Convergence of Limited Communications Gradient Methods

Sindri Magnússon, Kathryn Heal, Chinwendu Enyioha, Na Li, Carlo Fischione, and Vahid Tarokh

***Abstract*— Distributed control and decision making increasingly play a central role in economical and sustainable operation of cyber-physical systems. Nevertheless, the full potential of the technology has not yet been fully exploited in practice due to communication limitations of real-world infrastructures. This work investigates the fundamental properties of gradient methods for distributed optimization, where gradient information is communicated at every iteration, when using limited number of communicated bits. In particular, a general class of quantized gradient methods are studied where the gradient direction is approximated by a finite quantization set. Conditions on the quantization set are provided that are necessary and sufficient to guarantee the ability of these methods to minimize any convex objective function with Lipschitz continuous gradient and a nonempty, bounded set of optimizers. Moreover, a lower bound on the cardinality of the quantization set is provided, along with specific examples of minimal quantizations. Furthermore, convergence rate results are established that connect the fineness of the quantization and number of iterations needed to reach a predefined solution accuracy. The results provide a bound on the number of bits needed to achieve the desired accuracy. Finally, an application of the theory to resource allocation in power networks is demonstrated, and the theoretical results are substantiated by numerical simulations.**

I. Introduction

Recent advances in distributed control and optimization techniques have enabled more economical and sustainable operation of cyber-physical systems. However, the full potential of the technology has not been fully exploited in many application such as power networks due to inherent communication constraints. For example, although power networks are equipped with a natural communication infrastructure such as power line communications [1], it is currently not used for distributed decision making due to limited bandwidth/Shannon capacity. Instead, research efforts in distributed operation of power networks usually assume high data rates and low latency wireless communication technologies that might be integrated into the networks sometime in the future. Another component of cyber-physical systems where economic communication plays a central role are wireless sensor networks (WSN). WSNs are powered by battery sources for communication over wireless links; hence, are constrained in how much transmission they engage in to prolong battery life and operation time of the sensors. Motivated by the discussion above, the goal of this paper is to investigate fundamental communication limits of distributed optimization approaches based on gradient methods and decomposition.

Decomposition methods in optimization have been widely investigated in wired/wireless communication [2]–[5], power networks [6], [7], and WSNs [8], among others. These methods are typically based on communicating gradient information from a set of source nodes to users, which then solve a simple, local subproblem. The procedure can be performed using a) one-way communication where the source nodes estimate the gradient using available information [3], [9] or b) two-way communication where users and sources need to coordinate to evaluate the gradient. The main contribution of this paper is to investigate the performance of such decomposition methods where bandwidth is limited.

Limited bandwidth in distributed optimization has already received attention in the literature [10]–[12]. For example, [10] considers a variant of incremental gradient methods [13] over networks where each node projects its iterate to a grid before sending the iterate to the next node. Similar quantization ideas are explored in [11] in the context of consensus-type subgradient methods [14]. Our work differs from the aforementioned papers in that we consider decomposition methods where the gradient is communicated whereas in [10], [11] it is the decision variables that are communicated. The work in [12] studies the convergence of standard interference functions methods for power control in cellular wireless systems where base stations send binary signals to the cells. Unlike [12], in this work the gradient information is quantized and transmitted by a constrained number of bits.

A. Contributions of This Work

We consider quantized gradient methods (QGM) where at each iteration the gradient direction is projected to a finite quantization set $\mathcal{D}$. We begin by investigating conditions under which the quantization set $\mathcal{D}$ is *proper* in the sense that QGMs can minimize any convex function $f : \mathbb{R}^N \to \mathbb{R}$ with Lipschitz continuous gradients and non-empty, bounded set of minimizers. We provide necessary and sufficient conditions that characterize such proper quantization sets. We then use this characterization to provide examples of proper quantization sets $\mathcal{D}$. Further, we show that if $|\mathcal{D}| \leq N$ then $\mathcal{D}$ can not be proper, i.e., for every such $\mathcal{D}$ there exists an optimization problem which QGMs can not solve. In addition, we show that there exists a proper quantization with $|\mathcal{D}| = N + 1$. We show that the stopping criteria $||\nabla f(\mathbf{x})|| < \epsilon$ and $f(\mathbf{x}) - f^\star < \epsilon$ can be achieved for any $\epsilon > 0$ in finite number of iterations. Moreover, we provide a bound on the number of iterations needed to achieve these

This work was supported by the VR Chromos Project and NSF grant No. 1548204.

S. Magnússon and C. Fischione are with Electrical Engineering School, Access Linnaeus Center, KTH Royal Institute of Technology, Stockholm, Sweden. E-mail: sindrim@kth.se.

K. Heal, C. Enyioha, N. Li, and V. Tarokh are with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA USA.

stopping conditions; this bound depends on the fineness of the quantization set $\mathcal{D}$. Specifically, the bound on number of iterations decreases when the quantization set becomes finer. We also show that, when the step-sizes are non-summable but square summable, then the iterates of QGMs converge to the set of optimal values. We show how the theory presented in this paper can be applied to a resource allocation problem in electrical power grids. Finally, we numerically illustrate the performance of the algorithm.

B. Notation

Vectors and matrices are represented by boldface lower and upper case letters, respectively. The set of real and natural numbers are denoted by $\mathbb{R}$ and $\mathbb{N}$, respectively. The set of real n vectors and $n\times m$ matrices are denoted by $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$, respectively. Otherwise, we use calligraphy letters to represent sets. We let $\mathcal{S}^n=\{\mathbf{x}\in\mathbb{R}^n\big|1=||\mathbf{x}||\}$ denote the unit sphere. The superscript $(\cdot)^{\mathrm{T}}$ stands for transpose. $\mathrm{diag}(\mathbf{A}_1,\ldots,\mathbf{A}_n)$ denotes the diagonal block matrix with $\mathbf{A}_1,\ldots,\mathbf{A}_n$ on the diagonal. $||\cdot||$ denotes the 2-norm.

II. Preliminaries and Application Example

In this paper we consider optimization problems of the form

$$\underset{\mathbf{x}\in\mathbb{R}^N}{\text{minimize}} \quad f(\mathbf{x}). \tag{1}$$

We denote by $f^\star$ and $\mathcal{X}^\star$ the optimal value and the set of optimizers to Problem (1), respectively. We consider the following class of functions f:

Definition 1: Let $\mathcal{F}$ denote the set of convex and differentiable functions on $\mathbb{R}^N$ with L-Lipschitz continuous gradients where $\mathcal{X}^\star$ is nonempty and bounded.

For $f\in\mathcal{F}$ it is well known that the gradient method

$$\mathbf{x}(t+1)=\mathbf{x}(t)-\gamma(t)\nabla f(\mathbf{x}(t)), \tag{2}$$

converges to $\mathcal{X}^\star$ under appropriate step-size rules [15]. When only the gradient *direction* is known, recursion (2) becomes

$$\mathbf{x}(t+1)=\mathbf{x}(t)-\gamma(t)\frac{\nabla f(\mathbf{x}(t))}{||\nabla f(\mathbf{x}(t))||}. \tag{3}$$

(3) converges to $\mathcal{X}^\star$ under appropriate diminishing step-size rules, and for fixed step-size $\gamma(t)=\gamma$ the stopping condition $f(\mathbf{x}(t))-f^\star<\epsilon$ can be achieved for all $\epsilon>0$ [16].

Problems of the form (1) commonly appear as primal or dual master problems in distributed optimization methods such as primal or dual decomposition [4], [5]. In such methods, the gradient information needs to be communicated to perform the recursions (2) or (3). We now give an example of such a distributed procedure.

A. Application Example: Distributed Power Allocation

Consider a network consisting of N resources and M users. The generation of resource $j=1,\ldots N$ and the usage of user $i=1,\ldots,M$ are donated by $r_j\in\mathcal{R}_j\subseteq\mathbb{R}$ and $\mathbf{q}_i\in\mathcal{Q}_i\subseteq\mathbb{R}^N$, respectively. The local constraints $\mathcal{R}_j$ and $\mathcal{Q}_j$ represent generation limits and user preferences, respectively. The generation of resource j has cost function C_j, and the usage of user i has utility function U_i. The operation goal of the network is to optimize the social welfare of the system by solving the following maximization problem.

$$\begin{aligned}\underset{(\mathbf{q},\mathbf{r})\in\mathbb{R}^{MN\times N}}{\text{maximize}} \quad & \sum_{i=1}^M U_i(\mathbf{q}_i)-\sum_{j=1}^N C_j(r_j)\\ \text{subject to} \quad & \mathbf{q}_i\in\mathcal{Q}_i, \quad \text{for } i=1,\ldots,M\\ & r_j\in\mathcal{R}_j, \quad \text{for } j=1,\ldots,N\\ & \sum_{i=1}^N\mathbf{q}_i=\mathbf{r}.\end{aligned} \tag{RA}$$

For notational ease, we write $\mathbf{q}=(\mathbf{q}_1,\ldots,\mathbf{q}_M)$, $\mathbf{r}=(r_1,\ldots,r_N)$, $\mathcal{Q}=\mathcal{Q}_1\times\ldots\mathcal{Q}_M$, and $\mathcal{R}=\mathcal{R}_1\times\ldots\mathcal{R}_N$. The dual problem of (RA) is of the form given in (1) where the dual function $f:\mathbb{R}^N\to\mathbb{R}$ is given by

$$f(\mathbf{p})=\underset{(\mathbf{q},\mathbf{r})\in\mathcal{Q}\times\mathcal{R}}{\text{maximize}} L(\mathbf{q},\mathbf{r},\mathbf{p})=L(\mathbf{q}(\mathbf{p}),\mathbf{r}(\mathbf{p}),\mathbf{p}). \tag{4}$$

Here

$$L(\mathbf{q},\mathbf{r},\mathbf{p})=\sum_{i=1}^M U_i(\mathbf{q}_i)-\sum_{j=1}^N C_j(r_j)-\mathbf{p}^{\mathrm{T}}\left(\mathbf{r}-\sum_{i=1}^M\mathbf{q}_i\right),$$

and for all $i=1,\ldots,M$ and $j=1,\ldots,N$ we have

$$\mathbf{q}_i(\mathbf{p})=\underset{\mathbf{q}_i\in\mathcal{Q}_i}{\text{argmax}}\ U_i(\mathbf{q}_i)-\mathbf{p}^{\mathrm{T}}\mathbf{q}_i, \tag{5}$$

$$r_j(p_j)=\underset{r_j\in\mathcal{R}_j}{\text{argmax}}\ -C_i(r_i)+p_jr_j. \tag{6}$$

The following result is proved in the appendix of [17].

Lemma 1: Suppose there exists $\mu>0$ such that U_i and C_j are μ-strongly concave and μ-strongly convex, respectively, for all $i=1,\ldots,M$ and $j=1,\ldots,n$ and $\mathcal{Q}_i$ and $\mathcal{R}_i$ are convex and compact sets. Then, f is continuously differentiable on $\mathbb{R}^N$ and the gradient $\nabla f(\mathbf{p})=\mathbf{r}(\mathbf{p})-\sum_{i=1}^M\mathbf{q}_i(\mathbf{p})$ is $(M+1)/\mu$-Lipschitz continuous.

By Lemma 1 the update rules (2) and (3) apply here. However, the gradient information must be broadcasted to the users so they can solve their subproblems (5) and (6). In many applications, the gradient $\nabla f(\mathbf{p})=\mathbf{r}(t)-\sum_{i=1}^M\mathbf{q}_i(t)$, i.e., the amount used of each resource at time t, can be measured at the source. Hence, using only one-way communication is feasible.

III. Quantized Gradient Descent Methods

We consider general quantized gradient methods of the form

$$\mathbf{x}(t+1)=\mathbf{x}(t)-\gamma(t)\mathbf{d}(t), \tag{7}$$

with $\mathbf{d}(t)\in\mathcal{D}\subseteq\mathcal{S}^N$, where $\mathcal{D}$ is a finite set of quantized gradient directions. Clearly, we have the following relation between the cardinality of $\mathcal{D}$ and communicated bits at each iteration of (7).

Remark 1: The set $\mathcal{D}$ can be coded using $\log_2(|\mathcal{D}|)$ bits. We investigate what fundamental properties the set $\mathcal{D}$ needs for the recursion (7) to minimize any $f\in\mathcal{F}$. To formally assert the meaning of such a proper quantization we make the following definition.

Definition 2: A set $\mathcal{D}$ is a *proper quantization* if for every $f \in \mathcal{F}$ and every initialization $\mathbf{x}(0) \in \mathbb{R}^N$ we can choose $\mathbf{d}(t) \in \mathcal{D}$ and $\gamma(t) \in \mathbb{R}_+$, for all $t \in \mathbb{N}$, in the recursion (7) such that

$$\lim_{t\to\infty} \texttt{dist}(\mathbf{x}(t), \mathcal{X}^\star) = 0. \tag{8}$$

In other words, every limit point of $\mathbf{x}(t)$ is in $\mathcal{X}^\star$.

Definition 2 is not constructive since its validation requires testing the dynamics (7) on every function $f \in \mathcal{F}$. In addition, it does not provide an algorithm based on (7); it does not provide a strategy for choosing $\mathbf{d}(t)$ even when $\mathcal{D}$ is known to be a *proper quantization*. Moreover, Definition 2 gives no insight into other interesting properties of the quantization set $\mathcal{D}$. For example, it gives no information regarding the minimal size of $|\mathcal{D}|$ which ensures a proper quantization or how the structure of $\mathcal{D}$ affects the convergence behavior of potential quantized gradient methods. All of the points mentioned above are investigated in this paper. We specifically provide solutions to the following questions:

A) What are equivalent constructive conditions for the set $\mathcal{D}$ being *proper quantization* that can be used to determine whether $\mathcal{D}$ is a *proper quantization* or to construct such sets?
B) Given a proper quantization $\mathcal{D}$, how can we construct an algorithm from (7) such that $\lim_{t\to\infty} \texttt{dist}(\mathbf{x}(t), \mathcal{X}^\star) = 0$, i.e., choose the proper $\mathbf{d}(t) \in \mathcal{D}$ $\gamma(t) \in \mathbb{R}_+$?
C) What are the connections between the fineness of the quantization, i.e., the size of $|\mathcal{D}|$, to the possible convergence of the algorithm?
D) What is the minimal quantization, i.e., size $|\mathcal{D}|$, for which $\mathcal{D}$ is a *proper quantization*?

In the following subsections we answer each of the questions above, but refer to later sections for many of the technical details.

A. θ-Covers: Solution to Question A)

We now provide conditions that are equivalent to $\mathcal{D}$ being a proper quantization (Definition 2) but are *constructive* in the sense that they can be used to determine if a set $\mathcal{D}$ is a proper quantization or to construct such $\mathcal{D}$. We then use this condition to provide examples of proper quantization sets $\mathcal{D}$.

Definition 3: We say that the set $\mathcal{D}$ is a θ-cover if $\theta > 0$ and for every $\mathbf{g} \in \mathcal{S}^N$ there exists $\mathbf{d} \in \mathcal{D}$ such that

$$\cos(\texttt{ang}(\mathbf{g}, \mathbf{d})) \geq \theta. \tag{9}$$

We say that the θ-cover $\mathcal{D}$ is tight if there exists a vector $\mathbf{g} \in \mathcal{S}^R$ such that $\max_{\mathbf{d}\in\mathcal{D}} \cos(\texttt{ang}(\mathbf{g}, \mathbf{d})) = \theta$.

The following result asserts the equivalence between Definitions 2 and 3.

Theorem 1: Consider a quantization set $\mathcal{D}$. $\mathcal{D}$ is a proper quantization (Definition 2) if and only if there exists $\theta > 0$ such that $\mathcal{D}$ is a θ-cover for some $\theta > 0$ (Definition 3).

Proof: Let us start by showing via a contradiction that $\mathcal{D}$ being a proper quantization implies that there exists $\theta > 0$ such that $\mathcal{D}$ is θ-cover. Suppose there does not exists $\theta > 0$ for which $\mathcal{D}$ is a θ-cover. Then, since $\mathcal{D}$ is finite, there exists $\mathbf{a} \in \mathcal{S}^N$ such that $\cos(\texttt{ang}(\mathbf{a}, \mathbf{d})) \leq 0$ for all $\mathbf{d} \in \mathcal{D}$. In particular, we have for all $\mathbf{d} \in \mathcal{D}$ that

$$\langle \mathbf{a}, \mathbf{d} \rangle = \|\mathbf{a}\|\|\mathbf{d}\| \cos(\texttt{ang}(\mathbf{a}, \mathbf{d})) \leq 0. \tag{10}$$

By choosing $\mathbf{x}(0) = \mathbf{a}$, using Recursion (7) and Cauchy-Schwarz inequality we conclude that for all $t \in \mathbb{N}$

$$||\mathbf{x}(t)|| \geq \langle \mathbf{a}, \mathbf{x}(t) \rangle = \langle \mathbf{a}, \mathbf{a} \rangle - \sum_{i=0}^{t-1} \gamma(t) \langle \mathbf{a}, \mathbf{d}(t) \rangle \geq 1,$$

where the inequality follows from that $||\mathbf{a}|| = 1$ and that for all $\mathbf{d}(t) \in \mathcal{D}$ we have $\langle \mathbf{a}, \mathbf{d} \rangle \leq 0$. If we choose $f(\mathbf{x}) = ||\mathbf{x}||$ then $f \in \mathcal{F}$ and f has the unique optimizer $\mathbf{x}^\star = \mathbf{0}$, but

$$\texttt{dist}(\mathbf{x}(t), \mathcal{X}^\star) = ||\mathbf{x}(t)|| \geq 1, \tag{11}$$

for all $t \in \mathbb{N}$. Since (11) holds for all $\mathbf{d}(t) \in \mathcal{D}$ and $\gamma(t) \in \mathbb{R}_+$, we can conclude that $\mathcal{D}$ is not a proper quantization.

The fact that $\mathcal{D}$ being a θ-cover implies that $\mathcal{D}$ is a proper quantization, follows from Theorem 6 in Section IV-B, where we show that for all $f \in \mathcal{F}$ we can choose $\mathbf{d}(t) \in \mathcal{D}$ and $\gamma(t) \in \mathbb{R}_+$ such that $\lim_{t\to\infty} \texttt{dist}(\mathbf{x}(t), \mathcal{X}^\star) = 0$. ■

We now provide some examples of θ-covers.

Example 1 (Minimal Example: $|\mathcal{D}_1| = N + 1$): Set

$$\mathcal{D}_1 = \{\mathbf{e}_1, \ldots, \mathbf{e}_N, -\mathbf{1}/\sqrt{N}\}, \tag{12}$$

where $\mathbf{e}_i$ is the i-th element of the normal basis and $\mathbf{1}$ is N dimensional vector with 1 in every component. Clearly, $|\mathcal{D}| = N + 1$ and therefore $\mathcal{D}$ can be coded using only $\log_2(N + 1)$ bits. We show in Section III-C that this is a minimal quantization, since in general if $|\mathcal{D}| \leq N$, then $\mathcal{D}$ cannot be a proper quantization. We show in Lemma 4 in [17], that $\mathcal{D}_1$ is a θ-cover with

$$\theta = \frac{1}{\sqrt{N^2 + 2\sqrt{N}(N-1)}}. \tag{13}$$

Example 2 (Example in $\mathbb{R}^2$: $|\mathcal{D}_2| = n$): For every $n \in \mathbb{N}$ set

$$\mathcal{D}_n = \left\{ \begin{bmatrix} \cos(2\pi k/n) \\ \sin(2\pi k/n) \end{bmatrix} \in \mathbb{R}^2 \middle| k = 0, 1, \ldots, n-1 \right\}.$$

Clearly, if $n \geq 3$, $\mathcal{D}_n$ is a θ-cover with $\theta = \cos(\pi/n)$.

Example 3 ($\pm$ Normal Basis: $|\mathcal{D}_3| = 2N$): Let $\mathcal{D}_3 = \{\mathbf{e}_1, -\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_2, \ldots, \mathbf{e}_N, -\mathbf{e}_N\}$. Clearly, $|\mathcal{D}_3| = 2N$ and hence $\log_2(2N)$ bits are needed to broadcast the quantized gradient direction. Let us now show that $\mathcal{D}_3$ is θ-cover with $\theta = 1/\sqrt{N}$. Take $\mathbf{x} \in \mathcal{S}^N$, then if we choose $\mathbf{d} = \texttt{sign}(\mathbf{x}_i)\mathbf{e}_i$ where $i = \text{argmax}_{i=1,\ldots,N} |\mathbf{x}_i|$ then it holds that $\cos(\texttt{ang}(\mathbf{x}, \mathbf{d})) = \langle \mathbf{x}, \mathbf{d} \rangle = \mathbf{x}_i \cdot \texttt{sign}(\mathbf{x}_i) = |\mathbf{x}_i| \geq 1/\sqrt{N}$.

Example 4 (Signs of the gradients: $|\mathcal{D}_4|$=2^N, θ=$1/\sqrt{N}$): Let $\mathcal{D} = \{(1/\sqrt{N})(e_1, e_2, \ldots, e_N) \big| e_i \in \{-1, 1\}\}$. Here, each $\mathbf{d} \in \mathcal{D}$ represents one orthant of $\mathbb{R}^N$. Therefore, this choice is well suited when the sources cannot cooperate and each source updates its price based on local estimates of their part of the gradient, i.e., $\mathbf{d}(t) = \texttt{sign}(\nabla f(\mathbf{p}(t)))$. It can be checked that $|\mathcal{D}| = 2^N$ hence $\log_2(2^N) = N$ bits are needed to broadcast the quantized gradient direction. To

show that $\mathcal{D}_4$ is θ-cover with $\theta = 1/\sqrt{N}$, take any $\mathbf{x} \in \mathcal{S}^N$. Then it holds for $\mathbf{d} = (1/\sqrt{N})\texttt{sign}(\mathbf{x})$ that

$$\cos(\texttt{ang}(\mathbf{x}, \mathbf{d})) = \langle \mathbf{x}, \mathbf{d} \rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbf{x}_i \cdot \texttt{sign}(\mathbf{x}_i)$$
$$\geq \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbf{x}_i^2 = \frac{1}{\sqrt{N}} ||\mathbf{x}|| = \frac{1}{\sqrt{N}}.$$

B. Algorithm: Solution to Questions B) and C)

Unlike Definition 2, Definition 3 actually provides us with tools to construct algorithms for solving Problem (1) when $\mathcal{D}$ is a proper quantization. In particular, for $\mathbf{x}(t) \in \mathbb{R}^N$ we can quantize the gradient $\nabla f(\mathbf{x}(t))$ with a $\mathbf{d}(t) \in \mathcal{D}$ such that $\cos(\texttt{ang}(\nabla f(\mathbf{x}(t)), \mathbf{d}(t))) \geq \theta$, as seen in Algorithm 1. We study the convergence of Algorithm 1 in Section IV. In particular, we provide a bound on the number of iterations needed to achieve a specified accuracy that decreases as θ becomes closer to 1. Moreover, we show how to choose the step-sizes so that any limit point of the algorithm is an optimizer of Problem (1).

Algorithm 1: θ-Quantized Gradient Methods (θ-QGM)

Initialization: *Choose* $\mathbf{x}(0) \in \mathbb{R}^N$;
for $t = 0, 1 \ldots$ **do**
  Quantized Gradient: *Choose* $\mathbf{d}(t) \in \mathcal{D}$ *such that*

$$\cos(ang(\nabla f(\mathbf{x}(t)), \mathbf{d}(t))) \geq \theta$$

  Gradient Step: $\mathbf{x}(t+1) = \mathbf{x}(t) - \gamma(t)\mathbf{d}(t)$

C. Minimal Quantization: Solution to Question D)

In Example 1 we provided a proper quantization $\mathcal{D}$ where $|\mathcal{D}| = N + 1$. We now show that $N + 1$ is a minimal proper quantization in the sense that there does not exist a proper quantization set $\mathcal{D}$ with cardinality less than $N + 1$.

Theorem 2: Suppose that $|\mathcal{D}| \leq N$. Then $\mathcal{D}$ is not a proper quantization (Defintion 2).

Proof: First consider the case where either $|\mathcal{D}| < N$ or $|\mathcal{D}| = N$ and the elements of $\mathcal{D}$ are linearly dependent. Then $\texttt{Span}(\mathcal{D})$ is a proper subspace of $\mathbb{R}^N$, so there exists a normal $\mathbf{a} \in \mathcal{S}^N$ such that $\cos(\texttt{ang}(\mathbf{a}, \mathbf{d})) = \langle \mathbf{a}, \mathbf{d} \rangle \leq 0$ for all $\mathbf{d} \in \texttt{Span}(\mathcal{D})$. Since $\mathcal{D} \subseteq \texttt{Span}(\mathcal{D})$, $\mathcal{D}$ is not a θ-cover for any $\theta > 0$ and the result follows from Theorem 1.

Let us next consider the other case, where $|\mathcal{D}| = N$ and the vectors of $\mathcal{D}$ are linearly independent, i.e., $\texttt{Span}(\mathcal{D}) = \mathbb{R}^N$. Define $\mathbf{D} \in \mathbb{R}^{N \times N}$ such that for $i = 1, \ldots, N$ row i in $\mathbf{D}$ is the i-th elemnt of $\mathcal{D}$, where the elements have some arbitrary order. Then $\mathbf{D}$ is invertible and we can choose $\mathbf{a} = \mathbf{D}^{-1}(-\mathbf{1})$ where $\mathbf{1} \in \mathbb{R}^N$ is a vector of all ones. Then we have for $i = 1, \ldots, N$ that $\langle \mathbf{d}_i, \mathbf{a} \rangle = -\mathbf{d}_i \mathbf{D}^{-1} \mathbf{1} = -1$. Hence, as in the previous case, we get that $\langle \mathbf{a}, \mathbf{d} \rangle \leq 0$ for all $\mathbf{d} \in \mathcal{D}$ implying that $\mathcal{D}$ can not be a θ-cover for any $\theta > 0$, and the result follows from Theorem 1. ∎

We now study the convergence of the Algorithm 1.

IV. Convergence

We first investigate the convergence Algorithm 1 when the step-sizes are fixed, i.e., $\gamma(t) = \gamma$, in subsection IV-A. Then in subsection IV-B we consider diminishing step-size.

A. Constant Step Size

For constant step-size we consider the following two types of stopping criteria:

$$\text{Type-1:} \qquad ||\nabla f(\mathbf{x})|| < \epsilon, \tag{14}$$
$$\text{Type-2:} \qquad f(\mathbf{x}(t)) - f^\star < \epsilon. \tag{15}$$

We note that when performing primal or dual decomposition usually only the gradient of f is available but the objective function is distributed between different users, hence (14) tends to be a more practical stopping condition.

1) Stopping Condition of Type-1*:* We start by showing that the Type-1 stopping criterion can be achieved for any $\epsilon > 0$ in finitely many iterations. Further, we provide a bound on number of iterations that depends on θ outlined next.

Theorem 3: For $\epsilon > 0$ we define the set

$$\mathcal{X}(\epsilon) = \{\mathbf{x} \in \mathbb{R}^N \big| ||\nabla f(\mathbf{x})|| \leq \epsilon\}. \tag{16}$$

If $f \in \mathcal{F}$, $\mathcal{D}$ is a θ-cover, and the sequence $(\mathbf{x}(t))_{t \in \mathcal{N}}$ is generated using Algorithm 1, then the following holds:

a) For any $\epsilon > 0$, if $\gamma \in (0, 2\theta\epsilon/L)$ then there exists $T \in \mathbb{N}$ such that $\mathbf{x}(T) \in \mathcal{X}(\epsilon)$, with T bounded by

$$T \leq \left\lceil \frac{2(f(\mathbf{x}(0)) - f^\star)}{\gamma(2\theta\epsilon - L\gamma)} \right\rceil. \tag{17}$$

The upper bound (17) is minimized with the optimal step size $\gamma^\star = \theta\epsilon/L$.

b) For any step size $\gamma > 0$ and scalar $\kappa > 0$, if we choose

$$\epsilon(\kappa, \gamma) = \kappa + \gamma L/(2\theta) \tag{18}$$

then there exists $T \in \mathbb{N}$ such that $\mathbf{x}(T) \in \mathcal{X}(\epsilon(\kappa, \gamma))$, with T bounded by

$$T \leq \left\lceil \frac{f(\mathbf{x}(0)) - f^\star}{\theta\gamma\kappa} \right\rceil. \tag{19}$$

c) (Lower Bound on T) For any step-size $\gamma > 0$ and $\epsilon > 0$ if $\mathbf{x}(T) \in \mathcal{X}(\epsilon)$ then

$$\frac{||\nabla f(\mathbf{x}(0))|| - \epsilon}{\gamma L} \leq T \tag{20}$$

Proof: a) Let $\epsilon > 0$ be given and choose any $\gamma \in (0, 2\theta\epsilon/L)$. From Lemma 2 below we have for all $\mathbf{x}(t) \in \mathbb{R}^N \setminus \mathcal{X}(\epsilon)$ that

$$f(\mathbf{x}(t+1)) \leq f(\mathbf{x}(t)) - \delta(\epsilon, \gamma, \theta), \tag{21}$$

where $\delta(\epsilon, \gamma, \theta) > 0$ is defined in (27). By recursively using (21), it follows that if $\mathbf{x}(t) \in \mathbb{R}^N \setminus \mathcal{X}(\epsilon)$ for all $t < s$ then

$$f(\mathbf{x}(s)) \leq f(\mathbf{x}(0)) - s\ \delta(\epsilon, \gamma, \theta). \tag{22}$$

Therefore, there must exist $T \leq \lceil (f(\mathbf{x}(0)) - f^\star)/\delta(\epsilon, \gamma, \theta) \rceil$ such that $\mathbf{x}(T) \in \mathcal{X}(\epsilon)$; otherwise, we can use (21) with $s = \lceil (f(\mathbf{x}(0)) - f^\star)/\delta(\epsilon, \gamma, \theta) \rceil + 1$ to get the contradiction that $f(\mathbf{x}(s)) < f^\star$, which cannot be true since $f^\star$ is the optimal

solution to (1). By rearranging $\lceil (f(\mathbf{x}(0)) - D^\star)/\delta(\epsilon,\gamma,\theta)\rceil$, we get (17). The optimal step-size $\gamma^\star = \theta\epsilon/L$ comes by simply maximizing the denominator in (17).

b) This result can be obtained by using similar arguments as were used to prove part a). The only difference is that now we have an explicit form for ϵ when using (21), which results in

$$\delta(\epsilon,\gamma,\theta) = \delta\left(\kappa + \frac{\gamma L}{2\theta},\gamma,\theta\right) = \theta\gamma\kappa. \tag{23}$$

c) Using that the gradient ∇f is L-Lipschitz continuous and the triangle inequality, we have for all $t \in \mathbb{N}$ that

$$||\nabla f(\mathbf{x}(t))|| - L\gamma \leq ||\nabla f(\mathbf{x}(t+1))||. \tag{24}$$

Recursively applying (24) gives

$$||\nabla f(\mathbf{x}(0))|| - L\gamma t \leq ||\nabla f(\mathbf{x}(t))||. \tag{25}$$

Hence, from (25), $||\nabla f(\mathbf{p}(t))|| \leq \epsilon$ can only hold when $t \geq (||\nabla f(\mathbf{p}(0))|| - \epsilon)/(L\gamma)$. ■

Lemma 2: Let $f : \mathbb{R}^N \to \mathbb{R}$ be a convex and continuously differentiable function with L-continuous gradient. Suppose $\epsilon > 0$, $\gamma \in (0, 2\theta\epsilon/L)$, $\theta \in (0,1]$, $\mathbf{x} \in \mathbb{R}^N$, and $\mathbf{d} \in \mathcal{S}^N$ where $\cos(\texttt{ang}(\nabla f(\mathbf{x}), \mathbf{d})) \geq \theta$ and $||\nabla f(\mathbf{x})|| > \epsilon$. Then

$$f(\mathbf{x} - \gamma\mathbf{d}) \leq f(\mathbf{x}) - \delta(\epsilon,\gamma,\theta), \tag{26}$$

where

$$\delta(\epsilon,\gamma,\theta) = -\left(\frac{L}{2}\gamma - \theta\epsilon\right)\gamma > 0. \tag{27}$$

Proof: See Lemma 2 in the extended version [17]. ■

Theorem 3-a) proves that if $\mathcal{D}$ is a θ-cover then the Type-1 stopping condition, (eq. (14)) can be achieved with ϵ-accuracy in finitely many iterations, for all $\epsilon > 0$. Moreover, it gives a bound on the number of iterations needed to achieve such ϵ-accuracy depending on θ, where the bound decreases as θ approaches 1. Theorem 3-b) demonstrates what ϵ-accuracy can be achieved for a given step-size. The parameter κ captures a trade off between the ϵ-accuracy and the number of iterations. In fact, by optimizing over both γ and κ in Theorem 3-b), we find an optimal bound on the accuracy ϵ that can be guaranteed in $T^{\max}$ iterations. This idea is formalized in the following Theorem.

Theorem 4: Suppose an upper bound $T^{\max} \in \mathbb{N}$ on number of iterations is given. Then the minimal bound $\epsilon(\kappa,\gamma)$ achieved from (18) in $T^{\max}$ iterations is

$$\epsilon^\star = \frac{L}{\theta}\sqrt{\frac{2(f(\mathbf{x}(0)) - f^\star)}{LT^{\max}}}, \tag{28}$$

where the corresponding optimal γ and κ are

$$\gamma^\star = \sqrt{\frac{2(f(\mathbf{x}(0)) - f^\star)}{LT^{\max}}} \text{ and } \kappa^\star = \frac{\sqrt{L(f(\mathbf{x}(0)) - f^\star)}}{\theta\sqrt{2T^{\max}}}. \tag{29}$$

In other words, equations (28) and (29) give an optimal solution to the following optimization problem:

$$\begin{aligned} \underset{\kappa,\gamma}{\text{minimize}} \quad & \epsilon(\kappa,\gamma) = \kappa + (L/(2\theta))\gamma \\ \text{subject to} \quad & \frac{f(\mathbf{x}(0)) - f^\star}{\theta\gamma\kappa} \leq T^{\max}, \\ & \gamma,\kappa > 0. \end{aligned} \tag{30}$$

Proof: See Theorem 4 in the extended version [17]. ■

We next demonstrate how the convergence results translate to Type-2 stopping conditions (eq. (15)).

2) Stopping Condition of Type-2*:* We have the following result.

Theorem 5: Suppose $f \in \mathcal{F}$, $\mathcal{D}$ is a θ-cover, and the iterates $\mathbf{x}(t)$ come from Algorithm 1, then following holds:

a) For any $\epsilon > 0$, $\gamma \in (0, 2\theta\epsilon/L)$ and $T \in \mathbb{N}$ such that $\mathbf{x}(T) \in \mathcal{X}(\epsilon)$ it holds, for all $t \geq T$, that

$$f(\mathbf{x}(t)) \leq F(\epsilon) + \left(\epsilon + \frac{L}{2}\gamma\right)\gamma, \tag{31}$$

where $F : \mathbb{R}_+ \to \mathbb{R} \cup \{\infty\}$ is given by

$$F(\kappa) = \sup\{f(\mathbf{x}) | \mathbf{x} \in \mathcal{X}(\kappa)\}. \tag{32}$$

There exists $\bar{\kappa} > 0$ such that $F(\kappa) < \infty$ for all $\kappa < \bar{\kappa}$ and $\lim_{\kappa\to 0^+} F(\kappa) = f^\star$.

b) If f is μ-strongly convex then we have

$$F(\epsilon) \leq f^\star + \epsilon^2/(2\mu). \tag{33}$$

Proof: See Theorem 5 in the extended version [17]. ■

Theorem 5 yields the following immediate corollary.

Corollary 1: For any $\epsilon > 0$ there exists step-size $\gamma > 0$ and $T \in \mathbb{N}$ such that $f(\mathbf{x}(T)) - f^\star < \epsilon$. Moreover, if f is μ-strongly convex and $\epsilon_1, \gamma > 0$ are chosen such that $\gamma \in (0, 2\theta\epsilon_1)$ and $\epsilon_1/(2\mu) + (\epsilon_1 + L\gamma/2)\gamma < \epsilon$ then

$$T \leq \left\lceil \frac{2(f(\mathbf{x}(0)) - f^\star)}{\gamma(2\theta\epsilon_1 - L\gamma)} \right\rceil. \tag{34}$$

These results prove that the Type-2 stopping condition (eq. (15)) can be achieved in finitely many iterations. Moreover, when f is strongly concave the results provide a bound on the number of iterations..

B. Diminishing Step Size

We now consider the diminishing step-size case.

Theorem 6: Suppose that $f \in \mathcal{F}$, $\mathcal{X}^\star$ is bounded, $\mathcal{D}$ is a θ-cover, and the sequence $(\mathbf{x}(t))_{t\in\mathcal{N}}$ is generated using Algorithm 1. If the step-size $\gamma(t)$ is non-summable and square summable, i.e.,

$$\sum_{t=0}^{N} \gamma(t) = \infty \quad \text{and} \quad \sum_{t=0}^{N} \gamma(t)^2 < \infty, \tag{35}$$

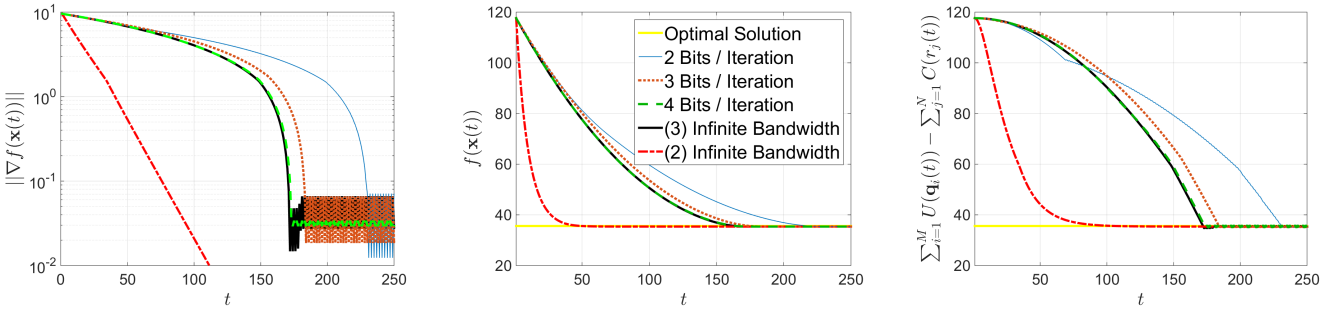
then $\lim_{t\to\infty} \texttt{dist}(\mathbf{x}(t), \mathcal{X}^\star) = 0$.

Proof: See Theorem 6 in the extended version [17]. ■

Theorem 6 shows that when $\mathcal{D}$ is a θ-cover then there exists a step-size rule such that every limit point of the quantized gradient methods is an optimal solution to (1). We next numerically illustrate the quantized gradient methods.

V. Simulation Results

We illustrate the performance of the quantized gradient methods on an instance of (RA) from Section II-A with M=4 users and N=2 resources. For i=1, 2, 3, 4 and $j = 1, 2$, we set $C_j(r_j) = -c_j r_j^2$, $U_i(\mathbf{q}_i) = a_{i1}\log(0.1 + q_{i1}) + a_{i2}\log(0.1 + q_{i2})$, $\mathcal{R}_j = [0, 10]$ and $\mathcal{Q}_i = \{(x,y) \in \mathbb{R}^2 | x, y \geq 0, x + y \leq 3\}$. Clearly, C_j and U_i are strongly concave on their domains with



(a) Gradient (b) Dual Objective Function Value (c) Primal Objective Function Value

Fig. 1: Gradient and objective function value over the course of the algorithm.

concavity parameters c_j and $\mu_i = \min\{a_{i1}, a_{i2}\}/3.1^2 = \min\{a_{i1}, a_{i2}\}/9.61$, respectively. By Lemma 1, the dual gradient is L-Lipschitz continuous with $L = 5/\mu$, where $\mu = \min\{c_1, c_2, \mu_1, \ldots, \mu_4\}$. We choose $c_1 = c_2 = 1$, and $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4 = (50, 14), (20, 11), (40, 12), (35, 10)$.

Figure 1 depicts the results when the step-size is $\gamma = 0.1$ and the initialization is $\mathbf{x}(0) = (0, 0)$ (recall that $\mathbf{x}$ is the dual variable here). We use the quantization set $\mathcal{D}$ from Example 2, which corresponds to the case where 2, 3, and 4 bits are communicated per iteration, i.e., $|\mathcal{D}| = 4, 8, 16$, see Remark 1. The norm of the gradient $||\nabla f||$ reaches the accuracy $\epsilon = 0.1$ in roughly 140, 180, and 240 using 560, 540, 480 bits when 2, 3, and 4 bits are communicated per iteration, respectively. We compare the results to recursions (2) and (3) where no quantization is done, i.e., infinite bandwidth is used. Figure 1a shows that by using 4 bits per iteration, the results achieved by QGM are almost as good as when the full gradient direction is communicated in (3). However, the QGMs do not perform as well as (2); this is to be expected, since in (2) the full direction and magnitude of the gradient is known. Our results illustrate that we can dramatically reduce the number of bits communicated without sacrificing much in performance.

VI. Conclusions and Future Work

In this paper quantized gradient methods for distributed optimization were investigated. Necessary and sufficient conditions were provided that ensure that the quantized methods can minimize any function with Lipschitz continuous gradient. These conditions were used to provide a lower bound on the number of bits needed to quantizes the gradients. Moreover, the results demonstrated how the number of iterations needed to achieve the desired solution accuracy are related to the fineness of the quantization. These results can be used to provide a bound on the number of bits needed to solve an optimization problem up to a desired accuracy. Future research directions will consider problems with non-differentiable objective functions and constraints. Moreover, for given θ it is desirable to find the optimal quantization, i.e., the one that can be coded using the fewest number of bits.

References

[1] H. C. Ferreira, H. M. Grové, O. Hooijen, and A. Han Vinck, *Power line communication*. Wiley Online Library, 2001.

[2] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *The Journal of the Operational Research Society*, vol. 49, no. 3, pp. pp. 237–252, 1998. [Online]. Available: http://www.jstor.org/stable/3010473

[3] S. Low and D. Lapsley, "Optimization flow control. I. basic algorithm and convergence," *Networking, IEEE/ACM Transactions on*, vol. 7, no. 6, pp. 861–874, Dec 1999.

[4] D. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 8, pp. 1439–1451, Aug 2006.

[5] ——, "Alternative distributed algorithms for network utility maximization: Framework and applications," *Automatic Control, IEEE Transactions on*, vol. 52, no. 12, pp. 2254–2269, Dec 2007.

[6] N. Li, L. Chen, and S. H. Low, "Optimal demand response based on utility maximization in power networks," in *Power and Energy Society General Meeting, 2011 IEEE*, 2011, pp. 1–8.

[7] L. Chen, N. Li, S. H. Low, and J. C. Doyle, "Two market models for demand response in power networks," *IEEE SmartGridComm*, vol. 10, pp. 397–402, 2010.

[8] R. Madan and S. Lall, "Distributed algorithms for maximum lifetime routing in wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 8, pp. 2185–2193, Aug 2006.

[9] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *Automatic Control, IEEE Transactions on*, vol. 59, no. 5, pp. 1177–1189, May 2014.

[10] M. Rabbat and R. Nowak, "Quantized incremental algorithms for distributed optimization," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 4, pp. 798–808, April 2005.

[11] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. Tsitsiklis, "Distributed subgradient methods and quantization effects," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, Dec 2008, pp. 4177–4184.

[12] J. Herdtner and E. Chong, "Analysis of a class of distributed asynchronous power control algorithms for cellular wireless systems," *Selected Areas in Communications, IEEE Journal on*, vol. 18, no. 3, pp. 436–446, March 2000.

[13] A. Nedic and D. P. Bertsekas, "Incremental subgradient methods for nondifferentiable optimization," *SIAM Journal on Optimization*, vol. 12, no. 1, pp. 109–138, 2001.

[14] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *Automatic Control, IEEE Transactions on*, vol. 54, no. 1, pp. 48–61, Jan 2009.

[15] D. P. Bertsekas, *Nonlinear Programming: 2nd Edition*. Athena Scientific, 1999.

[16] N. Z. Shor, *Minimization Methods for Non-Differentiable Functions*. Springer, 1985.

[17] S. Magnússon, E. C. L. N. F. C. Heal, K., and V. Tarokh, "Convergence of Limited Communications Gradient Methods," *ArXiv e-prints*, Mar. 2016, [Online]. Available: http://adsabs.harvard.edu/abs/2016arXiv160300316M.