

Optimization of Training and Feedback for Beamforming over a MIMO Channel

Wiroonsak Santipach

Department of Electrical Engineering
Faculty of Engineering, Kasetsart University
Chatuchak, Bangkok 10900, Thailand
Email: wiroonsak.s@ku.ac.th

Michael L. Honig

Department of Electrical Engineering and Computer Science
Northwestern University
Evanston, Illinois 60208, USA
Email: mh@eecs.northwestern.edu

Abstract—We examine the capacity of beamforming over a block Rayleigh fading Multi-Input/Multi-Output (MIMO) channel with finite training for channel estimation and limited feedback. A fixed-length packet is assumed, which is spanned by T training symbols, B feedback bits, and the data symbols. The training symbols are used to obtain a Minimum Mean Squared Error (MMSE) estimate of the channel matrix. Given this estimate, the receiver selects a transmit beamforming vector from a codebook containing 2^B *i.i.d.* random vectors, and relays the corresponding B -bit index back to the transmitter. We derive bounds on the large system capacity, i.e., as the number of transmit antennas $N_t \rightarrow \infty$ and receive antennas $N_r \rightarrow \infty$ with fixed ratio N_t/N_r . The bounds are used to show that the optimal T , which maximizes the capacity, increases as $N_t/\log N_t$, whereas the optimal B increases as $N_t/\log^2 N_t$.

I. INTRODUCTION

The increase in capacity offered by a Multi-Input/Multi-Output (MIMO) channel, relative to a Single-Input/Single-Output channel, depends on the number of antennas and also on available channel knowledge at the receiver and transmitter [1]. In practice, the receiver typically estimates channel coefficients from a known training sequence. The relation between the amount of training and achievable rate is studied in [2], [3]. In some situations, the transmitter may be able to obtain channel information from the receiver via a low-rate feedback channel. Several feedback schemes for this purpose have been proposed and analyzed recently [4]–[12]. Here we study the capacity of beamforming for a MIMO channel with limited feedback *and* training. This work is a continuation of prior work, which studies training and feedback for a Multi-Input/Single-Output (MISO) channel [13].

We consider an *i.i.d.* block Rayleigh fading channel with N_t transmit and N_r receive antennas. The $N_t N_r$ channel parameters are constant within each block, and are independent across blocks. The block size is assumed to be constant, and the transmitted codewords span many blocks, so that the maximum achievable rate is the ergodic capacity. Each coherence block contains T training symbols and D data symbols. Furthermore, we assume that after transmission of the training symbols, the transmitter waits for the receiver

to relay B bits over a feedback channel, which specify a particular beamforming vector. This delay, in addition to the T training symbols, is counted as part of the packet overhead.

We assume that the receiver computes a Minimum Mean Squared Error (MMSE) estimate of the channel, based on the training symbols, and uses the noisy channel estimate to choose a rank-one precoding matrix. The precoding matrix is specified by a transmit beamforming vector. *Random Vector Quantization (RVQ)* is assumed [14] in which the beamformer is selected from a codebook consisting of 2^B random vectors, which are independent and isotropically distributed, and known *a priori* at the transmitter and receiver. The associated codebook index is relayed using B bits via a noiseless feedback channel to the transmitter. The capacity of this scheme with perfect channel estimation is analyzed in [9], [10]. It is shown in [9], [10] that the RVQ codebook is optimal (i.e., maximizes the capacity) in the large system limit in which N_t , N_r , and B tend to infinity with fixed feedback bits per degree of freedom $\bar{B} = B/N_t$ and $\bar{N}_r = N_r/N_t$. RVQ has been observed to give excellent performance for systems with small N_t [15].

The capacity with MMSE channel estimates at the receiver (with or without limited feedback) is unknown. We derive upper and lower bounds on the capacity with RVQ and limited training, which are implicit functions of the number of training symbols T and feedback bits B . Given a fixed block size, or packet length L , we then optimize the capacity bounds over B and T . Namely, small T leads to a poor channel estimate, which decreases capacity, whereas large T leads to an accurate channel estimate, but leaves few symbols in the packet for transmitting the message. This tradeoff has been studied in [2], [16] for MIMO channels without feedback. Here there is also an optimal amount of feedback B , which increases with the training interval T . That is, more feedback is needed to quantize more accurate channel estimates. A similar optimization problem was studied for MISO channel in [13]. However, extension of those results (in particular, the capacity bounds) to the MIMO case is not straightforward.

We show that the optimal T/N_t and B/N_t , which maximize the capacity tend to zero at the rate of $1/\log(N_t)$ and $1/\log^2(N_t)$, respectively, as $L \rightarrow \infty$ with fixed L/N_t . For large N_t , the packet overhead devoted to feedback should

This work was supported in part by the U.S. Army Research Office under grant DAAD190310119 and the National Science Foundation under grantCCR-0310809.

be less than that devoted to training and the ratio between the optimal feedback and training lengths tends to zero as $1/\log(N_t)$. This is in contrast to the results for the MISO channel in [13], which state that the ratio between the optimal feedback and training lengths tends to one as $N_t \rightarrow \infty$. The difference is due to the fact that the amount of feedback required to achieve a target rate for a MIMO channel is less than that for a MISO channel.

II. SYSTEM MODEL

We consider a point-to-point *i.i.d.* block fading MIMO channel with N_t transmit antennas and N_r receive antennas. We assume a rich scattering environment in which the channel gains corresponding to different pairs of transmit/receive antennas are independent and Rayleigh distributed. The i th $N_r \times 1$ received vector in a particular block is given by

$$\mathbf{r}(i) = \mathbf{H}\mathbf{v}b(i) + \mathbf{n}(i) \quad \text{for } 1 \leq i \leq D \quad (1)$$

where \mathbf{H} is an $N_r \times N_t$ channel matrix whose elements are independent, complex Gaussian random variables with zero mean and unit variance, \mathbf{v} is an $N_t \times 1$ unit-norm beamforming vector, b is the transmitted symbol with unit variance, \mathbf{n} is additive white Gaussian noise (AWGN) with covariance $\sigma_n^2 \mathbf{I}$, and D is the number of data symbols in a block.

In prior work [9], [10], we have analyzed the channel capacity with perfect channel knowledge at the receiver, but with *limited* channel knowledge at the transmitter. Specifically, a quantized beamforming vector is relayed from the receiver to the transmitter, given by

$$\mathbf{v}_{\mathbf{H}} = \arg \max_{\mathbf{v}_j \in \mathcal{V}} \left\{ \log(1 + \rho \|\mathbf{H}\mathbf{v}_j\|^2) \mid \mathbf{H} \right\} \quad (2)$$

where $\rho = 1/\sigma_n^2$ is the background signal-to-noise ratio (SNR), and $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_{2^B}\}$ is the quantization codebook, which is known *a priori* at both the transmitter and receiver. The (uncoded) index corresponding to the best beamforming vector (i.e., which maximizes the achievable rate) is relayed to the transmitter via an error-free feedback link. The capacity depends on the beamforming codebook \mathcal{V} and B . As $B \rightarrow \infty$, the $\mathbf{v}_{\mathbf{H}}$ that maximizes the capacity is the eigenvector of $\mathbf{H}^\dagger \mathbf{H}$, which corresponds to the maximum eigenvalue.

We have shown in [9], [10] that RVQ, in which the codebook vectors are independent and isotropically distributed, is optimal (i.e. maximizes capacity) in the large system limit in which $(B, N_t, N_r) \rightarrow \infty$ with fixed normalized feedback $\bar{B} = B/N_t$ and $\bar{N}_r = N_r/N_t$. Furthermore, the corresponding capacity grows as $\log(\rho N_t)$. Although, strictly speaking, RVQ is suboptimal for a finite-size system, numerical results show that it gives excellent performance [15].

In addition to limited channel information at the transmitter, here we also account for channel estimation error at the receiver. Letting $\hat{\mathbf{H}}$ be the estimated channel matrix, we have

$$\mathbf{H} = \hat{\mathbf{H}} + \mathbf{w} \quad (3)$$

where \mathbf{w} is the error matrix whose elements are *i.i.d.* with zero mean and variance σ_w^2 . We assume that the receiver

computes the MMSE estimate of \mathbf{H} . As a result, $\hat{\mathbf{H}}$ and \mathbf{w} are independent and $\hat{\mathbf{H}}$ has zero mean and covariance $(1 - \sigma_w^2)\mathbf{I}$. The receiver then selects $\mathbf{v}_{\hat{\mathbf{H}}}$, assuming that $\hat{\mathbf{H}}$ is the actual channel, i.e.,

$$\mathbf{v}_{\hat{\mathbf{H}}} = \arg \max_{\mathbf{v}_j \in \mathcal{V}} \left\{ \log(1 + \rho \|\hat{\mathbf{H}}^\dagger \mathbf{v}_j\|^2) \mid \hat{\mathbf{H}} \right\}. \quad (4)$$

The quality of the channel estimate and the capacity depend on the number of training symbols T .

In what follows, we assume that the forward and feedback links are time-division multiplexed, and each block consists of T training symbols, B feedback bits, and D data symbols. Given that the size of each block is L symbols, we have the constraint

$$L = T + \mu B + D \quad (5)$$

where μ is a conversion factor, which relates bits to symbols. We would like to maximize the ergodic capacity, which is the maximum mutual information between b and \mathbf{r} ,

$$\max_{T, B} \{C = E[\max_{p_b} I(\mathbf{r}; b \mid \mathbf{H}, \mathbf{v}_{\hat{\mathbf{H}}})]\} \quad (6)$$

subject to (5), where p_b is the probability density function (pdf) for the transmitted symbol b .

III. CAPACITY OPTIMIZATION

Determining the ergodic capacity of RVQ beamforming with channel estimation in (6) appears to be intractable, so instead we derive upper and lower bounds, which are functions of D , B , and T . As in [13], we condition on the channel matrix, in addition to the estimate, to obtain an upper bound

$$C \leq E[\max_{p_b} I(\mathbf{r}; b \mid \hat{\mathbf{H}}, \mathbf{v}_{\hat{\mathbf{H}}}, \mathbf{H})] \quad (7)$$

$$= E[\log(1 + \rho \|\mathbf{H}\mathbf{v}_{\hat{\mathbf{H}}}\|^2)] \quad (8)$$

$$\leq \log(1 + \rho E[\|\mathbf{H}\mathbf{v}_{\hat{\mathbf{H}}}\|^2]) \quad (9)$$

where we use the fact that the maximizing pdf in (7) is Gaussian, and apply Jensen's inequality (9). Substituting (3) into the expectation in (9) gives

$$E[\|\mathbf{H}\mathbf{v}_{\hat{\mathbf{H}}}\|^2] = \sigma_w^2 + E[\eta] \quad (10)$$

where

$$\eta \triangleq \mathbf{v}_{\hat{\mathbf{H}}}^\dagger \hat{\mathbf{H}}^\dagger \hat{\mathbf{H}} \mathbf{v}_{\hat{\mathbf{H}}} \quad (11)$$

and we have used the fact that $\mathbf{v}_{\hat{\mathbf{H}}}$ and \mathbf{w} are independent. Substituting (10) into (9) gives the capacity upper bound

$$C \leq C_u = \log(1 + \rho \sigma_w^2 + \rho E[\eta]). \quad (12)$$

As in [13], [17], a lower bound on the capacity is obtained by assuming that the sum of estimation error and AWGN is Gaussian. That is,

$$C \geq E \left[\log \left(1 + \frac{1}{\sigma_w^2 + \sigma_n^2} \eta \right) \right] \quad (13)$$

$$\geq C_l = (1 - c_{N_t}) \log \left(1 + \frac{\rho}{1 + \rho \sigma_w^2} E[\eta] \right) \quad (14)$$

where $c_{N_t} \triangleq \sigma_\eta / (2E[\eta])$, σ_η is the standard deviation for η , and (14) is obtained by applying the inequality derived in

[18]. Note that both the upper and lower bounds, given by (12) and (14), are functions of σ_w^2 and $E[\eta]$.

The training symbols are used to compute an MMSE estimate of the channel matrix. It is shown in [2] that selecting the beamforming vectors during training to achieve the Welch bound minimizes the mean square error. The variance of the estimation error is given by [2]

$$\sigma_w^2 = \begin{cases} 1 - \frac{\bar{T}}{1+\rho^{-1}}, & \bar{T} < 1 \\ \frac{1}{1+\rho\bar{T}}, & \bar{T} \geq 1 \end{cases} \quad (15)$$

where the normalized training length $\bar{T} = T/N_t$.

To evaluate and maximize both bounds, we must evaluate

$$E[\eta] = E_{\hat{\mathbf{H}}} E_{\mathcal{V}} \left[\max_{1 \leq j \leq 2^B} \{ \mathbf{v}_j^\dagger \hat{\mathbf{H}}^\dagger \hat{\mathbf{H}} \mathbf{v}_j \} | \hat{\mathbf{H}} \right]. \quad (16)$$

Since the RVQ beamforming vectors \mathbf{v}_j , $j = 1, \dots, 2^B$, are *i.i.d.*, the corresponding received powers $\mathbf{v}_j^\dagger \hat{\mathbf{H}}^\dagger \hat{\mathbf{H}} \mathbf{v}_j$, $j = 1, \dots, 2^B$, are also *i.i.d.* However, the pdf for $\mathbf{v}_j^\dagger \hat{\mathbf{H}}^\dagger \hat{\mathbf{H}} \mathbf{v}_j$ for given $\hat{\mathbf{H}}$ is difficult to obtain [19]. As a result, computing $E[\eta]$ analytically for any N_t , N_r and B is not tractable. Maximizing the capacity bounds over training and feedback lengths for a finite-size system therefore remains an open problem. Instead, we perform an asymptotic analysis in the next section.

A. Asymptotic Analysis

To evaluate the bounds in (12) and (14), we must evaluate η in (22), where the beamforming vector $\mathbf{v}_{\hat{\mathbf{H}}}$ is given by (2). As $(N_t, N_r, B) \rightarrow \infty$ with fixed ratios $\bar{N}_r = N_r/N_t$ and $\bar{B} = B/N_t$, we can show that

$$\frac{1}{N_t} \eta \rightarrow (1 - \sigma_w^2) \gamma_{\text{rvq}} \quad (17)$$

where $\gamma_{\text{rvq}} = \lim_{(N_t, N_r, B) \rightarrow \infty} \gamma_{\text{rvq}}^{N_t}$ in the mean square sense [10] and

$$\gamma_{\text{rvq}}^{N_t} = \max_{\mathbf{v}_j \in \mathcal{V}} \mathbf{v}_j^\dagger \mathbf{H}^\dagger \mathbf{H} \mathbf{v}_j \quad (18)$$

is the received SINR for RVQ, given perfect channel knowledge at the receiver. The large system SINR can be evaluated explicitly, and is given by [10]

$$(\gamma_{\text{rvq}})^{\bar{N}_r} e^{-\gamma_{\text{rvq}}} = 2^{-\bar{B}} \left(\frac{\bar{N}_r}{e} \right)^{\bar{N}_r} \quad \text{for } 0 \leq \bar{B} \leq \bar{B}^* \quad (19)$$

$$\begin{aligned} \gamma_{\text{rvq}} &= (1 + \sqrt{\bar{N}_r})^2 - \exp \left\{ \frac{1}{2} \bar{N}_r \log(\bar{N}_r) - (\bar{N}_r - 1) \right. \\ &\quad \left. \times \log(1 + \sqrt{\bar{N}_r}) + \sqrt{\bar{N}_r} - \bar{B} \log(2) \right\} \quad \text{for } \bar{B} > \bar{B}^* \end{aligned} \quad (20)$$

where

$$\bar{B}^* = \frac{1}{\log(2)} \left(\bar{N}_r \log(\sqrt{\bar{N}_r}) - \bar{N}_r \log(1 + \sqrt{\bar{N}_r}) + \sqrt{\bar{N}_r} \right). \quad (21)$$

Note that γ_{rvq} is a function of only \bar{B} and \bar{N}_r . From (17) we can write

$$E[\eta] = (1 - \sigma_w^2) \gamma_{\text{rvq}} N_t + \kappa_{N_t} \quad (22)$$

where $\kappa_{N_t}/N_t \rightarrow 0$ as $N_t \rightarrow \infty$. Substituting (22) and (15) into (12) and (14) gives upper and lower bounds on the capacity as functions of \bar{T} and \bar{B} . With D transmitted symbols

in an L -symbol packet, the effective capacity $\mathcal{C} = (\bar{D}/\bar{L})\mathcal{C}$, where the normalized data symbols $\bar{D} = D/N_t$ and the normalized packet length $\bar{L} = L/N_t$.

We would like to maximize the capacity bounds subject to the constraint on normalized packet length, i.e.,

$$\max_{\bar{T}, \bar{B}, \bar{D}} \mathcal{C}_l = \frac{\bar{D}}{\bar{L}} \mathcal{C}_l, \quad (23)$$

$$\max_{\bar{T}, \bar{B}, \bar{D}} \mathcal{C}_u = \frac{\bar{D}}{\bar{L}} \mathcal{C}_u, \quad (24)$$

$$\text{subject to } \bar{T} + \mu \bar{B} + \bar{D} = \bar{L}. \quad (25)$$

It can be shown that the optimized training lengths for both capacity bounds have the same asymptotic behavior (as stated in the following theorem), which must apply to the optimized training length for the *actual* capacity. The analogous statement holds for the optimized feedback length. Solving these optimization problems therefore leads to the following Theorem, which applies to the actual capacity \mathcal{C} .

Theorem 1: Let $\{\bar{T}^\circ, \bar{B}^\circ, \bar{D}^\circ\} = \arg \max_{\{\bar{T}, \bar{B}, \bar{D}\}} \mathcal{C}$ subject to (25) and \mathcal{C}° be the maximized capacity. As $(N_t, N_r) \rightarrow \infty$ with fixed $\bar{N}_r = N_r/N_t$,

$$\bar{T}^\circ \log(N_t) \rightarrow \bar{L}, \quad (26)$$

$$\bar{B}^\circ \log^2(N_t) \rightarrow \frac{\bar{L}^2 \log(2)}{2\mu^2 \bar{N}_r}, \quad (27)$$

$$\frac{\bar{D}^\circ}{1 - \frac{1}{\log(N_t)} - \left(\frac{\bar{L} \log(2)}{2\bar{N}_r \mu} \right) \frac{1}{\log^2(N_t)}} \rightarrow \bar{L} \quad (28)$$

and the capacity satisfies

$$\mathcal{C}^\circ - \log(\rho N_t) + \log(\log(N_t)) \rightarrow \xi \quad (29)$$

where

$$\xi^* - \log(1 + \rho) \leq \xi \leq \xi^* \quad (30)$$

and $\xi^* = \log(\rho \bar{L} \bar{N}_r) - \log(\rho + 1) - 1$.

As $N_t \rightarrow \infty$, the capacity with limited training and feedback increases as $\log(\rho N_t) - \log(\log(N_t))$. In contrast, it is shown in [13] that for the MISO channel, the optimized rate increases as $\log(\rho N_t) - 2 \log(\log(N_t))$. Hence adding antennas at the receiver in proportion to N_t provides a gain of only $\log(\log(N_t))$ as $N_t \rightarrow \infty$.¹ The optimal training length for both MISO and MIMO channels increases at the same rate, namely $N_t/\log N_t$. However, the growth in optimal feedback length for the MISO channel is $N_t/\log N_t$, as opposed to the $N_t/\log^2 N_t$ growth shown in Theorem 1.

From Theorem 1, the ratio between optimized feedback and training lengths satisfies

$$\frac{\mu \bar{B}^\circ}{\bar{T}^\circ} \asymp \frac{\bar{L} \log(2)}{2\mu \bar{N}_r \log(N_t)} \rightarrow 0 \quad (31)$$

where $a(N_t) \asymp b(N_t)$ is used to denote $a(N_t)/b(N_t) \rightarrow 1$ as $N_t \rightarrow \infty$. That is, the fraction of the packet dedicated to training should be larger than that dedicated to feedback as

¹Of course, here we are restricted to beamforming, and are not considering any additional multiplexing gains.

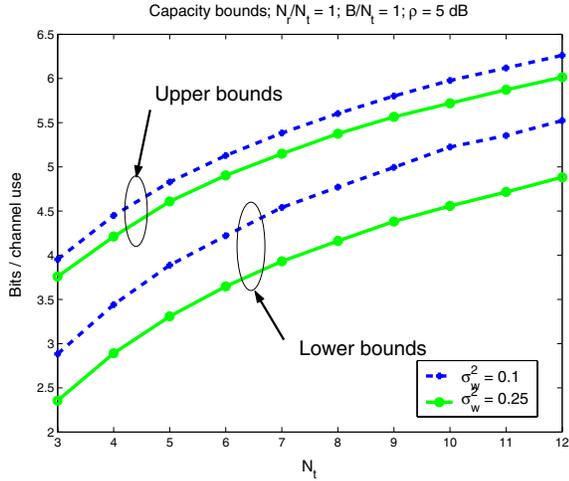


Fig. 1. Capacity bounds (12) and (14) (bits/channel use) versus number of transmit antennas with $\sigma_w^2 = 0.1$ and 0.25 .

N_t and N_r/N_t increase. In contrast, for the MISO channel $\mu\bar{B}^o \asymp \bar{T}^o$. The difference can be attributed to the fact that as N_t increases, the amount of normalized feedback needed to achieve a target rate for the MIMO channel is less than that needed for the MISO channel.

IV. NUMERICAL RESULTS

Fig. 1 shows the upper and lower bounds (12) and (14) with $\sigma_w^2 = 0.1$ and 0.25 , $\bar{B} = 1$, $\bar{N}_r = 1$, and $\rho = 5$ dB. Monte Carlo simulations were used to obtain c_{N_t} and $E[\eta]$. Both bounds grow logarithmically with N_t , as expected. The gap between the upper and lower bounds narrows as the variance of the estimation error σ_w^2 decreases. As $\sigma_w^2 \rightarrow 0$, the bounds converge to the actual capacity. We also observe that the gap between the bounds decreases with larger \bar{B} (although it does not converge to zero as $\bar{B} \rightarrow \infty$). Since RVQ requires an exhaustive search over the codebook, and the number of codebook entries grows exponentially with the number of antennas, simulation results are not shown for $N_t > 12$.

Fig. 2 shows the optimized training, feedback, and data portions of the packet (normalized by the packet length L) versus N_t . These values were obtained by numerically optimizing the capacity lower bound, and are therefore denoted as B_l^o , T_l^o , and D_l^o in the figure. System parameters are $\bar{N}_r = 2$, $\bar{L} = 50$, $\mu = 1$, and $\rho = 5$ dB. As predicted by Theorem 1, both the optimal \bar{T} and \bar{B} decrease to zero, and \bar{D} increases to \bar{L} as $N_t \rightarrow \infty$. The associated capacity lower bound is shown in Fig. 3 with a dashed line. Also shown is the capacity lower bound with the heuristic choice of parameters $\bar{B} = 1$ and $\bar{T} = 1.5$ (dashed-dot curve). For $N_t = 3$, the bound with optimized parameters is approximately 10% greater than that with the heuristic choice. Those results are compared with the capacity with perfect channel knowledge at both the transmitter and receiver, and the capacity with perfect channel knowledge at the receiver *only* with B_l^o feedback bits (solid curves). This comparison indicates how much of the loss in

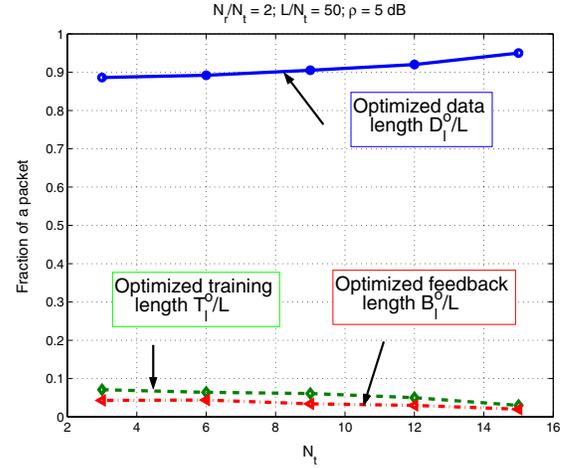


Fig. 2. $\{\bar{T}_l^o/\bar{L}, \bar{B}_l^o/\bar{L}, \bar{D}_l^o/\bar{L}\}$ versus number of transmit antennas N_t .

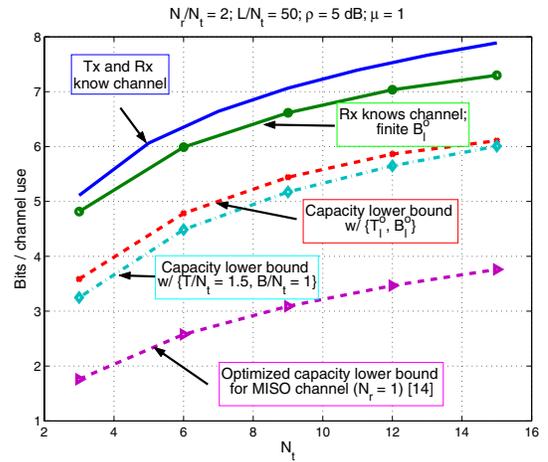


Fig. 3. Achievable rate versus number of transmit antennas N_t with different assumptions about channel knowledge at the receiver and transmitter. Also shown is the optimized capacity lower bound for the corresponding MISO channel [13]

achievable rate for the model considered is due to channel estimation at the receiver (including associated overhead), and how much is due to quantization of the precoding matrix.

The results show that for $N_t = 3$, the capacity with perfect channel knowledge at both the transmitter and receiver is about 40% larger than the rate with optimized feedback and training lengths. Knowing the channel at the receiver achieves most of this gain, largely due to the elimination of associated training overhead. Of course, this gap tends to zero as the block size $\bar{L} \rightarrow \infty$. Also shown in the figure for comparison is the capacity lower bound for a MISO channel with optimized training and feedback lengths, given in [13]. This is substantially lower than that shown for the MIMO channel. We note from Theorem 1 and [13] that the gap between the optimized lower bounds for the MISO and MIMO channels increases as $\log(\bar{N}_r)$.

Fig. 4 shows the capacity lower bound versus total overhead $(\bar{T} + \mu\bar{B})/\bar{L}$. The capacity is zero when $\bar{T} + \bar{B} = 0$,

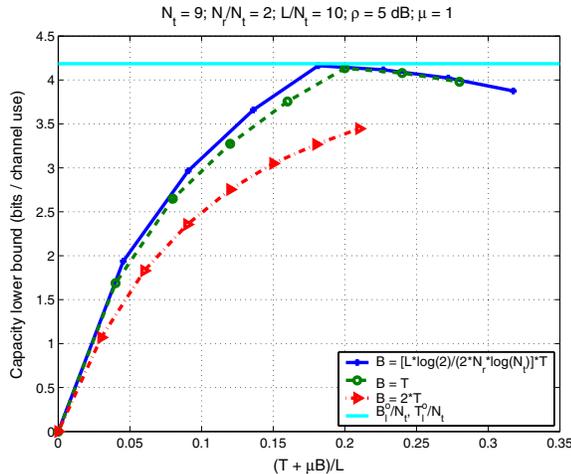


Fig. 4. Lower bound on capacity versus normalized training and feedback $(\bar{T} + \mu\bar{B})/\bar{L}$.

since the estimate is uncorrelated with the channel, and when $\bar{T} + \bar{B} = \bar{L}$, since $\bar{D} = 0$. The solid line corresponds to optimized parameters with $\bar{L} = 10$, $N_t = 9$, $N_r = 2$, $\mu = 1$, and $\rho = 5$ dB. Different curves correspond to different ratios between \bar{T} and \bar{B} . For the case considered, these results show that the rate achieved with equal portions of training and feedback is close to the maximum (corresponding to optimized training and feedback). Allocating the overhead according to the asymptotic results in Theorem 1, i.e., taking $\mu\bar{B}/\bar{T} = \bar{L} \log(2)/(2\mu N_r \log(N_t))$, performs marginally better than allocating equal training and feedback. The total optimized overhead in this case is $(\bar{T} + \bar{B})/\bar{L} \approx 0.2$. The performance degrades when \bar{B} deviates significantly from \bar{T} (e.g., $\bar{B} = 2\bar{T}$). The three curves shown are not extended to $(\bar{T} + \bar{B})/\bar{L} = 1$ since the simulation complexity associated with RVQ increases exponentially with \bar{B} .

V. CONCLUSION

We have presented bounds on the capacity of a MIMO block Rayleigh fading channel with beamforming, assuming limited training and feedback. For a large number of transmit antennas, we have characterized the optimal amount of training and feedback as a fraction of the packet duration, assuming linear MMSE estimation of the channel, and an RVQ codebook for quantizing the beamforming vector. Our results show that when optimized, the fractions of the packet devoted to training and feedback tend to zero at the rates of $1/\log N_t$ and $1/\log^2 N_t$, respectively, as $N_t \rightarrow \infty$. This is in contrast to the analogous results for the MISO channel in [13], which state that both the optimized training and feedback lengths tend to zero at the rate of $1/\log N_t$. For the numerical example considered, allocating packet overhead in a finite-size system according to the ratio $\mu\bar{B}/\bar{T} = \bar{L} \log(2)/(2\mu N_r \log(N_t))$, which is obtained from the asymptotic analysis, can achieve near-optimal performance.

Although the pilot-based scheme considered for channel estimation is practical, it is most likely suboptimal. Namely, in

the absence of feedback such a pilot-based scheme is strictly suboptimal (does not achieve capacity), although it is nearly optimal at high SNRs [2]. With feedback the capacity of the block fading MIMO channel considered (i.e., no channel knowledge at the receiver and transmitter) is unknown. Here we have assumed that the average power is constant throughout the entire packet. It may be possible to extend these results to allow different powers for the training and data portions, although based on prior related results in [2], [16], we suspect that this additional flexibility will provide only a marginal performance gain.

REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.
- [2] B. Hassibi and B. M. Hochwald, "How much training is needed in a multiple-antenna wireless link?" *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–964, Apr. 2003.
- [3] T. L. Marzetta, "BLAST training: Estimating channel characteristics for high-capacity space-time wireless," in *Proc. 37th Annu. Allerton Conf. on Communications, Control, and Computing*, Monticello, IL, Sept. 1999.
- [4] 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, no. 8, pp. 1423–1436, Oct. 1998.
- [5] D. J. Love and R. W. Heath Jr., "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2735–2745, Oct. 2003.
- [6] 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.
- [7] 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*, vol. 52, no. 3, pp. 1101–1112, Mar. 2006.
- [8] V. K. N. Lau, Y. Liu, and T.-A. Chen, "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.
- [9] W. Santipach and M. L. Honig, "Asymptotic capacity of beamforming with limited feedback," in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, Chicago, IL, June 2004, p. 290.
- [10] —, "Capacity of a multi-antenna fading channel with a quantized precoding matrix," in preparation.
- [11] W. Dai, Y. Liu, V. K. N. Lau, and B. Rider, "On the information rate of MIMO systems with finite rate channel state feedback and power on/off strategy," in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, Adelaide, Australia, Sept. 2005, pp. 1549–1553.
- [12] C. K. Au-Yeung and D. J. Love, "On the performance of random vector quantization limited feedback beamforming," *IEEE Trans. Wireless Commun.*, 2005, submitted for publication.
- [13] W. Santipach and M. L. Honig, "Capacity of beamforming with limited training and feedback," in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, Seattle, WA, July 2006.
- [14] —, "Signature optimization for CDMA with limited feedback," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, Oct. 2005.
- [15] D. J. Love, R. W. Heath Jr., W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 54–59, Oct. 2004.
- [16] Y. Sun and M. L. Honig, "Large system capacity of MIMO block fading channels with least squares linear adaptive receivers," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, St. Louis, MO, Nov. 2005.
- [17] M. Médard, "The effect upon channel capacity in wireless communication of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.
- [18] A. Ben-Tal and E. Hochman, "More bounds on the expectation of a convex function of a random variable," *Journal of Applied Probability*, vol. 9, pp. 803–812, 1972.
- [19] S. B. Provost and Y.-H. Cheong, "The distribution of Hermitian quadratic forms in elliptically contoured random vectors," *Journal of Statistical Planning and Inference*, vol. 102, no. 2, pp. 303–316, Apr. 2002.