Nonlinear Analysis: Modelling and Control, Vol. 28, No. 3, 412–430
https://doi.org/10.15388/namc.2023.28.31563

Distributed optimization for multi-agent systems with communication delays and external disturbances under a directed network*

Xuening Xu[a], **Zhiyong Yu**[a,1] **Da Huang**[b] **Haijun Jiang**[a]

[a]College of Mathematics and System Sciences,
Xinjiang University,
Urumqi 830017, China
yzygsts@163.com

[b]Department of Mathematics and Physics,
Xinjiang Institute of Engineering,
Urumqi 830023, China

Received: May 26, 2022 / **Revised:** February 7, 2023 / **Published online:** February 22, 2023

Abstract. This article studies the distributed optimization problem for multi-agent systems with communication delays and external disturbances in a directed network. Firstly, a distributed optimization algorithm is proposed based on the internal model principle in which the internal model term can effectively compensate for external environmental disturbances. Secondly, the relationship between the optimal solution and the equilibrium point of the system is discussed through the properties of the Laplacian matrix and graph theory. Some sufficient conditions are derived by using the Lyapunov–Razumikhin theory, which ensures all agents asymptotically reach the optimal value of the distributed optimization problem. Moreover, an aperiodic sampled-data control protocol is proposed, which can be well transformed into the proposed time-varying delay protocol and analyzed by using the Lyapunov–Razumikhin theory. Finally, an example is given to verify the effectiveness of the results.

Keywords: distributed optimization, internal model principle, directed network, multi-agent systems, Lyapunov–Razumikhin theory.

1 Introduction

During the last decade, the distributed optimization problem of multi-agent systems has been widely studied due to its potential applications in sensor networks [1], resource allocation over networks [5], economic dispatch [2], and so on. The core of the distributed

*This work was supported in part by the National Natural Science Foundation of China (grant Nos. 62003289, 62163035), in part by the China Postdoctoral Science Foundation (grant No. 2021M690400), in part by the Special Project for Local Science and Technology Development guided by the Central Government (grant No. ZYYD2022A05), and in part by Xinjiang Key Laboratory of Applied Mathematics (grant No. XJDX1401).

[1]Corresponding author.

optimization problem is to minimize the sum of the local objective functions of all agents in a distributed way in which each agent only accesses its local cost function. The distributed optimization problem, as contrasted with multi-agent system consensus [13, 18, 19], minimizes the optimization function while also forcing all agents to the same state.

In the study of distributed optimization, how to design effective algorithms has become a hot topic for various multi-agent systems. The existing distributed optimization algorithms can be roughly divided into two categories: discrete-time algorithms [10, 12, 23] and continuous-time algorithms [4, 6, 9]. It is well recognized that the design of the distributed optimization strategy for practical large-scale systems is apt to encounter two main issues simultaneously. On the one hand, the communication delays will affect the dynamics of the system and even destroy the stability due to the finite bandwidth of the network. To solve the distributed optimization problem with communication delays, some distributed algorithms were given in [21, 22]. In [21], the distributed optimization is addressed based on a continuous-time multi-agent system in the presence of time-varying communication delays by the Lyapunov stability theory. When the number of agents is large, verifying the solvability of the linear matrix inequality (LMI) conditions becomes difficult. Based on the Lyapunov–Razumikhin theorem, the authors propose a distributed controller [22] that allows agents to achieve consensus while enduring interconnection delays, avoiding the need to verify the LMI conditions, and reducing the computation burden. On the other hand, external disturbances are also unavoidable in many physical systems, which may degrade the performance of systems. In existing works, the internal model principle can effectively restrain some uncertain disturbances caused by the external environment. In [11], the heterogeneous linear multi-agent system subject to external disturbances was investigated, and it was proved that the protocol devised by the internal model principle can make all agents reach the optimal output. Furthermore, there are some distributed optimization algorithms in the current studies. For instance, motivated by the internal model idea, a distributed optimization protocol was proposed for a class of nonlinear multi-agent systems with external disturbances in [16]. To solve distributed optimization for continuous-time multi-agent systems with unknown frequency disturbances, a distributed optimization algorithm by using an adaptive internal model approach was designed in [17].

It is not difficult to find that all the above works [11, 16, 17, 21, 22] are focused only on one issue while ignoring the other. To the best of the authors' knowledge, the distributed optimization problem for multi-agent systems by simultaneously considering the communication delays and disturbances was rarely studied. Although the distributed optimal problem of the system with both communication delays and external disturbances was investigated and some sufficient conditions in terms of LMIs were derived based on the Lyapunov–Krasovskii functionals in [15], sufficient conditions are only applicable to undirected networks, and it is difficult to verify the LMI conditions when the network size is very large. Additionally, in some systems, delayed information is intentionally used to ensure the system performance and simultaneously reduce input cost. Because of this, it is crucial to implement consensus algorithms in multi-agent systems based on the sampled-data control method. A typical example is the sampled-data system [3,14], where the sampled data are used during sampling intervals and updated at sampling instants. Whereafter, the distributed consensus optimization problem of multi-agent systems with delayed sampled-data is considered in [20].

Motivated by the aforementioned discussion, this paper concentrates on the distributed optimization problem for multi-agent systems with both communication delays and external disturbances over a directed network. Based on the internal model principle, a new kind of distributed optimization protocol is proposed, and the convergence of the protocol is analyzed by using the Lyapunov–Razumikhin theory. The key contributions of this paper are listed as follows:

1. Unlike the distributed continuous-time optimization algorithms in [21, 22] and [11, 16, 17], which only consider communication delays or external disturbances. We propose a new distributed optimal protocol for multi-agent systems with both communication delays and external disturbances under a directed network in which each agent only uses its information.
2. Compared with [15], this paper utilizes the Lyapunov–Razumikhin theory to analyze the optimization problem of multi-agent systems with communication delays, which can enable the agents to achieve consensus and avoid verifying the LMIs. In addition, the stability conditions are derived via the construct Lyapunov–Razumikhin function, which can estimate the upper bound of communication delays and can also clearly find the relationship among the parameters in the algorithm.
3. In general, the distributed optimization problem with aperiodic sampling and external disturbances can be transformed into one with communication delays. Therefore, the protocol and theoretical method proposed in this paper can be extended to the aperiodic sampling data systems.

The remaining sections are organized as follows. Some preliminaries and problem statements are given in Section 2. In Section 3, the main results are detailed. An example is given to illustrate the effectiveness of the results in Section 3. Section 5 gives a conclusion to this paper.

Notations. Let $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{R}^n$ denote the sets of natural numbers, real numbers, and real vectors of dimension n, respectively. $I_n \in \mathbb{R}^{n\times n}$ represents $n\times n$ identity matrix, $\mathbf{1}_n$(or $\mathbf{0}_n$) denotes an n-dimensional column vector whose all entries are 1 (or 0). For a matrix A, A^{T} is its transpose, $\lambda_{\min}(A)$, and $\lambda_{\max}(A)$ denote the smallest and the largest eigenvalues of the matrix, respectively. For vectors $x_1, x_2, \dots, x_n$, $\mathrm{col}(x_1, x_2, \dots, x_n) = [{x_1}^{\mathrm{T}}, {x_2}^{\mathrm{T}}, \dots, {x_n}^{\mathrm{T}}]^{\mathrm{T}}$. Let $\|\cdot\|$ and $\otimes$ represent the Euclidean norm and the Kronecker product, respectively. The gradient of f is denoted by ∇f.

2 Preliminaries

2.1 Algebraic graph theory

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ represent a weighted digraph with the finite set of nodes $\mathcal{V} = \{1, 2, \dots, N\}$, edge set $\mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$, and weighted adjacency matrix $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N\times N}$. Let

$N_i = \{j \in \mathcal{V},\ (j,i) \in \mathcal{E}\}$ be the set of neighbors of node i, and the weighted adjacency matrix $\mathcal{A} = [a_{ij}]$ is defined as $a_{ii} = 0$, $a_{ij} > 0$ if $(j,i) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. If $\sum_{j=1}^{N} a_{ij} = \sum_{j=1}^{N} a_{ji}$ for all $i \in \mathcal{V}$, then the digraph $\mathcal{G}$ is called weighted-balanced. The digraph $\mathcal{G}$ is called strongly connected if there exists a directed path between any different nodes. The Laplacian matrix of digraph $\mathcal{G}$ is $L = [l_{ij}] \in \mathbb{R}^{N\times N}$, which is defined as $l_{ii} = \sum_{j=1,\, j\neq i}^{N} a_{ij}$ and $l_{ij} = -a_{ij}$ for $i \neq j$.

Lemma 1. (*See* [7].) *Let L be the Laplacian matrix of a directed graph $\mathcal{G}$, then it has at least one zero eigenvalues, and all the nonzero eigenvalues have positive real parts. In particular, if $\mathcal{G}$ is strongly connected, L has a simple zero eigenvalue with $\mathbf{1}_N = [1, 1, \dots, 1]^{\mathrm{T}} \in \mathbb{R}^N$ as its right eigenvector.*

2.2 Definitions and lemmas

Consider the following time-delay system:

$$\begin{aligned} \dot{x}(t) &= f(t, x_t), \quad t > t_0, \\ x(\theta) &= \Pi(\theta), \quad \theta \in [-\tau, t_0], \end{aligned} \tag{1}$$

where $x_t(\theta) = x(t+\theta)$ and $f(t,0) = 0$. In the sequel, suppose that $t_0 = 0$. Let $C([-\tau,0],\mathbb{R}^n)$ be a Banach space of continuous function $\Pi : [-\tau, 0] \to \mathbb{R}^n$ with the norm $\|\Pi\| = \sup_{-\tau\leqslant\theta\leqslant 0} \|\Pi(\theta)\|$.

Lemma 2. (*See* [8].) *Suppose that the function $f : \mathbb{R} \times C([-\tau,0],\mathbb{R}^n) \to \mathbb{R}^n$ is continuous, and f maps bounded sets of $C([-\tau,0],\mathbb{R}^n)$ to bounded sets of $\mathbb{R}^n$. Let $\Pi_1(t)$, $\Pi_2(t)$, and $\Pi_3(t)$ be continuous, nonnegative, nondecreasing functions with $\Pi_1(t) > 0$, $\Pi_2(t) > 0$, $\Pi_3(t) > 0$ for $t > 0$ and $\Pi_1(0) = \Pi_2(0) = 0$. There is a continuous function $W(t,x)$ such that*

$$\Pi_1\bigl(\|x\|\bigr) \leqslant W(t,x) \leqslant \Pi_2\bigl(\|x\|\bigr), \quad t \in \mathbb{R},\ x \in \mathbb{R}^n.$$

In addition, if there exists a continuous nondecreasing function $\Pi(t)$ with $\Pi(t) > t$, $t > 0$, and

$$W\bigl(t+\theta,\ x(t+\theta)\bigr) < \Pi\bigl(W(t,x)\bigr), \quad \theta \in [-\tau, 0], \tag{2}$$

such that

$$\dot{W}(t,x)|_{(1)} \leqslant -\Pi_3(t),$$

then the solution $x = 0$ of system (1) *is uniformly asymptotically stable.*

Lemma 3. (*See* [15].) *For a positive matrix $M \in \mathbb{R}^{N\times N}$ and for all $a, b \in \mathbb{R}^N$, it has*

$$2a^{\mathrm{T}}b \leqslant a^{\mathrm{T}}M^{-1}a + b^{\mathrm{T}}Mb.$$

In particular, when $M \in \mathbb{R}^{N\times N}$ is an identity matrix, it yields $2a^{\mathrm{T}}b \leqslant a^{\mathrm{T}}a + b^{\mathrm{T}}b$.

2.3 Problem formulation

We consider a multi-agent system consisting of N agents, and the interaction topology is described by a digraph $\mathcal{G}$. The dynamic of the ith agent is described by

$$\dot{x}_i(t) = u_i(t) + d_i(t), \quad i = 1, 2, \dots, N, \tag{3}$$

where $x_i(t), u_i(t), d_i(t) \in \mathbb{R}^n$ are the state, control input, and external disturbance of agent i, respectively. Suppose that $d_i(t)$ is governed by an exosystem

$$\dot{v}_i(t) = Bv_i(t), \qquad d_i(t) = Cv_i(t), \tag{4}$$

where $v_i(t) \in \mathbb{R}^s$ is the state of the exosystem state. $B \in \mathbb{R}^{s\times s}$ and $C \in \mathbb{R}^{n\times s}$ are the constant matrices with appropriate dimensions. Suppose that all eigenvalues of $B \in \mathbb{R}^{s\times s}$ are distinct lying on the imaginary axis, which means the disturbance is bounded.

The optimization problem of multi-agent systems (3) is defined as

$$\begin{aligned} &\text{minimize} \quad F\bigl(x(t)\bigr) = \sum_{i=1}^{N} f_i(x_i(t)), \quad x_i(t) \in \mathbb{R}^n, \\ &\text{subject to} \quad (L \otimes I_n)x(t) = \mathbf{0}_{Nn}, \end{aligned} \tag{5}$$

where $x(t) = \operatorname{col}(x_1(t), x_2(t), \dots, x_N(t)) \in \mathbb{R}^{Nn}$, L is the Laplacian matrix of the communication topology $\mathcal{G}$. In this paper, our goal is to design a distributed optimization algorithm such that all agents' states reach consensus and converge to the optimal solution of the optimization problem (5) via local communication.

Assumption 1. The digraph $\mathcal{G}$ is strongly connected and weighted-balanced.

Remark 1. From Assumption 1 zero is a simple eigenvalue of matrix L, and $\mathbf{1}_N^{\mathrm{T}} L = 0$. Moreover, there exists a matrix $Q_2 \in \mathbb{R}^{N\times(N-1)}$ with

$$\mathbf{1}_N^{\mathrm{T}} Q_2 = \mathbf{0}, \qquad Q_2^{\mathrm{T}} Q_2 = I_{N-1}, \qquad Q_2 Q_2^{\mathrm{T}} = I_N - \frac{1}{N}\mathbf{1}_N \mathbf{1}_N^{\mathrm{T}}$$

such that the matrix $Q_2^{\mathrm{T}} L Q_2 = J$ in which the real parts of all eigenvalues of J are positive, and $J + J^{\mathrm{T}}$ is a positive definite matrix.

Assumption 2. For each agent $i \in \mathcal{V}$, the function f_i is differentiable, and its gradient satisfies l_i-Lipschitz ($l_i > 0$) condition in $\mathbb{R}^n$, i.e.,

$$\bigl\| f_i(x) - f_i(y) \bigr\| \leqslant l_i \|x - y\| \quad \forall x, y \in \mathbb{R}^n.$$

Assumption 3. For each agent $i \in \mathcal{V}$, f_i is m_i-strongly convex ($m_i > 0$) that means

$$(x - y)^{\mathrm{T}} \bigl(\nabla f_i(x) - \nabla f_i(y)\bigr) > m_i \|x - y\|^2 \quad \forall x, y \in \mathbb{R}^n,\ x \neq y.$$

Remark 2. Based on Assumption 3, the global optimization function $F(\cdot)$ is strongly convex due to the strong convexity of local cost function $f_i(\cdot)$. Therefore, Assumption 3 guarantees the uniqueness of the optimal solution to (5).

3 Main results

Due to the existence of disturbance $d_i(t)$, some existing optimization algorithms are not applicable. Before designing the distributed optimization algorithm, we first give the following transformation.

Let

$$p(\lambda) = \lambda^s + a_{s-1}\lambda^{s-1} + \cdots + a_1\lambda + a_0$$

be the minimal zeroing polynomial of the matrix B and $q_i(t) = (q_{i1}(t), q_{i2}(t), \ldots, q_{in}(t))^{\mathrm{T}}$ with $q_{ij}(t) = (d_{ij}(t), \mathrm{d}d_{ij}(t)/\mathrm{d}t, \ldots, \mathrm{d}^{s-1}d_{ij}(t)/\mathrm{d}t^{s-1})^{\mathrm{T}}$, $j = 1, 2, \ldots, n$. It can be seen that $q_{ij}(t)$ satisfies $\dot{q}_{ij}(t) = \Phi q_{ij}(t)$, $d_{ij}(t) = \Psi q_{ij}(t)$, where

$$\Phi = \begin{bmatrix} 0 & I_{s-1} \\ -a_0 & -a_1 - a_2 \cdots - a_{s-1} \end{bmatrix}, \qquad \Psi = [1, \underbrace{0, \ldots, 0}_{s-1}].$$

Then the disturbance $d_i(t)$ can be rewritten as

$$\dot{q}_i(t) = (I_n \otimes \Phi)q_i(t), \qquad d_i(t) = (I_n \otimes \Psi)q_i(t). \tag{6}$$

Since the pair (Φ, Ψ) is observable, so there exists a matrix P such that $H = \Phi + P\Psi$ is Hurwitz stable.

3.1 Optimization protocol with communication delays

From above transformation the following internal-model-based optimization protocol is proposed:

$$\begin{aligned} \dot{w}_i(t) &= k \sum_{j \in N_i} a_{ij}\big(x_i\big(t - \tau(t)\big) - x_j\big(t - \tau(t)\big)\big), \\ \dot{\eta}_i(t) &= (I_n \otimes H)\eta_i(t) + (I_n \otimes P)u_i(t), \\ u_i(t) &= -\alpha \nabla f_i\big(x_i(t)\big) - w_i(t) - (I_n \otimes \Psi)\eta_i(t) \\ &\quad - \beta \sum_{j \in N_i} \big(x_i\big(t - \tau(t)\big) - x_j\big(t - \tau(t)\big)\big), \end{aligned} \tag{7}$$

where k, α, β are positive parameters, $w_i(t) \in \mathbb{R}^n$ is an auxiliary state of agent i, $\tau(t) > 0$ is the communication delay.

Remark 3. A new distributed optimization algorithm $u_i(t)$ is proposed with external disturbances and communication delays. Obviously, the proposed optimization protocol consists of four partsṪhe gradient-based term $-\alpha\nabla f_i(x_i(t))$ is used to drive the agents to the optimization point. The term $w_i(t)$ is introduced to eliminate error caused by the gradient differences. The internal model term $-(I_n \otimes \Psi)\eta_i(t)$ is employed to compensate the external disturbances asymptotically, and the term $-\beta \sum_{j \in N_i}(x_i(t - \tau(t)) - x_j(t - \tau(t))$ can make all agents achieve consensus.

Substituting equations (6) and (7) into (3), the following dynamic system is given:

$$\begin{aligned}
\dot{x}_i(t) &= -\alpha\nabla f_i\big(x_i(t)\big) - w_i(t) + (I_n\otimes\Psi)\big(q_i(t)-\eta_i(t)\big)\\
&\quad - \beta\sum_{j\in N_i} a_{ij}\big(x_i\big(t-\tau(t)\big) - x_j\big(t-\tau(t)\big)\big),\\
\dot{w}_i(t) &= k\sum_{j\in N_i} a_{ij}\big(x_i\big(t-\tau(t)\big) - x_j\big(t-\tau(t)\big)\big),\\
\dot{\eta}_i(t) &= (I_n\otimes H)\eta_i(t) + (I_n\otimes P)\Big\{-\alpha\nabla f_i\big(x_i(t)\big) - w_i(t) - (I_n\otimes\Psi)\eta_i(t)\\
&\quad - \beta\sum_{j\in N_i} a_{ij}\big(x_i(t-\tau(t)\big) - x_j\big(t-\tau(t)\big)\big)\Big\}.
\end{aligned} \tag{8}$$

Let $\gamma_i(t) = \eta_i(t) - q_i(t)$, then system (8) is rewritten as follows:

$$\begin{aligned}
\dot{x}_i(t) &= -\alpha\nabla f_i\big(x_i(t)\big) - w_i(t) - (I_n\otimes\Psi)\gamma_i(t)\\
&\quad - \beta\sum_{j\in N_i} a_{ij}\big(x_i\big(t-\tau(t)\big) - x_j\big(t-\tau(t)\big)\big),\\
\dot{w}_i(t) &= k\sum_{j\in N_i} a_{ij}\big(x_i\big(t-\tau(t)\big) - x_j\big(t-\tau(t)\big)\big),\\
\dot{\gamma}_i(t) &= (I_n\otimes H)\gamma_i(t) + (I_n\otimes P)\Big\{-\alpha\nabla f_i\big(x_i(t)\big) - w_i(t) - (I_n\otimes\Psi)\gamma_i(t)\\
&\quad - \beta\sum_{j\in N_i} a_{ij}\big(x_i\big(t-\tau(t)\big) - x_j\big(t-\tau(t)\big)\big)\Big\}.
\end{aligned} \tag{9}$$

Let $w(t) = \mathrm{col}(w_1(t), w_2(t), \dots, w_N(t))$, $\gamma(t) = \mathrm{col}(\gamma_1(t), \gamma_2(t), \dots, \gamma_N(t))$, and $\nabla f(x(t)) = \mathrm{col}(\nabla f_1(x_1(t)), \nabla f_2(x_2(t)), \dots, \nabla f_N(x_N(t)))$. Based on the definition of Laplacian matrix L, system (9) can be expressed in the following compact form:

$$\begin{aligned}
\dot{x}(t) &= -\alpha\nabla f\big(x(t)\big) - w(t) - (I_{Nn}\otimes\Psi)\gamma(t) - \beta(L\otimes I_n)x\big(t-\tau(t)\big),\\
\dot{w}(t) &= k(L\otimes I_n)x\big(t-\tau(t)\big),\\
\dot{\gamma}(t) &= (I_{Nn}\otimes H)\gamma(t) + (I_{Nn}\otimes P)\big\{-\alpha\nabla f\big(x(t)\big) - w(t)\\
&\quad - (I_{Nn}\otimes\Psi)\gamma(t) - \beta(L\otimes I_n)x\big(t-\tau(t)\big)\big\}.
\end{aligned} \tag{10}$$

The following theorem indicates the relationship between the equilibrium point of system (10) and the optimal solution of optimization problem (5).

Theorem 1. *Suppose that Assumptions* 1–3 *are satisfied, then the following two statements hold*:

(i) $\Omega(\epsilon) = \{(x, w, \gamma) | (\mathbf{1}_N\otimes I_n)^{\mathrm{T}} w = \epsilon\}$ *is a positive invariant set for any* $\epsilon\in\mathbb{R}^n$.
(ii) *If* $(x^*, -\alpha\nabla\tilde{f}(x^*), \gamma^*)\in\Omega(\mathbf{0}_n)$ *is an equilibrium point of system* (10)*, then* x^* *is an optimal solution of* (5).

Proof. (i) Since the directed graph $\mathcal{G}$ is weighted-balanced, we have

$$(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}(L \otimes I_n) = (\mathbf{0}_N \otimes I_n)^{\mathrm{T}}.$$

Therefore,

$$(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}\dot{w}(t) = k(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}(L \otimes I_n)x\bigl(t - \tau(t)\bigr) \equiv \mathbf{0}_n,$$

which implies $(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}w(t)$ is a constant. That means the set $\Omega(\epsilon)$ is invariant for any ϵ. Furthermore,

$$(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}w(t) = (\mathbf{1}_N \otimes I_n)^{\mathrm{T}}w(0),$$

which implies $\epsilon = (\mathbf{1}_N \otimes I_n)^{\mathrm{T}}w(0)$.

(ii) Let $(x^*, -\alpha\nabla\tilde{f}(x^*), \gamma^*) \in \Omega(\mathbf{0}_n)$ be the equilibrium point of system (10), then we have

$$\begin{aligned}
&-\alpha\nabla f(x^*) - w^* - (I_{Nn} \otimes \Psi)\gamma^* - \beta(L \otimes I_n)x^* = 0,\\
&k(L \otimes I_n)x^* = 0,\\
&(I_{Nn} \otimes H)\gamma^* + (I_{Nn} \otimes P)\bigl\{-\alpha\nabla f(x^*)\\
&\quad - w^* - (I_{Nn} \otimes \Psi)\gamma^* - \beta(L \otimes I_n)x^*\bigr\} = 0.
\end{aligned} \tag{11}$$

From the first equation and the third equation in (11) it yields $(I_{Nn} \otimes H)\gamma^* = 0$, which implies $\gamma^* = 0$. According to Lemma 1, one has

$$(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}(L \otimes I_n) = \mathbf{1}_N^{\mathrm{T}}L \otimes I_n = 0.$$

Since the directed graph $\mathcal{G}$ is strongly connected, it has $x^* = \mathbf{1}_N \otimes z$, where $z \in \mathbb{R}^n$ is a constant vector. Based on $\gamma^* = 0$ and $(x^*, -\alpha\nabla f(x^*), \gamma^*) \in \Omega(\mathbf{0}_n)$, so multiplying from the left by $(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}$ the first equation in (11), one can obtain $\alpha(\mathbf{1}_N \otimes I_n)^{\mathrm{T}} \times \nabla f(z) = 0$. Thus, the optimal condition $\nabla F(z) = 0$ is satisfied, which implies that x^* is the optional solution of (5). Substituting x^* into the first equation in (11), one obtains $w^* = -\alpha\nabla f(x^*)$. □

Remark 4. We can see that the optimal solution of (5) is located in the invariant set $\Omega(\mathbf{0}_n)$ from Theorem 1, so the initial value $w(0) \in \Omega(\mathbf{0}_n)$ is a necessary condition for solving the optimization problem. If the initial value is not in $\Omega(\mathbf{0}_n)$, i.e., $(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}w(0) \neq \mathbf{0}_n$, then $\alpha(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}\nabla f(z) \neq \mathbf{0}_n$, which implies that the optimal solution cannot be obtained.

Theorem 2. *Suppose that Assumptions* 1–3 *hold, then all states of agents in the multi-agent system* (3) *can reach consensus and converge to the optimal solution of the optimization problem* (5) *under the protocol* (7) *for any initial values* $x(0)$, $w(0)$, $\gamma(0)$ *satisfying* $(\mathbf{1}_N \otimes I_n)^{\mathrm{T}}w(0) = \mathbf{0}$ *and* $\tau(t) \in [0, \bar{\tau})$ *if the following conditions are satisfied*:

$$\begin{gathered}
4\sigma\alpha\check{m} - a_2^2\sigma - 4\alpha^2\hat{l}^2 - 4\sigma - 8a_2^2 - a_3a_4 - \underline{\lambda}' \geqslant 0,\\
\beta > \frac{2}{\underline{\lambda}} + k,
\end{gathered} \tag{12}$$

$$\bar{\tau} < \frac{\underline{\lambda}'}{2\sigma(\beta - k)[(\beta + k)\bar{\lambda} + \beta^3\bar{\lambda}'] + p\zeta(\mu + 1)}, \tag{13}$$

where $\sigma > 0$, $\zeta = 3\sigma + a_3\lambda + (4\sigma\delta + 2\delta a_3\lambda)(\alpha^2\hat{l}^2 + a_1^2 + a_2^2)$, $\mu = \lambda_{\max}(D_1)$, $\lambda = \lambda_{\max}(U)$, $\underline{\lambda} = \lambda_{\min}(J^{\mathrm{T}} + J)$, $\bar{\lambda} = \lambda_{\max}(JJ^{\mathrm{T}})$, $\underline{\lambda}' = \lambda_{\min}(\Lambda - 2I_{2N-2})$, $\bar{\lambda}' = \lambda_{\max}(J^2(J^2)^{\mathrm{T}})$, *and*

$$\Lambda = \begin{pmatrix} 2(\beta - k)(J^{\mathrm{T}} + J) & 2I_{N-1} \\ 2I_{N-1} & 4I_{N-1} \end{pmatrix} \otimes I_n.$$

Proof. Let $\bar{x}(t) = x(t) - x^*$, $\bar{w}(t) = w(t) - w^*$, and $\bar{\gamma}(t) = \gamma(t) - (I_{Nn} \otimes P)\bar{x}(t)$, then system (10) can be transformed into the following form:

$$\begin{aligned} \dot{\bar{x}}(t) &= -\alpha h\big(\bar{x}(t)\big) - \bar{w}(t) - (I_{Nn} \otimes \Psi)\big(\bar{\gamma}(t) + (I_{Nn} \otimes P)\bar{x}(t)\big) \\ &\quad - \beta(L \otimes I_n)\bar{x}\big(t - \tau(t)\big), \\ \dot{\bar{w}}(t) &= k(L \otimes I_n)\bar{x}\big(t - \tau(t)\big), \\ \dot{\bar{\gamma}}(t) &= (I_{Nn} \otimes H)\bar{\gamma}(t) + (I_{Nn} \otimes HP)\bar{x}(t), \end{aligned}$$

where $h(\bar{x}(t)) = \nabla f(x(t) + x^*) - \nabla f(x^*)$.

Let $e(t) = (Q^{\mathrm{T}} \otimes I_n)\bar{x}(t)$, $\xi(t) = (Q^{\mathrm{T}} \otimes I_n)\bar{w}(t)$, $Q = [\mathbf{1}_N/\sqrt{N}, Q_2]$. Denote $e(t) = \mathrm{col}(e_1(t), e_2(t))$ with $e_1^{\mathrm{T}}(t) \in \mathbb{R}^n$, $e_2^{\mathrm{T}}(t) \in \mathbb{R}^{(N-1)n}$, $\xi(t) = \mathrm{col}(\xi_1(t), \xi_2(t))$ with $\xi_1^{\mathrm{T}}(t) \in \mathbb{R}^n$ and $\xi_2^{\mathrm{T}}(t) \in \mathbb{R}^{(N-1)n}$. By the structure of Q and Remark 1 the above system can be transformed as follows:

$$\begin{aligned} \dot{e}_1(t) &= -\alpha\bigg(\frac{\mathbf{1}_N^{\mathrm{T}}}{\sqrt{N}} \otimes I_n\bigg)h\big(\bar{x}(t)\big) - \bigg(\frac{\mathbf{1}_N^{\mathrm{T}}}{\sqrt{N}} \otimes I_n\bigg)\Gamma(t), \\ \dot{e}_2(t) &= -\alpha\big(Q_2^{\mathrm{T}} \otimes I_n\big)h\big(\bar{x}(t)\big) - \xi_2(t) - \beta(J \otimes I_n)e_2\big(t - \tau(t)\big) \\ &\quad - \big(Q_2^{\mathrm{T}} \otimes I_n\big)\Gamma(t), \\ \dot{\xi}_1(t) &= 0, \qquad \dot{\xi}_2(t) = k(J \otimes I_n)e_2\big(t - \tau(t)\big), \\ \dot{\bar{\gamma}}(t) &= (I_{Nn} \otimes H)\bar{\gamma}(t) + (I_{Nn} \otimes HP)\bar{x}(t), \end{aligned} \tag{14}$$

where $\Gamma(t) = (I_{Nn} \otimes \Psi)(\bar{\gamma}(t) + (I_{Nn} \otimes P)(Q^{\mathrm{T}} \otimes I_n)\bar{x}(t))$. According to the properties of matrix norm, there is $\|\Gamma(t)\| \leqslant a_1\|\bar{\gamma}(t)\| + a_2\|\bar{x}(t)\|$ with $a_1 = \|\Psi\|$ and $a_2 = \|\Psi P\|$.

There exists a positive definite matrix U such that $UH^{\mathrm{T}} + H^{\mathrm{T}}U + 2I_s \leqslant 0$ due to H is Hurwitz stable. Let $\varepsilon(t) = \mathrm{col}(e(t), \xi(t), \bar{\gamma}(t)) = \mathrm{col}(e_1(t), e_2(t), \xi_1(t), \xi_2(t), \bar{\gamma}(t))$. We construct the Lyapunov–Razumikhin function as

$$V\big(\varepsilon(t)\big) = \varepsilon^{\mathrm{T}}(t)(D \otimes I_n)\varepsilon(t)$$

with

$$D = \begin{pmatrix} 2\sigma & 0 & 2 & 0 & 0 \\ 0 & 2\sigma I_{N-1} & 0 & 2I_{N-1} & 0 \\ 2 & 0 & \frac{2\beta}{k} & 0 & 0 \\ 0 & 2I_{N-1} & 0 & \frac{2\beta}{k}I_{N-1} & 0 \\ 0 & 0 & 0 & 0 & a_3(I_N \otimes U) \end{pmatrix},$$

where $\sigma > 1$, and there is a fact that D is positive definite since $\beta > k$.

The derivation of $V(\varepsilon(t))$ along with system (14) is given by

$$\begin{aligned}\dot{V}\big(\varepsilon(t)\big) &= 2\varepsilon^{\mathrm{T}}(t)(D\otimes I_n)\dot{\varepsilon}(t)\\ &= 4\big(\sigma e_1(t)+\xi_1(t)\big)^{\mathrm{T}}\dot{e}_1(t) + 4\big(\sigma e_2(t)+\xi_2(t)\big)^{\mathrm{T}}\dot{e}_2(t)\\ &\quad + 4\bigg(e_1(t)+\frac{\beta}{k}\xi_1(t)\bigg)^{\mathrm{T}}\dot{\xi}_1(t) + 4\bigg(e_2(t)+\frac{\beta}{k}\xi_2(t)\bigg)^{\mathrm{T}}\dot{\xi}_2(t)\\ &\quad + a_3\bar{\gamma}(t)\big((I_{Nn}\otimes U)+(I_{Nn}\otimes U^{\mathrm{T}})\big)\dot{\bar{\gamma}}(t).\end{aligned}$$

Due to $\dot{\xi}_1(t)=0$ and $w(0)\in\Omega(0)$, one has $\xi_1(t)=0$ for all $t\geqslant 0$, then

$$\begin{aligned}\dot{V}\big(\varepsilon(t)\big) &= a_3\bar{\gamma}(t)\big((I_{Nn}\otimes U)+(I_{Nn}\otimes U^{\mathrm{T}})\big)\dot{\bar{\gamma}}(t)\\ &\quad + 4\sigma e_1^{\mathrm{T}}(t)\dot{e}_1(t) + 2\varepsilon_2^{\mathrm{T}}(t)D_1\dot{\varepsilon}_2(t),\end{aligned} \tag{15}$$

where $\varepsilon_2(t)=\mathrm{col}(e_2(t),\xi_2(t))$,

$$D_1 = \begin{pmatrix} 2\sigma I_{N-1} & 2I_{N-1}\\ 2I_{N-1} & (2\beta/k)I_{N-1}\end{pmatrix}\otimes I_n.$$

From the second and fourth equations of system (14) we have

$$\dot{\varepsilon}_2(t) = E\varepsilon_2(t) + F\varepsilon_2\big(t-\tau(t)\big) + O(t), \tag{16}$$

where

$$O(t) = \begin{pmatrix} -\alpha(Q_2^{\mathrm{T}}\otimes I_n)h(\bar{x}(t)) - (Q_2^{\mathrm{T}}\otimes I_n)\Gamma(t)\\ 0\end{pmatrix},$$

$$E = \begin{pmatrix} 0 & -I_{N-1}\\ 0 & 0\end{pmatrix}\otimes I_n, \quad \text{and} \quad F = \begin{pmatrix} -\beta J & 0\\ kJ & 0\end{pmatrix}\otimes I_n.$$

Utilizing the Leibniz–Newton formula, it follows that

$$\begin{aligned}\varepsilon_2\big(t-\tau(t)\big) &= \varepsilon_2(t) - \int_{t-\tau(t)}^{t}\dot{\varepsilon}_2(s)\,\mathrm{d}s\\ &= \varepsilon_2(t) - E\int_{-\tau(t)}^{0}\varepsilon_2(t+s)\,\mathrm{d}s - F\int_{-2\tau(t)}^{-\tau(t)}\varepsilon(t+s)\,\mathrm{d}s\\ &\quad - \int_{-\tau(t)}^{0}O(t+s)\,\mathrm{d}s.\end{aligned}$$

Therefore, equation (16) can be rewritten as

$$\dot{\varepsilon}_2(t) = \bar{O}\varepsilon_2(t) - FE \int\limits_{-\tau(t)}^{0} \varepsilon_2(t+s)\,\mathrm{d}s - F^2 \int\limits_{-2\tau(t)}^{-\tau(t)} \varepsilon_2(t+s)\,\mathrm{d}s - F \int\limits_{-\tau(t)}^{0} O(t+s)\,\mathrm{d}s + O(t),$$

where $\bar{O} = E + F$.

Furthermore, one has

$$\begin{aligned} 2\varepsilon_2^{\mathrm{T}}(t)D_1\dot{\varepsilon}_2(t) &= 2\varepsilon_2^{\mathrm{T}}(t)D_1\bar{O}\varepsilon_2(t) - 2\varepsilon_2^{\mathrm{T}}(t)D_1FE \int\limits_{-\tau(t)}^{0} \varepsilon_2(t+s)\,\mathrm{d}s \\ &\quad - 2\varepsilon_2^{\mathrm{T}}(t)D_1F^2 \int\limits_{-2\tau(t)}^{-\tau(t)} \varepsilon_2(t+s)\,\mathrm{d}s + 2\varepsilon_2^{\mathrm{T}}(t)D_1O(t) \\ &\quad - 2\varepsilon_2(t)D_1F \int\limits_{-\tau(t)}^{0} O(t+s)\,\mathrm{d}s. \end{aligned} \tag{17}$$

Combined with (16) and (17), equation (15) is written as

$$\begin{aligned} \dot{V}(\varepsilon(t)) &= 2\varepsilon_2^{\mathrm{T}}(t)D_1\bar{O}\varepsilon_2(t) - 2\varepsilon_2^{\mathrm{T}}(t)D_1FE \int\limits_{-\tau(t)}^{0} \varepsilon_2(t+s)\,\mathrm{d}s \\ &\quad - 2\varepsilon_2^{\mathrm{T}}(t)D_1F^2 \int\limits_{-2\tau(t)}^{-\tau(t)} \varepsilon_2(t+s)\,\mathrm{d}s - 2\varepsilon_2(t)D_1F \int\limits_{-\tau(t)}^{0} O(t+s)\,\mathrm{d}s \\ &\quad + 2\varepsilon_2^{\mathrm{T}}(t)D_1O(t) + 4\sigma e_1^{\mathrm{T}}(t)\dot{e}_1(t) \\ &\quad + a_3\bar{\gamma}(t)\big((I_{Nn} \otimes U) + (I_{Nn} \otimes U^{\mathrm{T}})\big)\dot{\bar{\gamma}}(t). \end{aligned}$$

By Lemma 3 one can obtain

$$\begin{aligned} &-2\varepsilon_2^{\mathrm{T}}(t)D_1EF \int\limits_{-\tau(t)}^{0} \varepsilon_2(t+s)\,\mathrm{d}s \\ &\quad \leqslant \bar{\tau}\varepsilon_2^{\mathrm{T}}(t)D_1EFD_1^{-1}(D_1EF)^{\mathrm{T}}\varepsilon_2(t) + \int\limits_{-\tau(t)}^{0} \varepsilon_2^{\mathrm{T}}(t+s)D_1\varepsilon_2(t+s)\,\mathrm{d}s. \end{aligned} \tag{18}$$

Similarly, we have

$$-2\varepsilon_2^{\mathrm{T}}(t)D_1F^2\int_{-2\tau(t)}^{-\tau(t)}\varepsilon_2(t+s)\,\mathrm{d}s \leqslant \bar{\tau}\varepsilon_2^{\mathrm{T}}(t)D_1F^2D_1^{-1}(D_1F^2)^{\mathrm{T}}\varepsilon_2(t)+\int_{-2\tau(t)}^{-\tau(t)}\varepsilon_2^{\mathrm{T}}(t+s)D_1\varepsilon_2(t+s)\,\mathrm{d}s$$

and

$$-2\varepsilon_2^{\mathrm{T}}(t)D_1F\int_{-\tau(t)}^{0}O(t+s)\,\mathrm{d}s \leqslant \bar{\tau}\varepsilon_2^{\mathrm{T}}(t)D_1FD_1^{-1}(D_1F)^{\mathrm{T}}\varepsilon_2(t)+\int_{-\tau(t)}^{0}O^{\mathrm{T}}(t+s)D_1O(t+s)\,\mathrm{d}s. \tag{19}$$

Since

$$O(t)=\begin{pmatrix}-\alpha(Q_2^{\mathrm{T}}\otimes I_n)h(\bar{x}(t))-(Q_2^{\mathrm{T}}\otimes I_n)\Gamma(t)\\0\end{pmatrix},$$

then

$$\begin{aligned}\|O(t)\|^2&=\|-\alpha(Q_2^{\mathrm{T}}\otimes I_n)h(\bar{x}(t))-(Q_2^{\mathrm{T}}\otimes I_n)\Gamma(t)\|^2\\&\leqslant 2\alpha^2\|(Q_2^{\mathrm{T}}\otimes I_n)h(\bar{x}(t))\|^2+2(Q_2^{\mathrm{T}}\otimes I_n)\Gamma(t)\|^2\\&\leqslant 2(\alpha^2\hat{l}^2+a_1^2+a_2^2)\|\varepsilon(t)\|^2,\end{aligned}$$

where $\hat{l}=\max\{l_1,l_2,\dots,l_N\}$. From the transformation $e(t)=(Q^{\mathrm{T}}\otimes I_n)\bar{x}(t)$ it follows that

$$\begin{aligned}4\sigma e_1^{\mathrm{T}}(t)\dot{e}_1(t)&=-4\sigma e_1^{\mathrm{T}}(t)\left[\alpha\left(\frac{1_N^{\mathrm{T}}}{\sqrt{N}}\otimes I_n\right)h(\bar{x}(t))+\left(\frac{1_N^{\mathrm{T}}}{\sqrt{N}}\otimes I_n\right)\Gamma(t)\right]\\&=-4\sigma\alpha\bar{x}^{\mathrm{T}}(t)h(\bar{x}(t))+4\sigma\alpha e_2^{\mathrm{T}}(t)(Q_2\otimes I_n)h(\bar{x}(t))\\&\quad-4\sigma\bar{x}(t)^{\mathrm{T}}\Gamma(t)+4\sigma e_2^{\mathrm{T}}(t)(Q_2\otimes I_n)\Gamma(t).\end{aligned}$$

Therefore, according to Assumption 1, one has

$$\begin{aligned}&2\varepsilon_2^{\mathrm{T}}(t)D_1O(t)+4\sigma e_1^{\mathrm{T}}(t)\dot{e}_1(t)\\&\quad=-4\alpha\xi_2^{\mathrm{T}}(t)(Q_2\otimes I_n)h(\bar{x}(t))-4\sigma\alpha\bar{x}^{\mathrm{T}}(t)h(\bar{x}(t))-4\sigma\bar{x}^{\mathrm{T}}(t)\Gamma(t)\\&\qquad-4\xi_2^{\mathrm{T}}(t)(Q_2\otimes I_n)\Gamma(t)\\&\quad\leqslant\|\xi_2(t)\|^2+4\alpha^2\hat{l}^2\|\bar{x}(t)\|^2-4\sigma\alpha\check{m}\|\bar{x}(t)\|^2+a_1^2\sigma\|\bar{\gamma}(t)\|^2\\&\qquad+4\sigma\|\bar{x}(t)\|^2+a_2^2\sigma\|\bar{x}(t)\|^2+\|\xi_2(t)\|^2+8a_1^2\|\bar{\gamma}(t)\|^2+8a_2^2\|\bar{x}(t)\|^2\end{aligned}$$

$$\leqslant 2\varepsilon_2^{\mathrm{T}}(t)\varepsilon_2(t) + \left(8a_1^2 + a_1^2\sigma\right)\left\|\bar{\gamma}(t)\right\|^2 - \left(4\sigma\alpha\check{m} - a_2^2\sigma - 4\alpha^2\hat{l}^2 - 4\sigma - 8a_2^2\right)\left\|\bar{x}(t)\right\|^2,$$

where $\check{m} = \min\{m_1, m_2, \dots, m_N\}$.

Based on Lemma 2, take $\Pi(t) = pt$ for some constant $p > 1$. In case if

$$V\big(\varepsilon(t+\theta)\big) < pV\big(\varepsilon(t)\big), \quad \theta \in [-2\tau, 0],$$

then

$$\varepsilon_2^{\mathrm{T}}(t+s)D_1\varepsilon_2(t+s) \leqslant \varepsilon^{\mathrm{T}}(t+s)D\varepsilon(t+s) < p\varepsilon^{\mathrm{T}}(t)D\varepsilon(t). \tag{20}$$

Next, considering the integral term in (19), we can obtain

$$\begin{aligned}
&\int_{-\tau(t)}^{0} O^{\mathrm{T}}(t+s)D_1O(t+s)\,\mathrm{d}s \\
&\quad\leqslant 2\lambda_{\max}(D_1)(\alpha^2\hat{l}^2 + a_1^2 + a_2^2)\int_{-\tau(t)}^{0} \varepsilon^{\mathrm{T}}(t+s)\varepsilon(t+s)\,\mathrm{d}s \\
&\quad\leqslant 2\delta\big(\alpha^2\hat{l}^2 + a_1^2 + a_2^2\big)p\bar{\tau}\big(2\sigma e_1^{\mathrm{T}}(t)e_1(t) + \varepsilon_2^{\mathrm{T}}(t)D_1\varepsilon_2(t) + a_3\bar{\gamma}^{\mathrm{T}}(t)(I_{Nn}\otimes U)\bar{\gamma}(t)\big) \\
&\quad\leqslant 4\sigma\delta\big(\alpha^2\hat{l}^2 + a_1^2 + a_2^2\big)p\bar{\tau}\varepsilon^{\mathrm{T}}(t)\varepsilon(t) + 2\delta\big(\alpha^2\hat{l}^2 + a_1^2 + a_2^2\big)p\bar{\tau}\varepsilon_2^{\mathrm{T}}(t)D_1\varepsilon_2(t) \\
&\qquad + 2\delta a_3(\alpha^2\hat{l}^2 + a_1^2 + a_2^2)p\bar{\tau}\bar{\gamma}^{\mathrm{T}}(t)(I_{Nn}\otimes U)\bar{\gamma}(t),
\end{aligned}$$

where $\delta = \lambda_{\max}(D_1)/\lambda_{\min}(D)$.

In addition, we substitute (20) into the integral term (18), and it can be obtained that

$$\int_{-\tau(t)}^{0} \varepsilon_2^{\mathrm{T}}(t+s)D_1\varepsilon_2(t+s)\,\mathrm{d}s < 2\sigma p\bar{\tau}\varepsilon^{\mathrm{T}}(t)\varepsilon(t) + p\bar{\tau}\varepsilon_2^{\mathrm{T}}(t)D_1\varepsilon_2(t) + a_3p\bar{\tau}\bar{\gamma}^{\mathrm{T}}(t)(I_{Nn}\otimes U)\bar{\gamma}(t).$$

Similarly,

$$\int_{-2\tau(t)}^{-\tau(t)} \varepsilon_2^{\mathrm{T}}(t+s)D_1\varepsilon_2(t+s) < \sigma p\bar{\tau}\varepsilon^{\mathrm{T}}(t)\varepsilon(t) + p\bar{\tau}\varepsilon_2^{\mathrm{T}}(t)D_1\varepsilon_2(t) + a_3p\bar{\tau}\bar{\gamma}^{\mathrm{T}}(t)(I_{Nn}\otimes U)\bar{\gamma}(t).$$

Note that H is Hurwitz stable, and then there is

$$\begin{aligned}
\dot{V}\big(\varepsilon(t)\big) \leqslant{}& -\varepsilon_2^{\mathrm{T}}(t)(\Lambda - 2I_{2N-2})\varepsilon_2(t) \\
& + \bar{\tau}\varepsilon_2^{\mathrm{T}}(t)\big[D_1EFD_1^{-1}(D_1EF)^{\mathrm{T}} + D_1F^2D_1^{-1}(D_1F^2)^{\mathrm{T}} \\
& + D_1FD_1^{-1}(D_1F)^{\mathrm{T}} + 2\big(p + \delta p\alpha^2\hat{l}^2 + \delta pa_1^2 + \delta pa_2^2\big)D_1\big]\varepsilon_2(t)
\end{aligned}$$

$$+ p\bar{\tau}\big[(3\sigma + 2a_3\lambda) + (4\sigma\delta + 2\delta a_3\lambda)\big(\alpha^2\hat{l}^2 + a_1^2 + a_2^2\big)\big]\varepsilon^{\mathrm{T}}(t)\varepsilon(t)$$
$$- (\underline{\lambda}' + 1)\bar{\gamma}^{\mathrm{T}}(t)\bar{\gamma}(t) - \big(4\sigma\alpha\check{m} - a_2^2\sigma - 4\alpha^2\hat{l}^2 - 4\sigma - 8a_2^2 - a_3a_4\big)\big\|\bar{x}(t)\big\|^2,$$

where $a_4 = (8 + \sigma)a_1^2 + \underline{\lambda}' + 1$ and

$$D_1EFD_1^{-1}(D_1EF)^{\mathrm{T}} + D_1F^2D_1^{-1}\big(D_1F^2\big)^{\mathrm{T}} + D_1FD_1^{-1}(D_1F)^{\mathrm{T}}$$
$$= \begin{pmatrix} 2\sigma(\beta - k)[(\beta + k)JJ^{\mathrm{T}} + \beta^3J^2(J^2)^{\mathrm{T}}] & 0 \\ 0 & 0 \end{pmatrix} \otimes I_n.$$

If β satisfies condition (12), $\Lambda - 2I_{2N-2}$ is positive definite, one has

$$\varepsilon_2^{\mathrm{T}}(t)(\Lambda - 2I_{2N-2})\varepsilon_2(t) \geqslant \underline{\lambda}'\varepsilon_2^{\mathrm{T}}(t)\varepsilon_2(t)$$
$$\geqslant \underline{\lambda}'\varepsilon^{\mathrm{T}}(t)\varepsilon(t) - \underline{\lambda}'\big\|\bar{x}(t)\big\|^2 - \underline{\lambda}'\bar{\gamma}^{\mathrm{T}}(t)\bar{\gamma}(t),$$

and then

$$\dot{V}\big(\varepsilon(t)\big) \leqslant -\underline{\lambda}'\varepsilon^{\mathrm{T}}(t)\varepsilon(t) + \bar{\tau}\varepsilon_2^{\mathrm{T}}(t)\big\{2\sigma(\beta - k)\big[(\beta + k)\bar{\lambda} + \beta^3\bar{\lambda}'\big]$$
$$+ 2\big(p + \delta p\alpha^2\hat{l}^2 + \delta pa_1^2 + \delta pa_2^2\big)\mu\big\}\varepsilon_2(t)$$
$$+ p\bar{\tau}\big[(3\sigma + a_3\lambda) + (4\sigma\delta + 2\delta a_3\lambda)(\alpha^2\hat{l}^2 + a_1^2 + a_2^2)\big]\varepsilon^{\mathrm{T}}(t)\varepsilon(t)$$
$$- \big\|\bar{\gamma}(t)\big\|^2 - \big(4\sigma\alpha\check{m} - a_2^2\sigma - 4\alpha^2\hat{l}^2 - 4\sigma - 8a_2^2 - a_3a_4 - \underline{\lambda}'\big)\big\|\bar{x}(t)\big\|^2$$
$$\leqslant -\underline{\lambda}'\varepsilon^{\mathrm{T}}(t)\varepsilon(t) + \bar{\tau}\varepsilon^{\mathrm{T}}(t)\big\{2\sigma(\beta - k)\big[(\beta + k)\bar{\lambda} + \beta^3\bar{\lambda}'\big]$$
$$+ p\big[(3\sigma + a_3\lambda) + (4\sigma\delta + 2\delta a_3\lambda)\big(\alpha^2\hat{l}^2 + a_1^2 + a_2^2\big)\big](\mu + 1)\big\}\varepsilon^{\mathrm{T}}(t).$$

Since $\bar{\tau}$ satisfies condition (13), then $\dot{V}(\varepsilon(t)) \leqslant 0$. By the Lyapunov–Razumikhin theory we have $\lim_{t\to\infty} \varepsilon(t) = 0$, i.e.,

$$\lim_{t\to\infty} \bar{x}(t) = \lim_{t\to\infty} \bar{w}(t) = \lim_{t\to\infty} \bar{\gamma}(t) = 0.$$

Hence

$$\lim_{t\to\infty} x(t) = x^*, \quad \lim_{t\to\infty} w(t) = w^*,$$

which means that the optimization problem (5) can be solved by system (3) under the distributed optional algorithm (7). □

3.2 Optimization protocol with sampled-data communication

In this section, we consider the distributed optimization problem of aperiodic sampling control systems with external disturbances, and a sampled-data communication scheme is formulated.

Given a strictly increasing time sequence $\{t_k\}$, $k \in \mathbb{N}$, such that $\lim_{k\to\infty} t_k = \infty$. The state $x_i(t)$ of system (2) is assumed to be sampled at time instants t_k and available at

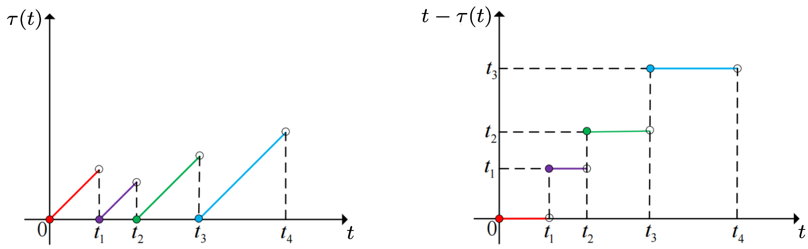

Figure 1. The graphic of $\tau(t)$ and corresponding graphic of $t-\tau(t)$.

$t_k+\tau(t)$ and $\tau(t)\geqslant 0$, that is, the sampled-data $x_i(t_k)$ is available with a time-varying delay $\tau(t)$. The sampling interval $[t_k, t_{k+1})$ satisfies

$$0 < T_{\min} \leqslant t_{k+1}-t_k = T_{k+1} \leqslant T_{\max} \quad \forall k\in\mathbb{N},$$

where T_k is the length of the kth sampling interval, $T_{\min}=\min\{T_k\}$ and $T_{\max}=\max\{T_k\}$. We use the following sampled-data-based communication scheme to solve the optimization problem (5):

$$\begin{aligned}
\dot{w}_i(t) &= k\sum_{j\in N_i} a_{ij}\big(x_i(t_k)-x_j(t_k)\big),\\
\dot{\eta}_i(t) &= (I_n\otimes H)\eta_i(t)+(I_n\otimes P)u_i(t),\\
u_i(t) &= -\alpha\nabla f_i\big(x_i(t)\big)-w_i(t)-(I_n\otimes\Psi)\eta_i(t)-\beta\sum_{j\in N_i}\big(x_i(t_k)-x_j(t_k)\big).
\end{aligned}\tag{21}$$

Suppose that $\bar{\tau}$ is the upper bound of $\tau(t)$, that is, $\tau(t)\leqslant\bar{\tau}$ and $\bar{\tau}\leqslant T_{\min}$. When sampled at time instants t_k be set as

$$t_k = t-\tau(t) \quad \forall t\in[t_k, t_{k+1}),\tag{22}$$

it is obvious from Fig. 1 that the $\tau(t)$ and $t-\tau(t)$ are piecewise and discontinuous. In addition, the sampled-data-based communication algorithms (21) can be transformed into communication protocol (7) with time delay input by utilizing (22). Therefore, the theoretical analysis of distributed optimization problems for aperiodic sampling control systems with external disturbances can be obtained as follows by Theorem 2.

Corollary 1. *Suppose that Assumptions* 1–3 *hold, then all states of agents in a multi-agent system* (3) *can reach consensus and converge to the optimal solution of the optimization problem* (5) *under protocol* (21) *for any initial values* $x(0)$, $w(0)$, $\gamma(0)$ *satisfying* $(\mathbf{1}_N\otimes I_n)^{\mathrm{T}}w(0)=\mathbf{0}$ *and* $\tau(t)\in[0,\bar{\tau})$ *if the conditions of Theorem* 2 *hold.*

Remark 5. The distributed optimization protocol with time-delay input proposed in this paper can be used in both time-delay and aperiodic sampling control systems, giving it a broad application background. Furthermore, the sampled-data communication technique requires all agents to communicate only during sampling instants, which can significantly reduce the amount of communication data.

4 Numerical example

Considers a multi-agent system consisting of five agents in which directed topology with all weights being $0-1$ is shown in Fig. 2.

The optimization problem is given by

$$\text{minimize} \quad F\big(x(t)\big) = \sum_{i=1}^{10} f_i\big(x_i(t)\big), \quad x_i(t) \in \mathbb{R},$$

where the local cost function is designed as follows:

$$f_1(x) = \frac{3}{2}(x-1)^2 + 10, \quad f_2(x) = (x-2)^2 + 11, \quad f_3(x) = \frac{1}{2}x^2 - 8x,$$
$$f_4(x) = x^2 - 7x + 17, \quad f_5(x) = 0.6x^2 - 2.4x + 9, \quad f_6(x) = \frac{3}{2}(x-1)^2 + 1,$$
$$f_7(x) = (x-2)^2, \quad f_8(x) = \frac{1}{2}x^2 + 3x, \quad f_9(x) = x^2 - 7x + 9,$$
$$f_{10}(x) = 0.6x^2 - 8x + 12.$$

Obviously, the local cost function $f_i(x)$ $(i = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10)$ satisfies Assumptions 2 and 3 in which $\hat{l} = 2$ and $\check{m} = 1$. The diagrams of local objective functions and global objective functions are shown in Figs. 3–4, respectively.

The disturbances are expressed by $d_i(t) = A_i \cos(\omega_i t + z_i)$ in agent dynamics, which can be generated by system (4) with

$$B = \begin{pmatrix} 0 & \varsigma \\ -\varsigma & 0 \end{pmatrix}, \qquad C = (1,\, 0), \qquad \omega_i(0) = \begin{pmatrix} A_i \cos z_i \\ A_i \sin z_i \end{pmatrix}.$$

Supposing that $\varsigma = 1$, $(A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}) = (\sqrt{2}/2, \sqrt{2}, 2\sqrt{2}, 2\sqrt{2}, 3\sqrt{3}, \sqrt{2}/2, \sqrt{2}, 2\sqrt{2}, 2\sqrt{2}, 3\sqrt{3})$ and $z_i = \pi/4$ $(i = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10)$, then

$$\Phi = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \Psi = \big(1,\, 0\big).$$

Furthermore, we select $P = (-1/8,\, -1/16)$ such that the matrix $H = \Phi + P\Psi$ is Hurwitz stable.

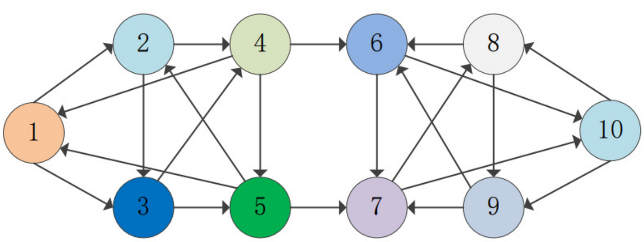

Figure 2. The directed topology.

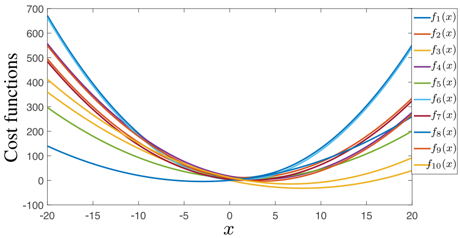

Figure 3. The diagrams of local functions $f_i(x)$.

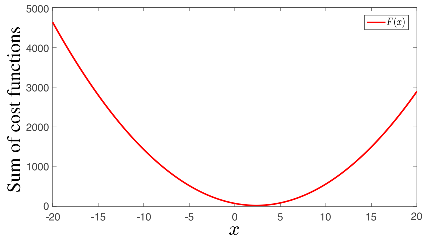

Figure 4. The diagrams of global function $F(x)$.

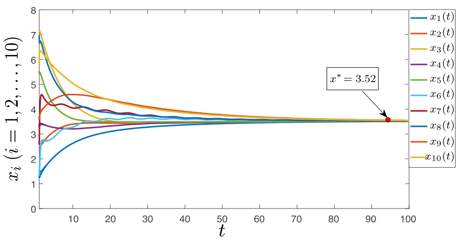

Figure 5. State trajectories of $x_i(t)$.

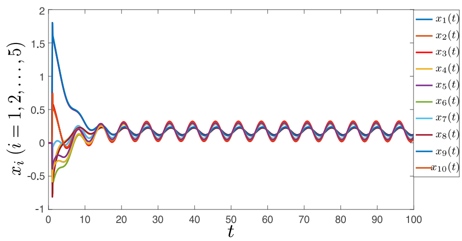

Figure 6. State trajectories of $w_i(t)$.

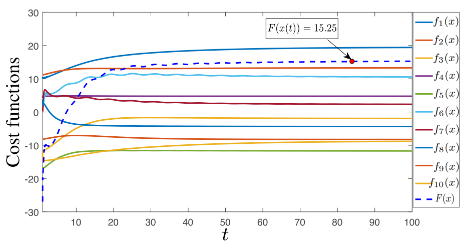

Figure 7. The diagrams optimization function.

In the proposed protocol (7), we choose $k = 2$, $\alpha = 18$, $\beta = 2.5$. In addition, $p = 2$ and $\sigma = 10$ that satisfy the hypothesis are selected. By calculating one can be obtained that

$$H = \begin{pmatrix} -1/8 & 1 \\ -17/16 & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} 33/2 & -1 \\ -1 & 266/17 \end{pmatrix},$$

$a_1 = 1$, $a_2 = 1$, $a_3 = 5.377$, $a_4 = 19$. Obviously, all conditions of Theorem 2 are satisfied. Let the initial values $x(0) = [1.3,\ 2.4,\ 7.1,\ 3.5,\ 5.6,\ 1.2,\ 2.4,\ 7,\ 3.2,\ 5.4]^{\mathrm{T}}$, $w(0) = [0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]^{\mathrm{T}}$, under algorithm (7), the simulation results are shown in Figs. 5–7. It is easy to observe that all agents' states gradually reach consensus and converge to optimization state $x^* = 3.52$ in Fig. 5. In addition, the evolution processes of the global optimization function and local objective functions are shown in Fig. 7, the value of global optimization function $F(x(t))$ is 15.25.

In control protocol (7), if we conduct data simulation on the result without considering the internal model term, it can be found that the states of all agents cannot converge to the same state in Fig. 6, thus the optimization problem (5) cannot be solved. This also fully shows that the internal mode principle can effectively offset the external interference of the system and can make the states of the system achieve the same.

5 Conclusion

In this paper, the distributed optimization problem is solved for multi-agent systems with communication delays and external disturbances in a directed network. Firstly, a continuous distributed optimization algorithm is proposed based on the internal model principle. To ensure the states of all agents converge to the optimal value of systems, some sufficient conditions are derived based on Lyapunov–Razumikhin theory and graph theory in which the upper bound of communication delays $\tau(t)$ can be effectively estimated. Finally, the effectiveness of the optimization algorithm can be illustrated by an example.

References

1. G. Amrita, H.Subir, K. Sajal, Distributed on-demand clustering algorithm for lifetime optimization in wireless sensor networks, *J. Parallel Distrib. Comput.*, **141**:129–142, 2020, https://doi.org/10.1016/j.jpdc.2020.03.014.
2. H. Azarhava, J. Niza, Energy efficient resource allocation in wireless energy harvesting sensor networks, *IEEE Wireless Commun. Lett.*, **7**:1000–1003, 2020, https://doi.org/10.1109/LWC.2020.2978049.
3. Y. Cao, L. Zhang, C. Li, Michael Z.Q. Chen, Observer-based consensus tracking of nonlinear agents in hybrid varying directed topology, *IEEE Trans. Cybern.*, **47**:2212–2222, 2017, https://doi.org/10.1109/TCYB.2016.2573138.
4. W. Chen, W. Ren, Event-triggered zero-gradient-sum distributed consensus optimization over directed networks, *Automatica*, **65**:90–97, 2016, https://doi.org/10.1016/j.automatica.2015.11.015.
5. Z. Deng, S. Liang, Y. Hong, Distributed continuous-time algorithms for resource allocation problems over weight balanced digraphs, *IEEE Trans. Cybern.*, **48**:3116–3125, 2018, https://doi.org/10.1109/TCYB.2017.2759141.
6. G. Foderaro, S. Ferrari, T. Wettergren, Distributed optimal control for multi-agent trajectory optimization, *Automatica*, **50**:149–154, 2014, https://doi.org/10.1016/j.automatica.2013.09.014.
7. C. Godsil, G. Royle, *Algebraic Graph Theory*, 2011, https://doi.org/10.1007/978-1-4613-0163-9.
8. J. Hale, *Theory of Functional Differential Equations*, 1991, https://doi.org/10.1007/978-1-4612-9892-2.
9. S. Kia, J. Cortés, S. Martínez, Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication, *Automatica*, **55**:254–264, 2015, https://doi.org/10.1016/j.automatica.2015.03.001.
10. C. Li, S. Chen, J. Li, F. Wang, Distributed multi-step subgradient optimization for multi-agent system, *Syst. Control Lett.*, **128**:26–33, 2019, https://doi.org/10.1016/j.sysconle.2019.04.008.

11. S. Li, X. Nian, Z. Deng, Distributed optimization of general linear multi-agent systems with external disturbance, *J. Franklin Inst.*, **358**:5951–5970, 2021, https://doi.org/10.1016/j.jfranklin.2021.05.024.

12. A. Nedic, A. Ozdaglar, Distributed subgradient methods for multi-agent optimization, *IEEE Trans. Autom. Control*, **54**:48–61, 2009, https://doi.org/10.1109/TAC.2008.2009515.

13. H. Psillakis, Q. Wang, Distributed adaptive consensus of nonlinear heterogeneous agents with delayed and sampled neighbor measurements, *IEEE Trans. Cybern.*, **52**:2340–2350, 2022, https://doi.org/10.1109/TCYB.2020.3009726.

14. R. Saravanakumar, A. Amini, R. Datta, Y. Cao, Reliable memory sampled-data consensus of multi-agent systems with nonlinear actuator faults, *IEEE Trans. Circuits Syst. II Express Briefs*, **69**:2201–2205, 2022, https://doi.org/10.1109/TCSII.2021.3124043.

15. N. Tran, J. Xiao, Y. Wang, W. Yang, Distributed optimization problem with communication delay and external disturbance, *Int. J. Syst. Sci.*, **48**:1–12, 2017, https://doi.org/10.1080/00207721.2017.1382605.

16. X. Wang, Y. Hong, H. Ji, Distributed optimization for a class of nonlinear multiagent systems with disturbance rejection, *IEEE Trans. Cybern.*, **46**:1655–1666, 2016, https://doi.org/10.1109/TCYB.2015.2453167.

17. X. Wang, Y. Hong, P. Yi, H. Ji, Y. Kang, Distributed optimization design of continuous-time multiagent systems with unknown-frequency disturbances, *IEEE Trans. Cybern.*, **47**:2058–2066, 2017, https://doi.org/10.1109/TCYB.2017.2703945.

18. J. Wei, T. Charalambous, Distributed consensus tracking of multi-agent systems with time-varying input/output delays and mismatched disturbances, *IFAC-PapersOnLine*, **54**:787–794, 2021, https://doi.org/10.1016/j.ifacol.2021.06.140.

19. Z. Xu, X. Liu, J. Cao, M. Song, Fixed-time bipartite consensus of nonlinear multi-agent systems under directed signed graphs with disturbances, *J. Franklin Inst.*, **359**:2693–2709, 2022, https://doi.org/10.1016/j.jfranklin.2022.02.023.

20. J. Yan, H. Yu, X. Xia, Distributed optimization of multi-agent systems with delayed sampled-data, *Neurocomputing*, **296**:100–108, 2018, https://doi.org/10.1016/j.neucom.2018.03.036.

21. S. Yang, Q. Liu, J. Wang, Distributed optimization based on a multiagent system in the presence of communication delays, *IEEE Trans. Syst. Man Cybern.*, **47**:717–728, 2017, https://doi.org/10.1109/TSMC.2016.2531649.

22. H. Yu, J. Yan, Distributed optimization of multiagent systems in directed networks with time-varying dela, *J. Control Sci. Eng.*, pp. 1–9, 2017, https://doi.org/10.1155/2017/7937916.

23. M. Zhu, S. Martínez, Sign projected gradient flow: A continuous-time approach to convex optimization with linear equality constraints, *Automatica*, **120**:109156, 2020, https://doi.org/10.1016/j.automatica.2020.109156.