L_p Quasi Norm State Estimator for Power Systems

Zhongliang Lyu, Hua Wei, Xiaoqing Bai, Daiyu Xie, Le Zhang, and Peijie Li

***Abstract*—This paper proposes an L_p ($0<p<1$) quasi norm state estimator for power system static state estimation. Compared with the existing L_1 and L_2 norm estimators, the proposed estimator can suppress the bad data more effectively. The robustness of the proposed estimator is discussed, and an analysis shows that its ability to suppress bad data increases as p decreases. Moreover, an algorithm is suggested to solve the nonconvex state estimation problem. By introducing a relaxation factor in the mathematical model of the proposed estimator, the algorithm can prevent the solution from converging to a local optimum as much as possible. Finally, simulations on a 3-bus DC system, the IEEE 14-bus and IEEE 300-bus systems as well as a 1204-bus provincial system verify the high computation efficiency and robustness of the proposed estimator.**

***Index Terms*—Power system bad data, quasi norm estimator, robustness, state estimation.**

I. Introduction

As one of the core applications of energy management systems (EMSs), state estimation (SE) is the basis of other advanced applications. This technique involves filtering random noises and provides an estimation of the most likely state of a system based on a redundant set of measurements that contain a certain degree of error.

The weighted least squares (WLS) [1]-[4], as a classic method implemented for power system SE, is an L_2 norm estimator, whose mathematical model is succinct and easy to solve. Unbiased estimation results can be obtained from measurements with Gaussian noise. However, the performance of the WLS may dramatically deteriorate in the presence of bad data, which have a significant influence on the estimation. Therefore, post-SE schemes for bad data identification are usually implemented with the WLS. The largest normalized residual (LNR) test [5]-[7] is one of those schemes, which can identify bad data effectively but cannot suitably handle multiple conforming errors.

In this regard, a series of robust SE algorithms were proposed [8]-[11]. One of them is the weighted least absolute value (WLAV) method, which is an L_1 norm estimator. However, the estimation results of this method are unsatisfactory when bad leveraged data exist. To solve this problem, the generalized-M (GM) estimators [12] and high-breakdown point estimators [13] have been proposed. Moreover, in [14], a scaling strategy was used in conjunction with the WLAV method to handle bad leveraged data. Simulation results indicated the effectiveness of the weighted least absolute value with scaling (WLAV-S) method in suppressing bad leveraged points. In [15], a robust iteratively reweighted least squares (IRLS) method for power system SE was proposed. Through the introduction of an adaptive weighting in the least squares (LS) estimation, the proposed method in [15] can mitigate the impact of bad data effectively.

In [16], based on the theory of uncertainty in measurement, a robust state estimator, known as the maximum normal measurement rate (MNMR) estimator, was proposed. Simulations showed that the MNMR estimator is highly robust to the existence of bad conforming or leveraged data, but it cannot filter the noise well for normal measurements. An estimation is deemed nonproblematic within the relevant interval. Based on the maximum correntropy criterion (MCC) [17], [18], the maximum exponential square (MES) estimator was proposed in [19], which can automatically suppress bad data, and its calculation is rapid. The maximum exponential absolute value (MEAV) estimator was defined in [20], [21], which can suppress different types of bad data robustly and efficiently. The MEAV and MES share the same theoretical foundation, and the main difference between them is that the former uses the Laplace kernel function, whereas the latter uses the Gaussian kernel function. However, the optimal solutions of the MEAV and MES are both sensitive to the Parzen window width. High-quality estimations cannot be obtained when this width is too large or too small.

In recent years, several signal analysis algorithms and image processing algorithms based on L_p quasi norm minimization have been widely studied and applied [22]-[26]. Compared with the L_1 and L_2 norms, the L_p quasi norm exhibits a higher rejection capability for invalid information in the aforementioned algorithms. However, this advantage of the L_p quasi norm has not been applied in the power system SE. Therefore, based on the existing literature [22]-[26], the L_p quasi norm is introduced into power system SE to solve the problem of low estimation quality caused by bad data in traditional norm estimators. Owing to the bad data rejection ca-

Manuscript received: June 15, 2020; revised: October 3, 2020; accepted: January 11, 2021. Date of CrossCheck: January 11, 2021. Date of online publication: May 17, 2021.

This work was supported by the National Natural Science Foundation of China (No. 51967002).

Z. Lyu, H. Wei (corresponding author), X. Bai, L. Zhang, and P. Li are with the School of Electrical Engineering, Guangxi University, Nanning, China, and they are also with the Guangxi Key Laboratory of Power System Optimization and Energy Technology, Guangxi University, Nanning, China (e-mail: lvzhongliang1992@163.com; weihuagxu@163.com; baixq@gxu.edu.cn; wjtclwwf@163.com; lipeijie@gxu.edu.cn).

D. Xie is with the Power Dispatching Control Center of Guangxi Power Grid, Nanning, China (e-mail: 371486756@qq.com).

DOI: 10.35833/MPCE.2020.000377

pability of the L_p ($p \to 0$) metric, the breakdown point of the proposed method is improved. As a result, the proposed method can perform better than the existing methods, such as least absolute value with scaling (LAV-S) method and IRLS method, in a low-redundancy environment with bad data. The main contributions of this paper are as follows.

1) An L_p quasi norm state estimator for power systems is presented, which can handle bad data more effectively than traditional norm estimators, and it can solve the problem of low estimation quality that results from bad data.

2) A relaxation algorithm for solving the non-convex L_p quasi norm SE problem is suggested. This algorithm can prevent the solution from converging to a local optimum because of measurement noise as much as possible. Hence, an effective implementation of the proposed estimator can be achieved.

3) After the analysis of computation efficiency and estimation quality of the proposed estimator, an approximately optimal value of p ($p = 0.1$) for practical implementation is provided.

The remainder of this paper is organized as follows. The L_p quasi norm SE problem is formulated in Section II. The characteristics of the proposed estimator are discussed and analyzed in Section III. Section IV presents the algorithm for solving the proposed SE problem. Simulation results are presented in Section V. Finally, the conclusions are given in Section VI.

II. Formulation of L_p Quasi Norm SE Problem

In the real vector space $\mathbb{R}^n$, the L_p quasi norm ($0 < p < 1$) [27] of the vector $\boldsymbol{x}$ ($\boldsymbol{x} \in \mathbb{R}^n$) is defined as:

$$\| \boldsymbol{x} \|_p^p = \sum_{i=1}^{n} |x_i|^p \tag{1}$$

where n is the number of elements in $\boldsymbol{x}$. Equation (1) describes a quasi norm as the expression violates the triangle inequality [28]-[30].

The measurement equation of SE problem can be described as:

$$\boldsymbol{z} = \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{r} \tag{2}$$

where $\boldsymbol{z} \in \mathbb{R}^m$ is the measurement vector, and m is the number of measurements; $\boldsymbol{r} \in \mathbb{R}^m$ is the measurement residual vector; and $\boldsymbol{g}(\boldsymbol{x}): \mathbb{R}^n \to \mathbb{R}^m$ is a vector of nonlinear functions.

According to (1) and (2), the nonlinear L_p quasi norm SE problem may be formulated as the following nonlinear programming problem:

$$\begin{cases} \min \boldsymbol{e}^{\mathrm{T}} |\boldsymbol{r}|^p \\ \text{s.t. } \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ \quad \boldsymbol{r} = \boldsymbol{z} - \boldsymbol{g}(\boldsymbol{x}) \end{cases} \tag{3}$$

where $\boldsymbol{h}(\boldsymbol{x}): \mathbb{R}^n \to \mathbb{R}^d$ is the zero-injection equality constraint; and $\boldsymbol{e} = [1, 1, \ldots, 1]^{\mathrm{T}} \in \mathbb{R}^m$.

III. Characteristics of L_p Quasi Norm State Estimator

In this section, we analyze the characteristics of the L_p quasi norm state estimator based on the definition of leveraged measurement, and we demonstrate that the proposed estimator is more robust than the least absolute value (LAV) estimator in the presence of bad leveraged data.

A. Robustness of L_p Quasi Norm State Estimator

When $p > 1$, the optimal condition of (3) can be expressed as:

$$\begin{aligned} &\frac{\partial(\boldsymbol{e}^{\mathrm{T}}|\boldsymbol{r}|^p)}{\partial \boldsymbol{x}} = \frac{\partial(\boldsymbol{e}^{\mathrm{T}}|\boldsymbol{r}|^p)}{\partial \boldsymbol{r}} \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}} = \frac{\partial(\boldsymbol{e}^{\mathrm{T}}|\boldsymbol{r}|^p)}{\partial \boldsymbol{r}} \frac{\partial(\boldsymbol{z} - \boldsymbol{g}(\boldsymbol{x}))}{\partial \boldsymbol{x}} = \\ &\quad -p \frac{\partial \boldsymbol{g}(\boldsymbol{x})}{\partial \boldsymbol{x}} \mathrm{diag}(\boldsymbol{e}) \mathrm{sign}(\boldsymbol{r}) |\boldsymbol{r}|^{p-1} = 0 \Rightarrow \\ &\quad \nabla \boldsymbol{g}(\boldsymbol{x}) \mathrm{diag}(\boldsymbol{e}) \mathrm{sign}(\boldsymbol{r}) |\boldsymbol{r}|^{p-1} = 0 \end{aligned} \tag{4}$$

where $\mathrm{sign}(\boldsymbol{r})$ is a diagonal function with elements $\mathrm{sign}(r_i)$, and $\mathrm{sign}(r_i) = 1$ when $r_i \geq 0$ and 0 otherwise.

When $p > 1 \to 1$, namely $p = 1^+$, (4) can be rewritten as (5), which represents the optimal condition of the LAV SE problem [31].

$$\nabla \boldsymbol{g}(\boldsymbol{x}) \mathrm{diag}(\boldsymbol{e}) \mathrm{sign}(\boldsymbol{r}) |\boldsymbol{r}|^{1^+ - 1} \approx \nabla \boldsymbol{g}(\boldsymbol{x}) \mathrm{sign}(\boldsymbol{r}) \boldsymbol{e} = 0 \tag{5}$$

The i^{th} column of $\nabla \boldsymbol{g}(\boldsymbol{x}) \mathrm{sign}(\boldsymbol{r})$ in (5) can be given as:

$$\nabla g_i(\boldsymbol{x}) \mathrm{sign}(r_i) \tag{6}$$

According to [19], if the values of elements in $\nabla g_i(\boldsymbol{x})$ are larger than those in other columns, the corresponding measurement is referred to as the leveraged measurement. As the LAV estimator cannot suppress the influence of $\nabla g_i(\boldsymbol{x})$ (the source of leveraged points), it cannot handle bad leveraged data.

When $p < 1$, (3) becomes an L_p quasi norm SE problem, which can be expressed as an LAV SE problem with self-adaptive weighting:

$$\begin{cases} \min \boldsymbol{e}^{\mathrm{T}} |\boldsymbol{r}|^p = \boldsymbol{e}^{\mathrm{T}} |\boldsymbol{r}|^{p-1} |\boldsymbol{r}| \\ \text{s.t. } \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ \quad \boldsymbol{r} = \boldsymbol{z} - \boldsymbol{g}(\boldsymbol{x}) \end{cases} \tag{7}$$

System (7) is similar in form to IRLS SE [15]. Self-adaptive weighting based on residuals is incorporated into both of their objective functions, which helps improve the robustness of the LAV and LS estimators. However, the difference between (7) and IRLS is clear. In IRLS, the self-adaptive weighting is implemented with a piecewise function and different weighting equations are set for measurements based on the values of residuals; consequently, bad data can be suppressed. In (7), $|\boldsymbol{r}|^{p-1}$ is a continuous weight function, which automatically changes with the values of residuals and filters out bad data at each iteration. In addition, (7) involves L_p quasi norm SE, whereby bad data can be suppressed. When $p \to 0$, the proposed estimator operates through an L_0 metric. This property enables the classification of normal and bad data effective through L_p quasi norm state estimator, thereby enhancing robustness. The purpose of extracting $|\boldsymbol{r}|^{p-1}$ from $|\boldsymbol{r}|^p$ is to analyze the characteristics of the LAV and L_p quasi norm state estimators, and then demonstrate that the proposed estimator is more robust than the LAV estimator. Moreover, L_p quasi norm SE problem cannot be solved directly in a noisy environment owing to the existence of multiple local optimums. Therefore, a relaxation model is proposed to handle this problem based on (7).

When $p<1$, (6) can be rewritten as:

$$\nabla g_i(\boldsymbol{x})\text{sign}(r_i)|r_i|^{p-1} \tag{8}$$

Figure 1 shows that the curves of $|r_i|^{p-1}$ varying with $|r_i|$ when $p<1$.

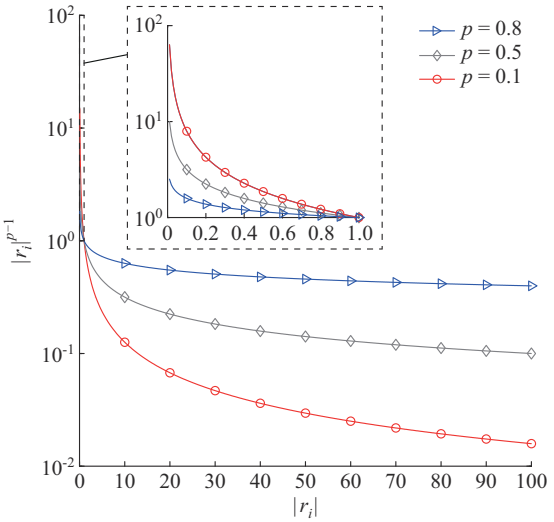

Fig. 1. Curves of $|r_i|^{p-1}$ varying with $|r_i|$ when $p<1$.

For a bad leveraged data, if $p=0.1$ and $r_i=100$, we have $|r_i|^{0.1-1}=100^{-0.9}=0.0158$, that is, the value of $\nabla g_i(\boldsymbol{x})|r_i|^{p-1}$ when $p=0.1$ is reduced by 0.0158 times than that when $p=1$. Hence, the influence of $\nabla g_i(\boldsymbol{x})$ can be suppressed, and the estimation results of L_p quasi norm SE are not affected by the bad leveraged data. For a good leveraged data, $\nabla g_i(\boldsymbol{x})$ is not downweighted by $|r_i|^{p-1}$. As a result, the improvement in estimation accuracy brought by the good leveraged data is not weakened but instead enhanced.

Based on the aforementioned discussion, $|r_i|^{p-1}$ when $p<1$ plays a role of self-adaptive weighting in (7). As p decreases, the effect of $|r_i|^{p-1}$ on $\nabla g_i(\boldsymbol{x})$ becomes stronger, and the estimator becomes more robust.

Since $|r_i|^{p-1}=1$ when $p=1$, the effect of self-adaptive weighting does not activate in the LAV estimator [32], [33]. Therefore, the proposed estimator is more robust than the LAV estimator in the presence of bad leveraged data.

To better demonstrate the robustness of the proposed estimator, a 3-bus DC system is used as an example. The topology of the system and the power flow results are shown in Fig. 2, where V_1, V_2, and V_3 are the node voltages; I_{12} and I_{23} are the line currents; P_1, P_2, and P_3 are the injected active power; and R_{12} and R_{23} are the line resistances.

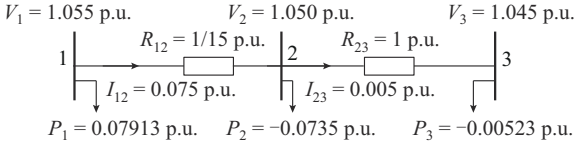

Fig. 2. Topology of 3-bus DC system and power flow results.

It is assumed that $z_1=1.045$, $z_2=-0.005$, $z_3=0.010$, and $z_4=14.175$ are the measurement values of V_2, I_{32}, I_{23}, and I_{21}, respectively, where z_4 ($R_{12}\ll R_{23}$) is the bad leveraged data. We aim to estimate the true state of the system.

To illustrate how the bad leveraged data affect the optimal system value on a two-dimension plane, the aforementioned SE problem is simplified without loss of generality. It is assumed that $V_1=1.055$ p.u. and $V_3=1.045$ p.u. are known. We estimate $V_2=x$ using the measurements z_1, z_2, z_3, and z_4.

Based on the power flow equation, we can obtain $r_1=z_1-g_1(x)=z_1-V_2=1.045-x$, $r_2=z_2-g_2(x)=z_2-(V_3-V_2)/R_{23}=x-1.050$, $r_3=z_3-g_3(x)=z_3-(V_2-V_3)/R_{23}=-x+1.055$, and $r_4=z_4-g_4(x)=z_4-(V_2-V_1)/R_{12}=-15x+30.000$ as the residual functions of V_2, I_{32}, I_{23}, and I_{21}, respectively.

According to (3), the aforementioned SE problem can be described as:

$$\min \overbrace{\left\{\overbrace{\sum_{i=1}^{3}|z_i-g_i(x)|^p}^{①}+\overbrace{|z_4-g_4(x)|^p}^{②}\right\}}^{③} \tag{9}$$

The curves of ①, ②, and ③, with an assumption that $p=1$, are depicted in Fig. 3. It can be observed that the minimum point of ③ resulting from the superposition of ① and ② deviates from the true value; hence, it does not appear in Fig. 3. The minimum point of ③ is determined by ② (bad leveraged data term). Therefore, the LAV estimator could not suppress the bad leveraged data.

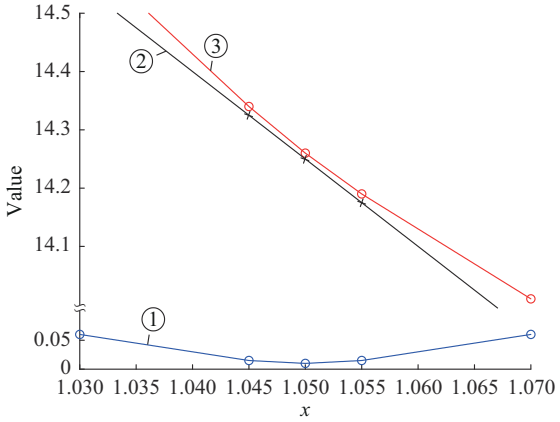

Fig. 3. Curves of ①, ②, and ③ when $p=1$.

The curves of ①, ②, and ③, with an assumption that $p=0.1$, are depicted in Fig. 4. It can be observed that the minimum point of curve ③ resulting from the superposition of ① and ② is close to the true value. This verifies that the breakdown points of L_p quasi norm state estimator are higher than those of the LAV estimator in the presence of bad leveraged data.

However, L_p quasi norm SE problem is a non-convex problem that possesses many local minimums when the measurements are noisy. As shown in Fig. 4, two local minimums appear at $x=1.045$ and $x=1.055$, respectively. It is hard to find the global optimum by solving this non-convex problem directly. More importantly, the statistical efficiency of the L_p quasi norm state estimator is relatively lower than those of the WLS and LAV estimators.

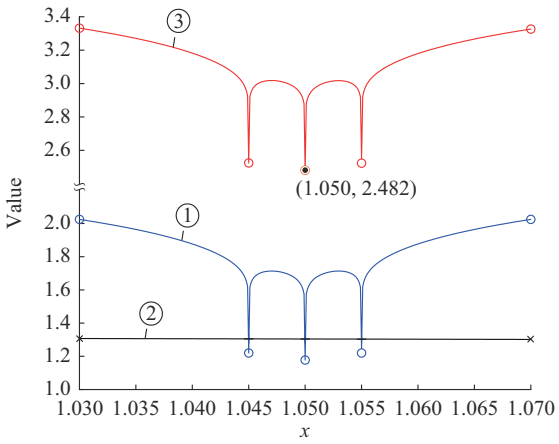

Fig. 4. Curves of ①, ②, and ③ when $p=0.1$.

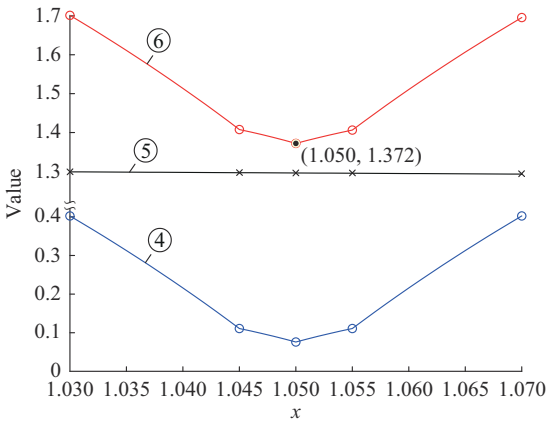

Fig. 5. Curves of ④, ⑤, and ⑥ when $p=0.1$.

B. Relaxation Model of L_p Quasi Norm SE Problem

To handle the aforementioned problems, a relaxation model of L_p quasi norm SE problem is proposed. By introducing a relaxation factor $\tau>0$, (7) can be rewritten as:

$$\begin{cases}\min \hat{\boldsymbol{\psi}}^{\mathrm{T}}(\boldsymbol{r})|\boldsymbol{r}| \\ \text{s.t. } \boldsymbol{h}(\boldsymbol{x})=\boldsymbol{0} \\ \quad \boldsymbol{r}=\boldsymbol{z}-\boldsymbol{g}(\boldsymbol{x})\end{cases} \tag{10}$$

where $\hat{\boldsymbol{\psi}}(\boldsymbol{r})=[\hat{\psi}(r_1), \hat{\psi}(r_2), \ldots, \hat{\psi}(r_i), \ldots, \hat{\psi}(r_m)]^{\mathrm{T}}$, and $\hat{\psi}(r_i)=(|r_i|+\tau)^{p-1}$.

When there are no gross errors, the vector $\boldsymbol{r}$ ($r_i \ll \tau$) becomes small and (11) can be obtained.

$$\begin{cases}\hat{\psi}(r_1)\approx\hat{\psi}(r_2)\approx\ldots\approx\hat{\psi}(r_m)\approx\hat{\psi}(\tau) \\ \hat{\boldsymbol{\psi}}(\boldsymbol{r})\approx[c_0, c_0, \ldots, c_0]^{\mathrm{T}}\end{cases} \tag{11}$$

Therefore, the modified L_p quasi norm SE problem is similar to the LAV SE problem when there are no gross errors. The statistical efficiency of the relaxation model is improved. Moreover, local optimums caused by the measurement noises do not appear in the relaxation model of L_p quasi norm SE problem.

To illustrate the effectiveness of the proposed relaxation model, we further analyze the simple example of the 3-bus DC system.

According to (10), the SE problem of the 3-bus DC system can be described as:

$$\min \overbrace{\left\{\overbrace{\sum_{i=1}^{3}\hat{\psi}(r_i)|z_i-g_i(x)|}^{④} + \overbrace{\hat{\psi}(r_4)|z_4-g_4(x)|}^{⑤}\right\}}^{⑥} \tag{12}$$

where $\hat{\psi}(r_i)=(|z_i-g_i(x)|+\tau)^{p-1}$.

The curves of ④, ⑤, and ⑥, with assumptions that $p=0.1$ and $\tau=0.1$, are depicted in Fig. 5. The minimum point of curve ⑥ resulting from the superposition of ④ and ⑤ is also close to the true value. More importantly, the local minimum does not appear in Fig. 5. This verifies that the proposed model can prevent a solution from converging to a local optimum caused by the measurement noises.

IV. Algorithm for Solving L_p Quasi Norm SE Problem

System (10) can be transformed into an equivalent model as:

$$\begin{cases}\min \hat{\boldsymbol{\psi}}^{\mathrm{T}}(\boldsymbol{r})\boldsymbol{r} \\ \text{s.t. } \boldsymbol{h}(\boldsymbol{x})=\boldsymbol{0} \\ \quad \boldsymbol{r}\geq\boldsymbol{z}-\boldsymbol{g}(\boldsymbol{x}) \\ \quad \boldsymbol{r}\geq\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{z}\end{cases} \tag{13}$$

By introducing two m-dimension vectors of positive slack variables $\boldsymbol{l}^+\in\mathbb{R}_+^m$ and $\boldsymbol{u}^+\in\mathbb{R}_+^m$, (13) can be transformed as:

$$\begin{cases}\min \hat{\boldsymbol{\psi}}^{\mathrm{T}}(\boldsymbol{r})\boldsymbol{r} \\ \text{s.t. } \boldsymbol{h}(\boldsymbol{x})=\boldsymbol{0} \\ \quad (\boldsymbol{l}^+,\boldsymbol{u}^+)\geq\boldsymbol{0} \\ \quad \boldsymbol{r}-\boldsymbol{z}+\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{l}^+=\boldsymbol{0} \\ \quad \boldsymbol{r}+\boldsymbol{z}-\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{u}^+=\boldsymbol{0}\end{cases} \tag{14}$$

From (14), we can obtain: $[(\boldsymbol{r}-\boldsymbol{z}+\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{l}^+)+(\boldsymbol{r}+\boldsymbol{z}-\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{u}^+)]/2=\boldsymbol{0}\Rightarrow\boldsymbol{r}=\boldsymbol{l}+\boldsymbol{u}$, $[(\boldsymbol{r}-\boldsymbol{z}+\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{l}^+)-(\boldsymbol{r}+\boldsymbol{z}-\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{u}^+)]/2=\boldsymbol{0}\Rightarrow\boldsymbol{z}-\boldsymbol{g}(\boldsymbol{x})+\boldsymbol{l}-\boldsymbol{u}=\boldsymbol{0}$. Hence, (14) can be transformed as:

$$\begin{cases}\min \hat{\boldsymbol{\psi}}^{\mathrm{T}}(\cdot)(\boldsymbol{l}+\boldsymbol{u}) \\ \text{s.t. } \boldsymbol{h}(\boldsymbol{x})=\boldsymbol{0} \\ \quad (\boldsymbol{l},\boldsymbol{u})\geq\boldsymbol{0} \\ \quad \boldsymbol{z}-\boldsymbol{g}(\boldsymbol{x})+\boldsymbol{l}-\boldsymbol{u}=\boldsymbol{0}\end{cases} \tag{15}$$

where $\hat{\boldsymbol{\psi}}(\cdot)=(\boldsymbol{l}+\boldsymbol{u}+\tau)^{p-1}$; $\boldsymbol{l}=\frac{1}{2}\boldsymbol{l}^+$; and $\boldsymbol{u}=\frac{1}{2}\boldsymbol{u}^+$.

The Lagrangian function associated with (15) can be defined as:

$$\begin{aligned}\boldsymbol{L}(\boldsymbol{x},\boldsymbol{l},\boldsymbol{u},\boldsymbol{\lambda},\boldsymbol{\pi},\boldsymbol{\alpha},\boldsymbol{\beta})=\hat{\boldsymbol{\psi}}^{\mathrm{T}}(\cdot)(\boldsymbol{l}+\boldsymbol{u})-\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x})- \\ \boldsymbol{\pi}^{\mathrm{T}}(\boldsymbol{z}-\boldsymbol{g}(\boldsymbol{x})+\boldsymbol{l}-\boldsymbol{u})-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{l}-\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{u}\end{aligned} \tag{16}$$

where $(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})\in\mathbb{R}^m$ and $\boldsymbol{\lambda}\in\mathbb{R}^d$ are the Lagrangian multipliers; and $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{l}$ and $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{u}$ are the complementarity terms, which can drive the complementarity equation of the Karush-Kuhn-Tucker (KKT) conditions. When $(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{l}+\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{u})\rightarrow 0$, the KKT conditions are satisfied, and the optimal solution of the primal problem is obtained [34].

According to the KKT conditions of (15), the KKT equations can be expressed as:

$$\begin{cases} \boldsymbol{L}_x = \partial \boldsymbol{L}/\partial \boldsymbol{x} = \nabla \boldsymbol{h}(\boldsymbol{x})\boldsymbol{\lambda} - \nabla \boldsymbol{g}(\boldsymbol{x})\boldsymbol{\pi} = \boldsymbol{0} \\ \boldsymbol{L}_l = \partial \boldsymbol{L}/\partial \boldsymbol{l} = p\hat{\boldsymbol{\psi}}(\cdot) - \tau\hat{\boldsymbol{\psi}}'(\cdot) - \boldsymbol{\pi} - \boldsymbol{\alpha} = \boldsymbol{0} \\ \boldsymbol{L}_u = \partial \boldsymbol{L}/\partial \boldsymbol{u} = p\hat{\boldsymbol{\psi}}(\cdot) - \tau\hat{\boldsymbol{\psi}}'(\cdot) + \boldsymbol{\pi} - \boldsymbol{\beta} = \boldsymbol{0} \\ \boldsymbol{L}_\lambda = \partial \boldsymbol{L}/\partial \boldsymbol{\lambda} = \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ \boldsymbol{L}_\pi = \partial \boldsymbol{L}/\partial \boldsymbol{\pi} = \boldsymbol{z} - \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{l} - \boldsymbol{u} = \boldsymbol{0} \end{cases} \tag{17}$$

$$\begin{cases} \boldsymbol{L}_\alpha = \boldsymbol{ALe} = \boldsymbol{0} \\ \boldsymbol{L}_\beta = \boldsymbol{BUe} = \boldsymbol{0} \end{cases} \tag{18}$$

$$(\boldsymbol{l}, \boldsymbol{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \geq \boldsymbol{0} \tag{19}$$

where $\boldsymbol{\psi}'(\cdot) = (p-1)(\boldsymbol{l}+\boldsymbol{u}+\tau)^{p-2}$; and $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{L}$, and $\boldsymbol{U}$ are the diagonal matrices with elements α_i, β_i, l_i, and u_i, respectively.

However, (18) cannot be solved directly using the Newton's method as:

$$\begin{cases} \boldsymbol{L}\Delta\boldsymbol{\alpha} + \boldsymbol{A}\Delta\boldsymbol{l} = -\boldsymbol{ALe} \\ \boldsymbol{U}\Delta\boldsymbol{\beta} + \boldsymbol{B}\Delta\boldsymbol{u} = -\boldsymbol{BUe} \end{cases} \tag{20}$$

At the k^{th} iteration, if $l_i^{(k)}$ becomes zero and $\alpha_i^{(k)} > 0$, we can obtain that $\Delta l_i^{(k)} = 0$ according to (20), indicating that l_i remains zero at all subsequent iterations. Such an undesirable attribute precludes the global convergence of the algorithm [34]. Hence, it is necessary to introduce a perturbed parameter $\mu > 0$ to relax (18), which is formulated as:

$$\begin{cases} \boldsymbol{L}_\alpha^\mu = \boldsymbol{ALe} - \mu\boldsymbol{e} = \boldsymbol{0} \\ \boldsymbol{L}_\beta^\mu = \boldsymbol{BUe} - \mu\boldsymbol{e} = \boldsymbol{0} \end{cases} \tag{21}$$

where $\mu = \sigma \cdot Gap/(2m)$, $Gap = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{l} + \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{u}$ [34], which is the complementarity gap, and $\sigma \in (0, 1)$ is the centering parameter.

By applying the Newton's method to solve the perturbed KKT equations, we can obtain:

$$\begin{cases} \boldsymbol{H}(\cdot)\Delta\boldsymbol{x} + \nabla\boldsymbol{h}(\boldsymbol{x})\Delta\boldsymbol{\lambda} - \nabla\boldsymbol{g}(\boldsymbol{x})\Delta\boldsymbol{\pi} = -\boldsymbol{L}_x \\ \boldsymbol{C}(\cdot)\Delta\boldsymbol{l} + \boldsymbol{C}(\cdot)\Delta\boldsymbol{u} - \Delta\boldsymbol{\pi} - \Delta\boldsymbol{\alpha} = -\boldsymbol{L}_l \\ \boldsymbol{C}(\cdot)\Delta\boldsymbol{l} + \boldsymbol{C}(\cdot)\Delta\boldsymbol{u} + \Delta\boldsymbol{\pi} - \Delta\boldsymbol{\beta} = -\boldsymbol{L}_u \\ \nabla^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x})\Delta\boldsymbol{x} = -\boldsymbol{L}_\lambda \\ -\nabla^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x})\Delta\boldsymbol{x} + \Delta\boldsymbol{l} - \Delta\boldsymbol{u} = -\boldsymbol{L}_\pi \end{cases} \tag{22}$$

$$\begin{cases} \boldsymbol{L}\Delta\boldsymbol{\alpha} + \boldsymbol{A}\Delta\boldsymbol{l} = -\boldsymbol{L}_\alpha^\mu \\ \boldsymbol{U}\Delta\boldsymbol{\beta} + \boldsymbol{B}\Delta\boldsymbol{u} = -\boldsymbol{L}_\beta^\mu \end{cases} \tag{23}$$

$$\begin{cases} \boldsymbol{H}(\cdot) = \nabla^2\boldsymbol{h}(\boldsymbol{x})\boldsymbol{\lambda} - \nabla^2\boldsymbol{g}(\boldsymbol{x})\boldsymbol{\pi} \\ \boldsymbol{C}(\cdot) = \mathrm{diag}(p\boldsymbol{\psi}'(\cdot) - \tau\hat{\boldsymbol{\psi}}''(\cdot)) \\ \hat{\boldsymbol{\psi}}''(\cdot) = (p-1)(p-2)(\boldsymbol{l}+\boldsymbol{u}+\tau)^{p-3} \end{cases} \tag{24}$$

We can obtain $[\Delta\boldsymbol{x}, \Delta\boldsymbol{\pi}, \Delta\boldsymbol{\lambda}, \Delta\boldsymbol{\alpha}, \Delta\boldsymbol{\beta}, \Delta\boldsymbol{l}, \Delta\boldsymbol{u}]^{\mathrm{T}}$ by solving (25).

$$\begin{bmatrix} \boldsymbol{H}(\cdot) & -\nabla\boldsymbol{g}(\boldsymbol{x}) & \nabla\boldsymbol{h}(\boldsymbol{x}) & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ -\nabla^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}) & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{E} & -\boldsymbol{E} \\ \nabla^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x}) & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{L} & \boldsymbol{0} & \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{U} & \boldsymbol{0} & \boldsymbol{B} \\ \boldsymbol{0} & -\boldsymbol{E} & \boldsymbol{0} & -\boldsymbol{E} & \boldsymbol{0} & \boldsymbol{C}(\cdot) & \boldsymbol{C}(\cdot) \\ \boldsymbol{0} & \boldsymbol{E} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{E} & \boldsymbol{C}(\cdot) & \boldsymbol{C}(\cdot) \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{x} \\ \Delta\boldsymbol{\pi} \\ \Delta\boldsymbol{\lambda} \\ \Delta\boldsymbol{\alpha} \\ \Delta\boldsymbol{\beta} \\ \Delta\boldsymbol{l} \\ \Delta\boldsymbol{u} \end{bmatrix} = -[\boldsymbol{L}_x \quad \boldsymbol{L}_\pi \quad \boldsymbol{L}_\lambda \quad \boldsymbol{L}_\alpha^\mu \quad \boldsymbol{L}_\beta^\mu \quad \boldsymbol{L}_l \quad \boldsymbol{L}_u]^{\mathrm{T}} \tag{25}$$

where $\boldsymbol{E}$ is a diagonal matrix with elements e_i.

To ensure $(\boldsymbol{l}, \boldsymbol{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}) > \boldsymbol{0}$, we can calculate the primal and dual step lengths (denoted by θ_P and θ_D, respectively) [8] by:

$$\begin{cases} \theta_P = 0.995\min\{\min_i(-l_i/\Delta l_i : \Delta l_i < 0; -u_i/\Delta u_i : \Delta u_i < 0), 1\} \\ \theta_D = 0.995\min\{\min_i(-\alpha_i/\Delta\alpha_i : \Delta\alpha_i < 0; -\beta_i/\Delta\beta_i : \Delta\beta_i < 0), 1\} \end{cases} \tag{26}$$

The procedure of solving the L_p quasi norm SE problem can be summarized as Algorithm 1.

Algorithm 1: procedure of solving L_p quasi norm SE problem

Initialization: choose the parameter p; set the iteration number $k = 0$ and the maximum iteration number $k_{\max}$; set the tolerance as $\varepsilon = 10^{-3}$ and the centering parameter as $\sigma \in (0, 1)$; choose $(\boldsymbol{l}, \boldsymbol{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}) > \boldsymbol{0}$, $\boldsymbol{\lambda} \neq \boldsymbol{0}$, $\boldsymbol{\pi} = \boldsymbol{0}$, and $\tau > 0$; set $\boldsymbol{x}$ as the flat start voltage of the generator bus.

Step 1: calculate the complementary gap $Gap = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{l} + \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{u}$. If $Gap < \varepsilon$ or $k > k_{\max}$, stop, and output the result; else, go to *Step 2*.

Step 2: calculate the perturbed parameter $\mu = \sigma \cdot Gap/2m$.

Step 3: obtain $[\Delta\boldsymbol{x}, \Delta\boldsymbol{\pi}, \Delta\boldsymbol{\lambda}, \Delta\boldsymbol{\alpha}, \Delta\boldsymbol{\beta}, \Delta\boldsymbol{l}, \Delta\boldsymbol{u}]^{\mathrm{T}}$ by solving (25).

Step 4: calculate θ_P and θ_D by (26).

Step 5: set the primal and dual variables as:

$$\begin{cases} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{l} \\ \boldsymbol{u} \end{bmatrix}^{(k+1)} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{l} \\ \boldsymbol{u} \end{bmatrix}^{(k)} + \theta_P \begin{bmatrix} \Delta\boldsymbol{x} \\ \Delta\boldsymbol{l} \\ \Delta\boldsymbol{u} \end{bmatrix} \\ \begin{bmatrix} \boldsymbol{\pi} \\ \boldsymbol{\lambda} \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix}^{(k+1)} = \begin{bmatrix} \boldsymbol{\pi} \\ \boldsymbol{\lambda} \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix}^{(k)} + \theta_D \begin{bmatrix} \Delta\boldsymbol{\pi} \\ \Delta\boldsymbol{\lambda} \\ \Delta\boldsymbol{\alpha} \\ \Delta\boldsymbol{\beta} \end{bmatrix} \end{cases}$$

Step 6: set $k = k + 1$, go to *Step 1*.

There are two remarks to be noticed as follows.

1) The parameter 0.995 is introduced for the primal and dual step lengths (θ_P, θ_D) to ensure the slack variables $\boldsymbol{l}$ and $\boldsymbol{u}$ satisfying $(\boldsymbol{l}, \boldsymbol{u}) > \boldsymbol{0}$ and the Lagrange multipliers $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ satisfying $(\boldsymbol{\alpha}, \boldsymbol{\beta}) > \boldsymbol{0}$. This mechanism can ensure the global convergence of the algorithm [34]. In theory, this parameter can be set to be any value between 0 and 1, and it is suggested between 0.99 and 0.9995 in [34]. For the L_p quasi norm SE problem, the algorithm performs well and almost the same when this parameter is in [0.99, 0.9995]. Hence, in this study, the value of 0.995 is used.

2) The relaxation factor τ is the key parameter in the L_p quasi norm SE algorithm. To ensure the performance of the proposed algorithm, a suitable value of $\tau > 0$ is needed.

From Section III-B, the relaxation factor τ can prevent the solution from converging to a local optimum caused by measurement noise. As a result, the statistical capabilities of the proposed algorithm are enhanced, and the convergence is improved. However, as τ increases, the L_p quasi norm state estimator reduces to an LAV estimator. Therefore, a suitable value of τ is needed to address this problem.

To illustrate how the relaxation factor τ affects the performance of the proposed algorithm, three cases are considered, namely nonconvergence, $k > 20$, and $k \leq 20$. Table I lists the number of the three cases when $\tau = 0.001$, 0.05, 0.1, and 0.2 on an IEEE 300-bus system (redundancy is 1.45) with a random Gaussian noise (standard deviation is 0.001 p.u.). The parameter p is set to be 0.1.

As shown in Table I, of the 1000 test cases, 978 cases are nonconvergent when $\tau = 0.001$, while all cases are convergent when $\tau = 0.05$, 0.1, and 0.2. Moreover, there are 56 and 0 cases of $k > 20$ when $\tau = 0.2$ and $\tau = 0.1$, respectively. Based on the result, τ can be set between 0.05 and 0.2. In this pa-

per, τ is set to be 0.1.

TABLE I
NUMBER OF DIFFERENT CASES WITH DIFFERENT τ

τ	Number of different cases		
	Nonconvergence	$k>20$	$k\leq 20$
0.001	978	22	0
0.050	0	56	944
0.100	0	0	1000
0.200	0	0	1000

As the relaxation model of L_p quasi norm SE is still considered as a non-convex optimal model, the proposed algorithm typically cannot guarantee a global optimum. However, in most cases, the quality of such a solution is comparable with that of the global optimal one.

V. SIMULATION RESULTS AND SELECTION OF p

In this section, simulations on a 3-bus DC system, the IEEE 14-bus and IEEE 300-bus systems as well as a 1204-bus provincial system are carried out to evaluate the performance of the proposed estimator. In addition, the procedure for choosing the exponent p is discussed. For comparison, WLS+LNR, quadratic tangent (QT) estimator (a GM estimator) [15], LAV-S [14], and the LAV are also implemented.

The algorithms are coded in MATLAB. The PC used for the simulations includes an Intel Core™ i5-4590S, 2.60 GHz processor, 4 GB RAM, and Win10 Ultimate.

A. 3-bus DC System with Bad Leveraged Data

The 3-bus DC system, which contains bad leveraged data, is simulated to evaluate the performance of the proposed estimator in suppressing bad leveraged data. The measurement configuration of the 3-bus DC system is shown in Fig. 6, where $P_{1\text{-}2}$ and $Q_{1\text{-}2}$ are the measurements of active and reactive power from bus 1 to bus 2, respectively, and set as bad data, which are the leveraged measurements; $\dot{V}_1$, $\dot{V}_2$, and $\dot{V}_3$ are the node voltages; and X_{12}, X_{13}, and X_{23} are the line reactances.

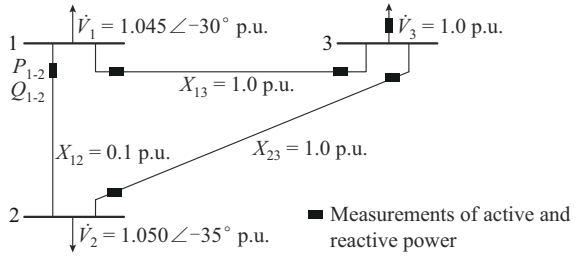

Fig. 6. Measurement configuration of 3-bus DC system.

The simulation results of 3-bus DC system are listed in Table II. As listed in Table II, since the WLS+LNR estimator fails to detect the bad leveraged data in this simulation, their adverse effect could not be mitigated in the estimation. The results obtained by the WLS+LNR estimator deviate from the true values. The estimation error of $Q_{1\text{-}2}$ obtained by this estimator is as high as 3189.29% of the true value and is similar to those of the measurement value. The LAV estimator also fails to obtain the correct result. The estimation error of $Q_{1\text{-}2}$ obtained by the LAV estimator is as high as 3273.28% of the true value. Instead, QT and LAV-S estimators successfully suppress the bad leveraged data.

TABLE II
SIMULATION RESULTS OF $P_{1\text{-}2}$ AND $Q_{1\text{-}2}$ IN 3-BUS DC SYSTEM

Estimator	$P_{1\text{-}2}$		$Q_{1\text{-}2}$	
	Value (p.u.)	Error (%)	Value (p.u.)	Error (%)
WLS+LNR	–4.1172	330.530	–2.9132	3189.290
QT	–0.9581	0.188	0.0950	0.742
LAV-S	–0.9573	0.105	0.0944	0.106
LAV	–4.1425	333.180	–2.9924	3273.280
$L_{0.9}$	–4.1369	332.600	–2.9853	3265.750
$L_{0.5}$	–4.1586	334.860	–2.9937	3274.660
$L_{0.1}$	–0.9567	0.042	0.0946	0.318
True value	–0.9563	0	0.0943	0
Measurement value	–4.1563	334.620	–2.9943	3275.290

For L_p quasi norm state estimator, owing to the self-adaptive weighting in (10), the bad leveraged data are downweighed sufficiently when $p=0.1$ (the robust mechanism is discussed in detail in Section III-A). Hence, the $L_{0.1}$ estimator obtains unbiased results, and the maximum estimation errors of $P_{1\text{-}2}$ and $Q_{1\text{-}2}$ do not exceed 0.318% of the true value.

B. IEEE 14-bus System with Conforming Bad Data

The IEEE 14-bus system, which contains conforming bad data, is simulated to evaluate the performance of the proposed estimator in suppresses conforming bad data. The measurement configuration of the IEEE 14-bus system is listed in Table III.

TABLE III
MEASUREMENT CONFIGURATION OF IEEE 14-BUS SYSTEM

Number of bus	Number of node voltage V_i	Number of lines	Number of line flow $P_{i\text{-}j}$ and $Q_{i\text{-}j}$	Number of injected power P_i and Q_i
14	14	20	68	28

P_1, Q_1, $P_{1\text{-}2}$, and $Q_{1\text{-}2}$ in the IEEE 14-bus system are the conforming measurements, which are all set to be half of their true values in this case. The simulation results of P_1, Q_1 and $P_{1\text{-}2}$, $Q_{1\text{-}2}$ are listed in Tables IV and V, respectively.

It can be observed that WLS+LNR estimator could not obtain the correct estimation results. In contrast, the other estimators successfully suppress the bad conforming data in the IEEE 14-bus system.

C. IEEE 14-bus System with Bad Critical Measurements

The IEEE 14-bus system, which contains bad critical measurements, is simulated to evaluate the performance of the proposed estimator in suppressing bad critical data.

To form critical measurements, only $P_{11\text{-}10}$ and $Q_{9\text{-}10}$ associated with state variables V_{10} and θ_{10} are used in the test system. Hence, $P_{11\text{-}10}$ and $Q_{9\text{-}10}$ become critical measurements, and $P_{11\text{-}10}$ is set as the bad data. The simulation results of $P_{11\text{-}10}$ and $Q_{9\text{-}10}$ are listed in Table VI.

TABLE IV
SIMULATION RESULTS OF P_1 AND Q_1 IN IEEE 14-BUS SYSTEM

Estimator	P_1		Q_1	
	Value (p.u.)	Error (%)	Value (p.u.)	Error (%)
WLS + LNR	1.5409	33.700	–0.1542	8.700
QT	2.3185	0.237	–0.1674	0.888
LAV-S	2.3230	0.043	–0.1692	0.178
LAV	2.3229	0.047	–0.1694	0.296
$L_{0.9}$	2.3245	0.022	–0.1699	0.592
$L_{0.5}$	2.3242	0.009	–0.1695	0.355
$L_{0.1}$	2.3250	0.043	–0.1687	0.118
True value	2.3240	0	–0.1689	0
Measurement value	1.1620	50.000	–0.0844	50.030

TABLE V
SIMULATION RESULTS OF $P_{1\text{-}2}$ AND $Q_{1\text{-}2}$ IN IEEE 14-BUS SYSTEM

Estimator	$P_{1\text{-}2}$		$Q_{1\text{-}2}$	
	Value (p.u.)	Error (%)	Value (p.u.)	Error (%)
WLS + LNR	0.9717	38.040	–0.1943	4.710
QT	1.5661	0.140	–0.2016	1.128
LAV-S	1.5677	0.038	–0.2034	0.245
LAV	1.5677	0.038	–0.2026	0.638
$L_{0.9}$	1.5687	0.026	–0.2035	0.196
$L_{0.5}$	1.5692	0.057	–0.2037	0.098
$L_{0.1}$	1.5694	0.070	–0.2038	0.049
True value	1.5683	0	–0.2039	0
Measurement value	0.7841	50.000	–0.1020	49.980

TABLE VI
SIMULATION RESULTS OF $P_{11\text{-}10}$ AND $Q_{9\text{-}10}$ IN IEEE 14-BUS SYSTEM

Estimator	$P_{11\text{-}10}$		$Q_{9\text{-}10}$	
	Value (p.u.)	Error (%)	Value (p.u.)	Error (%)
WLS + LNR	3.70700	9894.61	2.21100	5079.200
QT	3.70800	9897.30	1.56400	3563.620
LAV-S	3.70800	9897.30	2.12400	4875.400
LAV	3.70800	9897.30	3.75200	8688.940
$L_{0.9}$	3.70900	9900.00	2.54400	5859.240
$L_{0.5}$	3.70800	9897.30	3.00000	6927.400
$L_{0.1}$	3.70800	9897.30	1.54500	3519.110
True value	0.03709	0	0.04269	0
Measurement value	3.70900	9900.00	0.04249	0.468

As critical measurements are essential to the SE process, all the estimators in Table VI fail to provide correct estimation results in this case.

To further evaluate the performance of the proposed estimator in a low-redundancy environment, different numbers of measurements associated with state variables V_{10} and θ_{10} are used in the test system. Table VII shows the maximum amount of bad data associated with V_{10} and θ_{10} that the $L_{0.1}$ estimator can handle, which is denoted as N_{Bad}. In this simulation, N_{Bad} is close to round$((m_s - n_s)/2)$, where m_s is the number of measurements associated with the state variables V_{10} and θ_{10}; and $n_s = 2$ is the number of state variables V_{10} and θ_{10}.

TABLE VII
THE MAXIMUM AMOUNT OF BAD DATA

m_s	n_s	N_{Bad}
2	2	0
3	2	1
4	2	1
5	2	2
6	2	2
7	2	2
8	2	3

When $p = 0.1$, the cost function of the proposed estimator becomes an $L_{0.1}$ ($L_{0.1} \rightarrow L_0$) metric. This property is helpful in SE with bad data. Thus, the $L_{0.1}$ estimator can also perform well in a low-redundancy environment.

D. IEEE 300-bus System with Gaussian Noise

The statistical efficiency is an important criterion to evaluate the performance of an estimator. Thus, two groups (G1 and G2) of IEEE 300-bus system with Gaussian noise and different measurement redundancies are simulated. The standard deviation of noise is set as [0.0002,0.0010]p.u.. The measurement configurations of G1 and G2 are listed in Table VIII.

TABLE VIII
MEASUREMENT CONFIGURATIONS OF G1 AND G2

Group	Number of node voltage V_i	Number of line flow $P_{i\text{-}j}$ and $Q_{i\text{-}j}$	Number of injected power P_i and Q_i	Redundancy
G1	300	1208	600	3.51
G2	149	406	314	1.45

Formula (27) is used to evaluate the estimation results.

$$\begin{cases} \bar{S} = \dfrac{1}{m}\sum_{i=1}^{m}(\hat{z}_i - \check{z}_i)^2 \\ S_{\max} = \max\left|\hat{z}_i - \check{z}_i\right| \\ \bar{Q} = \dfrac{1}{n}\sum_{i=1}^{n}(\hat{x}_i - \check{x}_i)^2 \\ Q_{\max} = \max\left|\hat{x}_i - \check{x}_i\right| \end{cases} \tag{27}$$

where $\check{z}_i$ and $\hat{z}_i$ are the true and estimation values of z_i, respectively, and they do not contain phase angles; $\check{x}_i$ and $\hat{x}_i$ are the true and estimation values of the state variables, respectively; $\bar{S}$ and $\bar{Q}$ are the average squared estimation errors of measurements and state variables, respectively; and $S_{\max}$ and $Q_{\max}$ are the maximum estimation errors of measurements and state variables, respectively.

The average values of $\bar{S}$ and $S_{\max}$ from 1000 simulations on G1 and G2 with Gaussian noise are shown in Figs. 7 and 8, respectively.

The average values of $\bar{Q}$ and $Q_{\max}$ from 1000 simulations on G1 and G2 with Gaussian noise are shown in Figs. 9 and 10, respectively.

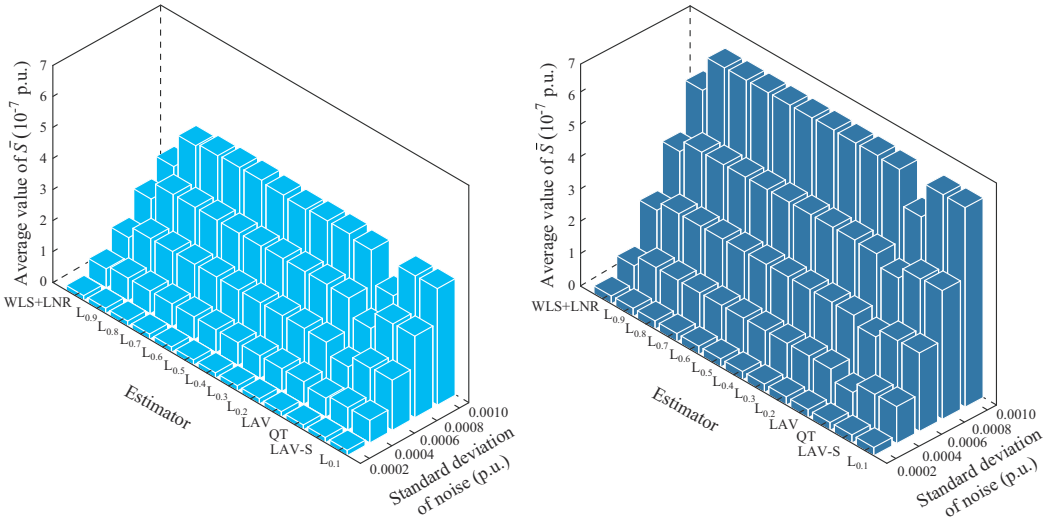

Fig. 7. Average value of $\bar{S}$ from 1000 simulations on G1 and G2. (a) G1. (b) G2.

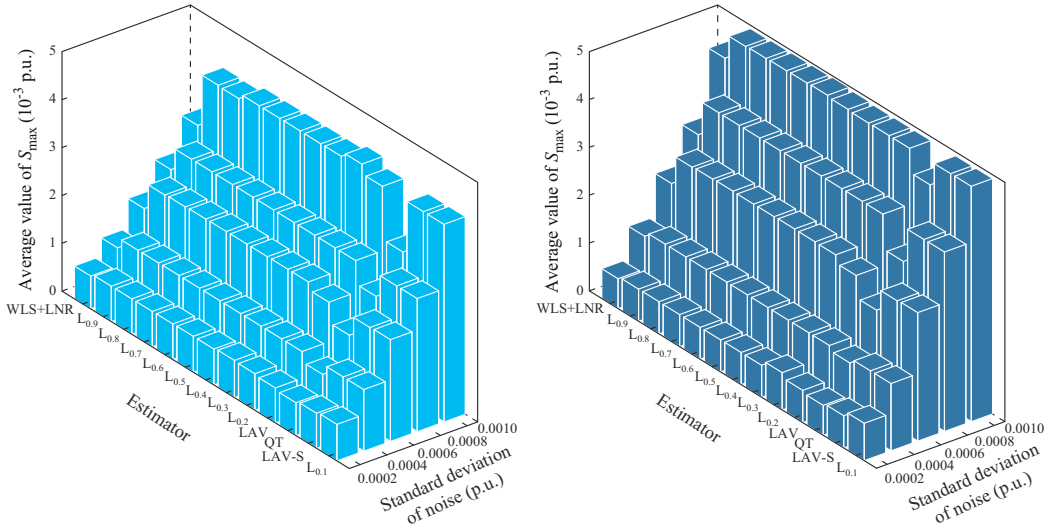

Fig. 8. Average value of S_{max} from 1000 simulations on G1 and G2. (a) G1. (b) G2.

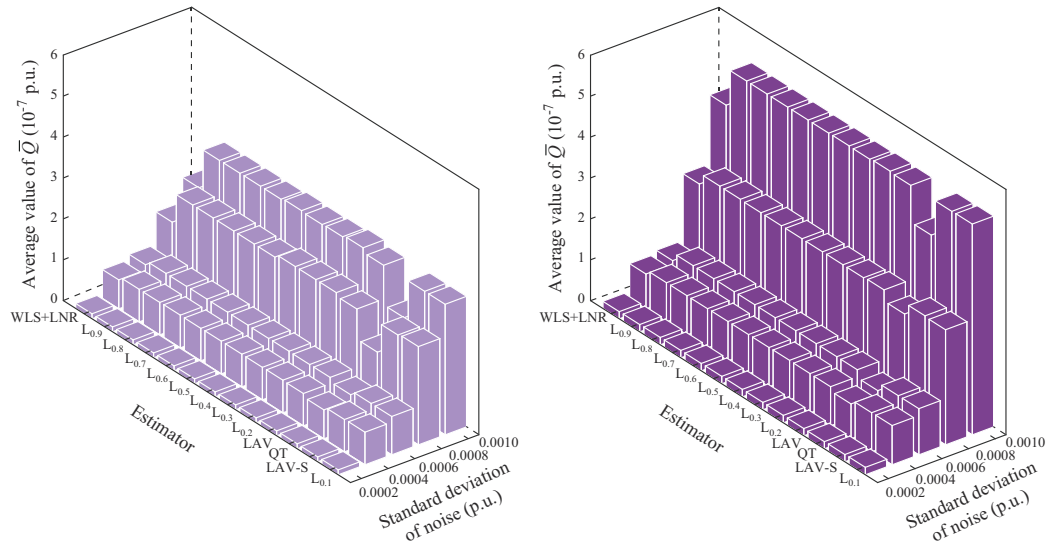

Fig. 9. Average value of $\bar{Q}$ from 1000 simulations on G1 and G2. (a) G1. (b) G2.

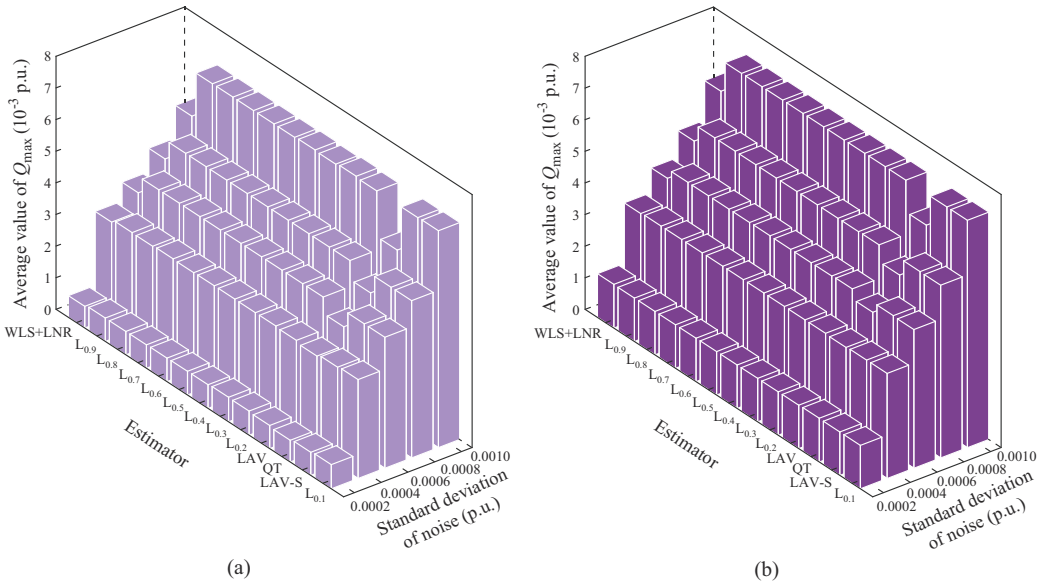

Fig. 10. Average value of Q_{max} from 1000 simulations on G1 and G2. (a) G1. (b) G2.

When the test systems contain Gaussian noise only, WLS + LNR and QT estimators perform better than the other estimators. When $0.1 < p < 0.9$, the result indices obtained by the proposed estimator are similar to those of the LAV estimator. Moreover, the estimators perform better in G1 with higher redundancy.

E. IEEE 300-bus System with Different Ratios of Bad Data

To further analyze the performance of the proposed estimator, G2 with 1%, 2%, 3%, 4%, and 5% bad data are simulated. Bad data are formed by increasing or decreasing the absolute value by more than 30% of measurements. The standard deviation of noise is set as 0.0010 p.u..

The average values of $\bar{S}$, $\bar{Q}$ and S_{max}, Q_{max} from 1000 simulations on G2 with different ratios of bad data are shown in Figs. 11 and 12, respectively.

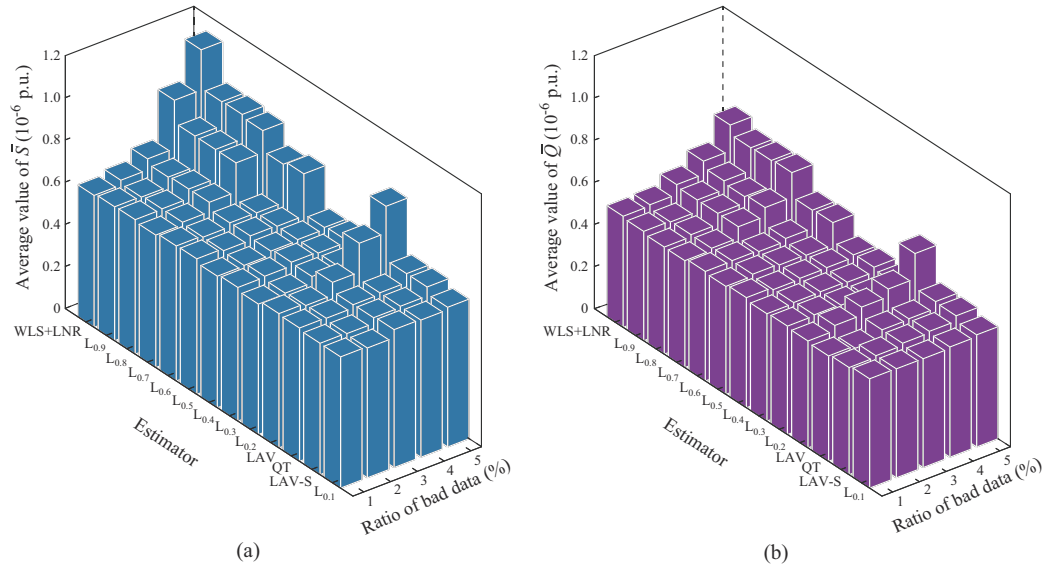

Fig. 11. Average values of $\bar{S}$ and $\bar{Q}$ from 1000 simulations on G2. (a) $\bar{S}$. (b) $\bar{Q}$.

The performance of the estimators can be summarized as follows.

1) Since some conforming measurements become bad data as the ratio of bad data increases in G2, the performance of the WLS + LNR estimator is degraded. Some of the bad data cannot be detected, and their estimation values deviate from true values. Thus, the WLS + LNR estimator has the worst performance. The result indices are significantly higher than those of the other estimators.

2) Compared with WLS + LNR, the LAV estimator could suppress conforming bad data more effectively and provide a more reliable estimation. However, as the ratio of bad data increases, some leveraged measurements become bad data. As the LAV estimator cannot handle them, the values of the result indices obtained by the LAV estimator also detectably increase.

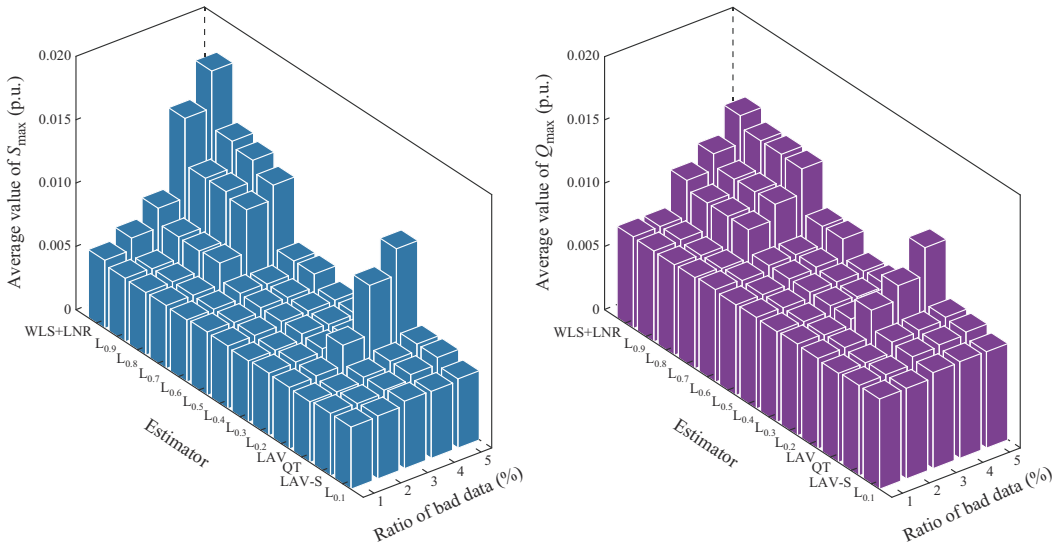

Fig. 12. Average values of S_{max} and Q_{max} from 1000 simulations on G2. (a) S_{max}. (b) Q_{max}.

3) The QT and LAV-S estimators can suppress both bad conforming and leveraged data effectively. These estimators perform better than WLS + LNR and LAV estimators.

4) For L_p quasi norm state estimator, the values of the indices show a downward trend when p varies from 0.9 to 0.1. Moreover, $L_{0.1}$ estimator can also suppress the bad data of some measurement subsets with low redundancy. Hence, by comparison, the $L_{0.1}$ estimator provides more accurate estimations.

F. Computation Efficiency

The computation efficiency of an estimator plays an important role in evaluating its application. In this section, we present the average values of k and computation time of the aforementioned estimators from the simulations on G2 with 5% bad data, as shown in Figs. 13 and 14, respectively.

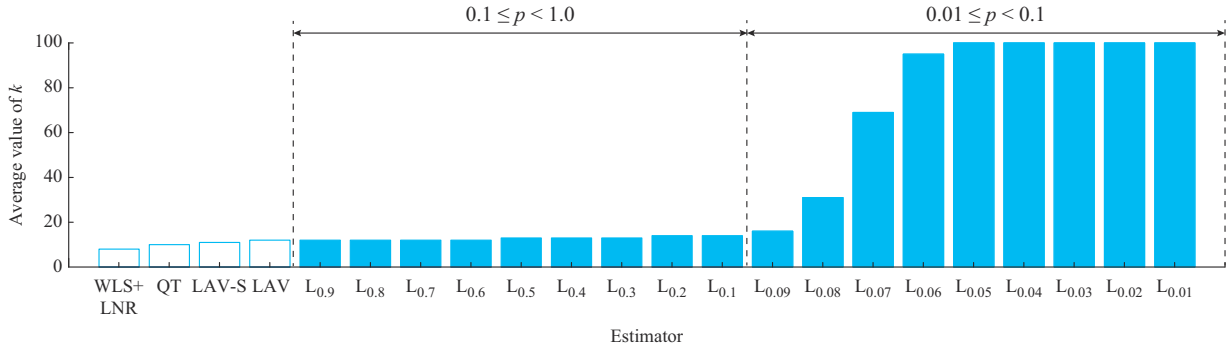

Fig. 13. Average value of k from simulations on G2 with 5% bad data.

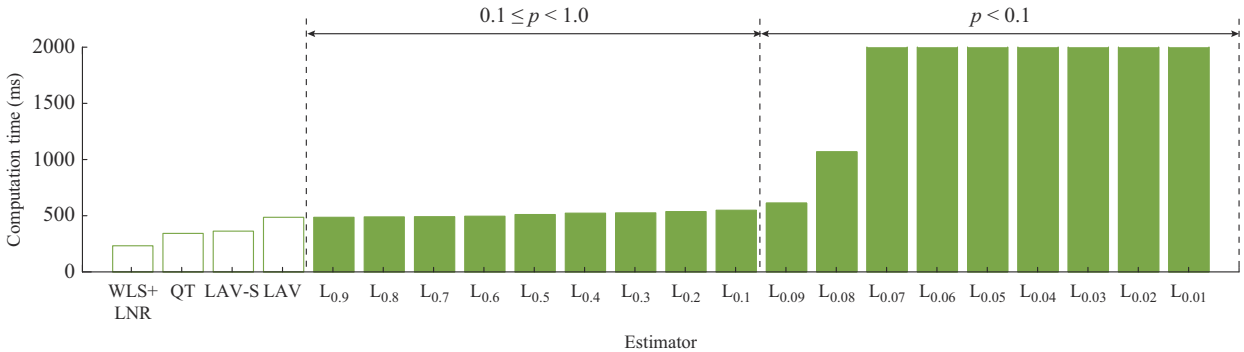

Fig. 14. Average computation time from simulations on G2 with 5% bad data.

As shown in Figs. 13 and 14, the WLS + LNR estimator shows the highest computation efficiency with the smallest values of k and computation time.

As p varies from 0.9 to 0.1, the computation efficiency of L_p quasi norm state estimator is close to that of the LAV estimator, and the average values of k and the computation time

of L_p quasi norm state estimator increase slightly.

As p varies from 0.07 to 0.05, the average value of k using L_p quasi norm state estimator increases rapidly.

Notably, the upper limit of iteration is set to be 100, so the nonconvergence occurs when $0.01 \le p \le 0.06$. Thus, the average computation time increases dramatically.

G. Choosing Value of p

In Sections V-A to V-F, the performance of the proposed estimator is discussed in detail. The robustness and convergence of L_p quasi norm estimator can be summarized as follows.

1) As p varies from 0.9 to 0.1, the robustness of L_p quasi norm state estimator is enhanced and the value of k increases slowly. Moreover, the computation efficiency of $L_{0.1}$ estimator is close to that of the LAV estimator. In a test system with low redundancy, this characteristic is well validated.

2) As p varies from 0.09 to 0.01, the value of k increases rapidly.

Therefore, considering the robustness of L_p quasi norm state estimator, for $0.1 \le p \le 0.9$, $p = 0.1$ is an appropriate choice. Considering the computation efficiency of L_p quasi norm state estimator, for $0.01 \le p \le 0.1$, $p = 0.1$ is also an appropriate choice. Thus, $p = 0.1$ is the approximately optimal value for the practical implementation of the proposed estimator.

H. Application in 1204-bus Provincial System

The $L_{0.1}$ estimator is applied to the 1204-bus provincial system in China. The measurement configuration of this system is shown in Table IX.

TABLE IX
MEASUREMENT CONFIGURATION OF 1204-BUS PROVINCIAL SYSTEM

Number of bus	Number of node voltage V_i	Number of lines	Number of line flow $P_{i\text{-}j}$ and $Q_{i\text{-}j}$	Number of injected power P_i and Q_i
1204	525	1376	2734	1228

Following the standards of the China Southern Power Grid Company Limited, the quality of the estimation results is evaluated based on measurement acceptance rate η [16], [21], which is defined as:

$$\eta = \frac{n_r}{m} \times 100\% \tag{29}$$

where n_r is the number of measurements whose estimation residuals are smaller than the threshold r_i. The voltage magnitude, measurements of active and reactive power, and r_i are set to be 0.5%, 2%, and 3% of the assessment standard values (depending on the measurement types and voltage levels), respectively.

Table X lists the average estimation results obtained for 1000 time sections using the WLS + LNR, QT, LAV-S, LAV, and $L_{0.1}$ estimators.

Generally, the WLS + LNR estimator could provide unbiased estimation results when the measurements only contain Gaussian noise and general bad data. Unfortunately, the 1204-bus provincial system has some conforming bad data in its measurements, and the WLS + LNR estimator fails to detect these bad data. This estimator could not provide unbiased estimation results for those measurements affected by these bad conforming data, and the estimation results deviate from their true values. Therefore, WLS + LNR estimator obtains the lowest η and exhibits the worst quality of estimation results.

TABLE X
ESTIMATION RESULTS OF 1204-BUS PROVINCIAL SYSTEM

Estimator	k	Computation time (s)	η (%)
WLS + LNR	8	1.41	88.52
QT	12	1.76	94.62
LAV-S	11	1.64	94.53
LAV	13	1.85	93.67
$L_{0.1}$	15	2.11	95.61

The LAV estimator can suppress the conforming bad data more effectively, with an η value higher than that of the WLS + LNR estimator. However, it could not handle the bad leveraged data. Instead, QT and LAV-S estimators obtain higher η values and perform better than the LAV estimator.

As $L_{0.1}$ estimator can suppress the bad data of some measurement subsets with low redundancy more effectively, it obtains the highest η value and generates the best estimation results.

The computation time of the $L_{0.1}$ estimator is 14% higher than that of the LAV estimator, which indicates that the $L_{0.1}$ estimator is slightly less computationally efficient, although it can still accommodate the real-time requirement of the power grid.

VI. CONCLUSION

In this study, an L_p quasi norm state estimator is proposed for power systems. By introducing a relaxation model, the problem of low statistical efficiency caused by the presence of many local minimums of L_p quasi norm SE in a noisy environment is addressed. Therefore, this non-convex problem can be solved by using the modern interior point method effectively. Theoretical analysis indicates that the robustness of the L_p quasi norm state estimator increases as p decreases when $0.1 \le p < 1$. The L_p quasi norm state estimator, with a small value of p, shows a higher capacity of suppressing the bad data compared with the WLS + LNR and LAV estimators. High-quality estimation results can also be obtained on a system with low redundancy. After considering the efficiency and estimation quality of the proposed estimator, an optimal value of $p = 0.1$ is suggested for practical implementation, which is verified in the 1204-bus provincial system in China.

REFERENCES

[1] F. C. Shweppe, "Power system static-state estimation, part I-III," *IEEE Transactions on Power Apparatus & Systems*, vol. 89, no. 1, pp. 104-135, Jan. 1970.

[2] A. Bose and K. A. Clements, "Real-time modelling of power net-

works," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1607-1622, Dec. 1987.
[3] F. L. Alvarado and W. F. Tinney, "State estimation using augmented block matrices," *IEEE Transactions on Power Systems*, vol. 5, no. 3, pp. 911-921, Aug. 1990.
[4] R. Nucera and M. Gilles, "A blocked sparse matrix formulation for the solution of equality-constrained state estimation," *IEEE Power Engineering Review*, vol. 11, no. 2, pp. 57-58, Feb. 1991.
[5] E. Handschin, F. C. Schweppe, J. Kohlas *et al.*, "Bad data analysis for power system state estimation," *IEEE Transactions on Power Apparatus Systems*, vol. 94, no. 2, pp. 329-337, Mar. 1975.
[6] F. Wu, W. Liu, L. Holten *et al.*, "Observability analysis and bad data processing for state estimation using Hachtel's augmented matrix method," *IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 604-611, May 1988.
[7] E. Caro, A. J. Conejo, R. Mínguez *et al.*, "Multiple bad data identification considering measurement dependencies," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1953-1961, Nov. 2011.
[8] H. Wei, H. Sasaki, J. Kubokawa *et al.*, "An interior point method for power system weighted nonlinear L_1 norm static state estimation," *IEEE Transactions on Power Systems*, vol. 13, no. 2, pp. 617-623, May 1998.
[9] Y. Chen, F. Liu, S. Mei *et al.*, "A robust WLAV state estimation using optimal transformations," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 2190-2191, Jul. 2015.
[10] M. Kabiri and N. Amjady, "Robust optimisation-based state estimation considering parameter errors for systems observed by phasor measurement units," *IET Generation*, *Transmission & Distribution*, vol. 12, no. 8, pp. 1925-1921, Apr. 2018.
[11] G. F. Piepel, "Robust regression and outlier detection," *Technometrics*, vol. 31, no. 2, pp. 260-261, Aug. 1989.
[12] L. Mili, M. G. Cheniae, N. S. Vichare *et al.*, "Robust state estimation based on projection statistics of power systems," *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 1118-1127, May 1996.
[13] L. Mili, V. Phaniraj, and P. J. Rousseeuw, "Least median of squares estimation in power systems," *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 511-523, May 1991.
[14] M. K. Celik and A. Abur, "Use of scaling in WLAV estimation of power system states," *IEEE Transactions on Power Systems*, vol. 7, no. 2, pp. 684-692, May 1992.
[15] R. C. Pires, A. S. Costa, and L. Mili, "Iteratively reweighted least squares state estimation through Givens rotations," *IEEE Transactions on Power Systems*, vol. 14, no. 4, pp. 1499-1507, Nov. 1999.
[16] G. He, S. Dong, J. Qi *et al.*, "Robust state estimator based on maximum normal measurement rate," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2058-2065, Nov. 2011.
[17] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: properties and applications in non-Gaussian signal processing," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5286-5291, Nov. 2007.
[18] V. Miranda, A. Santos, and J. Pereira, "State estimation based on correntropy: a proof of concept," *IEEE Transactions on Power Systems*, vol. 24, no. 4, pp. 1888-1889, Nov. 2009.
[19] W. Wu, Y. Guo, B. Zhang *et al.*, "Robust state estimation method based on maximum exponential square," *IET Generation, Transmission & Distribution*, vol. 24, no. 4, pp. 1165-1172, Nov. 2011.
[20] Y. Chen, F. Liu, G. He *et al.*, "Maximum exponential absolute value approach for robust state estimation," in *Proceedings of 2012 IEEE International Conference on Power System Technology (POWERCON)*, Auckland, New Zealand, Oct.-Nov. 2012, pp. 1-6.
[21] Y. Chen, J. Ma, P. Zhang *et al.*, "Robust state estimator based on maximum exponential absolute value," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1537-1544, Jul. 2017.
[22] H. Tsutsu and Y. Morikawa, "An l_p norm minimization using auxiliary function for compressed sensing," in *Proceedings of the International Multi Conference of Engineers and Computer Scientists*, Hong Kong, China, Mar. 2012, pp. 712-715.
[23] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707-710, Oct. 2007.
[24] P. Li, X. Huang, X. Zhu *et al.*, "Lp ($p \le 1$) norm partial directed coherence for directed network analysis of scalp EEGs," *Brain Topography*, vol. 31, no. 5, pp. 738-752, Jan. 2018.
[25] Y. Gao, J. Peng, S. Yue *et al.*, "On the null space property of lq-minimization for $0<q<1$ in compressed sensing," *Journal of Function Spaces*, vol. 2015, no. 5, p. 579853, Mar. 2015.
[26] J. Wright, A. Y. Yang, A. Ganesh *et al.*, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, Feb. 2009.
[27] C. Wang and J. Peng, "Analysis of the equivalence relationship between l_0-minimization and l_p-minimization," *Journal of Inequalities & Applications*, vol. 2017, no. 1, p. 313, Dec. 2017.
[28] Q. Lyu, Z. Lin, Y. She *et al.*, "A comparison of typical *lp* minimization algorithms," *Neurocomputing*, vol. 119, no. 7, pp. 413-424, Nov. 2013.
[29] A. Carmi, P. Gurfil, and D. Kanevsky, "Methods for sparse signal recovery using Kalman filtering with embedded pseudo-measurement norms and quasi-norms," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2405-2409, Apr. 2010.
[30] F. Nie, H. Wang, X. Cai *et al*., "Robust matrix completion via joint schatten *p*-norm and l_p-norm minimization," in *Proceedings of 2012 IEEE 12th International Conference on Data Mining*, Brussels, Belgium, Dec. 2012, pp. 566-574.
[31] R. Gonin and A. H. Money, "The nonlinear L_1-norm estimation problem," in *Nonlinear Lp-norm Estimation*, New York and Basel: Marcel Dekker, 1989, pp. 43-119.
[32] Y. Gu, Z. Yu, R. Diao *et al*., "Doubly-fed deep learning method for bad data identification in linear state estimation," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1140-1150, Nov. 2020.
[33] A. El-Bakry, R. A. Tapia, T. Tsuchiya *et al*., "On the formulation and theory of the Newton interior-point method for nonlinear programming," *Journal of Optimization Theory and Applications*, vol. 89, no. 3, pp. 507-541, Jun. 1996.
[34] G. Tian, Y. Gu, D. Shi *et al*., "Neural-network-based power system state estimation with extended observability," J*ournal of Modern Power Systems and Clean Energy*, vol. 9, no. 5, pp. 1043-1053, Sept. 2021.

Zhongliang Lyu received the B.S. degree in agricultural electrification and automation from Nanjing Agricultural University, Nanjing, China, in 2015. Currently, he is pursuing the Ph.D. degree in power engineering at Guangxi University, Nanning, China. His research interests include the application of optimization methods in power systems.

Hua Wei received the B.S. and M.S. degrees in power engineering from Guangxi University, Nanning, China, in 1981 and 1987, respectively, and the Ph.D. degree in power engineering from Hiroshima University, Hiroshima, Japan, in 2002. Currently, he is working as a Professor at Guangxi University. He is also the Director of the Institute of Power System Optimization, Guangxi University. His research interests include power system operation and planning, and the application of optimization theory and methods in power systems.

Xiaoqing Bai received the B.S. and Ph.D. degrees in power engineering from Guangxi University, Nanning, China, in 1991 and 2010, respectively. From 2012 to 2015, she was a Postdoctoral Research Assistant with the University of Nebraska Lincoln, Lincoln, USA. She is currently a Professor in the Institute of Power System Optimization, Guangxi University. Her research interests include power system optimization based on semidefinite programing (SDP) and robust optimization.

Daiyu Xie received the B.S. and M.S. degrees in power engineering from Guangxi University, Nanning, China, in 2014 and 2017, respectively. He currently works with the Power Dispatching Control Center of the Guangxi Power Grid, Nanning, China. His research interests include the application of optimization methods in power systems.

Le Zhang received the B.S. degree in power engineering from Beijing Jiaotong University, Beijing, China, in 2017. He is currently working towards the M.S. degree at Guangxi University, Nanning, China. His research interest includes power system optimization.

Peijie Li received the B.E. and Ph.D. degrees in electrical engineering from Guangxi University, Nanning, China, in 2006 and 2012, respectively. From 2015 to 2016, he was with Argonne National Laboratory, Lemont, USA, as a Visiting Scholar. He is currently an Associate Professor with Guangxi University. His research interests include optimal power flow, small-signal stability, and security-constrained economic dispatch and restoration.