Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2008, Article ID 594928, 15 pages
doi:10.1155/2008/594928

Research Article

A Diversity Guarantee and SNR Performance for Unitary Limited Feedback MIMO Systems

Bishwarup Mondal and Robert W. Heath Jr.

Department of Electrical and Computer Engineering, The University of Texas at Austin, University Station C0803, Austin, TX 78712, USA

Correspondence should be addressed to Robert W. Heath Jr., rheath@ece.utexas.edu

Received 16 June 2007; Accepted 26 October 2007

Recommended by David Gesbert

A multiple-input multiple-output (MIMO) wireless channel formed by antenna arrays at the transmitter and at the receiver offers high capacity and significant diversity. Linear precoding may be used, along with spatial multiplexing (SM) or space-time block coding (STBC), to realize these gains with low-complexity receivers. In the absence of perfect channel knowledge at the transmitter, the precoding matrices may be quantized at the receiver and informed to the transmitter using a feedback channel, constituting a limited feedback system. This can possibly lead to a performance degradation, both in terms of diversity and array gain, due to the mismatch between the quantized precoder and the downlink channel. In this paper, it is proven that if the feedback per channel realization is greater than a threshold, then there is no loss of diversity due to quantization. The threshold is completely determined by the number of transmit antennas and the number of transmitted symbol streams. This result applies to both SM and STBC with unitary precoding and confirms some conjectures made about antenna subset selection with linear receivers. A closed form characterization of the loss in SNR (transmit array gain) due to precoder quantization is presented that applies to a precoded orthogonal STBC system and generalizes earlier results for single-stream beamforming.

1. INTRODUCTION

Linear precoding uses channel state information (CSI) at the transmitter to provide high data rates and improved diversity with low complexity receivers in multiple-input multiple-output (MIMO) wireless channels [1, 2]. The main idea of linear precoding is to customize the array of transmit signals by premultiplication with a spatial precoding matrix [3–8]. While precoding can be performed based on instantaneous CSI [9–19] or statistical CSI [20–23], the benefits are more in the instantaneous case assuming the CSI is accurate at the transmitter. Unfortunately, the system performance in terms of diversity and signal-to-noise ratio (SNR) depends crucially on the accuracy of CSI at the transmitter. In a *limited feedback* system, precoder information is quantized at the receiver and sent to the transmitter via a feedback channel [9–17]. In such a system quantization errors significantly impact the system performance and this motivates the present investigation.

Prior work

In this paper, we consider an important special case of precoding called unitary precoding that forms the basis of a limited feedback system. In this case, the precoder matrix has orthonormal columns, which incurs a small loss versus the nonunitary case especially in dense scattering environments (unitary precoding allocates power uniformly to all the selected eigenmodes and can be thought of as a generalization to antenna subset selection [24–27]) [28]. There have been several efforts at characterizing the diversity performance (measured in terms of the gain asymptotic slope of the average probability of error in Rayleigh fading channels versus SISO systems) of different limited feedback MIMO systems. The diversity of orthogonal space-time block coding with transmit antenna subset selection is analyzed in [27]. Spatial multiplexing systems with receive antenna selection with a capacity metric were considered in [29] and shown to achieve full diversity. In the case of a spatial multiplexing system

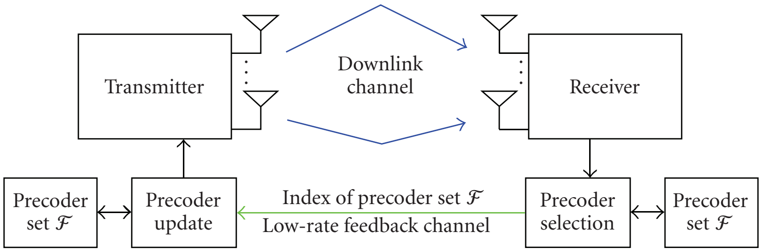

Figure 1: A quantized precoded MIMO system.

employing transmit antenna selection, conjectures on diversity order based on experimental evidence were presented in [30]. These conjectures were subsequently proved and generalized in [31]. In the special case of single-stream beamforming, the diversity order with limited feedback precoding was studied in [32] and a necessary and sufficient condition on the feedback rate for preserving full diversity is presented. A sufficient condition on the feedback rate for preserving diversity was derived for precoded orthogonal space-time block coding systems in [11, 33]. In the more complicated cases of limited feedback precoding in spatial multiplexing systems, experimental results were presented in [33, 34].

In summary, the diversity order for a quantized precoded spatial multiplexing system with linear receivers or a space-time block coding system (including nonorthogonal) is not characterized. This paper fills this gap by introducing an analysis approach based on matrix algebra and utilizing results from differential geometry. A sufficient condition on the number of feedback bits required per channel realization is derived that will guarantee full-CSI diversity for general limited feedback MIMO systems, which includes both spatial multiplexing as well as space-time block coding systems. The results for transmit antenna subset selection fall out as a special case.

An important implication of unitary precoding is the transmit array gain which is also affected due to precoder quantization. An analytical characterization of the loss in array gain due to quantization for single-stream beamforming in MISO systems was presented in [9, 16, 35] and the results for MIMO systems were presented by Mondal and Heath [36]. Analogous results, however, are not available for multistream transmission schemes. This paper takes a step forward by providing a closed-form characterization of the loss in array gain (or SNR of the received symbol) in the case of a precoded orthogonal space-time block coded system. This result simplifies to the beamforming scenario [9, 16, 35, 36] and naturally holds for antenna subset selection.

Detailed discussion of contributions

A pictorial description of a limited feedback system as considered in this paper is provided in Figure 1. A fixed, predetermined set of unitary precoding matrices is known to the transmitter and to the receiver. The receiver, for every instance of estimated downlink channel information, selects an element of the set and sends the index of the selected precoder to the transmitter using B bits of feedback. This precoder element is subsequently used by the transmitter for precoding. For analytic tractability we consider an uncorrelated Rayleigh flat-fading MIMO channel and we let M_t, M_r, and M_s denote the number of transmit antennas, receive antennas, and symbol streams transmitted, respectively. The uncorrelated Rayleigh channel is commonly used in rate distortion analysis for limited feedback systems [9, 12, 16, 35], including correlation along the lines of recent work is an interesting topic for future research [32]. Because discussions of diversity and array gain depend on transmitter and receiver structure, in this paper we consider explicitly two classes of systems—quantized precoded spatial-multiplexing (QPSM) and quantized precoded full-rank space-time block coding (QPSTBC) systems. A subclass of QPSTBC systems is due to orthogonal STBCs and is termed as QPOSTBC systems. The diversity analysis applies to both QPSM and QPSTBC systems, while the SNR result only applies to QPOSTBC systems. The detailed contributions of this paper may be summarized as follows.

(i) Diversity analysis: the diversity result applies to QPSTBC systems and to QPSM systems with zero-forcing (ZF) or minimum-mean-squared-error (MMSE) receivers. Leveraging a mathematical result due to Clark and Shekhtman [37] it is deduced that almost all (meaning with probability 1) sets of quantized precoding matrices, chosen at random, will guarantee no loss in diversity due to quantization if $2^B \geq M_s(M_t - M_s) + 1$. This is remarkable in the light of the fact that antenna subset selection known to preserve diversity in certain cases implies a feedback of $\log_2\binom{M_t}{M_s}$ bits which is an upper bound to $\log_2(M_s(M_t - M_s) + 1)$. This also means that for sufficiently large feedback, the design of the set of quantized precoders is irrelevant from the point of view of diversity.

(ii) SNR analysis: for a QPOSTBC system, the loss in SNR due to quantization reduces as $\sim 2^{-B/M_s(M_t-M_s)}$ with increasing feedback bits B. Thus, most of the channel gain is obtained at low values of feedback rate (bits per channel realization) and increasing feedback further leads to insignificant gains. Our characterization also shows that increasing M_r provides robustness to quantization error. Single-stream beamforming or maximum-ratio transmission and combining (MRT-MRC) results of [36] fall out as a special case of this result.

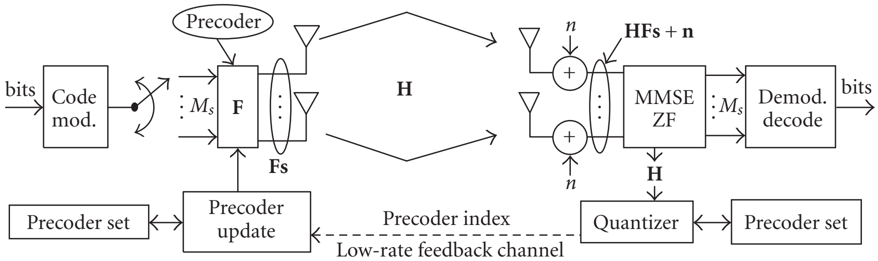

FIGURE 2: Discrete-time quantized precoded MIMO spatial multiplexing system.

This paper is organized as follows. The system model is described and the assumptions are mentioned in Section 2. The diversity of such systems and the effective channel gain are analyzed in Sections 3 and 4, respectively, before the results are summarized in Section 5.

Notation. Matrices are in bold capitals, vectors are in bold lower case. We use H to denote conjugate transpose, $\|\cdot\|_F$ to denote the Frobenius norm, $\|\cdot\|_2$ to denote matrix 2-norm, $[\mathbf{A}]_{ij}$ to denote the (i,j)th element of the matrix $\mathbf{A}$, -1 to denote matrix inverse, $\overset{d}{=}$ to denote equality in distribution, $\mathbf{I}$ to denote the identity matrix and, $E\{\cdot\}$ to denote expectation. We also denote the trace of $\mathbf{A}$ by $\text{tr}(\mathbf{A})$, the rank of $\mathbf{A}$ by $\text{rank}(\mathbf{A})$, a diagonal matrix with $\lambda_1, \lambda_2, \ldots, \lambda_n$ as its diagonal entries starting with the top left element by $\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. $\lambda_{\min}(\mathbf{A})$ denotes the minimum eigenvalue of the matrix $\mathbf{A}$. $\pi \oplus \omega$ denotes the direct sum of the subspaces π and ω of the space χ meaning $\chi = \pi + \omega$ and $\pi \cap \omega = \{0\}$. $\mathcal{CN}(0, N_0)$ denote a complex normal distribution with zero mean and N_0 variance with i.i.d. real and imaginary parts.

2. SYSTEM OVERVIEW

In this section, a precoded spatial multiplexing system and a precoded space-time block coding system, both with precoder quantization and feedback, are described. Then a brief motivation is provided for unitary precoding assuming perfect CSI at the transmitter. Subsequently limited feedback precoding is introduced and formulated as a quantization problem. Finally the main assumptions of the paper are summarized.

2.1. Quantized precoded spatial multiplexing system (QPSM)

As shown in Figure 2, in a spatial multiplexing system a single data stream is modulated before being demultiplexed into m_s symbol streams. This produces a symbol vector $\mathbf{s}$ of length m_s for a symbol period. We assume that $E\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}$. The symbol vector $\mathbf{s}$ is spread over M_t antennas by multiplying it with an $M_t \times M_s$ precoding matrix $\mathbf{F}$, where $M_s = m_s$. This process of linear precoding produces an M_t length vector $\mathbf{Fs}$ that is transmitted using M_t antennas. Then the discrete-time equivalent signal model for one symbol period at baseband with perfect synchronization can be written as

$$\mathbf{y} = \sqrt{\frac{E_s}{M_s}}\mathbf{HFs} + \mathbf{n}, \tag{1}$$

where $\mathbf{y}$ is the received signal vector at the M_r received antennas, E_s is the energy for one symbol period, $\mathbf{H}$ is a matrix with complex entries that represents the channel transfer function, and $\mathbf{n}$ represents an additive white Gaussian noise (AWGN) vector. For a QPSM system we assume $M_t > M_s$, $M_r \geq M_s$. In this paper we only concentrate on ZF and MMSE receivers that enable low-complexity implementation.

We also consider a fixed predetermined set of precoding matrices $\mathcal{F} = \{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_N\}$ that is known to both the transmitter and the receiver. Depending on the channel realization $\mathbf{H}$, the receiver selects an element of $\mathcal{F}$ and informs the transmitter of the selection through a feedback link. Note that $\lceil \log_2 N \rceil$ bits are sufficient to identify a precoding matrix in $\mathcal{F}$.

2.2. Quantized precoded STBC system (QPSTBC)

The second class of systems under consideration uses precoding along with space-time block coding as illustrated in Figure 3. At the transmitter, after the bit stream is modulated using a constellation of symbols, a block of m_s symbols $s_1, s_2, \ldots, s_{m_s}$ is mapped to construct a space-time code matrix $\mathbf{C}$. The code matrix $\mathbf{C}$ is of dimension $M_s \times T$ and this code is premultiplied by an $M_t \times M_s$ precoding matrix $\mathbf{F}$, resulting in a matrix $\mathbf{FC}$. Thus $\mathbf{FC}$ spreads over M_t antennas and T symbol periods. The channel matrix $\mathbf{H}$ is assumed to be constant for the T symbol periods and changes randomly in the next symbol period. The discrete-time baseband signal model for T symbol periods may be written as

$$\mathbf{Y} = \sqrt{\frac{E_s}{M_s}}\mathbf{HFC} + \mathbf{N}, \tag{2}$$

where $\mathbf{Y}$ is the received signal at the M_r receive antennas over T symbol periods, E_s is the energy over one symbol period, and $\mathbf{N}$ is the AWGN at the receiver for T symbol periods. We assume $M_t > M_s$, but there is no restriction on M_r. As before, we consider a set of precoding matrices $\mathcal{F}$ known to both the transmitter and the receiver. The receiver chooses an

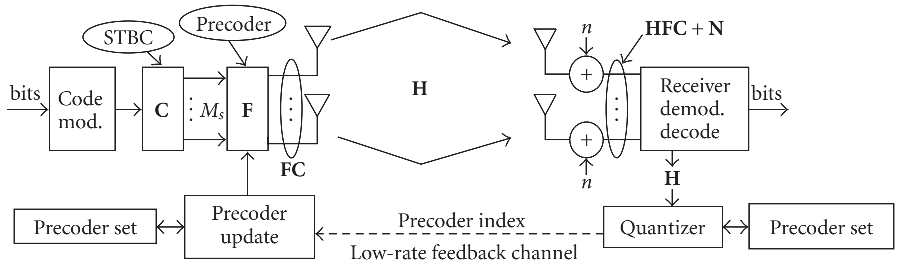

Figure 3: Discrete-time MIMO system with quantized precoded STBC.

element of $\mathcal{F}$ depending on $\mathbf{H}$ and sends this information to the transmitter using a feedback link. As mentioned before, we restrict ourselves to full-rank STBCs for which,

$$\lambda_{\min}\left(\mathbf{E}_{ij}\mathbf{E}_{ij}^H\right) > 0 \quad \forall i \neq j, \tag{3}$$

where $\mathbf{E}_{ij} = \mathbf{C}_i - \mathbf{C}_j$ is the codeword difference matrix between the ith and the jth block code. Full-rank STBCs encompass a wide variety of codes differing in rate and complexity, including orthogonal STBCs [38, 39], STBCs from division algebras [40], space-time group codes [41], and quasiorthogonal STBCs modified using rotation [42] or power-allocation [43]. A special class of QPSTBC systems is characterized by the property

$$\mathbf{C}\mathbf{C}^H = \left\{\sum_{i=1}^{m_s} |s_i|^2\right\}\mathbf{I}, \tag{4}$$

where $\mathbf{C}$ is a space-time code matrix. This implies that $\mathbf{C}$ is an orthogonal STBC [38, 39] and such systems form a subclass termed as QPOSTBC systems. In our analysis, an ML receiver is assumed for all QPSTBC systems.

Precoding for the special case of $M_s = 1$ represents beamforming where a single symbol is spread over M_t antennas by the beamforming vector. The ML receiver, in this case, becomes a maximum-ratio combiner (MRC).

2.3. Limited feedback unitary precoding

In the following sections, it will be of interest to define a perfect-CSI precoding matrix (or a precoding matrix with infinite feedback bits) as

$$\mathbf{F}_\infty = \overline{\mathbf{U}}, \tag{5}$$

where $\mathbf{H}^H\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^H$ denote the SVD of $\mathbf{H}^H\mathbf{H}$, $\boldsymbol{\Sigma} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{M_t})$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{M_t} \geq 0$ are the ordered eigenvalues of $\mathbf{H}^H\mathbf{H}$ and $\overline{\mathbf{U}}$ denotes the $M_t \times M_s$ submatrix of $\mathbf{U}$ with columns corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_{M_s}$. Thus $\mathbf{F}_\infty^H\mathbf{F}_\infty = \mathbf{I}$ such that $\mathbf{F}_\infty$ is tall and unitary. (The term unitary is used in a generic sense to represent matrices with the property $\mathbf{A}^H\mathbf{A} = \mathbf{I}$ where $\mathbf{A}$ can be either tall or square.)

At the receiver, corresponding to a channel realization $\mathbf{H}$, a precoding matrix is chosen from the set $\mathcal{F}$. This selection may be described by a map $\mathcal{Q}$ such that $\mathcal{Q}(\mathbf{F}_\infty) \in \mathcal{F}$, where $\mathbf{F}_\infty$ is obtained from $\mathbf{H}$ using (5). The map $\mathcal{Q}$ may also be visualized as a quantization process applied to the set of all perfect-CSI precoding matrices. Then borrowing vector quantization terminology, the map $\mathcal{Q}$ is a quantization function, $\mathbf{F}_\infty$ is the source random matrix, $\mathcal{F}$ is a codebook, elements of $\mathcal{F}$ are codewords (or quantization levels), and the cardinality of $\mathcal{F}$ is the number of quantization levels or the quantization rate. This justifies the "Q" in QPSM and QPSTBC systems. The quantization function $\mathcal{Q}$ is also referred to as the *precoder selection criterion* in the literature and we will use these terms interchangeably in this paper. It may be noted that assuming a feedback of $\lceil \log_2 N \rceil$ per channel realization $\mathbf{H}$, the precoding matrix $\mathbf{F}$ in (1) and (2) becomes an element of $\mathcal{F}$ chosen by a precoder selection criterion described by $\mathbf{F} = \mathcal{Q}(\mathbf{F}_\infty)$.

Antenna subset selection at the transmitter may be considered as a special case of quantized precoding [24–27]. In this case, the elements of $\mathcal{F}$ are submatrices of the $M_t \times M_t$ identity matrix. In particular, every combination of M_s columns of the identity matrix forms an element of $\mathcal{F}$ and thus $\operatorname{card}(\mathcal{F}) = \binom{M_t}{M_s}$.

2.4. Assumptions

The assumptions in this paper are summarized as follows. The elements of $\mathcal{F}$ are unitary implying $\mathbf{F}_i^H\mathbf{F}_i = \mathbf{I}$ for $i = 1, 2, \ldots, N$. The channel is uncorrelated Rayleigh fading and the elements of $\mathbf{H}$ are distributed as i.i.d. $\mathcal{CN}(0, 1)$. The i.i.d. assumption is typically used for the analysis of limited feedback systems [9, 12, 27, 31] mainly due to the tractable nature of the eigenvalues and eigenvectors in this case. The elements of $\mathbf{n}, \mathbf{N}$ represents AWGN, are distributed as i.i.d. $\mathcal{CN}(0, N_0)$. The feedback link is assumed to be error-free and having zero-delay, and we assume perfect channel knowledge at the receiver.

3. SUFFICIENT CONDITION FOR NO DIVERSITY LOSS

A concern for QPSM and QPSTBC systems is whether the diversity order is reduced due to quantization. The objective of this section is to provide a sufficient condition that will guarantee no loss in diversity due to precoder quantization for such systems.

As evidenced by simulation results it turns out (this will be proved in the following) that the diversity order of QPSM and QPSTBC systems does not change if, corresponding to a given channel realization $\mathbf{H}$, the precoding matrix $\mathbf{F}$ is substituted by $\mathbf{FQ}$, where $\mathbf{Q}$ is an arbitrary unitary matrix.

This motivates the representation of the precoding matrix $\mathbf{F}$ as a point on the complex Grassmann manifold which is introduced in the next subsection. In the following we outline a strategy for the proof and introduce the projection 2-norm distance and the chordal distance as analysis tools. In the course of the analysis, a special class of codebooks called *covering codebooks* is defined that satisfies a certain condition on its covering radius (measured in terms of projection 2-norm distance). It is proven that a covering codebook can guarantee full-CSI diversity for both QPSM and QPSTBC systems in Corollaries 1 and 2, respectively. These form the main results in this section of the paper. Finally, a connection between the covering radius characterization and the covering-by-complements problem [37] is discovered that allows us to identify the class of covering codebooks that can be employed in real systems, thereby preserving full diversity.

3.1. The complex grassmann manifold

First we provide an intuitive understanding of the complex Grassmann manifold similar to [44]. The complex Grassmann manifold denoted by $G_{n,p}$ is the set of all p-dimensional linear subspaces of $\mathbb{C}^n$. An element in $G_{n,p}$ is a linear subspace and may be represented by an arbitrary basis spanning the subspace. Given any n -by-p tall unitary matrix ($n > p$), the subspace spanned by its columns forms an element in $G_{n,p}$. Corresponding to a given precoding matrix $\mathbf{F}$ of dimensions $M_t \times M_s$, we can associate an element $\omega \in G_{M_t,M_s}$ such that ω is the column space of $\mathbf{F}$. We can explicitly write this relation as $\omega(\mathbf{F}) \in G_{M_t,M_s}$. Also since a rotation of the basis does not change its span, $\omega(\mathbf{FQ})$ is the same element in G_{M_t,M_s} for all M_s-by-M_s unitary matrices $\mathbf{Q}$. This models the fact that the precoding matrices $\mathbf{F}$ and $\mathbf{FQ}$ provide the same diversity irrespective of any $\mathbf{Q}$.

3.2. Proof strategy

This subsection provides an intuitive sketch of the proof ideas and not a rigorous treatment. In order to implement a limited feedback system, a precoder selection criterion $\mathcal{Q}$ needs to be in place. The choice of $\mathcal{Q}$ depends on the performance metric (e.g., SNR, capacity) and system parameters like the receiver type. The precoder selection criteria assumed in this paper for different systems are denoted by $\mathcal{Q}^*$ and mentioned in (14), (16), (17) and they target bit-error rate as the system performance metric.

To prove the diversity results, as a mathematical tool, we define another precoder selection criterion as

$$\mathcal{Q}_P(\mathbf{F}_\infty) = \arg\min_{\mathbf{F}_k \in \mathcal{F}} d_P(\mathbf{F}_\infty, \mathbf{F}_k), \tag{6}$$

where $d_P(\cdot,\cdot)$ is the projection 2-norm distance and is defined as [44]

$$d_P(\mathbf{F}_1, \mathbf{F}_2) = ||\mathbf{F}_1\mathbf{F}_1^H - \mathbf{F}_2\mathbf{F}_2^H||_2, \tag{7}$$

where $\mathbf{F}_1$ and $\mathbf{F}_2$ are two arbitrary precoding matrices of the same dimensions. Observe that $d_P(\mathbf{F}_1,\mathbf{F}_2) = d_P(\mathbf{F}_1\mathbf{Q}_1, \mathbf{F}_2\mathbf{Q}_2)$ for arbitrary unitary matrices $\mathbf{Q}_1$, $\mathbf{Q}_2$, and thus intuitively $d_P(\cdot,\cdot)$ can be used to measure the distance between $\omega(\mathbf{F}_1)$ and $\omega(\mathbf{F}_2)$ on G_{M_t,M_s}. It turns out that the $d_P(\cdot,\cdot)$ is a distance measure in G_{M_t,M_s}. The proofs leading up to the diversity results follow in two steps: (i) first, we assume that $\mathcal{Q}_P$ is used as the precoder selection criterion and prove that the diversity result is true for such a system; (ii) second, if $\mathcal{Q}^*$ as defined in (14), (16), (17) is used instead of $\mathcal{Q}_P$, the diversity performance of the system is identical or better, thus the result is true for systems using $\mathcal{Q}^*$. Note that in a real system, a precoder will be chosen based on $\mathcal{Q}^*$ and we prove our results for such a system. The introduction of $\mathcal{Q}_P$ is a mathematical tool and is not intended to be used in a real system.

Analogously for the SNR results, we introduce another precoder selection criterion expressed as

$$\mathcal{Q}_C(\mathbf{F}_\infty) = \arg\min_{\mathbf{F}_k \in \mathcal{F}} d_C(\mathbf{F}_\infty, \mathbf{F}_k), \tag{8}$$

where $d_C(\cdot,\cdot)$ is the chordal distance [44]

$$d_C(\mathbf{F}_1, \mathbf{F}_2) = ||\mathbf{F}_1\mathbf{F}_1^H - \mathbf{F}_2\mathbf{F}_2^H||_F, \tag{9}$$

where $\mathbf{F}_1$ and $\mathbf{F}_2$ are two arbitrary precoding matrices of the same dimensions. $d_C(\cdot,\cdot)$ is also a distance metric in G_{M_t,M_s}. The proofs for the SNR results in Section 4 follow the following steps: (i) if $\mathcal{Q}_C$ is used as the precoder selection criterion, then the SNR result is true; (ii) if $\mathcal{Q}^*$ as defined in (18) is used instead of $\mathcal{Q}_C$, the SNR performance of the system is identical for sufficiently large number of bits of feedback. Again it is worth mentioning that $\mathcal{Q}_C$ is introduced to aid analysis and is not intended to be used in a real system. (The introduction of $d_P(\cdot,\cdot)$ and $d_C(\cdot,\cdot)$ simplifies the proofs for diversity and SNR respectively but we were unable to discover any fundamental reason behind this. It is mentioned in passing that the distance measures $d_P(\cdot,\cdot)$ and $d_C(\cdot,\cdot)$ coincide in $G_{M_t,1}$, $d_P(\cdot,\cdot) \le d_C(\cdot,\cdot)$ and $d_P(\cdot,\cdot) \approx d_C(\cdot,\cdot)$ when either is close to zero.)

3.3. Covering codebook

The notion of a covering codebook is another mathematical aid. Covering codebooks define a subset of all possible codebooks and we show later that a covering codebook along with a precoder selection criterion $\mathcal{Q}^*$ is sufficient to guarantee full-CSI diversity. Note that, in a real system, a codebook may be designed according to various criteria [33, 34]; but according to a result in [37], it is deduced that any codebook, with a certain cardinality or higher but chosen at random, is a covering codebook with probability 1. In the following we show that a covering radius characterization of a codebook is equivalent to a covering-by-complements by the codebook in a complex Grassmann manifold.

Theorem 1. *The following are equivalent.*

(i) *The covering radius δ of $\mathcal{F} = \{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_N\}$ in terms of the projection 2-norm distance is strictly less than unity. This is expressed as*

$$\delta = \sup_{\mathbf{F}} \min_{\mathbf{F}_k} d_P(\mathbf{F}, \mathbf{F}_k) < 1, \tag{10}$$

where $\mathbf{F}_k \in \mathcal{F}$ and $\mathbf{F} \in G_{M_t,M_s}$.

(ii) *The complements of the elements of* $\mathcal{F}$ *provide a covering for* G_{M_t,M_t-M_s}. *This may be written as* $c(\mathbf{F}_1) \cup c(\mathbf{F}_2) \cup \cdots \cup c(\mathbf{F}_N) = G_{M_t,M_t-M_s}$, *where* $c(\mathbf{F}_k)$ *is the complement of* $\mathbf{F}_k$ *defined as* $c(\mathbf{F}_k) = \{\pi : \pi \in G_{M_t,M_t-M_s}, \pi \oplus \omega(\mathbf{F}_k) = \mathbb{C}^{Mt}\}$.

Proof. See Appendix A. □

Now let us define a codebook $\mathcal{F}$ with a covering radius strictly less than unity (that satisfies (10)) as a *covering codebook*. Since $d_P(\mathbf{F}, \mathbf{F}_k)$ takes values in $[0, 1]$, it is intuitive that a codebook, chosen at random, will be a covering codebook with probability 1. This is proved in the work by Clark and Shekhtman [37]. They have studied the problem of covering-by-complements for vector spaces over algebraically closed fields. Since $\mathbb{C}$ is algebraically closed, it follows from [37] that the least cardinality of $\mathcal{F}$ to be a covering codebook is $M_s(M_t - M_s) + 1$. It also follows from [37] that *almost all* (in probability sense) codebooks of cardinality larger than $M_s(M_t - M_s) + 1$ are covering codebooks.

3.4. Diversity of QPSM with linear receivers

The diversity of a QPSM or a QPSTBC system is the slope of the symbol-error-rate curve for asymptotically large SNRs defined as a limit expressed by

$$d = -\lim_{E_s/N_0 \to \infty} \frac{\log P_e}{\log E_s/N_0}, \tag{11}$$

where P_e is the probability of symbol error. Here we consider a QPSM system and focus on a ZF receiver. A ZF receive filter given by

$$\mathbf{G}^{(\mathrm{ZF})} = \left[\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}\right]^{-1}\mathbf{F}^H\mathbf{H}^H \tag{12}$$

is applied to the received signal vector $\mathbf{y}$ in (1) and the resulting M_s data streams (corresponding to $\mathbf{G}\mathbf{y}$) are independently detected. The postprocessing SNR for the ith data stream after receiving ZF filtering is given by [30]

$$\mathrm{SNR}_i^{(\mathrm{ZF})}(\mathbf{F}) = \frac{E_s}{M_s N_0 \left[\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}\right]_{ii}^{-1}}. \tag{13}$$

In the following we assume that the precoder selection criterion $\mathcal{Q}^*$ maximizes the minimum postprocessing substream SNR. The following result summarizes the diversity characteristics of such a QPSM-ZF system.

Corollary 1. *Assume a QPSM system with a ZF receiver and a precoder selection criterion given by*

$$\mathcal{Q}^*(\mathbf{F}_\infty) = \arg\max_{\mathbf{F}_k \in \mathcal{F}} \min_i \mathrm{SNR}_i^{(\mathrm{ZF})}(\mathbf{F}_k). \tag{14}$$

Then if $\mathcal{F}$ *is a covering codebook, the precoder* $\mathcal{Q}^*(\mathbf{F}_\infty)$ *provides the same diversity as provided by* $\mathbf{F}_\infty$.

Proof. The proof of Corollary 1 proceeds in two stages as described in Section 3.2. We prove that a precoder chosen from $\mathcal{F}$ according to $\mathcal{Q}_P$ as in (6) provides the same diversity as $\mathbf{F}_\infty$. Then we show that $\mathcal{Q}^*$ given by (14) provides a better diversity performance than $\mathcal{Q}_P$; for a detailed proof see Appendix B. □

Corollary 1 states that a covering codebook preserves the diversity order of a precoded spatial multiplexing system with a ZF receiver. (It is worth mentioning that the diversity order of a precoded spatial multiplexing system (using $\mathbf{F}_\infty$ as the precoder) with a ZF receiver is not available. As a supplementary result we establish the diversity order of such a system with the restriction $M_r = M_s$ in Appendix E.) An important example of a covering codebook is due to antenna subset selection. It is straightforward to show the following.

Lemma 1. *The antenna selection codebook of cardinality* $\binom{M_t}{M_s}$ *is a covering codebook.*

Proof. See Appendix C. □

It directly follows from Lemma 1 and Corollary 1 that transmit antenna subset selection for spatial-multiplexing systems with a ZF receiver can guarantee full-CSI diversity [31]. An MMSE receive filter converges to a ZF filter for high values of E_s/N_0 leading to the common understanding that both receivers achieve the same diversity order. This implies that the results presented above also apply to MMSE receivers.

3.5. Diversity of QPSTBC systems

Recall that in a QPSTBC system (2) the difference codewords $\mathbf{E}_{ij} = \mathbf{C}_i - \mathbf{C}_j$, $i \neq j$ are full rank. It is known that these systems provide a diversity order of $M_t M_r$. QPOSTBC systems are a subset of QPSTBC systems where $\mathbf{E}_{ij} = \alpha\mathbf{I}$, $i \neq j$, and $\alpha \in \mathbb{C}$. The Chernoff bound for pairwise error probability (PEP) for a QPSTBC system may be expressed as [45]

$$P(\mathbf{C}_i \longrightarrow \mathbf{C}_j \mid \mathbf{H}) \leq e^{-(E_s/N_0)\|\mathbf{H}\mathbf{F}\mathbf{E}_{ij}\|_F^2}, \tag{15}$$

where $P(\mathbf{C}_i \to \mathbf{C}_j \mid \mathbf{H})$ is the probability of detecting $\mathbf{C}_j$ given, $\mathbf{C}_i$ is transmitted and the channel realization being $\mathbf{H}$. From the expression of PEP (15) a precoder selection criterion can be obtained that minimizes the Chernoff bound. The following corollary assumes such a criterion and summarizes the diversity characterization.

Corollary 2. *Assume a QPSTBC system where the difference codewords are full rank and the precoder selection criterion is given by*

$$\mathcal{Q}^*(\mathbf{F}_\infty) = \arg\max_{\mathbf{F}_k \in \mathcal{F}} \min_{i,j} \left|\left|\mathbf{H}\mathbf{F}_k\mathbf{E}_{ij}\right|\right|_F^2. \tag{16}$$

Then if $\mathcal{F}$ *is a covering codebook, the precoder* $\mathcal{Q}^*(\mathbf{F}_\infty)$ *provides the same diversity as provided by* $\mathbf{F}_\infty$.

Proof. The proof of Corollary 2 proceeds in a way similar to Corollary 1 by assuming a precoder selection criterion $\mathcal{Q}_P$ given by (6) and then showing that $\mathcal{Q}^*$ given by (16) provides a diversity performance better than that by $\mathcal{Q}_P$; for a detailed proof see Appendix D. □

It may be noted that in the particular case of QPOSTBC, it easily follows from (16) that the precoder selection criterion simplifies to

$$\mathcal{Q}^*(\mathbf{F}_\infty) = \arg\max_{\mathbf{F}_k \in \mathcal{F}} ||\mathbf{H}\mathbf{F}_k||_F^2, \tag{17}$$

and from Corollary 2 it follows that a covering codebook provides full diversity. The special case of QPOSTBC has also been studied in [33] and a sufficient condition for preserving full diversity was derived. It follows from Corollary 2 and Lemma 1that a full-rank STBC system with transmit antenna subset selection is guaranteed to achieve full diversity.

3.6. Observations

It is proven that precoder selection criteria motivated by postprocessing SNR and the Chernoff bound on PEP preserve diversity order. This is a pleasing result for system designers. Diversity can be guaranteed by a codebook chosen at random of size determined only by M_t and M_s. The structure in the codebook or a particular element of a codebook is irrelevant and thus codebook design algorithms need not consider diversity as a criterion. It is also interesting to note that diversity can be preserved with less feedback than that for antenna subset selection.

4. CHARACTERIZATION OF SNR LOSS

The objective of this section is to quantify the loss in expected SNR of a received symbol due to quantization for a QPOSTBC system as a function of the feedback bits B or for convenience $N = 2^B$.

4.1. Relation of SNR loss with chordal distortion

Following the system model in (2) and considering a QPOSTBC system, the expected SNR for a received symbol may be written as $E\{\|\mathbf{HF}\|_F^2\}(E_s/M_sN_0)$. This naturally leads to a precoder selection criterion that maximizes the expected SNR and is expressed by

$$\mathcal{Q}^*(\mathbf{F}_\infty) = \arg\max_{\mathbf{F}_k \in \mathcal{F}} ||\mathbf{H}\mathbf{F}_k||_F^2. \tag{18}$$

Notice that the expected SNR of a system using a precoder $\mathbf{F}$ does not change if $\mathbf{F}$ is substituted by $\mathbf{FQ}$, where $\mathbf{Q}$ is an arbitrary square unitary matrix (of dimension $M_s \times M_s$). This fact, similar to the case of diversity, justifies the representation of a precoding matrix on a complex Grassmann manifold. Recall from Section 3.2 that the first step in the proof is to consider a precoder selection criterion based on $d_C(\cdot,\cdot)$ given by (8). Then we have the following result.

Theorem 2. *Assume a precoder selection criterion given by*

$$\mathcal{Q}_C(\mathbf{F}_\infty) = \arg\min_{\mathbf{F}_k \in \mathcal{F}} d_C(\mathbf{F}_\infty, \mathbf{F}_k). \tag{19}$$

Then

$$E\{||\mathbf{HF}_\infty||_F^2\} - E\{||\mathbf{H}\mathcal{Q}(\mathbf{F}_\infty)||_F^2\} = (\overline{\Lambda} - \underline{\Lambda})E\{d_C^2(\mathbf{F}_\infty, \mathcal{Q}(\mathbf{F}_\infty))\}, \tag{20}$$

where $\overline{\Lambda} = (1/M_s)\sum_{i=1}^{M_s} E\{\lambda_i\}$, $\underline{\Lambda} = (1/(M_t - M_s))\sum_{i=M_s+1}^{M_t} E\{\lambda_i\}$, *where* $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_{M_t} \geq 0$, *are the ordered eigenvalues of* $\mathbf{H}^H\mathbf{H}$.

Proof. See Appendix F. □

An intuitive understanding of the final SNR result follows directly from Theorem 2. It follows from a result in [46–48] that (8) defines a quantization problem with a distortion function as $d_C^2(\cdot,\cdot)$ and the expected distortion, $E\{d_C^2(\mathbf{F}_\infty, \mathcal{Q}(\mathbf{F}_\infty))\} \sim N^{-1/M_s(M_t - M_s)}$ in the asymptotic regime of large N. Then it follows from (20) that the SNR loss due to quantization also decays as $\sim N^{-1/M_s(M_t-M_s)}$. Now, as part of the second step of the proof, it is easy to see that the precoder selection criterion (18) results in an equal or better SNR compared to (19). Thus with (18), the SNR loss due to quantization decays at least as fast as $\sim N^{-1/M_s(M_t-M_s)}$. A precise set of arguments follows and our final result is presented in the following subsection.

4.2. Asymptotic characterization of SNR

Theorem 2 shows that the loss in expected SNR due to precoder quantization can be exactly captured by the expected chordal distance between $\mathbf{F}_\infty$ and its quantized version *assuming a precoder selection criterion given by* (8). Note that $E\{d_C^2(\mathbf{F}_\infty, \mathcal{Q}(\mathbf{F}_\infty))\}$ is the expected distortion for the quantization function $\mathcal{Q}$ defined by (8). This class of quantization problems with chordal distortion has been studied in [46–48]. In the particular case of an uncorrelated Rayleigh fading channel the probability distribution of $\mathbf{F}_\infty$ is known [49]. A lower bound on the expected distortion $E\{d^2(\mathbf{F}_\infty, \mathcal{Q}(\mathbf{F}_\infty))\}$ is derived in [36] for large N which takes the form

$$E\{d^2(\mathbf{F}_\infty, \mathcal{Q}(\mathbf{F}_\infty))\} \geq \left(\frac{M_s(M_t - M_s)}{M_s(M_t - M_s) + 2}\right)\{c(M_t, M_s)N\}^{-1/M_s(M_t-M_s)}, \tag{21}$$

where $c(M_t, M_s)$ is a constant and may be expressed as $c(M_t, M_s) = (1/(M_tM_s - M_s^2)!)\prod_{i=1}^{M_s}((M_t - i)!/(M_s - i)!)$ for $M_s \leq M_t/2$ and $c(M_t, M_s) = (1/(M_tM_s - M_s^2)!)\prod_{i=1}^{M_s}((M_t - i)!/(M_t - M_s - i)!)$ otherwise. Thus for large N and with precoder selection criterion given by (8) we can write

$$E\{||\mathbf{H}\mathcal{Q}(\mathbf{F}_\infty)||_F^2\} \leq E\{||\mathbf{HF}_\infty||_F^2\} - KN^{-1/M_s(M_t-M_s)}, \tag{22}$$

where K is independent of N and may be obtained from (20) and (21). It is also known from quantization theory [47, 50] that there exits a sequence of codebooks of cardinality $1, 2, \ldots, N, N + 1, \ldots$ such that

$$\lim_{N\to\infty} E\{d_C^2(\mathbf{F}_\infty, \mathcal{Q}(\mathbf{F}_\infty))\} = 0 \Longrightarrow \lim_{N\to\infty} E\{||\mathbf{H}\mathcal{Q}(\mathbf{F}_\infty)||_F^2\} = E\{||\mathbf{HF}_\infty||_F^2\}. \tag{23}$$

It directly follows from (23) that for sufficiently large N, the left-hand side and the right-hand side of (22) is contained within a ball of radius $\epsilon > 0$.

It is easy to observe that the precoder selection criterion given by (8), in general, does not maximize $E\{\|\mathbf{H}\mathcal{Q}(\mathbf{F}_\infty)\|_F^2\}$. On the other hand, a precoder selection criterion given by

$$\overline{\mathcal{Q}}(\mathbf{F}_\infty) = \arg\max_{\mathbf{F}_k \in \mathcal{F}} \|\mathbf{H}\mathbf{F}_k\|_F^2 \tag{24}$$

maximizes $E\{\|\mathbf{H}\overline{\mathcal{Q}}(\mathbf{F}_\infty)\|_F^2\}$. It is easy to see that for any given codebook $\mathcal{F}$, we have

$$E\{\|\mathbf{H}\mathcal{Q}(\mathbf{F}_\infty)\|_F^2\} \leq E\{\|\mathbf{H}\overline{\mathcal{Q}}(\mathbf{F}_\infty)\|_F^2\} \leq E\{\|\mathbf{H}\mathbf{F}_\infty\|_F^2\}; \tag{25}$$

and using the same sequence of codebooks as before, we have from (23) and (25)

$$\lim_{N\to\infty} E\{\|\mathbf{H}\overline{\mathcal{Q}}(\mathbf{F}_\infty)\|\} = E\{\|\mathbf{H}\mathbf{F}_\infty\|\}. \tag{26}$$

It follows from (22), (23), and (26) that for sufficiently large N,

$$\begin{aligned}\sup_{\mathcal{F}:\,\mathrm{card}(\mathcal{F})=N} & E\{\|\mathbf{H}\overline{\mathcal{Q}}(\mathbf{F}_\infty)\|\} \\ & \approx E\{\|\mathbf{H}\mathbf{F}_\infty\|_F^2\} - KN^{-1/M_s(M_t-M_s)},\end{aligned} \tag{27}$$

where the approximation in (27) means that the left-hand side and the right-hand side can be contained in a ball of radius $\epsilon > 0$.

4.3. Special case of MRT-MRC

In the special case of single-stream beamforming with $M_s = 1$, $\mathbf{F}_\infty$ reduces to maximum-ratio transmission (MRT). Considering a maximum-ratio combining (MRC) receiver, the loss in expected SNR of the received symbol due to quantization of the beamformer $\mathbf{F}_\infty$ may be expressed as $\Delta\mathrm{SNR} = E\{\|\mathbf{H}\mathbf{F}_\infty\|_F^2\} - \sup_{\mathcal{F}:\,\mathrm{card}(\mathcal{F})=N} E\{\|\mathbf{H}\overline{\mathcal{Q}}(\mathbf{F}_\infty)\|_F^2\}$. The approximation (22) simplifies to the form

$$\Delta\mathrm{SNR} \approx (E\{\lambda_1\} - M_r)N^{-1/(M_t-1)}. \tag{28}$$

This particular result has also been derived earlier by Mondal and Heath [36].

4.4. Experimental results

The utility of the approximation (22) is validated by simulations. A 4×4 QPOSTBC MIMO system is considered and precoding with $M_s = 1, 2$ is simulated. In both cases, the codebooks are designed using the FFT-based search algorithm proposed in [51]. The precoder selection criterion is given by (24) and $E\{\|\mathbf{H}\overline{\mathcal{Q}}(\mathbf{F}_\infty)\|_F^2\}$ is plotted in dB as a function of $\log_2 N$ in Figure 4. The experimental results show that the approximation in (27) is reasonably accurate even at small values of N and provides a practical characterization of performance.

4.5. Observations

To better understand the result in (27), we provide an analogous result from vector quantization theory [52, 53]. Consider a D-dimensional (complex dimension) random vector

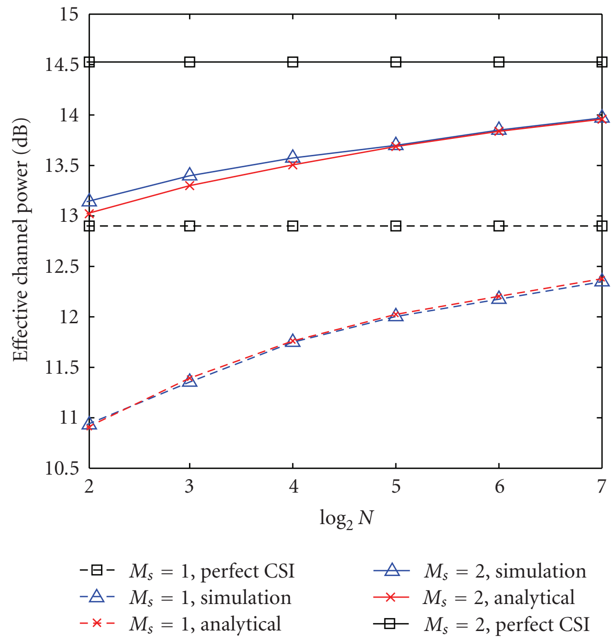

FIGURE 4: The expected SNR, $10\log_{10}(E\{\|\mathbf{H}\overline{\mathcal{Q}}(\mathbf{F}_\infty)\|_F^2\})$, is plotted against the number of bits used for quantization. The simulation results are compared against the closed-form approximation in (27). The system parameters are $M_t = 4$, $M_r = 4$, and the perfect CSI case meaning $E\{\|\mathbf{H}\mathbf{F}_\infty\|_F^2\}$ is also plotted for comparison.

and let every instance of the vector be quantized independently with B bits. Then the average error due to quantization measured in terms of square-Euclidean distance follows $\sim 2^{-B/D}$. The loss in expected SNR from (27) may be written as $\sim 2^{-B/M_s(M_t-M_s)}$ where $B = \log_2 N$ represents the number of quantization bits used for every instance of the precoding matrix. Comparing the two results, it appears that although the precoding matrices are of complex dimension M_tM_s, the dimension of the space that is getting quantized is much smaller, and of dimension $M_s(M_t - M_s)$. In fact, it can be shown that if the performance metric is unitarily invariant, the precoding matrices are unitary and the elements of $\mathcal{F}$ are also unitary, then the precoding matrices can be mapped to a bounded space of dimension $M_s(M_t - M_s)$, and then equivalently quantized. (The space of dimension $M_s(M_t - M_s)$ is the complex Grassmann manifold and this equivalent formulation of quantization is available in [47, 54].) The reduction in dimension (as well as the bounded nature of the space) implies that we are quantizing a much smaller region (compared to $\mathbb{C}^{M_tM_s}$) which is the precise reason why the loss in performance due to quantization is surprisingly small. This also justifies the quantized precoding matrices being unitary.

The loss in expected SNR reduces exponentially with the number of feedback bits B. Thus, most of the gains in channel power is obtained at low values of feedback rates and increasing feedback further leads to insignificant gains (also evident from Figure 4). It may be noted from (20) that the loss in expected SNR depends on the spread of the expected eigenvalues. The number of receive antennas M_r only affects the factor $(\overline{\Lambda} - \underline{\Lambda})$ in (20). It is observed from experiments that this factor decreases with increasing M_r and, thus, the loss in expected SNR also reduces for a fixed N.

5. CONCLUSIONS

In this paper a precoded spatial multiplexing system using a ZF or MMSE receiver and a precoded space-time block coding system are investigated. The focus was on precoding matrices that are unitary and quantized using a codebook of matrices. The main result states that there is no loss in diversity due to quantization as long as the cardinality of the codebook is above a certain threshold (determined only by the number of transmit antennas and the number of data streams) irrespective of the codebook structure. In precoded OSTBC systems, the loss in SNR due to quantization reduces exponentially with the number of feedback bits. Thus increasing the number of feedback bits beyond a certain threshold produces diminishing returns.

In this analysis, we have assumed perfect channel knowledge at the receiver and considered an uncorrelated Rayleigh fading channel. Performance analysis incorporating channel estimation errors and more general channel models is a possible direction of future research.

APPENDICES

A. PROOF OF THEOREM 1

In this proof we abuse notation and denote the column space of an arbitrary matrix **F** also by **F**. The connotations are obvious from context.

Claim 1. Let $\mathbf{S} \in G_{M_t,M_s}$ be any point and $\mathbf{F}_k$ be any element of $\mathcal{F}$ (both $\mathbf{S}$, $\mathbf{F}_k$ are unitary). Then

$$d_P(\mathbf{S},\mathbf{F}_k) < 1 \Longleftrightarrow \mathbf{S}^{\perp} \in c(\mathbf{F}_k), \tag{A.1}$$

where $\mathbf{S}^{\perp}$ denotes the orthogonal complement of the subspace $\mathbf{S}$ and $c(\mathbf{F}_k)$ denotes the complement of $\mathbf{F}_k$ as defined in Theorem 1,

$$d_P(\mathbf{S},\mathbf{F}_k) < 1 \Longleftrightarrow \|\mathbf{F}_k^H\mathbf{S}^{\perp}\|_2 < 1 \tag{A.2}$$

$$\Longleftrightarrow \max_{1\le i\le \min(M_s,M_t-M_s)} \cos\theta_i < 1 \tag{A.3}$$

$$\Longleftrightarrow \mathbf{F}_k \cap \mathbf{S}^{\perp} = \{0\}, \tag{A.4}$$

where (A.2) follows from the representation $d_P(\mathbf{S},\mathbf{F}_k) = \|\mathbf{F}_k^H\mathbf{S}^{\perp}\|_2$ mentioned in [55], (A.3) follows from the notation that $\cos\theta_i$ are the singular values of $\mathbf{F}_k^H\mathbf{S}^{\perp}$ which also means that θ_i are the critical angles between the subspaces $\mathbf{F}_k$ and $\mathbf{S}^{\perp}$, for a reference see [55], (A.4) follows from [55, Theorem 12.4.2] which states that if all the singular values ($\cos\theta_i$) are less than 1, then the subspaces have zero intersection.

Also, $\mathbf{F}_k + \mathbf{S}^{\perp} = \mathbb{C}^{M_t}$, thus $\mathbb{C}^{M_t} = \mathbf{F}_k \oplus \mathbf{S}^{\perp}$ and the claim follows. From Claim 1, it follows that the following are equivalent.

(i) $d_P(\mathbf{S},\mathbf{F}_k) < 1$ for some $\mathbf{F}_k \in \mathcal{F}$ for all $\mathbf{S} \in G_{M_t,M_s}$.
(ii) $c(\mathbf{F}_1) \cup c(\mathbf{F}_2) \cup \cdots \cup c(\mathbf{F}_N) = G_{M_t,M_t-M_s}$.

Now, define a function over G_{M_t,M_s} by the following:

$$f(\mathbf{F}) = \min_{\mathbf{F}_i\in\mathcal{F}} \|\mathbf{F}\mathbf{F}^H - \mathbf{F}_i\mathbf{F}_i^H\|_2. \tag{A.5}$$

Then $f(\mathbf{F})$ is continuous over G_{M_t,M_s}. This implies

$$\sup_{\mathbf{F}\in G_{M_t,M_s}} f(\mathbf{F}) = \delta < 1 \tag{A.6}$$

since $f(\mathbf{F}) < 1$ for $\mathbf{F} \in G_{M_t,M_s}$ and G_{M_t,M_s} is compact.

B. PROOF OF COROLLARY 1

Recall the definition of $\mathbf{U}$, $\mathbf{\Sigma}$, $\overline{\mathbf{U}}$ based on the SVD of $\mathbf{H}^H\mathbf{H}$ from Section 2.3. Let $\overline{\mathbf{\Sigma}} = \operatorname{diag}(\lambda_{M_t},\ldots,\lambda_{M_s})$, $\underline{\mathbf{\Sigma}} = \operatorname{diag}(\lambda_{M_s+1},\ldots,\lambda_{\min(M_t,M_r)})$ and $\underline{\mathbf{U}}$ be the $M_t \times (\min(M_t,M_r) - M_s)$ submatrix of $\mathbf{U}$ corresponding to $\{\lambda_{M_s+1},\ldots,\lambda_{\min(M_t,M_r)}\}$. Since $\mathbf{H}^H\mathbf{H}$ is of rank equal to $\min(M_t,M_r)$ with probability 1, in the following we consider $\underline{\mathbf{\Sigma}}$ to be full rank. It may be noted, however, that the rank depends on the value of M_t, M_r, and M_s and in case $M_r = M_s$, $\underline{\mathbf{\Sigma}}$ and $\underline{\mathbf{U}}$ are not defined and the following derivation remains valid while ignoring all terms involving $\underline{\mathbf{\Sigma}}$ and $\underline{\mathbf{U}}$.

Claim 2. Consider $\mathbf{F} = \mathbf{F}_{\infty}$. Then the diversity may be written as

$$d = -\lim_{\eta\to\infty} \frac{\log E\{e^{-\eta\lambda_{M_s}}\}}{\log \eta}, \tag{B.1}$$

where η is a constant.

The postprocessing SNR for the kth stream can be expressed as

$$\begin{aligned}\mathrm{SNR}_k^{(\mathrm{ZF})}(\mathbf{F}_{\infty}) &= \frac{E_s}{M_sN_0[\mathbf{F}_{\infty}^H\mathbf{H}^H\mathbf{H}\mathbf{F}_{\infty}]_{kk}^{-1}} \\ &= \frac{E_s}{M_sN_0[\overline{\mathbf{\Sigma}}]_{kk}^{-1}} = \frac{E_s}{M_sN_0}\lambda_k.\end{aligned} \tag{B.2}$$

The expected probability of symbol error can be written as

$$P_e \simeq \sum_{k=1}^{M_s} E\left\{\overline{N_e}Q\left(\sqrt{\frac{E_sd_{\min}^2\lambda_k}{2M_sN_0}}\right)\right\} \tag{B.3}$$

$$\le \sum_{k=1}^{M_s} E\{e^{-(E_sd_{\min}^2/4M_sN_0)\lambda_k}\}, \tag{B.4}$$

where $\overline{N_e}$ is the number of nearest neighbors and $Q(\cdot)$ is the Gaussian Q-function. Thus as $E_s/N_0 \to \infty$, we can write

$$P_e \le E\{e^{-(E_sd_{\min}^2/4M_sN_0)\lambda_{M_s}}\}. \tag{B.5}$$

Note that the upper bound in (B.4) stems from the Chernoff bound due to the inequality $Q(x) \le e^{-x^2/2}$. It is straightforward to show that $Q(x) \ge \eta_1 e^{-\eta_2 x^2}$ for some constants η_1,

η_2 and a lower bound to P_e could be derived using the same arguments as before. Thus the diversity can be expressed as

$$d = -\lim_{\eta_3 \to \infty} \frac{\log E\{e^{-\eta_3 \lambda_{M_s}}\}}{\log \eta_3} \tag{B.6}$$

for some constant η_3. This justifies the claim.

Claim 3 (cf. (6)). If $\mathbf{F} = \mathcal{Q}_P(\mathbf{F}_\infty) = \arg\min_{\mathbf{F}_i \in \mathcal{F}} d_P(\mathbf{F}_\infty, \mathbf{F}_i)$, then

$$\frac{1}{[\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}]^{-1}_{kk}} \geq \frac{1}{\left[\mathbf{F}^H\overline{\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}}^H\mathbf{F}\right]^{-1}_{kk}}. \tag{B.7}$$

Since $\mathcal{F}$ is a covering codebook, according to Theorem 1 we have $d_P(\mathbf{F}_\infty, \mathbf{F}) < 1$. Noting $\mathbf{F}_\infty = \overline{\mathbf{U}}$, it follows that $\mathbf{F}^H\overline{\mathbf{U}}$ is full rank. Also, $\overline{\boldsymbol{\Sigma}}$ and $\underline{\boldsymbol{\Sigma}}$ are full rank by definition. Then we can write

$$\begin{aligned}
[\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}]^{-1} &= [\mathbf{F}^H\overline{\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}}^H\mathbf{F} + \mathbf{F}^H\underline{\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}}^H\mathbf{F}]^{-1} && \text{(B.8)}\\
&= [\mathbf{A} + \mathbf{Y}\underline{\boldsymbol{\Sigma}}\mathbf{Y}^H]^{-1} && \text{(B.9)}\\
&= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{Y}(\underline{\boldsymbol{\Sigma}}^{-1} + \mathbf{Y}^H\mathbf{A}^{-1}\mathbf{Y})^{-1}\mathbf{Y}^H\mathbf{A}^{-1} && \text{(B.10)}\\
&= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{Y}\mathbf{V}\mathbf{S}\mathbf{V}^H\mathbf{Y}^H\mathbf{A}^{-1} && \text{(B.11)}\\
&= \mathbf{A}^{-1} - \mathbf{B}\mathbf{B}^H, && \text{(B.12)}
\end{aligned}$$

where (B.9) is just a change in notation by defining $\mathbf{A} = \mathbf{F}^H\overline{\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}}^H\mathbf{F}$ and $\mathbf{Y} = \mathbf{F}^H\underline{\mathbf{U}}$, (B.10) follows from a standard formula in [56], (B.11) is derived by an SVD decomposition given by $(\underline{\boldsymbol{\Sigma}}^{-1} + \mathbf{Y}^H\mathbf{A}^{-1}\mathbf{Y})^{-1} = \mathbf{V}\mathbf{S}\mathbf{V}^H$, and (B.12) is again a change in notation where $\mathbf{B} = \mathbf{A}^{-1}\mathbf{Y}\mathbf{V}\mathbf{S}^{1/2}$. Since $\mathbf{B}\mathbf{B}^H$ have real-positive diagonal entries, it follows from (B.12) that

$$[\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}]^{-1}_{kk} \leq \left[\mathbf{F}^H\overline{\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}}^H\mathbf{F}\right]^{-1}_{kk} \tag{B.13}$$

which justifies the claim.

Claim 4 (cf. (6)). If $\mathbf{F} = \mathcal{Q}_P(\mathbf{F}_\infty) = \arg\min_{\mathbf{F}_i \in \mathcal{F}} d_P(\mathbf{F}_\infty, \mathbf{F}_i)$, then

$$\frac{1}{\left[\mathbf{F}^H\overline{\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}}^H\mathbf{F}\right]^{-1}_{kk}} \geq \eta\lambda_{M_s}, \tag{B.14}$$

where η is a positive constant.

In the following $\mathbf{e}_k$ denotes a vector of unit magnitude where the kth element is unity:

$$\begin{aligned}
\mathbf{e}_k^H\left[\mathbf{F}^H\overline{\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}}^H\mathbf{F}\right]^{-1}\mathbf{e}_k &\leq \|\mathbf{e}_k^H(\mathbf{F}^H\overline{\mathbf{U}})^{-1}\|_F^2\lambda_{M_s}^{-1} && \text{(B.15)}\\
&\leq \|(\mathbf{F}^H\overline{\mathbf{U}})^{-1}\|_F^2\lambda_{M_s}^{-1} && \text{(B.16)}\\
&= \|\mathbf{W}\mathbf{S}\mathbf{V}^H\|_F^2\lambda_{M_s}^{-1} && \text{(B.17)}\\
&\leq \max_{1\leq i\leq M_s}\left(\frac{M_s}{\cos^2\theta_i}\right)\lambda_{M_s}^{-1} && \text{(B.18)}\\
&= \left(\frac{M_s}{1-\delta^2}\right)\lambda_{M_s}^{-1}. && \text{(B.19)}
\end{aligned}$$

In the above (B.15) holds due to the fact that $(\mathbf{x}^H\overline{\boldsymbol{\Sigma}}^{-1}\mathbf{x}/\|\mathbf{x}\|^2) \leq \lambda_{M_s}^{-1}$, (B.16) holds because $\|\mathbf{A}\mathbf{B}\|_F^2 \leq \|\mathbf{A}\|_F^2\|\mathbf{B}\|_F^2$, (B.17) follows from the SVD decomposition of $(\mathbf{F}^H\overline{\mathbf{U}})^{-1} = \mathbf{W}\mathbf{S}\mathbf{V}^H$, (B.18) holds due to the fact that $\mathbf{S} = \operatorname{diag}(1/\cos\theta_1, \ldots, 1/\cos\theta_{M_s})$, where θ_i are the critical angles between the column spaces of $\mathbf{F}$ and $\overline{\mathbf{U}}$, and finally (B.19) holds because the covering radius of the codebook is upper bounded by $\delta < 1$ from Theorem 1. Thus the claim is justified.

Let us define the selected precoder $\mathbf{E}$ as [cf. (14)]

$$\mathbf{E} = \mathcal{Q}^*(\mathbf{F}_\infty) = \arg\max_{\mathbf{F}\in\mathcal{F}}\min_k \text{SNR}_k^{(\text{ZF})}(\mathbf{F}). \tag{B.20}$$

Then we have the following:

$$\begin{aligned}
\text{SNR}_k^{(\text{ZF})}(\mathbf{E}) &= \frac{E_s}{M_sN_0[\mathbf{E}^H\mathbf{H}^H\mathbf{H}\mathbf{E}]^{-1}_{kk}} && \text{(B.21)}\\
&\geq \min_k \frac{E_s}{M_sN_0[\mathbf{E}^H\mathbf{H}^H\mathbf{H}\mathbf{E}]^{-1}_{kk}} && \text{(B.22)}\\
&\geq \min_k \frac{E_s}{M_sN_0[\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}]^{-1}_{kk}} && \text{(B.23)}\\
&\geq \min_k \frac{E_s}{M_sN_0\left[\mathbf{F}^H\overline{\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}}^H\mathbf{F}\right]^{-1}_{kk}} && \text{(B.24)}\\
&\geq \zeta\lambda_{M_s}, && \text{(B.25)}
\end{aligned}$$

where ζ is a constant. In the above (B.23) holds because $\mathbf{F}$ is chosen according to the criterion $\mathbf{F} = \arg\min_{\mathbf{F}_i\in\mathcal{F}} d_P(\mathbf{F}_\infty, \mathbf{F}_i)$ and is a suboptimal precoder [cf. (6)], (B.24) follows from Claim 3, and (B.25) holds due to Claim 4.

From (B.25) and Claim 2 it follows that the diversity is preserved when $\mathcal{F}$ is a covering codebook and the precoder selection criterion is (B.20).

C. PROOF OF LEMMA 1

An alternate representation for $d_P(\mathbf{F}_1, \mathbf{F}_2)$ for arbitrary $\mathbf{F}_1, \mathbf{F}_2 \in G_{M_t,M_s}$ is given by [44]

$$d_P(\mathbf{F}_1, \mathbf{F}_2) = \max_{1\leq i\leq M_s} \sin\theta_i, \tag{C.1}$$

where θ_i are the critical angles between the column spaces of $\mathbf{F}_1, \mathbf{F}_2$. Consider an arbitrary precoder $\mathbf{F} \in G_{M_t,M_s}$ and the antenna selection codebook $\mathcal{F} = \{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_N\}$, where $N = \binom{M_t}{M_s}$. Since $\operatorname{rank}(\mathbf{F}) = M_s$, $\exists$ a set of M_s linearly independent rows in $\mathbf{F}$. Suppose $\{i_1, i_2, \ldots, i_{M_s}\}$, $1 \leq i_k \leq M_t$ denote the set of rows. Also let $\mathbf{F}_* \in \mathcal{F}$ be the precoder that selects the antenna set $\{i_1, i_2, \ldots, i_{M_s}\}$. Then $\operatorname{rank}(\mathbf{F}_*^H\mathbf{F}) = M_s$. Thus $\max_{1\leq i\leq M_s}\theta_i < \pi/2$, where θ_i are the critical angles between the column spaces of $\mathbf{F}_*$ and $\mathbf{F}$. Then $\max_{1\leq i\leq M_s}\sin\theta_i < 1$. Then from (C.1) it follows that $d_P(\mathbf{F}_*, \mathbf{F}) < 1$.

D. PROOF OF COROLLARY 2

Let us define $\mathbf{U}$, $\mathbf{\Sigma}$, $\overline{\mathbf{U}}$, $\underline{\mathbf{U}}$, $\overline{\mathbf{\Sigma}}$, $\underline{\mathbf{\Sigma}}$ as in Appendix B. Also assume $\mathbf{E}_{ij}$ is the codeword difference matrix between the codewords i and j, is of size $M_t \times T$, and subsumes $i \neq j$. Also assume $\mathbf{E}_{ij}\mathbf{E}_{ij}^H$ is full-rank for all i, j. Recall that the chosen precoder $\mathbf{E}$ is expressed as [cf. (16)]

$$\mathbf{E} = \mathcal{Q}^*(\mathbf{F}_\infty) = \arg\max_{\mathbf{F}\in\mathcal{F}} \min_{ij} \left\|\mathbf{HFE}_{ij}\right\|_F^2, \tag{D.1}$$

and for purposes of analysis we also introduce a suboptimal precoder $\mathbf{F}$ given by [cf. (6)]

$$\mathbf{F} = \mathcal{Q}_P(\mathbf{F}_\infty) = \arg\min_{\mathbf{F}\in\mathcal{F}} d_P(\mathbf{F}_\infty, \mathbf{F}). \tag{D.2}$$

Note that in an unprecoded OSTBC system with M_t transmit and M_r receive antennas, the Chernoff bound on the pairwise error probability is of the form $e^{-\eta_1\|\mathbf{H}\|_F^2}$, where η_1 is a positive constant and this system can achieve a diversity order of M_tM_r. Then it follows from (15) that if $\|\mathbf{HEE}_{ij}\|_F^2 > \eta_2\|\mathbf{H}\|_F^2$ for some positive constant η_2, then the precoded system can also achieve a diversity order of M_tM_r with (D.1) as the precoder selection criterion.

Claim 5. If $\mathbf{F} = \arg\min_{\mathbf{F}\in\mathcal{F}} d_P(\mathbf{F}_\infty, \mathbf{F})$, then

$$\left\|\mathbf{HFE}_{ij}\right\|_F^2 \geq \eta\|\mathbf{H}\|_F^2 \tag{D.3}$$

for some positive constant η.

The left-hand side of (D.3) may be expressed as

$$\begin{aligned}
\left\|\mathbf{HFE}_{ij}\right\|_F^2 &= \operatorname{tr}(\mathbf{E}_{ij}^H\mathbf{F}^H\mathbf{U\Sigma U}^H\mathbf{FE}_{ij}) && \text{(D.4)}\\
&= \operatorname{tr}(\mathbf{E}_{ij}^H\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{\Sigma}}\overline{\mathbf{U}}^H\mathbf{FE}_{ij}) + \operatorname{tr}(\mathbf{E}_{ij}^H\mathbf{F}^H\underline{\mathbf{U}}\underline{\mathbf{\Sigma}}\underline{\mathbf{U}}^H\mathbf{FE}_{ij}) && \text{(D.5)}\\
&\geq \operatorname{tr}(\overline{\mathbf{\Sigma}}\overline{\mathbf{U}}^H\mathbf{FE}_{ij}\mathbf{E}_{ij}^H\mathbf{F}^H\overline{\mathbf{U}}) && \text{(D.6)}\\
&= \operatorname{tr}(\overline{\mathbf{\Sigma}}\mathbf{VSV}^H) && \text{(D.7)}\\
&= \operatorname{tr}(\mathbf{SV}^H\overline{\mathbf{\Sigma}}\mathbf{V}) && \text{(D.8)}\\
&\geq \operatorname{tr}(\overline{\mathbf{\Sigma}})\lambda_{\min}(\mathbf{S}) && \text{(D.9)}\\
&= \operatorname{tr}(\overline{\mathbf{\Sigma}})\lambda_{\min}(\overline{\mathbf{U}}^H\mathbf{FE}_{ij}\mathbf{E}_{ij}^H\mathbf{F}^H\overline{\mathbf{U}}) && \text{(D.10)}\\
&= \operatorname{tr}(\overline{\mathbf{\Sigma}})\lambda_{\min}(\mathbf{E}_{ij}\mathbf{E}_{ij}^H\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{U}}^H\mathbf{F}) && \text{(D.11)}\\
&\geq \operatorname{tr}(\overline{\mathbf{\Sigma}})\lambda_{\min}(\mathbf{E}_{ij}\mathbf{E}_{ij}^H)\lambda_{\min}(\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{U}}^H\mathbf{F}) && \text{(D.12)}\\
&\geq \operatorname{tr}(\overline{\mathbf{\Sigma}})\lambda_{\min}(\mathbf{E}_{ij}\mathbf{E}_{ij}^H)(1-\delta^2) && \text{(D.13)}\\
&\geq \eta\|\mathbf{H}\|_F^2. && \text{(D.14)}
\end{aligned}$$

In the above (D.4) follows from the definition of Frobenius norm, (D.5) holds because $\mathbf{U\Sigma U}^H = \overline{\mathbf{U}}\overline{\mathbf{\Sigma}}\overline{\mathbf{U}}^H + \underline{\mathbf{U}}\underline{\mathbf{\Sigma}}\underline{\mathbf{U}}^H$, (D.6) follows from the fact that $\mathbf{E}_{ij}^H\mathbf{F}^H\underline{\mathbf{U}}\underline{\mathbf{\Sigma}}\underline{\mathbf{U}}^H\mathbf{FE}_{ij}$ is hermitian nonnegative definite so its trace is nonnegative, (D.7) can be explained by the SVD decomposition given by $\overline{\mathbf{U}}^H\mathbf{FE}_{ij}\mathbf{E}_{ij}^H\mathbf{F}^H\overline{\mathbf{U}} = \mathbf{VSV}^H$, (D.8) follows from the repeated use of the property $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$, in (D.9) we use the fact $\operatorname{tr}(\mathbf{V}^H\overline{\mathbf{\Sigma}}\mathbf{V}) = \operatorname{tr}(\overline{\mathbf{\Sigma}})$, (D.10) holds because $\mathbf{S}$ is a diagonal matrix consisting of the eigenvalues of $\overline{\mathbf{U}}^H\mathbf{FE}_{ij}\mathbf{E}_{ij}^H\mathbf{F}^H\overline{\mathbf{U}}$, (D.11) follows from the fact that the eigenvalues of $\mathbf{AB}$ and the eigenvalues of $\mathbf{BA}$ are equal when $\mathbf{A}$, $\mathbf{B}$ are square [57], (D.12) follows from the fact that $\lambda_{\min}(\mathbf{AB}) \geq \lambda_{\min}(\mathbf{A})\lambda_{\min}(\mathbf{B})$ for positive definite Hermitian matrices $\mathbf{A}$, $\mathbf{B}$ also mentioned in [57], (D.13) holds because $\lambda_{\min}(\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{U}}^H\mathbf{F}) = \min_i \cos^2\theta_i$, where θ_i are the critical angles between the column spaces of $\mathbf{F}$ and $\overline{\mathbf{U}}$ along with the fact that the covering radius of $\mathcal{F}$ is δ and by Theorem 1 we have $\delta < 1$, and finally (D.14) is because $\min_{i,j,i\neq j}\lambda_{\min}(\mathbf{E}_{ij}\mathbf{E}_{ij}^H)$ is a positive constant by definition of a full-rank STBC and $\operatorname{tr}(\overline{\mathbf{\Sigma}}) \geq \lambda_1 \geq (1/M_t)\|\mathbf{H}\|_F^2$. This justifies the claim.

Then we have

$$\begin{aligned}
\left\|\mathbf{HEE}_{ij}\right\|_F^2 &\geq \min_{i,j,i\neq j}\left\|\mathbf{HEE}_{ij}\right\|_F^2 && \text{(D.15)}\\
&\geq \min_{i,j,i\neq j}\left\|\mathbf{HFE}_{ij}\right\|_F^2 && \text{(D.16)}\\
&\geq \eta\|\mathbf{H}\|_F^2, && \text{(D.17)}
\end{aligned}$$

where (D.16) is because of the suboptimality of the precoder $\mathbf{F}$ compared to $\mathbf{E}$ and (D.17) is due to Claim 5.

E. DIVERSITY ORDER FOR QPSM WITH $M_R = M_S$

Lemma 2. *A precoded spatial multiplexing system with a ZF receiver, $\mathbf{F}_\infty$ is the precoder and $M_r = M_s$ can attain a diversity order of $M_t - M_s + 1$.*

Recall that since $M_s = M_r$, $\mathbf{H}^H\mathbf{H}$ will have M_s nonzero eigenvalues with probability 1. Consider $\{\lambda_1, \lambda_2, \dots, \lambda_{M_s}\}$ as the set of ordered eigenvalues of $\mathbf{H}^H\mathbf{H}$, where $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_{M_s} \geq 0$. Now let $\{\lambda_1^{(u)}, \lambda_2^{(u)}, \dots, \lambda_{M_s}^{(u)}\}$ denote the set of unordered eigenvalues. Then the joint probability density functions of $\{\lambda_1, \lambda_2, \dots, \lambda_{M_s}\}$ and $\{\lambda_1^{(u)}, \lambda_2^{(u)}, \dots, \lambda_{M_s}^{(u)}\}$ differ only by a scaling factor of $M_s!$. The marginal probability densities of $\lambda_i^{(u)}$, $1 \leq i \leq M_s$ are identical and may be expressed as [58]

$$f(\lambda) = \frac{1}{M_s}\sum_{k=0}^{M_s-1}\frac{k!}{(k+M_t-M_s)!}\left[L_k^{M_t-M_s}(\lambda)\right]^2\lambda^{M_t-M_s}e^{-\lambda}, \tag{E.1}$$

where $L_k^{M_t-M_s}(\lambda) = (1/k!)e^{\lambda}\lambda^{M_t-M_s}(d^k/d\lambda^k)(e^{-\lambda}\lambda^{M_t-M_s+k})$ is an associated Laguerre polynomial of order k [59]. Define $\mathbf{\Sigma} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_{M_s})$ and $\mathbf{\Sigma}^{(u)} = \operatorname{diag}(\lambda_1^{(u)}, \lambda_2^{(u)}, \dots, \lambda_{M_s}^{(u)})$. Then the postprocessing SNR for the ZF receiver can be written as (see (B.2))

$$\mathrm{SNR}_k^{(\mathrm{ZF})} = \frac{E_s}{M_sN_0[\mathbf{\Sigma}]_{kk}^{-1}} \overset{d}{=} \frac{\eta}{[\mathbf{\Sigma}^{(u)}]_{kk}^{-1}} = \eta\lambda_k^{(u)}, \tag{E.2}$$

where the constant η absorbs the factors $E_s/N_0, M_s, M_s!$. Then the probability of symbol error may be expressed as

$$P_e \simeq \overline{N_e} E\left\{Q\left(\sqrt{\eta\lambda\frac{d_{\min}^2}{2}}\right)\right\} \tag{E.3}$$

$$\leq \overline{N_e} E\{e^{-\eta\lambda(d_{\min}^2/4)}\} \tag{E.4}$$

$$= \frac{\overline{N_e}}{M_s}\int_0^\infty e^{-\eta\lambda(d_{\min}^2/4)} \sum_{k=0}^{M_s-1} a_k\left[L_k^{M_t-M_s}(\lambda)\right]^2 \lambda^{M_t-M_s} e^{-\lambda} d\lambda \tag{E.5}$$

$$= \frac{\overline{N_e}}{M_s}\int_0^\infty e^{-\lambda[\eta(d_{\min}^2/4)+1]} \sum_{k=0}^{M_s-1} a_k \times\left[\sum_{p=0}^{k} b_p L_{2p}^{2(M_t-M_s)}(2\lambda)\right]\lambda^{M_t-M_s} d\lambda \tag{E.6}$$

$$= \frac{\overline{N_e}}{M_s}\int_0^\infty e^{-\lambda[\eta(d_{\min}^2/4)+1]} \sum_{k=0}^{M_s-1} a_k \times\left\{\sum_{p=0}^{k} b_p\left[\sum_{q=0}^{2p}(-1)^q c_q\lambda^q\right]\right\}\lambda^{M_t-M_s} d\lambda \tag{E.7}$$

$$= \frac{\overline{N_e}}{M_s}\sum_{k=0}^{M_s-1} a_k \times\left\{\sum_{p=0}^{k} b_p\left[\sum_{q=0}^{2p}(-1)^q c_q\int_0^\infty e^{-\lambda[\eta(d_{\min}^2/4)+1]}\lambda^{M_t-M_s+q} d\lambda\right]\right\} \tag{E.8}$$

$$= \frac{\overline{N_e}}{M_s}\sum_{k=0}^{M_s-1} a_k \times\left\{\sum_{p=0}^{k} b_p\left[\sum_{q=0}^{2p}(-1)^q c'_q\left(\eta\frac{d_{\min}^2}{4}+1\right)^{-(M_t-M_s+q+1)}\right]\right\} \tag{E.9}$$

$$\overset{\eta\to\infty}{\simeq} \zeta\left(\eta\frac{d_{\min}^2}{4}\right)^{-(M_t-M_s+1)}, \tag{E.10}$$

where ζ is a positive constant. In the above, in (E.3), $\overline{N_e}$ is the number of nearest neighbors and $Q(\cdot)$ is the Gaussian Q-function, λ denotes an unordered eigenvalue, $d_{\min}^2$ is the minimum distance of the constellation, (E.4) represents the Chernoff bound for $Q(\cdot)$, (E.5) follows from the form of the pdf of λ in (E.1) and a_k is a positive constant, (E.6) follows from an expansion of $[L_k^{M_t-M_s}(\lambda)]^2$ given in [59] and $b_p > 0$ are constants, (E.7) follows from an expansion of $L_{2p}^{2(M_t-M_s)}(2\lambda)$ given in [59] and $c_q > 0$ are constants, (E.8) is just a rearrangement of terms, (E.9) results from the relation $\int_0^\infty x^n e^{-\mu x}dx = n!\mu^{-n-1}$ for $\mu > 0$ and absorbing the constant $n!$ into c'_q, and (E.9) results from the observation that as $\eta \to \infty$ almost the entire contribution to sum is by the term with $q = 0$ and absorbing all positive constants into ζ. From (E.9) we conclude that the diversity order is $M_t - M_s + 1$.

F. PROOF OF THEOREM 2

Define [cf. (8), (19)]

$$\mathbf{F} = \mathcal{Q}_C(\mathbf{F}_\infty) = \arg\min_{\mathbf{F}_k\in\mathcal{F}} d_C(\mathbf{F}_\infty, \mathbf{F}_k). \tag{F.1}$$

Recall from Section 2.3 that the SVD of $\mathbf{H}^H\mathbf{H}$ is given by

$$\mathbf{H}^H\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H, \tag{F.2}$$

and $\overline{\mathbf{U}}$ is the $M_t \times M_s$ submatrix of $\mathbf{U}$ corresponding to the M_s dominant singular values and $\overline{\mathbf{\Sigma}}$ is defined in Appendix B. Note that we redefine $\underline{\mathbf{\Sigma}}$ and $\underline{\mathbf{U}}$ in the following. Let $\underline{\mathbf{\Sigma}} = \mathrm{diag}(\lambda_{M_s+1}, \lambda_2, \ldots, \lambda_{M_t})$ and let $\underline{\mathbf{U}}$ be the $M_t \times (M_t - M_s)$ submatrix of $\mathbf{U}$ with columns corresponding to $\lambda_{M_s+1}, \lambda_{M_s+2}, \ldots, \lambda_{M_t}$. Then we have the following:

$$E\{\|\mathbf{HF}\|_F^2\} = E\{\mathrm{tr}(\mathbf{F}^H\mathbf{U}\mathbf{\Sigma}^H\mathbf{U}^H\mathbf{F})\} \tag{F.3}$$

$$= E\{\mathrm{tr}(\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{\Sigma}}^H\overline{\mathbf{U}}^H\mathbf{F})\} + E\{\mathrm{tr}(\mathbf{F}^H\underline{\mathbf{U}}\underline{\mathbf{\Sigma}}^H\underline{\mathbf{U}}^H\mathbf{F})\}. \tag{F.4}$$

Claim 6.

$$E\{\mathrm{tr}(\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{\Sigma}}\overline{\mathbf{U}}^H\mathbf{F})\} = \overline{\Lambda}E\{\mathrm{tr}(\overline{\mathbf{U}}^H\mathbf{F}\mathbf{F}^H\overline{\mathbf{U}})\}, \tag{F.5}$$

where $\overline{\Lambda} = (1/M_s)\sum_{k=1}^{M_s}E\{\lambda_k\}$.

Now, $\overline{\mathbf{U}} \overset{d}{=} \overline{\mathbf{U}}\mathbf{\Pi}$, where $\mathbf{\Pi}$ is a $M_s\times M_s$ permutation matrix. This is because $\overline{\mathbf{U}}$ is isotropically distributed and $\mathbf{\Pi}$ is unitary. Also, $\mathcal{Q}(\overline{\mathbf{U}}) = \mathcal{Q}(\overline{\mathbf{U}}\mathbf{\Pi})$ since $d(\overline{\mathbf{U}}, \mathbf{F}_k) = d(\overline{\mathbf{U}}\mathbf{\Pi}, \mathbf{F}_k)$. Then we have

$$E\{\mathrm{tr}(\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{\Sigma}}\overline{\mathbf{U}}^H\mathbf{F})\} = E\{\mathrm{tr}(\mathbf{F}^H\overline{\mathbf{U}}\mathbf{\Pi}\overline{\mathbf{\Sigma}}\mathbf{\Pi}^H\overline{\mathbf{U}}^H\mathbf{F})\} \tag{F.6}$$

$$= E\{\mathrm{tr}(\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{\Sigma}}^\Pi\overline{\mathbf{U}}^H\mathbf{F})\}, \tag{F.7}$$

where we define $\overline{\mathbf{\Sigma}}^\Pi = \mathbf{\Pi}\overline{\mathbf{\Sigma}}\mathbf{\Pi}^H$. Now,

$$E\{\mathrm{tr}(\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{\Sigma}}\overline{\mathbf{U}}^H\mathbf{F})\} = \frac{1}{M_s!}\sum_{\mathbf{\Pi}}E\{\mathrm{tr}(\mathbf{F}^H\overline{\mathbf{U}}\overline{\mathbf{\Sigma}}^\Pi\overline{\mathbf{U}}^H\mathbf{F})\} \tag{F.8}$$

$$= E\left\{\mathrm{tr}\left(\mathbf{F}^H\overline{\mathbf{U}}\left\{\frac{1}{M_s!}\sum_{\mathbf{\Pi}}\overline{\mathbf{\Sigma}}^\Pi\right\}\overline{\mathbf{U}}^H\mathbf{F}\right)\right\} \tag{F.9}$$

$$= \mathrm{tr}\left(E\{\overline{\mathbf{U}}^H\mathbf{F}\mathbf{F}^H\overline{\mathbf{U}}\}E\left\{\frac{1}{M_s!}\sum_{\mathbf{\Pi}}\overline{\Sigma}^\Pi\right\}\right) \tag{F.10}$$

$$= \overline{\Lambda}E\{\mathrm{tr}(\overline{\mathbf{U}}^H\mathbf{F}\mathbf{F}^H\overline{\mathbf{U}})\}, \tag{F.11}$$

where $\sum_{\mathbf{\Pi}}$ denotes a summation over all permutation matrices of the same dimension as $\mathbf{\Pi}$. Then (F.8) follows from (F.7) and because there are $M_s!$ permutation matrices of size $M_s \times M_s$, (F.9) holds because expectation and trace are linear operators, (F.10) follows because expectation and trace can be interchanged and $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, and (F.11) holds because $E\{(1/M_s!)\sum_{\mathbf{\Pi}}\overline{\mathbf{\Sigma}}^\Pi\} = \overline{\Lambda}\mathbf{I}$. This justifies Claim 6.

Claim 7.

$$E\{\mathrm{tr}(\mathbf{F}^H\underline{\mathbf{U}}\underline{\mathbf{\Sigma}}\underline{\mathbf{U}}^H\mathbf{F})\} = \underline{\Lambda}M_s - \underline{\Lambda}E\{\mathrm{tr}(\overline{\mathbf{U}}^H\mathbf{F}\mathbf{F}^H\overline{\mathbf{U}})\}, \tag{F.12}$$

where $\underline{\Lambda} = (1/(M_t - M_s))\sum_{k=M_s+1}^{M_t}E\{\lambda_k\}$.

Now, $\underline{\mathbf{U}} \overset{d}{=} \underline{\mathbf{U}}\mathbf{\Pi}$, where $\mathbf{\Pi}$ is an $(M_t - M_s) \times (M_t - M_s)$ permutation matrix. Then

$$\begin{aligned} E\{\text{tr}(\mathbf{F}^H \underline{\mathbf{U}}\underline{\mathbf{\Sigma}}\underline{\mathbf{U}}^H \mathbf{F})\} &= E\{\text{tr}(\mathbf{F}^H \underline{\mathbf{U}}\mathbf{\Pi}\underline{\mathbf{\Sigma}}\mathbf{\Pi}^H \underline{\mathbf{U}}^H \mathbf{F})\} \\ &= E\{\text{tr}(\mathbf{F}^H \underline{\mathbf{U}}\underline{\mathbf{\Sigma}}^{\Pi} \underline{\mathbf{U}}^H \mathbf{F})\}, \end{aligned} \tag{F.13}$$

where we define $\mathbf{\Pi}\underline{\mathbf{\Sigma}}\mathbf{\Pi}^H = \underline{\mathbf{\Sigma}}^{\Pi}$. Then following the steps in the earlier derivation we have

$$E\{\text{tr}(\mathbf{F}^H \underline{\mathbf{U}}\underline{\mathbf{\Sigma}}\underline{\mathbf{U}}^H \mathbf{F})\} = \underline{\Lambda} E\{\text{tr}(\underline{\mathbf{U}}^H \mathbf{F}\mathbf{F}^H \underline{\mathbf{U}})\}. \tag{F.14}$$

Then we can write

$$E\{\text{tr}(\mathbf{F}^H \underline{\mathbf{U}}\underline{\mathbf{\Sigma}}\underline{\mathbf{U}}^H \mathbf{F})\} = \underline{\Lambda} E\{\text{tr}(\mathbf{F}^H \underline{\mathbf{U}}^H \underline{\mathbf{U}}\mathbf{F})\} \tag{F.15}$$

$$= \underline{\Lambda} E\{\text{tr}(\mathbf{F}^H (\mathbf{I} - \overline{\mathbf{U}}\overline{\mathbf{U}}^H)\mathbf{F})\} \tag{F.16}$$

$$= \underline{\Lambda} M_s - \Lambda E\{\text{tr}(\overline{\mathbf{U}}^H \mathbf{F}\mathbf{F}^H \overline{\mathbf{U}})\}, \tag{F.17}$$

where (F.15) holds because $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$, (F.16) follows from $\overline{\mathbf{U}}\overline{\mathbf{U}}^H + \underline{\mathbf{U}}\underline{\mathbf{U}}^H = \mathbf{I}$, and (F.17) holds also due to $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$. This justifies Claim 7.

Note that $\overline{\mathbf{U}} = \mathbf{F}_\infty$ and $E\{\text{tr}(\overline{\mathbf{U}}^H \mathbf{F}\mathbf{F}^H \overline{\mathbf{U}})\} = M_s - E\{d_C^2(\mathbf{F}_\infty, \mathbf{F})\}$ by definition (applying the relation $\|\mathbf{A}\|_F^2 = \text{tr}(\mathbf{A}\mathbf{A}^H)$). Then from (F.4), Claims 6, 7 we have

$$\begin{aligned} E\{\|\mathbf{HF}\|_F^2\} &= \overline{\Lambda} M_s + (\overline{\Lambda} - \underline{\Lambda}) E\{d^2(\mathbf{F}_\infty, \mathbf{F})\} \\ &= E\{\|\mathbf{HF}_\infty\|_F^2\} + (\overline{\Lambda} - \underline{\Lambda}) E\{d^2(\mathbf{F}_\infty, \mathbf{F})\}. \end{aligned} \tag{F.18}$$

ACKNOWLEDGMENTS

Bishwarup Mondal is the recipient of a Motorola Partnerships in Research Grant. This material is based in part upon work supported by the National Science Foundation under Grant CCF-514194. This work has appeared in part in the Proceedings of IEEE International Workshop on Signal Processing Advances, in Wireless Communications pages 1–5, July 2–5, Cannes, 2006.

REFERENCES

[1] H. Sampath and A. J. Paulraj, "Joint transmit and receive optimization for high data rate wireless communication using multiple antennas," in *Proceedings of the 33rd Asilomar Conference on Signals, Systems and Computers (ACSSC '99)*, vol. 1, pp. 215–219, Pacific Grove, Calif, USA, October 1999.

[2] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1051–1064, 2002.

[3] H. Kramer and M. Mathews, "A linear coding for transmitting a set of correlated signals," *IEEE Transactions on Information Theory*, vol. 2, no. 3, pp. 41–46, 1956.

[4] K.-H. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Transactions on Communications*, vol. 24, no. 12, pp. 1283–1290, 1976.

[5] J. Salz, "Digital transmission over cross-coupled linear channels," *AT & T Technical Journal*, vol. 64, no. 6, pp. 1147–1159, 1985.

[6] M. L. Honig, P. Crespo, and K. Steiglitz, "Suppression of near- and far-end crosstalk by linear pre- and post-filtering," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 614–629, 1992.

[7] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Transactions on Communications*, vol. 42, no. 12, pp. 3221–3231, 1994.

[8] W. M. Jang, B. R. Vojcic, and R. L. Pickholtz, "Joint transmitter-receiver optimization in synchronous multiuser communications over multipath channels," *IEEE Transactions on Communications*, vol. 46, no. 2, pp. 269–278, 1998.

[9] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, 1998.

[10] G. Jongren and M. Skoglund, "Utilizing quantized feedback information in orthogonal space-time block coding," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '00)*, vol. 2, pp. 995–999, San Francisco, Calif, USA, November-December 2000.

[11] E. G. Larsson, G. Ganesan, P. Stoica, and W.-H. Wong, "On the performance of orthogonal space-time block coding with quantized feedback," *IEEE Communications Letters*, vol. 6, no. 11, pp. 487–489, 2002.

[12] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2562–2579, 2003.

[13] B. C. Banister and J. R. Zeidler, "A simple gradient sign algorithm for transmit antenna weight adaptation with feedback," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1156–1171, 2003.

[14] D. J. Love, R. W. Heath Jr., W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54–59, 2004.

[15] J. C. Roh and B. D. Rao, "An efficient feedback method for MIMO systems with slowly time-varying channels," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC '04)*, vol. 2, pp. 760–764, Atlanta, Ga, USA, March 2004.

[16] P. Xia and G. B. Giannakis, "Design and analysis of transmit-beamforming based on limited-rate feedback," in *Proceedings of the 60th IEEE Vehicular Technology Conference*, vol. 60, pp. 1653–1657, Los Angeles, Calif, USA, September 2004.

[17] S. Zhou, Z. Wang, and G. B. Giannakis, "Quantifying the power loss when transmit beamforming relies on finite-rate feedback," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1948–1957, 2005.

[18] T. L. Marzetta and B. M. Hochwald, "Fast transfer of channel state information in wireless systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1268–1278, 2006.

[19] T. A. Thomas, K. L. Baum, and P. Sartori, "Obtaining channel knowledge for closed-loop multi-stream broadband MIMO-OFDM communications using direct channel feedback," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '05)*, vol. 6, pp. 3907–3911, St. Louis, Mo, USA, November-December 2005.

[20] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, 2001.

[21] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1165–1175, 2004.

[22] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2599–2613, 2002.

[23] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1673–1690, 2003.

[24] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio combining in rayleigh fading," *IEEE Transactions on Communications*, vol. 47, no. 12, pp. 1773–1776, 1999.

[25] D. A. Gore, R. U. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '00)*, vol. 5, pp. 2785–2788, Istanbul, Turkey, June 2000.

[26] A. F. Molisch, M. Z. Win, and J. H. Winters, "Capacity of MIMO systems with antenna selection," in *Proceedings of IEEE International Conference on Communications (ICC '01)*, vol. 2, pp. 570–574, Helsinki, Finland, June 2001.

[27] D. A. Gore and A. J. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2580–2588, 2002.

[28] Z. Hong, K. Liu, R. W. Heath Jr., and A. M. Sayeed, "Spatial multiplexing in correlated fading via the virtual channel representation," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 856–866, 2003.

[29] A. Gorokhov, D. A. Gore, and A. J. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: theory and algorithms," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2796–2807, 2003.

[30] R. W. Heath Jr., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Communications Letters*, vol. 5, no. 4, pp. 142–144, 2001.

[31] H. Zhang, H. Dai, Q. Zhou, and B. L. Hughes, "On the diversity order of transmit antenna selection for spatial multiplexing systems," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '05)*, vol. 3, pp. 1451–1455, St. Louis, Mo, USA, November-December 2005.

[32] D. J. Love and R. W. Heath Jr., "Necessary and sufficient conditions for full diversity order in correlated Rayleigh fading beamforming and combining systems," *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 20–23, 2005.

[33] D. J. Love and R. W. Heath Jr., "Limited feedback unitary precoding for orthogonal space-time block codes," *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp. 64–73, 2005.

[34] D. J. Love and R. W. Heath Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2967–2976, 2005.

[35] J. C. Roh and B. D. Rao, "Performance analysis of multiple antenna systems with VQ-based feedback," in *Proceedings of the 38th Asilomar Conference on Signals, Systems and Computers (ACSSC '04)*, vol. 2, pp. 1978–1982, Pacific Grove, Calif, USA, November 2004.

[36] B. Mondal and R. W. Heath Jr., "Performance analysis of quantized beamforming MIMO systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4753–4766, 2006.

[37] W. E. Clark and B. Shekhtman, "Covering by complements of subspaces, II," *Proceedings of the American Mathematical Society*, vol. 125, no. 1, pp. 251–254, 1997.

[38] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[39] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[40] B. A. Sethuraman, B. Sundar Rajan, and V. Shashidhar, "Full-diversity, high-rate space-time block codes from division algebras," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2596–2616, 2003.

[41] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2335–2367, 2001.

[42] W. Su and X.-G. Xia, "Signal constellations for quasi-orthogonal space-time block codes with full diversity," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2331–2347, 2004.

[43] S. Das, N. Al-Dhahir, and R. Calderbank, "Novel full-diversity high-rate STBC for 2 and 4 transmit antennas," *IEEE Communications Letters*, vol. 10, no. 3, pp. 171–173, 2006.

[44] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, no. 2, pp. 303–353, 1998.

[45] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, New York, NY, USA, 2003.

[46] B. Mondal, R. W. Heath Jr., and L. W. Hanlen, "Quantization on the grassmann manifold: Applications to precoded mimo wireless systems," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '05)*, vol. 5, pp. 1025–1028, Philadelphia, Pa, USA, March 2005.

[47] B. Mondal, S. Dutta, and R. W. Heath Jr., "Quantization on the Grassmann manifold," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4208–4216, 2007.

[48] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds of arbitrary dimensions and MIMO communications with feedback," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '05)*, vol. 3, pp. 1456–1460, St. Louis, Mo, USA, November-December 2005.

[49] A. T. James, "Normal multivariate analysis and the orthogonal group," *The Annals of Mathematical Statistics*, vol. 25, no. 1, pp. 40–75, 1954.

[50] S. Graf and H. Luschgy, *Foundations of Quantization for Probability Distributions*, vol. 1730 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 2000.

[51] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 1962–1973, 2000.

[52] A. Gersho, "Asymptotically optimal block quantization," *IEEE Transactions on Information Theory*, vol. 25, no. 4, pp. 373–380, 1979.

[53] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325–2384, 1998.

[54] B. Mondal, S. Dutta, and R. W. Heath Jr., "Quantization on the complex projective space," in *Proceedings of the Data Compression Conference (DCC '06)*, pp. 242–251, Snowbird, Utah, USA, March 2006.

[55] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Md, USA, 3rd edition, 1996.

[56] H. V. Henderson and S. R. Searle, "On deriving the inverse of a sum of matrices," *SIAM Review*, vol. 23, no. 1, pp. 53–60, 1981.

[57] L.-Z. Lu and C. E. M. Pearce, "Some new bounds for singular values and eigenvalues of matrix products," *Annals of Operations Research*, vol. 98, no. 1–4, pp. 141–148, 2000.

[58] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[59] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York, NY, USA, 1994.