

Broad Coverage Precoding for 3D Massive MIMO with Huge Uniform Planar Arrays

An-An Lu ^{1,2} , Yan Chen ¹ and Xiqi Gao ^{1,2,*}

¹ National Mobile Communications Research Laboratory (NCRL), Southeast University, Nanjing 210096, China; aalu@seu.edu.cn (A.-A.L.); 213160372@seu.edu.cn (Y.C.)

² Purple Mountain Laboratories, Nanjing 211111, China

* Correspondence: xqgao@seu.edu.cn

Abstract: In this paper, we propose a novel broad coverage precoder design for three-dimensional (3D) massive multi-input multi-output (MIMO) equipped with huge uniform planar arrays (UPAs). The desired two-dimensional (2D) angle power spectrum is assumed to be separable. We use the per-antenna constant power constraint and the semi-unitary constraint which are widely used in the literature. For normal broad coverage precoder design, the dimension of the optimization space is the product of the number of antennas at the base station (BS) and the number of transmit streams. With the proposed method, the design of the high-dimensional precoding matrices is reduced to that of a set of low-dimensional orthonormal vectors, and of a pair of low-dimensional vectors. The dimensions of the vectors in the set and the pair are the number of antennas per column and per row of the UPA, respectively. We then use optimization methods to generate the set of orthonormal vectors and the pair of vectors, respectively. Finally, simulation results show that the proposed broad coverage precoding matrices achieve nearly the same performance as the normal broad coverage precoder with much lower computational complexity.

Keywords: massive multi-input multi-output (MIMO); broad coverage precoder design; uniform planar array (UPA); high dimension



Citation: Lu, A.-A.; Chen, Y.; Gao, X. Broad Coverage Precoding for 3D Massive MIMO with Huge Uniform Planar Arrays. *Entropy* **2021**, *23*, 887. <https://doi.org/10.3390/e23070887>

Academic Editor: Song-Nam Hong

Received: 12 June 2021

Accepted: 6 July 2021

Published: 13 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Massive multiple-input multiple-output (MIMO) [1,2] is one of the key enabling technologies of fifth generation (5G) wireless communications systems. It provides huge spatial multiplexing gains and high energy efficiency by employing a large number of antennas at a base station (BS). In the process of establishing links between the BS and the users, the broadcasting and control information plays an important role. In practical systems, uniform planar array (UPA) antennas are preferred to uniform linear array (ULA) antennas due to their compact size. In this paper, we focus on the broad coverage precoder design for synchronization and control information transmission in a massive MIMO equipped with UPAs.

1.1. Existing Works

For massive MIMO systems equipped with ULAs, omnidirectional precoding has been proposed in the literature [3–5] recently. It is designed by assuming the channel state information (CSI) is unavailable. In [3], the received mean signal power needed to be constant at finite discrete angles under the per-antenna power constraint and the semi-unitary constraint. In [4,5], the space-time codes are used to design a precoder for omnidirectional transmission at any time. The received signal power is designed to be constant at finite discrete angles in [4] and any angle in [5].

In omnidirectional precoder design for massive MIMO systems with ULAs, Golay complementary sequences play a key role. These sequences were introduced by Golay [6] in 1951. The pair of complementary sequences was then generalized to a set of complementary

sequences [7], which are also called Welty codes [8]. Furthermore, Golay complementary sequences are generalized to Golay complementary arrays [9,10] or high-dimensional Welty codes, which have been investigated in [11,12]. Inspired by the Golay complementary arrays and the high-dimensional Welty codes, a novel and low-complexity array design is proposed in [13] to construct the omnidirectional precoding matrices for massive MIMO with UPAs. Omnidirectional precoder design for massive MIMO with UPAs has also been investigated in [14–16] based on complementary codes, numerical optimization methods and Golay complementary matrices.

In practical systems, omnidirectional precoding might cause interference with other cells or sectors. To solve this problem, broad coverage precoding is proposed in [17]. It is designed for massive MIMO with ULAs based on manifold optimization under the per-antenna constant power and semi-unitary constraints. By considering the synchronization performance, the work in [17] is then extended to [18,19] for broad coverage precoder design for three-dimensional (3D) massive MIMO in terrestrial and satellite communication systems, respectively. However, as the size of antennas is still increasing, the complexity of the broad coverage precoder design for massive MIMO with a large number of antennas will be a problem. For example, the optimization in [18] is performed over the space of multiple matrices and thus will be too complicated for huge antenna arrays.

1.2. Main Contribution

In this paper, we propose a broad coverage precoder design for 3D massive MIMO with huge UPAs. The two-dimensional (2D) angle power spectrum for the broad coverage precoding is assumed to be separable. We also use the per-antenna power constraint and the semi-unitary constraint. The number of antennas equipped at the BS is $M_t = M_z M_x$. The number of transmit streams is r . With the proposed design method, the design of the high-dimensional $M_t \times r$ precoding matrices is reduced to that of a pair of low-dimensional $M_x \times 1$ vectors, and of a set with length r of low-dimensional $M_z \times 1$ orthonormal vectors.

The main contributions of the paper are summarized as follows: Theorem 1 provides a method to generate high-dimensional precoding matrices with a separate 2D angle power spectrum from low-dimensional vectors. Theorem 2 proves the conditions for low-dimensional vectors to generate high-dimensional precoding matrices satisfying the per-antenna power constraint. Based on Theorems 1 and 2, a low-complexity broad coverage precoder design is proposed. Furthermore, manifold optimization is used to generate a set of orthonormal vectors with the required angle power spectrum. The design of two vectors with nonzero elements with a constant envelope and required angle power spectrum is also provided.

The rest of the paper is organized as follows. The system model and problem formulation are provided in Section 1. The proposed method for designing 2D broad coverage precoding matrices is given in Section 3. Simulation results are presented in Section 4. The conclusion is drawn in Section 5. A lemma and proofs of theorems are in the Appendix.

Notations: Throughout this paper, lowercase and uppercase boldface letters are used to denote column vectors and matrices, respectively. The superscripts $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ indicate the conjugate, transpose and conjugate transpose operation, respectively. The operator \otimes is used to denote the Kronecker product. The notation $\text{tr}(\cdot)$ represents the trace function of a matrix. The operator $\text{diag}(\mathbf{A})$ represents the diagonal matrix obtained by using the diagonal elements of \mathbf{A} . The identity matrix with dimension $N \times N$ is denoted by \mathbf{I}_N .

2. System Model and Problem Formulation

2.1. System Model

We consider a 3D massive MIMO system where a UPA is equipped at a BS. The number of antennas in the antenna array is $M_t = M_z M_x$, where M_z and M_x denote the numbers of antennas in each row and column, respectively. In this paper, we focus on the problem of broad coverage precoder design for 3D massive MIMO systems. We define \mathbf{h}

as the $M_t \times 1$ channel vector from the BS to a point $P = (\rho, \theta, \phi)$ at the free space, where ρ denotes the distance between the locations of the BS and the point P , and θ and ϕ represent the polar and azimuthal angles, respectively. We assume r transmitted data streams are used in precoding. At the point P , the received signal can be written as

$$y = \mathbf{h}\mathbf{P}\mathbf{x} + z, \tag{1}$$

where the symbol \mathbf{P} denotes the $M_t \times r$ precoding matrix, \mathbf{x} is the $r \times 1$ transmitted vector and z is a complex Gaussian noise variable. The mean and variance of z is 0 and σ_z^2 . Let \mathbf{p}_i denote the vector by extracting the i -th column of \mathbf{P} . We reshape the vector \mathbf{p}_i into a matrix of $\mathbb{C}^{M_z \times M_x}$ and use \mathbf{P}_i to denote it. Then, the matrix \mathbf{P} can be re-expressed as

$$\mathbf{P} = [\mathbf{p}_1 \ \mathbf{p}_2 \ \cdots \ \mathbf{p}_r] = [\text{vec}(\mathbf{P}_1) \ \text{vec}(\mathbf{P}_2) \ \cdots \ \text{vec}(\mathbf{P}_r)]. \tag{2}$$

2.2. Problem Formulation

In practical massive MIMO systems, the precoder for public information transmission should be specifically designed to provide realistic broad coverage. The required 2D angle power spectrum in the free space is usually given according to the system parameters.

Let $\mathbf{U}(\theta, \phi)$ be the steering matrix in the direction (θ, ϕ) with the elements being defined as

$$[\mathbf{U}(\theta, \phi)]_{kl} = e^{-(k-1)\frac{2j\pi d_z \cos \theta}{\lambda}} e^{-j(l-1)\frac{2j\pi d_x \sin \theta \cos \phi}{\lambda}}, \tag{3}$$

where $k = 1, 2, \dots, M_z$ and $l = 1, 2, \dots, M_x$. The spacings d_z and d_x are assumed to be equal to 0.5λ . For convenience, we set $u = \frac{1}{2}\cos \theta$ and $v = \frac{1}{2}\sin \theta \cos \phi$ as the scaled directional cosines with respect to the z and x axes, respectively. Then, the steering matrix $\mathbf{U}(\theta, \phi)$ in Equation (3) can be rewritten as

$$\mathbf{U}(u, v) = \mathbf{v}_z(u) \otimes \mathbf{v}_x(v)^T, \tag{4}$$

where $\mathbf{v}_z(u)$ and $\mathbf{v}_x(v)$ are the steering vectors in vertical and horizontal directions, defined by

$$\mathbf{v}_z(u) = \frac{1}{\sqrt{M_z}} [1 \ e^{-j2\pi u} \ \cdots \ e^{-j(M_z-1)2\pi u}]^T \tag{5}$$

$$\mathbf{v}_x(v) = \frac{1}{\sqrt{M_x}} [1 \ e^{-j2\pi v} \ \cdots \ e^{-j(M_x-1)2\pi v}]^T. \tag{6}$$

Let $\mathbf{a}(u, v)$ be the steering vector obtained by vectorizing the steering matrix $\mathbf{U}(u, v)$ as

$$\mathbf{a}(u, v) = \text{vec}(\mathbf{U}(u, v)). \tag{7}$$

Then, we have

$$\mathbf{a}(u, v) = \mathbf{v}_x(v) \otimes \mathbf{v}_z(u). \tag{8}$$

We assume the desired 2D angle power spectrum for broad coverage is separable in u, v , i.e., the 2D angle power spectrum to the angle (u, v) is given by

$$\sum_{i=1}^r |\mathbf{a}(u, v)^T \text{vec}(\mathbf{P}_i)|^2 = a_z(u)a_x(v), \tag{9}$$

where $a_z(u)$ and $a_x(v)$ denote the functions of the scaled directional cosines u and v . Substituting Equation (8) into Equation (9), we then obtain

$$\sum_{i=1}^r |\mathbf{v}_z(u)^T \mathbf{P}_i \mathbf{v}_x(v)|^2 = a_z(u)a_x(v). \tag{10}$$

For brevity, we define

$$s_{\mathbf{P}_i}(u, v) = |\mathbf{v}_z(u)^T \mathbf{P}_i \mathbf{v}_x(v)|^2 \quad (11)$$

then we have

$$\sum_{i=1}^r s_{\mathbf{P}_i}(u, v) = a_z(u) a_x(v). \quad (12)$$

After discussing the desired 2D angle power spectrum, we need to consider the constraints that the precoding matrices need to satisfy. To maximize the power efficiency of all the antennas, a commonly used constraint in the literature is the per-antenna constant power constraint. Let $p_{i,kl}$ denote the k, l -th element in the matrix \mathbf{P}_i . We formulate this constraint as

$$\sum_{i=1}^r |p_{i,kl}|^2 = 1, \quad \forall k, l. \quad (13)$$

For a public channel, the CSI is usually not known. To maximize the mutual information when the channel vector \mathbf{h} has independent and identically distributed (i.i.d.) entries with zero mean and unit variance [20], we then consider the semi-unitary constraint

$$\mathbf{P}^H \mathbf{P} = \frac{M_t}{r} \mathbf{I}_r. \quad (14)$$

In conclusion, the broad coverage precoding matrices need to be designed by using the following three conditions:

$$\sum_{i=1}^r s_{\mathbf{P}_i}(u, v) = a_z(u) a_x(v) \quad (15)$$

$$\sum_{i=1}^r |p_{i,kl}|^2 = 1, \quad \forall k, l \quad (16)$$

$$\mathbf{P}^H \mathbf{P} = \frac{M_t}{r} \mathbf{I}_r. \quad (17)$$

When $a_z(u) = 1$ and $a_x(v) = 1$, the problem is reduced to that for omnidirectional precoder design in [13]. In this paper, we aim to solve the problem for general $a_z(u)$ and $a_x(v)$.

3. Broad Coverage Precoding Matrix Design for UPAs

In this section, we propose the design of the broad coverage precoder for 3D massive MIMO with UPAs.

3.1. Preliminary Background

Before presenting the broad coverage precoder design, we give some definitions which will be used in the paper. Let $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r\}$ and $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_r\}$ be two sets of $M_z \times 1$ vectors, respectively. Their aperiodic cross-correlation is defined as

$$c_{\mathcal{D}, \mathcal{F}}(t) = \sum_{i=1}^r \sum_{j=1}^{M_z-t} [\mathbf{d}_i]_j [\mathbf{f}_i]_{j+t}^* = c_{\mathcal{F}, \mathcal{D}}(-t)^*, \quad (18)$$

where $t = 0, 1, \dots, M_z - 1$. For convenience, we define the angle power spectrum of column vectors and row vectors in a different way. The angle power spectrum of a column vector \mathbf{d}_i is defined as

$$s_{\mathbf{d}_i}(u) = |\mathbf{v}_z(u)^T \mathbf{d}_i|^2. \quad (19)$$

The angle power spectrum of a row vector $\boldsymbol{\phi}$ is defined as

$$s_{\boldsymbol{\phi}}(v) = |\boldsymbol{\phi}^T \mathbf{v}_x(v)|^2. \quad (20)$$

Let \mathbf{a} denote a vector $[a_1, a_2, \dots, a_n]^T$ and \mathbf{a}^\dagger denote its complex conjugate of the reversal, i.e., $\mathbf{a}^\dagger = [a_n^* a_{n-1}^* \dots a_1^*]^T$.

3.2. Broad Coverage Precoder Design

If we optimize \mathbf{P} directly, we need to search over an rM_zM_x dimensional space. For massive MIMO, rM_zM_x is usually very large and is still increasing as wireless communications systems evolve, thus the complexity of the optimization will be very high. To reduce the computational complexity, we need to find certain special structures that can be employed to transform the optimization problem with high dimensions into one with low dimensions. In [13], an omnidirectional precoder design for 3D massive MIMO with UPAs is proposed. Among the many theorems, Theorem 1 in [13] is an easy, but important, one. It will also play an important role in helping us to transform the optimization problem with high dimensions into one with low dimensions. Thus, we present it in Appendix A as Lemma A1.

With Lemma A1, we can prove the following theorem, which can be used to generate the precoding matrices having the first property in Equation (15) required by broad coverage precoding.

Theorem 1. Let the two sets \mathcal{D} and \mathcal{F} of $M_z \times 1$ vectors be defined as $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r\}$ and $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_r\}$ with

$$c_{\mathcal{D}, \mathcal{F}}(t) = 0, t = -M_z + 1, \dots, M_z - 1. \quad (21)$$

Furthermore, the sums of the angle power spectrum of the vectors in these two sets are equal, and are given as

$$\sum_{i=1}^r s_{\mathbf{d}_i}(u) = \sum_{i=1}^r s_{\mathbf{f}_i}(u) = a_z(u). \quad (22)$$

Let $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ denote two $M_x \times 1$ vectors with the sum of their angle power spectrum being given by

$$s_{\boldsymbol{\phi}}(v) + s_{\boldsymbol{\varphi}}(v) = a_x(v). \quad (23)$$

The precoding matrices \mathbf{P}_i are defined as

$$\mathbf{P}_i = \mathbf{d}_i \otimes \boldsymbol{\phi}^T + \mathbf{f}_i \otimes \boldsymbol{\varphi}^T, i = 1, 2, \dots, r. \quad (24)$$

Then, the sum of the 2D angle power spectrum of these precoding matrices is

$$\sum_{i=1}^r s_{\mathbf{P}_i}(u, v) = a_z(u)a_x(v). \quad (25)$$

Proof. The proof is provided in Appendix B. \square

Theorem 1 provides a method to generate high-dimensional precoding matrices having the first condition required by broad coverage precoding from low-dimensional vectors. When \mathcal{D} and \mathcal{F} are both sets of orthonormal vectors, it can be easily proved that the third condition required by broad coverage precoding is also satisfied. Thus, we only need to let the precoding matrices satisfy the second condition. We then obtain a sufficient condition for the precoding matrices to satisfy the second condition in the following theorem.

Theorem 2. Let \mathcal{D} and \mathcal{F} be two sets of $M_z \times 1$ vectors defined as $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r\}$ and $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_r\}$ with

$$\sum_{i=1}^r |d_{ik}|^2 = 1, k = 1, 2, \dots, M_z \tag{26}$$

$$\sum_{i=1}^r |f_{ik}|^2 = 1, k = 1, 2, \dots, M_z, \tag{27}$$

where d_{ik} and f_{ik} denote the k -th elements of \mathbf{d}_i and \mathbf{f}_i , respectively. Let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ be two vectors with binary entries and $\alpha_j + \beta_j = 1, j = 1, 2, \dots, M_x$. Let $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ be two $M_x \times 1$ vectors defined by $[\boldsymbol{\phi}]_j = e^{-j2\pi w_j} \alpha_j$ and $[\boldsymbol{\varphi}]_j = e^{-j2\pi w_j} \beta_j$. The precoding matrices \mathbf{P}_i are defined as

$$\mathbf{P}_i = \mathbf{d}_i \otimes \boldsymbol{\phi}^T + \mathbf{f}_i \otimes \boldsymbol{\varphi}^T, i = 1, 2, \dots, r. \tag{28}$$

Then, we have

$$\sum_{i=1}^r |p_{i,kl}|^2 = 1, \forall k, l. \tag{29}$$

Proof. The proof is provided in Appendix C. \square

Theorem 2 has been presented as part of Theorem 3 in [13]. However, putting it in an independent theorem makes it easier to understand what is the condition for the per-antenna constant power constraint. We define $\mathbf{D} = [\mathbf{d}_1 \ \mathbf{d}_2 \ \dots \ \mathbf{d}_r]$ and $\mathbf{F} = [\mathbf{f}_1 \ \mathbf{f}_2 \ \dots \ \mathbf{f}_r]$. The conditions in Equations (26) and (27) are equivalent to that the matrices \mathbf{D} and \mathbf{F} should satisfy the constraints

$$\text{diag}(\mathbf{D}\mathbf{D}^H) = \mathbf{I}_{M_z} \tag{30}$$

and

$$\text{diag}(\mathbf{F}\mathbf{F}^H) = \mathbf{I}_{M_z}. \tag{31}$$

The conditions for $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ in Theorem 2 can be rephrased as follows. The nonzero elements in $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ have a constant envelope, and their locations in $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ are complementary.

Combining Theorems 1 and 2, we can obtain a broad coverage precoder design, which is a generalized version of Theorem 3 in [13] for omnidirectional precoder design. When $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ are a pair of complementary sequences, and \mathcal{D} and \mathcal{F} are two sets of orthonormal complementary vectors, the combination of Theorems 1 and 2 reduces to Theorem 3 in [13]. From Theorems 1 and 2, we see that the broad coverage precoding matrices can be obtained from the pair $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ and the two sets \mathcal{D} and \mathcal{F} . Thus, instead of optimizing the high-dimensional precoding matrices directly, we can optimize the low-dimensional $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$, and the low-dimensional $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r$ and $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_r$, respectively.

In the following, we first investigate the design of \mathcal{D} and \mathcal{F} having zero aperiodic cross-correlation. Theorem 11 in [7] presents a method to construct two sets having zero cross-correlation when r is even. The method is used to generate sets of complementary vectors with bipolar elements, but can be extended to design sets of vectors with arbitrary elements easily, as shown in the following theorem.

Theorem 3. Let $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r\}$ be a set of vectors, where r is even. Let $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_r\}$ be a set of vectors defined by $\mathbf{f}_1 = \mathbf{d}_2^\dagger, \mathbf{f}_2 = -\mathbf{d}_1^\dagger, \mathbf{f}_3 = \mathbf{d}_4^\dagger, \mathbf{f}_4 = -\mathbf{d}_3^\dagger, \dots, \mathbf{f}_r = -\mathbf{d}_{r-1}^\dagger$. Then, we have

$$c_{\mathcal{D},\mathcal{F}}(t) = 0, t = -M_z + 1, \dots, M_z - 1. \tag{32}$$

Proof. The proof is similar to that in [7] and is omitted here for brevity. \square

Theorem 3 presents a method to obtain the set \mathcal{F} whose aperiodic cross-correlation with the set \mathcal{D} is zero when the latter is already given. We can also observe that the sums of

the angle power spectrum of these two sets are the same, and that the vectors in \mathcal{F} are also orthonormal vectors if the vectors in \mathcal{D} are orthonormal vectors. Thus, instead of designing two sets, we only need to design one set of orthonormal vectors with the required angle power spectrum.

3.3. Design of Set \mathcal{D}

In this subsection, we introduce the design of the set \mathcal{D} for broad coverage precoder design. Since $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r\}$ is a set of orthonormal column vectors, we have

$$\mathbf{D}^H \mathbf{D} = \frac{M_z}{r} \mathbf{I}_r. \quad (33)$$

From Equation (22), we then see that \mathbf{D} should also satisfy

$$\mathbf{v}_z(u)^T \mathbf{D} \mathbf{D}^H \mathbf{v}_z(u)^* = a_z(u). \quad (34)$$

To find the matrix \mathbf{D} that satisfies the above equation, we formulate the problem as

$$\begin{aligned} & \text{find } \mathbf{D} \\ & \text{s.t. } \mathbf{v}_z(u)^T \mathbf{D} \mathbf{D}^H \mathbf{v}_z(u)^* = a_z(u). \\ & \quad \text{diag}(\mathbf{D} \mathbf{D}^H) = \mathbf{I}_{M_z} \\ & \quad \mathbf{D}^H \mathbf{D} = \frac{M_z}{r} \mathbf{I}_r. \end{aligned} \quad (35)$$

It is very hard to find \mathbf{D} that satisfies Equation (35) directly because

$$\mathbf{v}_z(u)^T \mathbf{D} \mathbf{D}^H \mathbf{v}_z(u)^* = a_z(u)$$

might have no solution. To make the optimization problem valid, we sample the function $a_z(u)$ at the points

$$u_k = \frac{k}{N_z M_z}, k = 1, 2, \dots, N_z M_z,$$

where N_z is the sampling ratio. Let $a_{z,k} = a_z(u_k)$. Let \mathbf{V}_z^F be an $N_z M_z \times N_z M_z$ DFT matrix, and \mathbf{V}_z be an $N_z M_z \times M_z$ matrix defined by

$$\mathbf{V}_z = \mathbf{V}_z^F \begin{pmatrix} \mathbf{I}_{M_z} \\ \mathbf{0}_{(N_z-1)M_z} \end{pmatrix}. \quad (36)$$

We then try to find an approximate solution of

$$\mathbf{e}_k^T \mathbf{V}_z \mathbf{D} \mathbf{D}^H \mathbf{V}_z^H \mathbf{e}_k = a_{z,k}, k = 1, 2, \dots, N_z M_z$$

under the other constraints. To achieve this, we need to define a distance or divergence between the sequences $\mathbf{e}_k^T \mathbf{V}_z \mathbf{D} \mathbf{D}^H \mathbf{V}_z^H \mathbf{e}_k$ and $a_{z,k}$. Let $\hat{a}_{z,k}$ be defined as

$$\hat{a}_{z,k} = \mathbf{e}_k^T \mathbf{V}_z \mathbf{D} \mathbf{D}^H \mathbf{V}_z^H \mathbf{e}_k, k = 1, 2, \dots, N_z M_z. \quad (37)$$

Let \mathbf{a}_z and $\hat{\mathbf{a}}_z$ denote the sequences $\{a_{z,k}\}$ and $\{\hat{a}_{z,k}\}$, respectively. The KL divergence between the sequences \mathbf{a}_z and $\hat{\mathbf{a}}_z$ is defined as [21]

$$\begin{aligned} & D_{KL}(\mathbf{a}_z : \hat{\mathbf{a}}_z) \\ & = \sum_{k=1}^{N_z M_z} a_{z,k} \log \frac{a_{z,k}}{\hat{a}_{z,k}} - \sum_{k=1}^{N_z M_z} a_{z,k} + \sum_{k=1}^{N_z M_z} \hat{a}_{z,k}. \end{aligned} \quad (38)$$

From the definition of $\hat{a}_{z,k}$ and $\mathbf{V}_z^H \mathbf{V}_z = \mathbf{I}_{M_z}$, we have

$$\begin{aligned} \sum_{k=1}^{N_z M_z} \hat{a}_{z,k} &= \text{tr}(\mathbf{V}_z \mathbf{D} \mathbf{D}^H \mathbf{V}_z^H) \\ &= \text{tr}(\mathbf{D} \mathbf{D}^H \mathbf{V}_z^H \mathbf{V}_z) \\ &= \text{tr}(\mathbf{D} \mathbf{D}^H) \\ &= M_z. \end{aligned} \quad (39)$$

Then, the KL divergence can be rewritten as

$$D_{KL}(\mathbf{a}_z : \hat{\mathbf{a}}_z) = \sum_{k=1}^{N_z M_z} -a_{z,k} \log \hat{a}_{z,k} + c_k, \quad (40)$$

where

$$c_k = \sum_{k=1}^{N_z M_z} a_{z,k} \log a_{z,k} - \sum_{k=1}^{N_z M_z} a_{z,k} + M_z \quad (41)$$

is a constant not related to \mathbf{D} .

By using the KL divergence, we define $f(\mathbf{D}) = D_{KL}(\mathbf{a}_z : \hat{\mathbf{a}}_z)$. We then formulate an optimization problem approximate to that in Equation (35) as

$$\begin{aligned} \min_{\mathbf{D}} f(\mathbf{D}) \\ \text{s.t. } \text{diag}(\mathbf{D} \mathbf{D}^H) &= \mathbf{I}_{M_z} \\ \mathbf{D}^H \mathbf{D} &= \frac{M_z}{r} \mathbf{I}_r. \end{aligned} \quad (42)$$

The constrained optimization problem (42) is very hard to handle since its constraints are not as simple as linear or quadratic constraints. For such problems, manifold optimization is a powerful alternative [22]. It has already been widely used in MIMO systems to handle various complicated constraints. The most commonly addressed manifolds in the MIMO literature are Grassmann and Stiefel manifolds [23–26]. For the optimization problem (42), the two constraints are the Stiefel manifold

$$\mathcal{U} = \{\mathbf{D} \in \mathbb{C}^{M_z \times r} \mid \mathbf{D}^H \mathbf{D} = \frac{M_z}{r} \mathbf{I}_r\} \quad (43)$$

and the oblique manifold [27]

$$\mathcal{O} = \{\mathbf{D} \in \mathbb{C}^{M_z \times r} \mid \text{diag}(\mathbf{D} \mathbf{D}^H) = \mathbf{I}_{M_z}\}. \quad (44)$$

The optimization problem is then reformulated as

$$\begin{aligned} \arg \min_{\mathbf{D}} f(\mathbf{D}) \\ \text{s.t. } \mathbf{D} \in \mathcal{O} \cap \mathcal{U}. \end{aligned} \quad (45)$$

The obtained optimization problem is over the intersection of two manifolds. The problem (45) can also be used to design the broad coverage precoder for ULAs, which has been investigated in [17] with a different objective function based on manifold optimization. It has also been investigated in [13] with $f(\mathbf{D})$ being changed to the objective function for omnidirectional precoding. Thus, it can also be solved by using the proposed methods in [13,17]. Since the projected gradient method on the intersection of two manifolds provided in [13] is simpler, we also use it in this work.

In the following, we briefly introduce the projected gradient method. The gradient of the objective function f in the Euclidean space is obtained as

$$\nabla f(\mathbf{D}) = \frac{\partial f(\mathbf{D})}{\partial \mathbf{D}^*} = \sum_{k=1}^{N_z M_z} -a_{z,k} \hat{a}_{z,k}^{-1} \mathbf{V}^H \mathbf{e}_k \mathbf{e}_k^T \mathbf{V} \mathbf{D}. \tag{46}$$

With the obtained Euclidean gradient, we can search along the gradient direction $\nabla f(\mathbf{D})$ to find the optimal \mathbf{D} . Let μ be the step size. After moving a step ahead, the obtained point $\mathbf{D} - \mu \nabla f(\mathbf{D})$ is out of $\mathcal{O} \cap \mathcal{U}$. We need to project it back to $\mathcal{O} \cap \mathcal{U}$. To achieve this, we first define the projections onto the manifolds \mathcal{O} and \mathcal{U} as

$$P_{\mathcal{O}}(\mathbf{D}) = (\text{diag}(\mathbf{D}\mathbf{D}^H))^{-1/2} \mathbf{D} \tag{47}$$

and

$$P_{\mathcal{U}}(\mathbf{D}) = \sqrt{\frac{M_z}{r}} \mathbf{D} (\mathbf{D}^H \mathbf{D})^{-1/2}. \tag{48}$$

From [28], the projection onto $\mathcal{O} \cap \mathcal{U}$ is obtained as

$$P_{\mathcal{O} \cap \mathcal{U}}(\mathbf{D}) = P_{\mathcal{O}}(P_{\mathcal{U}}(\dots P_{\mathcal{O}}(P_{\mathcal{U}}(\mathbf{D}))))), \tag{49}$$

where \dots indicates that an infinite number of projections are needed theoretically. In practical use, the alternative projections proceed until reaching a pre-set target. By using the projection, the point $\mathbf{D} - \mu \nabla f(\mathbf{D})$ can be projected back to the intersection of two manifolds $\mathcal{O} \cap \mathcal{U}$ by $P_{\mathcal{O} \cap \mathcal{U}}(\mathbf{D} - \mu \nabla f(\mathbf{D}))$.

The projected gradient method is summarized as

$$\mathbf{D}^{d+1} = P_{\mathcal{O} \cap \mathcal{U}}(\mathbf{D}^d - \mu \nabla f(\mathbf{D}^d)), \tag{50}$$

where \mathbf{D}^d is the point obtained after $d - 1$ iterations. To guarantee the convergence, μ should have a very small value or needs to be determined by the line search method [22]. By using the projected gradient method, an optimal point of the problem (45) is obtained.

3.4. Design of the Pair

In this subsection, we investigate the design of $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$. To meet the condition provided in Theorem 2, the pair $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ should satisfy

$$s_{\boldsymbol{\phi}}(v) + s_{\boldsymbol{\varphi}}(v) = a_x(v) \tag{51}$$

and their elements should satisfy $\phi_j = e^{-j2\pi w_j} \alpha_j$ and $\varphi_j = e^{-j2\pi w_j} \beta_j$, where α_j and β_j are binary and $\alpha_j + \beta_j = 1$. For simplicity, we consider the case when M_x is even and fix the structures of $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ in this paper. We redefine $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ as

$$\boldsymbol{\phi} = [e^{-j2\pi\rho_1} \ e^{-j2\pi\rho_2} \ \dots \ e^{-j2\pi\rho_{M_x/2}} \ 0 \ \dots \ 0]^T \tag{52}$$

and

$$\boldsymbol{\varphi} = [0 \ \dots \ 0 \ e^{-j2\pi\lambda_1} \ e^{-j2\pi\lambda_2} \ \dots \ e^{-j2\pi\lambda_{M_x/2}}]^T. \tag{53}$$

It can be observed that the condition needed for the vectors $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ to ensure the per-antenna constant power constraint is still satisfied. Instead of optimizing $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ directly, we will optimize ρ_i and λ_i , $i = 1, 2, \dots, M_x/2$ since they have no constraint.

It is also very hard to obtain ρ_i and λ_i that satisfy Equation (51) because it might have no solution. To obtain a target than can make the optimization easier, we sample the $a_x(v)$ in the points

$$v_k = \frac{k}{N_x M_x}, k = 1, 2, \dots, N_x M_x$$

and try to find an approximate solution of

$$s_{\phi}(v_k) + s_{\varphi}(v_k) = a_{x,k}, k = 1, 2, \dots, N_x M_x, \tag{54}$$

where $a_{x,k} = a_x(v_k)$. Thus, we utilize the KL divergence again in the following. Let N_x be the sampling ratio. Let \mathbf{V}_x^F be an $N_x M_x \times N_x M_x$ DFT matrix and \mathbf{V}_x be an $N_x M_x \times M_x$ matrix defined by

$$\mathbf{V}_x = \mathbf{V}_x^F \begin{pmatrix} \mathbf{I}_{M_x} \\ \mathbf{0}_{(N_x-1)M_x} \end{pmatrix}. \tag{55}$$

We define $\hat{a}_{x,k}$ as

$$\begin{aligned} \hat{a}_{x,k} &= s_{\phi}(v_k) + s_{\varphi}(v_k) \\ &= \mathbf{e}_k^T \mathbf{V}_x \boldsymbol{\phi} \boldsymbol{\phi}^H \mathbf{V}_x^H \mathbf{e}_k + \mathbf{e}_k^T \mathbf{V}_x \boldsymbol{\varphi} \boldsymbol{\varphi}^H \mathbf{V}_x^H \mathbf{e}_k. \end{aligned} \tag{56}$$

The optimization problem can be formulated as

$$\min_{\{\rho_i, \lambda_i\}} \sum_{k=1}^{N_x M_x} a_{x,k} \log \frac{a_{x,k}}{\hat{a}_{x,k}} - \sum_{k=1}^{N_x M_x} a_{x,k} + \sum_{k=1}^{N_x M_x} \hat{a}_{x,k}. \tag{57}$$

It can be proved that $\sum_{k=1}^{N_x M_x} \hat{a}_{x,k} = M_x$. Thus, the problem can be rewritten as

$$\min_{\{\rho_i, \lambda_i\}} \sum_{k=1}^{N_x M_x} -a_{x,k} \log \hat{a}_{x,k} + d_k, \tag{58}$$

where

$$d_k = \sum_{k=1}^{N_x M_x} a_{x,k} \log a_{x,k} - \sum_{k=1}^{N_x M_x} a_{x,k} + M_x \tag{59}$$

is a constant not related to $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ or ρ_i and λ_i . The above optimization problem is an unconstrained optimization problem, and thus a solution can be obtained by using the gradient method. Let $g(\boldsymbol{\rho}, \boldsymbol{\lambda})$ denote the objective function in Equation (58) as

$$g(\boldsymbol{\rho}, \boldsymbol{\lambda}) = \sum_{k=1}^{N_x M_x} -a_{x,k} \log \hat{a}_{x,k} + d_k. \tag{60}$$

In the following theorem, we present the gradients of g with respect to ρ_i and λ_i , respectively.

Theorem 4. *The gradients of the objective function g with respect to ρ_i and λ_i are given by*

$$\frac{\partial g}{\partial \rho_i} = \sum_{k=1}^{N_x M_x} -a_{x,k} \hat{a}_{x,k}^{-1} \Re(j2\pi \boldsymbol{\phi}^H \mathbf{e}_i \mathbf{e}_i^T \mathbf{V}_x^H \mathbf{e}_k \mathbf{e}_k^T \mathbf{V}_x \boldsymbol{\phi}) \tag{61}$$

$$\frac{\partial g}{\partial \lambda_i} = \sum_{k=1}^{N_x M_x} -a_{x,k} \hat{a}_{x,k}^{-1} \Re(j2\pi \boldsymbol{\varphi}^H \mathbf{e}_i \mathbf{e}_i^T \mathbf{V}_x^H \mathbf{e}_k \mathbf{e}_k^T \mathbf{V}_x \boldsymbol{\varphi}). \tag{62}$$

Proof. The proof is provided in Appendix D. \square

With the gradient provided in Theorem 4, the gradient method can be performed and summarized as

$$\rho_i^{d+1} = \rho_i^d + \mu \frac{\partial g}{\partial \rho_i} \tag{63}$$

$$\lambda_i^{d+1} = \lambda_i^d + \mu \frac{\partial g}{\partial \lambda_i}, \tag{64}$$

where μ is the step size determined by the line search method to guarantee the convergence.

The steps of the proposed broad coverage precoder design are now summarized as follows:

1. Solve the problem (45) by using the projected gradient method provided in Equation (50) to obtain \mathbf{D} . Obtain the vectors $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r$ as the column of \mathbf{D} .
2. Compute $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_r$ as $\mathbf{f}_1 = \mathbf{d}_2^\dagger, \mathbf{f}_2 = -\mathbf{d}_1^\dagger, \mathbf{f}_3 = \mathbf{d}_4^\dagger, \mathbf{f}_4 = -\mathbf{d}_3^\dagger, \dots, \mathbf{f}_r = -\mathbf{d}_{r-1}^\dagger$.
3. Solve Equation (58) by using the gradient method provided in Equations (63) and (64) to obtain ρ_i and λ_i . Generate $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ according to Equations (52) and (53).
4. Generate the precoding matrices as $\mathbf{P}_i = \mathbf{d}_i \otimes \boldsymbol{\phi}^T + \mathbf{f}_i \otimes \boldsymbol{\varphi}^T, i = 1, 2, \dots, r$.

4. Simulation Results

In this section, we provide simulation results to verify the analytic results of the proposed broad coverage precoder design. For simplicity, the target angle power spectra $a_z(u)$ and $a_x(v)$ are defined according to raised-cosine filters as

$$a_z(u) = \begin{cases} 1, & |u| \leq \alpha_z - \beta \\ 0, & |u| \geq \alpha_z + \beta \\ \frac{1}{2}(1 + \cos(\frac{\pi}{\beta}(|u| - (\alpha_z - \beta))))), & \text{otherwise} \end{cases} \quad (65)$$

and

$$a_x(v) = \begin{cases} 1, & |v| \leq \alpha_x - \beta \\ 0, & |v| \geq \alpha_x + \beta \\ \frac{1}{2}(1 + \cos(\frac{\pi}{\beta}(|v| - (\alpha_x - \beta))))), & \text{otherwise,} \end{cases} \quad (66)$$

where α_x and α_z are values related to horizontal and vertical coverage, and β is related to the roll-off factor. We set the number of the data streams r to 2 in all simulations.

In the first simulation, we set $M_x = M_z = 32, \alpha_x = 1/3$ and $\alpha_z = 1/3$. We first obtain two sets of orthonormal sequences $\{\mathbf{d}_i\}, \{\mathbf{f}_i\}$ with zero cross-correlation and an angle power spectrum close to $a_z(u)$ by using the method provided in Section 3.3. We then use the gradient method provided in Section 3.4 to generate the pair $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$ with the sum of their angle power spectra being close to $a_x(v)$. Furthermore, we compute the broad coverage precoding matrices as $\mathbf{P}_i = \mathbf{d}_i \otimes \boldsymbol{\phi}^T + \mathbf{f}_i \otimes \boldsymbol{\varphi}^T, i = 1, 2$. The angle power spectra of the two precoding matrices are plotted in Figure 1a,b, and their sum is plotted in Figure 1c. From the figure, we observe that the shape of the generated power pattern of one stream has large perturbation, and whereas the sum of the generated power pattern of two streams has little perturbation. The sum is very close to that of the target power pattern $a_z(u)a_x(v)$. By using two streams, we are able to obtain a broad coverage precoding matrix whose angle power spectrum is very close to that of the target one.

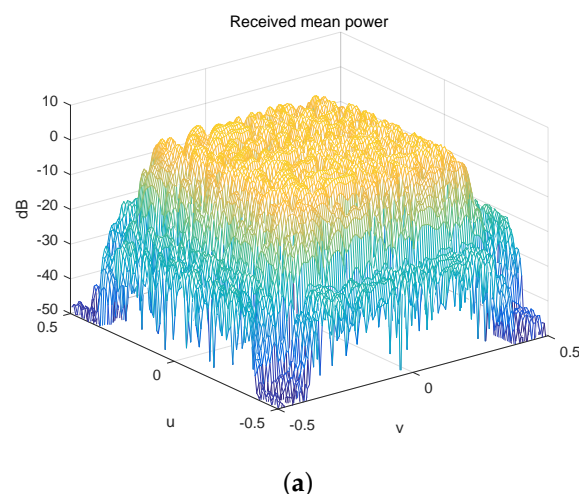


Figure 1. Cont.

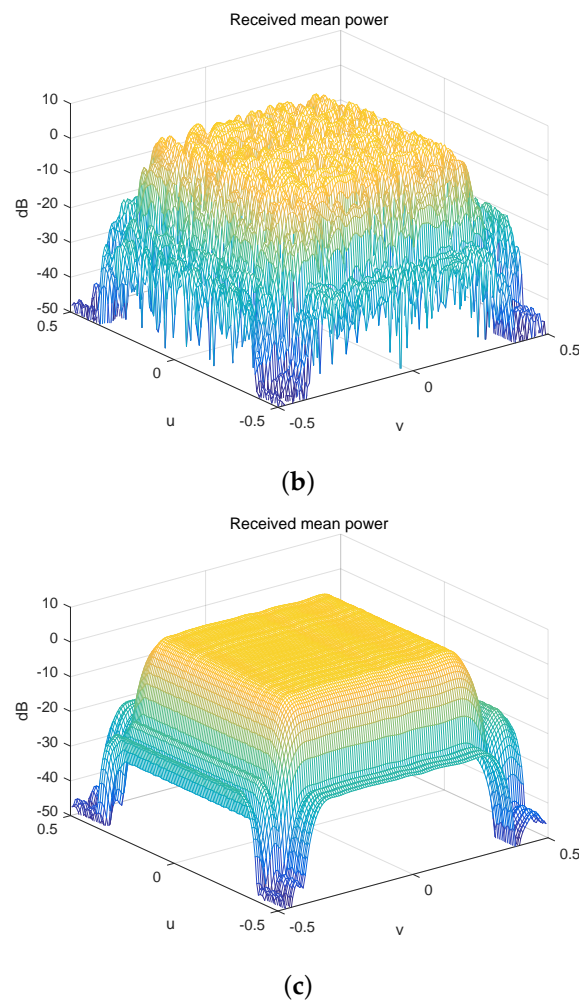


Figure 1. The anglepower spectrum of the proposed broad coverage precoder design. (a) The angle power spectrum of the 32×32 precoder \mathbf{P}_1 . (b) The angle power spectrum of the 32×32 precoder \mathbf{P}_2 . (c) The sum of the angle power spectrum of the 32×32 precoders \mathbf{P}_1 and \mathbf{P}_2 .

We then investigate the performance of the proposed broad coverage precoder in a 3D massive MIMO system. The simulation parameters are in Table 1. The base station of the considered system is equipped with a UPA with $M_z = 64$, $M_x = 32$ and is located at $(0, 0, 25)$. The locations of the users are generated randomly in a 120 degree sector with radius $R = 200$ m around the origin $(0, 0, 0)$ at 1.5 m height. The downtilt degree is set as $\alpha = \arctan(2h/R)$, where $h = 23.5$ m. For convenience, we define the virtual horizontal angle as φ such that $\cos \varphi = \cos \phi \sin \theta$. To cover the 120 degree sector, the value of α_x is set as $\alpha_x = \sqrt{3}/4$, whereas the value of α_z needs to be chosen according to the synchronization performance. In the simulations, we use $a_z = 0.05, 0.1$ and 0.2 , the radiation power patterns of the proposed precoders are shown in Figure 2b–d. The range of the elevation angles becomes larger as the value of a_z increases. Furthermore, all results show that the proposed method can generate broad coverage precoder having the radiation power pattern close to the target radiation power pattern. For comparison, the method from [18] is also used to generate the broad coverage precoder according to $a_z(u)a_x(v)$ with $\alpha_x = 0.1$. The radiation power pattern of this case is plotted in Figure 2a. From the comparison of Figure 2a,c, it can be observed that the radiation power patterns of the two methods are nearly the same, but the proposed precoder design has lower leakage outside the desired coverage area than the method from [18]. This means that the proposed method causes less interference than the latter. The reason is that the dimensions of the optimization problems in the proposed precoder design are smaller, and thus it is easier to obtain good results.

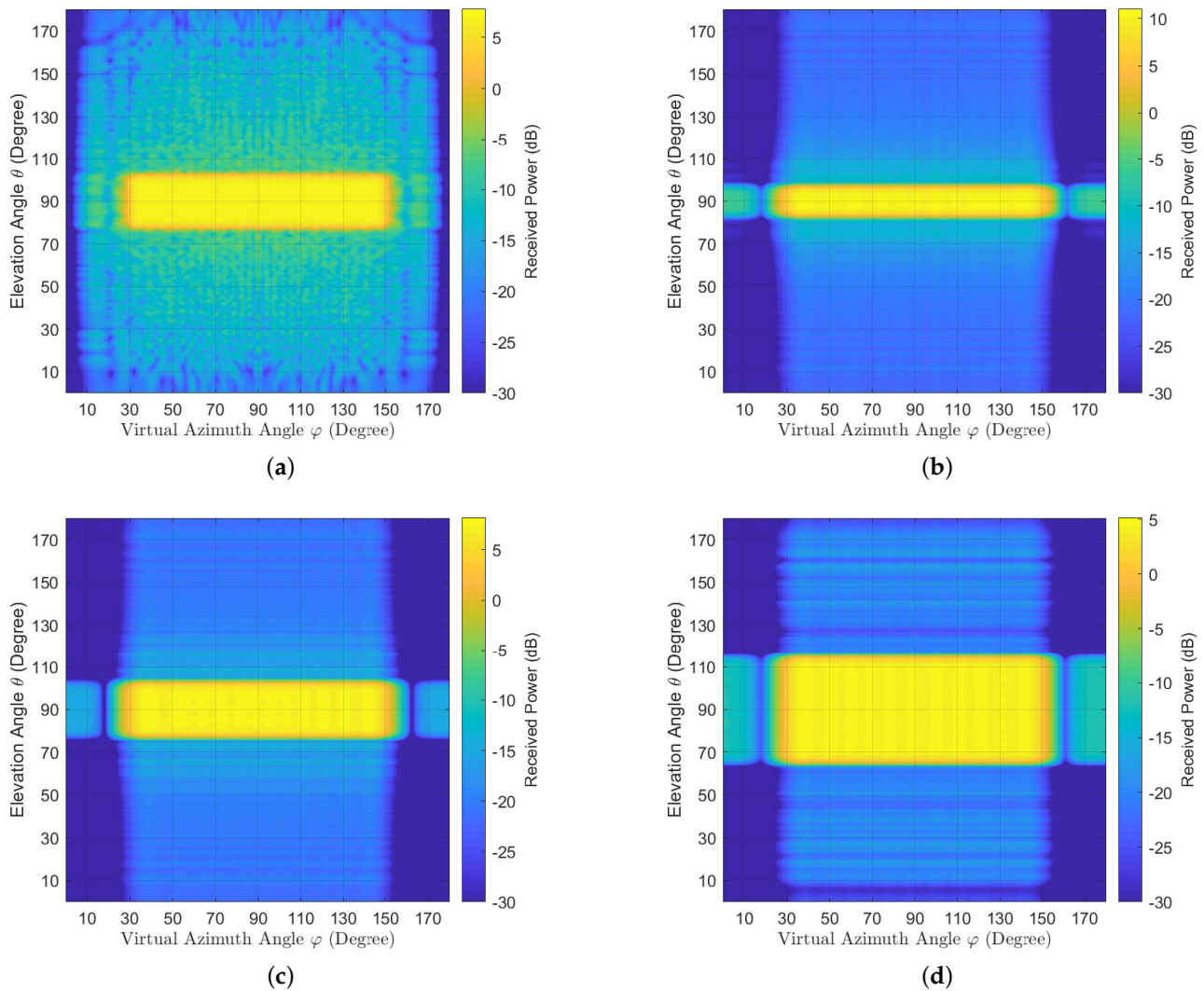


Figure 2. The radiation power pattern generated by the precoders when $r = 2$ and $M_z = 64, M_x = 32$. (a) The power pattern generated by the method from [18] with $\alpha_z = 0.1$. (b) The power pattern generated by the proposed method with $\alpha_z = 0.05$. (c) The power pattern generated by the proposed method with $\alpha_z = 0.1$. (d) The power pattern generated by the proposed method with $\alpha_z = 0.2$.

Table 1. Simulation parameters for 3D massive MIMO.

Parameter	Value
Antenna configuration	$M_z = 64, M_x = 32$
BS location	(0, 0, 25)
Radius	200 m
Downtilt degree	$\arctan(2h/R)$
Height of users	1.5 m

Furthermore, to show the coverage of the proposed precoders in the 120-degree sector, the received mean powers of the proposed precoders at height 1.5 m are shown in Figure 3b–d. They show that the proposed precoder with $\alpha_z = 0.05$ has more evenly received mean power than the proposed precoders in the other two cases. However, a certain area that is far from the BS is not covered when $\alpha_z = 0.05$. Meanwhile, the received mean power of the proposed precoder with $\alpha_z = 0.2$ focuses more on the areas near the BS than those areas at the cell edge. Thus, the proposed precoder with $\alpha_z = 0.1$ provides better coverage than the proposed precoders from the other two cases. From the results of

the received mean power, it seems the the proposed precoder with $\alpha_z = 0.1$ is preferred than the other two precoders. The received mean power of the method from [18] with $\alpha_x = 0.1$ is plotted in Figure 3a. From the comparison of Figure 3a,c, we also observe that the proposed method has lower leakage outside the desired coverage area than the method from [18], and the two methods have nearly the same coverage.

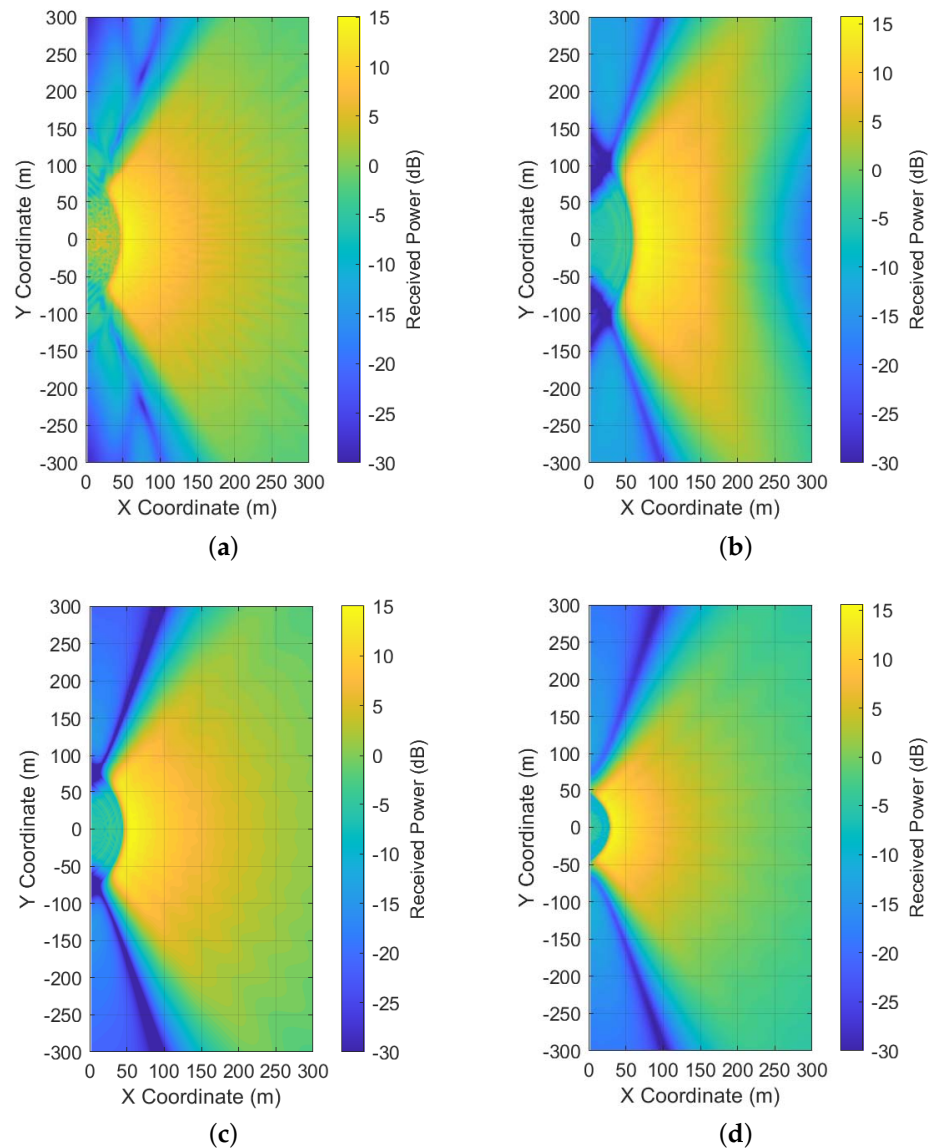


Figure 3. The received mean power in a 120-degree sector generated by the precoders when $r = 2$ and $M_z = 64, M_x = 32$. (a) The power pattern generated by the method from [18] with $\alpha_z = 0.1$. (b) The power pattern generated by the proposed method with $\alpha_z = 0.05$. (c) The power pattern generated by the proposed method with $\alpha_z = 0.1$. (d) The power pattern generated by the proposed method with $\alpha_z = 0.2$.

Finally, we investigate the synchronization performance of the proposed precoders in the 120-degree sector. For simplicity, we only consider the line of sight scenario. The users are randomly generated inside the sector, and then their channel coefficients are generated according to their location. The missed detection (MD) probability is one of the most important metrics that can characterize the synchronization performance. Thus, we use the MD probability to verify the synchronization performance of the proposed precoders. The MD probabilities of the proposed precoders are shown in Figure 4. The results of the omnidirectional precoder from [13] and the method from [18] are also pro-

vided as benchmarks. It shows that the proposed broad coverage precoder outperforms the omnidirectional precoder. This is because the omnidirectional precoder wastes part of its power by sending signals outside the sector. Among the three broad coverage precoders with different α_z s, the proposed precoder with $\alpha_z = 0.1$ has the best MD probability performance. This indicates that we can choose $\alpha_z = 0.1$ for the considered 3D massive MIMO system to ensure a better synchronization performance. Meanwhile, the differences between the MD probabilities of the method from [18] and the proposed precoder are negligible. Thus, the proposed precoder can achieve nearly the same performance as the broad coverage precoder from [18] with much lower computational complexity.

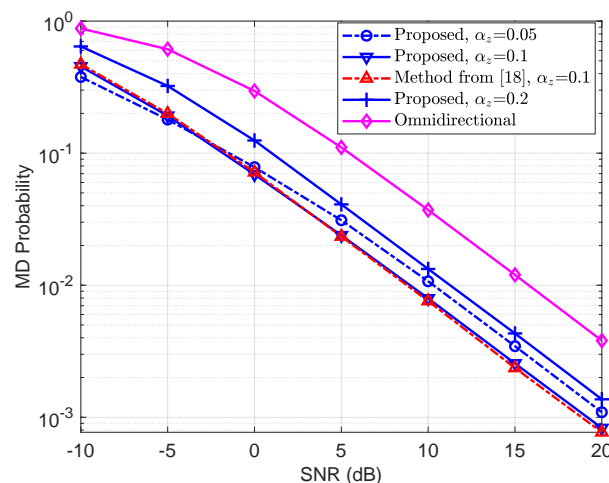


Figure 4. The MD probabilities of the proposed precoder for the line of sight scenario.

5. Conclusions

In this paper, we proposed a broad coverage precoder design for massive MIMO with huge UPAs. The desired 2D angle power spectrum is assumed to be separable, and the per-antenna constant power constraint and semi-unitary constraint were considered. We proposed a method to construct the high-dimensional broad coverage precoding matrices from a set of orthonormal vectors and two vectors with certain properties. We also presented optimization methods to generate the set of orthonormal vectors and the pair of vectors, respectively. Finally, simulation results showed that the proposed broad coverage precoding matrices achieve nearly the same performance as the normal broad coverage precoder with much lower computational complexity.

The proposed broad coverage precoder is designed by assuming the required 2D angle power spectrum is separable. In practical use, the target 2D angle power spectrum might be inseparable. In this case, we might need to consider the problem of approximating the inseparable power spectrum with the separable power spectrum if we still want to use the proposed precoder design. We can also try to find a lower-complexity precoder design for the inseparable 2D angle power spectrum. However, it might be very hard to obtain a solution. Meanwhile, the proposed broad coverage precoder is for massive MIMO with UPAs. The design for other more advanced antenna configurations is also an interesting problem.

Author Contributions: Conceptualization, A.-A.L. and X.G.; methodology, A.-A.L.; software, Y.C.; writing—original draft preparation, A.-A.L.; writing—review and editing, A.-A.L.; supervision, X.G.; 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 Grants 61801113, the Natural Science Foundation of Jiangsu Province under Grant BK20180362.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article.

Acknowledgments: The authors would like to thank the editor and the anonymous reviewers for their helpful comments and suggestions.

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

Appendix A. A Lemma

Lemma A1 (Theorem 1 in [13]). Let $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r\}$ and $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_r\}$ be two sets consisting of $M_z \times 1$ vectors. Let $\mathbf{D} = [\mathbf{d}_1 \ \mathbf{d}_2 \ \dots \ \mathbf{d}_r] \in \mathbb{C}^{M_z \times r}$ and $\mathbf{F} = [\mathbf{f}_1 \ \mathbf{f}_2 \ \dots \ \mathbf{f}_r] \in \mathbb{C}^{M_z \times r}$. We define the vector-valued functions $\tilde{\mathbf{d}}(u)$ and $\tilde{\mathbf{f}}(u)$ as

$$\tilde{\mathbf{d}}(u)^T = \mathbf{v}_z(u)^T \mathbf{D} \tag{A1}$$

$$\tilde{\mathbf{f}}(u)^T = \mathbf{v}_z(u)^T \mathbf{F}. \tag{A2}$$

If the aperiodic cross-correlation of \mathcal{D} and \mathcal{F} is zero, i.e.,

$$c_{\mathcal{D}, \mathcal{F}}(t) = 0, t = -M_z + 1, \dots, M_z - 1 \tag{A3}$$

then $\tilde{\mathbf{d}}(u)$ and $\tilde{\mathbf{f}}(u)$ are orthogonal for any u , i.e.,

$$\tilde{\mathbf{d}}(u)^H \tilde{\mathbf{f}}(u) = 0, \forall u. \tag{A4}$$

Appendix B. Proof of Theorem 1

The vector-valued function $\mathbf{q}_i(u)$ is defined as

$$\begin{aligned} \mathbf{q}_i(u)^T &= \mathbf{v}_z(u)^T \mathbf{P}_i \\ &= \mathbf{v}_z(u)^T (\mathbf{d}_i \otimes \boldsymbol{\phi}^T) + \mathbf{v}_z(u)^T (\mathbf{f}_i \otimes \boldsymbol{\phi}^T). \end{aligned} \tag{A5}$$

Let $\mathbf{Q}(u) = [\mathbf{q}_1(u) \ \mathbf{q}_2(u) \ \dots \ \mathbf{q}_r(u)]^T$, and we obtain

$$\sum_{i=1}^r |\mathbf{v}_z(u)^T \mathbf{P}_i \mathbf{v}_x(v)|^2 = \mathbf{v}_x(v)^H \mathbf{Q}(u)^H \mathbf{Q}(u) \mathbf{v}_x(v). \tag{A6}$$

We define $\tilde{\mathbf{d}}(u)$ and $\tilde{\mathbf{f}}(u)$ as

$$\begin{aligned} \tilde{\mathbf{d}}(u) &= [\tilde{d}_1(u) \ \tilde{d}_2(u) \ \dots \ \tilde{d}_r(u)]^T \\ \tilde{\mathbf{f}}(u) &= [\tilde{f}_1(u) \ \tilde{f}_2(u) \ \dots \ \tilde{f}_r(u)]^T, \end{aligned}$$

where $\tilde{d}_i(u) = \mathbf{v}_z(u)^T \mathbf{d}_i$ and $\tilde{f}_i(u) = \mathbf{v}_z(u)^T \mathbf{f}_i$. We then have $\mathbf{Q}(u) = \tilde{\mathbf{d}}(u) \otimes \boldsymbol{\phi}^T + \tilde{\mathbf{f}}(u) \otimes \boldsymbol{\phi}^T$. Furthermore, we have

$$\mathbf{Q}(u) \mathbf{v}_x(v) = (\tilde{\mathbf{d}}(u) \otimes \boldsymbol{\phi}^T) \mathbf{v}_x(v) + (\tilde{\mathbf{f}}(u) \otimes \boldsymbol{\phi}^T) \mathbf{v}_x(v). \tag{A7}$$

According to the mixed product property of the Kronecker product, we then obtain

$$\begin{aligned} \mathbf{Q}(u) \mathbf{v}_x(v) &= (\tilde{\mathbf{d}}(u) \otimes \boldsymbol{\phi}^T \mathbf{v}_x(v)) + (\tilde{\mathbf{f}}(u) \otimes \boldsymbol{\phi}^T \mathbf{v}_x(v)) \\ &= \tilde{\mathbf{d}}(u) (\boldsymbol{\phi}^T \mathbf{v}_x(v)) + \tilde{\mathbf{f}}(u) (\boldsymbol{\phi}^T \mathbf{v}_x(v)). \end{aligned} \tag{A8}$$

From Lemma A1, we see that $\tilde{\mathbf{d}}(u)$ and $\tilde{\mathbf{f}}(u)$ are orthogonal when the cross-correlation function $c_{\mathcal{D},\mathcal{F}}(t) = 0$. Using this condition, Equations (22), (A6) and (A8), we then have

$$\begin{aligned} & \sum_{i=1}^r |\mathbf{v}_z(u)^T \mathbf{P}_i \mathbf{v}_x(v)|^2 \\ &= e(u) |\boldsymbol{\phi}^T \mathbf{v}_x(v)|^2 + e(u) |\boldsymbol{\varphi}^T \mathbf{v}_x(v)|^2. \end{aligned} \tag{A9}$$

From (23), we obtain the result Equation (25).

Appendix C. Proof of Theorem 2

First, we obtain

$$\mathbf{e}_k^T \mathbf{P}_i \mathbf{e}_l = \mathbf{e}_k^T (\mathbf{d}_i \otimes \boldsymbol{\phi}^T) \mathbf{e}_l + \mathbf{e}_k^T (\mathbf{f}_i \otimes \boldsymbol{\varphi}^T) \mathbf{e}_l. \tag{A10}$$

According to the mixed product property of the Kronecker product, we then obtain

$$\begin{aligned} \mathbf{e}_k^T \mathbf{P}_i \mathbf{e}_l &= (\mathbf{e}_k^T \mathbf{d}_i \otimes \boldsymbol{\phi}^T \mathbf{e}_l) + (\mathbf{e}_k^T \mathbf{f}_i \otimes \boldsymbol{\varphi}^T \mathbf{e}_l) \\ &= d_{ik} \phi_l + f_{ik} \varphi_l. \end{aligned} \tag{A11}$$

Furthermore, we have

$$\sum_{i=1}^r |\mathbf{e}_k^T \mathbf{P}_i \mathbf{e}_l|^2 = \sum_{i=1}^r (|d_{ik}|^2 |\phi_l|^2 + |f_{ik}|^2 |\varphi_l|^2). \tag{A12}$$

Since $\sum_{i=1}^r |d_{ik}|^2 = \sum_{i=1}^r |f_{ik}|^2 = 1$, and the fact that one of $|\phi_l|^2$ and $|\varphi_l|^2$ must be zero and the other equal one, we have

$$\sum_{i=1}^r |\mathbf{e}_k^T \mathbf{P}_i \mathbf{e}_l|^2 = 1. \tag{A13}$$

From

$$|p_{i,kl}|^2 = |\mathbf{e}_k^T \mathbf{P}_i \mathbf{e}_l|^2 \tag{A14}$$

we then obtain the result in the theorem.

Appendix D. Proof of Theorem 4

From the chain rule, we have

$$\frac{\partial g}{\partial \rho_i} = \sum_{k=1}^{N_x M_x} -a_{x,k} \hat{a}_{x,k}^{-1} \frac{\partial \hat{a}_{x,k}}{\partial \rho_i}. \tag{A15}$$

From the definition of $\hat{a}_{x,k}$, we have

$$\begin{aligned} \frac{\partial \hat{a}_{x,k}}{\partial \rho_i} &= \frac{\partial \mathbf{e}_k^T \mathbf{V}_x \boldsymbol{\phi} \boldsymbol{\phi}^H \mathbf{V}_x^H \mathbf{e}_k}{\partial \rho_i} \\ &= \frac{\mathbf{e}_k^T \mathbf{V}_x \mathbf{e}_i (\partial \mathbf{e}_i^T \boldsymbol{\phi}) \boldsymbol{\phi}^H \mathbf{V}_x^H \mathbf{e}_k}{\partial \rho_i} \\ &\quad + \frac{\mathbf{e}_k^T \mathbf{V}_x \boldsymbol{\phi} (\partial \boldsymbol{\phi}^H \mathbf{e}_i) \mathbf{e}_i^T \mathbf{V}_x^H \mathbf{e}_k}{\partial \rho_i}. \end{aligned} \tag{A16}$$

According the definition of $\boldsymbol{\phi}$ in Equation (52), we obtain

$$\frac{\partial \mathbf{e}_i^T \boldsymbol{\phi}}{\partial \rho_i} = -j2\pi \mathbf{e}_i^T \boldsymbol{\phi} \tag{A17}$$

and

$$\frac{\partial \boldsymbol{\phi}^H \mathbf{e}_i}{\partial \rho_i} = j2\pi \boldsymbol{\phi}^H \mathbf{e}_i. \quad (\text{A18})$$

Substituting Equations (A17) and (A18) into Equation (A16), we obtain

$$\begin{aligned} \frac{\partial \hat{a}_{x,k}}{\partial \rho_i} &= -j2\pi \mathbf{e}_k^T \mathbf{V}_x \mathbf{e}_i \mathbf{e}_i^T \boldsymbol{\phi} \boldsymbol{\phi}^H \mathbf{V}_x^H \mathbf{e}_k \\ &\quad + j2\pi \mathbf{e}_k^T \mathbf{V}_x \boldsymbol{\phi} \boldsymbol{\phi}^H \mathbf{e}_i \mathbf{e}_i^T \mathbf{V}_x^H \mathbf{e}_k \\ &= \Re(j2\pi \boldsymbol{\phi}^H \mathbf{e}_i \mathbf{e}_i^T \mathbf{V}_x^H \mathbf{e}_k \mathbf{e}_k^T \mathbf{V}_x \boldsymbol{\phi}). \end{aligned} \quad (\text{A19})$$

Then, we obtain

$$\frac{\partial g}{\partial \rho_i} = \sum_{k=1}^{N_x M_x} -a_{x,k} \hat{a}_{x,k}^{-1} \Re(j2\pi \boldsymbol{\phi}^H \mathbf{e}_i \mathbf{e}_i^T \mathbf{V}_x^H \mathbf{e}_k \mathbf{e}_k^T \mathbf{V}_x \boldsymbol{\phi}). \quad (\text{A20})$$

Finally, we can obtain the gradient of g with respect to λ_i in a similar way.

References

- Marzetta, T.L.; Larsson, E.G.; Yang, H.; Ngo, H.Q. *Fundamentals of Massive MIMO*; Cambridge University Press: Cambridge, UK, 2016.
- Lu, A.A.; Gao, X.Q.; Zhong, W.; Xiao, C.; Meng, X. Robust Transmission for Massive MIMO Downlink with Imperfect CSI. *IEEE Trans. Commun.* **2019**, *67*, 5362–5376. [\[CrossRef\]](#)
- Meng, X.; Gao, X.Q.; Xia, X.G. Omnidirectional precoding based transmission in massive MIMO systems. *IEEE Trans. Commun.* **2016**, *64*, 174–186. [\[CrossRef\]](#)
- Meng, X.; Xia, X.G.; Gao, X.Q. Omnidirectional space-time block coding for common information broadcasting in massive MIMO systems. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 1407–1417. [\[CrossRef\]](#)
- Xia, X.G.; Gao, X.Q. A space-time code design for omnidirectional transmission in massive MIMO systems. *IEEE Wirel. Commun. Lett.* **2016**, *5*, 512–515. [\[CrossRef\]](#)
- Golay, M.J. Static multislit spectrometry and its application to the panoramic display of infrared spectra. *JOSA* **1951**, *41*, 468–472. [\[CrossRef\]](#) [\[PubMed\]](#)
- Tseng, C.C.; Liu, C. Complementary sets of sequences. *IEEE Trans. Inf. Theory*. **1972**, *18*, 644–652. [\[CrossRef\]](#)
- Wolti, G. Quaternary codes for pulsed radar. *IRE Trans. Inf. Theory* **1960**, *6*, 400–408. [\[CrossRef\]](#)
- Jedwab, J.; Parker, M.G. Golay complementary array pairs. *Des. Codes Cryptogr.* **2007**, *44*, 209–216. [\[CrossRef\]](#)
- Fiedler, F.; Jedwab, J.; Parker, M.G. A multi-dimensional approach to the construction and enumeration of Golay complementary sequences. *J. Comb. Theory Ser. A* **2008**, *115*, 753–776. [\[CrossRef\]](#)
- Luke, H.D. Sets of one and higher dimensional Wolti codes and complementary codes. *IEEE Trans. Aerosp. Electron. Syst.* **1985**, *AES-21*, 170–179. [\[CrossRef\]](#)
- Dymond, M. Barker Arrays: Existence, Generalization and Alternatives. Ph.D. Thesis, University of London, London, UK, 1992.
- Lu, A.A.; Gao, X.; Meng, X.; Xia, X.G. Omnidirectional Precoding for 3D Massive MIMO with Uniform Planar Arrays. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 2628–2642. [\[CrossRef\]](#)
- Su, D.; Jiang, Y.; Wang, X.; Gao, X. Omnidirectional precoding for massive MIMO with uniform rectangular array—Part I: Complementary codes-based schemes. *IEEE Trans. Signal Process.* **2019**, *67*, 4761–4771. [\[CrossRef\]](#)
- Su, D.; Jiang, Y.; Wang, X.; Gao, X. Omnidirectional precoding for massive MIMO with uniform rectangular array—Part II: Numerical optimization based schemes. *IEEE Trans. Signal Process.* **2019**, *67*, 4772–4781. [\[CrossRef\]](#)
- Li, F.; Jiang, Y.; Du, C.; Wang, X. Construction of Golay Complementary Matrices and Its Applications to MIMO Omnidirectional Transmission. *IEEE Trans. Signal Process.* **2021**, *69*, 2100–2113. [\[CrossRef\]](#)
- Guo, W.; Lu, A.A.; Meng, X.; Gao, X.Q.; Ma, N. Broad Coverage Precoding Design for Massive MIMO with Manifold Optimization. *IEEE Trans. Commun.* **2019**, *67*, 2933–2946. [\[CrossRef\]](#)
- Guo, W.; Lu, A.A.; Meng, X.; Gao, X.; Xia, X.G. Broad Coverage Precoder Design for 3D Massive MIMO System Synchronization. *IEEE Trans. Commun.* **2020**, *68*, 4233–4246. [\[CrossRef\]](#)
- Guo, W.; Lu, A.A.; Gao, X.; Xia, X.G. Broad Coverage Precoder Design for Synchronization in Satellite Massive MIMO Systems. *IEEE Trans. Commun.* **2021**. [\[CrossRef\]](#)
- Telatar, E. Capacity of multi-antenna Gaussian channels. *Eur. Trans. Telecommun.* **1999**, *10*, 585–595. [\[CrossRef\]](#)
- Amari, S.I. *Information Geometry and Its Applications*; Springer: Berlin/Heidelberg, Germany, 2016.
- Absil, P.A.; Mahony, R.; Sepulchre, R. *Optimization Algorithms on Matrix Manifolds*; Princeton University Press: Princeton, NJ, USA, 2009.
- Manton, J.H. Optimization algorithms exploiting unitary constraints. *IEEE Trans. Signal Process.* **2002**, *50*, 635–650. [\[CrossRef\]](#)

24. Henkel, O. Sphere-packing bounds in the Grassmann and Stiefel manifolds. *IEEE Trans. Inf. Theory* **2005**, *51*, 3445–3456. [[CrossRef](#)]
25. Dai, W.; Liu, Y.; Rider, B. Quantization bounds on Grassmann manifolds and applications to MIMO communications. *IEEE Trans. Inf. Theory* **2008**, *54*, 1108–1123. [[CrossRef](#)]
26. Lu, A.A.; Gao, X.; Zheng, Y.R.; Xiao, C. Linear Precoder Design for SWIPT in MIMO Broadcasting Systems With Discrete Input Signals: Manifold Optimization Approach. *IEEE Trans. Commun.* **2017**, *65*, 2877–2888. [[CrossRef](#)]
27. Selvan, S.E.; Amato, U.; Gallivan, K.A.; Qi, C.; Carfora, M.F.; Larobina, M.; Alfano, B. Descent algorithms on oblique manifold for source-adaptive ICA contrast. *IEEE Trans. Neural Netw. Learn. Syst.* **2012**, *23*, 1930–1947. [[CrossRef](#)] [[PubMed](#)]
28. Lewis, A.S.; Malick, J. Alternating projections on manifolds. *Math. Oper. Res.* **2008**, *33*, 216–234. [[CrossRef](#)]