Robust power system state estimation using t-distribution noise model

Chen, Tengpeng; Sun, Lu; Ling, Keck-Voon; Ho, Weng Khuen

2019

Chen, T., Sun, L., Ling, K.-V., & Ho, W. K. (2020). Robust power system state estimation using t-distribution noise model. IEEE Systems Journal, 14(1), 771-781.

doi:10.1109/JSYST.2018.2890106

<https://hdl.handle.net/10356/137155>

<https://doi.org/10.1109/JSYST.2018.2890106>

© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at:
<https://doi.org/10.1109/JSYST.2018.2890106>.

Downloaded on 26 Apr 2023 19:22:32 SGT

Robust Power System State Estimation Using t -Distribution Noise Model

Tengpeng Chen, Lu Sun, KV Ling, and W.K. Ho

Abstract—In this paper we propose an optimal robust state estimator using maximum likelihood optimization with the t -distribution noise model. In robust statistics literature, the t -distribution is used to model Gaussian and non-Gaussian statistics. The influence function, an analytical tool in robust statistics, is employed to obtain the solution to the resulting maximum likelihood estimation optimization problem so that the proposed estimator can be implemented within the framework of traditional robust estimators. Numerical results obtained from simulations of the IEEE 14-bus system, IEEE 118-bus system and experiment on a microgrid demonstrated the effectiveness and robustness of the proposed estimator. The proposed estimator could suppress the influence of outliers with smaller Average Mean Squared Errors (AMSE) than the traditional robust estimators such as Quadratic-Linear (QL), Square-Root (SR), Schweppe-Huber Generalized-M (SHGM), Multiple-Segment (MS) and Least Absolute Value (LAV) estimator. A new approximate AMSE formula is also derived for the proposed estimator to predict and evaluate its precision.

Index Terms—Robust state estimation; t -distribution; Maximum likelihood estimation; Influence function.

NOMENCLATURE

The notations used in the paper are summarized below for easy reference.

x	true state vector
$\hat{x}$	estimated state vector
$\hat{x}_{MLE}$	estimated state vector obtained by maximum likelihood estimation
$\hat{x}_{Tay}$	estimated state vector using Taylor series expansion
$\hat{x}_{IF}$	estimated state vector using influence function approximation
$\bar{x}$	operating point
V_i^r	real part of the voltage phasor at bus i

V_i^{im}	imaginary part of the voltage phasor at bus i
z	measurements
ϵ	measurement noise
σ_i	the i -th standard deviation of measurement noise ϵ_i
k	time instance
$e_i(k)$	the i -th measurement residual at time instance k
ρ	ρ -function
J	total cost function
$f_i(\epsilon_i)$	probability density function of measurement noise ϵ_i
H	measurement matrix
H_i	the i -th row of measurement matrix
$h(x)$	nonlinear function relating the state vector x
$h_i(x)$	nonlinear function relating the state vector x to measurement i
$w_i(k)$	weighting factor for measurement i at time k
Ψ	Ψ -function
$IF(\cdot)$	influence function
$\Gamma(\cdot)$	gamma function
m	number of measurements in 1 batch
n	number of states
N	number of batches
Nm	total measurement numbers for estimation
i	measurement index
a_i	estimator first threshold parameter for e_i
b_i	estimator second threshold parameter for e_i
r_i	estimator third threshold parameter for e_i
Ω, Λ	diagonal matrix
v_i, ξ_i	the t -distribution parameters

I. INTRODUCTION

Power system state estimation (PSSE) is one of the key functions in energy management system (EMS). The Gaussian noise assumption is commonly made in PSSE but this assumption is only an approximation to reality [1]. For example, there are studies in power systems where the Gaussian noises are corrupted by outliers with uniform [2] or Gaussian distribution [3]. The recent introduction of Phasor Measurement Unit (PMU) [4]–[6] makes possible the measurements of voltages and currents in a synchronized manner with respect to the Global Positioning System and the measurement model becomes linear. However, the PMU measurements are known to experience sudden changes or may become unavailable due to communication system malfunction or noise [7]. This would lead to occurrence of outliers and then generate non-Gaussian measurements errors [8]. Outliers that are far away from the expected Gaussian distribution function can give rise

Manuscript received [date to be filled]. This work was supported in part by Fujian Education and Scientific Research Funds for Middle-aged and Young Teachers under Grant JT180001, and in part by the Singapore National Research Foundation (NRF) under its Campus for Research Excellence And Technological Enterprise (CREATE) programme, and Cambridge Centre for Advanced Research in Energy Efficiency in Singapore (CARES), C4T project. (Corresponding author: Lu Sun.)

Tengpeng Chen is with the Department of Instrumental and Electrical Engineering, Xiamen University, Xiamen, China 361102. (e-mail: tpchen@xmu.edu.cn).

Lu Sun and W.K. Ho are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576. (e-mail: sunlu@nus.edu.sg, wk.ho@nus.edu.sg).

KV Ling is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798. (e-mail: ekvling@ntu.edu.sg).

to misleading estimation results [9].

The Weighted Least Squares (WLS) estimator is widely used but it is not robust. Therefore, the largest normalized residuals (LNR) test-based method [10] is usually used in conjunction with WLS. The negative impact of the outliers on the state estimates and several ways to reduce them have been discussed in [11] where the authors propose a new approach to screen and suppress the gross error in measurements by iterations. Following this approach many traditional M-estimators with non-quadratic cost functions such as Quadratic-Linear (QL), Square-Root (SR), Schweppe-Huber Generalized-M (SHGM), Multiple-Segment (MS) and Least Absolute Value (LAV) have been discussed in standard textbooks [12], [13] and papers [5], [14]–[17] where Gaussian noise assumption is commonly made.

In this paper, we propose an optimal robust state estimator using maximum likelihood optimization with the t -distribution noise model. It is well known that the t -distribution reduces to Gaussian distribution when the shape parameter tends to infinity [8]. Therefore, the t -distribution can be used to model both Gaussian and non-Gaussian noise distribution. Moreover, the t -distribution is commonly used in robust state estimation literature [18]–[20] to model noise with outliers. In Example 1, we use the t -distribution to model the Gaussian noise corrupted by outliers with uniform [2] or Gaussian distribution [3].

The main contributions of this paper are as follows:

- 1) Instead of the usual Gaussian noise distribution, the use of t -distribution in the proposed estimator could suppress the influence of outliers and provides smaller Average Mean Squared Errors (AMSE) than the traditional robust estimators mentioned above.
- 2) The Influence Function (IF) is employed to give the solution to the optimization problem arising from maximum likelihood estimation (MLE) so that the proposed estimator with t -distribution noise can be efficiently implemented within the framework of traditional robust estimators.
- 3) A new approximate AMSE formula is also derived for the proposed estimator.

Recently many new robust estimators are proposed [18]–[28]. However, to obtain the statistics such as the AMSE or variances of their state estimates, hundreds of simulation have to be run [29]. In contrast, for the proposed estimator, because of the influence function approximation, AMSE or variances can be obtained from an analytical formula. The formula as a mathematical function is more insightful than just a numerical answer from simulation. For example, in estimator design, instead of trial and error through simulations, the formula allows us to predict and evaluate its precision in terms of AMSE or variances for different choices of estimator parameter values leading to a more efficient and effective estimator design process.

The rest of paper is organized as follows. The robust

state estimation problem is formulated in Section II and the proposed estimator using IF approximation is discussed in Section III. Numerical results obtained from simulations of the IEEE 14-bus system, IEEE 118-bus system and experiment on a microgrid are presented in Section IV to verify the effectiveness of the proposed estimator in terms of AMSE and computational time. Conclusions are given in Section V.

II. ROBUST STATE ESTIMATION

Robust state estimation has been discussed elsewhere in [2], [5], [12], [17]. This section focuses on the equations necessary for the derivation of the results in this paper.

A. Measurement model

Consider the measurement model in power system:

$$z(k) = h(x(k)) + \epsilon(k) \quad (1)$$

where

$$\begin{aligned} z(k) &= [z_1(k), z_2(k), \dots, z_m(k)]^T \\ x(k) &= [x_1(k), x_2(k), \dots, x_n(k)]^T \\ h(x(k)) &= [h_1(x(k)), h_2(x(k)), \dots, h_m(x(k))]^T \\ \epsilon(k) &= [\epsilon_1(k), \epsilon_2(k), \dots, \epsilon_m(k)]^T \end{aligned}$$

The nonlinear case is applicable when the measurements are collected from SCADA. Measurement $z_i(k)$, $i = 1, \dots, m$, is taken at each time instance $k = 1, \dots, N$. $\epsilon_i(k)$ is the measurement noise and the rectangular coordinates of the bus voltage phasor, $x(k)$, is the states of the system. A traditional power system may be considered as a quasi-static system [30] because load demands change slowly and hence the state changes slowly. The typical measurement sampling interval is 2 to 4 seconds while the estimates are usually updated only once every few minutes [31]. Hence the estimated states during the time instants, $k = 1, \dots, N$ can be taken as $\hat{x}$.

The first-order Taylor series expansion about the operating point $\bar{x}$ is used to approximate the nonlinear function $h(x(k))$ in (1) to give

$$z(k) \approx h(\bar{x}) + H(x(k) - \bar{x}) + \epsilon(k) \quad (2)$$

where the Jacobian matrix

$$H = \left. \frac{\partial h(x)}{\partial x} \right|_{x=\bar{x}} \quad (3)$$

The i -th measurement residual is defined as

$$e_i(k) = z_i(k) - h_i(\hat{x}) \quad (4)$$

where $h_i(\hat{x})$ is approximated as $h_i(\bar{x}) + H_i(\hat{x} - \bar{x})$ and $e(k)$ is a vector, $e(k) = [e_1(k), e_2(k), \dots, e_m(k)]^T$.

When the entire measurements are collected from enough PMUs and we assume that the system is observable, model (1) reduces to [5], [32]

$$z(k) = Hx(k) + \epsilon(k) \quad (5)$$

where H is the measurement matrix. In this case the i -th measurement residual is

$$e_i(k) = z_i(k) - H_i \hat{x} \quad (6)$$

B. Traditional robust state estimation

Given N sets of m measurements, the state estimate $\hat{x}$ is obtained by minimizing the following cost function:

$$J = \sum_{i=1}^m \sum_{k=1}^N \rho(e_i(k)) \quad (7)$$

where ρ -function is a nonlinear chosen function of the residual $e_i(k)$.

Differentiating (7) wrt $\hat{x}$ gives

$$\begin{aligned} \frac{\partial J}{\partial \hat{x}} &= \frac{\partial J}{\partial e_i(k)} \frac{\partial e_i(k)}{\partial \hat{x}} \\ &= \sum_{i=1}^m \sum_{k=1}^N \frac{\partial \rho(e_i(k))}{\partial e_i(k)} \frac{1}{e_i(k)} e_i(k) \frac{\partial e_i(k)}{\partial \hat{x}} \\ &= - \sum_{i=1}^m \sum_{k=1}^N w_i(k) e_i(k) H_i^T \end{aligned} \quad (8)$$

where $(\cdot)^T$ represents the transpose operation and

$$\begin{aligned} w_i(k) &= \frac{\partial \rho(e_i(k))}{\partial e_i(k)} \frac{1}{e_i(k)} \\ \frac{\partial e_i(k)}{\partial \hat{x}} &= -H_i^T \end{aligned} \quad (9)$$

Using (6), equation (8) can be written as

$$\begin{aligned} \frac{\partial J}{\partial \hat{x}} &= - \sum_{i=1}^m \sum_{k=1}^N w_i(k) (z_i(k) - H_i \hat{x}) H_i^T \\ &= -\bar{H}^T W (Z - \bar{H} \hat{x}) \\ &= -\bar{H}^T W E \end{aligned} \quad (10)$$

where

$$\bar{H} = [H^T \dots H^T]^T \in \mathbb{R}^{Nm \times n} \quad (11)$$

$$Z = [z(1)^T \dots z(N)^T]^T \in \mathbb{R}^{Nm} \quad (12)$$

$$\begin{aligned} E &= Z - \bar{H} \hat{x} = [e(1)^T \dots e(N)^T]^T \in \mathbb{R}^{Nm} \\ W &= \text{diag}(w_1(1), \dots, w_m(1), \dots, w_1(N), \dots, w_m(N)) \\ &\in \mathbb{R}^{Nm \times Nm} \end{aligned}$$

When the measurement model is nonlinear, linearization is proceeded using (3) and (4). Moreover, $z(k)$ in (11) is replaced by $\tilde{z}(k) = (z(k) - h(\bar{x}) + H\bar{x})$, $k = 1, \dots, N$.

To minimize the cost function in (7), set $\frac{\partial J}{\partial \hat{x}} = 0$ in (10) and the solution for $\hat{x}$ is given by

$$\hat{x} = (\bar{H}^T W \bar{H})^{-1} \bar{H}^T W Z \quad (13)$$

Notice that the matrix $\bar{H}^T W \bar{H}$ is an invertible matrix when the system is assumed to be observable. For the WLS estimator, the element of diagonal matrix W is constant, i.e.,

the reciprocal of the error variance for measurements, $1/\sigma_i^2$. Note that σ_i is the standard deviation of noise ϵ_i . The largest normalized residuals (LNR) method [10] is usually used in WLS to deal with bad data. The normalized residuals are calculated as:

$$\begin{aligned} \bar{R} &= \text{diag}(\sigma_1^2, \dots, \sigma_m^2, \dots, \sigma_1^2, \dots, \sigma_m^2) \\ \bar{G} &= \bar{H}^T \bar{R}^{-1} \bar{H} \\ \bar{\Omega} &= \bar{R} - \bar{H} \bar{G}^{-1} \bar{H}^T \\ e_i^{norm}(k) &= \frac{|e_i(k)|}{\sqrt{\bar{\Omega}_{ii}}} \end{aligned}$$

The normalized residuals $e_i^{norm}(k)$ are calculated according to the residual covariance matrix $\bar{\Omega}$ and measurement residual $e_i(k)$. If the normalized residuals $e_i^{norm}(k)$ are larger than a pre-determined threshold, the largest one will correspond to the bad measurement Z_i^{bad} . Once the largest normalized residual is found, corresponding measurement is updated:

$$Z_i^{new} = Z_i^{bad} - \frac{\bar{R}_{ii}}{\bar{\Omega}_{ii}} e_i^{bad}(k) \quad (14)$$

The states will then be recalculated based on the updated measurements Z_i^{new} . Several iterations may be needed in order to make sure that all normalized residuals are less than pre-determined threshold, for example, 3.0 [5].

The matrix W in (13) for the MS, QL, SR and SHGM estimator is a re-weighted matrix and is updated according to Table I [12], [29].

TABLE I: The weighting matrices of robust estimators.

Estimator	$w_i(k)$	$e_i(k)$
MS	$\frac{1}{\sigma_i^2}$	$ e_i(k) \leq a_i \sigma_i$
	$\frac{a_i e_i(k) }{\sigma_i (r_i - b_i) \sigma_i^2 e_i(k) }$	$a_i \sigma_i < e_i(k) \leq b_i \sigma_i$
	0	$b_i \sigma_i < e_i(k) \leq r_i \sigma_i$
QL	$\frac{1}{\sigma_i^2}$	$ e_i(k) \leq a_i \sigma_i$
	0	otherwise
SR	$\frac{1}{\sigma_i^2}$	$ e_i(k) \leq a_i \sigma_i$
	$\sqrt{\frac{a_i^3}{\sigma_i e_i(k) ^3}}$	otherwise
SHGM	$\frac{1}{\sigma_i^2}$	$ e_i(k) \leq a_i \kappa_i \sigma_i$
	$\frac{a_i \kappa_i}{\sigma_i e_i(k) }$	otherwise

The parameters a_i, b_i and r_i are thresholds, and κ_i is the penalty factor chosen specifically to cancel the effect of any existing leverage points in the measurement set. Note that $a_i < b_i < r_i$. The QL, WLS and LAV estimators are special cases of MS estimator. The MS estimator reduces to the QL estimator when $b_i \rightarrow \infty$ and the WLS estimator when $a_i \rightarrow \infty$. The QL estimator reduces to LAV estimator when the threshold a_i is small [12].

Using (12), replace Z by $\bar{H} \hat{x} + E$ and (13) becomes

$$\hat{x} = (\bar{H}^T W \bar{H})^{-1} \bar{H}^T W (\bar{H} \hat{x} + E)$$

which can be solved iteratively as

$$\hat{x} \leftarrow \hat{x} + (\bar{H}^T W \bar{H})^{-1} \bar{H}^T W E \quad (15)$$

III. THE PROPOSED ESTIMATOR

In this section, a new robust estimator based on MLE is proposed where the t -distribution is used to model noise with outliers. Moreover, the IF is used to solve the optimization problem arising from MLE so that the proposed estimator can be easily implemented within the framework of the traditional robust estimators.

A. t -distribution

In the proposed state estimation algorithm, the following t -distribution probability density function (pdf) [18] is used to model measurement noise corrupted by outliers.

$$f_i(\epsilon_i) = \frac{\Gamma(\frac{v_i+1}{2})}{\sqrt{v_i\pi}(\xi_i)\Gamma(\frac{v_i}{2})} \left(1 + \frac{|\epsilon_i|^2}{(\xi_i)^2 v_i}\right)^{-\frac{v_i+1}{2}} \quad (16)$$

where $\Gamma(\cdot)$ is the gamma function, ξ_i is the scale parameter, and v_i is the shape parameter. It is well known that the t -distribution reduces to Gaussian distribution when the shape parameter tends to infinity [8], [18]. Therefore, the t -distribution can be used to model both Gaussian or non-Gaussian noise distribution. Moreover, the t -distribution is commonly used in robust state estimation literature [18]–[20] to model and study outliers. The t -distribution with $\xi_i = 0.01$ and different v_i are shown in Fig. 1. When the shape parameter v_i tends to ∞ , the t -distribution reduce to a Gaussian distribution (36) with $\sigma_i = 0.01$.

Fig. 1: The t -distribution with different shape parameters.

B. Maximum likelihood estimation

In MLE, we minimize the following performance index:

$$J_o = - \sum_{i=1}^m \sum_{k=1}^N \ln f_i(\epsilon_i(k)) \quad (17)$$

To minimize J_o , take the derivative of the log likelihood J_o and use $\frac{\partial \epsilon_i(k)}{\partial \hat{x}} = -H_i^T$ from (6):

$$\begin{aligned} \frac{\partial J_o}{\partial \hat{x}} &= \frac{\partial J_o}{\partial \epsilon_i(k)} \frac{\partial \epsilon_i(k)}{\partial \hat{x}} \\ &= - \sum_{i=1}^m \sum_{k=1}^N \frac{\partial \ln f_i(\epsilon_i(k))}{\partial \epsilon_i(k)} \frac{1}{\epsilon_i(k)} \epsilon_i(k) \frac{\partial \epsilon_i(k)}{\partial \hat{x}} \\ &= - \sum_{i=1}^m \sum_{k=1}^N w_i(k) \epsilon_i(k) H_i^T \\ &= - \sum_{i=1}^m \sum_{k=1}^N \frac{(v_i + 1) \epsilon_i(k)}{\xi_i^2 v_i + (\epsilon_i(k))^2} H_i^T \\ &\triangleq \Psi(\varepsilon) \end{aligned} \quad (18)$$

where $\varepsilon = [\epsilon(1)^T \dots \epsilon(N)^T]^T$ and

$$w_i(k) = \frac{(v_i + 1)}{\xi_i^2 v_i + (\epsilon_i(k))^2} \quad (19)$$

C. Influence function illustration

A textbook example [12] is used to illustrate the influence function. The measurement model is given as

$$z(k) = \begin{bmatrix} 1.0 & 1.5 \\ 0.5 & -2.5 \\ -1.5 & 0.25 \\ 0 & -1.0 \\ 1.0 & -0.5 \end{bmatrix} x(k) + \epsilon(k)$$

where the noise ϵ_k follows the t -distribution with $v_i = 3$, $\xi_i = 2$, $i = 1, \dots, 4$. The measurements are $z = [-3.01 \ 3.52 \ -5.49 \ 4.03 \ 5.01]^T$.

According to (18), the Ψ in this example is $\Psi = [\Psi_1, \Psi_2]^T$, where

$$\begin{aligned} \Psi_1 &= \frac{(v_1 + 1)(z_1 - \hat{x}_1 - 1.5\hat{x}_2)}{\xi_1^2 v_1 + (z_1 - \hat{x}_1 - 1.5\hat{x}_2)^2} \\ &\quad + 0.5 \frac{(v_2 + 1)(z_2 - 0.5\hat{x}_1 + 2.5\hat{x}_2)}{\xi_2^2 v_2 + (z_2 - 0.5\hat{x}_1 + 2.5\hat{x}_2)^2} \\ &\quad - 1.5 \frac{(v_3 + 1)(z_3 + 1.5\hat{x}_1 - 0.25\hat{x}_2)}{\xi_3^2 v_3 + (z_3 + 1.5\hat{x}_1 - 0.25\hat{x}_2)^2} \\ &\quad + \frac{(v_5 + 1)(z_5 - \hat{x}_1 + 0.5\hat{x}_2)}{\xi_5^2 v_5 + (z_5 - \hat{x}_1 + 0.5\hat{x}_2)^2} \end{aligned} \quad (20)$$

$$\begin{aligned} \Psi_2 &= 1.5 \frac{(v_1 + 1)(z_1 - \hat{x}_1 - 1.5\hat{x}_2)}{\xi_1^2 v_1 + (z_1 - \hat{x}_1 - 1.5\hat{x}_2)^2} \\ &\quad - 2.5 \frac{(v_2 + 1)(z_2 - 0.5\hat{x}_1 + 2.5\hat{x}_2)}{\xi_2^2 v_2 + (z_2 - 0.5\hat{x}_1 + 2.5\hat{x}_2)^2} \\ &\quad + 0.25 \frac{(v_3 + 1)(z_3 + 1.5\hat{x}_1 - 0.25\hat{x}_2)}{\xi_3^2 v_3 + (z_3 + 1.5\hat{x}_1 - 0.25\hat{x}_2)^2} \\ &\quad - \frac{(v_4 + 1)(z_4 - \hat{x}_2)}{\xi_4^2 v_4 + (z_4 - \hat{x}_2)^2} \\ &\quad - 0.5 \frac{(v_5 + 1)(z_5 - \hat{x}_1 + 0.5\hat{x}_2)}{\xi_5^2 v_5 + (z_5 - \hat{x}_1 + 0.5\hat{x}_2)^2} \end{aligned} \quad (21)$$

Letting $\Psi = 0$ gives the MLE solution $\hat{x}_{MLE} = [3.005, -4.010]^T$. The red surfaces at the top of Fig. 2 shows the plots of Ψ_1 and Ψ_2 as functions of $\hat{x}_1$ and $\hat{x}_2$, while the black surfaces represent $\Psi_1 = 0$ and $\Psi_2 = 0$. The two curves from the intersections of the red surfaces and the black surfaces are obtained by setting Equations (20) and (21) equal to zero. The intersection between the two curves as shown at the bottom of Fig. 2, is the solution to the MLE.

Fig. 2: The original maximum likelihood estimation.

An alternative approach to solve $\Psi = 0$ is given as follows. Take the first-order Taylor series approximation of Ψ about the operating point $\bar{x}$ giving

$$\Psi \approx \Psi|_{\hat{x}=\bar{x}} + \left. \frac{\partial \Psi}{\partial \hat{x}} \right|_{\hat{x}=\bar{x}} (\hat{x} - \bar{x}) \quad (22)$$

Set $\Psi = 0$ in (22) gives

$$\hat{x}_{Tay} \approx \bar{x} - \left(\left. \frac{\partial \Psi}{\partial \hat{x}} \right|_{\hat{x}=\bar{x}} \right)^{-1} \Psi|_{\hat{x}=\bar{x}} \quad (23)$$

If the operating point $\bar{x} = [2, -3]^T$ then (23) gives $\hat{x}_{Tay} = [3.839, -5.066]^T$. The green surfaces in the first row of Fig. 3 shows the plots Ψ as a function of $\hat{x}$ using (22). The two curves from the intersections of the green surfaces and the black surfaces are obtained by setting (22) equals to zero. The intersection between the two curves shown in the bottom of Fig. 3 is the approximate solution from Taylor series approximation.

If an outlier occurs, for example, 15.01 instead of 5.01 for the fifth measurement, then the MLE estimation result is $\hat{x}_{MLE} = [3.050, -4.079]^T$. The approximate result using Taylor series expansion is $\hat{x}_{Tay} = [5.046, -5.628]^T$ as shown in Fig. 4. This result is far from the MLE result.

The term $\left. \frac{\partial \Psi}{\partial \hat{x}} \right|_{\hat{x}=\bar{x}}$ in (22) is sensitive to the change from 5.01 to 15.01. Therefore, replacing the term with its expecta-

Fig. 3: Taylor series approximation when no outlier occurs.

Fig. 4: Taylor series approximation when outlier occurs.

tion, i.e., $\int_{-\infty}^{\infty} \frac{\partial \Psi}{\partial \hat{x}} f(\epsilon) d\epsilon$, (22) is changed to

$$\Psi \approx \Psi|_{\hat{x}=\bar{x}} + \left(\int_{-\infty}^{\infty} \frac{\partial \Psi}{\partial \hat{x}} f(\epsilon) d\epsilon \right) (\hat{x} - \bar{x}) \quad (24)$$

Set (24) equals to zero gives

$$\hat{x}_{IF} \approx \bar{x} - \left(\int_{-\infty}^{\infty} \frac{\partial \Psi}{\partial \hat{x}} f(\epsilon) d\epsilon \right)^{-1} \Psi|_{\hat{x}=\bar{x}} \quad (25)$$

where the second term on the right side is the influence function [33]. If the value of the fifth measurement is still 5.01 and the other measurements are unchanged, an approximation of $\hat{x}_{MLE}$ using (25) is $\hat{x}_{IF} = [3.370, -4.250]^T$. The blue surfaces at the top of Fig. 5 shows the plots Ψ as a function of $\hat{x}$ using (25). The two curves from the intersections of the blue surfaces and the black surfaces are obtained by setting (24) equals to zero. The intersection between these

two curves mentioned above also shows the approximate result $\hat{x}_{IF}$. Compared with $\hat{x}_{Tay} = [5.046, -5.628]^T$, the estimation result using (25) is $\hat{x}_{IF} = [3.347, -4.252]^T$ which is closer to the MLE estimation result of $\hat{x}_{MLE} = [3.050, -4.079]^T$.

Fig. 5: IF approximation.

D. Influence function approximation

The definition of the IF is given in [1], [33], [34] as

$$\begin{aligned} \Delta \hat{x} &= \hat{x} - x \approx \text{IF}(\varepsilon) \\ &= - \left[\int_{-\infty}^{\infty} \frac{\partial \Psi(\varepsilon)}{\partial x} dF(\varepsilon) \right]^{-1} \Psi(\varepsilon) \end{aligned} \quad (26)$$

where

$$\begin{aligned} dF(\varepsilon) &= f_1(\varepsilon_1(1)) \cdots f_m(\varepsilon_m(1)) \cdots f_1(\varepsilon_1(N)) \cdots f_m(\varepsilon_m(N)) \\ &\quad \times d\varepsilon_1(1) \cdots d\varepsilon_m(1) \cdots d\varepsilon_1(N) \cdots d\varepsilon_m(N) \end{aligned}$$

since the measurement noise ε_i is assumed to be independent and identically distributed.

Using (26), $\hat{x}$ is given as:

$$\hat{x} \approx x + \text{IF}(\varepsilon) \quad (27)$$

In practice, since x is unknown, the noise vector $\varepsilon = Z - \bar{H}x$ is replaced by $E = Z - \bar{H}\hat{x}$ and (28) is used to find $\hat{x}$ iteratively given an initial guess $\hat{x}$.

$$\hat{x} \leftarrow \hat{x} + \text{IF}(E) \quad (28)$$

where $\text{IF}(E)$ is given by

$$\text{IF}(E) = (\bar{H}^T \Omega \bar{H})^{-1} \bar{H}^T \Lambda E \quad (29)$$

The derivation of (29) and the definition of Ω, Λ are given in Appendix.

Fig. 6: The flowchart of the traditional robust estimators (MS, QL, SR and SHGM) and the proposed estimator.

The noise pdf can be obtained by analyzing the historical data. The robust estimators such as MS, QL, SR and SHGM estimator need to calculate the inverse of matrix $(\bar{H}^T W \bar{H})$ as given in (15) because the matrix W is updated during the iteration process. When the size of power system becomes large, the inverse computation will take a long time. However, the matrix $(\bar{H}^T \Omega \bar{H})^{-1}$ of the IF in (29) can be precomputed because Ω remains unchanged, thus saving computational load in real-time applications.

Unlike references [18]–[20], the contribution of this paper is to employ the Influence Function (IF) approximation so that the proposed estimator with t -distribution noise can be efficiently implemented within the framework (Fig. 6) of traditional robust estimators. This flowchart shows that the only difference between the algorithms of the traditional robust estimators and the proposed estimator is in the calculation of Δ . For the traditional robust estimator, $\Delta = (\bar{H}^T W \bar{H})^{-1} \bar{H}^T W E$ according to (15). For the proposed estimator, $\Delta = (\bar{H}^T \Omega \bar{H})^{-1} \bar{H}^T \Lambda E$ according to (29).

E. Estimation precision

Precision of estimation is traditionally given by its variances. As given in [33], the variance of the state estimates can be approximated as

$$\begin{aligned} \text{Var}(\hat{x}) &= \text{Var}(\hat{x} - \bar{x}) = \int_{-\infty}^{\infty} \text{IF}(\varepsilon) \text{IF}(\varepsilon)^T dF(\varepsilon) \\ &= \left[\int_{-\infty}^{\infty} \frac{\partial \Psi(\varepsilon)}{\partial \hat{x}} dF(\varepsilon) \right]^{-1} \int_{-\infty}^{\infty} \Psi(\varepsilon) (\Psi(\varepsilon))^T dF(\varepsilon) \\ &\quad \left[\int_{-\infty}^{\infty} \frac{\partial \Psi(\varepsilon)}{\partial \hat{x}} dF(\varepsilon) \right]^{-T} \end{aligned} \quad (30)$$

Let the residue E be given by the noise ε in (10), equation (30) gives

$$\begin{aligned} \text{Var}(\hat{x}) &= \left[\bar{H}^T \int_{-\infty}^{\infty} \frac{\partial W \varepsilon}{\partial \varepsilon} dF(\varepsilon) \bar{H} \right]^{-1} \\ &\quad \times \bar{H}^T \int_{-\infty}^{\infty} W \varepsilon \varepsilon^T W^T dF(\varepsilon) \bar{H} \\ &\quad \times \left[\bar{H}^T \int_{-\infty}^{\infty} \frac{\partial W \varepsilon}{\partial \varepsilon} dF(\varepsilon) \bar{H} \right]^{-T} \\ &= [\bar{H}^T \Omega \bar{H}]^{-1} [\bar{H}^T \Pi \bar{H}] [\bar{H}^T \Omega \bar{H}]^{-T} \quad (31) \end{aligned}$$

where

$$\begin{aligned} \Omega &= \int_{-\infty}^{\infty} \frac{\partial W \varepsilon}{\partial \varepsilon} dF(\varepsilon) \\ &= \text{diag}(\Omega_1(1), \dots, \Omega_m(1), \dots, \Omega_1(N), \dots, \Omega_m(N)) \\ \Pi &= \int_{-\infty}^{\infty} W \varepsilon \varepsilon^T W^T dF(\varepsilon) \\ &= \text{diag}(\Pi_1(1), \dots, \Pi_m(1), \dots, \Pi_1(N), \dots, \Pi_m(N)) \quad (32) \end{aligned}$$

Substituting (19) into (32) gives

$$\Pi_i(k) = 2 \int_0^{\infty} \frac{(v_i + 1)^2 \epsilon_i^2}{(\xi_i^2 v_i + \epsilon_i^2)^2} f_i(\epsilon_i) d\epsilon_i$$

The sum of the variances (SV) is the trace of the matrix

$$\text{SV} = \text{tr} \left([\bar{H}^T \Omega \bar{H}]^{-1} [\bar{H}^T \Pi \bar{H}] [\bar{H}^T \Omega \bar{H}]^{-T} \right) \quad (33)$$

and taking the average gives

$$\text{AMSE} = \frac{\text{SV}}{n} = \frac{\text{tr} \left([\bar{H}^T \Omega \bar{H}]^{-1} [\bar{H}^T \Pi \bar{H}] [\bar{H}^T \Omega \bar{H}]^{-T} \right)}{n} \quad (34)$$

which can be approximated from the simulation as the Average Mean Square Errors

$$\text{AMSE} = \frac{1}{N_r} \sum_{k=1}^{N_r} \frac{1}{n} \|\hat{x}_k - x_k\|^2 \quad (35)$$

where N_r is the number of simulation runs.

F. The weighted least squares (WLS) connection

The proposed estimator will reduce to WLS when the noise ϵ_i is Gaussian:

$$f_i(\epsilon_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\epsilon_i^2}{2\sigma_i^2}\right) \quad (36)$$

It is well known that (16) reduces to (36) when $v_i \rightarrow \infty$ and $\xi_i = \sigma_i$ [18]. According to (??) and (??), we would get

$$\begin{aligned} \Lambda_i(k) &= \lim_{v_i \rightarrow \infty} \frac{(v_i + 1)}{\xi_i^2 v_i + (\epsilon_i(k))^2} = \frac{1}{\xi_i^2} = \frac{1}{\sigma_i^2} \\ \Omega_i(k) &= \lim_{v_i \rightarrow \infty} \int_{-\infty}^{+\infty} \frac{(v_i + 1)(\xi_i^2 v_i - \epsilon_i^2)}{(\xi_i^2 v_i + \epsilon_i^2)^2} f_i(\epsilon_i) d\epsilon_i = \frac{1}{\sigma_i^2} \end{aligned}$$

Then the estimated results $\hat{x}$ is given as

$$\begin{aligned} \hat{x} &\leftarrow \hat{x} + (\bar{H}^T \Omega \bar{H})^{-1} \bar{H}^T \Lambda E \\ &= (\bar{H}^T W \bar{H})^{-1} \bar{H}^T W \bar{H} \hat{x} + (\bar{H}^T W \bar{H})^{-1} \bar{H}^T W E \\ &= (\bar{H} W \bar{H})^{-1} \bar{H} W Z \end{aligned}$$

which is the same as (13).

IV. SIMULATION AND EXPERIMENT RESULTS

In this section, simulation results obtained from the IEEE 14-bus and 118-bus systems are given in Example 1 and 2 respectively. The measurement noises are generated according to [2], [3]. Experimental results obtained from the 13-bus microgrid at the Nanyang Technological University (NTU) is given in Example 3 where the measurements are collected from the real system and the noise distributions are unknown. For easy reference, the pdfs of the Gaussian distribution and outlier, the t -distribution (to model noise with outliers), the Gaussian distribution (to model noise with outliers), and the parameters of the robust state estimators are summarized in Table II.

A. Example 1: IEEE 14-bus system with PMU measurements

The IEEE 14-bus system is modeled in RT-LAB, a real-time simulation platform for power systems [35], [36]. Meanwhile, the proposed estimator is programmed on a remote computer simulating the control center. The computer is connected to RT-LAB via cable.

The PMUs are placed according to [6]. Fifty-eight measurements, z_i , $i = 1, \dots, 58$, comprising of 12 voltages ($i = 1, \dots, 12$) and 46 currents ($i = 13, \dots, 58$) are taken at each time instance and $N = 3$ sets of measurements are used to give one set of estimates. The measurement matrix H is calculated according to the parameters in [37]. There are $n = 28$ states in the vector $x = [V_1^r \dots V_{14}^r V_1^{im} \dots V_{14}^{im}]^T$. The real and imaginary part of the voltage phasor are defined as V_i^r and V_i^{im} respectively. The parameters for the QL and SR estimators are chosen as $a_i = 3$ while the parameters for the MS estimator are $a_i = 3$, $b_i = 4$, $r_i = 5$ according to [5]. The penalty κ_i , $i = 1, \dots, 58$ of SHGM estimator is calculated according to [12], [16]. Three different types of measurement noise pdfs are considered in the simulations. Note that the actual noises need not be given by the t -distribution. There are studies in power systems where the Gaussian noises are corrupted by outliers with uniform or Gaussian distribution [2],

TABLE II: Parameters of noises and estimators used in the simulations and experiment.

Example	1			2	3
Model	IEEE 14-bus system			IEEE 118-bus system	Microgrid
Measurement noise with outliers	$95\%N(0.005) + 5\%U(0.005)$	$95\%N(0.005) + 5\%N(0.05)$	$t_3(0.005)$	$t_3(0.005)$	unknown
t -distribution	$v_i = 4, \xi_i = 0.0046$	$v_i = 2, \xi_i = 0.0042$	$v_i = 3, \xi_i = 0.005$	$v_i = 3, \xi_i = 0.005$	$v_i = 3, \xi_i = 0.000028$
Gaussian	$\sigma_i = 0.007$	$\sigma_i = 0.012$	$\sigma_i = 0.009$	$\sigma_i = 0.009$	$\sigma_i = 0.000042$
† Estimator	LAV, WLS with LNR, QL(3), SHGM(3), SR(3), MS(3,4,5), the proposed estimator				

$$N(\sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\epsilon_i^2}{2\sigma_i^2}\right); U(\sigma_i) = \frac{1}{4\sigma_i} \text{ for } 2\sigma_i \leq |\epsilon_i| \leq 6\sigma_i; t_3(\xi_i) = \frac{\Gamma(2)}{\sqrt{3\pi}(\xi_i)\Gamma(1.5)} \left(1 + \frac{|\epsilon_i|^2}{3(\xi_i)^2}\right)^{-2}.$$

† Parameter a_i is given in the bracket of QL(a_i), SHGM(a_i) and SR(a_i) while parameters a_i, b_i, r_i are given in the bracket of MS(a_i, b_i, r_i).

Fig. 7: The t -distribution with $v_i = 4, \xi_i = 0.0046$ and Gaussian distribution with $\sigma_i = 0.007$ are used to fit the noise generalized according to (37) with $\sigma_i = 0.005$.

[3]. In Example 1, we model the Gaussian noise corrupted by outliers with uniform or Gaussian distribution by the t -distribution.

1) *1 Gaussian + 1 Uniform (1G+1U)*: Firstly the noise is associated with the pdf

$$f_i(\epsilon_i) = \begin{cases} \frac{0.95}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\epsilon_i^2}{2\sigma_i^2}\right) + \frac{0.05}{8\sigma_i} & 2\sigma_i \leq |\epsilon_i| \leq 6\sigma_i \\ \frac{0.95}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\epsilon_i^2}{2\sigma_i^2}\right) & \text{otherwise} \end{cases} \quad (37)$$

where $\sigma_i = 0.005$. The pdf in the form of a mixture distribution in (37) is also used in [2]. The 5% of uniform distribution $\frac{0.05}{4\sigma_i}$ in (37) is useful for modeling initial conditions, disturbances, and measurement errors that are equally likely to occur anywhere within a given interval [38]. Fig. 7 shows how the Gaussian ($\sigma_i = 0.007$) and t -distribution ($v_i = 4, \xi_i = 0.0046$) can be fitted to the histogram using the maximum likelihood criterion according to [39]. The histogram represents the measurement noise generated by (37). In practice, the fitting process can be done off-line using historical measurements.

The threshold parameters δ and q_{max} are chosen as 10^{-8} and 200 respectively. The AMSE results obtained by $N_r = 10,000$ simulations runs using (35) and by the formula in Equation (34) are given in Table III. The AMSEs of the

proposed estimator obtained from simulation are verified by the formula in Equation (34) as given in the last two columns. In addition, It is clear that the AMSE of the proposed estimator is the smallest. Hence we have demonstrated that the t -distribution can be used as the noise model in the proposed estimator. Compared to the WLS estimator, the proposed estimator improves on the AMSE by $\frac{2.95-2.08}{2.95} \times 100\% = 29.5\%$. The improvement of proposed estimator over other robust estimators can be calculated similarly and the results are given in Table III.

2) *1 Gaussian + 1 Gaussian (1G+1G)*: In this subsection the noise ϵ_i is associated with the pdf [3], [33]

$$f_i(\epsilon_i) = \frac{0.95}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\epsilon_i^2}{2\sigma_i^2}\right) + \frac{0.05}{\sqrt{2\pi(10\sigma_i)^2}} \exp\left(-\frac{\epsilon_i^2}{2(10\sigma_i)^2}\right) \quad (38)$$

where $\sigma_i = 0.005$. The first term in the above pdf represents the 95% of Gaussian noise while the second term represents outliers by the 5% Gaussian distribution with standard deviation $10\sigma_i$. Following the above subsection, the t -distribution with $v_i = 2, \xi_i = 0.0042$ and the Gaussian distribution with $\sigma_i = 0.012$ are used to model the noise generated by (38). The AMSE results are given in Table III. It is clear that the AMSE = 2.20×10^{-6} of the proposed estimator is the smallest. More estimated results such as the Mean Squared Error ($MSE_k = \frac{1}{n} \|\hat{x}_k - x_k\|^2$), $\hat{V}_1^r$ and $\hat{V}_{14}^{im}$ of the MS estimator and the proposed estimator are shown in Figs. 8, 9 (a) and 9 (b). The ones obtained from the proposed estimator are closer to the true states.

3) *t_3 noise*: In this subsection, let the noise ϵ_i be associated with the following t -distribution which is commonly used to model noise with outliers [33]. The parameters in (16) are chosen as: $v_i = 3, \xi_i = 0.005$ for measurement noises and Table II shows the Gaussian parameters obtained by the maximum likelihood criterion. The AMSE results of the different estimators are given in Table III. It is clear that the AMSE = 2.55×10^{-6} of the proposed estimator is the smallest. Table III also shows the improvements of the proposed estimator over other robust estimators.

Finally it is clear from Table III that the AMSEs of the

TABLE III: The AMSE of the different estimators for the IEEE 14-bus system.

Noise with outliers	Estimator	WLS	WLS with LNR	QL	SR	MS	SHGM	LAV	Proposed estimator	
									Simulation using (35)	Formula using (34)
1G+1U	$\S$ AMSE	2.95	2.39	2.79	2.74	2.84	2.78	2.51	2.08	2.04
	$\dagger$ Proposed estimator's improvement (%)	$\frac{2.95-2.08}{2.95} \times 100 = 29.5$	13.0	25.4	24.1	26.7	25.2	17.1	#0	#0
1G+1G	AMSE	9.90	2.70	4.85	4.07	3.55	4.84	2.53	2.20	2.15
	Proposed estimator's improvement (%)	$\frac{9.90-2.20}{9.90} \times 100 = 77.3$	18.5	54.6	45.9	38.0	54.5	13.0	#0	#0
t_3	AMSE	4.73	3.29	3.65	3.52	3.49	3.65	2.78	2.55	2.46
	Proposed estimator's improvement (%)	$\frac{4.73-2.55}{4.73} \times 100 = 46.1$	22.5	30.1	27.5	26.9	30.1	8.3	#0	#0

$\S$ AMSE unit: $\times 10^{-6}$.

$\dagger$ Proposed estimator's improvement (%) = $\frac{(\text{Traditional estimator} - \text{Proposed estimator})}{\text{Traditional estimator}} \times 100$; unit: %.

Proposed estimator's improvement over itself is 0%.

Fig. 8: The MSEs of the MS estimator and the proposed estimator.

proposed estimator from simulation using (35) are verified by the formula in Equation (34) as shown in the last two columns and they are smaller than the AMSEs of all other estimators studied in this paper.

B. Example 2: IEEE 118-bus system with PMU measurements

To further verify the performance of the proposed estimator, the IEEE 118-bus system is used in the simulation. The PMUs are placed according to [6] where a total number of 108 voltage measurements and 366 current measurements are considered. The pdf of measurement noise is given in (16) and the t -distribution parameters are $v_i = 3$, $\xi_i = 0.005$. The AMSE results are given in Table IV and it is clear that the $\text{AMSE} = 0.58 \times 10^{-6}$ of the proposed estimator is the smallest.

The computational efficiency of the proposed estimator is analyzed and compared to that of other robust estimators. The simulation runs using MATLAB version R2012b on a Windows 10 computer configured with Intel® Core™, CPU

(a)

(b)

Fig. 9: The estimated result of the MS estimator and the proposed estimator. (a) $\hat{V}_1^r$. (b) $\hat{V}_{14}^{im}$.

i7-4500U, 1.80 GHz and 8 GB RAM, while the LAV optimization problem is expressed as an equivalent linear programming problem and is solved using GUROBI [5]. The computational time of the different estimators in the IEEE 118-bus system are given in Table IV. Except for WLS, WLS with LNR and LAV

Fig. 10: Microgrid with 1 programmable source and 1 load.

Fig. 11: The t -distribution with $v_i = 3, \xi_i = 2.8 \times 10^{-5}$ and Gaussian distribution with $\sigma_i = 4.2 \times 10^{-5}$ are used to fit the raw measurement (the real power flow from Bus 3 to Bus 2) collected from the microgrid.

estimator, the proposed estimator is faster. This is because the matrix $(\hat{H}^T \Omega \hat{H})^{-1}$ in the proposed estimator can be computed off-line.

C. Example 3: Experiment in microgrid

Experiments on a lab microgrid at Nanyang Technological University have been performed for isolated operations of the microgrid. The lab microgrid has a total of 13 buses. There are several generation sources and loads currently installed for the microgrid. For simplicity, we use the programmable source and one load in the microgrid, see Fig. 10. The programmable source and the load are connected to Bus 7 and Bus 4 respectively. Measurement data was sampled at intervals of 5 seconds with very low noise levels. We fit the measurement noises with the t -distribution and Gaussian distribution using the maximum likelihood criterion. Fig. 11 shows the raw measurement (the real power flow from Bus 3 to Bus 2). The proposed estimator estimates the bus voltage phasor every 30 seconds. The system provides a total of 31 measurements consisting of 1 voltage magnitude measurement, 2 pairs of power injection and 13 pairs of power flow.

According to [12]: “Treatment of bad data can be viewed as a robustness issue for an estimator. If the estimated state remains insensitive to major deviations in a limited number of measurements, then the corresponding estimator is considered statistically robust”. In Fig. 12 a measurement (the real power flow from Bus 3 to Bus 2) outlier occurs at $k = 300$. Compared to the WLS and MS estimators, the proposed estimator is more robust since smaller deviations of the estimation results are obtained. Note that the QL, WLS and LAV estimators are special cases of the MS estimator.

Fig. 12: The estimation result $\hat{V}_2^r$ with a measurement (the real power flow from Bus 3 to Bus 2) outlier at 300 seconds.

V. CONCLUSIONS

In this paper a robust state estimator using the t -distribution as the noise model is proposed. The influence function is employed to obtain the solution to the resulting MLE optimization problem so that the proposed estimator can be implemented within the framework of traditional robust estimators. Numerical results obtained from simulations of the IEEE 14-bus system, IEEE 118-bus system and experiment on the microgrid demonstrated the effectiveness and robustness of the proposed estimator. The proposed estimator could suppress the influence of outliers with smaller AMSE than the traditional robust estimators.

REFERENCES

- [1] W. K. Ho, K. V. Ling, H. D. Vu, and X. Wang, “Filtering of the armaz process with generalized t -distribution noise: The influence function approach,” *Industrial & Engineering Chemistry Research*, vol. 53, no. 17, pp. 7019–7028, 2014.
- [2] R. Baldick, K. Clements, Z. Pinjo-Dzagal, and P. Davis, “Implementing nonquadratic objective functions for state estimation and bad data rejection,” *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 376–382, Feb 1997.
- [3] J. Zhao and L. Mili, “Robust unscented kalman filter for power system dynamic state estimation with unknown noise statistics,” *IEEE Trans. Smart Grid*, pp. 1–1, 2018.
- [4] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. M. Begovic, and A. Phadke, “Wide-area monitoring, protection, and control of future electric power networks,” *Proceedings of the IEEE*, vol. 99, no. 1, pp. 80–93, Jan 2011.

TABLE IV: The AMSEs and average computational time for one simulation run of the estimators for the IEEE 118-bus system.

Estimator	WLS	WLS with LNR	QL	SR	MS	SHGM	LAV	Proposed estimator
AMSE	1.06	0.75	0.83	0.80	0.79	0.84	0.64	0.58
Improvement (%)	$\frac{1.06-0.58}{1.06} \times 100 = 45.3$	22.7	30.1	27.5	26.6	30.9	9.4	#0
Average computational time for one run (s)	0.105	0.126	0.646	1.012	0.831	1.134	0.084	0.446

AMSE unit: $\times 10^{-6}$.

Proposed estimator's improvement over itself is 0%.

- [5] M. Göl and A. Abur, "LAV based robust state estimation for systems measured by PMUs," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1808–1814, July 2014.
- [6] P. Yang, Z. Tan, A. Wiesel, and A. Nehorai, "Power system state estimation using PMUs with imperfect synchronization," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4162–4172, Nov. 2013.
- [7] A. Abur and A. Rouhani, "Linear phasor estimator assisted dynamic state estimation," *IEEE Trans. Smart Grid*, vol. PP, no. 99, pp. 1–1, 2016.
- [8] D. Wang and J. Romagnoli, "A framework for robust data reconciliation based on a generalized objective function," *Industrial & Engineering Chemistry Research*, vol. 42, no. 13, pp. 3075–3084, 2003.
- [9] V. Barnett and T. Lewis, *Outliers in statistical data*. Wiley, New York, 1994.
- [10] A. Monticelli and A. Garcia, "Reliable bad data processing for real-time state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 5, pp. 1126–1139, May 1983.
- [11] H. M. Merrill and F. C. Schweppe, "Bad data suppression in power system static state estimation," *IEEE Trans. Power App. Syst.*, no. 6, pp. 2718–2725, 1971.
- [12] A. Abur and A. Gómez-Expósito, *Power system state estimation: Theory and implementation*. CRC Press, 2004.
- [13] A. Monticelli, *State estimation in electric power systems: A generalized approach*. New York: Springer, 1999.
- [14] J. Zhao, L. Mili, and R. C. Pires, "Statistical and numerical robust state estimator for heavily loaded power systems," *IEEE Trans. Power Syst.*, DOI: 10.1109/TPWRS.2018.2849325, 2018.
- [15] Y. Ejima, H. Kondo, and S. Iwamoto, "New bad data rejection algorithm using nonquadratic objective function for state estimation," in *2007 IEEE Power Engineering Society General Meeting*, June 2007, pp. 1–8.
- [16] L. Mili, M. Cheniae, N. Vichare, and P. Rousseeuw, "Robust state estimation based on projection statistics," *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 1118–1127, May 1996.
- [17] M. Göl and A. Abur, "A robust PMU based three-phase state estimator using modal decoupling," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2292–2299, Sep. 2014.
- [18] S. Kotz and S. Nadarajah, *Multivariate t-distributions and their applications*. Cambridge University Press, 2004.
- [19] P. Theodossiou, "Financial data and the skewed generalized t distribution," *Management Science*, vol. 44, no. 12-Part-1, pp. 1650–1661, 1998.
- [20] D. Xu, C. Shen, and F. Shen, "A robust particle filtering algorithm with non-gaussian measurement noise using student-t distribution," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 30–34, Jan 2014.
- [21] W. Zheng, W. Wu, A. Gomez-Expósito, B. Zhang, and Y. Guo, "Distributed robust bilinear state estimation for power systems with nonlinear measurements," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 499–509, Jan 2017.
- [22] Y. Chen, F. Liu, S. Mei, and J. Ma, "A robust wlav state estimation using optimal transformations," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2190–2191, July 2015.
- [23] K. Mahmoud and M. Abdel-Nasser, "Efficient spf approach based on regression and correction models for active distribution systems," *IET Renew. Power Gener.*, vol. 11, no. 14, pp. 1778–1784, 2017.
- [24] Y. Chen, F. Liu, S. Mei, and J. Ma, "A robust WLAV state estimation using optimal transformations," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2190–2191, July 2015.
- [25] —, "Toward adaptive robust state estimation based on MCC by using the generalized gaussian density as kernel functions," *Int. J. Elect. Power Energy Syst.*, vol. 71, pp. 297–304, 2015.
- [26] Y. Chen, J. Ma, P. Zhang, F. Liu, and S. Mei, "Robust state estimator based on maximum exponential absolute value," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1537–1544, July 2017.
- [27] K. G. Boroojeni, M. H. Amini, and S. S. Iyengar, *Smart Grids: Security and Privacy Issues*. Springer, 2017.
- [28] M. H. Amini, M. Rahmani, K. G. Boroojeni, G. Atia, S. S. Iyengar, and O. Karabasoglu, "Sparsity-based error detection in dc power flow state estimation," in *IEEE International Conference on Electro Information Technology (EIT)*, May 2016, pp. 0263–0268.
- [29] W. K. Ho, T. Chen, K. V. Ling, and L. Sun, "Variance analysis of robust state estimation in power system using influence function," *International Journal of Electrical Power & Energy Systems*, vol. 92, pp. 53–62, 2017.
- [30] J. Zhang, G. Welch, and G. Bishop, "Observability and estimation uncertainty analysis for PMU placement alternatives," in *Proc. North Amer. Power Symp.*, 2010, pp. 1–8.
- [31] Y. F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, "State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 33–43, Sept 2012.
- [32] M. Ahmad, *Power system state estimation*. Artech House, 2012.
- [33] F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel, *Robust statistics: The approach based on influence functions*. Wiley, 2011.
- [34] P. Huber, *Robust statistics*. Wiley, New York, 2009.
- [35] C. Dufour and J. Blanger, "On the use of real-time simulation technology in smart grid research and development," *IEEE Trans. Ind. Applicat.*, vol. 50, no. 6, pp. 3963–3970, Nov 2014.
- [36] D. Bian, M. Kuzlu, M. Pipattanasomporn, S. Rahman, and Y. Wu, "Real-time co-simulation platform using opal-rt and opnet for analyzing smart grid performance," in *2015 IEEE Power Energy Society General Meeting*, July 2015, pp. 1–5.
- [37] I. Dabbagchi, "Power systems test case archive," University of Washington, Tech. Rep., 1993. [Online]. Available: <http://www.ee.washington.edu/research/pstca/>
- [38] D. G. Robertson and J. H. Lee, "On the use of constraints in least squares estimation and control," *Automatica*, vol. 38, no. 7, pp. 1113–1123, 2002.
- [39] D. Gao, Y. Sun, and Y. Lu, "A robust demand response control of commercial buildings for smart grid under load prediction uncertainty," *Energy*, vol. 93, pp. 275–283, 2015.