

Limited Feedback Unitary Precoding for Spatial Multiplexing Systems

David J. Love, *Member, IEEE*, Robert W. Heath, Jr., *Member, IEEE*,

Abstract—Multiple-input multiple-output (MIMO) wireless systems use antenna arrays at both the transmitter and receiver to provide communication links with substantial diversity and capacity. Spatial multiplexing is a common space-time modulation technique for MIMO communication systems where independent information streams are sent over different transmit antennas. Unfortunately, spatial multiplexing is sensitive to ill-conditioning of the channel matrix. Precoding can improve the resilience of spatial multiplexing at the expense of full channel knowledge at the transmitter - which is often not realistic. This correspondence proposes a quantized precoding system where the optimal precoder is chosen from a finite codebook known to both receiver and transmitter. The index of the optimal precoder is conveyed from the receiver to the transmitter over a low-delay feedback link. Criteria are presented for selecting the optimal precoding matrix based on the error rate and mutual information for different receiver designs. Codebook design criteria are proposed for each selection criterion by minimizing a bound on the average distortion assuming a Rayleigh fading matrix channel. The design criteria are shown to be equivalent to packing subspaces in the Grassmann manifold using the projection two-norm and Fubini-Study distances. Simulation results show that the proposed system outperforms antenna subset selection and performs close to optimal unitary precoding with a minimal amount of feedback.

Index Terms—Diversity methods, Grassmannian subspace packing, MIMO systems, Quantized precoding, Rayleigh channels, Spatial multiplexing, V-BLAST.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) wireless channels, created by exploiting antenna arrays at both the transmitter and receiver, promise high capacity and high quality wireless communication links [1], [2]. Exploiting the benefits offered by MIMO channels requires choosing a space-time modulation scheme and receiver algorithm that provide a sensible performance and complexity tradeoff. Along these lines, spatial multiplexing, where a bit stream is demultiplexed into multiple substreams that are sent over different antennas, is a practical space-time modulation technique that permits a choice of optimal, near-optimal, and suboptimal receivers. Unfortunately,

This work was supported in part upon work supported by the Texas Advanced Research (Technology) Program under Grant No. 003658-0614-2001, the National Instruments Foundation, the Samsung Institute of Applied Technology, and the SBC Foundation. This paper was presented in part at the 41st Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, Oct. 1-3, 2003, the IEEE Military Communication Conference, Boston, MA, Oct. 13-16, 2003, and the IEEE Global Telecommunications Conference, San Francisco, CA, Dec. 1-5, 2003.

D. J. Love is with the Center for Wireless Systems and Applications, School of Electrical and Computer Engineering, Purdue University, 465 Northwestern Ave., West Lafayette, IN 47907 USA. (email: djlove@ecn.purdue.edu)
R. W. Heath, Jr. is with the Wireless Networking and Communications Group, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712 USA. (email: rheath@ece.utexas.edu).

spatial multiplexing is sensitive to ill-conditioning of the channel matrix.

Premultiplying the transmitted data streams by a precoding matrix, chosen based on channel information, is one way to guard against rank deficiencies in the channel and to improve error rate performance. The basic idea of precoding is to use some form of channel knowledge at the transmitter to customize the transmitted signal to the eigen-structure of the matrix channel. Precoding has been proposed based on knowledge of the full channel state information at the transmitter [3], [4], first-order statistics [5], [6], or second-order statistics of the channel [7]–[11]. The full gains from precoding are achieved with full channel state information since this allows the transmitted signal to be customized based on the eigen-structure of the matrix channel. In time division duplex systems with suitable ping-pong time, full channel state information may be available. In frequency division duplex systems, however, full channel state information must be conveyed through a feedback channel. This is impractical, though, due to the number of channel coefficients that need to be quantized and transported to the transmitter over limited bandwidth control channels.

In this correspondence, we propose a solution to the problem of precoding for spatial multiplexing systems with limited feedback capacity. The essential idea is that the transmit precoder is chosen from a finite set of precoding matrices, called the codebook, known to both the receiver and the transmitter. The receiver chooses the optimal precoder from the codebook as a function of the current channel state information and sends the binary index of this matrix to the transmitter over a limited feedback channel. We address two key problems in this correspondence: (i) selection of the optimal precoder from the codebook and (ii) design of optimal codebooks. We assume that the channel is statistically described by the narrowband uncorrelated Rayleigh matrix fading model and that there is zero delay in the feedback channel. For codeword selection, we propose selection criteria based on the error probability [3], [12]–[14] and mutual information [15]–[17]. We consider both maximum likelihood (ML) (optimal but high implementation complexity) and linear receivers (suboptimal but lower implementation complexity). The optimal precoding matrix can be easily chosen using our selection criteria by simply searching through all codebook matrices. We address codebook design for each of the proposed selection criteria. By bounding the average distortion, we show that the codeword selection criteria imply that codebooks should be designed such that the constituent matrices are maximally spaced. The distance measure is not Euclidean distance but rather subspace distance on the Grassmann manifold. Specifically we show

that, depending on the receiver, optimal codebook designs are subspace packings in the Grassmann manifold using either the projection two-norm or the Fubini-Study distance.

Limited feedback precoding has been considered in the past extensively for the cases of transmit beamforming [18]–[20], precoding for space-time block codes [21]–[23], and covariance quantization [24], [25]. As well, antenna subset selection can be viewed as an important example of limited feedback precoding where the optimal subset of transmit antennas is computed and conveyed to the transmitter (see for example [12], [15], [17], [26]–[28]). The random vector quantization work in [29] is another form of limited feedback precoding. Our correspondence proposes a natural generalization of this existing limited feedback work.

This correspondence is organized as follows. Section II reviews the precoded spatial multiplexing system model. Criteria for choosing the optimal matrix from the codebook is presented in Section III. Design criteria for creation of the precoder codebook are derived in Section IV. Section V illustrates the performance improvements over no precoding, unquantized precoding, and antenna subset selection using Monte Carlo simulations of the symbol error rate. Conclusions are presented in Section VI.

Notation: We use T to denote transposition, $*$ to denote conjugate transposition, $^{-1}$ to denote matrix inversion, $^{+}$ to denote the matrix pseudo-inverse, \mathbf{I}_M to denote the $M \times M$ identity matrix, $\|\cdot\|_2$ to denote the matrix two-norm, $\|\cdot\|_F$ to denote the matrix Frobenius norm, $tr(\cdot)$ to denote the trace of a matrix, $det(\cdot)$ to denote the determinant of a matrix, \mathbb{C}^m to denote the m -dimensional complex vector space, $\mathcal{CN}(0, \sigma^2)$ to denote the complex normal distribution with independent real and imaginary parts distributed according to $\mathcal{N}(0, \sigma^2/2)$, $\mathcal{U}(M_t, M)$ to denote the set of $M_t \times M$ matrices with orthonormal columns, $\lambda_i\{\mathbf{A}\}$ denotes the i^{th} largest singular value of \mathbf{A} , $\text{argmin}_{\mathcal{A}}$ ($\text{argmax}_{\mathcal{A}}$) to denote a function that returns a global minimizer (maximizer) over the set \mathcal{A} , and $\text{card}(\cdot)$ to denote the cardinality of a set.

II. SYSTEM OVERVIEW

The proposed system is illustrated in Fig. 1. A bit stream is sent into a vector encoder and modulator block where it is demultiplexed into M different bit streams. Each of the M bit streams is then modulated independently using the same constellation \mathcal{W} . This yields a symbol vector at time k of $\mathbf{s}_k = [s_{k,1} \ s_{k,2} \ \dots \ s_{k,M}]^T$. For convenience, we will assume that $E_{\mathbf{s}_k}[\mathbf{s}_k \mathbf{s}_k^*] = \mathbf{I}_M$.

The symbol vector \mathbf{s}_k is then multiplied by an $M_t \times M$ precoding matrix \mathbf{F} (which is chosen as a function of the channel using criteria to be described) producing a length M_t vector $\mathbf{x}_k = \sqrt{\frac{\mathcal{E}_s}{M}} \mathbf{F} \mathbf{s}_k$ where \mathcal{E}_s is the total transmit energy, M_t is the number of transmit antennas, and $M_t > M$. We assume throughout the correspondence that $M_r \geq M$. Assuming perfect timing, synchronization, sampling, and a memoryless linear matrix channel, this formulation allows the baseband, discrete-time equivalent received signal to be written as

$$\mathbf{y}_k = \sqrt{\frac{\mathcal{E}_s}{M}} \mathbf{H} \mathbf{F} \mathbf{s}_k + \mathbf{v}_k \quad (1)$$

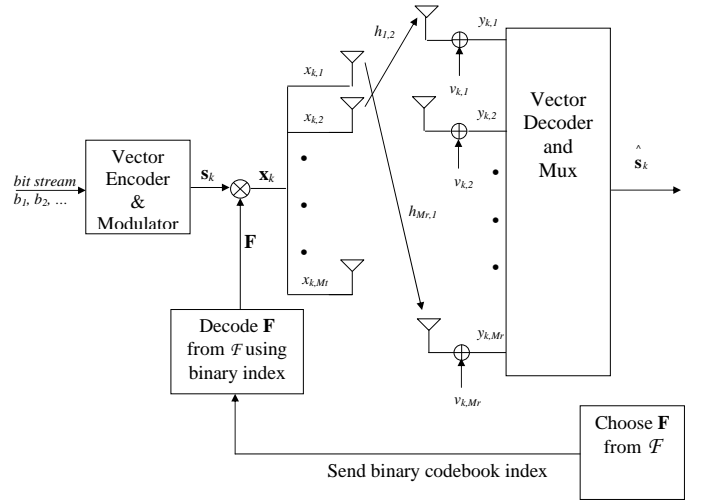


Fig. 1. Block diagram of a limited feedback precoding MIMO system.

where \mathbf{H} is the channel matrix and \mathbf{v}_k is the noise vector. We assume that the entries of \mathbf{H} are independent and identically distributed (i.i.d.) according to $\mathcal{CN}(0, 1)$ and the entries of \mathbf{v}_k are independent and distributed according to $\mathcal{CN}(0, N_0)$. The received vector is then decoded by a vector decoder, assuming perfect knowledge of $\mathbf{H}\mathbf{F}$, that produces a hard decoded symbol vector $\hat{\mathbf{s}}_k$.

In this correspondence, the receiver chooses a precoding matrix \mathbf{F} from a finite set of possible precoding matrices $\mathcal{F} = \{\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_N\}$ and conveys the index of the chosen precoding matrix back to the transmitter over a limited capacity, zero-delay feedback link. We assume that each $\mathbf{F} \in \mathcal{F}$ has unit column vectors that are orthogonal. This assumption is not especially restrictive since it follows from the form of the optimal, full channel knowledge precoders derived in [3] assuming a maximum singular value constraint on \mathbf{F} . Thus the proposed codebook will satisfy $\mathcal{F} \subseteq \mathcal{U}(M_t, M)$. To simplify implementation, we will typically assume that B bits of feedback are available; thus, the codebook consists of $N = 2^B$ matrices in $\mathcal{U}(M_t, M)$. The fact that the set \mathcal{F} is discrete allows the receiver to solve for \mathbf{F} by computing the selection metric of interest for each of the $N = 2^B$ codebook entries. The limitation of the codebook to 2^B matrices allows the system designer to constrain the precoding overhead and to take full advantage of the limited feedback channel.

To illustrate the concept, consider a codebook that corresponds to antenna subset selection [12]. Such a codebook would consist of the $\binom{M_t}{M}$ matrices consisting of M columns of \mathbf{I}_{M_t} . Notice that each set of M columns of \mathbf{I}_{M_t} is an element of $\mathcal{U}(M_t, M)$. Naturally, antenna selection precoding can be directly implemented in a limited feedback system because a total feedback of only $\lceil \log_2 \binom{M_t}{M} \rceil$ bits is required. Unfortunately, the performance is highly limited because (i) the columns of \mathbf{F} are restricted to being M columns of \mathbf{I}_{M_t} and (ii) the size of the codebook is limited by M and M_t . It is of interest to remove any restrictions about the nature of the elements of the codebook as well as the number of elements in an effort to come closer to the gains of approximately-optimal precoding.

The primary goal of this precoding is to improve the overall system performance using a sub-optimal receiver (though we do consider optimal ML decoding as well for completeness). For example, precoding can allow an easily implemented linear receiver to outperform optimal decoding at the expense of (i) using more transmit antennas and (ii) requiring channel information at the transmitter. It should be noted that depending on the SNR and the antenna configuration, feedback may not be needed. Using feedback to improve the data rate of the forward data path can cause a significant overhead on the reverse path. When the SNR is high, it is well-known that feedback is not needed to achieve capacity and the feedback overhead will be detrimental. We do not consider the effect of feedback overhead on the system.

As well, we assume that M is fixed. We do not vary the dimensionality of the transmitted data vector as a function of the channel matrix. This kind of adaptation will cause a large increase in both transmit and receive complexity. It should be noted, however, that varying M can dramatically improve the system performance. Readers are referred to [30] for discussion on this kind of adaptation.

Problem Statement. The objective of this correspondence is to solve the two key problems that are needed to effectively design and implement a limited feedback precoding system as proposed in Fig. 1. The first is to develop algorithms for selecting the optimal \mathbf{F} from \mathcal{F} as a function of the error probability or mutual information. This is the codeword or precoder selection problem. The second is to determine how to select a good codebook \mathcal{F} , based on a distortion measure that accounts for the fact that the channel is uncorrelated Rayleigh fading. This is the codebook design problem.

III. PRECODING CRITERIA

In this section, we discuss the criteria used for choosing the optimal precoding matrix from a given codebook. We outline criteria based on minimizing the error rate for the maximum likelihood or linear decoder and on maximizing the mutual information. When illustrative, we derive the optimal matrix over $\mathcal{U}(M_t, M)$.

A. Maximum Likelihood Receiver

The ML receiver solves the optimization problem

$$\hat{\mathbf{s}}_k = \underset{\mathbf{s} \in \mathcal{W}^M}{\operatorname{argmin}} \left\| \mathbf{y}_k - \sqrt{\frac{\mathcal{E}_s}{M}} \mathbf{H} \mathbf{F} \mathbf{s} \right\|_2^2. \quad (2)$$

A closed-form expression for the probability of symbol vector error is difficult to derive. One approach is to observe that the probability of symbol vector error can be upper bounded for high signal-to-noise ratios (SNR) using the vector Union Bound [31]. This approach is motivated by the fact that the Union Bound provides an adequately tight prediction of the probability of error for large SNR. Since we assume \mathcal{E}_s/N_0 to be fixed, the Union Bound is solely a function of the receive minimum distance $d_{\min,R}$ of the multidimensional constellation \mathcal{W}^M [13], which is given by

$$d_{\min,R} = \min_{\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{W}^M: \mathbf{s}_1 \neq \mathbf{s}_2} \sqrt{\frac{\mathcal{E}_s}{M}} \|\mathbf{H} \mathbf{F} (\mathbf{s}_1 - \mathbf{s}_2)\|_2. \quad (3)$$

The computation of $d_{\min,R}$ requires a search over $\binom{\operatorname{card}(\mathcal{W}^M)}{2}$ vectors.

Using (3), the minimum Euclidean distance criterion is to pick \mathbf{F} from the codebook \mathcal{F} for a given \mathbf{H} assuming that \mathcal{W} and \mathcal{E}_s/N_0 are fixed according to the following criterion.

ML-SC: ML Selection Criterion: Pick \mathbf{F} such that

$$\mathbf{F} = \underset{\mathbf{F}_i \in \mathcal{F}}{\operatorname{argmax}} d_{\min,R}. \quad (4)$$

Deriving a closed-form solution to *ML-SC* is difficult since the minimum distance depends on the constellation as well as the channel realization.

B. Linear Receiver

Linear receivers apply an $M \times M_r$ matrix \mathbf{G} , chosen according to some criterion, to produce $\hat{\mathbf{s}}_k = \mathbf{Q}(\mathbf{G} \mathbf{y}_k)$ where $\mathbf{Q}(\cdot)$ is a function that performs single-dimension ML decoding for each entry of a vector. Criteria will be presented for two different forms of \mathbf{G} : zero-forcing and minimum mean square error. For a zero-forcing (ZF) linear decoder, $\mathbf{G} = (\mathbf{H} \mathbf{F})^+$. When a minimum mean square error (MMSE) linear decoder is used $\mathbf{G} = [\mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F} + (N_0 M / \mathcal{E}_s) \mathbf{I}_M]^{-1} \mathbf{F}^* \mathbf{H}^*$.

1) *Minimum Singular Value:* We will characterize the average probability of symbol vector error performance using the substream with the minimum SNR following the results given in [12]. It was shown in [12] that the SNR of the k^{th} substream is given by

$$SNR_k^{(ZF)} = \frac{\mathcal{E}_s}{MN_0 [\mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F}]_{k,k}^{-1}} \quad (5)$$

for the ZF decoder and

$$SNR_k^{(MMSE)} = \frac{\mathcal{E}_s}{MN_0 [\mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F} + MN_0 / \mathcal{E}_s \mathbf{I}_M]_{k,k}^{-1}} - 1 \quad (6)$$

for the MMSE decoder, where $\mathbf{A}_{k,k}^{-1}$ is entry (k, k) of \mathbf{A}^{-1} . In [12] it is shown that in order to minimize a bound on the average probability of a symbol vector error, the minimum substream SNR must be maximized.

Using a selection criterion based on the minimum SNR requires the computation of the SNR of each of the M substreams and the estimation of the \mathcal{E}_s/N_0 . The computational complexity combined with the possibility of estimation error makes the minimum cumbersome to implement. For this reason, [12] shows that the minimum SNR for ZF can be bounded using

$$SNR_{\min}^{(ZF)} = \min_{1 \leq k \leq M} SNR_k^{(ZF)} \quad (7)$$

$$\geq \lambda_{\min}^2 \{\mathbf{H} \mathbf{F}\} \frac{\mathcal{E}_s}{MN_0} \quad (8)$$

where $\lambda_{\min}\{\mathbf{H} \mathbf{F}\}$ is the minimum singular value of $\mathbf{H} \mathbf{F}$.

We use (8) to obtain a requirement for choosing \mathbf{F} from \mathcal{F} for a given \mathbf{H} . We have assumed that \mathcal{F} and \mathcal{E}_s/N_0 are fixed.

MSV-SC: Minimum Singular Value Criterion: Pick \mathbf{F} such that

$$\mathbf{F} = \underset{\mathbf{F}_i \in \mathcal{F}}{\operatorname{argmax}} \lambda_{\min}\{\mathbf{H} \mathbf{F}_i\}. \quad (9)$$

This criterion provides a close approximation to maximizing the minimum SNR for dense constellations. The reason for

this is that as $\text{card}(W)$ grows large, the probability of an error vector lying collinear to the minimum singular value direction goes to one.

Optimal Unquantized Precoder

For comparison purposes, we also derive $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ that maximizes $\lambda_{\min}\{\mathbf{H}\mathbf{F}_{opt}\}$. Note that when the feasible set¹ is $\mathcal{U}(M_t, M)$, \mathbf{F}_{opt} is not unique. For example, if \mathbf{F}_{opt} maximizes $\lambda_{\min}\{\mathbf{H}\mathbf{F}_{opt}\}$ then so does $\mathbf{F}_{opt}\mathbf{U}$ for any $\mathcal{U}(M, M)$.

Let the singular value decomposition of \mathbf{H} be given by

$$\mathbf{H} = \mathbf{V}_L \mathbf{\Sigma} \mathbf{V}_R^* \quad (10)$$

where $\mathbf{V}_L \in \mathcal{U}(M_r, M_r)$, $\mathbf{V}_R \in \mathcal{U}(M_t, M_t)$, and $\mathbf{\Sigma}$ is an $M_r \times M_t$ diagonal matrix with $\lambda_k\{\mathbf{H}\}$ denoting the k^{th} largest singular value of \mathbf{H} , at entry (k, k) .

Lemma 1: An optimal precoder over $\mathcal{U}(M_t, M)$ for MSV-SC is $\mathbf{F}_{opt} = \bar{\mathbf{V}}_R$ where $\bar{\mathbf{V}}_R$ is a matrix constructed from the first M columns of \mathbf{V}_R .

Proof: Let $\tilde{\mathbf{F}} = [\tilde{\mathbf{F}}_{opt} \tilde{\mathbf{f}}_1 \dots \tilde{\mathbf{f}}_{M_t-M}]$ where $\tilde{\mathbf{F}}^* \tilde{\mathbf{F}} = \mathbf{I}_{M_t}$. It is clear that the matrix $\tilde{\mathbf{F}}^* \mathbf{H}^* \mathbf{H} \tilde{\mathbf{F}}$ is a Hermitian matrix and $\mathbf{F}_{opt}^* \mathbf{H}^* \mathbf{H} \mathbf{F}_{opt}$ is obtained from $\tilde{\mathbf{F}}^* \mathbf{H}^* \mathbf{H} \tilde{\mathbf{F}}$ by simply taking the principle submatrix corresponding to the first M rows. By the *Inclusion Principle* [32],

$$\begin{aligned} \lambda_{\min}\{\mathbf{H}\} &= \lambda_{M_t}\{\mathbf{H}\} \\ &\leq \lambda_{\min}\{\mathbf{H}\mathbf{F}_{opt}\} \\ &= \lambda_M\{\mathbf{H}\mathbf{F}_{opt}\} \\ &\leq \lambda_M\{\mathbf{H}\}. \end{aligned}$$

Here the M^{th} singular value refers to the singular value at entry (M, M) of the ordered diagonal matrix $\mathbf{\Sigma}$. For the case where $M_r \leq M_t$, there will be M_r different singular values. Because $M \leq M_r$, the M^{th} singular value is well-defined. This upper-bound can thus be achieved if $\mathbf{F}_{opt} = \bar{\mathbf{V}}_R$. ■

2) *Minimum Mean Squared Error:* Previous work [3] has considered improving the overall system performance by minimizing some function of the mean squared error (MSE) matrix

$$\overline{\text{MSE}}(\mathbf{F}, \mathbf{G}) = E \left[\left(\mathbf{G}\mathbf{y}_k - \sqrt{\frac{\mathcal{E}_s}{M}} \mathbf{s}_k \right) \left(\mathbf{G}\mathbf{y}_k - \sqrt{\frac{\mathcal{E}_s}{M}} \mathbf{s}_k \right)^* \right]$$

where the expectation is taken over \mathbf{s}_k and \mathbf{v}_k . When MMSE linear decoding is used, we express the MSE as

$$\overline{\text{MSE}}(\mathbf{F}) = \frac{\mathcal{E}_s}{M} \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F} \right)^{-1} \quad (11)$$

Using (11) we derive a selection criterion for choosing \mathbf{F} from \mathcal{F} .

MSE-SC: Mean Squared Error Criterion: Pick \mathbf{F} such that

$$\mathbf{F} = \underset{\mathbf{F}_i \in \mathcal{F}}{\text{argmin}} m(\overline{\text{MSE}}(\mathbf{F}_i)) \quad (12)$$

where $m(\cdot)$ is either $\text{tr}(\cdot)$ or $\text{det}(\cdot)$.

Note that minimizing the MSE does not specifically mean a reduction in the probability of error. In general, if the goal is to minimize the probability of error MSV-SC should be chosen.

Optimal Unquantized Precoder

Again, we present the optimal precoder over the unquantized

set $\mathcal{U}(M_t, M)$ for subsequent comparisons. In [3], various constraints on \mathbf{F}_{opt} were considered along with various mean squared error cost functions based on $\overline{\text{MSE}}(\mathbf{F}_{opt})$. Since we restrict our search to $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ we will consider the constraint in [3] where the maximum eigenvalue of $\mathbf{F}_{opt} \mathbf{F}_{opt}^*$ is unity. Note that all matrices in $\mathcal{U}(M_t, M)$ satisfy this constraint, but belonging to $\mathcal{U}(M_t, M)$ is not a necessary condition for this constraint.

It was shown in [3] that \mathbf{F}_{opt} that minimizes $\text{tr}(\overline{\text{MSE}}(\mathbf{F}_{opt}))$ or $\text{det}(\overline{\text{MSE}}(\mathbf{F}_{opt}))$ under this maximum eigenvalue constraint is $\mathbf{F}_{opt} = \bar{\mathbf{V}}_R$. Therefore we can state the following lemma as a consequence.

Lemma 2: (Scaglione et al [3]) A matrix $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ that minimizes either of the two cost functions $\text{tr}(\overline{\text{MSE}}(\mathbf{F}_{opt}))$ and $\text{det}(\overline{\text{MSE}}(\mathbf{F}_{opt}))$ is $\mathbf{F}_{opt} = \bar{\mathbf{V}}_R$.

Once again \mathbf{F}_{opt} is not unique because $\text{tr}(\overline{\text{MSE}}(\mathbf{F}_{opt})) = \text{tr}(\overline{\text{MSE}}(\mathbf{F}_{opt}\mathbf{U}))$ and $\text{det}(\overline{\text{MSE}}(\mathbf{F}_{opt})) = \text{det}(\overline{\text{MSE}}(\mathbf{F}_{opt}\mathbf{U}))$ for any $\mathbf{U} \in \mathcal{U}(M, M)$.

C. Capacity

In the context of antenna subset selection for spatial multiplexing systems, the mutual information (or capacity) has been used to formulate a precoder selection criterion [15], [16]. When the transmitter precodes with \mathbf{F} before transmission, the equivalent channel is $\mathbf{H}\mathbf{F}$. Thus the mutual information assuming an uncorrelated complex Gaussian source given \mathbf{H} and a fixed \mathbf{F} is

$$I(\mathbf{F}) = \log_2 \det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F} \right). \quad (13)$$

Therefore we can state a capacity inspired selection criterion as follows.

Capacity-SC: Capacity Selection Criterion: Pick \mathbf{F} such that

$$\mathbf{F} = \underset{\mathbf{F}_i \in \mathcal{F}}{\text{argmax}} I(\mathbf{F}_i). \quad (14)$$

Note that we call this selection criterion ‘‘Capacity-SC’’ for consistency with previous works [12], [15], [16].

Optimal Unquantized Precoder

It is possible to find the optimal unquantized precoder $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ for the *Capacity-SC* criterion.

Lemma 3: A precoder matrix $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ that maximizes $I(\mathbf{F}_{opt})$ is given by $\mathbf{F}_{opt} = \bar{\mathbf{V}}_R$.

Proof: Note that maximizing $\log_2 \det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}_{opt}^* \mathbf{H}^* \mathbf{H} \mathbf{F}_{opt} \right)$ is equivalent to maximizing $\det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}_{opt}^* \mathbf{H}^* \mathbf{H} \mathbf{F}_{opt} \right)$ and thus minimizing $\det \left(\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}_{opt}^* \mathbf{H}^* \mathbf{H} \mathbf{F}_{opt} \right)^{-1} \right)$. The

latter expression differs from $\text{det}(\overline{\text{MSE}}(\mathbf{F}))$ by a constant scale factor. It therefore follows from Lemma 2 that $\mathbf{F}_{opt} = \bar{\mathbf{V}}_R$ maximizes $I(\mathbf{F}_{opt})$. ■

Because of the relationship to *MSE-SC*, it is easily seen that $I(\mathbf{F}_{opt}) = I(\mathbf{F}_{opt}\mathbf{U})$ for any $\mathbf{U} \in \mathcal{U}(M, M)$.

IV. LIMITED FEEDBACK PRECODING: MOTIVATION AND CODEBOOK DESIGN

In the previous section we derived criteria for selecting the optimal precoding matrix. It is important that codebook \mathcal{F} is

¹The feasible set of an optimization is the domain that the cost function is optimized over.

designed specifically for the chosen criterion. To understand the codebook design problem, we first perform a probabilistic characterization of the optimal precoding matrix. We then use this characterization to derive codebooks that maximize average bounds on each of the performance criteria.

A. Probabilistic Characterization of Optimal Precoding Matrix

Let the eigenvalue decomposition of $\mathbf{H}^*\mathbf{H}$ be given by

$$\mathbf{H}^*\mathbf{H} = \mathbf{V}_R \Sigma^2 \mathbf{V}_R^* \quad (15)$$

with \mathbf{V}_R and Σ defined as in (10). In [33], it is shown that for a MIMO Rayleigh fading channel \mathbf{V}_R , the right singular vector matrix, is isotropically distributed on $\mathcal{U}(M_t, M_t)$, the group of unitary matrices. An *isotropically distributed* $M_t \times M$ matrix \mathbf{V} is a matrix where $\Theta^*\mathbf{V} \stackrel{d}{=} \mathbf{V}$ for all $\Theta \in \mathcal{U}(M_t, M_t)$ with $\stackrel{d}{=}$ denoting equivalence in distribution [34]. As stated in Section III, the optimal precoder for *MSV-SC*, *MSE-SC*, and *Capacity-SC* is constructed by simply taking the first M columns of \mathbf{V}_R . Using the isotropic distribution of \mathbf{V}_R , it is possible to derive the distribution of \mathbf{F}_{opt} .

Lemma 4: For a memoryless, i.i.d. Rayleigh fading channel \mathbf{H} , $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$ is isotropically distributed on $\mathcal{U}(M_t, M)$.

Proof: First note that

$$\mathbf{F}_{opt} = \overline{\mathbf{V}}_R = \mathbf{V}_R \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{(M_t-M) \times M} \end{bmatrix} \quad (16)$$

where $\mathbf{0}_{(M_t-M) \times M}$ is an $(M_t - M) \times M$ matrix of zeros. Since \mathbf{V}_R is isotropically distributed,

$$\begin{aligned} \Theta^* \mathbf{F}_{opt} &= \Theta^* \mathbf{V}_R \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{(M_t-M) \times M} \end{bmatrix} \\ &\stackrel{d}{=} \mathbf{V}_R \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{(M_t-M) \times M} \end{bmatrix} \\ &= \mathbf{F}_{opt}. \end{aligned}$$

■

Lemma 4 will allow the effect of the codebook on average distortion to be studied, with distortion to be defined later.

B. Grassmannian Subspace Packing

Before stating design criteria for each of the precoding matrix selection criteria, we present some relevant background about finite sets of matrices in $\mathcal{U}(M_t, M)$. The set $\mathcal{U}(M_t, M)$ defines the complex *Stiefel manifold* [35] of real dimension $2M_tM - M^2$. Each matrix in $\mathcal{U}(M_t, M)$ represents an M -dimensional subspace of \mathbb{C}^{M_t} . The set of all M -dimensional subspaces spanned by matrices in $\mathcal{U}(M_t, M)$ is the *complex Grassmann manifold*, denoted as $\mathcal{G}(M_t, M)$. Thus if $\mathbf{F}_1, \mathbf{F}_2 \in \mathcal{U}(M_t, M)$ then the column spaces of \mathbf{F}_1 and \mathbf{F}_2 , $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ respectively, are contained in $\mathcal{G}(M_t, M)$. Note that the Grassmann manifold can be analyzed using a real or complex Stiefel manifold [35], however, we will only make use of complex subspaces in this correspondence. Our codebook \mathcal{F} , which consists of a finite number of matrices chosen from $\mathcal{U}(M_t, M)$, thus represents a set, or packing, of subspaces in the Grassmann manifold. Designing sets of

N matrices that maximize the minimum subspace distance (where distance can be chosen in a number of different ways [36]) is known as *Grassmannian subspace packing*. We will use the interpretation of the precoding codebook \mathcal{F} as a packing of subspaces to simplify notation and analysis.

A normalized invariant measure μ is induced on $\mathcal{G}(M_t, M)$ by the Haar measure in $\mathcal{U}(M_t, M)$. This measure allows the computation of volumes within $\mathcal{G}(M_t, M)$. Subspaces within the Grassmann manifold can be related by their distance from each other [36]–[38]. A number of different distances can be defined [36], [39], but we will only make use of three. The *chordal distance* between the two subspaces $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ is

$$\begin{aligned} d_{chord}(\mathbf{F}_1, \mathbf{F}_2) &= \frac{1}{\sqrt{2}} \|\mathbf{F}_1 \mathbf{F}_1^* - \mathbf{F}_2 \mathbf{F}_2^*\|_F \\ &= \sqrt{M - \sum_{i=1}^M \lambda_i^2 \{\mathbf{F}_1^* \mathbf{F}_2\}}. \end{aligned}$$

The *projection two-norm distance* between two subspaces $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ is

$$d_{proj}(\mathbf{F}_1, \mathbf{F}_2) = \|\mathbf{F}_1 \mathbf{F}_1^* - \mathbf{F}_2 \mathbf{F}_2^*\|_2 = \sqrt{1 - \lambda_{\min}^2 \{\mathbf{F}_1^* \mathbf{F}_2\}}.$$

The *Fubini-Study distance* between two subspaces $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ is

$$d_{FS}(\mathbf{F}_1, \mathbf{F}_2) = \arccos |\det(\mathbf{F}_1 \mathbf{F}_2^*)|.$$

Each of these distances correspond to different ideas of distance between subspaces. The chordal distance generalizes the distance between points on the unit sphere through an isometric embedding from $\mathcal{G}(M_t, M)$ to the unit sphere [37]. Maximizing this distance corresponds to minimizing the sum of the eigenvalues of $\mathbf{F}_2^* \mathbf{F}_1 \mathbf{F}_1^* \mathbf{F}_2$ or similarly $\|\mathbf{F}_1^* \mathbf{F}_2\|_F^2$. The projection two-norm distance is maximized by minimizing the smallest singular value of $\mathbf{F}_1^* \mathbf{F}_2$, while the Fubini-Study distance is maximized by minimizing the product of the singular values of $\mathbf{F}_1^* \mathbf{F}_2$. Note that

$$\|\mathbf{F}_1^* \mathbf{F}_2\|_F^2 \geq M \lambda_{\min}^2 \{\mathbf{F}_1^* \mathbf{F}_2\} \geq M |\det(\mathbf{F}_1^* \mathbf{F}_2)|^2, \quad (17)$$

thus

$$\begin{aligned} d_{chord}(\mathbf{F}_1, \mathbf{F}_2) &\leq \sqrt{M} d_{proj}(\mathbf{F}_1, \mathbf{F}_2) \\ &\leq \sqrt{M} \sin(d_{FS}(\mathbf{F}_1, \mathbf{F}_2)). \end{aligned} \quad (18)$$

Let $\mathcal{S} = \{\mathcal{P}_{\mathbf{F}_1}, \mathcal{P}_{\mathbf{F}_2}, \dots, \mathcal{P}_{\mathbf{F}_N}\}$ be the packing of column spaces of the codebook matrices where $\mathcal{P}_{\mathbf{F}_i}$ is the column space of \mathbf{F}_i . Similarly to binary error correcting codes [36], a packing can be characterized by its minimum distance

$$\delta = \min_{1 \leq i < j \leq N} d(\mathbf{F}_i, \mathbf{F}_j)$$

where $d(\cdot, \cdot)$ is a distance function on $\mathcal{G}(M_t, M)$.

Consider the open ball in $\mathcal{G}(M_t, M)$ of radius $\gamma/2$ defined as

$$\mathcal{B}_{\mathbf{F}_i}(\gamma/2) = \{\mathcal{P}_{\mathbf{U}} \in \mathcal{G}(M_t, M) \mid d(\mathbf{U}, \mathbf{F}_i) < \gamma/2\}.$$

This metric ball can be defined with respect to any of the distance function on $\mathcal{G}(M_t, M)$. Note that if $d_{chord}(\mathbf{F}_1, \mathbf{F}_2) < \sqrt{1 - \rho^2}$, with $0 \leq \rho \leq 1$, then we are guaranteed that $d_{proj}(\mathbf{F}_1, \mathbf{F}_2) < \sqrt{1 - \rho^2}$ and $d_{FS}(\mathbf{F}_1, \mathbf{F}_2) < \arccos(\rho^M)$.

This follows by restricting the largest $M - 1$ singular values of $\mathbf{F}_1^* \mathbf{F}_2$ to be unity in order to find a lower bound on the minimum singular value. This observation yields

$$\mathcal{B}_{\mathbf{F}_i}^{chord}(\delta_{proj}/2) \subseteq \mathcal{B}_{\mathbf{F}_i}^{proj}(\delta_{proj}/2) \quad (19)$$

and

$$\mathcal{B}_{\mathbf{F}_i}^{chord} \left(\sqrt{1 - \cos^{2/M}(\delta_{FS}/2)} \right) \subseteq \mathcal{B}_{\mathbf{F}_i}^{FS}(\delta_{FS}/2). \quad (20)$$

The *density* of a subspace packing with respect to a distance γ ($\gamma \leq \delta$) is

$$\Delta(\gamma) = \mu \left(\bigcup_{i=1}^N \mathcal{B}_{\mathbf{F}_i}(\gamma/2) \right) = \sum_{i=1}^N \mu(\mathcal{B}_{\mathbf{F}_i}(\gamma/2))$$

where $\mathcal{B}_{\mathbf{F}_i}(\gamma/2)$ can be defined with respect to any distance function on the Grassmann manifold. The density of a packing is a measure of how well the codebook matrices “cover” $\mathcal{G}(M_t, M)$. The density allows the probability of the isotropically distributed $\bar{\mathbf{V}}_R$ falling in one of the set $\mathcal{B}_{\mathbf{F}_i}(\gamma/2)$, with $\gamma \leq \delta$, to be expressed as

$$Pr \left(\bar{\mathbf{V}}_R \in \bigcup_{i=1}^N \mathcal{B}_{\mathbf{F}_i}(\gamma/2) \right) = \Delta(\gamma). \quad (21)$$

Furthermore, (19) and (20) yield

$$\Delta_{chord}(\delta_{proj}) \leq \Delta_{proj}(\delta_{proj}) \quad (22)$$

and

$$\Delta_{chord} \left(2\sqrt{1 - \cos^{2/M}(\delta_{FS}/2)} \right) \leq \Delta_{FS}(\delta_{FS}) \quad (23)$$

where the subscript indicates the distance used. Notice that the factor of 2 in (23) follows from the fact that the Fubini-Study minimum distance is halved inside of the cosine function. For large M_t it has been shown in [36] that

$$\Delta_{chord}(\delta) \approx N \left(\frac{\delta}{2\sqrt{M}} \right)^{2M_t M + o(M_t)}. \quad (24)$$

C. Codebook Design Criteria

We now derive the codebook design criteria for each specific selection criterion using the distribution of the optimal unquantized precoding matrix derived in Section IV-A and the Grassmannian subspace packing results in Section IV-B.

ML-SC, MSV-SC, & MSE-SC (with Trace Cost Function)

We will show that the *ML-SC*, *MSV-SC*, and *MSE-SC* (with the trace cost function) selection criteria all relate to maximizing the average minimum singular value of the effective channel matrix $\mathbf{H}\mathbf{F}$. We will use this result to derive a codebook design criterion that relates to subspace packing on the Grassmann manifold using the projection two-norm distance.

Using (3), we can bound

$$d_{\min,R} \geq \sqrt{\frac{\mathcal{E}_s}{M}} \left(\min_{\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{W}^M: \mathbf{s}_1 \neq \mathbf{s}_2} \|\mathbf{s}_1 - \mathbf{s}_2\|_2 \right) \lambda_{\min}\{\mathbf{H}\mathbf{F}\}. \quad (25)$$

Thus maximizing the lower bound on $d_{\min,R}$ is equivalent to maximizing $\lambda_{\min}\{\mathbf{H}\mathbf{F}\}$. Thus this bound shows *ML-SC* requires maximizing $\lambda_{\min}\{\mathbf{H}\mathbf{F}\}$.

MSE-SC using the trace cost function chooses $\mathbf{F} \in \mathcal{F}$ that maximizes

$$tr \left(\frac{\mathcal{E}_s}{M} \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F} \right)^{-1} \right).$$

For high SNR, this can be approximated by $N_0 tr \left((\mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F})^{-1} \right)$, and we can bound

$$\min_{\mathbf{F}_i \in \mathcal{F}} N_0 tr \left((\mathbf{F}_i^* \mathbf{H}^* \mathbf{H} \mathbf{F}_i)^{-1} \right) \leq \min_{\mathbf{F}_i \in \mathcal{F}} \frac{MN_0}{\lambda_{\min}^2\{\mathbf{H}\mathbf{F}_i\}}. \quad (26)$$

The bound in (26) uses the fact that the maximum eigenvalue of $(\mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F})^{-1}$ is the inverse of the minimum eigenvalue of $\mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F}$. Minimizing the bound approximately minimizes the trace of the MSE matrix. Therefore, maximizing the $\lambda_{\min}\{\mathbf{H}\mathbf{F}\}$ is an approximate method for minimizing the trace of the MSE matrix.

Based on (25) and (26), we can relate the selection of the optimal codeword in a given codebook for the *ML-SC* and *MSE-SC* (using the trace cost function) cases to selection of the optimal codeword based on *MSV-SC*. To define a notion of an optimal codebook, we need a distortion measure with which to measure the average distortion. To design codebooks for the *ML-SC*, *MSE-SC*, and *MSV-SC* case we will use the error difference

$$\lambda_{\min}^2\{\mathbf{H}\mathbf{F}_{opt}\} - \lambda_{\min}^2\{\mathbf{H}\mathbf{F}_i\}$$

which is nonnegative for any choice of $\mathbf{F}_i \in \mathcal{F}$. Thus we will choose our codebook to minimize the average distortion

$$\begin{aligned} E_{\mathbf{H}} \left[\lambda_{\min}^2\{\mathbf{H}\mathbf{F}_{opt}\} - \max_{\mathbf{F}_i \in \mathcal{F}} \lambda_{\min}^2\{\mathbf{H}\mathbf{F}_i\} \right] \\ = E_{\mathbf{H}} \left[\lambda_M^2\{\mathbf{H}\} - \max_{\mathbf{F}_i \in \mathcal{F}} \lambda_{\min}^2\{\mathbf{H}\mathbf{F}_i\} \right]. \end{aligned} \quad (27)$$

Evaluating the expectation exactly in (27) is difficult; therefore, we will minimize an upper bound on the average distortion.

Using the singular value representation used in Section III and the properties of Grassmannian subspace packing,

$$\begin{aligned} E_{\mathbf{H}} \left[\max_{\mathbf{F}_i \in \mathcal{F}} \lambda_{\min}^2\{\mathbf{H}\mathbf{F}_i\} \right] &= E_{\mathbf{H}} \left[\max_{\mathbf{F}_i \in \mathcal{F}} \lambda_{\min}^2\{\mathbf{\Sigma}\mathbf{V}_R^* \mathbf{F}_i\} \right] \\ &\geq E_{\mathbf{H}} \left[\max_{\mathbf{F}_i \in \mathcal{F}} \lambda_{\min}^2\{\mathbf{\Sigma}\bar{\mathbf{V}}_R^* \mathbf{F}_i\} \right] \\ &\geq E_{\mathbf{H}} \left[\lambda_M^2\{\mathbf{H}\} \right] E_{\mathbf{H}} \left[\max_{\mathbf{F}_i \in \mathcal{F}} \lambda_{\min}^2\{\bar{\mathbf{V}}_R^* \mathbf{F}_i\} \right] \end{aligned} \quad (28)$$

where $\bar{\mathbf{\Sigma}}$ is the matrix constructed from the first M columns of $\mathbf{\Sigma}$. The result in (28) follows from the fact that singular values and singular vectors of complex normal matrices are independent [33], [35]. Due to the results in (27) and (28) we

obtain

$$\begin{aligned}
& E_{\mathbf{H}} \left[\lambda_{\min}^2 \{ \mathbf{H} \mathbf{F}_{opt} \} - \max_{\mathbf{F}_i \in \mathcal{F}} \lambda_{\min}^2 \{ \mathbf{H} \mathbf{F}_i \} \right] \\
& \leq E_{\mathbf{H}} \left[\lambda_M^2 \{ \mathbf{H} \} \right] E_{\mathbf{H}} \left[\left(1 - \max_{\mathbf{F}_i \in \mathcal{F}} \lambda_{\min}^2 \{ \bar{\mathbf{V}}_R^* \mathbf{F}_i \} \right) \right] \\
& \leq E_{\mathbf{H}} \left[\lambda_M^2 \{ \mathbf{H} \} \right] \cdot \\
& \quad \left(\frac{\delta_{proj}^2}{4} \Delta_{proj}(\delta_{proj}) + (1 - \Delta_{proj}(\delta_{proj})) \right) \\
& \quad \quad \quad (29) \\
& \lesssim E_{\mathbf{H}} \left[\lambda_M^2 \{ \mathbf{H} \} \right] \cdot \\
& \quad \left(1 + N \left(\frac{\delta_{proj}}{2\sqrt{M}} \right)^{2M_t M + o(M_t)} \left(\frac{\delta_{proj}^2}{4} - 1 \right) \right). \\
& \quad \quad \quad (30)
\end{aligned}$$

The bound in (29) is a result of partitioning the possible outcomes into two cases: (i) the subspace of $\bar{\mathbf{V}}_R$ falls within a codeword metric ball of radius δ_{proj} and (ii) the subspace of $\bar{\mathbf{V}}_R$ does not fall within a codeword metric ball. The codewords fall within a metric ball with probability $\Delta_{proj}(\delta_{proj})$ and must have distance less than $\delta_{proj}/2$ from some codeword when they fall within a metric ball. Substituting the density bound in (22) and the approximation in (24) results in (30). Differentiation and making the assumption that $2M_t M + o(M_t) > 2/3$, gives the following design criterion. We always have that $\delta_{proj} < 1$ so the function is a decreasing function of δ_{proj} .

Codebook Design Criterion: A codebook \mathcal{F} for a system using *ML-SC*, *MSV-SC*, or *MSE-SC* (using the trace cost function) to select \mathbf{F} from \mathcal{F} should be designed by maximizing the minimum projection two-norm distance between any pair of codeword matrix column spaces.

MSE-SC (with Determinant Cost Function) & Capacity-SC Similarly to the above, we will show that the *MSE-SC* (with the determinant cost function) and *Capacity-SC* selection criteria require maximizing the average system capacity. This result will be used to derive a codebook design criterion that relates to subspace packing on the Grassmann manifold using the Fubini-Study distance.

Selecting $\mathbf{F} \in \mathcal{F}$ using *MSE-SC* with the determinant cost function requires solving for \mathbf{F} that minimizes

$$\begin{aligned}
& \det \left(\frac{\mathcal{E}_s}{M} \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F} \right)^{-1} \right) \\
& = \left(\frac{\mathcal{E}_s}{M} \right)^M \left(\det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F} \right) \right)^{-1}.
\end{aligned}$$

This is equivalent to solving for the \mathbf{F} that maximizes $\tilde{I}(\mathbf{F}) \doteq \det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}^* \mathbf{H}^* \mathbf{H} \mathbf{F} \right)$, the same expression maximized in *Capacity-SC*. Using the SVD representation of \mathbf{H} ,

$$\begin{aligned}
\tilde{I}(\mathbf{F}) & = \det \left(\mathbf{F}^* \mathbf{V}_R \left(\mathbf{I}_{M_t} + \frac{\mathcal{E}_s}{MN_0} \boldsymbol{\Sigma}^T \boldsymbol{\Sigma} \right) \mathbf{V}_R^* \mathbf{F} \right) \\
& \geq \left| \det \left(\bar{\mathbf{V}}_R^* \mathbf{F} \right) \right|^2 \det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \bar{\boldsymbol{\Sigma}}^T \bar{\boldsymbol{\Sigma}} \right).
\end{aligned}$$

To define a notion of an optimal codebook, we need a distortion measure to measure the average performance loss in this case. Since \mathbf{F}_{opt} that maximizes the mutual information over $\mathcal{U}(M_t, M)$ gives $\tilde{I}(\mathbf{F}_{opt}) = \det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \bar{\boldsymbol{\Sigma}}^T \bar{\boldsymbol{\Sigma}} \right)$, we will use the error difference

$$\tilde{I}(\mathbf{F}_{opt}) - \left| \det \left(\bar{\mathbf{V}}_R^* \mathbf{F} \right) \right|^2 \det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \bar{\boldsymbol{\Sigma}}^T \bar{\boldsymbol{\Sigma}} \right)$$

which is nonnegative for any choice of $\mathbf{F}_i \in \mathcal{F}$. Thus we will choose our codebook to minimize the average distortion

$$E_{\mathbf{H}} \left[\det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \bar{\boldsymbol{\Sigma}}^T \bar{\boldsymbol{\Sigma}} \right) - \right. \quad (31)$$

$$\begin{aligned}
& \left. \max_{\mathbf{F}_i \in \mathcal{F}} \left| \det \left(\bar{\mathbf{V}}_R^* \mathbf{F}_i \right) \right|^2 \det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \bar{\boldsymbol{\Sigma}}^T \bar{\boldsymbol{\Sigma}} \right) \right] \\
& = E_{\mathbf{H}} \left[\det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \bar{\boldsymbol{\Sigma}}^T \bar{\boldsymbol{\Sigma}} \right) \cdot \right. \\
& \quad \left. \left(1 - E_{\mathbf{H}} \left[\max_{\mathbf{F}_i \in \mathcal{F}} \left| \det \left(\bar{\mathbf{V}}_R^* \mathbf{F}_i \right) \right|^2 \right] \right) \right]. \quad (32)
\end{aligned}$$

where (32) follows from the independence of $\boldsymbol{\Sigma}$ and \mathbf{V}_R [33], [35].

The distortion cost function can be bounded as

$$\begin{aligned}
& E_{\mathbf{H}} \left[\det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \bar{\boldsymbol{\Sigma}}^T \bar{\boldsymbol{\Sigma}} \right) \right] \cdot \\
& \quad \left(1 - E_{\mathbf{H}} \left[\max_{\mathbf{F}_i \in \mathcal{F}} \left| \det \left(\bar{\mathbf{V}}_R^* \mathbf{F}_i \right) \right|^2 \right] \right) \\
& \leq E_{\mathbf{H}} \left[\det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \bar{\boldsymbol{\Sigma}}^T \bar{\boldsymbol{\Sigma}} \right) \right] \cdot \\
& \quad \left(1 - \cos^2(\delta_{FS}/2) \Delta_{FS}(\delta_{FS}) \right) \quad (33) \\
& \lesssim E_{\mathbf{H}} \left[\det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \bar{\boldsymbol{\Sigma}}^T \bar{\boldsymbol{\Sigma}} \right) \right] \left(1 - N \cos^2(\delta_{FS}/2) \cdot \right. \\
& \quad \left. \left(\sqrt{\frac{1 - \cos^{2/M}(\delta_{FS}/2)}{M}} \right)^{2M_t M + o(M_t)} \right). \quad (34)
\end{aligned}$$

The result in (33) follows from the facts that the subspace of $\bar{\mathbf{V}}_R$ lies within a codeword metric ball with probability $\Delta_{FS}(\delta_{FS})$ and that all subspaces within the metric balls have distance less than $\delta_{FS}/2$. Using the density bound in (23) and the approximation in (24) yields (34). Differentiating this bound, and assuming that $M_t + o(M_t)/(2M) \geq (2^{1/M} - 1)$, tells us that we want to maximize δ_{FS} in order to minimize the distortion cost function. We can now state the following.

Codebook Design Criterion: A codebook \mathcal{F} for a system using *MSE-SC* with the determinant cost function or *Capacity-SC* to select \mathbf{F} from \mathcal{F} should be designed by maximizing the minimum Fubini-Study distance between any pair of codeword matrix column spaces.

Discussion

In summary, thinking of the codebook \mathcal{F} as a packing of M -dimensional subspaces rather than a set of $M_t \times M$ matrices allows us to bound the distortion for each of the selection criteria proposed in Section III. The distortion bound for *ML-SC*, *MSV-SC*, and *MSE-SC* with the trace cost function

is minimized by maximizing the minimum projection two-norm distance between any pair of codebook subspaces. The *Capacity-SC* and *MSE-SC* with the determinant cost function bound is minimized by maximizing the minimum Fubini-Study distance between any pair of codebook subspaces. Thus the codebook design is equivalent to subspace packing in the Grassmann manifold.

Observe that both design criteria make assumptions on the relation between M_t , M , and the $o(M_t)$ term. Numerical experiments have shown that for most $M_t > 2$ the assumptions are satisfied. When $M = 1$, it is also known that the $o(M_t)$ term is always -2 and (24) is an exact expression [20].

Finding good packings in the Grassmann manifold for arbitrary M_t , M , and N , and thus finding good codebooks, is difficult (see for example [37], [38], [40]). The problem is exasperated by the use of the projection two-norm and Fubini-Study distances instead of the more common chordal distance [36], [39]. For instance, in the simplest case of $M = 1$ where the Rankin lower bound on line packing correlation [38] can be employed, packings that achieve equality with the lower bound are often impossible to design. One simple method for designing good packings with arbitrary distance functions is to use the non-coherent constellation designs from [41]. We have found that the algorithms for constellation design in [41] yield codebooks with large minimum distances and can be easily modified to work with any distance function on the Grassmann manifold.

Note that this work generalizes the analysis in [20]. When $M = 1$, both the projection two-norm and Fubini-Study distance criteria minimize $\max_{1 \leq i < j \leq N} |\mathbf{f}_i^* \mathbf{f}_j|$ where the codebook is given by $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N\}$. This would be expected because the probability of error and capacity cost functions are both functions of the effective channel gain $\|\mathbf{H}\mathbf{f}\|_2^2$ when $M = 1$.

V. SIMULATIONS

Monte Carlo simulations were performed to illustrate the performance of Grassmannian precoders. The codebooks were designed using the criteria proposed in Section IV-C. To optimize the criteria, the codebooks were designed using the non-coherent constellations in [41]. Example codebooks can be downloaded at [42]. For each of the precoding systems using M_t transmit antennas and M substreams we also plotted the $M \times M_r$ spatial multiplexing results with both ZF and ML decoding. In addition we simulated the unquantized MMSE precoding using the trace cost function and both the sum power and maximum singular value constraints [3] and maximum minimum singular value antenna selection [12]. As well, we use the notation $M_t \times M_r$ to denote an M_t transmit and M_r receive antenna system.

Comparisons Between Receiver Architectures: This simulation used binary phase shift keying (BPSK) modulation and two substreams on a 4×2 wireless system. The results are shown in Fig. 2. We simulated six bit limited feedback precoding using *Capacity-SC*, *MSV-SC*, and *ML-SC*. ZF decoding and precoding using *Capacity-SC* provided more than a 4dB performance gain at a probability of symbol vector

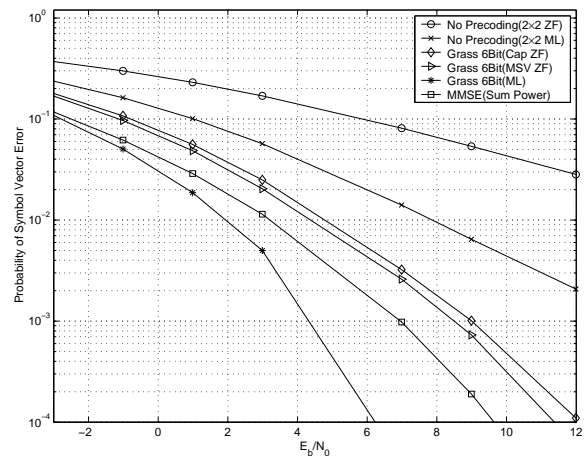


Fig. 2. Probability of symbol vector error comparison of various precoding schemes for a 2 substream 4×2 system using BPSK.

error of 2×10^{-3} over unprecoded decoding using ML decoding. Precoding using *MSV-SC* provided a 0.5dB gain over *Capacity-SC*. Unquantized MMSE precoding with the sum power constraint performs approximately 1.5dB better than limited feedback precoding using *MSV-SC*. As expected, ML decoding combined with *ML-SC* provided a large performance gain and outperformed unquantized MMSE precoding at a probability of symbol vector error of 10^{-3} by around 2.5dB.

Feedback Allows Lower Complexity Decoding: We simulated three substream precoding on a 6×3 wireless system in this experiment using 16QAM with results shown in Fig. 3. We used limited feedback precoding with six bits of feedback. Limited feedback precoding with *Capacity-SC* performed approximately the 0.25dB better than antenna selection. Limited feedback precoding using *MSV-SC* and *MSE-SC* with the trace cost function both performed approximately the same. They both provide around a 0.25dB improvement over *Capacity-SC* and perform within 1dB of unquantized optimal precoding using ZF decoding. Note that all selection criteria outperformed unprecoded spatial multiplexing using an ML receiver. This shows though the power of precoding: near ML or better than ML performance with low complexity receivers at the expense of feedback.

Comparison with Direct Channel Quantization: The purpose of this experiment is to demonstrate the problems associated with directly quantizing the matrix channel \mathbf{H} . The results are presented in Fig. 4. This experiment considered a 4×2 wireless system using two substreams and 16QAM. Directly quantizing the channel with sixteen bits of feedback performs approximately 4.7dB worse than a six bit limited feedback precoder at a probability of symbol vector error of 10^{-2} . The limited feedback precoder obtains performance approximately identical to that of the unquantized MMSE precoder with the maximum singular value power constraint.

VI. CONCLUSIONS AND FUTURE WORK

We proposed a system for precoding when only a limited feedback channel is available to convey channel state information. The essential component of this system is a codebook

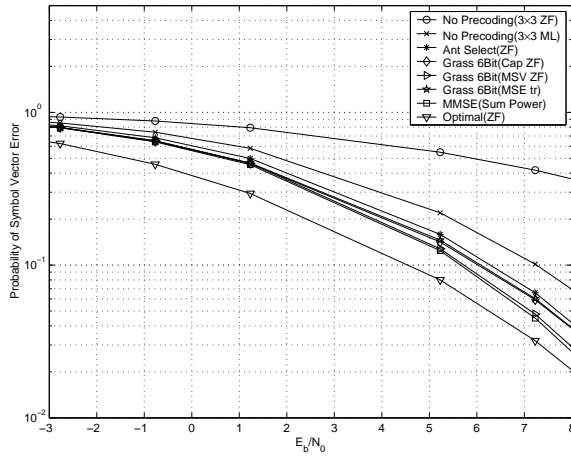


Fig. 3. Probability of symbol vector error comparison of various precoding schemes for a 3 substream 6×3 system using 16QAM.

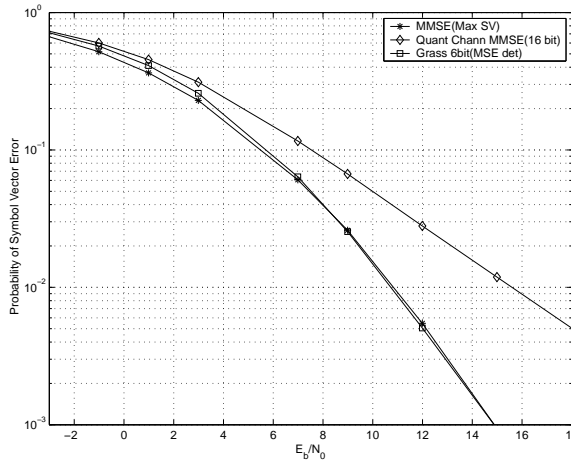


Fig. 4. Probability of symbol vector error comparison for direct channel quantization, unquantized MMSE, and limited feedback precoding on a 2 substream 4×2 system using 16QAM.

of precoding matrices that is known to both the transmitter and receiver. The codebook is designed offline based on the distribution of the channel. Observations of the channel at the receiver are used to determine the optimal precoder in the codebook and this index is conveyed to the transmitter.

In this correspondence, we presented solutions to the problems of precoder selection and codebook construction. We presented criteria for choosing the optimal precoder from the codebook according to different performance criteria and found that they were related directly to the unquantized codebook design criteria in [3] and the antenna selection criteria in [13]. We defined a notion of average distortion and used this to propose a codebook design criteria that minimizes a bound on the average distortion. We showed that the proposed design essentially relates to the problem of packing subspaces in the Grassmann manifold using the projection two-norm and the Fubini-Study distances. This problem of spacing subspaces is a famous problem in applied mathematics known as Grassmannian subspace packing [36], [37].

One important point of future work that is not addressed in this correspondence is the effect of delay and errors in

the feedback channel. This will lead to a degradation of the bit error rate performance compared with ideal channels. This analysis is currently being studied in the IEEE 802.16E standards body.

REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm.*, vol. 10, no. 6, pp. 585–595, 1999.
- [2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, March 1998.
- [3] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 1051–1064, May 2002.
- [4] H. Sampath, P. Stoica, and A. Paulraj, "Generalized precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, pp. 2198–2206, Dec. 2001.
- [5] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Info. Th.*, vol. 47, pp. 2632–2639, Sept. 2001.
- [6] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 2599–2613, Oct. 2002.
- [7] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Commun. Lett.*, vol. 6, pp. 239–241, June 2002.
- [8] Z. Hong, K. Liu, R. W. Heath, Jr., and A. M. Sayeed, "Spatial multiplexing in correlated fading via the virtual channel representation," *IEEE Jour. Select. Areas in Commun.*, vol. 21, pp. 856–866, June 2003.
- [9] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Int. Conf. on Commun.*, vol. 7, pp. 11–14, June 2001.
- [10] H. Bölcskei, R. U. Nabar, V. Erceg, D. Gesbert, and A. J. Paulraj, "Performance of multi-antenna signaling strategies in the presence of polarization diversity," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 2553–2562, Oct. 2002.
- [11] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Info. Th.*, vol. 49, pp. 1673–1690, July 2003.
- [12] R. W. Heath Jr., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, pp. 142–144, April 2001.
- [13] R. W. Heath Jr. and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. IEEE Int. Conf. on Commun.*, vol. 7, pp. 2276–2280, June 2001.
- [14] L. Collin, O. Berder, P. Rostaing, and G. Burel, "Optimal minimum distance-based precoder for MIMO spatial multiplexing systems," *IEEE Trans. Sig. Proc.*, vol. 52, pp. 617–627, March 2004.
- [15] D. A. Gore, R. U. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 5, pp. 2785–2788, June 2000.
- [16] S. Sandhu, R. U. Nabar, D. A. Gore, and A. Paulraj, "Near-optimal selection of transmit antennas for a MIMO channel based on Shannon capacity," in *Proc. 34th Asilomar Conf. on Sig., Sys., and Comp.*, vol. 1, Nov. 2000.
- [17] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, pp. 322–324, Aug. 2002.
- [18] 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 Jour. Select. Areas in Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.
- [19] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple antenna systems," *IEEE Trans. Info. Th.*, vol. 49, pp. 2562–2579, Oct. 2003.
- [20] D. J. Love, R. W. Heath Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Info. Th.*, vol. 49, pp. 2735–2747, Oct. 2003.
- [21] E. G. Larsson, G. Ganesan, P. Stoica, and W.-H. Wong, "On the performance of orthogonal space-time block coding with quantized feedback," *IEEE Commun. Lett.*, vol. 6, pp. 487–489, Nov. 2002.
- [22] D. J. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for orthogonal space-time block codes," *IEEE Trans. Sig. Proc.*, vol. 53, pp. 64–73, Jan. 2005.

- [23] G. Jöngren and M. Skoglund, "Quantized feedback information in orthogonal space-time block coding," *IEEE Trans. Info. Th.*, vol. 50, pp. 2473–2486, Oct. 2004.
- [24] R. S. Blum, "MIMO with limited feedback of channel state information," in *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 4, pp. 89–92, April 2003.
- [25] V. Lau, Y. Liu, and T.-A. Chen, "On the design of MIMO block-fading channels with feedback-link capacity constraint," *IEEE Trans. Commun.*, vol. 52, pp. 62–70, Jan. 2004.
- [26] M. Gharavi-Alkhansari and A. B. Gershman, "Fast antenna subset selection in wireless MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 5, pp. 57–60, April 2003.
- [27] A. Gorokhov, "Antenna selection algorithms for MEA transmission systems," in *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 3, pp. 2857–2860, May 2002.
- [28] A. F. Molisch, M. Z. Win, and J. H. Winters, "Capacity of MIMO systems with antenna selection," in *Proc. IEEE Int. Conf. on Commun.*, vol. 2, pp. 570–574, June 2001.
- [29] W. Santipach and M. L. Honig, "Asymptotic performance of MIMO wireless channels with limited feedback," in *Proc. IEEE Mil. Comm. Conf.*, vol. 1, pp. 141–146, Oct. 2003.
- [30] D. J. Love and R. W. Heath, Jr., "Multi-mode precoding using linear receivers for limited feedback MIMO systems," in *Proc. IEEE Int. Conf. on Commun.*, vol. 1, pp. 448–452, June 2004.
- [31] R. W. Heath Jr. and A. Paulraj, "Switching between spatial multiplexing and transmit diversity based on constellation distance," in *Proc. of Allerton Conf. on Comm. Cont. and Comp.*, Oct. 2000.
- [32] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, March 2001.
- [33] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Annals of Mathematical Statistics*, vol. 35, pp. 475–501, June 1964.
- [34] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Info. Th.*, vol. 45, pp. 139–157, Jan. 1999.
- [35] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Info. Th.*, vol. 48, pp. 359–383, Feb. 2002.
- [36] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Info. Th.*, vol. 48, pp. 2450–2454, Sept. 2002.
- [37] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, pp. 139–159, 1996.
- [38] T. Strohmer and R. W. Heath, Jr., "Grassmannian frames with applications to coding and communications," *Appl. Comput. Harmonic Analysis*, vol. 14, pp. 257–275, May 2003.
- [39] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthonormality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.
- [40] D. Agrawal, T. J. Richardson, and R. L. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Info. Th.*, vol. 47, pp. 2618–2626, Sept. 2001.
- [41] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Info. Th.*, vol. 46, pp. 1962–1973, Sept. 2000.
- [42] D. J. Love, "Tables of complex Grassmannian packings." available at <http://www.ece.purdue.edu/~djlove/grass.html>, 2004.

PLACE
PHOTO
HERE

David J. Love (S'98 - M'05) was born in Garland, Texas. He received the B.S. (with highest honors), M.S.E., and Ph.D. degrees in electrical engineering in 2000, 2002, and 2004, respectively, from The University of Texas at Austin. During the summers of 2000 and 2002, he was a summer researcher at the Texas Instruments DSPS R&D Center, Dallas, TX. Since August 2004, he has been with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, as an Assistant Professor. His research interests are in the design and

analysis of communication systems, coding theory, and information theory.

Prof. Love is a member of Tau Beta Pi and Eta Kappa Nu. In 2003, he was the recipient of the IEEE Vehicular Technology Society Daniel Noble Fellowship.

PLACE
PHOTO
HERE

Robert W. Heath, Jr. (S'96 - M'01) received the B.S. and M.S. degrees from the University of Virginia, Charlottesville, in 1996 and 1997, respectively, and the Ph.D. from Stanford University, Stanford, CA, in 2002, all in electrical engineering. From 1998 to 1999, he was a Senior Member of the Technical Staff at IospanWireless Inc, San Jose, CA, where he played a key role in the design and implementation of the physical and link layers of the first commercial MIMO-OFDM communication system. From 1999 to 2001, he served as a Senior

Consultant for IospanWireless Inc. In 2003, he founded MIMOWireless Inc., a consulting company dedicated to the advancement of MIMO technology. Since January 2002, he has been with the Department of Electrical and Computer Engineering, The University of Texas at Austin, where he serves as an Assistant Professor as part of the Wireless Networking and Communications Group. His research interests include interference management in wireless networks, sequence design, and all aspects of MIMO communication including antenna design, practical receiver architectures, limited feedback techniques, and scheduling algorithms.

Dr. Heath serves as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.