

# Event-Triggered Zero-Gradient-Sum Distributed Algorithm for Convex Optimization with Time-Varying Communication Delays and Switching Directed Topologies

Lei Ye

Department of Mathematics, College of Information Science and Technology, Jinan University, Guangzhou, China

Email: 1390545629@qq.com

**How to cite this paper:** Ye, L. (2022) Event-Triggered Zero-Gradient-Sum Distributed Algorithm for Convex Optimization with Time-Varying Communication Delays and Switching Directed Topologies. *Journal of Applied Mathematics and Physics*, **10**, 1247-1265.

<https://doi.org/10.4236/jamp.2022.104088>

**Received:** March 14, 2022

**Accepted:** April 22, 2022

**Published:** April 25, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Nowadays, distributed optimization algorithms are widely used in various complex networks. In order to expand the theory of distributed optimization algorithms in the direction of directed graph, the distributed convex optimization problem with time-varying delays and switching topologies in the case of directed graph topology is studied. The event-triggered communication mechanism is adopted, that is, the communication between agents is determined by the trigger conditions, and the information exchange is carried out only when the conditions are met. Compared with continuous communication, this greatly saves network resources and reduces communication cost. Using Lyapunov-Krasovskii function method and inequality analysis, a new sufficient condition is proposed to ensure that the agent state finally reaches the optimal state. The upper bound of the maximum allowable delay is given. In addition, Zeno behavior will be proved not to exist during the operation of the algorithm. Finally, a simulation example is given to illustrate the correctness of the results in this paper.

## Keywords

Distributed Convex Optimisation, Zero-Gradient-Sum Algorithm, Event-Triggered, Time-Varying Delays, Switching Network, Multi-Agent System, Directed Network

## 1. Introduction

In today's network information age, due to the rapid development of communication and sensing technology, the original point-to-point control system has

been reorganized, and a new network control system composed of a large number of interrelated subsystems came into being. Networked control system is divided into centralized and distributed structures. In order to deal with more and more complex practical problems, distributed systems have attracted more and more attention, and also derived hot topics such as distributed synchronization [1] and distributed optimization [2]. Among them, DCOP, that is, distributed convex optimization problem, plays an important role in networked control system, so scholars pay great attention to it. The core of this problem is to discuss the following optimization objective in a network with  $N$  nodes:

$$x^* \in \arg \min_{x \in \mathbb{R}^n} \sum_{i=1}^N f_i(x), \quad (1)$$

where  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$  is a local cost function and is assumed to be strongly convex;  $x^* \in \mathbb{R}^n$  is the optimal value of  $\sum_{i=1}^N f_i(x)$ . Optimization problem (1) has a wide range of application scenarios such as sensor scheduling [3] [4], source localization [5], distributed active power optimal control in power systems [6], parallel and distributed computation [7], distributed parameter estimation [8], distributed optimal resource allocation over networks [9], spectrum sensing for cognitive radio networks [10], distributed statistics and machine learning [11], emulation of swarms in biological networks [12], and distributed Lasso [13].

At present, there have been many research results on DCOP. Among them, literature [14] [15] [16] [17] give some consensus-based distributed optimization algorithms to solve the problem (1). However, this kind of algorithm has an obvious defect that its step size is attenuation step, which will lead to slow convergence speed. For this reason, researchers have proposed distributed optimization algorithms based on auxiliary-variables method [18] [19] [20] [21]. These algorithms adopt fixed step size, which improves the convergence speed and accuracy of the algorithm. But this also increases the cost of computing and communication.

In order to overcome the problems caused by attenuation step size and auxiliary-variables at the same time, Lu and Tang proposed a new algorithm called zero gradient sum algorithm (ZGS) in [22]. The main feature of the algorithm is that the initial state of each agent is its own optimal value, and in the subsequent process, the sum of the gradients of all local objective functions is always equal to 0. The advantage of ZGS algorithm is that it has fast convergence speed under the condition of ensuring asymptotic convergence or even exponential convergence. Therefore, researchers had done a lot of work to promote this result [23] [24] [25] [26]. In [23], the authors studied the distributed ZGS consensus problem with time-varying delay and the time delay is a factor that must be considered in practical application. In [24], Liu *et al.* studied the distributed ZGS consensus problem with time-varying topologies. And the finite-time convergence with ZGS algorithms was studied in [25] and [26].

The event-triggered mechanism can greatly reduce the communication cost, so the ZGS algorithm based on event-triggered has also been widely studied [27] [28] [29]. Sampled-data based distributed convex optimization problem with

event-triggered communication was studied in [27]. In [28], the authors studied the event-triggered ZGS distributed optimization problem with time-varying topologies. In [29], Liu and Xie prove the convergence of ZGS algorithm with time-varying delay based on event-triggered mechanism.

All the above studies assume that the topological graph is undirected. In fact, due to the complexity of the real situation, it is also meaningful to study ZGS in the case of directed graph. At present, there have been some researches on ZGS algorithm in the case of directed graph [30] [31] [32], among which in [30], Guo and Chen gave sufficient conditions for the convergence of ZGS algorithm with time-varying delay and switching topology, which has been a great expansion of ZGS algorithm. In [31] and [32], the authors studied the directed graph ZGS algorithm without and with time delays, respectively. However, compared with the case of undirected graph, the research based on directed graph is still relatively few. As far as we know, no one makes the work on event-triggered ZGS optimization algorithm with time-varying delay and switching directed topologies. Considering that time delays always exist and the network topology has the possibility of switching in reality, combined with the need to reduce network communication cost, this research is of significance.

Therefore, to generalize the continuous-time ZGS optimization consensus algorithm, we discuss the convergence of ZGS algorithm with time-varying delay and switching topologies based on event-triggered mechanism under directed networks. Compared with the above related literature, the main contributions of this paper can be summarized as follows:

1) Different from the previous ZGS optimization algorithm results [22]-[32], this paper takes the first step to study the ZGS optimization algorithm with time-varying delay and switching topologies based on event-triggered mechanism under directed networks, which is more challenging and practical.

2) Compared with the work in [29], we consider the possibility of topology switching instead of discussing fixed topology, which is reasonable due to the interference of various uncertain factors in reality. What's more, the scope of application of our algorithm changes from undirected graph to a wider balanced strongly connected directed graph. And by the Lyapunov-Krasovskii-based approach, the sufficient conditions about the maximum admissible time delays are derived.

3) Compared with the work in [30], we add event-triggered mechanism to it, which can better save network resources and reduce communication cost. To our best of knowledge, it is the first time to use event-trigger mechanism in ZGS algorithm considering time-varying delays and switching directed topologies.

The rest of this paper is organized as follows. In Section 2, we give some preliminaries about graph theory and strongly convex functions. The distributed ZGS optimization consensus protocol and convergence analysis are derived in Section 3. Some simulation studies are performed in Section 4 to validate the effectiveness of our proposed ZGS optimization algorithm. Section 5 concludes

this paper.

**Notations:** Let  $\mathbb{R}$  and  $\mathbb{Z}^+$  denote the set of real numbers and the set of positive integers, respectively.  $\mathbb{R}^n$  and  $\mathbb{R}^{n \times n}$  denote the set of  $n \times 1$  real vector and  $n \times n$  real matrix, respectively. Let  $\mathbf{1}_n$  and  $\mathbf{0}_n$  denote, respectively, the  $n \times 1$  column vector of all ones and zeros.  $I_n \in \mathbb{R}^{n \times n}$  denotes the identity matrix.  $A^T$  and  $x^T$  represent the transpose of matrix  $A$  and vector  $x$ , respectively. The Kronecker product of matrix  $A = [a_{ij}] \in \mathbb{R}^{n \times m}$  and  $B \in \mathbb{R}^{p \times q}$  is denoted as  $A \otimes B = \{a_{11}B, \dots, a_{1m}B; \dots; a_{n1}B, \dots, a_{nm}B\} \in \mathbb{R}^{nq \times mp}$ .  $\|x\|$  denotes the standard Euclidean norm of vector  $x$  and  $\inf(\cdot)$  denotes the greatest lower bound. For a continuously differentiable function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\nabla f$  and  $\nabla^2 f$  represent, respectively, the gradient and the Hessian matrix of  $f$ . For matrices  $A$  and  $B$ , the matrix inequalities  $A > B (A \geq B)$  and  $A < B (A \leq B)$  mean that  $A - B$  and  $B - A$  are positive (semi) definite, respectively. Besides, if not explicitly stated, matrices or vectors are assumed to have compatible dimensions.

## 2. Problem Description and Preliminaries

We will first introduce geometric graph theory. In this paper, we use  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$  to represent the directed fixed communication network, where  $\mathcal{V} = \{1, 2, \dots, N\}$  is a finite nonempty node set,  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  represents the edge set of ordered pairs of nodes, and  $\mathcal{A} = [a_{ij}]_{N \times N}$  ( $i, j \in \mathcal{V}$ ) is the adjacency matrix.  $(j, i) \in \mathcal{E}$  means that there is an arc from node  $j$  to node  $i$ . The entry  $a_{ij}$  of the adjacency matrix  $\mathcal{A}$  is greater than zero if and only if  $(j, i) \in \mathcal{E}$ , otherwise  $a_{ij} = 0$ .  $\mathcal{N}_i = \{j \in \mathcal{V} | (j, i) \in \mathcal{E}\}$  denotes the set of neighbors of the  $i$ th node. The in-degree of node  $i$  is defined as  $d_i = \sum_{j=1}^N a_{ij}$  and the in-degree matrix  $\mathcal{D}$  is defined as  $\mathcal{D} = \text{diag}\{d_1, \dots, d_N\}$ . The Laplacian matrix  $\mathcal{L}$  associated with the graph  $\mathcal{G}$  is defined as  $\mathcal{L} = \mathcal{D} - \mathcal{A}$ . For switching topology, let  $\Theta = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_l\} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}_{\sigma(t)}\}$  be a finite set of directed graphs. Define the switching signal  $\sigma(t): [0, +\infty) \rightarrow \Upsilon = \{1, 2, \dots, l\}$  ( $l \in \mathbb{Z}^+$  denotes the total number of all possible graphs). For any time interval in which the  $k$ th topology is activated, we have  $\mathcal{G}_{\sigma(t)} = \mathcal{G}_k \in \Theta$ , and the Laplacian matrix is described as  $\mathcal{L}_k$ .

Next, the strongly convex functions will be introduced.

**Definition 1.** [22] A twice continuously differentiable function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be locally strong convex on any convex and compact set  $S$  if there exists a constant  $m > 0$  such that the following three equivalent conditions hold for any  $x, y \in S$ :

$$\begin{cases} f(y) - f(x) - \nabla f^T(x)(y-x) \geq \frac{m}{2} \|y-x\|^2, \\ (\nabla f(y) - \nabla f(x))^T (y-x) \geq m \|y-x\|^2, \\ \nabla^2 f(x) \geq mI_n, \end{cases} \tag{2}$$

where  $m$  is called the convexity parameter of  $f$ .

And if the function  $f$  is strong convex for any  $S$ . We have the following propo-

sition:

**Proposition 1.** [22] [23] Define the subset  $Q = \{x \in S \mid f(x) \leq f(x(0))\}$ , where  $x(0)$  can be chosen such that  $Q$  is closed. Combining with (2), we can get that  $Q$  is a compact set. Then,  $\forall x, y \in Q$ , there exists a constant  $M$  such that the following equivalent conditions hold:

$$\begin{cases} f(y) - f(x) - \nabla f^T(x)(y-x) \leq \frac{M}{2} \|y-x\|^2, \\ (\nabla f(y) - \nabla f(x))^T (y-x) \leq M \|y-x\|^2, \\ \nabla^2 f(x) \leq MI_n. \end{cases} \quad (3)$$

Finally, We will list the important lemmas needed in this paper.

**Lemma 1.** [20] The following three notions are equivalent: i)  $\mathcal{G}$  is weight-balanced, ii)  $\mathbf{1}_N^T \mathcal{L} = \mathbf{0}$ , and iii)  $\mathcal{L} + \mathcal{L}^T$  is positive semi-definite. Moreover, if  $\mathcal{G}$  is weight-balanced and strongly connected, then zero is a simple eigenvalue of  $\mathcal{L} + \mathcal{L}^T$ .

**Lemma 2.** [34] For any vectors  $x, y \in \mathbb{R}^n$  and one positive matrix  $H \in \mathbb{R}^{n \times n}$ , the following inequality holds:

$$2x^T y \leq x^T Hx + y^T H^{-1}y.$$

**Lemma 3.** [35] If the positive constant matrix  $W \in \mathbb{R}^{n \times n}$ , the scalar  $\tau > 0$  and the concerned integrations of the vector function  $\omega(r): [t-\tau, t] \rightarrow \mathbb{R}^n$  are well defined, then the following inequality is satisfied:

$$\left( \int_{t-\tau}^t \omega(r) dr \right)^T W \left( \int_{t-\tau}^t \omega(r) dr \right) \leq \tau \int_{t-\tau}^t \omega(r)^T W \omega(r) dr. \quad (4)$$

**Lemma 4.** [32] Assume that the graph  $\mathcal{G}$  is strongly connected and balanced, then the following inequality is true for any vector  $x$  with appropriate dimension:

$$x^T \mathcal{L}x \geq \frac{\lambda_2(\mathcal{L} + \mathcal{L}^T)}{2\lambda_{\max}(\mathcal{L}^T \mathcal{L})} x^T \mathcal{L}^T \mathcal{L}x, \quad (5)$$

where  $\lambda_2(\mathcal{L} + \mathcal{L}^T)$  is the minimum nonzero eigenvalue of matrix  $\mathcal{L} + \mathcal{L}^T$ ,  $\lambda_{\max}(\mathcal{L}^T \mathcal{L})$  denotes the maximum eigenvalue of the matrix  $\mathcal{L}^T \mathcal{L}$ .

**Lemma 5.** [36] If a differential function  $f(t)$  satisfies  $f(t), \dot{f}(t) \in L_\infty$ , and  $f(t) \in L_p$  for some value of  $p \in [1, \infty)$ , then  $f(t) \rightarrow 0$  as  $t \rightarrow \infty$ .

### 3. Main Results

In this section, we will elaborate the event-trigger ZGS algorithm with communication delays under switching directed networks and analyze its convergence. Firstly, we need the following two basic assumptions:

**Assumption 1.** For the switching network, the topology  $\mathcal{G}_{\sigma(t)}$  is directed. What's more, the topology is also strongly connected and balanced for any time interval in which the  $k$ th topology is activated.

**Assumption 2.** The cost function  $f_i$  used in (1) is twice continuously differentiable, strongly convex for  $i = 1, \dots, N$ . We assume there exists a convexity

parameter  $m_i > 0$  such that the inequalities in (2) are satisfied.  $f_i$  has an invertible locally Lipschitz Hessian matrix  $\nabla^2 f_i(x)$ .

**Proposition 2.** *With assumption 2, there exists a unique  $x^* \in \mathbb{R}^n$  such that for any  $x \in \mathbb{R}^n$ ,  $F(x^*) \leq F(x)$  and  $\nabla F(x^*) = 0$ . Therefore, problem (1) is well-posed.*

In order to make the algorithm more practical, we take the ubiquitous communication delay in practical applications into account. In this paper, we assume there exists a time-varying communication delay  $\tau(t)$  among agents which satisfies  $\tau(t) \in [0, d]$  and  $\dot{\tau}(t) \leq h$ ,  $h \in [0, 1)$ . In the actual optimization process, the channel between agents may be disconnected, data packet lost, faulty or out of range due to network failure. At the same time, new communication links may appear between agents. For those reasons, we will consider both time delays and switching networks in problem (1).

Since avoiding continuous communication can greatly reduce the consumption of network resources, we want to adopt the event triggered communication mechanism in the algorithm, *i.e.* only when the predefined event-triggered condition satisfies, the agent  $i$  samples its new state and broadcasts it to its neighbours with transition delay  $\tau(t)$ .

Let  $\{t_k^i, k \in \mathbb{Z}^+\}$  denote the event-triggered instants where  $t_k^i \geq 0$  and  $t_0^i = 0$ . And  $\hat{x}_i$  denote the latest broadcast state of agent  $i \in \mathcal{V}$ , that is,

$$\hat{x}_i(t) \triangleq x_i(t_k^i), \quad t \in [t_k^i, t_{k+1}^i);$$

thus,  $\hat{x}_i(t)$  converts the discrete-time signal  $x_i(t_k^i)$  into the continuous-time signal simply by holding its constant until the next event occurs. To determine the trigger instants, we first define the measurement error for agent  $i$  as

$$e_i(t) = \hat{x}_i(t) - x_i(t), \quad t \in [t_k^i, t_{k+1}^i). \tag{6}$$

Then, the trigger instants for agent  $i$  are thus defined iteratively by

$$t_{k+1}^i = \inf \{t : t > t_k^i, E_i(t) \geq 0\}, \tag{7}$$

where the triggering function  $E_i(t)$  is defined as follows:

$$E_i(t) = \|e_i(t)\|^2 - \beta_i^2 \|\hat{z}_i(t)\|^2 - ce^{-2\alpha t}, \tag{8}$$

for some  $\beta_i, c, \alpha > 0$  and  $\hat{z}_i(t) = \sum_{j=1}^N a_{ij}^{\sigma(t)} (\hat{x}_i(t) - \hat{x}_j(t))$ , where  $\hat{x}_j(t) = x(t_k^j)$  represents the latest received states from its neighbour  $j$ . Therefore,  $E_i(t_k^i) = 0$  and  $e_i(t)$  is reset to 0. In this paper, we assume that each agent can obtain its neighbours' information at  $t_k^i$ .

Now, we propose the following event-triggered ZGS optimisation algorithm with time delays and switching topologies under directed networks:

$$\begin{cases} \dot{x}_i(t) = \gamma (\nabla^2 f_i(x_i(t)))^{-1} \sum_{j=1}^N a_{ij}^{\sigma(t)} (\hat{x}_j(t - \tau(t)) - \hat{x}_i(t - \tau(t))), \\ x_i(0) = x_i^*, \forall i \in \mathcal{V} \end{cases} \tag{9}$$

where  $x_i(t) \in \mathbb{R}^n$  denotes the  $i$ th agent's estimate of the unknown minimizer  $x^* \in \mathbb{R}^n$ ;  $x_i(0) \in \mathbb{R}^n$  is the initial state;  $x_i^* \in \mathbb{R}^n$  is an optimal value of the lo-

cal objective function  $f_i$  defined in (1);  $a_{ij}^{\sigma(t)}$  is the connection weight corresponding to the graph  $\mathcal{G}_k$ ;  $\sigma(t)$  is defined in Section 2;  $\gamma$  is a positive gain constant used to adjust the convergence rate;  $\tau(t)$  is the time-varying communication delays between agents. Combining with the definition of  $e_i(t)$  in (6), we can rewrite algorithm (9) as

$$\begin{cases} \dot{x}_i(t) = \gamma \left( \nabla^2 f_i(x_i(t)) \right)^{-1} \sum_{j=1}^N a_{ij}^{\sigma(t)} \left( e_j(t-\tau(t)) + x_j(t-\tau(t)) - e_i(t-\tau(t)) - x_i(t-\tau(t)) \right), \\ x_i(0) = x_i^*, \forall i \in \mathcal{V}. \end{cases} \quad (10)$$

**Remark 1.** Inspired by the work of Liu and Xie [29] and the work of Guo and Chen [30], we got the protocol (9). And for any weight-balanced and strongly connected graph, from (10), we can easily get

$$\begin{aligned} \frac{d}{dt} \sum_{i=1}^N \nabla f_i(x_i(t)) &= -\gamma \sum_{i=1}^N \sum_{j=1}^N a_{ij}^{\sigma(t)} \left( (e_i(t-\tau(t)) - e_j(t-\tau(t))) \right. \\ &\quad \left. + (x_i(t-\tau(t)) - x_j(t-\tau(t))) \right) \\ &= -\gamma \left( \mathbf{1}_N^T \mathcal{L}_k \otimes I_n \right) \left( e(t-\tau(t)) + x(t-\tau(t)) \right) = \mathbf{0}, \end{aligned} \quad (11)$$

where  $e(t) = [e_1^T(t), \dots, e_N^T(t)]^T$ ,  $x(t) = [x_1^T(t), \dots, x_N^T(t)]^T$ . From (11), we know that the gradient sum  $\sum_{i=1}^N \nabla f_i(x_i(t))$  would remain constant along the evolution of system (10). Furthermore, we have

$$\sum_{i=1}^N \nabla f_i(x_i(t)) = \sum_{i=1}^N \nabla f_i(x_i(0)) = \sum_{i=1}^N \nabla f_i(x_i^*) = \mathbf{0}, \forall t > 0. \quad (12)$$

Thus, algorithm (9) also satisfies the ZGS property.

Let  $\xi_i(t) = x_i(t) - x^*$  represent the error between the state of agent  $i$  and the optimisation value  $x^*$ . According to (10), we have

$$\begin{aligned} \dot{\xi}_i(t) &= -\gamma \left( \nabla^2 f_i(x_i(t)) \right)^{-1} \sum_{j=1}^N a_{ij}^{\sigma(t)} \left( (e_i(t-\tau(t)) - e_j(t-\tau(t))) \right. \\ &\quad \left. + (\xi_i(t-\tau(t)) - \xi_j(t-\tau(t))) \right) \\ &= -\gamma \left( \nabla^2 f_i(x_i(t)) \right)^{-1} \sum_{j=1}^N L_{ij}^{(k)} \left( e_j(t-\tau(t)) + \xi_j(t-\tau(t)) \right), \end{aligned} \quad (13)$$

where  $L_{ij}^{(k)}$  denotes the entry of the Laplacian matrix  $\mathcal{L}_k$ .

**Remark 2.** Guo and Chen [30] also studied ZGS algorithm with time-varying delay and switching topology for directed graphs. Different from them, this paper adopts the event triggered communication mechanism, which can reduce the communication cost. At the same time, Zeno behavior was avoided.

**Remark 3.** Compared with the conclusion of Liu *et al.* [29], we relax the condition from undirected graph to directed equilibrium graph. Because the Laplace matrix of directed graph is not symmetric matrix,  $\mathcal{L}^T$  and  $\mathcal{L}$  cannot be regarded as the same in the proof. This will bring us new challenges.

Next, we have the following analysis of the distributed optimisation algorithm (9) based upon the common Lyapunov theory.

**Theorem 1.** *Suppose that Assumptions 1 and 2 are satisfied. If the following*

inequality

$$d \leq \frac{m^2(1-h) \left[ \lambda_2(\mathcal{L}_k + \mathcal{L}_k^T) (1 - 2\hat{\beta} \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)) - 2\hat{\beta} \lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T) \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k) \right]}{4\varepsilon_k m^2(1-h) \lambda_2(\mathcal{L}_k + \mathcal{L}_k^T) (1 - 2\hat{\beta} \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)) + 8\gamma^2 \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)} \quad (14)$$

holds for  $k = 1, 2, \dots, N$ , where  $\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)$  and  $\lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T)$  respectively denote the maximum eigenvalues of the matrix  $\mathcal{L}_k^T \mathcal{L}_k$  and  $\mathcal{L}_k + \mathcal{L}_k^T$ ,  $\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)$  represents the minimum nonzero eigenvalue of matrix  $\mathcal{L}_k + \mathcal{L}_k^T$ ,  $\varepsilon_k = \inf \{v : \mathcal{L}_k \mathcal{L}_k^T \leq v(\mathcal{L}_k + \mathcal{L}_k^T)\} > 0$ ,  $m = \min_{i \in V} (m_i)$ ,  $m_i$  is the convexity parameter of the function  $f_i$  and

$$\hat{\beta} = \max \{\beta_1, \dots, \beta_N\} < \frac{\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)}{2\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k) (\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T) + \lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T))}, \text{ then algo-}$$

rithm (10) with event-triggered condition (7) (8) can solve optimisation problem (1) and the Zeno behaviour will be avoided.

**Proof.** In order to prove our conclusion, we choose the following Lyapunov-Krasovskii function

$$V(t) = V_1(t) + V_2(t) + V_3(t), \quad (15)$$

where

$$V_1(t) = 2 \sum_{i=1}^N (f_i(x^*) - f_i(x_i) - \nabla^T f_i(x_i)(x^* - x_i)), \quad (16)$$

$$V_2(t) = \gamma \int_{-d}^0 \int_{t+s}^t \sum_{i=1}^N (\xi_i^T(r) \dot{\xi}_i(r) + e_i^T(r) \dot{e}_i(r)) dr ds, \quad (17)$$

$$V_3(t) = \frac{2\gamma^3 d}{(1-h)m_i^2} \int_{t-\tau(t)}^t \sum_{i=1}^N \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(r) + L_{ij}^{(k)} \xi_j(r))^T \right) \times \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(r) + L_{ij}^{(k)} \xi_j(r)) \right) dr. \quad (18)$$

Firstly, from (2), we can get

$$V_1(t) \geq \sum_{i=1}^N m_i \|x^* - x_i\|^2. \quad (19)$$

What's more, it is easy to obtain that  $V_2(t) \geq 0, V_3(t) \geq 0$ , for any  $t \geq 0$ . So the Lyapunov function above is well defined.

Taking the time derivative of  $V_1(t)$  along the trajectory evolution of  $x(t)$  of system (10) gives

$$\begin{aligned} \dot{V}_1(t) &= 2 \sum_{i=1}^N x_i^T(t) \nabla^2 f_i(x_i(t)) \dot{x}_i(t) \\ &= -2\gamma \sum_{i=1}^N x_i^T(t) \sum_{j=1}^N a_{ij}^{\sigma(t)} \left( (e_i(t-\tau(t)) - e_j(t-\tau(t))) + (x_i(t-\tau(t)) - x_j(t-\tau(t))) \right) \\ &= -2\gamma \sum_{i=1}^N x_i^T(t) \sum_{j=1}^N a_{ij}^{\sigma(t)} \left( (e_i(t-\tau(t)) - e_j(t-\tau(t))) + (\xi_i(t-\tau(t)) - \xi_j(t-\tau(t))) \right) \\ &= -2\gamma \sum_{i=1}^N \sum_{j=1}^N L_{ij}^{(k)} \xi_i^T(t) \left( (e_j(t-\tau(t)) + \xi_j(t-\tau(t))) \right). \end{aligned} \quad (20)$$



By the Newton-Leibniz formula, we have  $e_j(t - \tau(t)) = e_j(t) - \int_{t-\tau(t)}^t \dot{e}_j(r) dr$ ,  $\xi_j(t - \tau(t)) = \xi_j(t) - \int_{t-\tau(t)}^t \dot{\xi}_j(r) dr$ . Let  $\xi(t) = [\xi_1^T(t), \dots, \xi_N^T(t)]^T$ , by using Kronecker product of matrix, (20) is rearranged as

$$\begin{aligned} \dot{V}_1(t) &= -2\gamma \sum_{i=1}^N \sum_{j=1}^N L_{ij}^{(k)} \xi_i^T(t) \left( e_j(t) - \int_{t-\tau(t)}^t \dot{e}_j(r) dr + \xi_j(t) - \int_{t-\tau(t)}^t \dot{\xi}_j(r) dr \right) \\ &= -2\gamma \sum_{i=1}^N \sum_{j=1}^N L_{ij}^{(k)} \xi_i^T(t) e_j(t) - 2\gamma \sum_{i=1}^N \sum_{j=1}^N L_{ij}^{(k)} \xi_i^T(t) \xi_j(t) \\ &\quad + 2\gamma \int_{t-\tau(t)}^t \sum_{i=1}^N \sum_{j=1}^N L_{ij}^{(k)} \xi_i^T(t) \dot{e}_j(r) dr + 2\gamma \int_{t-\tau(t)}^t \sum_{i=1}^N \sum_{j=1}^N L_{ij}^{(k)} \xi_i^T(t) \dot{\xi}_j(r) dr \quad (21) \\ &= -2\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) e(t) - 2\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) \\ &\quad + 2\gamma \int_{t-\tau(t)}^t \xi^T(t) (\mathcal{L}_k \otimes I_n) \dot{e}(r) dr + 2\gamma \int_{t-\tau(t)}^t \xi^T(t) (\mathcal{L}_k \otimes I_n) \dot{\xi}(r) dr. \end{aligned}$$

Using Young's inequality yields

$$\begin{aligned} \dot{V}_1(t) &\leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + \gamma e^T(t) (\mathcal{L}_k \otimes I_n) e(t) \\ &\quad + 2\gamma \tau(t) \xi^T(t) (\mathcal{L}_k \mathcal{L}_k^T \otimes I_n) \xi(t) + \gamma \int_{t-\tau(t)}^t \dot{e}^T(r) \dot{e}(r) dr \\ &\quad + \gamma \int_{t-\tau(t)}^t \dot{\xi}^T(r) \dot{\xi}(r) dr \quad (22) \\ &\leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + \gamma e^T(t) (\mathcal{L}_k \otimes I_n) e(t) \\ &\quad + 2\gamma d \xi^T(t) (\mathcal{L}_k \mathcal{L}_k^T \otimes I_n) \xi(t) + \gamma \int_{t-d}^t \dot{e}^T(r) \dot{e}(r) dr \\ &\quad + \gamma \int_{t-d}^t \dot{\xi}^T(r) \dot{\xi}(r) dr. \end{aligned}$$

Taking the time derivative of  $V_2(t)$  along the trajectory evolution of  $\xi(t)$  of system (13), we have

$$\begin{aligned} \dot{V}_2(t) &= \gamma \int_{-d}^0 \left( \sum_{i=1}^N \dot{\xi}_i^T(t) \dot{\xi}_i(t) - \dot{\xi}_i^T(t+s) \dot{\xi}_i(t+s) \right) ds \\ &\quad + \gamma \int_{-d}^0 \left( \sum_{i=1}^N \dot{e}_i^T(t) \dot{e}_i(t) - \dot{e}_i^T(t+s) \dot{e}_i(t+s) \right) ds \\ &= \gamma d \sum_{i=1}^N \dot{\xi}_i^T(t) \dot{\xi}_i(t) - \gamma \int_{-d}^0 \sum_{i=1}^N \dot{\xi}_i^T(t+s) \dot{\xi}_i(t+s) ds \\ &\quad + \gamma d \sum_{i=1}^N \dot{e}_i^T(t) \dot{e}_i(t) - \gamma \int_{-d}^0 \sum_{i=1}^N \dot{e}_i^T(t+s) \dot{e}_i(t+s) ds \quad (23) \\ &= \gamma d \sum_{i=1}^N \dot{\xi}_i^T(t) \dot{\xi}_i(t) - \gamma \int_{t-d}^t \sum_{i=1}^N \dot{\xi}_i^T(r) \dot{\xi}_i(r) dr \\ &\quad + \gamma d \sum_{i=1}^N \dot{e}_i^T(t) \dot{e}_i(t) - \gamma \int_{t-d}^t \sum_{i=1}^N \dot{e}_i^T(r) \dot{e}_i(r) dr. \end{aligned}$$

Since  $\nabla^2 f_i(x_i(t)) \geq m I_n$ , where  $m = \min_{i \in V} (m_i)$ , we know that

$$\left( \nabla^2 f_i(x_i(t)) \right)^{-1} \leq \frac{1}{m} I_n \text{ holds. It follows that}$$

$$\begin{aligned} \dot{V}_2(t) &\leq \frac{2\gamma^3 d}{m^2} \sum_{i=1}^N \left( \sum_{j=1}^N L_{ij}^{(k)} (e_j(t-\tau(t)) + \xi_j(t-\tau(t))) \right)^T \\ &\quad \times \left( \sum_{j=1}^N L_{ij}^{(k)} (e_j(t-\tau(t)) + \xi_j(t-\tau(t))) \right) \\ &\quad - \gamma \int_{t-d}^t \dot{\xi}^T(r) \xi(r) dr - \gamma \int_{t-d}^t \dot{e}^T(r) e(r) dr. \end{aligned} \tag{24}$$

Taking the time derivative of  $V_3(t)$  gives

$$\begin{aligned} \dot{V}_3(t) &= \frac{2\gamma^3 d}{(1-h)m^2} \sum_{i=1}^N \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t) + L_{ij}^{(k)} \xi_j(t)) \right)^T \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t) + L_{ij}^{(k)} \xi_j(t)) \right) \\ &\quad - \frac{2\gamma^3 d}{(1-h)m^2} \sum_{i=1}^N \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t-\tau(t)) + L_{ij}^{(k)} \xi_j(t-\tau(t))) \right)^T \\ &\quad \times \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t-\tau(t)) + L_{ij}^{(k)} \xi_j(t-\tau(t))) \right) (1-\dot{\tau}(t)). \end{aligned} \tag{25}$$

Since  $\dot{\tau}(t) \leq h < 1$ , we have  $\frac{1-\dot{\tau}(t)}{1-h} \geq 1$ . Then, we conclude that

$$\begin{aligned} \dot{V}_3(t) &\leq \frac{2\gamma^3 d}{(1-h)m^2} \sum_{i=1}^N \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t) + L_{ij}^{(k)} \xi_j(t)) \right)^T \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t) + L_{ij}^{(k)} \xi_j(t)) \right) \\ &\quad - \frac{2\gamma^3 d}{m^2} \sum_{i=1}^N \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t-\tau(t)) + L_{ij}^{(k)} \xi_j(t-\tau(t))) \right)^T \\ &\quad \times \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t-\tau(t)) + L_{ij}^{(k)} \xi_j(t-\tau(t))) \right). \end{aligned} \tag{26}$$

Together with (22), (24), and (26), one can obtain that

$$\begin{aligned} \dot{V}(t) &= \dot{V}_1(t) + \dot{V}_2(t) + \dot{V}_3(t) \\ &\leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + \gamma e^T(t) (\mathcal{L}_k \otimes I_n) e(t) + 2\gamma d \xi^T(t) (\mathcal{L}_k \mathcal{L}_k^T \otimes I_n) \xi(t) \\ &\quad + \frac{2\gamma^3 d}{(1-h)m^2} \sum_{i=1}^N \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t) + L_{ij}^{(k)} \xi_j(t)) \right)^T \left( \sum_{j=1}^N (L_{ij}^{(k)} e_j(t) + L_{ij}^{(k)} \xi_j(t)) \right) \\ &\leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + \gamma e^T(t) (\mathcal{L}_k \otimes I_n) e(t) + 2\gamma d \xi^T(t) (\mathcal{L}_k \mathcal{L}_k^T \otimes I_n) \xi(t) \\ &\quad + \frac{2\gamma^3 d}{(1-h)m^2} (e(t) + \xi(t))^T (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) (e(t) + \xi(t)) \\ &\leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + \gamma e^T(t) (\mathcal{L}_k \otimes I_n) e(t) + 2\gamma d \xi^T(t) (\mathcal{L}_k \mathcal{L}_k^T \otimes I_n) \xi(t) \\ &\quad + \frac{2\gamma^3 d}{(1-h)m^2} e^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) e(t) + \frac{2\gamma^3 d}{(1-h)m^2} \xi^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) \xi(t) \\ &\quad + \frac{4\gamma^3 d}{(1-h)m^2} e^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) \xi(t) \\ &\leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + \gamma e^T(t) (\mathcal{L}_k \otimes I_n) e(t) + 2\gamma d \xi^T(t) (\mathcal{L}_k \mathcal{L}_k^T \otimes I_n) \xi(t) \\ &\quad + \frac{2\gamma^3 d}{(1-h)m^2} e^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) e(t) + \frac{2\gamma^3 d}{(1-h)m^2} \xi^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) \xi(t) \end{aligned}$$

$$\begin{aligned}
& + \frac{2\gamma^3 d}{(1-h)m^2} e^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) e(t) + \frac{2\gamma^3 d}{(1-h)m^2} \xi^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) \xi(t) \\
& \leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + 2\gamma d \xi^T(t) (\mathcal{L}_k \mathcal{L}_k^T \otimes I_n) \xi(t) \\
& \quad + \frac{4\gamma^3 d}{(1-h)m^2} \xi^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) \xi(t) + \gamma e^T(t) (\mathcal{L}_k \otimes I_n) e(t) \\
& \quad + \frac{4\gamma^3 d}{(1-h)m^2} e^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) e(t) \\
& \leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + 2\gamma d \xi^T(t) (\mathcal{L}_k \mathcal{L}_k^T \otimes I_n) \xi(t) \\
& \quad + \frac{4\gamma^3 d}{(1-h)m^2} \xi^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) \xi(t) + \frac{\gamma}{2} e^T(t) ((\mathcal{L}_k + \mathcal{L}_k^T) \otimes I_n) e(t) \\
& \quad + \frac{4\gamma^3 d}{(1-h)m^2} e^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) e(t) \\
& \leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + 2\gamma d \varepsilon_k \xi^T(t) ((\mathcal{L}_k + \mathcal{L}_k^T) \otimes I_n) \xi(t) \\
& \quad + \frac{4\gamma^3 d}{(1-h)m^2} \xi^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) \xi(t) \\
& \quad + \left( \frac{\gamma}{2} \cdot \lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T) + \frac{4\gamma^3 d}{(1-h)m^2} \cdot \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k) \right) e^T(t) e(t).
\end{aligned} \tag{27}$$

Next, by Lemma 4, we can get that

$$\begin{aligned}
\dot{V}(t) & \leq -\gamma \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) + 2\gamma d \varepsilon_k \xi^T(t) ((\mathcal{L}_k + \mathcal{L}_k^T) \otimes I_n) \xi(t) \\
& \quad + \frac{4\gamma^3 d}{(1-h)m^2} \cdot \frac{2\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)} \xi^T(t) (\mathcal{L}_k \otimes I_n) \xi(t) \\
& \quad + \left( \frac{\gamma}{2} \cdot \lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T) + \frac{4\gamma^3 d}{(1-h)m^2} \cdot \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k) \right) e^T(t) e(t) \\
& = - \left( \frac{\gamma}{2} - 2\gamma \varepsilon_k d - \frac{4\gamma^3 d \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{(1-h)m^2 \lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)} \right) \xi^T(t) ((\mathcal{L}_k + \mathcal{L}_k^T) \otimes I_n) \xi(t) \\
& \quad + \left( \frac{\gamma}{2} \cdot \lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T) + \frac{4\gamma^3 d}{(1-h)m^2} \cdot \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k) \right) e^T(t) e(t).
\end{aligned} \tag{28}$$

Let  $\hat{\beta} = \max\{\beta_1, \dots, \beta_N\}$  and based on event triggered condition (8), we can deduce

$$\begin{aligned}
\|e(t)\|^2 & \leq \hat{\beta} \|\hat{z}(t)\|^2 + ce^{-2\alpha t} \\
& = \hat{\beta} \left( (x(t) + e(t))^T (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) (x(t) + e(t)) \right) + ce^{-2\alpha t} \\
& = \hat{\beta} \left( (\xi(t) + e(t))^T (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) (\xi(t) + e(t)) \right) + ce^{-2\alpha t} \\
& = \hat{\beta} \xi^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) \xi(t) + 2\hat{\beta} e^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) \xi(t) \\
& \quad + \hat{\beta} e^T(t) (\mathcal{L}_k^T \mathcal{L}_k \otimes I_n) e(t) + ce^{-2\alpha t}
\end{aligned}$$

$$\begin{aligned} &\leq 2\hat{\beta}\xi^T(t)(\mathcal{L}_k^T \mathcal{L}_k \otimes I_n)\xi(t) + 2\hat{\beta}e^T(t)(\mathcal{L}_k^T \mathcal{L}_k \otimes I_n)e(t) + ce^{-2\alpha t} \\ &\leq \frac{2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)} \xi^T(t) \left( (\mathcal{L}_k + \mathcal{L}_k^T) \otimes I_n \right) \xi(t) + 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k) \|e(t)\|^2 + ce^{-2\alpha t}. \end{aligned} \tag{29}$$

Suppose  $\hat{\beta} < \frac{\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)}{2\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)(\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T) + \lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T))} < \frac{1}{2\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}$ , it

follows from (29) that

$$\begin{aligned} \|e(t)\|^2 &\leq \frac{2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k))} \xi^T(t) \left( (\mathcal{L}_k + \mathcal{L}_k^T) \otimes I_n \right) \xi(t) \\ &\quad + \frac{c}{1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)} e^{-2\alpha t}. \end{aligned} \tag{30}$$

Substituting (30) into (28), we can obtain

$$\begin{aligned} \dot{V}(t) &\leq - \left( \frac{\gamma}{2} - 2\gamma\varepsilon_k d - \frac{4\gamma^3 d \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{(1-h)m^2 \lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)} - \frac{\gamma\hat{\beta}\lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T)\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k))} \right. \\ &\quad \left. - \frac{8\gamma^3 d \hat{\beta}\lambda_{\max}^2(\mathcal{L}_k^T \mathcal{L}_k)}{m^2(1-h)\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k))} \right) \xi^T(t) \left( (\mathcal{L}_k + \mathcal{L}_k^T) \otimes I_n \right) \xi(t) \\ &\quad + \left( \frac{\gamma c \lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T)}{2(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k))} + \frac{4c\gamma^3 d \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{m^2(1-h)(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k))} \right) e^{-2\alpha t} \end{aligned} \tag{31}$$

Let

$$\begin{aligned} \phi &= \frac{\gamma}{2} - 2\gamma\varepsilon_k d - \frac{4\gamma^3 d \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{(1-h)m^2 \lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)} - \frac{\gamma\hat{\beta}\lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T)\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k))} \\ &\quad - \frac{8\gamma^3 d \hat{\beta}\lambda_{\max}^2(\mathcal{L}_k^T \mathcal{L}_k)}{m^2(1-h)\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k))}, \\ \psi &= \frac{\gamma c \lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T)}{2(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k))} + \frac{4c\gamma^3 d \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}{m^2(1-h)(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k))}, \end{aligned}$$

so because of the conditions

$$d \leq \frac{m^2(1-h) \left[ \lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)) - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T)\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k) \right]}{4\varepsilon_k m^2(1-h)\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)(1 - 2\hat{\beta}\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)) + 8\gamma^2 \lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)}$$

and  $\hat{\beta} < \frac{\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T)}{2\lambda_{\max}(\mathcal{L}_k^T \mathcal{L}_k)(\lambda_2(\mathcal{L}_k + \mathcal{L}_k^T) + \lambda_{\max}(\mathcal{L}_k + \mathcal{L}_k^T))}$  we have  $\phi \geq 0, \psi > 0$ .

Thus,  $\dot{V}(t)$  can be simply expressed as

$$\dot{V}(t) \leq -\phi \xi^T(t) \left( (\mathcal{L}_k + \mathcal{L}_k^T) \otimes I_n \right) \xi(t) + \psi e^{-2\alpha t}. \tag{32}$$

Next, according to the proposition 2 proposed in the work of Chen and Ren

[31], we can get that there exists a positive constant  $\rho_k$  such that

$$\begin{aligned} 0 &\leq m \sum_{i=1}^N \|x^* - x_i(t)\|^2 \leq V_1(t) \\ &\leq \frac{2}{\rho_k} x^T(t) \left( (L_k + L_k^T) \otimes I_n \right) x(t) \\ &= \frac{2}{\rho_k} \xi^T(t) \left( (L_k + L_k^T) \otimes I_n \right) \xi(t) \end{aligned} \tag{33}$$

holds over the compact set

$$\Theta_i^{(k)} = \left\{ x(t) \in R \mid f_i(x^*) - f_i(x(t)) - \nabla f_i(x(t))^T (x^* - x(t)) \leq V_1(x(0)) + mN \left( \frac{1}{8} + \frac{\|L_k\|^2}{32\varepsilon_k^2} \right) \right\}, \text{ where}$$

$\varepsilon_k = \inf \{ v \mid \mathcal{L}_k \mathcal{L}_k^T \leq v(\mathcal{L}_k + \mathcal{L}_k^T) \}$ . Consequently, (33) can be rewritten as

$$\dot{V}(t) \leq -\frac{\phi\rho_k}{2} V_1(t) + \psi e^{-2\alpha t}. \tag{34}$$

Integrating both sides of (36) for any  $t$  yields

$$V(t) - V(0) \leq -\frac{\phi\rho_k}{2} \int_0^t V_1(s) ds + \frac{\psi}{2\alpha}, \tag{35}$$

*i.e.*

$$V(t) + \frac{\phi\rho_k}{2} \int_0^t V_1(s) ds \leq V(0) + \frac{\psi}{2\alpha}, \tag{36}$$

which implies that  $V(t)$  and  $\frac{\phi\rho_k}{2} \int_0^t V_1(s) ds$  are both bounded. It follows from  $V_1(t) \leq V(t)$  that  $V_1(t)$  is bounded. From (19), we get

$\sum_{i=1}^N m_i \|x^* - x_i(t)\|^2 = \sum_{i=1}^N m_i \|\xi_i(t)\|^2$  is bounded. Since  $\sum_{i=1}^N m_i \|\xi_i(t)\|^2 \geq m \|\xi(t)\|^2$ , we get  $\|\xi(t)\|$  is bounded. It follows from (31) that  $\|e(t)\|$  is bounded. Hence, we further get  $\|\dot{\xi}(t)\|$  is bounded according to (14).

Therefore,  $\frac{d\|\xi(t)\|}{dt}$  is bounded due to  $\frac{d\|\xi(t)\|}{dt} \leq \|\dot{\xi}(t)\|$ . By using Lemma 5, we can get that  $\|\xi(t)\| \rightarrow 0$  as  $t \rightarrow \infty$ , *i.e.*  $\|x(t) - x^*\| \rightarrow 0$ , which implies the distributed optimisation problem is solved in system (10).

In the following, we will show that the Zeno-behaviour of triggering time will be excluded through the whole process for  $\forall i \in V$ , *i.e.* there exists a constant  $\zeta > 0$  such that  $t_{k+1}^i - t_k^i \geq \zeta$ .

Note that when  $t \in (t_k^i, t_{k+1}^i)$ ,  $\|\dot{e}_i(t)\| = \|\dot{\xi}_i(t)\|$  and  $\|\dot{\xi}_i(t)\|$  is bounded; Thus, there exists a constant  $\eta > 0$  such that  $\|\dot{e}_i(t)\| \leq \eta$ . Combining with  $\|e(t_k^i)\| = 0$ , we have

$$\int_{t_k^i}^t \|\dot{e}_i(t)\| dt = \|e_i(t)\| \leq \int_{t_k^i}^t \eta dt = \eta(t - t_k^i). \tag{37}$$

From the definition of triggering time sequences, we know that  $E_i(t) \geq 0$  at the next trigger time instant  $t_{k+1}^i$ , i.e.

$$\beta_i \|\hat{z}_i(t)\|^2 + ce^{-2\alpha t} \leq \|e(t_{k+1}^i)\| \leq \eta(t_{k+1}^i - t_k^i). \tag{38}$$

In the evolution of the system, whether  $\|\hat{z}_i(t)\|^2 = 0$  or not, the left expression of (38) is always positive for the existence of the term  $ce^{-2\alpha t}$ . So for every  $t = t_0$  there will always exists a constant  $\zeta(t_0) > 0$  such that

$$\eta(t_{k+1}^i - t_k^i) \geq \zeta(t_0), \text{ i.e. } (t_{k+1}^i - t_k^i) \geq \frac{\zeta(t_0)}{\eta} > 0, \text{ which means that Zeno-behaviour}$$

is excluded for all agents. This completes the proof.

### 4. Numerical Simulations

In this section, we will show the effectiveness and feasibility of our proposed theoretical results in Theorem 1. Here, we assume that there are eight nodes in the directed graph, the node states are scalars, and the local objective functions corresponding to the nodes are as follows:

$$f_i(x) = (x - i)^4 + 8 \times i \times (x - i)^2 \tag{39}$$

with  $i = 1, 2, 3, 4, 5, 6, 7, 8$ . Obviously, the local objective function  $f_i(x)$  satisfies Assumption 2 and the convexity parameters  $m_i = 16, i = 1, 2, 3, 4, 5, 6, 7, 8$ . Our

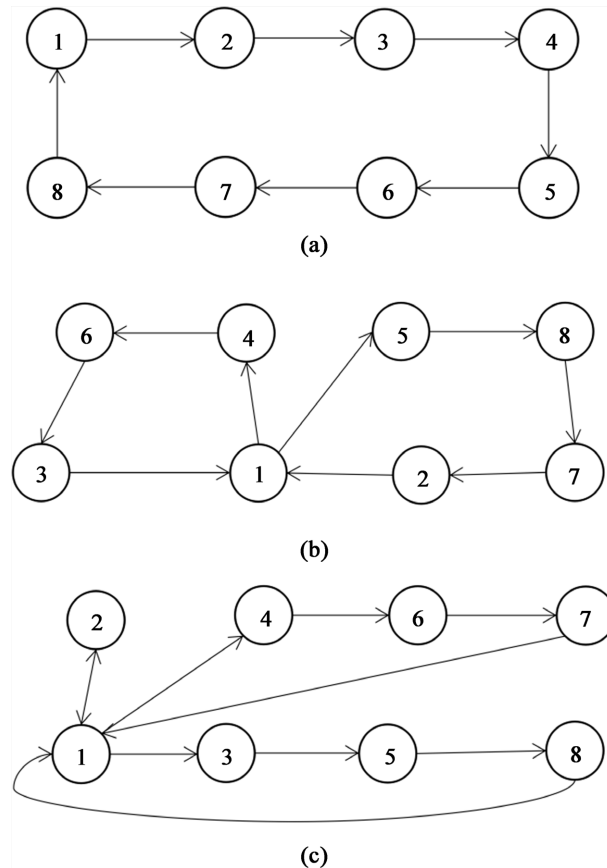
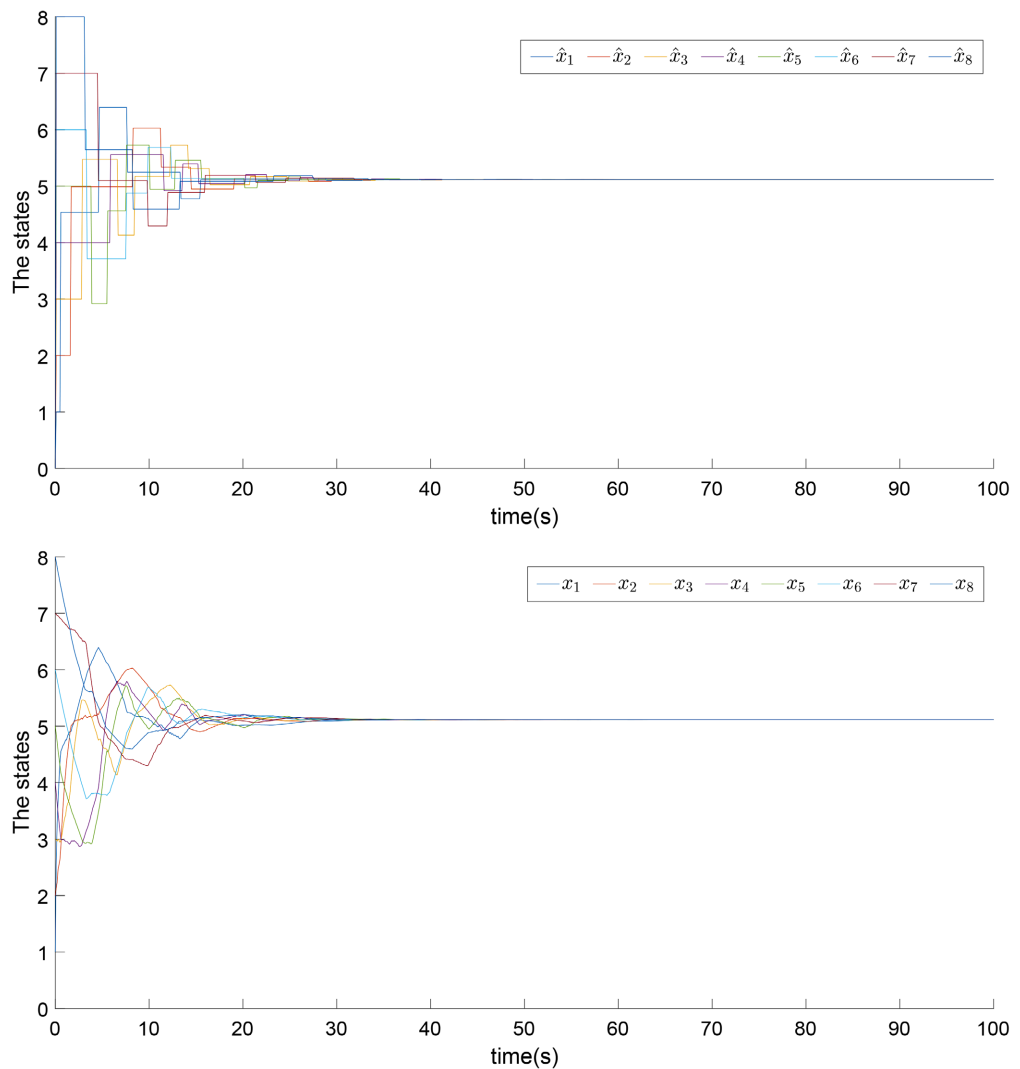


Figure 1. The directed switching topologies: (a)  $\mathcal{G}_1$ ; (b)  $\mathcal{G}_2$ ; (c)  $\mathcal{G}_3$ .

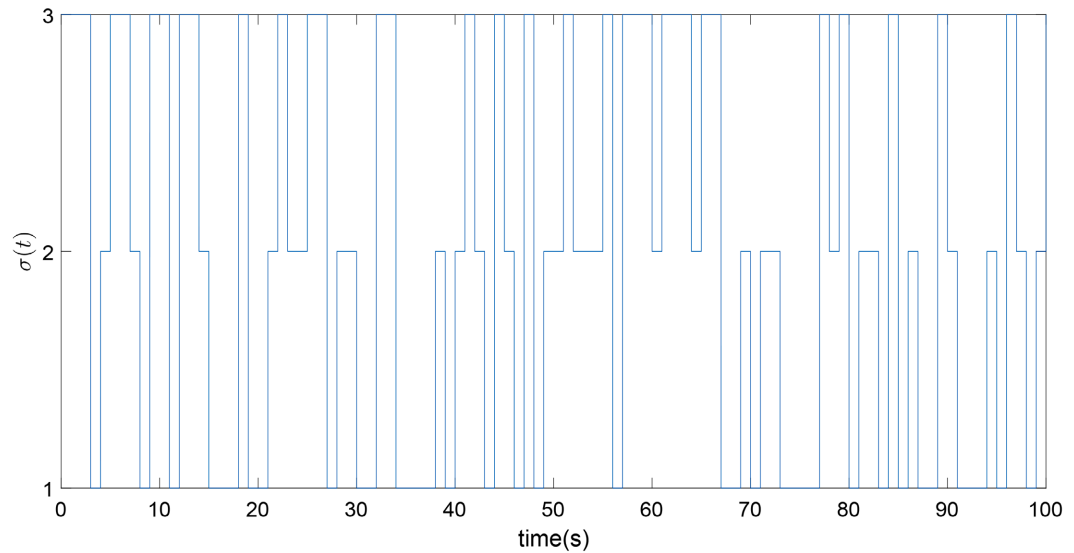
goal is to solve the global optimisation problem  $F(x) = \sum_{i=1}^8 f_i(x)$ . In other words, we need to prove that the states of all nodes will eventually converge to the global optimal value  $x^* = 5.1153$ . We select the initial state of each node as  $x_i^* = i$ ,  $i = 1, 2, \dots, 8$ , which is also the corresponding local optimal value.

As shown in **Figure 1**, three switching cases of directed topological graphs are given, all of which are strongly connected and balanced. According to the calculation, we can get  $\varepsilon_1 = 1$ ,  $\varepsilon_2 = 1.791$ ,  $\varepsilon_3 = 2.894$  and  $\hat{\beta} < 0.00238$ . So we choose the parameters  $\beta_1 = 0.0023$ ,  $\beta_2 = 0.0022$ ,  $\beta_3 = 0.0021$ ,  $\beta_4 = 0.0019$ ,  $\beta_5 = 0.0018$ ,  $\beta_6 = 0.002$ ,  $\beta_7 = 0.0021$ ,  $\beta_8 = 0.0017$ . What's more, we select the parameters  $\gamma = 50$ ,  $c = 15$ ,  $\alpha = 0.16$ , so we can select the time-varying delay as  $\tau(t) = 0.000008 + 0.000008 \sin(t)$ , which meets the condition (14).

All the simulation results are shown as follows and the sample time is 0.1. From **Figure 2**, we can see that the system state  $x_i(t)$  and the recently broadcast state  $\hat{x}_i(t)$  will eventually converge to the global optimum  $x^*$ , which proves our conclusion. **Figure 3** shows the switching signal  $\sigma(t)$ .



**Figure 2.** The trajectories of states of each node.



**Figure 3.** The switching signal  $\sigma(t)$ .

## 5. Conclusion

In this paper, ZGS algorithm with time-varying delays and switching topologies is extended from undirected graph network to directed equilibrium graph. Combined with event triggering mechanism, a new convergence result is proposed. It is proved that the agent state based on the algorithm will converge to the optimal state when the obtained conditions are satisfied. In addition, the algorithm avoids Zeno behavior. Finally, a simulation example is given to verify the effectiveness of the algorithm. In the future, we will try to solve the distributed optimization algorithm with constraints. Because constraints are often used in practical applications, this is a topic of practical significance.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Liu, X. and Qiu, X. (2021) Distributed Synchronization of Stochastic Complex Networks with Time-Varying Delays via Randomly Occurring Control. *Applied Mathematics*, **12**, 803-817. <https://doi.org/10.4236/am.2021.129054>
- [2] Chen, X.B., Yan, K.X., Gao, Y., Xu, X.F., Yan, K. and Wang, J. (2020) Push-Pull Finite-Time Convergence Distributed Optimization Algorithm. *American Journal of Computational Mathematics*, **10**, 118-146. <https://doi.org/10.4236/ajcm.2020.101008>
- [3] Li, C. and Elia, N. (2015) Stochastic Sensor Scheduling via Distributed Convex Optimization. *Automatica*, **58**, 173-182. <https://doi.org/10.1016/j.automatica.2015.05.014>
- [4] Puccinelli, D. and Haenggi, M. (2005) Wireless Sensor Networks: Applications and Challenges of Ubiquitous Sensing. *IEEE Circuits and Systems Magazine*, **5**, 19-31.



- <https://doi.org/10.1109/MCAS.2005.1507522>
- [5] Nedic, A. and Ozdaglar, A. (2009) Distributed Subgradient Methods for Multi-Agent Optimization. *IEEE Transactions on Automatic Control*, **54**, 48-61. <https://doi.org/10.1109/TAC.2008.2009515>
- [6] Chen, G., Ren, J. and Feng, E. (2017) Distributed Finite-Time Economic Dispatch of a Network of Energy Resources. *IEEE Transactions on Smart Grid*, **8**, 822-832.
- [7] Bertsekas, D. and Tsitsiklis, J. (1997) Parallel and Distributed Computation: Numerical Methods. Athena Scientific, Nashua.
- [8] Ram, S., Nedic, A. and Veeravalli, V. (2010) Distributed and Recursive Parameter Estimation in Parametrized Linear State-Space Models. *IEEE Transactions on Automatic Control*, **55**, 488-492. <https://doi.org/10.1109/TAC.2009.2037460>
- [9] Madan, R. and Lall, S. (2006) Distributed Algorithms for Maximum Lifetime Routing in Wireless Sensor Networks. *IEEE Transactions on Wireless Communications*, **5**, 2185-2193. <https://doi.org/10.1109/TWC.2006.1687734>
- [10] Bazerque, J. and Giannakis, G. (2010) Distributed Spectrum Sensing for Cognitive Radio Networks by Exploiting Sparsity. *IEEE Transactions on Signal Processing*, **58**, 1847-1862. <https://doi.org/10.1109/TSP.2009.2038417>
- [11] Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. (2011) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning*, **3**, 1-122. <https://doi.org/10.1561/22000000016>
- [12] Zhao, X., Tu, S. and Sayed, A. (2012) Diffusion Adaptation over Networks under Imperfect Information Exchange and Non-Stationary Data. *IEEE Transactions on Signal Processing*, **60**, 3460-3475. <https://doi.org/10.1109/TSP.2012.2192928>
- [13] Mateos, G., Bazerque, J. and Giannakis, G. (2010) Distributed Sparse Linear Regression. *IEEE Transactions on Signal Processing*, **58**, 5262-5276. <https://doi.org/10.1109/TSP.2010.2055862>
- [14] Nedic, A., Ozdaglar, A. and Parrilo, P.A. (2010) Constrained Consensus and Optimization in Multi-Agent Networks. *IEEE Transactions on Automatic Control*, **55**, 922-938. <https://doi.org/10.1109/TAC.2010.2041686>
- [15] Yuan, D., Xu, S., Zhang, B. and Rong, L. (2013) Distributed Primal-Dual Stochastic Subgradient Algorithms for Multi-Agent Optimization under Inequality Constraints. *International Journal of Robust and Nonlinear Control*, **23**, 1846-1868. <https://doi.org/10.1002/rnc.2856>
- [16] Qiu, Z., Liu, S. and Xie, L. (2018) Necessary and Sufficient Conditions for Distributed Constrained Optimal Consensus under Bounded Input. *International Journal of Robust and Nonlinear Control*, **28**, 2619-2635. <https://doi.org/10.1002/rnc.4040>
- [17] Rahili, S. and Ren, W. (2017) Distributed Continuous-Time Convex Optimization with Time-Varying Cost Functions. *IEEE Transactions on Automatic Control*, **62**, 1590-1605. <https://doi.org/10.1109/TAC.2016.2593899>
- [18] Jakovetic, D., Xavier, J. and Moura, J.M. (2011) Cooperative Convex Optimization in Networked Systems: Augmented Lagrangian Algorithms with Directed Gossip Communication. *IEEE Transactions on Signal Process*, **59**, 3889-3902. <https://doi.org/10.1109/TSP.2011.2146776>
- [19] Duchi, J.C., Agarwal, A. and Wainwright, M.J. (2012) Dual Averaging for Distributed Optimization: Convergence Analysis and Network Scaling. *IEEE Transactions on Automatic Control*, **57**, 592-606. <https://doi.org/10.1109/TAC.2011.2161027>

- [20] Gharesifard, B. and Cortés, J. (2014) Distributed Continuous-Time Convex Optimization on Weight-Balanced Digraphs. *IEEE Transactions on Automatic Control*, **59**, 781-786. <https://doi.org/10.1109/TAC.2013.2278132>
- [21] Yuan, D., Xu, S. and Lu, J. (2015) Gradient-Free Method for Distributed Multi-Agent Optimization via Push-Sum Algorithms. *International Journal of Robust and Nonlinear Control*, **25**, 1569-1580. <https://doi.org/10.1002/rnc.3164>
- [22] Lu, J. and Tang, C.Y. (2011) Zero-Gradient-Sum Algorithms for Distributed Convex Optimization: The Continuous-Time Case. *IEEE Transactions on Automatic Control*, **57**, 2348-2354. <https://doi.org/10.1109/TAC.2012.2184199>
- [23] Yang, Z., Pan, X., Zhang, Q. and Chen, Z. (2020) Distributed Optimization for Multi-Agent Systems with Time Delay. *IEEE Access*, **8**, 123019-123025. <https://doi.org/10.1109/ACCESS.2020.3007731>
- [24] Liu, J., Chen, W. and Dai, H. (2017) Distributed Zero-Gradient-Sum (ZGS) Consensus Optimisation over Networks with Time-Varying Topologies. *International Journal of Systems Science*, **48**, 1836-1843. <https://doi.org/10.1080/00207721.2017.1288840>
- [25] Song, Y. and Chen, W. (2016) Finite-Time Convergent Distributed Consensus Optimisation over Networks. *IET Control Theory & Applications*, **10**, 1314-1318. <https://doi.org/10.1049/iet-cta.2015.1051>
- [26] Wu, Z. and Li, Z. (2020) Finite-Time Distributed Convex Optimization with Zero-Gradient-Sum Algorithms. *IFAC-PapersOnLine*, **53**, 2495-2500.
- [27] Liu, J., Chen, W. and Dai, H. (2016) Sampled-Data Based Distributed Convex Optimization with Event-Triggered Communication. *International Journal of Control, Automation and Systems*, **14**, 1421-1429. <https://doi.org/10.1007/s12555-015-0133-9>
- [28] Liu, J., Chen, W. and Dai, H. (2018) Event-Triggered Zero-Gradient-Sum Distributed Convex Optimisation over Networks with Time-Varying Topologies. *International Journal of Control*, **92**, 2829-2841. <https://doi.org/10.1080/00207179.2018.1460693>
- [29] Liu, J. and Xie, J. (2021) Event-Triggered Zero-Gradient-Sum Distributed Optimization Algorithm with Time-Varying Communication Delays. *International Journal of Systems Science*, **52**, 110-125. <https://doi.org/10.1080/00207721.2020.1820622>
- [30] Guo, Z. and Chen, G. (2018) Distributed Zero-Gradient-Sum Algorithm for Convex Optimization with Time-Varying Communication Delays and Switching Networks. *International Journal of Robust and Nonlinear Control*, **28**, 4900-4915. <https://doi.org/10.1002/rnc.4289>
- [31] Chen, W. and Ren, W. (2016) Event-Triggered Zero-Gradient-Sum Distributed Consensus Optimization over Directed Networks. *Automatica*, **65**, 90-97. <https://doi.org/10.1016/j.automata.2015.11.015>
- [32] Zhao, Z. and Chen, G. (2021) Event-Triggered Scheme for Zero-Gradient-Sum Optimisation under Directed Networks with Time Delay. *International Journal of Systems Science*, **52**, 47-56. <https://doi.org/10.1080/00207721.2020.1819467>
- [33] Boyd, S. and Vandenberghe, L. (2004) Convex Optimization. Cambridge University Press, New York. <https://doi.org/10.1017/CBO9780511804441>
- [34] Dai, H., Jia, J., Yan, L., Wang, F. and Chen, W. (2019) Even-Triggered Exponential Synchronization of Complex Dynamical Networks with Cooperatively Directed Spanning Tree Topology. *Neurocomputing*, **330**, 355-368. <https://doi.org/10.1016/j.neucom.2018.11.013>

- [35] Jia, Q. and Tang, W.K.S. (2011) Leader Following of Nonlinear Agents with Switching Connective Network and Coupling Delay. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **58**, 2508-2519.  
<https://doi.org/10.1109/TCSI.2011.2131230>
- [36] Wang, Y., Cao, J., Wang, H. and Alsaadi, F.E. (2019) Event-Triggered Consensus of Multi-Agent Systems with Nonlinear Dynamics and Communication Delay. *Physica A: Statistical Mechanics and Its Applications*, **522**, 147-157.  
<https://doi.org/10.1016/j.physa.2019.01.124>