Automatica 148 (2023) 110741

Contents lists available at ScienceDirect

Automatica

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

Zeroth-order feedback optimization for cooperative multi-agent systems ☆

Yujie Tang [a,*], Zhaolin Ren [b], Na Li [b]

[a] Department of Industrial Engineering and Management, Peking University, Beijing, 100871, China
[b] School of Engineering and Applied Sciences, Harvard University, MA 02134, USA

ARTICLE INFO

Article history:
Received 23 August 2021
Received in revised form 27 August 2022
Accepted 1 October 2022
Available online 29 November 2022

Keywords:
Multi-agent systems
Distributed optimization
Zeroth-order optimization

ABSTRACT

We study a class of cooperative multi-agent optimization problems, where each agent is associated with a local action vector and a local cost, and the goal is to cooperatively find the joint action profile that minimizes the average of the local costs. We consider the setting where gradient information is not readily available, and the agents only observe their local costs incurred by their actions as a feedback to determine their new actions. We propose a zeroth-order feedback optimization scheme and provide explicit complexity bounds for the constrained convex setting with noiseless and noisy local cost observations. We also discuss briefly on the impacts of knowledge of local function dependence between agents. The algorithm's performance is justified by a numerical example of distributed routing control.

1. Introduction

In this paper, we study model-free decentralized optimization for a specific class of cooperative multi-agent systems. Specifically, the cooperative multi-agent system comprises a group of n decision-making agents connected by a communication network. Associated with each agent is a local action $x^i \in \mathbb{R}^{d_i}$, and after the agents take their actions, a local cost $f_i(x^1,\ldots,x^n)$ will be observed by agent i which reflects the impact of all agents' actions. The goal for the agents is to cooperatively seek their *local* actions that minimize their averaged cost as the *global* objective characterizing the system-wise performance:

$$\min_{x^1,\ldots,x^n} \frac{1}{n}\sum_{i=1}^{n} f_i(x^1,\ldots,x^n).$$

We shall focus on the *model-free* setting, where each agent can only utilize the observed (zeroth-order) feedback values of the associated local cost, but not (higher-order) derivatives thereof. We refer to such optimization problems as *(cooperative) multi-agent zeroth-order feedback optimization*.

Multi-agent zeroth-order feedback optimization and its variants can cover many real-world applications, such as optimal flow of routing games (Li & Marden, 2013; Nisan, Roughgarden, Tardos, & Vazirani, 2007), mobile sensor coverage (Cortés, Martínez, Karataş, & Bullo, 2004), wind farm power optimization problem (Marden, Ruben, & Pao, 2013), power control in wireless networks (Candogan, Menache, Ozdaglar, & Parrilo, 2010), etc. In these applications, decision-makers may not have access to a sufficiently accurate model of the underlying system, which motivates the use of zeroth-order/derivative free approaches.

We emphasize that in multi-agent zeroth-order feedback optimization, each agent i can only control its own action x^i, but each local cost f_i is a function of the joint action profile $x := (x^1,\ldots,x^n)$, i.e., agent i's local cost depends possibly on the actions of all agents (or a subset of agents). Such coupling in the local cost functions adds complexities in optimizing the global objective via local information, and requires carefully designed schemes of coordination among agents. We also point out that this problem setup is different from the more commonly studied *consensus optimization* setup, in which each agent maintains a local copy of the global decision variable, and is able to evaluate its local cost at its own local copy without being directly affected by other agents (see Nedić & Ozdaglar, 2010, for a survey).

Existing literature has investigated cooperative multi-agent zeroth-order feedback optimization and its variants from a number of different angles. One line of works (Marden, Young, & Pao,

☆ This work was supported by NSF, United States CNS-2003111, NSF, United States ECCS-2038603, NSF AI Institute in Dynamic Systems, United States CBET-2112085, NSF CAREER, United States ECCS-1553407, and ONR YIP, United States N00014-19-1-2217. This research was conducted when Yujie Tang was a postdoctoral fellow in the School of Engineering and Applied Sciences at Harvard University. The material in this paper was partially presented at the 59th IEEE Conference on Decision and Control, December 14–18, 2020, Jeju Island, Republic of Korea. This paper was recommended for publication in revised form by Associate Editor Luca Schenato under the direction of Editor Christos G. Cassandras.

* Corresponding author.
E-mail addresses: yujietang@pku.edu.cn (Y. Tang), zhaolinren@g.harvard.edu (Z. Ren), nali@seas.harvard.edu (N. Li).

https://doi.org/10.1016/j.automatica.2022.110741

2014; Menon & Baras, 2013a, 2013b, 2014) has been motivated by the wind farm power maximization problem and has developed algorithms for *social welfare maximization of multi-agent games*. Specifically, Marden et al. (2014) and Menon and Baras (2013a) studied welfare maximization of multi-agent games with discrete action spaces, which can be viewed as a discrete analog of our problem setup, and Marden et al. (2013) applied these methods to model-free wind farm power maximization; Menon and Baras (2013b) proposed a modified algorithm that incorporates exchange of information between agents to eliminate the restrictions on the payoff structure in previous works. Then in Menon and Baras (2014), the authors studied welfare maximization of multi-agent games with continuous action spaces, which is essentially identical to our problem setup; they developed a continuous-time decentralized payoff-based algorithm using extremum seeking control and consensus on the local payoffs. The paper Dougherty and Guay (2017) motivated its problem setup from distributed extremum seeking control over sensor networks, but can also be regarded as an extension of Menon and Baras (2014), which further handles coupled constraints on the actions by barrier functions. We point out that, apart from implementation issues of continuous-time algorithms, these works that were based on extremum seeking control have the limitation that they only established convergence to a neighborhood of an optimal joint action for limited situations, contrary to our work that establishes explicit complexity bounds that also reflect the impact of problem dimension and network structure for general constrained smooth convex problems. In another related direction, Li and Marden (2013) considered the problem of designing local objective functions so as to optimize global behavior in multi-agent games but it assumes the knowledge of the objective function structure.

Our contributions

In this paper, we propose a Zeroth-order Feedback Optimization (ZFO) algorithm for cooperative multi-agent systems. Our ZFO algorithm is based on local computation and communication of the two-point zeroth-order gradient estimators investigated in Nesterov and Spokoiny (2017) and Shamir (2017). More specifically, for each iteration, each agent first takes its own actions and observes the corresponding zeroth-order values of its own local cost, then collects and updates zeroth-order information of other agents' costs by exchanging data with its neighbors in the network, and finally constructs a two-point zeroth-order partial gradient estimate for updating its own action vector. The communication network could be subject to potential delays.

Furthermore, we conduct complexity analysis of our ZFO algorithm for smooth convex problems for both noise-free and noisy zeroth-order evaluations. A summary of the complexity bounds can be found in Table 1. Here we list the number of iterations needed for the proposed algorithm to converge with accuracy $\epsilon > 0$, where the accuracy is measured by the expected optimality gap in the objective value (see Section 4 for detailed definitions). These complexity bounds are also compared with the centralized counterparts. In addition, apart from the dependence on ϵ, we also provide the dependence of the complexity bounds on the problem's dimension d, and on the communication network's structure and delays quantified by $\bar{\mathfrak{b}}$. To the best of our knowledge, this work seems to be the first to provide explicit complexity bounds for algorithms of multi-agent zeroth-order feedback optimization with analysis on the impact of problem dimension and network structure.

Compared to the authors' conference paper Tang, Ren, and Li (2020a) which only analyzed the unconstrained nonconvex setting with noiseless zeroth-order evaluations, this journal article contains new results for (i) the constrained convex setting where the global objective function is convex and the feasible regions are compact and convex, and (ii) the situations where zeroth-order evaluations are corrupted by additive noise. In order to deal with the compact constraints in the convex setting, we introduce and analyze a new sampling procedure for the random perturbations in zeroth-order gradient estimation, which has its own merit for the general area of zeroth-order optimization. We also conduct a preliminary investigation on how knowledge of local function dependence can be exploited to improve convergence and reduce communication burden. We provide new numerical results on finding the optimal flow of a routing game with a convex global objective.

Table 1
Complexity bounds for our ZFO algorithm.

	$\nabla f(x^*) = 0$ known	$\nabla f(x^*) = 0$ not known
Noiseless	$\Theta\left(\frac{\bar{\mathfrak{b}} d}{\epsilon^2}\right)$	
Noisy	$\Theta\left(\frac{\bar{\mathfrak{b}}(d^2 + d\ln(1/\epsilon))}{\epsilon^3}\right)$	$\Theta\left(\frac{\bar{\mathfrak{b}}(d^2 + d\ln(1/\epsilon))}{\epsilon^4}\right)$

Other related work

Zeroth-order optimization. Our work employs zeroth-order optimization techniques to deal with the lack of model information. In the centralized setting, one line of research on zeroth-order optimization has focused on constructing gradient estimators using zeroth-order function values (Duchi, Jordan, Wainwright, & Wibisono, 2015; Flaxman, Kalai, & McMahan, 2005; Larson, Menickelly, & Wild, 2019; Nesterov & Spokoiny, 2017; Shamir, 2017), and there have also been works proposing direct-search methods that do not seek to approximate a gradient (Agarwal, Foster, Hsu, Kakade, & Rakhlin, 2013; Torczon, 1997). A survey can be found in Larson et al. (2019). In addition, there has been increasing interest recently in exploiting zeroth-order optimization methods in a distributed setting (Hajinezhad, Hong, & Garcia, 2019; Li, Tang, Zhang, & Li, 2021; Sahu, Jakovetic, Bajovic, & Kar, 2018; Tang, Zhang, & Li, 2021; Yu, Ho, & Yuan, 2022). However, to the best of our knowledge, most of them focus on the consensus optimization setup, rather than the cooperative multi-agent system setup discussed in this work.

Distributed optimization. Another related research area is distributed optimization. While our setting is different from consensus optimization (Chang, Hong, & Wang, 2015; Nedić & Ozdaglar, 2009; Pu & Nedić, 2021; Qu & Li, 2018; Shi, Ling, Wu, & Yin, 2015), we note that in both settings, collaborations among agents are needed for optimizing the global objective. In addition, in our problem setup, the agents will naturally experience delays when receiving information from other (possibly distant) agents in the network due to the local nature of communication. We shall see later that our algorithm and analysis share similarities with asynchronous/delayed distributed optimization (Agarwal & Duchi, 2011; Lian, Huang, Li, & Liu, 2015; Lian, Zhang, Hsieh, Huang, & Liu, 2016; Liu & Wright, 2015; Nedić, 2011; Zhang & Kwok, 2014). However, our work appears to be the first that studies the effects of delays in a decentralized zeroth-order setting.

Notation

We use $\|\cdot\|$ to denote the standard ℓ_2 norm, and use $\langle\cdot,\cdot\rangle$ to denote the standard inner product. For any differentiable function $h(x) = h(x^1, \ldots, x^n)$, we use $\nabla^i h(x)$ to denote the partial gradient of h with respect to x^i. For $x \in \mathbb{R}$, we let $[x]_+ := \max\{0, x\}$. For a finite set A, we use $|A|$ to denote its number of elements. For any $\mathcal{S} \subseteq \mathbb{R}^p$, we use $\operatorname{int}\mathcal{S}$ to denote its interior, use $\mathcal{S}+x$ to denote $\{s+x : s \in \mathcal{S}\}$ for any $x \in \mathbb{R}^p$, and use $u\mathcal{S}$ to denote $\{us : s \in \mathcal{S}\}$ for any $u \in \mathbb{R}$. The projection of x onto a closed convex set C will be denoted by $\mathcal{P}_C[x]$. The closed unit ball in $\mathbb{R}^p$ will be denoted by $\mathbb{B}_p$. The $p \times p$ identity matrix will be denoted by I_p. $\mathcal{N}(\mu, \Sigma)$ denotes the Gaussian distribution with mean μ and covariance Σ.

2. Problem formulation

Consider a group of n agents, where agent i is associated with an action vector $x^i \in \mathcal{X}_i \subseteq \mathbb{R}^{d_i}$ for each $i = 1, \ldots, n$. Each set $\mathcal{X}_i$ is convex and compact. The joint action profile of the group of agents is then $x := (x^1, x^2, \ldots, x^n) \in \mathcal{X}$, where $\mathcal{X} := \prod_{i=1}^n \mathcal{X}_i \subseteq \mathbb{R}^d$ and $d := \sum_{i=1}^n d_i$. Upon taking action jointly, each agent i receives a corresponding local cost $f_i(x) = f_i(x^1, \ldots, x^n)$ that depends on the joint action profile x, i.e., the actions of all agents. The goal of the agents is to cooperatively find the actions that minimize the average of all the local costs, i.e.,

$$\min_{x \in \mathcal{X}} \ f(x) := \frac{1}{n}\sum_{i=1}^n f_i(x^1, \ldots, x^n), \tag{1}$$

where $f(x)$ denotes the global objective function defined as the average cost among agents.

Since the local costs are affected by all agents' actions, when solving the problem (1), each agent will not only need to collect information on its own local cost, but also need to communicate and collaborate with other agents by exchanging necessary information. We further impose two assumptions for our problem setup; the first pertains to the type of information the agents can access, and the second to communication mechanism among agents:

1. **Access to only zeroth-order information.** Each agent i can only access zeroth-order function value of its local cost f_i, and derivatives of f_i of any order are not available. Moreover, the function values can only be obtained through observation of feedback cost after actions have been taken. Precisely, each agent i first determines its action vector x^i and takes the action, yielding a new joint action profile $x = (x^1, \ldots, x^n)$, and then observes its corresponding local cost f_i evaluated at $x = (x^1, \ldots, x^n)$. We shall also assume that the constraint $x \in \mathcal{X}$ is a hard constraint in the sense that each f_i is defined only on $\mathcal{X}$ and the agents can only explore their local costs within the set $\mathcal{X}$.
   In this paper, we consider two cases regarding the observation of local cost values:
   (i) Noiseless case: Each agent can observe its local cost accurately without being corrupted by noise.
   (ii) Noisy case: Each agent's observed local cost value is corrupted by additive random noise with zero mean and variance bounded by σ^2. We assume the noises are independent of each other and are also independent of x.
2. **Localized communication.** We let the n agents be connected by a communication network. The topology of the network is represented by an undirected, connected graph $\mathcal{G} = (\{1, \ldots, n\}, \mathcal{E})$, where the edges in $\mathcal{E}$ correspond to the bidirectional communication links. Each agent is only allowed to exchange messages directly with its neighbors in the network $\mathcal{G}$. We shall denote the distance (the length of the shortest path) between the pair of nodes (i, j) in the graph $\mathcal{G}$ by b_{ij}.

We adopt the following technical assumptions throughout the paper:

Assumption 1. We assume that for each $i = 1, \ldots, n$, the compact and convex set $\mathcal{X}_i$ has a nonempty interior. Without loss of generality we also assume $0 \in \operatorname{int}\mathcal{X}_i$, as we can always translate $\mathcal{X}_i$.

We define

$$\underline{r}_i := \sup\{r > 0 : r\mathbb{B}_{d_i} \subseteq \mathcal{X}_i\}, \qquad \underline{r} := \min_{1 \le i \le n} \underline{r}_i,$$

$$\bar{R}_i := \inf\{R > 0 : \mathcal{X}_i \subseteq R\mathbb{B}_{d_i}\}, \qquad \bar{R} := \sqrt{\sum_{i=1}^n \bar{R}_i^2}.$$

Assumption 1 guarantees that these quantities are well-defined positive real numbers.

Assumption 2. Each local cost function f_i is G-Lipschitz and L-smooth on $\mathcal{X}$, i.e.,

$$|f_i(x) - f_i(y)| \le G\|x - y\|, \ \ \|\nabla f_i(x) - \nabla f_i(y)\| \le L\|x - y\|$$

for any $x, y \in \mathcal{X}$ for each $i = 1, \ldots, n$. Furthermore, the global cost function f is convex.

In the following subsection, we present one example which fits the aforementioned formulation.

2.1. Example: Optimal flow of routing games

Consider finding the optimal flow of a nonatomic routing game formulated in Li and Marden (2013) and Nisan et al. (2007). We have n agents each seeking to send an amount of traffic $Q_i > 0$ through a network. Each agent i is able to use a set of paths $\mathcal{P}_i$ in the network, and here we allow the edges in the network to be shared by different paths for different agents. Each agent i is associated with an action vector $x^i \in \mathbb{R}^{|\mathcal{P}_i|}$ where x^i_p represents the proportion of traffic in Q_i allocated to the path $p \in \mathcal{P}_i$. The joint action profile is $x = (x^1, \ldots, x^n)$. Each edge e of the network has a cost function $c_e : [0, \infty) \to \mathbb{R}$ characterizing the congestion incurred by the total traffic through e, and the cost of a path $p \in \mathcal{P}_i$, denoted by c_p, is the sum of the costs of the constituent edges:

$$c_p(x) = \sum_{e \in p} c_e(q_e(x)), \qquad q_e(x) = \sum_{j=1}^n \sum_{p' \in \mathcal{P}_j : e \in p'} x^j_{p'} Q_j.$$

We assume that the function $t \mapsto t \cdot c_e(t)$ is smooth and convex for each edge e. The goal is to find the optimal joint action that minimizes the global cost defined by

$$f(x) = \frac{1}{n}\sum_{i=1}^n f_i(x), \quad f_i(x) = \sum_{p \in \mathcal{P}_i} x^i_p Q_i \cdot c_p(x),$$

subject to the constraints $x^i \in \mathcal{X}_i$ where $\mathcal{X}_i = \{x^i \in \mathbb{R}^{|\mathcal{P}_i|} : x^i_p \ge 0, \sum_{p \in \mathcal{P}_i} x^i_p = 1\}$. Observe that each local cost f_i is affected by other agents' actions $(x^j)_{j \ne i}$.

We assume the following mechanism of collecting and exchanging information among agents:

1. **Access to only zeroth-order information.** Each agent i does not know the specific form of the cost function c_p, and can only observe the local cost $f_i(x)$ corresponding to the currently implemented action x.
2. **Localized communication.** The agents are connected by a bidirectional communication network, and each agent can only directly talk to its neighbors.

Note that the set $\mathcal{X}_i$ has an empty interior and does not satisfy Assumption 1. One way to handle this issue is to arbitrarily select one route $\check{p}_i \in \mathcal{P}_i$, remove $x^i_{\check{p}_i}$ from the action vector x^i, and replace the constraints by

$$x^i_p \geq 0 \ \ \forall p \in \mathcal{P}_i \backslash \{\check{p}_i\}, \qquad \sum_{p \in \mathcal{P}_i \backslash \{\check{p}_i\}} x^i_p \leq 1.$$

After a further translation of the variables to include the origin in the interior, Assumption 1 will be satisfied. The variable $x^i_{\check{p}_i}$ can be recovered by $x^i_{\check{p}_i} = 1 - \sum_{p \in \mathcal{P}_i \backslash \{\check{p}_i\}} x^i_p$.

3. Algorithm

3.1. Zeroth-order gradient estimation

We first give a preliminary introduction to the zeroth-order optimization method adopted in this paper. Consider the following zeroth-order gradient estimator (Nesterov & Spokoiny, 2017):

$$\mathsf{G}_f(x; u, z) = \frac{f(x + uz) - f(x - uz)}{2u} z, \tag{2}$$

where $u > 0$ is called the *smoothing radius*, and z is a perturbation sampled from an isotropic distribution on $\mathbb{R}^d$ with finite second moment. Nesterov and Spokoiny (2017) shows that, if we let $z \sim \mathcal{N}(0, I_d)$, then $\mathbb{E}_z\left[\mathsf{G}_f(x; u, z)\right] = \nabla f^u(x)$ where $f^u(x) = \mathbb{E}_{y \sim \mathcal{N}(0, I_d)}[f(x + uy)]$, and one can also control the differences $|f^u(x) - f(x)|$ and $\|\nabla f^u(x) - \nabla f(x)\|$ by controlling u when f is Lipschitz continuous and smooth. In other words, $\mathsf{G}_f(x; u, z)$ can be viewed as a stochastic gradient with a nonzero bias controlled by the smoothing radius u.

3.2. Algorithm design

Our algorithm will be based on the zeroth-order gradient estimator (2) and the stochastic mirror descent algorithm

$$G(t) = \frac{1}{n} \sum_{j=1}^{n} \frac{\hat{f}_j^+(t) - \hat{f}_j^-(t)}{2u} z(t), \quad z(t) \sim \mathcal{N}(0, I_d),$$

$$x(t+1) = \arg\min_{x \in \mathcal{X}} \left\{ \langle G(t), x - x(t) \rangle + \frac{1}{\eta} \mathscr{D}_\psi(x|x(t)) \right\}.$$

Here, $\hat{f}_j^\pm(t) := f_j(x(t) \pm uz(t)) + \varepsilon_j^\pm(t)$ represent the observed local cost values after the agents take the actions $x(t) \pm uz(t)$, where $\varepsilon_j^+(t)$ and $\varepsilon_j^-(t)$ are the independent additive random noises with variance bounded above by σ^2 (setting $\sigma^2 = 0$ reduces to the noiseless case); $\mathscr{D}_\psi(x|y) := \psi(x) - \psi(y) - \langle \nabla \psi(y), x - y \rangle$ is the Bregman divergence associated with the function ψ that is convex and continuously differentiable (Beck & Teboulle, 2003).

In our multi-agent setting, we let ψ be given by $\psi(x^1, \ldots, x^n) = \sum_{i=1}^n \psi_i(x^i)$, and require each ψ_i to be 1-strongly convex. Then since $\mathcal{X} = \prod_i \mathcal{X}_i$, we observe that the mirror descent iteration can be decoupled among agents as follows:

$$x^i(t+1) = \arg\min_{x^i \in \mathcal{X}_i} \left\{ \langle G^i(t), x^i - x^i(t) \rangle + \frac{1}{\eta} \mathscr{D}_{\psi_i}(x^i|x^i(t)) \right\}, \tag{3}$$

where $G^i(t)$ is a zeroth-order estimate of the partial gradient $\nabla^i f(x(t))$ given by

$$G^i(t) = \frac{1}{n} \sum_{j=1}^{n} \frac{\hat{f}_j^+(t) - \hat{f}_j^-(t)}{2u} z^i(t), \quad z^i(t) \sim \mathcal{N}(0, I_{d_i}). \tag{4}$$

We can see that employing $\mathcal{N}(0, I_d)$ as the distribution of the perturbation z allows the agents to generate their associated subvectors z^i independently of each other without resorting to coordination strategies. However, we also notice the following two issues when adopting (3) and (4) in our setting:

1. In our setting, the agents can only take actions within $\mathcal{X}$. However, samples from $\mathcal{N}(0, I_d)$ are unbounded, and $x(t) + uz(t)$ or $x(t) - uz(t)$ may not lie in $\mathcal{X}$.
2. The computation of (4) requires agent i to collect $f_j^+(t) - f_j^-(t)$ for all j. While each agent can observe its own local cost, other agents' local cost information has to be transmitted via the communication network, which will result in delays.

We now discuss how to handle these two issues.

3.2.1. Sampling within the constraint set

Our idea of handling the first issue is to slightly modify the distribution of the perturbation z so that (i) $x(t) \pm uz(t)$ always lies in $\mathcal{X}$, (ii) each agent can still generate their associated z^i independently, and (iii) the resulting zeroth-order gradient estimator has comparable bias and variance with the original estimator (2).

For any $x^i \in \operatorname{int} \mathcal{X}_i$ and $u > 0$, we introduce the set

$$\mathcal{S}_i(x^i, u) := \frac{1}{u}\left(\mathcal{X}_i - x^i\right) \cap \left(-\frac{1}{u}\left(\mathcal{X}_i - x^i\right)\right);$$

see Fig. 1 for an illustrative description. Obviously, $\mathcal{S}_i(x^i, u)$ is a closed convex set with a nonempty interior, and $z^i \in \mathcal{S}_i(x^i, u)$ if and only if $-z^i \in \mathcal{S}_i(x^i, u)$. Moreover, we have $x^i + uz^i \in \mathcal{X}_i$ and $x^i - uz^i \in \mathcal{X}_i$ for any $z \in \mathcal{S}_i(x^i, u)$. Therefore we propose to generate z^i for each agent i by

$$z^i = \mathcal{P}_{\mathcal{S}_i(x^i, u)}\left[\tilde{z}^i\right], \qquad \tilde{z}^i \sim \mathcal{N}(0, I_{d_i}). \tag{5}$$

The resulting probability distribution of z^i and z will be denoted by $\mathcal{Z}_i(x^i, u)$ and $\mathcal{Z}(x, u)$ respectively.

The modified zeroth-order partial gradient estimator for agent i is then given by

$$G^i(t) = \frac{1}{n} \sum_{j=1}^{n} \frac{\hat{f}_j^+(t) - \hat{f}_j^-(t)}{2u} z^i(t), \quad z^i(t) \sim \mathcal{Z}_i(x^i(t), u). \tag{6}$$

In order for $G^i(t)$ to have comparable statistics with the original estimator (2), we require $\mathcal{S}_i(x^i(t), u)$ to contain a ball with a sufficiently large radius, so that the projection $\mathcal{P}_{\mathcal{S}_i(x^i, u)}$ in (5) happens rarely, i.e., the probability of $\tilde{z}^i \in \mathcal{S}_i(x^i, u)$ is close to 1. This further leads to the requirement that there should be sufficient distance from $x^i(t)$ to the boundary of $\mathcal{X}_i$. In order for $x^i(t)$ to satisfy this requirement, we modify the mirror descent step as in Agarwal, Dekel, and Xiao (2010) and Flaxman et al. (2005):

$$x^i(t+1) = \arg\min_{x^i \in (1-\delta)\mathcal{X}_i} \left\{ \langle G^i(t), x^i - x^i(t) \rangle + \frac{1}{\eta} \mathscr{D}_{\psi_i}(x^i|x^i(t)) \right\}, \tag{7}$$

where $\delta > 0$ is an algorithmic parameter to be determined later. In other words, we shrink the feasible set in the mirror descent step to be $(1 - \delta)\mathcal{X}$, so that a band along the boundary of $\mathcal{X}$ will be available for the sampling of $z(t)$. Indeed, Flaxman et al. (2005, Observation 3.2) shows that, for $\delta \in (0, 1)$ and any $x^i \in (1 - \delta)\mathcal{X}_i$, we have $x^i + \delta \underline{r}_i \mathbb{B}_{d_i} \subseteq \mathcal{X}_i$ (recall the definition of $\underline{r}_i$ after Assumption 1), i.e., the distance from x^i to the boundary of $\mathcal{X}_i$ is at least $\delta \underline{r}_i$. Moreover, $x^i + \delta \underline{r}_i \mathbb{B}_{d_i} \subseteq \mathcal{X}_i$ also implies $\frac{\delta \underline{r}_i}{u} \mathbb{B}_{d_i} \subseteq \mathcal{S}_i(x^i, u)$. Thus, if let $\delta \underline{r}_i / u$ be sufficiently large (say $\gtrsim 3\sqrt{d_i}$), then the set $\mathcal{S}_i(x^i, u)$ will correspondingly contain a sufficiently large ball, meaning that projection happens rarely when we sample z^i by (5).

Remark 1. Depending on the specific $\mathcal{X}_i$, the projection $\mathcal{P}_{\mathcal{S}_i(x^i, u)}$ either admits a closed-form expression that can be computed efficiently, or needs to be computed via some iterative method. One option for iterative methods is Dykstra's projection algorithm (Boyle & Dykstra, 1986) that finds the projection onto the intersection of finitely many convex sets; one may also consider

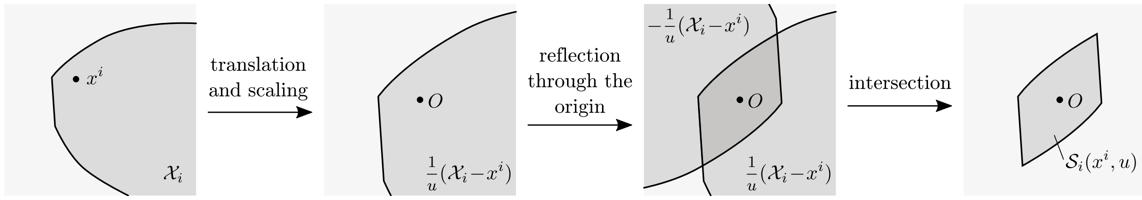

Fig. 1. Construction of the set $\mathcal{S}_i(x^i, u)$.

applying a fast optimization algorithm (such as ADMM) to the optimization problem

$$\min_{z^i \in \mathbb{R}^{d_i}} \|z^i - \tilde{z}^i\|^2 \quad \text{s.t.} \ \ x^i \pm u z^i \in \mathcal{X}_i.$$

We also point out that, for sufficiently large $\delta \underline{r}_i / u$, the probability that $\mathcal{P}_{\mathcal{S}(x^i,u)}[\tilde{z}^i] \neq \tilde{z}^i$ will be very small, meaning that one rarely needs to explicitly compute the projection in (5) for generating z^i. Therefore, we expect that the projection step in (5) will not have a major impact on the computational efficiency of our algorithm.

3.2.2. Collecting data from other agents

To ensure that each agent can obtain $\hat{f}_j^+(t) - \hat{f}_j^-(t)$ of all other agents as soon as possible, we develop a procedure for generating, distributing and utilizing the most up-to-date information among agents via the network. This procedure consists of the following parts:

1. **Generating new data:** At time step t, each agent i generates $z^i(t) \sim \mathcal{Z}_i(x^i(t), u)$, adjusts its actions to be $x^i(t) \pm u z^i(t)$ and observes the corresponding local costs $\hat{f}_i^{\pm}(t) = f_i(x(t) \pm u z(t)) + \varepsilon_i^{\pm}(t)$. Agent i then computes

   $$D_i^i(t) := \frac{\hat{f}_i^+(t) - \hat{f}_i^-(t)}{2u},$$

   and also records the timestamp $\tau_i^i(t) = t$ at which the data $D_i^i(t)$ is generated. This pair of newly-generated data $(D_i^i(t), \tau_i^i(t))$ is going to be distributed via the communication network among agents.
2. **Distributing and updating other agents' information:** Each agent i maintains a $2 \times n$ array that records the most up-to-date information on the difference quotients of all f_j at each time step t:

difference quotient	$D_1^i(t)$	$D_2^i(t)$	$\cdots$	$D_n^i(t)$
time instant	$\tau_1^i(t)$	$\tau_2^i(t)$	$\cdots$	$\tau_n^i(t)$

   (8)

   Here the quantity $D_j^i(t)$ records agent i's most up-to-date value of the difference quotient $\frac{\hat{f}_j^+(\tau) - \hat{f}_j^-(\tau)}{2u}$, and the quantity $\tau_j^i(t)$ records the time step at which $D_j^i(t)$ was generated by agent j. In other words,

   $$D_j^i(t) = D_j^j(\tau_j^i(t)) = \frac{\hat{f}_j^+(\tau_j^i(t)) - \hat{f}_j^-(\tau_j^i(t))}{2u}.$$

   Note that the entries $D_i^i(t)$ and $\tau_i^i(t)$ in the array (8) will be updated by agent i itself following the previous part. In order to update other entries in (8) at time t, each agent i first collects data that has been sent by its neighbors in the previous time step, to get their versions of the array (8). We use $(D_j^{k\to i}(t), \tau_j^{k\to i}(t))$ to denote the entries of the array on the difference quotient of f_j that agent i has received from its neighbor k at time t. In the situation when agent i does not receive the array from agent k at time t, we let $(D_j^{k\to i}(t), \tau_j^{k\to i}(t)) = (D_j^i(t-1), \tau_j^i(t-1))$. Then for each $j \neq i$, agent i compares all collected $\tau_j^{k\to i}(t)$ and finds the neighbor $k_j^i(t)$ that has sent the largest $\tau_j^{k\to i}(t)$, i.e.,

   $$k_j^i(t) = \operatorname*{arg\,max}_{k:(k,i)\in\mathcal{E}} \ \tau_j^{k\to i}(t).$$

   In other words, the difference quotient of f_j sent by the neighbor $k_j^i(t)$ is the most up-to-date among all of agent i's neighbors. We then update $(D_j^i(t), \tau_j^i(t))$ to be equal to the data sent by the neighbor $k_j^i(t)$.
   Finally, after agent i finishes updating the array (8), it sends this array to all of its neighbors.
   Each agent initializes the array (8) by setting $D_j^i(-1) = 0$ and $\tau_j^i(-1) = -1$.
3. **Constructing partial gradient estimator with delayed information:** Each agent i calculates the partial gradient estimator (6) but with delayed information. Specifically,

   $$G^i(t) = \frac{1}{n}\sum_{j=1}^{n} D_j^i(t)\, z^i(\tau_j^i(t)), \tag{9}$$

   where $z^i(t) \sim \mathcal{Z}_i(x^i(t), u)$ for each t, and the past perturbation $z^i(\tau_j^i(t))$ is used to pair with the delayed information $D_j^i(t)$ for $j \neq i$. The mirror descent step (7) is then applied to obtain $x^i(t+1)$.

We further elaborate on this procedure and the communication delays therein: Assuming that each round of communication takes one time step and there are no additional delays for all t, we see that agent i's received data $\left(D_j^{k\to i}(t), \tau_j^{k\to i}(t)\right)$ will be just $(D_j^k(t-1), \tau_j^k(t-1))$. As a result, it takes exactly b_{ij} communication rounds to transmit data from agent j to agent i (recall that b_{ij} is the distance between i and j in $\mathcal{G}$), and consequently $\tau_j^i(t) = t - b_{ij}$ and $D_j^i(t) = D_j^j(t - b_{ij})$ for $t \geq b_{ij}$. On the other hand, if some additional delay occurs during communication, then agent i may fail to receive new data from some neighbor k at some time step t, and in this case $\tau_j^i(t)$ may be smaller than $t - b_{ij}$. In Section 4, we shall see that as long as the additional delays during communication are bounded, our algorithm will still work with performance guarantees.

3.3. Our proposed algorithm

We are now ready to present our multi-agent Zeroth-order Feedback Optimization (ZFO) algorithm, which is given by Algorithm 1. In summary, each iteration of Algorithm 1 consists of the following steps:

1. Each agent i generates the associated random perturbation $z^i(t) \sim \mathcal{Z}_i(x^i(t), u)$ by (5) (Line 3).
2. Each agent takes the two perturbed actions $x^i(t) \pm u z^i(t)$ successively and observes the corresponding local cost values (Lines 4–5). Note that we require the agents to take each of the two perturbed actions synchronously.
3. Based on the new cost values, each agent i computes the difference quotient $D_i^i(t)$ of its own cost function and records the current time instant $\tau_i^i(t) = t$ (Line 6).

4. Based on the information received from the neighbors, each agent updates other columns of its array (8) by the procedure described in Section 3.2.2 (Line 7).
5. Finally, each agent sends the updated array (8) to its neighbors (Line 8) and performs stochastic mirror descent (Line 9).

Algorithm 1: Zeroth-order Feedback Optimization (ZFO) for cooperative multi-agent systems

Require: step size $\eta > 0$, smoothing radius $u > 0$, number of iterations T, initial action profile $(x_0^1, \ldots, x_0^n)$

1 **Initialize:** $x^i(0) = x_0^i$, $D_j^i(-1) = 0$, $\tau_j^i(-1) = -1$ for all $i, j = 1, \ldots, n$.
2 **for** $t = 0, \ldots, T-1$ **do**
3 Each agent i generates $z^i(t) \sim \mathcal{Z}_i(x^i(t), u)$ by (5).
4 Each agent i takes action $x^i(t) + uz^i(t)$ and observes its local cost $\hat{f}_i^+(t)$.
5 Each agent i takes action $x^i(t) - uz^i(t)$ and observes its local cost $\hat{f}_i^-(t)$.
6 Agent i computes and records
$$D_i^i(t) = \frac{\hat{f}_i^+(t) - \hat{f}_i^-(t)}{2u}, \quad \tau_i^i(t) = t.$$
7 Agent i receives data $\left(D_j^{k\to i}(t), \tau_j^{k\to i}(t)\right)_{j=1}^n$ from each neighbor $k : (k, i) \in \mathcal{E}$, and updates
$$k_j^i(t) = \mathop{\arg\max}_{k:(k,i)\in\mathcal{E}} \tau_j^{k\to i}(t), \quad \tau_j^i(t) = \tau_j^{k_j^i(t)\to i}(t),$$
$$D_j^i(t) = D_j^{k_j^i(t)\to i}(t), \qquad \forall j \neq i.$$
8 Agent i sends $\left(D_j^i(t), \tau_j^i(t)\right)_{j=1}^n$ to its neighbors.
9 Agent i updates
$$G^i(t) = \frac{1}{n}\sum\nolimits_{j=1}^n D_j^i(t) z^i(\tau_j^i(t)), \tag{9}$$
$$x^i(t+1) = \mathop{\mathrm{argmin}}_{x^i\in(1-\delta)\mathcal{X}_i} \left\{\langle G^i(t), x^i - x^i(t)\rangle + \frac{1}{\eta}\mathscr{D}_{\psi_i}(x^i | x^i(t))\right\}.$$
10 **end**

4. Complexity results

In this section, we present our main results on the complexity of Algorithm 1. We first make the following assumption on the delays occurred during the optimization procedure:

Assumption 3. There exists $\Delta \geq 0$ such that the delays are bounded above by $t - \tau_j^i(t) \leq b_{ij} + \Delta$ for every $t \geq 0$ and $i, j = 1, \ldots, n$.

We define
$$\bar{b} := \left(\frac{\sum_{i,j=1}^n (b_{ij} + \Delta)^2}{n^2}\right)^{1/2}, \tag{10}$$
$$\bar{\mathfrak{b}} := \left(\frac{\sum_{i,j=1}^n (b_{ij} + \Delta)^2 (d_i + d_j)}{\sum_{i,j=1}^n (d_i + d_j)}\right)^{1/2}. \tag{11}$$
Roughly speaking, these two quantities are (weighted) averages of pairwise distances of nodes plus the additional delay bound in the network, and characterize the connectivity of the network: Smaller $\bar{b}$ or $\bar{\mathfrak{b}}$ indicates that the nodes are more closely connected and information can be transferred over the network with fewer hops. We also define $B := \max_{i,j} b_{ij} + \Delta$.

We characterize the complexity of our algorithm by the number of iterations T needed to achieve $\mathbb{E}[f(\bar{x}(T))] - f(x^*) \leq \epsilon$ for sufficiently small ϵ, where x^* is a minimizer of $f(x)$ over $x \in \mathcal{X}$, and
$$\bar{x}(T) := \frac{1}{T - B + 1}\sum_{t=B}^{T} x(t).$$
Here we require the total number of iterations T to be greater than or equal to B to ensure that each agent i has updated the entries on agent j in the array (8) at least once.

The following theorems characterize the complexity results of Algorithm 1 for our constrained convex setting, whose proofs will be postponed to Section 5. Recall that σ^2 is the variance of the additive noise on the agents' observed local cost values. We also denote $\overline{\mathscr{D}} := \max_{x\in\mathcal{X}} \mathscr{D}_\psi(x^*|x)$.

Theorem 1 (*Convex, Noiseless*). *Suppose $\sigma = 0$. Let $\epsilon \in (0, \max_{x\in\mathcal{X}} f(x) - f(x^*)]$ be arbitrary. Then by choosing the parameters of Algorithm 1 to satisfy*
$$\delta \leq \frac{\epsilon}{5G\bar{R}}, \quad u \cdot \sqrt{d + \frac{4}{9}\left[\ln \frac{20G\bar{R}^2\sqrt{n}}{u\epsilon}\right]_+} \leq \frac{\delta r}{3},$$
$$\eta \leq \frac{\epsilon/18}{\left[G^2 + \left(\frac{L\bar{R}}{3}\right)^2\right]\left(\bar{\mathfrak{b}} + \frac{1}{2}\right)\left(\sqrt{d} + \frac{1}{5}\right)^2}, \quad T - B + 1 \geq \left\lceil \frac{15\overline{\mathscr{D}}}{2\eta\epsilon} \right\rceil,$$
we can guarantee that $\mathbb{E}[f(\bar{x}(T))] - f(x^*) \leq \epsilon$. Moreover, if all the conditions on the parameters are satisfied with equality, then
$$T = \Theta\left(\frac{\bar{\mathfrak{b}} d}{\epsilon^2}\right).$$

Theorem 2 (*Convex, Noisy*). *Suppose $\sigma > 0$, and let $\epsilon > 0$ be sufficiently small.*

1. *By choosing the parameters of Algorithm 1 to satisfy $\delta \leq \epsilon/(5G\bar{R})$ and*
$$\begin{aligned} &u \cdot \sqrt{d + \frac{4}{9}\left[\ln \frac{20G\bar{R}^2\sqrt{n}}{u\epsilon}\right]_+} \leq \frac{\delta r}{3}, \\ &\eta \leq \frac{u^2\epsilon}{2\sigma^2\left(\bar{\mathfrak{b}} + \frac{1}{2}\right)\left(\sqrt{d} + \frac{1}{5}\right)^2}, \quad T - B + 1 \geq \left\lceil \frac{15\overline{\mathscr{D}}}{2\eta\epsilon} \right\rceil, \end{aligned} \tag{12}$$
we can guarantee that $\mathbb{E}[f(\bar{x}(T))] - f(x^*) \leq \epsilon$. Moreover, if all the conditions on the parameters are satisfied with equality, then
$$T = \Theta\left(\frac{\bar{\mathfrak{b}}(d^2 + d\ln(1/\epsilon))}{\epsilon^4}\right).$$
2. *Suppose it is known that $\nabla f(x^*) = 0$. By choosing the parameters of Algorithm 1 to satisfy $\delta \leq \sqrt{\epsilon}/(\bar{R}\sqrt{2L})$ and the conditions in (12), we can guarantee that $\mathbb{E}[f(\bar{x}(T))] - f(x^*) \leq \epsilon$. Moreover, if all the conditions on the parameters are satisfied with equality, then*
$$T = \Theta\left(\frac{\bar{\mathfrak{b}}(d^2 + d\ln(1/\epsilon))}{\epsilon^3}\right).$$

We now provide some discussion on the two theorems:

1. **Existence of u.** Observe that the map
$$u \mapsto u \cdot \sqrt{d + \frac{4}{9}\left[\ln \frac{20G\bar{R}^2\sqrt{n}}{u\epsilon}\right]_+}$$

is continuous over $u \in (0, +\infty)$, goes to 0 as $u \to 0^+$ and diverges to $+\infty$ as $u \to +\infty$. Therefore given $\epsilon > 0$ and $\delta > 0$, there always exists some $u \in (0, +\infty)$ that satisfies the conditions in Theorems 1 and 2, and the condition can be achieved with equality.

2. **Complexity bound for the noiseless case.** It can be seen that in the convex noiseless case, the number of iterations needed for Algorithm 1 to achieve $\mathbb{E}[f(\bar{x}(T))] - f(x^*)$ is on the order of $O(\bar{\mathfrak{b}}d/\epsilon^2)$. The d/ϵ^2 part is in accordance with the centralized zeroth-order method (Nesterov & Spokoiny, 2017).
Equivalently, the convergence rate of Algorithm 1 can be represented as $\mathbb{E}[f(\bar{x}(T))] - f(x^*) \leq O\left(\sqrt{\bar{\mathfrak{b}}d/T}\right)$.
3. **Complexity bound for the noisy case.** For the convex noisy case, depending on whether or not we know an optimizer x^* lies in the interior of the feasible set $\mathcal{X}$, the complexity bound of Algorithm 1 can be different. Specifically, if we know $\nabla f(x^*) = 0$ (for which a sufficient condition is $x^* \in \operatorname{int}\mathcal{X}$), then the complexity has $O(\epsilon^{-3}\ln(1/\epsilon))$ dependence on ϵ and $O(d^2)$ dependence on the problem dimension d; they are in accordance with the centralized case in Bach and Perchet (2016) except for a logarithmic dependence on $1/\epsilon$. On the other hand, if we do not have $\nabla f(x^*) = 0$, the complexity bound becomes worse in terms of the dependence on ϵ. Here we provide a qualitative explanation of this difference: If one knows $\nabla f(x^*) = 0$, then by the smoothness of the objective function, we have $f(x)-f(x^*) \sim O(\|x-x^*\|^2)$, implying that the suboptimality caused by shrinking into a smaller set $(1-\delta)\mathcal{X}$ is on the order of $O((1-\delta)^2)$. Therefore, one can shrink the feasible set more aggressively, allowing a larger smoothing radius u that does not amplify the noise much, and consequently the number of iterations can be reduced. On the other hand, if we do not have $\nabla f(x^*) = 0$, then only $f(x) - f(x^*) \sim O(\|x-x^*\|)$ can be guaranteed by the Lipschitz continuity of the objective function, and the suboptimality caused by shrinkage is on the order of $O(1-\delta)$. Therefore the set $(1-\delta)\mathcal{X}$ needs to be sufficiently large to make sure that the suboptimality caused by shrinkage is small, resulting in more restricted size of the smoothing radius. Consequently, the additive noise in the gradient estimator can be more severely amplified, and one needs more iterations to average out the noise.
4. **Dependence on the network connectivity.** We can see that the complexity bounds of both the noiseless and noisy cases has an addition factor $\bar{\mathfrak{b}}$. This term upper bounds the influence of the connectivity of the communication network. We shall see in Section 7 that for the numerical test cases we have run, our algorithm achieves better empirical behavior than the bounds would suggest in terms of the dependence on the network connectivity. It would be interesting to investigate whether and how we can improve the theoretical analysis in the future.

5. Proofs of complexity results

Note that the iterations of Algorithm 1 can be written as

$$x(t+1) = \operatorname*{arg\,min}_{x \in (1-\delta)\mathcal{X}} \left\{ \langle G(t), x - x(t) \rangle + \frac{1}{\eta}\mathscr{D}_\psi(x|x(t)) \right\}, \tag{13}$$

where $G(t)$ is the d-dimensional vector that concatenates $G^1(t), \ldots, G^n(t)$, and $\psi(x) = \sum_{i=1}^n \psi_i(x^i)$. Recall that each $\psi_i(x)$ is 1-strongly convex, so that $\mathscr{D}_{\psi_i}(x|y) \geq \frac{1}{2}\|x-y\|^2$ for all $x, y \in \mathcal{X}_i$. For notational simplicity, we let $D_j(t)$ denote $D_j^i(t)$ for $t \geq 0$, and let each $D_j(t) = 0$ and $z(t) = 0$ for $t < 0$. We let $\mathcal{F}_t$ denote the σ-algebra generated by $x(\tau)$ for $\tau \leq t$ and all $\tau_j^i(s)$ for $1 \leq i, j \leq n$ and $0 \leq s \leq T$.

5.1. Auxiliary results on the gradient estimator

We first establish a lemma for bounding the bias of the gradient estimator (2) with $z \sim \mathcal{Z}(x, u)$.

Lemma 1. *Let $h : \mathcal{X} \to \mathbb{R}$ be a G-Lipschitz continuous and L-smooth function. Let $\delta \in (0, 1)$ and suppose $0 < u \leq \delta \underline{r}/(3\sqrt{d})$. Then there exists some $\kappa(u) \in [199/200, 1]$ such that for all $x \in (1-\delta)\mathcal{X}$,*

$$\left\| \mathbb{E}\left[\mathsf{G}_h(x; u, z)\right] - \kappa(u)\nabla h^u(x) \right\| \leq \frac{2G\bar{R}}{u}\exp\left(\frac{d}{2} - \frac{\delta^2 \underline{r}^2}{4u^2}\right),$$

where the expectation in $\mathbb{E}\left[\mathsf{G}_h(x; u, z)\right]$ is with respect to $z \sim \mathcal{Z}(x, u)$, and $h^u : (1-\delta)\mathcal{X} \to \mathbb{R}$ is given by

$$h^u(x) = \mathbb{E}_{y\sim\mathcal{Y}(u)}[h(x+uy)] \tag{14}$$

for some compactly supported and isotropic distribution $\mathcal{Y}(u)$ that does not depend on the function h. Moreover, h^u is a G-Lipschitz continuous and L-smooth function that satisfies

$$\left|h^u(x) - h(x)\right| \leq \min\left\{uG\sqrt{d}, \frac{1}{2}u^2 Ld\right\}$$

for all $x \in (1-\delta)\mathcal{X}$.

The proof of Lemma 1 is given in our online report (Tang, Ren, & Li, 2020b).

The following lemma handles the second moment of the (delayed) gradient estimation (9), whose proof is given in Appendix A.

Lemma 2. *For any $t \geq 0$, we have*

$$\mathbb{E}\left[\left\|D_j(t)z^i(t)\right\|^2 \middle| \mathcal{F}_t\right] \leq \left(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\right)d_i,$$

$$\mathbb{E}\left[\|G^i(t)\|^2\right] \leq \left(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\right)d_i.$$

The following lemma will be used for bounding the error in the gradient estimation (9) caused by delays.

Lemma 3. *For any $t \geq 0$ and $1 \leq i, j, l \leq n$,*

$$\mathbb{E}\left[\|x^i(t) - x^i(\tau_j^l(t))\|^2\right] \leq \eta^2(b_{lj} + \Delta)^2\left(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\right)d_i,$$

$$\mathbb{E}\left[\|x(t) - x(\tau_j^l(t))\|^2\right] \leq \eta^2(b_{lj} + \Delta)^2\left(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\right)d.$$

Proof. Since ψ_i is 1-strongly convex, we have $\|x^i(t+1)-x^i(t)\|^2 \leq \mathscr{D}_{\psi_i}(x^i(t)|x^i(t+1)) + \mathscr{D}_{\psi_i}(x^i(t+1)|x^i(t))$. The first-order optimality condition of (13) can be written as

$$\left\langle -\eta G^i(t) - \left(\nabla\psi_i(x^i(t+1)) - \nabla\psi_i(x^i(t))\right), \tilde{x}^i - x^i(t+1)\right\rangle \leq 0$$

for all $\tilde{x}^i \in (1-\delta)\mathcal{X}_i$, and together with the identity $\langle \nabla\psi_i(x) -$

$\nabla\psi_i(y), x-y\rangle = \mathscr{D}_{\psi_i}(y|x) + \mathscr{D}_{\psi_i}(x|y)$, we have

$$\begin{aligned}&\|x^i(t+1)-x^i(t)\|^2\\&\le \left\langle \nabla\psi_i(x^i(t+1)) - \nabla\psi_i(x^i(t)), x^i(t+1)-x^i(t)\right\rangle\\&\le -\eta\langle G^i(t), x^i(t+1)-x^i(t)\rangle \le \eta\|G^i(t)\|\,\|x^i(t+1)-x^i(t)\|,\end{aligned}$$

which implies $\|x^i(t+1)-x^i(t)\| \le \eta\|G^i(t)\|$. Then

$$\begin{aligned}&\mathbb{E}\left[\|x^i(t)-x^i(\tau^i_j(t))\|^2\right] \le \mathbb{E}\left[\left(\sum_{\tau=-b_{ij}-\Delta}^{-1}\left\|\eta G^i(\tau)\right\|\right)^2\right]\\&\le \eta^2(b_{ij}+\Delta)\sum_{\tau=-b_{ij}-\Delta}^{-1}\mathbb{E}\left[\|G^i(\tau)\|^2\right]\\&\le \eta^2(b_{ij}+\Delta)^2\left(4\sqrt{3}G^2+\frac{\sigma^2}{2u^2}\right)d_i,\end{aligned}$$

where we used the assumption $t-\tau^i_j(t) \le b_{ij}+\Delta$ in the first step, and used Lemma 2 in the last step. Finally, by taking the sum of the above bound over $i = 1,\ldots,n$, we get the second inequality. □

5.2. Analysis of the complexity

We now analyze the complexity of Algorithm 1 for our constrained convex setting. First, the following is a standard result for mirror descent (see Beck & Teboulle, 2003, Eq. (4.21)).

Lemma 4. *Let $\tilde{x} \in (1-\delta)\mathcal{X}$ be arbitrary. Then*

$$\begin{aligned}&\frac{1}{\eta}\left(\mathscr{D}_\psi(\tilde{x}|x(t+1)) - \mathscr{D}_\psi(\tilde{x}|x(t))\right)\\&\le \langle G(t), \tilde{x}-x(t)\rangle + \frac{\eta}{2}\|G(t)\|^2.\end{aligned} \tag{15}$$

Our analysis of the complexity consists of two steps: (1) bounding the expectation of the right-hand side of (15), and (2) taking the telescoping sum to cancel the terms from the left-hand side of (15).

Step 1: Bounding the expectation of the right-hand side of (15). It can be seen that the expectation of $\eta\|G(t)\|^2/2$ can be bounded via Lemma 2. In order to bound the expectation of $\langle G(t), \tilde{x}-x(t)\rangle$, we note that

$$\begin{aligned}&\mathbb{E}\left[\langle G(t), \tilde{x}-x(t)\rangle\right]\\&= \mathbb{E}\left[\frac{1}{n}\sum_{i,j=1}^n\left\langle D_j(\tau^i_j(t))z^i(\tau^i_j(t)), \tilde{x}^i - x^i(\tau^i_j(t))\right\rangle\right]\\&\quad + \mathbb{E}\left[\frac{1}{n}\sum_{i,j=1}^n\left\langle D_j(\tau^i_j(t))z^i(\tau^i_j(t)), x^i(\tau^i_j(t)) - x^i(t)\right\rangle\right].\end{aligned} \tag{16}$$

The following two lemmas bound the two terms on the right-hand side of (16) respectively.

Lemma 5. *Let $\tilde{x} = (\tilde{x}^1,\ldots,\tilde{x}^n) \in \mathcal{X}$ be arbitrary. Suppose $0 < u \le \delta\underline{r}/(3\sqrt{d})$. Then for $t \ge B$, we have*

$$\begin{aligned}&\frac{1}{\kappa(u)}\mathbb{E}\left[\frac{1}{n}\sum_{i,j=1}^n\left\langle D_j(\tau^i_j(t))z^i(\tau^i_j(t)), \tilde{x}^i - x^i(\tau^i_j(t))\right\rangle\right]\\&\le \mathbb{E}\left[f(\tilde{x}) - f(x(t))\right] + \min\left\{uG\sqrt{d}, \frac{u^2Ld}{2}\right\}\\&\quad + 2\sqrt[4]{3}\eta\bar{\mathfrak{b}}\left(G^2+\frac{\sigma^2}{8\sqrt{3}u^2}\right)\sqrt{d} + \eta L\bar{\mathfrak{b}}\sqrt{nd}\sqrt{4\sqrt{3}G^2+\frac{\sigma^2}{2u^2}}\cdot\overline{R}\\&\quad + \frac{2G\overline{R}^2\sqrt{n}}{\kappa(u)u}\exp\left(\frac{d}{2}-\frac{\delta^2\underline{r}^2}{4u^2}\right),\end{aligned}$$

where $\kappa(u) \in [199/200, 1]$.

The proof of Lemma 5 is given in Appendix B.

Lemma 6. *For any $t \ge 0$,*

$$\begin{aligned}&\mathbb{E}\left[\frac{1}{n}\sum_{i,j=1}^n\left\langle D_j(\tau^i_j(t))z^i(\tau^i_j(t)), x^i(\tau^i_j(t)) - x^i(t)\right\rangle\right]\\&\le \eta\left(4\sqrt{3}G^2+\frac{\sigma^2}{2u^2}\right)\bar{\mathfrak{b}}d.\end{aligned}$$

Proof. We have

$$\begin{aligned}&\mathbb{E}\left[\frac{1}{n}\sum_{i,j=1}^n\left\langle D_j(\tau^i_j(t))z^i(\tau^i_j(t)), x^i(\tau^i_j(t)) - x^i(t)\right\rangle\right]\\&\le \frac{1}{2n}\sum_{i,j=1}^n\mathbb{E}\Big[\eta\bar{\mathfrak{b}}\left\|D_j(\tau^i_j(t))z^i(\tau^i_j(t))\right\|^2\\&\qquad\qquad + \frac{1}{\eta\bar{\mathfrak{b}}}\|x^i(\tau^i_j(t)) - x^i(t)\|^2\Big]\\&\le \frac{1}{2n}\left(4\sqrt{3}G^2+\frac{\sigma^2}{2u^2}\right)\sum_{i,j=1}^n\left[\eta\bar{\mathfrak{b}}d_i + \eta\bar{\mathfrak{b}}^{-1}(b_{ij}+\Delta)^2d_i\right]\\&= \eta\left(4\sqrt{3}G^2+\frac{\sigma^2}{2u^2}\right)\bar{\mathfrak{b}}d,\end{aligned}$$

where we used Lemmas 2 and 3 in the second step. □

Step 2: Taking the telescoping sum. By taking the telescoping sum of (15) and summarizing the previous results, we get the following theorem.

Theorem 3. *Let x^* be a minimizer of $f(x)$ over $x \in \mathcal{X}$. Let $T \ge B$, and let $\bar{x}(T) = \frac{1}{T-B+1}\sum_{t=B}^T x(t)$. Denote $\overline{\mathscr{D}} := \max_{x\in\mathcal{X}}\mathscr{D}_\psi(x^*|x)$. Suppose $0 < u \le \frac{\delta\underline{r}}{3\sqrt{d}}$. Then*

$$\begin{aligned}&\mathbb{E}\left[f(\bar{x}(T))\right] - f(x^*)\\&\le \frac{5\overline{\mathscr{D}}}{4\eta(T-B+1)} + 9\eta\left[G^2+\left(\frac{L\overline{R}}{3}\right)^2\right]\left(\bar{\mathfrak{b}}+\frac{1}{2}\right)\left(\sqrt{d}+\frac{1}{5}\right)^2\\&\quad + \frac{2\eta\sigma^2}{3u^2}\left(\bar{\mathfrak{b}}+\frac{1}{2}\right)\left(\sqrt{d}+\frac{1}{5}\right)^2 + \frac{5G\overline{R}^2\sqrt{n}}{2u}\exp\left(\frac{d}{2}-\frac{\delta^2\underline{r}^2}{4u^2}\right)\\&\quad + uG\sqrt{d} + G\overline{R}\delta.\end{aligned}$$

If it is known that $\nabla f(x^*) = 0$, then

$$\begin{aligned}&\mathbb{E}\left[f(\bar{x}(T))\right] - f(x^*)\\&\le \frac{5\overline{\mathscr{D}}}{4\eta(T-B+1)} + 9\eta\left[G^2+\left(\frac{L\overline{R}}{3}\right)^2\right]\left(\bar{\mathfrak{b}}+\frac{1}{2}\right)\left(\sqrt{d}+\frac{1}{5}\right)^2\\&\quad + \frac{2\eta\sigma^2}{3u^2}\left(\bar{\mathfrak{b}}+\frac{1}{2}\right)\left(\sqrt{d}+\frac{1}{5}\right)^2 + \frac{5G\overline{R}^2\sqrt{n}}{2u}\exp\left(\frac{d}{2}-\frac{\delta^2\underline{r}^2}{4u^2}\right)\\&\quad + \frac{u^2Ld}{2} + \frac{L\overline{R}^2\delta^2}{2}.\end{aligned}$$

Proof. By the previous lemmas, we see that for $t \ge B$,

$$\begin{aligned}&\frac{1}{\kappa(u)\eta}\mathbb{E}\left[\mathscr{D}_\psi(\tilde{x}|x(t+1)) - \mathscr{D}_\psi(\tilde{x}|x(t))\right]\\&\le \frac{\mathbb{E}\left[\langle G(t), \tilde{x}-x(t)\rangle\right]}{\kappa(u)} + \frac{\eta\,\mathbb{E}\left[\|G(t)\|^2\right]}{2\kappa(u)}\end{aligned}$$

$$
\begin{aligned}
&\le \mathbb{E}\left[f(\tilde{x}) - f(x(t))\right] + \min\left\{uG\sqrt{d},\, \frac{u^2Ld}{2}\right\} \\
&\quad + 2\sqrt[4]{3}\eta\bar{\mathfrak{b}}\Big(G^2 + \frac{\sigma^2}{8\sqrt{3}u^2}\Big)\sqrt{d} + \eta L\bar{b}\sqrt{nd}\sqrt{4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}}\cdot\bar{R} \\
&\quad + \frac{2G\bar{R}^2\sqrt{n}}{\kappa(u)u}\exp\left(\frac{d}{2} - \frac{\delta^2\underline{r}^2}{4u^2}\right) + \frac{\eta}{\kappa(u)}\Big(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\Big)\bar{\mathfrak{b}}d \\
&\quad + \frac{\eta}{2\kappa(u)}\Big(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\Big)d.
\end{aligned}
$$

By taking the telescoping sum and noting that $\kappa(u) \in [199/200, 1]$, we get

$$
\begin{aligned}
&\mathbb{E}\left[\frac{1}{T-B+1}\sum_{t=B}^{T} f(x(t))\right] - f(\tilde{x}) \\
&\le \frac{\mathbb{E}\left[\mathcal{D}_\psi(\tilde{x}|x(B))\right]}{\kappa(u)\eta(T-B+1)} + \min\left\{uG\sqrt{d},\, \frac{u^2Ld}{2}\right\} \\
&\quad + 9\eta\left[G^2 + \left(\frac{L\bar{R}}{3}\right)^2\right]\left(\bar{\mathfrak{b}} + \frac{1}{2}\right)\left(\sqrt{d} + \frac{1}{5}\right)^2 \\
&\quad + \frac{2\eta}{3}\cdot\frac{\sigma^2}{u^2}\left(\bar{\mathfrak{b}} + \frac{1}{2}\right)\left(\sqrt{d} + \frac{1}{5}\right)^2 \\
&\quad + \frac{2G\bar{R}^2\sqrt{n}}{\kappa(u)u}\exp\left(\frac{d}{2} - \frac{\delta^2\underline{r}^2}{4u^2}\right),
\end{aligned}
$$

where we plugged in the following bounds derived by noting $\kappa(u) \in [199/200, 1]$ and some inequality manipulation:

$$
\begin{aligned}
&2\sqrt[4]{3}\eta\bar{\mathfrak{b}}\Big(G^2 + \frac{\sigma^2}{8\sqrt{3}u^2}\Big)\sqrt{d} + \frac{\eta}{\kappa(u)}\Big(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\Big)\bar{\mathfrak{b}}d \\
&\quad + \frac{\eta}{2\kappa(u)}\Big(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\Big)d \\
&\le 7\eta\left(\bar{\mathfrak{b}} + \frac{1}{2}\right)\left(\sqrt{d} + \frac{1}{5}\right)^2\Big(G^2 + \frac{\sigma^2}{8\sqrt{3}u^2}\Big), \\
&\eta L\bar{b}\sqrt{nd}\sqrt{4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}}\cdot\bar{R} \\
&\le \eta Ld\sqrt{\frac{\sum_{i,j}(b_{ij}+\Delta)^2}{nd}}\cdot\frac{1}{2}\left[\frac{1}{2L}\Big(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\Big) + 2L\bar{R}^2\right] \\
&\le \eta\left(\sqrt{d} + \frac{1}{5}\right)^2\left(\bar{\mathfrak{b}} + \frac{1}{2}\right)\Big(\sqrt{3}G^2 + \frac{\sigma^2}{8u^2} + L^2\bar{R}^2\Big).
\end{aligned}
$$

Now let $\tilde{x} = \mathcal{P}_{(1-\delta)\mathcal{X}}[x^*]$. We have $\|\tilde{x}-x^*\| \le \|(1-\delta)x^*-x^*\| \le \delta\bar{R}$, and so $f(\tilde{x}) - f(x^*) \le G\|\tilde{x}-x^*\| \le G\delta\bar{R}$; if we further know that $\nabla f(x^*) = 0$, then $f(\tilde{x}) - f(x^*) \le L\|\tilde{x}-x^*\|^2/2 \le L\delta^2\bar{R}^2/2$. Summarizing the above results and plugging in lower bounds of $\kappa(u)$, we can then get the desired results by further noting $\mathbb{E}\left[\frac{1}{T-B+1}\sum_{t=B}^{T} f(x(t))\right] \ge \mathbb{E}\left[f(\bar{x}(T))\right]$ since f is convex, and using $\mathbb{E}\left[\mathcal{D}_\psi(\tilde{x}|x(B))\right] \le \bar{\mathcal{D}}$. □

Now Theorems 1 and 2 can be proved.

Proof of Theorem 1. The condition on u implies

$$
u \le \frac{u}{\sqrt{d}}\cdot\sqrt{d + \frac{4}{9}\left[\ln\frac{20G\bar{R}^2\sqrt{n}}{u\epsilon}\right]_+} \le \frac{\delta\underline{r}}{3\sqrt{d}},
$$

meaning that the condition of Theorem 3 is satisfied. The condition on u also implies

$$
\begin{aligned}
&\frac{5G\bar{R}^2\sqrt{n}}{2u}\exp\Big(\frac{d}{2} - \frac{\delta^2\underline{r}^2}{4u^2}\Big) \\
&= \frac{e^{-7d/4}\epsilon}{8}\exp\left[\frac{9}{4}\left(d + \frac{4}{9}\ln\frac{20G\bar{R}^2\sqrt{n}}{u\epsilon} - \left(\frac{\delta\underline{r}}{3u}\right)^2\right)\right] \le \frac{e^{-7/4}\epsilon}{8}.
\end{aligned}
$$

The conditions on δ, η and T further guarantee

$$
\begin{aligned}
&\frac{5\bar{\mathcal{D}}}{4\eta(T-B+1)} + uG\sqrt{d} + G\bar{R}\delta \\
&\quad + 9\eta\left[G^2 + \left(\frac{L\bar{R}}{3}\right)^2\right]\left(\bar{\mathfrak{b}} + \frac{1}{2}\right)\left(\sqrt{d} + \frac{1}{5}\right)^2 \\
&\le \frac{\epsilon}{6} + \frac{\epsilon\underline{r}}{15\bar{R}} + \frac{\epsilon}{5} + \frac{\epsilon}{2} \le \frac{14}{15}\epsilon.
\end{aligned}
$$

Summarizing all these bounds and using Algorithm 1 with $\sigma = 0$ shows that $\mathbb{E}[f(\bar{x}(t))] - f(x^*) \le \epsilon$. □

Proof of Theorem 2. Just as in the proof of Theorem 1, We can similarly show that $u \le \delta\underline{r}/(3\sqrt{d})$ and that $\frac{5G\bar{R}^2\sqrt{n}}{2u}\exp\Big(\frac{d}{2} - \frac{\delta^2\underline{r}^2}{4u^2}\Big) \le \frac{e^{-7d/4}}{8}\epsilon$. Moreover, the condition on u implies that, for sufficiently small $\epsilon > 0$,

$$
u \le \Theta\Big(\frac{\delta}{\sqrt{d + \ln(1/\epsilon)}}\Big), \tag{17}
$$

in which equality can be achieved if the condition on u is satisfied with equality. By plugging in the conditions on the parameters, it can be established that

$$
\begin{aligned}
&\frac{5\bar{\mathcal{D}}}{4\eta(T-B+1)} + \frac{2\eta\sigma^2}{3u^2}\left(\bar{\mathfrak{b}} + \frac{1}{2}\right)\left(\sqrt{d} + \frac{1}{5}\right)^2 \\
&\quad + 9\eta\left[G^2 + \left(\frac{L\bar{R}}{3}\right)^2\right]\left(\bar{\mathfrak{b}} + \frac{1}{2}\right)\left(\sqrt{d} + \frac{1}{5}\right)^2 \\
&\le \frac{\epsilon}{6} + \frac{\epsilon}{3} + \frac{9u^2\epsilon}{2\sigma^2}\left[G^2 + \left(\frac{L\bar{R}}{3}\right)^2\right] \le \frac{2\epsilon}{3}
\end{aligned}
$$

for sufficiently small $\epsilon > 0$.

Now, if $\delta \le \epsilon/(5G\bar{R})$, then $uG\sqrt{d} + G\bar{R}\delta \le \frac{\epsilon\underline{r}}{15\bar{R}} + \frac{\epsilon}{5} \le \frac{4\epsilon}{15}$, and since $e^{-7d/4}/8 < 1/15$, by summarizing the above results and using Theorem 3, we get $\mathbb{E}[f(\bar{x}(T))] - f(x^*) \le \epsilon$.

If $\nabla f(x^*) = 0$ and $\delta \le \sqrt{\epsilon}/(\bar{R}\sqrt{2L})$, then $\frac{u^2Ld}{2} + \frac{L\bar{R}^2}{2}\delta^2 \le \frac{\epsilon\underline{r}^2}{36\bar{R}^2} + \frac{\epsilon}{4} \le \frac{5\epsilon}{18}$, and since $e^{-7d/4}/8 < 1/18$, by summarizing the above results and using Theorem 3, we get $\mathbb{E}[f(\bar{x}(T))] - f(x^*) \le \epsilon$.

The asymptotic behavior of T can be derived from (17) and the conditions on the parameters. □

6. Knowledge of local function dependence

In the previous sections we assume that each local cost f_j may be affected by any other agent's action, i.e., $\nabla^i f_j$ can be nonzero for any i. However, in some situations, f_j may only depend on the actions of a subset of agents, and the agents may have knowledge of this dependence. In this section, we briefly discuss the benefits when the agents have additional knowledge of such local function dependence information.

Let $\mathcal{A}_i$ be the set of agents whose local costs will be affected by agent i's action (i.e., $\nabla^i f_j$ is not always zero for each $j \in \mathcal{A}_i$). Then, if each agent i knows its associated set $\mathcal{A}_i$, due to the fact that $\nabla^i f(x) = \frac{1}{n}\sum_{j\in\mathcal{A}_i}\nabla^i f_j(x)$, the partial gradient estimator (9) can be further simplified as

$$
G^i(t) = \frac{1}{n}\sum_{j\in\mathcal{A}_i} D_j^i(t)\, z^i(\tau_j^i(t)). \tag{18}
$$

In this case, the following lemma shows that the second-moment of the gradient estimator will be reduced:

Lemma 7. *We have*

$$\mathbb{E}\left[\left\|\frac{1}{n}\sum_{j\in\mathcal{A}_i} D_j^i(t)z^i(\tau_j^i(t))\right\|^2\right] \le \left(4\sqrt{3}G^2+\frac{\sigma^2}{2u^2}\right)\frac{|\mathcal{A}_i|^2}{n^2}d_i.$$

Proof. By the first bound in Lemma 2, we have

$$\begin{aligned}&\mathbb{E}\left[\left\|\frac{1}{n}\sum_{j\in\mathcal{A}_i} D_j^i(t)\,z^i(\tau_j^i(t))\right\|^2\right]\\ &\le \frac{|\mathcal{A}_i|}{n^2}\sum_{j\in\mathcal{A}_i}\mathbb{E}\left[\left\|D_j^i(\tau_j^i(t))\,z^i(\tau_j^i(t))\right\|^2\right]\\ &\le \frac{|\mathcal{A}_i|}{n^2}\sum_{j\in\mathcal{A}_i}\left(4\sqrt{3}G^2+\frac{\sigma^2}{2u^2}\right)d_i = \frac{|\mathcal{A}_i|^2}{n^2}\left(4\sqrt{3}G^2+\frac{\sigma^2}{2u^2}\right)d_i,\end{aligned}$$

which completes the proof. □

Compared to Lemma 2, we see that $\mathbb{E}\left[\|G^i(t)\|^2\right]$ is reduced by a factor of $|\mathcal{A}_i|^2/n^2$. Consequently, the complexity of Algorithm 1 can be further improved with better dependence on the network topology and the number of agents. We omit detailed analysis here but provide brief numerical comparison in Section 7.

Another benefit brought by the knowledge of $\mathcal{A}_i$ is communication savings. Originally, in Algorithm 1, each agent needs to send the whole array (8) to its neighbors. On the other hand, the following theorem shows that, the communication burden can be relieved if $\mathcal{A}_i$ is known to each agent i and the communication network has a structure compatible with the sets $\mathcal{A}_i$.

Theorem 4. *Suppose for any i, j, l such that $j \in \mathcal{A}_l\backslash\mathcal{A}_i$ (i.e., f_j depends on x^l but not x^i), the following conditions hold:*

1. *There exists a path P_{lj} in $\mathcal{G}$ connecting l and j which does not contain i.*
2. *For any agent r on the path P_{lj}, f_j depends on x_r.*

Further, suppose no communication failures occur at any link. Then, in order for each agent to be able to construct the partial gradient estimator (18) with $t-\tau_j^i(t)$ being bounded, each agent i only needs to record, update and pass $(D_j^i(t), \tau_j^i(t))$ for $j \in \mathcal{A}_i$.

Proof. It suffices to show that for each i, agent i does not need to pass information about the difference quotient of f_j for any $j \notin \mathcal{A}_i$ for the sake of other agents' updates.

Let $i \in \{1,\dots,n\}$ and $j \notin \mathcal{A}_i$ be arbitrary, and let l be an arbitrary agent such that $j \in \mathcal{A}_l$. By the first condition stated in the theorem, we know that there exists a path P_{lj} not containing i. Moreover, by the second condition, for any agent r on the path P_{lj}, f_j is a function of x_r, so agent r receives and passes on information about f_j. This then implies that agent l can successfully receive the information it needs from f_j via the path P_{lj}, and further that $t-t_j^l(t)$ is upper bounded by the length of P_{lj}. Hence, agent i does not need to pass on information about f_j for agent l, and by the arbitrariness of i, j and l, we get the desired conclusion. □

Theorem 4 shows that, when the communication graph is "compatible" with the local function dependence (in the sense stated in the conditions of the theorem), the number of columns of the array (8) can then be reduced from n to $|\mathcal{A}_i|$ for each agent i, which also leads to reduced communication burden. We mention that Theorem 4 analyzes only one possibility of "compatibility" between the communication network and the local function dependence, and one can propose other compatibility conditions for the communication network so that the size of the array (8) and/or the communication burden can be reduced. Investigating other notions of compatibility between the communication network and the local function dependence for the ZFO algorithm will be an interesting direction which we leave as future work.

7. Numerical examples

We demonstrate the performance of our ZFO algorithm on finding the optimal flow of a nonatomic routing game introduced in Section 2.1. It can be shown that the global objective function is (Nisan et al., 2007)

$$f(x) = \frac{1}{n}\sum_e q_e(x)\cdot c_e(q_e(x)),$$

where $q_e(x) = \sum_j \sum_{p'\in\mathcal{P}_j:e\in p'} x_{p'}^j Q_j$ is the total traffic through the edge e. Since q_e is affine in x and the function $t \mapsto t\cdot c_e(t)$ is convex, $f(x)$ is a convex function of x.

Just as explained in Section 2.1, we eliminate one entry $x_{\check{p}_i}^i$ from each action vector and perform a translation so that the new feasible set has a nonempty interior containing the origin. We choose ψ_i to be the unnormalized negative entropy of $\left(x_p^i\right)_{p\in\mathcal{P}_i}$, i.e.,

$$\psi_i(x^i) = \sum_{p\in\mathcal{P}_i} x_p^i(\ln x_p^i - 1),$$

where $x_{\check{p}_i}^i := 1-\sum_{p\in\mathcal{P}_i\backslash\{\check{p}_i\}} x_p^i$.

The routing network of the test case consists of 28 vertices and 85 edges, and the congestion function $c_e(t)$ for each edge is a convex and increasing quadratic function. There are 60 agents in the test case. Each agent i is associated with a pair of vertices (s_i, t_i), and is allowed to use 4 paths from s_i to t_i in the routing network for sending its traffic Q_i. We consider three communication networks that connect the group of agents:

1. A linear chain network, with $\bar{\mathfrak{b}} = 24.4915$.
2. A 4×15 grid network, with $\bar{\mathfrak{b}} = 7.2303$.
3. A randomly generated Erdős–Rényi network, with $\bar{\mathfrak{b}} = 4.3522$.

We assume no additional delays in the communication network so that $\Delta = 0$. The optimal value for the test case is given by $f^* = 5.4530$. Details of the test cases and the code for our numerical experiments can be found at https://github.com/tyj518/ZFO_Distributed-Routing.

Noiseless setting. We first simulate the setting where the function value observations are noiseless and we do not assume knowledge of the function dependence. We also run the centralized zeroth-order optimization method for our test case as a benchmark. We use the same step size η and shrinkage factor δ for the three communication networks as well as the centralized method. The results are shown in Fig. 2, where the dark curves represent the average of $f(x(t))$, the light bands indicate 3.0-standard deviation confidence intervals computed from 50 random trials, and the red dash-dotted line indicates the optimal value f^*.

It can be seen that our ZFO algorithm is able to approach the optimal value with satisfactory convergence behavior. Moreover, while the three communication networks have different values of $\bar{\mathfrak{b}}$, our algorithm exhibits almost indistinguishable performance on the three communication networks and compared to the centralized algorithm. This suggests the theoretical complexity result in Theorem 1 might be more conservative compared to the real performance in terms of the dependence on the communication network's connectivity. It would be interesting to further investigate whether and how the theoretical analysis can be improved, but we leave it as future work.

Noisy setting. We then consider the setting where the function value observations are noisy, and simulate two cases $\sigma = 0.01f^*$

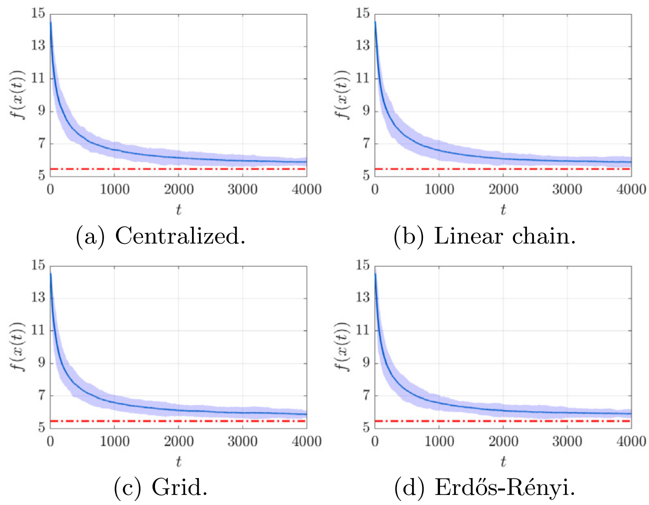

Fig. 2. Noiseless, no knowledge of local function dependence.

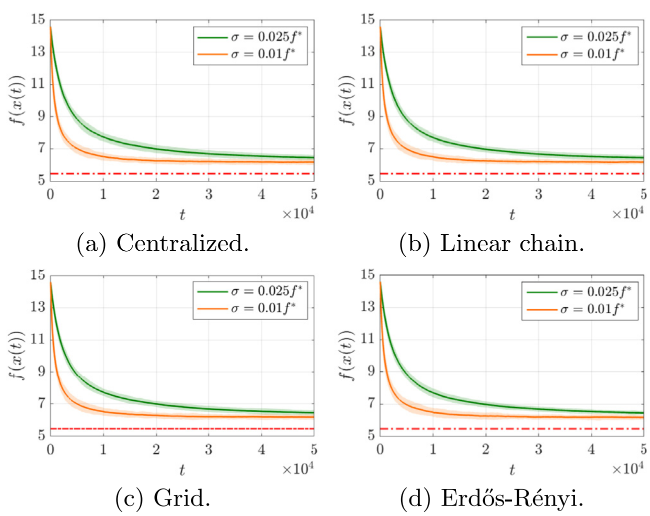

Fig. 3. Noisy, no knowledge of local function dependence.

and $\sigma = 0.025f^*$. We decrease the step size η and increase the smoothing radius u as well as the shrinkage factor δ as the noise level σ increases, in order to suppress the variance associated with the zeroth-order gradient estimator. We do not assume knowledge of the function dependence for both cases.

The results are shown in Fig. 3, where again the dark curves represent the average of $f(x(t))$, the light bands indicate 3.0-standard deviation confidence intervals computed from 50 random trials, and the red dash-dotted line indicates the optimal value f^*. Compared to the noiseless case, we see that the convergence is substantially slower. Also, as the noise level increases, the convergence becomes slower, and the final optimality gap becomes larger. On the other hand, we again observe that our algorithm achieves very similar convergence behavior for the three communication networks and compared with the centralized setting, which suggests that our algorithm may have better performance than indicated by Theorem 2 in terms of the dependence on the communication network's connectivity.

Known local function dependence. In this setting, we assume that each agent knows the set $\mathcal{A}_i$ of local function dependence, and employs (18) for gradient estimation. We enlarge the step sizes by 1.5 times compared to the corresponding scenarios without knowledge of $\mathcal{A}_i$. Fig. 4 compares the convergence behavior

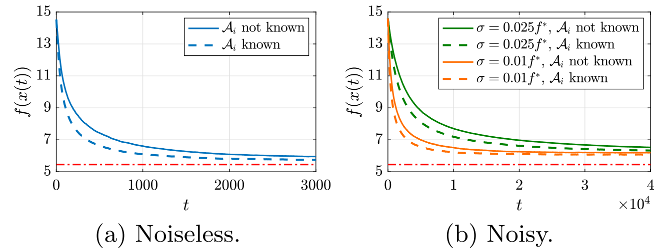

Fig. 4. With local function dependence as in (18), linear chain communication network.

of $f(x(t))$ with and without utilizing local function dependence, averaged over 50 random trials; here only the results for the linear chain communication network are presented, but we mention that the results for the other two communication networks are very similar. It is not surprising to see that both the convergence rates and the final optimality gaps are improved when knowledge of $\mathcal{A}_i$ has been utilized in our algorithm.

8. Conclusion and future directions

We consider the cooperative multi-agent optimization problem, where a group of agents determine their actions cooperatively through observations of only their local cost values, and each local cost is affected by all agents' actions. We propose a zeroth-order feedback optimization (ZFO) algorithm, and conduct theoretical analysis on its performance. Specifically, we provide complexity bounds of our algorithm for constrained convex problems with noiseless and noisy function value observations. We also briefly discuss the benefits of utilizing local function dependence. Some interesting future directions include (i) extending the algorithm to handle coupled constraints on the actions, (ii) analysis for constrained nonconvex problems, (iii) investigating whether the sampling procedure of the random perturbation can be simplified, (iv) improving the algorithm's complexity by incorporating, e.g., variance reduction techniques, (v) extending the algorithm to handle asynchronous local action updates, (vi) further investigation on how local function dependence can be exploited.

Appendix A. Proof of Lemma 2

We first provide some useful lemmas.

Lemma 8 (*Concentration Inequality, Boucheron, Lugosi, & Massart, 2013, Theorem 5.6*). *Let $h : \mathbb{R}^d \to \mathbb{R}$ be G-Lipschitz. Then we have*

$$\mathbb{P}_{z\sim\mathcal{N}(0,I_d)}\left(|h(z) - \mathbb{E}_z[h(z)]| \geq t\right) \leq 2\exp\left(-t^2/(2G^2)\right).$$

With the help of the concentration inequality, we can prove the following lemma.

Lemma 9. *Let $h : \mathbb{R}^d \to \mathbb{R}$ be G-Lipschitz. Then*

$$\mathbb{E}_z\left[\left|\frac{h(z)-h(-z)}{2}z_i\right|^2\right] \leq 4\sqrt{3}G^2,$$

where $z = (z_1, \ldots, z_d) \sim \mathcal{N}(0, I_d)$.

Proof. The proof follows Shamir (2017, Lemmas 9 & 10) closely. Denote $\bar{h} = \mathbb{E}_z[h(z)]$. We have

$$\mathbb{E}_z\left[\left|\frac{h(z)-h(-z)}{2}z_i\right|^2\right] = \frac{1}{4}\mathbb{E}_z\left[z_i^2(h(z)-h(-z))^2\right]$$

$$= \frac{1}{4}\mathbb{E}_z\left[z_i^2\left((h(z)-\bar{h})-(h(-z)-\bar{h})\right)^2\right]$$
$$\le \frac{1}{2}\mathbb{E}_z\left[z_i^2\left((h(z)-\bar{h})^2+(h(-z)-\bar{h})^2\right)\right]$$
$$= \mathbb{E}_z\left[z_i^2(h(z)-\bar{h})^2\right].$$

Then

$$\mathbb{E}_z\left[z_i^2(h(z)-\bar{h})^2\right] \le \sqrt{\mathbb{E}_z[z_i^4]}\cdot\sqrt{\mathbb{E}_z[(h(z)-\bar{h})^4]}$$
$$= \sqrt{3}\left(\int_0^{+\infty}\mathbb{P}_z\left((h(z)-\bar{h})^4\ge t\right)dt\right)^{1/2}$$
$$\le \sqrt{3}\left(\int_0^{+\infty}2\exp\left(-\frac{\sqrt{t}}{2G^2}\right)dt\right)^{1/2} = 4\sqrt{3}G^2,$$

where the first step follows from the Cauchy–Schwarz inequality, the second step uses the fact that $\mathbb{E}_z[z_i^4]=3$, and the third step uses Lemma 8. □

We then derive bounds on the second moment of the gradient estimator (2) with $z\sim\mathcal{Z}(x,u)$.

Lemma 10. *Let $h:\mathcal{X}\to\mathbb{R}$ be G-Lipschitz, and let $\delta\in(0,1)$ be arbitrary. Then for any $x\in(1-\delta)\mathcal{X}$ and $1\le i\le d$,*

$$\mathbb{E}_{z\sim\mathcal{Z}(x,u)}\left[\left\|\frac{h(x+uz)-h(x-uz)}{2u}z^i\right\|^2\right]\le 4\sqrt{3}G^2 d_i.$$

Proof. Let $\tilde{h}(z)=h\left(x+u\cdot\mathcal{P}_{\mathcal{S}(x,u)}[z]\right),\ \forall z\in\mathbb{R}^d$. Then

$$\left|\tilde{h}(z_1)-\tilde{h}(z_2)\right|$$
$$= \left|h\left(x+u\cdot\mathcal{P}_{\mathcal{S}(x,u)}[z_1]\right)-h\left(x+u\cdot\mathcal{P}_{\mathcal{S}(x,u)}[z_2]\right)\right|$$
$$\le uG\left\|\mathcal{P}_{\mathcal{S}(x,u)}[z_1]-\mathcal{P}_{\mathcal{S}(x,u)}[z_2]\right\|\le uG\left\|z_1-z_2\right\|,$$

showing that $\tilde{h}$ is a uG-Lipschitz continuous function on $\mathbb{R}^d$. Moreover, we have

$$\mathbb{E}_{z\sim\mathcal{Z}(x,u)}\left[\left\|(h(x+uz)-h(x-uz))z^i\right\|^2\right]$$
$$= \mathbb{E}_{z\sim\mathcal{N}(0,I_d)}\left[\left\|(\tilde{h}(z)-\tilde{h}(-z))\cdot\mathcal{P}_{\mathcal{S}_i(x^i,u)}[z^i]\right\|^2\right]$$
$$\le \mathbb{E}_{z\sim\mathcal{N}(0,I_d)}\left[\left\|(\tilde{h}(z)-\tilde{h}(-z))z^i\right\|^2\right],$$

where the last inequality follows from $\|\mathcal{P}_{\mathcal{S}_i(x^i,u)}[z^i]\|\le\|z^i\|$ as $\mathcal{S}_i(x^i,u)$ is a convex set containing the origin.

Then by Lemma 9, we have

$$\mathbb{E}_{z\sim\mathcal{N}(0,I_d)}\left[\left\|\frac{\tilde{h}(z)-\tilde{h}(-z)}{2u}z^i\right\|^2\right]\le\frac{1}{u^2}\cdot4\sqrt{3}u^2G^2\cdot d_i$$
$$=4\sqrt{3}G^2d_i,$$

which gives the desired bound. □

We are now ready to prove Lemma 2. Denoting $\varepsilon_j(t)=\varepsilon_j^+(t)-\varepsilon_j^-(t)$, we have

$$\mathbb{E}\left[\left\|D_j(t)z^i(t)\right\|^2\middle|\mathcal{F}_t\right]$$
$$=\mathbb{E}\left[\left\|\frac{f_j(x(t)+uz(t))-f_j(x(t)-uz(t))}{2u}z^i(t)\right\|^2\middle|\mathcal{F}_t\right]$$
$$+\frac{1}{4u^2}\mathbb{E}\left[\varepsilon_j(t)^2\|z^i(t)\|^2|\mathcal{F}_t\right]$$
$$\le 4\sqrt{3}G^2d_i+\frac{\sigma^2}{2u^2}d_i,$$

where we used Lemma 10, the independence between $\varepsilon_j(t)$ and $z^i(t)$, and the fact that $\mathbb{E}_{z^i\sim\mathcal{Z}^i(x,u)}\left[\|z^i\|^2\right]\le\mathbb{E}_{z^i\sim\mathcal{N}(0,I_{d_i})}\left[\|z^i\|^2\right]\le d_i$. Then,

$$\mathbb{E}\left[\|G^i(t)\|^2\right]\le\frac{1}{n}\sum_{j=1}^n\mathbb{E}\left[\left\|D_j(\tau_j^i(t))z^i(\tau_j^i(t))\right\|^2\right]$$
$$\le\left(4\sqrt{3}G^2+\frac{\sigma^2}{2u^2}\right)d_i.$$

Note that by summing over $i=1,\dots,n$, we can get a similar bound on $\mathbb{E}\left[\|G(t)\|^2\right]$.

Appendix B. Proof of Lemma 5

For each $\tau\ge0$, we have

$$\mathbb{E}\left[\frac{1}{n}\sum_{i,j}\left\langle D_j(\tau_j^i(t))z^i(\tau_j^i(t)),\tilde{x}^i-x^i(\tau_j^i(t))\right\rangle\cdot 1_{\tau_j^i(t)=\tau}\middle|\mathcal{F}_\tau\right]$$
$$=1_{\tau_j^i(t)=\tau}\cdot\left(\frac{1}{n}\sum_{i,j}\left\langle\kappa(u)\nabla^i f_j^u(x(\tau)),\tilde{x}^i-x^i(\tau)\right\rangle\right.$$
$$\left.+\frac{1}{n}\sum_{i,j}\left\langle\mathbb{E}\left[D_j(\tau)z^i(\tau)|\mathcal{F}_\tau\right]-\kappa(u)\nabla^i f_j^u(x(\tau)),\tilde{x}^i-x^i(\tau)\right\rangle\right),$$

where the second term can be bounded by Lemma 1 and $\sum_{i=1}^n\bar{R}_i\le\sqrt{n}\bar{R}$ as

$$\frac{1}{n}\sum_{i,j}\left\langle\mathbb{E}\left[D_j(\tau)z^i(\tau)|\mathcal{F}_\tau\right]-\kappa(u)\nabla^i f_j^u(x(\tau)),\tilde{x}^i-x^i(\tau)\right\rangle$$
$$\le\frac{1}{n}\sum_{i,j}\frac{2G\bar{R}}{u}\exp\left(\frac{d}{2}-\frac{\delta^2\underline{r}^2}{4u^2}\right)\bar{R}_i$$
$$\le\frac{2G\bar{R}}{u}\exp\left(\frac{d}{2}-\frac{\delta^2\underline{r}^2}{4u^2}\right)\sqrt{n}\cdot\bar{R}.$$

Therefore

$$\mathbb{E}\left[\frac{1}{n}\sum_{i,j}\left\langle D_j(\tau_j^i(t))z^i(\tau_j^i(t)),\tilde{x}^i-x^i(\tau_j^i(t))\right\rangle\right]$$
$$=\sum_\tau\mathbb{E}\left[\mathbb{E}\left[\frac{1}{n}\sum_{i,j}1_{\tau_j^i(t)=\tau}\cdot\left\langle D_j(\tau_j^i(t))z^i(\tau_j^i(t)),\right.\right.\right.$$
$$\left.\left.\left.\tilde{x}^i-x^i(\tau_j^i(t))\right\rangle\middle|\mathcal{F}_\tau\right]\right]$$
$$\le\frac{\kappa(u)}{n}\mathbb{E}\left[\sum_{i,j}\left\langle\nabla^i f_j^u\left(x(\tau_j^i(t))\right),\tilde{x}^i-x^i(\tau_j^i(t))\right\rangle\right]$$
$$+\frac{2G\bar{R}}{u}\exp\left(\frac{d}{2}-\frac{\delta^2\underline{r}^2}{4u^2}\right)\sqrt{n}\cdot\bar{R}.$$

Now,

$$\frac{1}{n}\sum_{i,j}\left\langle\nabla^i f_j^u\left(x(\tau_j^i(t))\right),\tilde{x}^i-x^i(\tau_j^i(t))\right\rangle$$
$$=\left\langle\nabla f^u(x(t)),\tilde{x}-x(t)\right\rangle$$
$$+\frac{1}{n}\sum_{i,j}\left\langle\nabla^i f_j^u(x(t)),x^i(t)-x^i(\tau_j^i(t))\right\rangle$$
$$+\frac{1}{n}\sum_{i,j}\left\langle\nabla^i f_j^u\left(x(\tau_j^i(t))\right)-\nabla^i f_j^u(x(t)),\tilde{x}^i-x^i(\tau_j^i(t))\right\rangle,$$

where $f^u(x):=\frac{1}{n}\sum_j f_j^u(x)$. Note that by (14), we have $f^u(x)=\mathbb{E}_{y\sim\mathcal{Y}(u)}[f(x+uy)]$, and together with the convexity of f, we see that f^u is convex and $f^u(x)\ge f(x)$. Then by Lemma 1,

$$\left\langle\nabla f^u(x(t)),\tilde{x}-x(t)\right\rangle\le f^u(\tilde{x})-f^u(x(t))$$

$$\leq f(\tilde{x}) - f(x(t)) + \min\left\{uG\sqrt{d},\, \frac{1}{2}u^2 Ld\right\},$$

and by Lemma 3, we have

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{n}\sum_{i,j}\left\langle \nabla^i f_j^u(x(t)), x^i(t) - x^i(\tau_j^i(t))\right\rangle\right] \\
&\leq \frac{1}{2n}\sum_{i,j}\Bigg(2\sqrt[4]{3}\eta\bar{\mathfrak{b}}\sqrt{d}\,\mathbb{E}\left[\|\nabla^i f_j^u(x(t))\|^2\right] \\
&\qquad + \frac{\mathbb{E}\left[\|x^i(t) - x^i(\tau_j^i(t))\|^2\right]}{2\sqrt[4]{3}\eta\bar{\mathfrak{b}}\sqrt{d}}\Bigg) \\
&\leq \frac{1}{2n}\Bigg(2\sqrt[4]{3}\eta\bar{\mathfrak{b}}\sqrt{d}\,nG^2 \\
&\qquad + \frac{1}{2\sqrt[4]{3}\eta\bar{\mathfrak{b}}\sqrt{d}}\,\eta^2\left(4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}\right)\cdot\sum_{i,j}(b_{ij} + \Delta)^2 d_i\Bigg) \\
&= \frac{1}{2}\left(2\sqrt[4]{3}\eta\bar{\mathfrak{b}}G^2\sqrt{d} + 2\sqrt[4]{3}\eta\bar{\mathfrak{b}}\left(G^2 + \frac{\sigma^2}{8\sqrt{3}u^2}\right)\sqrt{d}\right) \\
&\leq 2\sqrt[4]{3}\cdot\eta\bar{\mathfrak{b}}\left(G^2 + \frac{\sigma^2}{8\sqrt{3}u^2}\right)\sqrt{d},
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{n}\sum_{i,j}\left\langle \nabla^i f_j^u\left(x(\tau_j^i(t))\right) - \nabla^i f_j^u(x(t)), \tilde{x}^i - x^i(\tau_j^i(t))\right\rangle\right] \\
&\leq \frac{1}{n}\sum_{i,j}\mathbb{E}\left[\left\|\nabla^i f_j^u\left(x(\tau_j^i(t))\right) - \nabla^i f_j^u(x(t))\right\|\bar{R}_i\right] \\
&\leq \frac{L}{n}\sum_{i,j}\sqrt{\mathbb{E}\left[\|x(\tau_j^i(t)) - x(t)\|^2\right]}\cdot\bar{R}_i \\
&\leq \frac{\eta L\sqrt{d}}{n}\sqrt{4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}}\sum_{i,j}(b_{ij} + \Delta)\bar{R}_i \\
&\leq \eta L\bar{b}\sqrt{nd}\sqrt{4\sqrt{3}G^2 + \frac{\sigma^2}{2u^2}}\cdot\bar{R},
\end{aligned}$$

where the last step follows from Cauchy's inequality. Summarizing these results, we get the desired bound.

References

Agarwal, A., Dekel, O., & Xiao, L. (2010). Optimal algorithms for online convex optimization with multi-point bandit feedback. In *Proceedings of the 23rd conference on learning theory*.

Agarwal, A., & Duchi, J. C. (2011). Distributed delayed stochastic optimization. In *Advances in neural information processing systems*: *Vol. 24*.

Agarwal, A., Foster, D. P., Hsu, D., Kakade, S. M., & Rakhlin, A. (2013). Stochastic convex optimization with bandit feedback. *SIAM Journal on Optimization*, *23*(1), 213–240.

Bach, F., & Perchet, V. (2016). Highly-smooth zero-th order online optimization. In *Proceedings of the 29th conference on learning theory*.

Beck, A., & Teboulle, M. (2003). Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, *31*(3), 167–175.

Boucheron, S., Lugosi, G., & Massart, P. (2013). *Concentration inequalities: A nonasymptotic theory of independence*. Oxford University Press.

Boyle, J. P., & Dykstra, R. L. (1986). A method for finding projections onto the intersection of convex sets in Hilbert spaces. In R. Dykstra, T. Robertson, & F. T. Wright (Eds.), *Advances in order restricted statistical inference* (pp. 28–47). New York, NY: Springer.

Candogan, U. O., Menache, I., Ozdaglar, A., & Parrilo, P. A. (2010). Near-optimal power control in wireless networks: A potential game approach. In *2010 proceedings IEEE INFOCOM*.

Chang, T.-H., Hong, M., & Wang, X. (2015). Multi-agent distributed optimization via inexact consensus ADMM. *IEEE Transactions on Signal Processing*, *63*(2), 482–497.

Cortés, J., Martínez, S., Karataş, T., & Bullo, F. (2004). Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*, *20*(2), 243–255.

Dougherty, S., & Guay, M. (2017). An extremum-seeking controller for distributed optimization over sensor networks. *IEEE Transactions on Automatic Control*, *62*(2), 928–933.

Duchi, J. C., Jordan, M. I., Wainwright, M. J., & Wibisono, A. (2015). Optimal rates for zero-order convex optimization: The power of two function evaluations. *IEEE Transactions on Information Theory*, *61*(5), 2788–2806.

Flaxman, A. D., Kalai, A. T., & McMahan, H. B. (2005). Online convex optimization in the bandit setting: Gradient descent without a gradient. In *Proceedings of the sixteenth annual ACM-SIAM symposium on discrete algorithms*.

Hajinezhad, D., Hong, M., & Garcia, A. (2019). ZONE: Zeroth-order nonconvex multiagent optimization over networks. *IEEE Transactions on Automatic Control*, *64*(10), 3995–4010.

Larson, J., Menickelly, M., & Wild, S. M. (2019). Derivative-free optimization methods. *Acta Numerica*, *28*, 287–404.

Li, N., & Marden, J. R. (2013). Designing games for distributed optimization. *IEEE Journal of Selected Topics in Signal Processing*, *7*(2), 230–242.

Li, Y., Tang, Y., Zhang, R., & Li, N. (2021). Distributed reinforcement learning for decentralized linear quadratic control: A derivative-free policy optimization approach. *IEEE Transactions on Automatic Control*.

Lian, X., Huang, Y., Li, Y., & Liu, J. (2015). Asynchronous parallel stochastic gradient for nonconvex optimization. In *Advances in neural information processing systems*: *Vol. 28*.

Lian, X., Zhang, H., Hsieh, C.-J., Huang, Y., & Liu, J. (2016). A comprehensive linear speedup analysis for asynchronous stochastic parallel optimization from zeroth-order to first-order. In *Advances in neural information processing systems*: *Vol. 29*.

Liu, J., & Wright, S. J. (2015). Asynchronous stochastic coordinate descent: Parallelism and convergence properties. *SIAM Journal on Optimization*, *25*(1), 351–376.

Marden, J. R., Ruben, S. D., & Pao, L. Y. (2013). A model-free approach to wind farm control using game theoretic methods. *IEEE Transactions on Control Systems Technology*, *21*(4), 1207–1214.

Marden, J. R., Young, H. P., & Pao, L. Y. (2014). Achieving Pareto optimality through distributed learning. *SIAM Journal on Control and Optimization*, *52*(5), 2753–2770.

Menon, A., & Baras, J. S. (2013a). Convergence guarantees for a decentralized algorithm achieving Pareto optimality. In *Proceedings of the 2013 American control conference*.

Menon, A., & Baras, J. S. (2013b). A distributed learning algorithm with bit-valued communications for multi-agent welfare optimization. In *Proceedings of the 52nd IEEE conference on decision and control*.

Menon, A., & Baras, J. S. (2014). Collaborative extremum seeking for welfare optimization. In *Proceedings of the 53rd IEEE conference on decision and control*.

Nedić, A. (2011). Asynchronous broadcast-based convex optimization over a network. *IEEE Transactions on Automatic Control*, *56*(6), 1337–1351.

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). Cooperative distributed multi-agent optimization. In D. P. Palomar, & Y. C. Eldar (Eds.), *Convex optimization in signal processing and communications* (pp. 340–386). Cambridge University Press.

Nesterov, Y., & Spokoiny, V. (2017). Random gradient-free minimization of convex functions. *Foundations of Computational Mathematics*, *17*, 527–566.

Nisan, N., Roughgarden, T., Tardos, É., & Vazirani, V. V. (Eds.), (2007). *Algorithmic game theory*. Cambridge University Press.

Pu, S., & Nedić, A. (2021). Distributed stochastic gradient tracking methods. *Mathematical Programming*, *187*, 409–457.

Qu, G., & Li, N. (2018). Harnessing smoothness to accelerate distributed optimization. *IEEE Transactions on Control of Network Systems*, *5*(3), 1245–1260.

Sahu, A. K., Jakovetic, D., Bajovic, D., & Kar, S. (2018). Distributed zeroth order optimization over random networks: A Kiefer-Wolfowitz stochastic approximation approach. In *Proceedings of the 57th IEEE conference on decision and control*.

Shamir, O. (2017). An optimal algorithm for bandit and zero-order convex optimization with two-point feedback. *Journal of Machine Learning Research*, *18*(52), 1–11.

Shi, W., Ling, Q., Wu, G., & Yin, W. (2015). EXTRA: An exact first-order algorithm for decentralized consensus optimization. *SIAM Journal on Optimization*, *25*(2), 944–966.

Tang, Y., Ren, Z., & Li, N. (2020a). Zeroth-order feedback optimization for cooperative multi-agent systems. In *Proceedings of the 59th IEEE conference on decision and control*.

Tang, Yujie, Ren, Zhaolin, & Li, Na (2020b). Zeroth-order feedback optimization for cooperative multi-agent systems. arXiv preprint arXiv:2011.09728.
Tang, Y., Zhang, J., & Li, N. (2021). Distributed zero-order algorithms for nonconvex multi-agent optimization. *IEEE Transactions on Control of Network Systems*, *8*(1), 269–281.
Torczon, V. (1997). On the convergence of pattern search algorithms. *SIAM Journal on Optimization*, *7*(1), 1–25.
Yu, Z., Ho, D. W. C., & Yuan, D. (2022). Distributed randomized gradient-free mirror descent algorithm for constrained optimization. *IEEE Transactions on Automatic Control*, *67*(2), 957–964.
Zhang, R., & Kwok, J. (2014). Asynchronous distributed ADMM for consensus optimization. In *Proceedings of the 31st international conference on machine learning*.

Yujie Tang received his bachelor's degree in Electronic Engineering from Tsinghua University in 2013 and his Ph.D. degree in Electrical Engineering from the California Institute of Technology in 2019. He is currently an Assistant Professor in the Department of Industrial Engineering and Management at Peking University. Before joining Peking University, he was a postdoctoral fellow in the School of Engineering and Applied Sciences at Harvard University. His research interests lie in distributed optimization, control and learning and their applications in cyber–physical systems.

Zhaolin Ren received a Bachelor's degree in Mathematics and a Master's degree in Statistics from Stanford University in 2019. He is currently a Ph.D. student at Harvard University, working under the supervision of Prof. Na Li. His main research interests include zeroth-order optimization, control theory, and reinforcement learning.

Na Li received a bachelor's degree in mathematics from Zhejiang University in 2007 and a Ph.D. degree in control and dynamical systems from the California Institute of Technology in 2013. She is a Gordon McKay Professor of Electrical Engineering and Applied Mathematics at Harvard University. She was a Postdoctoral Associate with Massachusetts Institute of Technology from 2013 to 2014. Her research lies in the control, learning, and optimization of networked systems, including theory development, algorithm design, and applications to real-world cyber–physical societal systems. She received the NSF Career Award, the AFSOR Young Investigator Award, the ONR Young Investigator Award, the Donald P. Eckman Award, the McDonald Mentoring Award, the IFAC Thoma Manfred Medal, along with some other awards.