

# Capacity Analysis of MIMO Systems with Unknown Channel State Information

Jun Zheng and Bhaskar D. Rao

Dept. of Electrical and Computer Engineering  
University of California at San Diego

e-mail: juzheng@ucsd.edu, brao@ece.ucsd.edu

**Abstract** — We consider a mobile wireless communication system composed of  $M$  transmit and  $N$  receive antennas operating in a fading environment. Assuming channel state information is unavailable to the transmitter and the receiver, a capacity upper bound of the unknown MIMO channel under the assumption of restricted input distributions is provided. By analyzing the proposed capacity upper bounds, we reinforce the advantages of using an orthogonal pilot structure which minimizes the mean square estimation error, in that it also maximizes the capacity upper bounds. Interestingly, the capacity upper bound is shown to be a monotonically decreasing function with respect to the number of pilot symbols  $T_\tau$ . Numerical evaluations of the capacity upper bound further demonstrate that the capacity gain is insignificant when the number of pilot symbols  $T_\tau$  decreases below  $M$ , suggesting an optimum training duration of  $M$  time slots.

## I. INTRODUCTION

Communication systems using multiple antennas at both the transmitter and the receiver has recently received increased attention due to its capability of providing great capacity increases in a wireless fading environment, as reported by Telatar [1] and Foschini [2]. However, the capacity analysis provided is based on the underlying assumption that the fading channel coefficients between each transmit and receive antenna pairs are perfectly known at the receiver without any cost, which is not a reasonable assumption for most practical communication systems especially when the fading channel is changing fast.

Marzetta and Hochwald provide in [3] the capacity analysis of an unknown MIMO channel with a finite coherent time interval  $T$ . They showed that the capacity is achieved when the transmitted signal matrix is equal to the product of an isotropically distributed unitary matrix times a random diagonal matrix with real, nonnegative diagonal elements. Furthermore, Zheng and Tse [4] compute the asymptotic capacity of this channel at high signal to noise ratios.

However, in practice not only finding the optimal input distribution is an involved task and requires numerical optimizations, but also there are no known space-time codes that can approach this capacity. Hence, this paper takes a more pragmatic approach and focuses on systems that are able to take advantages of the existing channel estimation algorithms and

the powerful forward error correction coding techniques, like turbo or LDPC codes. Hassibi and Hochwald propose in [5] a channel model that separates one coherent block into two phases: training and data. Based on the two phase channel model, as well as by applying the MMSE channel estimation algorithm, they provide a capacity lower bound for the unknown MIMO channel, and prove that the optimal number of training symbols is equal to  $T_\tau = M$  when the training and data powers are allowed to vary. The capacity lower bound provided in [5] assumes that the channel estimation (LMMSE) is obtained by only using the training symbols, thereby not making use of the channel information contained in the received data symbols. Therefore, the lower bound is pessimistic and unable to represent the true capacity (or the maximum achievable information rate) accurately.

Without assuming any specific channel estimation algorithm, we propose in this article a capacity upper bound for the unknown MIMO channel with a two-phase transmitted signal structure given in [5]. By analyzing the proposed capacity upper bounds with respect to different system parameters, we show that the orthogonal pilots structure not only minimizes the mean square estimation error, but also maximizes the capacity upper bounds. Furthermore, we also prove that the channel capacity upper bound is a monotonically decreasing function with respect to the number of pilot symbols  $T_\tau$ , but with insignificant capacity increment when  $T_\tau$  decreases below the number of transmit antennas  $M$ , which is verified by numerical evaluation.

## II. SYSTEM MODEL

We consider in this article a MIMO system with  $M$  transmitter antennas and  $N$  receive antennas, signaling through a frequency flat fading channel with i.i.d channel coefficient between each transmit and receive antenna pairs. It is assumed that the fading coefficient  $\mathbf{H}$  remains static within a coherent time interval of  $T$  symbol periods, and varies independently from one coherent time block to another; and each element of  $\mathbf{H}$  is complex Gaussian distributed. Hence, the signaling model can be written as

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{H} + \mathbf{w} \quad , \quad (1)$$

where  $\mathbf{Y}$  is a  $T \times N$  received complex signal matrix,  $\mathbf{X}$  is a  $T \times M$  transmitted complex signal matrix,  $\mathbf{H}$  is a  $M \times N$  complex channel matrix, and  $\mathbf{w}$  is a  $T \times N$  matrix of additive Gaussian noise. Both matrix  $\mathbf{H}$  and  $\mathbf{w}$  have zero mean unit variance independent complex Gaussian entries. We also assume that

<sup>1</sup>This research was supported by CoRe grant No. 02-10109 sponsored by Ericsson.

the entries of the transmitted signal matrix  $\mathbf{X}$  have, on average, the following power constraint,

$$\frac{1}{T} \cdot E[\text{tr}(\mathbf{X}^H \mathbf{X})] = \rho, \quad (2)$$

where  $\rho$  is the average signal to noise ratio at each receive antenna. The transmitted signal matrix  $\mathbf{X}$  is further separated into two submatrixes: training followed by data, which is represented as

$$\mathbf{X} = \begin{bmatrix} (\rho_\tau/M)^{\frac{1}{2}} \cdot \mathbf{S}_\tau \\ (\rho_d/M)^{\frac{1}{2}} \cdot \mathbf{X}_d \end{bmatrix}, \quad (3)$$

where  $\mathbf{S}_\tau$  is the fixed pilot symbols and  $\mathbf{X}_d$  is the information bearing data symbols, whose structures are given by

$$\begin{aligned} \mathbf{S}_\tau &= [\mathbf{s}_1^H, \dots, \mathbf{s}_{T_\tau}^H]^H, & \mathbf{S}_\tau &\in \mathbb{C}^{T_\tau \times M}, \\ \mathbf{X}_d &= [\mathbf{x}_1^H, \dots, \mathbf{x}_{T_d}^H]^H, & \mathbf{X}_d &\in \mathbb{C}^{T_d \times M}. \end{aligned} \quad (4)$$

Conservation of time and energy leads to the following constraints,

$$\begin{aligned} \text{tr}(\mathbf{S}_\tau^H \cdot \mathbf{S}_\tau) &= MT_\tau, & E_{\mathbf{X}_d}[\text{tr}(\mathbf{X}_d^H \cdot \mathbf{X}_d)] &= MT_d, \\ T &= T_\tau + T_d, & \rho T &= \rho_\tau T_\tau + \rho_d T_d. \end{aligned} \quad (5)$$

### III. CAPACITY ANALYSIS

#### A. Restricted Capacity Upper Bound

From the capacity analysis result provided in [3], we know that not only finding the capacity achieving input distribution is an involved task and requires numerical optimization, but also there are no known viable space-time codes that can approach this capacity.

In this section we restrict our attention to a conventional MIMO system having an input signal structure, which is described in Section II. It is further assumed that the input data vectors  $\mathbf{x}_i$  follow i.i.d. Gaussian distributions, i.e.

$$E[\mathbf{x}_i^H \cdot \mathbf{x}_j] = \delta_{i,j} \cdot I_M. \quad (6)$$

Although input distribution (6) is not optimized to achieve the maximum mutual information rate, it is a reasonable assumption of a communication system with no channel state information available at the transmitter. Therefore, under this restricted input distribution assumption, we have the following MIMO capacity upper bound. All the proofs provided are brief due to length restrictions. Please refer to [6] for more details.

**Proposition 1** *Mutual information between  $\mathbf{X}$  and  $\mathbf{Y}$  are upper bounded by,*

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}) &\leq \bar{C} = N \left( \log_2 \left| I_M + \frac{\rho_\tau}{M} \mathbf{S}_\tau^H \mathbf{S}_\tau \right| + T_d \cdot \log_2(1 + \rho_d) \right. \\ &\quad \left. - E_{\mathbf{X}_d} \left[ \log_2 \left| I_M + \frac{\rho_\tau}{M} \mathbf{S}_\tau^H \mathbf{S}_\tau + \frac{\rho_d}{M} \mathbf{X}_d^H \mathbf{X}_d \right| \right] \right), \end{aligned} \quad (7)$$

where the expectation  $E_{\mathbf{X}_d}[\cdot]$  is taken with respect to data  $\mathbf{X}_d$ .

**Proof:** First, conditioned on any input data sequences  $\mathbf{X}_d$  (or  $\mathbf{X}$ ),  $\text{vec}(\mathbf{Y})$  is a Gaussian distributed vector of zero mean and variance

$$\Sigma_{\mathbf{Y}|\mathbf{X}} = \text{Cov}(\text{vec}(\mathbf{Y})|\mathbf{X}) = I_N \otimes (\mathbf{X}\mathbf{X}^H + I_T). \quad (8)$$

Taking expectation of (8) with respect to  $\mathbf{X}_d$ , the covariance matrix of  $\text{vec}(\mathbf{Y})$  is obtained as,

$$\Sigma_{\mathbf{Y}} = \text{Cov}(\text{vec}(\mathbf{Y})) = I_N \otimes (\Sigma + I_T), \quad (9)$$

where

$$\Sigma = \begin{bmatrix} \frac{\rho_\tau}{M} \cdot \mathbf{S}_\tau \mathbf{S}_\tau^H & \mathbf{0} \\ \mathbf{0} & \rho_d \cdot I_{T_d} \end{bmatrix}. \quad (10)$$

Due to the fact that Gaussian distribution has the maximum entropy among any vector distributions with the same covariance matrix, entropy  $h(\mathbf{Y})$  can be upper bounded by

$$h(\mathbf{Y}) \leq \log_2 \left( (\pi e)^{NT} \cdot |\Sigma_{\mathbf{Y}}| \right). \quad (11)$$

Therefore, we have the following capacity upper bound

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}) &= h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \stackrel{(a)}{\leq} \log_2 \left( (\pi e)^{NT} \cdot |\Sigma_{\mathbf{Y}}| \right) \\ &\quad - E_{\mathbf{X}_d} \left[ \log_2 \left( (\pi e)^{NT} \cdot |\Sigma_{\mathbf{Y}|\mathbf{X}}| \right) \right], \end{aligned} \quad (12)$$

where the second term of inequality (a) is from a direct expansion of the conditional entropy according to the definition. Substituting (8) and (9) into (12), capacity upper bound (7) can be obtained in a straightforward manner.

Since the received signal  $\mathbf{Y}$  can be viewed as a weighted sum of Gaussian random vectors, its distribution is close the Gaussian as long as it contains a large number of independent random variables according to the central limit theorem. Hence, the upper bound is tight and quite likely to be even less than the true unknown MIMO channel capacity provided in [3]. Therefore, maximizing the capacity upper bound is a reasonable approach and will not make the bound become loose due to the fact that both the capacity upper bound as well as the mutual information  $I(\mathbf{X}; \mathbf{Y})$  in (7) increases through the optimization with respect to different system parameters.

#### B. Pilot Structure Optimizations

The most commonly used pilots have an orthogonal structure. They are optimal in a sense that they minimize the mean square channel estimation error as well as achieve the Cramer-Rao lower bound. As a straightforward extension of the single input single output system, the covariance matrix of the MMSE estimation error  $\tilde{\mathbf{H}} = \mathbf{H} - \hat{\mathbf{H}}$  for the unknown MIMO channel is given by

$$\mathbf{C}_{\tilde{\mathbf{H}}, \tilde{\mathbf{H}}} = \text{Cov}(\text{vec}(\tilde{\mathbf{H}}), \text{vec}(\tilde{\mathbf{H}})) = I_N \otimes \left( I_M + \frac{\rho_\tau}{M} \mathbf{S}_\tau^H \mathbf{S}_\tau \right)^{-1}. \quad (13)$$

It is obvious that the average mean square error of the channel estimation

$$\text{tr}(\mathbf{C}_{\tilde{\mathbf{H}}, \tilde{\mathbf{H}}}) = N \times \text{tr} \left( \left( I_M + \frac{\rho_\tau}{M} \mathbf{S}_\tau^H \mathbf{S}_\tau \right)^{-1} \right), \quad (14)$$

is minimized when the non-zero eigenvalues of  $\mathbf{S}_\tau^H \mathbf{S}_\tau$  are all equal. Therefore, the following orthogonal pilot structure, represented as

$$\begin{aligned} \mathbf{S}_\tau^H \mathbf{S}_\tau &= T_\tau \cdot I_M, & T_\tau \geq M, \\ \mathbf{S}_\tau \mathbf{S}_\tau^H &= M \cdot I_{T_\tau}, & T_\tau < M, \end{aligned} \quad (15)$$

minimizes the MIMO MMSE channel mean square estimation error.

In order to obtain the optimal pilot structure with respect to the capacity upper bound  $\bar{C}$ , we utilize the following concavity property.

**Proposition 2** *The capacity upper bound obtained in Proposition 1 is concave with respect to matrix  $\mathbf{Q} = \mathbf{S}_\tau^H \mathbf{S}_\tau$ , i.e.,*

$$\lambda \bar{C}(\mathbf{Q}_1) + (1-\lambda) \bar{C}(\mathbf{Q}_2) \leq \bar{C}(\lambda \cdot \mathbf{Q}_1 + (1-\lambda) \cdot \mathbf{Q}_2), \quad (16)$$

where  $\lambda \in [0, 1]$ .

**Proof:** The Hessian of the capacity upper bound is given by

$$\frac{\partial^2 \bar{C}}{\partial^2 \text{vec}(\mathbf{Q})} = -\frac{N \rho_\tau^2}{M^2 \ln 2} \cdot E_{\mathbf{X}_d} \left[ \boldsymbol{\Sigma}_1^{-1} \otimes \boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\Sigma}_2^{-2} \otimes \boldsymbol{\Sigma}_2^{-1} \right], \quad (17)$$

where  $\boldsymbol{\Sigma}_1$  and  $\boldsymbol{\Sigma}_2$  are given by

$$\boldsymbol{\Sigma}_1 = I_M + \frac{\rho_\tau}{M} \mathbf{Q}, \quad \boldsymbol{\Sigma}_2 = I_M + \frac{\rho_\tau}{M} \mathbf{Q} + \frac{\rho_d}{M} \mathbf{X}_d^H \mathbf{X}_d. \quad (18)$$

It can be shown that the Hessian matrix is negative semi-definite, and hence the capacity upper bound is concave with respect to  $\mathbf{Q}$ .

As a direct result of Proposition 2, we have the following optimal pilot structure.

**Proposition 3** *The optimal pilot structure, which maximizes the capacity upper bound (7), satisfies the following orthogonal conditions*

$$\mathbf{Q} = \mathbf{S}_\tau^H \mathbf{S}_\tau = \frac{M T_\tau}{\min(T_\tau, M)} \left[ \begin{array}{c|c} I_{\min(T_\tau, M)} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right], \quad (19)$$

which is equivalent to (15).

**Proof:** First, substituting (18) into (7), the capacity upper bound can be represented as

$$\bar{C} = N \left( T_d \cdot \log_2(1 + \rho_d) + \log_2 |\boldsymbol{\Sigma}_1| - E_{\mathbf{X}_d} \left[ \log_2 |\boldsymbol{\Sigma}_2| \right] \right). \quad (20)$$

The following equality is true

$$\begin{aligned} E_{\mathbf{X}_d} \left[ \log_2 |\boldsymbol{\Sigma}_2(\mathbf{Q})| \right] &= E_{\mathbf{X}_d} \left[ \log_2 |\mathbf{U}^H \boldsymbol{\Sigma}_2 \mathbf{U}| \right] \\ &= E_{\mathbf{X}_d} \left[ \log_2 \left| I_M + \frac{\rho_\tau}{M} \mathbf{U}^H \mathbf{Q} \mathbf{U} + \frac{\rho_d}{M} (\mathbf{X}_d \mathbf{U})^H (\mathbf{X}_d \mathbf{U}) \right| \right] \\ &\stackrel{a}{=} E_{\mathbf{X}_d} \left[ \log_2 \left| I_M + \frac{\rho_\tau}{M} \mathbf{U}^H \mathbf{Q} \mathbf{U} + \frac{\rho_d}{M} \mathbf{X}_d^H \mathbf{X}_d \right| \right] \\ &= E_{\mathbf{X}_d} \left[ \log_2 |\boldsymbol{\Sigma}_2(\mathbf{U}^H \mathbf{Q} \mathbf{U})| \right], \end{aligned} \quad (21)$$

where  $\mathbf{U}$  is any unitary matrix, and (a) follows from the fact that  $\mathbf{X}_d \mathbf{U}$  has the same distribution as  $\mathbf{X}_d$ . Further due to the fact that

$$|\boldsymbol{\Sigma}_1(\mathbf{Q})| = |\boldsymbol{\Sigma}_1(\mathbf{U}^H \mathbf{Q} \mathbf{U})|, \quad (22)$$

the capacity upper bound (20) is hence invariant under the following transformation

$$\bar{C}(\mathbf{Q}) = \bar{C}(\mathbf{U}^H \mathbf{Q} \mathbf{U}). \quad (23)$$

If the unitary matrix  $\mathbf{U}$  is set to be composed of the eigenvectors of  $\mathbf{Q}$ , then according to (23) we only need to focus our attention on the case where  $\mathbf{Q}$  is a diagonal matrix.

Furthermore, it is also true that any permutations on the non-zero diagonal elements of  $\mathbf{Q}$  will not change the upper bound,

$$\bar{C}(\mathbf{Q}) = \frac{1}{K!} \sum_{\mathbf{P}} \bar{C}(\mathbf{P}^H \mathbf{Q} \mathbf{P}) \stackrel{a}{\leq} \bar{C} \left( \frac{1}{K!} \sum_{\mathbf{P}} \mathbf{P}^H \mathbf{Q} \mathbf{P} \right), \quad (24)$$

where  $K = \min(T_\tau, M)$ , matrix  $\mathbf{P}$  is any permutation matrix that permutes the first  $K$  rows (or columns) of the non-zero elements of  $\mathbf{Q}$ , and (a) follows from the concavity property of the upper bound. At this point, it is evident that the optimal pilot, which achieves the maximum capacity upper bound, has an orthogonal structure given by (19).

Therefore, although starting from different perspectives, orthogonal pilots structure not only minimizes the estimation mean square error, but also maximizes the capacity upper bounds. Substituting the optimal structure (19) into the equation (7), we obtain the following capacity upper bound

$$\bar{C} = N \left( T_d \cdot \log_2(1 + \rho_d) - E_{\mathbf{X}_d} \left[ \log_2 \left| I_M + \frac{\rho_d}{M} \boldsymbol{\Lambda}^{-1} \mathbf{X}_d^H \mathbf{X}_d \right| \right] \right), \quad (25)$$

where  $\mathbf{X}_d$  is of size  $\mathbb{C}^{T_d \times M}$ , and  $\boldsymbol{\Lambda}$  is given by

$$\boldsymbol{\Lambda} = \left[ \begin{array}{c|c} \left( 1 + \frac{\rho_\tau T_\tau}{\min(T_\tau, M)} \right) \cdot I_{\min(T_\tau, M)} & \mathbf{0} \\ \hline \mathbf{0} & I_{M - \min(T_\tau, M)} \end{array} \right]. \quad (26)$$

### C. Equal Training and Data Power Allocation

For some communication systems, it might not be possible to vary the power during the training slots and data slots. Hence the capacity upper bound, assuming training symbol and data symbol share the same power, is obtained by substituting the power allocations ( $\rho_\tau = \rho_d = \rho$ ) into (25). The capacity upper bound is further optimized with respect to the data allocation scheme ( $T_\tau, T_d$ ), and we have the following important result concerning the dependence on  $T_d$ , the number of data symbols.

**Proposition 4** *Capacity upper bounds under equal power allocation schemes are monotonically increasing with respect to the number of data slots  $T_d$ , i.e.*

$$\begin{aligned} \bar{C}(T_d = k) &\geq \bar{C}(T_d = k - 1), & k \leq T - M, \\ \bar{C}(T_d = T) &\geq \bar{C}(T_d = T - M). \end{aligned} \quad (27)$$

**Proof:** We begin with the first part of (27) when  $T_\tau \geq M$ , where equation (25) is reduced to

$$\bar{C} = N \left( T_d \cdot \log_2(1+\rho) - E_{\mathbf{X}_d} \left[ \log_2 \left| I_M + \frac{\rho'}{M} \cdot \mathbf{X}_d^H \mathbf{X}_d \right| \right] \right), \quad (28)$$

where  $\rho' = \rho/(1 + \rho T_\tau/M)$ . We further separate the data matrix  $\mathbf{X}_d$  into  $\mathbf{X}'_d$  and  $\mathbf{x}_{T_d}$ ,

$$\mathbf{X}_d = \begin{bmatrix} \mathbf{X}'_d \\ \mathbf{x}_{T_d} \end{bmatrix}, \quad \mathbf{X}'_d \in \mathbb{C}^{T_d-1 \times M}, \quad \mathbf{x}_{T_d} \in \mathbb{C}^{1 \times M}. \quad (29)$$

Then we have the following inequality,

$$\begin{aligned} & E_{\mathbf{X}_d} \left[ \log_2 \left| I_M + \frac{\rho'}{M} \mathbf{X}_d^H \mathbf{X}_d \right| \right] \\ &= E_{\mathbf{X}_d} \left[ \log_2 \left| I_M + \frac{\rho'}{M} (\mathbf{X}'_d{}^H \mathbf{X}'_d + \mathbf{x}_{T_d}^H \mathbf{x}_{T_d}) \right| \right] \\ &\stackrel{a}{\leq} E_{\mathbf{X}'_d} \left[ \log_2 \left| I_M + \frac{\rho'}{M} (\mathbf{X}'_d{}^H \mathbf{X}'_d + E_{\mathbf{x}_{T_d}}[\mathbf{x}_{T_d}^H \mathbf{x}_{T_d}]) \right| \right], \end{aligned} \quad (30)$$

where inequality (a) follows from the fact that  $\log|\cdot|$  is a concave function. Therefore, using assumption (6), we can obtain the following result

$$\bar{C}(T_d = k) - \bar{C}(T_d = k - 1) \geq \Delta f(\rho) \geq 0, \quad (31)$$

where

$$\Delta f(\rho) = \log_2(1+\rho) - M \log_2 \left( \frac{1 + \rho(T_\tau + 1)/M}{1 + \rho T_\tau/M} \right), \quad (32)$$

due to the fact that  $\Delta f(0) \geq 0$  and  $\Delta' f(\rho) \geq 0$ . A similar approach can be used to prove the second part of (27).

#### D. Optimization over $(\rho_\tau, \rho_d)$ and $(T_\tau, T_d)$

For communication systems where the power allocation can be varied between training and data symbols, the optimal capacity upper bound is obtained by solving the following constrained optimization problem

$$\bar{C}(T_\tau, T_d) = \max_{(\rho_\tau, \rho_d)} \bar{C}(\rho_\tau, \rho_d, T_\tau, T_d), \quad \rho_\tau T_\tau + \rho_d T_d = \rho T. \quad (33)$$

However, numerical results provided in Section IV indicate that there is insignificant capacity loss by using equal power allocations, which is much easier for implementation.

## IV. NUMERICAL AND SIMULATION RESULTS

### A. Optimal Pilot Structure

We know from Section III. B that orthogonal pilot structure not only minimizes the mean square estimation error, but also maximizes the capacity upper bound (7). Fig. 1 demonstrate the sensitivity of the capacity upper bound with respect to the pilots structures. As can be observed from the plot, the capacities using random pilot structure, which are denoted as dotted curves, are inferior to those of applying orthogonal pilot structure. And the capacity loss is significant in moderate to high SNR ranges.

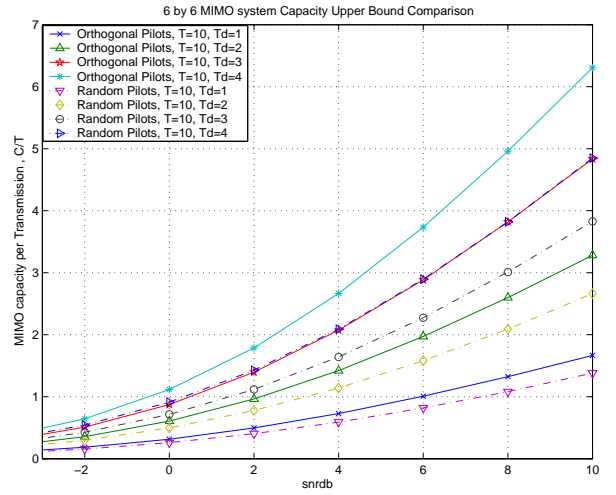


Figure 1: Capacity comparison between orthogonal pilot structures and random pilot structures under equal power allocation schemes of a  $6 \times 6$  MIMO system with coherent time intervals  $T = 10$ , and data interval  $T_d = 1, 2, 3, 4$

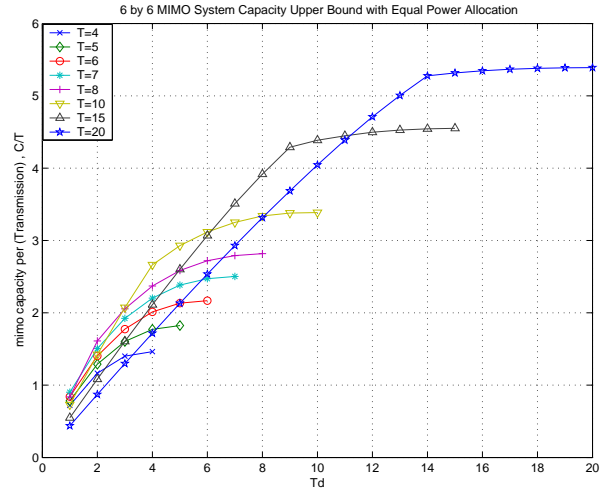


Figure 2: Capacity upper bound of a  $6 \times 6$  MIMO system under equal power allocation scheme of SNR  $\rho = 4\text{dB}$ , and with different coherent time intervals  $T = 4, 5, 6, 7, 8, 10, 15, 20$

### B. Equal Power Allocation

The capacity upper bounds of a  $6 \times 6$  MIMO system under equal power allocation scheme with different coherent time  $T$  and data slot allocation  $T_d$  are demonstrated in Fig. 2. From the plot, we can observe that the capacity upper bound is monotonically increasing with respect to  $T_d$  (even for the case  $T_d > T - M$ ). However, the capacity gain is insignificant for  $T_d$  beyond  $T - M$  especially when  $T$  is larger than  $M$ . Therefore,  $T_d = T - M, T > M$  is a good trade-off point between the achievable capacity and implementation complexity.

### C. Optimal Power Allocation

We demonstrate in Fig. 3 the capacity upper bound with

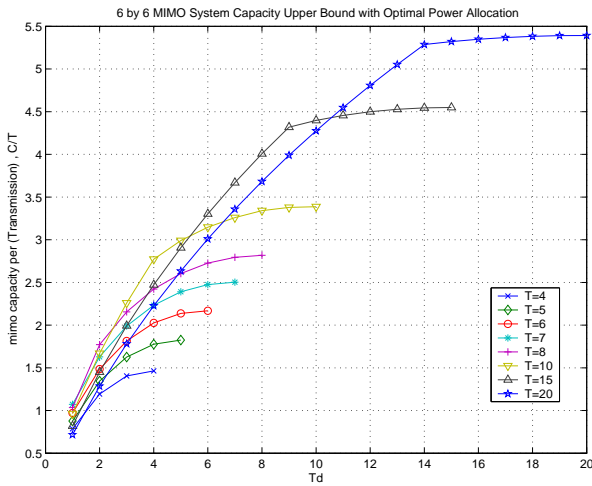


Figure 3: Capacity upper bound of a  $6 \times 6$  MIMO system under optimal power allocation scheme of SNR  $\rho = 4dB$ , and with different coherent time intervals  $T = 4, 5, 6, 7, 8, 10, 15, 20$

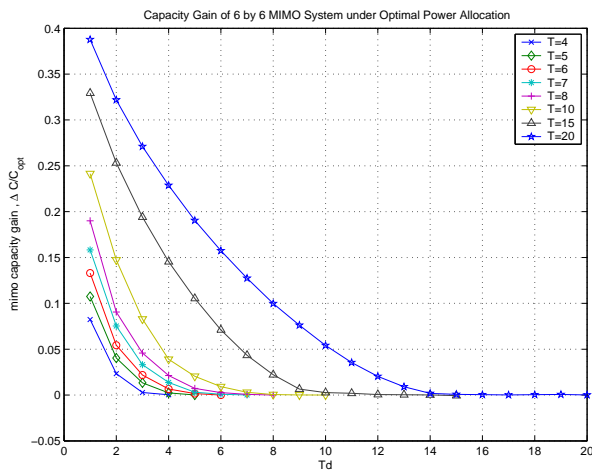


Figure 4: Capacity gain of optimal power allocations over equal power allocations of a  $6 \times 6$  MIMO system with SNR  $\rho = 4dB$ , under different coherent time intervals  $T = 4, 5, 6, 7, 8, 10, 15, 20$

optimal power allocation schemes for the same MIMO system used in the Fig. 2 with varying coherent time intervals. As a comparison, the capacity gain from using optimal power allocation over equal power allocation scheme is shown in Fig. 4. We can observe from the plot that, when the number of training symbols  $T_r$  is small, there is insignificant capacity loss due to using equal power allocations, which is much easier for implementation.

#### D. Comparison

In Fig. 5, for comparison purpose we plot the capacity lower bound provided in [5], the capacity upper bound provided in this paper, as well as the MIMO channel capacity with ideal channel state information at the receiver.

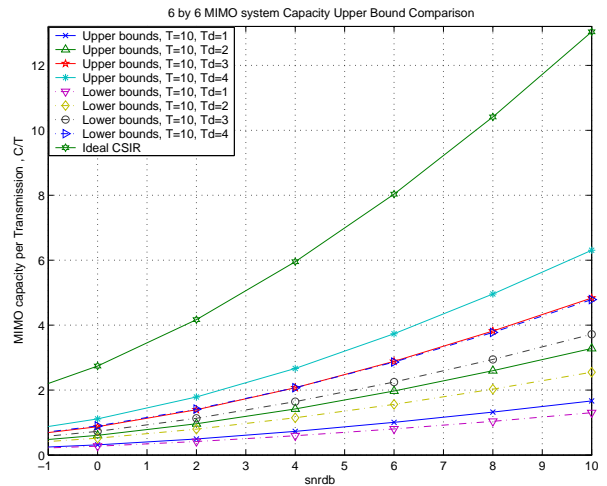


Figure 5: Capacity comparison between lower bounds and upper bounds under equal power allocation schemes of a  $6 \times 6$  MIMO system with coherent time intervals  $T = 10$ , and data interval  $T_d = 1, 2, 3, 4$

## V. CONCLUSION

In this paper, we propose a capacity upper bound of the unknown MIMO channel. Through the analysis of the proposed upper bound, we show that orthogonal pilot structure is optimal. It not only minimizes the mean square estimation error, but also maximizes the proposed capacity upper bound. We also proved that under equal power allocation scheme, capacity upper bound is a monotonically increasing function with respect to the number of data slots  $T_d$ . Through numerical evaluations, we further demonstrate that the capacity increment is insignificant when  $T_d$  is larger than  $T - M$ , and limited capacity gain can be achieved by using optimal power allocation between training and data symbols when compared to the simple equal power allocation scheme.

## REFERENCES

- [1] E. Telatar, "Capacity of multi-antenna Gaussian channels", *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585-595, 1999.
- [2] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 4159, Autumn, 1996.
- [3] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading", *IEEE Trans. on Information Theory*, vol. 45, Issue. 1, pp. 139-157, Jan, 1999.
- [4] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel", *IEEE Trans. on Information Theory*, vol. 48, Issue. 2, pp. 359-383, Feb, 2002.
- [5] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?", *IEEE Trans. on Information Theory*, vol. 49, Issue. 4, pp. 951-963, Apr, 2003.
- [6] J. Zheng and B. D.Rao, "Is Training Necessary?—Part I: Capacity Analysis of MIMO Systems with Unknown Channel State Information", *in preparation*, Apr, 2004.