This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the ICC 2008 proceedings.

Asymptotic BER Analysis for MIMO-BICM with Zero-Forcing Detectors Assuming Imperfect CSI

I-Wei Lai*†,Susanne Godtmann*, Tzi-Dar Chiueh†, Gerd Ascheid*, and Heinrich Meyr*

*Institute for Integrated Signal Processing Systems, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany

†Graduate Institute of Electronics Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan

***Abstract*—In this paper, we derive the asymptotic bit error rate (BER) for multiple-input multiple-output bit-interleaved coded modulation (MIMO-BICM) with linear zero-forcing (ZF) receivers for a temporally correlated flat Rayleigh fading channel. Pilot symbol assisted modulation (PSAM) in combination with linear minimum mean-squared error (LMMSE) channel estimation is considered. We also demonstrate that the deterioration due to imperfect channel state information (CSI) can be represented by a signal-to-noise ratio (SNR) degradation.**

I. Introduction

Bit-interleaved coded modulation (BICM), first introduced by Zehavi [1], can achieve large diversity orders in fading channels. Caire et al. [2] have later built the information-theoretical foundation of BICM. The performance of BICM has been evaluated under the assumption of perfect channel state information (CSI) or no CSI. Müller-Weinfurtner [3] has demonstrated that for the multiple-input multiple-output (MIMO) system, BICM shows excellent performance in fast-fading channel when maximum likelihood (ML) detection is used. However, as the complexity of ML detection is a heavy burden for implementation, a low complexity zero-forcing (ZF) detector has been studied in [4]. The results reveal the interesting fact that the performance gap between the ML and ZF detection is significantly reduced when either the number of receive antennas or the modulation order increases.

The capacity and BER performance of MIMO-BICM with ZF detection were thoroughly analyzed by McKay et al. [5]. The asymptotic bit error rate (BER) is analytically derived under the assumption of fast fading and perfect CSI. However, in a realistic environment the channel is unknown at receiver side, and pilot symbol insertion is obligatory to estimate the CSI. This is the so-called pilot symbol assisted modulation (PSAM) and is explained in e.g. [6]. Accurate channel estimation is a major issue in MIMO systems. In order to guarantee the same accuracy as in SISO systems, MIMO systems require M_t-times more pilot symbols, where M_t represents the number of transmit antennas. The trade-off between accurate channel estimation and the achievable data rate in MIMO systems was pointed out in [7].

In this paper, we explore the asymptotic BER performance for a spatially uncorrelated MIMO channel model, where each subchannel is a temporally correlated flat Rayleigh fading channel. Linear minimum mean-squared error (LMMSE) channel estimation that relies on periodically inserted orthogonal pilot symbols [7] is assumed. We base our derivations on the previous work in [5], where the asymptotic BER under the assumption of perfect CSI was calculated. Furthermore, we build on the results in [8] where a SISO-BICM system with imperfect CSI is regarded. Further analyses lead to interpret the estimation error as a signal-to-noise ratio (SNR) degradation, which is labeled as *estimation loss*. This interpretation was first applied in [7] and [9] in the context of the achievable rate.

This paper is organized as follows: in Section II the MIMO transmission model is provided. Section III briefly summarizes the results from [5] on the BER performance for a MIMO-BICM system with ZF detection assuming perfect CSI. The mathematical expressions for the asymptotic BER assuming imperfect CSI and LMMSE channel estimation are derived in Section IV. The approximated SNR degradation due to channel estimation inaccuracies is elaborated in Section V. Section VI validates the analytical results by means of Monte Carlo simulation and the conclusions are given in Section VII.

II. MIMO Transmission Model

The architecture of the system discussed in this paper is presented in Fig. 1 for a 2×2 MIMO system. In general, N information bits are first assembled into a packet and demultiplexed into M_t transmit streams. Each bit stream is then encoded with a binary convolutional code of rate $r = k_c/n$, interleaved (possibly across the streams) and mapped onto an M-ary constellation set χ based on the mapping rule μ. χ_b^i denotes the subset of χ whose labels have the binary value $b \in [0,1]$ at the ith bit position, with $i \in [1, I]$ and $I = \log_2(M)$. The resulting M_t symbols at time instant k are labeled as a *hypersymbol*. Here, we define a *pilot hypersymbol group* $\mathbf{P} \in C^{M_t \times M_t}$ that consists of concatenated orthogonal hypersymbols in order to allow for LMMSE channel estimation [7]. The pilot spacing P_s denotes the spacing between two successive pilot hypersymbol groups. The ratio between orthogonal pilot hypersymbols and data hypersymbols is given to P_s/M_t. Note that the number of pilot hypersymbols per P_s data symbols increases with M_t. Periodic pilot hypersymbols and data hypersymbols are then multiplexed and transmitted over a spatially-uncorrelated channel where each path is a temporally correlated flat Rayleigh fading channel. The maximum normalized Doppler spread with respect to the symbol duration is given by $F_d = f_d T_s$.

Assuming perfect timing synchronization, the received baseband signals of a MIMO system with M_t transmit antennas and M_r receive antenna after matched filtering and sampling

978-1-4244-2075-9/08/$25.00 ©2008 IEEE

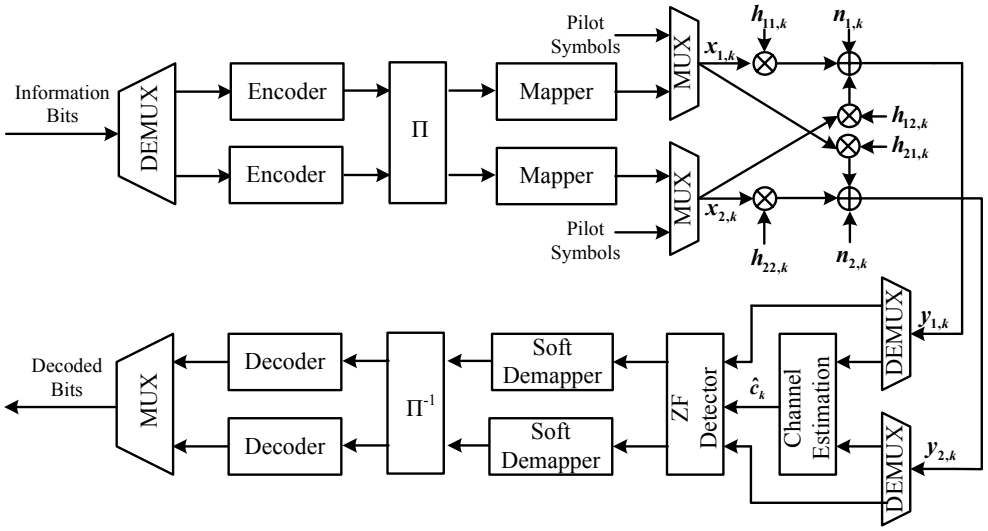


Fig. 1. transmission model of a 2×2 MIMO-BICM with ZF detection

can be modeled as

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

where $\mathbf{H}_k \in C^{M_r \times M_t}$ is the complex MIMO channel matrix with the (i,j)th element $h_{ij,k}$ indicating the path gain from the jth transmit antenna to the ith receive antenna at time instant k. The variance of each channel gain is $\mathrm{E}\{|h_{ij,k}|^2\} = \sigma_h^2$. The complex noise component $\mathbf{n}_k \in C^{M_r \times 1}$ is additive zero-mean white Gaussian noise. The covariance matrix $\mathrm{E}\{\mathbf{n}_k \mathbf{n}_k^H\}$ is an identity matrix $\sigma_n^2 \mathbf{I}_{M_r} = N_0 \mathbf{I}_{M_r}$, where $(\cdot)^H$ denotes conjugate transpose (hermitian). The transmitted hypersymbol $\mathbf{x}_k$ is spatially independent $\mathrm{E}\{\mathbf{x}_k \mathbf{x}_k^H\} = \sigma_x^2 \mathbf{I}_{M_t}$. The received signal power per transmit antenna is $\sigma_x^2 \sigma_h^2$, and the SNR $E_s/N_0 = M_t \cdot \sigma_x^2 \sigma_h^2 / N_0$.

At receiver side, the pilot hypersymbols are extracted from the received signals and employed for channel estimation. Since it is impossible to perform LMMSE estimation of $\mathbf{H}$ directly, we define the effective channel matrix $\mathbf{A}_k$ and estimate it first. The effective channel matrix is given by

$$\mathbf{A}_k = \mathbf{H}_k \mathbf{P}. \tag{2}$$

Due to the orthogonality of the columns of $\mathbf{P}$, $\mathbf{P}$ is of full rank and invertible. Therefore, it is possible to perform the Wiener filtering based on the observations $\tilde{\mathbf{A}}_k$ and to obtain the actual channel matrix $\hat{\mathbf{H}}_k$ as

$$\hat{\mathbf{H}}_k = \hat{\mathbf{A}}_k \mathbf{P}^{-1}, \tag{3}$$

where $\hat{\mathbf{A}}_k$ is the output of the Wiener filter. An illustrative example of MIMO channel estimation is depicted in Fig. 2 in [7]. The mean squared error (MSE) of this MIMO LMMSE channel estimator is identical to the MSE of a SISO LMMSE channel estimator assuming the same SNR E_s/N_0, the pilot spacing P_s, and the channel dynamics F_d.

After channel estimation, the ZF detector recovers the transmitted data hypersymbols

$$\hat{\mathbf{x}}_k = \hat{\mathbf{C}}_k \mathbf{y}_k = \mathbf{x}_k + \hat{\mathbf{C}}_k \mathbf{E}_k \mathbf{x}_k + \hat{\mathbf{C}}_k \mathbf{n}_k \tag{4}$$

with

$$\begin{aligned} \hat{\mathbf{C}}_k &= (\hat{\mathbf{H}}_k^H \hat{\mathbf{H}}_k)^{-1} \hat{\mathbf{H}}_k^H \\ \mathbf{E}_k &= \mathbf{H}_k - \hat{\mathbf{H}}_k. \end{aligned} \tag{5}$$

In conjunction with imperfect channel estimation, the ZF detection introduces interference $\hat{\mathbf{C}}_k \mathbf{E}_k \mathbf{x}_k$. When perfect CSI is assumed, $\hat{\mathbf{C}}_k$ is replaced by $\mathbf{C}_k = (\mathbf{H}_k^H \mathbf{H}_k)^{-1} \mathbf{H}_k^H$, and (4) reduces to

$$\hat{\mathbf{x}}_k = \mathbf{C}_k \mathbf{y}_k = \mathbf{x}_k + \mathbf{C}_k \mathbf{n}_k, \tag{6}$$

where only the colored noise $\mathbf{C}_k \mathbf{n}_k$ is left as distortion. In the remainder of this paper, We will drop the time index k where it is obvious.

III. Asymptotic BER for MIMO-BICM with Perfect CSI

In this section, we shortly summarize the derivations in [5]. For a MIMO-BICM system with ZF detector, in [5] it is shown that the expurgated bound for SISO systems derived in [2][1] can be applied to a single stream of a MIMO-BICM system with ZF detectors.

The upper bound of the coded BER at lth stream is given by

$$P_{b,l} \leq \frac{1}{k_c} \sum_{d=d_{\min,l}}^{\infty} W_{I,l}(d) f_l(d, \mu, \chi), \tag{7}$$

where $W_{I,l}(d)$ is the total input weight of error events and $f_l(d, \mu, \chi)$ denotes the pairwise error probability (PEP). $d_{\min,l}$ is the minimum Hamming distance of code employed on the lth stream. It should be stressed that since each transmit stream is encoded independently, different codes can be adopted for different streams. According to (7), the key quantity to be evaluated is the PEP since $W_{I,l}(d)$ is

[1] Sethuraman and Hajek in 2006 [10] have indicated that the proof that expurgated bound is an upper BER bound for QAM signal sets and Gray labeling is flawed. However, the results still represent a very tight approximation on the coded BER.

a known coefficient when a specific convolutional code is employed. The PEP can be approximated by the expurgated PEP $f_{\text{ex},l}(d,\mu,\chi)$ [2], [10]

$$f_l(d,\mu,\chi) \approx f_{\text{ex},l}(d,\mu,\chi)$$

with

$$f_{\text{ex},l}(d,\mu,\chi) = \frac{1}{2\pi j}\int_{v-j\infty}^{v+j\infty}[\psi_{\text{ex},l}(s)]^d\frac{ds}{s}, \tag{8}$$

where v is chosen such that the integral converges [2]. A numerical method called Gauss-Chebyshev quadratures can approximate the integral closely [11]. $\psi_{\text{ex},l}(s)$ is defined as

$$\psi_{\text{ex},l}(s) = \frac{1}{I2^I}\sum_{i=1}^{I}\sum_{b=0}^{1}\sum_{x\in\chi_b^i}\Phi_{\Delta_l(x,z)}(s), \tag{9}$$

where $\Phi_{\Delta_l(x,z)}(s)$ is the bilateral Laplace transform of the probability density function (pdf) of the metric difference $\Delta_l(x,z)$. x is the transmit symbol, whereas z represents the symbol that is erroneously detected because of an error event. The index l of x and z is dropped without loss of generality. For BICM, $z = z(x) \in \chi_b^i$ is the nearest neighboring constellation point of $x \in \chi_b^i$ for which χ_b^i holds. As Gray Mapping maximizes the average distance $\{x \to z\}$, it is considered here. $\Delta_l(x,z)$ is formally defined to be [5]

$$\Delta_l(x,z) = \log p(\hat{x}|x) - \log p(\hat{x}|z) \tag{10}$$

with

$$p(\hat{x}|x) = \frac{1}{\pi N_0\|\mathbf{c}_l\|^2}\exp\left\{-\frac{|\hat{x}-x|^2}{N_0\|\mathbf{c}_l\|^2}\right\} \tag{11}$$

for the perfect CSI case, where $\mathbf{c}_l \in C^{1\times M_r}$ is the lth row vector of $\mathbf{C}$. Inserting (11) into (10), the constant N_0 can be dropped, since it is only the probability $P(\Delta_l(x,z)<0)$ (Eq. (24,28) in [5]) that matters. The metric difference is given by

$$\Delta_l(x,z)_P = \frac{|\hat{x}-z|^2 - |\hat{x}-x|^2}{\|\mathbf{c}_l\|^2}, \tag{12}$$

where $(\cdot)_P$ denotes *perfect* CSI. We can then calculate the bilateral Laplace transform by following Eq.(29-33) in [5]. Normalizing the distance $d = x-z$ to σ_x^2, the result is expressed as

$$\begin{aligned}\Phi_{\Delta_l(x,z)_P}(s) &= \mathrm{E}_{\|\mathbf{c}_l\|^2}\left\{e^{-s\Delta_l(x,z)_P}\middle|\|\mathbf{c}_l\|^2\right\}\\ &= \left(1-\sigma_h^2\sigma_x^2|\tilde{d}|^2 s(N_0 s-1)\right)^{-(M_r-M_t+1)},\end{aligned} \tag{13}$$

where the exponent $(M_r - M_t + 1)$ indicates how the number of antennas effects the performance. If $M_t = M_r$, (13) reduces to the result obtained for a SISO system [2] with the same E_s/N_0.

IV. Asymptotic BER for MIMO-BICM with Imperfect CSI

The difference of the asymptotic BER analysis for the MIMO-BICM system with ZF detectors between the perfect and imperfect CSI case occurs when calculating the biliteral Laplace transform of the metric difference $\Phi_{\Delta_l(x,z)}$ and, of course, when deriving the metric difference $\Delta_l(x,z)_I$ itself. Therefore, we only focus on the calculations of the $\Phi_{\Delta_l(x,z)_I}$ in this section. $(\cdot)_I$ denotes "imperfect" CSI.

We assume that the receiver utilizes LMMSE channel estimation, which fulfills the orthogonal projection theorem

$$\mathrm{E}\left\{\hat{\mathbf{H}}_k\mathbf{E}_k^H\right\} = \mathbf{0}. \tag{14}$$

Substituting $\hat{\mathbf{c}}_l$ for $\mathbf{c}_l$ in (12), the metric difference for imperfect CSI is given as

$$\Delta_l(x,z)_I = \frac{|\hat{x}-z|^2-|\hat{x}-x|^2}{\|\hat{\mathbf{c}}_l\|^2} = \hat{\beta}_l(|\hat{x}-z|^2-|\hat{x}-x|^2), \tag{15}$$

where we define a new variable $\hat{\beta}_l$ as the inverse of the squared norm of the row vector $\hat{\mathbf{c}}_l$. Replacing $(x-z)$ by d and the detected signal $\hat{x}$ by $\hat{\mathbf{c}}_l(\mathbf{H}\mathbf{x}+\mathbf{n})$ yields

$$\Delta_l(x,z)_I = \hat{\beta}_l\left(2\,\mathrm{Re}\left\{\hat{\mathbf{c}}_l(\mathbf{H}\mathbf{x}+\mathbf{n})d^*\right\} - |x|^2 + |z|^2\right). \tag{16}$$

Making use of (14) results in

$$\begin{aligned}&\Delta_l(x,z)_I\\ &\quad= 2\hat{\beta}_l(\mathrm{Re}\left\{\hat{\mathbf{c}}_l\mathbf{n}d^*\right\} + \mathrm{Re}\left\{\hat{\mathbf{c}}_l\hat{\mathbf{H}}\mathbf{x}d^*\right\} + \mathrm{Re}\left\{\hat{\mathbf{c}}_l\mathbf{E}\mathbf{x}d^*\right\})\\ &\qquad- \hat{\beta}_l(|x|^2-|z|^2)\\ &\quad= 2\hat{\beta}_l\,\mathrm{Re}\left\{\hat{\mathbf{c}}_l\mathbf{n}d^*\right\} + \hat{\beta}_l(|x|^2 - 2\,\mathrm{Re}\left\{xz^*\right\} + |z|^2)\\ &\qquad+ 2\hat{\beta}_l\,\mathrm{Re}\left\{\hat{\mathbf{c}}_l\mathbf{E}\mathbf{x}d^*\right\})\\ &\quad= 2\hat{\beta}_l\,\mathrm{Re}\left\{\hat{\mathbf{c}}_l(\mathbf{n}+\mathbf{E}\mathbf{x})d^*\right\} + \hat{\beta}_l|d|^2.\end{aligned} \tag{17}$$

$\mathbf{n}$ and $\mathbf{E}\mathbf{x}$ are independent zero-mean complex Gaussian random variables. The covariance matrix of $\mathbf{E}\mathbf{x}$ is given by

$$\mathrm{E}\left\{\mathbf{E}\mathbf{x}\mathbf{x}^H\mathbf{E}^H\right\} = \sum_{i=1}^{M_t}|x_i|^2\sigma_e^2\mathbf{I}_{M_r} = M_t\sigma_x^2\sigma_e^2\mathbf{I}_{M_r}. \tag{18}$$

It should be emphasized that (18) only holds if σ_e^2 is independent of the time index k, and this independence exists when the channel sampling rate is faster than the Nyquist rate of the fading process [9]. Therefore, $\Delta_l(x,z)_I$ given $\hat{\mathbf{H}}$ is a Gaussian random variable with its mean value and variance given as

$$\mu_\Delta = \hat{\beta}_l|d|^2, \tag{19}$$

$$\begin{aligned}\sigma_\Delta^2 &= 4\hat{\beta}_l^2|d|^2\,\mathrm{E}\left\{\mathrm{Re}\left\{\hat{\mathbf{c}}_l(\mathbf{n}+\mathbf{E}\mathbf{x})\right\}^2\right\}\\ &= 2\hat{\beta}_l^2|d|^2(M_t\sigma_x^2\sigma_e^2 + N_0)\hat{\mathbf{c}}_l\mathbf{I}_{M_r}\hat{\mathbf{c}}_l^H\\ &= 2\hat{\beta}_l|d|^2(M_t\sigma_x^2\sigma_e^2 + N_0).\end{aligned} \tag{20}$$

In accordance with (13), we obtain the conditional bilateral Laplace transform of the pdf of $\Delta_l(x,z)_I$ as

$$\Phi_{\Delta_l(x,z)_I}(s) = \mathrm{E}_{\hat{\beta}_l}\left\{\exp\left(-\mu_\Delta s + \frac{1}{2}\sigma_\Delta^2 s^2\right)\middle|\hat{\beta}_l\right\}. \tag{21}$$

From (19) and (20), it follows that (21) can be expressed as

$$\begin{aligned}&\Phi_{\Delta_l(x,z)_I}(s)\\ &= \mathrm{E}_{\hat{\beta}_l}\left\{\exp\left(\hat{\beta}_l|d|^2\left((N_0 + M_t\sigma_x^2\sigma_e^2)s^2 - s\right)\right)\middle|\hat{\beta}_l\right\}.\end{aligned} \tag{22}$$

β_l is a Gamma random variable (Eq. (32) in [5]) and so is

$$\hat{\beta}_l \sim \ G(M_r - M_t + 1, \sigma_h^2 - \sigma_e^2)$$

with the moment generating function of its pdf given as

$$M_{\hat{\beta}_l}(t) = \mathrm{E}_{\hat{\beta}_l}\left\{\exp\left(\hat{\beta}_l t\right)\right\} = \left(1-(\sigma_h^2-\sigma_e^2)t\right)^{-(M_r-M_t+1)}. \tag{23}$$

In accordance with (13), the conditional expectation then can be calculated and expressed in a normalized way

$$\Phi_{\Delta_l(x,z)_I}(s) = \left(1-\sigma_x^2(\sigma_h^2-\sigma_e^2)|\tilde{d}|^2 s\left((N_0+M_t\sigma_x^2\sigma_e^2)s-1\right)\right)^{-(M_r-M_t+1)}. \tag{24}$$

V. Relation between Channel Estimation Error and SNR degradation

For the imperfect CSI case, the performance is clearly deteriorated because of the estimation error. By using the Chernoff bound approximation

$$f_l(d,\mu,\chi) \approx f_{ch,l}(d,\mu,\chi) = \min_{v>0}[\psi_{\mathrm{ex},l}(s)]^d, \tag{25}$$

where v is the real part of the smallest right-hand side pole of $\psi_{\mathrm{ex},l}(s)$. Therefore, v is $1/(2N_0 + 2M_t\sigma_x^2\sigma_e^2)$ for the imperfect CSI case and $1/(2N_0)$ for the perfect CSI case. Inserting this one into (25) and using (24) and (13), we get

$$f_{ch,l}(d,\mu,\chi)_I = \left(\frac{1}{I2^I}\sum_{i=1}^{I}\sum_{b=0}^{1}\sum_{x\in\chi_b^i}\left(1+\frac{|\tilde{d}|^2(\sigma_h^2-\sigma_e^2)\sigma_x^2}{4(N_0+M_t\sigma_x\sigma_e^2)}\right)^{-(M_r-M_t+1)}\right)^d$$

and

$$f_{ch,l}(d,\mu,\chi)_P = \left(\frac{1}{I2^I}\sum_{i=1}^{I}\sum_{b=0}^{1}\sum_{x\in\chi_b^i}\left(1+\frac{|\tilde{d}|^2\sigma_h^2\sigma_x^2}{4N_0}\right)^{-(M_r-M_t+1)}\right)^d. \tag{26}$$

Comparing $f_{ch,l}(d,\mu,\chi)_P$ and $f_{ch,l}(d,\mu,\chi)_I$ in (26), we can see that the reduction of the received signal power

$$\sigma_h^2\sigma_x^2 \to (\sigma_h^2-\sigma_e^2)\sigma_x^2 \tag{27}$$

and the enhancement of the interference and the noise

$$N_0 \to N_0 + M_t\sigma_x^2\sigma_e^2 \tag{28}$$

contributes to the degradation. This phenomenon can be interpreted as an additional SNR degradation and has been investigated when elaborating on the achievable rate of a SISO system [9] and a MIMO system [7]. Note that the comparison is made under the assumption that the transmit power per antenna σ_x^2 in (13) and (24) are the same. To make a fair comparison, we assume a constant power budget. This assumption leads to different **transmit signal power** per antenna σ_x^2 for the system with perfect and imperfect CSI because the constant energy per frame is shared between the data and the pilot symbols for the imperfect CSI case. As inserting pilot symbols reduces the power per transmit symbol, the E_s/N_0 decreases:

$$E_s/N_0 = (E_s/N_0)_P \cdot \frac{P_s - M_t}{P_s}, \tag{29}$$

where $(E_s/N_0)_P$ is the E_s/N_0 for the system without any pilot symbols. A completely fair comparison is based on both a constant transmit power and a constant number of transmitted symbols per frame. However, it requires adjustable code rates and modulation alphabet sizes to get a constant information rate. These factors greatly influence the performance, and jointly optimizing these parameters for certain P_s in order to guarantee a fair comparison is highly complicated. Therefore, we only stick to the "*constant power budget condition*". The overall SNR degradation factor with respect to the system with perfect and imperfect CSI can then be expressed as

$$L_{\mathrm{SNR}} = \left(\frac{P_s-M_t}{P_s}\right)\cdot\frac{1-\sigma_e^2/\sigma_h^2}{1+\sigma_e^2/\sigma_h^2(E_s/N_0)_P\cdot\left(\frac{P_s-M_t}{P_s}\right)}, \tag{30}$$

where the first term in the right-hand side is the energy loss caused by inserting pilot symbols and the second is the *estimation loss*.

For LMMSE channel estimation, σ_e^2 can be calculated analytically. Furthermore, by simplifying the fading spectrum to a rectangular spectrum and extending the filter length to infinite, the close-form expression of the estimation error is give by (cf. Eq. (29) in [7])

$$\sigma_e^2 = \frac{2F_dP_s}{2F_dP_s + (E_s/N_0)_P\cdot(\frac{P_s-M_t}{P_s})}\sigma_h^2. \tag{31}$$

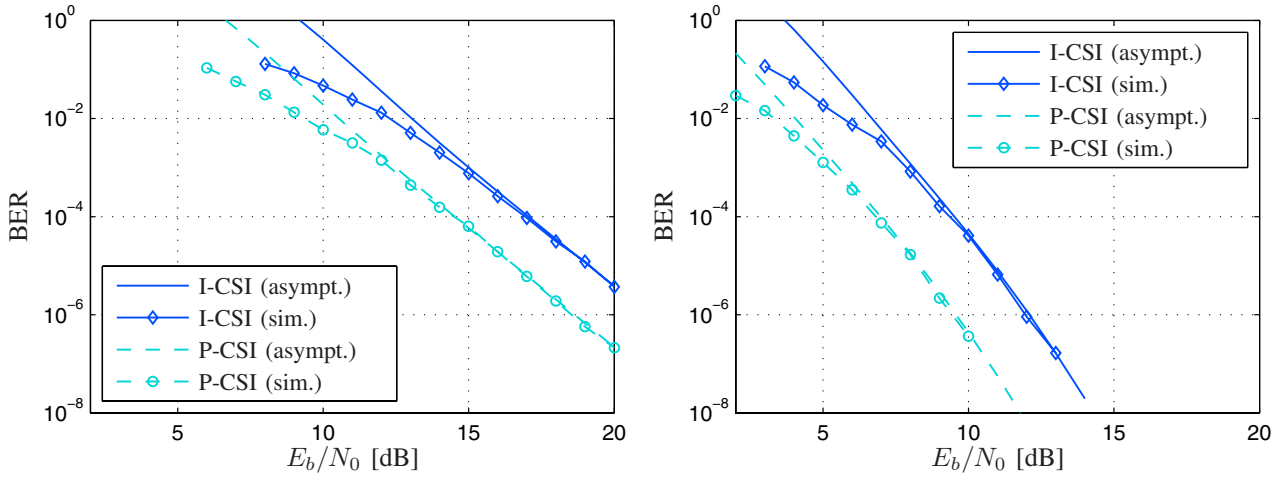
Making use of this expression, we can get

$$\tilde{L}_{\mathrm{SNR}} = \left(\frac{P_s-M_t}{P_s}\right)\cdot\frac{1}{2F_dP_s(1+\frac{P_s}{(E_s/N_0)_P(P_s-M_t)})+1}. \tag{32}$$

Then the optimum P_s for a given $(E_s/N_0)_P$, M_t, and F_d can be determined by using (32).

VI. Simulation Results

In this section, we validate the results on the asymptotic BER by means of Monte-Carlo simulation. The system is evaluated in a temporarily correlated flat Rayleigh fading channel with a Jakes' spectrum. The normalized Doppler spread is $F_d = 6.25\cdot10^{-3}$, the filter length $F_W = 10$, and the interleaver length $N = 6\cdot10^4$. The BCJR algorithm with its max-log MAP simplification is used for soft decoding. Fig. 2 gives the analytical results (lines without markers) and the results from the Monte-Carlo simulation (lines with markers). The solid lines represent the imperfect CSI case while the dashed lines correspond to the perfect CSI case. We denote them as "P-CSI" and "I-CSI" in the figure legends. We observe a asymptotically tight match between the analytical and the numerical result for both 2×2 and 2×4 MIMO systems. The performance gap between the P-CSI and I-CSI curves for the same MIMO system is exactly the estimation loss and energy loss caused by pilot inertion given in (30). Fig. 3 depicts the SNR loss for a MIMO system with different numbers of transmit and receive antennas vs. the pilot spacing. Different



(a) 2×2 MIMO-BICM with ZF detectors

(b) 2×4 MIMO-BICM with ZF detectors

Fig. 2. BER vs. E_b/N_0 for $M_t \times M_r$ MIMO systems; 8PSK, Gray Mapping, $N = 6 \cdot 10^4$, $f_d T_s = 6.25 \cdot 10^{-3}$, $P_s = 51$, $F_W = 10$, Conv. Code $\{5,7\}_8$, $r = 1/2$.

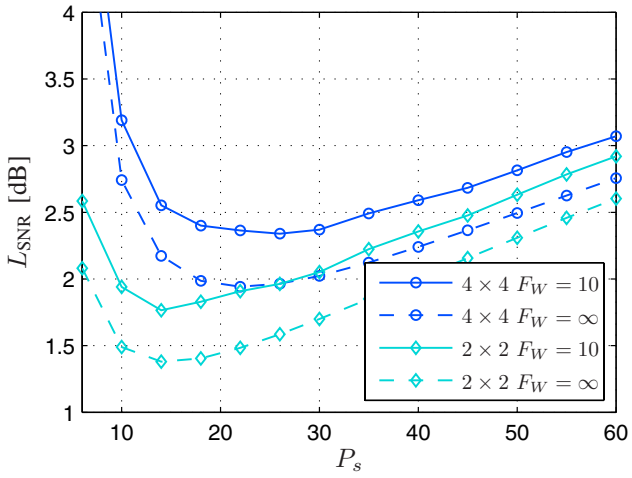


Fig. 3. SNR degradation vs. P_s for MIMO-BICM with ZF detectors; 8PSK, Gray Mapping, $E_b/N_0 = 17$ dB, $r = 1/2$, $f_d T_s = 6.25 \cdot 10^{-3}$, Conv. Code $\{5,7\}_8$

Wiener filter lengths are also considered. The solid lines correspond to a filter length $F_W = 10$ and the dashed lines correspond to an infinite filter length. The difference between them is the so-called *implementation loss*. The curves with circle markers represent a 4×4 MIMO system, whereas the curves with diamond markers represent a 2×2 MIMO system. The gap between these two systems results from inserting different pilot hypersymbol groups of different size. The gap obviously decreases as P_s increases. The curve of a 2×4 MIMO system not shown here coincides with that of a 2×2 MIMO system. We see that the optimum pilot spacing for a 2×2 MIMO-BICM with ZF detectors in this case is $P_s \approx 14$, while for a 4×4 system $P_s \approx 26$. It is worth noting that changing the filter length also slightly affects the optimum value of P_s, e.g., for the 4×4 MIMO system with infinite filter length $P_s \approx 22$ is optimal rather than $P_s \approx 26$.

VII. Conclusion

In this paper, we explore the asymptotic BER for a MIMO-BICM system with a ZF detector assuming imperfect CSI. We base our derivation on a spatially uncorrelated flat Rayleigh fading channel model with LMMSE channel estimation at receiver side. The deterioration with respect to perfect CSI can be well approximated by an SNR degradation that depends on the channel dynamics and the SNR. The intermediate performance, estimation error, and the final performance, coded BER, are linked together by this SNR degradation and the system design is significantly facilitated.

References

[1] E. Zehavi, "8-PSK Trellis codes for a Rayleigh channel," *IEEE Transactions on Communications*, vol. 40, no. 5, pp. 873–884, May 1992.

[2] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

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

[4] M. R. G. Bulter, I. B. Collings, "A zero-forcing approximate log-likelihood receiver for MIMO bit-interleaved coded modulation," *IEEE Communications Letters*, vol. 8, no. 2, pp. 105–107, Oct. 2004.

[5] M. R. McKay, I. B. Collings, "Capacity and performance of MIMO-BICM with zero-forcing receivers," *IEEE Transactions on Communications*, vol. 1, no. 1, pp. 74–83, Jan. 2005.

[6] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*, 1st ed. New York, NY: John Wiley & Sons, 1998.

[7] J. Baltersee, G. Fock, and H. Meyr, "Achievable rate of MIMO channels with data-aided channel estimation and perfect interleaving," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 12, pp. 2358–2368, Dec. 2001.

[8] S. Godtman, I. -W. Lai, T. -D. Chiueh, G. Ascheid, H. Meyr, "Analytical bit error rate calculation for BICM(-ID) assuming imperfect CSI," *to appear in Proceedings of 41th Asilomar Conference on Signals, Systems and Computers*, Nov. 2007.

[9] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2115–2123, Dec. 2001.

[10] V. Sethuraman and B. Hajek, "Comments on "Bit-interleaved coded modulation"," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1795–1797, Apr. 2006.

[11] E. Biglieri, G. Caire, G. Taricco, and J. Ventura-Traveset, "Simple method for evaluating error probabilities," *IEE Electronics Letters*, vol. 32, no. 3, pp. 191–192, Feb 1996.