Abstract — We consider a single-user, point-to-point communication system with $M$ transmit and $N$ receive antennas with independent flat Rayleigh fading between antenna pairs. The mutual information of the multi-input/multi-output (MIMO) channel is maximized when the transmitted symbol vector is a Gaussian random vector with covariance matrix $Q$. The optimal $Q$ depends on how much channel state information is available at the transmitter. Namely, in the absence of any channel state information, the optimal $Q$ is full-rank and isotropic, whereas with perfect channel knowledge, the optimal $Q$ has columns which are the eigenvectors of the channel, and has rank at most $\min\{M, N\}$.

We assume that the receiver can feed back $B$ bits to the transmitter (per codeword). The feedback bits are used to choose the columns of $Q$ from a random set of i.i.d. vectors. We compute the mutual information as a function of both $B$ and the rank of $Q$. Our results are asymptotic in the number of antennas, and show how much feedback is needed to achieve a rate, which is close to the capacity with perfect channel knowledge at the transmitter.

1. INTRODUCTION

Adding antennas to the transmitter and receiver of a single point-to-point fading link can substantially increase the achievable data rate [1, 2]. The channel capacity depends on the amount of channel state information (CSI) known at the transmitter and receiver. Assuming perfect CSI at the receiver, and independent fading across different transmitter-receiver antenna pairs, the channel capacity has been evaluated for the cases where there is no CSI and complete CSI at the transmitter.

Here we examine the capacity with complete CSI at the receiver, but partial CSI at the transmitter. We assume $M$ transmit antennas, $N$ receive antennas, and that the channels across different transmitter-receiver antenna pairs are i.i.d. Rayleigh. (The capacity with correlated fading across antenna pairs is studied in [3].) The mutual information is maximized when the transmitted symbol vector is Gaussian with covariance matrix $Q$, which depends on how much CSI is available at the transmitter. Specifically, in the absence of any CSI, the optimal $Q$ is full-rank and isotropic, whereas with perfect CSI the optimal $Q$ has rank at most $\min\{M, N\}$. We assume that the channel is stationary, and that the receiver relays $B$ bits to the transmitter (per codeword) via a reliable feedback channel (i.e., no feedback errors). The purpose of the feedback is to specify $Q$.

In this paper, we analyze the performance of a Random Vector Quantization (RVQ) scheme, which is used to specify a $Q$ with rank $D$, where $D$ can be optimized. The RVQ scheme is motivated by related work on signature optimization for CDMA with limited feedback, where it is shown to be optimal [4, 5]. Other related work on the performance of MIMO channels with limited feedback has been presented in [6–10]. Although our model is similar to that originally presented in [6], our asymptotic approach, based on RVQ, along with the associated results, differs substantially from prior work.

We compute the sum mutual information between the transmitted and received symbols per receive antenna as the number of transmit antennas $M$, the number of receive antennas $N$, the number of feedback bits $B$, and the rank $D$ all increase to infinity with fixed ratios. The limiting mutual information can be evaluated by using results from extreme order statistics and random matrix theory. We show how the optimal rank for the covariance matrix with RVQ decreases with the amount of feedback $B$, and observe that relatively little feedback is needed to achieve a rate close to the capacity with perfect CSI when $N/M$ is large. Optimizing the rank $D$ can also results in a substantial increase in the sum
The paper is organized as follows. In section 2, we describe the system model. In section 3, we review the behavior of the capacity without any CSI at the transmitter, and with perfect CSI. We present RVQ and the associated performance analysis in section 4 and the numerical results in section 5. Section 6 concludes the paper.

2. CHANNEL MODEL

We consider a single-user, point-to-point wireless channel with \( M \) transmit antennas and \( N \) receive antennas. We assume \( i.i.d. \) flat Rayleigh fading channels, so that the set of channel gains \( \{ h_{i,j} \} \) between the \( i \)th transmit antenna and the \( j \)th receive antenna contains \( i.i.d \) complex Gaussian random variables with zero mean and variance \( E[|h_{i,j}|^2] = 1 \). The received vector is given by

\[
y = Hx + w
\]

where \( H = [h_{i,j}] \) is the \( N \times M \) channel matrix, \( x = [x_i] \) is the \( M \times 1 \) transmitted symbol vector, \( y = [y_j] \) is the \( N \times 1 \) received symbol vector, and the noise \( w \) is complex Gaussian with covariance matrix \( \sigma_n^2 I_N \) where \( I_N \) is the \( N \times N \) identity matrix.

Our performance metric is sum mutual information between \( x \) and \( y \). The mutual information is maximized when the transmitted symbols \( x_i \) are complex Gaussian random variables, in which case the sum mutual information is given by

\[
I(x; y) = E_{H} \left[ \log \det \left( I_N + \rho H \Sigma H^\dagger \right) \right] \tag{1}
\]

where the covariance matrix \( \Sigma = E[xx^\dagger] \), \( \rho = E[\|x\|^2] / \sigma_n^2 \) is the background Signal-to-Noise Ratio (SNR), and \( \log \) denotes natural logarithm.

We are now interested in maximizing the sum mutual information over the covariance matrix \( \Sigma \) subject to a power constraint \( \text{tr} \{ \Sigma \} \leq 1 \).

3. PERFECT CSI VS. NO CSI

Here we consider the two extreme cases in which the transmitter has no CSI available, and the transmitter has perfect CSI. These correspond to normalized feedback values \( \bar{B} = B/N^2 = 0 \) and \( \bar{B} = \infty \), respectively.

3.1. Zero Feedback

If there is no CSI available at the transmitter (corresponding to zero feedback bits from the receiver), then the optimal \( \Sigma = \frac{\rho}{M} I_M \) \cite{2}. That is, each antenna transmits an independent symbol, and the transmitted power is allocated equally across transmit antenna. As \( (N, M) \to \infty \) with fixed \( N = N/M \), the capacity per receive antenna is given by

\[
\frac{1}{N} \det \log \left( I_N + \frac{\rho}{M} HH^\dagger \right) \to \int_0^\infty \log \left( 1 + \rho \lambda \right) g(\lambda) \, d\lambda
\]

\[
= \mathbb{C}^\infty(\bar{B} = 0)
\]

where convergence is almost surely, and \( g(\lambda) \) is the asymptotic probability density function for a randomly chosen eigenvalue of \( HH^\dagger \), and is given in \cite{11}. The integral in (2) has been evaluated in \cite{12}, and gives the following closed-form expression

\[
\mathbb{C}^\infty(\bar{B} = 0) = \log \rho y + \frac{1 - \bar{N}}{\bar{N}} \log \left( \frac{1}{1 - y} \right) - \frac{\bar{z}}{\bar{N}} \tag{3}
\]

where

\[
y = \frac{1}{2} \left( 1 + \bar{N} + \frac{1}{\rho} + \sqrt{1 + \bar{N} + \frac{1}{\rho}} \right)^{-4 \bar{N}} \frac{\bar{z}}{\bar{N}} \tag{5}
\]

\[
z = \frac{1}{2} \left( 1 + \bar{N} + \frac{1}{\rho} - \sqrt{1 + \bar{N} + \frac{1}{\rho}} \right)^{-4 \bar{N}} \tag{6}
\]

3.2. Infinite Feedback

If \( \bar{B} = \infty \), then the transmitter has perfect CSI, and the optimal \( \Sigma \) is given by

\[
\mathbb{D} = \mathbb{F} \mathbb{D} \mathbb{F}^\dagger \tag{4}
\]

where \( \mathbb{U} \) is an unitary matrix whose columns are eigenvectors of \( HH^\dagger \), and \( \mathbb{D} \) is a diagonal matrix whose elements are computed by water-filling over the available dimensions. The rank of the optimal \( \mathbb{D} \) is at most \( \min\{M, N\} \), and depends on the SNR. In contrast, with \( \bar{B} = 0 \) the optimal \( \mathbb{D} \) is full-rank.

Let \( \mathbb{C}^\infty(\bar{B} = \infty) \) denote the capacity per receive antenna in the limit as \( (N, M) \to \infty \). We have that

\[
\mathbb{C}^\infty(\bar{B} = \infty) = \int_0^\infty (\log \gamma)^+ g(\lambda) \, d\lambda
\]

where the water level \( \gamma \) is determined by the power constraint

\[
\bar{N} \int_0^\infty \left( \gamma - \frac{1}{\lambda} \right)^+ g(\lambda) \, d\lambda = \rho
\]

where \( x^+ = \max\{0, x\} \). We can verify that \( \mathbb{C}^\infty(\bar{B} = 0) \leq \mathbb{C}^\infty(\bar{B} = \infty) \), and now consider the ratio \( \mathbb{C}^\infty(\bar{B} = 0) / \mathbb{C}^\infty(\bar{B} = \infty) \).
is able to decode the symbol whether the power is allocated
to maximize the sum mutual information,
where the columns of contains more independent copies of a transmitted symbol and
power allocation asymptotically gives the same capacity as water pouring.

For small SNR, we have that
\[
\lim_{\rho \to 0} \frac{C^\infty(\tilde{B} = \infty)}{C^\infty(\tilde{B} = 0)} = 1 \quad \text{for } \tilde{N} > 0.
\]

Hence for large SNR, there is clearly little incentive to relay CSI back to the transmitter. This is because for large SNR, the transmitter uses all available dimensions, and an equal power allocation asymptotically gives the same capacity as water pouring.

For small SNR, we have that
\[
\lim_{\rho \to 0} \frac{C^\infty(\tilde{B} = \infty)}{C^\infty(\tilde{B} = 0)} = \left(1 + \frac{1}{\sqrt{\tilde{N}}} \right)^2 \quad \text{for } \tilde{N} > 0.
\] (7)

This limit is an extension of the analogous limit in [3], which is for \( \tilde{N} = 1 \). Hence the potential gain due to feedback decreases with \( \tilde{N} \). This is because for large \( \tilde{N} \), the receiver obtains more independent copies of a transmitted symbol and is able to decode the symbol whether the power is allocated equally or optimally.

4. RANDOM VECTOR QUANTIZATION

To maximize the sum mutual information, \( Q \) is given by (4) where the columns of \( \mathbf{U} \) are the eigenvectors of \( \mathbf{H}^H \mathbf{H} \). If \( Q \) has rank \( D < \min\{M, N\} \), then the eigenvectors correspond to the \( D \) largest eigenvalues. In what follows, we assume that the receiver quantizes the matrix \( \mathbf{U} \), and relays this back to the transmitter. We also assume that the transmitter distributes equal power and rate over the \( D \) eigenvectors, i.e., the optimal water pouring distribution is not relayed back to the transmitter. (The degradation in capacity due to using a constant power allocation with optimized rank \( D \), rather than the optimal water pouring power allocation has been shown to be small [13].)

Given \( B \) feedback bits, we can construct a vector quantizer, which chooses the matrix \( \mathbf{U} \) from the set, or codebook \( \{\mathbf{V}_1, \cdots, \mathbf{V}_{2^B}\} \). That is, assuming that \( \mathbf{H} \) is known at the receiver, the receiver chooses the \( \mathbf{V}_j \) that maximizes the sum mutual information, and relays the corresponding index back to the transmitter. Optimizing this codebook for finite \( M \) and \( N \) is quite difficult; however, as \( (M, N) \to \infty \), the eigenvectors of \( \mathbf{H} \) are isotropically distributed. This suggests that the columns of the codebook entries should also be isotropically distributed. Hence in what follows, we assume that the matrices \( \{\mathbf{V}_j\}, j = 1, \cdots, 2^B \) are independent and isotropic, and refer to this scheme as Random Vector Quantization (RVQ). An analogous RVQ scheme has been previously proposed for CDMA signature optimization with limited feedback in [4,5], and has been shown to maximize the SINR in the large system limit.

We are interested in the mutual information with RVQ in the limit as \( (N, M, B, D) \to \infty \) with fixed ratios \( \tilde{N} = N/M, \tilde{B} = B/N^2, \) and \( \tilde{D} = D/M \). (Note that the number of feedback bits is normalized by the number of matrix entries.) Given the rank \( D \leq M \), the covariance matrix is chosen from the set \( \{Q_j\}, j = 1, \cdots, 2^B \), where
\[
Q_j = \frac{1}{\tilde{D}} \mathbf{V}_j \mathbf{V}_j^H \quad \text{for } 1 \leq j \leq 2^B
\]
and \( \mathbf{V}_j \) is an \( M \times D \) random unitary matrix, i.e., \( \mathbf{V}_j^H \mathbf{V}_j = \mathbf{I}_D \). The matrices \( \mathbf{V}_j \) for all \( j \) are chosen independently, and the transmitted power \( \text{tr}\{Q_j\} = 1 \). The receiver selects the quantized covariance matrix
\[
\hat{Q} = \arg \max \limits_{1 \leq j \leq 2^B} \left\{ J_N^j = \frac{1}{N} \log \det \left( \mathbf{I}_N + \rho \mathbf{Q}_j \mathbf{H}^H \right) \right\}.
\]

Hence the sum mutual information per receive antenna with \( B \) feedback bits is given by
\[
I_{\text{rvq}}^N = E \left[ \max \limits_{1 \leq j \leq 2^B} J_N^j \right] = E \left[ \frac{1}{N} \log \det \left( \mathbf{I}_N + \rho \mathbf{Q}_j \mathbf{H}^H \right) \right]
\] (8)

where the expectation is over the channel and covariance matrix. To evaluate (8), we must know the cumulative distribution function (cdf) for \( J_N^j \). Since the matrices \( Q_j \) are i.i.d., the random variables \( \{J_N^j\} \) are also i.i.d.. Letting \( F_N \) denote the corresponding cdf, we have that
\[
I_{\text{rvq}}^N = \tilde{B} \int x F_N^{B-1}(x) f_N(x) \, dx
\]
where \( f_N(x) \) is the probability density function (pdf) for \( J_N^j \). In what follows we consider the limit as \( N, M, D, \tilde{B} \) all tend to infinity.

**Theorem 1** As \( (N, M, B, D) \to \infty \) with fixed \( \tilde{N} = N/M > 0, \tilde{B} = B/N^2, \) and \( \tilde{D} = D/M \), the sum mutual information per receive antenna \( I_{\text{rvq}}^N \) converges almost surely to
\[
I_{\text{rvq}}^\infty = \lim \limits_{(N,M,D,B) \to \infty} F_N^{-1} \left( 1 - \frac{1}{2^B} \right).
\] (9)

3 of 6
The inverse cdf $F_N^{-1}$ depends on $N$, $M$, and $D$. The proof relies on the asymptotic theory of extreme order statistics [14]. With appropriate shifting and scaling [14], $F_N$ converges in distribution to a Weibull cdf as $B \to \infty$. An analogous result in the context of CDMA signature optimization has been presented in [4].

Computing the exact cdf $F_N$ appears to be difficult. We therefore resort to a Gaussian approximation, which is motivated by the following lemma. In what follows, we drop the superscript in $J_N^*$ to simplify the notation.

**Lemma 1** As $(N, M, D) \to \infty$ with $N/M = \tilde{N}$ and $D/M = \tilde{D}$,

$$N(J_N - \mu) \to N(0, \sigma^2) \quad \text{in distribution}$$

where $N(0, \sigma^2)$ is the Gaussian cdf with zero mean and variance $\sigma^2$,

$$\mu = \frac{\tilde{D}}{N} \log \left( 1 + \frac{\tilde{N}}{\tilde{D} \rho} - \frac{\tilde{N}}{\tilde{D} \rho} \right) + \log \left( 1 + \rho - \rho v \right) - v$$

and

$$\sigma^2 = -\log \left( 1 - \frac{\tilde{N}}{\tilde{D} v^2} \right)$$

where

$$v = \frac{1}{2} + \frac{\tilde{D}}{2N} + \frac{\tilde{D}}{2N \rho} - \frac{1}{2} \sqrt{\left( 1 + \frac{\tilde{N}}{\tilde{D}} + \frac{\tilde{D}}{\tilde{N} \rho} \right)^2 - \frac{4\tilde{D}}{\tilde{N} \rho}}.$$ 

This lemma is an extension of a similar result in [15] with $\tilde{D} = 1$, and follows from Theorem 1.1 in [16]. To compute $\mu$ and $\sigma^2$, we write

$$J_N = \frac{1}{N} \sum_{k=1}^{N} \log \left( 1 + \rho \frac{\tilde{N}}{\tilde{D}} \lambda_k \right)$$

where $\lambda_k$ is the $k$th eigenvalue of $\frac{1}{\tilde{N}} \text{HVV}^\dagger \text{H}$. As $(N, M, D) \to \infty$, the empirical eigenvalue distribution converges to a deterministic function, which can be represented in terms of its Stieltjes transform [16]. The asymptotic mean and variance of $J_N$ can then be evaluated according to results in [16].

Although Lemma 1 states that the asymptotic distribution of $J_N$ is Gaussian, this assumption leads to an approximation for the asymptotic sum mutual information because the distribution for finite $N$ is still needed to compute $\tilde{I}_{\text{rvq}}^\infty$ in Theorem 1. Specifically, using the Gaussian assumption in Theorem 1 gives

$$\tilde{I}_{\text{rvq}}^\infty = \mu + \sigma \sqrt{2\tilde{B} \log 2} \quad (10)$$

which approximates $I_{\text{rvq}}^\infty$. As $\tilde{B} \to 0$, this approximation becomes exact. However, as $\tilde{B} \to \infty$, $\tilde{I}_{\text{rvq}}^\infty \to \infty$, whereas $C^\infty(\tilde{B} = \infty)$ is finite, and can be computed by (5) and (6). Therefore the Gaussian assumption gives an inaccurate estimate of $I_{\text{rvq}}^\infty$ for large $\tilde{B}$. This is because $J_N$ is bounded for all $N$, whereas the Gaussian pdf assumes that $J_N$ can assume arbitrarily large values.

Note that $I_{\text{rvq}}^\infty$ is a function of both the rank $\tilde{D}$ and feedback $\tilde{B}$. For a given $\tilde{B}$, we wish to select the $\tilde{D}$, which maximizes $I_{\text{rvq}}^\infty$. The optimal $\tilde{D}$ can be found by differentiating $I_{\text{rvq}}^\infty$, and depends on $\tilde{N}$, $\tilde{B}$, and $\rho$.

5. NUMERICAL RESULTS

Figure 1 shows $\tilde{I}_{\text{rvq}}^\infty$ with optimal rank $\tilde{D}_{\text{rvq}}$ versus $\tilde{B}$ with different values of $\rho$ and $\tilde{N} = 0.5$. The dashed lines show the water-filling capacity $C^\infty(\tilde{B} = \infty)$. As expected, the performance of the optimal power allocation, relative to equal power allocation, decreases as SNR increases. Also, the amount of feedback needed to reach $C^\infty(\tilde{B} = \infty)$ decreases with SNR. The dashed-dotted lines show $\tilde{I}_{\text{rvq}}^\infty$ with fixed rank $\tilde{D} = 1$ and different values of $\rho$. In this case, the rates with the optimized rank and full-rank do not differ much. Also shown in the plot are simulation results $I_{\text{rvq}}^\infty$ with $\tilde{D} = 1$. We can see that $\tilde{I}_{\text{rvq}}^\infty$ is a good estimate even for small $M$ (e.g., 8).

Figure 1: Sum mutual information per receive antenna for RVQ with optimized rank and full-rank versus normalized feedback.

Figure 2 shows the optimal RVQ rank $\tilde{D}_{\text{rvq}}^*$ versus $\tilde{B}$ for different values of $\rho$ and $\tilde{N} = 0.5$. The rank starts at $\tilde{D}_{\text{rvq}}^* = 1$ (full-rank) for $\tilde{B} = 0$, and decreases with $\tilde{B}$. The rank con-
verges to the value associated with the optimal covariance matrix with perfect CSI. For given \( \tilde{B} \), the optimal rank increases with SNR, reflecting the fact that the water pouring distribution is spread across more eigenvectors. In the figure, the optimal rank converges to \( \tilde{N} \) for \( \rho = 9 \) and \( 6 \) dB, but converges to a lower number for \( \rho = 3 \) and \( 0 \) dB.

Figure 2: The optimal RVQ rank versus normalized feedback.

Figure 3 shows the mutual information per receive antenna versus normalized feedback with different values of \( \tilde{N} \). The water-filling capacity is also shown. As \( \tilde{N} \) decreases, the performance gap associated with \( \tilde{B} = 0 \) and \( \tilde{B} = \infty \) widens, which is consistent with (7). Also, the sum mutual information per receive antenna decreases with \( \tilde{N} \).

Figure 3: Sum mutual information per receive antenna versus normalized feedback with different values of \( \tilde{N} \).

Figure 4 shows mutual information per receive antenna versus normalized rank from (10) with \( \tilde{N} = 0.1, \rho = 6 \) dB, and different values of \( \tilde{B} \). The maximal rates are attained at \( \tilde{B} = 0.1 \) for cases shown. As \( \tilde{B} \) increases, the rate increases and the difference between the rate with optimized rank and full-rank also increases.

Figure 4: Mutual information per receive antenna versus normalized rank with different feedback.

6. CONCLUSIONS

We have studied the achievable rate of a single-user MIMO fading channel with limited feedback. Our results are asymptotic in the number of antennas, and indicate the minimum feedback, which is needed to achieve the capacity associated with perfect CSI at the transmitter. The asymptotic mutual information was evaluated with a random vector quantization scheme in which the covariance matrices in the source codebook are independent and isotropic. This approach allows us to study the rank of the covariance matrix which maximizes the mutual information. Our numerical results show that less feedback is needed as the ratio between the number of receive antennas and transmit antennas increases. For the cases shown, the covariance with the optimized rank gives a large increase in capacity over the one with full-rank (e.g., 50-60%).

REFERENCES


