SPATIAL AND TEMPORAL POWER ADAPTATION FOR SPACE-TIME CODED MIMO SYSTEMS WITH IMPERFECT CSI

Quan Kuang[1], *Shu-Hung Leung*[2], *Xiangbin Yu*[3]

1.Institute of Telecommunications, University of Stuttgart, 70569 Stuttgart, Germany
2.Department of EE, City University of Hong Kong, Hong Kong, China
3.Department of EE, Nanjing University of Aeronautics and Astronautics, Nanjing, China

ABSTRACT

This paper presents a strategy to allocate power both in space and time for beamforming orthogonal space-time block coded multi-antenna systems with imperfect channel state information at the transmitter. Most of the existing schemes focus on spatial-only power allocation, while the temporal dimension is not well exploited. The proposed space-time strategy is developed by decoupling the temporal and spatial power calculation through asymptotic analysis, resulting in efficient computation. Numerical results demonstrate that the proposed scheme has considerable performance gains over the existing power adaptation schemes.

Index Terms— Power adaptation, imperfect CSI, orthogonal space-time block coding, beamforming, MIMO.

1. INTRODUCTION

Multiple-input multiple-output (MIMO) systems with orthogonal space-time block codes (OSTBC) can provide diversity gain without the knowledge of channel state information at the transmitter (CSIT). Nevertheless, the error performance can be significantly improved by exploiting the CSIT via precoding.

For many common forms of imperfect/partial CSIT, such as the channel mean [1], channel correlation [2], and compound channel model [3], the optimal precoder consists of adaptive power allocation and multiple eigen-beamforming. Among these transmitters, various techniques and criteria have been used to derive the power adaptation strategies. However, the strategies of most of the existing schemes are spatial-only power allocation, while the temporal dimension is not well exploited. In [4], a temporal power allocation was developed by treating the imperfect CSIT as perfect. As will be shown in this paper, this treatment results in inferior bit-error-rate (BER) performance. In fact, despite a lot of research works dealing with precoder or power allocation design, no joint spatial and temporal power adaptation for OSTBC MIMO systems to minimize the BER under imperfect CSIT is available in the literature.

This paper presents a strategy to allocate the power both in space and time to minimize the average BER subject to a long term (time) average of transmit power constraint. Through asymptotic analysis, we are able to decouple the temporal and spatial power calculation. At high SNR, the proposed strategy has closed-form formulae to compute the temporal power, while at low SNR, a simple bisection method is used instead. After the temporal power is obtained, a closed-form scheme is used to distribute the power spatially among multiple beams. The proposed strategy significantly reduces the BER in comparison with the spatial-only power allocation. It also provides considerable gains over the space-time scheme in [4], since the imperfection of the CSIT is explicitly considered in the proposed method.

This work was supported in part by a grant from CityU (7002459) and in part by Doctoral Fund of Ministry of Education of China (No.20093218120021) and NUAA Research Funding (NS2010113). Email: EEEUGSHL@cityu.edu.hk

2. SYSTEM MODEL

We consider a wireless multi-antenna 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 $\hat{\mathbf{H}} = \mathbf{H} + \mathbf{E}$ [5], where $\mathbf{E}$ is the channel error matrix independent of $\mathbf{H}$, and its entries, $[\mathbf{E}]_{ij} \sim \mathcal{CN}(0, \sigma_e^2)$, are assumed to be i.i.d. complex Gaussian r.v.s. According to this signal model, it is easy to verify that the channel gain $\mathbf{H}$ conditioned on the channel estimate $\hat{\mathbf{H}}$ follows the mean feedback model [1], and can be expressed as

$$\mathbf{H} = \frac{1}{1+\sigma_e^2}\hat{\mathbf{H}} + \sqrt{\frac{\sigma_e^2}{1+\sigma_e^2}}\mathbf{W} \tag{1}$$

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 cases such as delayed feedback, quantized feedback, and channel estimation at the transmitter using channel reciprocity in TDD systems [1].

As shown in Fig.1, the transmitter architecture, similar to [1, 2, 5, 6], is composed of an orthogonal space-time encoder, power allocation and a set of M beamformers. This OSTBC and beamforming combination is robust to channel fading and provides an efficient way to exploit the imperfect CSIT [6]. The space-time encoder, which is represented by an $M \times T$ OSTBC codeword matrix $\mathbf{D}$, is used to encode K input data symbols into an M-dimensional vector sequence of T time slots with code rate $r = K/T$. Based on the imperfect CSIT model considered in this paper, it has been proved that the optimal beam directions are the eigenmodes of the channel estimate [1, 6]. Using the eigen-beams, the received signals of the system can be expressed as

$$\mathbf{Y} = \sqrt{S}\mathbf{H}\hat{\mathbf{U}}\mathbf{P}\mathbf{D} + \mathbf{Z} = \sqrt{S}\bar{\mathbf{H}}\mathbf{P}\mathbf{D} + \mathbf{Z} \tag{2}$$

where $\mathbf{Y}$ is the $N \times T$ received signal matrix, S is the total transmit power radiated from the M transmit antennas, $\mathbf{Z}$ is

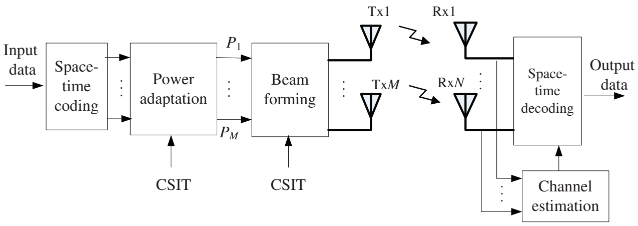

Fig. 1. System diagram

an $N \times T$ received Gaussian noise matrix with i.i.d. entries $[\mathbf{Z}]_{ij} \sim \mathcal{CN}(0, \sigma_n^2)$, $\bar{\mathbf{H}} \triangleq \mathbf{H}\hat{\mathbf{U}}$, $\hat{\mathbf{U}} = [\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \ldots, \hat{\mathbf{u}}_M]$ 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 spatial power allocation matrix with the spatial power constraints $\sum_{m=1}^{M} P_m = 1$ and $P_m \geq 0$.

It is assumed that the receiver perfectly knows the CSI. After OSTBC decoding, the instantaneous received SNR per symbol at the receiver is expressed as [7]

$$\rho = \frac{S}{r\sigma_n^2}\|\bar{\mathbf{H}}\mathbf{P}\|_F^2 = \frac{S}{r\sigma_n^2}\sum_{m=1}^{M} P_m\beta_m, \tag{3}$$

where β_m is the *m*-th eigen-channel power gain defined as

$$\beta_m = \sum_{n=1}^{N} |\bar{h}_{nm}|^2 = \sum_{n=1}^{N} |\sum_{i=1}^{M} h_{ni}\hat{u}_{im}|^2 \tag{4}$$

and $\hat{u}_{im}$ is the (i, m)-th entry of $\hat{\mathbf{U}}$.

According to (1), the elements $\{\bar{h}_{nm}\}$ of $\bar{\mathbf{H}}$ are i.i.d. complex Gaussian r.v.s, $\bar{h}_{nm} \sim \mathcal{CN}\left(\frac{\sum_{i=1}^{M}\hat{h}_{ni}\hat{u}_{im}}{1+\sigma_e^2}, \frac{\sigma_e^2}{1+\sigma_e^2}\right)$, conditioned on $\hat{\mathbf{H}}$. Thus given $\hat{\mathbf{H}}$, $\{\beta_m\}$ in (4) are independent noncentral chi-square r.v.s, with the conditional probability density function (pdf) given as [8]

$$f_{\beta_m|\hat{\mathbf{H}}}(\beta_m|\hat{\mathbf{H}}) = \frac{1}{\sigma^2}\left(\frac{\beta_m}{\tilde{\beta}_m}\right)^{\frac{N-1}{2}} \exp\left(-\frac{\tilde{\beta}_m + \beta_m}{\sigma^2}\right) \times I_{N-1}\left(\frac{2\sqrt{\tilde{\beta}_m\beta_m}}{\sigma^2}\right) \tag{5}$$

where $m = 1, \ldots, \min\{N, M\}$, the variance σ^2 and the sum of squared means, denoted by $\tilde{\beta}_m$, are given by

$$\sigma^2 = \frac{\sigma_e^2}{1+\sigma_e^2} \tag{6}$$

$$\tilde{\beta}_m = \frac{1}{(1+\sigma_e^2)^2}\sum_{n=1}^{N} |\sum_{i=1}^{M}\hat{h}_{ni}\hat{u}_{im}|^2 = \frac{\hat{\zeta}_m}{(1+\sigma_e^2)^2} \tag{7}$$

$I_v(x)$ is the *v*-th order modified Bessel function of the first kind [8]. For systems with $M > N$, there are $M - N$ beams with $\hat{\zeta}_m = 0$. The corresponding eigenvectors can be chosen arbitrarily as far as they are mutually orthogonal and orthogonal to $\{\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_N\}$. Given $\hat{\mathbf{H}}$, $\{\beta_m\}$ of those beams are independent central chi-square distributed, with the conditional pdf given as [8]

$$f_{\beta_m|\hat{\mathbf{H}}}(\beta_m|\hat{\mathbf{H}}) = \frac{1}{\sigma^{2N}\Gamma(N)}\beta_m^{N-1}\exp\left(-\frac{\beta_m}{\sigma^2}\right), \quad m = N+1, \ldots, M. \tag{8}$$

3. POWER ADAPTATION

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. Then the spatial power parameters $\{P_m\}$ are calculated to distribute the total power among the eigen-beams. Our design objective is to minimize the BER.

3.1. BER criterion

The BER for MQAM of modulation size Q with Gray mapping and received SNR ρ can be approximated by a tight bound [9]: $BER_\rho \approx 0.2\exp(-g\rho)$, where $g = 1.5/(Q-1)$ for square MQAM, and $g = 6/(5Q-4)$ for rectangular MQAM. With (3), the BER bound can be written as $P_{b|\hat{\mathbf{H}}}(\bar{\mathbf{H}}|\hat{\mathbf{H}}) = 0.2\exp\left(-\gamma\sum P_m\beta_m\right)$, where $\gamma \triangleq gS/(r\sigma_n^2)$.

Since the realization of $\bar{\mathbf{H}}$ is not available at the transmitter, the transmitter can use the following conditional average BER given $\hat{\mathbf{H}}$:

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

where $f_{\boldsymbol{\beta}|\hat{\mathbf{H}}}(\boldsymbol{\beta}|\hat{\mathbf{H}})$ is the joint conditonal pdf of $\{\beta_m\}$, given $\hat{\mathbf{H}}$. Applying (5), (8) and the independence of $\{\beta_m\}$ to (9), it can be shown that (9) is expressed as

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

3.2. Temporal power allocation

The optimization problem of minimizing the average BER subject to the average power constraint is formulated as

$$\min_{\gamma(\tilde{\boldsymbol{\beta}})\geq 0} E_{\tilde{\boldsymbol{\beta}}}[P_{b|\hat{\mathbf{H}}}] \quad \text{s.t. } E_{\tilde{\boldsymbol{\beta}}}[\gamma(\tilde{\boldsymbol{\beta}})] = g\bar{S}/(r\sigma_n^2) \triangleq \bar{\gamma} \tag{11}$$

where $P_{b|\hat{\mathbf{H}}}$ is given in (10), $\bar{S}$ is the average power budget, $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_1, \ldots, \tilde{\beta}_M)$ denotes the available CSIT defined as (7), and $E_{\tilde{\boldsymbol{\beta}}}[\cdot]$ is the expectation with respect to $\tilde{\boldsymbol{\beta}}$.

This is a calculus of variations problem with an isoperimetric constraint [10]. By introducing the Lagrange multiplier ξ (which is a constant), an auxiliary objective function is defined as

$$\min_{\gamma(\tilde{\boldsymbol{\beta}})\geq 0}\int\left(P_{b|\hat{\mathbf{H}}} + \xi[\gamma(\tilde{\boldsymbol{\beta}}) - \bar{\gamma}]\right) f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}}) d\tilde{\boldsymbol{\beta}} \tag{12}$$

where $f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}})$ is the joint pdf of $\tilde{\boldsymbol{\beta}}$, which can be obtained from the joint pdf of ordered eigenvalues of the Wishart matrix by transformation.

We define $J(\gamma(\tilde{\boldsymbol{\beta}})) \triangleq \left(P_{b|\hat{\mathbf{H}}} + \xi[\gamma(\tilde{\boldsymbol{\beta}}) - \bar{\gamma}]\right) f_{\tilde{\boldsymbol{\beta}}}(\tilde{\boldsymbol{\beta}})$. Since $P_{b|\hat{\mathbf{H}}}$ is convex in γ, the following Euler-Lagrange equation (13) is a necessary and sufficient condition for the optimal $\gamma(\tilde{\boldsymbol{\beta}})$ [10]

$$\frac{\partial J}{\partial\gamma(\tilde{\boldsymbol{\beta}})} = 0. \tag{13}$$

From (13), we have

$$0.2\prod_{m=1}^{M}\frac{1}{(1+\gamma\sigma^2P_m)^N}\exp\left(-\frac{\gamma\tilde{\beta}_mP_m}{1+\gamma\sigma^2P_m}\right) \times \sum_{m=1}^{M}\left(-\frac{N\sigma^2P_m}{1+\gamma\sigma^2P_m} - \frac{\tilde{\beta}_mP_m}{(1+\gamma\sigma^2P_m)^2}\right) + \xi = 0 \tag{14}$$

However, in order to obtain $\gamma(\tilde{\boldsymbol{\beta}})$ by solving Eq.(14) given the CSIT $\tilde{\boldsymbol{\beta}}$, σ^2 and the Lagrange multiplier ξ, we need to know the spatial power allocation $\{P_m\}$, which complicates the computation. To decouple the calculation of the temporal power from that of the spatial allocation, we propose an asymptotic analysis approach. The idea is that based on the optimal spatial power allocation for high/low SNR, we determine the temporal power distribution for given CSIT and then distribute the temporal power spatially in an optimal manner.

At low SNR, the optimal spatial power allocation strategy is to use only the largest eigenmode [1,5,6]. Thus, by substituting $P_1 = 1, P_2 = \cdots = P_M = 0$ into (14), we have

$$\frac{0.2\exp\left(-\frac{\gamma\tilde{\beta}_1}{1+\gamma\sigma^2}\right)}{(1+\gamma\sigma^2)^N}\left(\frac{N\sigma^2}{1+\gamma\sigma^2}+\frac{\tilde{\beta}_1}{(1+\gamma\sigma^2)^2}\right)=\xi. \tag{15}$$

Since the left-hand side of (15) is a monotonically decreasing function of γ, the unique positive γ satisfying (15) exists if $0.2(N\sigma^2+\tilde{\beta}_1) > \xi$, which can be found numerically by a bisection method. If $0.2(N\sigma^2+\tilde{\beta}_1) \le \xi$, the optimal γ would be zero.

At high SNR, the power should be allocated equally among all the eigen-beams [1,5,6]. Considering $P_m = 1/M$ and γ is large, (14) can be approximated as

$$0.2\left(\frac{M}{M+\gamma\sigma^2}\right)^{MN}\exp\left(-\frac{\gamma\sum_{m=1}^{M}\tilde{\beta}_m}{M+\gamma\sigma^2}\right)\frac{MN\sigma^2}{M+\gamma\sigma^2}=\xi. \tag{16}$$

Equation (16) has a unique positive solution when $0.2N\sigma^2 > \xi$. We define a parameter $\lambda \triangleq M/(M+\gamma\sigma^2)$ and rewrite (16) as

$$\lambda^{MN+1}e^{b\lambda}=\frac{5\xi}{N\sigma^2}e^b$$

where $b \triangleq \sum_{m=1}^{M}\tilde{\beta}_m/\sigma^2$. The solution of λ is thus given by the Lambert W function [11] as

$$\lambda=\frac{MN+1}{b}W\left(\frac{b}{MN+1}\left(\frac{5\xi e^b}{N\sigma^2}\right)^{\frac{1}{MN+1}}\right) \tag{17}$$

where $W(\cdot)$ denotes the principal branch of the Lambert W function. After we obtain λ from (17), γ is calculated as

$$\gamma=\frac{M}{\sigma^2}\left(\frac{1}{\lambda}-1\right). \tag{18}$$

Equations (17) and (18) provide the closed-form formulae to calculate the temporal power.

In regard to the Lagrange multiplier ξ in the above solutions, it can be determined off-line from the average power constraint in (11). At low SNR, the average power constraint can be expressed as

$$\bar{\gamma}=\int_0^\infty \gamma(\tilde{\beta}_1)f_{\tilde{\beta}_1}(\tilde{\beta}_1)d\tilde{\beta}_1=\int_0^\infty \gamma\left(\frac{\zeta_1}{1+\sigma_e^2}\right)f_{\zeta_1}(\zeta_1)d\zeta_1 \tag{19}$$

where ζ_1 is the largest eigenvalue of the Wishart matrix $\mathbf{H}^H\mathbf{H}$, with its pdf $f_{\zeta_1}(\zeta_1)$ given as [12] (assuming $M \ge N$ without loss of generality)

$$f_{\zeta_1}(\zeta_1)=\sum_{i=1}^{N}\sum_{j=M-N}^{(M+N)i-2i^2} d_{i,j}\frac{i^{j+1}\zeta_1^j e^{-i\zeta_1}}{j!}$$

where $d_{i,j}$'s are constants given in [12]. For a given ξ, we use the Gauss-Laguerre numerical integration formula [13] to calculate the integration in (19), where $\gamma(\frac{\zeta_1}{1+\sigma_e^2})$ is obtained by substituting $\tilde{\beta}_1 = \frac{\zeta_1}{1+\sigma_e^2}$ into (15) and solving the resulting equation for all ζ_1. Since the integration in (19) is a monotonic function of ξ, the value of ξ can be refined by a bisection method in comparing the computed expectation value with the target $\bar{\gamma}$. At high SNR, the average power constraint can be expressed as

$$\bar{\gamma}=\int_0^\infty \gamma(b)f_b(b)db=\int_0^\infty \gamma\left(\frac{t}{(1+\sigma_e^2)\sigma^2}\right)f_t(t)dt \tag{20}$$

where $t \triangleq \|\mathbf{H}\|_F^2$, which is central chi-square distributed with the degree of freedom of $2L$, and $L = M \times N$. The pdf of t is given as

$$f_t(t)=\frac{1}{(L-1)!}t^{L-1}e^{-t}.$$

Similarly, the integration in (20) is calculated by the Gauss-Laguerre method for a given ξ, where $\gamma\left(\frac{t}{(1+\sigma_e^2)\sigma^2}\right)$ is obtained based on (17) and (18) by replacing b in (17) with $\frac{t}{(1+\sigma_e^2)\sigma^2}$. ξ can be refined by a bisection method due to the monotonic property of (20).

3.3. Spatial power allocation

We can choose any existing spatial power allocation algorithms to distribute the obtained temporal power. To maintain computationally efficient, here we use the closed-form scheme proposed in [5]. The spatial power is calculated as

$$P_m=\frac{\hat{\tilde{\beta}}_m^{\mu}}{\sum_{l=1}^{M}\hat{\tilde{\beta}}_l^{\mu}},\ m=1,\ldots,M \tag{21}$$

where

$$\mu=\frac{M\sum_{m=1}^{M}\tilde{\beta}_m a_m}{N\gamma\sigma^4\sum_{m=1}^{M}a_m^2+[2M\gamma\sigma^2/(M+\gamma\sigma^2)]\sum_{m=1}^{M}\tilde{\beta}_m a_m^2}$$

and $\hat{\tilde{\beta}}_m=\tilde{\beta}_m+N\sigma^2$, $a_m=\log\hat{\tilde{\beta}}_m-\frac{1}{M}\sum_{l=1}^{M}\log\hat{\tilde{\beta}}_l$.

3.4. Switching between the two asymptotic solutions of γ

The average BER performance of the spatial-temporal power adaptation can be obtained by evaluating $E_{\tilde{\boldsymbol{\beta}}}[P_{b|\hat{\mathbf{H}}}]$ numerically, where $P_{b|\hat{\mathbf{H}}}$ is given in (10). For the given SNR ($\bar{S}/\sigma_n^2$), one of the asymptotic solutions of γ producing a lower average BER is selected. Actually, the switching point between low SNR and high SNR can be predetermined off-line.

3.5. Summary of the proposed algorithm

The online operations of the proposed spatial-temporal power adaptation are summarized as follows.

1) Compute the eigenvalues $\{\hat{\zeta}_m\}$ of $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ and then calculate $\{\tilde{\beta}_m\}$ using (7).
2) Based on the SNR ($\bar{S}/\sigma_n^2$), choose one of the asymptotic solutions to calculate γ.
3) Temporal power : At low SNR, if $0.2(N\sigma^2+\tilde{\beta}_1) \le \xi$, let $\gamma = 0$; otherwise use a bisection method to find γ by solving (15). At high SNR, use (17) and (18) to calculate γ.
4) Spatial power : Use (21) to compute $\{P_m\}$.

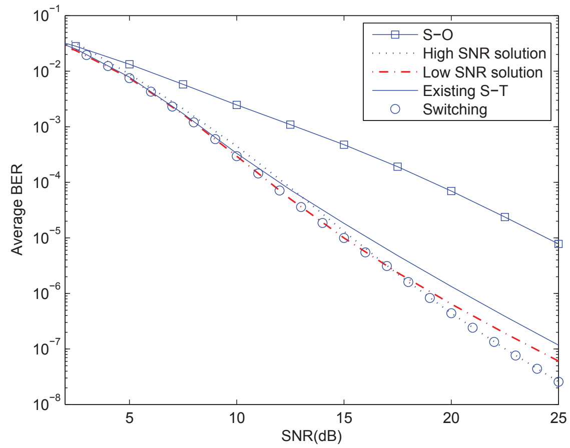

Fig. 2. BER performance of BPSK for different schemes, $M = 2, N = 1, \sigma_e^2 = 0.05, r = 1$ (Alamouti code).

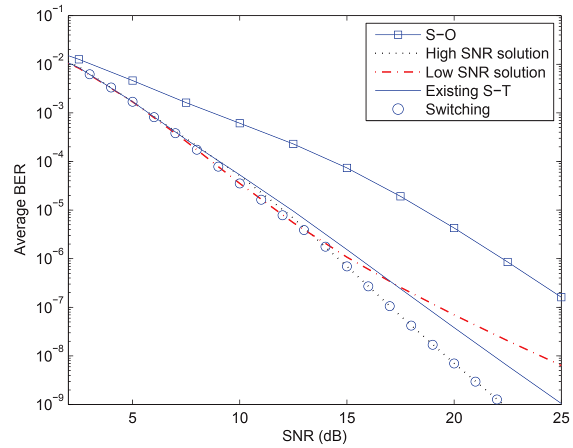

Fig. 3. BER performance of QPSK for different schemes, $M = 3, N = 1, \sigma_e^2 = 0.1, r = 1/2$ (G_3 code).

4. NUMERICAL RESULTS AND CONCLUSION

In Fig.2 and Fig.3, we compare the performance of the proposed power adaptation scheme with that of the existing spatial-only (S-O) scheme [5] and the spatial-temporal (S-T) scheme proposed in [4]. The configuration of the systems are shown in the captions. As shown in these figures, power adaptation both in time and space can significantly reduce the BER in comparison with spatial-only power allocation. The proposed S-T algorithm switches from the low SNR solution to the high SNR solution at SNR of 17.2dB in Fig.2 and at SNR of 13.4dB in Fig.3. Compared to the existing S-T method, our proposed algorithm provides considerable performance gains. For example, Fig.2 shows that the proposed S-T scheme begins to manifest its superiority at the BER of 10^{-3} and achieves the gains of 1dB and 1.8dB at BER levels of 10^{-5} and 10^{-6} respectively. For more stringent BER requirement, the power saving for the proposed algorithm is even more significant.

For the proposed S-T scheme, in order to reduce the computational complexity, the spatial power allocation is assumed known a priori when doing temporal power calculation. Actually, it is a suboptimal solution for medium SNRs. Performance can be further improved if we can jointly optimize the spatial and temporal power parameters.

5. REFERENCES

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

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

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

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

[5] Quan Kuang, Shu-Hung Leung, and Xiangbin Yu, "Closed-form power allocation scheme for space-time coded multiple-antenna systems with imperfect CSI," in *Proc. IEEE Int Acoustics Speech and Signal Processing (ICASSP) Conf*, 2010, pp. 3442–3445.

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

[7] Hyundong Shin and Jae Hong Lee, "Performance analysis of space-time block codes over keyhole Nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 351–362, March 2004.

[8] J. G. Proakis, *Digital Communications*, New York: McGraw-Hill, 4th edition, 2001.

[9] Shengli 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.

[10] Enid R. Pinch, *Optimal Control and the Calculus of Variations*, New York: Oxford University Press, 1993.

[11] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, vol. 5, pp. 329–359, 1996.

[12] P. A. Dighe, R. K. Mallik, and S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Trans. Commun.*, vol. 51, no. 4, pp. 694–703, 2003.

[13] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, New York: Academic Press, 1975.