SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

July 2026, Vol. 69, Iss. 7, 172207:1–172207:17
https://doi.org/10.1007/s11432-025-4919-0

# Cooperative resilient nonconvex optimization control for nonlinear MASs under DoS attacks

Sha FAN[1], Huaicheng YAN[2,3], Weiwei CHE[4] & Chao DENG[5*]

[1]*College of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210023, China*
[2]*Artificial Intelligence, Shanghai University of Electric Power, Shanghai 200090, China*
[3]*School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China*
[4]*College of Information Science and Engineering, Northeastern University, Shenyang 110819, China*
[5]*Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China*

Received 27 July 2025/Revised 27 October 2025/Accepted 17 November 2025/Published online 18 June 2026

**Abstract** In this paper, we investigate the neural-network-based cooperative resilient nonconvex optimization control issue for high-order nonlinear multi-agent systems (MASs) in the presence of denial-of-service (DoS) attacks. Compared to existing distributed optimization algorithms, a novel control framework for cooperative resilient nonconvex optimization is adopted, which is structured by designing a resilient distributed nonconvex optimization algorithm, developing a smooth-like trajectory, and proposing a neural network backstepping-based controller. Specifically, a resilient distributed nonconvex optimization algorithm is first proposed to ensure convergence under DoS attacks, where the global cost function satisfies the Polyak-Łojasiewicz (P-Ł) condition, which does not require the global cost function to be convex, and the global minimizer is not necessarily unique. Then, a smooth-like trajectory is constructed via Hermite interpolation to generate a new variable with well-defined high-order time derivatives. Furthermore, a neural network backstepping-based scheme is designed for high-order nonlinear MASs, ensuring that the system output tracks the optimal value. Finally, a simulation example is presented to verify the effectiveness of our proposed method.

**Keywords** adaptive control, backstepping control, nonlinear system, multi-agent systems, DoS attacks

**Citation** Fan S, Yan H C, Che W W, et al. Cooperative resilient nonconvex optimization control for nonlinear MASs under DoS attacks. Sci China Inf Sci, 2026, 69(7): 172207, https://doi.org/10.1007/s11432-025-4919-0

## 1 Introduction

Distributed optimization has attracted considerable interest in recent years owing to its extensive applications in multiple domains, such as power systems, sensor networks, and machine learning [1–3]. The essence of distributed optimization lies in the collaborative effort of multiple agents to minimize a global cost function by sharing local information, which enhances computational efficiency and scalability. Compared to centralized methods, distributed optimization algorithms offer advantages in privacy preservation and fault tolerance, but they also face unique challenges, including consensus enforcement, communication constraints, and asynchronous updates. Various distributed optimization algorithms have been proposed to address challenges in both discrete-time and continuous-time domains [4–8].

Note that the distributed algorithms mentioned above focus on the optimization of a sum of convex functions. However, in numerous practical applications, including resource allocation, traffic flow management, and machine learning, the objective functions are inherently nonconvex [9, 10]. Despite the inherent challenges posed by nonconvexity, significant progress has been made in developing distributed nonconvex optimization algorithms that ensure convergence to stationary points [11–13]. For instance, in [12], convergence is achieved by leveraging the Polyak-Łojasiewicz (P-Ł) condition, which is less stringent than strong convexity as it does not require the cost function to be convex, nor does it require a unique global minimizer. Additionally, to overcome the limitation of relying solely on finite function evaluations without gradient information, a derivative-free distributed method is introduced in [13].

In the realm of distributed control, network security has become a paramount concern. The vulnerability of distributed systems to cyber-attacks necessitates the development of robust algorithms capable of withstanding adversarial conditions. Among the common types of cyber attacks are denial-of-service (DoS) attacks, which

* Corresponding author (email: dengchao_neu@126.com)

 

aim to disrupt communication channels [14–16], and false data injection attacks, which corrupt the integrity of transmitted data [17]. Recent studies have focused on mitigating the effects of DoS attacks within distributed optimization frameworks, and have achieved noteworthy results in ensuring resilience against such disruptions [18–21]. Specifically, a comprehensive mathematical framework for modeling attacks and monitoring strategies in cyber-physical systems is presented in [19]. To address the limitation of assuming a fixed maximum number of tolerable attacks, a novel distributed optimization method is introduced in [22]. Furthermore, to analyze the effects of DoS-induced communication failures, a hybrid approach examining attack frequency and corresponding dwell time is proposed in [23]. In many practical scenarios, agents are physical systems, characterized by complex underlying dynamics. A distributed optimization algorithm determines the optimal state, but a control law is required to handle these nonlinear dynamics and ensure that the physical agents can track this optimal reference.

However, there remains a gap in the literature concerning the investigation of distributed nonconvex optimization issues for high-order nonlinear multi-agent systems (MASs) subject to DoS attacks. The primary challenges include the following statements. (i) How to develop a cooperative resilient optimization algorithm to achieve the distributed optimization objective with a nonconvex cost function condition under DoS attacks? (ii) How to develop a new smooth-like trajectory to ensure that the backstepping technique is available for high-order MASs when DoS attacks disrupt the smoothness of the cooperative resilient optimization algorithm?

Motivated by the above observations, this paper investigates the cooperative resilient nonconvex optimization control problem for high-order nonlinear MASs subject to DoS attacks. Compared with previous studies, the main contributions of our developed method are outlined as follows.

(1) It is the first trial to address the distributed nonconvex optimization problem for high-order nonlinear MASs subject to DoS attacks. To address this problem, a novel cooperative resilient nonconvex optimization control strategy is presented, including designing a resilient distributed nonconvex optimization algorithm, developing a smooth-like trajectory, and proposing a neural network backstepping-based controller.

(2) Compared with the distributed optimization results [7,24–26], in which the developed optimization algorithms work based on the strongly convex condition, our developed distributed optimization algorithm ensures exponential convergence under the nonconvex condition, i.e., the global cost function meets the P-Łcondition, which overcomes the limitation of convexity constraints.

(3) Different from the cooperative optimization results for high-order nonlinear MASs [27–29], our proposed approach is capable of achieving the cooperative optimization objective under DoS attacks. To overcome the challenge caused by discontinuous communication, a new smooth-like trajectory is designed using the Hermite interpolation method.

The remainder of this paper is structured as follows. Section 2 provides preliminaries and outlines of the resilient cooperative optimization problem (COP). The solution to COP is presented in Section 3. Section 4 includes a simulation example to illustrate the proposed approach. Finally, Section 5 concludes the paper.

## 2 Preliminaries and problem formulation

### 2.1 Preliminaries

**Notations.** The diagonal matrix consisting of matrices $S_1,\ldots,S_m$ on its main diagonal is represented as $\mathrm{diag}(S_1,\ldots,S_m)$. The notation $\mathrm{col}(\beta_1,\beta_2,\ldots,\beta_M) := [\beta_1^\top,\beta_2^\top,\ldots,\beta_M^\top]^\top$ and $\mathbf{1}_M := [1,\ldots,1]^\top \in \mathbb{R}^M$ with $M$ being the number of agents. The symbol $\mathbf{0}$ refers to the zero matrix/vector, and $I$ denotes the identity matrix. $\otimes$ denotes the Kronecker product. $\|\cdot\|$ is the Euclidean norm for vectors or the spectral norm for matrices. For any square matrix $\aleph$, with compatible dimensions vectors $w$ and $v$, let $\mathcal{Q}_\aleph(w) = w^\top \aleph w$, $\langle w, v\rangle_\aleph = w^\top \aleph v$.

**Graph theory.** In this paper, we examine an undirected graph denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{S})$, where $\mathcal{V} = \{v_1, v_2, \ldots, v_M\}$ represents the set of nodes and $\mathcal{S} = \{(v_i, v_j) | v_i, v_j \in \mathcal{V}\}$ represents the set of edges. If nodes $v_i$ and $v_j$ have the ability to exchange information, then $(v_i, v_j) \in \mathcal{S}$. The nodes $v_i$ and $v_j$ are considered neighbors if $(v_i, v_j) \in \mathcal{S}$. The set of neighbors for node $v_i$ is represented by $\mathcal{N}_i = \{v_j \in \mathcal{V} : (v_i, v_j) \in \mathcal{S}\}$. The undirected graph $\mathcal{G}$ is deemed connected when there exists a sequence of connected edges between any pair of nodes $v_j$ and $v_i$. The adjacency matrix $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{M\times M}$ is defined such that $a_{ii} = 0$, $a_{ij} = 0$ if $(v_i, v_j) \notin \mathcal{S}$, and $a_{ij} = 1$ if $(v_i, v_j) \in \mathcal{S}$. Let $D = \mathrm{diag}(d_1, d_2, \ldots, d_M)$ be the degree matrix with $d_i = \sum_{j=1}^{M} a_{ij}$. Then $\boldsymbol{\mathcal{L}} = D - \mathcal{A}$ is the Laplacian matrix of $\mathcal{G}$.

### 2.2 System dynamics

This paper considers the following high-order ($n$-th) MASs

$$\dot{x}_{i,1} = x_{i,2} + g_{i,1}(x_{i,1}), \tag{1}$$

$$\dot{x}_{i,2} = x_{i,3} + g_{i,2}(\bar{x}_{i,2}), \tag{2}$$

$$\vdots$$

$$\dot{x}_{i,n} = u_i + g_{i,n}(\bar{x}_{i,n}), \tag{3}$$

$$y_i = x_{i,1}, \tag{4}$$

where $\bar{x}_{i,n} = \mathrm{col}\{x_{i,1}, \ldots, x_{i,n-1}\} \in \mathbb{R}^{n-1}$ is the state. $u_i \in \mathbb{R}$ and $y_i \in \mathbb{R}$ represent the input and output of the MASs, respectively. According to the radial basis function neural network (RBFNN) in Lemma [30], the unknown smooth nonlinear function $g_{i,k}(\bar{x}_{i,k}) \in \mathbb{R}, k = 1, \ldots, n; i = 1, \ldots, M$ is approximated by the equation $g_{i,k}(\bar{x}_{i,k}) = \bar{\psi}_{i,k}^{\top}(\bar{x}_{i,k})\theta_i + \epsilon_{i,k}$, where $\bar{\psi}_{i,k}^{\top} = [0, \ldots, \psi_{i,k}^{\top}, 0, \ldots, 0]^{\top}$ and $\theta_i = [\theta_{i,1}, \ldots, \theta_{i,n}]^{\top}$ with $\theta_{i,k} \in \mathbb{R}$ represent the weight vector and the basis vector, respectively. $\epsilon_{i,k}(t)$ satisfies $|\epsilon_{i,k}(t)| \leqslant \bar{\epsilon}_{i,k}$, which is an arbitrarily small positive constant.

### 2.3 DoS attacks

In this paper, agents communicate periodically. However, disruptions in communication between agents can occur when attackers launch DoS attacks, which have the potential to target any channel within the graph $\mathcal{G}$. For an edge $(v_i, v_j)$, use $\mathcal{I}_k^{ij} = [h_k^{ij}, h_k^{ij} + \tau_k^{ij})$ to denote the $k$-th attack interval, where $h_k^{ij}$ and $h_k^{ij} + \tau_k^{ij}$ are respectively the starting and ending instants. Let $\mathcal{S}_{\langle i,j\rangle}^{A} = \bigcup_{k=1}^{\infty} \mathcal{I}_k^{ij}$ be the set when the edge $(v_i, v_j)$ is attacked. Define $\mathcal{S}_i^A = \cup_{j\in\mathcal{N}_i} \mathcal{S}_{\langle i,j\rangle}^{A}$ indicating the set where at least one edge connected to the $i$-th agent is targeted by attackers. In addition, we define $\mathcal{S}^A := \bigcup_{i=1}^{M} \mathcal{S}_i^A$ and $\bar{\mathcal{S}}^A := \complement_{(0,\infty)}\mathcal{S}_i^A$, where $\complement_{(0,\infty)}\mathcal{S}_i^A$ represents the complement of the set $\mathcal{S}_i^A$ concerning the interval $\complement_{(0,\infty)}$.

**Assumption 1** (DoS duration)**.** There exists a positive constant $p$ with $0 < p < 1$ satisfying that

$$\left|\mathcal{S}^A(0,t)\right| \leqslant pt, \ \forall\ t > 0.$$

**Remark 1.** It is worth noting that Assumption 1, utilized in the resilient distributed control work [31], is incorporated to guarantee that the duration of DoS attacks is limited to a certain proportion, specifically $100p\%$ of the total duration.

### 2.4 Resilient optimization objective

For the considered MASs (1)–(4) under DoS attacks satisfying Assumption 1, the main objective of this paper is to develop a cooperative resilient nonconvex optimization control method such that all the signals of the closed-loop system are bounded and all agents' output $\boldsymbol{y} = \mathrm{col}(y_1, y_2, \ldots, y_M)$ reach the optimal $y^*$, which satisfies

$$y^* = \frac{1}{M}\arg\min_{y\in\mathbb{R}^M} \sum_{i=1}^{M} f_i(y), \tag{5}$$

where $f_i(\cdot)$ corresponds to a local objective function.

**Remark 2.** It should be pointed out that achieving (5) involves tackling the following dual challenges: (i) susceptibility to DoS attacks within the network; (ii) local dynamics affected by an unknown function.

### 2.5 Useful assumptions and lemmas

Throughout this paper, the following assumptions and lemmas are given.

**Assumption 2.** The underlying graph $\mathcal{G}$ is undirected and connected.

**Assumption 3.** The function $f_i(x)$ $(i = 1, 2, \ldots, M)$ is twice continuously differentiable and smooth with constant $M_f > 0$, i.e.,

$$\|\nabla f_i(x) - \nabla f_i(y)\| \leqslant M_f \|x - y\|, \forall x, y \in \mathbb{R}. \tag{6}$$

**Assumption 4.** For a private local cost function $f_i$, the optimal set $\mathbb{X}^* = \arg\min_{x\in\mathbb{R}} f(x)$ is nonempty and $f^* = \min_{x\in\mathbb{R}} f(x) > -\infty$.

**Assumption 5.** The global cost function $f(x)$ satisfies the P-Ł condition with constant $\eta > 0$, i.e.,

$$\frac{1}{2}\|\nabla f(x)\|^2 \geqslant \eta(f(x) - f^*). \tag{7}$$

**Remark 3.** Assumptions 2 and 3 are common in the literature (e.g., [10,12]), and based on Lemma 1.2.2 in [32], it can be deduced from Assumption 3 that $\left\|\nabla^2 f_i(x)\right\| \leqslant M_f$ and $\left\|\nabla^2 f(x)\right\| \leqslant M_f$. In addition, Assumption 5 is weaker than strong convexity, in which the global minimizer is not necessarily unique, but each stationary point is a global minimizer.

**Lemma 1** ([25])**.** Given the undirected and connected graph $\mathcal{G}$, there exist a vector $r = \frac{1}{\sqrt{M}}\mathbf{1}_M \in \mathbb{R}^M$ and a matrix $R \in \mathbb{R}^{M\times(M-1)}$ forming the orthogonal matrix $[\,r\ R\,]$ such that (1) $\boldsymbol{\mathcal{L}}K_M = K_M\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{L}}$, $0 < \underline{\rho}(\boldsymbol{\mathcal{L}})K_M \leqslant \boldsymbol{\mathcal{L}} \leqslant \rho(\boldsymbol{\mathcal{L}})K_M$; (2) $RR^\top = K_M$; (3) $Q = R\Gamma_1^{-1}R^\top$, $Q\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{L}}Q = K_M$ and $\frac{K_M}{\rho(\boldsymbol{\mathcal{L}})} \leqslant Q \leqslant \underline{\rho}(\boldsymbol{\mathcal{L}})K_M$, where $K_M = I_M - \frac{1}{M}\mathbf{1}_M\mathbf{1}_M^\top$, $\Gamma_1 = \mathrm{diag}\,(\lambda_2,\ldots,\lambda_M)$ with $\lambda_2,\ldots,\lambda_M$ being the eigenvalues of the Laplacian matrix $\boldsymbol{\mathcal{L}}$.

**Lemma 2** ([10])**.** If the P-Ł condition is satisfied for $f(\cdot)$, the range of the projection operator $\mathcal{P}_{\mathbb{X}^*}(x)$ is set $\mathbb{X}^*$, i.e., $\mathcal{P}_{\mathbb{X}^*}(x) = \arg\min_{y\in\mathbb{X}^*}\|x-y\|^2$, then we have

$$f(x) - f^* \geqslant 2\eta\left\|\mathcal{P}_{\mathbb{X}^*}(x) - x\right\|^2 \quad \forall x \in \mathbb{R}. \tag{8}$$

## 3 Solutions of the resilient cooperative optimization problem

First, a resilient distributed nonconvex optimization algorithm is proposed, which ensures the convergence of the developed algorithm under DoS attacks. Then, a smooth-like trajectory is designed to generate a new variable with well-defined high-order time derivatives. Finally, a neural network backstepping-based controller is designed for high-order nonlinear MASs, which ensures that the output of the MASs can track the optimal value.

### 3.1 Hierarchical equivalence mechanism

Before being given the resilient optimal distributed algorithm, a hierarchical equivalence mechanism is developed in the following lemma to transform the resilient optimization problem into an equivalent problem, which consists of developing a resilient distributed optimization algorithm and designing a neural network controller.

**Lemma 3.** The optimization problem (5) is solvable iff there exists a vector $\boldsymbol{\xi} = [\xi_1,\ldots,\xi_M]^\top$ such that the following conclusions hold at the same time: (i) $\lim_{t\to\infty}(\xi_i(t) - \xi^*) = 0$ with $\xi^* = \frac{1}{M}\arg\min_{\xi\in\mathbb{R}^M}\sum_{i=1}^M f_i(\xi)$; (ii) $\lim_{t\to\infty}(\xi_i(t) - y_i(t)) = 0$.

*Proof.* Based on Lemma 3.1 in [26], the objective function (5) is equivalent to

$$\min\ \frac{1}{M}\sum_{i=1}^M f_i(y_i) \tag{9a}$$

$$\text{s.t. } \boldsymbol{\mathcal{L}}(t)\boldsymbol{y} = \mathbf{0}. \tag{9b}$$

Besides, we can further conclude that the optimization problem (9) is solvable iff there exists a $\xi$ such that the following optimization problem holds

$$\min\ \frac{1}{M}\sum_{i=1}^M f_i(\xi_i) \tag{10a}$$

$$\text{s.t. } \boldsymbol{\mathcal{L}}(t)\boldsymbol{\xi} = \mathbf{0},\ \xi - y = 0. \tag{10b}$$

Based on Lemma 3.1 in [26], it is concluded that the optimization problem (10) is equivalent to the COP (5).

Therefore, the subsequent subsections focus on developing a resilient optimal distributed algorithm to minimize $(1/M)\sum_{i=1}^M f_i(\xi_i)$ and designing a neural network adaptive controller such that the outputs of the systems track the state of the resilient optimal algorithm, respectively.

### 3.2 Resilient optimal distributed algorithm design

To minimize $(1/M)\sum_{i=1}^M f_i(\xi_i(t))$ for the communication network suffering from DoS attacks, based on the idea of [10], our emphasis is on formulating a resilient optimal distributed algorithm in this subsection.

For each agent, we introduce a virtual signal $\xi_i(t)$, and its updating process is governed by

$$\dot{\zeta}_i(t) = \alpha\sum_{j=1}^M a_{ij}^t\left(\xi_j(t) - \xi_i(t)\right), \sum_{i=1}^M \zeta_i(0) = 0, \tag{11a}$$

$$\dot{\xi}_i(t) = -\beta\sum_{j=1}^{M} a_{ij}^t(\xi_j(t)-\xi_i(t)) - \alpha\zeta_i(t) - \nabla f_i(\xi_i(t)). \tag{11b}$$

The introduction of the weight $a_{ij}^t$ is attributed to DoS attacks, with its definition provided as

$$a_{ij}^t = \begin{cases} -a_{ij}, & t \in \complement_{(0,\infty)}\mathcal{S}_{\langle i,j\rangle}^A, \\ 0, & t \in \mathcal{S}_{\langle i,j\rangle}^A, \end{cases}$$

where $\complement_{(0,\infty)}\mathcal{S}_{\langle i,j\rangle}^A$ represents the complement of the set $\mathcal{S}_{\langle i,j\rangle}^A$ concerning the interval $\complement_{(0,\infty)}$. Based on the definition of $a_{ij}^t$, the adjacency matrix under DoS attacks is defined as $\mathcal{A}^t = [a_{ij}^t] \in \mathbb{R}^{M\times M}$. Besides, the Laplacian matrix is defined as $\boldsymbol{\mathcal{L}}(t) = [\mathcal{L}_{ij}^t]$ with

$$\mathcal{L}_{ij}^t = \begin{cases} \sum_{j\neq i} a_{ij}^t, & i = j, \\ -a_{ij}^t, & i \neq j. \end{cases}$$

Define $\boldsymbol{\zeta} = \mathrm{col}(\zeta_1, \ldots, \zeta_M)$ and $\boldsymbol{\xi} = \mathrm{col}(\xi_1, \ldots, \xi_M)$. Subsequently, we can express (11) as follows:

$$\dot{\boldsymbol{\zeta}}(t) = \alpha\boldsymbol{\mathcal{L}}(t)\boldsymbol{\xi}(t), \tag{12a}$$

$$\dot{\boldsymbol{\xi}}(t) = -\alpha\boldsymbol{\zeta}(t) - \beta\boldsymbol{\mathcal{L}}(t)\boldsymbol{\xi}(t) - \nabla f(\boldsymbol{\xi}(t)), \tag{12b}$$

where $\nabla f(\boldsymbol{\xi}(t)) = \mathrm{col}(\nabla f_1(\xi_1), \nabla f_2(\xi_2), \ldots, \nabla f_M(\xi_M))$.

We now present the following convergence result.

**Theorem 1.** If Assumptions 2 and 3 hold, the resilient optimal distributed algorithm (11) is used and the parameters are chosen satisfying $\alpha \in \max\{\sigma_3, \frac{\sigma_1}{\sigma_2-1}\}$, $\beta \in [\alpha+\sigma_1, \sigma_2\alpha]$, then $\xi_i(t)$ converges to a stationary point asymptotically, i.e.,

$$\lim_{t\to\infty}\left(\frac{1}{M}\sum_{i=1}^{M}\|\xi_i(t)-\xi^o(t)\| + \|\nabla f(\xi^o(t))\|\right) = 0, \tag{13}$$

where $\xi^o(t) = \frac{1}{M}\sum_{i=1}^{M}\xi_i(t)$, $\varepsilon_1 = (\beta-\alpha)\underline{\rho}(\boldsymbol{\mathcal{L}}) - [\frac{1}{\alpha}(2M_f^4+\frac{1}{2}) + \frac{3}{2}M_f^2 + \frac{1}{2}]$, $\varepsilon_2 = \alpha - \frac{1}{2} - \frac{b_1b_2}{2\alpha}$, $\varepsilon_3 = \frac{1}{2} - \frac{2M_f^2}{\alpha}$, $\varepsilon_4 = \min\{\varepsilon_1,\varepsilon_2,\varepsilon_3\}$, $b_1 = \frac{\beta}{\alpha} + \frac{1}{\underline{\rho}(\boldsymbol{\mathcal{L}})}$, $b_2 = \frac{\beta}{\alpha} + \frac{1}{\underline{\rho}(\boldsymbol{\mathcal{L}})}$, and $\varepsilon_5 = \min\{\frac{1}{4}, \frac{b_1}{2}-1\}$ with $\sigma_1 = \frac{1}{\underline{\rho}(\boldsymbol{\mathcal{L}})}\left[\frac{1}{\sigma_3}(2M_f^4+\frac{1}{2}) + \frac{3}{2}M_f^2 + 1\right]$, $\sigma_2 = -\frac{1}{\underline{\rho}(\boldsymbol{\mathcal{L}})} + \sqrt{2\sigma_3^2-\sigma_3}$, and $\sigma_3 > \max\left\{4M_f^2, \frac{1}{4}\left[1+\left(1+2\left(2+\frac{2}{\underline{\rho}(\boldsymbol{\mathcal{L}})}\right)^2\right)^{\frac{1}{2}}\right]\right\}$.

*Proof.* The Lyapunov stability analysis is used to show that each $\xi_i(t)$ asymptotically converges to a stationary point. Then, the Lyapunov candidate is chosen as $W(t) = \sum_{i=1}^{M} W_i(t)$, where $W_1(t) = \frac{1}{2}\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))$, $W_2(t) = \frac{1}{2}\mathcal{Q}_{Q+\frac{\beta}{\alpha}K_M}(\boldsymbol{\zeta}(t) + \frac{1}{\alpha}\boldsymbol{h}^o(t))$, $W_3(t) = \langle\boldsymbol{\xi}(t), \boldsymbol{\zeta}(t) + \frac{1}{\alpha}\boldsymbol{h}^o(t)\rangle_{K_M}$, $W_4(t) = M(f(\xi^o(t)) - f^*)$. $\boldsymbol{\xi}^o(t) = \mathbf{1}_M \otimes \xi^o(t)$ and $\boldsymbol{h}^o(t) = \nabla f(\boldsymbol{\xi}^o(t))$, $\boldsymbol{h}(t) = \nabla f(\boldsymbol{\xi}(t))$, $\bar{\boldsymbol{h}}(t) = \frac{1}{M}\mathbf{1}_M\mathbf{1}_M^\top\boldsymbol{h}(t)$, $\bar{\boldsymbol{h}}^o(t) = \frac{1}{M}\mathbf{1}_M\mathbf{1}_M^\top\boldsymbol{h}^o(t) = \mathbf{1}_M \otimes \nabla f(\xi^o(t))$.

We analyze the asymptotic convergence of (12) in the attack-free scenario.

$$\dot{\boldsymbol{\zeta}}(t) = \alpha\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t), \forall\boldsymbol{\xi}(0) \in \mathbb{R}^M, \mathbf{1}_M^\top\boldsymbol{\zeta}(0) = 0, \tag{14a}$$

$$\dot{\boldsymbol{\xi}}(t) = -\alpha\boldsymbol{\zeta}(t) - \beta\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t) - \nabla f(\boldsymbol{\xi}(t)). \tag{14b}$$

Taking the derivative of $W_1(t)$ with respect to (14), one has

$$\begin{aligned}\dot{W}_1(t) =& \boldsymbol{\xi}^\top(t)K_M[-\alpha\boldsymbol{\zeta}(t) - \beta\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t) - \nabla f(\boldsymbol{\xi}(t))] - \left\langle\frac{1}{\alpha}\boldsymbol{h}^o(t) + \frac{1}{\alpha}\boldsymbol{h}^o(t)\right\rangle_{K_M} \\ =& -\beta\mathcal{Q}_{\mathcal{L}}(\boldsymbol{\xi}(t)) - \alpha\left\langle\boldsymbol{\xi}(t), \boldsymbol{\zeta}(t) + \frac{1}{\alpha}\boldsymbol{h}^o(t)\right\rangle_{K_M} + \langle\boldsymbol{\xi}(t), \boldsymbol{h}^o(t) - \boldsymbol{h}(t)\rangle_{K_M} \\ \leqslant& -\beta\mathcal{Q}_{\mathcal{L}}(\boldsymbol{\xi}(t)) - \alpha\left\langle\boldsymbol{\xi}(t), \boldsymbol{\zeta}(t) + \frac{1}{\alpha}\boldsymbol{h}^o(t)\right\rangle_{K_M} + \frac{1}{2}\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t)) + \frac{M_f^2}{2}\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t)),\end{aligned} \tag{15}$$

where the first equality follows from (14a), and the last inequality holds since the Cauchy-Schwarz inequality and (6).

Similarly, the derivatives of $W_2(t)$ are as follows:

$$\begin{aligned}
\dot{W}_2(t) =& \left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)^\top \left(Q+\frac{\beta}{\alpha}K_M\right)\times\left[\alpha\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t)-\frac{1}{\alpha}\nabla^2 f(\boldsymbol{\xi}^o(t))(\bar{\boldsymbol{h}}(t)-\bar{\boldsymbol{h}}^o(t)+\bar{\boldsymbol{h}}^o(t))\right]\\
=&\left\langle \boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\alpha\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t)\right\rangle_{Q+\frac{\beta}{\alpha}K_M}-\frac{1}{\alpha}\left\langle \boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\nabla^2 f(\boldsymbol{\xi}^o(t))(\bar{\boldsymbol{h}}(t)-\bar{\boldsymbol{h}}^o(t))\right\rangle_{Q+\frac{\beta}{\alpha}K_M}\\
&-\left\langle \boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\frac{1}{\alpha}\nabla^2 f(\boldsymbol{\xi}^o(t))\bar{\boldsymbol{h}}^o(t)\right\rangle_{Q+\frac{\beta}{\alpha}K_M}\\
\leqslant&\left\langle \boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\alpha\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t)\right\rangle_{Q+\frac{\beta}{\alpha}K_M}+\frac{b_1}{2\alpha}\times\mathcal{Q}_{Q+\frac{\beta}{\alpha}K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)\\
&+\frac{1}{\alpha b_1}\times\mathcal{Q}_{Q+\frac{\beta}{\alpha}K_M}\left(\nabla^2 f(\boldsymbol{\xi}^o(t))(\bar{\boldsymbol{h}}(t)-\bar{\boldsymbol{h}}^o(t))\right)+\frac{1}{\alpha b_1}\mathcal{Q}_{Q+\frac{\beta}{\alpha}K_M}\left(\nabla^2 f(\boldsymbol{\xi}^o(t)\bar{\boldsymbol{h}}^o(t))\right)\\
\leqslant&\left\langle \boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\alpha\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t)\right\rangle_{Q+\frac{\beta}{\alpha}K_M}+\frac{b_1}{2\alpha}\times\mathcal{Q}_{Q+\frac{\beta}{\alpha}K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)\\
&+\frac{M_f^2}{\alpha}\left\|\bar{\boldsymbol{h}}(t)-\bar{\boldsymbol{h}}^o(t)\right\|^2+\frac{M_f^2}{\alpha}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2\\
\leqslant&\left\langle \boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\alpha\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t)\right\rangle_{Q+\frac{\beta}{\alpha}K_M}+\frac{b_1}{2\alpha}\times\mathcal{Q}_{Q+\frac{\beta}{\alpha}K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)\\
&+\frac{M_f^4}{\alpha}\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+\frac{M_f^2}{\alpha}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2,
\end{aligned}\tag{16}$$

where the first equality holds due to (14b) and $\dot{\xi}^o(t)=-\frac{1}{M}\sum_{i=1}^M\nabla f_i(\xi_i(t))$. The first inequality is satisfied since the Cauchy-Schwarz inequality, and the second inequality holds due to the fact $\left\|\nabla^2 f_i(x)\right\|\leqslant M_f$ and $\left\|\nabla^2 f(x)\right\|\leqslant M_f$ inferring from Assumption 4 and $\rho(\frac{\beta}{\alpha}K_M+Q)\leqslant\rho(\frac{\beta}{\alpha}K_M)+\rho(Q)$, $\rho(K_M)=1$; the last inequality holds due to $\left\|\bar{\boldsymbol{h}}(t)-\bar{\boldsymbol{h}}^o(t)\right\|^2\leqslant\left\|\boldsymbol{h}(t)-\boldsymbol{h}^o(t)\right\|^2\leqslant M_f^2\left\|\boldsymbol{\xi}^o(t)-\boldsymbol{\xi}(t)\right\|^2=M_f^2\mathcal{Q}_{K_M}\big(\boldsymbol{\xi}(t)\big)$.
For $W_3(t)$, the derivatives are as follows:

$$\begin{aligned}
\dot{W}_3(t)=&\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)^\top K_M\times[-\alpha\boldsymbol{\zeta}(t)-\beta\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t)-\nabla f(\boldsymbol{\xi}(t))+\boldsymbol{h}^o(t)-\boldsymbol{h}^o(t)]\\
&+\boldsymbol{\xi}^\top(t)K_M\left[\alpha\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t)-\frac{1}{\alpha}\nabla^2 f(\boldsymbol{\xi}^o(t))\bar{\boldsymbol{h}}(t)\right]\\
=&-\beta\left\langle\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\boldsymbol{\xi}(t)\right\rangle_{\boldsymbol{\mathcal{L}}}-\alpha\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+\left\langle\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\boldsymbol{h}^o(t)-\boldsymbol{h}(t)\right\rangle_{K_M}\\
&+\alpha\mathcal{Q}_{\boldsymbol{\mathcal{L}}}(\boldsymbol{\xi}(t))-\left\langle\frac{1}{\alpha}\boldsymbol{\xi}(t),\nabla^2 f(\boldsymbol{\xi}^o(t))(\bar{\boldsymbol{h}}(t)-\bar{\boldsymbol{h}}^o(t)+\bar{\boldsymbol{h}}^o(t))\right\rangle_{K_M}\\
\leqslant&-\beta\left\langle\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\boldsymbol{\xi}(t)\right\rangle_{\boldsymbol{\mathcal{L}}}-\alpha\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+\frac{1}{2}\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+\frac{M_f^2}{2}\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))\\
&+\alpha\mathcal{Q}_{\boldsymbol{\mathcal{L}}}(\boldsymbol{\xi}(t))+\frac{1}{2\alpha}\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+\frac{M_f^4}{\alpha}\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+\frac{M_f^2}{\alpha}\mathcal{Q}_{K_M}\left(\bar{\boldsymbol{h}}^o(t)\right),
\end{aligned}\tag{17}$$

where the first equality follows from (14) and the inequality holds due to the Cauchy-Schwarz inequality and (6). For $W_4(t)$, the derivative is calculated as

$$\begin{aligned}
\dot{W}_4(t)=&M(\nabla f(\xi^o(t)))^\top\left(-\frac{1}{M}\sum_{i=1}^M\nabla f_i(\xi_i(t))\right)\\
=&-(\nabla f(\xi^o(t)))^\top\sum_{i=1}^M\left(\nabla f_i(\xi_i(t))-\nabla f_i(\xi^o(t))+\nabla f_i(\xi^o(t))\right)\\
=&(\nabla f(\xi^o(t)))^\top\sum_{i=1}^M\left(\nabla f_i(\xi^o(t))-\nabla f_i(\xi_i(t))\right)-\frac{1}{2M}\left(\sum_{i=1}^M\nabla f_i(\xi^o(t))\right)^\top\left(\sum_{i=1}^M\nabla f_i(\xi^o(t))\right)
\end{aligned}$$

$$
\begin{aligned}
&-\frac{M}{2}\left\|\nabla f(\xi^o(t))\right\|^2\\
\leqslant&\left\|\nabla f(\xi^o(t))\right\|\sum_{i=1}^{M}\left\|\nabla f_i(\xi^o(t))-\nabla f_i(\xi_i(t))\right\|-\frac{1}{2}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2-\frac{M}{2}\left\|\nabla f(\xi^o(t))\right\|^2\\
\leqslant&M_f\left\|\nabla f(\xi^o(t))\right\|\sum_{i=1}^{M}\left\|\xi^o(t)-\xi_i(t)\right\|-\frac{1}{2}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2-\frac{M}{2}\left\|\nabla f(\xi^o(t))\right\|^2\\
\leqslant&\frac{M}{2}\left\|\nabla f(\xi^o(t))\right\|^2+\frac{M_f^2}{2M}\left(\sum_{i=1}^{M}\left\|\xi^o(t)-\xi_i(t)\right\|\right)^2-\frac{M}{2}\left\|\nabla f(\xi^o(t))\right\|^2-\frac{1}{2}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2\\
=&\frac{M_f^2}{2M}\left(\sum_{i=1}^{M}\left\|\xi^o(t)-\xi_i(t)\right\|\right)^2-\frac{1}{2}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2\\
\leqslant&\frac{M_f^2}{2}\sum_{i=1}^{M}\left\|\xi^o(t)-\xi_i(t)\right\|^2-\frac{1}{2}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2\\
=&\frac{M_f^2}{2}\mathcal{Q}_{K_M}\big(\boldsymbol{\xi}(t)\big)-\frac{1}{2}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2,
\end{aligned}\tag{18}
$$

where $\nabla f(\xi^o(t))=\frac{1}{M}\sum_{i=1}^{M}\nabla f_i(\xi^o(t))$; the first equality holds due to $\dot{\xi}^o(t)=-\frac{1}{M}\sum_{i=1}^{M}\nabla f_i(\xi_i(t))$; the first and third inequalities hold since Cauchy-Schwarz inequality; the second inequality hold since $\left\|\nabla^2 f_i(x)\right\|\leqslant M_f$ and $\left\|\nabla^2 f(x)\right\|\leqslant M_f$; the last equality holds due to $\sum_{i=1}^{M}\left\|\xi^o(t)-\xi_i(t)\right\|^2=\left\|\boldsymbol{\xi}^o(t)-\boldsymbol{\xi}(t)\right\|^2=\mathcal{Q}_{K_M}\big(\boldsymbol{\xi}(t)\big)$. Then, based on (15)–(18), one has

$$
\begin{aligned}
\dot{W}(t)\leqslant&-\varepsilon_1\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right)-\varepsilon_2\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)-\varepsilon_3\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2\\
\leqslant&-\varepsilon_4\left[\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right)+\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2\right].
\end{aligned}\tag{19}
$$

From conditions $\beta\geqslant\alpha+\sigma_1$, $\alpha>\sigma_3$, $\sigma_1=\frac{1}{\underline{\rho}(\boldsymbol{\mathcal{L}})}[\frac{1}{\sigma_3}(2M_f^4+\frac{1}{2})+\frac{3}{2}M_f^2+1]$, and $\sigma_3>\max\big\{4M_f^2,\frac{1}{4}\big[1+\big(1+2\big(2+\frac{2}{\underline{\rho}(\boldsymbol{\mathcal{L}})}\big)^2\big)^{\frac{1}{2}}\big]\big\}$, one has

$$
\varepsilon_1>\sigma_1\underline{\rho}(\boldsymbol{\mathcal{L}})-\frac{1}{\sigma_3}\left(2M_f^4+\frac{1}{2}\right)-\frac{3}{2}M_f^2-1=0.\tag{20}
$$

From $\beta\leqslant\sigma_2\alpha$, $\alpha>\sigma_3$, one has

$$
\varepsilon_2>\sigma_3-\frac{1}{2\sigma_3}\left(\sigma_2+\frac{1}{\underline{\rho}(\boldsymbol{\mathcal{L}})}\right)^2-\frac{1}{2}=0.\tag{21}
$$

From $\alpha>\sigma_3$, $\sigma_3>4M_f^2$, one has

$$
\varepsilon_3>0.\tag{22}
$$

According to (20)–(22), we know that $\varepsilon_4>0$, and it can be further concluded that $W(t)$ is nonincreasing.

Next, it follows the Cauchy-Schwarz inequality, and we obtain

$$
W(t)\geqslant\varepsilon_5\left[\mathcal{Q}_{K_M}\big(\boldsymbol{\xi}(t)\big)+\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+M(f(\boldsymbol{\xi}^o(t))-f^*)\right].\tag{23}
$$

Then, from the fact that $W(t)$ is nonincreasing and (23), we have

$$
W(0)\geqslant W(t)\geqslant 0.\tag{24}
$$

It follows the way to obtain (15), we have

$$
\left(\frac{d\left\|\boldsymbol{\xi}(t)-\boldsymbol{\xi}^o(t)\right\|}{dt}\right)^2=\left(\frac{\left(\boldsymbol{\xi}(t)-\boldsymbol{\xi}^o(t)\right)^\top\left(\dot{\boldsymbol{\xi}}(t)-\dot{\boldsymbol{\xi}}^o(t)\right)}{\left\|\boldsymbol{\xi}(t)-\boldsymbol{\xi}^o(t)\right\|}\right)^2
$$

$$\leqslant \left\|\dot{\boldsymbol{\xi}}(t)-\dot{\boldsymbol{\xi}}^o(t)\right\|^2=\left\|\dot{\boldsymbol{\xi}}(t)\right\|_{K_M}^2=\|-\beta\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(t)-\alpha\boldsymbol{\zeta}(t)-\nabla f(\boldsymbol{\xi}(t))\|_{K_M}^2$$
$$\leqslant b_3\left[\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right)+\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)\right], \tag{25}$$

where $b_3=\max\{3(\beta^2\rho^2(\boldsymbol{\mathcal{L}})+M_f^2),3\alpha^2\}$. Similarly, one has

$$\left(\frac{d\left\|\bar{\boldsymbol{h}}^o(t)\right\|}{dt}\right)^2=\left(\frac{\left(\bar{\boldsymbol{h}}^o(t)\right)^\top\dot{\bar{\boldsymbol{h}}}^o(t)}{\left\|\bar{\boldsymbol{h}}^o(t)\right\|}\right)^2\leqslant\left\|\dot{\bar{\boldsymbol{h}}}^o(t)\right\|^2=M\left\|\nabla^2 f(\xi^o(t))\dot{\xi}^o(t)\right\|^2$$
$$\leqslant 2M_f^4\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right)+2M_f^2\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2. \tag{26}$$

According to (19) and (24), it follows that $\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right)+\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+\left\|\bar{\boldsymbol{h}}^o(t)\right\|$ is integrable. Then, from (25) and (26), it is known that the derivatives of $\left\|\boldsymbol{\xi}(t)-\bar{\boldsymbol{\xi}}^o(t)\right\|$ and $\left\|\bar{\boldsymbol{h}}^o(t)\right\|$ are square-integrable. Thus, from the Barbalat's lemma, one has $\lim_{t\to\infty}\left(\|\boldsymbol{\xi}(t)-\boldsymbol{\xi}^o(t)\|+\left\|\bar{\boldsymbol{h}}^o(t)\right\|\right)=0$, which is equivalent to (13).

Next, we are in a position to present the main convergence results.

**Theorem 2.** Considering a connected undirected graph $\mathcal{G}$ and assuming Assumptions 1–5 hold. If the resilient optimal distributed algorithm (11) is used with the same $\alpha$ and $\beta$ being given in Theorem 1, and the DoS duration parameter $p$ satisfies $p(\frac{\varepsilon_{12}}{\varepsilon_5}+\frac{\varepsilon_6}{\varepsilon_7})<\frac{\varepsilon_6}{\varepsilon_7}$, then one has $\frac{1}{M}\sum_{i=1}^M\|\xi_i(t)-\xi^o(t)\|^2+f(\xi^o(t))-f^*\leqslant b_4\mathrm{e}^{-\bar{\varepsilon}t}$. Besides, according to the definition of $\mathcal{P}_{\mathbb{X}^*}(\cdot)$, we have

$$\sum_{i=1}^M\|\xi_i(t)-\mathcal{P}_{\mathbb{X}^*}(\xi^o(t))\|^2\leqslant\left(1+\frac{1}{2\eta}\right)Mb_4\mathrm{e}^{-\bar{\varepsilon}t}, \tag{27}$$

where $b_4=\frac{W(0)}{M\varepsilon_5}>0$, $\bar{\varepsilon}=\frac{\varepsilon_6}{\varepsilon_7}-p(\frac{\varepsilon_6}{\varepsilon_7}+\frac{\varepsilon_{12}}{\varepsilon_5})>0$, $\varepsilon_6=\min\{\varepsilon_1,\varepsilon_2,2\eta\varepsilon_3\}>0$, $\varepsilon_7=\max\{1,\frac{1}{2}(b_2+1)\}$, $\varepsilon_8=-\{(\beta-\alpha)\underline{\rho}(\boldsymbol{\mathcal{L}}(t))-[\frac{1}{\alpha}(2M_f^4+\frac{1}{2})+\frac{3}{2}M_f^2+\frac{1}{2}]\}$, $\varepsilon_9=-(\alpha-\frac{1}{2}-\frac{\bar{b}_1\bar{b}_2}{2\alpha})$, $\varepsilon_{10}=-\varepsilon_3=-(\frac{1}{2}-\frac{2M_f^2}{\alpha})$, $\varepsilon_{11}=\max\{\varepsilon_8,\varepsilon_9,\varepsilon_{10}\}$, $\bar{b}_1=\frac{\beta}{\alpha}+\frac{1}{\rho(\boldsymbol{\mathcal{L}}(t))}$, $\bar{b}_2=\frac{\beta}{\alpha}+\frac{1}{\underline{\rho}(\boldsymbol{\mathcal{L}}(t))}$, $\varepsilon_{12}=\max\{\varepsilon_8,\varepsilon_9,2\eta\varepsilon_{10}\}$.

*Proof.* The proof is divided into the following three parts. (i) Consider the time interval during which communication remains attack-free. Combining (7) and (19), we obtain that

$$\dot{W}(t)\leqslant-\varepsilon_1\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))-2\varepsilon_3\eta M\left(f(\xi^o(t))-f^*\right)-\varepsilon_2\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)$$
$$\leqslant-\varepsilon_6\left[\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right)+\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+M\left(f(\xi^o(t))-f^*\right)\right]. \tag{28}$$

Besides, by using the Cauchy-Schwarz inequality, one has

$$W(t)\leqslant\varepsilon_7\left[\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right)+\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+M\left(f(\xi^o(t))-f^*\right)\right]. \tag{29}$$

Then, according to (28) and (29), we have

$$\dot{W}(t)\leqslant-\frac{\varepsilon_6}{\varepsilon_7}W(t). \tag{30}$$

(ii) Consider the time interval during which communication is compromised by DoS attacks, and the dynamics are given in (12). By utilizing the similar scaling rules in (15)–(18), then, the derivative of $W(t)$ with respect to (12) can be expressed as

$$\dot{W}_1(t)\leqslant-\beta\mathcal{Q}_{\boldsymbol{\mathcal{L}}(t)}\left(\boldsymbol{\xi}(t)\right)-\alpha\left\langle\boldsymbol{\xi}(t),\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right\rangle_{K_M}+\frac{1}{2}\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right)+\frac{M_f^2}{2}\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right). \tag{31}$$

Similarly, the derivatives of $W_2(t)$ along satisfy

$$\dot{W}_2(t)\leqslant\left\langle\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\alpha\boldsymbol{\mathcal{L}}(t)\boldsymbol{\xi}(t)\right\rangle_{Q+\frac{\beta}{\alpha}K_M}+\frac{\bar{b}_1}{2\alpha}\mathcal{Q}_{Q+\frac{\beta}{\alpha}K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)$$

$$+\frac{M_f^2}{\alpha}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2+\frac{M_f^4}{\alpha}\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right). \tag{32}$$

For $W_3(t)$, the derivative is

$$\begin{aligned}\dot{W}_3(t)\leqslant&-\beta\left\langle\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t),\boldsymbol{\mathcal{L}}(t)\boldsymbol{\xi}(t)\right\rangle_{K_M}-\alpha\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+\frac{M_f^2}{2}\mathcal{Q}_{K_M}\left(\boldsymbol{\xi}(t)\right)\\&+\frac{1}{2}\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+\alpha\mathcal{Q}_{\boldsymbol{\mathcal{L}}(t)}(\boldsymbol{\xi}(t))+\frac{1}{2\alpha}\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+\frac{M_f^4}{\alpha}\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+\frac{M_f^2}{\alpha}\mathcal{Q}_{K_M}(\bar{\boldsymbol{h}}^o(t)).\end{aligned} \tag{33}$$

The derivatives of $W_4(t)$ under attack-active circumstances remain the same as $\dot{W}_4(t)$ in the case of attack-free. Then, taking the sum of $\dot{W}_1(t)$ to $\dot{W}_4(t)$, one has

$$\begin{aligned}\dot{W}(t)\leqslant&-\mathcal{Q}_{(\beta-\alpha)\boldsymbol{\mathcal{L}}(t)-[\frac{1}{\alpha}(\frac{1}{2}+2M_f^4)+\frac{3}{2}M_f^2+\frac{1}{2}]K_M}(\boldsymbol{\xi}(t))-\mathcal{Q}_{(\alpha-\frac{\beta\bar{b}_1}{2\alpha^2}-\frac{1}{2})K_M-\frac{\bar{b}_1}{2\alpha}Q}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)\\&-\left(\frac{1}{2}-\frac{2M_f^2}{\alpha}\right)\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2\\\leqslant&\varepsilon_8\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+\varepsilon_9\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+\varepsilon_{10}\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2 \qquad (34)\\\leqslant&\varepsilon_{11}\left[\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+\left\|\bar{\boldsymbol{h}}^o(t)\right\|^2\right]. \qquad (35)\end{aligned}$$

Using similar solution means in (20)–(22), it is concluded from (35) that $W(t)$ is also nonincreasing. According to the P-Ł condition in (7) and (34), we have

$$\begin{aligned}\dot{W}(t)\leqslant&\varepsilon_8\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+2\varepsilon_{10}\eta M(f(\xi^o(t))-f^*)+\varepsilon_9\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)\\\leqslant&\varepsilon_{12}\left[\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+\mathcal{Q}_{K_M}\left(\boldsymbol{\zeta}(t)+\frac{1}{\alpha}\boldsymbol{h}^o(t)\right)+M(f(\xi^o(t))-f^*)\right].\end{aligned} \tag{36}$$

Then, based on (23) and (36), we have

$$\dot{W}(t)\leqslant\frac{\varepsilon_{12}}{\varepsilon_5}W(t). \tag{37}$$

(iii) Subsequently, combining (23), (30), (37), and Assumption 1, we have

$$\begin{aligned}\frac{1}{M}\sum_{i=1}^{M}\left\|\xi_i(t)-\xi^o(t)\right\|^2+f(\xi^o(t))-f^*\leqslant&\frac{W(t)}{M\varepsilon_5}\leqslant\mathrm{e}^{-\frac{\varepsilon_6}{\varepsilon_7}|\bar{\mathcal{S}}_A(0,t)|}\mathrm{e}^{\frac{\varepsilon_{12}}{\varepsilon_5}|\mathcal{S}_A(0,t)|}W(0)\\=&\frac{W(0)}{M\varepsilon_5}\mathrm{e}^{-\frac{\varepsilon_6}{\varepsilon_7}t+(\frac{\varepsilon_6}{\varepsilon_7}+\frac{\varepsilon_{12}}{\varepsilon_5})|\mathcal{S}_A(0,t)|}\\\leqslant&b_4\mathrm{e}^{-\bar{\varepsilon}t}.\end{aligned} \tag{38}$$

According to (8) and Cauchy-Schwarz inequality, we have

$$\begin{aligned}\left\|\boldsymbol{\xi}(t)-\mathbf{1}_M\mathcal{P}_{\mathbb{X}^*}(\xi^o(t))\right\|^2=&\left\|\boldsymbol{\xi}(t)-\boldsymbol{\xi}^o(t)+\boldsymbol{\xi}^o(t)-\mathbf{1}_M\mathcal{P}_{\mathbb{X}^*}(\xi^o(t))\right\|^2\\\leqslant&\left(1+\frac{1}{2\eta}\right)\left\|\boldsymbol{\xi}(t)-\boldsymbol{\xi}^o(t)\right\|^2+(1+2\eta)M\times\left\|\xi^o(t)-\mathcal{P}_{\mathbb{X}^*}(\xi^o(t))\right\|^2\\\leqslant&\left(1+\frac{1}{2\eta}\right)\left\|\boldsymbol{\xi}(t)-\boldsymbol{\xi}^o(t)\right\|^2+(1+2\eta)\frac{M}{2\eta}(f(\xi^o(t))-f^*)\\=&\left(1+\frac{1}{2\eta}\right)(\mathcal{Q}_{K_M}(\boldsymbol{\xi}(t))+M(f(\xi^o(t))-f^*))\\\leqslant&\left(1+\frac{1}{2\eta}\right)Mb_4\mathrm{e}^{-\bar{\varepsilon}t}.\end{aligned} \tag{39}$$

**Remark 4.** Compared with the distributed optimization results [24–26, 33], our results establish exponential convergence under reduced requirements. More specifically, the authors in [33] presented a resilient distributed optimization algorithm and ensured the exponential convergence objective under a strongly convex condition, while the developed algorithms exponentially converge to the optimal point under the assumption that the global cost function meets P-Ł condition, which eliminates the strict requirement of convexity and expands the scope of optimization task to more complex situations.

### 3.3 Smooth-like trajectory design

The controller design leverages the classical backstepping framework in [34–36], which requires reference derivatives up to the $n$-th order. To resolve the limitation, we propose the following smooth-like trajectory,

$$s_i(t)=\begin{cases}\xi_i(t_{n-1}),\ t=t_n,\\ \sum\limits_{\kappa=1}^{2(1+n)} v_{i,\kappa,n}t^{2(1+n)-\kappa},\quad t\in(t_n,t_{n+1}),\end{cases}\tag{40}$$

where $v_{i,\kappa,n}=\mathcal{Y}_{i,n}[\kappa]$, and $\mathcal{Y}_{i,n}[\kappa]$ is the $j$ row of $\mathcal{Y}_{i,n}$ with $\mathcal{Y}_{i,n}=\mathfrak{A}_{in}^{-1}\mathfrak{C}_{in}$. $\mathfrak{A}_{in}=\mathrm{col}(\mathfrak{A}_{in,1},\ldots,\mathfrak{A}_{in,2(n+1)})$ and $\mathfrak{C}_{in}=\mathrm{col}(\mathfrak{C}_{in,1},\ldots,\mathfrak{C}_{in,2(n+1)})$, where $\mathfrak{A}_{in,1}=[t_n^{2n+1},\ t_n^{2n},\ldots,t_n,1]$, $\mathfrak{A}_{in,2}=[t_{n+1}^{2n+1},\ldots,t_{n+1},1]$, $\mathfrak{A}_{in,2k-1}=[\frac{(2n+1)!}{(2(n+1)-k)!}t_n^{2(n+1)-k},\frac{(2n)!}{(2n+1-k)!}t_n^{2n+1-k},\ldots,\ t_n,1,0,0,\ldots]$ and $\mathfrak{A}_{in,2k}=[\frac{(2n+1)!}{(2(n+1)-k)!}t_{n+1}^{2(n+1)-k},\ldots,\ t_{n+1},\ 1,0,0,\ldots]$, for $k=2,\ldots,n+1$, $\mathfrak{C}_{in,1}=\xi_i(t_{n-1})$, $\mathfrak{C}_{in,2}=\xi_i(t_n)$, and $\mathfrak{C}_{in,2k-1}=\mathfrak{C}_{in,2k}=0$. Throughout the initial interval $[-h,0)$, the function $s_i(t)$ maintains the constant value $\xi_i(0)$. Discrete time instants are defined by $t_n=nh$, where $n\in\mathbb{Z}\geqslant 0$ and $h$ denotes a small positive constant.

**Lemma 4.** The trajectory $s_i(t)$ converges asymptotically to $\xi^*$ as $t\to\infty$ and possesses well-defined derivatives $s_i^{(n)}(t)$.

*Proof.* As shown in (40), the derivatives of $s_i^{(n)}(t)$ exist at all points except the discrete instants $\{t_n\}$, for $n\in\mathbb{N}+$, which constitute a measure-zero set. Consequently, demonstrating the derivative existence of $\dot{s}_i(t),\ddot{s}_i(t),\ldots,s_i^{(n)}(t)$ specifically at $\{t_n\}$ becomes sufficient. For $t\in(t_n,t_{n+1})$, we have

$$s_i(t_{n+1}^-)=\sum_{\kappa=1}^{2(1+n)} t_{n+1}^{2(1+n)-\kappa}v_{i,\kappa,n},\quad s_i(t_n^+)=\sum_{\kappa=1}^{2(1+n)} t_n^{2(1+n)-\kappa}v_{i,\kappa,n}.\tag{41}$$

Due to the fact that $\begin{bmatrix}\mathfrak{A}_{i,n,1}\\ \mathfrak{A}_{i,n,2}\end{bmatrix}\mathcal{Y}_{i,n}=\begin{bmatrix}\mathfrak{C}_{i,n,1}\\ \mathfrak{C}_{i,n,2}\end{bmatrix}$, it implies $s_i(t_n^+)=s_i(t_{n-1})$ and $s_i(t_{n+1}^-)=s_i(t_n)$. This establishes the continuity of $s_i(t)$ over $[t_n,t_{n+1})$ within this interval, the $k$-th derivative $s_i^{(k)}(t)$ for $k=1,\ldots,n$ is expressed as $s_i^{(k)}(t)=\sum_{\kappa=1}^{2(1+n)-k}\frac{(2(n+1)-\kappa)!}{(2(n+1)-\kappa-k)!}v_{i,\kappa,n}t^{2(1+n)-\kappa-k}$. According to $\begin{bmatrix}\mathfrak{A}_{i,n,2k-1}\\ \mathfrak{A}_{i,n,2k-2}\end{bmatrix}\mathcal{Y}_{i,n}=\begin{bmatrix}\mathfrak{C}_{i,n,2k-1}\\ \mathfrak{C}_{i,n,2k-2}\end{bmatrix}$, we can obtain the existence of $\dot{s}_i(t),\ddot{s}_i(t),\ldots,s_i^{(n)}(t)$.

**Remark 5.** Direct adoption of $\xi_i(t)$ as a virtual reference is problematic: its second-order derivative fails to exist during DoS attacks, which invalidates the backstepping technique. To resolve this limitation, based on the Hermite interpolation method, a new $n$-th order derivative reference $s_i(t)$ is constructed to replace $\xi_i(t)$.

### 3.4 Neural network adaptive controller design

An adaptive controller is designed in this subsection using the backstepping method and adaptive neural networks to ensure that the outputs $y_i$ of the MASs (1)–(4) can track the smooth trajectory $s_i$ for $i=1,2,\ldots,M$. The specific details are as follows.

**Step 1.** Define the tracking error $\phi_{i,1}$ as $\phi_{i,1}=y_i-s_i$. Subsequently, the time derivatives of $\phi_{i,1}$ along (1) and (4) are

$$\dot{\phi}_{i,1}=\phi_{i,2}+\pi_{i,1}+\psi_{i,1}^\top\theta_i+\epsilon_{i,1}-\dot{s}_i,\tag{42}$$

where $\phi_{i,2}$ denotes the error between the state $x_{i,2}$ of the MASs and the virtual control $\pi_{i,1}$, i.e., $\phi_{i,2}=x_{i,2}-\pi_{i,1}$. In this step, the virtual control $\pi_{i,1}$ can be constructed as follows:

$$\pi_{i,1}=-\delta_{i,1}\phi_{i,1}-\psi_{i,1}^\top\hat{\theta}_i-\phi_{i,1}\Upsilon_{i,1}\hat{\epsilon}_{i,1}+\dot{s}_i,\tag{43}$$

where $\delta_{i,1} > 0$ is a positive constant and $\Upsilon_{i,1} = \frac{1}{\sqrt{\phi_{i,1}^2+\mu_i^2(t)}}$. In addition, $\hat{\theta}_i$ and $\hat{\epsilon}_{i,1}$ are adaptive parameters designed based on RBFNN to estimate the unknown constants $\theta_i$ and $\bar{\epsilon}_{i,1}$. $\hat{\epsilon}_{i,1}$ is updated by

$$\dot{\hat{\epsilon}}_{i,1} = \gamma_{i,1}\phi_{i,1}^2\Upsilon_{i,1}, \tag{44}$$

where $\gamma_{i,1}$ is a positive constant. The parameter $\mu_i(t)$ satisfies (i) $\mu_i(t) > 0$; (ii) $\int_0^\infty \mu_i(\tau)d\tau \leqslant \bar{\mu}_i$ with $\bar{\mu}_i$ being a positive constant. Substituting (43) into (42) yields

$$\dot{\phi}_{i,1} = -\delta_{i,1}\phi_{i,1} + \phi_{i,2} - \psi_{i,1}^\top\tilde{\theta}_i + \epsilon_{i,1} - \phi_{i,1}\Upsilon_{i,1}\hat{\epsilon}_{i,1}, \tag{45}$$

where $\tilde{\theta}_i = \hat{\theta}_i - \theta_i$. Choose the following form of Lyapunov function $S_{i,1} = \frac{1}{2}\big(\phi_{i,1}^2 + \tilde{\theta}_i^\top\Pi_i^{-1}\tilde{\theta}_i + \frac{1}{\gamma_{i,1}}\tilde{\epsilon}_{i,1}^2\big)$, where $\tilde{\epsilon}_{i,1} = \hat{\epsilon}_{i,1} - \bar{\epsilon}_{i,1}$ represents the estimation error and $\Pi_i > 0$. Taking the derivative of $S_{i,1}$ with respect to (45) is

$$\dot{S}_{i,1} = -\delta_{i,1}\phi_{i,1}^2 + \phi_{i,1}\phi_{i,2} + \mathcal{T}_1 - \tilde{\theta}_i^\top\Pi_i^{-1}\big(\Pi_i\tau_{i,1} - \dot{\hat{\theta}}_i\big), \tag{46}$$

where $\tau_{i,1} = \psi_{i,1}\phi_{i,1}$ and $\mathcal{T}_1 = \phi_{i,1}\epsilon_{i,1} - \phi_{i,1}^2\Upsilon_{i,1}\hat{\epsilon}_{i,1} + \frac{1}{\gamma_{i,1}}\tilde{\epsilon}_{i,1}\dot{\tilde{\epsilon}}_{i,1}$. As indicated by (44), it yields $\mathcal{T}_1 = -\phi_{i,1}^2\Upsilon_{i,1}\bar{\epsilon}_{i,1} + \phi_{i,1}\epsilon_{i,1}$. Since $|\epsilon_{i,k}(t)| \leqslant \bar{\epsilon}_{i,k}$, one has

$$\mathcal{T}_1 \leqslant |\phi_{i,1}|\bar{\epsilon}_{i,1} - \phi_{i,1}^2\Upsilon_{i,1}\bar{\epsilon}_{i,1} \leqslant \bar{\epsilon}_{i,1}\frac{\mu_i^2(t)}{\sqrt{\phi_{i,1}^2+\mu_i^2(t)} + |\phi_{i,1}|} \leqslant \bar{\epsilon}_{i,1}\mu_i. \tag{47}$$

As a consequence, substituting (47) into (46), we obtain

$$\dot{S}_{i,1} \leqslant -\delta_{i,1}\phi_{i,1}^2 + \phi_{i,1}\phi_{i,2} - \tilde{\theta}_i^\top\Pi_i^{-1}\big(\Pi_i\tau_{i,1} - \dot{\hat{\theta}}_i\big) + \bar{\epsilon}_{i,1}\mu_i. \tag{48}$$

**Step 2.** Let $\phi_{i,3} = x_{i,3} - \pi_{i,2}$ with $\pi_{i,2}$ being a virtual variable. From (2), it can be inferred that

$$\dot{\phi}_{i,2} = \pi_{i,2} + \phi_{i,3} + \epsilon_{i,2} + \psi_{i,2}^\top(x_i)\theta_i - \dot{\pi}_{i,1}, \tag{49}$$

where $\pi_{i,1}$ relies on $x_{i,1}$, $\hat{\theta}_i$, $\hat{\epsilon}_{i,1}$, $\mu_i$ and $\dot{s}_i$. The derivative of $\pi_{i,1}$ can be expressed as

$$\dot{\pi}_{i,1} = \frac{\partial\pi_{i,1}}{\partial x_{i,1}}\dot{x}_{i,1} + \frac{\partial\pi_{i,1}}{\partial\hat{\theta}_i}\dot{\hat{\theta}}_i + \ddot{s}_i + \frac{\partial\pi_{i,1}}{\partial\hat{\epsilon}_{i,1}}\dot{\hat{\epsilon}}_{i,1} + \frac{\partial\pi_{i,1}}{\partial\mu_i}\dot{\mu}_i. \tag{50}$$

Hereafter, $\pi_{i,2}$ is devised as follows:

$$\begin{aligned}\pi_{i,2} = &-\phi_{i,1} - \delta_{i,2}\phi_{i,2} + \frac{\partial\pi_{i,1}}{\partial x_{i,1}}x_{i,2} + \ddot{s}_i + \frac{\partial\pi_{i,1}}{\partial\hat{\theta}_i}\Pi_i\tau_{i,2} - \left(\psi_{i,2} - \frac{\partial\pi_{i,1}}{\partial x_{i,1}}\psi_{i,1}\right)^\top\hat{\theta}_i \\ &+ \frac{\partial\pi_{i,1}}{\partial\hat{\epsilon}_{i,1}}\dot{\hat{\epsilon}}_{i,1} + \frac{\partial\pi_{i,1}}{\partial\mu_i}\dot{\mu}_i - \phi_{i,2}\Upsilon_{i,2}\hat{\epsilon}_{i,2},\end{aligned} \tag{51}$$

where $\delta_{i,2}$ is a positive constant. $\Upsilon_{i,2} = \frac{1}{\sqrt{\phi_{i,2}^2+\mu_i^2(t)}}$ and $\hat{\epsilon}_{i,2}$ represent the estimation of $\bar{\epsilon}_{i,2}$, which is driven by

$$\dot{\hat{\epsilon}}_{i,2} = \gamma_{i,2}\phi_{i,2}^2\Upsilon_{i,2} \tag{52}$$

with $\gamma_{i,2}$ being a positive constant, and $\tau_{i,2}$ being given by

$$\tau_{i,2} = \tau_{i,1} + \left(\psi_{i,2} - \frac{\partial\pi_{i,1}}{\partial x_{i,1}}\psi_{i,1}\right)\phi_{i,2}. \tag{53}$$

Substituting (50)–(53) into (49) yields

$$\dot{\phi}_{i,2} = -\phi_{i,1} - \delta_{i,2}\phi_{i,2} + \phi_{i,3} + \epsilon_{i,2} - \left(\psi_{i,2} - \frac{\partial\pi_{i,1}}{\partial x_{i,1}}\psi_{i,1}\right)^\top\tilde{\theta}_i + \frac{\partial\pi_{i,1}}{\partial\hat{\theta}_i}\big(\Pi_i\tau_{i,2} - \dot{\hat{\theta}}_i\big) - \phi_{i,2}\Upsilon_{i,2}\hat{\epsilon}_{i,2}. \tag{54}$$

Establish the following Lyapunov function as $S_{i,2} = S_{i,1} + \frac{1}{2}\big(\phi_{i,2}^2 + \frac{1}{\gamma_{i,2}}\tilde{\epsilon}_{i,1}^2\big)$, where $\tilde{\epsilon}_{i,2} = \hat{\epsilon}_{i,2} - \bar{\epsilon}_{i,2}$ indicates the estimation error. Using (48), we can derive that

$$\dot{S}_{i,2} \leqslant -\delta_{i,1}\phi_{i,1}^2 - \delta_{i,2}\phi_{i,2}^2 + \phi_{i,2}\phi_{i,3} + \mathcal{T}_2 - \tilde{\theta}_i^\top\big(\tau_{i,2} - \Pi_i^{-1}\dot{\hat{\theta}}_i\big) - \phi_{i,2}\frac{\partial \pi_{i,1}}{\partial \hat{\theta}_i}\big(\dot{\hat{\theta}}_i - \Pi_i\tau_{i,2}\big), \tag{55}$$

where $\mathcal{T}_2 = \phi_{i,2}\epsilon_{i,2} - \phi_{i,2}^2\Upsilon_{i,2}\cdot\hat{\epsilon}_{i,2} + \frac{1}{\gamma_{i,2}}\tilde{\epsilon}_{i,2}\dot{\tilde{\epsilon}}_{i,2}$. Similarly, it yields that

$$\mathcal{T}_2 \leqslant \bar{\epsilon}_{i,2}\mu_i. \tag{56}$$

Substituting (56) into (55) leads to

$$\dot{S}_{i,2} \leqslant -\delta_{i,1}\phi_{i,1}^2 - \delta_{i,2}\phi_{i,2}^2 + \phi_{i,2}\phi_{i,3} - \tilde{\theta}_i^\top\big(\tau_{i,2} - \Pi_i^{-1}\dot{\hat{\theta}}_i\big) + \phi_{i,2}\frac{\partial \pi_{i,1}}{\partial \hat{\theta}_i}\big(\Pi_i\tau_{i,2} - \dot{\hat{\theta}}_i\big) + \mu_i(\bar{\epsilon}_{i,1} + \bar{\epsilon}_{i,2}). \tag{57}$$

**Step** $l$ **(**$l = 3, 4, \ldots, n-1$**).** According to (1) and (4), we have

$$\dot{\phi}_{i,l} = \phi_{i,l} + \pi_{i,l} + \psi_{i,l}^\top\theta_i + \epsilon_{i,l} - \dot{\pi}_{i,l-1}, \tag{58}$$

where $\pi_{i,l-1}$ is a function of $\hat{\theta}_i$, $\hat{\epsilon}_{i,1}, \ldots, \hat{\epsilon}_{i,l}$, $s_i^{(l-1)}$, $x_{i,1}, \ldots, x_{i,l}$, and its derivative is indicated as

$$\dot{\pi}_{i,l-1} = \sum_{s=1}^{l-1}\frac{\partial \pi_{i,l-1}}{\partial x_{i,s}}\dot{x}_{i,s} + \frac{\partial \pi_{i,l-1}}{\partial \hat{\theta}_i}\dot{\hat{\theta}}_i + s_i^{(l)} + \sum_{j=1}^{l-1}\frac{\partial \pi_{i,l-1}}{\partial \hat{\epsilon}_{i,j}}\dot{\hat{\epsilon}}_{i,j} + \frac{\partial \pi_{i,l-1}}{\partial \mu_i}\dot{\mu}_i. \tag{59}$$

Based on adaptive neural networks, the virtual control $\pi_{i,l}$ can be constructed as

$$\begin{aligned}\pi_{i,l} =& -\phi_{i,l-1} - k_{i,l}\phi_{i,l} + \sum_{j=1}^{l-1}\frac{\partial \pi_{i,l-1}}{\partial x_{i,j}}x_{i,j+1} - \varpi_{i,l}^\top\hat{\theta}_i + \frac{\partial \pi_{i,l-1}}{\partial \hat{\theta}_i}\Pi_i\tau_{i,l} + \sum_{j=2}^{l-1}\phi_{i,j}\frac{\partial \pi_{i,j-1}}{\partial \hat{\theta}_i}\Pi_i\varpi_{i,l} + s_i^{(l)} \\ &+ \sum_{j=1}^{l-1}\frac{\partial \pi_{i,l-1}}{\partial \hat{\epsilon}_{i,j}}\dot{\hat{\epsilon}}_{i,j} + \frac{\partial \pi_{i,l-1}}{\partial \mu_i}\dot{\mu}_i - \phi_{i,l}\Upsilon_{i,l}\hat{\epsilon}_{i,l},\end{aligned} \tag{60}$$

where $k_{i,l}$ is a positive constant and $\Upsilon_{i,l} = \frac{1}{\sqrt{\phi_{i,l}^2 + \mu_i^2(t)}}$, $\hat{\epsilon}_{i,l}$ indicates the estimation of $\bar{\epsilon}_{i,l}$, which is updated by

$$\dot{\hat{\epsilon}}_{i,l} = \gamma_{i,l}\phi_{i,l}^2\Upsilon_{i,l} \tag{61}$$

with $\gamma_{i,l}$ being a positive constant. $\varpi_{i,l}$ and $\tau_{i,l}$ are given by $\varpi_{i,l} = \psi_{i,l} - \sum_{j=1}^{l-1}\frac{\partial \pi_{i,l-1}}{\partial x_{i,j}}\psi_{i,j}$ and $\tau_{i,l} = \tau_{i,l-1} + \big(\psi_{i,l} - \sum_{j=1}^{l-1}\frac{\partial \pi_{i,l-1}}{\partial x_{i,j}}\psi_{i,j}\big)\phi_{i,l}$. Define the Lyapunov function $S_{i,l} = \frac{1}{2}\big(\phi_{i,l}^2 + \frac{1}{\gamma_{i,l}}\tilde{\epsilon}_{i,l}^2\big) + S_{i,l-1}$, where $\tilde{\epsilon}_{i,l} = \hat{\epsilon}_{i,l} - \bar{\epsilon}_{i,l}$ indicates the estimation error. Thereafter, we have

$$\begin{aligned}\dot{S}_{i,l} \leqslant& -\sum_{s=1}^{l}k_{i,s}\phi_{i,s}^2 + \phi_{i,s}\phi_{i,s+1} + \tilde{\theta}_i^\top\Pi_i^{-1}(\Pi_i\tau_{i,s} - \dot{\hat{\theta}}_i) \\ &+ \left(\sum_{s=2}^{l}\phi_{i,s}\frac{\partial \pi_{i,s-1}}{\partial \hat{\theta}_i}\right)\Big(\Pi_i\tau_{i,l} - \dot{\hat{\theta}}_i\Big) + \mathcal{T}_l,\end{aligned} \tag{62}$$

where $\mathcal{T}_l = \phi_{i,l}\epsilon_{i,l} - \phi_{i,l}^2\Upsilon_{i,l}\hat{\epsilon}_{i,l} + \frac{1}{\gamma_{i,l}}\tilde{\epsilon}_{i,l}\dot{\tilde{\epsilon}}_{i,l}$. Analogous to the calculation of $\mathcal{T}_1$, we have

$$\mathcal{T}_l \leqslant \bar{\epsilon}_{i,l}\mu_i. \tag{63}$$

Substituting (63) into (62) yields

$$\dot{S}_{i,l} \leqslant -\sum_{s=1}^{l}k_{i,s}\phi_{i,s}^2 + \phi_{i,s}\phi_{i,s+1} + \tilde{\theta}_i^\top\Pi_i^{-1}(\Pi_i\tau_{i,l} - \dot{\hat{\theta}}_i)$$

$$+\left(\sum_{s=2}^{l}\phi_{i,s}\frac{\partial\pi_{i,s-1}}{\partial\hat{\theta}_i}\right)\left(\Pi_i\tau_{i,l}-\dot{\hat{\theta}}_i\right)+\mu_i\sum_{s=1}^{l}\bar{\epsilon}_{i,l}. \tag{64}$$

**Step** $n$**.** From (1)–(4), it yields that $\dot{\phi}_{i,n}=u_i+\psi_{i,n}^{\top}\theta_i+\epsilon_{i,n}-\dot{\pi}_{i,n-1}$, where $\pi_{i,n-1}$ relies on $\hat{\theta}_i$, $\hat{\epsilon}_{i,1}$, $\ldots$, $\hat{\epsilon}_{i,n-1}$, $s_i^{(n-1)}$, and $x_{i,1},\ldots,x_{i,n-1}$. The derivative of $\pi_{i,n-1}$ can be indicated as

$$\dot{\pi}_{i,n-1}=\sum_{s=1}^{n-1}\frac{\partial\pi_{i,n-1}}{\partial x_{i,s}}\dot{x}_{i,s}+\frac{\partial\pi_{i,n-1}}{\partial\hat{\theta}_i}\dot{\hat{\theta}}_i+s_i^{(n-1)}+\frac{\partial\pi_{i,n-1}}{\partial\mu_i}\dot{\mu}_i+\sum_{j=1}^{n-1}\frac{\partial\pi_{i,n-1}}{\partial\hat{\epsilon}_{i,j}}\dot{\hat{\epsilon}}_{i,j}. \tag{65}$$

With RBFNN, the control input $u_i$ is designed as

$$\begin{aligned}u_i=&-\phi_{i,n-1}-\delta_{i,n}\phi_{i,n}+\sum_{j=1}^{n-1}\frac{\partial\pi_{i,n-1}}{\partial x_{i,j}}x_{i,j+1}-\varpi_{i,n}^{\top}\hat{\theta}_i+\frac{\partial\pi_{i,n-1}}{\partial\hat{\theta}_i}\Pi_i\tau_{i,n}+\sum_{j=2}^{n-1}\phi_{i,j}\frac{\partial\pi_{i,j-1}}{\partial\hat{\theta}_i}\Pi_i\varpi_{i,n}+\frac{\partial\pi_{i,n-1}}{\partial\mu_i}\dot{\mu}_i\\&+\sum_{j=1}^{n-1}\frac{\partial\pi_{i,n-1}}{\partial\hat{\epsilon}_{i,j}}\dot{\hat{\epsilon}}_{i,j}+s_i^{(n)}-\phi_{i,n}\Upsilon_{i,n}\hat{\epsilon}_{i,n},\end{aligned} \tag{66}$$

where $\delta_{i,n}$ is a positive constant and $\Upsilon_{i,n}=\frac{\phi_{i,n}}{\sqrt{\phi_{i,n}^2+\mu_i^2(t)}}$, and $\hat{\epsilon}_{i,n}$ indicates the estimation of $\bar{\epsilon}_{i,n}$, which is driven by

$$\dot{\hat{\epsilon}}_{i,n}=\gamma_{i,n}\phi_{i,n}^2\Upsilon_{i,n} \tag{67}$$

with $\gamma_{i,n}>0$, and $\varpi_{i,n}$ and $\tau_{i,n}$ are provided by $\varpi_{i,n}=\psi_{i,n}-\sum_{s=1}^{n-1}\frac{\partial\pi_{i,n-1}}{\partial x_{i,s}}\psi_{i,s}$ and $\tau_{i,n}=\tau_{i,n-1}+\varpi_{i,n}\phi_{i,n}$. Define the Lyapunov function $S_{i,n}=S_{i,n-1}+\frac{1}{2}\big(\phi_{i,n}^2+\frac{1}{\gamma_{i,n}}\tilde{\epsilon}_{i,n}^2\big)$, where $\tilde{\epsilon}_{i,l}=\hat{\epsilon}_{i,l}-\bar{\epsilon}_{i,l}$. Furthermore, $\hat{\theta}_i$ is formulated as follows:

$$\dot{\hat{\theta}}_i=\Pi_i\tau_{i,n}. \tag{68}$$

Thus, we have

$$\dot{S}_{i,n}\leqslant-\sum_{s=1}^{n}k_{i,s}\phi_{i,s}^2+\tilde{\theta}_i^{\top}\Pi_i^{-1}(\Pi_i\tau_{i,n}-\dot{\hat{\theta}}_i)+\left(\sum_{s=2}^{n}\phi_{i,s}\frac{\partial\pi_{i,s-1}}{\partial\hat{\theta}_i}\right)\left(\Pi_i\tau_{i,n}-\dot{\hat{\theta}}_i\right)+\mathcal{T}_n, \tag{69}$$

where $\mathcal{T}_n=\phi_{i,n}\epsilon_{i,n}-\phi_{i,n}^2\Upsilon_{i,n}\hat{\epsilon}_{i,n}+\frac{1}{\gamma_{i,n}}\tilde{\epsilon}_{i,n}\dot{\tilde{\epsilon}}_{i,n}$. In a similar manner to the computation of $\mathcal{T}_1$, we obtain $\mathcal{T}_n\leqslant\bar{\epsilon}_{i,n}\mu_i$. By substituting the above inequality into (69), we obtain

$$\dot{S}_{i,n}\leqslant-\sum_{l=1}^{n}k_{i,s}\phi_{i,s}^2+\tilde{\theta}_i^{\top}\Pi_i^{-1}(\Pi_i\tau_{i,n}-\dot{\hat{\theta}}_i)+\left(\sum_{j=2}^{n}\phi_{i,j}\frac{\partial\pi_{i,j-1}}{\partial\hat{\theta}_i}\right)\left(\Pi_i\tau_{i,n}-\dot{\hat{\theta}}_i\right)+\mu_i\sum_{s=1}^{n}\bar{\epsilon}_{i,l}. \tag{70}$$

According to the controller (66), along with (67) and (68), we present the stability analysis in the following theorem.

**Theorem 3.** Considering high order nonlinear MASs (1)–(4) with Assumptions 1–3. If the developed neural network adaptive controllers (66), augmented with adaptive laws (67) and (68), are designed using the resilient optimal algorithm (11) and the smooth-like trajectory of the decision (40), and the conditions in Theorem 2 are satisfied, then the resilient cooperative objective can be achieved.

*Proof.* Define $S=\sum_{i=1}^{M}S_{i,n}$. It can be deduced from (68) and (70) that $\dot{S}\leqslant-\sum_{i=1}^{M}\sum_{s=1}^{n}k_{i,s}\phi_{i,s}^2+\mu_i\sum_{s=1}^{n}\bar{\epsilon}_{i,l}$. Applying the integration to both sides of the above inequality yields

$$S(t)+\int_0^t\sum_{i=1}^{M}\sum_{s=1}^{n}k_{i,s}\phi_{i,s}^2(s)ds\leqslant S(0)+\int_0^t\sum_{i=1}^{M}\sum_{s=1}^{n}\bar{\epsilon}_{i,l}\mu_i(s)ds. \tag{71}$$

Considering the definition of $\mu_i(t)$, it is evident that $\int_0^{\infty}\mu_i(\tau)d\tau\leqslant\bar{\mu}_i$. Using (71), we can establish that $\lim_{t\to\infty}\int_0^t\phi_{i,l}^2(s)ds$ is bounded. Applying Barbalat's lemma, it can be inferred that $\lim_{t\to\infty}\phi_{i,1}(t)=0$, indicating the asymptotic convergence of the tracking error $y_i-s_i$.

**Remark 6.** In contrast to [37], more general nonlinear MASs are considered in this paper. Moreover, our analysis encompasses the consideration of DoS attacks in this paper. To address this challenge, we introduce a new smooth-like trajectory $s_i$ with the existence of $\dot{s}_i(t),\ddot{s}_i(t),\ldots,s_i^{(n)}(t)$, which employs the Hermite interpolation technique. Subsequently, we apply the backstepping technique to track the optimal value.

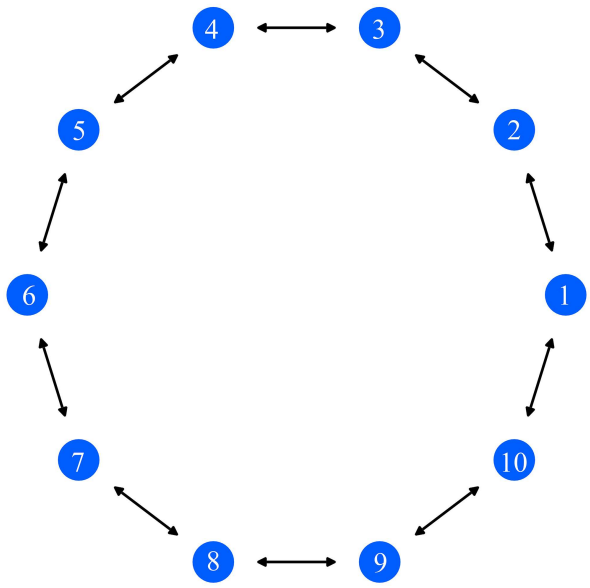


**Figure 1** (Color online) The network topology.

## 4 Simulation

This section presents a simulation example to validate the effectiveness of the proposed method. The simulation is organized into three subsections, i.e., parameter selection, verification of the proposed approach, and algorithm comparison.

### 4.1 Parameter selection

In this subsection, we first consider robotic manipulator systems $\mathcal{J}_i\ddot{q}_i + \mathcal{B}_i\dot{q}_i + \mathcal{M}_i g_i \mathcal{L}_i \sin(q_i(t)) = u_i(t)$, where $q_i$, $\dot{q}_i$, and $\ddot{q}_i$ denote the position, velocity, and acceleration of the $i$-th link. The parameters $g_i$, $\mathcal{L}_i$, $\mathcal{B}_i$, $\mathcal{J}_i$, $\mathcal{M}_i$, and $u_i$ represent the gravitational acceleration, geometry of the link, the dynamics, and the control input, respectively. Let $y_{i,1} = q_i$ and $y_{i,2} = \dot{q}_i$ denote the state variables. Then, the aforementioned MASs can be formulated as $y_i = y_{i,1}$, where $\dot{y}_{i,1} = y_{i,2}$ and $\dot{y}_{i,2} = \frac{1}{\mathcal{J}_i}u_i - \frac{\mathcal{B}_i}{\mathcal{J}_i}x_{i,2}(t) - \frac{\mathcal{M}_i g_i \mathcal{L}_i}{\mathcal{J}_i}\sin(x_{i,1}(t))$.

The network topology is presented in Figure 1. Inspired by [10], the following local cost functions are selected $f_1(y_1) = 0.2\sqrt{y_1^4+3} + 0.7\cos^2(y_1)$, $f_2(y_2) = 2\sin(y_2) - 0.1(y_2^2+2)^{\frac{1}{3}}$, $f_3(y_3) = \frac{0.3y_3^2}{\sqrt{y_3^2+1}}$, $f_4(y_4) = -0.1\sqrt{y_4^4+3} - \sin(y_4)$, $f_5(y_5) = \frac{-0.2y_5^2}{\sqrt{y_5^2+1}} + 2\sin^2(y_5)$, $f_6(y_6) = -0.1\sqrt{y_6^4+3} - \frac{0.1y_6^2}{\sqrt{y_6^2+1}}$, $f_7(y_7) = -\sin(y_7) - 1$, $f_8(y_8) = y_8^2 + 0.3\cos^2(y_8)$, $f_9(y_9) = 2\sin^2(y_9) + 0.2(y_9^2+2)^{\frac{1}{3}}$, $f_{10}(y_{10}) = -0.1(y_{10}^2+2)^{\frac{1}{3}}$. It can be readily confirmed that the global cost function is nonconvex and satisfies the P-Ł condition. According to Theorem 1, the parameters are set as $\alpha = 5$, $\beta = 2.5$. The initial state is chosen as $\boldsymbol{\xi}(0) = [3,\ 2,\ -1,\ -2,\ -2.2,\ 2,\ 3,\ 10,\ -1.6,\ 1.8]$, while $\boldsymbol{\zeta}(0)$ is initialized such that $\sum_{i=1}^{10}\zeta_i(0) = 0$. Throughout the simulation time interval of $[0,\ 20]$ s, several DoS attack periods are introduced, specifically within the intervals $[0.6,\ 1.8)$, $[2.8,\ 4.0)$, $[5.1,\ 6.2)$, and $[7.4,\ 8.8)$ s. The parameters of the aforementioned MASs are specified as $\mathcal{J}_i = 1$, $\mathcal{B}_i = 2$, and $\mathcal{M}_i g_i \mathcal{L}_i = 10 \times i$. The initial values of the adaptive parameters are given by $\hat{\theta}_{1,2}(0) = 12 \times 1_5, \hat{\theta}_{2,2}(0) = 30 \times 1_5, \hat{\theta}_{3,2}(0) = 9 \times 1_5, \hat{\theta}_{4,2}(0) = 4 \times 1_5, \hat{\theta}_{5,2}(0) = 5 \times 1_5, \hat{\theta}_{6,2}(0) = 72 \times 1_5, \hat{\theta}_{7,2}(0) = 105 \times 1_5, \hat{\theta}_{8,2}(0) = 24 \times 1_5, \hat{\theta}_{9,2}(0) = 9 \times 1_5, \hat{\theta}_{10,2}(0) = 10 \times 1_5$, and $\hat{\epsilon}_{i,2}(0) = i$ for all $i = 1, \ldots, 10$. In addition, the time-varying parameters are defined as $\gamma_{i,2} = 10i$, and $\Upsilon_i = 0.1 \times i \times I_5$. The basis function $\psi_{i,2}$ is constructed as $\psi_{i,2} = [\mathrm{e}^{-\frac{(y_{i,2}+2)^2}{4}},\ \mathrm{e}^{-\frac{(y_{i,2}+1)^2}{4}},\ \mathrm{e}^{-\frac{y_{i,2}^2}{4}},\ \mathrm{e}^{-\frac{(y_{i,2}-1)^2}{4}},\ \mathrm{e}^{-\frac{(y_{i,2}-2)^2}{4}}]^\top$.

### 4.2 Verification of the proposed approach

Given the selected parameters and initial conditions, the simulation results are depicted in Figures 2 and 3. As shown in Figure 2(a), the trajectories of $\xi_i$ under the proposed resilient distributed nonconvex optimization algorithm converge to the optimal value, which verifies the effectiveness of the proposed algorithm under DoS attacks. In Figure

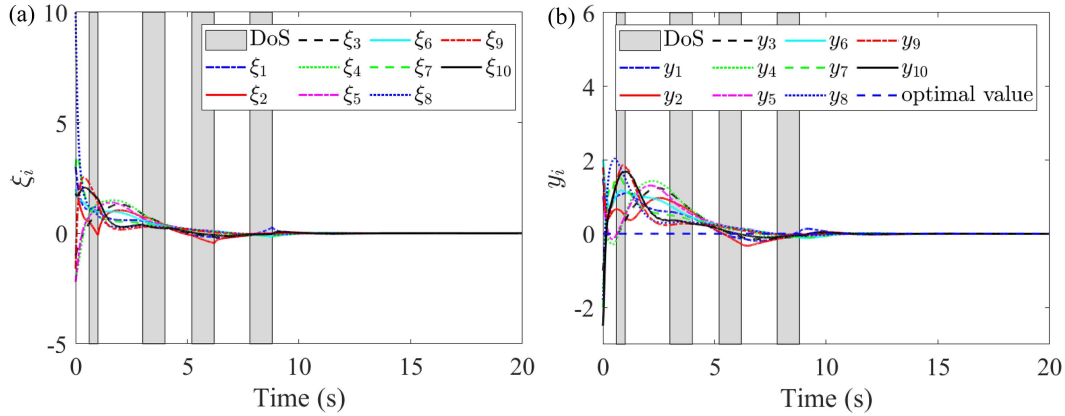


**Figure 2** (Color online) Trajectories of $\xi_i$ (a) and $y_i$ (b) for $i = 1, 2, \ldots, 10$.

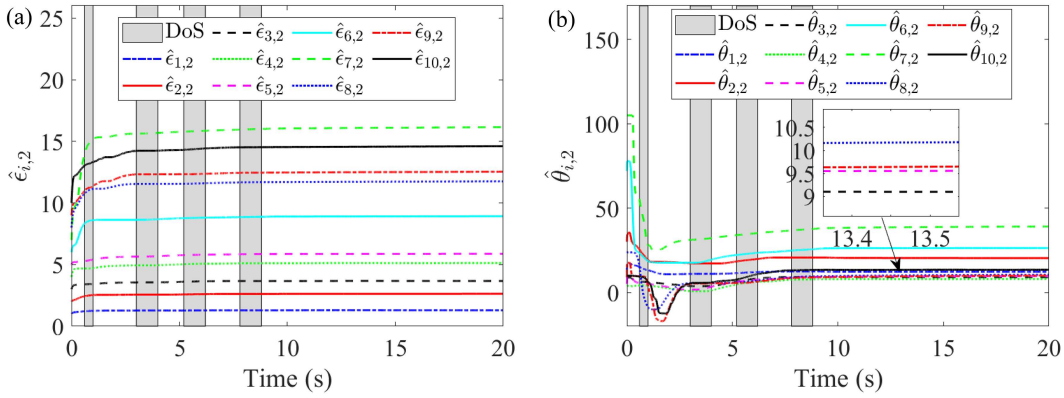


**Figure 3** (Color online) Trajectories of adaptive parameters $\hat{\epsilon}_{i,2}$ (a) and $\hat{\theta}_{i,2}$ (b) for $i = 1, 2, \ldots, 10$.

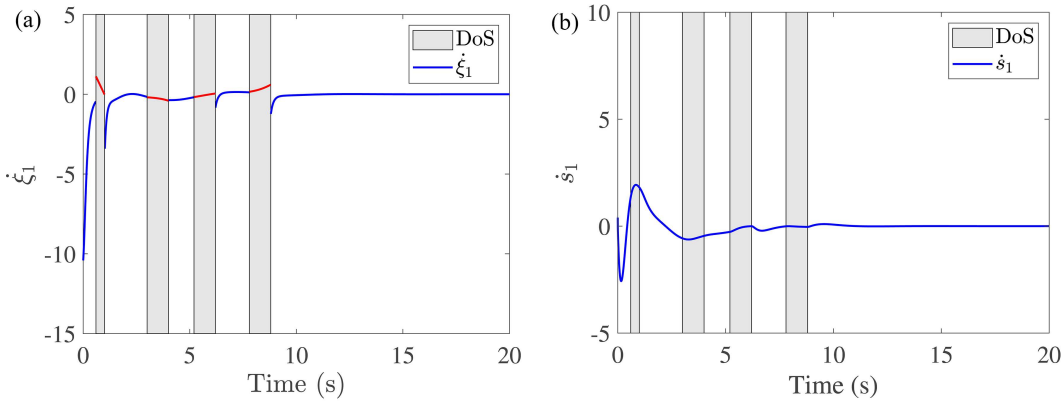


**Figure 4** (Color online) (a) Trajectory of $\dot{\xi}_1$ generated by the method in [10]; (b) trajectory of $\dot{s}_1$ generated by our proposed method.

2(b), the trajectories of $y_i$ converge to the optimal value, which demonstrates the effectiveness of the backstepping-based controller in handling nonlinear dynamics. To further validate the adaptive capability of the proposed approach, the trajectories of the adaptive parameters $\hat{\theta}_{i,2}$ and $\hat{\epsilon}_{i,2}$ are depicted in Figure 3.

### 4.3 Algorithm comparison

To validate the capability of the proposed smooth-like trajectory in preserving high-order smoothness, Figure 4(a) presents the trajectory of $\dot{\xi}_1$ by the methods in [10], where clear discontinuities appear at switching instants. In contrast, Figure 4(b) shows the trajectory of $\dot{s}_1$ generated by our proposed smooth-like trajectory, which exhibits significantly improved smoothness. These results confirm that the second-order derivatives of the signals generated by the proposed smooth-like trajectory remain well-defined even in the presence of DoS-induced disruptions.

## 5 Conclusion

In this study, we have addressed the cooperative resilient nonconvex optimization control problem of nonlinear MASs under DoS attacks. Specifically, a novel cooperative resilient nonconvex optimization control method has been proposed, which integrates a resilient distributed nonconvex optimization algorithm, a smooth-like trajectory, and a neural network backstepping-based controller. The developed algorithm ensures convergence under DoS attacks with the global cost function meeting the P-Łcondition, which relaxes the convexity requirements. Furthermore, a smooth-like trajectory is designed to generate variables with well-defined high-order derivatives, which facilitates the design of an adaptive controller using the backstepping technique. Finally, a simulation example is given to verify the effectiveness of our developed method. Note that the developed method is obtained under the assumption of undirected graphs. A challenging future work is to generalize this result to directed communication graphs among agents or the Nash equilibrium [38].

**Acknowledgements** This work was supported by National Natural Science Foundation of China (Grant Nos. 62303241, 62203233) and National Science Research Start-up Foundation of Nanjing University of Posts and Telecommunications (Grant Nos. NY223174, NY221007, NY222111).

**References**

1 Zhang K P, Xu L, Yi X L, et al. Predefined-time distributed multiobjective optimization for network resource allocation. Sci China Inf Sci, 2023, 66: 170204
2 Nedic A, Ozdaglar A. Distributed subgradient methods for multi-agent optimization. IEEE Trans Automat Contr, 2009, 54: 48–61
3 Wang Y H, Zeng X L, Zhao W X, et al. A zeroth-order algorithm for distributed optimization with stochastic stripe observations. Sci China Inf Sci, 2023, 66: 199202
4 Yang T, Yi X, Wu J, et al. A survey of distributed optimization. Annu Rev Control, 2019, 47: 278–305
5 Johansson B, Rabi M, Johansson M. A randomized incremental subgradient method for distributed optimization in networked systems. SIAM J Optim, 2009, 20: 1157–1170
6 Gong P, Wang Q G. Robust adaptive distributed optimization for heterogeneous unknown second-order nonlinear multiagent systems. Sci China Inf Sci, 2025, 58: 149202
7 Wei J, Mao Z, Hu X, et al. Optimal robust control of robotic arms using particle swarm optimisation: addressing uncertainties and enhancing stability. J Control Decis, 2025, 1–17
8 Cao L, Qin Y, Pan Y, et al. Prescribed performance-based optimal formation control for USVs with position constraints and yaw angle time-varying partial constraints. IEEE Trans Intell Transp Syst, 2025, 26: 4109–4121
9 Boyd S, Xiao L, Mutapcic A. Subgradient Methods. Stanford: Stanford University, 2003
10 Xu L, Yi X, Shi Y, et al. Distributed nonconvex optimization with event-triggered communication. IEEE Trans Automat Contr, 2024, 69: 2745–2752
11 Mokhtari A, Ling Q, Ribeiro A. Network Newton distributed optimization methods. IEEE Trans Signal Process, 2017, 65: 146–161
12 Yi X, Zhang S, Yang T, et al. Linear convergence of first- and zeroth-order primal-dual algorithms for distributed nonconvex optimization. IEEE Trans Automat Contr, 2022, 67: 4194–4201
13 Tang Y, Zhang J, Li N. Distributed zero-order algorithms for nonconvex multiagent optimization. IEEE Trans Control Netw Syst, 2021, 8: 269–281
14 Li Y, Tong S. Bumpless transfer distributed adaptive backstepping control of nonlinear multi-agent systems with circular filtering under DoS attacks. Automatica, 2023, 157: 111250
15 Li Y, Lu G, Li K. Fuzzy adaptive event-triggered consensus control for nonlinear multiagent systems with output constraints and DoS attacks. IEEE Trans Cybern, 2025, 55: 2–13
16 Yang H, Yu Z, Fu M, et al. Resilient consensus control for multiple UAVs with input saturation under DoS attacks. IEEE Trans Cybern, 2025, 55: 1159–1171
17 Mo Y, Sinopoli B. Secure control against replay attacks. Allerton Conf Commun Control Comput, 2009: 911–918
18 Sun Q, Chen J C, Shi Y. Event-triggered robust MPC of nonlinear cyber-physical systems against DoS attacks. Sci China Inf Sci, 2022, 65: 110202
19 Pasqualetti F, Dorfler F, Bullo F. Attack detection and identification in cyber-physical systems. IEEE Trans Automat Contr, 2013, 58: 2715–2729
20 Liu S, Jiang B, Mao Z, et al. Neural-network-based adaptive fault-tolerant cooperative control of heterogeneous multiagent systems with multiple faults and DoS attacks. IEEE Trans Neural Netw Learn Syst, 2024, 35: 6273–6285
21 Badie K, Dami L, El Abbadi R, et al. Cyber-attack detection in networked control systems: application to Markovian jump systems. J Control Decis, 2025. doi:10.1080/23307706.2025.2527399
22 Zhao C, He J, Wang Q G. Resilient distributed optimization algorithm against adversarial attacks. IEEE Trans Automat Contr, 2020, 65: 4308–4315
23 Wang X F, Teel A R, Liu K Z, et al. Stability analysis of distributed convex optimization under persistent attacks: a hybrid systems approach. Automatica, 2020, 111: 108607
24 Li Z, Wu Z, Li Z, et al. Distributed optimal coordination for heterogeneous linear multiagent systems with event-triggered mechanisms. IEEE Trans Automat Contr, 2020, 65: 1763–1770
25 Yi X, Yao L, Yang T, et al. Distributed optimization for second-order multi-agent systems with dynamic event-triggered communication. IEEE Conf Decis Control, 2018: 3397–3402
26 Gharesifard B, Cortes J. Distributed continuous-time convex optimization on weight-balanced digraphs. IEEE Trans Automat Contr, 2014, 59: 781–786
27 Liu T, Qin Z, Hong Y, et al. Distributed optimization of nonlinear multiagent systems: a small-gain approach. IEEE Trans Automat Contr, 2021, 67: 676–691
28 Qin Z, Liu T, Jiang Z P. Adaptive backstepping for distributed optimization. Automatica, 2022, 141: 110304
29 Qin Z, Liu T, Liu T, et al. Distributed feedback optimization of nonlinear uncertain systems subject to inequality constraints. IEEE Trans Automat Contr, 2023, 69: 3989–3996
30 Hou Z G, Cheng L, Tan M. Decentralized robust adaptive control for the multiagent system consensus problem using neural networks. IEEE Trans Syst Man Cybern B, 2009, 39: 636–647
31 Deng C, Wen C. Distributed resilient observer-based fault-tolerant control for heterogeneous multiagent systems under actuator faults and DoS attacks. IEEE Trans Control Netw Syst, 2020, 7: 1308–1318
32 Nesterov Y. Lectures on Convex Optimization. New York: Springer, 2018

33 Feng Z, Hu G. Attack-resilient distributed convex optimization of cyber-physical systems against malicious cyber-attacks over random digraphs. IEEE Int Things J, 2023, 10: 458–472
34 Kokotovic P V. The joy of feedback: nonlinear and adaptive. IEEE Control Syst Mag, 1992, 12: 7–17
35 Ren H, Cao L, Ma H, et al. Dynamic event-triggered-based fuzzy adaptive pinning control for multiagent systems with output saturation. IEEE Trans Fuzzy Syst, 2025, 33: 1277–1286
36 Zhao Z, Wang T, Yu J, et al. Bilateral cooperative control of nonlinear multiagent systems with state and output quantification. IEEE Trans Cybern, 2025, 55: 2949–2957
37 Ma H J, Yang G H, Chen T. Event-triggered optimal dynamic formation of heterogeneous affine nonlinear multi-agent systems. IEEE Trans Automat Contr, 2021, 66: 497–512
38 Xu W, Wang T, Qiu J, et al. Multiplayer pursuit-evasion games with distributed Nash equilibrium solution. IEEE Trans Netw Sci Eng, 2025, 12: 4152–4163