Quadrature Amplitude Modulation Division for Multiuser MISO Broadcast Channels
Zheng Dong, Yan-Yu Zhang, Jian-Kang Zhang, Senior Member, IEEE and Xiang-Chuan Gao

Abstract—This paper considers a discrete-time multiuser multiple-input single-output (MISO) Gaussian broadcast channel (BC), in which channel state information (CSI) is available at both the transmitter and the receivers. The flexible and explicit design of a uniquely decomposable constellation group (UDCG) is provided based on pulse amplitude modulation (PAM) and rectangular quadrature amplitude modulation (QAM) constellations. With this, a modulation division (MD) transmission scheme is developed for the MISO BC. The proposed MD scheme enables each receiver to uniquely and efficiently detect their desired signals from the superposition of mutually interfering cochannel signals in the absence of noise. In our design, the optimal transmitter beamforming problem is solved in a closed-form for two-user MISO BC using max-min fairness as a design criterion. Then, for a general case with more than two receivers, we develop a user-grouping-based beamforming scheme, where the grouping method, beamforming vector design and power allocation problems are addressed by using weighted max-min fairness. It is shown that our proposed approach has a lower probability of error compared with the zero-forcing (ZF) method when the Hermitian angle between the two channel vectors is small in a two-user case. In addition, simulation results also reveal that for the general channel model with more than two users, our user-grouping-based scheme significantly outperforms the ZF, time division (TD), minimum mean-square error (MMSE) and signal-to-leakage-and-noise ratio (SLNR) based techniques when the number of users approaches to the number of base station (BS) antennas and it degrades into the ZF scheme when the number of users is far less than the number of BS antennas in Rayleigh fading channels.

Index Terms—Broadcast channels (BCs), MISO, modulation division, uniquely decomposable constellation group (UDCG), grouping, non-orthogonal multiple access (NOMA).

I. INTRODUCTION

MULTIUSER systems in a downlink, also known as broadcast channels, have long been the main building block of modern wireless communication systems such as cellular system, telephone conference and digital TV broadcasting. In this paper, we concentrate on MISO BC, where one multi-antenna access point serves several single-antenna receivers at the same time. Hence, the design of the receiver can be significantly simplified, since it has only one antenna and the system still enjoys the multi-antenna diversity and multiuser diversity. BC was first introduced by Cover [1], who demonstrated the idea of superposition coding for both binary-symmetric and Gaussian BC. Since then, great efforts have been devoted to obtaining the capacity regions of different BCs and to seeking for the optimal transmission strategies under various constraints. Although the capacity region for the general discrete memoryless broadcast channel (DM-BC) is still unknown, much progress has been made since [1]. In particular, the achievability and converse of the capacity region for the degraded DM-BC were proved by Bergmans [2] and Gallager [3], respectively. Surveys of the literature on the BC can be found in [4]–[6]. On the other hand, if the transmitter and/or receiver nodes are allowed to have more than a single antenna, there will be a Gaussian vector channel in which much higher spectral efficiency (through spatial multiplexing) and reliability (by multi-path diversity) can be achieved by exploiting the scattering medium between the transmitter and receiver antenna arrays [7], [8]. Specifically for the MISO BCs, the achievable throughput was developed in [9] based on Costa’s writing on dirty paper approach [10] that achieves the sum-capacity for a two-user case with a single transmitting antenna. Then, the sum-capacity for a general multiuser MISO BC was established in [11]–[14] by exploiting the uplink-downlink duality between the multiaccess channel (MAC) and the corresponding BC. By using more practical finite-alphabet constellations rather than Gaussian input signals, the transmission schemes that maximize the mutual information between the BS and all the receivers of the multi-user BC are considered in [15]–[18].

The aforementioned information-theoretic analyses serve as a guideline for a general system design. For a transmitter design, non-linear precoding techniques such as the dirty paper coding (DPC) method can be used to approach the sum rate of the MISO BC [12], [13]. It was shown in [9] that a successive interference cancellation procedure, namely the ZF-DPC, can be performed at the transmitter side to completely remove mutual interference between receivers. Given the complexity of DPC, the Tomlinson-Harashima precoding (THP) method [19]–[22] serves as a suboptimal but practical approximation of DPC by introducing a modulo operation to transmitted symbols. Despite the fact that there is a modulo loss [23], the transmitted symbols are guaranteed to have a finite dynamic range. All these precoding methods were primarily devoted to improving the sum rate of multiuser MISO BCs. On the other hand, practical transmitter designs may also aim to improve signal quality at the receiver side. Among...
such transmitter designs, linear precoding techniques receive tremendous attention because of their potential and simplicity. Using signal-to-noise ratio (SNR) as a design criterion, it was shown that transmitter beamforming can increase the received SNR in the multiuser MISO channel by performing optimization on the beamformer design and the power allocation scheme \([24]–[26]\). In addition, by employing MMSE as a performance measure, an optimal precoder was proposed in \([27]\) with regularized channel inversion, which outperforms the ZF scheme when the channel condition number is large. By maximizing SLNR for all users simultaneously, a closed-form beamforming method was given in \([28]\). Even so, however, it was demonstrated that the ZF beamforming technique, which is simple to implement, can achieve most of the capacity in moderate and high SNR regimes \([29], [30]\). Comprehensive comparisons of these schemes can be found in \([31]\).

As we can see, interference has long been the central focus for meeting the ever increasing requirements on quality of service in modern and future wireless communication systems. The key to the understanding of multiuser communications is the understanding of interference. Traditional approaches, which treat interference as a detrimental phenomenon are, therefore, to suppress or eliminate it. The classical information-theoretic study on the two-user Gaussian interference channel \([32]\) suggests us that we should treat the interference as noise when it is weak and that the optimal strategy is to decode the interference when it is very strong. In addition, when the level of interference is of the same order of the power of a desired signal, one good strategy is to suppress all the undesired interferences into a smaller space that has no overlap with the signal space \([33]–[35]\). However, some recent innovative approaches, which consider interference as a useful resource, are, thus, to make use of it for developing energy efficient and secure 5G communication systems \([36]–[38]\). For example, interference can be used for boosting up the desired signal \([39]–[41]\) for energy harvesting \([42]–[49]\) or to deteriorate the signal of the eavesdropper for secure communication \([50]\).

Inspired by \([34]\), in this paper we consider the management of interference for BC by carefully designing communication signals. To better elaborate on our idea, we would like to revisit some early seminal work \([51], [52]\) of how to strategically take advantage of the finite alphabet properties of digital communication signals for managing interference for a two-user access binary channel. Essentially, the Kasami and Lin’s main idea is to carefully design such two finite length codes for the two users that when any sum binary signal of the two user codewords is received in a noiseless environment, each individual user codeword can be uniquely decoded, as well as in a noisy case, the resulting error is able to be correctable. Specifically, such uniquely decodable code (UDC) was explicitly constructed for a two-user binary ensure channel \([53], [54]\). Then, this important concept was extended to the design of UDC based on trellis modulation for an \(N\)-user binary multi-access channel \([55]\), which allows a number of users to access a common receiver simultaneously and outperforms the time sharing method in terms of the probability of error. Furthermore, the design of trellis-coded UDC was investigated in a complex number domain to extract the desired signal from the superposition of the signal and cochannel interferences \([56]–[58]\). In addition, the concept of UDC was also exploited to design variety of multi-resolution modulation schemes for BC and it was shown that they not only outperform the frequency division scheme by properly designing the resulting constellation \([59]–[62]\), but also reduce the transmission delay of the network at the cost of increased transmitting power for fading channels \([61]\). Recently, a pair of uniquely decodable constellations was designed to study the capacity region of a two user Gaussian multi-access channel \([63]\).

Indeed, it is the above aforementioned factors that greatly motivate and enlighten us to look into interference from the perspective of signal processing. In this paper, we are interested in exploring a novel signal processing technique to manage interference for BC, which allows strongly interfered user signals to cooperate with each other as a common desired sum signal from which each individual user signal is able to be uniquely and efficiently decoded. Specifically, our main contributions of this paper can be summarized as follows:

1) An explicit construction of UDCG, which can be considered as a UDC in the complex domain for a multiuser case, for general PAM and rectangular QAM constellations for \(n\) number of users is proposed. The main difference between our UDC design and all currently available UFC designs in literature is that in our UDC design, the sum-constellation and all the user constellations are PAM and QAM constellations with different scales. It is because of this nice geometric structure that once the sum signal is received, each individual user signal can be efficiently and uniquely decoded (see Algorithms 1 and 2).

2) Using the newly developed UDCG, we propose a novel NOMA transmission scheme called QAM-modulation division for the multiuser MISO BC. First, an optimal beamformer that maximizes the minimum SNR between the two receivers is obtained in a closed-form for a two-user case. Then, for a general network topology with more than two receivers, a grouping-based transmitter design problem is also investigated, with ZF eliminating the inter-group interference, where the grouping policy, the beamformer design and power allocation strategy are addressed. It is demonstrated that when the Hermitian angle of the two channel vectors is small, our proposed division method has a much lower probability of error, which confirms that the NOMA method with proper interference control is a promising technology for 5G communications.

Our work can be considered as a concrete, simple and systematic design of the constellation for NOMA \([64], [65]\), serving different users with different power levels, and it has a considerable spectral gain over the traditional methods.

Notations: Matrices and column vectors are denoted by boldface characters with uppercase (e.g., \(A\)) and lowercase (e.g., \(b\)), respectively. The transpose of \(A\) is denoted by \(A^T\) while the Hermitian transpose of \(A\) (i.e., the conjugate and
transpse of \( A \) is denoted by \( A^H \). \( \| b \| \) denotes the Euclidean norm of \( b \). We let \( j = \sqrt{-1} \) and \( \binom{n}{k} = \frac{n!}{k!(n-k)!} \) be the binomial coefficient. In addition, \( \prod \) stands for the production operation and \( \mathcal{A} \setminus \mathcal{B} = \{ x \in \mathcal{A} \text{ and } x \notin \mathcal{B} \} \). \( [a] \) is the floor function which represents the largest integer no more than \( a \) and \( \lceil b \rceil \) is the ceiling function that returns the smallest integer not less than \( b \). \( a \mod b = a - b \times \text{fix}(a/b) \) where \( \text{fix}(\cdot) \) rounds the variable to the nearest integer towards zero.

II. UNIQUELY DECOMPOSABLE CONSTELLATION GROUP

In this section, we introduce the definition of the UDCG and then, provide the flexible and the explicit construction of the UDCG using commonly-used PAM and QAM constellations and the corresponding efficient decoding algorithms.

**Definition 1 (UDCG):** A group of constellations \( \{ \mathcal{X}_i \}_{i=1}^N \) is said to form a UDCG, denoted by \( \{ \mathcal{X}_i \}_{i=1}^N = \mathcal{G} \), if there exist two groups of \( x_i, \tilde{x}_i \in \mathcal{X}_i \), for \( i = 1, 2, \ldots, N \) such that \( \sum_{i=1}^N x_i = \sum_{i=1}^N \tilde{x}_i \), then, we have \( x_i = \tilde{x}_i \) for \( i = 1, 2, \ldots, N \).

For presentation convenience, constellation \( \psi_{i=1}^N \mathcal{X}_i \) is called the sum-constellation of all \( \mathcal{X}_i \) and each \( \mathcal{X}_i \) is called the \( i \)-th sub-constellation of \( \psi_{i=1}^N \mathcal{X}_i \) or \( i \)-th user constellation.

The concept of UDCG can be considered as an extension of UDC in binary field to the complex number field. In this section, we introduce the definition of the UDCG and its construction, to emphasize a major difference between the definition of UDC and the traditional concept of UDC. In our Definition 1, we are interested in each analysis component of the decomposition, i.e., the geometrical structure of each user-constellation, as well as the synthesis component of the decomposition, i.e., the geometrical structure of the sum constellation.

The following property reveals such a fact that checking whether or not a group of constellations forms a UDCG is equivalent to checking whether or not the cardinality of the sum constellation is equal to the product of the cardinalities of the user-constellations.

**Property 1 (Unambiguity):** Given a group of constellations \( \{ \mathcal{X}_i \}_{i=1}^N \) with each having a finite size, if we let \( \mathcal{G} = \{ \sum_{i=1}^N x_i : x_i \in \mathcal{X}_i \} \), then, \( \mathcal{G} = \psi_{i=1}^N \mathcal{X}_i \) if and only if \( \| \mathcal{G} \| = \prod_{i=1}^N | \mathcal{X}_i | \).

**Proof:** Let \( \mathcal{Y} \) denote a set of \( N \)-tuples \( \mathcal{Y} = \{(x_1, x_2, \ldots, x_N) : x_i \in \mathcal{X}_i \} \). Then, \( |\mathcal{Y}| = \prod_{i=1}^N | \mathcal{X}_i | \) by the combinational rule of product and \( \mathcal{Y} \) is a finite set, since each \( \mathcal{X}_i \) is a finite set.

If \( |\mathcal{G}| = \prod_{i=1}^N | \mathcal{X}_i | \), then, we have \( |\mathcal{G}| = |\mathcal{Y}| \). Since \( \mathcal{G} \) and \( \mathcal{Y} \) are finite sets, there exists a bijection map between these two sets, which is denoted by \( f_{bij} : \mathcal{G} \rightarrow \mathcal{Y} \).

Without loss of generality, we let \( f_{bij} (\sum_{i=1}^N x_i) = (x_1, x_2, \ldots, x_N) \). As \( f_{bij} : \mathcal{G} \rightarrow \mathcal{Y} \) is a bijective map, then, \( \sum_{i=1}^N x_i = \sum_{i=1}^N \tilde{x}_i \) must have \( (x_1, x_2, \ldots, x_N) = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N) \) and hence, \( x_i = \tilde{x}_i \) for \( i = 1, 2, \ldots, N \). Therefore, \( \mathcal{G} = \psi_{i=1}^N \mathcal{X}_i \).

If \( \mathcal{G} = \psi_{i=1}^N \mathcal{X}_i \), by Definition 1 for any \( (x_1, x_2, \ldots, x_N), (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N) \in \mathcal{G} \) satisfying \( \sum_{i=1}^N x_i = \sum_{i=1}^N \tilde{x}_i \), we have \( x_i = \tilde{x}_i \) for \( i = 1, 2, \ldots, N \), or equivalently \( (x_1, x_2, \ldots, x_N) = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N) \). Therefore, there exists an injective function \( f_{bij} : \mathcal{Y} \rightarrow \mathcal{G} \).
and for $i \geq 2$,

$$
\hat{x}_i = \begin{cases} 0 & K_i = 0 \\
\left( \frac{\hat{g} + 2^{K_i - 1} - \hat{g} + 2^{K_i - 1}}{\mod 2^{K_i - 1}} \right) \mod 2K_i \cdot \frac{2^{K_i - 1}}{2} ) 2^{\sum_{i=1}^{K_i} K_i} & K_i \geq 1 
\end{cases}
$$

(3)

The proof of Algorithm 1 is given in Appendix A. As we know, a rectangular QAM constellation is generated from a pair of the PAM constellations. Hence, Theorem 1 can be extended to the PAM and QAM mixed case in a straightforward manner, whose proof, therefore, is omitted.

**Theorem 2 (QAM):** For two positive integers $N$ and $K = K^c + K^s$, with $K^c$ and $K^s$ being nonnegative integers, let $K^c_i$ and $K^s_i$ for $i = 1, 2, \cdots, N$ denote any two given nonnegative integer sequences satisfying $K^c = \sum_{i=1}^{N} K^c_i$ and $K^s = \sum_{i=1}^{N} K^s_i$ with $K^c_i + K^s_i > 0$. Then, there exists a PAM and QAM mixed constellation $Q$ such that $Q = \cup_{i=1}^{N} X_i$, where $X_i = \mathcal{X}_{K^c} \cup \mathcal{X}_{K^s}$ with $jX^c = \{ jx : x \in X^c \}$. In addition, $Q^{c} = \cup_{i=1}^{N} X^c_i$ and $Q^{s} = \cup_{i=1}^{N} X^s_i$ are two PAM UDCGs given in Theorem 1 according to the rate-allocation $K^c_i$ and $K^s_i$, respectively.

Similar to Algorithm 1, we also have an efficient detection method for a UDCG based on the QAM constellation below:

**Algorithm 2 (Fast detection of QAM UDCG):** Let a UDCG $\mathcal{G} = \cup_{i=1}^{N} X_i$ be generated from the QAM constellation by Theorem 2. Then, for an observed noisy complex signal $y = \sum_{i=1}^{N} x_i + \xi$, where $x_i \in \mathcal{X}_i$ and $\xi \sim \mathcal{CN}(0, \sigma^2)$ is an additive circularly-symmetric complex Gaussian noise, all the user-signals are efficiently estimated using the following two successive steps:

1) **Quantization of the sum signal:** Let $y = y^{c} + jy^{s}$. Find the quantized PAM signal $\hat{y}^{c} \in \mathcal{Q}^{c}$ and $\hat{y}^{s} \in \mathcal{Q}^{s}$ of $y^{c}$ and $y^{s}$, respectively by solving the following optimization problems

$$
\hat{y}^{c} = \arg\min_{y^{c} \in \mathcal{Q}^{c}} |y^{c} - g^{c}|, \\
\hat{y}^{s} = \arg\min_{y^{s} \in \mathcal{Q}^{s}} |y^{s} - g^{s}|
$$

2) **Decoding of the user-signals:** By Algorithm 1 the estimates of all the user real signals $\hat{x}^{c}_i \in \mathcal{X}^{c}_i$ and $\hat{x}^{s}_i \in \mathcal{X}^{s}_i$ for $i = 1, 2, \cdots, N$ can be efficiently obtained such that $\hat{g}^{c} = \sum_{i=1}^{N} \hat{x}^{c}_i$ and $\hat{g}^{s} = \sum_{i=1}^{N} \hat{x}^{s}_i$, and thus, $\hat{x}_i = \hat{x}^{c}_i + j \hat{x}^{s}_i$.

III. MODULATION DIVISION FOR TWO-USER MISO BC

Our primary purpose in this section is to apply the UDCG based on the QAM constellation to the design of an optimal beamformer for a two-user BC. Toward this goal, let us specifically consider a MISO BC having two single-antenna receivers and a BS equipped with $M$ antennas which transmits independent and identically distributed (i.i.d.) signals $s_1$ and $s_2$ simultaneously to the two receivers. The channel is assumed to be flat fading and quasi-static. Let $h_1 = [h_{1,1}, h_{2,1}, \cdots, h_{M,1}]^H$ and $h_2 = [h_{1,2}, h_{2,2}, \cdots, h_{M,2}]^H$ denote, respectively, the channel links between BS and user 1 and 2, which are perfectly available at the transmitter. Here, our main idea is that BS treats the two channels to be strongly interfered each other and hence, in order to serve the two receivers at the same time, the BS transmits a sum signal $s = s_1 + s_2$, with one common beamforming vector $w \in \mathbb{C}^{M \times 1}$ to be designed, where $s_1$ and $s_2$ are randomly chosen from an aforementioned UDCG $Q = \mathcal{X}_1 \cup \mathcal{X}_2$ such that $s_1 \in \mathcal{X}_1$ and $s_2 \in \mathcal{X}_2$. Then, the signal intended for each receiver can be decoded separately by using our fast detection method described in Algorithm 1 or 2.

A. Modulation Division for Two-User Case

The equivalent complex-baseband channel model for the received signals at the two receivers is given by

$$
y_1 = h_1^H w s + \xi_1, \\
y_2 = h_2^H w s + \xi_2,
$$

where $s$ is the information carrying symbol for both users with $E[|s|^2] = 1$ and hence the total transmitted power is $P = E[|s|^2] w^H w = w^H w$. Also, $\xi_1, \xi_2 \sim \mathcal{CN}(0, \sigma^2)$ are additive circularly-symmetric complex Gaussian noise arising at each receiver. It is worth noting that the case where different receivers have different noise levels can be incorporated into our model by performing a scaling operation on the channel coefficients. Hence, the noises are assumed to be of equal variance. The SNRs for the sum signal $s$ at each receiver are expressed by

$$
\text{SNR}_{md_1} = \frac{|h_1^H w|^2}{\sigma^2}, \\
\text{SNR}_{md_2} = \frac{|h_2^H w|^2}{\sigma^2}.
$$

By using a max-min fairness on the received SNR, we aim to solve the following optimization problem:

**Problem 1:** Find the beamforming vector $w$ such that

$$
\max_{w} \min \{ w^H h_1 h_1^H w, w^H h_2 h_2^H w \},
$$

s.t. $w^H w = P$.

Without loss of generality, we assume that $\|h_1\|, \|h_2\| \neq 0$, since otherwise, the solution is trivial and in fact we can not achieve reliable communication to both users simultaneously in this case. Now, let

$$
A = h_1 h_1^H - h_2 h_2^H,
$$

(5)

1) If $h_1$ and $h_2$ are linearly dependent (or equivalently, $h_1 = \tau h_2$ for some $\tau \in \mathbb{C}$), then $A = (|\tau|^2 - 1) h_2 h_2^H$. Hence, if $|\tau| = 1$, we have $A = 0$. Otherwise, $A$ has rank one.

2) If $h_1$ and $h_2$ are linearly independent (i.e., $h_1 \neq \tau h_2, \forall \tau \in \mathbb{C}$), then, the rank of $A$ is 2. Let the eigenvalue decomposition of $A$ be given by

$$
A = V \Sigma V^H,
$$

(6)

where $V$ is a unitary matrix and $\Sigma = \text{diag}(\lambda_1, -\lambda_2, 0, \ldots, 0)$ with $\lambda_1 > 0$ and $\lambda_2 > 0$. From (6), we can obtain

$$
\Sigma = V^H (h_1 h_1^H - h_2 h_2^H) V = \tilde{h}_1 \tilde{h}_1^H - \tilde{h}_2 \tilde{h}_2^H,
$$

(7)
where $\hat{h}_1 = V^H h_1 = [\hat{h}_{1,1}, \hat{h}_{1,2}, 0, \ldots, 0]^T$ and $\hat{h}_2 = V^H h_2 = [\hat{h}_{2,1}, \hat{h}_{2,2}, 0, \ldots, 0]^T$. We also denote $\tilde{w} = V^H \mathbf{w} = [\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_M]^T$ and 
\[ \mathbf{H} = V^H V = [\hat{h}_1 \ \hat{h}_2]. \] (8)

Equation (7) is equivalent to 
\[ |\hat{h}_{1,1}|^2 - |\hat{h}_{2,1}|^2 = \lambda_1, \] (9a)
\[ |\hat{h}_{1,2}|^2 - |\hat{h}_{2,2}|^2 = -\lambda_2, \] (9b)
\[ \frac{\hat{h}_{1,1}}{\hat{h}_{1,2}} = \frac{\hat{h}_{2,1}}{\hat{h}_{2,2}}. \] (9c)

The above relationships can be characterized by 
\[ \begin{bmatrix} \hat{h}_{1,1} \\
\hat{h}_{1,2} \\
\hat{h}_{2,1} \\
\hat{h}_{2,2} \end{bmatrix} = \begin{bmatrix} \sqrt{\lambda_1} \sec \theta e^{j\beta} \\
\sqrt{\lambda_2} \tan \theta e^{j\gamma} \\
\sqrt{\lambda_1} \tan \theta e^{j\gamma} \\
\sqrt{\lambda_2} \sec \theta e^{j\gamma} \end{bmatrix} \] \[ \text{where } \theta = \arccos \frac{\sqrt{\lambda_1}}{\lambda_1}, \quad 0 \leq \theta < \pi/2 \quad \text{and} \quad \beta = \arg(\hat{h}_{1,1}), \quad \gamma = \arg(\hat{h}_{2,2}) \quad \text{and} \quad \alpha = \arg(\hat{h}_{1,2}) - \arg(\hat{h}_{2,2}). \] (10)

Now, we are ready to state one of our main results in this paper, i.e., the optimal solution to two-user beamforming problem in [4].

**Theorem 3 (Optimal beamforming for two-user cases):** Let 
\[ f(\mathbf{w}) = \min \{ \| \mathbf{w}^H h_1^1 h_1^2 \|, \| \mathbf{w}^H h_2^1 h_2^2 \| \} \] Then, the optimal solution $\mathbf{w}^{opt}$ to Problem 1 is determined as follows:

Scenario 1: $h_1 = \tau h_2, \tau \in \mathbb{C}$.
\[
\max_{\mathbf{w}^H \mathbf{w} = P} f(\mathbf{w}) = \min \left\{ P \| h_1 \|^2, P \| h_2 \|^2 \right\},
\]
where $\mathbf{w}^{opt} = \sqrt{P h_1} / \| h_1 \|$.

Scenario 2: $h_1 \neq \tau h_2, \forall \tau \in \mathbb{C}$. The solution is given below:

1) $\lambda_1 \leq \lambda_2$.
   a) for $0 \leq \sin \theta \leq \frac{\lambda_1}{\lambda_2}$, we have
   \[
   \max_{\| \mathbf{w} \|^2 = P} f(\mathbf{w}) = \frac{P \lambda_1 \lambda_2 (1 + \sin \theta)^2}{\lambda_1 + \lambda_2 \cos^2 \theta},
   \]
   where $\mathbf{w}^{opt} = V \mathbf{w}^{opt} = \begin{bmatrix} P \lambda_1 \lambda_2 e^{j\beta} \\
\sqrt{\lambda_2} \lambda_1 \lambda_2 e^{j(\beta - \alpha)} \end{bmatrix}$. 
   b) for $\frac{\lambda_1}{\lambda_2} < \sin \theta < 1$, we have
   \[
   \max_{\| \mathbf{w} \|^2 = P} f(\mathbf{w}) = \frac{P (\lambda_1 + \lambda_2 \sin \theta)^2}{\cos^2 \theta},
   \]
   where $\mathbf{w}^{opt} = \sqrt{P \lambda_1 \lambda_2} \sin \theta e^{j\gamma} [\sqrt{\lambda_1 \lambda_2} \sin \theta e^{j(\beta - \alpha)}, 0, \ldots, 0]$. 

2) $\lambda_1 > \lambda_2$.
   a) for $0 \leq \sin \theta \leq \frac{\lambda_1}{\lambda_2}$, we have
   \[
   \max_{\| \mathbf{w} \|^2 = P} f(\mathbf{w}) = \frac{P \lambda_1 \lambda_2 (1 + \sin \theta)^2}{\lambda_1 + \lambda_2 \cos^2 \theta},
   \]
   where $\mathbf{w}^{opt} = V \mathbf{w}^{opt} = \begin{bmatrix} P \lambda_1 \lambda_2 e^{j\gamma} \\
\sqrt{\lambda_1 \lambda_2} \sin \theta e^{j(\beta - \alpha)} \end{bmatrix}$. 
   b) and for $\frac{\lambda_1}{\lambda_2} < \sin \theta < 1$, we have
   \[
   \max_{\| \mathbf{w} \|^2 = P} f(\mathbf{w}) = \frac{P \lambda_1 \lambda_2 \sin \theta + \lambda_2}{\cos^2 \theta},
   \]
where $\mathbf{w}^{opt} = V \mathbf{w}^{opt}$ with $\tilde{w}^{opt} = \begin{bmatrix} \sqrt{\frac{P \lambda_1 \lambda_2 \sin \theta + \lambda_2}{\lambda_1 + \lambda_2 \cos^2 \theta}} \\
\sqrt{\frac{P \lambda_1 \lambda_2 \sin \theta + \lambda_2}{\lambda_1 + \lambda_2 \cos^2 \theta}} e^{j(\gamma + \alpha)}, 0, \ldots, 0 \end{bmatrix}$. 

The proof of Theorem 3 can be found in Appendix B. We would like to make the following comments on Theorem 3:

1) The problem dealt with in Theorem 3 is different from the physical-layer multicasting problem discussed in [67], where a group of users is interested in a common message. However, in our model, the information symbols intended for separate users are different and form a UDCG. In addition, despite the fact that the optimization problem in [67] is more general, its solution is numerical and not necessarily global. Our Theorem 3 gives the global solution in the closed form for the two-user case.

2) Here, it is should mentioned clearly that when we have finished our manuscript, we realize work [68] dealing with the same optimization problem as ours for the optimal design of a multicast beamformer with superposition coding. Unfortunately, the optimal solution given in [68] holds only when the condition $0 \leq \sin \theta \leq \min(\lambda_1, \lambda_2)$ in Theorem 3 is satisfied. In other words, under that condition, the optimal solution is achieved in boundary $\mathbf{w}^{opt} h_1 h_1 w = \mathbf{w}^{opt} h_2 h_2 w$, which, however, is not true in general. In fact, the condition under which the solution to the maximin optimization problem is reached in the boundary is widely and deeply studied in [69].

**B. The Comparison between MD and ZF Method**

In this section, we compare the effective performance of our proposed MD beamforming with that of ZF beamforming [29]. For simplicity, we assume that the information rates of the two receivers are the same and that the channel matrix $\mathbf{H} = [h_1 \ h_2]$ has full column rank, whose singular values are $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$ with $\mu_1, \mu_2 > 0$. Then, the received SNR for ZF beamforming with the max-min fairness criterion is determined by

\[
\text{SNR}_{zf} = \frac{P}{\sigma^2 \sum_{i=1}^{\infty} (|\mathbf{H}^H \mathbf{H}|^{-1})_{i,i}} = \frac{P \mu_1 \mu_2}{\sigma^2 (\mu_1 + \mu_2)}. \] (11)

On the other hand, for the MD method, by Theorem 3, the minimum received SNR between the two users for the sum signal is given by

\[
\text{SNR}_{md} = \max_{\| \mathbf{w} \|^2 = P} f(\mathbf{w}). \] (12)

Jointly considering (8) and (10), we can obtain

\[
\mu_1 + \mu_2 = \text{tr}(\mathbf{H}^H \mathbf{H}) = \lambda_1 + \lambda_2 \frac{1 + \sin^2 \theta}{\cos^2 \theta}, \] (13a)
\[
\mu_1 \mu_2 = \det(\mathbf{H}^H \mathbf{H}) = \lambda_1 \lambda_2. \] (13b)

Hence, (11) can be further represented in terms of $\lambda_1$ and $\lambda_2$ as

\[
\text{SNR}_{zf} = \frac{P \lambda_1 \lambda_2 \cos^2 \theta}{\sigma^2 (\lambda_1 + \lambda_2) (1 + \sin^2 \theta)}. \] (14)
For discussion convenience, we define $\kappa = \lambda_1/\lambda_2$ and the SNR gain as $\eta(\kappa, \theta) = 10 \log_{10} \frac{SNR_{CH}}{SNR_{CF}}$. Then, by Theorem 3 the SNR gain as a function of $\kappa$ and $\theta$ is given by

**Corollary 1 (SNR Gain in terms of $\kappa$ and $\theta$):** The following statements are true:

1) If $0 < \kappa \leq 1$ and $0 \leq \sin \theta \leq \kappa$, then $\eta(\kappa, \theta) = 10 \log_{10} \frac{1}{1-\sin \theta \sin \phi}$.
2) If $0 < \kappa \leq 1$ and $\kappa < \sin \theta < 1$, then $\eta(\kappa, \theta) = 10 \log_{10} \frac{(1+\sin^2 \theta)(\kappa+\sin^2 \theta)(1+\kappa/\kappa)}{\cos^2 \theta}$.
3) If $1 < \kappa$, $0 \leq \sin \theta < 1$ and $\theta < \pi$, then $\eta(\kappa, \theta) = 10 \log_{10} \frac{1}{1-\sin \theta}$.
4) If $1 < \kappa$ and $1/\kappa < \sin \theta < 1$, then $\eta(\kappa, \theta) = 10 \log_{10} \frac{1}{1+\sin^2 \theta/(1+\kappa)}$.

By Corollary 1, the SNR gain can be evaluated once $H$ has been obtained. To further appreciate the physical meaning of the SNR gain, we have the following lemma.

**Lemma 1:** Given channel $H = [h_1 \ h_2]$, let $\sqrt{\rho_1}$ and $\sqrt{\rho_2}$ denote its two singular values, $\|h_1\|^2 = a$ and $\|h_2\|^2 = b$, and $\|h_1 h_2\|^2 = c$. Also we let $\lambda_1$ and $\lambda_2$ be defined in (6). Then, we have $\mu_1 = \frac{a+b+\sqrt{(a+b)^2+4c^2}}{2}$, $\mu_2 = \frac{a+b-\sqrt{(a+b)^2+4c^2}}{2}$, and $\lambda_1 = \frac{a-b+\sqrt{(a-b)^2-4c^2}}{2}$, $\lambda_2 = \frac{a-b-\sqrt{(a-b)^2-4c^2}}{2}$.

The proof of Lemma 1 can be found in Appendix C. By Lemma 1 we can immediately have the following corollary:

**Corollary 2:** Let $\theta$ be defined in (10), and $a, b$ and $c$ be defined in Lemma 1. Then, we have $\sin \theta = \frac{2c}{a+b+\sqrt{(a+b)^2-4c^2}}$.

To gain more physical meaning of the SNR gain, we now define $\rho = a/b$ and $\cos \varphi = \frac{b h_2^\dagger h_1}{\|h_1\| \|h_2\|}$, where $\varphi \in [0, \pi/2]$ is called the Hermitian angle [70] between two channel vectors $h_1$ and $h_2$. The SNR gain as a function of $\rho$ and $\varphi$ is defined by $\nu(\rho, \varphi) = 10 \log_{10} \frac{SNR_{CH}}{SNR_{CF}}$. Inserting $\theta$ in Corollary 2 into Corollary 1 and using Lemma 1 and Corollary 2, we can have the following corollary, whose proof is omitted.

**Corollary 3 (SNR Gain in terms of $\rho, \varphi$):** The following statements are true:

1) If $0 < \rho \leq 1$ and $0 \leq \cos \varphi \leq \sqrt{\rho}$ (i.e., $0 < \kappa \leq 1$, $0 \leq \sin \theta \leq \kappa$), then $\nu(\rho, \varphi) = 10 \log_{10} \frac{1}{1+\rho}$.
2) If $0 < \rho \leq 1$ and $\sqrt{\rho} < \cos \varphi \leq 1$ (i.e., $0 < \kappa \leq 1$, $\kappa < \sin \theta < 1$), then $\nu(\rho, \varphi) = 10 \log_{10} \frac{1}{\cos \varphi}$.
3) If $1 < \rho$ and $0 \leq \cos \varphi \leq 1/\sqrt{\rho}$ (i.e., $1 < \kappa < 1/\kappa$, $0 \leq \sin \theta \leq 1/\kappa$), then $\nu(\rho, \varphi) = 10 \log_{10} \frac{1}{1+\rho \cos \varphi}$.
4) If $1 < \rho$ and $1/\sqrt{\rho} < \cos \varphi \leq 1$ (i.e., $1 < \kappa < 1/\kappa$, $0 < \sin \theta < 1$), $\nu(\rho, \varphi) = 10 \log_{10} \frac{1}{1+\rho}$.

From Corollary 3 it is not hard to obtain $\nu(\rho, \varphi) \geq 0$ for all $\rho > 0, 0 \leq \varphi \leq \pi/2$. Hence, The SNR gain of our proposed MD beamforming is at least as good as that of ZF beamforming. To see it more clearly, the SNR gain in terms of $\rho$ and $\varphi$ are plotted in Fig. 1 for $0 < \rho < 2$ and $0 < \varphi < \frac{\pi}{2}$. It can be observed that for given $\rho$, the SNR gain is determined by the Hermitian angle $\varphi$ between two channel vectors. When $\varphi$ approaches zero, i.e., $h_1$ and $h_2$ are approximately aligned with each other, the SNR gain is extremely large. For more clarity, the SNR gains for some specific cases are shown in Table 1.

![Fig. 1. SNR Gain in terms of $\rho$, $\varphi$ in dB](image-url)

| $\varphi$ (rad) | 0 | 1/8 | 1/4 | 1/2 | $\pi$ |
|-----------------|---|-----|-----|-----|-----|
| $\rho$ = 1/16  | 35.43 | 21.46 | 12.00 | 6.28 | 3.27 |
| $\rho$ = 1/8   | 35.67 | 21.71 | 12.25 | 6.53 | 3.52 |
| $\rho$ = 1/4   | 36.13 | 22.16 | 12.71 | 6.99 | 3.98 |
| $\rho$ = 1/2   | 36.92 | 23.96 | 13.50 | 7.78 | 4.77 |
| $\rho$ = 1     | 38.17 | 24.29 | 14.68 | 8.73 | 5.33 |

**IV. GROUPED MODULATION DIVISION TRANSMISSION FOR MULTIUSER MISO BC**

In this section, a novel grouped modulation division transmission method is proposed for the multiuser MISO BC. The grouping algorithm is developed for cases where each group has at most two users. Then, the optimal beamforming vector and power allocation are all given in a closed-form.
A. System Model

We consider a communication system with a BS equipped with a set of $M$ transmitting antennas communicating with $N$ single antenna users $U = \{U_1, U_2, \ldots, U_N\}$ simultaneously in the downlink as illustrated in Fig. 2. The channel links from BS to all the receivers can be stacked together into a matrix $H \in \mathbb{C}^{M \times N}$, the $k$-th column of which is denoted by $h_k = [h_{1,k}, h_{2,k}, \ldots, h_{N,k}]^T$, representing the channel link from BS to the $k$-th receiver. The $N$ different receiver nodes can be further divided into $G \leq N$ groups, with $k$-th group containing $N_k$ users, say, $U_{k1}, U_{k2}, \ldots, U_{kN_k}$, such that $N = \sum_{k=1}^{G} N_k$. For clarity, all the users are relabelled to represent the grouping results. If we let $S$ denote a set consisting of all the users, then, it can be partitioned into $S = S_1 \cup S_2 \cup \ldots \cup S_G$, $S_k \cap S_l = \emptyset, \forall k \neq l$, where $S_k = \{U_{k1}, U_{k2}, \ldots, U_{kN_k}\}$. Correspondingly, the channel vector from BS to $U_{k} \in S_k$ is now denoted by $h_k$. For $k \in \{1, 2, \ldots, G\}$ and $\ell \in \{1, 2, \ldots, N_k\}$ and in turn, the channel matrix between BS and all the users in $S_k$ is represented by

$$H_k = [h_{k1}, h_{k2}, \ldots, h_{kN_k}].$$

Meanwhile, the matrix containing channel links from BS to all the other users in $S \setminus S_k$ is represented by

$$H_{\bar{k}} = [H_1, \ldots, H_{k-1}, H_{k+1}, \ldots, H_G].$$

The grouping strategy of dividing the original user set $U$ into $G$ mutually disjoint subsets $S_k$ of $S$ will be discussed later. For now, let us suppose that the grouping method has been given. Then, the communication process is carried out in the following two steps.

Firstly, we assume that all the users in the group $S_k$ use one UDCG $Q_k = \{\chi_{k1}, \chi_{k2}, \ldots, \chi_{kN_k}\}$ adopted by user $U_{k_1}$. The rate allocation of group $S_k$ is based on the sum decomposition $K_k = \sum_{\ell=1}^{N_k} K_{k\ell}, \forall k \in \{1, 2, \ldots, G\}$, where $K_k$ is the sum rate for all the users in group $S_k$ and $K_{k\ell} = \log_2(|\chi_{k\ell}|)$ is the rate of user $U_{k\ell}$.

Secondly, a normalized information carrying signal $s_{k\ell}$ intended for all the users in $S_k$ is generated by using the UDCG $Q_k = \{\chi_{k1}, \chi_{k2}, \ldots, \chi_{kN_k}\}$, i.e.,

$$s_{k\ell} = \frac{1}{\sqrt{\mathbb{E}[\sum_{l=1}^{N_k} s_{k_l}^2]}} \sum_{l=1}^{N_k} s_{k_l}, \quad \forall k \in \{1, 2, \ldots, G\},$$

where $s_{k_l}$ is assumed to be independently and uniformly drawn from the corresponding sub-constellation $\chi_{k_l}$. It can be showed that the sum signal $s_k$ is also uniformly distributed over the scaled sum-constellation $\frac{1}{\sqrt{\mathbb{E}[\sum_{l=1}^{N_k} s_{k_l}^2]}} Q_k$ such that $\mathbb{E}[|s_k|^2] = 1, \forall k$. It is worth pointing out that due to the power normalization, the minimum Euclidean distance of the constellation points of $s_k$ for different user group $S_k$ might be different. Since the probability of error for the sum signal $s_k$ is dominated by the minimum Euclidean distance in high SNR regimes, it is anticipated that more transmitting power is required for $s_k$ when the sum-constellation is large with the same target error performance.

Then, all users in the same group $S_k$ adopt the same precoding vector $w_k$ and the weighted signals are fed into $M$ transmitter antennas at BS. Hence, the received signal at user $U_{k_l}$ can be expressed by

$$y_{k_l} = \underbrace{H_{k_l}^H w_k s_k}_{\text{intra-group signal}} + \underbrace{\sum_{m=1, m\neq k}^{G} H_{k_l}^H w_m s_m + \xi_{k_l}}_{\text{out of group interference}},$$

in which $\xi_{k_l} \sim CN(0, \sigma^2)$ is the circularly-symmetric complex Gaussian noise arising at $U_{k_l}$. Here, the noise variance is assumed to be the same for all the users. The case with different receiver noise level can be incorporated into our model by performing a scaling operation on the channel coefficient $h_{k_l}$.

In the receiver side, all users $U_{k_l}$ can detect the information intended for themselves from the uniquely decomposable signal $s_k$ by using our fast detection method, i.e., Algorithms 1 and 2 while treating the out of group interference as additive noise. However, in high SNR regimes, the out of group interference is the dominant term that limits the error performance of our system. In what follows, a novel transmission scheme is proposed so that the out of group interference is completely cancelled out by using the ZF philosophy while the intra-group interference contained in $s_k$ can be eliminated by taking advantage of the uniquely decomposable property of the sum-constellation.

B. Weighted Max-Min Fairness Grouped Transmission with ZF and Modulation Division

From (18), we know that the cochannel interference for $U_{k_l}$ consists of two parts: 1) the inter-group interference (i.e., interference originated from users in $S \setminus S_k$) and 2) the intra-group interference (i.e., interference due to users in $S_k \setminus \{U_{k_l}\}$). In our scheme, we use a ZF method to cancel the out of group interference, i.e.,

$$H_{k}^H w_k = 0, \forall k \in \{1, 2, \ldots, G\},$$

where $\bar{H}_k$ is defined in (17). Now, user $U_{k_l}$ only suffers from intra-group interference which can be also eliminated later by utilizing uniquely decomposable property. Under the ZF constraint (19), the SNR for the sum signal $s_k$ at user $U_{k_l}$ can be expressed by $\text{SNR}_{s_k} = \frac{|H_{k_l}^H w_k|^2}{\sigma^2}$, $\forall k, \ell$. Let us denote the total transmitted power for all the users in $S_k$ at BS as

$$P_k = \mathbb{E}[|w_k s_k|^2] = w_k^H w_k, \quad \forall k \in \{1, 2, \ldots, G\}. \quad (20)$$

Therefore, we aim to solve the following weighted max-min grouped beamforming optimization problem:
Problem 2: Find the beamforming vectors $w_k$ such that the worst case weighted received signal power is maximized, i.e.,

$$\max_{w_k,\forall k} \min_{\varrho_k} \varrho_k |h_{k_1}^H w_k|^2$$  \quad (21a)

subject to

$$\sum_{k=1}^{G} w_k^H w_k = P.$$  \quad (21b)

The quantity $\varrho_k |h_{k_1}^H w_k|^2$ in Problem 2 is usually called the weighted receiving signal power for $s_k$ at $U_k$. Using weighted SNR is a common method to balance the QoS among different users [72, 73]. The resulting SNR for each user is mainly determined by the SNR of the sum signal $s_k$ and the minimum Euclidean distance of the sum-constellation of user groups $S_k$, in this paper we choose $\varrho_k$ to be

$$\varrho_k = \frac{1}{E[\sum_{k=1}^{N_k} s_k^2]}$$  \quad (22)

which reasonably balances the minimum Euclidean distance for the sum signal $s_k$ in different user groups. Other choices of $\varrho_k$ are possible based on different application requirements.

In order to solve Problem 2 we first examine whether or not its feasible domain, $W = \{w = (w_1^T, w_2^T, \ldots, w_G^T)^T : H_k w_k = 0 \text{ and } \sum_{k=1}^{G} w_k^H w_k = P\}$ is empty, i.e., Problem 2 is feasible. This essentially checks whether constraint $H_k^H w_k = 0, k \in \{1, 2, \ldots, G\}$, can be satisfied. Since $H_k \in \mathbb{C}^{M \times (N-N_k)}$ has a rank of $N-N_k$, where $N_k \geq 1$, the constraint can be satisfied if $M \geq N-N_k, \forall k$. This condition is indeed satisfied, since we assume $N \leq M + 1$ in this paper. Therefore, Problem 2 is always feasible. On the other hand, we observe an important fact on the feasible domain. For any fixed $P_k, 0 < P_k \leq P$ for $k = 1, 2, \ldots, G$, if we let $W(P_1, P_2, \ldots, P_G) = \{(w_1^T, w_2^T, \ldots, w_G^T)^T : H_k w_k = 0 \text{ and } w_k^H w_k = P_k, k = 1, 2, \ldots, G\}$.

Since $W$ can be decomposed into a union of all such $W(P_1, P_2, \ldots, P_G)$, i.e., $W = \bigcup_{k=1}^{G} W(P_k, P_2, \ldots, P_G)$. Therefore, the original optimization Problem 2 can be equivalently split into the following two kinds of sub-optimization problems:

**Sub-problem 2.1:** For any fixed $P_k, 0 < P_k < P$, find the beamforming vectors $w_k, \forall k \in \{1, 2, \ldots, G\}$ such that

$$\zeta(P_k) = \max_{w_k} \min_{\forall k} |h_{k_1}^H w_k|^2$$  \quad (23a)

subject to $H_k w_k = 0$ and $w_k^H w_k = P_k$.  \quad (23b)

**Sub-problem 2.2:** Once sub-problem 2.1 has been solved, find an optimal power allocation strategy for all user groups $S$ such that

$$\max_{P_k, \forall k} \min_{\forall k} \varrho_k \zeta(P_k) \quad \text{s.t.} \quad \sum_{k=1}^{G} P_k = P.$$  \quad (24)

In general, the optimization problem 2 for arbitrary $N_k \geq 3, \forall k$ is hard to solve. However, since the power required for using a large sum-constellation $Q_k = \cup_{i=1}^{N_k} \mathcal{X}_{ki}$ with certain error target is huge if $N_k$ is too large, in this paper we primarily restrict ourself in the case with $N_k \leq 2$. In this case, we assume that $N \leq M + 1$. Let us consider first, where $P_k$ is temporarily regarded as a fixed number. For group $S_k$ with $N_k = 1$, by the Cauchy-Swarz inequality we have $w_k = \sqrt{P_k h_{k_1}^H}$ for user group $S_k$ with $N_k = 2$, the sub-optimization problem 2.1 can be reformulated as

$$\zeta(P_k) = \max_{\|w_k\|^2 = P_k} \min_{\forall k} \{w_k^H h_{k_1} h_{k_1}^H w_k, w_k^H h_{k_2} h_{k_2}^H w_k\}$$  \quad (25)

subject to $H_k^H w_k = 0, \forall k \in \{1, 2, \ldots, G\}$. Let us consider the constraint of (25) first. For $N_k = 2$, we have $H_k \in \mathbb{C}^{M \times (N-N)}$, which is a tall matrix of full column rank, since $N \leq M + 2$. This constraint essentially requires that $w_k$ lies in the orthogonal complement subspace of span$(H_k)$. Since $H_k$ has full column rank, the orthogonal complement projector for span$(H_k)$ is determined by $P_k = I - H_k (H_k^H H_k)^{-1} H_k^H \in \mathbb{C}^{M \times M}$, where $H_k^H P_k = 0$.

We know that the rank of $P_k$ is $(M - N + 2)$. Now we want to find an orthonormal basis for $w_k$. To do that, let the QR-decomposition of $P_k$ be $P_k = Q_k R_k$, where $Q_k \in \mathbb{C}^{M \times (M-N+2)}$ is a column-wise unitary matrix. If we let $w_k = Q_k w_k$, then problem (25) is equivalent to

$$\zeta(P_k) = \max_{\|w_k\|^2 = P_k} \min_{\forall k} \{w_k^H h_{k_1} h_{k_1}^H w_k, w_k^H h_{k_2} h_{k_2}^H w_k\}$$

where $H_k = Q_k^H h_{k_1}, \forall k, l$. The above optimization problem can be solved by using Theorem 3 with the optimal value, i.e., $\zeta(P_k)$, being linear in terms of $P_k$, i.e.,

$$\zeta(P_k) = P_k \zeta_k,$$  \quad (26)

in which $\zeta_k$ is determined by channel coefficients and is independent of $P_k$. Our next goal is to solve subproblem 2.2 in this case. Substituting (26) into (24) yields

$$\max_{P_k, \forall k} \min_{\forall k} P_k \varrho_k \zeta_k \quad \text{s.t.} \quad \sum_{k=1}^{G} P_k = P$$  \quad (27)

Its optimal value is attained when all $P_k \varrho_k \zeta_k$ for $k = 1, 2, \ldots, G$ are equal to each other, hence, leading to

$$P_{opt} = \frac{P \prod_{k=1}^{G} \varrho_k \zeta_k}{\sum_{m=1}^{G} \prod_{n=1}^{G} \varrho_{n,m} \zeta_{n,m}}, \forall k \in \{1, 2, \ldots, G\}.$$  \quad (28)

Thus far, we have solved the problem (21) for given grouping method $S$ with $N_k \leq 2, \forall k \in \{1, 2, \ldots, G\}$.

**C. User Grouping for $N_k \leq 2, \forall k \in \{1, 2, \ldots, G\}$**

As we have mentioned before, the performance of our proposed transmission method is closely related to the user grouping strategy. In this subsection, we consider the user grouping method for cases with $N_k \leq 2, \forall k \in \{1, 2, \ldots, G\}$. In these cases, we require that $G \leq N \leq 2G$ and as a consequence, there are $N-G$ groups with each having $2$ users and $2G-N$ groups with each having one user. For example, if $G = N$, each group has only one user. If $G = N/2$, each group has exactly two users. Since $S$ and $S_k$ are all unordered sets,
for the given number of groups $G$, we have $\frac{N^G - N!}{(N - G)!}$ different grouping methods if $G \leq N - 1$ and only one method if $G = N$. Since for $N_k \leq 2$, $\lceil \frac{N}{2} \rceil \leq G \leq N$, we have in total $1 + \sum_{N_k}^{\lceil \frac{N}{2} \rceil} \frac{N^G - N!}{(N - G)!}$ different grouping ways. For small $N$, the optimal grouping method can be found by brute-force search. However, it would be prohibitively complicated for large $N$, which makes our design hard to implement in practice. Therefore, we now propose a suboptimal user-grouping method to make trade-off between performance and complexity by setting a threshold $\gamma_T$, which is a predefined level to balance the error performance among different groups.

Example 2: Consider a two-user MISO BC with ZF beamforming, where each user employs a square $K$-ary QAM constellation and hence, the sum-rate of this network is $2 \log_2 K$. In contrast, for the modulation division method, the sum-constellation is set to be a $K$-ary QAM constellation such that the sum-rate is $\log_2 K^2 = 2 \log_2 K$, which is the same as ZF method. Assume that the average power of the transmitted symbol $x_k$ for the ZF method and that of $s_k$ for the sum-constellation are all unified to 1. Then the minimum Euclidean distance of the constellation points of $x_k$ is $d_{zd}(K) = \sqrt{\frac{6}{K - 1}}$ and that of $s_k$ is $d_{md}(K^2) = \sqrt{\frac{6}{K^2 - 1}}$. As a consequence, we can set $\gamma_T = 10 \log_{10} \left( \frac{d_{zd}(K)}{d_{md}(K^2)} \right) = 10 \log_{10} (K + 1)$ to compensate the SNR loss due to using a larger constellation. For example, every user is using a 4-QAM, we would expect $\gamma_T = 6.99$dB, $\gamma_T = 12.3$dB for 16-QAM, $\gamma_T = 18.13$dB for 64-QAM and $\gamma_T = 24.10$dB for 256-QAM. ■

Algorithm 3 (Grouping method): The grouping method for $N_k \leq 2, \forall k \in \{1, 2, \ldots, G\}$ is given as follows for $N$ receivers such that $N = \sum_{k=1}^{G} N_k$. There are $\left( \begin{array}{c} N \rule{0cm}{0.3cm} \\
2 \end{array} \right) = \frac{N(N-1)}{2}$ possible grouping methods for one group with exactly two users.

1) Enumeration: Find all the possible groups with two users and generate the grouping index.
2) Grouping gain calculation: For all the grouping indexes, calculate the coding gain. For example, suppose that users $m$ and $n$ ($m < n$) are grouped together with channel matrix $H_{m,n} = [h_m, h_n]$ and another matrix containing all the channel links of users in $U \setminus \{U_m, U_n\}$, which is denoted by $H_{m,n} = [h_1, h_2, \ldots, h_{m-1}, h_{m+1}, \ldots, h_{n-1}, h_{n+1}, \ldots, h_N]$. Let $P_{m,n} = I - H_{m,n}^H H_{m,n}^{-1} H_{m,n}$ and its QR-decomposition be $P_{m,n} = Q_{m,n} R_{m,n}$. Compute $\rho_{m,n} = \left\| h_m \right\|_2$ and $\cos \varphi_{m,n} = \left\| h_m^H h_n \right\|_2 / \left\| h_m \right\|_2 \left\| h_n \right\|_2$, where $h_m = Q_{m,n}^H h_m$ and $h_n = Q_{m,n}^H h_n$. Then, by Corollary 3 calculate the grouping gain $\nu_{m,n} = \nu(\rho_{m,n}, \cos \varphi_{m,n})$ as a function of $\rho_{m,n}, \cos \varphi_{m,n}$ if users $m$ and $n$ are grouped together.
3) Sorting: Now the $\frac{N(N-1)}{2}$ grouping gains $\nu_{m,n}$ are sorted in descending order, forming a vector $[\nu_{m_1, n_1}, \nu_{m_2, n_2}, \ldots, \nu_{m_{N(N-1)/2}, n_{N(N-1)/2}}]$.
4) Grouping: If $\nu_{m_1, n_1} > \gamma_T$, then, users $m_1$ and $n_1$ are grouped together. Otherwise, go to the next step and no users are grouped together. Then, consider $\nu_{m_2, n_2}$. Again, if $\nu_{m_2, n_2} \leq \gamma_T$, go to the next step and terminate the grouping procedure. Otherwise, if $\nu_{m_2, n_2} > \gamma_T$ and either $m_2$ or $n_2$ has not been grouped yet, then, $m_2$ and $n_2$ are grouped together. Repeat this process until all the $\nu_{m,n}$ have already been considered. The remaining users is left ungrouped.
5) Stop and output the grouping index.

In our model, the grouping operation is carried out at the BS and then, the grouping indexes are informed to all the receivers. In fact, each receiver $h_k$, only needs to know the grouping index $k_l$ to obtain the sum-constellation used by the group and the corresponding sub-constellation of itself.

V. COMPUTER SIMULATIONS AND DISCUSSIONS

In this section, computer simulations are carried out to verify our theoretical analysis and to assess the effectiveness of our proposed modulation division transmission method in a multiuser MISO BC. Throughout our simulations, we assume that the channel links from the BS are potentially correlated, but are uncorrelated between different users, i.e.,

$$E[h_k h_k^H] = \begin{cases} 0 & \forall k \neq \ell, \\ \Sigma & \text{otherwise} \end{cases}$$

(28)

In addition, to help with our simulation, we assume that the channel covariance matrix $\Sigma$ is the commonly-used non-symmetric Kac-Murdock-Szegö (KMS) matrix [74], the $(m,n)$-th entry of which is denoted by $\sigma(m-n)$, i.e.,

$$\sigma(m-n) = [\Sigma]_{mn} = \begin{cases} \rho^{m-n} & m \leq n, \\ \rho^{n-m} & m > n, \end{cases}$$

(29)

where $0 \leq |\rho| < 1$ indicates the degree of correlation. In particular, if $\rho = 0$, then $\Sigma = I$, i.e., all the entries of $H$ are

$^1$Then our method essentially degrade into the ZF method.
i.i.d. Gaussian. Under all these assumptions, we perform five kinds of simulations to test our proposed MD transmission scheme in terms of uncoded BER.

The first kind of simulations is to test our MD method for the two user case. Its error performance comparison with the ZF beamforming method and the time division (TD) method is plotted in Fig. 3 where our grouping method for two-user case is examined, with \( \rho = 0 \) and \( M = 2, 4, 6 \) transmitting antennas, \( N = 2 \) receivers. For the ZF method, each user uses 4-QAM and for the MD and TD methods, the sum constellation is 16-QAM. It can be observed that the BER performance of the proposed MD method is always better than those of the ZF beamforming and the TD methods. Specifically, the SNR gain at BER \( 10^{-4} \) is approximately 15dB for \( M = 2 \) and \( N = 2 \). However, as the number of transmitter antennas is increased to \( M = 4 \), the BER performance of MD is still better than that of ZF, but the gap between the two methods decreases. Particularly when \( M = 6 \), the performance of our proposed method is almost the same as that of ZF.

The second kind of simulations is to examine that how the correlations among the transmitter antennas affect the error performance of our MD method in the two user situation. The simulation results are shown in Fig. 4. It can be seen that the error performance gap between the MD and the ZF method becomes large with \(|\rho|\) increasing. In this case with mild correlation, e.g., \( \rho = 0.3 \exp(0.5j) \), the performance gain of our method is not observable. However, when the transmitter antennas are severely correlated, e.g., \( \rho = 0.9 \exp(0.5j) \), our method attains at least 1.5dB gain at BER \( 10^{-4} \) over the ZF scheme.

The third kind of simulations is to test our proposed suboptimal grouping method for the multi-user \( (N \geq 3) \) MISO BC, as shown in Fig. 5 where \( M = 10, N = 7, 8, 9, 10 \) and the channels are i.i.d. Rayleigh fading. It can be observed that the MD method with user grouping always outperforms the ZF.
method. Specifically for the case of $M = N = 10$, the SNR gain is approximately 20dB at BER $10^{-4}$. We find that the closer the number of users $N$ is to the number of transmitters $M$, the larger the performance gap between the proposed MD strategy and ZF method becomes, since when $M$ is close to $N$, there is a higher probability that the Hermitian angle between the channel vector of two users is small and as a consequence, the grouping gain becomes large. A similar conclusion can be drawn from the case with $M = 20$, $N = 14, 17, 19, 20$, as shown in Fig. 8.

Similar to the second kind of simulations, the fourth kind of simulations is to investigate how the channel correlations among the transmitter antennas affect the error performance of our proposed MD-based grouping scheme. The simulation results are shown in Figs. 4 and 5. In Fig. 4, we consider the case with $M = N = 20$ and different correlation coefficients.

It can be expected that the BER performance of our proposed method is much better than that of ZF method. In addition, it is not surprising that the BER performance becomes worse when the channel links from the BS becomes more correlated. Similar observations are also verified for the case with $M = 20$ and $N = 18$, as shown in Fig. 8.

The last kind of simulation is to compare our proposed MD method with other existing precoding methods in Figs. 9 and 10. Toward this end, in Fig. 9 we compare the average BER of the MD approach with SLNR based scheme in [28] and MMSE method with equal power allocation as well as TD and ZF methods when $M = 4$, $N = 4$. For the MD method, we consider both the proposed suboptimal grouping method described in Algorithm 3 and an exhaustive search grouping scheme that enumerates all the possible grouping methods for seeking the best possible max-min weighted SNR in Problem 2. It can be observed that the TD method has the worst BER in low and moderate SNR regimes. It can also be noticed that the SLNR and MMSE methods with equal power allocation have the same BER performance as proved by [75] and both of them outperform the ZF method. In addition, we can see that in a low SNR regime, the MMSE and SLNR methods have a lower BER than the MD method. However, in a moderate and high SNR regime, the MD method outperforms all the other methods in terms of BER. Despite the fact that it has better error performance than the proposed MD approach with the suboptimal grouping method, the exhaustive grouping MD method obtains the marginal BER gain and costs much higher computational complexity. Therefore, the proposed suboptimal grouping method is greatly desirable in practice. To further demonstrate the error performance comparison of our proposed MD method with other precoding schemes, the scenario with $M = 5$, $N = 4$ is given in Fig. 10 where similar conclusion can be drawn.

Here, it should be pointed out clearly that our optimal beamformer design, the optimal grouping scheme and all the resulting simulations are based on the prior assumption that
the BS has the knowledge of the perfect CSI. However, in practice, it is difficult to obtain the perfect knowledge of the CSI at BS. Therefore, it would be very necessary to analyze the error performance of our proposed algorithm with the imperfect CSI, especially for the case of multiuser massive MIMO BC. Unfortunately, this problem is too big and too important to have space for any investigation in this paper, but will be further studied in our future research.

In addition, it also should be mentioned explicitly that in terms of computational complexity, channel state information required and other overheads, there are some drawbacks in our proposed MD scheme when compared with the ZF scheme, as listed in Table II.

| Aspects                  | MD            | ZF            |
|--------------------------|---------------|---------------|
| Complexity               | $O(M^2)$      | $O(M^2)$      |
| CSI                      | CSIT and CSIR | CSIT          |
| Overhead                 | Grouping Index needed | No additional overhead |

### VI. CONCLUSION

In this paper, we have revealed an important property on PAM and QAM constellations for multiuser communications that any PAM constellation or QAM constellation of large size can be uniquely decomposed into the sum of a group of the scaled version of the PAM or QAM constellations of variety of small flexible sizes. In addition, we have developed two detection algorithms, showing that one of significant advantages of such unique decomposition is that once the sum signal is detected, each individual user signal can be efficiently decoded. Then, we consider a MISO BC with two users. For the special case with two receivers, the optimal beamforming vector is given in a closed-form based on a max-min criterion on the received SNR for the sum-signal. In addition, the SNR gain compared with ZF method is also given explicitly, which can be used to evaluate the effectiveness of our MD method in different channel coefficients. In the case with more than two receivers, a novel low-complexity grouping transmission scheme based on MD, which aims at improving the condition number of the channel matrix, is proposed, with each group having one or two users.

Finally, the simulation results have demonstrated that for the Rayleigh channel, if the number of the receivers is far less than the number of BS antennas, our method has the same error performance as the ZF method. However, when the number of users is very close to that of the BS antennas, the error performance of our proposed MD scheme is substantially better than ZF. Moreover, our computer simulations have also shown that when the transmitter antennas are correlated, our presented method is still better than ZF, even if the number of transmitter antennas is larger than that of the receivers.

In conclusion, our QAM MD transmission scheme can be considered as a feasible and concrete approach to the general NOMA method that introduces interference with proper power level superimposed on the desired signal and is possible to be applied to multiuser networks, which would enable a new promising multiple access method.

### APPENDIX

#### A. Proof of Algorithm 7

For any given $y = g + \xi$, the ML detection of $g$ is exactly equivalent to that of $p = g + \frac{K-1}{2} \in \{k\}_{k=0}^{2K-1}$ from $p + \xi$, which is given by

$$\hat{p} = \begin{cases} 0, & y + \frac{K-1}{2} \leq 0; \\ (y + \frac{K-1}{2} + \frac{1}{2})/2, & 0 < y + \frac{K-1}{2} \leq 2K - 1; \\ 2K - 1, & y + \frac{K-1}{2} > 2K - 1. \end{cases}$$

leading us to (1). For $K_i = 0$, (2) and (3) indeed hold. Therefore, we only consider the case $K_i \neq 0$. For presentation simplicity, we define $p_i$ as follows: 1) for any $x_1 \in X_1$, $p_1 = x_1 + \frac{K-1}{2}$; and 2) for any $x_1 \in X_1$, $p_1 = x_1 \times 2^{-\sum_{i=1}^{K_i} K_i + \frac{K-1}{2}}$. From the definitions of $X_i$, we can attain that $p_1 \in \{k\}_{k=0}^{2K_i-1}$. In addition, Theorem 1 tells us that given $\hat{g} \in G$ defined by (1), $\hat{g}$ can be uniquely decomposed into $\hat{g} = \sum_{i=1}^{N} \hat{x}_i$, where $\hat{x}_i \in X_i$ and thus, be equivalently rewritten into

$$\hat{g} = \sum_{i=1}^{N} \hat{x}_i = \hat{p}_1 + \sum_{i=2}^{N} \hat{p}_i \times 2^{\sum_{i=1}^{K_i} K_i - \frac{K-1}{2}}$$

where $\hat{p}_i \in \{k\}_{k=0}^{2K_i-1}$. It is noticed that for any $\hat{g} \in G = \{(\pm(k-1)/2)\}_{k=0}^{2K_i-1}$, we can have $\{\hat{g} + 2K_i - 1/2 : \hat{g} \in G\} = \{k\}_{k=0}^{2K_i-1}$. Then, letting $\hat{p} = \hat{g} + \frac{K-1}{2}$ produces an equivalent form of (30) as $\hat{p} = \hat{p}_1 + \sum_{i=2}^{N} \hat{p}_i \times 2^{\sum_{i=1}^{K_i} K_i}$, where $\hat{p}_i \in \{k\}_{k=0}^{2K_i-1}$. Now, we notice that $p_1 = \hat{p} \mod 2K_i$ because of the fact that $\sum_{i=2}^{N} \hat{p}_i \times 2^{\sum_{i=1}^{K_i} K_i} \mod 2K_i = 0$. This result gives us that $\hat{x}_1 = (\hat{g} + \frac{K-1}{2}) \mod 2K_i - \frac{K-1}{2}$, proving (2).

Also, we observe that $\frac{\hat{p}_1}{2} = \hat{p}_2 + \sum_{i=3}^{N} \hat{p}_i \times 2^{\sum_{i=1}^{K_i} K_i}$ and thus, attain that $\hat{p}_2 = \hat{x}_2 \mod 2K_i$, leading to $\hat{x}_2 = \left(\frac{\hat{p}_2 - 2^{K_i - 1/2}}{2}\right) \mod 2K_i$. Following this process, we can attain that for $2 \leq i \leq N$, $\hat{p}_i = \frac{\hat{x}_i \times 2^{-\sum_{i=1}^{K_i} K_i - \frac{K-1}{2}} \mod 2K_i}$ allows us to arrive at the desired (3).

Therefore, this verifies Algorithm 1.

#### B. Proof of Theorem 3

**Scenario 1:** $h_1 = \tau h_2, \tau \in \mathbb{C}$. If $|\tau| > 1$, then, we have $w^H h_1 h_2 h_2^H w - w^H h_2 h_2^H w = (|\tau|^2 - 1)w^H h_1 h_2 h_2^H w \geq 0$ for any $w^H w = P$ and as a result, the original optimization problem degrades into $\max_{|w|^2 = P} \Re{w^H h_1 h_2^H w}$. Using Cauchy-Schwarz inequality, $w^{\text{opt}} = \frac{\tau h_2}{\|h_2\|}$ (or $w^{\text{opt}} = \frac{\sqrt{\tau} h_1}{\|h_1\|}$). The case when $|\tau| < 1$ is similar and we omit it here.

**Scenario 2:** $h_1 \neq \tau h_2, \forall \tau \in \mathbb{C}$. In this case, the overall optimization problem can be divided into the following two problems by introducing restrictions on the original feasible region and the global optimum is the maximum of the two.
• **Case 1:** SNR of user 1 is not worse than that of user 2

\[
\max_w w^H h_2^H w \quad \text{s.t. } w^H A w \geq 0 \text{ and } w^H w = P
\]

\[
\max_w w^H h_1^H w \quad \text{s.t. } w^H A w < 0 \text{ and } w^H w = P
\]

Let us consider **Case 1** first. Using the notation in [7], the optimization problem can be reformulated as

\[
\max \tilde{w}^H \tilde{h}_2^H \tilde{w} \quad \text{s.t. } \lambda_1 |\tilde{w}_1|^2 \geq \lambda_2 |\tilde{w}_2|^2 \text{ and } \sum_{\ell=1}^M |\tilde{w}_\ell|^2 = P
\]

We have the following two observations on the above optimization problem:

1) From the objective function \( \tilde{w}^H \tilde{h}_2^H \tilde{w} = |\sum_{\ell=1}^M \tilde{w}_\ell \tilde{h}_{2,\ell}|^2 \), the constraint, one optimal choice of the angle of \( \tilde{w}_\ell \) is \( \arg(\tilde{w}_\ell) = \arg(\tilde{h}_{2,\ell}), \forall \ell \).

2) Since \( \tilde{h}_2 = [\tilde{h}_{2,1}, \tilde{h}_{2,2}, 0, \ldots, 0]^T \), we should let \( |\tilde{w}_1| = 0, \forall \ell \geq 3 \). Otherwise we could let \( |\tilde{w}_1| = 0, \forall \ell \geq 3 \) and increase \( |\tilde{w}_1| \) and \( |\tilde{w}_2| \) slightly without violating the constraint, resulting in an increased objective function.

Hence, the optimization problem can be simplified into

\[
\max_{\tilde{w}_1} |\tilde{w}_1| |\tilde{h}_{2,1}| + \sqrt{P - |\tilde{w}_1|^2} |\tilde{h}_{2,2}| \quad \text{s.t. } \sqrt{P^2 - |\tilde{w}_1|^2} \leq |\tilde{w}_1| \leq \sqrt{P}
\]

Therefore, the solution to problem (34a) is given below:

1) \( 0 \leq \sin \theta \leq \frac{\lambda_1}{\lambda_2} \) (i.e., \( 0 \leq \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}} \cos \theta \leq \sqrt{P/2} \)), then

\[
g_1(|\tilde{w}_1|) = \begin{cases} 0 & |\tilde{w}_1| = \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}} \cos \theta \\ 0 \leq |\tilde{w}_1| < \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}} \cos \theta & |\tilde{w}_1| \leq 0 \end{cases}
\]

2) \( \frac{\lambda_1}{\lambda_2} < \sin \theta \leq 1 \) (i.e., \( \sqrt{P/2}, \lambda_1 + \lambda_2} \)), then

\[
g_1(|\tilde{w}_1|) = \begin{cases} 0 & |\tilde{w}_1| = \sqrt{\frac{P_{\lambda_2}}{\lambda_2} + \sin \theta} \cos \theta \\ 0 \leq |\tilde{w}_1| < \sqrt{\frac{P_{\lambda_2}}{\lambda_2} + \sin \theta} \cos \theta & |\tilde{w}_1| \leq 0 \end{cases}
\]

For **Case 2**, by a similar argument on the objective function \( \tilde{w}^H \tilde{h}_1^H \tilde{w} = |\sum_{\ell=1}^M \tilde{w}_\ell \tilde{h}_{1,\ell}|^2 \), we have \( \arg(\tilde{h}_{1,\ell}), \forall \ell \), and \( |\tilde{w}_\ell| = 0, \forall \ell \geq 3 \). The resulting optimization problem can be reformulated as

\[
\max_{\tilde{w}_1} |\tilde{w}_1| |\tilde{h}_{1,1}| + \sqrt{P - |\tilde{w}_1|^2} |\tilde{h}_{1,2}| \quad \text{s.t. } 0 \leq |\tilde{w}_1| \leq \sqrt{\frac{P_{\lambda_2}}{\lambda_1 + \lambda_2}}
\]

With the aid of (10), we can represent the objective function as \( g_2(|\tilde{w}_1|) = |\tilde{w}_1| \sqrt{\lambda_1 \sin \theta + \sqrt{P - |\tilde{w}_1|^2}} \lambda_2 \tan \theta \).

Since its first order derivative is \( g_2'(|\tilde{w}_1|) = \lambda_1 \sec \theta - \frac{|\tilde{w}_1|}{\sqrt{P - |\tilde{w}_1|^2}} \lambda_2 \tan \theta \), we have

\[
\begin{align*}
&g_2'(|\tilde{w}_1|) = \left\{ \begin{array}{ll}
0 & |\tilde{w}_1| = \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}} \\
> 0 & 0 \leq |\tilde{w}_1| < \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}} \\
< 0 & \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}} < |\tilde{w}_1| \leq 0
\end{array} \right.
\end{align*}
\]

Therefore, the solution to (35) can be determined as follows:

1) \( 0 \leq \sin \theta \leq \frac{\lambda_1}{\lambda_2} \) (i.e., \( \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}} \leq \sqrt{P} \)), then

\[
\max_{\tilde{w}_1} 0 \leq |\tilde{w}_1| \leq \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}}
\]

2) \( \frac{\lambda_1}{\lambda_2} \leq \sin \theta \leq 1 \) (i.e., \( 0 < \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}} \leq \sqrt{P} \)), then

\[
\max_{\tilde{w}_1} 0 \leq |\tilde{w}_1| \leq \sqrt{\frac{P_{\lambda_1}}{\lambda_1 + \lambda_2}}
\]

The overall maximum value of the original problem is the maximum of the two cases.

### C. Proof of Lemma

Let the singular value decomposition (SVD) of \( H = U_1 \Sigma_1 V_1^H \), where \( U_1 \in \mathbb{C}^{M \times 2} \) is a tall-column unitary matrix, \( V_1 = [v_1, v_2] \in \mathbb{C}^{2 \times 2} \) is a unitary matrix, and \( \Sigma_1 = \mathrm{diag}(\sqrt{\mu_1}, \sqrt{\mu_2}) \). Then, we have \( |h_1 - h_2| = U_1 \Sigma_1 \begin{bmatrix} v_{1,1} & v_{1,2} \\ v_{2,1} & v_{2,2} \end{bmatrix} \), and thus, equation (5) can be represented by \( A = [h_1 - h_2]^T h_2^H = U_1 B U_1^H \), where

\[
B = \Sigma_1 \begin{bmatrix} |v_{1,1}|^2 - |v_{1,2}|^2 & v_{1,1} v_{1,2}^* - v_{1,2} v_{1,1}^* \\ v_{2,1} v_{2,2}^* - v_{2,2} v_{2,1}^* & |v_{1,2}|^2 - |v_{2,2}|^2 \end{bmatrix} \Sigma_1.
\]

Notice that matrix \( A \) and \( B \) have the same non-zero eigenvalues, i.e., \( \text{diag}(\lambda_1 - \lambda_2) = \text{eig}(B) \). Hence, in order to use \( a, b \) and \( c \) to represent \( \lambda_1, \lambda_2, \mu_1 \) and \( \mu_2 \), we need to calculate \( V_1 \). Since \( |h_1|^2 = a, |h_2|^2 = b \) and \( |h_1^H h_2|^2 = c \), we can use the eigenvector of \( h_2^H h_2 \) to \( c \), for instance, we have

\[
H^H H = V_1^H \Sigma_1^H \Sigma_1 V_1^H = \begin{bmatrix} a & c e^{j\phi_c} \\ c e^{-j\phi_c} & b \end{bmatrix}.
\]

Hence, \( V_1 \) is the eigenvector matrix of \( H^H H \) and \( \mu_1 \) and \( \mu_2 \) are its eigenvalues, which must satisfy the following characteristic equation,

\[
\begin{align*}
&\begin{bmatrix} a - \mu_i & c e^{j\phi_c} \\ c e^{-j\phi_c} & b - \mu_i \end{bmatrix} = 0, \text{ for } i = 1, 2.
\end{align*}
\]

Therefore, we have \( \mu_1 = \frac{a + b + \sqrt{(a - b)^2 + 4c^2}}{2}, \mu_2 = \frac{a + b - \sqrt{(a - b)^2 + 4c^2}}{2} \).
Correspondingly, the two column vectors of \( \mathbf{v}_1 = [v_1, v_2] \) must satisfy the following equations:
\[
\begin{align*}
\begin{bmatrix}
a - \mu_1 c e^{j\phi_c} \\
b - \mu_1 e^{-j\phi_c}
\end{bmatrix}
\end{align*}
\]
\[
\begin{align*}
\begin{bmatrix}
a - \mu_2 c e^{j\phi_c} \\
b - \mu_2 e^{-j\phi_c}
\end{bmatrix}
\end{align*}
\]
\( \mathbf{v}_1 = 0, \quad \mathbf{v}_2 = 0. \)

Let \( d = \sqrt{(a-b)^2 + 4c^2}. \) Then, we have
\[
\begin{align*}
v_1 &= \begin{bmatrix}
\sqrt{2d^2 - 2(a-b)d} & a - b - d \\
\sqrt{2d^2 - 2(a-b)d} & a - b + d
\end{bmatrix}^T, \\
v_2 &= \begin{bmatrix}
\sqrt{2d^2 + 2(a-b)d} & a - b + d \\
\sqrt{2d^2 + 2(a-b)d} & a - b - d
\end{bmatrix}^T,
\end{align*}
\]
and \( \mu_1 = a+b+d, \quad \mu_2 = a+b-d \) and \( \sqrt{\mu_1 \mu_2} = \sqrt{ab-c^2}. \) Hence, (36) can be expressed by
\[
B = \frac{1}{2d} \begin{bmatrix}
(a-b)(a+b+d) & 4c\sqrt{ab-c^2} \\
4c\sqrt{ab-c^2} & -(a-b)(a+b-d)
\end{bmatrix}.
\]

On the other hand, the eigenvalues of \( \mathbf{A} \), i.e., \( \lambda_1, -\lambda_2 \) must satisfy the following characteristic equation:
\[
\begin{align*}
\begin{vmatrix}
(a-b)(a+b+d) & 4c\sqrt{ab-c^2} \\
4c\sqrt{ab-c^2} & -(a-b)(a+b-d)
\end{vmatrix} - \lambda &= 0.
\end{align*}
\]
Solving this equation and after some algebraic manipulations, we attain \( \lambda_1 = \frac{a-b\sqrt{(a+b)^2-4c^2}}{2} \) and \( \lambda_2 = \frac{-a+b\sqrt{(a+b)^2-4c^2}}{2}. \)

This completes the proof of Lemma [1].

REFERENCES

[1] T. Cover, “Broadcast channels,” IEEE Trans. Inf. Theory, vol. 18, pp. 2–14, Jan. 1972.
[2] P. Bergmans, “Random coding theorem for broadcast channels with degraded components,” IEEE Trans. Inf. Theory, vol. 19, pp. 197–207, Mar. 1973.
[3] R. G. Gallager, “Coding and capacity for degraded broadcast channels,” Probl. Peredachi Inf., vol. 10, pp. 185–193, Oct. 1974.
[4] E. van der Meulen, “A survey of multi-way channels in information theory: 1961-1976,” IEEE Trans. Inf. Theory, vol. 23, pp. 1–37, Jan. 1977.
[5] T. Cover, “Comments on broadcast channels,” IEEE Trans. Inf. Theory, vol. 44, pp. 2524–2530, Oct. 1998.
[6] A. E. Gamal and Y.-H. Kim, Network Information Theory. Cambridge Univ. Press, 2012.
[7] I. Telatar, “Capacity of multiple antenna Gaussian channels,” Europ. Trans. Telecommun., vol. 10, pp. 585–595, Nov.-Dec. 1999.
[8] G. Foschini and M. Gans, “On limits of wireless communications in a fading environment when using multiple antenna,” Wireless Personal Commun., vol. 6, pp. 311–335, Mar. 1998.
[9] Caire and S. Shamai, “The capacity region of the Gaussian multiple-input multiple-output broadcast channel,” IEEE Trans. Inf. Theory, vol. 52, pp. 1912–1921, Aug. 2003.
[10] W. Yu and J. Cioffi, “Sum capacity of Gaussian vector broadcast channels,” IEEE Trans. Inf. Theory, vol. 50, pp. 1875–1892, Sep. 2004.
[11] H. Weingarten, Y. Steinberg, and S. Shamai, “The capacity region of the Gaussian multiple-input multiple-output broadcast channel,” IEEE Trans. Inf. Theory, vol. 52, pp. 3936–3964, Sep. 2006.
[12] N. Deshpande and B. S. Rajan, “Constellation constrained capacity of two-user broadcast channels,” in Proc. IEEE Global Commun. Conf. (GLOBECOM’09), pp. 1–6, Nov. 2009.
[41] C. Masouros and T. Ratnarajah, “Interference as a source of green signal power in cognitive relay assisted co-existing MIMO wireless transmissions,” IEEE Trans. Commun., vol. 60, pp. 525–536, Feb. 2012.

[42] Y. Zeng and R. Zhang, “Optimized training design for wireless energy transfer,” IEEE Trans. Commun., vol. 63, pp. 536–550, Feb. 2015.

[43] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, “Wireless networks with RF energy harvesting: A contemporary survey,” IEEE Commun. Surveys Tuts., vol. 17, no. 2, pp. 757–789, 2015.

[44] I. Krikidis, S. Timotheou, S. Nikolau, G. Zheng, D. W. K. Ng, and R. Schober, “Simultaneous wireless information and power transfer in modern communication systems,” IEEE Commun. Mag., vol. 52, pp. 104–110, Nov. 2014.

[45] D. W. K. Ng, E. S. Lo, and R. Schober, “Wireless information and power transfer: Energy efficiency optimization in OFDMA systems,” IEEE Trans. Wireless Commun., vol. 12, pp. 6532–6570, Dec. 2013.

[46] R. Zhang and C. K. Ho, “MIMO broadcasting for simultaneous wireless information and power transfer,” IEEE Trans. Wireless Commun., vol. 12, pp. 1989–2001, May 2013.

[47] K. Huang and V. K. N. Lau, “Enabling wireless power transfer in cellular networks: Architecture, modeling and deployment,” IEEE Trans. Wireless Commun., vol. 13, pp. 902–912, Feb. 2014.

[48] Z. Ding, I. Krikidis, B. Sharif, and H. V. Poor, “Wireless information and power transfer in cooperative networks with spatially random relays,” IEEE Trans. Wireless Commun., vol. 13, pp. 4440–4453, Aug. 2014.

[49] E. Schimmerling, “Decoding for a multiple-access channel,” in Proc. IEEE Int. Symp. Inf. Theory, pp. 104–110, Nov. 2014.

[50] Y. Li, H. Murata, and S. Yoshida, “Coding for multi-user detection in MIMO systems,” in Proc. IEEE Int. Symp. Inf. Theory, pp. 525–529, Jun. 2004.

[51] W. Zhang, C. D’Amours, and A. Yongacoglu, “Trellis coded modulation design for multi-user systems on AWGN channels,” in Proc. IEEE 59th Vehi. Tech. Conf. (VTC Spring’04), vol. 3, pp. 1722–1726 Vol.3, May 2004.

[52] K. Ramchandran, A. Ortega, K. U., and M. Vetterli, “Multisolution broadcast for digital HDTV using joint source/channel coding,” IEEE J. Sel. Areas Commun., vol. 11, no. 5, pp. 1898–1904, Sep. 1993.

[53] Y. Li, H. Murata, and S. Yoshida, “Coding for multi-user detection in interference channel,” in Proc. IEEE Global Commun. Conf. (GLOBECOM’98), vol. 6, pp. 1596–1601 vol.6, 1998.

[54] W. Zhang, C. D’Amours, and A. Yongacoglu, “Trellis coded modulation design for multi-user systems on AWGN channels,” in Proc. IEEE 59th Vehi. Tech. Conf. (VTC Spring’04), vol. 3, pp. 1722–1726 Vol.3, May 2004.

[55] Y. Li, H. Murata, and S. Yoshida, “Coding for multi-user detection in interference channel,” in Proc. IEEE Global Commun. Conf. (GLOBECOM’98), vol. 6, pp. 1596–1601 vol.6, 1998.

[56] X. Wang and Y. Li, “Coding for multi-user detection in interference channel,” in Proc. IEEE Global Commun. Conf. (GLOBECOM’98), vol. 6, pp. 1596–1601 vol.6, 1998.

[57] K. Ramchandran, A. Ortega, K. U., and M. Vetterli, “Multisolution broadcast for digital HDTV using joint source/channel coding,” IEEE J. Sel. Areas Commun., vol. 11, no. 5, pp. 1898–1904, Sep. 1993.

[58] Y. Li, H. Murata, and S. Yoshida, “Coding for multi-user detection in interference channel,” in Proc. IEEE Global Commun. Conf. (GLOBECOM’98), vol. 6, pp. 1596–1601 vol.6, 1998.

[59] W. Zhang, C. D’Amours, and A. Yongacoglu, “Trellis coded modulation design for multi-user systems on AWGN channels,” in Proc. IEEE 59th Vehi. Tech. Conf. (VTC Spring’04), vol. 3, pp. 1722–1726 Vol.3, May 2004.

[60] K. Ramchandran, A. Ortega, K. U., and M. Vetterli, “Multisolution broadcast for digital HDTV using joint source/channel coding,” IEEE J. Sel. Areas Commun., vol. 11, no. 5, pp. 1898–1904, Sep. 1993.

[61] X. Wang and Y. Li, “Coding for multi-user detection in interference channel,” in Proc. IEEE Global Commun. Conf. (GLOBECOM’98), vol. 6, pp. 1596–1601 vol.6, 1998.

[62] K. Ramchandran, A. Ortega, K. U., and M. Vetterli, “Multisolution broadcast for digital HDTV using joint source/channel coding,” IEEE J. Sel. Areas Commun., vol. 11, no. 5, pp. 1898–1904, Sep. 1993.

[63] K. Ramchandran, A. Ortega, K. U., and M. Vetterli, “Multisolution broadcast for digital HDTV using joint source/channel coding,” IEEE J. Sel. Areas Commun., vol. 11, no. 5, pp. 1898–1904, Sep. 1993.

[64] X. Wang and Y. Li, “Coding for multi-user detection in interference channel,” in Proc. IEEE Global Commun. Conf. (GLOBECOM’98), vol. 6, pp. 1596–1601 vol.6, 1998.

[65] W. Zhang, C. D’Amours, and A. Yongacoglu, “Trellis coded modulation design for multi-user systems on AWGN channels,” in Proc. IEEE 59th Vehi. Tech. Conf. (VTC Spring’04), vol. 3, pp. 1722–1726 Vol.3, May 2004.

[66] K. Ramchandran, A. Ortega, K. U., and M. Vetterli, “Multisolution broadcast for digital HDTV using joint source/channel coding,” IEEE J. Sel. Areas Commun., vol. 11, no. 5, pp. 1898–1904, Sep. 1993.

[67] X. Wang and Y. Li, “Coding for multi-user detection in interference channel,” in Proc. IEEE Global Commun. Conf. (GLOBECOM’98), vol. 6, pp. 1596–1601 vol.6, 1998.