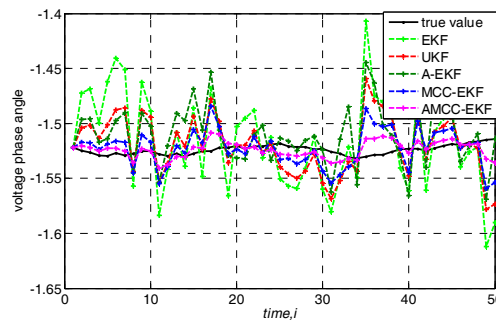

Figure 11. The true value of voltage angle of no.3 bus in IEEE 30-bus test system and estimated value of other algorithms when the loads change follows a random fluctuation.

To further demonstrate the superiority of the algorithm numerically, the MAE and RMSE of voltage amplitude of no.3 bus in IEEE 30-bus are shown in Table 6, and the results of the proposed AMCC-EKF algorithm is better than other algorithms.

Table 6. The MAE and RMSE of voltage amplitude of no.3 bus in IEEE 30-bus.

	EKF	UKF	A-EKF	MCC-EKF	AMCC-EKF
MAE	0.06	0.05	0.05	0.03	0.01
RMSE	0.25	0.21	0.23	0.18	0.12

5.5. Case 5: in Presence of Outliers

Now, measurement in the presence of outliers at sample time 30 is considered in this case and assuming that no measures are taken to verify and identify the bad data under Gaussian mixture measurement noise environment. In this case, four state estimation algorithms are adopted for simulation, resulting in the overall performance index changes of state estimation filter as shown in Figure 12. Firstly, we know that all algorithms are affected by outliers. In the presence of outliers, the filtering performance index will increase and the estimation accuracy will decrease. Secondly, the EKF, UKF, A-EKF and MCC-EKF algorithms are greatly affected by bad data at sample time 30. Although the proposed AMCC-EKF algorithm is also affected to some extent, but the filtering average performance index still fluctuates below 0.33, and the filtering estimation value is relatively correct. The average overall performance of all algorithms in standard IEEE 30-bus are given in Table 7.
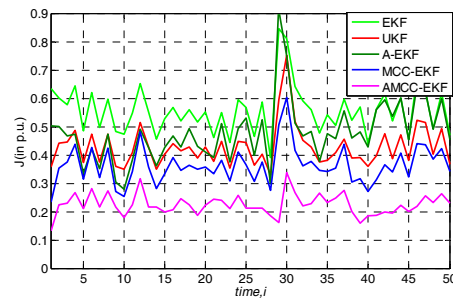
Figure 12. The overall performance of all algorithms in standard IEEE 30-bus in presence of outliers.

Table 7. The average overall performance of all algorithms in standard IEEE 30-bus in presence of outliers.

	EKF	UKF	A-EKF	MCC-EKF	AMCC-EKF
Index J (p.u.)	0.56	0.46	0.48	0.36	0.24

6. Conclusions

In this paper, a novel AMCC-EKF algorithm is developed to address the power system state estimation problem, and its effectiveness and robustness are verified by some scenarios under non-Gaussian noise environments. First, the actual power system is very susceptible to non-Gaussian noise and the MCC-EKF is employed to design a robust state estimation approach at first. Second, it is difficult to calculate the noise statistics in the most actual situation, and hence the adaptive MCC-EKF (AMCC-EKF) is proposed by introducing the adaptive mechanism into the MCC-EKF algorithm to continuously update the covariance matrix to improve the accuracy of the estimation results. We perform experiments on the IEEE 14 and 30 bus systems to test the performance of the proposed AMCC-EKF method, and the simulation results demonstrate that the filtering performance and estimation accuracy of the AMCC-EKF algorithm for state estimation is better than EKF, UKF, A-EKF and MCC-EKF methods and the estimation results are relatively stable in the presence of outliers.

Author Contributions: Z.Z. proposed the main idea; J.Q. wrote the draft and derived the algorithm; W.M. was in charge of technical checking.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. Stott, B.; Alsac, O.; Monticelli, A.J. Security analysis and optimization. *Proc. IEEE* **1987**, *75*, 1623–1644. [CrossRef]
2. Shivakumar, N.R.; Jain, A. A review of power system dynamic state estimation techniques. In Proceedings of the Joint International Conference on Power System Technology & IEEE Power India Conference, New Delhi, India, 12–15 October 2008; pp. 1–6.
3. Gomez-Exposito, A.; Conejo, A.J.; Canizares, C. *Electric Energy Systems: Analysis and Operation*; CRC Press: Boca Raton, FL, USA, 2018.
4. Green, R.C.; Wang, L.; Alam, M. Applications and trends of high performance computing for electric power systems: Focusing on smart grid. *IEEE Trans. Smart Grid* **2013**, *4*, 922–931. [CrossRef]

5. Li, W.; Li, J.; Gao, A. Review and research trends on state estimation of electrical power systems. In Proceedings of the Power and Energy Engineering Conference (APPEEC), 2011 Asia-Pacific, New York, NY, USA, 25–28 March 2011; pp. 1–4.

6. Ljung, L. Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. *IEEE Trans. Autom. Control* **2013**, *24*, 36–50. [CrossRef]

7. Ghahremani, E.; Kamwa, I. Dynamic state estimation in power system by applying the extended Kalman filter with unknown inputs to phasor measurements. *IEEE Trans. Power Syst.* **2011**, *26*, 2556–2566. [CrossRef]

8. Maybeck, P.S.; Jensen, R.L.; Harnly, D.A. An adaptive extended Kalman filter for target image tracking. *IEEE Trans. Aerosp. Electron. Syst.* **1981**, *2*, 173–180. [CrossRef]

9. Netto, M.; Zhao, J.; Mili, L. A robust extended Kalman filter for power system dynamic state estimation using PMU measurements. In Proceedings of the Power and Energy Society General Meeting (PESGM), Boston, MA, USA, 17–21 July 2016; pp. 1–5.

10. Gandhi, M.A.; Mili, L. Robust Kalman filter based on a generalized maximum-likelihood-type estimator. *IEEE Trans. Signal Process.* **2010**, *58*, 2509–2520. [CrossRef]

11. Zhao, J.B.; Netto, M.; Mili, L. A robust iterated extended Kalman filter for power system dynamic state estimation. *IEEE Trans. Power Syst.* **2017**, *32*, 3205–3216. [CrossRef]

12. Zhao, J.B. Dynamic State Estimation With Model Uncertainties Using H∞ Extended Kalman Filter. *IEEE Trans. Power Syst.* **2018**, *33*, 1099–1100. [CrossRef]

13. Wan, E.A.; Van, D.M.R. The unscented Kalman filter for nonlinear estimation. In Proceedings of the Adaptive Systems for Signal Processing, Communications, and Control Symposium, Lake Louise, AB, Canada, 4 October 2000; pp. 153–158.

14. Valverde, G.; Terzija, V. Unscented Kalman filter for power system dynamic state estimation. *IET Gener. Transm. Distrib.* **2011**, *5*, 29–37. [CrossRef]

15. Liu, W.; Puskal, P.; Jose, C. Principe Correntropy: Properties and Applications in Non-Gaussian Signal Processing. *IEEE Trans. Signal Process.* **2007**, *55*, 5286–5298. [CrossRef]

16. Chen, B.; Xing, L.; Liang, J. Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion. *IEEE Signal Process. Lett.* **2014**, *21*, 880–884.

17. Chen, B.; Wang, J.; Zhao, H.; Zheng, N. Convergence of a Fixed-Point Algorithm under Maximum Correntropy Criterion. *IEEE Signal Process. Lett.* **2015**, *22*, 1723–1727. [CrossRef]

18. Chen, B.; Príncipe, J.C. Maximum correntropy estimation is a smoothed MAP estimation. *IEEE Signal Process. Lett.* **2012**, *19*, 491–494. [CrossRef]

19. Chen, B.; Xing, L.; Zhao, H. Robustness of Maximum Correntropy Estimation Against Large Outliers. *arXiv* **2017**, arXiv:1703. 08065.

20. Chen, B.; Liu, X.; Zhao, H. Maximum correntropy Kalman filter. *Automatica* **2017**, *76*, 70–77. [CrossRef]

21. Liu, X.; Qu, H.; Zhao, J. Extended Kalman filter under maximum correntropy criterion. In Proceedings of the International Joint Conference on Neural Networks, Vancouver, BC, Canada, 24–29 July 2016; pp. 1733–1737.

22. Ma, W.; Chen, B.; Qu, H. Sparse least mean p-power algorithms for channel estimation in the presence of impulsive noise. *Signal Image Video Process.* **2016**, *10*, 503–510. [CrossRef]

23. Wang, S.; Dang, L.; Wang, W.; Qian, G. Kernel adaptive filters with feedback based on maximum correntropy. *IEEE Access* **2018**, *6*, 10540–10552. [CrossRef]

24. Dees, B.S.; Xia, Y.; Douglas, S.C. Correntropy-Based Adaptive Filtering of Noncircular Complex Data. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018.

25. Wang, G.; Gao, Z.; Zhang, Y. Adaptive Maximum Correntropy Gaussian Filter Based on Variational Bayes. *Sensors* **2018**, *18*, 1960. [CrossRef]

26. Xian, Z.; Hu, X.; Lian, J. Robust innovation-based adaptive Kalman filter for INS/GPS land navigation. In Proceedings of the Chinese Automation Congress IEEE, Changsha, China, 7–8 November 2013; pp. 374–379.

27. Hide, C.; Moore, T. Adaptive Kalman filtering for low-cost INS/GPS. *J. Navig.* **2003**, *56*, 143–152. [CrossRef]

28. Haykin, S.S. *Adaptive Filter Theory*; Publishing House of Electronics Industry: Beijing, China, 2002.

29. Elkeyi, A.; Kirubarajan, T.; Gershman, A.B. Robust adaptive beamforming based on the Kalman filter. *IEEE Signal Process Soc.* **2005**, *53*, 3032–3041. [CrossRef]

30. Conejo, A.J. Power System State Estimation-Theory and Implementations. *IEEE Power Energy Mag.* **2005**, *3*, 64–65. [CrossRef]

31.   Leite, A.M.; Silva, M.B.D.; Queiroz, J.F. State forecasting in electric power systems. *IEE Proc. C Gener. Transm. Distrib.* **1983**, *130*, 237–244. [CrossRef]