A Complex Domain Gaussian Belief Propagation Method for Fully Distributed State Estimation

Kang Sun, *Student Member, IEEE,* Zhinong Wei, *Member, IEEE,* Venkata Dinavahi, *Fellow, IEEE,* Manyun Huang, *Member, IEEE,* and Guoqiang Sun, *Member, IEEE*

***Abstract*—To alleviate the communication, storage, and computation burden on the control center and make full use of edge computing resources, fully distributed state estimation has received increasing interest recently. This paper intends to improve the efficiency and robustness of the fully distributed state estimation by introducing a meter-level method based on the Gaussian belief propagation theory. Specifically, we propose a complex domain factor graph, which extends the state variable vector from voltage phasors to multiple electrical quantities, including voltage phasors, current phasors, voltage magnitudes, and active/reactive power, enabling the direct processing of nonlinear measurement models and significantly reducing the number of iterations. Furthermore, based on the M-estimation theory, we innovatively incorporate multiple robust functions to the Gaussian belief propagation method to enhance the robustness of the proposed fully distributed estimator. The effectiveness of the proposed method is demonstrated under various operation conditions.**

***Index Terms*—Complex domain, fully distributed state estimation, factor graph, Gaussian belief propagation, M-estimation, power system analities.**

I. Introduction

With the deregulation of the power industry and the increasing penetration of distributed renewable energy generation, the interaction between regional power grids, as well as between transmission and distribution systems, has grown significantly. An important requirement arising from this transformation is the system-wide power system state estimation (SE) that is capable of unified analysis across the entire power system [1]. In this context, the conventional centralized SE method would lead to unprecedented communication, storage, and computation burdens on the control center, increasing the need for a distributed SE.

Distributed SE has received notable interest recently and can be categorized into two main groups based on the size of the minimum computation units: multi-area SE (MASE) and fully distributed SE (FDSE). The MASE method divides the system into several subsystems, each with an individual centralized estimator, while FDSE is capable of being implemented at either bus-level or meter-level, thereby eliminating the need for the centralized estimator. Their main goal is to achieve the same estimation accuracy as the centralized SE method while reducing the communication, storage, and computation burden on each control center.

The MASE method can be further divided into two groups: hierarchical and decentralized approaches. In the early literature on MASE, most research focused on the hierarchical approach, which requires a fusion center to coordinate the local estimates of all subsystems [2]–[5]. However, these methods would result in excessive communication and storage burden in the fusion center. Recent research on MASE has focused more on decentralized approaches, which eliminate the need for a fusion center through data exchange between subsystems. In this regard, the method presented in [6] combined the alternative direction method of multipliers (ADMM) and the semidefinite relaxation technique to resolve the non-convex SE problem. Similarly, the authors of [7] exploit proximal ADMM and matrix completion to address the low-observability problems in MASE. Further, [8] presented a model-free method based on distributed tensor completion to recover the state information of the whole system. In [9], a non-overlapping approach was presented considering the switching of inter-area communication graphs. However, whether it is hierarchical or decentralized MASE, the numerical stability, estimation accuracy, and computational efficiency heavily depend on the partitioning strategy [10], limiting the flexibility of algorithm application. Furthermore, with the ongoing expansion of the system scale, regular re-partitioning to ensure the performance of SE will be inevitable, which is impractical for field implementation.

The FDSE method takes buses or meters as the minimum computation units, thereby eliminating the impact of system scale on state estimators. Additionally, with the continuous integration of intelligent electronic devices (IEDs) into power systems [11], FDSE can even be implemented directly into these IEDs, completely eliminating the dependence on the control center. [11] proposed an agent-based FDSE method, which estimates the local state of the system through bus-level local computations and asynchronous message exchange between neighboring buses. The authors of [12] introduced a meter-level method based on the weighted least square (WLS) method and graph theory. Moreover, many recent FDSE approaches developed their estimators using belief propagation (BP), which is a highly efficient probabilistic inference algorithm and is free of ill-conditioning problems since no matrix operations are required [13], [14]. In this regard, the authors of [15] applied the BP to the power system

This work was supported in part by the National Natural Science Foundation of China under Grant U1966205 and in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

K. Sun, Z. Wei, M. Huang, and G. Sun are with the School of Electrical and Power Engineering, Hohai University, Nanjing 210098, China (e-mail: sunkang.real@outlook.com; wzn_nj@263.net; hmy_hhu@yeah.net; hhusunguoqiang@163.com).

V. Dinavahi is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: dinavahi@ualberta.ca)

SE problem for the first time. [16] proposed an alternating Gaussian belief propagation method for linear SE problems to reduce the iterations by dividing the factor graph into clusters. In [17], a Gauss-Newton BP (GN-BP) method was introduced to solve nonlinear SE problems by applying BP sequentially over a sequence of linear approximations of the SE model. However, these approaches face three common challenges: First, they generally require excessive iterations to converge for nonlinear SE problems, making it difficult to meet the requirements of accuracy and efficiency simultaneously. Second, while there is no centralized estimator, a global view (i.e., the capability to obtain information about the entire power system) is still needed to decide when to initiate/stop calculation and determine the order of message propagation. Finally, the robustness of these methods is insufficient, as they generally do not handle bad data (BD) [12], [15], [16], or rely on the largest normalized residual test (LNRT)-like procedures [11], [17], which may have low levels of accuracy and efficiency in BD processing, especially when dealing with a large number of BD [18]. Hence, further research on FDSE is necessary to achieve truly practical distributed computation.

To address the aforementioned challenges, this paper proposes a complex domain FDSE method based on Gaussian BP (GaBP) theory. The SE model is established by constructing a complex domain factor graph, where variables, measurements, and their relationships are represented as variable/factor nodes. Subsequently, each node exchanges messages with its neighboring nodes to get the final estimates. The main contributions and advantages are described below.

- We propose a complex domain factor graph and corresponding GaBP message passing rules for FDSE, allowing the direct processing of nonlinear measurement models without a global view. Compared to the conventional FDSE methods, the number of iterations is significantly reduced.
- We construct robust data factor nodes with multiple cost functions based on the M-estimation theory, achieving fully distributed bad data processing. Comprehensive comparative studies demonstrate the strong robustness of the proposed method.

The remainder of this paper is structured as follows. Section II presents an overview of GaBP. Section III introduces the proposed method in detail. Simulation results are described in Section IV. Conclusions are drawn in Section V.

II. Fundamentals

GaBP is a probabilistic graphical model-based inference algorithm, which is capable of solving an overdetermined linear system in a fully distributed way [13]. Consider the measurement functions of a linear system in the complex domain:

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathbf{z} = [z_1, z_2, \cdots, z_m]^T$ and $\mathbf{x} = [x_1, x_2, \cdots, x_n]^T$ denote the measurement vector and the state vector, respectively; $\mathbf{h}(\cdot)$ is the linear function mapping states to measurements through measurement matrix $\mathbf{H}$; $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m]^T$ denotes the measurement noise vector. The variance of a complex random variable X is generally defined as $\sigma_X^2 \triangleq \mathbb{E}[(X-\mu)(X-\mu)^*]$ [19], where μ denotes the mean value, $\mathbb{E}[\cdot]$ and $(\cdot)^*$ are the expected value and complex conjugate of the argument, respectively. With this definition, for a complex measurement z_c, if its real and imaginary parts are independent, the measurement variance can be obtained as $\sigma_c^2 = \sigma_r^2 + \sigma_i^2$ [19, Sec. 5.2], where σ_r and σ_i are the standard deviations of the real and imaginary parts of the complex measurement, respectively. Similarly, if its magnitude and phase angle are independent, the measurement variance can be obtained using its expected value conditioned on the measured value as $\mathbb{E}[\sigma_c^2|z_c] = \sigma_m^2(2 - e^{-\sigma_\theta^2}) + r_m^2(1 - e^{-\sigma_\theta^2})$ [20, Thm. 3], where r_m denotes the magnitude of the complex measurement, σ_m and σ_θ are the standard deviations of the magnitude measurement and the phase angle measurement, respectively. When GaBP is performed on the factor graph, a type of probabilistic graphical model that represents the relationship between variables [13], the state variables could be constructed as variable nodes, while the measurements and the corresponding measurement functions could be constructed as factor nodes. Each factor node is connected with the involved variable nodes according to the measurement function through edges. By passing messages, which contain probability distribution information such as mean and variance, between interconnected nodes, the GaBP algorithm could estimate the global system state and achieve the same accuracy as the centralized WLS estimator [13], [17].

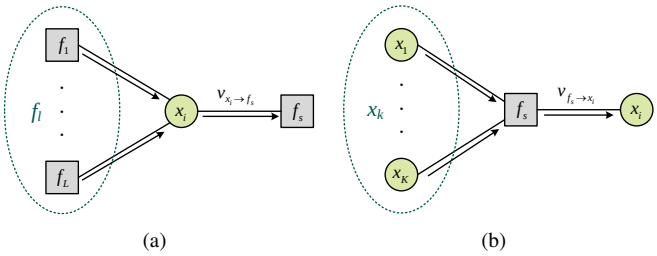


Fig. 1. Messages in the GaBP algorithm: (a) message from variable node to factor node; (b) message from factor node to variable node.

1) Messages from variable nodes to factor nodes: Consider a local factor graph that centers on the variable node x_i in Fig. 1(a). Let $v_{x_i \to f_s}$ denotes the message from variable node x_i to factor node f_s. When this message follows a Gaussian distribution, it could be represented by its mean $\mu_{x_i \to f_s}$ and variance $\sigma^2_{x_i \to f_s}$. According to the sum-product rule [21], the message from the variable node x_i to the factor node f_s is formulated as the product of incoming messages from all connected factor nodes except f_s:

$$\mu_{x_i \to f_s} = \Big(\sum_{f_l \in \mathcal{N}\{x_i\} \backslash f_s} \frac{\mu_{f_l \to x_i}}{\sigma^2_{f_l \to x_i}} \Big) \sigma^2_{x_i \to f_s}, \tag{2}$$

$$1/\sigma^2_{x_i \to f_s} = \sum_{f_l \in \mathcal{N}\{x_i\} \backslash f_s} 1/\sigma^2_{f_l \to x_i}, \tag{3}$$

where $\mu_{f_l \to x_i}$ and $\sigma^2_{f_l \to x_i}$ are the mean and variance of $v_{f_l \to x_i}$ (i.e., the message from factor node f_l to variable node x_i), respectively; $\mathcal{N}\{x_i\} \backslash f_s$ denotes the set of factor nodes that are neighbors of the variable node x_i except f_s.

2) Messages from factor nodes to variable nodes: Consider a local factor graph in Fig. 1(b), the factor node f_s represents a function of x_i as well as K other neighboring variable nodes $x_k \in \{x_1, \cdots, x_K\}$. According to (1), the mean and variance of the message $v_{f_l \to x_i}$ could be obtained as

$$\mu_{f_s \to x_i} = \frac{1}{H_{x_i,f_s}} \Big(\mu_s - \sum_{x_k \in \mathcal{N}\{f_s\} \backslash x_i} H_{x_k,f_s} \mu_{x_k \to f_s} \Big), \quad (4)$$

$$\sigma^2_{f_s \to x_i} = \frac{1}{|H_{x_i,f_s}|^2} \Big(\sigma_s^2 + \sum_{x_k \in \mathcal{N}\{f_s\} \backslash x_i} |H_{x_k,f_s}|^2 \, \sigma^2_{x_k \to f_s} \Big), \quad (5)$$

where μ_s and σ_s^2 denote the measured value and the corresponding variance of measurement noise, respectively; H_{x_i,f_s} and H_{x_k,f_s} are corresponding elements of the measurement matrix $\mathbf{H}$; $\mathcal{N}\{f_s\}\backslash x_i$ is the set of variable nodes neighboring the factor node f_s, excluding x_i.

When there are cycles in the factor graph, the GaBP algorithm has to be implemented in an iterative way, which requires a message-passing schedule. Generally, the messages can be passed in a synchronous or asynchronous manner [13]:

- Synchronous scheduling: All variable and factor nodes update outgoing messages at each iteration using the incoming messages of the previous iteration.
- Asynchronous scheduling: Messages could be updated in an event-triggered way, where each variable/factor node broadcasts the outgoing messages immediately as the incoming messages are updated.

We could find that synchronous scheduling requires a global view of the iteration process, while asynchronous scheduling only focuses on individual nodes and hence could better support the distributed calculation. For single-connected factor graphs, GaBP is guaranteed to converge, while for factor graphs with cycles, the convergence of GaBP could be determined by the spectral radius of the matrix that reflects mean updates [14], [17], [22]. Moreover, the convergence of GaBP could be improved by message damping, which is achieved by replacing the current message with a combination of the message at the current and previous iterations [22]. Readers may refer to [22] for the detailed description of the convergence analysis for GaBP. After convergence, the belief of each variable node, which represents the estimated values of the mean and variance of each state variable, is obtained by collecting incoming messages from all connected factor nodes as follows.

$$\mu_{x_i} = \Big(\sum_{f_l \in \mathcal{N}\{x_i\}} \frac{\mu_{f_l \to x_i}}{\sigma^2_{f_l \to x_i}} \Big) \sigma^2_{x_i}, \quad (6)$$

$$1/\sigma^2_{x_i} = \sum_{f_l \in \mathcal{N}\{x_i\}} 1/\sigma^2_{f_l \to x_i}, \quad (7)$$

where $\mathcal{N}\{x_i\}$ is the set of factor nodes that are neighbors of the variable node x_i. Note that the GaBP algorithm could only deal with linear estimation problems. However, for legacy measurements (i.e., voltage magnitudes, active/reactive power flows, active/reactive power injections, etc.), the relationships between measurements and state variables are nonlinear. To address this problem, the state-of-the-art method sequential linearizes the nonlinear state estimation model by the Gauss-Newton (GN) method and applies GaBP to each linear subproblem [17]. Then, the outer iteration loop of GN and the inner iteration loop of GaBP will exist concurrently, resulting in a significant increase in the number of iterations. Moreover, in order to ensure that the previous outer iteration has been completed before the next outer iteration begins, it is necessary to globally control the order of message passing, which requires a global view, and hence, makes it difficult for such algorithms to achieve fully distributed execution. We will introduce a complex domain GaBP (CD-GaBP) method to address these issues in the next section.

III. Proposed State Estimation Method

In this section, a complex domain factor graph for power system state estimation is established, which is capable of representing the nonlinear measurement functions of legacy measurements. Furthermore, the bad data processing (BDP) algorithm and distributed iteration strategy for the proposed method are presented in detail.

A. CD-GaBP

In the proposed CD-GaBP method, the factor nodes are divided into data factor nodes and smoothness factor nodes, which correspond to measurement data and measurement functions, respectively. It is worth noting that this modification is not a necessary condition for constructing the proposed estimator. In fact, all types of probabilistic graphical models, including these two types of factor graphs, can be mutually converted [13].

1) CD-GaBP with phasor measurements: Under the condition that only phasor measurements are utilized, the voltage phasors and current injections of all buses, as well as the current flows of all branches are constructed as variable nodes, each of which has a corresponding data factor node. Given that typically not all data factor nodes have measurement data, and there are no current injections in practical phasor measurements [23], data factor nodes without measurements will be assigned pseudo means (e.g., a mean of 0 for current flow/injection and 1 for voltage phasor) and near-infinite variances (e.g., 10^{50}) to mitigate their potential adverse impact on estimation accuracy. Note that the current injections are modeled for establishing the relationships between legacy and phasor measurements, which will be detailed in Section III-A2. In those cases where a bus lacks power injection and current injection magnitude measurements, the corresponding part of complex current injections can be removed from the factor graph. The smoothness factor nodes represent the functions of the connected variable nodes as

$$\dot{\mathbf{I}}_{fl} = \mathbf{h}_{(\dot{\mathbf{I}}_{fl})}(\dot{\mathbf{V}}) = \mathbf{H}_{(\dot{\mathbf{I}}_{fl})} \dot{\mathbf{V}}, \quad (8)$$

$$\dot{\mathbf{I}}_{inj} = \mathbf{h}_{(\dot{\mathbf{I}}_{inj})}(\dot{\mathbf{I}}_{fl}) = \mathbf{H}_{(\dot{\mathbf{I}}_{inj})} \dot{\mathbf{I}}_{fl}, \quad (9)$$

where $\dot{\mathbf{V}}$, $\dot{\mathbf{I}}_{fl}$, and $\dot{\mathbf{I}}_{inj}$ denote the vectors of voltage phasor, current flow, and current injection, respectively; $\mathbf{h}_{(\dot{\mathbf{I}}_{fl})}(\cdot)$ and $\mathbf{h}_{(\dot{\mathbf{I}}_{inj})}(\cdot)$ are the linear measurement functions, whose measurement matrices are $\mathbf{H}_{(\dot{\mathbf{I}}_{fl})}$ and $\mathbf{H}_{(\dot{\mathbf{I}}_{inj})}$, respectively.

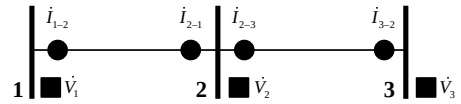


Fig. 2. 3-bus sample system with phasor measurements.

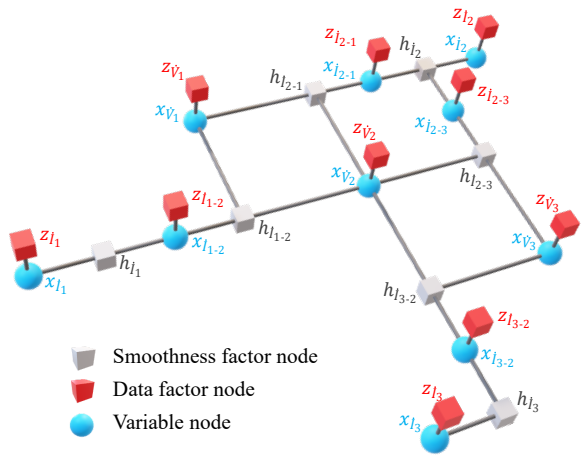


Fig. 3. Complex domain factor graph with phasor measurements for the 3-bus sample system.

For illustration consider a 3-bus sample system with phasor measurements in Fig. 2, whose complex domain factor graph is presented in Fig. 3. As depicted in Fig. 3, variable nodes are constructed for all voltage phasors, current injections, and current flows, which are denoted as $x_{\dot{V}}$, $x_{\dot{I}_{inj}}$, and $x_{\dot{I}_{fl}}$, respectively. Each variable node is paired with a corresponding data factor node, representing either measurement or pseudo measurement data, which are represented as $z_{\dot{V}}$, $z_{\dot{I}_{inj}}$, and $z_{\dot{I}_{fl}}$. Additionally, two types of smoothness factor nodes are constructed to model the relationships between the variable nodes, denoted herein as $h_{\dot{I}_{fl}}$ and $h_{\dot{I}_{inj}}$, corresponding to equations (8) and (9), respectively.

Based on the node type and the direction of message passing, messages are divided into four types: $v_{x\to z}$, $v_{x\to h}$, $v_{z\to x}$, and $v_{h\to x}$, which denote messages from a variable node to a data factor node, from a variable node to a smoothness factor node, from a data factor node to a variable node, and from a smoothness factor node to a variable node, respectively. The messages sent from the variable node, including $v_{x\to z}$ and $v_{x\to h}$, could be formulated according to (2) and (3), the message $v_{z\to x}$ is represented by the measured value and the corresponding variance. In the following, we will discuss the message $v_{h\to x}$, which could be further categorized into four types: $v_{h_{\dot{I}_{fl}}\to x_{\dot{V}}}$, $v_{h_{\dot{I}_{fl}}\to x_{\dot{I}_{fl}}}$, $v_{h_{\dot{I}_{inj}}\to x_{\dot{I}_{fl}}}$, and $v_{h_{\dot{I}_{inj}}\to x_{\dot{I}_{inj}}}$.

As shown in Fig. 4(a), the message $v_{h^{(i)}_{\dot{I}_{fl}}\to x^{(j)}_{\dot{V}}}$ from a current flow smoothness factor node $h^{(i)}_{\dot{I}_{fl}}$ to a voltage phasor variable node $x^{(j)}_{\dot{V}}$ could be obtained according to (8) as

$$\mu_{h^{(i)}_{\dot{I}_{fl}}\to x^{(j)}_{\dot{V}}} = \frac{1}{H_{x^{(j)}_{\dot{V}},h^{(i)}_{\dot{I}_{fl}}}}\Big(\mu_{x^{(l)}_{\dot{I}_{fl}}\to h^{(i)}_{\dot{I}_{fl}}} - \sum_{k\in\mathcal{X}_{\dot{V}}(i)\backslash j} H_{x^{(k)}_{\dot{V}},h^{(i)}_{\dot{I}_{fl}}}\mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{I}_{fl}}}\Big), \tag{10}$$

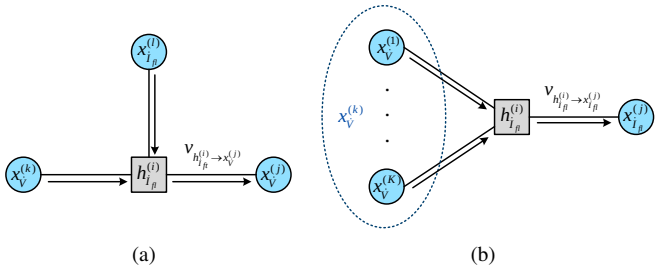


Fig. 4. Messages from current flow smoothness factor node (a) to voltage phasor variable node and (b) to current flow variable node.

$$\sigma^2_{h^{(i)}_{\dot{I}_{fl}}\to x^{(j)}_{\dot{V}}} = \frac{1}{\left|H_{x^{(j)}_{\dot{V}},h^{(i)}_{\dot{I}_{fl}}}\right|^2}\Big(\sigma^2_{x^{(l)}_{\dot{I}_{fl}}\to h^{(i)}_{\dot{I}_{fl}}} + \sum_{k\in\mathcal{X}_{\dot{V}}(i)\backslash j}\left|H_{x^{(k)}_{\dot{V}},h^{(i)}_{\dot{I}_{fl}}}\right|^2\sigma^2_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{I}_{fl}}}\Big), \tag{11}$$

where the superscripts $(\cdot)^{(i)}$, $(\cdot)^{(j)}$, $(\cdot)^{(k)}$, and $(\cdot)^{(l)}$ represent unique node numbers; $H_{x^{(k)}_{\dot{V}},h^{(i)}_{\dot{I}_{fl}}}$ is the corresponding element of $\mathbf{H}_{(\dot{\mathbf{I}}_{fl})}$; $\mathcal{X}_{\dot{V}}(i)\backslash j$ denotes the set of node numbers of the voltage phasor variable nodes incident to the smoothness factor node with node number i, excluding j. Similarly, refer to Fig. 4(b), the message $v_{h^{(i)}_{\dot{I}_{fl}}\to x^{(j)}_{\dot{I}_{fl}}}$ could be obtained as

$$\mu_{h^{(i)}_{\dot{I}_{fl}}\to x^{(j)}_{\dot{I}_{fl}}} = \sum_{k\in\mathcal{X}_{\dot{V}}(i)} H_{x^{(k)}_{\dot{V}},h^{(i)}_{\dot{I}_{fl}}}\mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{I}_{fl}}}, \tag{12}$$

$$\sigma^2_{h^{(i)}_{\dot{I}_{fl}}\to x^{(j)}_{\dot{I}_{fl}}} = \sum_{k\in\mathcal{X}_{\dot{V}}(i)}\left|H_{x^{(k)}_{\dot{V}},h^{(i)}_{\dot{I}_{fl}}}\right|^2\sigma^2_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{I}_{fl}}}, \tag{13}$$

where $\mathcal{X}_{\dot{V}}(i)$ denotes the set of node numbers of the voltage phasor variable nodes incident to the smoothness factor node with node number i. Analogously, the message $v_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{fl}}}$, which is represented by $\mu_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{fl}}}$ and $\sigma^2_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{fl}}}$, and $v_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{inj}}}$, which is represented by $\mu_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{inj}}}$ and $\sigma^2_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{inj}}}$, could be formulated as follows.

$$\mu_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{fl}}} = \mu_{x^{(l)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{I}_{inj}}} - \sum_{k\in\mathcal{X}_{\dot{I}_{fl}}(i)\backslash j}\mu_{x^{(k)}_{\dot{I}_{fl}}\to h^{(i)}_{\dot{I}_{inj}}}, \tag{14}$$

$$\sigma^2_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{fl}}} = \sigma^2_{x^{(l)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{I}_{inj}}} + \sum_{k\in\mathcal{X}_{\dot{I}_{fl}}(i)\backslash j}\sigma^2_{x^{(k)}_{\dot{I}_{fl}}\to h^{(i)}_{\dot{I}_{inj}}}, \tag{15}$$

$$\mu_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{inj}}} = \sum_{k\in\mathcal{X}_{\dot{I}_{fl}}(i)}\mu_{x^{(k)}_{\dot{I}_{fl}}\to h^{(i)}_{\dot{I}_{inj}}}, \tag{16}$$

$$\sigma^2_{h^{(i)}_{\dot{I}_{inj}}\to x^{(j)}_{\dot{I}_{inj}}} = \sum_{k\in\mathcal{X}_{\dot{I}_{fl}}(i)}\sigma^2_{x^{(k)}_{\dot{I}_{fl}}\to h^{(i)}_{\dot{I}_{inj}}}. \tag{17}$$

2) CD-GaBP with hybrid measurements: For most power systems, phasor measurements are insufficient to ensure full system observability. In this subsection, we present a novel factor graph structure to directly process the nonlinear models of legacy measurements, avoiding the requirement of a global view and additional computational burden. The relationships between legacy and phasor measurements are formulated as follows.

$$\dot{V}_f = |V_f|\mathrm{e}^{j\theta_f}, \tag{18}$$

$$\dot{I}_{ft} = (\dot{S}_{ft}/\dot{V}_f)^*, \tag{19}$$

$$\dot{I}_f = (\dot{S}_f/\dot{V}_f)^*, \tag{20}$$

where $\dot{V}_f$, $|V_f|$, θ_f, $\dot{I}_f$, and $\dot{S}_f$ denote the voltage phasor, voltage magnitude, phase angle, current injection, and complex power injection of bus f, respectively; $\dot{I}_{ft}$ and $\dot{S}_{ft}$ are the current flow and complex power flow of branch f-t, respectively. According to (18), (19), and (20), three types of smoothness factor nodes ($h_{|V|}$, $h_{\dot{S}_{fl}}$, and $h_{\dot{S}_{inj}}$), three types of data factor nodes ($z_{|V|}$, $z_{\dot{S}_{fl}}$, and $z_{\dot{S}_{inj}}$), and three types of variable nodes ($x_{|V|}$, $x_{\dot{S}_{fl}}$, and $x_{\dot{S}_{inj}}$) are added to the complex domain factor graph. Furthermore, the virtual variable node $x_{e^{j\theta}}$ is established to link voltage magnitude with voltage phasor.

For illustration consider the 3-bus sample system in Fig. 2. When the power injection and the voltage magnitude of all buses as well as the power flow of all branches are modeled, the complex domain factor graph of bus 2 is depicted in Fig. 5. Note that each variable node (excluding the virtual variable node) is connected to a corresponding data factor node, which is omitted in Fig. 5 for brevity. Aside from the messages from variable nodes and data factor nodes, seven additional types of messages are modeled, including messages related to voltage magnitude ($v_{x_{e^{j\theta}}\to h_{|V|}}$, $v_{h_{|V|}\to x_{\dot{V}}}$, $v_{h_{|V|}\to x_{|V|}}$, and $v_{h_{|V|}\to x_{e^{j\theta}}}$), as well as messages related to power injection and power flow ($v_{h_{\dot{S}}\to x_{\dot{V}}}$, $v_{h_{\dot{S}}\to x_{\dot{I}}}$, and $v_{h_{\dot{S}}\to x_{\dot{S}}}$), where the subscript $(\cdot)_S$ denotes both power flow and power injection. Since the virtual variable nodes are not connected to any of the data factor nodes, they will transmit their beliefs (i.e., means and variances) to the smoothness factor node, equivalent to directly sending it back after receiving the message (i.e., $v_{x_{e^{j\theta}}\to h_{|V|}} = v_{h_{|V|}\to x_{e^{j\theta}}}$). Note that according to Theorem 1, the magnitude of $\sigma^2_{x^{(i)}_{e^{j\theta}}}$ should be restricted to interval $(0,1)$. In this paper, $\sigma^2_{x^{(i)}_{e^{j\theta}}}$ is restricted to $[10^{-50}, 1-10^{-50}]$, for the sake of implementation convenience.

Theorem 1: For a Gaussian random variable $\theta \sim \mathcal{N}(\mu_\theta, \sigma^2_\theta)$, the variance of $e^{j\theta}$ is $\mathrm{Var}(e^{j\theta}) = 1 - e^{-\sigma^2_\theta} \in (0,1)$.

Proof: The characteristic function of θ is given by $\phi_\theta(t) = \mathbb{E}[e^{jt\theta}] = e^{jt\mu_\theta - \sigma^2_\theta t^2/2}$ [24]. By setting t to 1 and -1, we get $\mathbb{E}[e^{j\theta}] = e^{j\mu_\theta - \sigma^2_\theta/2}$ and $\mathbb{E}[e^{-j\theta}] = e^{-j\mu_\theta - \sigma^2_\theta/2}$, respectively. Then, the variance of $e^{j\theta}$ can be derived as

$$\begin{aligned}\mathrm{Var}(e^{j\theta}) &= \mathbb{E}\big[(e^{j\theta} - \mathbb{E}[e^{j\theta}])(e^{j\theta} - \mathbb{E}[e^{j\theta}])^*\big] \\ &= \mathbb{E}[(e^{j\theta} - e^{j\mu_\theta - \frac{\sigma^2_\theta}{2}})(e^{-j\theta} - e^{-j\mu_\theta - \frac{\sigma^2_\theta}{2}})] \\ &= 1 + e^{-\sigma^2_\theta} - e^{-j\mu_\theta - \frac{\sigma^2_\theta}{2}}\mathbb{E}[e^{j\theta}] \\ &\quad - e^{j\mu_\theta - \frac{\sigma^2_\theta}{2}}\mathbb{E}[e^{-j\theta}] \\ &= 1 - e^{-\sigma^2_\theta}.\end{aligned} \tag{21}$$

Here, $\mathrm{Var}(e^{j\theta}) = 1 - e^{-\sigma^2_\theta} \in (0,1)$ for any $\sigma_\theta > 0$. ■

The message $v_{h_{|V|}\to x_{\dot{V}}}$, whose local factor graph is depicted in Fig. 6(a), could be formulated as

$$\mu_{h^{(i)}_{|V|}\to x^{(k)}_{\dot{V}}} = \mu_{x^{(l)}_{|V|}\to h^{(i)}_{|V|}}\, \mu_{x^{(j)}_{e^{j\theta}}\to h^{(i)}_{|V|}}, \tag{22}$$

$$\begin{aligned}\sigma^2_{h^{(i)}_{|V|}\to x^{(k)}_{\dot{V}}} = &\left|\mu_{x^{(l)}_{|V|}\to h^{(i)}_{|V|}}\right|^2 \sigma^2_{x^{(j)}_{e^{j\theta}}\to h^{(i)}_{|V|}} + \sigma^2_{x^{(j)}_{|V|}\to h^{(i)}_{|V|}} \\ &+\sigma^2_{x^{(j)}_{e^{j\theta}}\to h^{(i)}_{|V|}}\sigma^2_{x^{(j)}_{|V|}\to h^{(i)}_{|V|}},\end{aligned} \tag{23}$$

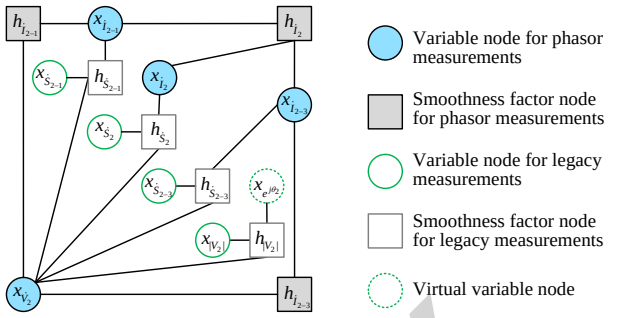


Fig. 5. Complex domain factor graph with hybrid phasor/legacy measurements for bus 2 of the 3-bus sample system. The solid circles and squares, which are already modeled in Fig. 3, represent the phasor measurements, while the hollow circles and squares represent the legacy measurements or virtual measurements (data factor nodes are omitted for brevity).

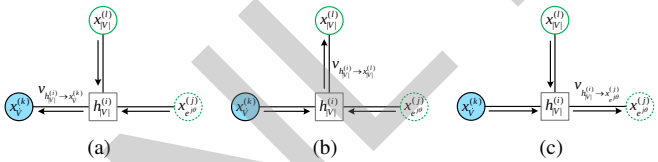


Fig. 6. Messages sent from voltage magnitude smoothness factor node $h^{(i)}_{|V|}$: (a) $v_{h^{(i)}_{|V|}\to x^{(k)}_{\dot{V}}}$; (b) $v_{h^{(i)}_{|V|}\to x^{(l)}_{|V|}}$; (c) $v_{h^{(i)}_{|V|}\to x^{(j)}_{e^{j\theta}}}$.

where the value of variance $\sigma^2_{h^{(i)}_{|V|}\to x^{(k)}_{\dot{V}}}$ is obtained according to Theorem 2 as follows.

Theorem 2: For independent complex random variables X and Y, the variance of XY is $\mathrm{Var}(XY) = |\mathbb{E}(X)|^2\mathrm{Var}(Y) + |\mathbb{E}(Y)|^2\mathrm{Var}(X) + \mathrm{Var}(X)\mathrm{Var}(Y)$.

Proof: For independent complex random variables X and Y, the variance of their product XY could be obtained as

$$\begin{aligned}\mathrm{Var}(XY) &= \mathbb{E}\big[(XY - \mathbb{E}[XY])(XY - \mathbb{E}[XY])^*\big] \\ &= \mathbb{E}\big[|X|^2|Y|^2 + |\mathbb{E}[XY]|^2 \\ &\quad -XY(\mathbb{E}[XY])^* - (XY)^*\mathbb{E}[XY]\big] \\ &= \mathbb{E}[|X|^2]\mathbb{E}[|Y|^2] - |\mathbb{E}[XY]|^2 \\ &= \big(\mathrm{Var}(X) + |\mathbb{E}[X]|^2\big)\big(\mathrm{Var}(Y) \\ &\quad +|\mathbb{E}[Y]|^2\big) - |\mathbb{E}[X]|^2|\mathbb{E}[Y]|^2 \\ &= |\mathbb{E}[X]|^2\mathrm{Var}(Y) + |\mathbb{E}[Y]|^2\mathrm{Var}(X) \\ &\quad +\mathrm{Var}(X)\mathrm{Var}(Y).\end{aligned} \tag{24}$$

We note that Theorem 2 is a complex-variable extension of [25], which employs real-valued random variables. ■

For the message $v_{h_{|V|}\to x_{|V|}}$, as shown in Fig. 6(b), the mean and variance could be obtained from $v_{x^{(k)}_{\dot{V}}\to h^{(i)}_{|V|}}$ as

$$\mu_{h^{(i)}_{|V|}\to x^{(l)}_{|V|}} = \left|\mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{|V|}}\right|, \tag{25}$$

$$\sigma^2_{h^{(i)}_{|V|}\to x^{(l)}_{|V|}} = \sigma^2_{x^{(k)}_{\dot{V}}\to h^{(i)}_{|V|}}. \tag{26}$$

Here, $\sigma^2_{h^{(i)}_{|V|}\to x^{(l)}_{|V|}}$ is larger than the theoretical value, as the error of the phase angle contained in $\sigma^2_{x^{(k)}_{\dot{V}}\to h^{(i)}_{|V|}}$ is counted into the variance of voltage magnitude $\sigma^2_{h^{(i)}_{|V|}\to x^{(l)}_{|V|}}$. However, it is worth noting that this approximation will not affect the estimation result since the only message sent from $x^{(l)}_{|V|}$ to smoothness factor node (i.e., $v_{x^{(l)}_{|V|}\to h^{(i)}_{|V|}}$) is solely determined

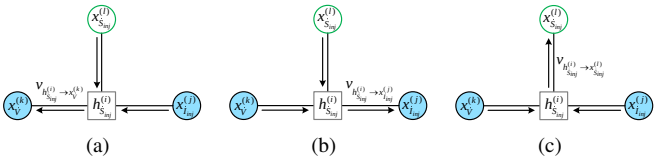


Fig. 7. Messages sent from power injection smoothness factor node $h^{(i)}_{\dot{S}_{inj}}$: (a) $v_{h^{(i)}_{\dot{S}_{inj}}\to x^{(k)}_{\dot{V}}}$; (b) $v_{h^{(i)}_{\dot{S}_{inj}}\to x^{(j)}_{\dot{I}_{inj}}}$; (c) $v_{h^{(i)}_{\dot{S}_{inj}}\to x^{(l)}_{\dot{S}_{inj}}}$.

by the measurement data of voltage magnitude. Similarly, the message $v_{h_{|V|}\to x_{e^{j\theta}}}$, whose local factor graph is shown in Fig. 6(c), could be formulated as

$$\mu_{h^{(i)}_{|V|}\to x^{(j)}_{e^{j\theta}}} = \mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{|V|}} \Big/ \left|\mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{|V|}}\right|, \tag{27}$$

$$\sigma^2_{h^{(i)}_{|V|}\to x^{(j)}_{e^{j\theta}}} = \sigma^2_{x^{(k)}_{\dot{V}}\to h^{(i)}_{|V|}} \Big/ \left|\mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{|V|}}\right|^2, \tag{28}$$

where the variance $\sigma^2_{h^{(i)}_{|V|}\to x^{(j)}_{e^{j\theta}}}$ is also larger than the theoretical value, but this is acceptable for a virtual variable.

The formulas for messages sent from $h_{\dot{S}_{fl}}$ and $h_{\dot{S}_{inj}}$ have the same structure. In this context, we will only discuss the messages sent from $h_{\dot{S}_{inj}}$ below due to space limitations. As shown in Fig. 7(a), the message $v_{h^{(i)}_{\dot{S}_{inj}}\to x^{(k)}_{\dot{V}}}$ could be obtained based on (20) as follows.

$$\mu_{h^{(i)}_{\dot{S}_{inj}}\to x^{(k)}_{\dot{V}}} = \mu_{x^{(l)}_{\dot{S}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}} \Big/ \mu^*_{x^{(j)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}}, \tag{29}$$

$$\begin{aligned}\sigma^2_{h^{(i)}_{\dot{S}_{inj}}\to x^{(k)}_{\dot{V}}} = \Big(&\left|\mu_{x^{(j)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}}\right|^2 \sigma^2_{x^{(l)}_{\dot{S}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}} \\ &+\left|\mu_{x^{(l)}_{\dot{S}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}}\right|^2 \sigma^2_{x^{(j)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}}\Big) \\ &\Big/ \left|\mu_{x^{(j)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}}\right|^4,\end{aligned} \tag{30}$$

where $\sigma^2_{h^{(i)}_{\dot{S}_{inj}}\to x^{(k)}_{\dot{V}}}$ is obtained through Theorem 3. Moreover, $v_{h^{(i)}_{\dot{S}_{inj}}\to x^{(k)}_{\dot{V}}}$ will be assigned a pseudo mean $\mu_{h^{(i)}_{\dot{S}_{inj}}\to x^{(k)}_{\dot{V}}} = 1$ and a near-infinite variance if $\mu_{x^{(j)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}} = 0$.

Theorem 3: For independent complex random variables X and Y, the variance of $Z = X/Y$ is $\mathrm{Var}(Z) = \big(|X|^2\mathrm{Var}(Y) + |Y|^2\mathrm{Var}(X)\big)/|Y|^4$.

Proof: we can estimate $Z - \mathbb{E}[Z]$ by the first-order Taylor series expansion of $Z = X/Y$ as

$$Z - \mathbb{E}[Z] = (X - \mathbb{E}[X])\frac{\partial Z}{\partial X} + (Y - \mathbb{E}[Y])\frac{\partial Z}{\partial Y}. \tag{31}$$

Then, the variance of Z could be formulated as

$$\begin{aligned}\mathrm{Var}(Z) &= \mathbb{E}\big[(Z - \mathbb{E}[Z])(Z - \mathbb{E}[Z])^*\big] \\ &= \mathbb{E}\big[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^* \frac{\partial Z}{\partial X}(\frac{\partial Z}{\partial X})^* \\ &\quad +(Y - \mathbb{E}[Y])(Y - \mathbb{E}[Y])^* \frac{\partial Z}{\partial Y}(\frac{\partial Z}{\partial Y})^* \\ &\quad +(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^* \frac{\partial Z}{\partial X}(\frac{\partial Z}{\partial Y})^* \\ &\quad +(X - \mathbb{E}[X])^*(Y - \mathbb{E}[Y])(\frac{\partial Z}{\partial X})^* \frac{\partial Z}{\partial Y}\big].\end{aligned} \tag{32}$$

Since X is independent of Y, the last two terms (i.e., the last two lines) of (32) are zero. Substituting $\partial Z/\partial X = 1/Y$ and $\partial Z/\partial Y = -X/Y^2$ into (32) yields

$$\begin{aligned}\mathrm{Var}(Z) &= \mathbb{E}\big[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^* \frac{\partial Z}{\partial X}(\frac{\partial Z}{\partial X})^* \\ &\quad +(Y - \mathbb{E}[Y])(Y - \mathbb{E}[Y])^* \frac{\partial Z}{\partial Y}(\frac{\partial Z}{\partial Y})^*\big] \\ &= \mathrm{Var}(X)/|Y|^2 + \mathrm{Var}(Y)|X|^2/|Y|^4 \\ &= \big(|X|^2\mathrm{Var}(Y) + |Y|^2\mathrm{Var}(X)\big)/|Y|^4.\end{aligned} \tag{33}$$

Please note that Theorem 3 could be developed from [26], which employs real-valued random variables. Furthermore, the variance obtained here is through a first-order Taylor expansion, indicating that while this variance is relatively accurate, it is still an approximation. Therefore, the proposed method may have a different convergence process than WLS, and theoretically, it does not converge to the WLS solution. However, this approximation has little impact on the estimation accuracy, as evidenced by the simulation results in Section IV. ■

Similar to $v_{h^{(i)}_{\dot{S}_{inj}}\to x^{(k)}_{\dot{V}}}$, the message $v_{h^{(i)}_{\dot{S}_{inj}}\to x^{(j)}_{\dot{I}_{inj}}}$, whose local factor graph is given in Fig. 7(b), could be formulated as

$$\mu_{h^{(i)}_{\dot{S}_{inj}}\to x^{(j)}_{\dot{I}_{inj}}} = \left(\mu_{x^{(l)}_{\dot{S}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}} \Big/ \mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{S}_{inj}}}\right)^*, \tag{34}$$

$$\begin{aligned}\sigma^2_{h^{(i)}_{\dot{S}_{inj}}\to x^{(j)}_{\dot{I}_{inj}}} = \Big(&\left|\mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{S}_{inj}}}\right|^2 \sigma^2_{x^{(l)}_{\dot{S}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}} \\ &+\left|\mu_{x^{(l)}_{\dot{S}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}}\right|^2 \sigma^2_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{S}_{inj}}}\Big) \\ &\Big/ \left|\mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{S}_{inj}}}\right|^4.\end{aligned} \tag{35}$$

Refer to the local factor graph in Fig. 7(c), the message $v_{h^{(i)}_{\dot{S}_{inj}}\to x^{(l)}_{\dot{S}_{inj}}}$ could be formulated as

$$\mu_{h^{(i)}_{\dot{S}_{inj}}\to x^{(l)}_{\dot{S}_{inj}}} = \mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{S}_{inj}}}\, \mu^*_{x^{(j)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}}, \tag{36}$$

$$\begin{aligned}\sigma^2_{h^{(i)}_{\dot{S}_{inj}}\to x^{(l)}_{\dot{S}_{inj}}} = &\left|\mu_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{S}_{inj}}}\right|^2 \sigma^2_{x^{(j)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}} \\ &+\left|\mu_{x^{(j)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}}\right|^2 \sigma^2_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{S}_{inj}}} \\ &+\sigma^2_{x^{(k)}_{\dot{V}}\to h^{(i)}_{\dot{S}_{inj}}}\, \sigma^2_{x^{(j)}_{\dot{I}_{inj}}\to h^{(i)}_{\dot{S}_{inj}}},\end{aligned} \tag{37}$$

where (37) is derived based on Theorem 2. Conclusively, the discussion above shows that the nonlinear relationships between legacy measurements and phasor measurements could be represented through the complex domain factor graph. Consequently, the proposed CD-GaBP method could be carried out in a distributed manner without requiring a global view.

Remark 1: The magnitude measurements of current flows $|I_{ft}|$ and current injections $|I_f|$ can also be incorporated into the proposed CD-GaBP method. The relationship between them and the phasor measurements can be expressed as $\dot{I}_f = |I_f|e^{j\varphi_f}$ and $\dot{I}_{ft} = |I_{ft}|e^{j\varphi_{ft}}$. The corresponding messages can be formulated in the same way as voltage magnitude (as shown in (22), (23), (25)-(28)). Due to the similarity of the formulas and space limitations, the detailed description is omitted.

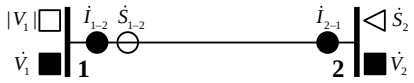


Fig. 8. 2-bus sample system with phasor and legacy measurements.

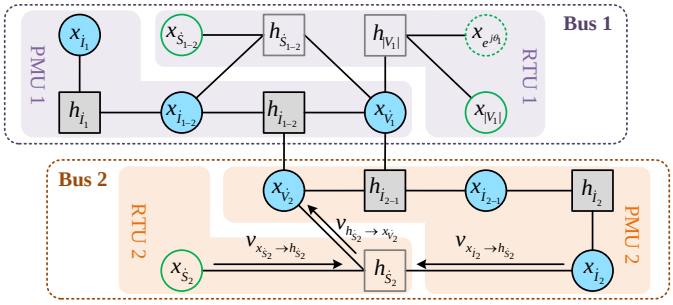


Fig. 9. Complex domain factor graph of CD-GaBP for the 2-bus sample system (data factor nodes are omitted for brevity).

3) A simple case for CD-GaBP: To illustrate the message propagation process consider a 2-bus sample system with hybrid measurements in Fig. 8, the corresponding complex domain factor graph of CD-GaBP is presented in Fig. 9. In the initialization step, each state variable node send the message that consist of initial mean value (flat start or warm start [18]) and near-infinite variance to incident smoothness factor nodes. Subsequently, the smoothness factor nodes will update their outgoing messages according to the incoming messages. For example, when messages $v_{x_{i_2}\to h_{\dot{S}_2}}$ and/or $v_{x_{\dot{S}_2}\to h_{\dot{S}_2}}$ are updated, the mean and variance of message $v_{h_{\dot{S}_2}\to x_{\dot{V}_2}}$ will be formulated according to (29) and (30) as $\mu_{h_{\dot{S}_2}\to x_{\dot{V}_2}} = \mu_{x_{\dot{S}_2}\to h_{\dot{S}_2}}/\mu^*_{x_{i_2}\to h_{\dot{S}_2}}$ and $\sigma^2_{h_{\dot{S}_2}\to x_{\dot{V}_2}} = (|\mu_{x_{i_2}\to h_{\dot{S}_2}}|^2\sigma^2_{x_{\dot{S}_2}\to h_{\dot{S}_2}} + |\mu_{x_{\dot{S}_2}\to h_{\dot{S}_2}}|^2\sigma^2_{x_{i_2}\to h_{\dot{S}_2}})/|\mu_{x_{i_2}\to h_{\dot{S}_2}}|^4$. Similarly, when the incoming messages of a state variable is updated, the corresponding outgoing message will be formulated according to (2) and (3). The proposed method takes meters as the minimum computation units. For illustration consider the 2-bus sample system in Fig. 8. When measurement data is collected from phasor measurement units (PMUs) and remote terminal units (RTUs) at buses 1 and 2, we can map the complex domain factor graph in Fig. 9 into four minimum computation units, corresponding to the four measurement devices. In practical implementation, we can group the minimum computation units according to actual requirements. For example, we can group the minimum computation units based on buses as depicted in Fig. 9. Further, we can even assign the nodes in the complex domain factor graph of multiple buses to a single computing device. It is worthwhile to note that the structural differences between complex domain factor graphs and communication systems should also be considered in practical implementation. This aspect will be further investigated in our future studies.

4) Complexity analysis: The computational complexity of GaBP depends on the following three aspects: the complexity involved in calculating each message, the number of edges in the factor graph, and the number of iterations required for convergence. Given the inherent sparsity in the topology of power systems, the complexity of each message could be considered $\mathcal{O}(1)$ [17]. In this way, the complexity of a GaBP algorithm can be expressed by $\mathcal{O}(\sum_{i=1}^{E} L_i)$, where E and L_i are the number of edges and the number of iterations for ith edge, respectively. In CD-GaBP and GN-BP, the number of edges is $3n + 10b + 4m_L$ and $\sum_{i=1}^{2n} d_i$, respectively, where n, b, and m_L denote the number of buses, branches, and legacy measurements, respectively, d_i denotes the degree of the variable node i. The above analysis indicates that the computational complexities for CD-GaBP and GN-BP are both $\mathcal{O}(n)$ per iteration. Furthermore, according to our tests, the total number of edges in the proposed method is comparable to that in GN-BP, and it may even be fewer, especially for power systems with high measurement redundancy. Moreover, compared to the conventional factor graph that only constructs voltage phasors as variable nodes, the proposed complex domain factor graph decouples the relationship between current injections and voltage phasors, leading to fewer short cycles, thereby enhancing convergence and reducing the number of iterations [13]. Furthermore, the GaBP model for GN-BP needs to be executed multiple times to achieve convergence in the outer iteration loop (as illustrated in Section II), thereby further increasing the total iteration number. Therefore, the overall computational complexity of the proposed method is expected to be lower than that of GN-BP.

B. Robust Data Factor Nodes based on M-estimation

Compared to centralized estimators, the proposed method faces greater challenges in processing bad data because each node can only leverage the information from itself and the connected edges. To deal with this challenge, we employ the M-estimation theory to construct robust data factor nodes by scaling the variance of the measurement data. Consider a data factor node whose noise follows a Gaussian distribution as

$$z_s(\mathbf{x}) = Ke^{-\frac{1}{2}M_s^2} = Ke^{-\frac{1}{2}(z_s^m - h_s(\mathbf{x}))^T w_s (z_s^m - h_s(\mathbf{x}))}, \quad (38)$$

where z_s^m is the sth measurement; $h_s(\cdot)$ and w_s are the corresponding measurement function and weight, respectively; K denotes the scaling factor for normalization, which does not need to be calculated [13]; $M_s = \sqrt{(z_s^m - h_s(\mathbf{x}))^T w_s (z_s^m - h_s(\mathbf{x}))}$ represents the Mahalanobis distance [27]. To mitigate the impact of outliers and non-Gaussian noises, we could replace M_s^2, which correspond to L2 cost, with a sub-quadratic cost function. Some commonly used cost functions, including L2, L1, Huber [28], and Hampel [29], are depicted in Fig. 10. The Huber and Hampel functions yield the same cost as L2 for small errors, indicating that they can provide accurate estimates similar to L2 under the Gaussian distribution. In this paper, the Huber and Hampel cost functions are utilized for handling bad data. Specifically, the Huber function is applied to the initial calculation after updating measurement data to mitigate the impact of outliers on convergence performance, given that the Hampel function is non-convex, while the Hampel function will be employed in the subsequent calculation processes for superior robustness.

A data factor node with Huber cost could be expressed as

$$z_s^{Hu}(\mathbf{x}) = \begin{cases} Ke^{-\frac{1}{2}M_s^2} & M_s \le c \\ Ke^{-cM_s + \frac{1}{2}c^2} & c < M_s, \end{cases} \quad (39)$$

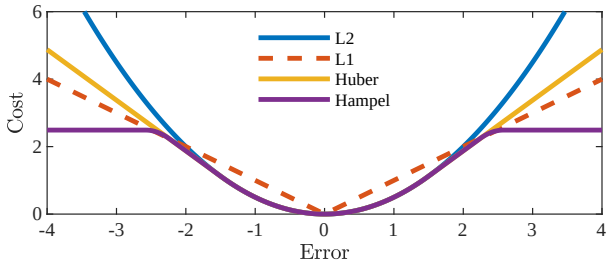


Fig. 10. Some commonly used cost functions.

where c is a positive parameter that adjusts the shape of the cost function and is set to 1.96 with confidence levels of 95% [29]. Then, the equivalent L2 cost for Huber, M_s^{Hu}, could be obtained through

$$\frac{1}{2}(M_s^{Hu})^2 = \begin{cases} \frac{1}{2}M_s^2 & M_s \le c \\ cM_s - \frac{1}{2}c^2 & c < M_s. \end{cases} \tag{40}$$

Then, the scaling ratio k_s^{Hu} can be obtained as

$$k_s^{Hu} = \frac{(M_s^{Hu})^2}{M_s^2} = \begin{cases} 1 & M_s < c \\ \frac{1}{M_s^2}(2cM_s - c^2) & c < M_s. \end{cases} \tag{41}$$

For the Hampel function, its data factor node is expressed as

$$z_s^{Ha}(\mathbf{x}) = \begin{cases} Ke^{-\frac{1}{2}M_s^2} & M_s \le c_1 \\ Ke^{-c_1 M_s + \frac{1}{2}c_1^2} & c_1 < M_s \le c_2 \\ Ke^{-\frac{1}{2}c_1(c_2+c_3-c_1+\frac{M_s-c_3}{c_2-c_3})} & c_2 < M_s \le c_3 \\ Ke^{-\frac{1}{2}c_1(c_2+c_3-c_1)} & c_3 < M_s, \end{cases} \tag{42}$$

where the thresholds c_1, c_2, and c_3 adjust the shape of the Hampel cost function. The smaller the value of these thresholds, the higher the robustness, but the lower the accuracy. In this paper, we follow the suggestions in [29] and [30] by setting these thresholds to $c_1 = 1.96$, $c_2 = 2.24$, and $c_3 = 2.58$ with confidence levels of 95%, 97.5%, and 99%, respectively. In the same way as k_s^{Hu}, the scaling ratio for the Hampel cost function could be obtained as follows

$$k_s^{Ha} = \begin{cases} 1 & M_s \le c_1 \\ \frac{1}{M_s^2}(2c_1 M_s - c_1^2) & c_1 < M_s \le c_2 \\ \frac{1}{M_s^2}c_1(c_2+c_3-c_1+\frac{M_s-c_3}{c_2-c_3}) & c_2 < M_s \le c_3 \\ \frac{1}{M_s^2}c_1(c_2+c_3-c_1) & c_3 < M_s. \end{cases} \tag{43}$$

For the sth data factor node, the variance of the outgoing message will be initially scaled by the scaling ratio k_s^{Hu} and subsequently by k_s^{Ha} to suppress the impact of bad data.

C. Distributed Iteration Strategy

In conventional GaBP methods, the convergence is determined by checking whether the changes in all state variables or messages between two iterations are smaller than a predefined threshold. However, this strategy requires a global view, and hence, is not feasible for CD-GaBP in distributed execution scenarios, where each variable node can only access information from itself and its neighbors. To address this issue, a distributed iteration strategy is proposed as follows.

TABLE I
MEASUREMENT STANDARD DEVIATIONS

Legacy measurements			Phasor measurements		
Voltage	Injected power	Power flows	Voltage	Current	Phase angle
0.5%	1%	1%	0.1%	0.1%	0.1°

1) Consecutive Convergence Test: For each variable node, the convergence counter is incremented when the change in its mean exceeds the convergence threshold α; otherwise, it is reset to 0. Convergence is confirmed only when the convergence counter reaches a preset value τ. This design aims to enhance the robustness of convergence determination, mitigating the risk of erroneous judgments caused by occasional fluctuations. In this paper, the α and τ are set to 10^{-6} and 10, respectively.

2) Reactivation of Computation: For each variable node that has already converged, if the change in its mean exceeds the reactivation threshold γ, this variable node will be reactivated and require further iterations. In this paper, the reactivation threshold γ is set to 10^{-5}.

With the distributed iteration strategy, each variable node can autonomously provide estimates without waiting for the entire SE model to converge. In the scenario of large-scale power system state estimation, the calculation of the proposed method can be distributed to various substations or edge computing devices, thereby alleviating the communication, storage, and computation burdens on the control center.

IV. PRACTICAL CASE STUDIES

The proposed CD-GaBP method was compared against the GN-BP method [17], the conventional WLS method [18], and several robust estimators. The performance index used was the mean absolute error (MAE), which is defined as

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\dot{V}_i^{esti} - \dot{V}_i^{true}|, \tag{44}$$

where n is the number of buses; $\dot{V}_i^{esti}$ and $\dot{V}_i^{true}$ represent the estimated and true values of the voltage phasor for bus i, respectively. Unless stated otherwise, 20% of the buses were equipped with PMUs, providing voltage phasor and current flow measurements of all connected branches. Meanwhile, the voltage magnitudes, injected powers of all buses, and power flows of all branches are monitored by legacy measurements. Further, independent zero-mean Gaussian noises with standard deviations σ, as specified in Table I, are applied to the measurements. The convergence threshold for all methods was set to 10^{-6}. Moreover, the total iteration limit for GN-BP was set to 6×10^4, with an upper limit of 10 for the outer iteration loop and 6000 for each inner iteration loop, to prevent endless iterations [17]. Unless otherwise specified, all results were obtained through 3000 Monte Carlo (MC) simulations conducted on a PC with an Intel Core i7-12700 CPU with 16GB of RAM. Furthermore, for centralized methods, sparse techniques were utilized to accelerate calculations.

TABLE II
AVERAGE MAE VALUES FOR DIFFERENT METHODS (P.U.)

System	CD-GaBP	GN-BP	WLS
IEEE 14-bus	6.21×10^{-4}	6.32×10^{-4}	6.19×10^{-4}
IEEE 118-bus	5.15×10^{-4}	6.17×10^{-4}	5.09×10^{-4}
IEEE 300-bus	8.58×10^{-4}	2.99×10^{-3}	8.66×10^{-4}

TABLE III
AVERAGE ITERATION NUMBERS FOR DIFFERENT METHODS

System	CD-GaBP	GN-BP	WLS
IEEE 14-bus	1.82×10^{2}	2.30×10^{4}	4.00
IEEE 118-bus	2.50×10^{2}	6.00×10^{4}	1.83×10^{1}
IEEE 300-bus	3.84×10^{2}	6.00×10^{4}	1.16×10^{1}

TABLE IV
AVERAGE EXECUTION TIME FOR DIFFERENT METHODS (S)

System	CD-GaBP	GN-BP	WLS
IEEE 14-bus	2.78×10^{-2}	4.64	3.51×10^{-3}
IEEE 118-bus	1.84×10^{-1}	1.60×10^{2}	4.94×10^{-2}
IEEE 300-bus	4.67×10^{-1}	3.15×10^{2}	6.49×10^{-2}

A. Assessment of Accuracy and Efficiency

In this section, the estimation accuracy and computational efficiency for the proposed CD-GaBP method, the GN-BP method, and the WLS method are compared. Table II shows the average MAE values for the three methods. It is clear that the estimation accuracy of the proposed algorithm is significantly higher than that of GN-BP and is comparable to WLS. This is mainly because the proposed complex domain factor graph has fewer short cycles, which improves the convergence and accuracy of CD-GaBP. It can be found from Tables III and IV that the number of iterations for CD-GaBP is 2 orders of magnitude fewer than that of GN-BP. Accordingly, the computational efficiency of CD-GaBP is also significantly superior to that of GN-BP. In addition, it is important to note that the average iteration number of CD-GaBP corresponds to the variable nodes with the maximum iteration number. Therefore, as illustrated by the convergence process in an arbitrarily selected MC run in Fig. 11, the iteration numbers for the majority of variable nodes in CD-GaBP are fewer than those shown in Table III. Moreover, as shown in Table IV, given that a centralized estimator can choose highly efficient solvers from a global perspective and utilize accelerated computational methods such as sparse Cholesky decomposition, the computational efficiency of the proposed method is inferior to that of the WLS method. This deficiency will be even more pronounced in large-scale power systems. For example, in the Polish 2383-bus system, the average execution time for WLS is 5.41×10^{-1} seconds, while that for CD-GaBP is 30.82 seconds, which is 3 orders of magnitude greater than that for WLS. However, we highlight again that the main advantage of the proposed method lies in its capability for fully distributed computing. Additionally, the computational efficiency of the BP-based algorithm can be directly improved by increasing the number of processors or threads [13]. To demonstrate this point, we increased the thread count of CD-GaBP to 3 in the IEEE 14-bus system. Specifically, calculations for buses 1 to 3 and 5 were assigned to thread 1, those for buses 4, 7 to 9 were assigned to thread 2, and those for the remaining buses were assigned to thread 3. After implementing multi-threaded calculations, the overall computation time decreased from 2.78×10^{-2} seconds to 1.33×10^{-2} seconds. In practical applications, the calculations of CD-GaBP could be distributed to various substations or edge computing devices to alleviate the computational burden on the control center.

B. Impact of Measurement Precision

To assess the impact of measurement precision on the proposed method, three standard deviation values are considered

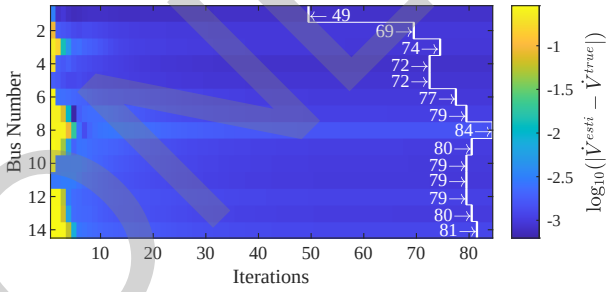


Fig. 11. Convergence process for each voltage phasor variable node. The white lines and numbers correspond to the number of iterations, while the colors represent the logarithm of absolute errors.

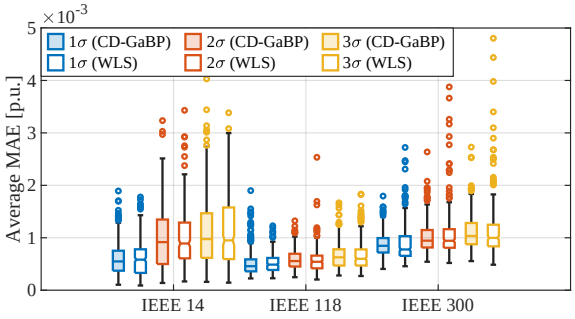


Fig. 12. Average MAE values for CD-GaBP and WLS with different measurement precisions.

for all measurements: in addition to the standard deviation σ shown in Table I, standard deviations of 2σ and 3σ were considered. Fig. 12 shows the average MAE values of CD-GaBP and WLS. As expected, a higher standard deviation led to lower estimation accuracy. Nevertheless, the increase in MAE for the proposed method is very close to that for the WLS method, both remaining within the order of 10^{-3}. This indicates that the estimated accuracy of CD-GaBP is not heavily dependent on the measurement precision.

C. Impact of Non-Gaussian Noise

Recent studies indicate that the noise in phasor measurements may follow a non-Gaussian distribution [31]. In this

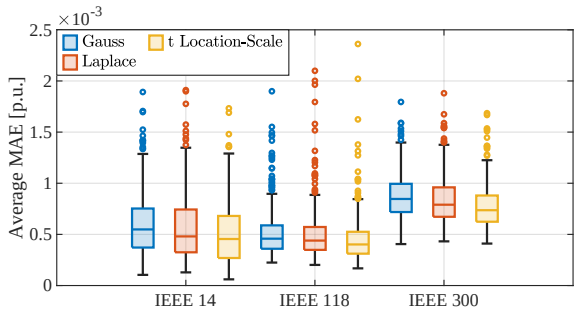


Fig. 13. Average MAE values for CD-GaBP with different noise distributions.

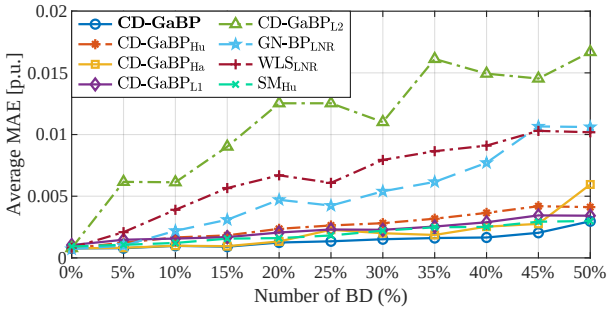


Fig. 14. Average MAE values with increases in the number of BD.

TABLE V
BAD DATA SET IN THE EXTREME BD SCENARIO

Measurement Type	Measurement
Phasor measurement	All current flow measurements, namely $\dot{I}_{5-1}$, $\dot{I}_{5-2}$, $\dot{I}_{5-4}$, $\dot{I}_{5-6}$, $\dot{I}_{11-6}$, $\dot{I}_{11-10}$
Legacy measurement	V_9, P_1, Q_1, P_2, Q_2, P_3, Q_3, P_9, Q_9, P_{13}, Q_{13}, P_{1-2}, Q_{1-2}, P_{2-4}, Q_{2-4}, P_{4-7}, Q_{4-7}, P_{4-5}, Q_{4-5}, P_{5-6}, Q_{5-6}, P_{6-12}, Q_{6-12}, P_{7-4}, Q_{7-4}

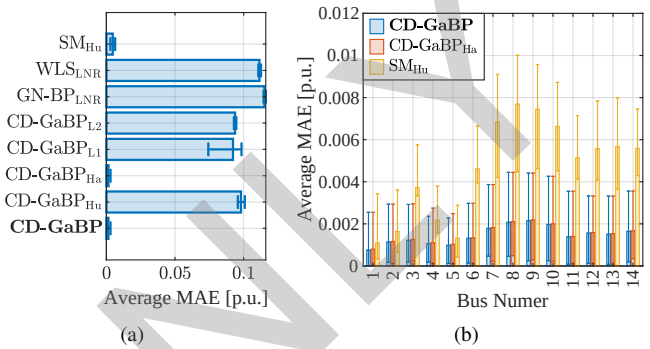


Fig. 15. Average MAE values for different methods in the extreme BD scenario. The error bars indicate the maximum and minimum values of MAE.

regard, we simulated the noise of phasor measurements with Laplace and t location-scale distributions, and the coefficients of these distributions are set to make the standard deviation of each distribution equal to σ (as in [32]). The Average MAE values of CD-GaBP for the IEEE 14-, 118-, and 300-bus systems are presented in Fig. 13, from which we can observe that the estimation error under non-Gaussian noise of CD-GaBP is not significantly different from, and even be lower than, that under Gaussian noise. This is mainly because the cost function for the data factor nodes in the proposed method is a combination of Huber and Hampel, both of which do not heavily rely on the Gaussian assumption.

D. Impact of Bad Data

In this section, the proposed CD-GaBP method was compared in the IEEE 14-bus system with GN-BP incorporating a LNRT-like procedure (GN-BP$_{\mathrm{LNR}}$) [17], WLS with LNRT (WLS$_{\mathrm{LNR}}$) [18], Schweppe-type M-estimator with Huber cost function (SM$_{\mathrm{Hu}}$) [33], and four variants of the proposed CD-GaBP method with L2 cost function (CD-GaBP$_{\mathrm{L2}}$), L1 cost function (CD-GaBP$_{\mathrm{L1}}$), Huber cost function (CD-GaBP$_{\mathrm{Hu}}$), and Hampel cost function (CD-GaBP$_{\mathrm{Ha}}$). The tests were conducted for both normal and extreme BD scenarios.

In the normal BD scenario, the standard deviations of a subset of the measurements were set to 40σ to simulate BD. The locations of BD were randomly selected in a certain quantity for each MC run. Fig. 14 depicts the average MAE values over the rate of the number of BD to the number of measurements. It is clear that the proposed CD-GaBP method consistently exhibits the lowest average MAE.

In the extreme BD scenario, PMUs were placed on buses 5 and 11. The magnitudes of the measurements listed in Table V were increased by 50% to simulate BD. As shown in Fig. 15(a), CD-GaBP, CD-GaBP$_{\mathrm{Ha}}$, and SM$_{\mathrm{Hu}}$ demonstrate significantly higher accuracy compared to other algorithms. The accuracy of CD-GaBP$_{\mathrm{Hu}}$ and CD-GaBP$_{\mathrm{L1}}$ are lower than CD-GaBP, mainly due to the relatively large costs of Huber and L1 cost functions when the error is large. Fig. 15(b) further illustrates the average MAE of each bus for CD-GaBP, CD-GaBP$_{\mathrm{Ha}}$, and SM$_{\mathrm{Hu}}$. It can be observed that CD-GaBP outperforms SM$_{\mathrm{Hu}}$ in accuracy for all buses. Moreover, the error of each bus for CD-GaBP$_{\mathrm{Ha}}$ is comparable to that of CD-GaBP, but the overall accuracy is lower (the average MAE for CD-GaBP and CD-GaBP$_{\mathrm{Ha}}$ are 1.47×10^{-3} and 1.50×10^{-3} , respectively). Additionally, please note that the BD included in critical measurement sets cannot be processed by any SE method, including the proposed method, if no redundant measurements are introduced [18].

E. Impact of Measurement Redundancy

The transmission system in some regions and most of the distribution systems still suffer from low measurement redundancy. In this context, we assessed CD-GaBP in the following four cases with different measurement redundancies.

Case 1: No phasor measurements, and about half of the legacy measurements (i.e., voltage magnitude measurements of all buses, power injection measurements of half of the buses, and power flow measurements at one terminal of all branches).

Case 2: 10% of the buses were deployed with PMUs, and about half of the legacy measurements.

Case 3: 30% of the buses were deployed with PMUs, and full legacy measurements (i.e., voltage magnitude and power injection measurements of all buses, and power flow measurements at both terminals of all branches).

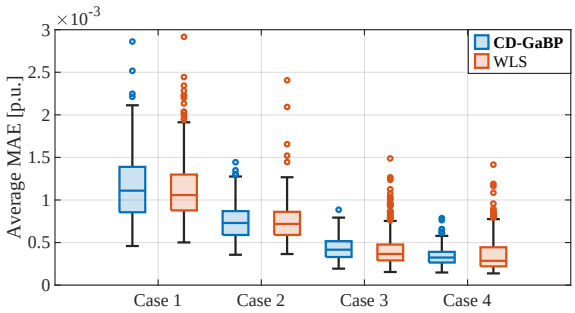


Fig. 16. Average MAE values for different measurement redundancies on the IEEE 118-bus system.

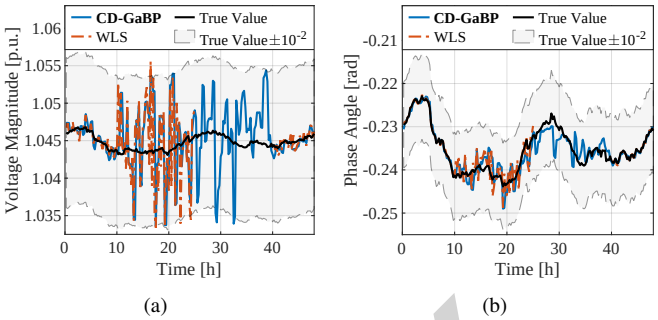


Fig. 17. State changes and estimates of bus 12 for the IEEE 14-bus system in the presence of missing data: (a) voltage magnitude; (b) phase angle.

Case 4: 50% of the buses were deployed with PMUs, full legacy measurements.

The measurement redundancy increases gradually from Case 1 to Case 4. As shown in Fig. 16, under various measurement redundancies, the estimation accuracy of the proposed CD-GaBP method is very close to that of WLS, indicating that the measurement redundancy does not have a significant impact on the proposed method.

F. Impact of Missing Data

The effectiveness of the proposed method in the presence of missing data was assessed in the IEEE 14-bus system through time-continuous simulations, employing an optimal power flow program [34] and real-world load profiles [35] with a sampling rate of one frame every 15 minutes. Two PMUs were deployed on buses 5 and 11. Meanwhile, 37 legacy measurements were considered, which include 1 voltage magnitude at bus 1, 22 power injections at buses 1 to 4 and 7 to 13, as well as 14 power flows at branches 1-2, 2-4, 4-7, 7-8, 7-9, 9-10, and 9-14 (all of the power measurements are in active/reactive pairs). All the phasor measurement data were missing between 10 to 40 hours, and the power injection measurements at bus 12, which are critical measurements without phasor measurements, were missing between 25 to 40 hours. The state estimates over time of bus 12 for an arbitrarily selected MC simulation run are depicted in Fig. 17, and the average MAE values for CD-GaBP and WLS are shown in Fig. 18(a). It can be observed that the accuracy of CD-GaBP decreases rapidly when all phasor measurements are missing, but remains similar to that of the WLS algorithm. Moreover, after the critical measurements P_{12} and Q_{12} were missing, WLS failed to converge, whereas CD-GaBP continues to provide estimates, thanks to the prior information provided by previous iterations. Additionally, a useful observation can be found from Fig. 18(b) that the convergence rate of CD-GaBP could be improved by the use of phasor measurements. It is important to note that when the system undergoes large state changes, the absence of crucial measurements may lead to poor tracking performance for unobservable buses. This aspect will be discussed in Section IV-G.

G. Impact of Sudden State Changes

To validate the effectiveness of the proposed method under sudden state changes, a time-domain simulation program [36]

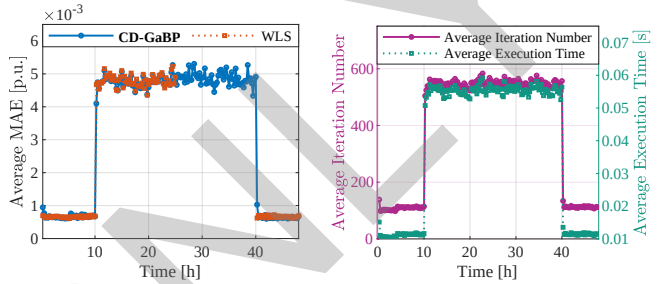


Fig. 18. Average MAE values, iteration numbers, and execution time for the IEEE 14-bus system in the presence of missing data: (a) average MAE values; (b) average iteration numbers and execution time.

was utilized to generate true states for the IEEE 14-bus system, where a doubly-fed induction generator with the Mexican hat wavelet wind model [37] was deployed at bus 1, a load shedding was simulated on bus 9 at 1st second, a generator shedding was simulated on bus 8 at 3rd second, and branch 1-5 was disconnected at 7th second. Legacy measurements were obtained every 2 seconds and PMUs were deployed at buses 5 and 11 with a sampling rate of 30 frames/s. In this context, buses 3, 7 to 9, and 12 to 14 can not be observed by PMUs. As can be seen from the state estimates of bus 9 in Fig. 19 and the average MAE values in Fig. 20(a), when a sudden state change occurs (at 1st, 3rd, and 7th seconds), the estimation accuracy is significantly affected until the next legacy measurement update (at 2nd, 4th, and 8th seconds). Nonetheless, the errors caused by sudden state changes are bounded. In cases where the observability of the entire power system cannot be guaranteed through phasor measurements, these estimates would be still useful for power system control and protection applications. Additionally, as shown in Fig. 20(b), the sudden state change results in an increase in the number of iterations, due to the significant inconsistency between the prior information in CD-GaBP and the true system states.

V. Conclusion

This paper presents a meter-level FDSE, which is capable of solving the nonlinear hybrid state estimation problems through local message exchange with few iterations. By incorporating the Huber and Hampel cost functions into the data factor node, the proposed method achieves comparable or even superior robustness to centralized robust estimators.

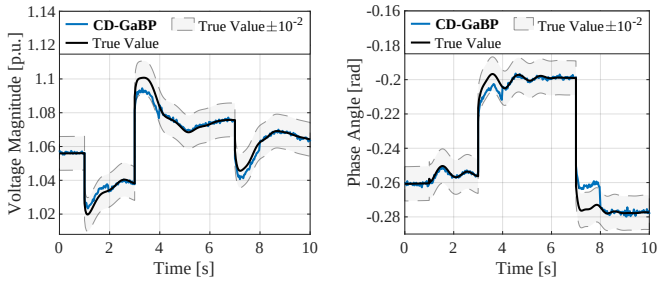


Fig. 19. State changes and estimates of bus 9 for the IEEE 14-bus system in the presence of sudden state changes: (a) voltage magnitude; (b) phase angle.

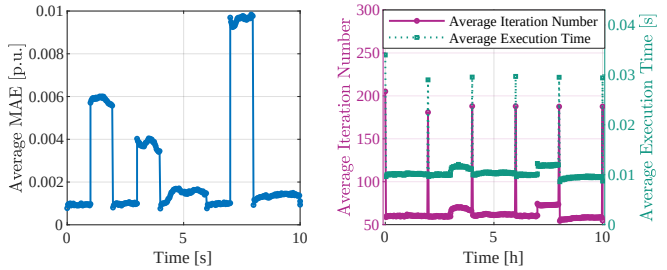


Fig. 20. Average MAE values, iteration numbers, and execution time for the IEEE 14-bus system in the presence of sudden state changes: (a) average MAE values; (b) average iteration numbers and execution time.

Moreover, the proposed distributed iteration strategy not only ensures the fully distributed nature of the CD-GaBP method but also improves its estimation accuracy by providing prior information. Extensive simulations validate the effectiveness of the proposed method in the presence of non-Gaussian noise, bad data, missing data, and sudden state changes.

REFERENCES

[1] V. Terzija, G. Valverde, C. Deyu, *et al.*, "Wide-area monitoring, protection, and control of future electric power networks," *Proc. IEEE*, vol. 99, no. 1, pp. 80–93, Jan. 2011.

[2] R. Ebrahimian and R. Baldick, "State estimation distributed processing [for power systems]," *IEEE Trans. Power Syst.*, vol. 15, no. 4, pp. 1240–1246, Nov. 2000.

[3] D. M. Falcao, F. F. Wu, and L. Murphy, "Parallel and distributed state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 2, pp. 724–730, May 1995.

[4] G. N. Korres, "A distributed multiarea state estimation," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 73–84, Feb. 2011.

[5] L. Zhao and A. Abur, "Multiarea state estimation using synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 611–617, May 2005.

[6] H. Zhu and G. B. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 6, pp. 1039–1050, Dec. 2014.

[7] A. Sagan, Y. Liu, and A. Bernstein, "Decentralized low-rank state estimation for power distribution systems," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3097–3106, Jul. 2021.

[8] Y. Liu, A. S. Zamzam, and A. Bernstein, "Multi-area distribution system state estimation via distributed tensor completion," *IEEE Trans. Smart Grid*, pp. 1–1, Nov. 2022.

[9] J. Wang and T. Li, "Distributed multi-area state estimation for power systems with switching communication graphs," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 787–797, Jan. 2021.

[10] U. C. Yilmaz and A. Abur, "A robust parallel distributed state estimation for large scale distribution systems," *IEEE Trans. Power Syst.*, vol. 39, no. 2, pp. 4437–4445, Mar. 2024.

[11] K. Saxena and A. R. Abhyankar, "Agent-based distributed computing for power system state estimation," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 5193–5202, Nov. 2020.

[12] Q. Li, L. Cheng, W. Gao, *et al.*, "Fully distributed state estimation for power system with information propagation algorithm," *J. Mod. Power Syst. Clean Energy*, vol. 8, no. 4, pp. 627–635, Jul. 2020.

[13] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, 2009.

[14] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*, ser. Representation and reasoning. San Francisco, California: Morgan Kaufmann, 2014.

[15] Y. Hu, A. Kuh, T. Yang, *et al.*, "A belief propagation based power distribution system state estimator," *IEEE Comput. Intell. Mag.*, vol. 6, no. 3, pp. 36–46, Aug. 2011.

[16] M. Cosovic, D. Miskovic, M. Delalic, *et al.*, "Distributed inference over linear models using alternating Gaussian belief propagation," *IEEE Internet Things J.*, vol. 10, no. 22, pp. 19 949–19 963, Nov. 2023.

[17] M. Cosovic and D. Vukobratovic, "Distributed Gauss–Newton method for state estimation using belief propagation," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 648–658, Jan. 2019.

[18] A. Abur and A. Gomez-Exposito, *Power System State Estimation: Theory and Implementation*. New York: Marcel Dekker, 2004.

[19] K. I. Park, *Fundamentals of Probability and Stochastic Processes with Applications to Communications*. Cham, Switzerland: Springer International Publishing, 2018.

[20] A. S. Dobakhshari, M. Abdolmaleki, V. Terzija, *et al.*, "Online non-iterative estimation of transmission line and transformer parameters by SCADA data," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2632–2641, May 2021.

[21] H.-A. Loeliger, J. Dauwels, J. Hu, *et al.*, "The factor graph approach to model-based signal processing," *Proc. IEEE*, vol. 95, no. 6, pp. 1295–1322, Jun. 2007.

[22] D. M. M. Willsky, J. K. Johnson, and A. S., "Walk-sums and belief propagation in Gaussian graphical models," *J. Mach. Learn. Res.*, vol. 7, no. 73, pp. 2031–2064, Dec. 2006.

[23] A. Gomez-Exposito, A. Abur, P. Rousseaux, *et al.*, "On the use of PMUs in power system state estimation," in *17th power system computation conference*, Aug. 2011.

[24] F. Oberhettinger, *Fourier transforms of distributions and their inverses: a collection of tables*. New York: Academic Press, 1973.

[25] L. A. Goodman, "On the exact variance of products," *J. Am. Stat. Assoc.*, vol. 55, no. 292, pp. 708–713, Dec. 1960.

[26] J. R. Taylor, *An Introduction To Error Analysis: The Study of Uncertainties in Physical Measurements*. University Science Books, 1997.

[27] P. C. Mahalanobis, "On the generalized distance in statistics," *Proc. Natl. Inst. Sci. India*, vol. 2, no. 1, pp. 49–55, Jan. 1936.

[28] J. Zhao, G. Zhang, Z. Y. Dong, *et al.*, "Robust forecasting aided power system state estimation considering state correlations," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2658–2666, Jul. 2018.

[29] Y. Zou, S. C. Chan, and T. S. Ng, "A recursive least M-estimate (RLM) adaptive filter for robust filtering in impulse noise," *IEEE Signal Process. Lett.*, vol. 7, no. 11, pp. 324–326, Nov. 2000.

[30] C. H. Ho, H. C. Wu, S. C. Chan, *et al.*, "A robust statistical approach to distributed power system state estimation with bad data," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 517–527, Jan. 2020.

[31] C. Huang, C. Thimmisetty, X. Chen, *et al.*, "Power distribution system synchrophasor measurements with non-Gaussian noises: Real-world data testing and analysis," *IEEE Open Access J. Power Energy*, vol. 8, pp. 223–228, May 2021.

[32] K. Sun, M. Huang, Z. Wei, *et al.*, "High-refresh-rate robust state estimation based on recursive correction for large-scale power systems," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–13, May 2023.

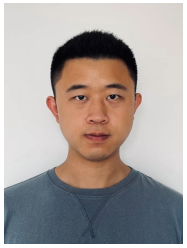
[33] A. S. Dobakhshari, M. Abdolmaleki, V. Terzija, *et al.*, "Robust hybrid linear state estimator utilizing SCADA and PMU measurements," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1264–1273, Mar. 2020.

[34] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[35] K. Sun, "A complex domain Gaussian belief propagation method for fully distributed state estimation: load profiles," Nov. 2023. [Online]. Available: https://github.com/Kang-S/CD-GaBP-load-profiles

[36] F. Milano, "An open source power system analysis toolbox," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1199–1206, Aug. 2005.

[37] "Wind turbines - part 1: Design requirements," *Int. Std. IEC 61400-1*, 2005.



Kang Sun (Student Member, IEEE) received the B.S. degree in electrical engineering and automation from the College of Energy and Electrical Engineering, Hohai University, Nanjing, China, in 2019. He is currently pursuing the Ph.D. degree in electrical engineering with the School of Electrical and Power Engineering, Hohai University.

From September 2023 to September 2024, he was a Visiting Scholar at the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. His current research interests include power system state estimation, cyber-physical systems, and high-performance computing.

Zhinong Wei (Member, IEEE) received the B.S. degree in power system automation from the Hefei University of Technology, Hefei, China, in 1984, the M.S. degree in power system automation from Southeast University, Nanjing, China, in 1987, and the Ph.D. degree in power system automation from Hohai University, Nanjing, China, in 2004.

He is currently a Professor of electrical engineering with the School of Electrical and Power Engineering, Hohai University, Nanjing, China. His research interests include power system state estimation, integrated energy systems, smart distribution systems, optimization and planning, load forecasting, and integration of distributed generation into electric power systems.

Venkata Dinavahi (Fellow, IEEE) received the B.Eng. degree in electrical engineering from Visvesvaraya National Institute of Technology (VNIT), Nagpur, India, the M.Tech. degree in electrical engineering from the Indian Institute of Technology (IIT) Kanpur, India, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Ontario, Canada.

He is currently a Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada. His research interests include real-time simulation of power systems and power electronic systems, electromagnetic transients, device-level modeling, artificial intelligence machine learning, large-scale systems, and parallel and distributed computing.

He is a Fellow of the Engineering Institute of Canada (EIC) and a Fellow of the Asia-Pacific Artificial Intelligence Association (AAIA). He is a Professional Engineer with the Province of Alberta, Canada.

Manyun Huang (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from the College of Energy and Electrical Engineering, Hohai University, Nanjing, China, in 2014 and 2019, respectively.

From September 2017 to September 2018, she was a joint Ph.D. Student with RWTH Aachen University, Aachen, Germany. She is currently an Associate Professor with the School of Electrical and Power Engineering, Hohai University, Nanjing, China. Her current research interests include the theory and algorithms of power system state estimation, robust state estimation, and Kalman filter.

Guoqiang Sun (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Hohai University, Nanjing, China, in 2001, 2005, and 2010, respectively.

He was a Visiting Scholar with North Carolina State University, Raleigh, NC, USA, from 2015 to 2016. He is currently a Professor with the School of Electrical and Power Engineering, Hohai University, Nanjing, China. His research interests include power system analysis and economic dispatch and optimal control of integrated energy systems.