Automatica 83 (2017) 116–123

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Optimal tradeoff between instantaneous and delayed neighbor information in consensus algorithms☆

Weiguo Xia [a], Ziyang Meng [b], Guodong Shi [c], Karl Henrik Johansson [d]

[a] School of Control Science and Engineering, Dalian University of Technology, China
[b] State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instrument, Tsinghua University, China
[c] Research School of Engineering, The Australian National University, Australia
[d] ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Sweden

ARTICLE INFO

Article history:
Received 9 June 2015
Received in revised form 21 May 2017
Accepted 27 May 2017
Available online 1 July 2017

Keywords:
Delayed information
Optimal tradeoff
Convergence rate
Consensus algorithms

ABSTRACT

We consider a distributed consensus problem over a network, where at each time instant every node receives two pieces of information from disjoint neighboring sets: a weighted average of current states of neighbors from a primary network, and a weighted average of one-hop delayed states of neighbors from a secondary network. The proposed algorithm makes each node update its state to a weighted average of these individual averages. We show that convergence to consensus is guaranteed with non-trivial weights. We also present an explicit formula for the weights allocated to each piece of the information for the optimal rate of convergence, when the secondary network is the complement of the primary network. Finally numerical examples are given to explore the case when the neighbor sets of the agents do not cover the whole network.

1. Introduction

Recent years have witnessed great advances in the development of distributed algorithms for multi-agent systems. One benchmark problem is the design of distributed consensus algorithms that aim at driving a group of agents to reach an agreement on a variable of interest (Jadbabaie, Lin, & Morse, 2003; Moreau, 2005; Olfati-Saber & Murray, 2007). Along this research line, the impact of directed communication topologies, high-order dynamics, nonlinear interactions has been considered (Lin, Francis, & Maggiore, 2007; Liu, Slotines, & Barabasi, 2011; Wieland, Kim, & Allgower, 2011) and fruitful results have been obtained on formation control, coverage control, and network controllability (Cortes, Martinez, Karatas, & Bullo, 2003; Ren & Atkins, 2007; Scardovi & Sepulchre, 2009; Tanner, Jadbabaie, & Pappas, 2007).

In wireless communication networks, the information exchange between agents can sometimes be affected by time delays, which can have great impact on system performance. In the study of multi-agent systems, there have been continuing efforts on disclosing the effect of time delays in the process of reaching an agreement (Olfati-Saber & Murray, 2004). In Blondel, Hendrickx, Olshevsky, and Tsitsiklis (2005); Moreau (2004), the authors showed that consensus algorithms are robust to arbitrary bounded communication delays in both continuous-time and discrete-time settings. These results were extended to the case with unbounded time-varying coupling delays in Liu, Lu, and Chen (2010). Time-domain and frequency-domain approaches have been adopted to derive convergence conditions in the presence of input delays (Tian & Liu, 2009). While negative impact of time delays on the system performance was studied in the literature above, there are also research efforts making use of delayed information to accelerate the convergence of consensus algorithms. By introducing memory for each node, it has been shown that the convergence process of the consensus algorithm can speed up (Oreshkin, Coates, & Rabbat, 2010; Sarlette, 2014). In the voter model, where each voter is equipped with an individual inertia to change their opinion depending on the persistence time of a voter's current opinion, this can counter-intuitively lead to faster consensus (Stark, Tessone, & Schweitzer, 2008). In Jin and Murray (2006), a multi-hop relay protocol has been proposed for fast consensus in a network where an agent can send not only its own state but also a collection of its instantaneous neighbors' states.

☆ This work was supported in part by the National Natural Science Foundation of China (61603071, 61503249), the Fundamental Research Funds for the Central Universities (DUT15RC(3)131), Beijing Municipal Natural Science Foundation under Grant 4173075, the Knut and Alice Wallenberg Foundation, the Swedish Foundation for Strategic Research, and the Swedish Research Council. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Wei Ren under the direction of Editor Christos G. Cassandras. Correspondence of this paper should be addressed to G. Shi (+61 02 6125 5856).

E-mail addresses: wgxiaseu@dlut.edu.cn (W. Xia), ziyangmeng@mail.tsinghua.edu.cn (Z. Meng), guodong.shi@anu.edu.au (G. Shi), kallej@kth.se (K.H. Johansson).

http://dx.doi.org/10.1016/j.automatica.2017.06.021

Motivated by that network nodes often have multiple radio interfaces in practice, we consider in this paper a scenario with two networks describing the interactions between nodes. The messages exchanged in the primary network are received immediately, while in the secondary network the messages are received with a one-hop delay. For such a prototypical setup, we ask under what circumstances the consensus protocol converges. In particular, we consider that every node receives two pieces of information from two disjoint neighboring sets in the primary and secondary networks: a weighted average of current states of the primary network neighborhood and a weighted average of one-hop delayed states of the neighborhood of the secondary network. The tradeoff between current and delayed information is characterized by a parameter in the system update equation. We give conditions on this parameter to ensure the convergence of the algorithm and explore the optimal value of the parameter that leads to the fastest convergence rate when the two neighbor sets of each agent cover the whole network.

The organization of this paper is as follows. We first formulate the problem in Section 2. The convergence conditions are given in Section 3. The optimal selection on the tradeoff parameter is presented in Section 4. Numerical results are given in Section 5 to study possible extensions. Concluding remarks are given in Section 6.

2. Problem Statement

Consider a network consisting of N agents (nodes) indexed in the set $\mathcal{V} = \{1, 2, \dots, N\}$. Interactions between nodes are carried out through two networks: the primary and secondary networks, which are described by two simple undirected graphs without self-loops. The messages exchange between nodes in the primary network is instantaneous, while in the secondary network the messages are received with a one-hop delay.

Let the undirected graphs $\mathcal{G}^{(1)} = (\mathcal{V}, \mathcal{E}^{(1)})$ and $\mathcal{G}^{(2)} = (\mathcal{V}, \mathcal{E}^{(2)})$ denote the primary and secondary networks, respectively. Let $\mathcal{G} = \mathcal{G}^{(1)} \cup \mathcal{G}^{(2)} = (\mathcal{V}, \mathcal{E}^{(1)} \cup \mathcal{E}^{(2)})$ and suppose that $\mathcal{E}^{(1)} \cap \mathcal{E}^{(2)} = \emptyset$. Let $\mathcal{N}_i^{(1)} = \{j : \{i,j\} \in \mathcal{E}^{(1)}\}$ be the set of neighbors of agent i in $\mathcal{G}^{(1)}$ and $\mathcal{N}_i^{(2)}$ defined similarly. Assume that each edge $\{i,j\} \in \mathcal{E}^{(1)}$ is associated with a weight $w_{ji} > 0$, each edge $\{i,j\} \in \mathcal{E}^{(2)}$ has weight $w_{ji}^{\dagger} > 0$, and assume that self-weights $w_{ii}, w_{ii}^{\dagger}$, $i = 1, \dots, N$, are non-negative (not necessarily all positive, in contrast to the literature (Blondel et al., 2005; Jadbabaie et al., 2003; Xiao & Wang, 2006)). Assume that the weights w_{ij} and $w_{ij}^{\dagger}$ satisfy the following assumption.

Assumption 1. For all i, $\sum_{j \in \{i\} \cup \mathcal{N}_i^{(1)}} w_{ij} = 1$ and $\sum_{j \in \{i\} \cup \mathcal{N}_i^{(2)}} w_{ij}^{\dagger} = 1$.

Time is slotted as $t = 0, 1, 2, \dots$, and each node i holds a scalar value $x_i(t)$. At each time t, agent i has access to two aggregated values:
(i) The instantaneous weighted average from neighbor set $\mathcal{N}_i^{(1)}$ and itself in the primary network given by

$$\mathcal{A}_i(t) := \sum_{j \in \{i\} \cup \mathcal{N}_i^{(1)}} w_{ij} x_j(t).$$

(ii) The one-step delayed weighted average from neighbor set $\mathcal{N}_i^{(2)}$ and itself in the secondary network given by

$$\mathcal{A}_i^{\dagger}(t) := \sum_{j \in \{i\} \cup \mathcal{N}_i^{(2)}} w_{ij}^{\dagger} x_j(t-1).$$

Let $x(t) = [x_1(t), \dots, x_N(t)]^T$ and define $x(-1) = x(0)$. The aim of the network is to reach a consensus making use of the two pieces of information at each node, $\mathcal{A}_i(t)$ and $\mathcal{A}_i^{\dagger}(t)$. We propose the following simple algorithm that makes a tradeoff between the current and the delayed information exchange:

$$x_i(t+1) = (1-\beta)\mathcal{A}_i(t) + \beta \mathcal{A}_i^{\dagger}(t), \tag{1}$$

where the parameter β is a constant weight given to the delayed information.

We aim to analyze the range of β for the convergence of Algorithm (1) and its optimal value leading to the fastest convergence rate for a given network $\mathcal{G}$. Algorithm (1) relates to the algorithm studied in Jin and Murray (2006), where each agent sends to its neighbors not only its own state but also a collection of its instantaneous neighbors' states. The collection of each agent's instantaneous neighbors' states can be regarded as one-hop delayed information for the receiver. The convergence speed is accelerated compared to the original system where each agent only makes use of its neighbors' information.

Remark 1. Here we assume that $\mathcal{A}_i(t)$ and $\mathcal{A}_i^{\dagger}(t)$ are the messages received at each node i, so node i can distinguish between $\mathcal{A}_i(t)$ and $\mathcal{A}_i^{\dagger}(t)$, while it cannot infer the value of $x_j(t)$ or $x_j(t-1)$ of a neighbor j in $\mathcal{N}_i^{(1)}$ or $\mathcal{N}_i^{(2)}$, respectively. Note that $\mathcal{A}_i(t)$ can be written as $\mathcal{A}_i(t) = x_i(t) + \sum_{j=\mathcal{N}_i^{(1)}} w_{ij}(x_j(t) - x_i(t))$ for all $i = 1, \dots, N$ and $\mathcal{A}_i^{\dagger}(t)$ can be written in a similar way. In such expressions $x_i(t)$ only provides a description of the state without assuming that it is known to node i. Therefore the nodes do not have to possess the values of their absolute states according to some global coordinate system and only relative or aggregated states can be communicated (Olfati-Saber & Murray, 2007).

3. Convergence conditions

In this section, convergence conditions for Algorithm (1) are given for the case when $\mathcal{G}$ is connected and then for the case when $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$.

3.1. $\mathcal{G}$ is connected

Define $y(t) = [x^T(t)\, x^T(t-1)]^T$ with $y(0) = [x^T(0)\, x^T(-1)]^T$. Let $W_1 \in \mathbb{R}^{N\times N}$ with $[W_1]_{ii} = w_{ii}$, $[W_1]_{ij} = w_{ij}$ for $\{j, i\} \in \mathcal{E}^{(1)}$, and $[W_1]_{ij} = 0$ otherwise. Similarly $W_2 \in \mathbb{R}^{N\times N}$ is defined by $[W_2]_{ii} = w_{ii}^{\dagger}$, $[W_2]_{ij} = w_{ij}^{\dagger}$ for $\{j, i\} \in \mathcal{E}^{(2)}$, and $[W_2]_{ij} = 0$ otherwise. Let

$$\Phi(\beta) = \begin{bmatrix} (1-\beta)W_1 & \beta W_2 \\ I & \mathbf{0} \end{bmatrix},$$

where I and $\mathbf{0}$ are the identity matrix and zero matrix with compatible dimension. It is clear that W_1 and W_2 are stochastic matrices (Horn & Johnson, 1985) from Assumption 1. Algorithm (1) can be rewritten as

$$y(t+1) = \begin{bmatrix} (1-\beta)W_1 & \beta W_2 \\ I & \mathbf{0} \end{bmatrix} y(t) := \Phi(\beta) y(t). \tag{2}$$

Similar to Theorem 1 in Xiao and Boyd (2004) and Theorem 1 in Johansson and Johansson (2008), it can be shown that the necessary and sufficient conditions for Algorithm (1) converging to the average of its initial condition are (C1) $\Phi(\beta)\mathbf{1} = \mathbf{1}$; (C2) $\boldsymbol{\alpha}^T \Phi(\beta) = \boldsymbol{\alpha}^T$ for vector $\boldsymbol{\alpha}^T = [\alpha_1 \mathbf{1}^T\ \alpha_2 \mathbf{1}^T]^T$ with α_1, α_2 satisfying $\alpha_1 + \alpha_2 = 1$; and (C3) $\rho(\Phi(\beta) - 1/N \mathbf{1}\boldsymbol{\alpha}^T) < 1$, where $\mathbf{1}$ is an all-one vector with compatible dimension and $\rho(\cdot)$ is the spectral radius of a matrix. If these three conditions are satisfied, then $\lim_{t\to\infty} \Phi(\beta)^t = 1/N \mathbf{1}\boldsymbol{\alpha}^T$. The conditions (C1)–(C3) are equivalent to the condition that one is a simple eigenvalue of $\Phi(\beta)$ with $\mathbf{1}$ and $\boldsymbol{\alpha}$ being its corresponding right and left eigenvectors, respectively, and all the other eigenvalues lie inside the unit circle.

Since W_1 and W_2 are stochastic matrices, it is easy to verify that $\Phi(\beta)\mathbf{1} = \mathbf{1}$. Let $\alpha_1 = 1/(1+\beta)$ and $\alpha_2 = \beta/(1+\beta)$. It is clear that $\boldsymbol{\alpha}^T\Phi(\beta) = \boldsymbol{\alpha}^T$ since W_1 and W_2 are symmetric matrices. For condition (C3), using similar arguments to that in Proposition 1 in Oreshkin et al. (2010), we are able to show that $\Phi(\beta) - 1/N\mathbf{1}\boldsymbol{\alpha}^T$ and $\Phi(\beta) - J$ have the same spectra where $J = 1/(2N)\mathbf{1}\mathbf{1}^T$ and consequently $\rho(\Phi(\beta)-1/N\mathbf{1}\boldsymbol{\alpha}^T) = \rho(\Phi(\beta)-J)$. As discussed in Oreshkin et al. (2010), minimizing the spectral radius $\rho(\Phi(\beta)-J)$ is an optimality criterion of the convergence time for approaching the average of Algorithm (1). A sufficient condition for the convergence of Algorithm (1) can be derived.

Theorem 1. *Suppose that there exists a node $i_0 \in \mathcal{V}$ such that $w_{i_0i_0} > 0$ and $\mathcal{G} = \mathcal{G}^{(1)} \cup \mathcal{G}^{(2)}$ is connected. If $\beta \in (0, 1)$, then Algorithm (1) converges to the average of the initial state.*

Proof. It suffices to prove that one is a simple eigenvalue of $\Phi(\beta)$ and all the other eigenvalues lie inside the unit circle. Define a directed graph $\mathcal{G}(\Phi(\beta))$ corresponding to the matrix $\Phi(\beta)$ as follows: $\mathcal{G}(\Phi(\beta))$ has $2N$ nodes and has no self-loops; there exists an edge $\{j, i\}$ in $\mathcal{G}(\Phi(\beta))$ if and only if the ij-th element $(\Phi(\beta))_{ij}$ of $\Phi(\beta)$ is nonzero for $i \neq j$. Then from the structure of $\Phi(\beta)$, one knows that $\{i, N+i\}$ is an edge of $\mathcal{G}(\Phi(\beta))$ for all $i = 1, \ldots, N$. In addition, if $\{j, k\}$ is an edge of $\mathcal{G}_1$, then $\{j, k\}$ is an edge of $\mathcal{G}(\Phi(\beta))$ and if $\{j, k\}$ is an edge of $\mathcal{G}_2$, then $\{N+j, k\}$ is an edge of $\mathcal{G}(\Phi(\beta))$. Since the union graph $\mathcal{G}$ of the subgraphs $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ is connected, it is easy to show that in $\mathcal{G}(\Phi(\beta))$ there exists a directed path from a node i to every other node for all $i = 1, \ldots, N$. Hence $\mathcal{G}(\Phi(\beta))$ contains a directed spanning tree and in particular $\mathcal{G}(\Phi(\beta))$ is rooted at i_0.[1] In view of Lemma 1 in Xiao and Wang (2006), it follows that 1 is a simple eigenvalue of $\Phi(\beta)$ and all the other eigenvalues lie inside the unit circle. □

The idea of the proof is inspired by Lemma 2 in Xiao and Wang (2006), but note that in Theorem 1 it is not required that all the diagonal elements of W_1 are positive, which is in contrast to the assumption of Lemma 2 in Xiao and Wang (2006).

Remark 2. In Theorem 1, we give a sufficient condition such that the convergence of Algorithm (1) is achieved. This shows that the existence of delayed information in the secondary network will not negatively affect the convergence of the consensus algorithm as long as the parameter β is properly chosen.

Next we focus on the case when the secondary graph $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$. It turns out that a larger range of β can be established for the convergence of Algorithm (1).

3.2. $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$

Theorem 1 indicates that as long as $\beta \in (0, 1)$, convergence can be guaranteed if $\mathcal{G}$ is connected. It is however unclear how sharp this condition on β is. We show in the following that for the case when $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$ and hence $\mathcal{G}$ is a complete graph, the condition can be considerably improved. The scenario on complete graphs has been proposed in different settings in the literature such as reaching consensus in a network consisting of cooperative and adversarial agents (LeBlanc & Koutsoukos, 2011) and the Byzantine generals problem (Lamport, Shostack, & Pease, 1982).

[1] A directed graph $\mathcal{G}$ is said to contain a directed spanning tree if there exists a node that has a directed path to every other node. Such a node is called a root and we say that $\mathcal{G}$ is rooted at this node.

Calculate the characteristic polynomial of $\Phi(\beta)$ as

$$|\lambda I - \Phi(\beta)| = \left|\lambda^2 I - (1-\beta)\lambda W_1 - \beta W_2\right|. \tag{3}$$

Assume that the eigenvalues $\lambda_1(W_1), \ldots, \lambda_N(W_1)$ of W_1 and those of W_2 are ordered in decreasing order as $\lambda_N(W_1) \le \cdots \le \lambda_2(W_1) \le \lambda_1(W_1) = 1$ and $\lambda_N(W_2) \le \cdots \le \lambda_2(W_2) \le \lambda_1(W_2) = 1$. In general, it is difficult to derive the explicit expression of the eigenvalues of $\Phi(\beta)$ as functions of eigenvalues of W_1 and W_2.

If $W_1W_2 = W_2W_1$, then there exists an orthogonal matrix P that simultaneously diagonalizes W_1 and W_2 ((Horn & Johnson, 1985). One has that

$$\begin{aligned} PW_1P^T &= \mathrm{diag}(\lambda_1(W_1), \lambda_2(W_1)\ldots, \lambda_N(W_1)), \\ PW_2P^T &= \mathrm{diag}(\lambda_{i_1}(W_2), \lambda_{i_2}(W_2)\ldots, \lambda_{i_N}(W_2)), \end{aligned} \tag{4}$$

where $1 \le i_1, \ldots, i_N \le N$. Note that i_k is not necessarily equal to k for $k = 1, \ldots, N$. Eq. (3) becomes

$$|\lambda I - \Phi(\beta)| = \prod_{k=1}^{N}\left(\lambda^2 - (1-\beta)\lambda_k(W_1)\lambda - \beta\lambda_{i_k}(W_2)\right). \tag{5}$$

It is clear from (5) that for $\beta = 0$, Algorithm (1) converges to an agreement when $\mathcal{G}^{(1)}$ is connected and W_1 has at least one positive diagonal element. For $\beta = 1$, Algorithm (1) does not converge to an agreement since both 1 and -1 are roots of the polynomial (5).

It is in general difficult to check whether two matrices W_1 and W_2 commute. Since the position of the nonzero elements of W_1 and W_2 are determined by the two graphs $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$, it is an open question which graphs of $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ can guarantee that $W_1W_2 = W_2W_1$. When the union graph $\mathcal{G}$ is a complete graph, then for a proper choice of the weights, W_1 and W_2 commute. This motivates us to impose the following assumption.

Assumption 2. $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$. Moreover, the weight matrices W_1 and W_2 satisfy $w_{ij} = 1/N$ for $\{i, j\} \in \mathcal{E}_1$, and $w_{ij} = 0$, otherwise, and $w^\dagger_{ij} = 1/N$ for $\{i, j\} \in \mathcal{E}_2$, and $w^\dagger_{ij} = 0$, otherwise.

It can be verified that when Assumption 2 holds, $W_1 + W_2 = I + J$, where J has all entries equal to $1/N$. Obviously $W_2 = I + J - W_1$, so $W_1W_2 = W_2W_1$. In addition, for the orthogonal matrix P that diagonalizes W_1 in (4), $PJP^T = \mathrm{diag}(1, 0, \ldots, 0)$. We have that

$$PW_2P^T = \mathrm{diag}(1, 1-\lambda_2(W_1), \ldots, 1-\lambda_N(W_1)). \tag{6}$$

From (5), the characteristic polynomial of $\Phi(\beta)$ is given by

$$\begin{aligned} |\lambda I - \Phi(\beta)| = (\lambda - 1)(\lambda + \beta)\prod_{i=2}^{N}\Big(\lambda^2 \\ - (1-\beta)\lambda_i(W_1)\lambda - \beta(1-\lambda_i(W_1))\Big). \end{aligned} \tag{7}$$

We present the following result based on this observation.

Theorem 2. *Assume that Assumption 2 holds and $\lambda_N(W_1) \ge 0$. When $\lambda_2(W_1) \in (\frac{2}{3}, 1]$, Algorithm (1) converges to the average of the initial states of all the agents if and only if $\beta \in (\frac{\lambda_2(W_1)-1}{2\lambda_2(W_1)-1}, 1)$; when $\lambda_2(W_1) \in [0, \frac{2}{3}]$, Algorithm (1) converges to the average if and only if $\beta \in (-1, 1)$.*

Proof. Denote the eigenvalues of $\Phi(\beta)$ as $\lambda^*_i(\Phi(\beta))$ and $\lambda^{**}_i(\Phi(\beta))$, $1 \le i \le N$. First consider the case $\lambda_2(W_1) < 1$. It is clear that $\lambda^*_1(\Phi(\beta)) = 1$ and $\lambda^{**}_1(\Phi(\beta)) = -\beta$ are always roots of (7). Note that when $|\beta| \ge 1$, $|\lambda^{**}_1(\Phi(\beta))| \ge 1$. So the parameter β should lie in the interval $(-1, 1)$ for the convergence of (2) to consensus.

$$\begin{aligned}\beta_i^* &= \frac{\lambda_i^2(W_1)+2\lambda_i(W_1)-2+2\sqrt{(\lambda_i(W_1)-1)(\lambda_i^2(W_1)+\lambda_i(W_1)-1)}}{\lambda_i^2(W_1)}\\ \beta_i^{**} &= \frac{\lambda_i^2(W_1)+2\lambda_i(W_1)-2-2\sqrt{(\lambda_i(W_1)-1)(\lambda_i^2(W_1)+\lambda_i(W_1)-1)}}{\lambda_i^2(W_1)}.\end{aligned} \tag{11}$$

Box I.

From $\lambda^2-(1-\beta)\lambda_i(W_1)\lambda-\beta(1-\lambda_i(W_1))=0$, one has

$$\begin{aligned}\lambda_i^*(\Phi(\beta)) &= \frac{1}{2}\Big((1-\beta)\lambda_i(W_1)\\ &\quad+\sqrt{(1-\beta)^2\lambda_i^2(W_1)+4\beta(1-\lambda_i(W_1))}\Big)\\ \lambda_i^{**}(\Phi(\beta)) &= \frac{1}{2}\Big((1-\beta)\lambda_i(W_1)\\ &\quad-\sqrt{(1-\beta)^2\lambda_i^2(W_1)+4\beta(1-\lambda_i(W_1))}\Big).\end{aligned} \tag{8}$$

Let

$$\mathcal{J}_i[\beta,\lambda_i(W_1)] = \begin{cases}|\lambda_1^{**}(\Phi(\beta))| = |\beta|, & i=1,\\ \max(|\lambda_i^{*}(\Phi(\beta))|, |\lambda_i^{**}(\Phi(\beta))|), & i\geq 2.\end{cases} \tag{9}$$

It can be verified that $0, \lambda_1^{**}(\Phi(\beta)), \ldots, \lambda_N^{*}(\Phi(\beta)), \lambda_N^{**}(\Phi(\beta))$ are eigenvalues of $\Phi(\beta)-J$ and thus $\rho(\Phi(\beta)-J) = \max_{i=1,\ldots,N} \mathcal{J}_i[\beta,\lambda_i(W_1)]$. We specify the range of β such that $\max_{i=1,\ldots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)] < 1$.

For $i=2,\ldots,N$, to check whether the eigenvalues $\lambda_i^*(\Phi(\beta))$ and $\lambda_i^{**}(\Phi(\beta))$ are complex or real, we solve the equation

$$(1-\beta)^2\lambda_i^2(W_1)+4\beta(1-\lambda_i(W_1))=0. \tag{10}$$

This gives us the two possible solutions β_i^* and β_i^{**} in (11) (see Box I).

The expression under the square root in (11) (see Box I) is non-negative if $\lambda_i(W_1)\in[0,\frac{\sqrt{5}-1}{2}]$ and negative if $\lambda_i(W_1)\in(\frac{\sqrt{5}-1}{2},1)$. Thus when $\lambda_i(W_1)\in(\frac{\sqrt{5}-1}{2},1)$, $(1-\beta)^2\lambda_i^2(W_1)+4\beta(1-\lambda_i(W_1))>0$ for $\beta\in(-1,1)$. It follows that the eigenvalues $\lambda_i^*(\Phi(\beta))$ and $\lambda_i^{**}(\Phi(\beta))$ in (8) are both real and

$$\begin{aligned}\mathcal{J}_i[\beta,\lambda_i(W_1)] &= \frac{1}{2}\Big((1-\beta)\lambda_i(W_1)\\ &\quad+\sqrt{(1-\beta)^2\lambda_i^2(W_1)+4\beta(1-\lambda_i(W_1))}\Big).\end{aligned} \tag{12}$$

When $\lambda_i(W_1)\in[0,\frac{\sqrt{5}-1}{2}]$, β_i^* and β_i^{**} are real and $\beta_i^{**}\leq\beta_i^*$. Algebraic manipulation shows that $\beta_i^{**}\leq -1\leq\beta_i^*\leq 0$. When $\beta\in(-1,\beta_i^*)$, the expression under the square root in (8) is negative and thus the eigenvalues $\lambda_i^*(\Phi(\beta))$ and $\lambda_i^{**}(\Phi(\beta))$ are complex. It follows that

$$\mathcal{J}_i[\beta,\lambda_i(W_1)] = \sqrt{-\beta(1-\lambda_i(W_1))}. \tag{13}$$

When $\beta\in[\beta_i^*,1)$, the eigenvalues $\lambda_i^*(\Phi(\beta))$ and $\lambda_i^{**}(\Phi(\beta))$ are real and $\mathcal{J}_i[\beta,\lambda_i(W_1)]$ is the same as in (12).

For $\mathcal{J}_i[\beta,\lambda_i(W_1)]$ given by (12), it can be shown that when $\lambda_i(W_1)\in(\frac{2}{3},1)$, $\mathcal{J}_i[\beta,\lambda_i(W_1)]<1$ if $\beta>\frac{\lambda_i(W_1)-1}{2\lambda_i(W_1)-1}$; when $\lambda_i(W_1)\in[0,\frac{2}{3}]$, $\mathcal{J}_i[\beta,\lambda_i(W_1)]<1$ if $\beta\in(-1,1)$. Since $\frac{\lambda-1}{2\lambda-1}$ is an increasing function of λ, if there exists some $\lambda_i(W_1)>\frac{2}{3}$, then $\mathcal{J}_i[\beta,\lambda_i(W_1)]<1$ when $\beta\in(\frac{\lambda_2(W_1)-1}{2\lambda_2(W_1)-1},1)$; otherwise, $\mathcal{J}_i[\beta,\lambda_i(W_1)]<1$ when $\beta\in(-1,1)$.

When $\mathcal{J}_i[\beta,\lambda_i(W_1)]=\sqrt{-\beta(1-\lambda_i(W_1))}$, $\mathcal{J}_i[\beta,\lambda_i(W_1)]<1$ when $\beta\in(-1,0]$ and $\lambda_i(W_1)\in[0,1)$.

From the above derivations, we conclude that when $\lambda_2(W_1)\in(\frac{2}{3},1)$, $\max_{i=1,\ldots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]<1$ if and only if $\beta\in(\frac{\lambda_2(W_1)-1}{2\lambda_2(W_1)-1},1)$; when $\lambda_2(W_1)\in[0,\frac{2}{3}]$, $\max_{i=1,\ldots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]<1$ if only if $\beta\in(-1,1)$.

When $\lambda_2(W_1)=1$, which implies that W_1 is not connected, it can be seen from (7) that β and $1-\beta$ are eigenvalues of $\Phi(\beta)$. It is clear that β should be in the interval $(0,1)$ to guarantee the convergence of Algorithm (1) to the average consensus. The proof is completed. □

Remark 3. In Theorem 2, we give a necessary and sufficient condition on β such that the convergence of Algorithm (1) is achieved for the case when $\mathcal{G}$ is a complete graph. It is clear that the range of tradeoff parameter β is enlarged due to the nice property of complete graphs.

4. Optimal tradeoff

In this section, we calculate the optimal value of β for Algorithm (1) to achieve the fastest convergence rate. To this end, the expression $\mathcal{J}_i[\beta,\lambda_i(W_1)]$ in (9) should be available. As discussed in the previous section, for general matrices W_1 and W_2, it is difficult to derive an explicit expression for the eigenvalues of $\Phi(\beta)$. In the following, we focus on the case when the union graph $\mathcal{G}$ is a complete graph. We conduct numerical examples to explore the case when $\mathcal{G}^{(2)}$ is not the complement of $\mathcal{G}^{(1)}$ in the next section.

Finding out the optimal value of the parameter β which achieves the fastest convergence rate is equivalent to finding the solution to the optimization problem

$$\bar{\beta} = \arg\min_{\beta\in(-1,1)}\max_{i=1,\ldots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]. \tag{14}$$

The following result holds.

Theorem 3. *Suppose that Assumption 2, $\lambda_N(W_1)\geq 0$, and $\lambda_2(W_1)<1$ hold. Algorithm (1) achieves the fastest convergence rate if β takes the value $\bar{\beta}$ given in (15) (see Box II) with $\tilde{\beta}$ given in (16) (see Box III).*

Proof. The outline of the proof is as follows. We first specify the expression of $\max_{i=1,\ldots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ for $\beta\geq 0$ and then explore the optimal value of β for three different cases: $\frac{\sqrt{5}-1}{2}<\lambda_N(W_1)\leq\lambda_2(W_1)$, $\lambda_N(W_1)\leq\frac{\sqrt{5}-1}{2}<\lambda_2(W_1)$, and $\lambda_N(W_1)\leq\lambda_2(W_1)\leq\frac{\sqrt{5}-1}{2}$. For the first two cases, the optimal value of β is achieved when $\beta>0$; for the third case, the optimal value of β is achieved when $\beta<0$ and hence we further explore the expression of $\max_{i=1,\ldots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ for $\beta<0$ to identify the optimal β.

We first specify the expression of $\max_{i=1,\ldots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ for $\beta\geq 0$. When $\beta\geq 0$, the eigenvalues of $\Phi(\beta)$ are given by (8) and $\mathcal{J}_i[\beta,\lambda_i(W_1)]$ is given by (12) for $i=2,\ldots,N$. Algebraic calculation of the derivative of $\mathcal{J}_i[\beta,\lambda_i(W_1)]$ with respect to $\lambda_i(W_1)$

$$\bar{\beta}=\begin{cases}\dfrac{1}{\lambda_N(W_1)+1}, & \text{if } \dfrac{\sqrt{5}-1}{2}<\lambda_N(W_1),\\ \dfrac{3-\sqrt{5}}{2}, & \text{if } \lambda_N(W_1)\le\dfrac{\sqrt{5}-1}{2}<\lambda_2(W_1),\\ \dfrac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)}, & \text{if } \dfrac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)}+1\le\lambda_N(W_1)\le\lambda_2(W_1)\le\dfrac{\sqrt{5}-1}{2},\\ \tilde{\beta}, & \text{if } \lambda_2(W_1)\le\dfrac{\sqrt{5}-1}{2} \text{ and } \lambda_N(W_1)<\dfrac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)}+1,\end{cases} \tag{15}$$

Box II.

$$\tilde{\beta}=\frac{1}{2\lambda_2^2(W_1)(1-\lambda_N(W_1))}\Big[-(2-\lambda_2(W_1)-\lambda_N(W_1))^2-2(\lambda_N(W_1)-1)\lambda_2^2(W_1) \\ +(2-\lambda_2(W_1)-\lambda_N(W_1))\sqrt{(2-\lambda_2(W_1)-\lambda_N(W_1))^2+4(\lambda_N(W_1)-1)\lambda_2^2(W_1)}\Big] \tag{16}$$

Box III.

gives

$$\frac{\partial\mathcal{J}_i[\beta,\lambda_i(W_1)]}{\partial\lambda_i(W_1)}=\frac{1}{2}\Bigg(1-\beta+\frac{\lambda_i(W_1)(1-\beta)^2-2\beta}{\sqrt{(1-\beta)^2\lambda_i^2(W_1)+4\beta(1-\lambda_i(W_1))}}\Bigg). \tag{17}$$

It can be verified that when $\beta\in[0,\frac{3-\sqrt{5}}{2}]$, $\frac{\partial\mathcal{J}_i[\beta,\lambda_i(W_1)]}{\partial\lambda_i(W_1)}\ge 0$; when $\beta\in(\frac{3-\sqrt{5}}{2},1)$, $\frac{\partial\mathcal{J}_i[\beta,\lambda_i(W_1)]}{\partial\lambda_i(W_1)}<0$.

When $\beta\in[0,\frac{3-\sqrt{5}}{2}]$, $\mathcal{J}_i[\beta,\lambda_i(W_1)]\le\mathcal{J}_2[\beta,\lambda_2(W_1)]$ for $i=3,\dots,N$. We compare $\mathcal{J}_2[\beta,\lambda_2(W_1)]$ with $\mathcal{J}_1[\beta,\lambda_1(W_1)]=\beta$. Straightforward calculation shows that $\mathcal{J}_2[\beta,\lambda_2(W_1)]\ge\beta$ if $\beta\in[0,\frac{3-\sqrt{5}}{2}]$. It follows that $\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]=\mathcal{J}_2[\beta,\lambda_i(W_1)]$.

When $\beta\in(\frac{3-\sqrt{5}}{2},1)$, $\mathcal{J}_i[\beta,\lambda_i(W_1)]\le\mathcal{J}_N[\beta,\lambda_N(W_1)]$ for $i=2,3,\dots,N-1$. We compare $\mathcal{J}_N[\beta,\lambda_N(W_1)]$ with β. It can be shown that when $\beta\ge\frac{1}{\lambda_N(W_1)+1}$, $\beta\ge\mathcal{J}_N[\beta,\lambda_N(W_1)]$; when $\beta<\frac{1}{\lambda_N(W_1)+1}$, $\beta<\mathcal{J}_N[\beta,\lambda_N(W_1)]$. Since $\frac{1}{\lambda_N(W_1)+1}>\frac{3-\sqrt{5}}{2}$, the expression of $\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ for $\beta\ge 0$ is given in (18) (see Box IV).

By examining the derivative of $\mathcal{J}_i[\beta,\lambda_i(W_1)]$ given in (12) with respect to β, algebraic manipulations show that for $\beta\in(-1,1)$, $\frac{\partial\mathcal{J}_i[\beta,\lambda_i(W_1)]}{\partial\beta}<0$ when $\lambda_i(W_1)\in(\frac{\sqrt{5}-1}{2},1)$, and $\frac{\partial\mathcal{J}_i[\beta,\lambda_i(W_1)]}{\partial\beta}\ge 0$ when $\lambda_i(W_1)\in[0,\frac{\sqrt{5}-1}{2}]$.

We consider three different cases depending on the eigenvalues of W_1: $\frac{\sqrt{5}-1}{2}<\lambda_N(W_1)\le\lambda_2(W_1)$, $\lambda_N(W_1)\le\frac{\sqrt{5}-1}{2}<\lambda_2(W_1)$, and $\lambda_N(W_1)\le\lambda_2(W_1)\le\frac{\sqrt{5}-1}{2}$.

Case 1: $\frac{\sqrt{5}-1}{2}<\lambda_N(W_1)\le\lambda_2(W_1)$: From the calculation of $\frac{\partial\mathcal{J}_i[\beta,\lambda_i(W_1)]}{\partial\beta}$ above, one knows that $\mathcal{J}_2[\beta,\lambda_2(W_1)]$ and $\mathcal{J}_N[\beta,\lambda_N(W_1)]$ are both decreasing functions on the interval $\beta\in[0,1)$. Obviously, β is an increasing function. We can conclude from (18) (see Box IV) that $\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ is a decreasing function on $\beta\in[0,\frac{1}{\lambda_N(W_1)+1}]$ and is an increasing function on $\beta\in[\frac{1}{\lambda_N(W_1)+1},1)$. It immediately follows that the smallest value of $\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ on $\beta\in[0,1)$ is $\frac{1}{\lambda_N(W_1)+1}$, which is no greater than $\lambda_2(W_1)$ and is achieved when $\beta=\frac{1}{\lambda_N(W_1)+1}$. In the following, we show that $\lambda_2(W_1)\le\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ for $\beta\in(-1,0)$.

When $\mathcal{J}_2[\beta,\lambda_2(W_1)]$ is given by (12) with $i=2$, we examine $\mathcal{J}_2[\beta,\lambda_2(W_1)]-\lambda_2(W_1)$. It can be verified that for $\beta\in(-1,0)$, $\mathcal{J}_2[\beta,\lambda_2(W_1)]>\lambda_2(W_1)$ if $\lambda_2(W_1)\in(\frac{\sqrt{5}-1}{2},1)$, and $\mathcal{J}_2[\beta,\lambda_2(W_1)]\le\lambda_2(W_1)$ if $\lambda_2(W_1)\in[0,\frac{\sqrt{5}-1}{2}]$.

Since $\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]\ge\mathcal{J}_2[\beta,\lambda_2(W_1)]$ for $\beta<0$, one has $\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_2(W_1)]\ge\lambda_2(W_1)$ if $\lambda_2(W_1)>\frac{\sqrt{5}-1}{2}$. We can conclude that the solution to the optimization problem (14) is $\bar{\beta}=\frac{1}{\lambda_N(W_1)+1}$ and $\rho(\Phi(\bar{\beta})-J)=\frac{1}{\lambda_N(W_1)+1}$.

Case 2: $\lambda_N(W_1)\le\frac{\sqrt{5}-1}{2}<\lambda_2(W_1)$: Since $\mathcal{J}_2[\beta,\lambda_2(W_1)]$ is decreasing on the interval $\beta\in[0,\frac{3-\sqrt{5}}{2}]$ and $\mathcal{J}_N[\beta,\lambda_N(W_1)]$ is increasing on the interval $\beta\in(\frac{3-\sqrt{5}}{2},\frac{1}{\lambda_N(W_1)+1})$, in view of (18) (see Box IV), $\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ is decreasing on $\beta\in[0,\frac{3-\sqrt{5}}{2}]$ and increasing on $\beta\in(\frac{3-\sqrt{5}}{2},1)$. Thus in this case, the optimal parameter $\bar{\beta}=\frac{3-\sqrt{5}}{2}$ and $\mathcal{J}_2[\bar{\beta},\lambda_2(W_1)]=\frac{\sqrt{5}-1}{2}$.

Case 3: $\lambda_N(W_1)\le\lambda_2(W_1)<\frac{\sqrt{5}-1}{2}$: In this case, Algorithm (1) converges to the average of the initial values for $\beta\in(-1,1)$. From previous discussions, one knows that $\mathcal{J}_2[\beta,\lambda_2(W_1)]$ is increasing on the interval $\beta\in[0,\frac{3-\sqrt{5}}{2}]$ and $\mathcal{J}_N[\beta,\lambda_N(W_1)]$ is increasing on the interval $\beta\in(\frac{3-\sqrt{5}}{2},\frac{1}{\lambda_N(W_1)+1})$. This implies that $\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]\ge\mathcal{J}_2[0,\lambda_2(W_1)]=\lambda_2(W_1)$ for $\beta\in[0,1)$. But for $\beta\in(-1,0)$, it is possible that $\mathcal{J}_i[\beta,\lambda_i(W_1)]<\lambda_2(W_1)$. We next examine $\max_{i=1,\dots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ for $\beta\in(-1,0)$.

Since all the eigenvalues $\lambda_i(W_1)$ lie in the interval $[0,\frac{\sqrt{5}-1}{2}]$, we obtain $\beta_i^{**}\le -1\le\beta_i^*\le 0$ given in (11) (see Box I). Calculating the derivative of β_i^* with respect to $\lambda_i(W_1)$ leads to

$$\frac{\partial\beta_i^*}{\partial\lambda_i(W_1)}=\frac{1}{\lambda_i^3(W_1)}\Bigg(4-2\lambda_i(W_1)+\frac{-\lambda_i^3(W_1)+6\lambda_i(W_1)-4}{\sqrt{\lambda_i^3(W_1)-2\lambda_i(W_1)+1}}\Bigg). \tag{19}$$

Algebraic manipulations show that $\frac{\partial\beta_i^*}{\partial\lambda_i(W_1)}\le 0$ when $\lambda_i(W_1)\in[0,\frac{\sqrt{5}-1}{2}]$.

Since $\lambda_N(W_1)\le\cdots\le\lambda_2(W_1)\le\frac{\sqrt{5}-1}{2}$, we have that $-1\le\beta_2^*\le\cdots\le\beta_N^*\le 0$. From (13), on the interval $(-1,\beta_2^*)$, $\mathcal{J}_i[\beta,\lambda_i(W_1)]=\sqrt{-\beta(1-\lambda_i(W_1))}$ for $i=2,\dots,N$ and it is

$$\max_{i=1,\ldots,N} \mathcal{J}_i[\beta,\lambda_i(W_1)] = \begin{cases} \frac{1}{2}\left((1-\beta)\lambda_2(W_1) - 2 + \sqrt{(1-\beta)^2\lambda_2^2(W_1) + 4\beta(1-\lambda_2(W_1))}\right), & \beta \in \left[0, \frac{3-\sqrt{5}}{2}\right], \\ \frac{1}{2}\left((1-\beta)\lambda_N(W_1) - 2 + \sqrt{(1-\beta)^2\lambda_N^2(W_1) + 4\beta(1-\lambda_N(W_1))}\right), & \beta \in \left(\frac{3-\sqrt{5}}{2}, \frac{1}{\lambda_N(W_1)+1}\right), \\ \beta, & \beta \in \left[\frac{1}{\lambda_N(W_1)+1}, 1\right). \end{cases} \quad (18)$$

Box IV.

obvious that $\sqrt{-\beta(1-\lambda_N(W_1))} \geq \sqrt{-\beta(1-\lambda_i(W_1))}$ for $i = 2,\ldots,N-1$. We conclude that on the interval $\beta \in (-1, \beta_2^*)$,

$$\max_{1\leq i\leq N} \mathcal{J}_i[\beta,\lambda_i(W_1)] = \max\left\{-\beta, \sqrt{-\beta(1-\lambda_i(W_1))}\right\}. \quad (20)$$

On the interval $\beta \in [\beta_N^*, 0)$, $\mathcal{J}_i[\beta,\lambda_i(W_1)]$ are given in (12) for $i = 2,\ldots,N$. It can be shown that the derivative of $\mathcal{J}_i[\beta,\lambda_i(W_1)]$ with respect to $\lambda_i(W_1)$, given in (17), is positive for $\beta < 0$. Thus $\mathcal{J}_2[\beta,\lambda_2(W_1)] \geq \mathcal{J}_i[\beta,\lambda_i(W_1)]$ for $i = 3,\ldots,N$, on the interval $\beta \in [\beta_N^*, 0)$. We conclude that on the interval $\beta \in [\beta_N^*, 0)$,

$$\begin{aligned} &\max_{i=1,\ldots,N} \mathcal{J}_i[\beta,\lambda_i(W_1)] \\ &= \max\Big\{-\beta, \frac{1}{2}\Big((1-\beta)\lambda_2(W_1) \\ &\quad + \sqrt{(1-\beta)^2\lambda_2^2(W_1) + 4\beta(1-\lambda_2(W_1))}\Big)\Big\}. \end{aligned} \quad (21)$$

On the interval $\beta \in (\beta_2^*, \beta_N^*)$, $\mathcal{J}_N[\beta,\lambda_N(W_1)] = \sqrt{-\beta(1-\lambda_N(W_1))}$ and $\mathcal{J}_2[\beta,\lambda_2(W_1)]$ is given by (12) with $i = 2$. For $\mathcal{J}_i[\beta,\lambda_i(W_1)], i = 3,\ldots,N-1$, they are given by either (12) or (13). Since $\mathcal{J}_i[\beta,\lambda_i(W_1)]$ in (12) is an increasing function of $\lambda_i(W_1)$ and $\sqrt{-\beta(1-\lambda_i(W_1))}$ is a decreasing function of $\lambda_i(W_1)$, on the interval $\beta \in (\beta_2^*, \beta_N^*)$, one has

$$\begin{aligned} &\max_{i=1,\ldots,N} \mathcal{J}_i[\beta,\lambda_i(W_1)] \\ &= \max\Big\{-\beta, \sqrt{-\beta(1-\lambda_N(W_1))}, \frac{1}{2}\Big((1-\beta)\lambda_2(W_1) \\ &\quad + \sqrt{(1-\beta)^2\lambda_2^2(W_1) + 4\beta(1-\lambda_2(W_1))}\Big)\Big\}. \end{aligned} \quad (22)$$

We examine $-\beta$, $\mathcal{J}_N[\beta,\lambda_N(W_1)]$ and $\mathcal{J}_2[\beta,\lambda_2(W_1)]$ given in (22) as functions of β. It is easy to see that when $\beta \in [\lambda_N(W_1)-1, 0)$, $-\beta \leq \sqrt{-\beta(1-\lambda_N(W_1))}$; when $\beta \in (-1, \lambda_N(W_1)-1)$, $-\beta > \sqrt{-\beta(1-\lambda_N(W_1))}$.

When $\lambda_2(W_1) \in [2-\sqrt{2}, \frac{\sqrt{5}-1}{2}]$, $\frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)} \leq \frac{-\lambda_2(W_1)}{2-\lambda_2(W_1)}$. It can be verified that $\mathcal{J}_2[\beta,\lambda_2(W_1)] \geq -\beta$ if $\beta \in [\frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)}, 0)$, and $\mathcal{J}_2[\beta,\lambda_2(W_1)] \leq -\beta$ if $\beta \in \left(-1, \frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)}\right)$.

We conclude that if $\lambda_N(W_1) \geq \frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)} + 1$, then

$$\begin{aligned} &\max_{i=1,\ldots,N} \mathcal{J}_i[\beta,\lambda_i(W_1)] \\ &= \begin{cases} -\beta, & \beta \in \left(-1, \frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)}\right], \\ \mathcal{J}_2[\beta,\lambda_2(W_1)], & \beta \in \left(\frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)}, 0\right). \end{cases} \end{aligned} \quad (23)$$

It is clear that if $\beta = \frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)}$, $\max_{i=1,\ldots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ achieves the smallest value which is less than $\lambda_2(W_1)$. Thus $\bar{\beta} = \frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)}$ is the optimal value on the interval $\beta \in (-1, 1)$ when $\lambda_2(W_1) \in \left[2-\sqrt{2}, \frac{\sqrt{5}-1}{2}\right]$.

If $\lambda_N(W_1) < \frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)} + 1$, then

$$\begin{aligned} &\max_{i=1,\ldots,N} \mathcal{J}_i[\beta,\lambda_i(W_1)] \\ &= \begin{cases} -\beta, & \beta \in (-1, \lambda_N(W_1)-1], \\ \mathcal{J}_N[\beta,\lambda_N(W_1)], & \beta \in (\lambda_N(W_1)-1, \tilde{\beta}], \\ \mathcal{J}_2[\beta,\lambda_2(W_1)], & \beta \in (\tilde{\beta}, 0), \end{cases} \end{aligned} \quad (24)$$

where $\tilde{\beta}$ is given by solving the equation $\mathcal{J}_N[\beta,\lambda_N(W_1)] = \mathcal{J}_2[\beta,\lambda_2(W_1)]$. This gives the value of $\tilde{\beta}$ in (16) (see Box III). It follows that the optimal $\bar{\beta}$ is given by $\tilde{\beta}$.

When $\lambda_2(W_1) \in [0, 2-\sqrt{2})$, $\frac{1-2\lambda_2(W_1)}{1-\lambda_2(W_1)} > \frac{-\lambda_2(W_1)}{2-\lambda_2(W_1)}$. It can be verified that $\mathcal{J}_2[\beta,\lambda_2(W_1)] \geq -\beta$ if $\beta \in [\frac{-\lambda_2(W_1)}{2-\lambda_2(W_1)}, 0)$, and $\mathcal{J}_2[\beta,\lambda_2(W_1)] \leq -\beta$ if $\beta \in \left(-1, \frac{-\lambda_2(W_1)}{2-\lambda_2(W_1)}\right)$. Since $\lambda_N - 1 \leq 1-\sqrt{2} \leq \frac{-\lambda_2(W_1)}{2-\lambda_2(W_1)}$, thus we conclude that $\max_{i=1,\ldots,N}\mathcal{J}_i[\beta,\lambda_i(W_1)]$ is the same as in (24). The optimal value of β that achieves the fastest convergence rate is given by $\bar{\beta} = \tilde{\beta}$ in (16) (see Box III). The proof is thus completed. □

Remark 4. The optimal selection of β depends on the values of $\lambda_2(W_1)$ and $\lambda_N(W_1)$. Note that $I - W_1$ is a Laplacian matrix and $1-\lambda_1(W_1),\ldots,1-\lambda_N(W_1)$ are the eigenvalues of $I - W_1$. Distributed algorithms exist to estimate the eigenvalues of a Laplacian matrix (Franceschelli, Gasparri, Giua, & Seatzu, 2013). Therefore the optimal value of β can be calculated in a distributed way.

5. Numerical examples

Our theoretical analysis has been focused on the case when $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$, where the optimal values of β and $\rho(\Phi(\beta) - J)$ can be computed analytically. In this section, we perform numerical studies for the case when $\mathcal{G}^{(2)}$ is not the complement of $\mathcal{G}^{(1)}$.

Assume that the primary network $\mathcal{G}^{(1)}$ is an Erdös–Rényi random graph, where the presence of a possible edge between a pair of distinct vertices is independent with probability p. For each p, we generate 6 graphs with the number of nodes being 30, 60, 90, 120, 150, and 180. For $p = 0.2$, first assume that $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$ and Assumption 2 holds. Fig. 1(a) depicts the eigenvalues $\lambda_2(W_1)$ and $\rho(\Phi(\bar{\beta})-J)$ with $\bar{\beta}$ given in (15) (see Box II); the optimal parameter $\bar{\beta}$ for each network is shown in Fig. 1(b). We also plot the corresponding optimal values of $\rho(\Phi(\beta) - J)$ for the case of $W_2 = I$, where Eq. (2) becomes

$$y(t+1) = \begin{bmatrix} (1-\beta)W_1 & \beta I \\ I & \mathbf{0} \end{bmatrix} y(t). \quad (25)$$

The above equation corresponds to the algorithm studied in Oreshkin et al. (2010), Sarlette (2014) and Yang, Freeman, and Lynch

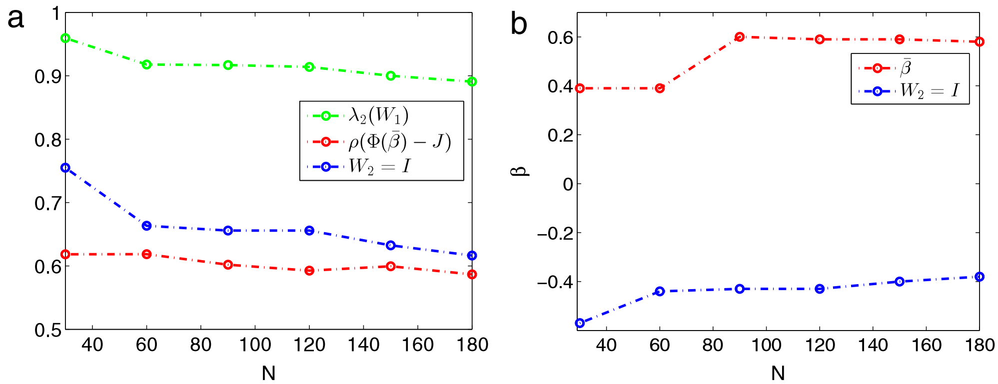


Fig. 1. (a) shows the eigenvalues $\lambda_2(W_1)$, $\rho(\Phi(\bar{\beta})-J)$ and the optimal values of $\rho(\Phi(\beta)-J)$ for $W_2 = I$ and (b) shows the corresponding optimal values of β.

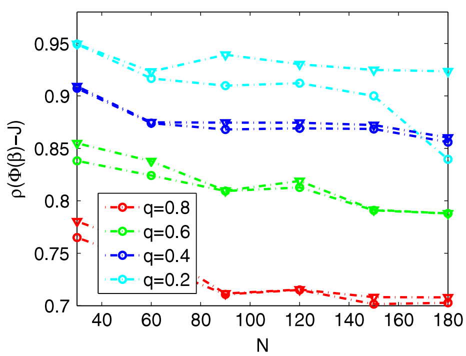


Fig. 2. $p = 0.2$, the dots with circles indicate the values of $\rho(\Phi(\beta)-J)$ with optimal choices of β, and the dots with down triangles indicate $\rho(\Phi(\bar{\beta})-J)$ with $\bar{\beta}$ calculated when $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$.

(2006) which introduces one memory register in each node to store its outdated state in order to accelerate the convergence speed. It can be seen from Fig. 1(b) that the optimal parameter β for the case of the matrix $W_2 = I$ is negative, while it takes positive values for the case when $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$.

Next, assume that an edge between a pair of distinct vertices in $\mathcal{G}^{(2)}$ exists independently with probability q. We consider four different values of q, i.e., $q = 0.2, 0.4, 0.6, 0.8$. For each q, the optimal values of $\rho(\Phi(\beta)-J)$ with optimal choices of β are plotted in Fig. 2, which are indicated with circles for different networks. It is clear that $\bar{\beta}$, which is derived for the case when $\mathcal{G}^{(2)}$ is the complement of $\mathcal{G}^{(1)}$, will not give us the optimal values of $\rho(\Phi(\beta)-J)$ in this case. To see the difference between the corresponding values of $\rho(\Phi(\bar{\beta})-J)$ and the optimal values of $\rho(\Phi(\beta)-J)$, $\rho(\Phi(\bar{\beta})-J)$ are calculated and indicated in Fig. 2 with down triangles. It can be seen that though taking $\beta = \bar{\beta}$ does not provide us the optimal value of $\rho(\Phi(\beta)-J)$, the difference between these two values are small when the probability q is no less than 0.4. When the network size becomes large and the probability q is small, $\rho(\Phi(\bar{\beta})-J)$ is not a good approximation of the optimal value of $\rho(\Phi(\beta)-J)$.

We plot two other cases of $p = 0.1$ and $p = 0.3$, which are illustrated in Fig. 3(a) and (b), respectively. It can be seen that when p is small ($p = 0.1$ and $q \le 0.6$), or when p is relatively large and q is large ($p = 0.3$ and $q \ge 0.6$), $\rho(\Phi(\bar{\beta})-J)$ is close to the optimal value of $\rho(\Phi(\beta)-J)$; while when p is relatively large and q is small ($p = 0.3$, $q \le 0.4$), the difference between $\rho(\Phi(\bar{\beta})-J)$ and the optimal value of $\rho(\Phi(\beta)-J)$ becomes large as the network size grows.

6. Conclusions

A simple algorithm has been proposed to reach an average consensus in a network by making a tradeoff between current and delayed neighbor information. It has been shown that as long as a nontrivial fraction is used of this information, consensus can be reached. The optimal way of making use of the information for the case when the secondary network is the complement of the primary network was also given. These results have shown the necessity and the optimal way of reaching consensus using both current and delayed information. Numerical examples have shown that when the secondary network is not the complement of the primary network, the optimal value $\bar{\beta}$, obtained for the case when the secondary network is the complement of the primary network, can lead to $\rho(\Phi(\bar{\beta})-J)$ close to the optimal one in some cases.

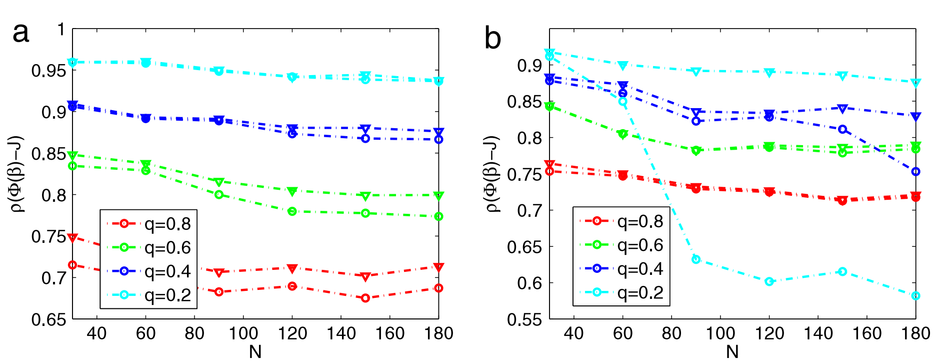


Fig. 3. Similar plots to Fig. 2 with (a) $p = 0.1$ and (b) $p = 0.3$.

References

Blondel, V. D., Hendrickx, J. M., Olshevsky, A., & Tsitsiklis, J. N. (2005). Convergence in multiagent coordination, consensus, and flocking. In *Proc. of the 44th IEEE conference on decision and control* (pp. 2996–3000).

Cortes, J., Martinez, S., Karatas, T., & Bullo, F. (2003). Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*, *20*(2), 243–255.

Franceschelli, M., Gasparri, A., Giua, A., & Seatzu, C. (2013). Decentralized estimation of Laplacian eigenvalues in multi-agent systems. *Automatica*, *49*, 1031–1036.

Horn, R. A., & Johnson, C. R. (1985). *Matrix analysis*. Cambridge, U.K: Cambridge University Press.

Jadbabaie, A., Lin, J., & Morse, A. S. (2003). Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, *48*(6), 985–1001.

Jin, Z., & Murray, R. M. (2006). Multi-hop relay protocols for fast consensus seeking. In *Proc. of the 45th IEEE conference on decision and control* (pp. 1001–1006).

Johansson, B., & Johansson, M. (2008). Faster linear interations for distributed averaging. In *Proc. of the 17th IFAC world congress* (pp. 2861–2866).

Lamport, L., Shostack, R., & Pease, M. (1982). The byzantine generals problem. *ACM Transactions on Programming Languages and Systems*, *4*, 382–401.

LeBlanc, H. J., & Koutsoukos, X. (2011). Consensus in networked multi-agent systems with adversaries. In *Proceedings of the 14th international conference on hybrid systems: Computation and control* (pp. 281–290).

Lin, Z., Francis, B., & Maggiore, M. (2007). State agreement for continuous-time coupled nonlinear systems. *SIAM Journal on Control and Optimization*, *46*(1), 288–307.

Liu, X., Lu, W., & Chen, T. (2010). Consensus of multi-agent systems with unbounded time-varying delays. *IEEE Transactions on Automatic Control*, *55*(10), 2396–2401.

Liu, Y.-Y., Slotines, J.-J., & Barabasi, A.-L. (2011). Controllability of complex networks. *Nature*, *473*, 167–173.

Moreau, L. (2004). Stability of continuous-time distributed consensus algorithms. In *Proc. of the IEEE conference on decision and control* (pp. 3998–4003).

Moreau, L. (2005). Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, *50*(2), 169–182.

Olfati-Saber, R., & Murray, R. M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, *49*(9), 1520–1533.

Olfati-Saber, R., & Murray, R. M. (2007). Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, *95*(1), 215–233.

Oreshkin, B. N., Coates, M. J., & Rabbat, M. G. (2010). Optimization and analysis of distributed averaging with short node memory. *IEEE Transactions on Signal Processing*, *58*(5), 2850–2865.

Ren, W., & Atkins, E. (2007). Distributed multi-vehicle coordinated control via local information exchange. *International Journal of Robust and Nonlinear Control*, *17*(10–11), 1002–1033.

Sarlette, A. (2014). Adding a single memory per agent gives the fastest average consensus. arXiv:1412.0402v1.

Scardovi, L., & Sepulchre, R. (2009). Synchronization in networks of identical linear systems. *Automatica*, *45*(11), 2557–2562.

Stark, H.-U., Tessone, C. J., & Schweitzer, F. (2008). Slower is faster: Fostering consensus formation by heterogeneous inertia. *Advances in Complex Systems*, *11*(4), 551–563.

Tanner, H. G., Jadbabaie, A., & Pappas, G. J. (2007). Flocking in fixed and switching networks. *IEEE Transactions on Automatic Control*, *52*(5), 863–868.

Tian, Y., & Liu, C. (2009). Robust consensus of multi-agent systems with diverse input delays and asymmetric interconnection perturbations. *Automatica*, *45*(5), 1347–1353.

Wieland, P., Kim, J. S., & Allgower, F. (2011). On topology and dynamics of consensus among linear high-order agents. *International Journal of Systems Science*, *42*(10), 1831–1842.

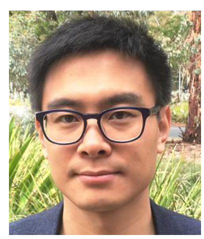
Xiao, F., & Wang, L. (2006). State consensus for multi-agent systems with switching topologies and time-varying delays. *International Journal of Control*, *79*(10), 1277–1284.

Xiao, L., & Boyd, S. (2004). Fast linear iterations for distributed averaging. *Systems & Control Letters*, *53*(1), 65–78.

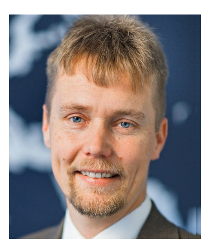
Yang, P., Freeman, R. A., & Lynch, K. M. (2006). Optimal information propogation in sensor networks. In *Proc. of the 2006 IEEE international conference on robotics and automation* (pp. 3122–3127).

Weiguo Xia received the B.Sc. and M.Sc. degrees in applied mathematics from Southeast University, Nanjing, China, in 2006 and 2009, respectively, and the Ph.D. degree from the Faculty of Mathematics and Natural Sciences, ENTEG, University of Groningen, Groningen, The Netherlands, in 2013. He is currently an Associate Professor with the School of Control Science and Engineering, Dalian University of Technology, Dalian, China. From 2013 to 2015, he was a Post-Doctoral Researcher with the ACCESS Linnaeus Centre, Royal Institute of Technology, Stockholm, Sweden. His current research interests include complex networks, social networks, and multiagent systems.

Ziyang Meng received his Bachelor degree with honors from Huazhong University of Science & Technology, Wuhan, China, in 2006, and Ph.D. degree from Tsinghua University, Beijing, China, in 2010. He was an exchange Ph.D. student at Utah State University, Logan, USA from Sept. 2008 to Sept. 2009. From 2010 to 2015, he held postdoc, researcher, and Humboldt research fellow positions at, respectively, Shanghai Jiao Tong University, Shanghai, China, KTH Royal Institute of Technology, Stockholm, Sweden, and Technical University of Munich, Munich, Germany. He joined Department of Precision Instrument, Tsinghua University, China as an associate professor since Sept. 2015. His research interests include multiagent systems, small satellite systems, distributed control and optimization, nonlinear systems and information fusion techniques. He was selected to the national 1000-Youth Talent Program of China in 2015.



Guodong Shi received his Ph.D. in Systems Theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in July 2010. From Aug. 2010 to Apr. 2014 he was a postdoctoral researcher at the ACCESS Linnaeus Centre, KTH Royal Institute of Technology, Stockholm, Sweden. Since May 2014 he has been with the Research School of Engineering, The Australian National University, Canberra, Australia, where he is now a Senior Lecturer and Future Engineering Research Leadership Fellow. His current research interests include distributed control systems, quantum networking and decisions, and social opinion dynamics.



Karl Henrik Johansson is Director of the Stockholm Strategic Research Area ICT The Next Generation and Professor at the School of Electrical Engineering, KTH Royal Institute of Technology. He received M.Sc. and Ph.D. degrees in Electrical Engineering from Lund University. He has held visiting positions at UC Berkeley, Caltech, NTU, HKUST Institute of Advanced Studies, and NTNU. His research interests are in networked control systems, cyber-physical systems, and applications in transportation, energy, and automation. He is a member of the IEEE Control Systems Society Board of Governors and the European Control Association Council. He has received several best paper awards and other distinctions, including a ten-year Wallenberg Scholar Grant, a Senior Researcher Position with the Swedish Research Council, the Future Research Leader Award from the Swedish Foundation for Strategic Research, and the triennial Young Author Prize from IFAC. He is Fellow of the IEEE and IEEE Distinguished Lecturer.