

MIMO Spatial Multiplexing Systems with Limited Feedback

June Chul Roh and Bhaskar D. Rao

Department of Electrical and Computer Engineering

University of California, San Diego

La Jolla, CA 92093-0407, USA

Email: jroh@ece.ucsd.edu, brao@ece.ucsd.edu

Abstract—This paper investigates the problem of transmit beamforming in MIMO spatial multiplexing (SM) systems with finite-rate feedback channel. Assuming a fixed number of spatial channels and equal power allocation, we propose a new design criterion for designing the codebook of beamforming matrices that is based on minimizing the capacity loss resulting from the limited rate in the feedback channel. Using the criterion, we develop an iterative design algorithm that converges to an optimum codebook. Under the *i.i.d.* channel and high SNR assumption, the effect on channel capacity of the finite-bit representation of beamforming matrix is analyzed. Central to this analysis is the complex multivariate beta distribution and tractable approximations to the Voronoi regions associated with the code points. Furthermore, to compensate for the degradation due to the equal power allocation assumption, we propose a multi-mode SM transmission strategy wherein the number of data streams is determined based on the average SNR. This approach is shown to allow for effective utilization of the feedback bits.

I. INTRODUCTION

The performance of a multiple-antenna communication system depends on the nature of channel state information (CSI) available at the transmitter and at the receiver. When the transmitter has perfect CSI, a higher capacity link can be achieved in the single user case, and there are other benefits such as lower-complexity receivers and better system throughput in a multiuser environment. However, the assumption that the transmitter has perfect knowledge of multi-dimensional channel could be unrealistic as in many practical systems the channel information is provided to the transmitter through a finite-rate feedback channel. There are several studies dealing with how to feed back the channel information. Some researchers have worked on feedback of channel information in vector forms, for example, for MISO channels [1], [2], [3] and for the principal eigen-mode of MIMO channels [4]. Only recently, feedback of channel information in matrix forms for MIMO channels have begun to be addressed [5], [6], [7]. In this paper, we develop efficient transmit beamforming techniques for MIMO spatial multiplexing (SM) systems with finite-rate feedback.

The contributions of the paper are as follows. First, a tractable measure of capacity loss due to finite feedback rate is derived and used to design an optimal codebook of beamforming matrices in the context of a fixed number of spatial channels with equal power allocation. The efficacy of the approach is demonstrated by using it to design quantizers in

a variety of MIMO contexts and evaluating their performance through computer simulations. Then the challenging task of analyzing the performance of MIMO systems employing such finite-rate feedback techniques is undertaken, and interesting analytical results are provided. A key feature of the analysis is the connection to the complex matrix-variate beta distribution and the approximation made of the Voronoi region associated with each code point for analytical tractability. To compensate for the degradation due to the equal power allocation assumption, we also propose a multi-mode SM transmission strategy wherein the number of data streams is determined based on the average SNR. This approach necessitates generalization of the codebook design methodology, and the overall approach is shown to allow for effective utilization of the feedback bits.

We use the following notations. A^\dagger and A^T indicate the conjugate transpose and the transpose of matrix A , respectively. $\|A\|_F$ denotes the Frobenius norm of matrix A . An $m \times n$ ($m \geq n$) matrix A with $A^\dagger A = I_n$ will be called *orthonormal column matrix*. $A > 0$ means that matrix A is positive definite. $\tilde{\mathcal{N}}_r(\mu, \Sigma)$ is the r -dimensional proper complex Gaussian random vector with mean μ and covariance Σ . Uniform distribution over a set S is denoted by $\mathcal{U}(S)$. The function $\log(\cdot)$ is the natural logarithm.

II. SYSTEM MODEL

A MIMO channel with t transmit and r receive antennas, assuming flat fading in each antenna pair, is modeled by the channel matrix $H \in \mathbb{C}^{r \times t}$. That is, the channel input $x \in \mathbb{C}^t$ and the channel output $y \in \mathbb{C}^r$ have the following relationship:

$$y = Hx + \eta \quad (1)$$

where η is the additive white Gaussian noise distributed by $\tilde{\mathcal{N}}_r(0, I)$. For this work, we assume¹ $t \geq r$ and the rank of H is denoted by m . The singular value decomposition (SVD) of H is given by $H = U_H \Sigma_H V_H^\dagger$, where $U_H \in \mathbb{C}^{r \times m}$ and $V_H \in \mathbb{C}^{t \times m}$ are orthonormal column matrices and $\Sigma_H \in \mathbb{R}^{m \times m}$ contains the singular values $\sigma_1 \geq \dots \geq \sigma_m > 0$ of H . The average transmit power is denoted by P_T , i.e., $E[x^\dagger x] = P_T$.

When perfect CSI is known at the receiver and the transmitter, a water-filling based SM strategy is known to be optimal. This means that information about V_H and the power allocation on the different channels/beams represented by the columns of V_H has to be fed back to the transmitter. With only finite number of bits available for channel information

This research was supported by UC Discovery Grant No. core02-10109 sponsored by Ericsson and in part by the U.S. Army Research Office under MURI Grant No. W911NF-04-1-0224.

¹The case $t \leq r$ can also be handled in a similar manner.

feedback, we concentrate first on representing the first n ($1 \leq n \leq m$) column vectors of V_H . For notational convenience, let us denote the i th column vector of V_H by v_i and define the first n column vectors of V_H as V , i.e., $V = [v_1, \dots, v_n]$. We also define $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n)$, an $n \times n$ submatrix of Σ_H .

To formalize the problem, we assume that the MIMO system has a feedback channel with a finite rate of B bits per channel update. A codebook \mathcal{C} containing N , where $N = 2^B$, candidate beamforming matrices, i.e., $\mathcal{C} = \{\hat{V}_1, \dots, \hat{V}_N\}$ with \hat{V}_i being $t \times n$ orthonormal column matrices, is assumed known to both the receiver and the transmitter. The design of the codebook is a topic of this paper. The receiver selects the optimum beamforming matrix $\hat{V} = \mathcal{Q}(H)$ from the codebook \mathcal{C} based on the current channel H , and sends the index of \hat{V} to the transmitter through the feedback channel. This paper assumes feedback with no error and no delay and focuses on the effect of finite-bit representation of the channel information.

The channel information $\hat{V} = \mathcal{Q}(H)$ is employed as the beamforming matrix at the transmitter. That is, an information-bearing symbol vector $s = [s_1, \dots, s_n]^T$ is transmitted as $x = \hat{V}s$, resulting in the received signal

$$y = H\hat{V}s + \eta. \quad (2)$$

Here we assume that $s \sim \tilde{\mathcal{N}}_n(0, \Phi)$ and $\Phi = P_T \cdot \text{diag}(\varphi_1, \dots, \varphi_n)$ with $\varphi_i \geq 0$ and $\sum_i \varphi_i = 1$. The vector $\varphi = [\varphi_1, \dots, \varphi_n]$ will be called *power allocation information*. With perfect CSI at receiver, the optimum power allocation for the equivalent channel $H\hat{V}$ can be calculated at the receiver. Instead of feeding back the power allocation information, which reduces the bit budget for the beamforming matrix, one can employ a simple equal power allocation strategy, i.e., $\varphi_i = 1/n \forall i$. For a given n , with equal power allocation, the mutual information between s and y is given by

$$I(H, \hat{V}) = \log \det(I + \rho \Sigma_H^2 V_H^\dagger \hat{V} \hat{V}^\dagger V_H) \quad (3)$$

where $\rho := P_T/n$. This paper mainly focuses on equal power allocation because it is more amenable to optimum codebook design, and results in minor performance degradation when coupled with an efficient multi-mode SM transmission wherein n is chosen based on the average SNR. The multi-mode SM scheme is discussed more fully in Section V.

III. CODEBOOK DESIGN FOR BEAMFORMING MATRIX

In this section, we develop a general matrix quantization (MQ) based design method for constructing the codebook of beamforming matrices. We consider the case of $n = m$ in this section, and $n < m$ is discussed in Section V.

A. Capacity Loss Due to Finite Rate Feedback

After a little manipulation, (3) can be written as

$$I(H, \hat{V}) = \log \det(I + \rho \Sigma_H^2) + \log \det[I - (I + \rho \Sigma_H^2)^{-1} \rho \Sigma_H^2 (I - V_H^\dagger \hat{V} \hat{V}^\dagger V_H)]. \quad (4)$$

Here we notice that when $n = m$, the first term is $I(H, V)$, the mutual information with the perfect beamforming matrix V at the transmitter, and the second term accounts for the loss due to finite-bit representation of V . Let us define the *capacity loss* as the difference between the ergodic capacities associated with V and \hat{V} , that is,

$$C_L(H, \hat{V}) = E[I(H, V)] - E[I(H, \hat{V})] = E[I_L(H, \hat{V})] \quad (5)$$

where $I_L(H, \hat{V}) := I(H, V) - I(H, \hat{V})$, which is from (4)

$$I_L(H, \hat{V}) = -\log \det[I - \tilde{\Sigma}_H^2 (I - V_H^\dagger \hat{V} \hat{V}^\dagger V_H)] \quad (6)$$

where $\tilde{\Sigma}_H^2 := (I + \rho \Sigma_H^2)^{-1} \rho \Sigma_H^2$.

The following two approximations to $I_L(H, \hat{V})$ will be considered: i) When $P_T \ll 1$ (low SNR), or when $V_H^\dagger \hat{V} \hat{V}^\dagger V_H$ is close to I (which is valid when B is reasonably large), we use the approximation $-\log \det(I - A) \simeq -\log(1 - \text{tr}A) \simeq \text{tr}A$ when A is small. More specifically, when the eigenvalues $\lambda_i(A) \ll 1 \forall i$, we have

$$I_L(H, \hat{V}) \simeq \text{tr}[\tilde{\Sigma}_H^2 (I - V_H^\dagger \hat{V} \hat{V}^\dagger V_H)]. \quad (7)$$

ii) When $P_T \gg 1$ (high SNR), since $\tilde{\Sigma}_H^2 \simeq I$,

$$I_L(H, \hat{V}) \simeq -\log \det(V_H^\dagger \hat{V} \hat{V}^\dagger V_H). \quad (8)$$

The first approximation will be used in codebook design next and the second for performance analysis in Section IV.

B. Codebook Design Criterion and Design Algorithm

For designing the beamforming-matrix codebook, a natural design criterion is maximizing the expected mutual information $E[I(H, \hat{V})]$ or, equivalently, minimizing the capacity loss defined in (5). However, unfortunately using it directly does not lead to an iterative design algorithm with monotonic convergence property², since generally there is no analytical expression for the optimum code matrix as a function of a given partition region in the channel space [5].

Instead of using the direct form of the capacity expression, we consider the approximation to the capacity loss: from (7), when $P_T \ll 1$ or when B is reasonably large,

$$C_L \simeq E[\text{tr} \tilde{\Sigma}_H^2 - \text{tr}(\tilde{\Sigma}_H^2 V_H^\dagger \hat{V} \hat{V}^\dagger V_H)]. \quad (9)$$

The second term inside the bracket in (9) can be written as $\|(V_H \tilde{\Sigma}_H)^\dagger \hat{V}\|_F^2$. Therefore, minimizing the expectation of (9) is equivalent to the following codebook design criterion.

New Design Criterion: Design a mapping \mathcal{Q} (mathematically, $\mathcal{Q} : \mathbb{C}^{r \times t} \rightarrow \mathcal{C}$) such that

$$\max_{\mathcal{Q}(\cdot)} E \left\| (V_H \tilde{\Sigma}_H)^\dagger \mathcal{Q}(H) \right\|_F^2 \quad (10)$$

where $\hat{V} = \mathcal{Q}(H)$ is the quantized beamforming matrix ($\hat{V}^\dagger \hat{V} = I_n$).

This design criterion will be called the *generalized MSwIP* (mean squared weighted inner-product) criterion since it can be viewed as a generalization for MIMO channels of the

²Monotonic convergence means that an improved design is guaranteed at every iteration.

MSwIP criterion that was developed for beamforming codebook design (for MISO systems) in [3].

One of the virtues of the new design criterion is that it *does* lead to an iterative design algorithm with guaranteed monotonic convergence. The design algorithm is essentially similar to the Lloyd algorithm in vector quantization (VQ) study, which is based on two necessary conditions for optimality: the neighborhood condition (NNC) and the centroid condition (CC) [8]. The same approach is used here for designing the codebook of beamforming matrices.

Design Algorithm:

1. *NNC:* For given code matrices $\{\hat{V}_i; i = 1, \dots, N\}$, the optimum partition cells satisfy

$$\mathcal{H}_i = \left\{ H \in \mathbb{C}^{r \times t} : \begin{aligned} & \|(V_H \tilde{\Sigma}_H)^\dagger \hat{V}_i\|_F^2 \\ & \geq \|(V_H \tilde{\Sigma}_H)^\dagger \hat{V}_j\|_F^2, \forall j \neq i \end{aligned} \right\}, \quad (11)$$

for $i = 1, \dots, N$, where \mathcal{H}_i is the partition cell of the channel matrix space $\mathbb{C}^{r \times t}$ for the i th code matrix \hat{V}_i .

2. *CC:* For a given partition $\{\mathcal{H}_i; i = 1, \dots, N\}$, the optimum code matrices satisfy

$$\hat{V}_i = \arg \max_{\hat{V} \in \mathbb{C}^{t \times n} : \hat{V}^\dagger \hat{V} = I_n} E \left[\|(V_H \tilde{\Sigma}_H)^\dagger \hat{V}\|_F^2 \mid H \in \mathcal{H}_i \right], \quad (12)$$

for $i = 1, \dots, N$. Fortunately, the above optimization problem has a closed-form solution as (see proof below)

$$\hat{V}_i = (n \text{ principal eigenvectors}) \text{ of } E[V_H \tilde{\Sigma}_H^2 V_H^\dagger \mid H \in \mathcal{H}_i]. \quad (13)$$

Proof: The Frobenius norm in (12) can be expressed as $\text{tr}(\hat{V}^\dagger V_H \tilde{\Sigma}_H^2 V_H^\dagger \hat{V})$. By noting that under a given condition $H \in \mathcal{H}_i$, \hat{V} is a constant matrix, (12) can be rewritten as

$$\hat{V}_i = \arg \max_{\hat{V} \in \mathbb{C}^{t \times n} : \hat{V}^\dagger \hat{V} = I_n} \text{tr} \left\{ \hat{V}^\dagger E[V_H \tilde{\Sigma}_H^2 V_H^\dagger \mid H \in \mathcal{H}_i] \hat{V} \right\}.$$

By applying Corollary 4.3.18 of [9, p. 191], we arrive at the desired solution given in (13). ■

The above two conditions are iterated until the design objective $E \|(V_H \tilde{\Sigma}_H)^\dagger \mathcal{Q}(H)\|_F^2$ converges. In practice, a codebook is designed off-line using a sufficiently large number of training samples (channel realizations). In that case, the statistical correlation matrix $E[V_H \tilde{\Sigma}_H^2 V_H^\dagger \mid H \in \mathcal{H}_i]$ in (13) is estimated with a sample average.

Encoding: For a given codebook $\mathcal{C} = \{\hat{V}_1, \dots, \hat{V}_N\}$, the receiver selects the optimum beamforming matrix from the codebook based on the observed channel H so that the mutual information is maximized, i.e., $\hat{V} = \arg \max_{\hat{V}_i \in \mathcal{C}} I(H, \hat{V}_i)$. It is useful later in performance analysis to note that this encoding scheme is equivalent to

$$\hat{V} = \mathcal{Q}(H) = \arg \min_{\hat{V}_i \in \mathcal{C}} I_L(H, \hat{V}_i). \quad (14)$$

By the encoding scheme, the channel matrix space $\mathbb{C}^{r \times t}$ is partitioned into $\{\mathcal{R}_i; i = 1, \dots, N\}$, where

$$\mathcal{R}_i = \{H \in \mathbb{C}^{r \times t} : I_L(H, \hat{V}_i) \leq I_L(H, \hat{V}_j), \forall j \neq i\}. \quad (15)$$

C. Two Related Design Methods

The generalized MSwIP design method optimizes the codebook for a particular SNR (or P_T). As a result, we may need more than one codebook if the system has multiple operating SNR points. Therefore, it would be interesting to find other design methods that do not depend on P_T .

1) *Low SNR Region:* When $\rho \ll 1$ (low SNR), $\tilde{\sigma}_i \simeq \sqrt{\rho} \sigma_i$; hence $\tilde{\Sigma}_H \simeq \sqrt{\rho} \Sigma_H$. Then, the original criterion (10) becomes $\max_{\mathcal{Q}(\cdot)} E \|(V_H \Sigma_H)^\dagger \mathcal{Q}(H)\|_F^2$. It is interesting to see that this design criterion is equivalent to maximizing the mean squared channel norm (MSCN), that is,

$$\max_{\mathcal{Q}(\cdot)} E \|H \mathcal{Q}(H)\|_F^2. \quad (16)$$

This design criterion (16) will be referred as the *MSCN criterion*. The MSCN criterion is a reasonable choice by itself, because naturally we can benefit by maximizing the gain of the composite channel ($H\hat{V}$).

2) *High SNR Region:* As the SNR increases ($\rho \rightarrow \infty$), $\tilde{\Sigma}_H \rightarrow I$; hence, the criterion (10) reduces to

$$\max_{\mathcal{Q}(\cdot)} E \|V_H^\dagger \mathcal{Q}(V_H)\|_F^2 \quad (17)$$

where $\hat{V} = \mathcal{Q}(V_H)$ is the quantized beamforming matrix ($\hat{V}^\dagger \hat{V} = I_n$). Interestingly, the above design criterion (17) can be intuitively explained using a distance between two subspaces. It is equivalent to $\min_{\mathcal{Q}(\cdot)} E [d_c^2(V_H, \mathcal{Q}(V_H))]$, where $d_c(V_H, \hat{V}_H)$ is the *chordal distance* between the two subspaces specified by V_H and \hat{V}_H [10].

D. Design Examples

With the design methods developed above, we can obtain an optimum codebook for any set of system parameters (t, r, n, B). The performances of codebooks designed with the three different design methods are compared in Fig. 1 in terms of the ergodic channel capacity, $C_{\mathcal{Q}(V)} = E[I(H, \hat{V})]$. For ease of comparison, all the capacities are normalized with respect to $C_V = E[I(H, V)]$, the ergodic capacity with the perfect beamforming matrix. As expected, in high SNR region, the codebook optimized for high SNR region performs better than that optimized for low SNR region; and in low SNR region, the reverse relation holds. Moreover, the generalized MSwIP method always results in a better performance than or equal to any of the two in all the SNR range. Additionally, we also compare with the codebook design method considered in [7], wherein a codebook is designed using the algorithm developed for the non-coherent space-time constellation design. We can see that the MSwIP method outperforms it in all the cases considered, especially in the low SNR range.

IV. CAPACITY LOSS WITH QUANTIZED BEAMFORMING

In this section, we will analytically quantify the effect of quantization of the beamforming matrix with a finite number of bits on the channel capacity for the *i.i.d.* MIMO channel. The considered MIMO system is modeled as

$$y = \sqrt{\frac{P_T}{n}} H \hat{V} s + \eta$$

where H has *i.i.d.* $\tilde{N}(0, 1)$ entries, $s \sim \tilde{N}_n(0, I)$, $\eta \sim \tilde{N}_r(0, I)$, and they are all independent. For analytical tractability, we consider the high-SNR approximation to the capacity loss using (8). We also confine our attention to the case where the number of data stream equals to the rank of channel matrix, i.e., $n = m$. For the rest of this section, for notational simplicity, the subscript in V_H will be dropped. The high-SNR approximation to the capacity loss will be denoted by

$$\bar{C}_L = E_H \left[-\log \det(V^\dagger \hat{V} \hat{V}^\dagger V) \right]. \quad (18)$$

In high SNR region, since $I_L(H, \hat{V}_i)$ is well approximated to $-\log \det(V^\dagger \hat{V}_i \hat{V}_i^\dagger V)$, the encoding scheme given in (14) also can be rewritten as

$$\hat{V} = \mathcal{Q}(V) = \arg \max_{\hat{V}_i \in \mathcal{C}} \det(V^\dagger \hat{V}_i \hat{V}_i^\dagger V). \quad (19)$$

We denote the Stiefel manifold where random matrix V lies, by $\mathcal{V}_{n,t} = \{V \in \mathbb{C}^{t \times n} : V^\dagger V = I_n\}$. With the encoding scheme (19), the Stiefel manifold is partitioned into $\{\bar{\mathcal{R}}_i; i = 1, \dots, N\}$ where

$$\bar{\mathcal{R}}_i = \{V \in \mathcal{V}_{n,t} : \det(V^\dagger \hat{V}_i \hat{V}_i^\dagger V) \geq \det(V^\dagger \hat{V}_j \hat{V}_j^\dagger V), \forall j \neq i\}. \quad (20)$$

Since from (19) $\hat{V} = \hat{V}_i, \forall V \in \bar{\mathcal{R}}_i$, (18) can be expressed as

$$\bar{C}_L = \sum_{i=1}^N P(V \in \bar{\mathcal{R}}_i) E_{V \in \bar{\mathcal{R}}_i} \left[-\log \det(V^\dagger \hat{V}_i \hat{V}_i^\dagger V) \right]. \quad (21)$$

In order to calculate (21), we need to know the conditional statistics of a random matrix $U_i := V^\dagger \hat{V}_i \hat{V}_i^\dagger V$ given $V \in \bar{\mathcal{R}}_i$. For that purpose, we start with the simpler but related problem of the *unconditional* statistics of a random matrix $U_0 := V^\dagger V_0 V_0^\dagger V$, where $V \sim \mathcal{U}(\mathcal{V}_{n,t})$ and V_0 is a *fixed* matrix in $\mathcal{V}_{n,t}$.³

A. *Statistics of $U_0 = V^\dagger V_0 V_0^\dagger V$ and $\gamma_0 = \det(U_0)$*

First we summarize related definitions and theorems from multivariate statistical analysis that are germane to the analysis. A $p \times p$ random Hermitian positive definite matrix U is said to have a multivariate beta distribution with parameter (a, b) , denoted as $U \sim \tilde{\mathcal{B}}_p(a, b)$, if its density is given by

$$f(U) = \frac{\tilde{\Gamma}_p(a+b)}{\tilde{\Gamma}_p(a)\tilde{\Gamma}_p(b)} (\det U)^{a-p} \det(I-U)^{b-p}, \quad 0 < U < I_p,$$

for $a \geq p, b \geq p$ and $f(U) = 0$ elsewhere, where $\tilde{\Gamma}_p(n)$ is the complex multivariate gamma function.

Theorem 1: If $U \sim \tilde{\mathcal{B}}_p(n_1, n_2)$ and $U = T^\dagger T$, where T is upper-triangular, then $t_{11}^2, \dots, t_{pp}^2$ are all independent and $t_{ii}^2 \sim \beta(n_1 - i + 1, n_2)$ for $i = 1, \dots, p$, where $\beta(a, b)$ means the beta distribution with parameters (a, b) .

Proof: This can be proved by extending Theorem 3.3.3 in Muirhead [11] to the complex matrix case. ■

Theorem 2 (Distribution of U_0): A random matrix H has *i.i.d.* $\tilde{N}(0, 1)$ entries, and V ($t \times m$) is the right singular matrix

³It can be shown that for the channel H with *i.i.d.* $\tilde{N}(0, 1)$ entries, the random matrix V is uniformly distributed over $\mathcal{V}_{n,t}$.

of H as defined in Section II. For a fixed orthonormal column matrix V_0 ($t \times m$), the random matrix $U_0 = V^\dagger V_0 V_0^\dagger V$ has a multivariate beta distribution $\tilde{\mathcal{B}}_m(m, t - m)$.

Proof: See [12] for detailed proof. ■

Corollary 1: $\det(U_0)$ is distributed as the product of independent beta variables, that is, $\gamma_0 = \det(U_0) \sim \prod_{i=1}^m \beta_i$, where $\beta_i \sim \beta(m - i + 1, t - m)$ and are independently distributed.

Proof: When $U = T^\dagger T$ where T is upper-triangular, $\det(U_0) = \prod_{i=1}^m t_{ii}^2$. Therefore, from Theorem 1 we arrive at the desired result. ■

We can obtain the density function for a given m although it has a long and complicated form (e.g., [13]). For $m = 2$, we have a relatively concise expression.

Corollary 2: When $m = 2$, γ_0 has the following density.

$$f_{\gamma_0}(x) = \frac{\Gamma(t)\Gamma(t-1)}{\Gamma(2t-4)} (1-x)^{2t-5} {}_2F_1(t-2, t-3; 2t-4; 1-x) \quad (22)$$

Proof: See [13]. ■

B. *Approximate Density of $\gamma = \det(V^\dagger \hat{V} \hat{V}^\dagger V)$*

Now let us look at the conditional density of $\gamma_i := \det(V^\dagger \hat{V}_i \hat{V}_i^\dagger V)$ given $V \in \bar{\mathcal{R}}_i$. From the high-SNR encoding given in (19) or (20), generally each partition cell has a complicated shape. This geometrical complexity in the partition cells makes it difficult to obtain the exact analytical form for the conditional density of γ_i .

However, when N is large, since V is uniformly distributed over $\mathcal{V}_{n,t}$, $P(V \in \bar{\mathcal{R}}_i) \simeq 1/N$ for all i , and furthermore the shapes of partition cells will be approximately identical. For analytical tractability, we consider the following approximation for the partition cell.

$$\bar{\mathcal{R}}_i \simeq \tilde{\mathcal{R}}_i := \{V \in \mathcal{V}_{n,t} : \det(V^\dagger \hat{V}_i \hat{V}_i^\dagger V) \geq 1 - \delta_B\} \quad (23)$$

for some $\delta_B > 0$, which will be determined as a function of B using $P(V \in \bar{\mathcal{R}}_i) = 1/N$ for all i , i.e.,

$$P(V \in \bar{\mathcal{R}}_i) = \int_{1-\delta_B}^1 f_{\gamma_i}(x) dx = \frac{1}{2^B}. \quad (24)$$

With the partition cell approximation (23) and from the symmetrical property in the partition cells, we can use f_{γ_0} in the place of f_{γ_i} . Using the density function for γ_0 (e.g., one derived in Corollary 2), δ_B can be numerically calculated for a given B . Although, in general there are overlaps in the approximated partition cells, the analytical results from the approximation turn out to be quite accurate even when N is small. Since the approximated cells $\tilde{\mathcal{R}}_i$ have the identical geometrical shape, we can focus on a particular partition cell, arriving at the following result.

Approximate Density: With the partition cell approximation described in (23), the density function for $\gamma := \det(V^\dagger \hat{V} \hat{V}^\dagger V)$ is approximated by a *truncated* density of $f_{\gamma_0}(x)$, that is,

$$f_\gamma(x) \simeq \tilde{f}_\gamma(x) = 2^B f_{\gamma_0}(x) 1_{[1-\delta_B, 1]}(x) \quad (25)$$

where $1_A(x)$ is the indicator function having 1 if $x \in A$ and 0 otherwise.

C. Capacity Loss with Quantized Beamforming

Now we utilize the statistical results developed above to analyze the problem of interest, the capacity loss analysis (18). This can be written as $\tilde{C}_L = \int_0^1 [-\log x] f_\gamma(x) dx$. For the expectation, we will use the approximate density \tilde{f}_γ given in (25) instead of the real density f_γ . And the new approximate is denoted by \tilde{C}_L , that is,

$$\tilde{C}_L = \int_{1-\delta_B}^1 [-\log x] \tilde{f}_\gamma(x) dx. \quad (26)$$

For $m = 2$ (i.e., when $t \geq 2$ and $r = 2$), using the density f_{γ_0} in Corollary 2 and the definition of the hypergeometric function, a closed-form expression for the capacity loss can be derived.

$$\tilde{C}_L = 2^B \frac{\Gamma(t)\Gamma(t-1)}{\Gamma(2t-4)} \sum_{k=0}^{\infty} \frac{(t-2)_k (t-3)_k}{(2t-4)_k k! (k+2t-4)} \left[(1-\delta_B^{k+2t-4}) \log(1-\delta_B) + \sum_{l=1}^{k+2t-4} \frac{\delta_B^{k-l+2t-3}}{k-l+2t-3} \right] \quad (27)$$

where $(a)_k := a(a+1)\cdots(a+k-1)$.

For a general case of $m \geq 3$, one can obtain the approximate capacity loss (26) easily with an efficient Monte Carlo integration method which is described as follows: i) First generate a large number of samples for the random variable γ , each is just a product of independent beta distributed random variables as shown in Corollary 1. ii) Take a subset $S = \{\gamma > 1 - \delta_B : P(\gamma > 1 - \delta_B) = 1/2^B\}$. iii) Average over the subset S for $E_S[-\log \gamma]$, which is an estimate (26).

D. Numerical Results

Fig. 2 shows the capacity loss in bits per channel use for ($t = 4, r = 2$) MIMO channels when the beamforming matrix is represented with $B = 1, \dots, 8$. The analytical result is from \tilde{C}_L given in (27). The figure also includes simulation results using the beamforming-matrix codebooks. It shows the analytical result is close to the simulation result at high SNR (e.g., $P_T = 20, 10$ dB). However, at low SNR (e.g., $P_T = 0$ dB) it deviates from the simulation result. This is because in our analysis, with the high-SNR assumption, the effect of the term $\tilde{\Sigma}_H^2 = (I + \rho \Sigma_H^2)^{-1} \rho \Sigma_H^2$ is ignored, which is on average quite different from I at low SNR (P_T). Therefore, the high-SNR assumption results in higher values for capacity loss at low SNR.

V. MULTI-MODE SM TRANSMISSION STRATEGY

Equal power allocation on each of the parallel channels is clearly inefficient at lower SNR. To overcome this limitation, in this section, we present a multi-mode MIMO spatial multiplexing transmission scheme that allows for efficient utilization of the feedback bits. The transmission strategy is described as follows:

1. The number of data streams n is determined based on the average SNR: n changes from 1 to m , the rank of the channel, as SNR increases (see the example in Section V-B).

2. In each mode, the simple equal power allocation over n spatial channels is employed, thereby the entire feedback bits are utilized in representing only useful beamforming vectors without concerning about the power allocation over the spatial channels.

The average SNR is assumed to change at a much slower rate than the beamforming vectors, and so has to be updated less frequently consuming much fewer bits. Therefore it can be assumed to be negligible overhead. The effectiveness of the transmission strategy can be understood with a water-filling argument. The transmission strategy can be understood with a water-filling argument. That is, in low SNR region, only the principal spatial channel is useful in most of the time; and as the SNR increases, more spatial channels are getting involved. The proposed scheme can be viewed as a rough and indirect implementation of water-filling power allocation over multiple spatial channels. With the multi-mode transmission scheme, with perfect knowledge of $V(t \times n)$ and equal power allocation, we can achieve most of the capacity with perfect CSI at transmitter over all the range of SNR. For finite-rate feedback, only the relevant beamforming vectors are encoded using the entire feedback bits. From the quantization point of view, the multi-mode scheme increases the quantization resolution by reducing the dimension of the beamforming matrix in low SNR region.

In [7], another form of multi-mode SM scheme was studied in which the mode is adapted using current channel condition and the feedback bits are divided among multiple codebooks, one for each mode. For the scenario envisioned, this is less effective since it lowers the quantization resolution compared to the scheme mentioned above.

A. Codebook Design for Beamforming Matrix (When $n < m$)

Let us define submatrices $\Sigma_H = \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma_2 \end{bmatrix}$ and $V_H = [V \ V_2]$ where Σ ($n \times n$) and V ($t \times n$). Using an approach similar to that in Section III (for details refer to [12]), we have the following codebook design criterion:

$$\max_{\mathcal{Q}(\cdot)} E \left\| (V \tilde{\Sigma})^\dagger \mathcal{Q}(H) \right\|_F^2 \quad (28)$$

where $\hat{V} = \mathcal{Q}(H)$ is the $t \times n$ quantized beamforming matrix. Also as in Section III-C, we have two related design methods, each optimized for low and high SNR region.

B. Example

In Fig. 3, the performance of the multi-mode SM scheme with a finite number of feedback bits ($B = 8$) is shown for ($t = 6, r = 4$) MIMO in different modes. The number of data streams n is determined as follows: $n = 1$ when $P_T < P_{T,1}$, $n = 2$ when $P_{T,1} \leq P_T < P_{T,2}$, and $n = 3$ when $P_T \geq P_{T,2}$, where $P_{T,1}$ and $P_{T,2}$ are the boundary points (see the figure). To determine the mode, the transmitter needs to know the operating SNR of the system, which is assumed to be changing at a much slower rate than the channel itself in most time-varying channel environments. Therefore, only small additional feedback is necessary (e.g., to notify the transmitter to increase or decrease the number of data streams).

VI. SUMMARY

We have investigated the codebook design problem associated with transmit beamforming in MIMO spatial multiplexing systems with finite-rate feedback. Assuming a fixed number of spatial channels and equal power allocation, we designed the beamforming codebook by minimizing the capacity loss resulting from the finite-rate feedback. The capacity loss, under the assumption of a reasonably large number of feedback bits or low SNR, was suitably approximated leading to an iterative codebook design algorithm with monotonic convergence property. The developed design algorithm is based on the Lloyd algorithm in vector quantization study, but now has as its objective matrix quantization to minimize capacity loss. With the proposed method, we can design the optimum beamforming codebook for arbitrary number of transmit and receive antennas, feedback bits, and any spatial correlation structure in the channel. The effect on the MIMO channel capacity of finite-rate feedback was analyzed assuming high SNR and equal power allocation over the spatial channels. Central to the analysis is the complex multivariate beta distribution and simplifications of the Voronoi regions resulting from the codebook. To compensate for the degradation due to the equal power allocation assumption, we also proposed a multi-mode spatial multiplexing transmission strategy that allows for efficient utilization of the feedback bits by quantizing only relevant beamforming vectors.

REFERENCES

- [1] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.
- [2] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple antenna systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.
- [3] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple antenna systems with finite rate feedback: A VQ-based approach," *IEEE Trans. Inform. Theory*, submitted for publication.
- [4] D. J. Love, R. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.
- [5] K. N. Lau, T. A. Chen, and Y. Liu, "On the design of MIMO block-fading channels with feedback-link capacity constraint," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 62–70, Jan. 2004.
- [6] W. Santipach and M. L. Honig, "Asymptotic performance of MIMO wireless channels with limited feedback," in *Proc. IEEE MILCOM 2003*, Boston, MA, Oct. 2003.
- [7] D. J. Love and R. Heath, Jr., "Multi-mode precoding using linear receivers for limited feedback MIMO systems," in *Proc. IEEE ICC 2004*, Paris, France, June 2004.
- [8] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic, 1992.
- [9] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
- [10] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2450–2454, Sept. 2002.
- [11] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, 1982.
- [12] J. C. Roh and B. D. Rao, "Design and analysis of MIMO spatial multiplexing systems with quantized feedback," *IEEE Trans. Signal Processing*, submitted for publication.
- [13] A. M. Mathai and D. S. Tracy, "Random beta matrix and the distribution of its determinant," *Metron*, vol. 44, no. 1–4, pp. 101–110, 1986.

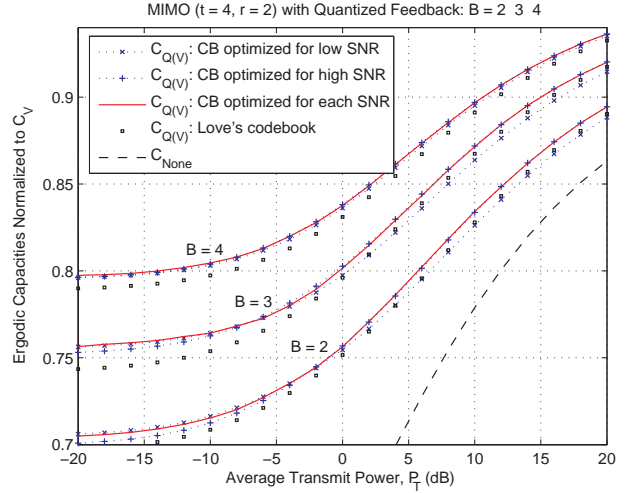


Fig. 1. Ergodic capacities for MIMO channels with quantized beamforming matrix for the different codebook design methods ($t = 4, r = 2, n = 2$, and $B = 2, 3, 4$).

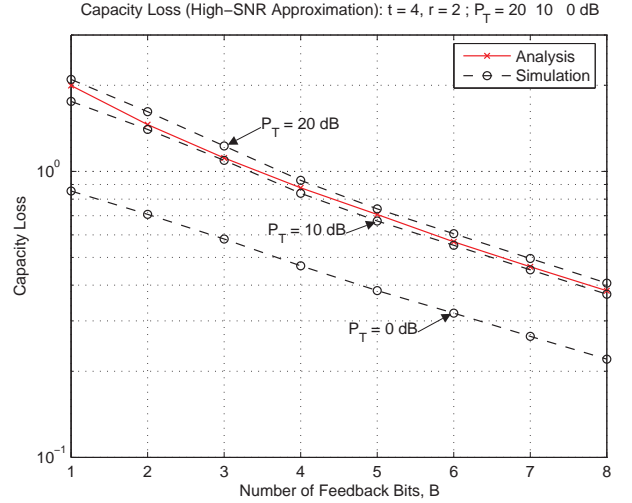


Fig. 2. Capacity loss of MIMO channel due to quantized beamforming matrix ($t = 4, r = 2; P_T = 20, 10, 0$ dB; $B = 1, 2, \dots, 8$).

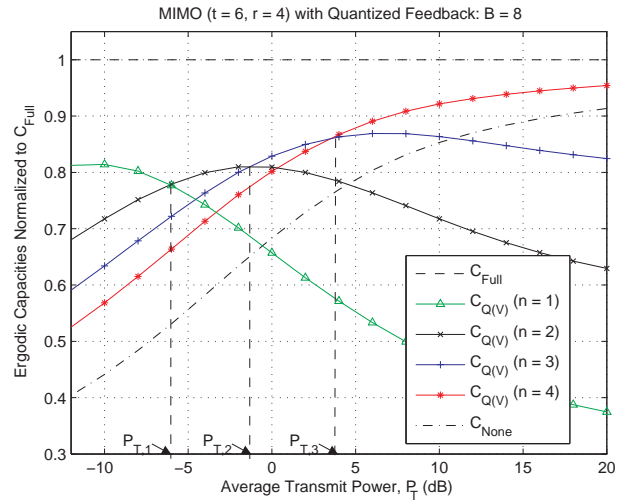


Fig. 3. Ergodic capacity of multi-mode MIMO spatial multiplexing systems with quantized beamforming matrix ($t = 6, r = 4, n = 1, \dots, 4; B = 8$).