Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Continuous-time Distributed Convex Optimization with Set Constraints ⋆

Shuai Liu *,** **Zhirong Qiu** ** **Lihua Xie** **

* *School of Control Science and Engineering, Shandong University, Jinan, P.R. China*
** *School of EEE, Nanyang Technological University, Singapore 639798 (e-mail: LIUS0025@ntu.edu.sg,QIUZ0005@e.ntu.edu.sg,elhxie@ntu.edu.sg)*

Abstract: We study a distributed convex optimization problem with set constraints. The objective function is a summation of strictly convex functions. Based on a multi-agent system formulation, we consider that each node is with continuous-time dynamics and can only access its local objective function. Meanwhile, each node is subject to a common convex set constraint. The nodes can exchange local information with their neighbor nodes. A distributed gradient-based control protocol is applied to each node. It is shown that when the nodes are connected as an undirected graph and the time-varying gains of the gradients satisfy a persistence condition, the states of all the nodes will converge to the unique optimal point subject to the set constraints. Numerical examples are provided to demonstrate the results.

Keywords: Convex Optimization; Set Constraint; Multi-agent System; Gradient; Distributed.

1. INTRODUCTION

Optimization is one of the hottest research topics in science and engineering due to its wide applications. Recently, distributed algorithm attracted more and more attentions of researchers. See, for example, distributed cooperative control Qu (2009), Su and Huang (2012), Liu et al. (2013); distributed estimation Michael and Robert (2004), Calafiore and Abrate (2009), distributed sensor deployment Khan et al. (2009) and distributed resource allocation Raynal (2013). Many advantages of distributed algorithm, such as scalability, high reliability and the ability of reducing communications, motivate researchers to explore distributed optimization algorithms.

In the past decade, many works have been dedicated to distributed optimization, Johansson et al. (2009), Nedic and Ozdaglar (2009), Nedic et al. (2010), Notarstefano and Bullo (2011), Zhu and Martinez (2012), Duchi et al. (2012). Distributed optimization is to optimize a function by connected nodes in a distributed way provided that each node can access partial information of the function. Since the nodes are connected and the cost function information can be broadcasted to every corner, the global optimum can be achieved under appropriate distributed algorithms. By introducing a Lagrangian function and applying the distributed dual averaging technique, the distributed optimization problem is investigated under different scenarios; see Yuan et al. (2011), Duchi et al. (2012), Zhu and Martinez (2012) and the reference therein. Based on the distributed primal-dual method, multi-agent optimization with inequality constraint is studied in Yuan et al. (2011). A distributed algorithm for optimization is provided in Zhu and Martinez (2012) by considering both equality and inequality constraints.

Recently, the gradient based method (Nedic and Bertsekas (2001), Rabbat and Nowak (2005)) has been widely used in distributed optimization since it can be often executed in a distributed fashion. By incorporating consensus approach, the distributed convex optimization problem is studied in Nedic and Ozdaglar (2009). An approximate optimal solution is given and the gap between the global optimal objective function and the one under the proposed algorithm is also derived. The optimization problem under random communication graph is considered in Lobel and Ozdaglar (2011). Under the condition that the link failures are independent and identically distributed, the subgradient algorithm can guarantee the almost sure convergence to the optimal set. In the recent years, there has been more studies on the constraint distributed optimization problem. Based on a randomized incremental subgradient method, in Johansson et al. (2009) a distributed convex optimization approach is proposed with each node subject to a common convex constraint. Asynchronous distributed gradient approach is investigated in Srivastava and Nedic (2011), Nedic (2011). Using broadcasting approach, the optimization problem is studied in Nedic (2011) based on the subgradient method. This approach is robust to link failure and is asynchronous which allows more freedom for each agent to apply it. In Srivastava et al. (2010), transmission noises are considered when solving the dis-

⋆ This research is funded by the Republic of Singapores National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore.The research is also supported by National Natural Science Foundation of China (NSFC) under Grants 61203045, 61304045.

tributed optimization problem. In Bianchi and Jakubowicz (2013), the non-convex optimization is solved by applying a stochastic gradient algorithm.

Most of the current works on gradient based distributed optimization are devoted to discrete-time systems. In some practical systems, the nodes are with continuous-time dynamics which requires continuous-time optimization algorithm. However, there are many challenges when developing continuous-time algorithms. For example, the solutions of differential equations may not exist or may not be unique. A linear programming problem is solved in Brockett (1988) by introducing continuous-time gradient method. In Wang and Elia (2010), a continuous-time algorithm is proposed for distributed convex optimization without constraint. It is proved that the proposed algorithm is robust to additive noises due to the introduction of an integral term.

In this paper, we shall consider the distributed convex optimization problem with set constraints. The multi-agent system is introduced to perform the distributed gradient method in a continuous-time manner. The objective function is a summation of different convex functions and each node can only access one convex function. The discrete-time case can be found in Nedic et al. (2010). By local information exchange, we can control the states of the nodes to converge to the unique optimum within the constraint set. The control input can be divided into three parts: local information exchange, local subgradient and set projection. Since the global optimum (without set constraints) may not be in the constraint set, the effect of local gradient needs to diminish such that the set projection is not offset and the state of the nodes can finally fall in the constraint set. On the other hand, the effect of local gradient needs to be persistent such that the optimum can be achieved. Therefore, the gain of the gradient in the algorithm should satisfy a persistence condition which is less restrictive than the stochastic approximation condition that requires square integrability in continuous-time setting Li and Zhang (2009) or square summability in discrete-time setting Nedic et al. (2010), Liu et al. (2011).

Some notations are listed below which will be used throughout this paper. Given an arbitrary matrix M, $[M]_{i,j}$ and M' denote its (i,j)-th entry and transpose respectively. $\mathrm{diag}\{x_1,\cdots,x_n\}$ is the diagonal matrix with the i-th diagonal component being x_i. Given a function f, the gradient of f at x is denoted by $\nabla f(x)$. $\mathcal{R}$, $\mathcal{R}^n$ and $\mathcal{R}^{m\times n}$ stand for the set of real numbers, the set of n dimensional real vectors and the set of $m\times n$ matrices. $\mathbf{1}_n$ is an n dimensional vector with each component being 1 and $J=\frac{1}{n}\mathbf{1}_n\mathbf{1}_n'$. Given $x,y\in\mathcal{R}^n$, $\langle x,y\rangle$ and $|x-y|$ are their inner product and the corresponding distance, i.e. $\langle x,y\rangle = x'y$, $|x-y|=\sqrt{\langle x-y,x-y\rangle}$. $B(x,r)$ means an open ball centered at x with radius r, i.e. $B(x,r)=\{y:|y-x|<r\}$. The norm induced by the above inner product in $\mathcal{R}^{n\times n}$ is defined as $\|M\|=\sup_{|x|\neq 0}\frac{|Mx|}{|x|}$ for $M\in\mathcal{R}^{n\times n}$. Kronecker product of X and Y is denoted by $X\otimes Y$.

2. PRELIMINARIES

Some preliminaries in graph theory and convex analysis will be reviewed in this section.

2.1 Preliminaries in Graph Theory

An undirected graph denoted by $\mathcal{G}$, with a node set $\mathcal{V}=\{1,2,\cdots,n\}$ and an unordered edge set $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ where self loop (i,i) is excluded, is often used to model communications among nodes. $(i,j)\in\mathcal{E}$ means that node i and node j can exchange information with each other. The set of neighbours of node i is denoted by $\mathcal{N}_i=\{j\mid j\in\mathcal{V},\ (j,i)\in\mathcal{E}\}$. A path from node i to node j is defined by a sequence $(i,l_1),(l_1,l_2),\ldots,(l_p,j)\in\mathcal{E}(\mathcal{G})$, where $i,j,l_1,\ldots,l_p$ are distinct nodes. $\mathcal{G}$ is called connected if there exists a path between any pair of distinct nodes. In addition, a weighted adjacency matrix $A\in\mathcal{R}^{n\times n}$ with $[A]_{i,j}=a_{i,j}$ will be used to describe $\mathcal{G}$, where $a_{i,j}=a_{j,i}\geq 0$ and $a_{i,j}>0$ if and only if $(i,j)\in\mathcal{E}$. Moreover, the Laplacian matrix is defined as $L=D-A$, where $D=diag\{D_1,\cdots,D_n\}$ and $D_i\triangleq\sum_{j\in\mathcal{N}_i}a_{i,j}$ is the in-degree of node i.

By denoting all the eigenvalues of L as λ_i, $i=1,2,\ldots,n$, some properties of the Laplacian matrix are recalled below:

Lemma 1. For an undirected graph $\mathcal{G}$, suppose that the eigenvalues of the Laplacian matrix $L\in\mathcal{R}^{n\times n}$ of $\mathcal{G}$ satisfy $\lambda_1\leq\cdots\leq\lambda_n$, we have the following properties:

(1) $\lambda_1=0$. $\lambda_2>0$ if and only if $\mathcal{G}$ is connected.
(2) When $\mathcal{G}$ is connected, it has $\lim_{t\to\infty}e^{-Lt}=J$ and $\left\|\int_0^\infty e^{-Lt}(I-J)dt\right\|\leq\frac{1}{\lambda_2}$.

Proof. The first property is straightforward by recalling the graph theory in Godsil and Royle (2001). We shall focus on the proof of (2). Since graph $\mathcal{G}$ is undirected, we know that L is symmetric and there exists an orthogonal matrix $P=[\frac{1}{\sqrt{n}}\mathbf{1}_n\ \phi]$ with $\phi\in\mathcal{R}^{n\times(n-1)}$ orthogonal to $\mathbf{1}_n$ such that $L=P\Lambda P'$, where $\Lambda=\mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}$. According to the first property, the connectedness of the graph implies $\lambda_i>0$, $i=2,\ldots,n$. Then we have

$$\lim_{t\to\infty}e^{-Lt}=P\lim_{t\to\infty}e^{-\Lambda t}P'=J.$$

The first equation has been proved.

Note that $P'(I-J)=[0\ \phi]'$, then we have

$$\begin{aligned}\int_0^\infty e^{-Lt}(I-J)dt &= \int_0^\infty Pe^{-\Lambda t}P'(I-J)dt\\ &= \int_0^\infty \phi\mathrm{diag}\{e^{-\lambda_2 t}\ \cdots\ e^{-\lambda_n t}\}\phi' dt\\ &= \phi\mathrm{diag}\left\{\frac{1}{\lambda_2}\ \cdots\ \frac{1}{\lambda_n}\right\}\phi'.\end{aligned}$$

Since $\|\phi\|=\|\phi'\|=1$, the above inequality implies the result.

2.2 Convex Analysis

Given a set $\mathcal{C}$, $\mathcal{C}$ is called convex if $\lambda x+(1-\lambda)y\in\mathcal{C}$ for any $x,y\in\mathcal{C}$ and $0\leq\lambda\leq 1$. For a closed convex set $\mathcal{C}\in\mathbb{R}^m$, $P_{\mathcal{C}}(x)\in\mathcal{C}$ is the projection of x onto $\mathcal{C}$, uniquely satisfying

$$|x-P_{\mathcal{C}}(x)|=\inf_{v\in\mathcal{C}}|x-v|\triangleq|x|_{\mathcal{C}}.$$

A function $f(\cdot):\mathcal{R}^m\to\mathcal{R}$ is convex if $\forall x\neq y\in\mathcal{R}^m$ and $0<\lambda<1$,

$$f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y). \quad (1)$$

f is strictly convex when (1) holds as a strict inequality. For a convex function f, if it is differentiable, then it is continuously differentiable. Moreover, its gradient ∇f satisfies that

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle, \ \forall x, y$$

and is said to be monotone in the sense that

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0, \ \forall x, y.$$

Before closing this section, the following lemmas are introduced.

Lemma 2. (Aubin and Frankowska (2009)) Given a closed convex set $\mathcal{C} \subset \mathcal{R}^m$, $\forall x \in \mathcal{R}^m$, $\forall y \in \mathcal{C}$, it has

$$\langle P_{\mathcal{C}}(x) - x, P_{\mathcal{C}}(x) - y \rangle \leq 0. \quad (2)$$

Lemma 3. (Shi et al. (2012)) Given a closed convex set $\mathcal{C} \subset \mathcal{R}^m$ and $x, y \in \mathcal{R}^m$, we have

$$\langle x - P_{\mathcal{C}}(x), y - x \rangle \leq |x|_{\mathcal{C}}(|y|_{\mathcal{C}} - |x|_{\mathcal{C}}). \quad (3)$$

2.3 Non-smooth Analysis

Given a continuous function $f : (t_0, t_f) \to \mathcal{R}$, the upper Dini derivative $D^+ f(t)$ is defined as

$$D^+ f(t) = \limsup_{h \to 0^+} \frac{f(t+h) - f(t)}{h}.$$

A continuous function $f(t)$ is non-increasing over (t_0, t_f) if and only if $D^+ f(t) \leq 0$, $t \in (t_0, t_f)$.

The following lemmas are introduced.

Lemma 4. (Shi et al. (2013)) Let $V_i(t, x) : \mathcal{R} \times \mathcal{R}^m \to \mathcal{R}$, $i = 1, \ldots, n$ be continuously differentiable and $V(t, x) = \max_{i=1,\ldots,n} V_i(t, x)$. If $\mathcal{I}(t) = \{i : V(t, x) = V_i(t, x)\}$ denotes the set of indices where the maximum is reached at t, then $D^+ V(t, x(t)) = \max_{i \in \mathcal{I}(t)} \dot{V}_i(t, x(t))$.

Lemma 5. Let E be an open set in $\mathbb{R}^2$ and $g(t, u)$ is a continuous scalar function defined on E. Assume that $v(t)$ and $w(t)$ are continuous on $[t_0, t_0 + a)$, with $(t, v(t))$, $(t, w(t)) \in E$. If $v(t_0) \leq w(t_0)$ and the inequalities below hold for $t \in [t_0, t_0 + a)$:

$$D^+ v(t) \leq g(t, v(t)), \ D^+ w(t) \geq g(t, w(t)),$$

then $v(t) \leq w(t), t \in (t_0, t_0 + a)$.

3. PROBLEM STATEMENT

We consider the following optimization problem

$$\min F(x) = \sum_{i=1}^{n} f_i(x)$$
$$s.t. \ x \in \mathcal{X}, \quad (4)$$

where $F(\cdot)$ is the cost function to be minimized, $\mathcal{X} \subset \mathcal{R}^m$ is a set constraint.

We shall solve this optimization problem in a distributed way by introducing multi-agent systems. Suppose we have n nodes with each node i assigned with a state value $x_i \in \mathcal{R}^m$. Node i can only access the information of f_i and the node needs to exchange information with its neighbor nodes such that the common optimum point in $\mathcal{X}$ is achieved. Then problem (4) can be reformulated as

$$\min \sum_{i=1}^{n} f_i(x_i)$$
$$s.t. \ x_1 = \cdots = x_n, \ x_i \in \mathcal{X}. \quad (5)$$

The assumptions of problem (5) are introduced as follows:

A1: $\mathcal{X}$ is convex and compact.

A2: $f_i(\cdot)$, $i = 1, \ldots, n$ is strictly convex and differentiable.

A3: All the nodes are connected as an undirected fixed graph.

We shall design a gradient based distributed algorithm for the nodes such that problem (4) is solved. Noticing that each node should solve a constrained optimization problem and meanwhile reach consensus with other nodes, the algorithm is thus proposed as follows.

$$\dot{x}_i(t) = \sum_{j=1}^{n} a_{i,j}(x_j - x_i) - \alpha(t) \nabla f_i(x_i) + P_{\mathcal{X}}(x_i) - x_i,$$
$$i = 1, \ldots, n, \quad (6)$$

where $a_{i,j} \geq 0$ is the weighting on the communication edge, $P_{\mathcal{X}}(\cdot)$ is the projection onto $\mathcal{X}$, and $\alpha(t) > 0$ is a continuous function satisfying the following assumption:

A4: Persistence condition

$$\int_0^\infty \alpha(t) dt = +\infty, \ \lim_{t \to \infty} \alpha(t) = 0. \quad (7)$$

Since the functions $f_i(\cdot)$, $i = 1, \ldots, n$ are strictly convex and the constraint set is closed and convex, there must exist a unique point $x^* \in \mathcal{X}$ such that

$$x^* = \arg\min_{v \in \mathcal{X}} F(v) \quad (8)$$

In the rest of this paper, we shall prove that under algorithm (6), $x_i(t)$ asymptotically converges to x^*, i.e.

$$\lim_{t \to \infty} x_i(t) = x^*, \ i = 1, \ldots, n. \quad (9)$$

4. CONVERGENCE ANALYSIS

In this section, we shall provide the main result and analyze the convergence.

Theorem 6. Suppose A1-A4 hold. The distributed optimization algorithm (6) converges to the optimum of (5) if the communication graph is connected.

The analysis of the convergence in Theorem 6 is divided into three steps: global set convergence, consensus analysis and optimal point convergence.

4.1 Global set convergence

In this section we will prove that

$$\lim_{t \to \infty} |x_i(t)|_{\mathcal{X}} = 0, \ i = 1, \ldots, n. \quad (10)$$

We shall first establish the global existence of $x_i(t)$, whose local existence has been guaranteed by the continuity of the righthand side of (6). Denote $d_i(t) = |x_i(t)|^2_{\mathcal{X}}$ and $d(t) = \max_{i=1,\ldots,n} d_i(t)$. According to (6) and Lemma 3, and recalling the monotonicity of ∇f_i, it has

$$\begin{aligned}\dot{d}_i(t) &= 2\langle x_i - P_{\mathcal{X}}(x_i), \dot{x}_i(t) \rangle \\ &\leq 2 \sum_{j=1}^{n} a_{i,j} |x_i|_{\mathcal{X}} (|x_j|_{\mathcal{X}} - |x_i|_{\mathcal{X}}) \\ &\quad - 2\alpha(t) \langle x_i - P_{\mathcal{X}}(x_i), \nabla f_i(P_{\mathcal{X}}(x_i)) \rangle - 2|x_i|^2_{\mathcal{X}}. \quad (11)\end{aligned}$$

By denoting $\mathcal{I}(t)$ as the index set containing the nodes which attain $d(t)$ at t, the following inequality holds according to Lemma 4,

$$\begin{aligned} D^+d(t) &= \max_{i\in\mathcal{I}(t)} \dot{d}_i(t) \\ &\le \max_{i\in\mathcal{I}(t)} [-2|x_i(t)|^2_{\mathcal{X}} \\ &\quad +2\alpha(t)\langle x_i - P_{\mathcal{X}}(x_i), \nabla f_i(P_{\mathcal{X}}(x_i))\rangle] \\ &\le -2d(t) + 2\alpha(t)s^*\sqrt{d(t)}, \end{aligned} \tag{12}$$

where $s^* = \max\{\nabla f_i(y) : y \in \mathcal{X}, i = 1,\dots,n\} < \infty$. In addition, by the continuity of $\alpha(t)$ and the assumption that $\lim_{t\to\infty}\alpha(t) = 0$, there exists $M > 0$ such that $\sup_{t\ge 0}\alpha(t) < M$. On the other hand, note that the solution of the following equation

$$\begin{cases} \dot{y}(t) = 2(-y + s^*M\sqrt{y}) \\ y(0) = d(0) \ge 0 \end{cases} \tag{13}$$

is given by

$$y(t) = \begin{cases} [s^*M + e^{-t}(\sqrt{d(0)} - s^*M)]^2, & d(0) > 0, \\ 0, & d(0) = 0. \end{cases} \tag{14}$$

Then by comparison principle in Lemma 5, we have

$$\begin{aligned} d(t) &< [s^*M + e^{-t}(\sqrt{d(t_0)} - s^*M)]^2 \\ &\le \max\{d(0), (s^*M)^2\}. \end{aligned} \tag{15}$$

Now the boundedness of $d(t)$ guarantees that the solution of (6) exists over $[0, +\infty)$. Furthermore, for any $\varepsilon > 0$, there exists a time constant t_ε such that $\forall t \ge t_\varepsilon$, $\alpha(t) \le \varepsilon$. By replacing M and 0 with ε and t_ε respectively in (13), we have $\forall t \ge t_\varepsilon$,

$$d(t) < [s^*\varepsilon + e^{-(t-t_\varepsilon)}(\sqrt{d(t_\varepsilon)} - s^*\varepsilon)]^2,$$

which implies that $\limsup_{t\to\infty} d(t) \le (s^*\varepsilon)^2$. The conclusion is established by the arbitrariness of ε.

4.2 Consensus analysis

In this section we will prove that

$$\lim_{t\to\infty}(x_i(t) - x_j(t)) = 0. \tag{16}$$

Denote the communication graph of the nodes by $\mathcal{G}$, its associated weighted adjacency matrix by $[A]_{i,j} = a_{i,j}$, and the associated Laplacian matrix by L. Letting $X(t) = [x_1'(t) \ \cdots \ x_n'(t)]' \in \mathcal{R}^{mn}$, it follows that

$$\dot{X}(t) = -(L\otimes I_m)X(t) + \delta(t), \tag{17}$$

where $\delta(t) = -\mathrm{col}\{\alpha(t)\nabla f_1(x_1) \ \cdots \ \alpha(t)\nabla f_n(x_n)\} + \mathrm{col}\{P_{\mathcal{X}}(x_1) - x_1 \ \cdots \ P_{\mathcal{X}}(x_n) - x_n\}$. In light of (10), we know that $\lim_{t\to\infty}(P_{\mathcal{X}}(x_i) - x_i) = 0$, $i = 1,\dots,n$, which together with $\lim_{t\to\infty}\alpha(t) = 0$ and the boundedness of the gradients implies that $\lim_{t\to\infty}|\delta(t)| = 0$. Then, $\forall\varepsilon > 0$ there exists a time instant t_ε such that $|\delta(t)| < \varepsilon$, $\forall t \ge t_\varepsilon$. Define $\tilde{X}(t) = [(I-J)\otimes I_m]X(t)$. Then we have

$$\dot{\tilde{X}}(t) = -(L\otimes I_m)\tilde{X}(t) + [(I-J)\otimes I_m]\delta(t). \tag{18}$$

According to Lemma 1, the following inequality holds

$$\begin{aligned} \lim_{t\to\infty}|\tilde{X}(t)| &= \lim_{t\to\infty}\left| e^{-(L\otimes I_m)t}\tilde{X}(0) \right. \\ &\quad \left. + \int_0^t e^{-(L\otimes I_m)(t-\tau)}[(I-J)\otimes I_m]\delta(\tau)d\tau \right| \\ &\le |(J\otimes I_m)\tilde{X}(0)| + \lim_{t\to\infty}\left|\int_0^{t_\varepsilon}[J(I-J)\otimes I_m]\delta(\tau)d\tau\right| \\ &\quad +\varepsilon\lim_{t\to\infty}\left\|\int_{t_\varepsilon}^t e^{-(L\otimes I_m)(t-\tau)}[(I-J)\otimes I_m]d\tau\right\| \\ &\le \frac{\varepsilon}{\lambda_2}, \end{aligned} \tag{19}$$

where λ_2 is the second smallest eigenvalue of the Laplacian matrix. Since the graph is connected, $\lambda_2 > 0$. The arbitrariness of ε implies that $\lim_{t\to\infty}\tilde{X}(t) = 0$. From the definition of $\tilde{X}(t)$ we know that

$$\lim_{t\to\infty} X(t) = (J\otimes I_m)\lim_{t\to\infty} X(t),$$

which implies (16).

4.3 Optimal point convergence

In this subsection we will prove (9). The proof is carried out by contradiction. According to the convergence results in the last two subsections, i.e. (10) and (16), for any $\varepsilon > 0$, there exists a time instant t_ε such that $\forall t \ge t_\varepsilon$,

$$|\bar{x}(t)|_{\mathcal{X}} + \max_{i=1,\dots,n}|x_i(t) - \bar{x}(t)| \le \frac{\varepsilon}{n},$$

where $\bar{x}(t) = \frac{1}{n}\sum_{i=1}^n x_i(t)$. The uniformly boundedness of $x_i(t)$, $i = 1,\dots,n$ implies that

$$|F(P_{\mathcal{X}}(\bar{x})) - F(\bar{x})| + \sum_{i=1}^n |f_i(\bar{x}) - f_i(x_i)| \le s\varepsilon, \tag{20}$$

where $s = \max\{\nabla f_i(y) : |y|_{\mathcal{X}} \le \max(d(0), (s^*M)^2), i = 1,\dots,n\} < \infty$.

Denote $l_i(t) = |x_i(t) - x^*|^2$ and $l(t) = \sum_{i=1}^n l_i(t)$. Then we have

$$\begin{aligned} \dot{l}_i(t) &= 2\langle x_i - x^*, \dot{x}_i(t)\rangle \\ &= 2\Big\langle x_i - x^*, \sum_{j=1}^n a_{i,j}(x_j - x_i) \\ &\quad -\alpha(t)\nabla f_i(x_i) + P_{\mathcal{X}}(x_i) - x_i \Big\rangle \\ &\le 2\sum_{j=1}^n a_{i,j}\langle x_i - x^*, x_j - x_i\rangle \\ &\quad +2\alpha(t)\left(f_i(x^*) - f_i(x_i)\right), \end{aligned}$$

and

$$\begin{aligned} \dot{l}(t) &\le 2\alpha(t)\sum_{i=1}^n (f_i(x^*) - f_i(x_i)) \\ &= 2\alpha(t)[F(x^*) - F(P_{\mathcal{X}}(\bar{x})) + F(P_{\mathcal{X}}(\bar{x})) - F(\bar{x}) \\ &\quad + \sum_{i=1}^n (f_i(\bar{x}) - f_i(x_i))] \\ &\le 2\alpha(t)[F(x^*) - F(P_{\mathcal{X}}(\bar{x}(t))) + s\varepsilon], \ t \ge t_\varepsilon. \end{aligned} \tag{21}$$

By considering that x^* is a unique global minimum of F on $\mathcal{X}$ and $\int_0^\infty \alpha(t)dt = +\infty$, we can prove that $\liminf_{t\to\infty}|P_{\mathcal{X}}(\bar{x}) - x^*| = 0$. Moreover, by global set convergence and consensus result we obtain $\liminf_{t\to\infty} l(t) = 0$.

On the other hand, assume that $\limsup_{t\to\infty} l(t) = \delta > 0$. Noticing that $\liminf_{t\to\infty} l(t) = 0$, $E = \{\tau : \dot{l}(\tau) = 0, l(\tau)$ is a local maximum$\}$ consists of infinitely many points. Besides, there exists an infinite sequence $\{t_{n_k}\} \subseteq E$ such that $\delta = \limsup_{k\to\infty} l(t_{n_k})$. However, by (21) $\dot{l}(t_{n_k}) = 0$ implies $F(P_{\mathcal{X}}(\bar{x}(t_{n_k}))) - F(x^*) \le s\varepsilon$ when $t_{n_k} \ge t_\varepsilon$. By the arbitrariness of ε it can be seen that $P_{\mathcal{X}}(\bar{x}(t_{n_k})) \to x^*$ and $\limsup_{k\to\infty} l(t_{n_k}) = 0$, a contradiction. Hence $\limsup_{t\to\infty} l(t) = 0$, which entails (9).

5. APPLICATION TO BUILDING TEMPERATURE REGULATION

In this section, we shall consider a building temperature regulation problem and apply the proposed algorithm for thermal comfort optimization. Due to the thermal conduction and convection among the adjacency zones, the temperature of each zone interacts with each other. By ignoring the heat transfer from external wall of the building and radiation, the indoor zone temperature dynamic model is given as follows Wang (1999):

$$M_i c_p \dot{T}_i(t) = \sum_{j\in\mathcal{N}_i} m_{i,j} c_p (T_j - T_i) + u_i(t),\ i = 1, \ldots, n, \tag{22}$$

where T_i (^oC) is the temperature of zone i, u_i (kJ) is the heat exchange rate provided by supply air through terminal unit in zone i, M_i (kg) is the mass of air in zone i, $m_{i,j}$ (kg/s) represents mass flow rate between zone i and j, c_p ($kJ/kg\cdot{}^oC$) is specific heat of air . We assume that $M_i = 100kg$, $m_{i,j} = 2kg/s$, $c_p = 1kJ/kg\cdot{}^oC$. The heat exchange rate is affected by the supply air flow rate and temperature. Therefore, we can control $u_i(t)$ by tuning the air handling unit and terminal unit. For simplicity, we consider only 4 zones and the configuration is given in Fig. 1. From the model it is clear that the connection

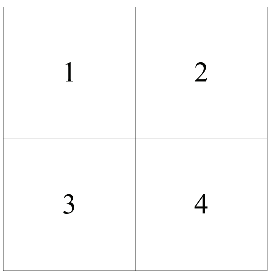

Fig. 1. Configuration of Zones.

graph is undirected and can be describe in Fig. 2. The

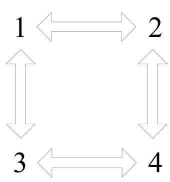

Fig. 2. Connection graph of Zones.

corresponding adjacency matrix A and Laplacian matrix L are given below:

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix},\ L = \begin{bmatrix} -2 & 1 & 1 & 0 \\ 1 & -2 & 0 & 1 \\ 1 & 0 & -2 & 1 \\ 0 & 1 & 1 & -2 \end{bmatrix}.$$

The reference temperature $T_{i,0}$, $i = 1, \ldots, 4$ are given below:

$$T_{1,0} = 25^oC,\ T_{2,0} = 25.5^oC,$$
$$T_{3,0} = 26^oC,\ T_{4,0} = 26.5^oC.$$

The cost function for each zone is defined as $f_i = (T_i - T_{i,0})^2$, which can be understood as thermal comfort penalty. On the other hand, due to thermal comfort constraint and cooling/heating capacity of air-conditioning system, there is temperature constraint for all of the zones, which is defined as $\mathcal{T} = [23^oC, 28^oC]$. We set the initial conditions as follows:

$$T_1(0) = 29^oC,\ T_2(0) = 30^oC,$$
$$T_3(0) = 31^oC,\ T_4(0) = 32^oC.$$

Due to the inherent temperature dynamic coupling of zones, according to (6), the input $u_i(t)$ can be designed as

$$u_i(t) = M_i c_p[-2\alpha(t)(T_i(t) - T_{i,0}) + P_{\mathcal{T}}(T_i) - T_i(t)].$$

We choose $\alpha(t) = \frac{1}{(t+1)^{0.8}}$ which satisfies (7) and is not square integrable. The total cost function to be minimized is $F = \sum_{i=1}^4 f_i(T_i)$. It is shown in Fig. 3 that all of the temperature converges to the global optimum 25.75^oC. Next, we change the temperature set-point of each room

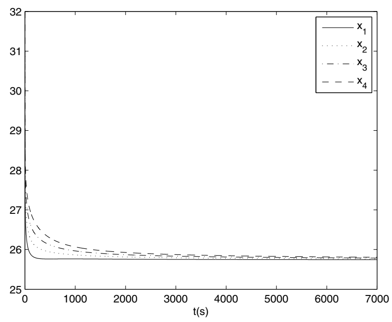

Fig. 3. Temperature trajectories with set-point $T_{1,0} = 25^oC,\ T_{2,0} = 25.5^oC,\ T_{3,0} = 26^oC,\ T_{4,0} = 26.5^oC$.

as follows:

$$T_{1,0} = 26^oC,\ T_{2,0} = 27^oC,$$
$$T_{3,0} = 32^oC,\ T_{4,0} = 33^oC.$$

Note that the global optimum is greater than the upper bound 28^oC. According to Theorem 6 we know that the temperature will converge to 28^oC, which is illustrated in Fig. 4.

6. CONCLUSION

In this paper we have studied the distributed convex optimization problem with set constraints. Each node is assigned with a state and all of the nodes are connected as an undirected graph. Under the proposed algorithm all the states were shown to asymptotically converge to the unique optimal value within the set constraint. In the future, we shall consider the case that different nodes are subject to different set constraints with nonempty intersection. Directed communication topology will also be considered in the future.

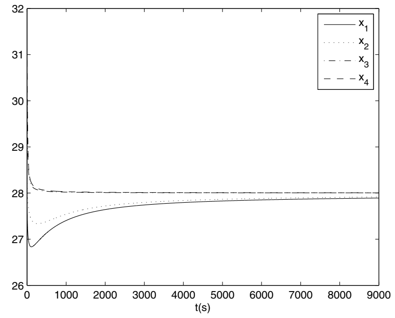

Fig. 4. Temperature trajectories with set-point $T_{1,0} = 26^oC,\ T_{2,0} = 27^oC,\ T_{3,0} = 32^oC,\ T_{4,0} = 33^oC$.

ACKNOWLEDGEMENTS

The valuable suggestions of Professor Yiguang Hong and Karl Henrik Johansson are gratefully acknowledged.

REFERENCES

Aubin, J.P. and Frankowska, H. (2009). Differential inclusions. In *Set-Valued Analysis*, Modern Birkhauser Classics. Birkhauser Boston.

Bianchi, P. and Jakubowicz, J. (2013). Convergence of a multi-agent projected stochastic gradient algorithm for non-convex optimization. 58(2), 391–405.

Brockett, R.W. (1988). Dynamical systems that sort lists, diagonalize matrices and solve linear programming problems. In *Proceedings of the 27th IEEE Conference on Decision and Control*, 799–803.

Calafiore, G.C. and Abrate, F. (2009). Distributed linear estimation over sensor networks. *International Journal of Control*, 82(5), 868 – 882.

Duchi, J.C., Agarwal, A., and Wainwright, M.J. (2012). Dual averaging for distributed optimization: Convergence analysis and network scaling. *IEEE Transactions on Automatic Control*, 57(3), 592–606.

Godsil, C.D. and Royle, G. (2001). *Algebraic Graph Theory*. Springer, New York.

Johansson, B., Rabi, M., and Johansson, M. (2009). A randomized incremental subgradient method for distributed optimization in networked systems. *SIAM Journal on Optimization*, 20(3), 1157–1170.

Khan, U.A., Kar, S., and Moura, J.M.F. (2009). Distributed sensor localization in random environments using minimal number of anchor nodes. *IEEE Transactions on Signal Processing*, 57(5), 2000–2016.

Li, T. and Zhang, J.F. (2009). Mean square average-consensus under measurement noises and fixed topologies: Necessary and sufficient conditions. *Automatica*, 45(8), 1929–1936.

Liu, S., Li, T., Xie, L., Fu, M., and Zhang, J.F. (2013). Continuous-time and sampled-data-based average consensus with logarithmic quantizers. *Automatica (in press)*.

Liu, S., Xie, L., and Zhang, H. (2011). Distributed consensus for multi-agent systems with delays and noises in transmission channels. *Automatica*, 47(5), 920–934.

Lobel, I. and Ozdaglar, A. (2011). Distributed subgradient methods for convex optimization over random networks. *IEEE Transactions on Automatic Control*, 56(6), 1291–1306.

Michael, R. and Robert, N. (2004). Distributed optimization in sensor networks.

Nedic, A. (2011). Asynchronous broadcast-based convex optimization over a network. *IEEE Transactions on Automatic Control*, 56(6), 1337–1351.

Nedic, A. and Bertsekas, D. (2001). Incremental subgradient methods for nondifferentiable optimization. *SIAM Journal on Optimization*, 12(1), 109–138.

Nedic, A. and Ozdaglar, A. (2009). Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1), 48–61.

Nedic, A., Ozdaglar, A., and Parrilo, P.A. (2010). Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control*, 55(4), 922–938.

Notarstefano, G. and Bullo, F. (2011). Distributed abstract optimization via constraints consensus: Theory and applications. *IEEE Transactions on Automatic Control*, 56(10), 2247–2261.

Qu, Z. (2009). *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. Springer Verlag.

Rabbat, M.G. and Nowak, R.D. (2005). Quantized incremental algorithms for distributed optimization. *IEEE Journal on Selected Areas in Communications*, 23(4), 798–808.

Raynal, M. (2013). *Distributed Resource Allocation*. Springer Berlin Heidelberg.

Shi, G., Hong, Y., and Johansson, K.H. (2012). Connectivity and set tracking of multi-agent systems guided by multiple moving leaders. *IEEE Transactions on Automatic Control*, 57(3), 663–676.

Shi, G., Johansson, K.H., and Hong, Y. (2013). Reaching an optimal consensus: Dynamical systems that compute intersections of convex sets. *IEEE Transactions on Automatic Control*, 58(3), 610–622.

Srivastava, K., Nedic, x, A., Stipanovic, and M., D. (2010). Distributed constrained optimization over noisy networks. In *Proceedings of the 49th IEEE Conference on Decision and Control*, 1945–1950.

Srivastava, K. and Nedic, A. (2011). Distributed asynchronous constrained stochastic optimization. *IEEE Journal of Selected Topics in Signal Processing*, 5(4), 772–790.

Su, Y. and Huang, J. (2012). Cooperative output regulation of linear multi-agent systems. *IEEE Transactions on Automatic Control*, 57(4), 1062–1066.

Wang, J. and Elia, N. (2010). Control approach to distributed optimization. In *Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computing*, 557–561. UIUC, Illinois, USA.

Wang, S. (1999). Dynamic simulation of building VAV air-conditioning system and evaluation of EMCS online control strategies. *Building and Environment*, 34(6), 681–705.

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

Zhu, M. and Martinez, S. (2012). On distributed convex optimization under inequality and equality constraints. *IEEE Transactions on Automatic Control*, 57(1), 151–164.