Kybernetika

Chenyang Liu; Xiaohua Dou; Yuan Fan; Songsong Cheng

A penalty ADMM with quantized communication for distributed optimization over multi-agent systems

Kybernetika, Vol. 59 (2023), No. 3, 392–417

Persistent URL: http://dml.cz/dmlcz/151722

KYBERNETIKA — VOLUME *59* (2023), NUMBER 3, PAGES 392–417

A PENALTY ADMM WITH QUANTIZED COMMUNICATION FOR DISTRIBUTED OPTIMIZATION OVER MULTI-AGENT SYSTEMS

Chenyang Liu, Xiaohua Dou, Yuan Fan, and Songsong Cheng

In this paper, we design a distributed penalty ADMM algorithm with quantized communication to solve distributed convex optimization problems over multi-agent systems. Firstly, we introduce a quantization scheme that reduces the bandwidth limitation of multi-agent systems without requiring an encoder or decoder, unlike existing quantized algorithms. This scheme also minimizes the computation burden. Moreover, with the aid of the quantization design, we propose a quantized penalty ADMM to obtain the suboptimal solution. Furthermore, the proposed algorithm converges to the suboptimal solution with an $O(\frac{1}{k})$ convergence rate for general convex objective functions, and with an R-linear rate for strongly convex objective functions.

Keywords: quantized communication, distributed optimization, alternating direction method of multipliers (ADMM), constrained optimization

Classification: 90C33

1. INTRODUCTION

In recent years, the distributed optimization algorithm has been widely used in various fields and has performed well, e. g., [24, 31]. Thus it has attracted the attention of many researchers. There are many excellent algorithms in the distributed optimization literature, including distributed ADMM [7, 20, 30], gradient tracking[17, 22], primal-dual [10, 14, 23, 29], and mirror descent [21, 28]. [7] proposed distributed dual consensus ADMM (DC-ADMM) and the distributed inexact DC-ADMM (IDC-ADMM) to solve the problem of network resource allocation. In [20], the authors showed the distributed ADMM algorithm with a linear rate when objective functions are strongly convex. They also studied the effect of the network topology, the condition number of the objective function, and the algorithm's parameters on the algorithm's convergence rate.

In practical applications, bandwidth limitation is one of the primary and essential topics of multi-agent systems. For distributed algorithms over multi-agent systems, many scholars studied quantized communication to overcome the impact of bandwidth limitation. In [4], the authors proposed a distributed quantized algorithm for seeking a Nash equilibrium of games. [26] developed a distributed subgradient algorithm with

DOI: 10.14736/kyb-2023-3-0392

dynamic quantization for distributed optimization problems. [9] proposed a quantized "projected+consensus" for distributedly solving linear algebraic equations with a linear (sublinear) convergence rate for exact (least squares) solutions. Other related references about quantized distributed optimization include still [3, 6, 15, 19]. However, all of the quantized algorithms in [3, 4, 9, 19, 26] need local encoders and decoders, which increase computation burdens for local agents.

Due to the resources limitation and uncertain environment of practical circumstances, the decision variables of optimization problems are usually constrained by feasible sets. In [11], the authors designed a distributed proximal point algorithm (DPPA) for a smooth constrained optimization problem over an unbalanced time-varying network and achieved an $O(1/\sqrt{k})$ convergence rate for general objective functions. By utilizing an exact penalty method and under strongly convex conditions, [13] developed a distributed projected subgradient algorithm for a nonsmooth optimization problem constrained by a feasible set and achieved an exponential convergence rate. However, the two results are developed based on the identical constraint. As discussed in [1, 5], the nonidentical feasible set constraints are more practical, increase challenges and degenerate convergence performance. In [8], the authors proposed a primal-dual distributed optimization algorithm for nonidentical feasible set constraints and showed an asymptotical convergence rate for general objective functions. Inspired by the exact penalty idea, [32] proposed a primal-dual algorithm for nonsmooth constrained optimization problems with an asymptotical convergence rate. Combining gradient tracking and projected dynamics, [5] proposed projected gradient tracking and achieved an $O(1/k)$ convergence rate for strongly convex objective functions. In [27], the authors designed a push-sum-based constrained optimization algorithm for optimization problems over time-varying directed graphs and showed $O(\ln k)/\sqrt{k}$ and $O(1/k)$ convergence rates for general convex and strongly convex objective functions respectively. According to existing results, we wonder that is it possible to improve the convergence rate by sacrificing little convergence accuracy.

Inspired by the above discussions, we propose a distributed penalty ADMM algorithm with quantized communication to solve distributed constrained optimization problems in this paper. The main contributions of this paper are summarized as follows.

1) In comparison with the ADMM algorithms for the unconstraint optimization problems in [20, 30], we developed a penalty ADMM for distributedly solving distributed constrained optimization problems.

2) We utilize a novel quantized communication scheme for the distributed algorithms and remove the local encoders and decoders, which are necessary in [3, 4, 9, 19, 26].

3) Compared with the quantizers with fixed quantization intervals in [15, 26], we apply the quantizer with dynamic quantization intervals to the penalty ADMM (P-ADMM), which reduces the quantized error and improves the convergence accuracy.

4) We analyze the convergence performance of the proposed algorithm and show it sublinearly (linearly) converges to the approximated solution of generally (strongly) convex optimization problems. Moreover, the convergence rates are faster than that of [5, 8, 27, 32]

The rest of our paper is organized as follows. In Section 2, we first explain the knowledge of graph theory and the meaning of some necessary symbols. Then we briefly introduce the original algorithm and a quantizer used in our proposed QP-ADMM. We show the derivation of the proposed QP-ADMM in Subsection 3.1. We analyze the convergence performance of our algorithm in general convex and strongly convex cases in Subsection 3.2 and Subsection 3.3, respectively. In Section 4, we demonstrate the performance of our proposed algorithm through some simulation examples. Section 5 concludes this paper.

2. PRELIMINARIES

2.1. Graph theory

In this paper, we consider distributed optimization problems over multi-agent systems connected by a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. For the graph $\mathcal{G}$, $\mathcal{N} \triangleq \{1, 2, \cdots, n\}$ is the set of agents and $\mathcal{E}$ is composed of a pair of two different agents in $\mathcal{N}$. $\mathcal{N}_i \triangleq \{j \in \mathcal{N} : (i, j) \in \mathcal{E}$ or $(i, j) \in \mathcal{E}\} \cup \{i\}$ is the set of neighboring agents of i and includes itself.

2.2. Notation

$x_i \in \mathbb{R}^m$ and $x \in \mathbb{R}^{n\times m}$ are the local variable held by each agent i and the global variable, respectively, i. e., $x \triangleq [x_1^{\mathrm{T}}; x_2^{\mathrm{T}}; \cdots; x_n^{\mathrm{T}}] \in \mathbb{R}^{n\times m}$. If all local variables satisfy the following condition, $x_1 = x_2 = \cdots = x_n$, then the global variable x is consensual. Denote x^k and x_i^k as the values of x and x_i at the kth iteration, respectively.

For a vector v, $\|v\|_1$ and $\|v\|$ denote its l_1 and l_2 norms, respectively. For a matrix $B = [b_{ij}] \in \mathbb{R}^{n\times m}$, b_{ij} is the element of B at the ith row and jth column and $\|B\|_{\mathrm{F}}$ is the Frobenius norm of the matrix B. For a symmetric positive semidefinite matrix $G \succ 0$, we denote $\|B\|_G \triangleq \sqrt{\mathrm{trace}(B^{\mathrm{T}}GB)}$ and $\langle A, B\rangle_G \triangleq \langle A, GB\rangle$ are the induced norm of B and the inner product of A and B, respectively.

We define $\mu(\cdot)$ $(\sigma(\cdot))$, $\mu_{\max}(\cdot)$ $(\sigma_{\max}(\cdot))$, and $\mu_{\min}(\cdot)$ $(\sigma_{\min}(\cdot))$ are the eigenvalue (singular value), the largest eigenvalue (singular value), and the smallest eigenvalue (singular value) of a given matrix, respectively. I_n and $\mathbf{1}_{n\times m}$ denote an n-dimensional identity matrix and $n \times m$ all-ones matrix, respectively.

2.3. Problem formulation

In this paper, we introduce a distributed penalty ADMM algorithm (P-ADMM) in [30] to solve the following distributed convex optimization problem with feasible constraint sets,

$$\begin{aligned} &\min_{x_i \in \mathbb{R}^m} \frac{1}{n}\sum_{i=1}^n f_i(x_i), \\ &\text{s.t. } x_i \in \mathbb{K}_i, \end{aligned} \tag{1}$$

where $f_i(x_i) : \mathbb{R}^m \to \mathbb{R}$ is a convex and continuously differentiable function which is only owned by the agent i and $\mathbb{K}_i \subseteq \mathbb{R}^m$ is a compact and convex constraint that is only held by the corresponding local variable x_i. For (1), we have the following mild assumptions in the distributed optimization literature.

Assumption 1. For f_i, its gradient $\nabla f_i(x)$ is Lipschitz continuous; i. e.,

$$\|\nabla f_i(x) - \nabla f_i(y)\| \leq L_f \|x - y\|, \quad \forall x, y \in \mathbb{R}^m,$$

where $L_f > 0$ is the Lipschitz constant.

Assumption 2. Each f_i is strongly convex; i. e.;

$$\langle x - y, \nabla f_i(x) - \nabla f_i(y)\rangle \geq \mu_f \|x - y\|^2, \quad \forall x, y \in \mathbb{R}^m,$$

where $\mu_f > 0$ is the strong convexity constant.

To reformulate (1) into the form with consensual constraint, we introduce a mixing matrix $W \in \mathbb{R}^{n\times n}$, where each element w_{ij} in W is a weight between agents i and j. For the mixing matrix W, the following assumption is adopted.

Assumption 3. The graph $\mathcal{G}$ is connected and undirected. The mixing matrix W is doubly stochastic and symmetric; i. e., $W\mathbf{1}_{n\times 1} = \mathbf{1}_{n\times 1}$ and $W = W^{\mathrm{T}}$. If $j \notin \mathcal{N}_i$, $w_{ij} = 0$; otherwise, $w_{ij} > 0$.

Remark 1. Based on the Perron-Frobenius theorem in [18], the eigenvalues of W lie in $(-1, 1]$ and the multiplicity of the largest eigenvalue is one which implies that $\mathrm{span}(\mathbf{1}_{\mathrm{n}\times 1})$ is the null space of $I - W$. Therefore the null space of the square root of $I - W$ is the same as it, i. e., $(I - W)^{\frac{1}{2}} x = 0$ if and only if $x_1^{\mathrm{T}} = x_2^{\mathrm{T}} = \cdots = x_n^{\mathrm{T}}$.

We rewrite (1) as the following form

$$\hat{x}^* \in \underset{x}{\operatorname{argmin}}\, f(x), \quad \text{s.t. } (I - W)^{\frac{1}{2}} x = 0,\ x \in \mathbb{K}, \tag{2}$$

where $\hat{x}^* \in \mathbb{R}^{n\times m}$ is the optimal solution, $f(x) \triangleq \sum_{i=1}^n f_i(x_i)$, and $\mathbb{K} = \prod_{i=1}^n \mathbb{K}_i \subseteq \mathbb{R}^{n\times m}$ is the global constraint set which is held by the global variable x.

For the convenience of solving the consensus constrained problem in (2), we transfer it as the following approximated form

$$x^* \in \underset{x}{\operatorname{argmin}}\ f(x) + \tfrac{1}{2\epsilon}\|(I - W)^{\frac{1}{2}} x\|_{\mathrm{F}}^2, \quad \text{s.t. } x \in \mathbb{K}, \tag{3}$$

where $\epsilon > 0$ is the penalty parameter and x^* is the optimal solution of (3).

Remark 2. It is worth emphasizing that the optimal solutions of (2) $\hat{x}^*$ and (3) x^* are not equal. Moreover, the error between $\hat{x}^*$ and (3) x^* is scaled with the parameter ϵ. Namely, a smaller ϵ yields a more accurate solution.

Assumption 4. The set of minimizers of (1) and (3) are nonempty.

2.4. Quantizer design

Define a quantizer (4) with a fixed number of b bits as follows.

$$Q(c) = \begin{cases} \bar{c} - \frac{l}{2} & \text{if } c \in (-\infty, \bar{c} - \frac{l}{2}), \\ \bar{c} + \mathrm{sgn}(c - \bar{c})\Delta \lfloor \frac{|c-\bar{c}|}{\Delta} \rfloor + \frac{\Delta}{2} & \text{if } c \in [\bar{c} - \frac{l}{2}, \bar{c} + \frac{l}{2}], \\ \bar{c} + \frac{l}{2} & \text{if } c \in (\bar{c} + \frac{l}{2}, \infty), \end{cases} \tag{4}$$

where c is a real number and $\mathrm{sgn}(\cdot)$ is a sign function. $\bar{c}$, l, and $\Delta = \frac{l}{2^b}$ are the middle value, the quantization interval, and the quantization step size of (4), respectively. For (4), we draw on Figure 1 to better illustrate the principle of the quantizer. Based on

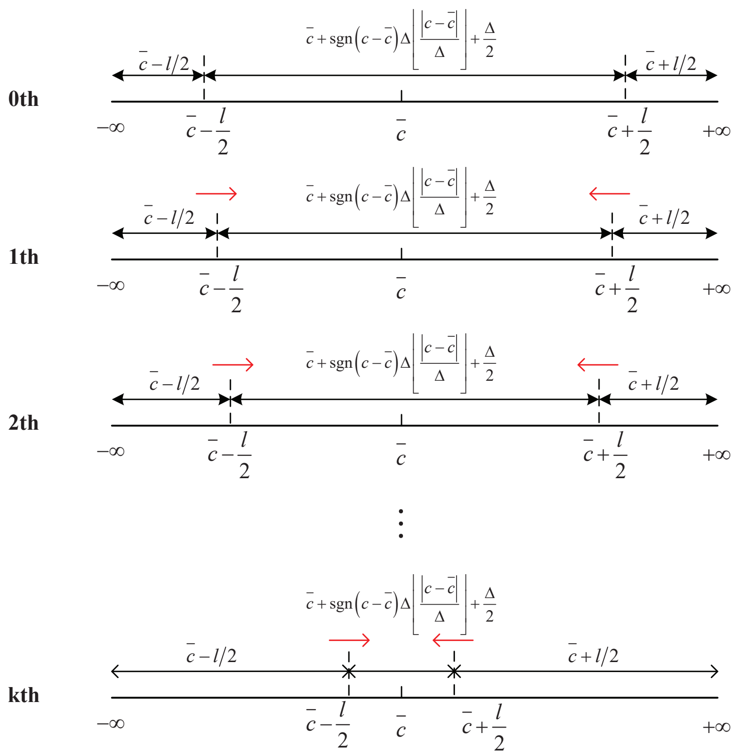

Fig. 1. The principle of the quantizer.

(4), if the quantized value c lies in the quantization interval $c \in [\bar{c} - \frac{l}{2}, \bar{c} + \frac{l}{2}]$, we bound the quantized error as follows

$$|c - Q(c)| \le \frac{\Delta}{2} = \frac{l}{2^{b+1}}. \tag{5}$$

In this paper, we denote the quantized value of x and x_i as x^Q and x_i^Q, $\forall i \in \mathcal{N}$, respectively. Since the middle value $\bar{c}$ and the quantization interval l change at each iteration. Thus at the k-th iteration, we define that the middle value $\bar{x}_i^k$ is previous quantized value $x_i^{Q,k-1}$, i. e.,

$$\bar{x}_i^k = x_i^{Q,k-1},$$

and the quantization interval is

$$l_i^k = C\theta^k,$$

where $\theta \in (0,1)$ and C is the initial quantization interval l_i^0. The error generated by the quantizer (4) is defined as

$$e_i^k = x_i^{Q,k} - x_i^k.$$

3. ALGORITHM DESIGN AND CONVERGENCE ANALYSIS

In this section, we propose a quantized penalty ADMM algorithm (QP-ADMM) based on P-ADMM to solve distributed constrained convex optimization problems. Firstly, we explain how our proposed QP-ADMM is derived from the original algorithm. Next we analyze the convergence performance of QP-ADMM under general convex and strongly convex cases, respectively.

3.1. Algorithm design

For (3), we define an auxiliary variable $z = (I - W)^{\frac{1}{2}} x \in \mathbb{R}^{n\times m}$ and substitute it into (3)

$$\begin{aligned} \min_{x,z} \quad & f(x) + \tfrac{1}{2\epsilon}\|z\|_{\mathrm{F}}^2, \\ \text{s.t.} \quad & L^{\frac{1}{2}} x = z, x \in \mathbb{K}, \end{aligned} \tag{6}$$

where $L \triangleq I - W$.

Based on (6), we have the following augmented Lagrangian function

$$L_\alpha(x,z,\lambda) \triangleq f(x) + \tfrac{1}{2\epsilon}\|z\|_{\mathrm{F}}^2 + \langle \lambda, L^{\frac{1}{2}} x - z\rangle + \tfrac{\alpha}{2}\|L^{\frac{1}{2}} x - z\|_{\mathrm{F}}^2, \tag{7}$$

where $\lambda \in \mathbb{R}^{n\times m}$ is the Lagrange multiplier and $\alpha > 0$.

When the proposed algorithm runs, values of x may fall outside feasible constraint sets. Using these can result in severe calculation errors. Therefore, we project x onto the constraint set $\mathbb{K}$ to avoid it. Based on (7), the updating law is given as follows.

$$x:\ x^{k+1} = \mathrm{Proj}_{\mathbb{K}}(x^k - c[\nabla f(x^k) + \alpha L x^k - \alpha \tilde{z}^k + \tilde{\lambda}^k]), \tag{8a}$$

$$\tilde{z}:\ \tilde{z}^{k+1} = \tfrac{1}{\alpha + \frac{1}{\epsilon}}[\tilde{\lambda}^k + \alpha L x^{k+1}], \tag{8b}$$

$$\tilde{\lambda}:\ \tilde{\lambda}^{k+1} = \tilde{\lambda}^k + \alpha[L x^{k+1} - \tilde{z}^{k+1}], \tag{8c}$$

where $\tilde{z}^k \triangleq L^{\frac{1}{2}} z^k$ and $\tilde{\lambda}^k \triangleq L^{\frac{1}{2}} \lambda^k$. Then we add the quantizer (4) after the x-update of (8a),

$$x: \ x^{k+1} = \mathrm{Proj}_{\mathbb{K}}(x^{Q,k} - c[\nabla f(x^{Q,k}) + \alpha L x^{Q,k} - \alpha \tilde{z}^k + \tilde{\lambda}^k]), \tag{9a}$$

$$x^Q: \ x^{Q,k+1} = Q(x^{k+1}), \tag{9b}$$

$$\tilde{z}: \ \tilde{z}^{k+1} = \frac{1}{\alpha + \frac{1}{\epsilon}}[\tilde{\lambda}^k + \alpha L x^{Q,k+1}], \tag{9c}$$

$$\tilde{\lambda}: \ \tilde{\lambda}^{k+1} = \tilde{\lambda}^k + \alpha[L x^{Q,k+1} - \tilde{z}^{k+1}]. \tag{9d}$$

Assumption 5. The set $\mathbb{K}$ is compact and convex. If $x \in \mathbb{K}$, the quantized value $x^{Q,k}$ is bounded as

$$\|x^{Q,k} - x^*\|_{\mathrm{F}} \leq A,$$

where $A > 0$ is a positive constant and x^* is the optimal solution of (6).

In summary, (9) is the updating law of QP-ADMM. The detailed implementation of the proposed algorithm is shown in Algorithm 1.

Algorithm 1. Quantized Penalty ADMM

Initialization: Choose the parameters ϵ, α, and c. Initialize $x_i^0 = \mathbf{1}_{m\times 1}$, $\tilde{z}_i^0 = \mathbf{1}_{m\times 1}$, and $\tilde{\lambda}_i^0 = \mathbf{1}_{m\times 1}$.

Update flows: For each $i \in \mathcal{N}$,

for $k = 1, 2, \ldots$ **do**

1 : Compute the projection of local variable x_i^{k+1} by
$x_i^{k+1} = \mathrm{Proj}_{\mathbb{K}_i}(x_i^{Q,k} - c[\nabla f_i(x_i^{Q,k}) + \alpha(x_i^{Q,k} - \sum_{j\in\mathcal{N}_i} w_{ij} x_j^{Q,k} - \tilde{z}_i^k) + \tilde{\lambda}_i^k])$.

2 : Update the parameters of the quantizer: Q_i^{k+1}: l_i^{k+1} and $\bar{x}_i^{k+1}$.

3 : Quantize the local variable: $x_i^{Q,k+1} = Q_i^{k+1}(x_i^{k+1})$.

4 : Transmit $x_i^{Q,k+1}$ to / receive $x_j^{Q,k+1}$ from neighbors $j \in \mathcal{N}_i$.

5 : Update local auxiliary variable $\tilde{z}_i^{k+1}$ by
$\tilde{z}_i^{k+1} = \frac{1}{\alpha+\frac{1}{\epsilon}}[\tilde{\lambda}_i^k + \alpha(x_i^{Q,k+1} - \sum_{j\in\mathcal{N}_i} w_{ij} x_j^{Q,k+1})]$.

6 : Update local dual variable $\tilde{\lambda}_i^{k+1}$ by
$\tilde{\lambda}_i^{k+1} = \tilde{\lambda}_i^k + \alpha[(x_i^{Q,k+1} - \sum_{j\in\mathcal{N}_i} w_{ij} x_j^{Q,k+1}) - \tilde{z}_i^{k+1}]$.

end for

Remark 3. To reduce the communication burdens of distributed solving (2), we design Algorithm 1 with the following two intuitions. Firstly, Algorithm 1 is designed with a quantized communication scheme, where each agent exchanges information with its neighbors in limited bandwidth; Secondly, inspired by [20, 30], each agent only exchanges the primal variables with its neighbors in Algorithm 1, which is more efficient in communication resources saving than the exchanging of both primal and dual counterparts in [8, 14].

3.2. Convergence analysis

In this subsection, we analyze the convergence of QP-ADMM under general convexity and strong convexity.

$\{\tilde{z}^k\}$ and $\{\tilde{\lambda}^k\}$ in (9) denote auxiliary variables and dual variables, respectively. To simplify the notation, we use $\{z^k\}$ and $\{\lambda^k\}$ to replace them. Then (9) becomes

$$x: \ x^{k+1} = \mathrm{Proj}_{\mathbb{K}}(x^{Q,k} - c[\nabla f(x^{Q,k}) + \alpha L x^{Q,k} - \alpha L^{\frac{1}{2}} z^k + L^{\frac{1}{2}}\lambda^k]), \tag{10a}$$

$$x^Q: \ x^{Q,k+1} = Q(x^{k+1}), \tag{10b}$$

$$z: \ z^{k+1} = \frac{1}{\alpha+\frac{1}{\epsilon}}[\lambda^k + \alpha L^{\frac{1}{2}} x^{Q,k+1}], \tag{10c}$$

$$\lambda: \ \lambda^{k+1} = \lambda^k + \alpha[L^{\frac{1}{2}} x^{Q,k+1} - z^{k+1}]. \tag{10d}$$

Rearranging (10d), we have

$$\lambda^k = \lambda^{k+1} - \alpha L^{\frac{1}{2}} x^{Q,k+1} + \alpha z^{k+1}. \tag{11}$$

Substituting (11) into (10c), we get

$$(\alpha + \tfrac{1}{\epsilon}) z^{k+1} = \lambda^{k+1} + \alpha z^{k+1},$$

which implies that

$$\tfrac{1}{\epsilon} z^{k+1} = \lambda^{k+1}. \tag{12}$$

As long as initializing with $\frac{1}{\epsilon} z^0 = \lambda^0$, then we have $\frac{1}{\epsilon} z^k = \lambda^k$.

3.2.1. General convex objective functions

We provide detailed proofs and some critical conclusions in this part to establish that the proposed algorithm converges to the optimal solution under the general convex case.

To simplify notations, we define

$$u^k_{[Q]} \triangleq (x^{Q,k}, z^k, \lambda^k), \quad u^* \triangleq (x^*, z^*, \lambda^*), \quad G \triangleq I - \alpha c(I - W), \quad H \triangleq (\tfrac{1}{2c} G, \tfrac{\alpha}{2} I, \tfrac{1}{2\alpha} I),$$

where $u^k_{[Q]}$ is generated by (10) and u^* is the optimal solution of (6). The square of the distance from $u^k_{[Q]}$ to u^* is defined as $\|u^k_{[Q]} - u^*\|^2_H \triangleq \|x^{Q,k} - x^*\|^2_{\frac{1}{2c}G} + \|z^k - z^*\|^2_{\frac{\alpha}{2}I} + \|\lambda^k - \lambda^*\|^2_{\frac{1}{2\alpha}I}$.

Theorem 3.1. Under Assumptions 1, 3, 4, and 5, if the parameters α and c are chosen such that $\frac{1}{2c}G - \frac{L_f}{2} I \succ 0$, then

$$\|u^k_{[Q]} - u^*\|^2_H - \|u^{k+1}_{[Q]} - u^*\|^2_H + \Lambda \geq \tau \|x^{Q,k} - x^{k+1}\|^2_G + \tfrac{1}{\epsilon}\|z^k - z^{k+1}\|^2_{\mathrm{F}} + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|^2_{\mathrm{F}}, \tag{13}$$

where $\tau > 0$ is a positive constant and Λ is the error term and

$$\Lambda \triangleq \frac{1 + 2\alpha\epsilon}{2\epsilon} \frac{npC^2}{2^{2(b+1)}} \theta^{2(k+1)} + \frac{\sqrt{np}AC}{2^{b+1}c} \theta^{k+1}.$$

Proof. See Appendix 6.1. □

Based on Theorem 3.1, $\|u^k_{[Q]} - u^*\|^2_H$ falls fast and tends to 0 when the selected α and c satisfy the condition $\frac{1}{2c}G - \frac{L_f}{2}I \succ 0$. However, Theorem 3.1 alone cannot show that QP-ADMM converges to the optimal solution of (6) under the case of general convexity. Thus we provide the following theorem, whose proof is similar to that of Theorem 3.2 in [16].

Theorem 3.2. Under Assumptions 1, 3, 4, and 5, the sequence $\{x^{Q,k}, z^k, \lambda^k\}$ produced by (10) converges to the optimal solutions (x^*, z^*, λ^*) of (6) from any starting point, when α and c are the same as chosen in Theorem 3.1.

Proof. See Appendix 6.2. □

Theorem 3.2 establishes that QP-ADMM converges to the optimal solution u^* of (6). Theorem 3.1 and Theorem 3.2 tell us how to choose α and c, i.e., $c < \frac{1}{L_f + \alpha\mu_{\max}(L)}$, based on $\frac{1}{2c}G - \frac{L_f}{2}I \succ 0$, where $\mu_{\max}(L)$ is the largest eigenvalue of L.

We have shown that QP-ADMM can converge to the optimal solution u^* of (6) under the generally convex case. We propose the following theorem to establish that the proposed algorithm converges at the rate of $O(\frac{1}{k})$ under general convexity.

We define the formula in (6) as $\psi(\nu) \triangleq f(x) + \frac{1}{2\epsilon}\|z\|^2_{\mathrm{F}}$, where $\nu = (x, z)$ is a triple. Let $\breve{\nu}^k \triangleq (\breve{x}^k, \breve{z}^k)$ with $\breve{x}^k \triangleq \frac{1}{k}\sum_{t=1}^k x^t$ and $\breve{z}^k \triangleq \frac{1}{k}\sum_{t=1}^k z^t$.

Theorem 3.3. Under Assumptions 1, 3, 4, and 5, if the parameters α and c are chosen as in Theorem 3.1, it holds for all $k \geq 1$ that

$$\max\left\{|\psi(\breve{\nu}^k) - \psi(\nu^*)|, \|\lambda^*\|_{\mathrm{F}}\|L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\|_{\mathrm{F}}\right\} \leq \frac{Q_1}{k},$$

where $Q_1 \triangleq \frac{1}{2c}\|x^* - x^{Q,0}\|^2_G + \frac{\alpha}{2}\|z^* - z^0\|^2_{\mathrm{F}} + \frac{4}{\alpha}\|\lambda^*\|^2_{\mathrm{F}} - \frac{1}{\alpha}\|\lambda^0\|^2_{\mathrm{F}} + \Lambda_1 + \Lambda_2$ with $\Lambda_1 \triangleq \frac{1+2\alpha\epsilon}{2\epsilon}\frac{npC^2}{2^{2(b+1)}}\frac{\theta^2}{1-\theta^2}$ and $\Lambda_2 \triangleq \frac{\sqrt{npAC}}{2^{b+1}c}\frac{\theta}{1-\theta}$.

Proof. See Appendix 6.3. □

3.2.2. Strongly convex objective functions

If the function $f(x)$ is strongly convex, QP-ADMM can achieve a linear convergence rate. The proof of this conclusion is as follows.

Theorem 3.4. Under Assumption 1-5, if the variables are initialized as $\frac{1}{\epsilon}z^0 = \lambda^0$ and $\delta \triangleq 1 - \frac{L_f^2 c}{\mu_{\min}(G)\mu_f} > 0$, the convergence performance of the sequence $\{x^{Q,k}, z^k\}$ which is generated by (10) satisfies

$$\begin{aligned} &c_2\|x^* - x^{Q,k+1}\|^2_G + \|z^* - z^{k+1}\|^2_{\mathrm{F}} \\ &\leq \tau_2^{k+1}(c_2\|x^* - x^{Q,0}\|^2_G + \|z^* - z^0\|^2_{\mathrm{F}} + \Lambda_3 + \Lambda_4), \end{aligned} \tag{14}$$

where $0 < \theta < \tau_2 < 1$ and $\tau_2 \triangleq \max\{\frac{1}{c_1 c+1}, \frac{2\alpha^2\epsilon^2+2}{\alpha\epsilon+2\alpha^2\epsilon^2+2}\} < 1$ with $c_1 \triangleq \frac{\mu_f \delta}{(1+\delta)\mu_{\max}(G)}$ and $c_2 \triangleq \frac{c_1+\frac{1}{c}}{\frac{1}{2\epsilon}+\alpha+\frac{1}{\alpha\epsilon^2}}$. Define $\Lambda_3 \triangleq \frac{a_1}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}} \frac{1}{1-\frac{\theta^2}{\tau_2}}$ with $a_1 \triangleq \frac{1+2\alpha\epsilon}{\epsilon} \frac{npC^2\theta^2}{2^{2(b+1)}}$ and $\Lambda_4 \triangleq \frac{a_2}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}} \frac{1}{1-\frac{\theta}{\tau_2}}$ with $a_2 \triangleq \frac{\sqrt{np}AC\theta}{c2^b}$, respectively.

P r o o f. See Appendix 6.4. □

Remark 4. Based on Theorem 3.4, the sequence $\{x^{Q,k}, z^k\}$ generated by (10) convergences to the optimal solutions x^* and z^* of (6) with R-linear rate under the strongly convex case.

4. NUMERICAL SIMULATIONS

This section demonstrates the effectiveness of the proposed algorithm with a numerical example. We utilize QP-ADMM to solve a distributed quadratic programming problem that satisfies Assumptions 1-5. We set the mixing matrix W based on the Metropolis-Hasting rule in [2]. The form of the problem is as follows.

$$\begin{aligned} &\min_{\{x_i\}} \sum_{i=1}^{n} x_i^{\mathrm{T}} a_i x_i - b_i^{\mathrm{T}} x_i, \\ &\text{s.t. } x_i \in \mathbb{K}_i,\ x_i = x_j,\ i,j \in \mathcal{N}. \end{aligned} \tag{15}$$

In the multi-agent system, the number of agents is $n = 4$. The network topology of this multi-agent system is shown in Figure 2. To use QP-ADMM to solve (15), the parameters a_i and b_i, $i \in \mathcal{N}$ are chosen as follows,

$$a_1 = \begin{bmatrix} -1.20 & 0.50 \\ -0.65 & 0.40 \end{bmatrix},\ b_1 = \begin{bmatrix} 0.0447 \\ -0.1342 \end{bmatrix},$$
$$a_2 = \begin{bmatrix} -0.40 & 1.15 \\ -1.00 & 1.66 \end{bmatrix},\ b_2 = \begin{bmatrix} 0.7155 \\ -0.9704 \end{bmatrix},$$
$$a_3 = \begin{bmatrix} -0.60 & 0.80 \\ -2.00 & 1.00 \end{bmatrix},\ b_3 = \begin{bmatrix} 1.5205 \\ -0.5366 \end{bmatrix},$$
$$a_4 = \begin{bmatrix} -1.25 & -2.47 \\ -1.00 & -0.76 \end{bmatrix},\ b_4 = \begin{bmatrix} 0.3354 \\ -0.4248 \end{bmatrix}.$$

The optimal solution $\hat{x}^*$ of (15) is

$$\hat{x}^* = \begin{bmatrix} 0.4472 & -0.8944 \\ 0.4472 & -0.8944 \\ 0.4472 & -0.8944 \\ 0.4472 & -0.8944 \end{bmatrix}.$$

All local constraint sets $\mathbb{K}_i$, $i \in \mathcal{N}$ and the global constraint set $\mathbb{K}$ are the unit circle.

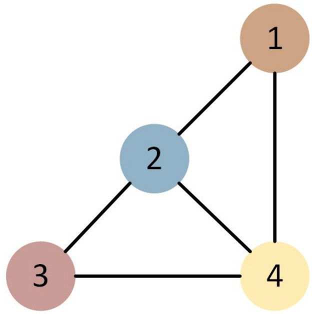

Fig. 2. The network topology of the multi-agent system.

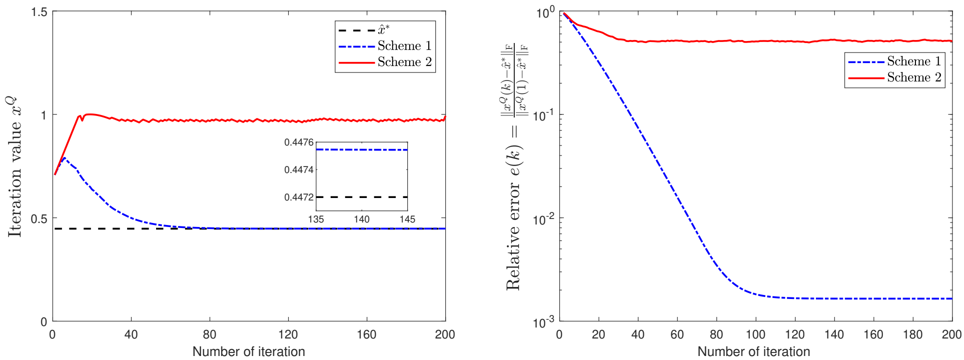

Fig. 3. The convergence performance of the QP-ADMM under schemes 1 and 2.

We design two schemes to demonstrate the dynamic quantizer scheme we use. One is QP-ADMM with the dynamic quantizer and the other is QP-ADMM with the traditional quantizer that uses fixed quantization intervals. We set step sizes α and c are 40 and 0.0093 for the penalty parameter $\epsilon = 10^{-2}$, respectively. We initialize quantizers with the following schemes,

$$\begin{cases} \text{Scheme 1}: l^0 = 6,\ b = 6,\ \Delta^0 = \frac{l^0}{2^b} = 0.094,\ l^k = l^0\theta^k,\ \theta = 0.89; \\ \text{Scheme 2}: l^0 = 6,\ b = 6,\ \Delta^0 = \frac{l^0}{2^b} = 0.094,\ l^k = l^0. \end{cases}$$

As shown in Figure 3, there are errors in both schemes. The errors are yielded from the penalty approximation and the quantized communication. But the final value of Scheme 1 is more accurate than that of Scheme 2. This accuracy does not sacrifice the convergence speed.

In Figure 4, all parameters are the same as in Scheme 1 except for the penalty parameter ϵ. Figure 4 shows that QP-ADMM can quickly converge to the neighborhood of the optimal value $\hat{x}^*$ even when the penalty parameter ϵ is small, its convergence rate

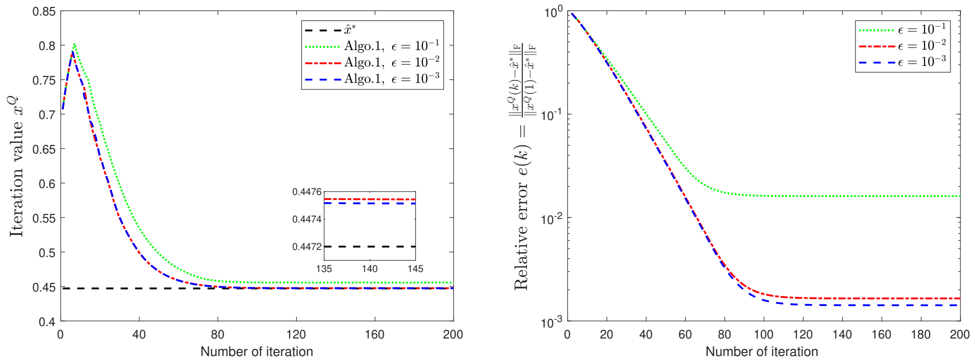

Fig. 4. The convergence performance of the QP_ADMM with different ϵ.

is linear, which is consistent with our previous Theorem 3.4. The excellent convergence performance of the proposed algorithm does not lead to a decrease in the accuracy of the multi-agent system.

5. CONCLUSION

In this paper, we proposed a distributed ADMM algorithm over quantized communication with feasible set constraints called QP-ADMM. We presented a detailed derivation of the proposed algorithm from the original algorithm. Firstly, we introduced a dynamic quantizer and applied it to the original algorithm. Then we analyzed our algorithm's convergence performance in general convexity and strong convexity and proposed some critical theorems. Finally, to illustrate the effectiveness of our proposed algorithm, we provided a numerical example.

6. APPENDIX

6.1. Proof of Theorem 3.1

Based on the property of variational inequalities, we rewrite (10a),

$$\langle x^* - x^{k+1}, x^{k+1} - (x^{Q,k} - c[\nabla f(x^{Q,k}) + \alpha L x^{Q,k} - \alpha L^{\frac{1}{2}} z^k + L^{\frac{1}{2}} \lambda^k])\rangle \geq 0, \tag{16}$$

since $x^* \in \mathbb{K}$. Substituting (11) into (16), we have

$$\begin{aligned} 0 \leq \langle x^* - x^{k+1}, \tfrac{1}{c}[x^{k+1} - x^{Q,k} - (\alpha c L x^{Q,k+1} - \alpha c L x^{Q,k})] + \alpha L^{\frac{1}{2}} z^{k+1} - \alpha L^{\frac{1}{2}} z^k \\ + \nabla f(x^{Q,k}) + L^{\frac{1}{2}} \lambda^{k+1}\rangle. \end{aligned}$$

Rearranging the above formulation,

$$0 \leq \underbrace{\tfrac{1}{c}\langle x^* - x^{k+1}, G(x^{Q,k+1} - x^{Q,k})\rangle}_{(\mathrm{T}_1)} + \underbrace{\langle x^* - x^{k+1}, \nabla f(x^{Q,k})\rangle}_{(\mathrm{T}_2)} - \tfrac{1}{c}\langle x^* - x^{k+1}, e^{k+1}\rangle$$
$$+ \langle x^* - x^{k+1}, L^{\frac{1}{2}}\lambda^{k+1}\rangle + \alpha\langle L^{\frac{1}{2}}(x^* - x^{k+1}), z^{k+1} - z^k\rangle. \tag{17}$$

For the term (T_1) in (17), we rewrite it by the identity $(v_1 - v_2)^{\mathrm{T}} G(v_3 - v_4) = \frac{1}{2}(\|v_1 - v_4\|_G^2 - \|v_1 - v_3\|_G^2) + \frac{1}{2}(\|v_2 - v_3\|_G^2 - \|v_2 - v_4\|_G^2)$,

$$\begin{aligned} &\tfrac{1}{c}\langle x^* - x^{k+1}, G(x^{Q,k+1} - x^{Q,k})\rangle \\ =\ &\tfrac{1}{2c}(\|x^* - x^{Q,k}\|_G^2 - \|x^* - x^{Q,k+1}\|_G^2) + \tfrac{1}{2c}(\|e^{k+1}\|_G^2 - \|x^{k+1} - x^{Q,k}\|_G^2). \end{aligned} \tag{18}$$

Under Assumption 1, we rewrite the term (T_2) in (17) as follows

$$\begin{aligned} &\langle x^* - x^{k+1}, \nabla f(x^{Q,k})\rangle \\ =\ &\langle x^* - x^{Q,k}, \nabla f(x^{Q,k})\rangle + \langle x^{Q,k} - x^{k+1}, \nabla f(x^{Q,k})\rangle \\ \leq\ &f(x^*) - f(x^{Q,k}) + f(x^{Q,k}) - f(x^{k+1}) + \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2 \\ =\ &f(x^*) - f(x^{k+1}) + \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2. \end{aligned} \tag{19}$$

Substituting (18) and (19) into (17), we have

$$\begin{aligned} 0 \leq\ &\tfrac{1}{2c}(\|x^* - x^{Q,k}\|_G^2 - \|x^* - x^{Q,k+1}\|_G^2) + \tfrac{1}{2c}\|e^{k+1}\|_G^2 + f(x^*) - f(x^{k+1}) \\ &+ \underbrace{\alpha\langle L^{\frac{1}{2}}(x^* - x^{k+1}), z^{k+1} - z^k\rangle}_{(\mathrm{T}_3)} - \tfrac{1}{2c}\|x^{k+1} - x^{Q,k}\|_G^2 + \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2 \\ &+ \langle x^* - x^{k+1}, L^{\frac{1}{2}}\lambda^{k+1}\rangle - \tfrac{1}{c}\langle x^* - x^{k+1}, e^{k+1}\rangle. \end{aligned} \tag{20}$$

With the help of another identity $(v_1 - v_2)^{\mathrm{T}}(v_3 - v_4) = \frac{1}{2}(\|v_1 - v_4\|^2 - \|v_1 - v_3\|^2) + \frac{1}{2}(\|v_2 - v_3\|^2 - \|v_2 - v_4\|^2)$, the term (T_3) in (20) becomes

$$\begin{aligned} &\alpha\langle L^{\frac{1}{2}}(x^* - x^{k+1}), z^{k+1} - z^k\rangle \\ =\ &\alpha\langle L^{\frac{1}{2}}(x^* - x^{Q,k+1}), z^{k+1} - z^k\rangle + \alpha\langle L^{\frac{1}{2}}e^{k+1}, z^{k+1} - z^k\rangle \\ =\ &\tfrac{\alpha}{2}(\|L^{\frac{1}{2}}x^* - z^k\|_{\mathrm{F}}^2 - \|L^{\frac{1}{2}}x^* - z^{k+1}\|_{\mathrm{F}}^2) + \alpha\langle L^{\frac{1}{2}}e^{k+1}, z^{k+1} - z^k\rangle \\ &+ \tfrac{\alpha}{2}(\|L^{\frac{1}{2}}x^{Q,k+1} - z^{k+1}\|_{\mathrm{F}}^2 - \|L^{\frac{1}{2}}x^{Q,k+1} - z^k\|_{\mathrm{F}}^2). \end{aligned} \tag{21}$$

Rearranging (10d),

$$\frac{\lambda^{k+1} - \lambda^k}{\alpha} = L^{\frac{1}{2}}x^{Q,k+1} - z^{k+1}. \tag{22}$$

Substituting the above formula into (21)

$$\begin{aligned}
&\alpha\langle L^{\frac{1}{2}}(x^* - x^{k+1}), z^{k+1} - z^k\rangle \\
&= \tfrac{\alpha}{2}(\|L^{\frac{1}{2}}x^* - z^k\|_{\mathrm{F}}^2 - \|L^{\frac{1}{2}}x^* - z^{k+1}\|_{\mathrm{F}}^2) + \alpha\langle L^{\frac{1}{2}}e^{k+1}, z^{k+1} - z^k\rangle \\
&\quad + \tfrac{\alpha}{2}(\|\tfrac{\lambda^{k+1}-\lambda^k}{\alpha}\|_{\mathrm{F}}^2 - \|\tfrac{\lambda^{k+1}-\lambda^k}{\alpha} + z^{k+1} - z^k\|_{\mathrm{F}}^2) \\
&= \tfrac{\alpha}{2}(\|L^{\frac{1}{2}}x^* - z^k\|_{\mathrm{F}}^2 - \|L^{\frac{1}{2}}x^* - z^{k+1}\|_{\mathrm{F}}^2) - \langle\lambda^{k+1} - \lambda^k, z^{k+1} - z^k\rangle - \tfrac{\alpha}{2}\|z^{k+1} - z^k\|_{\mathrm{F}}^2 \\
&\quad + \alpha\langle L^{\frac{1}{2}}e^{k+1}, z^{k+1} - z^k\rangle.
\end{aligned} \tag{23}$$

Then, we substitute three equations $\frac{1}{\epsilon}z^k = \lambda^k$, $\frac{1}{\epsilon}z^{k+1} = \lambda^{k+1}$ and $L^{\frac{1}{2}}x^* = z^*$ into (23),

$$\begin{aligned}
\alpha\langle L^{\frac{1}{2}}(x^* - x^{k+1}), z^{k+1} - z^k\rangle = \tfrac{\alpha}{2}(\|z^* - z^k\|_{\mathrm{F}}^2 - \|z^* - z^{k+1}\|_{\mathrm{F}}^2) - (\tfrac{\alpha}{2} + \tfrac{1}{\epsilon})\|z^{k+1} - z^k\|_{\mathrm{F}}^2 \\
+ \alpha\langle L^{\frac{1}{2}}e^{k+1}, z^{k+1} - z^k\rangle.
\end{aligned} \tag{24}$$

Substituting (24) into (20) and rearranging related terms,

$$\begin{aligned}
&f(x^*) - f(x^{k+1}) + \tfrac{1}{2c}\|e^{k+1}\|_G^2 + \tfrac{1}{2c}(\|x^* - x^{Q,k}\|_G^2 - \|x^* - x^{Q,k+1}\|_G^2) \\
&\quad + \tfrac{\alpha}{2}(\|z^* - z^k\|_{\mathrm{F}}^2 - \|z^* - z^{k+1}\|_{\mathrm{F}}^2) + \langle x^* - x^{k+1}, L^{\frac{1}{2}}\lambda_Q^{k+1}\rangle - \tfrac{1}{c}\langle x^* - x^{k+1}, e^{k+1}\rangle \\
&\geq \tfrac{1}{2c}\|x^{k+1} - x^{Q,k}\|_G^2 - \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2 + (\tfrac{\alpha}{2} + \tfrac{1}{\epsilon})\|z^{k+1} - z^k\|_{\mathrm{F}}^2 \\
&\quad - \alpha\langle L^{\frac{1}{2}}e^{k+1}, z^{k+1} - z^k\rangle.
\end{aligned} \tag{25}$$

Note that

$$\tfrac{1}{2\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 + \langle z^* - z^{k+1}, \tfrac{1}{\epsilon}z^{k+1}\rangle = \tfrac{1}{2\epsilon}\|z^*\|_{\mathrm{F}}^2 - \tfrac{1}{2\epsilon}\|z^{k+1}\|_{\mathrm{F}}^2. \tag{26}$$

Combining (12) with (26),

$$\tfrac{1}{2\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 + \langle z^* - z^{k+1}, \lambda^{k+1}\rangle = \tfrac{1}{2\epsilon}\|z^*\|_{\mathrm{F}}^2 - \tfrac{1}{2\epsilon}\|z^{k+1}\|_{\mathrm{F}}^2. \tag{27}$$

Substituting (27) into (25),

$$\begin{aligned}
&\underbrace{f(x^*) + \tfrac{1}{2\epsilon}\|z^*\|_{\mathrm{F}}^2 - f(x^{k+1}) - \tfrac{1}{2\epsilon}\|z^{k+1}\|_{\mathrm{F}}^2}_{(\mathrm{T}_4)} + \underbrace{\langle x^* - x^{k+1}, L^{\frac{1}{2}}\lambda^{k+1}\rangle - \langle z^* - z^{k+1}, \lambda^{k+1}\rangle}_{(\mathrm{T}_5)} \\
&\quad + \tfrac{1}{2c}(\|x^* - x^{Q,k}\|_G^2 - \|x^* - x^{Q,k+1}\|_G^2) + \tfrac{\alpha}{2}(\|z^* - z^k\|_{\mathrm{F}}^2 - \|z^* - z^{k+1}\|_{\mathrm{F}}^2) \\
&\geq \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 - \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2 + (\tfrac{\alpha}{2} + \tfrac{1}{\epsilon})\|z^{k+1} - z^k\|_{\mathrm{F}}^2 + \tfrac{1}{2\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 \\
&\quad + \tfrac{1}{c}\langle x^* - x^{k+1}, e^{k+1}\rangle + \alpha\langle L^{\frac{1}{2}}e^{k+1}, z^k - z^{k+1}\rangle - \tfrac{1}{2c}\|e^{k+1}\|_G^2.
\end{aligned} \tag{28}$$

We introduce an equation to further simplify (28)

$$\begin{aligned}
0 &= \big\langle\lambda^* - \lambda^{k+1}, \tfrac{\lambda^k - \lambda^{k+1}}{\alpha}\big\rangle - \big\langle\lambda^* - \lambda^{k+1}, \tfrac{\lambda^k - \lambda^{k+1}}{\alpha}\big\rangle \\
&= \big\langle\lambda^* - \lambda^{k+1}, -L^{\frac{1}{2}}x^{Q,k+1} + z^{k+1}\big\rangle - \big\langle\lambda^* - \lambda^{k+1}, \tfrac{\lambda^k - \lambda^{k+1}}{\alpha}\big\rangle \\
&= \underbrace{-\big\langle\lambda^*, L^{\frac{1}{2}}x^{k+1} - z^{k+1}\big\rangle}_{(\mathrm{T}_6)} + \underbrace{\big\langle\lambda^{k+1}, L^{\frac{1}{2}}x^{k+1} - z^{k+1}\big\rangle}_{(\mathrm{T}_7)} \underbrace{-\big\langle\lambda^* - \lambda^{k+1}, \tfrac{\lambda^k - \lambda^{k+1}}{\alpha}\big\rangle}_{(\mathrm{T}_8)} \\
&\quad - \big\langle L^{\frac{1}{2}}(\lambda^* - \lambda^{k+1}), e^{k+1}\big\rangle.
\end{aligned} \tag{29}$$

Combining the term (T_4) in (28) and the term (T_6) in (29), we have

$$\begin{aligned} &f(x^*) + \tfrac{1}{2\epsilon}\|z^*\|_F^2 - f(x^{k+1}) - \tfrac{1}{2\epsilon}\|z^{k+1}\|_F^2 - \langle \lambda^*, L^{\frac{1}{2}} x^{k+1} - z^{k+1}\rangle \\ &= L(x^*, z^*, \lambda^*) - L(x^{k+1}, z^{k+1}, \lambda^*) \le 0. \end{aligned} \tag{30}$$

For the term (T_5) in (28), we combine it with the term (T_7) in (29) and use $L^{\frac{1}{2}} x^* = z^*$ to further simplify it.

$$\begin{aligned} &\langle x^* - x^{k+1}, L^{\frac{1}{2}}\lambda^{k+1}\rangle - \langle z^* - z^{k+1}, \lambda^{k+1}\rangle + \langle \lambda^{k+1}, L^{\frac{1}{2}}x^{k+1} - z^{k+1}\rangle \\ &= \langle x^* - x^{k+1}, L^{\frac{1}{2}}\lambda^{k+1}\rangle + \langle \lambda^{k+1}, L^{\frac{1}{2}}x^{k+1} - z^{k+1} - z^* + z^{k+1}\rangle \\ &= \langle x^* - x^{k+1}, L^{\frac{1}{2}}\lambda^{k+1}\rangle + \langle \lambda^{k+1}, L^{\frac{1}{2}}x^{k+1} - L^{\frac{1}{2}}x^*\rangle = 0. \end{aligned} \tag{31}$$

Adding $\frac{1}{2\alpha}\|\lambda^{k+1} - \lambda^k\|_F^2$ into the term (T_8) in (29), we have

$$\tfrac{1}{2\alpha}\|\lambda^{k+1} - \lambda^k\|_F^2 - \left\langle \lambda^* - \lambda^{k+1}, \tfrac{\lambda^k - \lambda^{k+1}}{\alpha}\right\rangle = \tfrac{1}{2\alpha}\|\lambda^* - \lambda^k\|_F^2 - \tfrac{1}{2\alpha}\|\lambda^* - \lambda^{k+1}\|_F^2. \tag{32}$$

Substituting (30), (31), and (32) into (28) and rearranging it then yields

$$\begin{aligned} &L(x^*, z^*, \lambda^*) - L(x^{k+1}, z^{k+1}, \lambda^*) + \tfrac{1}{2c}\|x^* - x^{Q,k}\|_G^2 - \tfrac{1}{2c}\|x^* - x^{Q,k+1}\|_G^2 \\ &+ \tfrac{\alpha}{2}\|z^* - z^k\|_F^2 - \tfrac{\alpha}{2}\|z^* - z^{k+1}\|_F^2 + \tfrac{1}{2\alpha}\|\lambda^* - \lambda^k\|_F^2 - \tfrac{1}{2\alpha}\|\lambda^* - \lambda^{k+1}\|_F^2 \\ \ge\ & \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 - \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_F^2 + (\tfrac{\alpha}{2} + \tfrac{1}{\epsilon})\|z^{k+1} - z^k\|_F^2 + \tfrac{1}{2\epsilon}\|z^* - z^{k+1}\|_F^2 \\ &+ \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_F^2 + \tfrac{1}{c}\langle x^* - x^{k+1}, e^{k+1}\rangle + \alpha\langle L^{\frac{1}{2}}e^{k+1}, z^k - z^{k+1}\rangle - \tfrac{1}{2c}\|e^{k+1}\|_G^2 \\ &+ \langle L^{\frac{1}{2}}(\lambda^* - \lambda^{k+1}), e^{k+1}\rangle. \end{aligned} \tag{33}$$

Based on $\|u_{[Q]}^k - u^*\|_H^2 = \|x^{Q,k} - x^*\|^2_{\frac{1}{2c}G} + \|z^k - z^*\|^2_{\frac{\alpha}{2}I} + \|\lambda^k - \lambda^*\|^2_{\frac{1}{2\alpha}I}$, we further simplify (33),

$$\begin{aligned} &L(x^*, z^*, \lambda^*) - L(x^{k+1}, z^{k+1}, \lambda^*) + \|u_{[Q]}^k - u^*\|_H^2 - \|u_{[Q]}^{k+1} - u^*\|_H^2 + \tfrac{1}{2c}\|e^{k+1}\|_G^2 \\ \ge\ & \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 - \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_F^2 + (\tfrac{\alpha}{2} + \tfrac{1}{\epsilon})\|z^k - z^{k+1}\|_F^2 + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_F^2 \\ &+ \tfrac{1}{2\epsilon}\|z^* - z^{k+1}\|_F^2 + \tfrac{1}{c}\langle x^* - x^{k+1}, e^{k+1}\rangle + \alpha\langle L^{\frac{1}{2}}(z^k - z^{k+1}), e^{k+1}\rangle \\ &+ \underbrace{\langle \lambda^* - \lambda^{k+1}, L^{\frac{1}{2}}e^{k+1}\rangle}_{(T_9)}. \end{aligned} \tag{34}$$

Because $\frac{1}{\epsilon}z^* = \lambda^*$ and $\frac{1}{\epsilon}z^{k+1} = \lambda^{k+1}$, we substitute them into the term (T_9) in (34),

$$\langle \lambda^* - \lambda^{k+1}, L^{\frac{1}{2}}e^{k+1}\rangle = \tfrac{1}{\epsilon}\langle z^* - z^{k+1}, L^{\frac{1}{2}}e^{k+1}\rangle. \tag{35}$$

According to (34) and (35), we get

$$\begin{aligned} &L(x^*, z^*, \lambda^*) - L(x^{k+1}, z^{k+1}, \lambda^*) + \|u_{[Q]}^k - u^*\|_H^2 - \|u_{[Q]}^{k+1} - u^*\|_H^2 + \tfrac{1}{2c}\|e^{k+1}\|_G^2 \\ &+ \tfrac{1}{c}\langle x^{k+1} - x^*, e^{k+1}\rangle \\ \ge\ & \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 - \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_F^2 + (\tfrac{\alpha}{2} + \tfrac{1}{\epsilon})\|z^k - z^{k+1}\|_F^2 + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_F^2 \\ &+ \tfrac{1}{2\epsilon}\|z^* - z^{k+1}\|_F^2 + \alpha\langle z^k - z^{k+1}, L^{\frac{1}{2}}e^{k+1}\rangle + \tfrac{1}{\epsilon}\langle z^* - z^{k+1}, L^{\frac{1}{2}}e^{k+1}\rangle. \end{aligned} \tag{36}$$

Note that

$$\begin{aligned}\langle z^* - z^{k+1}, L^{\frac{1}{2}} e^{k+1}\rangle \geq & -\tfrac{1}{4}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 - \tfrac{1}{4}\|L^{\frac{1}{2}} e^{k+1}\|_{\mathrm{F}}^2 \\ \geq & -\tfrac{1}{4}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 - \tfrac{\sigma_{\max}^2(L^{\frac{1}{2}})}{4}\|e^{k+1}\|_{\mathrm{F}}^2, \end{aligned} \tag{37}$$

and

$$\begin{aligned}\langle z^k - z^{k+1}, L^{\frac{1}{2}} e^{k+1}\rangle \geq & -\tfrac{1}{2}\|z^k - z^{k+1}\|_{\mathrm{F}}^2 - \tfrac{1}{2}\|L^{\frac{1}{2}} e^{k+1}\|_{\mathrm{F}}^2 \\ \geq & -\tfrac{1}{2}\|z^k - z^{k+1}\|_{\mathrm{F}}^2 - \tfrac{\sigma_{\max}^2(L^{\frac{1}{2}})}{2}\|e^{k+1}\|_{\mathrm{F}}^2, \end{aligned} \tag{38}$$

where $\sigma_{\max}(L^{\frac{1}{2}})$ is the largest singular value of $L^{\frac{1}{2}}$. Next, substituting (37) and (38) into (36), we obtain

$$\begin{aligned} & L(x^*, z^*, \lambda^*) - L(x^{k+1}, z^{k+1}, \lambda^*) + \|u_{[Q]}^k - u^*\|_H^2 - \|u_{[Q]}^{k+1} - u^*\|_H^2 + \tfrac{1}{2c}\|e^{k+1}\|_G^2 \\ & + \tfrac{1}{c}\langle x^{k+1} - x^*, e^{k+1}\rangle \\ \geq & \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 - \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{\epsilon}\|z^k - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_{\mathrm{F}}^2 \\ & + \tfrac{1}{4\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 - \tfrac{(1+2\alpha\epsilon)\sigma_{\max}^2(L^{\frac{1}{2}})}{4\epsilon}\|e^{k+1}\|_{\mathrm{F}}^2. \end{aligned} \tag{39}$$

For the term $\|e^{k+1}\|_G^2$ in (39), we have the following inequality,

$$\|e^{k+1}\|_G^2 \leq \mu_{\max}(G)\|e^{k+1}\|_{\mathrm{F}}^2, \tag{40}$$

where $\mu_{\max}(G)$ is the largest eigenvalue of G.

Combining (39) with (40), we get

$$\begin{aligned} & L(x^*, z^*, \lambda^*) - L(x^{k+1}, z^{k+1}, \lambda^*) + \|u_{[Q]}^k - u^*\|_H^2 - \|u_{[Q]}^{k+1} - u^*\|_H^2 \\ & + \left(\tfrac{\mu_{\max}(G)}{2c} + \tfrac{(1+2\alpha\epsilon)\sigma_{\max}^2(L^{\frac{1}{2}})}{4\epsilon} - \tfrac{1}{c}\right)\|e^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{c}\langle x^{Q,k+1} - x^*, e^{k+1}\rangle \\ \geq & \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 - \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{\epsilon}\|z^k - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_{\mathrm{F}}^2 \\ & + \tfrac{1}{4\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2. \end{aligned} \tag{41}$$

Based on Cauchy–Schwarz inequality, we rewrite the term $\frac{1}{c}\langle x^{Q,k+1} - x^*, e^{k+1}\rangle$ under Assumption 5,

$$\tfrac{1}{c}\langle x^{Q,k+1} - x^*, e^{k+1}\rangle \leq \tfrac{1}{c}\|x^{Q,k+1} - x^*\|_{\mathrm{F}}\|e^{k+1}\|_{\mathrm{F}} \leq \tfrac{A}{c}\|e^{k+1}\|_{\mathrm{F}}. \tag{42}$$

Then substituting (42) into (41),

$$\begin{aligned} & \|u_{[Q]}^k - u^*\|_H^2 - \|u_{[Q]}^{k+1} - u^*\|_H^2 + L(x^*, z^*, \lambda^*) - L(x^{k+1}, z^{k+1}, \lambda^*) + \tfrac{A}{c}\|e^{k+1}\|_{\mathrm{F}} \\ & + \left(\tfrac{\mu_{\max}(G)}{2c} + \tfrac{(1+2\alpha\epsilon)\sigma_{\max}^2(L^{\frac{1}{2}})}{4\epsilon} - \tfrac{1}{c}\right)\|e^{k+1}\|_{\mathrm{F}}^2 \\ \geq & \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 - \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{\epsilon}\|z^k - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_{\mathrm{F}}^2 \\ & + \tfrac{1}{4\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2. \end{aligned} \tag{43}$$

For the error term in (43), based on $e^{k+1} \leq \frac{\Delta}{2} = \frac{l}{2^{b+1}}$, we have

$$\begin{aligned}
&\Big[\frac{\mu_{\max}(G)}{2c} + \frac{(1+2\alpha\epsilon)\sigma_{\max}^2(L^{\frac{1}{2}})}{4\epsilon} - \frac{1}{c}\Big]\Big\|e^{k+1}\Big\|_{\mathrm{F}}^2 + \frac{A}{c}\Big\|e^{k+1}\Big\|_{\mathrm{F}} \\
&\leq \Big[\frac{\mu_{\max}(G)}{2c} + \frac{(1+2\alpha\epsilon)\sigma_{\max}^2(L^{\frac{1}{2}})}{4\epsilon} - \frac{1}{c}\Big] np\Big(\frac{l}{2^{b+1}}\Big)^2 + \frac{A}{c}\sqrt{np}\Big(\frac{l}{2^{b+1}}\Big) \\
&= \Big[\frac{\mu_{\max}(G)}{2c} + \frac{(1+2\alpha\epsilon)\sigma_{\max}^2(L^{\frac{1}{2}})}{4\epsilon} - \frac{1}{c}\Big] np\Big(\frac{C\theta^{k+1}}{2^{b+1}}\Big)^2 + \frac{A}{c}\sqrt{np}\Big(\frac{C\theta^{k+1}}{2^{b+1}}\Big) \\
&= \Big[\frac{\mu_{\max}(G)}{2c} + \frac{(1+2\alpha\epsilon)\sigma_{\max}^2(L^{\frac{1}{2}})}{4\epsilon} - \frac{1}{c}\Big] np\frac{C^2\theta^{2(k+1)}}{2^{2(b+1)}} + \frac{A}{c}\sqrt{np}\frac{C\theta^{k+1}}{2^{b+1}}. \qquad (44)
\end{aligned}$$

Because the eigenvalues of the matrix G lie in $(1-2\alpha c, 1]$ and the singular values of $L^{\frac{1}{2}}$ lie in $(0, \sqrt{2})$, we further simplify (44)

$$\begin{aligned}
&\Big[\frac{\mu_{\max}(G)}{2c} + \frac{(1+2\alpha\epsilon)\sigma_{\max}^2(L^{\frac{1}{2}})}{4\epsilon} - \frac{1}{c}\Big] np\frac{C^2\theta^{2(k+1)}}{2^{2(b+1)}} + \frac{A}{c}\sqrt{np}\frac{C\theta^{k+1}}{2^{b+1}} \\
&< \Big(\frac{1+2\alpha\epsilon}{2\epsilon} - \frac{1}{2c}\Big) np\frac{C^2\theta^{2(k+1)}}{2^{2(b+1)}} + \frac{A}{c}\sqrt{np}\frac{C\theta^{k+1}}{2^{b+1}} \\
&< \frac{1+2\alpha\epsilon}{2\epsilon}\frac{npC^2}{2^{2(b+1)}}\theta^{2(k+1)} + \frac{\sqrt{np}AC}{2^{b+1}c}\theta^{k+1} \triangleq \Lambda. \qquad (45)
\end{aligned}$$

Substituting (45) into (43),

$$\begin{aligned}
&L(x^*, z^*, \lambda^*) - L(x^{k+1}, z^{k+1}, \lambda^*) + \|u_{[Q]}^k - u^*\|_H^2 - \|u_{[Q]}^{k+1} - u^*\|_H^2 + \Lambda \\
&\geq \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 - \tfrac{L_f}{2}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{\epsilon}\|z^k - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_{\mathrm{F}}^2 \\
&\quad + \tfrac{1}{4\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2. \qquad (46)
\end{aligned}$$

Choosing $\frac{1}{2c}G - \frac{L_f}{2}I \succ 0$, we can find $\tau > 0$ such that $\frac{1}{2c}G - \frac{L_f}{2}I \succ \tau I$, and it holds for all $k \geq 1$ that

$$\begin{aligned}
&\|u_{[Q]}^k - u^*\|_H^2 - \|u_{[Q]}^{k+1} - u^*\|_H^2 + \Lambda \\
&\geq \tau\|x^{Q,k} - x^{k+1}\|_G^2 + \tfrac{1}{\epsilon}\|z^k - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_{\mathrm{F}}^2, \qquad (47)
\end{aligned}$$

which completes the proof. □

6.2. Proof of Theorem 3.2

Based on Theorem 3.1, the terms of the right side of (13) all tend to 0, which implies that

$$\begin{cases} x^{Q,k} - x^{k+1} \to 0, \\ z^k - z^{k+1} \to 0, \\ \lambda^k - \lambda^{k+1} \to 0. \end{cases} \qquad (48)$$

Further, we obtain

$$L^{\frac{1}{2}}x^{Q,k} - z^k \to 0. \qquad (49)$$

Because the sequences $\{u_{[Q]}^k\}$ lies in a compact region, we set a subsequence $\{u_{[Q]}^{k_j}\}$ of $\{u_{[Q]}^k\}$ which convergences to $\breve{u} = (\breve{x}, \breve{z}, \breve{\lambda})$, where $\breve{u}$ is a limit point of $\{u_{[Q]}^k\}$ and $L^{\frac{1}{2}}\breve{x} - \breve{z} = 0$. And $\|u_{[Q]}^k - u^*\|_H^2$ is monotonically non-increasing and thus converges.

Note that (10a) implies that

$$\nabla f(\breve{x}) + L^{\frac{1}{2}}\breve{\lambda} = 0. \tag{50}$$

Note also that (12) implies that

$$\tfrac{1}{\epsilon}\breve{z} = \breve{\lambda}. \tag{51}$$

Next, taking $k \to +\infty$ in (10) yields the KKT conditions of (6),

$$\begin{aligned} &\nabla f(x^*) + L^{\frac{1}{2}}\lambda^* = 0,\\ &\tfrac{1}{\epsilon}z^* = \lambda^*,\\ &L^{\frac{1}{2}}x^* = z^*,\\ &x^* \in \mathbb{K}. \end{aligned} \tag{52}$$

Because $\breve{x} \in \mathbb{K}$ and $L^{\frac{1}{2}}\breve{x} = \breve{z}$, thus $(\breve{x}, \breve{z}, \breve{\lambda})$ satisfies the KKT conditions and is an optimal solution of (6). Therefore, any limit point of $\{x^{Q,k}, z^k, \lambda^k\}$ is an optimal solution.

After completing the above proof, we still need to show that $\{x^{Q,k}, z^k, \lambda^k\}$ has a unique limit point. We set $(\breve{x}_1, \breve{z}_1, \breve{\lambda}_1)$ and $(\breve{x}_2, \breve{z}_2, \breve{\lambda}_2)$ are any two limit points of $\{x^{Q,k}, z^k, \lambda^k\}$. In other words, they are optimal solutions of (6).

Replacing u^* in (47) with $\breve{u}_1 \triangleq (\breve{x}_1, \breve{z}_1, \breve{\lambda}_1)$ and $\breve{u}_2 \triangleq (\breve{x}_2, \breve{z}_2, \breve{\lambda}_2)$, we obtain

$$\|u_{[Q]}^{k+1} - \breve{u}_i\|_H^2 \le \|u_{[Q]}^k - \breve{u}_i\|_H^2 + \Lambda, \quad i = 1, 2. \tag{53}$$

For the right-hand side of (53),

$$\lim_{k\to+\infty}(\|u_{[Q]}^k - \breve{u}_i\|_H + \Lambda) = \lim_{k\to+\infty}\|u_{[Q]}^k - \breve{u}_i\|_H + \lim_{k\to+\infty}\Lambda, \quad i = 1, 2.$$

Because of $\theta \in (0, 1)$,

$$\lim_{k\to+\infty}\Lambda = 0.$$

Therefore,

$$\lim_{k\to+\infty}(\|u_{[Q]}^k - \breve{u}_i\|_H + \Lambda) = \lim_{k\to+\infty}\|u_{[Q]}^k - \breve{u}_i\|_H \triangleq \rho_i < +\infty, \quad i = 1, 2. \tag{54}$$

With the identity

$$\|u_{[Q]}^k - \breve{u}_1\|_H^2 - \|u_{[Q]}^k - \breve{u}_2\|_H^2 = \|\breve{u}_1\|_H^2 - \|\breve{u}_2\|_H^2 - 2\langle u_{[Q]}^k, \breve{u}_1 - \breve{u}_2\rangle_H. \tag{55}$$

Calculating the limits of (55) then yields

$$\rho_1^2 - \rho_2^2 = -2\langle\breve{u}_1, \breve{u}_1 - \breve{u}_2\rangle_H + \|\breve{u}_1\|_H^2 - \|\breve{u}_2\|_H^2 = -\|\breve{u}_1 - \breve{u}_2\|_H^2 \tag{56}$$

and

$$\rho_1^2 - \rho_2^2 = -2\langle \breve{u}_2, \breve{u}_1 - \breve{u}_2\rangle_H + \|\breve{u}_1\|_H^2 - \|\breve{u}_2\|_H^2 = \|\breve{u}_1 - \breve{u}_2\|_H^2. \tag{57}$$

Obviously, to want (56) = (57) if and only if $\|\breve{u}_1 - \breve{u}_2\|_H^2 = 0$. Thus the limit point of $\{x^{Q,k}, z^k, \lambda^k\}$ is unique. This completes the proof. □

6.3. Proof of Theorem 3.3

Based on the definition of the convex function, we have

$$\psi\Big(\frac{\nu^1 + \cdots + \nu^k}{k}\Big) \le \frac{\psi(\nu^1) + \cdots + \psi(\nu^k)}{k}. \tag{58}$$

Then,

$$\begin{aligned} &\psi(\breve{\nu}^k) - \psi(\nu^*) + \langle \lambda, L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\rangle \\ \le\; &\frac{\psi(\nu^1) + \cdots + \psi(\nu^k)}{k} - \psi(\nu^*) + \langle \lambda, L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\rangle \\ =\; &\frac{1}{k}\sum_{t=0}^{k-1}\Big[\psi(\nu^{t+1}) - \psi(\nu^*) + \Big\langle \lambda, L^{\frac{1}{2}}x^{t+1} - z^{k+1}\Big\rangle\Big],\ \forall\lambda. \end{aligned} \tag{59}$$

Rewriting (46) then yields

$$\begin{aligned} &\tfrac{1}{2c}\|x^* - x^{Q,k}\|_G^2 - \tfrac{1}{2c}\|x^* - x^{Q,k+1}\|_G^2 + \tfrac{\alpha}{2}\|z^* - z^k\|_{\mathrm{F}}^2 - \tfrac{\alpha}{2}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 \\ &+ \tfrac{1}{2\alpha}\|\lambda - \lambda^k\|_{\mathrm{F}}^2 - \tfrac{1}{2\alpha}\|\lambda - \lambda^{k+1}\|_{\mathrm{F}}^2 + \Lambda \\ \ge\; & L(x^{k+1}, z^{k+1}, \lambda) - L(x^*, z^*, \lambda) = \psi(\nu^{k+1}) - \psi(\nu^*) + \Big\langle \lambda, L^{\frac{1}{2}}x^{k+1} - z^{k+1}\Big\rangle,\ \forall\lambda. \end{aligned} \tag{60}$$

Combining (59) and (60), we obtain

$$\begin{aligned} &\psi(\breve{\nu}^k) - \psi(\nu^*) + \langle \lambda, L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\rangle \\ \le\; &\frac{1}{k}\sum_{t=0}^{k-1}\Big[\tfrac{1}{2c}\|x^* - x^{Q,k}\|_G^2 - \tfrac{1}{2c}\|x^* - x^{Q,k+1}\|_G^2 + \tfrac{\alpha}{2}\|z^* - z^k\|_{\mathrm{F}}^2 - \tfrac{\alpha}{2}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 \\ &+ \tfrac{1}{2\alpha}\|\lambda - \lambda^k\|_{\mathrm{F}}^2 - \tfrac{1}{2\alpha}\|\lambda - \lambda^{k+1}\|_{\mathrm{F}}^2 + \Lambda\Big] \\ =\; &\frac{1}{k}\Big[\tfrac{1}{2c}\|x^* - x^{Q,0}\|_G^2 - \tfrac{1}{2c}\|x^* - x^{Q,k}\|_G^2 + \tfrac{\alpha}{2}\|z^* - z^0\|_{\mathrm{F}}^2 - \tfrac{\alpha}{2}\|z^* - z^k\|_{\mathrm{F}}^2 \\ &+ \tfrac{1}{2\alpha}\|\lambda - \lambda^0\|_{\mathrm{F}}^2 - \tfrac{1}{2\alpha}\|\lambda - \lambda^k\|_{\mathrm{F}}^2\Big] + \frac{1}{k}\sum_{t=0}^{k-1}\Lambda \\ =\; &\frac{1}{k}\Big[\tfrac{1}{2c}\|x^* - x^{Q,0}\|_G^2 + \tfrac{\alpha}{2}\|z^* - z^0\|_{\mathrm{F}}^2 + \tfrac{1}{2\alpha}\|\lambda - \lambda^0\|_{\mathrm{F}}^2\Big] + \frac{1}{k}\sum_{t=0}^{k-1}\Lambda,\ \forall\lambda. \end{aligned} \tag{61}$$

For the term $\frac{1}{k}\sum_{t=0}^{k-1}\Lambda$ in (61), we have

$$\begin{aligned}\frac{1}{k}\sum_{t=0}^{k-1}\Lambda &= \frac{1}{k}\sum_{t=0}^{k-1}\Big[\frac{1+2\alpha\epsilon}{2\epsilon}\frac{npC^2}{2^{2(b+1)}}\theta^{2(k+1)} + \frac{\sqrt{np}AC}{2^{b+1}c}\theta^{k+1}\Big]\\ &= \frac{1}{k}\Big(\frac{1+2\alpha\epsilon}{2\epsilon}\frac{npC^2}{2^{2(b+1)}}\Big)\sum_{t=0}^{k-1}\theta^{2(k+1)} + \frac{1}{k}\Big(\frac{\sqrt{np}AC}{2^{b+1}c}\Big)\sum_{t=0}^{k-1}\theta^{k+1}\\ &\le \frac{1}{k}\Big(\frac{1+2\alpha\epsilon}{2\epsilon}\frac{npC^2}{2^{2(b+1)}}\Big)\frac{\theta^2}{1-\theta^2} + \frac{1}{k}\Big(\frac{\sqrt{np}AC}{2^{b+1}c}\Big)\frac{\theta}{1-\theta} \triangleq \frac{1}{k}\Lambda_1 + \frac{1}{k}\Lambda_2. \qquad (62)\end{aligned}$$

Based on (62), (61) is modified to

$$\begin{aligned}&\psi(\breve{\nu}^k) - \psi(\nu^*) + \langle\lambda, L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\rangle\\ &\le \frac{1}{k}\Big[\tfrac{1}{2c}\|x^* - x^{Q,0}\|_G^2 + \tfrac{\alpha}{2}\|z^* - z^0\|_{\mathrm{F}}^2 + \tfrac{1}{2\alpha}\|\lambda - \lambda^0\|_{\mathrm{F}}^2\Big] + \frac{1}{k}\Lambda_1 + \frac{1}{k}\Lambda_2,\ \forall\lambda. \qquad (63)\end{aligned}$$

Setting

$$\lambda = \frac{2\|\lambda^*\|_{\mathrm{F}}\Big(L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\Big)}{\|L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\|_{\mathrm{F}}}, \qquad (64)$$

and substituting it into the left-hand side of (63), we get

$$\begin{aligned}&\psi(\breve{\nu}^k) - \psi(\nu^*) + \langle\lambda, L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\rangle\\ &= \psi(\breve{\nu}^k) - \psi(\nu^*) + 2\|\lambda^*\|_{\mathrm{F}}\|L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\|_{\mathrm{F}}. \qquad (65)\end{aligned}$$

For the right hand side of (63), let $\|\lambda\|_{\mathrm{F}} = 2\|\lambda^*\|_{\mathrm{F}}$ and we use the triangle inequality to rewrite the term $\|\lambda - \lambda^0\|_{\mathrm{F}}^2$

$$\|\lambda - \lambda^0\|_{\mathrm{F}}^2 \le 8\|\lambda^*\|_{\mathrm{F}}^2 + 2\|\lambda^0\|_{\mathrm{F}}^2. \qquad (66)$$

Based on (63), (65), and (66), we get

$$\begin{aligned}\psi(\breve{\nu}^k) - \psi(\nu^*) &\le \frac{1}{k}\Big[\tfrac{1}{2c}\|x^* - x^{Q,0}\|_G^2 + \tfrac{\alpha}{2}\|z^* - z^0\|_{\mathrm{F}}^2 + \tfrac{4}{\alpha}\|\lambda^*\|_{\mathrm{F}}^2 - \tfrac{1}{\alpha}\|\lambda^0\|_{\mathrm{F}}^2 + \Lambda_1 + \Lambda_2\Big]\\ &\quad - 2\|\lambda^*\|_{\mathrm{F}}\|L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\|_{\mathrm{F}}\\ &\triangleq \frac{Q_1}{k} - 2\|\lambda^*\|_{\mathrm{F}}\|L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\|_{\mathrm{F}}. \qquad (67)\end{aligned}$$

For all ν, $L(\nu, \lambda^*) \ge L(\nu^*, \lambda^*)$, thus

$$\psi(\breve{\nu}^k) - \psi(\nu^*) \ge -\|\lambda^*\|_{\mathrm{F}}\|L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\|_{\mathrm{F}}. \qquad (68)$$

Combining (67) and (68), we obtain

$$\max\Big\{|\psi(\breve{\nu}^k) - \psi(\nu^*)|, \|\lambda^*\|_{\mathrm{F}}\|L^{\frac{1}{2}}\breve{x}^k - \breve{z}^k\|_{\mathrm{F}}\Big\} \le \frac{Q_1}{k}. \qquad (69)$$

Thus we completed the proof. □

6.4. Proof of Theorem 3.4

Under Assumption 4, we rewrite (19) in the proof of Theorem 3.1 as follows,

$$\begin{aligned} & \langle x^* - x^{k+1}, \nabla f(x^{Q,k}) \rangle \\ = & \ \langle x^* - x^{k+1}, \nabla f(x^{k+1}) \rangle + \langle x^* - x^{k+1}, \nabla f(x^{Q,k}) - \nabla f(x^{k+1}) \rangle \\ \leq & \ f(x^*) - f(x^{k+1}) - \frac{\mu_f}{2} \|x^* - x^{k+1}\|_F^2 + \langle x^* - x^{k+1}, \nabla f(x^{Q,k}) - \nabla f(x^{k+1}) \rangle. \end{aligned} \quad (70)$$

Then the following operations are the same as that of Theorem 3.1. Furthermore (46) becomes

$$\begin{aligned} & \tfrac{1}{2c}\|x^* - x^{Q,k}\|_G^2 - \tfrac{1}{2c}\|x^* - x^{Q,k+1}\|_G^2 + \tfrac{\alpha}{2}\|z^* - z^k\|_{\mathrm{F}}^2 - \tfrac{\alpha}{2}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 \\ & + \tfrac{1}{2\alpha}\|\lambda^* - \lambda^k\|_{\mathrm{F}}^2 - \tfrac{1}{2\alpha}\|\lambda^* - \lambda^{k+1}\|_{\mathrm{F}}^2 + \Lambda \\ \geq & \ \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 + \tfrac{1}{\epsilon}\|z^k - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{4\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_{\mathrm{F}}^2 \\ & + \tfrac{\mu_f}{2}\|x^* - x^{k+1}\|_{\mathrm{F}}^2 - \underbrace{\langle x^* - x^{k+1}, \nabla f(x^{Q,k}) - \nabla f(x^{k+1}) \rangle}_{(\mathrm{T}_{10})}. \end{aligned} \quad (71)$$

For the term (T_{10}) in (71), we have

$$-\langle x^* - x^{k+1}, \nabla f(x^{Q,k}) - \nabla f(x^{k+1}) \rangle \geq -\tfrac{\tau_1}{2}\|x^* - x^{k+1}\|_{\mathrm{F}}^2 - \tfrac{L_f^2}{2\tau_1}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2, \quad (72)$$

where $\tau_1 > 0$ is any constant. Substituting (72) into (71), we get

$$\begin{aligned} & \tfrac{1}{2c}\|x^* - x^{Q,k}\|_G^2 - \tfrac{1}{2c}\|x^* - x^{Q,k+1}\|_G^2 + \tfrac{\alpha}{2}\|z^* - z^k\|_{\mathrm{F}}^2 - \tfrac{\alpha}{2}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 \\ & + \tfrac{1}{2\alpha}\|\lambda^* - \lambda^k\|_{\mathrm{F}}^2 - \tfrac{1}{2\alpha}\|\lambda^* - \lambda^{k+1}\|_{\mathrm{F}}^2 + \Lambda \\ \geq & \ \tfrac{1}{2c}\|x^{Q,k} - x^{k+1}\|_G^2 + \tfrac{1}{\epsilon}\|z^k - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{4\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{2\alpha}\|\lambda^k - \lambda^{k+1}\|_{\mathrm{F}}^2 \\ & + \tfrac{\mu_f - \tau_1}{2}\|x^* - x^{k+1}\|_{\mathrm{F}}^2 - \tfrac{L_f^2}{2\tau_1}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2. \end{aligned} \quad (73)$$

We utilize $\frac{1}{\epsilon} z^k = \lambda^k$ to further simplify (73),

$$\begin{aligned} & \tfrac{1}{c}(\|x^* - x^{Q,k}\|_G^2 - \|x^* - x^{Q,k+1}\|_G^2) + (\alpha + \tfrac{1}{\alpha\epsilon^2})(\|z^* - z^k\|_{\mathrm{F}}^2 - \|z^* - z^{k+1}\|_{\mathrm{F}}^2) + 2\Lambda \\ \geq & \ \tfrac{1}{c}\|x^{Q,k} - x^{k+1}\|_G^2 - \tfrac{L_f^2}{\tau_1}\|x^{Q,k} - x^{k+1}\|_{\mathrm{F}}^2 + (\tfrac{2}{\epsilon} + \tfrac{1}{\alpha\epsilon^2})\|z^k - z^{k+1}\|_{\mathrm{F}}^2 + \tfrac{1}{2\epsilon}\|z^* - z^{k+1}\|_{\mathrm{F}}^2 \\ & + (\mu_f - \tau_1)\|x^* - x^{k+1}\|_{\mathrm{F}}^2. \end{aligned} \quad (74)$$

For the term $(\tau_1 - \mu_f)\|x^* - x^{k+1}\|_{\mathrm{F}}^2$ in (74), we deal with it as follows.

$$\begin{aligned} & (\tau_1 - \mu_f)\|x^* - x^{k+1}\|_{\mathrm{F}}^2 \\ = & \ (\tau_1 - \mu_f)\|x^* - x^{k+1} - e^{k+1} + e^{k+1}\|_{\mathrm{F}}^2 \\ = & \ (\tau_1 - \mu_f)(\|x^* - x^{Q,k+1}\|_{\mathrm{F}}^2 + 2\langle x^* - x^{Q,k+1}, e^{k+1} \rangle + \|e^{k+1}\|_{\mathrm{F}}^2) \\ \leq & \ (\tau_1 - \mu_f)(\|x^* - x^{Q,k+1}\|_{\mathrm{F}}^2 + 2\|x^* - x^{Q,k+1}\|_F \|e^{k+1}\|_{\mathrm{F}} + \|e^{k+1}\|_{\mathrm{F}}^2) \\ \leq & \ (\tau_1 - \mu_f)(\|x^* - x^{Q,k+1}\|_{\mathrm{F}}^2 + 2A\|e^{k+1}\|_{\mathrm{F}} + \|e^{k+1}\|_{\mathrm{F}}^2), \end{aligned} \quad (75)$$

where $2A\|e^{k+1}\|_{\mathrm{F}}$ in the second inequality follows from Assumption 5.

Combining (75) with (74) yields

$$\begin{aligned}
&\tfrac{1}{c}(\|x^*-x^{Q,k}\|_G^2-\|x^*-x^{Q,k+1}\|_G^2)+(\alpha+\tfrac{1}{\alpha\epsilon^2})(\|z^*-z^k\|_{\mathrm{F}}^2-\|z^*-z^{k+1}\|_{\mathrm{F}}^2)+2\Lambda\\
\geq\ &\tfrac{1}{c}\|x^{Q,k}-x^{k+1}\|_G^2-\tfrac{L_f^2}{\tau_1}\|x^{Q,k}-x^{k+1}\|_{\mathrm{F}}^2+(\tfrac{2}{\epsilon}+\tfrac{1}{\alpha\epsilon^2})\|z^k-z^{k+1}\|_{\mathrm{F}}^2+\tfrac{1}{2\epsilon}\|z^*-z^{k+1}\|_{\mathrm{F}}^2\\
&+(\mu_f-\tau_1)\|x^*-x^{Q,k+1}\|_{\mathrm{F}}^2+(\mu_f-\tau_1)(2A\|e^{k+1}\|_{\mathrm{F}}+\|e^{k+1}\|_{\mathrm{F}}^2)\\
\geq\ &(\tfrac{\mu_{\min}(G)}{c}-\tfrac{L_f^2}{\tau_1})\|x^{Q,k}-x^{k+1}\|_{\mathrm{F}}^2+(\tfrac{2}{\epsilon}+\tfrac{1}{\alpha\epsilon^2})\|z^k-z^{k+1}\|_{\mathrm{F}}^2+\tfrac{1}{2\epsilon}\|z^*-z^{k+1}\|_{\mathrm{F}}^2\\
&+(\mu_f-\tau_1)\|x^*-x^{Q,k+1}\|_{\mathrm{F}}^2+(\mu_f-\tau_1)(2A\|e^{k+1}\|_{\mathrm{F}}+\|e^{k+1}\|_{\mathrm{F}}^2).
\end{aligned}\tag{76}$$

To complete the proof, we need $\frac{\mu_{\min}(G)}{c}-\frac{L_f^2}{\tau_1}$ and $\mu_f-\tau_1$ to satisfy the following conditions,

$$\begin{cases}\frac{\mu_{\min}(G)}{c}-\frac{L_f^2}{\tau_1}>0,\\ \mu_f-\tau_1>0.\end{cases}\tag{77}$$

Therefore, let $\delta>0$ to make (77) true, where we define δ as

$$\delta\triangleq 1-\frac{L_f^2c}{\mu_{\min}(G)\mu_f}>0.\tag{78}$$

Choosing $\tau_1=\frac{\mu_f}{1+\delta}$ and substituting it and (78) into (76) then yields

$$\begin{aligned}
&\tfrac{1}{c}(\|x^*-x^{Q,k}\|_G^2-\|x^*-x^{Q,k+1}\|_G^2)+(\alpha+\tfrac{1}{\alpha\epsilon^2})(\|z^*-z^k\|_{\mathrm{F}}^2-\|z^*-z^{k+1}\|_{\mathrm{F}}^2)+2\Lambda\\
\geq\ &\tfrac{\mu_{\min}(G)}{c}\delta^2\|x^{Q,k}-x^{k+1}\|_{\mathrm{F}}^2+(\tfrac{2}{\epsilon}+\tfrac{1}{\alpha\epsilon^2})\|z^k-z^{k+1}\|_{\mathrm{F}}^2+\tfrac{1}{2\epsilon}\|z^*-z^{k+1}\|_{\mathrm{F}}^2\\
&+\tfrac{\mu_f}{1+\delta}\|x^*-x^{Q,k+1}\|_{\mathrm{F}}^2+\tfrac{\mu_f}{1+\delta}(2A\|e^{k+1}\|_{\mathrm{F}}+\|e^{k+1}\|_{\mathrm{F}}^2)\\
\geq\ &\tfrac{\mu_f}{1+\delta}\|x^*-x^{Q,k+1}\|_{\mathrm{F}}^2+\tfrac{1}{2\epsilon}\|z^*-z^{k+1}\|_{\mathrm{F}}^2.
\end{aligned}\tag{79}$$

We can find a parameter $c_1>0$ such that

$$\tfrac{\mu_f}{1+\delta}\|x^*-x^{Q,k+1}\|_{\mathrm{F}}^2\geq c_1\|x^*-x^{Q,k+1}\|_G^2,\tag{80}$$

as long as c_1 is sufficiently small. In this paper, let $c_1=\frac{\mu_f\delta}{(1+\delta)\mu_{\max}(G)}$. Substituting c_1 into (79) and rearranging it, we obtain

$$\begin{aligned}
&\tfrac{1}{c}\|x^*-x^{Q,k}\|_G^2+(\alpha+\tfrac{1}{\alpha\epsilon^2})\|z^*-z^k\|_{\mathrm{F}}^2+2\Lambda\\
\geq\ &(c_1+\tfrac{1}{c})\|x^*-x^{Q,k+1}\|_G^2+(\alpha+\tfrac{1}{2\epsilon}+\tfrac{1}{\alpha\epsilon^2})\|z^*-z^{k+1}\|_{\mathrm{F}}^2.
\end{aligned}\tag{81}$$

Rearranging (81), then it becomes

$$\begin{aligned}
&c_2\|x^*-x^{Q,k+1}\|_G^2+\|z^*-z^{k+1}\|_{\mathrm{F}}^2\\
\leq\ &\tau_2(c_2\|x^*-x^{Q,k}\|_G^2+\|z^*-z^k\|_{\mathrm{F}}^2)+\tau_2\frac{1}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}2\Lambda,
\end{aligned}\tag{82}$$

where $\tau_2 \triangleq \max\{\frac{1}{c_1 c+1}, \frac{2\alpha^2\epsilon^2+2}{\alpha\epsilon+2\alpha^2\epsilon^2+2}\} < 1$ and $c_2 = \frac{c_1+\frac{1}{c}}{\frac{1}{2\epsilon}+\alpha+\frac{1}{\alpha\epsilon^2}}$. Based on (45), we rearrange 2Λ.

$$\begin{aligned} 2\Lambda &= \Big(\frac{1+2\alpha\epsilon}{\epsilon}\Big)np\frac{C^2\theta^{2(k+1)}}{2^{2(b+1)}} + \frac{A}{c}\sqrt{np}\frac{C\theta^{k+1}}{2} \\ &= \underbrace{\Big(\frac{1+2\alpha\epsilon}{\epsilon}\Big)\frac{npC^2\theta^2}{2^{2(b+1)}}}_{\triangleq a_1}\theta^{2k} + \underbrace{\frac{\sqrt{np}AC\theta}{c2^b}}_{\triangleq a_2}\theta^k. \end{aligned} \tag{83}$$

.

Substituting (83) into (82), we get

$$\begin{aligned} &c_2\|x^* - x^{Q,k+1}\|_G^2 + \|z^* - z^{k+1}\|_{\mathrm{F}}^2 \\ &\le \tau_2(c_2\|x^* - x^{Q,k}\|_G^2 + \|z^* - z^k\|_{\mathrm{F}}^2) + \tau_2\tfrac{a_1}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\theta^{2k} + \tau_2\tfrac{a_2}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\theta^k \\ &\le \tau_2^2(c_2\|x^* - x^{Q,k-1}\|_G^2 + \|z^* - z^{k-1}\|_{\mathrm{F}}^2) + \tau_2^2\tfrac{a_1}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\theta^{2(k-1)} + \tau_2\tfrac{a_1}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\theta^{2k} \\ &\quad + \tau_2^2\tfrac{a_2}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\theta^{k-1} + \tau_2\tfrac{a_2}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\theta^k \\ &\le \tau_2^{k+1}\Big[(c_2\|x^* - x^{Q,0}\|_G^2 + \|z^* - z^0\|_{\mathrm{F}}^2) + \tfrac{a_1}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\sum_{j=0}^{k}\Big(\frac{\theta^2}{\tau_2}\Big)^j + \tfrac{a_2}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\sum_{j=0}^{k}\Big(\frac{\theta}{\tau_2}\Big)^j\Big] \\ &\le \tau_2^{k+1}(c_2\|x^* - x^{Q,0}\|_G^2 + \|z^* - z^0\|_{\mathrm{F}}^2 + \Lambda_3 + \Lambda_4), \end{aligned} \tag{84}$$

where $0 < \theta < \tau_2 < 1$, $\Lambda_3 \triangleq \frac{a_1}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\frac{1}{1-\frac{\theta^2}{\tau_2}}$, and $\Lambda_4 \triangleq \frac{a_2}{\alpha+\frac{1}{2\epsilon}+\frac{1}{\alpha\epsilon^2}}\frac{1}{1-\frac{\theta}{\tau_2}}$.

The proof is completed. □

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China under Grants (62103003, 61903027, and 61973002), and in part by the Anhui Provincial Natural Science Foundation (2008085J32).

(Received September 23, 2022)

REFERENCES

[1] S. A. Alghunaim, E. K. Ryu, K. Yuan, and A. H. Sayed: Decentralized proximal gradient algorithms with linear convergence rates. IEEE Trans. Automat. Control *66* (2020), 6, 2787–2794. DOI:10.1109/TAC.2020.3009363

[2] S. Boyd, D. Persi, and L. Xiao: Fastest mixing Markov chain on a graph. SIAM Rev. *46* (2004), 4, 667–689. DOI:10.1137/S0036144503423264

[3] Z. Chen and S. Liang: Distributed aggregative optimization with quantized communication. Kybernetika *58* (2022), 1, 123–144. DOI:10.14736/kyb-2022-1-0123

[4] Z. Chen, J. Ma, S. Liang, and L. Li: Distributed Nash equilibrium seeking under quantization communication. Automatica *141* (2022), 110318. DOI:10.1016/j.automatica.2022.110318

[5] S. Cheng, S. Liang, Y. Fan, and Y. Hong: Distributed gradient tracking for unbalanced optimization with different constraint sets. IEEE Trans. Automat. Control (2022). DOI:10.1109/TAC.2022.3192316

[6] T. Dorina, K. Effrosyni, Y. Pu, and F. Pascal: Distributed average consensus with quantization refinement. IEEE Trans. Signal Process. *61* (2013), 1, 194–205. DOI:10.1109/TSP.2012.2223692

[7] L. Jian, J. Hu, J. Wang, and K. Shi: Distributed inexact dual consensus ADMM for network resource allocation. Optimal Control Appl. Methods *40* (2019), 6, 1071–1087. DOI:10.1002/oca.2538

[8] J. Lei, H. Chen, and H. Fang: Primal–dual algorithm for distributed constrained optimization. Systems Control Lett. *96* (2016), 110–117. DOI:10.1016/j.sysconle.2016.07.009

[9] J. Lei, P. Yi, G. Shi, and D. O. A. Brian: Distributed algorithms with finite data rates that solve linear equations. SIAM J. Optim. *30* (2020), 2, 1191–1222. DOI:10.1137/19M1258864

[10] X. Li, G. Feng, and L. Xie: Distributed proximal algorithms for multi-agent optimization with coupled inequality constraints. IEEE Trans. Automat. Control *66* (2021), 3, 1223–1230. DOI:10.1109/TAC.2020.2989282

[11] X. Li, F. Gang, and X. Lihua: Distributed proximal point algorithm for constrained optimization over unbalanced graphs. 2019 IEEE 15th International Conference on Control and Automation (ICCA), IEEE, (2019), 824-829. DOI:10.1109/ICCA.2019.8899938

[12] P. Li, J. Hu, L. Qiu, Y. Zhao, and K. G. Bijoy: A distributed economic dispatch strategy for power-water networks. IEEE Trans. Control Network Systems *9* (2022), 1, 356–366. DOI:10.1109/TCNS.2021.3104103

[13] W. Li, X. Zeng, S. Liang, and Y. Hong: Exponentially convergent algorithm design for constrained distributed optimization via nonsmooth approach. IEEE Trans. Automat. Control *67* (2022), 2, 934–940. DOI:10.1109/TAC.2021.3075666

[14] S. Liang, L. Wang, and Y. George: Exponential convergence of distributed primal–dual convex optimization algorithm without strong convexity. Automatica *105* (2019), 298–306. DOI:10.1016/j.automatica.2019.04.004

[15] Y. Liu, G. Wu, Z. Tian, and Q. Ling: DQC-ADMM: decentralized dynamic ADMM with quantized and censored communications. IEEE Trans. Neural Networks Learn. Systems *33* (2022), 8, 3290–3304. DOI:10.1109/TNNLS.2021.3051638

[16] S. Ma: Alternating proximal gradient method for convex minimization. J. Scientific Computing *68* (2016), 2, 546–572. DOI:10.1007/s10915-015-0150-0

[17] X. Ma, P. Yi, and J. Chen: Distributed gradient tracking methods with finite data rates. J. Systems Science Complexity *34* (2021), 5, 1927–1952. DOI:10.1007/s11424-021-1231-9

[18] S. U. Pillai, S. Torsten, and Ch. Seunghun: The Perron-Frobenius theorem: some of its applications. IEEE Signal Process. Magazine *22* (2005), 2, 62–75. DOI:10.1109/MSP.2005.1406483

[19] Z. Qiu, L. Xie, and Y. Hong: Quantized leaderless and leader-following consensus of high-order multi-agent systems with limited data rate. IEEE Trans. Automat. Control *61* (2016), 9, 2432–2447. DOI:10.1109/TAC.2015.2495579

[20] W. Shi, Q. Ling, K. Yuan, G. Wu, and W. Yin: On the linear convergence of the ADMM in decentralized consensus optimization. IEEE Trans. Signal Process. *62* (2014), 7, 1750–1761. DOI:10.1109/TSP.2014.2304432

[21] C. Wang, S. Xu, D. Yuan, B. Zhang, and Z. Zhang: Distributed online convex optimization with a bandit primal-dual mirror descent push-sum algorithm. Neurocomputing *497* (2022), 204–215. DOI:10.1016/j.neucom.2022.05.024

[22] J. Wang, L. Fu, Y. Gu, and T. Li: Convergence of distributed gradient-tracking-based optimization algorithms with random graphs. J. Systems Science Complexity *34* (2021), 4, 1438–1453. DOI:10.1007/s11424-021-9355-5

[23] Y. Wei, H. Fang, X. Zeng, J. Chen, and P. Panos: A smooth double proximal primal-dual algorithm for a class of distributed nonsmooth optimization problems. IEEE Trans. Automat. Control *65* (2020), 4, 1800–1806. DOI:10.1109/TAC.2019.2936355

[24] X. Xie, Q. Ling, P. Lu, W. Xu, and Z. Zhu: Evacuate before too late: distributed backup in inter-DC networks with progressive disasters. IEEE Trans. Parallel Distributed Systems *29* (2018), 5, 1058–1074. DOI:10.1109/TPDS.2017.2785385

[25] T. Xu and W. Wu: Accelerated ADMM-based fully distributed inverter-based Volt/Var control strategy for active distribution networks. IEEE Trans. Industr. Inform. *16* (2020), 12, 7532–7543. DOI:10.1109/TII.2020.2966713

[26] P. Yi and Y. Hong: Quantized subgradient algorithm and data-rate analysis for distributed optimization. IEEE Trans. Contro Network Systems *1* (2014), 4, 380–392. DOI:10.1109/TCNS.2014.2357513

[27] W. Yu, H. Liu, W. Z. Zheng, and Y. Zhu: Distributed discrete-time convex optimization with nonidentical local constraints over time-varying unbalanced directed graphs Automatica *134* (2021), 11, 109899. DOI:10.1016/j.automatica.2021.109899

[28] D. Yuan, Y. Hong, W. C. H. Daniel, and S. Xu: Distributed mirror descent for online composite optimization. IEEE Trans. Automat. Control *66* (2021), 2, 714–729. DOI:10.1109/TAC.2020.2987379

[29] D. Yuan, S. Xu, B. Zhang, and L. Rong: Distributed primal-dual stochastic subgradient algorithms for multi-agent optimization under inequality constraints. Int. J. Robust Nonlinear Control *23* (2013), 15, 1846–1868. DOI:10.1002/rnc.2856

[30] J. Zhang, H. Liu, M.-Ch. S. Anthony, Man-Cho, and Q. Ling: A penalty alternating direction method of multipliers for convex composite optimization over decentralized networks. IEEE Trans. Signal Process. *69* (2021), 4282–4295. DOI:10.1109/TSP.2021.3092347

[31] X. Zhao, P. Yi, and L. Li: Distributed policy evaluation via inexact ADMM in multi-agent reinforcement learning. Control Theory Technol. *18* (2020), 4, 362–378. DOI:10.1007/s11768-020-00007-x

[32] H. Zhou, X. Zeng, and Y. Hong: Adaptive exact penalty design for constrained distributed optimization. IEEE Trans. Automat. Control *64* (2019), 11, 4661–4667. DOI:10.1109/TAC.2019.2902612

Chenyang Liu, Anhui Engineering Laboratory of Human-Robot Integration System and Intelligent Equipment, School of Electrical Engineering and Automation, Anhui University, Hefei 230601, Anhui. P. R. China.

e-mail: z21201001@stu.ahu.edu.cn

Xiaohua Dou, Department of Techniques, Jiuquan Satellite Launching Center, Lanzhou 732750, Gansu. P. R. China.
e-mail: xibeilang_20@126.com

Yuan Fan, Anhui Engineering Laboratory of Human-Robot Integration System and Intelligent Equipment, School of Electrical Engineering and Automation, Anhui University, Hefei 230601, Anhui. P. R. China.
e-mail: yfan@ahu.edu.cn

Songsong Cheng, (Corresponding author) Anhui Engineering Laboratory of Human-Robot Integration System and Intelligent Equipment, School of Electrical Engineering and Automation, Anhui University, Hefei 230601, Anhui. P. R. China.
e-mail: sscheng@ahu.edu.cn