OPTIMUM POWER ALLOCATION FOR MAXIMUM-LIKELIHOOD CHANNEL ESTIMATION IN SPACE-TIME CODED MIMO SYSTEMS

Chaiyod Pirak[1,3], *Z. Jane Wang*[2] *K. J. Ray Liu*[1] *Somchai Jitapunkul*[3]

[1]University of Maryland College Park
Department of Electrical &
Computer Engineering, MD, USA.
Email: chaiyod.p@gmail.com, kjrliu@eng.umd.edu

[2]University of British Columbia
Department of Electrical &
Computer Engineering, Vancouver, Canada
Email: zjanew@ece.ubc.ca

[3]Chulalongkorn University
Department of Electrical
Engineering, Bangkok, Thailand
Email: somchai.j@chula.ac.th

ABSTRACT

This paper presents an optimum power allocation strategy for the maximum likelihood based channel estimation in the space-time coded multiple-input multiple-output systems employing a data-bearing approach for pilot-embedding. The corresponding channel estimation error, the Chernoff's upper bound on the detection error probability, and a lower bound on channel capacity of such systems are analyzed. Based on such analysis, the relationship between these two bounds are revealed, then a unified optimum power allocation scheme is proposed based on jointly optimizing both bounds subject to an acceptable channel estimation error. Simulation results indicate that the proposed power allocation scheme yields a better performance in terms of the error probability, whereas the equal power allocation scheme can be reasonably used as a suboptimum approach with an acceptable performance degradation. Furthermore, the unequal power allocation with more power constantly allocated to the data part yields a much better performance than the one with more power constantly allocated to the pilot part.

1. INTRODUCTION

A space-time (ST) coded multiple-input multiple-output (MIMO) system is a prominent communication system for future wireless communications due to its diversity gain that significantly improves an error probability performance without increasing transmission power [1, 2]. Typically, for such a system employing a coherent receiver, channel state information is crucially needed. Hence, a research on channel estimation is of interest.

Recently, many channel estimation approaches [3]-[5] have been proposed for such a system by employing a pilot signal, a known signal transmitted from the transmitter to the receiver. In [4], a linear minimum mean-squared error (LMMSE) channel estimation has thoroughly been investigated, including an optimum pilot structure, an optimum power allocation for pilot and data parts, and an optimum number of the pilot signal, based on the analysis of a lower bound on channel capacity. In [5], we proposed a data-bearing approach for pilot-embedding for joint channel estimation and data detection, where a maximum-likelihood (ML) channel estimator was investigated. In [5], the optimum power allocation scheme based on minimizing a Chernoff's upper bound on error probability has been discovered for the case of square-matrix ST codes. However, there is a lack of a thorough investigation for the ML channel estimator in ST coded MIMO systems employing arbitrary-matrix ST codes, and a comprehensive comparison between different optimum power allocation strategies based on different criteria. To address these concerns, the main contributions of this paper are as follows,

- A unified optimum power allocation, based on jointly optimizing Chernoff's upper bound on error probability and the lower bound on channel capacity subject to an acceptable channel estimation error, is developed for the ML channel estimator in the ST coded MIMO systems employing arbitrary-matrix ST codes.
- The relationship between these two bounds is revealed through an effective signal-to-noise ratio (SNR).
- For the case that no information about the SNR of such systems is available to the transmitter, an equal power allocation scheme can be used as a reasonable suboptimum approach with an acceptable performance degradation.

The rest of this paper is organized as follows. In Section 2, we describe the MIMO channel and system models first, then briefly review the data-bearing approach for pilot-embedding. In section 3, performance analysis, including the channel estimation error, the Chernoff's upper bound on error probability, and the lower bound on channel capacity, is presented, and the optimum power allocation is developed in Section 4. The simulation results are shown in Section 5, and we conclude this paper in Section 6.

2. THE DATA-BEARING APPROACH FOR PILOT-EMBEDDING

2.1. MIMO Channel and System Models

We briefly describe the MIMO channel and system models used in this paper. We consider the MIMO communication system with L_t transmit antennas and L_r receive antennas. In general, for a given block index t, a ST symbol matrix $\mathbf{U}(t)$ is an $L_t \times M$ codeword matrix transmitted across the transmit antennas in M time slots. The received symbol matrix $\mathbf{Y}(t)$ at the receiver front-end can be described as follows [5],

$$\mathbf{Y}(t) = \mathbf{H}(t)\mathbf{U}(t) + \mathbf{N}(t), \tag{1}$$

where $\mathbf{H}(t)$ is the $L_r \times L_t$ channel coefficient matrix and $\mathbf{N}(t)$ is the $L_r \times M$ additive complex white Gaussian noise matrix with zero mean and variance $\frac{\sigma^2}{2}\mathbf{I}_{(ML_r \times ML_r)}$ per real dimension. The elements of channel coefficient matrix $\mathbf{H}(t)$ are assumed to be independent complex Gaussian random variables with zero mean and variance 0.5 per real dimension. Or equivalently, an independent Rayleigh fading channel is assumed. We first estimate the channel coefficient matrix $\mathbf{H}(t)$ and the ST symbol matrix $\mathbf{U}(t)$ by using the pilot or training signal embedded in the ST symbol matrix $\mathbf{U}(t)$.

2.2. The Data-Bearing Approach for Pilot-Embedding

In what follows, we summarize the data-bearing approach for pilot-embedding proposed in [5]. In this approach, the pilot signal is firstly added into the ST data, and then this signal combination is regarded as the ST symbol. The proposed pilot-embedded ST symbol matrix $\mathbf{U}(t)$ can be expressed as follows,

$$\mathbf{U}(t) = \mathbf{D}(t)\mathbf{A} + \mathbf{P}, \tag{2}$$

where $\mathbf{D}(t) \in \mathbb{C}^{L_t \times N}$ is the ST data matrix, $\mathbf{A} \in \mathbb{R}^{N \times M}$ is the data bearer matrix with N being the data time slots, and $\mathbf{P} \in \mathbb{R}^{L_t \times M}$ is the

This work was partially supported by a grant from the Cooperation Project between Department of Electrical Engineering and Private Sector for Research and Development, Chulalongkorn University, Thailand.

pilot matrix. In addition, the power constraint of $\mathrm{E}[\|\mathbf{D}(t)\|^2] = L_t$ is maintained. The necessary properties for the proposed data-bearing approach for pilot-embedding are as follows,

$$\mathbf{A}\mathbf{P}^T = \mathbf{0} \in \mathbb{R}^{N\times L_t}, \mathbf{P}\mathbf{A}^T = \mathbf{0} \in \mathbb{R}^{L_t\times N},$$
$$\mathbf{A}\mathbf{A}^T = \beta\mathbf{I} \in \mathbb{R}^{N\times N}, \mathbf{P}\mathbf{P}^T = \alpha\mathbf{I} \in \mathbb{R}^{L_t\times L_t}, \quad (3)$$

where β is a real-valued data-power factor, α is a real-valued pilot-power factor, $\mathbf{0}$ stands for an all-zero-element matrix, and $\mathbf{I}$ stands for an identity matrix.

There are, at least, three possible structures of data-bearer and pilot real-matrices that satisfy the properties in (3) including Time-Multiplexing (TM)-based matrices, ST-Block-Code (STBC)-based matrices, and Code-Multiplexing (CM)-based matrices [5]. In this paper, we study the optimum power allocation for the case of quasi-static fading channels, where $\mathbf{H}(t)$ in (1) remains constant over each symbol block but it changes block-by-block independently. This channel model makes sense in our study because we employ the channel estimate in each symbol block to decode the ST data matrix in the same symbol block, hence, the effect of power optimization on the estimation and decoding performances of receivers, i.e. the mean-squared error (MSE) and error probability, respectively, are correctly characterized. In addition, this channel model has also been studied in [4] for the LMMSE channel estimation. Since all three structures yield the same performances for quasi-static channel model [5], we use the TM-based matrices as their representative in this paper, given as

$$\mathbf{A} = \sqrt{\beta}\left[\mathbf{0}_{(N\times L_t)}; \mathbf{I}_{(N\times N)}\right],$$
$$\mathbf{P} = \sqrt{\alpha}\left[\mathbf{I}_{(L_t\times L_t)}; \mathbf{0}_{(L_t\times N)}\right],\ M = N + L_t, \quad (4)$$

where ; stands for matrix combining. In this structure, the $L_t \times L_t$ identity matrix $\mathbf{I}$ is used as a pilot or training symbol. In addition, it has been shown in [5] that the CM-based matrices are the best structure among all other structures for nonquasi-static fading channels, where $\mathbf{H}(t)$ changes according to a process whose dominant frequency is much faster than $\frac{1}{M}$ or, in the other words, $\mathbf{H}(t)$ is not constant over each symbol block. Also, it is worth mentioning that the power allocation scheme proposed later can be generally applied to different data-bearer structures in [5].

2.2.1. Maximum-Likelihood Channel Estimation

The ML channel estimation procedure can be summarized as follows:

1.) Post-multiplying the received symbol matrix $\mathbf{Y}(t)$ in (1) by the transpose of the pilot matrix $\mathbf{P}^T$; dividing the result by α; and using the properties in (3) to arrive at

$$\mathbf{y}(t) = \mathbf{h}(t) + \mathbf{n}(t), \quad (5)$$

where $\mathbf{y}(t) \triangleq \mathbf{vec}(\frac{\mathbf{Y}(t)\mathbf{P}^T}{\alpha}),\ \mathbf{h}(t) \triangleq \mathbf{vec}(\mathbf{H}(t)),\ \mathbf{n}(t) \triangleq \mathbf{vec}$ $(\frac{\mathbf{N}(t)\mathbf{P}^T}{\alpha})$ with $\mathbf{vec}(\cdot)$ being vectorizing conversion.

2.) The ML channel estimator that maximizes the log-likelihood function $\ln(p(\mathbf{y}(t)|\mathbf{h}(t)))$ is given by

$$\hat{\mathbf{h}}(t) = \mathbf{y}(t) \text{ or } \hat{\mathbf{H}}(t) = \frac{\mathbf{Y}(t)\mathbf{P}^T}{\alpha}. \quad (6)$$

2.2.2. Maximum-Likelihood Data Detection

The ML data detection procedure can be summarized as follows:

1.) Post-multiplying the received symbol matrix $\mathbf{Y}(t)$ in (1) by the transpose of the data bearer matrix $\mathbf{A}^T$; dividing the result by β; and using the properties in (3) to arrive at

$$\mathbf{Y}_1(t) = \mathbf{H}(t)\mathbf{D}(t) + \mathbf{N}_1(t), \quad (7)$$

where $\mathbf{Y}_1(t) = \frac{\mathbf{Y}(t)\mathbf{A}^T}{\beta}$ and $\mathbf{N}_1(t) = \frac{\mathbf{N}(t)\mathbf{A}^T}{\beta}$.

2.) The ML receiver is employed by computing the decision matric and deciding the codeword that minimizes this decision matric [6],

$$\{\hat{d}_k^i\} = \min_{\{d_k^i\}}\left\{\sum_{k=1}^{N}\sum_{j=1}^{L_r}|y_k^j - \sum_{i=1}^{L_t}\hat{h}_{j,i}d_k^i|^2\right\},$$
$$\forall d_k^i,\ i\in\{1,\dots,L_t\},\ k\in\{1,\dots,N\}, \quad (8)$$

where y_k^j denotes the $(j,k)^{th}$ element of $\mathbf{Y}_1(t)$ in (7), $\hat{h}_{j,i}$ denotes the $(j,i)^{th}$ element of $\hat{\mathbf{H}}(t)$ in (6), and $\hat{d}_k^i$ denotes the $(i,k)^{th}$ element of the estimated ST data matrix $\hat{\mathbf{D}}(t)$.

3. PERFORMANCE ANALYSIS

In this section, we first analyze the channel estimation error of the ML channel estimator. Then, we examine the detection performance of the ML receiver by analyzing the Chernoff's upper bound on error probability, in which that of channel estimation error is taken into account. Finally, we analyze the lower bound on channel capacity of the ST coded MIMO systems, in which that of channel estimation error is also taken into consideration.

3.1. Channel Estimation Error

A channel estimation error vector can be expressed as follows,

$$\tilde{\mathbf{h}}(t) = \mathbf{h}(t) - \hat{\mathbf{h}}(t). \quad (9)$$

Substituting (6) into (9), after some algebraic manipulations, the variance matrix of the channel estimation error is given by

$$\mathrm{Var}\left[\tilde{\mathbf{h}}(t)\right] = \frac{\sigma^2}{\alpha}\mathbf{I}_{(L_tL_r\times L_tL_r)}. \quad (10)$$

The Total MSE of the ML channel estimation is given by

$$\text{Total MSE} \quad = \quad \mathrm{trace}\left\{\mathrm{Var}\left[\tilde{\mathbf{h}}(t)\right]\right\} = \frac{\sigma^2 L_t L_r}{\alpha}. \quad (11)$$

3.2. Chernoff's Upper Bound on Error Probability

Assuming full rank ST codes are employed, which can be relaxed in general, and considering independent Rayleigh distributions of the channel, the Chernoff's upper bound of the average probability of transmitting a codeword $\mathbf{d} \triangleq (d_1^1 d_1^2 \cdots d_1^{L_t} \cdots d_N^1 d_N^2 \cdots d_N^{L_t})^T$ and deciding in favor of a different codeword $\mathbf{e} \triangleq (e_1^1 e_1^2 \cdots e_1^{L_t} \cdots e_N^1 e_N^2 \cdots e_N^{L_t})^T$ is given by [6] (see also [5]),

$$P(\mathbf{d}\to\mathbf{e})_{\hat{\mathbf{H}}(t)} \le \left(\prod_{i=1}^{L_t}\lambda_i\right)^{-L_r}\left(\underbrace{\frac{N}{4\sigma^2}\cdot\frac{\sigma_Q^2}{\left(\frac{N}{\beta}+\frac{L_t}{\alpha}\right)}}_{Effective\ \ SNR^1}\right)^{-L_tL_r} \quad (12)$$

where λ_i is the eigenvalue of the code-error matrix $\mathbf{C}(\mathbf{d},\mathbf{e})$ defined as $C_{p,q} = \mathbf{x}_q^H\mathbf{x}_p$ where $\mathbf{x}_p = (d_1^P - e_1^P,\dots,d_N^P - e_N^P)^T$ and $\sigma_Q^2 = 1 + \frac{\sigma^2}{\alpha}$ is the variance of the element of the estimated channel coefficient vector $\hat{\mathbf{h}}(t)$. Notice that the $Effective\ \ SNR^1$ in (12) includes the parameters α, through the channel estimation error, and β in its expression; therefore, it reveals the underlying effects of pilot- and data-power factors on the probability of detection error.

3.3. A Lower Bound on Channel Capacity

The received symbol matrix $\mathbf{Y}_1(t)$ in (7) can be alternatively expressed as follows,

$$\mathbf{Y}_1(t) = \hat{\mathbf{H}}(t)\mathbf{D}(t) + \underbrace{(\mathbf{H}(t) - \hat{\mathbf{H}}(t))\mathbf{D}(t) + \mathbf{N}_1(t)}_{\mathbf{N}_1'(t)}, \quad (13)$$

where $\hat{\mathbf{H}}(t)$ is the channel estimate. The variance of the element of $\mathbf{N}_1'(t)$ can be computed by

$$\sigma^2_{\mathbf{N}_1'(t)} = \frac{1}{L_rN}\mathrm{trace}\left\{\mathrm{E}[\mathbf{N}_1'(t)\mathbf{N}_1'(t)^H]\right\}, \quad (14)$$

where the factor L_rN is the number of elements of $\mathbf{N}_1'(t)$. Substituting $\mathbf{N}_1'(t)$ into (14); and using the fact that trace $\left\{\mathrm{E}[\mathbf{N}_1(t)\mathbf{N}_1(t)^H]\right\} = \frac{\sigma^2 L_r N}{\beta}$ and the channel estimation error matrix $\tilde{\mathbf{H}}(t)$, the ST coded data matrix $\mathbf{D}(t)$, and the noise matrix $\mathbf{N}_1(t)$ are statistically independent, $\sigma^2_{\mathbf{N}_1'(t)}$ can be expressed as follows,

$$\sigma^2_{\mathbf{N}_1'(t)} = \frac{1}{L_r N}\text{trace}\left\{\mathrm{E}[\tilde{\mathbf{H}}(t)^H\tilde{\mathbf{H}}(t)]\mathrm{E}[\mathbf{D}(t)\mathbf{D}(t)^H]\right\} + \frac{\sigma^2}{\beta}, \quad (15)$$

where $\tilde{\mathbf{H}}(t) = \mathbf{H}(t) - \hat{\mathbf{H}}(t)$.

Referring to Theorem 1 in [4], $\mathrm{E}[\mathbf{D}(t)\mathbf{D}(t)^H] = \mathbf{I}_{(L_t\times L_t)}$ is the best in maximizing the mutual information of the ST coded MIMO systems given $\hat{\mathbf{H}}(t)$. This is equivalent to using the transmit antennas independently with equal power, which is reasonable because no channel knowledge is assumed to the transmitter. From (11), it can be shown that trace $\left\{\mathrm{E}[\tilde{\mathbf{H}}(t)^H\tilde{\mathbf{H}}(t)]\right\} = \frac{\sigma^2 L_t L_r}{\alpha}$. Hence,

$$\sigma^2_{\mathbf{N}_1'(t)} = \frac{\sigma^2}{N}\left(\frac{L_t}{\alpha} + \frac{N}{\beta}\right). \quad (16)$$

Let us define a normalized channel estimate $\bar{\mathbf{H}}(t) = \frac{1}{\sigma_Q}\hat{\mathbf{H}}(t)$. By using Theorem 1 in [4], the lower bound on channel capacity with the worst case of uncorrelated additive noise can be described as follows,

$$C_{L_t} \geq C_{worst} = \mathrm{E}\ \frac{M-L_t}{M}\log$$

$$\cdot \det\left(\mathbf{I}_{L_r\times L_r} + \underbrace{\frac{N}{\sigma^2}\cdot\frac{\sigma_Q^2}{\left(\frac{N}{\beta}+\frac{L_t}{\alpha}\right)}}_{Effective\ \ SNR^2}\bar{\mathbf{H}}(t)\bar{\mathbf{H}}(t)^H\right) \quad (17)$$

where the bandwidth efficiency loss is taken into account by the factor $\frac{M-L_t}{M}$, resulting from the use of $\mathbf{I}_{L_t\times L_t}$ as the pilot signal. It is worth noticing that $Effective\ \ SNR^2$ is similar to $Effective\ \ SNR^1$ except a scaling factor. Therefore, maximizing these factors results in minimizing the Chernoff's upper bound on error probability while maximizing the lower bound on channel capacity of the ST coded MIMO systems employing the ML channel estimator. As a result, these factors can be used as an objective function for an optimum power allocation, because they can be optimized by the parameters α and β.

4. OPTIMUM POWER ALLOCATION FOR MAXIMUM-LIKELIHOOD CHANNEL ESTIMATION

In this section, we address the block power allocation problem in order to optimally allocate the power to the data and the pilot parts. We consider the case of the constant block power, where the power of the pilot-embedded ST symbol matrix $\mathbf{U}(t)$ is constant. The normalized block power allocated to the pilot-embedded ST symbol matrix $\mathbf{U}(t)$, which is normalized by the transmit antenna numbers L_t, can be expressed as follows,

$$P_s = \frac{\mathrm{E}\left[\|\mathbf{D}(t)\mathbf{A}\|^2\right]}{L_t} + \frac{\mathrm{E}\left[\|\mathbf{P}\|^2\right]}{L_t} = P_s^{'} + P_p = \beta + \alpha, \quad (18)$$

where $P_s^{'} = \beta$ is the normalized block power allocated to the data part and $P_p = \alpha$ is the normalized block power allocated to the pilot part.

The objective is to maximize $Effective\ \ SNR^1$ with respect to the pilot-power factor α subject to the constraints of constant block power and acceptable Total MSE of the channel estimator. The acceptable Total MSE is a threshold T that indicates the acceptable channel estimation accuracy for a reliable channel estimate, which in turn yields good performance in term of probability of error. Substituting $\beta = P_s - \alpha$ into $Effective\ \ SNR^1$, the equivalent problem formulation is given by

$$\max_{\alpha}\ln\left(\frac{(\alpha+\sigma^2)(P_s-\alpha)}{(N-L_t)\alpha + P_s L_t}\right), \quad (19)$$

where Total MSE $\leq T$.

Differentiating (19) and equating the result to zero, we have the optimum solution for the pilot-power factor α^* as follows,

$$\alpha^* = \begin{cases} \frac{P_s-\sigma^2}{2};\ N = L_t \\ \frac{P_s L_t - \sqrt{P_s N(P_s L_t + \sigma^2(L_t-N))}}{(L_t-N)};\ N \neq L_t, \end{cases} \quad (20)$$

where the Total MSE in (11) must satisfy the following inequality

$$\text{Total MSE} = \frac{\sigma^2 L_t L_r}{\alpha} \leq T. \quad (21)$$

4.1. The Case of $N = L_t$

Substituting (20) into (21), we have the feasible range of SNR, i.e. SNR $= \frac{P_s L_t}{\sigma^2}$, and the minimum optimum pilot-power factor α^*_{min}, when MSE $= T$, as follows,

$$\text{SNR} \geq L_t + \frac{2L_t^2 L_r}{T} \text{ and } \alpha^*_{\text{min}} = \frac{L_t L_r P_s}{T + 2L_t L_r}. \quad (22)$$

Accordingly, the range of the optimum pilot-power factor α^* obtained in (20), when the SNR satisfies the inequality in (22), is given by

$$\frac{L_t L_r P_s}{T + 2L_t L_r} \leq \alpha^* < \frac{P_s}{2}. \quad (23)$$

However, there is a case when the SNR does not satisfy the inequality in (22); as a result, the Total MSE of the channel estimation is not reliable and the probability of detection error is inevitably increased. This scenario is equivalent to the low-SNR scenario, where wireless communications are not reliable. According to the range of the optimum pilot-power factor α^* obtained in (23), we use the minimum value of α^*, e.g. $\alpha^* = \frac{L_t L_r P_s}{T+2L_t L_r}$, in this scenario because $Effective\ \ SNR^1$ is a monotonically increasing function of α, for α within this range. In summary, we propose to determine the optimum pilot-power factor α^* for this case under different SNR scenarios as follows,

$$\alpha^* = \begin{cases} \frac{L_t L_r P_s}{T+2L_t L_r};\ \text{SNR} < L_t + \frac{2L_t^2 L_r}{T} \\ \frac{P_s-\sigma^2}{2};\ \text{Otherwise.} \end{cases} \quad (24)$$

4.2. The Case of $N \neq L_t$

Substituting (20) into (21), we have the feasible range of SNR and the minimum optimum pilot-power factor α^*_{min} as follows,

$$\text{SNR} - \sqrt{\tfrac{N}{L_t}}\sqrt{\text{SNR}(\text{SNR} + (L_t - N))} \geq \tfrac{L_t L_r (L_t-N)}{T},$$

$$(L_t - N)\alpha^*_{\text{min}} + \sqrt{\tfrac{P_s N}{L_t L_r}}\sqrt{P_s L_t^2 L_r + (L_t-N)T\alpha^*_{\text{min}}} = P_s L_t \quad (25)$$

Accordingly, the range of the optimum pilot-power factor α^* obtained in (20), when the SNR satisfies the inequality in (25), is given by

$$\alpha^*_{\text{min}} \leq \alpha^* \leq \frac{P_s(L_t - \sqrt{NL_t})}{(L_t - N)}. \quad (26)$$

For the case that the SNR does not satisfy the inequality in (25), we use the minimum value of α^*, i.e. α^*_{min}, for the same reason as in the case of $N = L_t$. In summary, we propose to determine the optimum pilot-power factor α^* for this case under different SNR scenarios as follows,

$$\alpha^* = \begin{cases} \alpha^*_{\text{min}};\ \text{When the inequlity in (25) is not satisfied} \\ \frac{P_s L_t - \sqrt{P_s N(P_s L_t + \sigma^2(L_t-N))}}{(L_t-N)};\ \text{Otherwise.} \end{cases} \quad (27)$$

In addition, the acceptable threshold T for the Total MSE of the channel estimation is quite small and is determined by practice, e.g. the simulation results in Section 5.

5. SIMULATION RESULTS

In this section, based on simulations, we examine the performances of the ST coded MIMO systems employing the ML channel estimator with four different power allocation strategies, including the optimum power allocation, the unequal power allocation where $\alpha = 0.265$ W and $\beta = 0.735$ W, the unequal power allocation where $\alpha = 0.735$ W and $\beta = 0.265$ W, and the equal power allocation where $\alpha = \beta = 0.5$ W. The total power of P_s=1 W/pilot-embedded ST symbol block is set. The 3×8 orthogonal STBC proposed in [2] is examined. The number of transmit and receive antennas are L_t=3 and L_r=2, respectively, and the number of data time slot per pilot-embedded ST symbol block is N=11. QPSK modulation is employed in these experiments and the acceptable threshold for the Total MSE of the channel estimation is 3, which yields α^*_{min} of 0.265 W. In addition, quasi-static flat Rayleigh fading channels are investigated in these experiments. For the case of $N = L_t$, please refer to the experiments in [5].

In Fig.1, the patterns of powers allocated to pilot and data parts using different power allocation strategies are illustrated. It is worth noticing that the patterns of power for the optimum power allocation

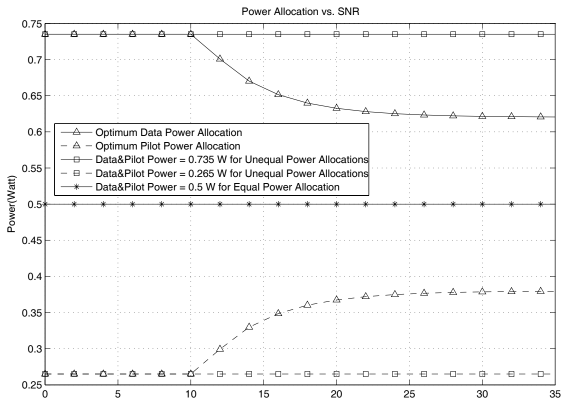

Fig. 1. the patterns of power allocated to pilot and data parts for different power allocation strategies.

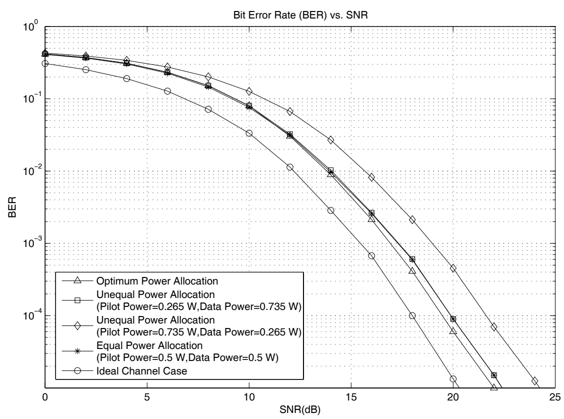

Fig. 2. The graph of BERs of the ST coded MIMO systems employing the ML channel estimator with different power allocation strategies.

strategy saturate at certain levels, e.g. $\alpha \approx 0.38$ W and $\beta \approx 0.62$ W, in high SNR regimes.

In Fig.2, the graph of BERs of the ST coded MIMO systems employing the ML channel estimator with different power allocation strategies are depicted. It is obvious that the optimum power allocation strategy yields the lowest bit error rate (BER), where the SNR difference of 2.2 dB is observed at BER of 10^{-4} compared with the ideal channel case, where the true channel is employed in the receiver. It is worth noticing that the equal power allocation and the unequal power allocation where $\alpha = 0.265$ W and $\beta = 0.735$ W yield the same BER whereas the unequal power allocation where $\alpha = 0.735$ W and $\beta = 0.265$ W is the worst, where SNR differences of 0.6 dB and 3 dB, respectively, are observed at BER of 10^{-4} compared with the optimum one. This results indicate that the good BER results from the proper power allocation strategies taking into account both of MSE of channel estimation and error probability. It is also worth noticing that the unequal power allocation strategy with more power constantly allocated to the data part provides a much better BER than the opposite one. In addition, if no information about the SNR is available to the transmitter, the equal power allocation can be reasonably used as a suboptimum approach with the acceptable performance degradation. In fact, the unequal power allocation where $\alpha = 0.735$ W and $\beta = 0.265$ W does provides the lowest MSE of channel estimation, however, this enhanced MSE does not result in the lowest probability of detection error. Since the total power is fixed, the more power allocated to the pilot part, the smaller remaining power is available for the data part. This power tradeoff results in error probability performance

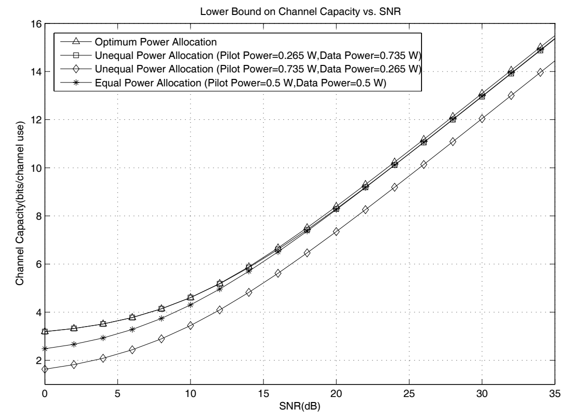

Fig. 3. The graph of lower bounds on channel capacity of the ST coded MIMO systems employing the ML channel estimator with different power allocation strategies.

degradation. Hence, the joint MSE and error probability optimization is an ultimate goal of such systems.

In Fig.3, the graph of the lower bounds on channel capacity when using different power allocation strategies are depicted. It is obvious that the optimum power allocation yields the highest lower bound on channel capacity whereas the unequal power allocation where $\alpha = 0.735$ W and $\beta = 0.265$ W is the worst. The same argument can be drawn in a similar way to Fig.2, where the reasonable suboptimal approach is the equal power allocation.

6. CONCLUSION

In this paper, we have proposed and unified the optimum power allocation for the ML channel estimation in ST coded MIMO systems based on jointly optimizing the Chernoff's upper bound on error probability and the lower bound on channel capacity. We have noted that these two bounds share the common effective SNR factor, thus, result in the same optimum power allocation strategy. Experimental results indicate that the proposed power allocation scheme yields a better performance in terms of bit error rate, while the unequal power allocation with more power constantly allocated to the pilot part performs the worst. Meanwhile, the equal power allocation and the unequal power allocation with more power constantly allocated to the data part yield close performance, where an SNR difference of 0.6 dB at BER of 10^{-4} compared with the optimum approach is observed. These results suggest that the equal power allocation can be reasonably used as the suboptimum approach when the optimum approach is not available.

7. REFERENCES

[1] G.J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," in *Bell Labs Tech. J.*, Vol. 1, pp. 41-59, 1996.

[2] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Infor. Theory*, Vol. 45, pp. 1456-1467, Jul. 1999.

[3] C. Budianu and L. Tong, "Channel estimation for space-time orthogonal block codes," *IEEE Trans. Signal Processing*, Vol. 50, pp. 2515-2528, Oct. 2002.

[4] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?, *IEEE Trans. Infor. Thoery*, Vol. 49, pp. 951-963, Apr. 2003.

[5] C. Pirak, Z.J. Wang, K.J.R. Liu, and S. Jitapunkul, "Performance analysis for pilot-embedded data-bearing approach in space-time coded MIMO systems," *Proc. IEEE ICASSP*, Philadelphia, PA, pp.593-596, Mar. 2005.

[6] J.G. Proakis, *Digital Communications, 3rd ed.*, McGraw-Hill, INC., NY, 1995.