MIMO-BICM WITH IMPERFECT CHANNEL STATE INFORMATION: EXIT CHART ANALYSIS AND LDPC CODE OPTIMIZATION

Clemens Novak[†]*, Gottfried Lechner*[*]*, and Gerald Matz*[†]

[†] Institut für Nachrichtentechnik und Hochfrequenztechnik,
Vienna University of Technology; E-mail: clemens.novak@nt.tuwien.ac.at

[*] Institute for Telecommunications Research
University of South Australia, Adelaide, Australia; E-mail: gottfried.lechner@unisa.edu.au

Abstract— We investigate MIMO systems employing bit-interleaved coded modulation and iterative decoding with imperfect channel estimates. An optimum soft demodulator for this scenario with i.i.d. Rayleigh fading was recently proposed by Sadough *et al.* We extend the optimum demodulator to correlated channel models and provide a code-independent performance comparison of this demodulator and the conventional (mismatched) demodulator for different symbol mappings by means of EXIT charts. Furthermore, we describe optimized LDPC code designs for these demodulators and verify the EXIT chart results in terms of bit error rate simulations.

1. INTRODUCTION

1.1. Background

Bit-interleaved coded modulation (BICM) has been introduced as a robust communication scheme in wireless channels. BICM systems with iterative decoding (BICM-ID) have been observed to yield excellent performance (see e.g. [1]). and BICM-ID can be employed in MIMO systems as well. The convergence properties of BICM-ID receivers has been studied in [2, 3] using EXIT charts [4]. BICM-ID systems that utilize LDPC codes [5] have been shown to be able to operate close to capacity limits. In addition, the EXIT charts of LDPC codes can be optimized by appropriately adjusting their variable and check node degree distributions.

In this paper, we consider BICM-ID within multiple-input multiple-output (MIMO) spatial multiplexing systems. Soft-in soft-out demodulators for MIMO-BICM-ID receivers are usually designed assuming perfect channel state information (CSI) and hence such conventional designs yield satisfactory performance only if accurate CSI is indeed available to the MIMO-BICM-ID receiver. In practical systems, additional pilot symbols are inserted at the transmitter to enable the receiver to perform channel estimation. However, inevitable estimation errors result in imperfect CSI. The resulting difference between the true channel and the estimated channel causes conventional MIMO demodulators to be mismatched which in turn degrades their performance. This motivated [6] to introduce an improved demodulator which explicitly takes the statistics of the channel estimate into account and thereby offers noticeable performance improvements for BICM-ID with imperfect CSI.

This work was supported by the STREP project MASCOT (IST-026905) within the Sixth Framework Programme of the European Commision. The authors wish to acknowledge the activity of the Network of Excellence in Wireless COMmunications NEWCOM++ of the European Commission (contract n. 216715) that motivated this work.

1.2. Contributions

The results in [6] were restricted to i.i.d. Rayleigh fading MIMO channels, MIMO-BICM with an off-the-shelf convolutional code, layer-wise Gray mapping, and a comparison of the performance of the improved and the mismatched demodulator in terms of bit error rate (BER). The contributions of this work are as follows:

- The improved demodulator of [6] is extended to MIMO channels with arbitrary spatial correlation.
- We use EXIT charts [4] to characterize the convergence behaviour of the MIMO-BICM-ID receivers employing the mismatched demodulator or the improved demodulator for different symbol mappings and different channel correlation models.
- We propose to use the maximum rates achievable with a specific demodulator as a code-independent performance metric. These rates are obtained by measuring the area under the EXIT function (cf. [7]). We compare the maximum achievable rates of the two demodulators for different mappings and channel models.
- Using the approach from [8], we design LDPC codes that are matched to a specific demodulator in terms of their EXIT functions.
- Finally, we provide BER comparisons of the various systems, using the optimized LDPC codes and a "standard" (i.e., non-optimized) LDPC code.

The paper is organized as follows: In Section 2 we present the system model, and in Section 3 the receiver is derived. In Section 4 we illustrate our results by numerical results. Conclusions are provided in Section 5.

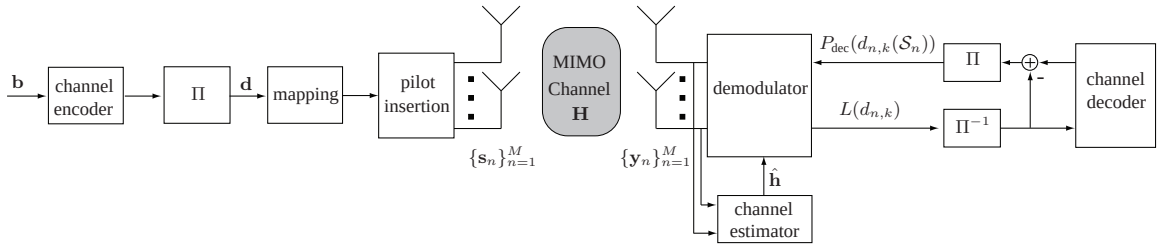

Fig. 1. *Block diagram of a pilot-assisted MIMO-BICM-ID system.*

2. SYSTEM MODEL

We consider the equivalent complex baseband representation of a MIMO-BICM system with M_T transmit antennas and M_R receive antennas. Assuming block flat fading, the length-M_R receive vector at symbol time n is given by

$$\mathbf{y}_n = \mathbf{H}\mathbf{s}_n + \mathbf{w}_n\,, \qquad n = 1, \ldots, N. \tag{1}$$

Here, $\mathbf{H}$ denotes the $M_R \times M_T$ MIMO channel matrix, $\mathbf{s}_n = (s_{n,1} \cdots s_{n,M_T})^T$ is the length-M_T transmit vector and $\mathbf{w}_n \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2\mathbf{I})$ denotes i.i.d. complex Gaussian noise and N is the block length. The symbols $s_{n,k}$ are taken from an alphabet $\mathcal{A}$ of size $|\mathcal{A}| = 2^B$. By stacking the columns of the channel matrix $\mathbf{H}$ into a vector $\mathbf{h} = \text{vec}\{\mathbf{H}\}$ and defining $\mathcal{S}_n = \mathbf{I} \otimes \mathbf{s}_n^T$, (1) can be rewritten as

$$\mathbf{y}_n = \mathcal{S}_n\mathbf{h} + \mathbf{w}_n\,, \qquad n = 1, \ldots, N. \tag{2}$$

The channel vector $\mathbf{h}$ is assumed zero-mean complex Gaussian with covariance matrix $\mathbf{C}_h$, $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_h)$.

The MIMO-BICM transmitter first encodes a length-K sequence of information bits $\mathbf{b} = (b_1 \cdots c_K)^T$ into a length-M sequence of code bits using an LDPC encoder (thus, the code rate equals $R = K/M$). The code bits are then passed through a random interleaver, yielding the interleaved code bit sequence $\mathbf{d} = (d_1 \cdots d_N)^T$. The kth element $s_{n,k}$ of The transmit vectors $\mathbf{s}_n$ are obtained by mapping groups of $L = BM_T$ successive bits $d_{n,k} = d_{l(n)+k}$, with $k = 1, \ldots, L$ and $l(n) = (n-1)L$, to M_T symbols from the alphabet $\mathcal{A}$. The number of transmit vectors is therefore $N = M/L$.

For channel estimation, a length-N_p pilot sequence with corresponding $N_\text{p}M_R \times M_TM_R$ matrix $\tilde{\mathcal{S}}_\text{p}$ is transmitted during a training phase. The received pilot sequence vector $\tilde{\mathbf{y}}_\text{p} = (\mathbf{y}_\text{p}^T[1] \cdots \mathbf{y}_\text{p}^T[N_\text{p}])^T$ of length $N_\text{p}M_R$ is (cf. (2))

$$\tilde{\mathbf{y}}_\text{p} = \tilde{\mathcal{S}}_\text{p}\mathbf{h} + \tilde{\mathbf{w}}, \tag{3}$$

with the stacked noise vector $\tilde{\mathbf{w}} = (\mathbf{w}^T[1] \cdots \mathbf{w}^T[N_\text{p}])^T$. We assume that the training sequence is orthogonal, $\tilde{\mathcal{S}}_\text{p}^H\tilde{\mathcal{S}}_\text{p} = \frac{N_\text{p}P}{M_TM_R}\mathbf{I}$, and has total power $\text{tr}(\tilde{\mathcal{S}}_\text{p}^H\tilde{\mathcal{S}}_\text{p}) = N_\text{p}P$ such that the effective training SNR equals $\text{SNR}_\text{p} = N_\text{p}P/\sigma_w^2$.

3. MIMO-BICM-ID RECEIVERS WITH IMPERFECT CSI

The iterative receiver structure employed is shown in Fig. 1. It consists of three blocks: a channel estimator, a demodulator and a channel decoder. Initially, the channel estimator provides an estimate $\hat{\mathbf{h}}$ of the channel based on the known pilot symbols. Demodulator and channel decoder are connected by interleavers and de-interleavers and form an iterative loop in which extrinsic information is exchanged.

The demodulator calculates posterior log-likelihood ratios (LLR)

$$L(d_{n,k}) = \log \frac{P_\text{dem}(d_{n,k} = 0)}{P_\text{dem}(d_{n,k} = 1)} \tag{4}$$

of the BM_T bits based on the observation of the received vector $\mathbf{y}_n$ and the extrinsic information $P_\text{dec}(d_{n,k})$ fed back by the channel decoder. The posterior LLRs in (4) are de-interleaved and passed to the channel decoder, which applies belief propagation [5] to decode the LDPC code. The decoder outputs LLRs for the code bits from which the demodulator LLRs are subtracted to obtain extrinsic LLRs of the code bits. These extrinsic LLRs are interleaved and provided as extrinsic information $P_\text{dec}(d_{n,k})$ to the demodulator for the next iteration step. In the last iteration, the signs of the posterior LLRs of the information bits provide the final bit decisions.

3.1. Channel Estimation

The channel matrix $\mathbf{H}$ is estimated based on the knwon pilot sequence $\tilde{\mathcal{S}}_\text{p}$ and the corresponding receive sequence $\tilde{\mathbf{y}}_\text{p}$ in (3). A general linear (homogeneous) estimator is given by

$$\hat{\mathbf{h}} = \mathbf{A}\tilde{\mathbf{y}}_\text{p} = \mathbf{A}(\tilde{\mathcal{S}}_\text{p}\mathbf{h} + \tilde{\mathbf{w}}), \tag{5}$$

with $\mathbf{A}$ being a $M_TM_R \times N_\text{p}M_R$ matrix. The least-squares (LS) channel estimate [9] $\hat{\mathbf{h}}_\text{LS}$ is obtained via the pseudo-inverse of the pilot matrix, i.e., $\mathbf{A} = \tilde{\mathcal{S}}_\text{p}^\# = \frac{M_TM_R}{N_pP}\tilde{\mathcal{S}}_\text{p}^H$ (here we used the orthogonality of the training sequence), and equals

$$\hat{\mathbf{h}}_\text{LS} = \tilde{\mathcal{S}}_\text{p}^H\mathbf{y}_\text{p} = \mathbf{h} + \mathbf{e}.$$

The elements of the error vector $\mathbf{e} = \frac{M_TM_R}{N_pP}\tilde{\mathcal{S}}_\text{p}^H\tilde{\mathbf{w}}$ are i.i.d. Gaussian with variance $\sigma_e^2 = \text{SNR}_\text{p}^{-1}$. The minimum mean square error (MMSE) estimator [9] $\hat{\mathbf{h}}_\text{MMSE}$ is obtained with

$$\mathbf{A} = \frac{1}{\sigma_w^2}\mathbf{\Sigma}\,\tilde{\mathcal{S}}_\text{p}^H, \qquad \text{with } \mathbf{\Sigma} = \left(\mathbf{C}_h^{-1} + \text{SNR}_\text{p}\mathbf{I}\right)^{-1}.$$

The posterior density $f(\mathbf{h}|\hat{\mathbf{h}})$ can be obtained as

$$f(\mathbf{h}|\hat{\mathbf{h}}) = \frac{f(\hat{\mathbf{h}}|\mathbf{h})f(\mathbf{h})}{f(\hat{\mathbf{h}})},$$

with $f(\hat{\mathbf{h}}) = \int f(\hat{\mathbf{h}}|\mathbf{h}) f(\mathbf{h}) d\mathbf{h}$. From (5), it follows that $\hat{\mathbf{h}}|\mathbf{h} \sim \mathcal{CN}(\mathbf{A}\tilde{\mathcal{S}}_{\mathrm{p}}\mathbf{h}, \sigma_w^2 \mathbf{A}\mathbf{A}^H)$. Using the channel model of Section 2 and the orthogonality of the training sequences $\tilde{\mathcal{S}}_{\mathrm{p}}$, $f(\mathbf{h}|\hat{\mathbf{h}})$ can be shown to be a complex Gaussian distribution:

$$\mathbf{h}|\hat{\mathbf{h}} \sim \mathcal{CN}\big(\hat{\mathbf{h}}_{\mathrm{MMSE}}, \boldsymbol{\Sigma}\big). \tag{6}$$

Note that this conditional distribution applies to *any* linear channel estimator.

3.2. Genie and Mismatched Demodulator

We first consider a genie demodulator that is in possession of perfect CSI. To simplify notation, we omit the time index n in what follows. The posterior LLR of the coded bits is calculated according to [6]

$$P_{\mathrm{dem}}(d_k = b) = \sum_{\mathbf{s} \in \chi_k^b} f(\mathbf{y}|\mathcal{S}, \mathbf{h}) \prod_{\substack{i=1 \\ i \neq k}}^{L} P_{\mathrm{dec}}(d_i(\mathcal{S})). \tag{7}$$

Here χ_k^b denotes the set of transmit vectors whose bit label at position k equals b and $P_{\mathrm{dec}}(d_i)$ denotes the extrinsic information provided by the decoder from the previous iteration (In the first iteration, $P_{\mathrm{dec}}(d_i) = \frac{1}{2}$). The conditional density $f(\mathbf{y}|\mathcal{S}, \mathbf{h})$ is obtained from the system model (2) as

$$f(\mathbf{y}|\mathcal{S}, \mathbf{h}) = \frac{1}{(\pi\sigma_w^2)^{M_R}} \exp\left(-\frac{\|\mathbf{y} - \mathcal{S}\mathbf{h}\|^2}{\sigma_w^2}\right), \tag{8}$$

abbreviated $\mathbf{y}|\mathcal{S}, \mathbf{h} \sim \mathcal{CN}(\mathcal{S}\mathbf{h}, \sigma_w^2 \mathbf{I})$. Since $\mathbf{h}$ is not available in practice, conventional receivers replace the actual channel $\mathbf{h}$ in (8) by the channel estimate $\hat{\mathbf{h}}$, i.e., instead of $\|\mathbf{y} - \mathcal{S}\mathbf{h}\|^2$ the metric $\|\mathbf{y} - \mathcal{S}\hat{\mathbf{h}}\|^2$ is evaluated. This approach is referred to as mismatched demodulation since in general $\hat{\mathbf{h}} \neq \mathbf{h}$.

3.3. Modified Demodulator

The problem with the mismatched demodulator is that it does not exploit the statistical information about $\mathbf{h}$ conveyed by the channel estimate $\hat{\mathbf{h}}$ (cf. (6)). In fact, what is available for demodulation is not $f(\mathbf{y}|\mathcal{S}, \mathbf{h})$ as in (7) but the conditional distribution $f(\mathbf{y}|\mathcal{S}, \hat{\mathbf{h}})$, which can be obtained by

$$f(\mathbf{y}|\mathcal{S}, \hat{\mathbf{h}}) = \int f(\mathbf{y}|\mathcal{S}, \mathbf{h}) f(\mathbf{h}|\hat{\mathbf{h}}) \, d\mathbf{h}.$$

This distribution takes the statistical properties of the channel estimate into account. Since both densities in the integral are Gaussian, $f(\mathbf{y}|\mathcal{S}, \hat{\mathbf{h}})$ is also Gaussian, i.e.,

$$\mathbf{y}|\,\mathcal{S}, \hat{\mathbf{h}} \sim \mathcal{CN}(\mathcal{S}\hat{\mathbf{h}}_{\mathrm{MMSE}}, \mathcal{S}\boldsymbol{\Sigma}\mathcal{S}^H + \sigma_w^2 \mathbf{I}).$$

Note that the mean of the conditional density $f(\mathbf{y}|\mathcal{S}, \hat{\mathbf{h}})$ is the same as that for a mismatched demodulator employing an MMSE channel estimate, but the covariance matrix is different and depends on the symbol $\mathcal{S}$. Using this expression in (7), leads to the improved demodulator which calculates the LLRs in (4) with $P_{\mathrm{dem}}(d_k = b)$ in (7) replaced by

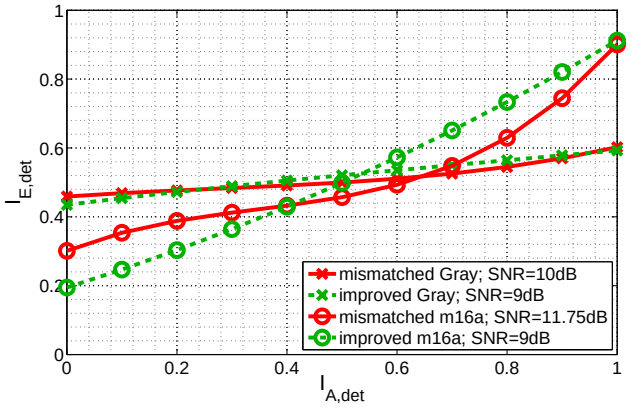

Fig. 2. *EXIT charts of mismatched and improved demodulator for Gray and m16a mapping.*

$$P_{\mathrm{dem}}(d_k = b) = \sum_{\mathbf{s} \in \chi_k^b} f(\mathbf{y}|\mathcal{S}, \hat{\mathbf{h}}) \prod_{\substack{i=1 \\ i \neq k}}^{L} P_{\mathrm{dec}}(d_i(\mathcal{S})). \tag{9}$$

In the special case of a spatially uncorrelated channel, $\mathbf{C}_h = \mathbf{I}$, the covariance matrix $\boldsymbol{\Sigma}$ reduces to

$$\boldsymbol{\Sigma} = \frac{1}{1 + \mathrm{SNR}_{\mathrm{p}}} \mathbf{I}$$

and the conditional density $f(\mathbf{y}|\mathcal{S}, \hat{\mathbf{h}})$ is

$$\mathbf{y}|\,\mathcal{S}, \hat{\mathbf{h}} \sim \mathcal{CN}\left(\mathcal{S}\hat{\mathbf{h}}_{\mathrm{MMSE}}, \left(\frac{\|\mathcal{S}\|^2}{M_T(1 + \mathrm{SNR}_{\mathrm{p}})} + \sigma_w^2\right)\mathbf{I}\right).$$

By inserting this into (9), the demodulator of [6] is reobtained.

4. NUMERICAL RESULTS

In the following, we present numerical results for a 2×2 BICM system with 16QAM symbol alphabet (normalized to unit power) and $N_{\mathrm{p}}P = 0.4$. The number of coded bits per channel use is $L = 8$. We considered two different mappings: layer-wise Gray mapping and layer-wise m16a mapping [10], which is a mapping specially optimized for BICM-ID with convolutional codes.

4.1. Demodulator EXIT Charts

The EXIT charts [4] of the demodulators were obtained by Monte Carlo simulations, using an AWGN channel for the a priori information. Fig. 4 shows the EXIT charts for the mismatched (with an LS channel estimator) and improved demodulator, both using Gray or m16a mapping. The SNRs have been chosen such that the area under the EXIT functions, which quantifies the maximum rate achievable with the respective demodulator [7], approximately equals $1/2$. With Gray mapping, the EXIT function of the mismatched

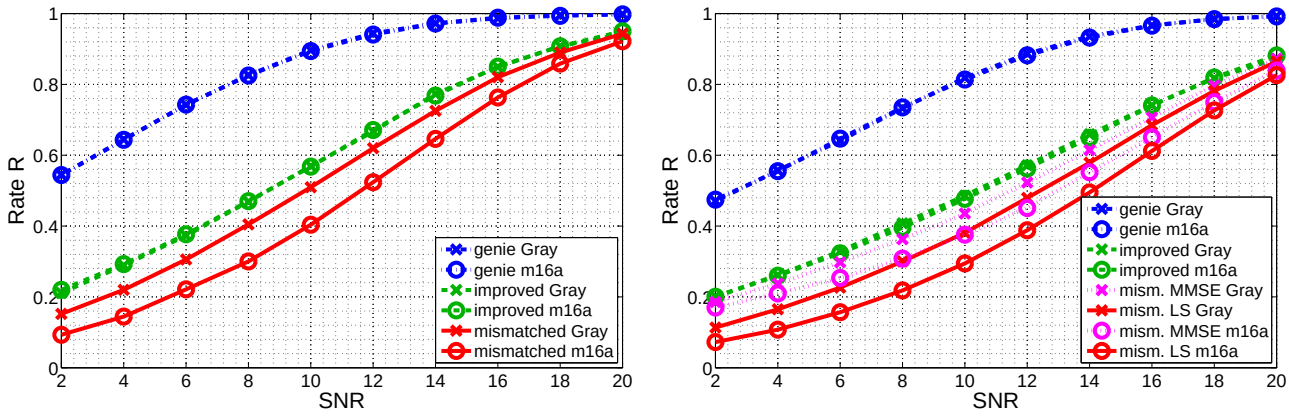

Fig. 3. *Rates achievable with genie, mismatched, and improved demodulator versus SNR for (a) uncorrelated Rayleigh fading and (b) correlated Rayleigh fading.*

demodulator at $10\,$dB SNR is almost identical to that of the improved detector at $9\,$dB. With m16a mapping, the EXIT functions of the two demodulators look quite different and the SNR required by mismatched and improved demodulation equals $11.75\,$dB and $9\,$dB, respectively. We conclude that the SNR threshold for the improved demodulator is identical under both mappings, even though different codes (matched to the respective EXIT function) are required to achieve this threshold. Furthermore, the SNR gain of the improved demodulator is about $1\,$dB for Gray mapping and $2.75\,$dB for m16a mapping.

Furthermore, codes designed for the mismatched demodulator with Gray mapping will also perform well for the improved detector with Gray mapping; however, with the improved demodulator the turbo cliff will occur at lower SNR. In contrast, with m16a mapping the pronounced difference between the EXIT functions of the mismatched and improved demodulator indicates that here the channel code should be matched to the respective demodulator used in order to avoid a large performance loss.

4.2. Achievable Rates

A code-independent measure for the performance of the various demodulators with different mappings is the maximum rate they allow to achieve with vanishing error probability. This rate can be measured via the area under the demodulator's EXIT chart when the *a priori* channel is a binary erasure channel [7]. For reasons of numerical stability, we used an AWGN channel as *a priori* channel for obtaining the EXIT charts, in which case the area yields a good approximation for the achievable rates.

The resulting maximum rates achievable with the genie, mismatched, and improved demodulator are plotted versus SNR for a spatially uncorrelated Rayleigh fading channel are shown in Fig. 3(a). It is seen that the improved demodulator is indeed superior to mismatched demodulation, even though there is still a significant gap to genie demodulation. For the genie and improved demodulator, the maximum rates are seen to be virtually the same for the two mappings used. Nevertheless, the corresponding EXIT charts (not shown) are different and achieving the maximum rates in an actual system thus requires matched code designs.

In contrast, for the mismatched demodulator the rate with m16a is lower than with Gray mapping. For a rate of $R = 1/2$, there is an SNR gap of about $2\,$dB between the mismatched demodulator with Gray and m16a. This indicates that the optimized m16a mapping is more sensitive to CSI inaccuracy than Gray mapping.

Fig. 3(b) shows similar results for the case of a channel with spatial correlation. We used a Kronecker model [11] for the correlation matrix of the channel, i.e., $\mathbf{C}_h = \mathbf{T}^{1/2} \otimes \mathbf{R}^{1/2}$, with the transmit and receive correlation matrices respectively chosen as

$$\mathbf{T} = \mathbf{R} = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}.$$

It is seen that the maximum achievable rates of the genie and modified demodulator are again almost independent of the mapping, albeit generaly smaller than in the uncorrelated case. The maximum rates achievable with the mismatched demodulator are shown for LS and MMSE channel estimation, both in conjunction with Gray and m16a mapping. Gray mapping is again preferable over m16a and in addition MMSE estimation is preferable over LS estimation due to its smaller channel estimation error. We conclude that m16a does not offer any advantage over Gray mapping in terms of maximum rates since the latter apparently is less sensitive to CSI inaccuracy.

4.3. BER Performance

We next present bit error rate (BER) results for LDPC codes with $5 \cdot 10^4$ code bits and code rate $R = 1/2$. With the same system parameters as before this amounts to 6250 transmit vectors. The MIMO channel was i.i.d. block fading, where the channel stays constant for 12 symbol periods (2 of which where used for training) and then a new, independent chan-

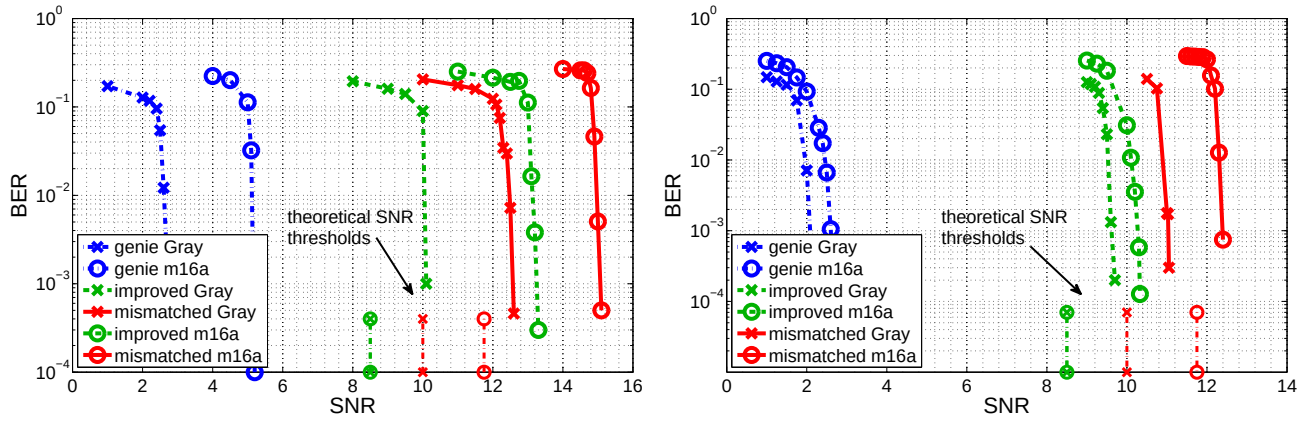

Fig. 4. *BER versus SNR for MIMO-BICM-ID employing genie, mismatched, or improved demodulator and Gray or m16a mapping for (a) a non-optimized LDPC code and (b) optimized LDPC codes.*

nel realization is drawn. The number of outer iterations (between demodulator and channel decoder) was 10, while the number of inner iterations (in the LDPC decoder) was 200. The mismatched demodulator was used with a LS channel estimator.

In Fig. 4(a), BER versus SNR for a non-optimized LDPC code with degree distribution $(3, 6)$ is shown together with the theoretical SNR thresholds. It is seen that there are significant gaps to the theoretical SNR thresholds, particularly for the genie and improved demodulator with m16a mapping. These gaps are caused by the mismatch between the EXIT functions of demodulator and code, which is significant when the m16a mapping is used. The EXIT charts of demodulators with Gray mapping are better matched to the code EXIT chart, therefore the gaps are smaller in this case. These BER results further confirm the superiority of the improved demodulator which outperforms the mismatched demodulator by about 2.4 dB (Gray) and 1.7 dB (m16a).

We further designed specifically optimized LDPC codes for each demodulator by matching the EXIT charts of the LDPC codes and of the demodulator according to [8]. The BER obtained with these optimized codes is plotted versus SNR in Fig. 4(b). All schemes are now much closer to the respective theoretical SNR thresholds. Furthermore, for genie and improved demodulation, Gray and m16a mapping now indeed feature approximately equal BER performance as predicted by Fig. 3(a) (m16a still performs slightly worse since here the code design does not achieve the theoretical optimum). The gain of improved demodulation over mismatched demodulation is about 1.5 dB (Gray mapping) and about 2 dB (m16a mapping).

5. CONCLUSIONS

We presented LDPC coded BICM-ID systems with imperfect CSI and derived an improved demodulator that accounts for the statistics of the channel estimate. The performance of the mismatched and improved demodulators using different mappings were characterized via EXIT charts and maximum achievable rates. BER performance was assessed using a regular $(3, 6)$ LDPC code and LDPC codes matched to the EXIT chart of the demodulator. The performance of the mismatched receiver was seen to depend strongly on the mapping (with m16a mapping performing worse than Gray mapping). In case of the improved demodulator, the maximum achievable rate noticeably larger and virtually independent of the mapping. However, with our code designs Gray mapping still performs better than m16a in terms of BER.

REFERENCES

[1] A. Chindapol and J. A. Ritcey, "Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE J. Sel. Areas Comm.*, vol. 19, pp. 944–957, May 2001.

[2] Y. Huang and J. Ritcey, "EXIT chart analysis of BICM-ID with imperfect channel state information," *IEEE Comm. Letters*, vol. 7, pp. 434–436, Sept. 2003.

[3] T. Clevorn, S. Godtmann, and P. Vary, "BER prediction using EXIT charts for BICM with iterative decoding," *IEEE Comm. Letters*, vol. 10, pp. 49–51, Jan. 2006.

[4] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Comm.*, vol. 49, pp. 1727–1737, Oct. 2001.

[5] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, 2001.

[6] S. Sadough, P. Piantanida, and P. Duhamel, "MIMO-OFDM optimal decoding and achievable information rates under imperfect channel estimation," in *Proc. IEEE-SP Workshop on Signal Processing Advances in Wireless Communications*, pp. 1–5, 2007.

[7] A. Ashikhmin, G. Kramer, and S. Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, 2004.

[8] G. Lechner, J. Sayir, and I. Land, "Optimization of LDPC Codes for Receiver Frontends," *Proc. IEEE ISIT*, pp. 2388–2392, 2006.

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs (NJ): Prentice Hall, 1993.

[10] F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Comm. Letters*, vol. 7, pp. 593–595, Dec. 2003.

[11] C. Oestges, B. Clerckx, D. Vanhoenacker-Janvier, and A. Paulraj, "Impact of fading correlations on MIMO communication systems in geometry-based statistical channel models," *IEEE Trans. Wireless Comm.*, vol. 4, pp. 1112–1120, May 2005.