Reduced Feedback MIMO-OFDM Precoding and Antenna Selection

Tarkesh Pande, David J. Love, and James V. Krogmeier

Center for Wireless Systems and Applications
Purdue University
West Lafayette, IN 47907
{pande, djlove, jvk}@ecn.purdue.edu

June 18, 2006

Abstract

Transmitter precoding for multiple-input multiple-output orthogonal frequency division multiplexing (MIMO-OFDM) is an effective way of leveraging the diversity gains afforded by a multiple transmit-multiple receive antenna system in a frequency selective environment. In the limited feedback scenario, optimal precoder representation for narrowband MIMO systems using moderately sized codebooks designed on the Grassmann manifold has been shown to perform remarkably well. In MIMO-OFDM systems precoder matrices have to be designed for all subcarriers and the amount of feedback can get prohibitively large. This is especially true for next generation wireless local area networks and wireless metropolitan area networks which have a large number of subcarriers. In this paper, we present techniques to reduce this feedback requirement and the performance of these algorithms is numerically shown to provide improvement over existing schemes.

Index Terms- Closed-loop MIMO systems, Orthogonal frequency division multiplexing, antenna subset selection, Limited feedback, Interpolation, Geodesics

This work was supported in part by the SBC Foundation, the NSF under Grants CCF0513916 and CCF0130599, and the Indiana Twenty-First Century Research and Technology Fund.

1 Introduction

In order to meet the ever increasing demand for higher data rates and better system performance, the coupling of orthogonal frequency division multiplexing (OFDM) and multiple-input multiple-output (MIMO) systems is seen as a promising solution for next generation wireless networks. OFDM is a spectrally efficient modulation scheme that transforms a frequency selective channel into a set of parallel frequency-flat channels and therefore permits simple equalization schemes if the channel length is smaller than the length of the cyclic prefix. Space-time signaling with multiple antennas per transceiver has been shown to result in a significant diversity and/or capacity improvement especially when there is channel knowledge at the transmitter [1].

One effective way of leveraging the gains available from transmitter channel knowledge is to use linear precoding in which the signal is transmitted on the dominant modes of the channel matrix. This is done by multiplying the signal to be transmitted by a precoder matrix that is a function of the channel matrix. In time division duplexing (TDD) systems, channel state information (CSI) may be faithfully estimated at the transmitter using channel reciprocity. Frequency division duplexing (FDD) systems lack channel reciprocity. This problem, however, can be overcome by conveying the forward-link CSI from the receiver to the transmitter over a limited rate feedback channel. In this case the precoding matrices can be designed at the receiver and conveyed back to the transmitter using precoder codebooks that are known to both the transmitter and the receiver.

Precoder design for narrowband, space-time signaling and its representation with as small a codebook as possible is a well studied problem [2]–[4]. By relating precoder codebook design to the famous Grassmann subspace packing problem [5], the best codebooks

were determined for narrowband systems under different performance criteria. With these codebooks, a low rate feedback channel can provide near optimal performance. MIMO-OFDM systems, however, require precoder knowledge for *all* subcarriers. This can generate a prohibitively large amount of feedback if not properly designed.

In this paper, we address the problem of how to best represent the set of precoder matrices for all of the subcarriers in a MIMO-OFDM system with as few bits of feedback as possible. We present two different algorithms and compare their performance with those existing in the literature [6]–[8]. In the first algorithm, by representing the precoder matrices as points in the Grassmann manifold, we reformulate the problem as that of finding a curve in this manifold, which when appropriately sampled gives the precoding matrices for all the subcarriers. Our solution to this curve representation problem is to find the optimal precoding matrices with respect to the precoder criterion being used for pilots interspersed in a MIMO-OFDM symbol and connect them via geodesics. A geodesic is the shortest line connecting two points in a manifold. Uniformly sampling the geodesic gives the precoder matrices for the other subcarriers. One can think of the geodesic sampling approach as a linear interpolation between two points on the Grassmann manifold. This method can be applied to beamforming, precoded spatial multiplexing, and precoded space-time block coding systems.

Previous work in interpolation of precoders for MIMO-OFDM includes [6]–[8] and [9]. In [6]–[8] precoders for the non-pilot subcarriers are determined at the transmitter using a modified spherical interpolation of the pilot precoders. This interpolation requires either extra phase [6], [8] or unitary subspace rotation matrix [7] information to be fedback in addition to that of the pilot precoders. In [9], precoders for the non-pilot subcarriers are

found by solving a weighted least squares problem on the Grassmann manifold.

The second algorithm we propose is an alternative way to implement current clustering based solutions where precoders for pilots are reused at the transmitter for neighboring subcarriers in order to simplify implementation complexity. If the receiver has channel knowledge available for all the subcarriers, instead of sending back the precoders for the pilots, a better approach is to determine the mean of all the precoders in a cluster and feed the index of the precoder nearest to the mean back to the transmitter. Since the precoder matrices are points in the Grassmann manifold, the standard method of finding the mean matrix in a vector space by averaging does not apply. Instead, by using the property that the mean matrix minimizes the sum of the squares of distances from the sample matrices one can obtain a mean precoder matrix which is in the Grassmann manifold. This is also called the Karcher mean and is typically computed via iterative optimization algorithms. However, since the codebook sizes for the precoders are finite, we show that a simple search suffices to find the precoder nearest to the mean. Finally, a small modification to the cost function for finding the mean results in the definition of the generalized median which can also be computed via a brute force search. Both the Karcher mean and the generalized median show significant performance improvements over current clustering methods.

We also show how the Karcher mean and the generalized median approach can be used for antenna subset selection in MIMO-OFDM. In single-carrier MIMO systems, antenna subset selection corresponds to using M out of M_{tx} transmit antennas for improving the capacity/diversity performance [10]–[15]. In this case, the transmitter only needs M radio frequency (RF) chains with M switches at the front-end to choose from the M_{tx} antennas. Typically in MIMO-OFDM, if antenna subset selection has to be separately done for all

subcarriers, we need to implement all M_{tx} RF chains. Alternatively, if only one subset of antennas is used for the whole MIMO-OFDM symbol then we get the same RF chain savings as in the single carrier case. Naturally, this will result in a diversity loss as compared to the case where optimal antenna selection is done for all subcarriers. We show that if the antenna subset selection on a per MIMO-OFDM symbol basis is done by using the Karcher mean/generalized median approach we are still able to get a significant performance gain with only M RF chains necessary at the transmitter as compared to the case where there are no extra antennas.

The rest of the paper is organized as follows. In Section II, we give a general overview of precoding for MIMO-OFDM and the current techniques used for reducing the feedback requirement. We explain some of the properties of geodesics and show how they are constructed in Section III. The proposed algorithms are then developed in Section IV and V followed by a complexity analysis in Section VI. Numerical results are presented in Section VII, and we conclude in Section VIII.

2 MIMO-OFDM Precoding

Consider an N subcarrier MIMO-OFDM system with M_{tx} transmit and M_{rx} receive antennas which uses limited feedback precoding as in Fig. 1. We assume that the channel between each transmit and receive antenna pair has a multipath structure with L taps. The impulse response for this frequency selective channel can therefore be written as

$$\mathbf{G}(k) = \sum_{l=0}^{L-1} \mathbf{G}(l)\delta(k-l) \tag{1}$$

where $\delta(\cdot)$ is the Kronecker delta function and $\mathbf{G}(l) \in \mathbb{C}^{M_{rx} \times M_{tx}}$ for all l.

If the cyclic prefix is designed to be of at least length $L-1$ then the MIMO channel

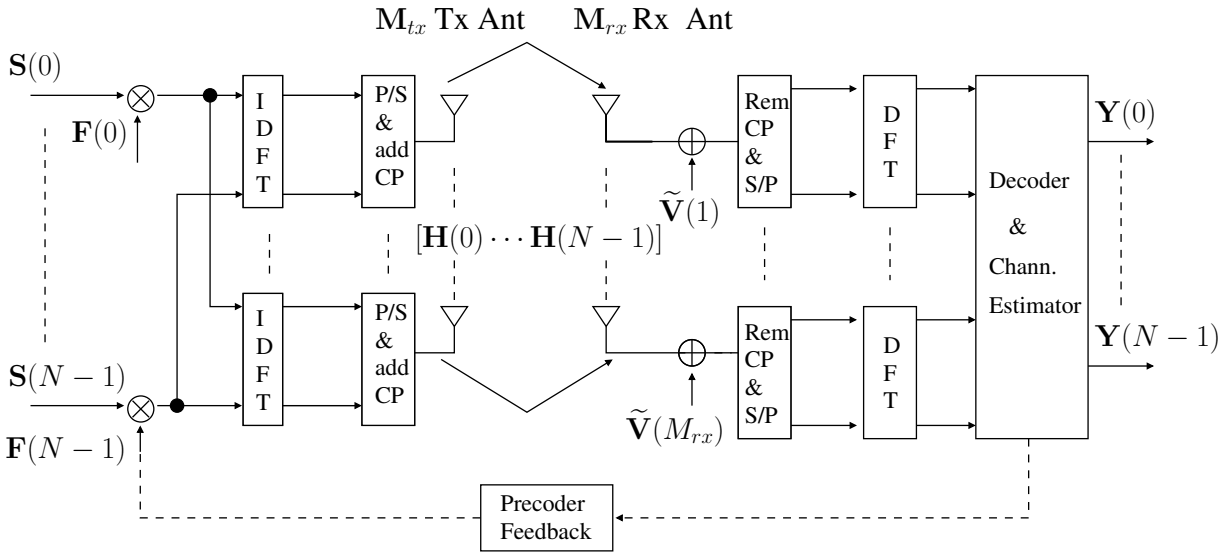

Figure 1: MIMO-OFDM system model with limited feedback precoding for all subcarriers. The tilde on the noise denotes that it is in the time domain.

seen by the data in the n^{th} subcarrier (where $n = 0, 1, \ldots, N-1$) is given by

$$\mathbf{H}(n) = \frac{1}{\sqrt{N}} \sum_{l=0}^{L-1} \mathbf{G}(l) \exp\left(-j\frac{2\pi nl}{N}\right) \tag{2}$$

where $j = \sqrt{-1}$. The received signal on the n^{th} subcarrier can then be written as

$$\mathbf{Y}(n) = \mathbf{H}(n)\mathbf{X}(n) + \mathbf{V}(n) \tag{3}$$

where $\mathbf{V}(n)$ is an $M_{rx} \times T$ matrix with $\mathcal{CN}(0,1)$ entries, $\mathbf{X}(n)$ is an $M_{tx} \times T$ transmitted matrix with $\text{tr}\left(E[\mathbf{X}(n)^H\mathbf{X}(n)]\right) = T\rho$ where ρ denotes the signal to noise ratio (SNR), and T is the number of channel uses involved in transmitting one space-time signal. In spatial multiplexing and beamforming the channel is used once ($T = 1$) to transmit a space-time signal [16], [17], while for orthogonal space-time block coding the channel has to be used multiple times ($T > 1$) [18], [19].

In linear precoding, the transmitted matrix $\mathbf{X}(n)$ can be written as

$$\mathbf{X}(n) = \mathbf{F}(n)\mathbf{S}(n) \tag{4}$$

where $\mathbf{S}(n)$ is an $M \times T$ space-time signal (with $1 \leq M \leq M_{tx}$) and the precoding matrix $\mathbf{F}(n)$ is $M_{tx} \times M$ (with $1 \leq M \leq M_{tx}$) and designed as a function of the channel $\mathbf{H}(n)$. To constrain the peak transmitted power, we will assume that $\mathbf{F}(n) \in \mathcal{U}(M_{tx}, M)$ where

$$\mathcal{U}(M_{tx}, M) = \{\mathbf{F} \in \mathbb{C}^{M_{tx} \times M} \mid \mathbf{F}^H \mathbf{F} = \mathbf{I}\}$$

with $\mathbf{I}$ denoting the identity matrix. This kind of linear precoding has been used for beamforming [2]–[4], precoded orthogonal space-time block codes [15], [20], and precoded spatial multiplexing [21], [22]. The precoder $\mathbf{F}(n)$ is designed to optimize some criteria examples of which are i) maximizing the minimum singular value of $\mathbf{H}(n)\mathbf{F}(n)$, ii) maximizing the capacity of the effective channel $\mathbf{H}(n)\mathbf{F}(n)$, or iii) maximizing the Frobenius norm of $\mathbf{H}(n)\mathbf{F}(n)$.

As an example, consider case iii) for beamforming with maximum ratio transmission/maximum ratio combining (MRT/MRC). Then the post-processed received signal on every subcarrier can be written as

$$\mathbf{Z}(n) = \mathbf{F}^H(n)\mathbf{H}^H(n)\{\mathbf{H}(n)\mathbf{F}(n)\mathbf{S}(n) + \mathbf{V}(n)\}. \tag{5}$$

The beamforming vector $\mathbf{F}(n)$ is an $M_{tx} \times 1$ unit norm vector while $\mathbf{S}(n)$ is a 1×1 complex data symbol chosen from a constellation set. In order to minimize the average error probability the signal to noise ratio (SNR)

$$\gamma(n) = \frac{\rho \mid \mathbf{F}^H(n)\mathbf{H}^H(n)\mathbf{H}(n)\mathbf{F}(n) \mid^2}{\|\mathbf{F}^H(n)\mathbf{H}^H(n)\|_2^2} \tag{6}$$

in every subcarrier should be maximized. The beamforming vector $\mathbf{F}(n)$ is chosen to maximize the two norm of $\mathbf{H}(n)\mathbf{F}(n)$ [3]. Equivalently, $\mathbf{F}(n)$ is the unit-norm right singular vector corresponding to the largest singular value of the channel matrix $\mathbf{H}(n)$.

All of the above precoding optimization criteria are invariant to right multiplication of $\mathbf{F}(n)$ by an $M \times M$ unitary matrix $\mathbf{U} \in \mathcal{U}(M, M)$. Thus, the optimization criteria depends only on the *column space* of $\mathbf{F}(n)$. The set of all M-dimensional subspaces spanned by matrices in $\mathcal{U}(M_{tx}, M)$ is the complex Grassmann manifold, denoted as $\mathcal{G}(M_{tx}, M)$. Since the Grassmann manifold is a quotient space[1], a point[2] $\mathbf{F} \in \mathcal{G}(M_{tx}, M)$ represents an equivalence class of $M_{tx} \times M$ orthogonal matrices. Two matrices are equivalent if and only if their columns span the same M-dimensional subspace i.e., the equivalence class of $\mathbf{F}$ is defined as $[\mathbf{F}] = \{\mathbf{FU} : \mathbf{U} \in \mathcal{U}(M, M)\}$ [24]. The precoders are therefore points on the Grassmann manifold.

In limited feedback design, the precoder $\mathbf{F}(n)$ is further restricted to lie in a codebook $\mathcal{F} = \{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_{2^B}\}$ where B is a positive integer and the codebook is known to both transmitter and receiver. Based on channel knowledge and precoding criteria, the receiver sends back B bits representing the codebook index of the precoder for each subcarrier. The design of approximately optimal codebooks for different space-time coding systems was analyzed in [3], [4], [20], [22], [25] and was shown to relate to the famous Grassmannian subspace packing problem. In MIMO-OFDM systems the linear precoding matrix $\mathbf{F}(n)$ must be designed as a function of $\mathbf{H}(n)$ for *every* subcarrier; therefore, the number of feedback bits $N \cdot B$ can become prohibitively large. Since neighboring subcarrier channel matrices are highly correlated, one may design algorithms which use knowledge of only a small set of $\mathbf{F}(n)'s$ to infer the others. The two main design methodologies are interpolation based schemes and cluster based approaches.

[1]The quotient space of a topological space Y is the set of equivalence classes relative to some given equivalence relation on a given topological space X [23].

[2]We interchangably use $\mathbf{F}$ to refer to the matrix $\mathbf{F}$ and the subspace defined by $\mathbf{F}$. We also use $|\cdot|$ to interchangeably refer to both the cardinality and absolute value.

A conventional interpolation scheme involves a system where K pilots with known data are transmitted in an N-subcarrier MIMO-OFDM symbol. The pilots are uniformly spaced in a comb-type arrangement. The received space-time signal on the pilot subcarriers can be written as

$$\mathbf{Y}(kN/K) = \mathbf{H}(kN/K)\underline{\mathbf{X}}(kN/K) + \mathbf{V}(kN/K) \tag{7}$$

for $k = 0, \ldots, K-1$ where N/K is an integer. The matrices $\{\underline{\mathbf{X}}(kN/K)\}$ comprise a training sequence that is known to both the transmitter and the receiver allowing the receiver to determine an estimate of $\mathbf{H}(kN/K)$ for $0 \leq k \leq K-1$. The receiver then uses the estimated channel to design precoders for each pilot and sends that information to the transmitter where interpolation is done to determine the precoders for the other subcarriers. To determine the minimum number of pilots required note that if N_{train} MIMO-OFDM symbols are used for channel estimation i.e., training is done over multiple symbols, then the minimum number of of pilot tones for minimizing the mean square error of the least squares channel estimate must satisfy the inequality $K \cdot N_{train} \geq L \cdot M_{tx}$ [26].

In [6]–[8] the conventional linear interpolator is modified for interpolating between pilot precoders in the MRT/MRC and spatial multiplexing MIMO-OFDM systems. An additional optimization parameter $\mathbf{Q}$ is introduced and the interpolator takes the form

$$\mathbf{Z}(kN/K + m) = (1 - \alpha_m)\mathbf{F}(kN/K) + \alpha_m \mathbf{Q}\mathbf{F}((k+1)N/K) \tag{8}$$

followed by the orthonormalization

$$\hat{\mathbf{F}}(kN/K + m; \mathbf{Q}) = \mathbf{Z}(kN/K + m)\{\mathbf{Z}(kN/K + m)^H \mathbf{Z}(kN/K + m)\}^{-1/2}. \tag{9}$$

where $1 \leq m \leq N/K - 1$, $\alpha_m = mK/N$ is the weighting coefficient and $\mathbf{F}(N) = \mathbf{F}(0)$. $\mathbf{Q}$ is an $M \times M$ unitary rotation matrix in the spatial multiplexing case and reduces to the

complex scalar parameter $e^{j\theta}$ in the beamforming scenario. Every pair of pilot precoders with which interpolation is done will have an optimizing $\mathbf{Q}$ value which must also be fedback. In the beamforming case, it is chosen from a codebook $\mathcal{Q}$ based on maximizing the effective channel gain for the subcarrier furthest away from the pilots

$$\mathbf{Q} = \underset{\mathbf{Q}\in\mathcal{Q}}{\operatorname{argmax}} \left\| \mathbf{H}(N/K(k+1/2))\hat{\mathbf{F}}(N/K(k+1/2)) \right\|_2^2. \tag{10}$$

Similarly, for the spatial multiplexing scenario, a minimizing the mean square error criteria [7] or maximizing the minimum singular value of the effective channel for the subcarrier furthest away from the pilots may be used to determine $\mathbf{Q}$. The above interpolator performs interpolation first on Euclidean space and then projects the solution back to the Grassmann manifold. A different way would be to directly construct the interpolator on the Grassmann manifold.

An alternative to interpolation is to reuse the pilot precoders for neighboring subcarriers. An N subcarrier MIMO-OFDM symbol can be divided into K groups or clusters with each cluster having N/K subcarriers. The precoder matrix for the center subcarrier in each cluster is reused for the remaining subcarriers in that cluster. This is simpler to implement at the transmitter than interpolation but suffers from a performance loss. In both cases a total of $K \cdot B$ bits of feedback are sent back for all the K precoders. If $N \gg K$, this results in a huge saving in the number of feedback bits.

To develop our interpolation and clustering algorithms which make use of the geometrical properties of the Grassmann manifold, we need to discuss the representation of geodesics.

3 Geodesics on the Grassmann Manifold

Given two points on a manifold, a geodesic is the curve lying on the manifold of shortest length connecting them. In Euclidean space this corresponds to a straight line; on a sphere it corresponds to a great circle. A geodesic starting from a point $\mathbf{F}(t_0)$ in $\mathcal{G}(M_{tx}, M)$ is given by the parametric equation [27]

$$\mathbf{F}(t) = \mathbf{Q}(t_0) \exp((t - t_0)\mathbf{B}) \begin{pmatrix} \mathbf{I}_{M \times M} \\ \mathbf{0}_{(M_{tx} - M) \times M} \end{pmatrix}. \tag{11}$$

The matrix $\mathbf{Q}(t_0) = [\mathbf{F}(t_0)\ \mathbf{F}_\perp(t_0)]$ is formed by some orthonormal completion of $\mathbf{F}(t_0)$, i.e., the column space spanned by $\mathbf{F}_\perp(t_0)$ where $\mathbf{F}_\perp(t_0) \in \mathcal{U}(M_{tx}, M_{tx} - M)$ spans the column null space of $\mathbf{F}^T(t_0)$. The skew symmetric matrix $\mathbf{B}$ has the form

$$\mathbf{B} = \begin{pmatrix} 0 & \mathbf{A}^H \\ -\mathbf{A} & 0 \end{pmatrix}, \qquad \mathbf{A} \in \mathbb{C}^{(M_{tx} - M) \times M} \tag{12}$$

where the sub-matrix $\mathbf{A}$ specifies the direction and movement of the geodesic curve with parameter t. The vector of principal subspace angles $\underline{\phi} = [\phi_1, \ldots, \phi_M] \in [0, \pi/2]^M$ between two points $\mathbf{F}(t_0)$ and $\mathbf{F}(t_1) \in \mathcal{G}(M_{tx}, M)$ can be computed via the singular value decomposition (SVD) [28]

$$\mathbf{F}^H(t_0)\mathbf{F}(t_1) = \mathbf{U}_1 \cos(\mathbf{\Phi}) \mathbf{V}_1^H. \tag{13}$$

In (13), $\mathbf{U}_1$ and $\mathbf{V}_1 \in \mathcal{U}(M, M)$ and $\underline{\phi}$ are the diagonal elements of $\mathbf{\Phi} \in \mathbb{R}^{M \times M}$. The $\cos(\cdot)$ matrix above corresponds to taking the cosine of only the diagonal entries of $\mathbf{\Phi}$.

The direction matrix $\mathbf{A}$ relates how the principal subspace angles between $\mathbf{F}(t)$ and

$\mathbf{F}(t_0)$ change with t. Let the compact SVD[3] of $\mathbf{A}$ be given by

$$\mathbf{A} = \mathbf{U}_2 \widehat{\mathbf{\Sigma}} \mathbf{U}_1^H \tag{14}$$

where $\mathbf{U}_2 \in \mathcal{U}(M_{tx} - M, M)$, $\mathbf{U}_1 \in \mathcal{U}(M, M)$ and $\widehat{\mathbf{\Sigma}} \in \mathbb{R}^{M \times M}$ is a real diagonal matrix. If ϕ_i $(i = 1, \ldots, M)$ are the principal angles between $\mathbf{F}(t)$ and $\mathbf{F}(t_0)$, then it can be shown that for t close to t_0, $\phi_i = (t - t_0)\sigma_i$ where σ_i are the diagonal elements of $\widehat{\mathbf{\Sigma}}$ [24].

All distance metrics between subspaces in $\mathcal{G}(M_{tx}, M)$ are functions of the principal subspace angles. Some examples are [24], [29]:

Geodesic distance

$$d_g(\mathbf{F}(t_1), \mathbf{F}(t_0)) = \left\| \underline{\phi} \right\|_2 ;$$

Chordal distance

$$d_c(\mathbf{F}(t_1), \mathbf{F}(t_0)) = \frac{1}{\sqrt{2}} \left\| \mathbf{F}^H(t_1)\mathbf{F}(t_1) - \mathbf{F}^H(t_0)\mathbf{F}(t_0) \right\|_F = \left\| \sin(\underline{\phi}) \right\|_2 ;$$

Fubini-Study distance

$$d_{FS}(\mathbf{F}(_1), \mathbf{F}(t_0)) = \arccos \left| \det \mathbf{F}(t_1)\mathbf{F}^H(t_0) \right| = \arccos \left(\prod_{i=1}^{M} \cos(\phi_i) \right) ;$$

Projection 2-norm

$$d_{F2}(\mathbf{F}(t_1), \mathbf{F}(t_0)) = \left\| \mathbf{F}^H(t_1)\mathbf{F}(t_1) - \mathbf{F}^H(t_0)\mathbf{F}(t_0) \right\|_2 = \left\| \sin(\underline{\phi}) \right\|_\infty .$$

Note that as we move along a geodesic curve given by (11) both the geodesic distance and principal angles between the initial and final subspaces increase linearly with t.

[3]The compact singular value decomposition (SVD) of a matrix $\mathbf{A} \in \mathbb{C}^{(n-k)\times k}$ with $n - k \geq k$ is $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}$ where $\mathbf{U} \in \mathcal{U}(n - k, k), \mathbf{V} \in \mathcal{U}(k, k)$ and $\mathbf{\Sigma} \in \mathbb{R}^{k \times k}$. This representation is useful when only the first k left singular vectors are required. Note that in a regular SVD $\mathbf{U} \in \mathcal{U}(n - k, n - k)$.

4 Reduced Feedback Algorithms

We now develop the different approaches of interpolation and cluster based precoding. For the geodesic interpolator, precoder knowledge for only the pilots is necessary. For the clustering scheme, we will assume the receiver knows the optimal precoders for all subcarriers. We will not focus on channel estimation techniques which can be found in [26],[30]–[35] and the references therein. Rather, we are interested in answering the question of given a fixed feedback rate, what is the best (in terms of error rate performance) we can do for beamforming, precoded spatial multiplexing systems, and precoded space-time coding.

4.1 Geodesic Interpolation

If the optimizing precoders for two consecutive pilot subcarriers are calculated, uniformly sampling the connecting geodesic gives the precoder matrices for the data subcarriers. The subspace angles between consecutive precoders then differ by a constant. Since precoders are designed with respect to the frequency domain subcarrier channel matrices, to avoid any ambiguity we will use the variable f instead of t for parameterizing geodesics. In this case with the two pilot precoders corresponding to points $\mathbf{F}(f_0)$ and $\mathbf{F}(f_1)$ in $\mathcal{G}(M_{tx}, M)$, determining the direction matrix $\mathbf{A}$ such that (11) is satisfied for $f = f_0$ and f_1, gives us the desired geodesic equation. Without loss of generality we may assume $f_0 = 0$ and $f_1 = N/K$ since pilots are spaced N/K subcarriers apart. In [27], three key observations are made to determine the connecting geodesic without explicitly computing $\mathbf{A}$.

1. The principle subspace angles can be calculated using knowledge of only $\mathbf{F}(N/K)$

and $\mathbf{F}(0)$ by noting that

$$\mathbf{F}^H(0)\mathbf{F}(N/K) = \mathbf{U}_1 \cos\left(\frac{N}{K}\mathbf{\Sigma}\right)\mathbf{V}_1^H \tag{15}$$

where $\mathbf{U}_1$ and $\mathbf{V}_1 \in \mathcal{U}(M, M)$ and $\mathbf{\Sigma} = \frac{K}{N}\widehat{\mathbf{\Sigma}}$ where $\widehat{\mathbf{\Sigma}}$ is defined in (14) . In the $\cos(\cdot)$ matrix above the non-diagonal entries are zero.

2. The matrix exponential $\exp(f\mathbf{B})$ can be written as

$$\exp(f\mathbf{B}) = \begin{pmatrix} \mathbf{U}_1 & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{U}}_2 \end{pmatrix} \begin{pmatrix} \cos(\mathbf{\Sigma}f) & \sin(\mathbf{\Sigma}f) & \mathbf{0} \\ -\sin(\mathbf{\Sigma}f) & \cos(\mathbf{\Sigma}f) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{M_{tx}-2M} \end{pmatrix} \begin{pmatrix} \mathbf{U}_1 & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{U}}_2 \end{pmatrix}^H \tag{16}$$

where $\widehat{\mathbf{U}}_2 \in \mathcal{U}(M_{tx} - M, M_{tx} - M)$ and $\mathbf{U}_1$ is defined in the previous equation.

3. Finally, we can force (15) to be a diagonal matrix by defining the canonical bases as $\overline{\mathbf{F}}(0) = \mathbf{F}(0)\mathbf{U}_1$ and $\overline{\mathbf{F}}(N/K) = \mathbf{F}(N/K)\mathbf{V}_1$. The geodesic connecting these two subspaces is found by right multiplication of (11) by $\mathbf{U}_1$ and is given by

$$\begin{aligned} \overline{\mathbf{F}}(f) &= \mathbf{Q}(0)\exp(f\mathbf{B}) \begin{pmatrix} \mathbf{I}_{M\times M} \\ \mathbf{0}_{(M_{tx}-M)\times M} \end{pmatrix} \mathbf{U}_1 \\ &= \mathbf{Q}(0) \begin{pmatrix} \mathbf{U}_1 \cos(\mathbf{\Sigma}f) \\ -\mathbf{U}_2 \sin(\mathbf{\Sigma}f) \end{pmatrix} \\ &= \mathbf{F}(0)\mathbf{U}_1 \cos(\mathbf{\Sigma}f) - \mathbf{F}_\perp(0)\mathbf{U}_2 \sin(\mathbf{\Sigma}f), \qquad 0 \le f \le N/K \end{aligned} \tag{17}$$

where $\mathbf{U}_2$ corresponds to the first M columns of $\widehat{\mathbf{U}}_2$. Evaluating (17) at $f = N/K$ allows us to solve for the unknown $\mathbf{F}_\perp(0)\mathbf{U}_2$

$$\begin{aligned} & \overline{\mathbf{F}}(N/K) = \mathbf{F}(0)\mathbf{U}_1 \cos\left(\frac{N}{K}\mathbf{\Sigma}\right) - \mathbf{F}_\perp(0)\mathbf{U}_2 \sin\left(\frac{N}{K}\mathbf{\Sigma}\right) \\ \Rightarrow \qquad & \mathbf{F}_\perp(0)\mathbf{U}_2 = \left[\overline{\mathbf{F}}(0)\cos\left(\frac{N}{K}\mathbf{\Sigma}\right) - \overline{\mathbf{F}}(N/K)\right] \sin^{-1}\left(\frac{N}{K}\mathbf{\Sigma}\right). \end{aligned} \tag{18}$$

Assuming there are $N/K - 1$ data subcarriers between two pilots, their precoder matrices can then be determined from

$$\overline{\mathbf{F}}\,(m) = \left[\overline{\mathbf{F}}(0)\cos\left(m\mathbf{\Sigma}\right) - \mathbf{F}_\perp(0)\mathbf{U}_2 \sin\left(m\mathbf{\Sigma}\right)\right], \quad 1 \le m \le N/K - 1. \tag{19}$$

and setting $\mathbf{F}(m) = \overline{\mathbf{F}}(m)$.

The geodesic approach will work well if the subspace angular variation of the optimal data subcarrier precoders with respect to the optimal pilot precoder varies approximately linearly with increasing subcarrier spacing. For the channel model given by (1) and (2), where each of the entries in $\mathbf{G}(l)$ are i.i.d. zero mean unit variance complex Gaussian random variables and reasonably spaced pilot precoders, numerical simulations show this to be true. For this channel model, we plot in Fig. 2 a single realization of the subspace angular variation for the right singular vector corresponding to the dominant singular value in a 4×4, 128 subcarrier MIMO-OFDM system for pilot spacing of 8 and 16 subcarriers.

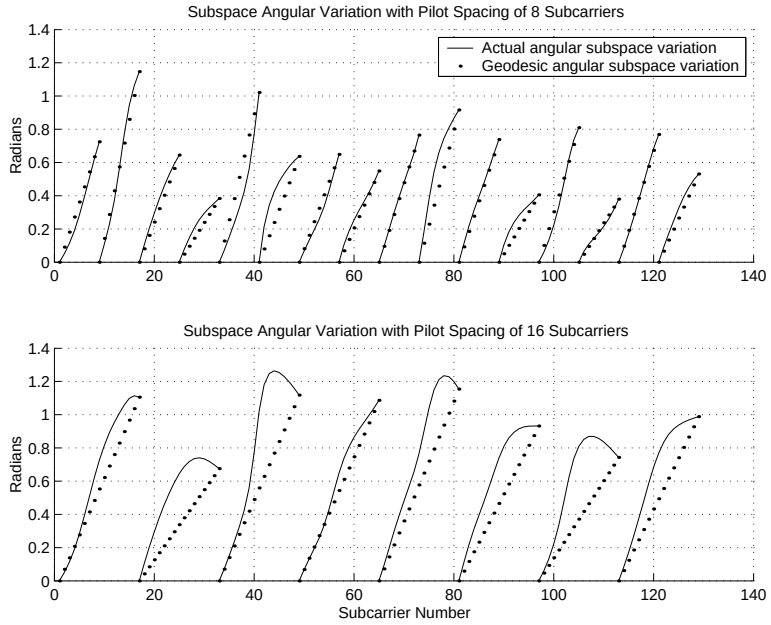

Figure 2: Subspace angular variation of the optimal data precoder matrices from the pilot precoder matrices.

The right and left singular vectors corresponding to the dominant singular value of the subcarrier channel matrix are the optimal beamforming and combining vectors in an MRT/MRC system. The discrete-time matrix channel impulse response has eight taps between each transmit and receive antenna pair with a uniform profile and i.i.d. complex Gaussian distribution as in (1). Notice that as the pilot spacing increases, the straight line

approximation of connecting the end points results in a less faithful representation of the actual subspace variation. Finally, we note that to determine the precoders for subcarriers $(K-1)N/K+1$ to $N-1$, we sample the geodesic connecting $\mathbf{F}((K-1)N/K)$ and $\mathbf{F}(0)$.

4.2 Karcher Mean and Generalized Median Clustering

In current clustering based approaches [6], [36], the precoder matrices for the center subcarriers in each cluster are reused for the other subcarriers in that cluster. In Fig. 2, the subspace angular variation of the optimal data subcarrier precoders with respect to the optimal pilot precoder (which is for the first subcarrier in a cluster) is shown. Recall from the previous section that distance is directly related to the principle angles between subspaces. Fig. 2 implies that if the spacing between two subcarriers increases so does the subspace distance between their respective optimal precoders. It has been shown in [20] that precoder performance is directly related to the distance between the optimal precoder and the assumed precoder. This increasing distance between the optimal precoder and assumed precoder as we move away from the center subcarrier explains the performance degradation exhibited in current clustering schemes.

An alternative approach would be for the receiver to find the mean of all the precoders in a cluster, determine the codebook entry closest to it, and send that information to the transmitter instead. However, the conventional method of finding the mean of matrices $\{\mathbf{F}(1), \ldots, \mathbf{F}(n)\} \in \mathcal{G}(M_{tx}, M)$ by averaging (i.e., $\frac{1}{n}\sum_{i=1}^{n}\mathbf{F}(i)$) does not work as the resulting solution is not a point on the Grassmann manifold. For points $\{y_1, \ldots, y_n\}$ in Euclidean space, an alternative way of viewing the mean is that it is the point which minimizes the

function

$$\frac{1}{2}\sum_{i=1}^{n}\|y-y_i\|_2^2\,. \tag{20}$$

Note that in the above, we are looking for the point y that minimizes the sum of the squares of distances from y to the sample points $\{y_1,\ldots,y_n\}$. In this case the distances are lengths of straight lines, which are the geodesics of Euclidean spaces. Extending this reasoning to manifolds, for precoders in the k^{th} cluster the Karcher mean $\mathbf{F}_k$ [37]–[40], can be defined as the point $\mathbf{F}$ which locally minimizes the objective function

$$\sum_{m=0}^{N/K-1} d^2(\mathbf{F},\mathbf{F}(kN/K+m)) \tag{21}$$

where $d(\cdot,\cdot)$ is a subspace distance. As a variety of distance metrics have been defined on the Grassmann manifold, different gradient based iterative optimization algorithms have been proposed in [41], [42], and [38] to numerically solve for the minimization of (21). These iterative techniques are computationally expensive and require testing for convergence. Since our goal is to find the codebook entry closest to $\mathbf{F}_k$ and given that our codebook size is finite, we can modify (21) to define our selection criterion as

$$\mathbf{F}_k = \underset{\mathbf{F}\in\mathcal{F}}{\operatorname{argmin}} \sum_{m=0}^{N/K-1} d^2(\mathbf{F},\mathbf{F}(kN/K+m)). \tag{22}$$

This can be accomplished by a simple brute force search. The distance metric that will be used in our simulations is the geodesic distance $d_g(\cdot,\cdot)$. In the same manner the generalized median is defined in [43] as the point $\mathbf{F}$ which minimizes the objective function

$$\sum_{m=0}^{N/K-1} d(\mathbf{F},\mathbf{F}(kN/K+m)). \tag{23}$$

Again, a brute force search like in (22) can be used to determine the representative precoder matrix from the codebook. Both these approaches can also be used for antenna subset selection as will be explained in the next section.

5 Antenna Subset Selection

Antenna subset selection is the simplest transmitter-precoding scheme for MIMO systems. It is a special case of limited feedback precoding where the columns of the precoder matrices are chosen from the columns of an $M_{tx} \times M_{tx}$ identity matrix. It has been shown to provide significant diversity and coding gains in space-time coded systems over an unprecoded system when both exact and statistical channel knowledge are used [10], [12], [14], [15], [44] and [45]. Consider the case of spatial multiplexing or OSTBC for M antennas with M_{tx} available transmit antennas. With antenna subset selection, the transmitter chooses the best M out of the M_{tx} possible transmit antennas depending on the precoding criterion. For single carrier MIMO systems this is easily performed by RF switches at the front-end. In this case the number of RF chains in a precoded MIMO system is the same as that for an unprecoded system. The only added hardware complexity at the transmitter is that of the extra $M_{tx} - M$ antennas and M RF switches to choose from the M_{tx} antennas.

In MIMO-OFDM systems, if the precoders are designed separately for all subcarriers, M_{tx} RF chains have to be implemented. For low-cost and widespread implementation, it is desirable to have a precoded MIMO-OFDM system that uses the same number of RF chains as an unprecoded system. This is possible only if one precoder matrix using an antenna subset selection codebook is chosen per MIMO-OFDM symbol. The RF switches which implement the precoder matrix multiplication choose the appropriate antennas as shown in Fig. 3. The Karcher mean and the generalized median approach present a natural way to design the precoder matrix.

For antenna subset selection, a natural way to represent the antenna subset is with a codebook $\mathcal{F} = \left\{\mathbf{E}_1, \ldots, \mathbf{E}_{\binom{M_{tx}}{M}}\right\}$ known at both transmitter and receiver where the

precoder matrices $\mathbf{E}_i$ are composed of M distinct columns from $\mathbf{I}_{M_{tx}}$. A total of $\left\lceil \log_2 \binom{M_{tx}}{M} \right\rceil$ bits of feedback are required per MIMO-OFDM symbol. The precoder selection criterion at the receiver for either the Karcher mean or the median can be written as

$$\mathbf{F} = \underset{\mathbf{E}\in\mathcal{F}}{\operatorname{argmin}} \sum_{m=0}^{N-1} d^{\nu}(\mathbf{E}, \mathbf{F}(m)), \qquad \nu = 1, 2 \tag{24}$$

where $\mathbf{F}(m)$ corresponds to the optimal precoder matrix for subcarrier m and $\mathbf{F}$ is the antenna subset to be used for the MIMO-OFDM symbol.

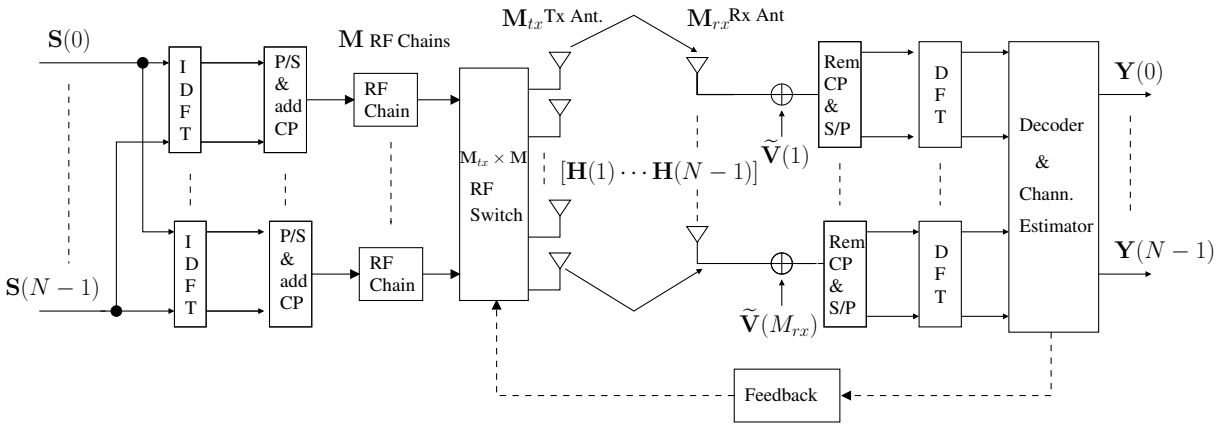

Figure 3: MIMO-OFDM system model for antenna selection on a per MIMO-OFDM symbol basis.

6 Complexity Analysis

In the geodesic approach an $O(K \cdot 2^B)$ search of the precoder codebook is required at the receiver to determine the precoders for the pilots in one MIMO-OFDM symbol. The transmitter determines the non-pilot precoders from (15), (18) and (19). Here, the computational cost at the transmitter is primarily governed by the eigen-decomposition in (15) which has to be done for consecutive pairs of pilot precoders resulting in an $O(K \cdot M^3)$

complexity. In contrast, the MS interpolator requires an $O(K \cdot 2^{B+B_1})$ search at the receiver where 2^{B_1} is the size of the codebook for the quantized phase θ or rotation matrix $\mathbf{Q}$ information. Furthermore, the complexity of computation of the non-pilot precoders at the transmitter is $O((N-K) \cdot M^3)$. To compute the Karcher mean or the median, first the optimal unquantized precoders for all the subcarriers are determined at the receiver at a computational cost $O(N \cdot M_{tx}^3)$. Then the principle subspace angles between the unquantized precoders and all the precoder codebook entries are computed. Using Chan's SVD algorithm [28] this has a computational cost of $O(N \cdot 2^B \cdot (M^3 + M_{tx}M^2))$. Finally, the receiver has to do an $O(K \cdot 2^B)$ search to determine the representative precoder for each of the K-clusters per MIMO-OFDM symbol.

7 Numerical Performance Analysis

In this section, we compare the performance of the proposed reduced feedback techniques with those existing in the literature. Simulations are performed for a MIMO-OFDM system using QPSK with 128 subcarriers ($N = 128$). The channel model (1) is used where we assume that the entries in $\mathbf{G}(l)$ are independent and identically distributed complex Gaussian random variables with zero mean unit variance and that the channel matrices $\mathbf{G}(l)$ are independent for different delays $0 \leq l < L$. Unless otherwise stated the channel impulse response is assumed to have eight taps ($L = 8$) between each pair of transmit and receive antennas with a uniform profile. We also assume that the feedback channel has no delay and no transmission errors, and that the receiver has perfect channel knowledge for all subcarriers $\mathbf{H}(n), n = 0, \ldots, N-1$. The codebooks used for the beamforming and spatial multiplexing simulations were generated using techniques in [46] and are given in [47].

Experiment 1: Fig. 4 shows the performance of using different cluster based approaches for a 4×4 system with beamforming at the transmitter and maximum ratio combining at the receiver.

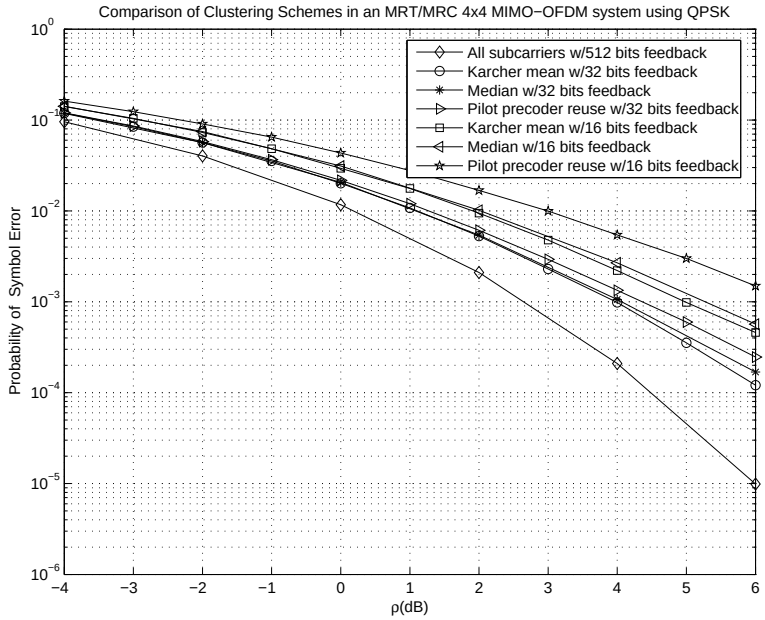

Figure 4: Probability of symbol error comparison for different clustering schemes in a 4×4 MIMO-OFDM system using QPSK.

A codebook of size $|\mathcal{F}|$=16 is assumed to be known to both transmitter and receiver. Performance curves with cluster sizes of 16 and 32 subcarriers corresponding to a feedback of 32 and 16 bits, respectively, per MIMO-OFDM symbol are shown for the Karcher mean and current pilot precoder reuse clustering methods. The benchmark for comparison is the ideal beamforming case where the indices from a codebook size $|\mathcal{F}| = 16$ are fed back for all beamforming vectors. This requires a total of 512 bits of feedback information. As can be seen, the degradation when increasing the cluster size is more prominent when using the pilot precoder reuse method. At a symbol error rate (SER) of 10^{-2}, with 32 subcarriers in a cluster, the Karcher mean approach outperforms the pilot precoder reuse method by more than 1 dB and is 2 dB away from the case where precoder matrices for all subcarriers are

fedback. The median approach performs slightly worse than the Karcher mean approach.

Experiment 2: In Fig. 5 we compare the performance of the different reduced feedback schemes in a 4×4 system with beamforming at the transmitter and maximum ratio combining at the receiver. For this MIMO-OFDM system, we have shown in Fig. 2 that the subspace angular variation of the ideal data subcarrier precoders with respect to the ideal pilot subcarrier precoders can be approximated as a straight line if the pilot subcarrier spacing was small. This was the justification for proposing a geodesic-sampling type interpolator. Fig. 5 shows that if the pilots are spaced eight subcarriers apart, the "ideal

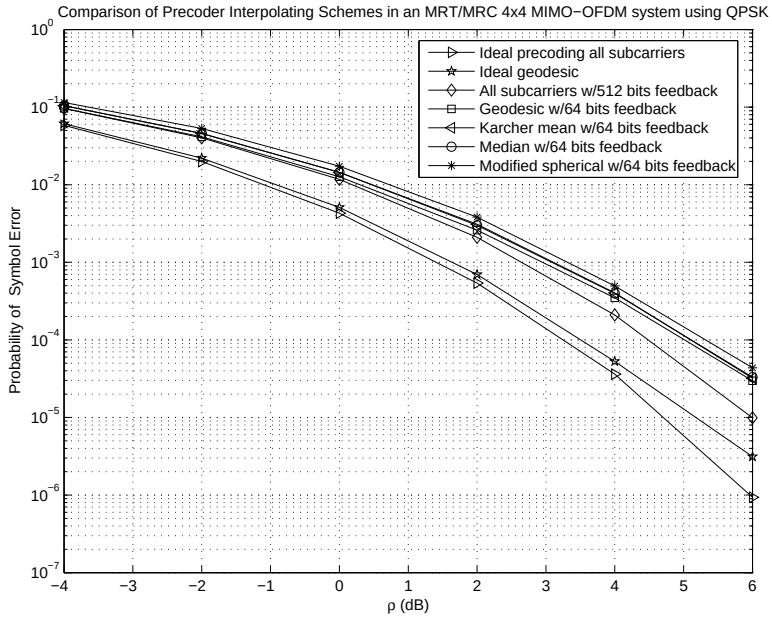

Figure 5: Probability of symbol error comparison for different precoder interpolation schemes in a 4×4 MIMO-OFDM system using QPSK.

geodesic" (i.e., the geodesic connecting the unquantized optimal pilot precoders) gives a performance close to that of ideal precoding where precoder knowledge is available for all subcarriers. To compare the performance in a limited feedback scenario for the MS interpolators, θ is uniformly quantized to 2 bits. A codebook of size $|\mathcal{F}| = 4$ is used for the beamforming vectors resulting in a total of $2K + 2K$ (K is the number of pilots) bits of

feedback information per MIMO-OFDM symbol. The geodesic interpolator uses a codebook of size $|\mathcal{F}| = 16$ for representing the pilot precoders which also requires $4K$ bits of feedback information. For interpolation with 16 pilots, at a symbol error rate (SER) of 10^{-4} the geodesic interpolator outperforms the MS interpolator by 0.3 dB and exhibits only 0.5 dB degradation from the case where precoder information is sent back for all subcarriers using a codebook of size $|\mathcal{F}| = 16$. The performance of the Karcher mean approach falls between that of the geodesic and MS interpolator.

Experiment 3: In Fig. 6 we investigate the performance of the different reduced feedback schemes when imperfect channel knowledge is available at the receiver of the above 4×4 beamforming system. The receiver estimates the channel for the n^{th} subcarrier as

$$\mathbf{H}_{est}(n) = \beta \mathbf{H}(n) + \sqrt{1-\beta^2}\mathbf{H}_{error}(n) \tag{25}$$

where $0 \leq \beta \leq 1$ and the entries of $\mathbf{H}_{error}(n)$ are modeled as i.i.d Gaussian random variables with zero mean unit variance. The beamforming and maximum ratio combining vectors are designed from the estimated channel $\mathbf{H}_{est}(n)$. The key observation from Fig. 6 is that for increasing degradation in channel estimates the Karcher mean and generalized median approach outperform both MS and geodesic interpolators.

Experiment 4: In this experiment, the performance of the interpolators are compared in a two substream precoded spatial multiplexing 4×2 system. The vector symbol error rate (VSER) is plotted in Fig. 7 to characterize the performance when using zero-forcing (ZF) receivers.

The precoding matrices are found using the minimum singular value criterion [22]. For the MS interpolator, as in [7], the unitary derotation matrix $\mathbf{Q}$ is chosen from a codebook size $|\mathcal{Q}| = 4$. With the pilot precoding matrices chosen from a codebook size $|\mathcal{F}| = 4$, the

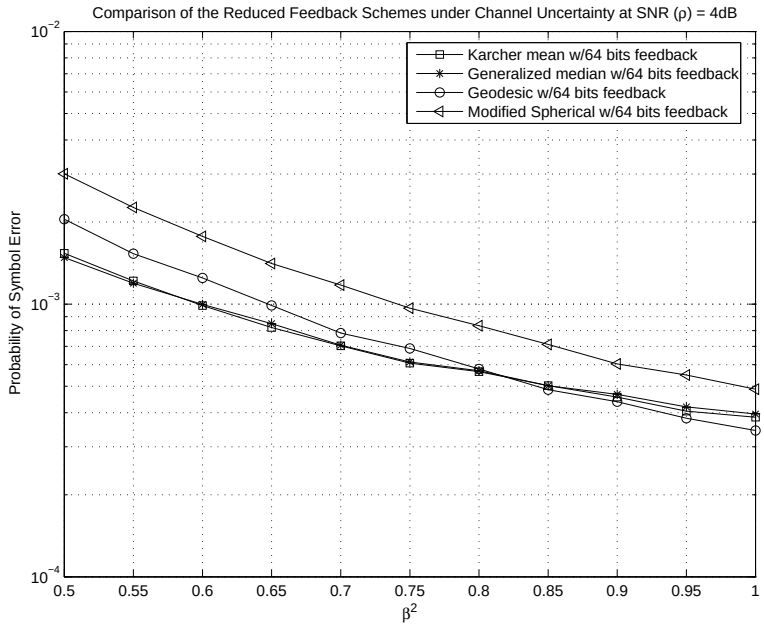

Figure 6: Probability of symbol error comparison as a function of channel estimation error for different reduced feedback schemes in a 4×4 MIMO-OFDM system using QPSK at SNR (ρ) = 4 dB.

MS scheme requires $2K + 2K$ bits of feedback. The corresponding geodesic interpolator with $|\mathcal{F}| = 16$ also uses $4K$ bits of feedback. The benchmark for comparison is when the indices of the optimal precoder matrices for all subcarriers are sent back from a codebook size of $|\mathcal{F}| = 16$ corresponding to 512 bits of feedback. As can be seen in Fig. 7, when interpolating with 16 pilots the geodesic sampling approach, the Karcher mean approach and the median approach perform better than the MS interpolator.

Experiment 5: In this experiment, we show the gains available when using antenna subset selection, the simplest precoding scheme. Simulations are done for a three substream precoded spatial multiplexing 6×3 system. In this case, the three columns of a precoder matrix $\mathbf{F}$ are chosen from the columns of a 6×6 identity matrix. The discrete-time channel impulse response between each pair of transmit and receive antennas is assumed to have four taps with a uniform profile and i.i.d. complex Gaussian distribution as in (1). The

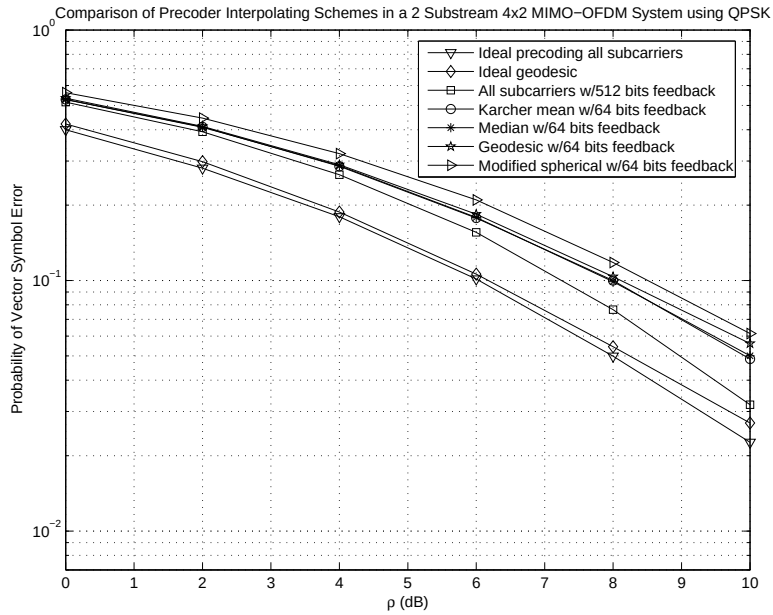

Figure 7: Probability of vector symbol error comparison for different precoder interpolation schemes in a two substream 4×2 MIMO-OFDM system using QPSK with a zero-forcing receiver.

VSER is plotted for ZF receivers in Fig. 8.

When antenna subset selection is performed for each subcarrier, $128 \cdot \left\lceil \log_2 \binom{6}{3} \right\rceil = 640$ bits of feedback are required. Although there is a significant improvement over an unprecoded 3×3 spatial multiplexing system, hardware complexity for implementation at the transmitter is prohibitively large as the antenna subset selection has to be done separately for every subcarrier. An alternative is to perform antenna subset selection on a per MIMO-OFDM symbol basis resulting in only a feedback of $\left\lceil \log_2 \binom{6}{3} \right\rceil = 5$ bits (i.e., only one precoder is chosen for all the subcarriers). Fig. 8 shows that if the precoder is chosen via the Karcher mean approach, at a VSER of 10^{-1} there is a performance improvement of 3 dB over the unprecoded case and a 5 dB degradation relative to precoding for all subcarriers. Using the median has a similar performance compared to the Karcher mean.

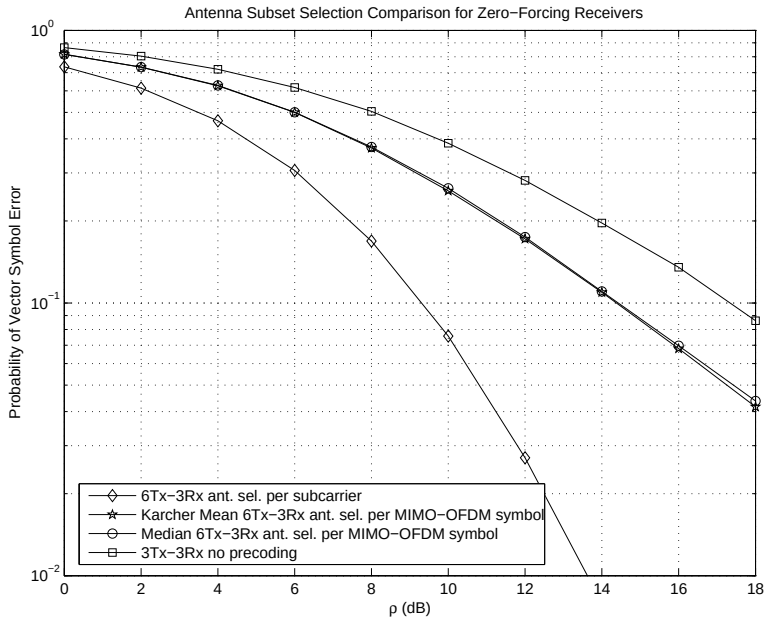

Figure 8: Probability of vector symbol error comparison for antenna subset selection in a three substream 6×3 MIMO-OFDM system using QPSK with a zero-forcing receiver.

8 Conclusions

In this paper two novel approaches for reducing the feedback requirement in a precoded MIMO-OFDM system were presented. In the first approach the receiver conveyed only pilot precoder matrix information to the transmitter. The transmitter then reconstructed the precoders for the data subcarriers by uniformly sampling the geodesic connecting the pilot precoder matrices. This approach was shown to work well for channel models where the subspace angular variation of the optimal precoder matrices for increasing subcarrier spacing was approximately linear. In the second method, the receiver conveyed to the transmitter either the Karcher mean or the generalized median of the precoders for a group of subcarriers. No further computation was necessary at the transmitter. Both methods were shown to perform better than existing reduced feedback precoding techniques. Finally, for the antenna subset selection scenario, we showed that the Karcher mean and the

generalized median approach can provide significant gains over an unprecoded system with very minimal feedback and minimal transmitter implementation complexity.

One direction for future work is to characterize the performance loss of the above two schemes as a function of the pilot spacing/cluster size and codebook size. In order to do this for the Rayleigh channel model given in (1), a thorough probabilistic analysis of how the subspace angles for the optimal precoders vary with increasing subcarrier spacing needs to be done.

References

[1] 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.

[2] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in mulitiple-antenna data transmission over fading channels," *IEEE Jour. Select. Areas in Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.

[3] 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.

[4] 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.

[5] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc. : Packings in Grassmannian spaces," *Experiment. Math.*, vol. 5, no. 2, pp. 139–159, 1996.

[6] J. Choi and R. W. Heath, Jr., "Interpolation based transmit beamforming for MIMO-OFDM with limited feedback," in *Proc. IEEE Int. Conf. on Commun.*, vol. 1, pp. 249–253, June 2004.

[7] J. Choi and R. W. Heath, Jr., "Interpolation based unitary precoding for spatial multiplexing MIMO-OFDM with limited feedback," in *Proc. IEEE Glob. Telecom. Conf.*, vol. 1, pp. 214–218, Nov.-Dec. 2004.

[8] J. Choi and R. W. Heath, Jr., "Interpolation based transmit beamforming for MIMO-OFDM with limited feedback," *IEEE Trans. Sig. Proc.*, vol. 53, pp. 4125–4135, Nov. 2005.

[9] T. Pande, D. J. Love, and J. V. Krogmeier, "A weighted least squares approach to precoding with pilots for MIMO-OFDM," *IEEE Trans. Sig. Proc.* To appear.

[10] 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.

[11] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microwave Mag.*, pp. 46–56, Mar. 2004.

[12] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, pp. 322–324, Aug. 2002.

[13] R. Nabar, D. Gore, and A. Paulraj, "Optimal selection and use of transmit antennas in wireless systems," in *Proc. ICT*, May 2000.

[14] D. A. Gore, R. W. Heath, Jr., and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, pp. 491–493, Nov. 2002.

[15] D. A. Gore and A. J. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 2580–2588, Oct. 2002.

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

[17] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, pp. 41–59, Autumn 1996.

[18] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Jour. Select. Areas in Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[19] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Info. Th.*, vol. 45, pp. 1456–1467, July 1999.

[20] 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.

[21] 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.

[22] D. J. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Info. Th.*, vol. 51, pp. 2967–2976, Aug. 2005.

[23] *Dictionary of Mathematics*. New York: McGraw-Hill, 1997.

[24] A. Edelman, T. Arias, and S. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM Journal of Matrix Analysis and Applications.*, vol. 20, pp. 303–353, 1998.

[25] S. Zhou, Z. Wang, and G. B. Giannakis, "Quantifying the power loss when transmit beamforming relies on finite rate feedback," *IEEE Trans. Wireless Comm.*, vol. 4, pp. 1948–1957, July 2005.

[26] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO-OFDM systems in mobile wireless channels," *IEEE Trans. Sig. Proc.*, vol. 51, pp. 1615–1624, June 2003.

[27] K. Gallivan, A. Srivastava, X. Liu, and P. Dooren, "Efficient algorithms for inferences on Grassmann Manifolds," in *Proc. IEEE Workshop on Statistical Signal Processing*, pp. 315–318, Sept. 2003.

[28] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore: Johns Hopkins University Press, third ed., 1996.

[29] 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.

[30] S. Sun, I. Wiemer, C. K. Ho, and T. T. Tjhung, "Training sequence assisted channel estimation for MIMO-OFDM," in *Proc. IEEE Wireless Comm. and Net. Conf.*, pp. 16–20, March 2003.

[31] W. Bai, C. He, L. Jiang, and H. Zhu, "Blind channel estimation in MIMO-OFDM systems," in *Proc. IEEE Glob. Telecom. Conf.*, vol. 1, pp. 317–321, Nov. 2002.

[32] C. Shin and E. Powers, "Blind channel estimation for MIMO-OFDM systems using virtual carriers," in *Proc. IEEE Glob. Telecom. Conf.*, vol. 4, pp. 2465–2469, Dec. 2004.

[33] H. Minn and N. Al-Dahir, "Optimal training signals for MIMO-OFDM channel estimation," in *Proc. IEEE Glob. Telecom. Conf.*, vol. 1, pp. 219–224, Dec. 2004.

[34] H. Bölcskei, R. W. Heath Jr., and A. Paulraj, "Blind channel identification and equalization in OFDM-based multi-antenna systems," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 96–109, Jan. 2002.

[35] T. Y. Al-Naffouri, O. Awoniyi, O. Oteri, and A. Paulraj, "Receiver design for MIMO-OFDM transmission over time variant channels," in *Proc. IEEE Glob. Telecom. Conf.*, pp. 2487–2492, 2004.

[36] R. Grunheid, E. Bolinth, and H. Rohling, "A blockwise loading algorithm for the adaptive modulation technique in OFDM systems," in *Proc. IEEE Veh. Technol. Conf.*, vol. 2, pp. 948–951, Oct. 2001.

[37] H. Karcher, "Riemannian center of mass and mollifier smoothing," *Comm. Pure Appl. Math.*, vol. 30, no. 5, pp. 509–541, 1977.

[38] E. Klassen, A. Srivastava, W. Mio, and S. H. Joshi, "Analysis of planar shapes using geodesic paths on shape spaces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, pp. 372–383, March 2004.

[39] W. Mio, A. Srivastava, and X. Liu, "Tutorial on nonlinear manifolds in pattern recognition and image analysis." available at *http://www.cavis.fsu.edu/manifold-tutorials/*.

[40] R. P. Woods, "Characterizing volume and surface deformations in an atlas framework: theory, applications and implementation," *NeuroImage*, vol. 18, pp. 769–788, 2003.

[41] P.-A. Absil, R. Mahony, and R. Sepulchrez, "Riemannian geometry of Grassmann manifolds with a view on algorithmic computation," *Acta Applicandae Mathematicae*, vol. 80, pp. 199–220, Jan. 2004.

[42] J. Manton, "A globally convergent numerical algorithm for computing the centre of mass on compact lie groups," in *Int. Conf. on Control, Automation, Robotics, and Vision*, 2004.

[43] D. Nowicki and O. Dekhtyarenko, "Averaging on Riemannian manifolds and unsupervised learning using neural associative memory," in *European Symp. on Artificial Neural Networks*, pp. 181–186, Apr. 2005.

[44] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Comm. Mag.*, vol. 42, pp. 74–80, Oct. 2004.

[45] 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.

[46] 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.

[47] D. J. Love, "Tables of complex Grassmannian packings." available at *http : //www.dynamo.ecn.purdue.edu/˜djlove*, 2006.