A Continuous-Time Multi-Agent Systems Based Algorithm for Constrained Distributed Optimization

Ping Liu[1], Huaqing Li[1, 2, *], Liping Feng[3]

[1]College of Electronic and Information Engineering, Southwest University, Chongqing, PR China

[2]State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University, Chongqing, PR China

[3]Department of Computer Science and Technology of Xinzhou Normal University, Xinzhou, Shanxi, PR China

Email address:
huaqingli@hotmail.com (Huaqing Li)
*Corresponding author

Abstract: This paper considers a second-order multi-agent system for solving the non-smooth convex optimization problem, where the global objective function is a sum of local convex objective functions within different bound constraints over undirected graphs. A novel distributed continuous-time optimization algorithm is designed, where each agent only has an access to its own objective function and bound constraint. All the agents cooperatively minimize the global objective function under some mild conditions. In virtue of the KKT condition and the Lagrange multiplier method, the convergence of the resultant dynamical system is ensured by involving the Lyapunov stability theory and the hybrid LaSalle invariance principle of differential inclusion. A numerical example is conducted to verify the theoretical results.

Keywords: Distributed Optimization, Multi-Agent Network, Lyapunov Method, Bound Constraint, Continuous-Time

1. Introduction

The distributed optimization of a sum of local convex functions has been widely investigated in a variety of scenarios in recent years. Examples include multi-agents system, resource allocation in communication networks and localization in sensor networks [1-18], to just name a few. Numerous distributed optimization algorithms are designed to be in a discrete-time fashion to search the optimal solutions of the optimization problem in [3, 5, 8], while continuous-time strategies due to its relatively complete theoretical framework have been widely applied to the distributed optimization problems in [10-12].

Distributed algorithms are characterized by high reliability, scalability and reduced communication capabilities, which attract many researchers to intensively study the distributed optimization algorithms (see e.g. [13-24]). Nedić and Ozdaglar [25] was the first to systematically put forward the distributed optimization problems. A projection-based distributed algorithm was developed in [7], and the further

investigations with respect to set constrained optimization were show in [26-27]. What is worth mentioning is that Bianchi and Jakubowicz [26] presented a distributed constraint non-convex optimization algorithm which consists of two steps: a local stochastic gradient descent at each agent and a gossip step that drives the network of agents to a consensus. Different from the above, a distributed optimization problem subject to the (in-)equality constraint or set constraint was investigated in [28]. The authors proposed two distributed subgradient algorithms for muilt-agent optimization problems, where the goal of agents is to minimize a sum of local objective functions. Motivated by [28], the primal-dual subgradient algorithm was studied in Yuan et al. [29] and Zhu et al. [24] for muilt-agent optimization problems with set constraints. Furthermore, in order to solve an un-constrained optimization problem, where the objective function is formed by a sum of convex functions available to individual agent, a second-order distributed dynamic was given in [10], while a similar second-order continuous-time distributed algorithm was proposed to solve the convex

optimization problem in [11].

Inspired by the works of [7, 10, 11, 19], a novel distributed second-order continuous-time multi-agent system is proposed to solve a distributed convex optimization problem, where the objective function is a sum of local objective functions, and each one can only know its local information. That is to say, all the agents cooperatively reach the optimal solution of the optimization problem. To tackle the optimization with box constraints, a logarithmic barrier penalty function is used, which is different from previous studies that are mainly based on the projection algorithm. In comparison with the existing distributed optimization methods, the proposed method in this paper has the following three advantages. Firstly, this paper designs a novel distributed continuous-time algorithm to solve more general distributed convex optimization problems. Secondly, the proposed algorithm can solve the convex optimization problem with a sum of convex objective functions with local bound constraints. Meanwhile, it does not require the objective function to be smooth, which is required in the most existing recurrent neural network algorithms. Thirdly, the box constraints are treated with a logarithmic barrier penalty function, which makes the proposed algorithm has a faster convergence speed to obtain the approximate solutions. This has a sufficient accuracy to satisfy most of the actual demands comparing with the projection algorithms.

The remainder of this paper is outlined as follows. Some preliminaries about the graph theory, the non-smooth analysis, and the stability of differential inclusions are presented in Section 2. Section 3 formulates a convex optimization problem and proposes a distributed continuous-time algorithm. In Section 4, a complete convergence proof is conducted to indicate that the dynamic system is convergent and stable meanwhile the agent's estimates are convergent to the same optimal solution. A numerical example for illustration is given in Section 5. Conclusions are finally drawn in Section 6.

2. Mathematical Preliminaries

In this section, some preliminaries about the graph theory, the non-smooth analysis, and the stability of differential inclusion are introduced.

A. Algebraic Graph Theory

A weighted undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ consists of a vertex set $\mathcal{V} = \{v_1, ..., v_n\}$, an undirected edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ and a weighted adjacency matrix $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{n \times n}$. Here, $a_{ij} > 0$ if and only if $(v_i, v_j) \in \mathcal{E}$; $a_{ij} = 0$, otherwise. An undirected path between v_i and v_j is denoted by $(v_i, v_j) \in \mathcal{E}$, which means that v_i and v_j can exchange information with each other [30]. We assume that the communications between agents are bidirectional and the weights are positive, which indicates that the connection weights between v_i and v_j in graph $\mathcal{G}$ satisfy $a_{ij} = a_{ji} > 0$ if and only if there exists an edge $(v_i, v_j) \in \mathcal{E}$; otherwise, $a_{ij} = a_{ji} = 0$. The degree matrix is defined as:

$D = diag\{\sum_{j=1}^{n} a_{1j}, \sum_{j=1}^{n} a_{2j}, ..., \sum_{j=1}^{n} a_{nj}\}$, so the Laplacian matrix of the undirected graph $\mathcal{G}$ is defined as $L = D - \mathcal{A}$. An undirected graph is connected if there exists a path between any pair of distinct nodes. Especially, L for an undirected graph is symmetric positive semi-definite. However, for a directed graph, L does not have this property.

B. Non-smooth Analysis

Let X be a space of Banach, $\|\cdot\|$ be Euclidean norm of X, X^* be the conjugate space of X, and D be a subset of X.

Definition 2.1: There exist a $\varepsilon > 0$ and a $x \in \sigma(x, \varepsilon)$ such that f satisfies the following:

$$| f(y) - f(y') | \leq L \| y - y' \|$$

The function $f : D \to \mathbb{R}$ is called Lipschitz near $x \in D$, where L represents the Lipschitz constant. If f is Lipschitz near any point $x \in D$, then f is also said to be locally Lipschitz in D.

Definition 2.2: Assume that $f : D \to \mathbb{R}$ is Lipschitz near x. The generalized directional derivative of f at x in the direction $v \in X$ is given by:

$$f^\circ(x; v) = \limsup_{y \to x, t \to 0^+} \frac{f(y + tv) - f(y)}{t}$$

Definition 2.3: The generalized gradient of f is defined as:

$$\partial f(x) = \{\zeta \mid \zeta \in X^*, f^\circ(x; v) \geq \langle \zeta, v \rangle, \forall v \in X\}$$

Lemma 2.1: If f is Lipschitz near any point $x \in D$, then

$$\partial f(x) = \overline{co}\{\lim_{i \to \infty} \nabla f(x_i) \mid x_i \to x, x_i \notin \Omega_f\}$$

where $\overline{co}$ is the convex closed hull, and Ω_f is the null measure set, which is composed by the undefined points of the generalized gradient of V.

C. Stability of Differential Inclusion

For an autonomous differential inclusion system:

$$\dot{x} \in F(x) \tag{1}$$

where $x \in \mathbb{R}^n$, $F : \mathbb{R}^n \to P(\mathbb{R}^n)$ is an upper semi-continuous set-valued mapping with compact convex values and 0 is a balance point of (1). That is to say, $0 \in F(0)$.

Definition 2.4: Let $x(t)$ be a solution of (1). If there is a sequence $\{t_i\}$ satisfying:

$$x(t_i) \to q \in \mathbb{R}^n, t_i \to +\infty$$

then q is the ω- limit point of the solution $x(t)$ of (1). All the ω- limit-points make up the limit-set, which is donated as $\Omega(x)$.

Definition 2.5: For any point x_0 in Ω, if there exists a maximal solution of the system in Ω, then Ω is called the weakly invariant set of the system (1).

Theorem 2.2: (Lasalle invariance principle of differential inclusion) Assuming $V : \mathbb{R}^n \to \mathbb{R}$ is a positive definite and locally Lipschitz regular function for almost all t, it satisfies

$$\frac{d}{dt}V(x) \le 0$$

If there exists a constant $l > 0$ such that $L_l = \{x \in \mathbb{R}^n \mid V(x) \le l\}$ bounded, then for any solution $x(t)$ passing through $x_0 \in L_l$, we have that $dist(x(t), M) \to 0, t \to +\infty$, where M is the largest weakly invariant subset of $\overline{Z}_V \cap L_l$. Here, $\overline{Z}_V$ is the closure of $Z_V = \{\overline{x} \in \mathbb{R}^{mn}, \overline{\lambda} \in \mathbb{R}^{mn} \mid 0 \in \dot{\overline{V}}(\overline{x}, \overline{\lambda})\}$.

Remark 2.1: If all the conditions of Theorem 2.2 hold, and $Z_V = \{0\}$, then the trivial solution of the autonomous differential inclusion systems (1) is asymptotically stable.

3. Problem Formulation and Optimization Algorithm

A. Problem Formulation

Consider a network of n agents that interact with each other over a connected graph $\mathcal{G}$. Each agent has a local objective function f^i and a local bound constraint Ω^i for all $i \in \{1, ..., n\}$. The muilt-agent group cooperatively solves the following distributed optimization problem:

$$\text{minimize} \quad f(x) = \sum_{i=1}^{n} f^i(x) \tag{2}$$
$$\text{subject to} \quad x \in \bigcap_{i=1}^{n} \Omega^i$$

where $\Omega^i \subseteq \mathbb{R}^n$ is the closed convex set, and $\Omega = \bigcap_{i=1}^{n} \Omega^i$ has nonempty interior points. The local objective function $f^i : \mathbb{R}^n \to \mathbb{R}$ is convex and not necessarily smooth. Here, $x = [x_1, x_2, ..., x_n]^T \in \mathbb{R}^n$ is a column vector.

We give some meaningful results, which will be used in this paper:

Assumption 3.1. The optimization problem (2) has at least one finite optimal solution x^*.

Assumption 3.2. At least one of the local objective functions $f^i(x), i = 1, 2, ..., n$ has a positive definite Hessian matrix.

Assumption 3.3. The weighted graph $\mathcal{G}$ is undirected and connected.

Under Assumption 3.3, we have that 0 is a simple eigenvalue of Laplacian L and 1_n is an eigenvector of L_n corresponding to the simple zero eigenvalue. Moreover, $L1_n = 0$, $1_n^T L = 0_n$.

B. Optimization algorithm

Denote $x_i \in \mathbb{R}^n, i = 1, 2, ..., n,$ as an estimation of agent i, $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^{m \times n}$ as a matrix with column vector x_i, and $\overline{x} = \text{vec}(x) = (x_1^T, x_2^T, ..., x_n^T)^T \in \mathbb{R}^{nm}$. The objective of optimization problem (2) is to achieve the global minimizer

$$x^* = \arg \min f(x).$$

Next, we will provide an equivalent optimization problem of (2).

Lemma 3.1: Let $L = L_n \otimes I_m \in \mathbb{R}^{nm \times nm}$, where $L_n \in \mathbb{R}^{n \times n}$ is a symmetric and positive semi-definite Laplacian matrix of connected graph $\mathcal{G}$, $\otimes$ is the Kronecker product and I_m denotes the identity matrix in $\mathbb{R}^{m \times m}$. Then, the equivalence problem of (2) is described as:

$$\text{minimize} \quad \overline{f}(\overline{x}) = \sum_{i=1}^{n} f^i(x_i)$$
$$\text{subject to} \quad L\overline{x} = 0 \tag{3}$$
$$\overline{x} \in \Omega$$

where $\Omega = \prod_{i=1}^{n} \Omega^i$ is the Cartesian product.

Proof: According to Assumption 3.3, let $\overline{x} = 1_n \otimes x$, we have the following:

$$L\overline{x} = (L_n \otimes I_m)\overline{x} = (L_n \otimes I_m)(1_n \otimes x)$$
$$= (L_n 1_n) \otimes (I_m x)$$
$$= 0$$

Denote A, B, and C as matrices with approximate dimensions. In virtue of the properties of Kronecker product $\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$, it is clearly that:

$$L\overline{x} = (L_n \otimes I_m)\overline{x}$$
$$= (L_n \otimes I_m)\text{vec}(x)$$
$$= \text{vec}(I_m x L_n)$$

If $L\overline{x} = 0$, then $I_m x L_n = 0$, which implies $L_n x^T = 0$. According to Assumption 3.3, we have $x^T = \alpha^T \otimes 1_n$, where $\alpha_i (i = 1, 2, ..., m)$ is any real number and $\alpha = (\alpha_1, \alpha_2, ..., \alpha_m)^T \in \mathbb{R}^m$ is a column vector. Then,

$$x = \alpha \otimes 1_n^T = \begin{pmatrix} \alpha_1 1_n^T \\ \alpha_2 1_n^T \\ ... \\ \alpha_m 1_n^T \end{pmatrix}$$

where each column vector x^i of x satisfies $x^i = \alpha$. We have $\overline{x} = \text{vec}(x) = 1_n \otimes \alpha$.

Based on the above analysis, we have that $L\overline{x} = 0$ if and only if $\overline{x} = 1_n \otimes x$ for some $x \in \mathbb{R}^n$. Since $1_n \otimes x \in \Omega$, we have $x^i \in \Omega^i$, which implies $x \in \bigcap_{i=1}^{n} \Omega^i$ and $\overline{f}(1_n \otimes x) = f(x)$. Therefore, the problem (2) is equivalent to the problem (3).

Remark 3.1: In this paper, we mainly consider the case that $\overline{x} \in \Omega$ is the box constraint $x_i^{k,\min} \le x_i^k \le x_i^{k,\max}$. Let $x_i = [x_i^1, x_i^2, ..., x_i^m]^T \in \mathbb{R}^m$, $x_i^{\min} = (x_i^{1,\min}, x_i^{2,\min}, ..., x_i^{m,\min})^T \in \mathbb{R}^m$, and $x_i^{\max} = (x_i^{1,\max}, x_i^{2,\max}, ..., x_i^{m,\max})^T \in \mathbb{R}^m$ for $i = 1, ..., n$.

Denote $\bar{x}^{\min} = [(x_1^{\min})^{\mathrm{T}}, (x_2^{\min})^{\mathrm{T}},...,(x_n^{\min})^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{mn}$ and $\bar{x}^{\max} = [(x_1^{\max})^{\mathrm{T}}, (x_2^{\max})^{\mathrm{T}},...,(x_n^{\max})^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{mn}$. Then the problem (3) can be rewritten as:

$$\begin{aligned} \text{minimize} \quad & \bar{f}(\bar{x}) = \sum_{i=1}^{n} f^i(x_i) \\ \text{subject to} \quad & L\bar{x} = 0 \\ & x_i^{k,\min} \leq x_i^k \leq x_i^{k,\max}, \quad k=1,...,m \end{aligned} \quad (4)$$

where x_i^k stands for the matrix entry in the k-th row and i-th column of x.

Then, combining with the optimization problems (3), a new optimization problem based on the method of augmented Lagrangian is proposed as follow:

$$\begin{aligned} \text{minimize} \quad & \bar{f}(\bar{x}) = \sum_{i=1}^{n} f^i(x_i) + \frac{1}{2}\bar{x}^{\mathrm{T}} L\bar{x} \\ \text{subject to} \quad & L\bar{x} = 0 \\ & x_i^{k,\min} \leq x_i^k \leq x_i^{k,\max}, \quad k=1,...,m \end{aligned} \quad (5)$$

where $L = L_n \otimes I_m \in \mathbb{R}^{nm \times nm}$, $L_n \in \mathbb{R}^{n \times n}$ is the Laplacian of graph $\mathcal{G}$, and the quadratic penalty term in the objective function plays a damping role.

Building on the Theorems 3.25 and 3.27 in [25], the following results are obtained.

Lemma 3.1: Assume that Assumption 3.1 and 3.2 hold. Denote $v^* \in \mathbb{R}^{mn}$ as the multiplier for the equality constraint, and $\eta^* \in \mathbb{R}^{mn}$, $\mu^* \in \mathbb{R}^{mn}$ as the Lagrangian multipliers for inequality constraints. Then, $\bar{x}^* \in \mathbb{R}^{mn}$ is the optimal solution of optimization problem (4) if and only if the following KKT condition is established:

$$\begin{aligned} &\nabla \bar{f}(\bar{x}^*) + v^{*\mathrm{T}} L\bar{x}^* + \mu^{*\mathrm{T}}(\bar{x}^* - \bar{x}^{\min}) + \eta^{*\mathrm{T}}(\bar{x}^{\max} - \bar{x}^*) = 0, \\ &L\bar{x}^* = 0, \qquad\qquad v^* \geq 0, \\ &\bar{x}^* - \bar{x}^{\min} \leq 0, \qquad \bar{x}^{\max} - \bar{x}^* \leq 0, \\ &\mu^{*\mathrm{T}}(\bar{x}^* - \bar{x}^{\min}) = 0, \quad \eta^{*\mathrm{T}}(\bar{x}^{\max} - \bar{x}^*) = 0, \\ &\mu^* \geq 0, \qquad\qquad \eta^* \geq 0. \end{aligned} \quad (6)$$

Lemma 3.2: Assume that Assumptions 1-3 hold. If optimization problem (4) has the optimal solution $\bar{x}^* = (x^*,...,x^*)$, then it is also the optimal solution of (5).

Let v_i^*, η_i^* and μ_i^* be defined as that in Lemma 3.1. Similar to Lemma 3.1, $\bar{x}^* = (x^*,...,x^*)$ is the optimal solution of optimization problem (5) if and only if the following KKT condition is established:

$$\begin{aligned} &\nabla \bar{f}(\bar{x}^*) + \frac{1}{2}\bar{x}^{*\mathrm{T}} L\bar{x}^* + v^{*\mathrm{T}} Lx^* + \mu^{*\mathrm{T}}(\bar{x}^* - \bar{x}^{\min}) + \eta^{*\mathrm{T}}(\bar{x}^{\max} - \bar{x}^*) = 0, \\ &L\bar{x}^* = 0, \qquad\qquad v^* \geq 0, \\ &\bar{x}^* - \bar{x}^{\min} \leq 0, \qquad \bar{x}^{\max} - \bar{x}^* \leq 0, \\ &\mu^{*\mathrm{T}}(\bar{x}^* - \bar{x}^{\min}) = 0, \quad \eta^{*\mathrm{T}}(\bar{x}^{\max} - \bar{x}^*) = 0, \\ &\mu^* \geq 0, \qquad\qquad \eta^* \geq 0. \end{aligned} \quad (7)$$

Remark 3.2: Supposing that Assumptions 1 and 3 hold, then the Slater condition is established for problem (5), implying that there exist Lagrangian multipliers satisfying KKT condition (7).

To deal with the inequality constraints of problem (5), by means of the logarithmic-barrier method, it follows that:

$$\begin{aligned} \text{minimize} \quad & \bar{f}(\bar{x}) = \sum_{i=1}^{n} f^i(x_i) + \frac{1}{2}\bar{x}^T L\bar{x} + \sum_{i=1}^{n} \theta(In[x_i^k - x_i^{k,\min}] + In[x_i^{k,\max} - x_i^k]) \\ \text{subject to} \quad & L\bar{x} = 0, \quad k=1,...,m \end{aligned} \quad (8)$$

where θ is a small enough positive real number.

Let $\bar{\lambda} \in \mathbb{R}^{mn}$ be the lagrangian multiplier of equality constraint $L\bar{x}$ in (8), then the Lagrange function of (8) is:

$$L(\bar{x}, \bar{\lambda}) = \sum_{i=1}^{n} f^i(x_i) + \frac{1}{2}\bar{x}^T L\bar{x} + \sum_{i=1}^{n} \theta(In[x_i^k - x_i^{k,\min}] + In[x_i^{k,\max} - x_i^k]) + \bar{\lambda}^T L\bar{x} \quad (9)$$

The corresponding Lagrange dual function is:

$$g(\bar{\lambda}) = \inf_{x \in \Omega} L(\bar{x}, \bar{\lambda}) = \inf_{x \in \Omega}\left[\sum_{i=1}^{n} f^i(x_i) + \frac{1}{2}\bar{x}^T L\bar{x} + \sum_{i=1}^{n} \theta(In[x_i^k - x_i^{k,\min}] + In[x_i^{k,\max} - x_i^k]) + \bar{\lambda}^T L\bar{x} \right] \quad (10)$$

Remark 3.3: Since the Slater condition deduced from Assumption 3.1 holds and there exists a nonempty interior point in Ω satisfying $x_i^{k,\min} < x_i^k < x_i^{k,\max}$, then the optimal solution of Lagrange dual problem (10) is equivalent to the optimization problem (8).

To solve the original optimization problem (2), the dynamics of multi-agent network is designed as:

$$\begin{aligned} \dot{x}_i^k \in {}& -\frac{\partial f(x_i)}{\partial x_i^k} - \sum_{i=1}^{n} a_{ij}(x_i^k - x_j^k) - \sum_{i=1}^{n} a_{ij}(\lambda_i^k - \lambda_j^k) + \theta\left(\frac{1}{x_i^k - x_i^{k,\min}} - \frac{1}{x_i^{k,\max} - x_i^k}\right) \\ \dot{\lambda}_i^k = {}& \sum_{i=1}^{n} a_{ij}(x_i^k - x_j^k), \quad i=1,...,n, \quad k=1,...,m \end{aligned} \quad (11)$$

Let $\overline{z} = \theta[(\frac{1}{x_1^1 - x_1^{1,min}} - \frac{1}{x_1^{1,max} - x_1^1}), (\frac{1}{x_1^2 - x_1^{2,min}} - \frac{1}{x_1^{2,max} - x_1^2}), ..., (\frac{1}{x_n^m - x_n^{m,min}} - \frac{1}{x_n^{m,max} - x_n^m})]^T$. Then (11) can be written in a compact form:

$$\dot{\overline{x}} \in -\nabla f(x) - L\overline{x} - L\overline{\lambda} + \overline{z}$$
$$\dot{\overline{\lambda}} = L\overline{x}$$

(12)

4. The Convergence Analysis

In this section, a complete convergence proof of dynamic system (11) [or (12)] is provided in the following theorems.

Lemma 4.1: Under the Assumptions 1-3, the trajectories of dynamic system (11) [or (12)] with any finite initial points are bounded.

Proof : In order to proof the stability of the dynamic system system (11) [or (12)], we construct a Lyapunov function as follow:

$$W(\overline{x}, \overline{\lambda}) = \frac{1}{2} \| \overline{x} - \overline{x}^* \|_2^2 + \frac{1}{2} \| \overline{\lambda} - \overline{\lambda}^* \|_2^2$$

(13)

Obviously, $W(\overline{x}, \overline{\lambda}) \geq 0$. In view of the chain rule, the time derivative of $W(\overline{x}, \overline{\lambda})$ along the trajectories of dynamic system (11) [or (12)] is :

$$\frac{dW(\overline{x}, \overline{\lambda})}{dt} = \dot{\overline{x}}^T (\overline{x} - \overline{x}^*) + \dot{\overline{\lambda}}^T (\overline{\lambda} - \overline{\lambda}^*)$$
$$= [-\nabla_{\overline{x}} L(\overline{x}, \overline{\lambda})]^T (\overline{x} - \overline{x}^*) + [\nabla_{\overline{\lambda}} L(\overline{x}, \overline{\lambda})]^T (\overline{\lambda} - \overline{\lambda}^*)$$

Since $L(\overline{x}, \overline{\lambda})$ is convex in $\overline{x}$, and concave in $\overline{\lambda}$, combining with the properties of convex function, we have:

$$[-\nabla_{\overline{x}} L(\overline{x}, \overline{\lambda})]^T (\overline{x} - \overline{x}^*) \leq L(\overline{x}^*, \overline{\lambda}) - L(\overline{x}, \overline{\lambda})$$

$$[\nabla_{\overline{\lambda}} L(\overline{x}, \overline{\lambda})]^T (\overline{\lambda} - \overline{\lambda}^*) \leq L(\overline{x}, \overline{\lambda}) - L(\overline{x}, \overline{\lambda}^*)$$

Based on Remark 3.3 and the Saddle point theorem, then

$$\frac{dW(\overline{x}, \overline{\lambda})}{dt} \leq L(\overline{x}^*, \overline{\lambda}) - L(\overline{x}, \overline{\lambda}) + L(\overline{x}, \overline{\lambda}) - L(\overline{x}, \overline{\lambda}^*)$$
$$= L(\overline{x}^*, \overline{\lambda}) - L(\overline{x}^*, \overline{\lambda}^*) + L(\overline{x}^*, \overline{\lambda}^*) - L(\overline{x}, \overline{\lambda}^*)$$
$$\leq 0$$

Hence, the Lyapunov function (13) is monotonic, non-increasing and has a lower bound, i.e., trajectories are bounded. Let $(\overline{x}(0), \overline{\lambda}(0))$ represent the initial point of $(\overline{x}, \overline{\lambda})$, then we have $W(\overline{x}, \overline{\lambda}) \leq W(\overline{x}(0), \overline{\lambda}(0))$, which indicates that there exists a positive invariant compact set such that the solution of (11) [or (12)] satisfy:

$$L_l = \{\overline{x} \in \mathbb{R}^{mn}, \overline{\lambda} \in \mathbb{R}^{mn} \mid W(\overline{x}, \overline{\lambda}) \leq l\}$$

(14)

Lemma 4.2: The trivial solution of the dynamic system (11) [or (12)] is asymptotically stable.

Proof: Define a function:

$$V(\dot{\overline{x}}, \dot{\overline{\lambda}}) = \frac{1}{2} <\dot{\overline{x}}, \dot{\overline{x}}> + \frac{1}{2} <\dot{\overline{\lambda}}, \dot{\overline{\lambda}}>$$

Then the time-derivative of $V(\dot{\overline{x}}, \dot{\overline{\lambda}})$ is:

$$\frac{dV(\dot{\overline{x}}, \dot{\overline{\lambda}})}{dt} = \frac{1}{2} <\dot{\overline{x}}, \ddot{\overline{x}}> + \frac{1}{2} <\dot{\overline{\lambda}}, \ddot{\overline{\lambda}}>$$
$$= \dot{\overline{x}}^T \{diag[-\nabla^2 f(x)]\dot{\overline{x}} - L\dot{\overline{x}} - L\dot{\overline{\lambda}} + diag[\nabla z]\dot{\overline{x}}\} + \dot{\overline{\lambda}}^T L\dot{\overline{x}}$$
$$= \dot{\overline{x}}^T diag[-\nabla^2 f(x)]\dot{\overline{x}} - \dot{\overline{x}}^T L\dot{\overline{x}} - \dot{\overline{x}}^T L\dot{\overline{\lambda}} + \dot{\overline{x}}^T diag[\nabla z]\dot{\overline{x}} + \dot{\overline{\lambda}}^T L\dot{\overline{x}}$$
$$= \dot{\overline{x}}^T diag[-\nabla^2 f(x)]\dot{\overline{x}} - \dot{\overline{x}}^T L\dot{\overline{x}} + \dot{\overline{x}}^T diag[\nabla z]\dot{\overline{x}}$$
$$< 0$$

(15)

where $diag[\nabla z] = diag[-\frac{\theta}{(x_1^1 - x_1^{1,min})^2} - \frac{\theta}{(x_1^{1,max} - x_1^1)^2}, ..., -\frac{\theta}{(x_n^m - x_n^{m,min})^2} - \frac{\theta}{(x_n^{m,max} - x_n^m)^2}]$.

With Assumption 3.2, at least one of the local objective functions $f^i(x), i = 1, ..., n$ of $f(x)$ has a positive definite Hessian matrix. Then $diag[-\nabla^2 f(x)]$ is negative definite, which implies $\dot{\overline{x}}^T \{diag[-\nabla^2 f(x)]\dot{\overline{x}} < 0$. Since L is a positive semi-definite matrix, then $-\dot{\overline{x}}^T L\dot{\overline{x}} \leq 0$. In addition, due to $diag[-\nabla z] < 0$, the last step of (14) holds.

Next, we shall prove $\Omega(\overline{x}, \overline{\lambda}) \subset \overline{Z}_V \bigcap L_l$. According to the definition of L_l , we have $\Omega(\overline{x}, \overline{\lambda}) \subset L_l$. Noting that $W(\overline{x}(t), \overline{\lambda}(t))$ is monotonic, non-increasing and has a lower bound, so $\lim_{t\to\infty} W(\overline{x}(t), \overline{\lambda}(t)) = c \geq 0$. According to the continuity of $W(\overline{x}, \overline{\lambda})$, we have $W(\overline{x}, \overline{\lambda}) = c$. Denoting $\psi(\overline{x}, \overline{\lambda}, t)$ as a solution of $\Omega(\overline{x}, \overline{\lambda})$, which satisfies $\psi(\overline{x}, \overline{\lambda}, 0) = y, \forall y \in \Omega(\overline{x}, \overline{\lambda})$, then $W(\psi(\overline{x}, \overline{\lambda}, t)) = c$ for any t . Thus, it gives

$$\frac{d}{dt} W(\psi(\overline{x}, \overline{\lambda}, t)) = 0$$

For almost all t , $0 \in \dot{\overline{V}}(\psi(\overline{x}, \overline{\lambda}, t))$. It can be concluded that there exists $\{t_i\}, t_i \to 0$ satisfying $\psi(\overline{x}, \overline{\lambda}, t_i) \in Z_V$. By the

continuity of ψ, it yields $\lim\limits_{i\to\infty}\psi(\bar{x},\bar{\lambda},t_i)=\psi(\bar{x},\bar{\lambda},0)=y\in\bar{Z}_V$. Hence, we have $\Omega(\bar{x},\bar{\lambda})\subset\bar{Z}_V$. Noting that $\Omega(\bar{x},\bar{\lambda})\subset M$ is weakly invariant set, then $dist((\dot{\bar{x}},\dot{\bar{\lambda}}),M)\to 0, t\to\infty$. From the above, according to Remark 2.1, we have that the trivial solution of the system (11) [or (12)] is asymptotically stable and it is also the optimal solution of (2).

5. Simulation

In this section, a simulation examples are presented to verify the theoretical analysis of the proposed second-order algorithm (11)[or (12)].

Example: Consider optimization problem (2) with $f^i(x)=ix_1-2ix_2-ix_3+|x_1+x_2+x_3|$, $x\in\mathbb{R}^3$ and $+\exp(ix_1-ix_2),(i=1,2,...,12)$

$\Omega^i=\{x\in\mathbb{R}^3:i-22\le x_1\le i+10, i/2-26\le x_2\le i/2+7, i/3-14\le x_3\le 2i+11\}$, where $f^i(x),(i=1,...,12)$ are non-smooth objective functions.

For the above optimization problem, we first assume that the network topology $\mathcal{G}$ is a cyclic connected network, as shown in Fig. 1(a). The connection weight is set to 1 if there exists a path between agent i and j, otherwise 0. The trajectories of twelve agents are shown in Fig. 1(b). It can be seen that all the agents converge to the same optimal solution $x^*=(-10,7.5,13)^T$ (approximate solution). Next, supposing that the network topology $\mathcal{G}$ is fully connected, as shown in Fig. 2(a), and the simulation results are shown in Fig. 2(b). It is clearly that the tighter the network connection, the faster the convergence rate is.
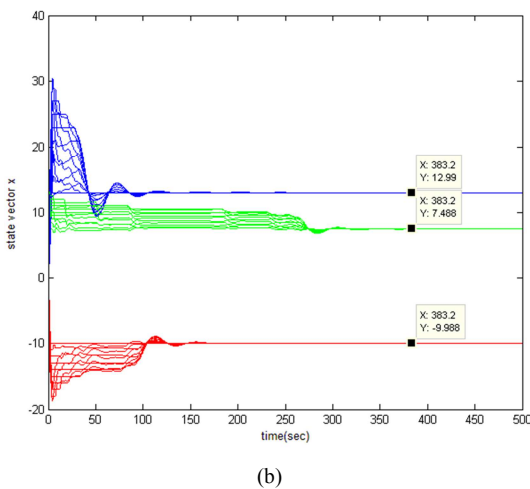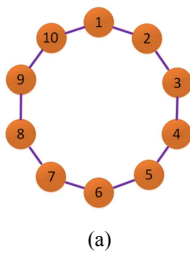

(a)

(b)

Fig. 1. *(a) A cyclic connected network. (b) The trajectories of the state vector* x *over a cyclic connected network.*
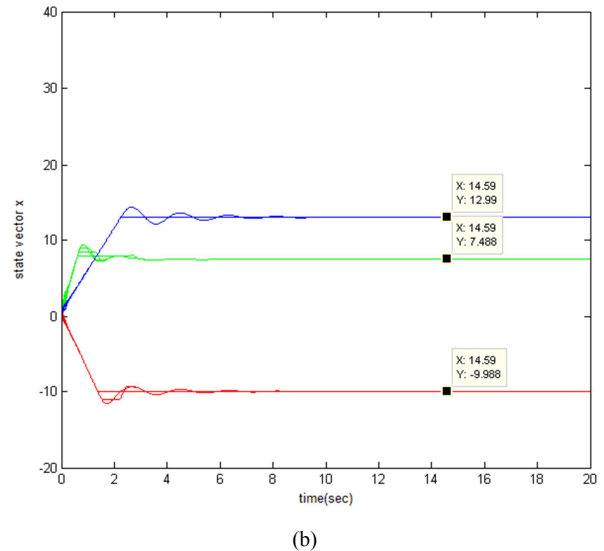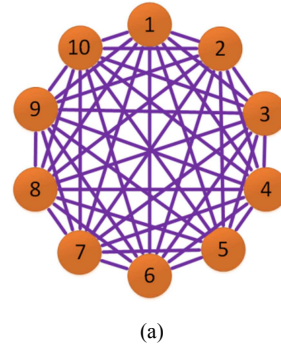

(a)

(b)

Fig. 2. *(a) A fully connected network. (b) The trajectories of the state vector* x *over a fully connected network.*

6. Conclusions

In this paper, a novel distributed continuous-time algorithm based on the KKT condition and the Lagrange multiplier method has been proposed for a distributed convex optimization problem. It aims to minimize the sum of the non-smooth local objective functions with local bound constraints over an undirected graph. Furthermore, the convergence analysis of the dynamical system is accomplished by using the Lyapunov stability theory and the hybrid LaSalle invariance principle of differential inclusion. The numerical simulation shows the performance of the proposed algorithm. In the future, our works may turn to the optimization problem with respect to directed topology and equality constraint, meanwhile analysis of its convergence speed.

Acknowledgements

References

[1] Han D, Mo Y, Wu J, et al. Stochastic event-triggered sensor schedule for remote state estimation. IEEE Transactions on Automatic Control, 2015, 60(10): 2661-2675.

[2] Li H, Liao X, Huang T, et al. Event-triggering sampling based leader-following consensus in second-order multi-agent systems. IEEE Transactions on Automatic Control, 2015, 60(7): 1998-2003.

[3] Olfati-Saber R, Fax A, Murray R M. Consensus and cooperation in networked multi-agent systems. Proceedings of the IEEE, 2007, 95(1): 215-233.

[4] Xia Z, Wang X, Sun X, et al. Steganalysis of least significant bit matching using multi-order differences. Security and Communication Networks, 2014, 7(8): 1283-1291.

[5] Lobel I, Ozdaglar A. Distributed subgradient methods for convex optimization over random networks. IEEE Transactions on Automatic Control, 2011, 56(6): 1291-1306.

[6] Guo P, Wang J, Geng X H, et al. A variable threshold-value authentication architecture for wireless mesh networks. Journal of Internet Technology, 2014, 15(6): 929-935.

[7] Nedić A, Ozdaglar A, Parrilo P A. Constrained consensus and optimization in multi-agent networks. IEEE Transactions on Automatic Control, 2010, 55(4): 922-938.

[8] Ren W, Beard R W. Distributed consensus in multi-vehicle cooperative control. Springer-Verlag, London, 2008.

[9] Li H, Chen G, Liao X, et al. Quantized data-based leader-following consensus of general discrete-time multi-agent systems, IEEE Transactions on Circuits and Systems II: Express Briefs, 2016, 63(4): 401-405.

[10] Wang J, Elia N. Control approach to distributed optimization. Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on. IEEE, 2010: 557-561.

[11] Gharesifard B, Cortés J. Distributed continuous-time convex optimization on weight-balanced digraphs. IEEE Transactions on Automatic Control, 2014, 59(3): 781-786.

[12] Liu S, Qiu Z, Xie L. Continuous-time distributed convex optimization with set constraints. Proceedings of the 19th IFAC World Congress, Cape Town, South Africa. 2014: 9762-9767.

[13] Li H, Liao X, Huang T, et al. Second-order global consensus in multi-agent networks with random directional link failure. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26(3): 565-575.

[14] Gungor V C, Hancke G P. Industrial wireless sensor networks: Challenges, design principles, and technical approaches. IEEE Transactions on Industrial Electronics, 2009, 56(10): 4258-4265.

[15] Shen J, Tan H W, Wang J, et al. A novel routing protocol providing good transmission reliability in underwater sensor networks. Journal of Internet Technology, 2015, 16(1): 171-178. 2015, 16(1): 171-178.

[16] Li H, Chen G, Huang T, et al. High-performance consensus control in networked systems with limited bandwidth communication and time-varying directed topologies. IEEE Transactions on Neural Networks and Learning Systems, 2016, DOI: 10.1109/TNNLS.2016.2519894.

[17] Blatt D, Hero III A O. Energy-based sensor network source localization via projection onto convex sets. IEEE Transactions on Signal Processing, 2006, 54(9): 3614-3619.

[18] Wen X, Shao L, Xue Y, et al. A rapid learning algorithm for vehicle classification. Information Sciences, 2015, 295: 395-406.

[19] Yi P, Hong Y, Liu F. Initialization-free distributed algorithms for optimal resource allocation with feasibility constraints and its application to economic dispatch of power systems. arXiv preprint arXiv:1510.08579, 2015.

[20] Li H, Liao X, Chen G, et al. Event-triggered asynchronous intermittent communication strategy for synchronization in complex dynamical networks. Neural Networks, 2015, 66: 1-10.

[21] Gu B, Sheng V S, Tay K Y, et al. Incremental support vector learning for ordinal regression. IEEE Transactions on Neural networks and learning systems, 2015, 26(7): 1403-1416.

[22] Wen G, Duan Z, Chen G, et al. Consensus tracking of multi-agent systems with Lipschitz-type node dynamics and switching topologies. IEEE Transactions on Circuits and Systems I: Regular Papers, 2014, 61(2): 499-511.

[23] Xie S, Wang Y. Construction of tree network with limited delivery latency in homogeneous wireless sensor networks. Wireless personal communications, 2014, 78(1): 231-246.

[24] Zhu M, Martínez S. On distributed convex optimization under inequality and equality constraints. IEEE Transactions on Automatic Control, 2012, 57(1): 151-164.

[25] Nedić A, Ozdaglar A. Distributed subgradient methods for multi-agent optimization. IEEE Transactions on Automatic Control, 2009, 54(1): 48-61.

[26] Bianchi P, Jakubowicz J. Convergence of a multi-agent projected stochastic gradient algorithm for non-convex optimization. IEEE Transactions on Automatic Control, 2013, 58(2): 391-405.

[27] Lou Y, Shi G, Johansson K H, et al. Approximate projected consensus for convex intersection computation: Convergence analysis and critical error angle. IEEE Transactions on Automatic Control, 2014, 59(7): 1722-1736.

[28] Zhu M, Martinez S. On distributed convex optimization under inequality and equality constraints via primal-dual subgradient methods. arXiv preprint arXiv:1001.2612, 2010.

[29] Yuan D, Xu S, Zhao H. Distributed primal–dual subgradient method for multiagent optimization via consensus algorithms. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2011, 41(6): 1715-1724.

[30] Li H, Chen G, Huang T, et al. Event-triggered distributed average consensus over directed digital networks with limited communication bandwidth. IEEE Transactions on Cybernetics, 2016, DOI: 10.1109/TCYB.2015.2496977.

[31] Ruszczyński A P. Nonlinear Optimization. Princeton university press, 2006.