Multiple Beamforming with Constellation Precoding: Diversity Analysis and Sphere Decoding

Hong Ju Park, Boyu Li, and Ender Ayanoglu

Center for Pervasive Communications and Computing
Department of Electrical Engineering and Computer Science
University of California, Irvine
Email: hjpark@uci.edu, boyul@uci.edu, ayanoglu@uci.edu

***Abstract*—Multi-Input Multi-Output (MIMO) wireless communication systems commonly employ beamforming techniques with Singular Value Decomposition (SVD). In such systems, if no channel encoding is employed, the full diversity order provided by the channel is achieved when a single symbol is transmitted over multiple channels; however, this property is lost whenever multiple symbols are simultaneously transmitted. The full diversity order can be restored when channel coding is added to such a system. For example, when Bit-Interleaved Coded Modulation (BICM) is combined with this technique, the full diversity order of NM in an $M \times N$ MIMO channel, transmitting S parallel streams is possible; provided $SR_c \leq 1$ where R_C is the BICM convolutional code rate. In this paper, we present multiple beamforming with constellation precoding which can achieve the full diversity order with both uncoded and BICM-coded SVD systems. An analytical proof of this property is provided. In addition, to reduce the computational complexity of Maximum Likelihood (ML) decoding, we introduce a Sphere Decoding (SD) technique. This technique achieves several orders of magnitude reduction in computational complexity not only with respect to conventional ML decoding, but also, with respect to conventional SD.**

***Index Terms*—MIMO systems, SVD, BICMB, constellation precoding, sphere decoding.**

I. Introduction

Beamforming is used to achieve spatial multiplexing and thereby increase the data rate, or to enhance the performance of a Multiple-Input Multiple-Output (MIMO) system when perfect channel state information is available at the transmitter [1]. For various design criteria, beamforming vectors are designed in [2], [3]. These vectors can be obtained by Singular Value Decomposition (SVD), leading to a channel-diagonalizing structure optimum in the sense of minimizing the average Bit Error Rate (BER) [3]. It was shown that Uncoded Single Beamforming (SB), which carries only one symbol at a time, achieves the full diversity order of NM where N and M are the number of transmit and receive antennas, respectively [4], [5]. However, uncoded multiple beamforming, which increases the throughput by sending multiple symbols at a time, has the diversity order of $(N-S+1)(M-S+1)$ where the symbols are transmitted on the subchannels with the largest S singular values. Although it increases the throughput, this system cannot achieve the full diversity order over a flat fading channel [4], [5]. Whereas, it is desirable to come up with a system that achieves both maximum diversity order and maximum spatial multiplexing provided by the channel.

An SVD subchannel with larger singular value provides larger diversity gain [5]. Similarly, when symbols are simultaneously transmitted in parallel on the diagonalized subchannels, the performance at high Signal-to-Noise Ratio (SNR) is dominated by the subchannel with the smallest singular value. To overcome this degradation of the diversity order in multiple beamforming, Bit-Interleaved Coded Multiple Beamforming (BICMB) was proposed [6], [7]. BICMB interleaves the codewords through the multiple subchannels with different diversity order, resulting in a better diversity order overall. Although it is a form of multiple beamforming, BICMB can achieve the full diversity order offered by the channel as long as the code rate R_c and the number of employed subchannels S satisfy the condition $R_cS \leq 1$ [8].

In this paper, we present a multiple beamforming technique that achieves the full diversity order in both of the coded and the uncoded systems. This technique employs the constellation precoding scheme [9], [10], [11], [12], [13], which is used for space-time or space-frequency block codes to increase the system data rate without losing the full diversity order. We show via Pairwise Error Probability (PEP) analysis that Fully Precoded Multiple Beamforming (FPMB) with Maximum Likelihood (ML) detection achieves the full diversity order even in the absence of any channel coding. We also present the diversity analysis of Bit-Interleaved Coded Multiple Beamforming with Constellation Precoding (BICMB-CP), which adds the constellation precoding stage to BICMB. We show that the addition of the constellation precoder to BICMB removes the requirement for BICMB that $R_cS \leq 1$ for full diversity, when the subchannels for the precoded symbols are properly chosen. Simulation results are provided to verify the analysis.

Multiple beamforming without constellation precoding separates the MIMO channel into independent parallel subchannels, enabling symbol-by-symbol detection on each subchannel. However, when a precoder is employed, this property is lost and the the parallel independent detection of the symbols on each subchannel is no longer possible. As a result, one needs to resort to ML detection for precoded symbols. On the other hand, the complexity of ML detection increases exponentially with the number of possible constellation points of the modulation scheme and the dimension of the constellation precoder. This complexity increase makes the receiver with ML detection unsuitable for practical purposes [14]. It is known that employing Sphere Decoding (SD) as an alternative for ML detection provides optimal performance with reduced computational complexity [15].

Furthermore, a number of complexity reduction techniques for SD have been proposed. For example, in [16] and [17], attention is drawn to the initial radius selection strategy, since an inappropriate initial radius can result in either a large number of lattice points to be searched, or a number of restarted searches with increased initial radius. In [18] and [19], the complexity is reduced by making a proper choice to update the sphere radius. Other methods, such as the K-best lattice decoder [20], [21], and a combination of SD and K-best decoder [22], can significantly reduce the complexity of low SNR at the cost of BER performance.

In this paper, we introduce an SD algorithm which efficiently improves the complexity of constellation precoded multiple beamforming over the flat fading channel by reducing the average number of multiplications required to obtain the optimal solution. This complexity reduction is accomplished by precalculating the multiplications at the beginning of decoding, and recycling them later for the repetitive calculations. Further reduction is achieved by using the lattice representation of our previous work presented in [23]. This representation introduces orthogonality between the real and imaginary parts of every detected symbol. Furthermore, we employ Zero-Forcing Decision Feedback Equalization (ZF-DFE), to determine the initial radius. This new technique reduces the average number of real multiplications needed to acquire one precoded bit metric for BICMB-CP. We illustrate by means of simulations that conventional SD reduces the complexity substantially compared with the exhaustive search, and the complexity can be further reduced effectively by our proposed SD. The complexity reduction increases as the constellation precoder dimension and the constellation size become larger.

Notation: Bold lower (upper) case letters denote vectors (matrices). The set of symbols diag$[\mathbf{B}_1, \cdots, \mathbf{B}_P]$ stands for a block diagonal matrix with matrices $\mathbf{B}_1, \cdots, \mathbf{B}_P$, and diag$[b_1, \cdots, b_P]$ is a diagonal matrix with diagonal entries $b_1, \cdots, b_P$. The symbols $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary part of a complex number, respectively. The superscripts $(\cdot)^H$, $(\cdot)^T$, $(\cdot)^*$, $\bar{(\cdot)}$ stand for conjugate transpose, transpose, complex conjugate, binary complement, respectively, and the symbol $\forall$ denotes "for all." The function $\lceil\cdot\rceil$ is the ceiling function that maps a real number to the next largest integer. The symbols $\mathbb{R}^+$ and $\mathbb{C}$ stand for the set of positive real numbers and the complex numbers, respectively. Finally, the symbol d_{min} represents the minimum Euclidean distance between two points in a constellation.

II. System Model

A. Uncoded Multiple Beamforming with Constellation Precoding

We introduce Uncoded Multiple Beamforming with Constellation Precoding (UMB-CP) as a system that transforms modulated symbols to precoded symbols via a precoding matrix. In this system, the $S \times 1$ symbol vector $\mathbf{x}$, where $S \leq \min(N, M)$, is precoded by a square matrix $\mathbf{\Theta}$. The elements of $\mathbf{x}$ belong to a signal set $\chi \subset \mathbb{C}$ of size $|\chi| = 2^m$, such as 2^m-QAM, where m is the number of input bits to the Gray encoder. We specify the precoder as

$$\mathbf{\Theta} = \begin{bmatrix} \tilde{\mathbf{\Theta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{S-P} \end{bmatrix} \tag{1}$$

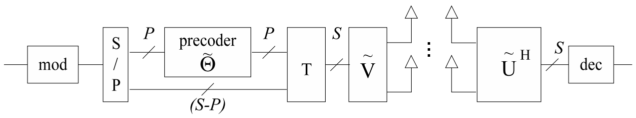
where $\tilde{\mathbf{\Theta}}$ is a $P \times P$ constellation precoding matrix that precodes the first P modulated symbols of the vector $\mathbf{x}$. When all of the S modulated symbols are precoded ($P = S$), we call the resulting system Fully Precoded Multiple Beamforming (FPMB), otherwise, we call it Partially Precoded Multiple Beamforming (PPMB). The permutation matrix $\mathbf{T}$ reorders the precoded P symbols and non-precoded $S - P$ symbols to be transmitted on the predefined subchannels created by the SVD of the MIMO channels. We define $\boldsymbol{\eta} = [\eta_1 \cdots \eta_P]$ as a vector whose element η_p is the index of the subchannel on which the precoded symbols are transmitted, and ordered increasingly such that $\eta_p < \eta_q$ for $p < q$. The vector $\boldsymbol{\omega} = [\omega_1 \cdots \omega_{(S-P)}]$ is defined in the same way as an increasingly ordered vector whose elements are the indices of the subchannels which carry the non-precoded symbols.

The MIMO channel $\mathbf{H} \in \mathbb{C}^{M \times N}$ is assumed to be quasi-static, Rayleigh, and flat fading, and perfectly known to both the transmitter and the receiver. The beamforming matrices are determined by the SVD of the MIMO channel, i.e., $\mathbf{H} = \mathbf{U\Lambda V^H}$ where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, and $\mathbf{\Lambda}$ is a diagonal matrix whose s^{th} diagonal element, $\lambda_s \in \mathbb{R}^+$, is a singular value of $\mathbf{H}$ in decreasing order. When S symbols are transmitted at the same time, then the first S vectors of $\mathbf{U}$ and $\mathbf{V}$ are chosen to be used as beamforming matrices at the receiver and the transmitter, respectively. Fig. 1(a) displays the structure of UMB-CP. In this figure, $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ denote the beamforming matrices picked from $\mathbf{U}$ and $\mathbf{V}$.

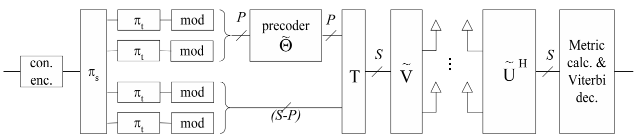
The serial-to-parallel converter organizes the symbol vector $\mathbf{x}$ as $\mathbf{x} = [\mathbf{x}_{\boldsymbol{\eta}}^T \vdots \mathbf{x}_{\boldsymbol{\omega}}^T]^T = [x_{\eta_1} \cdots x_{\eta_P} \vdots\ x_{\omega_1} \cdots x_{\omega_{(S-P)}}]^T$, where $\mathbf{x}_{\boldsymbol{\eta}}$ and $\mathbf{x}_{\boldsymbol{\omega}}$ consist of the modulated entries to be transmitted on the subchannels specified in $\boldsymbol{\eta}$ and $\boldsymbol{\omega}$, respectively. The $S \times 1$ detected symbol vector $\mathbf{y} = [\mathbf{y}_p^T \vdots \mathbf{y}_n^T]^T = [y_1 \cdots y_P \vdots\ y_{P+1} \cdots y_S]^T$ at the receiver is

$$\mathbf{y} = \mathbf{\Gamma\Theta x} + \mathbf{n} \tag{2}$$

where $\mathbf{\Gamma}$ is a block diagonal matrix, $\mathbf{\Gamma} = $ diag$[\mathbf{\Gamma}_p, \mathbf{\Gamma}_n]$, with diagonal matrices defined as $\mathbf{\Gamma}_p = $ diag$[\lambda_{\eta_1}, \cdots, \lambda_{\eta_P}]$, $\mathbf{\Gamma}_n = $ diag$[\lambda_{\omega_1}, \cdots, \lambda_{\omega_{(S-P)}}]$, and $\mathbf{n} = [\mathbf{n}_p^T \vdots \mathbf{n}_n^T]^T$ is an additive white Gaussian noise vector with zero mean and



(a) Uncoded Multiple Beamforming with Constellation Precoding.



(b) Bit-Interleaved Coded Multiple Beamforming with Constellation Precoding.

Fig. 1. Structure of Constellation Precoded Multiple Beamforming.

variance $N_0 = N/SNR$. The matrix $\mathbf{H}$ is complex Gaussian with zero mean and unit variance. To make the received signal-to-noise ratio SNR, the total transmitted power is scaled as N. The input-output relation in (2) is decomposed into

$$\begin{aligned} \mathbf{y}_p &= \mathbf{\Gamma}_p \tilde{\mathbf{\Theta}} \mathbf{x}_{\boldsymbol{\eta}} + \mathbf{n}_p \\ \mathbf{y}_n &= \mathbf{\Gamma}_n \mathbf{x}_{\boldsymbol{\omega}} + \mathbf{n}_n. \end{aligned} \tag{3}$$

The ML decoding of the detected symbol $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_{\boldsymbol{\eta}}^T \vdots \hat{\mathbf{x}}_{\boldsymbol{\omega}}^T]^T = [\hat{x}_{\eta_1} \cdots \hat{x}_{\eta_P} \vdots \hat{x}_{\omega_1} \cdots \hat{x}_{\omega_{(S-P)}}]^T$ is given by

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \chi^S} \|\mathbf{y} - \mathbf{\Gamma}\mathbf{\Theta}\mathbf{x}\|^2 \tag{4}$$

where χ^S represents the S-dimensional product space based on χ. For PPMB, the symbol can be detected in a parallel fashion as

$$\hat{\mathbf{x}}_{\boldsymbol{\eta}} = \arg \min_{\mathbf{x} \in \chi^P} \left\|\mathbf{y}_p - \mathbf{\Gamma}_p \tilde{\mathbf{\Theta}} \mathbf{x}\right\|^2 \tag{5}$$

for the precoded symbol, and

$$\hat{x}_l = \arg \min_{x \in \chi} |y_l - \lambda_{\tilde{l}} x|^2 \tag{6}$$

for the non-precoded symbol where $\tilde{l}$ is the corresponding index transformed by $\mathbf{T}$.

B. Bit-Interleaved Coded Multiple Beamforming with Constellation Precoding

Fig. 1(b) represents the structure of Bit-Interleaved Coded Multiple Beamforming with Constellation Precoding (BICMB-CP). In this system, first, the convolutional encoder with code rate $R_c = k_c/n_c$, possibly combined with a perforation matrix for a high rate punctured code, generates the codeword $\mathbf{c}$ from the information bits. Then, the spatial interleaver π_s distributes the coded bits into S streams, each of which is interleaved by an independent bit-wise interleaver π_t. The interleaved bits are mapped by Gray encoding onto the symbol sequence $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_K]$, where $\mathbf{x}_k$ is an $S \times 1$ symbol vector at the k^{th} time instant. Each entry of $\mathbf{x}_k$ belongs to a signal set χ.

The symbol vector $\mathbf{x}_k$ is multiplied by the $S \times S$ precoder $\mathbf{\Theta}$ in (1). When all of the S modulated entries are precoded ($P = S$), we call the resulting system Bit-Interleaved Coded Multiple Beamforming with Full Precoding (BICMB-FP), otherwise, we call it Bit-Interleaved Coded Multiple Beamforming with Partial Precoding (BICMB-PP). The precoded symbol vector is transmitted on the MIMO channel described in Section II-A.

As in UMB-CP, the spatial interleaver arranges the symbol vector $\mathbf{x}_k$ as $\mathbf{x}_k = [\mathbf{x}_{k,\boldsymbol{\eta}}^T \vdots \mathbf{x}_{k,\boldsymbol{\omega}}^T]^T = [x_{k,\eta_1} \cdots x_{k,\eta_P} \vdots x_{k,\omega_1} \cdots x_{k,\omega_{(S-P)}}]^T$. The $S \times 1$ detected symbol vector $\mathbf{r}_k = [(\mathbf{r}_k^p)^T \vdots (\mathbf{r}_k^n)^T]^T = [r_{k,1} \cdots r_{k,P} \vdots r_{k,P+1} \cdots r_{k,S}]^T$ at the k^{th} time instant is

$$\mathbf{r}_k = \mathbf{\Gamma}\mathbf{\Theta}\mathbf{x}_k + \mathbf{n}_k \tag{7}$$

where $\mathbf{n}_k = [(\mathbf{n}_k^p)^T \vdots (\mathbf{n}_k^n)^T]^T$ is an additive white Gaussian noise vector.

The location of the coded bit $c_{k'}$ within the symbol sequence $\mathbf{X}$ is known as $k' \rightarrow (k, l, i)$, where k, l, and i are the time instant in $\mathbf{X}$, the symbol position in $\mathbf{x}_k$, and the bit position on the label $x_{k,l}$, respectively. Let χ_b^i denote a subset of χ whose labels have $b \in \{0, 1\}$ in the i^{th} bit position. By using the location information and the input-output relation in (7), the receiver calculates the maximum likelihood bit metrics for the coded bit $c_{k'}$ as

$$\gamma^{l,i}(\mathbf{r}_k, c_{k'}) = \min_{\mathbf{x} \in \xi_{c_{k'}}^{l,i}} \|\mathbf{r}_k - \mathbf{\Gamma}\mathbf{\Theta}\mathbf{x}\|^2 \tag{8}$$

where $\xi_{c_{k'}}^{l,i}$ is a subset of χ^S, defined as

$$\xi_b^{l,i} = \{\mathbf{x} = [x_1 \cdots x_S]^T : x_{s|s=l} \in \chi_b^i, \text{ and } x_{s|s \neq l} \in \chi\}.$$

In particular, based on the decomposition of (7) similar to (5) and (6), the bit metrics, equivalent to (8) for partial precoding, are

$$\gamma^{l,i}(\mathbf{r}_k, c_{k'}) = \begin{cases} \min\limits_{\mathbf{x} \in \psi_{c_{k'}}^{l,i}} \|\mathbf{r}_k^p - \mathbf{\Gamma}_p \tilde{\mathbf{\Theta}} \mathbf{x}\|^2, & \text{if } 1 \leq l \leq P \\ \min\limits_{x \in \chi_{c_{k'}}^i} |r_{k,l} - \lambda_{\tilde{l}} x|^2, & \text{if } P+1 \leq l \leq S \end{cases} \tag{9}$$

where $\psi_b^{l,i}$ is a subset of χ^P, defined as

$$\psi_b^{l,i} = \{\mathbf{x} = [x_1 \cdots x_P]^T : x_{s|s=l} \in \chi_b^i, \text{ and } x_{s|s \neq l} \in \chi\},$$

and $\tilde{l}$ is an entry in $\boldsymbol{\omega}$, corresponding to the subchannel mapped by $\mathbf{T}$. Finally, the ML decoder makes decisions according to the rule

$$\hat{\mathbf{c}} = \arg \min_{\tilde{\mathbf{c}}} \sum_{k'} \gamma^{l,i}(\mathbf{r}_k, \tilde{c}_{k'}). \tag{10}$$

III. Diversity Analysis : UMB-CP

A. Fully Precoded Multiple Beamforming

Based on the ML decoding in (4), the upper bound to the instantaneous PEP between the transmitted symbol $\mathbf{x}$ and the detected symbol $\hat{\mathbf{x}}$ is calculated as

$$\Pr(\mathbf{x} \to \hat{\mathbf{x}} \mid \mathbf{H}) = \Pr\left(\|\mathbf{y} - \mathbf{\Gamma\Theta x}\|^2 \geq \|\mathbf{y} - \mathbf{\Gamma\Theta\hat{x}}\|^2 \mid \mathbf{H}\right) \leq \frac{1}{2}\exp\left(-\frac{\|\mathbf{\Gamma\Theta}(\mathbf{x}-\hat{\mathbf{x}})\|^2}{4N_0}\right). \quad (11)$$

Let $\mathbf{d} = [d_1 \cdots d_S]^T = \mathbf{\Theta}(\mathbf{x} - \hat{\mathbf{x}})$. Then, for FPMB, the average PEP becomes

$$\Pr(\mathbf{x} \to \hat{\mathbf{x}}) \leq E\left[\frac{1}{2}\exp\left(-\frac{\sum_{s=1}^{S}\lambda_s^2|d_s|^2}{4N_0}\right)\right]. \quad (12)$$

In [8], we showed that equations in the form of (12) have a closed form upper bound expression. We provide a formal statement below.

Theorem 1: Consider the $S \leq \min(N, M)$ ordered eigenvalues $\mu_1 > \cdots > \mu_S$ of the uncorrelated central Wishart matrix[1] [24], and a weight vector $\boldsymbol{\phi} = [\phi_1 \cdots \phi_S]^T$ with non-negative real elements. In the high signal-to-noise ratio regime, an upper bound for the expression $E[\exp(-\gamma\sum_{s=1}^{S}\phi_s\mu_s)]$ which is used in the diversity analysis of a number of MIMO systems is

$$E\left[\exp\left(-\gamma\sum_{s=1}^{S}\phi_s\mu_s\right)\right] \leq \zeta\left(\phi_{min}\gamma\right)^{-(N-\delta+1)(M-\delta+1)}$$

where γ is signal-to-noise ratio, ζ is a constant, $\phi_{min} = \min\{\phi_1, \cdots, \phi_S\}$, and δ is the index indicating the first nonzero element in the weight vector.

Proof: See [8]. □

Applying Theorem 1 to (12), we get the upper bound to PEP as

$$\Pr(\mathbf{x} \to \hat{\mathbf{x}}) \leq \tilde{\zeta}\left(\frac{\hat{d}_{min}}{4N}SNR\right)^{-(N-\delta+1)(M-\delta+1)} \quad (13)$$

where $\tilde{\zeta}$ is a constant, $\hat{d}_{min} = \min\{|d_1|^2, \cdots, |d_S|^2\}$, and δ is an index indicating the first nonzero element of the vector $\left[|d_1|^2 \cdots |d_S|^2\right]$. Therefore, FPMB achieves the full diversity order if δ from any distinct pair is equal to 1, which implies that $|d_1|^2 = |\boldsymbol{\theta}_1^T(\mathbf{x} - \hat{\mathbf{x}})|^2 > 0$ for any distinct pair, where $\boldsymbol{\theta}_1^T$ is the first row vector of $\mathbf{\Theta}$. Several methods to build the precoding matrix are described in [25] and [26].

B. Partially Precoded Multiple Beamforming

Generalizing (11) for PPMB, we get an upper bound to PEP as

$$\Pr(\mathbf{x} \to \hat{\mathbf{x}}) \leq E\left[\frac{1}{2}\exp\left(-\frac{\kappa}{4N_0}\right)\right] \quad (14)$$

[1] A central Wishart matrix is the Hermitian matrix $\mathbf{AA}^H$ where the entry of the matrix $\mathbf{A}$ is complex Gaussian with zero mean so that $E[\mathbf{A}] = \mathbf{0}$. The Wishart matrix $\mathbf{AA}^H$ is called uncorrelated if the common covariance matrix, defined as $\mathbf{C} = E[\mathbf{a}_s\mathbf{a}_s^H]\,\forall s$, where $\mathbf{a}_s$ is the s^{th} column vector of $\mathbf{A}$, satisfies $\mathbf{C} = \mathbf{I}$.

where

$$\kappa = \sum_{s=1}^{P}\lambda_{\eta_s}^2|\tilde{d}_s|^2 + \sum_{s=1}^{S-P}\lambda_{\omega_s}^2|x_{\omega_s} - \hat{x}_{\omega_s}|^2 \quad (15)$$

and $\tilde{d}_s$ is the s^{th} element of a vector $\tilde{\mathbf{d}} = \tilde{\mathbf{\Theta}}(\mathbf{x}_{\boldsymbol{\eta}} - \hat{\mathbf{x}}_{\boldsymbol{\eta}})$. Let us assume that the constellation precoding matrix $\tilde{\mathbf{\Theta}}$ meets the condition of FPMB to achieve the full diversity order. Since the expression (14) with (15) has a closed form expression similar to (13) as described in FPMB, the δ value needs to be obtained from a composite vector with the elements as $|\tilde{d}_s|^2$ and $|x_{\omega_s} - \hat{x}_{\omega_s}|^2$, to observe the diversity behavior of a given pairwise error. In addition, a different pair can lead to different diversity behavior. Therefore, we need to get the maximum δ out of all the possible pairwise errors to decide the diversity order of a given PPMB system.

All of the distinct pairs of $\mathbf{x}$ and $\hat{\mathbf{x}}$ are divided into three groups in terms of $\mathbf{x}_{\boldsymbol{\eta}}$, $\hat{\mathbf{x}}_{\boldsymbol{\eta}}$, $\mathbf{x}_{\boldsymbol{\omega}}$, and $\hat{\mathbf{x}}_{\boldsymbol{\omega}}$. The first group includes the pairs that have $\mathbf{x}_{\boldsymbol{\eta}} = \hat{\mathbf{x}}_{\boldsymbol{\eta}}$ but $\mathbf{x}_{\boldsymbol{\omega}} \neq \hat{\mathbf{x}}_{\boldsymbol{\omega}}$, and the second group comprises the pairs satisfying $\mathbf{x}_{\boldsymbol{\eta}} \neq \hat{\mathbf{x}}_{\boldsymbol{\eta}}$ but $\mathbf{x}_{\boldsymbol{\omega}} = \hat{\mathbf{x}}_{\boldsymbol{\omega}}$. Finally, the last group consists of the pairs for which $\mathbf{x}_{\boldsymbol{\eta}} \neq \hat{\mathbf{x}}_{\boldsymbol{\eta}}$ and $\mathbf{x}_{\boldsymbol{\omega}} \neq \hat{\mathbf{x}}_{\boldsymbol{\omega}}$. We will present the method to calculate the maximum δ for each group, and to find δ_{max} from the groups.

Since the vector $\tilde{\mathbf{d}}$ is a zero vector for the first group, the first summation of κ in (15) is zero, resulting in δ being equal to the minimum of $\boldsymbol{\omega}$. By considering all of the possible pairs, we easily see that $\omega_1 \leq \delta \leq \omega_{(S-P)}$. Therefore, the maximum value is $\delta_1 = \omega_{(S-P)}$ which corresponds to the pair satisfying $x_s = \hat{x}_s$ for all s except $s = \omega_{(S-P)}$. For any pair in the second group, the term with the first singular value survives in κ, according to the inherited property of the constellation precoding matrix, i.e., $|\tilde{d}_1|^2 > 0$. However, the second summation in κ disappears since $\mathbf{x}_{\boldsymbol{\omega}} = \hat{\mathbf{x}}_{\boldsymbol{\omega}}$. Therefore, the maximum value of this group is $\delta_2 = \eta_1$. Now, for the third group, both summations in κ exist. Then, δ is chosen to be the smaller value between the minimum of $\boldsymbol{\omega}$ and η_1. In the same manner as was already given in the analysis of the first group, the maximum of the minimum of $\boldsymbol{\omega}$ is found to be $\omega_{(S-P)}$. Therefore, the maximum δ for this group is $\delta_3 = \max\{\eta_1, \omega_{(S-P)}\}$. Finally, δ_{max} is decided as

$$\delta_{max} = \max\{\delta_1, \delta_2, \delta_3\} = \max\left(\eta_1, \omega_{(S-P)}\right). \quad (16)$$

Example: In Table I, we summarize the diversity order for all of the possible combinations of the 4×4 PPMB system $S = 4$ and $P = 2$. We will provide simulation results that verify this analysis in Section VI, specifically in Fig.4.

IV. Diversity Analysis : BICMB-CP

A. BICMB with Full Precoding

We assume that the d_H coded bits are interleaved such that they are placed in distinct symbols, where d_H denotes the Hamming distance between the transmitted codeword $\mathbf{c}$ and the decoded codeword $\hat{\mathbf{c}}$. Since the bit metrics in (8) are the same for the same coded bits between the pairwise errors, the

TABLE I
DIVERSITY ORDER (O_{div}) OF 4×4, $S = 4$ PARTIALLY PRECODED MULTIPLE BEAMFORMING SYSTEM

P	$\boldsymbol{\eta}$	$\boldsymbol{\omega}$	η_1	$\omega_{(S-P)}$	δ_{max}	O_{div}
2	[1 2]	[3 4]	1	4	4	1
	[1 3]	[2 4]	1	4	4	1
	[1 4]	[2 3]	1	3	3	4
	[2 3]	[1 4]	2	4	4	1
	[2 4]	[1 3]	2	3	3	4
	[3 4]	[1 2]	3	2	3	4
3	[1 2 3]	[4]	1	4	4	1
	[1 2 4]	[3]	1	3	3	4
	[1 3 4]	[2]	1	2	2	9
	[2 3 4]	[1]	2	1	2	9

original PEP is replaced by

$$\Pr(\mathbf{c} \to \hat{\mathbf{c}}|\mathbf{H}) = \Pr\left(\sum_{k,d_H} \min_{\mathbf{x}\in\xi^{l,i}_{c_{k'}}} \|\mathbf{r}_k - \mathbf{\Gamma\Theta x}\|^2 \geq \sum_{k,d_H} \min_{\mathbf{x}\in\xi^{l,i}_{\bar{c}_{k'}}} \|\mathbf{r}_k - \mathbf{\Gamma\Theta x}\|^2\right) \quad (17)$$

where the summation is restricted to the symbols corresponding to the different d_H coded bits.

Let us define $\tilde{\mathbf{x}}_k$ and $\hat{\mathbf{x}}_k$ as

$$\begin{aligned} \tilde{\mathbf{x}}_k &= \arg\min_{\mathbf{x}\in\xi^{l,i}_{c_{k'}}} \|\mathbf{r}_k - \mathbf{\Gamma\Theta x}\|^2 \\ \hat{\mathbf{x}}_k &= \arg\min_{\mathbf{x}\in\xi^{l,i}_{\bar{c}_{k'}}} \|\mathbf{r}_k - \mathbf{\Gamma\Theta x}\|^2 \end{aligned} \quad (18)$$

where $\bar{c}_{k'}$ is the complement of $c_{k'}$ in binary codes. It is easily found that $\tilde{\mathbf{x}}_k$ is different from $\hat{\mathbf{x}}_k$ since the sets that the l^{th} symbols belong to are disjoint, as can be seen from the definition of $\xi^{l,i}_{c_{k'}}$. In the same manner, we see that $\mathbf{x}_k$ is different from $\hat{\mathbf{x}}_k$. With $\tilde{\mathbf{x}}_k$ and $\hat{\mathbf{x}}_k$, we get, from (17),

$$\Pr(\mathbf{c} \to \hat{\mathbf{c}}|\mathbf{H}) = \Pr\left(\sum_{k,d_H} \|\mathbf{r}_k - \mathbf{\Gamma\Theta}\tilde{\mathbf{x}}_k\|^2 \geq \sum_{k,d_H} \|\mathbf{r}_k - \mathbf{\Gamma\Theta}\hat{\mathbf{x}}_k\|^2\right). \quad (19)$$

Based on the fact that $\|\mathbf{r}_k - \mathbf{\Gamma\Theta x}_k\|^2 \geq \|\mathbf{r}_k - \mathbf{\Gamma\Theta}\tilde{\mathbf{x}}_k\|^2$ and the relation in (7), equation (19) is upper-bounded by

$$\Pr(\mathbf{c} \to \hat{\mathbf{c}}|\mathbf{H}) \leq \Pr\left(\beta \geq \sum_{k,d_H} \|\mathbf{\Gamma\Theta}(\mathbf{x}_k - \hat{\mathbf{x}}_k)\|^2\right) \quad (20)$$

where

$$\beta = -\sum_{k,d_H} (\mathbf{x}_k - \hat{\mathbf{x}}_k)^H \mathbf{\Theta}^H \mathbf{\Gamma n}_k + \mathbf{n}_k^H \mathbf{\Gamma\Theta}(\mathbf{x}_k - \hat{\mathbf{x}}_k).$$

Since β is a zero mean Gaussian random variable with variance $2N_0 \sum_{k,d_H} \|\mathbf{\Gamma\Theta}(\mathbf{x}_k - \hat{\mathbf{x}}_k)\|^2$, the right hand side of (20) is replaced by the Q function as

$$\Pr(\mathbf{c} \to \hat{\mathbf{c}}|\mathbf{H}) \leq Q\left(\sqrt{\frac{\sum_{k,d_H} \|\mathbf{\Gamma\Theta}(\mathbf{x}_k - \hat{\mathbf{x}}_k)\|^2}{2N_0}}\right). \quad (21)$$

The numerator in (21) is rewritten as

$$\sum_{k,d_H} \|\mathbf{\Gamma\Theta}(\mathbf{x}_k - \hat{\mathbf{x}}_k)\|^2 = \sum_{s=1}^{S} \lambda_s^2 \sum_{k,d_H} |d_{k,s}|^2 \quad (22)$$

where $d_{k,s}$ is the s^{th} entry of the vector $\mathbf{d}_k = \mathbf{\Theta}(\mathbf{x}_k - \hat{\mathbf{x}}_k)$. Using an upper bound to the Q function, we calculate the average PEP as

$$\Pr(\mathbf{c} \to \hat{\mathbf{c}}) \leq E\left[\exp\left(-\frac{\sum_{s=1}^{S} \lambda_s^2 \sum_{k,d_H} |d_{k,s}|^2}{4N_0}\right)\right]. \quad (23)$$

According to Theorem 1, we can evaluate the diversity order of a given system by calculating the weight vector whose s^{th} element is $\sum_{k,d_H} |d_{k,s}|^2$. In particular, if the constellation precoder is designed such that

$$|d_{k,1}|^2 = |\boldsymbol{\theta}_1^T(\mathbf{x}_k - \hat{\mathbf{x}}_\mathbf{k})|^2 > 0, \forall(\mathbf{x}_k, \hat{\mathbf{x}}_k) \quad (24)$$

where $\boldsymbol{\theta}_1^T$ is the first row vector of the precoding matrix $\mathbf{\Theta}$, we see that $\sum_{k,d_H} |d_{k,1}|^2 > 0$, resulting in the full diversity order of NM. Therefore, (24) is a sufficient condition for the full diversity order of BICMB-FP.

B. BICMB with Partial Precoding

The bit metrics in (9) lead to the PEP calculation as

$$\Pr(\mathbf{c} \to \hat{\mathbf{c}}|\mathbf{H}) = \Pr(\tau_1 \geq \tau_2) \quad (25)$$

where

$$\tau_1 = \sum_{k,d_H^p} \min_{\mathbf{x}\in\psi^{l,i}_{c_{k'}}} \|\mathbf{r}_k^p - \mathbf{\Gamma}_p\tilde{\mathbf{\Theta}}\mathbf{x}\|^2 + \sum_{k,d_H^n} \min_{x\in\chi^{l,i}_{c_{k'}}} |r_{k,l} - \lambda_{\tilde{l}} x|^2$$

$$\tau_2 = \sum_{k,d_H^p} \min_{\mathbf{x}\in\psi^{l,i}_{\bar{c}_{k'}}} \|\mathbf{r}_k^p - \mathbf{\Gamma}_p\tilde{\mathbf{\Theta}}\mathbf{x}\|^2 + \sum_{k,d_H^n} \min_{x\in\chi^{l,i}_{\bar{c}_{k'}}} |r_{k,l} - \lambda_{\tilde{l}} x|^2$$

and $\sum_{k,d_H^p}$, $\sum_{k,d_H^n}$ stand for the summation over the d_H^p and d_H^n bit metrics, with d_H^p and d_H^n denoting the number of different coded bits between the two pairwise errors residing on the precoded and the non-precoded subchannels specified by $\boldsymbol{\eta}$ and $\boldsymbol{\omega}$, respectively. By using the appropriate system input-output relations, the PEP is written as

$$\Pr(\mathbf{c} \to \hat{\mathbf{c}}|\mathbf{H}) = \Pr\left(\hat{\beta} \geq \hat{\kappa}\right) \quad (26)$$

where $\hat{\beta} = \beta_p + \beta_n$,

$$\beta_p = -\sum_{k,d_H^p} (\mathbf{x}_{k,\boldsymbol{\eta}} - \hat{\mathbf{x}}_{k,\boldsymbol{\eta}})^H \tilde{\mathbf{\Theta}}^H \mathbf{\Gamma}_p \mathbf{n}_k^p + (\mathbf{n}_k^p)^H \mathbf{\Gamma}_p \tilde{\mathbf{\Theta}}(\mathbf{x}_{k,\boldsymbol{\eta}} - \hat{\mathbf{x}}_{k,\boldsymbol{\eta}}),$$

$$\beta_n = -\sum_{k,d_H^n} \lambda_{\tilde{l}}(x_{k,l} - \hat{x}_{k,l})^* n_{k,l} + \lambda_{\tilde{l}}(x_{k,l} - \hat{x}_{k,l}) n_{k,l}^*,$$

and

$$\hat{\kappa} = \sum_{k,d_H^p} \|\mathbf{\Gamma}_p \tilde{\mathbf{\Theta}} \left(\mathbf{x}_{k,\boldsymbol{\eta}} - \hat{\mathbf{x}}_{k,\boldsymbol{\eta}}\right) \|^2 + \sum_{k,d_H^n} |\lambda_{\tilde{l}} \left(x_{k,l} - \hat{x}_{k,l}\right)|^2.$$

Since $\hat{\beta}$ in (26) is a Gaussian random variable with zero mean and variance $2N_0\hat{\kappa}$, the PEP can be expressed in a way similar to (21) with the Q-function. In addition, if we define σ as

$$\sigma = \sum_{r=1}^{P} \lambda_{\eta_r}^2 \sum_{k,d_H^p} |\hat{d}_{k,r}|^2 + d_{min}^2 \sum_{r=1}^{S-P} \lambda_{\omega_r}^2 \alpha_{\omega_r} \qquad (27)$$

where $\hat{d}_{k,r}$ is the r^{th} entry of the vector $\hat{\mathbf{d}}_k = \tilde{\mathbf{\Theta}} \left(\mathbf{x}_{k,\boldsymbol{\eta}} - \hat{\mathbf{x}}_{k,\boldsymbol{\eta}}\right)$, and α_s is the number of times the s^{th} subchannel is used corresponding to d_H^n bits under consideration, then we can see that $\sigma \leq \hat{\kappa}$. Finally, the average PEP is calculated as

$$\Pr\left(\mathbf{c} \rightarrow \hat{\mathbf{c}}\right) \leq E\left[\frac{1}{2}\exp\left(-\frac{\sigma}{4N_0}\right)\right]. \qquad (28)$$

To determine the diversity order from σ, we need to find the index indicating the first nonzero element in an ordered composite vector which consists of $\sum_{k,d_H^p} |\hat{d}_{k,r}|^2$ and α_{ω_r} as in Theorem 1. If $d_H^p = 0$, the first summation part of σ vanishes. In this case, the first index is

$$\delta = \min\{s : \alpha_s > 0 \text{ for } s \in \{\omega_1, \cdots, \omega_{(S-P)}\}\}. \qquad (29)$$

In the other case of $d_H^p > 0$, we see that $\mathbf{x}_{k,\boldsymbol{\eta}}$ and $\hat{\mathbf{x}}_{k,\boldsymbol{\eta}}$ are obviously different for the same reason as in the previous section. If the constellation precoder satisfies the sufficient condition of (24), the term with $\lambda_{\eta_1}^2$ always exists in σ. By considering the second term of σ, we get δ for the case of $d_H^p > 0$

$$\delta = \begin{cases} \min(\eta_1, \delta') & \text{if } \delta' \text{ exists,} \\ \eta_1 & \text{otherwise.} \end{cases} \qquad (30)$$

where δ', if it exists, is obtained in the same way as (29). If, in search of δ', no s satisfying the right hand side of (29) exists, we state δ' does not exist and set $\delta = \eta_1$, as in (30).

Example: In this example, we employ 4-state $1/2$-rate convolutional code with generator polynomials $(5,7)$ in octal representation, in an $N = M = S = 3$ system. Two types of spatial interleavers are used to demonstrate the different results of the diversity order. A generalized transfer function of BICMB with the specific spatial interleaver and convolutional code provides the α-vectors for all of the pairwise errors, whose element indicates the number of times the stream is used for the erroneous bits [8]. In particular, due to the fact that $d_H^p = \sum_{r=1}^{P} \alpha_{\eta_r}$ and $d_H^n = \sum_{r=1}^{S-P} \alpha_{\omega_r}$ where α_s is the s^{th} element of the α-vector, the generalized transfer function approach in [8] is also useful in the analysis of BICMB-PP. Hence, we rewrite the transfer functions of the systems from [8], where a, b, and c are the symbolic representation of the $1^{st}, 2^{nd}, 3^{rd}$ streams, respectively. The spatial interleaver used in $\mathcal{T}_1$ is a simple rotating switch on 3 streams. For $\mathcal{T}_2$, the u^{th} coded bit is interleaved into the stream $s_{\mathrm{mod}(u-1,18)+1}$ where $s_1 = \cdots = s_6 = 1$, $s_7 = \cdots = s_{12} = 2$, $s_{13} = \cdots = s_{18} = 3$ and mod is the modulo operation. Each term represents an α-vector, and the powers of a, b, c in this term indicate the elements of the α-vector corresponding to that term.

$$\begin{aligned} \mathcal{T}_1 = {} & Z^5(a^2b^2c + a^2bc^2 + ab^2c^2) \\ & + Z^6(a^3b^2c + a^2b^3c + a^3bc^2 + \\ & \quad ab^3c^2 + a^2bc^3 + ab^2c^3) \\ & + Z^7(2a^3b^3c + 2a^3b^2c^2 + 2a^2b^3c^2 + \\ & \quad 2a^3bc^3 + 2a^2b^2c^3 + 2ab^3c^3) \\ & + Z^8(a^5b^3 + a^4b^3c + a^3b^4c + 2a^4b^2c^2 + \\ & \quad 3a^3b^3c^2 + 2a^2b^4c^2 + a^4bc^3 + 3a^3b^2c^3 + \\ & \quad 3a^2b^3c^3 + ab^4c^3 + b^5c^3 + a^3bc^4 + \\ & \quad 2a^2b^2c^4 + ab^3c^4 + a^3c^5) + \cdots \end{aligned} \qquad (31)$$

$$\begin{aligned} \mathcal{T}_2 = {} & Z^5(a^5 + a^3b^2 + a^2b^3 + \\ & \quad b^5 + a^3c^2 + b^3c^2 + a^2c^3 + b^2c^3 + c^5) \\ & + Z^6(a^4b^2 + 3a^3b^3 + a^2b^4 + a^4c^2 + 3a^2b^2c^2 + \\ & \quad b^4c^2 + 3a^3c^3 + 3b^3c^3 + a^2c^4 + b^2c^4) \\ & + Z^7(2a^4b^3 + 2a^3b^4 + a^3b^3c + 7a^3b^2c^2 + \\ & \quad 7a^2b^3c^2 + 2a^4c^3 + a^3bc^3 + 7a^2b^2c^3 + \\ & \quad ab^3c^3 + 2b^4c^3 + 2a^3c^4 + 2b^3c^4) + \cdots \end{aligned} \qquad (32)$$

Consider the case $\boldsymbol{\eta} = [1\,2]$. We see that all of the α-vectors of $\mathcal{T}_1$ have $d_H^p > 0$. Since $\eta_1 = 1$, δ equals 1 whether δ' exists or not. In fact, δ' does not exist for the term $Z^8a^5b^3$. Therefore, the $\mathcal{T}_1$ BICMB-PP system with $\boldsymbol{\eta} = [1\,2]$ achieves the full diversity order while BICMB without constellation precoding [8], or PPMB without Bit-Interleaved Coded Modulation (BICM) loses the full diversity order [25], [26]. For $\mathcal{T}_2$, the α-vector $[0\,0\,5]$ gives $d_H^p = 0$, resulting in $\delta = 3$. Therefore, the $\mathcal{T}_2$ BICMB-PP system with $\boldsymbol{\eta} = [1\,2]$ does not achieve the full diversity order.

The same analysis for $\boldsymbol{\eta} = [1\,3]$ results in the diversity order of 9, and $[2\,3]$ results in 4 for the transfer function $\mathcal{T}_1$. Similarly, both of $[1\,3]$ and $[2\,3]$ result in the diversity of 4 for $\mathcal{T}_2$. As a consequence, we find that proper selection of the subchannels for precoding, as well as the appropriate pattern of the spatial interleaver, is important to achieve the full diversity order of BICMB-PP. We will present simulation results that verify this analysis in Section VI, in particular, in Fig. 7.

V. Reduced Computational Complexity Sphere Detection

In this section, we will describe the reduced computational complexity SD for constellation precoded multiple beamforming employing square QAM. More specifically, we propose the SD technique to reduce the number of multiplications without losing the performance. Since detecting the transmitted non-precoded symbols for UMB-CP in (6) and finding the bit metrics of non-precoded symbols for BICMB-CP in (9) can be carried out independently of the symbols on the other subchannels, we focus on the precoded P symbols.

Given that a full search over the entire lattice space is performed [27], solving (5) for ML detection is well-known to be NP-hard. SD, on the other hand, solves (5) by searching only lattice points that lie inside a sphere of radius ρ centering around the received vector $\mathbf{y}_p$. A frequently used solution for the QAM-modulated complex signal model is to decompose the P-dimensional complex-valued problem (5) into a $2P$-dimensional real-valued problem, which is written as

$$\begin{aligned}\bar{\mathbf{y}} &= \begin{bmatrix}\Re\{\mathbf{y}_p\}\\ \Im\{\mathbf{y}_p\}\end{bmatrix} = \bar{\mathbf{F}}\bar{\mathbf{x}} + \bar{\mathbf{n}}\\ &= \begin{bmatrix}\Re\{\mathbf{F}\} & -\Im\{\mathbf{F}\}\\ \Im\{\mathbf{F}\} & \Re\{\mathbf{F}\}\end{bmatrix}\begin{bmatrix}\Re\{\mathbf{x}_{\boldsymbol{\eta}}\}\\ \Im\{\mathbf{x}_{\boldsymbol{\eta}}\}\end{bmatrix} + \begin{bmatrix}\Re\{\mathbf{n}_p\}\\ \Im\{\mathbf{n}_p\}\end{bmatrix}\end{aligned} \tag{33}$$

where $\mathbf{F} = \boldsymbol{\Gamma}_p\tilde{\boldsymbol{\Theta}}$ [15], [27]. The QR decomposition of the $2P \times 2P$ real-valued channel matrix turns (5) into the equivalent expression

$$\hat{\mathbf{x}}_{\boldsymbol{\eta}} = \arg\min_{\mathbf{x}\in\Psi}\left\|\bar{\mathbf{Q}}^H\bar{\mathbf{y}} - \bar{\mathbf{R}}\mathbf{x}\right\|^2 \tag{34}$$

where $\bar{\mathbf{Q}}$ and $\bar{\mathbf{R}}$ are the unitary matrix and the upper triangular matrix from the QR decomposition of $\bar{\mathbf{F}}$ [15], [27]. Let Ω denote the set of scalar symbols for one dimension of QAM, e.g., $\Omega = \{-3, -1, 1, 3\}$ for 16-QAM, then Ψ denotes a subset of Ω^{2P} whose elements satisfy $\|\bar{\mathbf{Q}}^H\bar{\mathbf{y}} - \bar{\mathbf{R}}\mathbf{x}\|^2 < \rho^2$. The initial radius ρ should be chosen properly so that it is neither too small nor too large. Too small an initial radius can result in too many unsuccessful searches by restarting the search and thus increasing the complexity, while too large an initial radius can result in too many lattice points to be searched.

The SD algorithm can be viewed as a pruning algorithm on a tree of depth $2P$, whose branches correspond to elements drawn from the set Ω [23], [27]. Conventional SD implements a Depth-First Search (DFS) strategy in the tree. This search achieves ML performance. The complexity of SD is measured in terms of the number of operations required per visited node multiplied by the number of visited nodes throughout the search algorithm [27]. The complexity can be reduced by either reducing the number of nodes to be visited, or the number of operations to be carried out at each node, or both. In order to reduce the number of visited nodes, one can either make a judicious choice of the initial radius to start the algorithm, or execute a proper sphere radius update strategy. The former strategy has been studied in [16] and [17], and the latter one has been discussed in [18] and [19]. In this paper, we propose methods to reduce the average number of real multiplications, which are the most expensive operations in terms of machine cycles required at each node for conventional SD. A proper choice of the initial radius for BICMB-CP will also be provided.

We start by writing the node weight as [23]

$$w(\bar{\mathbf{x}}^{(u)}) = w(\bar{\mathbf{x}}^{(u+1)}) + w_{pw}(\bar{\mathbf{x}}^{(u)}) \tag{35}$$

with $u = 2P, 2P-1, \cdots, 1$, $w(\bar{\mathbf{x}}^{(2P+1)}) = 0$, and $w_{pw}(\bar{\mathbf{x}}^{(2P+1)}) = 0$, where $\bar{\mathbf{x}}^{(u)}$ denotes the partial vector symbol at layer u. The partial weight $w(\bar{\mathbf{x}}^{(u)})$ is written as

$$w_{pw}(\bar{\mathbf{x}}^{(u)}) = |\tilde{y}_u - \sum_{v=u}^{2P}\bar{R}_{u,v}\bar{x}_v|^2 \tag{36}$$

where $\tilde{y}_u$ is the u^{th} element of $\bar{\mathbf{Q}}^H\bar{\mathbf{y}}$, $\bar{R}_{u,v}$ is the $(u,v)^{th}$ element of $\bar{\mathbf{R}}$, and $\bar{x}_v$ is the v^{th} element of $\bar{\mathbf{x}}$.

A. Precalculation of Multiplications

Note that for one channel realization, both $\bar{\mathbf{R}}$ and Ω are independent of time. In other words, to decode different received symbols for one channel realization, the only term in (36) which depends on time is $\tilde{y}_u$. Consequently, a table $\mathbb{T}$ can be constructed to store all terms of $\bar{R}_{u,v}\bar{x}$, where $\bar{R}_{u,v} \neq 0$ and $\bar{x} \in \Omega$, before starting the tree search procedure. Equations (35) and (36) imply that only one real multiplication is needed by using $\mathbb{T}$ instead of $2P-u+2$ for each node to calculate the node weight. As a result, the number of real multiplications can be significantly reduced.

Taking the square QAM structure into consideration, Ω can be divided into two smaller sets Ω_1 with negative elements and Ω_2 with positive elements. Take 16-QAM for example, $\Omega = \{-3, -1, 1, 3\}$, then $\Omega_1 = \{-3, -1\}$ and $\Omega_2 = \{1, 3\}$. Any negative element in Ω_1 has a positive element with the same absolute value in Ω_2. Consequently, in order to build $\mathbb{T}$, only terms of $\bar{R}_{u,v}\bar{x}$, where $\bar{R}_{u,v} \neq 0$ and $\bar{x} \in \Omega_1$, need to be calculated and stored. Hence, the size of $\mathbb{T}$ is

$$|\mathbb{T}| = \frac{N_R|\Omega|}{2} \tag{37}$$

where N_R denotes the number of nonzero elements in matrix $\bar{\mathbf{R}}$, and $|\Omega|$ denotes the size of Ω.

In order to build $\mathbb{T}$, both the number of terms that need to be stored and the number of real multiplications required are $|\mathbb{T}|$. Since the channel is assumed to be flat fading, only one $\mathbb{T}$ needs to be built in one burst. If the burst length is very long, the computational complexity of building $\mathbb{T}$ can be neglected.

B. Modified Depth First Search DFS Algorithm

The representation proposed in [23] replaces the conventional representation of (33) with

$$\check{\mathbf{y}} = \mathbf{G}\check{\mathbf{x}} + \check{\mathbf{n}} \tag{38}$$

where

$$\check{\mathbf{y}} = \left[\Re\{y_1\}\ \Im\{y_1\}\ \cdots\ \Re\{y_P\}\ \Im\{y_P\}\right]^T,$$

$$\mathbf{G} = \begin{bmatrix}\Re\{F_{1,1}\} & -\Im\{F_{1,1}\} & \cdots & \Re\{F_{1,P}\} & -\Im\{F_{1,P}\}\\ \Im\{F_{1,1}\} & \Re\{F_{1,1}\} & \cdots & \Im\{F_{1,P}\} & \Re\{F_{1,P}\}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \Re\{F_{P,1}\} & -\Im\{F_{P,1}\} & \cdots & \Re\{F_{P,P}\} & -\Im\{F_{P,P}\}\\ \Im\{F_{P,1}\} & \Re\{F_{P,1}\} & \cdots & \Im\{F_{P,P}\} & \Re\{F_{P,P}\}\end{bmatrix},$$

$$\check{\mathbf{x}} = \left[\Re\{x_{\eta_1}\}\ \Im\{x_{\eta_1}\}\ \cdots\ \Re\{x_{\eta_P}\}\ \Im\{x_{\eta_P}\}\right]^T,$$

$$\check{\mathbf{n}} = \left[\Re\{n_1\}\ \Im\{n_1\}\ \cdots\ \Re\{n_P\}\ \Im\{n_P\}\right]^T.$$

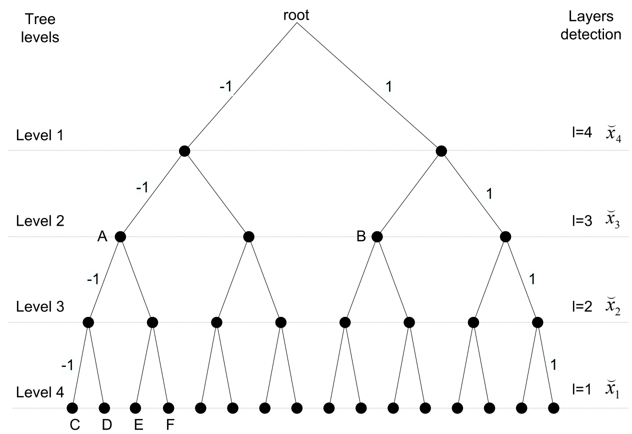


Fig. 2. Tree structure for a 2×2 FPMB system employing 4-QAM.

The structure of the lattice representation becomes advantageous after applying the QR decomposition to $\mathbf{G}$, i.e., $\mathbf{G} = \mathbf{QR}$. Due to a special form of orthogonality between each pair of columns, all elements $R_{u,u+1}$ for $u = 1, 3, \cdots, 2P-1$, in the upper triangular matrix $\mathbf{R}$ become zero [23]. The locations of these zeros introduce orthogonality between the real and imaginary parts of every detected symbol, which can be taken advantage of to reduce the computational complexity of SD. We provide the following example to explain this.

Consider a 2×2 $S = 2$ FPMB system employing 4-QAM. Then, SD constructs a tree with $2P = 4$ levels, where the branches coming out of each node represent the real values in the set $\Omega = \{-1, 1\}$. This tree is shown in Fig. 2. Based on the representation in (38), the input-output relation is given by

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \hat{y}_4 \end{bmatrix} = \begin{bmatrix} R_{1,1} & 0 & R_{1,3} & R_{1,4} \\ 0 & R_{2,2} & R_{2,3} & R_{2,4} \\ 0 & 0 & R_{3,3} & 0 \\ 0 & 0 & 0 & R_{4,4} \end{bmatrix} \begin{bmatrix} \breve{x}_1 \\ \breve{x}_2 \\ \breve{x}_3 \\ \breve{x}_4 \end{bmatrix} + \begin{bmatrix} \hat{n}_1 \\ \hat{n}_2 \\ \hat{n}_3 \\ \hat{n}_4 \end{bmatrix} \tag{39}$$

where $\hat{y}_u$, $\breve{x}_u$, $\hat{n}_u$ are the u^{th} element of the vectors $\mathbf{Q}^H\breve{\mathbf{y}}$, $\breve{\mathbf{x}}$, $\mathbf{Q}^H\breve{\mathbf{n}}$, respectively, and $R_{u,v}$ is the element of $\mathbf{R}$.

Calculating partial node weights of (39) for the first level and the second level are independent, same as the third level and the fourth level, because of the additional zeros in the $\mathbf{R}$ matrix. For instance, the partial weights of node A and B in Fig. 2 depend on only $\breve{x}_3$, and the partial weights of node C, D, E, and F depend on $\breve{x}_4$, $\breve{x}_3$, and $\breve{x}_1$ except $\breve{x}_2$. In other words, the partial weights of node A and B are equal, and need to be calculated once. Similarly, partial weights of node C and D can be used without an additional computation for the partial weights of node E and F, respectively.

Because of this feature, the DFS strategy is modified in the following way: for the u^{th} layer, where u is an odd number, partial weights of the nodes at the layer u (called *children nodes*) belonging to a node at the layer $u+1$ (called *a parent node*) are stored, and are used as partial weights of the nodes belonging to the same node at the layer $u+2$ (called *a grandparent node*), but to the different parent nodes. In other words, the weights of children nodes belonging to one of the parent nodes are recycled by the children's *cousins*.

By implementing the modified DFS algorithm, further complexity reduction is achieved beyond the reduction due to the precalculation table $\mathbb{T}$. We will show how many real multiplications are reduced to calculate all nodes at layers $u, u+1$ belonging to one grandparent node at layer $u+2$, where u is an odd number. Let us define $\nu \in [0, |\Omega|]$ as the number of non-pruned branches from the grandparent node, after calculating the node weights $\omega(\breve{\mathbf{x}}^{(u+1)})$ and comparing them with ρ^2. If $\nu = 0$, which means all branches from the grandparent node are pruned, the modified algorithm does not reduce computations from the original DFS algorithm. If $\nu > 0$, to get all of the weights at the layer u and $u+1$ under the grandparent node, the number of real multiplications reduces further from $(\nu+1)|\Omega|$ to $2|\Omega|$.

C. Initial Radius for BICMB-CP

The proposed SD algorithm for UMB-CP described in the previous sections can also be applied to BICMB-CP. The P-dimensional complex-valued input-output relation of the precoded part in (9) can be transformed into a $2P$-dimensional real-valued problem, based on the lattice representation in (38). Applying the QR decomposition to the $2P \times 2P$ dimensional matrix $\mathbf{G}$ in (38), the bit metrics of the precoded part in (9) are rewritten as

$$\gamma^{l,i}(\mathbf{r}_k, c_{k'}) = \min_{\mathbf{x} \in \Phi_{c_{k'}}} \|\hat{\mathbf{r}}_k - \mathbf{R}\mathbf{x}\|^2 \tag{40}$$

where $\hat{\mathbf{r}}_k$ is the product of $\mathbf{Q}^H$ and the transformed vector from $\mathbf{r}_k^p$. Due to the transformation, the position of $c_{k'}$ in the label of $\mathbf{x}$ needs to be acquired and stored in a new table $k' \rightarrow (k, \hat{l}, \hat{i})$, which means $c_{k'}$ lies in the $\hat{i}^{th}$ bit position of label for the $\hat{l}^{th}$ element of real-valued symbol vector $\mathbf{x}$. Let $\Omega_b^{\hat{i}}$ denote a subset of Ω whose labels have $b \in \{0, 1\}$ in the $\hat{i}^{th}$ bit position. If we define $\tilde{\xi}_b^{\hat{l},\hat{i}}$ as

$$\tilde{\xi}_b^{\hat{l},\hat{i}} = \{\mathbf{x} : x_{s|s=\hat{l}} \in \Omega_b^{\hat{i}}, \text{ and } x_{s|s \neq \hat{l}} \in \Omega\}$$

then, Φ_b denotes a subset of $\tilde{\xi}_b^{\hat{l},\hat{i}}$, whose elements satisfy $\|\hat{\mathbf{r}}_k - \mathbf{R}\mathbf{x}\|^2 \leq \rho_b^2$.

Similarly to UMB-CP, the SD algorithm for BICMB-CP now can be viewed as a pruning algorithm on a tree of depth $2P$. However, its branches of the layer $u = \hat{l}$ correspond to elements drawn only from the set $\chi_{c_{k'}}^{\hat{i}} \subset \chi$. To determine the initial radius for BICMB-CP, we use the ZF-DFE algorithm to acquire an estimated real-valued vector symbol $\mathbf{x}_k^b$ for $b = 0$ or 1, whose u^{th} element $x_{k,u}^b$ is detected successively from $x_{k,2P}^b$ to $x_{k,1}^b$ as

$$x_{k,u}^b = \arg\min_{x \in \Omega_b^{\hat{i}}} |\hat{r}_{k,u} - \sum_{v=u+1}^{2P} R_{u,v} x_{k,v}^b - R_{u,u} x| \tag{41}$$

for the element corresponding to $\hat{l}$ indicated by the table $k' \rightarrow$

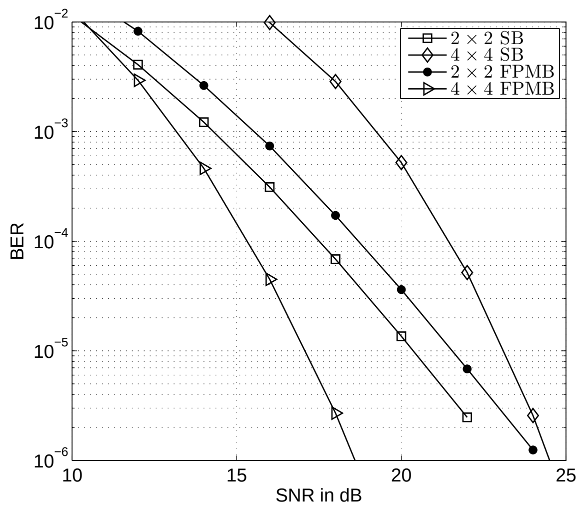


Fig. 3. BER vs. SNR comparison for 2×2, 4×4 SB and FPMB.

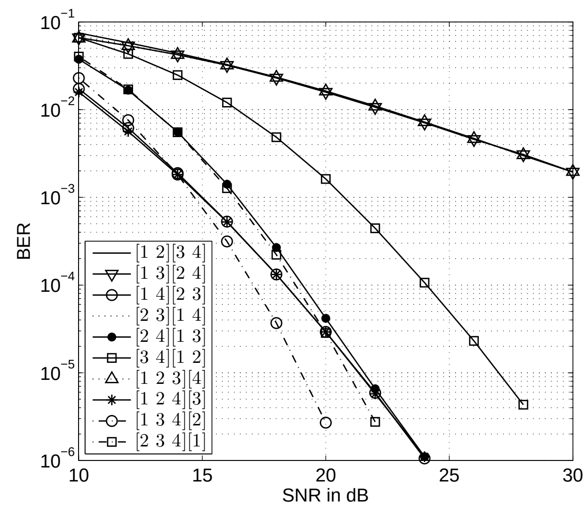


Fig. 4. BER vs. SNR for 4×4 $S = 4$, 4-QAM PPMB.

$(k, \hat{l}, \hat{i})$, and

$$x_{k,u}^b = \arg\min_{x \in \Omega} |\hat{r}_{k,u} - \sum_{v=u+1}^{2P} R_{u,v} x_{k,v}^b - R_{u,u} x| \quad (42)$$

for the rest of the elements. Then, the initial radius is calculated by

$$\rho_b^2 = \|\hat{\mathbf{r}}_k - \mathbf{R}\mathbf{x}_k^b\|^2. \quad (43)$$

With the initial radius acquired by the ZF-DFE algorithm, the SD guarantees no unsuccessful search for both of the bit metrics.

VI. Simulation Results

A. UMB-CP

We will now verify the diversity order analysis in Section III by means of simulation results using different system configurations. In Fig. 3, we present BER performance results for SB and FPMB. The curves with the legend FPMB are generated by the precoding matrices that outperform the others in [25], [26]. All of the FPMB systems employ 4-QAM modulation, and the system data rate for SB and FPMB is set to 4, 8 bits/channel use for a 2×2 and a 4×4 system, respectively. All of the FPMB systems achieve the full diversity order because each slope is parallel to that of the corresponding SB system, which is known to achieve the full diversity order of NM.

In Fig. 4, we present simulation results that support the diversity analysis of 4×4 $S = 4$ PPMB in Table I. The theoretical results in Table I are duplicated in the legend of Fig. 4. It can be observed that the diversity orders in the simulation results are the same as those in the analysis.

To verify the computational complexity reduction with sphere detection in Section V, we simulated a 4×4 $S = 4$ FPMB system using 4-QAM and 64-QAM with receivers employing the exhaustive search (EXH), the conventional SD (CSD), and the proposed SD (PSD). In these simulations, the initial radius is chosen to be $\rho^2 = 2N_0P$, inside which at least one lattice point lies with a high probability [18]. The average number of real multiplications for decoding one transmitted vector symbol is calculated at different SNR. Since the reductions in complexity are substantial, we will express them as orders of magnitude (in approximate terms) in the sequel. In Fig. 5 we present the simulation results of the 4×4 $S = 4$ FPMB system. For 4-QAM, the number of multiplications of CSD is reduced by 1.4 and 2.1 orders of magnitude at low and high SNR, respectively. PSD reduces the complexity by 2.1 orders of magnitude at low SNR, and 2.4 at high SNR. The reduction becomes larger as the constellation size increases in the 4×4 $S = 4$ FPMB system. For 64-QAM, the number of multiplications of CSD decreases by 3.3 and 6.4 orders of magnitude at low and high SNR, respectively. PSD gives a larger reduction by 4.3 orders of magnitude at low SNR, and 7.0 at high SNR. Simulation results clearly show that CSD reduces the complexity substantially compared with EXH, and the complexity can be further reduced effectively by our PSD. The complexity reduction becomes larger as the constellation precoder dimension or the constellation size becomes larger. For comparison, simulation results for the 2×2 $S = 2$ FPMB system are available in [28].

B. BICMB-CP

We present simulation results for 2×2, 3×3, and 4×4 BICMB and BICMB-FP in Fig.6. The convolutional code employed is a 64-state one punctured from the $1/2$-rate mother code with generator polynomials $(133, 171)$ in octal representation. These results verify the diversity analysis in Section IV. Previously, in [8], we showed the maximum achievable diversity order of BICMB with an R_c-rate convolutional code is $(N - \lceil S \cdot R_c \rceil + 1)(M - \lceil S \cdot R_c \rceil + 1)$. As a result, in this example, the maximum achievable diversity order of the three

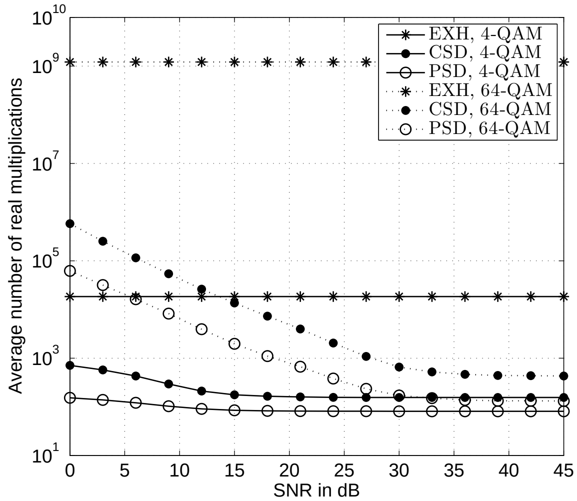


Fig. 5. Average number of real multiplications vs. SNR for the 4×4 FPMB systems with 4-QAM and 64-QAM.

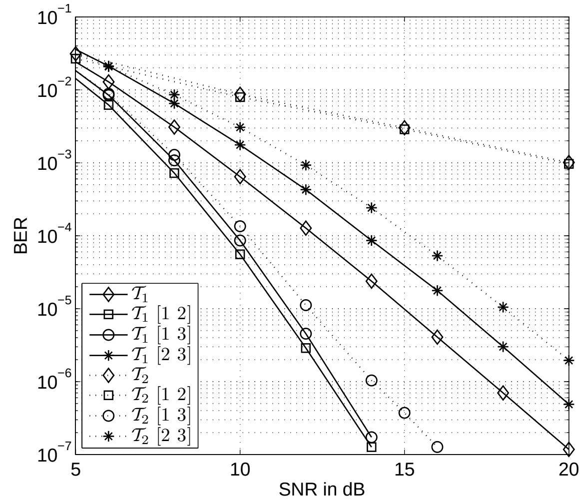


Fig. 7. BER vs. SNR for BICMB-PP with 3×3 $S = 3$, 4-QAM, and 4-state $1/2$-rate convolutional code.

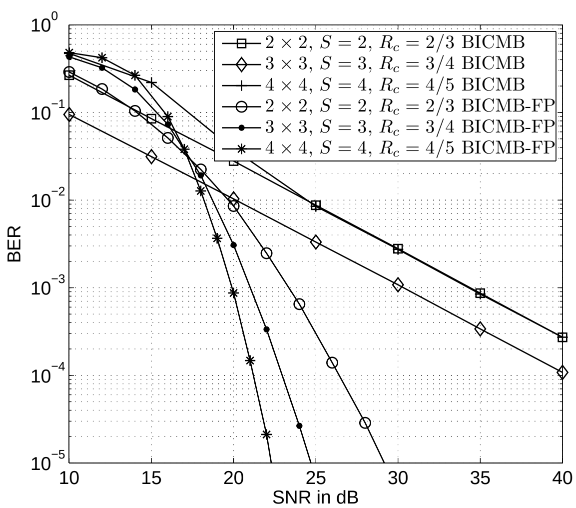


Fig. 6. BER comparison between BICMB and BICMB-FP with 16-QAM, and 64-state punctured convolutional code.

BICMB systems is 1. However, Fig. 6 shows that BICMB-FP achieves the full diversity order for any code rate.

In Fig. 7 we present the simulation results of BICMB-PP given in the example of Section III-B. The diversity orders of the BICMB systems, $\mathcal{T}_1$ and $\mathcal{T}_2$ are 4 and 1, respectively [8]. Comparing the slopes of BICMB-PP with BICMB, we see that the simulation results match the analysis in Section III-B.

To verify the proposed sphere decoding technique in this case for BICMB-FP, we simulated 4×4 $S = 4$, 64-state $R_c = 4/5$ BICMB-FP systems using 4-QAM and 64-QAM modulation with Gray mapping. The average number of real multiplications for acquiring one bit metric is calculated with receivers employing EXH, CSD, and PSD. Initial radii for both of CSD and PSD are determined by the ZF-DFE algorithm. Fig. 8 shows the number of multiplications of CSD for 4-QAM decreases by 1.3 and 1.5 orders of magnitude at low and high SNR, respectively. PSD gives bigger reductions by 2.1 orders of magnitude at low SNR, and 2.3 at high SNR. For the 64-QAM case, reductions between EXH and CSD by 3.2 and 4.4 orders of magnitude are observed at low and high SNR, respectively, while larger reductions by 4.2 and 5.4 are achieved by PSD. Similar to the uncoded case, the complexity reduction becomes larger as the constellation precoder dimension or the constellation size becomes larger. For comparison, simulation results for a 2×2 $S = 2$ 64-state $R_c = 2/3$ BICMB-FP system are available in [28].

One important property of our decoding technique needs to be emphasized: the substantial complexity reduction achieved causes no performance degradation.

VII. Conclusion

In this paper, we proposed constellation precoded multiple beamforming. This system achieves the full diversity order in both of the uncoded and coded MIMO multiple beamforming systems when the channel information is perfectly available at the transmitter as well as the receiver. This is achieved at different levels of spatial multiplexing, including the maximum $(\min(N, M))$ provided by the $N \times M$ channel. By employing the calculation of pairwise error probability and a theorem previously proved by the authors, an analysis of the diversity order was given for both of the multiple beamforming schemes. Examples of calculating the diversity orders of various multiple beamforming systems and simulation results supporting the analysis were given. To reduce the computational complexity of decoding, a sphere detection algorithm was proposed and simulation results were provided. The proposed SD algorithm in this paper can be applied to any MIMO system.

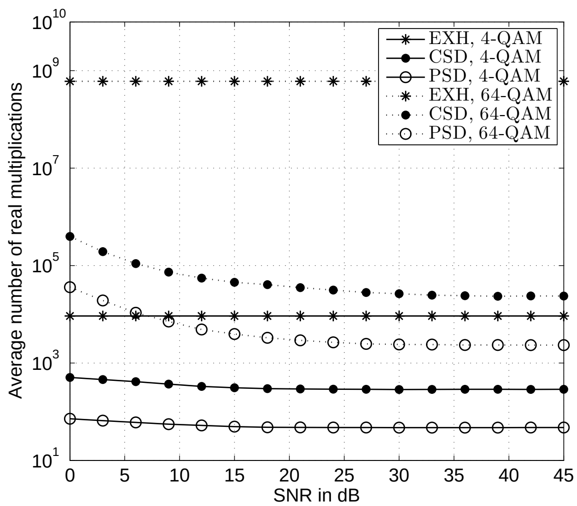


Fig. 8. Average number of real multiplications vs. SNR for the 4×4 BICMB-FP systems with 4-QAM and 64-QAM.

REFERENCES

[1] H. Jafarkhani, *Space-Time Coding: Theory and Practice*. Cambridge University Press, 2005.

[2] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, December 2001.

[3] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint tx-rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, September 2003.

[4] E. Sengul, E. Akay, and E. Ayanoglu, "Diversity analysis of single and multiple beamforming," *IEEE Trans. Commun.*, vol. 54, no. 6, pp. 990–993, June 2006.

[5] L. G. Ordonez, D. P. Palomar, A. Pages-Zamora, and J. R. Fonollosa, "High-SNR analytical performance of spatial multiplexing MIMO systems with CSI," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5447–5463, November 2007.

[6] E. Akay, E. Sengul, and E. Ayanoglu, "Bit interleaved coded multiple beamforming," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1802–1811, September 2007.

[7] E. Akay, H. J. Park, and E. Ayanoglu, "On bit-interleaved coded multiple beamforming," 2008, arXiv: 0807.2464. [Online]. Available: http://arxiv.org

[8] H. J. Park and E. Ayanoglu, "Diversity analysis of bit-interleaved coded multiple beamforming," in *Proc. IEEE ICC '09*, Dresden, Germany, June 2009.

[9] H. E. Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1097–1119, May 2003.

[10] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time diversity systems based on linear constellation precoding," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 294–309, March 2003.

[11] Z. Liu, Y. Xin, and G. B. Giannakis, "Linear constellation precoding for OFDM with maximum multipath diversity and coding gains," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 416–427, March 2003.

[12] W. Zhang, X.-G. Xia, and P. C. Ching, "High-rate full-diversity space-time-frequency codes for broadband MIMO block-fading channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 25–34, January 2007.

[13] N. Gresset and M. Khanfouci, "Precoded BICM design for MIMO transmit beamforming and associated low-complexity algebraic receivers," in *Proc. IEEE Globecom '08*, New Orleans, LA, November 2008.

[14] E. Zimmermann, W. Rave, and G. Fettweis, "On the complexity of sphere decoding," in *Proc. Wireless Personal Multimedia Communications (WPMC) '04*, Abano Terme, Italy, September 2004.

[15] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, April 2005.

[16] H. G. Han, S. K. Oh, S. J. Lee, and D. S. Kwon, "Computational complexities of sphere decoding according to initial radius selection schemes and an efficient initial radius reduction scheme," in *Proc. IEEE Globecom '05*, St. Louis, MO, November 2005, pp. 2354–2358.

[17] B. Cheng, W. Liu, Z. Yang, and Y. Li, "A new method for initial radius selection of sphere decoding," in *Proc. IEEE ISCC '07*, Aveiro, Portugal, July 2007, pp. 19–24.

[18] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. Expected complexity," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2806–2818, August 2005.

[19] W. Zhao and G. B. Giannakis, "Sphere decoding algorithms with improved radius search," *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1104–1109, July 2005.

[20] K.-W. Wong, C.-Y. Tsui, R. S.-K. Cheng, and W.-H. Mow, "A VLSI architecture of a K-Best lattice decoding algorithm for MIMO channels," in *Proc. IEEE ISCAS '02*, vol. 3, Scottsdale, Arizona, May 2002, pp. 273–276.

[21] T.-A. Huynh, D.-C. Hoang, M. R. Islam, and J. Kim, "Two-level-search sphere decoding algorithm for MIMO detection," in *Proc. IEEE ISWCS '08*, Reykjavik, Iceland, October 2008.

[22] J. Tang, A. H. Tewfik, and K. K. Parhi, "Reduced complexity sphere decoding and application to interfering IEEE 802.15.3a piconets," in *Proc. IEEE ICC '04*, vol. 5, Paris, France, June 2004.

[23] L. Azzam and E. Ayanoglu, "Reduced complexity sphere decoding for square QAM via a new lattice representation," in *Proc. IEEE Globecom '07*, Washington, D.C., November 2007.

[24] A. Zanella, M. Chiani, and M. Z. Win, "A general framework for the distribution of the eigenvalues of Wishart matrices," in *Proc. IEEE ICC '08*, May 2008, pp. 1271–1276.

[25] H. J. Park and E. Ayanoglu, "Constellation precoded beamforming," in *Proc. IEEE Globecom '09*, Honolulu, HI, November 2009.

[26] ——, "Constellation precoded beamforming," 2009, arXiv:0903.4738v1. [Online]. Available: http://arxiv.org

[27] B. Hassibi and H. Vikalo, "On the expected complexity of integer least-squares problems," in *Proc. IEEE ICASSP '02*, vol. 2, Orlando, FL, May 2002.

[28] H. J. Park, B. Li, and E. Ayanoglu, "Constellation precoded multiple beamforming," 2009, arXiv:0911.0709v1. [Online]. Available: http://arxiv.org