The Report Committee for Patrick Karl Sisterhen

Certifies that this is the approved version of the following report:

Quantized Successor Pre-coding: A Method for Spatial Multiplexing in MIMO Systems with Limited Feedback and Temporally-Correlated Channels

APPROVED BY

SUPERVISING COMMITTEE:

Supervisor: ________________________

R. W. Heath, Jr.

____

J.G. Andrews

Quantized Successor Pre-coding: A Method for Spatial Multiplexing in MIMO Systems with Limited Feedback and Temporally-Correlated Channels

by

Patrick Karl Sisterhen, B.S, B.E.

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

December 2010

Abstract

Quantized Successor Pre-coding: A Method for Spatial Multiplexing in MIMO Systems with Limited Feedback and Temporally-Correlated Channels

Patrick Karl Sisterhen, MSE

The University of Texas at Austin, 2010

Supervisor: R.W. Heath, Jr.

The use of feedback to provide channel state information to the transmitter can greatly improve the performance of a communication system. However, the amount of information required to characterize a time-varying MIMO channel can exceed the capacity of the feedback channel. This paper surveys research in limited feedback systems, which employ a number of methods to reduce the information and improve performance in multi-antenna communication systems. This paper also presents a new method, Quantized Successor Pre-coding (QSP), that exploits time-correlation to implement spatial multiplexing in a MIMO system using very little feedback. QSP uses an ordered codebook of pre-coders and transmission modes to reduce the feedback to a single bit. Simulations of QSP demonstrate a substantial performance improvement relative to open-loop spatial multiplexing.

Contents

List of Algorithms

List of Figures

1 Introduction

The use of feedback to provide channel state information to the transmitter (CSIT) improves the performance of a communication system across many metrics. With sufficient information about the channel, the transmitter can improve the instantaneous bit error rate, reduce the likelihood of outage, increase the single user data rate, and increase the throughput in a multi-user system. Adaptive modulation can optimally allocate power, adjust the rate, and select the coding scheme across both time and frequency. These techniques can be combined with scheduling, channel allocation, and network coding in multi-user systems. The expansion of interest in multiple antenna systems for both academic research and commercial applications highlighted further benefits (and challenges) of feedback to the transmitter. Both multi-input/single-output (MISO) and multi-input/multi-output (MIMO) systems can realize gains in both diversity and rate by using CSIT.

In modern communication systems, the receiver is normally assumed to possess near-perfect channel state information, deduced via training and pilot sequences in the transmitted packets. The transmitter cannot so easily determine this information. In some cases, reciprocity can be assumed to provide at least statistical information about the channel without explicit knowledge at the transmitter [1]. If CSIT is desired but reciprocity cannot be assumed, the receiver can provide a characterization of the channel through an explicit feedback mechanism. Such mechanisms include frequency-division duplexing (FDD) and time-division duplexing (TDD). The amount of channel state information becomes significant in multi-antenna systems. This information must then be updated at sufficient rate to continuously characterize a time-varying channel. A large amount of CSI may exceed the capacity of the feedback channel or take up channel resources (time, bandwidth) that could otherwise be allocated to the forward link. Reducing the amount of feedback is an important objective for optimizing modern communication systems, and has been the focus of a great deal of research [1].

Some research has focused on optimally quantizing the channel into a discrete set of states using vector quantization techniques [1]. Analysis of the capacity of a MIMO system leads to an alternative approach, which quantizes the covariance of the signals at the transmit antennas [1, 2, 3]. Another line of research has restated the problem as one of selecting the optimal beamforming vector or pre-coding matrix that the transmitter will use [4, 5, 6, 7, 8, 9]. From this perspective, the receiver selects the best beamforming or pre-coding scheme from its near-perfect knowledge of the channel and communicates that to the transmitter. The problem here is twofold: designing an optimal codebook (e.g.

a quantization of the space of beamforming vectors) and selecting the optimal method based on the current channel (e.g. selecting the best quantized beamforming vector). Recent work has focused on taking advantage of time-correlation in channels to adaptively optimize the codebook and mitigate delay in the feedback channel [10, 11, 12, 13].

This report presents a brief review of literature on limited feedback topics and also describes a new method that exploits time-correlation to implement spatial multiplexing in a MIMO system using 1-bit feedback. The fundamental idea is to create a codebook of quantized pre-coder matrices and also a successor list for each pre-coder. The successor list represents an ordered set of candidate pre-coders that the receiver will successively evaluate against the current pre-coder. Since the order of the successor list is known to both the transmitter and receiver, the transmitter knows which candidate pre-coder the receiver is currently evaluating. Consequently, a single bit of feedback suffices to indicate which pre-coder the transmitter should use for the next transmission. This paper also develops a multi-mode variation, allowing the receiver to indicate not only the correct pre-coder but also a more optimal power allocation with a single bit of feedback. Simulations show a substantial performance improvement relative to open-loop, equal power spatial multiplexing.

Section II presents a survey of research on limited feedback in MIMO systems. Section III details three different variations of the Quantized Successor Pre-coding (QSP) method. Section IV presents simulation results and discussion, and Section V contains conclusions and ideas for future work.

This report employs the following notation. The set of complex numbers is denoted as $\mathbb{C}$. The conjugate transpose of a matrix is denoted by a superscript *. The trace of a matrix is denoted $tr(.)$. The two-norm of a vector or matrix is denoted $\|.\|_2$ and the Frobenius-norm of a matrix is denoted $\|.\|_F$. The base-two logarithm is denoted $log_2(.)$. The determinant of a matrix is denoted $det(.)$. Vectors are indicated with bold, lowercase letters while matrices are indicated with bold, uppercase letters.

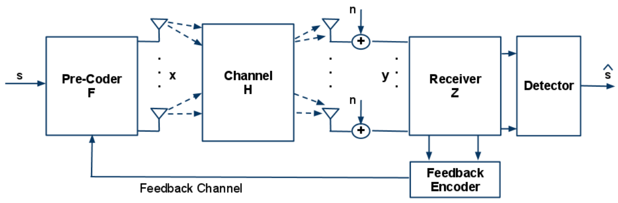

Figure 1: Generalized MIMO System Model

2 Limited Feedback Overview

In the general model depicted in Figure 1,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$$

where $\mathbf{y}$ is the N_R-dimensional receive vector, $\mathbf{x}$ is the M_t-dimensional transmit vector, $\mathbf{n}$ is the N_r-dimensional noise vector, and $\mathbf{H}$ is the $N_r \times M_t$ channel matrix. The channel is assumed to exhibit spatially uncorrelated Rayleigh fading. Except where otherwise noted, the channel is also assumed to be temporally uncorrelated. The noise is assumed to be i.i.d. complex Gaussian $\sim CN(0,1)$. The transmit signal is power constrained such that $\mathbb{E}_{\mathbf{H},\mathbf{x}}[||\mathbf{x}||^2] = \rho$.

At the transmitter, the transmit vector is formed by

$$\mathbf{x} = \sqrt{\rho}\mathbf{F}\mathbf{s}$$

where $\mathbf{s}$ is the input codeword (coded independently of the channel information) and $\mathbf{F}$ is the beamforming vector or pre-coder. For simplification of design and analysis, power is constrained such that $\mathbb{E}_{\mathbf{s}}[\mathbf{s}\mathbf{s}^*] = \mathbf{I}$. At the receiver, the channel is estimated ($\hat{\mathbf{H}}$) and $\hat{\mathbf{s}}$ detected after the receive vector is operated on by an equalizer/combiner/decoder such that

$$\tilde{\mathbf{y}} = \mathbf{Z}\mathbf{y}$$

$$\hat{\mathbf{s}} = detector_{ML}(\tilde{\mathbf{y}}, \hat{\mathbf{H}}).$$

Providing the channel information to the transmitter produces a number of benefits, but feedback capacity may be limited. For a M_t (Tx antenna) x N_r (Rx antenna) flat-fading system, the channel information consists of M_tN_r complex values. In a frequency-selective channel with L taps, this increases to M_rN_rL values. Multicarrier techniques such as OFDM conveniently decompose the channel into N frequency-flat subcarriers, but the resulting channel information is then M_tN_rN. The channel matrix $\mathbf{H}$ may be quantized using vector quantization techniques [1]. The following sections survey several alternative techniques that have been developed to reduce the required feedback.

2.1 Covariance Quantization

The ergodic capacity of this system is [14, 15]

$$C = \mathbb{E}_{\mathbf{H}}[\max_{\mathbf{Q}:tr(\mathbf{Q})\leq 1, \mathbf{Q}^*=\mathbf{Q}} \log_2 \det(\mathbf{I} + \rho\mathbf{H}\mathbf{Q}\mathbf{H}^*)]$$

where $\mathbf{Q}$ is the covariance matrix of the transmitted signal. From this perspective, the pre-coding block at the transmitter produces

$$\mathbf{x} = \sqrt{\rho}\mathbf{Q}^{\mathbf{1/2}}\mathbf{s}.$$

From the training information in the received packet, the receiver can deduce the current channel realization $\mathbf{H}$. The receiver can then solve an optimization problem to determine the capacity achieving $\mathbf{Q}$ for the current channel. This solution can be provided to the transmitter via the feedback path. However, high-resolution representation of any possible realization of $\mathbf{Q}$ may require more capacity than is available in the feedback path. To reduce the amount of feedback, the possible realizations of $\mathbf{Q}$ can be quantized to a discrete codebook Q [1, 2, 3]. With B bits of feedback, the receiver can determine the index n of the codebook that maximizes

$$\log_2 \det(\mathbf{I} + \rho\mathbf{H}\mathbf{Q_n}\mathbf{H}^*), n = 1, 2, ..., 2^B$$

for the current channel realization $\mathbf{H}$ and send this index to the transmitter. The problem is then one of determining optimal quantization strategies.

One approach to designing rate maximizing codebooks uses a form of the Lloyd algorithm for minimizing the average distortion of the quantization [2]. An outline of the algorithm is

0) Start with a random partition set of the channel matrix set

1) Determine the optimal transmission covariance matrix for each partition. The optimal choice is the one that minimizes the expected distortion across all

channels in the partition. The distortion metric is capacity distortion, and depends only on the channel and the transmission covariance matrix. This optimal transmission strategy is the partition centroid.

2) Re-partition the channel matrix set by allocating channels to partition centroids using the nearest neighbor rule, where the distance metric is the same distortion metric used in step 1.

Steps 1 and 2 are repeated to convergence. The authors of [2] note that this algorithm is not guaranteed to converge to the global optimum. Accordingly, the entire procedure is repeated across many randomly generated initial partitions and the codebook with the highest capacity is chosen.

Other research investigated generating random codebooks on the complex unit sphere [3]. The authors' results demonstrated that capacity loss decreases as an exponential function of the number of feedback bits. Specifically,

$$O(2^{-B/(2MM_t-2)})$$

where B is the number of feedback bits, M_t is the number of transmit antennas, N_r is the number of receive antennas, and $M = \min\{M_t, N_r\}$.

2.2 Beamformer Quantization

In a beamforming system, the covariance matrix $\mathbf{Q}$ is a rank 1 matrix, and pre-coding is merely an M_t-dimensional, unit-norm complex vector. A system of beamforming and combining vectors can achieve a diversity order of $M_t N_r$ [9]. Work along these lines treats the codebook design as one of selecting a set of optimal beamforming vectors $\{\mathbf{f}\}$ and combining vectors $\{\mathbf{z}\}$. One approach constrains the transmit antennas to identical gains but varies the phase. Specifically,

$$\mathbf{f} = (1/\sqrt{M_t})e^{j\theta}, \theta_k \in [0, 2\pi].$$

The work in [9] considered such a system, called equal gain transmission (EGT), coupled with several receiver designs: selection diversity combining (SDC), maximum ration combing (MRC), and equal gain combining (EGC). In EGC, the receive antenna weighting is constant $(1/\sqrt{N_r})$ and only the phase varies. The authors show that all of these transmitter/receiver combinations can achieve full diversity $(M_t N_r)$ in a memoryless, i.i.d. Rayleigh fading channel. They further develop an EGT beamformer quantization method that achieves the full diversity order, provided that

$$B \geq \log_2(M_t)/(M_t - 1).$$

Mukkavilli et al. produced a more general analysis of the beamforming problem for MISO systems [4]. This work used outage probability (the probability of the event that the channel realization cannot support the desired rate) as the performance metric. The authors first showed that outage probability is minimized by choosing a beamforming vector that maximizes $|\langle \mathbf{h}, \mathbf{f} \rangle|$. They further established a lower bound on outage probability that is asymptotically tight in the number beamforming vectors in the quantized codebook. An interesting result from this paper is that 8 bits of feedback performs very close to perfect (non-quantized) feedback. Using a distortion measure $d(B)$, defined as the loss of outage performance of a quantized codebook beamformer relative to perfect, non-quantized beamforming, the authors showed

$$d(B) \approx (M_t - 1) 2^{-\left(\frac{B}{M_t - 1}\right)}.$$

The authors of [4] also investigated construction of an optimal beamformer codebook. The design criterion for a good beamformer codebook was shown to be one that maximizes the angular distance between beamforming vectors. Specifically,

$$\min_{\{\mathbf{f}\} \in \mathbb{C}^{M_t}} \max_{i,j: 1 \leq i \neq j \leq 2^B} |\langle \mathbf{f}_i, \mathbf{f}_j \rangle| \, .$$

Interestingly, the design of beamformers in this way is equivalent to the design of unitary space-time constellations. Additionally, the design criterion can be restated as maximizing the minimum chordal distance between subspaces formed by the beamforming vectors [4].

The relation of codebook design to subspace packing is a concept that recurs in almost all of the literature on this topic. Due to the phase invariance of beamforming vectors, standard Euclidean vector quantization produces ambiguous results, so quantization on manifolds is a better approach [5, 7]. The formal mathematical backdrop is the complex Grassmann manifold $G(N, p)$, which is the set of p-dimensional subspaces in an N-dimensional complex Euclidean space. Techniques for optimal packing and traversals on this manifold have been applied to limited-feedback communication system design in a great many papers over the last decade [1].

Love et al. independently investigated the relationship of Grassmannian Manifold the-

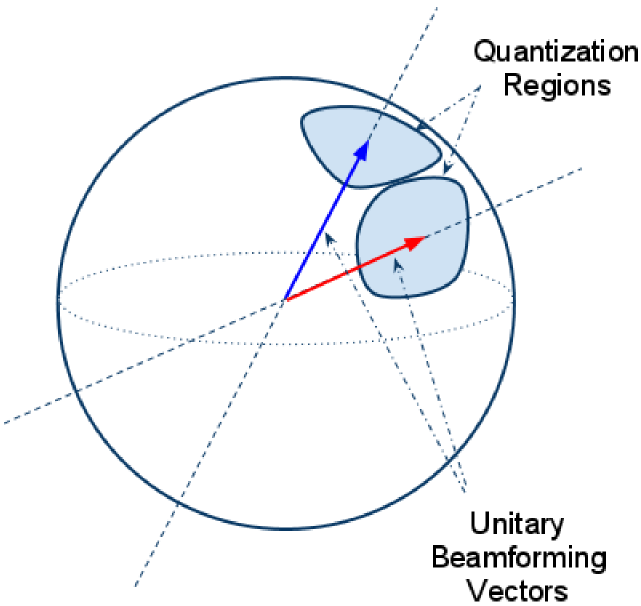

Figure 2: A visual representation of beamforming vector quantization on the G(3,1) Grassmann manifold

ory to beamforming in MIMO systems, and introduced many new and fundamental ideas [5]. An important revelation in this work is that the problem of finding quantized beamformers for MIMO is identical to that of MISO; there is no dependence on the number of receive antennas. The authors presented the Grassmann beamforming criterion: the codebook of beamforming vectors should be designed so that the minimum distance on the $G(M_t, 1)$ manifold is maximized. Formally, design the codebook of beamforming vectors $\mathcal{F} = \{\mathbf{f}_i\}_{i=1}^{2^B}$ to maximize

$$\delta(\mathcal{F}) = \min_{1 \leq k < l \leq 2^B} \sqrt{1- \mid \mathbf{f}_k^* \mathbf{f}_l \mid^2}.$$

The authors consider several transmission methods: quantized maximum ratio transmission (QMRT), quantized equal gain transmission (QEGT), and generalized subset selection (GSS). They demonstrate that the Grassmann beamforming criterion generates codebooks with full diversity order for all of these methods, provided that the size of the codebook ($N = 2^B$) is greater than or equal to M_t. They also provide analytic bounds that allow the codebook size (N) to be chosen to achieve the desired capacity loss or SNR loss relative to an unquantized system [5].

2.3 Spatial Multiplexing Pre-coder Quantization

Quantizing on the Grassmann manifold proved to be an extensible framework for expanding the problem from beamforming to spatial multiplexing. Further work by Love and Heath approached the spatial multiplexing pre-coder codebook $\mathcal{F}$ as a subset of (M_t x M) unitary matrices, where $M \leq M_t$ is the number of data streams [16]. The authors defined several criteria for selecting the optimal pre-coder for different receiver designs: Maximum Likelihood Selection Criterion (ML-SC), Minimum Singular Value Selection Criterion (MSV-SC), Mean Squared Error Selection Criterion (MSE-SC), and Capacity Selection Criterion (Capacity-SC). For the MSV-SC, the optimal (unquantized) pre-coder $\mathbf{F}_{opt}$ is one that maximizes the minimum singular value of $\mathbf{HF}$. $\mathbf{F}_{opt}$ was shown in this paper to be the first M columns of the right singular vector matrix of $\mathbf{H}$. The same construction of $\mathbf{F}_{opt}$ was also shown to be optimal for MSE-SC and Capacity-SC. For a memoryless, Rayleigh fading channel, $\mathbf{F}_{opt}$ is isotropically distributed on $\mathcal{U}(M_t, M)$, the set of (M_t x M) unitary matrices. The authors derived the following quantized codebook design criteria:

1) For ML-SC, MSV-SC, and MSE-SC (with a trace cost function): $\mathcal{F}$ should be designed by maximizing the minimum projection two-norm distance between any pair of codeword matrix column spaces. The projection two-norm distance is defined as:

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

2) For MSE-SC (with determinant cost function) and Capacity-SC: $\mathcal{F}$ should be designed by maximizing the minimum Fubini-Study distance between any pair of codeword matrix column spaces. The Fubini-Study distance is defined as:

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

Both criteria relate to subspace packing on the Grassmann manifold, albeit with separate distance metrics [16]. Further work by Mondal, Dutta, and Heath established lower bounds on the distortion-rate function for Grassmann quantifiers [6].

The systems discussed so far assume flat-fading channels. In a frequency selective channel, OFDM is often used to decompose the channel into flat-fading subcarriers. Khaled et al. developed a system for efficient quantized feedback in MIMO-OFDM spatial multiplexing systems [13]. Their method first generates a codebook of optimal unitary pre-coders for all subcarriers. The receiver determines the optimal pre-coder and the mode (i.e. how many data streams can be supported in the current channel) for each subcarrier and feeds this back to the transmitter. To reduce the amount of feedback, information for only a pilot subset of the subcarriers is fed back. An interpolator at the transmitter constructs the pre-coders for all subcarriers based on the information given for the pilot subcarriers. Further reduction of feedback is accomplished by clustering the mode selection for subcarriers. Because the techniques of Grassmannian subspace packing cannot be used for full-rank unitary matrices, a different method is proposed in this paper to generate the codebook of unitary pre-coders. Their method is an adaptation of the Lloyd algorithm, and is outlined as follows:

0) Start with a randomly generated codebook ($\mathcal{F}$) of M_t x M_t unitary pre-coding matrices.

1) Generate a set of independent, isotropic M_t x M_t unitary training matrices.

2) Assign each training matrix to the region associated with one of the pre-coders in $\mathcal{F}$ according to the nearest neighbor rule, using as the distance metric the squared Frobenius-norm of the difference of the pre-coder and training matrix.

3) For each region constructed in Step 2, find the optimal pre-coder, which is the centroid of the region. This optimal pre-coder is the orthogonal projection of the arithmetic mean of training matrices in the region onto the space of unitary M_t x M_t matrices. These pre-coders form the new set $\mathcal{F}$.

Steps 1-3 can be repeated to convergence. The authors note that since the Lloyd algorithm is not guaranteed to converge to the global optimum, the procedure should be repeated across many initial random codebooks [13].

2.4 Designs for Temporally Correlated Channels

Improvements to codebook design and reductions in feedback rate can be achieved assuming temporally correlated channels. When the channel is not Rayleigh uncorrelated, the quantization strategies discussed so far produce a degraded SNR [7]. Additionally the feedback can be reduced if the transmitter has some idea about future channels based on the current channel state. Mondal and Heath developed a switched-codebook algorithm that adapts to the channel distribution. In their paper, they considered a beamforming system in number of channel models, including the standard uncorrelated Rayleigh, correlated Rician, auto-regressive time correlated, and spatially correlated. Their approach creates a set of codebooks, each with a set of quantized beamforming vectors, and adaptively switches both the codebook and the beamforming vector within the codebook as the channel changes. Their simulations demonstrate that their algorithms perform substantially better in correlated channels than codebooks designed for uncorrelated channels, and no worse when the channel is uncorrelated [7].

A paper by Kim et al. extended the quantization concept to a temporally correlated MIMO channel, but approached the problem in a novel way. Their idea was to reduce feedback by quantizing only the pre-coder variation required to adapt to the channel variation [12]. The authors represented this variation as a set of unitary rotation matrices that would transform the current unitary pre-coding matrix to another unitary pre-coding matrix. The pre-coder for any timeτ is constructed as

$$\mathbf{F}_\tau = \Theta_i \mathbf{F}_{\tau-1}$$

where Θ_i is a unitary rotation matrix chosen from the codebook of rotation matrices $\mathcal{C}_\Theta$. The receiver chooses the optimal rotation matrix for the current channel and feeds back its index. The rotation matrix codebook is designed offline for a specific time correlation coefficient of the channel. Their simulation demonstrated throughput gains relative to Grassmannian quantization with less feedback [12].

Research by Huang, Heath, and Andrews modeled CSI as a first-order, finite-state Markov chain [10]. This work established the required bit-rate in the feedback channel and analyzed the effect of delay in this channel, showing significant throughput gain decay with feedback lag. Their paper also presented a feedback compression method using temporal correlation in the channel [10].

Papers such as [10] modeled temporal correlation in channels and demonstrated the significant impact of feedback delay. The following subsection surveys research that has exploited temporal correlation to improve performance and reduce feedback, including predictive coding methods that mitigate the impact of feedback lag.

2.4.1 Predictive Coding in Temporally Correlated Channels

While the preceding methods utilized quantized vectors on the Grassmann manifold, Banister and Zeidler proposed a different approach in an early line of research. Their algorithm uses 1-bit feedback to provide a gradient estimate to the transmitter [17]. The transmitter generates test beamforming vectors from the current beamforming vector, encodes pilot signals with these beamformers, and includes these pilots along with the transmission (the rest of which is encoded with the current beamformer). The receiver determines which test beamformer is more optimal for the current channel realization and provides feedback to indicate which one the transmitter should select as the next active beamformer. For each test interval, the transmitter generates the test beamforming vectors as

$$\mathbf{f}_1 = \mathbf{f} + \|\mathbf{f}\|\beta\mathbf{p}$$

$$\mathbf{f}_2 = \mathbf{f} - \|\mathbf{f}\|\beta\mathbf{p}$$

where $\mathbf{f}$ is the current beamforming vector, $\mathbf{p}$ is a zero mean complex random Gaussian vector with autocorrelation matrix $2\mathbf{I}$, and β is the adaptation rate. The pilot signal is encoded with $\mathbf{f}_1$ during even slots and $\mathbf{f}_2$ during odd slots. A single bit of feedback from the receiver in each test interval indicates which of $\mathbf{f}_1$or $\mathbf{f}_2$ should be used as $\mathbf{f}$ in the next test interval. Over time, the transmitter follows the gradient to the optimal beamforming vector. The algorithm converges in a static channel and tracks time correlated channels. The authors' simulations show that the algorithm has lower BER than early quantized beamforming and diversity space-time coding methods for moderate fading speeds [17].

Temporal correlation can also be exploited by predictive coding methods that use the Grassmannian techniques developed in early work. One approach partitions the surface of a spherical cap on the Grassmann manifold to design beamforming codebooks for the current channel realization using knowledge of preceding channel realizations [8].

A recently proposed algorithm exploits the differential geometric properties of the Grassmann manifold to predict optimal beamforming vectors in a temporally correlated channel [11]. The fundamental concept is to predict the next beamformer from a tangent vector from the previous beamformer to the current beamformer. The error in this prediction is measured at the receiver, quantized, and fed back to the transmitter. Let $\mathbf{x}[k]$ be the current beamforming vector, $\mathbf{x}[k-1]$ the previous, and $p = \mathbf{x}^*[k-1]\mathbf{x}[k]$ their inner product. This paper defines the chordal distance between beamforming vectors as

$$d = \sqrt{1- \mid p \mid^2}.$$

The tangent vector from $\mathbf{x}[k-1]$ to $\mathbf{x}[k]$ is constructed as

$$\mathbf{e}[k] = \arctan\left(\frac{d}{\mid p \mid}\right)\frac{(\mathbf{x}[k]/p) - \mathbf{x}[k-1]}{d/\mid p \mid}.$$

This tangent is interpreted as the unwrapping of the arc between $\mathbf{x}[k-1]$ and $\mathbf{x}[k]$ on the tangent space at $\mathbf{x}[k-1]$. The shortest path on the manifold between these beamformers (the geodesic) is

$$G(\mathbf{x}[k-1], \mathbf{e}[k], t) = \mathbf{x}[k-1]\cos(\|\mathbf{e}[k]\| t) + \frac{\mathbf{e}[k]}{\|\mathbf{e}[k]\|}\sin(\|\mathbf{e}[k]\| t)$$

where t is the step size parameter. The tangent can be transported to $\mathbf{x}[k]$ as

$$\hat{\mathbf{e}}[k] = \arctan\left(\frac{d}{\mid p \mid}\right)\frac{\mathbf{x}[k]p^* - \mathbf{x}[k-1]}{d}.$$

$\mathbf{x}[k+1]$ can be predicted by calculating the geodesic with this transported tangent vector as follows:

$$\hat{\mathbf{x}}[k+1] = \mathbf{x}[k]\cos(\|\hat{\mathbf{e}}[k]\| t) + \frac{\mathbf{x}[k]p^* - \mathbf{x}[k-1]}{d}\sin(\|\hat{\mathbf{e}}[k]\| t).$$

In the authors' predictive coding method, both the transmitter and receiver predict the next beamforming vector using the methods outlined above. After measuring the current channel, the receiver quantizes the error tangent vector from the predicted to actual optimal beamformer. The error tangent vector is quantized as two codewords (a real magnitude and unit-norm direction vector in $\mathbb{C}^{M_t}$) and fed back to the transmitter. The transmitter constructs its current and subsequent beamforming vectors using the geodesic and prediction formulas outlined above. This prediction method accounts for a unit delay in the feedback path. This algorithm demonstrated improvements in mean squared chordal distance error and sum rate in a delayed feedback system [11].

3 Quantized Successor Pre-coding

Much of the research in quantized feedback and predictive coding has analyzed beamforming systems. A system employing full-rank spatial multiplexing via square, unitary pre-coding matrices presents a new challenge. The techniques of Grassmannian subspace packing cannot, unfortunately, be used to develop a codebook of such pre-coders because they comprise a single, full-rank subspace. However, since the optimal pre-coder can be determined based on the channel information and a distance metric can be defined for full-rank pre-coders, a quantization and feedback method can be designed that allows the transmitter to choose a near-optimal pre-coder. If the channel matrix $\mathbf{H}$ is described via singular value decomposition as

$$\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^*,$$

the optimal pre-coder is $\mathbf{V}$ and the optimal shaping receiver matrix is $\mathbf{U}^*$. The sum rate of this system is

$$R_\Sigma = \max_{\rho_i : \sum_i \rho_i \leq \rho} \sum_{i=1}^{R_\mathbf{H}} \log_2(1 + \sigma_i^2 \rho_i),$$

where $\{\sigma\}$ is the set of singular values of $\mathbf{H}$, $R_\mathbf{H}$ is the rank of $\mathbf{H}$, and $\{\rho_i\}$ is the set of powers allocated to each stream. Capacity is achieved by waterfilling power allocation across the streams under the average power constraint, ρ. Provided that the receiver knows both $\mathbf{H}$ and the pre-coder $\mathbf{F}$ and that $\mathbf{F} = \mathbf{V}$, a linear, zero-forcing (ZF) receiver achieves capacity. If the shaping receiver $\mathbf{Z}$ is the ZF receiver

$$\begin{aligned}
\mathbf{Z} &= (\mathbf{HF})^+ \\
&= (\mathbf{F}^*\mathbf{H}^*\mathbf{HF})^{-1}\mathbf{F}^*\mathbf{H}^* \\
&= (\mathbf{V}^*\mathbf{V}\boldsymbol{\Sigma}\mathbf{U}^*\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\mathbf{V})^{-1}\mathbf{V}^*\mathbf{V}\Sigma\mathbf{U}^* \\
&= (\boldsymbol{\Sigma}^\mathbf{2})^{-1}\Sigma\mathbf{U}^* \\
&= \boldsymbol{\Sigma}^{-1}\mathbf{U}^*.
\end{aligned}$$

The recovered signal vector is

$$\begin{aligned}
\hat{\mathbf{s}} &= \mathbf{Zx} \\
&= \mathbf{ZHFs}
\end{aligned}$$

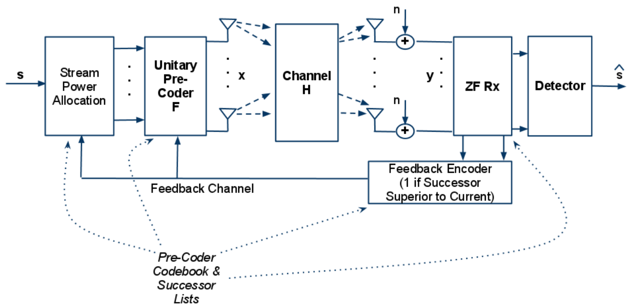

Figure 3: System Model for QSP

$$
\begin{aligned}
&= (\mathbf{HF})^{+}\mathbf{HFs} \\
&= \mathbf{s}.
\end{aligned}
$$

The enhanced noise vector in such a system is

$$\tilde{\mathbf{n}} = \mathbf{\Sigma}^{-1}\mathbf{U}^{*}\mathbf{n}.$$

Since the unitary matrix $\mathbf{U}^{*}$ does not change the statistical properties of the noise, the received SNR of the i^{th} stream is

$$
\begin{aligned}
SNR_i &= \rho_i/\tilde{n}_i \\
&= \rho_i/(n_i/\sigma_i^2) \\
&= \sigma_i^2\rho_i/n_i,
\end{aligned}
$$

where $\{n_i\}$ is the set of i.i.d. arbitrary white Gaussian noise (AWGN) on the streams. Capacity is then achieved by waterfilling power allocations. Using any other pre-coder with a ZF receiver will produce greater noise enhancement. If a quantized set of pre-coders is used, the transmitter should choose the quantized pre-coder "closest" to the optimal pre-coder for the current channel realization. The remainder of this paper describes a method that allows the transmitter to choose such a pre-coder from a codebook using very little feedback from the receiver. The intent is to provide a method that outperforms pre-coder-less equal power transmission by feeding back just enough information for the transmitter to choose a good pre-coder and decide which streams to allocate power to.

Quantized Successor Pre-coding (QSP) is a method for a spatial multiplexing system

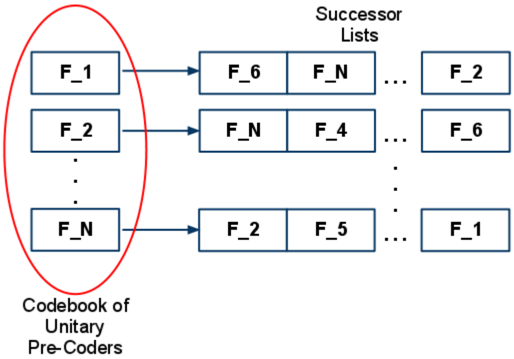

Figure 4: Example of QSP Codebook and Successor Lists

that allows the transmitter to iteratively converge to a near-optimal pre-coder for the current channel using periodic, 1-bit feedback from the receiver. The QSP codebook contains a set of square, unitary pre-coders as well as an ordered list of successors for each pre-coder. Each successor represents a candidate pre-coder that the transmitter may use as the next active pre-coder if the receiver indicates it would be better. The successors are ordered by increasing distance from the active pre-coder. In this way, the method takes advantage of temporal correlation of the channel, since because the channel changes only slightly between transmissions, a better pre-coder is more likely to be "close" to the active one. The codebook is generated offline and known to both the transmitter and receiver. Since the receiver knows the current channel, the active pre-coder, and the next successor pre-coder, it can indicate to the transmitter whether the successor would be better using a single bit.

3.1 Codebook Generation

Codebook generation follows an adaptation of the Lloyd algorithm employed in [2, 13]. The distance metric between pre-coders used in the codebook generation algorithm is the squared Frobenius-norm of their difference,

$$\delta(\mathbf{F}, \hat{\mathbf{F}}) = \|\mathbf{F} - \hat{\mathbf{F}}\|_F^2.$$

A single codebook generation is described in Algorithm 1. As noted in [2, 13], the algorithm is not guaranteed to converge to a global optimum. The entire algorithm is repeated a large number of times and the final codebook is selected as the one with the minimum mean distortion. Specifically, the minimum mean distance between training matrices and the optimal pre-coding matrix for the corresponding region.

After the set of pre-coders is chosen, a successor list is generated for each pre-coder. For each pre-coder $\mathbf{F}_i$ this is an ordered list of the other pre-coders $\mathbf{F}_j \in \mathcal{F}, i \neq j$. The list is ordered by increasing distance $\delta(\mathbf{F}_i, \mathbf{F}_j)$.

Algorithm 1 Codebook Generation

0: Generate a random initial codebook $\mathcal{F}$ of unitary $M_t \times M_t$ pre-coder matrices.
1: Generate a random set $\mathcal{T}$ of unitary $M_t \times M_t$ training matrices .
2: Assign each training matrix to the region corresponding to the "nearest neighbor" rule:
$\mathbf{T}_i \in \mathcal{R}_i$ if $\delta(\mathbf{F}_i, \mathbf{T}_i) \leq \delta(\mathbf{F}_j, \mathbf{T}_i), \forall j \neq i$.
3: Determine the optimal pre-coding matrix for each region as the orthogonal projection of the arithmetic mean of the region onto the set of unitary matrices $\mathcal{U}(M_t, M)$ as follows:
$[\mathbf{U}_i, \mathbf{S}_i, \mathbf{V}_i] = \mathrm{SVD}(\frac{1}{N_i} \sum_{\mathbf{T} \in \mathcal{R}_i} \mathbf{T})$,
where N_i is the number of training matrices assigned to $\mathcal{R}_i$. Then,
$\mathbf{F}_i = \mathbf{U}_i \mathbf{V}_i^*$
4: The resulting set $\mathbf{F}_i$ becomes the codebook $\mathcal{F}$ for the next iteration.
Steps **1-4** are repeated 100 times.

3.2 Codebook and Mode Selection at Runtime

3.2.1 Equal Power, Full Rank Option

For each pre-coder in the codebook, there is an ordered successor list. At each transmission, the receiver calculates whether the current pre-coder is achieves a higher rate than the next pre-coder in the successor list. For the ZF receiver, the sum rate is calculated as

$$R_\Sigma = \sum \log_2(1 + \gamma_i),$$

where

$$\gamma_i = \frac{SNR}{M_t \|\mathbf{z}_i\|^2}$$

and $\mathbf{z}_i$ is the i^{th} row of the zero-forcing receiver $\mathbf{Z} = (\mathbf{HF})^+$.

A single bit of feedback indicates whether the transmitter should begin using the successor pre-coder. If the pre-coder is not changed, the receiver will evaluate the next pre-coder in the successor list after the next transmission. If the pre-coder is changed, the receiver will begin evaluating at the head of the successor list of the new pre-coder. The transmitter sends M_t streams, each allocated equal power. This routine is formalized in Algorithm 2.

Algorithm 2 Equal Power, Full Rank Transmission

Initialization

```
1: ActivePreCoder := F_1
2: SuccessorIndex := 1
```

Transmitter Loop

```
1: Transmit using
     ActivePreCoder, allocating
     equal power to M_t streams
2: Successor :=
     SuccessorList(
       ActivePreCoder,
       SuccessorIndex)
3: if (Feedback == 0)
4:   SuccessorIndex :=
       (SuccessorIndex + 1) mod
       (SuccessorListSize)
5: else
6:   ActivePreCoder := Successor
7:   SuccessorIndex := 1
8: end if
```

Receiver Loop

```
1: Decode packet using
     ActivePreCoder and
     measures current channel
2: Successor :=
     SuccessorList(
       ActivePreCoder,
       SuccessorIndex)
3: if (Rate(Successor) >
       Rate(ActivePreCoder))
4:   Feedback := 1
5:   ActivePreCoder := Successor
6:   SuccessorIndex := 1
7: else
8:   Feedback := 0
9:   SuccessorIndex :=
       (SuccessorIndex + 1) mod
       (SuccessorListSize)
10: end if
```

In a static channel, this algorithm is guaranteed to converge to the optimum (quantized) pre-coder since the receiver will eventually compare every pre-coder in the successor list (which contains all pre-coders). In a temporally correlated channel, the initial convergence rate depends on the coherence time of the channel and the size of the codebook. If the coherence time is large relative to the transmission rate (which is also the rate of pre-coder innovation), the routine should converge to the neighborhood of best pre-coders. The routine will continuously track the channel, since the pre-coders are continuously evaluated by the receiver.

There is a trade-off between pre-coder optimality and convergence time. A larger codebook will have less distortion from the optimal pre-coders for each channel. However, a larger codebook will have a larger convergence time. This is mitigated by the ordering of the successor list, since pre-coder candidates in the neighborhood of the current pre-coder are evaluated first.

3.2.2 Equal Power, Adaptive Mode Option

In the preceding method, the transmitter attempts to send all M_t possible streams with equal power. This is sub-optimal from the perspective of sum rate maximization, since power may be be allocated to streams with small singular values. As an alternative, the method is modified so that the successor list represents both pre-coders and modes. This algorithm first converges to the best pre-coder, then converges to the best mode (how many streams will transmitted at once). The receiver begins by evaluating successor pre-coders. When a pre-coder is found that is better than its immediate successor, the receiver begins evaluating the sum rate using different numbers of streams. Whether the sum rate will be higher with a different number of streams (with equal power allocation) is indicated to the transmitter with a single bit of feedback. When the best mode for the current pre-coder is determined, the receiver resumes evaluating pre-coders in the successor list. This is formalized in Algorithms 3 and 4. At any given time, if M streams are active, the transmitter will use the first M columns of the current pre-coder (which should correspond to the M largest singular values).

Algorithm 3 Equal Power, Adaptive Mode (Tx Side)

Initialization

```
ActivePreCoder := F_1
SuccessorIndex := 1
ActiveStreams := M_t
SuccessorStreams := M_t
```

Transmitter Loop

```
Transmit using
  ActivePreCoder, allocating
  equal power across
  ActiveStreams
Successor :=
  SuccessorList(
    ActivePreCoder,
    SuccessorIndex)
if (Feedback == 0)
  SuccessorIndex :=
    (SuccessorIndex + 1) mod
    (SuccessorListSize)
  SuccessorStreams := M_t
  while (SuccessorStreams > 0)
    Transmit using
      ActivePreCoder, allocating
      equal power across
      ActiveStreams
    if (Feedback == 1)
      ActiveStreams :=
        SuccessorStreams
    end if
    decrement SuccessorStreams
  end while
else
  ActivePreCoder := Successor
  SuccessorIndex := 1
end if
```

Algorithm 4 Equal Power, Adaptive Mode (Rx Side)

Initialization

```
ActivePreCoder := F_1
SuccessorIndex := 1
ActiveStreams := M_t
SuccessorStreams := M_t
```

Receiver Loop

```
Decode packet using
  ActivePreCoder, ActiveStreams
  and measures current channel
Successor :=
  SuccessorList(
    ActivePreCoder,
    SuccessorIndex)
if (Rate(Successor) >
    Rate(ActivePreCoder))
  Feedback := 1
  ActivePreCoder := Successor
  SuccessorIndex := 1
else
  Feedback := 0
  SuccessorIndex :=
    (SuccessorIndex + 1) mod
    (SuccessorListSize)
  SuccessorStreams := M_t
  while (SuccessorStreams > 0)
    Decode packet using
      ActivePreCoder, ActiveStreams
      and measures current channel
    if (Rate(SuccessorStreams) >
        Rate(ActiveStreams))
      Feedback := 1
      ActiveStreams :=
        SuccessorStreams
    else
      Feedback := 0
    end if
    decrement SuccessorStreams
  end while
```

Although this routine mitigates the sub-optimal allocation of power to poor streams, it negatively affects the convergence time. By containing both pre-coders and modes, the successor list is lengthened to $M_t card(\mathcal{F})$.

3.2.3 Equal Power, Adaptive Mode with Increased Feedback Option

To mitigate the impact of the adaptive mode method on the successor list length, a third option is proposed. This routine is identical to Algorithm 2, except that the receiver also feeds back the optimal number of streams to transmit using the current pre-coder. This removes the transmission modes from the successor list, but adds a small amount of feedback ($\log_2(M_t)$ bits per feedback channel use).

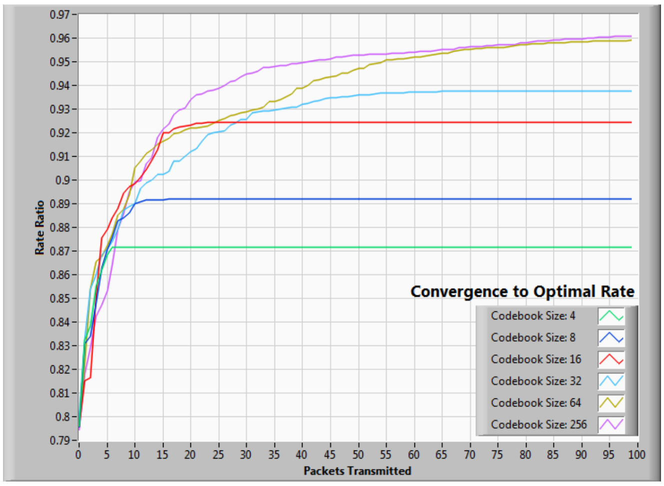

Figure 5: 2x2, Static Channel, Full Rank

4 Simulation and Results

The Equal Power, Full Rank and Equal Power, Adaptive Mode algorithms were implemented in LabVIEW. This section presents results for simulations on 2x2 and 4x4 MIMO systems with spatially independent Rayleigh fading channels. Temporally correlated channels were generated according to a first-order auto-regressive model,

$$\mathbf{H}[t] = \alpha \mathbf{H}[t-1] + \sqrt{1+\alpha^2}\mathbf{G}[t],$$

where $\mathbf{G}[t]$ is a $M_t \times N_r$ matrix of i.i.d. entries $\sim CN(0,1)$.

$$\alpha = J_0(2\pi\beta),$$

where J_0 is a Bessel function of zeroth order and β is the normalized Doppler frequency [10, 11].

4.1 Convergence toward Optimum Rate

An important metric is how well QSP converges toward the optimum rate in both static and time-varying channels. The rate ratio metric is simply the sum rate achieved using the active pre-coder and active number of streams relative to the capacity achievable with SVD pre-coding and waterfilling power allocation.

Figures 5-8 depict the convergence in a static channel for both methods and various codebook sizes. The algorithm takes longer to converge and converges to a lower rate in

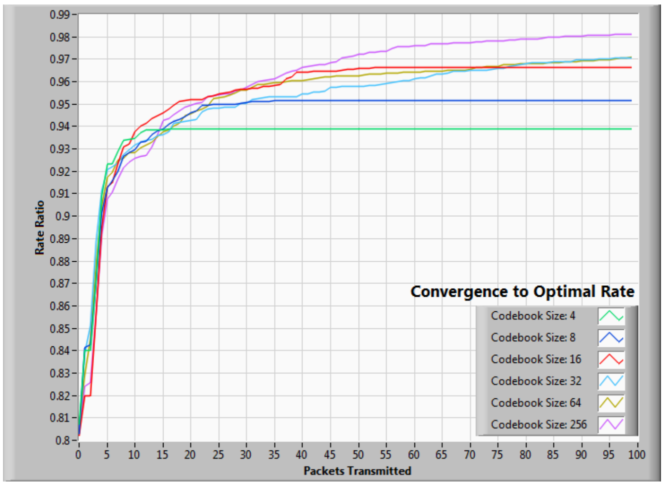

Figure 6: 2x2, Static Channel, Adaptive Mode

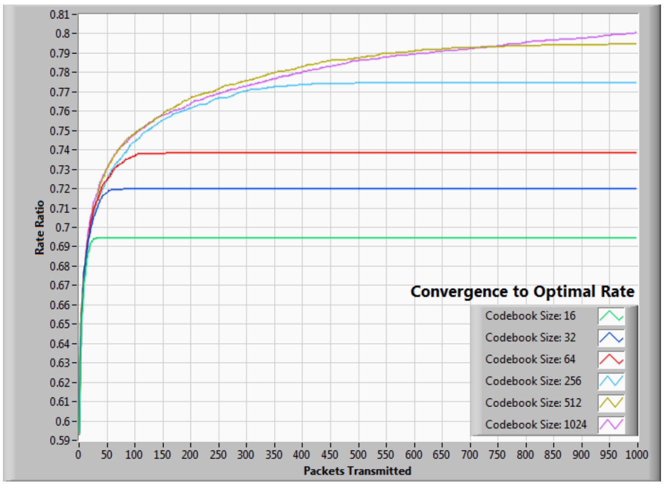

Figure 7: 4x4, Static Channel, Full Rank

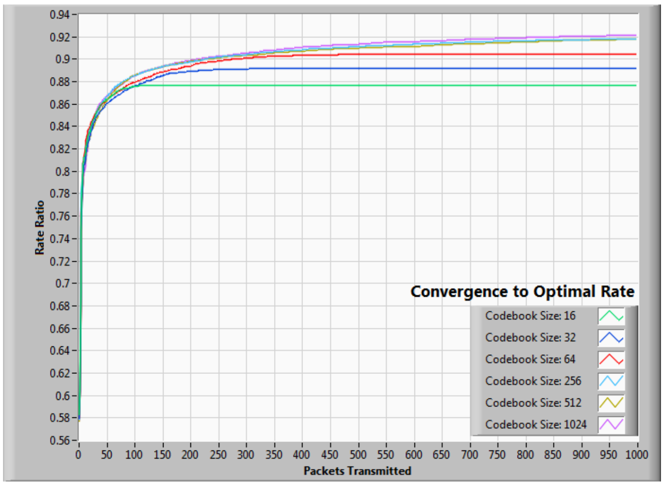

Figure 8: 4x4, Static Channel, Adaptive Mode

a 4x4 system relative a 2x2 system. The Full Rank method (which attempts to use M_t streams at all times) converges to a lower rate than the Adaptive Mode method (which adapts the number of streams to the channel). This rate gap is particularly evident in the 4x4 system.

Figures 9-14 depict the convergence in a correlated, time-varying channel. Once again, the Adaptive Mode method produces significantly better rate ratios than Full Rank method. With a more quickly varying channel, the Full Rank option performs very sub-optimally in a 4x4 system (Figure 11). However, the Adaptive Mode option performs much better in this channel (Figure 12). Both methods perform well in a 4x4 system with a more slowly varying channel.

4.2 Rate as a Function of SNR

This section reports the performance of the QSP methods against both a perfect CSIT system and an open-loop system that allocates power equally across M_t streams. The metric is the sum rate, and the plots show rate vs. SNR over a number of codebook sizes.

Figures 15-17 depict the Rate vs. SNR performance of the algorithms for a 2x2 system. Both algorithms outperform open-loop spatial multiplexing at all SNRs and all codebook sizes. At low SNR, the Adaptive Mode method outperforms the Full Rank method, and the performance is less dependent on the codebook size. For high SNR, Full Rank actually outperforms Adaptive Mode, which eventually performs below open-loop spatial multiplexing (at very high SNR). This is due to the fact that the Adaptive

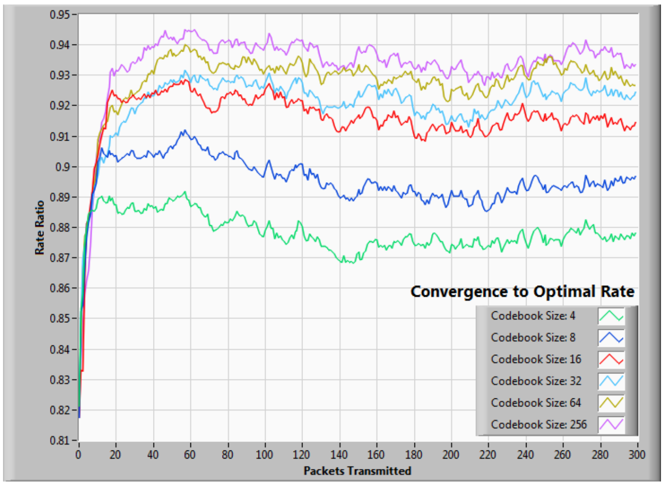

Figure 9: 2x2, $\beta = .01$, Full Rank

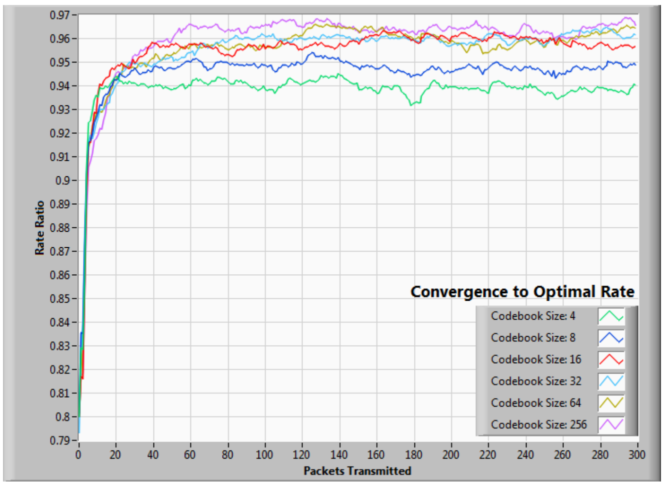

Figure 10: 2x2, $\beta = .01$, Adaptive Mode

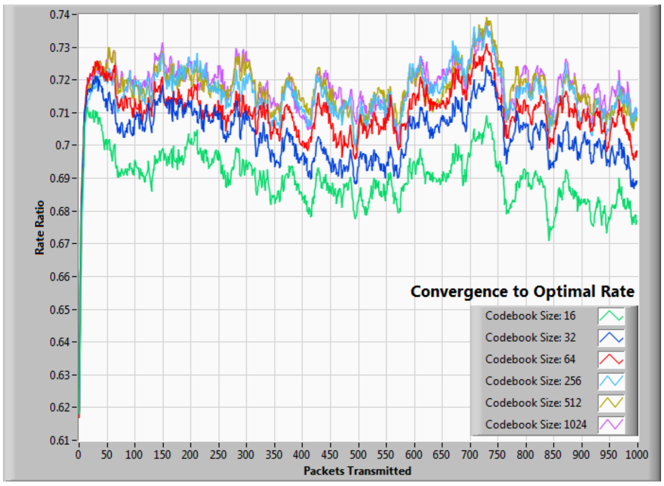

Figure 11: 4x4, $\beta = .01$, Full Rank

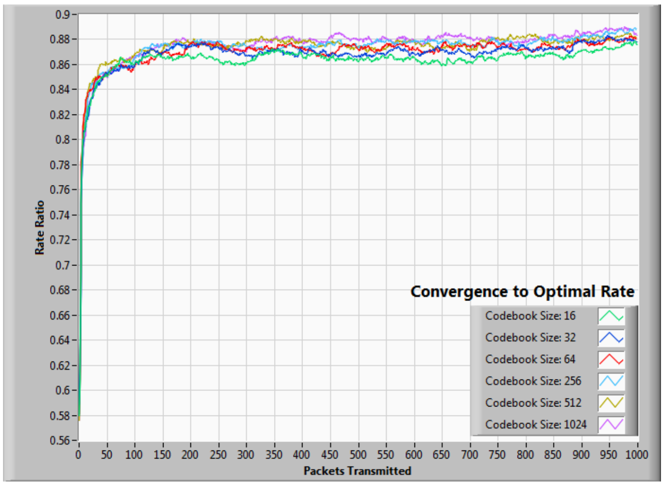

Figure 12: 4x4, $\beta = .01$, Adaptive Mode

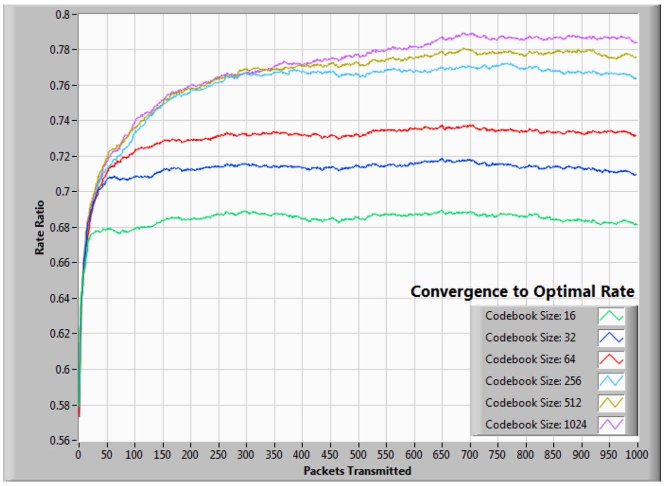

Figure 13: 4x4, $\beta = .001$, Full Rank

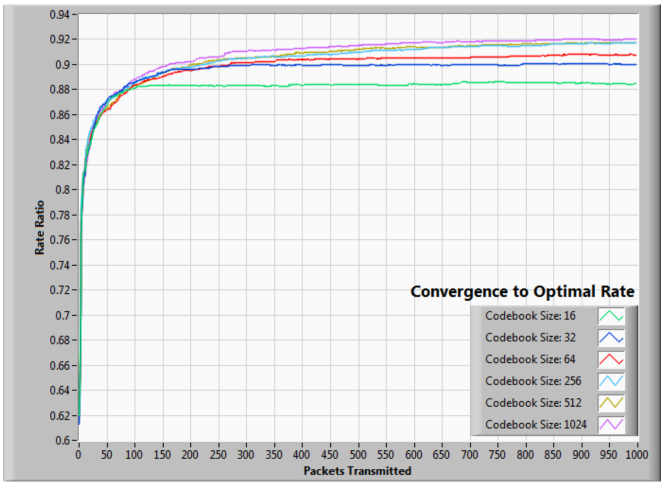

Figure 14: 4x4, $\beta = .001$, Adaptive Mode

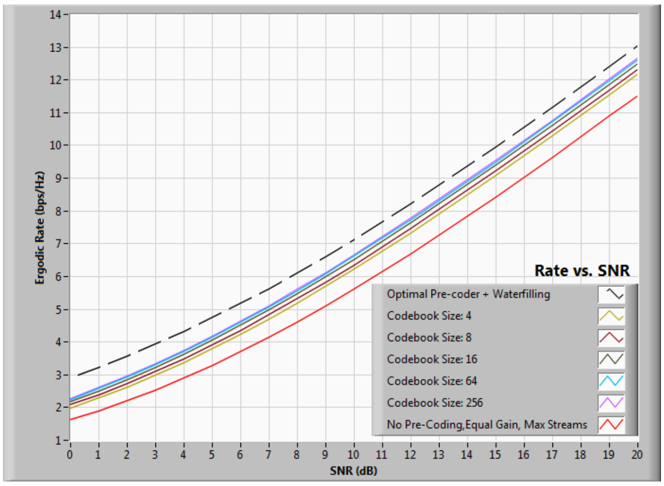

Figure 15: 2x2, $\beta = .01$, Full Rank

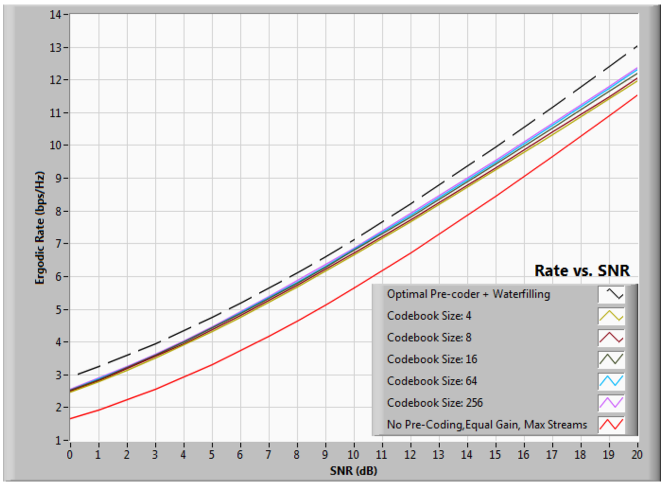

Figure 16: 2x2, $\beta = .01$, Adaptive Mode

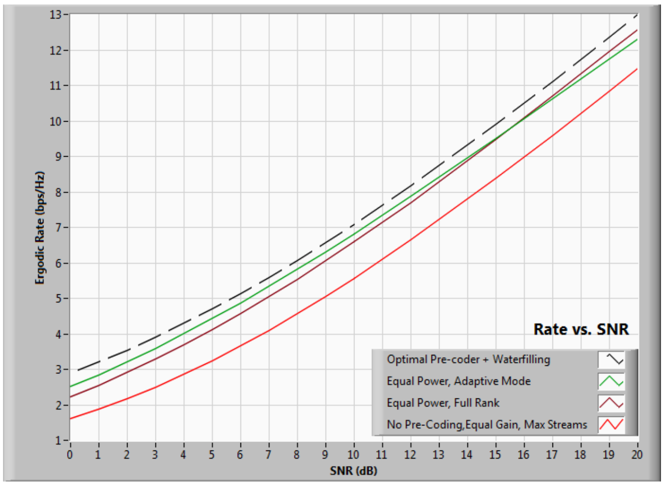

Figure 17: 2x2, $\beta = .01$, Comparison, Codebook Size=256

Mode algorithm may stick to a reduced number of streams for several transmissions while pre-coders are evaluated. This is less beneficial at high SNRs, when all streams can be used.

Figures 18-20 depict the Rate vs. SNR performance of the algorithms for a 4x4 system. The results are very similar to the 2x2 system. A notable result is that Adaptive Mode substantially outperforms Full Rank at low SNR, but they converge to similar performance at very high SNR.

The third option, Adaptive Mode with Increased Feedback, was not simulated, but should combine the improved sum rate performances of Adaptive Mode with the ability to respond to quickly-varying channels of Full Rank. This comes at the cost of an additional $\log_2(M_t)$ bits of feedback.

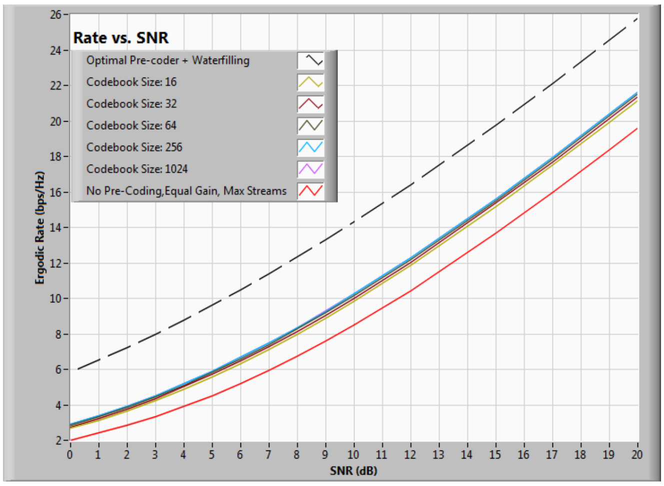

Figure 18: 4x4, $\beta = .01$, Full Rank

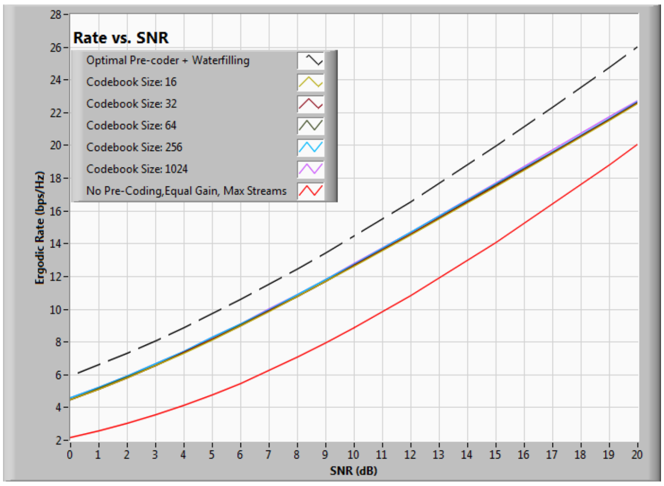

Figure 19: 4x4, $\beta = .01$, Adaptive Mode

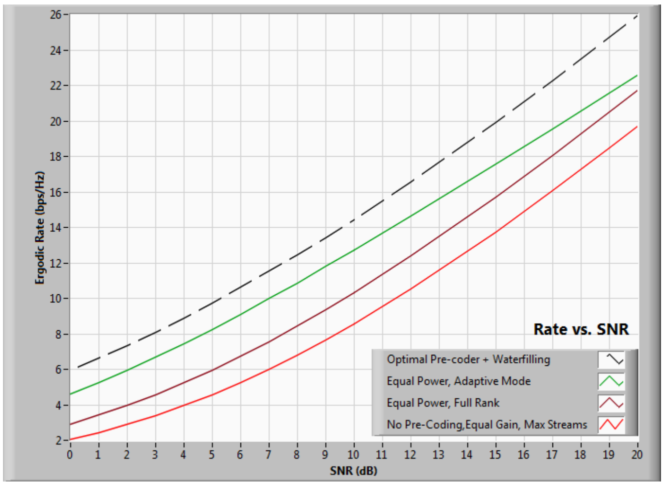

Figure 20: 4x4, $\beta = .01$, Comparison, Codebook Size=1024

5 Conclusion

The results demonstrate that QSP achieves significant improvement in sum rate in comparison to open-loop spatial multiplexing. Except at very high SNR, Adaptive Mode outperforms Full Rank regardless of the codebook size, number of antennas, and channel coherence time. In a 2x2 system with a channel coherence time of around 100 transmission blocks, both methods increase the ergodic rate by around 1 bps/Hz at a given SNR. For a 4x4 system, there is a more pronounced benefit using Adaptive Mode. For this system, the Full Rank method realizes an increase in roughly 2 bps/Hz at a given SNR, while the Adaptive Mode option realizes up to a 4 bps/Hz improvement.

The size of the codebook has an impact on the performance, especially in the higher antenna count systems. Luckily, the amount of feedback is not changed by the codebook size. The codebook size can be determined according to the desired trade off between sum rate performance and convergence time (smaller codebooks may be preferred for faster-varying channels).

5.1 Future Work

One obvious avenue for further research is optimizing the power allocation. The methods developed here all allocate equal power to the active streams. While this is done to reduce feedback, it is sub-optimal compared to a method that combines pre-coding with optimal power allocation. The power allocations could be quantized using vector quantization techniques, and the index of the best allocation could be included in the feedback to the transmitter. Alternatively, the Adaptive Mode option could be extended to include power allocations in the codebook successor lists. This would achieve a more optimal allocation of power to streams while requiring only a single bit of feedback.

In early experimentation, the receiver compared pre-coders by calculating their distance to the optimal (SVD) pre-coder for the current channel. The successor was considered better if it was "closer" to the optimal pre-coder. This comparison metric did not perform as well as the method discussed in this paper, which directly compares the rates achieved by the pre-coders. This implies that the distance metric used to order pre-coders in the successor list may be sub-optimal. Future research could refine this distance metric and shorten the convergence time to improve performance in faster-varying channels.

Future investigation could also evaluate the system in the presence of delay in the feedback path. As long as the feedback delay is small relative to the coherence time of the channel, the system should be fairly robust, but additional simulation could quantify the performance in the presence of feedback lag.

Another small optimization is possible due to the nature of single-bit feedback. Depending on the actual implementation of the feedback path, a 0 can be conveyed by simply not sending anything at all. This would reduce the power requirements of the receiver and possibly reduce the interference in multi-user systems.

References

[1] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, “An overview of limited feedback in wireless communication systems,” *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 8, pp. 1341 –1365, 2008.

[2] V. Lau, Y. Liu, and T.-A. Chen, “On the design of mimo block-fading channels with feedback-link capacity constraint,” *Communications, IEEE Transactions on*, vol. 52, no. 1, pp. 62 – 70, 2004.

[3] A. Dabbagh and D. Love, “Feedback rate-capacity loss tradeoff for limited feedback mimo systems,” *Information Theory, IEEE Transactions on*, vol. 52, no. 5, pp. 2190 – 2202, May 2006.

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

[5] D. Love, J. Heath, R.W., and T. Strohmer, “Grassmannian beamforming for multiple-input multiple-output wireless systems,” in *Communications, 2003. ICC '03. IEEE International Conference on*, vol. 4, May 2003, pp. 2618 – 2622 vol.4.

[6] B. Mondal, S. Dutta, and R. Heath, “Quantization on the grassmann manifold,” *Signal Processing, IEEE Transactions on*, vol. 55, no. 8, pp. 4208 –4216, 2007.

[7] B. Mondal and R. Heath, “Channel adaptive quantization for limited feedback mimo beamforming systems,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 12, pp. 4717 –4729, 2006.

[8] L. Liu and H. Jafarkhani, “Novel transmit beamforming schemes for time-selective fading multiantenna systems,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 12, pp. 4767 –4781, 2006.

[9] D. Love and J. Heath, R.W., “Equal gain transmission in multiple-input multiple-output wireless systems,” *Communications, IEEE Transactions on*, vol. 51, no. 7, pp. 1102 – 1110, 2003.

[10] K. Huang, R. Heath, and J. Andrews, “Limited feedback beamforming over temporally-correlated channels,” *Signal Processing, IEEE Transactions on*, vol. 57, no. 5, pp. 1959 –1975, May 2009.

[11] T. Inoue and R. Heath, "Grassmannian predictive coding for delayed limited feedback mimo systems," in *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on*, 30 2009.

[12] T. Kim, D. Love, B. Clerckx, and S. J. Kim, "Differential rotation feedback mimo system for temporally correlated channels," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, 30 2008.

[13] N. Khaled, B. Mondal, G. Leus, R. Heath, and F. Petre, "Interpolation-based multimode precoding for mimo-ofdm systems with limited feedback," *Wireless Communications, IEEE Transactions on*, vol. 6, no. 3, pp. 1003 –1013, 2007.

[14] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, pp. 311 – 335, 1998.

[15] E. Telatar, "Capacity of multi-antenna gaussian channels," *Euro. Trans. Telecommun.*, pp. 585 – 596, 1999.

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

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