2019

Applications of fixed point theory to distributed optimization, robust convex optimization, and stability of stochastic systems

Seyyed Shaho Alaviani
Iowa State University

Applications of fixed point theory to distributed optimization, robust convex optimization, and stability of stochastic systems

by

Seyyed Shaho Alaviani

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Electrical Engineering

Program of Study Committee:
Nicola Elia, Major Professor
Umesh Vaidya
Yongxin Chen
Domenico D'Alessandro
Soumik Sarkar

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2019

DEDICATION

To those who have moved the world.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF NOMENCLATURE

N	The set of all natural numbers
$\Re^n$	n-dimensional real space
$\Re^{m \times n}$	The set of all real m by n matrices
A^T	Transpose of the matrix A
I_n	Identity matrix of dimension n by n
$\mathcal{X}$	Topological space
$\mathcal{M}$	Metric space
$\mathcal{B}$	Real Banach space
$\mathcal{H}$	Real Hilbert space
B^n	A closed ball with Euclidean metric in $\Re^n$
ρ	A metric
$\|.\|_{\mathcal{H}}$	Norm of the real Hilbert space $\mathcal{H}$
$\|.\|_{\mathcal{B}}$	Norm of the real Banach space $\mathcal{B}$
$< .,. >$	Inner product of the real Hilbert space $\mathcal{H}$
$\mathcal{P}_C$	Projection onto the set C in the real Hilbert space $\mathcal{H}$
$\|A\|_1$	Induced 1-norm of matrix A
$\|A\|_2$	Induced 2-norm of matrix A
$\|A\|_\infty$	Induced ∞-norm of matrix A
$\mathbf{1}_n$	Vector of dimension n whose all entries are 1
$\mathbf{0}_n$	Vector of dimension n whose all entries are 0
$Re(a)$	Real part of the complex number a
$\lambda_2(A)$	Sorted in increasing order with respect to real parts, the second eigenvalue of the matrix A
$E[x]$	Expectation of the random variable x

L^1	The space of measurable functions
$\nabla f(x)$	Gradient of the function $f(x)$
$\emptyset$	The empty set
$A \succeq 0$	Positive semi-definite matrix

ACKNOWLEDGMENTS

I must thank God who has helped me to come to the U.S., selected Prof. Nicola Elia as my major professor, and helped me during my Ph.D. program as well as every second in my life. I would like to thank my major professor who has guided me to the world of optimization theory and networked systems, discussed interesting topics with me, and supported me during this research. I mention that Iowa State University and Ames, Iowa, are great places for studying and student life, respectively.

I would like to thank my parents, my uncle Habib Kharazmi, and my best friends Seyyed Ghafoor Barzanjeh, Mohammad Ali Saate', and Seyyed Taha Kamalizadeh for their emotional support.

This work was supported by National Science Foundation under Grants CCF-1320643, CNS-1239319, ECCS-1509372, and AFOSR under Grant FA95501510119.

ABSTRACT

Large-scale multi-agent networked systems are becoming more and more popular due to applications in robotics, machine learning, and signal processing. Although distributed algorithms have been proposed for efficient computations rather than centralized computations for large data optimization, existing algorithms are still suffering from some disadvantages such as distribution dependency or B-connectivity assumption of switching communication graphs. This study applies fixed point theory to analyze distributed optimization problems and to overcome existing difficulties such as distribution dependency or B-connectivity assumption of switching communication graphs. In this study, a new mathematical terminology and a new mathematical optimization problem are defined. It is shown that the optimization problem includes centralized optimization and distributed optimization problems over random networks. Centralized robust convex optimization is defined on Hilbert spaces that is included in the defined optimization problem. An algorithm using diminishing step size is proposed to solve the optimization problem under suitable assumptions. Consequently, as a special case, it results in an asynchronous algorithm for solving distributed optimization over random networks without distribution dependency or B-connectivity assumption of random communication graphs. It is shown that the random Picard iteration or the random Krasnoselskii-Mann iteration may be used for solving the feasibility problem of the defined optimization. Consequently, as special cases, they result in asynchronous algorithms for solving linear algebraic equations and average consensus over random networks without distribution dependency or B-connectivity assumption of switching communication graphs. As a generalization of the proposed algorithm for solving distributed optimization over random networks, an algorithm is proposed for solving distributed optimization with state-dependent interactions and time-varying topologies without B-connectivity assumption on communication graphs. So far these random algorithms are special cases of stochastic discrete-time systems. It is shown that difficulties such as

distribution dependency of random variable sequences that arise in using Lyapunov's and LaSalle's methods for stability analysis of stochastic nonlinear discrete-time systems may be overcome by means of fixed point theory.

CHAPTER 1. INTRODUCTION

Optimization has been a backbone of many problems in machine learning, energy efficiency, optimal control, signal processing etc. Optimization over networks has been a hot topic due to its applications in real life problems such as large-scale building energy systems [1]. This chapter aims at presenting motivation and contributions of this study.

1.1 Motivation of Problem

We consider an unconstrained collaborative optimization of a sum of convex functions where agents make decisions using local information. Each agent has its own private convex cost function that wishes to reach the minimizer of the sum of agents' cost functions by interacting with its neighbors, known as *distributed optimization*. Distributed optimization problems naturally appear in many distributed problems such as sensor networks, power grid control, source localization, distributed data regression [2]-[3], and large-scale building energy systems [1].

Consensus problems are long-established problems in automata theory and distributed computation [4] and management science and statistics [5]. In multi-agent networks, *consensus* means reaching an agreement which depends on all agents' states by interacting with neighbors. Motivated by the pioneering works of Borkar & Varaiya [6] and Tsitsiklis [7] (see also [8]), many researchers have paid much attention to consensus and distributed optimization problems [1]-[3], [9]-[134].

In most networks such as sensor networks, since nodes sometimes shut down their transmitters to save energy, or due to existing physical obstructions which block wireless channels, the availability of communication links in the network is typically random. Indeed, because of packet drops, link failures, or node failures, random graphs are suitable models for these kinds of networks. As a matter of fact, consensus problems over random networks have been a hot topic to research due to their applications in sensor networks [45]–[48]. Therefore, several researchers have investigated consensus problems and distributed convex optimization problems over random networks [27]-[72].

In a *synchronous* protocol, all nodes activate at the same time and perform communication updates. This protocol requires a common notion of time among the nodes. On the other hand, in an *asynchronous* protocol, each node has its own concept of time defined by a local timer which randomly triggers either by the local timer or by a message from neighboring nodes. The algorithms guaranteed to work with no bound on the time for updates are called *totally asynchronous*, and those that need B-connectivity assumption, namely there exists a bounded time interval such that union of the graphs is strongly connected and each edge transmits a message at least once, are called *partially asynchronous* (see [7] and [8, Ch. 6-7]). As the dimension of the network increases, synchronization becomes an issue; therefore, some investigators have considered asynchronous distributed optimization problems [70]-[84], to cite a few.

In practice, state-dependent networks appear in several systems such as mobile robotic networks (see [86] and references therein), wireless networks [87], and predator-prey interaction [88]. In mobile robotic networks or wireless networks, the quality of the link between two agents depends on the distance between them, resulting in state-dependent networks in reality (by considering the position as the state). Furthermore, opinion dynamics and flocking are modeled as state-dependent networks (see [89] and references therein). Also, the genome is viewed as a state-dependent network [90]. As stated, state-dependent networks appear in real networks. In [91], existence of consensus in a multi-robot network has been investigated. Since consensus problems are special cases of distributed optimization problems, solving distributed convex optimization problems over state-dependent networks are very important and useful. Therefore, some researcher has paid attention to solving distributed optimization with state-dependent interactions [52] and [85].

A special case of optimization problems is solving linear algebraic equations. Linear algebraic equations arise in modeling of many natural phenomena such as forecasting and estimation [92]. Since the processors are physically separated from each others, distributed computations to solve linear algebraic equations are important and useful. Several authors have proposed algorithms for solving the problem over non-random networks [93]-[134].

The algorithms for distributed optimization over random networks are special cases of stochastic systems. In these systems, almost sure and moment stability are the most popular notions of stability [135]-[136]. Lyapunov' direct method has been used for stability of stochastic discrete-time systems [137]-[152]. Moreover, Lyapunov measure which is dual to Lyapunov function has been introduced for stability analysis of stochastic discrete-time systems [153]-[154]. The converse Lyapunov's theorem for stochastic discrete-time systems has been studied in [155]. Recently, stochastic version of LaSalle's theorem has been developed for discrete-time systems [156].

As we have shown above, distributed optimization problems and stability of stochastic discrete-time systems are important and useful in practice.

1.2 Literature Review

Consensus problems: As stated in the previous section, several researchers have investigated consensus problems. To the best of our knowledge, in all existing results except [23] and [44], the distribution of random interconnection topologies or B-connectivity assumption is needed. In [23] for time-varying directed networks, the authors have improved the lengths of B-connectivity intervals to linear grow and tend to infinity. In [44], it has been shown that if the elements of a set of communication graphs whose union is strongly connected occur infinitely often almost surely, then distributed average consensus occurs[1]. In [44], the authors consider undirected links with Maximum-degree or Metropolis weights. In Maximum-degree weight, the number of the nodes is required to set the weights of links, while in Metropolis weights the degrees of an agent's neighbors are required.

Distributed optimization with state-independent interactions: As stated in the previous section, investigators have considered distributed optimization problems with state-independent interactions

[1]We mean that the authors in [44] have implicitly stated that.

over random or non-random networks with/without asynchronous protocols. The results are based on distribution dependency or B-connectivity assumptions of communication graphs.

Distributed optimization with state-dependent interactions: In [52] and [85], the authors have considered distributed multi-agent optimization problems over state-dependent networks. In [52], the authors assume that the weighted matrix of the graph is Markovian on the state variables at each iteration. In a non-random case of the network contemplated in [52], the Markovian assumption is not satisfied because the state-dependent weighted matrix of the graph at each iteration does depend on previous iterations. In [85], a continuous-time system is proposed for unconstrained optimal consensus of convex optimization problem over directed time-varying networks; the authors assume that the weight of each link has positive lower bound. Furthermore, they assume that the intersection of the set of optimal solutions of each agent's cost function should be nonempty. The continuous-time algorithm they propose needs to project each agent's state into its optimal solution set at each time.

Solving linear algebraic equations over networks: The linear algebraic equation considered in this study is of the form $Ax = b$ that is solved simultaneously by m agents assumed to know only a subset of the rows of the partitioned matrix $[A, b]$, by using local information from their neighbors; indeed, each agent only knows $A_i x_i = b_i, i = 1, 2, ..., m$, where the goal of them is to achieve a consensus $x_1 = x_2 = ... = x_m = \tilde{x}$ where $\tilde{x} \in \{\bar{x} | \bar{x} = \arg\min \|Ax - b\|\}$. Several authors have proposed algorithms for solving the problem over non-random networks [93]-[119]. Other distributed algorithms for solving linear algebraic equations have been proposed by some investigators [120]-[134] that the problems they consider are not the same as the problem considered here. Some approaches propose cooperative solution methods that exploit the matrix A interconnectivity and have each node in charge of one single solution variable or a dual variable [120]-[122]. One view of the problem is to formulate it as a constrained consensus problem over random networks and use the result in [50]; nevertheless, the result in [50] needs each agent to use projection onto its constraint set with some probability at each time and also needs weighted matrix of the graph to be independent at each time. Another view of the problem is to formulate it as a distributed convex optimization problem over random networks and use the results in [51], [52], [54], [70]. Nevertheless, the results in [51], [52], [54], [70] are based on subgradient descent or diminishing step size that have slow convergence as an optimal solution is approached. Furthermore, the results in [51], [52], [54], [70] need weighted matrix of the graph to be independent and identically distributed (i.i.d.). Recently, the authors of [111] have proposed asynchronous algorithms for solving the linear algebraic equation over time-varying networks where they impose B-connectivity assumption.

Stability of stochastic discrete-time systems: Although Lyapunov's and LaSalle's methods have ben useful tools for stability analysis of stochastic discrete-time systems, they need distribution dependency of random variable sequences. Lyapunov' direct method needs stochastic parameters to be i.i.d. [137]-[144] or stochastic parameters to be Markov processes [145]-[152]. Moreover, Lyapunov measure needs stochastic parameters to be i.i.d. [153]-[154]. The converse Lyapunov's theorem [155] needs random variable sequence to be an i.i.d. process. LaSalle's theorem for stochastic discrete-time systems [156] needs stochastic parameters to be independent. Quadratic Lyapunov functions have been useful to analyze stability of linear dynamical systems. Nevertheless, common quadratic Lyapunov functions may not exist for stability analysis of consensus problems in networked systems [157]. Furthermore, quadratic Lyapunov functions may not exist for stability analysis of switched linear systems [158]-[160]. For deterministic discrete-time systems, by proving a converse to the Banach's fixed point theorem and using the Banach's fixed point theorem, [161]-[162] prove necessary and sufficient conditions for global and local exponential stability of deterministic nonlinear systems that is locally continuously differentiable at its equilibrium point.

As we have shown above, distribution dependency or B-connectivity assumptions are the limitations of existing results for stability of stochastic discrete-time systems and distributed optimization over random networks with/without asynchronous updates. Furthermore, lower bound on nonlinear weight and nonempty intersection of optimal solutions of agents cost functions are limitations of existing works for distributed optimization with state-dependent interactions. We mention that in practice Cucker-Smale weight [163] does not have a positive lower bound that appears in biological networks.

1.3 Contributions

Before working on distributed optimization problems, the author published two papers [164] and [165] on applications of *fixed point theory* to overcome the fundamental difficulties mentioned in [164] which arise in using Lyapunov's and LaSalle's methods for stability analysis of time-varying systems with time delay. Here, our contribution in this study is a new perspective on distributed optimization by using the mathematical theory of random maps. We show that by applying fixed point theory, we are able to overcome the limitations of existing results for packet drops, synchrony, and state-dependent weights in distributed optimization. We state our proposed approaches in the remaining paragraphs.

Consensus problems: We consider the consensus problem over random networks. We show that this problem is to find a *fixed value point* of the random operator formed from the random weighted graph matrices. We assume that the random weighted graph matrices are doubly stochastic for all possible graphs. This assumption allows us to discard the distribution of random interconnection topologies. Consequently, this formulation includes asynchronous updates or/and unreliable communication protocols. Furthermore, this framework does not need distribution of random interconnection topologies or B-connectivity assumptions. Wireless sensor networks motivates this framework since interference among the sensors communication correlates the links' failures over probability space or time. We show that the random Krasnoselskii-Mann iterative algorithm converges almost surely and in mean square to the average consensus of initial states of the agents. We also show that the agents interact among themselves to approach the consensus subspace in such a way that the projection of their states onto consensus subspace at each time is equal to the average consensus of their initial states. Moreover, the algorithm is able to converge even if the interconnection weighted matrix is periodic and irreducible. We should mention that existing discrete-time algorithms for consensus problems are the algorithm proposed by Tsitsiklis [7] and its generalizations to random cases; this algorithm is in fact the Picard iteration.

Distributed optimization with state-independent interactions: We consider the problem of unconstrained distributed convex optimization over random networks. We approach the problem using a random operator formed from the random weighted graph matrices. We show that the distributed optimization problem can be formulated as minimization of a convex function over the set of fixed value points of the random operator. Since the random operator is nonexpansive, we define a mathematical optimization problem, namely minimization of a convex function over the set of fixed value points of a nonexpansive random operator, which includes the distributed optimization problem as a special case. The definition of fixed value point is a bridge from deterministic analysis to random analysis of the algorithm. With the help of fixed value point set and nonexpansivity property of the random operator, we are able to extend deterministic tools to random cases to prove boundedness, convergence to the feasible set, and convergence to the optimal solution of the generated sequence. This is very useful because we are able to analyze random processes by using extended deterministic tools. We propose a discrete-time algorithm using diminishing step size for almost sure and in mean square convergences to the optimal solution of the mathematical optimization problem. This framework does not need any assumption on distribution of random interconnection graphs. The proposed algorithm is also able to reach the optimal solution under asynchronous updates. Our algorithm is not comparable to existing algorithms since they need i.i.d. or B-connectivity assumptions.

Distributed optimization with state-dependent interactions: We consider an unconstrained distributed convex optimization problem over time-varying networks with state-dependent interactions. The union of graphs which occur infinitely often is assumed to be strongly connected, and the weights depend on the states continuously for each graph. We propose a framework for modeling multi-agent optimization problems over state-dependent networks with time-varying topologies, i.e., the minimization of sum of convex functions over the intersection of fixed point sets of operators constructed. We assume that each agent's cost function is strongly convex with Lipschitzian gradient and that weighted graph matrix of the network is doubly stochastic with respect to state variables at each time. We propose a gradient-based discrete-time algorithm using diminishing step size for converging to the optimal solution of the problem. Our algorithm does not require the weights to have positive lower bounds. This allows us to consider Cucker-Smale weights [163].

To the best of our knowledge, in contrast to existing results, our algorithm does not require B-connectivity assumption for convergence. Therefore, our results are not comparable with existing results even in state-independent case.

Solving linear algebraic equations over networks: Several authors in the literature have considered solving linear algebraic equations over switching networks with B-connectivity assumption such as [111]. However, B-connectivity assumption is not guaranteed to be satisfied for random networks. We formulate this problem such that this formulation does not need the distribution of random communication graphs or B-connectivity assumption if the weighted matrix of the graph is doubly stochastic. Thus this formulation includes asynchronous updates or unreliable communication protocols. We assume that the set $\mathcal{S} = \{x | \min_{x} \|Ax - b\| = 0\}$ is nonempty. Since the Picard iterative algorithm may not converge, we apply the random Krasnoselskii-Mann iterative algorithm for converging almost surely and in mean square to a point in $\mathcal{S}$ for *any* matrices A and b and *any* initial conditions. The proposed algorithm, like that of [111], requires that whole solution vector is computed and exchanged by each node over a network. Based on initial conditions of agents' states, we show that the limit point to which the agents' states converge is determined by the unique solution of a feasible convex optimization problem independent from the distribution of random communication graphs.

Stability of stochastic discrete-time systems: We apply fixed point theory to stability analysis of stochastic nonlinear discrete-time systems to overcome difficulties that arise in using Lyapunov's and LaSalle's approaches such as distribution dependency of random variable sequences.

1.4 Outline of Thesis

The outline of thesis is as follows. In Chapter 2, review of relevant fixed point theory is given. In Chapter 3, a new optimization problem and its special cases such as robust convex optimization and distributed optimization over random networks are given. Furthermore, two frameworks for distributed optimization over state-dependent networks with/without switching topologies are presented. In Chapter 4, an algorithm for solving the optimization problem defined in

Chapter 3, and its application to distributed optimization, is presented. In Chapter 5, solving linear algebraic equations, and its special case consensus problems, over random networks is considered. In Chapter 6, a generalization of the proposed algorithm to solve distributed optimization with state-dependent interactions and time-varying topologies is given. In Chapter 7, applied fixed point theory to stability analysis of stochastic discrete-time systems is presented. In Chapter 8, conclusions and future works are given.

CHAPTER 2. REVIEW OF RELEVANT FIXED POINT THEORY

In this chapter, we give relevant fixed point theorems and iterations that we use for our results. Before doing so, we present a history of relevant results. The following brief history of fixed point theory has been mentioned in [166].

The idea of fixed point of an operator was first flashed in the mind of Cauchy while dealing with the existence and uniqueness of solution of certain differential equation and by this notion, a new light in the research arena appeared as Fixed Point Theory. This has two-fold-valuation– one from the classical analysis point of view, and the other is its application on many branches of sciences and economics.

After Cauchy, R. Lipschitz simplified Cauchy's proof in 1877 (1876 in [167]) using *Lipschitz condition*, and G. Peano proved a deeper result in 1890 which relates mostly to the modern fixed point theory. In the same year Picard applied this method to ordinary and partial differential equations. In 1886, Poincaré proved a fixed point theorem for a continuous self-mapping f on $\Re^n$ satisfying condition $f(x) + \alpha x = r, \|x\| = r, \forall x \in \Re^n$, for some $r > 0$ and for every $\alpha > 0$. This theorem was rediscovered by P. Bohl in 1904. For a long period of time, this branch remained suppressed until it was redeemed and re-cultivated by the Dutch mathematician L. E. J. Brouwer who put this branch of mathematics in the front line of the research arena. In 1912 (1910 in [167]), he proved the well-known *Brouwer fixed point theorem* for a continuous self-map on a closed unit ball in $\Re^n$.

In 1922, S. Banach launched in this field with a new concept of mapping called *contraction mapping* and showed that a contraction self-mapping on a complete metric space has a unique fixed point. In 1930, R. Caccioppoli remarked on Banach contraction principle that the contraction condition may be replaced by the assumption of the convergence of the sequence of iterates, which led to open another direction of studying fixed point theory, known as *approximation of fixed point* *of an operator*. So, to speak on iterative sequence, the Picard iterative scheme has a wide range of applications in different branches of sciences. Nevertheless, it has found to have some crucial drawback that the iterative sequence obtained by this method may not always converge, which was pointed out and rectified by W. R. Mann in 1953 by introducing a new type of iteration scheme, called the *Mann iterative process*. In 1930, J. Schauder obtained the result for existing a fixed point for a continuous mapping on a Banach space. Under the assumption of the Schauder theorem, there was no method for approximating a fixed point of a mapping. However, Krasnoselskii showed in 1955 that a special type of iterative sequence converges to a fixed point of a nonexpansive mapping on a uniformly convex Banach space.

It is assumed that the reader is familiar with usual concepts of topological and metric spaces or is referred to [168]. Now we present relevant fixed point theorems and iterations in the following sections, respectively.

2.1 Relevant Fixed Point Theorems

Before we presents theorems, we need to give some definitions.

Definition 2.1: Let the operator $T : \mathcal{X} \longrightarrow \mathcal{X}$ be a self map. A $x \in \mathcal{X}$ is said to be a *fixed point* of T if $T(x) = x$, and $Fix(T)$ denotes the set of all fixed points of T.

Definition 2.2 [167]: A topological space $\mathcal{X}$ is said to possess *fixed point property* if every continuous mapping of $\mathcal{X}$ into $\mathcal{X}$ has a fixed point.

Definition 2.3 [167]: Let T be a mapping of a metric space $\mathcal{M} = (\mathcal{X}, \rho)$ into $\mathcal{M}$. T is called *contraction mapping* if there exists a number η such that $0 \leq \kappa < 1$ and

$$\rho(Tx, Ty) \leq \kappa\rho(x, y).$$

Definition 2.4: Let $\mathcal{H} = (\mathcal{X}, \|.\|_{\mathcal{H}})$ with inner product $< ., . >$ be a real Hilbert space. A self map operator $H : \mathcal{H} \longrightarrow \mathcal{H}$ is said to be *nonexpansive* if for any $x, y \in dom(H)$ we have

$$\|H(x) - H(y)\|_{\mathcal{H}} \leq \|x - y\|_{\mathcal{H}}.$$

Definition 2.5: The map $T : \mathcal{H} \longrightarrow \mathcal{H}$ is said to be *firmly nonexpansive* if for each $x, y \in \mathcal{H}$,

$$\|T(x) - T(y)\|_{\mathcal{H}}^2 \leq < T(x) - T(y), x - y > .$$

Remark 2.1 [169]: $\phi : \mathcal{H} \longrightarrow \mathcal{H}$ is a firmly nonexpansive mapping if $T : \mathcal{H} \longrightarrow \mathcal{H}$ is a nonexpansive mapping where

$$\phi(x) = \frac{1}{2}(x + T(x)).$$

Moreover, every firmly nonexpansive mapping is nonexpansive by Cauchy–Schwarz inequality.

Let (Ω^*, σ) be a measurable space (σ-sigma algebra) and C be a nonempty subset of a metric space $\mathcal{M}$. A mapping $x : \Omega^* \longrightarrow \mathcal{M}$ is *measurable* if $x^{-1}(U) \in \sigma$ for each open subset U of $\mathcal{M}$. The mapping $T : \Omega^* \times C \longrightarrow \mathcal{M}$ is a *random map* if for each fixed $z \in C$, the mapping $T(., z) : \Omega^* \longrightarrow \mathcal{M}$ is measurable, and it is *continuous* if for each $\omega^* \in \Omega^*$ the mapping $T(\omega^*, .) : C \longrightarrow \mathcal{M}$ is continuous.

Definition 2.6 [170]: A measurable mapping $x : \Omega^* \longrightarrow \mathcal{M}$ is a *random fixed point* of the random map $T : \Omega^* \times C \longrightarrow \mathcal{M}$ if $T(\omega^*, x(\omega^*)) = x(\omega^*)$ for each $\omega^* \in \Omega^*$.

Definition 2.7 [170]: Let C be a nonempty subset of a metric space $\mathcal{M}$ and $T : \Omega^* \times C \longrightarrow C$ be a random map. The map T is said to be *contraction random operator* if for each $\omega^* \in \Omega^*$ and for arbitrary $x, y \in C$ we have

$$\rho(T(\omega^*, x), T(\omega^*, y)) \leq \kappa \rho(x, y), \quad 0 \leq \kappa < 1.$$

Definition 2.8 [170]: Let C be a nonempty subset of a real Hilbert space $\mathcal{H}$ and $T : \Omega^* \times C \longrightarrow C$ be a random map. The map T is said to be *nonexpansive random operator* if for each $\omega^* \in \Omega^*$ and for arbitrary $x, y \in C$ we have

$$\|T(\omega^*, x) - T(\omega^*, y)\|_{\mathcal{H}} \leq \|x - y\|_{\mathcal{H}}.$$

Definition 2.8 [170]: Let C be a nonempty subset of a real Hilbert space $\mathcal{H}$ and $T : \Omega^* \times C \longrightarrow C$ be a random map. The map T is said to be *firmly nonexpansive random operator* if for each $\omega^* \in \Omega^*$ and for arbitrary $x, y \in C$ we have

$$\|T(\omega^*, x) - T(\omega^*, y)\|_{\mathcal{H}}^2 \leq < T(\omega^*, x) - T(\omega^*, y), x - y > .$$

Now we present the relevant fixed point theorems.

Theorem 2.1 [167] *(The Brouwer fixed point theorem)*: Every compact convex non-empty subset of $\Re^n$ has the fixed point property.

The following theorem is in fact equivalent to the Brouwer fixed point theorem.

Theorem 2.2 [167]: B^n has the fixed point property.

Theorem 2.3 [167] *(The Banach Fixed Point Theorem)*: Any contraction mapping of a complete non-empty metric space $\mathcal{M}$ into itself has a unique fixed point.

2.2 Relevant Fixed Point Iterations

The Picard iterative algorithm: The Picard iteration for finding a fixed point of an operator $T(x)$ is

$$x_{n+1} = T(x_n), \quad n \in N \cup \{0\}. \tag{2.1}$$

The Krasnoselskii-Mann iterative algorithm [171]-[172]: The Krasnoselskii-Mann iteration for finding a fixed point of an operator $T(x)$ is

$$x_{n+1} = (1 - \alpha_n)x_n + \alpha_n T(x_n), \quad n \in N \cup \{0\}, \tag{2.2}$$

where $\alpha_n \in [0, 1]$.

The Picard iteration may not always converge when $T(x)$ is nonexpansive on a real Hilbert space $\mathcal{H}$. For example, consider $T(x) := -x, x \in \Re$. However, Krasnoselskii [171] proved that Algorithm (2.2) when $\alpha_n = \frac{1}{2}$ always converges to a fixed point of a nonexpansive mapping on $\mathcal{H}$.

CHAPTER 3. TWO FRAMEWORKS FOR OPTIMIZATION

In this chapter, we define a new mathematical terminology called *fixed value point* and use it to define a new mathematical optimization framework. The framework includes centralized convex optimization, centralized robust convex optimization, and distributed convex optimization over random networks. Then we give a framework for distributed optimization problem with state-dependent interactions. The reader is assumed to be familiar with convex optimization concepts or is referred to [173].

Now we give the following definition.

Definition 3.1: If there exists a point $\hat{x} \in \mathcal{M}$ where $\hat{x} = T(\omega^*, \hat{x})$ for all $\omega^* \in \Omega^*$, we call it *fixed value point*, and $FVP(T)$ represents the set of all fixed value points of T.

Remark 3.1: A random mapping may have a random fixed point but may not have a fixed value point. For instance, if $\Omega^* = \{H, G\}$ and $T(H, x(H)) = 1, T(G, x(G)) = 0$, then the random variable $x(H) = 1, x(G) = 0$ is a random fixed point of T. However, T does not have any fixed value point.

Remark 3.2: A fixed value point of a nonexpansive random mapping is a common fixed point of a family of nonexpansive non-random mappings $T(\omega^*, .)$ for each ω^*. We refer the interested reader to [174]-[176] for existence theorems for a common fixed point of nonexpansive non-random mappings.

3.1 A Framework for Convex Optimization

Let $\mathcal{H}$ be a real Hilbert space. Given a convex function $f : \mathcal{H} \longrightarrow \Re$ and a nonexpansive random mapping $T : \Omega^* \times \mathcal{H} \longrightarrow \mathcal{H}$, the problem is to find $x^* \in \underset{x}{\mathrm{argmin}} f(x)$ such that x^* is a fixed value point of $T(\omega^*, x)$, i.e., we have the following minimization problem

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & x \in FVP(T) \end{aligned} \tag{3.1}$$

where $FVP(T)$ is the set of fixed value points of the random operator $T(\omega^*, x)$ (see Definition 3.1). We assume that the problem is feasible, namely $FVP(T) \neq \emptyset$.

The following preposition is a corollary of Preposition 5.3 in [177].

Preposition 3.1: Let C be a closed and convex subset of a real Hilbert space $\mathcal{H}$. If $T : C \longrightarrow C$ is nonexpansive, then the fixed point set of T is closed and convex.

Remark 3.3: From Preposition 3.1, the fixed point set of a nonexpansive non-random mapping $T : \Omega^* \times C \longrightarrow C$ where C is a closed convex set, for each ω^*, is a closed convex set. It is known that the intersection of closed convex sets (finite, countable, or uncountable) is closed and convex. Since, by Remark 3.2, fixed value points set of a nonexpansive random operator $T(\omega^*, x)$ is the intersection of fixed points set of nonexpansive non-random mappings $T(\omega^*, x)$ for each fixed $\omega^* \in \Omega^*$, i.e., $FVP(T) = \bigcap_{\omega^* \in \Omega^*} Fix(T(\omega^*, x))$, we have that $FVP(T)$ is a closed convex set. Therefore, Problem (3.1) is a convex optimization problem.

3.1.1 Application to Centralized Convex Optimization

Unconstrained optimization problem, i.e., $\min_{x} f(x)$, is included in the framework (3.1). In this case, the constraint set is $x = x$, or $x \in Fix(T)$ where $T(x) := x$. It is easy to check that $T(x) := x$ is nonexpansive. Moreover, constrained optimization problems are included in the framework (3.1). In this case, the constraint set is $x = \mathcal{P}_C(x)$ where C is a closed convex constraint set, and $\mathcal{P}_C(x)$ is the projection map onto C. It is known that $T(x) := \mathcal{P}_C(x)$ is nonexpansive.

3.1.2 Application to Centralized Robust Convex Optimization

Robust convex optimization has been investigated on Euclidean spaces [178]-[203]. In this subsection, we define centralized robust convex optimization (CRCO) on real Hilbert spaces. CRCO on $\mathcal{H}$ is in general of the form

$$\begin{aligned} \min_{x \in C} \quad & f(x) \\ \text{subject to} \quad & g(x, \omega^*) \leq 0, \forall \omega^* \in \Omega^*, \end{aligned} \tag{3.2}$$

where C is a nonempty closed convex subset of $\mathcal{H}$, $f : \mathcal{H} \longrightarrow \Re$ is a convex function, $g(x, .) : \Omega^* \longrightarrow \Re$ is measurable for each fixed $x \in \mathcal{H}$, $g(., \omega^*) : \mathcal{H} \longrightarrow \Re$ is a convex function for each fixed $\omega^* \in \Omega^*$, and the uncertainty ω^* enters into the constraint function $g(x, \omega^*)$. Assume the problem is feasible, i.e., there exists an $x^* \in C$ such that $g(x^*, \omega^*) \leq 0, \forall \omega^* \in \Omega^*$.

The constraint set of (3.2), i.e., $\{x | x \in C, g(x, \omega^*) \leq 0, \forall \omega^* \in \Omega^*\}$, can be converted to $\{x | x = \mathcal{P}^g(\omega^*, x), \forall \omega^* \in \Omega^*\}$ where $\mathcal{P}^g(\omega^*, x)$ is the projection onto the closed convex set $\{z | z \in C, g(z, \omega^*) \leq 0\}$ for each fixed $\omega^* \in \Omega^*$. The constraint set of the CRCO problem (3.2) is in fact $x \in FVP(\mathcal{P}^g)$. Therefore, (3.2) is equivalent to

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & x \in FVP(\mathcal{P}^g). \end{aligned} \tag{3.3}$$

Since the projection operator $\mathcal{P}^g(\omega^*, x)$ is nonexpansive, (3.1) includes (3.3) as a special case.

3.1.3 Application to Distributed Convex Optimization over Random Networks

In this subsection, we define the distributed convex optimization problem over random networks.

A network of m nodes labeled by the set $\mathcal{V} = \{1, 2, ..., m\}$ is considered. The topology of the interconnections among nodes is not fixed but defined by a set of graphs $\mathcal{G}(\omega^*) = (\mathcal{V}, \mathcal{E}(\omega^*))$ where $\mathcal{E}(\omega^*)$ is the ordered edge set $\mathcal{E}(\omega^*) \subseteq \mathcal{V} \times \mathcal{V}$ and $\omega^* \in \Omega^*$ where Ω^* is the set of all possible communication graphs, i.e., $\Omega^* = \{\mathcal{G}_1, \mathcal{G}_2, ..., \mathcal{G}_{\bar{N}}\}$. We write $\mathcal{N}_i^{in}(\omega^*)/\mathcal{N}_i^{out}(\omega^*)$ for the labels of agent i's in/out neighbors at graph $\mathcal{G}(\omega^*)$ so that there is an arc in $\mathcal{G}(\omega^*)$ from vertex j/i to vertex i/j only if agent i receives/sends information from/to agent j. We write $\mathcal{N}_i(\omega^*)$ when $\mathcal{N}_i^{in}(\omega^*) = \mathcal{N}_i^{out}(\omega^*)$. We assume that there are no self-looped arcs in the communication graphs.

We define the weighted graph matrix $\mathcal{W}(\omega^*) = [\mathcal{W}_{ij}(\omega^*)]$ with $\mathcal{W}_{ij}(\omega^*) = a_{ij}(\omega^*)$ for $j \in \mathcal{N}_i^{in}(\omega^*) \cup \{i\}$, and $\mathcal{W}_{ij}(\omega^*) = 0$ otherwise, where $a_{ij}(\omega^*) > 0$ is the scalar constant weight that agent i assigns to the information x_j received from agent j. For instance, if $\mathcal{W}(\mathcal{G}_k) = I_m$, for some $1 \leq k \leq \bar{N}$, implies that there are no edges in $\mathcal{G}_k$, or/and all nodes are not activated for communication updates in asynchronous protocol or both.

Now we define the distributed convex optimization problem as follows: for each node $i \in \mathcal{V}$, we associate a private convex cost function $f_i : \Re^n \longrightarrow \Re$ which is known to node i. The objective of each agent is to collaboratively seek the solution of the following optimization problem using local information exchange with the neighbors and switching communication topologies:

$$\min \sum_{i=1}^{m} f_i(x)$$

where $x \in \Re^n$. We assume that there is no communication delay or noise in delivering a message from agent j to agent i.

The full formulation of the above problem is as follows:

Problem 3.1: The distributed convex optimization is formulated as

$$\begin{aligned} \min_{x} \qquad & f(x) := \sum_{i=1}^{m} f_i(x_i) \\ \text{subject to} \quad & x_1 = x_2 = ... = x_m \end{aligned} \tag{3.4}$$

where $x = [x_1^T, ..., x_m^T]^T, x_i \in \Re^n, i = 1, 2, ..., m,\ f_i : \Re^n \longrightarrow \Re$ is a private cost function known to node i, and the constraint is achieved through random graph interactions.

Remark 3.4: The set $\mathcal{C} = \{x \in \Re^{mn} | x_i = x_j, 1 \leq i, j \leq m, x_i \in \Re^n\}$ is known as *consensus subspace.*

Now we impose the following assumptions.

Assumption 3.1: The weighted graph matrix $\mathcal{W}(\omega^*)$ is doubly stochastic for each $\omega^* \in \Omega^*$, i.e.,

i) $\sum_{j \in \mathcal{N}_i^{in}(\omega^*) \cup \{i\}} \mathcal{W}_{ij}(\omega^*) = 1, i = 1, 2, ..., m,$

ii) $\sum_{j \in \mathcal{N}_i^{out}(\omega^*) \cup \{i\}} \mathcal{W}_{ij}(\omega^*) = 1, i = 1, 2, ..., m.$

Note that any network with undirected links satisfies Assumption 3.1.

Assumption 3.2: The union of all of the graphs in Ω^* is strongly connected.

Now we give the following lemma regarding Assumption 3.2.

Lemma 3.1: The union of all of the graphs in Ω^* is strongly connected if and only if $Re[\lambda_2(\sum_{\omega^*\in\Omega^*}(I_m - \mathcal{W}(\omega^*)))] > 0$.

Proof: Since the union of all of the graphs is strongly connected, the matrix $\sum_{\omega^*\in\Omega^*}\mathcal{W}(\omega^*)$ is irreducible. Therefore, according to Perron-Frobenius theorem for irreducible matrices, it has a unique positive real largest eigenvalue. Since, by Assumption 3.1, $\mathcal{W}(\omega^*), \forall\omega^* \in \Omega^*$, is doubly stochastic, the unique largest eigenvalue of the matrix $\sum_{\omega^*\in\Omega^*}\mathcal{W}(\omega^*)$ is $\lambda^* = \bar{N}$. Thus $Re[\lambda_2(\sum_{\omega^*\in\Omega^*}(I_m - \mathcal{W}(\omega^*)))] > 0$. Conversely, we prove it by contradiction. Assume that the union of all of the graphs is not strongly connected. It is well-known that there exists a permutation matrix P such that

$$P(\sum_{\omega^*\in\Omega^*}\mathcal{W}(\omega^*))P^T = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

where A and C are square matrices. Therefore, $spec(\sum_{\omega^*\in\Omega^*}\mathcal{W}(\omega^*)) = spec(A) \cup spec(C)$. From Assumption 3.1, all columns of A has summation equal to $\bar{N}$, and all rows of C have summation equal to $\bar{N}$. Therefore, the eigenvalue $\lambda = \bar{N}$ has multiplicity 2 in $spec(\sum_{\omega^*\in\Omega^*}\mathcal{W}(\omega^*))$ which is a contradiction. Thus the proof is complete.

Assumption 3.2 ensures that the information sent from each node will be finally obtained by every other node through a directed path. Now, Problem 3.1 with Assumptions 3.1 and 3.2 can be reformulated as the following problem.

Problem 3.2: Problem 3.1 under Assumptions 3.1 and 3.2 can be formulated as

$$\begin{aligned} \min_x \quad & f(x) := \sum_{i=1}^m f_i(x_i) \\ \text{subject to} \quad & W(\omega^*)x = x, \forall\omega^* \in \Omega^*, \end{aligned} \tag{3.5}$$

where $W(\omega^*) = \mathcal{W}(\omega^*) \otimes I_n, \omega^* \in \Omega^*$.

Now we show that Problems 3.1 and 3.2 are equivalent. We obtain from $W(\omega^*)x = x, \forall\omega^* \in \Omega^*$ that

$$(I_{mn} - W(\omega^*))x = 0, \forall\omega^* \in \Omega^*,$$

which implies that

$$\sum_{\omega^*\in\Omega^*}(I_{mn} - W(\omega^*))x = 0. \tag{3.6}$$

Now we have

$$\begin{aligned}\sum_{\omega^*\in\Omega^*}(I_{mn} - W(\omega^*)) &= \sum_{\omega^*\in\Omega^*}((I_m - \mathcal{W}(\omega^*)) \otimes I_n)\\ &= (\sum_{\omega^*\in\Omega^*}(I_m - \mathcal{W}(\omega^*))) \otimes I_n\\ &= \Lambda \otimes I_n,\end{aligned}$$

where

$$\Lambda := \sum_{\omega^*\in\Omega^*}(I_m - \mathcal{W}(\omega^*)).$$

Λ has the following properties: the summation of all rows are equal to zero; the diagonal elements are non-negative; the off-diagonal elements are non-positive. Therefore, Λ has the Laplacian matrix structure. Since $Re[\lambda_2(\Lambda)] > 0$ (see the proof of Lemma 3.1), (3.6) implies that $x_1 = x_2 = \ldots = x_m$. Therefore, Problems 3.1 and 3.2 are equivalent.

Although double-stochasticity is restrictive in distributed setting [204], we show that Assumption 3.1 allows us to remove the distribution of random interconnection graphs. Now we show that the random operator $T(\omega^*, x) := W(\omega^*)x$ with Assumption 3.1 is nonexpansive in the Hilbert space $\mathcal{H} = (\Re^{mn}, \|.\|_2)$. For arbitrary $x, y \in \Re^{mn}$, we have

$$\begin{aligned}\|T(\omega^*, x) - T(\omega^*, y)\|_2 &= \|W(\omega^*)x - W(\omega^*)y\|_2\\ &= \|W(\omega^*)(x-y)\|_2\\ &\leq \|W(\omega^*)\|_2\|x-y\|_2.\end{aligned}$$

Now we have the following lemma.

Lemma 3.2 [205]: Let $W \in \Re^{m\times m}$. Then $\|W\|_2 \leq \sqrt{\|W\|_1\|W\|_\infty}$.

By Assumption 3.1 and Lemma 3.2, we obtain $\|W(\omega^*)\|_2 \leq 1, \forall \omega^* \in \Omega^*$. Thus we have

$$\begin{aligned}\|T(\omega^*, x) - T(\omega^*, y)\|_2 &\leq \|W(\omega^*)\|_2\|x-y\|_2\\ &\leq \|x-y\|_2 \qquad (3.7)\end{aligned}$$

which implies that T is nonexpasnive.

Now we give the following definition.

Definition 3.2: Given a weighted graph matrix $\mathcal{W}(\omega^*)$, we call $T(\omega^*, x) := W(\omega^*)x$, $\omega^* \in \Omega^*$, *weighted random operator of the graph.* Similarly, for non-random case, we call $T(x) := Wx$ *weighted operator of the graph.*

Remark 3.5: We show that since weighted random operator of the graph has nonexpansivity property in the Hilbert space $\mathcal{H} = (\Re^{mn}, \|.\|_2)$, any assumption on distribution of random communication topologies is not needed.

Remark 3.6: Consensus subspace is in fact the fixed value points set of weighted random operator of the graph with Assumption 3.2, i.e., $\mathcal{C} = FVP(T)$.

To conclude, we have shown that Problem 3.2 is a special case of (3.1) where $T(\omega^*, x) := W(\omega^*)x$.

3.2 A Framework for Distributed Convex Optimization with State-Dependent Interactions

Before we give a framework for distributed optimization with state-dependent interactions, we clarify that existing frameworks cannot be applied for this problem.

Fast Lipschitz Optimization has been introduced as a powerful method to capture the unique solution of convex or non-convex optimization problems [206]-[208]. If $\mathcal{W}$ does not depend on the states, i.e., the case of state-independent weighted graphs, the condition $\|W\| < 1$ is not satisfied for our problem because $\|W\|_2 = 1$; in fact, this condition makes the operator $T(x) := Wx$ a contraction, and the feasible set will be a unique point (see Theorem 2.3) instead of the set $\mathcal{C}$. Therefore, the results given in [206]-[208] cannot not be applied here.

Convex minimization over fixed point set of a nonexpansive mapping has been studies in [209] and references therein. This method has been usefully applied to signal processing, inverse problem, network bandwidth allocation and so on [210]-[214]. If W does not depend on the states, we have shown that the operator $T(x) := Wx$ is a nonexpansive mapping with Assumption 3.1. For any

$W(x)$, the operator T may not be a nonexpansive mapping. Therefore, the results given in [209] and references therein cannot be applied here. Similarly, the proposed framework in the previous section for non-random case cannot be applied here. Note that the framework of minimization over fixed point set of nonexpansive mapping considered in [209] and references therein includes the centralized convex optimization in Subsection 3.1.1 but does not include the problems in Subsections 3.1.2 and 3.1.3.

Similar to the previous subsection, distributed optimization with state-dependent interactions can be formulated as

$$\begin{aligned} \min_{x} \quad & f(x) := \sum_{i=1}^{m} f_i(x_i) \\ \text{subject to} \quad & W(x)x = x \end{aligned} \tag{3.8}$$

where $W(x) = \mathcal{W}(x) \otimes I_n$, and $\mathcal{W}(x) = [\mathcal{W}_{ij}(x_i, x_j)]$ is the state-dependent weighted matrix of the graph satisfying the following assumptions.

Assumption 3.3: The weights $\mathcal{W}_{ij} : \Re^n \times \Re^n \longrightarrow [0, 1]$ are continuous, and the state-dependent weighted matrix of the graph is doubly stochastic, i.e.,

i) $\sum_{j \in \mathcal{N}_i^{in} \cup \{i\}} \mathcal{W}_{ij}(x_i, x_j) = 1, i = 1, 2, ..., m,$

ii) $\sum_{j \in \mathcal{N}_i^{out} \cup \{i\}} \mathcal{W}_{ij}(x_i, x_j) = 1, j = 1, 2, ..., m.$

Assumption 3.4: The graph is strongly connected for all $x \in \Re^{mn}$.

Assumptions 3.3 and 3.4 ensure the connectivity of the graph, and that the information sent from each node will be finally obtained by every other node through a path. Note that Assumption 3.3 when applied to state-dependent weights would require undirected connections, and directed graphs are allowed in the special case of state-independent weights.

Now we show that the only solution of $W(x)x = x$ with Assumptions 3.3 and 3.4 is $x_1 = x_2 = ... = x_m$, i.e., the consensus subspace. Assume a $\tilde{x} = [\tilde{x}_1^T, \tilde{x}_2^T, ..., \tilde{x}_m^T]^T$ which satisfies $W(\tilde{x})\tilde{x} = \tilde{x}$. Since the summation of rows of $\mathcal{W}(\tilde{x}) - I_m$ is zero, the matrix $\mathcal{W}(\tilde{x}) - I_m$ has an eigenvalue zero; moreover, since the graph is strongly connected, this eigenvalue is unique. According to Assumptions 3.3 and 3.4, $W(\tilde{x})$ has some nonzero elements. Therefore, $\tilde{x}$ must be in the null space of $W(\tilde{x}) - I_{mn}$ which implies $\tilde{x}_1 = \tilde{x}_2 = ... = \tilde{x}_m$. Therefore, the only solution of $W(x)x = x$ with Assumptions 3.3 and 3.4 is $x_1 = x_2 = ... = x_m$.

Note that from Assumption 3.3 and Lemma 3.2, we have that $\|W(x)\|_2 \leq 1, \forall x \in \Re^{mn}$. The Hilbert space considered here is $\mathcal{H} = (\Re^{mn}, \|.\|_2)$.

Now, we introduce a framework for modeling multi-agent optimization problems.

Problem given by (3.8) with Assumptions 3.3 and 3.4 can be reformulated as

$$\begin{aligned} \min_{x} \quad & f(x) := \sum_{i=1}^{m} f_i(x_i) \\ \text{subject to} \quad & x \in Fix(T) \end{aligned} \tag{3.9}$$

where $T(x) := W(x)x$. We obtain $\|T(x)\|_2 \leq \|W(x)\|_2 \|x\|_2 \leq \|x\|_2$; in fact, the operator T maps every closed ball B^{mn} into itself. Thus, according to Theorem 2.2, we can guarantee that there exists a fixed point of T, i.e., there exists a point $\hat{x}$ such that $\hat{x} = W(\hat{x})\hat{x}$ which is the same as the constraint $W(x)x = x$.

Now we generalize the framework (3.9) to distributed optimization with state-dependent interactions and time-varying topologies.

The topology of the network is represented by $\mathcal{G}_n = (\mathcal{V}, \mathcal{E}_n)$ at time $n \in N \cup \{0\}$ with the ordered edge set $\mathcal{E}_n \subseteq \mathcal{V} \times \mathcal{V}$. Let consider the set $G = \{\mathcal{G}_n : n \in N \cup \{0\}\}$. Since $m \in N$, the cardinality of G, namely $|G| = \bar{N}$, is finite.

Now we impose the following assumptions.

Assumption 3.5: For each $\mathcal{G} \in G$, the weights $\mathcal{W}_{ij}(\mathcal{G}) : \Re^n \times \Re^n \longrightarrow [0,1]$ are continuous, and the state-dependent weighted matrix of the graph is doubly stochastic, i.e.,

i) $\sum_{j \in \mathcal{N}_i^{in} \cup \{i\}} \mathcal{W}_{ij}(x_i, x_j, \mathcal{G}) = 1, i = 1, 2, ..., m,$

ii) $\sum_{j \in \mathcal{N}_i^{out} \cup \{i\}} \mathcal{W}_{ij}(x_i, x_j, \mathcal{G}) = 1, j = 1, 2, ..., m.$

Assumption 3.6: The union of the graphs in G is strongly connected for all $x \in \Re^{mn}$.

Now we give the following lemma regarding Assumption 3.6.

Lemma 3.3: The union of the graphs in G is strongly connected for all $x \in \Re^{mn}$ if and only if $Re[\lambda_2(\sum_{\mathcal{G} \in G}(I_m - \mathcal{W}(x, \mathcal{G})))] > 0$ for all $x \in \Re^{mn}$.

Proof: Since the union of all of the graphs is strongly connected for all $x \in \Re^{mn}$, the matrix $\sum_{\mathcal{G}\in G} \mathcal{W}(x,\mathcal{G})$ is irreducible for all $x \in \Re^{mn}$. Therefore, according to Perron-Frobenius theorem for irreducible matrices, it has a unique positive real largest eigenvalue for each $x \in \Re^{mn}$. Since, by Assumption 3.5, $\mathcal{W}(x,\mathcal{G}), \forall \mathcal{G} \in G, \forall x \in \Re^{mn}$, is doubly stochastic, the unique largest eigenvalue of the matrix $\sum_{\mathcal{G}\in G} \mathcal{W}(x,\mathcal{G})$ is $\lambda^*(x) = \bar{N}$. Thus $Re[\lambda_2(\sum_{\mathcal{G}\in G}(I_m - \mathcal{W}(x,\mathcal{G})))] > 0$. Now we prove the opposite direction. We prove it by contradiction. Assume that the union of all of the graphs in G is not strongly connected for all $x \in \Re^{mn}$. It is well-known that there exists a permutation matrix P such that for some $\tilde{x} \in \Re^{mn}$

$$P(\sum_{\mathcal{G}\in G} \mathcal{W}(\tilde{x},\mathcal{G}))P^T = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

where A and C are square matrices. Therefore, $spec(\sum_{\mathcal{G}\in G} \mathcal{W}(\tilde{x},\mathcal{G})) = spec(A) \cup spec(C)$. From Assumption 3.5, all columns of A has summation equal to $\bar{N}$, and all rows of C has summation equal to $\bar{N}$. Therefore, the eigenvalue $\lambda = \bar{N}$ has multiplicity 2 in $spec(\sum_{\mathcal{G}\in G} \mathcal{W}(\tilde{x},\mathcal{G}))$ which is a contradiction. Thus the proof is complete.

Now we show that the only solution of $W(x,\mathcal{G})x = x, \forall \mathcal{G} \in G$, with Assumptions 3.5 and 3.6 is the constraint $x_1 = x_2 = ... = x_m$.

We obtain from $W(x,\mathcal{G})x = x, \forall \mathcal{G} \in G$, that

$$(I_{mn} - W(x,\mathcal{G}))x = 0, \forall \mathcal{G} \in G,$$

which implies that

$$\sum_{\mathcal{G}\in G}(I_{mn} - W(x,\mathcal{G}))x = 0. \tag{3.10}$$

Now we have

$$\begin{aligned}\sum_{\mathcal{G}\in G}(I_{mn} - W(x,\mathcal{G})) &= \sum_{\mathcal{G}\in G}((I_m - \mathcal{W}(x,\mathcal{G})) \otimes I_n) \\ &= \Lambda(x) \otimes I_n,\end{aligned}$$

where

$$\Lambda(x) := \sum_{\mathcal{G}\in G}(I_m - \mathcal{W}(x,\mathcal{G})).$$

$\Lambda(x)$ has the following properties: the summation of all rows are equal to zero; the diagonal elements are non-negative; the off-diagonal elements are non-positive for all $x \in \Re^{mn}$. Therefore, $\Lambda(x)$ has the Laplacian matrix structure. Since $Re[\lambda_2(\Lambda(x))] > 0, \forall x \in \Re^{mn}$, (see the proof of Lemma 3.3), (3.10) implies that $x_1 = x_2 = \ldots = x_m$.

Therefore, distributed convex optimization with state-dependent interactions and time-varying topologies with Assumptions 3.5 and 3.6 can be formulated as

$$\begin{aligned} \min_x \quad & f(x) := \sum_{i=1}^{m} f_i(x_i) \\ \text{subject to} \quad & W(x, \mathcal{G})x = x. \quad \forall \mathcal{G} \in G \end{aligned} \tag{3.11}$$

In the remaining part of this section, we introduce a framework for modeling multi-agent optimization problems.

Problem stated by (3.11) can be reformulated as

$$\begin{aligned} \min_x \quad & f(x) := \sum_{i=1}^{m} f_i(x_i) \\ \text{subject to} \quad & x \in \bigcap_{\mathcal{G} \in G} Fix(T(x, \mathcal{G})) \end{aligned} \tag{3.12}$$

where $T(x, \mathcal{G}) := W(x, \mathcal{G})x$.

Definition 3.3: We call $T(x) := W(x)x$ *state-dependent weighted operator of the graph.*

Remark 3.7: The consensus subspace $\mathcal{C}$ (see Remark 3.4) is the intersection of fixed points sets of state-dependent weighted operators of the graphs with Assumption 3.6, i.e.,

$$\mathcal{C} = \bigcap_{\mathcal{G} \in G} Fix(T(x, \mathcal{G})).$$

CHAPTER 4. ALGORITHMS TO SOLVE THE OPTIMIZATION

In this chapter, we propose an algorithm, in the first section, to solve the optimization problem (3.1). An application of the algorithm is to solve distributed convex optimization over random networks with/without asynchronous protocols. In the later sections, we show that the random Picard and the random Krasnoselskii-Mann iterations are useful to solve feasibility problem of (3.1) under suitable assumptions.

Before we present the algorithm, we need to give some definitions needed for the next section. For simplicity, we write $\|.\|_{\mathcal{H}} = \|.\|$ in this chapter.

Definition 4.1: An operator $A : \mathcal{H} \longrightarrow \mathcal{H}$ is said to be *monotone* if

$$< x - y, Ax - Ay >\geq 0$$

for all $x, y \in \mathcal{H}$.

Definition 4.2: $A : \mathcal{H} \longrightarrow \mathcal{H}$ is called ξ-*strongly monotone* if

$$< x - y, Ax - Ay >\geq \xi \|x - y\|^2$$

for all $x, y \in \mathcal{H}$.

Remark 4.1: A function is ξ-strongly convex if its gradient is ξ-strongly monotone.

Definition 4.3: A mapping $A : \mathcal{H} \longrightarrow \mathcal{H}$ is said to be K-*Lipschitz continuous* if there exists a $K > 0$ such that

$$\|Ax - Ay\| \leq K\|x - y\|$$

for all $x, y \in \mathcal{H}$.

Definition 4.4: A sequence of random variables x_n is said to *converge pointwise (surely)* to x if for every $\omega \in \Omega$

$$\lim_{n\longrightarrow\infty} \|x_n(\omega) - x(\omega)\| = 0.$$

Definition 4.5: A sequence of random variables x_n is said to *converge almost surely* to x if there exists a subset $\Theta \subseteq \Omega$ such that $Pr(\Theta) = 0$, and for every $\omega \notin \Theta$

$$\lim_{n \longrightarrow \infty} \|x_n(\omega) - x(\omega)\| = 0.$$

Definition 4.6: A sequence of random variables x_n is said to *converge in mean square* to x if

$$E[\|x_n - x\|^2] \longrightarrow 0 \quad as \quad n \longrightarrow \infty.$$

4.1 A Proposed Algorithm

Problem (3.1) is a static optimization problem that requires potentially large number of constraints as the cardinality of Ω^* can be big. Moreover, explicit knowledge of $T(\omega^*, x), \forall \omega^* \in \Omega^*$, is in principle necessary to formulate the constraints.

We are interested in obtaining an iterative solution to Problem (3.1) where the constraint set is not known a priory and randomly changes over time. In other words, we solve Problem (3.1) iteratively such that one set of constraint appears at each time. In what follows, we assume that the constraint changes randomly over time.

We propose the following algorithm

$$x_{n+1} = \alpha_n(x_n - \beta \nabla f(x_n)) + (1 - \alpha_n)\hat{T}(\omega_n^*, x_n), \tag{4.1}$$

where $\hat{T}(\omega_n^*, x_n) := (1-\eta)x_n + \eta T(\omega_n^*, x_n), \eta \in (0, 1), \alpha_n \in [0, 1]$, and ω_n^* denotes an outcome $\omega^* \in \Omega^*$ at iteration n. The convergence of the algorithm is proved under the following assumption.

Assumption 4.1: $f(x)$ is ξ-strongly convex, and $\nabla f(x)$ is K-Lipschitz continuous.

Remark 4.2: A key distinguishing feature of Algorithm (4.1) is the presence of α_n in the second term. As we will see, this with Assumption 4.1 will introduce nice convergence properties.

4.1.1 Convergence Analysis

Consider a probability measure μ defined on the space $(\Omega, \mathcal{F})$ where

$$\Omega = \Omega^* \times \Omega^* \times \Omega^* \times \dots$$

$$\mathcal{F} = \sigma \times \sigma \times \sigma \times \dots$$

such that $(\Omega, \mathcal{F}, \mu)$ forms a probability space. We denote a realization in this probabilty space by $\omega \in \Omega$. We have the following assumption.

Assumption 4.2: There exists a nonempty subset $\tilde{K} \subseteq \Omega^*$ such that $FVP(T) = \{\tilde{z}|\tilde{z} \in \mathcal{H}, \tilde{z} = T(\bar{\omega}, \tilde{z}), \forall \bar{\omega} \in \tilde{K}\}$, and each element of $\tilde{K}$ occurs infinitely often almost surely.

Remark 4.3: If the sequence $\{\bar{\omega}_n\}_{n=0}^{\infty}$ is mutually independent with $\sum_{n=0}^{\infty} Pr_n(\bar{\omega}) = \infty$ where $Pr_n(\bar{\omega})$ is the probability of $\bar{\omega}$ occurring at time n, then according to Borel-Cantelli lemma [215], Assumption 4.2 is satisfied. Consequently, any i.i.d. random sequence satisfies Assumption 4.2. Any ergodic stationary sequences $\{\omega_n^*\}_{n=0}^{\infty}, Pr(\bar{\omega}) > 0$, satisfy Assumption 4.2 (see proof of Lemma 1 in [49]). Consequently, any time-invariant Markov chain with its unique stationary distribution as the initial distribution satisfies Assumption 4.2 (see [49]).

Lemma 4.1: Let $\hat{T}(\omega^*, x) := (1-\eta)x + \eta T(\omega^*, x), \omega^* \in \Omega^*, x \in \mathcal{H}$, with a nonexpansive random operator T, $FVP(T) \neq \emptyset$, and $\eta \in (0, 1]$. Then

(i) $FVP(T) = FVP(\hat{T})$.

(ii) $< x - \hat{T}(\omega^*, x), x - z > \geq \frac{\eta}{2}\|x - T(\omega^*, x)\|^2, \quad \forall z \in FVP(T), \forall \omega^* \in \Omega^*$.

(iii) $\hat{T}(\omega^*, x)$ is nonexpasnive.

Proof: (i)

Consider a $\hat{x} \in FVP(T)$. Thus $\hat{x} = T(\omega^*, \hat{x}), \forall \omega^* \in \Omega^*$. Hence

$$\hat{T}(\omega^*, \hat{x}) = (1-\eta)\hat{x} + \eta T(\omega^*, \hat{x}) = \hat{x}, \forall \omega^* \in \Omega^*,$$

which implies that $FVP(T) \subseteq FVP(\hat{T})$. Conversely, consider a $\hat{x} \in FVP(\hat{T})$. Indeed, $\hat{x} = \hat{T}(\omega^*, \hat{x}), \forall \omega^* \in \Omega^*$. Thus we have

$$\hat{x} = \hat{T}(\omega^*, \hat{x}) = (1-\eta)\hat{x} + \eta T(\omega^*, \hat{x}), \forall \omega^* \in \Omega^*,$$

or

$$\hat{x} = T(\omega^*, \hat{x}), \forall \omega^* \in \Omega^*,$$

which implies that $FVP(\hat{T}) \subseteq FVP(T)$. Therefore, we can conclude that $FVP(\hat{T}) = FVP(T)$. Thus the proof of part *(i)* of Lemma 4.1 is complete.

(ii)

We have from nonexpansivity of $T(\omega^*, x)$ that

$$\|T(\omega^*, x) - z\|^2 \leq \|x - z\|^2, \quad \forall z \in FVP(T), \forall \omega^* \in \Omega^*. \tag{4.2}$$

From the fact that

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2 < u, v >, \forall u, v \in \mathcal{H}, \tag{4.3}$$

we obtain for all $z \in FVP(T)$ and for all $\omega^* \in \Omega^*$ that

$$\begin{aligned} \|T(\omega^*, x) - z\|^2 &= \|T(\omega^*, x) - x + x - z\|^2 \\ &= \|T(\omega^*, x) - x\|^2 + \|x - z\|^2 \\ &\quad + 2 < T(\omega^*, x) - x, x - z > . \end{aligned} \tag{4.4}$$

Substituting (4.4) for (4.2) yields

$$2 < x - T(\omega^*, x), x - z > \geq \|T(\omega^*, x) - x\|^2. \tag{4.5}$$

From the definition of $\hat{T}(\omega_n^*, x_n)$, substituting $x - T(\omega^*, x) = \frac{x - \hat{T}(\omega^*, x)}{\eta}$ for the left hand side of the inequality (4.5) implies *(ii)*. Thus the proof of part *(ii)* of Lemma 4.1 is complete.

(iii)

We have from nonexpansivity of $T(\omega^*, x)$ for arbitrary $x, y \in \mathcal{H}$ that

$$\begin{aligned} \|\hat{T}(\omega^*, x) - \hat{T}(\omega^*, y)\| &\leq (1 - \eta)\|x - y\| + \eta\|T(\omega^*, x) - T(\omega^*, y)\| \\ &\leq (1 - \eta)\|x - y\| + \eta\|x - y\| \\ &= \|x - y\|, \forall \omega^* \in \Omega^*. \end{aligned}$$

Therefore, $\hat{T}(\omega^*, x)$ is a nonexpansive random operator, and the proof of part *(iii)* of Lemma 4.1 is complete.

4.1.1.1 Almost Sure Convergence

Theorem 4.1: Consider Problem (3.1) with Assumptions 4.1 and 4.2. Let $\beta \in (0, \frac{2\xi}{K^2})$ and $\alpha_n \in [0,1], n \in N \cup \{0\}$ such that

(a) $\lim_{n \longrightarrow \infty} \alpha_n = 0$,

(b) $\sum_{n=0}^{\infty} \alpha_n = \infty$.

Then starting from any initial point, the sequence generated by (4.1) globally converges almost surely to the unique solution of the problem.

Remark 4.4: An example of α_n satisfying *(a)* and *(b)* of Theorem 4.1 is $\alpha_n := \frac{1}{(1+n)^\zeta}$ where $\zeta \in (0,1]$.

Proof of Theorem 4.1:

We prove Theorem 4.1 in three steps:

Step 1: $\{x_n\}_{n=0}^{\infty}, \forall \omega \in \Omega$, is bounded.

Step 2: $\{x_n\}_{n=0}^{\infty}$ converges almost surely to a random variable supported by the feasible set.

Step 3: $\{x_n\}_{n=0}^{\infty}$ converges almost surely to the optimal solution.

Remark 4.5: The definition of *fixed value point* is a bridge from deterministic analysis to random analysis of the algorithm. With the help of fixed value point set and nonexpansivity property of the random operator $T(\omega^*, x)$, we are able to: first, prove boundedness of the generated sequence $\{x_n\}$ in a deterministic way in Step 1; second, extend deterministic tools to random cases such as part *(ii)* of Lemma 4.1 and use it for proving the convergence to the feasible set in Step 2; third, apply deterministic tools to proving the convergence to the optimal solution in Step 3. Therefore, the definition of fixed value point set with nonexpansivity property of $T(\omega^*, x)$ makes analysis of random processes easier than those of existing results regardless of switching distributions.

Now we give th proofs of Steps 1-3 in details.

Step 1: $\{x_n\}_{n=0}^{\infty}, \forall \omega \in \Omega$, *is bounded.*

Since the cost function is strongly convex and the constraint set is closed, the problem has the unique solution. Let x^* be the unique solution of the problem. Since x^* is the solution, we have that $x^* = \hat{T}(\omega_n^*, x^*), \forall \omega_n^* \in \Omega^*, \forall n \in N \cup \{0\}$ (see part *(i)* of Lemma 4.1). Also, we can write $x^* = \alpha_n x^* + (1-\alpha_n)\hat{T}(\omega_n^*, x^*), \forall \omega_n^* \in \Omega^*, \forall n \in N \cup \{0\}$. Therefore, we have

$$\begin{aligned}\|x_{n+1} - x^*\| &= \|\alpha_n(x_n - \beta\nabla f(x_n)) + (1-\alpha_n)\hat{T}(\omega_n^*, x_n) - x^*\| \\ &= \|\alpha_n(x_n - \beta\nabla f(x_n) - x^*) + (1-\alpha_n)(\hat{T}(\omega_n^*, x_n) - \hat{T}(\omega_n^*, x^*))\| \\ &\leq \alpha_n\|x_n - \beta\nabla f(x_n) - x^*\| + (1-\alpha_n)\|\hat{T}(\omega_n^*, x_n) - \hat{T}(\omega_n^*, x^*)\|.\end{aligned}$$

Since $\hat{T}(\omega^*, x)$ is a nonexpansive random operator (see part *(iii)* of Lemma 4.1), the above can be written as

$$\begin{aligned}\|x_{n+1} - x^*\| &\leq \alpha_n\|x_n - \beta\nabla f(x_n) - x^*\| + (1-\alpha_n)\|\hat{T}(\omega_n^*, x_n) - \hat{T}(\omega_n^*, x^*)\| \\ &\leq \alpha_n\|x_n - \beta\nabla f(x_n) - x^*\| + (1-\alpha_n)\|x_n - x^*\|. \end{aligned} \tag{4.6}$$

Since $\nabla f(x)$ is ξ-strongly monotone, and $\nabla f(x)$ is K-Lipschitz continuous, we obtain from (4.3) for any $x, y \in \mathcal{H}$ that

$$\begin{aligned}\|x - y - \beta(\nabla f(x) - \nabla f(y))\|^2 &= \|x-y\|^2 - 2\beta < \nabla f(x) - \nabla f(y), x - y > + \beta^2\|\nabla f(x) - \nabla f(y)\|^2 \\ &\leq \|x-y\|^2 - 2\xi\beta\|x-y\|^2 + K^2\beta^2\|x-y\|^2 \\ &= (1 - 2\xi\beta + \beta^2K^2)\|x-y\|^2 \\ &= (1-\gamma)^2\|x-y\|^2\end{aligned}$$

where $\gamma = 1 - \sqrt{1 - \beta(2\xi - \beta K^2)}$, and selecting $\beta \in (0, \frac{2\xi}{K^2})$ implies $0 < \gamma \leq 1$. Indeed, we have

$$\|x - y - \beta(\nabla f(x) - \nabla f(y))\| \leq (1-\gamma)\|x-y\|. \tag{4.7}$$

We have that

$$\begin{aligned}\|x_n - \beta\nabla f(x_n) - x^*\| &= \|x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*)) - \beta\nabla f(x^*)\| \\ &\leq \|x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*))\| + \beta\|\nabla f(x^*)\|. \end{aligned} \tag{4.8}$$

Therefore, (4.7) and (4.8) implies

$$\begin{aligned}\|x_n - \beta\nabla f(x_n) - x^*\| &\leq \|x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*))\| + \beta\|\nabla f(x^*)\| \\ &\leq (1-\gamma)\|x_n - x^*\| + \beta\|\nabla f(x^*)\|. \end{aligned} \tag{4.9}$$

Substituting (4.9) for (4.6) yields

$$\begin{aligned}\|x_{n+1} - x^*\| &\leq (1 - \gamma\alpha_n)\|x_n - x^*\| + \alpha_n\beta\|\nabla f(x^*)\| \\ &= (1 - \gamma\alpha_n)\|x_n - x^*\| + \gamma\alpha_n(\frac{\beta\|\nabla f(x^*)\|}{\gamma})\end{aligned}$$

which by induction implies that

$$\|x_{n+1} - x^*\| \leq max\{\|x_0 - x^*\|, \frac{\beta\|\nabla f(x^*)\|}{\gamma}\}$$

that implies $\|x_n - x^*\|, n \in N \cup \{0\}, \forall\omega \in \Omega$, is bounded. Therefore, $\{x_n\}_{n=0}^{\infty}$ is bounded for all $\omega \in \Omega$.

As seen from above, we proved the boundedness of the sequence with the help of fixed value point set and nonexpansiveness of $T(\omega^*, x)$ as well as Assumption 4.1.

Step 2: $\{x_n\}_{n=0}^{\infty}$ converges almost surely to a random variable supported by the feasible set.

From (4.1) and $x_n = \alpha_n x_n + (1 - \alpha_n)x_n$, we have

$$x_{n+1} - x_n + \alpha_n\beta\nabla f(x_n) = (1 - \alpha_n)(\hat{T}(\omega_n^*, x_n) - x_n), \tag{4.10}$$

and hence

$$< x_{n+1} - x_n + \alpha_n\beta\nabla f(x_n), x_n - x^* >= -(1 - \alpha_n) < x_n - \hat{T}(\omega_n^*, x_n), x_n - x^* > . \tag{4.11}$$

Since $x^* \in FVP(T)$, we have from part *(ii)* of Lemma 4.1 that

$$< x_n - \hat{T}(\omega_n^*, x_n), x_n - x^* >\geq \frac{\eta}{2}\|x_n - T(\omega_n^*, x_n)\|^2. \tag{4.12}$$

From (4.11) and (4.12), we obtain

$$< x_{n+1} - x_n + \alpha_n\beta\nabla f(x_n), x_n - x^* >\leq -\frac{\eta}{2}(1 - \alpha_n)\|x_n - T(\omega_n^*, x_n)\|^2 \tag{4.13}$$

or equivalently

$$- < x_n - x_{n+1}, x_n - x^* >\leq -\alpha_n < \beta\nabla f(x_n), x_n - x^* > -\frac{\eta}{2}(1 - \alpha_n)\|x_n - T(\omega_n^*, x_n)\|^2. \tag{4.14}$$

For any $u, v \in \mathcal{H}$ we have

$$< u, v >= -\frac{1}{2}\|u - v\|^2 + \frac{1}{2}\|u\|^2 + \frac{1}{2}\|v\|^2. \tag{4.15}$$

From (4.15) we obtain

$$< x_n - x_{n+1}, x_n - x^* >= -C_{n+1} + C_n + \frac{1}{2}\|x_n - x_{n+1}\|^2 \tag{4.16}$$

where $C_n = \frac{1}{2}\|x_n - x^*\|^2$. From (4.14) and (4.16) we obtain

$$C_{n+1} - C_n - \frac{1}{2}\|x_n - x_{n+1}\|^2 \leq -\alpha_n < \beta\nabla f(x_n), x_n - x^* > -\frac{\eta}{2}(1-\alpha_n)\|x_n - T(\omega_n^*, x_n)\|^2. \tag{4.17}$$

From (4.10) and (4.3) we have

$$\begin{aligned} \|x_{n+1} - x_n\|^2 &= \| - \alpha_n\beta\nabla f(x_n) + (1-\alpha_n)(\hat{T}(\omega_n^*, x_n) - x_n)\|^2 \\ &= \alpha_n^2\|\beta\nabla f(x_n)\|^2 + (1-\alpha_n)^2\|\hat{T}(\omega_n^*, x_n) - x_n\|^2 \\ &\quad - 2\alpha_n(1-\alpha_n) < \beta\nabla f(x_n), \hat{T}(\omega_n^*, x_n) - x_n > . \end{aligned} \tag{4.18}$$

We know that $\|\hat{T}(\omega_n^*, x_n) - x_n\| = \eta\|x_n - T(\omega_n^*, x_n)\|$. Since $\alpha_n \in [0,1]$, we have also that $(1-\alpha_n)^2 \leq (1-\alpha_n)$. Using these facts as well as multiplying both sides of (4.18) by $\frac{1}{2}$ yield

$$\begin{aligned} \frac{1}{2}\|x_{n+1} - x_n\|^2 &= \frac{1}{2}\alpha_n^2\|\beta\nabla f(x_n)\|^2 + \frac{1}{2}(1-\alpha_n)^2\eta^2\|T(\omega_n^*, x_n) - x_n\|^2 \\ &\quad - \alpha_n(1-\alpha_n) < \beta\nabla f(x_n), \hat{T}(\omega_n^*, x_n) - x_n > . \\ &\leq \frac{1}{2}\alpha_n^2\|\beta\nabla f(x_n)\|^2 + \frac{1}{2}(1-\alpha_n)\eta^2\|T(\omega_n^*, x_n) - x_n\|^2 \\ &\quad - \alpha_n(1-\alpha_n) < \beta\nabla f(x_n), \hat{T}(\omega_n^*, x_n) - x_n > . \end{aligned} \tag{4.19}$$

From (4.17) and (4.19), we obtain

$$\begin{aligned} C_{n+1} - C_n &\leq \frac{1}{2}\|x_{n+1} - x_n\|^2 - \alpha_n < \beta\nabla f(x_n), x_n - x^* > \\ &\quad - \frac{\eta}{2}(1-\alpha_n)\|x_n - T(\omega_n^*, x_n)\|^2 \\ &\leq -(\frac{1}{2} - \frac{\eta}{2})\eta(1-\alpha_n)\|x_n - T(\omega_n^*, x_n)\|^2 + \alpha_n(\frac{1}{2}\alpha_n\|\beta\nabla f(x_n)\|^2 \\ &\quad - < \beta\nabla f(x_n), x_n - x^* > \\ &\quad - (1-\alpha_n) < \beta\nabla f(x_n), \hat{T}(\omega_n^*, x_n) - x_n >). \end{aligned} \tag{4.20}$$

Now we claim that there exists an $n_0 \in N$ such that the sequence $\{C_n\}$ is non-increasing for $n \geq n_0$. Assume by contradiction that this is not true. Then there exists a subsequence $\{C_{n_j}\}$ such that

$$C_{n_j+1} - C_{n_j} > 0$$

which together with (4.20) yields

$$\begin{aligned} 0 &< C_{n_j+1} - C_{n_j} \\ &\leq -(\frac{1}{2} - \frac{\eta}{2})\eta(1-\alpha_{n_j})\|x_{n_j} - T(\omega^*_{n_j}, x_{n_j})\|^2 \\ &\quad + \alpha_{n_j}(\frac{1}{2}\alpha_{n_j}\beta^2\|\nabla f(x_{n_j})\|^2 - < \beta\nabla f(x_{n_j}), x_{n_j} - x^* > \\ &\quad - (1-\alpha_{n_j}) < \beta\nabla f(x_{n_j}), \hat{T}(\omega^*_{n_j}, x_{n_j}) - x_{n_j} >). \end{aligned} \tag{4.21}$$

Since $\{x_n\}$ is bounded, $\nabla f(x)$ is continuous, and $\eta \in (0,1)$, we obtain from (4.21) by Theorem 4.1 *(a)* that

$$\begin{aligned} 0 &< \liminf_{j\longrightarrow\infty}[-(\frac{1}{2} - \frac{\eta}{2})\eta(1-\alpha_{n_j})\|x_{n_j} - T(\omega^*_{n_j}, x_{n_j})\|^2 \\ &\quad + \alpha_{n_j}(\frac{1}{2}\alpha_{n_j}\|\beta\nabla f(x_{n_j})\|^2 - < \beta\nabla f(x_{n_j}), x_{n_j} - x^* > \\ &\quad - (1-\alpha_{n_j}) < \beta\nabla f(x_{n_j}), \hat{T}(\omega^*_{n_j}, x_{n_j}) - x_{n_j} >)] \\ &\leq 0 \end{aligned} \tag{4.22}$$

which is a contradiction. Therefore, there exists an $n_0 \in N$ such that the sequence $\{C_n\}$ is non-increasing for $n \geq n_0$. Since $\{C_n\}$ is bounded below, it converges for all $\omega \in \Omega$.

Taking the limit of both sides of (4.20) and using the convergence of $\{C_n\}$, continuity of $\nabla f(x)$, Step 1, $\eta \in (0,1)$, and Theorem 4.1 *(a)* yield

$$\lim_{n\longrightarrow\infty}\|x_n - T(\omega^*_n, x_n)\| = 0, \quad \textit{pointwise}$$

which implies that $\{x_n\}_{n=0}^{\infty}$ converges for each $\omega \in \Omega$ since $FVP(T) \neq \emptyset$. Moreover, this together with Assumption 4.2 implies that $\{x_n\}$ converges almost surely to a random variable supported by $FVP(T)$.

As seen from above, we proved the convergence to the feasible set in a deterministic way with the help of fixed value point set and nonexpansivity of $T(\omega^*, x)$ as well as Lemma 4.1.

Step 3: $\{x_n\}_{n=0}^{\infty}$ *converges almost surely to the optimal solution.*

It remains to prove that $\{x_n\}_{n=0}^{\infty}$ converges almost surely to the optimal solution. Since $x^* \in FVP(T)$ is the optimal solution, we have

$$< \bar{x} - x^*, \nabla f(x^*) > \geq 0, \forall \bar{x} \in FVP(T). \tag{4.23}$$

We have from (4.3) that

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &= \|x_{n+1} - x^* + \alpha_n \beta \nabla f(x^*) - \alpha_n \beta \nabla f(x^*)\|^2 \\ &= \|x_{n+1} - x^* + \alpha_n \beta \nabla f(x^*)\|^2 + \alpha_n^2 \|\beta \nabla f(x^*)\|^2 \\ &\quad - 2\alpha_n < \beta \nabla f(x^*), x_{n+1} - x^* + \alpha_n \beta \nabla f(x^*) > . \end{aligned} \tag{4.24}$$

Since $x^* = \hat{T}(\omega_n^*, x^*), \forall \omega_n^* \in \Omega^*, \forall n \in N \cup \{0\}$, we have that $x^* = \alpha_n x^* + (1 - \alpha_n)\hat{T}(\omega_n^*, x^*), \forall \omega_n^* \in \Omega^*, \forall n \in N \cup \{0\}$; using this fact and (4.1), we obtain

$$\begin{aligned} \|x_{n+1} - x^* + \alpha_n \beta \nabla f(x^*)\|^2 &= \|\alpha_n [x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*))] \\ &\quad + (1 - \alpha_n)[\hat{T}(\omega_n^*, x_n) - \hat{T}(\omega_n^*, x^*)]\|^2. \end{aligned} \tag{4.25}$$

Furthermore, we have

$$\begin{aligned} < \beta \nabla f(x^*), x_{n+1} - x^* + \alpha_n \beta \nabla f(x^*) > &=< \beta \nabla f(x^*), x_{n+1} - x^* > + \alpha_n < \beta \nabla f(x^*), \beta \nabla f(x^*) > \\ &=< \beta \nabla f(x^*), x_{n+1} - x^* > + \alpha_n \|\beta \nabla f(x^*)\|^2. \end{aligned} \tag{4.26}$$

Substituting (4.25) and (4.26) for (4.24) yields

$$\begin{aligned}\|x_{n+1}-x^*\|^2 &= \|x_{n+1}-x^*+\alpha_n\beta\nabla f(x^*)\|^2+\alpha_n^2\|\beta\nabla f(x^*)\|^2\\&\quad -2\alpha_n < \beta\nabla f(x^*), x_{n+1}-x^*+\alpha_n\beta\nabla f(x^*) >\\&= \|\alpha_n[x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*))]+(1-\alpha_n)[\hat{T}(\omega_n^*,x_n)-\hat{T}(\omega_n^*,x^*)]\|^2\\&\quad -2\alpha_n < \beta\nabla f(x^*), x_{n+1}-x^* > -\alpha_n^2\|\beta\nabla f(x^*)\|^2\\&= \alpha_n^2\|x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*))\|^2\\&\quad +(1-\alpha_n)^2\|\hat{T}(\omega_n^*,x_n)-\hat{T}(\omega_n^*,x^*)\|^2\\&\quad +2\alpha_n(1-\alpha_n) < x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*)), \hat{T}(\omega_n^*,x_n)-\hat{T}(\omega_n^*,x^*) >\\&\quad -2\alpha_n < \beta\nabla f(x^*), x_{n+1}-x^* > -\alpha_n^2\|\beta\nabla f(x^*)\|^2.\end{aligned}$$

From (4.7), nonexpansivity property of $\hat{T}(\omega^*, x)$, and Cauchy–Schwarz inequality, we obtain

$$< x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*)), \hat{T}(\omega_n^*,x_n)-\hat{T}(\omega_n^*,x^*) >\leq (1-\gamma)\|x_n-x^*\|^2. \tag{4.27}$$

From (4.7), we also obtain

$$\|x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*))\|^2 \leq (1-\gamma)^2\|x_n-x^*\|^2. \tag{4.28}$$

Therefore, from (4.27), (4.28), and nonexpansivity property of $\hat{T}(\omega^*, x)$, we have

$$\begin{aligned}\|x_{n+1}-x^*\|^2 &= \alpha_n^2\|x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*))\|^2\\&\quad +(1-\alpha_n)^2\|\hat{T}(\omega_n^*,x_n)-\hat{T}(\omega_n^*,x^*)\|^2\\&\quad +2\alpha_n(1-\alpha_n) < x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*)), \hat{T}(\omega_n^*,x_n)-\hat{T}(\omega_n^*,x^*) >\\&\quad -2\alpha_n < \beta\nabla f(x^*), x_{n+1}-x^* > -\alpha_n^2\|\beta\nabla f(x^*)\|^2\\&\leq (1-2\gamma\alpha_n)\|x_n-x^*\|^2+\alpha_n(\gamma^2\alpha_n\|x_n-x^*\|^2-2 < \beta\nabla f(x^*), x_{n+1}-x^* >)\\&= (1-\gamma\alpha_n)\|x_n-x^*\|^2-\gamma\alpha_n\|x_n-x^*\|^2\\&\quad +\alpha_n(\gamma^2\alpha_n\|x_n-x^*\|^2-2 < \beta\nabla f(x^*), x_{n+1}-x^* >).\\&\leq (1-\gamma\alpha_n)\|x_n-x^*\|^2\\&\quad +\alpha_n(\gamma^2\alpha_n\|x_n-x^*\|^2-2 < \beta\nabla f(x^*), x_{n+1}-x^* >)\end{aligned}$$

or, finally,

$$\|x_{n+1} - x^*\|^2 \leq (1 - \gamma\alpha_n)\|x_n - x^*\|^2 + \gamma\alpha_n(\frac{\gamma^2\alpha_n\|x_n - x^*\|^2 - 2 < \beta\nabla f(x^*), x_{n+1} - x^* >}{\gamma}), \tag{4.29}$$

From Step 1, Step 2, (4.23), and Theorem 4.1 *(a)*, we obtain

$$\lim_{n\longrightarrow\infty}(\gamma^2\alpha_n\|x_n - x^*\|^2 - 2\beta < \nabla f(x^*), x_{n+1} - x^* >) \leq 0 \quad \textit{almost surely.} \tag{4.30}$$

Now we have the following lemma.

Lemma 4.2 [216]: Let $\{a_n\}_{n=0}^{\infty}$ be a sequence of non-negative real numbers satisfying

$$a_{n+1} \leq (1 - b_n)a_n + b_n h_n + c_n$$

where $b_n \in [0, 1], \sum_{n=0}^{\infty} b_n = \infty, \limsup_{n\longrightarrow\infty} h_n \leq 0$, and $\sum_{n=0}^{\infty} c_n < \infty$. Then

$$\lim_{n\longrightarrow\infty} a_n = 0.$$

According to Lemma 4.2 by setting

$$a_n = \|x_n - x^*\|^2,$$

$$b_n = \gamma\alpha_n,$$

$$h_n = (\frac{\gamma^2\alpha_n\|x_n - x^*\|^2 - 2\beta < \nabla f(x^*), x_{n+1} - x^* >}{\gamma}),$$

we obtain from (4.29), (4.30), and Theorem 4.1 *(b)* that

$$\lim_{n\longrightarrow\infty} \|x_n - x^*\|^2 = 0 \quad \textit{almost surely.}$$

Therefore, $\{x_n\}_{n=0}^{\infty}$ converges almost surely to x^*.

As seen from above, we proved the convergence to the optimal solution in a deterministic way (using Lemma 4.2) with the help of fixed value point set and nonexpansivity of $T(\omega^*, x)$ as well as Assumption 4.1.

4.1.1.2 Mean Square Convergence

Theorem 4.2: Consider Problem (3.1) with Assumptions 4.1 and 4.2. Suppose that $\beta \in (0, \frac{2\xi}{K^2})$ and $\alpha_n \in [0,1], n \in N \cup \{0\}$, satisfies *(a)* and *(b)* of Theorem 4.1. Then starting from any initial point, the sequence generated by (4.1) globally converges in mean square to the unique solution of the problem.

Proof: We have from Theorem 4.1 that

$$\lim_{n \longrightarrow \infty} \|x_n - x^*\| = 0 \quad \textit{almost surely},$$

or

$$\lim_{n \longrightarrow \infty} \|x_n - x^*\|^2 = 0 \quad \textit{almost surely}.$$

From Parallelogram Law, we have that

$$\|x_n - x^*\|^2 \leq 2(\|x_n\|^2 + \|x^*\|^2), \forall n \in N.$$

We define a non-negative measurable function

$$\tau_n = 2(\|x_n\|^2 + \|x^*\|^2) - \|x_n - x^*\|^2.$$

Hence, we obtain

$$\lim_{n \longrightarrow \infty} \tau_n = 4\|x^*\|^2 \quad \textit{almost surely}.$$

Now we have the following lemma.

Lemma 4.3 [217] *(Fatou's Lemma)*: If $\tau_n : \Omega \longrightarrow [0, \infty]$ is measurable, for each positive integer n, then

$$\int_\Omega (\liminf_{n \longrightarrow \infty} \tau_n) d\mu \leq \liminf_{n \longrightarrow \infty} \int_\Omega \tau_n d\mu.$$

Applying Lemma 4.3 yields

$$\int_\Omega (\liminf_{n \longrightarrow \infty} \tau_n) d\mu \leq \liminf_{n \longrightarrow \infty} \int_\Omega \tau_n d\mu$$

or

$$\begin{aligned} \int_\Omega 4\|x^*\|^2 d\mu &\leq \liminf_{n \longrightarrow \infty} (\int_\Omega 2\|x_n\|^2 d\mu + \int_\Omega 2\|x^*\|^2 d\mu - \int_\Omega \|x_n - x^*\|^2 d\mu) \\ &= \lim_{n \longrightarrow \infty} (\int_\Omega 2\|x_n\|^2 d\mu) + \int_\Omega 2\|x^*\|^2 d\mu - \limsup_{n \longrightarrow \infty} \int_\Omega \|x_n - x^*\|^2 d\mu. \qquad (4.31) \end{aligned}$$

Now we have the following lemma.

Lemma 4.4 [217] *(The Dominated Convergence Theorem)*: Let $\{\tau_n\}$ be a sequence in L^1 such that $\tau_n \longrightarrow \tau$ almost everywhere, and there exists a nonnegative $g \in L^1$ such that $|\tau_n| \leq g$ for all n. Then, $\tau \in L^1$ and $\int_\Omega \tau d\mu = \lim_{n\longrightarrow\infty} \int_\Omega \tau_n d\mu$.

Due to boundedness of $\{x_n\}_{n=0}^{\infty}, \forall\omega \in \Omega$, we obtain from Lemma 4.4 that

$$\lim_{n\longrightarrow\infty} \int_\Omega 2\|x_n\|^2 d\mu = \int_\Omega 2\|x^*\|^2 d\mu. \tag{4.32}$$

Thus, we obtain from (4.31) and (4.32) that

$$\begin{aligned}\int_\Omega 4\|x^*\|^2 d\mu &\leq \lim_{n\longrightarrow\infty} (\int_\Omega 2\|x_n\|^2 d\mu) + \int_\Omega 2\|x^*\|^2 d\mu - \limsup_{n\longrightarrow\infty} \int_\Omega \|x_n - x^*\|^2 d\mu \\ &= \int_\Omega 4\|x^*\|^2 d\mu - \limsup_{n\longrightarrow\infty} \int_\Omega \|x_n - x^*\|^2 d\mu,\end{aligned}$$

or

$$\limsup_{n\longrightarrow\infty} \int_\Omega \|x_n - x^*\|^2 d\mu = 0.$$

Thus we have

$$\begin{aligned}\lim_{n\longrightarrow\infty} E[\|x_n - x^*\|^2] &= \lim_{n\longrightarrow\infty} \int_\Omega \|x_n - x^*\|^2 d\mu \\ &\leq \limsup_{n\longrightarrow\infty} \int_\Omega \|x_n - x^*\|^2 d\mu \\ &= 0\end{aligned}$$

which implies that $\{x_n\}_{n=0}^{\infty}$ converges in mean square to x^*. Thus the proof of Theorem 4.2 is complete.

4.1.2 Application to Solve Distributed Optimization over Random Networks

So far we have provided the convergence of Algorithm (4.1) to the optimal solution of Problem (3.1). The algorithm can directly be applied to solving Problem 3.2 in a distributed fashion under the following considerations. We need to assume that each $f_i(x_i)$ is ξ-strongly convex and $\nabla f_i(x_i)$ is K-Lipschitz. Therefore, we arrive at the following corollaries of Theorems 4.1 and 4.2, respectively.

Corollary 4.1: Consider Problem 3.2 with Assumption 4.2. Assume that $f_i(x_i)$ is ξ-strongly convex, and $\nabla f_i(x_i)$ is K-Lipschitz continuous for $i = 1, 2, ..., m$. Suppose that $\beta \in (0, \frac{2\xi}{K^2}), \eta \in (0, 1)$, and $\alpha_n \in [0, 1], n \in N \cup \{0\}$, satisfies *(a)* and *(b)* of Theorem 4.1. Then starting from any initial point, the sequence generated by

$$x_{n+1} = \alpha_n(x_n - \beta \nabla f(x_n)) + (1 - \alpha_n)((1 - \eta)x_n + \eta W(\omega_n^*)x_n) \tag{4.33}$$

globally converges almost surely to the unique solution of the problem.

Corollary 4.2: Consider Problem 3.2 with Assumption 4.2. Assume that $f_i(x_i)$ is ξ-strongly convex, and $\nabla f_i(x_i)$ is K-Lipschitz continuous for $i = 1, 2, ..., m$. Suppose that $\beta \in (0, \frac{2\xi}{K^2}), \eta \in (0, 1)$, and $\alpha_n \in [0, 1], n \in N \cup \{0\}$, satisfies *(a)* and *(b)* of Theorem 4.1. Then starting from any initial point, the sequence generated by (4.33) globally converges in mean square to the unique solution of the problem.

Remark 4.6: The authors in [8] have presented a totally asynchronous algorithm for solving systems of equations of the form $x = f(x)$ where $f(x)$ is a contraction mapping on the Banach space $\mathcal{B}_\infty = (\Re^n, \|.\|_\infty)$ and a partially asynchronous algorithm for solving consensus system $x = Wx$ where W is nonexpansive on $\mathcal{B}_\infty$. Here, we are able to obtain a totally asynchronous distributed algorithm for solving distributed optimization problems (rather than systems of equations) constrained by the consensus system in the Hilbert space $\mathcal{H} = (\Re^n, \|.\|_2)$. Note that nonexpansivity (or contraction) property of an operator in general may not be preserved from a space to another.

4.1.2.1 Numerical Example

Now we give an instance of a distributed optimization problem over a random network in which there are distribution dependencies among communication graphs.

Example 4.1: Distributed estimation in wireless sensor networks (WSNs):

Consider a WSN with $m = 20$ sensors which measure the location of an object. The observation of the i^{th} sensor is described as $y_i = A_i\theta + \nu_i$ where $A_i \in \Re^{d\times n}$, $\theta \in \Re^n$ is the deterministic parameter to be estimated, and ν_i is the i.i.d. Gaussian observation noise. We use Maximum

Likelihood Estimator with regularization

$$\min_{\theta} \sum_{i=1}^{20} (\|A_i\theta - y_i\|_2^2 + \rho_i \|\theta\|_2^2),$$

where ρ_i is the regularization parameter of each sensor.

This problem can be formulated as the following distributed problem:

$$\begin{aligned} \min_{\theta_1,\dots,\theta_{20}} \quad & \sum_{i=1}^{20} (\|A_i\theta_i - y_i\|_2^2 + \rho_i \|\theta_i\|_2^2) \\ \text{subject to} \quad & \theta_1 = \theta_2 = \dots = \theta_{20}. \end{aligned} \tag{4.34}$$

We consider an undirected graph, i.e., $1 \longleftrightarrow 2 \ldots \longleftrightarrow 20$. Each sensor gives a weight $\mathcal{W}_{ij} = \frac{1}{|\mathcal{N}_i \cup \{i\}|}$ to information received from its neighbors. We select $y_i = [0.25, 0.25][2, 2]^T + \nu_i$ and $\rho_i = 0.2$ where ν_i is the i.i.d. Gaussian observation noise with zero mean and variance 0.01 for each sensor's measurement. One can see that $f_i(\theta_i) = \|A_i\theta_i - y_i\|_2^2, i = 1, 2, ..., 20$, are $2\rho_i$-strongly convex, and $\nabla f_i(\theta_i)$ are K_i-Lipschitz continuous where

$$K_i = \|2A_i^T A_i + 2\rho_i I_2\|_2 = 0.65.$$

Hence, $\xi = min\{2\rho_1, \ldots, 2\rho_{20}\} = 0.4$ and $K = max\{K_1, \ldots, K_{20}\} = 0.65$; consequently, we select $\beta = 1 \in (0, \frac{2\xi}{K^2})$. Also we select $\alpha_n = \frac{1}{n+1}, n \geq 0$ and $\eta = 0.5$ for simulation.

Here, $\Omega^* = \{\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_4\}$ where

$$\begin{aligned} \mathcal{G}_1 &= \{(2,3), (4,5), (6,7), (8,9)\}, \\ \mathcal{G}_2 &= \{(10,11), (12,13), (14,15), (16,17), (18,19)\}, \\ \mathcal{G}_3 &= \{(1,2), (3,4), (5,6), (7,8), (9,10), (11,12), (13,14), \\ &\quad (15,16), (17,18), (19,20)\}, \\ \mathcal{G}_4 &= \{\}. \end{aligned}$$

We assume that $\mathcal{G}_i, i = 1, 2, 3, 4$, have i.i.d. Bernoulli distribution with $Pr(\mathcal{G}_i) = \frac{1}{4}$ in every $\hat{N}-$interval, and at the iteration $k\hat{N}, k = 1, 2, \ldots$, a graph works that has worked the minimum number of times in the previous $\hat{N}-$interval. If some graphs $\mathcal{G}_i$ have the same number of minimum

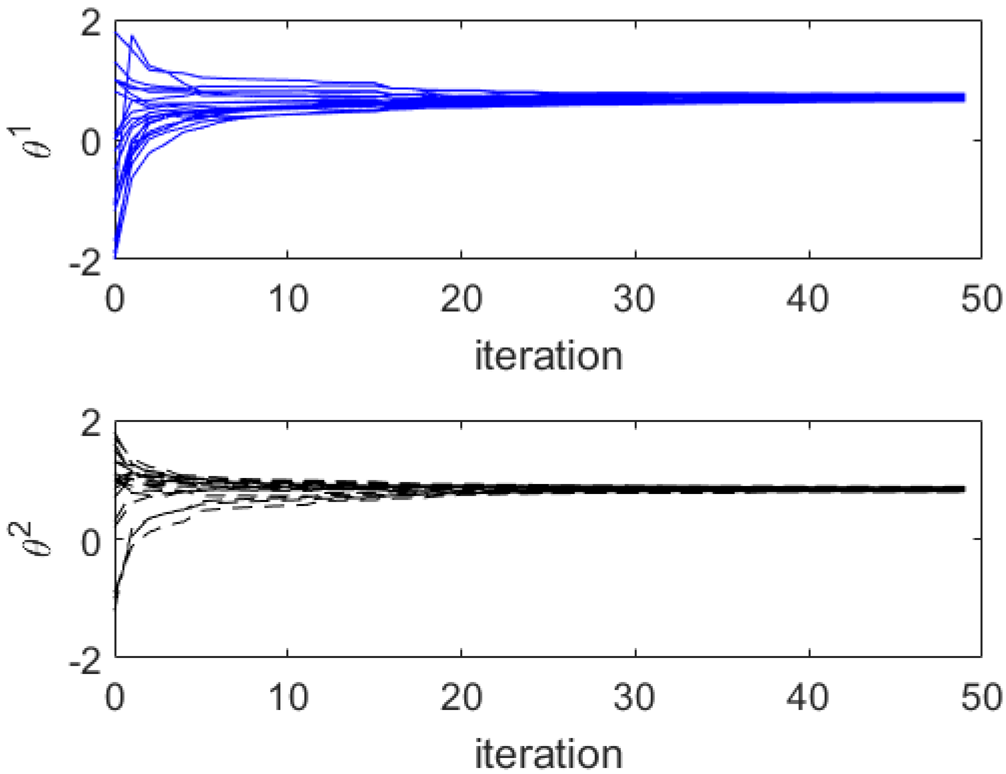

Figure 4.1 Variables θ^1 and θ^2 of 20 agents are shown by solid blue lines and dashed black lines, respectively. The figures show that they are approaching $\theta^* = [0.7417, 0.7417]^T$.

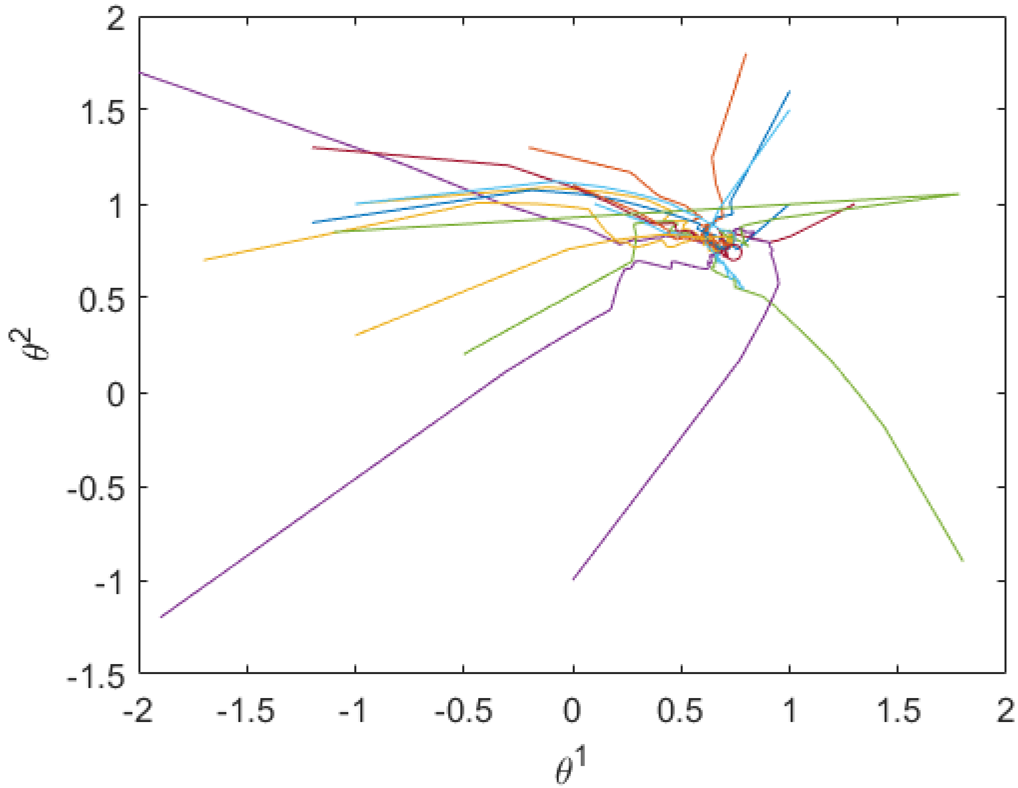

Figure 4.2 2D plot, where θ^* is shown by o, for 1000 iterations.

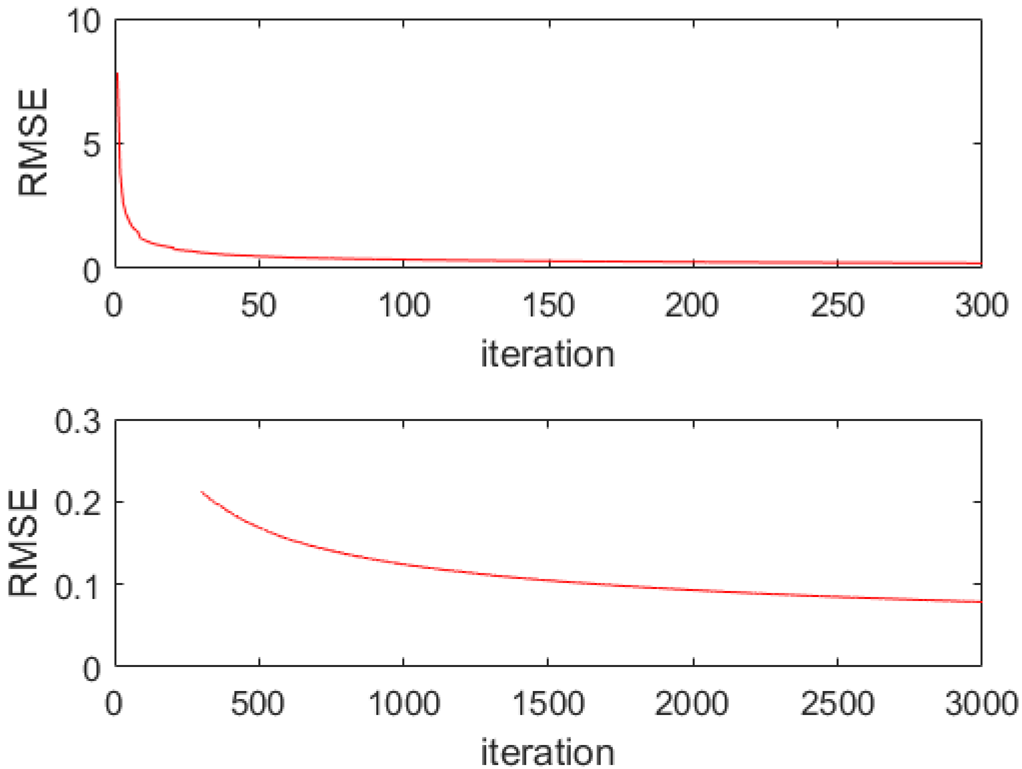

Figure 4.3 Root Mean Square Error (RMSE) for two intervals: $[0, 300]$ and $[301, 3000]$ iterations.

occurrences in the previous $\hat{N}-$interval, then one is chosen randomly. Thus the sequence $\{\omega_n^*\}_{n=0}^{\infty}$ is not independent and has time-varying distributions. In fact, it has a subsequence $\{\omega_{n_j}^*\}_{j=0}^{\infty}$ that is i.i.d. As a matter of fact, according to Borel-Cantelli lemma [215], $\mathcal{G}_i, i = 1,2,3,4$, occur infinitely often almost surely in the probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$ where

$$\bar{\Omega} = \{\mathcal{G}_1, \ldots, \mathcal{G}_4\} \times \{\mathcal{G}_1, \ldots, \mathcal{G}_4\} \times \{\mathcal{G}_1, \ldots, \mathcal{G}_4\} \times ...$$

Therefore, $\mathcal{G}_1, \ldots, \mathcal{G}_4$ occur infinitely often almost surely in the probability space $(\Omega, \mathcal{F}, \mu)$ in this example. Therefore, Assumption 4.2 is satisfied. Indeed, the conditions of Corollaries 4.1 and 4.2 are satisfied. We choose $\hat{N} = 20$ and random initial conditions for simulation. The results given by Algorithm (4.33) are shown in Figures 4.1.

We use CVX software of Matlab for solving optimization problem (4.34) and the solution is $\theta_i^* = [0.7417, 0.7417]^T, i = 1, \ldots, 20$. Note that θ_i^* may be different due to different observation noise. The two-dimensional plot is given in Figure 4.2, and the error $e_n = \|x_n - x^*\|_2$ is shown in Figure 4.3

As seen in this example, Algorithm (4.33) is able to solve distributed optimization problems in which there are distribution dependencies among possible graphs under mentioned assumptions.

4.2 The Random Picard Algorithm

Although the Picard iterative algorithm may not always converge to a fixed point of an operator (see Chapter 2), it converges for operators with special properties. This is useful for solving feasibility problem of (3.1). In the following two subsections, we show that the random Picard iteration can solve feasibility problems under suitable conditions.

4.2.1 Firmly Nonexpansive Random Maps

Consider a firmly nonexpansive random mapping $T(\omega^*, x)$ where $T : \Omega^* \times \Re^n \longrightarrow \Re^n, n \in N$, and $\mathcal{H} = (\Re^n, \|.\|_{\mathcal{H}})$. Now we have the following theorem.

Theorem 4.3: Consider the above firmly nonexpansive random map $T(\omega^*, x)$ where the cardinality of the set Ω^* is finite. Assume $FVP(T) \neq \emptyset$. If each $\omega^* \in \Omega^*$ occurs infinitely often almost surely, then the sequence $\{x_n\}_{n=0}^{\infty}$ generated by the random Picard iteration

$$x_{n+1} = T(\omega_n^*, x_n) \tag{4.35}$$

converges almost surely and in mean square to a random variable supported by $FVP(T)$.

Proof: We introduce the following lemma.

Lemma 4.5 [218]: Let $\phi_i : \mathcal{H} \longrightarrow \mathcal{H}, i = 1, 2, ..., \tilde{N}$, be firmly nonexpansive with $\cap_{i=1}^{\tilde{N}} Fix(\phi_i) \neq \emptyset$, where $\mathcal{H}$ is finite dimensional. Then the random sequence generated by

$$x_0 \in D \quad \text{arbitrary}, \quad x_{n+1} = \phi_{r(n)}(x_n), n \geq 0, \tag{4.36}$$

where each element of $\{1, ..., \tilde{N}\}$ appears in the sequence $\{r(0), r(1), ...\}$ an infinite number of times, converges to some point in $\cap_{i=1}^{\tilde{N}} Fix(\phi_i)$.

If each $\omega^* \in \Omega^*$ occurs infinitely often almost surely, we obtain from Lemma 4.5 that the sequence $\{x_n\}_{n=0}^{\infty}$ generated by (4.35) converges almost surely to a random variable supported by $FVP(T)$. From the proof of Theorem 4.2, the sequence also converges in mean square to the random variable. Thus the proof is complete.

4.2.2 Contraction Random Maps

Theorem 4.4: Consider a random operator $T(\omega^*, x)$ where $T : \Omega^* \times \mathcal{B} \longrightarrow \mathcal{B}, FVP(T) \neq \emptyset$, and T is a contraction random operator with constant $0 \leq \kappa < 1$. Then starting from any initial point, the sequence generated by the random Picard iteration (4.35) converges pointwise (surely) and in mean square to the solution of the problem with exponential rate of convergence.

Proof: We have that for each fixed $\omega^* \in \Omega^*$, the operator $T(\omega^*, x)$ is a contraction with constant κ. Thus, according to Theorem 2.3, it has a unique fixed point for each fixed $\omega^* \in \Omega^*$. Since $FVP(T) \neq \emptyset$, there exists a unique x^* such that $x^* = T(\omega^*, x^*), \forall \omega^* \in \Omega^*$. Hence, we obtain

$$\begin{aligned} \|x_{n+1} - x^*\|_{\mathcal{B}} &\leq \kappa \|x_n - x^*\|_{\mathcal{B}} \\ &\leq \kappa^2 \|x_{n-1} - x^*\|_{\mathcal{B}} \\ &\vdots \\ &\leq \kappa^{n+1} \|x_0 - x^*\|_{\mathcal{B}}, \quad \forall \omega \in \Omega. \end{aligned} \tag{4.37}$$

Now we show that the sequence $\{x_n\}_{n=0}^{\infty}$ is a Cauchy sequence in $\mathcal{B}$. We obtain by (4.37) that

$$\begin{aligned}\|x_{n+1} - x_n\|_{\mathcal{B}} &= \|x_{n+1} - x_n + x^* - x^*\|_{\mathcal{B}} \\ &\leq \|x_{n+1} - x^*\|_{\mathcal{B}} + \|x_n - x^*\|_{\mathcal{B}} \\ &\leq \kappa^{n+1}\|x_0 - x^*\|_{\mathcal{B}} + \kappa^n\|x_0 - x^*\|_{\mathcal{B}}, \quad \forall\omega \in \Omega.\end{aligned}$$

Therefore, $\{x_n\}_{n=0}^{\infty}$ is a Caushy sequence in $\mathcal{B}$ and, thus, converges pointwise (surely) to x^*. We also obtain from (4.37) that

$$\begin{aligned}\lim_{n\longrightarrow\infty} E[\|x_n - x^*\|_{\mathcal{B}}^2] &= \lim_{n\longrightarrow\infty} \int_{\Omega} \|x_n - x^*\|_{\mathcal{B}}^2 \\ &\leq \lim_{n\longrightarrow\infty} \kappa^{2n}\|x_0 - x^*\|_{\mathcal{B}}^2\mu(\Omega) \\ &= 0.\end{aligned}$$

Therefore, $\{x_n\}_{n=0}^{\infty}$ converges in mean square to x^*. One can see from (4.37) that the rate of convergence is exponential. Thus the proof of Theorem 4.4 is complete.

4.3 The Random Krasnoselskii-Mann Algorithm

In some cases when the Picard iteration may not converge, the Krasnoselskii-Mann iteration may be useful to solve a problem. In the following subsection, we show that the random Krasnoselskii-Mann iterative algorithm is useful to solve feasibility problem of (3.1).

4.3.1 Nonexpansive Random Maps

Consider a nonexpansive random map $T(\omega^*, x)$ where $T : \Omega^* \times \Re^n \longrightarrow \Re^n, n \in N$, and $\mathcal{H} = (\Re^n, \|.\|_{\mathcal{H}})$.

Theorem 4.5: Consider the above nonexpansive random map. Let the cardinality of the set Ω^* be finite. Assume $FVP(T) \neq \emptyset$. If each $\omega^* \in \Omega^*$ occurs infinitely often almost surely, then the sequence $\{x_n\}_{n=0}^{\infty}$ generated by the random Krasnoselskii-Mann iteration

$$x_{n+1} = \frac{1}{2}x_n + \frac{1}{2}T(\omega_n^*, x_n) \tag{4.38}$$

converges almost surely and in mean square to a random variable supported by the $FVP(T)$.

Proof: Since $T(\omega^*, x)$ is nonexpansive for each $\omega^* \in \Omega^*$, the random operator $\phi(\omega^*, x)$ where

$$\phi(\omega^*, x) := \frac{1}{2}(x + T(\omega^*, x)) \tag{4.39}$$

is, by Remark 2.1, firmly nonexpansive for each $\omega^* \in \Omega^*$. From Lemma 4.5, the sequence generated by the random Krasnoselskii-Mann algorithm (4.38) converges almost surely to a random variable supported by the $FVP(T)$. From the proof of Theorem 4.2, the sequence also converges in mean square to the random variable. Thus the proof is complete.

Remark 4.7: Algorithm (4.38) is a special case of Algorithm (2.2) for random case where $\alpha_n = \frac{1}{2}$. In this case, Algorithm (4.38) can be viewed as either the random Krasnoselskii-Mann iterative algorithm for finding a fixed value point of a nonexpansive random map $T(\omega^*, x)$ or the random Picard iterative algorithm for finding a fixed value point of a firmly nonexpansive random map $\phi(\omega^*, x)$ defined in (4.39).

CHAPTER 5. SOLVING LINEAR ALGEBRAIC EQUATIONS OVER RANDOM NETWORKS

In this chapter, we consider the problem of solving linear algebraic equations over random networks. This problem includes distributed consensus problem as a special case. We show that the random Krasnoselskii-Mann algorithm (4.38) is useful to solve this problem. The real Hilbert space considered in this chapter is $\mathcal{H} = (\Re^n, \|.\|_2), n \in N$. For simplicity we write $\|.\|_2 = \|.\|$ in this chapter.

5.1 A Distributed Algorithm for Solving Linear Algebraic Equations over Random Networks

Now we define the problem of solving linear algebraic equations over random network.

Consider m agents. The agents want to solve the problem $\min_x \|Ax - b\|, A \in \Re^{\mu \times q}, b \in \Re^{\mu}$, where each agent merely knows a subset of the rows of the partitioned matrix $[A, b]$; precisely, each agent knows a private equation $A_i x_i = b_i, i = 1, 2, ..., m$, where $A_i \in \Re^{\mu_i \times q}, b_i \in \Re^{\mu_i}, \sum_{i=1}^{m} \mu_i = \mu$. The objective of each agent is to collaboratively seek the solution of the following optimization problem using local information in presence of random interconnection graphs:

$$\min \sum_{i=1}^{m} \|A_i x - b_i\|^2$$

where $x \in \Re^q$.

Problem 5.1: Let the weighted random operator of the graph $T(\omega^*, x) := W(\omega^*)x$ be given (see Definition 3.2). Then the above problem under Assumptions 3.1 and 3.2 can be formulated as follows:

$$\begin{aligned} \min_x \quad & f(x) := \sum_{i=1}^{m} \|A_i x_i - b_i\|^2 \\ \text{subject to} \quad & x \in FVP(T), \end{aligned} \tag{5.1}$$

where $x = [x_1^T, ..., x_m^T]^T, x_i \in \Re^q, i = 1, 2, ..., m$.

Before presenting our main results, we impose the following assumption on the equation $Ax = b$.

Assumption 5.1: The linear algebraic equation $Ax = b$ has a solution, namely $\mathcal{S} := \{x | \min_x \|Ax - b\| = 0\} \neq \emptyset$.

Problem 5.1 with Assumption 5.1 can be reformulated as finding x such that

$$\bar{A}x = \bar{b}, \tag{5.2}$$

and

$$x \in FVP(T), \tag{5.3}$$

where

$$\bar{A} = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m \end{pmatrix}, \bar{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Lemma 5.1: The solution set of (5.2) is equal to the solution set of the following equation:

$$\tilde{A}x + \tilde{b} = x, \tag{5.4}$$

where

$$\tilde{A} = \begin{pmatrix} I_q - \theta_1 A_1^T A_1 & 0 & \cdots & 0 \\ 0 & I_q - \theta_2 A_2^T A_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_q - \theta_m A_m^T A_m \end{pmatrix}, \tag{5.5}$$

$$\tilde{b} = \begin{pmatrix} \theta_1 A_1^T b_1 \\ \theta_2 A_2^T b_2 \\ \vdots \\ \theta_m A_m^T b_m \end{pmatrix}, \tag{5.6}$$

and $\theta_i \in (0, \frac{2}{\lambda_{max}(A_i A_i^T)}), i = 1, 2, ..., m$.

Proof: Rows of (5.2) are written as $A_i x_i = b_i, i = 1, 2, ..., m$, which is equivalent to $x_i = x_i - \theta_i A_i^T(A_i x_i - b_i)$. Consequently, the solution sets of $A_i x_i = b_i$ and $x_i = x_i - \theta_i A_i^T(A_i x_i - b_i)$ are the same. This completes the proof of Lemma 5.1.

Now Problem 5.1 with Assumption 5.1 reduces to the following problem.

Problem 5.2: Consider Problem 5.1 with Assumption 5.1. Let $H(x) := \tilde{A}x + \tilde{b}$, where $\tilde{A}$ and $\tilde{b}$ are defined in (5.5)-(5.6), and let $T(\omega^*, x)$ be defined in Definition 3.2. The problem is to find x^* such that $x^* \in Fix(H) \cap FVP(T)$.

Now we give the following theorem.

Theorem 5.1: Consider Problem 5.2 with Assumption 4.2. Then starting from any initial condition, the sequence generated by the random Krasnoselskii-Mann algorithm

$$x_{n+1} = \frac{1}{2}x_n + \frac{1}{2}[(1-\varpi)W(\omega_n^*)x_n + \varpi(\tilde{A}x_n + \tilde{b})] \tag{5.7}$$

where $\varpi \in (0, 1)$ converges almost surely to x^* which is the unique solution of the following convex optimization problem:

$$\begin{aligned} &\min_x && \|x - x_0\| \\ &\text{subject to} && x = (1-\varpi)W(\omega^*)x + \varpi(\tilde{A}x + \tilde{b}),\ \forall \omega^* \in \Omega^*. \end{aligned} \tag{5.8}$$

Remark 5.1: Algorithm (5.7) cannot be derived from generalization of algorithms proposed in [93]-[119] and [120]-[134] to random case.

Before we give the proof of Theorem 5.1, we need to give some lemmas needed in the proof.

Lemma 5.2: Let $H(x)$ be defined in Problem 5.2. Then $H : \Re^{mq} \longrightarrow \Re^{mq}$ is nonexpansive.

Proof: We have that $\|H(z) - H(y)\| = \|\tilde{A}(z-y)\|, \forall z, y \in \Re^{mq}$. Now we prove that $\|\tilde{A}(z-y)\| \leq \|z - y\|$. Let $z = [z_1^T, z_2^T, ..., z_m^T]^T$ and $y = [y_1^T, y_2^T, ..., y_m^T]^T$. We have that

$$\|\tilde{A}(z-y)\|^2$$

$$= \left\| \begin{pmatrix} (I_q - \theta_1 A_1^T A_1)(z_1 - y_1) \\ (I_q - \theta_2 A_2^T A_2)(z_2 - y_2) \\ \vdots \\ (I_q - \theta_m A_m^T A_m)(z_m - y_m) \end{pmatrix} \right\|^2$$

$$= \sum_{j=1}^{m} \|(I_q - \theta_j A_j^T A_j)(z_j - y_j)\|^2.$$

Since $\theta_j \in (0, \frac{2}{\lambda_{max}(A_j A_j^T)})$, we have $\|I_q - \theta_j A_j^T A_j\| \leq 1$. Moreover, $\|(I_q - \theta_j A_j^T A_j)(z_j - y_j)\| \leq \|I_q - \theta_j A_j^T A_j\| \|z_j - y_j\|, j = 1, 2, ..., m$. Therefore, we obtain

$$\begin{aligned}\sum_{j=1}^{m} \|(I_q - \theta_j A_j^T A_j)(z_j - y_j)\|^2 &\leq \sum_{j=1}^{m} \|I_q - \theta_j A_j^T A_j\|^2 \|z_j - y_j\|^2 \\ &\leq \sum_{j=1}^{m} \|z_j - y_j\|^2 = \|z - y\|^2\end{aligned}$$

or

$$\|\tilde{A}(z - y)\| \leq \|z - y\|. \tag{5.9}$$

Thus the proof of Lemma 5.2 is complete.

Lemma 5.3: Let $T(\omega^*, x)$ and $H(x)$ be defined in Definition 3.2 and Problem 5.2, respectively, and

$$D(\omega^*, x) := (1 - \varpi)T(\omega^*, x) + \varpi H(x), \tag{5.10}$$

where $\omega^* \in \Omega^*, \varpi \in (0, 1)$. Then $FVP(D) = Fix(H) \cap FVP(T)$.

Proof: Assume a $\tilde{z} \in Fix(H) \cap FVP(T)$. In fact, $\tilde{z} = H(\tilde{z})$ and $\tilde{z} = T(\omega^*, \tilde{z}) = \tilde{z}, \forall \omega^* \in \Omega^*$. Therefore, we obtain from (5.10) that

$$\begin{aligned}D(\omega^*, \tilde{z}) &= (1 - \varpi)T(\omega^*, \tilde{z}) + \varpi H(\tilde{z}) \\ &= (1 - \varpi)\tilde{z} + \varpi \tilde{z} = \tilde{z}, \forall \omega^* \in \Omega^*,\end{aligned}$$

which implies that $Fix(H) \cap FVP(T) \subseteq FVP(D)$. Conversely, assume a $\tilde{z} \in FVP(D)$, i.e.,

$$D(\omega^*, \tilde{z}) = \tilde{z} = (1 - \varpi)T(\omega^*, \tilde{z}) + \varpi H(\tilde{z}), \forall \omega^* \in \Omega^*. \tag{5.11}$$

Since $Fix(H) \cap FVP(T) \neq \emptyset$, there exits a $y^* \in Fix(H) \cap FVP(T)$. Now by (5.11) we obtain

$$\|\tilde{z} - y^*\| = \|(1 - \varpi)T(\omega^*, \tilde{z}) + \varpi H(\tilde{z}) - y^*\|.$$

By the fact that $y^* = (1 - \varpi)y^* + \varpi y^*, \varpi \in (0, 1)$, we obtain

$$\begin{aligned}\|\tilde{z} - y^*\| &= \|(1 - \varpi)T(\omega^*, \tilde{z}) + \varpi H(\tilde{z}) - y^*\| \\ &= \|(1 - \varpi)(T(\omega^*, \tilde{z}) - y^*) + \varpi(H(\tilde{z}) - y^*)\|.\end{aligned} \tag{5.12}$$

Since $y^* = H(y^*)$ and $y^* = T(\omega^*, y^*), \forall \omega^* \in \Omega^*$, we obtain from (5.12) for all $\omega^* \in \Omega^*$ that

$$\|(1-\varpi)(T(\omega^*,\tilde{z})-y^*)+\varpi(H(\tilde{z})-y^*)\| = \|(1-\varpi)(T(\omega^*,\tilde{z})-T(\omega^*,y^*))+\varpi(H(\tilde{z})-H(y^*))\|. \quad (5.13)$$

Due to nonexpansivity property of $T(\omega^*, x)$ (see (3.7)), we have for all $\omega^* \in \Omega^*$ that

$$\|(1-\varpi)(T(\omega^*,\tilde{z})-T(\omega^*,y^*))+\varpi(H(\tilde{z})-H(y^*))\| \le (1-\varpi)\|\tilde{z}-y^*\|+\varpi\|H(\tilde{z})-H(y^*)\|. \quad (5.14)$$

By nonexpansivity property of $H(x)$ (see Lemma 5.2), we also have for all $\omega^* \in \Omega^*$ that

$$\|(1-\varpi)(T(\omega^*,\tilde{z})-T(\omega^*,y^*))+\varpi(H(\tilde{z})-H(y^*))\| \le (1-\varpi)\|T(\omega^*,\tilde{z})-T(\omega^*,y^*)\|+\varpi\|\tilde{z}-y^*\|. \quad (5.15)$$

Because of nonexpansivity property of $T(\omega^*, x)$, we obtain from (5.15) that

$$(1-\varpi)\|T(\omega^*,\tilde{z})-T(\omega^*,y^*)\|+\varpi\|\tilde{z}-y^*\| \le (1-\varpi)\|\tilde{z}-y^*\|+\varpi\|\tilde{z}-y^*\| = \|\tilde{z}-y^*\|, \forall \omega^* \in \Omega^*. \quad (5.16)$$

Due to nonexpansivity property of $H(x)$, we also obtain from (5.14) that

$$(1-\varpi)\|\tilde{z}-y^*\|+\varpi\|H(\tilde{z})-H(y^*)\| \le (1-\varpi)\|\tilde{z}-y^*\|+\varpi\|\tilde{z}-y^*\| = \|\tilde{z}-y^*\|. \quad (5.17)$$

From (5.12)-(5.17), we finally obtain

$$\begin{aligned}\|\tilde{z}-y^*\| &\le \|(1-\varpi)(T(\omega^*,\tilde{z})-T(\omega^*,y^*))+\varpi(H(\tilde{z})-H(y^*))\| \\ &\le (1-\varpi)\|\tilde{z}-y^*\|+\varpi\|H(\tilde{z})-H(y^*)\| \\ &\le \|\tilde{z}-y^*\|, \forall \omega^* \in \Omega^* \end{aligned} \quad (5.18)$$

and

$$\begin{aligned}\|\tilde{z}-y^*\| &\le \|(1-\varpi)(T(\omega^*,\tilde{z})-T(\omega^*,y^*))+\varpi(H(\tilde{z})-H(y^*))\| \\ &\le (1-\varpi)\|T(\omega^*,\tilde{z})-T(\omega^*,y^*)\|+\varpi\|\tilde{z}-y^*\| \\ &\le \|\tilde{z}-y^*\|, \forall \omega^* \in \Omega^*. \end{aligned} \quad (5.19)$$

Thus, the equalities hold in (5.18) and (5.19), that imply that

$$\begin{aligned}\|\tilde{z}-y^*\| &= \|(1-\varpi)(T(\omega^*,\tilde{z})-T(\omega^*,y^*))+\varpi(H(\tilde{z})-H(y^*))\| \\ &= \|H(\tilde{z})-H(y^*)\| \\ &= \|T(\omega^*,\tilde{z})-T(\omega^*,y^*)\|, \forall \omega^* \in \Omega^*. \end{aligned} \quad (5.20)$$

Now we have the following remark.

Remark 5.2 [219, Ch. 2]: Due to strict convexity of the norm $\|.\|$, if $\|x\| = \|y\| = \|(1-\varpi)x + \varpi y\|$ where $x, y \in X$ and $\varpi \in (0,1)$, then $x = y$.

Substituting $y^* = H(y^*)$ and $y^* = T(\omega^*, y^*), \forall \omega^* \in \Omega^*$, for (5.20) yields

$$\|H(\tilde{z}) - y^*\| = \|T(\omega^*, \tilde{z}) - y^*\| = \|(1-\varpi)(T(\omega^*, \tilde{z}) - y^*) + \varpi(H(\tilde{z}) - y^*)\|, \forall \omega^* \in \Omega^*,$$

which by Remark 5.2 implies that $H(\tilde{z}) - y^* = T(\omega^*, \tilde{z}) - y^*, \forall \omega^* \in \Omega^*$, or

$$H(\tilde{z}) = T(\omega^*, \tilde{z}), \forall \omega^* \in \Omega^*. \tag{5.21}$$

Substituting (5.21) for (5.11) yields

$$\tilde{z} = H(\tilde{z}) = T(\omega^*, \tilde{z}), \forall \omega^* \in \Omega^*,$$

which implies that $FVP(D) \subseteq Fix(H) \cap FVP(T)$. Therefore, $FVP(D) = Fix(H) \cap FVP(T)$. Thus the proof of Lemma 5.3 is complete.

Lemma 5.4: Let $D(\omega^*, x), \omega^* \in \Omega^*$, be defined in Lemma 5.3. Then $FVP(D)$ is a closed convex nonempty set.

Proof: For any $z, y \in \Re^{mq}$, we obtain

$$\begin{aligned}\|D(\omega^*, z) - D(\omega^*, y)\| &= \|(1-\varpi)(T(\omega^*, z) - T(\omega^*, y)) + \varpi(H(z) - H(y))\| \\ &\leq (1-\varpi)\|T(\omega^*, z) - T(\omega^*, y)\| + \varpi\|H(z) - H(y)\|. \end{aligned} \tag{5.22}$$

Because of nonexpansivity of both $T(\omega^*, x)$ and $H(x)$, we obtain from (5.22) that

$$\begin{aligned}\|D(\omega^*, z) - D(\omega^*, y)\| &\leq (1-\varpi)\|T(\omega^*, z) - T(\omega^*, y)\| + \varpi\|H(z) - H(y)\| \\ &\leq (1-\varpi)\|z - y\| + \varpi\|z - y\| = \|z - y\|\end{aligned}$$

that implies that $D(\omega^*, x)$ is nonexpansive. Indeed, since $\Re^{mq}$ is closed and convex, we obtain by Remark 3.3 that $FVP(D)$ is closed and convex. Furthermore, $FVP(D)$ is nonempty by Assumption 5.1 and Lemma 5.3. This completes the proof of Lemma 5.4.

Lemma 5.5: Let $T(\omega^*, x), \omega^* \in \Omega^*$, be defined in Definition 3.2, and

$$S(\omega, x) := (1 - \varpi)T(\omega^*, x) + \varpi\tilde{A}x, \omega^* \in \Omega^*, \tag{5.23}$$

where $\varpi \in (0, 1)$. Then $FVP(S)$ is nonempty, closed, and convex.

Proof: Since $\mathbf{0}_{mq}$ is a fixed value point of S, we can conclude that $FVP(S)$ is nonempty. Now for any $z, y \in \Re^{mq}$, we obtain

$$\begin{aligned}\|S(\omega^*, z) - S(\omega^*, y)\| &= \|(1 - \varpi)(T(\omega^*, z) - T(\omega^*, y)) + \varpi\tilde{A}(z - y)\| \\ &\leq (1 - \varpi)\|T(\omega^*, z) - T(\omega^*, y)\| + \varpi\|\tilde{A}(z - y)\|.\end{aligned} \tag{5.24}$$

Similar to the proof of Lemma 5.2, we obtain

$$\|\tilde{A}(z - y)\| \leq \|z - y\|. \tag{5.25}$$

Therefore, we obtain from (5.24) by nonexpansivity of $T(\omega^*, x)$ and (5.25) that

$$\begin{aligned}\|S(\omega^*, z) - S(\omega^*, y)\| &\leq (1 - \varpi)\|T(\omega^*, z) - T(\omega^*, y)\| + \varpi\|\tilde{A}(z - y)\| \\ &\leq (1 - \varpi)\|z - y\| + \varpi\|z - y\| \\ &= \|z - y\|\end{aligned}$$

which implies that $S(\omega^*, x), \omega^* \in \Omega^*$, is nonexpansive. Therefore, one can obtain by Remark 3.3 that $FVP(S)$ is closed and convex. Thus the proof of Lemma 5.5 is complete.

Lemma 5.6: Assume that the linear algebraic equation $Ax = b$ does not have a unique solution, i.e., $\mathcal{S}$ is not a singleton. Let $S(\omega^*, x)$ be defined in (5.23). Then $FVP(S)$ is a closed affine subspace.

Proof: By Lemma 5.5, we have that $FVP(S)$ is closed. Since $\mathcal{S}$ is not a singleton, $FVP(S)$ is not a singleton either. Consider two distinct points $\bar{z}, \bar{y} \in FVP(S)$, i.e.,

$$\bar{z} = S(\omega^*, \bar{z}), \bar{y} = S(\omega^*, \bar{y}), \forall\omega^* \in \Omega^*. \tag{5.26}$$

Now we obtain

$$\begin{aligned}S(\omega^*, \varsigma\bar{z} + (1 - \varsigma)\bar{y}) &= S(\omega^*, \varsigma\bar{z}) + S(\omega^*, (1 - \varsigma)\bar{y}) \\ &= \varsigma S(\omega^*, \bar{z}) + (1 - \varsigma)S(\omega^*, \bar{y}),\end{aligned} \tag{5.27}$$

where $\varsigma \in \Re$. Substituting (5.26) for (5.27) yields

$$S(\omega^*, \varsigma\bar{z} + (1-\varsigma)\bar{y}) = \varsigma S(\omega^*, \bar{z}) + (1-\varsigma)S(\omega^*, \bar{y}) = \varsigma\bar{z} + (1-\varsigma)\bar{y}$$

which implies that $\varsigma\bar{z} + (1-\varsigma)\bar{y} \in FVP(S)$. Therefore, $FVP(S)$ is an affine set.

Now we have the following remark.

Remark 5.3 [173]: If C is an affine set and $z_0 \in C$, then the set

$$C - z_0 = \{z - z_0 | z \in C\}$$

is a subspace.

Since $\mathbf{0}_{mq} \in FVP(S)$, we obtain by Remark 5.3 that the set

$$FVP(S) - \mathbf{0}_{mq} = FVP(S)$$

is a subspace. Thus the proof of Lemma 5.6 is complete.

Lemma 5.7: Let

$$Q_1(\omega^*, x) := \frac{1}{2}x + \frac{1}{2}D(\omega^*, x), \forall \omega^* \in \Omega^*, \tag{5.28}$$

$$Q_2(\omega^*, x) := \frac{1}{2}x + \frac{1}{2}S(\omega^*, x), \forall \omega^* \in \Omega^*. \tag{5.29}$$

Then $Q_1(\omega^*, x)$ and $Q_2(\omega^*, x)$ are nonexpansive and $FVP(Q_1) = FVP(D)$ and $FVP(Q_2) = FVP(S)$. Moreover, $Q_1(\omega^*, x)$ is firmly nonexpansive for each $\omega^* \in \Omega^*$.

Proof: Since $D(\omega^*, x)$ and $S(\omega^*, x)$ are nonexpansive, we obtain by Remark 2.1 that $Q_1(\omega^*, x)$ and $Q_1(\omega^*, x)$ are firmly nonexpansive for each $\omega^* \in \Omega^*$ and thus nonexpansive. Now consider a $\tilde{z} \in FVP(D)$. Thus $D(\omega^*, \tilde{z}) = \tilde{z}, \forall \omega^* \in \Omega^*$. Substituting this fact for (5.28) yields $Q_1(\omega^*, \tilde{z}) = \tilde{z}, \forall \omega^* \in \Omega^*$ which implies that $\tilde{z} \in FVP(Q_1)$. Now consider a $\tilde{z} \in FVP(Q_1)$. Similarly, one can obtain that $\tilde{z} \in FVP(D)$. Therefore, $FVP(Q_1) = FVP(D)$. With the same procedure, one can prove by using nonexpansivity of $S(\omega, x)$ (see proof of Lemma 5.5) that $FVP(Q_2) = FVP(S)$. Thus the proof of Lemma 5.7 is complete.

Remark 5.4: By Lemma 5.3 and Lemma 5.7, Assumption 5.1 guarantees that the set of equilibrium points of (5.7) is $Fix(H) \cap FVP(T) \neq \emptyset$. Also Assumption 5.1 guarantees the feasibility of the optimization problem (5.8).

Remark 5.5: Quadratic Lyapunov functions have been useful to analyze stability of linear dynamical systems. Nevertheless, common quadratic Lyapunov functions may not exist for consensus problems in networked systems [157]. Furthermore, common quadratic Lyapunov functions may not exist for switched linear systems [158]-[160]. Moreover, other difficulties mentioned in [164] may arise in using Lyapunov's direct method to analyze stability of dynamical systems. Also, LaSalle-type theorem for discrete-time stochastic systems (see [156] and references therein) needs $\{\omega_n^*\}_{n=0}^{\infty}$ to be independent. Therefore, we do not try Lyapunov's and LaSalle's approaches.

Proof of Theorem 5.1:

From Lemmas 5.3 and 5.7, we can write (5.7) as

$$x_{n+1} = Q_1(\omega_n^*, x_n). \tag{5.30}$$

Now we have the following definition and lemmas.

Definition 5.1 [220]: Suppose C is a closed convex nonempty set and $\{x_n\}_{n=0}^{\infty}$ is a sequence in $\mathcal{H}$. $\{x_n\}_{n=0}^{\infty}$ is said to be *Fejér monotone with respect to* C if

$$\|x_{n+1} - z\| \leq \|x_n - z\|, \quad \forall z \in C, n \geq 0.$$

Lemma 5.8 [220]: Suppose the sequence $\{x_n\}_{n=0}^{\infty}$ is Fejér monotone with respect to C. Then $\{x_n\}_{n=0}^{\infty}$ is bounded.

Lemma 5.9 [221]: Let $\{x_n\}_{n=0}^{\infty}$ be a sequence in $\mathcal{H}$ and let C be a closed affine subspace of $\mathcal{H}$. Suppose that $\{x_n\}_{n=0}^{\infty}$ is Fejér monotone with respect to C. Then $\mathcal{P}_C x_n = \mathcal{P}_C x_0, \forall n \in N$.

Consider a $\bar{c} \in FVP(D) = FVP(Q_1)$. From Lemma 5.7, we have $\bar{c} = Q_1(\omega^*, \bar{c})$. Hence, for all $\omega \in \Omega$, we have

$$\|x_{n+1} - \bar{c}\| = \|Q_1(\omega_n^*, x_n) - Q_1(\omega_n^*, \bar{c})\| \leq \|x_n - \bar{c}\|,$$

which implies that the sequence $\{x_n\}$ is Fejér monotone with respect to $FVP(D)$ (see Definition 5.1 and Lemma 5.4). Therefore, the sequence is bounded by Lemma 5.8 for all $\omega \in \Omega$. Since $m \in N$, $\bar{N}$ is finite. Thus we obtain from (5.30), Lemma 4.5, and Assumption 4.2 that $\{x_n\}_{n=0}^{\infty}$ converges almost surely to a random variable supported by $FVP(Q_1) = FVP(D)$ for any initial condition.

It remains to prove that $\{x_n\}_{n=0}^{\infty}$ converges almost surely to the unique solution x^*. If Problem 5.2 has a unique solution, then x^* is the only feasible point of the optimization (5.8); otherwise, $FVP(S)$ is a closed affine subspace by Lemma 5.6. Consider a fixed $\tilde{y} \in FVP(D) = FVP(Q_1)$. Thus $\tilde{y} = \frac{1}{2}\tilde{y} + \frac{1}{2}D(\omega^*, \tilde{y})$ and $D(\omega^*, \tilde{y}) = \tilde{y}, \forall \omega^* \in \Omega^*$. We obtain from these facts and (5.7) that

$$\begin{aligned} x_{n+1} - \tilde{y} &= \frac{1}{2}(x_n - \tilde{y}) + \frac{1}{2}(D(\omega_n^*, x_n) - \tilde{y}) \\ &= \frac{1}{2}(x_n - \tilde{y}) + \frac{1}{2}(D(\omega_n^*, x_n) - D(\omega_n^*, \tilde{y})) \\ &= \frac{1}{2}(x_n - \tilde{y}) + \frac{1}{2}(S(\omega_n^*, x_n) - S(\omega_n^*, \tilde{y})) \\ &= \frac{1}{2}(x_n - \tilde{y}) + \frac{1}{2}S(\omega_n^*, x_n - \tilde{y}) \\ &= Q_2(\omega_n^*, x_n - \tilde{y}). \end{aligned} \tag{5.31}$$

Now consider a $\bar{c} \in FVP(S) = FVP(Q_2)$. From (5.31) we obtain

$$\begin{aligned} \|x_{n+1} - \tilde{y} - \bar{c}\| &= \|Q_2(\omega_n^*, x_n - \tilde{y}) - \bar{c}\| \\ &= \|Q_2(\omega_n^*, x_n - \tilde{y}) - Q_2(\omega_n^*, \bar{c})\| \end{aligned}$$

which by nonexpansivity property of $Q_2(\omega^*, x)$ (see Lemma 5.7) implies

$$\begin{aligned} \|x_{n+1} - \tilde{y} - \bar{c}\| &= \|Q_2(\omega_n^*, x_n - \tilde{y}) - Q_2(\omega_n^*, \bar{c})\| \\ &\leq \|x_n - \tilde{y} - \bar{c}\|. \end{aligned} \tag{5.32}$$

Since $FVP(S) = FVP(Q_2)$ (by Lemma 5.7) is nonempty, closed, and convex (see Lemma 5.5), the sequence $\{x_n - \tilde{y}\}_{n=0}^{\infty}$ is Fejér monotone with respect to $FVP(Q_2) = FVP(S)$ for all $\omega \in \Omega$. Moreover, $FVP(S) = FVP(Q_2)$ (by Lemma 5.7) is a closed affine subspace by Lemma 5.6. Therefore, according to Lemma 5.9, we obtain

$$\lim_{n \longrightarrow \infty} x_n - \tilde{y} = \mathcal{P}_{FVP(S)}(x_0 - \tilde{y}).$$

As a matter of fact, $x^* = z^* + \tilde{y}$ where $z^* = \mathcal{P}_{FVP(S)}(x_0 - \tilde{y})$. Indeed, z^* can be considered as the solution of the following convex optimization problem:

$$\begin{aligned} \min_z \quad & \|z - (x_0 - \tilde{y})\| \\ \text{subject to} \quad & z = (1 - \varpi)W(\omega^*)z + \varpi\tilde{A}z, \ \forall \omega^* \in \Omega^*. \end{aligned} \tag{5.33}$$

By changing variable $x = z + \tilde{y}$ in optimization problem (5.33), (5.33) becomes

$$\begin{aligned} \min_{x} \quad & \|x - x_0\| \\ \text{subject to} \quad & x = (1-\varpi)W(\omega^*)(x-\tilde{y}) + \varpi\tilde{A}(x-\tilde{y}) + \tilde{y}, \forall\omega^* \in \Omega^*. \end{aligned} \tag{5.34}$$

where x^* is the solution of (5.34). By the fact that $\tilde{y} = (1-\varpi)\tilde{y} + \varpi\tilde{y}$, the constraint set in (5.34) becomes

$$x = (1-\varpi)(W(\omega^*)(x-\tilde{y}) + \tilde{y}) + \varpi(\tilde{A}(x-\tilde{y}) + \tilde{y}), \forall\omega^* \in \Omega^*. \tag{5.35}$$

Substituting $\tilde{y} = W(\omega^*)\tilde{y}, \forall\omega^* \in \Omega^*$, and $\tilde{y} = \tilde{A}\tilde{y} + \tilde{b}$ for (5.35) yields

$$x = (1-\varpi)W(\omega^*)x + \varpi(\tilde{A}x + \tilde{b}). \tag{5.36}$$

Substituting (5.36) for (5.34) yields (5.8). Because of strict convexity of $\|.\|$, the convex optimization problem (5.8) has the unique solution. Thus the proof of Theorem 5.1 is complete.

Theorem 5.2: Consider Problem 5.2 with Assumption 4.2. Then starting from any initial condition, the sequence generated by (5.7) converges in mean square to x^* which is the unique solution of the convex optimization problem (5.8).

Proof: One can obtain from Theorem 5.1 and the proof of Theorem 4.2.

5.1.1 Distributed Average Consensus over Random Networks

Now, we define the problem of reaching average consensus over random networks.

The agents want to reach the average of their initial states in presence of random interconnection topologies, i.e, $x_1 = x_2 = ... = x_m = \frac{1}{m}\sum_{i=1}^{m} x_i(0)$ where $x_i(0) \in \Re^d$ is an initial state of the agent i.

Before we present our results, we mention that the algorithm of Tsitsiklis [7] and its generalization to random case is

$$x_{n+1} = Wx_n. \tag{5.37}$$

Algorithm (5.37) is in fact the Picard iterative algorithm (2.1) for finding a fixed point of the nonexpansive operator $T(x) := Wx$. For periodic and irreducible matrices, the authors of [35]-[36]

prove that distributed consensus occurs with asynchronous updates. It is still a question if agents achieve consensus with synchronous updates. The answer is affirmative.

Remark 5.6: *Relaxation method* for convex feasibility problems was first investigated in [222]-[223]. It is shown in [224] that relaxation method is a special case of the Krasnoselskii-Mann iteration.

The random Krasnoselskii-Mann iterative algorithm (4.38) for consensus problems reduces to the following algorithm:

$$x_{n+1} = \frac{1}{2}x_n + \frac{1}{2}W(\omega_n^*)x_n. \tag{5.38}$$

Now we give the following theorem.

Theorem 5.3: Consider the average consensus problem where $\mathcal{W}(\omega^*)$ satisfies Assumptions 3.1, 3.2, and 4.2. Then the sequence generated by the random Krasnoselskii-Mann algorithm (5.38) converges almost surely to $x^* = \mathbf{1}_m \otimes \frac{1}{m}\sum_{i=1}^{m} x_i(0)$ so that $\mathcal{P}_{\mathcal{C}}x_n = \mathbf{1}_m \otimes \frac{1}{m}\sum_{i=1}^{m} x_i(0), \forall n \in N$.

Remark 5.7: We show here that the average consensus of initial states of the agents is in fact the projection of initial states of agents onto the consensus subspace in the Hilbert space $(\Re^{md}, \|.\|)$.

Before we give the proof of Theorem 5.3, we need to present the following lemma needed in the proof.

Lemma 5.10: Let

$$\tilde{D}(\omega^*, x) := \frac{1}{2}x + \frac{1}{2}T(\omega^*, x), \tag{5.39}$$

where $T(\omega^*, x)$ is defined in Definition 3.2. Then $FVP(\tilde{D}) = FVP(T)$.

Proof: One can prove from Lemma 5.3 where $\beta = \frac{1}{2}$ and $H(x) := x$.

Proof of Theorem 5.3: From (5.38), we have

$$\begin{aligned}\|x_{n+1}\| &= \|\frac{1}{2}x_n + \frac{1}{2}W(\omega_n^*)x_n\| \\ &\leq \frac{1}{2}\|x_n\| + \frac{1}{2}\|W(\omega_n^*)x_n\| \\ &\leq \frac{1}{2}\|x_n\| + \frac{1}{2}\|W(\omega_n^*)\|\|x_n\|.\end{aligned}$$

From (3.7), we have $\|W(\omega^*)\| \leq 1, \forall \omega^* \in \Omega^*$. Hence we obtain

$$\|x_{n+1}\| \leq \frac{1}{2}\|x_n\| + \frac{1}{2}\|W(\omega^*)\|\|x_n\| \leq \|x_n\|, \forall n \in N,$$

which implies that the sequence $\{x_n\}_{n=0}^{\infty}, \forall \omega \in \Omega$, is bounded.

Now consider a $c^* \in FVP(\tilde{D}) = \mathcal{C}$. Thus we have $c^* = \frac{1}{2}c^* + \frac{1}{2}c^*$. Using this fact and (5.38), we obtain

$$\begin{aligned}\|x_{n+1} - c^*\| &= \|\frac{1}{2}x_n + \frac{1}{2}W(\omega_n^*)x_n - c^*\| \\ &= \|\frac{1}{2}(x_n - c^*) + \frac{1}{2}(W(\omega_n^*)x_n - c^*)\|.\end{aligned} \tag{5.40}$$

Since $c^* \in FVP(\tilde{D})$, we have that $c^* = W(\omega_n^*)c^*, \forall \omega_n^* \in \Omega^*, \forall n \in N \cup \{0\}$. Therefore, we obtain

$$\begin{aligned}\|\frac{1}{2}(x_n - c^*) + \frac{1}{2}(W(\omega_n^*)x_n - c^*)\| &\leq \frac{1}{2}\|x_n - c^*\| + \frac{1}{2}\|W(\omega_n^*)(x_n - c^*)\| \\ &\leq \frac{1}{2}\|x_n - c^*\| + \frac{1}{2}\|W(\omega_n^*)\|\|x_n - c^*\|.\end{aligned} \tag{5.41}$$

Since $\|W(\omega^*)\| \leq 1, \forall \omega^* \in \Omega^*$, we obtain from (5.40)-(5.41) that

$$\|x_{n+1} - c^*\| \leq \frac{1}{2}\|x_n - c^*\| + \frac{1}{2}\|W(\omega_n^*)\|\|x_n - c^*\| \leq \|x_n - c^*\|. \tag{5.42}$$

Since the number of the agents, m, is finite, the number of the possible graphs $\bar{N}$ is finite, too. Due to nonexpansivity of $T(\omega^*, x)$ for each fixed $\omega^* \in \Omega^*$, we obtain by Remark 2.1 that

$$\tilde{S}(\omega^*, x) := \frac{1}{2}(x + T(\omega^*, x))$$

is firmly nonexpansive. Therefore, by Lemma 4.5, Lemma 5.10, and Assumption 4.2, (5.38) converges almost surely to a random variable supported by $\mathcal{C}$ since (5.38) is $x_{n+1} = \tilde{S}(\omega_n^*, x_n)$.

It remains to prove that the sequence $\{x_n\}_{n=0}^{\infty}$ converges almost surely to $x^* = \mathbf{1}_m \otimes \frac{1}{m}\sum_{j=1}^{m} x_j(0)$. We can see by (5.42) and Definition 5.1 that the sequence $\{x_n\}_{n=0}^{\infty}$ is Fejér monotone with respect to $\mathcal{C}$ for all $\omega \in \Omega$. Since $\mathcal{C}$ is a closed affine subspace, we conclude by Lemma 5.9 that the limit point of the sequence $\{x_n\}_{n=0}^{\infty}$ is $x^* = \mathcal{P}_{\mathcal{C}}x_0$, i.e., the solution of the following optimization problem:

$$\begin{aligned}&\underset{x}{\text{minimize}} \quad \|x - x_0\| \\ &\text{subject to} \quad x_1 = x_2 = ... = x_m.\end{aligned} \tag{5.43}$$

The optimization problem (5.43) is equivalent to the following optimization problem:

$$\begin{aligned}&\underset{x}{\text{minimize}} \quad \|x - x_0\|^2 \\ &\text{subject to} \quad x_1 = x_2 = ... = x_m.\end{aligned} \tag{5.44}$$

Indeed, the solution of the optimization problem (5.44) is $x^* = \mathbf{1}_m \otimes \frac{1}{m}\sum_{j=1}^{m} x_j(0)$ which implies that the sequence $\{x_n\}_{n=0}^{\infty}$ converges almost surely to the average of initial states of the agents. This completes the proof of Theorem 5.3.

Theorem 5.4: Consider the average consensus problem where $\mathcal{W}(\omega^*)$ satisfies Assumptions 3.1, 3.2, and 4.2. Then the sequence generated by (5.38) converges in mean square to $x^* = \mathbf{1}_m \otimes \frac{1}{m}\sum_{i=1}^{m} x_i(0)$ so that $\mathcal{P}_{\mathcal{C}} x_n = \mathbf{1}_m \otimes \frac{1}{m}\sum_{i=1}^{m} x_i(0), \forall n \in N$.

Proof: One can obtain from Theorem 5.3 and the proof of Theorem 4.2.

The random Krasnoselskii-Mann algorithm (5.7) for consensus problems reduces to the following algorithm:

$$x_{n+1} = \frac{1}{2}(1+\varpi)x_n + \frac{1}{2}(1-\varpi)W(\omega_n^*)x_n \tag{5.45}$$

where $\varpi \in (0,1)$. From Algorithms (5.7) and (5.38) and Theorems 5.1-5.4, we arrive at the following theorem.

Theorem 5.5: Consider the average consensus problem where $\mathcal{W}(\omega^*)$ satisfies Assumptions 3.1, 3.2, and 4.2. Then the sequence generated by (5.45) in which $\varpi \in [0,1)$ converges almost surely and in mean square to $x^* = \mathbf{1}_m \otimes \frac{1}{m}\sum_{i=1}^{m} x_i(0)$ so that $\mathcal{P}_{\mathcal{C}} x_n = \mathbf{1}_m \otimes \frac{1}{m}\sum_{i=1}^{m} x_i(0), \forall n \in N$.

Proof: Almost sure and mean square convergences of the sequence generated by Algorithm (5.45) where $\varpi \in (0,1)$ have been proved in Theorems 5.1 and 5.2, respectively. Almost sure and mean square convergences of the sequence generated by Algorithm (5.45) where $\varpi = 0$ have been proved in Theorems 5.3 and 5.4, respectively. Thus the proof of Theorem 5.5 is complete.

5.1.1.1 Numerical Example

Example 5.1: Consider three agents in the one-dimensional Euclidean space where $\Omega^* = \{\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3\}$ in which $\mathcal{G}_1 = \{\}, \mathcal{G}_2\{(1,2)\}$, and $\mathcal{G}_3\{(1,3)\}$ with undirected links where the weights of links are assumed to be $\mathcal{W}_{12} = 0.25, \mathcal{W}_{13} = 0.3$. One can see that $\mathcal{W}_{12}$ and $\mathcal{W}_{13}$ are neither Maximum-degree nor Metropolis weights. $\mathcal{W}(\omega^*), \forall \omega^* \in \Omega^*$ is doubly stochastic, and the union of all graphs in Ω^* is strongly connected. Therefore, Assumptions 3.1 and 3.2 are satisfied. We assume that $\mathcal{G}_1$ and $\mathcal{G}_2$ occur independently with probability $Pr(failure) = \frac{1}{2}$, and whenever $\mathcal{G}_2$ occurs and $\mathcal{G}_3$ did not occur in the previous iteration, $\mathcal{G}_3$ occurs after it. Thus the sequence $\{\omega_n^*\}_{n=0}^{\infty}$ is not independent and has time-varying distributions. In fact, it has a subsequence $\{\omega_{n_j}^*\}_{j=0}^{\infty}$ that is i.i.d. As a matter of fact, according to Borel-Cantelli lemma [215], $\mathcal{G}_1$ and $\mathcal{G}_2$ occur infinitely almost surely in the probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$ where

$$\bar{\Omega} = \{\mathcal{G}_1, \mathcal{G}_2\} \times \{\mathcal{G}_1, \mathcal{G}_2\} \times ...$$

Therefore, $\mathcal{G}_1$ and $\mathcal{G}_2$ occur infinitely almost surely in the probability space $(\Omega, \mathcal{F}, \mu)$ in this example. Thus $\mathcal{G}_3$ occurs infinitely almost surely, too. Therefore, Assumption 4.2 is satisfied. Indeed, the conditions of Theorem 5.3 are satisfied. We choose initial conditions $x_1(0) = -4, x_2(0) = 2$, and $x_3(0) = 5$ for simulation. In fact, the average of agents' initial states is $\frac{1}{3}\sum_{i=1}^{3} x_i(0) = 1$. We should mention that in the three-dimensional Euclidean space, we have that $\mathcal{P}_{\mathcal{C}}\zeta = [1, 1, 1]^T$ where $\zeta \in \{[x_1, x_2, x_3]^T \in \Re^3 | x_1 + x_2 + x_3 = 3\}$. As a matter of fact, the agents collaborate among themselves to approach the average of their initial states in such a way that they remain on the plane $\{[x_1, x_2, x_3]^T \in \Re^3 | x_1 + x_2 + x_3 = 3\}$ for all $n \in N$. The results are shown in Figure 5.1.

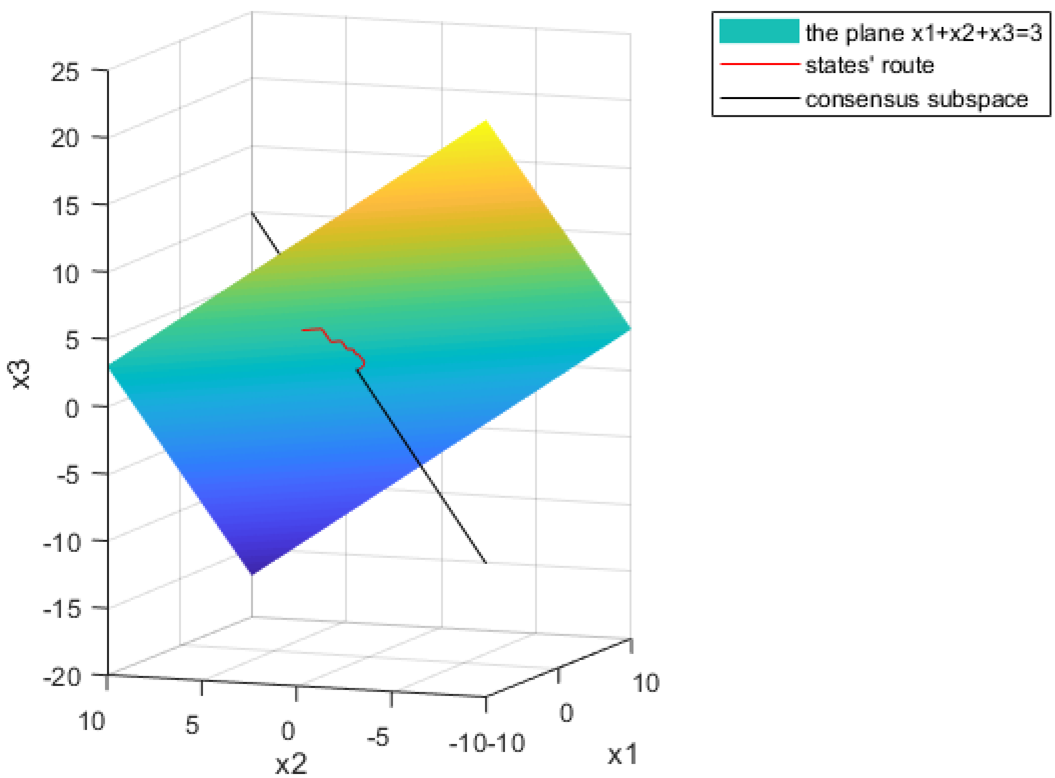

Figure 5.1 states' route

CHAPTER 6. A DISTRIBUTED ALGORITHM FOR DISTRIBUTED CONVEX OPTIMIZATION WITH STATE-DEPENDENT INTERACTIONS AND TIME-VARYING TOPOLOGIES

In this chapter, we show that a generalization of the proposed algorithm (4.33) can solve distributed optimization with state-dependent interactions and time-varying topologies. We consider the real Hilbert space $\mathcal{H} = (\Re^{mn}, \|.\|_2)$ in this chapter. For simplicity we write $\|.\|_2 = \|.\|$ in this chapter.

6.1 A Proposed Algorithm

We propose the following generalization of the distributed algorithm (4.33):

$$x_{n+1} = \alpha_n(x_n - \beta\nabla f(x_n)) + (1-\alpha_n)((1-\eta)x_n + \eta W(x_n, \mathcal{G}_n)x_n), \quad (6.1)$$

where $\eta \in (0,1)$.

Remark 6.1: The discrete algorithm proposed in [225] can solve consensus problems for a special weight form where $x_i \in \Re, i = 1, 2, ..., m$. Algorithm (6.1) is able to solve average consensus problems for weights satisfying Assumptions 3.5-3.6 while it is restricted to diminishing step size. Continuous algorithms have been proposed in [226] and [227] for solving consensus problems with state-dependent interactions.

6.1.1 Convergence Analysis

Related to Assumption 3.6, we have the following assumption.

Assumption 6.1: There exists a nonempty subset $\tilde{K} \subseteq G$ such that the union of all elements in $\tilde{K}$ is strongly connected for all $x \in \Re^{mn}$, and each element of $\tilde{K}$ occurs infinitely often.

Now we give the following theorem.

Theorem 6.1: Consider the problem (3.12) with Assumptions 3.5, 3.6, and 6.1. Let each $f_i(x_i), i = 1, \dots, m$, satisfies Assumption 4.1. Suppose that $\beta \in (0, \frac{2\xi}{K^2})$ and the sequence $\alpha_n \in [0,1], n \in N \cup \{0\}$, satisfies *(a)* and *(b)* of Theorem 4.1. Then the sequence generated by Algorithm (6.1) globally converges to the unique solution of the problem.

Before we give the proof of Theorem 6.1, we present the following lemma needed in the proof.

Lemma 6.1: Let $\hat{T}(x,\mathcal{G}) := (1-\eta)x + \eta T(x,\mathcal{G}), x, \mathcal{G} \in G, x \in \Re^{mn}$, with T defined in (3.12), and $\eta \in (0,1]$. Then

(i) $Fix(T(x,\mathcal{G})) = Fix(\hat{T}(x,\mathcal{G}))$.

(ii) $< x - \hat{T}(x,\mathcal{G}), x - z > \geq \frac{\eta}{2}\|x - T(x,\mathcal{G})\|^2, \quad \forall z \in \mathcal{C}, \forall \mathcal{G} \in G.$

(iii) $\|\hat{T}(x,\mathcal{G}) - z\| \leq \|x - z\|, \quad \forall z \in \mathcal{C}, \forall x \in \mathcal{H}, \forall \mathcal{G} \in G.$

Proof: (i)

Consider a $\hat{x} \in Fix(T(x,\mathcal{G}))$. Thus $\hat{x} = T(\hat{x},\mathcal{G})$. Hence

$$\hat{T}(\hat{x},\mathcal{G}) = (1-\eta)\hat{x} + \eta T(\hat{x},\mathcal{G}) = \hat{x},$$

which implies that $Fix(T(x,\mathcal{G})) \subseteq Fix(\hat{T}(x,\mathcal{G}))$. Conversely, consider a $\hat{x} \in Fix(\hat{T}(x,\mathcal{G}))$. Indeed, $\hat{x} = \hat{T}(\hat{x},\mathcal{G})$. Thus we have

$$\hat{x} = \hat{T}(\hat{x},\mathcal{G}) = (1-\eta)\hat{x} + \eta T(\hat{x},\mathcal{G}),$$

or $\hat{x} = T(\hat{x},\mathcal{G})$, which implies that $Fix(\hat{T}(x,\mathcal{G})) \subseteq Fix(T(x,\mathcal{G}))$. Therefore, we can conclude that $Fix(\hat{T}(x,\mathcal{G})) = Fix(T(x,\mathcal{G}))$. Thus the proof of *(i)* is complete.

(ii)

Since $z \in \mathcal{C}$, we have $W(x,\mathcal{G})z = z$. Therefore, we obtain

$$\begin{aligned}\|T(x,\mathcal{G}) - z\| &= \|W(x,\mathcal{G})x - W(x,\mathcal{G})z\| \\ &\leq \|W(x,\mathcal{G})\|\|x - z\|.\end{aligned}$$

Since by Assumption 3.5 $W(x,\mathcal{G})$ is doubly stochastic, we obtain from Lemma 3.2 that $\|W(x,\mathcal{G})\| \leq 1$. Hence,

$$\|T(x,\mathcal{G}) - z\| \leq \|W(x,\mathcal{G})\|\|x - z\| \leq \|x - z\| \qquad (6.2)$$

or

$$\|T(x,\mathcal{G})-z\|^2 \le \|x-z\|^2, \forall z \in \mathcal{C}, \forall \mathcal{G} \in G. \tag{6.3}$$

Also we have

$$\begin{aligned}\|T(x,\mathcal{G})-z\|^2 &= \|T(x,\mathcal{G})-x+x-z\|^2 \\ &= \|T(x,\mathcal{G})-x\|^2 + \|x-z\|^2 + 2 < T(x,\mathcal{G})-x, x-z > .\end{aligned} \tag{6.4}$$

Substituting (6.4) for (6.3) yields

$$2 < x-T(x,\mathcal{G}), x-z > \ge \|T(x,\mathcal{G})-x\|^2. \tag{6.5}$$

Substituting $x-T(x,\mathcal{G}) = \frac{x-\hat{T}(x,\mathcal{G})}{\eta}$ for the left hand side of the inequality (6.5) implies *(ii)*. Thus the proof of *(ii)* is complete.

(iii)

We have from (6.2) and $z=(1-\eta)z+\eta z$ that

$$\begin{aligned}\|\hat{T}(x,\mathcal{G})-z\| &\le (1-\eta)\|x-z\| + \eta\|T(x,\mathcal{G})-z\| \\ &\le (1-\eta)\|x-z\| + \eta\|x-z\| \\ &= \|x-z\|.\end{aligned}$$

Therefore, the proof of *(iii)* is complete.

Proof of Theorem 6.1:

We prove Theorem 6.1 in three steps.

Step 1: $\{x_n\}_{n=0}^{\infty}$ *is bounded.*

Step 2: $\{x_n\}_{n=0}^{\infty}$ *converges to an element in the feasible set.*

Step 3: $\{x_n\}_{n=0}^{\infty}$ *converges to the optimal solution.*

Now we give the proof of each step in details.

Proof of Step 1:

Since $f(x)$ is strongly convex and $\mathcal{C}$ is closed, the problem has a unique solution. Let x^* be the unique solution of the problem. We have $x^* = \alpha_n x^* + (1-\alpha_n)x^*$. Therefore, we have from (6.1) and $\hat{T}(x_n, \mathcal{G}_n) = (1-\eta)x_n + \eta T(x_n, \mathcal{G}_n)$ that

$$\begin{aligned}\|x_{n+1} - x^*\| &= \|\alpha_n(x_n - \beta\nabla f(x_n)) + (1-\alpha_n)\hat{T}(x_n, \mathcal{G}_n) - x^*\| \\ &= \|\alpha_n(x_n - \beta\nabla f(x_n) - x^*) + (1-\alpha_n)(\hat{T}(x_n, \mathcal{G}_n) - x^*)\|.\end{aligned}$$

We obtain from *(iii)* of Lemma 6.1 that

$$\begin{aligned}&\|\alpha_n(x_n - \beta\nabla f(x_n) - x^*) + (1-\alpha_n)(\hat{T}(x_n, \mathcal{G}_n) - x^*)\| \\ &\leq \alpha_n\|x_n - \beta\nabla f(x_n) - x^*\| + (1-\alpha_n)\|\hat{T}(x_n, \mathcal{G}_n) - x^*\| \\ &\leq \alpha_n\|x_n - \beta\nabla f(x_n) - x^*\| + (1-\alpha_n)\|x_n - x^*\|. \qquad (6.6)\end{aligned}$$

Since $\nabla f(x)$ is ξ-strongly monotone, and $\nabla f(x)$ is K-Lipschitz continuous, we obtain

$$\begin{aligned}\|x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*))\|^2 &= \|x_n - x^*\|^2 - 2\beta < \nabla f(x_n) - \nabla f(x^*), x_n - x^* > \\ &\quad + \beta^2\|\nabla f(x_n) - \nabla f(x^*)\|^2 \\ &\leq \|x_n - x^*\|^2 - 2\xi\beta\|x_n - x^*\|^2 + K^2\beta^2\|x_n - x^*\|^2 \\ &= (1 - 2\xi\beta + \beta^2K^2)\|x_n - x^*\|^2 \\ &= (1-\gamma)^2\|x_n - x^*\|^2\end{aligned}$$

where $\gamma = 1 - \sqrt{1 - \beta(2\xi - \beta K^2)}$, and selecting $\beta \in (0, \frac{2\xi}{K^2})$ implies $0 < \gamma \leq 1$. As a matter of fact, we have

$$\|x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*))\| \leq (1-\gamma)\|x_n - x^*\|. \qquad (6.7)$$

We have that

$$\begin{aligned}\|x_n - \beta\nabla f(x_n) - x^*\| &= \|x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*)) - \beta\nabla f(x^*)\| \\ &\leq \|x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*))\| + \beta\|\nabla f(x^*)\| \\ &\leq (1-\gamma)\|x_n - x^*\| + \beta\|\nabla f(x^*)\|. \qquad (6.8)\end{aligned}$$

Substituting (6.8) for (6.6) yields

$$\|x_{n+1} - x^*\| \leq (1-\gamma\alpha_n)\|x_n - x^*\| + \alpha_n\beta\|\nabla f(x^*)\| = (1-\gamma\alpha_n)\|x_n - x^*\| + \gamma\alpha_n(\frac{\beta\|\nabla f(x^*)\|}{\gamma})$$

which by induction implies

$$\|x_{n+1} - x^*\| \leq max\{\|x_0 - x^*\|, \frac{\beta\|\nabla f(x^*)\|}{\gamma}\}.$$

Thus $\{x_n\}_{n=0}^{\infty}$ is bounded.

Proof of Step 2:

From (6.1) and $x_n = \alpha_n x_n + (1-\alpha_n)x_n$, we have

$$x_{n+1} - x_n + \alpha_n \beta \nabla f(x_n) = (1-\alpha_n)(\hat{T}(x_n, \mathcal{G}_n) - x_n), \tag{6.9}$$

where $\hat{T}(x_n, \mathcal{G}_n) = (1-\eta)x_n + \eta T(x_n, \mathcal{G}_n)$. Hence

$$< x_{n+1} - x_n + \alpha_n \beta \nabla f(x_n), x_n - x^* >= -(1-\alpha_n) < x_n - \hat{T}(x_n, \mathcal{G}_n), x_n - x^* > . \tag{6.10}$$

Since $x^* \in \mathcal{C}$, we have from part *(ii)* of Lemma 6.1 that

$$< x_n - \hat{T}(x_n, \mathcal{G}_n), x_n - x^* >\geq \frac{\eta}{2}\|x_n - T(x_n, \mathcal{G}_n)\|^2. \tag{6.11}$$

From (6.10) and (6.11), we obtain

$$< x_{n+1} - x_n + \alpha_n \beta \nabla f(x_n), x_n - x^* >\leq -\frac{\eta}{2}(1-\alpha_n)\|x_n - T(x_n, \mathcal{G}_n)\|^2 \tag{6.12}$$

or equivalently

$$- < x_n - x_{n+1}, x_n - x^* >\leq -\alpha_n < \beta \nabla f(x_n), x_n - x^* > -\frac{\eta}{2}(1-\alpha_n)\|x_n - T(x_n, \mathcal{G}_n)\|^2. \tag{6.13}$$

From (4.15) we obtain

$$< x_n - x_{n+1}, x_n - x^* >= -C_{n+1} + C_n + \frac{1}{2}\|x_n - x_{n+1}\|^2, \tag{6.14}$$

where $C_n = \frac{1}{2}\|x_n - x^*\|^2$. From (6.13) and (6.14) we obtain

$$C_{n+1} - C_n - \frac{1}{2}\|x_n - x_{n+1}\|^2 \leq -\alpha_t < \beta \nabla f(x_n), x_n - x^* > -\frac{\eta}{2}(1-\alpha_n)\|x_n - T(x_n, \mathcal{G}_n)\|^2. \tag{6.15}$$

From (6.9) we have

$$\begin{aligned} \|x_{n+1} - x_n\|^2 &= \| - \alpha_n \beta \nabla f(x_n) + (1-\alpha_n)(\hat{T}(x_n, \mathcal{G}_n) - x_n)\|^2 \\ &= \alpha_n^2 \|\beta \nabla f(x_n)\|^2 + (1-\alpha_n)^2 \|\hat{T}(x_n, \mathcal{G}_n) - x_n\|^2 \\ &\quad - 2\alpha_n(1-\alpha_n) < \beta \nabla f(x_n), \hat{T}(\omega_n^*, x_n) - x_n > . \end{aligned} \tag{6.16}$$

We know that $\|\hat{T}(x_n,\mathcal{G}_n)-x_n\| = \eta\|x_n - T(x_n,\mathcal{G}_n)\|$. Since $\alpha_n \in [0,1]$, we have also that $(1-\alpha_n)^2 \leq (1-\alpha_n)$. Using these facts, (6.16) becomes

$$\frac{1}{2}\|x_{n+1}-x_n\|^2 \leq \frac{1}{2}\alpha_n^2\|\beta\nabla f(x_n)\|^2 + \frac{1}{2}(1-\alpha_n)\eta^2\|T(x_n,\mathcal{G}_n)-x_n\|^2 \\ -\alpha_n(1-\alpha_n) < \beta\nabla f(x_n), \hat{T}(x_n,\mathcal{G}_n) - x_n > . \tag{6.17}$$

From (6.15) and (6.17), we obtain

$$\begin{aligned} C_{n+1} - C_n &\leq \frac{1}{2}\|x_{n+1}-x_n\|^2 - \alpha_n < \beta\nabla f(x_n), x_n - x^* > \\ &\quad - \frac{\eta}{2}(1-\alpha_n)\|x_n - T(x_n,\mathcal{G}_n)\|^2 \\ &\leq -(\frac{1}{2}-\frac{\eta}{2})\eta(1-\alpha_n)\|x_n - T(x_n,\mathcal{G}_n)\|^2 \\ &\quad + \alpha_n(\frac{1}{2}\alpha_n\|\beta\nabla f(x_n)\|^2 - < \beta\nabla f(x_n), x_n - x^* > \\ &\quad - (1-\alpha_n) < \beta\nabla f(x_n), \hat{T}(x_n,\mathcal{G}_n) - x_n >). \end{aligned} \tag{6.18}$$

Now we claim that there exists an $n_0 \in N$ such that the sequence $\{C_n\}$ is non-increasing for $n \geq n_0$. Assume by contradiction that this is not true. Then there exists a subsequence $\{C_{n_j}\}$ such that

$$C_{n_j+1} - C_{n_j} > 0$$

which together with (6.18) yields

$$\begin{aligned} 0 &< C_{n_j+1} - C_{n_j} \\ &\leq -(\frac{1}{2}-\frac{\eta}{2})\eta(1-\alpha_{n_j})\|x_{n_j} - T(x_{n_j},\mathcal{G}_{n_j})\|^2 \\ &\quad + \alpha_{n_j}(\frac{1}{2}\alpha_{n_j}\beta^2\|\nabla f(x_{n_j})\|^2 - < \beta\nabla f(x_{n_j}), x_{n_j} - x^* > \\ &\quad - (1-\alpha_{n_j}) < \beta\nabla f(x_{n_j}), \hat{T}(x_{n_j},\mathcal{G}_{n_j}) - x_{n_j} >). \end{aligned} \tag{6.19}$$

Since $\{x_n\}$ is bounded, $\nabla f(x)$ is continuous, and $\eta \in (0,1)$, we obtain from (6.19) by Theorem 4.1 *(a)* that

$$\begin{aligned} 0 < \liminf_{j\longrightarrow\infty}(-(\frac{1}{2}-\frac{\eta}{2})\eta(1-\alpha_{n_j})\|x_{n_j}-T(x_{n_j},\mathcal{G}_{n_j})\|^2 \\ +\alpha_{n_j}(\frac{1}{2}\alpha_{n_j}\|\beta\nabla f(x_{n_j})\|^2 - <\beta\nabla f(x_{n_j}), x_{n_j}-x^*> \\ -(1-\alpha_{n_j})<\beta\nabla f(x_{n_j}), \hat{T}(\omega^*_{n_j}, x_{n_j})-x_{n_j}>)) \\ \leq 0 \end{aligned} \tag{6.20}$$

which is a contradiction. Therefore, there exists an $n_0 \in N$ such that the sequence $\{C_n\}$ is non-increasing for $n \geq n_0$. Since $\{C_n\}$ is bounded below, it converges.

Taking the limit of both sides of (6.18) and using the convergence of $\{C_n\}$, continuity of $\nabla f(x)$, Step 1, $\eta \in (0,1)$, and Theorem 4.1 *(a)* yield

$$\lim_{n\longrightarrow\infty}\|x_n - T(x_n,\mathcal{G}_n)\| = 0$$

or, by $\|\hat{T}(x_n,\mathcal{G}_n)-x_n\| = \eta\|x_n-T(x_n,\mathcal{G}_n)\|$,

$$\lim_{n\longrightarrow\infty}\|x_n - \hat{T}(x_n,\mathcal{G}_n)\| = 0. \tag{6.21}$$

We have from (6.1) and Step 1 that

$$\lim_{n\longrightarrow\infty}\|x_{n+1}-\hat{T}(x_n,\mathcal{G}_n)\| = \alpha_n\|x_n-\beta\nabla f(x_n)-\hat{T}(x_n,\mathcal{G}_n)\| = 0. \tag{6.22}$$

(6.21) and (6.22) together implies

$$\lim_{n\longrightarrow\infty}\|x_{n+1}-x_n\| \leq \lim_{n\longrightarrow\infty}\|x_{n+1}-\hat{T}(x_n,\mathcal{G}_n)\| + \lim_{n\longrightarrow\infty}\|x_n-\hat{T}(x_n,\mathcal{G}_n)\| = 0,$$

thus the sequence $\{x_n\}_{n=0}^{\infty}$ is convergent. From this fact, we obtain by Assumption 6.1, part *(i)* of Lemma 6.1, and (6.21) that $\{x_n\}_{n=0}^{\infty}$ converges to an element in the feasible set.

Proof of Step 3:

It remains to prove that $\lim_{n\longrightarrow\infty}\|x_n - x^*\| = 0$. We have that

$$\begin{aligned} \|x_{n+1}-x^*\|^2 &= \|x_{n+1}-x^*+\alpha_n\beta\nabla f(x^*)-\alpha_n\beta\nabla f(x^*)\|^2 \\ &= \|x_{n+1}-x^*+\alpha_n\beta\nabla f(x^*)\|^2+\alpha_n^2\|\beta\nabla f(x^*)\|^2 \\ &\quad -2\alpha_n<\beta\nabla f(x^*), x_{n+1}-x^*+\alpha_n\beta\nabla f(x^*)>. \end{aligned} \tag{6.23}$$

We have that $x^* = \alpha_n x^* + (1-\alpha_n)x^*, \forall n \in N \cup \{0\}$; using this fact and (6.1), we obtain

$$\|x_{n+1}-x^*+\alpha_n\beta\nabla f(x^*)\|^2 = \|\alpha_n[x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*))]+(1-\alpha_n)[\hat{T}(x_n,\mathcal{G}_n)-x^*]\|^2. \quad (6.24)$$

Moreover, we have

$$< \beta\nabla f(x^*), x_{n+1}-x^*+\alpha_n\beta\nabla f(x^*) >=< \beta\nabla f(x^*), x_{n+1}-x^* > +\alpha_n\|\beta\nabla f(x^*)\|^2. \quad (6.25)$$

Substituting (6.24) and (6.25) for (6.23) yields

$$\begin{aligned}
\|x_{n+1}-x^*\|^2 &= \|x_{n+1}-x^*+\alpha_n\beta\nabla f(x^*)\|^2 + \alpha_n^2\|\beta\nabla f(x^*)\|^2 \\
&\quad - 2\alpha_n < \beta\nabla f(x^*), x_{n+1}-x^*+\alpha_n\beta\nabla f(x^*) > \\
&= \|\alpha_n[x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*))] \\
&\quad + (1-\alpha_n)[\hat{T}(x_n,\mathcal{G}_n)-x^*]\|^2 \\
&\quad - 2\alpha_n < \beta\nabla f(x^*), x_{n+1}-x^* > -\alpha_n^2\|\beta\nabla f(x^*)\|^2 \\
&= \alpha_n^2\|x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*))\|^2 \\
&\quad + (1-\alpha_n)^2\|\hat{T}(x_n,\mathcal{G}_n)-x^*\|^2 \\
&\quad + 2\alpha_n(1-\alpha_n) < x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*)), \hat{T}(x_n,\mathcal{G}_n)-x^* > \\
&\quad - 2\alpha_n < \beta\nabla f(x^*), x_{n+1}-x^* > -\alpha_n^2\|\beta\nabla f(x^*)\|^2.
\end{aligned}$$

From (6.7), part *(iii)* of Lemma 6.1, and Cauchy–Schwarz inequality, we obtain

$$< x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*)), \hat{T}(x_n,\mathcal{G}_n)-x^* >\leq (1-\gamma)\|x_n-x^*\|^2. \quad (6.26)$$

From (6.7), we also obtain

$$\|x_n-x^*-\beta(\nabla f(x_n)-\nabla f(x^*))\|^2 \leq (1-\gamma)^2\|x_n-x^*\|^2. \quad (6.27)$$

Therefore, from (6.26), (6.27), and part *(iii)* of Lemma 6.1, we have

$$\begin{aligned}
\|x_{n+1}-x^*\|^2 &= \alpha_n^2\|x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*))\|^2 \\
&\quad + (1-\alpha_n)^2\|\hat{T}(x_n,\mathcal{G}_n) - x^*\|^2 \\
&\quad + 2\alpha_n(1-\alpha_n) < x_n - x^* - \beta(\nabla f(x_n) - \nabla f(x^*)), \hat{T}(x_n,\mathcal{G}_n) - x^* > \\
&\quad - 2\alpha_n < \beta\nabla f(x^*), x_{n+1} - x^* > -\alpha_n^2\|\beta\nabla f(x^*)\|^2 \\
&\leq (1-2\gamma\alpha_n)\|x_n - x^*\|^2 \\
&\quad + \alpha_n(\gamma^2\alpha_n\|x_n - x^*\|^2 - 2 < \beta\nabla f(x^*), x_{n+1} - x^* >) \\
&= (1-\gamma\alpha_n)\|x_n - x^*\|^2 - \gamma\alpha_n\|x_n - x^*\|^2 \\
&\quad + \alpha_n(\gamma^2\alpha_n\|x_n - x^*\|^2 - 2 < \beta\nabla f(x^*), x_{n+1} - x^* >). \\
&\leq (1-\gamma\alpha_n)\|x_n - x^*\|^2 \\
&\quad + \alpha_n(\gamma^2\alpha_n\|x_n - x^*\|^2 - 2 < \beta\nabla f(x^*), x_{n+1} - x^* >)
\end{aligned}$$

or finally

$$\|x_{n+1}-x^*\|^2 \leq (1-\gamma\alpha_n)\|x_n - x^*\|^2 + \gamma\alpha_n\left(\frac{\gamma^2\alpha_n\|x_n - x^*\|^2 - 2 < \beta\nabla f(x^*), x_{n+1} - x^* >}{\gamma}\right). \quad (6.28)$$

From Steps 1 and 2, (4.23), and Theorem 4.1 *(a)*, we obtain

$$\lim_{n\longrightarrow\infty}(\gamma^2\alpha_n\|x_n - x^*\|^2 - 2\beta < \nabla f(x^*), x_{n+1} - x^* >) \leq 0. \quad (6.29)$$

According to Lemma 4.2 by setting

$$\begin{aligned}
a_n &= \|x_n - x^*\|^2, \\
b_n &= \gamma\alpha_n, \\
h_n &= \left(\frac{\gamma^2\alpha_n\|x_n - x^*\|^2 - 2\beta < \nabla f(x^*), x_{n+1} - x^* >}{\gamma}\right),
\end{aligned}$$

we obtain from (6.28), (6.29), and Theorem 4.1 *(b)* that $\lim_{n\longrightarrow\infty}\|x_n - x^*\|^2 = 0$; therefore, $\{x_n\}$ converges to x^* as $n \longrightarrow \infty$. Thus the proof of Theorem 6.1 is complete.

6.1.1.1 Numerical Example

Now related to Remark 6.1, we give an instance of average consensus problem in the following example.

Example 6.1: Consider ten agents in a two-dimensional space that wish to reach average of their initial states. The state of each agent is its location in the two-dimensional space, i.e., $x_i = [y_i, z_i]^T$. It is known in this case that the cost functions of agents are

$$f_i(y, z) = 0.5(y - y_0^i)^2 + 0.5(z - z_0^i)^2$$

where $[y_0^i, z_0^i]^T$ is initial state of agent i.

The topology of the undirected graph is assumed to be $1 \longleftrightarrow 2 \ldots \longleftrightarrow 10$, and the weight of the link between agent i and j which is assumed to depend on the Euclidean distance of their states is considered of the form

$$\mathcal{W}_{ij} = \frac{0.25}{1 + \left\| \begin{pmatrix} y_i \\ z_i \end{pmatrix} - \begin{pmatrix} y_j \\ z_j \end{pmatrix} \right\|}.$$

The weight models the gain from j to i diminishing with the distance between the agents. One can see that $f_i(y, z), i = 1, 2, ..., 10$, are 1-strongly convex, and $\nabla f_i(y, z)$ are 1-Lipschitz continuous. Let Link 1=$\{(1, 2)\}, \ldots,$ Link 9=$\{(9, 10)\}$, where at each time n, the Link $t(mod)9 + 1$ works. Thus, the union of the graphs which occur infinitely often is strongly connected for all $x \in \Re^{20}$. Therefore, Assumption 6.1 is fulfilled. Thus the conditions of Theorem 6.1 are satisfied.

We use $\eta = 0.7, \alpha_n = \frac{1}{1+n}, n \geq 0, \beta = \frac{\xi}{K^2} = \frac{1}{1}$, and initial conditions $y_0^i = i, z_0^i = 2i$ for simulation. The results given by Algorithm (6.1) are shown in Figures 6.1 and 6.2. The error $e_n = \|x_n - x^*\|$ is shown in Figure 6.3. Figures 6.1-6.3 show that the positions of agents converge to the average of their initial positions.

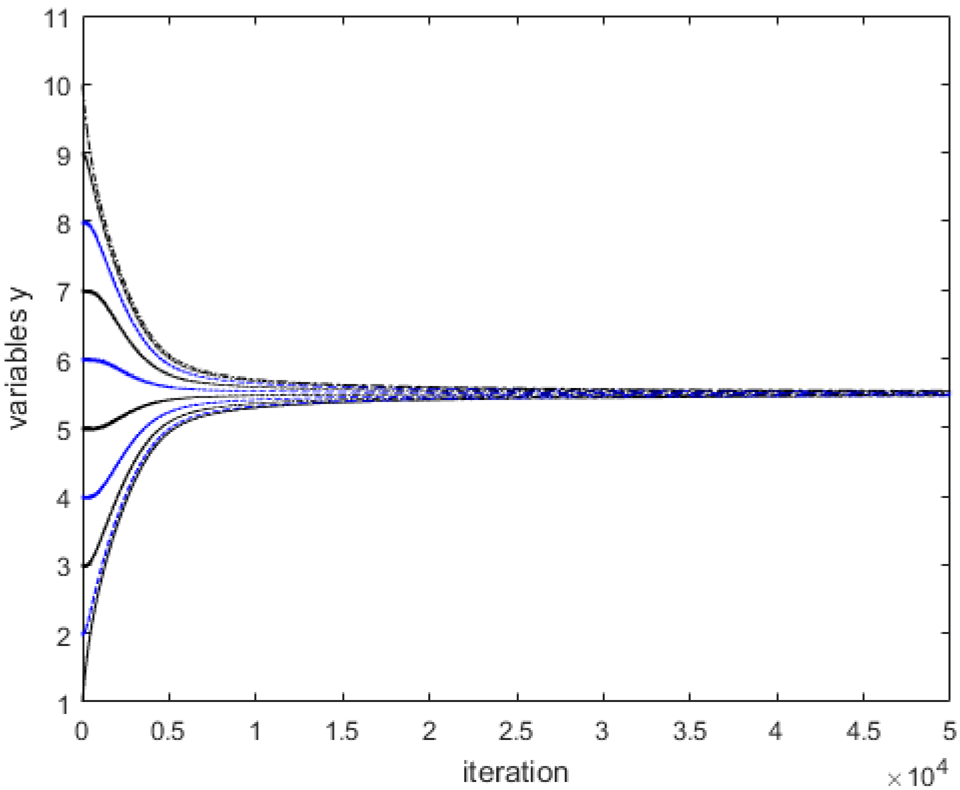

Figure 6.1 Variables y of agents in Example 6.1. This figure shows that the variables y converge to average of the initial positions of variables y.

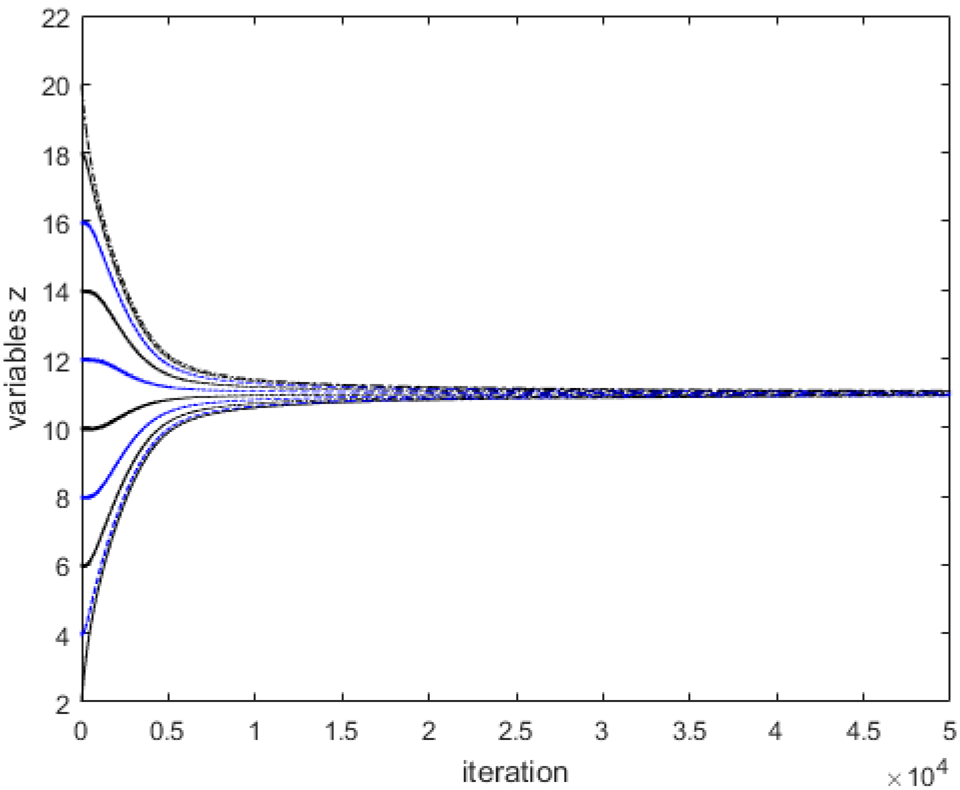

Figure 6.2 Variables z of agents in Example 6.1. This figure shows that the variables z converge to average of the initial positions of variables z.

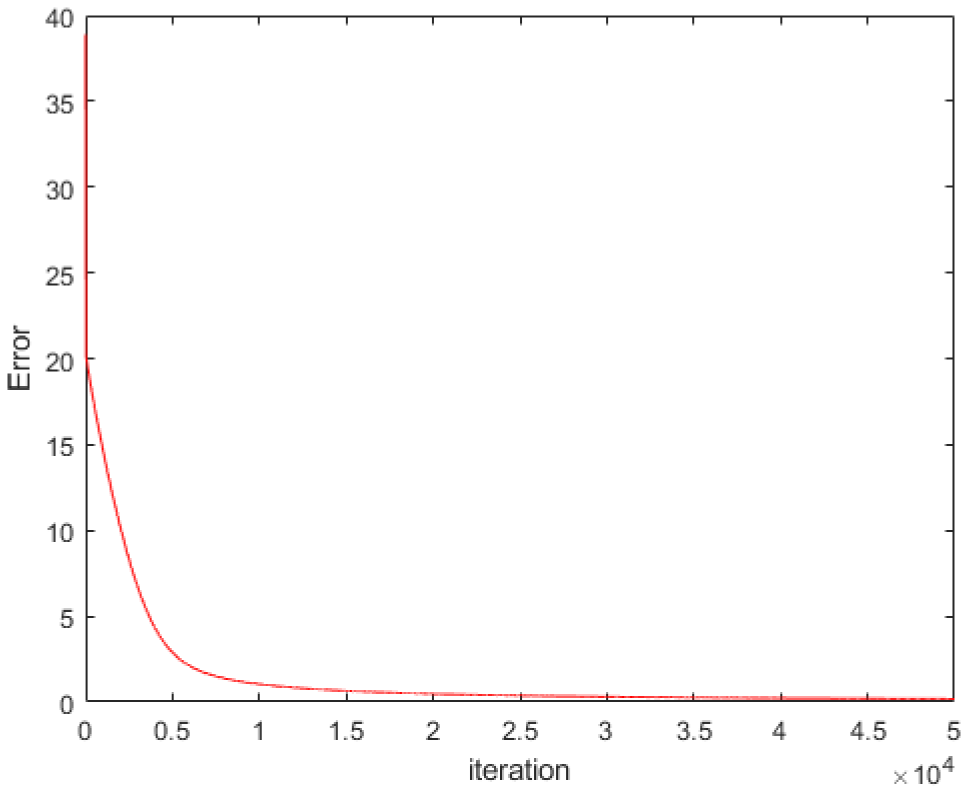

Figure 6.3 The error in Example 6.1. This figure shows that the positions of agents converge to the average of their initial positions.

CHAPTER 7. STABILITY OF STOCHASTIC NONLINEAR DISCRETE TIME SYSTEMS

So far the random Picard algorithm (4.35) and the random Krasnoselskii-Mann algorithm (4.38) for finding a fixed value point of a nonexpansive random operator are special cases of stochastic discrete-time systems. In this chapter, we analyze stability of stochastic nonlinear discrete-time systems by means of fixed point theory. We show that fixed point theory and the definition of fixed value point allow us to remove distribution dependency for stability of stochastic discrete-time systems by using Lyapunov's and LaSalle's approaches. We stress from Remark 5.5 that specific Lyapunov functions cannot exist for stability analysis of some stochastic systems.

Since the consensus subspace is a continuum set, the equilibrium set of stochastic systems (5.38) is continuum. In a continuum of equilibria, since every neighborhood of a non-isolated equilibrium contains another equilibrium, a non-isolated equilibrium cannot be asymptotically stable in the sense of Lyapunov. However, given a system that has a continuum of equilibria, it is natural to ask if the trajectories go to limit points and if the limit points are Lyapunov stable. These questions lead to consider properties of *convergence* and *semistability*. *Convergence* is the notion that every trajectory of the system goes to a limit point. The limit point, which is necessarily an equilibrium point, depends in general on the initial conditions. In a convergent system, the limit points of trajectories may or may not be Lyapunov stable. *Semistability* is the additional requirement that trajectories converges to limit points that are Lyapunov stable. Several authors have investigated semistability of deterministic and stochastic dynamical systems [228]-[244], to cite a few. Nevertheless, in this chapter, we only consider stability, but not semistability, of stochastic discrete-time systems.

7.1 Stability of Stochastic Systems

Now consider the following stochastic discrete-time system:

$$x_{t+1} = f(\omega_t^*, x_t) \tag{7.1}$$

where $f : \Omega^* \times \mathcal{R} \longrightarrow \mathcal{R}$ is a continuous random map (see Section 2.1), $\mathcal{R} \subseteq \Re^n$ is a closed set, and t represents time. We consider $(\Re^n, \|.\|_{\mathcal{B}})$. Note that $\Re^n$ equipped with any norm is a Banach space. For simplicity, we write $\|.\|_{\mathcal{B}} = \|.\|$ in this chapter.

Consider a probability measure P defined on the space $(\Omega, \mathcal{F})$ where

$$\Omega = \Omega^* \times \Omega^* \times \Omega^* \times \dots$$

$$\mathcal{F} = \sigma \times \sigma \times \sigma \times \dots$$

such that $(\Omega, \mathcal{F}, P)$ forms a probability space. We denote a realization in this probability space by $\omega \in \Omega$.

Now we have the following definitions of stability.

Definition 7.1 [245] *(Almost sure Lyapunov stability)*: The equilibrium point x^* of System (7.1) is said to be *almost surely Lyapunov stable* if $\forall \varepsilon > 0, \varrho > 0$, there exists $\delta = \delta(\varepsilon, \varrho) > 0$ such that $\|x_0 - x^*\| < \delta$ implies

$$P\{\sup_{t \geq 0} \|x_t - x^*\| \geq \varepsilon\} \leq \varrho.$$

Almost sure Lyapunov stability is also referred to as *Lyapunov stability with probability one.*

Definition 7.2 [245] *(Almost sure asymptotic stability)*: The equilibrium point of System (7.1) is said to be *almost surely asymptotically stable* if it is almost surely Lyapunov stable, and there exists $\delta > 0$ such that $\forall \varepsilon > 0, \|x_0 - x^*\| < \delta$ implies

$$\lim_{T \longrightarrow \infty} P\{\sup_{t \geq T} \|x_t - x^*\| \geq \varepsilon\} = 0.$$

Definition 7.3 [246] *(Mean square stability)*: The equilibrium point x^* of System (7.1) is said to be *mean square stable* if $\forall \varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that $E[\|x_t - x^*\|^2] < \varepsilon$ whenever $E[\|x_0 - x^*\|^2] < \delta$.

Definition 7.4 [246] *(Asymptotic mean square stability)*: The equilibrium point x^* of System (7.1) is said to be *asymptotically mean square stable* if it is mean square stable and there exists $\delta > 0$ such that $E[\|x_0 - x^*\|^2] < \delta$ implies

$$\lim_{t \longrightarrow \infty} E[\|x_t - x^*\|^2] = 0.$$

7.2 Stability Analysis of Stochastic Nonlinear Discrete-Time Systems

In this section, we show that with the help of fixed point theory and fixed value point, we can overcome distribution dependency of random variable sequences in existing results using Lyapunov's and LaSalle's approaches. Now we have the following theorem.

Theorem 7.1: Consider stochastic system (7.1). Assume that $f(\omega^*, x)$ is a contraction random mapping with constant $0 \leq \kappa < 1$ and $FVP(f) \neq \emptyset$. Then the equilibrium point x^* is both almost surely asymptotically stable and asymptotically mean square stable.

Proof: Since $f(\omega^*, x)$ is a contraction random map (see Definition 2.3), we obtain for each $\omega^* \in \Omega^*$ that

$$\|f(\omega^*, x) - f(\omega^*, y)\| \leq \kappa \|x - y\|, \quad \forall x, y \in \mathcal{R}.$$

According to Theorem 2.3, the random map $f(\omega^*, x)$ has a unique fixed point for each fixed $\omega^* \in \Omega^*$. Since $FVP(T) \neq \emptyset$, we have $x^* = f(\omega^*, x^*), \forall \omega^* \in \Omega^*$. Therefore, we obtain

$$\|x_{t+1} - x^*\| \leq \kappa \|x_t - x^*\|, \quad \forall \omega \in \Omega. \tag{7.2}$$

We have from (7.2) that

$$\|x_{t+1} - x^*\| \leq \kappa \|x_t - x^*\| \leq \|x_t - x^*\|, \quad \forall \omega \in \Omega,$$

or

$$\|x_t - x^*\| \leq \|x_0 - x^*\|, \quad \forall t \in N. \tag{7.3}$$

Hence, by setting $\varepsilon = \delta$ in Definition 7.1, we obtain

$$\|x_t - x^*\| \leq \|x_0 - x^*\| < \varepsilon$$

which implies that the equilibrium point x^* is almost surely Lyapunov stable. We also have

$$E[\|x_t - x^*\|^2] = \int_\Omega \|x_t - x^*\|^2 dP \leq \int_\Omega \|x_0 - x^*\|^2 dP = E[\|x_0 - x^*\|^2].$$

Thus by setting $\varepsilon = \delta$ in Definition 7.3, we obtain

$$E[\|x_t - x^*\|^2] \leq E[\|x_0 - x^*\|^2] < \varepsilon$$

which implies that the equilibrium point x^* is mean square stable.

We have from (7.2) that

$$\|x_{t+1} - x^*\| \leq \kappa^{t+1}\|x_0 - x^*\|, \quad \forall \omega \in \Omega. \tag{7.4}$$

Taking the limit of both sides of (7.4) yields

$$\lim_{t \longrightarrow \infty} \|x_{t+1} - x^*\| \leq \lim_{t \longrightarrow \infty} \kappa^{t+1}\|x_0 - x^*\| = 0, \quad \forall \omega \in \Omega$$

implying that the equilibrium point x^* is almost surely asymptotically stable. We also have

$$E[\|x_t - x^*\|^2] = \lim_{t \longrightarrow \infty} \int_\Omega \|x_t - x^*\|^2 dP \leq \lim_{t \longrightarrow \infty} \kappa^{2t}\|x_0 - x^*\|^2 P(\Omega) = 0$$

which implies that the equilibrium point x^* is asymptotically mean square stable. Thus the proof is complete.

Corollary 7.1: Consider the stochastic linear system

$$x_{t+1} = A(\omega_t^*)x_t, \quad \omega_t^* \in \Omega^*. \tag{7.5}$$

If there exists a norm $\|.\|_*$ such that

$$\|A(\omega^*)\|_* \leq \kappa, \quad \forall \omega^* \in \Omega^*,$$

where $0 \leq \kappa < 1$, then the origin is both almost surely asymptotically stable and asymptotically mean square stable.

Hint: $\|Ax\|_* \leq \|A\|_*\|x\|_*$ where $A \in \Re^{n \times n}$ and $x \in \Re^n$.

Remark 7.1: One may consider $\|.\|_1, \|.\|_2$, or $\|.\|_\infty$ for $\|.\|_*$ in Corollary 7.1.

Remark 7.2: Consider Theorem 7.1 where $f(x)$ is deterministic. Here, we do not assume that the system is locally continuously differentiable at its equilibrium point.

Theorem 7.2: Consider stochastic system (7.1) where the cardinality of the set Ω^* is finite and $(\Re^n, \|.\|_{\mathcal{H}})$. Let $f(\omega^*, x)$ be firmly nonexpansive random map, and $FVP(f) \neq \emptyset$. Assume that there exists a nonempty subset $\bar{K} \subseteq \Omega^*$ such that $\{x^*\} = \{\bar{z} | \bar{z} \in \Re^n, \bar{z} = f(\tilde{\omega}, \bar{z}), \forall \tilde{\omega} \in \bar{K}\}$, and each element of $\bar{K}$ occurs infinitely often almost surely. Then the equilibrium point x^* is both almost surely asymptotically stable and asymptotically mean square stable.

Remark 7.3: We recall Remark 4.3 here. If the sequence $\{\bar{\omega}_n\}_{n=0}^{\infty}$ is mutually independent with $\sum_{n=0}^{\infty} Pr_n(\bar{\omega}) = \infty$ where $Pr_n(\bar{\omega})$ is the probability of $\bar{\omega}$ occurring at time n, then according to Borel-Cantelli lemma [215], the assumption in Theorem 7.2 is satisfied. Consequently, any i.i.d. random sequence satisfies the assumption in Theorem 7.2. Any ergodic stationary sequences $\{\omega_n^*\}_{n=0}^{\infty}, Pr(\bar{\omega}) > 0$, satisfy the assumption in Theorem 7.2 (see proof of Lemma 1 in [49]). Consequently, any time-invariant Markov chain with its unique stationary distribution as the initial distribution satisfies the assumption in Theorem 7.2 (see [49]).

Proof: Since $f(\omega^*, x)$ is firmly nonexpansive random operator, we have by Remark 2.1 that it is nonexpansive. Thus we obtain for each $\omega^* \in \Omega^*$ that

$$\|f(\omega^*, x) - f(\omega^*, y)\|_{\mathcal{H}} \leq \|x - y\|_{\mathcal{H}}, \quad x, y \in \mathcal{R}.$$

Since $x^* = f(\omega^*, x^*), \forall \omega^* \in \Omega^*$, we obtain

$$\|x_{n+1} - x^*\|_{\mathcal{H}} \leq \|x_n - x^*\|_{\mathcal{H}}, \quad \forall \omega \in \Omega.$$

Therefore, similar to the proof of Theorem 7.1, we obtain that the equilibrium point x^* is both almost surely Lyapunov stable and mean square stable. From Lemma 4.5 and assumptions of Theorem 7.2, the equilibrium point x^* is almost surely asymptotically stable. Since the sequence $\{x_t\}_{t=0}^{\infty}$ is bounded for all $\omega \in \Omega$, we obtain from the proof of Theorem 4.2 that the equilibrium point x^* is asymptotically mean square stable. Thus the proof is complete.

Corollary 7.2: Consider the stochastic linear system (7.5) where the cardinality of the set Ω^* is finite. Assume that there exists a nonempty subset $\bar{K} \subseteq \Omega^*$ such that $\{\mathbf{0}_n\} = \{\bar{z} | \bar{z} \in \Re^n, \bar{z} =$

$A(\tilde{\omega})\bar{z}, \forall\tilde{\omega} \in \bar{K}\}$, and each element of $\bar{K}$ occurs infinitely often almost surely. If for each fixed $\omega^* \in \Omega^*$ either

I) $A(\omega^*) - A^T(\omega^*)A(\omega^*) \succeq 0$

or

II) $A(\omega^*) = \frac{1}{2}I_n + \frac{1}{2}\bar{A}(\omega^*)$

for some $\|\bar{A}(\omega^*)\|_2 \leq 1$, then the origin is both almost surely asymptotically stable and asymptotically mean square stable.

Hint: Use Definition 2.5 and Remark 2.1.

CHAPTER 8. CONCLUSIONS AND FUTURE WORKS

In this chapter, we summarize the contributions of this dissertation and give some directions for future research.

8.1 Contributions of This Dissertation

1- A new mathematical optimization problem based on a new mathematical terminology: We have defined a new mathematical terminology called *fixed value point* and defined a new mathematical optimization problem (3.1). This problem includes both centralized and distributed optimization problems. The results have been published and accepted in [247] and [248], respectively.

2- Centralized robust convex optimization on Hilbert spaces: We have defined robust convex optimization on real Hilbert spaces. We have shown that this problem is included in the problem (3.1).

3- A framework for distributed optimization over random networks: Based on the optimization problem (3.1), we have proposed a framework for unconstrained distributed optimization problems over random networks. The results are given in [247]-[248].

4- A framework for distributed optimization with state-dependent interactions: We have defined a framework for distributed optimization with state-dependent interactions. Then we have generalized it to give a framework for distributed optimization with state-dependent interactions and time-varying topologies. The preliminary results have been published in [249].

5- A proposed algorithm for the mathematical optimization: We have proposed an algorithm with diminishing step size to solve the mathematical optimization problem and analyzed its convergence with suitable assumptions. The results are given in [248].

6- An algorithm to solve unconstrained distributed optimization over random networks: As a special case of the proposed algorithm for solving the mathematical optimization problem, we have given an asynchronous algorithm with diminishing step size for solving distributed optimization over random networks that does not require distribution dependency or B-connectivity assumption on random communication graphs for convergence. The results are given in [248].

7- The random Picard and Krasnoselskii-Mann algorithms for solving the optimization problem: We have shown that the random Picard algorithm or the random Krasnoselskii-Mann algorithm which do not suffer from diminishing step size are useful to solve the feasibility problem of (3.1).

8- Solving linear algebraic equations over random networks: We have shown that the random Krasnoselskii-Mann iterative algorithm can be applied for solving linear algebraic equations over random networks without distribution dependency or B-connectivity assumption on random communication graphs for convergence. The algorithm is also an asynchronous algorithm. The preliminary results have been published in [250].

9- Distributed average consensus over random networks: We have shown that the random Krasnoselskii-Mann iterative algorithm can be applied for distributed average consensus over random networks without distribution dependency or B-connectivity assumption on random communication graphs for convergence. The algorithm is also an asynchronous algorithm. We have shown that the algorithm converges when the weighted matrix of the graph is periodic and irreducible. The results have been accepted in [251].

10- A distributed algorithm for convex optimization with state-dependent interactions and time-varying topologies: As a generalization of the proposed algorithm for distributed optimization over random networks, we have proposed an algorithm to solve distributed optimization with state-dependent interactions and time-varying topologies that does not require B-connectivity assumption on communication graphs for convergence. We have shown that this algorithm can be applied for solving distributed average consensus with state-dependent interactions and time-varying topologies.

11- Stability analysis of stochastic nonlinear discrete-time systems by means of fixed point theory: We have analyzed stability of stochastic nonlinear discrete-time systems by using fixed point theory to overcome difficulties that arise in using Lyapunov's and LaSalle's approaches such as distribution dependency of random variable sequences.

8.2 Future Works

Several future research directions based on the approaches of this dissertation are:

- Relaxing strong convexity or K-Lipschitz assumption on the cost function to only convex function in Assumption 4.1 to propose a distributed asynchronous algorithm which converges with any convex functions of agents.

- Distributed asynchronous algorithm without diminishing step size for solving least square problems.

- Relaxing doubly stochastic assumptions to only row stochastic assumption for distributed optimization over random networks or state-dependent interactions.

REFERENCES

[1] Z. Jiang, Distributed optimization for control and learning, Ph.D. dissertation, Dep. Mech. Eng., Iowa State University, Ames, Iowa, USA, 2018.

[2] S. Boyed, N. Parikh, E. Chu, B. Peleato, and J. Eckstin, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. Trends. Mach. Learn.*, vol. 3, pp. 1–122, 2010.

[3] S. S. Ram, A. Nedić, and V. V. Veeravalli, A new class of distributed optimization algorithms: Application to regression of distributed data, *Optim. Methods Softw.*, vol. 27, no. 1, pp. 71–88, 2012.

[4] N. A. Lynch, *Distributed Algorithms*, San Mateo, CA: Morgan Kaufmann, 1997.

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

[6] V. Borkar and P. P. Varaiya, Asymptotic agreement in distributed estimation, *IEEE Trans. on Automatic Control*, vol. AC-27, no. 3, pp. 650–655, 1982.

[7] J. N. Tsitsiklis, Problems in decentralized decision making and computation, Ph.D. dissertation, Dep. Elect. Eng. Comp. Sci., MIT, Cambridge, MA, 1984.

[8] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Englewood Cliffs, NJ: Prentice Hall, 1989.

[9] J. R. Marden, G. Arslan, and J. S. Shamma, Cooperative control and potential games, *IEEE Trans. on Systems, Man, and Cybernetics–Part B: Cybernetics*, vol. 39, no. 6, pp. 1393–1407, 2009.

[10] A. Nedić and A. Ozdaglar, Distributed subgradient methods for multi-agent optimization, *IEEE Trans. Automatic Control*, vol. 56, pp. 48–61, 2009.

[11] J. Wang and N. Elia, Control approach to distributed optimization, *Proc. of* 48^{th} *Annual Allerton Conf.*, Sep. 29–Oct. 1, Allerton House, UIUC, Illinois, USA, pp. 557–561, 2010.

[12] S. S. Kia, J. Cortés, and S. Mrtínez, Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication, *Automatica*, vol. 55, pp. 254–264, 2015.

[13] M. Franceschelli, A. Giua, and A. Pisano, Finite-time consensus on the median value with robustness properties, *IEEE Trans. on Automatic Control*, vol. 62, no. 4, pp. 1652–1667, 2017.

[14] D. Varagnolo, F. Zanella, A. Cenedese, G. Pillonetto, and L. Schenato, Newton-Raphson consensus for distributed convex optimization, *IEEE Trans. on Automatic Control*, vol. 61, no. 4, pp. 994–1009, 2016.

[15] T-H. Chang, M. Hong, and X. Wang, Multi-agent distributed optimization via inexact consensus ADMM, *IEEE Trans. on Signal Processing*, vol. 63, no. 2, pp. 482–497, 2015.

[16] N. Li and J. R. Marden, Designing games for distributed optimization, *IEEE J. of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 230–242, 2013.

[17] P. Chebotarev and R. Agaev, The forest consensus theorem, *IEEE Trans. on Automatic Control*, vol. 59, no. 9, pp. 2475–2479, 2014.

[18] G. Shi and K. H. Johansson, The role of persistent graphs in the agreement seeking of social networks, *IEEE Journal on Selected Areas in Communications/Supplement*, vol. 31, pp. 595–606, 2013.

[19] R. Olfati-Saber, J. A. Fax, and R. M. Murray, Consensus and cooperation in networked multi-agent systems, *Proceedings of The IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[20] D. Cheng, J. Wang, and X. Hu, An extension of LaSalle's invariance principle and its application to multi-agent consensus, *IEEE Trans. on Automatic Control*, vol. 53, no. 7, pp. 1765–1770, 2008.

[21] B. Ning, Q-L. Han, and Z. Zuo, Distributed optimization for multiagent systems: an edge-based fixed-time consensus approach, *IEEE Trans. on Cybernetics*, vol. 49, pp. 122-132, 2019.

[22] G. Notarstefano and F. Bullo, Distributed abstract optimization via constraints consensus: theory and applications, *IEEE Trans. on Automatic Control*, vol. 56, no. 10, pp. 2247–2261, 2011.

[23] Y. Zhang and Y. Hong, Multi-agent consensus convergence analysis with infinitely jointly connected and directed communication graphs, *Proc. of the* 11^{th} *World Congress on Intelligent Control and Automation*, June 29–July 4, Shenyang, China, pp. 2091–2096, 2014.

[24] J. Liu and A. Nedić, and T. Başar, Complex constrained consensus, *Proc. of* 53^{th} *IEEE Conf. on Decision and Control*, Dec. 15–17, Los Angeles, CA, USA, pp. 1464–1469, 2014.

[25] A. Nedić ans A. Olshevsky, Distributed optimization over time-varying directed graphs, *IEEE Trans. on Automatic Control*, vol. 60, no. 3, pp. 601–615, 2015.

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

[27] C. N. Hadjicostis, N. H. Vaidya, and A. D. Domínguez-García, Robust distributed average consensus via exchange of running sums, *IEEE Trans. on Automatic Control*, vol. 61, pp. 1492–1507, 2016.

[28] T. Li and J. Wang, Distributed averaging with random network graphs and noises, *IEEE Trans. on Information Theory*, vol. 64, pp. 7063–7080, 2018.

[29] W. Ren and R. W. Beard, Consensus seeking in multiagent systems under dynamically changing interaction topologies, *IEEE Trans. on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[30] Y. Hatano and M. Mesbahi, Agreement over random networks, *IEEE Trans. on Automatic Control*, vol. 50, no. 11, pp. 1867–1872, 2005.

[31] J. Wu, Z. Meng, T. Yang, G. Shi, and K. H. Johansson, Sampled-data consensus over random networks, *IEEE Trans. on Signal Processing*, vol. 64, no. 17, pp. 4479–4492, 2016.

[32] Y. Wang, L. Cheng, W. Ren, Z-G. Hou, and M. Tan, Seeking consensus in networks of linear agents: communication noises and Markovian switching topologies, *IEEE Trans. on Automatic Control*, vol. 60, no. 5, pp. 1374–1379, 2015.

[33] S. Patterson, B. Bamieh, and A. E. Abbadi, Distributed average consensus with stochastic communication failures, *Proc. of* 46^{th} *IEEE Conf. on Decision and Control*, Dec. 12–14, New Orleans, LA, USA, pp. 4215–4220, 2007.

[34] J. Wang and N. Elia, Distributed averaging algorithms resilient to communication noise and dropouts, *IEEE Trans. on Signal Processing*, vol. 61, no. 9, pp. 2231–2242, 2013.

[35] Y. Qin, M. Cao, and B. D. O. Anderson, Asynchronous agreement through distributed coordination algorithms associated with periodic matrices, *Proc. of* 20^{th} *IFAC World Congress*, July 9–14, Toulouse, France, pp. 1742–1747, 2017.

[36] G. Shi, B. D. O. Anderson, and K. H. Johansson, Consensus over random graph processes: network Borel-Cantelli lemmas for almost sure convergence, *IEEE Trans. on Information Theory*, vol. 61, pp. 5690–5707, 2015.

[37] F. Bénézit, V. Blondel, P. Thiran, J. Tsitsiklis, and M. Vetterli, Weighted gossip: distributed averaging using non-doubly stochastic matrices, *Proc. of IEEE International Symposium on Information Theory*, June 13–18, Austin, Texas, USA, pp. 1753–1757, 2010.

[38] A. D. Domínguez-García and C. N. Hadjicostis, Distributed startegies for average consensus in directed graphs, *Proc. of IEEE Conf. on Decision and Control and European Control Conf.*, Dec. 12–15, Orlando, FL, USA, pp. 2124–2129, 2011.

[39] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, Broadcast gossip algorithms for consensus, *IEEE Trans. on Signal Processing*, vol. 57, no. 7, pp. 2748–2761, 2009.

[40] F. Fagnani and S. Zampieri, Average consensus with packet drop communication, *SIAM J. Control and Optimization*, vol. 48, no. 1, pp. 102–133, 2009.

[41] N. Bof, R. Carli, and L. Schenato, Average consensus with asynchronous updates and unreliable communication, *Proc. of* 20^{th} *IFAC World Congress*, July 9–14, Toulouse, France, pp. 601–606, 2017.

[42] M. Huang, Stochastic approximation for consensus: A new approach via ergodic backward products, *IEEE Trans. on Automatic Control*, vol. 57, no. 12, pp. 2994–3008, 2012.

[43] N. Abaid and M. Porfiri, Consensus over numerosity-constrained random networks, *IEEE Trans. on Automatic Control*, vol. 56, no. 3, pp. 649–654, 2011.

[44] L. Xiao, S. Boyd, and S. Lall, A scheme for robust distributed sensor fusion based on average consensus, *Proc. of* 4^{th} *International Symposium on Information Processing in Sensor Networks*, April 25–27, UCLA, Los Angeles, California, USA, pp. 63–70, 2005.

[45] S. Kar, J. M. F. Moura, Sensor networks with random links: topology design for distributed consensus, *IEEE Trans. on Signal Pocessing*, vol. 56, pp. 3315–3326, 2008.

[46] S. Kar, J. M. F. Moura, Distributed average consensus in sensor networks with random link failures and communication channel noise, *Proc. of Conf. Record of the Forty-First Asilomar Conf. on Signals, Syetms and Computers*, pp. 676–680, 2007.

[47] S. Kar and J. M. F. Moura, Distributed average consensus in sensor networks with random link failures, *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 1013–1016, 2007.

[48] S. S. Pereira and A. Pagès-Zamora, Consensus in correlated random wireless sensor networks, *IEEE Trans. Signal Processing*, vol. 59, pp. 6279–6284, 2011.

[49] A. Tahbaz-Salehi and A. Jadbabaie, Consensus over ergodic stationary graph processes, *IEEE Trans. on Automatic Control*, vol. 55, pp. 225–230, 2010.

[50] G. Shi and K. H. Johansson, Randomized optimal consensus of multi-agent systems, *Automatica*, vol. 48, pp. 3018–3030, 2012.

[51] I. Lobel and A. Ozdaglar, Distributed subgradient methods for convex optimization over random networks, *IEEE Trans. on Automatic Control*, vol. 56, no. 6, pp. 1291–1306, 2011.

[52] I. Lobel, A. Ozdaglar, and D. Feiger, Distributed multi-agent optimization with state-dependent communication, *Math. Program. Ser. B*, vol. 129, pp. 255–284, 2011.

[53] D. Wang, D. Wang, W. Wang, Y. Liu, and F. E. Alsaadi, Distributed optimization for multi-agent systems with the first order integrals under markovian switching topologies, *International J. of Systems Science*, vol. 48, pp. 1787–1795, 2017.

[54] J. C. Duchi, A. Agarwal, and M. J. Wainwright, Dual averaging for distributed optimization: convergence analysis and network scaling, *IEEE Trans. on Automatic Control*, vol. 57, pp. 592–606, 2012.

[55] J. Jakovetić, J. M. F. Xavier, and J. M. F. Moura, Convergence rates of distributed Nesterov-like gradient methods on random networks, *IEEE Trans. on Signal Processing*, vol. 62, pp. 868–882, 2014.

[56] T. Wada, I. Masubuchi, K. Hanada, T. Asai, and Y. Fujisaki, Distributed multi-objective optimization over randomly varying unbalanced networks, *Proc. of the* 20^{th} *IFAC World Congress*, July 9–14, Toulouse, France, pp. 2403–2408, 2017.

[57] I. Matei and J. S. Baras, Performance evaluation of the consensus-based distributed subgradient method under random communication topologies, *IEEE J. of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 754–771, 2011.

[58] L. Su, On the convergence rate of average consensus and distributed optimization over unreliable networks, *Proc. of Asilomar Conf. on Signals, Systems, and Computers*, Oct. 28–31, Pacific Grove, CA, USA, pp. 43–47, 2018.

[59] D. Jakovetic, D. Bajovic, A. K. Sahu, and S. Kar, Convergence rates for distributed stochastic optimization over random networks, *Proc. of IEEE Conf. on Decision and Control*, Dec. 17–19, Miami Beach, FL, USA, pp. 4238–4245, 2018.

[60] A. K. Sahu, D. Jakovetic, D. Bajovic, and S. Kar, Distributed zeroth order optimization over random networks: a Kiefer-Wolfowitz stochastic approximation approach, *Proc. of IEEE Conf. on Decision and Control*, Dec. 17–19, Miami Beach, FL, USA, pp. 4951–4958, 2018.

[61] H. Iiduka, Fixed point optimization algorithms for distributed optimization in networked systems, *SIAM J. Optim.*, vol. 23, pp. 1–26, 2013.

[62] R. Carli, G. Notarstefano, L. Schenato, and D. Varagnolo, Distributed quadratic programming under asynchronous and lossy communications via Newton-Raphson consensus, *Proc. of European Control Conf.*, July 15–17, Linz, Austria, pp. 2514–2520, 2015.

[63] N. Bastianello, M. Todescato, R. Carli, L. Schenato, Distributed optimization over lossy networks via relaxed Peaceman-Rachford splitting: a robust ADMM approach, *Proc. of European Control Conf.*, July 12–15, Limassol, Cyprus, pp. 477–482, 2018.

[64] L. Majzoobi, V. Shah-Mansouri, and F. Lahouti, Analysis of distributed ADMM algorithm for consensus optimization over lossy networks, *IET Signal Processing*, vol. 12, pp. 786–794, 2018.

[65] J. Xu, S. Zhu, Y. C. Soh, and L. Xie, A Bregman splitting scheme for distributed optimization over networks, *IEEE Trans. on Automatic Control*, vol. 63, pp. 3809–3824, 2018.

[66] R. Carli, G. Notarstefano, L. Schenato, and D. Varagnolo, Analysis of Newton-Raphson consensus for multi-agent convex optimization under asynchronous and lossy communications, *Proc. of IEEE* 54^{th} *Annual Conf. on Decision and Control*, Dec. 15–18, Osaka, Japan, pp. 418–424, 2015.

[67] T-H. Chang, A randomized dual consensus ADMM method for multi-agent distributed optimization, *Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, April 19–24, Brisbane, Queensland, Australia, pp. 3541–3545, 2015.

[68] N. Bastianello, R. Carli, L. Schenato, and M. Todescato, A partition-based implementation of the relaxed ADMM distributed convex optimization over lossy networks, *Proc. of IEEE Conf. on Decision and Control*, Dec. 17–19, Miami Beach, FL, USA, pp. 3379–3384, 2018.

[69] J. Zhang, K. You, and T. Başar, Distributed discrete-time optimization by exchanging one bit of information, *Proc. of Annual Amer. Cont. Conf.*, June 27–29, Wisconsin Center, Milwaukee, USA, pp. 2065–2070, 2018.

[70] A. Nedić, Asynchronous broadcast-based convex optimization over a network, *IEEE Trans. on Automatic Control*, vol. 56, no. 6, pp. 1337–1351, 2011.

[71] K. Srivastava and A. Nedić, Distributed asynchronous constrained stochastic optimization, *IEEE J. of Selected Topics in Signal Processing*, vol. 5, pp. 772–790, 2011.

[72] B. Touri, A. Nedić, and S. S. Ram, Asynchronous stochastic convex optimization over random networks: error bounds, *Information Theory and Applications Workshop*, Jan. 31-Feb. 5, University of California, San Diego, USA, pp. 1–10, 2010.

[73] P. Bianchi, W. Hachem, and F. Iutzeler, A coordinate descent primal-dual algorithm and application to distributed asynchronous optimization, *IEEE Trans. on Automatic Control*, vol. 61, pp. 2947–2957, 2016.

[74] P. Bianchi, W. Hachem, and F. Iutzeler, A stochastic primal-dual algorithm for distributed asynchronous composite optimization, *Proc. of IEEE Global Conf. on Signal and Information*

Processing, Dec. 3–5, Georgia Tech Hotel & Conference Center, Atlanta, GA, USA, pp. 732–736, 2014.

[75] I. Notarnicola and G. Notarstefano, Asynchronous distributed optimization via randomized dual proximal gradient, *IEEE Trans. on Automatic Control*, vol. 62, pp. 2095–2106, 2017.

[76] F. Iutzeler, P. Bianchi, P. Ciblat, and W. Hachem, Asynchronous distributed optimization using a randomized Alternating Direction Method of Multipliers, *Proc. of 52th Conf. on Decision and Control*, Dec. 10–13, Florence, Italy, pp. 3671–3676, 2013.

[77] I. Notarnicola, R. Carli, and G. Notarstefano, Distributed partitioned big-data optimization via asynchronous dual decomposition, *IEEE Trans. on Control of Network Systems*, vol. 5, pp. 1910–1919, 2018.

[78] F. Zanella, D. Varagnolo, A. Cenedese, G. Pillonetto, and L. Schenato, Asynchronous Newton-Raphson consensus for distributed convex optimization, *Proc. of* 3^{rd} *IFAC Workshop on Distributed Estimation and Control in Networked Systems*, Sep. 14–15, Santa Barbara, CA, USA, pp. 133–138, 2012.

[79] M. Zhong and C. G. Cassandras, Asynchronous distributed optimization with minimal communication and connectivity preservation, *Proc. of the Joint* 48^{th} *Conf. on Decision and Control and* 28^{th} *Chinese Control Conf.*, Dec. 16–18, Shanghai, China, pp. 5396–5401, 2009.

[80] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, Distributed asynchronous deterministic and stochastic gradient optimization algorithms, *IEEE Trans. on Automatic Control*, vol. AC-31, no. 9, pp. 803–812, 1986.

[81] S. Lee and A. Nedić, Distributed random projection algorithm for convex optimization, *IEEE J. of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 221–229, 2013.

[82] M. Zhong and C. G. Cassandras, Asynchronous distributed optimization with event-driven communication, *IEEE Trans. on Automatic Control*, vol. 55, pp. 2735–2750, 2010.

[83] M. Zhong and C. G. Cassandras, Asynchronous distributed optimization with minimal communication, *Proc. of the* 47^{th} *Conf. on Decision and Control*, Dec. 9–11, Cancun, Mexico, pp. 363–368, 2008.

[84] W. Liu and Z. Hua, Asynchronous algorithms for distributed optimisation and application to distributed regression with robustness to outliers, *IET Control Theory and Applications*, vol. 7, pp. 2084–2089, 2013.

[85] G. Shi, K. H. Johansson, and Y. Hong, Reaching an optimal consensus: dynamical systems that compute intersections of convex sets, *IEEE Trans. on Automatic Control*, vol. 58, pp. 610–622, 2013.

[86] A. Simonetto, T. Kevicsky, and R. Babuška, Constrained distributed algebraic connectivity maximization in robotic networks, *Automatica*, vol. 49, pp. 1348–1357, 2013.

[87] Y. Kim and M. Mesbahi, On maximizing the second smallest eigenvalue of a state-dependent graph laplacian, *IEEE Trans. on Autom. Contr.*, vol. 51, pp. 116–120, 2006.

[88] D. D. Siljak, Dynamic graphs, *Nonlin. Analysis: Hybrid Syst.*, vol. 2, pp. 544–567, 2008.

[89] S. Totsch and E. Tadmor, Heterophilious dynamics enhances consensus, *SIAM Review*, vol. 56, pp. 577-621, 2014.

[90] I. Rajapakse, M. Groudine, and M. Mesbahi, Dynamics and control of state-dependent networks for probing genomic organization, *Proc. of Nation. Acad. of Scien. of the USA*, vol. 108, pp. 17257–17262, 2011.

[91] V. Trianni, D. D. Simone, A. Reina, and A. Baronchelli, Emergence of consensus in a multi-robot network: from abstract models to empirical validation, *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 348–353, 2016.

[92] R. Mehmood and J. Crowcroft, Parallel iterative solution method of large sparse linear equation systems, *Technical Report*, University of Cambridge, 2005.

[93] S. Mou and A. S. Morse, A fixed-neighbor, distributed algorithm for solving a linear algebraic equation, *Proc. of European Control Conf.*, July 17-19, Zürich, Switzerland, pp. 2269–2273, 2013.

[94] Y. Shang, A distributed memory parallel Gauss-Seidel algorithm for linear algebraic systems, *Computers and Mathematics with Applications*, vol. 57, pp. 1369–1376, 2009.

[95] G. Shi and B. D. O. Anderson, Distributed network flows solving linear algebraic equations, *Proc. of American Control Conf.*, Boston Marriott Copley Place, July 6-8, Boston, Massachusetts, USA, pp. 2864–2869, 2016.

[96] J. Liu, A. S. Morse, A. Nedić, and T. Başar, Stability of a distributed algorithm for solving linear algebraic equations, *Proc. of* 53^{rd} *IEEE Conf. on Decision and Control*, Dec. 15-17, Los Angeles, California, USA, pp. 3707–3712, 2014.

[97] X. Gao, J. Liu, and T. Başar, Stochastic communication-efficient distributed algorithms for solving linear algebraic equations, *Proc. of IEEE Conf. on Control Applications, part of IEEE Multi-Conference on Systems and Control*, Sept. 19-22, Buenos Aires, Argentina, pp. 380–385, 2016.

[98] J. Liu, X. Chen, T. Başar, and A. Nedić, A continuous-time distributed algorithm for solving linear equations, *Proc. of Amer. Cont. Conf.*, Boston Marriott Copley Place, July 6-8, Boston, MA, USA, pp. 5551–5556, 2016.

[99] S. Mou, A. S. Morse, Z. Lin, L. Wang, and D. Fullmer, A distributed algorithm for efficiently solving linear equations, *IEEE* 54^{th} *Annual Conf. on Decision and Control*, Dec. 15–18, Osaka, Japan, pp. 6791–6796, 2015.

[100] J. Liu, A. S. Morse, A. Nedić, and T. Başar, Exponential convergence of a distributed algorithm for solving linear algebraic equations, *Automatica*, vol. 83, pp. 37–46, 2017.

[101] S. Mou, J. Liu, and A. S. Morse, A distributed algorithm for solving a linear algebraic equation, *Proc. of* 51^{st} *Annual Allerton Conf.*, Allerton House, UIUC, Octob. 2-3, Illinois, USA, pp. 267–274, 2013.

[102] J. Liu, S. Mou, and A. S. Morse, An asynchronous distributed algorithm for solving a linear algebraic equation, *Proc. of* 52^{nd} *IEEE Conf. on Decision and Control*, Dec. 10-13, Florence, Italy, pp. 5409–5414, 2013.

[103] S. Mou, J. Liu, and A. S. Morse, A distributed algorithm for solving a linear algebraic equation, *IEEE Trans. on Automatic Control*, vol. 60, pp. 2863–2878, 2015.

[104] K. You, S. Song, and R. Tempo, A networked parallel algorithm for solving linear algebraic equations, *Proc. of* 55^{th} *IEEE Conf. on Decision and Control*, ARIA Resort & Casino, Dec. 12-14, Las Vegas, USA, pp. 1727–1732, 2016.

[105] X. Wang, S. Mou, and D. Sun, Further discussions on a distributed algorithm for solving linear algebra equations, *Proc. of Amer. Cont. Conf., Sheraton Seattle Hotel*, May 24-26, Seattle, WA, USA, pp. 4274–4278, 2017.

[106] P. Wang, W. Ren, and Z. Duan, Distributed solution to linear equations from arbitrary initializations, *Proc. of Amer. Cont. Conf., Sheraton Seattle Hotel*, May 24-26, Seattle, WA, USA, pp. 3986–3991, 2017.

[107] W. Shen, B. Yin, X. Cao, Y. Cheng, and X. Shen, A distributed secure outsourcing scheme for solving linear algebraic equations in ad hoc clouds, *IEEE Trans. on Cloud Computing*, to appear.

[108] M. Yang and C. Y. Tang, A distributed algorithm for solving general linear equations over networks, *IEEE* 54^{th} *Annual Conf. on Decision and Control*, Dec. 15–18, Osaka, Japan, pp. 3580–3585, 2015.

[109] P. Wang, Y. Gao, N. Yu, W. Ren, J. Lian, and D. Wu, Communication-efficient distributed solutions to a system of linear equations with Laplacian sparse structure, *Proc. of IEEE Conf. on Decision and Control*, Dec. 17–19, Miami Beach, FL, USA, pp. 3367–3372, 2018.

[110] J. Zhou, W. Xuan, S. Mou, and B. D. O. Anderson, Distributed algorithm for achieving minimum l_1 norm solutions of linear equations, *Proc. of Annual Amer. Cont. Conf.*, June 27–29, Wisconsin Center, Milwaukee, USA, pp. 5857–5862, 2018.

[111] J. Liu, S. Mou, and S. Morse, Asynchronous distributed algorithms for solving linear algebraic equations, *IEEE Trans. on Automatic Control*, vol. 63, pp. 372–385, 2018.

[112] X. Wang, S. Mou, and D. Sun, Improvement of a distributed algorithm for solving linear equations, *IEEE Trans. on Industrial Electronics*, vol. 64, no. 4, pp. 3113–3117, 2017.

[113] L. Wang, D. Fullmer, and A. S. Morse, A distributed algorithm with an arbitrary initialization for solving a linear algebraic equation, *Proc. of American Control Conf.*, Boston Marriott Copley Place, July 6-8, Boston, Massachusetts, USA, pp. 1078–1081, 2016.

[114] J. Liu, X. Gao, and T. Başar, A communication-efficient distributed algorithm for solving linear algebraic equations, *Proc. of* 7^{th} *International Conf. on Network Games, Control, and Optimization*, Oct. 29-31, Trento, Italy, pp. 62–69, 2014.

[115] D. Fullmer, L. Wang, and A. S. Morse, A distributed algorithm for computing a common fixed point of a family of paracontractions, *Proc. of* 10^{th} *IFAC Symposium on Nonlinear Control Systems*, August 23-25, Marriott Hotel Monterey, California, USA, pp. 552–557, 2016.

[116] D. Fullmer, J. Li, and A. S. Morse, An asynchronous distributed algorithm for computing a common fixed point of a family of paracontractions, *Proc. of* 55^{th} *IEEE Conf. on Decision and Control*, ARIA Resort & Casino, Dec. 12-14, Las Vegas, USA, pp. 2620–2625, 2016.

[117] P. Weng, W. Ren, and Z. Duan, Distributed minimum weighted norm solution to linear equations associates with weighted inner product, *Proc. of* 55^{th} *IEEE Conf. on Decision and Control*, ARIA Resort & Casino, Dec. 12-14, Las Vegas, USA, pp. 5220–5225, 2016.

[118] B. D. O. Anderson, S. Mou, A. S. Morse, and U. Helmke, Decentralized gradient algorithm for solution of a linear equation, *Numerical Algebra, Control and Optimization*, vol. 6, pp. 319–328, 2016.

[119] S. Mou, Z. Lin, L. Wang, D. Fullmer, and A. S. Morse, A distributed algorithm for efficiently solving linear equations and its applications (Special issue JCW), *System & Control Letters*, vol. 91, pp. 21–27, 2016.

[120] J. Wang and N. Elia, Distributed solution of linear equations over unreliable networks, *Proc. of Amer. Cont. Conf.*, Boston Marriott Copley Place, July 6-8, Boston, MA, USA, pp. 6471–6476, 2016.

[121] J. Wang and N. Elia, Solving systems of linear equations by distributed convex optimization in the presence of stochastic uncertainty, *Proc. of the* 19^{th} *IFAC World Congress*, August 24–29, Cape Town, South Africa, pp. 1210–1215, 2014.

[122] J. Wang and N. Elia, Distributed least square with intermittent communications, *Proc. of Amer. Cont. Conf.*, June 27–29, Fairmont Queen Elizabeth, Montréal, Canada, pp. 6479–4684, 2012.

[123] Z. Liu and C. Li, Distributed sparse recursive least-squares over networks, *IEEE Trans. on Signal Processing*, vol. 62, no. 6, pp. 1386–1395, 2014.

[124] D. E. Marelli and M. Fu, Distributed weighted least-squares estimation with fast convergence for large-scale systems, *Automatica*, vol. 51, pp. 27–39, 2015.

[125] J. Lu and C. Y. Tang, Distributed asynchronous algorithms for solving positive definite linear equations over networks–Part I: agent networks, *Proc. of the First IFAC Workshop on Estimation and Control of Networked Systems*, Sep. 24–26, Venice, Italy, pp. 252–257, 2009.

[126] J. Lu and C. Y. Tang, Distributed asynchronous algorithms for solving positive definite linear equations over networks–Part II: wireless networks, *Proc. of the First IFAC Workshop on Estimation and Control of Networked Systems*, Sep. 24–26, Venice, Italy, pp. 258–263, 2009.

[127] X. Wang, J. Zhou, S. Mou, and M. J. Corless, A distributed linear equation solver for least square solutions, *Proc. of the IEEE* 56th *Annual Conf. on Decision and Control*, Dec. 12–15, Melbourne, Australia, pp.5955–5960, 2017.

[128] J. Lu and C. Y. Tang, A distributed algorithm for solving positive definite linear equations over networks with membership dynamics, *IEEE Trans. on Control of Network Systems*, vol. 5, no. 1, pp. 215–227, 2018.

[129] X. Wang, S. Mou, and B. D. O. Anderson, A distributed algorithm with scalar states for solving linear equations, *Proc. of IEEE Conf. on Decision and Control*, Dec. 17–19, Miami Beach, FL, USA, pp. 2861–2865, 2018.

[130] A. H. Sayed and C. G. Poles, Distributed recursive least-squares over adaptive networks, *Proc. of fortieth Asilomar Conf. on Signals, Systems and Computers*, Oct. 29– Nov. 1, Pacific Grove, CA, USA, pp. 233–237, 2006.

[131] J. Wang, I. S. Ahn, Y. Lu, T. Yang, and G. Staskevich, A distributed least-squares algorithm in wireless sensor networks with limited communication, *Proc. of IEEE Int. Conf. Electro Information Technology*, May 14–17, Lincoln, NE, USA, pp. 467–471, 2017.

[132] S. Huang and C. Li, Distributed sparse total least-squares over networks, *IEEE Trans. on Signal Processing*, vol. 63, pp. 2986–2998, 2015.

[133] G. Mateos, I. D. Schizas, and G. B. Giannakis, Distributed recursive least-squares for consensus-based in-network adaptive estimation, *IEEE Trans. on Signal Processing*, vol. 57, pp. 4583–4588, 2009.

[134] G. Mateos and G. B. Giannakis, Distributed recursive least-squares: stability and performance analysis, *IEEE Trans. on Signal Processing*, vol. 60, pp. 3740–3754, 2012.

[135] R. Z. Has'minskiĭ, *Stability of Differential Equations*, Germantown, MD: Sijthoff & Noordhoff, 1980.

[136] L. Arnold, *Random Dynamical Systems*, Berlin, Heidenberg: Springer-Verlag, 1998.

[137] C. S. Kubrusly and O. L. V. Costa, Mean square stability conditions for discrete stochastic bilinear systems, *IEEE Trans. on Automatic Control*, vol. AC-30, no. 11, pp. 1082–1087, 1985.

[138] A. R. Teel, J. P. Hespanha, and A. Subbaraman, Equivalent characterizations of input-to-state stability for stochastic discrete-time systems, *IEEE Trans. on Automatic Control*, vol. 59, no. 2, pp. 516–522, 2014.

[139] F. Wei and T. Jie, Moment and sample path stability of switched discrete stochastic systems, *Proc. of the* 30^{th} *Chinese Control Conf.*, July 22–24, Yanrai, China, pp. 1810–1814, 2011.

[140] L. Huang, H. Hjalmarsson, and H. Koeppl, Almsot sure stability and stabilization of discrete-time stochastic systems, *Systems & Control Letters*, vol. 82, pp. 26–32, 2016.

[141] S. Sathananthan, M. J. Knap, A. Strong, and L. H. Keel, Robust stability and stabilization of a class of nonlinear discrete time stochastic systems: an LMI approach, *Applied mathematics and Computation*, vol. 219, pp. 1988–1997, 2012.

[142] C. Possieri and A. R. Teel, Asymptotic stability in probability for stochastic Boolean networks, *Automatica*, vol. 83, pp. 1–9, 2017.

[143] A. R. Teel, Lyapunov conditions certifying stability and recurrence for a class of stochastic hybrid systems, *Annual Reviews in Control*, vol. 37, pp. 1–24, 2013.

[144] Y. A. Phillis, On the stabilization of discrete linear time-varying stochastic systems, *IEEE Trans. on Syst. Man Cyber.*, vol. SMC-12, no. 3, pp. 415–417, 1982.

[145] W. Yijing, S. Yunna, and Z. Zhiqiang, Stochastic stability and stabilization of discrete-time markovian jump systems in the presence of incomplete knowledge of transition probabilities, *Proc. of the* 30^{th} *Chinese Control Conf.*, July 22–24, Yantai, China, pp. 1395–1399, 2011.

[146] T. Hou and H. Ma, Stability analysis of discrete-time stochastic systems with infinite markov jump parameter, *Proc. of Amer. Cont. Conf.*, palmer House Hilton, July 1–3, Chicago, IL, USA, pp. 4192–4197, 2015.

[147] Z. Zuo, Y. Liu, Y. Wang, and H. Li, Finite-time stochastic stability and stabilisation of linear discrete-time Markovian jump systems with partly unknown transition probabilities, *IET Control Theory and Applications*, vol. 6, pp. 1522–1526, 2012.

[148] O. L. V. Costa and D. Z. Figueiredo, Stochastic stability of jump discrete-time linear systems with Markov chain in a general Borel space, *IEEE Trans. on Automatic Control*, vol. 59, no. 1, pp. 223–227, 2014.

[149] C. A. C. Gonzaga and O. L. V. Costa, Stochastic stability for discrete-time Markov jump Lur'e systems, *Proc. of* 52^{nd} *IEEE Conf. on Decision and Control*, Dec. 10–13, Florence, Italy, pp. 5993–5998, 2013.

[150] Z. Zuo, H. Li, Y. Wang, and Y. Liu, Robust finite-time stochastic stability analysis and control synthesis of uncertain discrete-time Markovian jump linear systems, *Proc. of the* 10^{th} *World Congress on Intelligent Control and Automation*, July 6–8, Beijing, China, pp. 1925–1929, 2012.

[151] H. Linlin, Z. Guangdeng, and W. Yuqiang, Exponential $l_2 - l_\infty$ stochastic stability analysis for discrete-time switching Markov jump linear systems, *Proc. of the* 30^{th} *Chinese Control Conf.*, July 22–24, Yantai, China, pp. 1722–1727, 2011.

[152] C. A. C. Gonzaga and O. L. V. Costa, Stochastic stabilization and induced l_2-gain for discrete-time Markov jump Lur'e systems with control saturation, *Automatica*, vol. 50, pp. 2397–2404, 2014.

[153] U. Vaidya, Stochastic stability analysis of discrete-time system using Lyapunov measure, *Proc. of Amer. Cont. Conf.*, Palmer House Hilton, July 1–3, Chicago, Il, USA, pp. 4646–4651, 2015.

[154] U. Vaidya and V. Chinde, Computation of the Lyapunov measure for almost everywhere stochastic stability, *Proc. of the* 54^{th} *Annual Conf. on Decision and Control*, Dec. 15–18, Osaka, Japan, pp. 7042–7047, 2015.

[155] A. R. Teel, J. P. Hespanha, and A. Subbaraman, A converse Lyapunov theorem and robustness for asymptotic stability in probability, *IEEE Trans. on Autom. Cont.*, vol. 59, no. 9, pp. 2426–2441, 2014.

[156] W. Zhang, X. Lin, and B-S. Chen, LaSalle-type theorem and its applications to infinite horizon optimal control of discrete-time nonlinear stochastic systems, *IEEE Trans. on Automatic Control*, vol. 62, no. 1, pp. 250–261, 2017.

[157] A. Olshevsky and J. Tsitsiklis, On the nonexistence of quadratic Lyapunov functions for consensus algorithms, *IEEE Trans. on Automatic Control*, vol. 53, no. 11, pp. 2642–2645, 2008.

[158] J. M. Davis and G. Eisenbarth, On positivstellensatz and nonexistence of common quadratic Lyapunov functions, *Proc. of IEEE* 43^{th} *Southeastern Symposium on System Theory*, pp. 55–58, 2011.

[159] R. H. Ordóñez-Hurtado and M. A. Duarte-Mermoud, A methodology for determining the non-existence of common quadratic Lyapunov functions for pairs of stable systems, *Proc. of* 5^{th} *International Conf. on Genetic and Evolutionary Computing*, August 29–September 01, Kitakyushu Convention Center, Kitakyushu, Japan, pp. 127–130, 2011.

[160] M. A. Duarte-Mermoud, R. H. Ordóñez-Hurtado, and P. Zagalak, A method for determining the non-existence of common quadratic Lyapunov function for switched linear systems based on particle swarm optimisation, *International J. of Systems Science*, vol. 43, no. 11, pp. 2015–2029, 2012.

[161] L. Wang and Z. Xu, On characterizations of exponential stability of nonlinear discrete dynamcial systems on bounded regions, *IEEE Trans. on Autom. Cont.*, vol. 52, no. 10, pp. 1871–1881, 2007.

[162] Y. Fu, Q. Zhao, and L. Wang, A converse to global exponential stability of a class of difference equations, *Advances in Difference Equations*, vol. 9, pp. 1–10, 2013.

[163] F. Cucker and S. Smale, Emergent behavior in flocks, *IEEE Trans. on Automatic Control*, vol. 52, pp. 852–862, 2007.

[164] S. Sh. Alaviani, A necessary and sufficient condition for delay-independent stability of linear time-varying neutral delay systems, *J. of The Franklin Institute*, vol. 351, pp. 2574–2581, 2014.

[165] S. Sh. Alaviani, Delay-dependent exponential stability of linear time-varying neutral delay systems, *Proc. of the* 12^{th} *IFAC Workshop on Time Delay Systems*, June 28-30, University of Michigan, Ann Arbor, Michigan, USA, pp. 177–179, 2015.

[166] D. Bandyopadhyay, On a study of some results on the theory of fixed point of operators, Ph.D. dissertation, Dep. of Pure Math., University of Calcutta, 1997.

[167] D. R. Smart, *Fixed point theorems*, Cambridge Univ. Press, London, 1974.

[168] W. Rudin, *Functional Analysis*, McGraw-Hill: Singapore, 1991, second edition.

[169] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[170] H-K. Xu, Some random fixed point theorems for condensing and nonexpansive operators, *Proceedings of The American Mathematical Society*, vol. 110, no. 2, pp. 395–400, 1990.

[171] M. A. Krasnoselskii, Two remarks on the method of successive approximations, *Uspekhi Mat. Nauk*, vol. 10, pp. 123–127, 1955.

[172] W. R. Mann, Mean value methods in iteration, *Proc. of the Amer. Math. Soc.*, vol. 4, pp. 506–510, 1953.

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

[174] A. Kumar and S. Rathee, Some common fixed point and invariant approximation results for nonexpansive mappings in convex metric space, *Fixed Point Theory and Appl.*, vol. 182, pp. 1–14, 2014.

[175] R. Espinola, P. Lorenzo, and A. Nicolae, Fixed points, selections and common fixed points for nonexpansive mappings, *J. Math. Anal. Appli.*, vol. 382, pp. 503–515, 2011.

[176] I. Bula, Strictly convex metric spaces and fixed points, *Mathematica Moravica*, vol. 3, pp. 5–16, 1999.

[177] K. G. S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mapping*, New York: Marcel Dekker, 1984.

[178] G. C. Calafiore and M. C. Campi, The scenario approach to robust control design, *IEEE Trans. Automatic Control*, vol. 51, pp. 742–753, 2006.

[179] G. Clafiore and M. C. Campi, Robust convex program: randomized solutions and applications in control, *Proc. of* 42^{nd} *IEEE Conf. on Decision and Control*, Dec. 9–12, Maui, Hawaii, USA, pp. 2423–2428, 2003.

[180] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized Alorithms for Analysis and Control of Uncertain Systems with Applications*, Springer–Verlag: London, 2013.

[181] K. Yang, J. Huang, Y. Wu, X. Wang, and M. Chiang, Distributed robust optimization (DRO) part I: framework and example, *Optim. Eng.*, vol. 15, pp. 35–67, 2014.

[182] M. Bürger, G. Notarstefano, and F. Allgöwer, Distributed robust optimization via cutting-plane consensus, *Proc. of IEEE Conf. on Decision and Control*, Dec. 10–13, Maui, Hawaii, USA, pp. 7457–7463, 2012.

[183] M. Bürger, G. Notarstefano, and F. Allgöwer, A polyhedral approximation framework for convex and robust distributed optimization, *IEEE Trans. on Automatic Control*, vol. 59, pp. 384–395, 2014.

[184] L. Carlone, V. Srivastava, F. Bullo, and G. C. Calafiore, Distributed random convex programming via constraint consensus, *SIAM J. Control Optim.*, vol. 52, pp. 629–662, 2014.

[185] S. Wang and C. Li, Distributed random optimization in networked system, *IEEE Trans. on Cybernetics*, vol. 47, pp. 2321–2333, 2017.

[186] K. You and R. Tempo, Parallel algorithms for robust convex programs over networks, *Proc. of Amer. Cont. Conf.*, July 6–8, Boston Marriott Copley Place, Boston, MA, USA, pp. 2017–2022, 2016.

[187] K. You, R. Tempo, and P. Xie, Distributed algorithms for robust convex optimization via the scenario approach, *IEEE Trans. on Automatic Control*, vol. 64, pp. 880–895, 2019.

[188] A. Ben-Tal and A. Nemirovski, Robust convex optimization, *Mathematics of Operations Research*, vol. 23, pp. 769–805, 1998.

[189] V. Jeyakumar and G. Y. Li, Strong duality in robust convex programming: complete characterizations, *SIAM J. Optim.*, vol. 20, pp. 3384–3407, 2010.

[190] J. H. Lee and G. M. Lee, On approximate solutions for robust convex optimization problems, www.kurims.kyoto-u.ac.jp/ kyodo/kokyuroku/contents/pdf/2011-02.pdf

[191] K. Cai, W. Li, F. Ju, and X. Zhu, A scenario-based optimization approach to robust estimation of airport apron capacity, *Integrated Communications, Navigation, Surveillance Conf.,* April 10-12, Herndon, VA, USA, pp. 3A1-1–3A1-8, 2018.

[192] B. T. Polyak, Random algorithms for solving convex inequalities, *Studies in Computational Mathematics*, vol. 8, pp. 409–422, 2001.

[193] T. Wada and Y. Fujisaki, Sequential randomized algorithms for robust convex optimization, *IEEE Transactions on Automatic Control*, vol. 60, pp. 3356–3361, 2015.

[194] N. Ho-Nguyen and F. Kilinc-Karzan, Online first-order framework for robust convex optimization, *Operation Research*, vol. 66, pp. 1670–1692, 2018.

[195] C. Kroer, N. Ho-Nguyen, G. Lu, and F Kilinc-Karzan, Performance evaluation of iterative methods for solving robust convex quadratic problems, 10^{th} *NIPS Workshop on Optimization for Machine Learning*, Dec. 8, Long Beach, USA, 2017.

[196] A. Mutapcic and S. Boyd, Cutting-set methods for robust convex optimization with pessimizing oracles, *Optimization Methods & Software*, vol. 24, pp. 381–406, 2009.

[197] K. Margellos, A. Falsone, S. Garatti, and M. Prandini, Distributed constrained optimization and consensus in uncertain networks via proximal minimization, *IEEE Trans. on Automatic Control*, vol. 63, pp. 1372–1387, 2018.

[198] F. Alismail, P. Xiong, and C. Singh, Optimal wind farm allocation in multi-area power systems using distributionally robust optimization approach, *IEEE Trans. on Power Systems*, vol. 33, pp. 536–544, 2018.

[199] D. Apostolopoulou, Z. De Greve, and M. McCulloch, Robust optimization for Hydroelectric system operation under uncertainty, *IEEE Trans. on Power Systems*, vol. 33, pp. 3337–3348, 2018.

[200] C. Wang, J. Zhu, and T. Zhu, Decentralized robust optimization for real-time dispatch of power system based on approximate dynamic programming, *Int. Conf. on Power System Tech.*, Nov. 6–8, Guangzhou, China, pp. 1935–1941, 2018.

[201] R. Zhang and B. Zeng, Ambulance deployment with relocation through robust optimization, *IEEE Trans. on Automation Science and Engineering*, vol. 16, pp. 138–147, 2019.

[202] D. Yazdani, T. T. Nguyen, and J. Branke, Robust optimization over time by learning problem space characteristics, *IEEE Trans. on Evolutionary Computation*, vol. 23, pp. 143–155, 2019.

[203] M. Boulekchour, N. Aouf, and M. Richardson, Robust L_∞ convex optimization for monocular visual odometry trajectory estimation, *Robotica*, vol. 34, pp. 703–722, 2016.

———————

[204] B. Gharesifard and G. Cortés, When does a digraph admit a doubly stochastic adjacency matrix, *Proc. of Amer. Control Conf.*, Marriott Waterfront, Baltimore, MD, USA, June 30–July 2, pp. 2440–2445, 2010.

[205] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[206] C. Fischione, Fast-Lipschitz Optimization with Wireless Sensor Networks Applications, *IEEE Trans. on Automatic Control*, vol. 56, pp. 2319–2331, 2011.

[207] M. Jakobsson and C. Fischione, Extensions of Fast-Lipschitz optimization for convex and nonconvex problems, *Proc. of the* 3^{rd} *IFAC Workshop on Distributed estimation and Control in Networked Systems*, Sep. 14-15, Santa Barbara, California, USA, pp. 162–167, 2012.

[208] M. Jakobsson, S. Magnusson, C. Fischione, and P. C. Weeraddana, Extensions of Fast-Lipschitz optimization, *IEEE Trans. on Automatic Control*, vol. 61, pp. 861–876, 2016.

[209] H. Iiduka, Acceleration method for convex optimization over the fixed point set of a nonexpansive mapping, *Math. Program., Ser. A*, vol. 149, pp. 131–165, 2015.

[210] H. Iiduka, Convex optimization over fixed point sets of quasi-nonexpansive and nonexpansive mappings in utility-based bandwidth allocation problems with operational constraints, *J. Comput. Appl. Math.*, vol. 282, pp. 225–236, 2015.

[211] H. Iiduka, Fixed point optimization algorithm and its applications to network bandwidth allocation, *J. Comput. Appl. Math.*, vol. 236, pp. 1733–1742, 2012.

[212] K. Slavakis, J. Yamada, and K. Sakaniwa, Computation of symmetric positive definite Toeplitz matrices by the hybrid steepest descent method, *Signal Processing*, vol. 83, pp. 1135–1140, 2003.

[213] K. Slavakis and I. Yamada, Robust wideband beamforming by the hybrid steepest descent method, *IEEE Trans. on Signal Processing*, vol. 55, pp. 4511-4532, 2007.

[214] I. Yamada, N. Ogura, and N. Shirakawa, A numerical robust hybrid steepest descent method for the convexly constrained generalized inverse problems, *Contemp. Math.*, vol. 313, pp. 269–305, 2002.

[215] R. Durrett, *Probability: Theory and Examples*, New York: Cambridge Univ. Press, 2010, forth edition.

[216] H. K. Xu, Iterative algorithms for nonlinear operators, *J. London Math. Soc.*, vol. 66, pp. 240-256, 2002.

[217] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, John Wiley & Sons, Inc., USA, 1984.

[218] P. Tseng, On the convergence of the products of firmly nonexpansive mappings, *SIAM J. Optimization*, vol. 2, no. 3, pp. 425–434, 1992.

[219] I. Cioranescu, *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems*, Kluwer Academic Publishers: The Netherlands, 1990.

[220] H. H. Bauschke and J. M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review*, vol. 38, pp. 367–426, 1996.

[221] H. H. Bauschke and P. L. Combettes, *Covex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer: New York, 2011.

[222] S. Agmon, The relaxation method for linear inequalities, *Canadian Journal of Mathematics*, vol. 6, pp. 382–392, 1954.

[223] T. S. Motzkin and I. J. Schoenberg, The relaxation method for linear inequalities, *Canadian Journal of Mathematics*, vol. 6, pp. 393–404, 1954.

[224] I. M. Terfaloaga, The convex feasibility problem and Mann-type iteration, *Analele Universităţii "Eftimie Murgu" Reşiţa*, vol. 24, pp. 402–410, 2017.

[225] O. Slučiak and M. Rupp, Consensus algorithm with state-dependent weights, *IEEE Trans. on Signal Processing*, vol. 64, pp. 1972–1985, 2016.

[226] A. Bogojeska, M. Mirchev, I. Mishkovski, and L. Kocarev, Synchronization and consensus in state-dependent networks, *IEEE Trans. on Circuits and Systems-I: Regular Papers*, vol. 61, pp. 522–529, 2014.

[227] A. Awad, A. Chapman, E. Schoof, A. Narang-Siddarth, and M. Mesbahi, Time-scale separation on networks: consensus, tracking, and state-dependent interactions, *IEEE* 54^{th} *Annual Conf. on Decision and Control*, Dec. 15-18, Osaka, Japan, pp. 6172–6177, 2015.

[228] S. L. Campbell and N. J. Rose, Singular perturbation of autonomous linear systems, *SIAM J. Math. Anal.*, vol. 10, pp. 542–551, 1979.

[229] D. S. Bernstein and S. P. Bhat, Lyapunov stability, semi-stability, and asymptotic stability of matrix second-order systems, *ASME Trans. J. Vibr. Acoustics*, vol. 117, pp. 145–153, 1995.

[230] S. P. Bhat and D. S. Bernstein, Nontangency-based Lyapunov tests for convergence and stability in systems having a continuum of equilibria, *SIAM J. Control Optim.*, vol. 42, pp. 1745–1775, 2003.

[231] S. P. Bhat and D. S. Bernstein, Arc-length-based Lyapunov test for convergence and stability in systems having a continuum of equilibria, *Proc. of the Amer. Cont. Conf.*, June 4–6, Denver, CO, USA, pp. 2961–2966, 2003.

[232] Q. Hui, W. M. Hadda, and S. P. Bhat, Finite-time semistability and consensus for nonlinear dynamical networks, *IEEE Trans. on Automatic Control*, vol. 53, pp. 1887–1900, 2008.

[233] Q. Hui, W. M. Hadda, and S. P. Bhat, On robust control algorithms for nonlinear network consensus protocols, *Int. J. Robu. Nonlin. Cont.*, vol. 20, pp. 269–284, 2010.

[234] Q. Hui, W. M. Hadda, and S. P. Bhat, Lyapunov and converse Lyapunov theorems for semi-stability, *Proc. of IEEE Conf. on Decision and Control*, Dec. 12–14, New Orleans, LA, USA, pp. 5870–5874, 2007.

[235] Q. Hui, Convergence and stability analysis for iterative dynamics with application in balanced resource allocation: a trajectory distance Lyapunov approach, *Proc. of IEEE Conf. on Decision and Control and European Control Conf.*, Dec. 12–15, Orlando, FL, USA, pp. 7218–7223, 2011.

[236] Q. Hui and W. M. Haddad, Distributed nonlinear control algorithms for network consensus, *Automatica*, vol. 44, pp. 2375–2381, 2008.

[237] Q. Hui, Lyapunov-based semistability analysis for discrete-time switched network systems, *Proc. of Amer. Cont. Conf. on O'Farrell Street*, June 29–July 01, San Francisco, CA, USA, pp. 2602–2606, 2011.

[238] J. Shen, J. Hu, and Q. Hui, Semistability of switched linear systems with application to PageRank algorithms, *European J. of Control*, vol. 20, pp. 132–140, 2014.

[239] Q. Hui and H. Zhang, Optimal balanced coordinated network resource allocation using swarm optimization, *IEEE Trans. on Systems, Man, and Cybernetics: Systems*, vol. 45, pp. 770–787, 2015.

[240] J. Shen, J. Hu, and Q. Hui, Semistability of switched linear systems with applications to distributed sensor networks: A generating function approach, *Proc. of* 50^{th} *IEEE Conf. on Decision and Control and European Control Conf.*, Dec. 12–15, Orlando, FL, USA, pp. 8044–8049, 2011.

[241] Q. Hui, Convergence and stability analysis for iterative dynamics with application to compartmental networks: A trajectory distance based Lyapunov approach, *J. of The Franklin Institute*, vol. 350, pp. 679–697, 2013.

[242] J. Zhou and Q. Wang, Stochastic semistability with application to agreement problems over random networks, *Proc. of Amer. Cont. Conf.*, Marriott Waterfront, Baltimore, MD, USA, June 30-July 02, pp. 568–573, 2010.

[243] T. Rajpurohit and W. M. Haddad, Stochastic thermodynamics: a dynamical systems approach, *Entropy*, vol. 19, pp. 1–48, 2017.

[244] T. Rajpurohit and W. M. Haddad, Lyapunov and converse Lyapunov theorems for stochastic semistability, *Systems & Control Letters*, vol. 97, pp. 83–90, 2016.

[245] H. J. Kushner, *Introduction to Stochastic Control*, Reinhart and Winston: New York, 1971.

[246] K. Liu, *Stability of Infinite Dimensional Stochastic Differential Equations with Applications*, vol. 135 of Monographs and Surveys in pure and Applied Mathematics, Chapman & Hall/CRC, 2006.

[247] S. Sh. Alaviani and N. Elia, Distributed multi-agent convex optimization over random digraphs, *Proceedings of American Control Conference*, May 24–26, Sheraton Seattle Hotel, Seattle, USA, pp. 5288–5293, 2017.

[248] S. Sh. Alaviani and N. Elia, Distributed multi-agent convex optimization over random digraphs, *IEEE Transactions on Automatic Control*, to appear

[249] S. Sh. Alaviani and N. Elia, Distributed convex optimization with state-dependent interactions, *Proceedings of* 25^{th} *Mediterranean Conference on Control and Automation*, July 3–6, Valletta, Malta, pp. 129–134, 2017.

[250] S. Sh. Alaviani and N. Elia, A distributed algorithm for solving linear algebraic equations over random networks, *Proceedings of IEEE Conference on Decision and Control*, Dec. 17–19, Miami Beach, FL, USA, pp. 83–88, 2018.

[251] S. Sh. Alaviani and N. Elia, Distributed average consensus over random networks, *Proceedings of American Control Conference*, July 10–12, Philadelphia, USA, 2019, to appear