Automatica 69 (2016) 289–297

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Distributed continuous-time approximate projection protocols for shortest distance optimization problems ☆

Youcheng Lou [a,b], Yiguang Hong [b], Shouyang Wang [b]

[a] Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

[b] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 6 April 2014
Received in revised form
2 February 2016
Accepted 4 February 2016
Available online 21 March 2016

Keywords:
Distributed optimization
Convex intersection
Shortest distance optimization
Approximate projection

ABSTRACT

In this paper, we investigate a distributed shortest distance optimization problem for a multi-agent network to cooperatively minimize the sum of the quadratic distances from some convex sets, where each set is only associated with one agent. To deal with this optimization problem with projection uncertainties, we propose a distributed continuous-time dynamical protocol, where each agent can only obtain an approximate projection and communicate with its neighbors over a time-varying communication graph. First, we show that no matter how large the approximate angle is, system states are always bounded for any initial condition, and uniformly bounded with respect to all initial conditions if the inferior limit of the stepsize is greater than zero. Then, in both cases of nonempty and empty intersection of convex sets, we provide stepsize and approximate angle conditions to ensure the optimal convergence, respectively. Moreover, we also give some characterizations about the optimal solutions for the empty intersection case.

1. Introduction

In recent years, distributed optimization of a sum of convex functions has attracted much attention due to its wide applications in resource allocation, source localization and robust estimation (referring to Bertsekas & Tsitsiklis, 1989, Jakovetic, Xavier, & Moura, 2011, Johansson, Rabi, & Johansson, 2009, Lu & Tang, 2012, Lu, Tang, Regier, & Bow, 2011, Nedić & Ozdaglar, 2008, Nedić & Ozdaglar, 2009 and Nedić, Ozdaglar, & Parrilo, 2010). A whole optimization task can be accomplished cooperatively by a group of autonomous agents via simple local information exchange and distributed protocol design even when the communication graph among agents is time-varying.

Although many existing distributed optimization works have been done by discrete-time algorithms, more and more attention has been paid to continuous-time algorithms in recent years (Droge, Kawashima, & Egerstedt, 2014; Gharesifard & Cortés, 2014; Kvaternik & Pavel, 2012; Shi, Johansson, & Hong, 2013; Wang & Elia, 2010, 2011), partially because continuous-time models can be studied by various well-developed continuous-time methods or make the algorithms easily implemented in physical systems. A distributed continuous-time computation model was proposed to solve an optimization problem for a fixed undirected graph in Wang and Elia (2010), with the optimization achieved by controlling the sum of subgradients of convex functions, and later this model was generalized to weight balanced graph case in Gharesifard and Cortés (2013) for differentiable objective functions with globally Lipschitz continuous gradient. Another continuous-time distributed algorithm with a constant stepsize was developed in Kvaternik and Pavel (2012) for optimization problems with positivity constraints in a fixed undirected graph case, where a lower bound of convergence rate and an upper bound on the agents' estimate error were presented. Moreover, the relationship between the existing dual decomposition and consensus-based methods for distributed optimization was revealed in Droge et al. (2014), where both approaches were based on the subgradient method, but one with a proportional control term and the other with an integral control term.

When the optimal solution sets of agents' individual convex objective functions have a nonempty intersection, the distributed optimization problem is equivalent to convex intersection problems

☆ This work was supported by the National Natural Science Foundation of China under Grant 71401163 and 61333001, Beijing Natural Science Foundation under Grant 4152057, Hong Kong Research Grants Council under Grant 419511 and Hong Kong Scholars Program under Grant XJ2015049. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Claudio de Persis under the direction of Editor Christos G. Cassandras.

E-mail addresses: louyoucheng@amss.ac.cn (Y. Lou), yghong@iss.ac.cn (Y. Hong), sywang@amss.ac.cn (S. Wang).

http://dx.doi.org/10.1016/j.automatica.2016.02.019

(CIP) (Deutsch, 1983; Gubin, Polyak, & Raik, 1967; Lou, Shi, Johansson, & Hong, 2013, 2014; Nedić et al., 2010; Shi & Johansson, 2012; Shi et al., 2013). A projected consensus algorithm was proposed in Nedić et al. (2010) for a network to solve CIP, where all agents were shown to converge to a common point in the intersection set for weight-balanced and jointly connected communication graphs. Later, a continuous-time dynamical system was designed and various connectivity conditions were discussed in Shi et al. (2013). In addition, a random sleep algorithm was proposed with providing conditions to converge almost surely to a common point in the intersection set in Lou et al. (2013), with agents to randomly take the neighbor-based average or projection onto their individual sets based on a Bernoulli process. Almost all the existing optimization results were obtained based on the assumption that the exact projection point onto convex sets can be obtained (Deutsch, 1983; Gubin et al., 1967; Lin & Ren, 2012; Meng, Xiao, & Xie, 2013; Nedić et al., 2010; Shi & Johansson, 2012; Shi et al., 2013).

On the other hand, the intersection of convex optimal solution sets may be empty in practice. In this case, how to seek a point with the shortest (quadratic) distance to these sets is also important. For instance, the supply center location problem is concerned with how to seek the location of raw materials supply center so that the average transportation cost from this supply center to the multiple factories is minimal (Francis, McGinnis, & White, 1992; Pardalos & Romeijn, 2002); the source localization in a sensor network is related to estimate the location of the source emitting a signal based on the received signals of multiple sensors in a noisy environment (Meng et al., 2013; Zhang, Lou, Hong, & Xie, 2015). In fact, the problem for both empty and nonempty intersection cases is referred to as the shortest distance optimization problem (SDOP). Clearly, CIP is a special case of SDOP, and the average consensus problem is also a special case of SDOP since the optimal solution of the minimum of the sum of quadratic functions from some points is exactly the average of these points. Some distributed algorithms were proposed to discuss SDOP. For example, Meng et al. (2013) formulated a source localization problem as the SDOP in a plane and proposed a discrete-time distributed algorithm, with the adjacency matrices of communication graphs required to be doubly stochastic. Moreover, Lin and Ren (2012) proposed two distributed continuous-time algorithms to solve SDOP in the empty intersection case for connected graphs: the first one was designed for optimal consensus based on sign functions, and the second one was modified to avoid chattering but only to achieve the optimal consensus approximately.

The objective of this paper is to design a continuous-time distributed protocol to solve SDOP based on approximate projection. Note that the exact projection is usually hard to obtain in practice. Therefore, approximate projection issues were discussed in some situations, and for example, Lou et al. (2014) proposed a discrete-time approximate projected consensus algorithm to solve CIP. The motivation of the current research aims at cooperatively solving SDOP with projection uncertainties and continuous-time dynamics. For example, in a practical robotic network to solve the SDOP, a robot may not always obtain the exact projection point of its own convex set, but only spot some point on the set surface near the exact projection point. The contribution of this paper can be summarized as follows.

- We propose a new concept of approximate projection when the exact projection is hard to obtain. In fact, we consider an approximate projection related to set boundary surfaces, different from that defined in a "triangle" in Lou et al. (2014). To overcome the difficulties resulting from this new approximate projection, we employ a geometric method to convert the original problem into a heterogeneous stepsize problem.
- Given any approximate angle, we show that, with the proposed continuous-time algorithm, the agent states are always bounded for any initial condition, and uniformly bounded with respect to all initial conditions if the stepsize is not too small. The result with respect to the continuous-time algorithm is different from some results based on some discrete-time ones. In fact, $\pi/4$ was shown to be the critical approximate angle for the boundedness of the discrete-time algorithm with the approximate projection defined in a triangle in Lou et al. (2014).
- We study SDOP in both nonempty and empty intersection cases, and propose a unified protocol based on the approximate projection. In fact, the proposed convergence conditions and proofs in the two cases are quite different. Note that our result is different from that in Lin and Ren (2012) because we handle approximate projections without assuming that the communication graph is always connected, and ours tackles both nonempty and empty intersection cases, while Lou et al. (2014) only does the nonempty intersection case.

Notations: $\mathbf{1}$ denotes the vector with all ones; y^T denotes the transpose of a vector $y \in \mathbb{R}^m$; $|y|$ denotes the Euclidean norm of y; $[v, z]$ denotes the line segment connecting the two points v, z; line(v, z) denotes the line passing the two points v, z; for a set $K \subseteq \mathbb{R}^m$, int(K) and $bd(K) = K\backslash\text{int}(K)$ denote the sets of interior points and boundary points of K, respectively; $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product in $\mathbb{R}^m$; the angle between nonzero vectors y and z is denoted as $\angle(y, z) \in [0, \pi]$, where $\cos\angle(y, z) = \langle y, z\rangle/(|y|\,|z|)$; span$\{v_1, \ldots, v_p\}$ (aff$\{v_1, \ldots, v_p\}$) denotes the subspace (affine hull) generated by vectors $v_1, \ldots, v_p$.

2. Preliminaries

2.1. Graph theory

A multi-agent network can be described by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, \ldots, n\}$ is the node (or agent) set and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ the arc set with the arc $(j, i) \in \mathcal{E}$ describing that node i can receive the information of node j. Here $(i, i) \notin \mathcal{E}$ for all i. Let $\mathcal{N}_i = \{j \in \mathcal{V} | (j, i) \in \mathcal{E}\}$ be the set of neighbors of node i. A path from node i to node j in $\mathcal{G}$ is a sequence $(i, i_1), (i_1, i_2), \ldots, (i_p, j)$ of arcs in $\mathcal{E}$. Graph $\mathcal{G}$ is said to be strongly connected if there exists a path from i to j for each pair of nodes $i, j \in \mathcal{V}$. Graph $\mathcal{G}$ is undirected when $(j, i) \in \mathcal{E}$ if and only if $(i, j) \in \mathcal{E}$ (Godsil & Royle, 2001).

The communication over the network under consideration is switching and characterized by a directed graph process $\mathcal{G}_{\sigma(t)} = (\mathcal{V}, \mathcal{E}_{\sigma(t)})$, $t \geq 0$, with $\mathcal{E}_{\sigma(t)}$ the arc set of the graph at time t. Here $\sigma : [0, \infty) \to \mathcal{Q}$ is a piecewise constant function to describe the time-varying graph process, where $\mathcal{Q}$ is the index set of all possible graphs on $\mathcal{V}$. Let $\Delta := \{t_k, k \geq 0\}$ with $t_0 = 0$ denote the set of all switching moments of switching graph $\mathcal{G}_\sigma$. As usual, we assume there is a dwell time $\tau > 0$ between two consecutive graph switching moments, i.e., $t_{k+1} - t_k \geq \tau$ for all k. The switching graph $\mathcal{G}_\sigma$ is uniformly jointly strongly connected (UJSC) if there exists $T > 0$ such that the union graph $(\mathcal{V}, \cup_{t\leq s<t+T}\, \mathcal{E}(s))$ is strongly connected for $t \geq 0$.

2.2. Convex analysis

A set $K \subseteq \mathbb{R}^m$ is convex if $\lambda z_1 + (1 - \lambda)z_2 \in K$ for any $z_1, z_2 \in K$ and $0 < \lambda < 1$. For a closed convex set K in $\mathbb{R}^m$, we can associate with any $z \in \mathbb{R}^m$ a unique element $P_K(z) \in K$ satisfying $|z - P_K(z)| = \inf_{y\in K} |z - y| =: |z|_K$, where P_K is called the projection operator onto K. We have the following properties for the projection operator P_K (Rockafellar, 1972).

Lemma 2.1. *Let K be a closed convex set in $\mathbb{R}^m$. Then*

(i) $\langle y - P_K(y), z - P_K(y)\rangle \le 0 \quad \forall y,\ \forall z \in K$;
(ii) $|P_K(y) - z| \le |y - z| \quad \forall y,\ \forall z \in K$;
(iii) $\langle y - P_K(y), z - y\rangle \le |y|_K(|z|_K - |y|_K) \quad \forall y,\ z$;
(iv) $|P_K(y) - P_K(z)| \le |y - z| \quad \forall y,\ z$;
(v) $\langle y - z, P_K(y) - P_K(z)\rangle \ge |P_K(y) - P_K(z)|^2 \quad \forall y,\ z$;
(vi) $|P_K(y) - P_K(z)| = |y - z|$ *if and only if* $y - P_K(y) = z - P_K(z)$.

Proof. (i) is an equivalent definition of convex projection; (ii) comes from Lemma 1(b) in Nedić et al. (2010). We now show (iii). First of all, $\langle y - P_K(y), P_K(z) - P_K(y)\rangle \le 0$ by (i). Clearly, $\langle y - P_K(y), z - P_K(z)\rangle \le |y|_K|z|_K$. Then

$$\langle y - P_K(y), z - y\rangle = \langle y - P_K(y), z - P_K(z) + P_K(z) - P_K(y) + P_K(y) - y\rangle \le |y|_K|z|_K - |y|_K^2,$$

which implies (iii). (iv) is the standard non-expansive property, and (v) follows from

$$\begin{aligned}\langle y - z, P_K(y) - P_K(z)\rangle &= \langle y - P_K(y), P_K(y) - P_K(z)\rangle \\ &\quad + |P_K(y) - P_K(z)|^2 + \langle P_K(z) - z, P_K(y) - P_K(z)\rangle \\ &\ge |P_K(y) - P_K(z)|^2\end{aligned}$$

because $\langle y - P_K(y), P_K(y) - P_K(z)\rangle \ge 0$ and $\langle P_K(z) - z, P_K(y) - P_K(z)\rangle \ge 0$ by (i).

For (vi), the sufficiency is obvious. The necessity can be obtained from

$$\begin{aligned}|y - P_K(y) - (z - P_K(z))|^2 &= |y - z|^2 \\ &\quad + |P_K(z) - P_K(y)|^2 + 2\langle y - z, P_K(z) - P_K(y)\rangle \\ &= 2|y - z|^2 + 2\langle y - z, P_K(z) - P_K(y)\rangle \\ &\le 2|y - z|^2 - 2|P_K(y) - P_K(z)|^2 = 0,\end{aligned}$$

where the inequality follows from (v). □

The following lemma can be found on page 24 in Aubin and Cellina (1984).

Lemma 2.2. *For a closed convex set K in $\mathbb{R}^m$, $|x|_K^2$ is continuously differentiable and $\nabla |x|_K^2 = 2(x - P_K(x))$.*

A function $\varphi(\cdot) : \mathbb{R}^m \to \mathbb{R}$ is said to be convex if $\varphi(\lambda z_1 + (1 - \lambda)z_2) \le \lambda\varphi(z_1) + (1 - \lambda)\varphi(z_2)$ for any $z_1, z_2 \in \mathbb{R}^m$ and $0 < \lambda < 1$. For a continuously differentiable convex function φ, it holds that $\varphi(y) \ge \varphi(x) + \langle y - x, \nabla\varphi(x)\rangle,\ \forall x, y$.

The upper Dini derivative of function $g : (a, b) \to \mathbb{R}$ at $t \in (a, b)$ is defined as $D^+g(t) = \limsup_{s\to 0^+} \frac{g(t+s)-g(t)}{s}$. The following result was shown in Danskin (1966).

Lemma 2.3. *Let $g_i(t, x) : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, $i = 1, \dots, n$ be continuously differentiable. Then $D^+g(t, x(t)) = \max_{i\in\mathcal{I}(t)} \dot g_i(t, x(t))$, where $g(t, x) = \max_{1\le i\le n} g_i(t, x)$, $\mathcal{I}(t) = \{i | g_i(t, x(t)) = g(t, x(t)),\ 1 \le i \le n\}$.*

2.3. Consensus

Consider the following consensus model with disturbance w_i,

$$\dot z_i(t) = \sum_{j\in\mathcal{N}_i(t)} (z_j(t) - z_i(t)) + w_i(t), \quad i = 1, \dots, n, \tag{1}$$

where the disturbance $w_i(t) : [0, \infty) \to \mathbb{R}$ is continuous. System (1) has a continuous solution, which satisfies (1) for almost all t except at the switching moments of switching graph $\mathcal{G}_\sigma$. The next lemma can be obtained from the proof of Proposition 4.10 in Shi and Johansson (2013).

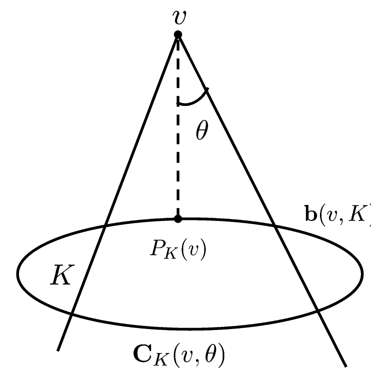

Fig. 1. The approximate projection of point v onto closed convex set K.

Lemma 2.4. *Suppose the switching graph $\mathcal{G}_\sigma$ of system (1) is UJSC and $\lim_{t\to\infty} w_i(t) = 0$ for all i. Then consensus is achieved for system (1), i.e., for any initial condition, $\lim_{t\to\infty} |z_i(t) - z_j(t)| = 0$ for all $1 \le i, j \le n$.*

3. Problem formulation and algorithm

In this section, we introduce the distributed SDOP and propose a distributed continuous-time approximate projected algorithm.

Consider a network of n agents (or nodes) and bounded closed convex sets $X_i \subseteq \mathbb{R}^m$ for $i = 1, \dots, n$, with X_i only associated with (or known by) agent i. The goal of the network is to cooperatively find a point x^* with the shortest quadratic distance from the n closed convex sets:

$$x^* \in \arg\min f(x), \quad f(x) = \sum_{i=1}^n |x|_{X_i}^2. \tag{2}$$

Projection-based methods have been widely adopted in the literature to solve CIP and constrained optimization problems, and almost all methods require that the exact projection can be obtained (Johansson et al., 2009; Lin & Ren, 2012; Meng et al., 2013; Nedić et al., 2010; Shi & Johansson, 2012; Shi et al., 2013). Since the exact projection may be difficult to obtain in practice, each agent may only obtain an approximate projection point located on the convex set surface and near the exact projection point. To be strict, we give the following definition.

Definition 1. Let $0 \le \theta < \pi/2$ and K be a closed convex set in $\mathbb{R}^m$. Define sets

$$\mathbf{C}_K(v, \theta) = v + \{z | \langle z, P_K(v) - v\rangle \ge |z|\,|v|_K \cos\theta\},$$
$$\mathbf{b}(v, K) = \{z | z \in bd(K), [v, z] \cap bd(K) = \{z\}\}.$$

The approximate projection $\mathbf{P}_K^a(v, \theta)$ of point v onto K is defined as the following set:

$$\mathbf{P}_K^a(v, \theta) = \begin{cases}\mathbf{C}_K(v, \theta) \cap \mathbf{b}(v, K), & \text{if } v \notin K;\\ \{v\}, & \text{otherwise.}\end{cases}$$

As shown in Fig. 1, the cone $\mathbf{C}_K(v, \theta) - v$ consists of all vectors having angle with the direction $P_K(v) - v$ not greater than θ, and $\mathbf{b}(v, K)$ is the region on the boundary of K that the agent can "see" starting from point v. Clearly, the exact projection $P_K(v) \in \mathbf{P}_K^a(v, \theta)$ for any $v \in \mathbb{R}^m$ and $0 \le \theta < \pi/2$ and $\mathbf{P}_K^a(v, 0) = \{P_K(v)\}$.

Remark 3.1. Approximate projections are more "practical" than the exact projection. For example, in reality, when a robot approaches its (convex) target set, it may not get the exact projection. Instead, it may select another point on the set surface as the exact one by mistake or to avoid expensive measurement or

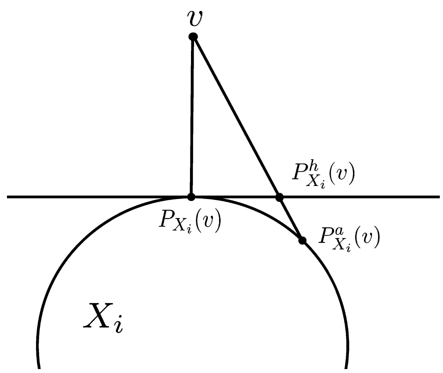

Fig. 2. An illustration for $P^h_{X_i}(v)$.

tedious computation. Then the selected projection point becomes an approximate one. In other words, this concept captures the situation when agents can only obtain some point on the set surface, which may not be but close to the exact projection point. Note that this concept is different from that given in Lou et al. (2014), where the approximate projection point is located in a "triangle" region specified by v, the hyperplane of K on $P_K(v)$ and the approximate angle θ.

We next give some basic assumptions for our following analysis.

A1 (Connectivity) The switching graph $\mathcal{G}_\sigma$ is UJSC.

A2 (Convex sets) (i) The boundary surfaces of convex sets X_i, $i = 1, \ldots, n$ are regular (or smooth);

(ii) The convex set X_i contains nonempty interior points for $i = 1, \ldots, n$.

The definition of regularity or smoothness of a manifold can be easily found (referring to Definition 1 on page 52 in do Carmo, 1976 for more details). Note that the Gaussian curvature of regular (or smooth) surfaces of closed bounded sets are bounded. In fact, **A2** is quite mild. The boundaries of many well-known sets, such as the surfaces of spheres, ellipsoids, are regular; and moreover, the assumption that set $X_i \subseteq \mathbb{R}^m$ contains nonempty interior points is equivalent to $\dim(X_i) = m$ ($\dim(X_i)$ denotes the dimension of the affine hull of set X_i), which was also widely used in the literature.

Let $P^a_{X_i}(\cdot) : \mathbb{R}^m \to \mathbb{R}^m$ be a continuous map with $P^a_{X_i}(v) \in \mathbf{P}^a_{X_i}(v, \theta_i(v))$ for any v, where

$$\theta_i(v) = \angle(P^a_{X_i}(v) - v, P_{X_i}(v) - v),$$

$0 \le \theta_i(v) < \pi/2$. Let $\theta_i(v) = 0$ for simplicity when $v \in X_i$. In this paper, $\theta_i(v)$ is referred to as the approximate angle of v onto X_i. The following assumption was used in Lou et al. (2014).

A3 (Approximate angle) There exists $0 < \theta^* < \pi/2$ such that $0 \le \theta_i(v) \le \theta^*$ for all i, v.

Here we propose a distributed continuous-time approximate projected algorithm:

$$\dot{x}_i(t) = \sum_{j \in \mathcal{N}_i(t)} (x_j(t) - x_i(t)) + \alpha_t (P^a_{X_i}(x_i(t)) - x_i(t)), \quad i = 1, \ldots, n, \tag{3}$$

where $x_i \in \mathbb{R}^m$ is the state estimate of agent i for optimal solutions, $\mathcal{N}_i(t)$ is the neighbor set of node i at time t, $\{\alpha_t\}$ is the stepsize ($0 \le \alpha_t \le \alpha^*, \alpha^* > 0$) and is uniformly continuous over t. The continuity of stepsize α_t and maps $P^a_{X_i}(\cdot)$ guarantees that (3) has a solution that is continuous over $[0, \infty)$ and continuously differentiable except at the switching moments of switching graph $\mathcal{G}_\sigma$.

The convergence analysis of (3) is not easy because the gradient term is corrupted with state-dependent approximation and there is no explicit expression to describe the relationship between the approximate projection point and the exact one. To handle the problem, we make some transformation. In the case of $v \notin X_i$, we define by $P^h_{X_i}(v)$ the intersection point of the hyperplane of X_i at $P_{X_i}(v)$ (the tangent plane of $bd(X_i)$ at $P_{X_i}(v)$) with $P_{X_i}(v) - v$ as the normal direction and the line segment $[v, P^a_{X_i}(v)]$, as shown in Fig. 2. Clearly, $P^h_{X_i}(v) = P_{X_i}(v)$ when $P^a_{X_i}(v) = P_{X_i}(v)$. In the case of $v \in X_i$, we define $P^h_{X_i}(v) = P^a_{X_i}(v) = v$. Then we can find that $P^h_{X_i}(v) = v$ if and only if $v \in X_i$. We write

$$P^a_{X_i}(v) - v = \gamma_{X_i}(v)(P^h_{X_i}(v) - v),$$

where $\gamma_{X_i}(v) = \frac{|P^a_{X_i}(v) - v|}{|P^h_{X_i}(v) - v|} \ge 1$ if $P^h_{X_i}(v) \ne v$, and $\gamma_{X_i}(v) = 1$ otherwise.

Rewrite $\alpha_t(P^a_{X_i}(x_i(t)) - x_i(t)) = \alpha_{i,t}(P^h_{X_i}(x_i(t)) - x_i(t))$, with the virtual stepsize of agent i defined as

$$\alpha_{i,t} = \begin{cases} \gamma_{X_i}(x_i(t))\alpha_t \\ \quad = \dfrac{|P^a_{X_i}(x_i(t)) - x_i(t)|}{|P^h_{X_i}(x_i(t)) - x_i(t)|}\alpha_t, & \text{if } P^h_{X_i}(x_i(t)) \ne x_i(t); \\ \alpha_t, & \text{otherwise.} \end{cases}$$

Clearly, $\alpha_{i,t} \ge \alpha_t$. Although the designed stepsize α_t is the same for all agents, agent i has its own virtual stepsize $\alpha_{i,t}$ based on its own approximate projection. We express (3) in another form with heterogeneous virtual stepsizes:

$$\dot{x}_i(t) = \sum_{j \in \mathcal{N}_i(t)} (x_j(t) - x_i(t)) + \alpha_{i,t}(P^h_{X_i}(x_i(t)) - x_i(t)), \quad i = 1, \ldots, n. \tag{4}$$

Then we describe our problem as follows.

Definition 2. The shortest distance optimization problem (SDOP) is solved by (3) or (4) if, for any initial condition $x_i(0) \in \mathbb{R}^m$, $i = 1, \ldots, n$, there exists $x^* \in \arg\min \sum_{i=1}^n |x|^2_{X_i}$ such that $\lim_{t\to\infty} x_i(t) = x^*$, $i = 1, \ldots, n$.

In the following three sections, we first establish some basic results, and then present the convergence results in nonempty intersection and empty intersection cases.

4. Discussions on boundedness and stepsizes

In this section, we show the state boundedness and establish an "equivalent" relationship between the designed stepsize α_t and the virtual stepsize $\alpha_{i,t}$.

4.1. Boundedness of system states

Denote $\theta_{i,t} = \theta_i(x_i(t))$ for simplicity. Note that $\theta_{i,t}$ is also equal to $\angle(P^h_{X_i}(x_i(t)) - x_i(t), P_{X_i}(x_i(t)) - x_i(t))$. Here we study the boundedness of $x_i(t)$, $i \in \mathcal{V}$, $t \ge 0$ of (4) with the approximate angle $\theta_{i,t}$.

Let $X_c = \text{co}\{X_i, i = 1, \ldots, n\}$ be the convex hull of the sets X_i, $i = 1, \ldots, n$, $\xi := \sup_{z_1, z_2 \in X_c} |z_1 - z_2|$, which is finite due to the boundedness of X_is.

Theorem 4.1. (i) *For any initial condition $x_i(0)$, $i \in \mathcal{V}$, the system states $x_i(t)$, $i \in \mathcal{V}$, $t \ge 0$ are bounded.*

(ii) *Suppose $\liminf_{t\to\infty} \alpha_t > 0$. Then, for any $0 < \theta < \pi/2$ and any initial condition $x_i(0)$, $i \in \mathcal{V}$,*

$$\limsup_{t\to\infty} |x_i(t)|_{X_c} \le \max\left\{\frac{\xi}{\sin\theta}, \xi\left(\tan\theta + \sqrt{(\tan\theta)^2 + 2\tan\theta}\right)\right\}.$$

*Furthermore, if **A3** holds, then for any initial condition $x_i(0)$, $i \in \mathcal{V}$, $\limsup_{t\to\infty} |x_i(t)|_{X_c} \le \xi\left(\tan\theta^* + \sqrt{(\tan\theta^*)^2 + 2\tan\theta^*}\right)$.*

Proof. Let $t \notin \Delta$. Denote $\hbar_i(t) = \frac{1}{2}|x_i(t)|^2_{X_c}$, $\hbar(t) = \max_{1\le i\le n} \hbar_i(t)$. By Lemmas 2.2 and 2.3, we have

$$D^+\hbar(t) = \max_{i\in\mathcal{I}(t)} \langle x_i(t) - P_{X_c}(x_i(t)), \dot{x}_i(t)\rangle$$
$$= \max_{i\in\mathcal{I}(t)} \Big\langle x_i(t) - P_{X_c}(x_i(t)), \sum_{j\in\mathcal{N}_i(t)} (x_j(t) - x_i(t)) + \alpha_{i,t}(P^h_{X_i}(x_i(t)) - x_i(t))\Big\rangle \quad (5)$$

where $\mathcal{I}(t) = \{i | i \in \mathcal{V}, \hbar_i(t) = \hbar(t)\}$. Take $i \in \mathcal{I}(t)$. Lemma 2.1(iii) implies that, for any j,

$$\langle x_i - P_{X_c}(x_i), x_j - x_i\rangle \le |x_i|_{X_c}(|x_j|_{X_c} - |x_i|_{X_c}) \le 0. \quad (6)$$

According to Lemma 2.1(i), $\langle x_i - P_{X_c}(x_i), P_{X_i}(x_i) - P_{X_c}(x_i)\rangle \le 0$ due to $X_i \subseteq X_c$. Therefore,

$$\langle x_i - P_{X_c}(x_i), P_{X_i}(x_i) - x_i\rangle \le -|x_i|^2_{X_c}. \quad (7)$$

Moreover, recalling the definitions of $P^h_{X_i}(x_i(t))$ and $\theta_{i,t}$, we have $\langle x_i(t) - P_{X_i}(x_i(t)), P^h_{X_i}(x_i(t)) - P_{X_i}(x_i(t))\rangle = 0$ and $|P^h_{X_i}(x_i(t)) - P_{X_i}(x_i(t))| = \tan\theta_{i,t}|x_i(t)|_{X_i}$. Then

$$\langle x_i(t) - P_{X_c}(x_i(t)), P^h_{X_i}(x_i(t)) - P_{X_i}(x_i(t))\rangle$$
$$\le |P_{X_i}(x_i(t)) - P_{X_c}(x_i(t))||P^h_{X_i}(x_i(t)) - P_{X_i}(x_i(t))|$$
$$\le \xi \tan\theta_{i,t}(|x_i(t)|_{X_c} + \xi). \quad (8)$$

Thus, based on (7) and (8), we have

$$\langle x_i(t) - P_{X_c}(x_i(t)), P^h_{X_i}(x_i(t)) - x_i(t)\rangle$$
$$\le -|x_i(t)|^2_{X_c} + \xi \tan\theta_{i,t}(|x_i(t)|_{X_c} + \xi). \quad (9)$$

With (5), (6), (9) and $i \in \mathcal{I}(t)$, we obtain $D^+\hbar(t) \le \alpha_{i,t}(-2\hbar(t) + \xi\tan\theta_{i,t}(\sqrt{2\hbar(t)} + \xi))$. We complete the proof by the following analysis.

(i) It is easy to see that for any $0 < \hat{\theta} < \pi/2$, $\theta_{i,t} \le \hat{\theta}$ when $|x_i(t)|_{X_c} \ge \frac{\xi}{\sin\hat{\theta}}$. Then $D^+\hbar(t) \le 0$ when $t \notin \Delta$ and $\hbar(t) \ge \max\big\{\frac{\xi^2}{2(\sin\hat{\theta})^2}, \frac{\xi^2(\tan\hat{\theta})^2}{4} + \frac{\xi^2\tan\hat{\theta}}{2}\big(1 + \frac{\sqrt{(\tan\hat{\theta})^2+4\tan\hat{\theta}}}{2}\big)\big\}$. Hence, system states are bounded for any initial condition.

(ii) Let $\alpha_* = (\liminf_{t\to\infty}\alpha_t)/2 > 0$ and $\hat{t}$ be the moment such that when $t \ge \hat{t}$, $\alpha_{i,t} \ge \alpha_t \ge \alpha_*$. Then $D^+\hbar(t) \le -\alpha_*\hbar(t)$ once $t \ge \hat{t}, t \notin \Delta$ and $\hbar(t) \ge \max\big\{\frac{\xi^2}{2(\sin\theta)^2}, \xi^2(\tan\theta)^2 + \xi^2\tan\theta\big(1 + \sqrt{(\tan\theta)^2 + 2\tan\theta}\big)\big\}$. Therefore, $\hbar(t)$ is not greater than the number in the preceding inequality when $t \ge \hat{t}$ and $t \notin \Delta$. The second conclusion can be obtained directly based on the above arguments.

Thus, the conclusion follows from the continuity of $x_i(t)$. □

A discrete-time algorithm was proposed in Lou et al. (2014) to solve CIP with approximate projection, where it was found that in the case of $\alpha_{i,k} \equiv 1$ and $\theta_{i,k} \equiv \theta$, the states are uniformly bounded with respect to all initial conditions when $\theta < \pi/4$ and unbounded for most all initial conditions when $\theta > \pi/4$. Different from this critical approximate angle result, Theorem 4.1 shows that the continuous-time system states are always bounded for any initial condition no matter how large θ is, and moreover, the states are uniformly bounded for all initial conditions with fixing $\alpha_{i,t} \equiv 1$.

4.2. Equivalence between stepsizes

To obtain the convergence conditions, we establish a relationship between the designed stepsize α_t and virtual stepsizes $\alpha_{i,t}$. Let $S = X_c + \mathbf{B}(0, r_0)$, where $\mathbf{B}(0, r_0)$ denotes the ball with center zero and radius $r_0 > 0$. Denote $\mu_i(v) = \angle(P_{X_i}(v) - P^a_{X_i}(v), v - P^a_{X_i}(v))$.

Lemma 4.1. *Suppose* **A2** *and* **A3** *hold. Then*

$$\inf_{v\in S\setminus X_i, P^a_{X_i}(v)\neq P_{X_i}(v)} \mu_i(v) > 0.$$

The proof is in the Appendix. Because the states of (4) are bounded by Theorem 4.1, we take sufficiently large r_0 such that S contains all the system states. We next show that **A2** and **A3** imply the equivalence between $\{\alpha_t\}$ and $\{\alpha_{i,t}\}$.

Clearly, $\alpha_t = \alpha_{i,t}$ when $x_i(t) \in X_i$ or $P^a_{X_i}(x_i(t)) = P_{X_i}(x_i(t))$. Then we only need to focus on the case when $x_i(t) \notin X_i$ and $P^a_{X_i}(x_i(t)) \neq P_{X_i}(x_i(t))$. Clearly, we can find that for any $v \in S\setminus X$ and $P^a_X(v) \neq P_X(v)$,

$$\gamma_X(v) = 1 + \frac{|P^a_X(v) - P^h_X(v)|}{|P^h_X(v) - P_X(v)|}\sin\theta(v) \le 1 + \frac{\sin\theta(v)}{\sin\mu(v)}.$$

Therefore, $\gamma_{X_i}(x_i(t)) \le 1 + \frac{\sin\theta_{i,t}}{\sin\mu_{i,t}}$ with $\mu_{i,t} := \mu_i(x_i(t))$. Then by Lemma 4.1 we have the following result.

Theorem 4.2. *Under* **A2** *and* **A3**,

$$\alpha_t \le \alpha_{i,t} \le C_{i,t}\alpha_t \le C_i\alpha_t, \quad \forall t, \quad (10)$$

where $C_{i,t} = 1 + \frac{\sin\theta_{i,t}}{\sin\mu_i}$, $C_i = 1 + \frac{1}{\sin\mu_i}$,

$$\mu_i := \inf_{t\ge 0, x_i(t)\notin X_i, P^a_{X_i}(x_i(t))\neq P_{X_i}(x_i(t))} \mu_{i,t}$$
$$\ge \inf_{v\in S\setminus X_i, P^a_{X_i}(v)\neq P_{X_i}(v)} \mu_i(v) > 0.$$

In fact, (10) somehow characterizes the bounded bending property of smooth surfaces, which helps convert the convergence conditions on $\alpha_{i,t}$ into the conditions on α_t.

Remark 4.1. As Theorem 4.2 shows, **A2** and **A3** guarantee the equivalence between the designed stepsize and the virtual stepsize. In fact, with (10), we found that under **A1**, the optimal convergence established in the next two sections holds for general convex sets (not necessary to satisfy **A2** and **A3**).

5. Nonempty intersection case

In this section, we show the convergence result in the nonempty intersection case, $\cap_{i=1}^n X_i \neq \emptyset$. Clearly, $X_0 := \cap_{i=1}^n X_i$ is the optimal solution set of $\min \sum_{i=1}^n |x|^2_{X_i}$.

Theorem 5.1. *Suppose* **A1**–**A3** *hold and* $\cap_{i=1}^n X_i \neq \emptyset$. *Then SDOP is solved by system (4) if* $\int_0^\infty \alpha_t dt = \infty$, $\int_0^\infty \alpha_t \tan\theta_t^+ dt < \infty$, *where* $\theta_t^+ = \max_{1\le i\le n}\theta_{i,t}$.

Proof. Denote $\alpha_t^+ = \max_{1\le i\le n}\alpha_{i,t}$. From (10), we find that $\int_0^\infty \alpha_t dt = \infty$, $\int_0^\infty \alpha_t \tan\theta_t^+ dt < \infty$ are equivalent to $\int_0^\infty \alpha_t^+ dt = \infty$, $\int_0^\infty \alpha_t^+ \tan\theta_t^+ dt < \infty$, respectively.

Denote the distance functions $h(t) = \max_{1\le i\le n} h_i(t)$, $h_i(t) = \frac{1}{2}|x_i(t)|^2_{X_0}$, $h_i^+ = \limsup_{t\to\infty} h_i(t)$, and $h_i^- = \liminf_{t\to\infty} h_i(t)$. Under the hypotheses, by Lemmas 2.1–2.3 and the similar arguments in Lemma 4.3 in Shi et al. (2013), we can successively show the following four conclusions: (i) $D^+h(t) \le 2\alpha_t^+\tan\theta_t^+ h(t)$ for any $t \notin \Delta$; (ii) the limit $\lim_{t\to\infty} h(t) =: h^*$ exists; (iii) if there is some node i_0 such that $h^-_{i_0} < h^*$, then $h^* = 0$; (iv) for any $z \in X_0$, $\lim_{t\to\infty}\max_{1\le i\le n}|x_i(t) - z|^2$ exists. It follows from the fourth conclusion that all agents will converge to a common point in X_0 if the consensus is achieved and $h^* = 0$. Thus, it suffices to show that consensus is achieved and $h^* = 0$.

Because

$$|P^h_{X_i}(x_i(t)) - x_i(t)| = \frac{|x_i(t)|_{X_i}}{\cos\theta_{i,t}} \le \frac{\sqrt{2h(t)}}{\cos\theta^*}, \tag{11}$$

it follows that, if $h^* = 0$, the second term on the right-hand side of (4) tends to zero as $t \to \infty$ and then the consensus is achieved for system (4) by Lemma 2.4. Therefore, it suffices to show $h^* = 0$ in what follows.

In fact, if there is some node i_0 with $h^-_{i_0} < h^*$, then $h^* = 0$ from the previous statements. Therefore, we need to prove $h^* = 0$ from $h^+_i = h^-_i = h^*$, $\forall i$ by contradiction. Clearly, for any $\varepsilon > 0$, there is $\bar t > 0$ such that when $t \ge \bar t$, $|x_i(t)|_{X_0} \le \sqrt{2h^*} + \varepsilon =: \phi$. We complete the proof by the following two steps.

Step (i). Suppose $h^+_i = h^-_i = h^* > 0$, $\forall i$. We claim that consensus can be achieved for system (4).

We first show that $\lim_{t\to\infty} \alpha_{i,t}|x_i(t)|^2_{X_i} = 0$ by contradiction. Hence suppose there exist i_0 and an increasing subsequence $\{s_k\}_{k\ge0}$ with $\lim_{k\to\infty} s_k = \infty$ such that $\alpha_{i_0,s_k}|x_{i_0}(s_k)|^2_{X_{i_0}} \ge c$ for some $c > 0$. Without loss of generality, we assume s_0 is sufficiently large such that $s_0 \ge \bar t$ and $\int_{s_0}^\infty \alpha^+_t \tan\theta^+_t dt \le \varepsilon/\sqrt{2h^*}$.

From Lemma 2.2, the boundedness of system states and (11), we know that $|x_{i_0}(t)|^2_{X_{i_0}}$ is uniformly continuous on $[0,\infty)$. This along with the uniform continuity of α_t again implies that $\alpha_t|x_{i_0}(t)|^2_{X_{i_0}}$ is also uniformly continuous on $[0,\infty)$. Therefore, there is $\delta > 0$ such that $\alpha_{i_0,t}|x_{i_0}(t)|^2_{X_{i_0}} \ge c/2$ when $s_k \le t \le s_k + \delta$. Without loss of generality, we assume $[s_k, s_k+\delta] \cap \varDelta = \emptyset$ for all k. We have

$$\frac{dh_{i_0}(t)}{dt} \le \sum_{j\in\mathcal{N}_{i_0}(t)} |x_{i_0}(t)|_{X_0}(|x_j(t)|_{X_0} - |x_{i_0}(t)|_{X_0}) - \alpha_{i_0,t}|x_{i_0}(t)|^2_{X_{i_0}} + \alpha_{i_0,t}\tan\theta_{i_0,t}|x_{i_0}(t)|^2_{X_0}$$

and then for $s_k \le t \le s_k + \delta$,

$$\begin{aligned}D^+|x_{i_0}(t)|_{X_0} &\le \sum_{j\in\mathcal{N}_{i_0}(t)} (|x_j(t)|_{X_0} - |x_{i_0}(t)|_{X_0}) - \frac{\alpha_{i_0,t}|x_{i_0}(t)|^2_{X_{i_0}}}{\phi} + \alpha_{i_0,t}\tan\theta^+_t\phi \\ &\le (n-1)\big(\phi - |x_{i_0}(t)|_{X_0}\big) - \frac{c}{2\phi} + \alpha^+_t\tan\theta^+_t\phi, \end{aligned} \tag{12}$$

which leads to

$$|x_{i_0}(s_k+\delta)|_{X_0} \le \zeta(\sqrt{2h^*}+\varepsilon) + (1-\zeta) \times \left(\sqrt{2h^*}+\varepsilon - \frac{c}{2(n-1)\phi}\right) + \varepsilon\frac{\sqrt{2h^*}+\varepsilon}{\sqrt{2h^*}}, \tag{13}$$

where $0 < \zeta = e^{-(n-1)\delta} < 1$. We can find that the right-hand side of (13) is less than $\sqrt{2h^*} - \frac{c(1-\zeta)}{4(n-1)\sqrt{2h^*}}$ when ε is sufficiently small, which contradicts $\lim_{t\to\infty} h_{i_0}(t) = h^*$. Thus, $\lim_{t\to\infty}\alpha_{i,t}|x_i(t)|^2_{X_i} = 0$, $\forall i$. From Theorem 4.2 we have $0 \le \alpha_{i,t} \le C_i\alpha^*$, and hence $\lim_{t\to\infty}\alpha_{i,t}|x_i(t)|_{X_i} = 0$, $\forall i$. According to the equality in (11) and Lemma 2.4, consensus is achieved for system (4).

Step (ii). Suppose $h^+_i = h^-_i = h^* > 0$, $\forall i$. We will show that $\liminf_{t\to\infty}\sum_{i=1}^n |x_i(t)|^2_{X_i} = 0$ by contradiction.

Hence suppose there is $b > 0$ such that $\sum_{i=1}^n |x_i(t)|^2_{X_i} \ge b$ for all sufficiently large t. Let $|x(t)|_{X_0} = (|x_1(t)|_{X_0}, \dots, |x_n(t)|_{X_0})^T$, $y(t) = (|x_1(t)|^2_{X_1}, \dots, |x_n(t)|^2_{X_n})^T$, $D(t) = \mathrm{diag}\{\alpha_{1,t}, \dots, \alpha_{n,t}\}$ (a diagonal matrix with diagonal entries $\alpha_{i,t}$). Then by (12) we have

$$D^+|x(t)|_{X_0} \le -\mathcal{L}_{\sigma(t)}|x(t)|_{X_0} - \frac{1}{\phi}D(t)y(t) + \phi\alpha^+_t\tan\theta^+_t\mathbf{1},$$

where $\mathcal{L}_{\sigma(t)}$ is the Laplacian of graph $\mathcal{G}_{\sigma(t)}$ with $(\mathcal{L}_{\sigma(t)})_{ij} = -1$ if $j \in \mathcal{N}_i(t)$, $(\mathcal{L}_{\sigma(t)})_{ij} = 0$ if $j \ne i, j \notin \mathcal{N}_i(t)$ and $(\mathcal{L}_{\sigma(t)})_{ii} = |\mathcal{N}_i(t)|$. Recall that t_k, $k \ge 0$ are all the switching moments of switching graph $\mathcal{G}_\sigma$ with $t_{k+1} - t_k \ge \tau$, $\forall k$. It is easy to see that we can add some new "switching moments" in $\{t_k\}_{k\ge0}$, denoted as $\{t'_k\}_{k\ge0}$, such that $2\tau \ge t'_{k+1} - t'_k \ge \tau$, $\forall k$. Therefore,

$$|x(t'_{k+1})|_{X_0} \le e^{-\mathcal{L}_{\sigma(t'_k)}(t'_{k+1}-t'_k)}|x(t'_k)|_{X_0} + \int_{t'_k}^{t'_{k+1}} e^{-\mathcal{L}_{\sigma(t'_k)}(t'_{k+1}-t)}\Big(-\frac{D(t)y(t)}{\phi} + \phi\alpha^+_t\tan\theta^+_t\mathbf{1}\Big)dt.$$

Note that for any $s > 0$, $e^{-s\mathcal{L}_{\sigma(t'_k)}}$ is a stochastic matrix (with nonnegative entries and all row sums are ones) and the graph $\mathcal{G}_{\sigma(t'_k)}$ is a subgraph of the graph associated with matrix $e^{-s\mathcal{L}_{\sigma(t'_k)}}$. Then applying the similar arguments given in the proof of Theorem 4.1 in Lou et al. (2014) we can show that $\liminf_{t\to\infty}\sum_{i=1}^n |x_i(t)|^2_{X_i} = 0$.

Then there is a subsequence $\{s_k\}_{k\ge0}$ with $\lim_{k\to\infty} s_k = \infty$ such that $\lim_{k\to\infty}|x_i(s_k)|_{X_i} = 0$ for all i. Because we have shown that consensus is achieved in Step (i), $\lim_{k\to\infty}|x_i(s_k)|_{X_j} = 0$ for all i, j, which leads to $\lim_{k\to\infty} h_i(s_k) = 0$ for all i. Thus, $h^* = \lim_{t\to\infty} h_i(t) = 0$, which contradicts $h^+_i = h^-_i = h^* > 0$. It follows that $h^+_i = h^-_i = h^* = 0$ and then the conclusion is proved. □

Remark 5.1. When the intersection set of X_is is nonempty, SDOP (2) is equivalent to CIP of finding a point in X_0 (Deutsch, 1983; Gubin et al., 1967; Lou et al., 2013, 2014; Meng et al., 2013; Nedić et al., 2010; Shi & Johansson, 2012; Shi et al., 2013). The optimal consensus algorithm based on the exact projection presented in Shi et al. (2013) is a special case of (3) with taking $\alpha_t \equiv 1$ and $\theta_{i,t} \equiv 0$, which is consistent with Theorem 5.1. Theorem 5.1 is also consistent with the convex intersection computation results of discrete-time algorithms in Lou et al. (2014), Meng et al. (2013) and Nedić et al. (2010).

6. Empty intersection case

At first we give a convergence result for the empty intersection case (that is, $\cap_{i=1}^n X_i = \emptyset$).

Theorem 6.1. *Suppose* **A1–A3** *hold,* $\mathcal{G}_{\sigma(t)}$, $t \ge 0$ *are undirected and* $\cap_{i=1}^n X_i = \emptyset$. *Then SDOP is solved by system* (4) *if* $\int_0^\infty \alpha_t dt = \infty$, $\int_0^\infty \alpha_t^2 dt < \infty$ *and* $\int_0^\infty \alpha_t\tan\theta^+_t dt < \infty$.

Proof. We rewrite (4) as $\dot x_i(t) = \sum_{j\in\mathcal{N}_i(t)}(x_j(t) - x_i(t)) + \alpha_t(P^h_{X_i}(x_i(t)) - x_i(t)) + \phi_i(t)$, where $\phi_i(t) = (\alpha_{i,t} - \alpha_t)(P^h_{X_i}(x_i(t)) - x_i(t))$. Denote $\bar x(t) = \frac{1}{n}\sum_{i=1}^n x_i(t)$. From the definition of $P^h_{X_i}$, we have $|P^h_{X_i}(x_i(t)) - x_i(t)| = \frac{|x_i(t)|_{X_i}}{\cos\theta_{i,t}} \le \frac{\eta}{\cos\theta_{i,t}}$, where $\eta = \sup_{i,j,t}\{|x_i(t) - x^*|, |\bar x(t)|_{X_i}, |x_i(t)|_{X_j}\}$ is a finite number by Theorem 4.1. Moreover, it follows from (10) that

$$|\alpha_{i,t} - \alpha_t| \le (C_{i,t} - 1)\alpha_t \le \frac{1}{\sin\mu_i}\alpha_t\sin\theta_{i,t}.$$

Therefore,

$$|\phi_i(t)| \le \frac{\eta}{\sin\mu_i}\alpha_t\tan\theta_{i,t}. \tag{14}$$

Take $x^* \in \arg\min\sum_{i=1}^n |x|^2_{X_i}$. Let $t \notin \varDelta$. Because $\mathcal{G}_{\sigma(t)}$ is undirected,

$$\frac{d\sum_{i=1}^n |x_i(t) - x^*|^2}{dt} \le 2\alpha_t\sum_{i=1}^n\Big\langle x_i(t) - x^*, P^h_{X_i}(x_i(t)) - x_i(t)\Big\rangle + 2\sum_{i=1}^n\langle x_i(t) - x^*, \phi_i(t)\rangle. \tag{15}$$

Then we estimate the first term in (15). Note that

$$\langle x_i(t)-x^*, P^h_{X_i}(x_i(t))-P_{X_i}(x_i(t))\rangle \le |x_i(t)-x^*|\tan\theta_{i,t}|x_i(t)|_{X_i} \le \eta^2\tan\theta_t^+. \quad (16)$$

We also have

$$\sum_{i=1}^n\langle x_i(t)-x^*, P_{X_i}(x_i(t))-x_i(t)\rangle = -\Big\langle \bar{x}(t)-x^*, \sum_{i=1}^n(\bar{x}(t)-P_{X_i}(\bar{x}(t)))\Big\rangle + \varrho(t), \quad (17)$$

where $\varrho(t)=\sum_{i=1}^n\langle x_i(t)-\bar{x}(t), P_{X_i}(\bar{x}(t))-\bar{x}(t)\rangle+\sum_{i=1}^n\langle x_i(t)-x^*, P_{X_i}(x_i(t))-P_{X_i}(\bar{x}(t))+\bar{x}(t)-x_i(t)\rangle$.

Clearly, the first term in $\varrho(t)$ is not greater than $\eta\sum_{i=1}^n|x_i(t)-\bar{x}(t)|$ and the second term in $\varrho(t)$ is not greater than $2\eta\sum_{i=1}^n|x_i(t)-\bar{x}(t)|$ by Lemma 2.1(iv). Moreover, by (14), the second term in (15) is not greater than $2n\eta^2\alpha_t\tan\theta_t^+/\sin\mu_-$, where $\mu_-=\min_{1\le i\le n}\mu_i$. Denote $\psi(t)=\langle\bar{x}(t)-x^*, \sum_{i=1}^n(\bar{x}(t)-P_{X_i}(\bar{x}(t)))\rangle$, $\varsigma(t)=6\eta(\sum_{i=1}^n\alpha_t|x_i(t)-\bar{x}(t)|)+2n\eta^2(1+\frac{1}{\sin\mu_-})\alpha_t\tan\theta_t^+$. In light of (15)–(17), we have

$$\frac{d\sum_{i=1}^n|x_i(t)-x^*|^2}{dt} \le -2\alpha_t\psi(t)+\varsigma(t)\le\varsigma(t) \quad (18)$$

because $\psi(t)$ is nonnegative, following from the convexity of objective function f, i.e., $\psi(t)=\langle\bar{x}(t)-x^*, \frac{1}{2}\nabla f(\bar{x}(t))\rangle\ge\frac{1}{2}(f(\bar{x}(t))-f(x^*))\ge 0$.

By a similar analysis technique given in the proof of Lemma 8 in Nedić et al. (2010), we can show $\int_0^\infty\sum_{i=1}^n\alpha_t|x_i(t)-\bar{x}(t)|dt<\infty$. This combined with $\int_0^\infty\alpha_t\tan\theta_t^+dt<\infty$ and (18) implies that $\lim_{t\to\infty}\sum_{i=1}^n|x_i(t)-x^*|^2$ is a finite number. Hence, it follows from (18) that $2\int_0^\infty\alpha_t\psi(t)dt=\int_0^\infty\alpha_t\langle\bar{x}(t)-x^*, \nabla f(\bar{x}(t))\rangle dt<\infty$. Due to $\int_0^\infty\alpha_t dt=\infty$, there exists a subsequence $\{s_r\}_{r\ge0}$ such that $\lim_{r\to\infty}\langle\bar{x}(s_r)-x^*, \nabla f(\bar{x}(s_r))\rangle=0$. Since system states are bounded, without loss of generality we assume $\lim_{r\to\infty}\bar{x}(s_r)=\hat{x}$ for some $\hat{x}$ (otherwise we can further find a subsequence of $\{s_r\}_{r\ge0}$). Since ∇f is continuous, $\langle\hat{x}-x^*, \nabla f(\hat{x})\rangle=0$, which leads to $f(x^*)\ge f(\hat{x})+\langle x^*-\hat{x}, \nabla f(\hat{x})\rangle=f(\hat{x})$. Thus, $\hat{x}\in\arg\min f$.

Replacing x^* with $\hat{x}$, we can similarly show that $\lim_{t\to\infty}\sum_{i=1}^n|x_i(t)-\hat{x}|^2$ is also a finite number denoted as ρ. Moreover, the uniform continuity of α_t and $\int_0^\infty\alpha_t^2dt<\infty$ imply $\lim_{t\to\infty}\alpha_t=0$. As a result, the consensus is achieved by Lemma 2.4 and then $\lim_{r\to\infty}x_i(s_r)=\hat{x}$. Hence, $\rho=0$, and therefore, $\lim_{t\to\infty}x_i(t)=\hat{x}$ for all i, which leads to the conclusion. □

From Theorems 5.1 and 6.1, we find that the sufficient optimal consensus conditions are essentially different in these two cases. In addition to the conditions in the nonempty intersection case, the square integrability condition is usually required in the empty intersection case.

Theorem 6.1 showed that all agents consensually converge to an optimal solution of $\min\sum_{i=1}^n|x|^2_{X_i}$ under certain conditions. Next, we show some properties of the optimal solution set of $\min\sum_{i=1}^n|x|^2_{X_i}$, denoted as X^*. According to Lemma 2.2, an optimal solution $x^*\in X^*$ must satisfy $\nabla\sum_{i=1}^n|x^*|^2_{X_i}=2\sum_{i=1}^n(x^*-P_{X_i}(x^*))=0$, or equivalently, $x^*=\frac{\sum_{i=1}^n P_{X_i}(x^*)}{n}$. Then we have the following results.

Theorem 6.2. (i) *For any $x^*\in X^*, y^*\in X^*$, we have $x^*-P_{X_i}(x^*)=y^*-P_{X_i}(y^*)$, $i=1,\dots,n$;*
(ii) *For any i, either $X^*\subseteq X_i$ or $X^*\cap X_i=\emptyset$;*
(iii) *Let $x^*\in X^*$, $x^*\notin X_i$ for some i. Then $X^*\cap$ line $(x^*, P_{X_i}(x^*))=\{x^*\}$.*

Proof. (i) It follows from $x^*=\frac{\sum_{i=1}^n P_{X_i}(x^*)}{n}$, $y^*=\frac{\sum_{i=1}^n P_{X_i}(y^*)}{n}$ and Lemma 2.1(iv) that

$$|x^*-y^*|=\left|\frac{\sum_{i=1}^n(P_{X_i}(x^*)-P_{X_i}(y^*))}{n}\right| \le\frac{\sum_{i=1}^n|P_{X_i}(x^*)-P_{X_i}(y^*)|}{n}\le|x^*-y^*|.$$

Therefore, $|P_{X_i}(x^*)-P_{X_i}(y^*)|=|x^*-y^*|$ for all i, which implies the conclusion by Lemma 2.1(vi).

(ii) This is straightforward from (i).

(iii) Let $z^*\in X^*\cap$ line$(x^*, P_{X_i}(x^*))$, $z^*\ne x^*$. If z^* locates the half-line with $P_{X_i}(x^*)$ as the starting point and $x^*-P_{X_i}(x^*)$ as the direction, then $P_{X_i}(z^*)=P_{X_i}(x^*)$. Therefore, $x^*-P_{X_i}(x^*)\ne z^*-P_{X_i}(z^*)$, which contradicts what was proven in (i) since both x^* and z^* are optimal solutions. If z^* locates the half-line with $P_{X_i}(x^*)$ as the starting point and $P_{X_i}(x^*)-x^*$ as the direction, then $P_{X_i}(x^*)$ is also an optimal solution since the optimal solution set X^* is a convex set and $P_{X_i}(x^*)$ can be written as a convex combination of x^*, z^*. Then $0=P_{X_i}(x^*)-P_{X_i}(P_{X_i}(x^*))\ne x^*-P_{X_i}(x^*)$, which also yields a contradiction since both x^* and $P_{X_i}(x^*)$ are optimal solutions. Thus, the conclusion follows. □

7. Numerical examples

In this section, we provide an example to illustrate the above results. Consider a network of three agents with node set $\mathcal{V}=\{1,2,3\}$. The convex set X_i of each agent i is the ball in $\mathbb{R}^2$ with center c_i and radius r_i. Let $\alpha_t=\frac{20}{t+20}$, $\theta_{i,t}=\frac{1}{t+50}$, which satisfy the conditions in Theorems 5.1 and 6.1. We next present the state trajectories of the three agents for the nonempty and empty intersection cases from time $t=0$ to $t=2000$, respectively.

(i) Nonempty intersection case with $c_1=(-1,0)$, $c_2=(1,0)$, $c_3=(0,-2)$, $r_1=2$, $r_2=1$, $r_3=2$. The graphs are periodically switching over the two directed graphs $G_1=(\mathcal{V},\mathcal{E}_1)$, $G_2=(\mathcal{V},\mathcal{E}_2)$ with period 1, where $\mathcal{E}_1=\{(2,1),(3,2)\}$, $\mathcal{E}_2=\{(1,3)\}$. The initial conditions are $x_1(0)=(-4,3)$, $x_2(0)=(3,5)$, $x_3(0)=(-6,-3)$, which are marked as ∘ in Fig. 3.

(ii) Empty intersection case with $c_1=(-\sqrt{3},0)$, $c_2=(\sqrt{3},0)$, $c_3=(0,-3)$, $r_1=r_2=r_3=1$. In this case, the (unique) optimal solution is $(0,-1)$. The graphs are periodically switching over the two undirected graphs $G_1=(\mathcal{V},\mathcal{E}_1)$, $G_2=(\mathcal{V},\mathcal{E}_2)$ with period 1, where $\mathcal{E}_1=\{(3,2)\}$, $\mathcal{E}_2=\{(1,2)\}$. The initial conditions are $x_1(0)=(-3,3)$, $x_2(0)=(4,2)$, $x_3(0)=(-5,-3)$, which are marked as ∘ in Fig. 4.

8. Conclusions

In this paper, a continuous-time method was proposed to cooperatively solve the SDOP by a group of agents that could only obtain their approximate projections and the communication graph among agents was UJSC. It was shown that system states were always bounded for any constant approximate angle, and uniformly bounded for any stepsize with inferior limit greater than zero. Both nonempty intersection and empty intersection cases of convex sets were investigated with respective sufficient conditions.

Acknowledgments

The authors would like to thank Dr. Guilin Yang for discussions about geometric analysis, Mr. Peng Yi for his generous help on numerical simulations, and reviewers for constructive comments.

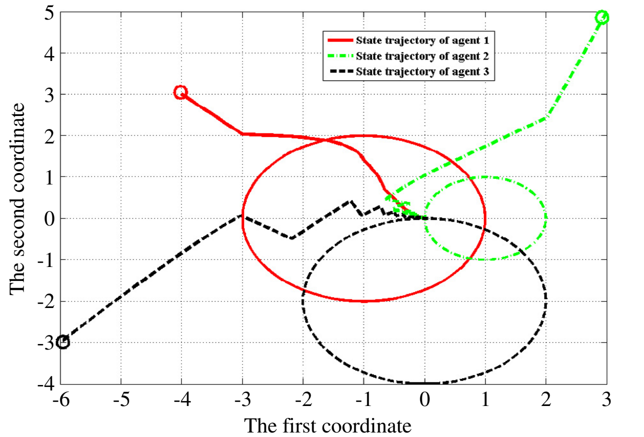

Fig. 3. In the nonempty intersection case, all agents converge to a common point in the intersection set.

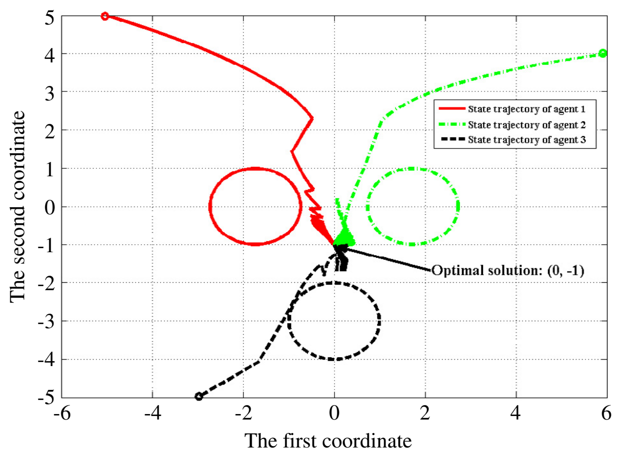

Fig. 4. In the empty intersection case, all agents converge to the unique optimal solution $(0, -1)$.

Appendix. Proof of Lemma 4.1

In this proof, we simplify θ_i, μ_i, and X_i as θ, μ, and X, respectively. Consider the following relation

$$cone(v, \mathbf{C}_X(v, \theta) \cap \mathbf{b}(v, X)) = \mathbf{C}_X(v, \theta) \tag{19}$$

where $\text{cone}(v, M) = \{v + \lambda(z - v) | \lambda \geq 0, z \in M\}$ for some set $M \subseteq \mathbb{R}^m$. We claim that (19) holds for θ^* and any $v \in S \backslash X$ with sufficiently small $|v|_X$, θ^* is the approximate angle given in **A3**.

We first show by contradiction $\text{int}(X) \cap \text{line}(v, P_X(v)) \neq \emptyset$ for any $v \in bd(S)$. For a regular surface, its tangent plane at boundary point z consists of the tangent vectors at point z of all curves passing z. Suppose that there is $v \in bd(S)$ with $\text{int}(X) \cap \text{line}(v, P_X(v)) = \emptyset$. Then, by convex set separation Theorem 11.3 on page 97 in Rockafellar (1972), there exists a hyperplane $\mathcal{H}$ separating X and $\text{line}(v, P_X(v))$ properly. As a result, $\mathcal{H}$ must contain $\text{line}(v, P_X(v))$. Let $\mathbf{n}$ be the unit normal vector of $\mathcal{H}$ with $\angle(\mathbf{n}, z - P_X(v)) \geq \pi/2$ for any $z \in X$, and $\mathcal{H}_v$ the tangent plane of $bd(X)$ at $P_X(v)$. Then $\mathbf{n} \in \mathcal{H}_v$ since $v - P_X(v)$ is a normal vector of tangent plane $\mathcal{H}_v$. However, it is not possible that there exists a curve on $bd(X)$ with tangent vector $\mathbf{n}$ at $P_X(v)$, which yields a contradiction.

Let $z \in \text{int}(X) \cap \text{line}(v, P_X(v))$. Then exists $\epsilon > 0$ such that $\mathbf{B}(z, \epsilon) \subseteq X$. Let $y \in bd(\mathbf{B}(z, \epsilon))$ be the point for which $\angle(y - z, v - z) = \pi/2$. Clearly, (19) holds for $v, \theta(v)$, where $\theta(v) = \angle(y - v, z - v) > 0$. Therefore, the claim follows.

We now show the conclusion by contradiction. Suppose that there is a sequence $\{v_k\}_{k\geq 0}$ with $v_k \in S\backslash X$ and $P_X^a(v_k) \neq P_X(v_k)$ such that $\lim_{k\to\infty} \mu(v_k) = 0$. Without loss of generality, we assume $\lim_{k\to\infty} v_k =: v^* \in S\backslash\text{int}(X)$.

We first consider the case of $v^* \in S\backslash X$. In the case of $P_X^a(v^*) \neq P_X(v^*)$, by the continuity we have $0 = \lim_{k\to\infty} \mu(v_k) = \mu(v^*) > 0$, which yields a contradiction. In the case of $P_X^a(v^*) = P_X(v^*)$, we have $\lim_{k\to\infty} \theta(v_k) = 0$, which implies $\lim_{k\to\infty} \angle(v_k - P_X(v_k), P_X^a(v_k) - P_X(v_k)) = \pi$ along with $\lim_{k\to\infty} \mu(v_k) = 0$. This, however, is impossible since the surface $bd(X)$ is regular.

We next consider the case of $v^* \in bd(X)$. Let $r > 0$ be a sufficiently small number such that (19) holds for θ^* and any $v + r\mathbf{n}(v)$ with $|v - v^*| \leq r$ and $v \in bd(X)$, where $\mathbf{n}(v)$ is the unit normal vector of the tangent plane of $bd(X)$ at v. Denote $z := v + r\mathbf{n}(v)$. Take arbitrarily a point $\hat{y} := \hat{y}(z) \in bd(\mathbf{C}_X(z, \theta^*)) \cap \mathbf{b}(z, X) \cap \text{aff}\{v, z, P_X^a(z)\}$ such that $\angle(v - z, \hat{y} - z) = \theta^*$. Then $\mu(z) \geq \angle(v - \hat{y}, z - \hat{y})$. Moreover, it is not hard to find that, for any z_1, z_2 such that $z_1 \notin X$, $z_2 \notin X$, $P_X(z_1) = P_X(z_2)$, $|z_2|_X > |z_1|_X$, and with (19) holding for both (z_1, θ^*) and (z_2, θ^*), we have $\beta(z_1) \geq \beta(z_2)$, where $\beta(z) = \inf_{\bar{y}\in bd(\mathbf{C}_X(z,\theta^*))\cap\mathbf{b}(z,X)} \angle(P_X(z) - \bar{y}, z - \bar{y})$. From the proceeding two inequalities we conclude that $\mu(v_k) \geq \inf_{v\in bd(X), |v-v^*|\leq r} \beta(v + r\mathbf{n}(v)) > 0$ for all sufficiently large k, which yields a contradiction. We complete the proof. □

References

Aubin, J., & Cellina, A. (1984). *Differential inclusions*. Berlin: Springer-Verlag.

Bertsekas, D. P., & Tsitsiklis, J. (1989). *Parallel and distributed computation: numerical methods*. Englewood Cliffs, NJ: Princeton Hall.

Danskin, J. (1966). The theory of max–min, with applications. *SIAM Journal on Applied Mathematics*, *14*(4), 641–664.

Deutsch, F. (1983). Rate of convergence of the method of alternating projections. In B. Brosowski, & F. Deutsch (Eds.), *Parametric optim. approx.: Vol. 76* (pp. 96–107). Basel, Switzerland: Birkhäuser.

do Carmo, M. (1976). *Differential geometry of curves and surfaces*. Englewood Cliffs, NJ: Prentice-Hall.

Droge, G., Kawashima, H., & Egerstedt, M. (2014). Continuous-time proportional-integral distributed optimization for networked systems. *Journal of Decision and Control*, *1*(3), 191–213.

Francis, R. L., McGinnis, L. F., & White, J. A. (1992). *Facility layout and location: an analytical approach* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Gharesifard, B., & Cortés, J. (2013). Distributed convergence to Nash equilibria in two-network zero-sum games. *Automatica*, *49*(6), 1683–1692.

Gharesifard, B., & Cortés, J. (2014). Distributed continuous-time convex optimization on weight-balanced digraphs. *IEEE Transactions on Automatic Control*, *59*(3), 781–786.

Godsil, C., & Royle, G. (2001). *Algebraic graph theory*. New York: Springer-Verlag.

Gubin, L., Polyak, B., & Raik, E. (1967). The method of projections for finding the common point of convex sets. *USSR Computational Mathematics and Mathematical Physics*, *7*(6), 1211–1228.

Jakovetic, D., Xavier, J., & Moura, J. M. F. (2011). Cooperative convex optimization in networked systems: Augmented lagrangian algorithms with directed gossip communication. *IEEE Transactions on Signal Processing*, *59*(8), 3889–3902.

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

Kvaternik, K., & Pavel, L. (2012). A continuous-time decentralized optimization scheme with positivity constraints. In *51st IEEE conference on decision and control* (pp. 6801–6807).

Lin, P., & Ren, W. (2012). Distributed shortest distance consensus problem in multi-agent systems. In *51st IEEE conference on decision and control, 2012* (pp. 4696–4701).

Lou, Y., Shi, G., Johansson, K. H., & Hong, Y. (2013). Convergence of random sleep algorithms for optimal consensus. *Systems & Control Letters*, *62*(12), 1196–1202.

Lou, Y., Shi, G., Johansson, K. H., & Hong, Y. (2014). Approximate projected consensus for convex intersection computation: Convergence analysis and critical error angle. *IEEE Transactions on Automatic Control*, *59*(7), 1722–1736.

Lu, J., & Tang, C. Y. (2012). Zero-gradient-sum algorithms for distributed convex optimization: The continuous-time case. *IEEE Transactions on Automatic Control*, *57*(9), 2348–2354.

Lu, J., Tang, C. Y., Regier, P. R., & Bow, T. D. (2011). Gossip algorithms for convex consensus optimization over networks. *IEEE Transactions on Automatic Control*, *56*(12), 2917–2923.

Meng, W., Xiao, W., & Xie, L. (2013). A projection based fully distributed approach for source localization in wireless sensor networks. *Adhoc & Sensor Wireless Networks*, *18*(1–2), 131–158.

Nedić, A., & Ozdaglar, A. (2008). Approximate primal solutions and rate analysis for dual subgradient methods. *SIAM Journal on Optimization*, *19*(4), 1757–1780.

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

Nedić, A., Ozdaglar, A., & Parrilo, P. A. (2010). Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control*, *55*(4), 922–938.
Pardalos, P. M., & Romeijn, H. E. (2002). *Handbook of global optimization. Vol. 2*. London: Kluwer Academic Publishers.
Rockafellar, R. T. (1972). *Convex analysis*. New Jersey: Princeton University Press.
Shi, G., & Johansson, K. H. (2012). Randomized optimal consensus of multi-agent systems. *Automatica*, *48*(12), 3018–3030.
Shi, G., & Johansson, K. H. (2013). Multi-agent robust consensus: Convergence analysis and application. *SIAM Journal on Control and Optimization*, *51*(5), 3673–3691.
Shi, G., Johansson, K. H., & Hong, Y. (2013). Reaching an optimal consensus: Dynamical systems that compute intersections of convex sets. *IEEE Transactions on Automatic Control*, *58*(3), 610–622.
Wang, J., & Elia, N. (2010). Control approach to distributed optimization. In *Forty-Eighth annual allerton conference* (pp. 557–561).
Wang, J., & Elia, N. (2011). A control perspective for centralized and distributed convex optimization. In *50th IEEE conference on decision and control and European control conference* (pp. 3800–3805).
Zhang, Y., Lou, Y., Hong, Y., & Xie, L. (2015). Distributed projection-based algorithms for source localization in wireless sensor networks. *IEEE Transactions on Wireless Communications*, *14*(6), 3131–3142.

Youcheng Lou received his Ph.D. degree in Complex Systems and Control from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences (CAS), Beijing in July 2013. From July 2013 he is a postdoctoral researcher of the National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, CAS. From Jan. 2016 he is a postdoctoral fellow of the Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, funded by Hong Kong Scholars Program. His research interests include multi-agent networks, distributed optimization and the applications of distributed methods in operations research.

Yiguang Hong received the B.Sc. and M.Sc. degrees from Peking University, Beijing, China, and the Ph.D. degree from the Chinese Academy of Sciences (CAS), Beijing. He is currently a Professor in the Institute of Systems Science, Academy of Mathematics and Systems Science, CAS, and serves as the Director of Key Lab of Systems and Control, CAS and the Director of the Information Technology Division, National Center for Mathematics and Interdisciplinary Sciences, CAS. His research interests include nonlinear dynamics and control, multi-agent systems, distributed optimization, social networks, and software reliability.

Prof. Hong serves as Editor-in-Chief of the Control Theory and Technology and Deputy Editor-in-Chief of Acta Automatica Sinca. Also he serves or served as Associate Editors for several journals including the IEEE Transactions on Automatic Control, IEEE Transactions on Control of Network Systems, IEEE Control Systems Magazine, Nonlinear Analysis: Hybrid Systems, and Kybernetika. Moreover, he was the chairs of IEEE Control Systems Society (CSS) membership and public information committee and CSS chapter activities committee. Additionally, he is a recipient of the Guan Zhao Zhi Award at the Chinese Control Conference, Young Author Prize of the IFAC World Congress, Young Scientist Award of CAS, the Youth Award for Science and Technology of China, and the National Natural Science Prize of China.

Shouyang Wang received his Ph.D. degree in Operations Research from the Institute of Systems Science, Chinese Academy of Sciences (CAS), Beijing in 1986. He is currently a Bairen Distinguished Professor of the Academy of Mathematics and Systems Science, CAS and a Changjiang Chair Professor of the School of Economics and Management, University of Chinese Academy of Sciences. He is the president of International Society of Knowledge and Systems Sciences, and has published over 30 monographs and 250 papers in leading journals including IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Automatic Control, Journal of Econometrics, Journal of Banking and Finance, and Quantitative Finance, etc. His research interests include data mining, supply chain management, decision support systems and financial engineering.