

Vector Channel Capacity with Quantized Feedback

Sudhir Srinivasa and Syed Ali Jafar
Electrical Engineering and Computer Science
University of California Irvine, Irvine, CA 92697-2625
Email: syed@ece.uci.edu, sudhirs@uci.edu

Abstract— We explore the capacity of an isotropic fading vector channel with multiple transmit antennas at the base station and a single antenna at the mobile receiver. Perfect channel knowledge is assumed to be available at the receiver while the transmitter has only partial knowledge of the direction of the user's channel vector based on quantized feedback. We determine a necessary and sufficient condition for optimality of beamforming, which turns out to be a precise relationship between the optimality of beamforming and the symmetry of the quantizer. For the cases where beamforming is not optimal, we develop upper and lower bounds. Numerical results are provided to help estimate the capacity.

I. INTRODUCTION

The amount of channel state information at the transmitter (CSIT) is a decisive factor governing data rates of wireless links with multiple antennas at the transmitter. This is especially true for typical cellular downlink scenarios where the base station can accommodate multiple (M) transmit antennas while the size-constrained mobile unit is restricted to one receive antenna. However, perfect channel knowledge at the transmitter (CSIT) is not practical, owing to the large number of time-varying channel coefficients in vector channels that have to be fed back to the transmitter. Perfect beamforming is therefore not feasible for most wireless channels. As a result, recent research efforts have focused on the performance of multiple antenna channels under limited feedback. A partial feedback scheme may be designed to convey some information about the “magnitude” and/or the “direction” of the channel to the transmitter. The knowledge of only the user's channel vector direction at the transmitter is much more useful than the knowledge of only the channel magnitude [1] because direction feedback allows the transmitter to direct its transmit power along the channel. Therefore at least partial knowledge about the channel direction is necessary for efficient use of the channel. Of late, partial direction feedback systems have attracted much research activity [2]–[7]. Especially relevant to this paper are [2]–[4], which concentrate on the performance of systems where a quantized version of the channel direction is fed back to the transmitter.

Previous work [2]–[4] on quantized feedback schemes assume beamforming as the transmit strategy. However, it is not known whether beamforming is the optimal (in the capacity sense) strategy with quantized direction feedback. This paper is motivated by a need to determine the capacity of these systems. In particular we seek to determine whether beamforming along the quantized direction is optimal for the capacity of the quantized direction feedback channel.

II. SYSTEM AND CHANNEL MODEL

We consider a point to point communication system consisting of a transmitter with M transmit antennas and a single antenna receiver. The input/output relationship is given by $Y = \langle \mathbf{H}, \mathbf{X} \rangle + Z = \mathbf{H}^H \mathbf{X} + Z$, where $\mathbf{X} \in \mathbb{C}^M$ is the $(M \times 1)$ transmitted vector of complex symbols, $Y \in \mathbb{C}$ is the received scalar signal, and $Z \sim \mathcal{CN}(0, 1)$ is complex additive white Gaussian noise (AWGN). The $M \times 1$ complex channel vector $\mathbf{H} = [H_1 \ H_2 \ \cdots \ H_M]^T$ represents a narrowband isotropic fading channel [1]. H_m is the channel gain from the m^{th} transmit antenna to the receiver. It is assumed that there is perfect channel state information at the receiver (CSIR). The receiver sends channel information to the transmitter by means of an error-free, limited capacity, zero-delay feedback channel. The limited data rate feedback allows B bits of quantized channel state information at the transmitter.

We assume that the B feedback bits are used to quantize the channel vector direction. A predetermined set $\mathcal{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N\}$ of $N = 2^B$ unit norm quantization vectors is available to the transmitter and receiver. We will assume that $M \leq N$ throughout this paper with the exception of Section V where the case $M > N$ is examined. Through the feedback channel the receiver sends only the index k of the quantization vector \mathbf{q}_k which is ‘closest’ to the channel vector \mathbf{H} . The transmitter uses the quantized channel information to adapt its transmit strategy. We use the following definition of distance between any two vectors: $d(\mathbf{v}_1, \mathbf{v}_2) \triangleq \sqrt{1 - \left| \left\langle \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} \right\rangle \right|^2} = \sin(\theta_{1,2})$. Given \mathcal{Q} , we define $d_{\min}(\mathcal{Q})$ as the minimum distance between any two vectors in \mathcal{Q} as $d_{\min}(\mathcal{Q}) = \min_{\mathbf{q}_i, \mathbf{q}_j \in \mathcal{Q}, i \neq j} d(\mathbf{q}_i, \mathbf{q}_j)$. Similar to [2] and [4], we assume that the quantization set \mathcal{Q} is designed such that the distance between the two closest vectors in the set is maximized, i.e., the design criterion for the quantization set is $\max_{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N} \min_{1 \leq i < j \leq N} d(\mathbf{q}_i, \mathbf{q}_j)$. We define the ‘decision region’ D_i for \mathbf{q}_i as

$$D_i \triangleq \{\mathbf{H} : d(\mathbf{H}, \mathbf{q}_i) < d(\mathbf{H}, \mathbf{q}_j) \ \forall j \neq i, 1 \leq j \leq N\}. \quad (1)$$

The name decision region follows from digital modulation nomenclature, if the channel $\mathbf{H} \in D_i$, the receiver feeds back a ‘decoded’ index i . We define the covariance matrix of the input as $K_i \triangleq \mathbb{E}[\mathbf{X}\mathbf{X}^H | \mathbf{H} \in D_i]$ and state our transmit power constraint as

$$\text{Tr}(K_i) \leq P, \quad \forall i \in \{1, 2, \dots, N\}. \quad (2)$$

Notice that the transmit power is the same for all decision regions. While this power constraint makes the analysis simpler by ruling out power adaptation at the transmitter, the results of our analysis are applicable even when the transmitter can choose a different transmit power P_i for each decision region D_i . Allowing power adaptation simply introduces another optimization problem on top of our problem statement, where the optimal transmit powers P_i^* for each decision region need to be determined. Ignoring power adaptation is also motivated by several practical considerations as discussed in [8]. So, in this paper we assume the power constraint of (2).

III. CAPACITY EXPRESSION AND PROBLEM STATEMENT

In this section, we will consider the capacity optimization problem for the quantized direction feedback system presented in Section II. We introduce the following definitions.

- 1) Define \mathcal{K} as the set of input covariance matrices given by $\mathcal{K} \triangleq \{K_1, K_2, \dots, K_N\}$.
- 2) Define the set of decision regions by $\mathcal{D} \triangleq \{D_1, D_2, \dots, D_N\}$.
- 3) Define $C(K, S)$ as

$$C(K, S) \triangleq \mathbb{E}_{\mathbf{H} \in S} [\log(1 + \mathbf{H}^\dagger K \mathbf{H})] \quad (3)$$

The following theorem describes the capacity of the quantized direction feedback system:

Theorem 1: *The capacity C_{QF} of the quantized feedback system is*

$$C_{QF} = \max_{\mathcal{K}: \text{Tr}(K_i) \leq P} \sum_{i=1}^N C(K_i, D_i) \text{Prob}(\mathbf{H} \in D_i) \quad (4)$$

The capacity is achieved by transmitting independent complex circular Gaussian symbols along the eigenvectors of K_i when $\mathbf{H} \in D_i$. The powers in the Gaussian symbols are the corresponding eigenvalues of K_i .

Proof: Since the channel knowledge at the transmitter ($\mathbf{H} \in D_i$) is a deterministic function of the CSIR (\mathbf{H}), it follows from Proposition 1 of [9] that the capacity of the vector channel with direction feedback described in Section II is

$$C_{QF} = \sum_{i=1}^N \text{Prob}(\mathbf{H} \in D_i) \max_{p(\mathbf{X}|\mathbf{H} \in D_i) \in \mathcal{P}} \mathcal{I}(\mathbf{X}; Y | \mathbf{H}, \mathbf{H} \in D_i) \quad (5)$$

where \mathcal{P} is the set of input distributions $p(\mathbf{X}|\mathbf{H} \in D_i)$ that satisfy the power constraint (2). Notice that for each i in $\{1, 2, \dots, N\}$, the capacity computation involves the following maximization.

$$\begin{aligned} & \max_{p(\mathbf{X}|\mathbf{H} \in D_i) \in \mathcal{P}} \mathcal{I}(\mathbf{X}; Y | \mathbf{H}, \mathbf{H} \in D_i) \\ &= \max_{P_i(\mathbf{X}) \in \mathcal{P}} \mathbb{E}_{\mathbf{H} \in D_i} \mathcal{I}(\mathbf{X}; Y | \mathbf{H} = \mathbb{H}) \end{aligned} \quad (6)$$

$$\begin{aligned} &= \max_{P_i(\mathbf{X}) \in \mathcal{P}} \mathbb{E}_{\mathbf{H} \in D_i} [\mathcal{H}(Y | \mathbf{H} = \mathbb{H}) - \mathcal{H}(Y | \mathbf{X}, \mathbf{H} = \mathbb{H})] \\ &= \max_{K_i: \text{Tr}(K_i) \leq P} \mathbb{E}_{\mathbf{H} \in D_i} \log(1 + \mathbb{H}^\dagger K_i \mathbb{H}) \end{aligned} \quad (7)$$

$$= \max_{K_i: \text{Tr}(K_i) \leq P} C(K_i, D_i) \quad (8)$$

where $P_i(\mathbf{X}) = p(\mathbf{X}|\mathbf{H} \in D_i)$. (7) results from the following observations. $\mathcal{H}(Y | \mathbf{X}, \mathbf{H} = \mathbb{H}) = \mathcal{H}(Z)$. Given that the input covariance matrix of \mathbf{X} is K_i and the channel is \mathbb{H} , the covariance of Y can be expressed as $\mathbb{H}^\dagger K_i \mathbb{H}$. Given the covariance matrix of Y the entropy is maximized when Y is a circularly symmetric complex Gaussian, which happens when \mathbf{X} is also a circularly symmetric complex Gaussian. Thus (8) follows and capacity is achieved by transmitting independent complex circular Gaussian symbols along the eigenvectors of the covariance matrix of \mathbf{X} , K_i . The only remaining optimization is to find the optimal K_i . Combining (5) and (8) we have the result of Theorem 1. ■

We notice from (4) that the capacity computation requires an optimization over the set of input covariance matrices \mathcal{K} . The following corollary states that the optimization can be broken down into individual optimizations over each K_i .

Corollary 1: *The capacity optimization problem (4) can be solved as N separate optimizations*

$$\max_{K_i: \text{Tr}(K_i) \leq P} C(K_i, D_i), \quad i \in \{1, 2, \dots, N\}. \quad (9)$$

While the optimal transmit strategy is given by the solution to the optimization problems (4) (9), we are specifically interested in the beamforming strategy.

A. Beamforming

For the beamforming strategy, whenever $\mathbf{H} \in D_i$, the transmit vector $\mathbf{X} = \mathbf{q}_i x$, where x is a scalar codeword symbol and \mathbf{q}_i , the quantization vector, is also the beamforming vector. In that case the input covariance matrix is constrained as $K_i = P \mathbf{q}_i \mathbf{q}_i^\dagger$. The capacity with this additional constraint is called the *beamforming capacity* and is given by

$$C_{BF} = \sum_{i=1}^N C(P \mathbf{q}_i \mathbf{q}_i^\dagger, D_i) \text{Prob}(\mathbf{H} \in D_i) \quad (10)$$

Using the definitions and expressions obtained above, next we state the problem definitions.

B. Problem Definition

Corollary 1 states that the N optimization problems in (9) can be solved separately. Without loss of generality we will consider the following case:

- 1) $i = 1$, i.e., $\mathbf{H} \in D_1$.
- 2) We choose a co-ordinate system such that $\mathbf{q}_1 = [1 \ 0 \ \dots \ 0]^T$.

This leads to the following problem statements.

- Determine the optimal input covariance matrix K_1 to maximize $C(K_1, D_1)$.
- When is beamforming the optimal transmit strategy?

IV. OPTIMALITY OF BEAMFORMING

For the problem definition of Section III-B, we start with a mathematical statement of the necessary and sufficient condition for optimality of beamforming in the following Theorem. The significance of these mathematical results will be explained in the following sections.

Theorem 2: Beamforming along \mathbf{q}_1 is the optimal transmit strategy if and only if $\lambda_{\max}(\Lambda_1) \leq \Lambda_1[1, 1]$, where we define

$$\Lambda_1 \triangleq E_{\mathbf{H} \in D_1} \left[\frac{\mathbf{H}\mathbf{H}^\dagger}{1 + P|H_1|^2} \right] \quad (11)$$

and $\Lambda_1[1, 1]$ is the element in the first row and first column of Λ_1 .

Proof: We start with the following Lemma:

Lemma 1: Given an $(M \times 1)$ unit vector \mathbf{q} and any $(M \times M)$ positive semidefinite matrix K , there exists a positive semidefinite matrix K' with $\text{rank}(K') \leq (M - 1)$ such that K can be expressed as $K = \rho \mathbf{q}\mathbf{q}^\dagger + K'$ for some value of $\rho \geq 0$.

Proof: Consider any arbitrary positive semidefinite matrix K of size $(M \times M)$. When K is singular, i.e., $\text{rank}(K) \leq (M - 1)$, then we set $\rho = 0$, $K' = K$ and Lemma 1 is trivially true. Therefore we assume that K is non-singular, and $\text{rank}(K) = M$. For any unit vector \mathbf{u} , we have $\mathbf{u}K\mathbf{u}^\dagger \geq \lambda_{\min}(K) > 0$ where $\lambda_{\min}(K)$ is the smallest eigenvalue of K . Since \mathbf{q} is a unit vector, the $M \times M$ matrix $\mathbf{q}\mathbf{q}^\dagger$ is positive-semidefinite with unit rank and the only eigenvalue is unity. Therefore, for any unit vector \mathbf{u} , we have $\mathbf{u}^\dagger(\rho \mathbf{q}\mathbf{q}^\dagger)\mathbf{u} \leq \rho$. Consequently, for all unit vectors \mathbf{u} ,

$$\mathbf{u}^\dagger(K - \rho \mathbf{q}\mathbf{q}^\dagger)\mathbf{u} \geq \lambda_{\min}(K) - \rho > 0 \quad (12)$$

for all $\rho < \lambda_{\min}(K)$. Thus the matrix $K - \rho \mathbf{q}\mathbf{q}^\dagger$ is a positive definite matrix for $\rho < \lambda_{\min}(K)$. However, as $\rho \rightarrow \infty$ $\mathbf{u}^\dagger(K - \rho \mathbf{q}\mathbf{q}^\dagger)\mathbf{u}$ is negative for $\mathbf{u} = \mathbf{q}$. Choose the smallest value of ρ , say ρ^* , for which $K - \rho \mathbf{q}\mathbf{q}^\dagger$ becomes singular. Then defining $K - \rho^* \mathbf{q}\mathbf{q}^\dagger$ to be our K' we have $\text{rank}(K') \leq (M - 1)$ and $K = \rho \mathbf{q}\mathbf{q}^\dagger + K'$. ■

According to Lemma 1 any input covariance matrix K_1 can be represented as $K_1 = (P - p)\mathbf{q}_1\mathbf{q}_1^\dagger + pK'$, where $p \in [0, 1]$ and $\text{rank}(K') \leq (M - 1)$. Since $\text{Tr}(K_1) = P$ and $\text{Tr}(\mathbf{q}_1\mathbf{q}_1^\dagger) = 1$, we have $\text{Tr}(K') = 1$. The optimality of beamforming is then equivalent to the following condition.

Lemma 2: Beamforming is the capacity achieving strategy if and only if, for all $M \times M$ positive semidefinite matrices K' with $\text{rank}(K') < M$ and $\text{Tr}(K') = 1$, the function

$$f(p) \triangleq C\left((P - p)\mathbf{q}_1\mathbf{q}_1^\dagger + pK', D_1\right) \quad (13)$$

defined for $p \in [0, 1]$, takes its maximum value at $p = 0$.

The proof of Lemma 2 is straightforward because $p = 0$ corresponds to beamforming along \mathbf{q}_1 [8]. Since $\text{rank}(K') \leq (M - 1)$, let the eigenvalues of K' be $\alpha_1, \alpha_2, \dots, \alpha_{M-1}$ and let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{M-1}$ be the corresponding unit norm eigenvectors. Clearly, the α_i are all nonnegative and sum to unity. Recall that $\mathbf{q}_1 = [1 \ 0 \ \dots \ 0]^T$. From Lemma 2 a necessary condition for optimality of beamforming is $\frac{d}{dp}f(p)|_{p=0} \leq 0$. This is also a sufficient condition because $f(p)$ is a concave function in p , as seen from the second derivative.

$$\frac{d}{dp}f(p) = E_{\mathbf{H} \in D_1} \frac{\sum_{j=1}^{M-1} \alpha_j \mathbf{H}^\dagger \mathbf{u}_j \mathbf{u}_j^\dagger \mathbf{H} - |H_1|^2}{1 + (P - p)|H_1|^2 + p \sum_{j=1}^{M-1} \alpha_j \mathbf{H}^\dagger \mathbf{u}_j \mathbf{u}_j^\dagger \mathbf{H}}$$

$$\frac{d^2}{dp^2}f(p) = -E_{\mathbf{H} \in D_1} \left[\frac{\sum_{j=1}^{M-1} \alpha_j \mathbf{H}^\dagger \mathbf{u}_j \mathbf{u}_j^\dagger \mathbf{H} - |H_1|^2}{1 + (P - p)|H_1|^2 + p \sum_{j=1}^{M-1} \alpha_j \mathbf{H}^\dagger \mathbf{u}_j \mathbf{u}_j^\dagger \mathbf{H}} \right]^2 < 0.$$

Thus, beamforming is optimal if and only if:

$$0 \geq \max_{K' \in \mathcal{K}'} \frac{d}{dp}f(p)|_{p=0} \quad (14)$$

$$= \max_{\alpha, \mathbf{u}} E_{\mathbf{H} \in D_1} \frac{\sum_{j=1}^{M-1} \alpha_j \mathbf{H}^\dagger \mathbf{u}_j \mathbf{u}_j^\dagger \mathbf{H} - |H_1|^2}{1 + P|H_1|^2} \quad (15)$$

$$= \max_{\alpha, \mathbf{u}} \sum_{j=1}^{M-1} \alpha_j \mathbf{u}_j^\dagger \Lambda_1 \mathbf{u}_j - E_{\mathbf{H} \in D_1} \left[\frac{|H_1|^2}{1 + P|H_1|^2} \right] \quad (16)$$

$$= \lambda_{\max}(\Lambda_1) - \Lambda_1[1, 1] \quad (17)$$

where $\Lambda_1 = E_{\mathbf{H} \in D_1} \left[\frac{\mathbf{H}\mathbf{H}^\dagger}{1 + P|H_1|^2} \right]$ as defined in (11) and $\Lambda_1[1, 1] = E_{\mathbf{H} \in D_1} \left[\frac{|H_1|^2}{1 + P|H_1|^2} \right]$ is the first diagonal element of Λ_1 . (17) follows from the Rayleigh-Ritz theorem [10] and because a convex combination of terms is less than or equal to the maximum term. This completes the proof of Theorem 2. ■

Simplifications of the necessary and sufficient condition (2) of Theorem 2 lead to the following conditions.

1) **Condition 1:**

$$E_{\mathbf{H} \in D_1} \left[\frac{1 + P|H_k|^2}{1 + P|H_1|^2} \right] \leq 1, \quad 1 \leq k \leq M. \quad (18)$$

2) **Condition 2:**

$$E_{\mathbf{H} \in D_1} \left(\frac{H_i H_j^*}{1 + P|H_1|^2} \right) = 0, \quad 1 \leq i < j \leq M. \quad (19)$$

In terms of these simplified conditions, we have the following corollaries.

Corollary 2: [Sufficient Condition] Beamforming along \mathbf{q}_1 is the optimal transmit strategy if both Condition 1 and Condition 2 are satisfied.

Corollary 3: [Necessary Condition] Beamforming along \mathbf{q}_1 is the optimal transmit strategy only if Condition 1 is satisfied.

Proof: We provide a sketch of the proof. Detailed proof is available in [8]. Given the result of Theorem 2 the corollaries 2 and 3 follow directly once we interpret conditions 1 and 2. Condition 2 implies that Λ_1 is a diagonal matrix. Condition 1 requires that the first diagonal element of Λ_1 is the largest diagonal element of the matrix. Clearly, when both conditions hold, Λ_1 is a diagonal matrix with largest diagonal element equal to $\Lambda_1[1, 1]$. Since the diagonal elements are also the eigenvalues of a diagonal matrix, the beamforming optimality condition 2 of Theorem 2 is satisfied and Corollary 2 follows. Similarly, if $\Lambda_1[1, 1] < \Lambda_1[k, k]$ for $k \neq 1$ then a higher capacity is achieved by transmitting power along the k^{th} transmit antenna. But beamforming along \mathbf{q}_1 implies putting all the power along transmit antenna 1. Therefore, if condition 1 is violated, beamforming along \mathbf{q}_1 is not optimal, and corollary 3 follows. ■

In [7], the necessary and sufficient conditions are used to establish optimality of beamforming for several cases when

the input, output and the noise are real valued. In particular, optimality of beamforming is demonstrated for:

- 1) Two transmit antenna cases ($M = 2$) with any arbitrary number N of quantization vectors.
- 2) Cases where the number of transmit antennas M is equal to the number of quantization vectors N .

We will now look at specific scenarios where the input, channel, output and noise are in the complex domain and determine if beamforming is optimal for such cases. We will use Corollary 2 to show that the sufficiency conditions are satisfied in such situations, hence proving that beamforming is the optimal transmit strategy. We will make use of the knowledge of the structure of the optimal quantization vectors and the inherent symmetry of the decision regions involved in order to prove the conditions (18) and (19). The following theorem considers the case where the number of transmit antennas is equal to the number of quantization vectors involved, i.e., $M = N$ and establishes the optimality of beamforming for this case.

Theorem 3: *Beamforming is optimal for M transmit antennas when the system has $N = M$ quantization vectors.*

Proof: The optimal packing of N quantization vectors in $M = N$ dimensions is to have the vectors along the axes, so that $\mathbf{q}_1 = [1 \ 0 \ \cdots \ 0]^T$, $\mathbf{q}_2 = [0 \ 1 \ \cdots \ 0]^T$ and so on. If $\mathbf{H} \in D_1$, we have $|\langle \mathbf{H}, \mathbf{q}_1 \rangle| \geq |\langle \mathbf{H}, \mathbf{q}_j \rangle| \ \forall j \in \{2, \dots, M\}$. Therefore $|H_1| \geq |H_j|$ for all $j \in \{2, \dots, M\}$. So if $\mathbf{H} \in D_1$

$$\frac{1 + P |H_j|^2}{1 + P |H_1|^2} \leq 1 \quad \forall j \in \{2, \dots, M\},$$

and condition 1 is satisfied. To show that condition 2 is satisfied, we make use of the symmetry of the decision regions in the optimal packing. For every channel vector $\mathbf{H} = [H_1 \ H_2 \ H_3 \ \cdots \ H_m \ \cdots \ H_M]^T \in D_1$ there exists a corresponding $\mathbf{H}' = [H_1 \ H_2 \ H_3 \ \cdots \ -H_m \ \cdots \ H_M]^T \in D_1$ for all $m \in \{2, 3, \dots, M\}$. Also note that \mathbf{H} and \mathbf{H}' both occur with equal probability. Similarly, if $\mathbf{H} = [H_1 \ H_2 \ \cdots \ H_i \ \cdots \ H_j \ \cdots \ H_M]^T \in D_1$ there exists a corresponding equiprobable channel vector $\mathbf{H}'' = [H_1 \ H_2 \ \cdots \ H_i \ \cdots \ -H_j \ \cdots \ H_M]^T \in D_1$. This gives us

$$\mathbb{E}_{\mathbf{H} \in D_1} \left(\frac{H_i H_j^*}{1 + P |H_1|^2} \right) = 0, \quad 1 \leq i < j \leq M, \quad (20)$$

so that condition 2 is satisfied. Since both conditions are satisfied, beamforming is optimal. ■

The optimality of beamforming has been proved above for the cases where $M = N$. It is also known that as $N \rightarrow \infty$, beamforming is optimal. With an increase in N , the decision regions become smaller in size, and it is more likely that condition 1 will be satisfied. In fact, it can be shown [8] that for $N \geq 2^{M-1}$, condition 1 will always be satisfied. Optimal packings tend to be symmetric, and one can hypothesize that condition 2 will also be satisfied in most cases. This suggests that beamforming is likely optimal for all $N \geq 2^{M-1}$.

V. CASES WHERE M IS GREATER THAN N

So far we have focused on cases where $M \leq N$. We will now consider cases where $M > N$ and examine optimality of beamforming in such scenarios. [7] discusses cases where $M > N$ and derives necessary and sufficient conditions for optimality when the channel gains are i.i.d Rayleigh variables. Although those conditions are derived in [7] for cases where the input, channel and output are real valued, the same conditions apply when the parameters are complex valued. We recall here the necessary and sufficient condition for the optimality of beamforming for the complex i.i.d channel [7] when $M > N$:

$$\mathbb{E}_{\mathbf{H}} \left[\frac{1}{1 + P \max_{1 \leq i \leq N} |H_i|^2} \right] \leq \frac{1}{1 + P}. \quad (21)$$

Based on this condition, the following theorem demonstrates the optimality of beamforming for complex i.i.d Rayleigh channels with $M > N > 2$ at high SNR.

Theorem 4: *At high SNR, when the channel gains H_i are i.i.d, beamforming is optimal for all $M > N > 2$.*

Proof: At high SNR ($P \gg 1$), when the channel gains H_i are i.i.d with $\mathbb{E}[|H_i|^2] = 1$, the necessary and sufficient condition for the optimality of beamforming when $M > N$ is (Equation (21))

$$\mathbb{E}_{H_i} \left[\frac{1}{\max_{1 \leq i \leq N} |H_i|^2} \right] \leq 1. \quad (22)$$

Let $s = \max_{1 \leq i \leq N} |H_i|^2$. Since H_i are i.i.d Rayleigh distributed variables, $|H_i|^2$ are exponentially distributed with mean and variance equal to 1. The cumulative distribution function (CDF) of the maximum of N i.i.d exponential variables is given by $F_S(s) = N(1 - e^{-s})^{N-1} e^{-s}$. Defining $f(N) \triangleq \mathbb{E}_{H_i} \left[\frac{1}{\max_{1 \leq i \leq N} |H_i|^2} \right]$, we have

$$\begin{aligned} f(N) &= N \int_0^\infty \frac{(1 - e^{-s})^{N-1} e^{-s}}{s} ds \\ &= N \sum_{k=0}^{N-1} (-1)^{N+k} \binom{N-1}{k} \ln(N - k), \end{aligned} \quad (23)$$

where equation (23) follows from [11] (p. 544; 4.267 41; BI (123) (12)). Equation (23) yields the value 1.386 for $N = 2$ and 0.86 for $N = 3$. Since $f(N)$ decreases with an increase in N , for all $N \geq 3$, we have $f(N) \leq 1$. Consequently, for high SNR, beamforming is optimal for all $M > N > 2$. ■

Note that we have assumed here that the quantization is across the entire M -dimensional space. If we have complete freedom in how the quantization vectors are selected, instead of quantization over all the M dimensions, we may choose to use the N vectors to cover only a smaller number (say m with $m < N$) of the M dimensions and still be able to achieve a higher capacity. Whether there is a potential benefit from straying away from quantization for all the M dimensions and

instead using the feedback bits to more definitively identify the projections of the channel in a smaller number of dimensions is to be investigated.

VI. SIMULATIONS

In this section, we present numerical results obtained from Monte Carlo simulations for the capacity of the quantized direction feedback system. Note that the beamforming capacity discussed in Section III is an achievable lower bound on the capacity (unless beamforming is optimal). We will provide plots for this lower bound and the upper bound explained in [8]. Perfect channel knowledge capacity plots are also shown for the sake of comparison. In all our simulations, we assume that the channel is isotropic with i.i.d Rayleigh fading and $E[H_i^2] = 1 \quad \forall i \in (1, 2, \dots, M)$. We will use the best known line packings provided by [12] for the quantization vectors.

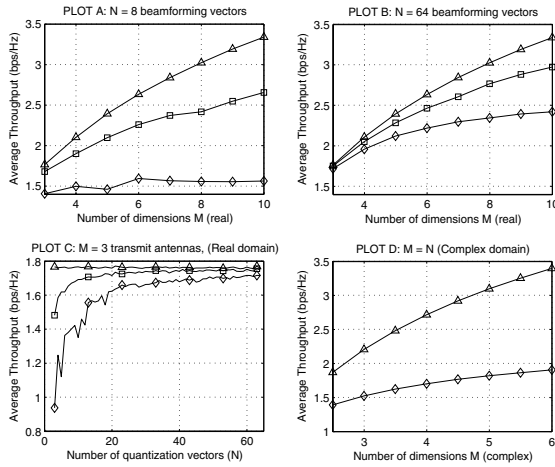


Fig. 1. Plots A-D show the average throughput (in bps/Hz) of the perfect feedback bound (Δ), the upper bound (\square) and the achievable lower bound (\diamond) for different scenarios.

Plot A and Plot B of Figure 1 compares the upper and lower bounds against the number of antennas with 8 and 64 beamforming vectors. As expected, for the same number of transmit antennas, the upper bound becomes tighter as the number of beamforming vectors increase. For the particular case of $N = 8$ quantization vectors for 8 and higher dimensions, the lower bound (beamforming capacity) does not change. This is due to the fact that in uniformly distributing 8 vectors in 8 and higher dimensions, we are effectively utilizing only 8 transmit antennas.

Plot C of Figure 1 shows the capacity bounds against the number of beamforming vectors with 3 transmit antennas. The perfect feedback capacity does not depend on the number of quantization vectors N . Consequently, given M , the perfect feedback capacity is constant. The upper and the lower bounds are expected to reach the perfect feedback bound as the number of quantization vectors goes to infinity. However note that the volume of the space increases exponentially with M .

As a result, for the same N , the gap between the upper bound and the perfect feedback capacity increases with M . The number of quantization vectors needed to have the same gap between the upper bound and the perfect channel knowledge capacity also grows exponentially in M .

Plot D of Figure 1 shows the capacity of quantized direction feedback systems with $M = N$ for complex i.i.d Rayleigh channels. The optimal quantization vectors in this case are along the M axes. Note that as the number of dimensions M increases, the number of quantization vectors needed to achieve a certain capacity increases exponentially. Consequently, in the $M = N$ plot, as M increases the gap between perfect feedback capacity and partial feedback ($\log N$ bits) capacity increases.

VII. CONCLUSION

We investigate the capacity of an isotropic fading vector channel with quantized channel feedback. We derive a necessary and sufficient condition for optimality of beamforming. We use the necessary/sufficient conditions to determine the precise relationship between the optimality of beamforming and the symmetry of the channel quantization. Using the necessary and sufficient condition, we then establish optimality of beamforming for several practical cases. While we argue that beamforming is optimal in ‘most’ cases, we also point out scenarios where it is sub-optimal (e.g. i.i.d Rayleigh channels with $M > N = 2$ at high SNR). Finally, we provide simulation plots for the beamforming capacity and the upper bound [8] for situations where the optimal vectors are known.

REFERENCES

- [1] S. Jafar and A. Goldsmith, “On the capacity region of the vector fading broadcast channel with no CSIT,” *IEEE International Conference on Communications*, June 2004.
- [2] K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, “On beamforming with finite rate feedback in multiple antenna systems,” *IEEE Transactions on Information Theory*, vol. 49, pp. 2562–2579, Oct 2003.
- [3] D. J. Love and R. W. Heath Jr., “Limited feedback precoding for spatial multiplexing systems,” *IEEE Global Communications Conference*, vol. 4, pp. 1857–1861, Dec. 2003.
- [4] D. J. Love, R. W. Heath Jr., and T. Strohmer, “Grassmannian beamforming for multiple-input multiple-output wireless systems,” *IEEE Transactions on Information Theory*, vol. 49, pp. 2735–2747, Oct 2003.
- [5] G. Jongren and M. Skoglund, “Improving orthogonal space-time block codes by utilizing quantized feedback information,” *IEEE International Symposium on Information Theory*, p. 220, June 2001.
- [6] E. G. Larsson, G. Ganesan, P. Stoica, and W. Wong, “On the performance of orthogonal space-time block coding with quantized feedback,” *IEEE Communications Letters*, vol. 6, pp. 487–489, Nov 2002.
- [7] S. A. Jafar and S. Srinivasa, “Capacity of the isotropic fading vector channel with quantized channel direction feedback,” *Asilomar Conference on Signals, Systems and Computers*, November 2004.
- [8] S. Srinivasa and S. A. Jafar, “Capacity of the isotropic fading vector channel with quantized channel direction feedback,” 2004. In preperation. Preprint available on request.
- [9] G. Caire and S. Shamai, “On the capacity of some channels with channel state information,” *IEEE Transactions on Information Theory*, vol. 45, pp. 2007–2019, Sept 1999.
- [10] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 1986.
- [11] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press, New York, 2000.
- [12] N. J. A. Sloane, “Table of best grassmannian packings,” <http://www.research.att.com/~njas/grass/>, January 1997. <http://www.research.att.com/~njas/grass/grassTab.html>.