

This document is downloaded from DR-NTU, Nanyang Technological University Library, Singapore.

Title	Robust state estimation for power systems via moving horizon strategy
Author(s)	Chen, Tengpeng
Citation	Chen, T. (2017). Robust state estimation for power systems via moving horizon strategy. Sustainable Energy, Grids and Networks, 10, 46-54.
Date	2017
URL	http://hdl.handle.net/10220/42960
Rights	© 2017 Elsevier. This is the author created version of a work that has been peer reviewed and accepted for publication by Sustainable Energy, Grids and Networks, Elsevier. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [http://dx.doi.org/10.1016/j.segan.2017.02.005].

Robust State Estimation for Power Systems via Moving Horizon Strategy

Tengpeng Chen

School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore

Abstract

In this paper, I propose a re-weighted moving horizon estimation (RMHE) to improve the robustness for power systems. The RMHE reduces its sensitivity to the outliers by updating their error variances real-time and re-weighting their contributions adaptively for robust power system state estimation (PSSE). Compared with the common robust state estimators such as the Quadratic-Constant (QC), Quadratic-Linear (QL), Square-Root (SR), Multiple-Segment (MS) and Least Absolute Value (LAV) estimator, one advance of RMHE is that the RMHE incorporates the uncertainty of process model and the arrival cost term during the optimization process. Constraints on states are also taken into account. The influence of the outliers can be further mitigated. Simulations on the IEEE 14-bus system show that the RMHE can obtain estimated results with smaller errors even when the outliers are present.

Keywords: Robust State Estimation, Moving Horizon Estimation (MHE), Re-weighted, Outlier.

Notation

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

0.1. State Variables

- x State vector
- 5 \hat{x} Estimated state
- V_i^r Real part of the voltage phasor at bus i
- V_i^{im} Imaginary part of the voltage phasor at bus i
- w Process noise

0.2. Measurements and Noise

- 10 z Measurements from Phasor Measurement Unit
- v Measurement noise

Email address: chen0887@e.ntu.edu.sg (Tengpeng Chen)

0.3. Functions

$e_{i,k}$ The i -th measurement residual at time step k

$\rho(e_{i,k})$ Chosen function of $e_{i,k}$

15 J Cost function

$f_i(v_i)$ Probability density function of v_i

H Measurement matrix

$W_{i,k}$ Weighting factor for i -th measurement at time step k

Ψ Derivative of J wrt \hat{x}

20 0.4. Numbers and Others

m Number of measurements in 1 batch

n Number of states

N Number of batches

i Measurement index

25 t, k Time index

q Iteration index

(k) Iteration index in ADMM

a_i First threshold for traditional estimator i

b_i Second threshold for traditional estimator i

30 r_i Third threshold for traditional estimator i

σ_i Standard deviation of measurement noise v_i

R Diagonal matrix

P state covariance matrix

\mathbf{x} Vector $[\hat{x}_{t-N}^T \cdots \hat{x}_t^T]^T$

35 Z Vector $[z_{t-N}^T \cdots z_t^T]^T$

x^0 Operating point

ρ_0 The penalty parameter in ADMM algorithm

$r^{(k+1)}$ The primal residuals in ADMM algorithm

$s^{(k+1)}$ The dual residuals in ADMM algorithm

40 1. Introduction

The most common assumption of measurement noise used in power system state estimation (PSSE) is Gaussian. However, the Gaussian noise assumption is only an approximation to reality [1]. When the system meets transient data in steady-state measurements, instrument failure, human error or model nonlinearity [1, 2], non-Gaussian measurement error could be generated. Such outliers that
45 are far away from the expected measuring data raise the potential risk of misleading the estimation result [3]. The influence of bad data or outliers on the estimated results and one method to suppress the bad measurements during the iterative process has been proposed in [4]. Robust estimators with different objective functions such as the Quadratic-Constant (QC), Quadratic-Linear (QL), Square-Root (SR) and Multiple-Segment (MS) estimator have also been introduced to solve this
50 kind of problem [5, 6, 7]. Moreover, robust estimation has also been applied to such systems that all measurements are collected from phasor measurement units (PMUs) [8, 9, 10].

The Moving Horizon Estimation (MHE) aims to solve at each time instant an optimization problem by using a limited amount of most recent information [11]. The states are estimated by minimizing an overall objective function which consists of sensor model error, process model error, and error in
55 the state estimate at the beginning of the window [12]. The constraints on states have also been exploited in the optimization process. This can overcome the issues such as the suboptimal estimates or instability of the error dynamics [13]. By having these constraints in the optimization, MHE is more robust to the measurement outliers. [14] propose one kind of robust MHE. It generates a robust estimate by separately minimizing a set of least-squares cost functions, where the measurements af-
60 fected by outliers are left out. Finally the estimation result associated with the lowest cost is chosen. One drawback of this method is that the observability of estimator can not be guaranteed when some measurements are deleted.

In this paper, after combining the advantages of the MHE and the robust estimators such as QC, QL, SR and MS, we propose a re-weighted MHE (RMHE) algorithm for robust PSSE. The
65 RMHE uses the same method proposed in [15, 4, 16] to deal with the outliers, where the variances of outliers are updated online based on the measurements. The weights of the outliers will be mitigated but the observability of estimator is not influenced. Moreover, the constraints are exploited in the optimization process in order to alleviate the influence of outliers. In order to accelerate the performance of RMHE, the Alternating Direction Method of Multipliers (ADMM) is adapted to solve
70 the quadratic problem based on RMHE. The ADMM is a powerful algorithm for solving structured convex optimization problems. It provides a structured way of decomposing very large problems into smaller-subproblems that can be solved efficiently [17]. Numerical simulations with the IEEE 14-bus benchmark system show the effectiveness of RMHE.

This paper is organized as follows. The robust state estimation problem is formulated in Section
75 2. The RMHE algorithm is proposed in Section 3. The simulations on IEEE 14-bus system is given in Section 4. Finally the conclusions are made in Section 5.

2. Robust State Estimation

2.1. Measurement Model and State Equation

This paper use rectangular coordinates. The linear measurement model based on the PMUs [8] is given by

$$z_t = Hx_t + v_t, \quad (1)$$

where t is the time step and $z \in \mathbb{R}^m$ is the measurement vector composed of the real and imaginary components of bus voltage (or the line current) phasors. The state vector is given as $x = [V_1^r \dots V_n^r \ V_1^{im} \dots V_n^{im}]^T \in \mathbb{R}^n$, in which V_i^r and V_i^{im} ($i = 1, \dots, \frac{n}{2}$) are the real and imaginary components of the bus voltage phasors, respectively. v is assumed to be noise with zero mean.

Specially the i -th measurement is given by

$$z_{i,t} = H_i x_t + v_{i,t}, \quad (2)$$

where the subscript i is the index and v_i is uncorrelated between different measurements.

In this paper, the following assumptions are held:

Assumption 1. *The local state estimation is performed using the measurement data collected within the same system-wide updating time interval.*

Assumption 2. *The system is observable by PMUs and matrix $G = H^T H$ is full rank.*

The following simplified process model is considered for the state estimation problem [18, 19]:

$$x_{t+1} = Ax_t + w_t, \quad (3)$$

where A is assumed to be an identity matrix [18] and w_t represents the zero-mean disturbance with variance $Q > 0$.

2.2. Robust Estimators

In this section we will discuss different types of M-estimators [5]. A traditional power system may be considered as a quasi-static system [18, 20] because load demands change slowly and hence the state changes slowly, i.e. $x_{t-N} \approx \dots \approx x_t \approx x$. Given $N + 1$ sets of measurements $z_{i,k}$, $k = t - N, \dots, t$ collected from $i = 1, \dots, m$ measurements, the state x can be estimated by minimizing the cost function as follows:

$$J = \sum_{i=1}^m \sum_{k=t-N}^t \rho(e_{i,k}), \quad (4)$$

where $e_{i,k}$ is the measurement residual,

$$e_{i,k} = z_{i,k} - H_i \hat{x}. \quad (5)$$

100 Equation (5) gives $\frac{\partial e_{i,k}}{\partial \hat{x}} = -(H_i)^T$. Differentiating the above cost function (4) with respect to \hat{x} ,

$$\begin{aligned}
\frac{\partial J}{\partial \hat{x}} = \Psi(E) &= \frac{\partial J}{\partial e_{i,k}} \frac{\partial e_{i,k}}{\partial \hat{x}} \\
&= \sum_{i=1}^m \sum_{k=t-N}^t \frac{\partial \rho(e_{i,k})}{\partial e_{i,k}} \frac{1}{e_{i,k}} e_{i,k} \frac{\partial e_{i,k}}{\partial \hat{x}_k} \\
&= - \sum_{i=1}^m \sum_{k=t-N}^t W_{i,k} e_{i,k} (H_i)^T,
\end{aligned} \tag{6}$$

where

$$W_{i,k} = \sum_{i=1}^m \sum_{k=t-N}^t \frac{\partial \rho(e_{i,k})}{\partial e_{i,k}} \frac{1}{e_{i,k}}. \tag{7}$$

Using Equation (5), $\Psi(E)$ can also be written by

$$\begin{aligned}
\Psi(E) &= - \sum_{i=1}^m \sum_{k=t-N}^t W_{i,k} (z_{i,k} - H_i \hat{x}) (H_i)^T \\
&= -\bar{H}^T W E \\
&= -\bar{H}^T W (Z - \bar{H} \hat{x}),
\end{aligned} \tag{8}$$

where

$$\begin{aligned}
\bar{H} &= \left[H^T \dots H^T \right]^T \in \mathbb{R}^{(N+1)m \times n}, \\
Z &= \left[z_{t-N}^T \dots z_t^T \right]^T \in \mathbb{R}^{(N+1)m}, \\
E &= \left[e_{t-N}^T \dots e_t^T \right]^T \in \mathbb{R}^{(N+1)m}, \\
W &= \text{diag}(W_{1,t-N}, \dots, W_{m,t-N}, \dots, W_{1,t}, \dots, W_{m,t}) \\
&\in \mathbb{R}^{(N+1)m \times (N+1)m}.
\end{aligned}$$

To minimize the cost function (4), set $\Psi(E) = 0$ in (8). We have

$$\hat{x} = (\bar{H}^T W \bar{H})^{-1} \bar{H}^T W Z \tag{9}$$

where the matrix $\bar{H}^T W \bar{H}$ in (9) is an invertible matrix since we assume that the system is observable.

105 Using (7), the matrix W are updated and (9) is solved iteratively until the difference between the current and previous iteration for \hat{x} is less than a specified tolerance [5]. The diagonal matrix W for the MS, QC, QL and SR estimator can be obtained as follows.

2.2.1. *The Multiple-Segment, Quadratic-Constant, Quadratic-Linear and Weighted-Least-Squares Estimators*

110 The cost function of the MS estimator, sometimes known as the Hampel estimator [21], is given as

$$\rho(e_{i,k}) = \begin{cases} \frac{(e_{i,k})^2}{2\sigma_i^2} & |e_{i,k}| \leq a_i\sigma_i \\ \frac{a_i|e_{i,k}|}{\sigma_i} - \frac{a_i^2}{2} & a_i\sigma_i < |e_{i,k}| \leq b_i\sigma_i \\ -\frac{a_i(r_i\sigma_i - |e_{i,k}|)^2}{2(r_i - b_i)\sigma_i^2} \\ + \frac{1}{2}a_i r_i + \frac{1}{2}a_i b_i - \frac{1}{2}a_i^2 & b_i\sigma_i < |e_{i,k}| \leq r_i\sigma_i \\ \frac{1}{2}(a_i r_i + a_i b_i - a_i^2) & r_i\sigma_i < |e_{i,k}| \end{cases}$$

where a_i, b_i and r_i are thresholds selected by the user and σ_i is the variance of measurement $z_{i,t}$.

Differentiating the above cost function wrt $e_{i,k}$ and substituting into Equation (7) gives

$$W_{i,k} = \begin{cases} \frac{1}{\sigma_i^2} & |e_{i,k}| \leq a_i\sigma_i \\ \frac{a_i}{\sigma_i|e_{i,k}|} & a_i\sigma_i < |e_{i,k}| \leq b_i\sigma_i \\ \frac{a_i(r_i\sigma_i - |e_{i,k}|)}{(r_i - b_i)\sigma_i^2|e_{i,k}|} & b_i\sigma_i < |e_{i,k}| \leq r_i\sigma_i \\ 0 & r_i\sigma_i < |e_{i,k}| \end{cases} \quad (10)$$

Remark: Note that $a_i < b_i < r_i$. The MS estimator reduces to the QC estimator when $b_i \rightarrow a_i, r_i \rightarrow a_i$, the QL estimator when $b_i \rightarrow \infty, r_i \rightarrow \infty$, the weighted-least-squares (WLS) estimator when $a_i \rightarrow \infty, b_i \rightarrow \infty, r_i \rightarrow \infty$ and the Least Absolute Value (LAV) estimator when the $a_i \rightarrow 0, b_i \rightarrow \infty$ [5].

2.2.2. *The Square-Root Estimator*

The M-estimator using Square-Root (SR) function is given by

$$\rho(e_{i,k}) = \begin{cases} \frac{(e_{i,k})^2}{2\sigma_i^2} & |e_{i,k}| \leq a_i\sigma_i \\ 2a_i^{3/2} \sqrt{\frac{|e_{i,k}|}{\sigma_i}} - \frac{3}{2}a_i^2 & \text{otherwise} \end{cases}$$

120 Differentiating the above cost function wrt $e_{i,k}$ and substituting into (7) gives

$$W_{i,k} = \begin{cases} \frac{1}{\sigma_i^2} & |e_{i,k}| \leq a_i\sigma_i \\ \frac{a_i^3}{\sigma_i|e_{i,k}|^3} & \text{otherwise} \end{cases} \quad (11)$$

The cost functions for different estimators and their derivatives are respectively shown in Figs. 1 and 2, where the thresholds are set as $a_i = 3, b_i = 4$ and $r_i = 5$.

$W_{i,k}$ is re-weighted according to the measurement residuals $e_{i,k}$, which is calculated after each iteration of the estimation. The steps to implement the robust estimators can be summarized in the following pseudo-code:

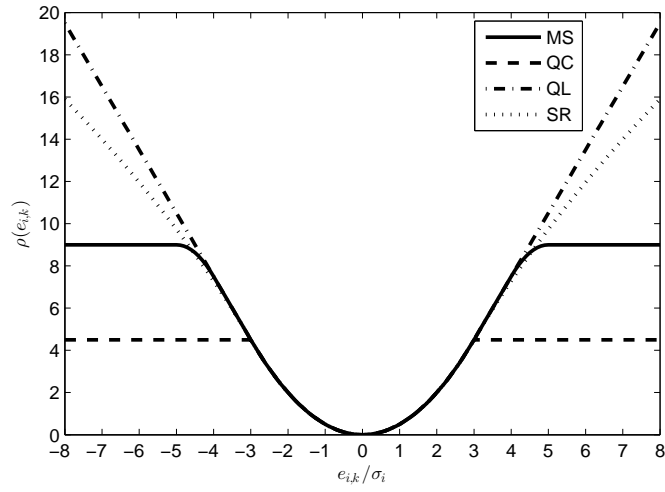


Figure 1: $\rho(e_{i,k})$ of the Quadratic-Constant (QC), Quadratic-Linear (QL), Multiple-Segment (MS) and Square-Root (SR) estimator.

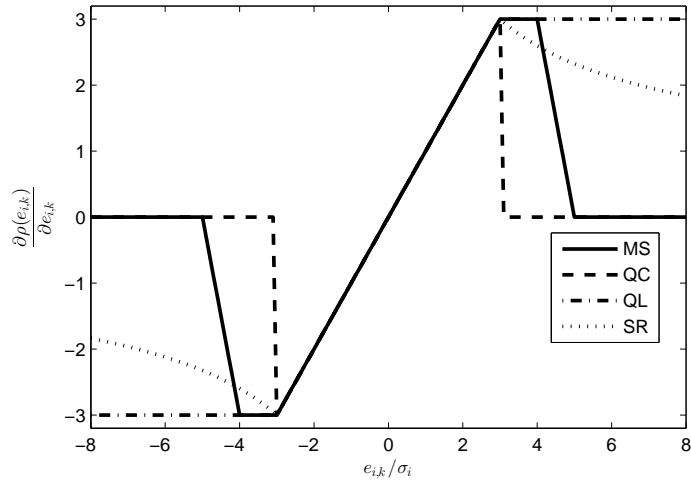


Figure 2: The derivative of $\rho(e_{i,k})$ for the Quadratic-Constant (QC), Quadratic-Linear (QL), Multiple-Segment (MS) and Square-Root (SR) estimator.

1. Initialization: Choose an initial estimate \hat{x}_0 .
2. Main procedure:
 - while** new measurement exists **do**
 - 130 (a) **if** $t \leq N$, set $t - N = t$, $\hat{x}_t = (H^T W H)^{-1} H^T W z_t$.
 - (b) **else**
 - (c) Set $q = 0$.
 - (d) Set $\hat{x}_t^q = \hat{x}_{t-1}$.
 - (e) **repeat**
 - 135 i. Set $q = q + 1$.
 - ii. Calculate $e_{j,k}$ using (5).
 - iii. Renew $W_{j,k}$ using (10) or (11).
 - iv. Calculate \hat{x}^q using (9).
 - (f) **until** $(\max(|\hat{x}_t^q - \hat{x}_t^{q-1}|) \leq \delta)$ or $q = q_{max}$.
 - 140 (g) **end if**
 - (h) Set $t \leftarrow t + 1$.
 - end while**

The iteration process will be terminated until the difference $(\hat{x}_t^q - \hat{x}_t^{q-1})$ reduces to a threshold δ or the iterative index q reaches a fixed number q_{max} . Similar methods are given in [15, 22].

3. Re-weighted Moving Horizon Estimation

3.1. RMHE

In this section we propose re-weighted MHE (RMHE) according to standard MHE discussed in [11] and then apply it to PSSE. The corresponding RMHE optimization problem is cast as

$$\Theta_t^* = \min_{\hat{x}_{t-N}^q, \dots, \hat{x}_t^q} \Psi_t(\hat{x}_{t-N}^q, \dots, \hat{x}_t^q), \quad (12)$$

150 subject to

$$\begin{aligned} \hat{x}_{k+1}^q &= A\hat{x}_k^q + \hat{w}_k^q, \quad k = t - N, \dots, t - 1 \\ z_k &= H\hat{x}_k^q + \hat{v}_k^q, \quad k = t - N, \dots, t \\ \hat{x}_k^q &\in \mathbb{X}, \end{aligned}$$

where \mathbb{X} are constraints defined by linear inequalities.

In the following, the notation $t - N|t - N - 1$ means the time step for prediction from step $t - N - 1$ to $t - N$. $\|\cdot\|_S^2$ indicates the square of the weighted Euclidean norm of a vector, defined as $\|x\|_S^2 = x^T S x$, where S is a positive definite matrix. The MHE objective function at the iteration q within the time step t is then calculated by

$$\Psi_t(\hat{x}_{t-N}^q, \dots, \hat{x}_t^q) = \frac{1}{2} \sum_{k=t-N}^t \|\hat{v}_k^q\|_{(R_k^q)^{-1}}^2 + \frac{1}{2} \sum_{k=t-N}^{t-1} \|\hat{w}_k^q\|_{Q^{-1}}^2 + \Phi_{t-N}, \quad (13)$$

Obviously the cost function consists of three error terms. The first term $\sum_{k=t-N}^t \|\hat{v}_k^q\|_{(R_k^q)^{-1}}^2$ is the error between the measurement model prediction and the raw measurement. The second term $\sum_{k=t-N}^{t-1} \|\hat{w}_k^q\|_{Q^{-1}}^2$ is the error between the estimated state $x_k, k = t-N, \dots, t-1$ and its process model. The third term Φ_{t-N} is named as the arrival cost and is given by

$$\Phi_{t-N} = \frac{1}{2} \|\hat{x}_{t-N} - \hat{x}_{t-N|t-N-1}\|_{P_{t-N|t-N-1}^{-1}}^2, \quad (14)$$

160 and R_k^q in (13) is given by

$$R_k^q = \text{diag}(R_{1,k}^q, \dots, R_{m,k}^q), \quad (15)$$

$$R_{j,k}^q = W_{j,k}^{-1} = \begin{cases} \sigma_j^2 & |e_{j,k}| \leq a_j \sigma_j \\ \frac{\sigma_j |e_{j,k}|}{a_j} & a_j \sigma_j < |e_{j,k}| \leq b_j \sigma_j \\ \frac{(r_j - b_j) \sigma_j^2 |e_{j,k}|}{a_j (r_j \sigma_j - |e_{j,k}|)} & b_j \sigma_j < |e_{j,k}| \leq r_j \sigma_j \\ +\infty & r_j \sigma_j < |e_{j,k}| \end{cases} \quad (16)$$

in which the measurement residuals are given by

$$e_{i,k} = z_{i,k} - H_i \hat{x}_k^{q-1}, k = t-N, \dots, t$$

For the arrival cost (14), we calculate $P_{t-N|t-N-1}$ from $P_{t-N-1|t-N-2}$ using the equation derived in [23]:

$$\begin{aligned} P_{t-N|t-N-1} &= AP_{t-N-1|t-N-2}A^T - AP_{t-N-1|t-N-2}H^T \\ &\quad \times (R + HP_{t-N-1|t-N-2}H^T)^{-1} \\ &\quad \times HP_{t-N-1|t-N-2}A^T + Q. \end{aligned} \quad (17)$$

3.2. ADMM for MHE

165 In this paper, we use the ADMM to accelerate the performance of RMHE. According to (1) and (3), separate (12) and define

$$\begin{aligned} \mathbf{x} &= \left[\hat{x}_{t-N}^T \cdots \hat{x}_t^T \right]^T \in \mathbb{R}^{(N+1)n} \\ Z &= \left[z_{t-N}^T \cdots z_t^T \right]^T \in \mathbb{R}^{(N+1)m} \\ \bar{Q} &= \mathcal{Y}_1^T (I_N \otimes Q^{-1}) \mathcal{Y}_1^T + \tilde{H}^T (I_{N+1} \otimes R^{-1}) \tilde{H} + \mathcal{Y}_2^T P_{t-N|t-N-1} \mathcal{Y}_2 \\ &\quad \in \mathbb{R}^{(N+1)n \times (N+1)n} \\ \bar{q}^T &= -Z^T (I_{N+1} \otimes R^{-1}) \tilde{H} - \mathcal{Y}_2^T P_{t-N|t-N-1} \hat{x}_{t-N|t-N-1} \in \mathbb{R}^{1 \times (N+1)n} \end{aligned}$$

$$\begin{aligned}
\tilde{H} &= I_{N+1} \otimes H \in \mathbb{R}^{(N+1)m \times (N+1)n} \\
\mathbb{K} &= 1_{N+1} \otimes \mathbb{X} \\
\mathcal{Y}_1 &= \begin{bmatrix} -A & I_n & 0 & \cdots & 0 & 0 \\ 0 & -A & I_n & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_n & 0 \\ 0 & 0 & 0 & \cdots & -A & I_n \end{bmatrix} \in \mathbb{R}^{(N+1)n \times (N+1)n} \\
\mathcal{Y}_2 &= \begin{bmatrix} I_n & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n \times (N+1)n}
\end{aligned}$$

where \otimes is the Kronecker product. Then the cost function (12) can be reformulated as a quadratic problem (QP) given as

$$\min_{\mathbf{x}} \frac{1}{2} \mathbf{x}^T \bar{Q} \mathbf{x} + \bar{q}^T \mathbf{x} \quad (18)$$

170 subject to

$$\mathbf{x} \in \mathbb{K} \text{ (or be defined as } C\mathbf{x} \leq c),$$

The QP problem (18) can be put on ADMM standard form by introducing a slack vector τ and putting an infinite penalty on negative components of τ , i.e.

$$\min_{\mathbf{x}} \frac{1}{2} \mathbf{x}^T \bar{Q} \mathbf{x} + \bar{q}^T \mathbf{x} + I_+(\tau) \quad (19)$$

subject to

$$C\mathbf{x} + \tau = c$$

where $I_+(\cdot)$ is the indicator function of the positive orthant defined as the following

$$I_+(\tau) = \begin{cases} 0 & \text{for } \tau \geq 0 \\ +\infty & \text{otherwise} \end{cases} \quad (20)$$

175 The associated augmented Lagrangian is

$$L_\rho(\mathbf{x}, \tau, u) = \frac{1}{2} \mathbf{x}^T \bar{Q} \mathbf{x} + \bar{q}^T \mathbf{x} + I_+(\tau) + \frac{\rho_0}{2} \|C\mathbf{x} + \tau - c + u\|^2 \quad (21)$$

where ρ_0 is the penalty parameter and $u = \phi/\rho_0$, which leads to the scaled ADMM iterations

$$\begin{aligned}
\mathbf{x}^{(k+1)} &= -(\bar{Q} + \rho_0 C^T C)^{-1} [\bar{q} + \rho_0 C^T (\tau^{(k)} + u^{(k)} - c)], \\
\tau^{(k+1)} &= \max\{0, -C\mathbf{x}^{(k+1)} - u^{(k)} + c\}, \\
u^{(k+1)} &= u^{(k)} + C\mathbf{x}^{(k+1)} - c + \tau^{(k+1)}
\end{aligned} \quad (22)$$

ADMM is particularly useful when the \mathbf{x} - and τ -minimizations can be carried out efficiently. The convergence of ADMM is often characterized in terms of the residuals

$$r^{(k+1)} = C\mathbf{x}^{(k+1)} + \tau^{(k+1)} - c, \quad (23)$$

$$s^{(k+1)} = \rho_0 C^T (\tau^{(k+1)} - \tau^{(k)}), \quad (24)$$

180 termed the primal and dual residuals, respectively [17]. The iteration of ADMM will be terminated until $r^{(k+1)}$ and $s^{(k+1)}$ reduce to some small tolerance.

In summary, the RMHE algorithm for power system is listed in the following pseudo code:

RMHE Algorithm for Power Systems

1. Initialization: Choose an initial estimate \hat{x}_0 . Set the length N and the variance matrices
185 Q , R , $P_{1|0}$, the iterative threshold δ_{RMHE} and number q_{max} .
2. Main procedure:
 - while** new measurement exists **do**
 - (a) **if** $t \leq N$, solve (12) with $\Phi_{t-N} = \frac{1}{2} \|\hat{x}_t - \hat{x}_{1|0}\|_{P_{1|0}^{-1}}^2$.
 - (b) **else**
 - (c) Set $q = 0$.
 - (d) Set $\hat{x}_t^q = \hat{x}_{t-1}$.
 - (e) **repeat**
 - i. Set $q = q + 1$.
 - ii. Calculate R_k^q using (16).
 - iii. Use (22) to solve (12) for \hat{x}_t^q .
 - (f) **until** ($\max(|\hat{x}_t^q - \hat{x}_t^{q-1}|) \leq \delta_{RMHE}$) or $q = q_{max}$.
 - (g) **end if**
 - (h) Set $t = t + 1$.
 - end while**

200

3.3. The Multiple-Segment Connection

The cost functions for the MS, QC, QL and SR estimators only include the error between the measurement model prediction and the raw measurement, while the cost function for RMHE are consisted of three parts. In this subsection we will connect the RMHE with the MS estimator to put
205 things in perspective.

3.3.1. $Q \rightarrow \infty$ and $N + 1 = 1$

If covariance matrix $Q \rightarrow \infty$, $P_{t-N|t-N-1}$ will also tend to ∞ according to (17). The arrival cost term in (13) will almost tend to 0. Moreover, if the horizon $N + 1 = 1$, then (13) will reduce to

$$\Psi_t(\hat{x}_t^q) = \frac{1}{2} \|\hat{v}_t^q\|_{(R_t^q)^{-1}}, \quad (25)$$

The solution to (25) can be given in the iteration form of

$$\hat{x}_t = (H^T(R_t^q)^{-1}H)^{-1}H^T(R_t^q)^{-1}z_t. \quad (26)$$

210 and it is totally the same as (9) for MS estimator. Note that the RMHE can also be reduced to the SR estimator if the matrix $(R_t^q)^{-1}$ is calculated according to (11).

3.3.2. $Q \rightarrow \infty$ and $N + 1 > 1$

In this case, the last two terms in the right side of (13) tend to 0 because both $P_{t-N|t-N-1}$ and Q become ∞ .

$$\Theta_t^* = \min_{\hat{x}_{t-N}^q, \dots, \hat{x}_t^q} \frac{1}{2} \sum_{k=t-N}^t \|\hat{v}_k^q\|_{(R_k^q)^{-1}}^2, \quad (27)$$

215 The solution is

$$\begin{bmatrix} \hat{x}_{t-N} \\ \vdots \\ \hat{x}_t \end{bmatrix} = (\tilde{H}^T W \tilde{H})^{-1} \tilde{H}^T W Z. \quad (28)$$

where W and Z are given in (8), and

$$\tilde{H} = \text{diag}(H, \dots, H) \quad (29)$$

The sequences $\hat{x}_{t-N}, \dots, \hat{x}_t$ are solved independently and actually \hat{x}_t is calculated using one set of measurement z_t , while the MS estimator uses $N + 1$ sets of measurements to get the estimated results \hat{x}_t . In this case the estimated results of MS estimator are better than RMHE's solution.

220 3.3.3. Q is a constant matrix and $N + 1 = 1$

In this case the second term in (13) are ignored and the cost function reduces to

$$\Psi_t(\hat{x}_{t-N}^q, \dots, \hat{x}_t^q) = \frac{1}{2} \sum_{k=t-N}^t \|\hat{v}_k^q\|_{(R_k^q)^{-1}}^2 + \Phi_{t-N} \quad (30)$$

As shown, the cost function includes the arrival cost term. The uncertainty of process model is expressed by the matrix Q . It will be better to combine the arrival cost because there will always be some uncertainty in practice.

225 3.3.4. Q is a constant matrix and $N + 1 > 1$

In this case the minimization problem estimates the states inside the window of data that spans the most recent $N + 1$ measurements. When new measurements come, RMHE shifts the window to discard oldest set of measurements and to include the new data. This is necessary to prevent the minimization problem from growing in size without bound and it can reduce the computational load.

230 Meanwhile, it is necessary to update the initial state in the window $x_{t-N+1|t-N}$ and its covariance $P_{t-N+1|t-N}$ from the previous window. The arrival cost is used to collect the influences of the historical information. Q is a constant matrix and links the sequences $\hat{x}_{t-N}, \dots, \hat{x}_t$ together and they are solved dependently.

3.4. WLS with Bad Data Processing

235 In this paper, the Largest normalized residuals (LNR) method [24] is used in WLS estimator to deal with bad data. The normalized residuals are calculated as:

$$\begin{aligned}
 \bar{R} &= \text{diag}(\sigma_1^2, \dots, \sigma_m^2, \dots, \sigma_1^2, \dots, \sigma_m^2) \\
 \bar{G} &= \bar{H}^T \bar{R}^{-1} \bar{H} \\
 \bar{\Omega} &= \bar{R} - \bar{H} \bar{G}^{-1} \bar{H}^T \\
 e_{i,k}^{norm} &= \frac{|e_{i,k}|}{\sqrt{\bar{\Omega}_{ii}}} \tag{31}
 \end{aligned}$$

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

$$Z_i^{new} = Z_i^{bad} - \frac{\bar{R}_{ii}}{\bar{\Omega}_{ii}} e_{i,k}^{bad} \tag{32}$$

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

4. Simulation Results

245 In this section, a simulation on IEEE 14-bus system using the RMHE algorithm is presented. The IEEE 14-bus system is shown in Figure 3 where the phase measurement units are placed according to [19]. Fifty-eight measurements, z_i , $i = 1, \dots, 58$, consisting of 12 voltages ($i = 1, \dots, 12$) and 46 currents ($i = 13, \dots, 58$) are taken at each time instance k . The measurement Jacobian matrix H is calculated according to the parameters in [25]. There are $n = 28$ states in the vector $x = [V_1^r \ V_2^r \ \dots \ V_{14}^r \ V_1^{im} \ \dots \ V_{14}^{im}]^T$. According to the references [18, 19], A is simplified as an identity matrix.

Define the Mean Square Error (MSE) according to [8] for the estimation results at time step k as

$$MSE_k = \sqrt{\frac{1}{n} \|\hat{x}_k - x_k\|^2}, \tag{33}$$

Denote the Average of Mean Square Error (AMSE) for $k \in [1, 60]$ to evaluate the estimation accuracy:

$$AMSE = \frac{1}{60} \sum_{k=1}^{60} \sqrt{\frac{1}{n} \|\hat{x}_k - x_k\|^2}. \tag{34}$$

255 For the voltage measurements $z_{i,t}$, $i = 1, \dots, 12$, the parameters for the the QC, QL and SR estimators are chosen as $a_i = 2.5$ [6, 7] while the parameters for MS estimator are $a_i = 2.5$, $b_i = 3.5$,

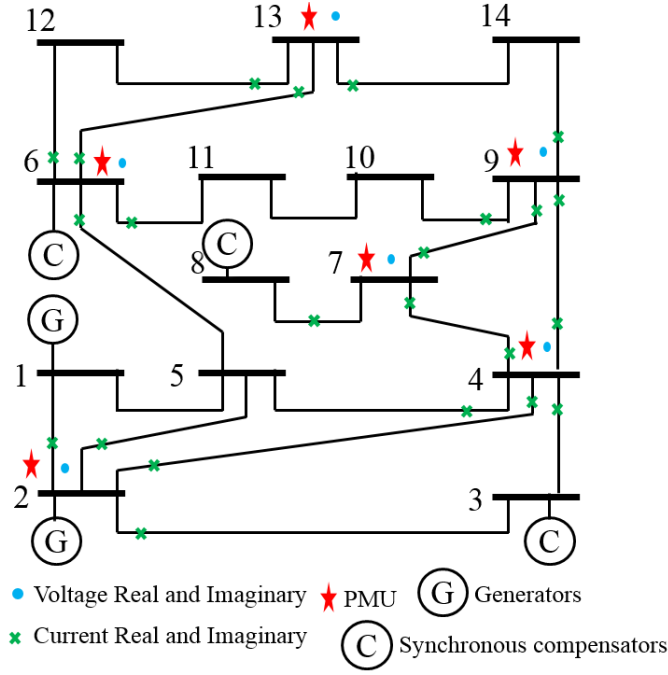


Figure 3: IEEE 14-bus system with Phase Measurement Units

$r_i = 4.5$. The noise v_i is associated with the probability density function [21] :

$$f_i(v_i) = \frac{0.97}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{v_i^2}{2\sigma_i^2}\right) + \frac{0.03}{\sqrt{2\pi(10\sigma_i)^2}} \exp\left(-\frac{v_i^2}{2(10\sigma_i)^2}\right)$$

where $\sigma_i = 0.005$ [26]. The first term accounted for 97% in $f_i(v_i)$ is the normal noise and the second part with larger standard deviations is assumed to be the outlier.

260 The uniform distribution is useful for modeling initial conditions, disturbances, and measurement errors that are equally likely to occur anywhere within a given interval [27]. For the current measurements $z_{i,t}$, $i = 13, \dots, 58$, the parameters for the the QC, QL and SR estimators are chosen as $a_i = 3$ while the parameters for MS estimator are $a_i = 3$, $b_i = 4$, $r_i = 5$. The noise v_i is associated with the probability density function [7] :

$$f_i(v_i) = \begin{cases} \frac{0.97}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{v_i^2}{2\sigma_i^2}\right) + \frac{0.03}{2 \times 10\sigma_i} & |v_i| \leq 10\sigma_i \\ \frac{0.97}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{v_i^2}{2\sigma_i^2}\right) & \text{otherwise} \end{cases}$$

265 where $\sigma_i = 0.01$ [26].

An n -dimensional column vector comprised of all ones or zeros is denoted by $\mathbf{1}_n$ and $\mathbf{0}_n$, respectively. The initialization parameters of RMHE algorithm are listed as follows:

- The initial state vectors $x_0 = [\mathbf{1}_{14}^T, \mathbf{0}_{14}^T]^T$.
- The initial variance matrix: $P_{1|0} = 10^2 \mathbf{I}_{28}$.
- 270 • The noise variances: $Q = 10^{-6} \mathbf{I}_{28}$; $R = \text{diag}(\sigma_1^2, \dots, \sigma_{58}^2)$.
- State constraints: $0.9 \leq \hat{V}_j^r \leq 1.2$, $-0.35 \leq \hat{V}_j^{im} \leq 0.01$, where $j = 1, \dots, 14$.

The comparison of AMSE under different estimators with different parameters are given in Tab. 1. Several examples are used to compare the propose RMHE and other estimators such as WLS and MS.

275 *4.1. Example 1*

4.1.1. $Q \rightarrow \infty$ and $N + 1 = 1$

In this case we assume that the measurement noises are Gaussian and the parameters a_i, b_i, r_i and Q are set to tend to ∞ . When the horizon length $N + 1$ is 1, The QC, QL, MS, SR, MHE and RMHE estimators reduce to WLS, where the AMSE is 2.4×10^{-3} . This verifies the results presented in Sections 2 and 3. The MSE results of WLS, MS and RMHE are shown in Fig. 4, where they are totally the same. In addition, the AMSE of LAV estimator is 2.5×10^{-3} and is larger than WLS. That verifies that the robust estimator LAV is not the best choice if the noise does not include many bad data.

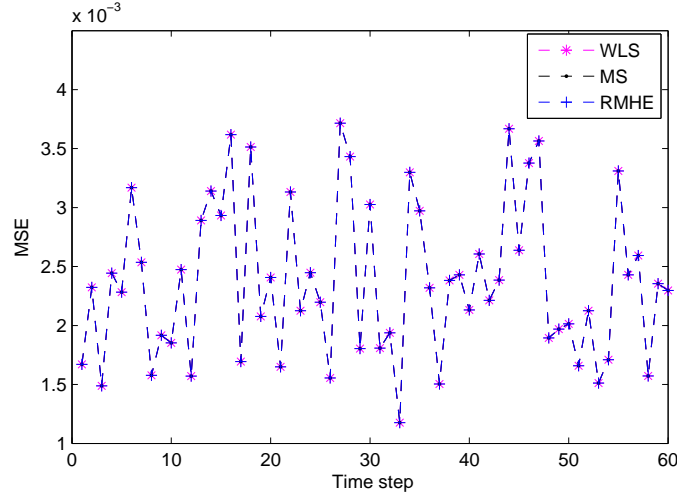


Figure 4: Mean square errors of WLS(1), MS(1) and RMHE ($Q \rightarrow \infty$ and $N + 1 = 1$) without constraints in Example 1.

4.1.2. $Q \rightarrow \infty$ and $N + 1 > 1$

285 When the horizon length is increased to 3 but matrix Q is still keep to ∞ , the estimated results of QC, QL, MS and SR are better than RMHE. In this case the RMHE is still a WLS estimator, as mentioned in previous section. However, the robust estimators use 3 sets of measurements at each time step and the QC, QL, MS and SR estimators use a potential assumption, $Q \rightarrow 0$, hence the estimated results are better. The MSE results of WLS without LNR, MS and RMHE are shown in Fig. 5, where MS is better than WLS and RMHE.

290

4.2. Example 2

Even though the parameters a_i, b_i and r_i are set to the limited numbers in Example 2, the AMSE for the QC, QL, MS and SR estimators are almost the same as the results in Example 1, because

Table 1: The AMSE[‡] for different estimators with different parameters in the simulation examples

Example	1		2		3	
Measurement i	1, ..., 12	13, ..., 58	1, ..., 12	13, ..., 58	1, ..., 12	13, ..., 58
Normal Noise	$*N(0.005)$	$N(0.01)$	$N(0.005)$	$N(0.01)$	$N(0.005)$	$N(0.01)$
Outlier	—	—	—	—	$N(0.05)$	$†U(0.1)$
Estimator	$‡(\infty, \infty, \infty)$	(∞, ∞, ∞)	$(2.5, 3.5, 4.5)$	$(3, 4, 5)$	$(2.5, 3.5, 4.5)$	$(3, 4, 5)$
$\#$ WLS (1)	2.4		2.4		3.7	
QC(1)	2.4		2.4		3.0	
QL(1)	2.4		2.4		3.0	
MS(1)	2.4		2.4		2.8	
SR(1)	2.4		2.4		2.8	
LAV(1)	2.5		2.5		3.0	
\S MHE($\infty, 1$)	2.4		2.4		3.7	
RMHE($\infty, 1$)	2.4		2.4		2.8	
MHE($10^{-5}I, 1$)	1.5		1.5		2.2	
RMHE($10^{-5}I, 1$)	1.5		1.5		1.6	
WLS with LNR (3)	1.6		1.6		1.8	
QC(3)	1.5		1.5		1.8	
QL(3)	1.5		1.5		1.8	
MS(3)	1.5		1.6		1.7	
SR(3)	1.5		1.5		1.7	
LAV(3)	1.6		1.6		1.8	
MHE($\infty, 3$)	2.4		2.4		3.7	
RMHE($\infty, 3$)	2.4		2.4		2.8	
MHE($10^{-6}I, 3$)	1.2		1.2		1.7	
RMHE($10^{-6}I, 3$)	1.2		1.2		1.4	
Fig.	4, 5		6		7, 8, 9	

$\#$ AMSE unit: $\times 10^{-3}$.

$$*N(\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{v_i^2}{2\sigma^2}}.$$

$$†U(\sigma) = \frac{1}{2\sigma} \text{ for } -\sigma \leq v_i \leq \sigma.$$

$‡$ Parameter a_i for QC, QL and SR are the first term of the bracket (a_i, b_i, r_i) , while parameters a_i, b_i, r_i for MS are given in the bracket of (a_i, b_i, r_i) .

$\#$ Parameter $N + 1$ for the traditional estimators are given in the bracket of $(N + 1)$.

\S Parameters Q and $N + 1$ for MHE and RMHE are given in the bracket of $(Q, N + 1)$.

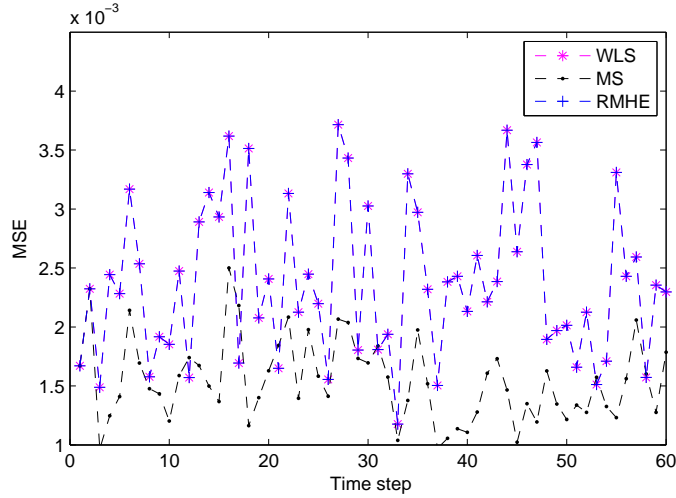


Figure 5: Mean square errors of WLS(1), MS(3) and RMHE ($Q \rightarrow \infty$ and $N + 1 > 1$) without constraints for Example 1.

outliers are not added in this example. When Q is equal to $10^{-5}I$ and $N + 1$ is 1, the AMSE results
of RMHE and MHE are better than the traditional estimators according to Table 1. The same
conclusion can also be draw when Q is equal to $10^{-6}I$ and $N + 1$ is 3. The AMSE results will be
smaller if more sets of measurements are used. The MSE results of WLS, MS and RMHE can also
be found in Fig. 6 and the results of RMHE is smallest, because the RMHE and MHE incorporate
more information such as the arrival cost and the error of process model. In addition, the re-weighted
process is not sensitive because the noise is still Gaussian, so the RMHE is equal to MHE.

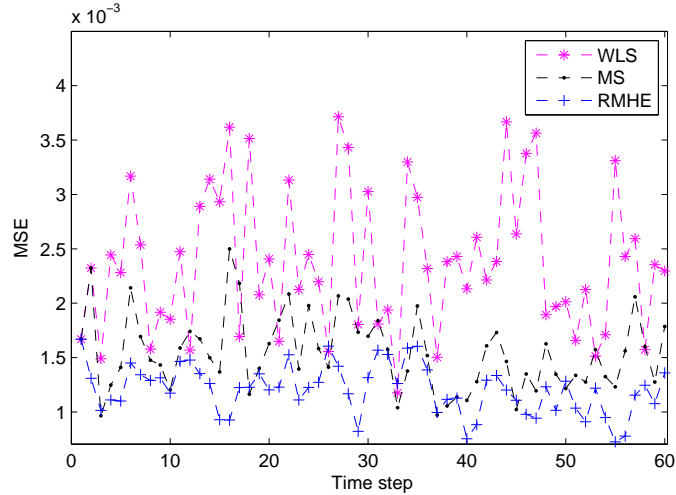


Figure 6: Mean square errors of WLS(1), MS(3) and RMHE ($Q = 10^{-6}I$, $N + 1 = 3$) without constraints for Example 2.

4.3. Example 3

In this example we assume that the measurements are added by non-Gaussian noise, the noise parameters can be found in Table 1. Moreover, we find that the AMSE results increase, compared

with the results of Examples 1 and 2. The re-weighted process of RMHE is efficient to reduce the outlier influence, compared with MHE. Even though the QC, QL, MS and SR also play well, all of them do not consider the information during the optimization process, i.e. the arrival cost and the constraints on states. The MSE results are given in Fig. 7 and 8. When the horizon length is set as $N + 1 = 1$, the AMSE of MS estimator is 2.8×10^{-3} and the result of RMHE is 1.6×10^{-3} through Table 1. The RMHE has an improvement of $(2.8 - 1.6)/2.8 \times 100\% = 43\%$ compared with MS estimator. The RMHE has an improvement of $(3.0 - 1.6)/3.0 \times 100\% = 47\%$ compared with LAV. Moreover, When the length is 3, the AMSE of MS estimator is 1.7×10^{-3} , the AMSE of LAV is 1.8×10^{-3} , while the results of RMHE is 1.4×10^{-3} . The RMHE has improvements of 18% and 22% to MS estimator and LAV estimator, respectively.

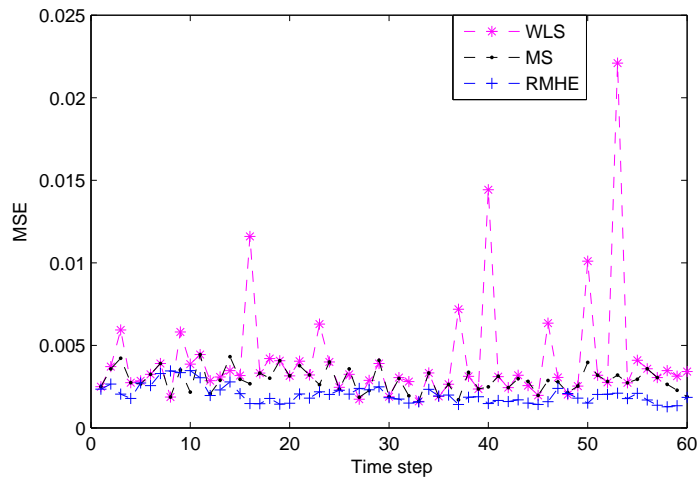


Figure 7: Mean square errors of WLS, MS and RMHE ($Q = 10^{-5}I, N + 1 = 1$) with constraints for Example 3.

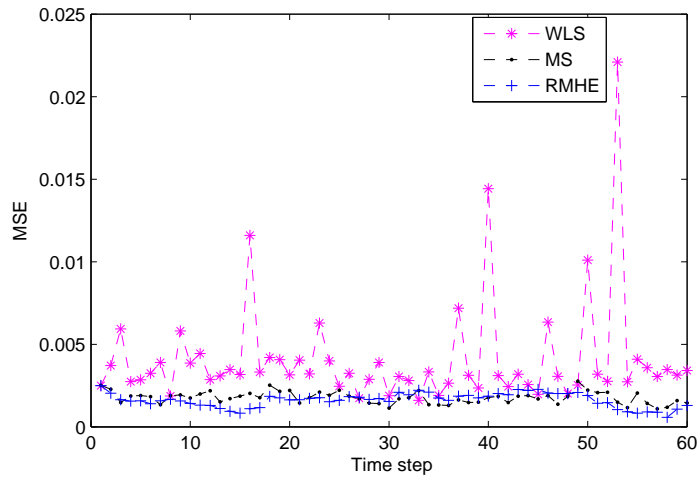


Figure 8: Mean square errors of WLS, MS and RMHE ($Q = 10^{-6}I, N + 1 = 3$) with constraints for Example 3.

Multiple bad-data is presented and happens at time step 23. Measurements, the real part of Bus

315 2 voltage phasor and the imaginary part of Bus 7 voltage phasor, are changed by 60%. From Fig. 9, we can see that WLS (without LNR) is affected seriously by the bad data, while robust estimators such as WLS with LNR (3), MS(3) and RMHE(3) can deal with the bad data and can still get good estimated results.

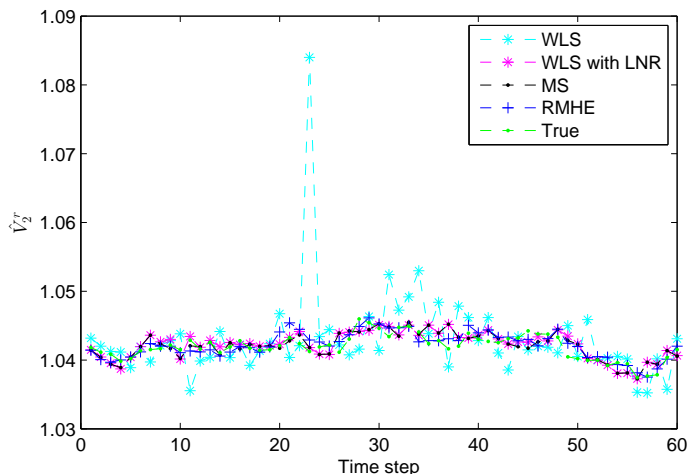


Figure 9: Real component of Bus 2 voltage phasor estimated by different estimators.

In addition, the simulation runs using Matlab version R2012b on a *i7* Windows 10 computer with
 320 8 GB RAM to test the computational time. The LAV estimator is built up according to the method provided by [8] and is conducted based on the matlab sub-function provided in GUROBI example. The computational time of different estimators are given in Table 2. The LAV takes the least time, where 0.012 seconds are needed for the LAV estimator to get one set of estimated results. But its results are worse than the results of RMHE, according to Table 1. The WLS with LNR takes 0.014
 325 seconds per step. For the traditional robust estimators, QC spends less time compared with QL, MS and SR estimator, due to the different chosen function ρ for estimators. The MHE takes only 0.024 seconds because no iteration process is considered. Finally we find that the RMHE estimator takes 0.05 seconds because it includes the iteration process to calculate the error variances online. Even though the RMHE takes more time than the WLS with LNR and LAV, the accuracy of RMHE is
 330 the highest. Due to the development of more powerful computation devices, it is able to faster the computation speed of RMHE. As a compromise between the computation speed and the accuracy of estimated results, the RMHE estimator to be implemented in power systems is acceptable.

5. Conclusion

In this paper a re-weighted moving horizon estimation (RMHE) is proposed for the robust power
 335 system state estimation, which adaptively estimates the system states along with an updated noise variance. ADMM is adapted to accelerate the performance of RMHE. Compared with the common estimators, two more error terms including the process model error and the error in the state estimate

Table 2: The computational time of different estimators in Example 3

Estimator	Total simulation time (seconds)	Average time(per step) (seconds)
WLS with LNR (3)	0.846	0.014
QC(3)	1.146	0.019
QL(3)	4.831	0.081
MS(3)	5.712	0.095
SR(3)	6.898	0.115
LAV(3)	0.738	0.012
MHE($10^{-6}I,3$)	1.453	0.024
RMHE($10^{-6}I,3$)	3.019	0.050

at the beginning of the window for RMHE are exploited to further mitigate the influence of the outliers. Simulations on the IEEE 14-bus system have shown that the RMHE is able to obtain estimation results with smaller errors, even in the presence of the outliers, compared with the results
340 obtained from the common robust state estimators such as the Quadratic-Constant (QC), Quadratic-Linear (QL), Square-Root (SR), Multiple-Segment (MS) and Least Absolute Value (LAV) estimator.

Acknowledgements

This work was supported by the Singapore National Research Foundation (NRF) under its Cam-
345 pus for Research Excellence And Technological Enterprize (CREATE) programme, and Cambridge Centre for Advanced Research in Energy Efficiency in Singapore (CARES), C4T project.

References

- [1] W. K. Ho, K. V. Ling, H. D. Vu, X. Wang, Filtering of the armax process with generalized t-distribution noise: The influence function approach, *Industrial & Engineering Chemistry Research* 53 (17) (2014) 7019–7028.
350
- [2] D. Wang, J. Romagnoli, A framework for robust data reconciliation based on a generalized objective function, *Industrial & engineering chemistry research* 42 (13) (2003) 3075–3084.
- [3] V. Barnett, T. Lewis, *Outliers in statistical data*, Wiley, New York, 1994.
- [4] H. Merrill, F. Schweppe, Bad data suppression in power system static state estimation, *IEEE Trans. Power App. Syst.* PAS-90 (6) (1971) 2718–2725. doi:10.1109/TPAS.1971.292925.
355
- [5] A. Abur, A. G. Expósito, *Power system state estimation: theory and implementation*, CRC Press, 2004.

- [6] E. Handschin, F. Schweppe, J. Kohlas, A. Fiechter, Bad data analysis for power system state estimation, *IEEE Trans. Power App. Syst.* 94 (2) (1975) 329–337. doi:10.1109/T-PAS.1975.31858.
- [7] R. Baldick, K. Clements, Z. Pinjo-Dzical, P. Davis, Implementing nonquadratic objective functions for state estimation and bad data rejection, *IEEE Trans. Power Syst.* 12 (1) (1997) 376–382. doi:10.1109/59.575722.
- [8] M. Göl, A. Abur, LAV based robust state estimation for systems measured by PMUs, *IEEE Trans. Smart Grid* 5 (4) (2014) 1808–1814. doi:10.1109/TSG.2014.2302213.
- [9] M. Göl, A. Abur, A robust PMU based three-phase state estimator using modal decoupling, *IEEE Trans. Power Syst.* 29 (5) (2014) 2292–2299. doi:10.1109/TPWRS.2014.2299756.
- [10] S. Chakrabarti, E. Kyriakides, PMU measurement uncertainty considerations in WLS state estimation, *IEEE Trans. Power Syst.* 24 (2) (2009) 1062–1071. doi:10.1109/TPWRS.2009.2016295.
- [11] C. V. Rao, J. B. Rawlings, J. H. Lee, Constrained linear state estimation moving horizon approach, *Automatica* 37 (10) (2001) 1619–1628.
- [12] R. Xiong, M. A. Grover, A modified moving horizon estimator for in situ sensing of a chemical vapor deposition process, *IEEE Trans. Contr. Syst. Technol.* 17 (5) (2009) 1228–1235. doi:10.1109/TCST.2009.2012519.
- [13] C. V. Rao, Moving horizon strategies for the constrained monitoring and control of nonlinear discrete-time systems, Ph.D. thesis, University of Wisconsin-Madison (2000).
- [14] M. Ellis, J. Zhang, J. Liu, P. D. Christofides, Robust moving horizon estimation based output feedback economic model predictive control, *Systems & Control Letters* 68 (2014) 101–109.
- [15] X. Li, A. Scaglione, Robust decentralized state estimation and tracking for power systems via network gossiping, *IEEE Journal on Selected Areas in Communications* 31 (7) (2013) 1184–1194. doi:10.1109/JSAC.2013.130703.
- [16] D. Falcao, P. Cooke, A. Brameller, Power system tracking state estimation and bad data processing, *IEEE Trans. Power App. Syst.* PAS-101 (2) (1982) 325–333. doi:10.1109/TPAS.1982.317110.
- [17] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Foundations and Trends® in Machine Learning* 3 (1) (2011) 1–122.
- [18] J. Zhang, G. Welch, G. Bishop, Observability and estimation uncertainty analysis for PMU placement alternatives, in: *Proc. North Amer. Power Symp.*, 2010, pp. 1–8.

- [19] P. Yang, Z. Tan, A. Wiesel, A. Nehora, Power system state estimation using PMUs with imperfect synchronization, *IEEE Trans. Power Syst.* 28 (4) (2013) 4162–4172. doi:10.1109/TPWRS.2013.2272220.
- [20] J. Machowski, J. Bialek, J. Bumby, Power system dynamics: stability and control, John Wiley & Sons, 2011.
- [21] F. Hampel, E. Ronchetti, P. Rousseeuw, W. Stahel, Robust statistics: The approach based on influence functions, Wiley, 2011.
- [22] E. Caro, R. Mínguez, A. J. Conejo, Robust WLS estimator using reweighting techniques for electric energy systems, *Electric Power Systems Research* 104 (2013) 9–17.
- [23] C. V. Rao, J. B. Rawlings, D. Q. Mayne, Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations, *IEEE Trans. Automat. Contr.* 48 (2) (2003) 246–258.
- [24] A. Monticelli, A. Garcia, Reliable bad data processing for real-time state estimation, *IEEE Trans. Power App. Syst.* PAS-102 (5) (1983) 1126–1139. doi:10.1109/TPAS.1983.318053.
- [25] I. Dabbaghi, Power systems test case archive, Tech. rep., University of Washington (1993). URL <http://www.ee.washington.edu/research/pstca/>
- [26] A. Gómez-Expósito, A. Abur, P. Rousseaux, A. de la Villa Jaén, C. Gómez-Quiles, On the use of PMUs in power system state estimation, in: *Proc. 17th Power Systems Computation Conference*, Vol. 22, 2011, p. 26.
- [27] D. G. Robertson, J. H. Lee, On the use of constraints in least squares estimation and control, *Automatica* 38 (7) (2002) 1113–1123.