Multi-agent Systems Reaching Optimal Consensus Based on Simple Bernoulli Decisions

Guodong Shi[1] and Karl Henrik Johansson[1]

1. ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Stockholm 10044, Sweden.
E-mail: guodongs, kallej@kth.se

Abstract: In this paper, we formulate and solve a randomized optimal consensus problem for multi-agent systems with stochastically time-varying interconnection topology. The considered multi-agent system with a simple randomized iterating rule achieves an almost sure consensus meanwhile solving the optimization problem $\min_{z\in\mathbb{R}^d} \sum_{i=1}^n f_i(z)$, in which the optimal solution set of objective function f_i corresponding to agent i can only be observed by agent i itself. At each time step, each agent independently and randomly chooses either taking an average among its neighbor set, or projecting onto the optimal solution set of its own optimization component. Both directed and bidirectional communication graphs are studied. Connectivity conditions are proposed to guarantee an optimal consensus almost surely with proper convexity and intersection assumptions. The convergence analysis is carried out using convex analysis. The results illustrate that a group of autonomous agents can reach an optimal opinion with probability one by each node simply making a randomized trade-off between following its neighbors or sticking to its own opinion at each time step.

Key Words: Multi-agent systems, Optimal consensus, Set convergence, Distributed optimization, Randomized algorithms

1 Introduction

In recent years, there have been considerable research efforts on multi-agent dynamics in application areas such as engineering, natural science, and social science. Cooperative control of multi-agent systems is an active research topic, where collective tasks are enabled by the recent developments of distributed control protocols via interconnected communication [6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 19]. However, fundamental difficulties remain in the search of suitable tools to describe and design the dynamical behavior of these systems and thus to provide insights in their basic principles. Unlike what is often the case in classical control design, multi-agent control systems aim at fully exploiting, rather than attenuating, the interconnection between subsystems. The distributed nature of the information processing and control requires completely new approaches to analysis and synthesis.

Minimizing a sum of functions, $\sum_{i=1}^n f_i(z)$, using distributed algorithms, where each component function f_i is known only to a particular agent i, has attracted much attention in recent years, due to its wide application in multi-agent systems and resource allocation in wireless networks [29, 30, 31, 32, 33, 34]. A class of subgradient-based incremental algorithms when some estimate of the optimal solution can be passed over the network via deterministic or randomized iteration, were studied in [29, 30, 38]. Then in [33] a non-gradient-based algorithm was proposed, where each node starts at its own optimal solution and updates using a pairwise equalizing protocol. The local information transmitted over the neighborhood is usually limited to a convex combination of its neighbors [6, 7, 8]. Combing the ideas of consensus algorithms and subgradient methods has resulted in a number of significant results. A subgradient method in combination with consensus steps was given for solving coupled optimization problems with fixed undirected topology in [32]. An important contribution on multi-agent optimization is [36], in which the presented decentralized algorithm was based on simply summing an averaging (consensus) part and a subgradient part, and convergence bounds for a distributed multi-agent computation model with time-varying communication graphs with various connectivity assumptions were shown. A constrained optimization problem was studied in [37], where each agent is assumed to always lie in a particular convex set, and consensus and optimization were shown to be guaranteed together by each agent taking projection onto its own set at each step. Then a convex-projection-based distributed control was presented for multi-agent systems with continuous-time dynamics to solve this optimization problem asymptotically [35].

In this paper, we present a randomized multi-agent optimization algorithm. Different from the existing results, we focus on the randomization of individual decision-making of each node. We assume that the optimal solution set of f_i, is a convex set, and can be observed only by node i. Then at each time step, there are two options for each agent: an average (consensus) part as a convex combination of its neighbors' state, and an projection part as the convex projection of its current state onto its own optimal solution set. In the algorithm, each agent independently makes a decision via a simple Bernoulli trial, i.e., chooses the averaging part with probability p, and the projection part with probability $1-p$. Viewing the state of each agent as its "opinion", one can interpret the randomized algorithm considered in this paper as a model of spread of information in social networks [28]. In this case, the averaging part of the iteration corresponds to an agent updating its opinion based on its neighbors' information, while the projection part corresponds to an agent updating its opinion based only on its own belief of what is the best move. The authors of [28] draw interesting conclusions from a model similar to ours on how misinformation can spread in a social network.

In our model, the communication graph is assumed to be a general random digraph process independent with the

This work has been supported in part by the Knut and Alice Wallenberg Foundation, the Swedish Research Council and KTH SRA TNG.

agents' decision making process. Instead of assuming that the communication graph is modeled by a sequence of i.i.d. random variables over time, we just require the connectivity-independence condition, which is essentially different with existing works [25, 27, 26]. Borrowing the ideas on uniform joint-connection [6, 7, 22] and $[t,\infty)$-joint connectedness [8, 18], we introduce connectivity conditions of stochastically uniformly (jointly) strongly connected (SUSC) and stochastically jointly connected (SJC) graphs, respectively. The results show that the considered multi-agent network can almost surely achieve a global optimal consensus, i.e., a global consensus within the optimal solution set of $\sum_{i=1}^{n} f_i(z)$, when the communication graph is SUSC with general directed graphs, or SJC with bidirectional information exchange. Convergence is derived with the help of convex analysis and probabilistic analysis.

The paper is organized as follows. In Section 2, some preliminary concepts are introduced. In Section 3, we formulate the considered multi-agent optimization model and present the optimization algorithm. We also establish some basic assumptions and lemmas in this section. Then the main result and convergence analysis are shown for directed and bidirectional graphs, respectively in Sections 4 and 5. Finally, concluding remarks are given in Section 6.

2 Preliminaries

Here we introduce some mathematical notations and tools on graph theory [5], convex analysis [2, 3] and Bernoulli trials [4].

2.1 Directed Graphs

A directed graph (digraph) $\mathcal{G}=(\mathcal{V},\mathcal{E})$ consists of a finite set $\mathcal{V}=\{1,\dots,n\}$ of nodes and an arc set $\mathcal{E}$. An element $e=(i,j)\in\mathcal{E}$, which is an ordered pair of nodes $i,j\in\mathcal{V}$, is called an *arc* leaving from node i and entering node j. If the e_j's are pairwise distinct in an alternating sequence $v_0e_1v_1e_2v_2\dots e_nv_n$ of nodes v_i and arcs $e_i=(v_{i-1},v_i)\in\mathcal{E}$ for $i=1,2,\dots,n$, the sequence is called a (directed) *path*. A path from i to j is denoted $i\to j$. $\mathcal{G}$ is said to be *strongly connected* if it contains paths $i\to j$ and $j\to i$ for every pair of nodes i and j.

A *weighted digraph* $\mathcal{G}$ is a digraph with weights assigned for its arcs. A weighted digraph $\mathcal{G}$ is called to be *bidirectional* if for any two nodes i and j, $(i,j)\in\mathcal{E}$ if and only if $(j,i)\in\mathcal{E}$, but the weights of (i,j) and (j,i) may be different. A bidirectional digraph is strongly connected if and only if it is connected as an undirected graph (ignoring the directions of the arcs).

The *adjacency matrix*, A, of digraph $\mathcal{G}$ is the $n\times n$ matrix whose ij-entry, A_{ij}, is 1 if there is an arc from i to j, and 0 otherwise. Additionally, if $\mathcal{G}_1=(\mathcal{V},\mathcal{E}_1)$ and $\mathcal{G}_2=(\mathcal{V},\mathcal{E}_2)$ have the same node set, the union of the two digraphs is defined as $\mathcal{G}_1\cup\mathcal{G}_2=(\mathcal{V},\mathcal{E}_1\cup\mathcal{E}_2)$.

2.2 Convex Analysis

A set $K\subset\mathbb{R}^d$ $(d>0)$ is said to be *convex* if $(1-\lambda)x+\lambda y\in K$ whenever $x,y\in K$ and $0\le\lambda\le 1$. For any set $S\subset\mathbb{R}^d$, the intersection of all convex sets containing S is called the *convex hull* of S, and is denoted by $co(S)$.

Let K be a closed convex set in $\mathbb{R}^d$ and denote $|x|_K\triangleq\inf_{y\in K}|x-y|$ as the distance between $x\in\mathbb{R}^d$ and K, where $|\cdot|$ denotes the Euclidean norm. Then we can associate to any $x\in\mathbb{R}^d$ a unique element $P_K(x)\in K$ satisfying $|x-P_K(x)|=|x|_K$, where the map P_K is called the *projector* onto K with

$$\langle P_K(x)-x, P_K(x)-y\rangle\le 0,\quad \forall y\in K. \tag{1}$$

Moreover, we have the following non-expansiveness property for P_K:

$$|P_K(x)-P_K(y)|\le|x-y|,\ x,y\in\mathbb{R}^d. \tag{2}$$

A function $f:\mathbb{R}^d\to\mathbb{R}$ is said to be convex if it satisfies

$$f(\alpha v+(1-\alpha)w)\le\alpha f(v)+(1-\alpha)f(w), \tag{3}$$

for all $v,w\in\mathbb{R}^d$ and $0\le\alpha\le 1$. The following conclusion holds.

Lemma 2.1 *Let K be a convex set in $\mathbb{R}^d$. Then $|x|_K$ is a convex function.*

The next lemma can be found in [1].

Lemma 2.2 *Let K be a subset of $\mathbb{R}^d$. The convex hull $co(K)$ of K is the set of elements of the form $x=\sum_{i=1}^{d+1}\lambda_i x_i$, where $\lambda_i\ge 0, i=1,\dots,d+1$ with $\sum_{i=1}^{d+1}\lambda_i=1$ and $x_i\in K$.*

Additionally, for every two vectors $0\ne v_1,v_2\in\mathbb{R}^d$, we define their angle as $\phi(v_1,v_2)\in[0,\pi]$ with $\cos\phi=\langle v_1,v_2\rangle/|v_1|\cdot|v_2|$.

2.3 Bernoulli Trials

A sequence of independent identically distributed (i.i.d.) Bernoulli trials is a finite or infinite sequence of independent random variables $Z_1,Z_2,Z_3,\dots$, such that

(i) For each i, Z_i equals either 0 or 1;

(ii) For each i, the probability that $Z_i=1$ is a constant p_0.

p_0 is called the success probability. The next lemma shows an important property of an infinite i.i.d. Bernoulli trials which will be useful in the sequent analysis. The proof is obvious, and therefore omitted.

Lemma 2.3 *Let $Z_k, k=1,2,\dots,$ be an infinite sequence of i.i.d. Bernoulli trials with success probability $p_0>0$. Denote $\{Z_k^\omega\}_{k=0}^\infty$ as a sample sequence. Then we can select a subsequence $\{Z_{k_m}^\omega\}_{m=0}^\infty$ of $\{Z_k^\omega\}_0^\infty$ with probability 1 such that $Z_{k_m}^\omega=1$ for all m.*

3 Problem Formulation

3.1 Multi-agent Model

Consider a multi-agent system with agent set $\mathcal{V}=\{1,2,\dots,n\}$. The objective of the network is to reach a consensus, and meanwhile to cooperatively solve the following optimization problem

$$\min_{z\in\mathbb{R}^d}\ F(z)=\sum_{i=1}^{n}f_i(z) \tag{4}$$

where $f_i:\mathbb{R}^d\to\mathbb{R}$ represents the cost function of agent i, observed by agent i only, and z is a decision vector.

Time is slotted, and the dynamics of the network is in discrete time. Each agent i starts with an arbitrary initial position, denoted $x_i(0) \in \mathbb{R}^d$, and updates its state $x_i(k)$ for $k = 0, 1, 2, \dots$, based on the information received from its neighbors and the information observed from its optimization component f_i.

3.1.1 Communication Graph

We suppose the communication graph over the multi-agent network is a stochastic digraph process $\mathcal{G}_k = (\mathcal{V}, \mathcal{E}_k), k = 0, 1, \dots$. To be precise, the ij-entry $A_{ij}(k)$ of the adjacency matrix, $A(k)$ of $\mathcal{G}_k$, is a general $\{0,1\}$-state stochastic process. We use the following assumption on the independence of $\mathcal{G}_k$.

A1 *(Connectivity Independence)* Events $\mathcal{C}_k = \{\mathcal{G}_k \text{ is connected (in certain sense)}\}, k = 0, 1, \dots,$ are independent.

Remark 3.1 *Connectivity independence means that a sequence of random variables $\varpi(k)$, which are defined by that $\varpi(k) = 1$ if $\mathcal{G}_k$ is connected (in certain sense) and $\varpi(k) = 0$ otherwise, are independent. Note that, different with existing works [25, 27, 26], we do not impose the assumption that $\varpi(k), k = 0, \dots,$ are identically distributed.*

At time k, node j is said to be a *neighbor* of i if there is an arc $(j, i) \in \mathcal{E}_k$. Particularly, we assume that each node is always a neighbor of itself. Let $\mathcal{N}_i(k)$ represent the set of agent i's neighbors at time k.

Denote the joint graph of $\mathcal{G}_k$ in time interval $[k_1, k_2]$ as $\mathcal{G}([k_1, k_2]) = (\mathcal{V}, \cup_{t\in[k_1,k_2]}\mathcal{E}(t))$, where $0 \le k_1 \le k_2 \le +\infty$. Then we have the following definition.

Definition 3.1 *(i) $\mathcal{G}_k$ is said to be stochastically uniformly (jointly) strongly connected (SUSC) if there exist two constants $B \ge 1$ and $0 < q < 1$ such that for any $k \ge 0$,*

$$\mathbf{P}\{\mathcal{G}([k, k+B-1]) \text{ is strongly connected}\} \ge q.$$

(ii) Assume that $\mathcal{G}_k$ is bidirectional for all $k \ge 0$. Then $\mathcal{G}_k$ is said to be stochastically jointly connected (SJC) if there exists a sequence $0 = k_0 < \dots < k_m < \dots$ and a constant $0 < q < 1$ such that

$$\mathbf{P}\{\mathcal{G}_{[k_m, k_{m+1})} \text{ is connected}\} \ge q, \quad m = 0, \dots.$$

3.1.2 Neighboring Information

The local information that each agent uses to update its state consists of two parts: the average and the projection parts. The average part is defined as

$$e_i(k) = \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k) x_j(k),$$

where $a_{ij}(k) > 0, i, j = 1, \dots, n$ are the arc weights. The weights fulfill the following assumption:

A2 *(Arc Weights)* (i) $\sum_{j\in\mathcal{N}_i(k)} a_{ij}(k) = 1$ for all i and k.

(ii) There exists a constant $\eta > 0$ such that $\eta \le a_{ij}(k)$ for all i, j and k.

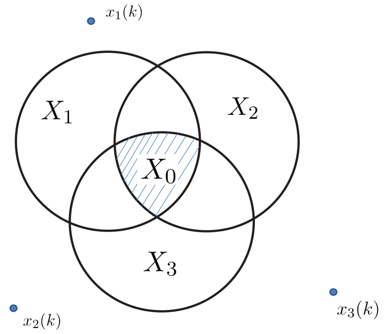

Fig. 1: The goal of the multi-agent network is to achieve a consensus in the optimal solution set X_0.

The projection part is defined as

$$g_i(k) = P_{X_i}(x_i(k)),$$

where $X_i \doteq \{v \mid f_i(v) = \min_{z\in\mathbb{R}^d} f_i(z)\}$ is the optimal solution set of each objective function $f_i, i = 1, \dots, n$. We use the following assumptions.

A3 *(Convex Solution Set)* $X_i, i = 1, \dots, n$, are closed convex sets.

A4 *(Nonempty Intersection)* $X_0 \doteq \bigcap_{i=1}^{n} X_i$ is nonempty.

In the rest of the paper, A1–A4 are our standing assumptions.

3.1.3 Randomized Algorithm

We are now ready to introduce the randomized optimization algorithm. At each time step, each agent independently and randomly either takes an average among its time-varying neighbor set, or projects onto the optimal solution set of its own objective function:

$$x_i(k+1) = \begin{cases} \sum_{j\in\mathcal{N}_i(k)} a_{ij}(k) x_j(k), & \text{with prob. } p \\ P_{X_i}(x_i(k)), & \text{with prob. } 1-p \end{cases} \tag{5}$$

where $0 < p < 1$ is a given constant.

Remark 3.2 *The constrained consensus algorithm studied in [37], can be viewed as a deterministic special case of (5), in which each node alternate between averaging and projection. Note that, we do not impose a double stochasticity assumption on the weights.*

Under assumptions A3 and A4, it is obvious that X_0 is the optimal solution set of cost function $F(z)$. Let $x^0 = (x_1^T(0), \dots, x_n^T(0))^T \in \mathbb{R}^{nd}$ be the initial condition. The considered optimal consensus problem is defined as follows. See Fig. 1 for an illustration.

Definition 3.2 *(i) A global optimal set aggregation is achieved almost surely for (5) if for all $x^0 \in \mathbb{R}^{nd}$, we have*

$$\mathbf{P}\{\lim_{k\to+\infty} |x_i(k)|_{X_0} = 0, i = 1, \dots, n\} = 1. \tag{6}$$

(ii) A global consensus is achieved almost surely for (5) if for all $x^0 \in \mathbb{R}^{nd}$, we have

$$\mathbf{P}\{\lim_{k\to+\infty} |x_i(k) - x_j(k)| = 0, i, j = 1, \dots, n\} = 1. \tag{7}$$

(iii) A global optimal consensus is achieved almost surely for (5) if both (6) and (7) hold.

3.2 Basic Properties

In this subsection, we establish two key lemmas on the algorithm (5).

Lemma 3.1 *Let K be a closed convex set in $\mathbb{R}^d$, and $K_0 \subseteq K$ be a convex subset of K. Then for any $y \in \mathbb{R}^d$, we have*

$$|P_K(y)|_{K_0}^2 + |y|_K^2 \leq |y|_{K_0}^2.$$

Proof. According to (1), we know that

$$\langle P_K(y) - y, P_K(y) - P_{K_0}(y) \rangle \leq 0.$$

Therefore, we obtain

$$\begin{aligned}&\langle P_K(y) - y, y - P_{K_0}(y) \rangle \\ &= \langle P_K(y) - y, y - P_K(y) + P_K(y) - P_{K_0}(y) \rangle \\ &\leq -|y|_K^2.\end{aligned}$$

Then,

$$\begin{aligned}|P_K(y)|_{K_0}^2 &= |P_K(y) - P_{K_0}(P_K(y))|^2 \\ &\leq |P_K(y) - P_{K_0}(y)|^2 \\ &\leq |y|_{K_0}^2 - |y|_K^2.\end{aligned}$$

The desired conclusion follows. □

Lemma 3.2 *Let $\{x(k) = (x_1^T(k), \ldots, x_n^T(k))^T\}_{k=0}^{\infty}$ be a sequence defined by (5). Then for any $k \geq 0$, we have*

$$\max_{i=1,\ldots,n} |x_i(k+1)|_{X_0} \leq \max_{i=1,\ldots,n} |x_i(k)|_{X_0}.$$

Proof. Take $l \in \mathcal{V}$. If node l follows average update rule at time k, we have

$$\begin{aligned}|x_l(k+1)|_{X_0} &= |P_{X_l}(x_l(k)) - P_{X_0}(P_{X_l}(x_l(k)))| \\ &\leq |x_l(k) - P_{X_0}(x_l(k))| \\ &\leq \max_{i=1,\ldots,n} |x_i(k)|_{X_0}. \end{aligned} \quad (8)$$

On the other hand, if node l follows projection update rule at time k, according to Lemma 2.1, we have

$$\begin{aligned}|x_l(k+1)|_{X_0} &= |\sum_{j \in N_l(k)} a_{lj}(k) x_j(k)|_{X_0} \\ &\leq \sum_{j \in N_l(k)} a_{lj}(k) |x_j(k)|_{X_0} \\ &\leq \max_{i=1,\ldots,n} |x_i(k)|_{X_0}. \end{aligned} \quad (9)$$

Hence, the conclusion holds. □

Based on Lemma 3.2, we know that the following limit exists:

$$\xi \doteq \lim_{k \to \infty} \max_{i=1,\ldots,n} |x_i(k)|_{X_0}.$$

It is immediate that the global optimal set aggregation is achieved almost surely if and only if $\mathbf{P}\{\xi = 0\} = 1$.

4 Main Results

Algorithm (5) is nonlinear and stochastic, and therefore quite challenging to analyze. In this section, we introduce the main results and convergence analysis. Due to space limitations, all the proofs which are skipped can be found in [39] for the rest of the paper.

4.1 Directed Graphs

The main result under general directed communications is stated as follows.

Theorem 4.1 *System (5) achieves a global optimal consensus almost surely if $\mathcal{G}_k$ is SUSC.*

Define

$$\delta_i \doteq \limsup_{k \to \infty} |x_i(k)|_{X_i}, \quad i = 1, \ldots, n.$$

Let $\mathcal{A} = \{\xi > 0\}$ and $\mathcal{M} = \{\exists i_0 \ s.t. \ \delta_{i_0} > 0\}$ be two events, indicating that convergence to X_0 for all the agents fails and convergence to X_{i_0} fails for some node i_0, respectively. The next lemma shows the relation between the two events.

Lemma 4.1 $\mathbf{P}[\mathcal{A} \cap \mathcal{M}] = 0$ *if $\mathcal{G}_k$ is SUSC.*

Proof. Let $\{x^\omega(k)\}_{k=0}^{\infty}$ be a sample sequence. Take an arbitrary node $i_0 \in \mathcal{V}$. Then there exists a time sequence $k_1 < \cdots < k_m < \ldots$ with $\lim_{m \to \infty} k_m = \infty$ such that

$$|x_{i_0}^\omega(k_m)|_{X_{i_0}} \geq \frac{1}{2}\delta_{i_0}(\omega) \geq 0. \quad (10)$$

Moreover, according to Lemma 3.2, $\forall \ell = 1, 2, \ldots$, $\exists T(\ell, \omega) > 0$ such that

$$k \geq T \ \Rightarrow \ 0 \leq |x_i^\omega(k)|_{X_0} \leq \xi(\omega) + \frac{1}{\ell}, \ i = 1, \ldots, n. \quad (11)$$

For any $k_m \geq T$, node i_0 projects onto X_i with probability p. Thus, Lemma 3.1 implies

$$\mathbf{P}\{|x_{i_0}(k_m+1)|_{X_0} \leq \sqrt{(\xi + \frac{1}{\ell})^2 - \frac{1}{4}\delta_{i_0}^2}\} \geq p. \quad (12)$$

At time $k_m + 2$, either one of two cases can happen in the update.

- If node i_0 chooses the projection option at time $k_m + 2$, we have
$$|x_{i_0}(k_m+2)|_{X_0} = |x_{i_0}(k_m+1)|_{X_0} \leq \sqrt{(\xi + \frac{1}{\ell})^2 - \frac{1}{4}\delta_{i_0}^2} \quad (13)$$
with probability at least p.
- If node i_0 chooses the average option at time $k_m + 2$, with (11), we can obtain from the weights rule and Lemma 2.1 that
$$\begin{aligned}&|x_{i_0}(k_m+2)|_{X_0} \\ &= |\sum_{j \in \mathcal{N}_{i_0}(k_m+1)} a_{i_0 j}(k_m+1) x_j(k_m+1)|_{X_0} \\ &\leq \eta \sqrt{(\xi + \frac{1}{\ell})^2 - \frac{1}{4}\delta_{i_0}^2} + (1-\eta)(\xi + \frac{1}{\ell}) \end{aligned} \quad (14)$$
with probability at least p.

Through similar analysis, we can also obtain that for $\tau = 1, 2, \ldots$,

$$\begin{aligned}\mathbf{P}\{|x_{i_0}(k_m+\tau)|_{X_0} &\leq \eta^{\tau-1}\sqrt{(\xi + \frac{1}{\ell})^2 - \frac{1}{4}\delta_{i_0}^2} \\ &+ (1-\eta^{\tau-1})(\xi + \frac{1}{\ell})\} \geq p. \end{aligned} \quad (15)$$

The upper analysis process can be carried out continuingly on intervals $[k_m+2B+1, k_m+3B], \dots, [k_m+(n-2)B+1, k_m+(n-1)B]$, and $i_3, \dots, i_{n-1}$ can be found until $\mathcal{V} = \{i_0, i_1, \dots, i_{n-1}\}$. Then one can obtain that for any $i \in \mathcal{V}$,

$$\mathbf{P}\{\max_{i=1,\dots,n} |x_i(k_m+(n-1)B+1)|_{X_0} \leq \eta^{(n-1)B}\sqrt{(\xi+\frac{1}{\ell})^2-\frac{1}{4}\delta_{i_0}^2}+(1-\eta^{(n-1)B})(\xi+\frac{1}{\ell})\} \geq p^n q^{n-1}. \tag{16}$$

Since (16) holds for any $k_m \geq T$ and $p^n q^{n-1}$ is a constant, and noting the fact the analysis on different time instances $\{k_m+(n-1)B+1, k_m \geq T\}$ is independent for different m, the events that

$$\max_{i=1,\dots,n} |x_i(k_m+(n-1)B+1)|_{X_0} \leq \eta^{(n-1)B}\sqrt{(\xi+\frac{1}{\ell})^2-\frac{1}{4}\delta_{i_0}^2}+(1-\eta^{(n-1)B})(\xi+\frac{1}{\ell})$$

can be viewed as an infinite sequence of i.i.d. Bernoulli trials with success probability $p^n q^{n-1}$. Then based on Lemma 2.3, we see that with probability 1, there is an infinite subsequence $\{\tilde{k}_j, j=1,2,\dots\}$ from $\{k_m+(n-1)B+1, k_m \geq T\}$ satisfying

$$\max_{i=1,\dots,n} |x_i(\tilde{k}_j)|_{X_0} \leq \eta^{(n-1)B}\sqrt{(\xi+\frac{1}{\ell})^2-\frac{1}{4}\delta_{i_0}^2}+(1-\eta^{(n-1)B})(\xi+\frac{1}{\ell}).$$

This implies

$$\mathbf{P}[\mathcal{R}_*] = 1, \tag{17}$$

where $\mathcal{R}_* = \lim_{\ell\to\infty}\mathcal{R}_\ell = \{\xi \leq \eta^{(n-1)B}\sqrt{\xi^2-\frac{1}{4}\delta_{i_0}^2}+(1-\eta^{(n-1)B})\xi\}$.

Finally, it is not hard to find that $\mathcal{A}\cap\mathcal{M} \subseteq \mathcal{R}_*^c$ because $0 < \eta^{(n-1)B} < 1$. Then the conclusion holds straightforwardly. □

Take a node $\alpha_0 \in \mathcal{V}$. Then define

$$z_{\alpha_0}(k) \doteq \max_{i=1,\dots,n} |x_i(k)|_{X_{\alpha_0}}.$$

We also need the following fact to prove the optimal set convergence.

Lemma 4.2 *We have*

$$z_{\alpha_0}(k+1) \leq z_{\alpha_0}(k) + \max_{i=1,\dots,n} |x_i(k)|_{X_i},\ k=0,1,\dots.$$

The optimal set convergence part of Theorem 4.1 can be proved in the following conclusion.

Proposition 4.1 *System (5) achieves a global optimal set aggregation almost surely if $\mathcal{G}_k$ is SUSC.*

In this subsection, we present the consensus analysis of the proof of Theorem 4.1. Let $x_{i,[\jmath]}(k)$ represent the $\jmath$'th coordinate of $x_i(k)$. Denote

$$h(k) = \min_{i=1,\dots,n} x_{i,[\jmath]}(k), \quad H(k) = \max_{i=1,\dots,n} x_{i,[\jmath]}(k).$$

The consensus proof of Theorem 4.1 will be built on the estimates of $S(k) = H(k) - h(k)$, which is summarized in the following conclusion.

Proposition 4.2 *System (5) achieves a global consensus almost surely if $\mathcal{G}_k$ is SUSC.*

4.2 Bidirectional Graphs

In this subsection, we discuss the randomized optimal consensus problem under more restrictive communication assumptions, that is, bidirectional communications.

To get the main result, we also need the following assumption.

A5 (Compactness) X_0 is compact.

Then we propose the main result on optimal consensus for the bidirectional case. It turns out that with bidirectional communications, the connectivity condition to ensure an optimal consensus is weaker.

Theorem 4.2 *Suppose $\mathcal{G}_k$ is bidirectional for all $k \geq 0$ and A5 holds. System (5) achieves a global optimal consensus almost surely if $\mathcal{G}_k$ is SJC.*

Remark 4.1 *Note that, although we assume that $\mathcal{G}_k, k \geq 0$ is bidirectional, the weight of arc (i,j) may not be equal to that of arc (j,i). In other words, we do not need the weight functions $a_{ij}(k)$ to be symmetric.*

In order to complete the proof of Theorem 4.2, we first need the following lemmas.

Lemma 4.3 *Assume that $\mathcal{G}_k$ is bidirectional for all $k \geq 0$. Then $\mathbf{P}[\mathcal{A}\cap\mathcal{M}] = 0$ if $\mathcal{G}_k$ is SJC.*

Lemma 4.4 *Define*

$$y_i = \liminf_{k\to\infty} |x_i(k)|_{X_0},\ i=1,\dots,n$$

and denote $\mathcal{D} = \{\exists i_0 \text{ s.t. } y_{i_0} < \xi\}$. Assume that $\mathcal{G}_k$ is bidirectional for all $k \geq 0$. Then $\mathbf{P}[\mathcal{A}\cap\mathcal{D}] = 0$ if $\mathcal{G}_k$ is SJC.

Then Theorem 4.2 follows from the following conclusions.

Proposition 4.3 *Assume $\mathcal{G}_k$ is bidirectional for all $k \geq 0$ and A5 holds. System (5) achieves a global optimal set aggregation almost surely if $\mathcal{G}_k$ is SJC.*

Proposition 4.4 *Assume that $\mathcal{G}_k$ is bidirectional for all $k \geq 0$ and A5 holds. System (5) achieves a global consensus almost surely if $\mathcal{G}_k$ is SJC.*

5 Conclusions

The paper investigated a randomized optimal consensus problem for multi-agent systems with stochastically time-varying interconnection topology. In this formulation, the decision process for each agent was a simple Bernoulli trial between following its neighbors or sticking to its own opinion at each time step. In terms of the optimization problem, each agent independently chose either taking an average among its time-varying neighbor set, or projecting onto the optimal solution set of its own objective function randomly with a fixed probability. Both directed and bidirectional communications were studied, and stochastically jointly connectivity conditions were proposed to guarantee an optimal consensus almost surely. The results showed that under this randomized decision making protocol, a group of autonomous agents can reach an optimal opinion with probability 1 with proper convex and nonempty intersection assumptions for the considered optimization problem.

REFERENCES

[1] J. Aubin and A. Cellina. *Differential Inclusions*. Berlin: Speringer-Verlag, 1984

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

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

[4] D. P. Bertsekas and J. N. Tsitsiklis. *Introduction to Probability*. Athena Scientific, Massachusetts, 2002.

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

[6] J. Tsitsiklis, D. Bertsekas, and M. Athans. Distributed Asynchronous Deterministic and Stochastic Gradient Optimization Algorithms. *IEEE Trans. Automatic Control*, 31, 803-812, 1986.

[7] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of Groups of Mobile Agents Using Nearest Neighbor Rules. *IEEE Trans. Automatic Control*, vol. 48, no. 6, 988-1001, 2003.

[8] L. Moreau. Stability of Multi-agent Systems with Time-dependent Communication Links. *IEEE Trans. Automatic Control*, 50, 169-182, 2005.

[9] S. Martinez, J. Cortés, and F. Bullo. Motion Coordination with Distributed Information. *IEEE Control Systems Magazine*, vol. 27, no. 4, 75-88, 2007.

[10] R. Olfati-Saber. Flocking for Multi-agent Dynamic Systems: Algorithms and Theory. *IEEE Trans. Automatic Control*, 51(3): 401-420, 2006.

[11] 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, 575-600, 2008.

[12] M. Cao, A. S. Morse and B. D. O. Anderson. Reaching a consensus in a dynamically changing environment: convergence rates, measurement delays, and asynchronous events. *SIAM J. Control Optim.*, vol. 47, no. 2, 601-623, 2008.

[13] R. Olfati-Saber and R. Murray. Consensus Problems in the Networks of Agents With Switching Topology and Time Dealys. *IEEE Trans. Automatic Control*, vol. 49, no. 9, 1520-1533, 2004.

[14] J. Fax and R. Murray. Information Flow and Cooperative Control of Vehicle Formations. *IEEE Trans. Automatic Control*, vol. 49, no. 9, 1465-1476, 2004.

[15] H. G. Tanner, A. Jadbabaie, G. J. Pappas. Flocking in fixed and switching networks. *IEEE Trans. Automatic Control*, 52(5): 863-868, 2007.

[16] F. Xiao and L. Wang. Asynchronous consensus in continuous-time multi-agent systems with switching topology and time-varying delays. *IEEE Trans. Automatic Control*, vol. 53, no. 8, 1804-1816, 2008.

[17] Y. Hong, J. Hu, and L. Gao. Tracking control for multi-agent consensus with an active leader and variable topology. *Automatica*, vol. 42, 1177-1182, 2006.

[18] G. Shi and Y. Hong. Global Target Aggregation and State Agreement of Nonlinear Multi-agent Systems with Switching Topologies. *Automatica*, vol. 45, 1165-1175, 2009.

[19] W. Ren and R. Beard. *Distributed Consensus in Multi-vehicle Cooperative Control*, Springer-Verlag, London, 2008.

[20] W. Ren and R. Beard. Consensus Seeking in Multi-agent Systems Under Dynamically Changing Interaction Topologies. *IEEE Trans. Automatic Control*, vol. 50, no. 5, 655-661.

[21] D. Cheng, J. Wang, and X. Hu. An Extension of LaSalle's Invariance Principle and Its Applciation to Multi-agents Consensus. *IEEE Trans. Automatic Control*, 53, 1765-1770, 2008.

[22] Z. Lin, B. Francis, and M. Maggiore. State Agreement for Continuous-time Coupled Nonlinear Systems. *SIAM J. Control Optim.*, vol. 46, no. 1, 288-307, 2007.

[23] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis. On distributed averaging algorithms and quantization effects. *IEEE Trans. Automatic Control*, vol. 54, no. 11, 2506-2517, 2009.

[24] S. Boyd, A. Ghosh, B. Prabhakar and D. Shah. Randomized Gossip Algorithms. *IEEE Trans. Information Theory*, vol. 52, no. 6, 2508-2530, 2006.

[25] Y. Hatano and M. Mesbahi. Agreement Over Random Networks. *IEEE Trans. on Automatic Control*, vol. 50, no. 11, pp. 18671872, 2005.

[26] A. Tahbaz-Salehi and A. Jadbabaie. A Necessary and Sufficient Condition for Consensus Over Random Networks. *IEEE Trans. on Automatic Control*, vol. 53, no. 3, 791-795, 2008.

[27] F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale net- works. *IEEE Journal on Selected Areas of Communications*, vol. 26, no.4, 634-649, 2008.

[28] D. Acemoglu, A. Ozdaglar and A. ParandehGheibi. Spread of (Mis)information in Social Networks. *Games and Economic Behavior*, vol. 70, no. 2, 194-227, 2010.

[29] M. Rabbat and R. Nowak. Distributed Optimization in Sensor Networks. *IPSN'04*, 20-27, 2004.

[30] B. Johansson, M. Rabi, and M. Johansson. A Simple Peer-to-Peer Algorithm for Distributed Optimization in Sensor Networks. *Proc. IEEE Conference on Decision and Control*, New Orleans, 4705-4710, 2007.

[31] S. S. Ram, A. Nedić, and V. V. Veeravalli. Stochastic Incremental Gradient Descent for Estimation in Sensor Networks. *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, pp. 582-586, 2007.

[32] B. Johansson, T. Keviczky, M. Johansson, and K. H. Johansson. Subgradient Methods and Consensus Algorithms for Solving Convex Optimization Problems. *Proc. IEEE Conference on Decision and Control*, Cancun, 4185-4190, 2008.

[33] J. Lu, C. Y. Tang, P. R. Regier, and T. D. Bow. A Gossip Algorithm for Convex Consensus Optimization over Networks. *Proc. American Control Conference*, Baltimore, 301-308, 2010.

[34] J. Lu, P. R. Regier, and C. Y. Tang. Control of Distributed Convex Optimization. *Proc. IEEE Conference on Decision and Control*, Atlanta, pp. 489-495, 2010.

[35] G. Shi, K. H. Johansson and Y. Hong. Multi-agent Systems Reaching Optimal Consensus with Directed Communication Graphs. *Proc. American Control Conference*, San Francisco, 5456-5461, 2011.

[36] A. Nedić and A. Ozdaglar. Distributed Subgradient Methods for Multi-agent Optimization. *IEEE Trans. on Automatic Control*, vol. 54, no. 1, 48-61, 2009.

[37] A. Nedić, A. Ozdaglar and P. A. Parrilo. Constrained Consensus and Optimization in Multi-Agent Networks. *IEEE Trans. on Automatic Control*, vol. 55, no. 4, 922-938, 2010.

[38] S. S. Ram, A. Nedić, and V. V. Veeravalli. Incremental Stochastic Subgradient Algorithms for Convex Optimization. *SIAM Journal on Optimization*, vol. 20, no. 2, 691-717, 2009.

[39] G. Shi and K. H. Johansson. Randomized Optimal Consensus of Multi-agent Systems. Available online from `arXiv:submit/0300543`.