

# Distributed Optimization with Finite Bit Adaptive Quantization for Efficient Communication and Precision Enhancement

Apostolos I. Rikos, Wei Jiang, Themistoklis Charalambous, and Karl H. Johansson

**Abstract**—In realistic distributed optimization scenarios, individual nodes possess only partial information and communicate over bandwidth constrained channels. For this reason, the development of efficient distributed algorithms is essential. In our paper we address the challenge of unconstrained distributed optimization. In our scenario each node's local function exhibits strong convexity with Lipschitz continuous gradients. The exchange of information between nodes occurs through 3-bit bandwidth-limited channels (i.e., nodes exchange messages represented by a only 3-bits). Our proposed algorithm respects the network's bandwidth constraints by leveraging zoom-in and zoom-out operations to adjust quantizer parameters dynamically. We show that during our algorithm's operation nodes are able to converge to the exact optimal solution. Furthermore, we show that our algorithm achieves a linear convergence rate to the optimal solution. We conclude the paper with simulations that highlight our algorithm's unique characteristics.

## I. INTRODUCTION

Distributed optimization has become increasingly important in various domains. Such domains range from large-scale machine learning [1], control [2], to other data-driven applications [3] that involve large amounts of data. In the distributed setting multiple nodes collaborate to collectively solve optimization problems with each node possessing only partial information about the global objective. Distributing the optimization process offers advantages in terms of scalability, fault tolerance, and adaptability to dynamic environments [4].

In current literature, numerous algorithms have been developed for distributed optimization [5]–[8]. These algorithms

Apostolos I. Rikos is with the Artificial Intelligence Thrust of the Information Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China. He is also affiliated with the Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China. E-mail: apostolosr@hkust-gz.edu.cn.

Wei Jiang was with the Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University, Espoo, Finland. Email: wjiang.lab@gmail.com.

Themistoklis Charalambous is with the Department of Electrical and Computer Engineering, School of Engineering, University of Cyprus, 1678 Nicosia, Cyprus. E-mail: charalambous.themistoklis@ucy.ac.cy. T. Charalambous is also with the Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University, Espoo, Finland. Email: themistoklis.charalambous@aalto.fi.

Karl H. Johansson is with the Division of Decision and Control Systems, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden, also affiliated with Digital Futures. E-mails: kallej@kth.se.

Part of this work was supported by the Knut and Alice Wallenberg Foundation, the Swedish Research Council, and the Swedish Foundation for Strategic Research. The work of T. Charalambous was partly supported by the European Research Council (ERC) Consolidator Grant MINERVA (Grant agreement No. 101044629).

typically necessitate nodes to exchange real-valued messages with high or infinite precision, representing local gradients or optimization variables. However, exchanging messages with infinite precision is often impractical or even impossible due to bandwidth constraints, as it requires the transmission of an infinite number of bits per second. Consequently, such algorithms tend to incur significant communication overhead, particularly in scenarios involving bandwidth-limited channels.

Due to the aforementioned limitations, there is a growing interest for distributed optimization algorithms that can operate efficiently under communication constraints. Specifically, there has been a surge of interest in distributed optimization approaches where nodes exchange compressed packets or quantized messages [9]–[21]. These approaches aim to reduce communication overhead while still enabling nodes to converge to the optimal solution or approximate it with high accuracy that depends on the utilized quantization level. However, in scenarios (such as resource-constrained networks or low-power devices) where communication among nodes is limited to only a small fixed amount of bits (e.g., 3, 4, or 5-bits per message), the aforementioned approaches face limitations. Algorithms that respect this constraint usually converge to a neighborhood of the optimal solution leading to an error floor that may be considerably large [22]. Refining the quantization level allows for approximation with higher precision (potentially enabling convergence to the exact optimal solution) [23]. But the refinement process increases the number of exchanged bits, leading to constraint violation and may compromise the efficiency gains achieved through quantization.

**Main Contributions.** Motivated by the aforementioned limitations and challenges, in our paper we propose a distributed optimization algorithm that exhibits efficient communication among nodes. The main advantage of our algorithm compared to the current literature is its ability to operate in scenarios where communication is limited to only a small fixed amount of bits. Additionally, due to its ability to adjust the quantizer parameters, it enables nodes to converge to the exact optimal solution. Our main contributions are the following:

- We propose a distributed optimization algorithm that exhibits efficient communication enabling nodes to exchange messages consisting only 3-bits; see Algorithm 1. Our algorithm's operation relies on zoom-in and zoom-out operations according to a set of conditions. These operations adjust the quantizer parameters (such as basis and range) to focus on the estimated

optimal solution while ensuring approximation with higher precision.

- We show that our algorithm's operation, enables nodes to estimate the *exact* optimal solution exhibiting linear convergence rate (thus avoiding convergence to a neighborhood of the optimal solution due to the quantization effect); see Theorem 1.

## II. NOTATION AND BACKGROUND

**Mathematical Notation and Symbols.** The symbols  $\mathbb{R}, \mathbb{Q}, \mathbb{Z}$ , and  $\mathbb{N}$  represent the sets of real, rational, integer, and natural numbers, respectively. The symbols  $\mathbb{Q}_{>1}, \mathbb{R}_{\geq 0}$  denote the set of rational numbers greater than one, and the set of positive real numbers, respectively. The symbol  $\mathbb{R}_{\geq 0}$  indicates the nonnegative orthant of the  $n$ -dimensional real space  $\mathbb{R}^n$ . Matrices are denoted by capital letters (e.g.,  $A$ ), and vectors by small letters (e.g.,  $x$ ). The transpose of a matrix  $A$  and a vector  $x$  are denoted as  $A^\top, x^\top$ , respectively. For any real number  $a \in \mathbb{R}$ ,  $\lfloor a \rfloor$  denotes the greatest integer less than or equal to  $a$ , and  $\lceil a \rceil$  denotes the least integer greater than or equal to  $a$ . For a matrix  $A \in \mathbb{R}^{n \times n}$ ,  $a_{ij}$  denotes the entry in row  $i$  and column  $j$ . The all-ones vector is denoted by  $\mathbb{1}$ , and the identity matrix by  $\mathbb{I}$  of appropriate dimensions. The Euclidean norm of a vector is represented by  $\|\cdot\|$ .

**Graph-Theoretic Notions.** The communication network is modeled as a directed graph (digraph), denoted by  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ . This digraph comprises  $n$  nodes ( $n \geq 2$ ), communicating solely with their immediate neighbors, and remains static over time. In  $\mathcal{G}$ , the node set is represented as  $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ , and the edge set as  $\mathcal{E} \subseteq \{\mathcal{V} \times \mathcal{V}\} \cup \{(v_i, v_i) \mid v_i \in \mathcal{V}\}$  (each node has a virtual self-edge). The number of nodes and edges are denoted by  $|\mathcal{V}| = n$  and  $|\mathcal{E}| = m$ , respectively. A directed edge from node  $v_i$  to node  $v_l$  is denoted by  $(v_l, v_i) \in \mathcal{E}$ , signifying that node  $v_l$  can receive information from node  $v_i$  at time step  $k$  (but not vice versa). The in-neighbors of  $v_i$  are the nodes that can directly transmit information to  $v_i$ , denoted by  $\mathcal{N}_i^- = \{v_j \in \mathcal{V}; \mid (v_i, v_j) \in \mathcal{E}\}$ . The out-neighbors of  $v_i$  are the nodes that can directly receive information from  $v_i$ , represented by  $\mathcal{N}_i^+ = \{v_l \in \mathcal{V}; \mid (v_l, v_i) \in \mathcal{E}\}$ . The in-degree and out-degree of  $v_j$  are denoted by  $\mathcal{D}_i^- = |\mathcal{N}_i^-|$  and  $\mathcal{D}_i^+ = |\mathcal{N}_i^+|$ , respectively. The diameter  $D$  of a digraph is the longest shortest path between any two nodes  $v_l, v_i \in \mathcal{V}$ . A directed path from  $v_i$  to  $v_l$  of length  $t$  exists if there's a sequence of nodes  $i \equiv l_0, l_1, \dots, l_t \equiv l$  such that  $(l_{\tau+1}, l_\tau) \in \mathcal{E}$  for  $\tau = 0, 1, \dots, t-1$ . A digraph is strongly connected if there's a directed path from every node  $v_i$  to every node  $v_l$  for all  $v_i, v_l \in \mathcal{V}$ .

**Node Operation.** During the operation of our proposed distributed optimization algorithm each node  $v_i \in \mathcal{V}$  executes a local optimization step and then, a distributed coordination algorithm. In the distributed optimization algorithm, at each time step  $k$ , each node  $v_i$  maintains (i) its local estimate variable  $x_i^{[k]} \in \mathbb{Q}$  for computing the optimal solution, (ii) a learning rate  $\alpha \in \mathbb{R}$  for optimizing its local function, (iii) the quantizer basis  $b_q \in \mathbb{Q}$  that can be used for shifting

and focusing on different ranges of values, (iv) the quantizer level  $\Delta^{[\nu_{\text{total}}]} \in \mathbb{Q}$  that can be quantizing its local estimate variable ( $\nu_{\text{total}}$  is explained below), (v) a zoom-in constant  $c^{\text{in}} \in \mathbb{Q}_{>1}$  for reducing the quantization level in order to estimate the optimal solution more precisely, (vi) a zoom-in counter  $\nu_{\text{in}} \in \mathbb{N}$  in order to count the times nodes performed zoom-in, (vii) a zoom-out constant  $c^{\text{out}} \in \mathbb{Q}_{>1}$  for increasing the quantization level so that the quantizer range covers a wider range of values that may include the optimal solution, (viii) a zoom-out counter  $\nu_{\text{out}} \in \mathbb{N}$  in order to count the times nodes performed zoom-out, and (ix) a total zoom counter  $\nu_{\text{total}} \in \mathbb{N}$  in order to count the times nodes performed zoom operation (note that  $\nu_{\text{total}} = \nu_{\text{in}} + \nu_{\text{out}}$ ). In the coordination algorithm, at each time step  $k$ , every node  $v_i$  maintains (i) stopping variables  $M_i$  and  $m_i \in \mathbb{N}$  to determine convergence, (ii) mass variables  $y_i^{[\lambda]}, z_i^{[\lambda]} \in \mathbb{Q}$  for communication by sending or receiving messages with other nodes, and (iii) state variables  $y_{i,(s)}^{[\lambda]}, z_{i,(s)}^{[\lambda]}, q_{i,(s)}^{[\lambda]} \in \mathbb{Q}$  to store received messages and compute the coordination algorithm's result.

### 3-bit Mid-Rise Uniform Quantizer with Base Shifting.

A quantizer discretizes a continuous range of values into a finite set of discrete values, reducing the number of bits needed to represent information. This compression minimizes bandwidth requirements for message transmission and enhances power and computational efficiency. Quantization is vital in scenarios like wireless communication, distributed control systems, and sensor networks, where communication constraints and imperfect information exchanges occur. The most common quantizers are uniform quantizers that partition the continuous value range into equally-sized intervals. Other types of quantizers include uniform midtread and non-uniform. For the results of this paper, we will utilize a 3-bit mid-rise uniform quantizer with base shifting capabilities, defined as follows:

$$Q^{3\text{MRU}}(b_q, \xi, \Delta^{[\nu_{\text{total}}]}) = \begin{cases} b_q - \frac{7\Delta^{[\nu_{\text{total}}]}}{2}, & \xi \in (-\infty, b_q - 3\Delta^{[\nu_{\text{total}}]}) \\ b_q - \frac{5\Delta^{[\nu_{\text{total}}]}}{2}, & \xi \in [b_q - 3\Delta^{[\nu_{\text{total}}]}, b_q - 2\Delta^{[\nu_{\text{total}}]}) \\ b_q - \frac{3\Delta^{[\nu_{\text{total}}]}}{2}, & \xi \in [b_q - 2\Delta^{[\nu_{\text{total}}]}, b_q - \Delta^{[\nu_{\text{total}}]}) \\ b_q - \frac{\Delta^{[\nu_{\text{total}}]}}{2}, & \xi \in [b_q - \Delta^{[\nu_{\text{total}}]}, b_q) \\ b_q + \frac{\Delta^{[\nu_{\text{total}}]}}{2}, & \xi \in [b_q, b_q + \Delta^{[\nu_{\text{total}}]}) \\ b_q + \frac{3\Delta^{[\nu_{\text{total}}]}}{2}, & \xi \in [b_q + \Delta^{[\nu_{\text{total}}]}, b_q + 2\Delta^{[\nu_{\text{total}}]}) \\ b_q + \frac{5\Delta^{[\nu_{\text{total}}]}}{2}, & \xi \in [b_q + 2\Delta^{[\nu_{\text{total}}]}, b_q + 3\Delta^{[\nu_{\text{total}}]}) \\ b_q + \frac{7\Delta^{[\nu_{\text{total}}]}}{2}, & \xi \in [b_q + 3\Delta^{[\nu_{\text{total}}]}, +\infty) \end{cases}$$

where  $\xi \in \mathbb{R}$  is the value to be quantized,  $b_q \in \mathbb{Q}$  is the basis of the quantizer,  $Q^{3\text{MRU}}(b_q, \xi, \Delta^{[\nu_{\text{total}}]})$  is the quantized

version of  $\xi$ , and  $\Delta^{[\nu_{\text{total}}]} \in \mathbb{Q}$  is the quantization level. Our 3-bit mid-rise uniform quantizer is also shown in Fig. 1. Note here that our findings can also be generalized to other types of quantizers.

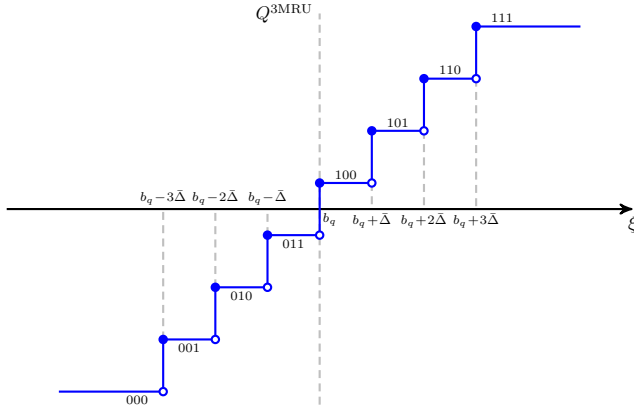


Fig. 1. A 3-bit mid-rise uniform quantizer,  $Q^{3MRU}(b_q, \xi, \Delta)$ .

Our quantizer (which is specified for exposition purposes only) utilizes 3 bits to represent quantized values within the output range  $\{b_q - \frac{7\Delta^{[\nu_{\text{total}}]}}{2}, b_q - \frac{5\Delta^{[\nu_{\text{total}}]}}{2}, \dots, b_q + \frac{7\Delta^{[\nu_{\text{total}}]}}{2}\}$ , and has  $2^3$  output levels (i.e., the output range contains  $2^3$  elements). Furthermore, note that our quantizer saturates for input values in the range  $\xi \in (-\infty, b_q - 3\Delta^{[\nu_{\text{total}}]}) \cup [b_q + 3\Delta^{[\nu_{\text{total}}]}, +\infty)$ . This means that the input signal  $\xi$  exceeds our quantizer's dynamic range, leading to clipping of the quantized output value to the nearest extreme value within the quantization range. This characteristic highlights the importance of selecting an appropriate quantization range to ensure an accurate representation of input signals without distortion due to saturation effects.

### III. PROBLEM FORMULATION

Let us consider a distributed network modeled as a digraph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with  $n = |\mathcal{V}|$  nodes. We assume that each node  $v_i$  is endowed with a local cost function  $f_i(x) : \mathbb{R}^p \mapsto \mathbb{R}$  only known to itself, and communication channels among nodes have limited capacity of 3 bits per channel, thus, only quantized values can be transmitted/communicated. In this paper we aim to develop a distributed algorithm which allows nodes, to cooperatively solve the following optimization problem called **P1** below:

$$\min_{x \in \mathcal{X}} F(x_1, x_2, \dots, x_n) \equiv \sum_{i=1}^n f_i(x_i), \quad (1a)$$

$$\text{s.t. } x_i = x_j, \forall v_i, v_j \in \mathcal{V}, \quad (1b)$$

$$x_i^{[0]} \in \mathcal{X} \subset \mathbb{Q}_{\geq 0}, \forall v_i \in \mathcal{V}, \quad (1c)$$

$$\text{nodes communicate via 3 bit channels,} \quad (1d)$$

where  $\mathcal{X}$  is the set of feasible values of variable  $x$ , and  $x^*$  is the optimal solution of the optimization problem. Eq. (1a) means that our objective is to minimize the global cost function, defined as the sum of the local cost functions in the network. Eq. (1b) imposes the requirement that nodes

need to calculate equal optimal solutions. Eq. (1c) means that the initial estimations of nodes must belong in the same set. However, it is not necessary for the initial values of nodes to be rational numbers (i.e.,  $x_i^{[0]} \in \mathcal{X} \subset \mathbb{Q}_{\geq 0}$ ), as they can generate a quantized version of their initial states by utilizing the Mid-Rise Uniform Quantizer presented in Section II. Eq. (1d) means that nodes are transmitting and receiving 3 bit quantized values with their neighbors. This constraint reflects the limited bandwidth of the network's communication channels.

**Remark 1** (Challenges of Optimization Problem). *The main challenges of our optimization problem **P1** are due to constraint (1d) (i.e., nodes are limited to exchanging messages represented by only 3 bits). Our proposed algorithm is able to address these challenges by enabling nodes to decide whether to zoom-in or zoom-out to the optimal solution, ensuring both communication efficiency and avoidance of saturation in the quantizer output range. Moreover, our algorithm enables nodes to calculate the exact optimal solution while exchanging 3-bit messages, mitigating quantization errors and saturation issues.*

### IV. DISTRIBUTED OPTIMIZATION WITH FINITE BIT ADAPTIVE QUANTIZATION

In this section, we present our distributed optimization algorithm which solves problem **P1** described in Section III. Before presenting our algorithm, we make the following assumptions that are crucial for the development of our results.

**Assumption 1.** *The communication network described as a digraph  $\mathcal{G}$  is strongly connected.*

**Assumption 2.** *The local cost function  $f_i(x)$  of every node  $v_i$  is smooth and strongly convex. This means that for every node  $v_i$ , for every  $x_1, x_2 \in \mathcal{X}$ ,*

- *there exists positive constant  $L_i$  such that*

$$\|\nabla f_i(x_1) - \nabla f_i(x_2)\|_2 \leq L_i \|x_1 - x_2\|_2, \quad (2)$$

- *there exists positive constant  $\mu_i$  such that*

$$f_i(x_2) \geq f_i(x_1) + \nabla f_i(x_1)^\top (x_2 - x_1) + \frac{\mu_i}{2} \|x_2 - x_1\|_2^2. \quad (3)$$

*This means that for the global cost function  $F$  (see (1a)), the Lipschitz-continuity constant of the gradient is  $L = \sum_i L_i$ , and the strong-convexity constant is  $\mu = \min_i \mu_i$ .*

**Assumption 3.** *The diameter  $D$  (or an upper bound of the diameter) of the digraph  $\mathcal{G}$  is known to every node  $v_i$  in the network.*

Assumption 1 ensures that information from each node propagates to every other node in the network. Thus all nodes are able to calculate the optimal solution  $x^*$  of **P1**. Assumption 2 is the Lipschitz-continuity condition (2), and strong-convexity condition (3). Lipschitz-continuity is a standard assumption in distributed first-order optimization problems (see [24], [25]). It guarantees that problem **P1**

has an optimal solution  $x^*$  for nodes to calculate, and that this solution is globally optimal. Strong-convexity guarantees linear convergence rate, and uniqueness of the optimal solution (i.e., the global function  $F$  has no more than one global minimum). Assumption 3 enables each node  $v_i \in \mathcal{V}$  to determine whether calculation of a solution  $x_i$  that fulfills (1b) has been achieved in a distributed manner. Note here that nodes can distributively compute the digraph diameter  $D$  in finite time by employing a graph diameter calculation protocol such as [26].

#### A. Distributed Optimization Algorithm with Finite Bit Adaptive Quantization

We now present our proposed algorithm detailed below as Algorithm 1.

---

#### Algorithm 1 Distributed Optimization Algorithm with Finite Bit Adaptive Quantization (DOFiBAQ)

---

**Input:** A strongly connected directed graph  $\mathcal{G}$  with  $n = |\mathcal{V}|$  nodes and  $m = |\mathcal{E}|$  edges. Every node  $v_i \in \mathcal{V}$  has: static step-size  $\alpha \in \mathbb{R}$ , digraph diameter  $D$ , initial value  $x_i^{[0]}$ , local cost function  $f_i$ , quantization level  $\Delta^{[0]} \in \mathbb{Q}$ , zoom-in constant  $c^{\text{in}} \in \mathbb{Q}_{>1}$ , zoom-out constant  $c^{\text{out}} \in \mathbb{Q}_{>1}$ . Assumptions 1, 2, 3 hold.

**Initialization:** Each node  $v_i \in \mathcal{V}$  sets  $b_q = 0$ ,  $\nu_{\text{in}} = 0$ ,  $\nu_{\text{out}} = 0$ ,  $\nu_{\text{total}} = 0$ .

**Iteration:** For  $k = 0, 1, 2, \dots$ , each node  $v_i \in \mathcal{V}$  does the following:

- 1)  $x_i^{[k+\frac{1}{2}]} = x_i^{[k]} - \alpha \nabla f_i(x_i^{[k]})$ ;
- 2)  $x_i^{[k+1]} = \text{Algorithm 2}(x_i^{[k+\frac{1}{2}]}, D, \Delta^{[\nu_{\text{total}}]})$ ;
- 3) **if**  $x_i^{[k+1]} = x_i^{[k]}$ , **then set**  $\nu_{\text{total}} = \nu_{\text{total}} + 1$  **and check**
  - **if**  $x_i^{[k+1]} \geq b_q + 3\Delta^{[\nu_{\text{total}}]}$ , **or**  $x_i^{[k+1]} < b_q - 3\Delta^{[\nu_{\text{total}}]}$  **then set**

$$\nu_{\text{out}} = \nu_{\text{out}} + 1, \quad (4a)$$

$$b_q = x_i^{[k+1]}, \quad (4b)$$

$$\Delta^{[\nu_{\text{total}}]} := c^{\text{out}} \Delta^{[\nu_{\text{total}}-1]}, \quad (4c)$$

- **else set**

$$\nu_{\text{in}} = \nu_{\text{in}} + 1, \quad (5a)$$

$$b_q = x_i^{[k+1]}, \quad (5b)$$

$$\Delta^{[\nu_{\text{total}}]} := \Delta^{[\nu_{\text{total}}-1]} / c^{\text{in}}, \quad (5c)$$

- 4) go to Step 1;

**Output:** Each node  $v_i \in \mathcal{V}$  calculates the optimal solution  $x_i^*$  of problem P1 in Section III.

---

The intuition of Algorithm 1 (DOFiBAQ) can be summarized in the following three parts.

*Part 1: Input and Initialization.* Each node  $v_i$  retains an estimate  $x_i^{[0]}$  of the optimal solution. Also, it has knowledge of (i) the network diameter  $D$ , (ii) a quantization level  $\Delta^{[0]}$ , (iii) a zoom-in constant  $c^{\text{in}} \in \mathbb{Q}_{>1}$ , (iv) a zoom-out constant  $c^{\text{out}} \in \mathbb{Q}_{>1}$ , and (v) a static step-size  $\alpha \in \mathbb{R}$ . The basis of the quantizer  $b_q$  (see Section II) is initialized to be equal

---

#### Algorithm 2 Finite-Time Quantized Average Consensus (FiTQuAC)

---

**Input:**  $x_i^{[k+\frac{1}{2}]}, D, \Delta^{[\nu_{\text{total}}]}$ .

**Initialization:** Each node  $v_i \in \mathcal{V}$  does the following:

- 1) Assigns probability  $b_{li}$  to each out-neighbor  $v_l \in \mathcal{N}_i^+ \cup \{v_i\}$ , as follows

$$b_{li} = \begin{cases} \frac{1}{1+\mathcal{D}_i^+}, & \text{if } l = i \text{ or } v_l \in \mathcal{N}_i^+, \\ 0, & \text{if } l \neq i \text{ and } v_l \notin \mathcal{N}_i^+; \end{cases}$$

- 2) sets  $z_i = 2$ ,  $y_i = 2 Q^{3\text{MRU}}(b_q, x_i^{[k+\frac{1}{2}]}, \Delta^{[\nu_{\text{total}}]}) / \Delta^{[\nu_{\text{total}}]}$ .

**Iteration:** For  $\lambda = 1, 2, \dots$ , each node  $v_i \in \mathcal{V}$ , does:

- 1) **if**  $\lambda \bmod (D) = 1$  **then**  $M_i = \lceil y_i / z_i \rceil$ ,  $m_i = \lfloor y_i / z_i \rfloor$ ;
- 2) broadcasts  $M_i, m_i$  to every  $v_l \in \mathcal{N}_i^+$ ; receives  $M_j, m_j$  from every  $v_j \in \mathcal{N}_i^-$ ; sets  $M_i = \max_{v_j \in \mathcal{N}_i^- \cup \{v_i\}} M_j$ ,  $m_i = \min_{v_j \in \mathcal{N}_i^- \cup \{v_i\}} m_j$ ;
- 3) sets  $c_i^z = z_i$ ;
- 4) **while**  $c_i^z > 1$  **do**
  - 4.1)  $c_i^y = \lfloor y_i / z_i \rfloor$ ;
  - 4.2) sets  $y_i = y_i - c_i^y$ ,  $z_i = z_i - 1$ , and  $c_i^z = c_i^z - 1$ ;
  - 4.3) transmits  $c_i^y$  to randomly chosen out-neighbor  $v_l \in \mathcal{N}_i^+ \cup \{v_i\}$  according to  $b_{li}$ ;
  - 4.4) receives  $c_j^y$  from  $v_j \in \mathcal{N}_i^-$  and sets

$$y_i = y_i + \sum_{j=1}^n w_{\lambda,ij}^{[r]} c_j^y, \quad (6)$$

$$z_i = z_i + \sum_{j=1}^n w_{\lambda,ij}^{[r]}, \quad (7)$$

where  $w_{\lambda,ij}^{[r]} = 1$  when node  $v_i$  receives  $c_j^y$ , 1 from  $v_j$  at time step  $\lambda$  (otherwise  $w_{\lambda,ij}^{[r]} = 0$  and  $v_i$  receives no message at time step  $\lambda$  from  $v_j$ );

- 5) **if**  $\lambda \bmod D = 0$  **and**  $M_i - m_i \leq 1$  **then sets**  $x_i^{[k+1]} = m_i \Delta^{[\nu_{\text{total}}]}$  **and stops operation.**

**Output:**  $x_i^{[k+1]}$ .

---

to zero (our algorithm remains functional also for different arbitrary initial values of  $b_q$ ). Also, the total zoom, zoom-in and zoom-out counters  $\nu_{\text{total}}, \nu_{\text{in}}, \nu_{\text{out}}$  are initialized equal to zero. The aforementioned parameters are common for every node. Note however that nodes can decide the parameter values in finite time by executing either a voting protocol [27], or a digraph diameter calculation protocol [26] (as mentioned in Assumption 3).

*Part 2: Iteration – Gradient Descent and Coordination.*

During each time step  $k$  of Algorithm 1, each node  $v_i$  updates the estimation  $x_i^{[k+\frac{1}{2}]}$  of its optimal solution by performing a gradient descent step towards the negative direction of the gradient of the node's local cost function. Then, each node employs Algorithm 2 (FiTQuAC). Algorithm 2 is a finite-time quantized averaging algorithm (details of its operation are presented below). This algorithm enables nodes to satisfy (1d) and compute in finite time an estimation of the optimal

solution  $x_i^{[k+1]}$  that satisfies (1b).

*Part 3: Iteration – Zoom-out or Zoom-in.* After executing Algorithm 2, nodes check if the calculated estimation of the optimal solution is the same as the previous step (i.e., if  $x_i^{[k+1]} = x_i^{[k]}$ ). The condition  $x_i^{[k+1]} = x_i^{[k]}$  means that nodes have converged to a neighborhood of the optimal solution defined by the current quantization level  $\Delta^{[\nu_{\text{total}}]}$  (see [22], [23]). Additionally, when this condition is met nodes are not able to enhance the precision of the estimated optimal solution using the existing quantization level. For this reason, they need to adjust the parameters of the utilized quantizer. Therefore, if the aforementioned condition holds, then nodes increase by one the total zoom counter  $\nu_{\text{total}}$  check:

- *Zoom-out:* If the estimation exceeds the quantizer’s dynamic range (i.e., if the quantizer has saturated because  $x_i^{[k+1]} \in (-\infty, b_q - 3\Delta^{[\nu_{\text{total}}]}) \cup [b_q + 3\Delta^{[\nu_{\text{total}}]}, +\infty)$ ), then the optimal solution of problem **P1** falls outside the current quantizer range. This means that nodes are not able to estimate the optimal solution with the current quantizer parameters. In this case nodes zoom-out. This is done by (i) setting the quantizer basis to be equal to the calculated estimation  $x_i^{[k+1]}$ , and (ii) increasing the quantization level by multiplying it with the zoom-out constant. Also, nodes increase by one the zoom-out counter  $\nu_{\text{out}}$  in order to keep track how many times they performed zoom-out. The primary purpose of zoom-out is to adapt the quantizer’s parameters to encompass a broader range of values that may potentially include the optimal solution.
- *Zoom-in:* If the estimate falls within the quantizer’s dynamic range (i.e., if the quantizer is not saturated and we have  $x_i^{[k+1]} \in [b_q - 3\Delta^{[\nu_{\text{total}}]}, b_q + 3\Delta^{[\nu_{\text{total}}]})$ ), then nodes are able to estimate the optimal solution of problem **P1** with the current quantizer parameters. In this case nodes zoom-in by (i) setting the quantizer basis to be equal to the calculated estimation  $x_i^{[k+1]}$ , and (ii) decreasing the quantization level by dividing it with the zoom-in constant. Also, nodes increase by one the zoom-in counter  $\nu_{\text{in}}$  to count the amount of zoom-in operations. By zoom-in nodes aim to adapt the quantizer’s parameters to “focus” at a smaller range of values that may potentially include the optimal solution. In this smaller range of values nodes are able to estimate with higher precision (since the updated  $\Delta^{[\nu_{\text{total}}]}$  is smaller than the original). Later in the paper, we will demonstrate how nodes are able to estimate the exact optimal solution by exchanging only 3-bit quantized messages with their neighbor nodes (thus fulfilling constraint (1d)).

Algorithm 2 (FiTQuAC) enables each node to calculate the quantized average of each node’s estimate in finite time. Additionally it incorporates a distributed stopping mechanism to notify nodes when calculation has converged so that nodes can return to Iteration Step 3 of Algorithm 1. During Algorithm 2, each node processes and transmits quantized messages, and the precision of the average calculation is

determined by the quantizer type (i.e., its quantization level, and its saturation limits). An analytical intuition of Algorithm 2 (FiTQuAC) is shown in [23, Algorithm 2]. We omit it due to space constraints.

### B. Convergence of Algorithm 1

We now analyze the convergence of Algorithm 1. We first introduce the following proposition to show that by zoom out nodes are able to adjust the quantizer dynamic range in finite time so it includes the optimal solution. We then present a theorem which establishes the linear convergence rate of our algorithm. Due to space constraints, we omit the proofs. They will be available in an extended version of our paper.

**Proposition 1** (Finite Zoom-out Instances). *Let us assume that during the operation of Algorithm 1,  $x^* \notin (b_q - 3\Delta^{[0]}, b_q + 3\Delta^{[0]})$ . Then, after a finite number of time steps  $\nu_0$  for which*

$$\nu_0 > \left\lceil \frac{x^* - \log(3\Delta)}{\log(c^{\text{out}})} \right\rceil, \quad (8)$$

*nodes are able to calculate  $\Delta^{[\nu_0]}$  such that  $x^* \in (b_q - 3\Delta^{[\nu_0]}, b_q + 3\Delta^{[\nu_0]})$ .*

We now show the linear convergence rate of Algorithm 1 in the following theorem. The proof follows from [22, Theorem 1] and will be available in an extended version of our paper.

**Theorem 1** (Linear Convergence Rate [22]). *Under Assumptions 1–3, when the step-size  $\alpha$  satisfies  $\alpha \in (0, \frac{2n}{\mu+L}]$  where  $L = \sum_i L_i$ , and  $\mu = \min_i \mu_i$  Algorithm 1 generates a sequence of points  $\{x^{[k]}\}$  (i.e., the variable  $x_i^{[k]}$  of each node  $v_i \in \mathcal{V}$ ) which satisfy*

$$\|\hat{x}^{[k+1]} - x^*\| \leq (1 - \frac{\alpha\mu}{n})\|\hat{x}^{[k]} - x^*\| + \mathcal{O}(\Delta^{[\nu_{\text{total}}]}), \quad (9)$$

where  $\Delta^{[\nu_{\text{total}}]}$  is the quantizer level and

$$\mathcal{O}(\Delta^{[\nu_{\text{total}}]}) = (\frac{4\alpha L}{n} + 2)\Delta^{[\nu_{\text{total}}]}. \quad (10)$$

## V. SIMULATION RESULTS

In this section, we present simulation results in order to illustrate Algorithm 1 and highlight its operational characteristics. We illustrate Algorithm 1 over a random digraph of 20 nodes (see Fig. 2). We show how nodes are able to adjust the quantizer’s parameters by performing zoom-out, or zoom-in. Subsequently, we show how nodes are able to converge linearly to the exact optimal solution of problem **P1** by exchanging messages consisted of only 3-bits.

During Algorithm 1 the parameters for each node  $v_i$  are:  $\alpha = 0.12$ , initial estimation of optimal solution  $x_i^{[0]} \in [1, 5]$ , quadratic local cost function of the form  $f_i(x) = \frac{1}{2}\beta_i(x - x_0)^2$  with  $\beta_i, x_0$  randomly chosen in the set  $\{1, 2, 3, 4, 5\}$ , initial quantization level  $\Delta^{[0]} = 0.5$ , zoom-in constant  $c^{\text{in}} = 4/3$ , zoom-out constant  $c^{\text{out}} = 2$ , and initial quantizer basis

$b_q = 0$ . We plot the error  $e^{[k]}$  in a logarithmic scale against the number of iterations defined as

$$e^{[k]} = \sqrt{\sum_{j=1}^n \frac{(x_j^{[k]} - x^*)^2}{(x_j^{[0]} - x^*)^2}}, \quad (11)$$

where  $x^*$  is the optimal solution of optimization problem **P1**.

In Fig. 2 during the operation of Algorithm 1 we can see that at time step  $k = 2$  the condition  $x_i^{[2]} = x_i^{[1]}$  holds for all nodes. Thus, nodes perform zoom-out during time step  $k = 2$ . They adjust the quantizer basis  $b_q$ , and increase the quantization level  $\Delta^{[0]}$  according to (4b) and (4c), thereby widening the quantizer range to encompass a broader range of values. At time step  $k = 4$  the condition  $x_i^{[4]} = x_i^{[3]}$  holds. Nodes again perform zoom out (i.e. adjust the quantizer parameters  $b_q, \Delta^{[1]}$ ) aiming for the adjusted quantizer range to include the optimal solution. Since the quantizer error depends on the quantization level then increasing  $\Delta^{[\nu_{\text{total}}]}$  increases the error introduced in the estimation of the optimal solution. This increases the distance between the optimal solution and each node's estimation of the optimal solution. For this reason, the error  $e^{[k]}$  increases at time steps  $k = 3, 5$ . At time step  $k = 6$  we have that  $x_i^{[6]} = x_i^{[5]}$  and nodes decide to zoom-in. Performing zoom-in at time step  $k = 6$  means that the zoom-out operations at time steps  $k = 2, 4$  adjusted  $b_q, \Delta^{[0]}$  in such a way that the quantizer range now includes the optimal solution. During the zoom-in at time step  $k = 6$ , nodes adjust  $b_q$ , and decrease  $\Delta^{[2]}$  according to (5b) (5c). This enables them to approximate the optimal solution with higher precision at time step  $k = 7$  (as we can see  $e^{[7]} < e^{[6]}$ ). Then, at time step  $k = 9$  nodes zoom in since  $x_i^{[9]} = x_i^{[8]}$ . This enables them to approximate the optimal solution with even higher precision at time step  $k = 10$ . By continuing this iterative process (i.e. perform zoom-in and approximate the optimal solution with even higher precision every time) we have that nodes are able to converge to the exact optimal solution during the operation of Algorithm 1 (as shown in Theorem 1).

**Remark 2** (Possible  $c^{\text{in}}$  Values). *Regarding the operation of Algorithm 1 in Fig. 2 we make the following two observations for the zoom-in constant  $c^{\text{in}}$ .*

- In Algorithm 1, nodes set  $c^{\text{in}} = 4/3$ . This means that after the zoom-in operation the adjusted quantizer will cover the range of values  $[b_q - 16\Delta^{[\nu_{\text{total}}-1]}/3, b_q + 16\Delta^{[\nu_{\text{total}}-1]}/3]$ . Note that Algorithm 2 (i) introduces a quantization error upper bounded by  $\Delta^{[\nu_{\text{total}}]}/2$  during its initialization, and (ii) gives the average of nodes states with a quantization error  $\Delta^{[\nu_{\text{total}}]}$ . Thus, the updated interval  $[b_q - 16\Delta^{[\nu_{\text{total}}-1]}/3, b_q + 16\Delta^{[\nu_{\text{total}}-1]}/3]$  includes the optimal solution despite the accumulated quantization error from Algorithm 2. Note also that setting  $c^{\text{in}} > 2$  for our 3 bit quantizer may lead to an adjusted quantizer range that does not include the optimal solution after reducing the quantization level  $\Delta^{[\nu_{\text{total}}]}$ . This may lead to nodes performing zoom-out so that the updated range will include again the

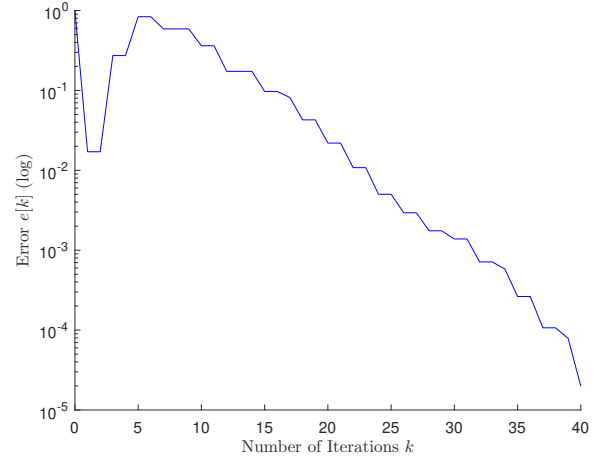


Fig. 2. Execution of Algorithm 1 over a random digraph of 20 nodes.

optimal solution. As a result, setting  $c^{\text{in}} > 2$  may lead to consecutive zoom-in and zoom-out operations that may prevent Algorithm 1 from converging to the exact optimal solution.

- In case we utilize a quantizer with higher resolution (i.e., with 4 or more bits) then setting the zoom-in constant equal to  $c^{\text{in}} = 4/3$  will also allow us to zoom towards an interval that includes the optimal solution. However, in this case we also could possibly increase the zooming constant to be greater than  $4/3$  in order to zoom in faster thus enhance the speed of convergence of the Algorithm 1.

Note that analyzing a possible range of values for  $c^{\text{in}}$  (according to the two aforementioned observations) is outside the scope of this paper but will be considered in an extended version.

## VI. CONCLUSIONS

In this paper, we focused on distributed optimization problem. In our problem each node's local function is strongly convex with lipschitz continuous gradients, and nodes exchange messages represented by 3-bits via bandwidth limited channels. We presented a distributed algorithm that enables nodes to estimate the exact optimal solution. We analyzed our algorithm's operation and established its linear convergence rate. We concluded with simulations in order to highlight our algorithm's operational advantages.

## REFERENCES

- [1] A. Nedic, "Distributed gradient methods for convex machine learning problems in networks: Distributed optimization," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 92–101, 2020.
- [2] G. S. Seyboth, D. V. Dimarogonas, and K. H. Johansson, "Event-based broadcasting for multi-agent average consensus," *Automatica*, vol. 49, no. 1, pp. 245–252, 2013.
- [3] S. U. Stich, J. B. Cordonnier, and M. Jaggi, "Sparsified SGD with memory," in *Advances in Neural Information Processing Systems*, vol. 31. Curran Associates, Inc., 2018, pp. 4447–4458.
- [4] A. Nedić and J. Liu, "Distributed optimization for control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 77–103, 2018.

- [5] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.
- [6] R. Xin and U. A. Khan, "A linear algorithm for optimization over directed graphs with geometric convergence," *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 315–320, 2018.
- [7] V. Khatana, G. Saraswat, S. Patel, and M. V. Salapaka, "Gradient-consensus method for distributed optimization in directed multi-agent networks," in *2020 American Control Conference (ACC)*, 2020, pp. 4689–4694.
- [8] T. Qin, S. R. Etesami, and C. A. Uribe, "Communication-efficient decentralized local SGD over undirected networks," in *2021 60th IEEE Conference on Decision and Control (CDC)*, 2021, pp. 3361–3366.
- [9] P. Yi and Y. Hong, "Quantized subgradient algorithm and data-rate analysis for distributed optimization," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 4, pp. 380–392, 2014.
- [10] C. Huang, H. Li, D. Xia, and L. Xiao, "Quantized subgradient algorithm with limited bandwidth communications for solving distributed optimization over general directed multi-agent networks," *Neurocomputing*, vol. 185, pp. 153–162, 2016.
- [11] Y. Pu, M. N. Zeilinger, and C. N. Jones, "Quantization design for distributed optimization," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2107–2120, 2017.
- [12] H. Li, S. Liu, Y. C. Soh, and L. Xie, "Event-triggered communication and data rate constraint for distributed optimization of multiagent systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 11, pp. 1908–1919, 2018.
- [13] T. T. Doan, S. T. Maguluri, and J. Romberg, "Fast convergence rates of distributed subgradient methods with adaptive quantization," *IEEE Transactions on Automatic Control*, vol. 66, no. 5, pp. 2191–2205, 2021.
- [14] S. Magnusson, H. Shokri-Ghadikolaei, and N. Li, "On maintaining linear convergence of distributed learning and optimization under limited communication," *IEEE Transactions on Signal Processing*, vol. 68, pp. 6101–6116, 2020.
- [15] D. Jhunjunwala, A. Gadhikar, G. Joshi, and Y. C. Eldar, "Adaptive quantization of model updates for communication-efficient federated learning," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 3110–3114.
- [16] Y. Kajiyama, N. Hayashi, and S. Takai, "Linear convergence of consensus-based quantized optimization for smooth and strongly convex cost functions," *IEEE Transactions on Automatic Control*, vol. 66, no. 3, pp. 1254–1261, 2021.
- [17] J. Liu, Z. Yu, and D. W. C. Ho, "Distributed constrained optimization with delayed subgradient information over time-varying network under adaptive quantization," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–14, 2022.
- [18] A. Koloskova, S. Stich, and M. Jaggi, "Decentralized stochastic optimization and gossip algorithms with compressed communication," in *Proceedings of the 36th International Conference on Machine Learning*, 2019, pp. 3478–3487.
- [19] D. Basu, D. Data, C. Karakus, and S. Diggavi, "Qsparse-local-SGD: Distributed SGD with quantization, sparsification and local computations," in *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [20] B. Li, S. Cen, Y. Chen, and Y. Chi, "Communication-efficient distributed optimization in networks with gradient tracking and variance reduction," in *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, 2020, pp. 1662–1672.
- [21] S. Horváth, D. Kovalev, K. Mishchenko, P. Richtárik, and S. Stich, "Stochastic distributed learning with gradient quantization and double-variance reduction," *Optimization Methods and Software*, vol. 38, no. 1, pp. 91–106, 2023.
- [22] A. I. Rikos, W. Jiang, T. Charalambous, and K. H. Johansson, "Distributed optimization with gradient descent and quantized communication," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 5900–5906, 2023.
- [23] —, "Distributed optimization via gradient descent with event-triggered zooming over quantized communication," in *IEEE Conference on Decision and Control*, 2023, pp. 6321–6327.
- [24] J. Xu, S. Zhu, Y. C. Soh, and L. Xie, "Convergence of asynchronous distributed gradient methods over stochastic networks," *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 434–448, 2018.
- [25] G. Qu and N. Li, "Harnessing smoothness to accelerate distributed optimization," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1245–1260, 2018.
- [26] G. Oliva, R. Setola, and C. N. Hadjicostis, "Distributed finite-time calculation of node eccentricities, graph radius and graph diameter," *Systems & Control Letters*, vol. 92, pp. 20–27, 2016.
- [27] J. Cortés, "Distributed algorithms for reaching consensus on general functions," *Automatica*, vol. 44, no. 3, pp. 726–737, March 2008.