SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •
September 2023, Vol. 66 190206:1–190206:16
https://doi.org/10.1007/s11432-022-3846-1
Special Topic: Control, Optimization, and Learning for Networked Systems

Accuracy analysis for distributed dynamic state estimation in large-scale systems with a cyclic network graph

Mingyan ZHU[1,5], Rui WANG[2,4*], Xiaodong MIAO[3*] & Tianju SUI[1,5]

[1]*School of Control Science and Engineering, Dalian University of Technology, Dalian* 116024, *China;*
[2]*School of Aeronautics and Astronautics, Dalian University of Technology, Dalian* 116024, *China;*
[3]*School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing* 211816, *China;*
[4]*Key Laboratory of Advanced Technology for Aerospace Vehicles, Liaoning Province, Dalian* 116024, *China;*
[5]*Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education, Dalian* 116024, *China*

Received 27 September 2022/Revised 21 March 2023/Accepted 11 August 2023/Published online 28 August 2023

Abstract This paper addresses a distributed dynamic state estimation problem in large-scale systems characterized by a cyclic network graph. The objective is to develop a distributed estimation algorithm for each node to generate local state estimations, based on the coupled measurements and boundary information exchanged with neighboring nodes. Our proposed approach is grounded in the maximum a posteriori (MAP) estimation method, which yields suboptimal results in acyclic network graphs compared with the centralized MAP approach. We extend this approach to systems with a cyclic network graph. Furthermore, we provide an accuracy analysis by deriving bounds for the differences in estimation error covariance and state estimation between the proposed distributed algorithm and the suboptimal centralized MAP method. These bounds apply to a specific category of systems that satisfy certain conditions, including cyclic topology and sparse connections. We demonstrate that these bounds converge asymptotically, with the rate of convergence determined by the loop-free depth of the graph. The loop-free depth of the graph refers to the maximum number of nodes that can be traversed in a cycle without revisiting any node. Finally, we demonstrate the validity of the algorithm through numerical examples.

Keywords distributed state estimation, maximum a posteriori estimation, cyclic network graph

Citation Zhu M Y, Wang R, Miao X D, et al. Accuracy analysis for distributed dynamic state estimation in large-scale systems with a cyclic network graph. Sci China Inf Sci, 2023, 66(9): 190206, https://doi.org/10.1007/s11432-022-3846-1

1 Introduction

Lately, there has been a growing fascination with large-scale systems, especially the networked systems comprising multiple sensors. This interest stems from their potential applications in diverse domains, including environmental monitoring, healthcare, and collaborative information processing [1–4]. Large-scale systems are susceptible to various uncertainties, such as measurement and communication uncertainties, which can adversely affect their stability and control efficiency. State estimation approaches can deal with multiple uncertainties in large-scale systems. State estimation algorithms are employed to obtain system estimations from noisy measurements that are not directly monitored due to various factors. Therefore, state estimation algorithms have gained the attention of researchers in multiple fields [5–8].

State estimation algorithms for large-scale systems can be classified into three main types: centralized, hierarchical, and distributed algorithms. The traditional centralized state estimation algorithm relies on a fusion center and experiences computational and communication burdens that scale with the system's size, making it unsuitable for large-scale systems. In hierarchical state estimation, information is sequentially processed from lower to higher levels, with the top-level nodes integrating fusion information from the

* Corresponding author (email: ruiwang@dlut.edu.cn, mxiaodong@njtech.edu.cn)

lower-level nodes. This approach requires a centralized coordinator to obtain global estimation, leading to high communication burdens. Furthermore, due to the characteristics of information fusion, centralized and hierarchical estimation methods can encounter communication bottlenecks and reliability issues. Consequently, distributed estimation algorithms that impose minimal computational and communication burdens are urgently required [9–12].

The existing distributed state estimation methods can be categorized into two types: static and dynamic. A quasi-static system refers to a system where the state changes smoothly over time with minor changes. Many studies have been conducted on the estimation problem for quasi-static systems [13–15]. Xie et al. [13] introduced a distributed state estimation algorithm where the local state estimation converges asymptotically to the global state estimation. To improve the convergence speed, Pasqualetti et al. [14] presented a distributed static estimation algorithm employing kernel projection, where each node obtains a global state estimate in finite time. Tai et al. [15] introduced a distributed weighted least-squares approach with the property that only a low communication load is required.

However, static estimation is insufficient to capture dynamic changes in a system, leading to the development of distributed dynamic estimation algorithms. One commonly used approach for dynamic estimation approaches is the average consensus strategy [16–18]. Despite its simplicity, this approach has several drawbacks. Firstly, achieving asymptotic convergence with average consensus requires infinite iterations. Secondly, it is challenging to determine a practical stopping criterion. Thirdly, the approach usually provides suboptimal estimations. Another limitation of existing distributed state estimation algorithms is their reliance on a fusion center, making them not fully distributed [19,20]. Furthermore, these algorithms do not consider the effect of communication topology on state estimation [20, 21]. Hence, a requirement exists for a completely distributed state estimation algorithm that offers favorable optimality characteristics while maintaining low computational, communication, and storage complexities.

This study addresses the distributed dynamic state estimation problem in large-scale systems with cyclic network graphs while considering finite-time convergence, which is crucial for such systems. Specifically, we examine the fully distributed algorithm proposed in [22], which only requires local computation and communication. Moreover, this algorithm ensures convergence in a finite number of steps for communication networks consisting of acyclic graphs. However, the algorithm cannot generate optimal results for cyclic network graphs, even though it often converges quickly and provides satisfactory solutions. Based on these observations, we investigate the accuracy of the estimation provided by [22] in large-scale systems with cyclic network graphs. The main contributions of this study are outlined as follows.

(1) This study provides the accuracy analysis for distributed dynamic state estimation algorithm, where the system satisfies certain conditions, including cyclic topology and sparse connections. The research shows that the distributed dynamic estimator achieves high accuracy and convergence.

(2) We demonstrate that the result of the distributed dynamic estimator asymptotically converges to that of the suboptimal centralized MAP method, which is a widely used method for estimating the system state. Furthermore, the convergence rate of the distributed algorithm is dependent on the loop-free depth.

The paper is structured as follows: Section 2 introduces the problem formulation. Section 3 explores the transformation of a given graph into other types, and Section 4 presents the preliminaries. The accuracy analysis for the estimation error covariance is presented in Section 5. Section 6 focuses on accuracy analysis for the state estimation. In Section 7, numerical examples are provided to illustrate the efficacy of the proposed method. Finally, Section 8 concludes the paper.

Notations. $\mathbb{R}^n$ is n-dimensional Euclidean space. $\mathbb{N}$ refers to the set of natural numbers. $\mathcal{N}$ denotes the Gaussian distribution. $\mathbb{E}(\cdot)$ denotes the expectation of random variables. Let $||x||$ denote the Euclidean norm for a matrix or vector x. The transpose operation is represented by x^{T} or x'. $\mathrm{diag}(x_1, \ldots, x_m)$ denotes a diagonal matrix with the elements on its diagonal. Let $\mathbb{S}_+^n$ denote the set of n by n positive definite matrices, and when $x \in \mathbb{S}_+^n$, we have $x > 0$. Let $\dim x$ denote the dimension of the vector x. Let $\sigma_i(x)$ denote the ith singular value of the matrix X, and $\sigma_{\min}(x)$ denote the smallest singular value of the matrix x.

2 Problem formulation

The state-space model of a system with I nodes is represented by

$$x_{k+1}^i = A_i x_k^i + \omega_k^i, \tag{1}$$

where $x_k^i \in \mathbb{R}^{n_i}$ denotes the state of node i at time k, $i \in 1, 2, \ldots, I$. Here, A_i is a known time-invariant real matrix, and ω_k^i represents the process noise with covariance $F_i > 0$. The initial state x_0^i is a zero-mean Gaussian random vector.

Node i can perform self-measurements and joint measurements with neighboring nodes. The self-measurement of node i is given by

$$z_k^i = C_i x_k^i + v_k^i, \tag{2}$$

where z_k^i is the self-measurement of node i, and v_k^i is the measurement noise with covariance $R_i > 0$. The matrix C_i is a known time-invariant real matrix.

Joint measurements between nodes i and j are expressed as

$$z_k^{i,j} = B_{i,j} x_k^i + B_{j,i} x_k^j + v_k^{i,j}, \tag{3}$$

where $z_k^{i,j}$ is the joint measurement between nodes i and j, and $v_k^{i,j}$ is the measurement noise with covariance $R_{i,j} > 0$. The matrices $B_{i,j}$ and $B_{j,i}$ are both known time-invariant real matrices. We assume that the pair (i, j) is unordered, implying that $z_{i,j} = z_{j,i}$ and $v_{i,j} = v_{j,i}$. Therefore, the communication graph composed of nodes is undirected. The notation $z_k^{i,j}$ indicates that node i can communicate with node j, and $j \in \mathcal{N}_i$ represents node j as a neighbor of node i. Additionally, we assume that the noise terms ω_k^i, v_k^i, $v_k^{i,j}$, and x_0^i are uncorrelated.

By individually combining the state and measurement vectors of all subsystems, the following outcome is obtained:

$$x_k = [(x_k^1)^{\mathrm{T}}, (x_k^2)^{\mathrm{T}}, \ldots, (x_k^I)^{\mathrm{T}}]^{\mathrm{T}}, \;\; z_k = [\ldots, (z_k^i)^{\mathrm{T}}, \ldots, (z_k^{i,j})^{\mathrm{T}}, \ldots]^{\mathrm{T}}.$$

Thus, the state and measurement equations for the entire system are represented as

$$x_{k+1} = A x_k + \omega_k, \tag{4}$$

$$z_k = H x_k + v_k, \tag{5}$$

where $A = \mathrm{diag}\{A_1, \ldots, A_I\}$. The mean value of the initial state is $x_0 = [(\bar{x}_0^1)^{\mathrm{T}}, (\bar{x}_0^2)^{\mathrm{T}}, \ldots, (\bar{x}_0^I)^{\mathrm{T}}]^{\mathrm{T}}$ with the variance $\Sigma_0 = \mathrm{diag}\{\Sigma_0^1, \ldots, \Sigma_0^I\}$.

$$H = \begin{bmatrix} & & \cdots & & \\ \cdots & 0 & C_i & 0 & \cdots \\ & & \cdots & & \\ \cdots & 0 & B_{i,j} & 0 & B_{j,i} \; 0 \; \cdots \\ & & \cdots & & \end{bmatrix}$$

is composed of the corresponding measurement matrix of each subsystem. 0 represents a zero matrix with proper dimensions. The covariance of $\omega_k = [(\omega_k^1)^{\mathrm{T}}, \ldots, (\omega_k^I)^{\mathrm{T}}]^{\mathrm{T}}$ and $v_k = [\ldots, (v_k^i)^{\mathrm{T}}, \ldots, (v_k^{i,j})^{\mathrm{T}}, \ldots]^{\mathrm{T}}$ is $F = \mathrm{cov}(\omega_k) = \mathrm{diag}\{F_1, \ldots, F_I\}$ and $R = \mathrm{cov}(v_k) = \mathrm{diag}\{\ldots, R_i, \ldots, R_{i,j}\}$, respectively.

This section explains the combination of state and measurement vectors from individual subsystems to represent the entire system. Eqs. (4) and (5) provide the state and measurement equations for the complete system, with the matrix A being a diagonal matrix composed of the individual subsystem matrices $A_1, \ldots, A_I$. Let $x_0 \sim \mathcal{N}(x_0, \Sigma_0)$ denote the initial state of the system.

This study presents definitions for expressing state estimations and the estimation error covariance of a system. Specifically, we define the predicted estimation of x_k as $\hat{x}_{k|k-1}$, along with the corresponding estimation error covariance $\Sigma_{k|k-1}$. Furthermore, we define $\hat{x}_{k|k}$ as the posterior estimation of x_k, along with the corresponding estimation error covariance $\Sigma_{k|k}$. These values are provided below:

$$\hat{x}_{k|k} \triangleq \mathbb{E}[x_k|Z_k], \;\; \Sigma_{k|k} \triangleq \mathbb{E}[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^{\mathrm{T}}|Z_k],$$
$$\hat{x}_{k|k-1} \triangleq \mathbb{E}[x_k|Z_{k-1}], \;\; \Sigma_{k|k-1} \triangleq \mathbb{E}[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^{\mathrm{T}}|Z_{k-1}],$$

where $Z_k = \{z_0, z_1, \ldots, z_k\}$.

Algorithm 1 Distributed dynamic state estimation algorithm

(1) **Initialization**: For each node $i \in 1, 2, \ldots, I$, $k = 0$,

$$\hat{x}^i_{0|0} = \bar{x}(0), \Sigma^i_{0|0} = \Sigma(0); \tag{6}$$

(2) **Main iteration:** At time $k = 1, 2, \ldots,$ do

(2.1) **Prediction:**

$$\hat{x}^i_{k|k-1} = A_i \hat{x}^i_{k-1|k-1}, \tag{7}$$

$$\Sigma^i_{k|k-1} = A_i \Sigma^i_{k-1|k-1} A_i^{\mathrm{T}} + F_i; \tag{8}$$

(2.2) **Data fusion:** At iteration $n = 1, 2, \ldots,$ for each node i,

(a) Local estimation:

$$\breve{\alpha}^i_k(0) = C_i^{\mathrm{T}} R_i^{-1} Z^i_k + (\Sigma^i_{k|k-1})^{-1} \hat{x}^i_{k|k-1}, \tag{9}$$

$$\breve{Q}^i_k(0) = C_i^{\mathrm{T}} R_i^{-1} C_i + (\Sigma^i_{k|k-1})^{-1}, \tag{10}$$

$$\alpha^{j\to i}_k(0) = 0, \quad Q^{j\to i}_k(0) = 0; \tag{11}$$

(b) Each node i calculates:

$$\alpha^i_k(n) = \breve{\alpha}^i_k(0) + \sum_{j\in\mathcal{N}_i} \alpha^{j\to i}_k(n-1), \tag{12}$$

$$Q^i_k(n) = \breve{Q}^i_k(0) + \sum_{j\in\mathcal{N}_i} Q^{j\to i}_k(n-1), \tag{13}$$

and

$$\hat{x}^i_{k|k}(n) = Q^{-i}_k(n)\alpha^i_k(n), \ \Sigma^i_{k|k}(n) = Q^{-i}_k(n); \tag{14}$$

(c) Information transmission to neighbor nodes $j \in \mathcal{N}_i$:

$$\alpha^{i\to j}_k(n) = B^{\mathrm{T}}_{j,i}(S^{i\to j}_k(n))^{-1} y^{i\to j}_k(n), \tag{15}$$

$$Q^{i\to j}_k(n) = B^{\mathrm{T}}_{j,i}(S^{i\to j}_k(n))^{-1} B_{j,i}, \tag{16}$$

$$y^{i\to j}_k(n) = z^{i,j}_k - B_{i,j}(Q^i_k(n) - Q^{j\to i}_k(n-1))^{-1}(\alpha^i_k(n) - \alpha^{i\to j}_k(n-1)), \tag{17}$$

$$S^{i\to j}_k(n) = R_{i,j} + B_{i,j}(Q^i_k(n) - Q^{j\to i}_k(n-1))^{-1} B^{\mathrm{T}}_{i,j}. \tag{18}$$

We aim to design a distributed dynamic state estimation algorithm that enables each subsystem to estimate its local state. In an acyclic network graph, the distributed maximum a posteriori (MAP) estimation algorithm, as introduced in [22], can minimize the following objective function:

$$J_1(x_k) = (z_k - Hx_k)^{\mathrm{T}} R^{-1}(z_k - Hx_k) + (x_k - \hat{x}_{k|k-1})^{\mathrm{T}} \breve{\Sigma}^{-1}_{k|k-1}(x_k - \hat{x}_{k|k-1}), \tag{19}$$

where $\breve{\Sigma}_{k|k-1} = \mathrm{diag}\{\Sigma^1_{k|k-1}, \ldots, \Sigma^I_{k|k-1}\}$. $\Sigma^i_{k|k-1}$ denotes the estimation error covariance for subsystem i, corresponding to state estimation $\hat{x}^i_{k|k-1}$. According to [22], the result of this distributed approach converges to the estimation as $\hat{x}^S_{k|k} = \arg\min_{x_k} J_1(x_k)$. The modification of the objective function concerning the optimal case, where the error covariance is non-diagonal, results in suboptimal local state estimations for each node. Nevertheless, as reported by [23], this suboptimal solution remains close to the optimal solution.

We apply the distributed MAP approach proposed in [22] to the system with cyclic network graphs, and summarize the method in Algorithm 1. In this algorithm, each node calculates the local state estimation $\hat{x}^i_{k|k}(N)$ and the corresponding error covariance $\Sigma^i_{k|k}(N)$ utilizing the local information vector $\alpha^i_k(N)$ and information matrix $Q^i_k(N)$ at time k and step N. Next, node i constructs the information vector $\alpha^{i\to j}_k(N)$ and the information matrix $Q^{i\to j}_k(N)$, which are transmitted to neighbor j in the current iteration. At the same time, node i removes $\alpha^{j\to i}_k(N-1)$ and $Q^{j\to i}_k(N-1)$ obtained from neighbor j in the previous iteration. In Algorithm 1, each node computes the local state estimation via its local measurements and the messages conveyed by neighbors.

In this study, we establish the following assumptions.

Assumption 1. The communication network among the nodes constitutes a cyclic graph defined by $\mathcal{G}$.

Assumption 2. H is of full column rank.

Assumption 3. For each $i = 1, 2, \ldots, I$, $C_i^{\mathrm{T}} R_i^{-1} C_i > 0$.

It is important to note that communication networks between nodes often exhibit looped graph structures [24, 25]. Assumption 2 assumes that measurements in the system are collectively available, which is generally valid and ensures the nonsingularity of $H'R^{-1}H$ [23]. Furthermore, Assumption 3, along

with $\Sigma(0) > 0$, guarantees the invertibility of $Q_k^i(N) - Q_k^{j\to i}(N-1)$. When $N = 1$, it is evident that $Q_k^i(1) - Q_k^{j\to i}(0) = C_i^{\mathrm{T}} R_i^{-1} C_i + \Sigma(0)$. If this condition is not satisfied, Algorithm 1 cannot yield accurate results.

Ref. [22] has shown that Algorithm 1 converges to local estimations in finite time, assuming the network topology is acyclic. However, a significant challenge arises when Algorithm 1 is applied to cyclic graphs. This study intends to analyze the accuracy of Algorithm 1 when the communication topology is cyclic, i.e., to quantify the mismatch of the estimation error covariance and state estimate between Algorithm 1 and suboptimal centralized MAP algorithms, respectively. In the forthcoming sections, without compromising generalizability, our analysis focuses on evaluating the accuracy of an arbitrary node, indicated by the label "node 1". The results of the accuracy analysis can be extended from node 1 to other nodes.

3 Representation of graphs

The communication network in this study is a cyclic graph $\mathcal{G}$. Due to the difficulty of directly analyzing a cyclic graph, we demonstrate converting the cyclic graph $\mathcal{G}$ into an acyclic graph in Subsection 3.1. We show the equivalence of two graphs using Algorithm 1; i.e., Algorithm 1 yields the same result for node 1 in both graphs. Subsequently, in Subsection 3.2, we convert this acyclic graph into an equivalent line graph whose communication topology can be characterized by a single line.

3.1 Representation of the acyclic graph

This subsection presents the process for converting a cyclic graph $\mathcal{G}$ to an acyclic graph.

The acyclic graph is obtained by converting the cyclic graph $\mathcal{G}$ into a root-tree graph, as illustrated in [26–28]. Since the accuracy for node 1 is investigated, we label any node i in the acyclic graph as node 1 and select it as the root node. The acyclic graph has the property that when Algorithm 1 is executed in the initial cyclic graph $\mathcal{G}$ and the transformed acyclic graph, it produces the same result at node 1. For nodes in the cyclic graph $\mathcal{G}$, let $\mathcal{G}(j)$ denote the node associated with node j in the acyclic graph. Graph $\mathcal{A}$ is the limit of the following iterative process. $\mathcal{A}_0$ is defined as the empty graph. A_1 is defined as the graph of a single node, and this node is associated with node 1 in graph $\mathcal{G}$. After $N \geqslant 2$, the following iterations are performed.

(1) Seek out all the leaf nodes j in the root-tree graph $\mathcal{A}_{N-1}$.

(2) For each leaf node j in $\mathcal{A}_{N-1}$, seek out all neighbors nodes of $\mathcal{G}(j)$ in $\mathcal{G}$, not including the parent node in the $\mathcal{A}_{N-1}$, and add these nodes as the children to this leaf node.

The summary of the leaf node, parent node, and child node can be found in [29]. Graph $\mathcal{A}_N$ represents N iterations of the above procedure. The states and observations for each node in the root-tree graph $\mathcal{A}_N$ are replicated from the corresponding nodes in the cyclic graph $\mathcal{G}$. To illustrate this equivalent conversion, we give an example, as shown in Figure 1. Notice that $1', 1'', 1''', 1\text{`}, 1\text{“}, 1\text{“`}$ all carry the same values.

In the sequel, we associate (1)–(3) of the system with the acyclic graph $\mathcal{A}_N$,

$$\bar{x}_{k+1}^i = \bar{A}_i \bar{x}_k^i + \bar{\omega}_k^i, \tag{20}$$

$$\bar{z}_k^i = \bar{C}_i \bar{x}_k^i + \bar{v}_k^i, \tag{21}$$

$$\bar{z}_k^{i,j} = \bar{B}_{i,j} \bar{x}_k^i + \bar{B}_{j,i} \bar{x}_k^j + \bar{v}_k^{i,j}, \tag{22}$$

for all $i \in \mathcal{A}_N$ and $j \in d_i$ (d_i denotes the child nodes of node i), $\bar{v}_k^i \sim \mathcal{N}(0, \bar{R}_i)$, $\bar{v}_k^{i,j} \sim \mathcal{N}(0, \bar{R}_{i,j})$. By substituting (1)–(3) in the cyclic graph $\mathcal{G}$, we get the values of the associated nodes in the acyclic graph $\mathcal{A}_N$, as follows:

$$\bar{x}_k^i = x_k^{\mathcal{G}(i)},\ \bar{\omega}_k^i = \omega_k^{\mathcal{G}(i)},\ \bar{A}_i = A_{\mathcal{G}(i)},\ \bar{F}_i = F_{\mathcal{G}(i)},\ \bar{v}_k^i = v_k^{\mathcal{G}(i)},\ \bar{C}_i = C_{\mathcal{G}(i)},\ \bar{\Sigma}_{k|k-1}^i = \Sigma_{k|k-1}^{\mathcal{G}(i)},$$
$$\bar{R}_i = R_{\mathcal{G}(i)},\ \bar{z}_k^{i,j} = z_k^{\mathcal{G}(i),\mathcal{G}(j)},\ \bar{v}_k^{i,j} = v_k^{\mathcal{G}(i),\mathcal{G}(j)},\ \bar{B}_{i,j} = B_{\mathcal{G}(i),\mathcal{G}(j)},\ \bar{R}_{i,j} = R_{\mathcal{G}(i),\mathcal{G}(j)},\ \bar{z}_k^i = z_k^{\mathcal{G}(i)}.$$

Remark 1. $\mathcal{G}$ and $\mathcal{A}_N$ are equivalent graphs in terms of Algorithm 1; i.e., Algorithm 1 yields the same result at node 1 in both graphs.

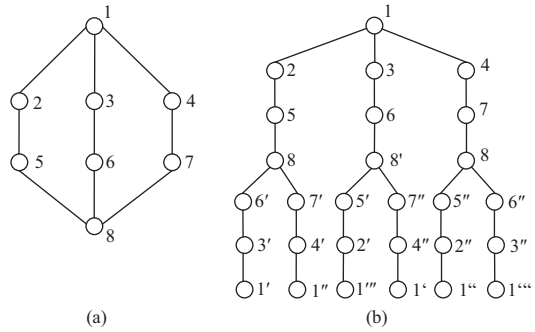

Figure 1 (a) The cyclic graph and (b) the acyclic graph with seven layers ($N = 7$).

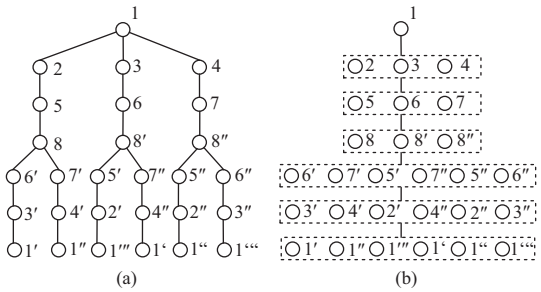

Figure 2 (a) The acyclic graph and (b) the line graph.

3.2 Representation of the line graph

The previous subsection shows that $\mathcal{A}_N$ is an acyclic graph with N layers, and its root node is 1. This subsection presents the equivalent conversion of $\mathcal{A}_N$ to the line graph $\mathcal{L}_N$ from the perspective of Algorithm 1, wherein both graphs produce identical results at node 1. The construction of $\mathcal{L}_N$ is a straightforward process. For all $n = 1, 2, \ldots, N$, the nodes in $\mathcal{A}_N$ that are located at a distance of $n-1$ steps from node 1 are grouped into a single node defined as τ_n. To clarify this transformation, we provide an example illustrated in Figure 2. Note that the nodes enclosed in the dashed box are merged into a single large node.

Then the key is to associate (20)–(22) of the system with the line graph $\mathcal{L}_N$, which we explain in detail subsequently. Define the size of any finite set τ as $|\tau|$ whose elements are $\tau(1), \tau(2), \ldots, \tau(|\tau|)$. For $n \in \mathbb{N}$, the state of τ_n at time k is provided by

$$\tilde{x}_k^n = [(\bar{x}_k^{\tau_n(1)})^{\mathrm{T}}, (\bar{x}_k^{\tau_n(2)})^{\mathrm{T}}, \ldots, (\bar{x}_k^{\tau_n(|\tau_n|)})^{\mathrm{T}}]^{\mathrm{T}}. \tag{23}$$

Also, the equations of τ_n are presented by

$$\tilde{x}_{k+1}^n = \tilde{A}_n \tilde{x}_k^n + \tilde{\omega}_k^n, \tag{24}$$

$$\tilde{z}_k^n = \tilde{C}_n \tilde{x}_k^n + \tilde{v}_k^n, \tag{25}$$

$$\tilde{z}_k^{n,n+1} = \tilde{B}_{n,n+1} \tilde{x}_k^n + \tilde{B}_{n+1,n} \tilde{x}_k^{n+1} + \tilde{v}_k^{n,n+1}. \tag{26}$$

That is, $\tilde{z}_k^n$ is composed of the measurements $\bar{z}_k^i$ in the acyclic graph, with $i \in \tau_n$. The elements of $\tilde{z}_k^{n,n+1}$ are obtained by aggregating the measurements $\bar{z}_k^{i,j}$ corresponding to nodes $i \in \tau_n$ and $j \in \tau_{n+1}$ in the acyclic graph. Notice that $\tilde{v}_k^n \sim N(0, \tilde{R}_n)$ and $\tilde{v}_k^{n,n+1} \sim N(0, \tilde{R}_{n,n+1})$. The parameters $\tilde{C}_n, \tilde{R}_n, \tilde{B}_{n,n+1}$ and $\tilde{R}_{n,n+1}$ are all associated with $\bar{C}_i, \bar{R}_i, \bar{B}_{i,j}$ and $\bar{R}_{i,j}$. Then, we have

$$\begin{aligned}
&\tilde{A}_n = \mathrm{diag}\{\bar{A}_{\tau_n(1)}, \bar{A}_{\tau_n(2)}, \ldots, \bar{A}_{\tau_n(|\tau_n|)}\},\ \tilde{F}_n = \mathrm{diag}\{\bar{F}_{\tau_n(1)},\ \bar{F}_{\tau_n(1)}, \ldots, \bar{F}_{\tau_n(|\tau_n|)}\},\\
&\tilde{z}_k^n = [(\bar{z}_k^{\tau_n(1)})^{\mathrm{T}}, (\bar{z}_k^{\tau_n(2)})^{\mathrm{T}}, \ldots, (\bar{z}_k^{\tau_n(|\tau_n|)})^{\mathrm{T}}]^{\mathrm{T}},\ \tilde{C}_n = \mathrm{diag}\{\bar{C}_{\tau_n(1)}, \bar{C}_{\tau_n(2)}, \ldots, \bar{C}_{\tau_n(|\tau_n|)}\},\\
&\tilde{R}_n = \mathrm{diag}\{\bar{R}_{\tau_n(1)}, \bar{R}_{\tau_n(2)}, \ldots, \bar{R}_{\tau_n(|\tau_n|)}\},\ \tilde{\Sigma}_{k|k-1}^n = \mathrm{diag}\{\bar{\Sigma}_{k|k-1}^{\tau_n(1)}, \bar{\Sigma}_{k|k-1}^{\tau_n(2)}, \ldots, \bar{\Sigma}_{k|k-1}^{\tau_n(|\tau_n|)}\}.
\end{aligned}$$

In addition, we get

$$\begin{aligned}
&\tilde{z}_k^{n,n+1} = [(\grave{z}_k^{\tau_n(1)})^{\mathrm{T}}, (\grave{z}_k^{\tau_n(2)})^{\mathrm{T}}, \ldots, (\grave{z}_k^{\tau_n(|\tau_n|)})^{\mathrm{T}}]^{\mathrm{T}},\ \tilde{B}_{n,n+1} = \mathrm{diag}\{\grave{B}_{\tau_n(1)}, \grave{B}_{\tau_n(2)}, \ldots, \grave{B}_{\tau_n(|\tau_n|)}\},\\
&\tilde{B}_{n+1,n} = \mathrm{diag}\{\acute{B}_{\tau_n(1)}, \acute{B}_{\tau_n(2)}, \ldots, \acute{B}_{\tau_n(|\tau_n|)}\},\ \tilde{R}_{n,n+1} = \mathrm{diag}\{\grave{R}_{\tau_n(1)}, \grave{R}_{\tau_n(2)}, \ldots, \grave{R}_{\tau_n(|\tau_n|)}\},
\end{aligned}$$

with

$$\begin{aligned}
&\grave{z}_k^i = [(\bar{z}_k^{i,d_i(1)})^{\mathrm{T}}, (\bar{z}_k^{i,d_i(2)})^{\mathrm{T}}, \ldots, (\bar{z}_k^{i,d_i(|d_i|)})^{\mathrm{T}}]^{\mathrm{T}},\ \grave{B}_i = [\bar{B}_{i,d_i(1)}^{\mathrm{T}}, \bar{B}_{i,d_i(2)}^{\mathrm{T}}, \ldots, \bar{B}_{i,d_i(|d_i|)}^{\mathrm{T}}]^{\mathrm{T}},\\
&\acute{B}_i = \mathrm{diag}\{\bar{B}_{d_i(1),1}^{\mathrm{T}}, \bar{B}_{d_i(1),2}^{\mathrm{T}}, \ldots, \bar{B}_{d_i(|d_i|),i}^{\mathrm{T}}\},\ \grave{R}_i = \mathrm{diag}\{\bar{R}_{i,d_i(1)}, \bar{R}_{i,d_i(2)}, \ldots, \bar{R}_{i,d_i(|d_i|)}\}.
\end{aligned}$$

Remark 2. Notably, Algorithm 1 is applicable to both graphs and produces the same results at node 1. Thus, the dynamic estimation problem at node 1 in the cyclic graph, addressed by Algorithm 1, can be simplified to the estimation problem in the linear graph.

4 Preliminaries

In this section, we present the concept of Riemann distance for positive definite matrices. In addition, we provide several related properties of this metric.

Definition 1 ([30]). The Riemann distance between two $n \times n$ positive definite matrices B and D can be represented as

$$\delta(B, D) = \sqrt{\sum_{i=1}^{n} \log^2 \sigma_i(BD^{-1})},$$

where $\sigma_1(BD^{-1}) \geqslant \cdots \geqslant \sigma_n(BD^{-1})$.

In [24], several properties of the Riemann distance for positive definite matrices are presented.

Proposition 1 ([24]). For any two $n \times n$ matrices $B > 0$ and $D > 0$, it follows that

(1) $\delta(B^{-1}, D^{-1}) = \delta(D, B) = \delta(B, D)$.

(2) $\delta(FBF^{\mathrm{T}}, FDF^{\mathrm{T}}) \leqslant \delta(B, D)$ if F is a matrix with full row rank. Moreover, if F is invertible, the equality holds.

(3) If $B \geqslant D$ and $E \geqslant 0$, then $\delta(B + E, D) \geqslant \delta(B, D)$.

(4) For any positive matrix $E \in \mathbb{R}^{m \times m}$, and matrix $F \in \mathbb{R}^{m \times n}$, we obtain

$$\delta(E + FB^{-1}F^{\mathrm{T}}, E + FD^{-1}F^{\mathrm{T}}) \leqslant \frac{\alpha}{\alpha + \beta}\delta(B, D),$$

where $\alpha = \max\{||FB^{-1}F^{\mathrm{T}}||, ||FD^{-1}F^{\mathrm{T}}||\}$ and $\beta = \sigma_{\min}(E)$.

(5) $||B - D|| \leqslant (\mathrm{e}^{\delta(B,D)} - 1)\min\{||B||, ||D||\}$.

Here, we present the notations necessary for this study.

Notation 1. In the given graph $\mathcal{C}$, we define the information vector $\alpha_k^i(\mathcal{C})$, information matrix $Q_k^i(\mathcal{C})$, and state estimation $\hat{x}_k^i(\mathcal{C})$ for node i at time k, which are obtained by executing Algorithm 1 on $\mathcal{C}$.

Notation 2. Let $\{O_k = Q_k^{j \to i}(0) : i = 1, \ldots, I, j \in \mathcal{N}_i\}$ denote the information received by neighbor node j, for each i at time k and step 0. In the cyclic network graph $\mathcal{G}$, we execute Algorithm 1 and obtain the state estimation and information matrix for node 1 at time k and step N, denoted by $\hat{x}_k^1(N, O_k)$ and $Q_k^1(N, O_k)$, respectively.

Notation 3. In a graph, a path is a sequence of edges that connect two vertices, where each edge connects two adjacent vertices in the sequence. The length of a path is the number of edges in the sequence. A circle is a path that starts and ends at the same vertex. The distance between two vertices is the length of the shortest path connecting them. We define $\mathcal{S}_1(f_1)$ as the subgraph of $\mathcal{G}$ consisting of nodes that are at most a distance of f_1 away from node 1. The loop-free depth f_1 is the largest value such that the subgraph $\mathcal{S}_1(f_1)$ is acyclic, meaning that there are no cycles in the subgraph.

Notation 4. $\bar{e} \triangleq \max_i |\mathcal{N}_i| - 1, \bar{n} \triangleq \max_i \dim x_i, \bar{m} \triangleq \max\{\max_i \dim z_k^i, \max_{i,j} \dim z_k^{i,j}\}$.

5 Accuracy analysis for the estimation error covariance

This section is dedicated to accuracy analysis for the estimation error covariance at node 1. Our primary contribution is outlined in Subsection 5.1, where we present the findings concerning the accuracy of the estimation error covariance. Subsequently, we provide the proof for this result in Subsection 5.2.

5.1 Main outcome

Theorem 1. Let $\mathrm{Cov}_1^M(k)$ be the error covariance acquired by node 1 at time k using the suboptimal centralized MAP method on $\mathcal{G}$. If $\varphi < 1$, it can be established that a constant $\bar{\theta}$ exists, which depends on various parameters including $\bar{e}$, $\bar{n}$, C_i, $B_{i,j}$, R_i, and $R_{i,j}$. Then for any $N \geqslant f_1 + 1$,

$$||\mathrm{Cov}_1^M(k) - (Q_k^1(N))^{-1}|| \leqslant \bar{\theta}\varphi^{f_1}, \tag{27}$$

with

$$\varphi = \lambda\sqrt{\bar{e}},\ \lambda = \frac{\mu_1}{\mu_1 + \upsilon_1}\frac{\mu_2}{\mu_2 + \upsilon_2},\ \mu_1 = \bar{e}\max_{i,j}||B_{i,j}^{\mathrm{T}}R_{i,j}^{-1}B_{i,j}||,$$

$$\upsilon_1 = \max_i \sigma_{\min}(C_i^{\mathrm{T}} R_i^{-1} C_i),\ \mu_2 = \max_{i,j} ||B_{i,j}(C_i^{\mathrm{T}} R_i^{-1} C_i) B_{i,j}^{\mathrm{T}}||,\ \upsilon_2 = \max_{i,j} \sigma_{\min}(R_{i,j}).$$

Remark 3. If $\varphi < 1$, the error covariance acquired by node 1 using Algorithm 1 converges exponentially to that obtained by the suboptimal centralized MAP method. The loop-free depth f_1 of node 1 determines the convergence rate.

Remark 4. Since $\varphi = \lambda\sqrt{\bar{e}}$, the result is applicable to a category of graphs characterized by sparse connections (small $\bar{e}$) and the small signal-to-noise ratio of both local measurements and joint measurements (small λ).

5.2 Proof of Theorem 1

The equivalence between $Q_k^1(N)$, $Q_k^1(\mathcal{A}_N)$, and $Q_k^1(\mathcal{L}_N)$ can be established through the conversions of the graph detailed in Section 3. To prove Theorem 1, it is essential to obtain the following lemmas.

Lemma 1. For any $N \in \mathbb{N}$, let O_k^1 and O_k^2 be the sets of initial values of $Q_k^{j\to i}(0)$ that satisfy the condition $0 \leqslant Q_k^{j\to i}(0) \leqslant B_{i,j}^{\mathrm{T}} R_{i,j}^{-1} B_{i,j}$ for all $Q_k^{j\to i}(0) \in O_k^a$ and $a \in 1, 2$, i.e., the message sent from node j to i at time k as an initial transmission. Then, it holds that

$$\delta(Q_k^1(N, O_k^1) - Q_k^1(N, O_k^2)) \leqslant \bar{\zeta}\varphi^{N-1}, \tag{28}$$

where $\bar{\zeta} = \sqrt{\bar{n}(\bar{e}+1)} \max_i \log ||I + (\sum_{j\in N_i} B_{i,j}^{\mathrm{T}} R_{i,j}^{-1} B_{i,j})(C_i^{\mathrm{T}} R_i^{-1} C_i)^{-1}||$.

Proof. See Appendix A.

Notation 5. Let $O_k^{\max} = \{Q_k^{j\to i}(0) = B_{i,j}^{\mathrm{T}} R_{i,j}^{-1} B_{i,j} : i = 1, \ldots, I \text{ and } j \in N_i\}$, $O_k^{\min} = \{Q_k^{j\to i}(0) = 0 : i = 1, \ldots, I \text{ and } j \in N_i\}$.

Similar to Lemma 16 in [24], we present the following result.

Lemma 2 ([24])**.** For any $N \geqslant f_1 + 1$ (f_1 is defined in Notation 3), we have

$$||\mathrm{Cov}_1^M(k) - (Q_k^1(N))^{-1}|| \leqslant ||(Q_k^1(f_1+1, O_k^{\max}))^{-1} - (Q_k^1(f_1+1, O_k^{\min}))^{-1}||.$$

The proof of Theorem 1 relies on Lemmas 1 and 2. These lemmas facilitate the derivation of an upper bound on the difference between $\mathrm{Cov}_1^M(k)$ and $(Q_k^1(N))^{-1}$.

Proof of Theorem 1. By satisfying the condition of Lemma 2 with $O_k^{\max}$ and $O_k^{\min}$, we obtain

$$||\mathrm{Cov}_1^M(k) - (Q_k^1(N))^{-1}|| \leqslant ||(Q_k^1(f_1+1, O_k^{\max}))^{-1} - (Q_k^1(f_1+1, O_k^{\min}))^{-1}||.$$

According to Lemma 1, $\delta(Q_k^1(f_1+1, O_k^{\max}) - Q_k^1(f_1+1, O_k^{\min})) \leqslant \bar{\zeta}\varphi^{f_1}$. From Proposition 1,

$$\begin{aligned} ||(Q_k^1(f_1, O_k^{\max}))^{-1} - (Q_k^1(f_1, O_k^{\min}))^{-1}|| &\leqslant (\mathrm{e}^{\bar{\zeta}\varphi^{f_1}} - 1)||(Q_k^1(f_1, O_k^{\min}))^{-1}|| \\ &\leqslant (\mathrm{e}^{\bar{\zeta}\varphi^{f_1}} - 1)||(C_1^{\mathrm{T}} R_1^{-1} C_1)^{-1}|| \leqslant \upsilon_1(\mathrm{e}^{\bar{\zeta}} - 1)\varphi^{f_1}. \end{aligned}$$

Since $\upsilon_1(\mathrm{e}^{\bar{\zeta}} - 1)$ only depends on $\bar{e}, \bar{n}, C_i, B_{i,j}, R_i, R_{i,j}$, inequality (27) holds.

6 Accuracy analysis for the state estimation

In this section, we present the accuracy analysis for state estimation at node 1. The main result is provided in Subsection 6.1, while the proof of the result is given in Subsection 6.2.

6.1 Main outcome

Theorem 2. We apply the suboptimal centralized MAP algorithm on $\mathcal{G}$ to obtain the state estimation $\hat{x}_1^M(k)$ for node 1 at time k. If $\gamma < 1$, it can be established that a constant $\bar{b}$ exists, which depends on various parameters, including $\bar{e}$, $\bar{n}$, $B_{i,j}$, C_i, $R_{i,j}$, R_i, F_i, and $\bar{z}$. For any $N \geqslant f_1 + 1$, the following inequality holds:

$$||\hat{x}_{k|k}^1(N) - \hat{x}_1^M(k)|| \leqslant \bar{b}\gamma^{f_1+1}, \tag{29}$$

where $\gamma = \max\{\bar{e}\sqrt{\phi}, \sqrt{\bar{e}}\iota^{1/\varsigma}\}$, $\phi = \frac{\alpha_1}{\alpha_1+\beta_1}\frac{\alpha_2}{\alpha_2+\beta_2}$, $\alpha_1 = \underline{r}^{-1}\bar{e}\max_{i,j}||C_{i,j}||^2$, $\alpha_2 = \max_{i,j}||C_{i,j}||^2\frac{\bar{e}\bar{r}}{\underline{\varepsilon}^2}$, $\beta_1 = \bar{r}^{-1}\underline{\varepsilon}^2, \beta_2 = \underline{r}, \iota = \frac{\sqrt{\bar{q}}-\sqrt{\underline{q}}}{\sqrt{\bar{q}}+\sqrt{\underline{q}}}, \bar{q} = \bar{\varepsilon}^2\underline{r}^{-1} + \bar{\Sigma}, \underline{q} = \underline{\varepsilon}^2\bar{r}^{-1}, \varsigma = 2 + \log_{\frac{1}{\sqrt{\phi}}}(\bar{q}/\underline{q}), \bar{r} = \max_i\{||R_i||, ||R_{i,j}||\}, \bar{\Sigma} = \max_{i\leqslant I}||F_i^{-1}||, \underline{r} = \min_i\{\sigma_{\min}(R_i), \sigma_{\min}(R_{i,j})\}, \bar{\varepsilon} = \max_{i,j}\sqrt{||C_i||^2 + 4\bar{e}||B_{i,j}||^2}, \underline{\varepsilon} = \min_i\sigma_{\min}(C_i), \bar{z} = \max_{i,j}\{||z_k^i||_\infty, ||z_k^{i,j}||_\infty\}$.

Remark 5. If the condition $\gamma < 1$ is satisfied, the state estimation obtained by node 1 utilizing Algorithm 1 converges exponentially to that acquired by the suboptimal centralized MAP method. The convergence rate relies on the loop-free depth f_1 of node 1.

Remark 6. Since γ is defined as $\{\bar{e}\sqrt{\phi}, \sqrt{\bar{e}}\iota^{1/\varsigma}\}$, the outcome is applicable to a category of graphs that exhibit sparse connections (small $\bar{e}$), small ϕ, and small ι. Moreover, the magnitude of ϕ is affected by the signal-to-noise ratio of both local measurements and joint measurements. The characteristics of the system parameters affect the magnitude of ι.

6.2 Proof of Theorem 2

The proof of Theorem 2 is divided into three parts. Firstly, we derive a bound for $\hat{x}^1_{k|k}(\mathcal{L}_{N+1}) - \hat{x}^1_{k|k}(\mathcal{L}_N)$ in the line graph. Secondly, we extend this outcome to an arbitrary graph. Finally, we provide an accuracy analysis.

6.2.1 *Bound of the augmentation in the line graph*

An analysis is conducted on the line graph $\mathcal{L}_N$ using (24)–(26). Consider

$$y_k^i = [(\tilde{z}_k^i)^{\mathrm{T}}, (\tilde{z}_k^{i,i+1})^{\mathrm{T}}]^{\mathrm{T}},\ \Psi_k^i = [(\tilde{v}_k^i)^{\mathrm{T}}, (\tilde{v}_k^{i,i+1})^{\mathrm{T}}]^{\mathrm{T}},$$
$$G_{i,i} = \begin{bmatrix} \tilde{C}_i \\ \tilde{B}_{i,i+1} \end{bmatrix},\ G_{i,i+1} = \begin{bmatrix} 0 \\ \tilde{B}_{i+1,i} \end{bmatrix},\ S_i = \begin{bmatrix} \tilde{R}_i & 0 \\ 0 & \tilde{R}_{i,i+1} \end{bmatrix},$$

for $i = 1, 2, \ldots, N-1$, and $y_k^N = \tilde{z}_k^N, \Psi_k^N = \tilde{v}_k^N, G_{N,N} = \tilde{C}_N, S_N = \tilde{R}_N$. Eqs. (24)–(26) become $\tilde{x}_{k+1}^i = \tilde{A}_i\tilde{x}_k^i + \tilde{\omega}_k^i, y_k^i = G_{i,i}\tilde{x}_k^i + G_{i,i+1}\tilde{x}_k^{i+1} + \Psi_k^i$ with $\Psi_k^i \sim \mathcal{N}(0, S_i)$. The estimate of $\tilde{x}_{k|k-1}^i$ is defined as $\hat{\tilde{x}}_{k|k-1}^i$. We also define $\boldsymbol{A}_N = \mathrm{diag}\{\tilde{A}_1, \ldots, \tilde{A}_N\}$, $\varpi_k^N = [(\tilde{\omega}_k^1)^{\mathrm{T}}, \ldots, (\tilde{\omega}_k^N)^{\mathrm{T}}], \boldsymbol{x}_k^N = [(\tilde{x}_k^1)^{\mathrm{T}}, \ldots, (\tilde{x}_k^N)^{\mathrm{T}}]^{\mathrm{T}}, \boldsymbol{y}_k^N = [(y_k^1)^{\mathrm{T}}, \ldots, (y_k^N)^{\mathrm{T}}]$, $\boldsymbol{\Psi}_k^N = [(\Psi_k^1)^{\mathrm{T}}, \ldots, (\Psi_k^N)^{\mathrm{T}}]$,

$$[\boldsymbol{G}_N]_{i,j} = \begin{cases} G_{i,j}, & 0 \leqslant j - i \leqslant 1, \\ 0, & \text{otherwise.} \end{cases}$$

Subsequently, we have

$$\boldsymbol{x}_{k+1}^N = \boldsymbol{A}_N\boldsymbol{x}_k^N + \varpi_k^N, \tag{30}$$
$$\boldsymbol{y}_k^N = \boldsymbol{G}_N\boldsymbol{x}_k^N + \boldsymbol{\Psi}_k^N, \tag{31}$$

where $\varpi_k^N \sim \mathcal{N}(0, T_N)$, $T_N = \mathrm{diag}[\tilde{F}_1, \ldots, \tilde{F}_N]$, $\boldsymbol{\Psi}_k^N \sim \mathcal{N}(0, \boldsymbol{S}_N)$, and $\boldsymbol{S}_N = \mathrm{diag}[S_1, \ldots, S_N]$. Moreover, $\check{\Sigma}_{k|k-1}^N = \mathrm{diag}\{\tilde{\Sigma}_{k|k-1}^1, \ldots, \tilde{\Sigma}_{k|k-1}^N\}$.

The suboptimal centralized MAP estimation $\hat{\boldsymbol{x}}_{k|k}^N$ of $\boldsymbol{x}_k^N$ is expressed as

$$\hat{\boldsymbol{x}}_{k|k}^N = (\boldsymbol{Q}_k^N)^{-1}\boldsymbol{q}_k^N,$$

where $\boldsymbol{q}_k^N = (\boldsymbol{G}_N)^{\mathrm{T}}\boldsymbol{S}_N^{-1}\boldsymbol{y}_k^N + (\check{\Sigma}_{k|k-1}^N)^{-1}\hat{\boldsymbol{x}}_{k|k-1}^N = [(q_k^1)^{\mathrm{T}}, \ldots, (q_k^N)^{\mathrm{T}}]^{\mathrm{T}}$ with

$$q_k^i = \begin{cases} G_{i,i}^{\mathrm{T}}S_i^{-1}y_k^i + (\tilde{\Sigma}_{k|k-1}^i)^{-1}\hat{\tilde{x}}_{k|k-1}^i, & i = 1, \\ G_{i,i}^{\mathrm{T}}S_i^{-1}y_k^i + G_{i-1,i}^{\mathrm{T}}S_{i-1}^{-1}y_k^{i-1} + (\tilde{\Sigma}_{k|k-1}^i)^{-1}\hat{\tilde{x}}_{k|k-1}^i, & i > 1. \end{cases} \tag{32}$$

The (i,j)-th entry of $\boldsymbol{Q}_k^N = (\boldsymbol{G}_N)^{\mathrm{T}}\boldsymbol{S}_N^{-1}\boldsymbol{G}_N + (\check{\Sigma}_{k|k-1}^N)^{-1}$ is provided by

$$\begin{aligned} Q_k^{i,i} &= \begin{cases} G_{i,i}^{\mathrm{T}}S_i^{-1}G_{i,i} + (\tilde{\Sigma}_{k|k-1}^i)^{-1}, & i = 1, \\ G_{i,i}^{\mathrm{T}}S_i^{-1}G_{i,i} + G_{i-1,i}^{\mathrm{T}}S_{i-1}^{-1}G_{i-1,i} + (\tilde{\Sigma}_{k|k-1}^i)^{-1}, & i > 1, \end{cases} \\ Q_k^{i,i+1} &= G_{i,i}^{\mathrm{T}}S_i^{-1}G_{i,i+1}, Q_k^{i+1,i} = (Q_k^{i,i+1})^{\mathrm{T}}, \\ Q_k^{i,j} &= 0,\ |i-j| \geqslant 2. \end{aligned} \tag{33}$$

Define $P_k^N = (\boldsymbol{Q}_k^N)^{-1}$ and $[P_k^N]^{i,j}$ as the (i,j)-th block. Applying the inverse equation from Theorem 3.1 in [31], we get the first block row of P_k^N as follows:

$$[P_k^N]^{1,j} = \left(\prod_{l=1}^{j-1} (\Delta_k^l)^{-1} Q_k^{l,l+1}\right) \Phi_k^j(N) \tag{34}$$

with

$$\Phi_k^j(N) = \Gamma_k^j(N) - Q_k^{j,j-1}(\Delta_k^{j-1})^{-1} Q_k^{j-1,j}, \tag{35}$$

$$\Gamma_k^l(N) = \begin{cases} Q_k^{l,l}, & l = N, \\ Q_k^{l,l} - Q_k^{l,l+1}(\Gamma_k^{l+1}(N))^{-1} Q_k^{l+1,l}, & l < N, \end{cases}$$
$$\Delta_k^l = \begin{cases} Q_k^{l,l}, & l = 1, \\ Q_k^{l,l} - Q_k^{l,l-1}(\Delta_k^{l-1})^{-1} Q_k^{l-1,l}, & l > 1, \end{cases}$$

for $j = 1, 2, \ldots, N$. The first entry $[\hat{\boldsymbol{x}}_{k|k}^N]_1$ can be given by

$$[\hat{\boldsymbol{x}}_{k|k}^N]_1 = \sum_{j=1}^{N} [P_k^N]^{1,j} q_k^j. \tag{36}$$

According to [22], we have $\hat{x}_{k|k}^1(\mathcal{L}_N) = [\hat{x}_{k|k}^N]_1$, representing the estimation at node 1. Using (36), we find

$$||\hat{x}_{k|k}^1(\mathcal{L}_{N+1}) - \hat{x}_{k|k}^1(\mathcal{L}_N)|| \leqslant \sum_{j=1}^{N} ||[P_k^{N+1}]^{1,j} - [P_k^N]^{1,j}||||q_k^j|| + ||[P_k^{N+1}]^{1,N+1}||||q_k^{N+1}||. \tag{37}$$

To establish the bound for $||\hat{x}_{k|k}^1(\mathcal{L}_{N+1}) - \hat{x}_{k|k}^1(\mathcal{L}_N)||$, we introduce a result in Lemma 6. The proof of this lemma depends on various intermediate results. We begin by providing bounds for certain quantities, i.e., $(\check{\Sigma}_{k|k-1}^N)^{-1}, \hat{\boldsymbol{x}}_{k|k-1}^N, \boldsymbol{G}_N, \boldsymbol{Q}_k^N, \Delta_k^l, \Gamma_k^l(N)$. The bounds for these quantities are provided below.

Lemma 3. For any $k = 1, 2, \ldots$, we can derive the following inequality:

$$||(\check{\Sigma}_{k|k-1}^N)^{-1}|| \leqslant \tilde{\bar{\Sigma}}, \tag{38}$$

where $\tilde{\bar{\Sigma}} = \max_{i \leqslant N} ||\tilde{F}_i^{-1}||$.

Proof. See Appendix B.

Lemma 4. For any $k = 1, 2, \ldots$, and $N \in \mathbb{N}$, there exists a positive scalar $\tilde{\bar{X}}$ such that

$$||\hat{\boldsymbol{x}}_{k|k-1}^N|| \leqslant \tilde{\bar{X}}. \tag{39}$$

Proof. See Appendix C.

Lemma 5 ([24])**.** For any $k = 1, 2, \ldots$, $N \in \mathbb{N}$, $1 \leqslant l \leqslant N$, and $1 \leqslant j \leqslant N$, we have

$$\underline{\tilde{\varepsilon}} I \leqslant \boldsymbol{G}_N \leqslant \tilde{\bar{\varepsilon}} I, \tag{40}$$
$$\underline{\tilde{q}} \leqslant ||\boldsymbol{Q}_k^N||, ||\Delta_k^l||, ||\Gamma_k^l(N)|| \leqslant \tilde{\bar{q}}, \tag{41}$$
$$||[P_k^N]^{1,j}|| \leqslant \tilde{c}\tilde{\iota}^j, \ \max_{n \leqslant N} ||q_n|| \leqslant \tilde{\eta}_N, \tag{42}$$
$$||[P_k^{N+1}]^{1,j} - [P_k^N]^{1,j}|| \leqslant \tilde{\bar{q}}^{-1} \tilde{r}^j (\mathrm{e}^{\tilde{\psi}_N \tilde{\lambda}_N^{N-j}} - 1), \tag{43}$$

with $\tilde{\bar{\varepsilon}} = \max_i(||\tilde{C}_i||^2 + 2\max\{||\tilde{B}_{i-1,i}||^2, ||\tilde{B}_{i,i-1}||^2\} + 2\max\{||\tilde{B}_{i,i+1}||^2, ||\tilde{B}_{i+1,i}||^2\})^{1/2}, \underline{\tilde{\varepsilon}} = \min_i \sigma_{\min}(\tilde{C}_i)$. $\underline{\tilde{q}} = \frac{\underline{\tilde{\varepsilon}}^2}{\tilde{\bar{r}}}, \tilde{\bar{r}} = \max_i\{||\tilde{R}_i||, ||\tilde{R}_{i,i+1}||\}, \tilde{\bar{q}} = \frac{\tilde{\bar{\varepsilon}}^2}{\underline{\tilde{r}}} + \tilde{\bar{\Sigma}}, \underline{\tilde{r}} = \min_i\{\sigma_{\min}(\tilde{R}_i), \sigma_{\min}(\tilde{R}_{i,i+1})\}, i = 1, 2, \ldots, N$. $\tilde{\psi}_N = \sqrt{\bar{n}|\tau_N|}\tilde{\bar{\xi}}_N$, $\tilde{\bar{\xi}}_N = \max_{i \leqslant N} \log \sigma_{\max}[I + (\tilde{B}_{i,i+1}^{\mathrm{T}} \tilde{R}_{i,i+1}^{-1} \tilde{B}_{i,i+1})(\tilde{C}_i^{\mathrm{T}} \tilde{R}_i^{-1} \tilde{C}_i)^{-1}]$, $\tilde{c} = \frac{\tilde{r}-1}{2\tilde{q}\tilde{\iota}}$, $\tilde{\iota} = \frac{\sqrt{\tilde{r}}-1}{\sqrt{\tilde{r}}+1}, \tilde{\lambda}_N =$

$\frac{\tilde{\mu}_{1,N}}{\tilde{\mu}_{1,N}+\tilde{\upsilon}_{1,N}}\frac{\tilde{\mu}_{2,N}}{\tilde{\mu}_{2,N}+\tilde{\upsilon}_{2,N}}, \tilde{\eta}_N = \max_{i\leqslant N}\frac{2^{3/2}\tilde{\bar{\varepsilon}}\sqrt{\bar{m}|\tau_i|}\tilde{\bar{z}}_N}{\underline{\tilde{r}}} + \tilde{\bar{\Sigma}}\tilde{\bar{X}}, \tilde{\bar{z}}_N = \max_{k\in\mathbb{N},i\leqslant N}\{||\tilde{z}_k^i||_\infty, ||\tilde{z}_k^{i,i+1}||_\infty\}, \tilde{r} = \tilde{\bar{q}}/\underline{\tilde{q}}$, where

$$\tilde{\mu}_{1,N} = \max_{i\leqslant N}||\tilde{B}_{i,i+1}^{\mathrm{T}}\tilde{R}_{i,i+1}^{-1}\tilde{B}_{i,i+1}||,\ \tilde{\mu}_{2,N} = \max_{i\leqslant N}||\tilde{B}_{i+1,i}(\tilde{C}_{i+1}^{\mathrm{T}}\tilde{R}_{i+1}^{-1}\tilde{C}_{i+1})^{-1}\tilde{B}_{i+1,i}^{\mathrm{T}}||,$$
$$\tilde{\upsilon}_{1,N} = \min_{i\leqslant N}\sigma_{\min}(\tilde{C}_i^{\mathrm{T}}\tilde{R}_i^{-1}\tilde{C}_i),\ \tilde{\upsilon}_{2,N} = \min_{i\leqslant N}\sigma_{\min}(\tilde{R}_{i,i+1}).$$

By combining Lemmas 3–5, we obtain the following result.

Lemma 6. For any $k = 1, 2, \ldots$, and $1 \leqslant U \leqslant N$,

$$||\hat{x}_{k|k}^1(\mathcal{L}_{N+1}) - \hat{x}_{k|k}^1(\mathcal{L}_N)|| \leqslant \tilde{\eta}_{N+1}\left(\frac{\tilde{r}^U}{(\tilde{r}-1)\tilde{\bar{q}}}(\mathrm{e}^{\tilde{\psi}_N\tilde{\lambda}_N^{N-U}}-1) + \frac{2\tilde{c}}{1-\tilde{\iota}}\tilde{\iota}^U\right).$$

Proof. Using (37) and Lemmas (3–5), it follows that

$$\begin{aligned}
||\hat{x}_{k|k}^1(\mathcal{L}_{N+1}) - \hat{x}_{k|k}^1(\mathcal{L}_N)|| \leqslant& \sum_{j=1}^{U-1}||[P_k^{N+1}]^{1,j} - [P_k^N]^{1,j}||||q_k^j|| + \sum_{j=U}^{N}||[P_k^{N+1}]^{1,j} - [P_k^N]^{1,j}||||q_k^j|| \\
&+ ||[P_k^{N+1}]^{1,N+1}||||q_k^{N+1}|| \\
\leqslant& \tilde{\eta}_{N+1}\left(\sum_{j=1}^{U-1}\tilde{\bar{q}}^{-1}\tilde{r}^j(\mathrm{e}^{\tilde{\psi}_N\tilde{\lambda}_N^{N-U}}-1) + 2\tilde{c}\sum_{j=U}^{N}\tilde{\iota}^j + \tilde{c}\tilde{\iota}^{N+1}\right) \\
\leqslant& \tilde{\eta}_{N+1}\left(\tilde{\bar{q}}^{-1}(\mathrm{e}^{\tilde{\psi}_N\tilde{\lambda}_N^{N-U}}-1)\frac{\tilde{r}^U-\tilde{r}}{\tilde{r}-1} + 2\tilde{c}\frac{\tilde{\iota}^U-\tilde{\iota}^{N+2}}{1-\tilde{\iota}}\right) \\
\leqslant& \tilde{\eta}_{N+1}\left(\frac{\tilde{r}^U}{(\tilde{r}-1)\tilde{\bar{q}}}(\mathrm{e}^{\tilde{\psi}_N\tilde{\lambda}_N^{N-U}}-1) + \frac{2\tilde{c}}{1-\tilde{\iota}}\tilde{\iota}^U\right).
\end{aligned}$$

6.2.2 *Bound of the augmentation in an arbitrary graph*

At time k and at the N-th iteration, the state estimation $\hat{x}_{k|k}^1(N)$ obtained by node 1 using Algorithm 1 on $\mathcal{G}$ is equivalent to the estimate $\hat{x}_{k|k}^1(\mathcal{A}_N)$ acquired by the same algorithm on $\mathcal{A}_N$ and $\hat{x}_{k|k}^1(\mathcal{L}_N)$ obtained by Algorithm 1 on $\mathcal{L}_N$. Hence, $||\hat{x}_{k|k}^1(\mathcal{L}_{N+1}) - \hat{x}_{k|k}^1(\mathcal{L}_N)|| = ||\hat{x}_{k|k}^1(N+1) - \hat{x}_{k|k}^1(N)||$. By employing Lemma 6, we can derive a similar result as Lemma 27 in [24].

Lemma 7 ([24])**.** If $\check{\gamma} < 1$, then, based on the definition of Theorem 2, at time k and step $N \in \mathbb{N}$,

$$||\hat{x}_{k|k}^1(N+1) - \hat{x}_{k|k}^1(N)|| \leqslant \check{\bar{\Omega}}\check{\gamma}^N,$$

with $\check{\bar{\Omega}} = \frac{\check{\psi}\check{\eta}}{(\bar{q}-\underline{q})\lambda} + \frac{2\check{\eta}c}{1-\iota}, \check{\gamma} = \max\{\bar{e}\sqrt{\lambda}, \sqrt{\bar{e}}\iota^{1/\check{\varsigma}}\}$, where $\check{\psi} = (\mathrm{e}^{\check{\xi}\sqrt{\bar{n}(\bar{e}+1)}}-1), \check{\eta} = \frac{\bar{\varepsilon}\bar{z}\sqrt{8\bar{m}(\bar{e}+1)}}{\underline{r}} + \bar{\Sigma}\bar{\bar{X}}, \check{\varsigma} = 2 + \log_{\frac{1}{\sqrt{\lambda}}}\left(\bar{q}/\underline{q}\right), c = \frac{\bar{q}-\underline{q}}{2\bar{q}\underline{q}\iota}, \bar{\xi} = \max_j \log||I + (\sum_{i\in\mathcal{N}_j}B_{j,i}^{\mathrm{T}}R_{j,i}^{-1}B_{j,i})(C_j^{\mathrm{T}}R_j^{-1}C_j)^{-1}||$.

6.2.3 *Accuracy analysis*

In this study, we introduce $\hat{x}_1^M(k,n)$ as the suboptimal centralized MAP estimation of x_k^1 (the state at node 1) on the subgraph $\mathcal{S}_1(n-1)$, which includes nodes within $n-1$ hops from node 1 in $\mathcal{G}$. If Algorithm 1 is initialized in the acyclic graph, its output at node 1 will converge to the suboptimal centralized MAP estimation, as demonstrated in [22]. Observing that for any $n \leqslant f_1 + 1$, as defined by Notation 3 for the acyclic depth f_1,

$$\hat{x}_{k|k}^1(n) = \hat{x}_1^M(k,n). \tag{44}$$

Let $\check{\tau}_n$ be the collection of nodes that are $n-1$ steps from node 1 in the original cyclic graph $\mathcal{G}$. By consolidating all the nodes in $\check{\tau}_n$ and their internal connections, we obtain a single node. Then we can construct a linear graph similar to the process described in Subsection 3.2, but we omit a specific analysis for simplicity. The number of nodes in the graph $\check{\mathcal{L}}$ is expressed as

$$r_1 = \max_{j\leqslant I} D_{1,j} + 1,$$

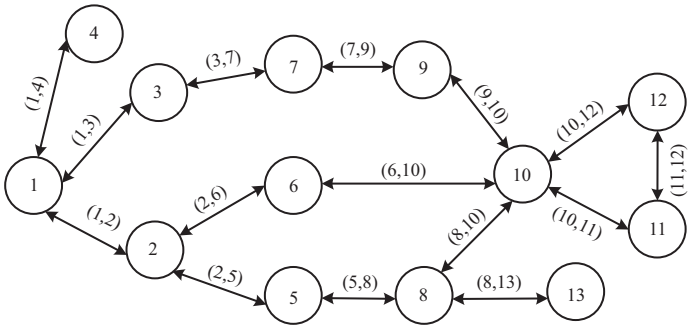

Figure 3 The cyclic network graph.

where $D_{1,j}$ represents the shortest path length between nodes 1 and j in the graph $\mathcal{G}$.

Lemma 7 establishes an upper bound on the estimation increment. We apply the proof procedure of this lemma for graph $\check{\mathcal{L}}$ instead of $\mathcal{L}_N$, then we get the following result.

Lemma 8. With the notation introduced in Theorem 2 and Lemma 7, we can obtain that for $n \in \mathbb{N}$, if $\gamma < 1$ and at time k, the following inequality holds:

$$||\hat{x}_1^M(k, n+1) - \hat{x}_1^M(k, n)|| \leqslant \begin{cases} \bar{\Omega}\gamma^n, & n \leqslant r_1, \\ 0, & n > r_1, \end{cases} \tag{45}$$

with $\bar{\Omega} = \frac{\bar{\psi}\bar{\eta}}{(\bar{q}-\underline{q})\phi} + \frac{2\bar{\eta}c}{1-\iota}$, $\bar{\psi} = \mathrm{e}^{\bar{\xi}(\bar{e}+1)\sqrt{\bar{n}}} - 1$, $\bar{\eta} = \bar{\varepsilon}\bar{z}(\bar{e}+1)\sqrt{8\bar{m}}\underline{r}^{-1} + \bar{\Sigma}\bar{\tilde{X}}$.

By combining the above inequality with (44), Lemma 7, and Lemma 8, we can establish a proof for Theorem 2.

Proof of Theorem 2.

$$\begin{aligned} ||\hat{x}_{k|k}^1(N) - \hat{x}_1^M(k)|| &= ||\hat{x}_{k|k}^1(N) - \hat{x}_{k|k}^1(f_1+1) + \hat{x}_1^M(k, f_1+1) - \hat{x}_1^M(k, r_1)|| \\ &\leqslant ||\hat{x}_{k|k}^1(N) - \hat{x}_{k|k}^1(f_1+1)|| + ||\hat{x}_1^M(k, r_1) - \hat{x}_1^M(k, f_1+1)|| \\ &\leqslant \check{\bar{\Omega}} \sum_{i=f_1+1}^{N-1} \check{\gamma}^i + \bar{\Omega} \sum_{i=f_1+1}^{r_1-1} \gamma^i \leqslant \frac{\check{\bar{\Omega}}}{1-\check{\gamma}} \check{\gamma}^{f_1+1} + \frac{\bar{\Omega}}{1-\gamma} \gamma^{f_1+1}. \end{aligned}$$

We can observe that $\alpha_1 \geqslant \mu_1, \alpha_2 \geqslant \mu_2, \beta_1 \leqslant \upsilon_1, \beta_2 \leqslant \upsilon_2$ imply $\phi \geqslant \lambda, \varsigma \geqslant \check{\varsigma}$, and therefore $\gamma \geqslant \check{\gamma}$. This enables us to derive an upper bound on the difference between $\hat{x}_{k|k}^1(N)$ and $\hat{x}_1^M(k, N)$. $||\hat{x}_{k|k}^1(N) - \hat{x}_1^M(k, N)|| \leqslant (\frac{\check{\bar{\Omega}}}{1-\check{\gamma}} + \frac{\bar{\Omega}}{1-\gamma})\gamma^{f_1+1}$, where $(\frac{\check{\bar{\Omega}}}{1-\check{\gamma}} + \frac{\bar{\Omega}}{1-\gamma})$ depends on $\bar{e}, \bar{n}, B_{i,j}, C_i, R_{i,j}$, R_i, F_i, and $\bar{z}$. Thus, we acquire the inequality (29).

7 Simulation

This section presents numerical examples to validate the theoretical analysis presented in this study. The objective is to demonstrate the effectiveness of Algorithm 1, a distributed dynamic state estimation algorithm, comparing its performance with the suboptimal centralized maximum a posteriori (SCMAP) state estimator [22] and the distributed static state estimator (DSSE) [15].

The communication network graph under consideration consists of 13 nodes with multiple loops, as depicted in Figure 3. All nodes in this network share identical state equations and measurement equations. The system parameters are listed below. $x_{k+1}^i = 0.98x_k^i + \omega_k^i$, $z_k^i = x_k^i + v_k^i$, $z_k^{i,j} = 0.5x_k^i + 0.5x_k^j + v_k^{i,j}$. $F_i = 0.01$, $R_i = R_{i,j} = 0.01$, $[x_0^1, x_0^2, \ldots, x_0^{13}]^{\mathrm{T}} = [2, 3, \ldots, 13, 14]^{\mathrm{T}}$, for all $i = 1, 2, \ldots, 13$ and $j \in \mathcal{N}_i$.

To ensure accurate results, 1000 Monte Carlo simulations are conducted to evaluate the performance of Algorithm 1. The time horizon is set to 100, and there are 50 iterations between each time steps t and $t+1$. The root mean square error (RMSE) is employed as the performance metric, defined as $e_k = \sqrt{\frac{1}{I}\sum_{i=1}^{I} ||\hat{x}_{k|k}^i - x_k^i||^2}$.

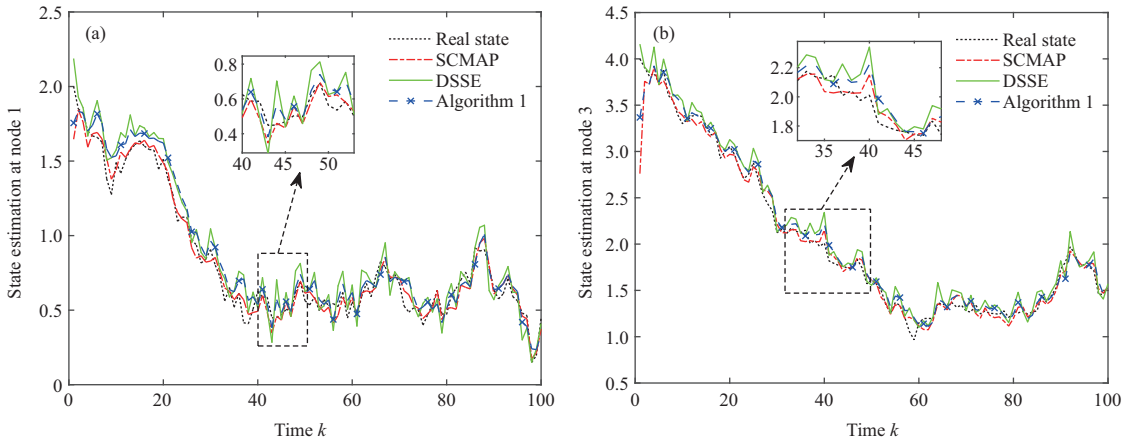

Figure 4 (Color online) The state estimation $\hat{x}^1_{k|k}$ and $\hat{x}^3_{k|k}$ for different state estimation algorithms. (a) Estimation $\hat{x}^1_{k|k}$ at node 1; (b) estimation $\hat{x}^3_{k|k}$ at node 3.

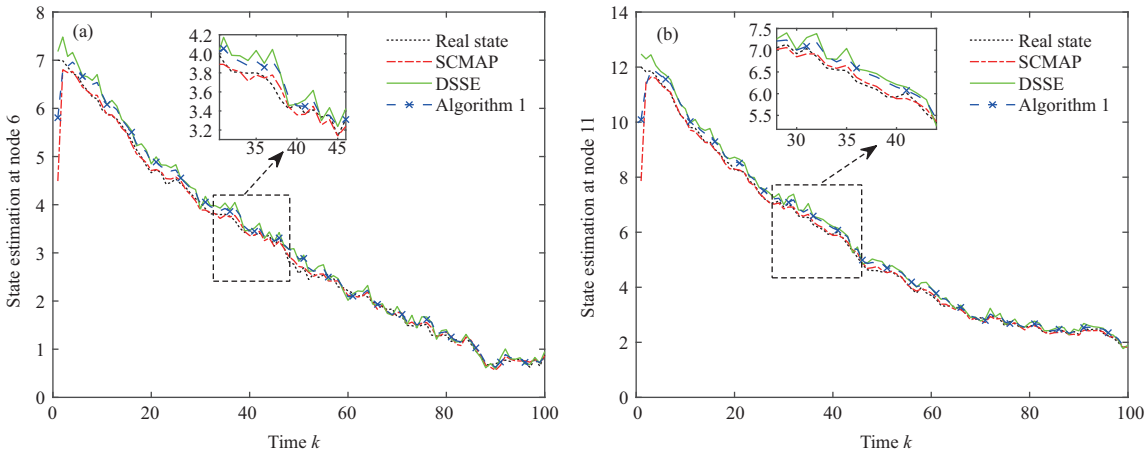

Figure 5 (Color online) The state estimation $\hat{x}^6_{k|k}$ and $\hat{x}^{11}_{k|k}$ for different state estimation algorithms. (a) Estimation $\hat{x}^6_{k|k}$ at node 6; (b) estimation $\hat{x}^{11}_{k|k}$ at node 11.

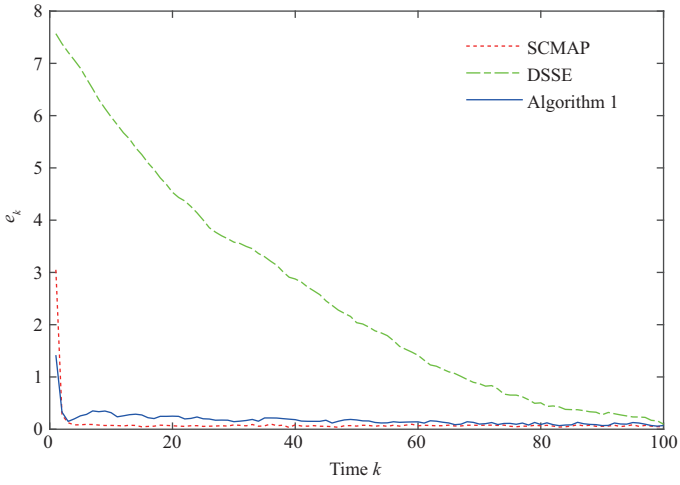

Figure 6 (Color online) RMSE e_k for different state estimation algorithms.

Figures 4 and 5 illustrate the dynamic estimation performance of nodes 1, 3, 6, and 11 for the SCMAP algorithm, Algorithm 1, and the DSSE algorithm. The results show that the SCMAP algorithm outperforms the DSSE algorithm and Algorithm 1 in terms of state estimation performance. Additionally,

Algorithm 1 exhibits satisfactory dynamic performance.

Figure 6 displays the RMSE values generated by Algorithm 1, the SCMAP algorithm, and the DSSE algorithm. Theorem 2 states that when the cyclic graph fulfills specific conditions, the results of Algorithm 1 converge asymptotically to those of the SCMAP algorithm. Figures 4–6 verify this result. Moreover, Algorithm 1 exhibits superior performance compared with the DSSE algorithm, thereby demonstrating the effectiveness of the proposed distributed dynamic state estimator.

8 Conclusion

In this study, we have investigated a distributed state estimation algorithm suitable for large-scale interconnected systems characterized by a cyclic network graph. The algorithm is based on the MAP state estimation, which employs local measurements and boundary information from neighboring nodes. We have provided the accuracy analysis by deriving bounds on the differences in estimation error covariance and state estimation between the proposed distributed algorithm and the suboptimal centralized MAP method, respectively, for a specific class of systems satisfying certain conditions, including cyclic topology and sparse connections. We have demonstrated that the results of the distributed dynamic estimator converge asymptotically to that of the suboptimal centralized MAP method, with the convergence rate depending on the loop-free depth. To validate the effectiveness of the proposed algorithm, we have presented simulation results. In future research, we will evaluate the effectiveness of the distributed state estimation algorithm by actual power system data. Moreover, future research directions will address challenges such as correlated measurement noises and transmission failures between nodes.

Acknowledgements This work was supported by National Key R&D Program of China (Grant Nos. 2018YFB1700100, 2020YFB1708600), National Natural Science Foundation of China (Grant Nos. 62173057, 62033006, 61906088), Liaoning Revitalization Talents Program (Grant No. XLYC2007187), Natural Science Foundation of Liaoning (Grant No. 2020-MS-122), and Dalian Young Talents Program (Grant No. 2021RQ038).

References

1 Hou N, Li J H, Liu H J, et al. Finite-horizon resilient state estimation for complex networks with integral measurements from partial nodes. Sci China Inf Sci, 2022, 65: 132205
2 Xiao H C, Ding D R, Dong H L, et al. Adaptive event-triggered state estimation for large-scale systems subject to deception attacks. Sci China Inf Sci, 2022, 65: 122207
3 Primadianto A, Lu C N. A review on distribution system state estimation. IEEE Trans Power Syst, 2016, 32: 3875–3883
4 Liu Y L, Chen W-H, Lu X M. Slow state estimation for singularly perturbed systems with discrete measurements. Sci China Inf Sci, 2021, 64: 129202
5 Olfati-Saber R, Murray R M. Consensus problems in networks of agents with switching topology and time-delays. IEEE Trans Automat Contr, 2004, 49: 1520–1533
6 Xie L, Choi D H, Kar S, et al. Fully distributed state estimation for wide-area monitoring systems. IEEE Trans Smart Grid, 2012, 3: 1154–1169
7 Gan D, Xie S Y, Liu Z X. Stability of the distributed Kalman filter using general random coefficients. Sci China Inf Sci, 2021, 64: 172204
8 Chen Y, Zhu M Z, Lu R Q, et al. Distributed H_∞ filtering of nonlinear systems with random topology by an event-triggered protocol. Sci China Inf Sci, 2021, 64: 202204
9 Battistelli G, Chisci L. Stability of consensus extended Kalman filter for distributed state estimation. Automatica, 2016, 68: 169–178
10 Mitra A, Richards J A, Bagchi S, et al. Resilient distributed state estimation with mobile agents: overcoming Byzantine adversaries, communication losses, and intermittent measurements. Auton Robot, 2019, 43: 743–768
11 Subbotin M V, Smith R S. Design of distributed decentralized estimators for formations with fixed and stochastic communication topologies. Automatica, 2009, 45: 2491–2501
12 Vu T T, Rahmani A R. Distributed consensus-based Kalman filter estimation and control of formation flying spacecraft: simulation and validation. In: Proceedings of AIAA Guidance, Navigation, and Control Conference, 2015. 1553
13 Xie L, Choi D H, Kar S. Cooperative distributed state estimation: local observability relaxed. In: Proceedings of IEEE Power and Energy Society General Meeting, 2011. 1–11
14 Pasqualetti F, Carli R, Bullo F. Distributed estimation via iterative projections with application to power network monitoring. Automatica, 2012, 48: 747–758
15 Tai X, Lin Z Y, Fu M Y, et al. A new distributed state estimation technique for power networks. In: Proceedings of American Control Conference, 2013. 3338–3343
16 Das S, Moura J M F. Distributed Kalman filtering with dynamic observations consensus. IEEE Trans Signal Process, 2015, 63: 4458–4473
17 Zhou Z W, Fang H T, Hong Y G. Distributed estimation for moving target based on state-consensus strategy. IEEE Trans Automat Contr, 2013, 58: 2096–2101
18 Battistelli G, Chisci L. Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability. Automatica, 2014, 50: 707–718
19 Chen B, Zhang W A, Yu L. Distributed fusion estimation with missing measurements, random transmission delays and packet dropouts. IEEE Trans Automat Contr, 2014, 59: 1961–1967
20 Chen B, Zhang W A, Yu L, et al. Distributed fusion estimation with communication bandwidth constraints. IEEE Trans Automat Contr, 2015, 60: 1398–1403

21 Yan J Q, Yang X, Mo Y L, et al. A distributed implementation of steady-state Kalman filter. IEEE Trans Automat Contr, 2023, 68: 2490–2497
22 Sun Y B, Fu M Y, Wang B C, et al. Dynamic state estimation for power networks using distributed MAP technique. Automatica, 2016, 73: 27–37
23 Anderson B D, Moore J B. Optimal Filtering. Mineola: Courier Corporation, 2005
24 Sui T J, Marelli D E, Fu M Y, et al. Accuracy analysis for distributed weighted least-squares estimation in finite steps and loopy networks. Automatica, 2018, 97: 82–91
25 Marelli D E, Fu M Y. Distributed weighted least-squares estimation with fast convergence for large-scale systems. Automatica, 2015, 51: 27–39
26 Weiss Y. Correctness of local probability propagation in graphical models with loops. Neural Computation, 2000, 12: 1–41
27 Ihler A T, Fisher III J W, Willsky A S, et al. Loopy belief propagation: convergence and effects of message errors. J Mach Learn Res, 2005, 6: 905–936
28 Xie K, Cai Q Q, Zhang Z R, et al. Distributed algorithms for average consensus of input data with fast convergence. IEEE Trans Syst Man Cybern Syst, 2021, 51: 2653–2664
29 Makki S A M, Havas G. Distributed algorithms for depth-first search. Inf Processing Lett, 1996, 60: 7–12
30 Bougerol P. Kalman filtering with random coefficients and contractions. SIAM J Control Optim, 1993, 31: 942–959
31 Meurant G. A review on the inverse of symmetric tridiagonal and block tridiagonal matrices. SIAM J Matrix Anal Appl, 1992, 13: 707–728

Appendix A Proof of Lemma 1

We focus on the line graph $\mathcal{L}_N$ derived from the original cyclic graph $\mathcal{G}$. We define $\tilde{Q}_k^i(n, O_k)$ as the information matrix acquired by node i at time k and step n while implementing Algorithm 1 on the line graph $\mathcal{L}_N$ with the initial set O_k. Similarly, we utilize $\tilde{Q}_k^{i\to i-1}(n, O_k)$ and $\tilde{Q}_k^{i\to i+1}(n, O_k)$ to represent the information matrices sent from node i to its left and right neighbors, respectively. Furthermore, we define $\tilde{Q}_k^{i\to i-1,i}(n, O_k) \triangleq \tilde{Q}_k^i(n, O_k) - \tilde{Q}_k^{i-1\to i}(n-1, O_k)$ as the message transmitted from node i to left neighbor $i-1$ without using the information from its left neighbor in the previous time step.

Using Proposition 1, we can calculate the Riemannian distance between $Q_k^1(N, O_k^1)$ and $Q_k^1(N, O_k^2)$ as follows:

$$\begin{aligned}
&\delta(\tilde{Q}_k^1(N, O_k^1) - \tilde{Q}_k^1(N, O_k^2))\\
&\quad= \delta(\tilde{C}_1^{\mathrm{T}}\tilde{R}_1^{-1}\tilde{C}_1 + (\Sigma_{k|k-1}^1)^{-1} + \tilde{B}_{1,2}^{\mathrm{T}}[\tilde{R}_{1,2} + \tilde{B}_{2,1}(\tilde{Q}_k^{2\to 1,2}(N-1, O_k^1))^{-1}\cdot \tilde{B}_{2,1}^{\mathrm{T}}]^{-1}\tilde{B}_{1,2},\ \tilde{C}_1^{\mathrm{T}}\tilde{R}_1^{-1}\tilde{C}_1\\
&\qquad + (\Sigma_{k|k-1}^1)^{-1} + \tilde{B}_{1,2}^{\mathrm{T}}[\tilde{R}_{1,2} + \tilde{B}_{2,1}(\tilde{Q}_k^{2\to 1,2}(N-1, O_k^2))^{-1}\tilde{B}_{2,1}^{\mathrm{T}}]^{-1}\tilde{B}_{1,2})\\
&\quad\leqslant \frac{\tilde{\mu}_{1,N}}{\tilde{\mu}_{1,N}+\tilde{v}_{1,N}}\delta(\tilde{R}_{1,2} + \tilde{B}_{2,1}(\tilde{Q}_k^{2\to 1,2}(N-1, O_k^1))^{-1}\tilde{B}_{2,1}^{\mathrm{T}}, \tilde{R}_{1,2} + \tilde{B}_{2,1}(\tilde{Q}_k^{2\to 1,2}(N-1, O_k^2))^{-1}\tilde{B}_{2,1}^{\mathrm{T}})\\
&\quad\leqslant \frac{\tilde{\mu}_{1,N}}{\tilde{\mu}_{1,N}+\tilde{v}_{1,N}}\frac{\tilde{\mu}_{2,N}}{\tilde{\mu}_{2,N}+\tilde{v}_{2,N}}\delta(\tilde{Q}_k^{2\to 1,2}(N-1, O_k^1), \tilde{Q}_k^{2\to 1,2}(N-1, O_k^2)),
\end{aligned}$$

where

$$\begin{aligned}
&\tilde{\mu}_{1,N} = \max_{i\leqslant N}||\tilde{B}_{i+1,i}^{\mathrm{T}}\tilde{R}_{i,i+1}^{-1}\tilde{B}_{i+1,i}||,\ \tilde{\mu}_{2,N} = \max_{i\leqslant N}||\tilde{B}_{i+1,i}(\tilde{C}_{i+1}^{\mathrm{T}}\tilde{R}_{i+1}^{-1}\tilde{C}_{i+1})\tilde{B}_{i+1,i}^{\mathrm{T}}||,\\
&\tilde{v}_{1,N} = \max_{i\leqslant N}\sigma_{\min}(\tilde{C}_i^{\mathrm{T}}\tilde{R}_i^{-1}\tilde{C}_i^{\mathrm{T}}),\ \tilde{v}_{2,N} = \max_{i\leqslant N}\sigma_{\min}(\tilde{R}_{i,i+1}).
\end{aligned}$$

By using (24) and (25), we can establish $\tilde{\mu}_{1,N}\leqslant\mu_1, \tilde{v}_{1,N}\geqslant v_1$. From this, we can derive

$$\begin{aligned}
\delta(\tilde{Q}_k^1(N, O_k^1), \tilde{Q}_k^2(N, O_k^2)) &\leqslant \lambda\delta(\tilde{Q}_k^{2\to 1,2}(N-1, O_k^1), \tilde{Q}_k^{2\to 1,2}(N-1, O_k^2))\\
&\leqslant \lambda^{N-1}\delta(\tilde{Q}_k^{N\to N-1,N}(1, O_k^1), \tilde{Q}_k^{N\to N-1,N}(1, O_k^2)).
\end{aligned}$$

For any initial condition O_k, the following equality holds: $\tilde{Q}_k^{N\to N-1,N}(1, O_k) = \tilde{C}_N^{\mathrm{T}}\tilde{R}_N^{-1}\tilde{C}_N + (\tilde{\Sigma}_{k|k-1}^N)^{-1} + Q_k^{N+1\to N}(0, O_k)$. Furthermore,

$$\begin{aligned}
&\delta(\tilde{C}_N^{\mathrm{T}}\tilde{R}_N^{-1}\tilde{C}_N + (\tilde{\Sigma}_{k|k-1}^N)^{-1} + Q_k^{N+1\to N}(0, O_k^1), \tilde{C}_N^{\mathrm{T}}\tilde{R}_N^{-1}\tilde{C}_N + (\tilde{\Sigma}_{k|k-1}^N)^{-1} + Q_k^{N+1\to N}(0, O_k^2))\\
&\quad\leqslant \delta(\tilde{C}_N^{\mathrm{T}}\tilde{R}_N^{-1}\tilde{C}_N + Q_k^{N+1\to N}(0, O_k^1), \tilde{C}_N^{\mathrm{T}}\tilde{R}_N^{-1}\tilde{C}_N + Q_k^{N+1\to N}(0, O_k^2))\\
&\quad\leqslant \delta(\tilde{C}_N^{\mathrm{T}}\tilde{R}_N^{-1}\tilde{C}_N, \tilde{C}_N^{\mathrm{T}}\tilde{R}_N^{-1}\tilde{C}_N + \tilde{B}_{N,N+1}^{\mathrm{T}}\tilde{R}_{N,N+1}^{-1}\tilde{B}_{N,N+1}).
\end{aligned}$$

From the findings in Subsection 3.2, it is evident that $|\tau_n|$ denotes the number of nodes in layer n of the graph $\mathcal{A}_N$. Considering that $|\tau_n|\leqslant(\bar{e}+1)\bar{e}^{n-1}$, we can infer

$$\begin{aligned}
&\delta(\tilde{Q}_k^{N\to N-1,N}(1, O_k^1), \tilde{Q}_k^{N\to N-1,N}(1, O_k^2))\\
&\quad\leqslant \delta(\tilde{C}_N^{\mathrm{T}}\tilde{R}_N^{-1}\tilde{C}_N, \tilde{C}_N^{\mathrm{T}}\tilde{R}_N^{-1}\tilde{C}_N + \tilde{B}_{N,N+1}^{\mathrm{T}}\tilde{R}_{N,N+1}^{-1}\tilde{B}_{N,N+1}) \leqslant \sqrt{\bar{n}|\tau_n|}\max_{i\leqslant N}\log||I + (\tilde{B}_{i,i+1}^{\mathrm{T}}\tilde{R}_{i,i+1}^{-1}\tilde{B}_{i,i+1})\\
&\quad\times(\tilde{C}_i^{\mathrm{T}}\tilde{R}_i^{-1}\tilde{C}_i)^{-1}|| \leqslant \sqrt{\bar{n}|\tau_n|}\max_j\log||I + \left(\sum_{k\in N_j}B_{j,k}^{\mathrm{T}}R_{j,k}^{-1}B_{j,k}\right)(C_i^{\mathrm{T}}R_i^{-1}C_i)^{-1}|| \leqslant (\sqrt{\bar{e}})^{N-1}\bar{\zeta}.
\end{aligned}$$

Considering the equivalent conversion of the graph, we establish the equality $\delta(\tilde{Q}_k^1(N, O_k^1), \tilde{Q}_k^2(N, O_k^2)) = \delta(Q_k^1(N, O_k^1), Q_k^2(N, O_k^2))$. Hence, we obtain the Riemannian distance between $Q_k^1(N, O_k^1)$ and $Q_k^1(N, O_k^2)$. $\delta(Q_k^1(N, O_k^1), Q_k^2(N, O_k^2)) = \delta(\tilde{Q}_k^1(N, O_k^1), \tilde{Q}_k^2(N, O_k^2))\leqslant\bar{\zeta}(\lambda\sqrt{\bar{e}})^{N-1} = \bar{\zeta}\varphi^{N-1}$.

Appendix B Proof of Lemma 3

We begin by considering the expression for $\tilde{\Sigma}^i_{k|k-1}$, which is given by $\tilde{\Sigma}^i_{k|k-1} = \tilde{A}_i\tilde{\Sigma}^i_{k-1|k-1}(N)\tilde{A}_i^{\mathrm{T}} + \tilde{F}_i$. By applying the matrix inversion lemma, we can derive $(\tilde{\Sigma}^i_{k|k-1})^{-1} = \tilde{F}_i^{-1} - \tilde{F}_i^{-1}\tilde{A}_i((\tilde{\Sigma}^i_{k-1|k-1}(N))^{-1} + \tilde{A}_i^{\mathrm{T}}\tilde{F}_i^{-1}\tilde{A}_i)^{-1}\tilde{A}_i^{\mathrm{T}}\tilde{F}_i^{-1}$. It follows that $||(\tilde{\Sigma}^i_{k|k-1})^{-1}|| < ||\tilde{F}_i^{-1}||$. Moreover, $\tilde{\Sigma}^N_{k|k-1} = \mathrm{diag}\{\tilde{\Sigma}^1_{k|k-1}, \ldots, \tilde{\Sigma}^N_{k|k-1}\}$. Finally, we can demonstrate that Eq. (38) holds.

Appendix C Proof of Lemma 4

$||\hat{\boldsymbol{x}}^N_{k|k-1}|| = [(\hat{\tilde{x}}^1_{k|k-1})^{\mathrm{T}}, \ldots, (\hat{\tilde{x}}^N_{k|k-1})^{\mathrm{T}}]^{\mathrm{T}}$. By utilizing (6), (7), and (30), we can derive the following inequality:

$$||\hat{\boldsymbol{x}}^N_{1|0}|| = ||\boldsymbol{A}_N\hat{\boldsymbol{x}}^N_{0|0}|| \leqslant ||\boldsymbol{A}_N||||\hat{\boldsymbol{x}}^N_{0|0}|| \leqslant \max_{i\leqslant N}||\tilde{A}_i||\sqrt{N\max_{i\leqslant N}||\hat{\tilde{x}}^i_{0|0}||}.$$

This implies that the inequality (39) holds when $k = 1$.

Assume the existence of a positive scalar $\bar{\sigma}$ such that

$$||\hat{\boldsymbol{x}}^N_{n|n-1}|| \leqslant \bar{\sigma}. \tag{C1}$$

It is necessary to prove that Eq. (39) holds at $k = n + 1$. By using (30), we obtain $||\hat{\boldsymbol{x}}^N_{n+1|n}|| = ||\boldsymbol{A}_N\hat{\boldsymbol{x}}^N_{n|n}||$. Then, we have

$$\begin{aligned}||\hat{\boldsymbol{x}}^N_{n|n}|| = ||\boldsymbol{Q}^N_k\boldsymbol{q}^N_k|| &\leqslant ||(\boldsymbol{Q}^N_n)^{-1}\boldsymbol{q}^N_n|| = ||((\boldsymbol{G}_N)^{\mathrm{T}}\boldsymbol{S}_N^{-1}\boldsymbol{G}_N + (\check{\Sigma}^N_{n|n-1})^{-1})^{-1}||||(\boldsymbol{G}_N)^{\mathrm{T}}\boldsymbol{S}_N^{-1}\boldsymbol{y}^N_n + (\check{\Sigma}^N_{n|n-1})^{-1} \times \hat{\boldsymbol{x}}^N_{k|k-1}|| \\ &\leqslant ||((\boldsymbol{G}_N)^{\mathrm{T}}\boldsymbol{S}_N^{-1}\boldsymbol{G}_N)||[||(\boldsymbol{G}_N)^{\mathrm{T}}\boldsymbol{S}_N^{-1}\boldsymbol{y}^N_n|| + ||(\check{\Sigma}^N_{n|n-1})^{-1}||||\hat{\boldsymbol{x}}^N_{n|n-1}||].\end{aligned}$$

By substituting (38) and (C1) into the above inequality, it reveals the existence of an upper bound that satisfies (39). This bound is determined by $\bar{e}, \bar{n}, B_{i,j}, C_i, R_{i,j}$, R_i, F_i, and $\bar{z}$.