Stochastic Dual Averaging for Decentralized Online Optimization on Time-Varying Communication Graphs

Soomin Lee, Angelia Nedić, and Maxim Raginsky

***Abstract*—We consider a decentralized online convex optimization problem in a network of agents, where each agent controls only a coordinate (or a part) of the global decision vector. For such a problem, we propose two decentralized stochastic variants (SODA-C and SODA-PS) of Nesterov's dual averaging method (DA), where each agent only uses a coordinate of the noise-corrupted gradient in the dual-averaging step. We show that the expected regret bounds for both algorithms have sublinear growth of $O(\sqrt{T})$, with the time horizon T, in scenarios when the underlying communication topology is time-varying. The sublinear regret can be obtained when the stepsize is of the form $1/\sqrt{t}$ and the objective functions are Lipschitz-continuous convex functions with Lipschitz gradients, and the variance of the noisy gradients is bounded. We also provide simulation results of the proposed algorithms on sensor networks to complement our theoretical analysis.**

***Index Terms*—Decentralized optimization, online optimization, stochastic dual-averaging method, pseudo-regret.**

I. Introduction

Decentralized optimization has recently been receiving a significant attention due to the emergence of large-scale distributed algorithms in machine learning, signal processing, and control applications for wireless communication networks, power networks, and sensor networks; see, for example, [1]–[6]. A central generic problem in such applications is decentralized resource allocation for a multi-agent system, where the network agents collectively solve an optimization problem in the absence of full knowledge about the overall problem structure. The agents are allowed to communicate with immediate neighbors, which enables them to learn the information needed for an efficient global resource allocation.

In recent literature on control and optimization, a decentralized resource allocation problem of the form where the objective function $f(\mathbf{x})$ is common to all agents but the resource allocation vector $\mathbf{x}$ is distributed among the agents, i.e., $\mathbf{x} = (x_1, \ldots, x_n)$ and each agent i is responsible for maintaining and updating only a coordinate (or a part) x_i of the whole vector $\mathbf{x}$, has been studied in [7]–[12] (see also the textbook [13]). Recently, Li and Marden [14] have proposed a circulation-based algorithm for this type of problems, where each agent i keeps estimates for the variables x_j, $j \neq i$, that are controlled by all the other agents in the network. The convergence of this algorithm relies on some contraction properties of the iteration dynamics.

Our work in this technical note is motivated by the ideas of Li and Marden [14] and also by the broadcast-based subgradient push, which was originally developed by Kempe *et. al.* [15], later extended in [16] (see also [17], [18]) for distributed optimization. Specifically, we employ the circulation-based protocol for undirected communications in [14] and also the broadcast-based push-sum protocol for directed and weight-imbalanced communications in [15]–[18] to propose online decentralized stochastic approximation variants of Nesterov's dual-averaging algorithm (DA) [19]. We call these algorithms SODA-C (Stochastic Online Dual Averaging with Circulation-based communication) and SODA-PS (Stochastic Online Dual Averaging with Push-Sum communication), respectively. We provide expected regret bounds in terms of the global resource allocation vector $\mathbf{x}$, when the signals received by the agents are subject to random noise. For both SODA-C and SODA-PS, we show that the expected regret has sublinear growth of order $O(\sqrt{T})$ in time T with the stepsize of the form $1/\sqrt{t+1}$.

This technical note is an extension of the work in [20], where ODA-C and ODA-PS were studied under deterministic signal models. In addition, ODA-C was designed for a time-invariant and connected communication graph, which is not practical for a wide variety of applications. In contrast with [20]: (1) We assume the gradients are evaluated with some random error, and these stochastic gradients are used for dual averaging instead of the exact gradients. This is motivated by many applications including distributed learning and recursive regression over networks. The extension is important yet nontrivial, as the stochastic gradient errors made by each agent propagate through the network to every other agent and also across time, making the iterates statistically dependent across time and agents. (2) Also, we propose an extension of ODA-C that allows relaxations in the structure of the communication graph.

In recent literature on decentralized online optimization, an extensively studied problem is one where the time-varying system objective function $f_t(\mathbf{x})$ decomposes into a sum of local objective functions, i.e., $f_t(\mathbf{x}) = \sum_{i=1}^{n} f_{i,t}(\mathbf{x})$ where $f_{i,t}$ is known only to agent i; see, for example [21]–[24]. In contrast with these existing methods, our algorithms have the following distinctive features, which essentially calls for different analysis techniques: (1) We consider an online convex optimization problem with a *nondecomposable* system objective, which is a function of a distributed resource allocation vector. (2) Each agent maintains and updates its private estimate of the best global allocation vector at each time, but contributes only one coordinate (or a coordinate block) to the network-wide decision vector. (3) We provide regret bounds in terms of the true global

Manuscript received June 5, 2016; revised November 16, 2016; accepted January 3, 2017. Date of publication January 9, 2017; date of current version December 1, 2017. The work has been partially supported by the National Science Foundation under Grant CCF 11-11342 and by the Office of Naval Research under Grant N00014-12-1-0998. Recommended by Associate Editor M. K. Camlibel.

S. Lee is with the ISyE, Georgia Institute of Technology, Atlanta, GA USA 30318 (e-mail: soomin.lee@isye.gatech.edu).

A. Nedić is with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ USA 85287-5706 (e-mail: angelia.nedich@asu.edu).

M. Raginsky is with the Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL USA 61801 (e-mail: maxim@illinois.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2017.2650563

resource allocation vector $\mathbf{x}$ (rather than some estimate of $\mathbf{x}$ by a single agent).

Notation: All vectors are column vectors. We use boldface to distinguish between vectors in $\mathbb{R}^n$ from scalars. For example, $\mathbf{z}_i(t)$ is a vector while $x_i(t)$ is a scalar. We will work with the Euclidean norm, denoted by $\|\cdot\|$. We will use $\mathbf{e}_1, \ldots, \mathbf{e}_n$ to denote the unit vectors in the standard Euclidean basis of $\mathbb{R}^n$. We use δ_i^k to denote the Kronecker delta symbol, i.e., $\delta_i^k = 1$ if $i = k$ and $\delta_i^k = 0$, otherwise. We use $\mathbf{1}$ to denote a vector with all entries equal to 1. For any $n \geq 1$, the set of integers $\{1, \ldots, n\}$ is denoted by $[n]$. We use $\sigma_2(A)$ to denote the second largest singular value of a matrix A. We use $\mathbb{E}[Z]$ to denote the expectation of a random variable Z. We denote the products of the time-varying matrices $A(t), \ldots, A(s)$ by $A(t:s)$, i.e., $A(t:s) \triangleq A(t) \cdots A(s)$, for all $t \geq s \geq 0$. Also, we let $A(t-1:t) \triangleq I$, for all $t \geq 1$.

II. Problem Formulation

Consider a multi-agent system (network) consisting of n agents, indexed by elements of the set $\mathcal{V} = [n]$. At each time t, each agent $i \in \mathcal{V}$ takes an action $x_i(t)$ in an action space X, which is a closed and bounded interval of the real line,[1] based only on the information available before time t. Let us denote the *network action* at time t by

$$\mathbf{x}(t) = (x_1(t), \ldots, x_n(t)) \in \mathsf{X}^n. \tag{1}$$

Then, each agent receives a random signal which depends on the time-varying cost function f_t, and the multi-agent system incurs a cost of $f_t(\mathbf{x}(t))$. We assume that the functions f_t come from a fixed class $\mathcal{F}$ of convex functions $f : \mathsf{X}^n \to \mathbb{R}$. Moreover, the environment may be adaptive, i.e., the choice of the function f_t may depend on all of the data generated by the network up to time t.

The communication among agents in the network is governed by either one of the two following graphs:

1) Time-varying undirected connected graphs $\mathcal{G}_1(t) = (\mathcal{V}, \mathcal{E}(t))$: If agents i and j are connected by an edge (which we denote by $i \leftrightarrow j$), then they may exchange information with one another. Thus, the neighborhood of each agent $i \in \mathcal{V}$ is defined as $\mathcal{N}_i(t) \triangleq \{j \in \mathcal{V} : i \leftrightarrow j \in \mathcal{E}(t)\} \cup \{i\}$.
2) Time-varying B-strongly connected digraphs $\mathcal{G}_2(t) = (\mathcal{V}, \mathcal{E}(t))$: If there exists a directed link from agent j to i at time t (which we denote by (j,i)), agent j may send its information to agent i. We use the notation $\mathcal{N}_i^{\text{in}}(t)$ and $\mathcal{N}_i^{\text{out}}(t)$ to denote the in and out neighbors of agent i at time t, respectively. That is, $\mathcal{N}_i^{\text{in}}(t) \triangleq \{j \mid (j,i) \in \mathcal{E}(t)\} \cup \{i\}$, and $\mathcal{N}_i^{\text{out}}(t) \triangleq \{j \mid (i,j) \in \mathcal{E}(t)\} \cup \{i\}$. Also, we use $d_i(t)$ to denote the out-degree of node i at time t, i.e., $d_i(t) \triangleq |\mathcal{N}_i^{\text{out}}(t)|$. We assume B-strong connectivity of the graphs $\mathcal{G}_2(t)$, where $B > 0$ is an integer. Namely, $(\mathcal{V}, \mathcal{E}_B(t))$ with $\mathcal{E}_B(t) = \bigcup_{i=(t-1)B+1}^{tB} \mathcal{E}(i)$ is strongly connected for every $t \geq 1$. In other words, the union of the edges appearing for B consecutive time instances periodically constructs a strongly connected graph. This assumption is required to ensure that there exists a path from one node to every other node within any bounded interval of length B even if the underlying network topology is time-varying. Overall, the network interacts with an environment according to the protocol shown in BOX I.

Since the vector sequence $\{\mathbf{x}(t)\}_{t=1}^{\infty}$ [cf. (1)] is generated based on noisy information, all elements in $\{\mathbf{x}(t)\}_{t=1}^{\infty}$ are random variables. For the same reason, the functions f_t selected by the environment are also random. Hence, the notion of the classical *regret* that is often used in online optimization needs to be modified accordingly. We will consider

[1]Everything easily generalizes to X being a compact convex subset of a multidimensional space $\mathbb{R}^d$; we work with the scalar case mainly for simplicity.

Box I:
Online Optimization Protocol

Parameters: base action space X; network graph $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$; function class $\mathcal{F}$
For each round $t = 1, 2, \ldots$:
1) Each agent $i \in \mathcal{V}$ selects an action $x_i(t) \in \mathsf{X}$
2) Each agent $i \in \mathcal{V}$ exchanges local information with its neighbors $\mathcal{N}_i(t)$
3) The environment selects the current objective $f_t \in \mathcal{F}$, and all agents receive noisy signals about f_t

the following quantity at an arbitrary time horizon $T \geq 1$

$$\bar{R}(T) \triangleq \mathbb{E}\left[\sum_{t=1}^{T} f_t(\mathbf{x}(t))\right] - \inf_{\mathbf{y} \in \mathsf{X}^n} \mathbb{E}\left[\sum_{t=1}^{T} f_t(\mathbf{y})\right]. \tag{2}$$

Adopting the terminology from the literature on bandit problems [25], we refer to $\bar{R}(T)$ as the *pseudo-regret* at time horizon T.[2] Thus, $\bar{R}(T)$ is the difference between the expectation of the total cost incurred by the network at time T and the smallest expected total cost that could have been achieved with a single action in X^n in hindsight (i.e., with perfect advance knowledge of the sequence $f_1, \ldots, f_T$), with perfect knowledge of the distributions of any random signals, and without any restriction on the communication between the agents.

Our goal is to design a policy that each agent $i \in \mathcal{V}$ should use to determine its action $x_i(t)$ based on the local information available to it until time t under the influence of the noise, such that the pseudo-regret in (2) is (a) sublinear as a function of the time horizon T and (b) exhibits "reasonable" dependence on the number of agents n and on the topology of the communication graphs.

III. Proposed Algorithms

Our proposed methods are based on Nesterov's dual averaging algorithm (DA) [19], originally designed for centralized minimization of a convex function over a convex set. We first introduce some notations and briefly introduce the DA algorithm. Then we present two stochastic variants of DA, namely SODA-C and SODA-PS, for decentralized online optimization.

The two algorithms are similar in that they both use DA as their optimization primitive, and each agent i only uses the i-th coordinate of the noisy gradient evaluated at its local primal estimate. However, they use different communication frameworks, i.e., SODA-C uses the circulation-based method over a sequence of time-varying undirected graphs $\{\mathcal{G}_1(t)\}_{t=0}^{\infty}$ and SODA-PS uses the broadcast-based push-sum method over a sequence of time-varying B-strongly connected digraphs $\{\mathcal{G}_2(t)\}_{t=0}^{\infty}$.

To be of practical interest, distributed optimization algorithms need to accommodate the constraints imposed by the underlying networked systems. Some systems allow only bidirectional (undirected) communications while some other systems also allow unidirectional (directed) communications. Unidirectional communication is usually considered to be more desirable, as in the bidirectional case deadlocks can occur when nodes block their communications while waiting for a response. Therefore, we choose either SODA-C or SODA-PS based on the structure of the networked system.

[2]Another frequenty used notion is that of an *expected regret*, which differs from the quantity in (2) by having the infimum over $\mathbf{y} \in \mathsf{X}^n$ *inside* the expectation. Pseudo-regret is always upper-bounded by the expected regret, as a consequence of Jensen's inequality.

A. DA for Centralized Minimization

DA is based on a nonnegative *proximal function* $\psi : \mathsf{Y} \to \mathbb{R}^+$, where Y is an arbitrary closed convex set in $\mathbb{R}^n$. The function ψ is assumed to be 1-strongly convex with respect to the Euclidean norm $\|\cdot\|$, i.e., for any $\mathbf{x}, \mathbf{y} \in \mathsf{Y}$ we have

$$\psi(\mathbf{y}) \geq \psi(\mathbf{x}) + \langle \bar{\nabla}\psi(\mathbf{x}), \mathbf{y} - \mathbf{x}\rangle + \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|^2 \quad (3)$$

where $\bar{\nabla}\psi$ denotes an arbitrary subgradient of ψ. In addition to this, we assume $\psi(0) = 0$. An example of such proximal functions includes the quadratic function $\psi(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$. We also define the *mapping* $\Pi^{\psi}_{\mathsf{Y}} : \mathbb{R}^n \times (0, \infty) \to \mathsf{Y}$ by

$$\Pi^{\psi}_{\mathsf{Y}}(\mathbf{z}, \alpha) \triangleq \underset{\mathbf{x}\in\mathsf{Y}}{\arg\min} \left\{ \langle \mathbf{z}, \mathbf{x}\rangle + \frac{1}{\alpha}\psi(\mathbf{x}) \right\} \quad (4)$$

which can be seen as a type of projection of $\mathbf{z}$ onto Y.

DA generates two iterate sequences $\{\boldsymbol{\xi}(t), \mathbf{z}(t)\}_{t=0}^{\infty} \subseteq \mathsf{Y} \times \mathbb{R}^n$ according to the following steps:

$$\mathbf{z}(t+1) = \mathbf{z}(t) + \mathbf{g}(t), \quad (5)$$

$$\boldsymbol{\xi}(t+1) = \Pi^{\psi}_{\mathsf{Y}}(\mathbf{z}(t+1), \alpha(t)), \quad (6)$$

where $\mathbf{g}(t)$ is a subgradient of f at $\boldsymbol{\xi}(t)$, i.e., $\mathbf{g}(t) \in \partial f(\boldsymbol{\xi}(t))$, and $\{\alpha(t)\}_{t=0}^{\infty}$ is a nonincreasing sequence of positive stepsizes. Roughly speaking, DA improves the intrinsic weakness of first-order subgradient methods for nonsmooth functions by maintaining an additional sequence $\{\mathbf{z}(t)\}_{t=0}^{\infty}$ in dual space with better averaging schemes. (See the original paper [19] for more details.)

B. SODA-C

Consider an $n \times n$ nonnegative row-stochastic weight matrix $M(t)$, i.e., for all $t \geq 0$, $\sum_{j=1}^n [M(t)]_{ij} = 1$, $i \in [n]$. In addition to this, $M(t)$ respects the topology of $\mathcal{G}_1(t)$, i.e., $[M(t)]_{ij} \geq \eta \in (0,1)$ if $i \neq j$ and $i \leftrightarrow j \in \mathcal{E}(t)$, $[M(t)]_{ii} = 1 - \sum_{j\in\mathcal{N}_i(t)\setminus\{i\}}[M(t)]_{ij}$, and $[M(t)]_{ij} = 0$, otherwise. Here, the property of η-*nondegeneracy*, which assumes a lower bound of η on all positive entries of $M(t)$ for all $t \geq 0$, is just required for convergence of the algorithm, and the agents need not know the exact value of η. For example, η can be just set to $\eta = 1/\max_{i\in\mathcal{V}}|\mathcal{N}_i(t)|$, where the maximum degree $\max_{i\in\mathcal{V}}|\mathcal{N}_i(t)|$ can be trivially upper estimated by n.

In Algorithm 1, we summarize our proposed stochastic online dual-averaging circulation-based algorithm. We assume that each agent $i \in \mathcal{V}$ generates a sequence $\{\boldsymbol{\xi}_i(t), \mathbf{z}_i(t)\}_{t=1}^{\infty} \subseteq \mathsf{X}^n \times \mathbb{R}^n$, where the primal and dual iterates

$$\boldsymbol{\xi}_i(t) = (\xi_i^1(t), \ldots, \xi_i^n(t)), \quad \mathbf{z}_i(t) = (z_i^1(t), \ldots, z_i^n(t))$$

are initialized with an arbitrary $\boldsymbol{\xi}_i(0) \in \mathsf{X}^n$ and $\mathbf{z}_i(0) = 0$. They are updated recursively using (6a)–(6b), where δ_i^k is the Kronecker delta symbol, $\{\alpha(t)\}_{t=0}^{\infty}$ is a nonincreasing sequence of positive step sizes, and $u_i(t) \in \mathbb{R}$ is a local update performed by agent i at time t, which is computed with stochastic errors. The action of agent i at time t is given by $x_i(t) = \xi_i^i(t)$, i.e., the ith coordinate of the vector $\boldsymbol{\xi}_i(t)$ in (6b) is set to be the action of agent i at time t. The dual update rule (6a) is inspired by the state dynamics proposed by Li and Marden [14], whereas the primal update rule (6b) is exactly what one has in DA [cf. (5b)].

C. SODA-PS

In Algorithm 2, we summarize our proposed stochastic online dual-averaging push-sum based algorithm. Here, each agent i maintains an additional scalar sequence $\{w_i(t)\}_{t=1}^{\infty} \subseteq \mathbb{R}$. Then, the three sequences $\{\mathbf{z}_i(t), \boldsymbol{\xi}_i(t), w_i(t)\} \subseteq \mathsf{X}^n \times \mathbb{R}^n \times \mathbb{R}$ are updated recursively using (7a)–(7c), where the weight matrix $A(t)$ is defined by the out-degrees of the agents at time t, i.e.,

$$[A(t)]_{ij} = \begin{cases} 1/d_j(t) & \text{whenever } j \in \mathcal{N}_i^{\text{in}}(t) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

which is column stochastic by construction.

Note that the update rules in Algorithm 2 are based on a simple broadcast communication. Each agent i broadcasts (or *pushes*) the quantities $w_i(t)/d_i(t)$ and $\mathbf{z}_i(t)/d_i(t)$ to all of the nodes in its out-neighborhood $\mathcal{N}_i^{\text{out}}(t)$. Then, in (7a)–(7b) each agent simply *sums* all the received messages to obtain $w_i(t+1)$ and $\mathbf{z}_i(t+1)$. The update rule (7c)–(7d) can be executed locally. Unlike SODA-C, the averaging matrix $A(t)$ in SODA-PS does not require symmetry due to this broadcast-based nature of the push-sum protocol.

Algorithm 1: SODA-C.

Require: Set $T \geq 1$ and $t = 0$. For every $i \in \mathcal{V}$, locally initialize $\boldsymbol{\xi}_i(0) \in \mathsf{X}^n$ and $\mathbf{z}_i(0) = 0$.

1: If $t = T$, then stop. Otherwise, set for every $i \in \mathcal{V}$

$$z_i^k(t+1) = z_i^k(t) + n\delta_i^k u_i(t) + \sum_{j=1}^n [M(t)]_{ij}\left(z_j^k(t) - z_i^k(t)\right), \; k \in [n] \quad (6a)$$

$$\boldsymbol{\xi}_i(t+1) = \Pi^{\psi}_{\mathsf{X}^n}(\mathbf{z}_i(t+1), \alpha(t)), \quad (6b)$$

$$x_i(t+1) = \xi_i^i(t+1). \quad (6c)$$

2: Increase t by one and return to Step 1.

Algorithm 2: SODA-PS.

Require: Set $T \geq 1$ and $t = 0$. For every $i \in \mathcal{V}$, locally initialize $w_i(0) = 1$, $\boldsymbol{\xi}_i(0) \in \mathsf{X}^n$ and $\mathbf{z}_i(0) = 0$.

1: If $t = T$, then stop. Otherwise, set for every $i \in \mathcal{V}$:

$$w_i(t+1) = \sum_{j=1}^n [A(t)]_{ij} w_j(t) \quad (7a)$$

$$z_i^k(t+1) = n\delta_i^k u_i(t) + \sum_{j=1}^n [A(t)]_{ij} z_j^k(t), \; k \in [n] \quad (7b)$$

$$\boldsymbol{\xi}_i(t+1) = \Pi^{\psi}_{\mathsf{X}^n}\left(\frac{\mathbf{z}_i(t+1)}{w_i(t+1)}, \alpha(t)\right) \quad (7c)$$

$$x_i(t+1) = \xi_i^i(t+1). \quad (7d)$$

2: Increase t by one and return to Step 1.

D. Feedback Policies $u_i(t)$

To complete the description of the algorithms, we must specify the feedback policies $u_i(t)$ in (6a) and (7b). We assume that at each time t, each agent $i \in \mathcal{V}$ receives a random signal $\widehat{g}_i(t)$ about the ith coordinate of the gradient of f_t at the agent's primal variable $\boldsymbol{\xi}_i(t)$. We emphasize that this signal model is of practical interest for decentralized systems as the agents need not receive signals about the entire gradient vector $\nabla f_t(\mathbf{x}(t))$.

Then, both SODA-C and SODA-PS feed this signal back into the dynamics (6a) and (7b) by letting

$$u_i(t) = \widehat{g}_i(t),\ \forall i \in [n],\ t \geq 0. \tag{9}$$

Note that the signals issued to different agents need not be independent.

Let $\mathscr{F}_t$ denote the σ-field generated by all the information up to time t (this includes all random signals related to $f_1, \ldots, f_t$, all iterates generated as well as all local messages exchanged in the network up to time t). Note that the actions determined by all agents at time t may only depend on $\mathscr{F}_{t-1}$.

In what follows, we assume that the stochastic gradient signals are well-behaved, i.e., the signals are unbiased and have finite second moments.

Assumption 1: The random signal $\widehat{g}_i(t)$ for $i \in \mathcal{V}$ and $t \geq 0$ is such that almost surely

$$\mathbb{E}\left[\widehat{g}_i(t)|\mathscr{F}_{t-1}\right] = \langle \nabla f_t(\boldsymbol{\xi}_i(t)), \mathbf{e}_i\rangle,\ \mathbb{E}[|\widehat{g}_i(t)|^2|\mathscr{F}_{t-1}] \leq \bar{L}^2$$

where $\bar{L} > 0$ is some constant.

This signal model is common in the literature on stochastic optimization (see, e.g., [21], [26]). Note that if all gradients are uniformly bounded, i.e., $\|\nabla f_t(\boldsymbol{\xi}_i(t))\| \leq L$ for all $i \in \mathcal{V}$ and $t \geq 0$, and $\mathbb{E}\left[|\widehat{g}_i(t) - \langle \nabla f_t(\boldsymbol{\xi}_i(t)), \mathbf{e}_i\rangle|^2|\mathscr{F}_{t-1}\right] \leq \nu^2$ holds almost surely for some $\nu > 0$, then Assumption 1 is satisfied with $\bar{L} = L^2 + \nu^2$.

IV. The Generic Regret Bound

In this section, we present a generic regret bound under the following additional assumptions. This generic bound will be later instantiated by SODA-C and SODA-PS in Sections V and VI for their full regret analysis.

Assumption 2: The proximal function ψ is bounded on X^n, i.e., there is a scalar $C > 0$ such that $\psi(\mathbf{y}) \leq C$ for all $\mathbf{y} \in \mathsf{X}^n$.

Assumption 3: All functions $f \in \mathcal{F}$ are Lipschitz continuous with a constant L

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq L\|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathsf{X}^n.$$

We also make the following assumption for later use although it is not required for the generic bound itself:

Assumption 4: All functions $f \in \mathcal{F}$ are differentiable and have Lipschitz continuous gradients with a constant G

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\| \leq G\|\mathbf{x} - \mathbf{y}\|, \quad \forall f \in \mathcal{F};\ \mathbf{x}, \mathbf{y} \in \mathsf{X}^n.$$

Theorem 1: Let $\{\mathbf{u}(t)\}_{t\geq 1} \subseteq \mathbb{R}^n$ be an arbitrary sequence of random vectors, and $\{\bar{\mathbf{x}}(t)\}_{t\geq 1} \subseteq \mathsf{X}^n$ be generated as

$$\bar{\mathbf{x}}(t+1) = \Pi^{\psi}_{\mathsf{X}^n}\left(\sum_{s=1}^{t}\mathbf{u}(s), \alpha(t)\right) \tag{10}$$

where $\{\alpha(t)\}_{t\geq 0}$ is a positive and nonincreasing sequence. In addition, let $\{\mathbf{g}(t)\}_{t\geq 1} \subseteq \mathbb{R}^n$ be an arbitrary sequence of deterministic vectors. Then, under Assumptions 2 and 3, the pseudo-regret $\bar{R}(T)$ in (2) can be upper-bounded in terms of $\mathbf{u}(t)$, $\bar{\mathbf{x}}(t)$ and $\mathbf{g}(t)$ as follows: for all $T \geq 1$

$$\begin{aligned}\bar{R}(T) \leq{}& \frac{1}{2}\sum_{t=1}^{T}\alpha(t-1)\mathbb{E}\|\mathbf{u}(t)\|^2 + \frac{C}{\alpha(T)}\\ &+ L\sum_{t=1}^{T}\mathbb{E}\|\mathbf{x}(t) - \bar{\mathbf{x}}(t)\| + \sqrt{n}D_{\mathsf{X}}\sum_{t=1}^{T}\mathbb{E}\|\nabla f_t(\bar{\mathbf{x}}(t)) - \mathbf{g}(t)\|\\ &+ \sup\nolimits_{\mathbf{y}\in\mathsf{X}^n}\sum_{t=1}^{T}\mathbb{E}\left[\langle \mathbf{g}(t) - \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\right]\end{aligned}$$

where $D_{\mathsf{X}} \triangleq \sup_{x,y\in\mathsf{X}}|x - y|$ is the diameter of the set X.[3]

Proof: For any t and any $\mathbf{y} \in \mathsf{X}^n$ we can write

$$\begin{aligned}&f_t(\mathbf{x}(t)) - f_t(\mathbf{y})\\ &= f_t(\mathbf{x}(t)) - f_t(\bar{\mathbf{x}}(t)) + f_t(\bar{\mathbf{x}}(t)) - f_t(\mathbf{y})\\ &\leq \langle \nabla f_t(\mathbf{x}(t)), \mathbf{x}(t) - \bar{\mathbf{x}}(t)\rangle + \langle \nabla f_t(\bar{\mathbf{x}}(t)), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\\ &\leq L\|\mathbf{x}(t) - \bar{\mathbf{x}}(t)\| + \langle \nabla f_t(\bar{\mathbf{x}}(t)), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\end{aligned} \tag{11}$$

where the second step follows from convexity of f_t, while the last step uses the fact that all $f \in \mathcal{F}$ are L-Lipschitz. The second term in (11) can be further expanded as

$$\begin{aligned}&\langle \nabla f_t(\bar{\mathbf{x}}(t)), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\\ &= \langle \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle + \langle \nabla f_t(\bar{\mathbf{x}}(t)) - \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle.\end{aligned} \tag{12}$$

Now, from relation (10) we obtain

$$\bar{\mathbf{x}}(t+1) = \arg\min_{\mathbf{x}\in\mathsf{X}^n}\left\{\sum_{s=1}^{t}\langle \mathbf{u}(s), \mathbf{x}\rangle + \frac{1}{\alpha(t)}\psi(\mathbf{x})\right\}.$$

Therefore, by [21, Lemma 3], we can write

$$\sum_{t=1}^{T}\langle \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle \leq \frac{1}{2}\sum_{t=1}^{T}\alpha(t-1)\|\mathbf{u}(t)\|^2 + \frac{\psi(\mathbf{y})}{\alpha(T)}. \tag{13}$$

Then, for an arbitrary vector $\mathbf{g}(t)$, the second term on the right-hand side of (12) can be rewritten as

$$\begin{aligned}&\langle \nabla f_t(\bar{\mathbf{x}}(t)) - \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\\ &= \langle \nabla f_t(\bar{\mathbf{x}}(t)) - \mathbf{g}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle + \langle \mathbf{g}(t) - \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\\ &\leq \|\nabla f_t(\bar{\mathbf{x}}(t)) - \mathbf{g}(t)\|\|\bar{\mathbf{x}}(t) - \mathbf{y}\| + \langle \mathbf{g}(t) - \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\\ &\leq \sqrt{n}D_{\mathsf{X}}\|\nabla f_t(\bar{\mathbf{x}}(t)) - \mathbf{g}(t)\| + \langle \mathbf{g}(t) - \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\end{aligned} \tag{14}$$

where the last step follows from the fact that $\bar{\mathbf{x}}(t), \mathbf{y} \in \mathsf{X}^n$ and the set X is bounded. Substituting the estimates (12)–(14) into (11), taking the total expectation, and then taking the supremum over the choice of $\mathbf{y} \in \mathsf{X}^n$, we get the desired result. ∎

V. Regret Analysis of SODA-C

We now show that the pseudo-regret of SODA-C grows sublinearly in T. Specifically, in Section V-A, we provide a lemma to describe the network-wide error due to the local primal-dual estimation using the algorithm definition in (6a)–(6c). In Section V-B, we provide full regret analysis by instantiating Theorem 1 using this lemma.

[3] This relation holds for any realization of any of the random variables, not just in expectation. We write this in expectation for later convenience.

A. Network Error of SODA-C

In the following lemma, we represent the recursive dynamics of the dual iterates $\mathbf{z}_i(t)$ and their mean field $\bar{\mathbf{z}}(t)$. Note that $\bar{\mathbf{z}}(t)$ is a fictitious quantity whose computation requires all the distributed information, and is only needed for the convergence analysis.

Lemma 1: Let $\mathbf{u}(t)$ be obtained by stacking up the individual feedback policies, i.e., $\mathbf{u}(t) = \big(u_1(t), \ldots, u_n(t)\big)^T$.

(a) The weighted sum $\bar{\mathbf{z}}(t) \triangleq \frac{1}{n}\sum_{i=1}^n \mathbf{z}_i(t)$ evolves according to the linear dynamics

$$\bar{\mathbf{z}}(t+1) = \bar{\mathbf{z}}(t) + \mathbf{u}(t). \tag{15}$$

(b) For any $i, k \in [n]$, the iterates in (7b) evolve according to the following dynamics

$$z_i^k(t) = n\sum_{s=0}^{t-1}[M(t-1:s+1)]_{ik}u_k(s). \tag{16}$$

Proof: (a) Let $V^k(t)$ denote the $n \times n$ matrix with entries $[V^k(t)]_{ij} = z_j^k(t) - z_i^k(t)$. Then

$$\begin{aligned}\bar{z}^k(t+1) &= \frac{1}{n}\sum_{i=1}^n z_i^k(t+1)\\ &= \frac{1}{n}\sum_{i=1}^n\left\{z_i^k(t) + n\delta_i^k u_i(t) + \sum_{j=1}^n [M(t)]_{ij}[V^k(t)]_{ij}\right\}\\ &= \bar{z}^k(t) + u_k(t) + \mathrm{tr}[\tilde{M}(t)V^k(t)]\end{aligned}$$

where $\tilde{M}(t)$ is an $n \times n$ matrix with entries $[\tilde{M}(t)]_{ij} = \frac{1}{n}[M(t)]_{ij}$. Since $\tilde{M}(t)$ is a symmetric matrix, and $V^k(t)$ is skew-symmetric, $\mathrm{tr}[\tilde{M}(t)V^k(t)] = 0$, so we obtain (15).

(b) Let $\mathbf{r}_k(t)$ be obtained by stacking up the kth coordinate of $\mathbf{z}_i(t)$'s, i.e., $\mathbf{r}_k(t) = (z_1^k(t), \ldots, z_n^k(t))^T$. Then, by stacking up the dynamics (6a) over i, we obtain for all $k \in [n]$ and $t \ge 0$

$$\mathbf{r}_k(t+1) = M(t)\mathbf{r}_k(t) + nu_k(t)\mathbf{e}_k. \tag{17}$$

By solving this from time 0 to t and recalling that $\mathbf{z}_i(0) = 0$ for all $i \in \mathcal{V}$ (hence, $\mathbf{r}_k(0) = 0$ for all $k \in [n]$), we obtain for any $t \ge 1$

$$\begin{aligned}\mathbf{r}_k(t) &= M(t-1:0)\mathbf{r}_k(0) + n\sum_{s=0}^{t-1}u_k(s)M(t-1:s+1)\mathbf{e}_k\\ &= n\sum_{s=0}^{t-1}u_k(s)M(t-1:s+1)\mathbf{e}_k,\end{aligned}$$

where we used $M(t-1:t) = I$. We get the desired result by taking the i-th component of this vector. ■

In the next lemma, we show the network-wide disagreement at any t can be bounded by some constant.

Lemma 2: Under Assumptions 1 and 3, for the dynamics in (6a)–(6b) over any time-varying sequence of undirected connected graphs $\{\mathcal{G}_1(t)\}_{t\ge0}$, we have for all $t \ge 1$

$$\sum_{i=1}^n \mathbb{E}\|\mathbf{z}_i(t) - \bar{\mathbf{z}}(t)\|^2 \le \frac{4n^4\bar{L}^2}{\theta^2(1-\theta)^2}$$

where $\theta = 1 - \frac{\eta}{4n^2}$.

Proof: From the definitions of $\mathbf{z}_i(t)$ and $\bar{\mathbf{z}}(t)$, we have

$$\sum_{i=1}^n\|\mathbf{z}_i(t) - \bar{\mathbf{z}}(t)\|^2 = \sum_{i=1}^n\sum_{k=1}^n|z_i^k(t) - \bar{z}^k(t)|^2. \tag{18}$$

Thus, we upper-bound the quantity on the right-hand side.

From relation (15) (i.e., $\bar{z}^k(t) = n\sum_{s=0}^{t-1}\frac{1}{n}u_k(s)$) and (16), we obtain

$$\begin{aligned}z_i^k(t) - \bar{z}^k(t) &= n\sum_{s=0}^{t-1}\left([M(t-1:s+1)]_{ik} - \frac{1}{n}\right)u_k(s)\\ &= n\sum_{s=1}^{t-1}\left([M(t-1:s)]_{ik} - \frac{1}{n}\right)u_k(s-1)\\ &\quad + n\left([M(t-1:t)]_{ik} - \frac{1}{n}\right)u_k(t-1).\end{aligned}$$

From the result in [27] (Corollary 1 with $B = 1$), we have for any $i, k \in [n]$ and t, s and $t \ge s$

$$\left|[M(t:s)]_{ik} - \frac{1}{n}\right| \le \theta^{t-s-1}$$

where $\theta = 1 - \frac{\eta}{4n^2} < 1$. Therefore, we obtain $|z_i^k(t) - \bar{z}^k(t)| \le n\sum_{s=1}^{t-1}\theta^{t-s-2}|u_k(s-1)| + n|u_k(t-1)|$, where we used $M(t-1:t) = I$ for bounding the last term. Using the inequality $(a+b)^2 \le 2a^2 + 2b^2$, we get

$$|z_i^k(t) - \bar{z}^k(t)|^2 \le \frac{2n^2}{\theta^2(1-\theta)^2}\max_s|u_k(s-1)|^2 + 2n^2|u_k(t-1)|^2.$$

By taking the expectation and using $u_k(t) = \widehat{g}_k(t)$ (cf. (9)), we obtain

$$\mathbb{E}|z_i^k(t) - \bar{z}^k(t)|^2 \le \frac{4n^2\bar{L}^2}{\theta^2(1-\theta)^2}.$$

From this and relation (18), the desired result follows. ■

B. Pseudo-Regret of SODA-C

We now show that SODA-C achieves the pseudo-regret of order $O(\sqrt{T})$.

Theorem 2: Suppose that Assumptions 1–4 are in force. Then, the expected regret (2) of SODA-C over any time-varying sequence of undirected connected graphs $\{\mathcal{G}_1(t)\}_{t\ge0}$ with stochastic local gradient signals $\{(\widehat{g}_i(t), i \in \mathcal{V})\}_{t=1}^\infty$ and the step-size choice of $\alpha(t) = \frac{1}{\sqrt{t+1}}$ achieves the following pseudo-regret: for all $T \ge 1$

$$\bar{R}(T) \le \left[n\bar{L}^2 + C + \frac{2n^{5/2}\bar{L}}{\theta(1-\theta)}(L + \sqrt{n}GD_{\mathsf{X}})\right]\sqrt{T+1}.$$

Proof: We let $\bar{\mathbf{x}}(t) \triangleq \Pi_{\mathsf{X}^n}^{\psi}(\bar{\mathbf{z}}(t), \alpha(t-1))$. In order to invoke the generic regret bound in Theorem 1, in what follows, we verify that our dual update rule in (6a) satisfies relation (10)

$$\bar{\mathbf{x}}(t) = \Pi_{\mathsf{X}^n}^{\psi}(\bar{\mathbf{z}}(t), \alpha(t-1)) = \Pi_{\mathsf{X}^n}^{\psi}\left(\sum_{s=1}^{t-1}\mathbf{u}(s), \alpha(t-1)\right)$$

where the second equality holds from the recursive relation (15).

Recalling that $\mathbf{x}(t)$ is the network action vector (see (1)), we have the following for the third term in Theorem 1:

$$\|\mathbf{x}(t) - \bar{\mathbf{x}}(t)\| = \left\|\sum_{i=1}^n\big(x_i(t) - \bar{x}^i(t)\big)\mathbf{e}_i\right\| \le \sum_{i=1}^n\|\boldsymbol{\xi}_i(t) - \bar{\mathbf{x}}(t)\| \tag{19}$$

where the equality follows from the definition of $\mathbf{x}(t)$ in (1) and $\bar{\mathbf{x}}(t) = (\bar{x}^1(t), \ldots, \bar{x}^n(t))$, and the inequality follows from the dynamics (6c).

Let $\mathbf{g}(t) = (g_1(t), \ldots, g_n(t))$ be the vector with coordinates

$$g_i(t) = \mathbb{E}\left[\widehat{g}_i(t)|\mathscr{F}_t\right] \triangleq \langle\nabla f_t(\boldsymbol{\xi}_i(t)), \mathbf{e}_i\rangle, \quad i \in [n].$$

Then we can write

$$\|\nabla f_t(\bar{\mathbf{x}}(t)) - \mathbf{g}(t)\| = \left\|\sum_{i=1}^n \langle \nabla f_t(\bar{\mathbf{x}}(t)) - \nabla f_t(\boldsymbol{\xi}_i(t)), \mathbf{e}_i\rangle \mathbf{e}_i\right\|$$
$$\leq \sum_{i=1}^n \|\nabla f_t(\bar{\mathbf{x}}(t)) - \nabla f_t(\boldsymbol{\xi}_i(t))\| \leq G\sum_{i=1}^n \|\bar{\mathbf{x}}(t) - \boldsymbol{\xi}_i(t)\|$$

where we have exploited the fact that the gradients of all $f \in \mathcal{F}$ are G-Lipschitz. Now, by construction

$$\|\bar{\mathbf{x}}(t) - \boldsymbol{\xi}_i(t)\| = \left\|\Pi^{\psi}_{\mathsf{X}^n}(\bar{\mathbf{z}}(t), \alpha(t-1)) - \Pi^{\psi}_{\mathsf{X}^n}(\mathbf{z}_i(t), \alpha(t-1))\right\|$$
$$\leq \alpha(t-1)\|\bar{\mathbf{z}}(t) - \mathbf{z}_i(t)\|$$

where the last step follows from the fact that the map $\mathbf{z} \mapsto \Pi^{\psi}_{\mathsf{X}^n}(\mathbf{z}, \alpha)$ is α-Lipschitz (see, e.g., [19, Lemma 1]). Substituting these estimates into Theorem 1, we obtain

$$\bar{R}(T) \leq \frac{1}{2}\sum_{t=1}^T \alpha(t-1)\mathbb{E}\|\mathbf{u}(t)\|^2 + \frac{C}{\alpha(T)}$$
$$+ (L + \sqrt{n}GD_{\mathsf{X}})\sum_{t=1}^T \alpha(t-1)\sum_{i=1}^n \mathbb{E}\|\mathbf{z}_i(t) - \bar{\mathbf{z}}(t)\|$$
$$+ \sup_{\mathbf{y}\in\mathsf{X}^n}\sum_{t=1}^T \mathbb{E}\left[\langle \mathbf{g}(t) - \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\right]. \tag{20}$$

By the law of iterated expectations, we have

$$\mathbb{E}\|\mathbf{u}(t)\|^2 = \mathbb{E}\left[\sum_{i=1}^n \mathbb{E}\left[|\widehat{g}_i(t)|^2 | \mathscr{F}_t\right]\right] \leq n\bar{L}^2 \tag{21}$$

and, for any deterministic $\mathbf{y} \in \mathsf{X}^n$

$$\mathbb{E}\left[\langle \mathbf{g}(t) - \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\right]$$
$$= \mathbb{E}\left[\mathbb{E}\left[\langle \mathbf{g}(t) - \mathbf{u}(t), \bar{\mathbf{x}}(t) - \mathbf{y}\rangle | \mathscr{F}_t\right]\right]$$
$$= \mathbb{E}\left[\langle \mathbb{E}\left[\mathbf{g}(t) - \mathbf{u}(t)|\mathscr{F}_t\right], \bar{\mathbf{x}}(t) - \mathbf{y}\rangle\right] = 0 \tag{22}$$

where we have also used the fact that $\bar{\mathbf{x}}(t)$ is $\mathscr{F}_t$-measurable and $\mathbf{y}$ is deterministic. Lastly, for bounding the network disagreement term, we use Jensen's inequality to write

$$\sum_{i=1}^n \mathbb{E}\|\mathbf{z}_i(t) - \bar{\mathbf{z}}(t)\| \leq \sqrt{n}\sqrt{\sum_{i=1}^n \mathbb{E}\|\mathbf{z}_i(t) - \bar{\mathbf{z}}(t)\|^2}.$$

From the bound of Lemma 2, we further have

$$\sum_{i=1}^n \mathbb{E}\|\mathbf{z}_i(t) - \bar{\mathbf{z}}(t)\| \leq \frac{2n^{5/2}\bar{L}}{\theta(1-\theta)}. \tag{23}$$

Substituting estimates (21)–(23) into (20) and using the relation $\sum_{t=1}^T \alpha(t) \leq 2\sqrt{T}$, we obtain the claimed pseudo-regret bound. ■

VI. Regret Analysis of SODA-PS

We now show that the pseudo-regret of SODA-PS grows sublinearly. Proofs that are similar to those of SODA-C are omitted due to space limitations.

A. Network Error of SODA-PS

Note that Lemma 1 applies to SODA-PS as well (with $M(t)$ replaced by $A(t)$), although part (a) requires a slightly different proof as the dual update rule of SODA-PS in (7b) is different from that of SODA-C in (6a). We omit this proof due to space limitation.

Using Lemma 1, the next lemma provides an upper bound for the network disagreement term $\sum_{i=1}^n \mathbb{E}\|\frac{\mathbf{z}_i(t)}{w_i(t)} - \bar{\mathbf{z}}(t)\|^2$.

Lemma 3: Let the sequences $\{\mathbf{z}_i(t)\}$ and $\{w_i(t)\}$ be generated according to the algorithm (7a)–(7b) over any time-varying sequence of B-strongly connected digraphs $\{\mathcal{G}_2(t)\}_{t\geq 0}$. Recall that $\bar{\mathbf{z}}(t) = \frac{1}{n}\sum_{i=1}^n \mathbf{z}_i(t)$. Then, we have for all $t \geq 1$

$$\sum_{i=1}^n \mathbb{E}\left\|\frac{\mathbf{z}_i(t)}{w_i(t)} - \bar{\mathbf{z}}(t)\right\|^2 \leq n^2\left(\frac{2\beta\bar{L}}{\gamma\lambda(\lambda-1)}\right)^2$$

where

$$\gamma = \inf_{t\geq 0}\left(\min_{1\leq i\leq n}[A(t:0)\mathbf{1}]_i\right), \beta = 4 \text{ and } \lambda = (1 - 1/n^{nB})^{\frac{1}{B}}.$$

If in addition each $\mathcal{G}_2(t)$ is regular, we may choose

$$\beta = 2\sqrt{2}, \lambda = (1 - 1/4n^3)^{\frac{1}{B}} \text{ or } \beta = \sqrt{2}, \lambda = \max_{t\geq 0}\sigma_2(A(t))$$

whenever $\max_{t\geq 0}\sigma_2(A(t)) < 1$.

Proof: From the definitions of $\mathbf{z}_i(t)$ and $\bar{\mathbf{z}}(t)$, we have

$$\sum_{i=1}^n \left\|\frac{\mathbf{z}_i(t)}{w_i(t)} - \bar{\mathbf{z}}(t)\right\|^2 = \sum_{i=1}^n\sum_{k=1}^n \left(\frac{z_i^k(t)}{w_i(t)} - \bar{z}^k(t)\right)^2. \tag{24}$$

Thus, we can upper-bound the quantity on the right-hand side.

By inspecting equation (7a), it is easy to see that for any $i \in \mathcal{V}$ and $t \geq 1$, we have

$$w_i(t) = \sum_{\ell=1}^n [A(t-1:0)]_{i\ell} w_i(0) = \sum_{\ell=1}^n [A(t-1:0)]_{i\ell}.$$

From this and Lemma 1 [with $M(t)$ replaced by $A(t)$ in part (b)], we have the following chain of relations:

$$\frac{z_i^k(t)}{w_i(t)} - \bar{z}^k(t)$$
$$= \frac{n\sum_{s=0}^{t-1}[A(t-1:s+1)]_{ik}u_k(s)}{\sum_{\ell=1}^n [A(t-1:0)]_{i\ell}} - \sum_{s=0}^{t-1} u_k(s)$$
$$= \sum_{s=0}^{t-1} u_k(s)\frac{\sum_{\ell=1}^n [A(t-1:s+1)]_{ik} - \sum_{\ell=1}^n [A(t-1:0)]_{i\ell}}{\sum_{\ell=1}^n [A(t-1:0)]_{i\ell}} \tag{25}$$

where we used $n[A(t-1:s+1)]_{ik} = \sum_{\ell=1}^n [A(t-1:s+1)]_{ik}$ in the last equality. We now use the convergence result of the column stochastic matrices $A(t)$ [18, Corollary 2], i.e., there exists a sequence $\{\boldsymbol{\phi}(t)\}$ of vectors such that

$$|[A(t:s)]_{ij} - \phi_i(t)| \leq \beta\lambda^{t-s}, \forall t \geq s \geq 0, \tag{26}$$

and

$$\gamma = \inf_{t\geq 0}\left(\min_{1\leq i\leq n}[A(t:0)\mathbf{1}]_i\right). \tag{27}$$

By adding and subtracting $\phi_i(t-1)$ from relation (25) and using the triangle inequality, we obtain

$$\left|\frac{z_i^k(t)}{w_i(t)} - \bar{z}^k(t)\right|$$
$$= \sum_{s=0}^{t-1} |u_k(s)| \left(\frac{\sum_{\ell=1}^n |[A(t-1:s+1)]_{ik} - \phi_i(t-1)|}{\sum_{\ell=1}^n [A(t-1:0)]_{i\ell}} + \frac{\sum_{\ell=1}^n |\phi_i(t-1) - [A(t-1:0)]_{i\ell}|}{\sum_{\ell=1}^n [A(t-1:0)]_{i\ell}} \right)$$
$$\leq \sum_{s=0}^{t-1} |u_k(s)| \left(\frac{\beta\lambda^{t-s-2}}{\gamma} + \frac{\beta\lambda^{t-1}}{\gamma} \right) \leq \sum_{s=0}^{t-1} |u_k(s)| \frac{2\beta\lambda^{t-s-2}}{\gamma} \tag{28}$$

where in the last two inequalities we used results (26)–(27) and the fact that $\beta\lambda^{t-s-2} \geq \beta\lambda^{t-1}$ for all $s = 0, \ldots, t-1$. From this, we obtain

$$\left|\frac{z_i^k(t)}{w_i(t)} - \bar{z}^k(t)\right|^2 \leq \left(\sum_{s=0}^{t-1} \frac{2\beta\lambda^{t-s-2}}{\gamma} \right)^2 \max_s |u_k(s)|^2 \leq \frac{4\beta^2 \bar{L}^2}{\gamma^2\lambda^2(1-\lambda)^2}$$

where the last inequality follows from Assumption 1. Substituting this estimate in relation (24), we get the desired result. ■

B. Pseudo-Regret of SODA-PS

We now show that SODA-PS achieves the pseudo-regret of order $O(\sqrt{T})$.

Theorem 3: Suppose that Assumptions 1–4 hold. Then, the expected regret (2) of SODA-PS over any time-varying sequence of B-strongly connected digraphs $\{\mathcal{G}_2(t)\}_{t\geq 0}$ with stochastic local gradient signals $u_i(t) = \widehat{g}_i(t)$ and with the step-size choice of $\alpha(t) = \frac{1}{\sqrt{t+1}}$ achieves the following pseudo-regret: for all $T \geq 1$

$$\bar{R}(T) \leq \left(n\bar{L}^2 + C + (L + \sqrt{n}GD_{\mathsf{X}}) \frac{2n\beta\bar{L}}{\gamma\lambda(1-\lambda)} \right) \sqrt{T+1}.$$

Proof: From Lemma 1(a) and the definition of $\mathbf{u}(t)$, it can be seen that

$$\bar{\mathbf{x}}(t) = \Pi_{\mathsf{X}^n}^{\psi}(\bar{\mathbf{z}}(t), \alpha(t-1)) = \Pi_{\mathsf{X}^n}^{\psi}\left(\sum_{s=1}^{t-1} \mathbf{u}(s), \alpha(t-1) \right)$$

from which we can invoke the generic regret bound in Theorem 1.

Using the same arguments as in Theorem 2 while using

$$\|\bar{\mathbf{x}}(t) - \boldsymbol{\xi}_i(t)\| = \left\| \Pi_{\mathsf{X}^n}^{\psi}(\bar{\mathbf{z}}(t), \alpha(t-1)) - \Pi_{\mathsf{X}^n}^{\psi}\left(\frac{\mathbf{z}_i(t)}{w_i(t)}, \alpha(t-1) \right) \right\| \leq \alpha(t-1) \left\| \bar{\mathbf{z}}(t) - \frac{\mathbf{z}_i(t)}{w_i(t)} \right\| \tag{29}$$

we arrive at

$$\bar{R}(T) \leq \frac{1}{2} \sum_{t=1}^T \alpha(t-1) \mathbb{E}\|\mathbf{u}(t)\|^2 + \frac{C}{\alpha(T)} + (L + \sqrt{n}GD_{\mathsf{X}}) \sum_{t=1}^T \alpha(t-1) \sum_{i=1}^n \mathbb{E}\left\| \frac{\mathbf{z}_i(t)}{w_i(t)} - \bar{\mathbf{z}}(t) \right\|.$$

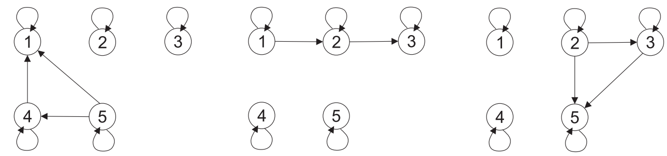

Fig. 1. Time-varying communication topology changing in cycle of three used for **SODA-PS** and its undirected version used for **SODA-C**.

For bounding the network disagreement term, we use Jensen's inequality to write

$$\sum_{i=1}^n \mathbb{E}\left\| \frac{\mathbf{z}_i(t)}{w_i(t)} - \bar{\mathbf{z}}(t) \right\| \leq \sqrt{n} \sqrt{\sum_{i=1}^n \mathbb{E}\left\| \frac{\mathbf{z}_i(t)}{w_i(t)} - \bar{\mathbf{z}}(t) \right\|^2}$$

and Lemma 3. Lastly, using the relation $\sum_{t=1}^T \alpha(t) \leq 2\sqrt{T}$, we obtain the claimed pseudo-regret bound. ■

VII. Simulation Results

Consider the problem of estimating a target vector $\mathbf{x} \in \mathbb{R}^p$ using measurements from a network of n sensors. Each sensor i is in charge of estimating a subvector $\mathbf{x}_i \in \mathbb{R}^{p_i}$ of $\mathbf{x}$, where $p_i \ll p$ and $p = \sum_{i=1}^n p_i$ is a large number. An example includes the localization of multiple targets, where in this case $\mathbf{x} \in \mathbb{R}^p$ becomes a stacked vector of all target locations. When there are a number of spatially dispersed targets, we can certainly benefit from distributed sensing.

The sensors are assumed to have a linear model of $r(\mathbf{x}) = A\mathbf{x}$, where $A \in \mathbb{R}^{m\times p}$ and $m < p$. At each time t, each sensor $i \in \mathcal{V}$ estimates its portion $\mathbf{x}_i(t) \in \mathbb{R}^{p_i}$ of the target vector $\mathbf{x} \in \mathbb{R}^p$, and then takes a measurement $q_t^i \in \mathbb{R}^{m_i}$, which is corrupted by observation error and possibly by modeling error. We use $\mathbf{q}_t \in \mathbb{R}^m$ with $m = \sum_{i=1}^n m_i$ to refer to the stacked vector of all q_t^i's. The regret is computed with respect to the least-squares estimate of the target locations at time T, i.e., $\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathsf{X}^p} \sum_{t=1}^T f_t(\mathbf{x})$, where $f_t(\mathbf{x}) = \frac{1}{2}\|A\mathbf{x} - \mathbf{q}_t\|^2$ and we set $\mathsf{X} \in [-20, 20]$. In the experiment, we use $u_i(t) = \langle \nabla f_t(\boldsymbol{\xi}_i(t)), \mathbf{e}_i \rangle + \delta$, where δ is uniform random noise in $[-0.5, 0.5]$. Note that small sized sensors usually adopt reduced range number systems such as fixed-point arithmetic to minimize computation effort, power consumption and chip size. Therefore, δ can be viewed as a quantization error in this case.

For SODA-PS, we experiment with a time-varying sequence of digraphs with $n = 5$ nodes whose communication topology is changing periodically with period 3 (see Fig. 1). The graph sequence is, therefore, 3-strongly connected. The averaging matrices $A(t)$ [cf. (8)] can be determined accordingly. For SODA-C, we also experiment with the graph in Fig. 1, but remove directions on the edges. We set $[M(t)]_{ij} = 1/5$ if $i \leftrightarrow j$, and $[M(t)]_{ii} = 1 - \sum_j [M(t)]_{ij}$ for all i. We ran our algorithms once for each $T \in [1000]$. That is, for a given T, the iterates in the algorithms are updated from $t = 1$ to $t = T$. We used step size $\alpha(t) = \frac{1}{\sqrt{t+1}}$ for both algorithms. We repeated the same experiment with 100 different random seeds and take the average to obtain $\bar{R}(T)$.

In Fig. 2, we depict the average pseudo-regret $\bar{R}(T)/T$ over time T of the distributed sensing problem when SODA-C and SODA-PS are used, respectively. It shows that the regret is sublinear for both algorithms and the average $\bar{R}(T)/T$ goes to zero as the time increases.

VIII. Conclusion

We have studied an online optimization problem for multi-agent systems in scenarios when the underlying network topology is

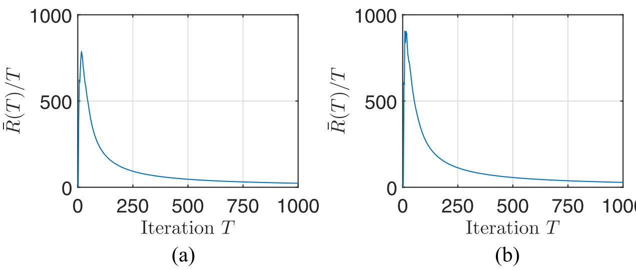

Fig. 2. Average pseudo-regret $\bar{R}(T)/T$ versus iterations for Online Distributed Active Sensing using **SODA-C** (left) and **SODA-PS** (right).

time-varying. We proposed two decentralized stochastic variants of Nesterov's dual-averaging method, namely, SODA-C using the information exchange dynamics by Li and Marden and SODA-PS using the broadcast-based push-sum protocol. The bounds on the expected regret for both algorithm are shown to grow as $O(\sqrt{T})$ when the step size is $\alpha(t) = 1/\sqrt{t+1}$. For SODA-C, the bound is valid for a sequence of undirected connected graphs and a row-stochastic matrix of weights $M(t)$. For SODA-PS, the bound is valid for a uniformly strongly connected sequence of digraphs and column-stochastic matrices of weights $A(t)$ whose components are based on the out-degrees of neighbors. We also provided simulation results of the proposed algorithms on sensor networks to complement our theoretical analysis.

References

[1] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, Applied Mathematics Series. Princeton, NJ: Princeton University Press, 2009.
[2] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods for Multiagent Networks*. Princeton, NJ: Princeton University Press, 2010.
[3] S. Kar and J. Moura, "Distributed consensus algorithms in sensor networks: Quantized data and random link failures," *IEEE Trans. Signal Process.*, vol. 58, pp. 1383–1400, 2010.
[4] A. Martinoli, F. Mondada, G. Mermoud, N. Correll, M. Egerstedt, A. Hsieh, L. Parker, and K. Stoy, *Distributed Autonomous Robotic Systems*. ser. Springer Tracts in Advanced Robotics. New York: Springer-Verlag, 2013.
[5] B. Zhang, A. Lam, A. Dominguez-Garcia, and D. Tse, "Optimal distributed voltage regulation in power distribution networks," 2015, http://arxiv.org/abs/1204.5226.
[6] T.-H. Chang, A. Nedić, and A. Scaglione, "Distributed constrained optimization by consensus-based primal-dual perturbation method," *IEEE Trans. Autom. Control*, vol. 59, pp. 1524–1538, 2014.
[7] J. Tsitsiklis, *Problems in Decentralized Decision Making and Computation*, Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., , Massachusetts Institute of Technology, , Cambridge, , MA, 1984.
[8] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. AC-31, pp. 803–812, 1986.
[9] J. N. Tsitsiklis and M. Athans, "Convergence and asymptotic agreement in distributed decision problems," *IEEE Trans. Autom. Control*, vol. AC-29, pp. 42–50, 1984.
[10] S. Li and T. Basar, "Distributed learning algorithms for the computation of noncooperative equilibria," *Automatica*, vol. 23, pp. 523–533, 1987.
[11] M. Raginsky, N. Kiarashi, and R. Willett, "Decentralized online convex programming with local information," in *Proc. Amer. Control Conf.*, 2011, pp. 5363–5369.
[12] K. Kvaternik, J. Llorca, D. Kilper, and L. Pavel, "A decentralized coordination strategy for networked multiagent systems," in *Proc. 50th Annu. Allerton Conf. Communication, Control, and Computing*, Oct. 2012, pp. 41–47.
[13] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Belmont, MA: Athena Scientific, 1997.
[14] N. Li and J. R. Marden, "Designing games for distributed optimization," *IEEE J. Select. Topics Signal Process.*, vol. 7, pp. 230–242, 2013.
[15] D. Kempe, A. Dobra, and J. Gehrke, "Gossip-based computation of aggregate information," in *Proc. 44th Annu. IEEE Symp. Foundations of Computer Science*, vol. 44, 2003, pp. 482–491.
[16] F. Benezit, V. Blondel, P. Thiran, J. Tsitsiklis, and M. Vetterli, "Weighted gossip: Distributed averaging using non-doubly stochastic matrices," in *Proc. IEEE Int. Symp. Information Theory Proceedings (ISIT)*, 2010, pp. 1753–1757.
[17] K. Tsianos, S. Lawlor, and M. Rabbat, "Push-sum distributed dual averaging for convex optimization," in *Proc. 51st Annu. Conf. Decision and Control*, 2012, pp. 5453–5458.
[18] A. Nedić and A. Olshevsky, "Distributed optimization over time-varying directed graphs," 2013, http://arxiv.org/abs/1303.2289.
[19] Y. Nesterov, "Primal-dual subgradient methods for convex problems," *Math. Program., ser. B*, vol. 120, pp. 221–259, 2009.
[20] S. Lee, A. Nedić, and M. Raginsky, "Coordinate dual averaging for decentralized online optimization with nonseparable global objectives," *IEEE Trans. Control Network Syst.* Apr.2016.
[21] J. C. Duchi, A. Agarwal, and M. J. Wainwright, "Dual averaging for distributed optimization: Convergence analysis and network scaling," *IEEE Trans. Autom. Control*, vol. 57, pp. 592–606, 2012.
[22] M. Akbari, B. Gharesifard, and T. Linder, "Distributed subgradient-push online convex optimization on time-varying directed graphs," in *Proc. 52nd Allerton Conf. Communication, Control, and Computing*, 2014, pp. 264–269.
[23] S. Hosseini, A. Chapman, and M. Mesbahi, "Online distributed optimization via dual averaging," in *Proc. IEEE 52nd Conf. Decision and Control*, 2013, pp. 1484–1489.
[24] D. Mateos-Nunez and J. Cortés, "Distributed online convex optimization over jointly connected digraphs," *IEEE Trans. Network Sci. Eng.*, vol. 1, no. 1, pp. 23–37, 2014.
[25] S. Bubeck and N. Cesa-Bianchi, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," in*Found. Trends Mach. Learn.*, vol. 5, no. 1, pp. 1–122, 2012.
[26] A. S. Nemirovsky and D. B. Yudin, *Problem Complexity and Method Efficiency in Optimization*. New York: Wiley, 1982.
[27] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "Distributed subgradient methods and quantization effects," in *Proc. IEEE Conf. Decision and Control*, Dec. 2008, pp. 4177–4184.