Distributed optimization for multiagent systems over general strongly connected digraph

Dong Wang[1], Zidong Wang[2], Dong Wang[1], Wei Wang[1]

1. School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, P. R. China.
Email: dwang@dlut.edu.cn, wangdong_dlut@163.com, wangwei@dlut.edu.cn

2. Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom.
E-mail: zidong.wang@brunel.ac.uk

Abstract: This paper proposes an "out-in degree" Laplacian matrix to dispose the distributed optimization problem for both the continuous-time and discrete-time multiagent systems with the first-order dynamics over a general strongly connected digraph. By making use of the out-degree and in-degree Laplacian matrices of the directed graph, we establish the parameter matrix which possesses some properties similar to the Laplacian matrix of the weight-balanced graph. Such a matrix is constructed to deal with the distributed optimization problem over a directed graph. First, we are concerned with the continuous-time case and sufficient condition for the existence of the distributed optimal protocol. Second, a similar result is established for the discrete-time case with a skillful design of Lyapunov function rather than usage of Young's inequality, which simplifies optimization and convergence analysis.

Key Words: Multiagent systems; Strongly connected graph; Out-in degree method; Distributed optimization; Lyapunov function

1 Introduction

In recent years, the distributed optimization problem for multiagent systems has drawn increasing research attention due to its extensive applications in engineering areas. In the framework of the distributed optimization, each agent possesses a local cost function which is only accessible to that agent, and the aim is to utilize the local data and the neighborhood information such that all the agents converge to the network optimal state. Such a problem is regarded as the extension of a consensus problem [1–8] in which the final state is the global optimal point. Up to now, there have been a surge of research efforts devoted to the distributed optimization problems [9–16]. References [9] and [10] have solved the unconstrained and constrained distributed optimization problems for discrete-time multiagent systems with the step-size converging to zero. Reference [11] has proposed a continuous-time algorithm based on the feedback-compensation method to solve the above problem. [12] further has provided a complete convergence proof for the continuous-time algorithm and solves the considered optimal problem for systems over a weight-balanced graph. Reference [13] has developed the event-triggered condition for a continuous-time algorithm to reduce the cost and the energy during the state information transmission.

In most of the existing literature relating to the distributed optimization problem, the communication topology among agents is assumed to be a directed strongly connected topology [17–24]. To be more specific, a push-sum algorithm has been put forward to solve the discrete-time distributed optimization problem over a directed graph in [17–20]. However, the step-size in these papers has to reduce to zero as time goes to infinity, which limits the application of the algorithm. On the other hand, most of the above results exclusively use the out-degree information of agents with which the optimal problem can be solved and consensus is achieved if and only if time converges to infinity, which leads to the algorithms conservative, complex and difficult to apply in practice. As such, there is a great need to establish some novel methods to investigate the distributed optimization problem. This constitutes the main motivation of this paper.

According to the above discussions, the distributed optimization problem for both continuous-time and discrete-time multi-agent systems based on out-degree and in-degree matrices over a general strongly connected graph is studied in this paper. The main contributions are threefold: 1) an approach to constructing "out-in degree" Laplacian matrix for the strongly connected graph is proposed, which possesses the similar properties of weight-balanced graph; 2) with the aid of the constructive "out-in degree" matrix, not only the criterion achieves optimization but also the convergence rate can be obtained by applying the Lyapunov method; and 3) the developed results are extended to the discrete-time case through an elaborate Lyapunov function. Compared with the existing results [13, 15], Young's inequality is no longer employed when dealing with the derivative of the selected Lyapunov function, thereby reducing the conservatism of the optimal condition and the difficulties handling the feedback-compensation term.

The rest of the paper is organized as follows. Section 2 formulates the distributed optimization problem over a general strongly connected graph. The algorithm for the continuous-time systems and its convergence analysis is presented in Section 3. Section 4 provides the corresponding results under the discrete-time distributed protocol. In Section 5, numerical examples are provided to show the validity of the established results. Finally, the summary and future work direction is concluded in Section 6.

Notations: B^T represents the transpose of the matrix B. 1_n and 0_n are the column vectors of n ones and n zeros respectively. I_n denotes the $n \times n$ identity matrix. The compact set is denoted by C. The real number is denoted by $\mathbb{R}$. The

This work is in part supported by the National Natural Science Foundation of China (61374070, 61473055), the Liaoning Province Science Foundation (2015020075) and Youth Star of Dalian Science and Technology (2015R052,2016RQ014).

dimension of the matrix is omitted if it is obvious from the context.

2 Problem Formulation

In this section, we formulate the distributed optimization problem over a general strongly connected graph, and provide some related assumptions and lemmas which will be used in the subsequent analysis.

Let $G = (v, \varepsilon)$ represent a graph, where $v = \{1, ..., N\}$ and $\varepsilon \subseteq v \times v$ denote the node set and the edge set, respectively. If there is an edge from i to j, denoted by (i, j), then agent i can transmit information to agent j. Here, agent i is the neighbor of agent j. If we can find at least one path between any two nodes, the graph is connected. If there is a directed path for every pair of nodes, then, the digraph is strongly connected. $A = [a_{ij}] \in \mathbb{R}^{N\times N}$ is a weighted adjacency matrix related to G such that $a_{ij} > 0$ if $(i, j) \in \varepsilon$, otherwise $a_{ij} = 0$. For each node $i \in v$, $N_i^- := \{h \in v : (i, h) \in \varepsilon\}$ and $N_i^+ := \{j \in v : (j, i) \in \varepsilon\}$ represent the set of its out-neighbors and in-neighbors, respectively. $d_i^{out} = \sum_{j=1}^{N} a_{ji}$ and $d_i^{in} = \sum_{j=1}^{N} a_{ij}$ denote the out-degree and in-degree of agent i, respectively. G is a weight-balanced graph if $d_i^{in} = d_i^{out}$ for all $i \in v$. The out-degree and in-degree Laplacian matrixes of G are defined as $L_1 = D_1 - A, L_2 = D_2 - A$, respectively, where $D_1 = diag(d_1^{out}, \ldots, d_N^{out}), D_2 = diag(d_1^{in}, \ldots, d_N^{in})$.

Consider a multiagent system with n agents labeled by $V = \{1, 2, ..., n\}$. The network topology is just strongly connected graph. Each agent $i \in V$ is assigned to a local cost function $f_i(x) : \mathbb{R}^m \to \mathbb{R}$ which is only known by agent i. Let $f(x) = \sum_{i=1}^{n} f_i(x)$. The distributed optimization is achieved if all the agents cooperatively accomplish

$$\min_{x\in R^m} \ f(x) \tag{1}$$

utilizing the local data and the communication with its neighbors.

The assumptions, lemmas and definitions are provided in what follows.

Assumption 1 *The local cost function $f_i(x_i)$ is ϖ_i-strongly convex ($\varpi_i > 0$), which indicates that it satisfies $(z - x)^T (\nabla f_i(z) - \nabla f_i(x)) \geq \varpi_i \|z - x\|^2$ for all $x, z \in R^m$.*

From Assumption 1, it is not difficult to obtain that the global cost function $f(x)$ is strictly convex, which means the uniqueness of the network optimizer.

Assumption 2 *The subgradient of the local cost function $f_i(x_i)$ is θ_i-Lipschitz ($\theta_i > 0$), which implies that $\|\nabla f_i(z) - \nabla f_i(x)\| \leq \theta_i \|z - x\|$ for all $x, z \in R^m$.*

Assumption 3 *The graph of the communication topology is strongly connected.*

Lemma 1 *(Young's inequality). Given $x, \ y \in \mathbb{R}^m$, for any $\zeta > 0$, we have*

$$x^T y \leq \frac{x^T x}{2\zeta} + \frac{\zeta}{2} y^T y$$

Lemma 2 *[4] The Laplacian matrix L of the strongly connected graph has a simple zero eigenvalue and the remaining eigenvalues with positive real parts.*

Assumption 4 *For the strongly connected graph, 0 is a simple eigenvalue of the "out-in degree" Laplacian matrix L_1L_2, and the rest eigenvalues of L_1L_2 have positive real part.*

Remark 1 *L_1L_2 is defined as the "out-in degree" Laplacian matrix for a general strongly connected graph. Generally, $L_1 \neq L_2$ for the directed graph. Note that L_1L_2 has some similar properties with the Laplacian matrix of the weight-balanced graph, such as $1_n^T L_1L_2 = 0, L_1L_21_n = 0$, which implies that we can substitute L_1L_2 for the weight-balanced Laplacian matrix to achieve some targets.*

It is not difficult to validate that 0 is the eigenvalue of L_1L_2 with the corresponding right eigenvector 1_n. Introduce an orthogonal matrix $Q = [r, R] \in \mathbb{R}^{n\times n}$ satisfying

$$r = \frac{\mathbf{1}_n}{\sqrt{n}}, \ \mathbf{1}_n^T R = \mathbf{0}_{n-1}, \ R^T R = I_{n-1},$$
$$RR^T = I_n - \frac{1_n 1_n^T}{n} \tag{2}$$

For $Z \in \mathbb{R}^n$, utilizing Lemma 2 and the properties of the above matrix yields

$$\begin{aligned} &Z^T Q^T L_1 L_2 Q Z \\ =&Z_{2:n}^T R^T L_1 L_2 R Z_{2:n} \\ =&Z_{2:n}^T R^T L_1 (RR^T + \frac{1_n 1_n^T}{n}) L_2 R Z_{2:n}. \end{aligned} \tag{3}$$

It is obtained that the eigenvalues for the product matrix of any two positive definite Hermite matrixes are positive in [25]. Although such a conclusion may not be extended to the case of the product of two generalized positive definite matrixes [26], matrix L_1L_2 is assumed to be a generalized positive semi-definite matrix with a simple zero eigenvalue.

Definition 1 *The distributed optimization (1) is achieved if the states of the agents satisfy*

$$\lim_{t\to\infty} ||x_i(t) - x^*|| = 0, \forall i \in v,$$

where x^* is a desired optimal point. Furthermore, the convergence rate for the continuous-time systems is a positive scalar δ if

$$||x_i(t) - x^*|| \leq Ce^{-\delta(t-t_0)}, \ \forall \mathrm{t} > \mathrm{t}_0$$

where C and t_0 are positive scalers. For the discrete-time case, the convergence rate is a positive scalar $\varphi \in (0, 1)$ if

$$||x_i(t) - x^*|| \leq C\varphi^{t-t_0}, \ \forall \mathrm{t} > \mathrm{t}_0.$$

The aim of this paper is to achieve consensus to the global optimal point through the local data of the agents and communicating with the neighbors over a general strongly connected graph.

3 Distributed optimization algorithm with continuous-time dynamics over a general strongly connected graph

In this section, a distributed algorithm for the continuous-time context is provided to dispose of the optimization problem described in (1) over a directed graph. Without loss of generality, the state dimension of each agent is set as 1.

Consider a continuous-time system with the following first-order dynamics:

$$\dot{x}_i = u_i, i = 1, \cdots, n \tag{4}$$

where x_i is the state of the i-th agents, and u_i is the corresponding control input.

The proposed algorithm is given as follows:

$$\begin{aligned} y_i &= \sum_{j \in N_i^+} a_{ij}(x_i - x_j) \\ u_i &= -\beta(\sum_{j \in N_i^-} a_{ji} y_i - \sum_{j \in N_i^+} a_{ij} y_j) - \alpha \nabla f_i(x_i) - \omega_i \\ \dot{\omega}_i &= \alpha\beta(\sum_{j \in N_i^-} a_{ji} y_i - \sum_{j \in N_i^+} a_{ij} y_j) \end{aligned} \tag{5}$$

where both α and β are positive scalars.

Remark 2 *It is noted that in (5) a_{ij} and a_{ji} are both employed from the Laplacian matrix. The distributed optimization cannot be achieved with the proposed algorithm in [13, 15] due to the variation of $\sum_{i=1}^{n} \omega_i(k)$. In fact, only the out-degree Laplacian matrix is used in [13, 15], hence, the graph information is not employed sufficiently. By contrast, the algorithms based on the out-degree and in-degree Laplacian matrixes are designed to achieve distributed optimization for both continuous-time and discrete-time multiagent systems over strongly connected graph in this paper.*

Theorem 1 *Suppose that Assumptions $1-4$ are satisfied. Distributed optimization with the first-order multi-agent systems (4) is reached under algorithm (5) given $\sum_{i=1}^{n} \omega_i(0) = 0$, $\alpha > 0$ and $\beta > 0$.*

Proof Combining (4) with (5) yields the following compact dynamics

$$\begin{aligned} Y &= L_2 X \\ \dot{X} &= -\beta L_1 Y - \alpha \nabla \tilde{f}(X) - W \\ \dot{W} &= \alpha\beta L_1 Y \end{aligned} \tag{6}$$

which implies

$$\begin{aligned} \dot{X} &= -\beta L_1 L_2 X - \alpha \nabla \tilde{f}(X) - W \\ \dot{W} &= \alpha\beta L_1 L_2 X \end{aligned} \tag{7}$$

where $X = [x_1, ..., x_n]^T$, $W = [w_1, ..., w_n]^T$ and $\tilde{f}(X) = \sum_{i=1}^{n} f_i(x_i)$.

Since all the column sum of L_1 are zero, we have $1_n^T L_1 = 0_n$. It follows from (7) that

$$\sum_{i=1}^{n} \dot{\omega}_i = 0 \tag{8}$$

which indicates

$$\sum_{i=1}^{n} \omega_i(t) = \sum_{i=1}^{n} \omega_i(0) = 0. \tag{9}$$

If the derivatives of closed-loop system (7) are zero, the equilibrium point (X^*, W^*) satisfies the following equalities:

$$\begin{aligned} L_1 L_2 X^* &= 0 \\ \beta L_1 L_2 X^* + W^* + \alpha \nabla \tilde{f}(X^*) &= 0. \end{aligned} \tag{10}$$

It follows (9) and (10) that

$$1_n^T \alpha \nabla \tilde{f}(X^*) = -1_n^T W^* = -\sum_{i=1}^{n} \omega_i(t) = 0. \tag{11}$$

Since L_2 is a Laplacian matrix whose row sums are zero, one has $L_2 1_n = 0$, which means 1_n is a right eigenvector for L_2. It is straightforward to obtain

$$\mathrm{X}^* = 1_n x^*. \tag{12}$$

For the stability analysis of (7), we utilize the following variable transformation to make the equilibrium point change to the origin

$$\eta = X - \mathrm{X}^*, \quad \tau = \mathrm{W} - \mathrm{W}^*. \tag{13}$$

Then, it follows from (7) that

$$\begin{cases} \dot{\eta} = -\beta L_1 L_2 \eta - \alpha h - \tau \\ \dot{\tau} = \alpha\beta L_1 L_2 \eta \end{cases} \tag{14}$$

where $h = \nabla \tilde{f}(\eta + \mathrm{X}^*) - \nabla \tilde{f}(\mathrm{X}^*)$.

Selecting the transformation in (2) yields

$$\hat{\eta} = Q^T \eta, \quad \hat{\tau} = Q^T \tau \tag{15}$$

and thus it follows from (14) that

$$\begin{cases} \dot{\hat{\eta}}_1 = -\alpha r^T h \\ \dot{\hat{\tau}}_1 = 0 \\ \dot{\hat{\eta}}_{2:n} = -\beta(R^T L_1 L_2 R)\hat{\eta}_{2:n} - \hat{\tau}_{2:n} - \alpha R^T h \\ \dot{\hat{\tau}}_{2:n} = \alpha\beta(R^T L_1 L_2 R)\hat{\eta}_{2:n} \end{cases} \tag{16}$$

where $\hat{\eta}_1$ and $\hat{\eta}_{2:n}$ denote the first row element and the rest elements of $\hat{\eta}$, respectively.

We choose the candidate Lyapunov function as follow

$$V = \frac{\hat{\eta}_1^T \hat{\eta}_1}{2} + \frac{\hat{\eta}_{2:n}^T \hat{\eta}_{2:n}}{2} + \frac{\hat{\tau}_{2:n}^T P^{-1} \hat{\tau}_{2:n}}{2\alpha\beta} \tag{17}$$

where $P = \frac{R^T(L_1 L_2 + L_2^T L_1^T)R}{2}$ is a symmetric and positive definite matrix from Assumption 4. The eigenvalues of P are arranged as $\lambda_2 \leq \cdots \leq \lambda_n$. Taking the derivative of (17) gives rise to

$$\dot{V} = -\alpha \hat{\eta}_1^T r^T h - \beta \hat{\eta}_{2:n}^T (R^T P R) \hat{\eta}_{2:n} - \alpha \hat{\eta}_{2:n}^T R^T h. \tag{18}$$

Recalling that the local cost functions satisfy Assumption 1, we have

$$-\alpha \hat{\eta}_1^T r^T h - \alpha \hat{\eta}_{2:n}^T R^T h \leq -\alpha \varpi \hat{\eta}^T \hat{\eta} \tag{19}$$

where $\varpi = min\{\varpi_1, \ldots, \varpi_n\}$. It is not difficult to obtain from (18) and (19) that

$$\dot{V} < 0. \tag{20}$$

Therefore, the solution of (14) asymptotically converges to the largest invariant set $\mathcal{M} = \{\hat{\eta}, \hat{\tau} | \hat{\eta} = 0, \hat{\tau} = 0\}$ based on the LaSalles Invariance Principle, which indicates

$$\lim_{t\to\infty} \hat{\eta}(t) = \lim_{t\to\infty} \hat{\tau}(t) = 0. \tag{21}$$

Then, it follows from (13) that

$$\lim_{t\to\infty} X(t) = X^* = 1_n \otimes x^* \tag{22}$$

which implies that the distributed optimization is achieved. ■

Remark 3 *It is worth mentioning that y_i in (5) acts as an intermediate variable after the first communication, which is omitted in the compact form (7), while it is lack of the effect to make all the agents converge to the same point. Therefore, the second communication is necessary to make the consensus be achieved.*

In Theorem 1, the agents asymptotically converge to the global optimizer for any $\alpha, \beta > 0$ with algorithm (5). We will discuss the convergence rate in the following theorem to show the performance of the algorithm.

Theorem 2 *Suppose that Assumptions $1-4$ are satisfied. All the agents under algorithm (5) for the first-order multi-agent systems (4) with $\sum_{i=1}^{n} \omega_i(0) = 0$ exponentially converge to the global optimizer with a rate no less than ψ if there exist positive scalers $\alpha > 1$ and β satisfying*

$$\begin{aligned} b_2 &= \alpha\varpi - \frac{\beta\lambda_n\mu + \theta\mu}{2\alpha} > 0 \\ b_3 &= \frac{2\mu - \beta\lambda_n - \theta}{2\alpha\mu} > 0 \\ 1 &- \beta\lambda_n > 0, \end{aligned} \tag{23}$$

where $\theta = max\{\theta_i\}$, λ_n is the largest eigenvalues of P, μ denotes a positive auxiliary parameter generated by employing Lemma 1.

Proof We utilize the equivalent transformation (16) of (7) to analyze the convergence rate of the system. Select the candidate Lyapunov function as

$$\begin{aligned} V =& \frac{\hat{\eta}_1^T\hat{\eta}_1}{2} + \frac{\hat{\eta}_{2:n}^T\hat{\eta}_{2:n}}{2} + \frac{\hat{\tau}_{2:n}^T P^{-1}\hat{\tau}_{2:n}}{2\alpha\beta} + \frac{1}{2\alpha}\hat{\eta}_{2:n}^T\hat{\tau}_{2:n} \\ &+ \frac{1}{2\alpha}\hat{\tau}_{2:n}^T\hat{\eta}_{2:n}. \end{aligned} \tag{24}$$

The conditions $1 - \beta\lambda_n > 0$ and $\alpha > 1$ are sufficient to guarantee $V > 0$. The derivative of (24) is as follow

$$\begin{aligned} \dot{V} =& -\alpha\hat{\eta}_1^T r^T h - \alpha\hat{\eta}_{2:n}^T R^T h - \frac{\beta}{\alpha}\hat{\eta}_{2:n}^T P\hat{\tau}_{2:n} \\ &- \frac{1}{\alpha}\hat{\tau}_{2:n}^T\hat{\tau}_{2:n}^T - \frac{1}{\alpha}\hat{\tau}_{2:n}^T R^T h. \end{aligned} \tag{25}$$

By making use of Lemma 1, we have

$$\begin{aligned} -\hat{\eta}_{2:n}^T P\hat{\tau}_{2:n} &\le \lambda_n|\hat{\eta}_{2:n}^T\hat{\tau}_{2:n}| \le \frac{\lambda_n\mu}{2}\hat{\eta}_{2:n}^T\hat{\eta}_{2:n} + \frac{\lambda_n}{2\mu}\hat{\tau}_{2:n}^T\hat{\tau}_{2:n} \\ -\hat{\tau}_{2:n}^T R^T h &\le \theta|\hat{\tau}_{2:n}^T\hat{\eta}_{2:n}| < \frac{\theta\mu}{2}\hat{\eta}_{2:n}^T\hat{\eta}_{2:n} + \frac{\theta}{2\mu}\hat{\tau}_{2:n}^T\hat{\tau}_{2:n}. \end{aligned} \tag{26}$$

Using (19), (25) and (26) leads to

$$\dot{V} \le -b_1\hat{\eta}_1^T\hat{\eta}_1 - b_2\hat{\eta}_{2:n}^T\hat{\eta}_{2:n} - b_3\hat{\tau}_{2:n}^T\hat{\tau}_{2:n} \tag{27}$$

where

$$\begin{aligned} b_1 &= \alpha\varpi \\ b_2 &= \alpha\varpi - \frac{\beta\lambda_n\mu + \theta\mu}{2\alpha} \\ b_3 &= \frac{2\mu - \beta\lambda_n - \theta}{2\alpha\mu}. \end{aligned}$$

According to (24) and Lemma 1, one has

$$V \le e_1\hat{\eta}_1^T\hat{\eta}_1 + e_2\hat{\eta}_{2:n}^T\hat{\eta}_{2:n} + e_3\hat{\tau}_{2:n}^T\hat{\tau}_{2:n} \tag{28}$$

where $e_1 = \frac{1}{2}, e_2 = \frac{\alpha+1}{2\alpha}, e_3 = \frac{\beta\lambda_2+1}{2\alpha\beta\lambda_2}$. It follows from (27) and (28) that

$$\dot{V} \le -\psi V \tag{29}$$

where $\psi = min\{\frac{b_i}{e_i}\}$,∨$i = 1, 2, 3$. Then, we have $V(t) \le V(0)\mathrm{e}^{-\psi\mathrm{t}}$, which implies the agents exponentially converge to the network optimizer with the rate no less than ψ from Definition 1. ■

Remark 4 *Most of the existing results [9, 10, 16] regarding the optimization problem are based on undirected graph or weight-balanced graph. Nevertheless, the distributed optimization problem in the present paper only requires that the directed graph is strongly connected, which is more general and weaker. In addition, the continuous-time algorithm based on in-degree and out-degree information has an advantage over those developed in [17–19]. Particularly, we utilize the double communication to compensate the lack of information resulting from the directed graph, which guarantees the validity of (8). This is the key point whether the distributed optimization based on feedback-compensation method is achieved or not.*

4 Distributed optimization algorithm with discrete-time dynamics over a general strongly connected graph

In this section, we propose the discrete-time distributed algorithm to address the optimization problem in (1) over a directed graph.

The dynamics of the discrete-time systems is described by

$$\frac{x_i(k+1) - x_i(k)}{\epsilon} = u_i(k), i = 1, \cdots, n \tag{30}$$

where $\epsilon > 0$ is the step-size of the discrete-time system. The proposed algorithm for the discrete-time systems is

$$\begin{aligned} u_i(k) =& -(\sum_{j\in N_i^-} a_{ji}y_i(k) - \sum_{j\in N_i^+} a_{ij}y_j(k)) \\ &- \nabla f_i(x_i(k)) - \omega_i(k) \\ y_i(k+1) =& \sum_{j\in N_i^+} a_{ij}(x_i(k+1) - x_j(k+1)) \end{aligned}$$

$$\omega_i(k+1)=\omega_i(k)+\epsilon(\sum_{j\in N_i^-}a_{ji}y_i(k+1)-\sum_{j\in N_i^+}a_{ij}y_j(k+1)). \quad (31)$$

Remark 5 *Actually,* (31) *can be written as the following equivalent form*

$$\begin{aligned}y_i(k)&=\sum_{j\in N_i^+}a_{ij}(x_i(k)-x_j(k))\\ \omega_i(k)&=\omega_i(k-1)+\epsilon(\sum_{j\in N_i^-}a_{ji}y_i(k)-\sum_{j\in N_i^+}a_{ij}y_j(k))\\ u_i(k)&=-(\sum_{j\in N_i^-}a_{ji}y_i(k)-\sum_{j\in N_i^+}a_{ij}y_j(k))-\nabla f_i(x_i(k))-\omega_i(k) \quad (32)\end{aligned}$$

The main difference between (31) and (32) lies in that the initial states in (31) are $x_i(0)$ and $w_i(0)$ while $x_i(1)$ and $w_i(0)$ in (32). Thus, we can observe from (32) that the times of network communication are twice for one iteration. We adopt the algorithm (31) to better explain the convergence of the algorithm.

Theorem 3 *Suppose that Assumptions $1-4$ are satisfied. The distributed optimization problem (1) over a general strongly connected graph is solved with protocol (31) for the first-order multi-agent systems (30) with $\sum_{i=1}^{n}\omega_i(0)=0$ under a convergence rate no less than ρ if $\epsilon<min\{\frac{1}{\sqrt{\kappa}},\frac{2\varpi}{\theta^2},\frac{2\lambda_2+2\varpi}{\gamma_n+2\theta\lambda_n+\theta^2}\}$.*

Proof The compact form for the discrete time case is

$$\begin{aligned}X(k+1)&=X(k)-\epsilon(L_1L_2X(k)+\nabla\tilde{f}(X(k))+W(k))\\ W(k+1)&=W(k)+\epsilon L_1L_2X(k+1). \quad (33)\end{aligned}$$

By making use of the following transformation and (2)

$$\phi(k)=X(k)-\mathrm{X}^*,\quad \sigma(k)=\mathrm{W(k)}-\mathrm{W}^* \quad (34)$$

$$\hat{\phi}(k)=Q^T\phi(k),\quad \hat{\sigma}(k)=Q^T\sigma(k), \quad (35)$$

one has

$$\begin{aligned}\hat{\phi}_1(k+1)&=\hat{\phi}_1(k)-\epsilon r^Th(k)\\ \hat{\sigma}_1(k+1)&=\hat{\sigma}_1(k)\\ \hat{\phi}_{2:n}(k+1)&=\hat{\phi}_{2:n}(k)-\epsilon R^TL_1L_2R\hat{\phi}_{2:n}(k)-\epsilon R^Th(k)-\epsilon\hat{\sigma}_{2:n}(k)\\ \hat{\sigma}_{2:n}(k+1)&=\hat{\sigma}_{2:n}(k)+\epsilon R^TL_1L_2R\hat{\phi}_{2:n}(k+1). \quad (36)\end{aligned}$$

Select the Lyapunov function as follows

$$V(k)=V_1(k)+(1-\kappa\epsilon^2)V_2(k)+V_3(k) \quad (37)$$

where

$$\begin{aligned}V_1(k)&=\hat{\phi}_1^T(k)\hat{\phi}_1(k)\\ V_2(k)&=\hat{\phi}_{2:n}^T(k)\hat{\phi}_{2:n}(k)\\ V_3(k)&=\hat{\sigma}_{2:n}^T(k)(R^TPR)^{-1}\hat{\sigma}_{2:n}(k).\end{aligned}$$

Taking the forward difference of (37) leads to

$$\begin{aligned}\nabla V_1(k)&=\hat{\phi}_1^T(k+1)\hat{\phi}_1(k+1)-\hat{\phi}_1^T(k)\hat{\phi}_1(k)\\ &=-2\epsilon\hat{\phi}_1^T(k)r^Th(k)+\epsilon^2h^T(k)rr^Th(k), \quad (38)\end{aligned}$$

$$\begin{aligned}\nabla V_2(k)&=\hat{\phi}_{2:n}^T(k+1)\hat{\phi}_{2:n}(k+1)-\hat{\phi}_{2:n}^T(k)\hat{\phi}_{2:n}(k)\\ &=-2\epsilon\hat{\phi}_{2:n}^T(k)R^TPR\hat{\phi}_{2:n}(k)-2\epsilon\hat{\phi}_{2:n}^T(k)R^Th(k)\\ &-2\epsilon\hat{\phi}_{2:n}^T(k)\hat{\sigma}_{2:n}(k)+\epsilon^2\hat{\phi}_{2:n}^T(k)R^TMR\hat{\phi}_{2:n}(k)\\ &+2\epsilon^2\hat{\phi}_{2:n}^T(k)R^TPR\hat{\sigma}_{2:n}(k)+\epsilon^2h^T(k)RR^Th(k)\\ &+2\epsilon^2\hat{\phi}_{2:n}^T(k)R^TPRR^Th(k)+\epsilon^2\hat{\sigma}_{2:n}^T(k)\hat{\sigma}_{2:n}(k)\\ &+2\epsilon^2h^T(k)R\hat{\sigma}_{2:n}(k), \quad (39)\end{aligned}$$

$$\begin{aligned}\nabla V_3(k)&=2\epsilon\hat{\phi}_{2:n}^T(k+1)\hat{\phi}_{2:n}(k)\\ &+\epsilon^2\hat{\phi}_{2:n}(k+1)^TR^TSR\hat{\phi}_{2:n}(k+1)\\ &\leq 2\epsilon\hat{\phi}_{2:n}^T(k)\hat{\sigma}_{2:n}(k)-2\epsilon^2\hat{\phi}_{2:n}^T(k)R^TPR\hat{\sigma}_{2:n}(k)\\ &-2\epsilon^2h^T(k)R\hat{\sigma}_{2:n}(k)-2\epsilon^2\hat{\sigma}_{2:n}^T(k)\hat{\sigma}_{2:n}(k)\\ &+\kappa\epsilon^2(\nabla V_2(k)+\hat{\phi}_{2:n}^T(k)\hat{\phi}_{2:n}(k)). \quad (40)\end{aligned}$$

By means of Assumption 1, we obtain

$$\hat{\phi}_1^T(k)r^Th(k)+\hat{\phi}_{2:n}^T(k)R^Th(k)\geq\varpi\hat{\phi}^T(k)\hat{\phi}(k). \quad (41)$$

From Assumption 2, it follows that

$$h^T(k)rr^Th(k)+h^T(k)RR^Th(k)\leq\theta^2\hat{\phi}^T(k)\hat{\phi}(k) \quad (42)$$

where

$$\begin{aligned}P&=(L_1L_2+L_2^TL_1^T)/2\\ M&=L_2^TL_1^TL_1L_2\\ S&=L_2^TL_1^TR(R^TPR)^{-1}R^TL_1L_2.\end{aligned}$$

From Assumption 4, $\lambda_1,\dots,\lambda_n$ satisfying $0=\lambda_1<\lambda_2\leq\cdots\leq\lambda_n$ denote the eigenvalues of P. The eigenvalues of M are denoted by $0=\gamma_1<\gamma_2\leq\cdots\leq\gamma_n$. The largest eigenvalue of S is denoted by κ. Combining with (37)-(42) yields

$$\begin{aligned}\nabla V(k)&\leq -l_1\hat{\phi}_1^T(k)\hat{\phi}_1(k)-l_2\hat{\phi}_{2:n}^T(k)\hat{\phi}_{2:n}(k)\\ &-l_3\hat{\sigma}_{2:n}^T(k)\hat{\sigma}_{2:n}(k) \quad (43)\end{aligned}$$

where

$$\begin{aligned}l_1&=2\epsilon\varpi-\epsilon^2\theta^2\\ l_2&=2\epsilon\lambda_2+2\epsilon\varpi-\epsilon^2\gamma_n-2\epsilon^2\theta\lambda_n-\epsilon^2\theta^2\\ l_3&=\epsilon^2.\end{aligned}$$

With the condition $\epsilon<min\{\frac{1}{\sqrt{\kappa}},\frac{2\varpi}{\theta^2},\frac{2\lambda_2+2\varpi}{\gamma_n+2\theta\lambda_n+\theta^2}\}$, we have $V(k)>0,\nabla V(k)<0$. Then, the distributed optimization is achieved. As for the convergence rate, with (37) and (43), we get the inequality $V(k+1)\leq\rho V(k)$, $\rho=1-min\{l_1,\frac{l_2}{1-\kappa\epsilon^2},l_3\lambda_2\}$. Furthermore, it is not difficult to obtain that $\|x_i(t)-x^*\|^2\leq C\rho^{t-t_0},\forall\mathrm{t}>\mathrm{t}_0$. Thus, the convergence rate of the discrete-time distributed optimization protocol is at least ρ from Definition 1. ■

Remark 6 *For the discrete time case, $X(k+1)$ is used to act as the iteration term of $W(k+1)$ instead of $X(k)$ in (33). The advantages are threefold:*

(1) *Motivated by the improved Jacobi iterative method, which uses the the former iterative result to calculate the latter value for each iteration, the iterative data is sufficiently used when $X(k)$ is replaced by $X(k+1)$ in (33). The convergence rate and stability of the system can be improved due to the introduction of the latest data in the undirected graph. Nevertheless, there is little difference between the use of $X(k)$ and $X(k+1)$ over the directed graph, which may be caused by the property variation of the Laplacian matrix L_1L_2.*

(2) *Although the Young's inequality is valid to deal with the cross terms, it also enhances the conservatism of the obtained conditions. This causes the difficulty in selecting the design parameters. In Theorem 3, the Young's inequality is discarded via constructing the proper candidate Lyapunov function. From this point of view, the derived condition has less conservatism.*

(3) *The terms containing $\hat{\sigma}_{2:n}$ except $\hat{\sigma}_{2:n}^T\hat{\sigma}_{2:n}$ are eliminated through design of Lyapunov function, which gets rid of the problem to analyze the effect of $\hat{\sigma}_{2:n}$. Generally, the terms with $\hat{\sigma}_{2:n}$ are dealt with Young's inequality, which causes certain challenges in the stability analysis of the multi-agent systems.*

Remark 7 *For the discrete-time multiagent systems, the Lyapunov method is employed to establish the stability criteria of the closed-loop system and achieve the distributed optimization. In comparison with the literature [9, 10], where the proof procedures are quite complex and mass, an appropriate Lyapunov function is selected in this paper to simplify the proof. In addition, we avoid utilizing the Young's inequality and hence the obtained results are less conservatism and the established condition is easy to achieve.*

Remark 8 *Theorem 3 illustrates that the system is definitely stable when the step-size ϵ is sufficiently small, which is corresponding to the result in Theorem 1 that the continuous-time system is stable for any positive parameters α and β.*

Remark 9 *In multiagent systems, all the agents move according to the control protocol, which could account for the variation of the communication topology due to the communication link failure when distances between agents go beyond the data transmission radius of sensors. Switching topologies with strongly connected property are discussed for consensus of multiagent systems in [1]. The distributed optimization problem over directed graph with similar requirements is capable of being solved with (31) if $\epsilon < min\{\frac{1}{\sqrt{\overline{\kappa}}}, \frac{2\varpi}{\theta^2}, \frac{2\underline{\lambda}_2+2\varpi}{\overline{\gamma}_n+2\theta\overline{\lambda}_n+\theta^2}\}$ and all the possible topologies are strongly connected, where $\overline{\kappa} = max\{\kappa_n(G_s)\}, \underline{\lambda}_2 = min\{\lambda_2(G_s)\}, \overline{\lambda}_n = max\{\lambda_n(G_s)\}, \overline{\gamma}_n = max\{\gamma_n(G_s)\}, s = 1, 2, \cdots, G_s$ represents the switching topology.*

Remark 10 *In the distributed optimization problem solved by the protocols (5) and (31), it is necessary to require that the graph is strongly connected. Such protocols are not applied to the connected graph. The reason is that there may exist an agent just receiving or sending information over the connected graph, which implies the elements of a row or a column for the corresponding Laplacian matrix L_1L_2 are all zero. The distributed optimization is not achieved owing to the inability of updating information or the lack of the information for this agent.*

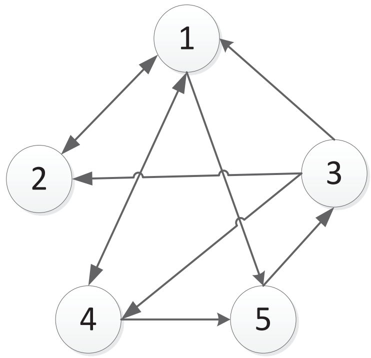

Fig. 1: The communication topology graph.

5 Simulation

In this section, numerical examples are given to testify the main results.

Consider a multiagent system with five agents over a general strongly connected graph. The local cost functions satisfying Assumptions 1-2 are selected as follows

$$\begin{aligned} f_1(x) &= x^2 + 0.1cos(x) + 3 \\ f_2(x) &= 0.1e^{-0.2x} + x + 5 \\ f_3(x) &= x^2 - 2x + 3 \\ f_4(x) &= 4x^2 + sin(0.2x) \\ f_5(x) &= 0.3e^{-0.3x} - 2x. \end{aligned}$$

It is observed that the whole cost function $f(x)$ is divided into five local cost functions and each local cost function is assigned to an agent such that all the agents cooperatively reach globally optimal point. Note that $f(x) = \sum_{i=1}^{5} f_i(x)$ is convex. The adjacency matrix for the strongly connected graph is formulated as a binary matrix, where each element is either 1 or 0. The communication topology is shown as Fig. 1. Then, the corresponding out-degree and in-degree Laplacian matrixes are described by

$$L_1 = \begin{bmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 3 & 0 & -1 \\ -1 & 0 & -1 & 2 & 0 \\ -1 & 0 & 0 & -1 & 1 \end{bmatrix}$$

and

$$L_2 = \begin{bmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ -1 & 0 & -1 & 2 & 0 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix}.$$

The initial state of x is $x(0) = [1,\ -1,\ 0,\ 3,\ 2]^T$. Parameters α and β are given as $\alpha = \beta = 1$. The trajectories of each agent controlled by algorithm (5) with continuous-time systems over a general strongly connected graph are shown

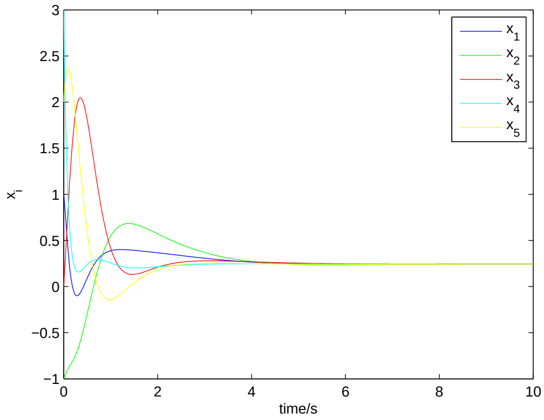

Fig. 2: The trajectories of the five agents for continuous-time systems over a general strongly connected graph with $\alpha = \beta = 1$.

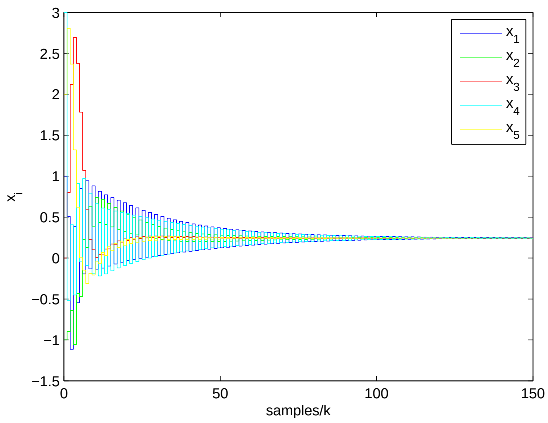

Fig. 3: The trajectories of the five agents for discrete-time systems over a general strongly connected graph with $\epsilon = 0.1$

in Fig. 2, where the states of the agents globally asymptotically converge to the network optimizer $x^* = 0.2439$. Fig. 3 provides the trajectories of the agents under discrete-time systems with $\epsilon = 0.1$. The simulation results illustrate the effectiveness of our results.

6 Conclusion

In this paper, the distributed optimization algorithm has been proposed for both the continuous-time and discrete-time multiagent systems over a general strongly connected graph. The optimization condition and convergence rate have been obtained for the considered systems based on the Lyapunov method. Moreover, numerical simulations have also been provided to testify the theoretical results. One of the future research topics is to establish the algorithm for average consensus over a general strongly connected graph.

References

[1] 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.

[2] H. Liu, L. Cheng, M. Tan, Z. Hou, and Y. Wang, "Distributed exponential finite-time coordination of multi-agent systems: containment control and consensus," *International Journal of Control*, vol. 88, no. 2, pp. 237–247, 2015.

[3] Y. Zhang and Y. P. Tian, "Allowable delay bound for consensus of linear multi-agent systems with communication delay," *International Journal of Systems Science*, vol. 45, no. 10, pp. 2172–2181, 2014.

[4] J. Mei, W. Ren, and J. Chen, "Distributed consensus of second-order multi-agent systems with heterogeneous unknown inertias and control gains under a directed graph," *IEEE Transactions on Automatic Control*, vol. 61, no. 8, pp. 2019–2034, 2016.

[5] Y. Hong, J. Hu, and L. Gao, "Tracking control for multi-agent consensus with an active leader and variable topology ," *Automatica*, vol. 42, no. 7, pp. 1177–1182, 2007.

[6] K. Cai and H. Ishii, "Average consensus on general strongly connected digraphs ," *Automatica*, vol. 48, no. 11, pp. 2750–2761, 2012.

[7] W. Zhu, Z. P. Jiang, and G. Feng, "Event-based consensus of multi-agent systems with general linear models ," *Automatica*, vol. 50, no. 2, pp. 552–558, 2014.

[8] Y. Su and J. Huang, "Cooperative global robust output regulation for nonlinear uncertain multi-agent systems in lower triangular form," *IEEE Transactions on Automatic Control*, vol. 60, no. 9, pp. 2378–2389, 2015.

[9] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[10] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.

[11] J. Wang and N. Elia, "Control approach to distributed optimization," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 557–561, 2010.

[12] B. Gharesifard and J. Corts, "Distributed continuous-time convex optimization on weight-balanced digraphs," *IEEE Transactions on Automatic Control*, pp. 781–786, 2014.

[13] S. S. Kia, J. Corts, and S. Martłnez, "Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication," *Automatica*, vol. 55, pp. 254–264, 2015.

[14] P. Yi, Y. Hong, and F. Liu, "Distributed gradient algorithm for constrained optimization with application to load sharing in power systems ," *Systems & Control Letters*, vol. 83, no. 711, pp. 45–52, 2015.

[15] X. Wang, Y. Hong, and H. Ji, "Distributed optimization for a class of nonlinear multiagent systems with disturbance rejection," *IEEE Transactions on Cybernetics*, vol. 46, pp. 1655–1666, July 2016.

[16] J. Lu and C. Y. Tang, "Zero-gradient-sum algorithms for distributed convex optimization: The continuous-time case," *IEEE Transactions on Automatic Control*, vol. 57, no. 9, pp. 5474–5479, 2011.

[17] A. Nedic and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 601–615, 2013.

[18] A. Nedic and A. Olshevsky, "Stochastic gradient-push for strongly convex functions on time-varying directed graphs," *IEEE Transactions on Automatic Control*, vol. to be published, pp. 1–12, 2016.

[19] K. I. Tsianos, S. Lawlor, and M. G. Rabbat, "Push-sum distributed dual averaging for convex optimization," in *Conference on Decision and Control*, pp. 5453–5458, 2012.

[20] D. Yuan, S. Xu, and J. Lu, "Gradient-free method for distributed multi-agent optimization via push-sum algorithms," *International Journal of Robust & Nonlinear Control*, vol. 25, no. 10, pp. 1569–1580, 2015.

[21] B. Touri and B. Gharesifard, "Continuous-time distributed convex optimization on time-varying directed networks," in *2015 IEEE 54th Conference on Decision and Control (CDC)*, pp. 724–729, 2015.

[22] B. Touri and B. Gharesifard, "Saddle-point dynamics for distributed convex optimization on general directed graphs," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 862–866, 2016.

[23] C. Xi, Q. Wu, and U. A. Khany, "Fast distributed optimization over directed graphs," in *American Control Conference*, pp. 6507–6512, 2016.

[24] A. Makhdoumi and A. Ozdaglar, "Graph balancing for distributed subgradient methods over directed graphs," in *2015 IEEE 54th Conference on Decision and Control (CDC)*, pp. 1364–1371, 2015.

[25] H. Y. Sha, "Estimation of the eigenvalues of AB for A>0, B>0," *Linear Algebra & Its Applications*, vol. 73, pp. 147–150, 1986.

[26] C. R. Johnson, "Positive definite matrices," *American Mathematical Monthly*, vol. 77, no. 3, pp. 69–75, 1970.