An Embedded Consensus ADMM Distribution Algorithm Based on Outer Approximation for Improved Robust State Estimation of Networked Microgrids

Zifeng Zhang and Yuntao Ju

***Abstract*—Networked microgrids (NMGs) are critical in the accommodation of distributed renewable energy. However, the existing centralized state estimation (SE) cannot meet the demands of NMGs in distributed energy management. The current estimator is also not robust against bad data. This study introduces the concepts of relative error to construct an improved robust SE (IRSE) optimization model with mixed-integer nonlinear programming (MINLP) that overcomes the disadvantage of inaccurate results derived from different measurements when the same tolerance range is considered in the robust SE (RSE). To improve the computation efficiency of the IRSE optimization model, the number of binary variables is reduced based on the projection statistics and normalized residual methods, which effectively avoid the problem of slow convergence or divergence of the algorithm caused by too many integer variables. Finally, an embedded consensus alternating direction of multiplier method (ADMM) distribution algorithm based on outer approximation (OA) is proposed to solve the IRSE optimization model. This algorithm can accurately detect bad data and obtain SE results that communicate only the boundary coupling information with neighbors. Numerical tests show that the proposed algorithm effectively detects bad data, obtains more accurate SE results, and ensures the protection of private information in all microgrids.**

***Index Terms*—Distributed optimization, alternating direction of multiplier methods (ADMM), robust state estimation (RSE), mixed-integer nonlinear programming (MINLP), networked microgrid (NMG).**

NOMENCLATURE

A. Indexes

0	Index of initial values
a, b	Indices of microgrid a (MGa) and microgrid b (MGb)
k	Index of iterations
r	Index of measurements
c	Index of suspicious measurements

B. Parameters

α	Step factor
$\boldsymbol{A}$	Coupling matrix
C	Total number of suspicious measurements
C_a	Number of measurement equations after filtering non-suspicious measurements
d_a	Dual residuals in MGa
$\boldsymbol{e}$	Vector of measurement errors
$\boldsymbol{h}(\boldsymbol{x})$	Vector of measurement equations
$h_r(\boldsymbol{x})$	The r^{th} measurement equation
k_r^+, k_r^-	The upper and lower values of tolerance for the r^{th} measurement k_r
$k_{\max}$	The maximum number of iterations
$L_{a,NIPF}$	The computational feasibility function in MGa
$L_{f_a}^k$, $U_{f_a}^k$	The lower and upper bounds of objective function for MGa
M	An arbitrarily large positive scalar value
m	Total number of measurements, i. e., number of binary variables
m_a	Number of binary variables in MGa
n_a	Number of buses in MGa
nb_a	Number of branches in MGa
$n_{a_{int}}$	Internal buses in MGa
$n_{a_{cou}}$	Number of coupling state variables in MGa
N	Total number of MGs
$\boldsymbol{z}$	Measurement vector, $\boldsymbol{z}=[z_r]$

Manuscript received: August 15, 2023; revised: November 3, 2023; accepted: January 17, 2024. Date of CrossCheck: January 17, 2024. Date of online publication: April 18, 2024.

This work was supported by the National Natural Science Foundation of China (No. 5217070269).

Z. Zhang (corresponding author) is with the State Grid Information & Telecommunication Group Co., Ltd., Beijing Branch, Beijing, China (e-mail: zzf_edu@126.com).

Y. Ju is with the School of Electrical and Control Engineering, North China University of Technology, Beijing, China (e-mail: juyuntaomail@163.com).

DOI: 10.35833/MPCE.2023.000565

C. Variables

$\boldsymbol{\lambda}$	Lagrange multiplier
χ	Feasible region
ρ	Penalty factor
β_a	Continuous variable of objective function for master problem
b_r	Binary variable for the r^{th} measurement
$\boldsymbol{b}$	Vector of binary viables
$\boldsymbol{x}_{s,a}$	State estimation variable of MGa
$\boldsymbol{x}$	Vector of state variables
x_a	State variable set of MGa
$x_{a_{int}}$	Internal state variable set in MGa
$x_{a_{cou}}$	Coupling state variable set in MGa

I. Introduction

As a greater number of distributed energy resources (DERs), including distributed generators, distributed energy storage systems, and flexible loads, are integrated with networked microgrids (NMGs) [1], the state estimation (SE) has become increasingly critical to the safety and reliable operation of NMGs. However, the conventional centralized SE methods have encountered the following significant technical challenges.

1) SE accuracy: the weighted least squares (WLS) estimator is used to find the best state estimates [2]. However, the WLS estimators are not robust against bad data, and even an erroneous measurement may significantly bias the estimation result [3]. To ensure the accurate estimation, a bad data detection (BDD) method is required. However, accurately detecting all bad data is a well-known challenge, particularly for bad leverage points.

2) Privacy issues: as microgrids are managed by different entities, the privacy-preserving is a concern with respect to managing NMGs. However, the centralized WLS estimators require global information, which leads to private information leakage from different microgrids and aggravates the burden of information communication.

The detection and elimination of bad data are critical for SE accuracy. An intelligent search strategy for conforming BDD methods was presented in [4], which addressed the issue of conforming error detection involving analog measurements and proposed a system topology with heuristic tool. Reference [5] proposed a robust SE (RSE) framework based on the Mahalanobis distance to filter out non-Gaussian noises measured by phasor measurement unit (PMU) and suppress outliers. Reference [6] introduced an SE method based on entropy-related concepts to detect and correct measurement errors. This method was verified by theoretical models and examples, where an algorithm was shown to filter out numerous errors. Reference [7] proposed a mixed-integer nonlinear programming (MINLP) method for RSE, where each measurement with its associated uncertainty interval is represented by the upper and lower limits of tolerance. This method was shown to be robust against gross errors and was not susceptible to bad leverage measurements. To further guarantee the global optimum theoretically and reduce the time consumption as the number of iterations increases, [8] used the linearized measurement equation to transform the MINLP formulation for RSE in [9] into a mixed-integer linear programming (MILP) one, which guaranteed that the global optimum would be found while avoiding convergence problems. These centralized BDD methods do not consider the distributed management required for NMGs to protect the information privacy.

Privacy issues have been widely studied for distributed SE [10]. In [11], a two-level SE approach was presented, where the local SE results at the first level were coordinated with those at the second level. Reference [12] proposed a hierarchical SE (HSE) method, which allowed more efficient information for faster convergence with lower communication costs. However, implementing a coordinator required stronger synchronization across operating regions. To reduce the amount of exchanged information among adjacent areas, a new multi-area HSE method was proposed in [13] based on the exchange of sensitivity functions of local SE instead of the exchange of boundary measurements. Additionally, the convergence speed is effectively improved. In [14], a distributed RSE method was developed based on the alternating direction of multiplier methods (ADMM), which respects privacy policies, exhibits low communication pressure, and converges to centralized SE results. Reference [15] proposed an adaptive distributed quasi-Newton algorithm with an optimal step-length tuning strategy to solve the multi-area SE problem. This algorithm obtains accurate estimation results and considers coupling information among areas while preserving the privacy and independence of each area. Reference [16] proposed a fully distributed integrated solution based on a distributed subgradient algorithm to solve multi-area topology identification and SE problems. The distributed SE method does not consider the detection and filtering of bad data, particularly bad leverage points, which might lead to inaccuracies in SE results and affect the stability of the system.

To improve the SE accuracy and meet the distributed management requirements of NMGs, this study constructs an improved RSE (IRSE) optimization model with MINLP and proposes an embedded consensus ADMM distribution algorithm based on outer approximation (OA). The proposed algorithm is susceptible to leverage points, accurately detects and eliminates bad data, and obtains accurate SE results for each microgrid. The main contributions of this study are as follows.

1) Because the elements of residual sensitivity matrix corresponding to the leverage measurement are very small, it is difficult for traditional methods to detect bad leverage points. An IRSE optimization model with MINLP is constructed in this study, which can find the optimal measurement set, unlike in statistics-based detection.

2) To improve the computational efficiency of the IRSE optimization model based on the projection statistics and normalized residual methods, the aggregate dimension of the suspicious measurement in the RSE is compressed, and the number of binary variables is reduced to ensure the optimal SE results.

3) In terms of the communication and convergence, the outer layer of the proposed algorithm exchanges only boundary coupling information with neighbors to protect information in a fully distributed manner. Additionally, the inner layer of the proposed algorithm solves each microgrid in parallel, which can accurately detect bad measurements and obtain accurate SE results for each microgrid under limited interactions.

The remainder of this paper is organized as follows. Section II introduces centralized and distributed IRSE optimization models with MINLP. Section III provides a detailed introduction to the embedded consensus ADMM distribution algorithm based on OA. Case studies proving the accuracy and convergence of the proposed algorithm are presented in Section IV. Finally, a conclusion is given in Section V.

II. Centralized and Distributed IRSE Optimization Models with MINLP

A. Centralized IRSE Optimization Model

For the traditional SE, the relationship between the state variable and measurements for the entire system is expressed as [2]:

$$\boldsymbol{z}=\boldsymbol{h}(\boldsymbol{x})+\boldsymbol{e} \tag{1}$$

where $\boldsymbol{x}$ usually includes the voltage magnitudes and phase angles of all buses.

To avoid the effects of outliers or erroneous measurements on the SE, a robust state estimator based on the maximum constraint satisfaction of uncertain measurements was proposed in [17], where each uncertain measurement is represented by the upper and lower limits of the bounds on the measurement errors as:

$$h_r(\boldsymbol{x})-z_r\le k_r^+ \tag{2}$$

$$h_r(\boldsymbol{x})-z_r\ge -k_r^- \tag{3}$$

where the values of k_r^+ and k_r^- are given in [2].

If all measurement errors are within the tolerance range $[-k_r^-,k_r^+]$, it is possible to find a solution that satisfies the aforementioned inequalities for all measurements. However, if some measurements have unexpectedly large errors, they will not satisfy (2) and (3). To effectively solve this problem, the concept of relative error is introduced in this study to avoid the effects of applying the same tolerance range to the accuracy of SE results, (2) and (3) can then be respectively reformulated as:

$$h_r(\boldsymbol{x})-z_r\le k_r^+\left|z_r\right|+Mb_r \tag{4}$$

$$h_r(\boldsymbol{x})-z_r\ge -k_r^-\left|z_r\right|-Mb_r \tag{5}$$

Therefore, the centralized IRSE optimization model with MINLP (Cen-model1) in [7] can be reformulated as:

$$\begin{cases}\min\limits_{\boldsymbol{x},b_r}\sum\limits_{r=1}^{m}b_r\\ \text{s.t. } \boldsymbol{g}(\boldsymbol{x})=\boldsymbol{0}\\ \quad h_r(\boldsymbol{x})-z_r\le k_r^+\left|z_r\right|+Mb_r\\ \quad h_r(\boldsymbol{x})-z_r\ge -k_r^-\left|z_r\right|-Mb_r\end{cases} \tag{6}$$

where $\boldsymbol{g}(\boldsymbol{x})$ includes the zero-injection constraint and phase angle constraints for slack buses. For a good measurement, $b_r=0$, indicating that the solution of the r^{th} measurement equation is in the interval $[-k_r^-,k_r^+]$; otherwise, for a bad measurement, $b_r=1$.

On this basis, we consider filtering out non-suspicious measurements to reduce the number of binary variables m, thereby reducing the complexity of Cen-model1. First, the projection statistics method is used to identify leverage measurements, and all leverage measurements are stored in the suspicious measurement set. Second, the normalized residual method [18] for unleveraged measurements identifies the measurement with bad data and stores the identified measurement in the suspicious measurement set. Finally, only the binary variable for the suspicious measurements b_c are retained, and Cen-model1 given in (6) is reformulated as Cen-model2:

$$\begin{cases}\min\limits_{\boldsymbol{x},b_c}\sum\limits_{c=1}^{C}b_c\\ \text{s.t. } \boldsymbol{g}(\boldsymbol{x})=\boldsymbol{0}\\ \quad h_r(\boldsymbol{x})-z_r\le k_r^+\left|z_r\right|\\ \quad h_r(\boldsymbol{x})-z_r\ge -k_r^-\left|z_r\right|\\ \quad h_r(\boldsymbol{x})-z_r\le k_r^+\left|z_r\right|+Mb_c\\ \quad h_r(\boldsymbol{x})-z_r\ge -k_r^-\left|z_r\right|-Mb_c\end{cases} \tag{7}$$

B. Distributed IRSE Optimization Model

1) Decoupling

As shown in Fig. 1, the NMGs are decoupled into Microgrid a (MGa) and Microgrid b (MGb) at the branch (a_i,b_j) by the bus replication method.

Taking MGa as an example, by copying the information of bus b_j to obtain a new coupling branch $(a_i,a_{j'})$, the internal and coupling bus sets are obtained as $a_{int}=\{a_i\}$ and $a_{cou}=\{a_{j'}\}$, respectively. The sets of internal state variables x_{a_i} and boundary coupling variables x_{a_j} are denoted as $x_{a_{int}}=\{x_{a_i}\}$ and $x_{a_{cou}}=\{x_{a_{j'}}\}$, respectively; and $x_a=x_{a_{int}}\bigcup x_{a_{cou}}$. The same definitions can be applied to MGb.

2) Coupling Constraints

To ensure that the solution of the above distributed IRSE optimization model is identical to the estimated solution of Cen-model2. For MGa, the coupling variables must satisfy the following coupling constraints:

$$\begin{cases}x_{a_i}=x_{b_{i'}}\\ x_{a_{j'}}=x_{b_j}\end{cases} \tag{8}$$

These coupling constraints are simply expressed as:

$$\boldsymbol{A}_a\boldsymbol{x}_a-\boldsymbol{A}_b\boldsymbol{x}_b=\boldsymbol{0} \tag{9}$$

where state variables $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$ include the voltage magnitude and phase angles of all buses in MGa and MGb, respectively.

The element of the coupling matrix $\boldsymbol{A}_a$ is expressed as:

$$A_a(c_a,i)=\begin{cases}1 & x_{a_i}\in x_{a_{int}}\\ 0 & x_{a_i}\in x_{a_{cou}}\end{cases} \tag{10}$$

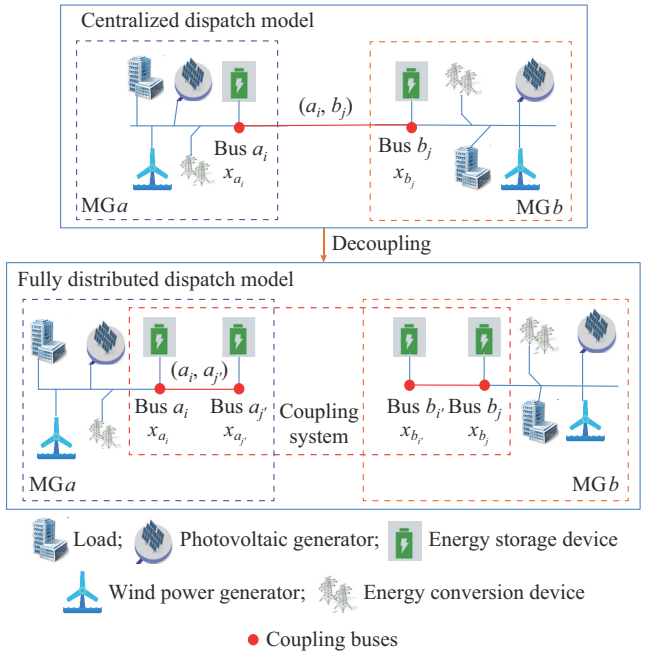

Fig. 1. Decoupling process of NMG.

where $c_a \in \{1, 2, ..., n_{a_{cou}}\}$ and $i = \{1, 2, ..., n_a\}$, indicating $\boldsymbol{A}_a$ is an $n_{a_{cou}} \times n_a$ matrix.

3) Formulation of Distributed IRSE Optimization Model

For MGa, the distributed IRSE optimization model, denoted as Dis-model1, with 0-1 binary variables $b_{a,r}$, is expressed as:

$$\min \sum_{r=1}^{m_a} b_{a,r} \tag{11}$$

s.t.

$$\begin{cases} \boldsymbol{g}_a(\boldsymbol{x}_a) = \boldsymbol{0} \\ h_{a,r}(\boldsymbol{x}_a) - z_{a,r} \le k_{a,r}^{+} \left| z_{a,r} \right| + M b_{a,r} \\ h_{a,r}(\boldsymbol{x}_a) - z_{a,r} \ge -k_{a,r}^{-} \left| z_{a,r} \right| - M b_{a,r} \\ \boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{A}_b \boldsymbol{x}_b = \boldsymbol{0} \end{cases} \tag{12}$$

where $\boldsymbol{g}_a(\boldsymbol{x}_a)$ includes zero-injection constraint and phase-angle constraints for slack buses in MGa. After non-suspicious measurements are filtered out, the distributed IRSE optimization model, denoted as Dis-model2, with binary variables $b_{a,c}$, is rewritten as:

$$\min \sum_{c=1}^{C_a} b_{a,c} \tag{13}$$

s.t.

$$\begin{cases} \boldsymbol{g}_a(\boldsymbol{x}_a) = \boldsymbol{0} \\ h_{a,r}(\boldsymbol{x}_a) - z_{a,r} \le k_{a,r}^{+} \left| z_{a,r} \right| \\ h_{a,r}(\boldsymbol{x}_a) - z_{a,r} \ge -k_{a,r}^{-} \left| z_{a,r} \right| \\ h_{a,r}(\boldsymbol{x}_a) - z_{a,r} \le k_{a,r}^{+} \left| z_{a,r} \right| + M b_{a,c} \\ h_{a,r}(\boldsymbol{x}_a) - z_{a,r} \ge -k_{a,r}^{-} \left| z_{a,r} \right| - M b_{a,c} \\ \boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{A}_b \boldsymbol{x}_b = \boldsymbol{0} \end{cases} \tag{14}$$

III. Embedded Consensus ADMM Distribution Algorithm Based on OA

We describe the embedded consensus ADMM distribution algorithm based on OA to solve the SE problem of Dis-model1 and Dis-model2.

A. Framework of Embedded Consensus ADMM Distribution Algorithm

Combining the decomposition of the dual ascent method and multiplier method, we add the coupling constraint to the Dis-model1 to obtain its augmented Lagrangian objective function, which is expressed as:

$$L(\boldsymbol{x}_a, \boldsymbol{b}_a, \boldsymbol{x}_b, \boldsymbol{b}_b, \boldsymbol{\lambda}_a) = \sum_{a=1}^{N} \left[\sum_{r=1}^{m_a} b_{a,r} + \boldsymbol{\lambda}_a^{\mathrm{T}} (\boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{A}_b \boldsymbol{x}_b) + \frac{\rho_a^k}{2} \left\| \boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{A}_b \boldsymbol{x}_b \right\|_2^2 \right] \tag{15}$$

The iterative form of the ADMM distributed algorithm in [19] is expressed as:

$$\boldsymbol{x}_a^{k+1} = \arg\min_{x_a \in \chi_a} L(\boldsymbol{x}_a, \boldsymbol{b}_a, \boldsymbol{x}_b^k, \boldsymbol{b}_b^k, \boldsymbol{\lambda}_a^k) \tag{16}$$

$$\boldsymbol{x}_b^{k+1} = \arg\min_{x_b \in \chi_b} L(\boldsymbol{x}_a^{k+1}, \boldsymbol{b}_a^{k+1}, \boldsymbol{x}_b, \boldsymbol{b}_b, \boldsymbol{\lambda}_b^k) \tag{17}$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \rho^k (\boldsymbol{A}_a \boldsymbol{x}_a^{k+1} - \boldsymbol{A}_b \boldsymbol{x}_b^{k+1}) \tag{18}$$

The computation of (17) for MGb cannot start until (16) for MGa is computed, which means the problems of MGa and MGb are solved sequentially rather than in parallel. This transforms the calculation time of each iteration into the total calculation time for microgrid. For each microgrid to be calculated independently in parallel, the consistent variable $\boldsymbol{y}_a$ is introduced in [20] to replace the coupled variables of adjacent microgrids.

$$\boldsymbol{y}_a^k = \boldsymbol{A}_b \boldsymbol{x}_b - (1/\rho_a^k) \boldsymbol{\lambda}_a^k \tag{19}$$

Therefore, (15) can be rewritten as:

$$L(\boldsymbol{x}_a, \boldsymbol{b}_a, \boldsymbol{y}_a, \boldsymbol{\lambda}_a) = \sum_{a=1}^{N} \left[\sum_{r=1}^{m_a} b_{a,r} + \boldsymbol{\lambda}_a^{\mathrm{T}} (\boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{y}_a) + \frac{\rho_a}{2} \left\| \boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{y}_a \right\|_2^2 \right] \tag{20}$$

s.t.

$$\boldsymbol{x}_a \in \chi_a \quad a = 1, 2, \ldots, N \tag{21}$$

Then, the objective function (20) is decoupled as:

$$L(\boldsymbol{x}_a, \boldsymbol{b}_a, \boldsymbol{y}_a, \boldsymbol{\lambda}_a) = \sum_{r=1}^{m_a} b_{a,r} + \boldsymbol{\lambda}_a^{\mathrm{T}} (\boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{y}_a) + \frac{\rho_a}{2} \left\| \boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{y}_a \right\|_2^2 \tag{22}$$

Finally, the iterative process of the ADMM distributed algorithm is obtained as:

$$\boldsymbol{x}_a^{k+1} = \arg\min_{x_a \in \chi_a} L(\boldsymbol{x}_a, \boldsymbol{b}_a, \boldsymbol{y}_a^k, \boldsymbol{\lambda}_a^k) \tag{23}$$

$$\boldsymbol{y}_a^{k+1} = \frac{\boldsymbol{A}_a \boldsymbol{x}_a^{k+1} + \boldsymbol{A}_b \boldsymbol{x}_b^{k+1}}{2} \quad a = 1, 2, ..., N \tag{24}$$

$$\boldsymbol{\lambda}_a^{k+1} = \boldsymbol{\lambda}_a^k + \rho_a^k \frac{\boldsymbol{A}_a \boldsymbol{x}_a^{k+1} - \boldsymbol{y}_a^{k+1}}{2} \tag{25}$$

where each microgrid is solved in parallel to effectively

overcome the disadvantage of the sequential iteration in (23). In addition, the Lagrangian multipliers in (24) and (25) only need to exchange the boundary coupling information between the adjacent microgrids, as shown in Fig. 2.

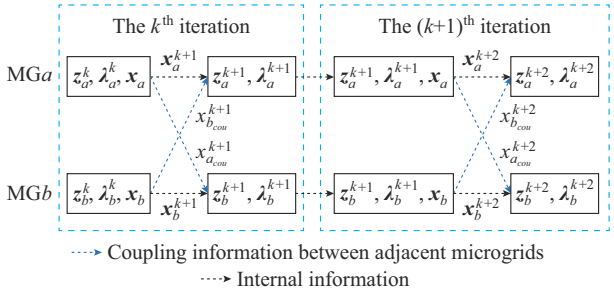

Fig. 2. Information communication between two adjacent microgrids.

The proposed algorithm converges when the dual residuals d_a^{k+1} are less than a small convergence threshold ε, which is expressed as:

$$d_a^{k+1} = \left\| \boldsymbol{A}_a \boldsymbol{x}_a^{k+1} - \boldsymbol{A}_b \boldsymbol{x}_b^{k+1} \right\|_2 \tag{26}$$

The pseudocode of the ADMM distributed algorithm for solving the problem in (23)-(25) for MGa is presented as Algorithm 1.

Algorithm 1: pseudocode of ADMM distributed algorithm

function ADMM $(\boldsymbol{x}_a^0, \boldsymbol{x}_b^0, z^0, \boldsymbol{\lambda}^0, \rho^0, \varepsilon)$
  $k \leftarrow 0$
  while $k < k_{\max}$ **or** $d_a^k \le \varepsilon$
    $\boldsymbol{x}_a^{k+1} \leftarrow \arg\min_{\boldsymbol{x}_a} L(\boldsymbol{x}_a, \boldsymbol{b}_a, \boldsymbol{y}_a^k, \boldsymbol{\lambda}_a^k),\ a = 1, 2, \ldots, N$
    $\boldsymbol{y}_a^{k+1} \leftarrow (\boldsymbol{A}_a \boldsymbol{x}_a^{k+1} + \boldsymbol{A}_b \boldsymbol{x}_b^{k+1})/2$
    $\boldsymbol{\lambda}_a^{k+1} \leftarrow \boldsymbol{\lambda}_a^k + \rho_a^k (\boldsymbol{A}_a \boldsymbol{x}_a^{k+1} - \boldsymbol{y}_a^{k+1})$
    $d_a^{k+1} \leftarrow \left\| \boldsymbol{A}_a \boldsymbol{x}_a^{k+1} - \boldsymbol{A}_b \boldsymbol{x}_b^{k+1} \right\|_2$
    $k \leftarrow k+1$
  end while
  return $\boldsymbol{x}_a^{k+1}$, $\boldsymbol{y}_a^{k+1}$, and $\boldsymbol{\lambda}_a^{k+1}$
end function

B. OA Algorithm for Solving SE Problem at MGa

An OA algorithm based on the relaxation principles [21] is next introduced to solve the problem in (23) for MGa.

Step 1: initialization. Set the minimum and maximum values of the objective function of MGa ($a = 1, 2, \ldots, N$) as $L_{f_a}^0 = -\infty$ and $U_{f_a}^0 = +\infty$, respectively. The initial value of the SE variable $\boldsymbol{x}_a^0$ is set as the flat start value. Initialize $\boldsymbol{b}_a^0$, $\boldsymbol{y}_a^0$, and $\boldsymbol{\lambda}_a^0$.

Step 2: solve the nonlinear subproblem.

1) The nonlinear subproblem is feasible. Substitute the k^{th} integer variable into MGa of the problems in (20) and (21). Here, the MINLP in (23) is relaxed into a nonlinear problem $L_{a,NLP}$ as:

$$\min L_{a,NLP}(\boldsymbol{x}_a, \boldsymbol{y}_a^k, \boldsymbol{b}_a^k, \boldsymbol{\lambda}_a^k) = \sum_{r=1}^{m_a} b_{a,r} + \boldsymbol{\lambda}_a^k (\boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{y}_a^k) + \frac{\rho_a^k}{2} \left\| \boldsymbol{A}_a \boldsymbol{x}_a - \boldsymbol{y}_a^k \right\|^2 \tag{27}$$

s.t.

$$\boldsymbol{x}_a \in \chi_a \quad a = 1, 2, \ldots, N \tag{28}$$

where the optimal solution $\boldsymbol{x}_{s,a}^{k+1}$ of the SE variable and the maximum value of the objective function $U_{f_a}^{k+1}$ are obtained by solving (27) and (28).

2) If (27) and (28) are infeasible, a continuous variable μ_{ar} is introduced as the objective function in MGa, and the computational feasibility problem $L_{a,NIPF}$ is expressed as:

$$\begin{cases} \min L_{a,NLPF}(\boldsymbol{x}_a, \boldsymbol{y}_a^k, \boldsymbol{b}_a^k, \boldsymbol{\lambda}_a^k, \mu_{ar}) = \sum_{r=1}^{m_a} \mu_{ar} U_{f_a}^{k+1} \\ \text{s.t. } h_{a,r}(\boldsymbol{x}_a^k) + \nabla h_{a,r}(\boldsymbol{x}_a^k)(\boldsymbol{x}_a - \boldsymbol{x}_a^k) - z_{a,r} - (k_{a,r} \left| z_{a,r} \right| + M b_{a,r}^k) \le \mu_{ar} \\ \quad -k_{a,r} \left| z_{a,r} \right| - M b_{a,r}^k - [z_{a,r}(\boldsymbol{x}_a^k) + \nabla z_{a,r}(\boldsymbol{x}_a^k)(\boldsymbol{x}_a - \boldsymbol{x}_a^k) - z_{a,r}] \le \mu_{ar} \\ \quad \mu_{ar} \ge 0 \\ \quad a = 1, 2, \ldots, N \end{cases} \tag{29}$$

After solving (29), we can yield the optimal $\boldsymbol{x}_{s,a}^{k+1}$ and $U_{f_a}^{k+1}$.

Step 3: solve the master problem.

1) The nonlinear subproblem is feasible. The MINLPs in (13) and (14) are approximated as mixed-integer linear problems by Taylor's first-order expansion method. In addition, the secant at point $\boldsymbol{x}_{s,a}^{k+1}$ is used as the new constraint condition, and the continuous variable β_a is introduced as the objective function of the master problem, which can be expressed as:

$$\begin{cases} \min \beta_a \\ \text{s.t. } L_a^k(\boldsymbol{x}_{s,a}^{k+1}, b_{a,r}^k, \boldsymbol{\lambda}_a^k, \boldsymbol{y}_a^k) + \nabla L_a^k(\boldsymbol{x}_{s,a}^{k+1}, b_{a,r}^k, \boldsymbol{\lambda}_a^k, \boldsymbol{y}_a^k)(\boldsymbol{x}_a - \boldsymbol{x}_{s,a}^{k+1}) - \\ \qquad \beta_a \le 0 \\ \quad h_{a,r}^k(\boldsymbol{x}_{s,a}^{k+1}) + \nabla h_{a,r}^k(\boldsymbol{x}_{s,a}^{k+1})(\boldsymbol{x}_a - \boldsymbol{x}_{s,a}^{k+1}) - z_{a,r} \le k_{a,r}^+ + M b_{a,r}^k \\ \quad h_{a,r}^k(\boldsymbol{x}_{s,a}^{k+1}) + \nabla h_{a,r}^k(\boldsymbol{x}_{s,a}^{k+1})(\boldsymbol{x}_a - \boldsymbol{x}_{s,a}^{k+1}) - z_{a,r} \ge -k_{a,r}^- - M b_{a,r}^k \\ \quad a \in 1, 2, \ldots, N \\ \quad b_{a,r}^k \in \{0, 1\} \end{cases} \tag{30}$$

$$\nabla L_a^k(\boldsymbol{x}_{s,a}^{k+1}, b_{a,r}^k, \boldsymbol{\lambda}_a^k, \boldsymbol{y}_a^k) = \frac{\partial L_a(\boldsymbol{x}_a, b_{a,r}, \boldsymbol{\lambda}_a^k, \boldsymbol{y}_a^k)}{\partial \boldsymbol{x}_a} \boldsymbol{x}_{s,a}^{k+1} \tag{31}$$

$$\nabla h_{a,r}^k(\boldsymbol{x}_{s,a}^{k+1}) = \frac{\partial h_{a,r}(\boldsymbol{x}_a)}{\partial \boldsymbol{x}_a} \boldsymbol{x}_{s,a}^{k+1} \tag{32}$$

The continuous variable $\boldsymbol{x}_a^{k+1}$ and binary variable $b_{a,r}^{k+1}$ are updated for the $(k+1)^{\text{th}}$ master problem, and the minimum value of the objective function for MGa is obtained as $L_{f_a}^k = \beta_a$.

2) The master problem in (30) is infeasible. As shown in Fig. 3, the feasible set of the original problem in (13) and (14) is enlarged by using the approximation method of (30). If the continuous variable $\boldsymbol{x}_a^{k+1}$ obtained by (30) satisfies the original problem, it also solves the MINLP. If it is feasible to approximation method but is not feasible to the r^{th} inequation constraint in the original problem, we can add a new valid inequation derived from [22] and [23] to tighten the feasible region of the approximation problem and thus exclude the solution.

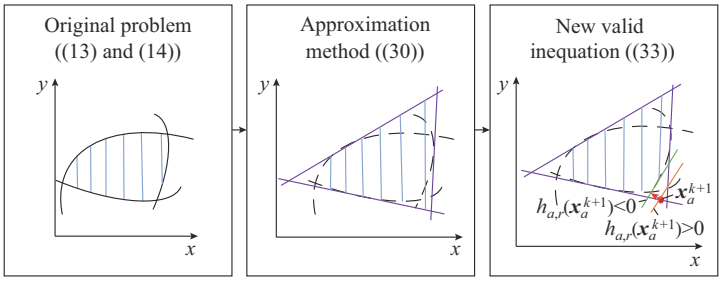

Fig. 3. Excluding infeasible points by adding a new valid inequation.

This means that the OA algorithm can finally converge to a solution that satisfies all constraints in the original problem.

In MGa, if the r^{th} measurement inequation $h_{a,r}^{k}(\boldsymbol{x}_a^{k+1})>0$, new inequality constraints are appended to (30), and the following mixed-integer linearized feasibility problem is solved:

$$\begin{cases}(30)\\ \text{s.t. } h_{a,r}^{k}(\boldsymbol{x}_a^{k+1})+\nabla h_{a,r}^{k}(\boldsymbol{x}_a^{k+1})(\boldsymbol{x}_a-\boldsymbol{x}_a^{k+1})\le 0\end{cases} \tag{33}$$

The solution update yields $\boldsymbol{x}_a^{k+1}, b_{a,r}^{k+1}$, and $L_{f_a}^{k}=\beta_\alpha$.

Step 4: convergence criteria. If $\left|U_{f_a}^{k}-L_{f_a}^{k}\right|<\varepsilon_{in}$ is not satisfied, return to *Step 2*; otherwise, end the iteration to obtain the optimal $\boldsymbol{x}_a^{k+1}$ and $b_{a,r}^{k+1}$ for each microgrid.

The pseudocode of the OA algorithm for solving the problem in (23) of BDD with MINLP for MGa is presented as Algorithm 2.

Algorithm 2: pseudocode of OA algorithm

Input: MINLP problem for each microgrid and convergence criterion ε_{in}.
Output: $(\boldsymbol{x}_a^{k+1}, b_{a,r}^{k+1})$, i.e., the optimal solution to the problem in (25).

function OA $(\boldsymbol{x}_a^0, b_{a,r}^0, \boldsymbol{y}_a^0, \rho_a^0, L_{f_a}^0, U_{f_a}^0, k_{\max})$
  $k\leftarrow 0$
  while $k<k_{\max}$ **or** $\left|U_{f_a}^{k}-L_{f_a}^{k}\right|>\varepsilon_{in}$
    if $\min L_{a,NLP}(\cdot)$ is feasible then
      Obtain $\boldsymbol{x}_{s,a}^{k+1}$ and $U_{f_a}^{k+1}$ by solving (27) and (28)
    else
      Obtain $\boldsymbol{x}_{s,a}^{k+1}$ and $U_{f_a}^{k+1}$ by solving (29)
    end if
    if (30) and (31) are feasible then
      Obtain $\boldsymbol{x}_a^{k+1}, b_{a,r}^{k+1}$, and $L_{f_a}^{k+1}$ by solving (30) and (31)
    else
      Obtain $\boldsymbol{x}_a^{k+1}, b_{a,r}^{k+1}$, and $L_{f_a}^{k+1}$ by solving (33)
    end if
    $k\leftarrow k+1$
  end while
  return $\boldsymbol{x}_a^{k+1}$ and $b_{b,r}^{k+1}$
end function

C. Modification Principle of Penalty Factor ρ

The penalty factor ρ is updated as a step factor to avoid algorithmic oscillations and slow convergence during the iteration. The study proposes the following amendment principles for the updated ρ:

$$\rho^{k+1}=\begin{cases}\rho^{k}\alpha & d^{k}-d^{k-1}\le 0.001\\ \rho^{k}/\alpha & d^{k}-d^{k-1}>0.001\end{cases} \tag{34}$$

where the initial and maximum values of ρ are set as $\rho^0=10$ and $\rho_{\max}=10^8$, respectively. The method for updating the penalty factor in (34) is adjusted appropriately according to the change size of the gradient d, which is suitable for any actual scenario.

IV. Case Studies

The 2-microgrid 5-bus and multi-microgrid 69-bus systems are used to verify the accuracy and convergence of the proposed algorithm. Numerical simulations are performed on a laptop with an Intel i7-8550u (1.99 GHz) CPU and 16 GB of RAM using MATLAB. The CasADi [24] toolbox, IPOPT solver [25], and Bonmin solver [26] are used for optimizations.

A. Case 1: 2-microgrid 5-bus System

A 10 kV single-phase 5-bus grid-connected NMG is constructed by interconnecting Microgrid 1 (MG1) and Microgrid 2 (MG2); the set of buses in MG1 is $\{1,2,3,4'\}$, where Bus 1 is a slack bus, and the set of buses in MG2 is $\{3',4,5\}$. The state variables of MG1 and MG2 are the voltage magnitudes and phase angles of all the buses, respectively. For the 2-microgrid 5-bus system, the measurement configuration is given in Table I, and the numbers of binary variables and locations of bad measurements are given in Table II.

TABLE I
Measurement Configuration of 2-microgrid 5-bus System

All measurement configuration	Bad measurement configuration	Leverage measurement
$U_1, P_{1\text{-}2}, P_{2\text{-}3}, P_{3\text{-}4}, P_{4\text{-}5}, Q_{1\text{-}2}, Q_{2\text{-}3}, Q_{3\text{-}4}, Q_{4\text{-}5}, P_1, P_2, P_3, P_4, P_5, Q_1, Q_2, Q_3, Q_4, Q_5$	$P_{1\text{-}2}, Q_{2\text{-}3}$	$P_{2\text{-}3}, Q_{1\text{-}2}, Q_{2\text{-}3}, P_1, Q_1$

TABLE II
Numbers of Binary Variables and Locations of Bad Measurements in 2-microgrid 5-bus System

Model	Number of binary variables	Location of bad measurement
Dis-model2	[5]; [1]	[7, 9]; [9]
Dis-model1	[9]; [12]	[7, 10]; [14]
Cen-model2	5	[11, 13]
Cen-model1	23	[21, 26]

Note: for an example, "[5]; [1]" means the numbers of binary variables in MG1 and MG2 are 5 and 1, respectively; "[7, 9]; [9]" means in MG1, the 7th and 9th measurement equations have bad measurement and in MG2, the 9th measurement equation has bad measurement.

1) Performance Comparison of Different Models

To verify that the improved Cen-model2 and Dis-model2 are not affected by the values of k_r^- and k_r^+, the calculation solutions of the centralized and distributed IRSE optimization models are provided, as listed in Table III, where L is the location where the calculated binary variable is 1, and the error value *err* is the difference between the SE results and the true value $\boldsymbol{x}_a^{true}$.

$$err = \sum_{a=1}^{N} \left\| \boldsymbol{x}_a - \boldsymbol{x}_a^{true} \right\|_2 \tag{35}$$

TABLE III
CALCULATION SOLUTIONS OF CENTRALIZED AND DISTRIBUTED IRSE OPTIMIZATION MODELS

Values of k_r^- and k_r^+	Model	L	*err*	Is it accurately identified?
0.100	Cen-model2	[21, 26]	4.50×10^{-5}	Yes
	Cen-model1	[11]	7.10×10^{-3}	No
	Dis-model2	[7, 10]; [14]	3.14×10^{-3}	Yes
	Dis-model1	[7]; []	2.21×10^{-2}	No
0.500	Cen-model2	[21, 26]	1.27×10^{-6}	Yes
	Cen-model1	[]; []	1.10×10^{-3}	No
	Dis-model2	[7, 10]; [14]	6.37×10^{-4}	Yes
	Dis-model1	[]; []	2.68×10^{-2}	No
0.010	Cen-model2	[21, 26]	7.79×10^{-8}	Yes
	Cen-model1	[21, 26]	2.77×10^{-2}	Yes
	Dis-model2	[7, 10]; [14]	6.30×10^{-5}	Yes
	Dis-model1	[7]; []	2.86×10^{-2}	No
0.001	Cen-model2	[21, 26]	9.27×10^{-7}	Yes
	Cen-model1	[21, 26]	3.61×10^{-6}	Yes
	Dis-model2	[7, 10]; [14]	3.00×10^{-6}	Yes
	Dis-model1	[7, 10]; [14]	2.86×10^{-2}	Yes

Note: [] indicates that there is no bad measurement, which means there are no binary variables.

Table III shows that Cen-model2 and Dis-model2 that introduce relative errors can accurately identify bad measurements and obtain SE results regardless of the values of k_r^- and k_r^+. However, Cen-model1 and Dis-model1 cannot accurately identify bad measurements and obtain accurate estimation results when k_r^- and k_r^+ are 0.01, 0.1, and 0.5, and the error value is also larger than the true value. Therefore, the IRSE optimization model can provide reasonable values of k_r^- and k_r^+ for different measurements, which ensures bad measurements are identified and accurate SE results are obtained.

2) Accuracy and Efficiency of Proposed Algorithm

The proposed algorithm is applied to perform distributed calculations on Dis-model1 and Dis-model2. Its state variables have the flat start values, $y^0=0$, $k_r^-=k_r^+=0.1$, and $\rho_0=100$. Under the limits of convergence accuracy, $\varepsilon=10^{-4}$ and $k_{\max}=50$ are set. Figure 4 shows the comparison between the true value of voltage amplitude and the corresponding results obtained by the proposed algorithm in MG1 and MG2, where the error value *err* is only 0.00045 p.u.. The results reveal that the proposed algorithm can accurately identify the location of bad data and obtain accurate SE results for each microgrid.

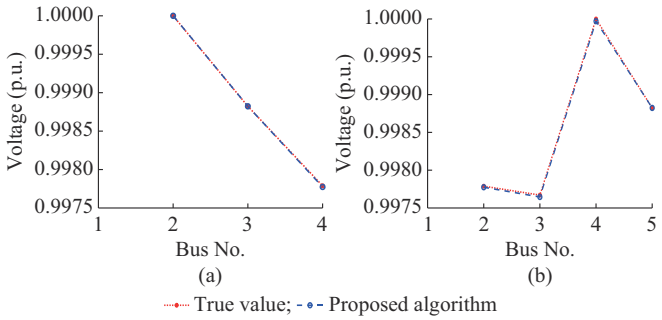

Fig. 4. Comparison between true value of voltage amplitude and corresponding results obtained by proposed algorithm in MG1 and MG2. (a) MG1. (b) MG2.

To investigate whether the proposed algorithm can provide an accurate solution, we first compare the optimal solutions under different models, as shown in Table IV. It can be observed that the location where the calculated binary variable is 1 is identical to that of the bad measurement setting. Therefore, the proposed algorithm is confirmed as effectively identifying bad measurements.

TABLE IV
OPTIMAL SOLUTIOIN UNDER DIFFERENT MODELS

Model	L	Number of iterations	T (s)	*err*
Dis-model2	[7, 9], [9]	11	2.46	4.50×10^{-4}
Dis-model1	[7, 10]; [14]	3	1.45	6.52×10^{-3}
Cen-model2	[11, 13]	6	0.25	5.10×10^{-8}
Cen-model1	[21, 26]	48	1.19	5.10×10^{-8}

In addition, the proposed algorithm can converge with a limited time and limited number of iterations under Dis-model2 to calculate the SE results. The iterative process of the proposed algorithm is shown in Fig. 5. The calculation time T of the proposed algorithm is longer than that of the centralized one. However, this slight increase in calculation time is acceptable in exchange for ensuring information privacy for each microgrid.

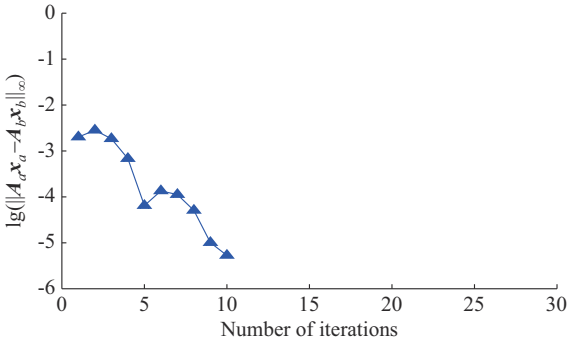

Fig. 5. Iterative process of proposed algorithm.

3) Sensitivity Analysis of Step Factor α

The size of the penalty factor ρ in the multiplier update step significantly affects the convergence of the proposed al-

gorithm. Therefore, to improve the convergence speed, the modified principle in (33) is used to dynamically adjust the coupling constraint penalty factor. As shown in Fig. 6, the step factor $\alpha=1$ is too small such that the numerical calculation is relatively stable, and no convergence occurs.

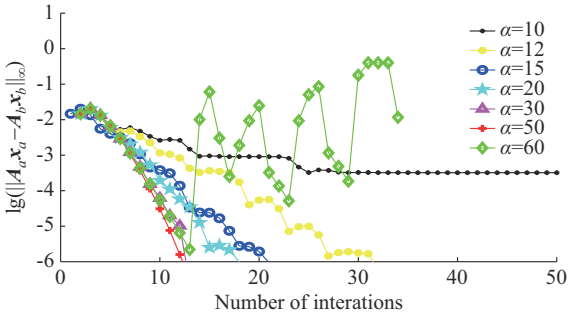

Fig. 6. Convergence procedures with α.

When $\alpha=6$ is too large, even though the proposed algorithm can converge quickly, the updated value ρ in the last iteration exceeds the set maximum value $\rho_{\max}$, which results in a false convergence to the wrong results. When $\alpha=3$ and $\alpha=5$, the proposed algorithm is more efficient and ensures that the value is stable during iteration. Finally, a reasonable range of α is set.

B. Case 2: Multi-microgrid 69-bus System

The calculation for a small-scale system with only two microgrids is relatively simple. Therefore, we use the multi-microgrid 69-bus system to verify the accuracy and efficiency of the proposed algorithm. The original 69-bus system is divided into two, three, four, and seven MGs (denoted as 2-MG, 3-MG, 4-MG, and 7-MG systems for short) using the bus replication method, respectively. The numbers of binary variables as well as the locations and numbers of bad measurements in the centralized and distributed IRSE optimization models of each system are listed in Table V.

TABLE V
NUMBERS OF BINARY VARIABLES AS WELL AS LOCATIONS AND NUMBERS OF BAD MEASUREMENTS OF EACH SYSTEM

System	Model	Number of binary variables	Bad measurement	
			Location	Number
2-MG	Dis-model2	[40]; [26]	[139, 143, 144, 150, 174, 176, 188, 192]	8
	Dis-model1	[100]; [108]	[278, 293, 296, 313, 371, 396, 377, 404]	8
3-MG	Dis-model2	[40]; [22]; [6]	[69, 77, 90, 92]; [75, 76, 89, 93]	8
	Dis-model1	[133]; [74]; [72]	[70, 86, 109, 115]; [83, 87, 129, 137]	8
4-MG	Dis-model2	[27]; [22]; [6]; [16]	[69, 77, 90, 92]; [50, 54]; [39, 40]	8
	Dis-model1	[91]; [74]; [72]; [44]	[70, 86, 109, 115]; [62, 70]; [44, 48]	8
7-MG	Dis-model2	[20]; [14]; [4]; [16]; [2]; [8]; [8]	[49, 64, 66]; [50, 54]; [39, 40]; [25]	8
	Dis-model1	[65]; [40]; [46]; [44]; [28]; [36]; [28]	[50, 79, 85]; [62, 70]; [44, 48]; [26]	8
Original	Cen-model2	70	[37, 49]; [31]; [27]; [25]; [17]; [24]; [21]	8
	Cen-model1	275		8

1) Accuracy of Proposed Algorithm

The calculation results of the binary variables under different models are given in Table VI.

TABLE VI
CALCULATION RESULTS OF BINARY VARIABLES UNDER DIFFERENT MODELS

System	Model	Binary variables	
		L	Number
2-MG	Dis-model2	[139, 143, 144, 150, 174, 176, 188, 192]	8
	Dis-model1	[278, 293, 296, 313, 371, 396, 377, 404]	8
3-MG	Dis-model2	[69, 77, 90, 92]; [75, 76, 89, 93]	8
	Dis-model1	[70, 86, 109, 115]; [83, 87, 129, 137]	8
4-MG	Dis-model2	[69, 77, 90, 92]; [50, 54]; [39, 40]	8
	Dis-model1	[70, 86, 109, 115]; [62, 70]; [44, 48]	8
7-MG	Dis-model2	[49, 64, 66]; [50, 54]; [39, 40]; [25]	8
	Dis-model1	[50, 79, 85]; [62, 70]; [44, 48]; [26]	8
Original	Cen-model2	[37, 49]; [31]; [27]; [25]; [17]; [24]; [21]	8
	Cen-model1		8

It can be observed that the location where the calculated binary variable is 1 corresponds to the bad measurement setting given in Table V, and the number of binary variables with 1 indicates the amount of bad measurements. Thus, the proposed algorithm is confirmed as obtaining the correct binary solution, and the bad measurement can then be effectively identified.

In addition, Fig. 7 shows the comparison between true values of voltage amplitude and corresponding results obtained by proposed algorithm for MG1 in each system. It shows that the proposed algorithm has a smaller error, and the obtained results can be directly used as accurate power flow results. The proposed algorithm can accurately identify the location of bad measurement and simultaneously obtain reasonable and accurate power flow results.

2) Efficiency of Proposed Algorithm

In the proposed algorithm, after each microgrid implements the independent solution in parallel through the interactive boundary coupling information to update their own multiplier and consistency variable information, the dual residual value is less than the convergence criterion value ε under a limited number of iterations. The convergence performance results and the iterative process of the proposed algorithm in each system are presented in Table VII and Fig. 8, respectively.

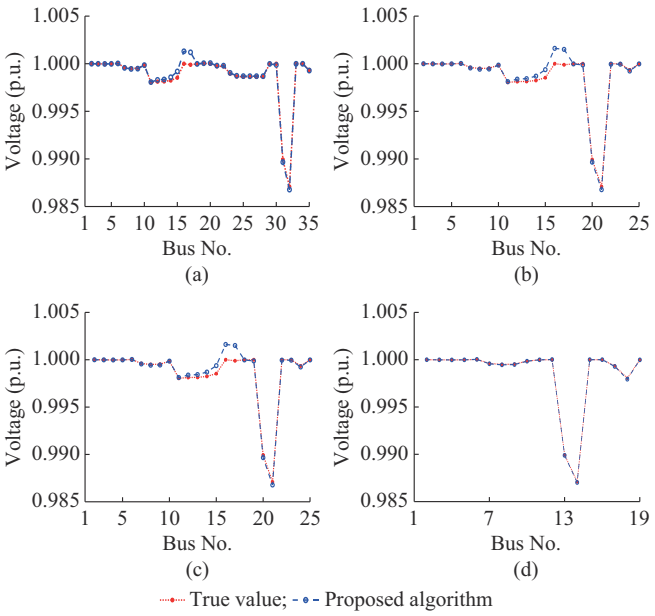

Fig. 7. Comparison between true values of voltage amplitude and corresponding results obtained by proposed algorithm for MG1 in each system. (a) 2-MG system. (b) 3-MG system. (c) 4-MG system. (d) 7-MG system.

TABLE VII
CONVERGENCE PERFORMANCE RESULTS UNDER DIFFERENT MODELS

System	Model	Number of iterations	T (s)	*err*
Original	Cen-model2	37	3.704	0.000094
	Cen-model1			
2-MG	Dis-model2	9	5.850	0.000220
	Dis-model1	3	65.360	0.012190
3-MG	Dis-model2	11	19.240	0.014580
	Dis-model1	12	389.780	0.030590
4-MG	Dis-model2	11	21.360	0.000360
	Dis-model1	12	331.010	0.033320
7-MG	Dis-model2	42	70.730	0.004920
	Dis-model1			

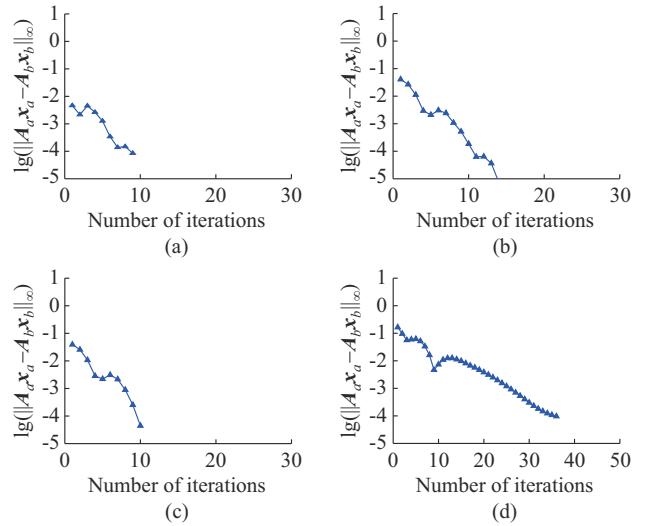

Fig. 8. Iterative process of proposed algorithm in each system. (a) 2-MG system. (b) 3-MG system. (c) 4-MG system. (d) 7-MG system.

The Cen-model1 fails to converge to the correct solution within a finite time due to the presence of numerous integer variables. By contrast, by filtering out some binary variables, the Cen-model2 iterates 37 times within 3.704 s to converge to the accurate solution. Therefore, the distributed solutions can be used for binary problems, and large-scale centralized decomposition with numerous binary variables can be considered for use in multiple small-scale systems containing a small number of binary variables for computation. Accordingly, the efficiency of the distributed IRSE optimization model is higher than that of the centralized one.

V. CONCLUSION

This study constructs an IRSE optimization model with MINLP, which overcomes the disadvantage of inaccurate results derived from different measurements when considering the same error as in the RSE optimization model. Second, to reduce the complexity of the IRSE optimization model, the number of binary variables is reduced based on projection statistics and normalized residual methods. This effectively avoids the problem of slow convergence or divergence of the algorithm caused by too many binary variables. In addition, to solve the aforementioned problem and meet the needs of NMGs in distributed energy management, an embedded consensus ADMM distribution algorithm based on OA is proposed, which can accurately identify bad measurements and obtain accurate SE results by communicating only boundary coupling information with neighbors. Finally, the simulation results verify that the IRSE optimization model and the proposed algorithm are accurate, computationally fast, and robust.

Although the IRSE optimization model can perform independent calculations for each microgrid, the process of updating each microgrid multiplier and consistent variable must be delayed until all microgrids are calculated in parallel. Because the calculation time of each microgrid is different, a microgrid with fast calculation cannot commence until a microgrid with slow calculation is completed. Therefore, we will consider asynchronously distributed calculations in future to further improve the effectiveness of the algorithm.

REFERENCES

[1] A. Majumdar and B. C. Pal, "Bad data detection in the context of leverage point attacks in modern power networks," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 2042-2054, May 2018.
[2] A. Abu and A. G. Expósito, *Power System State Estimation: Theory and Implementation*. New York: Marcel Dekker, 2004.
[3] A. Zakerian, A. Maleki, Y. Mohammadnian *et al*., "Bad data detection in state estimation using decision tree technique," in *Proceedings of 2017 Iranian Conference on Electrical Engineering* (*ICEE*), Tehran, Iran, May 2017, pp. 1037-1042.
[4] E. N. Asada, A. V. Garcia, and R. Romero, "Identifying multiple interacting bad data in power system state estimation," in *Proceedings of IEEE PES General Meeting*, San Francisco, USA, Jun. 2005, pp. 571-577.
[5] J. Zhao and L. Mili, "A framework for robust hybrid state estimation with unknown measurement noise statistics," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 1866-1875, May 2018.
[6] 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.
[7] M. R. Irving, "Robust state estimation using mixed integer program-

ming," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1519-1520, Aug. 2008.
[8] Y. Chen and J. Ma, "A mixed-integer linear programming approach for robust state estimation," *Journal of Modern Power Systems and Clean Energy*, vol. 2, no. 4, pp. 366-373, Jul. 2014.
[9] H. Zhang, B. Zhang, A. Bose *et al*., "A distributed multi-control-center dynamic power flow algorithm based on asynchronous iteration scheme," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1716-1724, Mar. 2018.
[10] S. Xia, Q. Zhang, J. Jing *et al*., "Distributed state estimation of multi-region power system based on consensus theory," *Energies*, vol. 12, no. 5, p. 900, Mar. 2019.
[11] T. Van Cutsem, J. L. Horward, and M. Ribbens-Pavella, "A two-level static state estimator for electric power systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 8, pp. 3722-3732, Aug. 1981.
[12] T. V. Cutsem and M. Ribbens-Pavella, "Critical survey of hierarchical methods for state estimation of electric power systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-102, no. 10, pp. 3415-3424, Oct. 1983.
[13] Y. Guo, L. Tong, W. Wu *et al*., "Hierarchical multi-area state estimation via sensitivity function exchanges," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 442-453, Jan. 2017.
[14] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617-1626, May 2013.
[15] W. Zheng and W. Wu, "An adaptive distributed quasi-newton method for power system state estimation," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5114-5124, Sept. 2019.
[16] W. Zhang, W. Liu, C. Zang *et al*., "Multiagent system-based integrated solution for topology identification and state estimation," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 714-724, Apr. 2017.
[17] A. K. Al-Othman and M. R. Irving, "Robust state estimator based on maximum constraints satisfaction of uncertain measurements," *Measurement*, vol. 40, no. 3, pp. 347-359, Apr. 2007.
[18] Z. Khan, R. B. Razali, H. Daud *et al*., "The largest studentized residual test for bad data identification state estimation of a power system," *ARPN Journal of Engineering and Applied Sciences*, vol. 10, no. 21, pp. 1-8, Nov. 2015.
[19] S. Boyd, N. Parikh, E. Chu *et al*., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1-122, Jan. 2011.
[20] T. Xu and W. Wu, "Accelerated ADMM-based fully distributed inverter-based volt/var control strategy for active distribution networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7532-7543, Dec. 2020.
[21] M. Kuchlbauer, F. Liers, and M. Stingl, "Outer approximation for mixed-integer nonlinear robust optimization," *Journal of Optimization Theory and Applications*, vol. 195, no. 3, pp. 1056-1086, Dec. 2022.
[22] J. Kronqvist, A. Lundell, and T. Westerlund, "The extended supporting hyperplane algorithm for convex mixed-integer nonlinear programming," *Journal of Global Optimization*, vol. 64, no. 2, pp. 249-272, Feb. 2016.
[23] A. Lundell and J. Kronqvist, "Polyhedral approximation strategies for nonconvex mixed-integer nonlinear programming in SHOT," *Journal of Global Optimization*, vol. 82, no. 4, pp. 863-896, Apr. 2022.
[24] J. A. E. Andersson, J. Gillis, G. Horn *et al*., "CasADi: a software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1-36, Jan. 2019.
[25] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, no. 1, pp. 25-57, Mar. 2006.
[26] P. Bonami, L. T. Biegler, A. R. Conn *et al*., "An algorithmic framework for convex mixed integer nonlinear programs," *Discrete Optimization*, vol. 5, no. 2, pp. 186-204, May 2008.

Zifeng Zhang received the B.S degree in information management system from North China Electric Power University Science & Technology College, Beijing, China, in 2016, and the Ph.D. degree in College of Information and Electrical Engineering, China Agricultural University, Beijing, China, in 2023. She is currently with the State Grid Information & Telecommunication Group Co., Ltd., Beijing Branch, Beijing, China. Her research interests include power system optimization, operation and control, and computer vision.

Yuntao Ju received the B.Sc. degree in mechanical engineering and the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2008 and 2013, respectively. In 2013, he was a Visiting Scholar with the University of Toronto, Toronto, Canada. In 2015, he joined China Electric Power Research Institute as a Research Fellow. He is currently a Professor with the School of Electrical and Control Engineering, North China University of Technology, Beijing, China. His research interests include hybrid energy system modeling, high-speed dynamic simulation, large-scale system parameter identification, state estimation, and uncertainty optimization.