

# Quantized Constant-Envelope Waveform Design for Massive MIMO DFRC Systems

Zheyu Wu, Ya-Feng Liu, Wei-Kun Chen, and Christos Masouros

**Abstract**—Both dual-functional radar-communication (DFRC) and massive multiple-input multiple-output (MIMO) have been recognized as enabling technologies for 6G wireless networks. This paper considers the advanced waveform design for hardware-efficient massive MIMO DFRC systems. Specifically, the transmit waveform is imposed with the quantized constant-envelope (QCE) constraint, which facilitates the employment of low-resolution digital-to-analog converters (DACs) and power-efficient amplifiers. The waveform design problem is formulated as the minimization of the mean square error (MSE) between the designed and desired beampatterns subject to the constructive interference (CI)-based communication quality of service (QoS) constraints and the QCE constraint. To solve the formulated problem, we first utilize the penalty technique to transform the discrete problem into an equivalent continuous penalty model. Then, we propose an inexact augmented Lagrangian method (ALM) algorithm for solving the penalty model. In particular, the ALM subproblem at each iteration is solved by a custom-built block successive upper-bound minimization (BSUM) algorithm, which admits closed-form updates, making the proposed inexact ALM algorithm computationally efficient. Simulation results demonstrate the superiority of the proposed approach over existing state-of-the-art ones. In addition, extensive simulations are conducted to examine the impact of various system parameters on the trade-off between communication and radar performances.

**Index Terms**—Augmented Lagrangian method, constructive interference, dual-functional radar communication, massive multiple-input multiple-output, quantized constant envelope.

## I. INTRODUCTION

Communication-radar spectrum sharing (CRSS) has emerged as a promising paradigm to address the spectrum congestion issue in today's ever-expanding wireless networks [1]. A straightforward approach to achieve CRSS is to allow individual radar and communication systems to share the frequency band, which, however, necessitates the exchange of side-information for managing cross-interference, leading to a prohibitive cooperation cost [2]. As a step ahead from the coexistence scheme, the dual-functional radar-communication (DFRC) system integrates both radar and communication functionalities on a single platform, which not only eliminates the need for information exchange but also greatly reduces the hardware cost and system complexity. The DFRC system

has gained significant attention recently in both academia and industry. In particular, the DFRC system has been considered as an enabling technology for future IoT applications and integrated sensing and communications (ISAC) has been listed as one of the six major usage scenarios in future 6G wireless networks [3], [4], [5], [6], [7], [8], [9], [10].

The performance of the DFRC system can be further enhanced when equipped with massive multiple-input multiple-output (MIMO) [11], another key technology for 5G and 6G wireless networks, as the increased number of antennas offers greater flexibility to mitigate the multiuser interference and shape the spectral beampattern for radar applications. Despite the great potential of massive MIMO DFRC systems, the high hardware cost and energy consumption pose major practical challenges. In particular, the numbers of digital-to-analog converters (DACs) and power amplifiers (PAs), which are the most power-hungry and costly components in fully digital systems, both scale with the number of transmit antennas. Therefore, it is crucial to employ low-resolution DACs and power-efficient PAs in massive MIMO DFRC systems. It is well-known that when constant-envelope (CE) waveforms are transmitted, the PAs can operate close to saturation without distortion and achieve their highest efficiency [12]. Quantized CE (QCE) transmission is a novel scheme that combines the CE requirement on the transmit waveform with the use of low-resolution DACs, in which case the transmit signals are constrained to have a fixed amplitude and their phases are limited to a finite set of values. The QCE transmission facilitates the use of both low-resolution DACs and the most power-efficient PAs, making it a favorable transmission scheme in massive MIMO systems. In this paper, we focus on the QCE transmit waveform design for massive MIMO DFRC systems.

### A. Related Works

Extensive research efforts have been devoted to the design and analysis of QCE precoding in massive MIMO communication-only systems [13], [14], [15], [16], [17], [18], [19]. The QCE precoding schemes can be broadly categorized into two classes: linear-quantized precoding and nonlinear precoding. Linear-quantized precoding schemes directly quantize the output of linear precoders, which, although simple, often fail to provide satisfactory performance [13], [14], [15]. In contrast, nonlinear precoding schemes optimize the transmit signal based on both the channel and the data symbols by solving appropriate optimization problems, which generally yield significantly superior performance compared to linear-quantized schemes [16], [17], [18], [19].

Z. Wu and Y.-F. Liu are with the State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (e-mail: {wuzy, yafliu}@lsec.cc.ac.cn). W.-K. Chen is with the School of Mathematics and Statistics/Beijing Key Laboratory on MCAACI, Beijing Institute of Technology, Beijing 100081, China (e-mail: chenweikun@bit.edu.cn). C. Masouros is with the Department of Electronic and Electrical Engineering, University College London, Torrington Place, London, WC1E 7JE, UK (e-mail: c.masouros@ucl.ac.uk).

Various nonlinear precoders have been proposed based on different criteria. In particular, the concept of constructive interference (CI) [20], [21], [22] has been incorporated into nonlinear precoding design [18], [19]. The main idea is to shape the multiuser interference (MUI) such that it aligns constructively with the signal of interest, thereby maximizing the useful signal power [23]. Due to the ability to exploit MUI, CI-based nonlinear precoders generally outperform those designed based on the classical mean square error (MSE) metric, which treats MUI as a detrimental factor. In addition to numerous notable works on QCE precoding for communication systems, there are also a few works investigating the QCE waveform design for massive MIMO radar systems; see [24], [25], [26] and the references therein. It is worth mentioning that as an extreme case of QCE transmission, one-bit waveform design (where one-bit DACs are employed) has attracted independent research interests in both communication systems [27], [28], [29], [30], [31] and radar systems [32], [33], [34].

Note that, the aforementioned works on QCE waveform design focus on either communication-only or radar-only scenarios. To date, there have been very few studies on QCE waveform design for DFRC systems. As far as we know, the only existing works are [35], [36], [37], all of which consider the extreme one-bit case. Specifically, in [35], the one-bit DFRC waveform is designed to minimize the communication MSE under the Cramér-Rao bound (CRB) constraint for radar direction-of-arrive estimation. To encompass wider applications for radar (e.g., target tracking), the works [36] and [37] adopt the MSE between the designed and desired beampatterns and the similarity between the designed and desired waveforms as radar metrics, respectively, which are minimized (in weight) jointly with the communication MSE.

There are two main limitations in the current studies on QCE waveform design for DFRC systems. First, the existing works [35], [36], [37] all adopt the MSE as the communication metric, which has been shown to be less effective than the CI metric in the context of nonlinear precoding. Although the works [38], [39], [40] have adopted the CI metric in the waveform design for DFRC systems, they do not consider the QCE constraint. Second, as mentioned earlier, the existing works [35], [36], [37] all focus on the one-bit case, and the proposed algorithms therein cannot be directly extended to handle the general QCE constraint. To the best of our knowledge, the general QCE waveform design for DFRC systems has not been explored in the literature, possibly due to the technical challenges introduced by both the QCE constraint and the DFRC scenario.

## B. Our Contributions

In this paper, we consider the general QCE waveform design for massive MIMO DFRC systems. The main contributions are summarized as follows.

- *Novel CI-based QoS-constrained general QCE waveform design problem formulation.* For the first time, we investigate the waveform design problem for DFRC systems under general QCE constraints. We employ CI as the communication metric and formulate the problem as the

minimization of the MSE between the designed and desired beampatterns subject to the CI-based communication QoS constraints and the general QCE constraint. This results in a large-scale optimization problem with a nonconvex quartic objective function, numerous linear constraints, and discrete QCE constraints.

- *Efficient approach for solving the formulated problem.* We propose an efficient approach for solving the formulated problem. Specifically, we first apply the penalty technique to transform the discrete model with complicated QCE constraints into a continuous penalty model, and theoretically establish the equivalence between these two models. Then, we propose an inexact augmented Lagrangian method (ALM) algorithm for solving the penalty model and prove its convergence. In particular, we propose a custom-built successive upper-bound minimization (BSUM) algorithm for solving the ALM subproblem at each iteration. The proposed BSUM algorithm admits closed-form updates, making the inexact ALM algorithm computationally efficient and suitable for solving large-scale problems.

We provide extensive simulation results to verify the effectiveness of the proposed approach. For the one-bit case, the proposed approach demonstrates significantly superior performance compared to existing state-of-the-art (SOTA) algorithms [4], [37]. In particular, with the same communication symbol error rate (SER), the beampattern MSE achieved by the proposed approach is generally more than three times lower than those achieved by the SOTA algorithms. We also examine the impact of various system parameters, including the numbers of transmit antennas and users in the system, the quantization level, and the safety margin threshold involved in the CI constraint, on the trade-off between communication and radar performances. In particular, we demonstrate that the shape and power level of the achieved beampattern are strongly influenced by the number of transmit antennas. Additionally, we show that increasing the resolution of DACs from one bit to three bits can greatly enhance the radar and communication performance, but further increases beyond three bits yield only marginal gains. This highlights the advantages of equipping massive MIMO and low-resolution DACs in DFRC systems.

## C. Organization and Notations

The rest of the paper is organized as follows. In Section II, we introduce the system model and formulate the problem. In Section III, we propose an efficient approach for solving the formulated problem. Extensive simulation results are presented in Section IV and the paper is concluded in Section V.

*Notation:* We use  $x$ ,  $\mathbf{x}$ ,  $\mathbf{X}$ , and  $\mathcal{X}$  to denote scalar, column vector, matrix, and set, respectively. For a matrix  $\mathbf{X}$ , we denote  $\mathbf{x}_n$  as its  $n$ -th column and  $x_{n,t}$  as its  $(n, t)$ -th entry. For a set  $\mathcal{X}$ ,  $\text{conv}(\mathcal{X})$  is the convex hull of  $\mathcal{X}$  and  $\mathcal{P}_{\mathcal{X}}(\cdot)$  is the projection operator onto  $\mathcal{X}$  (if  $\mathcal{X}$  is convex).  $\|\cdot\|$  denotes the  $\ell_2$  norm of the corresponding vector or the spectrum norm of the corresponding matrix. The operators  $(\cdot)^T$ ,  $(\cdot)^H$ ,  $\mathcal{R}(\cdot)$ ,  $\mathcal{I}(\cdot)$ , and  $|\cdot|$  return the transpose, the Hermitian transpose, the real part, the imaginary part, and the modular of their corresponding argument, respectively. The notation  $\langle \mathbf{x}, \mathbf{y} \rangle$  is also

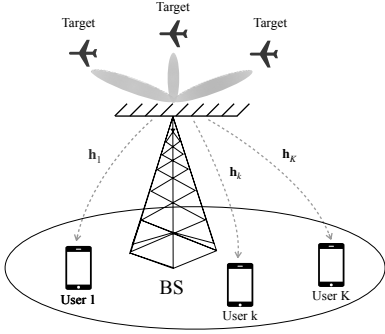


Fig. 1: A MIMO DFRC system, where there are one BS equipped with  $N$  transmit antennas, each employed with a pair of low-resolution DACs,  $K$  communication users, and many targets.

used in some contexts to denote the inner product of vectors  $\mathbf{x}$  and  $\mathbf{y}$  for clarity, as an alternative to  $\mathbf{x}^T \mathbf{y}$ . The symbol  $\mathbf{0}$  represents the all-zero vector with appropriate dimension.  $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$  represents the zero-mean circularly symmetric complex Gaussian distribution with covariance matrix  $\sigma^2 \mathbf{I}$ , where  $\mathbf{I}$  denotes the identity matrix. For a positive integer  $n$ , we use  $[n]$  to denote  $\{1, 2, \dots, n\}$ . Finally,  $\otimes$  represents the Kronecker product and  $j \triangleq \sqrt{-1}$  is the imaginary unit.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

As depicted in Fig. 1, we consider a MIMO DFRC system that simultaneously sends communication symbols to  $K$  single-antenna users and transmits probing waveform to the targets for sensing purposes. The system is equipped with a uniform linear array (ULA) with  $N$  transmit antennas, each employed with a pair of low-resolution DACs to enhance the hardware efficiency.

Let  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T] \in \mathbb{C}^{N \times T}$  be the transmit signal matrix, where  $T$  is the block length and  $\mathbf{x}_t$  is the transmit signal vector at the  $t$ -th time slot. Typically, the transmit signal matrix should satisfy the CE constraint to achieve the highest power efficiency of PAs [4], [38], i.e.,

$$|x_{t,n}| = \eta, \quad \forall n \in [N], \quad \forall t \in [T],$$

where  $\eta = \sqrt{P/N}$  with  $P$  being the maximum available power at each time slot. As low-resolution DACs are employed, each transmit signal is further restricted to be selected from only a finite set, i.e.,

$$x_{t,n} \in \mathcal{X}_L := \left\{ \eta e^{j \frac{(2\ell-1)\pi}{L}}, \ell \in [L] \right\}, \quad \forall n \in [N], \quad \forall t \in [T].$$

By convention, we refer to  $\mathcal{X}_L$  as the QCE set (as it is a quantized version of the CE signal set) and  $L$  as the number of quantization levels. Clearly, an  $L$ -level quantization can be realized by a pair of  $(\log_2 L - 1)$ -bit DACs [15]. In this paper, the transmit signal matrix  $\mathbf{X}$  is designed to optimize the radar performance with a guaranteed communication QoS requirement. Next, we will introduce the models and performance metrics for communication and radar sensing, respectively.

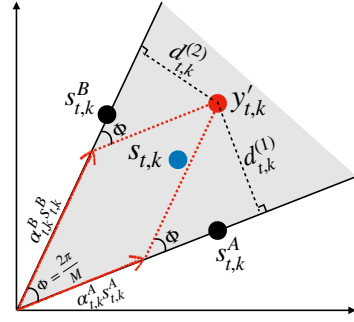


Fig. 2: An illustration of the CI metric.

### A. Communication Model and Performance Metric

The received signal matrix  $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T] \in \mathbb{C}^{K \times T}$  at the communication users is given by

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N}, \quad (1)$$

where  $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K]^T \in \mathbb{C}^{K \times N}$  is the channel matrix between the transmitter and the users, and  $\mathbf{N} = [\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_T] \in \mathbb{C}^{K \times T}$  is the additive white Gaussian noise matrix with  $\mathbf{n}_t \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$  for all  $t \in [T]$ . Let  $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_T] \in \mathbb{C}^{K \times T}$  be the data symbol matrix for the users. The transmit signal matrix  $\mathbf{X}$  is designed based on both the channel  $\mathbf{H}$  and the data symbol matrix  $\mathbf{S}$  through nonlinear precoding. This kind of precoding scheme can achieve significantly better performance than conventional linear precoding schemes designed solely based on  $\mathbf{H}$ , especially when low-resolution DACs are employed [13].

In this paper, we adopt the CI metric as the communication performance metric and focus on the  $M$ -ary phase shift keying (PSK) constellation<sup>1</sup>. CI refers to the interference that aligns constructively with the signal of interest, thereby enhancing the useful signal power and pushing the signal away from its corresponding decision boundary [22]. Exploiting the idea of CI, the CI metric aims to maximize the distance between the received noise-free signal and its closest decision boundary of the data symbol, a crucial quantity known as the safety margin [29]. The safety margin is closely related to the symbol error probability (SEP). Intuitively, with a larger safety margin, the received noise-free signal would be less susceptible to the additive noise, resulting in a lower SEP. More specifically, it has been shown in [14] that the SEP can be both lower and upper bounded by a function of the safety margin:

$$Q\left(\frac{\sqrt{2}d}{\sigma}\right) \leq \text{SEP} \leq 2Q\left(\frac{\sqrt{2}d}{\sigma}\right), \quad (2)$$

where  $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{1}{2}x^2} dx$ ,  $d$  is the safety margin, and  $\sigma$  is the standard variance of the additive noise. Therefore, the CI-based communication QoS constraint can be formulated as the form  $d \geq b$ , where  $b$  is the lower bound of the safety margin determined by the prespecified SEP threshold.

<sup>1</sup>We focus on PSK constellation for ease of presentation, but the results in this paper can be straightforwardly generalized to other constellation schemes, e.g., quadrature amplitude modulation (QAM).

Next, we derive the CI-based communication QoS constraint for model (1). For illustration, we depict a piece of the decision region of 8-PSK constellation in Fig. 2, where  $s_{t,k}$  and  $y'_{t,k} := \mathbf{h}_k^\top \mathbf{x}_t$  denote the data symbol and the noise-free received signal of user  $k$  at the  $t$ -th time slot, respectively. Clearly, the safety margin, i.e., the distance from  $y'_{t,k}$  to its closest decision boundary of  $s_{t,k}$ , is given by  $\min\{d_{t,k}^{(1)}, d_{t,k}^{(2)}\}$ . Therefore, the block-level CI-based communication QoS constraint can be formulated as

$$\min\{d_{t,k}^{(1)}, d_{t,k}^{(2)}\} \geq b_{t,k}, \quad \forall k \in [K], \quad \forall t \in [T], \quad (3)$$

where  $b_{t,k} > 0$  is the preset safety margin threshold for the  $k$ -th user at the  $t$ -th time slot.

We still need to derive the explicit formula of  $\min\{d_{t,k}^{(1)}, d_{t,k}^{(2)}\}$ . Following [29], we decompose  $y'_{t,k}$  along the decision boundaries as

$$y'_{t,k} = \alpha_{t,k}^A s_{t,k}^A + \alpha_{t,k}^B s_{t,k}^B, \quad (4)$$

where  $s_{t,k}^A = s_{t,k} e^{-j\frac{\pi}{M}}$  and  $s_{t,k}^B = s_{t,k} e^{j\frac{\pi}{M}}$  are the unit vectors on the decision boundaries, as shown in Fig. 2. Then, we have

$$\min\{d_{t,k}^{(1)}, d_{t,k}^{(2)}\} = \min\{\alpha_{t,k}^A \sin \Phi, \alpha_{t,k}^B \sin \Phi\},$$

where  $\Phi = \frac{2\pi}{M}$ . By further rewriting (4) into the real space and noting that  $y'_{t,k} = \mathbf{h}_k^\top \mathbf{x}_t$ , we can express  $\alpha_{t,k}^A$  and  $\alpha_{t,k}^B$  as

$$\begin{aligned} \begin{pmatrix} \alpha_{t,k}^A \\ \alpha_{t,k}^B \end{pmatrix} &= \frac{1}{\sin \Phi} \begin{pmatrix} \mathcal{I}(s_{t,k}^B) & -\mathcal{R}(s_{t,k}^B) \\ -\mathcal{I}(s_{t,k}^A) & \mathcal{R}(s_{t,k}^A) \end{pmatrix} \begin{pmatrix} \mathcal{R}(\mathbf{h}_k^\top) & -\mathcal{I}(\mathbf{h}_k^\top) \\ \mathcal{I}(\mathbf{h}_k^\top) & \mathcal{R}(\mathbf{h}_k^\top) \end{pmatrix} \begin{pmatrix} \mathcal{R}(\mathbf{x}_t) \\ \mathcal{I}(\mathbf{x}_t) \end{pmatrix} \\ &:= \frac{1}{\sin \Phi} \begin{pmatrix} \mathbf{c}_{t,2k-1}^\top \\ \mathbf{c}_{t,2k}^\top \end{pmatrix} \begin{pmatrix} \mathcal{R}(\mathbf{x}_t) \\ \mathcal{I}(\mathbf{x}_t) \end{pmatrix}, \end{aligned} \quad (5)$$

where we have used the fact that  $\mathcal{R}(s_{t,k}^A)\mathcal{I}(s_{t,k}^B) - \mathcal{R}(s_{t,k}^B)\mathcal{I}(s_{t,k}^A) = \sin(\arg(s_{t,k}^B) - \arg(s_{t,k}^A)) = \sin \Phi$ ; see [29] for a detailed derivation. Let  $\mathbf{x}_t^\mathcal{R} = [\mathcal{R}(\mathbf{x}_t)^\top \quad \mathcal{I}(\mathbf{x}_t)^\top]^\top \in \mathbb{R}^{2N}$  be the real space representation of  $\mathbf{x}_t$ ,  $\mathbf{C}_t = [\mathbf{c}_{t,1}, \mathbf{c}_{t,2}, \dots, \mathbf{c}_{t,2K-1}, \mathbf{c}_{t,2K}]^\top \in \mathbb{R}^{2K \times 2N}$  with  $\mathbf{c}_{t,k}$  defined in (5), and  $\mathbf{b}_t = [b_{t,1}, b_{t,2}, \dots, b_{t,K}]^\top \otimes [1, 1]^\top \in \mathbb{R}^{2K}$ . Then the constraint in (3) can be expressed as

$$\mathbf{C}_t \mathbf{x}_t^\mathcal{R} \geq \mathbf{b}_t, \quad \forall t \in [T].$$

### B. Radar Model and Performance Metric

Under the assumption that the transmitted probing waveform  $\mathbf{X}$  is narrow-band and that the propagation is line of sight (LoS), the baseband signal at angular  $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$  at the  $t$ -th time slot is

$$r_t(\theta) = \mathbf{a}(\theta)^\mathbf{H} \mathbf{x}_t,$$

where  $\mathbf{a}(\theta) := [1, e^{j\pi \sin \theta}, \dots, e^{j\pi(N-1) \sin \theta}]^\top$  is the steering vector and the antenna spacing is assumed to be half of the wavelength [41]. The beampattern, which describes the power distribution of the probing waveform in the spacial domain, is given by

$$P(\theta) = \frac{1}{T} \sum_{t=1}^T |r_t(\theta)|^2 = \frac{1}{T} \mathbf{a}(\theta)^\mathbf{H} \mathbf{X} \mathbf{X}^\mathbf{H} \mathbf{a}(\theta).$$

For radar applications such as detection and tracking, the probing waveform should exhibit strong power in the directions of potential targets while keeping power levels minimal in other directions. To achieve this goal, a common design criterion is to match the designed beampattern  $P(\theta)$  with a desired beampattern  $d(\theta)$ . Specifically, we adopt the MSE between the designed and desired beampatterns as the radar metric:

$$\frac{1}{Q} \sum_{q=1}^Q \left| \alpha d(\theta_q) - \frac{1}{T} \mathbf{a}(\theta_q)^\mathbf{H} \mathbf{X} \mathbf{X}^\mathbf{H} \mathbf{a}(\theta_q) \right|^2,$$

where  $\alpha > 0$  is a scaling factor and  $\{\theta_q\}_{q=1}^Q$  are the sampled angle grids.

### C. Problem Formulation

Based on the above discussions, the QCE transmit waveform design problem, aiming to minimize the MSE between the designed and desired beampatterns under the CI-based communication QoS constraint, can be formulated as follows:

$$\begin{aligned} \min_{\alpha > 0, \mathbf{X}} \quad & \frac{1}{Q} \sum_{q=1}^Q \left| \alpha d(\theta_q) - \frac{1}{T} \mathbf{a}(\theta_q)^\mathbf{H} \mathbf{X} \mathbf{X}^\mathbf{H} \mathbf{a}(\theta_q) \right|^2 \\ \text{s.t.} \quad & \mathbf{C}_t \mathbf{x}_t^\mathcal{R} \geq \mathbf{b}_t, \quad \forall t \in [T], \\ & x_{t,n} \in \mathcal{X}_L, \quad \forall n \in [N], \quad \forall t \in [T], \end{aligned} \quad (6)$$

where we recall that  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T]$  and  $\mathbf{x}_t^\mathcal{R}$  is the real-space representation of  $\mathbf{x}_t$ . To solve (6), we first note that for a given  $\mathbf{X}$ , the problem is quadratic in  $\alpha$ , which admits a closed-form solution as

$$\alpha^*(\mathbf{X}) = \frac{\frac{1}{T} \sum_{q=1}^Q d(\theta_q) \mathbf{a}(\theta_q)^\mathbf{H} \mathbf{X} \mathbf{X}^\mathbf{H} \mathbf{a}(\theta_q)}{\sum_{q=1}^Q d^2(\theta_q)}.$$

Substituting  $\alpha^*(\mathbf{X})$  into the objective function, problem (6) is transformed into the following equivalent form (where we have omitted a constant scaling factor  $\frac{1}{Q^2 T^2}$  in the objective function):

$$\begin{aligned} \min_{\{\mathbf{x}_t\}} \quad & \sum_{q=1}^Q \left( \sum_{t=1}^T |\mathbf{a}(\theta_q)^\mathbf{H} \mathbf{x}_t|^2 \right)^2 - \left( \sum_{q=1}^Q \sum_{t=1}^T c_q |\mathbf{a}(\theta_q)^\mathbf{H} \mathbf{x}_t|^2 \right)^2 \\ \text{s.t.} \quad & \mathbf{C}_t \mathbf{x}_t^\mathcal{R} \geq \mathbf{b}_t, \quad \forall t \in [T], \\ & x_{t,n} \in \mathcal{X}_L, \quad \forall n \in [N], \quad \forall t \in [T], \end{aligned} \quad (7)$$

where

$$c_q = \frac{d(\theta_q)}{\sqrt{\sum_{q=1}^Q d^2(\theta_q)}}, \quad \forall q \in [Q]. \quad (8)$$

By further rewriting (7) into the real space, we obtain

$$\min_{\{\mathbf{x}_t^\mathcal{R}\}} \quad \sum_{q=1}^Q \left( \sum_{t=1}^T \|\mathbf{A}_q \mathbf{x}_t^\mathcal{R}\|^2 \right)^2 - \left( \sum_{q=1}^Q \sum_{t=1}^T c_q \|\mathbf{A}_q \mathbf{x}_t^\mathcal{R}\|^2 \right)^2 \quad (9a)$$

$$\text{s.t.} \quad \mathbf{C}_t \mathbf{x}_t^\mathcal{R} \geq \mathbf{b}_t, \quad \forall t \in [T], \quad (9b)$$

$$\mathbf{x}_{t,n}^\mathcal{R} \in \mathcal{X}_L^\mathcal{R}, \quad \forall n \in [N], \quad \forall t \in [T], \quad (9c)$$

where in (9a),

$$\mathbf{A}_q = \begin{pmatrix} \mathcal{R}(\mathbf{a}(\theta_q)) & \mathcal{I}(\mathbf{a}(\theta_q)) \\ -\mathcal{I}(\mathbf{a}(\theta_q)) & \mathcal{R}(\mathbf{a}(\theta_q)) \end{pmatrix} \in \mathbb{R}^{2 \times 2N},$$

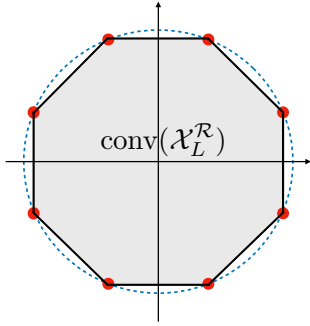


Fig. 3: An illustration of  $\mathcal{X}_L^{\mathcal{R}}$  and  $\text{conv}(\mathcal{X}_L^{\mathcal{R}})$  with  $L = 8$ , where  $\mathcal{X}_L^{\mathcal{R}}$  is the set of 8 red QCE points and its convex hull  $\text{conv}(\mathcal{X}_L^{\mathcal{R}})$  is the shadow region.

and in (9c),

$$\mathbf{x}_{t,n}^{\mathcal{R}} = \begin{bmatrix} x_{t,n}^{\mathcal{R}} \\ x_{t,n+N}^{\mathcal{R}} \end{bmatrix} \in \mathbb{R}^2, \quad \mathcal{X}_L^{\mathcal{R}} = \left\{ \eta \begin{bmatrix} \cos \frac{(2\ell-1)\pi}{L} \\ \sin \frac{(2\ell-1)\pi}{L} \end{bmatrix}, \ell \in [L] \right\}.$$

In the rest of this paper, we will focus on problem (9). From an optimization viewpoint, there are significant technical difficulties and challenges in solving problem (9). First, the discrete QCE constraint (9c) is complicated, which poses a major difficulty. Second, the linear constraints in (9b) further complicates the solution of the problem. As shown in [31], it is even NP-hard to check whether the problem is feasible with constraints (9b) and (9c). Moreover, the objective function (9a) is nonconvex and has a complicated expression. Finally, the problem size is large, especially when massive MIMO DFRC systems are considered. In the next section, we shall propose an efficient approach for solving (9) by judiciously exploiting the problem structure.

### III. PROPOSED APPROACH

In this section, we propose an efficient approach for solving problem (9). First, we employ the penalty technique to tackle the discrete QCE constraint in (9c), transforming the discrete problem into a continuous penalty model in Section III-A. Then, we propose an inexact ALM algorithm for solving the penalty model and establish its convergence in Section III-B. In particular, we propose a custom-designed BSUM algorithm for solving the ALM subproblem in Section III-C. The updates of all blocks of variables in the proposed BSUM admit closed-form solutions, making the proposed inexact ALM algorithm computationally very efficient.

#### A. Penalty Model

The penalty technique is a powerful tool to deal with complicated constraints involved in optimization problems [42]. Here, we use the penalty technique to handle the discrete QCE constraint. Specifically, we relax the QCE constraint into

its convex hull and include a negative square penalty term into the objective function as follows:

$$\begin{aligned} \min_{\{\mathbf{x}_t^{\mathcal{R}}\}} & \sum_{q=1}^Q \left( \sum_{t=1}^T \|\mathbf{A}_q \mathbf{x}_t^{\mathcal{R}}\|^2 \right)^2 \\ & - \left( \sum_{q=1}^Q \sum_{t=1}^T c_q \|\mathbf{A}_q \mathbf{x}_t^{\mathcal{R}}\|^2 \right)^2 - \lambda \sum_{t=1}^T \|\mathbf{x}_t^{\mathcal{R}}\|^2 \quad (10) \\ \text{s.t.} & \quad \mathbf{C}_t \mathbf{x}_t^{\mathcal{R}} \geq \mathbf{b}_t, \quad \forall t \in [T], \\ & \quad \mathbf{x}_{t,n}^{\mathcal{R}} \in \text{conv}(\mathcal{X}_L^{\mathcal{R}}), \quad \forall n \in [N], \quad \forall t \in [T], \end{aligned}$$

where  $\lambda$  is the penalty parameter. The idea is based on the observation that the discrete QCE points in  $\mathcal{X}_L^{\mathcal{R}}$  are the vertices of  $\text{conv}(\mathcal{X}_L^{\mathcal{R}})$ , which are the points in  $\text{conv}(\mathcal{X}_L^{\mathcal{R}})$  with the largest  $\ell_2$  norm; see Fig. 3. Hence, by promoting a larger  $\ell_2$  norm, the negative square penalty term, i.e.,  $-\lambda \sum_{t=1}^T \|\mathbf{x}_t^{\mathcal{R}}\|^2$ , encourages the solution of problem (10) to approach a discrete QCE point as the penalty parameter  $\lambda$  increases. In the following proposition, we theoretically establish the equivalence between the continuous penalty model (10) and the original discrete model (9).

**Proposition 1.** *There exists  $\lambda_0 > 0$  such that for all  $\lambda > \lambda_0$ , problems (9) and (10) share the same optimal solutions.*

*Proof.* See Appendix A.  $\square$

In practice, it is difficult to determine the explicit value of  $\lambda_0$  in Proposition 1. In our implementation, we employ a common technique in penalty-based approaches known as the homotopy technique [43], which initializes the penalty parameter  $\lambda$  with a small value and gradually increases it, tracking the solution path of the corresponding penalty models. Through this process, we will eventually find the required  $\lambda$  in Proposition 1. Moreover, this technique can greatly enhance the numerical performance of the penalty approaches [17], [31], compared to directly solving the penalty model (10) with a very large  $\lambda$ .

We remark here that similar penalty strategies have been employed to deal with the QCE constraint in communication-only QCE precoding design [17], [19]. However, the optimization problems in [17], [19] do not include any constraints other than the QCE constraint, while our problem involves a large number of linear constraints. These linear constraints exclude many QCE points from the feasible set and destroy the symmetry of set  $\text{conv}(\mathcal{X}_L^{\mathcal{R}})$ , making the feasible region of problem (10) more complicated and the equivalence between the original and penalty problems more challenging to be established.

Up to now, we have transformed the original discrete model (9) into an equivalent continuous penalty model (10), which is more amenable to the algorithmic design. However, problem (10) is still a nonconvex problem with a complicated objective function and many linear constraints. In the following subsection, we propose an efficient inexact ALM algorithm for solving problem (10).

### B. Inexact ALM Algorithm for Solving (10)

In this subsection, we propose an inexact ALM algorithm for solving the penalty model (10). In Section III-B1, we first reformulate problem (10) into a form that facilitates the use of the ALM framework. Then, we present the proposed inexact ALM algorithm and prove its convergence in Sections III-B2 and III-B3, respectively.

1) *Reformulation of (10)*: For problem (10), we introduce auxiliary variables  $\mathbf{w}_{t,q} = \mathbf{A}_q \mathbf{x}_t^{\mathcal{R}} \in \mathbb{R}^2$  to simplify its objective function and introduce auxiliary variables  $\mathbf{z}_t \geq \mathbf{0}$  to transform its inequality linear constraints into equality ones as follows:

$$\begin{aligned} \min_{\{\mathbf{x}_t^{\mathcal{R}}\}, \{\mathbf{w}_t\}, \{\mathbf{z}_t\}} & \sum_{q=1}^Q \left( \sum_{t=1}^T \|\mathbf{w}_{t,q}\|^2 \right)^2 \\ & - \left( \sum_{q=1}^Q \sum_{t=1}^T c_q \|\mathbf{w}_{t,q}\|^2 \right)^2 - \lambda \sum_{t=1}^T \|\mathbf{x}_t^{\mathcal{R}}\|^2 \quad (11) \\ \text{s.t.} & \quad \mathbf{C}_t \mathbf{x}_t^{\mathcal{R}} - \mathbf{z}_t = \mathbf{b}_t, \quad \mathbf{z}_t \geq \mathbf{0}, \quad \forall t \in [T], \\ & \quad \tilde{\mathbf{A}} \mathbf{x}_t^{\mathcal{R}} - \mathbf{w}_t = \mathbf{0}, \quad \forall t \in [T], \\ & \quad \mathbf{x}_{t,n}^{\mathcal{R}} \in \text{conv}(\mathcal{X}_L^{\mathcal{R}}), \quad \forall n \in [N], \forall t \in [T], \end{aligned}$$

where  $\tilde{\mathbf{A}} = [\mathbf{A}_1^{\top}, \mathbf{A}_2^{\top}, \dots, \mathbf{A}_Q^{\top}]^{\top} \in \mathbb{R}^{2Q \times 2N}$  and  $\mathbf{w}_t = [\mathbf{w}_{t,1}^{\top}, \mathbf{w}_{t,2}^{\top}, \dots, \mathbf{w}_{t,Q}^{\top}]^{\top} \in \mathbb{R}^{2Q \times 1}$ . Compared to (10), the ALM subproblem of (11) is more tractable, which will become more clear in Section III-C.

For ease of presentation, we further express problem (11) into the following more compact form:

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{w}, \mathbf{z}} & f(\mathbf{w}) + g(\mathbf{w}) + h(\mathbf{x}) \\ \text{s.t.} & \quad \mathbf{C}\mathbf{x} - \mathbf{z} = \mathbf{b}, \\ & \quad \mathbf{A}\mathbf{x} - \mathbf{w} = \mathbf{0}, \\ & \quad \mathbf{x} \in \mathcal{X}, \quad \mathbf{z} \in \mathcal{Z}. \end{aligned} \quad (12)$$

where  $\mathbf{x} = [(\mathbf{x}_1^{\mathcal{R}})^{\top}, (\mathbf{x}_2^{\mathcal{R}})^{\top}, \dots, (\mathbf{x}_T^{\mathcal{R}})^{\top}]^{\top}$ ,  $\mathbf{w} = [\mathbf{w}_1^{\top}, \mathbf{w}_2^{\top}, \dots, \mathbf{w}_T^{\top}]^{\top}$ ,  $\mathbf{z} = [\mathbf{z}_1^{\top}, \mathbf{z}_2^{\top}, \dots, \mathbf{z}_T^{\top}]^{\top}$ ,  $\mathbf{b} = [\mathbf{b}_1^{\top}, \mathbf{b}_2^{\top}, \dots, \mathbf{b}_T^{\top}]^{\top}$ ,  $\mathbf{C} = \text{diag}(\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_T)$ ,  $\mathbf{A} = \mathbf{I}_T \otimes \tilde{\mathbf{A}}$ ,  $\mathcal{X} = \{\mathbf{x} \mid \mathbf{x}_{t,n}^{\mathcal{R}} \in \text{conv}(\mathcal{X}_L^{\mathcal{R}}), \forall n \in [N], \forall t \in [T]\}$ ,  $\mathcal{Z} = \{\mathbf{z} \mid \mathbf{z} \geq \mathbf{0}\}$ , and  $f(\mathbf{w})$ ,  $g(\mathbf{w})$ , and  $h(\mathbf{x})$  are the first, second, and third terms of the objective function in (11), respectively.

2) *Inexact ALM Algorithm*: The ALM [43] is a popular and powerful method for solving (equality) constrained optimization problems like (12). It transforms the complicated constrained problem into a sequence of relatively easy unconstrained subproblems or subproblems with easy constraints (e.g., nonnegative constraints) via the augmented Lagrangian function. Then, by iteratively solving the ALM subproblems and updating the Lagrange multipliers and penalty parameters, the algorithm converges to a (stationary) solution of the original constrained problem. Next, we propose an inexact ALM algorithm for solving problem (12), which *inexactly* solves the ALM subproblem at each iteration in order to reduce the computational cost.

The augmented Lagrangian function of (12) is given by

$$\begin{aligned} \mathcal{L}_{\rho}(\mathbf{x}, \mathbf{w}, \mathbf{z}; \boldsymbol{\mu}, \boldsymbol{\nu}) &= f(\mathbf{w}) + g(\mathbf{w}) + h(\mathbf{x}) \\ &+ \boldsymbol{\mu}^{\top}(\mathbf{C}\mathbf{x} - \mathbf{z} - \mathbf{b}) + \boldsymbol{\nu}^{\top}(\mathbf{A}\mathbf{x} - \mathbf{w}) \\ &+ \frac{\rho_{\mu}}{2} \|\mathbf{C}\mathbf{x} - \mathbf{z} - \mathbf{b}\|^2 + \frac{\rho_{\nu}}{2} \|\mathbf{A}\mathbf{x} - \mathbf{w}\|^2, \end{aligned} \quad (13)$$

where  $(\boldsymbol{\mu}, \boldsymbol{\nu})$  are the Lagrange multipliers corresponding to the first and second equality constraints in (12) and  $\boldsymbol{\rho} := (\rho_{\mu}, \rho_{\nu})$  are the corresponding penalty parameters. The proposed inexact ALM algorithm iteratively performs the following three steps.

**Step 1: Inexactly solve the ALM subproblem over primal variables  $(\mathbf{x}, \mathbf{w}, \mathbf{z})$ .** In the classical ALM algorithm, one needs to solve the following ALM subproblem

$$(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1}) \in \arg \min_{\mathbf{x} \in \mathcal{X}, \mathbf{w}, \mathbf{z} \in \mathcal{Z}} \mathcal{L}_{\rho^m}(\mathbf{x}, \mathbf{w}, \mathbf{z}; \boldsymbol{\mu}^m, \boldsymbol{\nu}^m) \quad (14)$$

at each iteration [43]. To avoid the high computational cost of exactly solving (14), the proposed inexact ALM algorithm only requires (14) to be solved to an  $\epsilon_m$ -stationary point. Here  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})$  is called an  $\epsilon_m$ -stationary point of the ALM subproblem if

$$\begin{aligned} \text{dist}(\nabla \mathcal{L}_{\rho^m}(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1}; \boldsymbol{\mu}^m, \boldsymbol{\nu}^m) \\ + \partial \mathbb{I}_{\mathcal{X}}(\mathbf{x}^{m+1}) + \partial \mathbb{I}_{\mathcal{Z}}(\mathbf{z}^{m+1}), \mathbf{0}) \leq \epsilon_m, \end{aligned} \quad (15)$$

where  $\mathbb{I}_{\mathcal{X}}$  and  $\mathbb{I}_{\mathcal{Z}}$  denote the indicator functions of  $\mathcal{X}$  and  $\mathcal{Z}$ , respectively, and  $\partial \mathbb{I}_{\mathcal{X}}$  and  $\partial \mathbb{I}_{\mathcal{Z}}$  denote their subdifferentials, respectively. When  $\epsilon_m = 0$  in (15), then  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})$  is a stationary point of problem (14). To guarantee the convergence of the inexact ALM algorithm, the accuracy of solving the ALM subproblem should be gradually increased.

In our proposed algorithm, the ALM subproblem (14) is (approximately) solved by a custom-built BSUM algorithm, which admits closed-form updates and is guaranteed to find an  $\epsilon_m$ -stationary point within a finite number of iterations. To maintain coherence in presenting the proposed inexact ALM algorithm, we defer detailed discussions of the BSUM algorithm for solving (14) to Section III-C.

**Step 2: Update the Lagrange multipliers  $(\boldsymbol{\mu}, \boldsymbol{\nu})$ .** After obtaining  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})$ , the Lagrange multipliers  $(\boldsymbol{\mu}, \boldsymbol{\nu})$  are updated as

$$\begin{aligned} \boldsymbol{\mu}^{m+1} &= \min\{\max\{\boldsymbol{\mu}^m + \rho_{\mu}^m(\mathbf{C}\mathbf{x}^{m+1} - \mathbf{z}^{m+1} - \mathbf{b}), \boldsymbol{\mu}_{\min}\}, \boldsymbol{\mu}_{\max}\}, \\ \boldsymbol{\nu}^{m+1} &= \min\{\max\{\boldsymbol{\nu}^m + \rho_{\nu}^m(\mathbf{A}\mathbf{x}^{m+1} - \mathbf{w}^{m+1}), \boldsymbol{\nu}_{\min}\}, \boldsymbol{\nu}_{\max}\}, \end{aligned} \quad (16)$$

where  $(\boldsymbol{\mu}_{\min}, \boldsymbol{\nu}_{\min})$  and  $(\boldsymbol{\mu}_{\max}, \boldsymbol{\nu}_{\max})$  are the imposed lower and upper bounds on the Lagrange multipliers  $(\boldsymbol{\mu}, \boldsymbol{\nu})$ , respectively, and the operators  $\min$  and  $\max$  both operate component-wise on their entries. Different from the classical ALM, we apply safeguards in our method to ensure that the Lagrange multipliers are bounded, which is crucial for the convergence of the algorithm [44].

**Step 3: Update the penalty parameters  $\boldsymbol{\rho} = (\rho_{\mu}, \rho_{\nu})$ .** Finally, we update the penalty parameters  $\boldsymbol{\rho} = (\rho_{\mu}, \rho_{\nu})$  as in (17), which is given on the top of the next page, where  $\tau > 1$  and  $\delta > 0$ . The criterion in (17) increases the penalty parameters if the violation of the constraints is not sufficiently

$$\rho^{m+1} = \begin{cases} \tau \rho^m, & \text{if } \left\| \frac{\sqrt{\rho_\mu^m}(\mathbf{C}\mathbf{x}^{m+1} - \mathbf{z}^{m+1} - \mathbf{b})}{\sqrt{\rho_\nu^m}(\mathbf{A}\mathbf{x}^{m+1} - \mathbf{w}^{m+1})} \right\| \geq \delta \left\| \frac{\sqrt{\rho_\mu^m}(\mathbf{C}\mathbf{x}^m - \mathbf{z}^m - \mathbf{b})}{\sqrt{\rho_\nu^m}(\mathbf{A}\mathbf{x}^m - \mathbf{w}^m)} \right\|; \\ \rho^m, & \text{otherwise.} \end{cases} \quad (17)$$

---

**Algorithm 1** Inexact ALM algorithm for solving (12)

---

- 1: **Input:**  $(\mathbf{x}^0, \mathbf{w}^0, \mathbf{z}^0), (\boldsymbol{\mu}^0, \boldsymbol{\nu}^0), \boldsymbol{\rho}^0 = (\rho_\mu^0, \rho_\nu^0), -\infty < \mu_{\min} < \mu_{\max} < \infty, -\infty < \nu_{\min} < \nu_{\max} < \infty, \tau > 1, \delta > 0$ , a positive sequence  $\epsilon_m \rightarrow 0$ .
  - 2: **Initialize:**  $m = 0$ .
  - 3: **repeat**
  - 4:   *(Inexactly solve the ALM subproblem):* Obtain an  $\epsilon_m$ -stationary point  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})$  of (14) by Algorithm 2 (see further ahead).
  - 5:   *(Update the Lagrange multipliers):* Obtain  $(\boldsymbol{\mu}^{m+1}, \boldsymbol{\nu}^{m+1})$  by (16).
  - 6:   *(Update the penalty parameter):* Obtain  $\boldsymbol{\rho}^{m+1}$  by (17).
  - 7:   Set  $m = m + 1$ .
  - 8: **until** some stopping criterion is satisfied.
  - 9: **Output:**  $(\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m)$ .
- 

reduced and keep them fixed otherwise. We remark that the penalty parameters  $\rho_\mu$  and  $\rho_\nu$  should be increased at the same rate as in (17) to ensure the convergence of the algorithm.

We summarize the proposed inexact ALM algorithm for solving problem (12) in Algorithm 1.

3) *Convergence Analysis:* Before showing the convergence guarantee of the proposed inexact ALM algorithm, we first give the formal definition of the stationary point of problem (12) [43, Proposition 2.1.2].

**Definition 1** (Stationary point of (12)). *A point  $(\hat{\mathbf{x}}, \hat{\mathbf{w}}, \hat{\mathbf{z}})$  is called a stationary point of (12) if it satisfies  $(\hat{\mathbf{x}}, \hat{\mathbf{w}}, \hat{\mathbf{z}}) \in \mathcal{F}$  and*

$$\begin{aligned} (\nabla f(\mathbf{w}) + \nabla g(\mathbf{w}))^\top (\mathbf{w} - \hat{\mathbf{w}}) + \nabla h(\mathbf{x})^\top (\mathbf{x} - \hat{\mathbf{x}}) &\geq 0, \\ \forall (\mathbf{x}, \mathbf{w}) \in \{(\mathbf{x}, \mathbf{w}) \mid \exists \mathbf{z}, (\mathbf{x}, \mathbf{w}, \mathbf{z}) \in \mathcal{F}\}, \end{aligned} \quad (18)$$

where

$$\mathcal{F} := \{(\mathbf{x}, \mathbf{w}, \mathbf{z}) \mid \mathbf{C}\mathbf{x} - \mathbf{z} = \mathbf{b}, \mathbf{A}\mathbf{x} - \mathbf{w} = \mathbf{0}, \mathbf{x} \in \mathcal{X}, \mathbf{z} \in \mathcal{Z}\}$$

is the feasible region of (12).

Now we present the convergence result of Algorithm 1 in the following Theorem 1.

**Theorem 1.** *Any limit point  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}})$  of  $\{(\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m)\}$  generated by Algorithm 1 is a stationary point of (12). Moreover,*

$$\lim_{m \rightarrow \infty} \text{dist}((\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m), \mathcal{S}^*) = 0,$$

where  $\mathcal{S}^*$  is the set of stationary points of (12).

*Proof.* See Appendix C.  $\square$

We remark that the above result highly depends on the boundedness of the sequences  $\{(\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m)\}$  and  $\{(\boldsymbol{\mu}^m, \boldsymbol{\nu}^m)\}$  generated by Algorithm 1. In sharp contrast to classical convergence results for the ALM where the boundedness is typically imposed as an uncheckable assumption [43], in the proposed inexact ALM algorithm, the boundedness of the Lagrange multipliers is naturally ensured by the update

rule in (16), and the boundedness of the primal variables is guaranteed by the specialized BSUM algorithm for solving the ALM subproblem; see Section III-C and Appendix B for detailed discussions.

It is worth mentioning that a similar inexact ALM framework has been proposed in [44]. However, the problem considered in [44] is much simpler than (12). In addition, the condition that guarantees the convergence in [44] is more restrictive than that in the proposed inexact ALM algorithm. In particular, in [44],  $\epsilon_m$  is required to satisfy  $\epsilon_m \rho^m \rightarrow 0$  (where  $\rho^m$  is the corresponding penalty parameter), while we only require  $\epsilon_m \rightarrow 0$ .

### C. BSUM Algorithm for Solving Subproblem (14)

The efficiency of the proposed inexact ALM algorithm is mainly determined by the efficiency of solving the ALM subproblem (14) to obtain an  $\epsilon_m$ -stationary point. In this subsection, we propose an efficient BSUM algorithm for solving (14) by judiciously exploiting the problem's special structure. Since  $(\boldsymbol{\mu}^m, \boldsymbol{\nu}^m)$  are treated as fixed parameters when solving (14), we denote

$$\mathcal{L}_m(\mathbf{x}, \mathbf{w}, \mathbf{z}) := \mathcal{L}_{\rho^m}(\mathbf{x}, \mathbf{w}, \mathbf{z}; \boldsymbol{\mu}^m, \boldsymbol{\nu}^m) \quad (19)$$

in this subsection for simplicity of notations.

Noting that the constraints in (14) are separable within each block of variables  $\{\mathbf{x}, \mathbf{w}, \mathbf{z}\}$ , it is convenient to update the variables in an alternating fashion to solve the problem. However, the objective function is complicated (even with respect to each block of variables) and cannot be efficiently optimized within each block. This motivates us to apply a BSUM framework, which minimizes an upper bound of the objective function for each block at each iteration. Specifically, the BSUM framework is given by

$$\mathbf{x}^{(r+1)} \in \arg \min_{\mathbf{x} \in \mathcal{X}} u_x(\mathbf{x}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}); \quad (20a)$$

$$\mathbf{w}^{(r+1)} \in \arg \min u_w(\mathbf{w}; \mathbf{x}^{(r+1)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}); \quad (20b)$$

$$\mathbf{z}^{(r+1)} \in \arg \min_{\mathbf{z} \in \mathcal{Z}} u_z(\mathbf{z}; \mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r)}), \quad (20c)$$

where  $u_x, u_w$ , and  $u_z$  are locally tight upper bounds of  $\mathcal{L}_m$  in (19), i.e.,  $u_x(\mathbf{x}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) \geq \mathcal{L}_m(\mathbf{x}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  for all  $\mathbf{x} \in \mathcal{X}$  with equality holding at  $\mathbf{x} = \mathbf{x}^{(r)}$  (similar properties hold for  $u_w$  and  $u_z$ ). Please see [45] for more detailed discussions on BSUM. Here we use  $\{(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})\}$  to denote the inner sequence generated by the BSUM framework for solving the ALM subproblem, which is different from the outer sequence  $\{(\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m)\}$  generated by the inexact ALM.

In the following, we tailor a BSUM algorithm for solving problem (14). The key is to design the upper bounds  $u_x, u_w, u_z$  appropriately such that they are good approximations of the objective function  $\mathcal{L}_m$  and at the same time the subproblems (20a)–(20c) admit closed-form solutions or can be efficiently



solved. Next we discuss the updates of the  $\mathbf{x}$ -,  $\mathbf{w}$ -, and  $\mathbf{z}$ -blocks separately.

**Update of variable  $\mathbf{x}$ .** The difficulty in updating variable  $\mathbf{x}$  mainly arises from the complicated constraint  $\mathbf{x} \in \mathcal{X}$  defined below (12). Notably, the set  $\mathcal{X}$  is decoupled over  $\{\mathbf{x}_{i,n}^{\mathcal{R}}\}$ . To leverage this property, we construct a separable upper bound function of  $\mathcal{L}_m(\mathbf{x}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  as follows:

$$u_x(\mathbf{x}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) = \mathcal{L}_m(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) + \nabla_{\mathbf{x}} \mathcal{L}_m(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})^{\top} (\mathbf{x} - \mathbf{x}^{(r)}) + \frac{\gamma^m}{2} \|\mathbf{x} - \mathbf{x}^{(r)}\|^2. \quad (21)$$

The above  $u_x(\mathbf{x}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  is exactly the first-order Taylor expansion of  $\mathcal{L}_m(\mathbf{x}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  at  $\mathbf{x} = \mathbf{x}^{(r)}$  plus a quadratic regularization term, which equals  $\mathcal{L}_m(\mathbf{x}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  at  $\mathbf{x} = \mathbf{x}^{(r)}$ . To ensure that it serves as an upper bound of  $\mathcal{L}_m(\mathbf{x}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$ , we set  $\gamma^m = \|\rho_{\mu}^m \mathbf{C}^{\top} \mathbf{C} + \rho_{\nu}^m \mathbf{A}^{\top} \mathbf{A}\|$ .

With  $u_x(\mathbf{x}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  defined in (21) and by ignoring the constant terms, the  $\mathbf{x}$ -subproblem can be expressed as

$$\min_{\mathbf{x} \in \mathcal{X}} \nabla_{\mathbf{x}} \mathcal{L}_m(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})^{\top} \mathbf{x} + \frac{\gamma^m}{2} \|\mathbf{x} - \mathbf{x}^{(r)}\|^2, \quad (22)$$

whose solution is given by

$$\mathbf{x}^{(r+1)} = \mathcal{P}_{\mathcal{X}} \left( \mathbf{x}^{(r)} - \frac{\nabla \mathcal{L}_m(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})}{\gamma^m} \right). \quad (23)$$

According to the definition of  $\mathcal{X}$ , the projection onto  $\mathcal{X}$  can be decomposed into  $NT$  of projections onto  $\text{conv}(\mathcal{X}_L^{\mathcal{R}})$ , which admit closed-form expressions [17].

**Update of variable  $\mathbf{w}$ .** The functions  $f(\mathbf{w})$  and  $g(\mathbf{w})$  involved in the ALM subproblem are both quartic in  $\mathbf{w}$ , making the  $\mathbf{w}$ -update challenging. In particular,  $g(\mathbf{w})$  is concave and couples all the components in  $\mathbf{w}$ . To simplify the objective function while best preserving its structure, we linearize the concave quartic term  $g(\mathbf{w})$  and keep all the convex terms unchanged, which gives the following locally tight upper bound of  $\mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}, \mathbf{z}^{(r)})$ :

$$\begin{aligned} u_w(\mathbf{w}; \mathbf{x}^{(r+1)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) &= f(\mathbf{w}) + g(\mathbf{w}^{(r)}) + \nabla g(\mathbf{w}^{(r)})^{\top} (\mathbf{w} - \mathbf{w}^{(r)}) + h(\mathbf{x}^{(r+1)}) \\ &\quad + (\boldsymbol{\mu}^m)^{\top} (\mathbf{C}\mathbf{x}^{(r+1)} - \mathbf{z}^{(r)} - \mathbf{b}) + (\boldsymbol{\nu}^m)^{\top} (\mathbf{A}\mathbf{x}^{(r+1)} - \mathbf{w}) \\ &\quad + \frac{\rho_{\mu}^m}{2} \|\mathbf{C}\mathbf{x}^{(r+1)} - \mathbf{z}^{(r)} - \mathbf{b}\|^2 + \frac{\rho_{\nu}^m}{2} \|\mathbf{A}\mathbf{x}^{(r+1)} - \mathbf{w}\|^2. \end{aligned}$$

Then, by omitting the constant terms, the  $\mathbf{w}$ -subproblem reads

$$\min_{\mathbf{w}} f(\mathbf{w}) + (\nabla g(\mathbf{w}^{(r)}) - \boldsymbol{\nu}^m)^{\top} \mathbf{w} + \frac{\rho_{\nu}^m}{2} \|\mathbf{A}\mathbf{x}^{(r+1)} - \mathbf{w}\|^2. \quad (24)$$

Note that in the objective function of (24),  $f(\mathbf{w})$ , i.e., the first term of the objective function of (11), is separable over the index  $q \in [Q]$ , and both of the last two terms are fully separable. Hence, the problem can be decomposed into  $Q$  individual subproblems.

Let  $\mathbf{w}_{(q)} := [\mathbf{w}_{1,q}^{\top}, \mathbf{w}_{2,q}^{\top}, \dots, \mathbf{w}_{T,q}^{\top}]^{\top}$  denote the vector collecting all the components in  $\mathbf{w}$  related to  $q$  and let  $\boldsymbol{\xi}_q^{(r)}$  be the coefficient of the linear term in (24) related to  $\mathbf{w}_{(q)}$ . Then for each  $q \in [Q]$ , the  $\mathbf{w}_{(q)}$ -subproblem has the following form:

$$\arg \min_{\mathbf{w}_{(q)}} \|\mathbf{w}_{(q)}\|^4 + \frac{\rho_{\nu}^m}{2} \|\mathbf{w}_{(q)}\|^2 + (\boldsymbol{\xi}_q^{(r)})^{\top} \mathbf{w}_{(q)}. \quad (25)$$

Since the first two terms in the objective function of (25) are irrelevant to the angle of  $\mathbf{w}_{(q)}$ , it follows that the optimal angle of  $\mathbf{w}_{(q)}$  is  $-\boldsymbol{\xi}_q^{(r)} / \|\boldsymbol{\xi}_q^{(r)}\|$ . Therefore, we only need to optimize the norm of  $\mathbf{w}_{(q)}$ , i.e., solving a univariate quartic problem as follows:

$$\beta_q^{(r+1)} \in \arg \min_{\beta \geq 0} \left\{ \beta^4 + \frac{\rho_{\nu}^m}{2} \beta^2 - \|\boldsymbol{\xi}_q^{(r)}\| \beta \right\}. \quad (26)$$

By setting the derivative of the quartic objective function in (26) to zero, we obtain the following cubic equation:

$$4\beta^3 + \rho_{\nu}^m \beta - \|\boldsymbol{\xi}_q^{(r)}\| = 0.$$

The solution of (26) is the unique positive root of the above equation, whose explicit expression is given by

$$\begin{aligned} \beta_q^{(r+1)} &= \left( \frac{\|\boldsymbol{\xi}_q^{(r)}\|}{8} + \left( \left( \frac{\|\boldsymbol{\xi}_q^{(r)}\|}{8} \right)^2 + \left( \frac{\rho_{\nu}^m}{12} \right)^3 \right)^{\frac{1}{2}} \right)^{\frac{1}{3}} \\ &\quad + \left( \frac{\|\boldsymbol{\xi}_q^{(r)}\|}{8} - \left( \left( \frac{\|\boldsymbol{\xi}_q^{(r)}\|}{8} \right)^2 + \left( \frac{\rho_{\nu}^m}{12} \right)^3 \right)^{\frac{1}{2}} \right)^{\frac{1}{3}}. \end{aligned}$$

Combining the above discussions, we obtain the optimal solution of (25):

$$\mathbf{w}_{(q)}^{(r+1)} = -\frac{\beta_q^{(r+1)}}{\|\boldsymbol{\xi}_q^{(r)}\|} \boldsymbol{\xi}_q^{(r)}, \quad \forall q \in [Q]. \quad (27)$$

**Update of variable  $\mathbf{z}$ .** The ALM subproblem with respect to  $\mathbf{z}$  is straightforward to solve. In particular, we simply set

$$u_z(\mathbf{z}; \mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r)}) = \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}).$$

The  $\mathbf{z}$ -subproblem

$$\min_{\mathbf{z} \in \mathcal{Z}} -(\boldsymbol{\mu}^m)^{\top} \mathbf{z} + \frac{\rho_{\mu}^m}{2} \|\mathbf{z} - \mathbf{C}\mathbf{x}^{(r+1)} + \mathbf{b}\|^2 \quad (28)$$

admits a closed-form solution as

$$\mathbf{z}^{(r+1)} = \mathcal{P}_{\{\mathbf{z} \geq \mathbf{0}\}} \left( \mathbf{C}\mathbf{x}^{(r+1)} - \mathbf{b} + \frac{\boldsymbol{\mu}^m}{\rho_{\mu}^m} \right). \quad (29)$$

The proposed BSUM algorithm for solving the ALM subproblem (14) is summarized in Algorithm 2. Since all variables admit closed-form updates, the BSUM algorithm can be performed efficiently. Note that in Algorithm 2, we set the initial point  $(\mathbf{x}^{(0)}, \mathbf{w}^{(0)}, \mathbf{z}^{(0)})$  as the approximate stationary point of the previous ALM subproblem. This warm-start strategy is able to accelerate the convergence of BSUM, thereby enhancing the numerical efficiency of the inexact ALM algorithm. As will be demonstrated in the simulation, Algorithm 2 equipped with the warm-start strategy can generally terminate within only a few iterations.

In the following proposition, we theoretically show that Algorithm 2 can terminate within a finite number of iterations, i.e., an  $\epsilon_m$ -stationary point of (14) can be found within a finite number of iterations. Note that according to the update rule of  $\boldsymbol{\rho}$  in (17),  $\rho_{\mu}$  and  $\rho_{\nu}$  are increased consistently, and thus we can express  $\boldsymbol{\rho}^m$  as  $\boldsymbol{\rho}^m = (a_{\mu}, a_{\nu}) \boldsymbol{\rho}^m$  with  $a_{\mu} > 0$  and  $a_{\nu} > 0$  being two constants independent of  $m$ . With this notation, the iteration complexity of Algorithm 2 is given as follows.



---

**Algorithm 2** BSUM algorithm for solving (14)
 

---

- 1: **Input:**  $(\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m)$ ,  $(\boldsymbol{\mu}^m, \boldsymbol{\nu}^m)$ ,  $\boldsymbol{\rho}^m = (\rho_\mu^m, \rho_\nu^m)$ ,  $\gamma^m = \|\rho_\mu^m \mathbf{C}^\top \mathbf{C} + \rho_\nu^m \mathbf{A}^\top \mathbf{A}\|$ ,  $\epsilon_m$ .
  - 2: **Initialize:**  $(\mathbf{x}^{(0)}, \mathbf{w}^{(0)}, \mathbf{z}^{(0)}) = (\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m)$ ,  $r = 0$ .
  - 3: **repeat**
  - 4:   Update  $\mathbf{x}^{(r+1)}$ ,  $\mathbf{w}^{(r+1)}$ , and  $\mathbf{z}^{(r+1)}$  by (23), (27), and (29), respectively.
  - 5:   Set  $r = r + 1$ .
  - 6: **until**  $(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  is an  $\epsilon_m$ -stationary point of (14).
  - 7: **Output:**  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1}) = (\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$ .
- 

**Proposition 2.** Algorithm 2 can return an  $\epsilon_m$ -stationary point of (14) within  $\mathcal{O}\left(\frac{(\rho^m)^2}{\epsilon_m^2}\right)$  iterations.

*Proof.* See Appendix D.  $\square$

#### IV. NUMERICAL RESULTS

In this section, we present simulation results to evaluate the performance of the proposed approach under various system configurations. Each system configuration is represented by a tuple  $(N, K, M, L)$ , where  $N$  and  $K$  are the numbers of transmit antennas and users in the system, respectively,  $M$  is the constellation order, and  $L$  is the quantization level. We generate the simulation parameters in the same way as in [4], [38]. Specifically, the transmit power is fixed as  $P = 1$  and the SNR is defined as  $1/\sigma^2$ . We assume a Rayleigh fading communication channel  $\mathbf{H}$ , whose elements are independently drawn from  $\mathcal{CN}(0, 1)$ . We assume that there are three interested targets located at  $\bar{\theta}_1 = -40^\circ$ ,  $\bar{\theta}_2 = 0^\circ$ , and  $\bar{\theta}_3 = 40^\circ$ , respectively, and the desired beam pattern is given by

$$d(\theta) = \begin{cases} 1, & \text{if } \theta \in \left[\bar{\theta}_i - \frac{\Delta\theta}{2}, \bar{\theta}_i + \frac{\Delta\theta}{2}\right], i = 1, 2, 3; \\ 0, & \text{otherwise,} \end{cases} \quad (30)$$

where  $\Delta\theta = 10^\circ$  is the beam width. The direction grids  $\{\theta_q\}_{q \in [Q]}$  are uniformly sampled from  $-90^\circ$  to  $90^\circ$  with a resolution of  $1^\circ$ . The safety margin thresholds for different users at different time slots are set as the same, i.e.,  $b_{t,k} = b$  for all  $t \in [T]$  and  $k \in [K]$ . The block length is set as  $T = 50$ .

The implementation details of the proposed Algorithm 1 are as follows. We set  $\rho_\mu^0 = 0.01\sqrt{\lambda}$  and  $\rho_\nu^0 = \frac{1}{3}\rho_\mu^0$ . Both  $(\mathbf{x}^0, \mathbf{w}^0, \mathbf{z}^0)$  and  $(\boldsymbol{\mu}^0, \boldsymbol{\nu}^0)$  are set as all-zero vectors. The bounds on the Lagrange multipliers are given by  $\mu_{\min} = \nu_{\min} = -10^3$ ,  $\mu_{\max} = \nu_{\max} = 10^3$ . Finally,  $\tau = 1.01$ ,  $\delta = 0.95$ , and  $\epsilon_m = \frac{1}{m}$ . Algorithm 1 is terminated when the iteration number is larger than 500 or when

$$\max\{e^{m-1}, \|\mathbf{C}\mathbf{x}^m - \mathbf{z}^m - \mathbf{b}\|, \|\mathbf{A}\mathbf{x}^m - \mathbf{w}^m\|\} \leq \sqrt{T} \times 10^{-3}, \quad (31)$$

where

$$e^{m-1} := \|\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m) - \nabla_{\mathbf{x}} \mathcal{L}_m(\mathbf{x}^{m-1}, \mathbf{w}^{m-1}, \mathbf{z}^{m-1}) - \gamma^m(\mathbf{x}^m - \mathbf{x}^{m-1})\| + \|\nabla g(\mathbf{w}^m) - \nabla g(\mathbf{w}^{m-1})\|$$

is defined to characterize the accuracy of solving the ALM subproblem (see (46) and (47)). The above stopping criterion is employed based on the fact that when  $\max\{e^{m-1}, \|\mathbf{C}\mathbf{x}^m -$

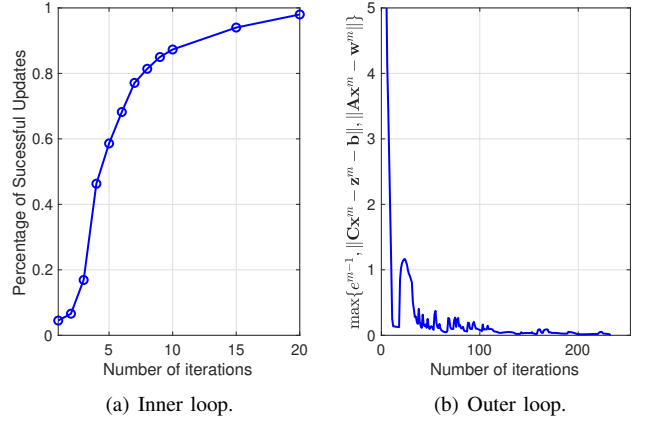


Fig. 4: Convergence behaviors of proposed Algorithms 1 and 2.

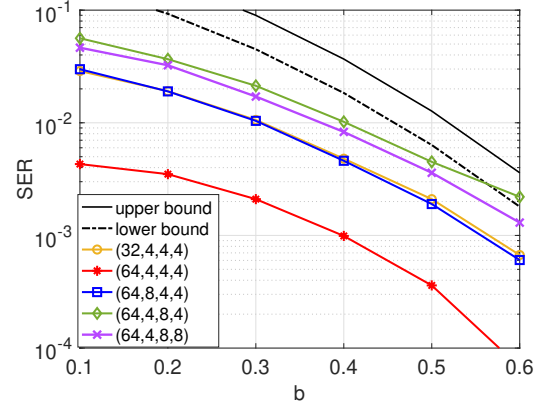


Fig. 5: Achieved SERs by the proposed approach versus the safety margin threshold  $b$  under different system configurations, where the tuple in the legend represents the system parameter  $(N, K, M, L)$  and the SNR is set as 10 dB. The curves labeled “upper bound” and “lower bound” correspond to the theoretical upper and lower bounds on the SEP in (2), respectively.

$\mathbf{z}^m - \mathbf{b}\|, \|\mathbf{A}\mathbf{x}^m - \mathbf{w}^m\|\}$  is small, the output of Algorithm 1 approaches a stationary point of the penalty model (12).

In the following, we first show the convergence behavior of the proposed inexact ALM and BSUM algorithms in Section IV-A. Then, we examine the impact of various problem parameters on the communication and radar performance in Section IV-B. Finally, we compare the performance of the proposed approach with existing SOTA ones in Section IV-C.

##### A. Convergence Behavior

We first test the convergence behaviors of proposed Algorithms 1 and 2 in Fig. 4, where  $(N, K, M, L) = (64, 4, 4, 4)$  and  $b = 0.8$ . Specifically, Fig. 4 (a) shows the behavior of the inner loop, where proposed BSUM Algorithm 2 is applied to solve the ALM subproblem. Here, we say that a successful update is achieved if Algorithm 2 reaches an  $\epsilon_m$ -stationary point of the ALM subproblem within the specified number

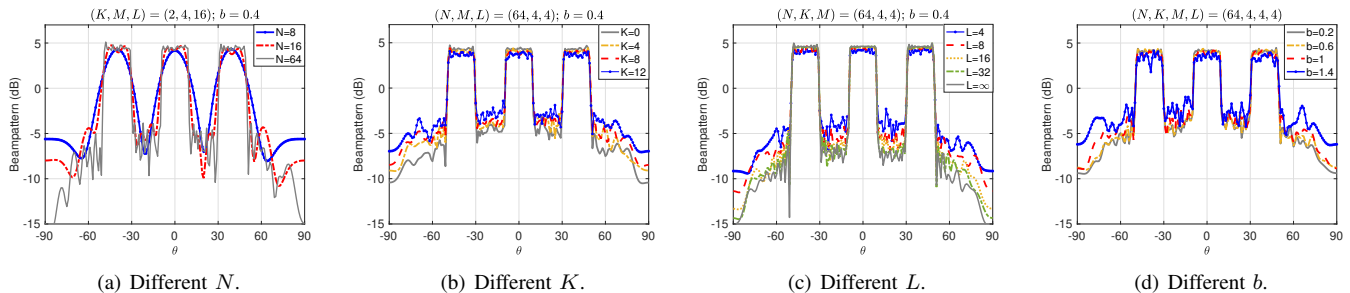


Fig. 6: Achieved beampatterns by the proposed approach under different system configurations.

of iterations. As shown Fig. 4 (a), an  $\epsilon_m$ -stationary point can be obtained within a very small number of inner iterations. Notably, over 80% of the inner loops terminate in fewer than 10 iterations and almost all of the inner loops terminates within 20 iterations. This, combined with the fact that all variables admit closed-form updates, demonstrates the efficiency of proposed BSUM Algorithm 2. The convergence behavior of the outer loop is presented in Fig. 4 (b). As demonstrated, the proposed Algorithm 1 converges rapidly at the early iterations and successfully meet the stopping criterion in (31) within the preset maximum of 500 iterations. The good convergence behaviors of both the inner and outer loops make the proposed inexact ALM algorithm computationally very efficient.

### B. Communication and Radar Performance under Different System Configurations

In this subsection, we evaluate the communication and radar performance achieved by the proposed approach under various system configurations. In particular, we aim at investigating the impact of several problem parameters, including the system parameters  $(N, K, M, L)$  and the safety margin threshold  $b$ , on the communication SER and the radar beampattern.

We first depict the SER with respect to the safety margin threshold  $b$  under different system configurations in Fig. 5, where the communication SNR is set as 10 dB. The theoretical upper and lower bounds on the SEP in (2) are included as benchmarks. As shown in the figure, the achieved SERs by the proposed approach under all system configurations decrease as  $b$  increases and remain much lower than the theoretical bound in most cases. This is because some of the CI constraints in (9b) are satisfied with strict inequality at the output solution, which makes the achieved SER much better than the theoretical SER derived based on the fixed  $b$ . However, when the problem becomes difficult (i.e., when  $N/K$  is small,  $L$  is small,  $M$  is large, and  $b$  is large), more CI constraints will be satisfied with equality or even violated (even the problem itself might be infeasible), leading to a reduction in the gap between the achieved and theoretical SERs, e.g., in the configuration  $(N, K, M, L) = (64, 4, 8, 4)$ . This also explains why different system configurations exhibit dramatically distinct SER performances with the same  $b$ .

In Fig. 6, we depict the achieved beampatterns by the proposed approach under various system configurations and explore the effects of different problem parameters on the

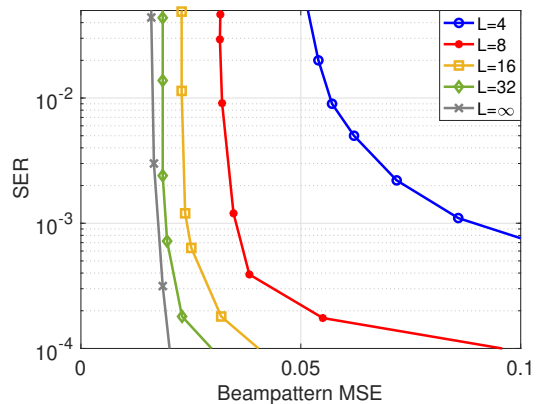


Fig. 7: Achieved radar beampattern MSE and communication SER trade-offs achieved by the proposed approach, where  $(N, K, M, L) = (64, 4, 8, 4)$  and SNR= 10 dB.

achieved beampattern. We first test the impact of the number of transmit antennas on the obtained beampatterns in Fig. 6 (a). As demonstrated, the number of transmit antennas has a strong influence on the shape of the obtained beampattern. With more antennas, the obtained beampattern can better approximate the desired square shape given in (30), emitting equal power in the main beam. This demonstrates the advantage of employing massive MIMO technology in the DFRC system.

In contrast, the number of users  $K$ , the quantization level  $L$ , and the safety margin threshold  $b$  have a weak impact on the shape and power level of the main beam; see Fig. 6 (b)-(d). Instead, they affect the power level of undesirable angles. Specifically, the power level of undesirable angles decreases as  $K$  and  $b$  decrease and  $L$  increases — that is, when the constraints in the considered problem are easier to be satisfied. An important engineering observation that can be drawn from Fig. 6 (c) is that the achieved beampattern with only  $L = 16$ , which corresponds to 3-bit DACs, is similar to that achieved by infinite-resolution DACs. Finally, we note that the constellation order  $M$  affects the problem in a similar manner to  $b$  (as it only influences the parameter  $C_t$  involved in the CI constraint). As a result, the impact of  $M$  on the achieved beampattern is also similar to that of  $b$ . As such, we omit the detailed simulation result of  $M$ .

To illustrate the impact of the quantization level  $L$  on the system performance more clearly, we plot the trade-off curves

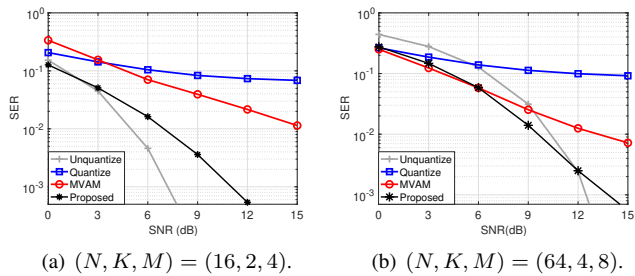


Fig. 8: Achieved SERs by different algorithms, where  $b = 0.4$ .

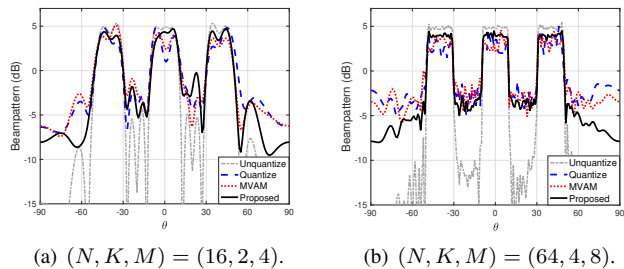


Fig. 9: Achieved beampatterns by different algorithms, where  $b = 0.4$ .

between the radar beampattern MSE and the communication SER for different quantization levels in Fig. 7. We consider a system with  $(N, K, M, L) = (64, 4, 8, 4)$  and set the SNR as 10 dB. The curves are obtained by varying the safety margin threshold  $b$ . It can be observed from the figure that increasing the quantization level from  $L = 4$  to  $L = 16$ , which corresponds to increasing the resolution of DACs from 1 bit to 3 bits, can significantly enhance both the communication and radar performance of the DFRC system. However, only marginal gains can be obtained if the quantization level is further increased beyond  $L = 16$ , i.e., the resolution of DACs is increased beyond 3 bits.

### C. Comparison with Existing SOTA Approaches

In this subsection, we compare both the radar and communication performance of the proposed approach with existing SOTA ones. As no prior work has addressed the general QCE transmit waveform design problem for DFRC systems, we shall focus on the one-bit case, i.e.,  $L = 4$ . We compare the proposed approach with the algorithm<sup>2</sup> in [4] and the MVAM algorithm in [37]. Note that the algorithm in [4] is designed for the infinite-resolution case. To obtain one-bit transmit signal, we directly quantize its output to satisfy the one-bit constraint, which is termed as “Quantize”. We also include the unquantized version of the algorithm in [4] as a benchmark, which is termed as “Unquantize”.

<sup>2</sup>In [4], several algorithms have been proposed to deal with different DFRC waveform design problems that are formulated based on different criteria. Here, we adopt the one designed for solving the weighted radar and communication optimization problem with the total power constraint as the benchmark, as it demonstrates the best radar and communication performance with the presence of one-bit quantization.

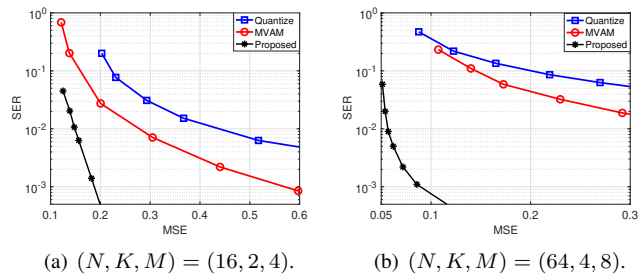


Fig. 10: Achieved radar beampattern MSE and communication SER trade-offs by different algorithms, where SNR= 10 dB.

In Figs. 8 and 9, we compare the SER and beampattern achieved by the considered four approaches for both a small system with  $(N, K, M) = (16, 2, 4)$  and a large system with  $(N, K, M) = (64, 4, 8)$ . The safety margin threshold is set as  $b = 0.4$ . Note that both of the algorithms in [4] and [37] focus on the optimization of a weighted radar and communication performance. In our implementation, we tune the weighting factor to achieve the best balanced radar and communication performance for these algorithms. As shown in Fig. 8, the SER performance of the algorithm in [4] suffers from a severe degradation with the direct one-bit quantization. In comparison, the MVAM algorithm [37] and the proposed algorithm, which are designed specifically for the one-bit scenario, achieve better SER performance. Due to superior problem formulations, they even outperform the unquantized version of the algorithm in [4] in large systems and at low SNRs; see Fig. 8 (b). Of these two approaches, the proposed one exhibits significantly lower SER, especially at high SNRs. Regarding the radar performance, it can be observed from Fig. 9 that the three one-bit approaches, though experience degradation due to the one-bit quantization, still produce satisfactory beampatterns, featuring strong mainlobes around the desired angles. Among them, the proposed approach demonstrates the best beampattern.

To give a more fair comparison, we plot in Fig. 10 the trade-off curves between the radar beampattern MSE and the communication SER of the three one-bit approaches. Specifically, the curves for the algorithms in [4] and [37] are obtained by ranging the weighting factors from 0 to 1 and the curve for the proposed algorithm is obtained by varying the safety margin threshold  $b$ . As demonstrated in the figure, the proposed algorithm achieves substantially lower radar beampattern MSE and communication SER compared to the other two algorithms, especially when the system dimension is large. In particular, with the same communication SER, the radar beampattern MSE achieved by the proposed approach is generally more than three times lower than those achieved by the SOTA algorithms. This is attributed to both the superiority of the CI-based problem formulation (compared to the MSE-based formulation) and the effectiveness of the proposed approach.

## V. CONCLUSION

In this paper, we investigated the general QCE waveform design for massive MIMO DFRC systems. The transmit waveform was optimized to minimize the MSE between the designed and desired beampatterns while satisfying the CI-based communication QoS constraints and the QCE constraint. We proposed an efficient approach for solving the formulated discrete optimization problem, which consists of two main steps. First, the discrete problem was transformed into an equivalent continuous problem using the penalty technique. Second, an inexact ALM algorithm was developed for solving the nonconvex continuous penalty model, where the ALM subproblem at each iteration was efficiently solved by a custom-designed BSUM algorithm in an inexact fashion. Simulation results demonstrated the good convergence behaviors of the proposed inexact ALM and BSUM algorithms and the superiority of the proposed approach compared to the existing SOTA ones in both radar and communication performance. An interesting future work is to optimize the overall radar performance rather than solely focusing on the transmitter design.

### APPENDIX A PROOF OF PROPOSITION 1

In this appendix, we prove the following proposition, which covers Proposition 1 as a special case.

**Proposition 3.** *Consider the following problem:*

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{A}} F(\mathbf{x}) \\ \text{s.t. } \mathbf{x}_n \in \mathcal{X}_L^{\mathcal{R}}, \forall n \in [m], \end{aligned} \quad (32)$$

where  $\mathbf{x} \in \mathbb{R}^{2m}$ ,  $F: \mathbb{R}^{2m} \rightarrow \mathbb{R}$  is a continuously differentiable function,  $\mathcal{A} \subseteq \mathbb{R}^{2m}$  is a closed set, and  $\mathbf{x}_n := [x_n; x_{m+n}]$ . Then there exists  $\lambda_0 > 0$  such that for all  $\lambda > \lambda_0$ , problem (32) shares the same optimal solution with the following penalty model:

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{A}} F(\mathbf{x}) - \lambda \|\mathbf{x}\|^2 \\ \text{s.t. } \mathbf{x}_n \in \text{conv}(\mathcal{X}_L^{\mathcal{R}}), \forall n \in [m]. \end{aligned} \quad (33)$$

*Proof.* It suffices to show that for all  $\lambda > \lambda_0$ , each optimal solution of (33) is feasible for (32). For ease of the notation, let

$$\mathcal{F}_d := \mathcal{A} \cap \{\mathbf{x} \mid \mathbf{x}_n \in \mathcal{X}_L^{\mathcal{R}}, \forall n \in [m]\}$$

and

$$\mathcal{F}_c := \mathcal{A} \cap \{\mathbf{x} \mid \mathbf{x}_n \in \text{conv}(\mathcal{X}_L^{\mathcal{R}}), \forall n \in [m]\}$$

be the feasible sets of the discrete problem (32) and the continuous penalty problem (33), respectively. In addition, let  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_B\}$  denote all the points in  $\mathcal{F}_d$  and define

$$\mathcal{S}_1 := \bigcup_{b=1}^B \left\{ \mathbf{x} \in \mathcal{F}_c \mid 0 < \|\mathbf{x} - \mathbf{v}_b\| < \eta \sin \frac{\pi}{L} \right\}$$

and

$$\mathcal{S}_2 := \bigcup_{b=1}^B \left\{ \mathbf{x} \in \mathcal{F}_c \mid \|\mathbf{x} - \mathbf{v}_b\| \geq \eta \sin \frac{\pi}{L} \right\},$$

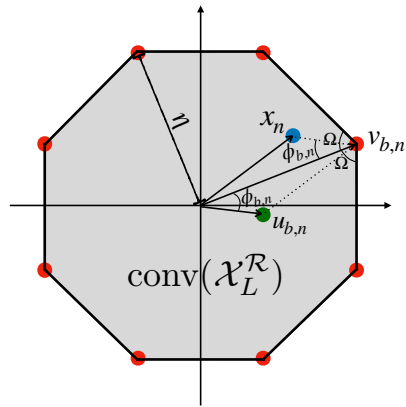


Fig. 11: An illustration of  $\{\mathbf{x}_n, \mathbf{v}_{b,n}, \mathbf{u}_{b,n}\}$  and  $\phi_{b,n}$ , where  $\Omega = \frac{\pi}{2} - \frac{\pi}{L}$ . Since  $\mathbf{x}_n$  lies in  $\text{conv}(\mathcal{X}_L^{\mathcal{R}})$ , i.e., the shaded area, we have  $\phi_{b,n} \leq \Omega$ . The equality holds when  $\mathbf{x}_n$  lies on the two edges of  $\text{conv}(\mathcal{X}_L^{\mathcal{R}})$  that are incident with  $\mathbf{v}_{b,n}$ .

where  $\mathcal{S}_1$  collects the points in  $\mathcal{F}_c$  that are sufficiently close to, but not within,  $\mathcal{F}_d$ , and  $\mathcal{S}_2$  collects the points in  $\mathcal{F}_c$  that are away from  $\mathcal{F}_d$ . To prove our claim, we need to show that all points in  $\mathcal{F}_c \setminus \mathcal{F}_d$  cannot be optimal for problem (33). For this purpose, we next examine the two cases  $\mathbf{x} \in \mathcal{S}_1$  and  $\mathbf{x} \in \mathcal{S}_2$ , separately.

**Case I:  $\mathbf{x} \in \mathcal{S}_1$ .** For any  $\mathbf{x} \in \mathcal{S}_1$ , there exists  $\mathbf{v}_b \in \mathcal{F}_d$  such that  $\|\mathbf{x} - \mathbf{v}_b\| = \rho < \eta \sin \frac{\pi}{L}$ . Let  $\mathbf{u}_b = \mathbf{v}_b - \mathbf{x}$  and  $\phi_{b,n}$  denote the angle between  $\mathbf{u}_{b,n}$  and  $\mathbf{v}_{b,n}$  for all  $n$ . Then we have

$$\begin{aligned} \|\mathbf{v}_b\|^2 - \|\mathbf{x}\|^2 &= \|\mathbf{v}_b\|^2 - \|\mathbf{v}_b - \mathbf{u}_b\|^2 \\ &= 2\mathbf{u}_b^T \mathbf{v}_b - \|\mathbf{u}_b\|^2 \\ &= 2 \sum_{n=1}^m \mathbf{u}_{b,n}^T \mathbf{v}_{b,n} - \rho^2 \\ &\stackrel{(a)}{=} 2\eta \sum_{n=1}^m \|\mathbf{u}_{b,n}\| \cos \phi_{b,n} - \rho^2 \\ &\stackrel{(b)}{\geq} 2\eta \sin \frac{\pi}{L} \rho - \rho^2, \end{aligned} \quad (34)$$

where (a) uses the fact that  $\|\mathbf{v}_{b,n}\| = \eta$ , and (b) holds since  $\sum_{n=1}^m \|\mathbf{u}_{b,n}\| \geq \|\mathbf{u}_b\| = \rho$  and  $\phi_{b,n} \leq \frac{\pi}{2} - \frac{\pi}{L}$ ; see Fig. 11 for an illustration. Since  $F(\mathbf{x})$  is continuously differentiable on the bounded set  $\mathcal{F}_c$ , it is Lipschitz continuous on  $\mathcal{F}_c$ . Let  $L_F$  be the Lipschitz constant, then we get

$$|F(\mathbf{x}) - F(\mathbf{v}_b)| \leq L_F \|\mathbf{x} - \mathbf{v}_b\| = L_F \rho. \quad (35)$$

Combing (35) with (34) and noting that  $\rho < \eta \sin \frac{\pi}{L}$ , we have that for any  $\lambda > \frac{L_F}{\eta \sin \frac{\pi}{L}}$ ,

$$\begin{aligned} F(\mathbf{x}) - \lambda \|\mathbf{x}\|^2 - (F(\mathbf{v}_b) - \lambda \|\mathbf{v}_b\|^2) \\ \geq \rho(-L_F + 2\lambda \eta \sin \frac{\pi}{L} - \lambda \rho) \\ \geq \rho(-L_F + \lambda \eta \sin \frac{\pi}{L}) > 0, \end{aligned}$$

i.e.,  $\mathbf{x}$  is not optimal for (33).

**Case II:  $\mathbf{x} \in \mathcal{S}_2$ .** Now consider any  $\mathbf{x} \in \mathcal{S}_2$ . It is easy to check that the set  $\mathcal{S}_2$  is compact, and thus

$$\max_{\mathbf{x} \in \mathcal{S}_2} \|\mathbf{x}\| := c_0 < \eta.$$

Let  $\mathbf{v}^* \in \mathcal{F}_d$  and  $F^* := F(\mathbf{v}^*)$  be an optimal solution and the optimal value of problem (32), respectively, and let  $\hat{F} := \min_{\mathbf{x} \in \mathcal{S}_2} F(\mathbf{x})$ . Then for all  $\mathbf{x} \in \mathcal{S}_2$  and  $\lambda > \frac{F^* - \hat{F}}{\eta - c_0}$ , we have

$$F(\mathbf{x}) - \lambda \|\mathbf{x}\|^2 - (F(\mathbf{v}^*) - \lambda \|\mathbf{v}^*\|^2) \geq \hat{F} - F^* + \lambda(\eta^2 - c_0^2) > 0,$$

where the first inequality holds since  $\|\mathbf{v}^*\| = \eta$ . Hence,  $\mathbf{x}$  is not optimal for (33).

Combining the above two cases, we can now conclude that the desired result holds for  $\lambda_0 = \max\left\{\frac{L_F}{\eta \sin \frac{\pi}{L}}, \frac{F^* - \hat{F}}{\eta - c_0}\right\}$ .  $\square$

## APPENDIX B

### BOUNDEDNESS OF THE SEQUENCES GENERATED BY ALGORITHMS 1 AND 2

In this section, we show that the sequences generated by Algorithms 1 and 2 are bounded. This result is important to the proof of both Theorem 1 and Proposition 2.

**Lemma 1.** *There exists a constant  $\bar{M} > 0$  (independent of  $m$  and  $r$ ) such that*

- 1) *The sequence  $(\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m)$  generated by Algorithm 1 satisfies*

$$\|\mathbf{x}^m\| \leq \bar{M}, \|\mathbf{w}^m\| \leq \bar{M}, \|\mathbf{z}^m\| \leq \bar{M}, \forall m.$$

- 2) *For any  $m > 0$ , the sequence  $(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  generated by Algorithm 2 satisfies*

$$\|\mathbf{x}^{(r)}\| \leq \bar{M}, \|\mathbf{w}^{(r)}\| \leq \bar{M}, \|\mathbf{z}^{(r)}\| \leq \bar{M}, \forall r.$$

*Proof.* Note that for any  $m \geq 0$ ,  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})$  is the output of Algorithm 2 for solving the ALM subproblem (14). Hence, it suffices to prove the second assertion of the above lemma.

Given any  $m \geq 0$  and consider the sequence  $\{(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)})\}$  generated by Algorithm 2 for solving the ALM subproblem (14). First, since  $\mathbf{x}^{(r+1)} \in \mathcal{X}$  and noting the definition of  $\mathcal{X}$ , we have

$$\|\mathbf{x}^{(r+1)}\| \leq \sqrt{NT}\eta := M_x, \forall r.$$

In addition, it follows from the update rule of variable  $\mathbf{z}$  in (29) that

$$\begin{aligned} & \|\mathbf{z}^{(r+1)}\| \\ & \leq \|\mathbf{C}\| \|\mathbf{x}^{(r+1)}\| + \|\mathbf{b}\| + \frac{\|\boldsymbol{\mu}^m\|}{\rho_\mu^m} \\ & \leq \|\mathbf{C}\| M_x + \|\mathbf{b}\| + \frac{\sqrt{2KT} \max\{|\mu_{\min}|, |\mu_{\max}|\}}{\rho_\mu^0} := M_z, \forall r, \end{aligned}$$

where the second inequality holds since the elements of  $\boldsymbol{\mu}$  are lower and upper bounded by  $\mu_{\min}$  and  $\mu_{\max}$ , respectively, and  $\rho_\mu^m$  is increasing in  $m$  and thus lower bounded by  $\rho_\mu^0$ ; see (16) and (17). The remaining task is to show that  $\{\mathbf{w}^{(r)}\}$  is bounded by a constant (independent of  $m$  and  $r$ ). According to the update rule of  $\mathbf{w}^{(r)}$  in (27), we only need to prove that  $\{\beta_q^{(r+1)}\}$  is bounded for all  $q \in [Q]$ .

Recall that  $\beta_q^{(r+1)}$  is the unique positive root of the cubic equation:

$$P(\beta) := 4\beta^3 + \rho_\nu^m \beta - \|\boldsymbol{\xi}_q^{(r)}\| = 0,$$

where<sup>3</sup>  $\boldsymbol{\xi}_q^{(r)} = (\nabla g(\mathbf{w}^{(r)}) - \boldsymbol{\nu}^m - \rho_\nu^m \mathbf{A}\mathbf{x}^{(r+1)})_{(q)}$ . Using the definitions of  $g(\mathbf{w})$  and  $\beta_q^{(r)}$ , we can further express  $\boldsymbol{\xi}_q^{(r)}$  as

$$\boldsymbol{\xi}_q^{(r)} = \left( 4 \sum_{q \in [Q]} c_q (\beta_q^{(r)})^2 \right) c_q \mathbf{w}_{(q)}^{(r)} - \boldsymbol{\nu}_{(q)}^m - \rho_\nu^m (\mathbf{A}\mathbf{x}^{(r+1)})_{(q)}.$$

It is obvious that  $P(\beta)$  is increasing in  $\beta$  with  $P(0) < 0$  and  $P(+\infty) = +\infty$ . Hence, if we can find a positive constant  $c$  such that  $P(c) > 0$ , then  $\beta_q^{(r+1)} \leq c$ . Note that  $P(\beta)$  can be lower bounded as

$$\begin{aligned} P(\beta) & \geq 4\beta^3 + \rho_\nu^m \beta - 4 \left( \sum_{q=1}^Q c_q (\beta_q^{(r)})^2 \right) c_q \beta_q^{(r)} \\ & \quad - \left\| \boldsymbol{\nu}_{(q)}^m \right\| - \rho_\nu^m \left\| (\mathbf{A}\mathbf{x}^{(r+1)})_{(q)} \right\|, \end{aligned}$$

where we have used the fact that  $\|\mathbf{w}_{(q)}^{(r)}\| = \beta_q^{(r)}$ . For the last two terms in the above lower bound, there exists a constant  $M_{w,0} > 0$  independent of  $m$  and  $r$  such that

$$\left\| \boldsymbol{\nu}_{(q)}^m \right\| + \left\| \rho_\nu^m (\mathbf{A}\mathbf{x}^{(r+1)})_{(q)} \right\| \leq \rho_\nu^m M_{w,0}.$$

Let  $\mathcal{I} = \{q \in [Q] \mid c_q > 0\}$ , where  $c_q$  is given in (8). We next investigate the two cases  $q \notin \mathcal{I}$  and  $q \in \mathcal{I}$  separately.

In case of  $q \notin \mathcal{I}$ , i.e.,  $c_q = 0$ , we have

$$P(\beta) \geq 4\beta^3 + \rho_\nu^m (\beta - M_{w,0}),$$

which further implies that  $P(M_{w,0}) > 0$  and  $\beta_q^{(r+1)} \leq M_{w,0}$ .

In case of  $q \in \mathcal{I}$ , it is straightforward to verify that  $P(\beta) > 0$  if

$$\beta > \max \left\{ \left( \sum_{q \in \mathcal{I}_q} c_q (\beta_q^{(r)})^2 \right)^{\frac{1}{3}}, (c_q \beta_q^{(r)})^{\frac{1}{3}}, M_{w,0} \right\}.$$

Then we immediately get

$$\beta_q^{(r+1)} \leq \max \left\{ \left( \sum_{q \in \mathcal{I}_q} c_q (\beta_q^{(r)})^2 \right)^{\frac{1}{3}}, (c_q \beta_q^{(r)})^{\frac{1}{3}}, M_{w,0} \right\}.$$

Squaring the above inequality, multiplying both sides by  $c_q$ , and then summing over  $q \in \mathcal{I}$ , we obtain

$$\begin{aligned} & \sum_{q \in \mathcal{I}} c_q (\beta_q^{(r+1)})^2 \\ & \leq \max \left\{ \left( \sum_{q \in \mathcal{I}} c_q (\beta_q^{(r)})^2 \right)^{\frac{2}{3}}, \left( \sum_{q \in \mathcal{I}} c_q (c_q \beta_q^{(r)})^{\frac{2}{3}} \right), \sum_{q \in \mathcal{I}} c_q M_{w,0}^2 \right\}. \end{aligned} \quad (36)$$

In addition, using  $\sum_{q \in \mathcal{I}} c_q^2 = 1$  (see (8)) and the (extended) Holder inequality, i.e.,

$$\left( \sum_{q \in \mathcal{I}} |u_q|^a |v_q|^b \right)^{a+b} \leq \left( \sum_{q \in \mathcal{I}} |u_q|^{a+b} \right)^a \left( \sum_{q \in \mathcal{I}} |v_q|^{a+b} \right)^b,$$

<sup>3</sup>Here and after, for a vector  $\boldsymbol{\xi}$ , we use  $\boldsymbol{\xi}_{(q)}$  to denote its subvector that collects the elements in  $\boldsymbol{\xi}$  at the same positions as  $\mathbf{w}_{(q)}$  in  $\mathbf{w}$ .



with  $u_q = c_q^2$ ,  $v_q = c_q(\beta_q^{(r)})^2$ ,  $a = \frac{2}{3}$ , and  $b = \frac{1}{3}$ , we have

$$\sum_{q \in \mathcal{I}} c_q (c_q \beta_q^{(r)})^{\frac{2}{3}} \leq \left( \sum_{q \in \mathcal{I}} c_q (\beta_q^{(r)})^2 \right)^{\frac{1}{3}}.$$

Plugging the above inequality into (36) yields

$$\sum_{q \in \mathcal{I}} c_q (\beta_q^{(r+1)})^2 \leq \max \left\{ \sum_{q \in \mathcal{I}} c_q (\beta_q^{(r)})^2, \sum_{q \in \mathcal{I}} c_q M_{w,0}^2 \right\},$$

which further implies

$$\sum_{q \in \mathcal{I}} c_q (\beta_q^{(r+1)})^2 \leq \max \left\{ \sum_{q \in \mathcal{I}} c_q (\beta_q^{(0)})^2, \sum_{q \in \mathcal{I}} c_q M_{w,0}^2 \right\}, \quad \forall r. \quad (37)$$

Note that  $\beta_q^{(0)} = \|\mathbf{w}_{(q)}^m\|$  (since the initial point of Algorithm 2 is set as  $(\mathbf{x}^{(0)}, \mathbf{w}^{(0)}, \mathbf{z}^{(0)}) = (\mathbf{x}^m, \mathbf{w}^m, \mathbf{z}^m)$ ) and  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})$  is the output of Algorithm 2. Then, it follows from (37) that

$$\sum_{q \in \mathcal{I}} c_q \|\mathbf{w}_{(q)}^{m+1}\|^2 \leq \max \left\{ \sum_{q \in \mathcal{I}} c_q \|\mathbf{w}_{(q)}^m\|^2, \sum_{q \in \mathcal{I}} c_q M_{w,0}^2 \right\}, \quad \forall m. \quad (38)$$

Combining (38) with (37), we get

$$\begin{aligned} \sum_{q \in \mathcal{I}} c_q (\beta_q^{(r+1)})^2 &\stackrel{(a)}{\leq} \max \left\{ \sum_{q \in \mathcal{I}} c_q \|\mathbf{w}_{(q)}^m\|^2, \sum_{q \in \mathcal{I}} c_q M_{w,0}^2 \right\} \\ &\stackrel{(b)}{\leq} \max \left\{ \sum_{q \in \mathcal{I}} c_q \|\mathbf{w}_{(q)}^0\|^2, \sum_{q \in \mathcal{I}} c_q M_{w,0}^2 \right\}, \quad \forall r, \end{aligned}$$

where (a) is obtained by replacing  $\beta_q^{(0)}$  in (37) with  $\|\mathbf{w}_{(q)}^m\|$  and (b) is obtained by iteratively applying (38) for  $m-1, \dots, 0$ . Hence, since  $c_q > 0$  for  $q \notin \mathcal{I}$ , we further have

$$\begin{aligned} &\sum_{q \in \mathcal{I}} (\beta_q^{(r+1)})^2 \\ &\leq \frac{\max \left\{ \sum_{q \in \mathcal{I}} c_q \|\mathbf{w}_{(q)}^0\|^2, \sum_{q \in \mathcal{I}} c_q M_{w,0}^2 \right\}}{\min_{q \in \mathcal{I}} c_q} := M_{w,1}, \quad \forall r. \end{aligned}$$

Let  $M_w := \sqrt{(Q - |\mathcal{I}|)M_{w,0}^2 + M_{w,1}}$ , where  $|\mathcal{I}|$  denotes the number of elements in  $\mathcal{I}$ . Then it is simple to check that  $\|\mathbf{w}^{(r)}\| \leq M_w$  for all  $r$ .

Combing the above discussions, we can now conclude that  $\bar{M} := \max\{M_x, M_z, M_w\}$  satisfies the condition in Lemma 1, which completes the proof.  $\square$

#### APPENDIX C PROOF OF THEOREM 1

According to Lemma 1, the sequence  $\{(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})\}$  generated by Algorithm 1 is bounded. As a result, to prove Theorem 1, we only need to prove the first assertion in it (and the second assertion follows directly from the first assertion and the fact that  $\{(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})\}$  lies in a compact set). The following

proof follows a similar argument to that in [44, Theorems 3.3-3.4].

Let  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}})$  be any limit point of  $\{(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})\}$ . We first show that  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}})$  is feasible, i.e.,  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}}) \in \mathcal{F}$ . Since  $\mathcal{X}$  and  $\mathcal{Z}$  are closed sets and  $\mathbf{x}^{m+1} \in \mathcal{X}$  and  $\mathbf{z}^{m+1} \in \mathcal{Z}$  for all  $m$ , we have  $\bar{\mathbf{x}} \in \mathcal{X}$  and  $\bar{\mathbf{z}} \in \mathcal{Z}$ . The remaining task is to show that  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}})$  satisfies the linear constraints in  $\mathcal{F}$ , i.e.,  $\mathbf{C}\bar{\mathbf{x}} - \bar{\mathbf{z}} = \mathbf{b}$  and  $\mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{w}}$ . From the update rule in (17), we can express  $(\rho_\mu^m, \rho_\nu^m) = (a_\mu \rho^m, a_\nu \rho^m)$ , where  $a_\mu > 0$ ,  $a_\nu > 0$ , and  $\rho^m$  is increasing in  $m$ . Hence, either  $\rho^m \rightarrow \infty$  or  $\rho^m$  keeps fixed after a finite number of iterations. In the latter case, the violation of the linear constraints will be decreased by a factor of  $\delta < 1$  at each iteration and thus tends to zero as  $m \rightarrow \infty$ , which further implies that  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}}) \in \mathcal{F}$ .

We next consider the case where  $\rho^m \rightarrow \infty$ . Assume, without loss of generality, that  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1}) \rightarrow (\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}})$ . Since  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})$  is an  $\epsilon_m$ -stationary point of the ALM subproblem, there exists  $\mathbf{e}^m := [\mathbf{e}_x^{m\top}, \mathbf{e}_y^{m\top}, \mathbf{e}_z^{m\top}]^\top$  such that

$$\begin{aligned} \mathbf{e}^m \in &\nabla \mathcal{L}_m(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1}) \\ &+ \partial \mathbb{I}_{\mathcal{X}}(\mathbf{x}^{m+1}) + \partial \mathbb{I}_{\mathcal{Z}}(\mathbf{z}^{m+1}) \end{aligned} \quad (39)$$

with  $\|\mathbf{e}^m\| \leq \epsilon_m$  (see (15)). Substituting  $\mathcal{L}_m(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})$  in (19) into (39), we get

$$\begin{aligned} &\langle \nabla h(\mathbf{x}^{m+1}) + \mathbf{C}^\top \boldsymbol{\mu}^m + \mathbf{A}^\top \boldsymbol{\nu}^m + \rho_\mu^m \mathbf{C}^\top (\mathbf{C}\mathbf{x}^{m+1} - \mathbf{z}^{m+1} - \mathbf{b}) \\ &+ \rho_\nu^m \mathbf{A}^\top (\mathbf{A}\mathbf{x}^{m+1} - \mathbf{w}^{m+1}) - \mathbf{e}_x^m, \mathbf{x} - \mathbf{x}^{m+1} \rangle \geq 0, \quad \forall \mathbf{x} \in \mathcal{X}, \end{aligned} \quad (40a)$$

$$\begin{aligned} &\nabla f(\mathbf{w}^{m+1}) + \nabla g(\mathbf{w}^{m+1}) - \boldsymbol{\nu}^m \\ &- \rho_\nu^m (\mathbf{A}\mathbf{x}^{m+1} - \mathbf{w}^{m+1}) - \mathbf{e}_w^m = \mathbf{0}, \end{aligned} \quad (40b)$$

$$\begin{aligned} &\langle -\boldsymbol{\mu}^m - \rho_\mu^m (\mathbf{C}\mathbf{x}^{m+1} - \mathbf{z}^{m+1} - \mathbf{b}) - \mathbf{e}_z^m, \mathbf{z} - \mathbf{z}^{m+1} \rangle \geq 0, \\ &\quad \forall \mathbf{z} \in \mathcal{Z}, \end{aligned} \quad (40c)$$

where we have used the fact that for a convex set  $\mathcal{X}$ ,  $\boldsymbol{\eta} \in \partial \mathbb{I}_{\mathcal{X}}(\bar{\mathbf{x}})$  is equivalent to  $\boldsymbol{\eta}^\top (\mathbf{x} - \bar{\mathbf{x}}) \leq 0$  for all  $\mathbf{x} \in \mathcal{X}$ . By further dividing (40a)–(40c) by  $\rho^m$ , we obtain

$$\begin{aligned} &\left\langle \frac{\nabla h(\mathbf{x}^{m+1}) + \mathbf{C}^\top \boldsymbol{\mu}^m + \mathbf{A}^\top \boldsymbol{\nu}^m - \mathbf{e}_x^m}{\rho^m} + a_\mu \mathbf{C}^\top (\mathbf{C}\mathbf{x}^{m+1} - \mathbf{z}^{m+1} - \mathbf{b}) \right. \\ &\left. + a_\nu \mathbf{A}^\top (\mathbf{A}\mathbf{x}^{m+1} - \mathbf{w}^{m+1}), \mathbf{x} - \mathbf{x}^{m+1} \right\rangle \geq 0, \quad \forall \mathbf{x} \in \mathcal{X}, \end{aligned} \quad (41a)$$

$$\begin{aligned} &\frac{\nabla f(\mathbf{w}^{m+1}) + \nabla g(\mathbf{w}^{m+1}) - \boldsymbol{\nu}^m - \mathbf{e}_w^m}{\rho^m} \\ &- a_\nu (\mathbf{A}\mathbf{x}^{m+1} - \mathbf{w}^{m+1}) = \mathbf{0}, \end{aligned} \quad (41b)$$

$$\begin{aligned} &\left\langle -\frac{\boldsymbol{\mu}^m - \mathbf{e}_z^m}{\rho^m} - a_\mu (\mathbf{C}\mathbf{x}^{m+1} - \mathbf{z}^{m+1} - \mathbf{b}), \mathbf{z} - \mathbf{z}^{m+1} \right\rangle \geq 0, \\ &\quad \forall \mathbf{z} \in \mathcal{Z}. \end{aligned} \quad (41c)$$

Recall that  $(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})$  lies in a compact set. Since  $\nabla h(\mathbf{x}), \nabla f(\mathbf{w})$ , and  $\nabla g(\mathbf{w})$  are all continuous, they are bounded on that compact set. In addition, the multipliers

$(\boldsymbol{\mu}^m, \boldsymbol{\nu}^m)$  are bounded and  $\|\mathbf{e}^m\| \leq \epsilon_m$  with  $\epsilon_m \rightarrow 0$ . Therefore, by letting  $m \rightarrow \infty$  in (41a)–(41c), we have

$$\langle a_\mu \mathbf{C}^\top (\mathbf{C}\bar{\mathbf{x}} - \bar{\mathbf{z}} - \mathbf{b}) + a_\nu \mathbf{A}^\top (\mathbf{A}\bar{\mathbf{x}} - \bar{\mathbf{w}}), \mathbf{x} - \bar{\mathbf{x}} \rangle \geq 0, \quad \forall \mathbf{x} \in \mathcal{X}, \quad (42a)$$

$$a_\nu (\bar{\mathbf{w}} - \mathbf{A}\bar{\mathbf{x}}) = \mathbf{0}, \quad (42b)$$

$$\langle a_\mu (\bar{\mathbf{z}} - \mathbf{C}\bar{\mathbf{x}} + \mathbf{b}), \mathbf{z} - \bar{\mathbf{z}} \rangle \geq 0, \quad \forall \mathbf{z} \in \mathcal{Z}. \quad (42c)$$

From (42b),  $\bar{\mathbf{w}} - \mathbf{A}\bar{\mathbf{x}} = \mathbf{0}$ . Substituting this into (42a) and combining (42a) with (42c), we get

$$\langle \mathbf{C}\bar{\mathbf{x}} - \bar{\mathbf{z}} - \mathbf{b}, (\mathbf{C}\mathbf{x} - \mathbf{z}) - (\mathbf{C}\bar{\mathbf{x}} - \bar{\mathbf{z}}) \rangle \geq 0, \quad \forall \mathbf{x} \in \mathcal{X}, \quad \mathbf{z} \in \mathcal{Z},$$

which further implies that

$$\begin{aligned} \|\mathbf{C}\bar{\mathbf{x}} - \bar{\mathbf{z}} - \mathbf{b}\|^2 &\leq \langle \mathbf{C}\bar{\mathbf{x}} - \bar{\mathbf{z}} - \mathbf{b}, \mathbf{C}\mathbf{x} - \mathbf{z} - \mathbf{b} \rangle \\ &\leq \|\mathbf{C}\mathbf{x} - \mathbf{z} - \mathbf{b}\| \|\mathbf{C}\bar{\mathbf{x}} - \bar{\mathbf{z}} - \mathbf{b}\|, \quad \forall \mathbf{x} \in \mathcal{X}, \quad \mathbf{z} \in \mathcal{Z}, \end{aligned}$$

i.e.,  $\|\mathbf{C}\bar{\mathbf{x}} - \bar{\mathbf{z}} - \mathbf{b}\| \leq \|\mathbf{C}\mathbf{x} - \mathbf{z} - \mathbf{b}\|$  for all  $\mathbf{x} \in \mathcal{X}$  and  $\mathbf{z} \in \mathcal{Z}$ . Hence,  $\mathbf{C}\bar{\mathbf{x}} - \bar{\mathbf{z}} - \mathbf{b} = \mathbf{0}$  (as long as  $\mathcal{F}$  is nonempty), and  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}}) \in \mathcal{F}$ .

Finally, we show that  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}})$  satisfies (18). Given any  $(\mathbf{x}, \mathbf{w}, \mathbf{z}) \in \mathcal{F}$ , by taking an inner product of (40b) with  $\mathbf{w} - \mathbf{w}^{m+1}$  and then summing it with (40a) and (40c), we obtain

$$\begin{aligned} &\langle \nabla f(\mathbf{w}^{m+1}) + \nabla g(\mathbf{w}^{m+1}), \mathbf{w} - \mathbf{w}^{m+1} \rangle \\ &\quad + \langle \nabla h(\mathbf{x}^{m+1}), \mathbf{x} - \mathbf{x}^{m+1} \rangle \\ (a) \quad &\geq \langle \boldsymbol{\mu}^m, \mathbf{C}\mathbf{x}^{m+1} - \mathbf{z}^{m+1} - \mathbf{b} \rangle + \langle \boldsymbol{\nu}^m, \mathbf{A}\mathbf{x}^{m+1} - \mathbf{w}^{m+1} \rangle \\ &\quad + \rho_\mu^m \|\mathbf{C}\mathbf{x}^{m+1} - \mathbf{z}^{m+1} - \mathbf{b}\|^2 + \rho_\nu^m \|\mathbf{A}\mathbf{x}^{m+1} - \mathbf{w}^{m+1}\|^2 \\ &\quad + \langle \mathbf{e}_x^m, \mathbf{x} - \mathbf{x}^{m+1} \rangle + \langle \mathbf{e}_w^m, \mathbf{w} - \mathbf{w}^{m+1} \rangle + \langle \mathbf{e}_z^m, \mathbf{z} - \mathbf{z}^{m+1} \rangle \\ &\geq -\|\boldsymbol{\mu}^m\| \|\mathbf{C}\mathbf{x}^{m+1} - \mathbf{z}^{m+1} - \mathbf{b}\| - \|\boldsymbol{\nu}^m\| \|\mathbf{A}\mathbf{x}^{m+1} - \mathbf{w}^{m+1}\| \\ &\quad - \|\mathbf{e}_x^m\| \|\mathbf{x} - \mathbf{x}^{m+1}\| - \|\mathbf{e}_w^m\| \|\mathbf{w} - \mathbf{w}^{m+1}\| - \|\mathbf{e}_z^m\| \|\mathbf{z} - \mathbf{z}^{m+1}\|, \end{aligned} \quad (43)$$

where (a) follows from the fact  $\mathbf{C}\mathbf{x} - \mathbf{z} = \mathbf{b}$  and  $\mathbf{A}\mathbf{x} - \mathbf{w} = \mathbf{0}$ . It is simple to verify that  $\mathcal{F}$  is a compact set, given that  $\mathcal{X}$  is compact. Furthermore, note that  $(\boldsymbol{\mu}^m, \boldsymbol{\nu}^m)$  are bounded,  $(\bar{\mathbf{x}}, \bar{\mathbf{w}}, \bar{\mathbf{z}}) \in \mathcal{F}$ ,  $\|\mathbf{e}^m\| \rightarrow 0$ , and  $\{(\mathbf{x}^{m+1}, \mathbf{w}^{m+1}, \mathbf{z}^{m+1})\}$  is bounded. Then, by letting  $m \rightarrow \infty$  in (43), we get

$$\langle \nabla f(\bar{\mathbf{w}}) + \nabla g(\bar{\mathbf{w}}), \mathbf{w} - \bar{\mathbf{w}} \rangle + \langle \nabla h(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \rangle \geq 0,$$

which gives (18) and completes the proof.

## APPENDIX D PROOF OF PROPOSITION 2

In this appendix, we give the proof of Proposition 2. We first give an outline of the proof, which contains three main steps.

**Step 1: Control the violation of the first-order optimality condition.** For  $(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)})$ , we bound its violation of the first-order optimality condition by the difference between successive iterations. Specifically, based on the update rules of the proposed BSUM Algorithm 2, we find an  $\mathbf{e}^{(r+1)}$  with

$$\begin{aligned} \mathbf{e}^{(r+1)} &\in \nabla \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)}) \\ &\quad + \partial \mathbb{I}_{\mathcal{X}}(\mathbf{x}^{(r+1)}) + \partial \mathbb{I}_{\mathcal{Z}}(\mathbf{z}^{(r+1)}) \end{aligned}$$

that satisfies

$$\begin{aligned} \|\mathbf{e}^{(r+1)}\|^2 &\leq C_0 (\rho^m)^2 \left( \|\mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}\|^2 \right. \\ &\quad \left. + \|\mathbf{w}^{(r+1)} - \mathbf{w}^{(r)}\|^2 + \|\mathbf{z}^{(r+1)} - \mathbf{z}^{(r)}\|^2 \right), \end{aligned} \quad (44)$$

where  $C_0$  is a constant independent of  $m$ .

**Step 2: Establish the sufficient decrease and lower boundedness of the objective function.** We show that the objective function is sufficiently decreased at each iteration:

$$\begin{aligned} &\mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)}) - \mathcal{L}_m(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) \\ &\leq -C_1 \rho^m \left( \|\mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}\|^2 - \|\mathbf{w}^{(r+1)} - \mathbf{w}^{(r)}\|^2 \right. \\ &\quad \left. - \|\mathbf{z}^{(r+1)} - \mathbf{z}^{(r)}\|^2 \right) \end{aligned} \quad (45)$$

and is bounded from below. The sufficient decrease is established based on the fact that a strongly convex upper bound is minimized for each block at each iteration, and the boundedness is due to the specific structure of  $\mathcal{L}_m$  and the boundedness of the sequence.

**Step 3: Identify an upper bound of  $r$  with  $\|\mathbf{e}^{(r)}\| \leq \epsilon_m$  by combining Steps 1 and 2.**

We next give the detailed proof and consider the above three steps separately.

**Step 1:** From the update rules of the BSUM algorithm given by (23), (27), (29), we have

$$\begin{aligned} \mathbf{0} &\in \nabla_{\mathbf{x}} \mathcal{L}_m(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) + \gamma^m (\mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}) + \partial \mathbb{I}_{\mathcal{X}}(\mathbf{x}^{(r+1)}), \\ \mathbf{0} &= \nabla_{\mathbf{w}} \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)}) + \nabla g(\mathbf{w}^{(r)}) - \nabla g(\mathbf{w}^{(r+1)}), \\ \mathbf{0} &\in \nabla_{\mathbf{z}} \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)}) + \partial \mathbb{I}_{\mathcal{Z}}(\mathbf{z}^{(r+1)}). \end{aligned}$$

Define

$$\begin{aligned} \mathbf{e}_x^{(r+1)} &= \nabla_{\mathbf{x}} \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)}) \\ &\quad - \nabla_{\mathbf{x}} \mathcal{L}_m(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) - \gamma^m (\mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}), \\ \mathbf{e}_w^{(r+1)} &= \nabla g(\mathbf{w}^{(r+1)}) - \nabla g(\mathbf{w}^{(r)}). \end{aligned} \quad (46)$$

Then

$$\begin{aligned} \mathbf{e}^{(r+1)} &:= \begin{bmatrix} \mathbf{e}_x^{(r+1)} \\ \mathbf{e}_w^{(r+1)} \\ \mathbf{0} \end{bmatrix} \in \nabla \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)}) \\ &\quad + \partial \mathbb{I}_{\mathcal{X}}(\mathbf{x}^{(r+1)}) + \partial \mathbb{I}_{\mathcal{Z}}(\mathbf{z}^{(r+1)}). \end{aligned} \quad (47)$$

From Lemma 1,  $\{\mathbf{w}^{(r+1)}\}$  lies in a bounded set, which we denote by  $\mathcal{W}$ . Let  $L_g$  be the Lipschitz constant of  $\nabla g(\mathbf{w})$  on  $\mathcal{W}$ . Then using the definitions of  $\mathbf{e}^{(r+1)}$  and  $\mathcal{L}_m$ , we have

$$\begin{aligned} &\|\mathbf{e}^{(r+1)}\|^2 \\ &\leq 3 \left\| (\gamma^m + 2\lambda) \mathbf{I} - \rho_\mu^m \mathbf{C}^\top \mathbf{C} - \rho_\nu^m \mathbf{A}^\top \mathbf{A} \right\|^2 \|\mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}\|^2 \\ &\quad + (3(\rho_\nu^m)^2 \|\mathbf{A}\|^2 + L_g^2) \|\mathbf{w}^{(r+1)} - \mathbf{w}^{(r)}\|^2 \\ &\quad + 3(\rho_\mu^m)^2 \|\mathbf{C}\|^2 \|\mathbf{z}^{(r+1)} - \mathbf{z}^{(r)}\|^2. \end{aligned} \quad (48)$$

Note that  $\rho_\mu^m = a_\mu \rho^m$ ,  $\rho_\nu^m = a_\nu \rho^m$ , and  $\gamma^m = \rho^m \|a_\mu \mathbf{C}^\top \mathbf{C} + a_\nu \mathbf{A}^\top \mathbf{A}\|$ , where  $\rho^m$  is increasing in  $m$  and thus lower-bounded by  $\rho^0 > 0$ . Hence, according to (48), we can find a constant  $C_0 > 0$  independent of  $m$  such that (44) holds, which completes the first step.



**Step 2:** We next establish the sufficient decrease of  $\mathcal{L}_m$ . Specifically,

$$\begin{aligned} & \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) - \mathcal{L}_m(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) \\ & \stackrel{(a)}{\leq} u_x(\mathbf{x}^{(r+1)}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) - u_x(\mathbf{x}^{(r)}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) \\ & \stackrel{(b)}{\leq} -\frac{\gamma^m}{2} \|\mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}\|^2, \end{aligned}$$

where (a) is due to the facts that  $u_x(\mathbf{x}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  is an upper bound of  $\mathcal{L}_m(\mathbf{x}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  and  $u_x(\mathbf{x}^{(r)}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) = \mathcal{L}_m(\mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  and (b) follows from the update rule  $\mathbf{x}^{(r+1)} \in \arg \min_{\mathbf{x} \in \mathcal{X}} u_x(\mathbf{x}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  and the fact that  $u_x(\mathbf{x}; \mathbf{x}^{(r)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)})$  is strongly convex with modulus  $\gamma^m$ . Similarly, we can show that

$$\begin{aligned} & \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r)}) - \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r)}, \mathbf{z}^{(r)}) \\ & \leq -\frac{\rho_\nu^m}{2} \|\mathbf{w}^{(r+1)} - \mathbf{w}^{(r)}\|^2 \end{aligned}$$

and

$$\begin{aligned} & \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)}) - \mathcal{L}_m(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r)}) \\ & \leq -\frac{\rho_\mu^m}{2} \|\mathbf{z}^{(r+1)} - \mathbf{z}^{(r)}\|^2. \end{aligned}$$

Summing the above inequalities and letting  $C_1 := \min\{a_\mu, a_\nu, \|a_\mu \mathbf{C}^T \mathbf{C} + a_\nu \mathbf{A}^T \mathbf{A}\|/2\}$ , we get the sufficient decrease of the objective function in (45). Since the sequence  $\{(\mathbf{x}^{(r+1)}, \mathbf{w}^{(r+1)}, \mathbf{z}^{(r+1)})\}$  and the multipliers  $(\mu^m, \nu^m)$  are bounded, the objective function  $\mathcal{L}_m$  is bounded from below. We denote the lower bound of  $\mathcal{L}_m$  by  $\underline{L}$ .

**Step 3:** Let

$$r(\epsilon_m) := \min\{r \geq 0 \mid \|\mathbf{e}^{(r)}\| \leq \epsilon_m\}. \quad (49)$$

Our goal is to determine an upper bound of  $r(\epsilon_m)$ . Given any  $R \geq 1$ , combining (44) and (45) and summing them from  $r = 0$  to  $r = R - 1$ , we have

$$\begin{aligned} \sum_{r=1}^R \|\mathbf{e}^{(r)}\|^2 & \leq \frac{C_0 \rho^m}{C_1} (\mathcal{L}_m(\mathbf{x}^{(0)}, \mathbf{w}^{(0)}, \mathbf{z}^{(0)}) \\ & \quad - \mathcal{L}_m(\mathbf{x}^{(R)}, \mathbf{w}^{(R)}, \mathbf{z}^{(R)})). \end{aligned} \quad (50)$$

Recalling the definition of  $\mathcal{L}_m$  in (19), we can upper bound  $\mathcal{L}_m(\mathbf{x}^{(0)}, \mathbf{w}^{(0)}, \mathbf{z}^{(0)})$  as

$$\mathcal{L}_m(\mathbf{x}^{(0)}, \mathbf{w}^{(0)}, \mathbf{z}^{(0)}) \leq C_2 \rho^m,$$

where  $C_2$  is a positive constant independent of  $m$ . Then, letting  $R = r(\epsilon_m) - 1$  in (50) and noting the definition of  $r(\epsilon_m)$ , we have

$$(r(\epsilon_m) - 1) \epsilon_m^2 \leq \sum_{r=1}^{r(\epsilon_m)-1} \|\mathbf{e}^{(r)}\|^2 \leq \frac{C_0 C_2}{C_1} (\rho^m)^2 - \underline{L},$$

which implies that  $r(\epsilon_m) = \mathcal{O}\left(\frac{(\rho^m)^2}{\epsilon_m^2}\right)$ .

## REFERENCES

- [1] L. Zheng, M. Lops, Y. C. Eldar, and X. Wang, "Radar and communication coexistence: An overview," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 85–99, Sept. 2019.
- [2] F. Liu, C. Masouros, A. Li, T. Ratnarajah, and J. Zhou, "MIMO radar and cellular coexistence: A power-efficient approach enabled by interference exploitation," *IEEE Trans. Signal Process.*, vol. 66, no. 14, pp. 3681–3695, Jul. 2018.
- [3] P. Kumari, J. Choi, N. González-Prelcic, and R. W. Heath, "IEEE 802.11ad-based radar: An approach to joint vehicular communication-radar system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3012–3027, Apr. 2018.
- [4] F. Liu, L. Zhou, C. Masouros, A. Li, W. Luo, and A. Petropulu, "Toward dual-functional radar-communication systems: Optimal waveform design," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4264–4279, Aug. 2018.
- [5] X. Liu, T. Huang, N. Shlezinger, Y. Liu, J. Zhou, and Y. C. Eldar, "Joint transmit beamforming for multiuser MIMO communications and MIMO radar," *IEEE Trans. Signal Process.*, vol. 68, pp. 3929–3944, 2020.
- [6] F. Liu, Y.-F. Liu, A. Li, C. Masouros, and Y. C. Eldar, "Cramér-Rao bound optimization for joint radar-communication beamforming," *IEEE Trans. Signal Process.*, vol. 70, pp. 240–253, 2022.
- [7] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.
- [8] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, "An overview of signal processing techniques for joint communication and radar sensing," *IEEE J. Sel. Top. Signal Process.*, vol. 15, no. 6, pp. 1295–1315, Nov. 2021.
- [9] Z. Wang, J. Wu, Y.-F. Liu, and F. Liu, "Globally optimal beamforming design for integrated sensing and communication systems," 2023. [Online]. Available: <http://arxiv.org/abs/2309.06674>
- [10] I.-R. WPSD, "Draft new recommendation ITU-RM.[IMT.FRAMEWORK FOR 2030 AND BEYOND]," 2023.
- [11] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [12] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*, 2nd ed. Norwood, MA, USA: Artech House, 2006.
- [13] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized precoding for massive MU-MIMO," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4670–4684, Nov. 2017.
- [14] Z. Wu, J. Wu, W.-K. Chen, and Y.-F. Liu, "Diversity order analysis for quantized constant envelope transmission," *IEEE Open J. Signal Process.*, vol. 4, pp. 21–30, 2023.
- [15] Z. Wu, J. Ma, Y.-F. Liu, and A. L. Swindlehurst, "Asymptotic SEP analysis and optimization of linear-quantized precoding in massive MIMO systems," *IEEE Trans. Inf. Theory (early access)*, 2023.
- [16] M. Kazemi, H. Aghaeinia, and T. M. Duman, "Discrete-phase constant envelope precoding for massive MIMO systems," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2011–2021, May 2017.
- [17] M. Shao, Q. Li, W.-K. Ma, and A. M.-C. So, "A framework for one-bit and constant-envelope precoding over multiuser massive MISO channels," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5309–5324, Oct. 2019.
- [18] H. Jedda, A. Mezghani, A. L. Swindlehurst, and J. A. Nossek, "Quantized constant envelope precoding with PSK and QAM signaling," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8022–8034, Dec. 2018.
- [19] Z. Wu, Y.-F. Liu, B. Jiang, and Y.-H. Dai, "Efficient quantized constant envelope precoding for multiuser downlink massive MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Rhodes Island, Greece, Jun. 2023, pp. 1–5.
- [20] C. Masouros, "Correlation rotation linear precoding for MIMO broadcast communications," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 252–262, Jan. 2011.
- [21] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Vector perturbation based on symbol scaling for limited feedback MISO downlinks," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 562–571, Feb. 2014.
- [22] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3628–3640, Jul. 2015.
- [23] A. Li, D. Spano, J. Krivochiza, S. Domouchtsidis, C. G. Tsinos, C. Masouros, S. Chatzinotas, Y. Li, B. Vucetic, and B. Ottersten, "A tutorial

- on interference exploitation via symbol-level precoding: Overview, state-of-the-art and future directions," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 796–839, 2nd Quart. 2020.
- [24] S. Ahmed, J. S. Thompson, Y. R. Petillot, and B. Mulgrew, "Finite alphabet constant-envelope waveform design for MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5326–5337, Nov. 2011.
- [25] A. Mezghani and R. W. Heath, "MIMO beam pattern and waveform design with low resolution DACs," in *Proc. IEEE Radar Conf.*, Boston, MA, USA, Apr. 2019, pp. 1–6.
- [26] J. Wen and B. Liao, "A block coordinate descent approach to design constant modulus waveform for MIMO radar with low-resolution DACs," in *Proc. IEEE Int. Conf. Inf. Commun. Signal Process.*, Shenzhen, China, Nov. 2022, pp. 349–354.
- [27] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4624–4634, Sept. 2017.
- [28] H. Jedda, A. Mezghani, J. A. Nossek, and A. L. Swindlehurst, "Massive MIMO downlink 1-bit precoding with linear programming for PSK signaling," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun.*, Jul. 2017, pp. 1–5.
- [29] A. Li, C. Masouros, F. Liu, and A. L. Swindlehurst, "Massive MIMO 1-bit DAC transmission: A low-complexity symbol scaling approach," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7559–7575, Nov. 2018.
- [30] F. Sohrabi, Y.-F. Liu, and W. Yu, "One-bit precoding and constellation range design for massive MIMO with QAM signaling," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 557–570, Jun. 2018.
- [31] Z. Wu, B. Jiang, Y.-F. Liu, M. Shao, and Y.-H. Dai, "Efficient CI-based one-bit precoding for multiuser downlink massive MIMO systems with PSK modulation," *IEEE Trans. Wireless Commun. (early access)*, 2023.
- [32] M. Deng, Z. Cheng, L. Wu, B. Shankar, and Z. He, "One-bit ADCs/DACs based MIMO radar: Performance analysis and joint design," *IEEE Trans. Signal Process.*, vol. 70, pp. 2609–2624, 2022.
- [33] Z. Cheng, B. Liao, Z. He, and J. Li, "Transmit signal design for large-scale MIMO system with 1-bit DACs," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4466–4478, Sept. 2019.
- [34] M. Deng, Z. Cheng, and Z. He, "Spectrally compatible waveform design for large-scale MIMO radar beam pattern synthesis with one-bit DACs," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 5, pp. 4729–4744, Oct. 2022.
- [35] Z. Cheng, S. Shi, Z. He, and B. Liao, "Transmit sequence design for dual-function radar-communication system with one-bit DACs," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5846–5860, Sept. 2021.
- [36] J. Yan and J. Zheng, "Transmit signal design of MIMO dual-function radar communication with 1-bit DACs," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun.*, Jul. 2022, pp. 1–5.
- [37] X. Yu, Q. Yang, Z. Xiao, H. Chen, V. Havyarimana, and Z. Han, "A precoding approach for dual-functional radar-communication system with one-bit DACs," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1965–1977, Jun. 2022.
- [38] R. Liu, M. Li, Q. Liu, and A. L. Swindlehurst, "Dual-functional radar-communication waveform design: A symbol-level precoding approach," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1316–1331, Nov. 2021.
- [39] Z. Zhang, Q. Chang, F. Liu, and S. Yang, "Dual-functional radar-communication waveform design: Interference reduction versus exploitation," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 148–152, Jan. 2022.
- [40] N. Su, F. Liu, Z. Wei, Y.-F. Liu, and C. Masouros, "Secure dual-functional radar-communication transmission: Exploiting interference for resilience against target eavesdropping," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7238–7252, Sept. 2022.
- [41] P. Stoica, J. Li, and Y. Xie, "On probing signal design for MIMO radar," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4151–4161, Aug. 2007.
- [42] Y.-F. Liu, T.-H. Chang, M. Hong, Z. Wu, A. M.-C. So, E. A. Jorswieck, and W. Yu, "A survey of advances in optimization methods for wireless communication system design," 2024. [Online]. Available: <http://arxiv.org/abs/2401.12025>
- [43] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 1999.
- [44] G. Galvan, M. Lapucci, T. Levato, and M. Sciandrone, "An alternating augmented Lagrangian method for constrained nonconvex optimization," *Optim. Methods Softw.*, vol. 35, no. 3, pp. 502–520, 2020.
- [45] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.