One-Bit Downlink Precoding for Massive MIMO OFDM System

Liyuan Wen, Hua Qian, *Senior Member, IEEE*, Yunbo Hu, Zhicheng Deng, and Xiliang Luo, *Senior Member, IEEE*

***Abstract*— Massive multiple-input multiple-output (MIMO) is a key technology in next generation wireless communication. However, the increasing number of radio frequency (RF) chains results in higher cost and power consumption. Given that hundreds or even thousands of transmit antennas are equipped at the base station (BS), low resolution digital-to-analog converters (DACs) are preferred to reduce the power consumption on both DACs and power amplifiers (PAs). Currently, there have been some studies about the application of low-resolution DACs for single-carrier systems. For multi-carrier systems, this problem hasn't been fully investigated. This paper aims to design a 1-bit downlink precoding algorithm for massive multi-user MIMO orthogonal frequency division multiplexing (OFDM) systems. A nonlinear precoding algorithm is proposed, which can address the non-convex optimization problem with discrete output constraint and guarantee convergence. Meanwhile, the proposed algorithm factors in the different path-losses experienced by different users. Furthermore, some approximation schemes can be applied to bring down the computational complexity of the proposed algorithm further. Simulation results illustrate that our algorithm performs the best among other nonlinear precoding methods in OFDM systems.**

***Index Terms*— Massive MIMO, OFDM, nonlinear precoding, 1-bit DAC, non-convex optimization.**

I. Introduction

MASSIVE multiple-input multiple-output (MIMO) is a key technology in the next generation wireless communication. Compared with conventional MIMO systems, hundreds or even thousands of antennas are equipped at the base station (BS) in the massive MIMO system, which improves the capacity significantly. Particularly, the spectral efficiency of the massive MIMO system is greatly improved when serving multiple user equipments (UEs) at the same time and frequency band [1], [2].

Despite the above advantages of the massive MIMO, some hardware restrictions are essential in practice. A dedicated radio frequency (RF) chain is needed for each antenna [3]. The cost and complexity of the digital signal processing are much higher in massive MIMO systems compared with conventional scenarios. Components such as power amplifiers (PAs), digital to analog converters (DACs) and analog to digital converters (ADCs) are power consuming [4].

In conventional MIMO systems, the power consumption of PAs is dominant compared to RF chains, while in massive MIMO systems, the consumption on each PA is greatly reduced with the squared number of antennas naturally [5]. Thus, the consumption of RF chains becomes relatively important in massive MIMO systems, especially with wideband signals. Therefore, some technologies such as the hybrid beamforming architecture have been proposed. A small number of RF chains are connected to the analog phase-shifter network after digital signal processing in the hybrid architecture. This technology cuts down the cost and power consumption by reducing the number of RF chains [6].

A. Benefits of Low Resolution DACs

Besides the hybrid beamforming, the use of low resolution ADCs/DACs can be another approach in practice. In contrast to the employment of fewer RF chains, the introduction of low resolution DACs reduces the power consumption on each RF chain [7]. The power consumption on DACs increases exponentially with the number of quantization bits [8]. Thus, the system power consumption is greatly reduced with the use of low resolution DACs. In addition, the PA efficiency can also be greatly improved with constant envelope signals since the nonlinear distortion of PA does not affect the quality of transmit signals [9]. The power consumption on both DACs and PAs can be reduced with 1-bit DACs.

Many wireless signals, such as orthogonal frequency division multiplexing (OFDM) signals, have large dynamic range, which require high resolution DACs to precisely represent the signal. This in turn leads to the requirement of highly linear PAs with low efficiency. Low resolution DACs, in its naive format, sacrifice the system performance.

Manuscript received 28 December 2021; revised 30 June 2022 and 10 November 2022; accepted 11 January 2023. Date of publication 26 January 2023; date of current version 12 September 2023. This work was supported in part by the National Key Research and Development Program of China under Grant 2020YFB2205603 and in part by the National Natural Science Foundation of China under Grant 61971286. The associate editor coordinating the review of this article and approving it for publication was M. Guillaud. *(Corresponding authors: Hua Qian; Xiliang Luo.)*

Liyuan Wen is with the Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China, also with the School of Information and Science Technology, ShanghaiTech University, Shanghai 201210, China, and also with the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: wenly@sari.ac.cn).

Hua Qian and Yunbo Hu are with the Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China (e-mail: qianh@sari.ac.cn; huyunbo2018@sari.ac.cn).

Zhicheng Deng and Xiliang Luo are with the School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China (e-mail: luoxiliang@ieee.org; dengzhch@shanghaitech.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2023.3238380.

Digital Object Identifier 10.1109/TWC.2023.3238380

Advanced signal processing is needed to manipulate the input signal with small dynamic range.

B. Related Works

In [10], some basic linear precoding methods in massive MIMO systems were introduced. The maximum ratio transmission (MRT) algorithm aimed to maximize the signal gain in a specific direction. The zero forcing (ZF) algorithm was used to mitigate the inter-user interference. The minimum mean square error (MMSE) algorithm was proposed to minimize the MSE at the receiver side. Considering low resolution DACs in massive MIMO systems, it is debatable whether linear precoding methods can perform well. Recently, some works have investigated the system with low resolution DACs. The performance of traditional linear precoding for the system with 1-bit DACs was analyzed in [11] and [12], where the linear precoding such as MMSE and ZF were implemented. In [13], the performance of the linear precoding was analyzed in terms of OFDM systems. The performance of such linear designs is not satisfactory as the distortion caused by the DACs appears after the precoding and is not compensated for.

The nonlinear precoding, on the other hand, could significantly improve the system performance compared with the linear precoding [14], especially in the system with low-resolution DACs. Most current works studied the single-carrier system. In [15], the authors proposed a nonlinear precoding method considering the constraint of 1-bit DACs, which minimized the mean square error (MSE) and outperformed the 1-bit linear precoding at the cost of higher complexity. In [16], the authors proposed a low complexity nonlinear precoding algorithm at much lower complexity. The error rate performance, on the other hand, was somewhat worse. In [17], the proposed algorithm exhibited both better performance and lower computational complexity. Besides, the algorithm proposed in [18] achieved satisfactory performance with convergence guarantee. In [19], the authors used the coordinate descent framework for the precoding design with 1-bit DACs. The performance was comparable with other methods, while the computational complexity was reduced. In [20], the authors studied both the linear and nonlinear precoding design with 1-bit DACs based on the constructive interference. A low-complexity scheme was proposed, which achieved a satisfactory trade-off between the performance and the complexity. In [21], the authors proposed a quadrature amplitude modulation (QAM) constellation range design for the system with one-bit precoding, which aimed to minimize the symbol error probability. In [22], a constellation-dependent method for both one-bit and constant envelope precoding was proposed. A symbol error probability minimization problem was formulated, and better performance was achieved. However, all above works focused on single-carrier systems, which could not be applied to the OFDM system directly. The conversion between the time and frequency domain cannot be omitted due to the nonlinear operation introduced by finite-resolution DACs.

Currently, only a few works have investigated the impact of low resolution DACs in multi-carrier systems. In [23], the authors proposed to maximize the distance between the received signals and the decision thresholds of the phase shift keying (PSK) signals in multi-user OFDM systems with 1-bit DACs. In [24], the authors extended the algorithm in [15] for flat-fading channels to the case with more general frequency-selective fading channels. In particular, the non-convex constraint was first dropped in [24] and then a constant envelope phase quantized signal was obtained by discretizing the solution of the relaxed problem. Besides the performance loss, there was no convergence analysis of the method for the original non-convex problem in [24]. In [25], the cyclic coordinate descent method was used to solve the MSE minimization problem formulated in massive multi-user MIMO (MU-MIMO) OFDM systems. The matrix dimension was huge with a large number of subcarriers. The convergence of existing approaches in OFDM systems was neither analyzed, nor guaranteed. Moreover, the case of users with different path-losses was not considered in the above studies. It is worthwhile further studying the nonlinear precoding design with low resolution DACs in multi-carrier systems.

C. Contributions

In this paper, we propose a novel precoding design for downlink MU-MIMO multi-carrier systems, which can compensate for the distortion introduced by 1-bit DACs. The proposed algorithm enjoys guaranteed convergence and outperforms other state-of-the-art nonlinear algorithms. Main contributions are summarized in the following.

- We propose a nonlinear precoding framework for multi-carrier systems. With the employment of 1-bit DACs, the operation of the discrete Fourier transform (DFT) and the inverse DFT (IDFT) cannot be canceled in multi-carrier systems. Our proposed design addresses this problem in a systematic way. Moreover, comparing to existing works, we consider users with different path-losses.
- The optimization problem is formulated as an MSE minimization problem, which is non-convex due to the discrete output constraint of 1-bit DACs. An efficient algorithm based on the modified framework of alternative direction method of multipliers (ADMM) is proposed with convergence guarantee. The proposed method can be extended to systems with arbitrary resolution DACs. Furthermore, the main complexity of the proposed algorithm comes from the operations of DFT and IDFT. The complexity of the remaining part is linear with respect to the number of subcarriers, which is affordable.
- Various approximation schemes are also proposed to reduce the computational complexity further. Specifically, by exploiting a series of lower complexity operations and the law of large numbers, matrix inversion can be avoided. The resulting precoding performance becomes close to that of the original algorithm as the number of antennas gets large.

The remainder of this paper is organized as follows. The system model is shown in Section II. In Section III, the problem is formulated. In Section IV, an efficient algorithm is proposed with convergence analysis. Simulation results are

illustrated in Section V. Finally, the paper is summarized in Section VI.

D. Notations

In this paper, symbols a, $\mathbf{a}$ and $\mathbf{A}$ denote the scalar, vector and matrix, respectively. The $\mathbf{A}^{\mathrm{H}}$ is the Hermitian transpose of matrix $\mathbf{A}$. Notations $\Re(a)$ and $\Im(a)$ represent the real and imaginary part of the scalar a, respectively. Notation $\|\cdot\|_p$ denotes the l_p-norm. Notation $\lfloor\cdot\rfloor$ denotes to take the floor of a real number. Notation $\mathrm{Diag}(\mathbf{a})$ denotes the diagonal matrix, and ∇ denotes the differential operator.

II. System Model

In this section, we introduce the system with 1-bit DACs at the transmitter side in MU-MIMO OFDM systems. The OFDM system in the presence of frequency-selective channels is studied. By proper system configurations, each subcarrier undergoes a flat fading channel effectively. In the following derivations, we assume that the BS has accurate knowledge about the downlink channel, a.k.a. transmitter side channel state information (CSI). The impact of channel estimation error will be characterized in Section V.

Fig. 1 illustrates the architecture of downlink massive MU-MIMO multi-carrier systems with 1-bit DACs at the transmitter. Data streams are firstly mapped to constellation points through different types of modulations. Then, symbols are processed by the nonlinear precoding module with binary-leveled time domain outputs at transmit antennas. Assume that the BS is equipped with N_t transmit antennas to serve U single-antenna users. Suppose each user owns only one stream. Ideal ADCs are assumed at UEs for simplicity. The IDFT is included in the module of nonlinear precoding, which transforms the frequency domain precoded vector into 1-bit time domain output vector at transmit antennas.

A. Time Domain

With the assumption of unit total transmit power at the BS, the received signal $\mathbf{y}_T[n] \in \mathbb{C}^{U\times 1}$ at all UEs at discrete time n can be represented as

$$\mathbf{y}_T[n] = \sum_{i=0}^{I-1} \mathbf{H}_T[i]\mathbf{x}_T[n-i] + \mathbf{z}_T[n], \quad n = 0, 1, \ldots, N-1, \tag{1}$$

where $\mathbf{x}_T[n] \in \mathbb{C}^{N_t\times 1}$ represents the time domain signal at transmit antennas, $\mathbf{z}_T[n] \in \mathbb{C}^{U\times 1}$ and $\mathbf{z}_T[n] \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_U)$ denotes the time domain independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN), and $\mathbf{H}_T[i] \in \mathbb{C}^{U\times N_t}$ denotes the complex channel filter tap at discrete time i, $i = 0, 1, \ldots, I-1$, which is assumed to remain constant within the symbol time. Notation N denotes the length of the OFDM symbol, which is also the size of DFT.

The time domain signal at transmit antennas $\mathbf{x}_T[n]$ is assumed to lie in the discrete set $\mathcal{A}^{N_t\times 1}$. With the employment of 1-bit DACs at the BS, $\mathcal{A}$ is expressed as follows,

$$\mathcal{A} = \sqrt{\gamma}\Big\{-1-j, -1+j, 1-j, 1+j\Big\}, \tag{2}$$

where γ denotes the power normalization factor. Besides, the unit total power constraint at the BS is assumed, which means that $\|\mathbf{x}_T[n]\|_2^2 = 1$ for $n = 0, 1, \ldots, N-1$.

B. Frequency Domain

Let matrices $\mathbf{X}_T = \big[\mathbf{x}_T[0], \mathbf{x}_T[1], \ldots, \mathbf{x}_T[N-1]\big]$, $\mathbf{Y}_T = \big[\mathbf{y}_T[0], \mathbf{y}_T[1], \ldots, \mathbf{y}_T[N-1]\big]$ and $\mathbf{Z}_T = \big[\mathbf{z}_T[0], \mathbf{z}_T[1], \ldots, \mathbf{z}_T[N-1]\big]$ represent the transmitted signal, received signal, and AWGN at time domain for simplicity, respectively. The corresponding frequency domain signals can be expressed in the following,

$$\begin{aligned}\mathbf{X}_F &= \mathbf{X}_T\mathbf{F}_N,\\ \mathbf{Y}_F &= \mathbf{Y}_T\mathbf{F}_N,\\ \mathbf{Z}_F &= \mathbf{Z}_T\mathbf{F}_N,\end{aligned} \tag{3}$$

where $\mathbf{F}_N \in \mathbb{C}^{N\times N}$ is the N point DFT matrix with (m, n)-th entry $(\mathbf{F}_N)_{mn} = e^{-j\frac{2\pi}{N}(m-1)(n-1)}$.

Suppose each OFDM symbol contains N subcarriers, and each subcarrier contains U data streams. Then, the k-th subcarrier received at UEs can be represented as

$$\mathbf{y}[k] = \mathbf{H}[k]\mathbf{x}[k] + \mathbf{z}[k], \quad k = 0, 1, \ldots, N-1, \tag{4}$$

where $\mathbf{x}[k]$, $\mathbf{y}[k]$, $\mathbf{z}[k]$ represent the k-th column of corresponding frequency domain matrices $\mathbf{X}_F$, $\mathbf{Y}_F$, $\mathbf{Z}_F$, respectively. The matrix $\mathbf{H}[k] \in \mathbb{C}^{U\times N_t}$ denotes the frequency domain channel of the k-th subcarrier, which satisfies

$$\mathbf{H}[k] = \sum_{i=0}^{I-1} \mathbf{H}_T[i]e^{-jk\frac{2\pi}{N}i}, \quad k = 0, 1, \ldots, N-1. \tag{5}$$

C. Precoding Design

Let $\mathbf{s}[k] \in \mathcal{M}^{U\times 1}$ denote the transmitted symbol on the k-th subcarrier, which has been mapped to the set of constellation points $\mathcal{M}$. Precoding methods can be broadly classified as linear and nonlinear precoding. These two precoding schemes can be distinguished by justifying whether the design is related to the input signals.

1) Linear Precoding: The linear precoding is determined by the CSI. Denote $\mathbf{P}[k]$ as the linear precoding on the k-th subcarrier. The ZF and Wiener-filter (WF) precoding are commonly used linear precoding algorithms. For example, the WF precoding is designed to minimize the MSE at the receiver side [26], which is given by

$$\mathbf{P}_{WF}[k] = \sqrt{\gamma}\mathbf{H}^H[k]\left(\mathbf{H}[k]\mathbf{H}^H[k] + U\sigma^2\mathbf{I}_U\right)^{-1}. \tag{6}$$

In this case, the precoded signal on the k-th subcarrier can be written as $\mathbf{x}[k] = \mathbf{P}_{WF}[k]\mathbf{s}[k]$. It should be noted that the system performance is excellent through WF precoding with infinite resolution DACs. However, the system may achieve poor performance with low resolution DACs at the BS [13]. Main reason is that the quantization procedure appears after the design of linear precoding. Thus, the distortion caused by quantization cannot be compensated by the linear precoding.

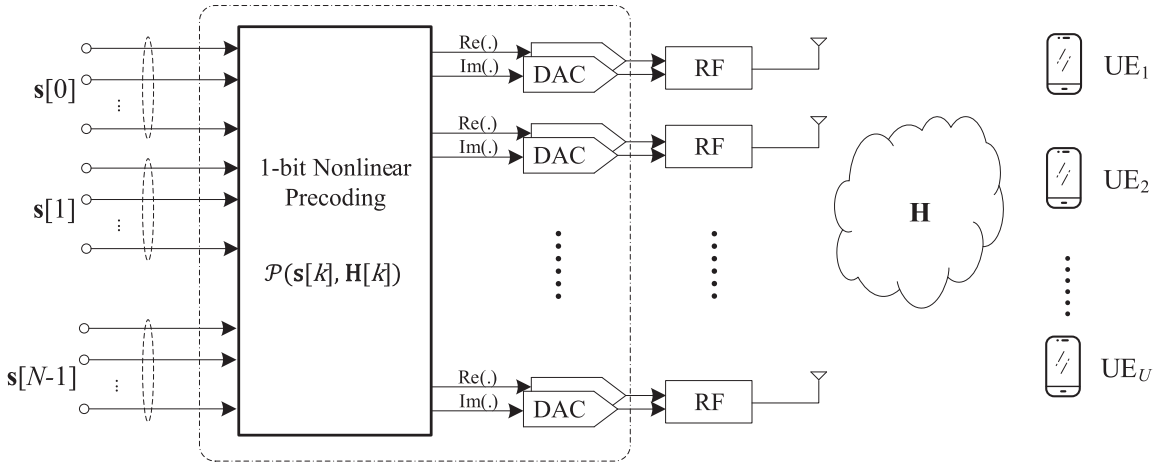


Fig. 1. Massive MU-MIMO OFDM systems with the nonlinear precoding considering 1-bit DACs.

2) Nonlinear Precoding: The design of nonlinear precoding considers both CSI and input signals. The precoded signal on the k-th subcarrier can be expressed as $\mathbf{x}[k] = \mathcal{P}(\mathbf{s}[k], \mathbf{H}[k])$. Moreover, the discrete output can be taken into account when designing the nonlinear precoding. The distortion caused by quantization can be compensated by the nonlinear precoding to some extent. Thus, the system can achieve better performance with the nonlinear precoding. In the following section, we aim to study the nonlinear precoding with the use of 1-bit DACs.

III. Problem Formulation

In order to evaluate the system performance of the quantized output, a general approach is to evaluate the signal to interference plus noise ratio (SINR) at the receiver side. However, the representation of SINR is not directly available for the OFDM signal with 1-bit DACs. One approach is to consider maximizing the signal to noise ratio (SNR) constrained by the level of harmful interference [7], [27], [28]. This approach, on the other hand, only works for phase shift keying (PSK) modulations. Another approach is to evaluate the MSE between the input and output, which is general and applicable to all types of modulations. The MSE on the k-th subcarrier can be expressed as

$$\mathrm{MSE}[k] = \mathbb{E}_{\mathbf{z}[k]}\left[\|\mathbf{s}[k] - \mathbf{A}\mathbf{y}[k]\|_2^2\right], \tag{7}$$

where $\mathbf{A} = \mathrm{Diag}[\alpha_1, \alpha_2, \ldots, \alpha_U]$, and $\alpha_i \in \mathbb{R}^+$ represents the adjustment factor of the i-th user. These adjustment factors turn out to be critical to deal with the scenario where different users experience different path-losses. By setting these parameters appropriately, e.g., as in (24), a lower overall MSE can be obtained. Comparing (7) to related works such as [17], [24] and [25] where α_i is assumed the same for $i = 1, 2, \ldots, U$, our work generalizes the feasibility for the case of users experiencing different path-losses. With the given statistical information $\mathbf{z}[k] \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_U)$, the total MSE can be further written as

$$\mathrm{MSE} = \sum_{k=0}^{N-1} \|\mathbf{s}[k] - \mathbf{A}\mathbf{H}[k]\mathbf{x}[k]\|_2^2 + N\sigma^2 \sum_{i=1}^{U} \alpha_i^2. \tag{8}$$

Then, we formulate the problem to minimize the total MSE on all subcarriers as problem (P1),

$$\begin{aligned}(\mathrm{P1}): \underset{\mathbf{X}_T, \mathbf{A}}{\text{minimize}} \quad & \sum_{k=0}^{N-1} \|\mathbf{s}[k] - \mathbf{A}\mathbf{H}[k]\mathbf{x}[k]\|_2^2 + N\sigma^2 \sum_{i=1}^{U} \alpha_i^2, \\ \text{s.t.} \quad & \mathbf{X}_F = \mathbf{X}_T\mathbf{F}_N, \qquad (9a) \\ & \mathbf{X}_T \in \mathcal{A}^{N_t \times N}, \qquad (9b)\end{aligned}$$

where the use of 1-bit DACs results in the discrete time domain outputs.

The form of (P1) is intractable, which contains signals at both time and frequency domain. This results in the difficulty for further system analysis. To simplify the following analysis, we aim to solve the problem at frequency domain. Combining constraints (9a) and (9b), we conclude that the frequency domain signal satisfies

$$\frac{1}{N}\mathbf{X}_F\mathbf{F}_N^H \in \mathcal{A}^{N_t \times N}. \tag{10}$$

Thus, problem (P1) is equivalent to problem (P2), which is

$$\begin{aligned}(\mathrm{P2}): \underset{\mathbf{X}_F, \mathbf{A}}{\text{minimize}} \quad & \sum_{k=0}^{N-1} \|\mathbf{s}[k] - \mathbf{A}\mathbf{H}[k]\mathbf{x}[k]\|_2^2 + N\sigma^2 \sum_{i=1}^{U} \alpha_i^2, \\ \text{s.t.} \quad & \frac{1}{N}\mathbf{X}_F\mathbf{F}_N^H \in \mathcal{A}^{N_t \times N},\end{aligned}$$

where the optimization variable is changed into the frequency domain signal $\mathbf{X}_F$.

The problem (P2) can be seen as a combinatorial optimization problem, which is a non-convex problem due to the discrete output constraint. A brute force method can be used for (P2) at the cost of extremely high computational complexity. It is infeasible especially in massive MU-MIMO systems with a large number of antennas. Hence, more efficient algorithms need to be studied.

IV. Design of 1-Bit Nonlinear Precoding

In this section, an algorithm is proposed based on the ADMM, which solves the non-convex optimization problem (P2) formulated in Section III. Furthermore, the convergence of the proposed algorithm is analyzed.

The problem (P2) without the constraint is a traditional quadratic programming problem, where the optimal solution can be easily obtained. However, the discrete constraint results in difficulties to solve the problem. In this paper, we propose a modified ADMM algorithm to address this problem.

The conventional ADMM algorithm is applicable to the optimization problem of two variables with equality constraint. The general form of the optimization problem to use ADMM [29] is presented as,

$$\begin{aligned} &\underset{\mathbf{p},\mathbf{q}}{\text{minimize}} \ \ g_1(\mathbf{p}) + g_2(\mathbf{q}), \\ &\text{s.t. } \mathbf{C}_1\mathbf{p} + \mathbf{C}_2\mathbf{q} = \mathbf{d}, \end{aligned} \tag{11}$$

where $\mathbf{p} \in \mathbb{R}^{m\times 1}$ and $\mathbf{q} \in \mathbb{R}^{n\times 1}$ are variables needed to be optimized, respectively. Notations $\mathbf{C}_1 \in \mathbb{R}^{l\times m}$, $\mathbf{C}_2 \in \mathbb{R}^{l\times n}$ and $\mathbf{d} \in \mathbb{R}^{l\times 1}$ are the given matrices and vector, respectively. Besides, functions $g_1(\cdot)$ and $g_2(\cdot)$ are required to be convex. With the above constraints, the ADMM algorithm can obtain the global optimal solution.

The constraint of (P2) can be integrated into the objective function as

$$\begin{aligned} f(\mathbf{X}_F, \mathbf{A}) = &\sum_{k=0}^{N-1} \|\mathbf{s}[k] - \mathbf{A}\mathbf{H}[k]\mathbf{x}[k]\|_2^2 + N\sigma^2 \sum_{i=1}^{U} \alpha_i^2 \\ &+ \delta_{\mathcal{A}}\left(\frac{1}{N}\mathbf{X}_F\mathbf{F}_N^H\right), \end{aligned} \tag{12}$$

where $\delta_{\mathcal{A}}(\cdot)$ is the indicator function given as

$$\delta_{\mathcal{A}}(\mathbf{X}) = \begin{cases} 0, & \mathbf{X} \in \mathcal{A}, \\ \infty, & \text{others}. \end{cases} \tag{13}$$

Then, the problem (P2) can be transformed into the following problem (P3)

$$(\text{P3}): \ \underset{\mathbf{X}_F,\mathbf{A}}{\text{minimize}} \quad f(\mathbf{X}_F, \mathbf{A}).$$

The problem (P3) can be further rewritten as

$$\begin{aligned} (\text{P4}): \ \underset{\mathbf{X},\mathbf{R},\mathbf{A}}{\text{minimize}} \ &\sum_{k=0}^{N-1} \|\mathbf{s}[k] - \mathbf{A}\mathbf{H}[k]\bar{\mathbf{x}}[k]\|_2^2 + N\sigma^2 \sum_{i=1}^{U} \alpha_i^2 \\ &+ \delta_{\mathcal{A}}\left(\frac{1}{N}\mathbf{R}\mathbf{F}_N^H\right), \\ \text{s.t. } &\mathbf{X} - \mathbf{R} = \mathbf{0}, \end{aligned}$$

where $\mathbf{R} \in \mathbb{C}^{N_t\times N}$ is the auxiliary variable, $\mathbf{X} \in \mathbb{C}^{N_t\times N}$, and $\bar{\mathbf{x}}[k]$ is the k-th column of $\mathbf{X}$. It is obvious that the form of (P4) is similar to the standard form of ADMM in (11). The difference is that the indicator function lies in a non-convex set, where the convergence cannot be guaranteed naturally. Therefore, the convergence of the proposed modified ADMM algorithm needs to be analyzed later.

The augmented Lagrangian adds a penalty term to the traditional Lagrangian function, which can ensure the algorithm converge faster without the strict convexity assumption [30]. The augmented Lagrangian function of (P4) is expressed as

$$\begin{aligned} \mathcal{L}_\lambda(\mathbf{X}, \mathbf{R}, \mathbf{V}) = &\sum_{k=0}^{N-1} \|\mathbf{s}[k] - \mathbf{A}\mathbf{H}[k]\bar{\mathbf{x}}[k]\|_2^2 + N\sigma^2 \sum_{i=1}^{U} \alpha_i^2 \\ &+ \delta_{\mathcal{A}}\left(\frac{1}{N}\mathbf{R}\mathbf{F}_N^H\right) + \langle \mathbf{V}, \mathbf{X} - \mathbf{R}\rangle + \frac{\lambda}{2}\|\mathbf{X} - \mathbf{R}\|_F^2, \end{aligned} \tag{14}$$

where $\lambda \in \mathbb{R}^+$ is the introduced penalty factor, and $\mathbf{V} \in \mathbb{C}^{N_t\times N}$ denotes the dual variable.

Then, the problem (P4) can be solved via the following procedure. At the $(t+1)$-th iteration, the update equations are

$$\mathbf{X}^{t+1} = \underset{\mathbf{X}}{\text{argmin}}\, \mathcal{L}_\lambda(\mathbf{X}, \mathbf{R}^t, \mathbf{V}^t), \tag{15a}$$

$$\mathbf{R}^{t+1} = \underset{\mathbf{R}}{\text{argmin}}\, \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}, \mathbf{V}^t), \tag{15b}$$

$$\mathbf{A}^{t+1} = \underset{\mathbf{A}}{\text{argmin}}\, f(\mathbf{R}^{t+1}, \mathbf{A}), \tag{15c}$$

$$\mathbf{V}^{t+1} = \mathbf{V}^t + \lambda(\mathbf{X}^{t+1} - \mathbf{R}^{t+1}). \tag{15d}$$

The updating order of primal variables have no impact on the convergence [31]. We firstly solve the minimization problem with respect to the variable $\mathbf{X}$. Then, the auxiliary variable $\mathbf{R}$ is updated, which satisfies the discrete constraint. The dual variable $\mathbf{V}$ is updated eventually.

Firstly, we rewrite (15a) with respect to $\mathbf{X}$ as

$$\begin{aligned} \mathbf{X}^{t+1} = \underset{\mathbf{X}}{\text{argmin}} &\sum_{k=0}^{N-1} \|\mathbf{s}[k] - \mathbf{A}\mathbf{H}[k]\bar{\mathbf{x}}[k]\|_2^2 + N\sigma^2 \sum_{i=1}^{U} \alpha_i^2 \\ &+ \frac{\lambda}{2}\left\|\mathbf{X} - \mathbf{R}^t + \frac{\mathbf{V}^t}{\lambda}\right\|_F^2, \end{aligned} \tag{16}$$

where the optimal solution of $\mathbf{X}^{t+1}$ can be obtained by least square method. The closed-form optimal solution of the k-th subcarrier $\bar{\mathbf{x}}^{t+1}[k]$ is expressed as

$$\begin{aligned} &\bar{\mathbf{x}}^{t+1}[k] \\ &= \left[2\tilde{\mathbf{H}}^H[k]\tilde{\mathbf{H}}[k] + \lambda\mathbf{I}_{N_t}\right]^{-1}\left(2\tilde{\mathbf{H}}^H[k]\mathbf{s}[k] + \lambda\mathbf{R}^t[k] - \mathbf{V}^t[k]\right) \\ &= \frac{1}{\lambda}\left[\mathbf{I}_{N_t} - \tilde{\mathbf{H}}^H[k]\left(\tilde{\mathbf{H}}[k]\tilde{\mathbf{H}}^H[k] + \frac{\lambda}{2}\mathbf{I}_U\right)^{-1}\tilde{\mathbf{H}}[k]\right] \\ &\quad \left(2\tilde{\mathbf{H}}^H[k]\mathbf{s}[k] + \lambda\mathbf{R}^t[k] - \mathbf{V}^t[k]\right), \end{aligned} \tag{17}$$

where $\tilde{\mathbf{H}}[k] = \mathbf{A}\mathbf{H}[k]$, and $\mathbf{R}[k], \mathbf{V}[k] \in \mathbb{C}^{N_t\times 1}$ denote the k-th column of $\mathbf{R}$ and $\mathbf{V}$, respectively.

The update of primal variable $\mathbf{R}$ in (15b) can be further solved by

$$\begin{aligned} \mathbf{R}^{t+1} &= \underset{\mathbf{r}}{\text{argmin}}\ \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}, \mathbf{V}^t) \\ &= \underset{\mathbf{R}}{\text{argmin}}\ \delta_{\mathcal{A}}\left(\frac{1}{N}\mathbf{R}\mathbf{F}_N^H\right) + \langle \mathbf{V}^t, \mathbf{X}^{t+1} - \mathbf{R}\rangle \\ &\qquad + \frac{\lambda}{2}\|\mathbf{X}^{t+1} - \mathbf{R}\|_F^2 \\ &= \underset{\mathbf{R}}{\text{argmin}}\ \delta_{\mathcal{A}}\left(\frac{1}{N}\mathbf{R}\mathbf{F}_N^H\right) + \frac{\lambda}{2}\left\|\mathbf{R} - \mathbf{X}^{t+1} - \frac{1}{\lambda}\mathbf{V}^t\right\|_F^2 \\ &\qquad - \frac{1}{2\lambda}\|\mathbf{V}^t\|_F^2 \\ &= \left\{\mathcal{P}_{\mathcal{A}}\left[\frac{1}{N}\left(\mathbf{X}^{t+1} + \frac{1}{\lambda}\mathbf{V}^t\right)\mathbf{F}_N^H\right]\right\}\mathbf{F}_N. \end{aligned} \tag{18}$$

Equation (18) suggests to firstly project the time domain signal to the given discrete set $\mathcal{A}$, and convert to the frequency domain. In general, for the system with B-bit DACs, the

projection of the i-th element of the complex signal $\mathbf{w}$ can be expressed as

$$\mathcal{P}_{\mathcal{A}}(w_i) = \|\mathbf{w}\|_\infty \left[\mathcal{P}'_{\mathcal{A}}\left(\Re(\bar{w}_i)\right) + j\mathcal{P}'_{\mathcal{A}}\left(\Im(\bar{w}_i)\right)\right], \quad (19)$$

where $\bar{\mathbf{w}} = \frac{\mathbf{w}}{\|\mathbf{w}\|_\infty}$ represents the normalized signal, and $\mathcal{P}'_{\mathcal{A}}(z)$ for $z \in \mathbb{R}$ is calculated as

$$\mathcal{P}'_{\mathcal{A}}(z) = -1 + \frac{2}{2^B - 1}\left[(z+1)(2^B-1) - \left\lfloor \frac{(z+1)(2^B-1)}{2} \right\rfloor\right]. \quad (20)$$

In particular, for the 1-bit DACs, the projection can be simplified as

$$\mathcal{P}_{\mathcal{A}}(w_i) = \text{sign}\left(\Re(w_i)\right) + j\text{sign}\left(\Im(w_i)\right). \quad (21)$$

The adjustment factor α_i for $i = 1, 2, \ldots, U$ in (15c) can be updated through least square method. The closed-form expression of α_i^{t+1} is derived as

$$\alpha_i^{t+1} = \frac{\Re\left[\sum_{k=0}^{N-1}\left(\mathbf{R}^{t+1}[k]\right)^H \mathbf{h}_i[k]\mathbf{s}_i[k]\right]}{\sum_{k=0}^{N-1}\left(\mathbf{R}^{t+1}[k]\right)^H \mathbf{h}_i[k]\mathbf{h}_i^H[k]\mathbf{R}^{t+1}[k] + N\sigma^2}, \quad (22)$$

where $\mathbf{h}_i[k] \in \mathbb{C}^{N_t \times 1}$, and $\mathbf{h}_i^T[k]$ represents the i-th row of $\mathbf{H}[k]$.

The proposed algorithm is summarized in Algorithm 1. Denote $\mathbf{R}^{\text{out}}$ as the output after T iterations, we obtain the frequency domain signal $\mathbf{X}_F$ as

$$\mathbf{X}_F = \sqrt{\gamma}\mathbf{R}^{\text{out}}, \quad (23)$$

$$\alpha_i = \frac{\sum_{k=0}^{N-1}\mathbf{x}^H[k]\mathbf{h}_i[k]\mathbf{s}_i[k]}{\sum_{k=0}^{N-1}\mathbf{x}^H[k]\mathbf{h}_i[k]\mathbf{h}_i^H[k]\mathbf{x}[k] + N\sigma^2}, \quad (24)$$

where $\gamma = \frac{N}{\|\mathbf{R}^{\text{out}}\|_F^2}$ is the power normalization factor.

Algorithm 1 Convergence-Guaranteed Multi-Carrier One-Bit Precoding (CG-MC1bit)

Initialization: $\mathbf{X}^0 = \mathbf{1}$, $\mathbf{R}^0 = \mathbf{X}$, $\mathbf{V}^0 = \mathbf{0}$, $\alpha_i = 0.01$, for $i = 1, 2, \ldots, U$;
for $t = 0 : T-1$ **do**
  for $k = 0 : N-1$ **do**
    Update $\bar{\mathbf{x}}^{t+1}[k]$ according to (17);
  end
  Update $\mathbf{R}^{t+1}$ according to (18);
  Update $\mathbf{A}^{t+1}$ according to (22);
  Update dual variable $\mathbf{V}^{t+1}$ according to (15d);
  $t = t + 1$;
end

A. Convergence Analysis

For the convex problem, the ADMM algorithm can converge to the optimal solution [30]. Since the optimization problem (P4) is non-convex due to the discrete constraint, the convergence of the algorithm in this paper depends on the penalty factor and the objective function [18]. Hence, the convergence of the proposed algorithm is necessary to be discussed. A general method for convergence analysis of ADMM was proposed in [31]. in [31], we next derive the condition to guarantee the convergence of the proposed CG-MC1bit in Algorithm 1.

With given $\mathbf{A}$, denote $g(\mathbf{X}) = f(\mathbf{X}, \mathbf{A})$. We know that $g(\mathbf{X})$ is convex when $\mathbf{X} \in \mathcal{A}$. We will show some lemmas at first, which are useful for the following analysis of convergence behavior.

Lemma 1: The augmented Lagrangian function is lower bounded as

$$\mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t) \geq g_{min} + \frac{\lambda - L_\phi}{2}\|\mathbf{R}^t - \mathbf{X}^t\|_F^2, \quad (25)$$

where L_ϕ is the Lipschitz constant, and $g_{min} := \min_{\mathbf{X}^t} g(\mathbf{X}^t)$ represents the lower bound of function $g(\cdot)$.

The proof of Lemma 1 is shown in Appendix B.

Lemma 2: The difference of the augmented Lagrangian function between successive iterations can be represented as

$$\mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^{t+1}) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t) \leq \left(\frac{1}{\lambda}L_\phi^2 - \frac{\lambda}{2}\right)\|\mathbf{X}^{t+1} - \mathbf{X}^t\|_F^2. \quad (26)$$

The proof of Lemma 2 is shown in Appendix C. According to Lemma 2, when the penalty factor satisfies $\lambda > \sqrt{2}L_\phi$, it can be observed that the augmented Lagrangian decreases monotonically.

Note that the projected gradient descent (PGD) is used to update variable $\mathbf{X}$, and the update policy of each iteration can be rewritten as

$$\mathbf{X}^{t+1} \in \mathcal{P}_{\mathcal{A}}\left(\mathbf{X}^t - \frac{1}{\lambda}\nabla g\left(\mathbf{X}^t\right)\right) \in \underset{\boldsymbol{\epsilon}\in\mathcal{A}}{\text{argmin}}\left\|\boldsymbol{\epsilon} - \left(\mathbf{X}^t - \frac{1}{\lambda}\nabla g\left(\mathbf{X}^t\right)\right)\right\|_F^2. \quad (27)$$

A common practice to show the convergence of one non-convex problem is to prove that the algorithm can converge to a stationary point instead. In the following theorem, we conclude that our proposed algorithm indeed can converge to a stationary point with respect to λ.

Theorem 1: Let $(\mathbf{X}^*, \mathbf{R}^*, \mathbf{V}^*)$ denote a limit point. Given $\lambda > \sqrt{2}L_\phi$, we can show that $\mathbf{X}^*$ is a stationary point of the optimization problem (P5) with respect to factor λ.

Proof: Let $(\mathbf{X}^*, \mathbf{R}^*, \mathbf{V}^*)$ be the limit point of $(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t)$, and $(\mathbf{X}^{t_m}, \mathbf{R}^{t_m}, \mathbf{V}^{t_m})$ be the sub-sequence that converges to $(\mathbf{X}^*, \mathbf{R}^*, \mathbf{V}^*)$, for $m = 0, 1, \ldots, M$. Then, we have

$$\lim_{m\to\infty}\mathbf{X}^{t_m} = \mathbf{X}^*, \quad (28a)$$

$$\lim_{m\to\infty}\mathbf{R}^{t_m} = \mathbf{R}^*, \quad (28b)$$

$$\lim_{m\to\infty}\mathbf{V}^{t_m} = \mathbf{V}^*. \quad (28c)$$

First, we can rewrite the $(T+1)$-th iteration of the augmented Lagrangian function as

$$\mathcal{L}_\lambda(\mathbf{X}^{T+1}, \mathbf{R}^{T+1}, \mathbf{V}^{T+1}) = \mathcal{L}_\lambda(\mathbf{X}^0, \mathbf{R}^0, \mathbf{V}^0) + \sum_{t=0}^{T}\left[\mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^{t+1}) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t)\right]$$

$$\overset{(*)}{\leq} \mathcal{L}_\lambda(\mathbf{X}^0,\mathbf{R}^0,\mathbf{V}^0)+\left(\frac{1}{\lambda}L_\phi^2-\frac{\lambda}{2}\right)\sum_{t=0}^{T}\|\mathbf{X}^{t+1}-\mathbf{X}^t\|_F^2, \tag{29}$$

where $\overset{(*)}{\leq}$ is obtained according to Lemma 2. Since the augmented Lagrangian is lower bounded by Lemma 1, we can further obtain

$$g_{min} \leq \mathcal{L}_\lambda(\mathbf{X}^0,\mathbf{R}^0,\mathbf{V}^0)+\left(\frac{1}{\lambda}L_\phi^2-\frac{\lambda}{2}\right)\sum_{t=0}^{T}\|\mathbf{X}^{t+1}-\mathbf{X}^t\|_F^2. \tag{30}$$

With the condition that $\lambda>\sqrt{2}L_\phi$, we find that the right hand side of the inequality (30) is always decreasing. The only way to satisfy the inequality is that $\lim_{T\to\infty}\sum_{t=0}^{T}\|\mathbf{X}^{t+1}-\mathbf{X}^t\|_F^2=0$. Therefore, we obtain $\lim_{t\to\infty}\|\mathbf{X}^{t+1}-\mathbf{X}^t\|_F=0$. Combining with (28a), we derive

$$\lim_{m\to\infty}\mathbf{X}^{t_m+1}=\mathbf{X}^*. \tag{31}$$

Besides, utilizing Lemma 3 in Appendix A, we have

$$\lim_{m\to\infty}\mathbf{V}^{t_m+1}=-\lim_{m\to\infty}\nabla g\left(\mathbf{X}^{t_m+1}\right)=-\nabla g\left(\mathbf{X}^*\right)=\mathbf{V}^*. \tag{32}$$

Therefore, we obtain

$$\lim_{m\to\infty}\mathbf{V}^{t_m+1}=\mathbf{V}^*. \tag{33}$$

According to (27), we have

$$\mathbf{R}^{t_m+1}\in\underset{\boldsymbol{\epsilon}\in\mathcal{A}}{\operatorname{argmin}}\left\|\boldsymbol{\epsilon}-\left(\mathbf{X}^{t_m}+\frac{1}{\lambda}\mathbf{V}^{t_m}\right)\right\|_F^2. \tag{34}$$

Besides, (15d) can be rewritten as

$$\mathbf{V}^{t_m+1}-\mathbf{V}^{t_m}=\lambda\left(\mathbf{X}^{t_m+1}-\mathbf{R}^{t_m+1}\right). \tag{35}$$

Take the limit of both sides in (35). Since we obtain that $\mathbf{V}^{t_m+1}$ can converge to $\mathbf{V}^*$, the limit of left hand side in (35) is zero. Hence, we obtain

$$\lim_{m\to\infty}\mathbf{R}^{t_m+1}=\mathbf{X}^*. \tag{36}$$

Therefore, taking the limit of (34), we obtain that

$$\mathbf{X}^*\in\underset{\boldsymbol{\epsilon}\in\mathcal{A}}{\operatorname{argmin}}\left\|\boldsymbol{\epsilon}-\left(\mathbf{X}^*-\frac{1}{\lambda}\nabla g\left(\mathbf{X}^*\right)\right)\right\|_F^2, \tag{37}$$

which means that the limit point cannot be locally improved with step size $\frac{1}{\lambda}$. Then, we can conclude that any limit point is a stationary point related to penalty factor λ. ■

According to Theorem 1, we conclude that our proposed algorithm can converge to a stationary point with the condition $\lambda>\sqrt{2}L_\phi$.

B. Complexity Analysis

Next, we analyze the computational complexity of the proposed algorithm. For simplicity, the computational complexity is calculated by the number of multiplications. The complexities of the proposed algorithm and other baselines are shown in Table I.

TABLE I
COMPUTATIONAL COMPLEXITY

Precoding	The number of multiplications
WF	$N\left(\frac{1}{3}U^3+N_tU^2+2U^2-\frac{1}{3}U+N_t\log_2 N\right)$
SQUID	$N\left(\frac{5}{6}U^3+\frac{3}{2}N_tU^2+3N_tU-\frac{1}{3}U\right)+2TNN_t(U+\log_2 N)$
QCESLP	$N^2N_tU+T\left(2N^2N_t^2+N^2N_tU+NU\right)$
CG-MC1bit	$NN_tU^2+TN\left(\frac{2}{3}U^3+4N_tU+2U^2+3U+2N_t\log_2 N\right)$

It can be observed that the main complexity of our proposed algorithm is caused by operations of the fast Fourier transform (FFT) and the inverse FFT (IFFT) ($\mathcal{O}(N\log_2 N)$), which cannot be avoided in multi-carrier systems. The algorithm is of affordable computational complexity. Compared to other baselines, the complexity of the proposed CG-MC1bit is at the same level with the SQUID method, and greatly lower than the QCESLP method.

In Appendix D, we describe different approximation methods to bring down the computational complexity further. Note that those approximations are especially useful in the case of large scale antenna arrays.

V. NUMERICAL RESULTS

In this section, the performance of the proposed 1-bit nonlinear precoding algorithm is analyzed through simulations.

Rayleigh fading channel model is assumed at each discrete time n, i.e., each element in the channel obeys the zero mean complex Gaussian distribution. Denote $\mathbf{h}_i^T[n]$ as the i-th row of $\mathbf{H}_T[n]$, which represents the channel of the i-th user, and $\mathbf{h}_i[n]\in\mathbb{C}^{N_t\times 1}$. Each element in $\mathbf{h}_i[n]$ follows i.i.d. $\mathcal{CN}(0,\sqrt{\rho_i}/I)$, where ρ_i denotes the corresponding path-loss parameter. Simulation parameters are summarized in Table II, and the path-loss parameters are set according to the 3GPP specification [32]. Besides, variables $\mathbf{X}$ and $\mathbf{R}$ are initialized to all-one matrices, and the dual variable $\mathbf{V}$ is initialized to all-zero matrix. We assume the maximum likelihood detector is used at the receiver side.

Nonlinear algorithms proposed in [24] and [25] are denoted as "SQUID" and "QCESLP" for performance benchmarks. Our proposed algorithm is denoted as "CG-MC1bit". The WF precoding with infinite resolution DACs (denoted as "WF-inf") and with 1-bit DACs (denoted as "WF-1bit") serve as baselines corresponding to the ideal case and the worst case, respectively.

A. Convergence Performance

First, we intend to verify the convergence of the proposed algorithm. The average MSE on all subcarriers in (8) is used to describe the convergence performance of the proposed algorithm. Different numbers of transmit antennas $N_t=64$, $N_t=256$ and $N_t=1024$ are used to test the convergence of the proposed algorithm. The 16QAM signals with the number of subcarriers $N=1024$, the number of users $U=8$ are adopted. The result is shown in Fig. 2, which corresponds to SNR of 0 dB. It can be observed that the proposed algorithm converges under different cases, which also verifies

TABLE II
SIMULATION PARAMETERS

Parameters	Values
Path-loss L_i of $\mathbf{h}_i^T[n]$	$35.6 + 36.7 \lg d_i$ dB
Bandwidth BW	180 kHz
Noise power spectral density N_0	-170 dBm/Hz
Path-loss parameter ρ_i	$N_0 BW + L_i$ dB
Penalty factor λ	0.01
Number of subcarriers N	1024
Number of transmit antennas N_t	$64, 256$
Taps of channels I	10
Number of users U	8
Distance of users d_i	$100 - 200$ m
Type of modulation	QPSK, 16QAM
Monte-Carlo Iterations	10^3

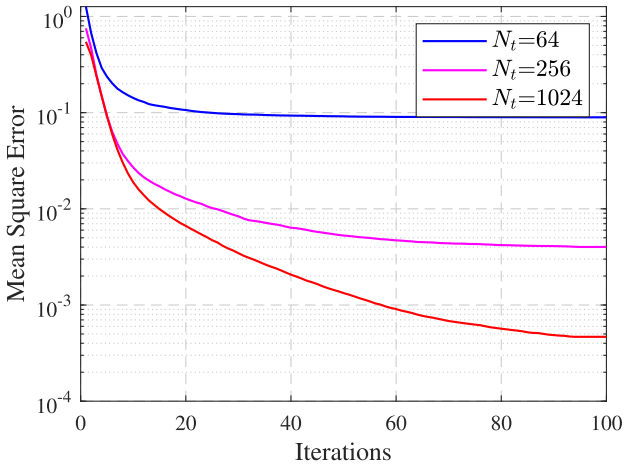
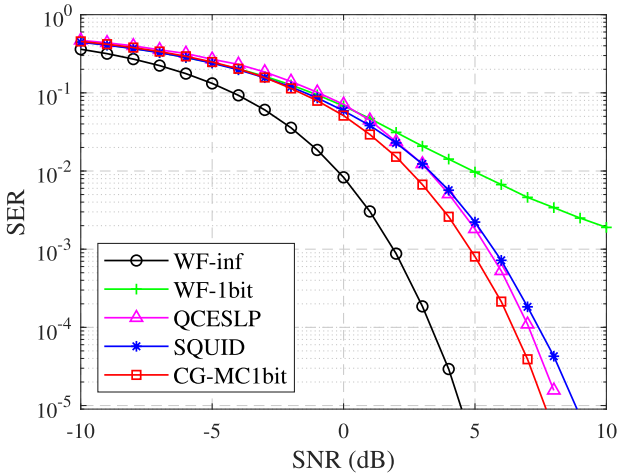
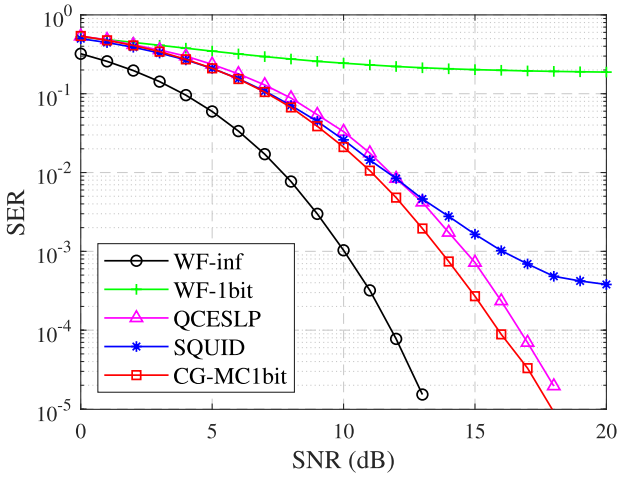


Fig. 2. Convergence analysis.



(a) $N_t = 64, N = 1024, U = 8$, QPSK.



(b) $N_t = 64, N = 1024, U = 8$, 16QAM.

Fig. 3. SER performance with $N_t = 64, N = 1024, U = 8$, QPSK and 16QAM signals.

the theoretical convergence analysis in Section IV-A. The more transmit antennas the BS uses, the lower level total MSE on all subcarriers can converge to. The result indicates that the proposed algorithm can converge with different numbers of antennas. Besides, more transmit antennas are beneficial to the MSE performance.

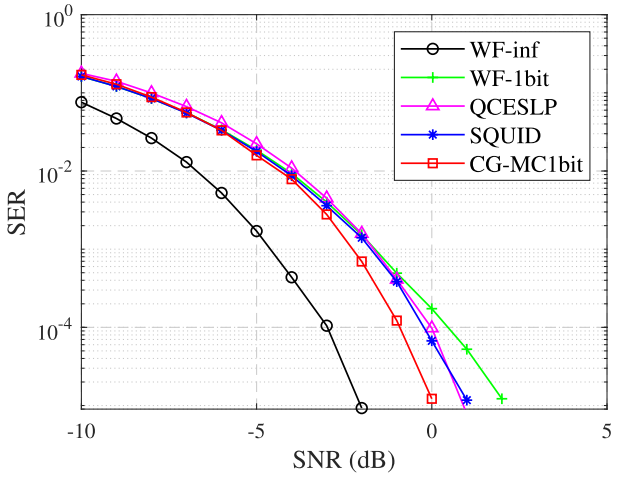
B. Demodulation Performance

In this part, we present the demodulation performance in terms of symbol error rate (SER) at the receiver side. In the simulations, the proposed CG-MC1bit algorithm is drawn by the red line with square. The black line with circle represents the method of WF with infinite resolution DACs, and the green line with plus sign denotes the WF with 1-bit DACs. The magenta line with triangle and blue line with asterisk represent the method of QCESLP and SQUID, respectively.

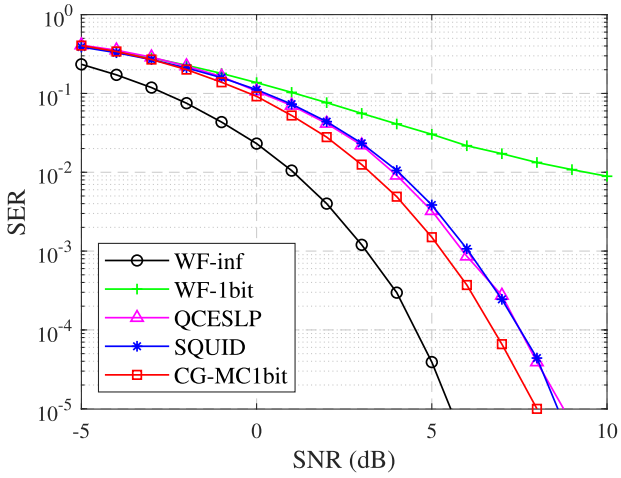
First, we intend to show the system performance of the proposed algorithm in terms of SER. The performance is analyzed with the set of parameters $N_t = 64$, $N = 1024$, $U = 8$. The distance between the BS and all users is assumed similar, so channels of different users own similar path-losses. Both the QPSK and 16QAM signals are simulated, results are shown in Fig. 3. In Fig 3(a), the QPSK signals are adopted. It is observed that the linear precoding WF-1bit performs the worst and the error floor appears. The SER cannot be further reduced at higher SNR. Poor performance is obtained since the design of linear precoding doesn't take account of the severe degradation caused by 1-bit DACs. In contrast, the SERs of nonlinear precoding methods always decrease with SNR in this case. The trends are consistent with the infinite resolution WF precoding, which shows that the nonlinear precoding can compensate for the loss of 1-bit quantization at high SNR.

Moreover, the proposed CG-MC1bit shows the best performance among existing approaches. The SERs of CG-MC1bit, QCESLP and SQUID are 3.91×10^{-5}, 1.09×10^{-4} and 1.83×10^{-4} at SNR of 7 dB, respectively. The performance is about 1 dB better than SQUID, and about 0.5 dB better than QCESLP. At SNR of 8 dB in the above setting, the SER of our proposed algorithm can drop below 10^{-5}. This better performance with our proposed scheme indeed aligns with the previous theoretical analysis, and the main discriminator is that our proposed algorithm handles the non-convex constraint in a more systematic manner. The gap between our proposed algorithm and WF-inf is about 3 dB.

In Fig. 3(b), the 16QAM signals are adopted. Compared to the case of QPSK signals in Fig. 3(a), the degradation caused by 1-bit DACs becomes more severe. In this case, the WF precoding with 1-bit DACs does little help in terms of the SER performance. The gap between our proposed algorithm and WF precoding with infinite resolution DACs is about 4 dB,



(a) $N_t = 256, N = 1024, U = 8$, QPSK.



(b) $N_t = 256, N = 1024, U = 8$, 16QAM.

Fig. 4. SER performance with $N_t = 256, N = 1024, U = 8$, QPSK and 16QAM signals.

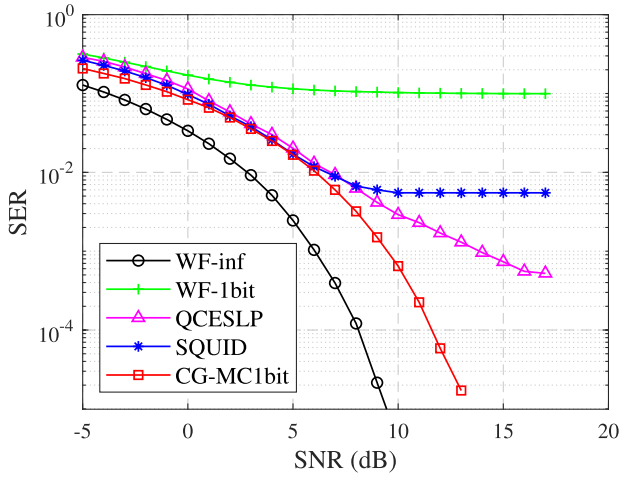


Fig. 5. SER performance with different path-losses.

which is larger compared to the case with the QPSK signals. This makes sense since higher order modulations demand better signal quality. We observe that error floor appears in SQUID around the level of 10^{-4}. This is due to the fact that final quantization of the relaxed solutions in the SQUID scheme leads to extra SER losses. Besides, the performance of CG-MC1bit is 0.8 dB better than QCESLP, which is the best among nonlinear methods. The SER can drop below the level of 10^{-5} at high SNR. The result indicates that the proposed CG-MC1bit algorithm is applicable to different types of modulation.

Next, we study the impact of transmit antennas on the SER performance. The parameter setting is $N_t = 256$, $N = 1024$, $U = 8$. The result is demonstrated in Fig. 4. Fig. 4(a) and Fig. 4(b) correspond to QPSK and 16QAM signals, respectively. Better performance is achieved compared to the case of $N_t = 64$ in Fig. 3. It indicates that more antennas at the BS benefit the SER performance. In Fig. 4(a), even the WF precoding with 1-bit DACs can achieve satisfactory performance. The error floor disappears at the level of 10^{-5}, which is consistent with the result in [13]. The gap between CG-MC1bit and WF-inf is about 2 dB, which is closer than that with $N_t = 64$ in Fig. 3(a). The algorithm CG-MC1bit is almost 0.8 dB better than QCESLP and SQUID. It can be inferred that, with sufficiently large number of transmit antennas, the degradation due to 1-bit DACs can be small enough and the SER performance can be closer to the ideal case of WF with infinite resolution DACs.

Moreover, the performance with 16QAM signals is shown in Fig. 4(b). The linear WF precoding with 1-bit DACs achieves poor performance with higher order modulation compared to QPSK signals in Fig. 4(a). It is clear that a simple superposition of 1-bit DACs and linear precoding gives the worst performance in the case of higher order modulations. Moreover, our proposed CG-MC1bit algorithm outperforms other existing approaches. The gap between CG-MC1bit and WF precoding with infinite resolution DACs is nearly the same compared to the QPSK signal in Fig. 4(a). It implies that our proposed algorithm is applicable to different modulation schemes, and more antennas are beneficial to the SER performance.

Next, we analyze the system performance when UEs experience different path-losses. In this simulation, the parameter setting is $N_t = 64$, $N = 1024$, $U = 8$ with QPSK signals. Furthermore, the path-losses to these $U = 8$ users are set such that user i has a SNR of $(\text{SNR}_0 + 2i)$ dB, where $i \in \{0, 1, \ldots, 7\}$. In Fig. 5, the x-axis corresponds to the value of SNR_0. Fig. 5 shows the SER performance with different path-losses of users. We see the performance diverges at high SNRs. We also observe that all algorithms with 1-bit DACs exhibit error floors except our proposed CG-MC1bit. In particular, the proposed CG-MC1bit algorithm has the same trend as Fig. 3(a), and the SER always decreases with SNR. By taking into account the different path-losses to different users, the proposed CG-MC1bit algorithm offers a satisfactory performance as in Fig. 5. Besides, the gap between CG-MC1bit and WF precoding with infinite resolution DACs is about 3.5 dB. The result validates that our proposed algorithm is applicable to the case of users with different path-losses.

In the following, we study the impact of the resolution level of DACs. The parameter setting is $N_t = 64$, $N = 1024$, $U = 8$ with QPSK signals, which is the same as Fig. 3(a). The cases with 1-bit, 2-bit, and 4-bit DACs are tested. The SER performance is shown in Fig. 6. We observe

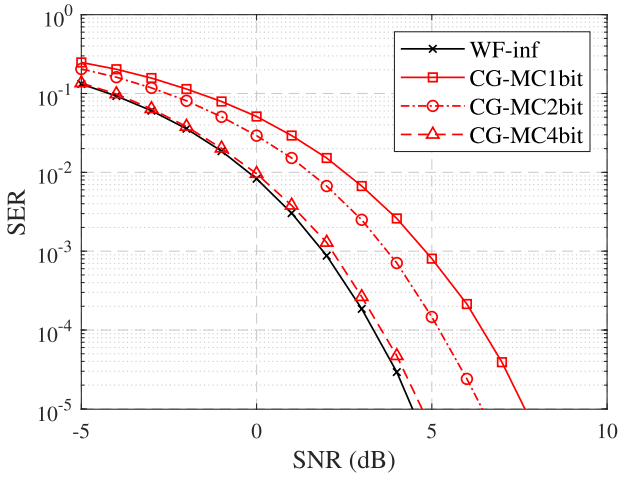


Fig. 6. SER performance with different resolution levels of DACs.

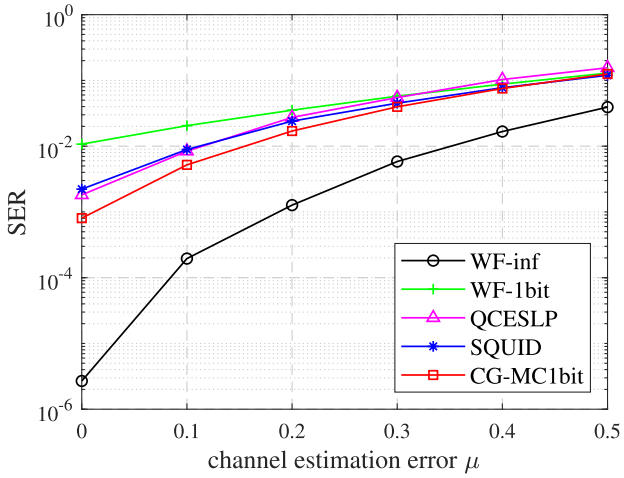


Fig. 7. SER performance with channel estimation error.

TABLE III
RUNNING TIME COMPARISON

Precoding	Running time of each Monte-Carlo iteration
WF	0.128s
SQUID	0.608s
QCESLP	68.924s
CG-MC1bit	0.773s

that our proposed algorithm achieves better SER performance when DACs have finer resolutions. This corroborates that our proposed framework indeed applies to multi-carrier systems with multi-bit DACs as well. Moreover, with 4-bit DACs, the proposed algorithm can nearly achieve the performance of optimal WF precoding. The result indicates that systems with 4-bit DACs are enough for the proposed CG-MC1-bit algorithm.

Next, we want to test the computational complexity of the proposed algorithm via running time. The CPU running time is used as the performance metric, where the dual-core CPU is of 3.60 GHz, and the memory is of 8 GB. Table III shows the average running time, where SNR$=5$ dB and other simulation conditions are the same as Fig. 3(a). The running time is compared for each Monte-Carlo iteration. It is observed that the WF precoding runs the fastest since there are no iterations. However, the WF precoding shows poor performance with 1-bit DACs. The running time of CG-MC1bit is slightly slower than SQUID, since our proposed method considers to update adjustment factors for each user in the iteration. Besides, the running time of QCESLP is much longer, since the matrix dimension is huge in QCESLP. The result indicates that our proposed algorithm can achieve better performance with lower or competitive computational complexity.

C. Robustness Comparison

In this subsection, we test the robustness of the proposed algorithm in terms of channel estimation error. The channel with error at discrete time n can be represented as follows,

$$\check{\mathbf{H}}_T[n] = \sqrt{1-\mu}\mathbf{H}_T[n] + \sqrt{\mu}\Delta\mathbf{H}_T[n], \tag{38}$$

where $\check{\mathbf{H}}_T[n]$ is the actual channel, $\Delta\mathbf{H}_T[n] \sim \mathcal{CN}(0,1)$ denotes the channel estimation error, and μ denotes the severity of channel estimation error. The parameter $\mu = 0$ represents perfect channel estimation, and $\mu = 1$ means that the obtained CSI is totally wrong.

We investigate the robustness of our proposed CG-MC1bit algorithm against channel estimation error. The parameter setting is the same as those in Fig. 3(a), and we focus on the SER performance at SNR $= 5$ dB. The simulation result is shown in Fig. 7. We observe that the SERs of all algorithms become worse with more severe channel estimation error. The performance of all methods is similar with that under perfect CSI. CG-MC1bit still performs the best among existing nonlinear precoding methods. It validates the robustness of the proposed algorithm against channel estimation error.

Besides, the gap among all algorithms becomes closer when the channel estimation error increases. Comparing to Fig. 3(a), we find that there is 1.5 dB performance degradation with the channel estimation error level $\mu = 0.1$. Similarly, with channel error $\mu = \{0.2,\ 0.3,\ 0.4,\ 0.5\}$, about $\{3,\ 4.5,\ 5.5,\ 7.5\}$ dB performance loss will come up, respectively. The result is in accordance with intuition, where the channel estimation error always degrades the system performance.

VI. CONCLUSION

In this paper, we have studied the precoding design for the downlink of massive MU-MIMO OFDM systems tailored to 1-bit DACs. A novel nonlinear precoding algorithm, i.e., the CG-MC1bit in Algorithm 1, has been proposed to compensate for the distortion caused by the 1-bit quantization of the DACs. We have also demonstrated that the proposed CG-MC1bit algorithm would converge to a stationary point under certain conditions. Moreover, our proposed algorithm is applicable to users with different path-losses. Meanwhile, we have made appropriate approximations and been able to reduce the computational complexity further. Through comprehensive simulations with different parameter settings and modulation schemes, we have further shown that our proposed CG-MC1bit algorithm exhibited the best performance compared to other state-of-the-art nonlinear 1-bit precoding schemes. With massive MIMO being deployed more and more, we envision that the proposed CG-MC1bit algorithm in this paper will have

great potential in facilitating cost reduction while ensuring superior system performance.

APPENDIX A
AUXILIARY LEMMA

Here, we introduce a useful lemma, which is helpful to the following analysis.

Lemma 3: For $t > 0$, we have

$$\mathbf{V}^t = -\nabla g\left(\mathbf{X}^t\right). \tag{39}$$

Proof: The update of $\mathbf{X}$ in equation (15a) is equal to solve the problem as (40) in the following,

$$\nabla \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^t) = 0. \tag{40}$$

Then, we obtain that

$$\nabla g\left(\mathbf{X}^{t+1}\right) + \mathbf{V}^t + \lambda\left(\mathbf{X}^{t+1} - \mathbf{R}^{t+1}\right) = 0. \tag{41}$$

Utilizing equation (15d),

$$\nabla g\left(\mathbf{X}^{t+1}\right) + \mathbf{V}^{t+1} = 0. \tag{42}$$

Hence, we derive the relationship as shown in equation (39). ∎

APPENDIX B
PROOF OF LEMMA 1

Proof: Utilizing Lemma 3, the augmented Lagrangian function (14) with variable $\mathbf{R}^t$ in discrete set $\mathcal{A}$ can be rewritten as

$$\mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t) = g\left(\mathbf{X}^t\right) + \left\langle \nabla g(\mathbf{X}^t), \mathbf{R}^t - \mathbf{X}^t \right\rangle + \frac{\lambda}{2}\|\mathbf{X}^t - \mathbf{R}^t\|_F^2. \tag{43}$$

By the Lagrange mean value theorem, we have

$$\begin{aligned} & g\left(\mathbf{R}^t\right) \\ &= g\left(\mathbf{X}^t\right) + \int_0^1 \left\langle \nabla g\left[\mathbf{X}^t + \tau\left(\mathbf{R}^t - \mathbf{X}^t\right)\right], \mathbf{R}^t - \mathbf{X}^t \right\rangle d\tau \\ &= g\left(\mathbf{X}^t\right) + \left\langle \nabla g(\mathbf{X}^t), \mathbf{R}^t - \mathbf{X}^t \right\rangle \\ &\quad + \int_0^1 \left\langle \nabla g\left[\mathbf{X}^t + \tau\left(\mathbf{R}^t - \mathbf{X}^t\right)\right] - \nabla g\left(\mathbf{X}^t\right), \mathbf{R}^t - \mathbf{X}^t \right\rangle d\tau. \end{aligned}$$

Then,

$$\begin{aligned} & \left| g\left(\mathbf{R}^t\right) - g\left(\mathbf{X}^t\right) - \left\langle \nabla g(\mathbf{X}^t), \mathbf{R}^t - \mathbf{X}^t \right\rangle \right| \\ &\le \int_0^1 \left| \left\langle \nabla g\left[\mathbf{X}^t + \tau\left(\mathbf{R}^t - \mathbf{X}^t\right)\right] - \nabla g\left(\mathbf{X}^t\right), \mathbf{R}^t - \mathbf{X}^t \right\rangle \right| d\tau \\ &\le \int_0^1 \left\| \nabla g\left[\mathbf{X}^t + \tau\left(\mathbf{R}^t - \mathbf{X}^t\right)\right] - \nabla g\left(\mathbf{X}^t\right) \right\| \|\mathbf{R}^t - \mathbf{X}^t\| d\tau \\ &\overset{(*)}{\le} \int_0^1 \tau L_\phi \|\mathbf{R}^t - \mathbf{X}^t\|_F^2 d\tau \\ &= \frac{L_\phi}{2}\|\mathbf{R}^t - \mathbf{X}^t\|_F^2, \end{aligned}$$

where $\overset{(*)}{\le}$ holds due to the definition of Lipschitz condition,

$$\|\nabla g(\mathbf{X}^{t+1}) - \nabla g(\mathbf{X}^t)\|_2 \le L_\phi \left\|\mathbf{X}^{t+1} - \mathbf{X}^t\right\|_F. \tag{44}$$

Hence, we obtain that

$$\mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t) \ge g\left(\mathbf{R}^t\right) + \frac{\lambda - L_\phi}{2}\|\mathbf{R}^t - \mathbf{X}^t\|_F^2. \tag{45}$$

Since we have already known that the function $g(\cdot)$ is lower bounded on discrete set $\mathcal{A}$ according to Lemma 1. Therefore, we obtain that the augmented Lagrangian function is lower bounded, where

$$\mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t) \ge g_{min} + \frac{\lambda - L_\phi}{2}\|\mathbf{R}^t - \mathbf{X}^t\|_F^2. \tag{46}$$

∎

APPENDIX C
PROOF OF LEMMA 2

Proof: First, the difference of the augmented Lagrangian function between the successive $(t+1)$-th and t-th iteration is written as

$$\begin{aligned} & \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^{t+1}) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t) \\ &= \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^{t+1}) - \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^t) \\ &\quad + \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^t) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t). \end{aligned} \tag{47}$$

Denote $\mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^{t+1}) - \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^t)$ as A_1 and $\mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^t) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t)$ as A_2. Then, A_1 can be further simplified in the following,

$$\begin{aligned} A_1 &= \left\langle \mathbf{V}^{t+1}, \mathbf{X}^{t+1} - \mathbf{R}^{t+1} \right\rangle - \left\langle \mathbf{V}^t, \mathbf{X}^{t+1} - \mathbf{R}^{t+1} \right\rangle \\ &= \left\langle \mathbf{V}^{t+1} - \mathbf{V}^t, \mathbf{X}^{t+1} - \mathbf{R}^{t+1} \right\rangle \\ &= \left\langle \mathbf{V}^{t+1} - \mathbf{V}^t, \frac{1}{\lambda}\left(\mathbf{V}^{t+1} - \mathbf{V}^t\right) \right\rangle \\ &= \frac{1}{\lambda}\left\|\mathbf{V}^{t+1} - \mathbf{V}^t\right\|_F^2 \\ &= \frac{1}{\lambda}\left\|\nabla g(\mathbf{X}^{t+1}) - \nabla g(\mathbf{X}^t)\right\|_F^2 \\ &\le \frac{1}{\lambda} L_\phi^2 \left\|\mathbf{X}^{t+1} - \mathbf{X}^t\right\|_F^2. \end{aligned} \tag{48}$$

Term A_2 can also be rewritten similarly as the form of (47),

$$\begin{aligned} A_2 &= \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^t) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^{t+1}, \mathbf{V}^t) \\ &\quad + \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^{t+1}, \mathbf{V}^t) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t). \end{aligned} \tag{49}$$

Since $\mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^{t+1}, \mathbf{V}^t) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t) \le 0$, we further obtain that

$$\begin{aligned} A_2 &\le \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^t) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^{t+1}, \mathbf{V}^t) \\ &\le -\frac{\lambda}{2}\left\|\mathbf{X}^{t+1} - \mathbf{X}^t\right\|_F^2. \end{aligned} \tag{50}$$

Hence, we obtain that

$$\begin{aligned} & \mathcal{L}_\lambda(\mathbf{X}^{t+1}, \mathbf{R}^{t+1}, \mathbf{V}^{t+1}) - \mathcal{L}_\lambda(\mathbf{X}^t, \mathbf{R}^t, \mathbf{V}^t) \\ &= A_1 + A_2 \\ &\le \left(\frac{1}{\lambda}L_\phi^2 - \frac{\lambda}{2}\right)\|\mathbf{X}^{t+1} - \mathbf{X}^t\|_F^2. \end{aligned} \tag{51}$$

∎

APPENDIX D
APPROXIMATION METHOD

We also propose some approximation methods to further reduce the computational complexity. The key complexity is caused by the FFT and IFFT calculation, which cannot be avoided in multi-carrier systems. Besides, the calculation of (17) is another time-consuming procedure due to the matrix inversion. Thus, we intend to reduce the complexity of the matrix inversion in (17) through approximation.

Here the Gauss-Seidel iterative method [33] is used to approximate the matrix inversion. Denote $\mathbf{W}[k] = \tilde{\mathbf{H}}[k]\tilde{\mathbf{H}}^H[k] + \frac{\lambda}{2}\mathbf{I}_U$. The equation (17) can be rewritten as

$$\bar{\mathbf{x}}^{t+1}[k] = \frac{1}{\lambda}\mathbf{b}[k] - \frac{1}{\lambda}\tilde{\mathbf{H}}^H[k]\mathbf{t}[k], \tag{52}$$

where $\mathbf{b}[k] = 2\tilde{\mathbf{H}}^H[k]\mathbf{s}[k] + \lambda\mathbf{R}^t[k] - \mathbf{V}^t[k]$, and $\mathbf{t}[k] = \mathbf{W}^{-1}[k]\tilde{\mathbf{H}}[k]\mathbf{b}[k]$. Then, according to the LDL decomposition, $\mathbf{W}[k]$ can be decomposed as

$$\mathbf{W}[k] = \mathbf{L}[k] + \mathbf{D}[k] + \mathbf{L}^H[k], \tag{53}$$

where $\mathbf{L}[k], \mathbf{D}[k] \in \mathbb{C}^{U\times U}$ denote the lower triangular matrix and diagonal matrix, respectively. Thus, $\mathbf{t}[k]$ can be approximated via iterations, where the $(i+1)$-th iteration is written as

$$\mathbf{t}^{i+1}[k] = (\mathbf{D}[k] + \mathbf{L}[k])^{-1}\left(\tilde{\mathbf{H}}[k]\mathbf{b}[k] - \mathbf{L}^H[k]\mathbf{t}^{(i)}[k]\right). \tag{54}$$

By the law of large numbers in [34], $\mathbf{W}[k]$ is a diagonal dominant matrix. With more transmit antennas used at the BS, $\mathbf{W}[k]$ becomes more diagonal dominant, and more rapid the iterative method converges [33].

The complexity is reduced regarding to the matrix inversion in (17), where the number of multiplications is $\frac{2}{3}U^3$ without approximation according to [35]. During the i-th iteration, the number of multiplications can be reduced to $iU^2 + N_tU + N_t$. Accordingly, the total number of multiplications becomes $NN_tU^2 + TN(5N_tU + (2+i)U^2 + N_t + 3U + 2N_t\log_2 N)$, and the running time of each Monte-Carlo iteration is 0.658s.

Furthermore, with sufficiently large numbers of antennas, the matrix $\mathbf{W}[k]$ can be simply approximated as follows according to the law of large numbers,

$$\begin{aligned}\mathbf{W}[k] &\xrightarrow{a.s.} \mathbf{D}[k] \\ &= \mathrm{Diag}\left[\alpha_1^2 N_t + \frac{\lambda}{2}, \alpha_2^2 N_t + \frac{\lambda}{2}, \ldots, \alpha_U^2 N_t + \frac{\lambda}{2}\right].\end{aligned} \tag{55}$$

By exploiting the above approximation, the total number of multiplications becomes $TN(3N_tU + 5U + 2N_t\log_2 N)$, and the running time of each Monte-Carlo iteration is 0.583s.

The performance of approximation methods is tested with $N = 1024, U = 8$, QPSK signals. The number of transmit antennas $N_t = 64$ and $N_t = 256$ are simulated. The simulation result is shown in Fig. 8. The "appr. 1" and "appr. 2" denote the approximation schemes using the iteration method and the law of large numbers, respectively. It can be observed that these two schemes have similar performance loss. With more transmit antennas, the degradation caused by approximation becomes smaller. Note that the SER can be further reduced with proper coding [36]. The result indicates that these approximation methods are meaningful in massive MIMO systems.

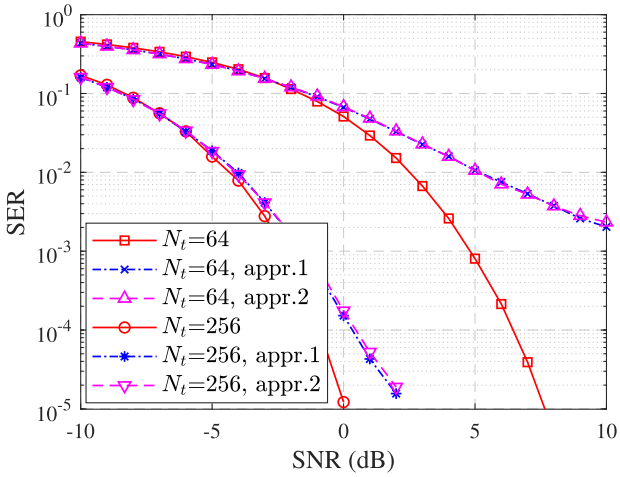


Fig. 8. SER performance of approximation methods.

REFERENCES

[1] T. L. Marzetta, "Massive MIMO: An introduction," *Bell Labs Tech. J.*, vol. 20, pp. 11–22, 2015.

[2] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 836–869, 2nd Quart., 2018.

[3] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Feb. 2016.

[4] T. S. Rappaport, R. W. Heath, R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Upper Saddle River, NJ, USA: Prentice-Hall, 2015.

[5] F. Rusek et al., "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[6] A. F. Molisch et al., "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, Sep. 2017.

[7] A. Li, C. Masouros, A. L. Swindlehurst, and W. Yu, "1-bit massive MIMO transmission: Embracing interference with symbol-level precoding," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 121–127, May 2021.

[8] L. N. Ribeiro, S. Schwarz, M. Rupp, and A. L. F. de Almeida, "Energy efficiency of mmWave massive MIMO precoding with low-resolution DACs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 298–312, May 2018.

[9] M. Parker, *Digital Signal Processing 101: Everything You Need to Know to Get Started*, 2nd ed. Amsterdam, The Netherlands: Newnes, 2017.

[10] M. A. Albreem, A. H. A. Habbash, A. M. Abu-Hudrouss, and S. S. Ikki, "Overview of precoding techniques for massive MIMO," *IEEE Access*, vol. 9, pp. 60764–60801, 2021.

[11] O. Bin Usman, H. Jedda, A. Mezghani, and J. A. Nossek, "MMSE precoder for massive MIMO using 1-bit quantization," in *Proc. IEEE ICASSP*, Shanghai, China, May 2016, pp. 3381–3385.

[12] 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, Jun. 2017.

[13] S. Jacobsson, G. Durisi, M. Coldrey, and C. Studer, "Massive MU-MIMO-OFDM downlink with one-bit DACs and linear precoding," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[14] R. Liu, H. Li, and M. Li, "Symbol-level hybrid precoding in mmWave multiuser MISO systems," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1636–1639, Sep. 2019.

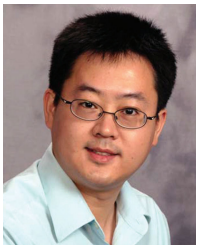
[15] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized precoding for massive MU-MIMO," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4670–4684, Nov. 2017.
[16] O. Castañeda, S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "1-bit massive MU-MIMO precoding in VLSI," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 4, pp. 508–522, Dec. 2017.
[17] C.-J. Wang, C.-K. Wen, S. Jin, and S.-H. Tsai, "Finite-alphabet precoding for massive MU-MIMO with low-resolution DACs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4706–4720, Jul. 2018.
[18] L. Chu, F. Wen, L. Li, and R. Qiu, "Efficient nonlinear precoding for massive MIMO downlink systems with 1-bit DACs," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4213–4224, Sep. 2019.
[19] J.-C. Chen, "Alternating minimization algorithms for one-bit precoding in massive multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7394–7406, Aug. 2018.
[20] 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, Sep. 2018.
[21] 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.
[22] 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.
[23] F. Askerbeyli, H. Jedda, and J. A. Nossek, "1-bit precoding in massive MU-MISO-OFDM downlink with linear programming," in *Proc. 23rd Int. ITG Workshop Smart Antennas (WSA)*, Apr. 2019, pp. 257–261.
[24] S. Jacobsson, O. Castañeda, C. Jeon, G. Durisi, and C. Studer, "Nonlinear precoding for phase-quantized constant-envelope massive MU-MIMO-OFDM," in *Proc. IEEE Int. Conf. Telecommun. (ICT)*, Jun. 2018, pp. 367–372.
[25] C. G. Tsinos, S. Domouchtsidis, S. Chatzinotas, and B. Ottersten, "Symbol level precoding with low resolution DACs for constant envelope OFDM MU-MIMO systems," *IEEE Access*, vol. 8, pp. 12856–12866, 2020.
[26] O. Castaneda, S. Jacobsson, G. Durisi, T. Goldstein, and C. Studer, "Finite-alphabet Wiener filter precoding for mmWave massive MU-MIMO systems," in *Proc. 53rd Asilomar Conf. Signals, Syst., Comput.*, Nov. 2019, pp. 178–183.
[27] A. Tabeshnezhad, A. L. Swindlehurst, and T. Svensson, "Reduced complexity precoding for one-bit signaling," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1967–1971, Feb. 2021.
[28] 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.
[29] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[30] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Nov. 2011.
[31] T. Huang, P. Singhania, M. Sanjabi, P. Mitra, and M. Razaviyayn, "Alternating direction method of multipliers for quantization," in *Proc. 24th Int. Conf. Artif. Intell. Statist. (AISTATS)*, Apr. 2021, pp. 208–216.
[32] *Further Advancements for E-UTRA Physical Layer Aspects (Release 9)*, document 3GPP TS 36.814, Mar. 2010.
[33] Y. Liu, J. Liu, Q. Wu, Y. Zhang, and M. Jin, "A near-optimal iterative linear precoding with low complexity for massive MIMO systems," *IEEE Commun. Lett.*, vol. 23, no. 6, pp. 1105–1108, Jun. 2019.
[34] Q. Zhang, S. Jin, K.-K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.
[35] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1987.
[36] B. C. Jung and D. K. Sung, "Performance analysis of orthogonal-code hopping multiplexing systems with repetition, convolutional, and turbo coding schemes," *IEEE Trans. Veh. Technol.*, vol. 57, no. 2, pp. 932–944, Mar. 2008.

Liyuan Wen received the B.S. degree from Shanghai Maritime University, Shanghai, China, in 2019. She is currently pursuing the joint Ph.D. degree with the School of Information Science and Technology, ShanghaiTech University, Shanghai, the Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, and the University of Chinese Academy of Sciences, Beijing, China. Her current research interests include massive MIMO systems and hybrid beamforming.

Hua Qian (Senior Member, IEEE) received the B.S. and M.S. degrees from Tsinghua University, Beijing, China, in 1998 and 2000, respectively, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta, GA, USA, in 2005. He has more than 20 years of research and development experience in signal processing, wireless communications, and ASIC design. He is currently a Professor with the Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, China. He is also an Adjunct Professor with ShanghaiTech University, Shanghai. He has coauthored two book chapters, published more than 130 SCI/EI indexed papers, and applied for 60 patents, including five granted U.S. patents. His current research interests include nonlinear signal processing, distributed signal processing, and system design of wireless communications.

Yunbo Hu received the B.S. degree from Tongji University, Shanghai, China, in 2018. He is currently pursuing the joint Ph.D. degree with the Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, and the University of Chinese Academy of Sciences, Beijing, China. His current research interests include signal processing for hybrid beamforming, intelligent reflecting surface, and massive MIMO systems.

Zhicheng Deng received the B.S. degree in automation from Northeastern University, Shenyang, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Information Science and Technology, ShanghaiTech University, Shanghai, China. His research interests include distributed optimization, asynchronous optimization, and tight complexity analysis for optimization algorithm.



Xiliang Luo (Senior Member, IEEE) received the B.S. degree in physics from Peking University, Beijing, China, in 2001, and the M.S. and Ph.D. degrees in electrical engineering from the University of Minnesota, Minneapolis, MN, USA, in 2003 and 2006, respectively. After finishing his Ph.D. studies, he joined Qualcomm Research and was involved in the system designs, analyses, and standardization of 4G LTE. He was a Designer of various enhancements to Qualcomm's LTE Solutions and led the designs of heterogeneous networks from initial concepts to successful modem development. Since 2014, he has been with the School of Information Science and Technology, ShanghaiTech University, Shanghai, China. He has authored or coauthored more than 100 research papers. He is the co-inventor of more than 70 U.S. and international patents and majority of those have been adopted into current 4G and 5G wireless communication standards. His research interests include signal processing, communications, and machine learning. Particularly, he is interested in researches combining information theory and machine learning theory that can shape and guide the designs of next generation data and information processing networks.