entropy

MDPI

Article

Efficient Constant Envelope Precoding for Massive MU-MIMO Downlink via Majorization-Minimization Method

Rui Liang, Hui Li *, Yingli Dong and Guodong Xue

School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China; liangrui@mail.nwpu.edu.cn (R.L.); d.y.l@mail.nwpu.edu.cn (Y.D.); xgd520@mail.nwpu.edu.cn (G.X.)
* Correspondence: lh@nwpu.edu.cn

Abstract: The practical implementation of massive multi-user multi-input–multi-output (MU-MIMO) downlink communication systems power amplifiers that are energy efficient; otherwise, the power consumption of the base station (BS) will be prohibitive. Constant envelope (CE) precoding is gaining increasing interest for its capability to utilize low-cost, high-efficiency nonlinear radio frequency amplifiers. Our work focuses on the topic of CE precoding in massive MU-MIMO systems and presents an efficient CE precoding algorithm. This algorithm uses an alternating minimization (AltMin) framework to optimize the CE precoded signal and precoding factor, aiming to minimize the difference between the received signal and the transmit symbol. For the optimization of the CE precoded signal, we provide a powerful approach that integrates the majorization-minimization (MM) method and the fast iterative shrinkage-thresholding (FISTA) method. This algorithm combines the characteristics of the massive MU-MIMO channel with the second-order Taylor expansion to construct the surrogate function in the MM method, in which minimizing this surrogate function is the worst-case of the system. Specifically, we expand the suggested CE precoding algorithm to involve the discrete constant envelope (DCE) precoding case. In addition, we thoroughly examine the exact property, convergence, and computational complexity of the proposed algorithm. Simulation results demonstrate that the proposed CE precoding algorithm can achievean uncoded biterror rate (BER) performance gain of roughly 1 dB compared to the existing CE precoding algorithm and has an acceptable computational complexity. This performance advantage also exists when it comes to DCE precoding.

Keywords: alternating minimization; constant envelope precoding; majorization-minimization method; fast iterative shrinkage-thresholding method; second-order Taylor expansion

check for updates

Citation: Liang, R.; Li, H.; Dong, Y.; Xue, G. Efficient Constant Envelope Precoding for Massive MU-MIMO Downlink via Majorization-Minimization Method. *Entropy* **2024**, *26*, 349. https://doi.org/10.3390/e26040349

Academic Editors: Boris Ryabko and Luca Barletta

Received: 4 January 2024
Revised: 30 March 2024
Accepted: 19 April 2024
Published: 21 April 2024

1. Introduction

Massive MIMO technology is widely regarded as a revolutionary advancement in the fifth generation mobile communication system [1]. Compared with small-scale MIMO systems, large-scale antenna arrays at transmitters have been demonstrated to offer substantial benefits in terms of spectrum efficiency, energy efficiency, and reliable transmission [2–4]. In contrast to conventional small-scale MIMO systems that employ highly linear and power-inefficient radio frequency (RF) amplifiers, the practical implementation of massive MIMO systems requires the RF amplifiers to be power-efficient. Otherwise, the power consumption of the base station (BS) would be prohibitive. Therefore, it is crucial to use RF amplifiers with a certain power efficiency to avoid excessive power consumption at the BS [5]. Regrettably, energy-efficient RF amplifiers frequently exhibit inferior linearity characteristics, necessitating a lower peak-to-average power ratio (PAPR) for the input signals.

In response to the demand for energy-efficient and cost-effective RF components in wireless communication systems, constant envelope (CE) transmission that forces each antenna unit to transmit a constant envelope signal and allows the use of the most energy-efficient and cheapest power amplifiers (PAs) has attracted great attention from researchers.

CE precoding was initially introduced by Mohammed and Larsson in [6]. Specifically, in the single-antenna CE constraint, the equivalent complex baseband signal of each transmit antenna is restricted to have a CE [7]. By combining with instantaneous CSI, CE precoding only transmits the phase of the desired information symbol to multiple antennas. CE precoding reduces the PAPR compared to non-CE precoding by providing a quasi-CE continuous-time RF signal to each PA. On the other hand, when highly efficient PAs with nonlinear amplitude transfer functions are used, the output distortion can be greatly reduced [8]. Furthermore, the low PAPR properties of the CE signals enable the utilization of cost-effective RF components that have a decreased dynamic range.

Based on all the advantages of CE precoding mentioned above, CE precoding has become a potential research direction. However, it is important to mention that the CE characteristics of the transmit signal are typically non-convex, which brings challenges to precoding design. By adopting multi-user interference (MUI) as the design goal, the CE precoding problem can be modeled as a non-convex nonlinear least squares (NLS) problem, and its local minimum can be obtained using the gradient descent method [9]. Nevertheless, the CE constraint can lead to the error-floor issue in situations with a high signal-to-noise-ratio (SNR). On this basis, the cross-entropy optimization (CEO) method was proposed in [10] to improve the CE precoding design optimal search. Furthermore, an alternating minimization projected gradient (GP-AltMin) method was proposed in [11]. This method improves the performance at high SNR by ignoring the noise of the system. Zhang et al. [12] studied a single-user multiple-input single-output (MISO) system with CE constraint for each antenna and proposed constellation designs for fixed-rate and variable-rate adaptive receivers. In [8], the joint design of transmit antenna grouping and receive beamforming vectors was conducted with the objective of minimizing the maximum symbol error rate (SER) in the data stream. Nevertheless, this typically necessitates a significant computational expenditure. Shao et al. [13] introduced a novel first-order algorithm that utilizes a projected gradient (PG) method in order to effectively minimize the SER of the system. To further accelerate the convergence rate of the PG method, a low-complexity fast gradient projection (FPG) algorithm similar to the fast iterative shrinkage-thresholding (FISTA) algorithm was considered. However, this approach necessitates a greater quantity of iterations in order to attain convergence. In [14], a novel algorithm was provided that integrates gradient extrapolation with the majorization-minimization technique (GEMM). This algorithm is not only suitable for CE precoding, but can also be applied to one-bit DACs precoding. Nevertheless, this approach must be used to approximate the SER expression of the system when formulating the optimization objective. Wang et al. [15] rotated and scaled each constellation point individually to take advantage of the additional degrees of freedom by jointly optimizing the transmit signal matrix and complex scaling factors to maximize the signal-to-interference-to-noise ratio (SINR) at the receiver.

Recently, researchers have found that including the concept of constructive interference (CI) in the design of precoding can significantly enhance system performance. Inspired by this, the authors in [16] investigated the utilization of CI to take advantage of the MUI in the system, with the aim of reducing the SER. In [17], a CEO-based method was proposed for PSK modulations, which achieved significant performance improvements over the classic CE precoding method based on interference minimization in [9]. Furthermore, Liu et al. [18] proposed an effective Riemannian conjugate gradient (RCG) method to address the CE precoding problem that takes into account CI and achieves a balance between performance and complexity. However, in fact, this CI-based CE precoding design is only suitable for PSK modulations. The CE characteristic of the signal assumes that the phase of the phase shifter is continuous or has nearly continuous phase resolution. Often, this is unsatisfactory. Therefore, it is necessary to study the discrete constant envelope (DCE) precoding methods in which the phase shifters have finite phase resolution. The investigation in [19] focused on the transmit signals in the DCE case. In [19], a PG-based symbol-level mean square error (MSE) precoding algorithm was proposed, which is not only applicable to the strict CE constraint, but also to the polygon constraint of DCE. In [20], a greedy precoding design

using the MSE of system as the design criterion was proposed. The DCE precoder was solved in [21] when using a single common PA and separate digital phase shifters. This method combines DCE with CI to improve performance. Moreover, the authors in [11,14] have expanded their investigations to accommodate the DCE transmit signals.

As previously stated, CE precoding methods typically exhibit favorable performance. Motivated by this, our study focuses on the development of a CE precoding design that aims to minimize the difference between the received signal and transmit symbol in massive MU-MIMO systems. In comparison with [14], which focuses on designing the minimum SER of the system under the worst-case for a single user, our method aims to optimize the overall performance of the system. We address this by formulating a CE precoding problem using the minimum mean square error (MMSE) criterion. It is important to note that in [14], it is required to make an approximation of the objective problem, and we intentionally avoid doing so. The main contributions are as follows:

1. One of the main challenges in solving the CE precoding problem is the interdependence between the CE precoded signal and the precoding factor. To address this problem, we employ a two-stage iterative procedure involving an alternating minimization (AltMin) framework. When addressing the CE precoded signal, the CE constraint is simplified and transformed into unit modulus constraints by introducing an auxiliary variable. Additionally, the unit modulus constraint is converted to continuous by adding a penalty term to the objective function.
2. The optimal precoded signal is obtained using the majorization-minimization (MM) framework. In the MM framework, the key is how to construct the surrogate function. We exploit the channel characteristics of massive MU-MIMO systems and combine them with a second-order Taylor expansion to obtain an efficient surrogate function. Unlike the one-step GEMM method described in [14], we obtain the precise values of the auxiliary variables through multiple iterations. In addition, we derive the L-Lipschiz constant and analyze the exact property, convergence, and computational complexity of the proposed algorithm.
3. The proposed method is extended to DCE precoding schemes that have finite phase resolution. At first, we manipulate the continuous phase of the CE signal to align with the PSK constellation by performing a straightforward rotation. Then, we employ algebraic knowledge to derive the DCE precoded signal by making secondary decisions.
4. Simulation results demonstrate that in the CE precoding case, the proposed algorithm exhibits superior uncoded BER performance and a lower computational complexity when compared to existing approaches. In both PSK modulation and QAM modulation, the suggested CE precoding method can achieve a performance gain of about 1 dB. In the 3-phase case, the proposed algorithm also has better performance.

The remainder of this paper is organized as follows: In Section 2, we present the model of the CE precoding system for the massive MU-MIMO system and the CE precoding problem based on the MMSE criterion. In Section 3, we give the detailed process of the algorithm for solving the CE precoding problem. Furthermore, the proposed algorithm is extended to the DCE case. The performance of the proposed algorithm is illustrated by analyzing its exact property, convergence, and computational complexity in Section 4. Section 5 presents the simulation, numerical results, and analysis. Section 6 is a summary of this paper.

Notations: In this paper, a, $\mathbf{a}$, and $\mathbf{A}$ are the scalars, vectors, and matrices. For matrices and vectors, $\|\cdot\|_2$ stands for the spectral norm and Euclidean norm, respectively. Operator $|\cdot|$ represents the absolute value of a scalar or the cardinality of a set. The transpose and its conjugate transpose of a vector or matrix are denoted by $(\cdot)^T$ and $(\cdot)^H$. $\langle \mathbf{a}, \mathbf{b} \rangle$ is the Euclidean inner product. $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of the vector or matrix. The $\lceil \cdot \rceil$ means rounding up to an integer. The set of complex numbers is denoted by $\mathbb{C}$. The matrix $\mathbf{I}$ is an identity matrix with the appropriate dimensions.

2. System Model and Problem Formulation

2.1. System Model

As shown in Figure 1, we consider a single-cell, single-carrier massive MU-MIMO downlink transmission system in TDD mode, in which the BS is equipped with a large-scale antenna array N_{TX} and communicates with N_u single-antenna users at the same time, that is $N_{\mathrm{TX}} \gg N_u$. At the BS, each RF chain is connected to a phase shifter, and it is assumed that the phase shifter can produce continuous phase values throughout the entire phase range. Therefore, the input constellation symbol vector $\mathbf{s} = [s_1, \ldots, s_{N_u}]^T \in \mathcal{O}$ passes through the precoder and phase shifters to form a transmit signal $\mathbf{x} = \left[x_1, \ldots, x_{N_{\mathrm{TX}}}\right]^T \in \mathbb{C}^{N_{\mathrm{TX}}}$ with a constant envelope, where $\mathcal{O}$ is the set of constellation points. The CE precoded signal transmitted by the t-th antenna at the BS is expressed as

$$x_t = \sqrt{\frac{P_T}{N_{\mathrm{TX}}}} e^{j\theta_t}, \forall t \in \{1, 2, \ldots, N_{\mathrm{TX}}\}, \tag{1}$$

where $\theta_t \in [0, 2\pi]$ is the phase of the CE precoded signal x_t. P_T is the instantaneous transmit power, and this shows that the CE precoded signal satisfies $\|\mathbf{x}\|_2^2 \leq P_T$. The CE property forces the transmit signal $\mathbf{x}$ to satisfy the CE constraint, that is

$$\mathcal{V}_1 = \left\{ \mathbf{x} \in \mathbb{C}^{N_{\mathrm{TX}}} \middle| |x_t| = \sqrt{\frac{P_T}{N_t}}, t = 1, \ldots, N_{\mathrm{TX}} \right\}. \tag{2}$$

Assuming a transmission time duration not exceeding the channel coherence time. For such a system, the discrete-time complex baseband signal received at users during the downlink of an arbitrary coherence interval can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{3}$$

where $\mathbf{y} = [y_1, \ldots, y_{N_u}]^T \in \mathbb{C}^{N_u}$ denotes the received signal vector of all users. The matrix $\mathbf{H} \in \mathbb{C}^{N_u \times N_{\mathrm{TX}}}$ denotes the downlink channel, which is perfectly known at the BS. The vector $\mathbf{n} \in \mathbb{C}^{N_u}$ is an additive noise and $\mathbf{n} \sim \mathcal{CN}\left(0, \sigma^2 \mathbf{I}\right)$.

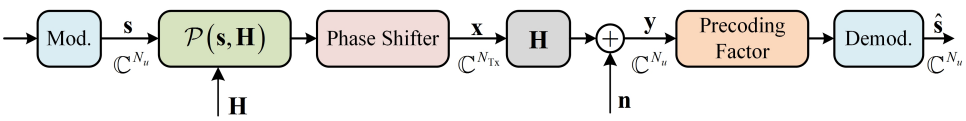


Figure 1. The CE precoder for the massive MU-MIMO downlink system.

2.2. Problem Formulation

In the CE precoding design, the task is to design the transmit signal $\mathbf{x}$ under the CE constraint so that the MSE between the transmit symbol $\mathbf{s}$ and its estimated value $\hat{\mathbf{s}}$ is minimized. In order to facilitate the use of channel gain, an additional precoding factor $\psi \in \mathbb{R}$ [22] is introduced in the CE precoding design. The users can use the precoding factor ψ to obtain an estimate of the transmit symbol $\mathbf{s}$ from the received signal $\mathbf{y}$, i.e., $\hat{\mathbf{s}} = \psi \mathbf{y}$. At the receiving end, the users can estimate the precoding factor for the block-fading channel using either pilot-based estimation or blind estimation [23], where a direct way to obtain the estimated precoding factor for the user is to use the pilots known at the user side. The MSE between the transmit symbol and its estimated value $\hat{\mathbf{s}}$ can be obtained by

$$\mathbb{E}_{\mathrm{s}}\left[\|\mathbf{s} - \hat{\mathbf{s}}\|_2^2\right] = \|\mathbf{s} - \psi \mathbf{H}\mathbf{x}\|_2^2 + \psi^2 N_u \sigma^2, \tag{4}$$

where we restrict the precoder results in the same precoding factor ψ for all the users [24]. With this assumption, the MSE after precoding will be roughly the same for all users,

which guarantees a certain degree of fairness among the users [22]. Thus, the CE precoding problem based on the MMSE criterion is stated as,

$$\begin{aligned} &\min_{\mathbf{x},\psi} \|\mathbf{s} - \psi\mathbf{H}\mathbf{x}\|_2^2 + \psi^2 N_u \sigma^2 \\ &\text{s.t. } \mathbf{x} \in \mathcal{V}_1,\ \psi \in \mathbb{R}. \end{aligned} \tag{5}$$

Since the CE constraint of the CE precoded signal is non-convex, the optimization problem (5) is non-convex. Generally, the optimization problem (5) is NP-hard. In addition, the mutually coupled CE precoded signal and precoding factor also bring difficulties for directly optimizing the CE precoding problem (5). Next, the CE precoded signal and precoding factor will be solved separately in an AltMin method, and an effective algorithm based on the MM framework will be proposed to obtain the CE precoded signal.

3. Majorization-Minimization Method for Constant Envelope Precoding

In order to effectively solve the coupling between the CE precoded signal $\mathbf{x}$ and precoding factors ψ, the AltMin framework is used to decouple the optimization problem (5) into two subproblems, in which ψ or $\mathbf{x}$ are solved alternately while maintaining correspondingly another variable $\mathbf{x}$ or ψ which is fixed. Specifically, each iterative process of applying the AltMin framework to solving the optimization problem (5) is expressed as the following two steps:

$$\psi^{i+1} = \arg\min_{\psi\in\mathbb{R}} \left\|\mathbf{s} - \psi\mathbf{H}\mathbf{x}^i\right\|_2^2 + \psi^2 N_u \sigma^2, \tag{6a}$$

$$\mathbf{x}^{i+1} = \arg\min_{\mathbf{x}\in\mathcal{V}_1} \left\|\mathbf{s} - \psi^{i+1}\mathbf{H}\mathbf{x}\right\|_2^2 + \left(\psi^{i+1}\right)^2 N_u \sigma^2, \tag{6b}$$

where the algorithm alternately solves the precoding factor and the CE precoded signal until the stopping condition $\left|\mathrm{MSE}^{i+1} - \mathrm{MSE}^i\right| / \mathrm{MSE}^{i+1} \le \epsilon_O$ is met.

In the first step of AltMin, that is to solve subproblem (6a), given the CE precoded signal $\mathbf{x}$, expand and rewrite subproblem (6a) as a quadratic function about the precoding factor ψ

$$\begin{aligned} &\|\mathbf{s} - \psi\mathbf{H}\mathbf{x}\|_2^2 + \psi^2 N_u \sigma^2 \\ &= \mathbf{s}^H\mathbf{s} - 2\psi\mathbf{s}^H\mathbf{H}\mathbf{x} + \psi^2\|\mathbf{H}\mathbf{x}\|_2^2 + \psi^2 N_u \sigma^2. \end{aligned} \tag{7}$$

Setting the partial derivative of (7) with respect to the precoding factor ψ to 0, we can obtain

$$\psi = \frac{\Re\{\mathbf{s}^H\mathbf{H}\mathbf{x}\}}{\|\mathbf{H}\mathbf{x}\|_2^2 + N_u\sigma^2}. \tag{8}$$

Next, the CE precoded signal is optimized by solving subproblem (6b), taking into account the precoding factor ψ. When the precoding factor ψ is provided as a constant, $\psi^2 N_u \sigma^2$ will be eliminated from Equation (6b). To simplify further processing, define the auxiliary variable $\mathbf{u}$ and make $\mathbf{x} = \sqrt{P_T/N_{\mathrm{TX}}}\,\mathbf{u}$. Let $\tilde{\mathbf{H}} = \sqrt{P_T/N_{\mathrm{TX}}}\,\psi\mathbf{H}$. Hence, the subproblem (6b) can be reformulated as

$$\begin{aligned} &\min_{\mathbf{u}} \left\|\mathbf{s} - \tilde{\mathbf{H}}\mathbf{u}\right\|_2^2 \\ &\text{s.t. } \mathbf{u} \in \mathcal{V}_2, \end{aligned} \tag{9}$$

where $\mathcal{V}_2 = \left\{\mathbf{u} \in \mathbb{C}^{N_{\mathrm{TX}}} \middle| |u_t| = 1, t = 1, \ldots, N_{\mathrm{TX}}\right\}$. This means that any u_t in $\mathcal{V}_2$ is on the unit circle.

It can be seen from (9) that although the objective function to be optimized is quadratic, the constraint $\mathcal{V}_2$ is a non-convex unit modulus constraint, so the problem is still non-convex. The generally employed approach for solving problems with a quadratic objective function is the semi-definite relaxation (SDR) method [25–27]. SDR has the advantage of employing

a non-convex optimization problem to approximate the objective function. Although SDR is capable of calculating approximate solutions to non-convex optimization problems in polynomial time, the worst-case computational complexity is proportional to $N_{\mathrm{Tx}}^{4.5}$ [27], hindering its application to large-scale applications. To improve the effectiveness of solving the optimization problem (9), we can explore the implementation of the penalty method in the unit modulus optimization problem. This involves relaxing the unit modulus constraint set for solving the CE precoded signal $\mathbf{x}$ and incorporating a penalty function into the objective function to ensure that the solution lies on the unit circle [28].

Let $f(\mathbf{u}) = \left\|\mathbf{s} - \tilde{\mathbf{H}}\mathbf{u}\right\|_2^2$, the optimization problem can be written as

$$\min_{\mathbf{u}\in\mathcal{V}_3} \mathcal{F}_\rho(\mathbf{u}) = f(\mathbf{u}) - \rho\|\mathbf{u}\|_2^2, \tag{10}$$

where $\mathcal{V}_3 = \left\{\mathbf{u} \in \mathbb{C}^{N_{\mathrm{TX}}} \middle| |u_t| \leq 1, t = 1, \ldots, N_{\mathrm{TX}}\right\}$. The $\rho > 0$ is a penalty parameter. Since CE precoded signals are complex, the penalty term $-\rho\|\mathbf{u}\|_2^2$ is used to push each u_t to any position on the unit circle. It is important to note that while the constraint of the optimization problem (10) exhibits convexity, the objective function is non-convex. Hence, the optimization problem (10) may be classified as a convex constrained minimization problem with a non-convex objective function. This problem can be effectively tackled by employing first-order optimization techniques like the PG method. In general, first-order optimization approaches exhibit a modest level of iterative complexity; however, the iterations needed to attain convergence can be significant. In the theory of convex optimization, it is usual to employ Nesterov- or FISTA-type acceleration algorithms [29] as a way to minimize the number of iterations required. In addition, in large-scale MU-MIMO systems, the optimization problem (10) is usually a large-scale problem, which poses a challenge to using classic algorithms to solve this optimization problem, and the MM architecture can solve this problem [30]. Next, a novel algorithm derived from the MM framework and FISTA method is introduced to efficiently address the optimization problem (10).

3.1. Surrogate Function Using Second-Order Taylor Expansion

Before providing the proposed method, an overview of the fundamental concepts underlying the MM framework for addressing minimization problems is provided. The MM framework is used to iteratively solve a series of simpler problems to replace non-convex optimization problems that are difficult to solve directly [31–33]. For example, consider minimizing the function $J(\mathbf{w})$ within the feasible set $\mathbf{w} \in \mathcal{W}$. Minimizing the function $J(\mathbf{w})$ becomes challenging when the objective function or constraint are non-convex. Thus, rather than directly minimizing the function $J(\mathbf{w})$, the surrogate function $\mathcal{J}\left(\mathbf{w}\middle|\mathbf{w}^k\right)$ of the original objective function is minimized during the k-th iteration. A valid surrogate function has the following properties:

$$\mathcal{J}\left(\mathbf{w}\middle|\mathbf{w}^k\right) > J(\mathbf{w}), \forall \mathbf{w} \in \mathcal{W}, \tag{11a}$$

$$\mathcal{J}\left(\mathbf{w}^k\middle|\mathbf{w}^k\right) = J\left(\mathbf{w}^k\right), \tag{11b}$$

$$\nabla\mathcal{J}\left(\mathbf{w}^k\middle|\mathbf{w}^k\right) = \nabla J\left(\mathbf{w}^k\right). \tag{11c}$$

The above properties indicate that the surrogate function is a tight upper bound of the original objective function. Therefore, the algorithm based on the MM framework starts from the feasible initial point $\mathbf{w}^0 \in \mathcal{W}$ and iteratively minimizes the surrogate function

$$\mathbf{w}^{k+1} = \arg\min_{\mathbf{w}\in\mathcal{W}} \mathcal{J}\left(\mathbf{w}\middle|\mathbf{w}^k\right). \tag{12}$$

To effectively utilize the MM framework for solving problem (10), the crucial aspect is to construct the surrogate function for the objective function. The following Lemma 1 serves as the foundation for the efficient development of surrogate functions [34]. The detailed proofs are described in [35].

Lemma 1. *Consider a quadratic function of the form* $\mathbf{a}^H\mathbf{S}\mathbf{a}$*, where* $\mathbf{S}$ *is a positive semi-definite matrix, then the surrogate function of the* $\mathbf{a}^H\mathbf{S}\mathbf{a}$ *function at point* $\mathbf{a}^k$ *is* $\mathbf{a}^H\mathbf{T}\mathbf{a} + 2\Re\left(\left\langle \mathbf{a}, (\mathbf{S}-\mathbf{T})\mathbf{a}^k \right\rangle\right) + \left\langle \mathbf{a}^k, (\mathbf{T}-\mathbf{S})\mathbf{a}^k \right\rangle$*, where* $\mathbf{T}$ *is a positive semi-definite matrix and* $\mathbf{T} \geq \mathbf{S}$*.*

For any quadratic differentiable function with bounded curvature, Lemma 1 is also known as the quadratic upper bound principle [32]. We first construct the surrogate function of $\mathcal{F}_\rho(\mathbf{u})$. By expanding $\mathcal{F}_\rho(\mathbf{u})$, we can obtain

$$\mathcal{F}_\rho(\mathbf{u}) = \mathbf{s}^H\mathbf{s} - \mathbf{s}^H\tilde{\mathbf{H}}\mathbf{u} - \mathbf{u}^H\tilde{\mathbf{H}}^H\mathbf{s} + \mathbf{u}^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{u} - \rho\|\mathbf{u}\|_2^2. \tag{13}$$

According to Lemma 1, we consider alternatives $\mathbf{u}^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{u}$ in (13). We define $\mathbf{S} = \tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$, and consider the second-order Taylor expansion of $\mathbf{u}^H\mathbf{S}\mathbf{u}$ around $\mathbf{u}^k$ as

$$\mathbf{u}^H\mathbf{S}\mathbf{u} = \left\langle \mathbf{u}^k, \mathbf{S}\mathbf{u}^k \right\rangle + \left\langle \mathbf{u}^k, \mathbf{S}\left(\mathbf{u}-\mathbf{u}^k\right) \right\rangle + \left(\mathbf{u}-\mathbf{u}^k\right)^H\mathbf{S}\mathbf{u}^k + \left(\mathbf{u}-\mathbf{u}^k\right)^H\mathbf{S}\left(\mathbf{u}-\mathbf{u}^k\right). \tag{14}$$

For the massive MIMO system, the channel matrix $\mathbf{H}$ is a fat matrix, and $\mathbf{S}$ is a positive semi-definite Hermitian matrix. Based on the Lemma 1, we replace the matrix $\mathbf{S}$ with a matrix $\mathbf{T}$, where $\mathbf{T} \geq \mathbf{S}$. This implies that (14) is rewritten as

$$\begin{aligned} \mathbf{u}^H\mathbf{S}\mathbf{u} &\leq \left\langle \mathbf{u}^k, \mathbf{S}\mathbf{u}^k \right\rangle + \left\langle \mathbf{u}^k, \mathbf{S}\left(\mathbf{u}-\mathbf{u}^k\right) \right\rangle + \left(\mathbf{u}-\mathbf{u}^k\right)^H\mathbf{S}\mathbf{u}^k + \left(\mathbf{u}-\mathbf{u}^k\right)^H\mathbf{S}\left(\mathbf{u}-\mathbf{u}^k\right) \\ &= \mathbf{u}^H\mathbf{T}\mathbf{u} + \mathbf{u}^H(\mathbf{S}-\mathbf{T})\mathbf{u}^k + \left\langle \mathbf{u}^k, (\mathbf{S}-\mathbf{T})\mathbf{u} \right\rangle + \left\langle \mathbf{u}^k, (\mathbf{T}-\mathbf{S})\mathbf{u}^k \right\rangle. \end{aligned} \tag{15}$$

Since $\mathbf{S}$ is a positive semi-definite Hermitian matrix, we perform eigenvalue decomposition on $\mathbf{S}$ and extract the maximum eigenvalue $\lambda_{\max}$, that is $\lambda_{\max} = \mathrm{eig}(\mathbf{S})$. We choose $\mathbf{T} = \lambda_{\max}\mathbf{I}$, then $\mathbf{T} \geq \mathbf{S}$ can be satisfied. Putting $\mathbf{T} = \lambda_{\max}\mathbf{I}$ into (15), we can obtain

$$\mathbf{u}^H\mathbf{S}\mathbf{u} \leq \lambda_{\max}\mathbf{u}^H\mathbf{u} + \mathbf{u}^H(\mathbf{S}-\lambda_{\max}\mathbf{I})\mathbf{u}^k + \left\langle \mathbf{u}^k, (\mathbf{S}-\lambda_{\max}\mathbf{I})\mathbf{u} \right\rangle + \left\langle \mathbf{u}^k, (\lambda_{\max}\mathbf{I}-\mathbf{S})\mathbf{u}^k \right\rangle. \tag{16}$$

We define

$$g\left(\mathbf{u}\middle|\mathbf{u}^k\right) = \lambda_{\max}\mathbf{u}^H\mathbf{u} + \mathbf{u}^H(\mathbf{S}-\lambda_{\max}\mathbf{I})\mathbf{u}^k + \left\langle \mathbf{u}^k, (\mathbf{S}-\lambda_{\max}\mathbf{I})\mathbf{u} \right\rangle + \left\langle \mathbf{u}^k, (\lambda_{\max}\mathbf{I}-\mathbf{S})\mathbf{u}^k \right\rangle. \tag{17}$$

Substituting $g\left(\mathbf{u}\middle|\mathbf{u}^k\right)$ into (13), we obtain the surrogate function $\mathcal{G}\left(\mathbf{u}|\mathbf{u}^k\right)$ of $\mathcal{F}_\rho(\mathbf{u})$

$$\mathcal{G}_\rho\left(\mathbf{u}|\mathbf{u}^k\right) = \mathbf{s}^H\mathbf{s} - \mathbf{s}^H\tilde{\mathbf{H}}\mathbf{u} - \mathbf{u}^H\tilde{\mathbf{H}}^H\mathbf{s} + g\left(\mathbf{u}|\mathbf{u}^k\right) - \rho\|\mathbf{u}\|_2^2. \tag{18}$$

It should be noted that, unlike the classical MM framework that approximates the non-convex part of the objective function, we use $\mathbf{T} = \lambda_{\max}\mathbf{I}$ to process $\mathbf{u}^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{u}$ in (13) according to Lemma 1, and the resulting surrogate function $\mathcal{G}_\rho\left(\mathbf{u}|\mathbf{u}^k\right)$ is an upper bound on $\mathcal{F}_\rho(\mathbf{u})$, which is a worst-case. The surrogate function $\mathcal{G}_\rho\left(\mathbf{u}|\mathbf{u}^k\right)$ satisfies the properties (11). Next, we will use the above surrogate function and combine MM framework with the FISTA algorithm to iteratively solve the optimization problem (10).

3.2. MM Method For Solving CE Precoding

We take the general form of the MM method as shown below to find the minimum of the surrogate function $\mathcal{G}_\rho\left(\mathbf{u}|\mathbf{u}^k\right)$

$$\mathbf{u}^{k+1} = \arg\min_{\mathbf{u}\in\mathcal{V}_3} \mathcal{G}_\rho\left(\mathbf{u}\middle|\mathbf{u}^k\right), k = 0, 1, 2, \ldots \tag{19}$$

In the process of solving (19), we use the FISTA method [36] to solve. The FISTA method for solving $\min_{\mathbf{u}\in\mathcal{V}_3} \mathcal{G}_\rho\left(\mathbf{u}\middle|\mathbf{u}^k\right)$ is

$$\mathbf{u}^{k+1} = \prod_{\mathcal{V}_3}\left(\mathbf{z}^k - \mu^{-1}\nabla\mathcal{G}_\rho\left(\mathbf{z}^k|\mathbf{u}^k\right)\right), k = 0, 1, 2, \ldots, \tag{20}$$

where μ is the step size. Notably, $\mathbf{z}^k - \mu^{-1}\nabla\mathcal{G}_\rho\left(\mathbf{z}^k\middle|\mathbf{u}^k\right)$ does not always satisfy the CE constraint. Consequently, we project $\mathbf{z}^k - \mu^{-1}\nabla\mathcal{G}_\rho\left(\mathbf{z}^k\middle|\mathbf{u}^k\right)$ to the CE constraint set $\mathcal{V}_3$, i.e.,

$$\prod_{\mathcal{V}_3}\left(\mathbf{z}^k - \mu^{-1}\nabla\mathcal{G}_\rho\left(\mathbf{z}^k\middle|\mathbf{u}^k\right)\right) = e^{j\measuredangle\left(\mathbf{z}^k - \mu^{-1}\nabla\mathcal{G}_\rho\left(\mathbf{z}^k|\mathbf{u}^k\right)\right)}, \tag{21}$$

where j is the imaginary unit. The $\measuredangle$ is the corresponding phase. The gradient vector of $\mathcal{G}_\rho\left(\mathbf{z}^k\middle|\mathbf{u}^k\right)$ is

$$\nabla\mathcal{G}_\rho\left(\mathbf{z}^k|\mathbf{u}^k\right) = 2\left((\lambda_{\max}-\rho)\mathbf{z}^k + (\mathbf{S}-\lambda_{\max}\mathbf{I})\mathbf{u}^k - \tilde{\mathbf{H}}^H\mathbf{s}\right). \tag{22}$$

The $\mathbf{z}^k$ is an extrapolated point and is updated with respect to $\mathbf{u}^{k-1}$ from the previous iteration and $\mathbf{u}^k$ from the current iteration

$$\mathbf{z}^k = \mathbf{u}^k + \alpha_k\left(\mathbf{u}^k - \mathbf{u}^{k-1}\right), \tag{23}$$

with

$$\alpha_k = \frac{\xi_{k-1} - 1}{\xi_k}, \quad \xi_k = \frac{1 + \sqrt{1 + 4\xi_{k-1}^2}}{2}, \tag{24}$$

and with $\xi_{-1} = 0$, $\mathbf{u}^{-1} = \mathbf{u}^0$. The $\{\alpha_k\}_{k\geq 0}$ is the extrapolation sequence. In particular, when $\alpha_k = 0$ in (23), the FISTA method is simplified to the PG method.

For the step size, we choose to use the L-Lipschiz constant as the step size μ. When $f(\mathbf{u})$ is a convex function on $[-1, 1]^{N_{\mathrm{TX}}}$, the Lipschitz continuity condition of $f(\mathbf{u})$ holds according to the following lemma:

Lemma 2. *For a function $f(\mathbf{u})$ that is L-Lipschitz continuous in the domain $[-1, 1]^{N_{TX}}$, its L-Lipschitz constant is*

$$L = 2\sqrt{N_{TX}}\|\tilde{\mathbf{H}}\|_2^2. \tag{25}$$

Proof. See Appendix A for a proof. □

In summary, the proposed MM (SoTMM) algorithm using the second-order Taylor expansion as the surrogate function uses (21) to (23) to iteratively minimize the upper bound of the objective function to solve the non-convex CE precoding problem (9) until the stopping condition $\left\|\mathbf{u}^{k+1} - \mathbf{u}^k\right\|_2^2 \leq \epsilon_I$ is met. After obtaining the optimal value of $\mathbf{u}$, the CE precoded signal can be obtained using the relationship $\mathbf{x} = \sqrt{P_T/N_{\mathrm{TX}}}\,\mathbf{u}$. As a convenience, Algorithm 1 provides a summary of the detailed procedures for resolving the optimization problem (5), which is divided into two iteration loops: the inner iteration loop used to solve the CE precoded signal $\mathbf{x}$, with k as the index; and the outer iteration loop used to optimize the precoding factor ψ, with i as the index.

Algorithm 1 SoTMM method for solving problem (5)

Input: $\mathbf{s}$, $\mathbf{H}$, σ^2.
Initialization: $\mathbf{x}^0 = \mathbf{0}$;
Set: $i = 0$, $\epsilon_O > 0$.
repeat
  Compute the precoding factor ψ by (8);
  Let $\tilde{\mathbf{H}} = \sqrt{P_T/N_{\text{TX}}}\,\psi\mathbf{H}$, and use (25) to calculate the Lipschitz constant;
  Set: $\mathbf{u}^0 = \mathbf{u}^{-1} = \left(1/\sqrt{P_T/N_{\text{TX}}}\right)\mathbf{x}^i$, $\xi_{-1} = 0$, penalty parameter $\rho > L$, step size $\mu = L$, $k = 0$, $\epsilon_I > 0$.
  Define $\mathbf{S} = \tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$ and extract the largest eigenvalue $\lambda_{\max}$ by eigenvalue decomposition of $\mathbf{S}$;
  repeat
    Compute α_k and ξ_k by (24);
    Compute the extrapolated point $\mathbf{z}^k$ by (23);
    Compute the gradient vector $\nabla\mathcal{G}_\rho\left(\mathbf{z}^k\middle|\mathbf{u}^k\right)$ by (22);
    Compute the $\mathbf{z}^k - \mu^{-1}\nabla\mathcal{G}_\rho\left(\mathbf{z}^k|\mathbf{u}^k\right)$ and update $\mathbf{u}^{k+1} = e^{j\angle\left(\mathbf{z}^k - \mu^{-1}\nabla\mathcal{G}_\rho\left(\mathbf{z}^k|\mathbf{u}^k\right)\right)}$;
    $k \leftarrow k + 1$;
  until *A stopping criterion triggers.*
  Reconstruction $\mathbf{x} = \sqrt{P_T/N_{\text{TX}}}\,\mathbf{u}$;
  $i \leftarrow i + 1$;
until *A stopping criterion triggers.*
Output: $\mathbf{x}$, ψ.

3.3. DCE Precoding

In previous studies, we assume that the phase shifters of the CE precoder can produce continuous phase values throughout the entire phase range, or that the phase shifters have approximately continuous phase resolutions. If the above two situations are not the case, we should consider the DCE precoding for phase shifters with a finite phase resolution. In what follows, we will extend the proposed SoTMM algorithm to DCE precoding design. In the DCE precoding case, the CE constraint $\mathcal{V}_1$ of the transmit signal $\mathbf{x}$ will be discretized as

$$\mathcal{X} = \left\{e^{j\frac{2\pi}{2^\kappa}p} \middle| p = 1, 2, \cdots 2^\kappa\right\}, \tag{26}$$

where κ is the discrete resolution of the phase shifter, that is, κ is a positive integer of $\kappa \geq 2$. Therefore, after using Algorithm 1 to obtain the optimal CE precoded signal $\mathbf{x}$, each signal element in the CE precoded signal $\mathbf{x}$ needs to be discretized to the closest DCE constraint set $\mathcal{X}$ points.

Figure 2 shows a diagram of the CE constraint set and the DCE constraint set, where the discrete resolution of the DCE constraint set is $\kappa = 3$. The red parts represent the CE constraint set of CE precoding and the DCE constraint of DCE precoding, respectively, and the shaded parts represent the relaxed constraint set. As can be seen from Figure 2, the CE constraint can be regarded as a continuous point on a circle with a radius of $\sqrt{P_T/N_{\text{TX}}}$, while the DCE constraint is a discrete point on the circle. Therefore, DCE precoding design is converted into a problem of how to design discrete phases on a circle, that is, designing a mapping method to discretize the continuous CE precoded signal into the DCE precoded signal. In general, it is difficult to obtain a strict algebraic expression of this mapping relationship. Existing research shows that in DCE precoding design, the CE precoded signal can be mapped to the 2^κ-PSK constellation to obtain the DCE precoded signal [19].

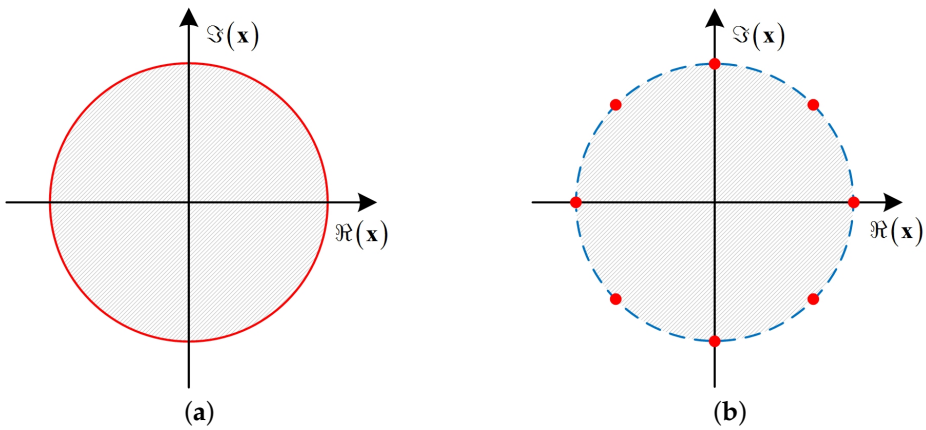


Figure 2. Illustration of constraint set. The red parts are the constraint set and the shaded parts are the relaxed constraint set. (**a**) CE set; (**b**) DCE set.

Using $\kappa = 3$ as an example, Figure 3 illustrates the process of DCE precoding projection. The red points represent the DCE precoded signal, the green circles and points represent the projected DCE precoded signal, and the orange squares represent the 8-PSK constellation points. As can be seen from Figure 3a, the CE precoded signal obtained using Algorithm 1 satisfies the CE constraint, that is, the CE precoded signal all falls on the circle. When $\kappa = 3$, the DCE precoded signal can be projected to the 8-PSK constellation point by simply rotating $\pi/8$ counterclockwise. According to Figure 3b, we will give a detailed projection process. During the projection process, a quadratic decision is used to make the mapping result more accurate. First, project the point located on $\overset{\frown}{AB}$ to the nearest point $\hat{\mathbf{x}}$ on the straight line AB, that is

$$\hat{\mathbf{x}}^{i+1} = \cos\left(\frac{2\pi}{2^{\kappa}}\right) + j\Im\left(\mathbf{x}^{i+1}\right). \tag{27}$$

In order to determine the final projection point, the straight line equation is used to divide the straight line AB. Combining algebraic knowledge, the equation of the straight line ℓ_1 can be obtained as

$$\ell_1 : \Im(\mathbf{x}) = \Re(\mathbf{x}). \tag{28}$$

Therefore, we can obtain the division of the straight line AB

$$\text{prox}\left(\hat{\mathbf{x}}^{i+1}\right) = \begin{cases} e^{j\frac{3\pi}{8}}, \Im\left(\hat{\mathbf{x}}^{i+1}\right) \geq \ell_1, \\ e^{j\frac{\pi}{8}}, \Im\left(\hat{\mathbf{x}}^{i+1}\right) < \ell_1, \end{cases} \tag{29}$$

where prox$(\cdot)$ means discretizing the input signal. Using the above formula, the CE precoded signal can be discretized into an 8-PSK constellation. Combining (27)–(29), the discrete expression of DCE precoding for arbitrary phases is given as

$$\text{prox}\left(\hat{\mathbf{x}}^{i+1}\right) = \begin{cases} e^{j\left(\frac{2\pi}{2^{\kappa}} + \frac{(n-1)\pi}{2^{\kappa}}\right)}, \Im\left(\hat{\mathbf{x}}^{i+1}\right) \geq \ell_{n-1}, \\ e^{j\left(\frac{2\pi}{2^{\kappa}} - \frac{(n-1)\pi}{2^{\kappa}}\right)}, \Im\left(\hat{\mathbf{x}}^{i+1}\right) < \ell_{n-1}, \end{cases} \tag{30}$$

where $n = \left\lceil \frac{\measuredangle \mathbf{x} + \pi/2^{\kappa}}{2\pi/2^{\kappa}} \right\rceil$. The point on the arc projects to the nearest point $\hat{\mathbf{x}}$ on the corresponding straight line as

$$\hat{\mathbf{x}}^{i+1} = \cos\left(\frac{n\pi}{2^{\kappa}}\right) + j\Im\left(\mathbf{x}^{i+1}\right). \tag{31}$$

The straight line ℓ_{n-1} is given by

$$\Im(\mathbf{x}) = \cos\left(\frac{n\pi}{2^{\kappa}}\right)\Re(\mathbf{x}). \tag{32}$$

To summarize, by utilizing (30)–(32), it is possible to obtain the DCE precoded signal with arbitrary phases. Finally, the DCE precoding factor is recalculated according to (8).

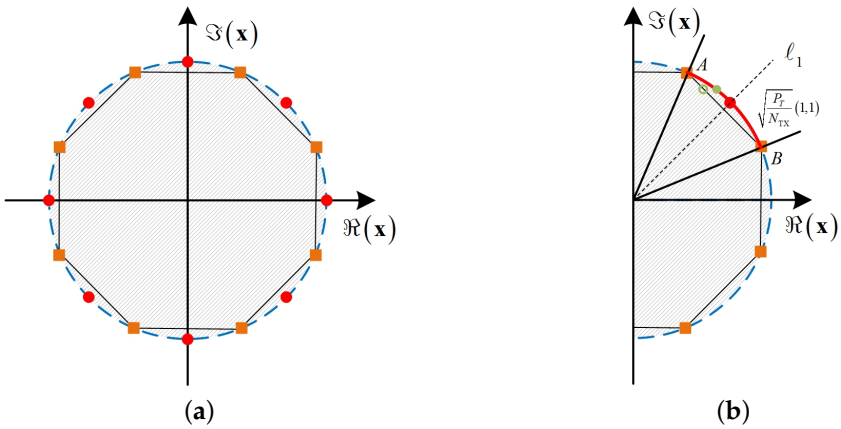


Figure 3. The projection of DCE precoding, $\kappa = 3$. (**a**) DCE set, $\kappa = 3$; (**b**) The projection onto the right half plane.

4. Performance Analysis

In this section, the exact property of the penalty optimization problem (10), the convergence performance, and the computational complexity of the proposed SoTMM algorithm are analyzed in detail.

4.1. The Exact Property of Problem (10)

For the minimization problem (10), it is natural to question whether the penalty optimization problem is an exact restatement of the original optimization problem (9). The following Theorem 1 [14] can illustrate this problem.

Theorem 1. *Assume that the function f is Lipschitz continuous in the feasible set $\mathcal{V}_3$. Then, there is a constant $\bar{\rho} > 0$ such that for any $\rho > \bar{\rho}$, any (global) optimal solution to the optimization problem* (10) *is also the (global) optimal solution to the optimization problem* (9)*. Especially, in the CE precoding case, $\bar{\rho} = L$, where L is the Lipschitz constant of function f in $\mathcal{V}_3$.*

Theorem 1 shows that when the penalty parameter is large enough, the optimal solutions of problems (10) and (9) are equivalent. In particular, this equivalent result does not require additional dynamic adjustment of the penalty parameter ρ, which also provides a theoretical basis for the selection of the penalty parameter ρ.

4.2. Convergence Analysis

In fact, the convergence analysis of non-convex first-order methods involving the accelerated proximal gradient method or the FISTA method is challenging. Here, we are inspired by [14] to prove the convergence performance of the SoTMM algorithm using gradient extrapolation. Theorem 2 describes the convergence performance of the proposed SoTMM algorithm.

Theorem 2. *Suppose there is a Lipschitz constant $L_{\mathcal{F}}$ such that the function $\mathcal{F}$ has a Lipschitz continuous gradient. And for any $\bar{\mathbf{u}} \in \mathcal{V}_3$, there is a Lipschitz constant $L_{\mathcal{G}}$ such that the surrogate*

function $\mathcal{G}_\rho(\cdot|\bar{\mathbf{u}})$ has a Lipschitz continuous gradient. In addition, assuming that α_k obtained from the (21) to (23) satisfies $0 \le \alpha_k \le \bar{\alpha}$, making

$$\begin{aligned} &\min_{m=0,\cdots,k} dist\left(\mathbf{0}, \nabla\mathcal{F}\left(\mathbf{u}^{m+1}\right) + \partial\mathcal{I}_{\mathcal{Y}_3}\left(\mathbf{u}^{m+1}\right)\right) \\ &\le C\sqrt{\frac{8}{(k+1)(1-\bar{\alpha}^2)\mu}(\mathcal{F}(\mathbf{u}^0) - \mathcal{F}^*)} \end{aligned} \tag{33}$$

is true, then the proposed SoTMM is guaranteed to find a stationary point, where

$$C = \max((L_{\mathcal{G}} + \mu)\bar{\alpha}, L_{\mathcal{F}} + \mu). \tag{34}$$

Proof. See Appendix B for a proof. □

4.3. Complexity Analysis

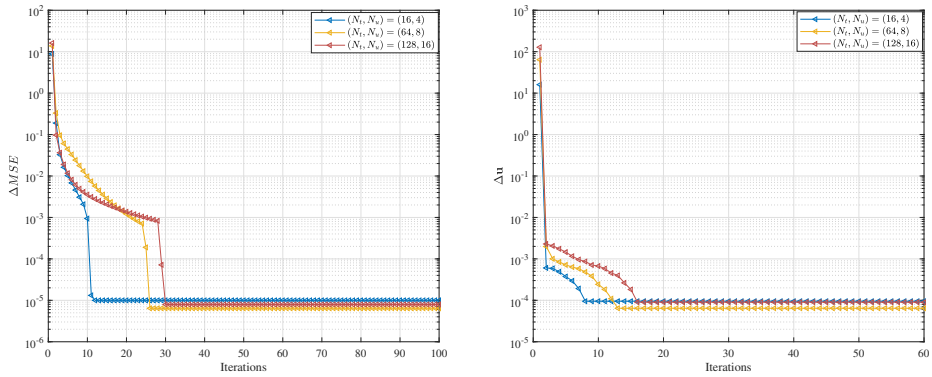
To illustrate the computational complexity of the SoTMM algorithm, we discuss the number of multiplications performed by the algorithm. First of all, it needs to be made clear that the SoTMM algorithm requires two iteration loops, inner and outer. The main complexity of the inner iteration loop comes from computing the gradient vector $\nabla\mathcal{G}_\rho\left(\mathbf{z}^k\middle|\mathbf{u}^k\right)$ of the surrogate function $\mathcal{G}_\rho\left(\mathbf{z}^k\middle|\mathbf{u}^k\right)$. The gradient computation in Lines 11–12 of Algorithm 1 is $N_{\text{TX}}^2 + N_{\text{TX}}N_u + 3N_{\text{TX}}$, where the computational complexity of the projection operation is ignored. In the inner iteration loop, the eigenvalue decomposition of matrix $\mathbf{S}$ and the calculation of the extrapolation point $\mathbf{z}^k$ require N_{TX}^2 and N_{TX} multiplications, respectively. In Algorithm 1, lines 5, 6, and 14 are the relevant steps of the outer iteration loop, requiring a total of $N_{\text{TX}}N_u + 2N_{\text{TX}}$ multiplication operations. In addition, in Algorithm 1, calculating the precoding factor ψ requires $N_{\text{TX}}^2 + N_{\text{TX}}N_u + N_{\text{TX}}$ complex multiplications. Hence, the overall computational complexity required to execute the proposed SoTMM algorithm once to obtain the optimal CE precoded signal and precoding factor is equal to

$$\mathcal{O}\left(K_1\left(2N_{\text{TX}}^2 + 2N_{\text{TX}}N_u + 3N_{\text{TX}} + K_2\left(N_{\text{TX}}^2 + N_{\text{TX}}N_u + 4N_{\text{TX}}\right)\right)\right), \tag{35}$$

where K_1 and K_2 are the maximum number of iterations for the inner and outer iteration loops, respectively.

5. Simulation Results and Discussions

We conduct simulation experiments to validate the performance of the proposed SoTMM algorithm and compare it to several existing CE precoding schemes. Among the involved comparison algorithms are the ZF precoding scheme employing direct projection to the CE constraint (ZF-CE), the GP-AltMin [11], the FPG method [13], and the GEMM algorithm [14] (simulation evaluations are carried out utilizing the simulation parameters as suggested in [11,13,14]). Moreover, as a benchmark, we consider a ZF precoding scheme without CE constraint (ZF-non). The simulations consider the commonly used massive MU-MIMO downlink wireless communication system. The specific simulation conditions are set as follows: Assuming that the communication channel $\mathbf{H}$ between the BS and the users is a standard complex Gaussian channel, that is, $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. SNR is defined as $\text{SNR} = P_T/\sigma^2$, where the transmit power is normalized to 1. All the simulation results are the average of 10^3 Monte Carlo simulations. According to Theorem 1, the penalty parameter $\rho > L$ and the step size $\mu = L$ are set in the proposed SoTMM algorithm. In particular, in Figure 4, one simulation experiment is enough to illustrate the convergence performance of the SoTMM algorithm.



(**a**) (**b**)

Figure 4. Convergence performance of the SoTMM algorithm with different system scales. (**a**) Outer iteration convergence performance; (**b**) Inner iteration convergence performance.

5.1. Convergence Analysis

Figure 4 depicts the inner and outer convergence performance of the proposed SoTMM algorithm in different system settings when SNR = 5 dB and 16QAM modulation. The convergence performance of the SoTMM algorithm is explained by checking the iteration gap of the inner and outer iteration stop conditions, respectively, that is,

$$\Delta \text{MSE} = \frac{\text{MSE}^{i+1} - \text{MSE}^{i}}{\text{MSE}^{i+1}}, \tag{36}$$

$$\Delta \mathbf{u} = \left\| \mathbf{u}^{k+1} - \mathbf{u}^{k} \right\|_2^2. \tag{37}$$

It can be concluded from Figure 4a that the SoTMM algorithm using the AltMin framework can achieve convergence in systems with different scales. As the number of BS antennas increases, the number of iterations required for the SoTMM algorithm to achieve convergence also increases. When the iterations exceed 30, the iteration gap of the SoTMM algorithm with different system sizes converge to 10^{-5}. Figure 4b illustrates the convergence of the inner iteration loop using the MM framework to solve $\mathbf{u}$. As can be seen from Figure 4b, similar to the outer iteration situation, as the number of BS antennas increases, more iterations are required for $\Delta\mathbf{u}$ to converge to 10^{-4}. Nonetheless, it only takes about 20 iterations for $\Delta\mathbf{u}$ to converge to 10^{-4} across different system sizes. Therefore, in the following simulations, the maximum iteration of the outer iteration loop is set to 40, and the maximum iteration of the inner iteration loop is set to 20.

5.2. CE Precoding

We compare the uncoded BER performance of the proposed SoTMM algorithm in the massive MU-MIMO system, where the BS is equipped with 128 transmit antennas to communicate with 16 single antenna users. First, we compare the uncoded BER performance of the algorithm when the transmit symbol is generated by the constant modulus constellation, that is, 16PSK. It can be seen from Figure 5 that the performance of the proposed SoTMM algorithm is better than that of FPG, GEMM, and GP-AltMin, and the SNR gap between the SoTMM and the ideal ZF is only 1 dB, which is promising. Different from PSK modulation, which can easily generate CE transmission signals, we pay more attention to the performance of the proposed algorithm in non-constant modulus modulation. Therefore, Figures 6 and 7 verify the performance of the proposed algorithm when the input constellation symbols $\mathbf{s}$ are generated by 16QAM and 64QAM modulation, respectively. From Figure 6, several conclusions can be drawn. First of all, compared with the transmit symbols using 16PSK modulation, the performance of all algorithms in the non-constant

modulus of 16QAM is improved, possibly because the distance between the constellation points of 16PSK is closer than that of 16QAM in the case of normalized transmit power, so the system performance will be reduced. Second, the ZF-CE precoding that directly projects the ZF precoded signal to the CE constraint has the worst performance, which is related to the limitations of the linear ZF precoding itself. Compared with the FPG and GEMM precoders, the proposed SoTMM precoder has the best performance. As can be seen from Figure 6, when the SNR exceeds 8 dB, as the SNR further increases, the performance advantage of the SoTMM precoder gradually becomes prominent. Compared with the ideal ZF precoding, the performance gap between the two precoders is about 1.2 dB. Finally, it can be seen from Figure 6 that the performance of the SoTMM and the GP-AltMin are similar. However, the GP-AltMin ignores the impact of noise when solving the CE precoded signals, thus improving performance. When considering the impact of noise on the CE precoded signals, the performance of the GP-AltMin needs to be further confirmed. When the constellation symbols $\mathbf{s}$ are generated using 64QAM modulation, the performance of all CE precoders decreases, among which the ZF-CE suffers an obvious performance loss. Similar to the 16QAM modulation situation, the proposed SoTMM precoder still has the lowest BER performance. When the uncoded BER is 10^{-4}, the uncoded BER performance gap between the ideal ZF precoding and the SoTMM is only 1.7 dB. This shows that the proposed SoTMM precoder has superior gains in uncoded BER performance compared with other CE precoders.

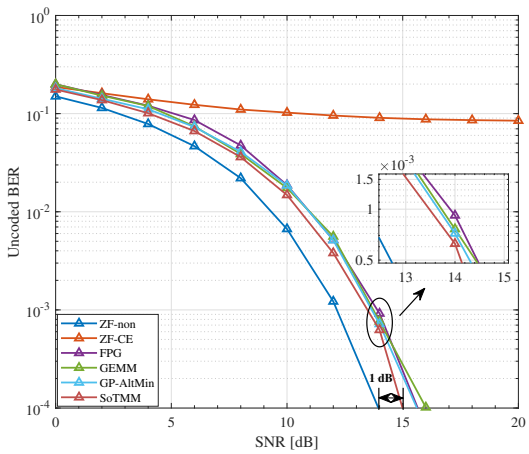


Figure 5. The uncoded BER performance for different CE precoders in $(N_{\mathrm{TX}}, N_u) = (128, 16)$ system, 16PSK.

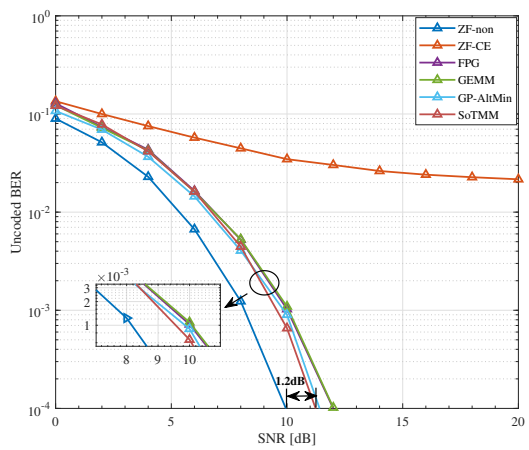


Figure 6. The uncoded BER performance for different CE precoders in $(N_{\mathrm{TX}}, N_u) = (128, 16)$ system, 16QAM.

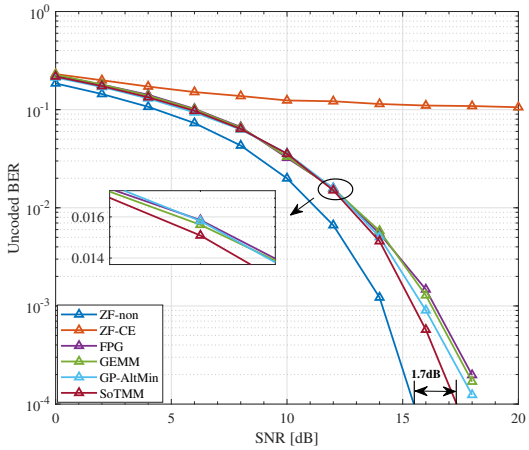


Figure 7. The uncoded BER performance for different CE precoders in $(N_{\mathrm{TX}}, N_u) = (128, 16)$ system, 64QAM.

5.3. DCE Precoding

In the previous simulations, the performance of the proposed SoTMM algorithm under CE constraint was verified. Next, we apply the SoTMM algorithm to the DCE constraint case to further elaborate on the performance of the proposed SoTMM algorithm. Figures 8–10 verify the uncoded BER performance of the SoTMM algorithm in different DCE cases. The simulations consider that the BS is equipped with 128 transmit antennas to communicate with 16 single antenna users. The constellation symbols $\mathbf{s}$ are generated by 16PSK, 16QAM, and 64QAM modulations, respectively. Some conclusions can be drawn from Figures 8–10. First, it can be seen from Figure 8 that, similar to the CE precoding case, the uncoded BER of the algorithm is slightly higher when the input symbol $\mathbf{s}$ is generated by 16PSK compared with 16QAM. Secondly, as the discrete resolution κ of the phase shifter increases, the uncoded BER performance of all algorithms continues to improve. Secondly, when the discrete resolutions are $\kappa = 2$ and $\kappa = 3$, the uncoded BER performance of the C2PO (C3PO) and the GP-AltMin are almost the same, and are significantly worse than the SoTMM. In particular, in Figure 9, when the discrete resolution is $\kappa = 3$, the uncoded BER performance of the proposed SoTMM is improved by about 3 dB compared with the C3PO and GP-AltMin. Third, when the discrete resolution of the phase shifter is raised to $\kappa = 4$, both the GP-AltMin and the SoTMM exhibit comparable uncoded BER performance, which closely approximates the performance achieved by the method when using a phase shifter with infinite resolution. Fourth, Figure 10 demonstrates that when the modulation is converted from 16QAM to 64QAM modulation, the uncoded BER performance of the 2-phase and 3-phase GP-AltMin and C2PO (C3PO) will change to unacceptable levels and gradually approach saturation. This error-floor problem can be greatly mitigated by increasing the discrete resolution. Even in the 2-phase and 3-phase cases of 64QAM modulation, the proposed SoTMM still has better uncoded BER performance. Finally, unlike the C2PO (C3PO) algorithm for specific discrete phases, the SoTMM algorithm is suitable for DCE precoding design with arbitrary phase shifter discrete resolution. In conclusion, the suggested SoTMM algorithm, as demonstrated in Figures 8–10, is more suitable for practical applications compared to the existing CE precoding technique and can be more flexibly extended to DCE precoding design.

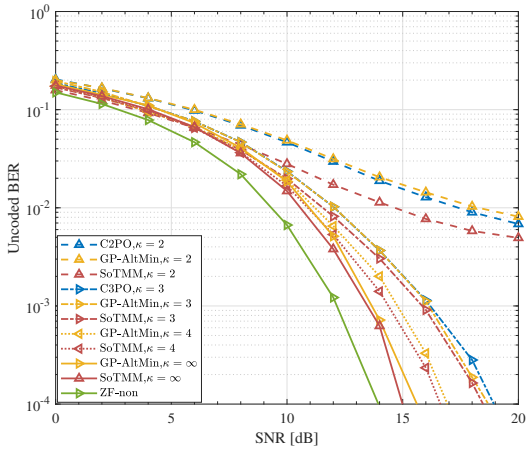


Figure 8. The uncoded BER performance for different CE precoders in different DCE cases. $(N_{\mathrm{TX}}, N_u) = (128, 16)$ system, 16PSK.

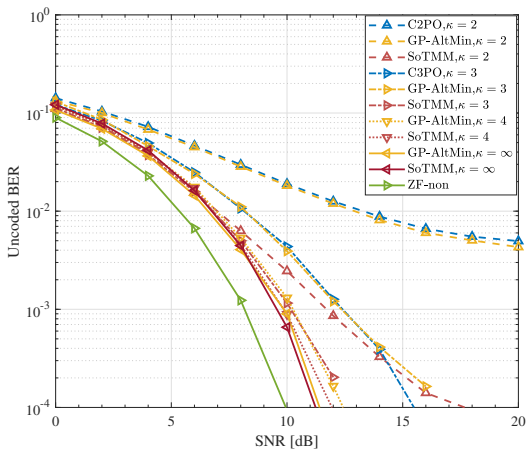


Figure 9. The uncoded BER performance for different CE precoders in different DCE cases. $(N_{\mathrm{TX}}, N_u) = (128, 16)$ system, 16QAM.

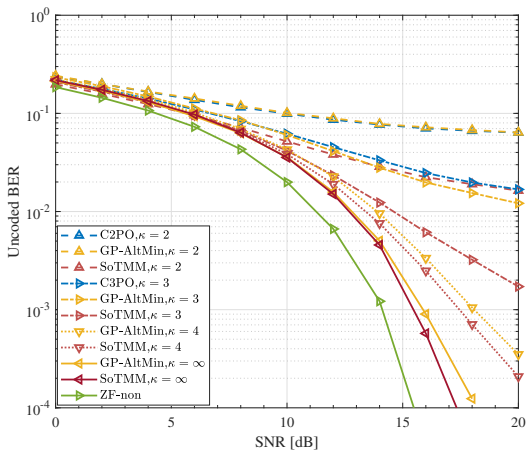


Figure 10. The uncoded BER performance for different CE precoders in different DCE cases. $(N_{\mathrm{TX}}, N_u) = (128, 16)$ system, 64QAM.

5.4. Complexity Analysis

In this subsection, we elucidate the complexity of the proposed SoTMM algorithm by quantifying the number of complex multiplications involved in the SoTMM precoding method within MU-MIMO systems of varying dimensions. The number of users in the massive MU-MIMO systems is set to 16. Table 1 displays the computational complexity of several CE precoding technologies. Figure 11 presents a comparison of the number of complex multiplications used by various CE precoding techniques in variable BS transmit antenna systems. Based on the findings presented in Figure 11, it is evident that the GP-AltMin exhibits the least computational complexity, while the GEMM follows closely behind. This is because only one-step projected gradient method is used in the GP-AltMin to optimize the CE precoded signal. However, the performance of algorithms using one-step projected gradient methods is usually unsatisfactory. The GP-AltMin artificially ignores the noise in the system to improve performance. However, improving the GP-AltMin in systems affected by noise may bring additional computational complexity. It is worth noting that while the GP-AltMin technique has low computational complexity, the SoTMM approach, as seen in Figures 9 and 10, can be more efficiently used for DCE case design. Compared with the GEMM algorithm, although the proposed SoTMM algorithm requires more computational complexity in the $(N_{\mathrm{TX}}, N_u) = (128, 16)$ system. Nevertheless, when considering the prior examination of uncoded BER performance, it is evident that the SoTMM method outperforms the GEMM algorithm in terms of uncoded BER performance (as shown in Figure 7, the performance gap is about 1 dB). Out of all the CE precoding techniques, FPG has the greatest computational complexity. This is because, as can be seen from Table 1, although the computational complexity of one iteration of the FPG algorithm is modest, in order to obtain the best performance, the FPG requires more iterations. This leads directly to the overall computational complexity of the FPG algorithm being too high. As a conclusion, the proposed SoTMM algorithm can efficiently strike a balance between uncoded BER performance and complexity.

Table 1. The complex multiplication of different CE precoding algorithms.

Methods	Maximum Iterations	Computational Complexity
FPG	$K = 5000$	$\mathcal{O}\big(K\big(5N_{\mathrm{TX}}N_u + 2N_u^2 + 2N_{\mathrm{TX}} + 5N_u\big)\big)$
GEMM	$K = 400$	$\mathcal{O}(K(N_{\mathrm{TX}}N_u + 3N_{\mathrm{TX}} + 8N_u))$
GP-AltMin	$K_1 = 20, K_2 = 5$	$\mathcal{O}(N_{\mathrm{TX}}N_u + K_1(N_{\mathrm{TX}}N_u + N_{\mathrm{TX}} + K_2(2N_{\mathrm{TX}}N_u + N_u)))$
SoTMM	$K_1 = 40, K_2 = 20$	$\mathcal{O}\big(K_1\big(2N_{\mathrm{TX}}^2 + 2N_{\mathrm{TX}}N_u + 3N_{\mathrm{TX}} + K_2\big(N_{\mathrm{TX}}^2 + N_{\mathrm{TX}}N_u + 4N_{\mathrm{TX}}\big)\big)\big)$

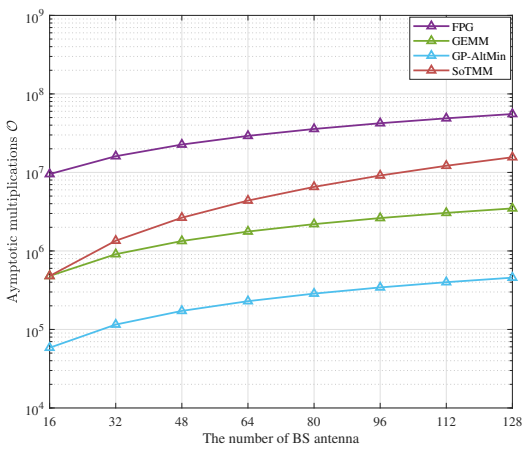


Figure 11. Comparison of computational complexity of different CE precoding algorithms, $N_u = 16$.

6. Conclusions

In this paper, a novel CE precoding scheme for massive MU-MIMO downlink systems is proposed. Different from existing methods, the CE precoded signal and precoding factor are designed to make the received signal approximate the transmit symbol as much as possible. This algorithm employs an AltMin framework that combines the MM method and the FISTA method to iteratively optimize variables. In particular, the second-order Taylor expansion and the properties of the massive MU-MIMO channel are employed to formulate a surrogate function that is effective in facilitating implementation of the MM technique. Furthermore, this approach is expanded into the DCE precoding design. This work thoroughly examines the exact property, convergence, and computational complexity of the suggested algorithm. The simulation findings demonstrate that this algorithm exhibits favorable uncoded BER performance and possesses computational efficiency, whether in the CE precoding case or DCE precoding case. In future, we intend to expand the scope of our work to a wider range of applications, such as combining low-resolution digital-to-analog converters precoding and CE transmission to further reduce the power consumption of the system.

Author Contributions: Methodology, formal analysis, and writing—review, R.L.; supervision and project administration, H.L.; formal analysis, Y.D.; editing, G.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant (No. 62271395) and the Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University (No. CX2021079).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Proof of Lemma 2

First, we can find that $f_\rho(\mathbf{u})$ is continuous in $[-1,1]^{N_{\mathrm{TX}}}$, so it is differentiable everywhere. Suppose $\mathbf{u}_1$ and $\mathbf{u}_2$ are two different input variables of $f_\rho(\mathbf{u})$, and satisfy $-\mathbf{1} \le \mathbf{u}_1 \le \mathbf{1}$ and $-\mathbf{1} \le \mathbf{u}_2 \le \mathbf{1}$, then, we can obtain

$$|f(\mathbf{u}_1) - f(\mathbf{u}_2)| = \left| \|\mathbf{s} - \tilde{\mathbf{H}}\mathbf{u}_1\|_2^2 - \|\mathbf{s} - \tilde{\mathbf{H}}\mathbf{u}_2\|_2^2 \right|. \tag{A1}$$

Using the matrix/vector operator norm inequality, we can obtain the following inequality

$$\begin{aligned} |f(\mathbf{u}_1) - f(\mathbf{u}_2)| &= \left| \|\mathbf{s} - \tilde{\mathbf{H}}\mathbf{u}_1\|_2^2 - \|\mathbf{s} - \tilde{\mathbf{H}}\mathbf{u}_2\|_2^2 \right| \\ &\le \left\| (\mathbf{s} - \tilde{\mathbf{H}}\mathbf{u}_1) - (\mathbf{s} - \tilde{\mathbf{H}}\mathbf{u}_2) \right\|_2^2 \\ &\le \|\tilde{\mathbf{H}}\|_2^2 \|\mathbf{u}_1 - \mathbf{u}_2\|_2^2. \end{aligned} \tag{A2}$$

On the other hand, since $-\mathbf{1} \le \mathbf{u}_1 \le \mathbf{1}$ and $-\mathbf{1} \le \mathbf{u}_2 \le \mathbf{1}$ hold, there is always

$$\|\mathbf{u}_1 - \mathbf{u}_2\|_2 \le 2\sqrt{N_{\mathrm{TX}}}. \tag{A3}$$

The equality holds if, and only if, $\mathbf{u}_1$ or $\mathbf{u}_2$ equals $\mathbf{1}$ and the other variable equals $-\mathbf{1}$. Substituting (A3) into the last inequality of (A2), we can obtain the upper bound of $|f(\mathbf{u}_1) - f(\mathbf{u}_2)|$ as

$$|f(\mathbf{u}_1) - f(\mathbf{u}_2)| \le 2\sqrt{N_t} \|\tilde{\mathbf{H}}\|_2^2 \|\mathbf{u}_1 - \mathbf{u}_2\|_2. \tag{A4}$$

According to $\left|f_\rho(\mathbf{u}_1) - f_\rho(\mathbf{u}_2)\right| / \|\mathbf{u}_1 - \mathbf{u}_2\|_2$ [37], we can obtain that (A4) is bounded as

$$L = 2\sqrt{N_{\mathrm{TX}}}\|\tilde{\mathbf{H}}\|_2^2, \tag{A5}$$

where L is a positive constant with $\|\tilde{\mathbf{H}}\|_2^2 \neq 0$.

Appendix B. Proof of Theorem 2

It is pointed out in [38] that gradient iteration $\mathbf{u}^{k+1} = \mathbf{z}^k - \mu^{-1}\nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right)$ can be regarded as the proximal regularization of the linearized function $\mathcal{G}$ at $\mathbf{x}^k$ [39]. Therefore, in the proposed SoTMM algorithm, the updated $\mathbf{u}^{k+1}$ in (20) can be equivalently rewritten as

$$\mathbf{u}^{k+1} = \min_{\mathbf{u}} \frac{\mu_k}{2}\left\|\mathbf{u}^k - \left(\mathbf{z}^k - \frac{1}{\mu}\nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right)\right)\right\|_2^2 + \mathcal{I}_{\mathcal{V}_3}(\mathbf{u}), \tag{A6}$$

where $\mathcal{I}_{\mathcal{V}_3}(\cdot)$ is the indicator function, that is, when $\mathbf{u} \in \mathcal{V}_3$, $\mathcal{I}_{\mathcal{V}_3}(\mathbf{u}) = 0$; when $\mathbf{u} \notin \mathcal{V}_3$, $\mathcal{I}_{\mathcal{V}_3}(\mathbf{u}) = \infty$. The first-order optimality condition of (A6) at point $\mathbf{u}^{k+1}$ is

$$\mu\left(\mathbf{u}^{k+1} - \mathbf{z}^k\right) + \nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right) + \partial\mathcal{I}_{\mathcal{V}_3}\left(\mathbf{u}^{k+1}\right) = \mathbf{0}. \tag{A7}$$

We assume $\mathbf{v}^{k+1} \in \partial\mathcal{I}_{\mathcal{V}_3}\left(\mathbf{u}^{k+1}\right)$, and define $\nabla\mathcal{F}\left(\mathbf{u}^{k+1}\right) = \mu\left(\mathbf{u}^{k+1} - \mathbf{z}^k\right) + \nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right)$ according to (20), then, $\mathbf{v}^{k+1} = -\nabla\mathcal{F}\left(\mathbf{u}^{k+1}\right)$. We can further rewrite (A7) as

$$\begin{aligned}
&\mathrm{dist}\left(\mathbf{0}, \mu\left(\mathbf{u}^{k+1} - \mathbf{z}^k\right) + \nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right) + \partial\mathcal{I}_{\mathcal{V}_3}\left(\mathbf{u}^{k+1}\right)\right) \\
\leq\ & \left\|\mu\left(\mathbf{u}^{k+1} - \mathbf{z}^k\right) + \nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right) + \mathbf{v}^{k+1}\right\|_2 \\
=\ & \left\|\mu\left(\mathbf{u}^{k+1} - \mathbf{z}^k\right) + \nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right) - \nabla\mathcal{F}\left(\mathbf{u}^{k+1}\right)\right\|_2 \\
\leq\ & \underbrace{\left\|\nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right) - \nabla\mathcal{F}\left(\mathbf{u}^{k+1}\right)\right\|_2}_{\text{first term}} + \underbrace{\left\|\mu\left(\mathbf{u}^{k+1} - \mathbf{z}^k\right)\right\|_2}_{\text{second term}},
\end{aligned} \tag{A8}$$

where $\mathrm{dist}(\mathbf{a},\chi) = \inf_{\mathbf{b}\in\chi}\|\mathbf{a} - \mathbf{b}\|_2$. Next, we will illustrate the convergence of the algorithm by analyzing the characteristics of the first and second terms of (A8). First, the first term of (A8) can be written as

$$\begin{aligned}
&\left\|\nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right) - \nabla\mathcal{F}\left(\mathbf{u}^{k+1}\right)\right\|_2 \\
\overset{(a)}{=}\ & \left\|\nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right) - \nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{u}^{k+1} \middle| \mathbf{u}^{k+1}\right)\right\|_2 \\
\leq\ & \left\|\nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{z}^k \middle| \mathbf{u}^k\right) - \nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{u}^k \middle| \mathbf{u}^k\right)\right\|_2 + \left\|\nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{u}^k \middle| \mathbf{u}^k\right) - \nabla_{\mathbf{u}}\mathcal{G}_\rho\left(\mathbf{u}^{k+1} \middle| \mathbf{u}^{k+1}\right)\right\|_2 \\
\overset{(b)}{\leq}\ & L_{\mathcal{G}}\left\|\mathbf{z}^k - \mathbf{u}^k\right\|_2 + L_{\mathcal{F}}\left\|\mathbf{u}^k - \mathbf{u}^{k+1}\right\|_2 \\
\overset{(c)}{=}\ & L_{\mathcal{G}}\alpha_k\left\|\mathbf{u}^k - \mathbf{u}^{k-1}\right\|_2 + L_{\mathcal{F}}\left\|\mathbf{u}^k - \mathbf{u}^{k+1}\right\|_2,
\end{aligned} \tag{A9}$$

where $L_{\mathcal{G}}$ is the Lipschitz constant of $\mathcal{G}_\rho$, and $L_{\mathcal{F}}$ is the Lipschitz constant of $\mathcal{F}$. In the derivation of (A9), (a) is derived from $\nabla_{\mathbf{u}}\mathcal{G}_\rho(\mathbf{u}|\mathbf{u}) = \nabla\mathcal{F}(\mathbf{u})$; (b) is derived from the Lipschitz continuity of $\nabla\mathcal{F}(\mathbf{u})$ and $\nabla_{\mathbf{u}}\mathcal{G}_\rho(\mathbf{u}|\mathbf{u})$; (c) is derived from (23).

The second term in (A8) is analyzed below. According to (23), we can obtain

$$\begin{aligned}\left\|\mu\left(\mathbf{u}^{k+1}-\mathbf{z}^{k}\right)\right\|_2 &= \mu\left\|\mathbf{u}^{k}-\mathbf{u}^{k+1}+\alpha_k\left(\mathbf{u}^{k}-\mathbf{u}^{k-1}\right)\right\|_2 \\ &\leq \mu\left\|\mathbf{u}^{k}-\mathbf{u}^{k+1}\right\|_2+\mu\alpha_k\left\|\mathbf{u}^{k}-\mathbf{u}^{k-1}\right\|_2.\end{aligned} \tag{A10}$$

Combining (A8)–(A10), we can obtain

$$\begin{aligned}&\operatorname{dist}\left(\mathbf{0},\mu\left(\mathbf{u}^{k+1}-\mathbf{z}^{k}\right)+\nabla_{\mathbf{u}}\mathcal{G}_{\rho}\left(\mathbf{z}^{k}|\mathbf{u}^{k}\right)+\partial\mathcal{I}_{\mathcal{V}_3}\left(\mathbf{u}^{k+1}\right)\right) \\ &\leq L_{\mathcal{G}}\alpha_k\left\|\mathbf{u}^{k}-\mathbf{u}^{k-1}\right\|_2+L_{\mathcal{F}}\left\|\mathbf{u}^{k}-\mathbf{u}^{k+1}\right\|_2+\mu\left\|\mathbf{u}^{k}-\mathbf{u}^{k+1}\right\|_2+\mu\alpha_k\left\|\mathbf{u}^{k}-\mathbf{u}^{k-1}\right\|_2 \\ &\leq C\left(\left\|\mathbf{u}^{k}-\mathbf{u}^{k-1}\right\|_2+\left\|\mathbf{u}^{k}-\mathbf{u}^{k+1}\right\|_2\right),\end{aligned} \tag{A11}$$

where $C=\max((L_{\mathcal{G}}+\mu)\bar{\alpha},L_{\mathcal{F}}+\mu)$. Therefore, it can be seen that $\mathcal{F}$ has a finite lower bound, so let $\mathcal{F}^{*}\triangleq\inf_{\mathbf{u}\in\mathcal{V}_3}\mathcal{F}(\mathbf{u})$. Consider the following Lemma A1 [40].

Lemma A1. *Define* $\mathbf{u}^{+}=\prod_{\chi}\left(\mathbf{z}-\mu^{-1}\nabla\mathcal{H}(\mathbf{z})\right)$, *where* $\mathbf{z}=\mathbf{u}+\alpha(\mathbf{u}-\bar{\mathbf{u}})$, $\mathbf{u},\bar{\mathbf{u}}\in\chi$, $0\leq\alpha\leq\bar{\alpha}$, *and* $\mathcal{H}(\cdot)$ *have Lipschitz continuous gradients; the feasible set* χ *is convex; the gradient step size* μ *makes the* $\mathcal{H}(\cdot)$ *satisfy the descent property during the iteration process, that is,* $\mathcal{H}(\mathbf{u}^{+})\leq\mathcal{H}(\mathbf{z})+\langle\nabla\mathcal{H}(\mathbf{z}),\mathbf{u}^{+}-\mathbf{z}\rangle+\frac{\mu}{2}\|\mathbf{u}^{+}-\mathbf{z}\|_2^2$. *Then, we obtain* $\mathcal{H}(\mathbf{z})-\mathcal{H}(\mathbf{u}^{+})\geq\frac{\mu}{2}\|\mathbf{u}^{+}-\mathbf{u}\|_2^2-\alpha^2\|\mathbf{u}-\bar{\mathbf{u}}\|_2^2$.

According to the update rule (21) to (23) of the SoTMM algorithm, we can obtain

$$\begin{aligned}\mathcal{F}\left(\mathbf{u}^{k}\right)-\mathcal{F}\left(\mathbf{u}^{k+1}\right) &\overset{(a)}{\geq}\mathcal{G}_{\rho}\left(\mathbf{u}^{k}\middle|\mathbf{u}^{k}\right)-\mathcal{G}_{\rho}\left(\mathbf{u}^{k+1}\middle|\mathbf{u}^{k}\right) \\ &\overset{(b)}{\geq}\frac{\mu}{2}\left(\left\|\mathbf{u}^{k+1}-\mathbf{u}^{k}\right\|_2^2-\bar{\alpha}^2\left\|\mathbf{u}^{k}-\mathbf{u}^{k-1}\right\|_2^2\right),\end{aligned} \tag{A12}$$

where (a) is obtained from $\mathcal{F}(\mathbf{u})=\mathcal{G}_{\rho}(\mathbf{u}|\mathbf{u})$ and $\mathcal{G}_{\rho}(\mathbf{u}|\bar{\mathbf{u}})\geq\mathcal{F}(\bar{\mathbf{u}})$; (b) is obtained from Lemma A1 and $\alpha\leq\bar{\alpha}$. Therefore, it can be obtained

$$\begin{aligned}&\mathcal{F}\left(\mathbf{u}^{0}\right)-\mathcal{F}\left(\mathbf{u}^{k+1}\right) \\ &=\sum_{m=0}^{k}\left(\mathcal{F}(\mathbf{u}^{m})-\mathcal{F}\left(\mathbf{u}^{m+1}\right)\right) \\ &\geq\sum_{m=0}^{k}\frac{\mu}{2}\left(\left\|\mathbf{u}^{m+1}-\mathbf{u}^{n}\right\|_2^2-\bar{\alpha}^2\left\|\mathbf{u}^{m}-\mathbf{u}^{m-1}\right\|_2^2\right) \\ &=\sum_{m=0}^{k-1}\frac{\mu}{2}\left\|\mathbf{u}^{m+1}-\mathbf{u}^{m}\right\|_2^2-\sum_{m=-1}^{k-1}\frac{\mu\bar{\alpha}^2}{2}\left\|\mathbf{u}^{m+1}-\mathbf{u}^{m}\right\|_2^2+\frac{\mu}{2}\left\|\mathbf{u}^{k+1}-\mathbf{u}^{k}\right\|_2^2 \\ &=\sum_{m=0}^{k-1}\frac{(1-\bar{\alpha}^2)\mu}{2}\left\|\mathbf{u}^{m+1}-\mathbf{u}^{m}\right\|_2^2+\frac{\mu}{2}\left\|\mathbf{u}^{k+1}-\mathbf{u}^{k}\right\|_2^2 \\ &\geq\sum_{m=0}^{k}\frac{(1-\bar{\alpha}^2)\mu}{2}\left\|\mathbf{u}^{m+1}-\mathbf{u}^{m}\right\|_2^2.\end{aligned} \tag{A13}$$

We can further rewrite (A13) as

$$\begin{aligned}\mathcal{F}\left(\mathbf{u}^{0}\right)-\mathcal{F}\left(\mathbf{u}^{k+1}\right) &\geq\frac{1}{2}\sum_{m=0}^{k}C_1\left\|\mathbf{u}^{m+1}-\mathbf{u}^{m}\right\|_2^2+\frac{1}{2}\sum_{m=1}^{k+1}C_1\left\|\mathbf{u}^{m}-\mathbf{u}^{m-1}\right\|_2^2 \\ &\overset{(a)}{=}C_2\min_{m=0,\cdots,k}\left(\left\|\mathbf{u}^{m+1}-\mathbf{u}^{m}\right\|_2^2+\left\|\mathbf{u}^{m}-\mathbf{u}^{m-1}\right\|_2^2\right),\end{aligned} \tag{A14}$$

where $C_1 = (1-\bar{\alpha}^2)\mu/2$, $C_2 = (k+1)C_1/2$. (a) is obtained from $\mathbf{x}^0 = \mathbf{x}^{-1}$. Based on the relationship $a+b \leq \sqrt{2(a^2+b^2)}$, (A14) can be further rewritten as

$$\min_{m=0,\cdots,k} \left(\left\| \mathbf{u}^{m+1} - \mathbf{u}^m \right\|_2^2 + \left\| \mathbf{u}^m - \mathbf{u}^{m-1} \right\|_2^2 \right) \leq \sqrt{\frac{8}{(k+1)(1-\bar{\alpha}^2)\mu}(\mathcal{F}(\mathbf{u}^0) - \mathcal{F}^*)}. \quad \text{(A15)}$$

Substituting (A15) into (A11), we can obtain

$$\begin{aligned} \min_{m=0,\cdots,k} \operatorname{dist}\left(\mathbf{0}, \nabla\mathcal{F}\left(\mathbf{u}^{m+1}\right) + \partial\mathcal{I}_{\mathcal{V}_3}\left(\mathbf{u}^{m+1}\right)\right) &\leq C \min_{m=0,\cdots,k} \left(\left\| \mathbf{u}^m - \mathbf{u}^{m-1} \right\|_2 + \left\| \mathbf{u}^m - \mathbf{u}^{m+1} \right\|_2 \right) \\ &\leq C\sqrt{\frac{8}{(k+1)(1-\bar{\alpha}^2)\mu}(\mathcal{F}(\mathbf{u}^0) - \mathcal{F}^*)}. \end{aligned} \quad \text{(A16)}$$

In summary, it can be seen that when (A16) is established, $\mathbf{u}^{m+1} \in \mathcal{V}_3$ is the stationary point of the proposed algorithm.

References

1. Marzetta, T.L. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans. Wirel. Commun.* **2010**, *9*, 3590–3600. [CrossRef]
2. Larsson, E.G.; Edfors, O.; Tufvesson, F.; Marzetta, T.L. Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **2014**, *52*, 186–195. [CrossRef]
3. Lu, L.; Li, G.Y.; Swindlehurst, A.L.; Ashikhmin, A.; Zhang, R. An overview of massive MIMO: Benefits and challenges. *IEEE J. Sel. Top. Signal Process.* **2014**, *8*, 742–758. [CrossRef]
4. Rezaei, F.; Tadaion, A. Multi-layer beamforming in uplink/downlink massive MIMO systems with multi-antenna users. *Signal Process.* **2019**, *164*, 58–66. [CrossRef]
5. Li, A.; Spano, D.; Krivochiza, J.; Domouchtsidis, S.; Tsinos, C.G.; Masouros, C.; Chatzinotas, S.; Li, Y.; Vucetic, B.; Ottersten, B. A tutorial on interference exploitation via symbol-level precoding: Overview, state-of-the-art and future directions. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 796–839. [CrossRef]
6. Mohammed, S.K.; Larsson, E.G. Single-user beamforming in large-scale MISO systems with per-antenna constant-envelope constraints: The doughnut channel. *IEEE Trans. Wirel. Commun.* **2012**, *11*, 3992–4005. [CrossRef]
7. Zhang, S.; Zhang, R.; Lim, T.J. MISO multicasting with constant envelope precoding. *IEEE Wirel. Commun. Lett.* **2016**, *5*, 588–591. [CrossRef]
8. Zhang, S.; Zhang, R.; Lim, T.J. Constant envelope precoding for MIMO systems. *IEEE Trans. Commun.* **2018**, *66*, 149–162. [CrossRef]
9. Mohammed, S.K.; Larsson, E.G. Per-antenna constant envelope precoding for large multi-user MIMO systems. *IEEE Trans. Commun.* **2013**, *61*, 1059–1071. [CrossRef]
10. Chen, J.C.; Wen, C.K.; Wong, K.K. Improved constant envelope multiuser precoding for massive MIMO systems. *IEEE Commun. Lett.* **2014**, *18*, 1311–1314. [CrossRef]
11. Chen, J.C. Efficient constant envelope precoding with quantized phases for massive MU-MIMO downlink systems. *IEEE Trans. Veh. Technol.* **2019**, *68*, 4059–4063. [CrossRef]
12. Zhang, S.; Zhang, R.; Lim, T.J. Constant envelope precoding with adaptive receiver constellation in MISO fading channel. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 6871–6882. [CrossRef]
13. Shao, M.; Li, Q.; Ma, W.K.; So, A.M.C. Minimum symbol error rate-based constant envelope precoding for multiuser massive MISO downlink. In Proceedings of the IEEE Statistical Signal Processing Workshop, SSP, Freiburg im Breisgau, Germany, 10–13 June 2018; pp. 727–731.
14. Shao, M.; Li, Q.; Ma, W.K.; So, A.M.C. A framework for one-bit and constant-envelope precoding over multiuser massive MISO channels. *IEEE Trans. Signal Process.* **2019**, *67*, 5309–5324. [CrossRef]
15. Wang, Y.; Chen, N.; Liu, F.; Li, A.; Zhou, J.; Masouros, C. Constant Envelope Precoding With Extended Degrees of Freedom Through Per-User Symbol Scaling. *IEEE Commun. Lett.* **2021**, *25*, 1620–1624. [CrossRef]
16. Amadori, P.V.; Masouros, C. Constant envelope precoding by interference exploitation in phase shift keying-modulated multiuser transmission. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 538–550. [CrossRef]
17. Amadori, P.V.; Masouros, C. Constructive interference based constant envelope precoding. In Proceedings of the 2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Edinburgh, UK, 3–6 July 2016; pp. 1–5.
18. Liu, F.; Masouros, C.; Amadori, P.V.; Sun, H. An efficient manifold algorithm for constructive interference based constant envelope precoding. *IEEE Signal Process. Lett.* **2017**, *24*, 1542–1546. [CrossRef]
19. Noll, A.; Jedda, H.; Nossek, J. PSK precoding in multi-user MISO systems. In Proceedings of the 1th IWSA—International ITG Workshop Smart Antennas, Berlin, Germany, 15–17 March 2017; pp. 1–7.

20. Nedelcu, A.; Steiner, F.; Staudacher, M.; Kramer, G.; Zirwas, W.; Ganesan, R.S.; Baracca, P.; Wesemann, S. Quantized precoding for multi-antenna downlink channels with MAGIQ. In Proceedings of the WSA 2018; 22nd International ITG Workshop on Smart Antennas, Bochum, Germany, 14–16 March 2018; pp. 1–8.
21. Jedda, H.; Mezghani, A.; Swindlehurst, A.L.; Nossek, J.A. Quantized constant envelope precoding with PSK and QAM signaling. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 8022–8034. [CrossRef]
22. Jacobsson, S.; Durisi, G.; Coldrey, M.; Goldstein, T.; Studer, C. Quantized precoding for massive MU-MIMO. *IEEE Trans. Commun.* **2017**, *65*, 4670–4684. [CrossRef]
23. Jacobsson, S.; Durisi, G.; Coldrey, M.; Goldstein, T.; Studer, C. Nonlinear 1-bit precoding for massive MU-MIMO with higher-order modulation. In Proceedings of the 2016 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 6–9 November 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 763–767.
24. Castañeda, O.; Jacobsson, S.; Durisi, G.; Coldrey, M.; Goldstein, T.; Studer, C. 1-bit massive MU-MIMO precoding in VLSI. *IEEE J. Emer. Select. Top. Circu. Syste.* **2017**, *7*, 508–522. [CrossRef]
25. Waldspurger, I.; d'Aspremont, A.; Mallat, S. Phase recovery, maxcut and complex semidefinite programming. *Math. Program.* **2015**, *149*, 47–81. [CrossRef]
26. So, A.M.C.; Zhang, J.; Ye, Y. On approximating complex quadratic optimization problems via semidefinite programming relaxations. *Math. Program.* **2007**, *110*, 93–110. [CrossRef]
27. Luo, Z.Q.; Ma, W.K.; So, A.M.C.; Ye, Y.; Zhang, S. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Process. Mag.* **2010**, *27*, 20–34. [CrossRef]
28. Shao, M.; Dai, Q.; Ma, W.K. Extreme-Point Pursuit for Unit-Modulus Optimization. In Proceedings of the ICASSP IEEE International Conference on Acoustics, Speech and Signal Processing, Singapore, 23–27 May 2022; pp. 5548–5552.
29. Tranter, J.; Sidiropoulos, N.D.; Fu, X.; Swami, A. Fast unit-modulus least squares with applications in beamforming. *IEEE Trans. Signal Process.* **2017**, *65*, 2875–2887. [CrossRef]
30. Fan, J.; Han, F.; Liu, H. Challenges of big data analysis. *Natl. Sci. Rev.* **2014**, *1*, 293–314. [CrossRef] [PubMed]
31. Sun, Y.; Babu, P.; Palomar, D.P. Majorization-minimization algorithms in signal processing, communications, and machine learning. *IEEE Trans. Signal Process.* **2016**, *65*, 794–816. [CrossRef]
32. Hunter, D.R.; Lange, K. A tutorial on MM algorithms. *Am. Stat.* **2004**, *58*, 30–37. [CrossRef]
33. Lange, K. *MM Optimization Algorithms*; SIAM: Philadelphia, PA, USA, 2016.
34. Song, J.; Babu, P.; Palomar, D.P. Optimization methods for designing sequences with low autocorrelation sidelobes. *IEEE Trans. Signal Process.* **2015**, *63*, 3998–4009. [CrossRef]
35. Arora, A.; Tsinos, C.G.; Rao, B.S.M.R.; Chatzinotas, S.; Ottersten, B. Hybrid transceivers design for large-scale antenna arrays using majorization-minimization algorithms. *IEEE Trans. Signal Process.* **2019**, *68*, 701–714. [CrossRef]
36. Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2009**, *2*, 183–202. [CrossRef]
37. Haqiqatnejad, A.; Kayhan, F.; ShahbazPanahi, S.; Ottersten, B. Finite-Alphabet Symbol-Level Multiuser Precoding for Massive MU-MIMO Downlink. *IEEE Trans. Signal Process.* **2021**, *69*, 5595–5610. [CrossRef]
38. Pedregal, P. *Introduction to Optimization*; Springer: Berlin/Heidelberg, Germany, 2004; Volume 46.
39. Martinet, B. Regularisation d'inequations variationelles par approximations successives. *Rev. Fr. D'informatique Rech. Oper.* **1970**, *4*, 154–159.
40. Xu, Y.; Yin, W. A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. *SIAM J. Imaging Sci.* **2013**, *6*, 1758–1789. [CrossRef]