iEnergy

ARTICLE

https://doi.org/10.23919/IEN.2023.0035

Topology and admittance estimation: Precision limits and algorithms

Ning Zhang[1], Yuxiao Liu[1] ✉, Fangyuan Si[2], Qingchun Hou[1], Audun Botterud[3] and Chongqing Kang[1]

ABSTRACT

Distribution grid topology and admittance information are essential for system planning, operation, and protection. In many distribution grids, missing or inaccurate topology and admittance data call for efficient estimation methods. However, measurement data may be insufficient or contaminated with large noise, which will fundamentally limit the estimation accuracy. This work explores the theoretical precision limits of the topology and admittance estimation (TAE) problem with different measurement devices, noise levels, and numbers of measurements. On this basis, we propose a conservative progressive self-adaptive (CPS) algorithm to estimate the topology and admittance. The results on IEEE 33 and 141-bus systems validate that the proposed CPS method can approach the theoretical precision limits under various measurement settings.

KEYWORDS

Distribution grids, topology estimation, admittance estimation, data-driven, Newton−Raphson method.

The vast integration of distributed energy resources and electric vehicles raises both economic and security concerns for modern distribution grids[1]. Smart grid operations such as state estimation (SE), demand response, voltage control, and pricing are increasingly implemented at the distribution level. However, accurate grid topology and line admittance, which are prerequisites for the above operations, are often unavailable in many medium- and low-voltage distribution grids. Therefore, efficient and accurate topology and admittance estimation (TAE) is essential in future distribution grids.

Many efforts have been made to estimate distribution grid topology and admittance. Most of these works use data collected from advanced metering infrastructures (AMIs) or micro-phasor measurement units (μPMUs). Some researches that only focus on topology identification use the statistical information of voltage magnitudes, such as covariances[2], mutual information[3], and the conditional independence test[4]. Other researches address the joint estimation of topology and line admittance by formulating the problem as maximum likelihood estimation[5–9]. The above works often make some assumptions to simplify the problem, including uncorrelated nodal power/current injections[3, 4, 10, 11], radial network topologies[2, 8, 10–13], sufficient phasor measurements[5, 6, 8, 9], or accurate voltage measurements[7]. These assumptions may hold in some cases but hinder practical implementation under more general distribution system cases. For instance, power/current injections can be highly correlated because of similar electricity consumption or rooftop solar PV generation patterns[14]. The distribution grids may contain loops or even be heavily meshed[15]. The measurement devices may not be sufficient at the distribution level, especially for μPMUs that contain phasor information[16].

In fact, TAE in distribution grids is extremely challenging because of the poor measurements and nonconvexity of power flow models[17, 18]. The TAE problem is also regards as generalized power system state estiamtion[19]. The best way to identify the precision limits of topology and admittance of distribution networks based on limited measurement devices and measurement precisions remains unresolved. Some recent works discuss the fundamental limits of the TAE problem. Moffat et al. proved that the admittance matrix cannot be estimated without any prior knowledge when the system contains some zero power injection buses[8]. However, they did not consider any physical knowledge (e.g., the admittance matrix is symmetric and sparse), which could greatly improve the estimation results. Li et al. provided a theoretical relationship among the number of measurements, prior knowledge, and probability of estimation error[9]. They proved a worst-case sample complexity for the linear graph learning task, i.e., how much data are required to guarantee a certain accuracy (with probability). The above researches[8, 9] only address the estimation limits for linear learning tasks. That is, the voltage and current magnitudes and angles are assumed to be available at all buses so that the estimation problem can be formulated by the linear Ohm's law. Grotas et al. proposed the lower bound for transmission grids under the DC power flow (DCPF) setting[20]. The DCPF approximation simplifies the problem but introduces great error to the estimation problem; this error is even larger for distribution grids because of the high R/X ratio and the requirement for phasor measurements.

This work addresses a more practical setting: to derive the theoretical precision limits with different measurement devices and measurement precisions. We first derive the Cramér-Rao Lower Bound (CRLB)[21] that can evaluate the best possible TAE precisions under given measurement devices, noise levels, and the number of data. Then, we show that designing an efficient algorithm for the TAE problem is mathematically difficult. The TAE problem is far more memory consuming, ill-conditioned, and nonconvex than the traditional SE problem[22]. Some traditionally well-behaved methods can easily diverge (Newton's method) or suffer from an extremely slow convergence speed (gradient-based methods), even

[1]Department of Electrical Engineering, Tsinghua University, Beijing 100084, China; [2]School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China; [3]Laboratory of Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA

Address correspondence to Yuxiao Liu, lyx12@tsinghua.org.cn

under a simple 3-bus case. Furthermore, we propose an algorithm for the TAE problem. The method contains a first-order optimizer and a second-order optimizer to combine stability (or conservative) and fast convergence (or progressive) features within one method. The method also contains a hybrid line search strategy to self-adaptively tune the weights of the first-order and second-order optimizers. To this end, we name the proposed algorithm the conservative progressive self-adaption (CPS) algorithm. Case studies on IEEE 33 and 141-bus systems show that the proposed CPS method can approach the theoretical precision limits under different experimental settings.

This work focuses on systems with a balanced power flow setting and Gaussian measurement noise. The framework can be extended to address the unbalanced three-phase estimation problem and some non-Gaussian noise. We also do not consider active injection approaches[23, 24] because they require sufficient active devices (e.g., smart inverters) and may not be compliant with the grid code[25].

In short, the contributions of this work are as follows:

(1) We quantify the theoretical precision limits for the distribution grid TAE problem. The proposed method identifies precision limits given different measurement devices, noise levels, numbers of measurements, and prior topology knowledge.

(2) We propose the CPS algorithm, which is specifically designed for the memory-heavy, ill-conditioned, and non-convex estimation problem. The method can approach the theoretical precision limits under a lack of voltage angle measurements and different levels of measurement noise.

The remainder of this paper is organized as follows. Section 1 introduces the method for evaluating the precision limits for the TAE problem. Section 2 demonstrates the proposed CPS algorithm. Section 3 provides case studies. Finally, Section 4 draws conclusions.

1 Precision limits for the TAE problem

1.1 Problem formulation

The TAE of a distribution grid can be formulated as a maximum likelihood problem using AC power flow equations.

$$\min_{G_{ij},B_{ij},V_i^t,\theta_i^t} \text{TAE} = \sum_{t=1}^{T}\sum_{i\in M_P}\frac{(P_i^t-\hat{P}_i^t)^2}{\sigma_{P_i}^2}+\sum_{i\in M_Q}\frac{(Q_i^t-\hat{Q}_i^t)^2}{\sigma_{Q_i}^2} + \sum_{i\in M_V}\frac{(V_i^t-\hat{V}_i^t)^2}{\sigma_{V_i}^2}+\sum_{i\in M_\theta}\frac{(\theta_i^t-\hat{\theta}_i^t)^2}{\sigma_{\theta_i}^2}, \quad (1a)$$

$$\text{with } \hat{P}_i^t = \hat{V}_i^t\sum_{j=1}^{N}\hat{V}_j^t\left(G_{ij}\cos\hat{\theta}_{ij}^t + B_{ij}\sin\hat{\theta}_{ij}^t\right),$$

$$\hat{Q}_i^t = \hat{V}_i^t\sum_{j=1}^{N}\hat{V}_j^t\left(G_{ij}\sin\hat{\theta}_{ij}^t - B_{ij}\cos\hat{\theta}_{ij}^t\right), \quad (1b)$$

where G_{ij}/B_{ij} denote the (i, j)th element of the conductance/susceptance matrix, P_i^t/Q_i^t denote the active/reactive power injection measurements of bus i snapshot t, V_i^t/θ_i^t denote the voltage magnitude/angle measurements of bus i snapshot t, $\hat{P}_i^t/\hat{Q}_i^t/\hat{V}_i^t/\hat{\theta}_i^t$ denote the evaluated values of $P_i^t/Q_i^t/V_i^t/\theta_i^t$, $\sigma_{P_i}/\sigma_{Q_i}/\sigma_{V_i}/\sigma_{\theta_i}$ denote the standard deviations of the corresponding measurements, and $M_P/M_Q/M_V/M_\theta$ denote the bus sets for which the corresponding $P/Q/V/\theta$ measurements are available. In the above problem (1), the estimation of G_{ij} and B_{ij} is equivalent to the estimation of topology and admittance, respectively, where zero values denote the disconnection of two buses and nonzero values denote admittance. We also recover all the state variables V_i^t and θ_i^t as a byproduct. The objective is to minimize the loss function (1a), which is the weighted squared loss of the data from the measurement sets $M_P/M_Q/M_V/M_\theta$. The measurement sets can be any combinations of the buses in the distribution grid, from the universe to the empty set. The standard deviations $\sigma_{P_i}/\sigma_{Q_i}/\sigma_{V_i}/\sigma_{\theta_i}$ reflect the precision of different measurements. The estimated values $\hat{P}_i^t/\hat{Q}_i^t/V_i^t/\theta_i^t$ satisfy the AC power flow equations in (1b). By neglecting the shunt admittance in the distribution grid[5,8], G_{ij} and B_{ij} become:

$$G_{ij} = G_{ji} = -g_{ij},\ G_{ii} = \sum_j g_{ij}, \quad (2a)$$

$$B_{ij} = B_{ji} = -b_{ij},\ B_{ii} = \sum_j b_{ij}, \quad (2b)$$

where g_{ij}/b_{ij} denotes the conductance/susceptance of line (i, j).

We then provide a general formulation of the estimation model. We formulate all the measurements as an $M\times 1$ vector $\boldsymbol{z}$:

$$\boldsymbol{z} = \left[\left\{\{P_i^t\}_{i\in M_P}\{Q_i^t\}_{i\in M_Q}\{V_i^t\}_{i\in M_V}\{\theta_i^t\}_{i\in M_\theta}\right\}_{t=1\sim T}\right]^{\mathrm{T}}. \quad (3)$$

Similarly, we define vector $\boldsymbol{\sigma}$ with each element corresponding to the standard deviation of measurement $\boldsymbol{z}$. The estimated variables contain not only the voltage states but also the network admittances, with the $S\times 1$ state vector $\boldsymbol{x}$ formulated as

$$\boldsymbol{x} = \left[\{g_{ij}\}_{(i,j)\in S_T}\{b_{ij}\}_{(i,j)\in S_T}\left\{\{\hat{V}_i^t\}_{i\in S_B}\{\hat{\theta}_i^t\}_{i\in S_B}\right\}_{t=1\sim T}\right]^{\mathrm{T}}, \quad (4)$$

where S_T denotes the set of possible connected buses and S_B denotes the set of all the buses. If no prior topology information is available, S_T contains all of the possible branches between any two buses. From Eqs. (1)–(4), we obtain the measurement model of the TAE problem:

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{\varepsilon}, \quad (5)$$

where $\boldsymbol{\varepsilon}$ denotes the $M\times 1$ vector of measurement error.

1.2 Cramér-Rao lower bound

Equation (5) formulates the TAE problem into a standard unbiased estimation problem. In this subsection, we use the Cramér-Rao lower bound to estimate the precision limit of the problem. The purpose of estimating the precision limit is to guide the algorithm of the TAE problem. CRLB is a classic theory to calculate the lower bound of the estimation variances of any unbiased estimation problem[21] and has rarely been applied in power system analysis.

Wang et al. derived the CRLB for the power system SE problem[26]. Grotas et al. derived the CRLB of TAE using a simplified linear DCPF model[20]. Damavandi et al.[27] and Xygkis et al.[28] developed the Fisher information-based approach, which is closely related to CRLB theory, for the meter placement problem. In this section, we derive the CRLB for the TAE problem under the accurate nonlinear AC power flow formulation. Compared with existing problems[20, 26–28], the present problem is a large-scale multiple snapshot nonlinear estimation problem with unknown models and state variables.

Theorem 1 (Cramér-Rao lower bound for topology and admittance estimation)**.** For the estimation problem (5) with $\boldsymbol{\varepsilon}\sim N(0,\boldsymbol{\sigma})$, the covariance matrix $\boldsymbol{C}$ of any unbiased estimator of $\boldsymbol{x}$ satisfies:

$$\boldsymbol{C}-\boldsymbol{F}^{-1}\geqslant 0, \tag{6a}$$

$$\text{with } \boldsymbol{F}=\boldsymbol{H}^{\mathrm{T}}diag(\boldsymbol{\sigma}^{-2})\boldsymbol{H},\ \boldsymbol{H}=\frac{\partial \boldsymbol{h}(\boldsymbol{x})}{\partial \boldsymbol{x}^{\mathrm{T}}}, \tag{6b}$$

where $\boldsymbol{F}$ denotes the $S\times S$ symmetric Fisher information matrix[21] and $\boldsymbol{\sigma}$ denotes the $M\times 1$ vector of measurement noise standard deviations.

See Appendix A.1 for the proof. Since $\boldsymbol{C}-\boldsymbol{F}^{-1}$ is semidefinite, the variance of $\boldsymbol{x}$, which reflects the estimation precision, is lower bounded by the diagonal elements of $\boldsymbol{F}^{-1}$:

$$\boldsymbol{\sigma}_x\geqslant \boldsymbol{\sigma}_x^{cr}=diag(\boldsymbol{F}^{-1}). \tag{7}$$

It can be intuitively interpreted that the more information the measurement $\boldsymbol{z}$ carries about the estimated state variable $\boldsymbol{x}$, the lower variance the estimators can obtain.

The CRLB provides an effective way of estimating the variance (in other words, precision limit) of the estimated variable x, which includes the admittances. The problem now becomes how to calculate $\boldsymbol{F}^{-1}$.

1.3 Fisher information matrix partition

The challenge of using the CRLB to calculate the precision limit of the TAE problem is that the dimension of the Fisher information matrix is very large. For an N-bus distribution grid system, S_T contains at most $C_N^2=N(N-1)/2$ elements, and M_V and M_θ contain at most N elements. In this case, the number of state variables is $S=N(N-1)+2NT$. Simply estimating the 33-bus distribution system using 200 snapshots of measurement data, we have to calculate the Fisher information matrix and its inverse with 14256×14256 dimensions.

However, when calculating the precision limit of the TAE problem, we only need to calculate the $\boldsymbol{F}^{-1}$ with respect to the admittances to be estimated, which is a small part of the inverse of the Fisher information matrix.

To address the heavy memory consumption challenge, we partition the Fisher information matrix and calculate its inverse in a memory-saving manner.

Before analyzing the structure of $\boldsymbol{F}$, we first analyze the structure of $\boldsymbol{H}$. Figure 1 shows the structure of the partitioned $\boldsymbol{H}$. $\boldsymbol{H}$ is the Jacobi matrix of $\boldsymbol{h}(\boldsymbol{x})$ with respect to $\boldsymbol{x}$. It has a row number of M (the same as $\boldsymbol{z}$) and a column number of S (the same as $\boldsymbol{x}$). $\boldsymbol{H}$ can be decomposed into a block matrix by the partition of $\boldsymbol{x}$ and $\boldsymbol{z}$, where the subscript $\boldsymbol{a}$ corresponds to the admittance $[\{g_{ij}\}\{b_{ij}\}]$ and the subscript $\boldsymbol{v}$ (from 1 to $\boldsymbol{T}$) corresponds to the voltage magnitudes and angles from all the snapshots $\left[\{\hat{V}_i^t\}_{i\in S_B}\{\hat{\theta}_i^t\}_{i\in S_B}\right]$. According to the partition of $\boldsymbol{z}$ (measurement $[\{P_i^t\}_{i\in M_P}\{Q_i^t\}_{i\in M_Q}\{V_i^t\}_{i\in M_V}\{\theta_i^t\}_{i\in M_\theta}]$ from 1 to $\boldsymbol{T}$), the block corresponding to the subscript $\boldsymbol{a}$ is a full matrix since the admittance appears in each $\boldsymbol{h}(\boldsymbol{x})$. However, the block corresponding to the subscript $\boldsymbol{v}$ is a sparse matrix since only the measurement from one snapshot has the state variables $\hat{V}_i^t$ and $\hat{\theta}_i^t$ on the same snapshot. Therefore, the blocks that correspond to different snapshots are zero.

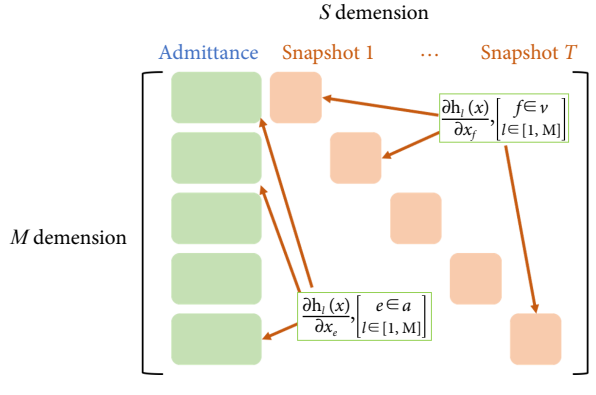

Fig. 1 The structure of the partitioned *H*.

According to the expression of $\boldsymbol{F}$ with respect to $\boldsymbol{H}$ (Equation (6b)), the nonzero blocks of $\boldsymbol{F}$ can be written as follows:

$$\boldsymbol{F}=\begin{bmatrix}\boldsymbol{F}_{aa} & \boldsymbol{F}_{av}\\ \boldsymbol{F}_{av}^{\mathrm{T}} & \boldsymbol{F}_{vv}\end{bmatrix}=\begin{bmatrix}\boldsymbol{F}_{aa} & \boldsymbol{F}_{a1} & \boldsymbol{F}_{a2} & \cdots & \boldsymbol{F}_{aT}\\ \boldsymbol{F}_{a1}^{\mathrm{T}} & \boldsymbol{F}_{11} & & & \\ \boldsymbol{F}_{a2}^{\mathrm{T}} & & \boldsymbol{F}_{22} & & \\ \cdots & & & \cdots & \\ \boldsymbol{F}_{aT}^{\mathrm{T}} & & & & \boldsymbol{F}_{TT}\end{bmatrix}, \tag{8}$$

Note that we formulate (8) in a symmetric way. From (8), we can observe that the Fisher information matrix is a sparse matrix because the voltage states in different snapshots are decoupled. In other words, the voltage measurements in one snapshot do not contain any information about the voltage in any other snapshots. We can take advantage of this feature and calculate the CRLB of $[\{g_{ij}\}\{b_{ij}\}]$ in a memory-saving manner.

Theorem 2 (Cramér-Rao lower bound with matrix partition). The covariance matrix $\boldsymbol{C}_a$ of any unbiased estimator of $\boldsymbol{x}_a=[\{g_{ij}\}_{(i,j)\in S_T}\{b_{ij}\}_{(i,j)\in S_T}]$ satisfies the following:

$$\boldsymbol{C}_a-\boldsymbol{F}_a^{-1}\geqslant 0, \tag{9a}$$

$$\text{with } \boldsymbol{F}_a=\boldsymbol{F}_{aa}-\sum_{t=1}^{T}\boldsymbol{F}_{at}\boldsymbol{F}_{tt}^{-1}\boldsymbol{F}_{at}^{\mathrm{T}}. \tag{9b}$$

Proof. The inverse of the Fisher information matrix is

$$\boldsymbol{F}^{-1}=\begin{bmatrix}\boldsymbol{F}_{aa} & \boldsymbol{F}_{av}\\ \boldsymbol{F}_{av}^{\mathrm{T}} & \boldsymbol{F}_{vv}\end{bmatrix}^{-1}=\begin{bmatrix}\boldsymbol{I}_{aa} & \boldsymbol{I}_{av}\\ \boldsymbol{I}_{av}^{\mathrm{T}} & \boldsymbol{I}_{vv}\end{bmatrix}. \tag{10}$$

From **Theorem 1**, the CRLB of $\boldsymbol{x}_a$ is given by

$$\boldsymbol{C}_a-\boldsymbol{F}_a^{-1}\geqslant 0,\ \text{with } \boldsymbol{F}_a^{-1}=\boldsymbol{I}_{aa}. \tag{11}$$

where $\boldsymbol{F}_a$ is the Schur complement of block $\boldsymbol{F}_{vv}$:

$$\boldsymbol{F}_a=\boldsymbol{F}_{aa}-\boldsymbol{F}_{av}\boldsymbol{F}_{vv}^{-1}\boldsymbol{F}_{av}^{\mathrm{T}}. \tag{12}$$

Since $\boldsymbol{F}_{vv}$ is a block diagonal matrix, we can compute the inverse in a decoupled way and finish the proof:

$$\boldsymbol{F}_{av}\boldsymbol{F}_{vv}^{-1}\boldsymbol{F}_{av}^{\mathrm{T}}=\sum_{t=1}^{T}\boldsymbol{F}_{at}\boldsymbol{F}_{tt}^{-1}\boldsymbol{F}_{at}^{\mathrm{T}}. \tag{13}$$

The structure of the Fisher information matrix and the method of calculating the inverse are shown in Figure 2. The required space to calculate the CRLB is largely reduced to $S'\times S'$, $S'=N(N-1)+2N$. Since we decouple the states from different snapshots, the space complexity does not increase with the increase in snapshots. We only need to compute the inverse of matrices with dimensions of $N(N-1)\times N(N-1)$ ($\boldsymbol{F}_{aa}$) and $2N\times 2N$ ($\boldsymbol{F}_{tt}$). Recalling the 33-bus example, the dimensions of the matrices are 1056×1056 and 66×66, much smaller than 14256×14256.

In this way, we can efficiently calculate the inverse of the Fisher Information matrix to estimate the precision limit of the admit-

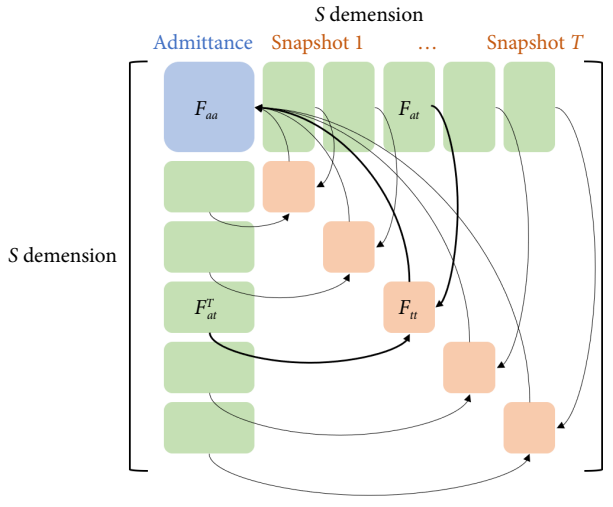

Fig. 2 The structure of the partitioned Fisher information matrix F.

tance. It is further used in the TAE problem to guide the topology search and the optimization of the estimation.

2 CPS algorithm

2.1 Geometric illustration of the TAE problem

As shown in Section 1.1, the estimation problem is an unconstrained nonlinear optimization problem and thus belongs to the same category as the power system SE problem. However, the TAE problem is much more difficult to solve, and the traditional methods used for SE problems have very poor performance on TAE problems.

In Figure 3, we provide an illustrative example of why the TAE problem is much more difficult to solve than the SE problem: (1) The TAE problem is more nonconvex because many more variables (both the model parameters and the state variables) are unknown and should be optimized. The multiplications of more decision variables, as shown in (1b), incur more severe nonconvexity. (2) The TAE problem is more ill-conditioned than the SE problem. The decision variables in the TAE problem have very different scales. For example, the voltage magnitudes are often approximately 1 p.u., and the voltage angles are often below 1.0 rad. However, the admittance in distribution grids can be very large (from 10^2 to more than 10^4), especially when the lines are short. As a result, the parameter space can have very different scales in different directions, which makes the problem ill-conditioned. Hence, we use a slightly nonconvex contour map to represent the SE problem in Figures 3(a) and 3(c). For comparison, we use a highly nonconvex and ill-conditioned contour map to represent the TAE problem in Figures 3(b) and 3(d).

There are two types of methods to solve the unconstrained differentiable optimization problem: first-order optimizations (gradient-based methods) and second-order optimizations (Newton or quasi-Newton-based methods)[29]. The first-order optimizations use the local gradient information and iteratively search the solution. With proper step length, the first-order optimization is stable because the direction of the gradient can reduce the loss function (1a). Such a strategy may work for SE problems, as shown in Figure 3(a), but may become trapped for TAE problems, as shown in Figure 3(b). On the other hand, the second-order optimizations search the solution iteratively by using the information from the second-order Taylor approximations. This kind of method is widely adopted and effective in SE problems[22]. It has a fast convergence speed for the problem in Figure 3(c). However, second-order optimizations can easily diverge, as shown in Figure 3(d).

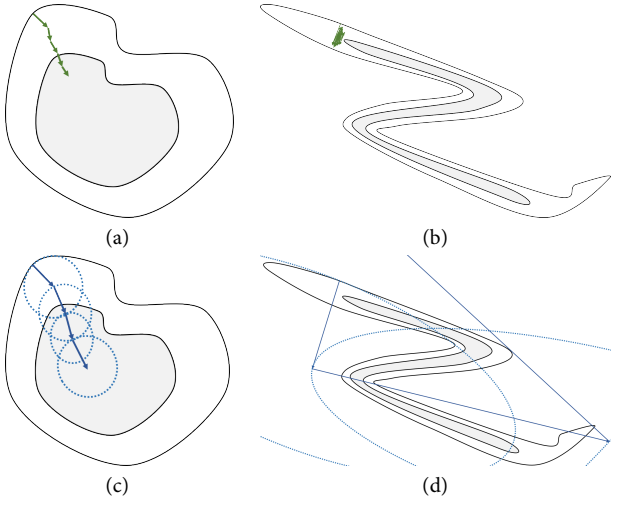

Fig. 3 Geometric illustration of the SE problem and TAE. The 2-D space denotes the parameter space. The black lines are contour lines of the loss function value in (1a). The blue dashed lines represent the local second-order Taylor approximations. (a) First-order optimization for SE. **(b)** First-order optimization for TAE. **(c)** Second-order optimization for SE. **(d)** Second-order optimization for TAE.

2.2 Framework of the algorithm

The framework of the proposed CPS algorithm is shown in Figure 4. We first obtain an initial value of the admittance by evaluating a simplified model. Then, we propose an optimizer that combines the advantages of both first-order optimizations and second-order optimizations. We take advantage of the stability (or conservative) and the fast convergence (or progressive) from the two methods. We also propose a hybrid search strategy to self-adaptively tune the weights of the first-order optimizer and the second-order optimizer. Next, we update the topology each time after the iteration converges. Finally, we stop the algorithm when there is no need to update the topology. The following sections will describe the details of each step of the framework.

In Figure 5, we also provide a geometric illustration of why the CPS method can have good performance for the aforementioned highly nonconvex and ill-conditioned problem. The CPS method can be both conservative, as it takes the gradient descent direction

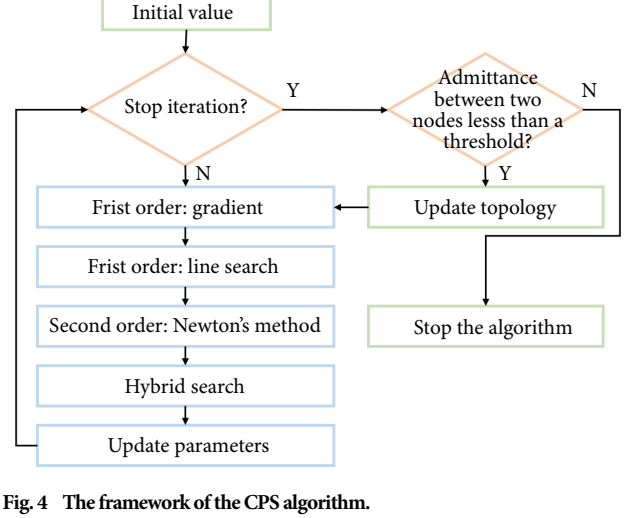

Fig. 4 The framework of the CPS algorithm.

when the second-order solution is not stable, and progressive, as the second-order solution greatly reduces the loss function (1a). In other words, the CPS method can automatically choose a proper "mode" between the first-order solution and the second-order solution.

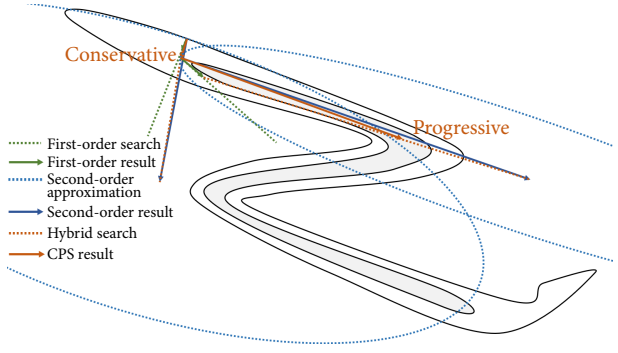

Fig. 5 Geometric illustration of the CPS method.

2.3 Initial value

The proposed CPS algorithm starts by obtaining an initial value of state vector $\boldsymbol{x}$ from the current measurements. Note that measurement availabilities can vary greatly, and there are no general methods to obtain the initial value under all circumstances. Since the initial value does not require a mesh network setup or accurate parameters, we refer to Refs. [5, 7, 8, 10–12] for initial value estimation under various conditions. Even the initial values can be set as the values from the grid planning files[5]. In this work, we extend the method in Ref. [7], which is explained in Appendix A.2.

2.4 First-order optimization

The first-order optimization used in this work is modified from the adaptive moment estimation (Adam) optimizer[30], which is widely applied and most favored in the training of deep learning networks. The Adam optimizer combines the advantages of the momentum method and the root mean square propagation (RMSProp) method[30]. On the one hand, Adam uses the moving average of the gradient to prevent oscillations during iterations. On the other hand, Adam rescales the gradient by dividing the moving average of the squared gradients so that the gradient of each variable is normalized to one or minus one.

On this basis, we improve Adam by rescaling the gradient. Adam rescales every gradient equally to one or minus one. However, the voltage magnitudes, angles, and admittances have very different scales; and thus, Adam can have poor performance. Instead, we first normalize the gradient and use the approximated CRLB to rescale the gradient. The details are shown in Algorithm 1 of the Appendix A.3.

2.5 Second-order optimization

The second-order optimization used in this work is modified from Newton's method[22, 29].

$$\boldsymbol{x}^k = \boldsymbol{x}^{k-1} - \boldsymbol{g}^k(\boldsymbol{F}^k)^{-1}. \tag{14}$$

Different from the SE problem, the TAE problem is a multiple snapshot estimation problem. The challenge for Newton's method is the large memory consumption when calculating $\boldsymbol{g}^k(\boldsymbol{F}^k)^{-1}$. We can derive from **Theorem 2** that $\boldsymbol{g}^k(\boldsymbol{F}^k)^{-1}$ can also be calculated in a memory-saving manner.

Theorem 3 (Second-order update with low memory consumption)**.** For the Fisher information matrix $\boldsymbol{F}$ in the form of (8) and the gradient vector in the form of

$$\boldsymbol{g} = \begin{bmatrix}\boldsymbol{g}_a^{\mathrm{T}} & \boldsymbol{g}_v^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}} = \begin{bmatrix}\boldsymbol{g}_a^{\mathrm{T}} & \boldsymbol{g}_1^{\mathrm{T}} & \boldsymbol{g}_2^{\mathrm{T}} & \cdots & \boldsymbol{g}_T^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}, \tag{15}$$

$\boldsymbol{g}(\boldsymbol{F})^{-1}$ can be calculated in a memory saving manner:

$$\boldsymbol{d} = -\boldsymbol{g}(\boldsymbol{F})^{-1} = \begin{bmatrix}\boldsymbol{d}_a^{\mathrm{T}} & \boldsymbol{d}_1^{\mathrm{T}} & \boldsymbol{d}_2^{\mathrm{T}} & \cdots & \boldsymbol{d}_T^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}, \tag{16a}$$

$$\text{with } \boldsymbol{d}_a = (\boldsymbol{F}_{aa} - \sum_{t=1}^{T}\boldsymbol{F}_{at}\boldsymbol{F}_{tt}^{-1}\boldsymbol{F}_{at}^{\mathrm{T}})^{-1}(\boldsymbol{g}_a - \sum_{t=1}^{T}\boldsymbol{F}_{at}\boldsymbol{F}_{tt}^{-1}\boldsymbol{g}_t), \tag{16b}$$

$$\boldsymbol{d}_t = (\boldsymbol{F}_{tt} - \boldsymbol{F}_{at}^{\mathrm{T}}\boldsymbol{F}_{aa}^{-1}\boldsymbol{F}_{at})^{-1}(\boldsymbol{g}_t - \boldsymbol{F}_{at}^{\mathrm{T}}\boldsymbol{F}_{aa}^{-1}\boldsymbol{g}_a). \tag{16c}$$

Note that for simplicity of notation, we drop the superscript of k that denotes the kth iteration.

Recall that the TAE problem is ill-conditioned due to different scales of admittance, voltage magnitudes and angles. The inversion of $(\boldsymbol{F}_{aa} - \sum_{t=1}^{T}\boldsymbol{F}_{at}\boldsymbol{F}_{tt}\boldsymbol{F}_{at}^{T})$ may incur numerical instability in (16b). Therefore, we use the Moore Penrose inverse [31] instead.

2.6 Hybrid line search

Neither the first-order optimization nor the second-order optimization can provide a stable and efficient search for the highly nonconvex and ill-conditioned TAE problem. We provide a hybrid line search strategy to combine the advantages of the two optimization methods. This strategy is inspired by the Armijo Goldstein condition[32], a.k.a., the backtracking line search. We modify the backtracking line search by adding the hybrid search strategy of first-order and second-order directions.

In detail, we first search for a step along the first-order direction that satisfies both criteria of the Armijo Goldstein condition. That is, we start with a sufficiently large step (thus satisfying criterion 1) and then iteratively decrease the step size until the decrease in the loss function is sufficiently large to correspond to the step size (thus satisfying criterion 2). This approach guarantees a stable searching solution because the gradient direction will decrease the loss function. However, compared with the second-order direction, the step size of the first-order direction may sometimes be small, which may jeopardize criterion 1. We then search the step along the line from the second-order result to the first-order result. We start by assigning a larger weight to the second-order result and then iteratively decrease the weight of the second-order direction and increase the first-order direction. Such a search keeps criterion 2 (since the loss function drop is at least as large as the first-order direction drop) while improving criterion 1 since we open the possibility of other search directions to avoid the drawback of the first-order direction. In this way, we maintain a sufficiently large proportion of the second-order optimizer as long as it is stable.

The geometric illustration is shown in Figure 6. The details of the hybrid line search are shown in Algorithm 2 of the Appendix A.3.

2.7 Update the topology

We estimate the admittances of all possible bus pairs. Some pairs are disconnected and have zero admittances. Therefore, we should update the topology and set the branches with small admittances to zero. We set small admittances to zero according to the CRLB calculation. Specifically, after each hybrid line search converges, we compare the admittance of the branches to the self-admittance of the bus that the branches connect and its CRLB. When the mutual admittance of two buses is small enough and its CRLB is

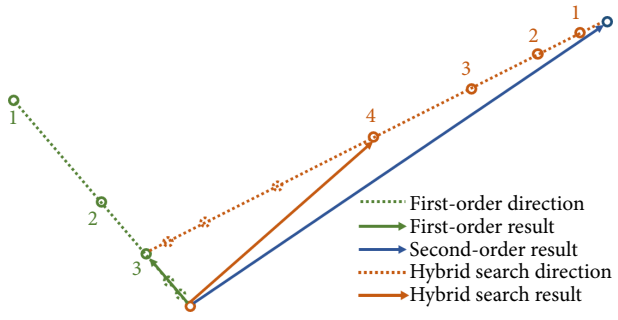

Fig. 6 Geometric illustration of hybrid search.

small enough (which assures that the estimate is confident), there is considered to be no branch connection between the corresponding two buses. Numerical experiments show that using line susceptance is more effective than using line conductance. Specifically, the ratio of line susceptance to self-admittance is used to judge whether the admittance is small enough.

It is common that some prior topology information of the distribution grids is available[7, 9], i.e., some bus pairs are impossible to connect. Consequently, we can set the corresponding admittances to zero from the beginning to accelerate the iterations. However, it is not a prerequisite of these algorithms.

3 Case study

The power load data are from the Commission for Energy Regulation in Ireland[33]. We simulate the power system operational data with the aid of MATPOWER 7.0[34]. The simulation strategy is the same as in Ref. [35]. White Gaussian noise is generated using the Monte Carlo simulation and is then added to the data to simulate the measurement error. In all our case studies, we set our hyperparameters as $\alpha = 0.9$, $r_0 = -5$, $r_{\max} = 20$, $\beta = 5$, $\eta = 0.01$, $\gamma = 10^{-5}$, $\gamma_{\text{topo}} = 0.05$, and $\gamma_{\text{cr}} = 0.1$.

3.1 Convergence of the optimizer

We first use a simple 3-bus system in Figure 7 to compare the convergence of the proposed method with other methods. Each of the load profiles of buses 2 and 3 is generated using a per-unit value of 100 aggregated residential loads. We assume that the topology of the 3-bus system is known and only estimate the admittance. We estimate the admittance with 50 snapshots of data, with P, Q, and V measurements under 0.1% noise.

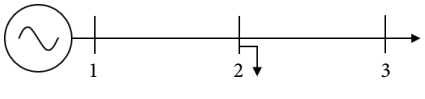

Fig. 7 The 3-bus system.

We compare the losses during iterations with three methods: the first-order optimization in Section 2.4, the second-order optimization in Section 2.5, and the proposed CPS algorithm. As shown in Figure 8, both the first-order and the second-order optimization have very poor performances even with a very simple 3-bus system. The first-order optimization has a very slow convergence speed. The loss is over 10^3 and fails to decrease after 10000 iterations. The second-order optimization is not stable, and the loss values oscillate in the range of $10^8 \sim 10^{10}$. However, the proposed CPS method converges in only 63 iterations with loss approaching 3×10^2. The performance of the compared three methods is thoroughly explained by the illustrations in Figures 3 and 5. The proposed CPS method shows extraordinary performance in the TAE problem.

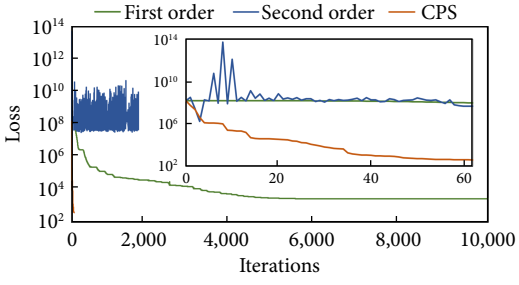

Fig. 8 The loss of the 3-bus system during the iterations.

3.2 The bound attainability

We then use the 12.66 kV 33-bus system[36] to demonstrate how well the CPS method can approach the theoretical precision limits in Section 1. The load profiles of each bus are generated using the per-unit value of 100 aggregated residential loads. We assume we have the P, Q, V, and θ measurement from all buses with 0.1% measurement error. The snapshot of the measurement is 120. We run the CPS algorithm 100 times under different randomly generated measurement noise and obtain the mean absolute error of admittances compared against their true value. We also compare the admittance estimation error of the CPS algorithm with the first-order optimization and the second-order estimation. Furthermore, we calculate the Cramer Rao bound-based theoretical precision limit calculated by the proposed Fisher information matrix partition method in Section 1. All of the admittance estimation errors and the theoretical precision limits are compared in Figure 9. The results show that the proposed CPS method is 1-3 orders of magnitude more accurate than the methods that rely on the first-order or second-order algorithm only (note that the ordinate is logarithmic). In particular, the first-order algorithm has the highest error on average due to its weakness in the complex nonconvex problem. The second-order algorithm has a slightly smaller error; however, it has higher variance under different randomly generated measurement noise due to its instability of convergence. By comparison, the proposed CPS algorithm has the lowest error, especially its worst case error, which is significantly smaller than for the first-order and second-order algorithms, showing its robustness among different cases. Its mean error is close to the theoretical precision limit. It should be noted that in some cases, the admittance estimation error of some buses is smaller than the theoretical precision limit because the precision limit is theoretical, where the value denotes the precision of infinite times of estimation with infinite measurements, while the estimation error is only a one-shot estimate based on a finite number of measurements.

3.3 Performance under different settings

We use the 33-bus system and the 141-bus system[37] to show the precision limits and the performance of the CPS algorithm under different settings. The load profiles of each bus are generated using the per-unit value of 100 aggregated residential loads. We compare the following different settings, as shown in Table 1. The experimental settings include the types of sensors, the noise level, and prior topology knowledge.

We use the ratio of the wrong branch number to the true branch number as the error of topology estimation:

$$E^t = \frac{N_{\text{wrong}}}{N_{\text{true}}} \times 100\% \tag{17}$$

The topology precision limit denotes the minimum wrong

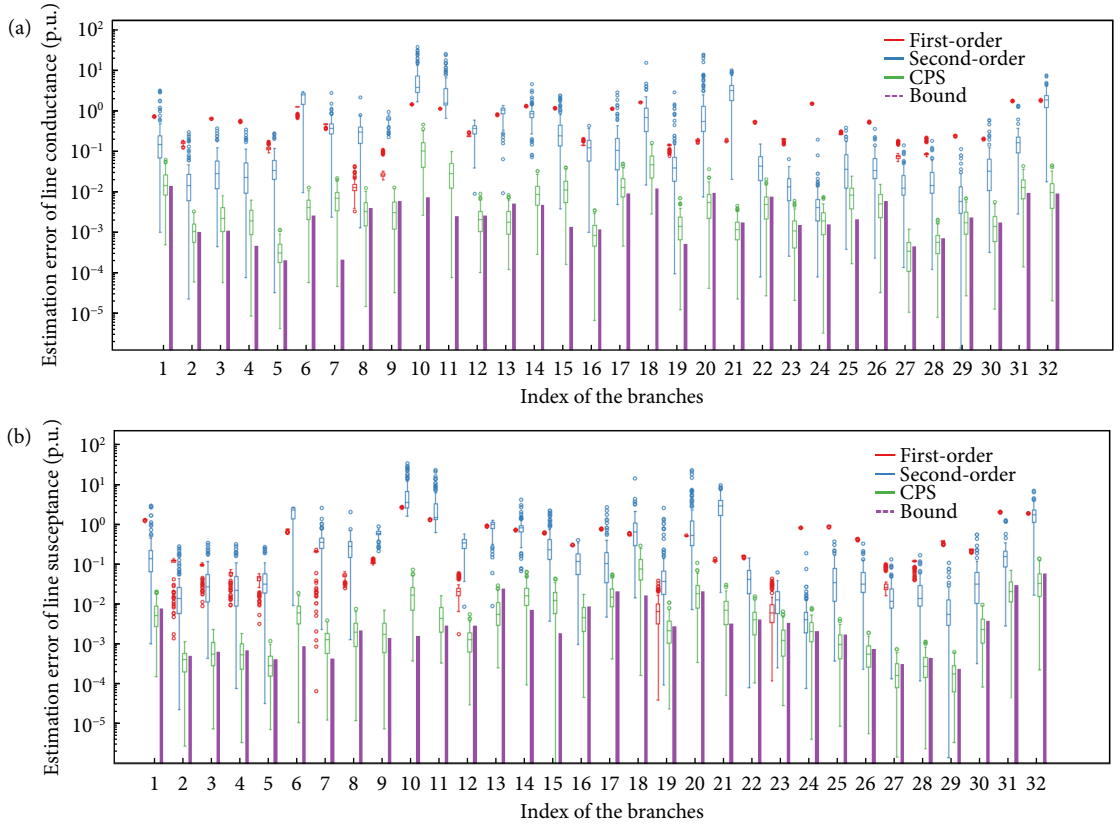

Fig. 9 The histogram of the admittance estimation error. (a) The error of line conductance. **(b)** The error of line susceptance.

Table 1 The theoretical precision limits and performance of CPS under different settings

System	Snapshots	Sensors	Noise	Prior topology knowledge	Topology precisionlimits (%)	Topology estimation error (%)	Admittance precisionlimits (%)	Admittance estimation error (%)
33-bus	120	P, Q, V, θ	0.1	No	0	0	0.264	0.236
33-bus mesh	120	P, Q, V, θ	0.1	No	0	0	0.296	0.281
33-bus	120	P, Q, V, θ	0.2	Yes	0	0	0.475	0.521
33-bus mesh	120	P, Q, V, θ	0.2	Yes	0	0	0.483	0.539
33-bus	120	P, Q, V, θ	0.2	No	3.125	3.125	0.701	0.746
33-bus	120	P, Q, V, θ	0.5	Yes	0	0	1.188	1.370
33-bus	120	P, Q, V	0.1	Yes	0	0	28.711	29.805
141-bus	200	P, Q, V, θ	0.1	Yes	0	0	2.346	2.823
141-bus	200	P, Q, V	0.01	Yes	0	0	21.747	33.625

branch number that the method can obtain when using the CRLB-based criterion shown in Algorithm 3 of the Appendix A.3 to identify the topology and thus provides a benchmark of how well the topology can be estimated. We use the relative geometric mean as the error of admittance estimation:

$$E^p = \left(\prod_{n=1}^{N} e_n\right)^{\frac{1}{N}} \tag{18}$$

where N is the number of variables to be estimated and e_n denotes the relative error of the nth variable:

$$e_n = \frac{|\hat{y}_n - y_n|}{y_n} \times 100\% \tag{19}$$

where $\hat{y}_n$ denotes the th estimated value and y_n denotes its real value. The admittance precision limit is calculated using the CRLB. All of the errors and the precision limit are transformed to a scale of 100%.

As can be concluded from Table 1, the proposed CPS method can approach the precision limits under various settings. Nevertheless, there are some fundamental limits from the results: (1) the estimation error will greatly increase (more than one degree of magnitude) without phasor measurements; (2) prior topology knowledge is good compensation for poor measurements; (3) the estimation of the 141-bus system with larger admittances and more buses is far more difficult than the estimation of the 33-bus system and requires more accurate measurements.

3.4 Comparison of different methods

We compare the proposed method with the methods proposed by Ref. [7] (denoted as M1) and Ref. [8] (denoted as M2). It should be noted that many methods are currently based on the measurement of P, Q, V, and θ. However, assuming that each bus has PMU measurements is not very realistic in a distribution network. We compare one method that uses only P, Q, and V measurements (M1) and one that uses P, Q, V, and measurements (M2). The results are shown in Table 2.

As shown in Table 2, the proposed method (M3) can approach the theoretical precision limit for different systems, different measuring equipment, different noise levels and different topological prior knowledge. By comparison, the errors of topology and admittance identification of M1 and M2 are larger. M1 does not have special treatment or modeling for the noise of voltage, which makes it difficult for M1 to deal with the situation of large measurement noise of voltage amplitude. Therefore, the parameter identification accuracy of method M1 can approach the theoretical parameter identification accuracy limit only when the noise is very small (0.01%). Although M2 uses the total least squares regression method to handle the noise of measurement equipment, it is difficult for the identification results to approach the theoretical accuracy limit. In addition, method M2 must rely on voltage phase angle measurements. The comparison illustrates that the proposed method outperforms the methods in Ref. [7] (denoted as M1) and Ref. [8] in terms of accuracy and adaptability of different measurements.

4 Conclusions

Our research advances the field by rigorously quantifying the theoretical precision boundaries inherent in the distribution grid TAE problem. We systematically evaluate these precision limits across diverse measurement scenarios encompassing varying measurement devices, noise levels, measurement quantities, and prior knowledge availability. Notably, we unveil the profound intricacies associated with the TAE problem, characterized by nonconvexity and ill-conditioning, distinguishing it from the SE problem.

In the context of TAE, conventional optimization techniques such as Adam or Newton's methods falter, even when confronted with a relatively straightforward 3-bus system. In response, we introduce a novel CPS algorithm adept at dynamically fusing the stability of first-order optimization and the rapid convergence of second-order optimization. We further facilitate large-scale second-order optimization by decoupling calculations across different snapshots.

Through comprehensive case studies conducted on IEEE 33- and 141-bus systems, we empirically validate the efficacy of our proposed CPS method in approaching the theoretically established precision limits across diverse operational scenarios. Our work also sheds light on critical insights, notably the substantial impact of the absence of phasor measurements and the compensatory role of prior topology knowledge.

In the future, our research agenda will extend toward exploring the intricate interplay among measurement devices, measurement noise levels, the number of snapshots, and the availability of prior knowledge, thereby advancing our understanding of the underlying theoretical relationships within this domain.

Appendix

A.1 Proof of Theorem 1

The negative log-likelihood distribution of problem (5) is the product of the following Gaussian distributions:

$$p(\boldsymbol{z},\boldsymbol{x}) = -\ln\prod_{m=1}^{M}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{(z_m-h_m(\boldsymbol{x}))^2}{2\sigma_m^2}\right] \tag{20a}$$

$$=\sum_{m=1}^{M}\left[\frac{(z_m-h_m(\boldsymbol{x}))^2}{2\sigma_m^2}+\frac{\ln(2\pi\sigma_m^2)}{2}\right]. \tag{20b}$$

The Fisher information matrix is defined as the expectation of the Hessian matrix of $p(\boldsymbol{z},\boldsymbol{x})$:

$$\boldsymbol{F}=\mathbb{E}_x\left[\frac{\partial^2}{\partial\boldsymbol{x}^{\mathrm{T}}\partial\boldsymbol{x}}p(\boldsymbol{z},\boldsymbol{x})\right] \tag{21a}$$

$$=\mathbb{E}_x\left[\frac{\partial}{\partial\boldsymbol{x}^{\mathrm{T}}}\sum_{m=1}^{M}\left(-\frac{z_m-h_m(\boldsymbol{x})}{\sigma_m^2}\frac{\partial h_m(\boldsymbol{x})}{\partial\boldsymbol{x}}\right)\right] \tag{21b}$$

$$=\mathbb{E}_x\left[\sum_{m=1}^{M}\frac{1}{\sigma_m^2}\frac{\partial^2 h_m(\boldsymbol{x})}{\partial\boldsymbol{x}^{\mathrm{T}}\partial\boldsymbol{x}}[h_m(\boldsymbol{x})-z_m]+\frac{1}{\sigma_m^2}\frac{\partial h^{\mathrm{T}}(\boldsymbol{x})}{\partial\boldsymbol{x}}\frac{\partial h(\boldsymbol{x})}{\partial\boldsymbol{x}^{\mathrm{T}}}\right]. \tag{21c}$$

Since the Gaussian noise distribution is symmetric to zero, the expectation of $z_m-h_m(\boldsymbol{x})$ over the whole distribution is zero. Therefore, we have

$$\boldsymbol{F}=\frac{\partial h^{\mathrm{T}}(\boldsymbol{x})}{\partial\boldsymbol{x}}\times\frac{1}{\boldsymbol{\sigma}_m^2}\times\frac{\partial h(\boldsymbol{x})}{\partial\boldsymbol{x}^{\mathrm{T}}}. \tag{22}$$

According to the CRLB[21], the covariance matrix is lower bounded by the inverse of the Fisher information matrix.

A.2 Method to obtain the initial value

We extend the method in Ref. [7] to obtain the initial value, assuming there are power injection measurements (P_i^t, Q_i^t) and voltage magnitude measurements V_i^t. The reason for this setting is that phasor measurements θ_i^t are usually not available in distribution grids. We also assume that the topology is radial at the initial stage, which can be relaxed in the following iteration stages. The initial value of $\boldsymbol{x}$ begins by setting the voltage magnitude measurement as the initial value of $[\{\hat{V}_i^t\}]$. Then, we estimate an initial value of $[\{g_{ij}\}\{b_{ij}\}]$ from the P_i^t, Q_i^t, and V_i^t measurements. Finally, we obtain the initial value of $[\{\hat{\theta}_i^t\}]$ by the DC power flow calculation.

The main difficulty is how to estimate $[\{g_{ij}\}\{b_{ij}\}]$ without voltage angle measurements.

We start by estimating a radial topology using the voltage mag-

Table 2 Comparison of the performances of different methods

System	Snapshots	Sensors	Prior topology knowledge	Topology precisionlimits (%)	Topology estimation error (%)	Admittance precisionlimits (%)	Admittance estimation error (%)	Noise(%)
33-bus	120	P, Q, V, θ	0.1	No	0.264	87.64	0.514	0.236
33-bus	120	P, Q, V, θ	0.2	Yes	0.475	96.37	3.382	0.521
33-bus	120	P, Q, V, θ	0.2	No	0.701	103.2	18.96	0.746
33-bus	120	P, Q, V, θ	0.5	Yes	1.188	63.73	20.32	1.370
33-bus	120	P, Q, V	0.1	Yes	28.71	87.64	â€”	29.81
141-bus	200	P, Q, V, θ	0.1	Yes	2.346	93.41	48.83	2.823
141-bus	200	P, Q, V	0.01	Yes	21.75	37.87	â€”	33.63

nitudes by a tree structure topology construction. We use the average value of the voltage magnitudes to obtain the root-leaf relationship of the distribution grids. We assume that the voltage magnitudes decrease from the root to the leaf buses. We only use the measurements at night to avoid the possible bidirectional power flow caused by distributed PV generation. We then calculate the moving average of the voltage magnitudes to decrease the influence of measurement noise:

$$\hat{V}_i^t = \frac{1}{l}\sum_{t=1}^{l} V_i^t. \tag{23}$$

Finally, we use the correlation coefficient to construct the topology. We assume that the voltage magnitudes of the connected buses are highly correlated[3, 10]. See Algorithm 4 of the Appendix A.3 for details. Note that the proposed method may not obtain an accurate result because the assumptions may not be satisfied in practice. However, Algorithm 4 only provides an initial value, and the inaccuracy can be corrected in the following iterations. Furthermore, some recent works also provide some topology estimation methods[3, 10, 13] that can be used to obtain the initial topology.

We then estimate line admittance by introducing an approximate power flow formulation without voltage angles.

Theorem 4 (Phasor-free power flow). The power flow equations can be approximated as (24) under the assumption that $\alpha_{ij} = P_{ij}/Q_{ij}$ is a constant and $\sin\theta_{ij} \approx \theta_{ij}$.

$$P_i/V_i = (V_i - V_j)/z_{ij}^{\#}, \tag{24a}$$

$$Q_i/V_i = (V_i - V_j)/(z_{ij}^{\#}\alpha ij), \tag{24b}$$

with the augmented impedance $z_{ij}^{\#}$ defined as

$$z_{ij}^{\#} = \frac{g_{ij} - b_{ij}/\alpha_{ij}}{g_{ij}^2 + b_{ij}^2}. \tag{25}$$

Proof. The branch flow equations are formulated as

$$\begin{aligned} P_i/V_i &= (V_i - V_j\cos\theta_{ij})g_{ij} - V_j\sin\theta_{ij}b_{ij} \\ &\approx (V_i - V_j\cos\theta_{ij})g_{ij} - V_j\theta_{ij}b_{ij}, \end{aligned} \tag{26a}$$

$$\begin{aligned} Q_i/V_i &= -(V_i - V_j\cos\theta_{ij})b_{ij} - V_j\sin\theta_{ij}g_{ij} \\ &\approx -(V_i - V_j\cos\theta_{ij})b_{ij} - V_j\theta_{ij}g_{ij}, \end{aligned} \tag{26b}$$

where we use the assumption $\sin\theta_{ij} \approx \theta_{ij}$. $(26a) \times b_{ij} + (26b) \times g_{ij}$ is

$$\frac{P_{ij}b_{ij}}{V_i} + \frac{Q_{ij}g_{ij}}{V_i} = -(b_{ij}^2 + g_{ij}^2)\theta_{ij}, \tag{27a}$$

$$\theta_{ij} = -\frac{1}{V_i}\left(\frac{b_{ij}P_{ij}}{b_{ij}^2 + g_{ij}^2} + \frac{g_{ij}Q_{ij}}{b_{ij}^2 + g_{ij}^2}\right). \tag{27b}$$

Substitute (27b) into (26a) and substitute (27b) into (26b):

$$\frac{P_{ij}}{V_i} = \frac{g_{ij}^2 + b_{ij}^2}{g_{ij} - \alpha_{ij}^{-1}b_{ij}}(V_i - V_j), \tag{28a}$$

$$\frac{Q_{ij}}{V_i} = -\frac{g_{ij}^2 + b_{ij}^2}{b_{ij} - \alpha_{ij}g_{ij}}(V_i - V_j), \tag{28b}$$

where $\alpha_{ij} = P_{ij}/Q_{ij}$.

Afterwards, we can estimate the values of $1/z_{ij}^{\#}$ and $1/(z_{\lfloor}^{\#}ij]\alpha_{ij})$ by least squares estimation and set them as the initial values of g_{ij} and b_{ij}, respectively.

$$g_{ij} = (P_i/V_i)^{\mathrm{T}}(V_i - V_j)/[(V_i - V_j)(V_i - V_j)^{\mathrm{T}})], \tag{29a}$$

$$b_{ij} = (Q_i/V_i)^{\mathrm{T}}(V_i - V_j)/[(V_i - V_j)(V_i - V_j)^{\mathrm{T}})]. \tag{29b}$$

A.3 Pseudo codes

Algorithm 1 The first-order optimization

Inputs: The last state vector $\boldsymbol{x}^{k-1}$ and the moment of the last iteration $\boldsymbol{m}^{k-1}$.

1: if the iteration step $k = 1$ then
2: Calculate the empirical CRLB using the initial value $\boldsymbol{x}^0$ and get $\boldsymbol{\sigma}_{\boldsymbol{x}}^{cr}$. Initialize $\boldsymbol{m}^0 = \boldsymbol{x}^0$. Get the mean value of $\boldsymbol{\sigma}_{\boldsymbol{x}}^{cr}$ among admittances, voltage magnitudes, voltage angles, and get $[w_a^{cr}, w_V^{cr}, w_\theta^{cr}]$.
3: Set the moment ratio α.
4: end if
5: Calculate the gradient by $\boldsymbol{g}^k = \nabla_{\boldsymbol{x}_{k-1}} Loss$. Get the corresponding mean absolute value $[w_a^g, w_V^g, w_\theta^g]$.
6: Normalize and rescale the gradient by $\boldsymbol{m}^k \leftarrow [\boldsymbol{g}_{\boldsymbol{a}}^k/w_a^g \times w_a^{cr}, \boldsymbol{g}_V^k/w_V^g \times w_V^{cr}, \boldsymbol{g}_\theta^k/w_\theta^g \times w_\theta^{cr}]$.
7: Update the moment $\boldsymbol{m}^k = \alpha\boldsymbol{m}^{k-1} + (\alpha - 1)\boldsymbol{g}^k$.

Outputs: The moment $\boldsymbol{m}^k$ and the gradient $\boldsymbol{g}^k$.

Algorithm 2 The hybrid line search

Inputs: The last state vector $\boldsymbol{x}^{k-1}$.

1: if the iteration step $k = 1$ then
2: Set the start ratio r_0, the maximum ratio r_{max}, the incremental ratio β, and the stop threshold η.
3: end if
4: Do Algorithm 1.
5: Get $\boldsymbol{m}^k$ and $\boldsymbol{g}^k$.
6: $r \leftarrow r_0$.
7: while $r \le r_{max}$ do
8: $\boldsymbol{x}^k \leftarrow \boldsymbol{x}^{k-1} + \boldsymbol{m}^k/\beta^r$.
9: $r \leftarrow r + 1$.
10: if $Loss(\boldsymbol{x}^{k-1}) - Loss(\boldsymbol{x}^k) \leqslant \eta(\boldsymbol{g}^k)^{\mathrm{T}}\boldsymbol{m}^k/\beta^r$ then
11: $\boldsymbol{m}^k \leftarrow \boldsymbol{m}^k/\beta^r$.
12: Break.
13: end if
14: end while
15: Calculate $\boldsymbol{d}^k$ from (16).
16: $r \leftarrow r_0$.
17: while $r \leqslant r_{max}$ do
18: $w_1 = 1/(1 + \beta^r)$. $w_2 = \beta^r/(1 + \beta^r)$.
19: $\boldsymbol{x}^k \leftarrow \boldsymbol{x}^{k-1} + w_1\boldsymbol{m}^k + w_2\boldsymbol{d}^k$.
20: if $Loss(\boldsymbol{x}^{k-1}) - Loss(\boldsymbol{x}^k) \leqslant \eta(\boldsymbol{g}^k)^{\mathrm{T}}(w_1\boldsymbol{m}^k + w_2\boldsymbol{d}^k)$ then
21: $\boldsymbol{x}^{k-1} \leftarrow \boldsymbol{x}^k$.
22: $\boldsymbol{m}^k \leftarrow \boldsymbol{m}^k/\beta^r$.
23: Break.
24: end if
25: $r \leftarrow r + 1$.
26: end while

Outputs: The vector of this state $\boldsymbol{x}^k$.

Algorithm 3: Topology updating algorithm based on the CRLB

Inputs Initial value of state vector $\boldsymbol{x}^0$. Current possible branch set S_T.

sets the maximum number of iteration steps to k.
sets the current number of steps $k = 0$.
sets the iteration convergence coefficient γ_{iter}.
sets the admittance threshold γ_{topo}.
sets the CRLB threshold γ_{cr}.
while $k \leqslant K$ do
  executes the algorithm 2.
  if $\max(|\boldsymbol{x}^{\boldsymbol{k}} - \boldsymbol{x}^{\boldsymbol{k-1}}|) \leqslant \gamma_{\text{iter}}$ then
    Sets the variable isChange = false denoting whether there is vtopology change.
    calculates the empirical CRLB $\boldsymbol{\sigma}^{CR}$, according to the currently identified parameter values. See section section â…¡ for the specific method.
    if there exists $b_{ij}/\max(b_{ii}, b_{jj}) \leqslant \gamma_{\text{topo}}$, and $\sigma^{cr}_{b_{ij}}/max(b_{ii}, b_{jj}) \leqslant \gamma_{\text{cr}}$ then
      set $b_{ij} = 0$, $g_{ij} = 0$.
      moves all (i,j) that meet certain conditions out of the set S_T.
      sets isChange = true.
    end if
    if isChange==FALSE then
      Break.
    end if
  end if
  $k \leftarrow k+1$.
end while

Outputs: Iterated state vector value $\boldsymbol{x}^{\boldsymbol{k}}$. The current topology collection S_T.

Algorithm 4 Tree structure topology construction

Inputs: The voltage magnitude measurements of different buses $[V_1, ..., V_i, ..., V_N]$.

Initialize the topology set $S_T = \emptyset$. Sort the average voltage magnitudes in a descending order $[V_{d1}, ..., V_{di}, ..., V_{dN}]$.
Get the l moving average of the voltage magnitudes $[\hat{V}_{d1}, ..., \hat{V}_{di}, ..., \hat{V}_{dN}]$.
for $i = 2 \sim N$ do
  Get the bus j with maximum correlation coefficient $\arg\max_{j \in 1 \sim i-1} Corr(\hat{V}_{di}, \hat{V}_{dj})$.
  $S_T \leftarrow (di, dj)$.
end for

Outputs: The topology set S_T.

Acknowledgements

This work is funded by the Science and Technology Project of State Grid Corporation of China (5100-202199519A-0-5-ZN).

Article history

Received: 10 July 2023; Revised: 23 September 2023; Accepted: 29 October 2023

Additional information

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] Muruganantham, B., Gnanadass, R., Padhy, N. P. (2017). Challenges with renewable energy sources and storage in practical distribution systems. *Renewable and Sustainable Energy Reviews*, 73: 125–134.
[2] Bolognani, S., Bof, N., Michelotti, D., Muraro, R., Schenato, L. (2013). Identification of power distribution network topology via voltage correlation analysis. In: Proceedings of the 52nd IEEE Conference on Decision and Control, Firenze, Italy.
[3] Weng, Y., Liao, Y., Rajagopal, R. (2017). Distributed energy resources topology identification via graphical modeling. *IEEE Transactions on Power Systems*, 32: 2682–2694.
[4] Deka, D., Backhaus, S., Chertkov, M. (2016). Estimating distribution grid topologies: A graphical learning based approach. In: Proceedings of the 2016 Power Systems Computation Conference (PSCC), Genoa, Italy.
[5] Yu, J., Weng, Y., Rajagopal, R. (2018). PaToPa: A data-driven parameter and topology joint estimation framework in distribution grids. *IEEE Transactions on Power Systems*, 33: 4335–4347.
[6] Yu, J., Weng, Y., Rajagopal, R. (2019). PaToPaEM: A data-driven parameter and topology joint estimation framework for time-varying system in distribution grids. *IEEE Transactions on Power Systems*, 34: 1682–1692.
[7] Zhang, J., Wang, Y., Weng, Y., Zhang, N. (2020). Topology identification and line parameter estimation for non-PMU distribution network: A numerical method. *IEEE Transactions on Smart Grid*, 11: 4440–4453.
[8] Moffat, K., Bariya, M., Von Meier, A. (2020). Unsupervised impedance and topology estimation of distribution networksâ€"Limitations and tools. *IEEE Transactions on Smart Grid*, 11: 846–856.
[9] Li, T., Werner, L., Low, S. H. (2020). Learning graphs from linear measurements: Fundamental trade-offs and applications. *IEEE Transactions on Signal and Information Processing Over Networks*, 6: 163–178.
[10] Deka, D., Chertkov, M., Backhaus, S. (2020). Joint estimation of topology and injection statistics in distribution grids with missing nodes. *IEEE Transactions on Control of Network Systems*, 7: 1391–1403.
[11] Park, S., Deka, D., Backhaus, S., Chertkov, M. (2020). Learning with end-users in distribution grids: Topology and parameter estimation. *IEEE Transactions on Control of Network Systems*, 7: 1428–1440.
[12] Miao, X., Ilic, M., Wu, X., Munz, U. (2019). Distribution grid admittance estimation with limited non-synchronized measurements. In: Proceedings of the 2019 IEEE Power & Energy Society General Meeting (PESGM), Atlanta, GA, USA.
[13] Zhao, J., Li, L., Xu, Z., Wang, X., Wang, H., Shao, X. (2020). Full-scale distribution system topology identification using Markov random field. *IEEE Transactions on Smart Grid*, 11: 4714–4726.
[14] Xu, H., Dominguez-Garcia, A. D., Sauer, P. W. (2019). Data-driven coordination of distributed energy resources for active power provision. *IEEE Transactions on Power Systems*, 34: 3047–3058.
[15] Sandraz, J., Macwan, R., Diaz-Aguilo, M., McClelland, J., de Leon, F., Czarkowski, D., Comack, C. (2014). Energy and economic impacts of the application of CVR in heavily meshed secondary distribution networks. *IEEE Transactions on Power Delivery*, 29: 1692–1700.
[16] Bhela, S., Kekatos, V., Veeramachaneni, S. (2018). Enhancing observability in distribution grids using smart meter data. *IEEE Transactions on Smart Grid*, 9: 5953–5961.
[17] Guo, Y., Zhang, B., Wu, W., Guo, Q., Sun, H. (2013). Solvability and solutions for bus-type extended load flow. *International Journal*

of Electrical Power & Energy Systems, 51: 89–97.
[18] Bhela, S., Kekatos, V., Veeramachaneni, S. (2019). Smart inverter grid probing for learning loads: Part Iâ€"Identifiability analysis. *IEEE Transactions on Power Systems*, 34: 3527–3536.
[19] Kekatos, V., Vlahos, E., Ampeliotis, D., Giannakis, G. B., Berberidis, K. (2013). A decentralized approach to generalized power system state estimation. In: Proceedings of the 2013 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), St. Martin, France.
[20] Grotas, S., Yakoby, Y., Gera, I., Routtenberg, T. (2019). Power systems topology and state estimation by graph blind source separation. *IEEE Transactions on Signal Processing*, 67: 2036–2051.
[21] Kay, S. M. (1993). Fundamentals of statistical signal processing: Estimation theory. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.
[22] A. Abur and A. G. Exposito, Power system state estimation: theory and implementation. CRC press, 2004.
[23] Liserre, M., Blaabjerg, F., Teodorescu, R. (2007). Grid impedance estimation via excitation of LCL-filter resonance. *IEEE Transactions on Industry Applications*, 43: 1401–1407.
[24] Cavraro, G., Kekatos, V. (2019). Inverter probing for power distribution network topology processing. *IEEE Transactions on Control of Network Systems*, 6: 980–992.
[25] Miao, X., Wu, X., Munz, U., Ilic, M. (2019). Multi-layered grid admittance matrix estimation for electric power systems with partial measurements. In: Proceedings of the 2019 American Control Conference (ACC), Philadelphia, PA, USA.
[26] Wang, G., Zamzam, A. S., Giannakis, G. B., Sidiropoulos, N. D. (2018). Power system state estimation via feasible point pursuit: Algorithms and Cramér-Rao bound. *IEEE Transactions on Signal Processing*, 66: 1649–1658.
[27] Ghasemi Damavandi, M., Krishnamurthy, V., Marti, J. R. (2015). Robust meter placement for state estimation in active distribution systems. *IEEE Transactions on Smart Grid*, 6: 1972–1982.
[28] Xygkis, T. C., Korres, G. N., Manousakis, N. M. (2018). Fisher information-based meter placement in distribution grids via the D-optimal experimental design. *IEEE Transactions on Smart Grid*, 9: 1452–1461.
[29] Boyd, S., Vandenberghe, L. (2004). *Convex Optimization*. Cambridge, UK: Cambridge University Press.
[30] Kingma, D. P., Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv Print*, 1412.6980.
[31] Ben-Israel, A., Greville, T. N. (2003). *Generalized Inverses: Theory and applications*. New York, NY: Springer.
[32] Armijo, L. (1966). Minimization of functions having lipschitz continuous first partial derivatives. *Pacific Journal of mathematics*, 16: 1–3.
[33] CER (2012). *CER Smart Metering Project-Electricity Customer Behaviour Trial, 2009–2010*. Commission for Energy Regulation (CER), Dublin, Ireland. Available at https://www.ucd.ie/issda/data/commissionforenergyregulationcer/.
[34] Zimmerman, R. D., Murillo-Sanchez, C. E., Thomas, R. J. (2011). MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26: 12–19.
[35] Liu, Y., Zhang, N., Wang, Y., Yang, J., Kang, C. (2019). Data-driven power flow linearization: A regression approach. *IEEE Transactions on Smart Grid*, 10: 2569–2580.
[36] Baran, M. E., Wu, F. F. (1989). Network reconfiguration in distribution systems for loss reduction and load balancing. *IEEE Power Engineering Review*, 9: 101–102.
[37] Khodr, H. M., Olsina, F. G., De Oliveira-De Jesus, P. M., Yusta, J. M. (2008). Maximum savings approach for location and sizing of capacitors in distribution systems. *Electric Power Systems Research*, 78: 1192–1203.