Preprints of the 23rd IFAC World Congress
Busan, Republic of Korea, August 23-28, 2026

Affine-coupled Distributed Optimization via Distributed Proximal Jacobian ADMM with Quantized Communication

Xu Du * **Boyu Han** * **Ivano Notarnicola** **
Karl H. Johansson *** **Apostolos I. Rikos**† *,****

* *The Artificial Intelligence Thrust of the Information Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China, (e-mail: {michaelxudu,boyuhan,apostolosr}@hkust-gz.edu.cn).*
** *Department of Electrical, Electronic, and Information Engineering, University of Bologna, Italy, (e-mail: ivano.notarnicola@unibo.it)*
*** *The Division of Decision and Control Systems, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden, (e-mail: kallej@kth.se)*
**** *The Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China.*

Abstract: This paper investigates distributed resource allocation optimization over directed graphs with limited communication bandwidth. We develop a novel distributed algorithm that integrates the centralized Proximal Jacobian Alternating Direction Method of Multipliers (PJ-ADMM) with a finite-level quantized consensus scheme, enabling nodes to cooperatively solve the optimization in a distributed fashion. Under the assumption of convex objective functions, we establish that the proposed algorithm achieves sublinear convergence to a neighborhood of the optimal solution, with the convergence accuracy explicitly bounded by the quantization level. Numerical experiments validate that the algorithm achieves competitive performance compared to existing approaches while exhibiting communication efficiency.

Keywords: Distributed Optimization, PJ-ADMM, Quantized Communication, Finite-time Consensus

1. INTRODUCTION

Distributed optimization has gained significant attention with the rise of federated learning Ren et al. (2025), robotics and model predictive control (MPC) Stomberg et al. (2025), power grids Du et al. (2019) and wireless communication Yang et al. (2024). These applications highlight that, as data volumes grow and optimization variables scale, solving large-scale problems on centralized devices becomes impractical (Doostmohammadian et al. (2025)).

Distributed optimization enables large-scale problems with high-dimensional variables to be partitioned across multiple devices, where each device solves a local subproblem and exchanges information until the global solution is obtained. Broadly, distributed optimization methods fall into two categories Ling et al. (2015): (i) primal decomposition and (ii) dual decomposition. In general, dual decomposition methods exhibit faster convergence and higher accuracy compared to primal approaches Ling et al. (2015). In this work, we focus on dual decomposition, with particular emphasis on the Alternating Direction Method of Multipliers (ADMM) (Boyd et al. (2011)) for resource allocation problems.

Existing Literature. ADMM was originally introduced in Glowinski (1975); Gabay and Mercier (1976). For comprehensive surveys and recent advances in the field, we refer readers to Boyd et al. (2011); Lin et al. (2022). However, in the context of resource allocation problems, He et al. (2015) demonstrated that the Jacobian variant of ADMM does not always guarantee convergence. To overcome this limitation, subsequent research has evolved along four primary directions:

- **Dual consensus ADMM**, see, e.g., Chang (2016); Jiang et al. (2022);
- **Variable Splitting ADMM**, such as Wang et al. (2013); Notarnicola and Falsone (2022) and (Houska et al., 2016, Algorithm 1);
- **Tracking-based ADMM over networks**, e.g., Falsone et al. (2020);
- **Proximal Jacobian ADMM with step-size modification**, including proximal regularization variants He et al. (2016, 2015); Deng et al. (2017); Choi and Choi (2025); Li and Yuan (2018), and accelerated dis-

⋆ † Corresponding author
**The work of X.D. and A.I.R. was supported by the Guangzhou-HKUST(GZ) Joint Funding Scheme (Grant No. 2025A03J3960). The work of A.I.R. was also supported by the Guangdong Provincial Project (Grant No. 2024QN11G109).

tributed augmented Lagrangian (ADAL)-type methods Chatzipanagiotis and Zavlanos (2017).

However, except for Falsone et al. (2020), which considers fully distributed information exchange over networks, the aforementioned approaches suffer from two main limitations: (i) their reliance on centralized coordination mechanisms, which constrains scalability in distributed systems, and (ii) the need for nodes to transmit real-valued messages, imposing significant bandwidth requirements and creating a scalability bottleneck. To the best of our knowledge, for Consensus ADMM-type algorithms, Rikos et al. (2023a) and Du et al. (2025) leverage the network-based quantized consensus algorithm proposed in Rikos et al. (2022) to design fully distributed quantized schemes. However, extending such analysis to resource allocation type ADMM algorithms remains largely unexplored. The main difficulty lies in the fact that resource allocation problems involve affine coupling constraints rather than simple consensus constraints, which breaks the symmetry required by most quantized consensus protocols and thus prevents the direct application of network-based quantized consensus algorithms.

Main Contributions. Motivated by the aforementioned challenges, we propose a novel distributed optimization algorithm based on the centralized Proximal Jacobian Alternating Direction Method of Multipliers (PJ-ADMM) framework Deng et al. (2017), which to the best of our knowledge represents the first integration of PJ-ADMM achieving (i) fully distributed operation, (ii) communication over directed graphs, and (iii) quantized information exchange among nodes. Drawing inspiration from Rikos et al. (2023a) and Du et al. (2025)), our approach employs a two-layer architecture that decouples optimization steps from distributed averaging, enabling faster convergence compared to single-layer methods. Specifically, we develop the Quantized distributed Proximal Jacobian ADMM (QDPJ-ADMM) detailed in Algorithm 1, where the inner layer implements the quantized consensus scheme from Rikos et al. (2023a) (Algorithm 2) to manage communication through quantized message exchanges, while the outer layer handles primal-dual variable updates. Theoretical analysis in Theorem 1 establishes sublinear convergence to a neighborhood of the optimal solution for convex objectives, with numerical experiments confirming convergence accuracy dependent on the utilized quantization level due to exclusive use of quantized data for inter-subproblem communication.

2. NOTATION AND PRELIMINARIES

Notation. The symbols $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$, and $\mathbb{N}$ denote the sets of real, rational, integer, and natural numbers, respectively. Matrices and vectors are denoted by uppercase (e.g., A) and lowercase (e.g., a) letters, respectively. The transpose of a matrix $A \in \mathbb{R}^{n\times n}$ or a vector $a \in \mathbb{R}^n$ is written as $A^\top$ or $a^\top$. For a scalar $a \in \mathbb{R}$, $\lfloor a \rfloor$ and $\lceil a \rceil$ represent the floor and ceiling functions, respectively. These operations apply element-wise to vectors $a \in \mathbb{R}^n$, yielding $\lfloor a \rfloor, \lceil a \rceil \in \mathbb{R}^n$. The vector of all ones is denoted by $\mathbf{1}$, and I is the identity matrix of appropriate dimension. The Euclidean norm of a vector a is written as $|a|$. The value of a variable x at node i and iteration k is denoted by $x_i^{[k]}$. Furthermore, $|\mathcal{S}|$ denotes the cardinality of a set $\mathcal{S}$ (e.g., $|\mathcal{V}| = N$, as introduced subsequently). The notation $a \mid b$ indicates that $b \in \mathbb{R}^n$ is the dual variable associated with the constraint $a \in \mathbb{R}^n$.

Graph Theory. The communication network is modeled as a directed graph (digraph) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = 1, \dots, N$ ($N \geq 2$) denotes the set of agents (nodes). The edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ includes a virtual self-edge (i,i) for every node $i \in \mathcal{V}$. A directed edge from node i to j is denoted $e_{ji} = (j,i) \in \mathcal{E}$. The set of in-neighbors of node i, denoted $\mathcal{N}_i^{[k]} = \{j \in \mathcal{V} \mid (i,j) \in \mathcal{E}\}$, consists of nodes that can transmit information directly to i, with its cardinality $\mathcal{D}_i^{[k]} = |\mathcal{N}_i^{[k]}|$ representing the in-degree. Similarly, the out-neighbor set $\mathcal{N}_i^{[k+1]} = \{l \in \mathcal{V} \mid (l,i) \in \mathcal{E}\}$ contains nodes that directly receive information from i, and its cardinality $\mathcal{D}_i^{[k+1]} = |\mathcal{N}_i^{[k+1]}|$ defines the out-degree. The diameter D of $\mathcal{G}$ is the length of the longest shortest directed path between any node pair $i, j \in \mathcal{V}$, and the digraph is strongly connected if a directed path exists between every pair of nodes $i, j \in \mathcal{V}$.

Quantization. Quantization plays a key role in digital communication networks by lowering bandwidth requirements and increasing communication efficiency. Its representation of signals with a finite number of bits enables the implementation of error-correcting codes such as Reed–Solomon or LDPC, thereby allowing significant gains in robustness against interference Proakis and Salehi (2002). While asymmetric, uniform, and logarithmic quantizers are all well-documented Wei et al. (2019), the analysis in this paper employs an asymmetric mid-rise quantizer with infinite range. The quantizer is defined as

$$q_\Delta^a(b) := \left\lfloor \frac{b}{\Delta} \right\rfloor, \tag{1}$$

for input vector $b \in \mathbb{R}^n$ and quantization level $\Delta \in \mathbb{Q}$, with the superscript a indicating its asymmetric nature. It is noted that our results hold for other quantizer types as well.

3. PROBLEM FORMULATION AND CENTRALIZED PJ-ADMM

3.1 Resource Allocation Optimization

Consider a communication network represented by a directed graph (digraph) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $N = |\mathcal{V}|$ nodes. We consider the following resource allocation problem defined over this network

$$\begin{aligned} \min_{x_i \in \mathbb{R}^{n_i},\ i\in\mathcal{V}} \quad & \sum_{i=1}^N f_i(x_i) \\ \text{s.t.} \quad & \sum_{i=1}^N A_i x_i = b, \end{aligned} \tag{2}$$

where each node $i \in \mathcal{V}$ is associated with a local objective $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$. The local decision variables $x_i \in \mathbb{R}^{n_i}$ are affinely coupled through the global equality constraint $\sum_{i=1}^N A_i x_i = b \in \mathbb{R}^m$, where $A_i \in \mathbb{R}^{m\times n_i}$ and $b \in \mathbb{R}^m$.

Now let us assume that communication over the network is subject to limited bandwidth. In order to solve problem (2) while ensuring communication efficiency, we reformulate it as a quantized resource allocation optimization problem

$$\begin{aligned} \min_{x_i \in \mathbb{R}^{n_i},\ i \in \mathcal{V}} \quad & \sum_{i=1}^{N} f_i(x_i) \\ \text{s.t.} \quad & \sum_{i=1}^{N} A_i x_i = b, \\ & \text{nodes communicate with quantized values,} \end{aligned} \tag{3}$$

where inter-node communications are restricted to quantized messages due to limited bandwidth.

3.2 Centralized PJ-ADMM

The Lagrangian of problem (2) is defined as

$$\mathcal{L}(x, \lambda) := \sum_{i=1}^{N} f_i(x_i) + \lambda^\top \left(\sum_{i=1}^{N} A_i x_i - b \right), \tag{4}$$

where $x = [x_1^\top, x_2^\top, \cdots, x_N^\top]^\top$ collects all local decision variables, and $\lambda \in \mathbb{R}^m$ denotes the dual multiplier associated with the coupling constraint. The corresponding augmented Lagrangian is given by,

$$\mathcal{L}_\rho(x, \lambda) := \sum_{i=1}^{N} f_i(x_i) + \lambda^\top \left(\sum_{i=1}^{N} A_i x_i - b \right) + \frac{\rho}{2} \left\| \sum_{i=1}^{N} A_i x_i - b \right\|^2, \tag{5}$$

where $\rho > 0$ denotes the penalty parameter.

Focusing on the augmented Lagrangian in (5), the centralized PJ-ADMM for solving problem (2) was proposed in Deng et al. (2017). The centralized PJ-ADMM iteration consists of the following two steps:

$$\begin{cases} x_i^{[k+1]} = \arg\min_{x_i} f_i(x_i) + \frac{1}{2} \left\| x_i - x_i^{[k]} \right\|_{P_i}^2 \\ \qquad + \frac{\rho}{2} \left\| A_i x_i + \sum_{j \neq i}^{N} A_j x_j^{[k]} - b + \frac{\lambda^{[k]}}{\rho} \right\|^2, \quad \forall i \in \mathcal{V}, \\ \lambda^{[k+1]} = \lambda^{[k]} + \gamma\rho \left(\sum_{i=1}^{N} A_i x_i^{[k+1]} - b \right), \end{cases} \tag{6}$$

where P_i satisfies $P_i \succ \rho \left(\frac{N}{2-\gamma} - 1 \right) \|A_i\|^2$ and $0 < \gamma < 2$, see (Deng et al., 2017, Lemma 2.2). In the first step of (6), each node minimizes the augmented Lagrangian with an additional proximal regularization term $\frac{1}{2} \left\| x_i - x_i^{[k]} \right\|_{P_i}^2$ with respect to its local variable x_i. In the second step, the dual variable λ is updated. For convex local objectives $f_i, \forall i \in \mathcal{V}$, the PJ-ADMM algorithm admits global sublinear convergence guarantees, as shown in (Deng et al., 2017, Section 2).

It is important to emphasize that the centralized PJ-ADMM formulation presented in (6) still depends critically on a centralized coordination mechanism, as the dual variable update necessitates the aggregation of global information from all agents. Such centralized dependency may hinder scalability and robustness in large-scale networked systems. To address this limitation, the next section introduces a fully distributed variant that eliminates the need for a central coordinator. Compared with PJ-ADMM, the proposed algorithm possesses two distinct features: (i) each node interacts only with its local neighbors while collaboratively solving the global problem, and (ii) communication efficiency is enhanced through the incorporation of quantized information exchange mechanisms.

4. DISTRIBUTED PROXIMAL JACOBIAN ADMM WITH EFFICIENT COMMUNICATION

In this section, we propose a novel algorithm to solve problem (3) in a fully distributed manner. Before presenting the algorithmic details, we introduce several standard assumptions that will be essential for the subsequent analysis.

Assumption 1. The communication network is represented by a *strongly connected* directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Each node $i \in \mathcal{V}$ is assumed to know the network diameter D and a common quantization level Δ.

Assumption 2. For each node $i \in \mathcal{V}$, the local objective function $f_i : \mathbb{R}^n \to \mathbb{R}$ is closed, proper, and convex. In particular, for all $x_\alpha, x_\beta \in \mathbb{R}^n$, the following inequality holds:

$$f_i(x_\alpha) + \partial f_i(x_\alpha)^\top (x_\beta - x_\alpha) \leq f_i(x_\beta), \tag{7}$$

where $\partial f_i(x_\alpha)$ denotes a sub-gradient of f_i at the point x_α. Moreover, the solution set of problem (3) is assumed to be non-empty.

Assumption 1 provides a necessary condition for the convergence of Algorithm 2 Rikos et al. (2023b), which serves as the coordination step of Algorithm 1 as described in (10). Moreover, the digraph diameter D can be computed by the nodes in a distributed manner within finite time using a protocol for graph diameter calculation, e.g., Oliva et al. (2016). Assumption 2 ensures the convexity of each local objective, enabling the establishment of global convergence results for Algorithm 1. Furthermore, Assumption 2 guarantees the existence of an optimal solution for problem (3).

4.1 Algorithm Development

In this section, we present our proposed distributed algorithm, detailed below as Algorithm 1.

To develop the distributed algorithm, we first introduce an auxiliary variable $d^{[k]}$ defined as

$$d^{[k]} := \sum_{i=1}^{N} A_i x_i^{[k]} - b. \tag{8}$$

The quantity $d^{[k]}$ represents the global residual of the affine coupled constraints in (3), which remains constant in a centralized setting. In the fully distributed implementation, it will be replaced by a locally maintained estimate $\hat{d}_i^{[k]}$ (see (10)), which is dynamically updated through local communication among neighboring agents.

Algorithm 1 is structured into three main phases: i) local optimization, ii) coordination among nodes, iii) dual variable update. In the first phase, each node $i \in \mathcal{V}$ minimizes its own local augmented Lagrangian function, see (9). In the second phase, all nodes exchange the information $\phi_i^{[k+1]} = N(A_i x_i^{[k+1]} - b)$ via the distributed quantized averaging Algorithm 2, thereby obtaining the updated auxiliary variable estimation $\hat{d}_i^{[k+1]}$ as described in (10). In the third phase, the dual variable $\hat{\lambda}_i^{[k+1]}$ is updated and subsequently used in the next iteration's local optimization step, see (11). Note that $\hat{\lambda}_i^{[k+1]}$ can be interpreted as a local copy of the global dual variable $\lambda^{[k+1]}$ in (6) for each agent i. These three phases are repeated iteratively until convergence.

Algorithm 2, first proposed in Rikos et al. (2022), consists of three main components: (i) quantization, (ii) averaging, and (iii) a stopping criterion. Firstly, each node initializes its local information as $\phi_i^{[k+1]}$ and introduces χ_i to denote the quantized representation of $\phi_i^{[k+1]}$. Secondly, in the averaging step, χ_i is split into ξ_i equal integer pieces (some pieces may differ by one). Each node retains the piece with the smallest value and transmits the remaining $\xi_i - 1$ pieces to randomly selected out-neighbors $l \in \mathcal{N}_i^{[k+1]}$ or to itself. Simultaneously, it receives pieces c_j from all in-neighbors $j \in \mathcal{N}_i^{[k]}$ and updates χ_i and ξ_i as in (13). Thirdly, the algorithm performs max- and min-consensus operations every D time steps. If the difference between the max-consensus M_i and min-consensus m_i satisfies $\|M_i - m_i\|_\infty \le 1$, then each node i scales its solution according to the quantization level to compute $\hat{z}_i^{[k+1]}$. Once the stopping condition is satisfied, the solution is scaled accordingly. At this point, Algorithm 2 terminates, and each node i proceeds to the second phase of Algorithm 1. The convergence analysis of Algorithm 2 is provided in (Rikos et al., 2022, Theorem 1).

Algorithm 1 QDPJ-ADMM: Quantized Distributed Proximal Jacobian ADMM

Input. Strongly connected digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, parameter ρ, network diameter D, quantization level Δ, for each node $i \in \mathcal{V}$. Each node $i \in \mathcal{V}$ has a local cost function f_i. Assumptions 1 and 2 hold.
Initialization. Randomly chosen dual variable $\hat{\lambda}_i \in \mathbb{R}^n$, and global variable estimation $\hat{d}_i \in \mathbb{R}^m$, for each node $i \in \mathcal{V}$.
Iteration.

1. Paralleled solve local NLP:
$$\begin{aligned} x_i^{[k+1]} = &\arg\min_{x_i} f_i(x_i) + \frac{1}{2}\left\|x_i - x_i^{[k]}\right\|_{P_i}^2 \\ &+ \frac{\rho}{2}\left\|A_i x_i - A_i x_i^{[k]} + \hat{d}_i^{[k]} + \frac{\hat{\lambda}_i^{[k]}}{\rho}\right\|^2. \end{aligned} \tag{9}$$
2. Calculate the global variable estimation executing the distributed following algorithm until convergence:
$$\hat{d}_i^{[k+1]} = \text{Algorithm 2}\left(\phi_i^{[k+1]}, D, \Delta\right), \tag{10}$$
where $\phi_i^{[k+1]} = N(A_i x_i^{[k+1]} - b)$.
3. Update the dual:
$$\hat{\lambda}_i^{[k+1]} = \hat{\lambda}_i^{[k]} + \gamma\rho\hat{d}_i^{[k+1]}. \tag{11}$$

Output. Each node i calculates x_i^* that solves problem (3).

4.2 Convergence Analysis

In this section, we provide the convergence analysis of Algorithm 1. First, we introduce a lemma that is important for our analysis. Then, we prove our main result via a theorem.

Lemma 1. The update of the local estimate $\hat{d}_i$ for each node $i \in \mathcal{V}$ is given by (10) in Algorithm 1. We note that the operation of Algorithm 2, as a consequence of quantized communication, satisfies the inequality,

$$\begin{cases} \hat{d}_i^{[k+1]} = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} \Delta \left\lfloor \dfrac{\phi_j^{[k+1]}}{\Delta} \right\rfloor + \kappa_i, \\ \left\|\hat{d}_i^{[k+1]} - d^{[k+1]}\right\| \le 2\sqrt{m}\Delta, \end{cases} \tag{14}$$

where $\|\kappa_i\|_\infty \le \Delta$, $\phi_i^{[k+1]} = N(A_i x_i^{[k+1]} - b), \forall i \in \mathcal{V}$ and $d^{[k+1]}$ is defined in (8).

Proof. See (Rikos et al., 2023b, Lemma 1), (Jiang et al., 2021, Appendix) . ■

Note that the proof of the theorem below is available in Du et al. (2026). It will also be available in an extended version of our paper.

Theorem 1. Let us consider a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Each node $i \in \mathcal{V}$ has a local cost function f_i, and Assumptions 1 and 2 hold. Each node $i \in \mathcal{V}$ in the network executes Algorithm 1 for solving the resource allocation optimization problem in (3) in a distributed fashion. Given parameter $\rho > 0$, during the operation of Algorithm 1 there always

Algorithm 2 DFQAC: Distributed Finite-time Quantized Average Consensus

Input. $\phi_i^{[k+1]} = N(A_i x_i^{[k+1]} - b), D, \Delta$.
Initialization. Each node $i \in \mathcal{V}$:

1. Assigns probability
$$p_{li} = \begin{cases} \dfrac{1}{1 + \mathcal{D}_i^{[k+1]}}, & \text{if } l \in \mathcal{N}_i^{[k+1]} \cup \{i\}, \\ 0, & \text{if } l \notin \mathcal{N}_i^{[k+1]} \cup \{i\}, \end{cases} \tag{12}$$
to each out-neighbor of node i.
2. Sets $\xi_i = 2$, $\chi_i = 2q_\Delta^a(\phi_i^{[k+1]})$ (see (1)).

Iteration. For time steps $t = 1, 2, \cdots$ each node $i \in \mathcal{V}$ does:

1. **If** $t \bmod(D) = 1$, sets $M_i = \left\lceil \frac{\chi_i}{\xi_i} \right\rceil$ and $m_i = \left\lfloor \frac{\chi_i}{\xi_i} \right\rfloor$.
2. Broadcasts M_i, m_i to each out-neighbor $l \in \mathcal{N}_i^{[k+1]}$ and receives M_j, m_j from each in-neighbor $j \in \mathcal{N}_i^{[k]}$. Then, sets $M_i = \max_{j \in \mathcal{N}_i^{[k]} \cup \{i\}} M_j$, $m_i = \min_{j \in \mathcal{N}_i^{[k]} \cup \{i\}} m_j$.
3. Sets $\tau_i = \xi_i$.
4. **While** $\tau_i > 1$ **do**
   (a) $c_i = \left\lfloor \frac{\chi_i}{\xi_i} \right\rfloor$.
   (b) Sets $\chi_i = \chi_i - c_i$, $\xi_i = \xi_i - 1$, $\tau_i = \tau - 1$.
   (c) Transmits c_i to randomly chosen out-neighbor $l \in \mathcal{N}_i^{[k+1]} \cup \{i\}$ with probability p_{li}.
   (d) Receives c_i from $j \in \mathcal{N}_i^{[k]}$ and updates
$$\chi_i^{[t+1]} = \chi_i^{[t]} + \sum_{j \in \mathcal{N}_i^{[k]}} w_{ij}^{[t]} c_j^{[t]}, \tag{13a}$$
$$\xi_i^{[t+1]} = \xi_i^{[t]} + \sum_{j \in \mathcal{N}_i^{[k]}} w_{ij}^{[t]}. \tag{13b}$$
   Here $w_{ij}^{[t]} = 1$ if node i receives $c_j^{[t]}$ from node j at step t. Otherwise $w_{ij}^{[t]} = 0$ and node i does not receive information from node j.
5. **if** t mod $(D) = 0$ and $\|M_i - m_i\|_\infty \le 1$, set $\hat{d}_i^{[k+1]} = m_i\Delta$, and stop the operation of the algorithm.

Output. $\hat{d}_i^{[k+1]}$.

exists a $0 < \gamma < 2$ and $P_i, \forall i \in \mathcal{V}$ such that

$$P_i \succ \rho \left(\frac{N}{2-\gamma} - 1 \right) \|A_i\|^2. \tag{15}$$

From (15), we have that during the operation of Algorithm 1 the following inequality is satisfied

$$\begin{aligned} &\mathcal{L}\left(\frac{\sum_{k=1}^{K} x^{[k]}}{K}, \lambda^* \right) - \mathcal{L}(x^*, \lambda^*) \\ \leq &\frac{1}{K} \sum_{k=1}^{K} \left(\mathcal{L}\left(x^{[k]}, \lambda^*\right) - \mathcal{L}(x^*, \lambda^*) \right) \leq \frac{C}{K} + \mathcal{O}(\Delta), \end{aligned} \tag{16}$$

where

$$C := \frac{1}{2} \sum_{i=1}^{N} \left\| x_i^{[1]} - x_i^* \right\|^2_{\rho A_i^\top A_i + P_i} + \frac{1}{2\gamma\rho} \left\| \hat{\lambda}_i^{[1]} - \lambda^* \right\|^2, \tag{17}$$

Δ denotes the quantization level and (x^*, λ^*) is a saddle point of (3).

The convergence result in Theorem 1 is inherited from Deng et al. (2017). Based on this theorem, we derive a convergence result for the decision variables x_i, as stated in the following corollary.

Corollary 1. Under the same conditions as in Theorem 1, it holds that

$$\begin{aligned} &\lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \frac{1}{2} \sum_{i=1}^{N} \left\| x_i^{[k]} - x_i^{[k+1]} \right\|^2_{\rho A_i^\top A_i + P_i - \frac{\rho}{\epsilon_i} A_i^\top A_i} \\ \leq &\frac{C}{K} + \mathcal{O}(\Delta), \end{aligned} \tag{18}$$

where the constant C is defined in (17), and each $\epsilon_i > 0$ for $i \in \mathcal{V}$.

Remark 1. Under the convexity of the objective functions f_i, Algorithm 1 guarantees that the function values converge to a neighborhood of the optimal value, with the size of this neighborhood ultimately determined by the quantization level Δ. Since C is a constant, the term $\frac{C}{K}$ tends to zero as $K \to \infty$. This implies that the left-hand side of (18) is upper bounded by $\mathcal{O}(\Delta)$. A more rigorous analysis of variable convergence, including the convergence of $\left\| x_i^{[k]} - x_i^* \right\|^2$ and the feasibility of coupling constraints in (3), will be addressed in an extended version of this paper.

5. NUMERICAL SIMULATION

In this section, we present numerical simulations to illustrate the performance of Algorithm 1 and highlight its improvements over existing optimization methods.

Our experiments are based on the setup in (Deng et al., 2017, Section 3.7), with implementation details available at `https://github.com/ZhiminPeng/Jacobi-ADMM`. We consider the following resource allocation problem:

$$\min_{x_i \in \mathbb{R}^{n_i},\ i \in \mathcal{V}} \sum_{i=1}^{N} \frac{1}{2} \|C_i x_i - e_i\|^2 \qquad \text{s.t.} \sum_{i=1}^{N} x_i = 0, \tag{19}$$

defined over a random digraph with $N = 100$ nodes, where $x_i \in \mathbb{R}^{100}$. Here $C_i \in \mathbb{R}^{120\times 100}$, and $e_i \in \mathbb{R}^{120}$ for all $i \in \mathcal{V}$ are randomly generated. For further details, please refer to (Deng et al., 2017, Section 3.7). We set $\rho = 0.01$, $\gamma = 1$, and $P_i = \tau I \succ 0$ with $\tau = \rho(N - 1 + 10^{-3})$ to satisfy the condition (15). Algorithm 1 is executed for quantization levels $\Delta = 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}$. We evaluate convergence by plotting the error $\sum_{i=1}^{N} \|x_i^{[k]} - x_i^*\|_1$, $\forall i \in \mathcal{V}$, where $x^* = [(x_1^*)^\top, (x_2^*)^\top, \dots, (x_N^*)^\top]^\top$ denotes the optimal solution of (3).

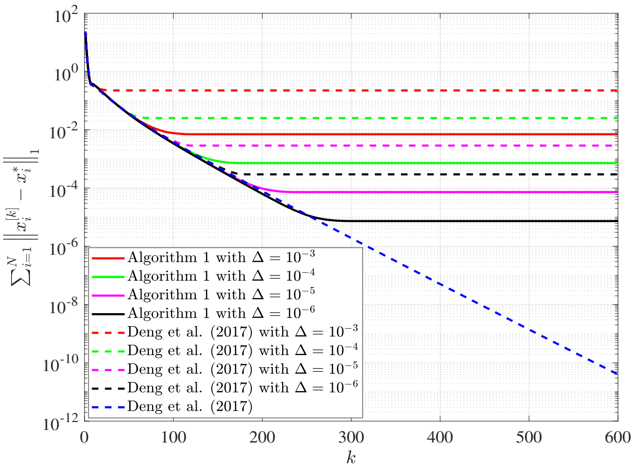

Fig. 1. Comparison of Algorithm 1 with PJ-ADMM (Deng et al. (2017)) over a directed graph with quantization level $\Delta = 10^{-3}, 10^{-4}, 10^{-5}$ and 10^{-6}.

From Fig. 1, it can be observed that Algorithm 1 converges to a neighborhood of the optimal solution of (3). In particular, a smaller quantization level Δ yields higher accuracy, which is consistent with the theoretical results in Theorem 1. This demonstrates that Algorithm 1 achieves a tradeoff between communication efficiency and solution accuracy, as a larger Δ reduces communication bandwidth requirements among nodes. Moreover, the algorithm operates in a fully distributed manner, without relying on a central server.

For comparison, the dashed lines in Fig. 1 show the convergence of the centralized PJ-ADMM (6) under different quantization levels Δ, where the last dashed line corresponds to real-valued transmission. Since each iteration of (6) requires exchanging both $x_i^{[k+1]}$ and $\lambda^{[k+1]}$, the centralized scheme accumulates larger quantization errors than Algorithm 1.

6. CONCLUSION

In this work, we proposed DQPJ-ADMM, a novel distributed algorithm that computes optimal solutions over directed networks under quantized communication. Operating without a central coordinator, it achieves sublinear convergence to a neighborhood of the optimum for convex resource allocation problems, with accuracy dependent on the quantization level. Theoretical analysis ensures convergence, while numerical tests validate the results and demonstrate improved communication efficiency over existing baselines. Future work will include a rigorous convergence analysis quantifying how the presence of quantization noise affects the accuracy of the final optimal solution under stronger assumptions, along with extending the proposed method to open network environments with dynamic topologies and agent participation.

REFERENCES

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1), 1–122.

Chang, T.H. (2016). A proximal dual consensus ADMM method for multi-agent constrained optimization. *IEEE Transactions on Signal Processing*, 64(14), 3719–3734.

Chatzipanagiotis, N. and Zavlanos, M.M. (2017). On the convergence of a distributed augmented Lagrangian method for nonconvex optimization. *IEEE Transactions on Automatic Control*, 62(9), 4405–4420.

Choi, H. and Choi, W. (2025). A linear convergence result for the Jacobi-proximal Alternating Direction Method of Multipliers. *arXiv preprint arXiv:2503.18601*.

Deng, W., Lai, M.J., Peng, Z., and Yin, W. (2017). Parallel multi-block ADMM with o (1/k) convergence. *Journal of Scientific Computing*, 71, 712–736.

Doostmohammadian, M., Aghasi, A., Pirani, M., Nekouei, E., Zarrabi, H., Keypour, R., Rikos, A.I., and Johansson, K.H. (2025). Survey of distributed algorithms for resource allocation over multi-agent systems. *Annual Reviews in Control*, 59, 100983.

Du, X., Engelmann, A., Jiang, Y., Faulwasser, T., and Houska, B. (2019). Distributed state estimation for AC power systems using Gauss-Newton ALADIN. In *IEEE Conference on Decision and Control*, 1919–1924.

Du, X., Han, B., Johansson, K.H., and Rikos, A.I. (2026). Affine-coupled distributed optimization via distributed proximal Jacobian ADMM with quantized communication. *arXiv preprint arXiv:2604.14861*.

Du, X., Johansson, K.H., and Rikos, A.I. (2025). Decentralized optimization via RC-ALADIN with efficient quantized communication. In *2025 IEEE 64th Conference on Decision and Control (CDC)*, 4357–4363. IEEE.

Falsone, A., Notarnicola, I., Notarstefano, G., and Prandini, M. (2020). Tracking-ADMM for distributed constraint-coupled optimization. *Automatica*, 117, 108962.

Gabay, D. and Mercier, B. (1976). A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2(1), 17 – 40.

Glowinski, R.and Marroco, A. (1975). Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 9(R2), 41–76.

He, B., Hou, L., and Yuan, X. (2015). On full Jacobian decomposition of the augmented Lagrangian method for separable convex programming. *SIAM Journal on Optimization*, 25(4), 2274–2312.

He, B., Xu, H.K., and Yuan, X. (2016). On the proximal Jacobian decomposition of ALM for multiple-block separable convex minimization problems and its relationship to ADMM. *Journal of Scientific Computing*, 66(3), 1204–1217.

Houska, B., Frasch, J., and Diehl, M. (2016). An augmented Lagrangian based algorithm for distributed nonconvex optimization. *SIAM Journal on Optimization*, 26(2), 1101–1127.

Jiang, W., Doostmohammadian, M., and Charalambous, T. (2022). Distributed resource allocation via ADMM over digraphs. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, 5645–5651. IEEE.

Jiang, W., Grammenos, A., Kalyvianaki, E., and Charalambous, T. (2021). An asynchronous approximate distributed alternating direction method of multipliers in digraphs. In *IEEE Conference on Decision and Control*, 3406–3413.

Li, M. and Yuan, X. (2018). The augmented Lagrangian method with full Jacobian decomposition and logarithmic-quadratic proximal regularization for multiple-block separable convex programming. *The SMAI Journal of computational mathematics*, 4, 81–120.

Lin, Z., Li, H., and Fang, C. (2022). *Alternating direction method of multipliers for machine learning*. Springer.

Ling, Q., Shi, W., Wu, G., and Ribeiro, A. (2015). DLM: Decentralized linearized alternating direction method of multipliers. *IEEE Transactions on Signal Processing*, 63(15), 4051–4064.

Notarnicola, I. and Falsone, A. (2022). Passivity-based analysis of the ADMM algorithm for constraint-coupled optimization. *Automatica*, 146, 110552.

Oliva, G., Setola, R., and Hadjicostis, C.N. (2016). Distributed finite-time calculation of node eccentricities, graph radius and graph diameter. *Systems & Control Letters*, 92, 20–27.

Proakis, J.G. and Salehi, M. (2002). *Communication Systems Engineering*. Prentice Hall, Upper Saddle River, N.J., 2nd edition.

Ren, X., Bastianello, N., Johansson, K.H., and Parisini, T. (2025). Communication-efficient stochastic distributed learning. *arXiv preprint arXiv:2501.13516*.

Rikos, A.I., Hadjicostis, C.N., and Johansson, K.H. (2022). Non-oscillating quantized average consensus over dynamic directed topologies. *Automatica*, 146, 110621.

Rikos, A.I., Jiang, W., Charalambous, T., and Johansson, K.H. (2023a). Asynchronous distributed optimization via ADMM with efficient communication. In *IEEE Conference on Decision and Control*, 7002–7008.

Rikos, A.I., Jiang, W., Charalambous, T., and Johansson, K.H. (2023b). Distributed optimization with gradient descent and quantized communication. *IFAC-PapersOnLine*, 56(2), 5900–5906.

Stomberg, G., Engelmann, A., Diehl, M., and Faulwasser, T. (2025). Decentralized real-time iterations for distributed nmpc. *IEEE Transactions on Automatic Control*.

Wang, X., Hong, M., Ma, S., and Luo, Z.Q. (2013). Solving multiple-block separable convex minimization problems using two-block alternating direction method of multipliers. *arXiv preprint arXiv:1308.5294*.

Wei, J., Yi, X., Sandberg, H., and Johansson, K.H. (2019). Nonlinear consensus protocols with applications to quantized communication and actuation. *IEEE Transactions on Control of Network Systems*, 6(2), 598–608.

Yang, J., Li, A., Liao, X., and Masouros, C. (2024). Low complexity slp: An inversion-free, parallelizable ADMM approach. *IEEE Transactions on Wireless Communications*, 23(9), 12424–12439.