Distributed neurodynamic approaches to nonsmooth optimization problems with inequality and set constraints

Linhua Luan · Xingnan Wen · Sitian Qin

Received: 3 July 2021 / Accepted: 3 May 2022 / Published online: 30 May 2022
© The Author(s) 2022

Abstract
In this paper, neurodynamic approaches are proposed for solving nonsmooth distributed optimization problems under inequality and set constraints, that is to find the solution that minimizes the sum of local cost functions. A continuous-time neurodynamic approach is designed and its state solution exists globally and converges to an optimal solution of the corresponding distributed optimization problem. Then, a neurodynamic approach with event-triggered mechanism is considered for the purpose of saving communication costs, and then, the convergence and its Zeno-free property are proved. Moreover, to realize the practical application of the neurodynamic approach, a discrete-time neurodynamic approach is proposed to solve nonsmooth distributed optimization problems under inequality and set constraints. It is rigorously proved that the iterative sequence generated by the discrete-time neurodynamic approach converges to the optimal solution set of the distributed optimization problem. Finally, numerical examples are solved to demonstrate the effectiveness of the proposed neurodynamic approaches, and the neurodynamic approach is further applied to solve the ill-conditioned Least Absolute Deviation problem and the load sharing optimization problem.

Keywords Distributed convex optimization · Penalty method · Multi-agent system · Projection operator · Event-triggered mechanism

Introduction
Distributed optimization problems over multi-agent systems have received much attention and the main principle of solving distributed optimization problems is to optimize the global cost function through the cooperation of interrelated agents [31,41,45]. The cost function of a distributed optimization problem is generally the summation of local cost functions which is only acquired by an individual agent. In particular, to protect the privacy of communication, agents in the multi-agent system only share particular information with their neighbors. Distributed optimization problems can be found in fields of traffic balance [41], multi-objective optimization [45], and collaborative control, such as the realization of formation maintenance between multiple vehicles [31]. Considering the limitations of communication bandwidth and communication range in many applications, it becomes necessary to study distributed neurodynamic approaches that require only partial information exchange between neighbors. In fact, problems in the engineering field can usually be abstracted into mathematical models and then solved using corresponding neurodynamic approaches, such as [2,8,11,13,27,35].

In recent years, kinds of neurodynamic approaches for solving distributed optimization problems have been published (see [11,13–15,27–29,36,46]). By means of collaborative control between agents, these neurodynamic approaches can not only protect privacy, but also solve the server overload issue. For unconstrained distributed optimization problems, authors in [36] constructed a neurodynamic approach under the assumption that the optimal solution set is a compact set. However, authors in [11] proposed a neurodynamic approach based on a strongly connected and directed communication graph to weaken this assumption. Actually, there are few unconstrained optimization problems in the actual
situation, so scholars pay more attention to distributed optimization problems with constraints [13, 28, 29, 46]. Authors in [46] constructed a distributed gradient neurodynamic approach to solve the distributed optimization problem over a multi-agent system where each agent has its own private constraint function, while authors in [29] proposed a continuous-time projection neurodynamic approach sharing the same constraint set among agents in the multi-agent system. It should be noted that the neurodynamic approach in [46] requires the local cost functions which are twice continuously differentiable. To relax this differentiability hypothesis, authors in [27] proposed a subgradient neurodynamic approach for nonsmooth distributed optimization problem, which was later developed into a subgradient projection neurodynamic approach in [28] to resolve distributed optimization problems with consensus constraints. Inspired by the works above, authors in [13] proposed a subgradient neurodynamic approach for solving the nonsmooth distributed optimization problem. In addition, authors in [23] presented a primal-dual projection neurodynamic approach for nonsmooth distributed optimization problems under local bounded constraints. Because instant messaging is an ideal state, authors in [21] considered the time delay in information exchange and designed a subgradient projection neurodynamic approach. Furthermore, by combining primal-dual methods for searching saddle points with projection operators for meeting set constraints, a continuous-time neurodynamic approach for nonsmooth constrained convex optimization was designed in [49].

All neurodynamic approaches mentioned above require that the communication process of the multi-agent system takes place at each time. Therefore, to avoid the high-energy consumption caused by frequent communication and in keeping with the fact that each node is usually equipped with a limited amount of energy, a large number of neurodynamic approaches with event-triggered mechanisms have emerged (see [7, 16, 25, 33, 51]). Among them, authors in [51] proposed an event-triggered neurodynamic approach to solve distributed optimization problems with equality constraints when the local objective functions are quadratic functions, and authors in [33] designed a neurodynamic approach for the second-order multi-agent systems with event-triggered and time-triggered communication. Furthermore, to consider the more general case, authors in [25] designed a gradient neurodynamic approach with event-triggered mechanism to solve distributed convex problems under set constraints.

As we all know, whether in computers or in practical applications, discrete-time neurodynamic approaches are more practical. For resolving distributed optimization problems with different constraints, a variety of discrete-time neurodynamic approaches have been proposed in public. With the help of properties of the projection operator, authors in [27] proposed a discrete-time neurodynamic approach for constrained nonsmooth optimization problems based on the parallel and distributed computing general framework proposed by [34]. On the basis of [27], authors in [21] proposed a subgradient neurodynamic approach for the same nonsmooth optimization problem. However, the neurodynamic approach in [21] required that the constraint set is compact. To weaken this assumption, authors in [37] introduced a discrete-time neurodynamic approach under switching topologies for distributed optimization problems with strongly convex local cost functions. Recently, authors in [24] have presented a discrete-time neurodynamic approach with fixed step and analyzed its convergence rate. Later, authors in [38] proposed a subgradient neurodynamic approach with various step sizes. Authors in [50] further extended smooth local cost functions to the summation of smooth convex functions and nonsmooth $L_1$-norm functions.

However, most of the above-mentioned neurodynamic approaches require some inevitably strict assumptions. For example, the local cost functions of the considered distributed optimization problem are smooth, or the structure of the constraint set is simple or even has no constraints. To overcome these limitations, we propose neurodynamic approaches to solve the nonsmooth distributed optimization problem under inequality and set constraints. Here are detailed contributions.

1. Unlike the existing neurodynamic approaches where the Laplacian matrixes of communication graph are used to ensure consensus constraints (see [13, 17, 46]), one of the highlights of this paper is using the exact penalty method to deal with uniform constraints, which reduces the dimension of solution space.

2. The continuous-time neurodynamic approach proposed herein has better convergence property than approaches in [6, 18] because it can ensure that state solutions converge to an optimal point of the distributed optimization problem rather than the optimal solution set. Furthermore, the proposed continuous-time neurodynamic approach can solve the nonsmooth distributed optimization problem under inequality and set constraints, and it can be expended to solve distributed optimization problems in [11, 12, 22, 42].

3. Compared with the neurodynamic approaches under continuous communication in [11, 13, 29, 46] and so on, the advantages of the event-triggered neurodynamic approach designed in this paper are saving the communication energy consumption between nodes and reducing the number of controller updates. Thus, avoiding unnecessary consumption of network resources and improving the utilization rate of bandwidth. Moreover, the event-triggered neurodynamic approach can solve distributed optimization problems under inequality and set con-
straints, which has the ability to solve more general problems than [25,33,51].

4. A discrete-time variable step-size neurodynamic approach is designed for the nonsmooth convex distributed optimization problems with inequality and set constraints. As far as we know, existing discrete-time neurodynamic approaches in [32,45,50] can only be applied to cases that have differentiable local cost functions or affine inequality constraints, so the proposed discrete-time neurodynamic approach herein can solve more general problems. In addition, discrete-time neurodynamic approaches are easier to implement and apply than continuous-time neurodynamic approaches.

This article is organized as follows. “Preliminaries” lists some notations and basic knowledge about graph theory, convex analysis, etc. In “Problem description and equivalent form”, the distributed optimization problem with inequality and set constraints is transformed into a problem with set constraints equivalently. In “Main results”, a continuous-time distributed neurodynamic approach and an event-triggered neurodynamic approach are discussed; meanwhile, the convergence of two neurodynamic approaches is proved, respectively. Moreover, a discrete-time neurodynamic approach is presented and its convergence is analyzed. “Simulations and applications” illustrates the effectiveness and performance of neurodynamic approaches through numerical examples. The last section summarizes the thesis and looks forward to the research direction in the future.

Preliminaries

First of all, here are some terminologies and symbols that will be utilized. \( \mathbb{R}^n \) is the set of \( n \)-dimension real vectors, \( \mathbb{R}^{m\times n} \) is the set of \( m \times n \) real matrices, and \( \mathbb{N} \) is the set of natural numbers. \( x^T \) and \( \|x\| = \sqrt{x^T x} \) denote the transpose and norm of \( x \in \mathbb{R}^n \), respectively. The open ball centered on \( x_0 \in \mathbb{R}^n \) is described by \( B(x_0, r) = \{ x \in \mathbb{R}^n : \|x - x_0\| < r \} \) with \( r > 0 \). Furthermore, \( \text{col}(x, y) = (x^T, y^T)^T \in \mathbb{R}^{n+m} \) for \( x \in \mathbb{R}^n \) and \( y \in \mathbb{R}^m \). \( I_n \) is an \( n \)-dimensional vector where all the components are ones, and \( 0_n \) is an \( n \)-dimensional vector where all the components are zero. The sign \( \otimes \) stands for the Kronecker product. For two sets \( A, B \subseteq \mathbb{R}^n \), \( \text{int}(A) \) and \( \text{bd}(A) \) mean the interior and the boundary of set \( A, A + B = \{ x + y : x \in A, y \in B \} \) and \( A + x_0 = \{ x + x_0 : x \in A \} \).

Next, some basic definitions of graph theory are as follows. Let \( G = (\mathcal{V}, \mathcal{E}) \) be a communication topology between several nodes in a multi-agent system consisting of \( N \) agents, where \( \mathcal{V} = \{1, 2, \ldots, N\} \) and \( \mathcal{E} \subseteq \mathcal{V} \times \mathcal{V} \) are node set and edge set, respectively. If \( (i, j) \in \mathcal{E} \), Node \( j \) is a neighbor of node \( i \). If \( A = (a_{ij})_{N \times N} \) satisfies that \( a_{ii} \geq 0 \), and \( a_{ij} > 0 \) when \( (i, j) \in \mathcal{E} \), \( A \) is a weighted adjacency matrix of \( G \). If there are distinct nodes \( i_j \in \mathcal{V} \), such that \( \mathcal{L} = \{(i, i_1), (i_1, i_2), \ldots, (i_k, j)\} \), then \( i \) and \( j \) are connected and \( \mathcal{L} \) is a path between nodes \( i \) and \( j \). Furthermore, the communication topology \( G \) is called undirected, if its weighted adjacency matrix is symmetric. And an undirected graph \( G \) is called connected, if there exists a path between any pair of nodes.

Subsequently, several necessary definitions and propositions of convex analysis and projection operators are given.

**Definition 1** [10] If \( f(x) \) is convex on \( \mathbb{R}^n \), then the subdifferential of \( f(x) \) at \( x_0 \in \mathbb{R}^n \) is

\[
\partial f(x_0) = \{ \xi \in \mathbb{R}^n : f(x) - f(x_0) \geq \langle \xi, x - x_0 \rangle, \forall x \in \mathbb{R}^n \}.
\]

**Proposition 1** [10] If \( f(x) \) is convex on \( \mathbb{R}^n \), then \( \partial f(x) \) is nonempty, convex, compact, and upper semicontinuous.

**Definition 2** [10] The differentiable function \( f(x) \) is \( u \)-strongly convex on \( \mathbb{R}^n \) \((u > 0)\), if for any \( x, y \in \mathbb{R}^n \)

\[
f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{u}{2} \|y - x\|^2.
\]

Let \( C \subseteq \mathbb{R}^n \) be a nonempty closed and convex set, then the projection operator on \( C \) is a function \( P_C : \mathbb{R}^n \rightarrow C \), which is defined by

\[
P_C(x) = \arg \min_{y \in C} \|x - y\|.
\]

**Proposition 2** [27] For the projection operator \( P_C \), it follows that:

1. \( \langle P_C(x) - x, y - P_C(x) \rangle \geq 0 \), \( \forall x \in \mathbb{R}^n \) and \( \forall y \in C \);
2. \( \|P_C(x) - P_C(y)\| \leq \|x - y\|, \forall x, y \in \mathbb{R}^n \). 

**Definition 3** Suppose that \( C \subseteq \mathbb{R}^n \) is a nonempty closed convex set. The normal cone of \( C \) at \( x \), denoted by \( N_C(x) \), is defined as

\[
N_C(x) = \{ v \in \mathbb{R}^n : \langle v, y - x \rangle \leq 0, \forall y \in C \}.
\]

**Proposition 3** [10] Suppose that \( C, C_1, C_2 \subseteq \mathbb{R}^n \) are nonempty closed convex sets.

1. If \( f(x) \) is local Lipschitz and attains a minimum over \( C \) at \( x \in C \), then \( 0 \in \partial f(x) + N_C(x) \);
2. If \( 0 \in \text{int}(C_1 - C_2) \), then for any \( x \in C_1 \cap C_2 \), \( N_{C_1\cap C_2}(x) = N_{C_1}(x) + N_{C_2}(x) \).
Problem description and equivalent form

In this section, we consider the following constrained optimization problem:

\[
\begin{align*}
\min & \quad f(x) = \sum_{i=1}^{N} f_i(x) \\
\text{s.t.} & \quad g_i(x) \leq 0, \quad i = 1, 2, \ldots, N \\
& \quad x \in \Omega,
\end{align*}
\]  

(1)

where \( x \in \mathbb{R}^n \) is the decision vector, \( f_i(x) : \mathbb{R}^n \to \mathbb{R} \) is convex but not necessarily smooth, and constraint functions \( g_i(x) : \mathbb{R}^n \to \mathbb{R}^{m_i} \) \( (i = 1, 2, \ldots, N) \) are convex. \( \Omega \subseteq \mathbb{R}^n \) is a bounded closed convex set. Without loss of generality, the optimal solution set of optimization problem (1) is nonempty.

Remark 1 Optimization problems have been discussed by [11–13, 17, 22, 42, 46] and so on, since they can be applied in various engineering and control areas. Actually, the optimization problem discussed here is under inequality and set constraints, so optimization problem (1) has a wider range of applications than problems discussed in [12, 22] and [42].

Let \( m = \sum_{i=1}^{N} m_i \), and note \( \bar{g}(x) = \text{col}(g_1(x), g_2(x), \ldots, g_N(x)) : \mathbb{R}^n \to \mathbb{R}^m \)

\[\mathcal{S} = \{ x \in \mathbb{R}^n : \bar{g}(x) \leq 0 \},\]

where \( \bar{g}(x) \leq 0 \) means each of its components \( \bar{g}_i(x) \leq 0, \quad i \in \{1, 2, \ldots, m\} \). To simplify the constraints in optimization problem (1), some necessary assumptions are given below.

Assumption 1

1. There exists a point \( \hat{x} \in \mathbb{R}^n \), such that \( \hat{x} \in \text{int}(\mathcal{S}) \cap \Omega \).
2. The set \( \mathcal{S} = \{ x \in \mathbb{R}^n : \bar{g}(x) \leq 0 \} \) is bounded, i.e., there exists \( M > 0 \), such that \( \mathcal{S} \subseteq B(\hat{x}, M) \).

Assumption 1 is actually common preconditions and have been used in [26, 48]. It is easy to see that the feasible region \( \mathcal{S} \cap \Omega \) of optimization problem (1) is a bounded set, which implies that there exists a positive number \( L \), such that

\[\|\xi\| \leq L \quad \text{and} \quad \|\eta\| \leq L, \quad (2)\]

for any \( \xi \in \partial f(x), \eta \in \partial \bar{g}(x), x \in \mathcal{S} \cap \Omega \). Besides, define a function

\[D(x) = \sum_{l=1}^{m} \max\{0, \bar{g}_l(x)\} = \sum_{i=1}^{N} \sum_{j=1}^{m_i} \max\{0, g_{ij}(x)\},\]

where \( \bar{g}_l(x) \) is the \( l \)th component of \( \bar{g}(x) \) and \( g_{ij}(x) \) means the \( j \)th component of \( g_i(x) \). Due to the convexity of \( g_i(x) \), it can be concluded that \( D(x) \) is convex. Specially, the closed form for \( \partial D(x) \) can be calculated [43] that

\[
\partial D(x) = \begin{cases}
\sum_{l \in I^0(x)} [0, 1] \partial \bar{g}_l(x), & x \in \text{bd}(\mathcal{S}), \\
[0], & x \in \text{int}(\mathcal{S}), \\
\sum_{l \in I^0(x)} [0, 1] \partial \bar{g}_l(x) + \sum_{l \in I^+(x)} \partial \bar{g}_l(x), & x \notin \mathcal{S},
\end{cases}
\]

(3)

where \( I = \{1, 2, \ldots, m\} \), \( I^0(x) = \{ l \in I : \bar{g}_l(x) = 0 \} \) and \( I^+(x) = \{ l \in I : \bar{g}_l(x) > 0 \} \). Let

\[\hat{g} = \min\{-\bar{g}_l(\hat{x}) : i \in I\}.
\]

(4)

Inspired by the exact penalty method mentioned in [20] and [44], we can set up the equivalent transformation of optimization problem (1) as follows:

Theorem 1 Under Assumption 1, \( x^* \in \mathbb{R}^n \) is an optimal solution of optimization problem (1) if and only if \( x^* \) is an optimal solution of the following optimization problem:

\[
\begin{align*}
\min & \quad f(x) + \sigma D(x) \\
\text{s.t.} & \quad x \in \Omega,
\end{align*}
\]

(5)

where the penalty parameter \( \sigma > LM/\hat{g} \), and \( M, \hat{g} \) are from Assumption 1, (2) and (4) respectively.

Proof Let \( x^* \) be an optimal solution of optimization problem (1), then due to Proposition 3, it derives that there exists \( \xi^* \in \partial f(x^*), \eta \in N_\mathcal{S}(x^*), \) such that \( -\xi^* \in N_\mathcal{S}(x^*) \). Based on Assumption 1, \( N_{\mathcal{S} \cap \Omega}(x^*) = N_\mathcal{S}(x^*) + N_\Omega(x^*) \). Next, the proof will be divided into cases as follows.

Case 1: \( x^* \in \text{int}(\mathcal{S}) \cap \Omega \). If this case holds, then \( N_\mathcal{S}(x^*) = \{0\} \). Hence, by the definition of \( \partial D(x) \) in (3), it has

\[-\xi^* \in N_\Omega(x^*) = \sigma \partial D(x^*) + N_\Omega(x^*), \]

(6)

which means that \( 0 \in \partial f(x^*) + \sigma \partial D(x^*) + N_\Omega(x^*) \). It is also known by convexity that \( x^* \) is an optimal solution of optimization problem (5).

Case 2: \( x^* \in \text{bd}(\mathcal{S}) \cap \Omega \). In this case, we have

\[\partial D(x^*) = \sum_{l \in I^0(x^*)} [0, 1] \partial \bar{g}_l(x^*).\]

In view of \( N_\mathcal{S}(x^*) = \bigcup_{\theta \geq 0} \partial \theta D(x^*) \) [10], it obtains that \( N_\mathcal{S}(x^*) = \sum_{l \in I^0(x^*)} [0, +\infty] \partial \bar{g}_l(x^*). \)

\[-\xi^* \in \sum_{l \in I^0(x^*)} [0, +\infty] \partial \bar{g}_l(x^*) + N_\Omega(x^*). \]

(7)
Since $x^* \in \mathrm{bd}(\mathcal{X}) \cap \Omega$, $\|x^*\| \leq L$ for any $\xi^* \in \partial f(x^*)$.

And from (7), there are $\sigma_1 \in [0, +\infty)$, $\eta^* \in \partial \tilde{g}_i(x^*)$ and $\gamma^* \in N_\Omega(x^*)$, such that

$$\xi^* = \sum_{l \in \mathcal{P}(x^*)} \sigma_l \eta^*_l + \gamma^*.$$ 

We claim that for any $l \in I^0(x^*)$, $\sigma_l \leq LM/\tilde{g}$. If not, there is at least one $\tilde{l} \in I^0(x^*)$, such that $\sigma_{\tilde{l}} > LM/\tilde{g}$. Thus, by the convexity of $\tilde{g}_i(x)$ and $\gamma^* \in N_\Omega(x^*)$, it has

$$\langle -\xi^*, x^* - \hat{x} \rangle = \sum_{l \in \mathcal{P}(x^*)} \sigma_l \langle \eta^*_l, x^* - \hat{x} \rangle + \langle \gamma^*, x^* - \hat{x} \rangle \geq \sum_{l \in \mathcal{P}(x^*)} \sigma_l (\tilde{g}_l(x^*) - \tilde{g}_l(\hat{x})) \geq \sigma_{\tilde{l}} \tilde{g} > LM,$$

which implies that $L < \|\xi^*\|$ by Assumption 1. Obviously, it is contradict with $\|\xi^*\| \leq L$. Therefore, when $\sigma \geq LM/\tilde{g}$,

$$\sum_{l \in \mathcal{P}(x^*)} \sigma_l \eta^*_l + \gamma^* \in \sigma \partial D(x^*) + N_\Omega(x^*).$$

Therefore, $0 \in \partial f(x^*) + \sigma \partial D(x^*) + N_\Omega(x^*)$ for all $\sigma \geq LM/\tilde{g}$. That is, $x^*$ is an optimal solution of optimization problem (5), because $f(x) + \sigma D(x)$ is a convex function.

On the other hand, if $x^*$ is an optimal solution of optimization problem (5), we are going to show that $x^*$ is an optimal solution of optimization problem (1). Without loss of generality, assume that $\tilde{x}$ is an optimal solution of (1). Let $\sigma_0 = LM/\tilde{g}$, then similar to (6)--(8), and it turns out that $\tilde{x}$ is also an optimal solution of

$$\text{min} \quad f(x) + \sigma_0 D(x)$$
$$\text{s.t.} \quad x \in \Omega.$$ (9)

Therefore, since $x^* \in \Omega$, it has

$$f(\tilde{x}) + \sigma_0 D(\tilde{x}) \leq f(x^*) + \sigma_0 D(x^*).$$ (10)

In addition, $x^*$ is an optimal solution of optimization problem (5), so

$$f(x^*) + \sigma D(x^*) \leq f(\tilde{x}) + \sigma D(\tilde{x});$$ (11)

when $\sigma > \sigma_0$. By combining (10) and (11), we get that

$$(\sigma - \sigma_0) D(x^*) \leq (\sigma - \sigma_0) D(\tilde{x}).$$

From the definition of $D(x)$, it holds that $x \in \mathcal{X}$ if and only if $D(x) = 0$. Thereby, it is clear that $D(x^*) = 0$ if $D(\tilde{x}) = 0$, which means $x^* \in \mathcal{X}$. As a result, $f(x^*) = f(x^*) + \sigma D(x^*) = \min_{x \in \mathcal{X} \cap \Omega} \{f(x) + \sigma D(x)\} \leq f(x)$, $\forall x \in \mathcal{X} \cap \Omega$. Thus, $x^*$ is an optimal solution of optimization problem (1).

Now, we consider a multi-agent system composed of $N$ agents, whose communication topology is $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\mathcal{V} = \{1, 2, \ldots, N\}$. For $i \in \mathcal{V}$, let $x_i$ be the decision variable of agent $i$ and $N_i = \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$ be the set of neighbors of agent $i$, then the following lemma is essential to the equivalent transformation of optimization problem (5).

**Assumption 2** The communication graph between agents is undirected and connected.

**Lemma 1** [20] The optimization problem (5) can be written as a distributed optimization problem with communication graph $\mathcal{G}$

$$\text{min} \quad p(x) \triangleq \sum_{i=1}^N f_i(x_i) + \sigma \sum_{i=1}^N D_i(x_i)$$
$$\text{s.t.} \quad x \in \Omega,$$

where $x = \text{col}\{x_1, x_2, \ldots, x_N\} \in \mathbb{R}^{N_n}$, $\Omega = \Omega \times \Omega \times \ldots \times \Omega$ and $D_i(x_i) = \sum_{j=1}^{n_i} \max\{0, g_{ij}(x_i)\}$.

**Remark 2** The problem (12) is a typical distributed optimization problem, which aims to find the optimal solution to global cost function through cooperation between agents. Finding the solution to distributed optimization problems has aroused great interest among scholars due to its privacy protection characteristics, intelligence, and flexibility such as [13, 17, 46]. The above papers ensure that consensus constraints hold by introducing the Laplacian matrices of communication graphs, but we equivalently transform the optimization problems with consensus constraints by the exact penalty method shown as the next theorem, which reduces the dimension of solution space.

Under Assumption 1, it is obvious from (2) that there is $L > 0$, such that for any $y \in \partial h(x)$, $\|y\| \leq (1 + \sigma) L$.

Therefore, based on the following theorem, neurodynamic approaches for solving the distributed optimization problem are proposed in the next section.

**Theorem 2** [39] Under Assumptions 1 and 2 as well as $\sigma > \max\{LM/\tilde{g}, NL + \sqrt{NL}(NL + 2i)\}$, $x^*$ is an optimal solution of optimization problem (1) if and only if $x^* = \text{col}\{x^*, x^*, \ldots, x^*\}$ is an optimal solution of
\[
\begin{aligned}
\min \quad & h(x) = \sum_{i=1}^{N} f_i(x_i) + \sigma \sum_{i=1}^{N} D_i(x_i) \\
& \quad + \frac{\sigma^2}{2} \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \|x_i - x_j\| \\
\text{s.t.} \quad & x \in \Omega.
\end{aligned}
\] (13)

**Main results**

In this section, a continuous-time neurodynamic approach, an event-triggered neurodynamic approach, and a discrete-time neurodynamic approach are proposed to solve the distributed optimization problem with inequality and set constraints.

**Continuous-time neurodynamic approach**

In this subsection, we propose a projection subgradient neurodynamic approach for nonsmooth convex distributed optimization problem (13) as follows:

\[
\dot{x}(t) \in -x(t) + \mathcal{P}_\Omega \left( x(t) - \partial h(x(t)) \right),
\] (14)

where \( h(x) \) is defined in (13). Noting that

\[
\partial f(x) = \text{col}\{\partial f_1(x_1), \partial f_2(x_2), \ldots, \partial f_N(x_N)\},
\]
\[
\partial D(x) = \text{col}\{\partial D_1(x_1), \partial D_2(x_2), \ldots, \partial D_N(x_N)\},
\]
\[
\mathcal{P}_\Omega(x) = \text{col}\{\mathcal{P}_\Omega(x_1), \mathcal{P}_\Omega(x_2), \ldots, \mathcal{P}_\Omega(x_N)\}.
\]

Then, the neurodynamic approach for agent \( i \in \mathcal{V} \) can be rewritten as

\[
\dot{x}_i(t) \in -x_i(t) + \mathcal{P}_\Omega \left( x_i(t) - \partial f_i(x_i(t)) - \sigma \partial D_i(x_i(t)) \right) - \sigma^2 \sum_{j \in \mathcal{N}_i} \partial \|x_i(t) - x_j(t)\|.
\] (15)

**Remark 3** In detail, the projection operator \( \mathcal{P}_\Omega \) in (15) is to ensure that the state solution stays at \( \Omega \), the term \(-\partial f_i(x_i(t)) - \partial D_i(x_i(t))\) ensures the state solution converges to an optimal solution of neurodynamic approach (15), and \(-\sum_{j \in \mathcal{N}_i} \partial \|x_i(t) - x_j(t)\|\) makes the state solution reach consensus. To explain neurodynamic approach (15) more clearly, its circuit diagram is shown in Fig. 1.

In fact, the subgradients of \( f_i \) and \( g_i \) (\( i \in \mathcal{V} \)) are represented by piecewise functions.

Actually, the combination of projection method and subgradient method is very common when solving distributed optimization problems, such as [24,30,48,49].

**Theorem 3** For any initial value \( x_0 \in \Omega \), there exists a local solution \( x(t) \) of neurodynamic approach (14) on \([0, T]\), \( T > 0 \) and

\[
x(t) \in \Omega, \ t \in [0, T)
\]

**Proof** According to Propositions 1, 2, and Assumption 1 in this paper, it can be concluded that \( h(x) \) is bounded from below on \( \mathbb{R}^{Nn} \) and Lipschitz on any bounded subset of \( \Omega \). Thus, by applying the Theorem 5.2 in [5], there exists a local solution \( x(t) \) of neurodynamic approach (14) on \([0, T]\) for any \( x_0 \in \Omega \), where \([0, T] \) is the maximum existence interval of \( x(t) \).

By the definition of differential inclusion, there is \( y(t) \) in \( \partial h(x(t)) \) for a.e. \( t \in [0, T] \), such that

\[
\dot{x}(t) = -x(t) + \mathcal{P}_\Omega \left( x(t) - y(t) \right),
\]

where \( h(x) \) is defined in distributed optimization problem (13). Let \( \beta(t) = \mathcal{P}_\Omega(x(t) - y(t)) \), then \( \beta(t) \in \mathcal{P}_\Omega(x(t) - \partial h(x(t))) \) for a.e. \( t \in [0, T] \) and

\[
\dot{x}(t) + x(t) = \beta(t), \text{ a.e. } t \in [0, T].
\] (16)

Inspired by the proof of Lemma 2.4 in [5], through integrating (16) on \([0, t], t \in [0, T] \), it has

\[
x(t) = e^{-t}x(0) + e^{-t} \int_0^t \beta(s) e^s ds = e^{-t}x(0) + (1 - e^{-t}) \int_0^t \beta(s) \frac{e^s}{e^t - 1} ds.
\] (17)

Since \( \int_0^t \frac{e^s}{e^t - 1} ds = 1 \) and \( x_0 \in \Omega \), it is easy to deduce that \( x(t) \in \Omega, t \in [0, T] \), because \( \Omega \) is convex. \( \square \)

**Theorem 4** Under Assumptions 1 and 2, if \( \sigma > \max\{LM/\hat{L}, NL + \sqrt{NL(NL + 2)}\} \), the state solution \( x(t) \) of neurodynamic neurodynamic approach (14) with initial value \( x(0) = x_0 \) exists globally and converges to an optimal solution of distributed optimization problem (13), that is, \( \lim_{t \to \infty} x(t) = x^* = \text{col}(x_1^*, x_2^*, \ldots, x_N^*) \), where \( x_i^* = x_j^* \) (\( i, j \in \mathcal{V} \)) are the optimal solution to optimization problem (1).

**Proof** Let \( x^* = \text{col}(x_1^*, x_2^*, \ldots, x_N^*) \) be an optimal solution of distributed optimization problem (13), and \( x_i(t) \) (\( i \in \mathcal{V} \)) be the state solution from initial value \( x_0(t) \) of neurodynamic approach (14). Then, according to the definition of differential inclusion, there exist measurable functions \( \xi_i(t) \in \partial f_i(x_i(t)), \eta_i(t) \in \partial D_i(x_i(t)) \) and \( \zeta_i(t) \in \sum_{j \in \mathcal{N}_i} \partial \|x_i(t) - x_j(t)\| \), such that

\[
\dot{x}_i(t) = -x_i(t) + \mathcal{P}_\Omega \left( x_i(t) - \xi_i(t) - \sigma \eta_i(t) - \sigma^2 \zeta_i(t) \right),
\]

for a.e. \( t \in [0, T] \).
Construct a function as follows:

\[ H(x) = \frac{1}{2} \sum_{i=1}^{N} \| x_i - x_i^* \|^2 + h(x) - h(x^*). \]  

(18)

The derivative of \( H(x(t)) \) with respect to time \( t \) is

\[
\frac{dH(x(t))}{dt} = \sum_{i=1}^{N} \left\{ \xi_i(t) + \sigma \eta_i(t) + \sigma^2 \zeta_i(t) + x_i(t) - x_i^* + \xi_i(t) \right\}
\]

\[
= \sum_{i=1}^{N} \left\{ \xi_i(t) + \sigma \eta_i(t) + \sigma^2 \zeta_i(t) + x_i(t) - x_i^* - x_i(t) \right\}
\]

\[
+ P_\Omega \left\{ x_i(t) - \xi_i(t) - \sigma \eta_i(t) - \sigma^2 \zeta_i(t) \right\}. \quad (19)
\]

It can be seen from Proposition 2 that

\[
\left\{ x_i(t) - \xi_i(t) - \sigma \eta_i(t) - \sigma^2 \zeta_i(t) - P_\Omega \right\}
\]

\[
(x_i(t) - \xi_i(t) - \sigma \eta_i(t) - \sigma^2 \zeta_i(t)),
\]

\[
P_\Omega \left\{ x_i(t) - \xi_i(t) - \sigma \eta_i(t) - \sigma^2 \zeta_i(t) \right\}
\]

\[
- x_i(t) + x_i(t) - x_i^* \geq 0,
\]

which implies that

\[
\left\{ \xi_i(t) + \sigma \eta_i(t) + \sigma^2 \zeta_i(t) + x_i(t) - x_i^*, P_\Omega \right\}
\]

\[
\left\{ x_i(t) - \xi_i(t) - \sigma \eta_i(t) - \sigma^2 \zeta_i(t) - x_i(t) \right\}
\]

\[
\leq - \sum_{i=1}^{N} \left\{ \xi_i(t) + \sigma \eta_i(t) + \sigma^2 \zeta_i(t), x_i(t) - x_i^* \right\}. \quad (20)
\]

Therefore, it has

\[
\frac{dH(x(t))}{dt} \leq - \sum_{i=1}^{N} \| x_i(t) - P_\Omega \left( x_i(t) - \xi_i(t) \right) - \sigma \eta_i(t) - \sigma^2 \zeta_i(t) \|^2
\]

\[
- \sum_{i=1}^{N} \left\{ \xi_i(t) + \sigma \eta_i(t) + \sigma^2 \zeta_i(t), x_i(t) - x_i^* \right\}. \quad (21)
\]

According to Definition 1, we have \( \sum_{i=1}^{N} \left\{ \xi_i(t) + \sigma \eta_i(t) + \sigma^2 \zeta_i(t), x_i(t) - x_i^* \right\} \geq 0, \) so

\[
\frac{dH(x(t))}{dt} \leq - \sum_{i=1}^{N} \| x_i(t) - P_\Omega \left( x_i(t) - \xi_i(t) \right) - \sigma \eta_i(t) - \sigma^2 \zeta_i(t) \|^2 \leq 0. \quad (21)
\]

From \( x(t) \in \Omega \) and (21), it can be obtained that \( 0 \leq H(x(t)) \leq H(x(0)) \) for \( t \geq 0. \) Then, \( x(t) \) is bounded. According to Theorem 3, we get that \( x(t) \) exists globally by the extension theorem of solutions in [1], that is

\[
x(t) \in \Omega, \; t \in [0, +\infty).
\]
Define another function

\[ G(x) = \inf \left\{ \sum_{i=1}^{N} \| x_i - P_{\Omega} \left( x_i - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right) \|^2 : \eta_i \in \partial D_i(x_i), \xi_i \in \partial f_i(x_i) \right\} \]

for all \( y_i \in \Omega \), and

\[ \langle \xi_i + \sigma \eta_i + \sigma^2 \xi_i, P_{\Omega} \left( x_i(t) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right) - y_i \rangle \leq \langle x_i(t) - P_{\Omega} \left( x_i(t) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right), P_{\Omega} \left( x_i(t) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right) - y_i \rangle \]

\[ = \langle x_i(t) - P_{\Omega} \left( x_i(t) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right), P_{\Omega} \left( x_i(t) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right) - x_i(t) + x_i(t) - y_i \rangle \]

\[ \leq \| x_i(t) - P_{\Omega} \left( x_i(t) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right) \| \| x_i(t) - y_i \|. \]

From (21), we have

\[ \frac{dH(x(t))}{dt} \leq -G(x(t)) \leq 0. \]  

(22)

Since \( H(x(t)) \geq 0 \) for \( t \geq 0 \), then there is \( H_0 \geq 0 \), such that

\[ \lim_{t \to +\infty} H(x(t)) = H_0. \]

Integrating (22) on \([0, t]\) yields

\[ H(x(t)) - H(x(0)) \leq -\int_0^t G(x(s))ds. \]

As \( t \to +\infty \), it derives that

\[ \int_0^{+\infty} G(x(s))ds \leq H(x(0)) - H_0 < +\infty. \]

Therefore, there exists a subsequence \( \{t_k\} \to 0 \), such that

\[ \lim_{k \to +\infty} G(x(t_k)) = 0. \]  

(23)

Due to 2) in Assumption 1 and the boundedness of \( x_i(t) \), \( \{x_i(t_k)\} \) has a convergent subsequence \( \{x_i(t)\} \) satisfying

\[ \lim_{t \to +\infty} x_i(t) = \hat{x}_i \]

and \( \hat{x}_i \in \Omega \) since \( x_i(t) \in \Omega \) and \( \Omega \) is closed. Next, we will prove that \( \hat{x} = \text{col} [\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_N] \) is an optimal solution of problem (13). From (23), there are \( \tilde{\xi}_i \in \partial f_i(x_i(t_k)), \eta_i \in \partial D_i(x_i(t_k)) \) and \( \tilde{\zeta}_i \in \sum_{j \in A_i} \partial \| x_i(t_k) - x_j(t_k) \| \), such that

\[ \lim_{t \to +\infty} \sum_{i=1}^{N} \| x_i(t_k) - P_{\Omega} \left( x_i(t_k) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right) \|^2 = 0. \]  

(24)

Then by the proposition of \( P_{\Omega} (\cdot) \), it yields that

\[ \langle x_i(t) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i, \phi \rangle \leq \langle x_i(t) - P_{\Omega} \left( x_i(t) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right), P_{\Omega} \left( x_i(t) - \xi_i - \sigma \eta_i - \sigma^2 \xi_i \right) - x_i(t_k) \rangle \]

\[ \geq 0 \quad \forall y_i \in \Omega. \]

Thus, \( \hat{x}_i - \tilde{\xi}_i - \tilde{\eta}_i - \tilde{\zeta}_i \in N_{\Omega} (\hat{x}_i) \) by Definition 3. Therefore, \( \hat{x} = \text{col} [\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_N] \) is an optimal solution to distributed optimization problem (13).

In the last step, we prove \( x_i(t) \to \hat{x}_i \) as \( t \to +\infty \). Let

\[ \hat{H}(t, x(t)) = \frac{1}{2} \sum_{i=1}^{N} \| x_i(t) - \hat{x}_i \|^2 + h(x(t)) - h(\hat{x}). \]

Similar to what we did before, it derives that

\[ \frac{d\hat{H}(t, x(t))}{dt} = -\sum_{i=1}^{N} \| \hat{x}_i(t) \|^2 \leq 0 \]

for a.e. \( t \in [0, +\infty) \). Together with the fact \( \hat{H}(t, x(t)) \geq 0 \), then \( \lim_{t \to +\infty} \hat{H}(t, x(t)) \) exists. Therefore

\[ \lim_{t \to +\infty} \hat{H}(t, x(t)) = \lim_{t \to +\infty} \hat{H}(t_i, x(t_i)) = 0. \]  

(25)
Since \( \frac{1}{2} \sum_{i=1}^{N} \| x_i(t) - \hat{x}_i \|^2 \leq \hat{H}(t, x(t)) \), then
\[
\lim_{t \to +\infty} \frac{1}{2} \sum_{i=1}^{N} \| x_i(t) - \hat{x}_i \|^2 = 0.\]
That is
\[
\lim_{t \to +\infty} x_i(t) = \hat{x}_i.
\]
By means of Theorem 2 and \( \sigma \geq \max\{LM/\bar{g}, NL + \sqrt{NL(NL + 2)}\} \), it follows \( \sigma \geq LM/\bar{g} \) and \( \sigma^2 > 2(1 + \sigma)NL \), thus \( x_i(t) \) converges to the optimal solution to optimization problem (1).

**Remark 4** The continuous-time neurodynamic approach proposed in this subsection can be used to solve nonsmooth distributed optimization problems with inequality and set constraints, and the state solution converges to an optimal solution. Thus, the continuous-time neurodynamic approach has better convergence property than [6] and [18] in which the state solution only converges to the optimal solution set. In addition, the proposed continuous-time neurodynamic approach can be extended to solve distributed optimization problems in [9,11,13,17,46]. See Table 1 for details.

### Event-triggered neurodynamic approach

It must be pointed out that in the process of solving large-scale optimization problems, the information transmission of agents may consume a lot of energy. Therefore, the addition of an event-triggered mechanism will greatly reduce energy consumption and save communication costs, which has been studied in [16,25,51]. Therefore, in this subsection, a distributed event-triggered neurodynamic approach is presented for solving distributed optimization problems and the corresponding event-triggered condition is designed by the Lyapunov approach. Before introducing the event-triggered mechanism, the following theorem should be shown.

**Theorem 5** Assume that \( p(x) \) is a strongly convex function and \( x^* \) is the optimal solution of the following optimization problem:

\[
\begin{align*}
\min & \quad p(x) \\
\text{s.t.} & \quad \varphi(x) = 0, \quad x \in \Omega,
\end{align*}
\]

where \( \Omega \subseteq \mathbb{R}^n \) is a compact set. Let \( x^{[v]} \) is the optimal solution of

\[
\begin{align*}
\min & \quad G^{[v]}(x) = p(x) + \frac{\nu}{2} \| \varphi(x) \|^2 \\
\text{s.t.} & \quad x \in \Omega;
\end{align*}
\]

then
\[
\lim_{v \to +\infty} x^{[v]} = x^*.
\]

**Proof** According to the penalty method mentioned in [4], if \( x^* \) is the optimal solution of the constrained optimization problem (26) and \( y^{[v]} \) is the optimal solution of the following unconstrained optimization problem:

\[
\min \quad F^{[v]}(y) = p(y) + \frac{\nu}{2} \| \varphi(y) \|^2 + \frac{\alpha}{2} \| y - x^* \|^2
\]

for any \( v \in \mathbb{N} \) and \( \alpha > 0 \), then \( \lim_{v \to +\infty} y^{[v]} = x^* \).

Since \( x^{[v]} \) is the optimal solution of optimization problem (27), from the definitions of \( G^{[v]}(x) \), \( F^{[v]}(y) \) and \( x^{[v]} \), we have

\[
G^{[v]}(x^{[v]}) \leq G^{[v]}(y^{[v]}) \leq F^{[v]}(y^{[v]}) \leq F^{[v]}(x^*) = p(x^*).
\]

Namely, \( p(x^{[v]}) + \frac{\nu}{2} \| \varphi(x^{[v]}) \|^2 \leq p(x^*) \).

Due to \( \Omega \) is a compact set, then \( p(x^{[v]}) \) is bounded on \( \Omega \). Thus, \( \lim_{v \to +\infty} \| \varphi(x^{[v]}) \|^2 = 0 \), which implies \( \varphi(x) = 0 \) for any limit point \( x \) of \( \{x^{[v]}\} \). Let \( v \to +\infty \), and it derives \( p(x) = p(x^*) \). From the strongly convexity of \( p(x) \), it is easy to get that \( \lim_{v \to +\infty} x^{[v]} = \bar{x} = x^* \).

According to Theorem 5, we construct a new distributed optimization problem

\[
\begin{align*}
\min & \quad H^{[v]}(x) = \sum_{i=1}^{N} f_i(x_i) + v \sum_{i=1}^{N} G_i^2(x_i) \\
& \quad + v \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \| x_i - x_j \|^2
\end{align*}
\]

s.t. \( x \in \Omega \),

where \( f_i(\cdot) \), \( G_i(\cdot) \) \( (i \in \mathcal{V}) \) and \( \Omega \) are from the distributed optimization problem (12). Here, we further assume that \( f_i(\cdot) \) is differentiable strongly convex, so it is obvious that \( H^{[v]}(\cdot) \) is a differentiable strongly convex function for any \( v \in \mathbb{N} \). Thus, distributed optimization problem (29) has a unique optimal solution, denoted by \( x^{[v]} \). Now, a neurodynamic approach with event-triggered mechanism is proposed as follows:

\[
\begin{align*}
\dot{x}_i(t) &= -x_i(t) + P_{\Omega}(x_i(t)) \\
& \quad - \beta_i \nabla H_i^{[v]}(x_i(t)) \quad (t \in t_k^{(i)}, t_k^{(i)}), \quad i \in \mathcal{V},
\end{align*}
\]

where \( t_k^{(i)} \in [0, +\infty) \) \((k \in \mathbb{N})\) is the \( k \)th triggering instant, \( \beta_i \) is a positive parameter, and \( H_i^{[v]}(x_i) = f_i(x_i) + v G_i^2(x_i) + v \sum_{j \in \mathcal{N}_i} \| x_i - x_j \|^2 \).

**Remark 5** In fact, an event-triggering algorithm consists of two parts: a neurodynamic algorithm and an event-triggering mechanism. Combining with published literatures about event-triggering mechanism in [7,16,25,33,51], we are
inspired to construct the event-triggering algorithm in this paper. It is worth knowing that the core of event-triggering algorithms is data updating and communication happened at trigger instants.

Zeno behavior refers to the phenomenon that the system triggers countless events within a limited time. It is well known that Zeno behavior is not desirable for control implementation, since physical devices cannot sample infinitely fast. Therefore, it is not negligible to determine whether there is Zeno behavior in an event-triggered neurodynamic approach.

Definition 4 [19] Under an event-triggered neurodynamic approach, agent $i$ ($i \in \mathcal{V}$) is said to be Zeno if

$$\lim_{k \to +\infty} t_{k}^{(i)} = \sum_{k=0}^{+\infty} (t_{k+1}^{(i)} - t_{k}^{(i)}) = T_0,$$

where $T_0$ is a finite constant, and Zeno-free otherwise. Moreover, the event-triggered neurodynamic approach is said to be Zeno if there is an agent to be Zeno, and Zeno-free otherwise.

Remark 6 A crucial reason for the popularity of event-triggered neurodynamic approaches is that it can reduce unnecessary interaction, at the same time, save limited network resources. More specifically, the information interaction occurs only when the trigger condition is met. Inspired by [47], an event-triggering condition can be designed based on Lyapunov functions to ensure convergence of the neurodynamic approach on the one hand and prevent Zeno behavior on the other.

Assumption 3 The gradient $\nabla H_i^{[v]}(\cdot)$ is $P_i^{[v]}$-Lipschitz with $P_i^{[v]} > 0$ on $\Omega$, that is

$$\|\nabla H_i^{[v]}(x) - \nabla H_i^{[v]}(y)\| \leq P_i^{[v]} \|x - y\|, \forall x, y \in \Omega.$$

Therefore, if the initial value $x_i(0) \in \Omega$, it can be obtained that $x_i(t) \in \Omega$. Let $x^* = \col{x_1^*, x_2^*, \ldots, x_N^*}$ and $x^{[v]} = \col{x_1^{[v]}, x_2^{[v]}, \ldots, x_N^{[v]}}$ be the optimal solutions for distributed optimization problem (13) and (29), respectively, $e_i(t) = x_i(t) - x_i(t_k^{(i)})$ be the measurement error for agent $i \in \mathcal{V}$. Then, we specify the event-triggered condition for storing or updating the sampled data

$$t_{k+1}^{(i)} = \inf_{t > t_k^{(i)}} \left\{ t > t_k^{(i)} \middle| \frac{q_i(t)}{P_i^{[v]}(\|P_i^{[v]}(\|x(t) - x_i(t_k^{(i)})\|)$

where $\xi_i = \col{P_i \{x_i(t) - \beta_i \nabla H_i^{[v]}(x_i(t_k^{(i)})), x_i(t)\}}, P_i = \left( \begin{array}{cc} 1 - \frac{\beta_i P_i^{[v]}}{2} & -\frac{3}{2} \\ -\frac{3}{2} & 2 - \frac{\beta_i P_i^{[v]}}{2} \end{array} \right) \otimes I_n$. Define $\bar{y}_i(t) = x_i(t) - \beta_i \nabla H_i^{[v]}(x_i(t_k^{(i)}))$ to facilitate the following description.

Theorem 6 If Assumptions 1–3 hold and $x_0 \in \Omega$, then the state solution $x(t)$ of event-triggered neurodynamic approach (30) converges to $x^*$ when the parameters $\beta_i$ and $P_i^{[v]}$ satisfy $\beta_i P_i^{[v]} > 3 + \sqrt{10}$ for each agent $i \in \mathcal{V}$. Furthermore, the event-triggered neurodynamic approach (30) is Zeno-free.
Proof Consider a Lyapunov function
\[ V^v(x) = \frac{1}{2} \|x\|^2 + H^v(x) - H^v(x^*), \]
where \( H^v(\cdot) \) is defined in (29). Since \( x^* \) is the optimal solution of distributed optimization problem (29), then \( V^v(x) \geq 0 \) for \( x \in \Omega \). Therefore, for \( t \in [t_k^{(i)}, t_{k+1}^{(i)}) \), taking the time derivative of \( V^v(x(t)) \), it has
\[ \dot{V}^v(x(t)) = \sum_{i=1}^{N} \langle \dot{x}_i(t), \dot{x}_i(t) \rangle + \sum_{i=1}^{N} \langle \nabla H^v_i(x_i(t)), \dot{x}_i(t) \rangle. \]

Since
\[ \langle x_i(t), \dot{x}_i(t) \rangle + \langle \nabla H^v_i(x_i(t)), \dot{x}_i(t) \rangle \]
\[ = \langle x_i(t), -x_i(t) + P_\Omega(\tilde{y}_i(t)) \rangle + \langle \nabla^2 H^v_i(x_i(t)) \rangle, \]
\[ = \langle x_i(t) + \beta_i \nabla H^v_i(x_i(t_k^{(i)})), -x_i(t) + P_\Omega(\tilde{y}_i(t)) \rangle + 2\langle x_i(t), -x_i(t) + P_\Omega(\tilde{y}_i(t)) \rangle \]
\[ + \beta_i \langle \nabla H^v_i(x_i(t)) \rangle - \nabla H^v_i(x_i(t_k^{(i)})), -x_i(t) + P_\Omega(\tilde{y}_i(t)) \rangle \]
\[ = \langle P_\Omega(\tilde{y}_i(t)) \rangle, -\dot{y}_i(t), -x_i(t) + P_\Omega(\tilde{y}_i(t)) \rangle \]
\[ + \beta_i \langle \nabla H^v_i(x_i(t)) \rangle - \nabla H^v_i(x_i(t_k^{(i)})), -x_i(t) + P_\Omega(\tilde{y}_i(t)) \rangle \]
\[ + \beta_i \langle \nabla H^v_i(x_i(t)) \rangle - \nabla H^v_i(x_i(t_k^{(i)})), -x_i(t) + P_\Omega(\tilde{y}_i(t)) \rangle. \]

According to 1) in Proposition 2, it follows that
\[ \langle P_\Omega(\tilde{y}_i(t)) \rangle, -\dot{y}_i(t), -x_i(t) + P_\Omega(\tilde{y}_i(t)) \rangle \leq 0. \]

Furthermore, based on Assumption 3 and the Cauchy–Schwarz inequality, we have
\[ \langle x_i(t), \dot{x}_i(t) \rangle + \langle \nabla H^v_i(x_i(t)), \dot{x}_i(t) \rangle \]
\[ \leq -\|P_\Omega(\tilde{y}_i(t))\|^2 - 2\|x_i(t)\|^2 + 3\langle x_i(t), P_\Omega(\tilde{y}_i(t)) \rangle \]
\[ + \beta_i P_i^v \|e_i(t)\| \left( \|x_i(t)\| + \|P_\Omega(\tilde{y}_i(t))\| \right) \]
\[ \leq \left( -1 + \frac{\beta_i P_i^v}{2} \right) \|P_\Omega(\tilde{y}_i(t))\|^2 + \left( -2 + \frac{\beta_i P_i^v}{2} \right) \]
\[ + 2\|x_i(t)\|^2 + 3\langle x_i(t), P_\Omega(\tilde{y}_i(t)) \rangle + \beta_i P_i^v \|e_i(t)\|^2 \]
\[ = -\xi^2_j \Pi_i \xi_j + \beta_i P_i^v \|e_i(t)\|^2. \]

Therefore, \( \dot{V}^v(x(t)) < 0 \) when \( t \in [t_k^{(i)}, t_{k+1}^{(i)}) \) and \( \beta_i P_i^v > 3 + \sqrt{10} \). Then, the Lyapunov stability theorem and the strongly convexity of \( H^v(\cdot) \) imply that \( x(t) \) is convergent to \( x^* \), which is defined as the optimal solution of (29).

Next, since \( e_i(t_k^{(i)}) = 0 \) and \( \dot{e}_i(t) = \dot{x}_i(t) = -x_i(t) + P_\Omega(x_i(t) - \beta_i \nabla H^v_i(x_i(t_k^{(i)}))) \), then for \( t \in [t_k^{(i)}, t_{k+1}^{(i)}) \)
\[ \|e_i(t)\| \leq \int_{t_k^{(i)}}^{t_{k+1}^{(i)}} \|e_i(s)\|ds \leq \tilde{M}_i(t)(t - t_k^{(i)}), \]
where \( \tilde{M}_i(t) = \max_{t \leq s \leq t} \|x_i(s) + P_\Omega(x_i(s) - \beta_i \nabla H^v_i(x_i(t_k^{(i)})))\| \). From event-triggering condition (31), we can get that the next event will not be triggered before \( q_i(t) = 0 \). Thus
\[ \|e_i(t_k^{(i)} - t_{k+1}^{(i)}) = \sqrt{\xi_j^2 \Pi_i \xi_j} \leq \tilde{M}_i(t_k^{(i)} - t_{k+1}^{(i)}), \]
where \( t_k^{(i)} - t_{k+1}^{(i)} \) is the left limit of \( t_k^{(i)} \). Then, \( t_k^{(i)} - t_{k+1}^{(i)} \geq \frac{\sqrt{\xi_j^2 \Pi_i \xi_j}}{M_i(t_k^{(i)} - t_{k+1}^{(i)})} > 0 \). It is obvious that \( \lim_{k \rightarrow +\infty} (t_k^{(i)} - t_{k+1}^{(i)}) = +\infty \), which implies that event-triggered neurodynamic approach (30) is Zeno-free.

From Theorem 5, the following corollary can be drawn.

Corollary 1 If Assumptions 1–3 hold and \( x_0 \in \Omega \), then the state solution \( x(t) \) of event-triggered neurodynamic approach (30) converges to the optimal solution \( x^* \) of (2) when \( v \) is large enough and the parameters \( \beta_i \) and \( P_i^v \) satisfy \( \beta_i P_i^v > 3 + \sqrt{10} \) for each agent \( i \in V \). Furthermore, the event-triggered neurodynamic approach (30) is Zeno-free.

Remark 7 Event-triggering mechanism is outstanding in alleviating bandwidth pressure and saving energy consumption, because agents do not need to communicate information all the time as required in [11, 13, 29, 46], but communicate their local information intermittently. In addition, compared to works in [25, 33, 51], the event-triggered neurodynamic approach (30) in this paper can solve distributed optimization problems under inequality and set constraints.

Discrete-time neurodynamic approach

Since solving the distributed optimization problem relies on the information exchange between agents in the multi-agent system, and the communication of agents is essentially a discrete process in actual operation, so studies on discrete-time neurodynamic approaches are also achieved widespread attention. In this subsection, the discrete-time neurodynamic approach in [3] is extended for solving nonsmooth convex distributed optimization problem (13) under assumptions mentioned in Sect. 4.1.
Proposition 4  
Dynamic approaches in [28] and [38].

Proposition 4 Under Assumptions 1 and 2, \( \| \tau_k \| \leq \ell_k \) and \( \| x_{k+1} - x_k \| \leq \ell_k \).

Proof  It is clear that the conclusion holds if \( \tau_k = 0 \). Now, we suppose \( \tau_k \neq 0 \). In this case, consider the convex function

\[
\Phi_{x_k}(w) = \frac{1}{2} \| w \|^2 + \frac{\ell_k}{\| \varsigma_k \|} \varsigma_k^T w, \quad \varsigma_k \in \partial h(x_k).
\]

Let

\[
\gamma_k = \tau_k + \frac{\ell_k}{\| \varsigma_k \|} \varsigma_k,
\]

then \( \gamma_k \in \partial \Phi_{x_k}(\tau_k) \). Since \( \tau_k \) is the minimum point of \( \Phi_{x_k}(w) \) on \( \Omega - x_k \), we have

\[
\langle u_k, z - \tau_k \rangle \geq 0,
\]

for all \( z \in \Omega - x_k \). Then, \( x_k \in \Omega \) implies that \( 0 \in \Omega - x_k \). Take \( z = 0 \), so

\[
\langle u_k, \tau_k \rangle \leq 0.
\]

By substituting (34) into (35), we can derive

\[
\| \tau_k \|^2 + \frac{\ell_k}{\| \varsigma_k \|} \varsigma_k^T \tau_k \leq 0,
\]

that is \( \| \tau_k \|^2 \leq -\frac{\ell_k}{\| \varsigma_k \|} \varsigma_k^T \tau_k \). Therefore

\[
\| \tau_k \|^2 \leq \frac{\ell_k}{\| \varsigma_k \|} \| \varsigma_k \| \| \tau_k \| \leq \ell_k \| \tau_k \|. \tag{37}
\]

The result can be obtained directly, since \( \| \tau_k \| \neq 0 \).

Proposition 5  If Assumptions 1 and 2 hold, then for any \( x \in \Omega \)

\[
3 \ell_k^2 + \| x - x_k \|^2 + \| x - x_{k+1} \|^2 \geq 2 \frac{\ell_k}{\| \varsigma_k \|} (h(x_k) - h(x)).
\]

Proof  From Proposition 4, we have

\[
\ell_k^2 + \| x - x_k \|^2 + \| x - x_{k+1} \|^2 \\
\geq \| x_{k+1} - x_k \|^2 + \| x - x_k \|^2 - \| x - x_{k+1} \|^2 \\
= 2(x - x_k, x_{k+1} - x_k) \\
= 2(\tau_k, x - x_k). \tag{38}
\]

By (35) and (34), it has \( u_k = \tau_k + \frac{\ell_k}{\| \varsigma_k \|} \varsigma_k \) and \( \{ u_k, x - x_k - \tau_k \} \geq 0, \forall x \in \Omega \). Then

\[
\langle \tau_k, x - x_k \rangle \\
= \langle u_k, x - x_k \rangle + \frac{\ell_k}{\| \varsigma_k \|} \langle \varsigma_k, x - x_k \rangle \\
\geq \langle u_k, \tau_k \rangle + \frac{\ell_k}{\| \varsigma_k \|} \langle \varsigma_k, x - x_k \rangle \\
= \langle \tau_k + \frac{\ell_k}{\| \varsigma_k \|} \varsigma_k, \tau_k \rangle + \frac{\ell_k}{\| \varsigma_k \|} \langle \varsigma_k, x - x_k \rangle \\
= \| \tau_k \|^2 + \frac{\ell_k}{\| \varsigma_k \|} \langle \varsigma_k, \tau_k \rangle + \frac{\ell_k}{\| \varsigma_k \|} \langle \varsigma_k, x - x_k \rangle \\
\geq \frac{\ell_k}{\| \varsigma_k \|} \langle \varsigma_k, \tau_k \rangle + \frac{\ell_k}{\| \varsigma_k \|} \langle \varsigma_k, x - x_k \rangle. \tag{39}
\]
In combination with the above inequalities

\[
\ell_k^2 + \|x - x_k\|^2 - \|x - x_{k+1}\|^2 \\
\geq 2\ell_k \|\varsigma_k\| (\varsigma_k, \tau_k) + 2\ell_k \|\varsigma_k\| (\varsigma_k, x_k - x) \\
\geq -2\ell_k \|\varsigma_k\| \|\tau_k\| + 2\ell_k \|\varsigma_k\| (\varsigma_k, x_k - x) \\
= -2\ell_k \|\tau_k\| + 2\ell_k \|\varsigma_k\| (\varsigma_k, x_k - x) \\
\geq -2\ell_k^2 + 2\ell_k \|\varsigma_k\| (\varsigma_k, x_k - x). \tag{40}
\]

By means of the convexity of \(h(x)\), then \((\varsigma_k, x_k - x) \geq h(x_k) - h(x)\), since \(\varsigma_k \in \partial h(x_k)\). Hence, the conclusion holds. \(\square\)

**Lemma 2** If Assumptions 1 and 2 hold, then \(\{x_k\}\) is bounded.

**Proof** According to Proposition 5, for \(x^* \in \mathcal{O}\), we have \(3\ell_i^2 + \|x^* - x_i\|^2 - \|x^* - x_{i+1}\|^2 \geq 2\ell_k \|h(x_k) - h(x^*)\| \geq 0\).

Therefore, it can be obtained that

\[
\|x^* - x_{i+1}\|^2 \leq \|x^* - x_i\|^2 + 3\ell_i^2 \\
\leq \|x^* - x_{i-1}\|^2 + 3\ell_{i-1}^2 + 3\ell_i^2 \\
\leq \cdots \leq \|x^* - x_0\|^2 + 3 \sum_{i=0}^{\infty} \ell_i^2. \tag{41}
\]

As a result, \(\{x_k\}\) is bounded by \(\sum_{k=0}^{\infty} \ell_k^2 < +\infty\). \(\square\)

**Theorem 7** Under Assumptions 1 and 2, there exists a cluster point of \(\{x_k\}\), noted as \(x^*\), and \(x^* \in \mathcal{O}\).

**Proof** From Proposition 5, it is obvious that for any \(x \in \mathcal{O}\)

\[
2\frac{\ell_k}{\|\varsigma_k\|} h(x_k) - h(x) \\
\leq \|x_k - x\|^2 - \|x_{k+1} - x\|^2 + 3\ell_k^2, \quad k \in \mathbb{N}. \tag{42}
\]

Since \(\{x_k\}\) is bounded, then \(\{\varsigma_k\}\) is bounded. Without lose of generality, we suppose that \(\|\varsigma_k\| \leq K, \quad k \in \mathbb{N}\) and \(\gamma_k = \gamma_k(x) = h(x_k) - h(x)\) for any \(x \in \mathcal{O}\). Then

\[
2\frac{\ell_k}{K} \gamma_k \leq \|x_k - x\|^2 - \|x_{k+1} - x\|^2 + 3\ell_k^2. \tag{43}
\]

It is easy to calculate that

\[
\frac{2}{K} \sum_{k=0}^{m} \ell_k \gamma_k \leq \sum_{k=0}^{m} \left(\|x_k - x\|^2 - \|x_{k+1} - x\|^2 + 3\ell_k^2\right) \\
= \|x_0 - x\|^2 - \|x_{m+1} - x\|^2 + 3 \sum_{k=0}^{m} \ell_k^2 \\
\leq \|x_0 - x\|^2 + 3 \sum_{k=0}^{m} \ell_k^2. \tag{44}
\]

Let \(m \to +\infty\), we have

\[
\sum_{k=0}^{+\infty} \ell_k \gamma_k < +\infty. \tag{45}
\]

Therefore, it has a subsequence \(\{\gamma_{i_k}\}\), such that \(\lim_{k \to +\infty} \gamma_{i_k} \leq 0\). Otherwise, there are \(\rho > 0, k \geq 0\), such that

\[
+\infty \sum_{k=0}^{+\infty} \ell_k \gamma_k \geq +\infty \sum_{k=k}^{+\infty} \ell_k, \quad k \geq k, \tag{47}
\]

which contradicts with \(\sum_{k=0}^{+\infty} \ell_k = +\infty\). Therefore, \(\{x_{i_k}\}\) converges to a point, noted by \(x^*\). Since \(\mathcal{O}\) is a closed set, then \(x^* \in \mathcal{O}\).

Next we prove \(x^* \in \mathcal{O}\). If \(x^* \notin \mathcal{O}\), then there exists \(\hat{x} \in \mathcal{O}\), such that \(h(\hat{x}) < h(x^*)\). It has been proved above that \(\lim_{k \to +\infty} \gamma_{i_k} = 0\). It yields

\[
\lim_{k \to +\infty} h(x_{i_k}) - h(\hat{x}) = h(x^*) - h(\hat{x}) \leq 0, \tag{46}
\]

that is, \(h(\hat{x}) \geq h(x^*)\), which results in contradiction. \(\square\)

**Theorem 8** Under Assumptions 1 and 2, all cluster points of \(\{x_k\}\) are optimal solutions to distributed optimization problem (13).

**Proof** According to the proof of Theorem 7, it is sufficient to prove for any \(x \in \mathcal{O}\), all cluster points of \(\gamma_k = h(x_k) - h(x)\) are nonpositive. Since \(\mathcal{O}\) is a bounded set and \(\varsigma_k \in \partial h(x_k)\), there is a \(K > 0\), such that \(\|\varsigma_k\| \leq K\). Then, for \(k \in \mathbb{N}\)

\[
\gamma_k - \gamma_{k+1} = h(x_k) - h(x) - (h(x_{k+1}) - h(x)) \\
\leq \varsigma_k^T (x_k - x_{k+1}) \\
\leq \|\varsigma_k\| \|x_k - x_{k+1}\| \leq K \ell_k. \tag{47}
\]

By Theorem 7, there is a subsequence \(\{\gamma_{i_k}\}\) such that \(\lim_{k \to +\infty} \gamma_{i_k} \leq 0\). If there exist \(\delta > 0\) and another subsequence \(\{\gamma_{i_\delta}\}\), such that \(\gamma_{i_\delta} \geq \delta\). Therefore, we consider the
subsequence $\gamma_{j_k}$, where $j_k (k \in \mathbb{N})$ is defined as follows:

\[
\begin{align*}
    j_0 &= \min\{m \geq 0 : \gamma_m \geq \delta\}, \\
    j_{2k+1} &= \min \left\{ m \geq j_{2k} : \gamma_m \geq \frac{\delta}{2} \right\}, \\
    j_{2k+2} &= \min \{m \geq j_{2k+1} : \gamma_m \geq \delta\}.
\end{align*}
\]

Obviously, $\gamma_{j_k}$ is well defined. Next, it comes to the conclusion that

\[
\begin{align*}
    \gamma_m &\geq \delta, \quad j_{2k} \leq m \leq j_{2k+1} - 1 \\
    \gamma_m &\leq \frac{\delta}{2}, \quad j_{2k+1} \leq m \leq j_{2k+2} - 1
\end{align*}
\]

and

\[
\gamma_{j_{2k}} - \gamma_{j_{2k+1}} \geq \frac{\delta}{2}, \quad \forall k \in \mathbb{Z}_{\geq 0}.
\]

By (45) and (47), it can be derived that

\[
+\infty > \sum_{k=0}^{+\infty} \ell_k \gamma_k = \sum_{k:j_k > 0} \ell_k \gamma_k + \sum_{k:j_k \leq 0} \ell_k \gamma_k \geq \sum_{m=j_{2k}}^{+\infty} \ell_m \gamma_m + \sum_{k:j_k \leq 0} \ell_k \gamma_k
\]

\[
\geq \frac{\delta}{K} \sum_{k=0}^{+\infty} \sum_{m=j_k}^{+\infty} \ell_m \gamma_m + \sum_{k:j_k \leq 0} \ell_k \gamma_k
\]

\[
\geq \frac{\delta}{K} \sum_{k=0}^{+\infty} \sum_{m=j_k}^{+\infty} (\gamma_m - \gamma_{m+1}) + \sum_{k:j_k \leq 0} \ell_k \gamma_k
\]

\[
= \frac{\delta}{K} \sum_{k=0}^{+\infty} (\gamma_{j_{2k}} - \gamma_{j_{2k+1}}) + \sum_{k:j_k \leq 0} \ell_k \gamma_k.
\]

Because $\lim_{k \to +\infty} \gamma_k \leq 0$, so $\{k : \gamma_k \leq 0\}$ is an infinite set. Known from $\sum_{k=0}^{+\infty} \ell_k \gamma_k < +\infty$, then there is $\tilde{k} > 0$, such that

\[
\ell_k \gamma_k \geq -\frac{\delta^2}{4K},
\]

for $k \geq \tilde{k}$. Furthermore, define

\[
\tilde{S} = \sum_{\{k : k \gamma_k \leq 0\}} (\gamma_{j_{2k}} - \gamma_{j_{2k+1}} + \frac{K}{\delta} \ell_k \gamma_k).
\]

Obviously $\tilde{S}$ is finite. Based on (49), it gets

\[
\begin{align*}
\frac{\delta}{K} \sum_{k=0}^{+\infty} (\gamma_{j_{2k}} - \gamma_{j_{2k+1}}) + \sum_{\{k : \gamma_k \leq 0\}} \ell_k \gamma_k &\geq \frac{\delta}{K} \sum_{\{k : \gamma_k \leq 0\}} (\gamma_{j_{2k}} - \gamma_{j_{2k+1}} + \frac{K}{\delta} \ell_k \gamma_k) \\
&= \frac{\delta}{K} \sum_{\{k : \gamma_k \leq 0\}} (\gamma_{j_{2k}} - \gamma_{j_{2k+1}} + \frac{K}{\delta} \ell_k \gamma_k) \geq \frac{\delta}{K} \tilde{S} \\
&= +\infty, \quad (50)
\end{align*}
\]

which contradicts to (48). To sum up, all cluster points of $\{x_k\}$ belong to $C$. \hfill \Box

**Corollary 2** Assume $\{x_k\}$ is generated by Approach I, then $\{x_k\}$ converges to the optimal solution set of distributed optimization problem (13). Furthermore, if there is $\sigma > \max\{LM/\hat{g}, NL + \sqrt{NL(NL + 2)}\}$, then $\{x_i(k)\} (i \in V)$ converges to the optimal solution set of optimization problem (1).

**Remark 9** The variable step-size neurodynamic approach designed as Approach I can resolve the nonsmooth distributed optimization problem with convex inequality and set constraints. Compared with existing discrete-time neurodynamic approaches in [32,45] and [50] which can only solve smooth distributed optimization problems with equality constraints or affine inequality constraints, Approach I has the ability to solve the nonsmooth distributed optimization problem with inequality and set constraints. Additionally, the Approach I is convergent without the assumption that local cost functions are strongly convex.

### Simulations and applications

In this section, the continuous-time neurodynamic approach and the discrete-time neurodynamic approach in this paper are utilized to solve a numerical example. Moreover, an ill-conditioned Least Absolute Deviation problem and a load sharing problem are considered and solved. Therefore, the feasibility of proposed neurodynamic approaches is verified.

**Numerical simulations**

**Example 1** Consider a system of 20 agents interacting over an undirected circled communication graph to collaboratively
solve the following optimization problem:

$$\min f(x) = \sum_{i=1}^{20} |i x_{i1} - x_{i2}|$$

s.t. \(x_{i1} + i x_{i2} \leq 2; \ x_{i1} \geq 0; \ x_{i1}^2 - x_{i2} \leq 4; \ \|x_i\| \leq 1, \) \hspace{1cm} (51)

where \(x_i = (x_{i1}, x_{i2})^T \) for \(i = 1, 2, \ldots, 20 \) and \(x = \text{col}\{x_1, x_2, \ldots, x_{20}\} \in \mathbb{R}^{40} \). Obviously, problem (51) is a nonsmooth convex distributed optimization problem under inequality constraint set \(\mathcal{E}_i = \{x \in \mathbb{R}^2 : x_{i1}^2 + i x_{i2} \leq 2, \ x_{i1} \geq 0, \ x_{i1}^2 - x_{i2} \leq 4\} \) and set constraint \(\Omega_i = \mathcal{B}(0, 1)\).

(I) Continuous-time neurodynamic approach (14) for (51)

It is easy to identify the inequality constraint set \(\mathcal{E}_i\) is bounded and there is \(\hat{x}_i = (1, 0) \in \text{int}(\mathcal{E}_i) \cap \Omega_i\) such that Assumption 1 is satisfied. Furthermore, for any \(x_i \in \mathcal{E}_i \cap \Omega_i\)

$$\partial f_i(x_i) = \begin{cases} (i, -1)^T, & ix_{i1} - x_{i2} > 0 \\
-1, 1\end{cases} (1, 1)^T, \quad (i, 1)^T, \quad (i x_{i1} - x_{i2} = 0 \\
(-i, 1)^T, \quad i x_{i1} - x_{i2} < 0 \quad \square$$

and \(\nabla g_1(x) = (1, i)^T, \ \nabla g_2(x) = (-1, 0)^T, \ \nabla g_3(x) = (2x_1, -1)^T\). Thus, the parameters are chosen as \(M = 10\) and \(L = 300\). According to utilizing continuous-time neurodynamic approach (14) from an initial point, we get that the state solution is convergent to an optimal solution \(x^* = [0, 0]^T\) of optimization problem (1), which is shown in Fig. 2a, and the trend of the global cost function is shown in Fig. 2b. Therefore, the effectiveness of the neurodynamic approach (14) can be guaranteed.

(II) Approach I for (51)

According to Approach I with the same initial value in (I), take \(\sigma = 121000\) and \(\alpha_k = 1/(k + 5000)\). Through 5000 iterations, then the convergence of the iteration sequence and the value of the global objective function \(f(x)\) can be seen intuitively in Fig. 3 and the result is same as that in(I).

(III) A higher dimensional simulation for (51)

Now, we come to consider a multi-agent system that \(N = 100\) and agents interact over an undirected communication graph showed in Fig. 4. to collaboratively solve the optimization problem (51). The 100 agents are divided into five floors, in which 20 agents in each layer are connected in a circular manner, and each two floors are connected by agents whose coordinate were \((x, y, z) = (1, 0, z)\). Therefore, this communication graph is undirected and connected. Therefore, by applying continuous-time neurodynamic approach (14) and Approach I proposed in this paper, respectively. Figure 5 shows the convergence of these two algorithms and illustrates their effectiveness for higher dimensional distributed optimization problems.

In Example 1, the dimension of the solution obtained by the neurodynamic approach 14 proposed in this paper is 40. Actually, the approaches proposed in [40] and [52] can be implemented to solve problem (51), and the dimensions of their solution are as high as 100 and 140, respectively. In addition, the approach in [51] is only suitable for solving distributed optimization problems with affine inequalities, so it cannot solve problem (51). However, when solving optimization problems of the same dimension, the solution dimension generated by the approach in [51] is up to 160, which is three times larger than the solution dimension obtained by (14). It is worth pointing out that this phenomenon will become more obvious when dealing with problems under larger dimensional constraints. Thus, the neurodynamic approach (14) can greatly reduce the amount of calculation and reduce the consumption of computer CPU.

**Least absolute deviation problem**

**Example 2** The system in [50] and Approach I proposed in this paper are used to solve the following ill-conditioned Least Absolute Deviation problem:

$$\min \sum_{i=1}^{N} \|Dx - c\|_1$$

(52)

where \(D = \begin{pmatrix} 1.1 & 4 & 7 & -3.1 & 5.5 \\ 2 & 5.1 & 8 & -4 & 6.9 \\ 3.1 & 6 & 8.9 & -5 & 8.1 \end{pmatrix}^T\) and \(c = (-2, 1.1, 4.2, -0.5, 2.1)^T\).

Since the system in [50] and Approach I are able to solve the nonsmooth distributed convex optimization problems. As [50] mentioned, the condition number of the matrix is about 200, so it is difficult to solve problem (52) using traditional optimization algorithms. Here, we consider a multi-agent network consisting of five nodes, each of which is assigned an objective function containing \(D\) and \(c\). The communication lines between agents can be seen in Fig. 6. Then, from the same initial point, both algorithms produce sequences of iterations that converge to the optimal solution \((2, 1, -2)^T\) and the results are given in Fig. 7. It is clear from Fig. 7 that one of the advantages of the algorithm in this paper is much faster convergence.

**Load sharing problem**

In the field of power system, the load sharing problem is frequently considered. We now want to find the optimal generation allocation to share the load under the constraints
Fig. 2  Convergence behavior of neurodynamic approach (14)

(a) Trajectory of the state solution $x(t)$

(b) The trend of global cost function $f(x(t))$

Fig. 3  Convergence behavior of approach I

(a) Trajectory of the iteration sequence $\{x_k\}$

(b) The trend of global cost function $f(x_k)$

Fig. 4  Communication topology of the multi-agent system in Example 1
(a) Trajectory of the state solution $\mathbf{x}(t)$  \hspace{1cm} (b) Trajectory of the iteration sequence $\{\mathbf{x}_k\}$

Fig. 5 The trajectories of 100 agents

Fig. 6 Communication topology of the multi-agent system in Example 2 and the interaction of the buses in Example 3

of generation and transmission capacity. Suppose the transmission network $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where the buses set is $\mathcal{V} = \{1, 2, \ldots, N\}$, the transmission lines set is $\mathcal{E} = \{1, \ldots, d\}$. $\mathcal{G}$ is connected and let a incidence matrix $D = (D_{ij}) \in \mathbb{R}^{N \times d}$ with

$$D_{ij} = \begin{cases} 1, & \text{line } j \text{ enters the bus } i \\ -1, & \text{line } j \text{ departs from the bus } i \\ 0, & \text{otherwise.} \end{cases}$$

By the work of [46], the load sharing optimal problem can be abstracted into the following mathematical model:

Fig. 7 Trajectory of the iteration sequence $\{x_k\}$
\[ \begin{align*} 
\text{min} & \quad \sum_{i \in \mathcal{V}} f_i(p_i^{\text{gene}}) \\
\text{s.t.} & \quad p_i^{\text{gene}} - Dv - p_i^{\text{load}} = 0_n \\
& \quad L_i^{\text{gene}} \leq p_i^{\text{gene}} \leq \bar{p}_i^{\text{gene}}, \forall i \in \mathcal{V} \\
& \quad v_j \leq v \leq \bar{v}_j, \forall j \in \mathcal{E}, 
\end{align*} \]

where \( p_i^{\text{load}} \in \mathbb{R}^n \) is the constant load demands at buses, \( f_i \) is the local cost function at bus \( i \), \( L_i^{\text{gene}} \leq p_i^{\text{gene}} \leq \bar{p}_i^{\text{gene}} \) is the generation capacity constraint at bus \( i \), \( v_j \leq v \leq \bar{v}_j \) is the power flow constraint in line \( j \). Notice that \( p_i^{\text{load}} = [p_1^{\text{load}}, p_2^{\text{load}}, \ldots, p_N^{\text{load}}]^T \) and \( v = [v_1, v_2, \ldots, v_d]^T \). We need to know that the upper and lower bounds of the constraint are given constants.

Equivalently, the problem (53) can be reformulated as the following distributed optimization problem:

\[ \begin{align*} 
\text{min} & \quad \sum_{i \in \mathcal{V}} f_i(p_i^{\text{gene}}) = \sum_{i \in \mathcal{V}} f_i \left( \sum_{j=1}^{d} D_{ij} v_j + p_i^{\text{load}} \right) \\
\text{s.t.} & \quad p_i^{\text{gene}} \leq \sum_{j=1}^{d} D_{ij} v_j + p_i^{\text{load}} \leq \bar{p}_i^{\text{gene}}, \forall i \in \mathcal{V} \\
& \quad v_j \leq v \leq \bar{v}_j, \forall j \in \mathcal{E}. 
\end{align*} \]

**Example 3** To solve the load sharing problem (54), we consider a five-bus and five-line system, and the interaction of the buses is shown in Fig. 6, i.e. \( N = d = 5 \). Assuming that \( p_i^{\text{load}} = [3, 1, 4, 2, 3]^T \), \( \bar{p}_i^{\text{gene}} = [4, 3, 1, 2, 5]^T \), \( \bar{p}_i^{\text{gene}} = [0, 0, 0, 0, 0]^T \), and \( \bar{v} = [4, 5, 7, 3, 1]^T \), \( \bar{v} = [-4, -3, -6, -5, -10]^T \)

\[ f_i = 2(p_i^{\text{gene}})^2 + 2p_i^{\text{gene}}, \quad i = 1, 2, \ldots, 5 \]

After digital simulation, we get the final result \( v = [3.3453, 2.9926, 0.9915, 1.9883, 3.6823]^T \), \( p_i^{\text{gene}} = [0.6488, -1.3438, 1.6647, 1.6764, 0.9941]^T \). The convergence process is shown in Fig. 8. Obviously, the solution satisfies the given constraints and minimizes the cost of load sharing.

**Conclusion**

In this paper, we proposed three neurodynamic approaches for solving distributed optimization problems with inequality and set constraints. At first, a continuous-time neurodynamic approach was proposed without using the Laplacian matrix of the communication topology, and the state solution of the neurodynamic approach was proved to converge to an optimal solution of the nonsmooth distributed optimization problem under several mild assumptions. Then, an event-triggered neurodynamic approach was designed to reduce the communication burden and it was verified that Zeno behavior does not occur. Furthermore, we proposed a discrete-time neurodynamic approach and proved the iteration sequence converges to the optimal solution set of the nonsmooth distributed optimization problem. At last, numerical examples were realized to demonstrate the effectiveness and superiority of neurodynamic approaches. From an application point of view, we plan to extend neurodynamic approaches...
for solving distributed optimization problems on directed or time-varying communication graphs in the next step.

**Declarations**

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest. We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, and there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled “Distributed Neurodynamic Approaches to Nonsmooth Optimization Problems with Inequality and Set Constraints”. This research is supported by the National Natural Science Foundation of China (62176073, 11871178).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

**References**

1. Aubin J, Cellina A (1984) Differential inclusions. Springer, Berlin
2. Barreiro-Gomez J, Obando G, Quijano N (2017) Distributed population dynamics: optimization and control applications. IEEE Trans Syst Man Cybern Syst 47(2):304–314
3. Bello Cruz JY (2013) A subgradient method for vector optimization problems. SIAM J Optim 23(4):2169–2182
4. Bertsekas DP (1999) Nonlinear programming. Athena Scientific, Belmont
5. Bian W, Xue X (2015) Asymptotic behavior analysis on multivalued evolution inclusion with projection in Hilbert space. Optimization 64(4):853–875
6. Bian W, Ma L, Qin S, Xue X (2018) Neural network for nonsmooth pseudocvx optimization with general convex constraints. Neural Netw 101:1–14
7. Cao X, Tamer B (2021) Decentralized online convex optimization with event-triggered communications. IEEE Trans Signal Process 59:284–299
8. Chang TH, Hong M, Wang X (2014) Multi-agent distributed optimization via inexact consensus admn. IEEE Trans Signal Process 63(2):482–497
9. Cherukuri A, Cortes J (2016) Initialization-free distributed coordination for economic dispatch under varying loads and generator commitment. Automatica 74:183–193
10. Clarke FH (1983) Optimization and nonsmooth analysis. Wiley, New York
11. Gharesifard B, Corts J (2014) Distributed continuous-time convex optimization on weight-balanced digraphs. IEEE Trans Autom Control 59(3):781–786
12. Jusem AN (2003) On the convergence properties of the projected gradient method for convex optimization. Comput Appl Math 22(1):37–52
13. Jia W, Qin S, Xue X (2019) A generalized neural network for distributed nonsmooth optimization with inequality constraint. Neural Netw 119:46–56
14. Jiang X, Qin S, Xue X (2020) A subgradient-based continuous-time algorithm for constrained distributed quadratic programming. J Franklin Inst 357(9):5570–5590
15. Jiang X, Qin S, Xue X (2021) A continuous-time algorithm for approximate distributed optimization with affine equality and convex inequality constraints. IEEE Trans Syst Man Cybern Syst 51(9):5809–5818
16. Kia SS, Cortes J, Martinez S (2014) Periodic and event-triggered communication for distributed continuous-time convex optimization. In: 2014 American Control Conference, pp 5010–5015
17. Kia SS, Cortés J, Martínez S (2015) Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication. Automatica 55:254–264
18. Li Q, Liu Y, Zhu L (2014) Neural network for nonsmooth pseudocvx optimization with general constraints. Neurocomputing 131:336–347
19. Li X, Sun ZM, Tang Y, Karimi HR (2020) Adaptive event-triggered consensus of multi-agent systems on directed graphs. IEEE Trans Autom Control 99:1–16
20. Li W, Zeng X, Liang S, Hong Y (2021) Exponentially convergent algorithm design for constrained distributed optimization via nonsmooth approach. IEEE Trans Autom Control. https://doi.org/10.1109/TAC.2020.3000819
21. Lin P, Ren W, Song Y (2016) Distributed multi-agent optimization subject to nonidentical constraints and communication delays. Automatica 65:120–131
22. Lin P, Ren W, Yang C, Gui W (2019) Distributed continuous-time and discrete-time optimization with nonuniform unbounded convex constraint sets and nonuniform stepizes. IEEE Trans Autom Control 64(12):5148–5155
23. Liu Q, Wang J (2015) A second-order multi-agent network for bound-constrained distributed optimization. IEEE Trans Autom Control 60(12):3310–3315
24. Liu S, Qiu Z, Xie L (2017) Convergence rate analysis of distributed optimization with projected subgradient algorithm. Automatica 83:162–169
25. Liu C, Li H, Shi Y, Xu D (2020) Distributed event-triggered gradient method for constrained convex minimization. IEEE Trans Autom Control 65(2):778–785
26. Nedic A, Ozdaglar A (2008) Approximate primal solutions and rate analysis for dual subgradient methods. SIAM J Optim 19(4):1757–1780
27. Nedic A, Ozdaglar A (2009) Distributed subgradient methods for multi-agent optimization. IEEE Trans Autom Control 54(1):48–61
28. Nedic A, Ozdaglar A, Parrilo P (2010) Constrained consensus and optimization in multi-agent networks. IEEE Trans Autom Control 55(4):922–938
29. Qiu Z, Liu S, Xie L (2016) Distributed constrained optimal consensus of multi-agent systems. Automatica 68:209–215
30. Ram S, Nedic A, Veeravalli V (2010) Distributed stochastic subgradient projection algorithms for convex optimization. J Optim Theory Appl 147(3):516–545
31. Ren W, Sorensen N (2008) Distributed coordination architecture for multi-robot formation control. Robot Auton Syst 56(4):324–333
32. Tang Y, Zhu H, Lv X (2019) Distributed output optimization for discrete-time linear multi-agent systems. In: 2019 Chinese control conference, pp 5665–5669
33. Tran N, Wang Y, Liu X, Xiao J, Lei Y (2019) Distributed optimization problem for second-order multi-agent systems with
event-triggered and time-triggered communication. J Franklin Inst Eng Appl Math 356(17):10196–10215
34. Tsitsiklis JN (1984) Problems in decentralized decision making and computation. Master’s thesis Massachusetts Institute
35. Tsitsiklis JN, Bertsekas DP, Athans M (1986) Gradient optimization algorithms. IEEE Trans Autom Control 31(9):803–812
36. Wang J, Elia N (2010) Control approach to distributed optimization. In: 48th Annual Allerton conference on communication, control, and computing, pp 557–561
37. Wang D, Wang J, Wang W (2017) Discrete-time distributed optimization for multi-agent systems under Markovian switching topologies. In: 13th IEEE international conference on control & automation (ICCA), pp 747–752
38. Wang P, Lin P, Ren W, Song Y (2018) Distributed subgradient-based multiagent optimization with more general step sizes. IEEE Trans Autom Control 63(7):2295–2302
39. Wen X, Qin S (2021) A projection-based continuous-time algorithm for distributed optimization over multi-agent systems. Complex & Intelligent Systems. https://doi.org/10.1007/s40747-020-00265-x
40. Wen X, Luan L, Qin S (2021) A continuous-time neurodynamic approach and its discretization for distributed convex optimization over multi-agent systems. Neural Netw 143(4):52–65
41. Wuthishuwong C, Traechtler A (2020) A projection-based continuous-time algorithm for distributed optimization over multi-agent systems. Complex Intell Syst 1(6):162–187
42. Xia Y, Leung H, Wang J (2002) A projection neural network and its application to constrained optimization problems. IEEE Trans Circ Syst 49(4):447–458
43. Xue X, Bian W (2008) Subgradient-based neural networks for non-smooth convex optimization problems. IEEE Trans Circuits Syst 55(8):2378–2391
44. Yan Z, Fan J, Wang J (2017) A collective neurodynamic approach to constrained global optimization. IEEE Trans Neural Netw Learn Syst 28(5):1206–1215
45. Yang S, Liu Q, Wang J (2018) A collaborative neurodynamic approach to multiple-objective distributed optimization. IEEE Trans Neural Netw Learn Syst 29(4):981–992
46. Yi P, Hong Y, Liu F (2015) Distributed gradient algorithm for constrained optimization with application to load sharing in power systems. Syst Control Lett 83:45–52
47. Yu Z, Yu S, Jiang H, Mei X (2021) Distributed fixed-time optimization for multi-agent systems over a directed network. Nonlinear Dyn 103(1):775–789
48. Yuan D, Xu S, Zhao H (2011) Distributed primal-dual subgradient method for multiagent optimization via consensus algorithms. IEEE Trans Syst Man Cybern Syst 41(6):1715–1724
49. Zeng X, Yi P, Hong Y (2017) Distributed continuous-time algorithm for constrained convex optimizations via nonsmooth analysis approach. IEEE Trans Autom Control 62(10):5227–5233
50. Zhao Y, Liu Q (2020) A consensus algorithm based on collective neurodynamic system for distributed optimization with linear and bound constraints. Neural Netw 122:144–151
51. Zhao Z, Chen G, Dai M (2018) Distributed event-triggered scheme for a convex optimization problem in multi-agent systems. Neurocomputing 284:90–98
52. Zhu Y, Yu W, Wen G, Chen G, Ren W (2019) Continuous-time distributed subgradient algorithm for convex optimization with general constraints. IEEE Trans Autom Control 64(4):1694–1701

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.