Robust Power System State Estimation Using $t$-Distribution Noise Model

Tengpeng Chen, Lu Sun, Keck-Voon Ling, and Weng Khuen 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 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, square-root, Schweppe–Huber generalized-M, multiple-segment, and least absolute value estimators. A new approximate AMSE formula is also derived for the proposed estimator to predict and evaluate its precision.

Index Terms—Influence function (IF), maximum likelihood estimation (MLE), robust state estimation, $t$-distribution.

NOMENCLATURE

$x$ True state vector.
$\hat{x}$ Estimated state vector.
$\hat{x}_{MLE}$ Estimated state vector obtained by maximum likelihood estimation.
$\hat{x}_{SE}$ Estimated state vector using Taylor series expansion.
$\hat{x}_{IF}$ Estimated state vector using influence function approximation.
$x_t$ 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.
$\epsilon$ Measurement noise.
$\sigma_i$ $i$th standard deviation of measurement noise $\epsilon_i$.

\[ k \] Time instance.
\[ e_i(k) \] $i$th measurement residual at time instance $k$.
\[ \rho \] $\rho$-function.
\[ J \] Total cost function.
\[ f_i(\epsilon_i) \] Probability density function of measurement noise $\epsilon_i$.
$H$ Measurement matrix.
$H_i$ $i$th row of measurement matrix.
$h_i(x)$ Nonlinear function relating the state vector $x$.
$\bar{h}_i(x)$ Nonlinear function relating the state vector $x$ to measurement $i$.
$w_i(k)$ Weighting factor for measurement $i$ at time $k$.
$\Psi$ $\Psi$-function.
$\Phi(\cdot)$ Influence function.
$\Gamma(\cdot)$ Gamma function.
$m$ Number of measurements in a 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 $e_i$.
$\Omega, \Lambda$ Diagonal matrix.
$\nu_i, \xi_i$ $t$-distribution parameters.

I. INTRODUCTION

POWER system state estimation (PSSE) is one of the key functions in an energy management system. 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 the 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 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 Merrill and Schweppe propose a new approach to screen and suppress the gross error in measurements by iterations. Following this approach, many traditional M-estimators with nonquadratic 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]–[16] 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 the Gaussian distribution when the shape parameter tends to infinity [8]. Therefore, the t-distribution can be used to model both the Gaussian and non-Gaussian noise distributions. Moreover, the t-distribution is commonly used in robust state estimation literature [17]–[19] to model noise with outliers. In Example 1, we use the t-distribution to model the Gaussian noise corrupted by outliers with the 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 earlier.

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 [17]–[23]. However, to obtain the statistics such as the AMSE or variances of their state estimates, hundreds of simulations have to be run [24]. In contrast, for the proposed estimator, because of the IF 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 the 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 the 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], and [16]. This section focuses on the equations necessary for the derivation of the results in this paper.

A. Measurement Model

Consider the following measurement model in the power system:

\[ z(k) = h(x(k)) + \epsilon(k) \]  

(1)

where

\[ z(k) = [z_1(k), z_2(k), \ldots, z_m(k)]^T \]

\[ x(k) = [x_1(k), x_2(k), \ldots, x_n(k)]^T \]

\[ h(x(k)) = [h_1(x(k)), h_2(x(k)), \ldots, h_m(x(k))]^T \]

\[ \epsilon(k) = [\epsilon_1(k), \epsilon_2(k), \ldots, \epsilon_m(k)]^T. \]

The nonlinear case is applicable when the measurements are collected from Supervisory Control and Data Acquisition (SCADA). Measurement \( z_i(k), i = 1, \ldots, m \), is taken at each time instance \( k = 1, \ldots, N \). \( \epsilon_i(k) \) is the measurement noise and the rectangular coordinates of the bus voltage phasor \( x(k) \) are the states of the system. A traditional power system may be considered as a quasi-static system [25] because load demands change slowly and, hence, the state changes slowly. The typical measurement sampling interval is 2–4 s while the estimates are usually updated only once every few minutes [26]. Hence, the estimated states during the time instants \( k = 1, \ldots, N \) can be taken as \( \hat{x} \).

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

\[ z(k) \approx h(\hat{x}) + H(x(k) - \hat{x}) + \epsilon(k) \]  

(2)

where the Jacobian matrix

\[ H = \frac{\partial h(x)}{\partial x}_{|x=\hat{x}}. \]  

(3)

The \( i \)th measurement residual is defined as

\[ e_i(k) = z_i(k) - h_i(\hat{x}) \]  

(4)

where \( h_i(\hat{x}) \) is approximated as \( h_i(\hat{x}) \). \( e_i(k) \) is a vector, \( e(k) = [e_1(k), e_2(k), \ldots, e_m(k)]^T \).

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

\[ z(k) = Hx(k) + \epsilon(k) \]  

(5)

where \( H \) is the measurement matrix. In this case, the \( i \)th measurement residual is

\[ e_i(k) = z_i(k) - H_i\hat{x}. \]  

(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:

\[ J = \sum_{i=1}^{m} \sum_{k=1}^{N} \rho(e_i(k)) \]  

(7)
where \( \rho \)-function is a nonlinear chosen function of the residual \( e_i(k) \).

Differentiating (7) w.r.t. \( \hat{x} \) gives

\[
\frac{\partial J}{\partial \hat{x}} = \frac{\partial J}{\partial e_i(k)} \frac{\partial e_i(k)}{\partial \hat{x}} = \sum_{i=1}^{m} \sum_{k=1}^{N} \frac{\partial \rho(e_i(k))}{\partial e_i(k)} \frac{1}{e_i(k)} \frac{\partial e_i(k)}{\partial \hat{x}} = -\sum_{i=1}^{m} \sum_{k=1}^{N} w_i(k)e_i(k)H_i^T
\]

(8)

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

\[
\frac{\partial e_i(k)}{\partial \hat{x}} = -H_i^T.
\]

(9)

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

\[
\frac{\partial J}{\partial \hat{x}} = -\sum_{i=1}^{m} \sum_{k=1}^{N} w_i(k)(z_i(k) - H_i\hat{x})H_i^T = -H_i^TWE
\]

(10)

where

\[
H = \begin{bmatrix} H_1 & \cdots & H_m \end{bmatrix}^T \in \mathbb{R}^{n_m \times n},
\]

\[
Z = \begin{bmatrix} z(1)^T & \cdots & z(N)^T \end{bmatrix}^T \in \mathbb{R}^{n_m}
\]

(11)

\[
E = Z - \hat{H}\hat{x} = \begin{bmatrix} e(1)^T & \cdots & e(N)^T \end{bmatrix}^T \in \mathbb{R}^{n_m}
\]

(12)

\[
W = \text{diag}(w_1(1), \ldots, w_m(1), \ldots, w_1(N), \ldots, w_m(N)) \in \mathbb{R}^{n_m \times n_m}.
\]

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

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 as

\[
\hat{x} = (\hat{H}^TWE\hat{H})^{-1}\hat{H}^TWEZ.
\]

(13)

Notice that the matrix \( \hat{H}^TWE\hat{H} \) 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 \( \sigma_i \) is the standard deviation of noise \( \epsilon_i \). The LNR method [10] is usually used in WLS to deal with bad data. The normalized residuals are calculated as

\[
\hat{R} = \text{diag}(\sigma_1^2, \ldots, \sigma_m^2, \ldots, \sigma_1^2, \ldots, \sigma_m^2)
\]

\[
\hat{G} = \hat{H}^T\hat{R}^{-1}\hat{H}
\]

\[
\hat{\Omega} = \hat{R} - \hat{H}G^{-1}\hat{H}^T
\]

\[
e_i^{\text{norm}}(k) = \frac{|e_i(k)|}{\sqrt{\Omega_{ii}}}.
\]

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

\[
Z_i^{\text{new}} = Z_i^{\text{bad}} - \frac{\hat{R}_{ii}}{\Omega_{ii}}e_i^{\text{bad}}(k).
\]

(14)

The states will then be re-calculated based on the updated measurements \( Z_i^{\text{new}} \). 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 estimators is a re-weighted matrix and is updated according to Table I [12], [24].

The parameters \( a_i, b_i, r_i \) are thresholds, and \( \kappa_i \) 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 an MS estimator. The MS estimator reduces to the QL estimator when \( b_i \to \infty \) and the WLS estimator when \( b_i \to \infty \). The QL estimator reduces to the LAV estimator when the threshold \( a_i \) is small [12].

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

\[
\hat{x} = (\hat{H}^TWE\hat{H})^{-1}\hat{H}^TWE(\hat{H}\hat{x} + E)
\]

which can be solved iteratively as

\[
\hat{x} \leftarrow \hat{x} + (\hat{H}^TWE\hat{H})^{-1}\hat{H}^TWE.
\]

(15)

III. 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) [17] is used to
model measurement noise corrupted by outliers:
\[
f_i(\epsilon_i) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi} \xi_i^{\nu+1}} \left(1 + \frac{|\epsilon_i|^2}{\xi_i^2 v_i}\right)^{-\frac{\nu+1}{2}}
\]
where \(\Gamma(\cdot)\) is the gamma function, \(\xi_i\) is the scale parameter, and \(v_i\) is the shape parameter. It is well known that the \(t\)-distribution tends to infinity \([8, 17]\). Therefore, the \(t\)-distribution can be used to model both the Gaussian and non-Gaussian noise distribution. Moreover, the \(t\)-distribution is commonly used in robust state estimation literature \([17]–[19]\) 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 \(\infty\), the \(t\)-distribution reduce to a Gaussian distribution (36) with \(\sigma_i = 0.01\).

**B. Maximum Likelihood Estimation**

In MLE, we minimize the following performance index:
\[
J_o = -\sum_{i=1}^{m} \sum_{k=1}^{N} \ln f_i(\epsilon_i(k)).
\]
To minimize \(J_o\), take the derivative of the log likelihood \(J_o\) and use
\[
\frac{\partial J_o}{\partial \epsilon_i(k)} = -\frac{\partial J_o}{\partial \epsilon_i(k)} \frac{\partial \epsilon_i(k)}{\partial \epsilon_i(k)} - H_i^T
\]
\[
= -\sum_{i=1}^{m} \sum_{k=1}^{N} \frac{\partial \ln f_i(\epsilon_i(k))}{\partial \epsilon_i(k)} \frac{1}{\epsilon_i(k)} \epsilon_i(k) \frac{\partial \epsilon_i(k)}{\partial \epsilon_i(k)} = -\sum_{i=1}^{m} \sum_{k=1}^{N} w_i(k) \epsilon_i(k) H_i^T
\]
\[
= -\sum_{i=1}^{m} \sum_{k=1}^{N} \frac{(v_i + 1) \epsilon_i(k)}{\xi_i^2 v_i + (\epsilon_i(k))^2} H_i^T
\]
\[
= \Psi(\epsilon)
\]
where \(\epsilon = [\epsilon(1)^T \ldots \epsilon(N)^T]^T\) and
\[
w_i(k) = \frac{(v_i + 1)}{\xi_i^2 v_i + (\epsilon_i(k))^2}.
\]

**C. IF Illustration**

A textbook example \([12]\) is used to illustrate the IF. The measurement model is given as
\[
z(k) = \begin{bmatrix}
1.0 & 1.5 \\
0.5 & -2.5 \\
0 & 1.0 \\
1.0 & -0.5
\end{bmatrix} x(k) + \epsilon(k)
\]
where the noise \(\epsilon_k\) follows the \(t\)-distribution with \(v_i = 3\), \(\xi_i = 2\), and \(i = 1, \ldots, 4\). The measurements are \(z = [-3.01 3.52 -5.49 4.03 5.01]^T\).

According to (18), \(\Psi\) in this example is \(\Psi = [\Psi_1, \Psi_2]^T\), where
\[
\Psi_1 = \frac{(v_1 + 1) (z_1 - \hat{x}_1 - 1.5\hat{x}_2)}{\xi_1^2 v_1 + (z_1 - \hat{x}_1 - 1.5\hat{x}_2)^2}
+ 0.5 \frac{(v_2 + 1) (z_2 - 0.5\hat{x}_1 + 2.5\hat{x}_2)}{\xi_2^2 v_2 + (z_2 - 0.5\hat{x}_1 + 2.5\hat{x}_2)^2}
- 1.5 \frac{(v_3 + 1) (z_3 + 1.5\hat{x}_1 - 0.25\hat{x}_2)}{\xi_3^2 v_3 + (z_3 + 1.5\hat{x}_1 - 0.25\hat{x}_2)^2}
+ \frac{(v_5 + 1) (z_5 - \hat{x}_1 + 0.5\hat{x}_2)}{\xi_5^2 v_5 + (z_5 - \hat{x}_1 + 0.5\hat{x}_2)^2}
\]
\[
\Psi_2 = 1.5 \frac{(v_1 + 1) (z_1 - \hat{x}_1 - 1.5\hat{x}_2)}{\xi_1^2 v_1 + (z_1 - \hat{x}_1 - 1.5\hat{x}_2)^2}
- 2.5 \frac{(v_2 + 1) (z_2 - 0.5\hat{x}_1 + 2.5\hat{x}_2)}{\xi_2^2 v_2 + (z_2 - 0.5\hat{x}_1 + 2.5\hat{x}_2)^2}
+ 0.25 \frac{(v_4 + 1) (z_4 + 1.5\hat{x}_1 - 0.25\hat{x}_2)}{\xi_4^2 v_4 + (z_4 + 1.5\hat{x}_1 - 0.25\hat{x}_2)^2}
- \frac{(v_4 + 1) (z_4 - \hat{x}_2)}{\xi_4^2 v_4 + (z_4 - \hat{x}_2)^2}
\]
\[
\Psi_2 = 0.5 \frac{(v_5 + 1) (z_5 - \hat{x}_1 + 0.5\hat{x}_2)}{\xi_5^2 v_5 + (z_5 - \hat{x}_1 + 0.5\hat{x}_2)^2}
\]

Letting \(\Psi = 0\) gives the MLE solution \(\hat{x}_{\text{MLE}} = [3.005, -4.010]^T\). The red surfaces at the top of Fig. 2 shows the plots of \(\Psi_1\) and \(\Psi_2\) as functions of \(\hat{x}_1\) and \(\hat{x}_2\), whereas 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 (20) and (21) equals zero. The intersection between the two curves, as shown at the bottom of Fig. 2, is the solution to the MLE.

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

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

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

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

If the operating point $\bar{x} = [2, -3]^T$, then (23) gives $\hat{x}_{\text{Tay}} = [3.839, -5.066]^T$. The green surfaces in the first row of Fig. 3 shows the plots $\Psi$ 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 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 $\hat{x}_{\text{MLE}} = [3.050, -4.079]^T$. The approximate result using Taylor series expansion is $\hat{x}_{\text{Tay}} = [5.046, -5.628]^T$ 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 expectation, i.e., $\int_{-\infty}^{\infty} \frac{\partial \Psi}{\partial \hat{x}} f(\epsilon) d\epsilon$, (22) is changed to

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

Setting (24) equals zero gives

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

where the second term on the right-hand side is the IF [28]. If the value of the fifth measurement is still 5.01 and the other measurements are unchanged, an approximation of $\hat{x}_{\text{MLE}}$ using (25) is $\hat{x}_{\text{IF}} = [3.370, -4.250]^T$. The blue surfaces at the top of Fig. 5 shows the plots $\Psi$ 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 zero. The intersection between these two curves mentioned earlier 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 \( \hat{x}_{MLE} = [3.050, -4.079]^T \).

D. IF Approximation

The definition of the IF is given in [1], [28], and [29] as

\[
\Delta \hat{x} = \hat{x} - x \approx \text{IF}(\varepsilon)
\]

\[
= - \left[ \int_{-\infty}^{\infty} \frac{\partial \Psi(\varepsilon)}{\partial x} dF(\varepsilon) \right]^{-1} \Psi(\varepsilon)
\]  \hspace{1cm} (26)

where

\[
dF(\varepsilon) = f_1(\varepsilon_1(1)) \cdots f_m(\varepsilon_m(1)) \cdots f_1(\varepsilon_1(N)) \cdots \times f_m(\varepsilon_m(N))d\varepsilon_1(1) \cdots d\varepsilon_m(1) \cdots d\varepsilon_1(N) \cdots d\varepsilon_m(N)
\]

since the measurement noise \( \varepsilon_i \) is assumed to be independent and identically distributed.

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

\[
\hat{x} \approx x + \text{IF}(\varepsilon).
\]  \hspace{1cm} (27)

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

\[
\hat{x} \leftarrow \hat{x} + \text{IF}(E)
\]  \hspace{1cm} (28)

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

\[
\text{IF}(E) = (\hat{H}^T \Omega \hat{H})^{-1} \hat{H}^T \Lambda E.
\]  \hspace{1cm} (29)

The derivation of (29) and the definition of \( \Omega \) and \( \Lambda \) are given in Appendix.

The noise pdf can be obtained by analyzing the historical data. The robust estimators, such as MS, QL, SR, and SHGM estimators, need to calculate the inverse of matrix \( (\hat{H}^T W \hat{H}) \) 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 \( (\hat{H}^T \Omega \hat{H})^{-1} \) of the IF in (29) can be pre-computed because \( \Omega \) remains unchanged, thus saving computational load in real-time applications.

Unlike [17]–[19], the contribution of this paper is to employ the IF approximation so that the proposed estimator with \( t \)-distribution noise can be efficiently implemented within the framework (see 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 \( \Delta \). For the traditional robust estimator, \( \Delta = (\hat{H}^T W \hat{H})^{-1} \hat{H}^T W E \) according to (15). For the proposed estimator, \( \Delta = (\hat{H}^T \Omega \hat{H})^{-1} \hat{H}^T \Lambda E \) according to (29).

E. Estimation Precision

Precision of estimation is traditionally given by its variances. As given in [28], the variance of the state estimates can be approximated as

\[
\text{Var}(\hat{x}) = \text{Var}(\hat{x} - \bar{x}) = \int_{-\infty}^{\infty} \text{IF}(\varepsilon)\text{IF}(\varepsilon)^T dF(\varepsilon)
\]

\[
= \left[ \int_{-\infty}^{\infty} \frac{\partial \Psi(\varepsilon)}{\partial \bar{x}} dF(\varepsilon) \right]^{-1} \int_{-\infty}^{\infty} \Psi(\varepsilon)(\Psi(\varepsilon))^T dF(\varepsilon)
\]

\[
= \left[ \int_{-\infty}^{\infty} \frac{\partial \Psi(\varepsilon)}{\partial \bar{x}} dF(\varepsilon) \right]^{-T}.
\]  \hspace{1cm} (30)

Let the residue \( E \) be given by the noise \( \varepsilon \) in (10), (30) gives

\[
\text{Var} (\hat{x}) = \left[ \hat{H}^T \int_{-\infty}^{\infty} \frac{\partial W \varepsilon}{\partial \varepsilon} dF(\varepsilon) \hat{H} \right]^{-1}
\]

\[
\times \hat{H}^T \int_{-\infty}^{\infty} W \varepsilon \varepsilon^T W^T dF(\varepsilon) \hat{H}
\]

\[
\times \left[ \hat{H}^T \int_{-\infty}^{\infty} \frac{\partial W \varepsilon}{\partial \varepsilon} dF(\varepsilon) \hat{H} \right]^{-T}
\]

\[
= [\hat{H}^T \Omega \hat{H}]^{-1} [\hat{H}^T \Pi \hat{H}] [\hat{H}^T \Omega \hat{H}]^{-T}
\]  \hspace{1cm} (31)

where

\[
\Omega = \int_{-\infty}^{\infty} \frac{\partial W \varepsilon}{\partial \varepsilon} dF(\varepsilon)
\]

\[
= \text{diag}(\Omega_1(1), \ldots, \Omega_m(1), \ldots, \Omega_1(N), \ldots, \Omega_m(N))
\]

\[
\Pi = \int_{-\infty}^{\infty} W \varepsilon \varepsilon^T W^T dF(\varepsilon)
\]

\[
= \text{diag}(\Pi_1(1), \ldots, \Pi_m(1), \ldots, \Pi_1(N), \ldots, \Pi_m(N)).
\]  \hspace{1cm} (32)
Table II

| 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) | t₃(0.005) | t₃(0.005) |
| t-distribution | v₁ = 4, ξ₁ = 0.0046 | v₁ = 2, ξ₁ = 0.0042 | v₁ = 3, ξ₁ = 0.005 | v₁ = 3, ξ₁ = 0.000028 |
| Gaussian | σ₁ = 0.007 | σ₁ = 0.012 | σ₁ = 0.009 | σ₁ = 0.009 |

1 Estimator: LAV, WLS with LNR, QL(3), SHGM(3), SR(3), MS(3,4,5), the proposed estimator

\[ N(\sigma_i) = \frac{1}{\sqrt{2\pi \sigma_i^2}} \exp \left( -\frac{\xi_i^2}{2\sigma_i^2} \right) \]

Substituting (19) into (32) gives

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

where \( N_r \) is the number of simulation runs.

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

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

Taking the average gives

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

which can be approximated from the simulation as the AMSE

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

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] and [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 [30], [31]. 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 \( z_i, i = 1, \ldots, 58 \), comprising of 12 voltages \( (i = 1, \ldots, 12) \) and 46 currents \( (i = 13, \ldots, 58) \) are taken at each time instance and \( N = 3 \) sets of measurements are used to give one set of estimates. The measurement matrix \( \hat{H} \) is calculated according to the parameters in [31]. There are \( n = 28 \) states in the vector \( x = [V_{11}^r \cdots V_{14}^r V_{11}^i \cdots V_{14}^i]^T \). The real and imaginary part of the voltage phasor are defined as \( V_r \) and \( V_i \), respectively. The parameters for the QL and SR estimators are chosen as \( a_i = 3 \), whereas the parameters for the MS estimator are \( a_i = 3, b_i = 4 \), and \( r_i = 5 \) according to [5]. The penalty \( \kappa_i, i = 1, \ldots, 58 \), of the SHGM estimator is calculated according to [12] and [15]. Three different types of measurement noise pdfs are considered in the simulations. Note that the actual noises need not to be given by the t-distribution. There are studies in power systems where the Gaussian noises are corrupted by outliers with the uniform or Gaussian distribution [2], [3]. In Example 1, we model the Gaussian noise corrupted by outliers with the uniform or Gaussian distribution by the t-distribution.
TABLE III
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 (35)   |
|                     |           |     |             |    |    |    |      |     | Formula (34)      |
| 1G+1U               | \(^{3}\text{AMSE}\) | 2.95 | 2.39        | 2.79 | 2.74 | 2.84 | 2.78 | 2.51 | 2.08 | 2.04 |
|                     | \(^{1}\text{Proposed estimator’s improvement (\%)}\) | \(2.95-2.08\times100 = 29.5\) | 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 (\%) | \(9.90-2.25\times100 = 77.3\) | 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 | 2.35 | 2.46 |
|                     | Proposed estimator’s improvement (\%) | \(4.73-2.55\times100 = 46.1\) | 22.5 | 30.1 | 27.5 | 26.9 | 30.1 | 8.3 | 0 | 0 |

\(^{3}\text{AMSE unit: } \times10^{-6}\).
\(^{1}\text{Proposed estimator’s improvement (\%)} = \frac{(\text{Traditional estimator} - \text{Proposed estimator})}{\text{Traditional estimator}} \times 100; \text{ unit: } \%\).
\(^{2}\text{Proposed estimator’s improvement over itself is } 0\%\).

![Graph](image)

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

1) 1 Gaussian + 1 Uniform (1G+1U): First, the noise is associated with the pdf

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

(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 the 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 [33]. Fig. 7 shows how the Gaussian (\(\sigma_i = 0.007\)) and \(t\)-distribution (\(v_i = 4\) and \(\xi_i = 0.0046\)) can be fitted to the histogram using the maximum likelihood criterion according to [34]. The histogram represents the measurement noise generated by (37). In practice, the fitting process can be done offline using historical measurements.

The threshold parameters \(\delta\) and \(q_{\text{max}}\) are chosen as \(10^{-8}\) and 200, respectively. The AMSE results obtained by \(N_r = 10000\) simulations run using (35) and by the formula in (34) are given in Table III. The AMSEs of the proposed estimator obtained from simulation are verified by the formula in (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 \(2.95 - 2.08 \times 100\% = 29.5\%\). The improvement of the 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 section, the noise \(\epsilon_i\) is associated with the pdf [3], [28]

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

(38)

where \(\sigma_i = 0.005\). The first term in the above mentioned pdf represents the 95\% of the Gaussian noise, whereas the second term represents outliers by the 5\% Gaussian distribution with standard deviation \(10\sigma_i\). Following the above-mentioned section, the \(t\)-distribution with \(v_i = 2\) and \(\xi_i = 0.0042\) 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 = \(2.20 \times 10^{-6}\) of the proposed estimator is the smallest. More estimated results such as the mean-squared error (MSE) \(\text{MSE} = \frac{1}{N}\|\hat{x}_k - x_k\|^2\), \(\hat{V}_i\), and \(\hat{V}_{\text{im}}\) of the MS estimator and the proposed estimator are shown in Figs. 8 and 9(a) and (b), respectively. The ones obtained from the proposed estimator are closer to the true states.

3) \(t_3\) Noise: In this section, let the noise \(\epsilon_i\) be associated with the following \(t\)-distribution, which is commonly used to model noise with outliers [28]. The parameters in (16) are chosen as \(v_i = 3\) and \(\xi_i = 0.005\) for measurement noises and Table II
TABLE IV
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 (%) | $\frac{1.06 - 0.58}{1.06} \times 100$ | 22.7  | 30.1   | 27.5   | 26.6   | 30.9    | 9.4     | 30                 |
| 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: $\times 10^{-6}$.  
$\dagger$ Proposed estimator’s improvement over itself is 0%.

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 $= 2.55 \times 10^{-6}$ 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 proposed estimator from simulation using (35) are verified by the formula in (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 $v_i = 3$ and $\xi_i = 0.005$. The AMSE results are given in Table IV and it is clear that the 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 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 is given in Table IV. Except for WLS, WLS with LNR, and LAV estimators, the proposed estimator is faster. This is because the matrix $(\bar{H}^T \Omega \bar{H})^{-1}$ in the proposed estimator can be computed offline.

C. Example 3: Experiment in Microgrid

Experiments on a lab microgrid at NTU 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 were sampled at intervals of 5 s 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 s. 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.

**V. CONCLUSION**

In this paper, a robust state estimator using the $t$-distribution as the noise model is proposed. The IF 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.

**APPENDIX**

**DERIVATION OF (29)**

According to (10), (12), (18), and (26), $\varepsilon$ is replaced by $E$

$$\Psi(E) = \bar{H}^T WE$$

and then IF ($E$) is given as

$$IF(E) = - \left( \int_{-\infty}^{\infty} \frac{\partial \Psi(E)}{\partial E} dF(E) \right)^{-1} \Psi(E)$$

$$= \left[ \int_{-\infty}^{\infty} \frac{\partial (\bar{H}^T WE)}{\partial E} dF(E) \right]^{-1} \bar{H}^T WE$$

$$= \left( \bar{H}^T \int_{-\infty}^{\infty} \frac{\partial (WE)}{\partial E} dF(E) \bar{H} \right)^{-1} \bar{H}^T WE$$

$$\triangleq (\bar{H}^T \Omega \bar{H})^{-1} \bar{H} \Lambda E$$

where

$$\Omega = \int_{-\infty}^{\infty} \frac{\partial (WE)}{\partial E} dF(E)$$

$$= \text{diag}(\Omega_1(1), \ldots, \Omega_m(1), \ldots, \Omega_1(N), \ldots, \Omega_m(N))$$

$$\Omega_i(k) = \int_{-\infty}^{\infty} \frac{\partial \psi_i e_i}{\partial e_i} f_i(e_i) de_i$$

$$= \int_{-\infty}^{+\infty} \left( v_i + 1 \right) \left( \xi_i^2 v_i - e_i^2 \right) f_i(e_i) de_i$$

$$= \int_{-\infty}^{+\infty} \frac{\left( v_i + 1 \right) \left( \xi_i^2 v_i - e_i^2 \right)^2}{\left( \xi_i^2 v_i + e_i^2 \right)^2} f_i(e_i) de_i$$

(39)
\[ \Lambda = \text{diag}(\Lambda_\text{l}(1), \ldots, \Lambda_\text{l}(1), \ldots, \Lambda_\text{l}(N), \ldots, \Lambda_\text{l}(N)) \]
\[ \Lambda_\text{l}(k) = w_\text{l}(k) = \frac{e_i + 1}{(\epsilon_\text{i})^2 + (e_i(k))^2} \]

(40)

REFERENCES

[1] W. K. Ho, K. V. Ling, H. D. Vu, and X. Wang, “Filtering of the ARMAX process with generalized t-distribution noise: The influence function approach,” Ind. Eng. Chem. Res., vol. 53, no. 17, pp. 7019–7028, 2014.
[2] R. Baldick, K. Clements, Z. Pinjo-Dziga, 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, vol. 10, no. 2, pp. 1215–1224, Mar. 2019.
[4] V. Terzija et al., “Wide-area monitoring, protection, and control of future electric power networks,” Proc. IEEE, vol. 99, no. 1, pp. 80–93, Jan. 2011.
[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, Jul. 2014.
[6] P. Yang, Z. Tan, A. Wiesel, and A. Nehora, “Power system state estimation using 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. 9, no. 1, pp. 211–219, Jan. 2018.
[8] D. Wang and J. Romagnoli, “A framework for robust data reconciliation based on a generalized objective function,” Ind. Eng. Chem. Res., vol. 42, no. 13, pp. 3075–3084, 2003.
[9] V. Barnett and T. Lewis, Outliers in Statistical Data. New York, NY, USA: Wiley, 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 Syst., vol. PAS-90, no. 6, pp. 2718–2725, Nov. 1971.
[12] A. Abur and A. Gómez-Expósito, Power System State Estimation: Theory and Implementation. Boca Raton, FL, USA: CRC Press, 2004.
[13] A. Monticelli, State Estimation in Electric Power Systems: A Generalized Approach. New York, NY, USA: Springer-Verlag, 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., vol. 33, no. 5, pp. 6904–6914, Nov. 2018, doi: 10.1109/TPWRS.2018.2849325.
[15] 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.
[16] 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.
[17] S. Kotz and S. Nadarajah, Multivariate t-Distributions and Their Applications. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[18] P. Theodossiou, “Financial data and the skewed generalized t-distribution,” Manage. Sci., vol. 44, no. 12, pt. 1, pp. 1650–1661, 1998.
[19] D. Xu, C. Shen, and F. Shen, “A robust particle filtering algorithm with non-Gaussian measurement noise using student-t distribution,” IEEE Signal Process. Lett., vol. 21, no. 1, pp. 30–34, Jan. 2014.
[20] W. Zheng, W. Wu, A. Gomez-Exposito, 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.
[21] Y. Chen, F. Liu, S. Mei, and J. Ma, “A robust WLV state estimation using epsilon,” IEEE Trans. Power Syst., vol. 30, no. 4, pp. 2190–2191, Jul. 2015.
[22] K. Mahmoud and M. Abdel-Nasser, “Efficient SPF approach based on regression and correction models for active distribution systems,” IET Renewable Power Gener., vol. 11, no. 14, pp. 1778–1784, 2017.
[23] 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, Jul. 2017.
[24] W. K. Ho, T. Chen, K. V. Ling, and L. Sun, “Variance analysis of robust state estimation in power system using influence function,” Int. J. Elect. Power Energy Syst., vol. 92, pp. 53–62, 2017.
[25] 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.
[26] 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, Sep. 2012.
[27] J. Rashidzadeh, Power System State Estimation. Norwood, MA, USA: Artech House, 2012.
[28] F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel, Robust Statistics: The Approach Based on Influence Functions. Hoboken, NJ, USA: Wiley, 2011.
[29] P. Huber, Robust Statistics. New York, NY, USA: Wiley, 2009.
[30] C. Dufour and J. Blangero, “On the use of real-time simulation technology in smart grid research and development,” IEEE Trans. Ind. Appl., vol. 50, no. 6, pp. 3963–3970, Nov. 2014.
[31] D. Biau, 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 Proc. IEEE Power Energy Soc. Gen. Meeting, Jul. 2015, pp. 1–5.
[32] L. Dabbaghchi, “Power systems test case archive,” Univ. Washington, Seattle, WA, USA. Tech. Rep., 1993. [Online]. Available: http://www.ee.washington.edu/research/pstca/
[33] 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.
[34] 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. 273–283, 2015.

Tengpeng Chen was born in China and received the Ph.D. degree from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2017. He is currently an Assistant Professor with the Department of Instrumental and Electronic Engineering, Xi’an University, Xi’an, China. His research interests include the distributed algorithms with application in large-scale systems, power system state estimation, and power system dynamics.

Lu Sun is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, Faculty of Engineering, National University of Singapore, Singapore. He is currently a Student Research Fellow with the Cambridge Centre for Advanced Research and Education in Singapore, a wholly-owned subsidiary of the University of Cambridge, Cambridge, U.K., where his work focuses on the analysis and design of robust state estimators in modernized power systems. His research interests include the study of advanced control theory, robust statistics, and robust state estimation algorithms.

Keck-Voon Ling received the B.Eng. degree in electrical engineering from the National University of Singapore, Singapore, and the Dr. Phil. degree in control engineering from Oxford University, Oxford, U.K., in 1988 and 1992, respectively. He is currently an Associate Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research interests include model predictive control and moving horizon estimation, their embedded implementation, and applications to fast dynamic systems.

Weng Khuon Ho received the B.Eng. and Ph.D. degrees in electrical engineering from the National University of Singapore, Singapore, in 1987 and 1992, respectively. He was an Engineer with the Chartered Industry of Singapore from 1987 to 1988. He joined the Department of Electrical and Computer Engineering, National University of Singapore, in 1992. He served full-time National Service in the Singapore Armed Forces as an Infantry Officer from 1979 to 1981 and in the army reserve from 1981 to 2012. He has supervised more than 30 graduate students. His research interests include control and signal processing in semiconductor manufacturing and process control.

Dr. Ho was the recipient of the 2003 Best Paper Award for the IEEE Transactions on Semiconductor Manufacturing.