

Convergence Analysis of Consensus-ADMM for General QCQP

Huiping Huang^a, Hing Cheung So^b, Abdelhak M. Zoubir^a

^aDepartment of Electrical Engineering and Information Technology, Technische Universität Darmstadt, Darmstadt, Germany

^bDepartment of Electrical Engineering, City University of Hong Kong, Hong Kong, China

Abstract

We analyze the convergence properties of the consensus-alternating direction method of multipliers (ADMM) for solving general quadratically constrained quadratic programs. We prove that the augmented Lagrangian function value is monotonically non-increasing as long as the augmented Lagrangian parameter is chosen to be sufficiently large. Simulation results show that the augmented Lagrangian function is bounded from below when the matrix in the quadratic term of the objective function is positive definite. In such a case, the consensus-ADMM is convergent.

Keywords: Augmented Lagrangian function, Consensus-ADMM, quadratically constrained quadratic program

1. Introduction

An attractive algorithm, called consensus-alternating direction method of multipliers (ADMM), has been recently proposed in [1], for solving general quadratically constrained quadratic programs (QCQPs). Note that these problems could be non-convex and thus NP-hard, including several special cases in real-world applications, see e.g. [2–10]. The consensus-ADMM, primarily enjoys the following two advantages: i) it formulates the general QCQP in such a way that each subproblem is a QCQP with only one constraint, referred to as QCQP-1, which can be effectively solved irrespective of (non)convexity; and ii) memory efficient implementation and parallel/distributed processing are possible, which greatly save the effort of solving general QCQPs, especially when the dimensions of the resulting matrices and/or the number of constraints are large.

Despite the effectiveness of consensus-ADMM, only a weak convergence result was presented in [1]. That is, the authors claim that the consensus-ADMM converges to a Karush-Kuhn-Tucker (KKT) stationary point, under the key assumptions that the limit points do exist, and that the difference between two successive iterations converges to zero, see Theorem 1 in [1]. Note that such assumptions are nonstandard and overly restrictive [11]. In this paper, we analyze the convergence of the consensus-ADMM, and show that the

*The work of H. Huang is supported by the Graduate School CE within the Centre for Computational Engineering at Technische Universität Darmstadt.

Email addresses: h.huang@spg.tu-darmstadt.de (Huiping Huang), hcs0@ee.cityu.edu.hk (Hing Cheung So), zoubir@spg.tu-darmstadt.de (Abdelhak M. Zoubir)

augmented Lagrangian function is monotonically non-increasing as long as the Lagrangian parameter is larger than a certain value.

The remainder of the paper is organized as follows. The problem formulation and the consensus-ADMM are presented in Section 2. Convergence analyses are established in Section 3. Simulation results are shown in Section 4, while Section 5 concludes the paper.

Notation: Bold-faced lower-case and upper-case letters denote vectors and matrices, respectively. \mathbf{I}_n denotes the $n \times n$ identity matrix, and $\mathbf{1}$ and $\mathbf{0}$ are the all-one and all-zero vectors of appropriate size, respectively. Superscripts \cdot^T , \cdot^H and \cdot^{-1} stand for transpose, Hermitian transpose and inverse operators, respectively. $|\cdot|$ is the absolute operator. $\|\cdot\|_2$ denotes the ℓ_2 -norm of a vector. $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ denote the maximum and minimum eigenvalues, respectively. \mathbb{C} is the set of complex numbers, and $\Re\{\cdot\}$ returns the real part of its input complex-valued variable. $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^H \mathbf{b}$ is the inner product of vectors \mathbf{a} and \mathbf{b} of same sizes. $\max\{a, b\}$ returns the maximum between a and b . $\mathbf{X} \succ 0$ means that \mathbf{X} is positive definite. $E\{\cdot\}$ and $\text{Var}\{\cdot\}$ denote the expected value and variance of a random variable, respectively.

2. Problem Statement and Consensus-ADMM

General QCQP is an optimization problem that minimizes a quadratic function subject to quadratic inequality and equality constraints [12], which can be typically formulated as [1]

$$\min_{\mathbf{x} \in \mathbb{C}^n} \mathbf{x}^H \mathbf{A}_0 \mathbf{x} - 2\Re\{\mathbf{b}_0^H \mathbf{x}\} \quad (1a)$$

$$\text{s.t. } \mathbf{x}^H \mathbf{A}_i \mathbf{x} - 2\Re\{\mathbf{b}_i^H \mathbf{x}\} \leq c_i, \quad \forall i = 1, 2, \dots, m, \quad (1b)$$

where the known matrices $\mathbf{A}_i \in \mathbb{C}^{n \times n}$ are assumed to be general Hermitian matrices (possibly indefinite), and $\mathbf{b}_i \in \mathbb{C}^n$, for all $i = 0, 1, \dots, m$. Note that any quadratic equality constraint can always be reformulated as two inequality constraints of the form in (1b). Therefore, for simplicity, only inequality constraints are given in Problem (1). The following standard steps [13] guide us to the consensus-ADMM for solving the general QCQP, i.e., Problem (1).

Step i): We reformulate Problem (1) by introducing auxiliary variables $\{\mathbf{z}_i\}_{i=1}^m$, and settling the original variable \mathbf{x} and the auxiliary variables $\{\mathbf{z}_i\}_{i=1}^m$ in a separable manner, as

$$\min_{\mathbf{x}, \{\mathbf{z}_i\}} \mathbf{x}^H \mathbf{A}_0 \mathbf{x} - 2\Re\{\mathbf{b}_0^H \mathbf{x}\} \quad (2a)$$

$$\text{s.t. } \mathbf{z}_i^H \mathbf{A}_i \mathbf{z}_i - 2\Re\{\mathbf{b}_i^H \mathbf{z}_i\} \leq c_i, \quad (2b)$$

$$\mathbf{z}_i = \mathbf{x}, \quad \forall i = 1, 2, \dots, m. \quad (2c)$$

Step ii): We form the scaled-form augmented Lagrangian function according to Problem (2), by dealing with the equality constraints therein, i.e., $\mathbf{z}_i = \mathbf{x}$, as

$$\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\}) \triangleq \mathbf{x}^H \mathbf{A}_0 \mathbf{x} - 2\Re\{\mathbf{b}_0^H \mathbf{x}\} + \rho \sum_{i=1}^m (\|\mathbf{z}_i - \mathbf{x} + \mathbf{u}_i\|_2^2 - \|\mathbf{u}_i\|_2^2) \quad (3)$$

35 where $\rho > 0$ is the augmented Lagrangian parameter, and \mathbf{u}_i is the scaled dual variable corresponding to the equality constraint $\mathbf{z}_i = \mathbf{x}$ in Problem (2).

Step iii): The consensus-ADMM updating equations can be written down by separately solving $\{\mathbf{z}_i\}$ and \mathbf{x} , as

$$\mathbf{z}_i^{(k+1)} = \begin{cases} \arg \min_{\mathbf{z}_i} & \|\mathbf{z}_i - \mathbf{x}^{(k)} + \mathbf{u}_i^{(k)}\|_2^2 \\ \text{s.t.} & \mathbf{z}_i^H \mathbf{A}_i \mathbf{z}_i - 2\Re\{\mathbf{b}_i^H \mathbf{z}_i\} \leq c_i, \end{cases} \quad (4a)$$

$$\mathbf{x}^{(k+1)} = (\mathbf{A}_0 + m\rho\mathbf{I}_n)^{-1} \left[\mathbf{b}_0 + \rho \sum_{i=1}^m (\mathbf{z}_i^{(k+1)} + \mathbf{u}_i^{(k)}) \right], \quad (4b)$$

$$\mathbf{u}_i^{(k+1)} = \mathbf{u}_i^{(k)} + \mathbf{z}_i^{(k+1)} - \mathbf{x}^{(k+1)}, \quad (4c)$$

where the superscript $\cdot^{(k)}$ denotes the corresponding variable obtained at the k -th iteration. As has been mentioned in [1], we put the consensus-ADMM into this form due to the fact that each update of \mathbf{z}_i is a QCQP-1, which can be computed optimally despite the fact that the matrices \mathbf{A}_i may be indefinite. The 40 readers are referred to [1] for details of solving QCQP-1. Furthermore, in the form of (4), $\{\mathbf{z}_i\}_{i=1}^m$ can be updated in parallel, and the result of $(\mathbf{A}_0 + m\rho\mathbf{I}_n)^{-1}$ can be cached to save computations in the subsequent iterations. On the other hand, it is worth pointing out that, different from [1] in which the algorithm was presented by updating \mathbf{x} before $\{\mathbf{z}_i\}$, we update $\{\mathbf{z}_i\}$ before \mathbf{x} . This shall be shown to be very important when analyzing the convergence of the algorithm.

45 3. Convergence Analysis

This section is devoted to a theorem showing the monotonicity of the augmented Lagrangian function. First of all, from (4c), we have $\mathbf{z}_i^{(k+1)} + \mathbf{u}_i^{(k)} = \mathbf{u}_i^{(k+1)} + \mathbf{x}^{(k+1)}$. Substituting it into (4b) yields

$$\rho \sum_{i=1}^m \mathbf{u}_i^{(k+1)} = \mathbf{A}_0 \mathbf{x}^{(k+1)} - \mathbf{b}_0. \quad (5)$$

Remark 1. Equality (5) is one of the keys to the success of proving the monotonicity of the augmented Lagrangian function, which will be shown subsequently. Moreover, such an equality or other similar results are not available if we update \mathbf{x} before $\{\mathbf{z}_i\}$, as done in [1].

Now, we show the monotonicity of the augmented Lagrangian function, $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\})$, defined in (3).

Theorem. As long as the augmented Lagrangian parameter

$$\rho > \max \left\{ -\frac{\lambda_{\min}}{m}, \frac{\sqrt{mC}\lambda_{\max} + \max\{-\lambda_{\min}, 0\}}{m} \right\}, \quad (6)$$

50 where C stands for some constant, and λ_{\min} and λ_{\max} are short for $\lambda_{\min}(\mathbf{A}_0)$ and $\lambda_{\max}(\mathbf{A}_0)$, respectively, we have:

- $\mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k+1)}\}) \leq \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k)}\}, \{\mathbf{u}_i^{(k)}\}), \forall k = 0, 1, 2, \dots;$

- $\lim_{k \rightarrow \infty} \left\| \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right\|_2 = 0.$

Proof. To show $\mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k+1)}\}) \leq \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k)}\}, \{\mathbf{u}_i^{(k)}\})$ holds $\forall k = 0, 1, 2, \dots$, we formulate their difference as

$$\begin{aligned} & \mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k+1)}\}) - \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k)}\}, \{\mathbf{u}_i^{(k)}\}) \\ &= \left[\mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k+1)}\}) - \mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}) \right] \end{aligned} \quad (7a)$$

$$+ \left[\mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}) - \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}) \right] \quad (7b)$$

$$+ \left[\mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}) - \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k)}\}, \{\mathbf{u}_i^{(k)}\}) \right]. \quad (7c)$$

In the sequel, we successively deal with (7a), (7b), and (7c). Firstly, for (7a), it is calculated as

$$\begin{aligned} & \mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k+1)}\}) - \mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}) \\ & \stackrel{(8a)}{=} \rho \sum_{i=1}^m \left(\left\| \mathbf{z}_i^{(k+1)} - \mathbf{x}^{(k+1)} + \mathbf{u}_i^{(k+1)} \right\|_2^2 - \left\| \mathbf{z}_i^{(k+1)} - \mathbf{x}^{(k+1)} + \mathbf{u}_i^{(k)} \right\|_2^2 - \left\| \mathbf{u}_i^{(k+1)} \right\|_2^2 + \left\| \mathbf{u}_i^{(k)} \right\|_2^2 \right) \\ & \stackrel{(8b)}{=} \rho \sum_{i=1}^m \left(\left\| 2\mathbf{u}_i^{(k+1)} - \mathbf{u}_i^{(k)} \right\|_2^2 - 2 \left\| \mathbf{u}_i^{(k+1)} \right\|_2^2 + \left\| \mathbf{u}_i^{(k)} \right\|_2^2 \right) \\ & = 2\rho \sum_{i=1}^m \left\| \mathbf{u}_i^{(k+1)} - \mathbf{u}_i^{(k)} \right\|_2^2 \stackrel{(8c)}{\leq} \rho C \left\| \sum_{i=1}^m (\mathbf{u}_i^{(k+1)} - \mathbf{u}_i^{(k)}) \right\|_2^2 \stackrel{(8d)}{\leq} \frac{C}{\rho} \lambda_{\max}^2(\mathbf{A}_0) \left\| \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right\|_2^2, \end{aligned} \quad (8)$$

where in (8a) we use the definition of $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\})$; in (8b) we use (4c); in (8c) we use the fact discussed in Remark 2 as below; in (8d) we employ (5) and the inequality $\|\mathbf{H}\mathbf{a}\|_2^2 \leq \lambda_{\max}^2(\mathbf{H})\|\mathbf{a}\|_2^2$ for any Hermitian matrix $\mathbf{H} \in \mathbb{C}^{n \times n}$ and any vector $\mathbf{a} \in \mathbb{C}^n$.

Next, we move on to (7b), which is calculated as

$$\begin{aligned} & \mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}) - \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}) \\ & \stackrel{(9a)}{\leq} \Re \left\{ \left\langle \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}), \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right\rangle \right\} - \frac{\gamma_{\mathbf{x}}}{2} \left\| \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right\|_2^2 \\ & \stackrel{(9b)}{=} -(\lambda_{\min}(\mathbf{A}_0) + m\rho) \left\| \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right\|_2^2, \end{aligned} \quad (9)$$

where in (9a) we utilize the fact that function $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\})$ is strongly convex (see Remark 3 below) with respect to (w.r.t.) \mathbf{x} with parameter $\gamma_{\mathbf{x}} > 0$ [14]; in (9b) we employ the optimality condition, namely, $\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}) = \mathbf{0}$, and $\gamma_{\mathbf{x}} = 2\lambda_{\min}(\mathbf{A}_0) + 2m\rho$ (see Remark 3).

Lastly, (7c) is calculated as

$$\begin{aligned} & \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k)}\}) - \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k)}\}, \{\mathbf{u}_i^{(k)}\}) \\ & \stackrel{(10a)}{=} \rho \sum_{i=1}^m \left(\left\| \mathbf{z}_i^{(k+1)} - \mathbf{x}^{(k)} + \mathbf{u}_i^{(k)} \right\|_2^2 - \left\| \mathbf{z}_i^{(k)} - \mathbf{x}^{(k)} + \mathbf{u}_i^{(k)} \right\|_2^2 \right) \stackrel{(10b)}{\leq} 0, \end{aligned} \quad (10)$$

where we utilize the definition of $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\})$ in (10a); in (10b) we use the fact that $\mathbf{z}_i^{(k+1)}$ is optimal to (4a).

Therefore, by substituting inequalities (8), (9), and (10) back into (7), we have

$$\begin{aligned} & \mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k+1)}\}) - \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k)}\}, \{\mathbf{u}_i^{(k)}\}) \\ & \leq \underbrace{\left(\frac{C}{\rho}\lambda_{\max}^2 - \lambda_{\min} - m\rho\right)}_{(i)} \left\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\right\|_2^2. \end{aligned}$$

We provide the following discussion regarding the term (i) of the above inequality. It is seen that if

$$\rho < \left(-\sqrt{\lambda_{\min}^2 + 4mC\lambda_{\max}^2} - \lambda_{\min}\right)/(2m)$$

(which should be deleted because of $\rho > 0$) or

$$\rho > \left(\sqrt{\lambda_{\min}^2 + 4mC\lambda_{\max}^2} - \lambda_{\min}\right)/(2m),$$

the coefficient $\frac{C}{\rho}\lambda_{\max}^2 - \lambda_{\min} - m\rho < 0$ and hence (i) ≤ 0 . Furthermore, as $\sqrt{\lambda_{\min}^2 + 4mC\lambda_{\max}^2} - \lambda_{\min} \leq 2\sqrt{mC}\lambda_{\max} + |\lambda_{\min}| - \lambda_{\min} = 2\sqrt{mC}\lambda_{\max} + 2\max\{-\lambda_{\min}, 0\}$, we reach the following conclusion: (i) ≤ 0 , as long as

$$\rho > \left(\sqrt{mC}\lambda_{\max} + \max\{-\lambda_{\min}, 0\}\right)/m. \quad (11)$$

Therefore, combining (11) and (20), we conclude that $\mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k+1)}\}) - \mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k)}\}, \{\mathbf{u}_i^{(k)}\}) \leq$ (i) ≤ 0 as long as ρ satisfies (6).

To prove the second part of the theorem, we first define $\Phi \triangleq -\frac{C}{\rho}\lambda_{\max}^2 + \lambda_{\min} + m\rho$, and denote $\mathcal{L}(\mathbf{x}^{(k+1)}, \{\mathbf{z}_i^{(k+1)}\}, \{\mathbf{u}_i^{(k+1)}\})$ and $\mathcal{L}(\mathbf{x}^{(k)}, \{\mathbf{z}_i^{(k)}\}, \{\mathbf{u}_i^{(k)}\})$ by $\mathcal{L}^{(k+1)}$ and $\mathcal{L}^{(k)}$, respectively. Then, when (6) holds, we have $\Phi > 0$ and

$$\mathcal{L}^{(k)} - \mathcal{L}^{(k+1)} \geq \Phi \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2^2, \quad \forall k = 0, 1, 2, \dots$$

Summing all the above inequalities over all $k \geq 0$, we obtain

$$\mathcal{L}^{(0)} \geq \Phi \sum_{k=0}^{\infty} \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2^2,$$

which implies [15] that $\lim_{k \rightarrow \infty} \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2 = 0$. □

Remark 2. It is easy to see that $\left\|\mathbf{u}_i^{(k+1)} - \mathbf{u}_i^{(k)}\right\|_2^2, \forall i = 1, 2, \dots, m$, are bounded from above. Hence, there exists an upper bound for $\sum_{i=1}^m \left\|\mathbf{u}_i^{(k+1)} - \mathbf{u}_i^{(k)}\right\|_2^2$. We can find a real-valued constant C , such that

$$\sum_{i=1}^m \left\|\mathbf{u}_i^{(k+1)} - \mathbf{u}_i^{(k)}\right\|_2^2 \leq C \left\|\sum_{i=1}^m \left(\mathbf{u}_i^{(k+1)} - \mathbf{u}_i^{(k)}\right)\right\|_2^2. \quad (12)$$

In the case where $m = 1$ (i.e., only one constraint in Problem (1)), the above inequality reduces to

$$\left\|\mathbf{u}_1^{(k+1)} - \mathbf{u}_1^{(k)}\right\|_2^2 \leq C \left\|\mathbf{u}_1^{(k+1)} - \mathbf{u}_1^{(k)}\right\|_2^2, \quad (13)$$

which implies that the constant C can be selected as

$$\begin{cases} C \text{ is arbitrary,} & \text{if } \mathbf{u}_1^{(k+1)} = \mathbf{u}_1^{(k)}, \\ C \geq 1, & \text{otherwise.} \end{cases} \quad (14)$$

65 In what follows, we provide an explanation on how to choose C from a statistical perspective. For the sake of notational simplicity, we define $\mathbf{d}_i \triangleq \mathbf{u}_i^{(k+1)} - \mathbf{u}_i^{(k)} \in \mathbb{C}^n$, $\forall i = 1, 2, \dots, m$, and $\mathbf{d}_i = [d_{i(1)}, d_{i(2)}, \dots, d_{i(n)}]^T$ where $d_{i(j)}$ is the j -th entry of \mathbf{d}_i . We make the following assumptions:

A1) \mathbf{d}_i 's are independently and identically distributed (i.i.d.) complex random variables;

A2) for any \mathbf{d}_i , all its entries are i.i.d.;

70 A3) $\mathbb{E}\{\mathbf{d}_i\} = \mu \mathbf{1}$ and $\text{Var}\{\mathbf{d}_i\} = \sigma^2 \mathbf{I}_n$, $\forall i = 1, 2, \dots, m$.

These assumptions indicate that $\mathbb{E}\{d_{i(j)}\} = \mu$ and $\text{Var}\{d_{i(j)}\} = \sigma^2$, for all $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$.

We focus on the expected value of (12), which is given as

$$\mathbb{E} \left\{ \sum_{i=1}^m \|\mathbf{d}_i\|_2^2 \right\} \leq \mathbb{E} \left\{ C \left\| \sum_{i=1}^m \mathbf{d}_i \right\|_2^2 \right\}. \quad (15)$$

The left-hand side of (15) is determined as

$$\begin{aligned} \mathbb{E} \left\{ \sum_{i=1}^m \|\mathbf{d}_i\|_2^2 \right\} &= \sum_{i=1}^m \mathbb{E} \left\{ \|\mathbf{d}_i\|_2^2 \right\} \\ &= \sum_{i=1}^m \mathbb{E} \left\{ \sum_{j=1}^n |d_{i(j)}|^2 \right\} \\ &= \sum_{i=1}^m \sum_{j=1}^n \mathbb{E} \{ |d_{i(j)}|^2 \} \\ &\stackrel{(16a)}{=} \sum_{i=1}^m \sum_{j=1}^n \left[\text{Var}\{d_{i(j)}\} + |\mathbb{E}\{d_{i(j)}\}|^2 \right] \\ &= \sum_{i=1}^m \sum_{j=1}^n (\sigma^2 + |\mu|^2) \\ &= mn(\sigma^2 + |\mu|^2), \end{aligned} \quad (16)$$

where in (16a) we use the relationship between expected value and variance of a complex random variable, see for example Pages 117 and 118 in [16]. Similarly, the right-hand side of (15) is computed as

$$\begin{aligned} \mathbb{E} \left\{ C \left\| \sum_{i=1}^m \mathbf{d}_i \right\|_2^2 \right\} &= C \times \mathbb{E} \left\{ \left\| \sum_{i=1}^m \mathbf{d}_i \right\|_2^2 \right\} \\ &= C \times \mathbb{E} \left\{ \sum_{j=1}^n \left| \sum_{i=1}^m d_{i(j)} \right|^2 \right\} \end{aligned}$$

$$\begin{aligned}
&= C \times \sum_{j=1}^n \mathbb{E} \left\{ \left| \sum_{i=1}^m d_{i(j)} \right|^2 \right\} \\
&\stackrel{(17a)}{=} C \times \sum_{j=1}^n \left[\text{Var} \left\{ \sum_{i=1}^m d_{i(j)} \right\} + \left| \mathbb{E} \left\{ \sum_{i=1}^m d_{i(j)} \right\} \right|^2 \right] \\
&= C \times \sum_{j=1}^n \left[\sum_{i=1}^m \text{Var} \{d_{i(j)}\} + \left| \sum_{i=1}^m \mathbb{E} \{d_{i(j)}\} \right|^2 \right] \\
&= C \times \sum_{j=1}^n (m\sigma^2 + m^2|\mu|^2) \\
&= Cmn(\sigma^2 + m|\mu|^2), \tag{17}
\end{aligned}$$

where in (17a) we use the same strategy as in (16a). Substituting (16) and (17) into (15) yields

$$\sigma^2 + |\mu|^2 \leq C(\sigma^2 + m|\mu|^2), \tag{18}$$

which implies that

$$\begin{cases} C \text{ is arbitrary,} & \text{if } \mu = 0 \text{ and } \sigma^2 = 0, \\ C \geq \frac{\sigma^2 + |\mu|^2}{\sigma^2 + m|\mu|^2}, & \text{otherwise.} \end{cases} \tag{19}$$

Note that: i) the case where $\mu = 0$ and $\sigma^2 = 0$ indicates that $\mathbf{d}_i = \mathbf{0}$ (i.e., $\mathbf{u}_i^{(k+1)} = \mathbf{u}_i^{(k)}$), $\forall i = 1, 2, \dots, m$; and ii) by assigning $m = 1$, we have $\frac{\sigma^2 + |\mu|^2}{\sigma^2 + m|\mu|^2} = 1$. Therefore, when $m = 1$, (19) is consistent with (14).

Remark 3. The statement that $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\})$ is strongly convex w.r.t. \mathbf{x} with parameter $\gamma_{\mathbf{x}}$, is equivalent to the statement that $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\}) - \frac{\gamma_{\mathbf{x}}}{2} \|\mathbf{x}\|_2^2$ is convex w.r.t. \mathbf{x} , see [17]. Note that $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\}) - \frac{\gamma_{\mathbf{x}}}{2} \|\mathbf{x}\|_2^2 = \mathbf{x}^H [\mathbf{A}_0 + (m\rho - \frac{\gamma_{\mathbf{x}}}{2})\mathbf{I}_n] \mathbf{x} - 2\Re \left\{ (\mathbf{b}_0 + \rho \sum_{i=1}^m (\mathbf{z}_i + \mathbf{u}_i))^H \mathbf{x} \right\} + C'$, where C' is a constant independent of \mathbf{x} . $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\}) - \frac{\gamma_{\mathbf{x}}}{2} \|\mathbf{x}\|_2^2$ is convex if and only if $\mathbf{A}_0 + (m\rho - \frac{\gamma_{\mathbf{x}}}{2})\mathbf{I}_n \succeq 0$, which indicates that parameter $\gamma_{\mathbf{x}} = 2\lambda_{\min}(\mathbf{A}_0) + 2m\rho$. This equality together with $\gamma_{\mathbf{x}} > 0$ yields $\rho > -\lambda_{\min}(\mathbf{A}_0)/m$. Since $\rho > 0$, we have

$$\rho > \max \{-\lambda_{\min}(\mathbf{A}_0)/m, 0\}. \tag{20}$$

Remark 4. When $\mathbf{A}_0 \succ 0$, our simulation results show that the augmented Lagrangian function $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\})$ is bounded from below during the consensus-ADMM iteration, i.e., (4). This observation together with the result in Theorem indicates the convergence of the consensus-ADMM. However, when \mathbf{A}_0 is indefinite, our simulation results show that $\mathcal{L}(\mathbf{x}, \{\mathbf{z}_i\}, \{\mathbf{u}_i\})$ is unbounded.

4. Simulation Results

In this section, we provide numerical simulations to confirm our theoretical analyses in Section 3. After fixing the problem dimensions n and m , we first generate $\mathbf{x}_{\text{feas}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_n)$. A Hermitian indefinite matrix \mathbf{A}_i ($i = 0, 1, \dots, m$) is generated by first randomly drawing an $n \times n$ matrix from $\mathcal{CN}(0, 1)$, and then taking

the average of its Hermitian and itself. In the situation where we need $\mathbf{A}_0 \succ 0$, we add $\epsilon \mathbf{I}_n$ to its indefinite counterpart, with ϵ larger than the absolute of the smallest eigenvalue of the indefinite matrix. We generate $\mathbf{b}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_n)$, $\forall i = 0, 1, \dots, m$, and $c_i = \mathbf{x}_{\text{feas}}^H \mathbf{A}_i \mathbf{x}_{\text{feas}} - 2\Re\{\mathbf{b}_i^H \mathbf{x}_{\text{feas}}\} + |v_i|$, $\forall i = 1, 2, \dots, m$, where v_i is randomly drawn from $\mathcal{N}(0, 1)$. The constructed constraint set is thus guaranteed to be non-empty, because at least \mathbf{x}_{feas} is feasible. The above-described problem setting is similar to those considered in [1] and [2]. We initiate the consensus-ADMM algorithm with $\mathbf{x}^{(0)} = \mathbf{x}_{\text{feas}}$ and $\mathbf{u}_i^{(0)} = \mathbf{0}$, $\forall i = 1, 2, \dots, m$.

In the first example, we set $n = 10$ and $m = 5$. All \mathbf{A}_i 's are indefinite. The augmented Lagrangian function value versus iteration index is plotted in Fig. 1. It is seen that the function value is monotonically non-increasing, but unbounded.

In all the remaining examples, \mathbf{A}_i ($i = 1, 2, \dots, m$) are indefinite and $\mathbf{A}_0 \succ 0$. In the second example, $n = 10$ and $m = 5$. The augmented Lagrangian function value versus iteration index is shown in Fig. 2. It is observed that the function value is not monotonic when $\rho = 2$ and $\rho = 5$; while it is monotonically non-increasing and bounded from below when $\rho \geq 10$. In Fig. 3, we plot the function value versus k in 20 independent runs, with $n = 10$, $m = 5$, and $\rho = 10$. The function value versus k is depicted in Fig. 4 for different number of constraints, with $n = 10$ and $\rho = 20$. All the curves in Figs. 3 and 4 indicate the convergence of the augmented Lagrangian function value sequence $\{\mathcal{L}^{(k)}\}$.

In the last example, we test the distance of the variables, by using the ℓ_2 -norm as performance metric. We set $n = 10$, $m = 5$, and $\rho = 10$. The results are shown in Fig. 5. It is seen that, both $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2$ and $\sum_{i=1}^m \|\mathbf{z}_i^{(k)} - \mathbf{x}^{(k)}\|_2$ approach 0 when k is large enough.

5. Conclusion

We studied the convergence property of the consensus-alternating direction method of multipliers (ADMM) for general quadratically constrained quadratic programs (QCQPs). We have proved that the augmented Lagrangian function is monotonically non-increasing if the augmented Lagrangian parameter is larger than a certain value. Simulation results showed that the augmented Lagrangian function is bounded from below when the matrix in the quadratic term of the objective function is positive definite. In such cases, the consensus-ADMM was shown to generate a convergent Lagrangian function value sequence. However, the augmented Lagrangian function is unbounded when the matrix in the quadratic term of the objective function is indefinite. We further proved that the distance of the variables between two successive iterations converges to 0. Numerical simulations with definite as well as indefinite matrices in the QCQP were conducted. The simulation results verified the theoretical development.

References

- [1] K. Huang, N. D. Sidiropoulos, Consensus-ADMM for general quadratically constrained quadratic programming, IEEE Transactions on Signal Processing 64 (20) (2016) 5297–5310.

115

doi:10.1109/TSP.2016.2593681.

120

- [2] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar, N. D. Sidiropoulos, Feasible point pursuit and successive approximation of non-convex QCQPs, *IEEE Signal Processing Letters* 22 (7) (2015) 804–808. doi:10.1109/LSP.2014.2370033.

125

- [3] J. Park, S. Boyd, General heuristics for nonconvex quadratically constrained quadratic programming, *arXiv* (May 2017).
URL <https://arxiv.org/abs/1703.07870>

- [4] O. Mehanna, N. D. Sidiropoulos, G. B. Giannakis, Joint multicast beamforming and antenna selection, *IEEE Transactions on Signal Processing* 61 (10) (2013) 2660–2674. doi:10.1109/TSP.2013.2252167.

- [5] F. Liu, C. Masouros, A. Li, H. Sun, L. Hanzo, MU-MIMO communications with MIMO radar: From co-existence to joint transmission, *IEEE Transactions on Wireless Communications* 17 (4) (2018) 2755–2770. doi:10.1109/TWC.2018.2803045.

- [6] S. A. Hamza, M. G. Amin, Hybrid sparse array beamforming design for general rank signal models, *IEEE Transactions on Signal Processing* 67 (24) (2019) 6215–6226. doi:10.1109/TSP.2019.2952052.

130

- [7] Z. Zheng, Y. Fu, W.-Q. Wang, H. C. So, Sparse array design for adaptive beamforming via semidefinite relaxation, *IEEE Signal Processing Letters* 27 (2020) 925–929. doi:10.1109/LSP.2020.2997253.

- [8] Z. Cheng, Z. He, B. Liao, Hybrid beamforming design for OFDM dual-function radar-communication system, *IEEE Journal of Selected Topics in Signal Processing* 15 (6) (2021) 1455–1467. doi:10.1109/JSTSP.2021.3118276.

135

- [9] J. Dan, S. Geirnaert, A. Bertrand, Grouped variable selection for generalized eigenvalue problems, *Signal Processing* 195 (2022) 108476. doi:<https://doi.org/10.1016/j.sigpro.2022.108476>.

- [10] H. Huang, H. C. So, A. M. Zoubir, Off-grid direction-of-arrival estimation using second-order Taylor approximation, *Signal Processing* 196 (2022) 108513. doi:<https://doi.org/10.1016/j.sigpro.2022.108513>.

140

- [11] M. Hong, Z.-Q. Luo, M. Razaviyayn, Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems, *SIAM Journal on Optimization* 26 (1) (2016) 337–364. doi:10.1137/140990309.

- [12] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, S. Zhang, Semidefinite relaxation of quadratic optimization problems, *IEEE Signal Processing Magazine* 27 (3) (2010) 20–34. doi:10.1109/MSP.2010.936019.

145

- [13] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Foundation and Trends in Machine Learning* 3 (1) (2011) 1–122. doi:10.1561/22000000016.

- [14] E. K. Ryu, S. P. Boyd, A primer on monotone operator methods, *Applied and Computational Mathematics* 15 (1) (2016) 3–43.
- [15] C. Lu, Z. Lin, S. Yan, Smoothed low rank and sparse matrix recovery by iteratively reweighted least squares minimization, *IEEE Transactions on Image Processing* 24 (2) (2015) 646–654. doi:10.1109/TIP.2014.2380155.
- [16] K. I. Park, Characterization of random variables, in: *Fundamentals of Probability and Stochastic Processes with Applications to Communications*, Springer International Publishing, Cham, 2018, pp. 109–133.
URL https://doi.org/10.1007/978-3-319-68075-0_5
- [17] X. Zhou, On the Fenchel duality between strong convexity and Lipschitz continuous gradient, arXiv (March 2018).
URL <https://arxiv.org/abs/1803.06573>

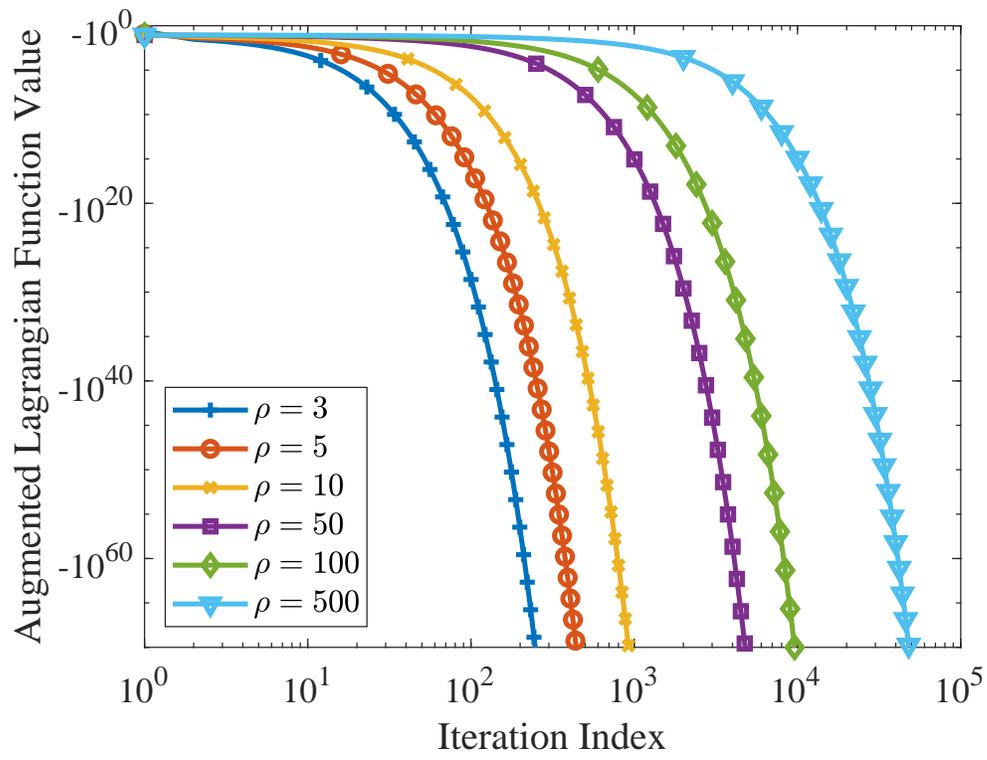


Figure 1: Function value $\mathcal{L}^{(k)}$ versus iteration index k , with $n = 10$, $m = 5$, and indefinite \mathbf{A}_0 .

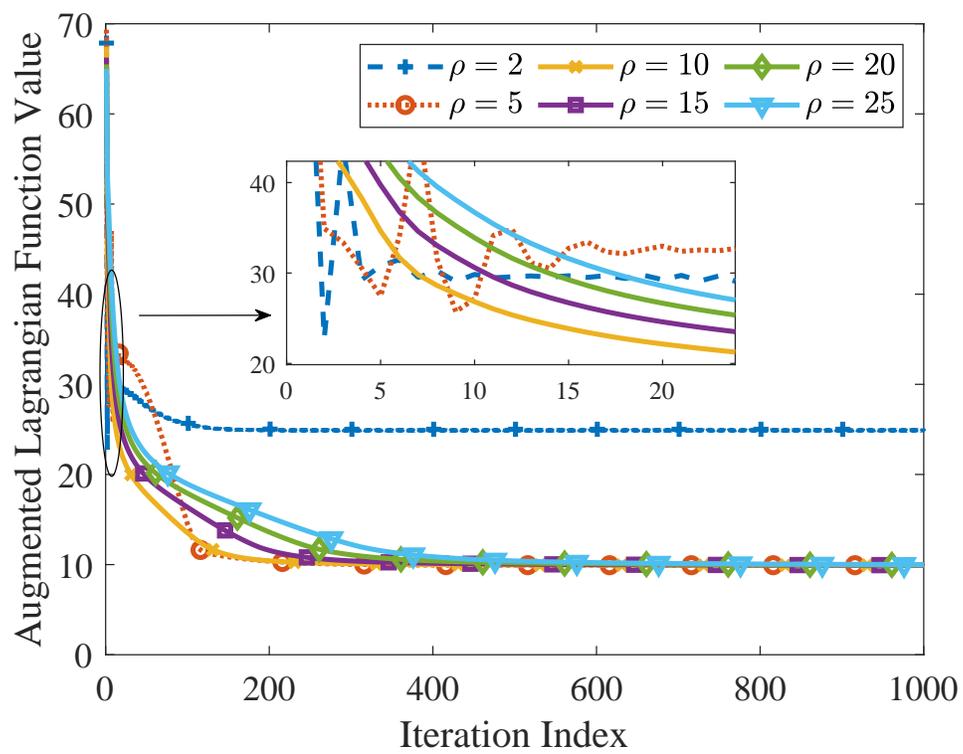


Figure 2: Function value $\mathcal{L}^{(k)}$ versus iteration index k , with $n = 10$, $m = 5$, and $\mathbf{A}_0 \succ 0$.

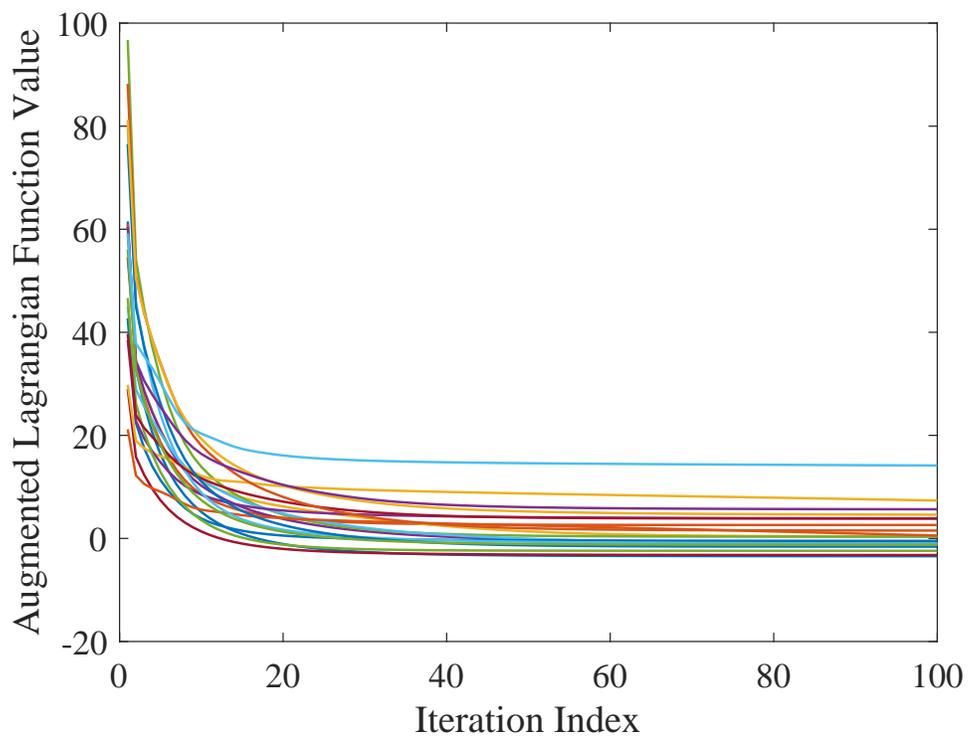


Figure 3: Function value $\mathcal{L}^{(k)}$ versus iteration index k in 20 runs, with $n = 10$, $m = 5$, $\rho = 10$, and $\mathbf{A}_0 \succ 0$.

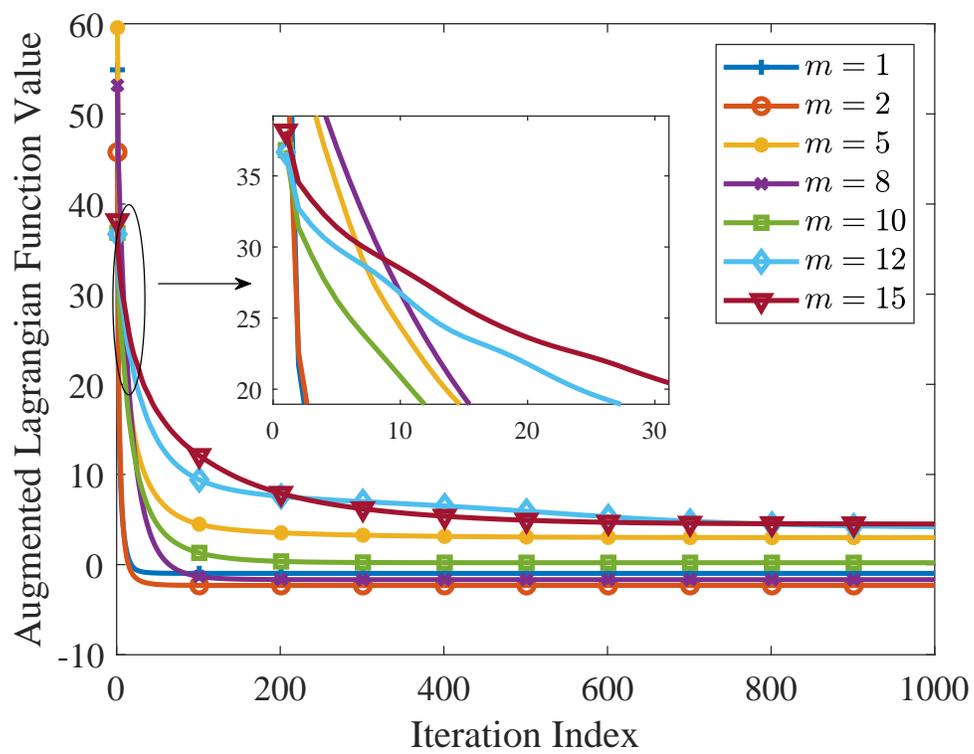


Figure 4: Function value $\mathcal{L}^{(k)}$ versus iteration index k for different number of constraints, with $n = 10$, $\rho = 20$, and $\mathbf{A}_0 \succ 0$.

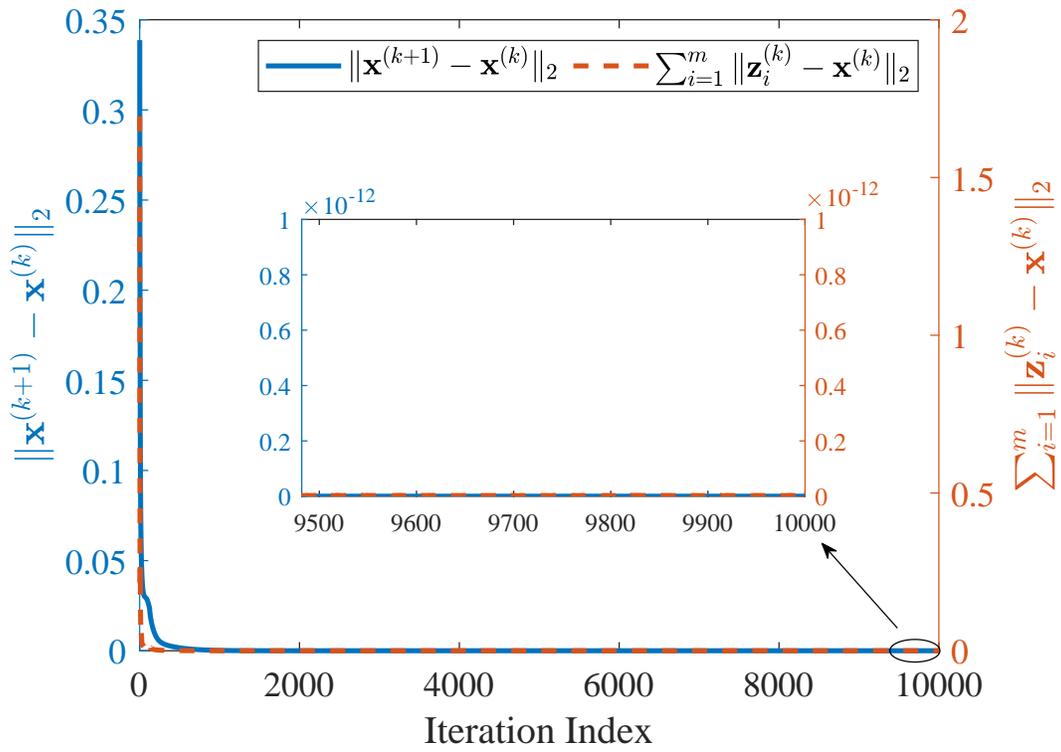


Figure 5: Distance of variables versus iteration index k , with $n = 10$, $m = 5$, $\rho = 10$, and $\mathbf{A}_0 \succ 0$.