ARTICLE IN PRESS
JID: JARAP [m5G;May 22, 2019;12:34]
Annual Reviews in Control xxx (xxxx) xxx

Contents lists available at ScienceDirect

Annual Reviews in Control

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

Review

A survey of distributed optimization☆

Tao Yang[a], Xinlei Yi[b], Junfeng Wu[c], Ye Yuan[d,*], Di Wu[e], Ziyang Meng[f], Yiguang Hong[g], Hong Wang[h], Zongli Lin[i], Karl H. Johansson[b]

[a] *Department of Electrical Engineering, University of North Texas, Denton, TX 76203, USA*
[b] *Division of Decision and Control Systems, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, 100 44, Sweden*
[c] *College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China*
[d] *School of Artificial Intelligence and Automation and State Key Lab of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China*
[e] *Pacific Northwest National Laboratory, Richland, WA 99352, USA*
[f] *Department of Precision Instrument and the State Key Laboratory of Precision Measurement Technology and Instruments, Tsinghua University, Beijing 100084, China*
[g] *Key Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Science, Beijing 100190, China*
[h] *Oak Ridge Laboratory, Oak Ridge, TN 37932, USA*
[i] *Charles L. Brown Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22904, USA*

A R T I C L E I N F O

Article history:
Received 14 January 2019
Revised 10 April 2019
Accepted 13 May 2019
Available online xxx

Keywords:
Distributed optimization
Coordination of distributed energy resources

A B S T R A C T

In distributed optimization of multi-agent systems, agents cooperate to minimize a global function which is a sum of local objective functions. Motivated by applications including power systems, sensor networks, smart buildings, and smart manufacturing, various distributed optimization algorithms have been developed. In these algorithms, each agent performs local computation based on its own information and information received from its neighboring agents through the underlying communication network, so that the optimization problem can be solved in a distributed manner. This survey paper aims to offer a detailed overview of existing distributed optimization algorithms and their applications in power systems. More specifically, we first review discrete-time and continuous-time distributed optimization algorithms for undirected graphs. We then discuss how to extend these algorithms in various directions to handle more realistic scenarios. Finally, we focus on the application of distributed optimization in the optimal coordination of distributed energy resources.

Contents

☆ This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).
* Corresponding author.
E-mail addresses: Tao.Yang@unt.edu (T. Yang), xinleiy@kth.se (X. Yi), jfwu@zju.edu.cn (J. Wu), yye@hust.edu.cn (Y. Yuan), di.wu@pnnl.gov (D. Wu), ziyangmeng@mail.tsinghua.edu.cn (Z. Meng), yghong@iss.ac.cn (Y. Hong), wangh6@ornl.gov (H. Wang), zl5y@virginia.edu (Z. Lin), kallej@kth.se (K.H. Johansson).

https://doi.org/10.1016/j.arcontrol.2019.05.006
1367-5788/

Please cite this article as: T. Yang, X. Yi and J. Wu et al., A survey of distributed optimization, Annual Reviews in Control, https://doi.org/10.1016/j.arcontrol.2019.05.006

1. Introduction

In recent years, rapid developments in digital systems, communication and sensing technologies have led to the emergence of networked systems. These networked systems consist of a large number of interconnected subsystems (agents), which are required to cooperate in order to achieve a desirable global objective. Applications for such networked systems include but are not limited to power systems, sensor networks, smart buildings, and smart manufacturing (Bullo, 2017; Dong, Hua, Zhou, Ren, & Zhong, 2018; Dörfler, Simpson-Porco, & Bullo, 2016; Gao, Gao, Ozbay, & Jiang, 2019; Lamnabhi-Lagarrigue et al., 2017; Meng & Moore, 2016; Ren & Cao, 2011; Yuan, Zhang, Wu, Zhu, & Ding, 2017; Zhu & Martínez, 2015). Many problems in networked systems can be posed in the framework of convex optimization (Boyd, Parikh, Chu, Peleato, & Eckstein, 2011). Due to the distributed nature of networked systems, the traditional centralized strategies are not suitable to solve these optimization problems. Moreover, the centralized framework is subject to performance limitations, such as a single point of failure, high communication requirement, substantial computation burden, and limited flexibility and scalability. All of these have made imperative the use of distributed approaches to solve these optimization problems (Boyd et al., 2011; Nedić, 2015; Sayed, 2014a).

1.1. Distributed optimization

In a networked system of N agents, each of which has a local private convex objective function $f_i(x)$, where $x \in \mathbb{R}^n$ is the optimization variable. The objective of distributed optimization is to minimize a global objective function, which is a sum of the objective functions of all agents:

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^{N} f_i(x), \tag{1}$$

in a distributed manner by local computation and communication.

The distributed optimization problem has been studied for a long time and can be traced back to the seminal works (Bertsekas & Tsitsiklis, 1989; Tsitsiklis, 1984; Tsitsiklis, Bertsekas, & Athans, 1986) in the context of parallel and distributed computation. It has gained growing renewed interest over the last decade due to its various applications in power systems, communication networks, machine learning, and sensor networks, just to name a few (Cao, Yu, Ren, & Chen, 2013; Nedić, 2015; Nedić, Olshevsky, & Rabbat, 2018; Sayed, 2014a). Various distributed algorithms have been proposed in the literature. For the recent review and progress, please refer to the surveys (Boyd et al., 2011; Nedić, 2015; Nedić & Liu, 2018; Nedić et al., 2018; Sayed, 2014a; Yang & Johansson, 2010) and the books (Bullo, 2017; Giselsson, 2018; Ren & Cao, 2011; Zhu & Martínez, 2015).

Most existing distributed algorithms are consensus based and can be divided into two categories depending on whether an algorithm is in discrete-time or in continuous-time. In these algorithms, each agent holds a dynamic state, which is the estimate of the optimization variable, and updates its value based on its own information and the information received from its neighbors through the underlying communication network.

1.2. Coordination of distributed energy resources

Although distributed optimization finds various applications, in this paper, we focus on the optimal coordination of distributed energy resources (DERs) in power systems (see, e.g., Bidram, Lewis, & Davoudi, 2014; Hadjicostis, Domínguez-García, & Charalambous, 2018; Kraning, Chu, Lavaei, & Boyd, 2014; Molzahn et al., 2017; Nedić & Liu, 2018; Qin, Ma, Shi, & Wang, 2017).

In the past decades, the power system has been undergoing a transition from a system with conventional generation power plants and inflexible loads to a system with a large number of distributed generators, energy storages, and flexible loads, often referred to as distributed energy resources (DERs) (Lasseter et al., 2003; Pedrasa, Spooner, & MacGill, 2010; Rahimi & Ipakchi, 2010). These resources are small and highly flexible compared with conventional generators and can be aggregated to provide power necessary to meet the regular demand. As the electricity grid continues to modernize, DERs can facilitate the transition to a smarter grid.

In order to achieve an effective and efficient deployment among DERs, one needs to properly design the coordination among them. The objective of the optimal DER coordination problem is to minimize the total production cost while meeting the total demand and satisfying the individual generator output limits. One approach is through a completely centralized control strategy, which suffers from the limitations as pointed out before. To overcome these limitations, recently, by using the results developed in the field of distributed optimization, various distributed strategies have been proposed for solving the optimal DER coordination problem, see, e.g., Dominguez-Garcia, Cady, and Hadjicostis (2012); Kar and Hug (2012); Rahbari-Asr, Ojha, Zhang, and Chow (2014); Xing, Mou, Fu, and Lin (2015); Yang, Tan, and Xu (2013); Yang, Wu, Sun, and Lian (2016); Zhang and Chow (2012); Zhao, He, Cheng, and Chen (2017) and Hadjicostis et al. (2018).

1.3. Contributions and outline

Recent survey papers on distributed optimization (Nedić, 2015; Nedić & Liu, 2018; Nedić et al., 2018) mainly focused on discrete-time algorithms with diminishing step-sizes. This paper aims to provide a detailed and comprehensive overview of discrete-time algorithms with both diminishing step-sizes and fixed step-sizes as well as continuous-time algorithms. In addition, we also discuss their applications to the optimal DER coordination problem, and offer some future research directions.

In Section 2, we provide some background on graph theory, convex analysis, and convergence analysis.

In Section 3, we review various basic discrete-time optimization algorithms as well as continuous-time optimization algorithms for undirected graphs. Regarding discrete-time algorithms, we focus on the recently developed algorithms with fixed step-sizes for the case where the local objective functions are strongly convex and smooth (have Lispchitz continuous gradients). Compared to algorithms with diminishing step-sizes, these algorithms have faster convergence, which is desirable in practical time critical applications, such as the DER coordination (Tang, Hill, & Liu, 2018). Moreover, due to the faster convergence, their computational costs and communication overheads are much smaller compared to the algorithms with diminishing step-sizes.

With these discrete-time and continuous-time algorithms in hand, in Section 4, we discuss various extensions of these algorithms to handle more realistic scenarios. These extensions are not independent but actually may overlap to some extent.

In Section 5, we focus on the application of distributed optimization to DER coordination, which has received substantial attention in both control and power system fields in recent years (see, e.g.,Bidram et al., 2014; Hadjicostis et al., 2018; Kraning et al., 2014; Molzahn et al., 2017; Nedić & Liu, 2018; Qin et al., 2017). The specific requirements and structures for power systems application provide more challenges and opportunities for the development of distributed optimization algorithms. As such, various general distributed optimization algorithms reviewed in Sections 3 and 4 have been applied and/or adapted to solve the DER coordination problem in power systems.

Finally, Section 6 concludes the paper with some discussion on future research directions.

Notations: Given a matrix A, $A^\top$ denotes its transpose and A^{-1} denotes its inverse. We denote by $A\otimes B$ the Kronecker product between matrices A and B. I_n denotes the identity matrix of dimension $n\times n$. $\mathbf{1}_n$ and $\mathbf{0}_n$ denote the column vector with each entry being 1 and 0, respectively.

2. Preliminaries

This section introduces some background on graph theory, convex analysis, and convergence analysis.

2.1. Graph theory

We first recall some basic concepts in graph theory (Godsi & Royle, 2001). Let $\mathcal{G}=(\mathcal{V},\mathcal{E})$ denote a directed graph (digraph) with the set of nodes (agents) $\mathcal{V}=\{1,\dots,N\}$ and the set of edges $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$. A directed edge from node i to node j is denoted by $(i,j)\in\mathcal{E}$. A digraph is undirected if and only $(i,j)\in\mathcal{E}$ implies $(j,i)\in\mathcal{E}$. For notational simplification, we assume that the digraph does not have any self loop, i.e., $(i,i)\notin\mathcal{E}$ for all $i\in\mathcal{V}$ although each node i has an access to its own information. A directed *path* from node i_1 to node i_k is a sequence of nodes $\{i_1,\dots,i_k\}$ such that $(i_j,i_{j+1})\in\mathcal{E}$ for $j=1,\dots,k-1$. If there exists a directed path from node i to node j, then node j is said to be reachable from node i. A digraph $\mathcal{G}$ is *strongly connected* if every node is reachable from every other node. Let $A=[a_{ij}]\in\mathbb{R}^{N\times N}$ be adjacency matrix associated with the digraph $\mathcal{G}$, where $a_{ij}>0$ is the weight for edge $(j,i)\in\mathcal{V}$ and $a_{ij}=0$ otherwise. A digraph is *weigh-balanced* if $\sum_{j=1}^N a_{ij}=\sum_{j=1}^N a_{ji}$ for all $i\in\mathcal{V}$. A digraph is *detailed balanced* if there exist some real numbers $\omega_i>0$, $i=1,2,\dots,N$, such that the coupling weights of the graph satisfy $\omega_i a_{ij}=\omega_j a_{ji}$ for all $i,j=1,2,\dots,N$.

The underlying communication network may also be modeled as a time-varying directed graph $\mathcal{G}(k)=(\mathcal{V},\mathcal{E}(k))$, where the edge set changes over time due to unexpected loss of communication links. All nodes which transmit information to node i directly at time t are said to be its in-neighbors and belong to the set $\mathcal{N}_i^{\text{in}}(k)=\{j\in\mathcal{V}\mid(j,i)\in\mathcal{E}(k)\}$. All nodes which receive information from agent i at time k belong to the set of its out-neighbors, denoted by $\mathcal{N}_i^{\text{out}}(k)=\{j\in\mathcal{V}\mid(i,j)\in\mathcal{E}(k)\}$. The joint graph of $\mathcal{G}(k)$ in the time interval $[k_1,\,k_2)$ with $k_1<k_2\le\infty$ is defined as $\mathcal{G}([k_1,k_2))=\cup_{k\in[k_1,k_2)}\mathcal{G}(k)=(\mathcal{V},\cup_{k\in[k_1,k_2)}\mathcal{E}(k))$. A time-varying directed graph $\mathcal{G}(k)$ is said to be *uniformly jointly strongly connected* if there exists a constant $B>0$ such that $\mathcal{G}([k_0,k_0+B))$ is strongly connected for any $k_0\ge 0$.

2.2. Convex analysis

Next, we provide some background on basic convex analysis (Boyd & Vandenberghe, 2004; Nesterov, 2004). A set $X \subset \mathbb{R}^n$ is *convex* if for all $x, y \in X$, and for all $\theta \in [0, 1]$, we have $\theta x + (1-\theta)y \in X$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$ for all $x, y \in \mathbb{R}^n$ and for all $\theta \in (0, 1)$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *strictly convex* if $f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y)$ for all $x \neq y \in \mathbb{R}^n$ and for all $\theta \in (0, 1)$. A continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is *strongly convex* with convexity parameter μ if $(\nabla f(y) - \nabla f(x))^\top (y-x) \geq \mu \|y-x\|^2$ for all $x, y \in \mathbb{R}^n$. A continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is *restricted strongly convex* with respect to a point $x \in \mathbb{R}^n$ if there exists a constant $\mu > 0$ such that $(\nabla f(y) - \nabla f(x))^\top (y-x) \geq \mu \|y-x\|^2$ for all $y \in \mathbb{R}^n$. A continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is *smooth* if it has a globally Lipschitz continuous gradient, i.e., there exists a constant $L > 0$ such that $\|\nabla f(y) - \nabla f(x)\| \leq L\|y-x\|$ for all $x, y \in \mathbb{R}^n$.

2.3. Convergence analysis

Finally, we recall the definition of linear (exponential in the language of control theory) convergence of a sequence from Nedić, Olshevsky, and Shi (2017a). Suppose that a sequence $\{x(k)\}$ converges to x^* in some norm $\|\cdot\|$. Then the convergence is (i) Q-linear if there exists a constant $0 < \rho < 1$ such that $\frac{\|x(k+1)-x^*\|}{\|x(k)-x^*\|} \leq \rho$ for all k; (ii) R-linear if there exist some constants $C > 0$ and $0 < \rho < 1$ such that $\|x(k) - x^*\| \leq C\rho^k$ for all k. Both of these rates are exponential or geometric. The difference is that Q-linear implies a monotonic decrease of $\|x(k) - x^*\|$, while the R-linear does not.

3. Basic distributed optimization algorithms

In this section, we provide an overview of some existing discrete-time and continuous-time distributed algorithms for solving the distributed unconstrained optimization problem over undirected and fixed communication networks. Various extensions of these algorithms will be discussed in Section 4.

3.1. Discrete-time Algorithm

Most existing distributed algorithms are in the discrete-time setting (see Nedić, 2015; Nedić and Liu, 2018; Nedić et al., 2018; Sayed, 2014a, and references therein). Among these algorithms, the first-order algorithms based on the consensus theory and the (sub)gradient method have received much attention since they are simple, amenable to implementation and easily generalizable. The existing distributed first-order discrete-time algorithms can be categorized into two classes depending on whether the step-sizes are diminishing or fixed.

3.1.1. Algorithms with diminishing step-sizes

We begin by reviewing discrete-time algorithms with diminishing step-sizes. A seminal work in this direction is Nedić and Ozdaglar (2009), where the authors proposed a simple distributed first-order (sub)gradient descent algorithm. In the proposed algorithm, each agent performs a consensus step and then a descent step along the local (sub)gradient direction of its own convex objective function. In particular, at time instant (step) k, each agent i runs the following update:

$$x_i(k+1) = \sum_{j=1}^{N} w_{ij}(k)x_j(k) - \alpha(k)s_i(k), \tag{2}$$

where $x_i(k) \in \mathbb{R}^n$ is agent i's estimate of the optimal solution at time instant k, $w_{ij}(k)$ is the edge weight of communication link (j, i) at time instant k such that $W(k) = [w_{ij}(k)] \in \mathbb{R}^{N\times N}$ is doubly stochastic for all k, $s_i(k)$ is the (sub)gradient of the local objective function $f_i(x)$ which is convex and possibly non-differentiable at $x = x_i(k)$, and $\alpha(k) > 0$ is the diminishing step-size satisfying the following conditions:

$$\sum_{k=0}^{\infty} \alpha(k) = \infty, \ \sum_{k=0}^{\infty} \alpha^2(k) < \infty,$$
$$\alpha(k) \leq \alpha(s) \text{ for all } k > s \geq 0. \tag{3}$$

Under the assumption that subgradients are bounded, it is shown in Nedić and Ozdaglar (2009) that the distributed (sub)gradient descent (DGD) algorithm with diminishing step-sizes asymptotically converges to one of the optimal solutions, if the fixed undirected graph is connected or the time-varying undirected graph is uniformly jointly connected. The authors of Johansson, Keviczky, Johansson, and Johansson (2008) modified of the DGD algorithm (2) by changing the order in which the consensus-step and the subgradient descent step are executed.

Although the simple DGD algorithm (2) is applicable to nonsmooth convex functions and has been extended in several directions to handle more realistic scenarios, which will be reviewed in details in the next section, the convergence is rather slow due to the diminishing step-sizes. For the case where local objective functions are smooth, the authors of Jakovetić, Xavier, and Moura (2014b) developed a fast distributed algorithm based on the centralized Nesterov gradient method. They showed that the proposed algorithm with diminishing step-sizes has a faster convergence rate compared to the DGD algorithm and the dual averaging algorithm (Duchi, Agarwal, & Wainwright, 2012), but is still slower than the centralized gradient descent algorithm. With a fixed step-size, it is shown in Matei and Baras (2011) and Yuan, Ling, and Yin (2015) that the DGD algorithm only converges to a point in the neighborhood of an optimal solution.

In order to achieve the convergence rate that matches the centralized gradient descent algorithm and to reduce communication overheads, recent studies focused on developing distributed accelerated algorithms with fixed step-sizes for the case where local objective functions are strongly convex and smooth. Both discrete-time and continuous-time distributed algorithms have been proposed and will be reviewed in Section 3.1.2 and Section 3.2, respectively.

3.1.2. Algorithms with a fixed step-size

In this subsection, we present several recently developed discrete-time distributed algorithms with fixed step-sizes which linearly (exponentially in the language of control theory) converge to an optimal solution. A common strategy of these proposed algorithms is to use some sort of historical information.

EXTRA. The earliest work is perhaps (Shi, Ling, Wu, & Yin, 2015a), where the authors developed an exact first-order algorithm (abbreviated as EXTRA). More specifically, the EXTRA contains two steps. In the first step, agent i performs the following update:

$$x_i(1) = \sum_{j=1}^{N} w_{ij}x_j(0) - \alpha\nabla f_i(x_i(0)), \tag{4}$$

where $\alpha > 0$ is a fixed step-size, the weight mixing matrix $W = [w_{ij}] \in \mathbb{R}^{N\times N}$ is doubly stochastic and $\nabla f_i(\cdot)$ is the gradient of the local objective function $f_i(\cdot)$. In the second step, agent i performs the following update:

$$x_i(k+2) = x_i(k+1) + \sum_{j=1}^{N} w_{ij}x_j(k+1) - \sum_{j=1}^{N} \tilde{w}_{ij}x_j(k) - \alpha(\nabla f_i(x_i(k+1)) - \nabla f_i(x_i(k))), \ k = 0, 1, \ldots, \tag{5}$$

where the weight mixing matrix $\tilde{W} = [\tilde{w}_{ij}] \in \mathbb{R}^{N\times N}$ is also doubly stochastic. Compared to the DGD algorithm (2), which only uses the estimate of the optimal solution and the gradient at the previous iteration, the EXTRA uses the estimates of the optimal solution and the gradients at the previous two iterations. It is shown in Shi et al. (2015a) that the EXTRA can be viewed as the DGD with a cumulative correction term to correct the error caused by the DGD with a fixed step-size.

For an undirected connected network, under certain conditions on the doubly stochastic weight mixing matrices, the linear convergence of the EXTRA with the step-size less than a certain critical value has been established if the global objective function is restricted strongly convex with respect to the global minimizer and local convex objective functions are smooth. The primal-dual interpretation of the EXTRA has been offered in Mokhtari and Ribeiro (2016) and Mokhtari, Shi, Ling, and Ribeiro (2016) based on the augmented Lagrangian function. The EXTRA has also been extended to composite convex problems where the local objective functions have the smooth and nonsmooth composite convex form. In particular, the authors of Shi, Ling, Wu, and Yin (2015b) developed a proximal gradient exact first-order algorithm (PG-EXTRA) by using proximal operations (Parikh & Boyd, 2014).

Other proximal algorithms for composite optimization are also available in the literature, see, e.g., Aybat, Wang, and Iyengar (2015); Aybat, Wang, Lin, and Ma (2018); Dhingra, Khong, and Jovanović (2019) and Xu, Zhu, Sohy, and Xie (2018b). In particular, the authors of Aybat et al. (2015) developed a distributed augmented Lagrangian algorithm with a double-loop structure. The authors of Aybat et al. (2018) proposed a distributed proximal gradient algorithm and its stochastic variant with noisy gradients. The proposed algorithms only contain a single-loop and therefore they are easy to be implemented.

The authors of Dhingra et al. (2019) developed a distributed proximal augmented Lagrangian method for the distributed composite convex optimization. The authors of Xu et al. (2018b) developed a distributed algorithm based on the Bregman method and operator splitting, referred to as Distributed Forward-Backward Bregman Splitting (D-FBBS). The proposed distributed algorithm provides a unified framework which recovers most existing distributed algorithms for fixed graphs, such as EXTRA (Shi et al., 2015a) and P-EXTRA (Shi et al., 2015b). Moreover, the proposed algorithm allows agents to communicate asynchronously, and thus is applicable to stochastic networks.

DIGing. Next, we review another class of distributed algorithms with fixed step-sizes, based on the combination of the distributed inexact gradient method and the gradient tracking technique (abbreviated as DIGing), developed independently in Nedić et al. (2017a); Qu and Li (2018); Xu, Zhu, Soh, and Xie (2015, 2018a) and Nedić, Olshevsky, Shi, and Uribe (2017b). In these algorithms, in addition to $x_i(k)$, which is agent i's estimate of the optimal solution at time step k, each agent i also holds another state $y_i(k)$, which is the estimate of the average gradient at time step k. In algorithms proposed in Qu and Li (2018) and Nedić et al. (2017a), these states are updated as follows:

$$x_i(k+1) = \sum_{j=1}^{N} w_{ij} x_j(k) - \alpha y_i(k), \tag{6a}$$

$$y_i(k+1) = \sum_{j=1}^{N} w_{ij} y_j(k) + \nabla f_i(x_i(k+1)) - \nabla f_i(x_i(k)), \tag{6b}$$

where the weight mixing matrix $W = [w_{ij}] \in \mathbb{R}^{N\times N}$ is doubly stochastic and $\alpha > 0$ is a fixed step-size. The algorithm is initialized with any $x_i(0)$ and $y_i(0) = \nabla f_i(x_i(0))$. The algorithm is based on the combination of the distributed inexact gradient method and the gradient tracking technique. More specifically, the update (6a) is a distributed inexact gradient method, where the variable $y_i(k)$ is used instead of the average gradient, and in the update (6b), $y_i(k)$ tracks the average gradient by employing dynamic average consensus (Zhu & Martínez, 2010).

Under the assumption that local objective functions are strongly convex and smooth, the linear convergence of the DIGing algorithm has been established for undirected connected networks if the fixed step-size is chosen properly (Nedić et al., 2017a; Qu & Li, 2018). For fixed undirected networks, it is shown in Nedić et al. (2017a) that the DIGing algorithm is equivalent to the EXTRA by properly choosing the two mixing matrices in the EXTRA. Moreover, Nedić et al. (2017a) also provided a primal-dual interpretation for the DIGing algorithm based on the augmented Lagrangian function.

Note that the algorithms proposed in Qu and Li (2018) and Nedić et al. (2017a) require an identical step-size for all agents. A few algorithms with different (uncoordinated) step-sizes have been proposed, see, e.g., Aug-DGM (augmented distributed gradient methods) (Xu et al., 2015), AsynDGM (asynchronous distributed gradient method) (Xu, Zhu, Soh, & Xie, 2018a), and ATC-DIGing (adapt-then-combine distributed inexact gradient tracking) (Nedić et al., 2017b). More specifically, each agent i performs the following update rule[1]:

$$x_i(k+1) = \sum_{j=1}^{N} w_{ij} (x_j(k) - \alpha_j y_j(k)), \tag{7a}$$

$$y_j(k+1) = \sum_{j=1}^{N} w_{ij} \big(y_j(k) + \nabla f_j(x_j(k+1)) - \nabla f_j(x_j(k))\big), \tag{7b}$$

where $\alpha_j > 0$ is the step-size of agent j.

Compared with the DIGing algorithm (6), which employs a combine-then-adapt (CTA) structure (Sayed, 2014b), where the states are combined with neighboring agents' states and then adapted, algorithm (7) utilizes an adapt-then-combine (ATC) structure (Sayed, 2014b), in which the states are first adapted and then combined with the adapted states of neighboring agents. As shown in Nedić et al. (2017b); Xu et al. (2015, 2018a), the ATC structure is capable of employing uncoordinated step-sizes. Another nice feature of the ATC structure is that the algorithms with ATC sturcture enjoy faster convergence (Nedić et al., 2017a; Nedić et al., 2017b).

Recently, the authors of Jakovetić (2019) provided a unified primal-dual analysis for both the EXTRA (Shi et al., 2015a) and the DIGing (Nedić et al., 2017a; Qu & Li, 2018) and revealed that a major difference between these two methods is on the effect of the primal error on the dual error. They then generalized these methods by deriving a new method which reduces the negative effect of primal error on the dual error. The authors of Sundararajan, Hu, and Lessard (2017) and Sundararajan, Scoy, and Lessard (2018) proposed a unified framework for analyzing the EXTRA, the DIGing, the algorithms proposed in Jakovetić (2019), and other existing first-order distributed algorithms for strongly convex and smooth objective functions by formulating a semidefinite program (SDP) which can be efficiently solved to provide a numerical certificate of the linear convergence. These works (Sundararajan et al., 2017; Sundararajan et al., 2018) can be viewed as an extension of their earlier work (Lessard, Recht, & Packard, 2016) for optimization, where the authors developed a framework to analyze and design iterative optimization algorithms by using integral quadratic constraints (IQC) from robust control theory (Megretski

[1] Note that the update for the variable y_i in AsynDGM proposed in Xu et al. (2018a) is slighly different, i.e., (7b) is replaced by $y_j(k+1) = \sum_{j=1}^{N} w_{ij} y_j(k) + \nabla f_i(x_i(k+1)) - \nabla f_i(x_i(k))$.

& Rantzer, 1997), to distributed multi-agent optimization. By using the IQC and dissipativity theory, the author of Han (2019) provided a computational proof for the convergence analysis of the distributed algorithm for undirected fixed graphs proposed in Qu and Li (2018) when local objective functions are smooth and convex but not strongly convex.

Distributed PI algorithm. To correct the error caused by the distributed gradient based algorithms with fixed step-sizes, distributed algorithms based on the proportional-integral (PI) control strategy have been developed in Lei, Chen, and Fang (2016); Yao, Yuan, Sundaram, and Yang (2018) and Yang, Wan, Wang, and Lin (2018). These algorithms are discrete-time counterparts of the continuous-time distributed PI algorithms proposed in Gharesifard and Cortés (2014); Kia, Cortés, and Martínez (2015a); Wang and Elia (2010) and Xie and Lin (2017), which will be reviewed in Section 3.2. In particular, in these algorithms, each agent performs the following update:

$$x_i(k+1) = x_i(k) - v_i(k) - \alpha \nabla f_i(x_i(k)) - \beta \sum_{j \in \mathcal{N}_i} a_{ij}(x_i(k) - x_j(k)), \tag{8a}$$

$$v_i(k+1) = v_i(k) + \alpha\beta \sum_{j \in \mathcal{N}_i} a_{ij}(x_i(k) - x_j(k)), \tag{8b}$$

where $x_i(k) \in \mathbb{R}^n$ is the local estimate of the global minimizer x^* of agent i at time step k, α, $\beta > 0$ are gain parameters, and $\mathcal{N}_i = \{j \in \mathcal{V} \mid (j,i) \in \mathcal{E}\}$ is the set of neighbors for agent i. The algorithm is initialized with any $x_i(0)$ and $v_i(0)$ such that $\sum_{i=1}^N v_i(0) = \mathbf{0}_n$.

It is shown in Yao et al. (2018) that the distributed PI algorithm (8) is equivalent to the EXTRA by properly choosing the two mixing matrices in the EXTRA. Therefore, the convergence analysis follows from the proof in Shi et al. (2015a). For the case where local objective functions are quadratic, a less conservative convergence condition was also established in Yao et al. (2018) via the Lyapunov stability analysis.

Distributed Newton–Raphson algorithm. Although we focus on first-order gradient-based algorithms, it is worthy to mention that a few second-order distributed algorithms based on the Newton method have been proposed, see, e.g., Varagnolo, Zanella, Cenedese, Pillonetto, and Schenato (2016); Wei, Ozdaglar, and Jadbabaie (2013a,b); Zanella, Varagnolo, Cenedese, Pillonetto, and Schenato (2011). In particular, the distributed Newton–Raphson algorithm proposed in Zanella et al. (2011) and Varagnolo et al. (2016) is based on the average consensus algorithm and the separation of time-scale idea. Intuitively, each agent computes and sequentially updates an approximated Newton–Raphson direction by means of suitable average consensus ratios. Although the algorithm proposed in Zanella et al. (2011) and Varagnolo et al. (2016) is actually a discrete-time algorithm, in order to establish its convergence, the algorithm is considered as a forward-Euler discretization of a continuous dynamics. The authors established the exponential stability of the continuous dynamics, which in turn implies the stability of the Euler discretization provided that the Euler discretization step-size is sufficiently small.

3.2. Continuous-time algorithms

Although classical distributed optimization algorithms are in the discrete-time setting, with the development of cyber-physical systems, much attention has been paid to the continuous-time setting, mainly because many practical systems such as robots and unmanned vehicles operate in continuous-time and the well-developed continuous-time control techniques (in particular Lyapunov stability theory) may facilitate the analysis. Various continuous-time distributed algorithms have been developed and can be categorized into two groups depending on whether an algorithm uses the first-order gradient information or the second-order Hessian information, which will be reviewed in Section 3.2.1 and Section 3.2.2, respectively. Similar to the discrete-time case, we focus on the distributed algorithms for fixed undirected graphs in this subsection and will review various extensions in Section 4.

3.2.1. First-order gradient-based algorithms

Distributed PI algorithm. By using the PI control strategy, various continuous-time distributed algorithms have been developed in the literature. The earliest work is perhaps (Wang & Elia, 2010), where the authors proposed the following algorithm:

$$\dot{x}_i(t) = \sum_{j=1}^N a_{ij}(x_j(t) - x_i(t)) + \sum_{j=1}^N a_{ij}(v_j(t) - v_i(t)) - \nabla f_i(x_i(t)), \tag{9a}$$

$$\dot{v}_i(t) = \sum_{j=1}^N a_{ij}(x_i(t) - x_j(t)), \tag{9b}$$

where $a_{ij} > 0$ is the edge weight for edge $(j,i) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise.

The algorithm is motivated by a feedback mechanism and is a proportional-integral control. More specifically, in (9a), the term $-\nabla f_i(x_i(t))$ ensures that each agent follows its local gradient descent while the term $\sum_{j=1}^N a_{ij}(x_j(t) - x_i(t))$ ensures that consensus is achieved among agents. However, if the dynamics just contains these two terms, the agents' states would not converge since the local gradients are not the same in general. Thus, to correct the error, the additional integral feedback term $\sum_{j=1}^N a_{ij}(v_j(t) - v_i(t))$ is included, where the dynamics of $v_i(t)$ is governed by (9b).

By defining $x(t) = [x_1^\top(t), x_2^\top(t), \dots, x_N^\top(t)]^\top$ and $v(t) = [v_1^\top(t), v_2^\top(t), \dots, v_N^\top(t)]^\top$, algorithm (9) can be rewritten in a more compact form:

$$\dot{x}(t) = -(L \otimes I_n)x(t) - (L \otimes I_n)v(t) - \nabla f(x(t)), \tag{10a}$$

$$\dot{v}(t) = (L \otimes I_n)x(t), \tag{10b}$$

where $L = [\ell_{ij}] \in \mathbb{R}^{N \times N}$, with $\ell_{ii} = \sum_{j=1}^N a_{ij}$ and $\ell_{ij} = -a_{ij}$ for $j \neq i$, is the Laplacian matrix associated with the underlying communication graph, $f(x(t)) = \sum_{i=1}^N f_i(x_i(t))$ and $\nabla f(x(t)) = [\nabla f_1^\top(x_1(t)), \dots, \nabla f_N^\top(x_N(t))]^\top$.

Note that algorithm (10) can also be interpreted as a primal-dual algorithm as shown in Gharesifard and Cortés (2014); Wang and Elia (2011) and Dörfler (2019). More specifically, as shown in Gharesifard and Cortés (2014, Lemma 3.1), the distributed optimization problem (1) is equivalent to the following optimization problem:

$$\min_{x \in \mathbb{R}^{Nn}} \quad \sum_{i=1}^N f_i(x_i) \tag{11a}$$

$$\text{s.t.} \quad (L \otimes I_n)x = \mathbf{0}_{Nn}. \tag{11b}$$

For the above optimization problem (11), consider the augmented Lagrangian function

$$\mathcal{L}(x, v) = f(x) + v^\top (L \otimes I_n)x + \frac{1}{2} x^\top (L \otimes I_n)x.$$

Then algorithm (10) is the associated saddle-point flow, which is also the continuous-time gradient flow algorithm developed in Arrow, Huwicz, and Uzawa (1958).

For undirected connected graphs, it is shown in Wang and Elia (2010) that algorithm (10) solves the distributed optimization problem asymptotically, that is, $\lim_{t\to\infty} x(t) = \mathbf{1}_N \otimes x^*$. Moreover, $\lim_{t\to\infty} v(t) = v^*$, where $(L \otimes I_n)v^* = \nabla f(\mathbf{1}_N \otimes x^*)$.

3.2.2. Second-order algorithms

In this subsection, we review the existing second-order distributed algorithms based on the second-order Hessian information. Compared to the first-order distributed algorithms, these algorithms result in a faster convergence rate by exploring the Hessian information.

Zero-Gradient-Sum Algorithm. The authors of Lu and Tang (2012) developed the following algorithm:

$$\dot{x}_i(t) = \gamma \left(\nabla^2 f_i(x_i(t))\right)^{-1} \sum_{j\in\mathcal{N}_i} a_{ij}(x_j(t) - x_i(t)), \tag{12}$$

where $\gamma > 0$ is a gain parameter. The algorithm is initialized with $x_i(0) = x_i^*$, where x_i^* is the minimizer of the local objective function $f_i(x)$.

Note that it is shown that $\sum_{i\in\mathcal{V}} \nabla f_i(x_i(t)) = 0$ for all $t \geq 0$ if the undirected network is connected. For this reason, the distributed algorithm (12) is called Zero-Gradient-Sum (ZGS) algorithm. For undirected connected graphs, the authors of Lu and Tang (2012) showed that algorithm (12) exponentially converges to the global minimizer via Lyapunov stability analysis if local objective functions are twice continuously differentiable, strongly convex, and have locally Lipschitz Hessians.

4. Extensions

In Section 3, we have reviewed various existing discrete-time and continuous-time distributed algorithms for solving the unconstrained optimization problem over undirected graphs. In this section, we discuss how these algorithms have been extended in various directions to handle more realistic scenarios. In particular, Section 4.1 describes how these algorithms can be extended to directed communication networks, which may be time-varying. Section 4.2 reviews distributed optimization algorithms for the case where agent dynamics are more complicated. Section 4.3 discusses how to handle the distributed optimization problem with constraints. We then focus on the case where the communication networks may be subject to communication constraints, such as time-delays in Section 4.4 and random graphs in Section 4.5. In Section 4.6, we describe how to design event-triggered communication strategies for distributed optimization algorithms to avoid continuous communication and to reduce communication overheads. Section 4.7 reviews distributed finite-time optimization algorithms.

In these subsections for various extensions, we first provide an overview of both discrete-time and continuous-time optimization algorithms, if both exist in the literature. Although we only focus on one extension in each of following subsections, these extensions are not independent but actually may overlap one another to some extent. We will also briefly discuss the future research directions at the end of each subsection.

4.1. Directed graphs

In this subsection, we focus on the case where the communication networks are directed and possibly time-varying, since in practice, the information exchange may be unidirectional due to nonuniform communication powers, and the network topology may vary due to unexpected loss of communication links. Various discrete-time and continuous-time distributed optimization algorithms have been proposed for directed networks, which will be discussed in Section 4.1.1 and Section 4.1.2, respectively.

4.1.1. Discrete-time

Existing discrete-time distributed algorithms can be divided into two classes depending on the method an algorithm uses to handle directed communication networks.

Distributed Push-Sum Based Algorithms. Most existing discrete-time distributed algorithms for directed graphs are based on the push-sum method (Bénézit, Blondel, Thiran, Tsitsiklis, & Vetterli, 2010; Kempe, Dobra, & Gehrke, 2003; Nedić & Olshevsky, 2015) or ratio consensus (Charalambous et al., 2015; Domínguez-García & Hadjicostis, 2011; 2015; Hadjicostis & Charalambous, 2014), which relaxes the requirement of doubly stochastic mixing matrices in the distributed algorithms reviewed in Section 3.1 for undirected graphs to column stochastic matrices. In the push-sum method, each agent runs two linear iterations simultaneously with a column stochastic matrix, where the initial conditions of these two iterations are the initial estimations for the optimal solution and the vector of all ones, respectively, and the ratio of the two states converges to the initial average for strongly connected directed graphs which are not necessarily weight-balanced. By using this method together with the (sub)gradient method, various distributed algorithms have been developed.

The earlier studies focused on developing distributed algorithms with diminishing step-sizes for the case where the local convex objective functions are not necessarily differentiable but with bounded subgradients. The authors of Tsianos, Lawlor, and Rabbat (2012) developed a distributed algorithm based on the push-sum method and dual averaging (Duchi et al., 2012) for fixed directed graphs which are strongly connected. Later, the authors of Nedić and Olshevsky (2015) studied general directed time-varying networks which are uniformly jointly strongly connected. The convergences of these proposed algorithms are rather slow due to the diminishing step-sizes.

To accelerate the convergence process, distributed algorithms with fixed step-sizes by using the push-sum method have been proposed in Nedić et al. (2017a, 2017b); Xi and Khan (2017); Zeng and Yin (2017). In particular, the authors of Zeng and Yin (2017) developed an algorithm, termed ExtraPush, for the case where the local objective functions are quasi-strongly convex and smooth. the authors of Xi and Khan (2017) developed an algorithm, termed DEXTRA (Directed-EXTRA), for the case where the local objective functions are restricted strongly convex with respect to the global minimizer and smooth. The step-sizes of both ExtraPush and DEXTRA are restrictive in the sense that their lower bounds are strictly greater than zero. the authors of Xi, Xin, and Khan (2018) proposed an algorithm, termed ADD-OPT (Accelerated Distributed Directed Optimization), in which the lower bound of the step-size is zero, and thus supports a wider range of step-sizes. The linear convergence was established for the case where the local objective functions are strongly convex and smooth. It is shown in Nedić et al. (2017a) that the ATC variant of the DIGing algorithm with the push-sum method is applicable to time-varying directed graphs which are uniformly jointly strongly connected. This algorithm was extended with uncoordinated step-sizes in Nedić et al. (2017b) and the linear convergence was established for connected undirected graphs even if the step-sizes are not identical.

Distributed Push–Pull Based Algorithms. The second class of algorithms is based on the push–pull method (see, e.g., Du, Yao et al., 2018; Pu, Shi, Xu, & Nedić, 2018; Xin & Khan, 2018; Yang et al., 2013). In Yang et al. (2013), the authors focused on the case where local objective functions are quadratic and developed a distributed algorithm based on the surplus idea (Cai & Ishii, 2012). The asymptotic convergence was established for strongly connected directed graphs if the learning gain parameter is sufficiently small. Recently, to extend the DIGing algorithm (6) to directed networks and general non-quadratic objective functions, distributed algorithms developed in Du, Yao et al. (2018); Xin and Khan (2018) and Pu et al. (2018) use a row stochastic matrix for the mixing of estimates of the optimal point, while they employ a column stochastic matrix for tracking the average gradient. From the viewpoint of an agent, the information about the estimates of the optimal point is pushed to the neighbors, while the information about the gradients is pulled (collected) from the neighbors, hence giving the name push–pull gradient method (Pu et al., 2018). For the case where local objective functions are strongly convex and smooth, the linear convergence has been established for these push–pull based algorithms over strongly connected directed graphs (Du, Yao et al., 2018; Pu et al., 2018; Xin & Khan, 2018).

Compared with push-sum based algorithms, distributed push–pull based algorithms require less computation and communication, since in push-sum based algorithms, the division operators are used, which causes additional computation (Pu et al., 2018). Recently, the authors of Zhang, Yi, George, and Yang (2019) presented a unified framework based on IQC for analyzing the linear convergence of various push–pull based algorithms proposed in Du, Yao et al. (2018); Xin and Khan (2018) and Pu et al. (2018) and the push–pull variant of the algorithm proposed in (Xu et al., 2018a) when local objective functions are strongly convex and smooth. The work can be viewed as an extension of the IQC framework recently proposed in (Sundararajan et al., 2017) for undirected graphs to directed graphs. the authors of Saadatniaki, Xin, and Khan (2018) developed a push–pull algorithm for time varying directed graphs which are uniformly jointly strongly connected.

To implement both distributed push-sum based and push–pull based algorithms, each agent is required to know its out-degree in order to construct a column stochastic matrix. This may be impractical when agents use broadcast-based communication (Hendrickx & Tsitsiklis, 2015). On the other hand, the construction of a row stochastic matrix is much easier since each agent could assign the edge weight to its in-neighbors. As shown in Hendrickx and Tsitsiklis (2015), in order to compute the optimal solution for the distributed optimization of a sum of local functions in the context of broadcast-based algorithms, each agent needs to know either its out-degree or its unique identifier. Recently, by assuming that each agent knows its unique identifier, the authors of Xi, Mai, Xin, Abed, and Khan (2018) developed a distributed algorithm which only uses a row stochastic matrix. To cancel the imbalance caused by employing only the row-stochastic matrix, each agent holds an additional variable that converges asymptotically to the left eigenvector corresponding to the eigenvalue at 1 of the row-stochastic matrix. The gradient is then divided by this additional variable to cancel the imbalance. The linear convergence was established for strongly connected directed graphs if the step-size is less than a certain value. This algorithm was later extended to uncoordinated step-sizes in Xin, Xi, and Khan (2019) and the linear convergence was established even if the step-sizes are not identical as long as the largest step-size is sufficiently small.

4.1.2. Continuous-time

In this subsection, we discuss how to extend continuous-time algorithms reviewed in Section 3.2 for undirected graphs to directed graphs.

Algorithm (10) originally proposed in Wang and Elia (2010) for undirected graphs was extended to directed graphs in Gharesifard and Cortés (2014). In particular, the authors developed the following algorithm:

$$\dot{x}(t) = -\alpha(L \otimes I_n)x(t) - Lv(t) - \nabla f(x(t)), \tag{13a}$$

$$\dot{v}(t) = (L \otimes I_n)x(t), \tag{13b}$$

where $\alpha > 0$ is a design parameter. Under the assumptions that the directed graph is strongly connected and weight-balanced and the local convex objective functions are smooth, it is shown that algorithm (13) asymptotically converges to the global minimizer if the gain parameter α is appropriately chosen.

Note that both algorithm (10) and algorithm (13) require communication for both variables x and v. In order to reduce the overall communication, the authors of Kia et al. (2015a) developed the following algorithm:

$$\dot{x}(t) = -\beta(L \otimes I_n)x(t) - v(t) - \alpha\nabla f(x(t)), \tag{14a}$$

$$\dot{v}(t) = \alpha\beta(L \otimes I_n)x(t), \quad (\mathbf{1}_N^\top \otimes I_n)v(0) = \mathbf{0}_n, \tag{14b}$$

where α, $\beta > 0$ are gain parameters.

Compared with algorithms (10) and (13), algorithm (14) only needs the communication for the variable x, but not for the variable v. On the other hand, in order to remove the communication among the variable v, it requires a special initialization for $v(0)$ as given in (14b). Under assumptions that the directed graph is strongly connected and weight-balanced, and the local objective functions are strongly convex and smooth, the authors of Kia et al. (2015a) established the exponential convergence of algorithm (14) with properly chosen gain parameters.

Note that gain parameters in most existing algorithms often depend on some global information, such as the Lipschitz constant and the network connectivity. To remove such a requirement, several distributed adaptive algorithms have been developed (see, e.g., Li, Ding, Sun, & Li, 2018; Lin, Ren, & Farrell, 2017; Zhao, Liu, Wen, & Chen, 2017). Continuous-time distributed algorithms discussed so far are based on the PI control strategy and are first-order algorithms based on the gradient information. On the other hand, the ZGS algorithm (12) which uses the second-order Hessian information proposed in Lu and Tang (2012) has also been extended to strongly connected and weight-balanced directed graphs in Chen and Ren (2016) and Guo and Chen (2018).

In summary, various discrete-time and continuous-time distributed optimization algorithms have been extended to directed graphs. In the discrete-time setting, common techniques for the convergence analysis are linear systems of inequalities and the small-gain theory. These techniques have been extended in Nedić et al. (2017a, 2017b) to time-varying directed graphs which are uniformly jointly strongly connected but not necessarily weight-balanced. In the continuous-time setting, a common technique used for the convergence analysis is the Lyapunov stability theory. For fixed directed graphs which are strongly connected and weight-balanced, various Lyapunov functions have been developed in the existing studies. However, it is difficult to construct a Lyapunov function for general unbalanced directed graphs. Moreover, it is also challenging to construct a common or multiple Lyapunov function for time-varying graphs (Liberzon & Morse, 1999). Therefore, the Lyapunov analysis has been only extended to time-varying directed graphs for certain special cases. For example, for the case where all local objective functions have a common minimizer, the common Lyapunov function has been constructed in Shi, Johansson, and Hong (2013). An interesting research direction is to extend the Lyapunov analysis to general unbalanced fixed directed graphs and time-varying directed graphs.

4.2. Complex agent dynamics

Most existing distributed optimization algorithms discussed so far can be viewed as distributed optimal coordination algorithms for multi-agent systems with single integrator agent dynamics. However, in many practical applications, the agent dynamics may be more complicated, such as those of double integrators, high-order systems, and Euler–Lagrangian (EL) systems. For example, EL systems have been used to describe many mechanical systems, such as mobile robots, rigid bodies, and autonomous vehicles (Lynch & Park, 2017; Spong, Hutchinson, & Vidyasagar, 2006). Moreover, any EL system with exact knowledge of nonlinearities can be transformed into a double-integrator system.

Recently, focusing on undirected connected graphs or strongly connected and weight-balanced directed graphs, distributed optimization algorithms for multi-agent systems with more complicated agent dynamics have been developed. Most studies focused on continuous-time agent dynamics. The research is progressing with an increasing complexity of agent dynamics. The double integrator agent dynamics were considered in Liu and Wang (2015); Yi, Yao, Yang, George, and Johansson (2018); Zhang and Hong (2014) and Wang, Gupta, and Wang (2018). In particular, the authors of Zhang and Hong (2014) developed a distributed algorithm and showed that the proposed algorithm asymptotically converges to the optimal solution if the local objective functions are strongly convex and smooth. Note that the parameters of the algorithm proposed in Zhang and Hong (2014) depend on some global information. To relax such dependence, the authors of Yi et al. (2018) developed an alternative distributed algorithm where no global information is needed to be known in advance. Moreover, the exponential convergence was established for the case where each local objective function is locally smooth and the global objective function is restricted strongly convex with respect to the global minimizer, which is more general compared to the strong convexity required in Zhang and Hong (2014).

There are a few works which studied the case where the agent dynamics are high-order systems. For agents described by chains of integrators, the authors of Zhang and Hong (2015) developed a continuous-time distributed algorithm based on the gradient method and the integral feedback idea. For the case where agents are general linear-time invariant systems and the local objective functions are quadratic-like, two distributed adaptive algorithms based on dynamic coupling gains have been developed in Zhao, Liu et al. (2017). For the case where the agent dynamics are EL systems, continuous-time distributed algorithms have been developed in Meng et al. (2017); Qiu, Hong, and Xie (2016) and Zhang, Deng, and Hong (2017). In particular, the authors of Meng et al. (2017) considered a special case where the intersection of local optimal convex sets is non-empty, while the general case was considered in Zhang et al. (2017). the authors of Zhang et al. (2017) developed two continuous-time distributed algorithms for the case without parametric uncertainties and the case with parametric uncertainties, respectively. The exponential convergence was established for the case without parametric uncertainties, while the asymptotic convergence was established for the case with parametric uncertainties.

Apart from more complex agent dynamics, another important issue is the physical constraints of the actuator since every actuator is subject to saturation due to its physical limitations. Although various distributed control laws (algorithms) have been developed, these proposed control laws designed in the absence of actuator saturation may fail to solve the distributed optimization problem in the presence of actuator saturation. Note that global consensus without the optimality concern has been widely studied (see, e.g., Li, Xiang, & Wei, 2011; Meng, Zhao, & Lin, 2013; Yang, Meng, Dimarogonas, & Johansson, 2014; Yi, Yang, Wu, & Johansson, 2019a; Zhao & Lin, 2016). However only a limited number of results are available on global optimal consensus (or equivalently distributed optimization) for multi-agent systems in the presence of actuator saturation. the authors of Xie and Lin (2017) developed distributed protocols for single integrator agents and double integrator agents, and showed that global optimal consensus is achieved in the presence of actuator saturation when the underlying network is strongly connected and detailed balanced. The design was later extended to high-order integrator agents in Xie and Lin (2019). For EL systems, the authors (Qiu, Hong et al., 2016) developed a distributed protocol to achieve global optimal consensus under given constraints on velocity with the requirement that the control input is bounded.

In summary, most studies in this direction assume that the graph is either undirected and connected or directed strongly connected and weight-balanced. Some of these results have been extended to the case with local constraints in Liu and Wang (2015); Qiu, Hong et al. (2016); Qiu, Liu, and Xie (2016) and Qiu, Liu, and Xie (2018). To avoid the continuous communication and to reduce the communication overheads, several event-triggered algorithms for the case where the agent dynamics are second-order have developed in Yi et al. (2018) and Wang, Gupta et al. (2018). These two extensions (local constraints and event-triggered communication schemes) will be discussed in Section 4.3 and Section 4.6, respectively.

There are a few other future research directions: i) extend the commonly used Lyapunov stability theory in these existing studies for high-order multi-agent systems to unbalanced directed graphs. Compared with the Lyapunov analysis for first-order multi-agent systems, the complex agent dynamic makes such an extension more challenging; ii) investigate the robustness properties of these existing distributed algorithms, since various disturbances, arising from either environment or communication, are ubiquitous in reality. Some preliminary results are available in Mateos-Núñez and Cortés (2016); Wang, Hong, and Ji (2016). In particular, the authors of Mateos-Núñez and Cortés (2016) studied the distributed PI algorithms for single integrator agents with persistent Gaussian white noise, and showed that the resulting stochastic dynamics is noise-to-state exponentially stable in the second moment, while the authors of Wang, Hong et al. (2016) considered the case where the agent dynamics are heterogeneous nonlinear high-order systems perturbed by external disturbances, and developed a distributed protocol based on the internal model principle; iii) study the multi-agent systems with discrete-time agent dynamics as such models are relevant for many practical sampled-data systems. Some studies are available in this direction but mainly focused on discrete-time single integrator agents. For example, the authors of Yang et al. (2018) extended the results of Xie and Lin (2017) for the continuous-time single-integrator agents to the discrete-time setting, while the authors of Qiu et al. (2018) developed a distributed bounded control protocol with time-varying gain parameters based on the local subgradient descent and the projection method for solving the distributed constrained optimization problem.

4.3. Constrained optimization

So far, we have only focused on the distributed unconstrained optimization problem. However, in physical applications, there may exist various constraints such as local constraints, global inequality and equality constraints. Such a constrained optimization problem can be formulated as follows:

$$\min_{x\in\mathbb{R}^n} \quad \sum_{i=1}^{N} f_i(x) \tag{15a}$$

$$\text{s.t.} \quad x \in \cap_{i=1}^{N} \Omega_i, \tag{15b}$$

$$g(x) \le \mathbf{0}_m, \tag{15c}$$

$$h(x) = \mathbf{0}_p, \tag{15d}$$

where the function $f_i : \mathbb{R}^n \to \mathbb{R}$ is the local convex objective function for agent i, and Ω_i is the local constraint set known to agent i only and is assumed to be closed and convex. The function $g : \mathbb{R}^n \to \mathbb{R}^m$ with each component g_ℓ, $\ell \in \{1, 2, \ldots, m\}$, being convex. The inequality $g(x) \le \mathbf{0}_m$ is understood to be component-wise, i.e., $g_\ell(x) \le 0$ for all $\ell \in \{1, 2, \ldots, m\}$, and represents a global inequality constraint. The function $h : \mathbb{R}^n \to \mathbb{R}^p$ with each component h_ℓ, $\ell \in \{1, 2, \ldots, p\}$, being convex, represents a global equality constraint.

4.3.1. Local constraint sets

We begin with a simple case where each node is subject to a local constraint set and there is no global inequality and equality constraints, that is, constraints (15c) and (15d) in (15) are nonexistent. The constrained optimization problem (15) then reduces to:

$$\min_{x\in\mathbb{R}^n} \quad \sum_{i=1}^N f_i(x) \tag{16a}$$

$$\text{s.t.} \quad x \in \cap_{i=1}^N \Omega_i. \tag{16b}$$

The earlier studies focused on the special case when the constraint sets are identical (common), i.e., Ω_i for $i = 1, 2, \ldots, N$, are the same. In the discrete-time setting, the authors of Nedić, Ozdaglar, and Parrilo (2010) developed a distributed projected (sub)gradient algorithm with diminishing step-sizes, in which each agent performs distributed algorithm (2) and takes the projection of this vector on the common constraint set. In the continuous-time setting, the authors of Qiu, Liu et al. (2016) developed a distributed protocol containing three terms: local averaging, local projection, and local subgradient with a diminishing but persistent gain. Both algorithms were shown to converge to an optimal point if the undirected graph is uniformly jointly connected.

For the general case when the constraint sets are different, the convergence of the projected subgradient algorithm was only established for fixed fully connected graphs in Nedić et al. (2010). Such a result was extended to time-varying directed graphs in Zhu and Martínez (2012) and Lin, Ren, and Song (2016), where the authors established the convergence results when time-varying directed graph is uniformly jointly strongly connected and weight-balanced at all times. Recently, the authors of Wang, Lin, Ren, and Song (2018) established the convergence under the same condition without the square summable requirement of the step-sizes in (3). Note that these existing methods require each node to compute a projection on the local constraint set, which may not be easy. the authors of Aybat and Hamedani (2016) and Hamedani and Aybat (2017) developed distributed primal-dual algorithms for the case where the objective functions are composite convex and the constraint sets are conic.

Although the aforementioned optimization algorithms are applicable to time-varying graphs, their convergence is rather slow due to the required diminishing step-sizes. To accelerate the convergence rate, a few distributed algorithms have been developed for fixed undirected graphs. In the discrete-time setting, the authors of Lei et al. (2016) developed a distributed primal-dual algorithm with a fixed step-size. The proposed algorithm can be viewed as an extension of the EXTRA (Shi et al., 2015a) and distributed PI algorithm (Yao et al., 2018) to the constrained case.

In the continuous-time setting, a few distributed algorithms are also available. the authors of Liu and Wang (2015) extended the distributed PI algorithm originally proposed in Gharesifard and Cortés (2014) to the constrained case. Note that the auxiliary variables of the proposed algorithm may be asymptotically unbounded, which makes the proposed algorithm hard to implement in practice. Such an issue was removed later in Zeng, Yi, and Hong (2017), where the authors developed a distributed algorithm which has bounded states while seeking the optimal solutions. These proposed algorithms in Lei et al. (2016); Liu and Wang (2015) and Zeng et al. (2017) are applicable to nonsmooth local convex objective functions and the asymptotic convergence result was established for connected undirected graphs. For the special case where local objective functions are strongly convex and locally smooth on the constraint set, the authors of Wang, Wang, Sun, and Wang (2018) proposed a distributed algorithm based on the integral control strategy and the projection method, and established its exponential convergence if the parameters are properly chosen.

4.3.2. Global constraints

In Section 4.3.1, we have discussed the distributed constrained optimization problem when each agent is subject to a local constraint set. In this subsection, we consider the general case when global inequality and equality constraints are also present in addition to the local constraints.

We first consider the case where there are only global inequality constraints, that is, constraints (15d) in (15) are nonexistent. The constrained optimization problem (15) then reduces to:

$$\min_{x\in\mathbb{R}^n} \quad \sum_{i=1}^N f_i(x) \tag{17a}$$

$$\text{s.t.} \quad x \in \cap_{i=1}^N \Omega_i, \tag{17b}$$

$$g(x) \le \mathbf{0}_m, \tag{17c}$$

By assuming that the global inequality constraints are known to all the agents, a few discrete-time distributed algorithms with diminishing step-sizes have been proposed. In particular, the authors of Zhu and Martínez (2012) developed a primal-dual projected subgradient algorithm with diminishing step-sizes based on the characterization of saddle points of the augmented Lagrangian function. It is shown that the algorithm asymptotically converges to an optimal solution for time-varying directed graphs which are weight-balanced and uniformly jointly strongly connected.

For the case where local constraints set are identical, the authors of Yuan, Xu, and Zhao (2011) proposed a distributed primal-dual subgradient algorithm with a fixed step-size and multiple updates for each consensus step. For connected undirected graphs, it is shown that the algorithm converges to the optimal point within the error level depending on the number of consensus updates.

Note that aforementioned algorithms in this subsection require each agent projecting its primal-dual estimate onto some convex sets at every iteration, which results in high computational cost. To reduce such a cost, a distributed regularized primal-dual subgradient algorithm with a fixed step-size was developed in Yuan, Ho, and Xu (2016) for the case without local constraints, whose implementation requires only one projection at the last iteration. Moreover, the explicit convergence rate for the objective error was also established for time-varying directed graphs which are uniformly jointly strongly connected.

Next, we consider the case where there are only global linear equality constraints, that is, constraints (15c) in (15) are nonexistent. The constrained optimization problem (15) then reduces to:

$$\min_{x\in\mathbb{R}^n} \quad \sum_{i=1}^N f_i(x) \tag{18a}$$

$$\text{s.t.} \quad x \in \cap_{i=1}^N \Omega_i, \tag{18b}$$

$$h(x) = \mathbf{0}_p, \tag{18c}$$

where the function $h : \mathbb{R}^n \to \mathbb{R}^p$ is defined as $h(x) = Bx - b$ with $B \in \mathbb{R}^{p\times n}$.

For the case where local constraint sets are identical, the authors of Zhu and Martínez (2012) developed a distributed penalty primal-dual subgradient algorithm with a diminishing step-size. It is shown that the algorithm asymptotically converges to an optimal solution for time-varying directed graphs which are uniformly jointly strongly connected and weight-balanced at all times.

For the general case where local constraint sets are different, the authors of Liu, Yang, and Hong (2017) developed discrete-time distributed algorithm with a fixed step-size. Asymptotic convergence to the global minimizer was established for undirected connected graphs if the step-size is less than an estimable upper bound.

In the continuous-time setting, the authors (Yang, Liu, & Wang, 2017b; Zhu, Yu, Wen, Chen, & Ren, 2018) developed distributed algorithms based on the PI control strategy and the subgradient method, respectively. By using the nonsmooth analysis and the Lyapunov stability theory, the asymptotic convergence results were established for undirected connected graphs.

In summary, most existing discrete-time and continuous-time distributed algorithms for solving the distributed constrained optimization problem are only applicable to undirected graphs or weight-balanced digraphs. One future research direction is to develop distributed algorithms for unbalanced digraphs. An interesting result in this direction has been obtained by Xie, You, Tempo, Song, and Wu (2018), where the authors considered a distributed optimization problem with a local inequality constraint and developed a discrete-time distributed algorithm with a diminishing step-size using the epigraph form of the constrained optimization. It is shown that the proposed algorithm asymptotically converges to the common optimal point even for time-varying unbalanced digraphs which are uniformly jointly strongly connected. It is worthy to investigate if such a result can be extended to the distributed algorithms with fixed step-sizes.

Another interesting future direction is to consider the optimization problem with a general set of equality and inequality constraints that couple all the agents' decision variables. In the presence of a coupled constraint, the feasible region of one agent's decision variable is influenced by some other agents' decision variables. If such a constraint is known by all agents, various algorithms have been proposed (see, e.g., primal-dual algorithms proposed in Cherukuri, Mallada, & Cortés, 2016; Feijer & Paganini, 2010; Qu & Li, 2019). However, in practice, such a coupled constraint may not be known to each agent. To meet this challenge, a few distributed algorithms have been developed, among which discrete-time algorithms are given in Chang, Nedić, and Scaglione (2014); Chatzipanagiotis, Dentcheva, and Zavlanos (2015); Chatzipanagiotis and Zavlanos (2016); Falsone, Margellos, Garatti, and Prandini (2017); Margellos, Falsone, Garatti, and Prandini (2018); Nedić, Olshevsky, and Shi (2018) and continuous-time algorithms are presented in Cherukuri and Cortés (2015, 2016); Deng, Liang, and Yu (2018); Liang, Zeng, and Hong (2018); Yi, Hong, and Liu (2016) and Deng, Liang, and Hong (2018).

Also note that most of these existing studies do not consider the communication network imperfections, such as varying topologies, time-delays, and packet drops. It is worthy to study the effects of these network imperfections to the existing algorithms and to develop distributed algorithms for solving the distributed constrained optimization problem even in the presence of these communication imperfections. In the next two subsections, we provide an overview of the existing distributed optimization algorithms that deal with communication network imperfections.

4.4. Time delays

In this subsection and the next subsection, we respectively discuss the effects of two communication imperfections, time-delays and random graphs, on the existing distributed optimization algorithms.

Time delays are ubiquitous in practical systems especially multi-agent systems, due to extra time required to get information, limited communication speed, computation time and execution time required to implement the control input (Aström & Kumar, 2014; Cao et al., 2013; Hespanha, Naghshtabrizi, & Xu, 2007). It is known that for a single system time delays might degrade the system performance or even destroy the stability (Richard, 2003; Zhu, Qi, Ma, & Chen, 2018). The effects of time delays on various existing discrete-time and continuous-time distributed optimization have been investigated, and will be reviewed in Section 4.4.1 and Section 4.4.2, respectively.

4.4.1. Discrete-time

In the discrete-time setting, the effects of both constant delays and time-varying bounded delays have been studied for existing distributed optimization algorithms. To model the bounded communication time delays, a common approach for convergence analysis is to use the augmented graph idea originated from Tsitsiklis (1984), which has also been used in the literature on distributed consensus (Cao, Morse, & Anderson, 2008; Charalambous et al., 2015; Hadjicostis & Charalambous, 2014; Nedić & Ozdaglar, 2010). More specifically, for each node in the original graph, extra virtual nodes are introduced to capture the effect of delays on various communication links. The distributed optimization problem with time delays in the original communication network is then reduced to the problem without delays in the augmented graph.

The earlier studies focused on distributed algorithms with diminishing step-sizes. For fixed directed networks, it is shown in Tsianos et al. (2012); Tsianos and Rabbat (2011) that the distributed algorithm based on dual averaging proposed in Duchi et al. (2012) can be extended to case with both fixed constant delays and random bounded delays, for undirected fixed graphs and strongly connected directed graphs, respectively.

For time-varying directed networks, the weight-balanced case and the unbalanced case have been considered in Lin et al. (2016) and Yang, Lu et al. (2017), respectively. In particular, the authors of Lin et al. (2016) extended the distributed algorithm based on the subgradient projection originally proposed in Nedić et al. (2010) for the distributed constrained optimization problem to handle communication delays and non-identical constraint sets. It is shown that the proposed subgradient projection algorithm solves the distributed constrained optimization problem if the time-varying directed network is uniformly jointly strongly connected and weight-balanced at all times, even in the presence of arbitrarily large bounded time-varying delays. the authors of Yang, Lu et al. (2017) showed that the distributed algorithm based on the push-sum method originally proposed in Nedić and Olshevsky (2015) still converges to the optimal solution if the time-varying directed network is uniformly jointly strongly connected but not necessarily weight-balanced, even in the presence of arbitrarily large bounded time-varying delays.

Note that the above distributed algorithms use diminishing step-sizes, which results in slow convergence. Recently, the effects of constant time delays to the distributed algorithm with a fixed learning gain (step-size) proposed in Yang et al. (2013) for fixed directed strongly connected graphs have been investigated in Yang, Wu, Sun, and Lian (2015); Zhao, Duan, and Shi (2019). It is shown in Zhao, Duan et al. (2019) that the algorithm still converges to the global minimizer even in the presence of nonuniform constant time delays provided that the fixed step-size is sufficiently small. The explicit upper bound on the step-size to ensure the convergence was also established for uniform constant delays.

4.4.2. Continuous-time

The effects of communication time delays on the existing distributed continuous-time optimization algorithms have also been studied, and the findings are reviewed in this subsection. the authors of Yang, Liu, and Wang (2017a) investigated the effects of delays on the distributed PI algorithm proposed in Kia et al. (2015a) for strongly connected and weight-balanced directed graphs. By using a Lyapunov–Krasovskii functional for time-delay systems, the authors established the delay-dependent and delay-independent sufficient conditions in the form of linear matrix inequalities (LMIs) for both slow varying delays and fast varying delays. The results were later extended to the case where the agent dynamics are double integrators in Tran, Wang, Liu, and Xiao (2017) and when each agent is subject to a local bounded constraint set in Wang, Wang, Chen, and Wang (2018). in Hatanaka, Chopra, Ishizaki, and Li (2018), the authors provided a passivity-based perspective for the distributed PI algorithm proposed in Wang and Elia (2010). By using the Lyapunov analysis and passivity, the authors showed that the algorithm is capable of handling arbitrarily large unknown constant time delays.

The authors of Guo and Chen (2018) studied the performance of the distributed zero-gradient-sum algorithm proposed in Lu and Tang (2012) in the face of time-varying communication delays. By constructing a Lyapunov–Krasovskii functionals, they established explicit sufficient conditions for the maximum admissible time delay to guarantee the convergence for fixed undirected connected graphs. The extension to time-varying undirected graphs that are connected at all times was also studied by using a common Lyapunov function.

For connected undirected graphs, the authors of Doan, Beck, and Srikant (2017) studied the effects of uniform constant communication time delays on the continuous-time version of the DGD algorithm proposed in Nedić et al. (2010). By constructing a Lyapunov–Razumikhin function, the authors provided an explicit analysis of the convergence rate of the algorithm as a function of the network size, topology, and time delay.

In summary, various discrete-time and continuous-time algorithms have been extended to solve the distributed optimization problem in the presence of communication time delays. In the discrete-time setting, a common approach for the convergence analysis is to use the augmented graph idea (Tsitsiklis, 1984) to convert the distributed optimization problem with bounded time delays in the original graph to the distributed optimization problem without time delays in the augmented graph. In the continuous-time setting, a common technique for the convergence analysis is to use the Lyapunov analysis based on Lyapunov–Krasovskii functionals or Lyapunov–Razumikhin functions (Fridman, 2014). Most existing studies focused on the case where the agent dynamics are single-integrator. It is an interesting future research direction to study the effects of communication time delays on the existing distributed algorithms for multi-agent systems with high-order agent dynamics. The objective here is to explicitly characterize the delay margin for high-order multi-agent systems subject to unknown communication time delays, which can be reviewed as a non-trivial extension of the recent studies (Ma & Chen, 2019) for a single agent and (Ma, Tian, Zulfiqar, Chen, & Chai, 2019) for distributed consensus to distributed optimization. This in turn provides an upper bound on the time delays under which the convergence of the distributed continuous-time algorithms is still guaranteed albeit slower.

4.5. Random graphs

In addition to communication time delays, there exist other uncertainties in communication networks, such as packet drops and link failures. Although time delays can be used to model unreliable communication networks with these uncertainties, a more realistic approach is to model such communication networks as random graphs (Schenato, Sinopoli, Franceschetti, Poolla, & Sastry, 2007). Thus, it is important to investigate the performance of the existing distributed optimization algorithms in the face of these communication uncertainties.

The earlier studies focused on discrete-time distributed algorithms for undirected graphs. the authors of Lobel and Ozdaglar (2011) considered the case where communication links fail according to a given stochastic process and showed that the DGD algorithm with a diminishing step-size proposed in Nedić and Ozdaglar (2009) almost surely converges to an optimal solution if the link failures are independent and identically distributed (i.i.d.) over time and the expected graph is connected. With a fixed step-size, the authors of Matei and Baras (2011) showed that for twice continuously differentiable strongly convex objective functions with bounded Hessians, the algorithm converges in the mean square sense to the global minimizer with a guaranteed distance, which can be made arbitrarily small by an appropriate choice of the step-size.

Recently, a few algorithms have been developed for solving the distributed optimization problem over directed random networks resulting from packet-dropping communication links. the authors of Carli, Notarstefano, Schenato, and Varagnolo (2015) developed a distributed algorithm based on the Newton–Raphson consensus algorithm (Varagnolo et al., 2016), the push-sum method (Bénézit et al., 2010; Charalambous et al., 2015; Hadjicostis & Charalambous, 2014; Kempe et al., 2003), and the robustified strategy (Dominguez-Garcia, Hadjicostis, & Vaidya, 2012; Hadjicostis, Vaidya, & Dominguez-Garcia, 2016). It is shown that the algorithm converges to the global minimizer almost surely if the Euler discretization step-size is sufficiently small. Note that the above algorithm uses the second-order Hessian information. When the Hessian information is not available, the authors of Wu, Yang, Wu, Kalsi, and Johansson (2017) developed a distributed algorithm by integrating the distributed algorithm based on the push sum method (Nedić & Olshevsky, 2015; Yang, Lu et al., 2017) and the robustified strategy proposed in Dominguez-Garcia, Hadjicostis et al. (2012); Hadjicostis et al. (2016). The almost sure convergence was established with diminishing step-sizes. Although the above algorithms are resilient to link failures, the drawback is that their convergence rates are rather small since (Carli et al., 2015) requires the Euler discretization step-size to be sufficiently small and (Wu, Yang, Wu et al., 2017) uses diminishing step-sizes. In order to accelerate the convergence rate, the authors of Jakovetić, Xavier, and Moura (2014a) modified the two distributed Nesterov-like gradient methods proposed in their earlier work (Jakovetić et al., 2014b) for fixed networks to handle random graphs and established their convergence rates in terms of the expected optimality gap in the cost function at an arbitrary node.

The aforementioned studies assumed that the subgradient is perfectly measured. However, in certain cases, the subgradient may only be measured with noises. Recently, the effects of both random graphs and noisy stochastic subgradients were studied simultaneously by focusing on undirected graphs (see, e.g., Jakovetić, Bajovic, Sahu, & Kar, 2018; Lei, Chen, & Fang, 2018; Srivastava & Nedić, 2011). In particular, the authors of Srivastava and Nedić (2011) and Jakovetić et al. (2018) developed distributed algorithms with two sets of diminishing step-sizes, one for the consensus part and the other one for the noisy subgradient part. The convergence results were established if the step-sizes for the subgradeint part decays to zero at a faster rate than those for the consensus part. Later, the authors of Lei et al. (2018) considered the distributed constrained optimization problem with local constraint sets and developed a distributed primal-dual algorithm which only uses the same diminishing step-sizes for both the consensus part and the subgradient part. The almost sure convergence was established by using the stochastic approximation theory.

In summary, the existing distributed algorithms for random graphs either use the gradient information (see, e.g., Jakovetić et al., 2014a; Lobel & Ozdaglar, 2011; Wu, Yang, Wu et al., 2017) and diminishing step-sizes or fixed step-sizes and the Hessian information (see, e.g., Carli et al., 2015; Matei & Baras, 2011). It is an interesting future research direction to develop accelerated gradient-based distributed algorithms with fixed step-sizes for random networks. It is worthy to mention that some efforts have been devoted to this direction. the authors of Xu et al. (2018a) developed a distributed algorithm based on the gradient tracking method and the dynamic average consensus technique for undirected random graphs. In contrast to most existing studies with diminishing step-sizes, the algorithm uses a constant uncoordinated step-size and allows for asynchronous implementation. Moreover, the authors showed that the algorithm converges almost surely to the optimal point if the largest step-size is less than a certain critical value. We believe that this algorithm can be extended to directed random graphs by using the push-sum method (Bénézit et al., 2010; Charalambous et al., 2015; Hadjicostis & Charalambous, 2014; Kempe et al., 2003; Nedić & Olshevsky, 2015) and the push–pull method (Du, Yao et al., 2018; Pu et al., 2018; Xin & Khan, 2018; Yang et al., 2013).

Also existing studies that investigated the effects of both random graphs and noisy stochastic subgradients are restricted to distributed algorithms for undirected graphs. It is also worthy to investigate both of these effects simultaneously on the existing push-sum based and push–pull based distributed algorithms for directed graphs.

4.6. Event-triggered communication

Most existing continuous-time distributed optimization algorithms require continuous information exchange among agents, which may be impractical in physical applications. Moreover, distributed networks are usually resources constrained and communication is energy-consuming. In order to avoid continuous communication and to reduce communication overheads, the idea of event-triggered communication and control has been proposed. The early works focused on a single system (Aström & Bernhardsson, 2002; Girard, 2015; Tabuada, 2007) and have been extended to multi-agent systems (Heemels, Johansson, & Tabuada, 2012; Wan & Lemmon, 2009; Wang & Lemmon, 2011; Zhang, Li, Sun, & He, 2019; Zhang, Sun, Liang, & Li, 2019; Zhong & Cassandras, 2010) and distributed consensus (Dimarogonas, Frazzoli, & Johansson, 2012; Meng, Xie, Soh, Nowzari, & Pappas, 2015; Seyboth, Dimarogonas, & Johansson, 2013; Yi et al., 2019a) Zhong and He (2019). For more details, please refer to Ding, Han, Ge, and Zhang (2018); Nowzari, Cortés, and Pappas (2018); Nowzari, Garcia, and Cortés (2019); Yi (2017, and references therein).

Recently, the event-triggered communication strategies have been extended to distributed optimization. Several studies developed event-triggered communication mechanisms for implementing distributed continuous-time algorithms over connected undirected graphs or strongly connected and weight-balanced directed graphs. The key challenge is to design the event-trigger communication strategies, such that the event-triggered algorithm converges to the optimal solution and is free of Zeno behavior, an infinite number of triggered events in a finite period of time (Johansson, Egerstedt, Lygeros, & Sastry, 1999).

Most studies focused on the first-order multi-agent systems with single integrator agent dynamics. in Kia et al. (2015a), the authors developed a distributed PI algorithm with an event-triggered communication strategy and established its exponential convergence to a neighborhood of an optimal point. To achieve the exact convergence, the distributed ZGS algorithm proposed in Lu and Tang (2012) has been extended with a periodical time-triggered communication mechanism in Chen and Ren (2016) and an event-triggered scheme in Liu and Chen (2016). Inspired by the distributed dynamic event-triggered strategy proposed in Girard (2015); Yi (2017), the authors of Du, Yi, George, Johansson, and Yang (2018) extended the distributed ZGS algorithm with a dynamic event-triggered communication strategy, which is more efficient compared to the time-triggered strategy and the static event-triggered strategy. the authors of Du, Yi, Zhang, George, and Yang (2019) equipped distributed PI algorithms proposed in Gharesifard and Cortés (2014); Wang and Elia (2010) and Kia et al. (2015a) with static event-triggered communication schemes.

For the second-order multi-agent systems with double integrator agent dynamics, the authors of Tran, Wang, Liu, Xiao, and Lei (2019) modified the distributed algorithm proposed in Zhang and Hong (2014) and equipped it with an event-triggered communication scheme. Note that the existing distributed optimization algorithms for second-order multi-agent systems are not fully distributed since the gain parameters of the algorithms depend on some global parameters, such as the eigenvalues of the graph Laplacian matrix. the authors of Yi et al. (2018) developed a fully distributed algorithm for second-order multi-agent systems where no global information is needed and extended it with a dynamic event-triggered communication mechanism.

For the high-order multi-agent systems, the authors of Wang, Wang, and Wang (2016) modified the distributed optimization algorithm proposed in Zhang and Hong (2015) and extended it with an event-triggered communication scheme.

Recently, the design of event-triggered communication strategies has taken communication effects, such as time-varying topologies, time delays, packet drops, and limited bandwidth, into consideration. in Liu, Chen, and Dai (2019), the authors developed an event-triggered algorithms based on the ZGS algorithm proposed in Lu and Tang (2012) for undirected time-varying networks. For first-order multi-agent systems, the authors of Liu, Xie, and Quevedo (2018) considered both the event-triggered communication scheme and the limited communication bandwidth, and developed a distributed event-triggered optimization algorithm with dynamic encoder and decoder pairs. The proposed algorithm is based on the algorithm proposed in Shi et al. (2013) which uses diminishing step-sizes and is applicable only to a special case where all local objective functions have a common minimizer.

So far most existing studies focused on developing event-triggered communication schemes for information exchange among agents. Recently, both the event-triggered communication strategy and the event-triggered gradient measurement strategy have been developed for distributed gradient based algorithms for first-order multi-agent systems. in Tran, Wang, Liu, and Xiao (2018), the authors developed identical periodical time-triggered strategies for both communication and gradient information. Based on the internal model principle, the authors of Deng, Wang, and Hong (2017) proposed a distributed algorithm with both the event-triggered communication strategy and the event-triggered gradient measurement strategy and established the convergence in the presence of the external disturbances. It is interesting to develop both the event-triggered communication strategy and the event-triggered gradient measurement strategy for high-order multi-agent systems.

Most existing distributed event-triggered optimization algorithms are in the continuous-time setting. Recently, a limited number of studies developed distributed event-triggered optimization algorithms in the discrete-time setting. An event-triggered algorithm based on the discretization of the ZGS algorithm (Lu & Tang, 2012) was proposed in Chen and Ren (2016). The hybrid event-time-driven optimization algorithm was developed in Hu, Guan, Chen, and Shen (2019). In order to ensure the effectiveness of the proposed event-triggered communication strategies, it is desirable to show that the inter event-times are at least lower bounded by a non-trivial bound, which has been largely ignored in the literature even for distributed consensus. Most event-triggered schemes for discrete-time distributed consensus and optimization algorithms only admitted a trivial lower bound of inter-event times, that is, the lower bound is one. We are only aware of one work (Meng, Xie, & Soh, 2016) in the context of the network utility maximization, which ensures a non-trivial lower bound of two. In the future, it is worthy to develop event-triggered communication strategies for distributed discrete-time optimization algorithms with fixed step-sizes, which admit a non-trivial lower bound of inter-event times.

4.7. Finite-time convergence

All the distributed optimization algorithms discussed so far converge to an optimal solution either asymptotically or exponentially. However, in some time-critical applications, such as DER coordination to be considered in Section 5, it is highly desirable to develop distributed optimization algorithms which solve the problem in a finite-time. An added advantage of the finite-time convergence is that the overall computation and communication overheads can be greatly reduced.

In the literature, several distributed finite-time optimization algorithms have been developed in both discrete-time and continuous-time settings.

4.7.1. Discrete-time

In the discrete-time setting, only a limited number of results are available (Mai & Abed, 2018; Yao et al., 2018). Both studies used the finite-time computation technique for distributed consensus originally proposed in Sundaram and Hadjicostis (2007); Yuan, Stan, Shi, Barahona, and Goncalves (2013) and Yuan (2012), which enables an arbitrarily chosen agent to compute the final consensus value within a finite number of time steps, by using its local successive states. In particular, the authors of Yao et al. (2018) studied the case where the communication graph is undirected. The authors first proposed a distributed discrete-time PI algorithm and established its linear convergence to the global minimizer for quadratic local objective functions. The proposed distributed PI algorithm was then equipped with the finite-time consensus technique to enable agents to compute the global minimizer in a finite number of time steps.

Independently, for quadratic local objective functions, a different distributed algorithm based on the finite-time consensus technique (Sundaram & Hadjicostis, 2007; Yuan, 2012; Yuan et al., 2013) and the ratio consensus algorithm proposed in Charalambous et al. (2015); Domínguez-García and Hadjicostis (2011); Hadjicostis and Charalambous (2014) was developed for directed graphs in Mai and Abed (2018) by exchanging the parameters of the objective functions. For general local objective functions, the finite-time consensus technique was used in Mai and Abed (2018) to periodically reset the distributed subgradient algorithm with a fixed step-size. It is shown that all agents reach consensus on an identical estimate of an optimal solution in a finite-time if the global objective function is strongly convex. The algorithm then behaves like the centralized subgradient method, i.e., all local estimates continue to agree and approach the global minimizer together.

4.7.2. Continuous-time

In the continuous-time setting, several distributed finite-time optimization algorithms have also been developed (see, e.g., Chen & Li, 2018; Feng & Hu, 2017; Lin et al., 2017; Pilloni, Pisano, Franceschelli, & Usai, 2016; Song & Chen, 2016). These algorithms are motivated by the discontinuous finite-time consensus protocols. In order to establish the convergence, the finite-time Lyapunov analysis (Bhat & Bernstein, 2000) is commonly used. the authors of Lin et al. (2017) developed a distributed finite-time algorithm based on a distributed tracking algorithm and a dynamic averaging estimator for the distributed constrained optimization problem where the local constraint sets are identical. For quadratic local objective functions, the authors of Pilloni et al. (2016) proposed a discontinuous signum-function based algorithm by modifying the distributed PI algorithms originally developed in Gharesifard and Cortés (2014); Wang and Elia (2010) and Kia et al. (2015a). Motivated by the finite-time consensus protocol (Yu & Long, 2015), the authors of Feng and Hu (2017) developed a distributed finite-time optimization algorithm to solve the distributed optimization problem with communication and computation uncertainties. Inspired by the finite-time consensus protocol (Wang & Xiao, 2010) and the distributed ZGS algorithm (Lu & Tang, 2012), the authors of Song and Chen (2016) developed a distributed finite-time ZGS algorithm.

The settling time in the existing continuous-time finite-time optimization algorithms depends on the initial conditions, which may be hard to preassign off-line. Recently, to overcome this limitation, the authors of Li, Yu, Zhou, and Ren (2017) and Chen and Li (2018) developed distributed algorithms which converge to the optimal solution within a fixed time independent of the initial condition, by using the fixed-time stability theory (Polyakov, 2018; Polyakov, Efimov, & Perruquetti, 2015a; 2015b).

In summary, distributed finite-time optimization algorithms in both discrete-time and continuous-time settings have been developed by using different techniques. In the discrete-time setting, the existing algorithms are based on the decentralized finite-time computation mechanism. In the continuous-time setting, the proposed algorithms are based on discontinuous finite-time consensus protocols. Note that the existing distributed finite-time algorithms in the continuous-time setting are only applicable to undirected networks. However, directed communication networks are more realistic due to nonuniform communication powers. Thus, it is worthy to develop distributed finite-time continuous-time optimization algorithms for directed networks in the future.

5. Application to coordination of distributed energy resources

In this section, we focus on the application of distributed optimization algorithms to solve the optimal coordination problem of distributed energy resources (DERs), which has received substantial attention in both control and power system fields in recent years (see, e.g., Bidram et al., 2014; Hadjicostis et al., 2018; Kraning et al., 2014; Molzahn et al., 2017; Nedić & Liu, 2018; Qin et al., 2017). In particular, Section 5.1 presents the problem formulation of DER coordination. In Section 5.2, we review existing works on distributed DER coordination for undirected graphs. The extension to directed graphs is discussed in Section 5.3. The communication constraints, such as time delays and packet drops, on the existing DER coordination algorithms are discussed in Section 5.4. Distributed DER coordination algorithms with event-triggered communication mechanisms and the finite-time convergence are reviewed in Section 5.5 and Section 5.6, respectively.

5.1. Problem formulation of DER coordination

The objective of the optimal DER coordination problem is to minimize the total production cost while meeting the total demand and simultaneously satisfying the individual generator output limits. Mathematically, it is formulated as the following optimization problem:

$$\min_{x_i \in X_i} \quad \sum_{i=1}^{N} C_i(x_i) \tag{19a}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} x_i = D, \tag{19b}$$

$$x_i \in X_i := [x_i^{\min}, x_i^{\max}], \; i = 1, 2, \ldots, N, \tag{19c}$$

where N is the number of distributed generators (DGs), x_i is the power generation of DG i, D is the total demand, and $C_i(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ is the cost function of DG i, $\mathbb{R}_+$ is the set of positive real numbers. Constraint (19b) is the power balance constraint, and constraints (19c) are generators' capacity constraints, where $x_i^{\min}$ and $x_i^{\max}$ are respectively the lower and upper bounds for DG i. The cost function $C_i(x_i)$ is assumed to be strictly convex and twice continuously differentiable over X_i.

We note that there are various power system applications where distributed control and optimization play an important role, such as distributed algorithms for optimal power flow (OPF), distributed optimal frequency control, distributed optimal voltage control, and optimal wide-area control. We refer the readers to the recent survey Molzahn et al. (2017), which provides a comprehensive review for these problems, Lam, Zhang, and Tse (2012); Low (2014) and Madani, Sojoudi, and Lavaei (2015) for the semidefinite program (SDP) relaxation of OPF, Zheng, Wu, Zhang, Sun, and Liu (2016) and Peng and Low (2018) for the second order cone program (SOCP) relaxation of OPF, and Kim and Baldick (1997); Sun, Phan, and Ghosh (2013) and Guo, Hug, and Tonguz (2017) for the non-convex OPF.

The problem formulation in (19) though simple, is very representative. For example, as pointed out in Nedić and Liu (2018), a similar problem formulation has been applied to other power system applications, including economic dispatch for microgrids (Hug, Kar, & Wu, 2015; Tang et al., 2018; Zhang & Chow, 2012), optimal load control (Mallada, Zhao, & Low, 2017; Zhao, Topcu, & Low, 2013), optimal load sharing (Yi, Hong, & Liu, 2015), and distributed energy management (Du, Yao et al., 2018; He, Yu, Huang, & Li, 2019; Zhang, Xu, Liu, Zang, & Yu, 2015; Zhao, He, Cheng et al., 2017). Due to the wide representation of the problem formulation in (19), a few survey papers on distributed consensus and optimization also briefly discussed the DER coordination problem, see, e.g., Qin et al. (2017) and Nedić and Liu (2018). The purpose of this section is to review distributed DER coordination algorithms in details.

One approach to solving the optimal DER coordination problem is through a completely centralized control strategy, which requires a single control center that accesses the entire network's information, computes the optimal generations, and sends the information back to generators. This centralized approach may be subject to performance limitations, such as a single point of failure, high communication requirement, substantial computation burden, and limited flexibility and scalability. To overcome these limitations, by using the results developed in the field of distributed optimization, various distributed algorithms have been proposed recently to solve the optimal DER coordination problem.

Before providing a detailed review for these distributed algorithms, we provide a brief discussion on how to solve the optimal DER coordination problem in (19) by distributed optimization algorithms developed for the problem of the form (1). Please refer to Yang, Lu et al. (2017) and Nedić and Liu (2018) for details.

Since i) each cost function $C_i(\cdot)$ is convex, ii) constraint (19b) is affine, and iii) the set $X_1 \times \cdots \times X_N$ is a polyhedral set, if we dualize problem (19) with respect to constraint (19b), there is zero duality gap. Moreover, the dual optimal set is nonempty (Bertsekas, Nedić, & Ozdaglar, 2003). Consequently, solutions of the optimal DER coordination problem can be obtained by solving its equivalent dual problem.

For convenience, let $\mathbf{x} = [x_1, \ldots, x_N]^\top \in \mathbb{R}_+^N$. Then, define the Lagrangian function

$$\mathcal{L}(\mathbf{x}, \lambda) = \sum_{i=1}^{N} C_i(x_i) - \lambda \left(\sum_{i=1}^{N} x_i - D \right), \tag{20}$$

where λ is the dual variable (incremental cost).

The corresponding Lagrangian dual problem can be written as

$$\max_{\lambda \in \mathbb{R}} \sum_{i=1}^{N} \Phi_i(\lambda), \tag{21}$$

where

$$\Phi_i(\lambda) = \min_{x_i \in X_i} C_i(x_i) - \lambda (x_i - D_i), \tag{22}$$

and D_i is a virtual local demand at each bus such that $\sum_{i=1}^{N} D_i = D$.

Therefore, the optimal DER coordination problem (19), which is a constrained optimization problem, can be solved by solving its equivalent dual problem (21), which has the same form as the distributed unconstrained optimization problem (1). The only difference is that the dual problem (21) is a maximization problem for concave functions while problem (1) is a minimization problem for convex functions. Therefore, distributed algorithms reviewed in Sections 3 and 4 for the distributed optimization problem (1) can be readily applied.

We are now ready to review existing distributed algorithms for solving the optimal DER coordination problem.

5.2. Undirected graphs

The earlier studies focused on the case where the communication network is modeled as an undirected graph and the generation cost functions are quadratic. the authors of Zhang and Chow (2012) proposed a leader–follower consensus based algorithm where the leader collects the mismatch between the demand and the total generation, and then leads the updates of the incremental cost. Through various case studies, it is shown that this distributed algorithm converges to the optimal generations for connected undirected graphs if the feedback gain parameter for the mismatch term is less than a certain critical value.

An alternative distributed algorithm with two sets of diminishing step-sizes based on the consensus and innovation approach (Kar & Moura, 2013) was developed in Kar and Hug (2012), where the consensus part guarantees that consensus is achieved, and the innovation term ensures the balance between the total generation and the total demand. The convergence result was established if the consensus term dominates the innovation term, i.e., if the diminishing step-size for the consensus term decays to zero at a faster rate compared to that of the innovation term. Motivated by the EXTRA (Shi et al., 2015a) the authors of Tang et al. (2018) developed a distributed DER coordination algorithm with a fixed step-size which leads to a fast convergence speed.

In summary, all the aforementioned distributed algorithms are only applicable to undirected communication networks. In the literature, various distributed algorithms for DER coordination have been developed by taking more realistic scenarios into consideration and will be reviewed in the following subsections.

5.3. Directed graphs

In this subsection, we review the existing discrete-time and continuous-time distributed algorithms for solving the optimal DER coordination problem over directed graphs which may be time-varying.

5.3.1. Discrete-time

In the discrete-time setting, the earlier studies focused on the case where the digraph is fixed and the cost functions are quadratic. The authors of Dominguez-Garcia, Cady et al. (2012) developed a distributed algorithm based on ratio consensus (Bénézit et al., 2010; Domínguez-García & Hadjicostis, 2011; Kempe et al., 2003). It is shown that the ratio consensus-based algorithm asymptotically converges to the optimal generations for strongly connected digraphs. A distributed algorithm with a fixed learning gain (step-size) based on the surplus idea (Cai & Ishii, 2012) was proposed in Yang et al. (2013). The algorithm asymptotically converges to the optimal generations for strongly connected digraphs if the step-size is sufficiently small.

Recently, several distributed DER coordination algorithms have been developed for general non-quadratic convex cost functions. In particular, the authors of Xing et al. (2015) proposed a distributed bisection method based on a consensus-like iteration. The asymptotic convergence was established for the case where the cost functions are strictly convex and twice continuously differentiable and the directed graph is strongly connected. The authors (Yang, Lu et al., 2017) developed a distributed algorithm based on the push-sum method (Nedić & Olshevsky, 2015) for time-varying directed graphs. It is shown that the algorithm converges to the optimal generations if the time-varying directed graph is uniformly jointly strongly connected and the step-sizes are diminishing. A nonnegative surplus-based distributed algorithm was developed in Xu et al. (2017). The asymptotic convergence was established if the time-varying directed graph is uniformly jointly strongly connected and the time-varying parameters are chosen appropriately. Recently, focusing on undirected graphs, the strictly convexity assumption on the cost functions has recently been relaxed to only convexity in Doan and Beck (2017, 2018). In particular, the authors of Doan and Beck (2017) developed a distributed Lagrangian algorithm with a diminishing step-size for connected undirected graphs and established the asymptotic convergence, while the authors of Doan and Beck (2018) focused on time-varying loads and uniformly jointly connected time-varying undirected graphs and established the almost sure convergence.

The aforementioned DER coordination algorithms for time-varying directed graphs require diminishing step-sizes, which results in rather slow convergence. Recently, the authors of Du, Yao et al. (2018) developed a distributed push–pull based algorithm with a fixed step-size for solving the DER coordination problem over both fixed strongly connected directed graphs and time-varying uniformly jointly strongly connected directed graphs. The proposed distributed push–pull based algorithm with a fixed step-size can be viewed as an extension of the distributed algorithm proposed in Yang et al. (2013) for fixed graphs and quadratic cost functions to time-varying graphs and general non-quadratic strictly convex cost functions.

5.3.2. Continuous-time

In the continuous-time setting, most studies focused on fixed graphs. the authors of Cherukuri and Cortés (2015) developed a distributed algorithm based on the Laplacian nonsmooth gradient dynamics for solving the DER coordination problem over fixed directed graphs. It is shown that the algorithm asymptotically converges to the optimal generations for strongly connected and weight-balanced directed graphs.

In order to ensure the convergence of the above algorithm, the initial conditions need to satisfy the load constraints. Such an initialization requirement was later removed in Cherukuri and Cortés (2016), where the authors developed a distributed algorithm based on the Laplacian nonsmooth gradient dynamics and dynamic average consensus (Freeman, Yang, & Lynch, 2006; Kia, Cortés, & Martínez, 2015b). In particular, each generator estimates the average mismatch between the demand and the total generation by employing the dynamic average consensus algorithm, which was then fedback to the Laplacian nonsmooth gradient dynamics. It is shown that the algorithm asymptotically converges to the optimal generations for any initial condition.

The proposed distributed algorithms in Cherukuri and Cortés (2015, 2016) require each agent to share its own gradient information with its neighboring agents, which may be private information that cannot be shared. To overcome this issue, the authors of Yi et al. (2016) developed two distributed initialization-free algorithms based on either projection or differentiated projection to solve the DER coordination problem over connected undirected graphs, which can be easily extended to strongly connected and weight-balanced directed graphs. Moreover, the proposed algorithms are also capable of handling more general local convex constraints other than box constraints.

By using an augmented Lagrangian function with the generation capacity constraints, the authors of Bai, Ye, Sun, and Hu (2019) developed an alternative initialization-free distributed DER coordination algorithm. The proposed algorithm is based on the saddle point dynamics, dynamic average consensus and leader–follower consensus, and is applicable to solve the DER coordination problem with transmission line constraints.

5.4. Communication imperfections

Since communication imperfections, such as time delays and packet drops, are ubiquitous in communication networks (Aström & Kumar, 2014; Cao et al., 2013; Hespanha et al., 2007), it is desirable to investigate the potential effects of these communication imperfections on the existing distributed DER coordination algorithms, and to develop distributed algorithms that are robust to these imperfections.

First, we consider the effects of communication time delays on both discrete-time and continuous-time distributed DER coordination algorithms.

In the discrete-time setting, the authors of Zhang, Chow, and Chakrabortty (2012) and Yang et al. (2015) respectively investigated the performance of the algorithms proposed in Zhang, Ying, and Chow (2011) and Yang et al. (2013) in the presence of uniform constant time delays via numerical simulations. Both studies found that there exists a critical value for delays, below which these DER coordination algorithms with given gain parameters still converge, and above which these algorithms fail to converge. It is shown in Zhao, Duan et al. (2019) that the algorithm proposed in Yang et al. (2013) still converges to the optimal generations even in the presence of nonuniform constant time delays provided that the learning gain parameter is sufficiently small. The explicit upper bound on the learning gain parameter to ensure the convergence was established for uniform constant delays. Note that the above studies considered constant time delays. In the case of time-varying delays, the authors of Yang, Lu et al. (2017) developed a distributed push-sum based algorithm with a diminishing step-size (Nedić & Olshevsky, 2015) and showed that the proposed algorithm converges to the optimal generations for time-varying uniformly jointly strongly connected digraphs, even in the presence of arbitrarily large bounded time-varying delays.

In the continuous-time setting, the authors of Zhu, Yu, and Wen (2016) developed a distributed DER coordination algorithm and studied the effects of uniform constant time delays. It is shown that the proposed algorithm still converges to the optimal generations if uniform constant delays are less than some threshold. The authors of Chen and Zhao (2018) developed a distributed DER coordination algorithm and investigated its performance in the presence of nonuniform constant time delays. The maximum allowable delay bound was obtained by the Generalized Nyquist Criterion. The above studies focused on the case where the communication graph is fixed. In the case of time-varying graphs, the authors of Somarakis and Baras (2015) and Somarakis, Maity, and Baras (2016) developed a distributed DER coordination algorithm and investigated its performance in the presence of time-varying delays. Sufficient conditions under which the proposed algorithm still solves the DER coordination problem were established.

Next, we consider another common communication imperfection in the communication networks – packet drops. Although time-varying communication networks may be used to model packet drops, a more realistic modeling approach is based on the probability framework, i.e., the communication link fails with a certain probability. In such a probability setting, most existing DER coordination algorithms are not able to handle packet drops. the authors of Wu, Yang, Wu et al. (2017) developed a robustified extension of the distributed algorithm proposed in Yang, Lu et al. (2017) by using the robustified strategy proposed in Dominguez-Garcia, Hadjicostis et al. (2012) and Hadjicostis et al. (2016). Under the assumption that the underlying communication network is strongly connected with a positive probability and the packet drops are i.i.d., it is shown that the robustified distributed algorithm solves the DER coordination problem almost surely even in the presence of packet drops.

5.5. Event-triggered communication

Most distributed algorithms for solving the optimal DER coordination problem discussed so far require the continuous information exchange among DERs. To avoid such continuous communication and thus to reduce the communication burden, distributed event-triggered DER coordination algorithms have been developed.

The authors of Li, Yu, Yu, Huang, and Liu (2016) first developed a distributed DER coordination algorithm based on θ-logarithmic barrier-based method and then equipped it with an event-triggered communication scheme. It is shown that the event-triggered DER coordination algorithm converges to the optimal generations if the undirected graph is connected. An event-triggered algorithm was proposed in Ding, Wang, Yin, Zheng, and Han (2019) to solve the DER coordination problem including both distributed generators and demand response. Recently, the authors of Zhao, Li, and Ding (2019) developed an event-triggered algorithm to solve the DER coordination problem with transmission losses. It is shown that the event-triggered DER coordination algorithm is free of Zeno behavior and converges to the optimal generations for undirected connected graphs. For the case where the communication network is modeled as a directed graph, the authors of Shi, Wang, Song, and Yan (2018) developed a distributed algorithm with a diminishing step-size and an event-triggered scheme to solve the DER coordination problem over strongly connected directed graphs.

5.6. Finite-time convergence

All the distributed algorithms discussed so far only solve the DER coordination problem asymptotically. However, power and energy systems require time-critical and fast response when new energy needs are demanded (Yu & Xue, 2016). Thus, it is highly desirable to develop distributed algorithms which solve the DER coordination problem in a finite-time. In the literature, a few distributed finite-time algorithms have been developed in both the continuous-time setting and the discrete-time setting. Most existing studies focused on quadratic generator cost functions.

In the discrete-time setting, motivated by the decentralized finite-time computation technique proposed in Sundaram and Hadjicostis (2007) and Yuan et al. (2013), the authors of Yang, Wu, Sun et al. (2016) developed a decentralized algorithm which enables each distributed generator to compute its optimal generation in a minimum number of time steps, by using its local successive states obtained from the underlying DER coordination algorithm proposed in Yang et al. (2013) for the case where the generation cost functions are quadratic and the directed graph is strongly connected.

In the continuous-time setting, based on distributed finite-time consensus protocols, several distributed DER coordination algorithms with finite-time convergence have been proposed. Motivated by discontinuous consensus protocol proposed in Chen, Lewis, and Xie (2011), the authors of Chen, Ren, and Feng (2017) developed a distributed DER coordination algorithm and established its finite-time convergence to the optimal generations for connected undirected graphs. Note that uncertain information commonly exists in the communication network and the computation process. the authors of Feng and Hu (2017) developed a distributed finite-time DER coordination algorithm and established its finite-time convergence for connected undirected graphs and general convex cost functions, even in the presence of communication and computation uncertainties.

Although these continuous-time DER coordination algorithms converge to the optimal generations in a finite-time, the settling time depends on the initial condition, which may be difficult to preassign off-line. To overcome this limitation, the authors of Li et al. (2017) and Chen and Li (2018) have recently developed distributed algorithms for solving the DER coordination problem over connected undirected graphs, without generation capacity constraints and with generation capacity constraints, respectively. It is shown that these algorithms converge to the optimal generations within a fixed time independent of the initial conditions.

6. Summary

In this paper, we have provided a comprehensive survey of existing discrete-time and continuous-time distributed optimization algorithms. Moreover, we have discussed how these distributed algorithms are applied/or adapted to solve the optimal DER coordination problem in power systems.

6.1. Current state

Tables 1–3 show comprehensive lists of existing distributed algorithms reviewed in Sections 3 and 4. In particular, Table 1 and Table 2 compare the existing discrete-time algorithms with diminishing step-sizes and fixed step-sizes, respectively, while the existing continuous-time algorithms are summarized in Table 3. Table 4 provides a list of existing distributed algorithms for solving the optimal DER coordination problem.

6.2. Future research directions

Although we have pointed out some future research directions for distributed optimization in each subsections of Section 4, there are still other important and yet challenging future research direc-

Table 1
Discrete-time distributed algorithms with diminishing step-sizes.

Study	Approach	Constraints type	Cost function	Communication graph	Convergence
(Nedić & Ozdaglar, 2009)	distributed (sub)graident method	unconstrained	convex but nonsmooth	uniformly jointly connected	$O(\ln k/\sqrt{k})$
(Duchi et al., 2012)	distributed dual averaging	common convex constraint set	convex but nonsmooth	connected undirected	$O(\ln k/\sqrt{k})$
(Jakovetić et al., 2014b)	distributed augmented Lagrangian with multiple consensus step in the inner loop	unconstrained	convex and smooth with bounded gradient	connected undirected	$O(1/k^2)$
(Nedić et al., 2010)	distributed projected (sub)gradient method	common convex constraint set	convex but nonsmooth	uniformly jointly connected	asymptotic
		different convex constraint sets		fully connected undirected	
(Tsianos & Rabbat, 2011)	dual averaging method (Duchi et al., 2012) with delays	common convex constraint set	convex but nonsmooth	connected undirected graphs with fixed and random bounded delays	$O(1/\sqrt{k})$
(Tsianos et al., 2012)	a combination of dual averaging and the push-sum method	common convex constraint set	convex but nonsmooth	strongly connected digraphs with fixed and random bounded delays	$O(1/\sqrt{k})$
(Nedić & Olshevsky, 2015)	a combination of the distributed subgradient method and the push-sum method	unconstrained	convex but nonsmooth	uniformly jointly strongly connected	$O(\ln k/\sqrt{k})$
(Zhu & Martínez, 2012)	primal-dual projected subgradient method	different local constraint sets and global inequality constraint	convex but nonsmooth	weight-balanced and uniformly jointly strongly connected	asymptotic
		identical local constraint sets and global equality constraint			
(Lin et al., 2016)	projected subgradient method	different local constraint sets	convex but nonsmooth	weight-balanced and uniformly jointly strongly connected digraphs without/with arbitrarily bounded delays	asymptotic
(Wang, Lin et al., 2018)	projected subgradient method with step-sizes which are not square summable	unconstrained	convex but nonsmooth	weight-balanced and uniformly jointly strongly connected	$O(1/\sqrt{k})$
		different convex constraint sets			
(Lobel & Ozdaglar, 2011)	distributed subgradient method	unconstrained	convex but nonsmooth	undirected graphs with random link failures	almost sure
(Srivastava & Nedić, 2011)	distributed stochastic gradient method	different local constraint sets	convex but nonsmooth	undirected random graphs	almost sure
(Jakovetić et al., 2018)	distributed stochastic subgradient method	unconstrained	strongly convex and bounded Hessians	undirected random graphs with noisy communications	$O(1/k)$ in mean square
(Lei et al., 2018)	stochastic approximation-based distributed primal-dual algorithm	different local constraint sets	convex and smooth	undirected random graphs with noisy communications	almost sure

Table 2
Discrete-time distributed algorithms with fixed step-sizes.

Study	Approach	Step-size	Constraints type	Cost function	Communication graph	Convergence
(Shi et al., 2015a)	distributed PI	identical	unconstrained	convex and smooth restricted strongly convex and smooth	connected undirected	$O(1/k)$ linear
(Qu & Li, 2018)	a combination of the distributed inexact method and the gradient tracking	identical	unconstrained	convex and smooth strongly convex and smooth	connected undirected	$O(1/k)$ linear
(Jakovetić, 2019)	a unified framework for (Qu & Li, 2018; Shi et al., 2015a)	identical	unconstrained	strongly convex and smooth	connected undirected	linear
(Yao et al., 2018)	distributed PI and finite-time consensus (Sundaram & Hadjicostis, 2007; Yuan et al., 2013)	identical	unconstrained	quadratic	connected undirected	finite-time computation
(Mai & Abed, 2018)	diffusion the coefficients of objective functions and finite-time consensus computation (Sundaram & Hadjicostis, 2007; Yuan et al., 2013)	identical	unconstrained	quadratic	strongly connected	finite-time computation
(Zeng & Yin, 2017)		identical	unconstrained	quasi-strong convex and smooth	strongly connected	linear
(Xi & Khan, 2017)	EXTRA and push-sum based method			restricted strongly convex and smooth		
(Xi, Xin et al., 2018)				strongly convex and smooth		
(Pu et al., 2018; Xin & Khan, 2018)	push–pull based method	identical	unconstrained	strongly convex and smooth	strongly connected	linear
(Nedić et al., 2017a)	DIGing DIGing and push-sum method	identical	unconstrained	strong convex and smooth	uniformly jointly connected uniformly jointly strongly connected	linear
(Xu et al., 2018b)	Bregman splitting method	identical	unconstrained	strongly convex and smooth	undirected stochastic	$O(1/k)$
(Aybat et al., 2018)	distributed proximal method	adaptive	unconstrained	composite convex	connected undirected	$O(1/k)$
(Lei et al., 2016)	a primal-dual algorithm with the projection method	identical	different constraint sets	convex and locally smooth	connected undirected	asymptotic
(Yuan et al., 2011)	a primal-dual subgradient algorithm with a fixed step-size	identical	global inequality constraint	convex but nonsmooth	connected undirected	–
(Yuan et al., 2016)						$O(k^{-1/4})$
(Liu et al., 2017)	a projection algorithm with a fixed step-size	identical	global equality constraint	convex but nonsmooth	connected undirected	asymptotic
(Xu et al., 2015)	augmented distributed gradient method	uncoordinated	unconstrained	convex and smooth	connected undirected	asymptotic
(Xu et al., 2018a)	asynchronous distributed gradient method	uncoordinated	unconstrained	convex and smooth strongly convex and smooth	undirected stochastic graphs with random failures	almost sure linear
(Nedić et al., 2017b)	DIGing-ATC	uncoordinated	unconstrained	strongly convex and smooth	connected undirected	linear
(Saadatniaki et al., 2018)	push–pull based method	uncoordinated	unconstrained	strongly convex and smooth	uniformly jointly strongly connected	linear

Please cite this article as: T. Yang, X. Yi and J. Wu et al., A survey of distributed optimization, Annual Reviews in Control, https://doi.org/10.1016/j.arcontrol.2019.05.006

Table 3
Continuous-time distributed algorithms.

Study	Approach	Arbitrary initialization	Constraints type	Cost function	Communication graph	Triggering	Convergence
(Wang & Elia, 2010)	distributed PI	yes	unconstrained	convex but nonsmooth	connected undirected	–	asymptotic
(Gharesifard & Cortés, 2014)	distributed PI	yes	unconstrained	convex but nonsmooth	connected undirected	–	asymptotic
				convex and smooth	strongly connected and weight-balanced		
(Kia et al., 2015a)	distributed PI	partial	unconstrained	convex local functions and a strictly convex global function	connected undirected	–	asymptotic
				strongly convex and smooth		–	exponential
				strongly convex and smooth		static event-triggered	exponential to a neighborhood
				strongly convex and smooth	strongly connected and weight-balanced	–	exponential
				strongly convex and smooth	strongly connected and weight-balanced at all times	–	exponential
(Du et al., 2019)	distributed PI (Gharesifard & Cortés, 2014; Wang & Elia, 2010)	yes	unconstrained	convex and continuously differentiable	connected undirected		asymptotic
				restricted strongly convex and locally smooth		static event-triggered	exponential
	distributed PI (Kia et al., 2015a)	partial		convex and continuously differentiable			asymptotic
				restricted strongly convex and locally smooth			exponential
(Lu & Tang, 2012)	zero gradient sum method	no	unconstrained	strongly convex with locally Lipschitz Hessian	connected and unidrected	–	exponential
(Liu & Wang, 2015)	two-layer projection neural network	yes	different local constraint sets	convex but nonsmooth	connected and unidrected	–	asymptotic
(Qiu, Liu et al., 2016)	consensus, subgraient, and projection	yes	identical local constraint set	convex	uniformly jointly connected	–	asymptotic
(Yang et al., 2017b)	distributed PI	yes	different local constraint sets and global equality and inequality constraints	convex on locally bounded feasible region	connected and unidrected	–	asymptotic
(Zhu, Yu et al., 2018)	distributed subgradient method	yes	different local constraint sets and global equality and inequality constraints	convex and nonsmooth local functions and a strictly convex global functions	connected and unidrected	–	asymptotic
(Yang et al., 2017a)	(Kia et al., 2015a) in the presence of delays	partial	unconstrained	strongly convex and smooth	strongly connected and weight-balanced digraphs with time-varying delays	–	asymptotic
(Hatanaka et al., 2018)	distributed PI and passivity	yes	unconstrained	strongly convex and smooth	connected undirected graphs with constant delays	–	asymptotic

(*continued on next page*)

Please cite this article as: T. Yang, X. Yi and J. Wu et al., A survey of distributed optimization, Annual Reviews in Control, https://doi.org/10.1016/j.arcontrol.2019.05.006

Table 3 (*continued*)

Study	Approach	Arbitrary initialization	Constraints type	Cost function	Communication graph	Triggering	Convergence
(Guo & Chen, 2018)	(Lu & Tang, 2012) under delays	no	unconstrained	strongly convex with locally Lipschitz Hessians	undirected connected graphs/strongly connected and weight-balanced digraphs at all times with time-varying delays	–	asymptotic
(Doan et al., 2017)	distributed gradient method under delays	yes	different local constraint sets	convex and differentiable	undirected connected graphs with uniform constant delays	–	$O(\ln k/\sqrt{k})$
(Chen & Ren, 2016)	(Lu & Tang, 2012) with ETM	no	unconstrained		strongly connected and weight-balanced	periodic triggered	exponential
(Liu & Chen, 2016)				strongly convex with locally Lipschitz Hessians		static event-triggered	
(Du, Yi et al., 2018)					connected undirected	dynamic event-triggered	
(Lin et al., 2017)	a distributed algorithm with nonuniform state-dependent gradient gains	no	identical local constraint sets	convex and differentiable	connected undirected at all times	–	asymptotic
	a combination of a distributed tracking algorithm and a dynamic averaging estimator	no		convex and twice differentiable		–	finite-time
(Pilloni et al., 2016)	distributed PI-like	no	unconstrained	quadratic	connected undirected	–	finite-time
(Feng & Hu, 2017)	finite-time consensus protocol in Yu and Long (2015)	yes	unconstrained	quadratic	connected undirected graphs with communication uncertainties	–	finite-time

Table 4
Distributed algorithms for the optimal DER coordination problem.

Study	Algorithm Type	Approach	Cost function	Communication graph	Communication imperfection	Triggering	Convergence
(Zhang & Chow, 2012)	discrete-time fixed step-size	leader–follower consensus-based	quadratic	connected undirected	–	–	asymptotic
(Kar & Hug, 2012)	discrete-time diminishing step-size	consensus and innovation approach (Kar & Moura, 2013)	quadratic	connected undirected	–	–	asymptotic
(Tang et al., 2018)	discrete-time fixed step-size	EXTRA (Shi et al., 2015a)	quadratic	connected undirected	–	–	asymptotic
(Dominguez-Garcia, Cady et al., 2012)	discrete-time fixed step-size	ratio consensus (Domínguez-García & Hadjicostis, 2011)	quadratic	strongly connected	–	–	asymptotic
(Yang et al., 2013)	discrete-time fixed step-size	surplus (Cai & Ishii, 2012)	quadratic	strongly connected	–	–	asymptotic
(Xing et al., 2015)	discrete-time fixed step-size	bisection method and consensus-like	strictly convex	strongly connected	–	–	asymptotic
(Yang, Lu et al., 2017)	discrete-time diminishing step-size	push-sum based algorithm (Nedić & Olshevsky, 2015)	strictly convex	uniformly jointly strongly connected	bounded time-varying delays	–	asymptotic
(Du, Yao et al., 2018)	discrete-time fixed step-size	push–pull method	strictly convex	strongly connected uniformly jointly strongly connected	–	–	asymptotic
(Doan & Beck, 2017)	discrete-time diminishing step-size	distributed Lagrangian method	convex but nonsmooth	connected undirected	–	–	asymptotic
Doan and Beck (2018)				uniformly jointly connected			
(Cherukuri & Cortés, 2015; 2016)	continuous-time	Laplacian nonsmooth gradient dynamics	convex, continuous, and locally Lipschitz	strongly connected and weight-balanced	–	–	asymptotic
(Yi et al., 2016)	continuous-time	projected dynamics	strictly convex and smooth strongly convex	connected undirected	–	–	asymptotic exponential
(Bai et al., 2019)	continuous-time	saddle point dynamics	quadratic	connected undirected	–	–	finite-time to a neighborhood
Zhao, Duan et al. (2019)	discrete-time fixed step-size	the algorithm proposed in Yang et al. (2013)	quadratic	strongly connected	uniform constant delays	–	asymptotic
(Zhu et al., 2016)	continuous-time	consensus-based	quadratic	connected undirected	uniform constant delays	–	asymptotic
(Chen & Zhao, 2018)	continuous-time	consensus-based	quadratic	connected undirected	nonuniform constant delays	–	asymptotic
(Somarakis & Baras, 2015; Somarakis et al., 2016)	continuous-time	consensus-based	quadratic	uniformly jointly connected	time-varying bounded delays	–	asymptotic
(Wu, Yang, Wu et al., 2017)	discrete-time diminishing step-size	push sum method (Nedić & Olshevsky, 2015) and running sum method (Dominguez-Garcia, Hadjicostis et al., 2012; Hadjicostis et al., 2016)	quadratic	strongly connected with a positive probability	packet-drops	–	almost sure
(Li et al., 2016)	discrete-time	θ-logarithmic barrier-based	quadratic	connected undirected	–	yes	asymptotic
(Zhao, Li et al., 2019)	continuous-time	saddle-point dynamics	strongly convex	connected undirected	–	yes	asymptotic
(Yang, Wu, Sun et al., 2016)	discrete-time	finite-time consensus computation (Sundaram & Hadjicostis, 2007; Yuan et al., 2013)	quadratic	strongly connected	–	–	finite-time
(Chen et al., 2017)	continuous-time	distributed finite-time algorithm based on (Chen et al., 2011)	quadratic	connected undirected	–	–	asymptotic
(Chen & Li, 2018; Li et al., 2017)	continuous-time	sliding mode control	quadratic	connected undirected	–	–	fixed-time

tions which have received much attention recently. Some of them are briefly summarized as follows:

- *Distributed non-convex optimization.* Most existing studies focused on the distributed convex optimization problem. However, in many physical applications, the optimization problem is non-convex. Thus, it is important to develop distributed algorithms to solve the non-convex optimization problem. This challenging yet important problem has drawn attention recently from various communities, such as control, signal processing, and machine learning (see, e.g., Chatzipanagiotis & Zavlanos, 2017; Hong, Hajinezhad, & Zhao, 2017; Lorenzo & Scutari, 2016; Matei & Baras, 2017; Tatarenko & Touri, 2017; Tian, Sun, Du, & Scutari, 2018; Wai, Lafond, Scaglione, & Moulines, 2017; Zeng & Yin, 2018; Zhu & Martínez, 2013). In these studies, distributed algorithms have been developed to solve unconstrained and constrained non-convex optimization problems over either fixed or time-varying graphs, either undirected or directed. However, this direction is far from being complete. For example, it is interesting to investigate the performance of these existing algorithms in case of noisy gradients. Another interesting direction is to develop event-trigger communication schemes for these existing algorithms to reduce the communication overheads. It is also worthy to apply these existing distributed algorithms to the non-convex optimization problem with theoretical guarantees to solve the non-convex optimal power flow problem.
- *Distributed resilient optimization.* The common assumption in the existing distributed optimization literature is that all agents cooperate to learn the optimal soultion in a collaborative manner. However, in networked cyber-physical systems, some agents may become adversarial due to failures or malicious attacks. Therefore, it is important to investigate the performance of the existing distributed optimization algorithms in the presence of adversarial agents. Although distributed resilient consensus has been well studied (see, e.g., LeBlanc, Zhang, Koutsoukos, & Sundaram, 2013; Pasqualetti, Bicchi, & Bullo, 2012; Sundaram & Hadjicostis, 2011), distributed resilient optimization with adversarial agents is less studied and there are only a limited number of studies (see, e.g., Su & Vaidya, 2016; Sundaram & Gharesifard, 2019; Zhao, He, & Wang, 2017). These results established sufficient and/or necessary conditions under which the proposed distributed algorithms ensure that the non-adversarial agents converge to the convex hull of the local minimizers even in the presence of adversarial agents. However, the results focused on distributed algorithms with diminishing step-sizes. It is interesting to develop distributed resilient optimization algorithms with fixed step-sizes.
- *Distributed online convex optimization.* Another common assumption in the existing distributed optimization literature is that every agent knows its local private convex objective function in advance. However, in many applications, there is no prior knowledge of the objective functions since the information is highly uncertain and unpredictable. For example, in a microgrid with a high penetration of DERs such as wind generators and solar panels, there is high uncertainty of power generation. Thus, the uncertain and unpredictable features of DERs need to be taken into account to design a more accurate energy management system for microgrids (Ma, Wang, Gupta, & Chen, 2018). This issue can be addressed within the framework of distributed online convex optimization (DOCO). Discrete-time DOCO with the time-invariant constraint set and inequality constraints has been studied in Koppel, Jakubiec, and Ribeiro (2015); Lee and Zavlanos (2016); Li, Yi, and Xie (2018); Tsianos and Rabbat (2012) and Yi, Yang, Wu, and Johansson (2019b) and continuous-time DOCO without constraints has been considered in Zhang et al. (2017). It is an important future research direction to consider DOCO with time-varying constraint sets, which is more challenging and practical.
- *Coordination of multi-type of DERs.* Most existing studies for DER coordination only considered coordination of distributed generators. It is desirable to consider other types of resources, such as energy storages and demand response. In fact, some results are available in this direction. For example, the authors of Cherukuri and Cortés (2018); Hug et al. (2015); Wu, Yang, Stoorvogel, and Stoustrup (2017); Yang, Wu, Stoorvogel, and Stoustrup (2016) and He et al. (2019) studied the coordination between distributed generators and energy storages, while the authors of Li, Chen, and Low (2011); Wu, Lian, Sun, Yang, and Hansen (2017) and Qin, Wan, Yu, Li, and Li (2019) studied the coordination between distributed generators and demand response. Most of these existing studies assumed reliable communication networks. However, the communication network in the distribution network is still under-deployed and has limited capabilities compared to that for the bulk power transmission network (Magnússon, Fischione, & Li, 2017; Zhang, Shi, Zhu, Dall'Anese, & Başar, 2018). Such a communication network could suffer from communication imperfections, such as switching topologies, time delays, and packet drops. An interesting future research direction is to develop distributed algorithms to optimally coordinate various distributed energy resources in the presence of these communication imperfections.

Conflict of interest

None.

References

Arrow, K., Huwicz, L., & Uzawa, H. (1958). *Studies in linear and non-linear programming*. Stanford University Press.
Aström, K., & Kumar, P. R. (2014). Control: A perspective. *Automatica, 50*(1), 3–43.
Aström, K. J., & Bernhardsson, B. M. (2002). Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In *Proceedings of the 41st IEEE conference on decision and control* (pp. 2011–2016).
Aybat, N., Wang, Z., & Iyengar, G. (2015). An asynchronous distributed proximal gradient method for composite convex optimization. In *Proceedings of the 32nd international conference on machine learning (ICML)* (pp. 2454–2462).
Aybat, N. S., & Hamedani, E. Y. (2016). A primal-dual method for conic constrained distributed optimization problems. In *Proceedings of the 30th international conference on neural information processing systems* (pp. 5056–5064).
Aybat, N. S., Wang, Z., Lin, T., & Ma, S. (2018). Distributed linearized alternating direction method of multipliers for composite convex consensus optimization. *IEEE Transactions on Automatic Control, 63*(1), 5–20.
Bai, L., Ye, M., Sun, C., & Hu, G. (2019). Distributed economic dispatch control via saddle point dynamics and consensus algorithms. *IEEE Transactions on Control Systems Technology, 27*(2), 898–905.
Bénézit, F., Blondel, V., Thiran, P., Tsitsiklis, J. N., & Vetterli, M. (2010). Weighted gossip: Distributed averaging using non-doubly stochastic matrices. In *Proceedings of the IEEE international symposium on information theory* (pp. 1753–1757).
Bertsekas, D., & Tsitsiklis, J. (1989). *Parallel and distributed computation: Numerical methods*. Prentice-Hall.
Bertsekas, D. P., Nedić, A., & Ozdaglar, A. (2003). *Convex analysis and optimization*. Belmont, MA: Athena Scientific.
Bhat, S. P., & Bernstein, D. S. (2000). Finite-time stability of continuous autonomous systems. *SIAM Journal on Control and Optimization, 38*(3), 751–766.
Bidram, A., Lewis, F. L., & Davoudi, A. (2014). Distributed control systems for small--scale power networks: Using multiagent cooperative control theory. *IEEE Control Systems Magazine, 34*(6), 56–77.
Boyd, S., Parikh, N., Chu, E., Peleato, B., & Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning, 3*(1), 1–122.
Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.
Bullo, F. (2017). *Lectures on network systems, version 0.95*. with contributions by Cortes, J., Dörfler, F., Martinez, S.
Cai, K., & Ishii, H. (2012). Average consensus on general strongly connected digraphs. *Automatica, 48*(11), 2750–2761.
Cao, M., Morse, A. S., & Anderson, B. D. O. (2008). Reaching a consensus in a dynamically changing environment: convergence rates, measurement delays, and asynchronous events. *SIAM Journal of Control and Optimization, 47*(2), 601–623.

Cao, Y., Yu, W., Ren, W., & Chen, G. (2013). An overview of recent progress in the study of distributed multi-agent coordination. *IEEE Transactions on Industrial Informatics, 9*(1), 427–438.
Carli, R., Notarstefano, G., Schenato, L., & Varagnolo, D. (2015). Analysis of Newton-Raphson consensus for multi-agent convex optimization under asynchronous and lossy communication. In *Proceedings of the 54th IEEE conference on decision and control (CDC)* (pp. 418–424).
Chang, T. H., Nedić, A., & Scaglione, A. (2014). Distributed constrained optimization by consensus-based primal-dual perturbation method. *IEEE Transactions on Automatic Control, 59*(6), 1524–1538.
Charalambous, T., Yuan, Y., Yang, T., Pan, W., Hadjicostis, C. N., & Johansson, M. (2015). Distributed finite-time average consensus in digraphs in the presence of time-delays. *IEEE Transactions on Control of Network Systems, 2*(4), 370–381.
Chatzipanagiotis, N., Dentcheva, D., & Zavlanos, M. M. (2015). An augmented Lagrangian method for distributed optimization. *Mathematical Programming, 152*(1), 405–434.
Chatzipanagiotis, N., & Zavlanos, M. M. (2016). A distributed algorithm for convex constrained optimization under noise. *IEEE Transactions on Automatic Control, 61*(9), 2496–2511.
Chatzipanagiotis, N., & Zavlanos, M. M. (2017). On the convergence of a distributed Augmented Lagrangian method for non-convex optimization. *IEEE Transactions on Automatic Control, 62*(9), 4405–4420.
Chen, G., Lewis, F. L., & Xie, L. (2011). Finite-time distributed consensus via binary control protocols. *Automatica, 47*(9), 1962–1968.
Chen, G., & Li, Z. (2018). A fixed-time convergent algorithm for distributed convex optimization in multi-agent systems. *Automatica, 95*, 539–543.
Chen, G., Ren, J., & Feng, E. N. (2017). Distributed finite-time economic dispatch of a network of energy resources. *IEEE Transactions on Smart Grid, 8*(2), 822–832.
Chen, G., & Zhao, Z. (2018). Delay effects on consensus-based distributed economic dispatch algorithm in microgrid. *IEEE Transactions on Power Systems, 33*(1), 602–612.
Chen, W., & Ren, W. (2016). Event-triggered zero-gradient-sum distributed consensus optimization over directed networks. *Automatica, 65*, 90–97.
Cherukuri, A., & Cortés, J. (2015). Distributed generator coordination for initialization and anytime optimization in economic dispatch. *IEEE Transactions on Control of Network Systems, 2*(3), 226–237.
Cherukuri, A., & Cortés, J. (2016). Initialization-free distributed coordination for economic dispatch under varying loads and generator commitment. *Automatica, 74*, 183–193.
Cherukuri, A., & Cortés, J. (2018). Distributed coordination of DERs with storage for dynamic economic dispatch. *IEEE Transactions on Automatic Control, 63*(3), 835–842.
Cherukuri, A., Mallada, E., & Cortés, J. (2016). Asymptotic convergence of constrained primal-dual dynamics. *Systems & Control Letters, 96*, 110–117.
Deng, Z., Liang, S., & Hong, Y. (2018). Distributed continuous-time algorithms for resource allocation problems over weight-balanced digraphs. *IEEE Transactions on Cybernetics, 48*(11), 3116–3125.
Deng, Z., Liang, S., & Yu, W. (2018). Distributed optimal resource allocation of second-order multiagent systems. *International Journal of Robust and Nonlinear Control, 28*(14), 4246–4260.
Deng, Z., Wang, X., & Hong, Y. (2017). Distributed optimisation design with triggers for disturbed continuous-time multi-agent systems. *IET Control Theory & Applications, 11*(2), 282–290.
Dhingra, N. K., Khong, S. Z., & Jovanović, M. R. (2019). The proximal augmented Lagrangian method for nonsmooth composite optimization. *IEEE Transactions on Automatic Control*. https://ieeexplore.ieee.org/abstract/document/8449971.
Dimarogonas, D. V., Frazzoli, E., & Johansson, K. H. (2012). Distributed event-triggered control for multi-agent systems. *IEEE Transactions on Automatic Control, 57*(5), 1291–1297.
Ding, L., Han, Q., Ge, X., & Zhang, X. (2018). An overview of recent advances in event-triggered consensus of multiagent systems. *IEEE Transactions on Cybernetics, 48*(4), 1110–1123.
Ding, L., Wang, L. Y., Yin, G., Zheng, W. X., & Han, Q. (2019). Distributed energy management for smart grids with an event-triggered communication scheme. *IEEE Transactions on Control Systems Technology*. to appear
Doan, T. T., & Beck, C. L. (2018). Distributed resource allocation over dynamic networks with uncertainty. arXiv:1708.03543.
Doan, T. T., & Beck, C. L. (2017). Distributed Lagrangian methods for network resource allocation. In *Proceedings of the IEEE conference on control technology and applications (CCTA)* (pp. 650–655).
Doan, T. T., Beck, C. L., & Srikant, R. (2017). On the convergence rate of distributed gradient methods for finite-sum optimization under communication delays. In *Proceedings of the ACM on Measurement and Analysis of Computing Systems - SIGMETRICS: 1* (pp. 37:1–37:27).
Dominguez-Garcia, A., Cady, S., & Hadjicostis, C. N. (2012). Decentralized optimal dispatch of distributed energy resources. In *Proceedings of the 51st IEEE conference decision and control (CDC)* (pp. 3688–3693).
Domínguez-García, A. D., & Hadjicostis, C. N. (2011). Distributed strategies for average consensus in directed graphs. In *Proc. IEEE decision control european control conf.* (pp. 2124–2129).
Domínguez-García, A. D., & Hadjicostis, C. N. (2015). Distributed resource coordination in networked systems described by digraphs. *Systems & Control Letters, 82*, 33–39.
Dominguez-Garcia, A. D., Hadjicostis, C. N., & Vaidya, N. H. (2012). Resilient networked control of distributed energy resources. *IEEE Journal on Selected Areas in Communications, 30*(6), 1137–1148.
Dong, X., Hua, Y., Zhou, Y., Ren, Z., & Zhong, Y. (2018). Theory and experiment on formation-containment control of multiple multirotor unmanned aerial vehicle systems. *IEEE Transactions on Automation Science and Engineering, 16*(1), 229–240.
Dörfler, F. (2019). Distributed consensus-based optimization. *Lecture note on advanced topics in control 2017: Distributed systems & control*. http://people.ee.ethz.ch/~floriand/docs/Teaching/ATIC_2018/Optimization_Lecture.pdf(last accessed: April 4)
Dörfler, F., Simpson-Porco, J., & Bullo, F. (2016). Breaking the hierarchy: distributed control and economic optimality in microgrids. *IEEE Transactions on Control of Network Systems, 3*(3), 241–253.
Du, W., Yao, L., Wu, D., Li, X., Liu, G., & Yang, T. (2018). Accelerated distributed energy management for microgrids. In *Proceedings of the IEEE power and energy society general meeting*.
Du, W., Yi, X., George, J., Johansson, K. H., & Yang, T. (2018). Distributed optimization with dynamic event-triggered mechanisms. In *Proceedings of the 57th IEEE conference on decision and control (CDC)* (pp. 969–974).
Du, W., Yi, X., Zhang, S., George, J., & Yang, T. (2019). Event-triggered proportional-integral algorithms for distributed optimization. *submitted to the 58th IEEE conference on decision and control (CDC)*.
Duchi, J. C., Agarwal, A., & Wainwright, M. J. (2012). Dual averaging for distributed optimization: Convergence analysis and network scaling. *IEEE Transactions on Automatic Control, 57*(3), 592–606.
Falsone, A., Margellos, K., Garatti, S., & Prandini, M. (2017). Dual decomposition for multi-agent distributed optimization with coupling constraints. *Automatica, 84*, 149–158.
Feijer, D., & Paganini, F. (2010). Stability of primal-dual gradient dynamicsand applications to network optimization. *Automatica, 46*(12), 1974–1981.
Feng, Z., & Hu, G. (2017). Finite-time distributed optimization with quadratic objective functions under uncertain information. In *Proceedings of the 56th IEEE conference on decision and control (CDC)* (pp. 208–213).
Freeman, R. A., Yang, P., & Lynch, K. M. (2006). Stability and convergence properties of dynamic average consensus estimators. In *Proceedings of the 45th IEEE conference on decision and control (CDC)* (pp. 338–343).
Fridman, E. (2014). Tutorial on Lyapunov-based methods for time-delay systems. *European Journal of Control, 20*(6), 271–283.
Gao, W., Gao, J., Ozbay, K., & Jiang, Z. (2019). Reinforcement-learning-based cooperative adaptive cruise control of buses in the lincoln tunnel corridor with time–varying topology. *IEEE Transactions on Intelligent Transportation Systems*. to appear
Gharesifard, B., & Cortés, J. (2014). Distributed continuous-time convex optimization on weight-balanced digraphs. *IEEE Transactions on Automatic Control, 59*(3), 781–786.
Girard, A. (2015). Dynamic triggering mechanisms for event-triggered control. *IEEE Transactions on Automatic Control, 60*(7), 1992–1997.
Giselsson, P. (2018). *Large-scale and distributed optimization*. Springer International Publishing.
Godsi, C., & Royle, G. F. (2001). *Algebraic graph theory*: 207. New York: Springer-Verlag.
Guo, J., Hug, G., & Tonguz, O. K. (2017). A case for nonconvex distributed optimization in large-scale power systems. *IEEE Transactions on Power Systems, 32*(5), 3842–3851.
Guo, Z., & Chen, G. (2018). Distributed zero-gradient-sum algorithm for convex optimization with time-varying communication delays and switching networks. *International Journal of Robust and Nonlinear Control, 28*(16), 4900–4915.
Hadjicostis, C. N., & Charalambous, T. (2014). Average consensus in the presence of delays in directed graph topologies. *IEEE Transactions on Automatic Control, 59*(3), 763–768.
Hadjicostis, C. N., Domínguez-García, A. D., & Charalambous, T. (2018). Distributed averaging and balancing in network systems with applications to coordination and control. *Foundations and Trends in Systems and Control, 5*(2–3), 99–292.
Hadjicostis, C. N., Vaidya, N., & Dominguez-Garcia, A. (2016). Robust distributed average consensus via exchange of running sums. *IEEE Transactions on Automatic Control, 61*(6), 1492–1507.
Hamedani, E. Y., & Aybat, N. S. (2017). Multi-agent constrained optimization of a strongly convex function over time-varying directed networks. In *Proceedings of the 55th annual Allerton conference on communication control, and computing (Allerton)* (pp. 518–525).
Han, S. (2019). Computational convergence analysis of distributed gradient descent for smooth convex objective function. In *Proceedings of the IEEE american control conference*. To appear
Hatanaka, T., Chopra, N., Ishizaki, T., & Li, N. (2018). Passivity-based distributed optimization with communication delays using PI consensus algorithm. *IEEE Transactions on Automatic Control, 63*(12), 4421–4428.
He, X., Yu, J., Huang, T., & Li, C. (2019). Distributed power management for dynamic economic dispatch in the multimicrogrids environment. *IEEE Transactions on Control Systems Technology*. https://ieeexplore.ieee.org/abstract/document/8329211.
Heemels, W. P. M. H., Johansson, K. H., & Tabuada, P. (2012). An introduction to event-triggered and self-triggered control. In *Proceedings of the 51st IEEE conference on decision and control* (pp. 3270–3285).

Hendrickx, J. M., & Tsitsiklis, J. N. (2015). Fundamental limitations for anonymous distributed systems with broadcast communications. In *Proceedings of the 53rd annual Allerton conference on communication control, and computing (Allerton)* (pp. 9–16).

Hespanha, J., Naghshtabrizi, P., & Xu, Y. (2007). A survey of recent results in networked control systems. *Proceedings of the IEEE, 95*(1), 138–162.

Hong, M., Hajinezhad, D., & Zhao, M. M. (2017). Prox-PDA: The proximal primal–dual algorithm for fast distributed nonconvex optimization and learning over networks. In *Proceedings of the 34th international conference on machine learning (ICML)* (pp. 1529–1538).

Hu, B., Guan, Z., Chen, G. R., & Shen, X. S. (2019). A distributed hybrid event-time–driven scheme for optimization over sensor networks. *IEEE Transactions on Industrial Electronics, 66*(9), 7199–7208.

Hug, G., Kar, S., & Wu, C. (2015). Consensus + innovations approach for distributed multiagent coordination in a microgrid. *IEEE Transactions on Smart Grid, 6*(4), 1893–1903.

Jakovetić, D. (2019). A unification and generalization of exact distributed first order methods. *IEEE Transactions on Signal and Information Processing over Networks, 5*(1), 31–46.

Jakovetić, D., Bajovic, D., Sahu, A., & Kar, S. (2018). Convergence rates for distributed stochastic optimization over random networks. In *Proceedings of the 57th IEEE conference on decision and control* (pp. 4238–4245).

Jakovetić, D., Xavier, J., & Moura, J. M. F. (2014a). Convergence rates of distributed Nesterov-like gradient methods on random networks. *IEEE Transactions on Signal Processing, 62*(4), 868–882.

Jakovetić, D., Xavier, J., & Moura, J. M. F. (2014b). Fast distributed gradient methods. *IEEE Transactions on Automatic Control, 59*(5), 1131–1146.

Johansson, B., Keviczky, T., Johansson, M., & Johansson, K. H. (2008). Subgradient methods and consensus algorithms for solving convex optimization problems. In *Proceedings of the 47th IEEE conference on decision and control (CDC)* (pp. 4185–4190).

Johansson, K. H., Egerstedt, M., Lygeros, J., & Sastry, S. (1999). On the regularization of Zeno hybrid automata. *Systems & Control Letters, 38*(3), 141–150.

Kar, S., & Hug, G. (2012). Distributed robust economic dispatch in power systems: A consensus+innovations approach. In *Proceedings of the IEEE power and energy society general meeting*.

Kar, S., & Moura, J. (2013). Consensus + innovations distributed inference over networks. *IEEE Signal Processing Magazine, 30*(3), 99–109.

Kempe, D., Dobra, A., & Gehrke, J. (2003). Gossip-based computation of aggregate information. In *Proceedings of the 44th annual IEEE symposium on foundations of computer science* (pp. 482–491).

Kia, S. S., Cortés, J., & Martínez, S. (2015a). Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication. *Automatica, 55*, 254–264.

Kia, S. S., Cortés, J., & Martínez, S. (2015b). Dynamic average consensus under limited control authority and privacy requirements. *International Journal of Robust and Nonlinear Control, 23*(13), 1941–1966.

Kim, B. H., & Baldick, R. (1997). Coarse-grained distributed optimal power flow. *IEEE Transactions on Power Systems, 12*(2), 932–939.

Koppel, A., Jakubiec, F. Y., & Ribeiro, A. (2015). A saddle point algorithm for networked online convex optimization. *IEEE Transactions on Signal Processing, 63*(19), 5149–5164.

Kraning, M., Chu, E., Lavaei, J., & Boyd, S. (2014). Dynamic network energy management via proximal message passing. *Foundations and Trends in Optimization, 1*(2), 70–122.

Lam, A., Zhang, B., & Tse, D. (2012). Distributed algorithms for optimal power flow problem. In *Proceedings of the IEEE conference on decision and control* (pp. 430–437).

Lamnabhi-Lagarrigue, F., Annaswamy, A., Engell, S., Isaksson, A., Khargonekar, P., Murray, R. M., et al. (2017). Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges. *Annual Reviews in Control, 43*, 1–64.

Lasseter, R., Akhil, A., Marnay, C., Stephens, J., Dagle, J., Guttromson, R., et al. (2003). Integration of distributed energy resources: The CERTS Microgrid concept. *Lawrence Berkeley National Laboratory.*.

LeBlanc, H. J., Zhang, H., Koutsoukos, X., & Sundaram, S. (2013). Resilient asymptotic consensus in robust networks. *IEEE Journal on Selected Areas in Communications, 31*(4), 766–781.

Lee, S., & Zavlanos, M. M. (2016). Distributed primal-dual methods for online constrained optimization. In *Proceedings of the IEEE american control conference* (pp. 7171–7176).

Lei, J., Chen, H. F., & Fang, H. T. (2016). Primal-dual algorithm for distributed constrained optimization. *Systems & Control Letters, 96*, 110–117.

Lei, J., Chen, H. F., & Fang, H. T. (2018). Asymptotic properties of primal-dual algorithm for distributed stochastic optimization over random networks with imperfect communications. *SIAM Journal of Control and Optimization, 56*(3), 2159–2188.

Lessard, L., Recht, B., & Packard, A. (2016). Analysis and design of optimization algorithms via integral quadratic constraints. *SIAM Journal on Optimization, 26*(1), 57–95.

Li, C., Yu, X., Yu, W., Huang, T., & Liu, Z. W. (2016). Distributed event-triggered scheme for economic dispatch in smart grids. *IEEE Transactions on Industrial Informatics, 12*(5), 1775–1785.

Li, C., Yu, X., Zhou, X., & Ren, W. (2017). A fixed time distributed optimization: A sliding mode perspective. In *Proceedings of the 43rd annual conference of the IEEE industrial electronics society* (pp. 8201–8207).

Li, N., Chen, L., & Low, S. H. (2011). Optimal demand response based on utility maximization in power networks. In *Proceedings of the IEEE power and energy society general meeting*.

Li, X., Yi, X., & Xie, L. (2018) Distributed online optimization for multi-agent networks with coupled inequality constraints. arXiv:1805.05573.

Li, Y., Xiang, J., & Wei, W. (2011). Consensus problems for linear time-invariant multi-agent systems with saturation constraints. *IET Control Theory & Applications, 5*(6), 823–829.

Li, Z., Ding, Z., Sun, J., & Li, Z. (2018). Distributed adaptive convex optimization on directed graphs via continuous-time algorithms. *IEEE Transactions on Automatic Control, 63*(5), 1434–1441.

Liang, S., Zeng, X., & Hong, Y. (2018). Distributed nonsmooth optimization with coupled inequality constraints via modified Lagrangian function. *IEEE Transactions on Automatic Control, 63*(6), 1753–1759.

Liberzon, D., & Morse, A. S. (1999). Basic problems in stability and design of switched systems. *IEEE Control Systems Magazine, 19*(5), 59–70.

Lin, P., Ren, W., & Farrell, J. A. (2017). Distributed continuous-time optimization: Nonuniform gradient gains, finite-time convergence, and convex constraint set. *IEEE Transactions on Automatic Control, 62*(5), 2239–2253.

Lin, P., Ren, W., & Song, Y. (2016). Distributed multi-agent optimization subject to nonidentical constraints and communication delays. *Automatica, 65*, 120–131.

Liu, J., & Chen, W. (2016). Distributed convex optimisation with event-triggered communication in networked systems. *International Journal of Systems Science, 47*(16), 3876–3887.

Liu, J., Chen, W., & Dai, H. (2019). Event-triggered zero-gradient-sum distributed convex optimization over networks with time-varying topologies. *International Journal of Control*. doi:10.1080/00207179.2018.1460693.

Liu, Q., & Wang, J. (2015). A second-order multi-agent network for bound-constrained distributed optimization. *IEEE Transactions on Automatic Control, 62*(7), 3461–3467.

Liu, Q., Yang, S., & Hong, Y. (2017). Constrained consensus algorithms with fixed step size for distributed convex optimization over multiagent networks. *IEEE Transactions on Automatic Control, 62*(8), 4259–4265.

Liu, S., Xie, L., & Quevedo, D. E. (2018). Event-triggered quantized communication-based distributed convex optimization. *IEEE Transactions on Control of Network Systems, 5*(1), 167–178.

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

Lorenzo, P. D., & Scutari, G. (2016). NEXT: In-network nonconvex optimization. *IEEE Transactions on Signal and Information Processing over Networks, 2*(2), 120–136.

Low, S. (2014). Convex relaxation of optimal power flow part I: Formulations and equivalence. *IEEE Transactions on Control of Network Systems, 1*(1), 15–27.

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

Lynch, K. M., & Park, F. C. (2017). *Modern robotics: Mechanics, planning, and control*. Cambridge University Press.

Ma, D., & Chen, J. (2019). Delay margin of low-order systems achievable by PID controllers. *IEEE Transactions on Automatic Control, 64*(5), 1958–1973.

Ma, D., Tian, R., Zulfiqar, A., Chen, J., & Chai, T. (2019). Bounds on delay consensus margin of second-order multi-agent systems with robust position and velocity feedback protocol. *IEEE Transactions on Automatic Control*. https://ieeexplore.ieee.org/abstract/document/8552432.

Ma, W. J., Wang, J., Gupta, V., & Chen, C. (2018). Distributed energy management for networked microgrids using online alternating direction method of multipliers with regret. *IEEE Transactions on Smart Grid, 9*(2), 847–856.

Madani, R., Sojoudi, S., & Lavaei, J. (2015). Convex relaxation for optimal power flow problem: Mesh networks. *IEEE Transactions on Power Systems, 30*(1), 199–211.

Magnússon, S., Fischione, C., & Li, N. (2017). Voltage control using limited communication. *IFAC-PapersOnLine, 50*(1), 1–6. 20th IFAC World Congress

Mai, V. S., & Abed, E. H. (2018). Local prediction for enhanced convergence of distributed optimization algorithms. *IEEE Transactions on Control of Network Systems, 5*, 1962–1975.

Mallada, E., Zhao, C., & Low, S. (2017). Optimal load-side control for frequency regulation in smart grids. *IEEE Transactions on Automatic Control, 62*(12), 6294–6309.

Margellos, K., Falsone, A., Garatti, S., & Prandini, M. (2018). Distributed constrained optimization and consensus in uncertain networks via proximal minimization. *IEEE Transactions on Automatic Control, 63*(5), 1372–1387.

Matei, I., & Baras, J. S. (2017) Nonlinear programming methods for distributed optimization. arXiv:1707.04598.

Matei, I., & Baras, J. S. (2011). Performance evaluation of the consensus-based distributed subgradient method under random communication topologies. *IEEE Journal of Selected Topics in Signal Processing, 5*(4), 754–771.

Mateos-Núñez, D., & Cortés, J. (2016). Noise-to-state exponentially stable distributed convex optimization on weight-balanced digraphs. *SIAM Journal on Control and Optimization, 54*(1), 266–290.

Megretski, A., & Rantzer, A. (1997). System analysis via integral quadratic constraints. *IEEE Transactions on Automatic Control, 42*(6), 819–830.

Meng, D., & Moore, K. L. (2016). Learning to cooperate: Networks of formation agents with switching topologies. *Automatica, 64*, 278–293.

Meng, X., Xie, L., & Soh, Y. C. (2016). Distributed event driven optimization for network utility maximization. In *Proceedings of the 55th IEEE conference on decision and control (CDC)* (pp. 2221–2226).

Meng, X., Xie, L., Soh, Y. C., Nowzari, C., & Pappas, G. J. (2015). Periodic event-triggered average consensus over directed graphs. In *Proceedings of the 54th IEEE conference on decision and control (CDC)* (pp. 2055–2060).
Meng, Z., Yang, T., Shi, G., Dimarogonas, D. V., Hong, Y., & Johansson, K. H. (2017). Targeted agreement of multiple Lagrangian systems. *Automatica, 84*, 109–116.
Meng, Z., Zhao, Z., & Lin, Z. (2013). On global leader-following consensus of identical linear dynamic systems subject to actuator saturation. *Systems & Control Letters, 62*, 132–142.
Mokhtari, A., & Ribeiro, A. (2016). DSA: Decentralized double stochastic averaging gradient algorithm. *Journal of Machine Learning Research, 17*, 1–35.
Mokhtari, A., Shi, W., Ling, Q., & Ribeiro, A. (2016). DQM: Decentralized quadratically approximated alternating direction method of multipliers. *IEEE Transactions on Signal Processing, 64*(19), 5158–5173.
Molzahn, D. K., Dörfler, F., Sandberg, H., Low, S. H., Chakrabarti, S., Baldick, R., & Lavaei, J. (2017). A survey of distributed optimization and control algorithms for electric power systems. *IEEE Transactions on Smart Grid, 8*(6), 2941–2962.
Nedić, A. (2015). Convergence rate of distributed averaging dynamics and optimization in networks. *Foundations and Trends in Systems and Control, 2*(1), 1–100.
Nedić, A., & Liu, J. (2018). Distributed optimization for control. *Annual Review of Control, Robotics, and Autonomous Systems, 1*, 77–103.
Nedić, A., & Olshevsky, A. (2015). Distributed optimization over time-varying directed graphs. *IEEE Transactions on Automatic Control, 60*(3), 601–615.
Nedić, A., Olshevsky, A., & Rabbat, M. G. (2018). Network topology and communication-computation tradeoffs in decentralized optimization. *Proceedings of the IEEE, 106*(5), 953–976.
Nedić, A., Olshevsky, A., & Shi, W. (2017a). Achieving geometric convergence for distributed optimization over time-varying graphs. *SIAM Journal on Optimization, 27*(4), 2597–2633.
Nedić, A., Olshevsky, A., Shi, W., & Uribe, C. A. (2017b). Geometrically convergent distributed optimization with uncoordinated step-sizes. In *Proceedings of the IEEE American control conference* (pp. 3950–3955).
Nedić, A., Olshevsky, A., & Shi, W. (2018). Improved convergence rates for distributed resource allocation. In *Proceedings of 57th IEEE 2018 IEEE conference on decision and control (CDC)* (pp. 172–177).
Nedić, A., & Ozdaglar, A. (2009). Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control, 54*(1), 48–61.
Nedić, A., & Ozdaglar, A. (2010). Convergence rate for consensus with delays. *Journal of Global Optimization, 47*(3), 437–456.
Nedić, A., Ozdaglar, A., & Parrilo, P. A. (2010). Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control, 55*(4), 922–938.
Nesterov, Y. (2004). *Introductory lectures on convex optimization: A basic course*. Kluwer Academic Publishers.
Nowzari, C., Cortés, J., & Pappas, G. J. (2018). *Event-triggered control for multi-agent average consensus* (pp. 177–208). Wiley. Ch. 7
Nowzari, C., Garcia, E., & Cortés, J. (2019). Event-triggered communication and control of network systems for multi-agent consensus. *Automatica, 105*, 1–27.
Parikh, N., & Boyd, S. (2014). Proximal algorithms. *Foundations and Trends in Optimization, 1*(3), 123–231.
Pasqualetti, F., Bicchi, A., & Bullo, F. (2012). Consensus computation in unreliable networks: A system theoretic approach. *IEEE Transactions on Automatic Control, 57*(1), 90–104.
Pedrasa, M. A. A., Spooner, T. D., & MacGill, I. F. (2010). Coordinated scheduling of residential distributed energy resources to optimize smart home energy services. *IEEE Transactions on Smart Grid, 1*(2), 134–143.
Peng, Q., & Low, S. H. (2018). Distributed optimal power flow algorithm for radial networks, I: Balanced single phase case. *IEEE Transactions on Smart Grid, 9*(1), 111–121.
Pilloni, A., Pisano, A., Franceschelli, M., & Usai, E. (2016). A discontinuous algorithm for distributed convex optimization. In *Proceedings of the 14th international workshop on variable structure systems (VSS)* (pp. 22–27).
Polyakov, A. (2018). Nonlinear feedback design for fixed-time stabilization of linear control system. *IEEE Transactions on Automatic Control, 57*(8), 2106–2109.
Polyakov, A., Efimov, D., & Perruquetti, W. (2015a). Finite-time and fixed-time stabilization: Implicit Lyapunov function approach. *Automatica, 51*, 332–340.
Polyakov, A., Efimov, D., & Perruquetti, W. (2015b). Nonsingular fixed-time consensus tracking for second-order multi-agent networks. *Automatica, 54*, 305–309.
Pu, S., Shi, W., Xu, J., & Nedić, A. (2018). A push–pull gradient method for distributed optimization in networks. In *Proceedings of the 57th IEEE conference on decision and control (CDC)* (pp. 3385–3390).
Qin, J., Ma, Q., Shi, Y., & Wang, L. (2017). Recent advances in consensus of multi--agent systems: A brief survey. *IEEE Transactions on Industrial Electronics, 64*(6), 4972–4983.
Qin, J., Wan, Y., Yu, X., Li, F., & Li, C. (2019). Consensus-based distributed coordination between economic dispatch and demand response. *IEEE Transactions on Smart Grid*. https://ieeexplore.ieee.org/abstract/document/8356102.
Qiu, Z., Hong, Y., & Xie, L. (2016). Optimal consensus of Euler-Lagrangian systems with kinematic constraints. *IFAC-papersonline, 49*(22), 327–332. 6th IFAC Workshop on Distributed Estimation and Control in Networked Systems NECSYS 2016
Qiu, Z., Liu, S., & Xie, L. (2016). Distributed constrained optimal consensus of multi--agent systems. *Automatica, 68*, 209–215.
Qiu, Z., Liu, S., & Xie, L. (2018). Necessary and sufficient conditions for distributed constrained optimal consensus under bounded input. *International Journal of Robust and Nonlinear Control, 28*(6), 2619–2635.
Qu, G., & Li, N. (2018). Harnessing smoothness to accelerate distributed optimization. *IEEE Transactions on Control of Network Systems, 5*(3), 1245–1260.
Qu, G., & Li, N. (2019). On the exponential stability of primal-dual gradient dynamics. *IEEE Control Systems Letters, 3*(1), 43–48.
Rahbari-Asr, N., Ojha, U., Zhang, Z., & Chow, M. Y. (2014). Incremental welfare consensus algorithm for cooperative distributed generation/demand response in smart grid. *IEEE Transactions on Smart Grid, 6*(5), 2836–2845.
Rahimi, F., & Ipakchi, A. (2010). Demand response as a market resource under the smart grid paradigm. *IEEE Transactions on Smart Grid, 1*(1), 82–88.
Ren, W., & Cao, Y. (2011). Distributed coordination of multi-agent networks. *Communications and control engineering series*. London: Springer Verlag.
Richard, J. P. (2003). Time-delay systems:an overview of some recent advances and open problems. *Automatica, 39*(10), 1667–1694.
Saadatniaki, F., Xin, R., & Khan, U. A. (2018). Optimization over time-varying directed graphs withrow and column-stochastic matrices. arXiv:1810.07393.
Sayed, A. (2014a). Adaptation, learning, and optimization over networks. *Foundations and Trends in Machine Learning, 7*(4–5), 311–801.
Sayed, A. H. (2014b). Diffusion adaptation over networks. In *Academic press library in signal processing, vol. 3 of academic press library in signal processing* (pp. 323–453). Elsevier.
Schenato, L., Sinopoli, B., Franceschetti, M., Poolla, K., & Sastry, S. S. (2007). Foundations of control and estimation over lossy networks. *Proceedings of the IEEE, 95*(1), 163–187.
Seyboth, G. S., Dimarogonas, D. V., & Johansson, K. H. (2013). Event-based broadcasting for multi-agent average consensus. *Automatica, 49*(1), 245–252.
Shi, G., Johansson, K. H., & Hong, Y. (2013). Reaching an optimal consensus: Dynamical systems that compute intersections of convex sets. *IEEE Transactions on Automatic Control, 58*(3), 610–622.
Shi, W., Ling, Q., Wu, G., & Yin, W. (2015a). EXTRA: An exact first-order algorithm for decentralized consensus optimization. *SIAM Journal on Optimization, 25*(2), 944–966.
Shi, W., Ling, Q., Wu, G., & Yin, W. (2015b). A proximal gradient algorithm for decentralized composite optimization. *IEEE Transactions on Signal Processing, 63*(22), 6013–6023.
Shi, X., Wang, Y., Song, S., & Yan, G. (2018). Distributed optimisation for resource allocation with event-triggered communication over general directed topology. *International Journal of Systems Science, 49*(6), 1119–1130.
Somarakis, C., & Baras, J. S. (2015). Distributed solution of the economic dispatch problem in smart grid power systems framework with delays. In *Proceedings of the 54th IEEE conference on decision and control (CDC)* (pp. 6577–6582).
Somarakis, C., Maity, D., & Baras, J. S. (2016). The effect of delays in the economic dispatch problem for smart grid architectures. In *Proceedings of the IEEE American control conference* (pp. 3533–3538).
Song, Y., & Chen, W. (2016). Finite-time convergent distributed consensus optimisation over networks. *IET Control Theory and Applications, 10*(11), 1314–1318.
Spong, M. W., Hutchinson, S., & Vidyasagar, M. (2006). *Robot dynamics and control*. John Wiley & Sons, Inc..
Srivastava, K., & Nedić, A. (2011). Distributed asynchronous constrained stochastic optimization. *IEEE Journal of Selected Topics in Signal Processing, 5*(4), 772–790.
Su, L., & Vaidya, N. (2016). Multi-agent optimization in the presence of byzantine adversaries: Fundamental limits. In *Proceedings of the IEEE American control conference* (pp. 7183–7188).
Sun, A., Phan, D., & Ghosh, S. (2013). Fully decentralized ac optimal power flow algorithms. In *Proceedings of the IEEE power and energy society general meeting.*
Sundaram, S., & Gharesifard, B. (2019). Distributed optimization under adversarial nodes. *IEEE Transactions on Automatic Control, 64*(3), 1063–1076.
Sundaram, S., & Hadjicostis, C. N. (2007). Finite-time distributed consensus in graphs with time-invariant topologies. In *Proceedings of the IEEE American control conference* (pp. 711–716).
Sundaram, S., & Hadjicostis, C. N. (2011). Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Transactions on Automatic Control, 56*(7), 1495–1508.
Sundararajan, A., Hu, B., & Lessard, L. (2017). Robust convergence analysis of distributed optimization algorithms. In *Proceedings of the 55th annual Allerton conference on communication control, and computing (Allerton)* (pp. 1206–1212).
Sundararajan, A., Scoy, B. V., & Lessard, L. 2018. A canonical form for first-order distributed optimization algorithms. arXiv:1809.08709.
Tabuada, P. (2007). Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control, 52*(9), 1680–1685.
Tang, Z., Hill, D. J., & Liu, T. (2018). A novel consensus-based economic dispatch for microgrids. *IEEE Transactions on Smart Grid, 9*(4), 3920–3922.
Tatarenko, T., & Touri, B. (2017). Non-convex distributed optimization. *IEEE Transactions on Automatic Control, 62*(8), 3744–3757.
Tian, Y., Sun, Y., Du, B., & Scutari, G. 2018. ASY-SONATA: Achieving geometric convergence for distributed asynchronous optimization. arXiv:1803.10359.
Tran, N. T., Wang, Y. W., Liu, X. K., & Xiao, J. W. (2017). Distributed optimisation of second-order multi-agent systems by control algorithm using position-only interaction with time-varying delay. *IET Control Theory & Applications, 11*(15), 2549–2558.
Tran, N. T., Wang, Y. W., Liu, X. K., & Xiao, J. W. (2018). Event-triggered gradient-based distributed optimisation for multi-agent systems with state consensus constraint. *IET Control Theory & Applications, 12*(10), 1515–1519.
Tran, N. T., Wang, Y. W., Liu, X. K., Xiao, J. W., & Lei, Y. (2019). Distributed optimization problem for second-order multi-agent systems with event-triggered and time-triggered communication. *Journal of the Franklin Institute*. doi:10.1016/j.jfranklin.2018.02.009.

Tsianos, K. I., Lawlor, S., & Rabbat, M. G. (2012). Push-sum distributed dual averaging for convex optimization. In *Proceedings of the 51st IEEE conference on decision and control (CDC* (pp. 5453–5458).
Tsianos, K. I., & Rabbat, M. G. (2011). Distributed consensus and optimization under communication delays. In *Proceedings of the 49th annual Allerton conference on communication control, and computing (Allerton)* (pp. 974–982).
Tsianos, K. I., & Rabbat, M. G. (2012). Distributed strongly convex optimization. In *Proceedings of the 50th annual Allerton conference on communication control, and computing (Allerton)* (pp. 593–600).
Tsitsiklis, J. N. (1984). *Problems in decentralized decision making and computation*. MIT, Cambridge, MA Ph.d. thesis.
Tsitsiklis, J. N., Bertsekas, D. P., & Athans, M. (1986). Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control, 31*(9), 803–812.
Varagnolo, D., Zanella, F., Cenedese, A., Pillonetto, G., & Schenato, L. (2016). Newton-Raphson consensus for distributed convex optimization. *IEEE Transactions on Automatic Control, 61*(4), 994–1009.
Wai, H., Lafond, J., Scaglione, A., & Moulines, E. (2017). Decentralized Frank-Wolfe algorithm for convex and nonconvex problems. *IEEE Transactions on Automatic Control, 62*(11), 5522–5537.
Wan, P., & Lemmon, M. D. (2009). Event-triggered distributed optimization in sensor networks. In *Proceedings of the international conference on information processing in sensor networks* (pp. 49–60).
Wang, D., Gupta, V., & Wang, W. (2018). An event-triggered protocol for distributed optimal coordination of double-integrator multi-agent systems. *Neurocomputing, 319*, 34–41.
Wang, D., Wang, D., & Wang, W. (2016). An optimal algorithm for high-order multi-agent systems with event-triggered communication. In *31st youth academic annual conference of Chinese association of automation (YAC)* (pp. 235–240).
Wang, D., Wang, Z., Chen, M., & Wang, W. (2018). Distributed optimization for multi-agent systems with constraints set and communication time-delay over a directed graph. *Information Sciences, 438*, 1–14.
Wang, J., & Elia, N. (2010). Control approach to distributed optimization. In *Proceedings of the 48th annual Allerton conference on communication control, and computing (Allerton)* (pp. 557–561).
Wang, J., & Elia, N. (2011). A control perspective for centralized and distributed convex optimization. In *Proceedings of the 50th IEEE conference on decision and control and european control conference* (pp. 3800–3805).
Wang, L., & Xiao, F. (2010). Finite-time consensus problems for networks of dynamic agents. *IEEE Transactions on Automatic Control, 55*(4), 950–955.
Wang, P., Lin, P., Ren, W., & Song, Y. (2018). Distributed subgradient-based multiagent optimization with more general step sizes. *IEEE Transactions on Automatic Control, 63*(7), 2295–2302.
Wang, X., Hong, Y., & Ji, H. (2016). Distributed optimization for a class of nonlinear multiagent systems with disturbance rejection. *IEEE Transactions on Cybernetics, 46*(7), 1655–1666.
Wang, X., & Lemmon, M. D. (2011). Event-triggering in distributed networked control systems. *IEEE Transactions on Automatic Control, 56*(3), 586–601.
Wang, Z., Wang, D., Sun, J., & Wang, W. (2018). Exponential convergence rate of distributed optimisation for multi-agent systems with constraints set over a directed graph. *IET Control Theory & Applications, 12*(9), 1201–1207.
Wei, E., Ozdaglar, A., & Jadbabaie, A. (2013a). A distributed Newton method for network utility maximization-I: Algorithm. *IEEE Transactions on Automatic Control, 58*(9), 2162–2175.
Wei, E., Ozdaglar, A., & Jadbabaie, A. (2013b). A distributed Newton method for network utility maximization-II: Convergence. *IEEE Transactions on Automatic Control, 58*(9), 2176–2188.
Wu, D., Lian, J., Sun, Y., Yang, T., & Hansen, J. (2017). Hierarchical control framework for integrated coordination between distributed energy resources and demand response. *Electric Power Systems Research, 150*, 45–54.
Wu, D., Yang, T., Stoorvogel, A. A., & Stoustrup, J. (2017). Distributed optimal coordination for distributed energy resources in power systems. *IEEE Transactions on Automation Science and Engineering, 14*(2), 414–424.
Wu, J., Yang, T., Wu, D., Kalsi, K., & Johansson, K. H. (2017). Distributed optimal dispatch of distributed energy resources over lossy communication networks. *IEEE Transactions on Smart Grid, 8*(6), 3125–3137.
Xi, C., & Khan, U. A. (2017). DEXTRA: A fast algorithm for optimization over directed graphs. *IEEE Transactions on Automatic Control, 62*(10), 4980–4993.
Xi, C., Mai, V. S., Xin, R., Abed, E., & Khan, U. A. (2018). Linear convergence in optimization over directed graphs with row-stochastic matrices. *IEEE Transactions on Automatic Control, 63*(10), 3558–3565.
Xi, C., Xin, R., & Khan, U. A. (2018). ADD-OPT: Accelerated distributed directed optimization. *IEEE Transactions on Automatic Control, 63*(5), 1329–1339.
Xie, P., You, K., Tempo, R., Song, S., & Wu, C. (2018). Distributed convex optimization with inequality constraints over time-varying unbalanced digraphs. *IEEE Transactions on Automatic Control, 63*(12), 4331–4337.
Xie, Y., & Lin, Z. (2017). Global optimal consensus of multi-agent systems with bounded controls. *Systems & Control Letters, 102*, 104–111.
Xie, Y., & Lin, Z. (2019). Global optimal consensus for higher-order multi-agent systems with bounded controls. *Automatica, 99*, 301–307.
Xin, R., & Khan, U. A. (2018). A linear algorithm for optimization over directed graphs with geometric convergence. *IEEE Control Systems Letters, 2*(3), 325–330.
Xin, R., Xi, C., & Khan, U. A. (2019). FROST-fast row-stochastic optimization with uncoordinated step-sizes. *EURASIP Journal on Advances in Signal Processing, 1*, 1–14.
Xing, H., Mou, Y., Fu, M., & Lin, Z. (2015). Distributed bisection method for economic power dispatch in smart grid. *IEEE Transactions on Power Systems, 30*(6), 3024–3035.
Xu, J., Zhu, S., Soh, Y. C., & Xie, L. (2015). Augmented distributed gradient methods for multi-agent optimization under uncoordinated constant stepsizes. In *Proceedings of 54th IEEE conference on decision and control (CDC)pp. 2055–2060*.
Xu, J., Zhu, S., Soh, Y. C., & Xie, L. (2018a). Convergence of asynchronous distributed gradient methods over stochastic networks. *IEEE Transactions on Automatic Control, 63*(2), 434–448.
Xu, J., Zhu, S., Soh, Y. C., & Xie, L. (2018b). A Bregman splitting scheme for distributed optimization over networks. *IEEE Transactions on Automatic Control, 63*(11), 3809–3824.
Xu, Y., Han, T., Cai, K., Lin, Z., Yan, G., & Fu, M. (2017). A distributed algorithm for resource allocation over dynamic digraphs. *IEEE Transactions on Signal Processing, 65*(10), 2600–2612.
Yang, B., & Johansson, M. (2010). *Distributed optimization and games: A tutorial overview* (pp. 109–148). London: Springer.
Yang, S., Liu, Q., & Wang, J. (2017a). Distributed optimization based on a multiagent system in the presence of communication delays. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 47*(5), 717–728.
Yang, S., Liu, Q., & Wang, J. (2017b). A multi-agent system with a proportional-integral protocol for distributed constrained optimization. *IEEE Transactions on Automatic Control, 60*(12), 3310–3315.
Yang, S., Tan, S., & Xu, J. X. (2013). Consensus based approach for economic dispatch problem in a smart grid. *IEEE Transactions on Power Systems, 28*(4), 4416–4426.
Yang, T., Lu, J., Wu, D., Wu, J., Shi, G., Meng, Z., et al. (2017). A distributed algorithm for economic dispatch over time-varying directed networks with delays. *IEEE Transactions on Industrial Electronics, 64*(6), 5095–5106.
Yang, T., Meng, Z., Dimarogonas, D. V., & Johansson, K. H. (2014). Global consensus for discrete-time multi-agent systems with input saturation constraints. *Automatica, 50*(2), 499–506.
Yang, T., Wan, Y., Wang, H., & Lin, Z. (2018). Global optimal consensus for discrete-time multi-agent systems with bounded controls. *Automatica, 102*, 104–111.
Yang, T., Wu, D., Stoorvogel, A., & Stoustrup, J. (2016). Distributed coordination of energy storage with distributed generators. In *Proceedings of the IEEE power and energy society general meeting*.
Yang, T., Wu, D., Sun, Y., & Lian, J. (2015). Impacts of time delays on distributed algorithms for economic dispatch. In *Proceedings of the IEEE power and energy society general meeting*.
Yang, T., Wu, D., Sun, Y., & Lian, J. (2016). Minimum-time consensus based approach for power system applications. *IEEE Transactions on Industrial Electronics, 63*(2), 1318–1328.
Yao, L., Yuan, Y., Sundaram, S., & Yang, T. (2018). Distributed finite-time optimization. In *Proceedings of the 14th IEEE international conference on control and automation (ICCA)* (pp. 147–154).
Yi, P., Hong, Y., & Liu, F. (2015). Distributed gradient algorithm for constrained optimization with application to load sharing in power systems. *Systems & Control Letters, 83*, 45–52.
Yi, P., Hong, Y., & Liu, F. (2016). Initialization-free distributed algorithms for optimal resource allocation with feasibility constraints and application to economic dispatch of power systems. *Automatica, 74*, 259–269.
Yi, X. (2017). *Resource-constrained multi-agent control systems: Dynamic event-triggering, input saturation, and connectivity preservation*. Royal Institute of Technology, Sweden Licentiate thesis.
Yi, X., Yang, T., Wu, J., & Johansson, K. H. (2019a). Distributed event-triggered control for global consensus of multi-agent systems with input saturation. *Automatica, 100*, 1–9.
Yi, X., Yang, T., Wu, J., & Johansson, K. H. (2019b). Distributed online convex optimization with long term coupled constraints. In *Proceedings of the 38th Chinese control conference*.
Yi, X., Yao, L., Yang, T., George, J., & Johansson, K. H. (2018). Distributed optimization for second-order multi-agent systems with dynamic event-triggered communication. In *Proceedings of the 57th IEEE conference on decision and control (CDC)* (pp. 3397–3402).
Yu, S., & Long, X. (2015). Finite-time consensus for second-order multi-agent systems with disturbances by integral sliding mode. *Automatica, 54*, 158–165.
Yu, X., & Xue, Y. (2016). Smart grids: A cyber-physical systems perspective. *Proceedings of the IEEE, 104*(5), 1058–1070.
Yuan, D., Ho, D. W. C., & Xu, S. (2016). Regularized primal-dual subgradient method for distributed constrained optimization. *IEEE Transactions on Cybernetics, 46*(9), 2109–2118.
Yuan, D., Xu, S., & Zhao, H. (2011). Distributed primal-dual subgradient method for multiagent optimization via consensus algorithms. *IEEE Transactions on Systems, Man, and Cybernetics, Part B, 41*(6), 1715–1724.
Yuan, K., Ling, Q., & Yin, W. (2015). On the convergence of decentralized gradient descent. *SIAM Journal on Optimization, 26*(3), 1835–1854.
Yuan, Y. (2012). *Decentralised network prediction and reconstruction algorithms*. University of Cambridge Ph.d. thesis.
Yuan, Y., Stan, G. B., Shi, L., Barahona, M., & Goncalves, J. (2013). Decentralised minimum-time consensus. *Automatica, 49*(5), 1227–1235.
Yuan, Y., Zhang, H., Wu, Y., Zhu, T., & Ding, H. (2017). Bayesian learning-based model-predictive vibration control for thin-walled workpiece machining processes. *IEEE/ASME Transactions on Mechatronics, 22*(1), 509–520.
Zanella, F., Varagnolo, D., Cenedese, A., Pillonetto, G., & Schenato, L. (2011). Newton-Raphson consensus for distributed convex optimization. In *Proceedings of the 50th IEEE conference on decision and control and European control conference* (pp. 5917–5922).

Zeng, J., & Yin, W. (2017). Extrapush for convex smooth decentralized optimization over directed networks. *Journal of Computational Mathematics, 35*(4), 381–394.
Zeng, J., & Yin, W. (2018). On nonconvex decentralized gradient descent. *IEEE Transactions on Signal Processing, 66*(11), 2834–2848.
Zeng, X., Yi, P., & Hong, Y. (2017). Distributed continuous-time algorithm for constrained convex optimizations via nonsmooth analysis approach. *IEEE Transactions on Automatic Control, 62*(10), 5227–5233.
Zhang, K., Shi, W., Zhu, H., Dall'Anese, E., & Başar, T. (2018). Dynamic power distribution system management with a locally connected communication network. *IEEE Journal of Selected Topics in Signal Processing, 12*(4), 673–687.
Zhang, S., Yi, X., George, J., & Yang, T. (2019). Computational convergence analysis of distributed optimization algorithms for directed graphs. In *Proceedings of the 15th IEEE international conference on control and automation (ICCA)*.
Zhang, W., Xu, Y., Liu, W., Zang, C., & Yu, H. (2015). Distributed online optimal energy management for smart grids. *IEEE Transactions on Industrial Informatics, 11*(3), 717–727.
Zhang, Y., Deng, Z., & Hong, Y. (2017). Distributed optimal coordination for multiple heterogeneous Euler–Lagrangian systems. *Automatica, 79*, 207–213.
Zhang, Y., & Hong, Y. (2014). Distributed optimization design for second-order multi-agent systems. In *Proceedings of the 33rd Chinese control conference* (pp. 1755–1760).
Zhang, Y., & Hong, Y. (2015). Distributed optimization design for high-order multi-agent systems. In *Proceedings of the 34th Chinese control conference* (pp. 7251–7256).
Zhang, Y., Li, H., Sun, J., & He, W. (2019). Cooperative adaptive event-triggered control for multiagent systems with actuator failures. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. https://ieeexplore.ieee.org/abstract/document/8587139.
Zhang, Y., Sun, J., Liang, H., & Li, H. (2019). Event-triggered adaptive tracking control for multiagent systems with unknown disturbances. *IEEE Transactions on Cybernetics*. https://ieeexplore.ieee.org/abstract/document/8474310.
Zhang, Z., & Chow, M. Y. (2012). Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid. *IEEE Transactions on Power Systems, 27*(4), 1761–1768.
Zhang, Z., Chow, M. Y., & Chakrabortty, A. (2012). The influence of time delays on decentralized economic dispatch by using incremental cost consensus algorithm. In M. D. Ilić (Ed.), *Control and optimization methods for electric smart grids, vol. 3 of power electronics and power systems* (pp. 313–326). New York: Springer.
Zhang, Z., Ying, X., & Chow, M. Y. (2011). Decentralizing the economic dispatch problem using a two-level incremental cost consensus algorithm in a smart grid environment. In *Proceedings of the North American power symposium (NAPS)*.
Zhao, C., Duan, X., & Shi, Y. (2019). Analysis of consensus-based economic dispatch algorithm under time delays. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. https://ieeexplore.ieee.org/abstract/document/8387460.
Zhao, C., He, J., Cheng, P., & Chen, J. (2017). Consensus-based energy management in smart grid with transmission losses and directed communication. *IEEE Transactions on Smart Grid, 8*(5), 2049–2061.
Zhao, C., He, J., & Wang, Q. (2017). Resilient distributed optimization algorithm against adversary attacks. In *Proceedings of 13th IEEE international conference on control automation (ICCA)* (pp. 473–478).
Zhao, C., Topcu, U., & Low, S. (2013). Optimal load control via frequency measurement and neighborhood area communication. *IEEE Transactions on Power Systems, 28*(4), 3576–3587.
Zhao, T., Li, Z., & Ding, Z. (2019). Consensus-based distributed optimal energy management with less communication in a microgrid. *IEEE Transactions on Industrial Informatics*. https://ieeexplore.ieee.org/abstract/document/8469107.
Zhao, Y., Liu, Y., Wen, G., & Chen, G. (2017). Distributed optimization for linear multiagent systems: Edge- and node-based adaptive designs. *IEEE Transactions on Automatic Control, 62*(7), 3602–3609.
Zhao, Z., & Lin, Z. (2016). Global leader-following consensus of a group of general linear systems using bounded controls. *Automatica, 68*, 294–304.
Zheng, W., Wu, W., Zhang, B., Sun, H., & Liu, Y. (2016). A fully distributed reactive power optimization and control method for active distribution networks. *IEEE Transactions on Smart Grid, 7*(2), 1021–1033.
Zhong, M., & Cassandras, C. G. (2010). Asynchronous distributed optimization with event-driven communication. *IEEE Transactions on Automatic Control, 55*(12), 2735–2750.
Zhong, X., & He, H. (2019). GrHDP solution for optimal consensus control of multiagent discrete-time systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. https://ieeexplore.ieee.org/abstract/document/8330750.
Zhu, J., Qi, T., Ma, D., & Chen, J. (2018). *Limits of stability and stabilization of time-delay systems: A small-gain approach*. Springer International Publishing.
Zhu, M., & Martínez, S. (2010). Discrete-time dynamic average consensus. *Automatica, 46*(2), 322–329.
Zhu, M., & Martínez, S. (2012). On distributed convex optimization under inequality and equality constraints. *IEEE Transactions on Automatic Control, 57*(1), 151–164.
Zhu, M., & Martínez, S. (2013). An approximate dual subgradient algorithm for multi-agent non-convex optimization. *IEEE Transactions on Automatic Control, 58*(6), 1534–1539.
Zhu, M., & Martínez, S. (2015). *Distributed optimization-based control of multi-agent networks in complex environments*. Springer.
Zhu, Y., Yu, W., & Wen, G. (2016). Distributed consensus strategy for economic power dispatch in a smart grid with communication time delays. In *Proceedings of the 2016 IEEE international conference on industrial technology (ICIT)* (pp. 1384–1389).
Zhu, Y., Yu, W., Wen, G., Chen, G., & Ren, W. (2018). Continuous-time distributed subgradient algorithm for convex optimization with general constraints. *IEEE Transactions on Automatic Control, 64*(4), 1694–1701.

Please cite this article as: T. Yang, X. Yi and J. Wu et al., A survey of distributed optimization, Annual Reviews in Control, https://doi.org/10.1016/j.arcontrol.2019.05.006