

# OPTIMAL PILOT SPACING AND POWER IN RATE-ADAPTIVE MIMO DIVERSITY SYSTEMS WITH IMPERFECT TRANSMITTER CSI

Duc V. Duong, Bengt Holter, and Geir E. Øien

Norwegian University of Science and Technology, Dept. of Electronics and Telecommunications  
O.S. Bragstads pl. 2B, N-7491 Trondheim, Norway. Email: {duong,bholter,oien}@iet.ntnu.no

## ABSTRACT

We analyze and optimize the performance of an adaptive coded modulation system with imperfect channel state information, operating on a multiple-input multiple-output channel where space-time block coding is utilized to exploit spatial diversity and, thereby combat fading in wireless channel communications. The pilot symbol period and the power allocated to pilot and data symbols are optimized in such a way that the *average spectral efficiency* is maximized. At the same time, the *bit error rate* is maintained below a desired level. A numerical example is given for independent flat Rayleigh fading subchannels with Jakes spectrum.

## 1. INTRODUCTION

Optimization of the pilot period and the power allocation between pilot and data symbols has been shown to be crucial for increasing the average spectral efficiency (ASE) of adaptive (coded) modulation (ACM) scheme, under a bit error rate (BER) constraint, both on single-input single-output channels [1], [2] and on single-input multiple-output (SIMO) channels [3]. In this paper we investigate the effect of such optimization when the system has multiple antennas to transmit from in addition to the multiple reception of data signals, i.e., we have a multiple-input multiple-output (MIMO) system. Space-time block coding (STBC) is used to exploit the available spatial diversity of the channel. Note that there is an inherent trade-off between the use of diversity and the need for adaptivity. The higher the diversity order, the more stable is the channel and the less need there is for link adaptation. We therefore restricted ourselves to moderate diversity orders to demonstrate the power of link adaptation in this context. The work is motivated by previous work by Holter *et al.* [4], which analyzed ACM in an MIMO diversity context, but without the kind of optimization mentioned above. It is also an extension of the papers by Duong *et al.* [1], [3], and Cai and Giannakis [2].

Throughout the paper we use the superscripts  $(\cdot)^T$ ,  $(\cdot)^*$ , and  $(\cdot)^H$  to denote transpose, complex conjugate, and Hermitian transpose, respectively.  $\mathcal{D}(\mathbf{z})$  is the diagonal matrix with vector  $\mathbf{z}$  on its diagonal.  $\mathbf{I}_K$  is the  $K \times K$  identity matrix,  $[\mathbf{A}]_{mn}$  denotes the  $m$ th element in the matrix  $\mathbf{A}$ ,  $\lfloor z \rfloor$  means the integer part of  $z$ , and  $E[z]$  denotes the expectation of  $z$ .

The remainder of the paper is organized as follows: In Section 2, the investigated system is described. The channel estimation and prediction is presented in Section 3. Section 4 analyzes the BER performance in the presence of imperfect channel estimation and prediction. The optimization of the ASE is dealt with

Orthogonal design	$n_T$	$R_s$	$K$	$T$
$G_1$ —No STBC	1	1	1	1
$G_2$ —Alamouti [6]	2	1	2	2
$H_4$ —Matrix (40) in [7]	4	3/4	3	4

Tab. 1. Orthogonal designs for STBC

in Section 5. A numerical example is given in Section 6, and the conclusions are given in Section 7.

## 2. SYSTEM MODEL

We analyze the system illustrated in Fig. 1. We have  $n_T$  transmit and  $n_R$  receive antennas, and all the subchannels are assumed to be mutually independent. The adaptive encoder always chooses to transmit symbols from the symbol constellation of size  $M_n$  which is best suited to the predicted channel, out of a set of constellations of sizes  $\{M_n\}_{n=1}^N$ —corresponding to a set of *spectral efficiencies* (SEs)  $\{R_n\}_{n=1}^N$ —where  $N$  is the number of constellations. The SEs are ordered such that  $R_1 < R_2 < \dots < R_N$ , and we also have  $M_1 < M_2 < \dots < M_N$ . The choice of  $n$  is based on the channel state information (CSI) fed back from the receiver:  $M_n$  is chosen if the predicted *channel-signal-to-noise ratio* (CSNR) falls between the *switching thresholds*  $\hat{\gamma}_n$  and  $\hat{\gamma}_{n+1}$ . By letting  $\hat{\gamma}_0 = 0$  and  $\hat{\gamma}_{N+1} = \infty$ , we have  $\hat{\gamma}_n < \hat{\gamma}_{n+1}$  for all  $n \in \{0, 1, \dots, N\}$ . The power allocation between data and pilot symbols will be chosen in an optimal way. The criterion for choosing the code and the transmit power distribution is to maximize the ASE while fulfilling a target BER,  $\text{BER}_0$ .

The channel estimation and prediction is performed using a pilot-symbol-assisted modulation (PSAM) scheme [5], meaning that known symbols—pilot symbols—must be transmitted along with the data symbols. When performing channel estimation and prediction using PSAM in a MIMO diversity system with  $n_T$  transmit antennas, the number of pilot symbols to be transmitted is  $n_T$  times the number of pilot symbols that are needed in the SIMO case [6]. As in [4], we assume that the datastream prior to the STBC is divided into frames of length  $L = mK + 1$ , where  $K$  is the number of data symbols to be used by the selected space-time block code, and  $m$  is a non-zero positive integer. Note that pilot symbol is not space-time block coded. Thus, when a pilot symbol enters the space-time encoder, it is cyclically shifted and transmitted once from each antenna. While a pilot symbol is transmitted from one antenna, the other antennas are silent such that each receiver branch can estimate/predict the channel between itself and the transmitting antenna.

Furthermore, the space-time encoder maps  $K$  QAM symbols

This work is supported by the Research Council of Norway under the project BEATS (URL: <http://www.tele.ntnu.no/projects/beats/>).

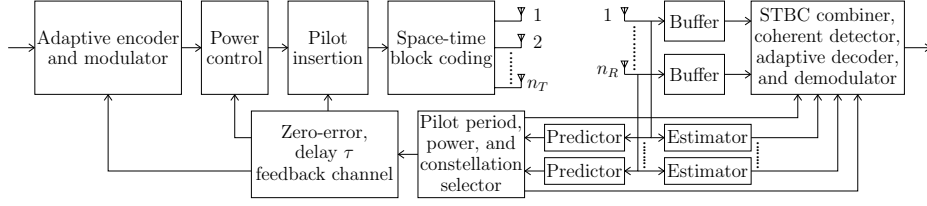


Fig. 1. The adaptive PSAM system model operating on a MIMO channel.

Antenna no. 1	$P$	$\circ$	$\circ$	$\circ$	$D_1$	$-D_2^*$	$\frac{D_3^*}{\sqrt{2}}$	$\frac{D_4^*}{\sqrt{2}}$
Antenna no. 2	$\circ$	$P$	$\circ$	$\circ$	$D_2$	$D_1^*$	$\frac{D_3^*}{\sqrt{2}}$	$-\frac{D_4^*}{\sqrt{2}}$
Antenna no. 3	$\circ$	$\circ$	$P$	$\circ$	$\frac{D_3}{\sqrt{2}}$	$\frac{D_3}{\sqrt{2}}$	$\dots$	$\dots$
Antenna no. 4	$\circ$	$\circ$	$\circ$	$P$	$\frac{D_3}{\sqrt{2}}$	$-\frac{D_3}{\sqrt{2}}$	$\dots$	$\dots$

Fig. 2. Example of a frame structure after STBC where  $n_T = 4$ , and the orthogonal design  $H_4$  in Tab. 1 is used.  $P$  stands for pilot and  $\circ$  denotes that the system does not send anything, while  $D_s$  are data symbols. To reduce the size of the figure we avoid to write out the four last data symbols, and demonstrate only the smallest frame size.

into  $n_T$  orthogonal sequences of length  $T$ —defined as  $(K/R_s)T_s$ , where  $T_s$  is the channel symbol interval and  $R_s$  is the rate of the employed space-time block code. The orthogonal space-time block codes used in this paper are listed in Tab. 1. These orthogonal designs and some other orthogonal designs can be found in [7]. An example of the frame structure is given in Fig. 2.

Let  $\mathbb{C}$  be the set of complex numbers and let  $\mathbf{y}_d \in \mathbb{C}^{n_R \times 1}$  be the received, noisy, and faded data symbol vector in complex baseband. It can be written as

$$\mathbf{y}_d(k; l) = \sqrt{\frac{\mathcal{E}_d}{n_T}} \mathbf{H}(k; l) \mathbf{s}(k; l) + \mathbf{n}(k; l), \quad l \in [n_T, \dots, L_b - 1], \quad (1)$$

where  $L_b = mK/R_s + n_T = (L-1)/R_s + n_T$  is the pilot symbol spacing on a single antenna branch after STBC [4]. Furthermore, let  $\mathbf{y}_{pl} \in \mathbb{C}^{n_R \times 1}$  be the received pilot symbol vector:

$$\mathbf{y}_{pl}(k; l) = \sqrt{\frac{\mathcal{E}_{pl}}{n_T}} \mathbf{H}(k; l) \mathbf{s}(k; l) + \mathbf{n}(k; l), \quad l \in [0, \dots, n_T - 1]. \quad (2)$$

In both (1) and (2),  $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$  is the channel matrix,  $\mathbf{s} \in \mathbb{C}^{n_T \times 1}$  is the vector of transmitted symbols, and  $\mathbf{n} \in \mathbb{C}^{n_R \times 1}$  is the channel AWGN. For simplicity we only consider the case when the power is equally allocated to different transmit antennas, thus  $\mathcal{E}_d$  and  $\mathcal{E}_{pl}$  is the total transmit power per pilot and per data symbol, respectively; these are to be optimized later. The index  $k$  counts frames, while  $l$  is the symbol index in that frame. We assume that  $E[|s_i|^2] = 1$ , where  $i \in [1, \dots, n_T]$ , and the pilot symbol has unity magnitude. Furthermore, elements in  $\mathbf{n}(k; l)$  is zero mean complex Gaussian distributed with variance  $N_0/2$  per dimension and dimensions being uncorrelated. The fading envelope of the  $[\mathbf{H}]_{ji}$ th branch is assumed to be a stationary complex Gaussian random process with zero mean and variance  $\sigma_{h_{ji}}^2 = 1$ .

Using the approach in [2] and [3], the average power per data symbol and per pilot symbol can be found to be  $\bar{\mathcal{E}}_d = \alpha L \mathcal{E} / (L-1)$  and  $\bar{\mathcal{E}}_{pl} = (1-\alpha)L\mathcal{E}$ , respectively, where  $\mathcal{E}$  is the average

transmit power for both pilot and data symbols. Equal data and pilot power is obtained if  $\alpha = 1 - 1/L$ .

### 3. CHANNEL ESTIMATION AND PREDICTION

Each subchannel is assumed to be slowly varying so that its fading remains constant over many channel symbols, and we perform estimation and prediction independently on each subchannel. Also, the correlation function is assumed known.

Both the estimator and predictor are restricted to be linear, and are made optimal in the *maximum a posteriori* (MAP) sense [8].

#### 3.1. Channel Estimation

Based on  $K_e$  received pilot symbols, the non-causal MAP-optimal estimator can estimate each fading channel in  $\{h_{ij}(k; l)\}_{l=n_T}^{L_b-1}$  as

$$h_{e;ij}(k; l) = \mathbf{w}_e^H \mathbf{y}_{ij}(k; i-1), \quad i \in [1, \dots, n_T], \quad j \in [1, \dots, n_R], \quad (3)$$

where  $\mathbf{y}_{ij}(k; i-1) = [y_{pl;ij}(k - \lfloor K_e/2 \rfloor; i-1), \dots, y_{pl;ij}(k + \lfloor (K_e-1)/2 \rfloor; i-1)]^T$ —corresponding to the vector of received pilot symbols transmitted from the  $i$ th antenna to the  $j$ th antenna—and  $\mathbf{w}_e$  is the estimator given by  $\mathbf{w}_e = \sqrt{\mathcal{E}_{pl}/n_T} (\mathcal{E}_{pl}/n_T (\mathcal{D}(\mathbf{s}) \mathbf{R}_e \mathcal{D}^*(\mathbf{s}) + N_0 \mathbf{I}_{K_e}))^{-1} \mathcal{D}(\mathbf{s}) \mathbf{r}_e$  [2]. The mean square estimation error (MSE) is given by  $\sigma_{e;ij}^2(l) = E[|h_{ij} - h_{e;ij}|^2]$ . With the chosen estimator we achieve the minimum MSE (MMSE) of any subchannel as

$$\sigma_{e;ij}^2(l) = 1 - \sum_{\kappa=1}^{K_e} \frac{|\mathbf{u}_\kappa^H \mathbf{r}_e|^2 (1-\alpha) L \bar{\gamma}_{ij}}{(1-\alpha) L \bar{\gamma}_{ij} \lambda_\kappa + 1}, \quad (4)$$

in which  $\bar{\gamma}_{ij} = \mathcal{E}/(n_T N_0)$  is the expected CSNR on any subchannel,  $\{\mathbf{u}_\kappa\}$  denotes the eigenvectors of the covariance matrix  $\mathbf{R}_e = E[\mathbf{h}\mathbf{h}^H]$  where  $\mathbf{h} = [h_{ij}(k - \lfloor K_e/2 \rfloor; i-1), \dots, h_{ij}(k + \lfloor (K_e-1)/2 \rfloor; i-1)]^T$ ,  $\{\lambda_\kappa\}$  are the corresponding eigenvalues, and  $\mathbf{r}_e$  is the covariance vector;  $\mathbf{r}_e = E[\mathbf{h}h_{ij}^*(k; l)]$ . We can see that  $\sigma_{e;ij}^2(l)$  only depends on the location of the symbol to be estimated  $l$ , and not on the frame index  $k$ .

Due to the orthogonality principle, the estimation error  $\epsilon_{e;ij} = h_{ij} - h_{e;ij}$ , and the estimated channel  $h_{e;ij}$  are uncorrelated.

#### 3.2. Channel Prediction

For notational simplicity, we restrict the system feedback delay<sup>1</sup> to be  $\tau = DL_b T_s$ , where  $D$  is a positive integer. The predictor uses

<sup>1</sup>The feedback delay here includes the time it takes to perform prediction, actual transmission delay on the feedback channel, and the processing time needed by the transmitter to activate the code to be transmitted.

$K_p$  pilot symbols from the past to predict one sample in the set  $\{h_{ij}(k; l)\}_{l=n_T}^{L_b-1}$ . Defining the channel gain vector  $\mathbf{h} = [h_{ij}(k-D; i-1), \dots, h_{ij}(k-D-K_p+1; i-1)]^T$ —corresponding to the pilot instances—the covariance matrix is  $\mathbf{R}_p = E[\mathbf{h}\mathbf{h}^H]$ , and the covariance vector is  $\mathbf{r}_p = E[\mathbf{h}h_{ij}^*(k; l)]$ . Similar to the estimation case, the predicted channel is given as  $h_{p;ij}(k; l) = \mathbf{w}_p^H \mathbf{y}_{ij}(k; i-1)$ , where  $\mathbf{y}_{ij}(k, i-1) = [y_{pl;ij}(k-D; i-1), \dots, y_{pl;ij}(k-D-K_p+1; i-1)]^T$ , and  $\mathbf{w}_p = \sqrt{\mathcal{E}_{pl}/n_T} (\mathcal{E}_{pl}/n_T (\mathcal{D}(\mathbf{s})\mathbf{R}_p\mathcal{D}^*(\mathbf{s}) + N_0\mathbf{I}_{K_p}))^{-1} \mathcal{D}(\mathbf{s})\mathbf{r}_p$ . As a result, similarly to (4), the MMSE of the prediction error is

$$\sigma_{p;ij}^2(l) = 1 - \sum_{\kappa=1}^{K_p} \frac{|\mathbf{u}_\kappa^H \mathbf{r}_e|^2 (1-\alpha)L\bar{\gamma}_{ij}}{(1-\alpha)L\bar{\gamma}_{ij}\lambda_\kappa + 1}, \quad (5)$$

where  $\{\mathbf{u}_\kappa\}$  and  $\{\lambda_\kappa\}$  is now the set of eigenvectors and the corresponding eigenvalues of the covariance matrix  $\mathbf{R}_p$ , respectively.

#### 4. BER PERFORMANCE ANALYSIS

The combined data signal is found by  $\mathbf{z}(k; l) = \mathbf{H}_e^H \mathbf{y}_d(k; l)$ . Using results from [2], [3] the total received CSNR can be written

$$\gamma(k; l) = \frac{\mathcal{E}_d \|\mathbf{H}_e\|_F^2}{n_T(N_0 + g\mathcal{E}_d\sigma_e^2(l))} \quad (6)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm. To obtain (6), we assume that the total estimation error variance at the  $j$ th receive branch is  $n_T$  times the estimation error variance from any transmit antenna to that branch. Furthermore, the total estimation error is the same for all branches  $j \in [1, \dots, n_R]$ . The constant  $g = 1$  for 4-QAM and  $g = 1.3$  for higher QAM constellations [2].

Similarly to in [3] and [4], we will utilize a set of 4-dimensional trellis codes as our component codes when giving numerical examples in Section 6. Tight approximations of BER performance of these codes on AWGN channels can be found in [9]. However, in order to obtain a closed-form and mathematically tractable solution we here use a somewhat looser BER approximation. It has been demonstrated—by comparing the results in [1] and [3]—that the impact of this on our performance analysis is negligible in practice. By using the BER approximations in [3, Eq. 8]—with (6) inserted—the BER is given as

$$\text{BER}(M_n | \{h_{e;ij}\}) = \sum_{\ell=1}^{\mathcal{L}} a_n(\ell) \exp(-A_n \mathcal{E}_d \|\mathbf{H}_e\|_F^2), \quad (7)$$

where  $A_n = b_n(\ell)/(n_T M_n (N_0 + g\mathcal{E}_d\sigma_e^2(l)))$  and  $\mathcal{L}$  is the number of exponential functions which approximates the simulated BER. The constants  $a_n(\ell)$ ,  $b_n(\ell)$  are code dependent constants and can be found in [3, Tab. I].

We may express the estimate of any subchannel as

$$h_{e;ij}(k; l) = h_{p;ij}(k; l) + \epsilon_{p;ij}(k; l) - \epsilon_{e;ij}(k; l). \quad (8)$$

With the assumption that the subchannels are independent we now use the results in [3] for the SIMO case, and extend them further to

the MIMO case. Given  $\{h_{p;ij}\}$ , the overall pdf of  $\{|h_{e;ij}|\} | \{h_{p;ij}\}$  is  $p(\{|h_{e;ij}|\} | \{h_{p;ij}\}) = \prod_{i=1}^{n_T} \prod_{j=1}^{n_R} p(|h_{e;ij}| | h_{p;ij})$ . Each subchannel is Ricean distributed with the Rice factor  $K = |(1-\rho)h_{p;ij}(k; l)|^2 / \tilde{\sigma}_{h_{e;ij}}^2$ , where  $\rho$  is the normalized correlation between  $\epsilon_{e;ij}(k; l)$  and  $h_{p;ij}(k; l)$ . It is shown in [2] that  $\rho$  typically takes on very small values, so we choose to set it equal to zero in our analysis. On the other hand, in contrary to the same paper, the estimation error  $\epsilon_{e;ij}(k; l)$  is *correlated* with the prediction error  $\epsilon_{p;ij}(k; l)$ . Hence, as a result,  $\tilde{\sigma}_{h_{e;ij}}^2 = \sigma_{p;ij}^2 - \sigma_{e;ij}^2$  and  $K = |h_{p;ij}(k; l)|^2 / \tilde{\sigma}_{h_{e;ij}}^2$  [3].

BER conditioned on the predicted channels is found by averaging (7) over the overall pdf found above:

$$\begin{aligned} \text{BER}(M_n | \{h_{p;ij}\}) &= \underbrace{\int_0^\infty \dots \int_0^\infty}_{n_T n_R \text{-fold}} \text{BER}(M_n | \{|h_{e;ij}|\}) \\ &\times p(\{|h_{e;ij}|\} | \{h_{p;ij}\}) d|h_{e;11}| \dots d|h_{e;n_T n_R}|. \quad (9) \end{aligned}$$

After some straightforward integrations, we obtain

$$\text{BER}(M_n | \{h_{p;ij}\}) = \sum_{\ell=1}^{\mathcal{L}} a_n(\ell) d_n^{n_T n_R} \exp(-A_n d_n \mathcal{E}_d \|\mathbf{H}_p\|_F^2), \quad (10)$$

where  $\tilde{\sigma}_{h_{e;ij}}^2 = \tilde{\sigma}_{h_e}^2 \forall i, j$  has been assumed due to the independence of the subchannels, and  $d_n \triangleq 1/(A_n \mathcal{E}_d \tilde{\sigma}_{h_e}^2 + 1)$ .<sup>2</sup>

As in [10], we define the predicted CSNR per symbol as

$$\hat{\gamma} = \frac{\bar{\mathcal{E}}_d \|\mathbf{H}_p\|_F^2}{n_T N_0} = \frac{\bar{\gamma}_j \bar{\mathcal{E}}_d \|\mathbf{H}_p\|_F^2}{n_T \mathcal{E}}, \quad (11)$$

where  $\bar{\gamma}_j = n_T \bar{\gamma}_{ij} = \mathcal{E}/N_0$  is the expected CSNR per symbol of one receive branch. Inserting (11) into (10) we obtain  $\text{BER}(M_n | \hat{\gamma})$  such that the optimal switching thresholds  $\{\hat{\gamma}_n\}$  are found by solving  $\text{BER}(M_n | \hat{\gamma}) = \text{BER}_0$ .

From (11) the average predicted CSNR per symbol is  $\bar{\hat{\gamma}} = r\bar{\gamma}_j n_T n_R / n_T$ —assuming<sup>2</sup>  $\sigma_{p;ij}^2(l) = \sigma_p^2(l) \forall i, j$  and letting  $r = \bar{\mathcal{E}}_d(1 - \sigma_p^2)/\mathcal{E}$  [3]—and it follows the Gamma distribution [4];  $\hat{\gamma} \sim \mathcal{G}(n_T n_R, r\bar{\gamma}_j/n_T) = \frac{1}{\Gamma(n_T n_R)} \left(\frac{n_T}{r\bar{\gamma}_j}\right)^{n_T n_R} \exp\left(-\frac{\hat{\gamma} n_T}{r\bar{\gamma}_j}\right)$ .

The average BER is the ratio between the average number of bits in error, and the number of bits transmitted in total [2]:

$$\text{BER} = \frac{\sum_{n=1}^N \text{BER}(M_n) R_n^{\text{STBC}}}{\sum_{n=1}^N P_n R_n^{\text{STBC}}}, \quad (12)$$

where  $R_n^{\text{STBC}}$  is the SE of the  $n$ th constellation after STBC—to be derived in Section 5.1. In (12),  $\text{BER}(M_n) = \int_{\hat{\gamma}_n}^{\hat{\gamma}_{n+1}} \text{BER}(M_n | \hat{\gamma}) p(\hat{\gamma}) d\hat{\gamma}$  and the result is shown in (13) at the bottom of this page, where  $\bar{\Gamma}(\cdot, \cdot)$  is the normalized incomplete gamma function [3].

<sup>2</sup>Actually we will use the prediction error variance stemming from transmit antenna 1,  $\sigma_{p;1j}^2(l)$ , in our calculation. The reason will be explained in Section 5.

$$\sum_{\ell=1}^{\mathcal{L}} a_n(\ell) \left(\frac{d_n \bar{\mathcal{E}}_d}{r d_n A_n \mathcal{E} \mathcal{E}_d + \bar{\mathcal{E}}_d}\right)^{n_T n_R} \left\{ \bar{\Gamma}\left(n_T n_R, n_T \hat{\gamma}_n \frac{r d_n A_n \mathcal{E} \mathcal{E}_d + \bar{\mathcal{E}}_d}{r \bar{\gamma}_j \bar{\mathcal{E}}_d}\right) - \bar{\Gamma}\left(n_T n_R, n_T \hat{\gamma}_{n+1} \frac{r d_n A_n \mathcal{E} \mathcal{E}_d + \bar{\mathcal{E}}_d}{r \bar{\gamma}_j \bar{\mathcal{E}}_d}\right) \right\} \quad (13)$$

## 5. OPTIMIZATION OF ASE

It is obvious that the variance of the *prediction* error is largest when predicting the last symbol in a frame—i.e.,  $l = L_b - 1$ . As a result, prediction of the symbol located at the end of the frame based on the pilots transmitted from the fourth antenna is slightly more accurate than it would be if based on pilots from the first antenna (cf. Fig. 2). On the other hand, the variance of the *estimation* error is almost the same for all  $l$ , if the order of the estimator is  $K_e \geq 20$  [2]. Thus, we use the estimation error variance  $\sigma_{e;ij}^2(L_b - 1)$ , and the conservative choice of the prediction error variance  $\sigma_{p;1j}^2(L_b - 1)$ —note that the subscript index here is  $i = 1$ —when finding the optimal switching thresholds  $\{\hat{\gamma}_n\}_{n=1}^N$ , as well as in the further optimization process.

### 5.1. ASE Performance Analysis

The system is experiencing an *outage* when the predicted CSNR falls below  $\hat{\gamma}_1$ , since there is no code in our code set which then guarantees the BER performance. In that case the system does not send anything but the pilots—in order to perform the channel estimation and prediction—while data is buffered at the transmitter. Since no transmission is allowed when  $\hat{\gamma} < \hat{\gamma}_1$ , we do not use any data power when this occurs. Therefore the actual transmitted power per data symbol is

$$\mathcal{E}_d = \frac{\bar{\mathcal{E}}_d}{\int_{\hat{\gamma}_1}^{\infty} p(\hat{\gamma}) d\hat{\gamma}} = \frac{\bar{\mathcal{E}}_d}{\bar{\Gamma}(n_T n_R, n_T \hat{\gamma}_1 / r \bar{\gamma}_j)}. \quad (14)$$

The SE of the  $n$ th constellation used by a  $2G$ -dimensional trellis code is  $R_n = (1 - 1/L)(\log_2(M_n) - 1/G)$  [9], where  $G$  is equal to 2 for our example codes. The effective SE becomes

$$R_n^{\text{STBC}} = (\log_2(M_n) - 1/G) \frac{(L - 1)R_s}{L - 1 + n_T R_s} \quad (15)$$

after STBC using the orthogonal designs in Tab. 1. Hence, the overall ASE is given by

$$\text{ASE} = \sum_{n=1}^N R_n^{\text{STBC}} P_n, \quad (16)$$

where  $P_n$  is the probability that  $\hat{\gamma} \in [\hat{\gamma}_n, \hat{\gamma}_{n+1})$ . I.e.,  $P_n = \bar{\Gamma}(n_T n_R, n_T \hat{\gamma}_n / r \bar{\gamma}_j) - \bar{\Gamma}(n_T n_R, n_T \hat{\gamma}_{n+1} / r \bar{\gamma}_j)$ .

When using Nyquist sampling,  $L$  must be less than  $L_{\max} = \lfloor 1/(2f_d T_s) \rfloor$  [5] where  $f_d$  is the maximum Doppler shift. Thus, for  $L \in [2, \dots, L_{\max}]$  we have the following optimization problem:

$$\begin{aligned} & \max_{\alpha, L} \text{ASE} \\ & \text{subject to } 0 < \alpha < 1. \end{aligned} \quad (17)$$

The optimization algorithm used is described in [1], [3]. Thus we choose not to go into details about it here and, instead, refer to these papers.

## 6. NUMERICAL EXAMPLE

At this point, we consider an example ACM system, which has a set of  $N = 8$  QAM signal constellations of sizes  $\{M_n\} = \{4, 8, 16, 32, 64, 128, 256, 512\}$  to switch between. As mentioned in Section 4, those constellations are used to code and decode eight

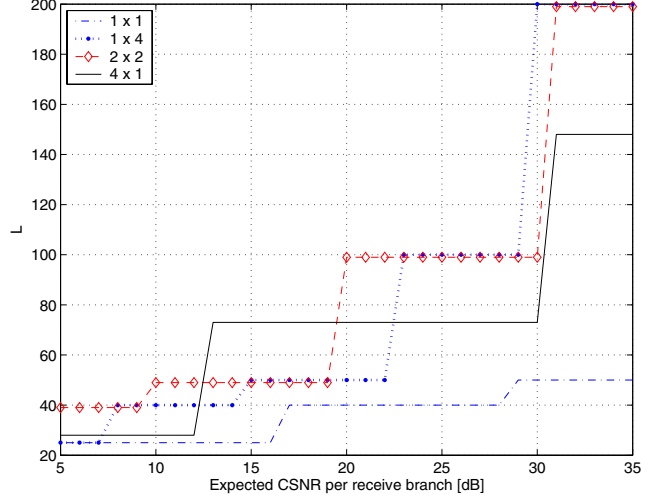


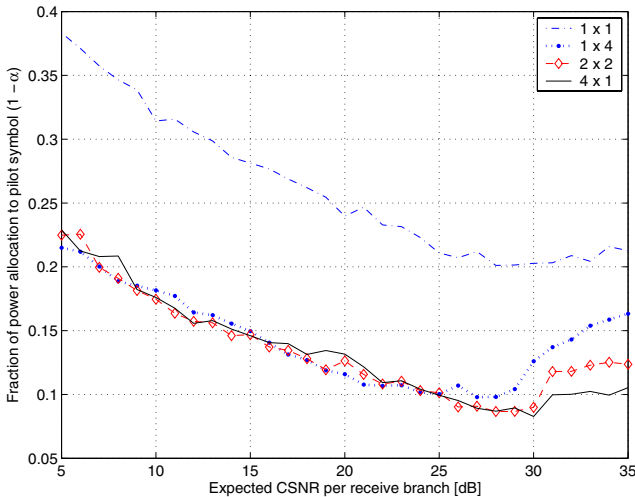
Fig. 3. Optimum pilot spacing  $L$  when the power is optimally allocated between pilot and data symbols. This figure is generated for different combinations of transmit and receive antennas.

4-dimensional trellis codes. The carrier frequency is 2 GHz and the length of a channel symbol is  $5 \mu\text{s}$ —corresponding to a channel bandwidth of 200 kHz using Nyquist sampling. The system delay considered is  $\tau = DL_b T_s = 1 \text{ ms}$  (or  $\tau f_d = 0.2$ ). Contrary to the results given in [2], for this delay,  $L_b$ —and hence  $L$ —can only take on certain discrete values. This is due to the fact that  $D = 200/L_b$  must be an integer. As a result,  $L_b \in \{1, 2, 4, 5, 8, 10, 20, 25, 40, 50, 100, 200\}$ . With the mobile velocity  $v = 30 \text{ m/s}$  and the given carrier frequency, the Doppler frequency is  $f_d = 200 \text{ Hz}$ . We require the system to tolerate a  $\text{BER}_0 = 10^{-5}$ , and we choose the order of the estimator and predictor to be  $K_e = 20$  and  $K_p = 250$ , respectively. Furthermore, we assume that the average subchannel CSNR is the same for all the branches:  $\bar{\gamma}_j = \bar{\gamma}$ .

Fig. 3 shows how the optimal pilot symbol spacing is distributed with the average CSNR on each receive branch. As expected the period increases with increasing CSNR. It increases faster with higher  $n_R$  for a given  $n_T$ . How power is optimally allocated to pilot symbols can be read from Fig. 4, from which the amount of power allocated to data symbols is also easily found. The more antennas there are—either on the transmitter side or on the receiver side, or on both sides—the less power is allocated to pilot symbols.

In order to have a good channel estimation/prediction to maintain the BER when the pilot symbol spacing is steeply increased in the high CSNR regions, however, more power is put on the pilots, which explains why the curves in Fig. 4 increase again at high CSNRs. More or less the same amount of power is allocated to the pilot symbols for average CSNRs up to about 25 dB, when the diversity order is the same.

In [4], the ASE was reduced when having two transmit antennas compared to when only one transmit antenna is equipped. In contradiction to this, Fig. 5 shows that the ASE is substantially increased by going from 1 transmit antenna to 2 transmit antennas, when the pilot spacing and the power are optimal. In general, when using STBC, the channel capacity is reduced, except when either the rate of the employed STBC is one, or the channel is rank



**Fig. 4.** Optimal fraction of power allocation to pilot symbols (i.e.  $1 - \alpha$ ) when the pilot period  $L$  is optimal. The figure is generated for different combinations of transmit and receive antennas.

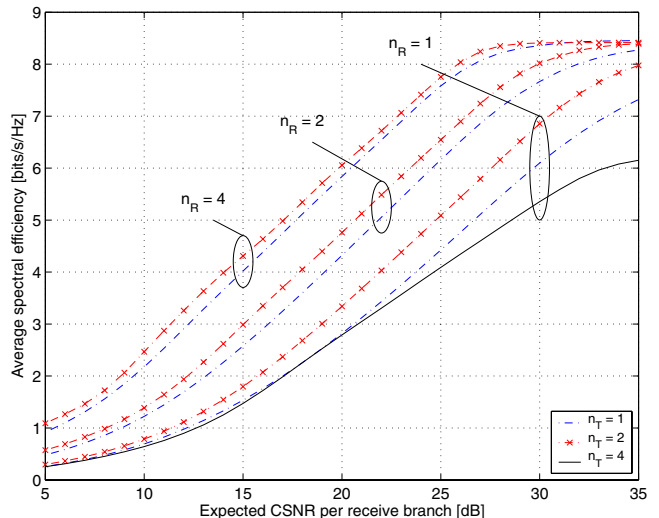
one [10]. In our example, only  $G_1$  and the orthogonal design  $G_2$  has rate one. Optimization of the system with 2 transmit antennas also gives a larger pilot spacing. As a result—in this case—the ASE becomes higher compared to the system with only one transmit antenna. When we have 4 transmit antennas, however, the employed STBC only has rate  $3/4$ . The throughput is significantly reduced due to that loss.

Due to space limitations we do not include the plot for average BER, but we should mention that the average BER is always satisfied the requirement of  $BER_0$ .

For the clarity of the graphs, we avoid to plot for the equal power case. However, it is worth mentioning that, when the power is equally distributed between pilot and data symbols, the pilot spacing  $L$  is smaller, and the fraction of power allocation to the pilots is also smaller (cf. the last paragraph in Section 2). The ASE corresponding to the optimal power and optimal  $L$  always outperforms the equal power case, with up to 0.5 bits/s/Hz, and without any sacrifice of BER.

## 7. CONCLUSIONS

We have analyzed and optimized an ACM system operating on a MIMO diversity channel. The throughput in terms of ASE, when the transmitter is equipped with 2 antennas, outperforms the same system with only one transmit antenna. This is due to the diversity gain and the *rate one* of the employed orthogonal design for STBC. This result is in contrast to what was obtained in [4], where pilot and data power were equal, and the pilot spacing was fixed. Having more than 2 transmit antennas gives even higher diversity order, but the overall rate is reduced. Using our pilot transmission scheme, the more transmit antennas the system has, the more bandwidth is devoted to send pilot symbols. Furthermore, each time when one transmitter sends its pilot, the  $n_T - 1$  other antennas are silent—meaning that the system is losing  $n_T - 1$  symbol intervals, during which data symbols can be transmitted. In addition, when the number of transmit antennas is greater than 2, the rate of the orthogonal design for STBC is less than one; hence,



**Fig. 5.** Average spectral efficiency for different combinations of transmit and receive antennas.

some symbols must be transmitted several times and this degrades the throughput. In conclusion, for a fixed product  $n_T \times n_R$ , the ASE is still always highest for  $n_T = 1$ . BER performance is satisfied (always below the pre-defined value  $10^{-5}$ ) in all cases.

## 8. REFERENCES

- [1] D. V. Duong and G. E. Øien, “Adaptive trellis-coded modulation with imperfect channel state information at the receiver and transmitter,” in *Proc. Nordic Radio Symposium*, Oulu, Finland, Aug 2004.
- [2] X. Cai and G. B. Giannakis, “Adaptive PSAM accounting for channel estimation and prediction errors,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 246–256, Jan 2005.
- [3] D. V. Duong, G. E. Øien, and K. J. Hole, “Adaptive coded modulation with receive antenna diversity and imperfect channel knowledge at receiver and transmitter,” Submitted to *IEEE Trans. on Vehicular Technol. (under revision)*. Also to appear in *Proc. European Signal Processing Conf. (EUSIPCO)*, Antalya, Turkey, September, 2005.
- [4] B. Holter, G. E. Øien, K. J. Hole, and H. Holm, “Limitations in spectral efficiency of a rate adaptive MIMO system utilizing pilot-aided channel prediction,” in *Proc. IEEE Vehicular Technology Conference (VTC)*, Jeju, Korea, April 2003.
- [5] J. K. Cavers, “An analysis of pilot symbol assisted modulation for Rayleigh fading channels,” *IEEE Transactions on Vehicular Technology*, vol. 40, no. 4, pp. 686–693, Nov 1991.
- [6] S. M. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct 1998.
- [7] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, “Space-time block codes from orthogonal designs,” *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.
- [8] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*, John Wiley & Sons, 1998.
- [9] K. J. Hole, H. Holm, and G. E. Øien, “Adaptive multi-dimensional coded modulation over flat fading channels,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 7, pp. 1153–1158, July 2000.
- [10] S. Sandhu and A. Paulraj, “Space-time block codes: A capacity perspective,” *IEEE Communication Letters*, vol. 4, no. 12, pp. 384–386, Dec 2000.