CAPACITY-BASED PERFORMANCE COMPARISON OF MIMO-BICM DEMODULATORS

Peter Fertl, Joakim Jaldén, and Gerald Matz

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
phone: +43 1 58801 38942, fax: +43 1 58801 38999, email: {pfertl, jjalden, gmatz}@nt.tuwien.ac.at

ABSTRACT

This paper provides a performance comparison of multiple-input multiple-output (MIMO) demodulators for bit-interleaved coded modulation (BICM) systems with non-iterative demodulation and decoding. We propose to use the capacity of an equivalent "modulation" channel as a performance measure that has the advantage of not depending on the outer error correcting code. Based on this approach, we conclude that a universal ranking of MIMO (soft and hard) demodulation algorithms is not possible. This result is confirmed via bit error rate simulations for a practical system involving low-density parity-check codes. Our approach also allows to derive practical guidelines for MIMO-BICM system design.

1. INTRODUCTION

1.1. Background

Bit-interleaved coded modulation (BICM) [1] has been conceived as a pragmatic approach to coded modulation. It has received a lot of attention in wireless communications due to its bandwidth and power efficiency and its robustness against fading. For single-antenna systems, BICM with Gray labeling can approach channel capacity [1, 2]. These advantages have motivated extensions of BICM to multiple-input multiple-output (MIMO) systems [3, 4].

In MIMO-BICM systems the optimal inner demodulator is the soft-output maximum a posteriori probability (MAP) detector, which provides the channel decoder with log-likelihood ratios (LLR) for the code bits. Its high computational complexity can be reduced without significant performance loss by using the so-called max-log approximation [4]. An exact implementation of the max-log MAP detector based on sphere decoding was presented in [5]. This implementation still involves a search set that grows exponentially with the number of transmit antennas

A computationally efficient approximation of the max-log demodulator with polynomial worst-case complexity has been proposed in [6] based on a semi-definite relaxation approach. Other approaches obtain approximate LLR values from a candidate list that constitutes a reduced search set. Several demodulators employ tree search techniques to generate this list, a well-known example being the list sphere decoder [7]. Alternatively, the candidate list can be obtained using lattice-reduction techniques (cf. [8]).

MIMO soft demodulators with very small complexity employ zero-forcing (ZF) equalization [9] or minimum mean-square error (MMSE) equalization [10] followed by scalar per-antenna soft demodulation. In a similar vein, successive interference cancelation can be extended to the soft output case [11].

This work was supported by the STREP project MASCOT (IST-026905) within the Sixth Framework Programme of the European Commission and by the WWTF project MOHAWI (MA 44).

The performance of MIMO soft demodulators is usually illustrated via numerical results showing coded bit error rate (BER) versus signal-to-noise ratio (SNR). However, these results can depend strongly on the outer error correcting code and on several other system parameters (symbol constellation, labeling, etc.). This renders a meaningful performance comparison extremely difficult.

1.2. Contributions and Organization of Paper

In this paper, we advocate an information theoretic approach for measuring the performance of MIMO (soft and hard) demodulators in the context of non-iterative (single-shot) receivers. Specifically, following [3], we propose to use the mutual information between the interleaved code bits and the MIMO demodulator output to assess the demodulator performance independently of the outer code. We note that ZF-based demodulation and max-log demodulation have been compared in a similar spirit in [9].

For space reasons, we restrict ourselves to a baseline selection of MIMO-BICM soft demodulators for which we performed extensive Monte-Carlo simulations in order to estimate the proposed mutual information performance measure. Our results allow for a number of interesting conclusions. Most importantly, we observed that the performance ranking of MIMO demodulators is *rate-dependent*. As an example, MMSE soft demodulation outperforms optimum hard decision maximum likelihood (ML) detection for low rates while for high rates it is just the other way around. This somewhat surprising behavior is verified in terms of BER simulations using low-density parity-check (LDPC) codes with different code rates. Our framework provides practical design guidelines for MIMO-BICM systems, e.g., which demodulator to prefer at a specific code rate.

The rest of this paper is organized as follows. Section 2 discusses the MIMO-BICM system model and Section 3 reviews MIMO soft demodulation. The mutual information based performance measure is presented in Section 4 and Monte-Carlo simulation results are presented in Section 5. Finally, our conclusions are summarized in Section 6.

2. MIMO-BICM SYSTEM MODEL

The MIMO-BICM model we consider is shown in Fig. 1. A sequence of information bits $b[q]$ is encoded using an error-correcting code and then passed through a *bitwise interleaver* Π. The interleaved code bits are demultiplexed into M_{T} antenna streams ("layers"). In each layer, groups of m bits are mapped bijectively to (complex) data symbols $x_k[n] \in \mathcal{A}$, $k = 1, \ldots, M_{\mathrm{T}}$; here, $\mathcal{A}$ denotes the symbol alphabet of size $|\mathcal{A}| = 2^m$. The transmit vector at symbol time n is given by[1] $\mathbf{x}[n] \triangleq (x_1[n], \ldots, x_{M_{\mathrm{T}}}[n])^T$ and carries $R_0 = mM_{\mathrm{T}}$ interleaved code bits $c_l[n]$, $l = 1, \ldots, R_0$.

[1] Superscript ^{T} (^{H}) denotes the (Hermitian) transpose.

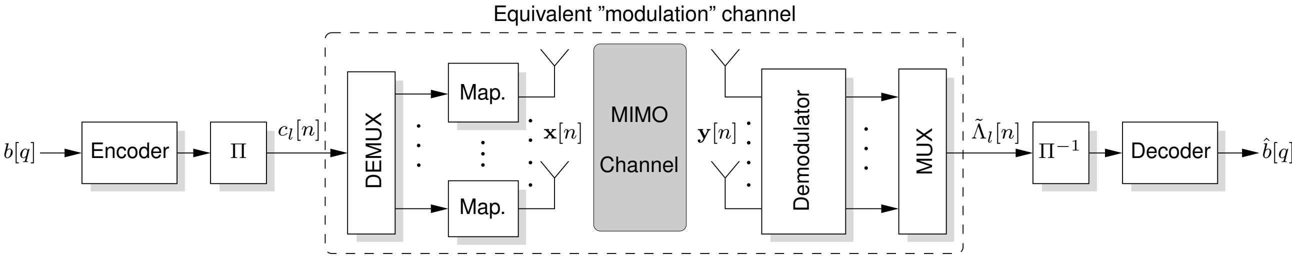


Fig. 1. MIMO-BICM transmission model.

Assuming M_{R} receive antennas and a fast fading MIMO channel, the receive vector $\mathbf{y}[n] \triangleq (y_1[n], \ldots, y_{M_{\mathrm{R}}}[n])^T$ is given by

$$\mathbf{y}[n] = \mathbf{H}[n]\mathbf{x}[n] + \mathbf{w}[n], \qquad n = 1, \ldots, N, \tag{1}$$

where $\mathbf{H}[n]$ is the $M_{\mathrm{R}} \times M_{\mathrm{T}}$ channel matrix, and the components of the noise vector $\mathbf{w}[n] \triangleq (w_1[n], \ldots, w_{M_{\mathrm{R}}}[n])^T$ are modeled as i.i.d. circularly symmetric complex Gaussian with zero mean and variance σ_w^2. The vector[2] $\mathbf{x}[n] \in \mathcal{A}^{M_{\mathrm{T}}}$ obeys the power constraint $\mathbb{E}\{\|\mathbf{x}[n]\|^2\} = E_s$ ($\mathbb{E}\{\cdot\}$ denotes expectation). Assuming a channel matrix with normalized entries, the SNR at each receive antenna equals $\rho \triangleq E_s/\sigma_w^2$. In the remainder of this paper, we will omit the time index n for convenience.

At the receiver, the demodulator uses the received vector $\mathbf{y}$ and the channel matrix $\mathbf{H}$ (assumed perfectly known) to calculate LLRs Λ_l (or approximate LLRs $\tilde{\Lambda}_l$ in case of a suboptimal demodulator) for all code bits c_l, $l = 1, \ldots, R_0$, carried by $\mathbf{x}$. The resulting LLRs are deinterleaved using Π^{-1} and then passed to the channel decoder that delivers detected bits $\hat{b}[q]$.

3. REVIEW OF MIMO SOFT DEMODULATORS

For space reasons, we focus on a small set of MIMO soft demodulators, which we will briefly review next. Our selection includes demodulators from opposite extremes with respect to complexity.

3.1. Optimum and Max-Log Demodulator

Assuming that all transmit vectors are equally likely and using the conditional probability density function (pdf) (cf. (1))

$$f(\mathbf{y}|\mathbf{x}, \mathbf{H}) = \frac{1}{(\pi\sigma_w^2)^{M_{\mathrm{R}}}} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2}{\sigma_w^2}\right), \tag{2}$$

the optimal soft MAP demodulator calculates the exact LLR for c_l according to [4]

$$\Lambda_l \triangleq \log \frac{f(c_l = 1|\mathbf{y}, \mathbf{H})}{f(c_l = 0|\mathbf{y}, \mathbf{H})} = \log \frac{\sum\limits_{\mathbf{x} \in \mathcal{X}_l^1} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma_w^2}\right)}{\sum\limits_{\mathbf{x} \in \mathcal{X}_l^0} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma_w^2}\right)}. \tag{3}$$

Here, $\mathcal{X}_l^1$ and $\mathcal{X}_l^0$ denote the sets of transmit vectors for which $c_l = 1$ and $c_l = 0$, respectively (note that $\mathcal{A}^{M_{\mathrm{T}}} = \mathcal{X}_l^1 \cup \mathcal{X}_l^0$). While giving exact LLRs, computation of (3) has complexity $\mathcal{O}(|\mathcal{A}|^{M_{\mathrm{T}}})$, i.e., increasing exponentially with the number of transmit antennas.

[2] Here, $\mathcal{A}^{M_{\mathrm{T}}} \triangleq \mathcal{A} \times \ldots \times \mathcal{A}$ denotes the M_{T}-fold Cartesian product of $\mathcal{A}$.

Applying the max-log approximation [4] to (3) results in

$$\Lambda_l \approx \frac{1}{\sigma_w^2}\left[\min_{\mathbf{x} \in \mathcal{X}_l^0} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 - \min_{\mathbf{x} \in \mathcal{X}_l^1} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2\right]. \tag{4}$$

The right-hand side of (4) is easier to compute than (3) but still involves a search over a set of size $|\mathcal{A}|^{M_{\mathrm{T}}}$.

3.2. Linear Demodulators

For convenience, in the following we denote the ith bit in the bit label of the kth symbol x_k by $c_k^{(i)}$ and the corresponding LLR by $\Lambda_k^{(i)}$. Low-complexity demodulators can be obtained by using a linear (ZF or MMSE) equalizer followed by *per-layer* (max-log) LLR calculation according to

$$\tilde{\Lambda}_k^{(i)} = \frac{1}{\sigma_k^2}\left[\min_{x \in \mathcal{A}_i^0} |\hat{x}_k - x|^2 - \min_{x \in \mathcal{A}_i^1} |\hat{x}_k - x|^2\right]. \tag{5}$$

Here, $\mathcal{A}_i^b \subset \mathcal{A}$ denotes the set of (scalar) symbols whose bit label at position i equals b, $\hat{x}_k$ is an estimate of the symbol in layer k provided by the equalizer, and σ_k^2 is an equalizer-specific weight. We emphasize that calculating LLRs separately for each layer results in a significant complexity reduction. In fact, ZF- and MMSE-based demodulation (discussed below in more detail) have a complexity of $\mathcal{O}(M_{\mathrm{R}}^3)$, which is much smaller than the complexity of max-log demodulation according to (4).

ZF-based Demodulator [9]. Here, the first stage consists of ZF equalization, i.e.,

$$\hat{\mathbf{x}}_{\mathrm{ZF}} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{y} = \mathbf{x} + \mathbf{z}, \tag{6}$$

where the noise $\mathbf{z}$ has correlation matrix

$$\mathbf{R}_{\mathbf{z}} = \sigma_w^2\,(\mathbf{H}^H\mathbf{H})^{-1}. \tag{7}$$

Subsequently, approximate bit LLRs are obtained according to (5) with symbol estimate $\hat{x}_k = (\hat{\mathbf{x}}_{\mathrm{ZF}})_k$ and weight factor $\sigma_k^2 = (\mathbf{R}_{\mathbf{z}})_{k,k}$.

MMSE-based Demodulator [10]. Here, the first stage is an MMSE equalizer that can be written as (cf. (6) and (7))

$$\hat{\mathbf{x}}_{\mathrm{MMSE}} = \mathbf{W}\hat{\mathbf{x}}_{\mathrm{ZF}}, \qquad \text{with} \quad \mathbf{W} = (\mathbf{I} + \mathbf{R}_{\mathbf{z}})^{-1}.$$

Letting $W_{k,k} = (\mathbf{W})_{k,k}$ denote the kth diagonal element of $\mathbf{W}$, the approximate LLR values in (5) are calculated using

$$\hat{x}_k = \frac{(\hat{\mathbf{x}}_{\mathrm{MMSE}})_k}{W_{k,k}} \quad \text{and} \quad \sigma_k^2 = \frac{1 - W_{k,k}}{W_{k,k}}.$$

We note that in spite of similar complexity, MMSE-based demodulation substantially outperforms ZF-based demodulation [10].

4. CAPACITY MEASURES

In order for the information rates discussed below to have interpretations as ergodic capacities, we assume that the channel $\mathbf{H}[n]$ is a stationary, finite-memory process.

4.1. Capacity of MIMO Coded Modulation

In a coded modulation (CM) system with equally likely transmit vectors $\mathbf{x} \in \mathcal{A}^{M_\mathrm{T}}$ and no CSI at the transmitter (which implies $I(\mathbf{x};\mathbf{H}) = 0$), the average mutual information in bits per channel use is given by[3] (cf. [3])

$$C_\mathrm{CM} \triangleq I(\mathbf{x};\mathbf{y}|\mathbf{H}) = R_0 - \mathbb{E}_{\mathbf{x},\mathbf{y},\mathbf{H}}\left\{\log_2 \frac{\sum_{\mathbf{x}' \in \mathcal{A}^{M_\mathrm{T}}} f(\mathbf{y}|\mathbf{x}',\mathbf{H})}{f(\mathbf{y}|\mathbf{x},\mathbf{H})}\right\}. \tag{8}$$

This expression involves the conditional pdf in (2). In the following, we will refer to C_CM as *CM capacity* [1]. It is seen from (8) that $C_\mathrm{CM} \leq R_0$ with the "raw" bit rate $R_0 = mM_\mathrm{T}$; in fact, the last term in (8) is a penalty term resulting from noise and MIMO interference.

Using the fact that the mapping between $\mathbf{x}$ and the bit label $c_1, \ldots, c_{R_0}$ is bijective and applying the chain rule of mutual information [12] to (8), we obtain

$$C_\mathrm{CM} = I(c_1, \ldots, c_{R_0};\mathbf{y}|\mathbf{H}) = \sum_{l=1}^{R_0} I(c_l;\mathbf{y}|c_1, \ldots, c_{l-1},\mathbf{H}). \tag{9}$$

We note that the single-antenna equivalent of (9) served as a motivation for multilevel coding and multistage decoding which has been shown to achieve CM capacity [2].

4.2. Capacity of MIMO-BICM

The capacity obtained by using BICM can be written as (cf. [3])

$$\begin{aligned} C_\mathrm{BICM} &\triangleq \sum_{l=1}^{R_0} I(c_l;\mathbf{y}|\mathbf{H}) \qquad (10)\\ &= R_0 - \sum_{l=1}^{R_0} \mathbb{E}_{\mathbf{x},\mathbf{y},\mathbf{H},b}\left\{\log_2 \frac{\sum_{\mathbf{x}' \in \mathcal{A}^{M_\mathrm{T}}} f(\mathbf{y}|\mathbf{x}',\mathbf{H})}{\sum_{\mathbf{x}' \in \mathcal{X}_l^b} f(\mathbf{y}|\mathbf{x}',\mathbf{H})}\right\}, \end{aligned}$$

where $b \in \{0,1\}$ is equiprobable, and the code bits c_l are assumed uniformly distributed and statistically independent (as guaranteed, e.g., by an ideal interleaver). Since conditioning cannot reduce mutual information, comparison of (10) and (9) reveals that

$$C_\mathrm{BICM} \leq C_\mathrm{CM}.$$

For single-antenna systems, it has been shown that the gap between BICM with Gray labeling and CM capacity is negligible. However, for MIMO the gap $C_\mathrm{CM} - C_\mathrm{BICM}$ increases with $|\mathcal{A}|$ and M_T and can thus be considerably larger. It furthermore depends strongly on the chosen constellation labeling [4].

[3] Sometimes, C_CM is referred to as constellation-constrained capacity. We recall that the ergodic capacity with Gaussian inputs is given by $C = \mathbb{E}_\mathbf{H}\{\log_2 \det\left(\mathbf{I} + \frac{\rho}{M_\mathrm{T}}\mathbf{H}\mathbf{H}^H\right)\}$.

4.3. Demodulator Performance

Observing that Λ_l in (3) is a sufficient statistic [12] for c_l given $\mathbf{y}$ and $\mathbf{H}$, we obtain[4]

$$C_\mathrm{BICM} = \sum_{l=1}^{R_0} I(c_l;\Lambda_l).$$

Thus, C_BICM can be interpreted as the capacity of an equivalent channel with input c_l, output Λ_l, and channel transition function $f(\Lambda_l|c_l)$.

This finally motivates us to measure the performance of a specific MIMO-BICM demodulator in terms of the mutual information of an equivalent "modulation" channel (cf. Fig. 1) with *discrete* input c_l and a *continuous* output that is constituted by the approximate LLRs $\tilde{\Lambda}_l$. The capacity of this channel, which is characterized by $f(\tilde{\Lambda}_l|c_l)$, can be shown to equal

$$\begin{aligned} R &\triangleq \sum_{l=1}^{R_0} I(c_l;\tilde{\Lambda}_l) \qquad (11)\\ &= R_0 - \sum_{l=1}^{R_0}\sum_{b=0}^{1}\int_{-\infty}^{\infty} \frac{1}{2} f(\tilde{\Lambda}_l|c_l = b)\log_2 \frac{f(\tilde{\Lambda}_l)}{\frac{1}{2}f(\tilde{\Lambda}_l|c_l = b)} d\tilde{\Lambda}_l\,, \end{aligned}$$

where $f(\tilde{\Lambda}_l) = \frac{1}{2}\sum_{b=0}^{1} f(\tilde{\Lambda}_l|c_l = b)$. The capacity R provides a code-independent performance measure for (approximate) soft demodulators. In fact, it has the intuitive operational interpretation as the highest rate achievable with a BICM system (in the sense of [1, Section III.A]) that employs the specific demodulator which produces $\tilde{\Lambda}_l$. Since $\tilde{\Lambda}_l$ is derived from $\mathbf{y}$ and $\mathbf{H}$, the data processing inequality [12] implies that $R \leq C_\mathrm{BICM}$ with equality if $\tilde{\Lambda}_l = \Lambda_l$. The performance of a soft demodulator can thus be measured in terms of the gap $C_\mathrm{BICM} - R$. Of course, the information theoretic performance measure in (11) does not take into account complexity issues and it has to be expected that there is a tradeoff between R and computational complexity.

We note that the performance of hard detectors for MIMO-BICM can be quantified in a similar manner. For any soft demodulator, the associated hard detector is given by $\hat{c}_l = \mathrm{u}(\tilde{\Lambda}_l)$, where $\mathrm{u}(\cdot)$ denotes the unit step function (although in most cases this will not be the way the hard detector is actually implemented). With respect to capacity, the only difference is that here the equivalent "modulation" channel is a binary symmetric channel (BSC) with *discrete* output $\hat{c}_l$. Consequently, the integral over $\tilde{\Lambda}_l$ in (11) is replaced with a summation over $\hat{c}_l \in \{0,1\}$.

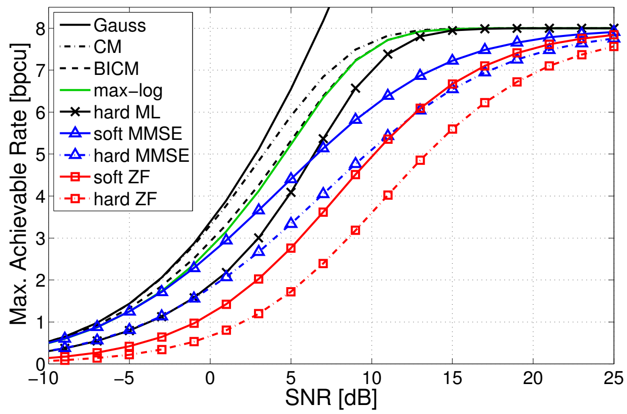
5. NUMERICAL RESULTS

In what follows we will compare the performance of the soft demodulators discussed in Section 3 as well as their hard detection variants for the case of Gray labeling and ergodic i.i.d. fast Rayleigh fading. We first provide a comparison in terms of R in (11), with C_CM in (8) and C_BICM in (10) as well as the channel capacity with Gaussian input serving as benchmarks. The conclusions obtained from this comparison are afterwards verified via BER simulations using outer LDPC codes.

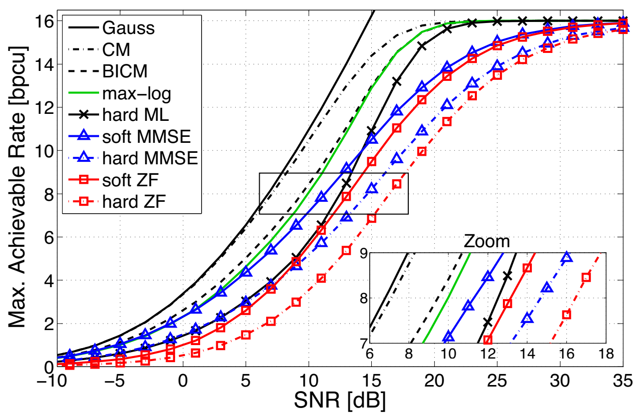
5.1. Maximum Achievable Information Rates

Since the pdfs required for computing the maximum rate R in (11) are hard to obtain in closed form, we used Monte-Carlo simulations

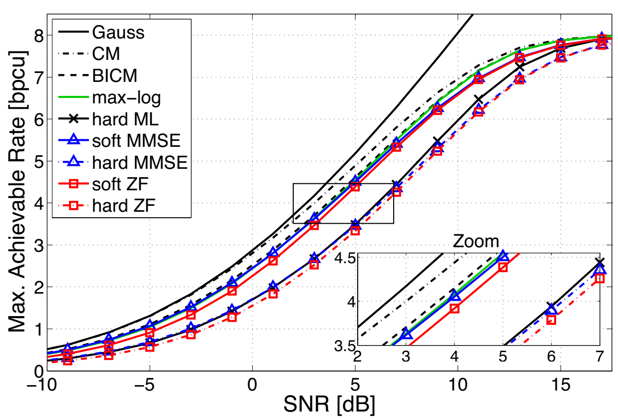
[4] Note, however, that $\{\Lambda_1, \ldots, \Lambda_{R_0}\}$ is *not* a sufficient statistic for $\{c_1, \ldots, c_{R_0}\}$.



(a)



(b)



(c)

Fig. 2. Demodulator comparison for (a) 4×4 MIMO with 4-QAM, (b) 4×4 MIMO with 16-QAM, and (c) 4×2 MIMO with 16-QAM.

to measure these pdfs and perform numerical integration. The results obtained for the various demodulators are shown in Fig. 2 in bit per channel use (bpcu) versus SNR ρ for the case of (a) 4×4 MIMO system with 4-QAM ($R_0 = 8$), (b) 4×4 with 16-QAM ($R_0 = 16$), and (c) 4×2 with 16-QAM ($R_0 = 8$). We note that the zoomed areas in Fig. 2 allow to assess the SNR gaps between the various schemes for a target rate of $R_0/2$ bpcu.

It is seen from Fig. 2(a) that (hard and soft) ZF demodulation performs significantly poorer than all other demodulators. At 4 bpcu, CM capacity is virtually identical to Gaussian capacity and BICM capacity looses about 1.3 dB. The penalty of using the max-log approximation is only about 0.3 dB. Furthermore, at this rate hard vector ML detection is 2.1 dB away from max-log demodulation whereas the corresponding gap for soft and hard MMSE demodulation equals 1.2 dB and 4.1 dB, respectively. An interesting observation in this scenario is the fact that at low rates, soft and hard MMSE detection essentially coincide with max-log and hard ML demodulation, respectively, whereas at high rates MMSE demodulation becomes highly suboptimal and approaches ZF performance. This points to the somewhat surprising result that the performance ranking is not universal but depends on the target rate (even for fixed number of antennas, constellation size, and labeling). Specifically, soft MMSE demodulation outperforms hard ML detection at low rates whereas at high rates it is just the other way around. Such a behavior can also be seen for hard MMSE and soft ZF demodulation. Very much the same observations apply when using 16-QAM instead of 4-QAM (cf. Fig. 2(b)) or when using set-partitioning labeling instead of Gray labeling (not shown here due to space reasons). Apart from a general shift to higher SNR values, the main effect of using the larger constellation is an increase of the gap between CM capacity and BICM capacity.

The above observations no longer hold true for the 4×2 case shown in Fig. 2(c). The increased SNR gap between CM and BICM capacity implied by the larger constellation is compensated by having more receive than transmit antennas (this agrees with observations in [4]). In addition, the performance differences between the individual demodulators are significantly reduced, the essential distinction being between soft and hard demodulators. In this scenario, it is also seen that both MMSE- and ZF-based soft demodulation outperform hard output ML detection for all rates.

5.2. BER Performance

To verify the foregoing observations for practical systems, we show BER results for a 4×4 MIMO-BICM system with 4-QAM in conjunction with regular LDPC codes of block length 51200 [13], designed[5] for rates of $1/4$ and $3/4$ (corresponding to 2 and 6 bpcu, respectively). We focus on MMSE demodulation (hard and soft) and hard output ML detection as well as max-log demodulation. For the case of soft demodulation, we used LDPC codes designed for additive white Gaussian noise channels whereas with hard detectors LDPC codes designed for a BSC were employed. At the receiver, message-passing LDPC decoding [13] was performed.[6] In the case of hard detection, the decoder was provided with the approximate LLRs

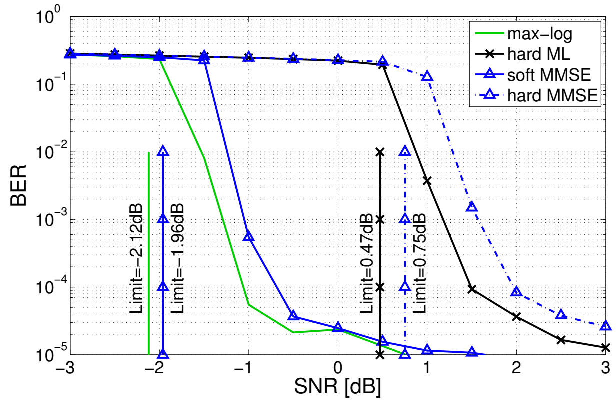
$$\tilde{\Lambda}_l \triangleq (2\hat{c}_l - 1)\,\log\left(\frac{1-p_0}{p_0}\right), \tag{12}$$

where p_0 is the cross-over probability of the equivalent BSC (determined via Monte-Carlo simulations). The choice in (12) is critical to provide the LDPC decoder with accurate reliability information.

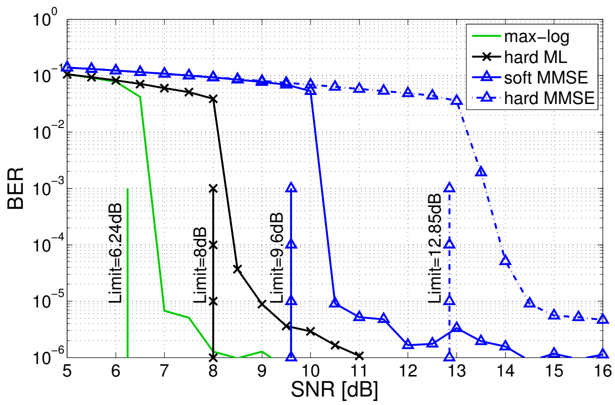
The BERs obtained are shown in Fig. 3 where also the capacity limit is indicated in terms of the minimal SNRs required for the respective rate according to Fig. 2(a). Note that the BER curves have an error floor typical for LDPC codes with message-passing decoding [14]. Generally speaking, our LDPC code designs are roughly

[5]The LDPC code design was performed using the EPFL web-tool at `http://lthcwww.epfl.ch/research/ldpcopt/`.

[6]We thank Gottfried Lechner for kindly providing his LDPC decoder implementation.



(a)



(b)

Fig. 3. BER performance for a 4×4 MIMO system with 4-QAM constellation using a (a) rate $1/4$ and (b) rate $3/4$ LDPC code.

1 dB away from the capacity limits. Furthermore, for both rates and all SNRs max-log demodulation performs best whereas hard MMSE detection performs worst. However, such universal statements cannot be made for hard ML and soft MMSE demodulation. Specifically, it can be seen in Fig. 3(a) that at rate $1/4$, the soft MMSE demodulator outperforms the hard ML detector by about 2.3 dB and is only 0.2 dB away from max-log demodulation. However, for rate $3/4$ (see Fig. 3(b)), soft MMSE demodulation performs about 1.9 dB poorer than hard ML detection and looses 3.5 dB with respect to max-log demodulation. These results agree well with the conclusions in the previous section. We note that similar observations can be inferred from the block error rate simulations in [15].

6. CONCLUSIONS

We have investigated the performance of various soft and hard demodulation schemes for non-iterative MIMO-BICM systems based on an information-theoretic measure which can be interpreted as maximum achievable rate of an equivalent "modulation" channel and which is independent of the outer code. Our simulations reveal that there is no universal demodulator performance ranking, since such a ranking depends strongly on the rate (or equivalently the SNR) at which the system operates. Our performance measure further provides guidelines how to select demodulator, number of antennas, and constellation size, to achieve a certain target information rate with a certain trade-off between required SNR and demodulator complexity. As an example, 4 bpcu can be achieved with an SNR of 2.8 dB using a $4{\times}4$ system with 4-QAM, rate $1/2$ code, and max-log demodulation. In this scenario the less complex soft ZF demodulator requires an SNR of 7.9 dB. However, the same rate can be obtained using a 4×2 system with 16-QAM for which the soft ZF demodulator requires only an SNR of 4.1 dB. A more comprehensive comparison of other MIMO soft demodulators mentioned in the introduction is left to future work.

7. REFERENCES

[1] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, pp. 927–945, May 1998.

[2] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *European Trans. Telecomm.*, vol. 6, pp. 557–567, Sept./Oct. 1995.

[3] E. Biglieri, G. Taricco, and E. Viterbo, "Bit-interleaved time-space codes for fading channels," in *Proc. Conf. on Information Sciences and Systems*, Princeton, NJ, March 2000, pp. WA 4.1–4.6.

[4] S. H. Müller-Weinfurtner, "Coding approaches for multiple antenna transmission in fast fading and OFDM," *IEEE Trans. Signal Processing*, vol. 50, pp. 2442–2450, Oct. 2002.

[5] J. Jaldén and B. Ottersten, "Parallel implementation of a soft output sphere decoder," in *Proc. 39th Asilomar Conf. Signals, Systems, Computers*, 2005, pp. 581– 585.

[6] B. Steingrimsson, Z.-Q. Luo, and K. M. Wong, "Soft quasi-maximum-likelihood detection for multiple-antenna wireless channels," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2710–2719, Nov. 2003.

[7] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 389–399, March 2003.

[8] V. Ponnampalam, D. McNamara, A. Lillie, and M. Sandell, "On generating soft outputs for lattice-reduction-aided MIMO detection," in *Proc. IEEE ICC-07*, June 2007, pp. 4144–4149.

[9] M. R. McKay and I. B. Collings, "Capacity and performance of MIMO-BICM with zero-forcing receivers," *IEEE Trans. Comm.*, vol. 53, no. 1, pp. 74–83, Jan. 2005.

[10] D. Seethaler, G. Matz, and F. Hlawatsch, "An efficient MMSE-based demodulator for MIMO bit-interleaved coded modulation," in *Proc. IEEE Globecom 2004*, Dallas, Texas, Dec. 2004, vol. IV, pp. 2455–2459.

[11] Won-Joon Choi, Kok-Wui Cheong, and J.M. Cioffi, "Iterative soft interference cancellation for multiple antenna systems," in *Proc. IEEE WCNC-2000*, Sept. 2000, vol. 1, pp. 304–309.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[13] 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.

[14] T. Richardson, "Error floors of LDPC codes," in *Proc. Allerton Conference on Communication, Control, and Computing*, Illinois, USA, Oct. 2003, pp. 1426–1435.

[15] C. Michalke, E. Zimmermann, and G. Fettweis, "Linear MIMO receivers vs. tree search detection: A performance comparison overview," in *Proc. IEEE PIMRC-06*, Dresden, Germany, Sept. 2006.