Approximate Projected Consensus for Convex Intersection Computation: Convergence Analysis and Critical Error Angle

Youcheng Lou, Guodong Shi, Karl Henrik Johansson, *Fellow, IEEE*, and Yiguang Hong, *Senior Member, IEEE*

***Abstract*—In this paper, we study an approximate projected consensus algorithm for a network to cooperatively compute the intersection of convex sets, where each set corresponds to one network node. Instead of assuming exact convex projection that each node can compute, we allow each node to compute an approximate projection with respect to its own set. After receiving the approximate projection information, nodes update their states by weighted averaging with the neighbors over a directed and time-varying communication graph. The approximate projections are related to projection angle errors, which introduces state-dependent disturbance in the iterative algorithm. Projection accuracy conditions are presented for the considered algorithm to converge. The results indicate how much projection accuracy is required to ensure global consensus to a point in the intersection set when the communication graph is uniformly jointly strongly connected. In addition, we show that $\pi/4$ is a critical angle for the error of the projection approximation to ensure the boundedness. Finally, the results are illustrated by simulations.**

***Index Terms*—Approximate projection, intersection computation, multi-agent systems, optimal consensus.**

I. Introduction

DISTRIBUTED analysis and control have drawn increasing research attention in various areas of engineering, physics, computer science, and social science. Collective tasks can be accomplished cooperatively for a group of autonomous agents via local information exchange and distributed protocol design. Distributed solutions to some simple but global goals such as consensus, formation, and aggregation have been extensively studied in the literature [8]–[15], [18]–[22]. Recently, distributed optimization problems have also arisen due to its wide applications in networked control systems and wireless communication networks [23]–[30], [32]–[37].

Manuscript received January 3, 2013; revised February 3, 2013, July 14, 2013, December 8, 2013, and January 24, 2014; accepted February 23, 2014. Date of publication February 28, 2014; date of current version June 19, 2014. This work was supported in part by the NNSF of China under Grants 61174071, 61120106011, by the Knut and Alice Wallenberg Foundation, and by the Swedish Research Council. This paper was presented in part at the 10th World Congress on Intelligent Control and Automation, Beijing, China, 2012. Recommended by Associate Editor S. Zampieri.

Y. Lou and Y. Hong are with Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (e-mail: louyoucheng@amss.ac.cn; yghong@iss.ac.cn).

G. Shi and K. H. Johansson are with ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Stockholm 10044, Sweden (e-mail: guodongs@kth.se; kallej@ee.kth.se).

Digital Object Identifier 10.1109/TAC.2014.2309261

A standard distributed optimization formulation for a multi-agent network is to cooperatively minimize a sum of convex functions, where each component is known only by a particular node. Subgradient-based incremental methods were established via deterministic or randomized iteration, where each node is assumed to be able to compute a local subgradient value of its objective function [23]–[25], [27]–[29]. Non-subgradient-based methods also showed up. For instance, a non-gradient-based algorithm was proposed in [33], [34], where each node starts at its own optimal solution and updates using a pairwise equalizing protocol, and an augmented Lagrangian method was introduced in [37].

If the optimal solution set of its own objective function can be obtained for each node, the optimization problem can be converted to a set intersection computation problem somehow [35], [36]. Moreover, in many distributed optimization problems the state of each node is restricted within a particular constrained set, and then the intersection of these constraint sets becomes important because it is where the global optimal solutions for the network locate [30]. All of these lead to a question on how to compute the intersection of several convex sets in a distributed manner. In fact, convex intersection computation problem is a classical problem in optimization [40]–[43]. An extensive survey about the convex intersection computation problem can be found in [43]. An "alternating projection algorithm" was used to be a standard centralized solution, where the projection is carried out alternatively onto each set [40]–[42]. Several years ago, the "projected consensus algorithm" was presented as a decentralized version of the alternating projection algorithm, where each node alternatively projects onto its own set and averages with its neighbors, with convergence analysis under time-varying directed interconnections [30]. Following this work, a flip-coin algorithm was introduced when each node randomly chooses projection or averaging based on Bernoulli processes, with almost sure convergence to an optimal consensus [36]. A deterministic dynamical system solution was then given in [35], where the network reaches a global optimal consensus using simple continuous-time local control laws. In all these algorithms, each node knows the exact convex projection of its current state onto its objective set.

However, in many practical applications, the exact convex projection is hard to compute due to measurement and computation inaccuracy. The objective of this paper is to solve the convex intersection problem when the exact projection cannot be obtained and how much error can be tolerated by the optimization algorithm. The contribution of this paper is summarized as follows:

- At first, we propose an approximate projected consensus algorithm to solve the convex intersection computation problem. Instead of assuming the exact convex projection, each node can only compute an approximate projection point located in a bounded convex projection cone determined by the current state and a projection angle error. In fact, the approximate projection algorithm extends the projected consensus algorithm, but it introduces a challenging state-dependent disturbance in the iterative algorithm.
- The communication graph for the distributed optimization is supposed to be directed and time-varying. With uniformly jointly strongly connected conditions, we show that the whole network can achieve a global consensus within the intersection of all convex sets when sufficient projection accuracy can be guaranteed. Moreover, a necessary and sufficient condition is given to ensure a global intersection computation in a special case.
- The critical angle error of the projection approximation is discussed if the angle is a constant. It is shown that $\pi/4$ is a critical angle error to ensure the boundedness of the algorithm.

The paper is organized as follows. Section II gives some basic concepts on graph theory and convex analysis. Section III introduces a distributed model and formulates the problem of interest. Section IV presents the main results of the proposed approximate projected consensus algorithm, while Section V provides all the proofs. Section VI gives some numerical examples and, finally, Section VII shows some concluding remarks.

II. Preliminaries

In this section, we introduce preliminary knowledge on graph theory [6] and convex analysis [2].

A digraph (directed graph) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is an ordered pair of node set $\mathcal{V} = \{1, 2, \ldots, n\}$ and arc set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. We call node j a neighbor of node i if $(j, i) \in \mathcal{E}$. We denote $\mathcal{N}_i$ as the set of neighbors of node i, that is, $\mathcal{N}_i = \{j \in \mathcal{V} | (j, i) \in \mathcal{E}\}$. In this paper, we assume $(i, i) \in \mathcal{E}$ for all i. A path from i to j in digraph $\mathcal{G}$ is an alternating sequence $i_1 e_1 i_2 e_2 \cdots i_{p-1} e_{p-1} i_p$ of nodes $i_r, 1 \le r \le p$ and arcs $e_r = (i_r, i_{r+1}) \in \mathcal{E}, 1 \le r \le p-1$, $i_1 = i, i_p = j$. Graph $\mathcal{G}$ is said to be strongly connected if there exists a path from i to j for each pair of nodes $i, j \in \mathcal{V}$.

A set $K \subseteq \Re^m$ is said to be convex if $\lambda x + (1-\lambda) y \in K$ for any $x, y \in K$ and $0 < \lambda < 1$, and is said to be a convex cone if $\lambda_1 x + \lambda_2 y \in K$ for any $x, y \in K$ and $\lambda_1, \lambda_2 \ge 0$. For a closed convex set K in $\Re^m$, we can associate to any $x \in \Re^m$ a unique element $P_K(x) \in K$ satisfying $|x - P_K(x)| = \inf_{y \in K} |x - y|$, which is denoted as $|x|_K$, where $|\cdot|$ denotes the Euclidean norm and P_K is the projection operator onto K. For a closed convex set K, if $x \notin K$, there is a supporting hyperplane to K at $P_K(x)$ with normal direction $x - P_K(x)$. The angle between vectors y and z is denoted as $\angle(y, z) \in [0, \pi]$ with $\cos \angle(y, z) = \langle y, z \rangle / (|y||z|)$, where $\langle y, z \rangle$ denotes the Euclidean inner product of y and z.

The following properties hold for the projection operator P_K.

Lemma 2.1: Let K be a closed convex set in $\Re^m$. Then:

(i) $|P_K(x) - P_K(y)| \le |x - y|$ for any x and y;
(ii) $||x|_K - |y|_K| \le |x - y|$ for any x and y;
(iii) $P_K(\lambda x + (1-\lambda) P_K(x)) = P_K(x)$ for any x and $0 < \lambda < 1$;
(iv) $|P_K(x) - y|^2 \le |x - y|^2 - |x|_K^2$ for any $x \in \Re^m$ and $y \in K$.

Here (i) is the standard non-expansiveness property; (ii) comes from Exercise 1.2 (c) on page 23 [3]; (iii) is a special case of Proposition 1.3 on page 24 [3], and (iv) is taken from Lemma 1 (b) in [30].

The next lemma can be found in [36].

Lemma 2.2: Let K and $K_0 \subseteq K$ be two closed convex sets in $\Re^m$. We have

$$|P_K(x)|_{K_0}^2 + |x|_K^2 \le |x|_{K_0}^2 \text{ for any } x.$$

A function $f(\cdot) : \Re^m \to \Re$ is said to be convex if $f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y)$ for any $x, y \in \Re^m$ and $0 < \lambda < 1$. A function f is said to be concave if $-f$ is convex. Here is a useful lemma for the following analysis (see Example 3.16 on page 88 [4]).

Lemma 2.3: Let K be a closed convex set in $\Re^m$. Then $f(z) = |z|_K$ is a convex function.

III. Problem Formulation

In this section, we introduce the intersection computation problem and the approximate projected consensus algorithm.

Consider a network consisting of n agents. Each agent i is associated with a set $X_i \subseteq \Re^m$ and X_i is known only by agent i. All these sets have a nonempty intersection, i.e., $\bigcap_{i=1}^n X_i \neq \emptyset$. The objective of the network is to find a point in the intersection set in a distributed way.

Remark 3.1: The intersection computation problem can be equivalently converted into the following distributed optimization problem: a group of n agents should reach a consensus and cooperatively solve

$$\min_{x \in \Re^m} \sum_{i=1}^n f_i(x)$$

where $f_i : \Re^m \to \Re$ is the cost function of agent i and known only by agent i. The problems are equivalent if $X_i = \{y | f_i(y) = \min_{x \in \Re^m} f_i(x)\}, 1 \le i \le n$ are nonempty and have a nonempty intersection.

The communication over the network is described by a sequence of digraphs, $\mathcal{G}_k = (\mathcal{V}, \mathcal{E}(k)), k \ge 0$ with $\mathcal{V} = \{1, 2, \ldots, n\}$. Recall that node j is a neighbor of node i at time k if there is an arc $(j, i) \in \mathcal{E}(k)$. Let $\mathcal{N}_i(k)$ denote the set of neighbors of node i at time k, and $a_{ij}(k)$ represent the weight of arc (j, i) at time k.

A. Approximate Projection

Projection-based methods have been widely used to solve various problems in the literature, e.g., in projected consensus [30], convex intersection computation [41], [42], and distributed computation [5]. In most of the literature, the projection

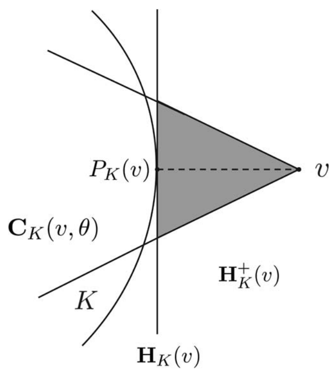


Fig. 1. Set marked by the shaded area is the approximate projection of v onto K.

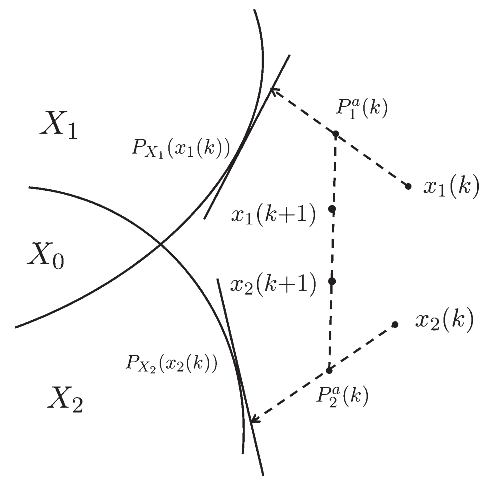


Fig. 2. Approximate projected consensus algorithm.

point $P_K(z)$ of z onto the closed convex set K is required to be accurate in order to achieve desired convergence. In practice, it is hard to obtain exact computation but the projection needs to be computed approximately. We introduce the following definition.

Definition 3.1: (Approximate Projection) Suppose K is a closed convex set in $\Re^m$ and $0 < \theta < \pi/2$. Define

$$\mathbf{C}_K(v,\theta) = v + \{z | \langle z, P_K(v) - v\rangle \geq |z||v|_K \cos\theta\}$$

$$\mathbf{H}_K^+(v) = \{z | \langle v - P_K(v), z\rangle \geq \langle v - P_K(v), P_K(v)\rangle\}.$$

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

$$\mathcal{P}_K^a(v,\theta) = \begin{cases} \mathbf{C}_K(v,\theta) \bigcap \mathbf{H}_K^+(v), & \text{if } v \notin K; \\ \{v\}, & \text{if } v \in K. \end{cases} \quad (1)$$

Remark 3.2: In biological and engineering systems, in many cases the visual field of a given agent, i.e., a bird or a robot camera, is a sector area in front of the agent [16], [17]. Therefore, in this case, compared to the exact projection, the approximate projection area $\mathcal{P}_K^a(v,\theta)$ is much easier to obtain. Basically, the approximate projection is an area which is closer to the convex set of interest in certain sense, instead of a single point.

In fact, $\mathbf{C}_K(v,\theta) - v$ is a convex cone generated by all vectors having angle with $P_K(v) - v$ less than θ and $\mathbf{H}_K^+(v)$ is a closed half-space containing point v with

$$\mathbf{H}_K(v) \triangleq \{z | \langle v - P_K(v), z\rangle = \langle v - P_K(v), P_K(v)\rangle\}$$

a supporting hyperplane to K at $P_K(v)$ with normal direction $v - P_K(v)$. Set $\mathcal{P}_K^a(v,\theta)$ is illustrated in Fig. 1.

Remark 3.3: Since the exact projection may be hard to obtain in practice, an approximate projection is used for its estimation. As defined, when a point is not in the closed convex set, its approximate projection onto the closed convex set is based on a convex projection region containing an infinite number of points.

Definition 3.2: The supporting approximate projection $\mathcal{P}_K^{sa}(v,\theta)$ of point v onto K with approximate angle θ is defined as

$$\mathcal{P}_K^{sa}(v,\theta) = \begin{cases} \mathbf{C}_K(v,\theta) \bigcap \mathbf{H}_K(v), & \text{if } v \notin K; \\ \{v\}, & \text{if } v \in K. \end{cases}$$

According to Definition 3.2, for any $y \in \mathcal{P}_K^a(v,\theta)$, we can associate y with $\hat{y} \in \mathcal{P}_K^{sa}(v,\theta)$ such that

$$y = (1-\beta)v + \beta\hat{y} \text{ for some } 0 \leq \beta \leq 1. \quad (2)$$

Moreover, it is easy to see that if $y \neq v$, $\hat{y}$ satisfying (2) is unique.

B. Distributed Iterative Algorithm

To solve the intersection computation problem, we propose the following approximate projected consensus algorithm:

$$x_i(k+1) = \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k) P_j^a(k),\ i = 1,\ldots,n, \quad (3)$$

where $P_i^a(k) \in \mathcal{P}_{X_i}^a(x_i(k),\theta_{i,k})$ for all i and k, and $\theta_{i,k}$ is a given accuracy parameter for the angle error away from the projection direction.

According to the definition of supporting approximate projection, there exist $0 \leq \alpha_{i,k} \leq 1$ and $P_i^{sa}(k) \in \mathcal{P}_{X_i}^{sa}(x_i(k),\theta_{i,k})$ such that

$$P_i^a(k) = (1-\alpha_{i,k})x_i(k) + \alpha_{i,k}P_i^{sa}(k). \quad (4)$$

Combining (3) and (4), we have

$$x_i(k+1) = \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\left((1-\alpha_{j,k})x_j(k) + \alpha_{j,k}P_j^{sa}(k)\right) \quad (5)$$

where, if $x_i(k) \notin X_i$, $P_i^{sa}(k) \in \mathbf{H}_{X_i}(x_i(k))$, and $\angle(P_i^{sa}(k) - x_i(k), P_{X_i}(x_i(k)) - x_i(k)) \leq \theta_{i,k}$. We illustrate the one-step iteration process of algorithm (3) in Fig. 2.

Remark 3.4: Clearly, the "projected consensus algorithm" presented in [30] is a special case of the approximate projected consensus algorithm discussed here when $\alpha_{i,k} \equiv 1$ and $\theta_{i,k} \equiv 0$. In fact, algorithm (3) can also be viewed as a disturbed version of the projected consensus algorithm in [30] with disturbance (projection error) $\sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)(P_j(k) - P_{X_j}^a(x_j(k)))$. Different from the conventional disturbance analysis, here the projection error is state-dependent, which brings new challenge to the convergence analysis.

Denote $X_0 = \bigcap_{i=1}^n X_i$. It is time to introduce our problem.

Definition 3.3: A global optimal consensus is achieved for the approximate projected consensus algorithm if, for any initial condition $x_i(0) \in \Re^m, i = 1, \ldots, n$, there exists $x^* \in X_0$ such that

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

Since a global optimal consensus x^* necessarily belongs to the intersection set X_0, a possible algorithm to find a point in X_0 is to employ the approximate projected consensus algorithm (3). We discuss the convergence of this algorithm in the next section.

C. Assumptions

Here we list all the assumptions about the convexity, the arc weights [29], [30], the connectivity [30], [35], the approximate angle and the boundedness of convex sets.

A1 *(Convexity)* $X_i, i = 1, \ldots, n$ are closed convex sets.

A2 *(Weights Rule)* (i) $\sum_{j\in\mathcal{N}_i(k)} a_{ij}(k) = 1$ for all i and k; (ii) There exists a constant $0 < \eta < 1$ such that $a_{ij}(k) \geq \eta$ for all i, k and $j \in \mathcal{N}_i(k)$.

A3 *(Connectivity)* The communication graph is uniformly jointly strongly connected $(UJSC)$, i.e., there exists a positive integer T such that $\mathcal{G}([k, k+T))$ is strongly connected for $k \geq 0$, where $\mathcal{G}([k, k+T))$ denotes the union graph with node set $\mathcal{V}$ and arc set $\bigcup_{k\leq s<k+T} \mathcal{E}(s)$.

A4 *(Approximate Angle)* $0 \leq \theta_{i,k} \leq \theta^* < \pi/2$ for all i and k.

A5 *(Bounded Sets)* $X_i, i = 1, \ldots, n$ are bounded sets.

Remark 3.5: Assumption **A4** is reasonable since we usually have a basically correct direction though we do not know what is the exact direction, especially for a bounded closed convex set. In fact, the assumption that the approximate angles are uniformly less than $\pi/2$ is equivalent to $\limsup_{k\to\infty} \theta_{i,k} < \pi/2$ for each i and hence, precludes the case there exist i_0 and time sequence $\{k_r\}_{r=0}^\infty$ such that $\limsup_{r\to\infty} \theta_{i_0,k_r} = \pi/2$.

IV. Main Results

In this section, we first present the main convergence results of the approximate projected consensus algorithm and then consider the convergence rate problem, finally discuss the critical angle error of the approximate projection. All proofs are provided in the next section.

A. Convergence Results

Denote $\alpha_k^- = \min_{1\leq i\leq n} \alpha_{i,k}$, $\alpha_k^+ = \max_{1\leq i\leq n} \alpha_{i,k}$, and $\theta_k^+ = \max_{1\leq i\leq n} \theta_{i,k}, k \geq 0$.

Theorem 4.1: Suppose **A1–A4** hold. Global optimal consensus is achieved for the approximate projected consensus algorithm if $\sum_{k=0}^\infty \alpha_k^- = \infty$ and $\sum_{k=0}^\infty \alpha_k^+ \tan\theta_k^+ < \infty$.

Remark 4.1: Theorem 4.1 gives robustness conditions on the projection error to guarantee a global optimal consensus. In fact, the projection accuracy and approximate angle conditions can be satisfied easily in various situations, for example, $\alpha_{i,k} = O(1/k^\gamma)$ and $\theta_{i,k} = O(1/k^\gamma)$ for each i and $1/2 < \gamma \leq 1$.

Remark 4.2: Generally, $\sum_{k=0}^\infty \alpha_k^- = \infty$ is somewhat fundamental for the optimal consensus convergence (see Remark 4.5), while $\sum_{k=0}^\infty \alpha_k^+ \tan\theta_k^+ < \infty$ is key for guaranteeing the boundedness of system states and then that of the disturbance term caused by the approximate projection (see Lemma 5.4 and Remark 3.4). In most of subgradient-based algorithms with subgradient corrupted by disturbance or stochastic noise, usually the noise is assumed to be bounded or have bounded variance [32], [38]. Under this setting, except the un-summable stepsize condition, generally the square summable stepsize condition is required to guarantee the optimal convergence. In fact, the optimal consensus can also be guaranteed for our approximate algorithm under the conditions $\sum_{k=0}^\infty \alpha_k^- = \infty$ and $\sum_{k=0}^\infty (\alpha_k^+)^2 < \infty$ provided that the system states are bounded.

Remark 4.3: Compared to the convergence results given in [30], Theorem 4.1 does not require the doubly stochastic assumption on the arc weights $a_{ij}(k)$ $(\sum_{j=1}^n a_{ij}(k) = \sum_{j=1}^n a_{ji}(k) = 1$ for all $i, k)$. This is important because the double stochasticity is hard to guarantee for the arc weights in a distributed way, especially when the communication between agents is directed.

Moreover, the connectivity assumption in [30] requires that $\mathcal{G}([k, k+T))$ is a fixed graph for sufficiently large k, which is more restrictive than our UJSC assumption. However, the assumption in [30] can be relaxed to UJSC graphs, as indicated by the authors.

Remark 4.4: If the arc weights are doubly stochastic, i.e., $\sum_{j=1}^n a_{ij}(k) = \sum_{j=1}^n a_{ji}(k) = 1$ for all i, k, the convergence analysis of Theorem 4.1 for the optimal consensus can be largely simplified. The simplified proof is provided in the next section.

Moreover, clearly Theorem 4.1 relaxes the doubly stochastic assumption as the general stochastic assumption. In our algorithm all agents make the (approximate) projection in order to get close to their own sets, while the weighted average mechanism guarantees that all agents converge to their intersection set. Intuitively, the UJSC connectivity (without double stochasticity) is enough to achieve the desired convergence since all agents can get directly or indirectly the information about all other sets over all time intervals with a certain length.

To investigate the necessity of the divergence condition in Theorem 4.1, we present the following result under the boundedness assumption, which shows a necessary projection accuracy condition.

Theorem 4.2: Suppose **A1–A5** hold, $\theta_{i,k} \equiv 0$ and $\alpha_k^+ < 1$ for all k. Then

i) Global optimal consensus is achieved for the approximate projected consensus algorithm only if $\sum_{k=0}^\infty \alpha_k^+ = \infty$;
ii) If $\sum_{k=0}^\infty \alpha_k^+ < \infty$, then, for initial condition $x_i(0) = z^*$, $i = 1, \ldots, n$ with $|z^*|_{X_0} > \sum_{k=0}^\infty \alpha_k^+ d^* / \prod_{k=0}^\infty (1-\alpha_k^+)$, there exists $y^* = y^*(z^*) \notin X_0$ such that $\lim_{k\to\infty} x_i(k) = y^*, \forall i$, where $d^* = \sup_{\omega_1,\omega_2\in\bigcup_{i=1}^n X_i} |\omega_1 - \omega_2|$.

Remark 4.5: Suppose **A1–A5** hold, $\theta_{i,k} \equiv 0$ and there exists a sequence $\{\alpha_k\}_{k=0}^\infty$ with $\alpha_k < 1$ for $k \geq 0$ such that $\alpha_{i,k} = \alpha_{j,k} = \alpha_k$ for each i, j and k. Then from Theorems 4.1 and 4.2, the global optimal consensus is achieved for the approximate projected consensus algorithm if and only if $\sum_{k=0}^\infty \alpha_k = \infty$.

B. Discussions: Convergence Rate

The preceding two theorems provide the sufficient conditions guaranteeing the optimal consensus. In this subsection, we discuss the convergence rate problem. The authors in [29] provide convergence bound for distributed sum objective function optimization problem in term of the objective function iteration value and the optimal value for the doubly stochastic graphs and constant stepsizes, while [38] provides various sharp convergence bounds as a function of the network size and topology for the distributed dual averaging algorithm developed by the authors.

Under the exact projection environment, that is, when $\alpha_{i,k}\equiv 1$ and $\theta_{i,k}\equiv 0$, the authors in [30] show that if the communication graph is completely connected with uniform weights and the nonempty intersection set has an interior point, then the network achieves an optimal consensus with exponential convergence rate. However, the convergence rate estimate problem for general directed graphs is still open. The following example illustrates that the interior assumption is a basic assumption guaranteeing the exponential convergence rate.

Example 4.1: Consider a network consisting of two agents 1, 2 in $\Re^2$. Their respective convex sets are the unit balls with centers $(-1,0)^T$ and $(1,0)^T$ (and hence, $X_0=\{(0,0)^T\}$). The communication graph is fixed as the complete graph on the two nodes. Suppose $a_{ij}=1/2, i,j=1,2$. Here we assume $\alpha_{i,k}\equiv 1$ and $\theta_{i,k}\equiv 0$. Let the initial condition be $x_1(0)=x_2(0)=(0,b)^T, b>0$ (see Fig. 3).

We find from Theorem 4.1 that the two agents converge to unique optimal point $(0,0)^T$. We next claim that the convergence rate is impossible to be exponential. Let $x_1(k)=(x_{11}(k),x_{12}(k))^T$ and $x_2(k)=(x_{21}(k),x_{22}(k))^T$. It is not hard to see that $x_{11}(k)=x_{21}(k)=0, \forall k$ and $x_{12}(k+1)=x_{22}(k+1)=x_{12}(k)/\sqrt{1+(x_{12}(k))^2}, \forall k$. Therefore, noticing that $\lim_{k\to\infty}x_{12}(k)=0$, we have

$$\lim_{k\to\infty}\frac{|x_1(k+1)|}{|x_1(k)|}=\lim_{k\to\infty}\frac{1}{\sqrt{1+(x_{12}(k))^2}}=1,$$

which implies that there is no $B>0, 0<\gamma<1$ such that $|x_1(k)|=|x_2(k)|\le B\gamma^k$ holds, that is, the exponential convergence is impossible and then the claim follows.

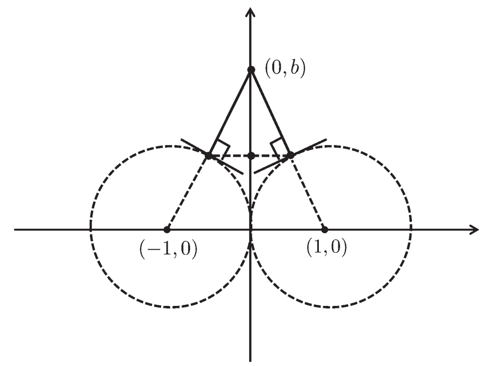


Fig. 3. The interior assumption is basic for the exponential convergence rate.

Next, we investigate the non-conservativeness of $\pi/4$ in Proposition 4.1. We focus on a special case with only one node in the network. Its set is denoted as X_*. Denote the states of the node as $\{x_*(k)\}_{k=0}^{\infty}$ driven by the approximate projected consensus algorithm:

$$x_*(k+1)\in\mathcal{P}^{sa}_{X_*}(x_*(k),\theta), \tag{6}$$

where $\mathcal{P}_{X_*^{sa}}(x_*(k),\theta)=\mathbf{C}_{X_*}(x_*(k),\theta)\bigcap\mathbf{H}_{X_*}(x_*(k))$.

The following conclusion holds.

Proposition 4.2: Suppose $\theta=\pi/4$. Then we have

i) For any bounded closed convex set X_* and initial condition $x_*(0)\in\Re^m$, we have
$$\limsup_{k\to\infty}|x_*(k)|_{X_*}\le|x_*(0)|_{X_*};$$
ii) There exists an approximate projection sequence $\{P^{sa}_*(k)\}_{k=0}^{\infty}$ with $P^{sa}_*(k)\in\mathcal{P}^{sa}_{X_*}(x_*(k),\pi/4)$ such that
  (ii.1) $\limsup_{k\to\infty}|x_*(k)|_{X_*}=0$ when X_* is a ball with radius $r>0$;
  (ii.2) $\limsup_{k\to\infty}|x_*(k)|_{X_*}=|x_*(0)|_{X_*}$ when X_* is a single point.

We present another result when $\theta>\pi/4$ in order to reveal that, in this case, the node states will be unbounded as long as the distance between the initial condition and X_* is larger than a certain threshold.

Proposition 4.3: Suppose $\theta>\pi/4$. Then for any bounded closed convex set X_*, there exists an approximate projection sequence $\{P^{sa}_*(k)\}_{k=0}^{\infty}$ such that

$$\limsup_{k\to\infty}|x_*(k)|_{X_*}=\infty$$

for all initial conditions satisfying

$$|x_*(0)|_{X_*}>\sup_{\omega_1,\omega_2\in X_*}|\omega_1-\omega_2|/(\tan\theta-1).$$

Combining Propositions 4.1, 4.2, and 4.3, we see that $\pi/4$ is a critical value of the angle error in the approximate projection for maintaining bounded states. If $\theta<\pi/4$, the system trajectories are uniformly bounded; if $\theta>\pi/4$, the trajectories diverge for a special case with one single node and particular approximate projection points; if $\theta=\pi/4$, the trajectories of the algorithm with one node are bounded (no longer uniformly with respect to initial conditions) and the property of the trajectories highly depend on the shape of the convex set.

C. Critical Angle Error

The boundedness of system states plays a key role for optimization methods [29], [31]. In this section, we consider the effect of the angle error $\theta_{i,k}$ on the boundedness of the states $\{x_i(k), i\in\mathcal{V}\}_{k=0}^{\infty}$ caused by the approximate projected consensus algorithm.

Suppose $\alpha_{i,k}\equiv 1$ and $\theta_{i,k}\equiv\theta$ with $0<\theta<\pi/2$. First, we give the following conclusion to show that when $\theta<\pi/4$, the trajectories of the algorithm are uniformly bounded with respect to all initial conditions.

Proposition 4.1: Suppose **A1**, **A2**, **A5** hold and $0<\theta<\pi/4$. Then we have

$$\sup_{x(0)}\limsup_{k\to\infty}|x_i(k)|_{X_0}<\infty$$

for $i=1,\dots,n$.

V. PROOFS

In this section, we present all the proofs of the various statements. Some auxiliary lemmas will be provided first, with detailed proofs following for each result.

A. Supporting Lemmas

In this subsection, we first establish several useful lemmas. Let $\{x_i(k)\}_{k=0}^{\infty}$ be the states of node i generated by algorithm (3). Denote $|x(k)|_{X_0} = (|x_1(k)|_{X_0} \cdots |x_n(k)|_{X_0})^T$, $y(k) = (y_1(k) \cdots y_n(k))^T$ with

$$y_i(k) = |x_i(k)|_{X_0} - \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2}.$$

Denote $A(k) = [a_{ij}(k)]$ and $D_k = \mathrm{diag}\{\alpha_{1,k} \cdots \alpha_{n,k}\}$. The following lemma holds.

Lemma 5.1: Suppose **A1**, **A2**, and **A4** hold. Then

$$|x(k+1)|_{X_0} \leq A(k)\,|x(k)|_{X_0} - A(k)D_k y(k) + \tan\theta_k^+ A(k) D_k\, |x(k)|_{X_0}. \quad (7)$$

Proof: According to Lemma 2.3, (5) implies

$$|x_i(k+1)|_{X_0} \leq \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\Big((1-\alpha_{j,k})\,|x_j(k)|_{X_0} + \alpha_{j,k}\left|P_j^{sa}(k)\right|_{X_0}\Big). \quad (8)$$

By Lemma 2.1 (ii), we have

$$\left|P_j^{sa}(k)\right|_{X_0} \leq \left|P_j^{sa}(k) - P_{X_j}(x_j(k))\right| + \left|P_{X_j}(x_j(k))\right|_{X_0}. \quad (9)$$

The definition of $P_j^{sa}(k)$ ensures that

$$\left|P_j^{sa}(k) - P_{X_j}(x_j(k))\right| \leq \tan\theta_{j,k}\, |x_j(k)|_{X_j}. \quad (10)$$

Moreover, it follows from Lemma 2.2 that for any $j \in \mathcal{V}$

$$\left|P_{X_j}(x_j(k))\right|_{X_0} \leq \sqrt{|x_j(k)|_{X_0}^2 - |x_j(k)|_{X_j}^2}. \quad (11)$$

It follows from inequalities (8), (9), (10), (11) and the relation $|x_j(k)|_{X_j} \leq |x_j(k)|_{X_0}$ that

$$\begin{aligned} |x_i(k+1)|_{X_0} &\leq \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\Big((1-\alpha_{j,k})\,|x_j(k)|_{X_0} + \alpha_{j,k}\sqrt{|x_j(k)|_{X_0}^2 - |x_j(k)|_{X_j}^2}\Big) \\ &\quad + \tan\theta_k^+ \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\alpha_{j,k}\,|x_j(k)|_{X_0} \quad (12) \\ &= \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\Big(|x_j(k)|_{X_0} - \alpha_{j,k}\Big(|x_j(k)|_{X_0} - \sqrt{|x_j(k)|_{X_0}^2 - |x_j(k)|_{X_j}^2}\Big)\Big) \\ &\quad + \tan\theta_k^+ \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\alpha_{j,k}\,|x_j(k)|_{X_0}. \quad (13) \end{aligned}$$

Then the conclusion follows. □

Lemma 5.2: Suppose **A1**, **A2**, **A4** hold, $\sum_{k=0}^{\infty}\alpha_k^+ \tan\theta_k^+ < \infty$. Then $\{x_i(k)\}_{k=0}^{\infty}$ is bounded for each i.

Proof: By taking $K = X_j$ and $z \in X_0 \subseteq X_j$, Lemma 2.1 (iv) leads to

$$\left|P_{X_j}(x_j(k)) - z\right| \leq \sqrt{|x_j(k) - z|^2 - |x_j(k)|_{X_j}^2}. \quad (14)$$

By considering $|x_i(k+1) - z|$ instead of $|x_i(k+1)|_{X_0}$, following similar procedures with (8), (9), (10), and substituting (11) with (14), we can show that, for any $i \in \mathcal{V}$,

$$\begin{aligned} |x_i(k+1) - z| &\leq \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\,|x_j(k) - z| - \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k) \times \alpha_{j,k}\Big(|x_j(k) - z| - \sqrt{|x_j(k) - z|^2 - |x_j(k)|_{X_j}^2}\Big) \\ &\quad + \tan\theta_k^+ \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\alpha_{j,k}\,|x_j(k) - z|. \quad (15) \end{aligned}$$

By dropping the non-positive term of the right-hand side in (15), and then based on **A2** (i) we have

$$\max_{1\leq i\leq n} |x_i(k+1) - z| \leq (1 + \alpha_k^+ \tan\theta_k^+) \max_{1\leq i\leq n} |x_i(k) - z|. \quad (16)$$

Therefore,

$$\begin{aligned} \max_{1\leq i\leq n} |x_i(k+1) - z| &\leq \prod_{l=0}^{k}\left(1 + \alpha_l^+ \tan\theta_l^+\right) \max_{1\leq i\leq n} |x_i(0) - z| \\ &\leq e^{\sum_{l=0}^{k}\alpha_l^+ \tan\theta_l^+} \max_{1\leq i\leq n} |x_i(0) - z| \\ &\leq e^{\sum_{l=0}^{\infty}\alpha_l^+ \tan\theta_l^+} \max_{1\leq i\leq n} |x_i(0) - z|, \quad (17) \end{aligned}$$

where the second inequality follows from $1 + b \leq e^b$ for $b \geq 0$. Then the conclusion follows. □

The next lemma is a modified version of Lemma 11 in [7] (page 50).

Lemma 5.3: Let $\{a_k\}_{k=0}^{\infty}$ and $\{b_k\}_{k=0}^{\infty}$ be non-negative sequences with $\sum_{k=0}^{\infty} b_k < \infty$. Suppose $a_{k+1} \leq a_k + b_k$ for all k. Then $\lim_{k\to\infty} a_k$ is a finite number.

The following result is about the existence of a limit.

Lemma 5.4: Suppose **A1**, **A2**, **A4** hold, $\sum_{k=0}^{\infty}\alpha_k^+ \tan\theta_k^+ < \infty$. Then the following limit exists:

$$\lim_{k\to\infty} \max_{1\leq i\leq n} |x_i(k)|_{X_0} := \vartheta.$$

Proof: Take $z \in X_0$. Based on (13), (17), and $|x_j(k)|_{X_0} \leq |x_j(k) - z|$, we have

$$\max_{1\leq i\leq n} |x_i(k+1)|_{X_0} \leq \max_{1\leq i\leq n} |x_i(k)|_{X_0} + \alpha_k^+ \tan\theta_k^+ e^{\sum_{l=0}^{\infty}\alpha_l^+ \tan\theta_l^+} \max_{1\leq i\leq n} |x_i(0) - z|.$$

The conclusion follows from the last inequality and Lemma 5.3. □

Denote

$$\eta_i^+ = \limsup_{k\to\infty} |x_i(k)|_{X_0}, \quad \eta_i^- = \liminf_{k\to\infty} |x_i(k)|_{X_0}, \ i \in \mathcal{V}.$$

Obviously, $0 \le \eta_i^- \le \eta_i^+ \le \vartheta$ for all i.

Lemma 5.5: Suppose **A1–A4** hold. If $\sum_{k=0}^{\infty} \alpha_k^+ \tan\theta_k^+ < \infty$ and there exists some agent $i_0 \in \mathcal{V}$ such that $\eta_{i_0}^- < \vartheta$, then $\vartheta = 0$.

Proof: Motivated by the idea of Lemma 4.3 in [35], we prove this lemma by contradiction.

Denote

$$\varrho_i = (\eta_i^- + \eta_i^+)/2, \ i \in \mathcal{V}.$$

Since $\eta_{i_0}^- < \vartheta$ and $\eta_{i_0}^+ \le \vartheta$, $\varrho_{i_0} < \vartheta$. Let $0 < \varepsilon < \vartheta - \varrho_{i_0}$. Then there exists an increasing sequence $\{k_l\}_{l=0}^{\infty}$ such that $|x_{i_0}(k_l)|_{X_0} \le \varrho_{i_0} + \varepsilon < \vartheta$ for $l \ge 0$. Moreover, there exists $K_0 = K_0(\varepsilon)$ such that $d_0 \sum_{k=K_0}^{\infty} \alpha_k^+ \tan\theta_k^+ \le \varepsilon$ and $|x_i(k)|_{X_0} \le \vartheta + \varepsilon$ for $k \ge K_0$ and each i, where

$$d_0 = \sup_{1\le i\le n, k\ge 0} |x_i(k)|_{X_0}, \tag{18}$$

which is a finite number by Lemma 5.2. Without loss of generality, we assume $k_0 \ge K_0$.

Based on inequality (13), we have

$$|x_{i_0}(k_0+1)|_{X_0} \le \sum_{j\in\mathcal{N}_{i_0}(k_0)\setminus i_0} a_{i_0 j}(k_0)\, |x_j(k_0)|_{X_0} + a_{i_0 i_0}(k_0)\, |x_{i_0}(k_0)|_{X_0} + d_0 \alpha_{k_0}^+ \tan\theta_{k_0}^+.$$

Therefore, $|x_{i_0}(k_0+1)|_{X_0} \le (1-\eta)(\vartheta+\varepsilon) + \eta(\varrho_{i_0}+\varepsilon) + d_0\alpha_{k_0}^+ \tan\theta_{k_0}^+$ and then

$$\begin{aligned}|x_{i_0}(k_0+2)|_{X_0} &\le (1-\eta)(\vartheta+\varepsilon) \\ &\quad + \eta\left[(1-\eta)(\vartheta+\varepsilon) + \eta\left(\varrho_{i_0}+\varepsilon\right) + d_0\alpha_{k_0}^+ \tan\theta_{k_0}^+\right] \\ &\quad + d_0\alpha_{k_0+1}^+ \tan\theta_{k_0+1}^+ \\ &\le (1-\eta^2)(\vartheta+\varepsilon) + \eta^2\left(\varrho_{i_0}+\varepsilon\right) + d_0 \sum_{k=k_0}^{k_0+1} \alpha_k^+ \tan\theta_k^+,\end{aligned}$$

where the second inequality follows from $0<\eta<1$. Similarly, we can show by induction that for $r \ge 1$

$$|x_{i_0}(k_0+r)|_{X_0} \le (1-\eta^r)(\vartheta+\varepsilon) + \eta^r\left(\varrho_{i_0}+\varepsilon\right) + d_0 \sum_{k=k_0}^{k_0+r-1} \alpha_k^+ \tan\theta_k^+.$$

Since the communication graph is UJSC, there exist agent $i_1 \ne i_0$ and time $k_0^1 \in [k_0, k_0+T)$ such that $(i_0, i_1) \in \mathcal{E}(k_0^1)$. As the above estimate for $|x_{i_0}(k_0+r)|_{X_0}$ with $x_{i_0}(k_0)$, by considering $|x_{i_1}(k_0^1+r)|_{X_0}$ with $|x_{i_0}(k_0^1)|_{X_0}$, we can show similarly that, for $r \ge 1$,

$$\begin{aligned}\left|x_{i_1}\left(k_0^1+r\right)\right|_{X_0} &\le (1-\eta^r)(\vartheta+\varepsilon) + \eta^r\left|x_{i_0}\left(k_0^1\right)\right|_{X_0} \\ &\quad + d_0 \sum_{k=k_0^1}^{k_0^1+r-1} \alpha_k^+ \tan\theta_k^+ \\ &\le (1-\eta^{k_0^1-k_0+r})(\vartheta+\varepsilon) + \eta^{k_0^1-k_0+r}\left(\varrho_{i_0}+\varepsilon\right) \\ &\quad + d_0 \sum_{k=k_0}^{k_0^1+r-1} \alpha_k^+ \tan\theta_k^+.\end{aligned}$$

Repeating the previous procedure on intervals $[k_0+pT, k_0+(p+1)T), 1 \le p \le n-2$, we obtain nodes $i_2, i_3, \ldots, i_{n-1}$ such that $\{i_j, 0 \le j \le n-1\} = \mathcal{V}$ and

$$\begin{aligned}\max_{1\le i\le n} |x_i(k_0+\hat{T})|_{X_0} &\le (1-\eta^{\hat{T}})(\vartheta+\varepsilon) + \eta^{\hat{T}}\left(\varrho_{i_0}+\varepsilon\right) + d_0 \sum_{k=k_0}^{\infty} \alpha_k^+ \tan\theta_k^+ \\ &\le (1-\eta^{\hat{T}})(\vartheta+\varepsilon) + \eta^{\hat{T}}\left(\varrho_{i_0}+\varepsilon\right) + \varepsilon,\end{aligned}$$

where $\hat{T} = (n-1)T$. Moreover, we can make similar analysis for $k_1, k_2, \ldots$ and obtain that for $l \ge 0$,

$$\max_{1\le i\le n} |x_i(k_l+\hat{T})|_{X_0} \le (1-\eta^{\hat{T}})(\vartheta+\varepsilon) + \eta^{\hat{T}}\left(\varrho_{i_0}+\varepsilon\right) + \varepsilon,$$

which yields a contradiction since $(1-\eta^{\hat{T}})(\vartheta+\varepsilon) + \eta^{\hat{T}}(\varrho_{i_0}+\varepsilon) + \varepsilon < \vartheta$ provided that ε is sufficiently small. □

We introduce the transition matrices

$$\Phi(k,s) = A(k)\cdots A(s+1)A(s) \text{ for all } k \text{ and } s \text{ with } k \ge s.$$

Recall that η and T were defined in **A2** and **A3**, respectively and $\hat{T} = (n-1)T$. The next lemma generalizes Lemma 2 in [29] on the lower bound of the entries of the transition matrices.

Lemma 5.6: Suppose **A2** and **A3** hold. Then $\Phi(k,s)_{ij} \ge \eta^{\hat{T}}$ for all i, j, s and $k \ge s+\hat{T}-1$.

Proof: By Lemma 2 in [29], $\Phi(s+\hat{T}-1, s)_{ij} \ge \eta^{\hat{T}}$ for all i, j and $s \ge 0$. Moreover, according to **A2** (i), $\sum_{l=1}^{n} A(k)_{il} = 1$ and then $\sum_{l=1}^{n} \Phi(k, s+\hat{T})_{il} = 1$ for all i, k and s. Thus, for all i, j and $k \ge s+\hat{T}-1$,

$$\begin{aligned}\Phi(k,s)_{ij} &= \left(\Phi\left(k, s+\hat{T}\right)\Phi\left(s+\hat{T}-1, s\right)\right)_{ij} \\ &\ge \sum_{l=1}^{n} \Phi\left(k, s+\hat{T}\right)_{il} \min_{1\le p,q\le n} \Phi\left(s+\hat{T}-1, s\right)_{pq} \\ &= \eta^{\hat{T}}.\end{aligned}$$

The conclusion follows. □

Lemma 5.7:

$$\frac{1}{n}\sum_{i=1}^{n}\sqrt{v_0^2-v_i^2}\le\sqrt{v_0^2-\left(\frac{\sum_{i=1}^{n}v_i}{n}\right)^2}$$

where $v_0\ge v_i\ge 0$ for all i.

Proof: The conclusion follows from that $f(z)=\sqrt{c^2-z^2}$ with domain $[-c,c]$ is a concave function for $c>0$. □

Consider the following consensus model with disturbance w_i,

$$z_i(k+1)=\sum_{j\in\mathcal{N}_i(k)}b_{ij}(k)z_j(k)+w_i(k),\ i=1,\dots,n,\quad (19)$$

where the weights $b_{ij}(k),i,j\in\mathcal{N}_i(k),k\ge 0$ satisfy **A2**. Consensus is said to be achieved for system (19) if for any initial conditions, $\lim_{k\to\infty}|z_i(k)-z_j(k)|=0$ for all $1\le i,j\le n$. The next lemma can be obtained from Theorem 1 in [39].

Lemma 5.8: If the graph of system (19) is UJSC with $\lim_{k\to\infty}w_i(k)=0$ for all i, then consensus is achieved for system (19).

In the following three subsections, we will present the proofs of Theorems 4.1, 4.2 and a simplified proof of Theorem 4.1 under the double stochasticity graph assumption, respectively.

B. Proof of Theorem 4.1

Rewrite (5) as

$$\begin{aligned}x_i(k+1)=&\sum_{j\in\mathcal{N}_i(k)}a_{ij}(k)x_j(k)\\&+\sum_{j\in\mathcal{N}_i(k)}a_{ij}(k)\alpha_{j,k}\big(P_{X_j}(x_j(k))-x_j(k)\big)\\&+\sum_{j\in\mathcal{N}_i(k)}a_{ij}(k)\alpha_{j,k}\big(P_j^{sa}(k)-P_{X_j}(x_j(k))\big).\quad (20)\end{aligned}$$

Based on (10), the sum of second and third terms in (20) is not larger than

$$\max_{1\le i\le n}\alpha_{i,k}\,|x_i(k)|_{X_i}+\alpha_k^+\tan\theta_k^+\max_{1\le i\le n}|x_i(k)|_{X_i}.\quad (21)$$

Recall that $\vartheta=\lim_{k\to\infty}\max_{1\le i\le n}|x_i(k)|_{X_0}$. Therefore, $\vartheta=0$ leads to $\lim_{k\to\infty}\max_{1\le i\le n}|x_i(k)|_{X_i}\le\lim_{k\to\infty}\max_{1\le i\le n}|x_i(k)|_{X_0}=0$, which implies that (21) tends to zero as $k\to\infty$. Therefore, by applying Lemma 5.8 to (20), consensus is achieved if $\vartheta=0$.

Moreover, we claim that, if $\vartheta=0$ and a consensus is achieved, all agents converge to a point in X_0. Since $\{x_i(k)\}_{k=0}^{\infty},i=1,\dots,n$ are bounded by Lemma 5.2 and the consensus is achieved, there exist $x^*\in X_0$ and a subsequence $\{k_l\}_{l=0}^{\infty}$ to make $\lim_{l\to\infty}x_i(k_l)=x^*$. Similar to (17), we have

$$\begin{aligned}\max_{1\le i\le n}|x_i(k)-x^*|&\le e^{\sum_{p=k_l}^{\infty}\alpha_p^+\tan\theta_p^+}\max_{1\le i\le n}|x_i(k_l)-x^*|\\&\le e^{\sum_{p=0}^{\infty}\alpha_p^+\tan\theta_p^+}\max_{1\le i\le n}|x_i(k_l)-x^*|\end{aligned}$$

for $k\ge k_l$, which implies $\lim_{k\to\infty}x_i(k)=x^*$ for all i.

If there exists some agent i_0 such that $\eta_{i_0}^-<\vartheta$, then $\vartheta=0$ by Lemma 5.5. Therefore, we only need to prove

$$\eta_i^+=\eta_i^-=\vartheta\text{ for all }i\quad\Longrightarrow\ \vartheta=0,$$

which shall be proven by contradiction. If $\vartheta>0$, then for any $\varepsilon>0$, there exists $K_1=K_1(\varepsilon)$ such that $|x_i(k)|_{X_0}\le\vartheta+\varepsilon$ and $d_0\alpha_k^+\tan\theta_k^+\le\varepsilon$ for $k\ge K_1$ and all i. We complete the proof in the following two steps.

(i). Suppose $\eta_i^+=\eta_i^-=\vartheta$ for all i. Consensus is achieved: $\lim_{k\to\infty}|x_i(k)-x_j(k)|=0$ for all i,j.

Denote

$$\varsigma_i=\limsup_{k\to\infty}\alpha_{i,k}\,|x_i(k)|_{X_i},\ i\in\mathcal{V}.$$

We next prove $\varsigma_i=0$ for all i by contradiction. If there exists some agent i_0 such that $\varsigma_{i_0}>0$, then there is an increasing time subsequence $\{k_l\}_{l=0}^{\infty}$ with $k_0\ge K_1$ such that $\alpha_{i_0,k_l}|x_{i_0}(k_l)|_{X_{i_0}}\ge c\varsigma_{i_0}$ for all l and some $0<c<1$. Therefore, by (12)

$$\begin{aligned}&|x_{i_0}(k_l+1)|_{X_0}\\&\le a_{i_0i_0}(k)\Big((1-\alpha_{i_0,k_l})\,|x_{i_0}(k_l)|_{X_0}\\&\qquad+\alpha_{i_0,k_l}\sqrt{|x_{i_0}(k_l)|_{X_0}^2-|x_{i_0}(k_l)|_{X_{i_0}}^2}\Big)\\&\quad+\sum_{j\in\mathcal{N}_{i_0}(k_l)\setminus i_0}a_{i_0j}(k_l)\,|x_j(k_l)|_{X_0}+d_0\alpha_{k_l}^+\tan\theta_{k_l}^+\\&\le\eta\left((1-\alpha_{i_0,k_l})\,(\vartheta+\varepsilon)+\sqrt{\alpha_{i_0,k_l}^2(\vartheta+\varepsilon)^2-c^2\varsigma_{i_0}^2}\right)\\&\quad+(1-\eta)(\vartheta+\varepsilon)+\varepsilon\\&=(1-\eta\alpha_{i_0,k_l})(\vartheta+\varepsilon)\\&\quad+\eta\sqrt{\alpha_{i_0,k_l}^2(\vartheta+\varepsilon)^2-c^2\varsigma_{i_0}^2}+\varepsilon,\qquad (22)\end{aligned}$$

where d_0 is the one given in (18), which yields a contradiction since the right-hand side of (22) is less than ϑ for a sufficiently small ε and sufficiently large l.

Consequently, $\lim_{k\to\infty}\alpha_{i,k}|x_i(k)|_{X_i}=0$ for all i. Moreover, since $\sum_{k=0}^{\infty}\alpha_k^+\tan\theta_k^+<\infty$, $\lim_{k\to\infty}\alpha_k^+\tan\theta_k^+=0$. The two preceding conclusions and the boundedness of $\{x_i(k)\}_{k=0}^{\infty}$ imply that the term in (21) tends to zero and then consensus is achieved by applying Lemma 5.8 to (20) again.

(ii). Suppose $\eta_i^+=\eta_i^-=\vartheta$ for all i. All agents converge to the nonempty intersection set X_0: $\lim_{k\to\infty}|x_i(k)|_{X_0}=0$ for all i.

Denote

$$\delta=\liminf_{k\to\infty}\sum_{i=1}^{n}|x_i(k)|_{X_i}.$$

We prove $\delta=0$ by contradiction. Hence, suppose $\delta>0$.

By (7), we obtain that, for $k \geq s$,

$$\begin{aligned}|x(k+1)|_{X_0} &\leq \Phi(k,s)|x(s)|_{X_0} - \sum_{l=s}^{k}\Phi(k,l)D_l y(l) + d_0\sum_{l=s}^{k}\alpha_l^+ \tan\theta_l^+ \mathbf{1}\\ &= \Phi(k,s)\,|x(s)|_{X_0} - \sum_{l=s}^{k-\hat{T}+1}\Phi(k,l)D_l y(l)\\ &\quad - \sum_{l=k-\hat{T}+2}^{k}\Phi(k,l)D_l y(l) + d_0\sum_{l=s}^{k}\alpha_l^+ \tan\theta_l^+ \mathbf{1}, \qquad (23)\end{aligned}$$

where $\mathbf{1}$ is the vector of all ones and $\hat{T} = (n-1)T$. Dropping the third term (nonpositive) on the right-hand side in (23) yields

$$|x(k+1)|_{X_0} \leq \Phi(k,s)\,|x(s)|_{X_0} - \sum_{l=s}^{k-\hat{T}+1}\Phi(k,l)D_l y(l) + d_0\sum_{l=s}^{k}\alpha_l^+ \tan\theta_l^+ \mathbf{1}. \qquad (24)$$

For $\bar{\varepsilon} = \delta^2/(4n^2\vartheta + 2\delta)$, there is a sufficiently large K_2 such that $\sum_{i=1}^{n}|x_i(k)|_{X_i} > \delta - \bar{\varepsilon}$ and $\vartheta - \bar{\varepsilon} \leq |x_i(k)|_{X_0} \leq \vartheta + \bar{\varepsilon}$ for $k \geq K_2$. For $k \geq K_2$, we have

$$\begin{aligned}\sum_{i=1}^{n}\sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2} &\leq \sum_{i=1}^{n}\sqrt{(\vartheta+\bar{\varepsilon})^2 - |x_i(k)|_{X_i}^2}\\ &\leq n\sqrt{(\vartheta+\bar{\varepsilon})^2 - \left(\frac{\sum_{i=1}^{n}|x_i(k)|_{X_i}}{n}\right)^2}\\ &\leq n\sqrt{(\vartheta+\bar{\varepsilon})^2 - \left(\frac{\delta-\bar{\varepsilon}}{n}\right)^2}\end{aligned}$$

where the second inequality follows from Lemma 5.7 and then

$$\begin{aligned}\sum_{i=1}^{n}\left(|x_i(k)|_{X_0} - \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2}\right) &\geq n\left(\vartheta - \bar{\varepsilon} - \sqrt{(\vartheta+\bar{\varepsilon})^2 - \left(\frac{\delta-\bar{\varepsilon}}{n}\right)^2}\right)\\ &:= \zeta > 0.\end{aligned}$$

Namely, $\sum_{i=1}^{n} y_i(l) \geq \zeta$ for $l \geq K_2$. Combining the preceding inequality with Lemma 5.6, we have that every component of $\Phi(k,l)D_l y(l)$ is not less than $\eta^{\hat{T}}\zeta\alpha_l^-$ for $K_2 \leq l \leq k - \hat{T} + 1$ and $k \geq K_2 + \hat{T} - 1$. Then by (24) with taking $s = K_2$

$$|x(k+1)|_{X_0} \leq \Phi(k,K_2)\,|x(K_2)|_{X_0} - \eta^{\hat{T}}\zeta\sum_{l=K_2}^{k-\hat{T}+1}\alpha_l^-\mathbf{1} + d_0\sum_{l=K_2}^{k}\alpha_l^+ \tan\theta_l^+ \mathbf{1}. \qquad (25)$$

Observing that $\sum_{l=K_2}^{\infty}\alpha_l^- = \infty$, $\sum_{l=K_2}^{\infty}\alpha_l^+ \tan\theta_l^+ < \infty$, and $\lim_{k\to\infty}|x(k)|_{X_0} = \vartheta\mathbf{1}$, a contradiction arises as $k \to \infty$ in (25).

Therefore, $\delta = 0$, that is, there is a subsequence $\{k_l\}_{l=0}^{\infty}$ such that $\lim_{l\to\infty}\sum_{i=1}^{n}|x_i(k_l)|_{X_i} = 0$. Since consensus is achieved by what we have proven in step (i), we have

$$\lim_{l\to\infty}\sum_{i=1}^{n}|x_i(k_l)|_{X_j} = 0 \text{ for all } j \in \mathcal{V},$$

which implies $\vartheta = \lim_{l\to\infty}\max_{1\leq i\leq n}|x_i(k_l)|_{X_0} = 0$.

This completes the proof. □

C. Proof of Theorem 4.2

Clearly, if $\theta_{i,k} \equiv 0$, the intersection set in (1) is the line segment from $x_i(k)$ to $P_{X_i}(x_i(k))$ and then $P_i^{sa}(k) = P_{X_i}(x_i(k))$. Then the approximate projected consensus algorithm can be written as

$$x_i(k+1) = \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\Big((1-\alpha_{j,k})x_j(k) + \alpha_{j,k}P_{X_j}(x_j(k))\Big), i = 1,\ldots,n. \qquad (26)$$

We complete the proof by analyzing the following two parts. (i). Here we prove that if $\sum_{k=0}^{\infty}\alpha_k^+ < \infty$, then there exist initial conditions from which all agents will not converge to X_0 and then the conclusion follows. Let $z^* \in \Re^m$. By (26), $x_i(1)$ can be rewritten as

$$\begin{aligned}x_i(1) &= \sum_{j\in\mathcal{N}_i(0)} a_{ij}(0)\left((1-\alpha_{j,0})x_j(0) + \alpha_{j,0}P_{X_j}(x_j(0))\right)\\ &= \sum_{j\in\mathcal{N}_i(0)} a_{ij}(0)(1-\alpha_{j,0})z^*\\ &\quad + \sum_{j\in\mathcal{N}_i(0)} a_{ij}(0)\alpha_{j,0}P_{X_0}(z^*) + \Delta_{i,0}\\ &= (1-\beta_{i,0})z^* + \beta_{i,0}P_{X_0}(z^*) + \Delta_{i,0},\end{aligned}$$

where $\beta_{i,0} = \sum_{j\in\mathcal{N}_i(0)} a_{ij}(0)\alpha_{j,0}$, $\Delta_{i,0} = \sum_{j\in\mathcal{N}_i(0)} a_{ij}(0)$ $(\alpha_{j,0}(P_{X_j}(z^*) - P_{X_0}(z^*)) + (1-\alpha_{j,0})(x_j(0) - z^*) + \alpha_{j,0}(P_{X_j}(x_j(0)) - P_{X_j}(z^*)))$ with $|\Delta_{i,0}| \leq \alpha_0^+ d^* + \max_{1\leq j\leq n}|x_j(0) - z^*|$ for all i.

We also have

$$\begin{aligned}x_i(2) &= \sum_{j\in\mathcal{N}_i(1)} a_{ij}(1)\left((1-\alpha_{j,1})x_j(1) + \alpha_{j,1}P_{X_j}(x_j(1))\right)\\ &= \sum_{j\in\mathcal{N}_i(1)} a_{ij}(1)(1-\alpha_{j,1})\left((1-\beta_{j,0})z^* + \beta_{j,0}P_{X_0}(z^*)\right)\\ &\quad + \sum_{j\in\mathcal{N}_i(1)} a_{ij}(1)\alpha_{j,1}P_{X_0}((1-\beta_{j,0})z^* + \beta_{j,0}P_{X_0}(z^*))\\ &\quad + \Delta_{i,1}\\ &= (1-\beta_{i,1})z^* + \beta_{i,1}P_{X_0}(z^*) + \Delta_{i,1},\end{aligned}$$

where $1-\beta_{i,1} = \sum_{j\in\mathcal{N}_i(1)} a_{ij}(1)(1-\alpha_{j,1})(1-\beta_{j,0})$, the third equality follows from Lemma 2.1 (iii) and $\Delta_{i,1} = \Delta_{i,1}^1 +$

$\Delta_{i,1}^2 + \Delta_{i,1}^3$ with

$$\Delta_{i,1}^1 = \sum_{j\in\mathcal{N}_i(1)} a_{ij}(1)(1-\alpha_{j,1})\Delta_{j,0};$$
$$\Delta_{i,1}^2 = \sum_{j\in\mathcal{N}_i(1)} a_{ij}(1)\alpha_{j,1}\left(P_{X_j}(x_j(1)) - P_{X_0}(x_j(1))\right);$$
$$\Delta_{i,1}^3 = \sum_{j\in\mathcal{N}_i(1)} a_{ij}(1)\alpha_{j,1} \times \left(P_{X_0}(x_j(1)) - P_{X_0}((1-\beta_{j,0})z^*+\beta_{j,0}P_{X_0}(z^*))\right). \tag{27}$$

We now give an estimation for the upper bound of $\Delta_{i,1}$. By Lemma 2.1 (i)

$$\begin{aligned}|P_{X_0}(x_j(1)) - P_{X_0}((1-\beta_{j,0})z^* + \beta_{j,0}P_{X_0}(z^*))| \\ \le |x_j(1) - ((1-\beta_{j,0})z^* + \beta_{j,0}P_{X_0}(z^*))| \\ = |\Delta_{j,0}|,\end{aligned}$$

which implies $|\Delta_{i,1}^1| + |\Delta_{i,1}^3| \le \max_{1\le i\le n}|\Delta_{i,0}| \le \alpha_0^+ d^* + \max_{1\le j\le n}|x_j(0) - z^*|$, and then $|\Delta_{i,1}| \le |\Delta_{i,1}^1| + |\Delta_{i,1}^3| + |\Delta_{i,1}^2| \le (\alpha_0^+ + \alpha_1^+)d^* + \max_{1\le j\le n}|x_j(0) - z^*|$.

Similarly, we can show by induction that for all i and k, $x_i(k+1)$ can be expressed as

$$x_i(k+1) = (1-\beta_{i,k})z^* + \beta_{i,k}P_{X_0}(z^*) + \Delta_{i,k}, \tag{28}$$

where $|\Delta_{i,k}| \le \sum_{l=0}^k \alpha_l^+ d^* + \max_{1\le j\le n}|x_j(0) - z^*|$ and $\{\beta_{i,k}, i\in\mathcal{V}\}_{k=0}^\infty$ satisfies

$$1-\beta_{i,k} = \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)(1-\alpha_{j,k})(1-\beta_{j,k-1}). \tag{29}$$

Based on (9), we can show by induction that

$$1-\beta_{i,k} \ge \prod_{l=0}^k \left(1-\alpha_l^+\right) \text{ for all } i \text{ and } k. \tag{30}$$

It follows from (28) and Lemma 2.1 (ii) that

$$|x_i(k+1)|_{X_0} \ge |(1-\beta_{i,k})z^* + \beta_{i,k}P_{X_0}(z^*)|_{X_0} - |\Delta_{i,k}|. \tag{31}$$

Moreover, for a convex set K and any $0\le\lambda\le1$,

$$\begin{aligned}&|(1-\lambda)x + \lambda P_K(x)|_K \\ &= |(1-\lambda)x + \lambda P_K(x) - P_K((1-\lambda)x + \lambda P_K(x))| \\ &= |(1-\lambda)x + \lambda P_K(x) - P_K(x)| \\ &= (1-\lambda)|x|_K,\end{aligned} \tag{32}$$

where the second equality follows from Lemma 2.1 (iii). Taking $\lambda = \beta_{i,k}$ in (32) gives

$$|(1-\beta_{i,k})z^* + \beta_{i,k}P_{X_0}(z^*)|_{X_0} = (1-\beta_{i,k})|z^*|_{X_0}.$$

Combining the last equality with (31) and (30), we obtain

$$|x_i(k+1)|_{X_0} \ge \prod_{l=0}^k \left(1-\alpha_l^+\right)|z^*|_{X_0} - \left(\sum_{l=0}^k \alpha_l^+ d^* + \max_{1\le j\le n}|x_j(0) - z^*|\right). \tag{33}$$

Taking the inferior limit on the two sides of (33), we have that, for all i,

$$\liminf_{k\to\infty}|x_i(k)|_{X_0} \ge \prod_{l=0}^\infty \left(1-\alpha_l^+\right)|z^*|_{X_0} - \left(\sum_{l=0}^\infty \alpha_l^+ d^* + \max_{1\le j\le n}|x_j(0) - z^*|\right), \tag{34}$$

which is positive provided that

$$|z^*|_{X_0} > \frac{\sum_{l=0}^\infty \alpha_l^+ d^* + \max_{1\le j\le n}|x_j(0) - z^*|}{\prod_{l=0}^\infty \left(1-\alpha_l^+\right)}, \tag{35}$$

where $\prod_{l=0}^\infty(1-\alpha_l^+) > 0$ since $0\le\alpha_l^+<1$ for all l and $\sum_{l=0}^\infty \alpha_l^+ < \infty$. Thus, all agents will not converge to the set X_0 for all initial conditions for which (35) has a solution z^*. Clearly, for the initial condition $x_i(0) = z^*, i = 1,\dots,n$ with $|z^*|_{X_0} > \sum_{l=0}^\infty \alpha_l^+ d^* / \prod_{l=0}^\infty(1-\alpha_l^+)$, all agents will not converge to X_0 and then the optimal consensus cannot be achieved.

(ii). We will prove that, for the initial condition $x_i(0) = z^*, i = 1,\dots,n$, if $\sum_{k=0}^\infty \alpha_k^+ < \infty$, there is $y^* = y^*(z^*) \notin X_0$ such that $\lim_{k\to\infty} x_i(k) = y^*$ for all i provided that z^* satisfies $|z^*|_{X_0} > \sum_{l=0}^\infty \alpha_l^+ d^* / \prod_{l=0}^\infty(1-\alpha_l^+)$.

Denote $d_1 = \sup_{1\le i\le n, k\ge0}|x_i(k)|_{X_i}$, which is finite by Lemma 5.2. From (26)

$$x_i(k+1) = \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)x_j(k) + \Gamma_{i,k}, k\ge0, \tag{36}$$

where $\Gamma_{i,k} = \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\alpha_{j,k}(P_{X_j}(x_j(k)) - x_j(k))$ with $|\Gamma_{i,k}| \le d_1\alpha_k^+$.

Take $i_0\in\mathcal{V}$. It follows from (36) that, for each j and k,

$$|x_j(k+1) - x_{i_0}(k)| \le \max_{1\le p,q\le n}|x_p(k) - x_q(k)| + \max_{1\le r\le n}|\Gamma_{r,k}|. \tag{37}$$

Then

$$\begin{aligned}&|x_{i_0}(k+2) - x_{i_0}(k)| \\ &= \left|\sum_{j\in\mathcal{N}_{i_0}(k+1)} a_{i_0j}(k+1)x_j(k+1) + \sum_{j\in\mathcal{N}_{i_0}(k+1)} a_{i_0j}(k+1)\Gamma_{j,k+1} - x_{i_0}(k)\right| \\ &\le \max_{1\le j\le n}|x_j(k+1) - x_{i_0}(k)| + \max_{1\le r\le n}|\Gamma_{r,k+1}| \\ &\le \max_{1\le p,q\le n}|x_p(k) - x_q(k)| + \max_{1\le r\le n}|\Gamma_{r,k}| + \max_{1\le r\le n}|\Gamma_{r,k+1}| \\ &\le \max_{1\le p,q\le n}|x_p(k) - x_q(k)| + d_1(\alpha_k^+ + \alpha_{k+1}^+).\end{aligned}$$

We can similarly show by induction that for all k and l,

$$\begin{aligned}&|x_{i_0}(k+l)-x_{i_0}(k)|\\&\quad\le \max_{1\le p,q\le n}|x_p(k)-x_q(k)|+d_1\sum_{s=k}^{k+l-1}\alpha_s^+\\&\quad\le \max_{1\le p,q\le n}|x_p(k)-x_q(k)|+d_1\sum_{s=k}^{\infty}\alpha_s^+. \qquad (38)\end{aligned}$$

Since $\lim_{k\to\infty}\alpha_k^+=0$ and $\{x_i(k)\}_{k=0}^{\infty}$ is bounded for all i, consensus is achieved for (36) by Lemma 5.8. Combining the consensus and the boundedness of $\{x_i(k)\}_{k=0}^{\infty}$, there exist a subsequence $\{k_l\}_{l=0}^{\infty}$ and y^* such that $\lim_{l\to\infty}x_i(k_l)=y^*$ for all i. Therefore, since $\sum_{k=0}^{\infty}\alpha_k^+<\infty$, for any $\varepsilon>0$, there exists $l_0=l_0(\varepsilon)$ such that $|x_i(k_l)-x_j(k_l)|\le\varepsilon/2$ and $d_1\sum_{s=k_l}^{\infty}\alpha_s^+\le\varepsilon/2$ for each i,j and $l\ge l_0$. Thus, for $l\ge l_0$ and $k\ge k_l$,

$$\begin{aligned}|x_{i_0}(k)-x_{i_0}(k_l)|&\le \max_{1\le p,q\le n}|x_p(k_l)-x_q(k_l)|+d_1\sum_{s=k_l}^{\infty}\alpha_s^+\\&\le\varepsilon,\end{aligned}$$

which implies $\lim_{k\to\infty}x_{i_0}(k)=y^*$ since ε can be arbitrarily small. By the analysis in the first part (i), $y^*\notin X_0$. The conclusion follows because i_0 is taken from $\mathcal{V}$ arbitrarily.

The proof is complete. □

D. A Simplified Proof of Theorem 4.1 With Doubly Stochastic Graphs

Suppose the arc weights are doubly stochastic, that is, $\sum_{j=1}^{n}a_{ij}(k)=\sum_{j=1}^{n}a_{ji}(k)=1$ for all i,k, and $\sum_{k=0}^{\infty}\alpha_k^-=\infty$ and $\sum_{k=0}^{\infty}\alpha_k^+\tan\theta_k^+<\infty$. By summing up the two sides in (13) over $i=1,\dots,n$, we obtain

$$\begin{aligned}\sum_{i=1}^{n}|x_i(k+1)|_{X_0}&\le\left(1+\alpha_k^+\tan\theta_k^+\right)\sum_{i=1}^{n}|x_i(k)|_{X_0}\\&\quad-\sum_{i=1}^{n}\alpha_{i,k}\left(|x_i(k)|_{X_0}-\sqrt{|x_i(k)|_{X_0}^2-|x_i(k)|_{X_i}^2}\right). \qquad (39)\end{aligned}$$

Summing the two sides in (39) over $k\ge0$ and rearranging the terms, we have

$$\begin{aligned}&\sum_{k=0}^{\infty}\alpha_k^-\sum_{i=1}^{n}\left(|x_i(k)|_{X_0}-\sqrt{|x_i(k)|_{X_0}^2-|x_i(k)|_{X_i}^2}\right)\\&\quad\le\sum_{k=0}^{\infty}\sum_{i=1}^{n}\alpha_{i,k}\left(|x_i(k)|_{X_0}-\sqrt{|x_i(k)|_{X_0}^2-|x_i(k)|_{X_i}^2}\right)\\&\quad\le\sum_{i=1}^{n}|x_i(0)|_{X_0}+\sum_{k=0}^{\infty}\alpha_k^+\tan\theta_k^+\sum_{i=1}^{n}|x_i(k)|_{X_0}\\&\quad<\infty. \qquad (40)\end{aligned}$$

The assumption $\sum_{k=0}^{\infty}\alpha_k^-=\infty$ and (40) imply that

$$\liminf_{k\to\infty}\sum_{i=1}^{n}\left(|x_i(k)|_{X_0}-\sqrt{|x_i(k)|_{X_0}^2-|x_i(k)|_{X_i}^2}\right)=0$$

and then

$$\liminf_{k\to\infty}\sum_{i=1}^{n}|x_i(k)|_{X_i}=0. \qquad (41)$$

From (40) we also have

$$\lim_{k\to\infty}\sum_{i=1}^{n}\alpha_{i,k}\left(|x_i(k)|_{X_0}-\sqrt{|x_i(k)|_{X_0}^2-|x_i(k)|_{X_i}^2}\right)=0 \qquad (42)$$

and then, for $i\in\mathcal{V}$,

$$\limsup_{k\to\infty}\alpha_{i,k}|x_i(k)|_{X_i}=0 \qquad (43)$$

since if there exist node i_0 and subsequence $\{k_l\}_{l=0}^{\infty}$ such that $\alpha_{i_0,k_l}|x_{i_0}(k_l)|_{X_{i_0}}\ge\varepsilon$ for all l and some $0<\varepsilon<d_0$, then for all l,

$$\begin{aligned}&\alpha_{i_0,k_l}\left(|x_{i_0}(k_l)|_{X_0}-\sqrt{|x_{i_0}(k_l)|_{X_0}^2-|x_i(k_l)|_{X_{i_0}}^2}\right)\\&\quad\ge\alpha_{i_0,k_l}|x_{i_0}(k_l)|_{X_0}-\sqrt{\alpha_{i_0,k_l}^2|x_{i_0}(k_l)|_{X_0}^2-\varepsilon^2}\\&\quad\ge d_0-\sqrt{d_0^2-\varepsilon^2}\\&\quad>0,\end{aligned}$$

which contradicts (42), where the second inequality follows from the function $f(b)=b-\sqrt{b^2-\varepsilon^2}$ is non-increasing on $[\varepsilon,\infty)$ and $\varepsilon\le\alpha_{i_0,k_l}|x_{i_0}(k_l)|_{X_{i_0}}\le\alpha_{i_0,k_l}|x_{i_0}(k_l)|_{X_0}\le d_0$ for all l. As shown in Theorem 4.1, (41) and (43) imply the optimal consensus. □

In the following three subsections, we will present the proofs of Propositions 4.1, 4.2, and 4.3, respectively.

E. Proof of Proposition 4.1

Recall $d^*=\sup_{\omega_1,\omega_2\in\cup_{i=1}^{n}X_i}|\omega_1-\omega_2|$, which is finite due to the boundedness of X_i, $i=1,\dots,n$. We claim that, for each i and all initial conditions $x(0)=(x_1^T(0)\cdots x_n^T(0))^T$,

$$\limsup_{k\to\infty}|x_i(k)|_{X_0}\le\frac{2d^*}{1-\tan\theta},$$

which implies the conclusion.

Based on (9), (10) and the definition of d^*, we have

$$\begin{aligned}|P_i^{sa}(k)|_{X_0}&\le\tan\theta\,|x_i(k)|_{X_i}+|P_{X_i}(x_i(k))|_{X_0}\\&\le\tan\theta\,|x_i(k)|_{X_0}+d^*. \qquad (44)\end{aligned}$$

Furthermore, in the case of $|x_i(k)|_{X_0}\ge2d^*/(1-\tan\theta)$, we obtain

$$\begin{aligned}|P_i^{sa}(k)|_{X_0}&\le\tan\theta\,|x_i(k)|_{X_0}+d^*\\&\le\frac{1+\tan\theta}{2}|x_i(k)|_{X_0}. \qquad (45)\end{aligned}$$

We next consider $h(k) := \max_{1\le i\le n} |x_i(k)|_{X_0}$ for the following two cases:

(i) $h(k) \le 2d^*/(1-\tan\theta)$. By Lemma 2.3 and (44), we have for all i,

$$\begin{aligned} |x_i(k+1)|_{X_0} &\le \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\left|P_j^{sa}(k)\right|_{X_0} \\ &= \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k)\left(\tan\theta\,|x_j(k)|_{X_0}+d^*\right) \\ &\le \frac{2d^*\tan\theta}{1-\tan\theta}+d^* \\ &\le \frac{2d^*}{1-\tan\theta}. \end{aligned} \tag{46}$$

Therefore, $h(k+1) \le 2d^*/(1-\tan\theta)$.

(ii) $h(k) > 2d^*/(1-\tan\theta)$. Define $\mathcal{V}_1(k) = \{i|\ |x_j(k)|_{X_0} \le 2d^*/(1-\tan\theta)$ for all $j\in\mathcal{N}_i(k)\}$ and $\mathcal{V}_2(k) = \mathcal{V}\setminus\mathcal{V}_1(k)$, where $\mathcal{V}_2(k)$ is nonempty.

1) If $i\in\mathcal{V}_1(k)$, similar to (46), we obtain $|x_i(k+1)|_{X_0} \le 2d^*/(1-\tan\theta)$;

2) If $i\in\mathcal{V}_2(k)$, we can define $\mathcal{N}_i^1(k) = \{j|\ j\in\mathcal{N}_i(k), |x_j(k)|_{X_0} \le 2d^*/(1-\tan\theta)\}$ and $\mathcal{N}_i^2(k) = \mathcal{N}_i(k)\setminus\mathcal{N}_i^1(k)$, where $\mathcal{N}_i^2(k)$ is nonempty.

According to Lemma 2.3,

$$\begin{aligned} |x_i(k+1)|_{X_0} \le &\sum_{j\in\mathcal{N}_i^1(k)} a_{ij}(k)\left|P_j^{sa}(k)\right|_{X_0} \\ &+\sum_{j\in\mathcal{N}_i^2(k)} a_{ij}(k)\left|P_j^{sa}(k)\right|_{X_0}. \end{aligned} \tag{47}$$

For $j\in\mathcal{N}_i^1(k)$, by (44) we have $|P_j^{sa}(k)|_{X_0} \le \tan\theta|x_j(k)|_{X_0}+d^* \le 2d^*\tan\theta/(1-\tan\theta)+d^* \le 2d^*/(1-\tan\theta)$. As a result, by (45) and (47), we have

$$\begin{aligned} |x_i(k+1)|_{X_0} \le &\frac{2d^*}{1-\tan\theta}\sum_{j\in\mathcal{N}_i^1(k)} a_{ij}(k) \\ &+\frac{1+\tan\theta}{2}\sum_{j\in\mathcal{N}_i^2(k)} a_{ij}(k)\,|x_j(k)|_{X_0}. \end{aligned} \tag{48}$$

Since $|x_j(k)|_{X_0} \ge 2d^*/(1-\tan\theta)$ for all $j\in\mathcal{N}_i^2(k)$, $\max_{j\in\mathcal{N}_i(k)}|x_j(k)|_{X_0} \ge 2d^*/(1-\tan\theta)$. Thus, it follows from (48) that

$$\begin{aligned} &|x_i(k+1)|_{X_0} \\ &\le \left(\sum_{j\in\mathcal{N}_i^1(k)} a_{ij}(k)+\frac{1+\tan\theta}{2}\right. \\ &\qquad\left.\times\sum_{j\in\mathcal{N}_i^2(k)} a_{ij}(k)\right)\max_{j\in\mathcal{N}_i(k)}|x_j(k)|_{X_0} \\ &\le \left(1-\eta+\frac{1+\tan\theta}{2}\eta\right)h(k+1) \\ &= \left(1-\frac{\eta(1-\tan\theta)}{2}\right)h(k+1), \end{aligned}$$

where $0<1-\eta(1-\tan\theta)/2<1$ and η is the lower bound of weights in **A2**.

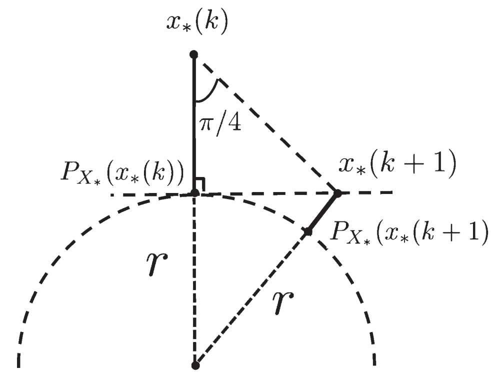


Fig. 4. Approximate projection with respect to a ball.

Therefore, based on the cases 1) and 2), we show that, if $h(k) > 2d^*/(1-\tan\theta)$, then

$$\begin{aligned} h(k+1) \le \max\{&2d^*/(1-\tan\theta)\,, \\ &(1-\eta(1-\tan\theta)/2)\,h(k)\}. \end{aligned}$$

Combining the two cases (i) and (ii), we have

$$h(k+1) \le \begin{cases} 2d^*/(1-\tan\theta), \text{ if } h(k)\le 2d^*/(1-\tan\theta); \\ \max\{2d^*/(1-\tan\theta), \\ \qquad (1-\eta(1-\tan\theta)/2)\,h(k)\}, \text{ otherwise.} \end{cases}$$

Thus, the conclusion follows. □

F. Proof of Proposition 4.2

(i) The conclusion follows from that

$$\begin{aligned} |x_*(k+1)|_{X_*} &\le |x_*(k+1)-P_{X_*}(x_*(k))| \\ &\le \tan\theta\,|x_*(k)|_{X_*} = |x_*(k)|_{X_*}. \end{aligned}$$

(ii) We select $\{P_*^{sa}(k)\}_{k=0}^\infty$ for which $x_*(k+1) = P_*^{sa}(k) \in \mathcal{P}_{X_*}^{sa}(x_*(k),\pi/4)$, $\angle(P_*^{sa}(k)-x_*(k), P_{X_*}(x_*(k))-x_*(k)) = \pi/4$ for $k\ge 0$ such that the solution of (6) with $\theta=\pi/4$ satisfies the following two cases.

(ii.1) It is easy to find that

$$|x_*(k+1)|_{X_*} = \sqrt{|x_*(k)|_{X_*}^2+r^2}-r, k\ge 0, \tag{49}$$

which implies that the sequence $\{|x_*(k)|_{X_*}\}_{k=0}^\infty$ is nonincreasing and then converges to some nonnegative number μ. As $k\to\infty$ in (49), we have $\mu = \sqrt{\mu^2+r^2}-r$ and therefore, $\mu = 0$ (see Fig. 4).

(ii.2) The conclusion is straightforward. □

G. Proof of Proposition 4.3

Select $\{P_*^{sa}(k)\}_{k=0}^\infty$, for which $x_*(k+1) = P_*^{sa}(k) \in \mathcal{P}_{X_*}^{sa}(x_*(k),\theta)$ and $\angle(P_*^{sa}(k)-x_*(k), P_{X_*}(x_*(k))-$

$x_*(k)) = \theta$ for all k. Then we have

$$\begin{aligned}|x_*(k+1)|_{X_*} &\geq |x_*(k+1) - P_{X_*}(x_*(k))| \\ &\quad - |P_{X_*}(x_*(k)) - P_{X_*}(x_*(k+1))| \\ &= \tan\theta\,|x_*(k)|_{X_*} - |P_{X_*}(x_*(k)) - P_{X_*}(x_*(k+1))| \\ &\geq \tan\theta\,|x_*(k)|_{X_*} - \sup_{\omega_1,\omega_2\in X_*} |\omega_1 - \omega_2|. \end{aligned} \tag{50}$$

At this point, we need the following conclusion: consider a nonnegative sequence $\{z_k\}_{k=0}^{\infty}$ with $z_{k+1} \geq (\tan\theta)z_k - \hat{d}$ and $\theta > \pi/4$. Then $\lim_{k\to\infty} z_k = \infty$ if $(\tan\theta - 1)z_0 - \hat{d} > 0$. Note that $z_1 - z_0 \geq (\tan\theta - 1)z_0 - \hat{d} > 0$ and then $z_2 - z_1 \geq (\tan\theta - 1)z_1 - \hat{d} \geq (\tan\theta - 1)z_0 - \hat{d} > 0$. We can show similarly by induction that $z_{k+1} - z_k \geq (\tan\theta - 1)z_0 - \hat{d}$ for all k. Thus, the conclusion follows from (50). □

VI. Numerical Examples

In this section, we first provide a numerical example comparing our approximate projected consensus algorithm with the projected consensus algorithm presented in [30] for three classes of convex sets and four classes of communication graphs from the perspective of convergence rate, and then a numerical example validating the approximate error angle results.

Example 6.1: Consider a network consisting of n agents with node set $\{1, 2, \ldots, n\}$ in $\Re^2$, their convex sets are $X_1, X_2, \ldots, X_n$. All agents have the same initial conditions $x_1(0) = x_2(0) = \cdots = x_n(0)$. The communication graph is fixed. Here we consider three classes of convex sets:

- Unit balls: the center of unit ball X_i is $(\cos(2\pi(i-1)/n), \sin(2\pi(i-1)/n))^T$, $i = 1, \ldots, n$;
- Lines: all lines pass the origin point with tangent angle $\pi(i-1)/n$, that is, $X_i = \{(z_1, z_2)^T \,|\, \sin(\pi(i-1)/n)z_1 = \cos(\pi(i-1)/n)z_2\}$, $i = 1, \ldots, n$;
- Half-spaces: all half-spaces pass the origin with normal direction $(-\sin(2\pi(i-1)/n), \cos(2\pi(i-1)/n))^T$, that is, $X_i = \{(z_1, z_2)^T \,|\, \cos(2\pi(i-1)/n)z_2 - \sin(2\pi(i-1)/n)z_1 \leq 0\}$, $i = 1, \ldots, n$,

and four classes of connected communication graphs

- Completely connected graphs (CCG): $a_{ii} = 1/n$, $i = 1, \ldots, n$;
- Chains: $a_{11} = a_{22} = 1/2$, $a_{i(i-1)} = a_{ii} = a_{i(i+1)} = 1/3$, $2 \leq i \leq n-1$, $a_{n(n-1)} = a_{nn} = 1/2$;
- Stars: $a_{1i} = 1/n, 1 \leq i \leq n$; $a_{ii} = a_{i1} = 1/2, 2 \leq i \leq n$;
- Cycles: $a_{i(i-1)} = a_{ii} = a_{i(i+1)} = 1/3, 2 \leq i \leq n-1$, $a_{1n} = a_{11} = a_{12} = a_{n(n-1)} = a_{nn} = a_{n1} = 1/3$.

Let $n = 50$. Note that for the three classes of convex sets, $X_0 = \{(0, 0)^T\}$. We consider 400 initial conditions $\{(-1.9 + 0.2p, -1.9 + 0.2q)^T | 0 \leq p \leq 19, 0 \leq q \leq 19\}$ equally spaced over the square $\{(z_1, z_2)^T \,||z_1| \leq 2, |z_2| \leq 2\}$. The following Tables I and II give the fraction of the initial conditions in the 400 initial conditions from which the approximate projected consensus algorithm (APCA) with $\alpha_{i,k} \equiv 0.5$ and $\theta_{i,k} \equiv 0$ converges faster than the projected consensus algorithm (PCA) presented in [30] at time 600 and 1000, respectively (compare their distance function $h(k) = \max_{1\leq i\leq n} |x_i(k)|_{X_0}$).

TABLE I
Fraction of the Initial Conditions From Which APCA Converges Faster Than PCA at Time 600

	Unit Balls	Lines	Half-spaces
CCG	0	0	0
Chains	0	0	0
Stars	0.9325	0	0
Cycles	0	0	0

TABLE II
Fraction of the Initial Conditions From Which APCA Converges Faster Than PCA at Time 1000

	Unit Balls	Lines	Half-spaces
CCG	0	0	0
Chains	0	0	0
Stars	1	0	0
Cycles	0	0	0

TABLE III
Average of the Needed Iteration Steps of APCA and PCA Such That $h(k) \leq 0.05$

	Unit Balls	Lines	Half-spaces
CCG	(3.415, 3)	(3.7475, 3.135)	(3.8525, 3.49)
Chains	(4288.4, 4268.8)	(3046.1, 2993.8)	(7251.7, 7220.3)
Stars	(49.535, 30.9525)	(33.45, 22.05)	(71.75, 47.72)
Cycles	(1936.6, 1922.1)	(1783.8, 1778.1)	(2727.3, 2712)

TABLE IV
Average of the Needed Iteration Steps of APCA and PCA Such That $h(k) \leq 0.01$

	Unit Balls	Lines	Half-spaces
CCG	(4, 4)	(4, 4)	(4, 4)
Chains	(6734.1, 6707.4)	(4275.6, 4205.2)	(9740, 9700.1)
Stars	(75.0725, 47.22)	(45, 29.5)	(97.775, 64.8425)
Cycles	(2831.4, 2812.1)	(2396.9, 2389.6)	(3624.6, 3604.3)

Tables III and IV present the average of needed iteration steps for the 400 initial conditions from which the APCA with $\alpha_{i,k} \equiv 0.5$ and $\theta_{i,k} \equiv 0$ and the PCA achieve the optimal convergence with errors 0.05 and 0.01, respectively.

From the first two tables, we find that when the convex sets are lines and half-spaces, our algorithm with the four classes of communication graphs converges slower for all the 400 initial conditions, while when the convex sets are unit balls, beyond 93% initial conditions, our algorithm with stars converges faster than the PCA in [30]. From the second two tables, we find that for all the three classes of convex sets and the communication graph chains and cycles, the needed iteration steps to achieve a certain convergence error for the APCA and the PCA are almost the same. Moreover, from the simulation results for unit balls and star graphs, we can infer that the APCA converges faster than PCA when agents are close to the origin. In fact, which algorithm converges faster highly depends on the shape of convex sets, communication graph, the computation time and the initial conditions.

Example 6.2: In this example, we consider algorithm (6) with X_* being the segment $\{(0, z)^T | 0 \leq z \leq 1\}$, where the initial condition is $x_*(0) = (15, 15)^T$. The convergence processes of algorithm (6) with angle errors $\tan\theta = 0.9$, $\tan\theta = 1$, $\tan\theta = 1.1$ are shown in Fig. 5. The simulation thus confirms the conclusions from Propositions 4.2, 4.3. From Fig. 5 we can find that the system states grow to infinity when the angle error is grater than $\pi/4$, while stay bounded when the angle error is less than or equal to $\pi/4$.

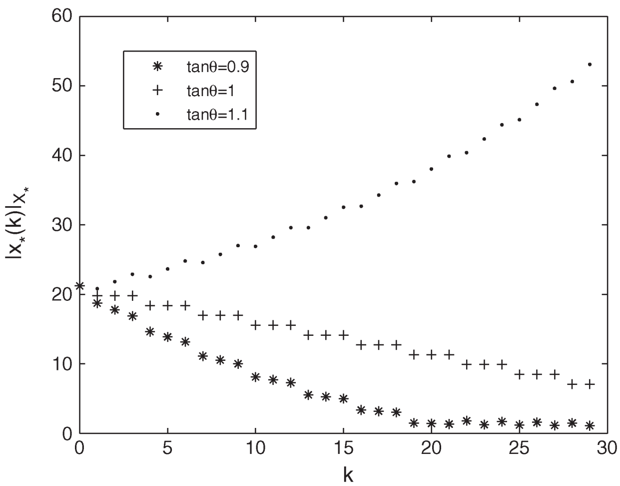


Fig. 5. $\pi/4$ is a critical angle error of maintaining the boundedness of system states.

VII. Conclusions

In this paper, we proposed an approximate projected consensus algorithm for a network to cooperatively compute the intersection of a sequence of convex sets, each of which is known only to one node. Each node computes an approximate projection located in its convex projection cone. Sufficient and/or necessary conditions were obtained for the considered algorithm on how much projection accuracy is required to ensure a global consensus within the intersection set, under the assumption that the communication graph is uniformly jointly strongly connected. We showed that $\pi/4$ is a threshold for the angle error in the projection approximation to ensure a bounded solution for the iterative projections. Examples was provided to verify the obtained results and compare the convergence rate of the approximate projected consensus algorithm and the projected consensus algorithm. Future works might include the effect of communication delay or packet drop, and the communication complexity for distributed optimization.

Acknowledgment

The authors would like to thank the Editor, the Associate Editor and the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

References

[1] Y. Lou, G. Shi, K. H. Johansson, and Y. Hong, "Reaching optimal consensus for multi-agent systems based on approximate projection," in *Proc. 10th World Congr. Intelligent Control and Automation*, Beijing, China, 2012, pp. 2794–2800.

[2] R. T. Rockafellar, *Convex Analysis*. New Jersey: Princeton University Press, 1972.

[3] F. Clarke, Y.S. Ledyaev, R. Stern, and P. Wolenski, *Nonsmooth Analysis and Control Theory*. Berlin, Germany: Speringer-Verlag, 1998.

[4] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.

[5] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Princeton, NJ: Princeton, 1989.

[6] C. Godsil and G. Royle, *Algebraic Graph Theory*. New York: Springer-Verlag, 2001.

[7] B. T. Polyak, *Introduction to Optimization*. New York: Optimization Software, Inc., 1987.

[8] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. 31, no. 9, pp. 803–812, Sep. 1986.

[9] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, Jun. 2003.

[10] S. Muthukrishnan, B. Ghosh, and M. Schultz, "First and second order diffusive methods for rapid, coarse, distributed load balancing," *Theory of Comput. Syst.*, vol. 31, no. 4, pp. 331–354, 1998.

[11] R. Diekmann, A. Frommer, and B. Monien, "Efficient schemes for nearest neighbor load balancing," *Parallel Comput.*, vol. 25, no. 7, pp. 789–812, 1999.

[12] G. Shi and Y. Hong, "Global target aggregation and state agreement of nonlinear multi-agent systems with switching topologies," *Automatica*, vol. 45, no. 5, pp. 1165–1175, May 2009.

[13] G. Shi, Y. Hong, and K. H. Johansson, "Connectivity and set tracking of multi-agent systems guided by multiple moving leaders," *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 663–676, Mar. 2012.

[14] Y. Lou and Y. Hong, "Target containment control of multi-agent systems with random switching interconnection topologies," *Automatica*, vol. 48, no. 5, pp. 879–885, 2012.

[15] M. H. DeGroot, "Reaching a consensus," *J. Amer. Statist. Assoc.*, vol. 69, no. 345, pp. 118–121, 1974.

[16] C. Reynolds, "Flocks, herds, schools: A distributed behavioral model," *Comput. Graph.*, vol. 21, no. 4, pp. 25–34, 1987.

[17] R. Arkin, *Behaivor-Based Robotics*. Cambridge, MA: MIT Press, 1998.

[18] W. Ren and R. Beard, "Consensus seeking in multi-agent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.

[19] S. Martinez, J. Cortes, and F. Bullo, "Motion coordination with distributed information," *IEEE Control Syst. Mag.*, vol. 27, no. 4, pp. 75–88, 2007.

[20] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401–420, Mar. 2006.

[21] M. Cao, A. S. Morse, and B. D. O. Anderson, "Reaching a consensus in a dynamically changing environment: A graphical approach," *SIAM J. Control Optim*, vol. 47, no. 2, pp. 575–600, 2008.

[22] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Flocking in fixed and switching networks," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 863–868, May 2007.

[23] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in *Proc. IPSN'04*, 2004, pp. 20–27.

[24] S. S. Ram, A. Nedić, and V. V. Veeravalli, "Stochastic incremental gradient descent for estimation in sensor networks," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, 2007, pp. 582–586.

[25] B. Johansson, T. Keviczky, M. Johansson, and K. H. Johansson, "Subgradient methods and consensus algorithms for solving convex optimization problems," in *Proc. IEEE Conf. Decision and Control*, Cancun, Mexico, 2008, pp. 4185–4190.

[26] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "Distributed subgradient methods and quantization effects," in *Proc. IEEE Conf. Decision and Control, Cancun, Mexico*, 2008, pp. 4177–4184.

[27] S. S. Ram, A. Nedić, and V. V. Veeravalli, "Incremental stochastic subgradient algorithms for convex optimization," *SIAM J. Optimiz.*, vol. 20, no. 2, pp. 691–717, 2009.

[28] B. Johansson, M. Rabi, and M. Johansson, "A randomized incremental subgradient method for distributed optimization in networked systems," *SIAM J. Optimiz.*, vol. 20, no. 3, pp. 1157–1170, 2009.

[29] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48–61, Jan. 2009.

[30] A. Nedić, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922–938, Apr. 2010.

[31] A. Nedić and A. Ozdaglar, "Approximate primal solutions and rate analysis for dual subgradient methods," *SIAM J. Optimiz.*, vol. 19, no. 4, pp. 1757–1780, 2008.

[32] K. Srivastava and A. Nedić, "Distributed asynchronous constrained stochastic optimization," *IEEE J. Select. Topics Signal Process.*, vol. 5, no. 4, pp. 772–790, 2011.

[33] J. Lu, C. Y. Tang, P. R. Regier, and T. D. Bow, "Gossip algorithms for convex consensus optimization over networks," *IEEE Trans. Autom. Control*, vol. 56, no. 12, pp. 2917–2923, Dec. 2011.

[34] J. Lu, P. R. Regier, and C. Y. Tang, "Control of distributed convex optimization," in *Proc. IEEE Conf. Decision and Control*, Atlanta, GA, 2010, pp. 489–495.

[35] G. Shi, K. H. Johansson, and Y. Hong, "Reaching an optimal consensus: Dynamical systems that compute intersections of convex sets," *IEEE Trans. Autom. Control*, vol. 58, no. 3, pp. 610–622, Mar. 2013.

[36] G. Shi and K. H. Johansson, "Randomized optimal consensus of multi-agent systems," *Automatica*, vol. 48, no. 12, pp. 3018–3030, 2012.

[37] D. Jakovetic, J. Xavier, and J. M. F Moura, "Cooperative convex optimization in networked systems: Augmented lagrangian algorithms with directed gossip communication," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3889–3902, 2011.

[38] J. C. Duchi, A. Agarwal, and M. J. Wainwright, "Dual averaging for distributed optimization: Convergence analysis and network scaling," *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 592–606, Mar. 2012.

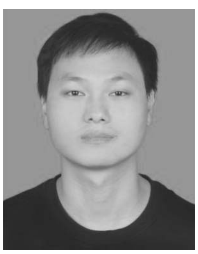
[39] L. Wang and L. Guo, "Robust consensus and soft control of multi-agent systems with noises," *J. Syst. Sci. and Complex.*, vol. 21, no. 3, pp. 406–415, 2008.

[40] N. Aronszajn, "Theory of reproducing kernels," *Trans. Amer. Math. Soc*, vol. 68, no. 3, pp. 337–404, 1950.

[41] L. Gubin, B. Polyak, and E. Raik, "The method of projections for finding the common point of convex sets," *U.S.S.R Comput. Math. Math. Phys*, vol. 7, no. 6, pp. 1211–1228, 1967.

[42] F. Deutsch, "Rate of convergence of the method of alternating projections," in *Parametric Optimization and Approximation*, B. Brosowski and F. Deutsch, Eds., 1983, vol. 76, pp. 96–107, Basel, Switzerland: Birkhäuser.

[43] H. H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Rev.*, vol. 38, no. 3, pp. 367–426, 1996.



Youcheng Lou received the B.Sc. degree in mathematics and applied mathematics from Zhengzhou University in 2008, and Ph.D. degree in complex systems and control from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, in 2013.

Currently, he is a postdoctoral researcher in the National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences. His research interests include multi-agent networks, distributed optimization, and the applications of distributed methods in operations research.

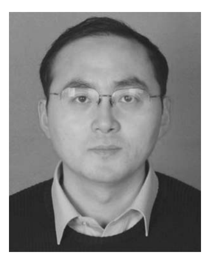
Guodong Shi received the B.Sc. degree in pure and applied mathematics from Shandong University, Jinan, China, in 2005, and the Ph.D. degree in systems theory from the Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in 2010, respectively.

Currently, he is a postdoctoral researcher at the ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden. His research interests include distributed control, computation, and optimization methods over deterministic or random graphs motivated by engineering and social networks.

Karl Henrik Johansson (SM'08–F'13) received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Lund, Sweden.

He is Director of the KTH ACCESS Linnaeus Centre and Professor at the School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden. He is a Wallenberg Scholar and has held a six-year Senior Researcher Position with the Swedish Research Council. He is Director of the Stockholm Strategic Research Area ICT The Next Generation. He has held visiting positions at the University of California at Berkeley (1998–2000) and California Institute of Technology (2006–2007). His research interests are in networked control systems, hybrid and embedded system, and applications in transportation, energy, and automation systems. He has served on the Executive Committees of several European research projects in the area of networked embedded systems.

Dr. Johansson has been a member of the IEEE Control Systems Society Board of Governors and the Chair of the IFAC Technical Committee on Networked Systems. He has been on the Editorial Boards of several journals, including *Automatica*, the IEEE Transactions on Automatic Control, and *IET Control Theory and Applications*. He is currently on the Editorial Board of IEEE Transactions on Control of Network Systems and the *European Journal of Control*. He has been Guest Editor for special issues, including the one on "Wireless Sensor and Actuator Networks" of IEEE Transactions on Automatic Control (2011). He was the General Chair of the ACM/IEEE Cyber-Physical Systems Week 2010 in Stockholm and IPC Chair of many conferences. In 2009, he received the Best Paper Award of the IEEE International Conference on Mobile Ad-hoc and Sensor Systems. In 2009, he was also awarded Wallenberg Scholar, as one of the first ten scholars from all sciences, by the Knut and Alice Wallenberg Foundation. He was awarded an Individual Grant for the Advancement of Research Leaders from the Swedish Foundation for Strategic Research in 2005. He received the triennial Young Author Prize from IFAC in 1996 and the Peccei Award from the International Institute of System Analysis, Austria, in 1993. He received Young Researcher Awards from Scania in 1996 and from Ericsson in 1998 and 1999.



Yiguang Hong (M'99–SM'02) 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 as 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, robotics, and reliability of software and communication systems.

Dr. Hong is Deputy Editor-in-Chief of *Acta Automatica Sinca*, Editor-at-Large of the *Journal of Control Theory and Applications*, and Associate Editors for several journals, including the IEEE Transactions on Automatic Control, *Journal of Systems Science and Complexity*, and *Cybernetika*. He is a recipient of the Guangzhaozhi Award at the Chinese Control Conference (1997), Young Author Prize of the IFAC World Congress (1999), Young Scientist Award of CAS (2001), and the Youth Award for Science and Technology of China (2006), and the National Natural Science Prize of China (2008), jointly with D. Cheng, *et al.*).