Optimal Joint Spatial and Temporal Power Adaptation for Space-Time-Coded Systems With Imperfect CSI

Quan Kuang, Shu-Hung Leung, and Xiangbin Yu

***Abstract*—In this correspondence, a joint optimization of spatio-temporal (S-T) power allocation to minimize the average bit error rate (BER) subject to an average power constraint is developed for beamforming orthogonal space-time block coded multiantenna systems over flat Rayleigh fading channels in the presence of imperfect channel state information at the transmitter (CSIT). Compared to the spatial-only (S-O) power allocation methods, the optimal S-T strategy reduces BER significantly. It can be used as a benchmark to evaluate the existing suboptimal S-T schemes. A necessary and sufficient condition, which can determine the number of eigenbeams being used for optimal S-T power allocation without computing power allocation, is derived. This result can greatly simplify the power allocation algorithm. Numerical results show that the proposed optimal S-T algorithm has not only the best BER performance, but also the lowest computational complexity in comparison with the existing S-T strategies.**

***Index Terms*—Beamforming, multiple-input multiple-output (MIMO), spatio-temporal power allocation.**

I. Introduction

Multiple-input multiple-output (MIMO) systems with space-time block codes (STBC) can increase diversity gain without the knowledge of channel state information (CSI) at the transmitter (CSIT). If (partial/imperfect) CSIT is available, the error performance of MIMO systems with STBC can be significantly improved via precoding [1]–[7]. For many common forms of imperfect/partial CSIT, such as the channel mean [4], [5], channel correlation [6], compound channel model [8], the optimal precoder consists of adaptive power allocation and multiple eigenbeamforming. Various techniques and criteria have been used to derive the power adaptation strategies for precoders. However, most of the existing power adaptation schemes perform power allocation only in the spatial domain, while the temporal dimension is not well exploited. To further utilize the degrees of freedom of power adaptation, it is beneficial to exploit the power allocation jointly in the spatial and temporal domains. In [9], a spatial-temporal (S-T) power allocation was developed that treats the imperfect CSIT as perfect in deriving the temporal power allocation. In fact, despite a lot of research works dealing with the precoder or power allocation design, no systematic development of joint spatial and temporal power adaptation for STBC systems to minimize the bit error rate (BER) under imperfect CSIT is available in the literature. Regarding the number of eigenbeams to be used in power allocation, most of power adaptation design procedures need to determine it together with the calculation of power allocation iteratively. This way of determination will have large computational complexity for joint S-T algorithms.

Two S-T algorithms based on heuristic parameterization in [10] and asymptotic analysis in [11] have been proposed by the authors. These S-T methods have been shown to outperform the spatial-only (S-O) power strategies and the existing S-T approach in [9] significantly. However, the joint S-T power optimization problem has not yet been optimally solved.

The main contribution of this work is to derive the optimal S-T power allocation scheme. The optimum BER performance can be used as a benchmark for the evaluation of various power adaptation algorithms. In addition, a necessary and sufficient condition is derived that provides an efficient way to determine the number of eigenbeams to be used for the optimal power allocation. Numerical results show that our proposed optimal S-T algorithm has the lowest computational complexity and BER among the power allocation algorithms in comparison.

Notations: Bold upper case and lower case letters denote matrices and vectors, respectively. The superscript $(\cdot)^H$ denotes the Hermitian transposition. $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. $[\cdot]_{ij}$ denotes (i,j)th entry of a matrix. $E[\cdot]$ denotes the expectation. $\mathrm{tr}(\cdot)$ denotes the trace of a matrix.

II. System Model

We consider a wireless multiantenna communication system with M transmit antennas and N receive antennas operating over a flat and quasi-static Rayleigh fading channel, represented by an $N \times M$ matrix $\mathbf{H} = \{h_{nm}\}$, where the channel gains $\{h_{nm}\}$ are modeled as independent identically distributed (i.i.d.) complex Gaussian random variables (r.v.s) with zero-mean and unit variance, i.e., $h_{nm} \sim \mathcal{CN}(0,1)$. At the transmitter, only the channel estimate $\hat{\mathbf{H}}$ is available, modeled as $\hat{\mathbf{H}} = \mathbf{H} + \mathbf{E}$ [10], [12], where $\mathbf{E}$ is the channel error matrix independent of $\mathbf{H}$ and is composed of i.i.d. complex Gaussian entries, $[\mathbf{E}]_{ij} \sim \mathcal{CN}(0, \sigma_e^2)$. According to this distribution, the original channel can be rewritten as $\mathbf{H} = \frac{1}{1+\sigma_e^2}\hat{\mathbf{H}} + \sqrt{\frac{\sigma_e^2}{1+\sigma_e^2}}\mathbf{W}$ [10], where $\mathbf{W}$ is independent of $\hat{\mathbf{H}}$, and its entries $[\mathbf{W}]_{ij} \sim \mathcal{CN}(0,1)$ are i.i.d. Gaussian r.v.s. This imperfect CSIT model has been adopted to cover many realistic cases, such as delay feedback, quantized feedback and channel estimation at the transmitter in TDD systems [5].

The transmitter is composed of an orthogonal STBC (OSTBC) encoder, power allocation and a set of M beamformers (see [10] for a detailed description). The received signals can be expressed as

$$\mathbf{Y} = \sqrt{S}\mathbf{H}\hat{\mathbf{U}}\mathbf{P}\mathbf{D} + \mathbf{Z} = \sqrt{S}\bar{\mathbf{H}}\mathbf{P}\mathbf{D} + \mathbf{Z} \tag{1}$$

where $\mathbf{D}$ is the $M \times T$ OSTBC codeword matrix with normalized average power as $\frac{E[\mathrm{tr}(\mathbf{D}\mathbf{D}^H)]}{T} = 1$, which is used to encode K input data symbols into an M-dimensional vector sequence of T time slots with code rate $r = \frac{K}{T}$, S is the total transmit power radiated from the M transmit antennas, $\mathbf{Z}$ is an $N \times T$ received noise matrix with i.i.d. entries $[\mathbf{Z}]_{ij} \sim \mathcal{CN}(0, \sigma_n^2)$, $\mathbf{Y}$ is the $N \times T$ received signal matrix, $\bar{\mathbf{H}} \triangleq \mathbf{H}\hat{\mathbf{U}}$, $\hat{\mathbf{U}} = \{\hat{u}_{ij},\ i,\ j = 1,\ldots,M\}$ is an $M \times M$ unitary matrix containing the M eigenvectors of $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ corresponding to the eigenvalues $\{\hat{\zeta}_m\}$ sorted in decreasing order, and $\mathbf{P} = \mathrm{diag}(\sqrt{P_1}, \sqrt{P_2}, \ldots, \sqrt{P_M})$ denotes a diagonal power allocation matrix with the power constraints $\sum_{m=1}^{M} P_m = 1$ and $P_m \geq 0,\ m = 1,\ldots,M$.

Manuscript received September 04, 2011; revised January 05, 2012; accepted February 23, 2012. Date of publication April 12, 2012; date of current version May 11, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Xavier Mestre. The work described in this correspondence was substantially supported by a grant from CityU (7002459). The work of X. Yu was supported in part by the Doctoral Fund of the Ministry of Education of China under Grant 20093218120021 and in part by the National Natural Science Foundation of China under Grant 61172077.

Q. Kuang was with the State Key Laboratory of Millimeter Wave and Department of Electronic Engineering, City University of Hong Kong, Hong Kong. He is now with the Institute of Telecommunications, University of Stuttgart, 70569 Stuttgart, Germany (e-mail: quan.kuang.box@gmail.com).

S.-H. Leung is with the State Key Laboratory of Millimeter Wave and Department of Electronic Engineering, City University of Hong Kong, Hong Kong (e-mail: eeeugshl@cityu.edu.hk).

X. Yu is with the Nanjing University of Aeronautics and Astronautics, 210016 Nanjing, China (e-mail: yxbxwy@gmail.com).

Color versions of one or more of the figures in this correspondence are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2012.2190409

We assume that the receiver perfectly knows the CSI. After OSTBC decoding, the instantaneous received SNR per symbol at the receiver is expressed as [10]

$$\rho = \frac{S}{r\sigma_n^2}\|\bar{\mathbf{H}}\mathbf{P}\|_F^2 = \frac{S}{r\sigma_n^2}\sum_{m=1}^{M} P_m\beta_m \tag{2}$$

where β_m is the $m\text{th}$ eigenchannel power gain defined as

$$\beta_m = \sum_{n=1}^{N}|\bar{h}_{nm}|^2 = \sum_{n=1}^{N}|\sum_{i=1}^{M}h_{ni}\hat{u}_{im}|^2. \tag{3}$$

As shown in [10], given $\hat{\mathbf{H}}$, $\{\beta_m\}$ in (3) are independent noncentral χ^2 r.v.s. with the following conditional probability density function (pdf)

$$f_{\beta_m|\hat{\mathbf{H}}}(\beta_m|\hat{\mathbf{H}}) = \frac{1}{\sigma^2}\left(\frac{\beta_m}{\tilde{\beta}_m}\right)^{\frac{N-1}{2}}\exp\left(-\frac{\tilde{\beta}_m+\beta_m}{\sigma^2}\right) \times I_{N-1}\left(\frac{2\sqrt{\tilde{\beta}_m\beta_m}}{\sigma^2}\right) \tag{4}$$

where $m = 1,\ldots,\min\{M,N\}$, the variance $\sigma^2 = \frac{\sigma_e^2}{1+\sigma_e^2}$, and the sum of squared means $\tilde{\beta}_m$ is given by

$$\tilde{\beta}_m = \frac{1}{(1+\sigma_e^2)^2}\sum_{n=1}^{N}\left|\sum_{i=1}^{M}\hat{h}_{ni}\hat{u}_{im}\right|^2 = \frac{\hat{\zeta}_m}{(1+\sigma_e^2)^2}, \tag{5}$$

$\tilde{\beta}_1 \geq \tilde{\beta}_2 \geq \cdots \geq \tilde{\beta}_M$ since $\hat{\zeta}_1 \geq \hat{\zeta}_2 \geq \cdots \geq \hat{\zeta}_M$, $I_v(x)$ is the $v\text{th}$ order modified Bessel function of the first kind. For systems with $M > N$, there are $M-N$ beams with $\hat{\zeta}_m = 0$. Hence, given $\hat{\mathbf{H}}$, $\{\beta_m\}$ of those beams are independent central χ^2 distributed with the conditional pdf given as

$$f_{\beta_m|\hat{\mathbf{H}}}(\beta_m|\hat{\mathbf{H}}) = \frac{1}{\sigma^{2N}\Gamma(N)}\beta_m^{N-1} \times \exp\left(-\frac{\beta_m}{\sigma^2}\right), \quad m = N+1,\ldots,M. \tag{6}$$

III. Problem Formulation

In the proposed power adaptation policy, the total power S is subject to the long term (time) average constraint and can be varied from one OSTBC block to another according to the updated CSIT, which is referred to as temporal power allocation. $\{P_m\}$'s are the spatial power allocation parameters. Our design objective is to find the optimal temporal power function $S = S(\tilde{\boldsymbol{\beta}})$ and spatial power function $P_m = P_m(\tilde{\boldsymbol{\beta}})$ jointly to minimize the average BER, where $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_1,\ldots,\tilde{\beta}_M)$ is the imperfect CSIT, as defined in (5).

The BER of multiple quadrature amplitude modulation (MQAM) of constellation size Q and Gray mapping can be approximated by a tight bound [13]: $\text{BER}_\rho \approx 0.2\exp(-g\rho)$, where ρ is the received SNR, $g = \frac{1.5}{Q-1}$ for square MQAM, and $g = \frac{6}{5Q-4}$ for rectangular MQAM. With (2), the BER approximation can be written as

$$\text{BER}_\rho \approx 0.2\exp\left(-\gamma\sum_{m=1}^{M}P_m\beta_m\right) \tag{7}$$

where

$$\gamma \triangleq \frac{gS}{r\sigma_n^2}. \tag{8}$$

Given $\hat{\mathbf{H}}$, the conditional average BER is expressed as

$$P_{b|\hat{\mathbf{H}}} = \int \text{BER}_\rho f_{\boldsymbol{\beta}|\hat{\mathbf{H}}}(\boldsymbol{\beta}|\hat{\mathbf{H}})d\boldsymbol{\beta} \tag{9}$$

where $f_{\boldsymbol{\beta}|\hat{\mathbf{H}}}(\boldsymbol{\beta}|\hat{\mathbf{H}})$ is the joint conditional pdf of $\{\beta_m\}$, given $\hat{\mathbf{H}}$. Applying (4), (6) and the independence of $\{\beta_m\}$ to (9), we can obtain the conditional average BER as

$$P_{b|\hat{\mathbf{H}}} = 0.2\prod_{m=1}^{M}\frac{1}{(1+\gamma\sigma^2P_m)^N}\exp\left(-\frac{\gamma\tilde{\beta}_mP_m}{1+\gamma\sigma^2P_m}\right). \tag{10}$$

According to (8), we define $\gamma = \gamma(\tilde{\boldsymbol{\beta}}) = \frac{gS(\tilde{\boldsymbol{\beta}})}{r\sigma_n^2}$. Now the joint spatio-temporal power adaptation problem can be formulated as

$$\underset{\gamma,P_m}{\text{minimize}}\ L(\gamma,P_m) = \int P_{b|\hat{\mathbf{H}}}\ f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}})d\tilde{\boldsymbol{\beta}} \tag{11a}$$

$$\text{s.t.}\quad \int \gamma\ f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}})d\tilde{\boldsymbol{\beta}} = \frac{g\bar{S}}{r\sigma_n^2} \triangleq \bar{\gamma},\ \gamma \geq 0 \tag{11b}$$

$$\sum_{m=1}^{M}P_m = 1,\ P_m \geq 0,\ m = 1,\ldots,M \tag{11c}$$

where $P_{b|\hat{\mathbf{H}}}$ is given in (10), $\bar{\gamma} \triangleq \frac{g\bar{S}}{r\sigma_n^2}$, $\bar{S}$ is the power budget. $f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}})$ is the joint pdf of $\tilde{\boldsymbol{\beta}}$, which can be obtained from the pdf of ordered eigenvalues of the Wishart matrix by transformation of random variables (see [10, App. A]).

In the above formulation, the temporal power parameter γ and spatial power parameters $\{P_m\}$ are functions of the random CSIT $\tilde{\boldsymbol{\beta}}$. We note that the constraints are separable in the sense that (11b) depends on γ while (11c) on $\{P_m\}$. Hence, we can solve the problem (11) as the following inner-outer formulation

$$\underset{\gamma}{\text{minimize}}\quad \tilde{L}(\gamma) \tag{12a}$$

$$\text{s.t.}\quad \int \gamma\ f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}})d\tilde{\boldsymbol{\beta}} = \bar{\gamma},\quad \gamma \geq 0 \tag{12b}$$

where

$$\tilde{L}(\gamma) = \inf_{P_m}\left\{L(\gamma,P_m)|\sum_{m=1}^{M}P_m = 1,\ P_m \geq 0\right\}. \tag{13}$$

IV. Joint Spatio-Temporal Power Adaptation

A. Solution to the Inner Problem

Solving (13) is equivalent to minimizing $P_{b|\hat{\mathbf{H}}}$ for given $\tilde{\boldsymbol{\beta}}$ and γ subject to the spatial power constraint. The objective function of the inner minimization problem is now defined as the logarithm of $P_{b|\hat{\mathbf{H}}}$ with the parameter-independent term disregarded. The inner problem can be rewritten as

$$\underset{P_m}{\text{minimize}}\quad -\sum_{m=1}^{M}\left[N\log(1+\gamma\sigma^2P_m) + \frac{\gamma\tilde{\beta}_mP_m}{1+\gamma\sigma^2P_m}\right] \tag{14a}$$

$$\text{s.t.}\quad \sum_{m=1}^{M}P_m = 1,\ P_m \geq 0,\ m = 1,\ldots,M. \tag{14b}$$

This is a convex optimization problem, where the optimal $\{P_m\}$ can be obtained by the Karush–Kuhn–Tucker (KKT) conditions [14]:

$$\sum_{m=1}^{M}P_m = 1,\quad P_m \geq 0 \tag{15}$$

$$-\frac{N\gamma\sigma^2}{1+\gamma\sigma^2 P_m}-\frac{\gamma\tilde{\beta}_m}{(1+\gamma\sigma^2 P_m)^2}+\eta-\mu_m=0, \quad \mu_m\geq 0 \quad (16)$$

$$\mu_m P_m = 0 \quad (17)$$

for $m=1,\ldots,M$, where η and $\{\mu_m\}$ are the Lagrange multipliers for the power equality constraint and inequality constraints in (14b) respectively. From (17), if $P_m>0$, then $\mu_m=0$. Equation (16) gives

$$\frac{N\gamma\sigma^2}{1+\gamma\sigma^2 P_m}+\frac{\gamma\tilde{\beta}_m}{(1+\gamma\sigma^2 P_m)^2}=\eta. \quad (18)$$

Let us define λ parameter from (18) as

$$\lambda=\frac{\eta}{\gamma}=\frac{N\sigma^2}{1+\gamma\sigma^2 P_m}+\frac{\tilde{\beta}_m}{(1+\gamma\sigma^2 P_m)^2}. \quad (19)$$

Solving (19) for P_m and combining the nonnegative condition give

$$P_m=\max\left\{0,\frac{1}{\gamma\sigma^2}\left[\mathcal{I}_m(\lambda)-1\right]\right\} \quad (20)$$

where

$$\mathcal{I}_m(\lambda)=\frac{\sqrt{N^2\sigma^4+4\tilde{\beta}_m\lambda}+N\sigma^2}{2\lambda}. \quad (21)$$

The λ parameter is obtained by substituting (20) into the sum power constraint in (15) and solving the resulting equation described as

$$G_{M_0}(\lambda)=\sum_{m=1}^{M_0}\mathcal{I}_m(\lambda)-(M_0+\gamma\sigma^2)=0 \quad (22)$$

where M_0 is the number of eigenbeams with nonzero power allocation, whose value can be determined iteratively. In the next subsection, we will provide a simple mechanism to determine the value of M_0 without iterative calculation. Notice that $G_{M_0}(\lambda)$ is a monotonically decreasing function of λ for $\lambda\in(0,\infty)$ with $G_{M_0}(0^+)=\infty$ and $G_{M_0}(\infty)=-(M_0+\gamma\sigma^2)<0$, the root does exist and is unique.

After substituting the P_m solution of (20) into (10), we obtain

$$\tilde{L}(\gamma)=\int P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}})d\tilde{\boldsymbol{\beta}} \quad (23)$$

where

$$P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})=0.2\prod_{m=1}^{M_0}\frac{\exp\left(\frac{b_m}{\mathcal{I}_m(\lambda(\gamma,\tilde{\boldsymbol{\beta}}))}-b_m\right)}{\mathcal{I}_m(\lambda(\gamma,\tilde{\boldsymbol{\beta}}))^N} \quad (24)$$

with $b_m\triangleq\frac{\tilde{\beta}_m}{\sigma^2}$. $\mathcal{I}_m(\lambda)$ are given in (21), and $\lambda(\gamma,\tilde{\boldsymbol{\beta}})$ is associated with γ and $\tilde{\boldsymbol{\beta}}$ according to (22).

B. Solution to the Outer Problem

Solving (12) for the optimal temporal power function is a calculus of variations problem with an isoperimetric constraint [15]. By introducing the Lagrange multiplier ξ, the objective functional is expressed as

$$\underset{\gamma\geq 0}{\text{minimize}}\int\left(P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})+\xi[\gamma-\bar{\gamma}]\right)f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}})d\tilde{\boldsymbol{\beta}} \quad (25)$$

where ξ is a constant. If we define $J\triangleq(P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})+\xi[\gamma-\bar{\gamma}])f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}})$, the optimal γ should satisfy $\frac{\partial J}{\partial\gamma}=0$, resulting in

$$\frac{\partial P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})}{\partial\gamma}+\xi=0. \quad (26)$$

Using (24), (19), and (21), $\frac{\partial P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})}{\partial\lambda}$ and $\frac{\partial^2 P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})}{\partial\lambda^2}$ can be given as

$$\frac{\partial P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})}{\partial\gamma}=-\lambda(\gamma,\tilde{\boldsymbol{\beta}})P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}}) \quad (27)$$

$$\frac{\partial^2 P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})}{\partial\gamma^2}=\left(\lambda^2(\gamma,\tilde{\boldsymbol{\beta}})-\frac{\sigma^2}{\sum_{m=1}^{M_0}\frac{\partial\mathcal{I}_m(\lambda)}{\partial\lambda}}\right)P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}}). \quad (28)$$

Since $\{\frac{\partial\mathcal{I}_m}{\partial\lambda}\}$ in (28) are negative (see property 1 of the Appendix), we have $\frac{\partial^2 P_{b|\hat{\mathbf{H}}}}{\partial\gamma^2}>0$. This shows that $P_{b|\hat{\mathbf{H}}}(\gamma,\tilde{\boldsymbol{\beta}})$ is convex in γ. Thus, (26) is both necessary and sufficient to achieve the global minimum of (12).

From (24) and (27), we can rewrite (26) as

$$0.2\prod_{m=1}^{M_0}\frac{1}{\mathcal{I}_m(\lambda)^N}\exp\left(\frac{b_m}{\mathcal{I}_m(\lambda)}-b_m\right)\lambda=\xi. \quad (29)$$

Taking logarithm of both sides of (29), we define $F_{M_0}(\lambda)$ as

$$F_{M_0}(\lambda)=\log 0.2+\sum_{m=1}^{M_0}\left(-N\log\mathcal{I}_m(\lambda)+\frac{b_m}{\mathcal{I}_m(\lambda)}-b_m\right)+\log\lambda-\log\xi \quad (30)$$

which is a monotonically increasing function of λ with $F_{M_0}(0^+)=-\infty$ and $F_{M_0}(\infty)=\infty$. Hence, the root is unique and can be found numerically, for example by Newton's method or bisection method [16]. Once λ is found, γ can be calculated from (22) as

$$\gamma=\frac{1}{\sigma^2}\left(\sum_{m=1}^{M_0}\mathcal{I}_m(\lambda)-M_0\right). \quad (31)$$

However, before using (30) to calculate λ, we should determine the value of M_0 to make sure the corresponding γ calculated by (31) is positive. The following Theorem provides a necessary and sufficient condition to determine M_0 without iterative calculation of λ.

Theorem 1: The optimal $\{P_m, m=1,\ldots,M_0\}$ and γ values for the beamforming OSTBC system with joint spatio-temporal power allocation are positive if and only if

$$F_{M_0}\left(\lambda_{M_0}^{(u)}\right)>0 \quad (32)$$

where $M_0=\arg\max_{\nu} F_\nu(\lambda_\nu^{(u)})>0$, $F_\nu(\lambda)$ is given by (30), and $\lambda_\nu^{(u)}$ is an upper bound of λ defined as $\lambda<N\sigma^2+\tilde{\beta}_\nu\triangleq\lambda_\nu^{(u)}$, $\nu=1,\ldots,M$.

Proof: The proof is shown in the Appendix. ■

The Lagrange multiplier ξ in (30) is determined offline through numerical search such that the average power constraint in (11b) is satisfied (see [10, App. A]).

C. Summary of the Algorithm

Now, we summarize the whole on-line algorithm of the joint S-T power adaptation as follows.

1) Based on the CSIT $\hat{\mathbf{H}}$ of the current OSTBC block, compute the eigenvalues $\{\hat{\zeta}_m\}$ of $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ and then calculate $\{\tilde{\beta}_m\}$ using (5).
2) Based on the CSIT $\{\tilde{\beta}_m\}$, starting from $\nu=M$ to $\nu=1$, determine the largest ν such that $F_\nu(\lambda_\nu^{(u)})>0$ and let $M_0=\nu$.
3) If the above condition cannot be satisfied, let $\gamma=0$; otherwise go to step 4).

TABLE I
AVERAGE RUN TIME OF DIFFERENT POWER ADAPTATION SCHEMES

	Optimal S-T	Sharma S-T	CSNR-based S-T	Asymptotic S-T	Jongren S-O
Time (μs)	357	833	1837	855	808

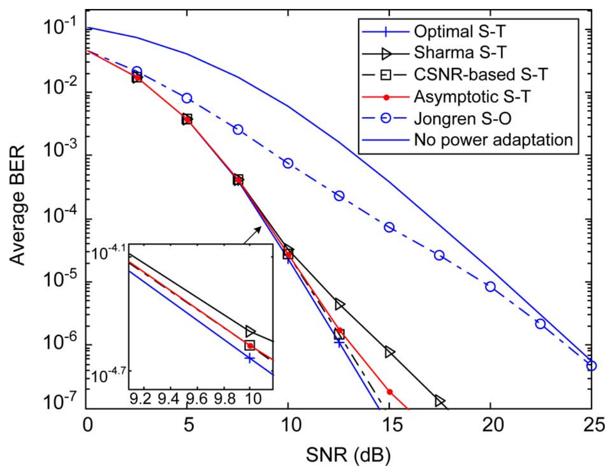

Fig. 1. BER performance for different power adaptation schemes, $M = 3$, $N = 1, \sigma_e^2 = 0.05$, H_3 code and QPSK. No power adaptation means constant power in time and equal power allocation in space.

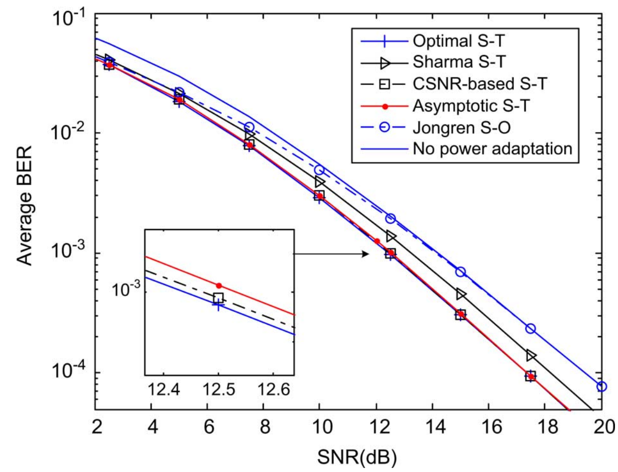

Fig. 2. BER performance for different power adaptation schemes, $M = 2$, $N = 1, \sigma_e^2 = 0.4$, G_2 code and BPSK. No power adaptation means constant power in time and equal power allocation in space.

4) Using Newton's method (or bisection method) to solve $F_{M_0}(\lambda) = 0$ for λ.
5) Compute γ (the temporal power parameter) using (31).
6) Compute $\{P_m\}$ (the spatial power parameter) using (20).

V. NUMERICAL RESULTS

In this section, we compare the proposed joint optimal power adaptation scheme with various strategies, including the spatial-only power allocation [2] (Jongren S-O), the S-T power allocation proposed in [9] (Sharma S-T), and our previously proposed two S-T approaches: the compressed SNR (CSNR) based S-T [10] and the asymptotic S-T [11]. We denote a system with M transmit and N receive antennas as $M \times N$ system and SNR is defined as $\frac{\bar{S}}{\sigma_n^2}$. For illustration, G_2 code (Alamouti code) with $r = 1$ and H_3 with $r = \frac{3}{4}$ are used for 2×1 and 3×1 systems respectively [17].

Figs. 1 and 2 plot the average BERs of the above-mentioned schemes for σ_e^2 equal to 0.05 and 0.4, respectively. As shown in the two figures, the S-T schemes reduce the BER significantly in comparison with the Jongren S-O. Our proposed optimal S-T achieves the best performance and can act as a benchmark for S-T schemes. As for the Sharma S-T, the temporal power is adapted by treating the imperfect CSIT as perfect. So when σ_e^2 is small, it performs closely to the optimal S-T at low SNRs. However, when SNR increases (larger than 10 dB in Fig. 1) or σ_e^2 is large (as shown in Fig. 2), the effect of errors in CSIT on the BER becomes prominent, which degrades the performance of the Sharma S-T scheme. Thus the optimal S-T is more suitable for practical systems than the Sharma S-T where the imperfection in the CSIT is non-negligible. Figs. 1 and 2 also show that the CSNR-based S-T provides the BER close to the optimal S-T. The reason is that the CSNR criterion being used for designing spatial power allocation can produce almost the same performance as the optimal spatial-only scheme (see [10, Figs. 2 and 3]) over a wide range of σ_e^2. Thus, when it is used for the development of S-T approach, it is expected to have similar performance as the optimal S-T. Finally, the asymptotic S-T has observable performance deterioration in the medium SNR region, i.e., 10–16 dB in Figs. 1 and 2, since it was developed based on asymptotic analysis at low SNR and high SNR.

Regarding the computational complexity, the above methods require to compute the eigendecomposition of $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ of dimension $M \times M$ in addition to the power calculation. The computational complexity is asymptotically dominated by the eigendecomposition as M increases because its complexity is $\mathcal{O}(M^3)$ and the rest of calculations are $\mathcal{O}(M^2)$. However, for the system with a small number of transmit antennas, the proposed optimal S-T will reduce the complexity considerably. In fact, the power calculation in all the above methods needs to perform one-dimensional numerical root finding. For the optimal S-T, Sharma S-T and Jongren S-O, the root finding is to obtain the value of a parameter. With the obtained parameter, all these three methods have closed-form formulae to calculate the power allocation. However, for the Sharma S-T and Jongren S-O schemes, the root finding should be performed iteratively together with the determination of the number of eigenbeams with positive power. In contrast, thanks to the necessary and sufficient condition we derived, the determination of the number of eigenbeams does not need to calculate the power allocation for the optimal S-T, making it simpler than the former two schemes. For the CSNR-based S-T and asymptotic S-T, the numerical root finding is to calculate the value of the temporal power directly, and the spatial power is then calculated by closed-form formulae. However, this numerical search requires more computation since the functions are more complicated. The run time of different power adaptation schemes averaged over 10 000 channel realizations for a 3×1 system with QPSK and H_3 code is listed in Table I, where SNR $= 5$ dB and $\sigma_e^2 = 0.05$. The computer we used in the experiment is equipped with Intel 1.60 GHz Core 2 CPU and 3G RAM. The simulation software is Matlab 7.1. As shown in the table, the optimal S-T in fact requires the least run time.

VI. CONCLUSION

In this correspondence, an optimal joint spatial and temporal power allocation algorithm is proposed. The optimal algorithm is derived from the minimization of a tight BER bound subject to an average power constraint. A necessary and sufficient condition is derived to provide an efficient procedure to determine the number of eigenbeams with positive power directly. The proposed optimal scheme has not only the best performance, but also the lowest computational complexity among all the power allocation schemes we compare.

In this correspondence, the constant-rate transmission for delay sensitive applications is considered, as in [9]–[11], where the transmission rate is kept constant over all fading conditions. The design objective is to minimize the *average* BER. Hence, for certain channel realizations, the instantaneous BER might be greater than a given threshold, causing an *outage*. It would be an interesting study to include an outage probability constraint into our formulation.

APPENDIX
PROOF OF THE THEOREM 1

It is easy to show that $\mathcal{I}_m(\lambda)$ defined in (21) has the following properties (the proof is omitted):

1) $\mathcal{I}_m(\lambda)$ is a strictly monotonically decreasing function of λ.
2) $\mathcal{I}_m(\lambda_m^{(u)}) = 1$.
3) $\mathcal{I}_{m_1}(\lambda) \geq \mathcal{I}_{m_2}(\lambda)$ for $m_1 < m_2$.

where $\lambda_m^{(u)}$ is defined as

$$\lambda_m^{(u)} = N\sigma^2 + \tilde{\beta}_m, \quad m = 1, \ldots, M. \tag{33}$$

Using (19), the λ parameter is bounded from above by $\lambda_m^{(u)}$ when setting $P_m = 0$.

In the first part of the proof, we show that if $F_{M_0}(\lambda_{M_0}^{(u)}) > 0$, then $\gamma > 0$ and $P_m > 0$ for $m = 1, \ldots, M_0$. Let λ satisfy $F_{M_0}(\lambda) = 0$. Due to the monotonically increasing property of $F_{M_0}(\lambda)$, we have $\lambda < \lambda_{M_0}^{(u)}$. Further, we have $\lambda < \lambda_{M_0}^{(u)} \leq \lambda_m^{(u)}$ for $m < M_0$, according to (33) and $\tilde{\beta}_1 \geq \tilde{\beta}_2 \ldots \geq \tilde{\beta}_M$. Applying $\lambda < \lambda_m^{(u)}$ and the properties 1) and 2) of $\mathcal{I}_m(\lambda)$ into (31), we show γ is positive as

$$\gamma = \frac{1}{\sigma^2}\left(\sum_{m=1}^{M_0} \mathcal{I}_m(\lambda) - M_0\right) > \frac{1}{\sigma^2}\left(\sum_{m=1}^{M_0} \mathcal{I}_m(\lambda_m^{(u)}) - M_0\right) = 0$$

and P_m is shown positive as

$$P_m = \frac{1}{\gamma\sigma^2}\left[\mathcal{I}_m(\lambda) - 1\right] > \frac{1}{\gamma\sigma^2}[\mathcal{I}_m(\lambda_m^{(u)}) - 1] = 0 \quad m = 1, \ldots, M_0.$$

In the second part of the proof, we assume $\gamma > 0$ and $P_m > 0$ for $m = 1, \ldots, M_0$, then we prove $F_{M_0}(\lambda_{M_0}^{(u)}) > 0$. From the definition of λ given in (19), we immediately have $\lambda < \lambda_{M_0}^{(u)}$ due to the positive γ and P_{M_0}. Since λ satisfies $F_{M_0}(\lambda) = 0$, we conclude that $F_{M_0}(\lambda_{M_0}^{(u)}) > 0$ due to the monotonically increasing property of $F_{M_0}(\lambda)$.

REFERENCES

[1] A. Pascual-Iserte, D. P. Palomar, A. I. Perez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 346–360, Jan. 2006.

[2] G. Jongren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.

[3] M. Vu and A. Paulraj, "Optimal linear precoders for MIMO wireless correlated channels with nonzero mean in space-time coded systems," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2318–2332, Jun. 2006.

[4] J. W. Huang, E. K. S. Au, and V. K. N. Lau, "Precoder design for space-time coded MIMO systems with imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 1977–1981, Jun. 2008.

[5] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2599–2613, Oct. 2002.

[6] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1673–1690, Jul. 2003.

[7] M. Vu and A. Paulraj, "MIMO wireless linear precoding," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 86–105, 2007.

[8] J. Wang and D. P. Palomar, "Worst-case robust MIMO transmission with imperfect channel knowledge," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3086–3100, 2009.

[9] V. Sharma, K. Premkumar, and R. N. Swamy, "Exponential diversity achieving spatio-temporal power allocation scheme for fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 188–208, Jan. 2008.

[10] Q. Kuang, S.-H. Leung, and X. Yu, "Novel power adaptation strategy for space-time-coded multiantenna systems with imperfect channel state information," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1227–1233, 2011.

[11] Q. Kuang, S.-H. Leung, and X. Yu, "Spatial and temporal power adaptation for space-time coded MIMO systems with imperfect CSI," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 22–27, 2011, pp. 3260–3263.

[12] E. G. Larsson, "Diversity and channel estimation errors," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 205–208, Feb. 2004.

[13] S. Zhou and G. B. Giannakis, "Adaptive modulation for multiantenna transmissions with channel mean feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1626–1636, 2004.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[15] E. R. Pinch, *Optimal Control and the Calculus of Variations*. New York: Oxford Univ. Press, 1993.

[16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1997.

[17] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 451–460, Mar. 1999.