

Distributed Subgradient Projection Algorithm for Convex Optimization

S. Sundhar Ram, A. Nedić and V. V. Veeravalli

University of Illinois at Urbana-Champaign

Email: {ssriniv5,angelia,vvv}@illinois.edu

Abstract—We consider constrained minimization of a sum of convex functions over a convex and compact set, when each component function is known only to a specific agent in a time-varying peer to peer network. We study an iterative optimization algorithm in which each agent obtains a weighted average of its own iterate with the iterates of its neighbors, updates the average using the subgradient of its local function and then projects onto the constraint set to generate the new iterate. We obtain error bounds on the limit of the function value when a constant stepsize is used.

Index Terms—distributed optimization, time-varying network, subgradient algorithm.

I. INTRODUCTION

A number of problems arising in the context of wired and wireless networks can be posed as the minimization of a sum of functions, when each component function is available only to a specific agent [1]–[3]. In many scenarios, it is not efficient, or not possible, for the network agents to share their objective functions with each other or with a central coordinator. For example, in a large wireless network, energy is a scarce resource and it may not be efficient for a central coordinator to learn the individual objective functions from each and every agent [1]. In a network of databases from which information is to be mined, privacy considerations may prohibit the sharing of the objective functions [4]. In a distributed network on a single chip, for the chip to be fault tolerant, it is desirable to perform the processing in a distributed manner to account for the statistical process variations [5]. In these settings, distributed algorithms are preferable that only require the agents to locally exchange limited and high level information.

We consider constrained minimization of a sum of convex functions, where each component function is known only to a specific network agent. The algorithm is a modification of the distributed algorithm proposed in [6] for unconstrained minimization. Each network agent maintains and updates its own iterate, and communicates the iterate to a subset of neighboring agents. The agent information exchange model is a delayless version of the computational model proposed by Tsitsiklis [7]. Each agent combines the received iterates with its own iterate, and then adjusts the iterate by using information of its own objective function. The algorithm is *distributed* since there is no central coordinator. Also, the algorithm is *local* since each agent uses only locally available information of its function and communicates only with its neighbors.

Related distributed and local algorithms are incremental methods, where the network agents sequentially update the

iterate in a cyclic or a random order [8], [2] (and references therein). In an incremental algorithm, only one agent updates the iterate at a given time. While distributed and local, incremental algorithms differ fundamentally from the algorithm studied in this paper, where all agents update simultaneously. Also related are the optimization algorithms in [9], which however, are not local as the objective function information is available to each agent.

This paper contributes to the literature on distributed optimization. The novelty of the paper is the consideration of constrained optimization problem within the distributed multi-agent setting. The presence of the constraint set complicates the model by introducing non-linearities in the agent system dynamics. The closest related work is [6], where unconstrained problem has been considered. Another closely related work is [10], where pure consensus problem has been investigated in the presence of constraints.

The rest of the paper is organized as follows. In Section II, we formulate the problem, describe the algorithm, and state our basic assumptions on the agent connectivity and information exchange. In Section III, we obtain an error bound on the algorithm's performance. We provide some concluding remarks in Section IV.

II. PROBLEM FORMULATION AND ALGORITHM

We consider a network of m agents that are indexed by $1, \dots, m$. Often, when convenient, we index the agents by using set $V = \{1, \dots, m\}$. The network objective is to solve the following constrained optimization problem:

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^m f_i(x) \\ & \text{subject to} && x \in X, \end{aligned} \quad (1)$$

where $X \subseteq \mathbb{R}^n$ is a constraint set and $f_i : X \rightarrow \mathbb{R}$ for all i . The function f_i is known only to agent i .

The goal is to solve problem (1) using a distributed algorithm in which the agents do not share the component functions with each other, but only exchange the iterates with their immediate neighbors in each iteration.

A. Algorithm

Let $w_{i,k}$ be the iterate with agent i at the end of iteration k . At the beginning of iteration $k + 1$, agent i receives the iterates $w_{j,k}$ of a subset of neighboring agents j . Then, agent

i computes a weighted average of these iterates and adjusts this average along the negative gradient direction of f_i . Formally, each agent i generates an iterate sequence $\{w_{i,k}\}$ as follows:

$$w_{i,k+1} = P_X [v_{i,k} - \alpha \nabla f_i(v_{i,k})], \quad (2)$$

where $w_{i,0} \in X$ is some initial iterate, α is the stepsize, $\nabla f_i(x)$ denotes the gradient of f_i at x , and P_X denotes the Euclidean projection onto the set X . The vector $v_{i,k}$ is the weighted average computed by agent i , given by

$$v_{i,k} = \sum_{j \in N_i(k+1)} a_{i,j}(k+1) w_{j,k}, \quad (3)$$

where $N_i(k+1)$ denotes the set of agents whose current iterates are available to agent i in the $(k+1)$ -th iteration. We assume that $i \in N_i(k+1)$ for all agents i and at all times k . The scalars $a_{i,j}(k+1)$ are the non-negative weights that agent i assigns to agent j 's iterate. For convenience, we define $a_{i,j}(k+1) = 0$ for $j \notin N_i(k+1)$ and rewrite (3) as

$$v_{i,k} = \sum_{j=1}^m a_{i,j}(k+1) w_{j,k}. \quad (4)$$

The rule for selecting the weights is discussed, among other assumptions, in the next section.

B. Assumptions

We define $f(x) = \sum_{i=1}^m f_i(x)$, $f^* = \min_{x \in X} f(x)$, and $X^* = \{x \in X : f(x) = f^*\}$. We make the following basic assumptions on the constraint set X and the functions f_i .

Assumption 1: The set $X \subseteq \mathbb{R}^n$ is convex and compact. The functions f_i , $i \in V$, are defined and convex over an open set that contains the set X .

This assumption is satisfied for example when each function f_i is defined and convex over the entire space \mathbb{R}^n . An immediate consequence of the assumption is that the functions f_i are continuous over the compact set X . Thus, by Weierstrass theorem, the optimal value f^* of the problem is finite and the optimal set X^* is nonempty.

We make no assumption on the differentiability of the functions f_i . At points where the gradient does not exist, $\nabla f_i(x)$ will instead denote a subgradient of f_i at x . A vector $\nabla f_i(x)$ is a subgradient of f_i at a point x if

$$\nabla f_i(x)^T(y - x) \leq f_i(y) - f_i(x) \quad \text{for all } y \in X. \quad (5)$$

A subgradient of f_i exists at any point of the set X since X is contained in an open set over which each f_i is assumed to be convex (see [11]). When f_i is differentiable at a point, the only subgradient is the gradient. To keep the discussion general, we refer to ∇f_i as a subgradient in the rest of the paper. Further, the subgradients are bounded since the set X is bounded¹. We use C_i to denote a bound on the subgradients of f_i over the set X , i.e., for each i ,

$$\|\nabla f_i(x)\| \leq C_i \quad \text{for all } x \in X.$$

¹See [11], Proposition 4.2.3.

The next assumption ensures that the agents communicate sufficiently often so that all component functions, directly or indirectly, influence the iterate sequence of any agent. Using the set $N_i(k+1)$ of agents communicating with agent i at time $k+1$, we define (V, E_{k+1}) to be the graph with edges

$$E_{k+1} = \{(j, i) : j \in N_i(k+1), i \in V\}.$$

Assumption 2: There exists an integer $Q \geq 1$ such that the graph $(V, \cup_{l=1, \dots, Q} E_{k+l})$ is strongly connected for all k .

It is important that the influence of agent i is “equal” in a long run, so that the sum of the functions f_i is minimized rather than a weighted sum of them. To ensure equal influence, we make the following assumption on the weights.

Assumption 3: For all $i \in V$ and all k ,

- (a) $a_{i,j}(k+1) \geq 0$, and $a_{i,j}(k+1) = 0$ when $j \notin N_i(k+1)$,
- (b) $\sum_{j=1}^m a_{i,j}(k+1) = 1$,
- (c) There exists a scalar η , $0 < \eta < 1$, such that $a_{i,j}(k+1) \geq \eta$ when $j \in N_i(k+1)$,
- (d) $\sum_{i=1}^m a_{i,j}(k+1) = 1$.

Assumptions 3(a) and 3(b) state that each agent uses a weighted average of all the iterates it has access to. Assumption 3(c) ensures that each agent gives a sufficient weight to its iterate and all the iterates it receives.² Assumption 3(d), together with Assumption 2, ensures that all agents are equally influential in a long run. To satisfy Assumption 3(d), the agents need to coordinate their weights, as discussed in [2], [6].

C. Preliminaries

Let $A(k)$ be the matrix with (i, j) -th entry equal to $a_{i,j}(k)$. As a consequence of Assumptions 3(a), 3(b) and 3(d), the matrix $A(k)$ is doubly stochastic³. Define, for all k, s with $k \geq s$,

$$\Phi(k, s) = A(k)A(k-1) \cdots A(s+1). \quad (6)$$

We next state a result from [12] (Corollary 1) on the convergence of the matrix $\Phi(k, s)$. Let $[\Phi(k, s)]_{i,j}$ denote the (i, j) -th entry of the matrix $\Phi(k, s)$, and let $e \in \mathbb{R}^m$ be the column vector with all entries equal to 1.

Lemma 1: Let Assumptions 2 and 3 hold. Then

- 1) $\lim_{k \rightarrow \infty} \Phi(k, s) = \frac{1}{m} ee^T$ for all s .
- 2) Further, the convergence is geometric and the rate of convergence is given by

$$\left| [\Phi(k, s)]_{i,j} - \frac{1}{m} \right| \leq \theta \beta^{k-s},$$

$$\text{where } \theta = \left(1 - \frac{\eta}{4m^2}\right)^{-2} \text{ and } \beta = \left(1 - \frac{\eta}{4m^2}\right)^{\frac{1}{Q}}.$$

III. ALGORITHM ANALYSIS

We first use Lemma 1 to estimate the difference between the iterates across agents. For this, we introduce auxiliary vectors

$$y_k = \frac{1}{m} \sum_{i=1}^m w_{i,k} \quad \text{for all } k. \quad (7)$$

²Agents need not be aware of the common bound η .

³The sum of its entries in every row and in every column is equal to 1.

The following result relates these vectors to the agent iterates.

Lemma 2: Under Assumptions 1–3, we have for all $j \in V$,

$$\limsup_{k \rightarrow \infty} \|y_k - w_{j,k}\| \leq \alpha \left(2 + \frac{m\theta\beta}{1-\beta}\right) \max_{i \in V} C_i.$$

Proof: Define for all $i \in V$ and all k ,

$$p_{i,k+1} = w_{i,k+1} - \sum_{j=1}^m a_{i,j}(k+1)w_{j,k}. \quad (8)$$

Using the matrices $\Phi(k, s)$ in (6), we can write (as in [13])

$$\begin{aligned} w_{i,k+1} &= \sum_{j=1}^m [\Phi(k+1, s)]_{i,j} w_{j,s} + p_{i,k+1} \\ &\quad + \sum_{\ell=s+1}^k \left(\sum_{j=1}^m [\Phi(k+1, \ell)]_{i,j} p_{j,\ell} \right). \end{aligned} \quad (9)$$

We can also rewrite y_k , defined in (7), as follows

$$y_{k+1} = \frac{1}{m} \left(\sum_{j=1}^m \sum_{i=1}^m a_{i,j}(k+1)w_{j,k} + \sum_{i=1}^m p_{i,k+1} \right).$$

In the view of $\sum_{i=1}^m a_{i,j}(k+1) = 1$, we have

$$y_{k+1} = \frac{1}{m} \left(\sum_{j=1}^m w_{j,k} + \sum_{i=1}^m p_{i,k+1} \right) = y_k + \frac{1}{m} \sum_{i=1}^m p_{i,k+1}.$$

Therefore, for any $s \leq k+1$,

$$\begin{aligned} y_{k+1} &= y_s + \frac{1}{m} \sum_{\ell=s+1}^{k+1} \sum_{j=1}^m p_{j,\ell} \\ &= \frac{1}{m} \sum_{j=1}^m w_{j,s} + \frac{1}{m} \sum_{\ell=s+1}^{k+1} \sum_{j=1}^m p_{j,\ell}. \end{aligned} \quad (10)$$

Substituting for y_{k+1} from (10) and for $w_{j,k+1}$ from (9), we obtain

$$\begin{aligned} \|y_{k+1} - w_{j,k+1}\| &= \left\| \sum_{i=1}^m \left(\frac{1}{m} - [\Phi(k+1, s)]_{j,i} \right) w_{i,s} \right. \\ &\quad + \sum_{\ell=s+1}^k \sum_{i=1}^m \left(\frac{1}{m} - [\Phi(k+1, \ell)]_{j,i} \right) p_{i,\ell} \\ &\quad \left. + \left(\frac{1}{m} \sum_{i=1}^m p_{i,k+1} - p_{j,k+1} \right) \right\|. \end{aligned}$$

Therefore, for all $j \in V$ and all k, s with $s \leq k+1$,

$$\begin{aligned} \|y_{k+1} - w_{j,k+1}\| &\leq \sum_{i=1}^m \left| \frac{1}{m} - [\Phi(k+1, s)]_{j,i} \right| \|w_{i,s}\| \\ &\quad + \sum_{\ell=s+1}^k \sum_{i=1}^m \left| \frac{1}{m} - [\Phi(k+1, \ell)]_{j,i} \right| \|p_{i,\ell}\| \\ &\quad + \frac{1}{m} \sum_{i=1}^m \|p_{i,k+1}\| + \|p_{j,k+1}\|. \end{aligned}$$

We have $\|w_{i,s}\| \leq \max_{x \in X} \|x\|$. Further, by using the result from Lemma 1 to bound $\left| \frac{1}{m} - [\Phi(k, \ell)]_{j,i} \right|$, we obtain

$$\begin{aligned} \|y_{k+1} - w_{j,k+1}\| &\leq m\theta\beta^{k+1-s} \max_{x \in X} \|x\| \\ &\quad + \theta \sum_{\ell=s+1}^k \beta^{k+1-\ell} \sum_{i=1}^m \|p_{i,\ell}\| \\ &\quad + \frac{1}{m} \sum_{i=1}^m \|p_{i,k+1}\| + \|p_{j,k+1}\|. \end{aligned} \quad (11)$$

We next estimate the norms $\|p_{i,k}\|$ for any k . From the definition of $p_{i,k+1}$ in (8) and the definition of the vector $v_{i,k}$ in (3), we have $p_{i,k+1} = w_{i,k+1} - v_{i,k}$. Being a weighted average of vectors $w_{j,k}$ in the convex set X , the vector $v_{i,k}$ is in the set X . By the definition of $w_{i,k+1}$ in (2) and the non-expansive property of the Euclidean projection, we have

$$\begin{aligned} \|p_{i,k+1}\| &= \|P_X[v_{i,k} - \alpha \nabla f_i(v_{i,k})] - v_{i,k}\| \\ &\leq \alpha \|\nabla f_i(v_{i,k})\|. \end{aligned}$$

Since the subgradients of f_i are bounded by C_i over the set X , we have $\|p_{i,k+1}\| \leq \alpha C_i$. Using this and $\sum_{\ell=s+1}^k \beta^{k+1-\ell} = \sum_{t=1}^{k-s} \beta^t$ in (11), we obtain

$$\begin{aligned} \|y_{k+1} - w_{j,k+1}\| &\leq m\theta\beta^{k+1-s} \max_{x \in X} \|x\| + \alpha\theta \sum_{t=1}^{k-s} \beta^t \sum_{i=1}^m C_i \\ &\quad + \alpha \frac{1}{m} \sum_{i=1}^m C_i + \alpha C_j. \end{aligned} \quad (12)$$

The result follows by setting $s = 1$ and letting $k \rightarrow \infty$. ■

Using Lemma 2, we now provide an error bound on the values of the objective function f at the agent iterates.

Theorem 1: Under Assumptions 1–3 we have for all $j \in V$,

$$\liminf_{k \rightarrow \infty} f(w_{j,k}) \leq f^* + \alpha \max_{\ell \in V} C_\ell \left(\sum_{i=1}^m C_i \right) \left(\frac{9}{2} + \frac{2m\theta\beta}{1-\beta} \right).$$

Proof: Using the non-expansive property of the Euclidean projection, the definition of the iterate $w_{i,k+1}$ in (2) and the subgradient inequality in (5), we have for any $x^* \in X^*$ and k ,

$$\begin{aligned} \|w_{i,k+1} - x^*\|^2 &\leq \|v_{i,k} - x^*\|^2 - 2\alpha (f_i(v_{i,k}) - f_i(x^*)) \\ &\quad + \alpha^2 \|\nabla f_i(v_{i,k})\|^2. \end{aligned}$$

Summing over $i \in V$ and using $\|\nabla f_i(v_{i,k})\| \leq C_i$, we obtain

$$\begin{aligned} \sum_{i=1}^m \|w_{i,k+1} - x^*\|^2 &\leq \sum_{i=1}^m \|v_{i,k} - x^*\|^2 + \alpha^2 \sum_{i=1}^m C_i^2 \\ &\quad - 2\alpha \sum_{i=1}^m (f_i(v_{i,k}) - f_i(x^*)). \end{aligned}$$

Convexity of the Euclidean norm and Assumption 3 imply $\sum_{i=1}^m \|v_{i,k+1} - x^*\|^2 \leq \sum_{j=1}^m \|w_{j,k+1} - x^*\|^2$. Combined with the preceding relation, this yields for any $x^* \in X$ and all k ,

$$\begin{aligned} \sum_{i=1}^m \|v_{i,k+1} - x^*\|^2 &\leq \sum_{i=1}^m \|v_{i,k} - x^*\|^2 + \alpha^2 \sum_{i=1}^m C_i^2 \\ &\quad - 2\alpha \sum_{i=1}^m (f_i(v_{i,k}) - f_i(x^*)). \end{aligned}$$

In view of the subgradient boundedness, from (5) we have

$$\begin{aligned} f_i(v_{i,k}) - f_i(x^*) &\geq (f_i(v_{i,k}) - f_i(y_k)) + (f_i(y_k) - f_i(x^*)) \\ &\geq -C_i \|y_k - v_{i,k}\| + (f_i(y_k) - f_i(x^*)). \end{aligned}$$

Substituting for $v_{i,k+1}$ from (4) and using the convexity of the norm, we have $\|y_k - v_{i,k}\| = \|y_k - \sum_{j=1}^m a_{i,j}(k+1)w_{j,k}\| \leq \sum_{j=1}^m a_{i,j}(k+1)\|y_k - w_{j,k}\|$. Using this and $\sum_{i=1}^m f_i(x^*) = f^*$, from the preceding relation we obtain

$$\begin{aligned} &\sum_{i=1}^m (f_i(v_{i,k}) - f_i(x^*)) \\ &\geq -\left(\max_{\ell \in V} C_\ell\right) \sum_{j=1}^m \sum_{i=1}^m a_{i,j}(k+1) \|y_k - w_{j,k}\| + f(y_k) - f^* \\ &= -\left(\max_{\ell \in V} C_\ell\right) \sum_{j=1}^m \|y_k - w_{j,k}\| + f(y_k) - f^*. \end{aligned}$$

Therefore, it follows

$$\begin{aligned} \sum_{i=1}^m \|v_{i,k+1} - x^*\|^2 &\leq \sum_{i=1}^m \|v_{i,k} - x^*\|^2 - 2\alpha(f(y_k) - f^*) \\ &\quad + \alpha^2 \sum_{i=1}^m C_i^2 + 2\alpha \left(\max_{\ell \in V} C_\ell\right) \sum_{j=1}^m \|y_k - w_{j,k}\|. \end{aligned}$$

Using (12) with appropriate k , and $s = k-1-\tau$ for $\tau \leq k-1$, we obtain

$$\sum_{j=1}^m \|y_k - w_{j,k}\| \leq m^2 \theta \beta^\tau \max_{x \in X} \|x\| + \left(\sum_{i=1}^m C_i\right) \left(\frac{m\alpha\theta\beta}{1-\beta} + 2\alpha\right).$$

From the preceding two relations it follows for $\tau \leq k-1$,

$$\begin{aligned} &\sum_{i=1}^m \|v_{i,k+1} - x^*\|^2 \\ &\leq \sum_{i=1}^m \|v_{i,k} - x^*\|^2 - 2\alpha(f(y_k) - f^*) + \alpha^2 \sum_{i=1}^m C_i^2 \\ &\quad + 2\alpha \left(\max_{\ell \in V} C_\ell\right) m^2 \theta \beta^\tau \max_{x \in X} \|x\| \\ &\quad + 2\alpha^2 \left(\max_{\ell \in V} C_\ell\right) \left(\sum_{i=1}^m C_i\right) \left(\frac{m\theta\beta}{1-\beta} + 2\right). \end{aligned}$$

By using $\sum_{i=1}^m C_i^2 \leq \max_{\ell \in V} C_\ell \sum_{i=1}^m C_i$ and a line of argument similar to that in the proof of Theorem 3.4 of [2], we can see that for all τ ,

$$\begin{aligned} \liminf_{k \rightarrow \infty} f(y_k) &\leq f^* + m^2 \theta \beta^\tau \max_{x \in X} \|x\| \max_{\ell \in V} C_\ell \\ &\quad + \alpha \max_{\ell \in V} C_\ell \left(\sum_{i=1}^m C_i\right) \left(\frac{5}{2} + \frac{m\theta\beta}{1-\beta}\right). \end{aligned}$$

In the limit as $\tau \rightarrow \infty$, we obtain

$$\begin{aligned} &\liminf_{k \rightarrow \infty} f(y_k) \\ &\leq f^* + \alpha \max_{\ell \in V} C_\ell \left(\sum_{i=1}^m C_i\right) \left(\frac{5}{2} + \frac{m\theta\beta}{1-\beta}\right). \end{aligned} \quad (13)$$

From the subgradient inequality in (5) and the subgradient boundedness we can see that for all $j \in V$,

$$f(w_{j,k}) - f(y_k) \leq \left(\sum_{i=1}^m C_i\right) \|y_k - w_{j,k}\|. \quad (14)$$

The result follows from Lemma 2, and Eqs. (13) and (14). ■
Theorem 1 implies that for all j ,

$$0 \leq \inf_k f(w_{j,k}) - f^* \leq \alpha m \left(\max_{\ell \in V} C_\ell\right)^2 \left(\frac{9}{2} + \frac{2m\theta\beta}{1-\beta}\right).$$

When the parameter η does not depend on the number m of agents, the error term is of the order of m^4 . Note however, that the error bound is for a general network topology. We believe that the bound is not tight and expect to obtain better bounds for special topologies such as spanning trees.

IV. CONCLUSION

We studied the performance of a distributed algorithm to minimize the sum of convex functions, when each function was known only to a specific network agent. We obtained an error bound on the performance of the algorithm for a constant stepsize. Our further interest is to investigate convergence when diminishing and other stepsize rules are used.

REFERENCES

- [1] M. G. Rabbat and R. D. Nowak, "Quantized incremental algorithms for distributed optimization," *IEEE Journal on Select Areas in Communications*, vol. 23, no. 4, pp. 798–808, 2005.
- [2] S. Sundhar Ram, A. Nedić, and V. V. Veeravalli, "Incremental stochastic sub-gradient algorithms for convex optimization," Available at <http://arxiv.org/abs/0806.1092>, 2008.
- [3] S. Sundhar Ram, V. V. Veeravalli, and A. Nedić, "Distributed and non-autonomous power control through distributed convex optimization," accepted at INFOCOM Mini-Conference, 2009.
- [4] N. Ye, Ed., *Data mining handbook*, chapter Distributed data mining: Algorithms, systems, and applications, Lawrence Erlbaum Associates, 2002, Authored by B. Park and H. Kargupta.
- [5] G. Varatkar, S. Narayanan, N. Shanbhag, and D. Jones, "Sensor network-on-chip," in *International Symposium on Systems-on-Chip*, 2007, pp. 1–4.
- [6] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *To appear in Transactions on Automatic Control*, 2009.
- [7] J. N. Tsitsiklis, *Problems in decentralized decision making and computation*, Ph.D. thesis, Massachusetts Institute of Technology, 1984.
- [8] A. Nedić and D. P. Bertsekas, "Incremental subgradient method for nondifferentiable optimization," *SIAM Journal of Optimization*, vol. 12, no. 1, pp. 109–138, 2001.
- [9] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: Numerical methods*, Athena Scientific, 1997.
- [10] A. Nedić, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus," LIDS Report 2779, MIT, 2008.
- [11] D. P. Bertsekas, A. Nedić, and A. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, 2003.
- [12] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "Distributed subgradient algorithms and quantization effects," <http://arxiv.org/abs/0803.1202>, 2008.
- [13] A. Nedić and A. Ozdaglar, "On the rate of convergence of distributed asynchronous subgradient methods for multi-agent optimization," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 4711–4716.