Robustness and Scalability of Consensus Networks: the Role of Memory Information

Jiamin Wang, Jian Liu, Feng Xiao, *Senior Member, IEEE*, and Yuanshi Zheng, *Senior Member, IEEE*

***Abstract*— It has been reported that local memory information could enhance certain consensus performance of multi-agent networks, such as protecting privacy and accelerating consensus. This article aims to investigate whether memory information can improve the robustness and scalability of consensus networks. The robustness is measured by the ℓ_2 gains from disturbances to consensus errors, and the scalability means that consensus can be preserved without re-tuning control parameters as the network scale increases. Using the linear combination of previous and current iteration states of agents and their neighbors, a memory-based consensus protocol is developed and we provide a necessary and sufficient condition for achieving consensus. Then, we establish the analytic expression of the ℓ_2 gain, which is exclusively determined by control parameters and non-zero minimum and maximum Laplacian eigenvalues. Furthermore, we show how tuning the memory coefficient can improve both robustness and scalability, and the optimal control parameters are further derived. Interestingly, we observe a positive correlation between robustness and scalability.**

***Index Terms*— Consensus networks, memory information, robustness, scalability, graph Laplacian**

I. Introduction

In recent decades, the control community has dedicated significant efforts to studying the consensus problems in multi-agent networks. In retrospect, since early insightful works [1]–[3] opened the gates of research on consensus problems, abundant studies constantly sprang up, including consensus for agents with various dynamics [4]–[7], event-based consensus [8], etc.

Another significant topic is to investigate the robustness of consensus with respect to disturbances. In the literature, the $\mathcal{L}_2$ gain is typically used to evaluate the robustness of continuous-time multi-agent networks, while the counterpart of discrete-time multi-agent networks is the ℓ_2 gain. Most researchers concentrated on achieving consensus with a prescribed $\mathcal{L}_2$ or ℓ_2 gain by using linear matrix inequality (LMI) technique [9]–[13]. In a quite different research avenue, some investigations explored the inherent relationship between the robustness and the network properties [14]–[19]. Specifically, it was shown in [14] that the $\mathcal{L}_2$ gain of the consensus network, integrating the first-order continuous-time multi-agent model and the undirected communication topology, is inversely proportional to the non-zero minimum Laplacian eigenvalue. It is worth highlighting that the non-zero minimum Laplacian eigenvalue of an undirected graph is also known as the *Fiedler eigenvalue* and it is used to measure the algebraic connectivity of the graph. Authors in [15] proved that the $\mathcal{L}_2$ gains for the first-order leader-follower multi-agent networks with undirected communication topologies are determined by the minimum eigenvalues of the grounded Laplacian matrices. In their subsequent work [16], it was proved that for the same multi-agent networks with directed communication topologies, the $\mathcal{L}_2$ gains hinge on the minimum singular values of the grounded Laplacian matrices. In addition, authors in [17] considered more general $\mathcal{L}_2$ and ℓ_2 performance indices for continuous-time consensus networks and the corresponding sampled-data consensus networks, respectively, and focused on the effect of periodic sampling on robustness. It was proved that the ℓ_2 gain of the sampled-data multi-agent network is totally reliant on the non-zero minimum Laplacian eigenvalue, the maximum Laplacian eigenvalue, and the sampling period, and the optimal sampling period was derived. In [18], the robustness of the first-order consensus against stochastic noises and deterministic disturbances was respectively analyzed in the presence of communication delays. In a recent work [19], the authors established the analytic expressions of $\mathcal{L}_2$ gains for second-order multi-agent networks and proposed a concise graph condition to ascertain which one of the absolute velocity protocol and the relative velocity protocol provides better robustness. These works mainly focused on the robustness analysis of continuous-time multi-agent consensus networks. The connection between the robustness of discrete-time consensus networks and the network features remains inadequately explored, and hence deserves to be further studied.

A striking feature in the existing literature is that the design of certain protocol parameters in consensus networks heavily depends on some global network properties, for example, the Laplacian spectrum. Once the network structure changes or the network size grows, we have to re-tune the control parameters to recapture consensus. However, it is impractical and costly for a multi-agent network with finite network resources. On the other hand, the Laplacian spectrum, as a kind of global key information for achieving consensus, can only be obtained in large-scale network scenarios through

This work is funded by the National Natural Science Foundation of China (62273267). The material in this paper was not presented at any conference. *(Corresponding author: Yuanshi Zheng)*

Jiamin Wang, Jian Liu and Yuanshi Zheng are with the Center for Multi-Agent, School of Mechano-electronic Engineering, Xidian University, Xi'an, 710071, China (E-mail: wangjiamin21@outlook.com; liujianzym@outlook.com; zhengyuanshi2005@163.com).

F. Xiao is with the School of Control and Computer Engineering, North China Electric Power University, Beijing, 102206, China (E-mail: fengxiao@ncepu.edu.cn).

decentralized or distributed estimation [20], [21], which further occupies scarce computational resources. Therefore, as more agents join in multi-agent networks, control gains are supposed to be preset and cannot be re-tuned. This feature was termed as *scalability* in [22]. If the consensus can be preserved as the network size grows without re-tuning the control parameters, it is called *scalable consensus*. In fact, preserving consensus is a fundamental requirement of scalability, networks should also maintain high connectivity as they grow in practical applications for the sake of rapid consensus, as reported in [23]. However, addressing this problem is in need of graph-theoretical endeavor which is outside the scope of the present article. We restrict attention here to study basic scalable consensus. Inspired by these observations, we delve into the scalability of discrete-time multi-agent networks.

Typically, consensus algorithms are constructed by using real-time information of agents, and once agents update the local information, their previous states are forgotten. Even if some memory information, such as the data shift resulting from delays, is sometimes inevitable, the majority of researchers just trivially incorporated it into the algorithm design as usual (see [24]–[26]), putting no weight on its potential advantages to consensus performance. In contrast, authors in [27], [28] assumed that state information can be stored in the on-board memories of agents and subsequently invoked at future time-steps. It was shown in [27] that repeatedly employing the memory information at two time-steps reduces the conservatism of consensus conditions given in [1]. For second-order continuous-time multi-agent networks, authors in [28] utilized the position-like states of agents from the past two moments to estimate the current velocity-like states. Moreover, some studies explored the superiority of memory information on consensus performances [29]–[36]. Specifically, in [29], agents employed past information to predict their neighbors' states, which increased the consensus convergence rate for continuous-time multi-agent networks. The authors in [30] claimed that delays did not affect the equidistant deployment of mobile agents. For first-order discrete-time multi-agent networks, the authors in [31] designed the following consensus algorithm

$$x_i(t+1) = \frac{1}{2|\mathcal{N}_i|}\left[\sum_{j\in\mathcal{N}_i} x_j(t) + \sum_{j\in\mathcal{N}_i} x_j(t-1)\right] \tag{1}$$

to obtain higher privacy index — an index based on the structural observability to characterize the confidentiality of privacy, where $x_i(t) \in \mathbb{R}$ is the state of the i-th agent, $|\mathcal{N}_i|$ is the number of neighbors of the i-th agent, and each agent shared its previous and current iteration states with its neighbors. In order to accelerate the convergence rate of consensus, the authors in [32] designed a predictor-based distributed consensus algorithm depending on the linear combination of current and history iteration information. Subsequently, the authors in [33] considered the one-tap memory version of the algorithm proposed in [32], provided the first theoretical verification for its accelerated convergence rate of consensus, and derived the optimal mixing parameter. In fact, the algorithms presented in [32], [33] can be regarded as special cases of the following memory-based algorithm given in [34].

$$\begin{aligned} x_i(t+1) = x_i(t) &+ \sum_{\tau=0}^{M} \varrho_\tau x_i(t-\tau) \\ &+ \vartheta \sum_{j\in\mathcal{N}_i} a_{ij}[x_j(t) - x_i(t)], \end{aligned} \tag{2}$$

where ϑ and ϱ_τ $(\tau = 0, \ldots, M)$ are constant parameters, $a_{ij} \geq 0$ is the weight of the edge between the i-th agent and the j-th agent, and $\sum_{\tau=0}^{M} \varrho_\tau = 0$ is a necessary condition for reaching average consensus. The authors in [34] also built analytic expressions of the optimal convergence rate and the corresponding optimal control parameters for (2) with different memory sizes. However, in aforementioned algorithms, memory information is private for each agent and cannot be shared with other agents. In this regard, a novel protocol was proposed in [35] which is shown as follows

$$x_i(t+1) = x_i(t) + \sum_{\tau=0}^{M} \vartheta_\tau \sum_{j\in\mathcal{N}_i} a_{ij}[x_j(t-\tau) - x_i(t-\tau)], \tag{3}$$

where each agent could also receive neighbors' memory information besides its own memory. In a recent advance [36], the authors gave the following general form of existing memory-based consensus algorithms.

$$\begin{aligned} x_i(t+1) = x_i(t) &+ \sum_{\tau=0}^{M} \varrho_\tau x_i(t-\tau) \\ &+ \sum_{\tau=0}^{M} \vartheta_\tau \sum_{j\in\mathcal{N}_i} a_{ij}[x_j(t-\tau) - x_i(t-\tau)], \end{aligned} \tag{4}$$

where ϑ_τ and ϱ_τ $(\tau = 0, \ldots, M)$ are constant parameters.

Note that the above works unveiled that memory information is able to protect privacy and accelerate the convergence rate of consensus. Impressed by these facts, we naturally wonder: *Whether the memory information can improve robustness and scalability of consensus networks?* On the other hand, although the role of memory information on consensus performance was reported, the consensus analysis has not been exhaustively discussed.

Inspired by all aforementioned observations, this article intends to explore the role of memory information on the robustness and the scalability of first-order discrete-time multi-agent consensus networks. We provide comprehensive consensus analysis for the multi-agent network with a memory-based consensus protocol. Then we investigate the effect of control parameters on the robustness and the scalability, where the robustness is characterized by the ℓ_2 gains from disturbances to consensus errors. Our contributions are summarized as follows:

- We propose a memory-based consensus protocol which is the linear combination of previous and current iteration states of agents and their neighbors. This protocol generalizes the parameter-free protocol in [31] and the protocol structures in [25]–[27]. On the other hand, it is also an elementary form extracted from [32]–[36], while we perform a more thorough consensus analysis that is necessary for the discussion of the robustness and scalability of the consensus property. Specifically, the necessary and sufficient condition for achieving consensus is established and we derive a specific consensus region that has not been reported yet. Interestingly, we also discover that the memory-based consensus protocol produces diverse consensus behaviors by adjusting the memory coefficient.

- By using the orthogonal transformation technique, we build the quantitative relations between ℓ_2 gains, control parameters, and the non-zero minimum and maximum Laplacian eigenvalues for first-order discrete-time multi-agent networks with the considered memory-based consensus protocols. In addition, we determine the optimal control parameters such that the consensus networks obtain the optimal robustness in the entire consensus region.
- The optimal memory coefficient with regard to the scalability is derived. It is shown that the memory information can simultaneously improve the robustness and the scalability as long as the memory coefficient is set as $0 < \alpha < 1$. More interestingly, it is shown that robustness and scalability are positively correlated.

The remainder of this article is organized as follows. In Section II, we present some preliminaries. The problem formulation is provided in Section III. Comprehensive consensus analysis, robustness analysis, and scalability analysis are given in Section IV. Further implications of our results are discussed in Section V. We conclude our work in Section VI.

Notation: Throughout this paper, $\mathbb{N}^+$ denotes the set of positive integers, $\mathbb{R}$ is the set of real numbers, $\mathbb{R}^n$ represents the n-dimensional real column vector space, and $\mathbb{R}^{n\times m}$ denotes the $n \times m$ real matrix space. For $x \in \mathbb{R}$, $\lfloor x \rfloor$ denotes the floor function and gives the largest integer value less than or equal to x. Let $\mathbf{0}$ and $\mathbf{1}$ be the all-zero and all-one matrices with appropriate dimensions, respectively. In particular, $n \times 1$ all zeros and all ones column vectors are denoted as $\mathbf{0}_n$ and $\mathbf{1}_n$, respectively. I_n indicates the n-dimensional identity matrix. For a matrix M, $M^\top$ denotes its transpose, M^H labels its conjugate transpose, and $\sigma_{\max}(M)$ represents its maximum singular value. For a Hermitian matrix X, $\lambda_{\max}(M)$ denotes its maximum eigenvalue. For a matrix $M \in \mathbb{R}^{n\times n}$, $\det(M)$ denotes its determinant and M is said to be orthogonal if $M^\top M = MM^\top = I_n$. $\mathrm{diag}\{a_1, a_2, \dots, a_n\}$ designates a diagonal matrix whose i_{th} diagonal element is a_i. For a finite set $\mathcal{S}$, $|\mathcal{S}|$ represents its cardinality. Define a set $\mathcal{I}_n = \{1, 2, \dots, n\}$. The null set is denoted as $\emptyset$. $\mathbf{j}$ is the imaginary unit. $\ell_2\,[0,\infty)$ dictates the space of square-summable infinite vector sequences, i.e., $f(t) \in \ell_2\,[0,\infty)$ if and only if $\sum_{t=0}^{\infty} f^\top(t) f(t) < \infty$.

II. Preliminaries

Consider a group of n agents. Their communication network is represented as a weighted undirected graph $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n, \mathcal{A}_n)$, where $\mathcal{V}_n = \{v_1, v_2, \dots, v_n\}$ is the set of vertices, $\mathcal{E}_n \subseteq \mathcal{V}_n \times \mathcal{V}_n$ is the set of edges, and $\mathcal{A}_n = [a_{ij}] \in \mathbb{R}^{n\times n}$ is the weighted adjacency matrix with $a_{ij} = a_{ji} \geq 0$. $\varepsilon_{ij} = (v_i, v_j) \in \mathcal{E}_n$ indicates that there exists an edge between vertices v_i and v_j, which means that they can receive each other's information. The adjacency element a_{ij} denotes the weight of the edge ε_{ij}, and $a_{ij} > 0$ if and only if $\varepsilon_{ij} \in \mathcal{E}_n$, otherwise $a_{ij} = 0$. Furthermore, assume that $a_{ii} = 0$. A *path* between two distinct vertices v_i and v_j is a sequence of distinct edges in $\mathcal{G}_n$ in the form $(v_i, v_{k_1}), (v_{k_1}, v_{k_2}), \dots, (v_{k_l}, v_j)$. An undirected graph is said to be *connected* if there is a path between any two distinct vertices of the graph. The degree matrix $\mathcal{D}_n = [d_{ij}] \in \mathbb{R}^{n\times n}$ is a diagonal matrix with $d_{ii} = \sum_{j=1}^{n} a_{ij}$. The Laplacian matrix $L_n = [l_{ij}] \in \mathbb{R}^{n\times n}$ associated with graph $\mathcal{G}_n$ is defined by $l_{ii} = \sum_{j=1, j\neq i}^{n} a_{ij}$ and $l_{ij} = -a_{ij}$, $j \neq i$. $\mathcal{N}_i = \{v_j \in \mathcal{V}_n | (v_i, v_j) \in \mathcal{E}_n\}$ denotes the neighbor set of vertex v_i.

The following definitions and lemmas will be utilized to establish our main results.

Lemma 1 ([1]): An undirected graph $\mathcal{G}_n$ is connected if and only if its Laplacian matrix L_n possesses a simple zero eigenvalue with $\mathbf{1}_n$ being the eigenvector and the other eigenvalues of L_n are positive.

For convenience, if the undirected graph $\mathcal{G}_n$ is connected, the eigenvalues λ_i, $i \in \mathcal{I}_n$ of Laplacian matrix L_n assume the form

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n \tag{5}$$

and the set of all non-zero eigenvalues of L_n is denoted by $\Gamma = \{\lambda_2, \dots, \lambda_n\}$.

Lemma 2 ([10]): For a connected undirected graph $\mathcal{G}_n$ with a Laplacian matrix $L_n \in \mathbb{R}^{n\times n}$, denote a symmetric matrix $\Theta_n = I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top$. Then, there exists an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ such that

$$U^\top \Theta_n U = \bar{\Theta}_n = \begin{bmatrix} 0 & \mathbf{0}_{n-1}^\top \\ \mathbf{0}_{n-1} & I_{n-1} \end{bmatrix} \tag{6}$$

and

$$U^\top L_n U = \bar{L}_n = \begin{bmatrix} 0 & \mathbf{0}_{n-1}^\top \\ \mathbf{0}_{n-1} & \tilde{L} \end{bmatrix}, \tag{7}$$

where $\tilde{L} \in \mathbb{R}^{(n-1)\times(n-1)}$ is positive definite.

Definition 1 ([37]): The H_∞ norm of an asymptotically stable discrete-time transfer matrix $T(z)$ is defined as $\|T(z)\|_\infty = \sup_{\upsilon\in[0,2\pi]} \sigma_{\max}\big[T(e^{\mathbf{j}\upsilon})\big]$.

III. Problem Statement

Consider a multi-agent network consisting of n agents with the discrete-time dynamics

$$x_i(t+1) = u_i(t), i \in \mathcal{I}_n, \tag{8}$$

where $x_i(t) \in \mathbb{R}$ is the state and $u_i(t) \in \mathbb{R}$ denotes the control input. These agents communicate on a network described as a weighted undirected graph $\mathcal{G}_n$. Here, we propose a memory-based consensus protocol

$$u_i(t) = \alpha\psi_i(t) + (1-\alpha)\psi_i(t-1), \tag{9}$$

where α is termed the memory coefficient and

$$\psi_i(t) = x_i(t) + \beta \sum_{j\in\mathcal{N}_i} a_{ij}\big[x_j(t) - x_i(t)\big]. \tag{10}$$

In (10), β is the control gain and a_{ij} is the weight of the edge ε_{ij}. Multi-agent network (8)-(9) is said to reach consensus if for any initial conditions $x_i(-1)$ and $x_i(0)$, we have $\lim_{t\to\infty}[x_i(t) - x_j(t)] = 0$, $\forall i, j \in \mathcal{I}_n$. Obviously, each agent retrieves its own and neighbors' current states and memory information of previous iteration to update its state. In particular, in the memoryless case with $\alpha = 1$, protocol (9) reduces to the classic first-order discrete-time consensus protocol [38].

Define a vector $\bar{x}(t) = \begin{bmatrix} x(t-1) \\ x(t) \end{bmatrix} \in \mathbb{R}^{2n}$, where $x(t) = [x_1(t), \ldots, x_n(t)]^\top \in \mathbb{R}^n$. Then, substituting (9) and (10) into (8) gives the augmented collective dynamics

$$\bar{x}(t+1) = \underbrace{\begin{bmatrix} \mathbf{0} & I_n \\ (1-\alpha)(I_n - \beta L_n) & \alpha(I_n - \beta L_n) \end{bmatrix}}_{\Phi_n} \bar{x}(t), \quad (11)$$

where L_n is the Laplacian matrix associated with the graph $\mathcal{G}_n$.

Different from existing works, we aim to provide the convergence analysis of consensus for the discrete-time multi-agent network (8)-(9) and explore the effect of memory information on the robustness and the scalability. Specifically, compared with the memoryless case (i.e., $\alpha = 1$), we try to investigate whether the memory information could improve the robustness and the scalability. Before establishing our main results, we introduce the notions of robustness and scalability of the multi-agent network.

Remark 1: The consensus analysis in related works [31]–[36] is far from perfect, spurring us to bridge this gap for the memory-based multi-agent networks. Although (11) can be treated as a class of time-delay systems, we try to utilize time-delays to enhance network performance rather than perceiving it as a detrimental factor like existing studies.

Remark 2: As shown in the memory-based consensus protocol (9), besides the real-time information, each agent is required to store its past state and share this memory information with its neighbors, which may occupy scarce storage spaces and impose additional communication burdens on networks in practice. In order to alleviate storage burdens and address additional communication demands, we can equip agents with extended physical registers and expand the transmission bandwidth.

A. Robustness

In this article, we consider the robustness of consensus networks to external disturbances. The disturbed multi-agent network (8) is shown as

$$x_i(t+1) = u_i(t) + \omega_i(t), i \in \mathcal{I}_n, \quad (12)$$

where $\omega_i(t) \in \mathbb{R}$ is the external disturbance of the i_{th} agent and we suppose that $\sum_{t=0}^{\infty} \omega_i^2(t) < \infty$. Note that disturbances may prevent the multi-agent network (12)-(9) from achieving consensus when $\omega_i(t) \neq 0$, $\exists i \in \mathcal{I}_n$. In order to measure the disagreement, we define the consensus error as $\epsilon_i(t) = x_i(t) - \frac{1}{n}\sum_{j=1}^{n} x_j(t), i \in \mathcal{I}_n$ whose compact form called the collective consensus error is given by

$$\epsilon(t) = x(t) - \frac{\mathbf{1}_n}{n}\sum_{j=1}^{n} x_j(t). \quad (13)$$

Substituting (9) into (12) and taking (13) as the nominal output yield the following augmented closed-loop system

$$\begin{cases} \bar{x}(t+1) = \Phi_n \bar{x}(t) + \begin{bmatrix} \mathbf{0} \\ I_n \end{bmatrix} \omega(t), \\ \epsilon(t) = \begin{bmatrix} \mathbf{0} & \Theta_n \end{bmatrix} \bar{x}(t), \end{cases} \quad (14)$$

where $\omega(t) = [\omega_1(t), \ldots, \omega_n(t)]^\top$.

Here, the robustness is measured by the ℓ_2 gain [39] from disturbance $\omega(t)$ to collective consensus error $\epsilon(t)$ of the system (14), which is defined by

$$\sup_{\substack{\omega(t)\neq 0 \\ \omega(t)\in\ell_2[0,\infty)}} \sqrt{\frac{\sum_{t=0}^{\infty} \epsilon^\top(t)\epsilon(t)}{\sum_{t=0}^{\infty} \omega^\top(t)\omega(t)}}.$$

Nevertheless, it is impossible to enumerate all disturbances in $\ell_2[0,\infty)$, making it unfeasible to directly use ℓ_2 gain to analyze the robustness. Denote the transfer matrix from $\omega(t)$ to $\epsilon(t)$ of system (14) as $T_{\alpha,\beta}(z)$. According to [37], the ℓ_2 gain is equivalent to the H_∞ norm of $T_{\alpha,\beta}(z)$, i.e., $\|T_{\alpha,\beta}(z)\|_\infty$. Then, as shown in Definition 1, we can determine $\|T_{\alpha,\beta}(z)\|_\infty$ through frequency-domain analysis. Note that smaller $\|T_{\alpha,\beta}(z)\|_\infty$ indicates better robustness. Hereafter, we use $\|T_{\alpha,\beta}(z)\|_\infty$ to characterize the robustness of multi-agent network (12)-(9) directly.

B. Scalability

The scalability refers to the capabilities of multi-agent networks to preserve the consensus property as the network scale grows.

Specifically, consider that the size of the multi-agent network (8)-(9) increases. As more agents join, undirected graphs of varying sizes constitute a *graph family* denoted as $\mathbb{G} = \{\mathcal{G}_{\iota_0}, \ldots, \mathcal{G}_{\iota_m}\}$, where $\iota_k \in \mathbb{N}^+$ $(k = 0, \ldots, m)$ represents the network size, each undirected graph $\mathcal{G}_{\iota_k}$ is connected, and ι_0 and ι_m $(\iota_0 < \iota_m)$ denote the incipient group size and the final group size, respectively. For generality, $\mathcal{G}_{\iota_k}$ needs not be a subgraph of $\mathcal{G}_{\iota_{k+1}}$ for all $k = 0, \ldots, m-1$. To avoid ambiguity, let $\lambda_j^{(\iota_k)}$, $j \in \{1, \ldots, \iota_k\}$ indicate the eigenvalues of the Laplacian matrix L_{ι_k} associated with the graph $\mathcal{G}_{\iota_k}$ in the family $\mathbb{G}$, which assume the form $0 = \lambda_1^{(\iota_k)} < \lambda_2^{(\iota_k)} \leq \cdots \leq \lambda_{\iota_k}^{(\iota_k)}$. Furthermore, it should be stressed that the memory coefficient α and the control gain β are preset and cannot be re-tuned throughout $\mathbb{G}$.

With these configurations, the scalable consensus is defined as follows.

Definition 2 (Scalable consensus): Given a family of connected undirected graphs $\mathbb{G}$, the multi-agent network (8)-(9) is said to achieve the scalable consensus on $\mathbb{G}$ if the consensus can be achieved on each graph in $\mathbb{G}$ for fixed memory coefficient α and control gain β.

Remark 3: The scalable consensus is also termed as scalable stability in [22] and the authors made some assumptions for the communication network. In contrast, we do not impose any assumptions or restrictions on graphs.

IV. Main Results

In this section, we first establish consensus criteria and determine the consensus state for the multi-agent network (8)-(9) on the fixed graph (see Section IV-A). Then, we build the analytic expression of $\|T_{\alpha,\beta}(z)\|_\infty$ and obtain the optimal control parameters in Section IV-B. Finally, based on the consensus criteria for the fixed graph, we show how to preset memory coefficient to improve the scalability and derive the optimal memory coefficient in Section IV-C. It is proved that the memory information can improve the robustness and the scalability simultaneously.

A. Consensus Analysis

Firstly, for the multi-agent network (11) on a fixed undirected graph $\mathcal{G}_n$, we should obtain the eigenvalues of Φ_n. Consider

$$\gamma I_{2n} - \Phi_n = \begin{bmatrix} \gamma I_n & -I_n \\ (\alpha-1)(I_n - \beta L_n) & \gamma I_n - \alpha(I_n - \beta L_n) \end{bmatrix}.$$

Since the Laplacian matrix L_n of the undirected graph $\mathcal{G}_n$ is symmetric, according to the spectral theorem [40], there exists an orthogonal matrix $V \in \mathbb{R}^{n\times n}$ such that $V^\top L_n V = \Lambda$, where $\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$ consists of all eigenvalues of L_n. In addition, matrices γI_n and $(\alpha-1)(I_n - \beta L_n)$ are commutative. Based on the above facts, the characteristic polynomial of Φ_n can be given by

$$\begin{aligned} &\det(\gamma I_{2n} - \Phi_n) \\ &= \left|\gamma(\gamma I_n - \alpha(I_n - \beta L_n)) - (1-\alpha)(I_n - \beta L_n)\right| \\ &= \left|\gamma^2 I_n - \alpha\gamma(I_n - \beta\Lambda) - (1-\alpha)(I_n - \beta\Lambda)\right| \\ &= \prod_{i=1}^{n} p(\gamma, \lambda_i), \end{aligned} \tag{15}$$

where $p(\gamma,\lambda_i) = \gamma^2 - \alpha(1-\beta\lambda_i)\gamma - (1-\alpha)(1-\beta\lambda_i)$. Thus, the eigenvalues of Φ_n are

$$\gamma_{i,1} = \frac{\alpha(1-\beta\lambda_i) + \sqrt{\alpha^2(1-\beta\lambda_i)^2 + 4(1-\alpha)(1-\beta\lambda_i)}}{2},$$
$$\gamma_{i,2} = \frac{\alpha(1-\beta\lambda_i) - \sqrt{\alpha^2(1-\beta\lambda_i)^2 + 4(1-\alpha)(1-\beta\lambda_i)}}{2},$$

where $\gamma_{i,1}$ and $\gamma_{i,2}$ are termed the eigenvalues of Φ_n associated with λ_i.

Then the following Lemma gives a necessary and sufficient condition of achieving consensus for multi-agent network (8) with the memory-based protocol (9).

Lemma 3: Multi-agent network (8)-(9) achieves consensus if and only if all eigenvalues of Φ_n excluding $\gamma_{1,1}$ and $\gamma_{1,2}$ are inside the unit circle in the complex plane.

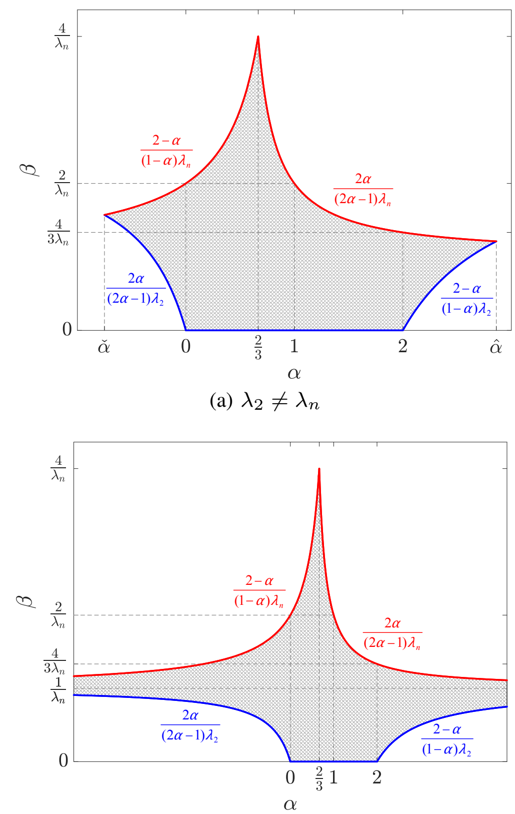
Proof: The proof is provided in Appendix A. ■

We further provide a more specific consensus criterion with respect to the control parameters and the graphical features for multi-agent network (8)-(9), which is shown as follows.

Theorem 1: Multi-agent network (8)-(9) achieves consensus if and only if the undirected graph $\mathcal{G}_n$ is connected and $(\alpha,\beta) \in \Omega = \bigcup_{j=1}^4 \Omega_j$, where

$$\Omega_1 = \left\{(\alpha,\beta)\ \middle|\ \alpha \le 0, \frac{2\alpha}{(2\alpha-1)\lambda_2} < \beta < \frac{2-\alpha}{(1-\alpha)\lambda_n}\right\},$$
$$\Omega_2 = \left\{(\alpha,\beta)\ \middle|\ 0 < \alpha \le \frac{2}{3}, 0 < \beta < \frac{2-\alpha}{(1-\alpha)\lambda_n}\right\},$$
$$\Omega_3 = \left\{(\alpha,\beta)\ \middle|\ \frac{2}{3} \le \alpha < 2, 0 < \beta < \frac{2\alpha}{(2\alpha-1)\lambda_n}\right\},$$
$$\Omega_4 = \left\{(\alpha,\beta)\ \middle|\ \alpha \ge 2, \frac{2-\alpha}{(1-\alpha)\lambda_2} < \beta < \frac{2\alpha}{(2\alpha-1)\lambda_n}\right\}.$$

Proof: The proof is given in Appendix B. ■



(a) $\lambda_2 \neq \lambda_n$

(b) $\lambda_2 = \lambda_n$

Fig. 1: Diagrams of the consensus region Ω.

We call Ω the consensus region. It is the set of parameter pairs (α,β) ensuring consensus. In order to better illustrate our results, we draw two cases of diagrams of Ω in Fig. 1. In both panels, the blue boundaries and the red boundaries respectively represent the lower bounds and the upper bounds of β given in Ω_i, $i = 1,2,3,4$, and the shadow areas sandwiched between them denote the consensus regions. In particular, when $\lambda_2 \neq \lambda_n$, the consensus region Ω is closed and it can be readily obtained that

$$\breve{\alpha} = \frac{2\lambda_n - 5\lambda_2 - \sqrt{(5\lambda_2 - 2\lambda_n)^2 + 16\lambda_2(\lambda_n - \lambda_2)}}{4(\lambda_n - \lambda_2)}$$

and

$$\hat{\alpha} = \frac{5\lambda_n - 2\lambda_2 + \sqrt{(2\lambda_2 - 5\lambda_n)^2 - 16\lambda_n(\lambda_n - \lambda_2)}}{4(\lambda_n - \lambda_2)}.$$

It should be noted that if the initial conditions are set as $x(-1) = x(0)$, we can derive that (30), (31), and (32) all equal to $\lim_{t\to\infty} \bar{x}(t) = \frac{1}{n}\sum_{i=1}^{n} x_i(0) \cdot \mathbf{1}_{2n}$. Thus, we have $\lim_{t\to\infty} x_i(t) = \frac{1}{n}\sum_{i=1}^{n} x_i(0)$, $\forall i \in \mathcal{I}_n$, which means that multi-agent network (8)-(9) achieves average consensus. Generally, when $x(-1) \neq x(0)$, it is observed from (30), (31), and (32) that multi-agent network (8)-(9) exhibits different consensus behaviors. Specifically,

- If $0 < \alpha < 2$, the consensus value is constant;
- If $\alpha = 0$, multi-agent network (8)-(9) only uses memory information $\psi_i(t-1)$ and achieves the periodically oscillatory consensus, i.e., the consensus value repeatedly fluctuates between $\frac{1}{n}\sum_{i=1}^{n} x_i(-1)$ and $\frac{1}{n}\sum_{i=1}^{n} x_i(0)$ as $t \to \infty$;
- If $\alpha < 0$, the consensus value is in oscillatory exponential growth;
- If $\alpha = 2$, we can get $x_i(t) \to \frac{1}{n}\sum_{i=1}^{n} x_i(0) + \frac{t}{n}\sum_{i=1}^{n}[x_i(0) - x_i(-1)]$ for all $i = 1, \ldots, n$ as $t \to \infty$. Note that it resembles the state trajectory of the second-order multi-agent consensus network with relative velocity protocol [41], though all agents in multi-agent network (8)-(9) employ the first-order dynamics. Thus, we call it quasi-second-order consensus;
- If $\alpha > 2$, the consensus value exponentially grows.

Remark 4: We are surprised that the memory information facilitates producing such manifold consensus behaviors, which is in sharp contrast with tedious consensus phenomena with memoryless information in most literature.

B. Robustness Analysis

In this subsection, we first build the analytic expression of $\|T_{\alpha,\beta}(z)\|_\infty$. Then, the optimal parameters α and β that make the multi-agent network (12)-(9) obtain the optimal robustness are given. Finally, we show how to tune the memory coefficient to improve the robustness.

Prior to stating the main results, we propose a useful lemma.

Lemma 4: Let $g_1(t) = \frac{1}{(\beta t)^2}$, $g_2(t) = \frac{1}{[2\alpha+(1-2\alpha)\beta t]^2}$, and $g_3(t) = \frac{1}{\Delta_1 \cdot \Delta_2^2}$, where $\beta, t > 0$, α is constant, $\Delta_1 = [1 + \frac{\alpha^2}{4(1-\alpha)}(1-\beta t)]$, and $\Delta_2 = \alpha - 2 + (1-\alpha)\beta t$. For convenience, denote $\check{\theta}(t) = \frac{-\alpha^2-4\alpha+4}{\alpha(1-\alpha)t}$, $\hat{\theta}(t) = \frac{-\alpha^2+4\alpha-4}{\alpha(1-\alpha)t}$, $\check{\eta}(t) = \frac{2\alpha}{(2\alpha-1)t}$, and $\hat{\eta}(t) = \frac{2-\alpha}{(1-\alpha)t}$. Then, we can obtain

1) $g_1(t)$ is decreasing on $(0, +\infty)$;
2) $g_2(t)$ is decreasing on
$$\begin{cases} (\check{\eta}(\beta), +\infty) & \text{if } \alpha \leq 0, \\ (0, \hat{\eta}(\beta)) & \text{if } 0 < \alpha < \frac{1}{2}; \end{cases}$$
3) $g_2(t)$ is constant if $\alpha = \frac{1}{2}$;
4) $g_2(t)$ is increasing on
$$\begin{cases} (0, \hat{\eta}(\beta)) & \text{if } \frac{1}{2} < \alpha \leq \frac{2}{3}, \\ (0, \check{\eta}(\beta)) & \text{if } \frac{2}{3} < \alpha \leq 2, \\ (\hat{\eta}(\beta), \check{\eta}(\beta)) & \text{if } \alpha > 2; \end{cases}$$
5) $g_3(t)$ is increasing on
$$\begin{cases} (\check{\eta}(\beta), \hat{\eta}(\beta)) & \text{if } \alpha \leq 0, \\ (0, \hat{\eta}(\beta)) & \text{if } 0 < \alpha \leq \frac{2}{3}, \\ (0, \check{\theta}(\beta)) & \text{if } \frac{2}{3} < \alpha < 1; \end{cases}$$
6) $g_3(t)$ is decreasing on
$$\begin{cases} (\hat{\theta}(\beta), +\infty) & \text{if } 1 < \alpha \leq 2, \\ (\hat{\eta}(\beta), \check{\eta}(\beta)) & \text{if } \alpha > 2. \end{cases}$$

Proof: The monotonicity of $g_1(t)$ and $g_2(t)$ can be readily proved by general analysis for their first derivations. The results for $g_3(t)$ can be obtained by respectively studying the properties of the elementary functions Δ_1 and Δ_2. ■

The following Theorem establishes the analytic expression of $\|T_{\alpha,\beta}(z)\|_\infty$.

Theorem 2: Consider the multi-agent network (12)-(9) on a connected undirected graph $\mathcal{G}_n$ with the Laplacian matrix L_n. If $(\alpha, \beta) \in \Omega$, we get

$$\|T_{\alpha,\beta}(z)\|_\infty = \begin{cases} \max\{\sqrt{g_2(\lambda_2)}, \sqrt{g_3(\lambda_n)}\} & \text{if } \alpha \leq 0, \\ \max\{\sqrt{g_1(\lambda_2)}, \sqrt{g_3(\lambda_n)}\} & \text{if } 0 < \alpha \leq \frac{2}{3}, \\ \mathcal{Q}_1(\alpha, \beta) & \text{if } \frac{2}{3} < \alpha \leq \alpha_1, \\ \max\{\sqrt{g_1(\lambda_2)}, \sqrt{g_2(\lambda_n)}\} & \text{if } \alpha_1 < \alpha \leq \alpha_2, \\ \mathcal{Q}_2(\alpha, \beta) & \text{if } \alpha_2 < \alpha < 2, \\ \max\{\sqrt{g_3(\lambda_2)}, \sqrt{g_2(\lambda_n)}\} & \text{if } \alpha \geq 2, \end{cases} \tag{16}$$

where

$$\alpha_1 = \frac{6\lambda_n - 2\lambda_2 + 2\sqrt{2\lambda_2(\lambda_2+\lambda_n)}}{9\lambda_n + \lambda_2} < 1,$$

$\alpha_2 \in (1, 2)$ is the unique real root of the cubic equation $(2\lambda_2 - 2\lambda_n)\alpha^3 + (9\lambda_n - 3\lambda_2)\alpha^2 + (4\lambda_2 - 12\lambda_n)\alpha + 4\lambda_n - 4\lambda_2 = 0$,

$$\mathcal{Q}_1(\alpha, \beta) = \begin{cases} \max\{\sqrt{g_1(\lambda_2)}, \sqrt{g_3(\lambda_n)}\} & \text{if } 0 < \beta \leq \check{\theta}(\lambda_n), \\ \sqrt{g_2(\lambda_n)} & \text{if } \check{\theta}(\lambda_n) < \beta < \check{\eta}(\lambda_n), \end{cases}$$

and

$$\mathcal{Q}_2(\alpha, \beta) = \begin{cases} \sqrt{g_1(\lambda_2)} & \text{if } 0 < \beta \leq \hat{\theta}(\lambda_2), \\ \max\{\sqrt{g_3(\lambda_2)}, \sqrt{g_2(\lambda_n)}\} & \text{if } \hat{\theta}(\lambda_2) < \beta < \check{\eta}(\lambda_n). \end{cases}$$

Proof: The proof is given in Appendix C. ■

As shown in Theorem 2, the robustness is entirely determined by parameters α, β and graphic features λ_2, λ_n. However, we are more interested in optimizing robustness by tuning the memory coefficient and the control gain, and we want to ascertain whether there are the optimal robustness and the optimal parameters.

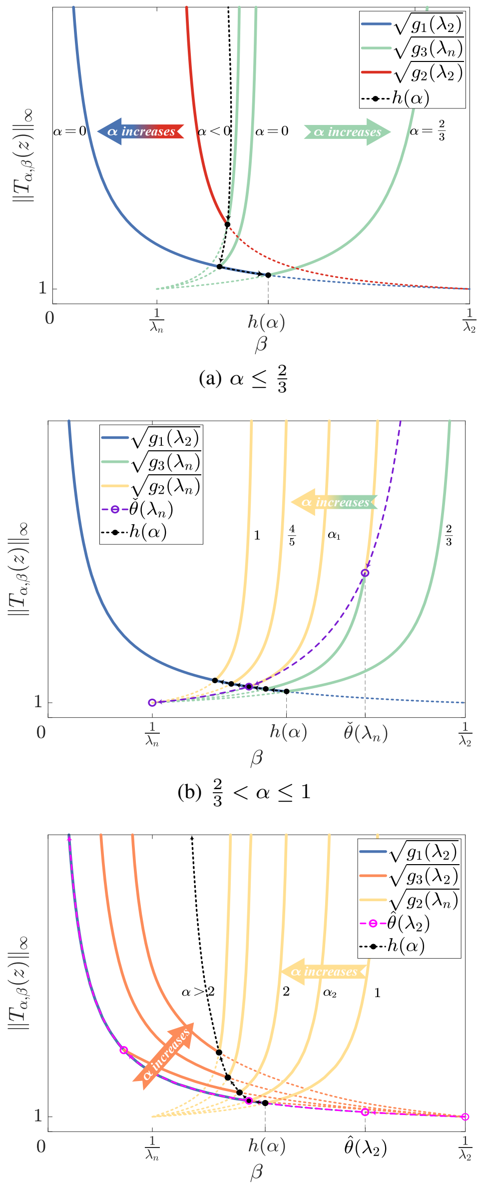
Definition 3: Consider the multi-agent network (12)-(9) on a connected undirected graph $\mathcal{G}_n$. The multi-agent network (12)-(9) is said to achieve the optimal robustness if the memory coefficient α and the control gain β solve the optimal problem

$$\min_{(\alpha,\beta)\in\Omega} \|T_{\alpha,\beta}(z)\|_\infty.$$

Moreover, let $\|T_{\alpha,\beta}(z)\|_\infty^* = \min_{(\alpha,\beta)\in\Omega} \|T_{\alpha,\beta}(z)\|_\infty$ and

$$(\alpha_r^*, \beta_r^*) = \arg\min_{(\alpha,\beta)\in\Omega} \|T_{\alpha,\beta}(z)\|_\infty,$$

where α_r^* is called the optimal memory coefficient and β_r^* indicates the optimal control gain.



(a) $\alpha \le \frac{2}{3}$

(b) $\frac{2}{3} < \alpha \le 1$

(c) $\alpha \ge 1$

Fig. 2: The trends of $\|T_{\alpha,\beta}(z)\|_\infty$ with respect to β.

Theorem 3: For a connected undirected graph $\mathcal{G}_n$, there exist unique optimal memory coefficient and control gain

$$(\alpha_r^*, \beta_r^*) = \left(\frac{2}{3}, \frac{4}{\lambda_n} - \frac{9\lambda_2^2 + 3\sqrt[3]{\chi_1} + 3\sqrt[3]{\chi_2}}{\lambda_n^3}\right)$$

such that the multi-agent network (12)-(9) obtains the optimal robustness

$$\|T_{\alpha,\beta}(z)\|_\infty^* = \frac{\lambda_n^3}{\lambda_2(4\lambda_n^2 - 9\lambda_2^2 - 3\sqrt[3]{\chi_1} - 3\sqrt[3]{\chi_2})}, \tag{17}$$

where

$$\chi_i = 27\lambda_2^6 \pm 8\lambda_2^2\lambda_n^3\sqrt{\lambda_n^2 - \lambda_2^2} + 8\lambda_2^2\lambda_n^4 - 36\lambda_2^4\lambda_n^2, i = 1, 2.$$

Proof: As shown in Theorem 1, $\|T_{\alpha,\beta}(z)\|_\infty$ can be deemed as a function with respect to α and β. Let $h(\alpha) = \arg\min_\beta \|T_{\alpha,\beta}(z)\|_\infty$ be the optimal control gain for a given α. We have to analyze the optimal robustness on each Ω_i, $i = 1, \ldots, 4$. However, the proof is quite similar. For saving space, we only show the proof for two representative cases.

Case (a): $0 \le \alpha \le \frac{2}{3}$, $0 < \beta < \hat{\eta}(\lambda_n)$

In this case, we have $\|T_{\alpha,\beta}(z)\|_\infty = \max\{\sqrt{g_1(\lambda_2)}, \sqrt{g_3(\lambda_n)}\}$. Denote $\rho_1(\alpha,\beta) = \sqrt{g_1(\lambda_2)}$ and $\rho_2(\alpha,\beta) = \sqrt{g_3(\lambda_n)}$. Along β-axis, $\rho_1(\alpha,\beta)$ and $\rho_2(\alpha,\beta)$ are decreasing and increasing, respectively. It follows from $\frac{1}{\lambda_n} < \frac{1}{\lambda_2}$, $\frac{1}{\lambda_n} < \hat{\eta}(\lambda_n)$, $\lim_{\beta\to 0}\rho_1(\alpha,\beta) = +\infty$, $\lim_{\beta\to\hat{\eta}(\lambda_n)}\rho_2(\alpha,\beta) = +\infty$, and $\rho_1(\alpha, \frac{1}{\lambda_2}) = \rho_2(\alpha, \frac{1}{\lambda_n}) = 1$ that $\rho_1(\alpha,\beta) - \rho_2(\alpha,\beta) = 0$ has a unique positive real root which is the minimum point $\beta = h(\alpha)$ of $\|T_{\alpha,\beta}(z)\|_\infty$ for a given α. Thus, $\|T_{\alpha,\beta}(z)\|_\infty$ can be rewritten as

$$\|T_{\alpha,\beta}(z)\|_\infty = \begin{cases} \sqrt{g_1(\lambda_2)} & \text{if } 0 < \beta \le h(\alpha), \\ \sqrt{g_3(\lambda_n)} & \text{if } h(\alpha) \le \beta < \hat{\eta}(\lambda_n). \end{cases} \tag{18}$$

Obviously, we have $\min_{\beta\in(0,\hat{\eta}(\lambda_n))}\|T_{\alpha,\beta}(z)\|_\infty = \|T_{\alpha,h(\alpha)}(z)\|_\infty$ that can be considered as the function of α. We can verify that $\rho_1(\alpha,\beta)$ is constant with respect to α and $\rho_2(\alpha,\beta)$ is decreasing with α when $\frac{1}{\lambda_n} < \beta < \hat{\eta}(\lambda_n)$, which means that $\|T_{\alpha,h(\alpha)}(z)\|_\infty$ decreases as α grows (see Fig. 2a).

Case (b): $\frac{2}{3} < \alpha \le 1$, $0 < \beta < \check{\eta}(\lambda_n)$

Combining with Theorem 2 and similar analysis in Case (a), one can deduce that

$$h(\alpha) = \begin{cases} h_1(\alpha) & \text{if } \frac{2}{3} < \alpha \le \alpha_1, \\ h_2(\alpha) & \text{if } \alpha_1 < \alpha \le 1, \end{cases}$$

where $h_1(\alpha)$ is the unique real root of $g_1(\lambda_2) = g_3(\lambda_n)$ and $h_2(\alpha)$ is the unique real root of $g_1(\lambda_2) = g_2(\lambda_n)$. Therefore, when $\frac{2}{3} < \alpha \le \alpha_1$, we have

$$\|T_{\alpha,\beta}(z)\|_\infty = \begin{cases} \sqrt{g_1(\lambda_2)} & \text{if } 0 < \beta \le h_1(\alpha), \\ \sqrt{g_3(\lambda_n)} & \text{if } h_1(\alpha) \le \beta \le \check{\theta}(\lambda_n), \\ \sqrt{g_2(\lambda_n)} & \text{if } \check{\theta}(\lambda_n) \le \beta < \check{\eta}(\lambda_n); \end{cases} \tag{19}$$

when $\alpha_1 < \alpha \le 1$, we can get

$$\|T_{\alpha,\beta}(z)\|_\infty = \begin{cases} \sqrt{g_1(\lambda_2)} & \text{if } 0 < \beta \le h_2(\alpha), \\ \sqrt{g_2(\lambda_n)} & \text{if } h_2(\alpha) \le \beta < \check{\eta}(\lambda_n). \end{cases} \tag{20}$$

Different with Case (a), $\|T_{\alpha,h(\alpha)}(z)\|_\infty$ increases as α grows because $\sqrt{g_2(\lambda_n)}$ and $\sqrt{g_3(\lambda_n)}$ are both increasing with α, which is well explained in Fig. 2b. The proof of cases $\alpha < 0$ and $\alpha > 1$ are similar.

To sum up, $\|T_{\alpha,h(\alpha)}(z)\|_\infty$ is monotonously decreasing and increasing when $\alpha \le \frac{2}{3}$ and $\alpha \ge \frac{2}{3}$, respectively. Therefore, it follows that $(\alpha_r^*, \beta_r^*) = (\frac{2}{3}, h(\frac{2}{3}))$. Note that $h(\frac{2}{3})$ is the unique real root of $\sqrt{g_1(\lambda_2)} = \sqrt{g_3(\lambda_n)}$ that is equivalent to the following cubic equation

$$\lambda_n^3\beta^3 - (12\lambda_n^2 - 27\lambda_2^2)\beta^2 + 48\lambda_n\beta - 64 = 0. \tag{21}$$

Then, β_r^* can be readily obtained by invoking the root-finding formula. Finally, substituting (α_r^*, β_r^*) into $\sqrt{g_1(\lambda_2)}$ yields the optimal robustness. ■

As discussed in the proof of Theorem 3, we can get more explicit expressions for $\|T_{\alpha,\beta}(z)\|_\infty$, (e.g., (18), (19), and (20)), which are visualized in Fig. 3. Specifically, we plot the two-dimensional surface $\Pi := \{(\alpha,\beta,\rho) \in \mathbb{R}^3 \mid (\alpha,\beta) \in$

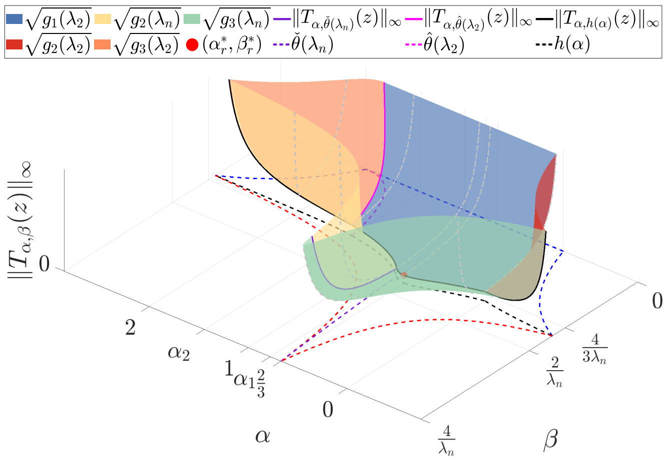


Fig. 3: The surface of $\|T_{\alpha,\beta}(z)\|_\infty$ with respect to α and β in Ω when $\lambda_2 \neq \lambda_n$.

$\Omega, \rho = \|T_{\alpha,\beta}(z)\|_\infty\}$ of $\|T_{\alpha,\beta}(z)\|_\infty$ with respect to α and β in Fig. 3, where analytic expressions of $\|T_{\alpha,\beta}(z)\|_\infty$ are distinguished by different colors. The gray dash lines represent the boundaries of different regions divided in terms of α, as seen in (16). Furthermore, the purple solid curve and its projection on Ω, i.e., the purple dash curve, denote the trend of $\|T_{\alpha,\beta}(z)\|_\infty$ subject to $\beta = \check{\theta}(\lambda_n)$ and the trajectory of $\check{\theta}(\lambda_n)$, respectively. The counterparts for $\hat{\theta}(\lambda_2)$ are shown as the pink curves. We mark the optimal parameter pair (α_r^*, β_r^*) by the red circle point. Recall that $h(\alpha) = \arg\min_\beta \|T_{\alpha,\beta}(z)\|_\infty$ is the optimal control gain for a given α. Then, the black solid curve and its projection on Ω, i.e., the black dash curve, indicate the trend of $\|T_{\alpha,\beta}(z)\|_\infty$ subject to $\beta = h(\alpha)$ and the trajectory of $h(\alpha)$, respectively. It is worth mentioning that Fig. 2 depicts the α-axial cross-section curves of the surface $\|T_{\alpha,\beta}(z)\|_\infty$ given in Fig. 3.

Remark 5: We further derive concise lower bound and upper bound of the optimal robustness (17). On the one hand, as evidence from Fig. 2b, $\|T_{\alpha,\beta}(z)\|_\infty^* = \|T_{\frac{2}{3},h(\frac{2}{3})}(z)\|_\infty > \|T_{\frac{2}{3},\frac{1}{\lambda_2}}(z)\|_\infty = 1$ is straightforward. On the other hand, if $\lambda_2 \neq \lambda_n$, according to (22) and Fig. 2b, we have $\|T_{\alpha,\beta}(z)\|_\infty^* < \min_{\beta\in(0,\frac{2}{\lambda_n})} \|T_{1,\beta}(z)\|_\infty = \|T_{1,\frac{2}{\lambda_2+\lambda_n}}(z)\|_\infty = \frac{\lambda_2+\lambda_n}{2\lambda_2}$. Otherwise, substituting $\lambda_2 = \lambda_n$ into (17) gives rise to $\|T_{\alpha,\beta}(z)\|_\infty^* = \frac{\lambda_2+\lambda_n}{2\lambda_2}$.

Remark 6: It is worth noting that the optimal memory coefficient is an invariant, which is in sharp contrast with the optimal control parameter. On the one hand, the ℓ_2 gain is used to measure the robustness of consensus networks in this paper, which inherently determines such outcome. Utilizing alternative performance indices may yield different results. On the other hand, it may be related to the way of using memory information. In protocol (9), for example, we replace the memory information $\psi_i(t-1)$ by the counterpart at earlier time-step. Then, the optimal memory coefficient, if any, may be reliant on the whole Laplacian spectrum.

In particular, $\|T_{1,\beta}(z)\|_\infty$ (i.e., $\alpha = 1$) reflects the robustness of the multi-agent network (12)-(9) without memory information, which is given by

$$\|T_{1,\beta}(z)\|_\infty = \begin{cases} \frac{1}{\beta\lambda_2} & \text{if } 0 < \beta \leq \frac{2}{\lambda_2+\lambda_n}, \\ \frac{1}{2-\beta\lambda_n} & \text{if } \frac{2}{\lambda_2+\lambda_n} < \beta < \frac{2}{\lambda_n}, \end{cases} \tag{22}$$

according to Theorem 2. It is worth mentioning that the similar form of (22) was reported in [17] (see the proof of Corollary 6 therein) that examined the robustness of sampled-data multi-agent consensus networks.

If we want to validate that the memory information can improve the robustness, we just need to prove $\|T_{\alpha,\beta}(z)\|_\infty \leq \|T_{1,\beta}(z)\|_\infty$, $\forall \beta \in (0, \frac{2}{\lambda_n})$. Accordingly, we have the following conclusion.

Corollary 1: Consider the multi-agent network (12)-(9) on a connected undirected graph $\mathcal{G}_n$. Then, $\|T_{\alpha,\beta}(z)\|_\infty \leq \|T_{1,\beta}(z)\|_\infty$, $\forall \beta \in (0, \frac{2}{\lambda_n})$ if and only if $0 < \alpha < 1$.

Proof: The sufficiency is straightforward from Fig. 2 and the monotonicity of $g_j(\lambda_i)$ $(j = 1, 2, 3)$ regarding to α, which has been discussed in the proof of Theorem 3. Similarly, we can also demonstrate that $\|T_{\alpha,\beta}(z)\|_\infty \geq \|T_{1,\beta}(z)\|_\infty$ as long as $\alpha \leq 0$ or $\alpha \geq 1$. Therefore, the necessity is proved. ■

Remark 7: We also obtain a more explicit result for the comparison between $\|T_{\alpha,\beta}(z)\|_\infty$ and $\|T_{1,\beta}(z)\|_\infty$ shown in Corollary 1. If $0 < \alpha < 1$, it follows from (22), Figs. 2a, 2b and $\|T_{0,\beta}(z)\|_\infty \equiv \|T_{1,\beta}(z)\|_\infty$ that

$$\begin{cases} \|T_{\alpha,\beta}(z)\|_\infty = \|T_{1,\beta}(z)\|_\infty & \text{if } 0 < \beta \leq \frac{2}{\lambda_2+\lambda_n}, \\ \|T_{\alpha,\beta}(z)\|_\infty < \|T_{1,\beta}(z)\|_\infty & \text{if } \frac{2}{\lambda_2+\lambda_n} < \beta < \frac{2}{\lambda_n}. \end{cases}$$

C. Scalability Analysis

Building on previous consensus analysis for fixed graphs in Section IV-A, we reinspect the consensus from the perspective of scalability and explore the role of memory information on the scalability.

As stated in Definition 2, it is trivial to verify the scalable consensus — enumerating each graph in the graph family $\mathbb{G}$ and following the consensus criterion given in Theorem 1. However, the point is that as the network size increases and the network topology changes, the non-zero minimum and maximum eigenvalues of the Laplacian matrices vary, thereby the consensus may fail without re-tuning the parameters α and β. Naturally, we say that better scalability means that the consensus can be preserved to more graphs. In this case, it is prominent to study how the preset and fixed parameter pair (α, β), especially the memory coefficient, affects the scalability. Here, we display some instructive examples.

Consider the following families of undirected graphs with $m \geq 3$ vertices and unit edge weights: the family of complete graphs $\mathbb{K} = \{K_3, \ldots, K_m\}$, the family of star graphs $\mathbb{S} = \{S_3, \ldots, S_m\}$, the family of path graph $\mathbb{P} = \{P_3, \ldots, P_m\}$, and the family of $2d$-regular ring lattices $\mathbb{C}^d = \{C_{2d+1}^d, \ldots, C_m^d\}$, where each C_i^d $(2d+1 \leq i \leq m)$ is highly structured network with agents arranged on a ring, each connecting to its $2d$ nearest neighbors. More details about these graphs are available in [42]. For above graphs, Table I lists the analytic expressions of the non-zero minimum and maximum eigenvalues of Laplacian matrices.

TABLE I: The analytic expressions of the non-zero minimum and maximum eigenvalues of Laplacian matrices.

Graph	Value of $\lambda_2^{(k)}$	Value of $\lambda_k^{(k)}$
K_k	k	k
S_k	1	k
P_k	$4\sin^2\left(\frac{\pi}{2k}\right)$	$4\sin^2(\frac{k-1}{2k}\pi)$
C_k^1	$3-\dfrac{\sin(\frac{3\pi}{k})}{\sin(\frac{\pi}{k})}$	$3-\dfrac{\sin(\lfloor\frac{k}{2}\rfloor\cdot\frac{3\pi}{k})}{\sin(\lfloor\frac{k}{2}\rfloor\cdot\frac{\pi}{k})}$
C_k^2	$5-\dfrac{\sin(\frac{5\pi}{k})}{\sin(\frac{\pi}{k})}$	$\varpi_2(k)^\dagger$
C_k^3	$7-\dfrac{\sin(\frac{7\pi}{k})}{\sin(\frac{\pi}{k})}$	$\varpi_3(k)$

† $\varpi_d(k)=\max_{j=2,\ldots,k}\left\{2d+1-\dfrac{\sin\left[\frac{(2d+1)(j-1)\pi}{k}\right]}{\sin\left[\frac{(j-1)\pi}{k}\right]}\right\}$, $d=2,3$.

Then we adjust memory coefficient under the same control gain to observe the effect of memory information on scalability. Combining Table I with the derived consensus criteria in Theorem 1, all graph families of achieving scalable consensus are summarized in Table II.

As illustrated in Table II, we find three striking features:

- For the same control gain and $\alpha>\frac{2}{3}$ ($\alpha<\frac{2}{3}$), consensus can be preserved to larger group size as the memory coefficient α decreases (increases). The trends are most evident on $\mathbb{C}^d$;
- For each type of graph families, consensus can be always preserved over the largest scales if we preset the memory coefficient as $\alpha=\frac{2}{3}$;
- For each type of graph families under the same control gain β, scalability is exactly the same for both $\alpha=0$ and $\alpha=1$.

These phenomena result from the variation of the consensus region. As seen in Fig. 1, when $\alpha>\frac{2}{3}$ ($\alpha<\frac{2}{3}$), the restriction of the control gain β ensuring consensus gradually becomes less conservative as α decreases (increases). Specially, the consensus region under $\alpha=0$ is identical to the counterpart under $\alpha=1$. Furthermore, the consensus region regarding β always remains maximum under $\alpha=\frac{2}{3}$.

Based on above observations and Fig. 1, we can summarize a general scenario. For any connected undirected graph, let $\Pi_1=\{(\alpha,\beta)\in\Omega\mid\alpha_0\in\mathbb{R},\alpha=\alpha_0\}$ and $\Pi_2=\{(\alpha,\beta)\in\Omega\mid\alpha_0'\in\mathbb{R},\alpha=\alpha_0'\}$ be two subsets of the consensus region, where $\alpha_0\neq\alpha_0'$, $\max_{\beta\in\Pi_1}\geq\max_{\beta\in\Pi_2}$, and $\min_{\beta\in\Pi_1}\leq\min_{\beta\in\Pi_2}$. It means that the range of β under $\alpha=\alpha_0$ that guarantees consensus completely covers the counterpart under $\alpha=\alpha_0'$. Therefore, if the multi-agent network (8)-(9) achieves scalable consensus on a graph family $\mathbb{G}$ with the fixed parameter pair (α_0',β_0), then the scalable consensus can be still preserved on $\mathbb{G}$ with (α_0,β_0) but not necessarily vice versa. In other words, if the memory coefficient is selected as α_0, it is more flexible to preset control gain β such that the consensus can be extended to more graphs.

Based on the above analysis and Fig. 1, we eventually have the following statement.

Proposition 1: Compared with the memoryless case (i.e., $\alpha=1$), memory information improves the scalability of the multi-agent network (8) as long as $0<\alpha<1$. Furthermore, the optimal memory coefficient is $\alpha_s^*=\frac{2}{3}$ in terms of scalability.

Remark 8: Our results provide a one-size-fits-all approach to improve the scalability rather than only for a specific graph family. Essentially, memory information enables the memoryless multi-agent network (8) that was originally unable to achieve consensus to recapture the consensus, thereby the scalability is enhanced. Similarly, we can also conclude that compared with the memoryless case, memory information deteriorates the scalability of the multi-agent network (8) if and only if $\alpha<0$ or $\alpha>2$. However, the scalabilities under $\alpha<0$ and $\alpha>2$ are basically not comparable because on most network topologies, the consensus regions regarding β under the two cases are not subsets of each other. The evidence is straightforward from Table II. Note that if $(\alpha,\beta)=(-\frac{1}{3},\frac{1}{3})$, the scalable consensus can be achieved on $\{C_3^1,C_4^1,C_5^1\}$, but cannot be reached on $\{C_7^1\}$. On the contrary, if $(\alpha,\beta)=(\frac{11}{5},\frac{1}{3})$, the scalable consensus is available on $\{C_3^1,C_5^1,C_7^1\}$, but cannot achieved on $\{C_4^1\}$.

V. Discussion

A. The Positive Correlation Between Robustness and Scalability With Respect to the Memory Coefficient

As shown in Section IV-C, if $\alpha<\frac{2}{3}$, the consensus convergence interval for β gradually expands as α increases, which means that the scalability is gradually enhanced. In contrast, the scalability deteriorates as α grows when $\alpha>\frac{2}{3}$. On the other hand, as shown in Fig. 2, $g_2(\lambda_i)$ and $g_3(\lambda_i)$ are both decreasing with α when $\alpha<\frac{2}{3}$, which implies that the robustness is improving with α. Similarly, if $\alpha>\frac{2}{3}$, we can verify that $g_2(\lambda_i)$ and $g_3(\lambda_i)$ both increase as α grows, thereby the robustness gradually deteriorates. Obviously, robustness and scalability are positively correlated as the memory coefficient changes.

B. Trade-Offs Between Memory Usage and Performance Gains

Following the above discussion and combining with Corollary 1 and Proposition 1, we can observe the trade-offs between memory usage and performance gains. As shown in (9), memory coefficient serves as a pivotal parameter to weigh the real-time information and the memory information. Both robustness and scalability are improved as long as the memory coefficient satisfies $0<\alpha<1$. Furthermore, the optimal memory coefficient is $\alpha=\frac{2}{3}$ for both robustness and scalability, which means that only moderate usage of memory information is more beneficial to performance gains. In contrast, incorporating memory information into consensus protocol (9) by setting $\alpha<0$ or $\alpha>2$ is detrimental to performance gains.

C. Memory Information of Multiple Time-Steps

In (9), the multi-agent network evolves by only invoking previous two time-steps information. A natural issue comes

TABLE II: Scalable graph families under various parameter pairs.

β	α	$\mathbb{K}$	$\mathbb{S}$	$\mathbb{P}$	$\mathbb{C}^1$	$\mathbb{C}^2$	$\mathbb{C}^3$
$\frac{1}{3}$	$-\frac{1}{3}$	$\{K_3, K_4\}$	✗	✗	$\{C_3^1, C_4^1, C_5^1\}$	$\{C_5^2\}$	✗
	0	$\{K_3, K_4, K_5\}$	$\{S_3, S_4, S_5\}$	$\{P_3, \dots\}$	$\{C_3^1, \dots\}$	$\{C_5^2\}$	✗
	$\frac{2}{3}$	$\{K_3, \dots, K_{11}\}$	$\{S_3, \dots, S_{11}\}$	$\{P_3, \dots\}$	$\{C_3^1, \dots\}$	$\{C_5^2, \dots\}$	$\{C_7^3, \dots\}$
	1	$\{K_3, K_4, K_5\}$	$\{S_3, S_4, S_5\}$	$\{P_3, \dots\}$	$\{C_3^1, \dots\}$	$\{C_5^2\}$	✗
	2	$\{K_3\}$	$\{S_3\}$	$\{P_3, \dots\}$	$\{C_3^1, C_5^1, C_7^1, \dots\}$	✗	✗
	$\frac{11}{5}$	$\{K_3\}$	$\{S_3\}$	$\{P_3, P_4\}$	$\{C_3^1, C_5^1, C_7^1\}$	✗	✗
$\frac{1}{2}$	$-\frac{1}{3}$	$\{K_3\}$	$\{S_3\}$	$\{P_3\}$	$\{C_3^1\}$	✗	✗
	0	$\{K_3\}$	$\{S_3\}$	$\{P_3, \dots\}$	$\{C_3^1, C_5^1, C_7^1, \dots\}$	✗	✗
	$\frac{2}{3}$	$\{K_3, \dots, K_7\}$	$\{S_3, \dots, S_7\}$	$\{P_3, \dots\}$	$\{C_3^1, \dots\}$	$\{C_5^2, \dots\}$	$\{C_7^3\}$
	1	$\{K_3\}$	$\{S_3\}$	$\{P_3, \dots\}$	$\{C_3^1, C_5^1, C_7^1, \dots\}$	✗	✗
	2	✗	✗	✗	✗	✗	✗
	$\frac{11}{5}$	✗	✗	✗	✗	✗	✗

to the fore — whether more steps of memory information could further improve the consensus performance. Although some state-of-the-art consensus algorithms with memory information of multiple time-steps have been developed [32]–[35], the effect of the amount of memory information was rarely discussed. It is intuitively true that more information is more beneficial to the performance. But a recent work [36] claimed that "*adding more taps of past information from neighbors cannot improve the rate (i.e., the convergence rate of consensus) further*". On this account, we believe that this issue should be extensively studied more in the context of multifarious performance for multi-agent networks, e.g., the robustness and scalability considered in this paper.

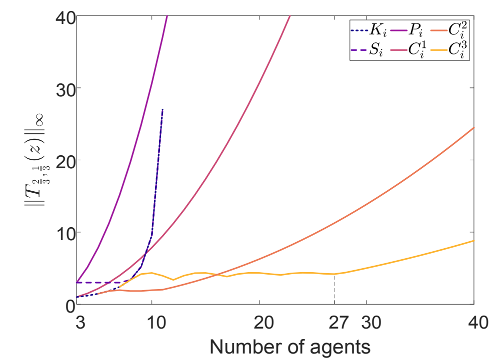


Fig. 4: The trends of $\|T_{\frac{2}{3},\frac{1}{3}}(z)\|_\infty$ on various graph families.

D. Performance Guarantee as the Multi-Agent Network Grows

Note that our robustness analysis is on fixed graphs, while the scalability analysis involves kind of long-term performance over a series of graphs. Although it has been known from Theorem 3 that we can select the optimal control parameters for any graphs, it is no more feasible under the rule of scalability — preset and fixed control parameters. One hopes that the consensus network can maintain good performance as the network grows [23] . Therefore, as the network size increases, we further test the trends of robustness $\|T_{\alpha,\beta}(z)\|_\infty$ with a preset fixed parameter pair $(\frac{2}{3}, \frac{1}{3})$ on the well-known graph families mentioned in Section IV-C. Results can be found in Fig. 4. It is shown that on complete graph families and star graph families, the consensus of multi-agent network (8)-(9) can be preserved up to $i = 11$ without re-tuning the parameters. By contrast, multi-agent network (8)-(9) is able to achieve scalable consensus on other graph families of any network size. More importantly, we observe from Fig. 4 that although robustness is basically getting worse on each graph, the robustness $\|T_{\frac{2}{3},\frac{1}{3}}(z)\|_\infty$ on 6-regular ring lattice C_i^3 is not greater than 4.34 if $i \leq 27$. This phenomenon appears to facilitate designing the topology structure of scalable consensus network with performance guarantee, and deserves to be further investigated.

E. Potential Index to Evaluate the Performance of Multi-Agent Networks

The consensus region Ω plays a vital role in our performance analysis. Although this concept was first reported in [43], we can reinspect it under the framework of scalability. Here, we consider the area of Ω ($\lambda_2 \neq \lambda_n$), which can be computed by

$$
\begin{aligned}
\mathcal{S}_\Omega =& \int_{\check{\alpha}}^{0} \left[\tfrac{2-\alpha}{(1-\alpha)\lambda_n} - \tfrac{2\alpha}{(2\alpha-1)\lambda_2}\right] \mathrm{d}\alpha + \int_0^{\frac{2}{3}} \tfrac{2-\alpha}{(1-\alpha)\lambda_n}\, \mathrm{d}\alpha \\
&+ \int_{\frac{2}{3}}^{2} \tfrac{2\alpha}{(2\alpha-1)\lambda_n}\, \mathrm{d}\alpha + \int_2^{\hat{\alpha}} \left[\tfrac{2\alpha}{(2\alpha-1)\lambda_n} - \tfrac{2-\alpha}{(1-\alpha)\lambda_2}\right] \mathrm{d}\alpha \\
=& \tfrac{1}{\lambda_n} \ln\left[(1-\check{\alpha})\sqrt{2\hat{\alpha}-1}\right] + \tfrac{1}{\lambda_2} \ln\left[(\hat{\alpha}-1)\sqrt{1-2\check{\alpha}}\right] \\
&+ \tfrac{3\ln 3}{2\lambda_n} + \tfrac{2}{\lambda_2} + \left(\tfrac{1}{\lambda_n} - \tfrac{1}{\lambda_2}\right)(\hat{\alpha} - \check{\alpha}),
\end{aligned}
$$

where $\check{\alpha}$ and $\hat{\alpha}$ are given in Section IV-A. The index $\mathcal{S}_\mathcal{G}$ is completely determined by the non-zero minimum and maximum eigenvalues of the Laplacian matrices. Combining with Table I, it is easy to verify that as the network size increases, $\mathcal{S}_\Omega \to 0$ on $\mathbb{K}$ and $\mathbb{S}$, $\mathcal{S}_\Omega \to \frac{1+\ln 3}{2}$ on $\mathbb{P}$ and $\mathbb{C}^1$, $\mathcal{S}_\Omega \to \frac{8+8\ln 3}{25}$ on $\mathbb{C}^2$, and $\mathcal{S}_\Omega \to \frac{450+450\ln 3}{1942}$ on $\mathbb{C}^3$. Obviously, except for $\mathbb{K}$ and $\mathbb{S}$, there always exists a fixed minimum region for each graph family, in which the consensus network can be infinitely expanded by presetting the control parameters. This is consistent with the observation in Table II. It seems that the index $\mathcal{S}_\Omega$ is effective to evaluate the scalability of a certain type of graph family. Furthermore, with graph-theoretical endeavor, it is likely to build an estimation of $\mathcal{S}_\Omega$ without Laplacian spectrum, thereby making the index more tractable to design a general graph family that always guarantees a common consensus region irrespective of the communication network.

Meanwhile, larger $\mathcal{S}_\Omega$ implies that multi-agent network (8)-(9) is less fragile to maintain a desired robustness margin with respect to the variation of communication topology, which coincides with the opinion in [9].

F. Performance Comparison With Existing Memory-Based Consensus Protocols

Compared with the algorithm (4) with one-tap memory (i.e., $M = 1$) proposed in [36], our algorithm (8)-(9) has less parameters and thus reduces the operational complexity. It should be pointed that in this paper, we focus on the one-tap memory versions of other three representative memory-based schemes (1), (2), and (3), which all contains no more than two control parameters and whose augmented compact forms are given as follows:

1) The parameter-free scheme (PF-A) proposed in [31]

$$\bar{x}(t+1) = \begin{bmatrix} \mathbf{0} & I_n \\ \frac{1}{2}\mathcal{D}_n^{-1}\mathcal{A}_n & \frac{1}{2}\mathcal{D}_n^{-1}\mathcal{A}_n \end{bmatrix} \bar{x}(t); \tag{23}$$

2) The one-tap partial memory scheme without neighbor's memory (PM-L) proposed in [34]

$$\bar{x}(t+1) = \begin{bmatrix} \mathbf{0} & I_n \\ -\varrho_0 I_n & (1+\varrho_0)I_n - \vartheta L_n \end{bmatrix} \bar{x}(t); \tag{24}$$

3) The one-tap full memory scheme (FM-L) designed in [35]

$$\bar{x}(t+1) = \begin{bmatrix} \mathbf{0} & I_n \\ -\vartheta_1 L_n & I_n - \vartheta_0 L_n \end{bmatrix} \bar{x}(t). \tag{25}$$

Under PF-A algorithm, consensus is always achieved. Similar to the proof of Theorem 1, the consensus regions of PM-L and FM-L algorithms for a given connected undirected graph can be obtained as $\Omega_{\mathrm{PM-L}} = \{(\varrho_0, \vartheta) \mid -1 < \varrho_0 < 1, 0 < \vartheta < \frac{2+2\varrho_0}{\lambda_n}\}$ and $\Omega_{\mathrm{FM-L}} = \{(\vartheta_0, \vartheta_1) \mid -\vartheta_0 < \vartheta_1 < \frac{1}{\lambda_n}, \vartheta_1 > \vartheta_0 - \frac{2}{\lambda_n}\}$, respectively.

In order to perform a fair comparison, suppose that systems (23), (24), and (25) take the same disturbance input matrix and the same output matrix as (14). Denote the input-output transfer function matrices of the obtained closed-loop systems as $T_{\mathrm{PF-A}}(z)$, $T_{\mathrm{PM-L}}(z)$, and $T_{\mathrm{FM-L}}(z)$, respectively. We compare the optimal robustness of these closed-loop systems with the optimal robustness of (14). Firstly, we obtain the asymptotically stable and observable subsystems of these closed-loop systems. Then, we can invoke the *mincx solver* in Matlab Robust Control Toolbox and solve a series of LMIs to obtain $\|T_{\mathrm{PF-A}}(z)\|_\infty$, $\|T_{\mathrm{PM-L}}(z)\|_\infty$, and $\|T_{\mathrm{FM-L}}(z)\|_\infty$ for any given parameters in consensus regions, where the LMIs are derived by imposing the well-known bounded-real lemma [44] on the obtained subsystems and the solving accuracy of the solver is set to be 10^{-16}. For PM-L and FM-L schemes, by repeatedly numerating and testing parameter pairs in consensus regions with higher searching accuracy, we can get the numerical approximations of the optimal robustness $\|T_{\mathrm{PM-L}}(z)\|_\infty^*$ and $\|T_{\mathrm{FM-L}}(z)\|_\infty^*$. Table III summarized the numerical results on various graphs. Obviously, our memory-based consensus protocol performs as well as FM-L algorithm and outperforms other algorithms on most graphs.

TABLE III: Robustness comparison with existing memory-based algorithms.

Robustness	K_{10}	S_{10}	P_{10}	C_{10}^1	C_{10}^2	C_{10}^3
$\|T_{\mathrm{PF-A}}(z)\|_\infty$	1.0663	**2.1381**	17.2030	5.2361	2.2677	1.3692
$\|T_{\mathrm{PM-L}}(z)\|_\infty^*$	1	3.8829	12.4246	4.0252	1.8687	1.3394
$\|T_{\mathrm{FM-L}}(z)\|_\infty^*$	1	3.6553	11.6669	3.7871	1.7954	1.3129
$\|T_{\alpha,\beta}(z)\|_\infty^*$	**1**	3.6553	**11.6669**	**3.7871**	**1.7954**	**1.3129**

According to the qualitative analysis in Section IV-C, we can summarize that a protocol has better scalability if its consensus region completely encompasses that of another protocol. Unfortunately, this property does not hold for consensus regions Ω, $\Omega_{\mathrm{PM-L}}$, and $\Omega_{\mathrm{FM-L}}$. Therefore, we cannot give a general comparison result. Furthermore, the degree matrix $\mathcal{D}_n$ in PF-A algorithm, serving as a coupling parameter, is varying with communication graphs, hence violates the setting of fixed parameters for scalability and makes PF-A algorithm not comparable with other algorithms. However, we can compare the scalability of our algorithm with PM-L and FM-L algorithms on specific graph families from the perspective of minimum consensus regions as suggested in Section V-E. For a fixed connected undirected graph, the area of consensus regions associated with $\mathrm{PM-L}$ and $\mathrm{FM-L}$ algorithms can be computed as $\mathcal{S}_{\Omega_{\mathrm{PM-L}}} = \frac{4}{\lambda_n}$ and $\mathcal{S}_{\Omega_{\mathrm{FM-L}}} = \frac{4}{\lambda_n^2}$, respectively. Table IV summarizes the infimums of $\mathcal{S}_\Omega$, $\mathcal{S}_{\Omega_{\mathrm{PM-L}}}$, and $\mathcal{S}_{\Omega_{\mathrm{FM-L}}}$ over various graph families. Obviously, except for $\mathbb{K}$ and $\mathbb{S}$, the infimums of the consensus region associated with our memory-based protocol are larger on other graph families, which means that parameters can be preset more easily for the scalable consensus over the whole graph family.

TABLE IV: Infimums of the area of consensus regions.

	$\mathbb{K}$	$\mathbb{S}$	$\mathbb{P}$	$\mathbb{C}^1$	$\mathbb{C}^2$	$\mathbb{C}^3$
$\inf \mathcal{S}_{\Omega_{\mathrm{PM-L}}}$	0	0	1	1	0.6400	0.4634
$\inf \mathcal{S}_{\Omega_{\mathrm{FM-L}}}$	0	0	0.2500	0.2500	0.1024	0.0537
$\inf \mathcal{S}_\Omega$	0	0	**1.0493**	**1.0493**	**0.6716**	**0.4863**

VI. Conclusion

In this article, we designed a memory-based consensus protocol and investigated the role of memory information on robustness and scalability of first-order discrete-time multi-agent consensus networks. We first built the consensus region that is a necessary and sufficient consensus criterion. Then, the quantitative relation between the ℓ_2 gain, the memory coefficient, the control gain, and the non-zero minimum and maximum Laplacian eigenvalues was obtained. Compared with memoryless case ($\alpha = 1$), it was proved that the memory information is able to improve both robustness and scalability if and only if the memory coefficient is set as $0 < \alpha < 1$.

We also derived the optimal memory coefficient and control gain in terms of robustness. Moreover, we showed that better scalability essentially implies broader consensus convergence interval of the control gain. Building on this fact, we gave the optimal memory coefficient with respect to scalability. Interestingly, it was observed that robustness and scalability are positive correlated. The inherent relation between them deserves to be further studied in future works.

Appendix

A. Proof of Lemma 3

Proof: (Sufficiency.) Since $\lambda_1 = 0$, we get

$$\gamma_{1,1} = \frac{\alpha + |\alpha - 2|}{2},\ \gamma_{1,2} = \frac{\alpha - |\alpha - 2|}{2}. \tag{26}$$

One can verify that

$$\begin{cases} \gamma_{1,1} = 1, \gamma_{1,2} = \alpha - 1 < -1 & \text{if } \alpha < 0, \\ \gamma_{1,1} = 1, \gamma_{1,2} = -1 & \text{if } \alpha = 0, \\ \gamma_{1,1} = 1, -1 < \gamma_{1,2} = \alpha - 1 < 1 & \text{if } 0 < \alpha < 2, \\ \gamma_{1,1} = 1, \gamma_{1,2} = 1 & \text{if } \alpha = 2, \\ \gamma_{1,1} = \alpha - 1 > 1, \gamma_{1,2} = 1 & \text{if } \alpha > 2. \end{cases} \tag{27}$$

It is worth mentioning that the geometric multiplicity of eigenvalue 1 equals to one when $\alpha = 2$. Noticing that all eigenvalues of Φ_n except $\gamma_{1,1}$ and $\gamma_{1,2}$ are inside the unit circle in the complex plane, Φ_n can be decomposed as the following form

$$\Phi_n = SJS^{-1} = [\zeta_1, \zeta_2, \ldots, \zeta_{2n}] \begin{bmatrix} J_1 & \mathbf{0} \\ \mathbf{0} & \bar{J} \end{bmatrix} \begin{bmatrix} \xi_1^\top \\ \xi_2^\top \\ \vdots \\ \xi_{2n}^\top \end{bmatrix}, \tag{28}$$

where

$$J_1 = \begin{cases} \gamma_{1,1} & \text{if } 0 < \alpha < 2, \\ \begin{bmatrix} \gamma_{1,1} & 0 \\ 0 & \gamma_{1,2} \end{bmatrix} & \text{if } \alpha \le 0 \text{ or } \alpha > 2, \\ \begin{bmatrix} \gamma_{1,1} & 1 \\ 0 & \gamma_{1,2} \end{bmatrix} & \text{if } \alpha = 2, \end{cases} \tag{29}$$

is the Jordan block matrix with respect to the eigenvalues $\gamma_{1,1}$ and $\gamma_{1,2}$ (only $\gamma_{1,1}$ when $0 < \alpha < 2$), $\bar{J}$ is the upper triangular Jordan block matrix associated with remaining eigenvalues of Φ_n, ζ_i denotes the right or generalized right eigenvector of the corresponding eigenvalue, and ξ_i represents the generalized left or left eigenvector of the corresponding eigenvalue. In addition, it is obvious that $\lim_{t\to\infty} \bar{J}^t = \mathbf{0}$.

Based on the above observations, we complete the proof by enumeration.

Case (a): $0 < \alpha < 2$

In this case, we have $\gamma_{1,1} = 1$ whose right and the left eigenvectors can be respectively selected as $\zeta_1 = \mathbf{1}_{2n}$ and $\xi_1 = [\frac{1-\alpha}{(2-\alpha)n}\mathbf{1}_n^\top\ \frac{1}{(2-\alpha)n}\mathbf{1}_n^\top]^\top$, where $\zeta_1^\top \xi_1 = 1$. Then, according to (28), (29) and the fact $\lim_{t\to\infty} \bar{J}^t = \mathbf{0}$, one can obtain that

$$\begin{aligned} \lim_{t\to\infty} \bar{x}(t) &= \lim_{t\to\infty} \Phi_n^t \bar{x}(0) = \lim_{t\to\infty} SJ^tS^{-1}\bar{x}(0) \\ &= \zeta_1\xi_1^\top \begin{bmatrix} x(-1) \\ x(0) \end{bmatrix} \\ &= \frac{(1-\alpha)\sum_{i=1}^n x_i(-1) + \sum_{i=1}^n x_i(0)}{(2-\alpha)n} \cdot \mathbf{1}_{2n}. \end{aligned} \tag{30}$$

Case (b): $\alpha \le 0$ or $\alpha > 2$

Under this circumstance, we get $\gamma_{1,1} = 1$ and $\gamma_{1,2} = \alpha - 1$. Here, the corresponding eigenvectors can be set as $\zeta_1 = \mathbf{1}_{2n}$, $\xi_1 = [\frac{1-\alpha}{(2-\alpha)n}\mathbf{1}_n^\top\ \frac{1}{(2-\alpha)n}\mathbf{1}_n^\top]^\top$, $\zeta_2 = [\mathbf{1}_n^\top\ (\alpha-1)\mathbf{1}_n^\top]^\top$ and $\xi_2 = [-\frac{1}{(\alpha-2)n}\mathbf{1}_n^\top\ \frac{1}{(\alpha-2)n}\mathbf{1}_n^\top]^\top$, where $\zeta_1^\top\xi_1 = \zeta_2^\top\xi_2 = 1$.

Then, it follows from $\lim_{t\to\infty} \bar{J}^t = \mathbf{0}$, (28) and (29) that

$$\begin{aligned} \lim_{t\to\infty} \bar{x}(t) &= \lim_{t\to\infty} \Phi_n^t \bar{x}(0) = \lim_{t\to\infty} SJ^tS^{-1}\bar{x}(0) \\ &= [\zeta_1, \zeta_2] \begin{bmatrix} 1 & 0 \\ 0 & (\alpha-1)^t \end{bmatrix} \begin{bmatrix} \xi_1^\top \\ \xi_2^\top \end{bmatrix} \begin{bmatrix} x(-1) \\ x(0) \end{bmatrix} \\ &= [\zeta_1\xi_1^\top + (\alpha-1)^t\zeta_2\xi_2^\top] \begin{bmatrix} x(-1) \\ x(0) \end{bmatrix} \\ &= \begin{bmatrix} \mu_1 \cdot \mathbf{1}_n \\ \mu_2 \cdot \mathbf{1}_n \end{bmatrix}, \end{aligned} \tag{31}$$

where $\mu_1 = \frac{[1-\alpha+(\alpha-1)^t]\sum_{i=1}^n x_i(-1) + [1-(\alpha-1)^t]\sum_{i=1}^n x_i(0)}{(2-\alpha)n}$ and $\mu_2 = \frac{[1-\alpha+(\alpha-1)^{t+1}]\sum_{i=1}^n x_i(-1) + [1-(\alpha-1)^{t+1}]\sum_{i=1}^n x_i(0)}{(2-\alpha)n}$.

Case (c): $\alpha = 2$

In this case, $\gamma_{1,1} = \gamma_{1,2} = 1$ whose geometric multiplicity equals to one. The right eigenvector and the generalized right eigenvector of eigenvalue 1 can be chosen as $\zeta_1 = -\mathbf{1}_{2n}$ and $\zeta_2 = [\mathbf{1}_n^\top\ \mathbf{0}_n^\top]^\top$, the corresponding generalized left eigenvector and the left eigenvector are set as $\xi_1 = -\frac{1}{n}[\mathbf{0}_n^\top\ \mathbf{1}_n^\top]^\top$ and $\xi_2 = \frac{1}{n}[\mathbf{1}_n^\top\ -\mathbf{1}_n^\top]^\top$, where $\zeta_1^\top\xi_1 = \zeta_2^\top\xi_2 = 1$.

Therefore, it follows from $\lim_{t\to\infty} \bar{J}^t = \mathbf{0}$, (28) and (29) that

$$\begin{aligned} \lim_{t\to\infty} \bar{x}(t) &= \lim_{t\to\infty} \Phi_n^t \bar{x}(0) = \lim_{t\to\infty} SJ^tS^{-1}\bar{x}(0) \\ &= [\zeta_1, \zeta_2] \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1^\top \\ \xi_2^\top \end{bmatrix} \begin{bmatrix} x(-1) \\ x(0) \end{bmatrix} \\ &= \begin{bmatrix} \frac{(1-t)\sum_{i=1}^n x_i(-1) + t\sum_{i=1}^n x_i(0)}{n} \cdot \mathbf{1}_n \\ \frac{-t\sum_{i=1}^n x_i(-1) + (1+t)\sum_{i=1}^n x_i(0)}{n} \cdot \mathbf{1}_n \end{bmatrix}. \end{aligned} \tag{32}$$

To sum up the above cases, it is always observed that for any initial conditions $x_i(-1)$ and $x_i(0)$, where $i = 1, 2, \ldots, n$, we have $\lim_{t\to\infty} |x_i(t) - x_j(t)| = 0$, $\forall i, j \in \mathcal{I}_n$. Namely, the consensus is achieved.

(Necessity.) Here, we prove the necessity by contradiction. Let $\delta_i(t) = x_i(t) - x_1(t)$, $\delta(t) = [\delta_2(t), \ldots, \delta_n(t)]^\top$ and $\bar{\delta}(t) = [\delta^\top(t-1)\ \delta^\top(t)]^\top$, $i = 2, \ldots, n$, then multi-agent network (8)-(9) achieves consensus if and only if the following system is stable

$$\bar{\delta}(t+1) = \underbrace{\begin{bmatrix} \mathbf{0} & I_{n-1} \\ (1-\alpha)(I_n - \beta\hat{L}) & \alpha(I_n - \beta\hat{L}) \end{bmatrix}}_{\hat{\Phi}} \bar{\delta}(t), \tag{33}$$

where

$$\hat{L} = \begin{bmatrix} l_{22} - l_{12} & \cdots & l_{2n} - l_{1n} \\ \vdots & \ddots & \vdots \\ l_{n2} - l_{12} & \cdots & l_{nn} - l_{1n} \end{bmatrix}$$

has the same eigenvalues of L_n excluding a zero eigenvalue (see the proof of Lemma 1 in [45]). It follows that $\hat{\Phi}$ has all eigenvalues of Φ_n except $\gamma_{1,1}$ and $\gamma_{1,2}$. Therefore, if at least one eigenvalue of Φ_n excluding $\gamma_{1,1}$ and $\gamma_{1,2}$ are not inside the unit circle, $\hat{\Phi}$ also has at least one eigenvalue not in the unit circle, which means that system (33) is unstable. Namely, multi-agent network (8)-(9) cannot achieves consensus, which leads to a contradiction. ■

B. Proof of Theorem 1

Proof: Applying the bilinear transformation $\gamma = \frac{\sigma+1}{\sigma-1}$ to $p(\gamma, \lambda_i)$, $i \in \{2, \dots, n\}$ gives rise to a new polynomial with real coefficients

$$\begin{aligned} &f(\sigma, \lambda_i) = (\sigma-1)^2 p(\frac{\sigma+1}{\sigma-1}, \lambda_i) \\ &= \beta\lambda_i\sigma^2 + [4 - 2\alpha + (2\alpha-2)\beta\lambda_i]\sigma + (1-2\alpha)\beta\lambda_i + 2\alpha. \end{aligned}$$

Then the Schur stability of $p(\gamma, \lambda_i)$ is equivalent to the Hurwitz stability of $f(\sigma, \lambda_i)$. Apparently, $f(\sigma, \lambda_i)$ is Hurwitz stable if and only if all its coefficients are positive, which implies that the following inequalities simultaneously hold

$$\begin{cases} \beta\lambda_i > 0, \\ 4 - 2\alpha + (2\alpha-2)\beta\lambda_i > 0, \\ (1-2\alpha)\beta\lambda_i + 2\alpha > 0. \end{cases} \tag{34}$$

Furthermore, as shown in (27), at least one of the eigenvalues $\gamma_{1,1}$ and $\gamma_{1,2}$ is equal to 1. Combining with Lemma 1 and the above observations, all eigenvalues of Φ_n excluding $\gamma_{1,1}$ and $\gamma_{1,2}$ are in the unit circle if and only if L_n has a simple zero eigenvalue (i.e., undirected graph $\mathcal{G}_n$ is connected) and (34) holds for all non-zero eigenvalues of L_n. Therefore, according to Lemma 3, consensus is achieved if and only if $\mathcal{G}_n$ is connected and (34) holds for any $i \in \{2, \dots, n\}$. Next, we only need to solve (34), which can be enumerated as the following cases.

Case (a): $0 < \alpha \le \frac{2}{3}$

Note that $\frac{2\alpha}{2\alpha-1} < 0 < \frac{2-\alpha}{1-\alpha}$ when $0 < \alpha \le \frac{1}{2}$ and $0 < \frac{2-\alpha}{1-\alpha} \le \frac{2\alpha}{2\alpha-1}$ when $\frac{1}{2} < \alpha \le \frac{2}{3}$. For both cases, we can eventually get $0 < \beta < \frac{2-\alpha}{(1-\alpha)\lambda_i}$ by solving (34). It follows from $0 = \lambda_1 < \lambda_2 \le \cdots \le \lambda_n$ that (34) holds for any $i \in \{2, \dots, n\}$ if and only if $0 < \beta < \frac{2-\alpha}{(1-\alpha)\lambda_n}$.

Case (b): $\frac{2}{3} \le \alpha < 2$

Under this circumstance, $0 < \frac{2\alpha}{2\alpha-1} \le \frac{2-\alpha}{1-\alpha}$ when $\frac{2}{3} \le \alpha \le 1$ and $\frac{2-\alpha}{1-\alpha} < 0 < \frac{2\alpha}{2\alpha-1}$ when $1 < \alpha < 2$. For either case, solving (34) yields $0 < \beta < \frac{2\alpha}{(2\alpha-1)\lambda_i}$. Therefore, (34) is satisfied for any $i \in \{2, \dots, n\}$ if and only if $0 < \beta < \frac{2\alpha}{(2\alpha-1)\lambda_n}$.

Case (c): $\alpha \le 0$

In this scenario, we have $\frac{2-\alpha}{1-\alpha} > \frac{2\alpha}{2\alpha-1} \ge 0$. Solving (34) gives $\frac{2\alpha}{(2\alpha-1)\lambda_i} < \beta < \frac{2-\alpha}{(1-\alpha)\lambda_i}$. It is inferred from $0 = \lambda_1 < \lambda_2 \le \cdots \le \lambda_n$ that (34) holds for any $i \in \{2, \dots, n\}$ if and only if $\frac{2\alpha}{(2\alpha-1)\lambda_2} < \beta < \frac{2-\alpha}{(1-\alpha)\lambda_n}$.

Case (d): $\alpha \ge 2$

It is obvious that $\frac{2\alpha}{2\alpha-1} > \frac{2-\alpha}{1-\alpha} \ge 0$. Likewise, we can easily obtain that (34) holds for any $i \in \{2, \dots, n\}$ if and only if $\frac{2-\alpha}{(1-\alpha)\lambda_2} < \beta < \frac{2\alpha}{(2\alpha-1)\lambda_n}$.

To sum up, multi-agent network (8)-(9) achieves consensus if and only if the undirected graph $\mathcal{G}_n$ is connected and $(\alpha, \beta) \in \Omega$. ■

C. Proof of Theorem 2

Proof: The proof can be decomposed into three steps.

Step I: Proving the existence of $\|T_{\alpha,\beta}(z)\|_\infty$.

Note that the H_∞ norm only exists for asymptotically stable systems as described in Definition 1, while system (14) contains at least one marginally stable mode and sometimes even includes an unstable mode, which is discussed in (27). To proceed further, we need first to prove the existence of $\|T_{\alpha,\beta}(z)\|_\infty$.

Because the undirected graph $\mathcal{G}_n$ is connected, according to Lemma 2, there is an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ such that (6) and (7) hold, where $\tilde{L} \in \mathbb{R}^{(n-1)\times(n-1)}$ is positive definite and contains all non-zero eigenvalues of L_n. Therefore, according to the spectral theorem [40], there is an orthogonal matrix $Q_1 \in \mathbb{R}^{(n-1)\times(n-1)}$ such that $Q_1^\top \tilde{L} Q_1 = \Lambda = \mathrm{diag}\{\lambda_2, \dots, \lambda_n\}$. Let $Q_2 = \begin{bmatrix} 1 & \mathbf{0}_{n-1}^\top \\ \mathbf{0}_{n-1} & Q_1 \end{bmatrix}$ and $Q = UQ_2$. Then, applying the following orthogonal transformation

$$\hat{x}(t) = \begin{bmatrix} Q^\top & \mathbf{0} \\ \mathbf{0} & Q^\top \end{bmatrix} \bar{x}(t), \hat{\omega}(t) = Q^\top \omega(t), \hat{\epsilon}(t) = Q^\top \epsilon(t), \tag{35}$$

for the system (14) yields

$$\begin{cases} \hat{x}(t+1) = \begin{bmatrix} \mathbf{0} & \bar{\Theta}_n \\ (1-\alpha)\Xi & \alpha\Xi \end{bmatrix} \hat{x}(t) + \begin{bmatrix} \mathbf{0} \\ I_n \end{bmatrix} \hat{\omega}(t), \\ \hat{\epsilon}(t) = \begin{bmatrix} \mathbf{0} & \bar{\Theta}_n \end{bmatrix} \hat{x}(t), \end{cases} \tag{36}$$

where $\Xi = \begin{bmatrix} 1 & \mathbf{0}_{n-1}^\top \\ \mathbf{0}_{n-1} & \Xi_1 \end{bmatrix}$ and $\Xi_1 = I_{n-1} - \beta\Lambda$. Obviously, (36) is composed of an observable subsystem of order $2n-2$ and an unobservable subsystem of order 2. The observable subsystem of (36) is

$$\begin{cases} \tilde{x}(t+1) = \begin{bmatrix} \mathbf{0} & I_{n-1} \\ (1-\alpha)\Xi_1 & \alpha\Xi_1 \end{bmatrix} \tilde{x}(t) + \begin{bmatrix} \mathbf{0} \\ I_{n-1} \end{bmatrix} \tilde{\omega}(t), \\ \tilde{\epsilon}(t) = \begin{bmatrix} \mathbf{0} & I_{n-1} \end{bmatrix} \tilde{x}(t). \end{cases} \tag{37}$$

It is easy to verify that the system matrix of (37) possesses all eigenvalues of Φ_n except $\gamma_{1,1}$ and $\gamma_{1,2}$. Combining Lemma 3 and Theorem 1, one can deduce that subsystem (37) is asymptotically stable if and only if $\mathcal{G}_n$ is connected and $(\alpha, \beta) \in \Omega$. Therefore, all observable modes of (14) are asymptotically stable, which means that $\|T_{\alpha,\beta}(z)\|_\infty$ exists and is completely determined by (37). Let $T^{(1)}_{\alpha,\beta}(z)$ and $T^{(2)}_{\alpha,\beta}(z)$ be the transfer matrices of systems (36) and (37), respectively.

Step II: Verifying that $\|T_{\alpha,\beta}(z)\|_\infty = \|T^{(2)}_{\alpha,\beta}(z)\|_\infty$.

We can obtain

$$T^{(1)}_{\alpha,\beta}(z) = Q^\top T_{\alpha,\beta}(z) Q = \begin{bmatrix} 0 & \mathbf{0}_{n-1}^\top \\ \mathbf{0}_{n-1} & T^{(2)}_{\alpha,\beta}(z) \end{bmatrix}.$$

It follows from Definition 1 and

$$\begin{aligned} \lambda_{\max}\left[(T^{(1)}_{\alpha,\beta}(e^{\mathbf{j}\upsilon}))^H T^{(1)}_{\alpha,\beta}(e^{\mathbf{j}\upsilon})\right] &= \lambda_{\max}\left[(T_{\alpha,\beta}(e^{\mathbf{j}\upsilon}))^H T_{\alpha,\beta}(e^{\mathbf{j}\upsilon})\right] \\ &= \lambda_{\max}\left[(T^{(2)}_{\alpha,\beta}(e^{\mathbf{j}\upsilon}))^H T^{(2)}_{\alpha,\beta}(e^{\mathbf{j}\upsilon})\right] \end{aligned}$$

that $\|T_{\alpha,\beta}(z)\|_\infty = \|T^{(1)}_{\alpha,\beta}(z)\|_\infty = \|T^{(2)}_{\alpha,\beta}(z)\|_\infty$.

Step III: Calculating $\|T^{(2)}_{\alpha,\beta}(z)\|_\infty$.

Combining with the fact $e^{\mathbf{j}\upsilon} = \cos\upsilon + \mathbf{j}\sin\upsilon$, we have $[T^{(2)}_{\alpha,\beta}(e^{\mathbf{j}\upsilon})]^H T^{(2)}_{\alpha,\beta}(e^{\mathbf{j}\upsilon}) = \mathrm{diag}\{\phi_2(\upsilon), \dots, \phi_n(\upsilon)\}$, where

$$\phi_i(\upsilon) = \frac{1}{[\cos 2\upsilon + \alpha\kappa_i \cos\upsilon + (1-\alpha)\kappa_i]^2 + [\sin 2\upsilon + \alpha\kappa_i \sin\upsilon]^2},$$

and $\kappa_i = \beta\lambda_i - 1$, $i = 2, \dots, n$. In particular, $\phi_i(\upsilon) \equiv 1$ if $\beta = \frac{1}{\lambda_i}$. According to $\|T_{\alpha,\beta}(z)\|_\infty = \|T^{(2)}_{\alpha,\beta}(z)\|_\infty$ and Definition 1, $\|T_{\alpha,\beta}(z)\|_\infty$ can be derived by

$$\begin{aligned}\|T_{\alpha,\beta}(z)\|_\infty &= \sup_{\upsilon\in[0,2\pi]} \sqrt{\lambda_{\max}\left[\left(T^{(2)}_{\alpha,\beta}(e^{\mathbf{j}\upsilon})\right)^H T^{(2)}_{\alpha,\beta}(e^{\mathbf{j}\upsilon})\right]} \\ &= \max_{i=2,\dots,n} \sqrt{\sup_{\upsilon\in[0,2\pi]} \phi_i(\upsilon)}.\end{aligned} \tag{38}$$

Solving $\frac{\mathrm{d}\phi_i(\upsilon)}{\mathrm{d}\upsilon} = 0$ yields

$$2(1-\alpha)\kappa_i \sin 2\upsilon + \alpha\kappa_i[(1-\alpha)\kappa_i + 1]\sin\upsilon = 0 \tag{39}$$

that always have three roots $\upsilon_{i,1} = 0$, $\upsilon_{i,2} = 2\pi$, and $\upsilon_{i,3} = \pi$ within $[0, 2\pi]$. Moreover, by solving (39), other possible roots $\upsilon \in [0, 2\pi]$ can be given by $\upsilon_{i,4}, \upsilon_{i,5} = \arccos \mu_i$ if and only if

$$-1 \le \mu_i \le 1, \tag{40}$$

where $\mu_i = \frac{\alpha[(1-\alpha)\beta\lambda_i + \alpha]}{4(\alpha-1)}$. Further computation provides

$$\begin{aligned}&\phi_i(\upsilon_{i,1}) = \phi_i(\upsilon_{i,2}) = g_1(\lambda_i), \\ &\phi_i(\upsilon_{i,3}) = g_2(\lambda_i), \\ &\phi_i(\upsilon_{i,4}) = \phi_i(\upsilon_{i,5}) = g_3(\lambda_i).\end{aligned} \tag{41}$$

However, depending on the location of (α, β) in the consensus region Ω, (40) may not hold, that is, roots $\upsilon_{i,4}$ and $\upsilon_{i,5}$ may not exist. Therefore, we have to complete the proof by enumeration. For notational convenience, let $\Upsilon_{i,q} = \{\upsilon_{i,1}, \dots, \upsilon_{i,q}\}$ be the set of all feasible roots of (39), where $i \in \mathcal{T}_n = \{2, \dots, n\}$ and $q \in \{3, 5\}$. In the following, we only show two representative scenarios.

Case (a): $0 \le \alpha \le \frac{2}{3}$, $0 < \beta < \hat{\eta}(\lambda_n)$

Solving (40) yields that the roots $\upsilon_{i,4}$ and $\upsilon_{i,5}$ exist if and only if $\beta \le \check{\theta}(\lambda_i)$. Since $\check{\theta}(\lambda_i) \ge \hat{\eta}(\lambda_i) \ge \hat{\eta}(\lambda_n)$ and $0 < \beta < \hat{\eta}(\lambda_n)$, there are always five roots of (39) for all $i \in \mathcal{T}_n$. Moreover, we obtain

$$\begin{cases} g_1(\lambda_i) \ge g_2(\lambda_i) \ge g_3(\lambda_i) & \text{if } 0 < \lambda_i \le \frac{1}{\beta}, \\ g_3(\lambda_i) > g_2(\lambda_i) > g_1(\lambda_i) & \text{if } \frac{1}{\beta} < \lambda_i < \hat{\eta}(\beta), \end{cases} \tag{42}$$

where the relation between $g_1(\lambda_i)$ and $g_2(\lambda_i)$ is obvious, and the relation between $g_3(\lambda_i)$ and $g_2(\lambda_i)$ can be easily derived by analyzing the function $g_3(\lambda_i) - g_2(\lambda_i)$ and its first derivation.

Recall that $\Gamma = \{\lambda_2, \dots, \lambda_n\}$. Here, we define two sets as $\mathcal{R}_1 = \{r \in \Gamma \mid 0 < r \le \frac{1}{\beta}\}$ and $\mathcal{R}_2 = \{r \in \Gamma \mid \frac{1}{\beta} < r < \hat{\eta}(\beta)\}$. Obviously, we have $\mathcal{R}_1 \cup \mathcal{R}_2 = \Gamma$ and $\mathcal{R}_1 \cap \mathcal{R}_2 = \emptyset$. Therefore, it is deduced from Lemma 4, (41) and (42) that if $\mathcal{R}_1 = \Gamma$ and $\mathcal{R}_2 = \emptyset$, we have $0 < \lambda_i \le \frac{1}{\beta}$, $\forall i \in \mathcal{T}_n$, and then $\max_{i\in\mathcal{T}_n} \sup_{\upsilon\in\Upsilon_{i,5}} \phi_i(\upsilon) = g_1(\lambda_2) \ge g_1(\lambda_n) \ge g_3(\lambda_n)$; and if $\mathcal{R}_1 = \emptyset$ and $\mathcal{R}_2 = \Gamma$, we get $\frac{1}{\beta} < \lambda_i < \hat{\eta}(\beta)$, $\forall i \in \mathcal{T}_n$, and then $\max_{i\in\mathcal{T}_n} \sup_{\upsilon\in\Upsilon_{i,5}} \phi_i(\upsilon) = g_3(\lambda_n) > g_3(\lambda_2) > g_1(\lambda_2)$. In addition, if $\mathcal{R}_1 \ne \emptyset$ and $\mathcal{R}_2 \ne \emptyset$, there must exists an eigenvalue λ_l $(2 \le l \le n-1)$ such that $\mathcal{R}_1 = \{\lambda_2, \dots, \lambda_l\}$ and $\mathcal{R}_2 = \{\lambda_{l+1}, \dots, \lambda_n\}$, and then $\max_{i\in\mathcal{T}_n} \sup_{\upsilon\in\Upsilon_{i,5}} \phi_i(\upsilon) = \max\{\max_{i\in\mathcal{R}_1} g_1(\lambda_i), \max_{i\in\mathcal{R}_2} g_3(\lambda_i)\} = \max\{g_1(\lambda_2), g_3(\lambda_n)\}$.

To summarize the above cases, (38) becomes

$$\begin{aligned}\|T_{\alpha,\beta}(z)\|_\infty &= \max_{i\in\mathcal{T}_n} \sqrt{\sup_{\upsilon\in\Upsilon_{i,5}} \phi_i(\upsilon)} \\ &= \max\{\sqrt{g_1(\lambda_2)}, \sqrt{g_3(\lambda_n)}\}.\end{aligned}$$

Case (b): $\frac{2}{3} \le \alpha \le 1$, $0 < \beta < \check{\eta}(\lambda_n)$

Note that depending on $\check{\theta}(\lambda_i)$, two cases is further classified: $\frac{1}{\lambda_i} < \check{\theta}(\lambda_i) < \check{\eta}(\lambda_i)$ when $\frac{2}{3} < \alpha < \frac{4}{5}$; and $\check{\theta}(\lambda_i) \le \frac{1}{\lambda_i} < \check{\eta}(\lambda_i)$ when $\frac{4}{5} \le \alpha \le 1$. We mainly prove the former case, which is the most complicated. Solving (40) provides that the roots $\upsilon_{i,4}$ and $\upsilon_{i,5}$ exist if and only if $\beta \le \check{\theta}(\lambda_i)$. It is easy to verify that

$$\begin{cases} \sup_{\upsilon\in\Upsilon_{i,5}} \phi_i(\upsilon) = g_1(\lambda_i) & \text{if } 0 < \beta \le \frac{1}{\lambda_i}, \\ \sup_{\upsilon\in\Upsilon_{i,5}} \phi_i(\upsilon) = g_3(\lambda_i) & \text{if } \frac{1}{\lambda_i} < \beta \le \check{\theta}(\lambda_i), \\ \sup_{\upsilon\in\Upsilon_{i,3}} \phi_i(\upsilon) = g_2(\lambda_i) & \text{if } \check{\theta}(\lambda_i) < \beta < \check{\eta}(\lambda_i). \end{cases} \tag{43}$$

Thus, if $\beta \le \check{\theta}(\lambda_n)$, it follows from (41) and (43) that $\sup_{\upsilon\in[0,2\pi]} \phi_i(\upsilon) = \sup_{\upsilon\in\Upsilon_{i,5}} \phi_i(\upsilon) = \max\{g_1(\lambda_i), g_3(\lambda_i)\}$, $\forall i \in \mathcal{T}_n$. Furthermore, $\beta \le \check{\theta}(\lambda_n) \le \cdots \le \check{\theta}(\lambda_2)$ implies $0 < \lambda_2 \le \cdots \le \lambda_n \le \check{\theta}(\beta) \le \check{\eta}(\beta)$. As a result, according to Lemma 4, we can obtain

$$\begin{aligned}\max_{i\in\mathcal{T}_n} \sup_{\upsilon\in[0,2\pi]} \phi_i(\upsilon) &= \max_{i\in\mathcal{T}_n} \sup_{\upsilon\in\Upsilon_{i,5}} \phi_i(\upsilon) \\ &= \max_{i\in\mathcal{T}_n} \max\{g_1(\lambda_i), g_3(\lambda_i)\} \\ &= \max\{\max_{i\in\mathcal{T}_n} g_1(\lambda_i), \max_{i\in\mathcal{T}_n} g_3(\lambda_i)\} \\ &= \max\{g_1(\lambda_2), g_3(\lambda_n)\}.\end{aligned} \tag{44}$$

On the other hand, consider $\check{\theta}(\lambda_n) \le \beta < \check{\eta}(\lambda_n)$. Then, there are no roots $\upsilon_{n,4}$ and $\upsilon_{n,5}$ for ϕ_n, hence $\sup_{\upsilon\in[0,2\pi]} \phi_n(\upsilon) = \sup_{\upsilon\in\Upsilon_{n,3}} \phi_n(\upsilon) = g_2(\lambda_n)$. For each ϕ_i, $i \in \mathcal{T}_{n-1}$, according to Lemma 4, we have $g_2(\lambda_n) > g_2(\lambda_i)$, $\forall\beta \in [\check{\theta}(\lambda_n), \check{\eta}(\lambda_n))$. Next, we can verify that $g_2(\lambda_n) = \frac{\alpha(1-\alpha)}{(3\alpha-2)^2}$ when $\beta = \check{\theta}(\lambda_n)$ and $g_3(\lambda_i) = \frac{\alpha(1-\alpha)}{(3\alpha-2)^2}$ when $\beta = \check{\theta}(\lambda_i)$, where $\check{\theta}(\lambda_n) \le \check{\theta}(\lambda_i)$. Furthermore, by inspecting the functions in Lemma 4, β and t are commutative, which means that $g_2(\lambda_n)$ and $g_3(\lambda_i)$ are increasing with β on $[\check{\theta}(\lambda_n), \check{\eta}(\lambda_n))$ and $[\check{\theta}(\lambda_n), \check{\theta}(\lambda_i))$, respectively. Therefore, we can deduce that $g_2(\lambda_n) \ge g_3(\lambda_i)$, $\forall\beta \in [\check{\theta}(\lambda_n), \min\{\check{\theta}(\lambda_i), \check{\eta}(\lambda_n)\})$. Based on above facts and (43), we conclude that

$$\begin{aligned}&\max_{i\in\mathcal{T}_n} \sup_{\upsilon\in[0,2\pi]} \phi_i(\upsilon) \\ =& \max_{i\in\mathcal{T}_{n-1}} \{\sup_{\upsilon\in[0,2\pi]} \phi_i(\upsilon), \sup_{\upsilon\in\Upsilon_{n,3}} \phi_n(\upsilon)\} \\ =& \max_{i\in\mathcal{T}_{n-1}} \{\max\{g_1(\lambda_i), g_3(\lambda_i)\}, g_2(\lambda_n)\} \\ =& \max_{i\in\mathcal{T}_{n-1}} \{g_1(\lambda_i), g_2(\lambda_n)\} \\ =& \max\{g_1(\lambda_2), g_2(\lambda_n)\}.\end{aligned} \tag{45}$$

In brief, combining (44) and (45), (38) becomes

$$\|T_{\alpha,\beta}(z)\|_\infty = \begin{cases} \max\{\sqrt{g_1(\lambda_2)}, \sqrt{g_3(\lambda_n)}\} & \text{if } 0 < \beta \leq \check{\theta}(\lambda_n), \\ \max\{\sqrt{g_1(\lambda_2)}, \sqrt{g_2(\lambda_n)}\} & \text{if } \check{\theta}(\lambda_n) \leq \beta < \check{\eta}(\lambda_n), \end{cases}$$

where $g_2(\lambda_n) = g_3(\lambda_n)$ when $\beta = \check{\theta}(\lambda_n)$. The trends of $\|T_{\alpha,\beta}(z)\|_\infty$ with respect to α and β are depicted in Fig. 2b. As we can see, $g_1(\lambda_2)$ is constant with respect to α, $g_3(\lambda_n)$ and $g_2(\lambda_n)$ are both increasing with α, and $\check{\theta}(\lambda_n)$ decreases with α. Obviously, as α grows, $\|T_{\alpha,\beta}(z)\|_\infty$ gradually exhibits different expressions. By solving $\sqrt{g_1(\lambda_2)} = \sqrt{g_2(\lambda_n)}$ with $\beta = \check{\theta}(\lambda_n)$, which means that $g_1(\lambda_2)$, $g_3(\lambda_n)$, and $g_2(\lambda_n)$ exactly intersect at $\beta = \check{\theta}(\lambda_n)$, we can get $\alpha = \frac{6\lambda_n - 2\lambda_2 + 2\sqrt{2\lambda_2(\lambda_2+\lambda_n)}}{9\lambda_n+\lambda_2} \triangleq \alpha_1$. Therefore, (38) can be further written as

$$\|T_{\alpha,\beta}(z)\|_\infty = \begin{cases} \mathcal{Q}_1(\alpha,\beta) & \text{if } \frac{2}{3} < \alpha \leq \alpha_1, \\ \max\{\sqrt{g_1(\lambda_2)}, \sqrt{g_2(\lambda_n)}\} & \text{if } \alpha_1 < \alpha < \frac{4}{5}. \end{cases}$$

Likewise, if $\frac{4}{5} \leq \alpha \leq 1$, as detailed in Fig. 2b, we can easily conclude that $\|T_{\alpha,\beta}(z)\|_\infty = \max\{\sqrt{g_1(\lambda_2)}, \sqrt{g_2(\lambda_n)}\}$.

Furthermore, the proof for cases $\alpha < 0$ and $\alpha > 1$ can be completed in a similar vein. For saving space, we omit the proof here. ■

References

[1] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, 2003.

[2] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1465–1476, 2004.

[3] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[4] W. Yu, G. Chen, and M. Cao, "Some necessary and sufficient conditions for second-order consensus in multi-agent dynamical systems," *Automatica*, vol. 46, no. 6, pp. 1089–1095, 2010.

[5] H. Rezaee and F. Abdollahi, "Average consensus over high-order multiagent systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 3047–3052, 2015.

[6] Y. Zheng, Y. Zhu, and L. Wang, "Consensus of heterogeneous multi-agent systems," *IET Control Theory & Applications*, vol. 5, no. 16, pp. 1881–1888, 2011.

[7] Y. Zheng, Q. Zhao, J. Ma, and L. Wang, "Second-order consensus of hybrid multi-agent systems," *Systems & Control Letters*, vol. 125, pp. 51–58, 2019.

[8] M. Cao, F. Xiao, and L. Wang, "Event-based second-order consensus control for multi-agent systems via synchronous periodic event detection," *IEEE Transactions on Automatic Control*, vol. 60, no. 9, pp. 2452–2457, 2015.

[9] Z. Li, Z. Duan, and G. Chen, "On H_∞ and H_2 performance regions of multi-agent systems," *Automatica*, vol. 47, no. 4, pp. 797–803, 2011.

[10] Y. Liu and Y. Jia, "H_∞ consensus control for multi-agent systems with linear coupling dynamics and communication delays," *International Journal of Systems Science*, vol. 43, no. 1, pp. 50–62, 2012.

[11] S. Chen, Z. Zhang, and Y. Zheng, "H_∞ scaled consensus for mass with mixed time delays and disturbances via observer-based output feedback," *IEEE Transactions on Cybernetics*, vol. 52, no. 2, pp. 1321–1334, 2020.

[12] H. Zhang, H. Yan, and J. Qiu, "Finite-horizon H_∞ consensus control of discrete time-varying multi-agent systems: A dynamic event-based learning approach," *IEEE Transactions on Automatic Control*, vol. 68, no. 10, pp. 6270–6276, 2023.

[13] Z. Li, J. Jiao, and X. Chen, "Distributed optimal control with recovered robustness for uncertain network systems: A complementary design approach," *IEEE Transactions on Automatic Control*, vol. 69, no. 4, pp. 2484–2491, 2023.

[14] M. Siami and N. Motee, "Systemic measures for performance and robustness of large-scale interconnected dynamical networks," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, 2014, pp. 5119–5124.

[15] M. Pirani, E. M. Shahrivar, B. Fidan, and S. Sundaram, "Robustness of leader–follower networked dynamical systems," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1752–1763, 2018.

[16] M. Pirani, H. Sandberg, and K. H. Johansson, "A graph-theoretic approach to the H_∞ performance of leader–follower consensus on directed networks," *IEEE Control Systems Letters*, vol. 3, no. 4, pp. 954–959, 2019.

[17] J. Wang, J. Liu, Y. Zheng, and J. Xi, "Analysis of H_∞ performance for multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 69, no. 8, pp. 5125–5140, 2024.

[18] Y. Ding, H. Peng, T. Qi, and J. Chen, "Consensus performance of first-order agents," *IEEE Transactions on Automatic Control*, vol. 69, no. 8, pp. 5446–5453, 2024.

[19] J. Wang, L. Zhou, D. Zhang, J. Liu, F. Xiao, and Y. Zheng, "Protocol selection for second-order consensus against disturbance," *Automatica*, vol. 161, p. 111497, 2024.

[20] M. Franceschelli, A. Gasparri, A. Giua, and C. Seatzu, "Decentralized estimation of laplacian eigenvalues in multi-agent systems," *Automatica*, vol. 49, no. 4, pp. 1031–1036, 2013.

[21] P. Di Lorenzo and S. Barbarossa, "Distributed estimation and control of algebraic connectivity over random graphs," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5615–5628, 2014.

[22] E. Tegling, B. Bamieh, and H. Sandberg, "Scale fragilities in localized consensus dynamics," *Automatica*, vol. 153, p. 111046, 2023.

[23] S. Studli, Y. Yan, M. M. Seron, and R. H. Middleton, "'Plug-and-play' style connection methods for graphs with an application in expanding multi-agent consensus networks," *IEEE Transactions on Automatic Control*, vol. 69, no. 5, pp. 2883–2898, 2024.

[24] W. Hou, M. Fu, H. Zhang, and Z. Wu, "Consensus conditions for general second-order multi-agent systems with communication delay," *Automatica*, vol. 75, pp. 293–298, 2017.

[25] Y. Chen and J. Lü, "Delay-induced discrete-time consensus," *Automatica*, vol. 85, pp. 356–361, 2017.

[26] X. Li, J.-C. Liu, and X.-G. Li, "Consensus of first-order discrete-time multi-agent systems with time delays," *Journal of the Franklin Institute*, vol. 356, no. 10, pp. 5315–5331, 2019.

[27] F. Xiao and L. Wang, "Consensus protocols for discrete-time multi-agent systems with time-varying delays," *Automatica*, vol. 44, no. 10, pp. 2577–2582, 2008.

[28] F. Xiao, L. Wang, and J. Chen, "Partial state consensus for networks of second-order dynamic agents," *Systems & Control Letters*, vol. 59, no. 12, pp. 775–781, 2010.

[29] D. Irofti, "An anticipatory protocol to reach fast consensus in multi-agent systems," *Automatica*, vol. 113, p. 108776, 2020.

[30] A. Aleksandrov, A. Fradkov, and A. Semenov, "Delayed and switched deployment of second-order agents on a line segment: delays and switches do not matter," *IEEE Transactions on Automatic Control*, vol. 68, no. 8, pp. 4956–4961, 2023.

[31] G. Ramos and S. Pequito, "Designing communication networks for discrete-time consensus for performance and privacy guarantees," *Systems & Control Letters*, vol. 180, p. 105608, 2023.

[32] T. C. Aysal, B. N. Oreshkin, and M. J. Coates, "Accelerated distributed average consensus via localized node state prediction," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1563–1576, 2009.

[33] B. N. Oreshkin, M. J. Coates, and M. G. Rabbat, "Optimization and analysis of distributed averaging with short node memory," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2850–2865, 2010.

[34] J.-W. Yi, L. Chai, and J. Zhang, "Convergence rate of accelerated average consensus with local node memory: Optimization and analytic solutions," *IEEE Transactions on Automatic Control*, vol. 68, no. 12, pp. 7254–7269, 2023.

[35] G. Pasolini, D. Dardari, and M. Kieffer, "Exploiting the agent's memory in asymptotic and finite-time consensus over multi-agent networks," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 6, pp. 479–490, 2020.

[36] J. Dai, J.-W. Yi, and L. Chai, "Optimal memory scheme for accelerated consensus over multi-agent networks," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 8, pp. 344–352, 2022.

[37] B. M. Chen, *Robust and H_∞ Control*. London, UK: Springer-Verlag London, 2000.

[38] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[39] W. Lin and C. I. Byrnes, "H_∞-control of discrete-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 41, no. 4, pp. 494–510, 1996.

[40] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, USA: Cambridge University Press, 2012.

[41] G. Xie and L. Wang, "Consensus control for a class of networks of dynamic agents," *International Journal of Robust and Nonlinear Control*, vol. 17, no. 10-11, pp. 941–959, 2007.

[42] P. Van Mieghem, *Graph Spectra for Complex Networks*. New York, USA: Cambridge University Press, 2023.

[43] Z. Li, Z. Duan, G. Chen, and L. Huang, "Consensus of multiagent systems and synchronization of complex networks: A unified viewpoint," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 1, pp. 213–224, 2010.

[44] L. Yu, *Robust Control: LMI Approach*. Beijing, China: Tsinghua University Press, 2002, (in Chinese).

[45] Y. Zhang and Y.-P. Tian, "Consentability and protocol design of multi-agent systems with stochastic switching topology," *Automatica*, vol. 45, no. 5, pp. 1195–1201, 2009.

Jiamin Wang was born in Yinchuan, China. He received the B.S. degree in automation from Chang'an University, Xi'an, China, in 2019, and the Ph.D. degree in control science and engineering from Xidian University, Xi'an, China, in 2024.

He is currently a Postdoctoral Fellow with the School of Mechano-Electronic Engineering, Xidian University. His research interests are currently focused on robust control, multi-agent systems, complex networks, and collective intelligence.

Jian Liu received the B.S. degree in mathematics from Fuyang Normal University, Fuyang, China, in 2007, the M.S. degree in mathematics from Donghua University, Shanghai, China, in 2010, and the Ph.D. degree in mathematics from Xi'an Jiaotong University, Xi'an, China, in 2017.

Since 2024, he has been with the Center for Multi-Agent, School of Mechano-Electronic Engineering, Xidian University, Xi'an, China. His research interests include iterative learning control, networked control systems, and complex networks.

Feng Xiao (Senior Member, IEEE) received the B.S. and M.S. degrees in mathematics from Inner Mongolia University, Hohhot, China, in 2001 and 2004, respectively, and the Ph.D. degree in systems and control from Peking University, Beijing, China, in 2008.

In 2008, he became a Faculty Member of the School of Automation, Beijing Institute of Technology, Beijing. From June 2010 to May 2013, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. From January 2016 to January 2017, he was a Visiting Professor with the Department of Mechanical Engineering, University of Victoria, Victoria, BC, Canada. He also served as a Professor with the Harbin Institute of Technology, Harbin, China. He is currently a Professor with the School of Control and Computer Engineering, North China Electric Power University, Beijing. His current research interests include group intelligence, coordination control, and networked systems.

Prof. Xiao was a recipient of the Izaak Walton Killam Postdoctoral Fellowship, the Dorothy J. Killam Memorial Postdoctoral Fellow Prize at the University of Alberta in 2010, the Program for New Century Excellent Talents in University, China, and the Excellent Young Scientists Fund by NSFC, China.

Yuanshi Zheng (Senior Member, IEEE) was born in Jiangshan, China. He received the B.S. and M.S. degrees in applied mathematics from Ningxia University, Yinchuan, China, in 2006 and 2009, respectively, and the Ph.D. degree in control science and engineering from Xidian University, Xi'an, China, in 2012.

He is currently a Professor with Xidian University. His research interests include coordination of multi-agent systems, game theory, and collective intelligence.