Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2011, Article ID 785437, 9 pages
doi:10.1155/2011/785437

Research Article

Power Allocation between Pilot and Data Symbols for MIMO Systems with MMSE Detection under MMSE Channel Estimation

Jun Wang,[1] Oliver Yu Wen,[2] Hongyang Chen,[3] and Shaoqian Li[1]

[1] *National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China*
[2] *Qualcomm Inc., Santa Clara, CA 95051, USA*
[3] *Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan*

Correspondence should be addressed to Jun Wang, junwang@uestc.edu.cn

Received 12 June 2010; Revised 6 December 2010; Accepted 3 January 2011

Academic Editor: Marc Moonen

For multiple-input multiple-output (MIMO) wireless communication system with minimum mean square error (MMSE) detection, a new scheme of power allocation between pilot and data symbols is investigated under MMSE channel estimation in this paper. First, we propose a novel soft-output MMSE MIMO detector by taking into consideration the channel estimation error. Then, through the application of random matrix theorem, we propose an efficient scheme for power allocation between pilot and data symbols which maximizes the lower bound of postprocessing signal-to-interference-and-noise ratio (SINR) for MIMO systems with equal number of transmitter and receiver antennas. We have proven the existence and uniqueness of the proposed optimal power allocation settings. Furthermore, our analysis shows that the proposed power allocation is also valid and applicable for those MIMO systems with unequal number of transmitter and receiver antennas. Finally, our extensive simulation results have validated this novel power allocation scheme.

1. Introduction

For wireless multiple-input multiple-output (MIMO) communication systems, most of the popular MIMO channel estimation approaches are based on maximum likelihood (ML) or minimum mean square error (MMSE) criteria while employing pilot symbols [1–4]. As the estimated MIMO channel matrix is inevitably imperfect, the channel estimation error should be taken into account while designing MIMO detectors in order to achieve better performances. This design philosophy has been investigated in [4] for ML detector and in [5] for V-BLAST detector.

For MIMO systems with channel coding, it is well known that the MIMO detector needs to output log-likelihood ratio (LLR), that is, soft information of each coded bit to the channel decoder so that more desirable performance can be achieved by soft decoding over hard decoding [6]. Soft-output linear minimum mean square error (MMSE) detection is widely applied in practical systems, because it has relatively lower, and therefore, manageable complexity while suppressing both coantenna interference (CAI) and noise at the same time [6]. Existing soft-output MMSE MIMO detectors are derived based on the assumption of the perfect channel estimation. However, its LLR generation needs to be modified to take the imperfect channel estimation into consideration in order to achieve better performances as in single antenna systems [7]. Consequently, we herein propose a novel soft-output MMSE MIMO detector with this consideration.

For pilot-symbols-assisted channel estimation, a fundamental question is how to allocate the transmission power between pilot symbols and data symbols in order to optimize system performances. In the case of the block-fading channel, this question has been investigated with the objective of maximizing the achievable channel capacity [2, 8] or for precoder design [9]. Based on our new MIMO detector, we address this fundamental question of optimal power allocation with the objective of optimizing bit error rate

(BER) performance, through the maximization of the mean value of the minimum postprocessing signal-to-interference-and-noise ratio (SINR). As a result, we have obtained an explicit and concise closed-form solution for MIMO systems with equal numbers of transmitter and receiver antennas. One very desirable property of our proposed scheme is that it does not need any form of channel coefficients feedback and has negligible added complexity. We also prove the existence and uniqueness of the optimized power allocation analytically. Meanwhile, we have also proven that our new scheme is valid and applicable for the MIMO systems which have unequal number of transmitter and receiver antennas. Extensive simulation results validate the advantages of our proposed power allocation scheme in various scenarios.

Compared with our scheme, it is worth noting that the power allocation scheme in [9] is relatively simplified by precoding. As a result of precoding, each spatial data stream has the same mean square error (MSE); therefore their power allocation can be implemented by minimizing this single MSE. However, if precoding is not applied, the power allocation in [9] will not be valid any more.

The rest of this paper is organized as follows. In Section 2, we describe the system model used in this paper. In Section 3, we derive our new soft-output MMSE MIMO detector for MMSE channel estimation. The power allocation scheme and corresponding remarks are provided in Section 4. The simulation results are given in Section 5. Finally, we conclude this paper in Section 6.

2. System Model

In a MIMO system with M_T transmitter antennas and M_R receiver antennas, the received signal can be modeled as follows:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{y} = [y_1, y_2, \ldots, y_{M_R}]^T$ is the received vector symbol and $\mathbf{H}$ is the $M_R \times M_T$ MIMO channel coefficient matrix. The element $h_{i,j}$ of $\mathbf{H}$ denotes the channel fading coefficient between the jth transmitter antenna and the ith receiver antenna. We assume that the elements of $\mathbf{H}$ are independent and identically distributed (i.i.d.) zero-mean circularly symmetric complex Gaussian (ZMCSCG) random variables with variance $\mathbb{E}\{|h_{i,j}|^2\} = 1$ and the notation $\mathbb{E}(\cdot)$ denotes the expectation operator. Meanwhile, we assume that there are $N_P + N_D$ vector symbols in each frame, where N_D vector symbols are used for data transmission and the remaining N_P vector symbols for pilot symbols transmission. In this paper, to make mathematical derivations tractable, the $h_{i,j}$ is assumed to be flat-fading and constant within a frame of $N_P + N_D$ MIMO vector symbols, which varies randomly from one frame to another, that is, block-fading channel. The performance results for time-varying fading channel will be provided through simulation, and $\mathbf{s} = [s_1, s_2, \ldots, s_{M_T}]^T$ is the complex vector symbol whose element s_i is taken from a complex modulation constellation $\mathcal{A}$ by mapping channel coded bit vector $\mathbf{b}_i$. Furthermore, we assume that each transmitted symbol is independently taken from the same modulation constellation and has the same mean power, that is, $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = P_d\mathbf{I}_{M_T}$ and $\mathbf{I}_{M_T}$ is a $M_T \times M_T$ identity matrix. Finally, $\mathbf{n}$ is an i.i.d. ZMCSCG noise vector with covariance matrix $\mathbb{E}\{\mathbf{n}\mathbf{n}^H\} = \sigma_n^2\mathbf{I}_{M_R}$.

3. MMSE MIMO Detection under MMSE Channel Estimation

3.1. MMSE MIMO Channel Estimation. We herein consider pilot symbols (or training sequences) based channel estimation. As mentioned in the previous section, for each frame of $N_P + N_D$ MIMO vector symbols, the MIMO channel realization is constant, during which N_P symbol intervals are dedicated to pilot symbols, and the remaining N_D to data. The total power available for each frame is P_t. We further assume that power allocated to pilot symbols is distributed uniformly among them, and the same is true for power allocated to data symbols, that is,

$$\left(N_P P_p + N_D P_d\right) M_T = P_t, \tag{2}$$

where P_p and P_d are the mean transmission power of each pilot and data symbol, respectively.

For MMSE MIMO channel estimation, the estimated MIMO channel matrix can be expressed as [1–4]

$$\begin{aligned}\hat{\mathbf{H}} &= (\mathbf{H}\mathbf{S}_P + \mathbf{N})\mathbf{S}_P^H\left(\sigma_n^2\mathbf{I}_{M_T} + \mathbf{S}_P\mathbf{S}_P^H\right)^{-1} \\ &= \mathbf{H} - \Delta\mathbf{H},\end{aligned} \tag{3}$$

where $\mathbf{S}_P$ is a $M_T \times N_P$ pilot symbol matrix and $N_P \geq M_T$; $\Delta\mathbf{H}$ is the zero-mean channel estimation error matrix, which is uncorrelated with $\hat{\mathbf{H}}$ by orthogonality principle of MMSE estimation [10]; $\mathbf{N} = [\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_{N_P}]$ is a $M_R \times N_P$ ZMCSCG noise matrix and $\mathbb{E}\{\mathbf{N}\mathbf{N}^H\} = N_P\sigma_n^2\mathbf{I}_{M_R}$. For i.i.d. MIMO channels, it has been proven that the optimal $\mathbf{S}_P$ minimizing MSE satisfies the following equation [1, 2]:

$$\mathbf{S}_P\mathbf{S}_P^H = N_P P_p \mathbf{I}_{M_T}. \tag{4}$$

In this case, $\hat{\mathbf{H}}$ has independent zero mean complex Gaussian entries with covariance matrix [2]

$$\mathbb{E}\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right\} = M_T\left(1 - \sigma_{\Delta h}^2\right)\mathbf{I}_{M_R} = M_T\sigma_{\hat{h}}^2\mathbf{I}_{M_R}, \tag{5}$$

where $\sigma_{\Delta h}^2 = \sigma_n^2/(\sigma_n^2 + N_P P_p)$ is the variance of each element in $\Delta\mathbf{H}$.

3.2. MMSE Filter and LLR Computation under Imperfect Channel Estimation. When we consider the channel estimation error, (1) can be rewritten as

$$\mathbf{y} = \left(\hat{\mathbf{H}} + \Delta\mathbf{H}\right)\mathbf{s} + \mathbf{n}. \tag{6}$$

Then, conditionally on $\hat{\mathbf{H}}$, the MMSE filter is given as in [10]

$$\widetilde{\mathbf{W}} = \mathbb{E}\left\{\mathbf{s}\mathbf{y}^H \mid \hat{\mathbf{H}}\right\}\mathbb{E}\left\{\mathbf{y}\mathbf{y}^H \mid \hat{\mathbf{H}}\right\}^{-1}, \tag{7}$$

where

$$\mathbb{E}\left\{\mathbf{y}\mathbf{y}^H \mid \hat{\mathbf{H}}\right\} = P_d\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \hat{\mathbf{H}}\mathbb{E}\left\{\Delta\mathbf{H}^H \mid \hat{\mathbf{H}}\right\} + \mathbb{E}\left\{\Delta\mathbf{H} \mid \hat{\mathbf{H}}\right\}\hat{\mathbf{H}}^H + \mathbb{E}\left\{\Delta\mathbf{H}\Delta\mathbf{H}^H \mid \hat{\mathbf{H}}\right\}\right) + \sigma_n^2\mathbf{I}_{M_R}. \tag{8}$$

As $\Delta\mathbf{H}$ is uncorrelated with $\hat{\mathbf{H}}$ according to the analysis in the previous subsection, we have

$$\mathbb{E}\left\{\Delta\mathbf{H} \mid \hat{\mathbf{H}}\right\} = \mathbb{E}\left\{\Delta\mathbf{H}^H \mid \hat{\mathbf{H}}\right\} = \mathbf{0},$$
$$\mathbb{E}\left\{\Delta\mathbf{H}\Delta\mathbf{H}^H \mid \hat{\mathbf{H}}\right\} = M_T\sigma_{\Delta h}^2\mathbf{I}_{M_R}. \tag{9}$$

Then, by substituting (9) into (8), we have

$$\mathbb{E}\left\{\mathbf{y}\mathbf{y}^H \mid \hat{\mathbf{H}}\right\} = P_d\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \left(\sigma_n^2 + M_T P_d\sigma_{\Delta h}^2\right)\mathbf{I}_{M_R}. \tag{10}$$

Meanwhile, through similar manipulations described above, we obtain

$$\mathbb{E}\left\{\mathbf{s}\mathbf{y}^H \mid \hat{\mathbf{H}}\right\} = \mathbb{E}\left\{\mathbf{s}\mathbf{s}^H\left(\hat{\mathbf{H}} + \Delta\mathbf{H}\right)^H \mid \hat{\mathbf{H}}\right\} = P_d\hat{\mathbf{H}}^H. \tag{11}$$

Finally, by combining (10) and (11) into (7), we reach the MMSE filter conditionally on $\hat{\mathbf{H}}$ as

$$\widetilde{\mathbf{W}} = P_d\hat{\mathbf{H}}^H\left[P_d\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \left(\sigma_n^2 + M_T P_d\sigma_{\Delta h}^2\right)\mathbf{I}_{M_R}\right]^{-1}. \tag{12}$$

Note that when the term $M_T P_d\sigma_{\Delta h}^2$ is dropped, it is reduced to the MMSE filter with perfect channel estimation [6].

By applying MMSE filter $\widetilde{\mathbf{W}}$ to $\mathbf{y}$, we have the output associated with s_i as

$$\hat{s}_i = \widetilde{\mathbf{W}}^i\mathbf{y} = P_d\hat{\mathbf{h}}_i^H\left[P_d\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \left(\sigma_n^2 + M_T P_d\sigma_{\Delta h}^2\right)\mathbf{I}_{M_R}\right]^{-1}\mathbf{y}, \tag{13}$$

where $\widetilde{\mathbf{W}}^i$ denotes the ith row of $\widetilde{\mathbf{W}}$. By applying the Gaussian approximation of the outputs of MMSE filter [11], we have

$$\hat{s}_i \approx \mu_i s_i + \xi_i, \tag{14}$$

where $\mu_i = \widetilde{\mathbf{W}}^i\hat{\mathbf{h}}_i$ and ξ_i is a zero-mean complex Gaussian random variable with variance $P_d(\mu_i - \mu_i^2)$. Therefore, the LLR value of the λth coded bit of $\mathbf{b}_i$ can be approximated as [12]

$$L\left(\mathbf{b}_i^\lambda\right) \approx \frac{\min_{a_i\in\mathcal{A}_\lambda^0}\left|\hat{s}_i - \mu_i a_i\right|^2 - \min_{a_i\in\mathcal{A}_\lambda^1}\left|\hat{s}_i - \mu_i a_i\right|^2}{P_d(\mu_i - \mu_i^2)}, \tag{15}$$

where $\mathcal{A}_\lambda^0$ and $\mathcal{A}_\lambda^1$ denote the modulation constellation symbols subset of $\mathcal{A}$, whose λth bit equals 0 or 1, respectively.

Note that the differences between our scheme and the method used in [5] are two fold. First, ML channel estimation is adopted in [5] while MMSE channel estimator is investigated in our scheme. Secondly, the method of [5] is derived by decomposing MMSE filter into two stages. The 1st stage is the original MMSE filter without channel estimation error, and the 2nd stage is the deviation from the original MMSE filter of the 1st stage due to channel estimation error. The deviation is derived based on MMSE criteria. Whereas our scheme is an unified derivation covering these two stages in a solution.

Finally, it is worth pointing out that we have actually taken into account the channel estimation error in the computation of LLR value. Thus, similar to the results of single antenna system in [7], better BER performance can be expected as a result.

4. Power Allocation between Pilot and Data Symbols

In this paper, the objective of our power allocation optimization is the minimization of the cost function of BER. Even though [13, 14] have provided a closed-form BER expressions for MMSE MIMO detection, which is based on the moment generating function and exponential integral function, it is unfortunately not easy to be treated. On the other hand, as BER is a monotonically decreasing function of SINR we can reach our goal through maximizing the SINR. Although the distribution of the SINR has been obtained in [13, 15], it is too complex to be handled in our situation. Therefore, in this section, we first develop a new approximation of the minimum SINR, and then derive an effective power allocation scheme by maximizing the average lower bound of minimum SINR when $M_T = M_R$. We prove that an optimal power allocation exits and is unique. Furthermore, we show that the proposed power allocation can also be applied to the case of $M_T \neq M_R$ through approximation.

4.1. Power Allocation for the Case of $M_T = M_R$. As in the case of perfect channel estimation [6, 16], by considering the model of (6), the postprocessing SINR for the ith substream ($i \in \{1, 2, \ldots, M_T\}$) is given by

$$\gamma_i = \frac{1}{\left[\left(\mathbf{I}_{M_T} + (P_d/(\sigma_n^2 + P_d M_T\sigma_{\Delta h}^2))\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right)^{-1}\right]_{i,i}} - 1. \tag{16}$$

Our goal is to reach a power allocation between pilot and data symbols which maximizes $(\gamma_i)_{\min}$ subject to the total transmitted power constraint, that is,

$$P_{p,\mathrm{op}} = \arg\max_{P_p}(\gamma_i)_{\min},$$
$$\text{s.t.} \quad N_P P_p + N_D P_d = \frac{P_t}{M_T} = \widetilde{P}. \tag{17}$$

In order to tackle this complex optimization problem in a more tractable manner, we resort to maximize the lower bound of $(\gamma_i)_{\min}$. Based on the fact that the largest eigenvalue

majorizes the largest diagonal term of a square matrix [17], we have

$$\begin{aligned}(\gamma_i)_{\min} &\geq \lambda_{\min}\left(\mathbf{I}_{M_T} + \frac{P_d}{\sigma_n^2 + P_d M_T \sigma_{\Delta h}^2}\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right) - 1 \\ &= \frac{P_d}{\sigma_n^2 + P_d M_T \sigma_{\Delta h}^2}\lambda_{\min}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right),\end{aligned} \tag{18}$$

where the notation $\lambda_{\min}(\mathbf{M})$ denotes the minimum eigenvalue of matrix $\mathbf{M}$. Then, we can rewrite the optimization problem (17) as

$$\begin{aligned}P_{p,\text{op}} &= \arg\max_{P_p}\left(\frac{P_d}{\sigma_n^2 + P_d M_T \sigma_{\Delta h}^2}\lambda_{\min}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right)\right) \\ &\text{s.t.} \quad N_P P_p + N_D P_d = \widetilde{P}.\end{aligned} \tag{19}$$

As $\lambda_{\min}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)$ is of exponential distribution with parameter $M_T/\sigma_{\hat{h}}^2$ in the case of $M_T = M_R$ [18], we have

$$\begin{aligned}\mathbb{E}\left\{\lambda_{\min}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right)\right\} &= \frac{\sigma_{\hat{h}}^2}{M_T} = \frac{N_P P_p}{\left(M_T\left(\sigma_n^2 + M_T P_p\right)\right)}, \\ \text{Var}\left\{\lambda_{\min}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right)\right\} &= \left(\mathbb{E}\left\{\lambda_{\min}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}\right)^H\right\}\right)^2 < \left(\frac{N_P}{M_T^2}\right)^2.\end{aligned} \tag{20}$$

Considering that N_P usually satisfies $N_P < M_T^2$ in practical applications, we can reasonably assume that $\text{Var}\{\lambda_{\min}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)\}$ is relatively small. For example, $\text{Var}\{\lambda_{\min}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)\} < 1$ when $M_T = 2$, $N_P = 2M_T$, and $\text{Var}\{\lambda_{\min}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)\} < 0.25$ in the case of $M_T = 4$, $N_P = 2M_T$. This means that the value of $\lambda_{\min}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)$ is close to its mean value $\mathbb{E}\{\lambda_{\min}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)\}$. Therefore, we can replace $\lambda_{\min}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)$ by $\mathbb{E}\{\lambda_{\min}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)\}$ and further re-write the optimization problem (19) as

$$\begin{aligned}P_{p,\text{op}} &= \arg\max_{P_p}\left(\frac{P_d}{\sigma_n^2 + P_d M_T \sigma_{\Delta h}^2}\mathbb{E}\left\{\lambda_{\min}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right)\right\}\right) \\ &\text{s.t.} \quad N_P P_p + N_D P_d = \widetilde{P}.\end{aligned} \tag{21}$$

Based on the results in the previous section, and also considering $A = \sigma_n^2 M_T N_P (N_D - M_T)$, $B = M_T \sigma_n^2 (M_T \widetilde{P} + N_D \sigma_n^2)$, we can finally express our optimization objective function as

$$P_{p,\text{op}} = \arg\max_{P_p}\left(\frac{\left(\widetilde{P} - N_P P_p\right)N_P P_p}{AP_p + B}\right), \quad P_p \in \left\{0, \frac{\widetilde{P}}{N_P}\right\}. \tag{22}$$

Then, it can be solved by differentiating the objective function. As a result, we get

$$P_{p,\text{op}} = \begin{cases}\eta\left(\sqrt{1 + \dfrac{\widetilde{P}}{(\eta N_P)}} - 1\right), & \text{if } M_T \neq N_D, \\ \dfrac{\widetilde{P}}{(2N_P)}, & \text{if } M_T = N_D,\end{cases} \tag{23}$$

where $\eta = B/A$. As A and B are all scalars, $P_{p,\text{op}}$ can be easily computed. In some cases, we are interested in the ratio $\alpha = P_d/P_p$. Based on (23), the optimal ratio is given as

$$\alpha_{\text{op}} = \begin{cases}\dfrac{\widetilde{P}}{\left(N_D P_{p,\text{op}}\right)} - \dfrac{N_P}{N_D}, & \text{if } M_T \neq N_D, \\ \dfrac{N_D}{N_P}, & \text{if } M_T = N_D.\end{cases} \tag{24}$$

4.2. Remarks on the Proposed Power Allocation Scheme. In this subsection, we further discuss the existence and uniqueness of the optimal power allocation $P_{p,\text{op}}$. Our analysis is carried out with respect to two cases, that is, $M_T = N_D$ and $M_T \neq N_D$, respectively.

(1) Case of $M_T = N_D$. Let

$$f\left(P_p\right) = \frac{\left(\widetilde{P} - N_P P_p\right)N_P P_p}{AP_p + B}, \tag{25}$$

we have the first-order differentiation of $f(P_p)$ as

$$f'\left(P_p\right) = \frac{BN_P\left(\widetilde{P} - 2N_P P_p\right)}{\left(AP_p + B\right)^2}\begin{cases}> 0, & \text{if } P_p < \dfrac{\widetilde{P}}{(2N_P)} \\ = 0, & \text{if } P_p = \dfrac{\widetilde{P}}{(2N_P)} \\ < 0, & \text{if } P_p > \dfrac{\widetilde{P}}{(2N_P)}.\end{cases} \tag{26}$$

Then, $f(P_p)$ is a monotonically increasing function in the region $P_p < \widetilde{P}/(2N_P)$, and a monotonically decreasing function in the region $P_p > \widetilde{P}/(2N_P)$. Therefore, $f(P_p)$ has an unique maximal value in the region $(0, \widetilde{P}/N_P)$.

(2) Case of $M_T \neq N_D$. For this case, we have the following proposition.

Proposition 1. *$f(P_p)$ is a convex function when $M_T \neq N_D$.*

Proof. To prove this proposition, we need to prove that the second-order differentiation of $f(P_p)$ satisfies $f''(P_p) < 0$ when $M_T \neq N_D$ [19]. The second-order differentiation of $f(P_p)$ is given as

$$f''\left(P_p\right) = \frac{-2BN_P\left(BN_P + A\widetilde{P}\right)}{\left(AP_p + B\right)^3}, \quad P_p \in \left(0, \frac{\widetilde{P}}{N_P}\right). \tag{27}$$

First, let us consider the case of $N_D > M_T$. For this case, we have $A > 0$. Consequently, we have $(AP_p + B)^3 > 0$ and $BN_P(BN_P + A\widetilde{P}) > 0$. Therefore, we have $f''(P_p) < 0$.

On the other hand, for the case of $N_D < M_T$, we have $A < 0$. Moreover, we have

$$AP_p + B = \sigma_n^2 M_T N_D\left(N_P P_p + M_T\right) + \sigma_n^2 M_T^2\left(\widetilde{P} - N_P P_p\right). \tag{28}$$

By considering $\widetilde{P} = N_P P_p + N_D P_s > N_P P_p$, we have $(AP_p + B)^3 > 0$. Meanwhile, we have

$$BN_P + A\widetilde{P} = \sigma_n^2 M_T N_P N_D \left(\sigma_n^2 + \widetilde{P}\right) > 0. \tag{29}$$

Therefore, we always have $f''(P_p) < 0$. □

It follows Proposition 1 that $f(P_p)$ has an unique maximal value which is given by (23). In what follows, we present the condition under which this maximal value lies in the region $(0, \widetilde{P}/N_P)$. Equivalently, the condition ensures $0 < \eta(\sqrt{1 + \widetilde{P}/(\eta N_P)} - 1) < \widetilde{P}/N_P$.

Proposition 2. *When $N_D \neq M_T$, one always has $0 < \eta(\sqrt{1 + \widetilde{P}/(\eta N_P)} - 1) < \widetilde{P}/N_P$.*

Proof. First, let us consider the case of $N_D > M_T$. For this case, it is obvious that $0 < \eta(\sqrt{1 + \widetilde{P}/(\eta N_P)} - 1)$ as $\eta > 0$. Also, as

$$\eta\left(\sqrt{1 + \frac{\widetilde{P}}{(\eta N_P)}} - 1\right) = \frac{\widetilde{P}}{N_P} \underbrace{\left[\frac{M_T + N_D\left(\sigma_n^2/\widetilde{P}\right)}{N_D - M_T}\left(\sqrt{1 + \frac{N_D - M_T}{M_T + N_D\left(\sigma_n^2/\widetilde{P}\right)}} - 1\right)\right]}_{\kappa}, \tag{30}$$

we have $\kappa < 1$ to satisfy $\eta(\sqrt{1 + \widetilde{P}/(\eta N_P)} - 1) < \widetilde{P}/N_P$. As for the condition which leads to $\kappa < 1$, for the sake of notational simplicity, let $\varsigma = \sigma_n^2/\widetilde{P}$, it is therefore $\varsigma > 0$. As

$$\kappa = \frac{1}{N_D - M_T}\left[N_D\left(\sqrt{(1+\varsigma)\left(\frac{M_T}{N_D} + \varsigma\right)} - \varsigma\right) - M_T\right], \tag{31}$$

which results in $\sqrt{(1+\varsigma)(M_T/N_D + \varsigma)} - \varsigma < 1$ to satisfy $\kappa < 1$. Consequently, we reach

$$\sqrt{(1+\varsigma)\left(\frac{M_T}{N_D} + \varsigma\right)} < 1 + \varsigma \Longleftrightarrow \frac{M_T}{N_D} < 1 \Longleftrightarrow M_T < N_D. \tag{32}$$

Therefore, we have $0 < \eta(\sqrt{1 + \widetilde{P}/(\eta N_P)} - 1) < \widetilde{P}/N_P$ for the case of $N_D > M_T$.

Subsequently, we consider the case of $N_D < M_T$. For this case, we have $A < 0$ and then $\eta < 0$. As

$$\frac{\widetilde{P}}{\eta N_P} = \frac{(N_D - M_T)\widetilde{P}}{M_T\widetilde{P} + N_D\sigma_n^2} = -\frac{M_T\widetilde{P} - N_D\widetilde{P}}{M_T\widetilde{P} + N_D\sigma_n^2}, \tag{33}$$

we further have

$$-1 < \frac{\widetilde{P}}{(\eta N_P)} < 0 \Longrightarrow \sqrt{1 + \frac{\widetilde{P}}{(\eta N_P)}} - 1 < 0. \tag{34}$$

Therefore, we get $0 < \eta(\sqrt{1 + \widetilde{P}/(\eta N_P)} - 1)$. On the other hand, let

$$\rho = \sqrt{(1+\varsigma)\left(\frac{M_T}{N_D} + \varsigma\right)} - \varsigma, \tag{35}$$

which satisfies $\rho > 1$ in the case of $N_D < M_T$. We then rewrite (31) as

$$\kappa = \frac{\rho N_D - M_T}{N_D - M_T}. \tag{36}$$

Since $\kappa > 0$, we have $\rho N_D - M_T < 0$, which immediately leads to $\kappa < 1$. □

Based on Propositions 1 and 2, we can conclude that our proposed optimal power allocation always exists and is unique.

Finally, it is worthy of pointing out that (23) can only achieve suboptimal performance as it falls short on our optimization target, which is to maximize the average lower bound of the minimal postprocessing SINR. Nevertheless, our simulation results provided in Section 5 show that (23) can achieve near-optimum performance.

4.3. Power Allocation for the Case of $M_T \neq M_R$. When $M_T \neq M_R$, the distribution of the smallest eigenvalue of $\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H$ cannot be expressed in a concise form [18, 20]. As a result, the power allocation cannot be treated in the same way as in the case of $M_T = M_R$. Since $\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H$ is a Wishart matrix, we first recall the distribution of the smallest eigenvalue of a Wishart matrix, then derive an approximated power allocation.

The cumulative distribution function (cdf) of the smallest eigenvalue $\lambda_{\min}$ of a complex Wishart matrix $\mathbf{X} \sim \mathrm{CW}_m(n, \mathbf{I})$ is given by [18]

$$F_{\lambda_{\min}}(x) = 1 - \mathrm{etr}(-x\mathbf{I}) \sum_{k=0}^{m(n-m)} \widehat{\sum_{\kappa}} \frac{C_\kappa(x\mathbf{I})}{k!}, \tag{37}$$

where $\widehat{\sum}_\kappa$ denotes summation over the partitions $\kappa = (k_1, \ldots, k_m)$ of k with $k_1 \leq n - m$; the $C_\kappa(\mathbf{X})$ denotes the complex zonal polynomials of a complex matrix $\mathbf{X}$, and $\mathrm{etr}(\mathbf{X}) = \exp(\mathbf{tr}(\mathbf{X}))$. Equation (37) can be further rewritten as

$$F_{\lambda_{\min}}(x) = 1 - \mathrm{etr}(-x\mathbf{I}) - \Delta, \tag{38}$$

where

$$\Delta = \mathrm{etr}(-x\mathbf{I}) \sum_{k=1}^{m(n-m)} \widehat{\sum_{\kappa}} \frac{C_\kappa(x\mathbf{I})}{k!}. \tag{39}$$

So, if we omit the item Δ in (38), the distribution of the smallest eigenvalue of a Wishart matrix with $n \neq m$ can be approximated as an exponential distribution. Therefore, the proposed power allocation for the case of $M_T = M_R$ can still be applied to the case of $M_T \neq M_R$. According to (39), the error resulted from this approximation is determined by the difference between M_T and M_R. For example, the cdf of the

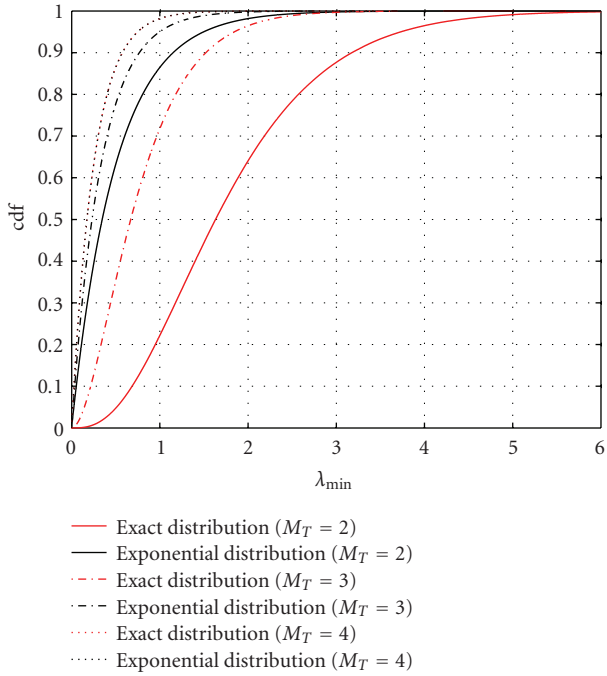

Figure 1: The cdf of the smallest eigenvalue of $\hat{\mathbf{H}}\hat{\mathbf{H}}^H$ and corresponding exponential distribution approximation ($M_R = 4$, $N_P = 4$, $P_d/\sigma_n^2 = 15$ dB).

smallest eigenvalue of $\hat{\mathbf{H}}\hat{\mathbf{H}}^H$ with $M_R = 4$ and $M_T = 2, 3, 4$ is presented in Figure 1. For the sake of comparison, the corresponding exponential distribution approximations are also plotted in the same figure. It can be seen that larger value of $M_R - M_T$ leads to larger difference between the exact distribution and its exponential distribution approximation. Fortunately, the difference is not significant. Simulation results in Section 5 verify that the exponential distribution approximation, when applied with our proposed power allocation, can retain significant performance gain. Therefore, the effectiveness of our proposed power allocation scheme is still valid and evident in the case of $M_R \neq M_T$.

5. Simulation Results

In our simulations, we adopt the 1/2 rate binary convolutional code with polynomials (133,171) in octal notation. M-PSK and 16-QAM modulations with Gray mapping are used in the simulations, and the simulations are performed for the MIMO systems with $M_R = 4$ and $M_T = 2, 3, 4$. For block-fading channel, the length of each frame is set to be 1000 MIMO vector symbols, that is, the channel fading is kept constant during $(N_D + N_P) = 1000$ MIMO vector symbol intervals and varied randomly to another channel realization for every subsequent 1000 MIMO vector symbol intervals. Each channel realization is generated as a random matrix with complex Gaussian entries of $\mathcal{CN}(0, 1)$. For all antenna configurations of MIMO systems, we set $N_P = 4$. Furthermore, each channel codeword is transmitted only during one channel fading block. Therefore, the channel codeword length is 15936 bits for 16-QAM, 7968 bits for QPSK, and 3984 bits for BPSK. SNR is defined as SNR $= P_t/((N_D+N_P)\sigma_n^2)$. In the following figures, the curves marked by "EPA" and "PPA" are the performance results obtained by "equal power allocation" and "proposed power allocation," respectively.

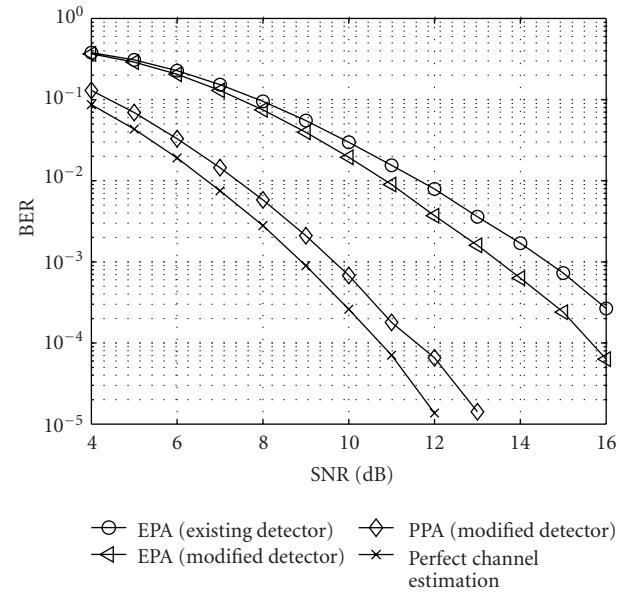

Figure 2: BER performance of different power allocation schemes (QPSK).

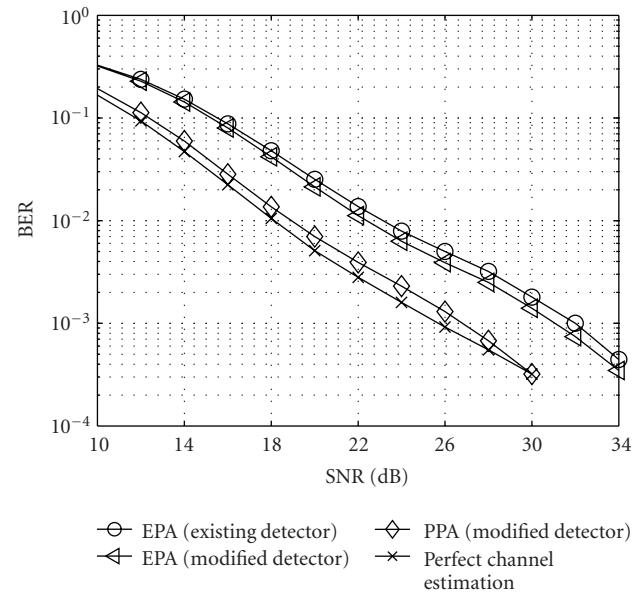

Figure 3: BER performance of different power allocation schemes (16-QAM).

5.1. BER Performance under Block-Fading Channel. Figures 2 and 3 show the BER performance comparison among our proposed soft-output MMSE MIMO detector with the new power allocation scheme based on (23), our proposed soft-output MMSE MIMO detector with equal power allocation,

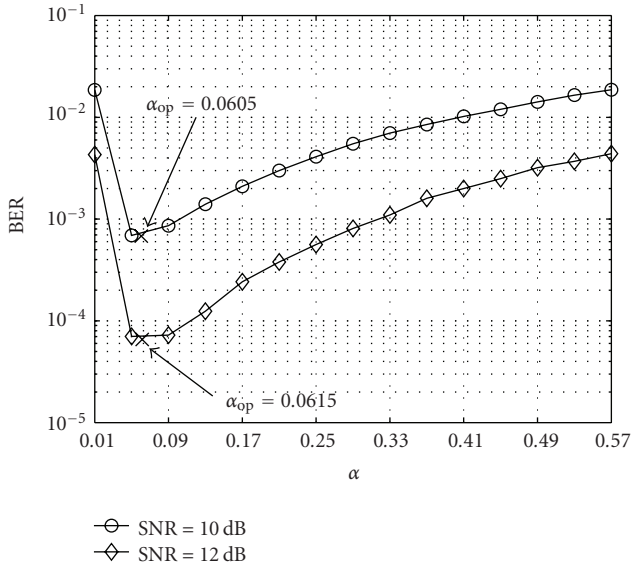

Figure 4: BER performance under different power allocation ratio (QPSK).

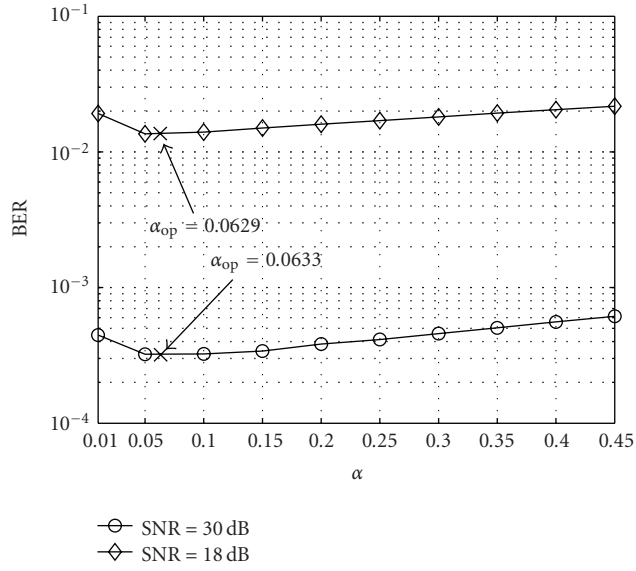

Figure 5: BER performance under different power allocation ratio (16-QAM).

and the existing MMSE MIMO detector with equal power allocation. Here, the existing MMSE MIMO detector refers to the MMSE detector derived under the perfect channel estimation. For equal power allocation, the same transmission power is assigned to pilot and data symbols. It can be seen that our proposed soft-output MMSE MIMO detector with the new power allocation scheme outperforms the existing MMSE MIMO detector by about 4 dB at BER = 10^{-3}. When equal power allocation is applied, it can be seen that our proposed soft-output MMSE MIMO detector also outperforms the existing MMSE MIMO detector with a gain of 1 dB for QPSK and 0.5 dB for 16-QAM at BER = 10^{-3}. As higher SNR is required for 16-QAM to achieve the same BER performance as QPSK, which leads to relatively smaller channel estimation error, the performance gain of our proposed detector becomes smaller for 16-QAM than that of QPSK. For the sake of comparison, we also present the performance under perfect channel estimation in these figures. For perfect channel estimation, power is equally allocated to data symbols. It can be seen that the performance of our proposed detector with the new power allocation is very close to that of perfect channel estimation. For the reason described previously, the performance gap is smaller for 16-QAM than that of QPSK.

For QPSK and 16-QAM, Figures 4 and 5 show the BER performance of our proposed soft-output MMSE MIMO detector under different power allocation ratio α for specific SNR value, respectively. The obtained power allocation ratios α_{op} based on our proposed power allocation scheme are also plotted in these figures. As expected, our proposed power allocation scheme approaches optimal allocation ratio.

Figures 6 and 7 present the BER performance comparison for the case of $M_R = 4$, $M_T = 2$, and $M_T = 3$, respectively. From these two figures, we can conclude that our proposed detector with the new power allocation is also superior over equal power allocation with significant performance gain for the case of $M_R \neq M_T$. Meanwhile, our proposed detector with the new power allocation can achieve nearly the same performance as perfect channel estimation. So, it can be concluded that the proposed power allocation is also applicable for the MIMO systems of $M_R \neq M_T$.

5.2. Throughput Performance. Figure 8 presents the throughput comparison for the MIMO system with $M_R = M_T = 4$. The normalized throughput is defined as

$$C_N = \begin{cases} \dfrac{rN_D M_T \log_2(|\mathcal{A}|)}{T_f B} \times \dfrac{N_D}{N_D + N_P}, & \text{if BER} \le \text{BER}_t, \\ 0, & \text{if BER} > \text{BER}_t, \end{cases} \tag{40}$$

where r denotes the code rate of the applied channel coder, T_f is the duration of a frame B is the system bandwidth BER_t stands for the target BER performance and $|\mathcal{A}|$ denotes the size of modulation constellation $\mathcal{A}$. The item $N_D/(N_D + N_P)$ accounts for the overhead of pilot symbols. The normalized throughput is obtained by setting the modulation constellation as BPSK, QPSK, and 16-QAM with target BER, $\text{BER}_t = 10^{-3}$. We simulate the BER performance of these three modulation modes and always choose the modulation mode with the highest constellation size, which meets the target BER performance, in the computation of the normalized throughput. We find that our proposed detector with the new power allocation obtains significant throughput gain over equal power allocation.

5.3. Performance under Time-Varying Channels. Figure 9 shows the BER performance comparison between different power allocation schemes under time-varying channels.

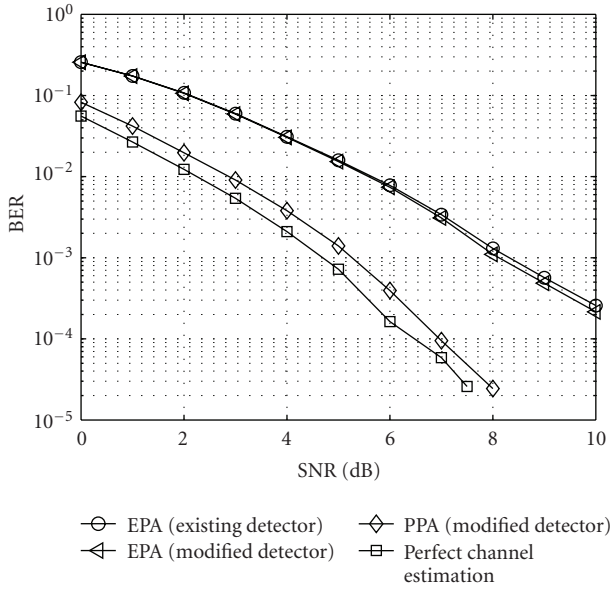

Figure 6: BER performance of different power allocation schemes (QPSK, $M_R = 4$, $M_T = 2$).

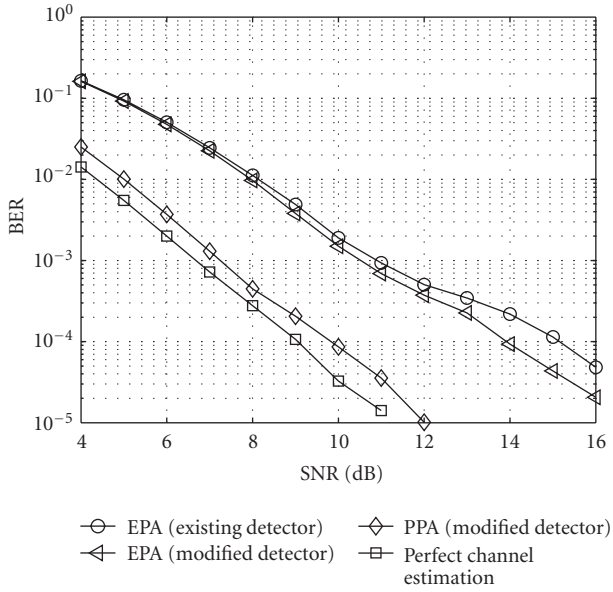

Figure 7: BER performance of different power allocation schemes (QPSK, $M_R = 4$, $M_T = 3$).

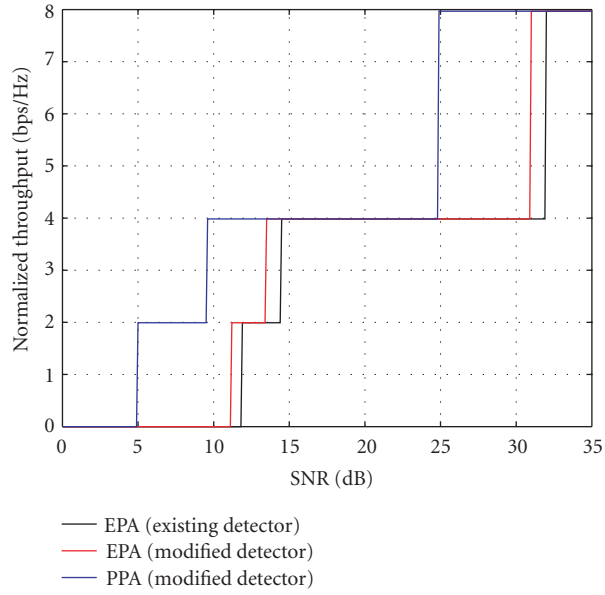

Figure 8: Normalized throughput of different power allocation schemes ($M_R = 4$, $M_T = 4$).

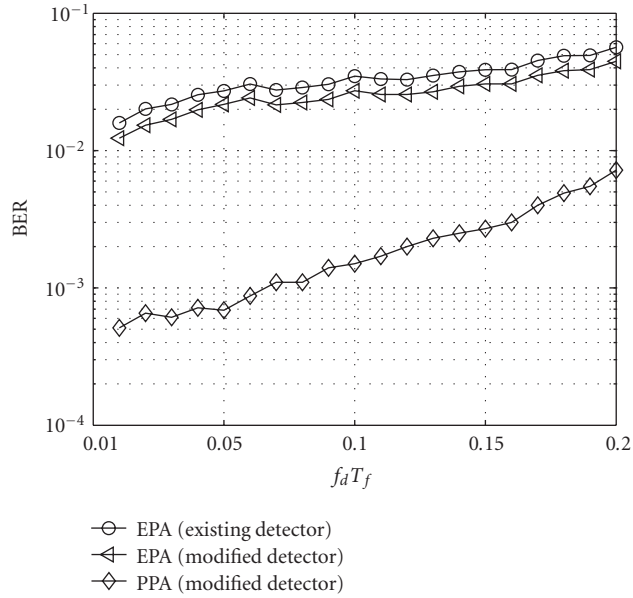

Figure 9: BER performance of different power allocation schemes under time-varying channels (QPSK, SNR = 10 dB, $M_R = 4$, $M_T = 4$).

The time-varying characteristic is measured by normalized Doppler frequency shift $f_D T_f$, where f_D denotes the maximal Doppler frequency shift and T_f is the duration of a frame [21]. Basically, the larger $f_D T_f$ is the faster the channel fading changes. In the simulation, we allocate the pilot symbols at the middle of a frame in order to alleviate the adverse effects of the time-varying characteristics of the channel. It can be seen that the BER performance degrades significantly for all detectors and power allocation schemes as $f_D T_f$ increases. The reason is that fast-fading channel will incur an increased channel estimation error for data vector symbols, which have relatively larger distance from the pilot symbols. As our new power allocation is more dependent on the estimated channel, the rate of the performance degradation is larger than that of equal power allocation. However, we find that the performance of our proposed power allocation is still better than that of equal power allocation in the region of desired BER. Finally, it is worthy of noting that the channel estimation can be improved by inserting more pilot symbols and tracking the time-varying channel fading. However, this topic is out of the scope of this paper.

6. Conclusion

In this paper, we have derived a novel soft-output MMSE MIMO detector under MMSE channel estimation by taking the channel estimation error into consideration. Based on this new detector, we then proposed an effective power allocation scheme between pilot and data symbols by maximizing the average lower bound of postprocessing SINR when $M_T = M_R$. This power allocation scheme has negligible complexity and does not need any form of channel coefficients feedback. We also prove the existence and uniqueness of our optimal power allocation. Furthermore, we point out that the proposed power allocation scheme can also be applied to the MIMO systems of $M_T \neq M_R$ through proper approximation. Compared with existing detector with equal power allocation, simulation results show that the new detector with the proposed power allocation scheme can obtain significant performance gain in terms of BER and throughput under block-fading and time-varying fading channels.

Acknowledgments

This work was supported in part by the Fundamental Research Funds for the Central Universities, the National Natural Science Foundation of China (NSFC) under Grant no. 61071102, the High-Tech Research and Development Program (863) of China under Grants nos. 2009AA011801, 2009AA012002, the National Basic Research Program (973) of China under Grant no. 2009CB320405, in part by National Key Technology R&D Program under Grant no, 2008BAH30B09, the National Fundamental Research Program of China under Grant nos. A1420080150, and the Major National Science and Technology Special Project of China under Grant no. 2008ZX03005-001, 2009ZX03007-004, 2009ZX03005-002, and 2009ZX03005-004. Part of this work was presented at IEEE PIMRC'10 [22].

References

[1] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884–893, 2006.

[2] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.

[3] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, 2006.

[4] G. Taricco and E. Biglieri, "Space-time decoding with imperfect channel estimation," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1874–1888, 2005.

[5] K. Lee and J. Chun, "Symbol detection in V-BLAST architectures under channel estimation errors," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 593–597, 2007.

[6] D. Seethaler, G. Matz, and F. Hlawatsch, "An efficient MMSE-based demodulator for MIMO bit-interleaved coded modulation," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 4, pp. 2455–2459, Dallas, Texas, USA, 2004.

[7] M. M. Wang, W. Xiao, and T. Brown, "Soft decision metric generation for QAM with channel estimation error," *IEEE Transactions on Communications*, vol. 50, no. 7, pp. 1058–1061, 2002.

[8] E. Baccarelli and M. Biagi, "Power-allocation policy and optimized design of multiple-antenna systems with imperfect channel estimation," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 136–145, 2004.

[9] S. Serbetli and A. Yener, "MMSE transmitter design for correlated MIMO systems with imperfect channel estimates: power allocation trade-offs," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2295–2304, 2006.

[10] S. M. Kay, *Fundaments of Statistical Signal Processing: Estimation Theory*, Prentice Hall, Englewood Cliffs, NJ, USA, 1993.

[11] X. Wang and H. V. Poor, *Wireless Communications: Advanced Techniques for Signal Reception*, Prentice Hall, New York, NY, USA, 2003.

[12] S. H. M. Weinfurtner, "Coding approaches for multiple antenna transmission in fast fading and OFDM," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2442–2450, 2002.

[13] N. Kim, Y. Lee, and H. Park, "Performance analysis of MIMO system with linear MMSE receiver," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, Article ID 4686826, pp. 4474–4478, 2008.

[14] N. Kim and H. Park, "Bit error performance of convolutional coded MIMO system with linear MMSE receiver," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3420–3424, 2009.

[15] P. Li, D. Paul, R. Narasimhan, and J. Cioffi, "On the distribution of SINR for the MMSE MIMO receiver and performance analysis," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 271–286, 2006.

[16] R. Narasimhan, "Spatial multiplexing with transmit antenna and constellation selection for correlated MIMO fading channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2829–2838, 2003.

[17] J. H. Kotecha and A. M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 546–557, 2004.

[18] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Eigenvalues and condition numbers of complex random matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 2, pp. 441–456, 2005.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.

[20] H. Zhang, F. Niu, H. Yang, and D. Yang, "Polynomial PDF of the smallest eigenvalue of complex central Wishart matrix with correlation at the side with the smallest number of antennas," *IEEE Communications Letters*, vol. 12, no. 10, pp. 749–751, 2008.

[21] G. L. Süber, *Principles of Mobile Communications*, Kluwer Academic, Norwell, Mass, USA, 2nd edition, 2001.

[22] J. Wang, H. Chen, and S. Li, "An effective power allocation scheme for MMSE MIMO channel estimation with MMSE detection," in *Proceedings of the IEEE PIMRC'10*, Istanbul, Turkey, 2010.