Capacity of Beamforming with Limited Training and Feedback

Wiroonsak Santipach and Michael L. Honig
Department of Electrical Engineering and Computer Science
Northwestern University
Evanston, IL 60208–3118 USA
{sak,mh}@eecs.northwestern.edu

***Abstract*— We examine the capacity of beamforming over a Multi-Input/Single-Output block Rayleigh fading 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 vector. 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 bits back to the transmitter. We derive bounds on the capacity and show that for a large number of transmit antennas N_t, the optimal T and B, which maximize the bounds, are approximately equal and both increase as $N_t/\log N_t$. We conclude that with limited training and feedback, the optimal number of antennas to activate also increases as $N_t/\log N_t$.**

I. Introduction

The capacity of a multi-antenna system with independent Rayleigh fading and perfect channel knowledge at the transmitter and receiver increases with number of antennas [1], [2]. In practice, the channel estimate at the receiver will not be perfect, and furthermore, this estimate must be quantized before it is relayed back to the transmitter. This has motivated recent work on the performance of feedback schemes with imperfect channel knowledge [3]–[5], and the design and performance of limited feedback schemes for Multi-Input/Multi-Output (MIMO) and Multi-Input/Single-Output (MISO) channels [6]–[14]. All of the preceding work on limited feedback assumes perfect channel knowledge at the receiver. Here we consider the performance of beamforming for a MISO channel with *both* an imperfect channel estimate at the receiver *and* limited feedback.

We consider an *i.i.d.* block Rayleigh fading channel in which the channel parameters are stationary within each block, and are independent from block to block. The block size is assumed 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 transmit beamforming vector. A *Random Vector Quantization (RVQ)* scheme is assumed [15] 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], [12]. It is shown in [9], [12] that the RVQ codebook is optimal (i.e., maximizes the capacity) in the large system limit in which number of transmit antennas N_t and B tend to infinity with fixed ratio $\bar{B} = B/N_t$. RVQ has been observed to give excellent performance for systems with small N_t [16]. Furthermore, for the MISO channel considered, the performance averaged over the random codebooks can be explicitly computed [14].

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 feedback, which are 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 [17], [18] 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.

As the packet length $L \to \infty$ with fixed $\bar{L} = L/N_t$, we show that the optimal $\bar{T} = T/N_t$ and $\bar{B} = B/N_t$, which maximize the bounds on capacity, both tend to zero at the rate $1/\log N_t$. Consequently, T increases as $N_t/\log N_t$, and we observe that the associated capacity can be achieved by activating only $N_t/\log N_t$ antennas. Equivalently, for this pilot-based scheme with limited feedback, the optimal number of (active) transmit antennas increases as $L/\log L$.

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

II. System Model

We consider a point-to-point MISO *i.i.d.* block fading channel with N_t transmit antennas. We assume a rich scattering environment in which the channel gains across transmit antennas are independent and Rayleigh distributed. The ith received symbol of a particular block is given by

$$r(i) = (\boldsymbol{h}^\dagger \boldsymbol{v}) b(i) + n(i) \qquad \text{for} \quad 1 \le i \le D \tag{1}$$

where $\boldsymbol{h}$ is an $N_t \times 1$ channel vector whose elements are independent, complex Gaussian random variables with zero mean and unit variance, $\boldsymbol{v}$ is an $N_t \times 1$ unit-norm beamforming vector, b is the transmitted symbol with unit variance, and n is additive white Gaussian noise with variance σ_n^2.

In prior work [9], [12], 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

$$\boldsymbol{v}_{\boldsymbol{h}} = \arg\max_{\boldsymbol{v}_j \in \mathcal{V}} \left\{ \log(1 + \rho |\boldsymbol{h}^\dagger \boldsymbol{v}_j|^2) \,\middle|\, \boldsymbol{h} \right\} \tag{2}$$

where $\rho = 1/\sigma_n^2$, and $\mathcal{V} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{2^B}\}$ is the quantization codebook, which is known at both the transmitter and receiver *a priori*. The (uncoded) index corresponding to the best beamforming vector is relayed to the transmitter via an error-free feedback link. The capacity depends on the beamforming codebook $\mathcal{V}$ and B. As $B \to \infty$, the $\boldsymbol{v}_{\boldsymbol{h}}$ that maximizes the capacity is the normalized channel vector $\boldsymbol{h}/\|\boldsymbol{h}\|$.

We have shown in [9], [12] 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 $(B, N_t) \to \infty$ with fixed normalized feedback $\bar{B} = B/N_t$. The resulting capacity was shown to grow as $\log(\rho N_t)$. Although, strictly speaking, RVQ is suboptimal for a finite-size system, numerical results show that it gives excellent performance [16].

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

$$\boldsymbol{h} = \hat{\boldsymbol{h}} + \boldsymbol{w} \tag{3}$$

where $\boldsymbol{w}$ is the error vector whose elements are *i.i.d.* with zero mean and variance σ_w^2. Here we assume that the receiver computes the MMSE estimate of $\boldsymbol{h}$. As a result, $\hat{\boldsymbol{h}}$ and $\boldsymbol{w}$ are independent and $\hat{\boldsymbol{h}}$ has zero mean and covariance $(1 - \sigma_w^2)\boldsymbol{I}$. The receiver then selects $\boldsymbol{v}_{\hat{\boldsymbol{h}}}$, assuming that $\hat{\boldsymbol{h}}$ is the actual channel, i.e.,

$$\boldsymbol{v}_{\hat{\boldsymbol{h}}} = \arg\max_{\boldsymbol{v}_j \in \mathcal{V}} \left\{ \log(1 + \rho |\hat{\boldsymbol{h}}^\dagger \boldsymbol{v}_j|^2) \,\middle|\, \hat{\boldsymbol{h}} \right\} \tag{4}$$

The quality of the channel estimate depends on the number of training symbols T, and so does the capacity.

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

$$L = T + \mu B + D \tag{5}$$

where μ is a conversion factor, which relates bits to symbols. Determining the ergodic capacity of RVQ with channel estimation appears to be intractable, so instead we derive upper and lower bounds, which are functions of D, B, and T. We would like to optimize both bounds over $\{D, B, T\}$, subject to (5).

III. Capacity Bounds

The ergodic capacity with channel estimation and quantized beamforming is the maximum mutual information between the received and transmitted symbols, and is given by

$$C = E[\max_{p_b} I(r; b | \hat{\boldsymbol{h}}, \boldsymbol{v}_{\hat{\boldsymbol{h}}})] \tag{6}$$

where p_b is the probability density function (pdf) for b and the expectation is over $\hat{\boldsymbol{h}}$ and $\boldsymbol{v}_{\hat{\boldsymbol{h}}}$. Conditioning on the actual channel vector, instead of the estimate, gives the upper bound

$$C \le E[\max_{p_b} I(r; b | \hat{\boldsymbol{h}}, \boldsymbol{v}_{\hat{\boldsymbol{h}}}, \boldsymbol{h})] \quad = E[\log(1 + \rho (\boldsymbol{v}_{\hat{\boldsymbol{h}}}^\dagger \boldsymbol{h})^2)] \tag{7}$$

$$\le \log(1 + \rho E[(\boldsymbol{v}_{\hat{\boldsymbol{h}}}^\dagger \boldsymbol{h})^2]) \tag{8}$$

where we use the fact that the maximizing pdf is Gaussian, and apply Jensen's inequality (8). Substituting (3) into the expectation in (8), and simplifying gives

$$E[(\boldsymbol{v}_{\hat{\boldsymbol{h}}}^\dagger \boldsymbol{h})^2] = \sigma_w^2 + E[(\boldsymbol{v}_{\hat{\boldsymbol{h}}}^\dagger \hat{\boldsymbol{h}})^2]. \tag{9}$$

Since $\|\hat{\boldsymbol{h}}\|^2$ and $\nu \triangleq (\boldsymbol{v}_{\hat{\boldsymbol{h}}}^\dagger \hat{\boldsymbol{h}}/\|\hat{\boldsymbol{h}}\|)^2$ are independent [10], [14], we have

$$E[(\boldsymbol{v}_{\hat{\boldsymbol{h}}}^\dagger \hat{\boldsymbol{h}})^2] = E[\|\hat{\boldsymbol{h}}\|^2] E[\nu] = (1 - \sigma_w^2) N_t E[\nu] \tag{10}$$

where $\nu = \max_{1 \le j \le 2^B} \{\nu_j = (\boldsymbol{v}_j^\dagger \hat{\boldsymbol{h}}/\|\hat{\boldsymbol{h}}\|)^2\}$. With RVQ the ν_j's are *i.i.d.* with pdf given in [8]. The pdf for ν and associated mean can be explicitly computed [14]. The mean is given by

$$E[\nu] = 1 - 2^B B \left(2^B, \frac{N_t}{N_t - 1} \right) \tag{11}$$

where the beta function $B(m, n) = \int_0^1 t^{m-1}(1-t)^{n-1}\,\mathrm{d}t$ for m and $n > 0$. We can bound $E[\nu]$ as follows.

Lemma 1: For $\bar{B} \ge 0$ and $N_t \ge 2$,

$$E[\nu] \quad \le \quad 1 - 2^{-\bar{B}} + \frac{1 + (\gamma - 1) 2^{-\bar{B}} + 2^{-\bar{B} N_t}}{N_t - 1} \tag{12}$$

$$E[\nu] \quad \ge \quad 1 - 2^{-\bar{B}} \tag{13}$$

where γ is Euler's number.

The proof is based on the inequality derived in [19]. We note that $E[\nu] \to 1 - 2^{-\bar{B}}$ as $N_t \to \infty$. Substituting (9)-(12) into (8) gives an upper bound on capacity.

To derive a lower bound on capacity, we substitute (3) into (1) and obtain

$$r(i) = (\boldsymbol{v}_{\hat{\boldsymbol{h}}}^\dagger \hat{\boldsymbol{h}}) b(i) + \underbrace{(\boldsymbol{v}_{\hat{\boldsymbol{h}}}^\dagger \boldsymbol{w}) b(i) + n(i)}_{z(i)}. \tag{14}$$

Since $\boldsymbol{w}$ and $\hat{\boldsymbol{h}}$ are independent, it can be shown that $E[z(i)b(i)] = 0$. It is shown in [17], [20] that replacing $z(i)$ with a zero-mean Gaussian random variable minimizes the mutual information $I(r; b|\hat{\boldsymbol{h}}, \boldsymbol{v}_{\hat{\boldsymbol{h}}})$ and therefore gives a lower bound on the capacity with channel estimation and quantized beamforming. The lower bound is maximized when $b(i)$ has a Gaussian pdf, i.e.,

$$C \geq E[\max_{p_b} \min_{p_z} I(r; b|\hat{\boldsymbol{h}}, \boldsymbol{v}_{\hat{\boldsymbol{h}}})] = E\left[\log\left(1 + \frac{(\hat{\boldsymbol{h}}^\dagger \boldsymbol{v}_{\hat{\boldsymbol{h}}})^2}{\sigma_z^2}\right)\right] \tag{15}$$

where p_z and σ_z^2 denote the pdf and variance for z, respectively. We derive the following lower bound on C by applying the inequality in [21].

Lemma 2:

$$\begin{aligned} &E\left[\log\left(1 + \frac{1}{\sigma_z^2}(\hat{\boldsymbol{h}}^\dagger \boldsymbol{v}_{\hat{\boldsymbol{h}}})^2\right)\right] \\ &\quad \geq \left(1 - \frac{\sigma_\nu}{2E[\nu]}\right)\log\left(1 + \frac{1}{\sigma_z^2}E[(\hat{\boldsymbol{h}}^\dagger \boldsymbol{v}_{\hat{\boldsymbol{h}}})^2]\right) \end{aligned} \tag{16}$$

where σ_ν^2 denotes the variance of ν.

Exact evaluation of σ_ν appears to be intractable; however, we are able to derive the upper bound

$$\frac{\sigma_\nu}{2E[\nu]} \leq \frac{\sqrt{\Gamma\left(1 + \frac{2}{N_t - 1}\right) - \Gamma^2\left(1 + \frac{1}{N_t - 1}\right)}}{2^{1 + \bar{B} + \frac{\bar{B}}{N_t - 1}} - 2\Gamma\left(1 + \frac{1}{N_t - 1}\right)} \triangleq d_{N_t} \tag{17}$$

where $\Gamma(\cdot)$ is the gamma function. We note that $d_{N_t} \to 0$ as $N_t \to \infty$.

To obtain a lower bound on capacity C, we substitute $\sigma_z^2 = \sigma_w^2 + \sigma_n^2$, (13), and (16)-(17) into (15). The capacity bounds are summarized as follows.

Theorem 1: The capacity with channel estimation variance σ_w^2 and normalized feedback $\bar{B}$ satisfies

$$C_l \leq C \leq C_u \quad \text{for} \quad \bar{B} \geq 0 \text{ and } N_t \geq 2 \tag{18}$$

where

$$C_l = (1 - d_{N_t})\log\left(1 + \rho\frac{1 - \sigma_w^2}{1 + \rho\sigma_w^2}(1 - 2^{-\bar{B}})N_t\right), \tag{19}$$

$$\begin{aligned} C_u = \log\Bigg(&1 + \rho\sigma_w^2 + \rho(1 - \sigma_w^2)N_t \\ &\times\left(1 - 2^{-\bar{B}} + \frac{1 + (\gamma - 1)2^{-\bar{B}} + 2^{-\bar{B}N_t}}{N_t - 1}\right)\Bigg). \end{aligned} \tag{20}$$

The gap between the two bounds tends to zero as σ_w^2 or ρ tend to zero. With fixed $\bar{B}$ and σ_w^2 both bounds (and the capacity) grow as $O(\log(N_t))$ as $N_t \to \infty$.

IV. Optimized Training and Feedback length

A. Channel Estimation Error

We first evaluate the channel estimation error in terms of the training length $\bar{T}$ and feedback $\bar{B}$. We assume that the transmitter transmits T training symbols $b_T(1), \cdots, b_T(T)$, and that the training symbol $b_T(i)$ modulates the corresponding beamforming vector $\boldsymbol{v}_T(i)$. The vector of T received samples from (1) is given by

$$\boldsymbol{r} = \boldsymbol{B}_T \boldsymbol{V}_T^\dagger \boldsymbol{h} + \boldsymbol{n} \tag{21}$$

where $\boldsymbol{B}_T = \text{diag}\{b_T(i)\}$ is a $T \times T$ matrix, $\boldsymbol{V}_T = [\boldsymbol{v}_T(1) \cdots \boldsymbol{v}_T(T)]$, and $\boldsymbol{n} = [n(1) \cdots n(T)]^T$. The $T \times N_t$ linear MMSE channel estimation filter is given by

$$\begin{aligned} \boldsymbol{C} &= \arg\min_{\tilde{\boldsymbol{C}}} E[\|\boldsymbol{h} - \tilde{\boldsymbol{C}}^\dagger \boldsymbol{r}\|^2] &(22) \\ &= (\boldsymbol{V}_T^\dagger \boldsymbol{V}_T + \sigma_n^2 \boldsymbol{I})^{-1} \boldsymbol{B}_T \boldsymbol{V}_T^\dagger &(23) \end{aligned}$$

and the MSE $\sigma_w^2 = 1 - \text{trace}\{\boldsymbol{C}^\dagger \boldsymbol{R} \boldsymbol{C}\}/N_t$ where $\boldsymbol{R} = E[\boldsymbol{r}\boldsymbol{r}^\dagger] = \boldsymbol{B}_T \boldsymbol{V}_T^\dagger \boldsymbol{V}_T \boldsymbol{B}_T^\dagger + \sigma_n^2 \boldsymbol{I}$ is the received covariance matrix. Note that the matrix of beamforming vectors during training, $\boldsymbol{V}_T$, is known to the transmitter and receiver, and can be chosen *a priori*. It is shown in [17] that the set of (unit-norm) beamforming vectors, which achieve the Welch bound with equality, minimizes the Mean Squared Error (MSE). We therefore have that [22]

$$\boldsymbol{V}_T \boldsymbol{V}_T^\dagger = \bar{T}\boldsymbol{I} \quad \text{if} \quad T > N_t, \tag{24}$$

$$\boldsymbol{V}_T^\dagger \boldsymbol{V}_T = \boldsymbol{I} \quad \text{if} \quad T \leq N_t. \tag{25}$$

Applying (23)-(25), we obtain the variance of the estimation error

$$\sigma_w^2 = \begin{cases} 1 - \frac{\bar{T}}{1 + \rho^{-1}}, & \bar{T} < 1 \\ \frac{1}{1 + \rho\bar{T}}, & \bar{T} \geq 1 \end{cases} \tag{26}$$

B. Asymptotic Behavior

We now study the behavior of the optimal T, B and D, and the capacity as $N_t \to \infty$. With D transmitted symbols in an L-symbol packet the effective capacity $\mathcal{C} = (\bar{D}/\bar{L})C$ where $\bar{D} = D/N_t$ and $\bar{L} = L/N_t$. The associated bounds are $\mathcal{C}_u = (\bar{D}/\bar{L})C_u$ and $\mathcal{C}_l = (\bar{D}/\bar{L})C_l$. From Theorem 1 and (26), we can write $\mathcal{C}_l$ and $\mathcal{C}_u$ as functions of $\{\bar{T}, \bar{B}, \bar{D}\}$ and optimize, i.e., for the lower bound we wish to

$$\max_{\bar{T}, \bar{B}, \bar{D}} \mathcal{C}_l \tag{27}$$

$$\text{subject to} \quad \bar{T} + \mu\bar{B} + \bar{D} = \bar{L}. \tag{28}$$

Let $\{\bar{T}_l^o, \bar{B}_l^o, \bar{D}_l^o\}$ denote the optimal values of $\bar{T}$, $\bar{B}$, and $\bar{D}$, respectively, and let $\mathcal{C}_l^o$ denote the maximized lower bound on capacity. Similarly, maximizing the upper bound gives the optimal parameters $\{\bar{T}_u^o, \bar{B}_u^o, \bar{D}_u^o\}$ and the corresponding bound $\mathcal{C}_u^o$. These solutions can be easily computed numerically, and also allow us to characterize the asymptotic behavior of the *actual* capacity.

Theorem 2: Let $\{\bar{T}^o, \bar{B}^o, D^o\} = \arg\max_{\{\bar{T}, \bar{B}, \bar{D}\}} \mathcal{C}$ subject to (28). As $N_t \to \infty$,

$$\bar{T}^o \log(N_t) \to \bar{L} \tag{29}$$

$$\bar{B}^o \log(N_t) \to \frac{1}{\mu}\bar{L} \tag{30}$$

$$\frac{\bar{D}^o}{1 - \frac{2}{\log(N_t)}} \to \bar{L} \tag{31}$$

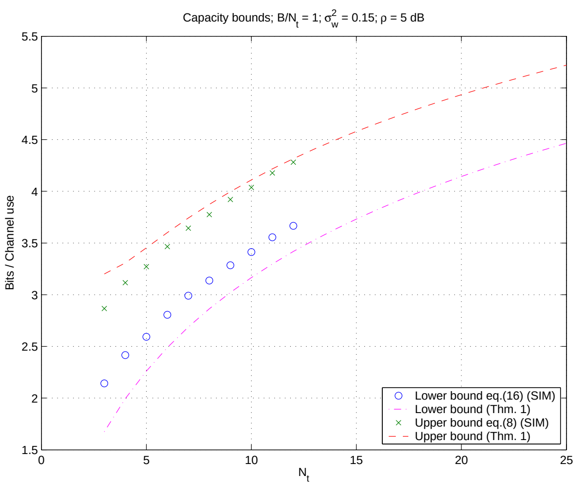

Fig. 1. The capacity bounds in Theorem 1 (bits/channel use) versus number of transmit antennas.

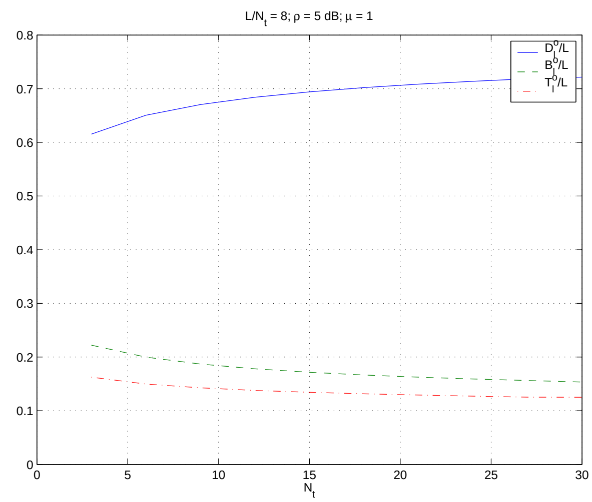

Fig. 2. $\{\bar{T}_l^o/\bar{L}, \bar{B}_l^o/\bar{L}, \bar{D}_l^o/\bar{L}\}$ versus number of transmit antennas N_t.

and the capacity satisfies

$$\mathcal{C}^o - \log(\rho N_t) - 2\log(\log(N_t)) \to \zeta \tag{32}$$

where

$$\zeta^* - \log(1+\rho) \le \zeta \le \zeta^* \tag{33}$$

and $\zeta^* = \log(\bar{L}^2(1+\rho)\log(2)) - \log(\mu\rho) - 2$.

According to the theorem, as N_t becomes large, to maximize the achievable rate the fraction of $\bar{L}$ devoted to training and feedback tends to zero, in which case the rate increases as $\log(\rho N_t) - 2\log(\log(N_t))$. Recall that the achievable rate with RVQ and perfect channel estimation grows as $\log(\rho N_t)$. Hence the loss of $2\log(\log(N_t))$ is due to imperfect channel estimation. Theorem 2 also implies that $\mu B/T \to 1$, i.e., the fraction of the packet devoted to feedback is asymptotically the same as that for training.

We observe that the preceding analysis applies if the beamforming vectors during training are chosen to be unit vectors. Namely, the matrix $\boldsymbol{V}_T$ can be taken to be diagonal, which corresponds to transmitting the sequence of training symbols over the transmit antennas successively one at a time. Hence the fact that the optimal T increases as $N_t/\log N_t$ implies that only $N_t/\log N_t$ antennas are activated. Equivalently, we conclude that as the packet size L increases, the optimal number of transmit antennas should increase as $L/\log L$.

V. Numerical Results

In Fig. 1, we compare the analytical bounds in Theorem 1 with the tighter bounds in (7) and (15). The tighter bounds, which are analytically intractable, are evaluated by Monte Carlo simulation and shown as $\circ$'s and $\times$'s in the figure. The plots show that the bounds in Theorem 1 are close to (7) and (15) even for small N_t. Since RVQ requires an exhaustive search, and the number of entries in the codebook grows exponentially with the number of antennas, simulation results are not shown for $N_t > 12$. As expected, both the upper and lower bounds grow at the same rate as N_t increases.

Fig. 2 shows the set of optimal values $\{\bar{T}_l^o/\bar{L}, \bar{B}_l^o/\bar{L}, \bar{D}_l^o/\bar{L}\}$, which maximizes $\mathcal{C}_l$, versus N_t with normalized block length $\bar{L} = 8$ and $\mu = 1$. As expected from Theorem 2, the optimal training and feedback lengths decrease to zero. The associated rate with this set of parameters is shown in Fig. 3 with a solid line. The dots correspond to simulation results with the same set of parameters as in Fig. 2. The numerical results for the bound in (15) nearly match the analytical lower bound $\mathcal{C}_l$, even for $N_t = 3$. We also compare this performance with optimized parameters to that with $\bar{T} = 2$, $\bar{B} = 2$, and $\bar{D} = 4$, which may be a reasonable heuristic choice of parameters. The rate loss at $N_t = 10$ is about 10%. Both rates are substantially less than the rate with perfect channel information at the transmitter and receiver, which is displayed by the dashed line. The dash-dot curve is the capacity with perfect channel estimation, where $\bar{B}$ is taken to be the optimized value corresponding to the solid line. Here we see a substantial gain relative to the solid line, since with perfect channel knowledge the receiver does not require training overhead.

Fig. 4 shows the capacity lower bound versus total overhead $(\bar{T}+\mu\bar{B})/\bar{L}$ with $\mu = 1$. The capacity is zero when $\bar{T}+\bar{B} = 0$, 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 $\bar{L} = 5$, $N_t = 5$, $\mu = 1$, and $\rho = 5$ dB. Different curves correspond to different ratios between $\bar{T}$ and $\bar{B}$. With equal amounts of training and feedback, the rate is essentially equal to that with optimized parameters. The peak is achieved when $(\bar{T}+\bar{B})/\bar{L} = 0.35$. The performance degrades when $\bar{B}$ deviates significantly from $\bar{T}$. Also shown are the simulation results for (15) when $T = B$. The analytical bound is quite close to the bound in (15) for $(\bar{T}+\bar{B})/\bar{L} \ge 0.5$.

VI. Conclusions

We have presented bounds on the capacity of a MISO block Rayleigh fading channel with beamforming, assuming

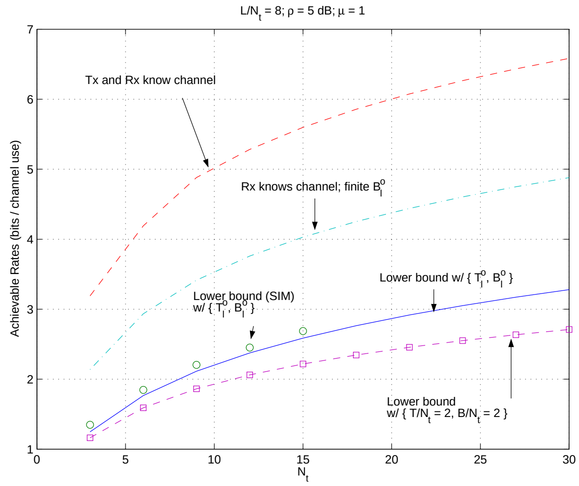

Fig. 3. Achievable rate versus number of transmit antennas N_t with different assumptions about channel knowledge at the receiver and transmitter.

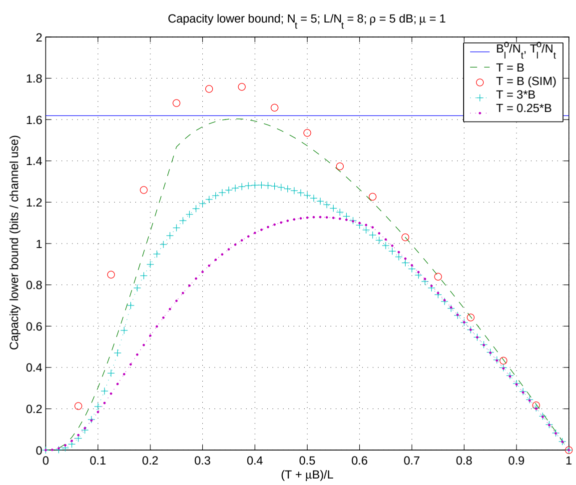

Fig. 4. Lower bound on capacity versus normalized training and feedback $(\bar{T} + \mu\bar{B})/\bar{L}$.

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 fraction of the packet devoted to training is asymptotically the same as the fraction of the packet devoted to feedback. Furthermore, the optimal training length increases as $N_t/\log(N_t)$, which can be interpreted as the optimal number of transmit antennas to activate.

Although the pilot-based scheme considered is practical, it is most likely suboptimal. Namely, in the absence of feedback such a pilot-based scheme is strictly suboptimal, although it is nearly optimal at high SNRs [17]. With feedback the capacity of the block fading MISO channel considered (i.e., no channel knowledge at the receiver and transmitter) is unknown. Extensions of the model presented here, which we intend to study, include allocating different powers for the training and data portions, and beamforming for a MIMO block fading channel.

REFERENCES

[1] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[3] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2632–2639, Sept. 2001.

[4] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2599–2613, Oct. 2003.

[5] S. H. Simon and A. L. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 406–417, Apr. 2003.

[6] 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.

[7] 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.

[8] 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.

[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] 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*, Aug. 2004, submitted for publication.

[11] 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.

[12] W. Santipach and M. L. Honig, "Capacity of multiple-antenna fading channel with quantized precoding matrix," in preparation.

[13] 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.

[14] 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.

[15] W. Santipach and M. L. Honig, "Signature optimization for CDMA with limited feedback," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, Oct. 2005.

[16] 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.

[17] 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.

[18] 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.

[19] D. Kershaw, "Some extensions of W. Gautschi's inequalities for the gamma function," *Mathematics of Computation*, vol. 41, no. 164, pp. 607–611, 1983.

[20] 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.

[21] 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.

[22] M. Rupf and J. L. Massey, "Optimum sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1261–1266, July 1994.