On Channel State Information in Multiple Antenna Block Fading Channels

Ashutosh Sabharwal†, Elza Erkip‡ and Behnaam Aazhang†

†Department of Electrical and Computer Engineering
Rice University
Houston TX 77005

‡Department of Electrical Engineering
Polytechnic University
Brooklyn, NY 11201

Abstract
Motivated by large capacity gains in multiple antenna systems when perfect channel state information is available at both receiver and transmitter, we examine the achievable rates of channel measurement based techniques. We assume that instantaneous channel state is not known a priori and measured at the receiver and transmitter using part of the total system resources. The receiver side information is obtained using a preamble. For transmitter side information, two classes of feedback channel are studied, a time-division duplex mode and a two-way channel mode. For each class, we study both low-rate quantized and reverse preamble feedback. Lower bounds for achievable rates are derived for the proposed methods, and these bounds are used for resource allocation among the preamble and the feedback. For slowly fading channels, significant gains over non-measurement schemes are possible even with a single bit of feedback.

1. Introduction

Block fading channels form a good approximation of the time-varying channels encountered in narrowband mobile communications, and hence have been a topic of active research [1–4]. In general, both transmitter and receiver are unaware of the instantaneous channel state, hence a capacity analysis à la [3, 5] is an accurate indicator of the achievable performance for methods without receiver feedback. The capacity achieving distribution in the absence of channel state information at transmitter (CSIT) and receiver (CSIR) is non-Gaussian, thereby precluding the use of rich literature on codes for additive Gaussian channels.

The analysis in [3] clearly shows the suboptimality of training based schemes; capacity achieving codes do not waste any resources to estimate the channel. But the analysis in [3] is limited to coding schemes which do not send any feedback to the receiver. Feedback in the form of channel state information has been shown to increase the capacity of fading channels [1, 2]. But the results in [1, 2] do not account for the fact that the feedback channel resources are part of the total spectral allocation for a system, and that CSIR is rarely perfect. If both feedback and non-feedback based schemes use the same spectral allocation, then it is not immediate that implementing a feedback channel with imperfect CSIR is ever beneficial.

In this paper, we analyze block fading channels with no CSIR and no CSIT, and study achievable rates of measurement based coding schemes. We show that carefully designed power control policies can perform better than the optimal coding schemes without feedback [3] for some channels. We assume that a partial CSIR is obtained via a preamble. Two classes of feedback channels are considered: low-rate causal feedback in a time-division duplex (TDD) mode †, and feedback in two-way communication mode [6]. For each of the two feedback classes, two different methods of obtaining partial CSIT are introduced: CSIT obtained from quantized low-bit rate CSIR, and CSIT obtained from a preamble transmitted by the receiver. For all schemes, lower bounds on mutual information are derived, which provide guidelines for preamble and feedback design; details of derivation can be found in [7]. Representative simulation results are provided.

The rest of the paper is outlined as follows. The channel model is given in Section 2. In Section 3, bounds on channel capacity with only imperfect CSIR, obtained via a preamble, are derived. In Section 4, the proposed feedback strategies are introduced and lower bounds are derived. Simulation results are presented in Section 5.

2. Channel Model

We will adopt the channel model considered in [3]. Assume that the system is equipped with $M$ transmit and $N$ receive antennas. We consider an i.i.d. Rayleigh fading model, which implies that the channel coefficients form a random $N \times M$ matrix, which is unknown to both transmitter and receiver. The coefficient matrix is assumed to stay constant for $T$ consecutive symbols, and change independently to a new realization in the next coherence interval. With the above assumptions, the complex baseband representation of the received signal is given by

$$
Y = HS + W
$$

where $S \in \mathbb{C}^{M \times T}$ is the matrix of the transmitted signal block and $H \in \mathbb{C}^{N \times M}$ is the channel gain matrix. The channel gain from the $m^{th}$ transmit antenna to the $n^{th}$ receive antenna, $h_{mn}$, $m = 1, \ldots, M$, $n = 1, \ldots, N$ are assumed to be i.i.d. circularly symmetric Gaussian variables, $\mathcal{CN}(0, 1)$,

$$
p(h_{mn}) = \frac{1}{\pi} \exp \left\{ -|h_{mn}|^2 \right\}
$$

Finally, the additive noise, $W \in \mathbb{C}^{N \times T}$, is assumed to have i.i.d. entries, $w_{nt} \sim \mathcal{CN}(0, \sigma^2)$. The transmitter is assumed to be equipped with a total expected power equal to $P$, so the power constraint can be written as

$$
\mathbb{E}\left\{ \text{trace}(S^H S) \right\} \leq TP
$$

where $S^H$ denotes the Hermitian conjugate of the complex matrix $S$. Throughout the sequel, random variables are denoted by boldface letters (e.g., $H$), and instances of random variables will be represented by normal face letters (e.g., $H$).

†Results for feedback channels in frequency division duplex mode can be derived easily from the results in this paper.
3. Channel Capacity with Imperfect State Information at Receiver

In this section, we will derive lower bounds on mutual information when the receiver has imprecise channel state information. The effect of imperfect side information has also been considered recently in [8] for $M = N = 1$; our results were derived independently and extend the results in [8] to multiple antennas. Following result from [9] for perfect CSIR and no CSIT will be used in the sequel.

Theorem 1 ([9]) The capacity of the channel (1) is achieved by a complex Gaussian input with zero mean and covariance $\frac{1}{\sigma_p^2}I_M$. The capacity is given by $E[\log \det (I_N + \frac{1}{\sigma_p^2}HH^H)]$.

3.1. Imperfect Channel Information at Receiver

To obtain channel state information at the receiver, a commonly used method is to transmit a preamble at the start of each coherence interval. The receiver uses the preamble to estimate the unknown channel coefficients. The estimated channel coefficients are then used to decode the transmitted codewords. In [3], it was shown that there is no capacity gain in making the number of transmit antennas, $M$, more than the length of the coherence interval, $T$; the same applies to a preamble based scheme and hence we assume $M \leq T$.

Since there are $NM$ unknown channel coefficients, at least $NM$ independent measurements are needed to determine the channel coefficients with finite estimation variance. Assume that $T_p$ symbols are used for preamble, and denote the preamble symbol matrix by $G_{M \times T_p}$. If $M \leq T$, then from (1), it immediately follows that the column rank of $G$ has to be greater than $M$ to ensure that all channel coefficients can be estimated with finite variance. If $T_p = M$, then the $M \times M$ identity matrix, $I_M$, is a possible choice for the preamble matrix $G$, which ensures that all channel coefficients are estimated with the same variance.

Since the capacity decreases logarithmically in power and linearly in time, we choose to transmit only the minimum required length of the preamble, $M$. To tighten our lower bound on mutual information, we will transmit the preamble with different power as compared to the rest of the data. For some cases, this may increase the peak to average power requirements of the transmitter\(^2\). We will restrict our attention to the case when $G = \sqrt{\frac{T_p}{M}}I$. This implies that the received signal during the preamble phase of the transmission is given by

$$Y_{N \times M} = H_{N \times M} G_{M \times M} + W_{N \times M}$$

$$= \sqrt{\frac{T_p}{M}}H + W$$

(4)

The minimum mean square channel estimate is given by

$$\hat{h}_{m,n} = \sqrt{\frac{T_p}{M}}y_{m,n}.$$

(5)

The estimation error, $\delta_{m,n} = h_{m,n} - \hat{h}_{m,n}$ is Gaussian distributed with zero mean and the variance;

$$E[\delta_{m,n}^2] = \frac{\sigma_p^2}{\frac{T_p}{M} + \sigma^2} = \frac{\sigma_p^2}{1 + \sigma^2}.$$  \hspace{1cm} (6)

\(^2\)If the transmitter has to operate under a strict peak to average power constraint, then a longer preamble may be required to tighten the lower bounds on mutual information.

The quantity $\sigma_p^2 = \frac{4}{\gamma_m^2}$ is the Cramér-Rao bound (CRB) for estimating $H_{m,n}$ for the above specified preamble. Note that the particular choice of preamble matrix, $G = \sqrt{T_p}I$, transmits BPSK symbols. Since the channel is complex, QPSK symbols can also be used as a preamble; a QPSK preamble will reduce the CRB by half.

Having estimated the channel coefficients, the channel decoder uses the estimates to decode the transmitted codewords. The channel decoder is assumed to implement the optimal decoding rule with the complete knowledge of joint distribution of the fading process and the channel estimates\(^3\). The capacity of the fading channel with noisy side information is given by [2, 11],

$$C_1 = \max_{p \in [0, 1]} \mathbb{E}[\log (I(Y; S|\hat{H} = \hat{H})].$$

Note that $S \in \mathbb{C}^{M \times (T - M)}$ since $M$ symbols have been used for the preamble. In each coherence interval, the preamble matrix is followed by coded data, and the received signal can be written as

$$Y = \hat{H}S + (H - \hat{H})S + W = \hat{H}S + \tilde{W}.$$  

(7)

Since $\tilde{W}$ is not Gaussian, computation of capacity is non-trivial. We lower bound the capacity by using an i.i.d. input $S$ with $s_m \sim \mathcal{CN}(0, \frac{1}{\sigma_p^2})$, where $P_d$ is the total power used to transmit the data and $P_d = \frac{\gamma_p^2}{\sigma_p^2}$. The entries of $\tilde{W}$ are uncorrelated with each other and uncorrelated with $\tilde{H}S$, with zero mean and variance,

$$\sigma^2 = \sigma^2 + \frac{\sigma_p^2}{1 + \gamma_p^2}P_d = \sigma^2 + \gamma_p^2P_d.$$  

(8)

Using Theorem 1, a lower bound on mutual information $I(Y; S|\hat{H})$ can be derived and is given by,

$$C_1 \geq \frac{T - M}{T} \mathbb{E} \left[ \log \det \left( I_N + \frac{\gamma_p^2}{M(\sigma^2 + \gamma_p^2P_d)}HH^H \right) \right].$$  

(9)

Further simplification of (9) can be obtained using results in [9]. The lower bound in (9) can be made tighter by choosing $P_d$ to maximize the right hand side in the inequality.

4. Channel Capacity with Imperfect State Information at Transmitter and Receiver

Our primary motivation of considering feedback based schemes comes from the capacity gains achievable in fading channels, when both transmitter and receiver have perfect channel state information. In fact, the capacity of the fading channel with perfect CSIR and CSIT (CSIR&T) goes to infinity as the number of transmit antennas goes to infinity, i.e., $M \to \infty$. In Figure 1, the behaviour of capacity with different amount of channel side information at the receiver and transmitter is shown. The bottom solid curve (with circles) in Figure 1 is the perfect CSIR&T capacity multiplied by the factor $(T - M - 1)/T$; this factor accounts for $M$ symbols of preamble and one symbol for feedback. Even with this adjustment for preamble and feedback, the adjusted perfect CSIR&T shows significant gains over perfect CSIR capacity. Thus, we seek coding schemes, with feedback from receiver,\(^3\) the effect of decoder mismatch is analyzed for imperfect CSIR and no CSIT.
which may achieve higher rates compared to the perfect CSIR case.

For low mobility applications, like indoor wireless large area networks and Bluetooth\(^4\), large values of coherence intervals are possible. The following result provides an additional motivation to investigate feedback based coding schemes for systems with large coherence intervals.

**Theorem 2** (Large Coherence Interval Limit,[7])
The achievable capacity of a multiple antenna system with no CSIR&T using feedback coding schemes approaches the perfect CSIR&T bound as \( T \to \infty \).

Thus, as the coherence interval becomes large, we are not limited by perfect CSIR channel capacity if the receiver feedback is available. With the availability of receiver feedback, the achievable rates of a system with no CSI is upper bounded by perfect CSIR and CSIT capacity.

From [9, 12], the capacity achieving scheme for perfect CSIR&T can be seen as a concatenation of the optimal encoder for an additive white Gaussian channel and an optimal beamformer which is different for each coherence interval. But optimal coding scheme for the case of imperfect CSIR&T is no longer a decoupled Gaussian encoder and beamformer [11, 13]. However, we will only investigate simpler suboptimal techniques which use a Gaussian encoder followed by a beamformer.

The receiver is assumed to allocate a finite average power, \( P_r \), to assist the transmitter in learning the channel. Since the feedback is required to be causal, it is clear that all the feedback channel resources should be used to send the CSI in the shortest possible time.

### 4.1. Feedback in TDD Mode

The feedback schemes considered in this section are most appropriate when \( N + M \ll T \). For simplicity of presentation, we restrict ourselves to the case of \( M \geq 1 \) transmit and \( N = 1 \) receive antenna.

#### 4.1.1. Low-rate quantized feedback

After measuring the channel using a preamble, the receiver has estimates of all \( M \) channel coefficients, \( \hat{\mathbf{H}} \). The receiver uses one symbol to transmit back the feedback information, which is denoted by \( f(\hat{\mathbf{H}}) \in \{1, \ldots, 2^b\} \). We begin by assuming that the entropy of the feedback \( f(\cdot) \) is low enough that it can be received at the transmitter with negligibly low probability of error. We further assume a specific form of feedback information, motivated by observing (9). For \( N = 1 \), the capacity lower bound depends on channel estimates only via \( \lambda = \hat{\mathbf{H}}^H \mathbf{H} = \sum_{m=1}^M |\hat{h}_m|^2 \), if equal power is transmitted on all the transmitter antennas. Thus, we assume that the receiver quantizes the total received energy \( \lambda \) and sends it to the transmitter\(^5\).

With the above assumed form of the vector quantizer, the transmitter uses equal power on all antennas, \( P_r \), when the feedback information is \( f(\hat{\mathbf{H}}) = i \). The computation of mutual information in closed form appears to be intractable, so we resort to the lower bounds derived in the previous section. Following [1, 2], we conclude

\[
C_{1,\text{quant}} \geq \frac{T - M - 1}{T} \max \lambda \left[ \log \left( 1 + \frac{P_r}{M(\sigma^2 + \gamma^2 P_r)}\lambda \right) \right].
\]  

\(^4\)Bluetooth is the new inter-device communication protocol.

\(^5\)Geometrically, the above form of feedback corresponds to vector quantization of the positive orthonormal, \( \mathbb{R}^+ \), by hyperplanes parallel to the simplex \( \sum_{i=1}^M |\hat{h}_m|^2 = 1 \).

where \( C_{1,\text{quant}} \) is the capacity with imperfect CSIR and quantized noiseless CSIT, and the maximum is computed over all power control policies which satisfy the power constraint, \( E_f(\hat{\mathbf{H}})(P_r) \leq P \). The multiplicative factor \( M^{M-1} \) results from using \( M \) preamble symbols every coherence interval and allocating one symbol for feedback. Similar to the suggestion in Section 3, the lower bound can be made tighter by optimizing over the transmit power used for preamble.

#### 4.1.2. Preamble transmitted by the receiver

In this scheme, instead of transmitting quantized CSIR, the receiver transmits a single training symbol. Since \( N = 1 \), one preamble symbol is enough to obtain finite variance channel estimates at the transmitter. The transmitter forms the MMSE estimate of the channel, \( \mathbf{H} \), similar to (5), and uses it for power control. Let \( \lambda = \mathbf{HH}^H \), \( \lambda = \mathbf{HH}^H \) and \( \gamma = \mathbf{HH} \). Assuming that the transmitter uses the beamforming similar to the optimal [9], the received signal in the communication phase is given by

\[
Y = \sqrt{\frac{P(\lambda)}{M}} \mathbf{H} \mathbf{H}^H S + W
\]

\[
= \sqrt{\frac{P(\lambda)}{M}} \mathbf{H} \mathbf{H}^H S + \left( \sqrt{\frac{P(\lambda)}{M}} \mathbf{H} \mathbf{H}^H - \sqrt{\frac{P(\lambda)}{M}} \mathbf{H} \mathbf{H}^H \right) S + W
\]

\[
= \sqrt{\frac{P(\lambda)\lambda}{M}} S + \frac{1}{\sqrt{M}} \left( \sqrt{P(\lambda)\gamma} - \sqrt{P(\lambda)\lambda} \right) S + W. \quad (11)
\]

where \( P(\lambda) \) is the transmitted power and \( P(\lambda) \) is the power estimate at the receiver. The resulting capacity lower bound assuming \( S \sim \mathcal{CN}(0, 1) \) is given by

\[
C_{1,\text{preamble}} \geq \frac{T - M - 1}{T} \max \lambda \left[ \log \left( 1 + \frac{P_r}{M(\sigma^2 + \gamma^2 P_r)}\lambda \right) \right]
\]

where \( \sigma^2_{\mathbf{H|H}} = E \{|U - E\{U\}|^2|\mathbf{H}\} \). The maximum in (12) is computed over all possible power control policies, \( P(\lambda) \) such that \( E\{P(\lambda)\} \leq P_r \). The lower bound can be further tightened by optimizing for preamble power. Further simplifications of (12) will be provided in [7].

### 4.2. Feedback in Two-way Communication Mode

An alternate causal feedback channel can be obtained by formulating the problem as a two-way communication problem, first introduced by Shannon [6]. In two-way communication, both ends simultaneously transmit and receive information. Note that both ends use the same spectrum for their respective transmissions. Unlike the two-way channel studied in [6], the two-way channel studied in this section sends data only in one direction (from transmitter to receiver). The reverse link, from receiver to transmitter, is used to transmit only the channel state information.

#### 4.2.1. Low-rate successively refined quantized feedback

The feedback information is successively refined leading to a power control which is continuously updated in a coherence interval. After sending a preamble of length \( M \), the transmitter starts sending coded data immediately. After each transmitted symbol, the newly received feedback information is used to update the power level for the next symbol. The lower bound for this method can be derived similar to (10) and is thus not presented here (see [7] for details).
4.2.2. Simultaneous preamble transmission

In this scheme, both transmitter and receiver start sending preambles simultaneously at the start of each coherence interval. After $M$ symbols, the receiver has an estimate of the channel, $\mathbf{H}$, as derived in (3). After $N$ symbol periods, the transmitter has an estimate $\hat{\mathbf{H}}$. Thus after $\max\{M, N\}$ symbol durations, both transmitter and receiver have an estimate of the channel. Note that the transmitter starts transmitting data immediately after sending the preamble with power adaptation based on partial channel knowledge. The channel information at the transmitter is incomplete for only the first $\min\{N, 0 - M\}$ data symbols. Again, the derivation of the lower bounds is similar (12) and is given in [7].

4.3. Numerical Results

In this section, two representative examples of the derived bounds are presented. In Figure 2-3, we compare four bounds, (a) capacity with perfect CSIR and CSIT (9), (b) capacity with perfect CSIR (9), (c) lower bound with 1-bit of feedback (10), and (d) lower bound (9) with no feedback and using a preamble. The 1-bit feedback bound is computed using a quantizer, which chooses the threshold 'a' for each such that $\text{Prob}(\lambda > a) = \text{Prob}(\lambda < a)$. Note that as $T$ increases, the 1-bit noiseless feedback scheme can gain compared to the perfect CSIR capacity, which bounds the capacity with no CSIR&T [3]. This is in accordance with the intuition derived from Proposition 2.

Feedback also has an additional benefit, evident from Figures 2-3. For large $T$, to achieve a desired capacity, a feedback based scheme can use fewer antennas compared to a non-feedback based method, thereby reducing system cost considerably; we label this gain as feedback antenna gain. Equivalently, it leads to a reduced SNR requirements to achieve a given capacity level, which we label as feedback SNR gain.

5. Conclusions

Motivated by gains of perfect CSIT&R in multiple antenna fading channels, we investigated the achievable rates of methods based on practical channel estimation and feedback strategies. A significant advantage of the proposed feedback scheme is not only a gain in achievable rates, but the codebooks used to achieve it. Bounds based on Gaussian codebooks were derived, since practical codes based on Gaussian