

# Performance Assessment of MIMO-BICM Demodulators based on System Capacity

Peter Fertl\*, *Student Member, IEEE*, Joakim Jaldén, *Member, IEEE*,  
and Gerald Matz, *Senior Member, IEEE*

## Abstract

In this paper we provide a comprehensive performance assessment and comparison of various soft-output and hard-output demodulation schemes for multiple-input multiple-output (MIMO) bit-interleaved coded modulation (BICM). Even though widely used in literature for demodulator comparison, coded bit error rate (BER) has the drawback of being strongly dependent on the outer error correcting code. This motivates us to propose a code-independent performance measure in terms of *system capacity*, i.e., mutual information of the equivalent modulation channel that comprises modulator, wireless channel, and demodulator. We present extensive numerical results for ergodic and quasi-static channels as well as for the case of imperfect channel state information. These results reveal that the performance ranking of MIMO demodulators is rate-dependent. Furthermore, they provide new insights regarding MIMO-BICM system design, specifically the joint selection of demodulation scheme, antenna configuration, and symbol constellation for a prescribed target rate.

## Index Terms

MIMO, BICM, performance limits, soft demodulation, system capacity, log-likelihood ratio

## EDICS Category: MSP-CAPC, MSP-CODR

The authors are with the 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, e-mail: {pfertl, jjalden, gmatz}@nt.tuwien.ac.at).

Part of this work has been previously presented at IEEE 9th Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2008) [1].

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

## I. INTRODUCTION

### A. Background

Bit-interleaved coded modulation (BICM) [2], [3] 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 [2], [4]. These advantages have motivated extensions of BICM to multiple-input multiple-output (MIMO) systems [5]–[7].

In MIMO-BICM systems, the optimum demodulator is the soft-output maximum a posteriori (MAP) demodulator, which provides the channel decoder with log-likelihood ratios (LLRs) for the code bits. Due to its high computational complexity, numerous alternative demodulators have been proposed in the literature, the most important of which are briefly describe below. Applying the max-log approximation [7] to the MAP demodulator reduces complexity without significant performance loss and leads to a search for data vectors minimizing a Euclidean norm. Efficient exact implementations of the max-log MAP detector based on sphere decoding have been presented in [8]–[10]. The implementation complexity can be further reduced by replacing the Euclidean norm with the  $\ell^\infty$  norm [11], [12]. However, the complexity of these methods still grows exponentially with the number of transmit antennas. An alternative demodulator that approximates the max-log LLRs is based on semidefinite relaxation (SDR) and has polynomial worst-case complexity [13].

Several demodulation schemes use a list of candidate data vectors to deliver approximate LLR values. Decreasing the size of the candidate list allows to trade off performance for complexity savings. A popular way to generate such lists are tree search techniques, with the list sphere decoder (LSD) [14] being a well-known example. Another approach to obtain a candidate list is based on lattice reduction (LR) [15], [16]. Variants of lattice-reduction aided list demodulators were proposed in [17]–[19]. Alternatively, candidate vectors can be obtained by “flipping” some of the bits in the bit label of the data vector obtained by (hard) maximum likelihood (ML) demodulation [20] (variants are obtained by replacing the ML detector with other hard-output detectors).

MIMO demodulators with still smaller complexity consist of a linear equalizer followed by per-layer scalar soft demodulators. This approach has been studied using zero-forcing (ZF) [21], [22] and minimum mean-square error (MMSE) [23] equalization. The soft interference canceler (SoftIC) proposed in [24] iteratively performs per-layer demodulation, preceded by a subtraction of an interference estimate. This estimate is generated using soft symbols for the other layers obtained in the previous iteration.

Low-complexity alternatives to soft demodulators are given by hard-output MIMO detectors that provide tentative decisions for the code bits but no associated reliability information. Among the best-known schemes here are ML, ZF, and MMSE demodulation [25] and successive interference cancelation (SIC) [26]–[28].

### B. Contributions

In the context of MIMO-BICM, the performance of the above MIMO demodulators has usually been assessed numerically in terms of coded bit error rate (BER), with a specific choice for the outer channel code. However, such numerical BER results can change dramatically with different channel codes<sup>1</sup>, thereby rendering a meaningful demodulator comparison difficult.

In this paper, we advocate an information theoretic approach for assessing the performance of (soft and hard) MIMO demodulators in the context of non-iterative (single-shot) BICM receivers (see also [1]). Following [5], we propose the mutual information between the code bits and the associated MIMO demodulator output as a *code-independent* performance measure. This quantity can be interpreted as system capacity (maximum rate allowing for error-free information recovery) for an equivalent “modulation” channel that comprises modulator and soft/hard demodulator in addition to the physical channel. This approach establishes a systematic framework for the assessment of current and future MIMO demodulation schemes. We note that ZF-based and max-log demodulation have been compared in a similar spirit in [22].

Based on Monte-Carlo simulations of the proposed system capacity, this paper provides extensive performance evaluations and comparisons for the above-mentioned MIMO demodulators for various system configurations in fast fading and in quasi-static fading. We also investigate the performance loss of the various demodulation schemes in non-idealized scenarios where only estimates of the channel matrix and the noise variance are available at the receiver. Due to lack of space, only part of our numerical results are shown here. Additional results for other antenna configurations, symbol constellations, and symbol labelings can be found in a supporting document [29].

Our results allow for several interesting conclusions. Most importantly, we found that there is no universal performance ranking of MIMO demodulators, i.e., such a ranking depends on the code rate or, equivalently, on the signal-to-noise ratio (SNR). As an example, MMSE soft demodulation outperforms hard ML demodulation at low rates while at high rates it is the other way around. We also verify this

<sup>1</sup>Of course, performance also depends on system parameters, like antenna configuration, symbol constellation, labeling, etc.

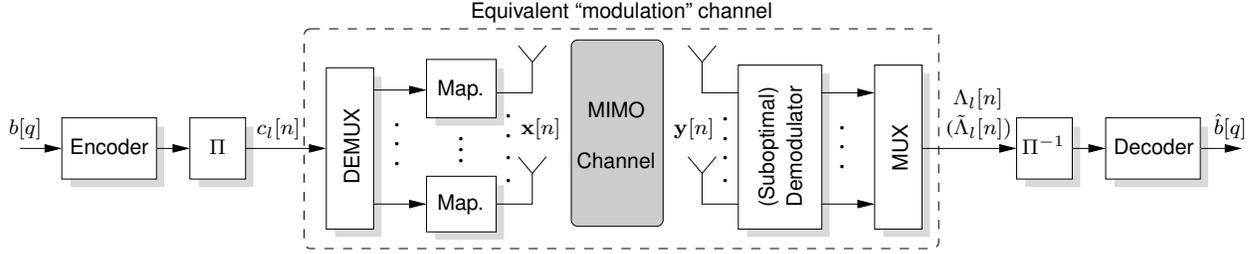


Fig. 1. Block diagram of a MIMO-BICM system.

surprising observation in terms of BER simulations using low-density parity-check (LDPC) codes. Finally, we use our numerical results to develop practical guidelines for the design of MIMO-BICM systems, i.e., which antenna configuration, symbol constellation, and demodulator to prefer in order to achieve a certain rate with minimum SNR.

### C. Paper Organization

The rest of this paper is organized as follows. Section II discusses the MIMO-BICM system model and Section III proposes system capacity as performance measure. In Sections IV and V, we assess the system capacity achievable with the MIMO-BICM demodulators referred to above for the case of fast fading. Section VI analyzes the impact of imperfect channel state information (CSI) on demodulator performance, and Section VII investigates the rate-versus-outage tradeoff of selected demodulators in quasi-static environments. In Section VIII, we summarize key observations and infer practical system design guidelines. Finally, conclusions are provided in Section IX.

## II. SYSTEM MODEL

### A. MIMO-BICM Transmission Model

A block diagram of our MIMO-BICM model is shown in Fig. 1. A sequence of information bits  $b[q]$  is encoded using an error-correcting code and is then passed through a *bitwise* interleaver  $\Pi$ . The interleaved code bits are demultiplexed into  $M_T$  antenna streams (“layers”). In each layer  $k = 1, \dots, M_T$ , groups of  $Q$  code bits  $c_k^{(i)}[n]$ ,  $i = 1, \dots, Q$ , ( $n$  denotes symbol time) are mapped via a one-to-one function  $\mu(\cdot)$  to (complex) data symbols  $x_k[n]$  from a symbol alphabet  $\mathcal{A}$  of size  $|\mathcal{A}| = 2^Q$ . Specifically,  $x_k[n] = \mu(c_k^{(1)}[n], \dots, c_k^{(Q)}[n])$ , where  $\{c_k^{(1)}[n], \dots, c_k^{(Q)}[n]\} = \mu^{-1}(x_k[n])$  is referred to as the bit label of  $x_k[n]$ .

The transmit vector is given by<sup>2</sup>  $\mathbf{x}[n] \triangleq (x_1[n] \dots x_{M_T}[n])^T \in \mathcal{A}^{M_T}$  and satisfies the power constraint  $\mathbb{E}\{\|\mathbf{x}[n]\|^2\} = E_s$ . It carries  $R_0 = Q M_T$  interleaved code bits  $c_l[n]$ ,  $l = 1, \dots, R_0$ , with  $c_k^{(i)}[n] = c_{kQ+i}[n]$ . We will for simplicity write  $\mathbf{x}[n] = \mu(c_1[n], \dots, c_{R_0}[n])$  and  $\mathbf{c}[n] \triangleq (c_1[n] \dots c_{R_0}[n])^T = \mu^{-1}(\mathbf{x}[n])$  as shorthand for the mapping  $\mathbf{x}[n] = (\mu(c_1^{(1)}[n], \dots, c_1^{(Q)}[n]) \dots \mu(c_{M_T}^{(1)}[n], \dots, c_{M_T}^{(Q)}[n]))^T$  and its inverse.

Assuming flat fading, the receive vector  $\mathbf{y}[n] \triangleq (y_1[n] \dots y_{M_R}[n])^T$  ( $M_R$  denotes the number of receive antennas) is given by

$$\mathbf{y}[n] = \mathbf{H}[n]\mathbf{x}[n] + \mathbf{v}[n], \quad n = 1, \dots, N, \quad (1)$$

where  $\mathbf{H}[n]$  is the  $M_R \times M_T$  channel matrix, and  $\mathbf{v}[n] \triangleq (v_1[n] \dots v_{M_R}[n])^T$  is a noise vector with independent identically distributed (i.i.d.) circularly symmetric complex Gaussian elements with zero mean and variance  $\sigma_v^2$ . In most of what follows, we will omit the time index  $n$  for convenience.

At the receiver, the optimum demodulator uses the received vector  $\mathbf{y}$  and the channel matrix  $\mathbf{H}$  to calculate LLRs  $\Lambda_l$  for all code bits  $c_l$ ,  $l = 1, \dots, R_0$ , carried by  $\mathbf{x}$ . In practice, the use of suboptimal demodulators or of a channel estimate  $\hat{\mathbf{H}}$  will result in approximate LLRs  $\tilde{\Lambda}_l$ . The LLRs are passed through the deinterleaver  $\Pi^{-1}$  and then on to the channel decoder that delivers the detected bits  $\hat{b}[q]$ .

### B. Optimum Soft MAP Demodulation

Assuming i.i.d. uniform code bits (as guaranteed, e.g., by an ideal interleaver), the optimum soft MAP demodulator calculates the exact LLR for  $c_l$  based on  $(\mathbf{y}, \mathbf{H})$  according to [7]

$$\Lambda_l \triangleq \log \frac{p(c_l=1|\mathbf{y}, \mathbf{H})}{p(c_l=0|\mathbf{y}, \mathbf{H})} = \log \frac{\sum_{\mathbf{x} \in \mathcal{X}_l^1} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2}{\sigma_v^2}\right)}{\sum_{\mathbf{x} \in \mathcal{X}_l^0} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2}{\sigma_v^2}\right)}. \quad (2)$$

Here,  $p(c_l|\mathbf{y}, \mathbf{H})$  is the probability mass function (pmf) of the code bits conditioned on  $\mathbf{y}$  and  $\mathbf{H}$ ,  $\mathcal{X}_l^1$  and  $\mathcal{X}_l^0$  denote the complementary sets of transmit vectors for which  $c_l=1$  and  $c_l=0$ , respectively (note that  $\mathcal{A}^{M_T} = \mathcal{X}_l^1 \cup \mathcal{X}_l^0$ ). Unfortunately, computation of (2) has complexity  $\mathcal{O}(|\mathcal{A}|^{M_T}) = \mathcal{O}(2^{R_0})$ , i.e., exponential in the number of transmit antennas. For this reason, several suboptimal demodulators have been proposed which promise near-optimal performance while requiring a lower computational complexity. The aim of this work is to provide a fair performance comparison of these demodulators.

<sup>2</sup>The superscripts <sup>T</sup> and <sup>H</sup> denote transposition and Hermitian transposition, respectively. Furthermore,  $\mathcal{A}^{M_T} \triangleq \mathcal{A} \times \dots \times \mathcal{A}$  is the  $M_T$ -fold Cartesian product of  $\mathcal{A}$ ,  $\mathbb{E}\{\cdot\}$  denotes expectation, and  $\|\cdot\|$  is the  $\ell^2$  (Euclidean) norm.

### III. SYSTEM CAPACITY

In order for the information rates discussed below to have interpretations as ergodic capacities, we consider a fast fading scenario where the channel  $\mathbf{H}[n]$  is a stationary, finite-memory process. We recall that the ergodic capacity with Gaussian inputs is given by

$$C_G = \mathbb{E}_{\mathbf{H}} \left\{ \log_2 \det \left( \mathbf{I} + \frac{E_s}{M_T \sigma_v^2} \mathbf{H} \mathbf{H}^H \right) \right\} \quad (3)$$

(here,  $\mathbf{I}$  denotes the identity matrix). The non-ergodic regime (slow fading) is discussed in Section III-D.

#### A. Capacity of MIMO Coded Modulation

In a coded modulation (CM) system with equally likely transmit vectors  $\mathbf{x} \in \mathcal{A}^{M_T}$  and no CSI at the transmitter, the average mutual information in bits per channel use (bpcu) is given by (cf. [5])

$$C_{\text{CM}} \triangleq I(\mathbf{x}; \mathbf{y} | \mathbf{H}) = R_0 - \mathbb{E}_{\mathbf{x}, \mathbf{y}, \mathbf{H}} \left\{ \log_2 \frac{\sum_{\mathbf{x}' \in \mathcal{A}^{M_T}} f(\mathbf{y} | \mathbf{x}', \mathbf{H})}{f(\mathbf{y} | \mathbf{x}, \mathbf{H})} \right\}. \quad (4)$$

This expression involves the conditional probability density function (pdf) (cf. (1))

$$f(\mathbf{y} | \mathbf{x}, \mathbf{H}) = \frac{1}{(\pi \sigma_v^2)^{M_R}} \exp \left( -\frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2}{\sigma_v^2} \right).$$

In the following, we will refer to  $C_{\text{CM}}$  as *CM capacity* [2] (sometimes,  $C_{\text{CM}}$  is alternatively termed “constellation-constrained capacity”). It is seen from (4) that  $C_{\text{CM}} \leq R_0$  with the “raw” bit rate  $R_0 = Q M_T$ ; in fact, the last term in (4) may be interpreted as a penalty term resulting from the noise and MIMO interference.

Using the fact that the mapping between the symbol vector  $\mathbf{x}$  and the associated bit label  $\{c_1, \dots, c_{R_0}\}$  is one-to-one and applying the chain rule for mutual information [30] to (4) leads to

$$C_{\text{CM}} = I(c_1, \dots, c_{R_0}; \mathbf{y} | \mathbf{H}) = \sum_{l=1}^{R_0} I(c_l; \mathbf{y} | c_1, \dots, c_{l-1}, \mathbf{H}). \quad (5)$$

The single-antenna equivalent of (5) served as a motivation for multilevel coding and multistage decoding, which can indeed achieve CM capacity [4].

### B. Capacity of MIMO-BICM

Using the assumption of i.i.d. uniform code bits, the maximum rate achievable with BICM can be shown to be given by (cf. [5])

$$\begin{aligned}
 C_{\text{BICM}} &\triangleq \sum_{l=1}^{R_0} I(c_l; \mathbf{y} | \mathbf{H}) \\
 &= R_0 - \sum_{l=1}^{R_0} \mathbb{E}_{\mathbf{x}, \mathbf{y}, \mathbf{H}, b} \left\{ \log_2 \frac{\sum_{\mathbf{x}' \in \mathcal{A}^{M_T}} f(\mathbf{y} | \mathbf{x}', \mathbf{H})}{\sum_{\mathbf{x}' \in \mathcal{X}_l^b} f(\mathbf{y} | \mathbf{x}', \mathbf{H})} \right\},
 \end{aligned} \tag{6}$$

where  $b \in \{0, 1\}$  with equal probability. Since conditioning cannot reduce mutual information, a comparison of (6) and (5) reveals that

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

The gap  $C_{\text{CM}} - C_{\text{BICM}}$  increases with  $|\mathcal{A}|$  and  $M_T$  and depends strongly on the symbol labeling [7]. For single-antenna BICM systems with Gray labeling, this gap has been shown to be negligible [2], [4]; however, for MIMO it can be considerably larger (see Section IV).

It can be shown that  $\Lambda_l$  in (2) is a sufficient statistic [30] for  $c_l$  conditioned on  $\mathbf{y}$  and  $\mathbf{H}$ .<sup>3</sup> Therefore, (6) can be rewritten as

$$C_{\text{BICM}} = \sum_{l=1}^{R_0} I(c_l; \Lambda_l) \tag{7}$$

Hence,  $C_{\text{BICM}}$  can be interpreted as the capacity of an equivalent channel with input  $c_l$  and output  $\Lambda_l$ . This channel is characterized by the conditional pdf  $f(\Lambda_l | c_l)$ , which usually is hard to obtain analytically, however.

### C. System Capacity and Demodulator Performance

Motivated by the interpretation of  $C_{\text{BICM}}$  as the system capacity of BICM using the optimum MAP demodulator, we propose to measure the performance of sub-optimal MIMO-BICM demodulators via the system capacity of the associated equivalent “modulation” channel with *discrete* input  $c_l$  and *continuous* output  $\tilde{\Lambda}_l$  (cf. Fig. 1). This channel is described by the conditional pdf  $f(\tilde{\Lambda}_l | c_l)$ . Its system capacity is

<sup>3</sup>Note, however, that  $\{\Lambda_1, \dots, \Lambda_{R_0}\}$  is *not* a sufficient statistic for  $\{c_1, \dots, c_{R_0}\}$ .

defined as the mutual information between  $c_l$  and  $\tilde{\Lambda}_l$ , which can be shown to equal

$$C \triangleq \sum_{l=1}^{R_0} I(c_l; \tilde{\Lambda}_l) \quad (8a)$$

$$= R_0 - \sum_{l=1}^{R_0} \sum_{b=0}^1 \int_{-\infty}^{\infty} \frac{1}{2} f(\tilde{\Lambda}_l | c_l = b) \log_2 \frac{f(\tilde{\Lambda}_l)}{\frac{1}{2} f(\tilde{\Lambda}_l | c_l = b)} d\tilde{\Lambda}_l, \quad (8b)$$

where  $f(\tilde{\Lambda}_l) = \frac{1}{2} \sum_{b=0}^1 f(\tilde{\Lambda}_l | c_l = b)$ . We emphasize that the system capacity  $C$  provides a performance measure for MIMO soft demodulators that is independent of the outer channel code. In fact, it has an intuitive operational interpretation as the highest rate achievable (in the sense of asymptotically vanishing error probability) in a BICM system (cf. [2, Section III.A]) using the specific demodulator which produces  $\tilde{\Lambda}_l$ . Since  $\tilde{\Lambda}_l$  is derived from  $\mathbf{y}$  and  $\mathbf{H}$ , the data processing inequality [30] implies that  $C \leq C_{\text{BICM}}$  with equality if  $\tilde{\Lambda}_l$  is a one-to-one function of  $\Lambda_l$ . The performance of a soft demodulator can thus be measured in terms of the gap  $C_{\text{BICM}} - C$ . Of course, the information theoretic performance measure in (8) does not take into account complexity issues and it has to be expected that a reduction of the gap  $C_{\text{BICM}} - C$  in general can only be achieved at the expense of increasing computational complexity.

We caution the reader that the rates in (7) and (8) are sums of mutual informations for the individual code bits  $c_1, \dots, c_{R_0}$  carried by one symbol vector. Indeed, the pdfs  $f(\Lambda_l | c_l)$  and  $f(\tilde{\Lambda}_l | c_l)$  in general depend on the code bit position  $l$ , even though for certain systems (e.g. 4-QAM modulation) the code bit protection and LLR statistics are independent of the bit position  $l$  for reasons of symmetry. Achieving (7) and (8) thus requires channel encoders and decoders that take the bit position into account. When the channel code fails to use this information, the rate loss is small provided that the mapping protects different code bits  $c_l$  roughly equally against noise and interference.

#### D. Non-ergodic channels

In the case of quasi-static or slow fading [31], the channel  $\mathbf{H}$  is random but constant over time, i.e., each codeword can extend over only one channel realization. In this regime, an ergodic system capacity of the equivalent modulation channel is no longer meaningful [31], [32]. Instead we consider the *outage probability*

$$P_{\text{out}}(R) \triangleq \mathbb{P}\{R_{\mathbf{H}} < R\}, \quad (9)$$

where  $R_{\mathbf{H}}$  is a random variable defined as

$$R_{\mathbf{H}} \triangleq \sum_{l=1}^{R_0} I_{\mathbf{H}}(c_l; \tilde{\Lambda}_l).$$

Here,  $I_{\mathbf{H}}(c_l; \tilde{\Lambda}_l)$  denotes conditional mutual information, which is evaluated with  $f(\tilde{\Lambda}_l|c_l, \mathbf{H})$  in place of  $f(\tilde{\Lambda}_l|c_l)$  (cf. (8)). Note that the ergodic system capacity  $C$  in (8) equals  $C = \mathbb{E}_{\mathbf{H}}\{R_{\mathbf{H}}\}$ . The outage probability  $P_{\text{out}}(R)$  can be interpreted as the smallest probability of error achievable at rate  $R$  [32]. A closely related concept is given by the  $\epsilon$ -capacity of the equivalent modulation channel, defined as

$$C_{\epsilon} \triangleq \sup \{R \mid \mathbf{P}\{R_{\mathbf{H}} < R\} < \epsilon\}. \quad (10)$$

The  $\epsilon$ -capacity may be interpreted as the maximum rate for which a probability of error less than  $\epsilon$  can be achieved. Rates smaller than  $C_{\epsilon}$  are referred to as  $\epsilon$ -achievable rates [32]. If  $P_{\text{out}}(R)$  is a continuous and increasing function of  $R$  (which is usually the case in practice), it holds that  $P_{\text{out}}(C_{\epsilon}) = \epsilon$ .

#### IV. BASELINE MIMO-BICM DEMODULATORS

In this section, we first review max-log and hard ML demodulation as well as linear MIMO demodulators and then we provide results illustrating their performance in terms of system capacities. These demodulators serve as baseline systems for later demodulator performance comparisons in Section V. We note that max-log and hard ML MIMO demodulators have the highest complexity among all soft and hard demodulation schemes, respectively, whereas linear MIMO demodulators are computationally the most efficient ones.

##### A. Max-Log and Hard ML Demodulator

Applying the max-log approximation to (2) simplifies the LLR computation to a minimum distance problem, resulting in the approximate LLRs [7]

$$\tilde{\Lambda}_l = \frac{1}{\sigma_v^2} \left[ \min_{\mathbf{x} \in \mathcal{X}_l^0} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 - \min_{\mathbf{x} \in \mathcal{X}_l^1} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \right]. \quad (11)$$

Eq. (11) facilitates (hardware) implementation since it avoids the logarithm and exponential functions in (2). However, computation of  $\tilde{\Lambda}_l$  in (11) still requires two searches over sets of size  $|\mathcal{A}|^{M_{\text{T}}/2} = 2^{R_0-1}$ . Efficient sphere decoder implementations of (11) are presented in [8], [10].

Hard vector ML demodulation can be performed by solving

$$\hat{\mathbf{x}}_{\text{ML}} = \arg \min_{\mathbf{x} \in \mathcal{A}^{M_{\text{T}}}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2. \quad (12)$$

The corresponding detected code bits  $\hat{c}_l$  are then obtained via the one-to-one mapping between code bits and symbol vectors, i.e.,  $\hat{\mathbf{c}} = (\hat{c}_1 \dots \hat{c}_{R_0})^{\text{T}} = \mu^{-1}(\hat{\mathbf{x}}_{\text{ML}})$ . It can be shown that the code bits  $\hat{c}_l$  obtained by the hard ML detector correspond to the sign of the corresponding max-log LLRs in (11), i.e.,  $\hat{c}_l = \text{u}(\tilde{\Lambda}_l)$  where  $\text{u}(\cdot)$  denotes the unit step function. When it comes to evaluating the system capacity

with hard-output demodulators, the only difference to soft-output demodulation is the *discrete* nature of the outputs  $\hat{c}_l$  of the equivalent “modulation” channel, which here becomes a binary symmetric channel (BSC). Consequently, the integral over  $\tilde{\Lambda}_l$  in (8b) is replaced with a summation over  $\hat{c}_l \in \{0, 1\}$ .

### B. Linear Demodulators

In the following,  $\Lambda_k^{(i)}$  is the LLR corresponding to  $c_k^{(i)}$  (the  $i$ th bit in the bit label of the  $k$ th symbol  $x_k$ ). Soft demodulators with extremely low complexity can be obtained by using a linear (ZF or MMSE) equalizer followed by *per-layer* max-log LLR calculation according to

$$\tilde{\Lambda}_k^{(i)} = \frac{1}{\sigma_k^2} \left[ \min_{x \in \mathcal{A}_i^0} |\hat{x}_k - x|^2 - \min_{x \in \mathcal{A}_i^1} |\hat{x}_k - x|^2 \right], \quad i = 1, \dots, Q, \quad k = 1, \dots, M_T. \quad (13)$$

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  $\sigma_k^2$  is an equalizer-specific weight (see below). We emphasize that calculating LLRs separately for each layer results in a significant complexity reduction. In fact, calculating the symbol estimates  $\hat{x}_k$  using a ZF or MMSE equalizer requires  $\mathcal{O}(M_R^3)$  operations; furthermore, the complexity of evaluating (13) for all code bits scales as  $\mathcal{O}(M_T 2^Q)$ , i.e., linearly in the number of transmit antennas.

Equalization-based hard bit decisions  $\hat{c}_k^{(i)}$  can be obtained by quantization of the equalizer output  $\hat{x}_k$  with respect to  $\mathcal{A}$  (denoted by  $\mathcal{Q}(\cdot)$ ), followed by the demapping, i.e.,  $(\hat{c}_k^{(1)} \dots \hat{c}_k^{(Q)})^T = \mu^{-1}(\mathcal{Q}(\hat{x}_k))$ . Again, the detected code bits correspond to the sign of the LLRs, i.e.,  $\hat{c}_k^{(i)} = \mathbf{u}(\tilde{\Lambda}_k^{(i)})$ .

1) *ZF-based Demodulator* [21], [22]: Here, the first stage consists of ZF equalization, i.e.,

$$\hat{\mathbf{x}}_{\text{ZF}} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{y} = \mathbf{x} + \tilde{\mathbf{v}}, \quad (14)$$

where the post-equalization noise vector  $\tilde{\mathbf{v}}$  has correlation matrix

$$\mathbf{R}_{\tilde{\mathbf{v}}} = \mathbb{E}\{\tilde{\mathbf{v}}\tilde{\mathbf{v}}^H\} = \sigma_v^2 (\mathbf{H}^H \mathbf{H})^{-1}. \quad (15)$$

Subsequently, approximate bit LLRs are obtained according to (13) with symbol estimate<sup>4</sup>  $\hat{x}_k = (\hat{\mathbf{x}}_{\text{ZF}})_k$  and weight factor  $\sigma_k^2 = (\mathbf{R}_{\tilde{\mathbf{v}}})_{k,k}$ .

<sup>4</sup>By  $(\mathbf{x})_k$  and  $(\mathbf{X})_{k,l}$  we respectively denote the  $k$ th element of the vector  $\mathbf{x}$  and the element in row  $k$  and column  $l$  of the matrix  $\mathbf{X}$ .

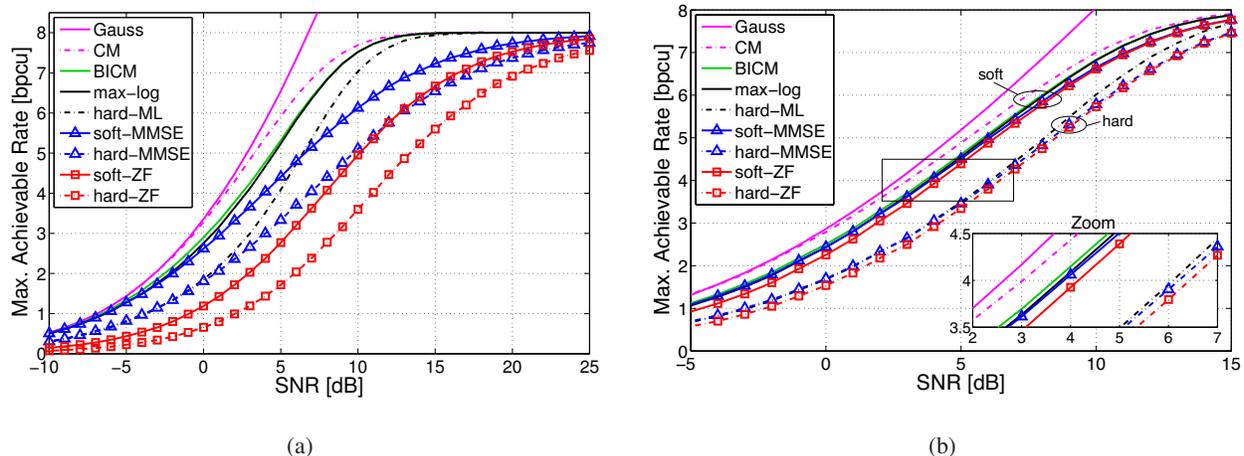


Fig. 2. Numerical capacity results for (a) a  $4 \times 4$  MIMO system with 4-QAM, and (b) a  $2 \times 4$  MIMO system with 16-QAM (in both cases with Gray labeling).

2) *MMSE-based Demodulator* [23]: Here, the first stage is an MMSE equalizer that can be written as (cf. (14) and (15))

$$\hat{\mathbf{x}}_{\text{MMSE}} = \mathbf{W} \hat{\mathbf{x}}_{\text{ZF}}, \quad \text{with } \mathbf{W} = (\mathbf{I} + \mathbf{R}_{\tilde{\mathbf{v}}})^{-1}. \quad (16)$$

Approximate LLRs are then calculated according to (13) with

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

where  $W_k = (\mathbf{W})_{k,k}$ . We emphasize that throughout this paper the results for soft (and hard) MMSE demodulation are based on the *unbiased* MMSE equalizer output  $\frac{(\hat{\mathbf{x}}_{\text{MMSE}})_k}{W_k}$ . In spite of similar complexity, MMSE-based demodulation outperforms ZF-based demodulation substantially [23].

### C. Capacity Results

We next compare the performance of the above baseline demodulators (i.e., max-log, hard ML, MMSE and ZF) in terms of their associated maximum achievable rate  $C$  (ergodic system capacity) in (8). In addition, the CM capacity  $C_{\text{CM}}$  in (4), the MIMO-BICM capacity  $C_{\text{BICM}}$  in (6), and the Gaussian input channel capacity  $C_G$  in (3) are shown as benchmarks. Throughout the paper, all capacity results have been obtained for spatially i.i.d. Rayleigh fading, with all fading coefficients normalized to unit variance.

The pdfs required for evaluating (8b) are generally hard to obtain in closed form. Thus, we measured these pdfs using Monte-Carlo simulations and then evaluated all integrals numerically. Based on the results in [33], we numerically optimized the binning (used to measure the pdfs) in order to reduce the

bias and variance of the mutual information estimates. The capacity results (obtained with  $10^5$  fading realizations) are shown in Fig. 2 in bits per channel use versus SNR  $\rho \triangleq E_s/\sigma_v^2$ . In the following, in some of the plots we show insets that provide zooms of the capacity curves around a target rate of  $R_0/2$  bpcu.

Fig. 2(a) pertains to the case of  $M_T \times M_R = 4 \times 4$  MIMO with Gray-labeled 4-QAM (here,  $R_0 = 8$ ). At a target rate of 4 bpcu, the SNR required for CM and Gaussian capacity is virtually the same, whereas that for BICM is larger by about 1.3 dB. The SNR penalty of using max-log demodulation instead of soft MAP is about 0.3 dB. Furthermore, hard ML demodulation requires 2.1 dB higher SNR to achieve this rate than max-log demodulation; for soft and hard MMSE demodulation the SNR gaps to max-log are 1.2 dB and 4.1 dB, respectively, while for soft and hard ZF demodulation they respectively equal 5.1 dB and 8.2 dB. An interesting observation in this scenario is the fact that at low rates, soft and hard MMSE demodulation essentially coincides with max-log and hard ML demodulation, respectively, whereas at high rates MMSE demodulation approaches ZF performance. Moreover, soft MMSE demodulation outperforms hard ML demodulation at low rates whereas at high rates it is the other way around. Surprisingly, at very low rates soft MMSE even performs slightly better than max-log demodulation. A similar cross-over of the capacity curves occurs with hard MMSE and soft ZF demodulation. These observations reveal the somewhat unexpected fact that the demodulator performance ranking is not universal but depends on the target rate (equivalently, the target SNR), even if the number of antennas, the symbol constellation, and the labeling are fixed. Similar observations apply to 16-QAM instead of 4-QAM and to set-partitioning labeling instead of Gray labeling (see [29]). Apart from a general shift of all curves to higher SNRs, the larger constellation and/or the different labeling strategy causes an increase of the gap between CM capacity and BICM capacity. When decreasing the antenna configuration to a  $2 \times 2$  system, we observed that soft ZF outperforms hard ML demodulation for low-to-medium rates, e.g., by about 1.7 dB at 4 bpcu with 16-QAM [29]. In this rate regime, even hard MMSE performs slightly better than hard ML demodulation.

The situation changes for the case of a  $2 \times 4$  MIMO system with Gray-labeled 16-QAM (again  $R_0 = 8$ ), shown in Fig. 2(b). 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 [7]). In addition, the performance differences between the individual demodulators are significantly reduced, revealing an essential distinction being between soft and hard demodulators. Having  $M_R > M_T$  helps the linear demodulators approach their non-linear counterparts even at larger rates, i.e., soft ZF/MMSE perform close to max-log and hard ZF/MMSE perform close to hard ML, with an SNR gap of about 2.3 dB between hard and soft demodulators. Note that in this scenario soft MMSE and soft

ZF both outperform hard ML demodulation at all rates.

#### D. BER Performance

Even though we advocate a demodulator comparison in terms of system capacity, the cross-over of some of the capacity curves prompts a verification in terms of the BER of soft and hard MMSE demodulation as well as max-log and hard ML demodulation. We consider a  $4 \times 4$  MIMO-BICM system with Gray-labeled 4-QAM in conjunction with irregular LDPC codes<sup>5</sup> [34] of block length 64000. For the case of soft demodulation, the LDPC codes were designed for an additive white Gaussian noise channel whereas for the case of hard demodulators the design was for a BSC. At the receiver, message-passing LDPC decoding [34] was performed. In the case of hard demodulation, the message-passing decoder was provided with the LLRs

$$\tilde{\Lambda}_l = (2\hat{c}_l - 1) \log \frac{1-p_0}{p_0}, \quad (17)$$

where  $p_0 = P\{\hat{c}_l \neq c_l\}$  is the cross-over probability of the equivalent BSC, which was determined via Monte-Carlo simulations. Using (17) is critical in order to provide the LDPC decoder with appropriate reliability information.

The BERs obtained are shown in Fig. 3(a) for a code rate of  $1/4$  (corresponding to 2 bpcu) and Fig. 3(b) for a code rates of  $3/4$  (6 bpcu). We also indicate by vertical lines the respective capacity limits, i.e., the minimum SNR required for the target rate according to Fig. 2(a). It is seen that all our LDPC code designs are less than 1 dB away from the capacity limits. For both rates and at all SNRs, max-log demodulation performs best and hard MMSE demodulation performs worst. No such universal statements can be made for hard ML and soft MMSE demodulation. Specifically, at rate  $1/4$  soft MMSE outperforms hard ML detector by 2.3 dB and is only 0.3 dB away from max-log (cf. Fig. 3(a)); however, at rate  $3/4$  soft MMSE performs 2 dB poorer than hard ML and 3.6 dB poorer than max-log (see Fig. 3(b)). These results confirm the capacity-based observation that there is no universal (i.e., rate- and SNR-independent) demodulator performance ranking. We note that the block error rate simulations in [35] allow for similar conclusions, even though not explicitly mentioned in that paper.

## V. OTHER DEMODULATORS

In the following, we study the system capacity of several other MIMO-BICM demodulators that differ in their underlying principle and their computational complexity. Unless stated otherwise, capacity results

<sup>5</sup>The LDPC code design was performed using the web-tool at <http://lthcwww.epfl.ch/research/ldpcopt>.

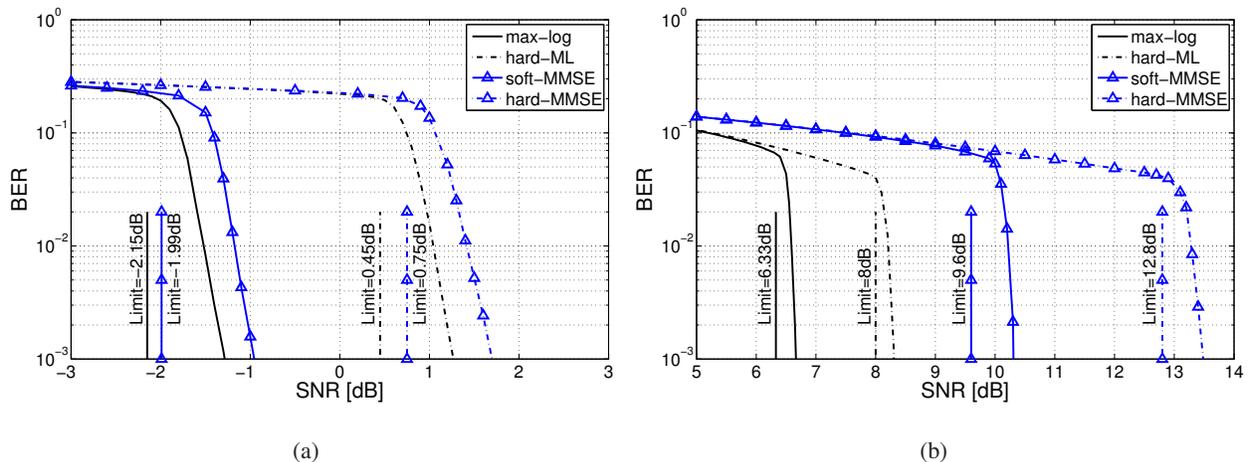


Fig. 3. BER vs. SNR for a  $4 \times 4$  MIMO system with Gray-labeled 4-QAM and LDPC codes with (a) rate  $1/4$  and (b) rate  $3/4$ .

in this section pertain to a  $4 \times 4$  MIMO system with 4-QAM using Gray labeling ( $R_0 = 8$ ). We note that the results for asymmetric  $2 \times 4$  MIMO systems with 16-QAM (shown in [29] but not here) essentially confirm the general distinction between hard and soft demodulators observed in connection with Fig. 2(b).

### A. List-based Demodulators

In order to save computational complexity, (11) can be approximated by decreasing the size of the search set  $\mathcal{A}^{M_T}$ . Usually, this is achieved by generating a (non-empty) candidate list  $\mathcal{L} \subset \mathcal{A}^{M_T}$  and restricting the search in (11) to this list, i.e.,

$$\tilde{\Lambda}_l = \frac{1}{\sigma_v^2} \left[ \min_{\mathbf{x} \in \mathcal{L} \cap \mathcal{X}_l^0} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 - \min_{\mathbf{x} \in \mathcal{L} \cap \mathcal{X}_l^1} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \right]. \quad (18)$$

The computational complexity of the metric evaluations and minima searches in (18) scales as  $\mathcal{O}(M_T M_R |\mathcal{L}|)$ . Thus, the list size  $|\mathcal{L}|$  allows to trade off performance for complexity savings. A larger list size generally incurs higher complexity but yields more accurate approximations of the max-log LLRs. For fixed list size, the performance can further strongly depend on which symbol vectors are actually included in  $\mathcal{L}$ . In the following, we consider two types of list-based demodulation where the candidate list is obtained using sphere decoding and bit flipping, respectively.

1) *List Sphere Decoder (LSD)*: The LSD proposed in [14] uses a simple modification of the hard-decision sphere decoder [36] to generate the candidate list  $\mathcal{L}$  such that it contains the  $|\mathcal{L}|$  symbol vectors  $\mathbf{x}$  with the smallest ML metric  $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$  (thus, by definition  $\mathcal{L}$  contains the hard ML solution  $\hat{\mathbf{x}}_{ML}$  in

(12)). If the  $l$ th bit in the labels of *all*  $\mathbf{x} \in \mathcal{L}$  equals 1, the set  $\mathcal{L} \cap \mathcal{X}_l^0$  is empty and (18) cannot be evaluated. Since in this case there is strong evidence for  $c_l = 1$  (at least if  $|\mathcal{L}|$  is not too small), the corresponding LLR  $\tilde{\Lambda}_l$  is set to a prescribed positive value  $\check{\Lambda} \gg 0$ . Analogously,  $\tilde{\Lambda}_l = -\check{\Lambda}$  in case  $\mathcal{L} \cap \mathcal{X}_l^1$  equals the empty set.

While the LSD may offer significant complexity savings compared to max-log demodulation, statements about its computational complexity are difficult and depend strongly on the actual implementation of the sphere decoder as well as the choice of the list size. We note that the case  $|\mathcal{L}| = 2^{R_0} = |\mathcal{A}^{M_T}|$  implies  $\mathcal{L} = \mathcal{A}^{M_T}$ ; thus,  $\mathcal{L} \cap \mathcal{X}_l^b = \mathcal{X}_l^b$  such that (18) equals the max-log demodulator in (11). The other extreme is a list size of one, i.e.,  $\mathcal{L} = \{\hat{\mathbf{x}}_{\text{ML}}\}$  (cf. (12)), in which case either  $\mathcal{L} \cap \mathcal{X}_l^0$  or  $\mathcal{L} \cap \mathcal{X}_l^1$  is empty (depending on the bit label of  $\hat{\mathbf{x}}_{\text{ML}}$ ); here,  $\tilde{\Lambda}_l = (2\hat{c}_l - 1)\check{\Lambda}$  where  $\hat{\mathbf{c}} = (\hat{c}_1 \dots \hat{c}_{R_0})^T = \mu^{-1}(\hat{\mathbf{x}}_{\text{ML}})$  and thus the LSD output is equivalent to hard ML demodulation (except for the choice of  $\check{\Lambda}$ , which is irrelevant, however, for capacity).

*Capacity Results.* Fig. 4 shows the maximum rates achievable with an LSD for various list sizes. BICM and soft MMSE capacity are shown for comparison. Note that with 4-QAM and  $M_T = 4$ ,  $|\mathcal{L}| = 256$  and  $|\mathcal{L}| = 1$  correspond to max-log and hard ML demodulation, respectively. It is seen that with increasing list size the gap between LSD and max-log decreases rapidly, specifically at high rates. In particular, the LSD with list sizes of  $|\mathcal{L}| \geq 8$  is already quite close to max-log performance. However, even with large list sizes LSD is outperformed by soft MMSE demodulation at sufficiently low rates. Specifically, below 3.6 dB, 1.5 dB, and  $-0.2$  dB the system capacity of soft MMSE demodulation is higher than that of LSD with list size 2, 4, and 8, respectively. Similar observations apply to other antenna configurations and symbol constellations (see [29]).

2) *Bit Flipping Demodulators:* Another way of generating the candidate list  $\mathcal{L}$ , proposed in [20], is to flip some of the bits in the label of the hard ML symbol vector estimate  $\hat{\mathbf{x}}_{\text{ML}}$  in (12). More generally, the ML solution  $\hat{\mathbf{x}}_{\text{ML}}$  can be replaced by a symbol vector  $\hat{\mathbf{x}} \in \mathcal{A}^{M_T}$  obtained with an arbitrary hard-output demodulator (e.g., hard ZF and MMSE demodulation). Let  $\hat{\mathbf{c}} = \mu^{-1}(\hat{\mathbf{x}})$  denote the bit label of  $\hat{\mathbf{x}}$ . The candidate list then consists of all symbol vectors whose bit label has Hamming distance at most  $D \leq R_0$  from  $\hat{\mathbf{c}}$ , i.e.,  $\mathcal{L} = \{\mathbf{x}: d_{\text{H}}(\mu^{-1}(\mathbf{x}), \hat{\mathbf{c}}) \leq D\}$ . Here,  $d_{\text{H}}(\mathbf{c}_1, \mathbf{c}_2)$  denotes the Hamming distance between two bit labels  $\mathbf{c}_1$  and  $\mathbf{c}_2$ . This list can be generated by systematically flipping up to  $D$  bits in  $\hat{\mathbf{c}}$  and mapping the results to symbol vectors. The resulting list size is given by  $|\mathcal{L}| = \sum_{d=0}^D \binom{R_0}{d}$ . Here, the structure of the list generated with bit flipping allows to reduce the complexity per candidate to  $\mathcal{O}(M_R)$  giving an overall complexity of  $\mathcal{O}(M_R|\mathcal{L}|)$  (plus the operations required for the initial estimate). For  $D = R_0$ ,

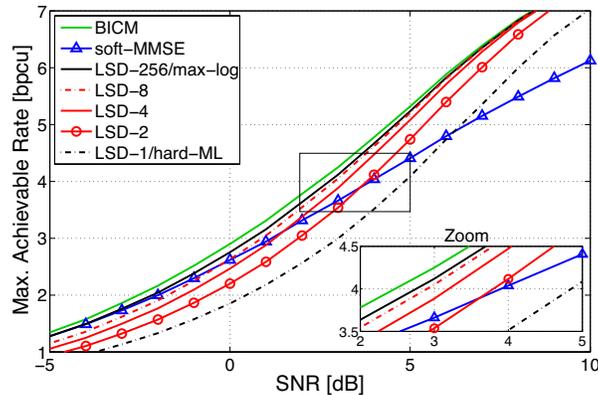


Fig. 4. System capacity of LSD with list size  $|\mathcal{L}| \in \{1, 2, 4, 8, 256\}$  ( $4 \times 4$  MIMO, 4-QAM, Gray labeling).

$\mathcal{L} = \mathcal{A}^{M_T}$  and (18) reduces to max-log demodulation; furthermore, with  $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{ML}$  and  $D = 0$  there is  $\mathcal{L} = \{\hat{\mathbf{x}}_{ML}\}$  and (18) becomes equivalent to hard ML demodulation. Note that in contrast to the LSD, bit flipping with  $D > 0$  ensures that  $\mathcal{L} \cap \mathcal{X}_l^0$  and  $\mathcal{L} \cap \mathcal{X}_l^1$  are nonempty such that (18) can always be evaluated.

*Capacity Results.* Fig. 5 shows the maximum rates achievable with bit flipping demodulation where the initial symbol vector estimate is chosen either as the hard ML solution  $\hat{\mathbf{x}}_{ML}$  in (12) or the hard MMSE estimate  $\mathcal{Q}(\hat{\mathbf{x}}_{MMSE})$  (cf. (16)). For  $D = 1$  ( $|\mathcal{L}| = 9$ ), Fig. 5(a) reveals that flipping 1 bit (labeled 'flip-1') allows for significant performance improvements over the respective initial hard demodulator (about 2.5 dB at 2 bpsu). For rates below 4 bpsu, hard ML and hard MMSE initialization yield effectively identical results, close to soft MMSE demodulation. At higher rates MMSE-based bit flipping even outperforms soft MMSE demodulation slightly. For  $D = 2$  ( $|\mathcal{L}| = 37$ ), it can be seen from Fig. 5(b) that bit flipping demodulation performs close to max-log at low rates and that hard ML and hard MMSE initialization are very close to each other for rates up to 6 bpsu (SNR up to 7 dB); in fact, below 5 bpsu hard MMSE initialization performs slightly better than hard ML initialization while at higher rates ML initialization gives significant gains. To maintain this behavior for larger constellations and more antennas, the maximum Hamming distance  $D$  has to increase with increasing  $R_0$  (see [29]).

### B. Lattice-Reduction-Aided Demodulation

Lattice reduction (LR) is an important technique for improving the performance or complexity of MIMO demodulators [15], [16] for the case of QAM symbol constellations. The basic underlying idea is to view the columns of the channel matrix  $\mathbf{H}$  as basis vectors of a point lattice. LR then yields

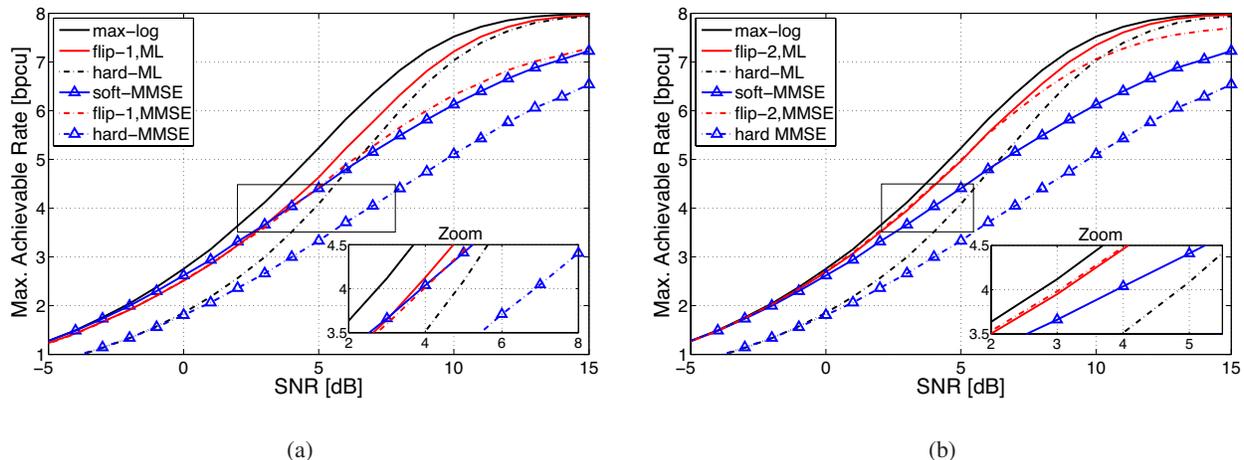


Fig. 5. System capacity of bit flipping demodulator with (a)  $D = 1$  and (b)  $D = 2$  ( $4 \times 4$  MIMO, 4-QAM, Gray labeling).

an alternative basis which amounts to a transformation of the system model (1) prior to demodulation; the advantage of such an approach is that the transformed channel matrix (i.e., the reduced basis) has improved properties (e.g., smaller condition number). An efficient algorithm to obtain a reduced basis was proposed by Lenstra, Lenstra, and Lovász (LLL) [37]. The overall computational complexity of LR-aided demodulation depends strongly on the complexity of the LLL algorithm which has been assessed in [38]. A comparison of different LR methods in the context of MIMO hard demodulation was provided in [39].

LR is usually formulated for the equivalent real-valued model; hence, for now we assume all quantities to be real-valued. Any lattice basis transformation can be formulated in terms of a unimodular transformation matrix  $\mathbf{T}$ , i.e., a matrix with integer entries and  $\det(\mathbf{T}) = \pm 1$ . Denoting the “reduced channel” by  $\tilde{\mathbf{H}} = \mathbf{H}\mathbf{T}$  and defining  $\mathbf{z} = \mathbf{T}^{-1}\mathbf{x}$ , the system model (1) can be rewritten as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} = \tilde{\mathbf{H}}\mathbf{z} + \mathbf{v}. \quad (19)$$

Under the assumption  $\mathbf{x} \in \mathbb{Z}^{M_T}$  (which for QAM can be ensured by an appropriate offset and scaling), the unimodularity of  $\mathbf{T}$  guarantees  $\mathbf{z} \in \mathbb{Z}^{M_T}$  and hence any demodulator can be applied to the better-behaved transformed system model on the right-hand side of (19). Here, we adopt the LR-aided hard-output MMSE demodulator developed in [16]. LR-aided soft demodulators (cf. [17]) are essentially list-based [18], [19]. In fact, many of these methods apply bit flipping (cf. Section V-A2) to an LR-aided hard demodulator output.

*Capacity Results.* Fig. 6 shows the capacity results for hard and soft LR-aided MMSE demodulation. Soft outputs are obtained by applying bit flipping with  $D = 1$  and  $D = 2$  to the LR-aided hard MMSE

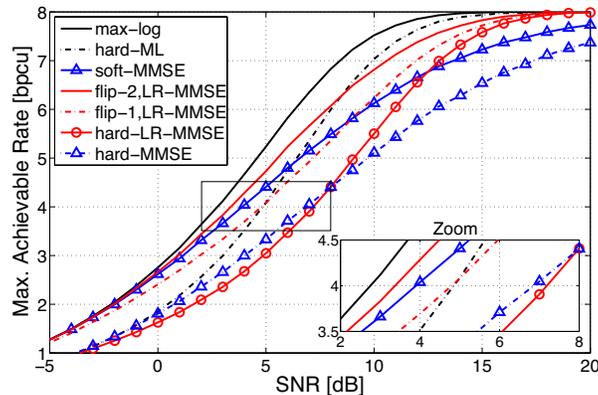


Fig. 6. System capacity of LR-aided hard and soft MMSE demodulation ( $4 \times 4$  MIMO, 4-QAM, Gray labeling).

demodulator output (cf. Section V-A2). It is seen that with hard MMSE demodulation, LR is beneficial only at SNRs above 8 dB (rates higher than 4.4 bpsu). At rates higher than about 6.8 bpsu, LR-aided hard demodulation even outperforms soft MMSE demodulation. Bit flipping is helpful particularly at low rates. Thus, at low rates LR-aided soft MMSE demodulation approaches max-log performance and outperforms hard ML. When flipping up to  $D = 2$  bits, LR-aided soft demodulation reveals a significant performance advantage over soft MMSE demodulation without LR which is especially pronounced in the high-rate regime.

### C. Semidefinite Relaxation (SDR) Demodulation

Semidefinite relaxation allows to approximately solve the hard ML demodulation problem in (12) [40], [41] with polynomial worst-case complexity based on convex optimization techniques. In this paper, we focus on the quasi-ML hard-output demodulator and its soft extension proposed in [13] that approximates max-log demodulation (with an overall worst-case complexity of  $\mathcal{O}(R_0^{4.5})$ ). We note that this approach applies only to BPSK or 4-QAM alphabets and employs a randomization procedure described in detail in [41].

*Capacity Results.* In Fig. 7 we show the system capacity for hard and soft SDR demodulation (as described in [13]) using randomization with 25 trials. It can be seen that hard SDR demodulation coincides with hard ML at low rates; at high rates it performs worse than hard ML but still better than soft MMSE. Soft SDR performs consistently better than hard SDR; while at low rates it coincides with soft MMSE demodulation, at rates higher than 6.2 bpsu its capacity lies between hard ML and soft MMSE.

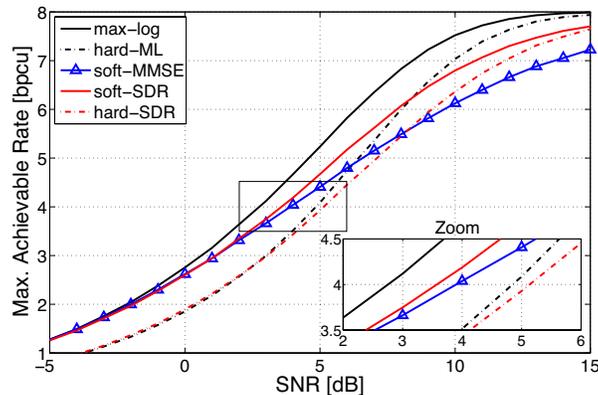


Fig. 7. System capacity of hard and soft SDR demodulation ( $4 \times 4$  MIMO, 4-QAM, Gray labeling).

#### D. $\ell^\infty$ -Norm Demodulator

It was shown in [11], [12] that the VLSI implementation complexity of hard ML demodulation (cf. (12)) can be significantly reduced by replacing the  $\ell^2$  norm in the sphere decoder with the  $\ell^\infty$  norm  $\|\mathbf{a}\|_\infty = \max\{\text{Re}\{a_1\}, \dots, \text{Re}\{a_M\}, \text{Im}\{a_1\}, \dots, \text{Im}\{a_M\}\}$  (here,  $\mathbf{a}$  is a complex-valued vector of length  $M$ ). Specifically, the use of the  $\ell^\infty$  norm avoids expensive squaring operations. The hard  $\ell^\infty$ -norm demodulator delivers the symbol vector estimate

$$\hat{\mathbf{x}}_{\ell^\infty} = \arg \min_{\mathbf{x} \in \mathcal{A}^{M_T}} \|\mathbf{Q}^H \mathbf{y} - \mathbf{R}\mathbf{x}\|_\infty, \quad (20)$$

Here,  $\mathbf{Q}$  and  $\mathbf{R}$  are the  $M_R \times M_T$  unitary and  $M_T \times M_T$  upper triangular factors in the QR decomposition  $\mathbf{H} = \mathbf{Q}\mathbf{R}$  of the channel matrix. Note that the solution to (20) may not be unique, in which case one solution is selected at random.

We propose to generate soft outputs by using the  $\ell^\infty$ -norm sphere decoder to determine

$$\tilde{\mathbf{x}}_l^b = \arg \min_{\mathbf{x} \in \mathcal{X}_l^b} \|\mathbf{Q}^H \mathbf{y} - \mathbf{R}\mathbf{x}\|_\infty$$

for  $b \in \{0, 1\}$  and then evaluating the approximate LLRs using the  $\ell^2$  norm:

$$\tilde{\Lambda}_l = \frac{1}{\sigma_v^2} \left[ \|\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}_l^0\|^2 - \|\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}_l^1\|^2 \right].$$

*Capacity Results.* Fig. 8 shows the system capacity for hard and soft  $\ell^\infty$ -norm demodulation. For the  $4 \times 4$  case with 4-QAM in Fig. 8(a), hard and soft  $\ell^\infty$ -norm demodulation perform within 1 dB of hard ML and max-log, respectively, particularly for high rates. However, at low rates (below 4 bpsu)  $\ell^\infty$ -norm demodulation is even outperformed by its corresponding MMSE counterpart. An interesting observation

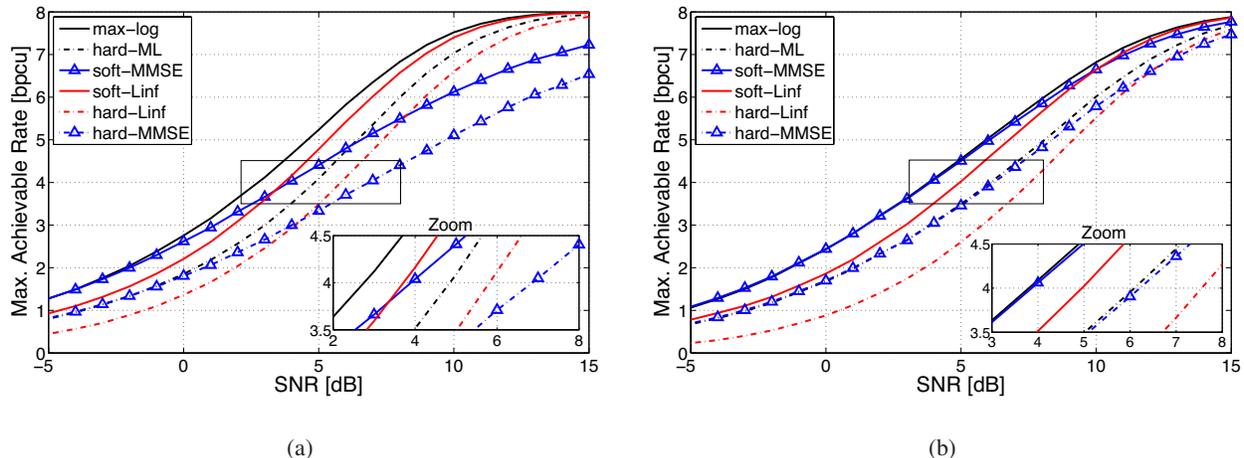


Fig. 8. System capacity of hard and soft  $\ell^\infty$ -norm demodulation for (a)  $4 \times 4$  MIMO with 4-QAM and (b)  $2 \times 4$  MIMO with 16-QAM (Gray labeling in both cases).

applies to the  $2 \times 4$  case with 16-QAM depicted in Fig. 8(b). All soft-output baseline demodulators perform almost identical and the same is true for all hard-output baseline demodulators, i.e., there is only a distinction between soft and hard demodulation (cf. Fig. 2(b)). However, soft and hard  $\ell^\infty$ -norm demodulation perform significantly worse in this asymmetric setup, specifically at low to medium rates. At 2 bpsu, soft  $\ell^\infty$ -norm demodulation requires 1.75 dB higher SNR than max-log and soft MMSE and hard  $\ell^\infty$ -norm demodulation requires 2.6 dB higher SNR than hard ML/MMSE.

### E. Successive and Soft Interference Cancellation

Successive interference cancellation (SIC) is a hard-output demodulation approach that became popular with the V-BLAST (*Vertical Bell Labs Layered Space-Time*) system [26]. Within one SIC iteration, only that layer with the largest post-equalization SNR is detected and its contribution to the receive signal is subtracted (canceled). A SIC implementation that replaces the ZF-based algorithm from [26] with an MMSE-based scheme and performs efficient layer ordering according to signal-to-interference-plus-noise-ratio (SINR) was presented in [27]. Suboptimal but more efficient SIC schemes are discussed in [28].

In order to mitigate the error propagation inherent to SIC, [24] proposed a parallel soft interference cancellation (SoftIC) scheme. SoftIC is an iterative method that alternately performs (i) parallel MIMO interference cancellation based on soft symbols and (ii) computation of improved soft symbols using the output of the interference cancellation stage. Each iteration of this scheme has a complexity that scales linearly with the number of antennas. Here, we use a modification that builds upon bit-LLRs. Let  $\tilde{\Lambda}_k^{(i)}[j]$

denote the LLR for the  $i$ th bit in layer  $k$  obtained in the  $j$ th iteration. Symbol probabilities can then be obtained as

$$P_k^{(j)}(x) = \prod_{i=1}^Q \frac{\exp\left(b_i(x)\tilde{\Lambda}_k^{(i)}[j]\right)}{1 + \exp\left(\tilde{\Lambda}_k^{(i)}[j]\right)},$$

with  $b_i(x)$  denoting the  $i$ th bit in the label of  $x \in \mathcal{A}$ , leading to the soft symbol estimate

$$\hat{x}_k^{(j)} = \sum_{x \in \mathcal{A}} x P_k^{(j)}(x).$$

Soft interference cancelation for each layer then yields

$$\mathbf{y}_k^{(j)} = \mathbf{y} - \sum_{k' \neq k} \mathbf{h}_{k'} \hat{x}_{k'}^{(j)} = \mathbf{h}_k x_k + \sum_{k' \neq k} \mathbf{h}_{k'} (x_{k'} - \hat{x}_{k'}^{(j)}) + \mathbf{v}, \quad (21)$$

where  $\mathbf{h}_k$  denotes the  $k$ th column of  $\mathbf{H}$ . Finally, updated LLRs  $\tilde{\Lambda}_k^{(i)}[j+1]$  are calculated from (21) based on a Gaussian assumption for the residual interference plus noise (for implementation details we refer to [24]). In contrast to [24], we suggest to initialize the scheme with the LLRs obtained by a low-complexity soft demodulator, e.g., the soft ZF detector described in Section IV-B. The complexity per iteration of the above SoftIC algorithm is given by  $\mathcal{O}(2^Q M_T(Q + M_R))$  (plus the operations required for the initial LLRs).

*Capacity Results.* In Fig. 9, we display capacity results for (hard) MMSE-SIC with detection ordering as in [28] (there referred to as ‘MMSE-BLAST’) and for SoftIC demodulation (initialized using soft ZF demodulation, also shown for reference). Hard MMSE-SIC demodulation is seen to perform similarly to hard ML demodulation at low rates. While at high rates MMSE-SIC shows a noticeable gap to hard ML, it can outperform both soft MMSE and SoftIC in this regime. When using 16-QAM instead of 4-QAM, at very low rates MMSE-SIC tends to even slightly outperform hard ML demodulation (see [29]). SoftIC (with 8 iterations) is superior to MMSE-SIC up to rates of 6.7 bpcu. At low rates, SoftIC even performs slightly better than max-log demodulation and essentially coincides with BICM capacity. For the chosen system parameters, SoftIC furthermore beats soft MMSE over the whole SNR range shown. This statement does not hold in general, however. For example, with 16-QAM the SoftIC performance drops below that of soft MMSE demodulation at high rates (see [29]).

## VI. IMPERFECT CHANNEL STATE INFORMATION

We next investigate the ergodic system capacity (8) for the case of imperfect channel state information (CSI). In particular, we consider training-based estimation of the channel matrix  $\mathbf{H}$  and the noise variance  $\sigma_v^2$  and assess how the amount of training influences the performance of the various demodulators.

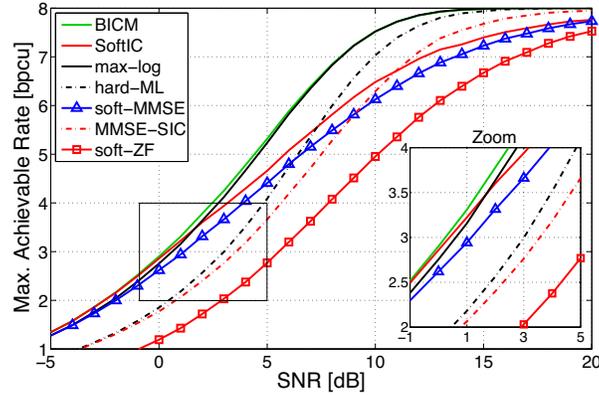


Fig. 9. System capacity of MMSE-SIC and SoftIC ( $4 \times 4$  MIMO, 4-QAM, Gray labeling).

*Training-based Channel Estimation.* In order to estimate the channel, the transmitter sends  $N_p > M_T$  training vectors<sup>6</sup> which are arranged into a full rank  $M_T \times N_p$  training matrix  $\mathbf{X}_p$ . We assume that the transmit power per channel use for training and actual data is the same such that the Frobenius norm [42] of  $\mathbf{X}_p$  equals  $\|\mathbf{X}_p\|_F^2 = N_p E_s$ .

Assuming that the channel stays constant for the duration of one block (which contains training and actual data), the  $M_R \times N_p$  receive matrix  $\mathbf{Y}_p$  induced by the training is given by

$$\mathbf{Y}_p = \mathbf{H} \mathbf{X}_p + \mathbf{V}. \quad (22)$$

Here, the  $M_R \times N_p$  matrix  $\mathbf{V}$  contains the noise received during the training period.

Using (22), the least-squares channel estimate (identical to the ML estimate under a Gaussian i.i.d. assumption for the noise) is computed as [43]

$$\hat{\mathbf{H}} = \mathbf{Y}_p \mathbf{X}_p^H (\mathbf{X}_p \mathbf{X}_p^H)^{-1}. \quad (23)$$

This estimate is unbiased and its mean square error equals

$$\mathbb{E}\{\|\hat{\mathbf{H}} - \mathbf{H}\|_F^2\} = M_R \sigma_v^2 \text{tr}\{(\mathbf{X}_p \mathbf{X}_p^H)^{-1}\} \geq \frac{M_R M_T^2}{N_p} \frac{1}{\rho},$$

where the lower bound is attained with orthogonal training sequences, i.e.,  $\mathbf{X}_p \mathbf{X}_p^H = \frac{N_p E_s}{M_T} \mathbf{I}$  (we recall that here  $\rho = E_s / \sigma_v^2$  denotes the SNR).

The estimated channel matrix  $\hat{\mathbf{H}}$  is then used to obtain the noise variance estimate

$$\hat{\sigma}_v^2 = \frac{1}{M_R(N_p - M_T)} \|\mathbf{Y}_p - \hat{\mathbf{H}} \mathbf{X}_p\|_F^2, \quad (24)$$

<sup>6</sup>While  $N_p = M_T$  is sufficient to estimate  $\mathbf{H}$ , extra training is required for estimation of  $\sigma_v^2$ .

which essentially amounts to measuring the mean power of  $\mathbf{Y}_p$  in the  $(N_p - M_T)$ -dimensional orthogonal complement of the range space of  $\mathbf{X}_p^H$ . The noise variance estimate is also unbiased and has MSE

$$\mathbb{E}\{|\hat{\sigma}_v^2 - \sigma_v^2|^2\} = \frac{\sigma_v^4}{M_R(N_p - M_T)},$$

which should be noted to be independent of the transmit power.

The estimates  $\hat{\mathbf{H}}$  in (23) and  $\hat{\sigma}_v^2$  in (24) then replace the true channel matrix and noise variance in the computation of the demodulator outputs, i.e., we consider *mismatched* demodulation. One could also change the metric in order to take into account the channel imperfections as in [44]; however, this is beyond the scope of this paper.

*Capacity Results.* The section provides numerical results for the ergodic system capacity of max-log, hard ML, and soft MMSE demodulation with perfect and estimated CSI (using orthogonal training sequences). Throughout, a  $4 \times 4$  MIMO system with 4-QAM and Gray labeling is considered ( $R_0 = 8$ ). Results for other demodulators with imperfect CSI are provided in [29].

Fig. 10 shows the maximum achievable rates versus SNR for a fixed training sequence length of  $N_p = 5$  (this corresponds to the worst case with minimum amount of training). It is seen that for all three detectors imperfect CSI results in a significant performance loss, e.g., at 4 bpcu the SNR loss for max-log, hard ML, and soft MMSE is 3.9 dB, 3.1 dB, and 4.95 dB, respectively. In fact, in the considered worst case imperfect CSI setup the performance advantage of soft MMSE demodulation over hard ML demodulation at low rates is much less pronounced; note that the cross-over between hard ML and soft MMSE performance at an SNR of about 6.3 dB shifts from 4.9 bpcu for perfect CSI to 3 bpcu for the case of imperfect CSI. The performance losses for all demodulators tend to be smaller at high rates, which may be partly attributed to the fact that CSI becomes more accurate with increasing SNR whereas this is not the case for the noise variance estimate. In general it can be observed that the performance loss of hard ML is the smallest while soft MMSE and max-log performance deteriorates much stronger; note that hard ML does not use the noise variance and hence is more robust to estimation errors in  $\sigma_v^2$ . Here, hard ML comes within less than 1.3 dB of max-log performance. In contrast, soft MMSE uses the imperfect channel and the noise variance estimate in the MMSE equalization stage (cf. (16)) and in the LLR calculation and is thus most strongly affected.

To investigate the impact of the amount of training, Fig. 11(a) and (b) depict the minimum SNRs required by the individual demodulators to achieve target rates of 2 bpcu and 6 bpcu, respectively, versus the number of training vectors  $N_p$ . It is seen that for all demodulators, the required SNR decreases rapidly with increasing amount of training. Yet, even for  $N_p = 20$  there is a significant gap of 1 to 2 dB to perfect

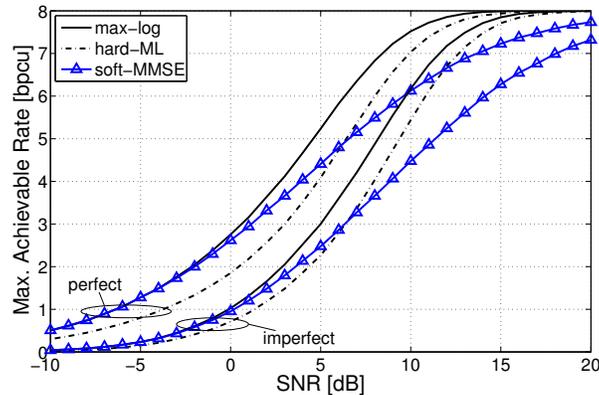


Fig. 10. System capacity of baseline demodulators with perfect and imperfect CSI ( $4 \times 4$  MIMO, 4-QAM, Gray labeling,  $N_p = 5$  training vectors).

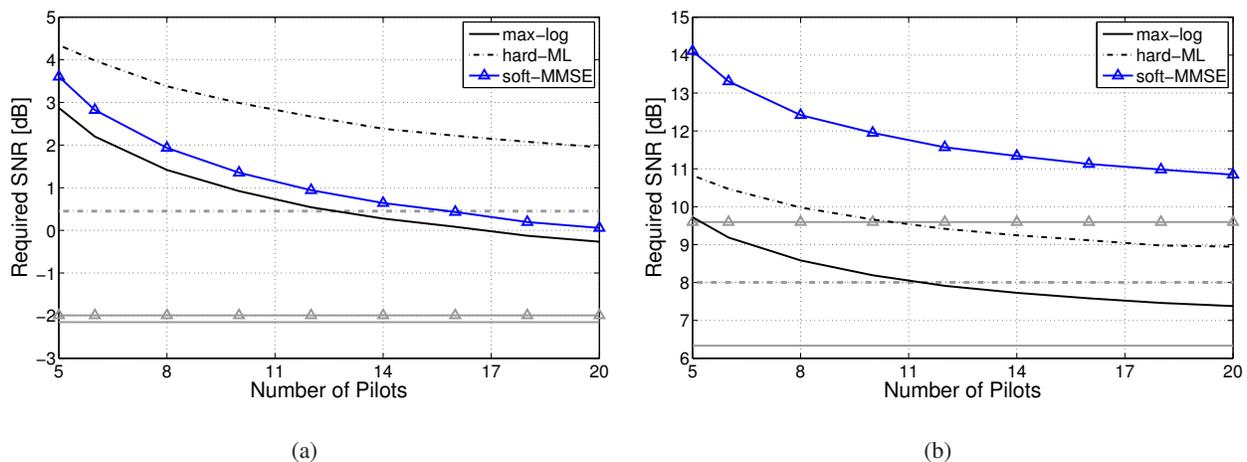


Fig. 11. SNR required by baseline demodulators versus number of pilots  $N_p$  for a target rate of (a)  $C = 2$  bpsu and (b)  $C = 6$  bpsu ( $4 \times 4$  MIMO, 4-QAM,  $R_0 = 8$ , Gray labeling).

CSI performance (indicated by corresponding horizontal, gray lines). Here, soft MMSE outperforms hard ML at 2 bpsu and gets closer to max-log performance with a larger amount of training. In contrast, at 6 bpsu hard ML performs consistently better than soft MMSE.

We note that further capacity results for the demodulators discussed in Section V are provided in [29]. One of the more interesting observations obtained from those results is that for training duration  $N_p = 5$ , the LSD with list size  $|\mathcal{L}| \geq 8$  consistently outperforms max-log (at least in the MIMO setup considered). Although this is not the case for larger training duration, LSD with  $|\mathcal{L}| = 8$  performs mostly very close to max-log.

## VII. QUASI-STATIC FADING

In this section we provide a demodulator performance comparison for quasi-static fading MIMO channels based on the outage probability  $P_{\text{out}}(R)$  in (9) and the  $\epsilon$ -capacity  $C_\epsilon$  in (10). The setup considered ( $4 \times 4$  MIMO with Gray-labeled 4-QAM) is the same as before apart from the spatially i.i.d. Rayleigh fading channel which now is assumed to be quasi-static. The outage probability  $P_{\text{out}}(R)$  was measured using  $10^5$  blocks (affected by independent fading realizations), each consisting of  $10^5$  symbol vectors. To keep the presentation concise, we restrict to the baseline demodulators from Section IV.

Fig. 12(a) shows the outage probability versus SNR  $\rho$  for target rates of  $R=2$  bpcu and  $R=6$  bpcu. For  $R=2$  bpcu, it is seen that the gap between optimum soft MAP demodulation (labeled ‘BICM’ for consistency with previous sections) and max-log demodulation equals about 0.5 dB. In such a low-rate regime, soft MMSE performs as well as max-log and about 2.5 dB better than hard ML. While max-log, hard ML, and soft MMSE demodulation all achieve full diversity (cf. the slope of the corresponding outage probability curves), soft ZF only has diversity order one, resulting in a huge performance loss (almost 19 dB at  $P_{\text{out}}(R) = 10^{-2}$ ). At  $R=6$  bpcu the situation is quite different: here, max-log coincides with soft-MAP and hard-ML loses only 1.4 dB (again, those three demodulators achieve full diversity). Surprisingly, soft MMSE breaks down significantly at this rate, losing all diversity. At  $P_{\text{out}}(R) = 10^{-2}$ , the SNR loss of soft MMSE and soft ZF relative to max-log equals about 11 dB and 19 dB, respectively.

The degradation of soft MMSE with increasing rate is clearly visible in Fig. 12(b), which shows  $\epsilon$ -capacity versus SNR for a maximum outage probability of  $\epsilon = 10^{-2}$ . Some of the behavior of the  $\epsilon$ -capacity curves is qualitatively very similar to the results obtained for the ergodic capacity (cf. Fig. 2(a)): soft MMSE demodulation outperforms hard ML demodulation (by up to 2.2 dB for rates less than 3.8 bpcu) while at high rates it is the other way round. Furthermore, for low rates soft MMSE demodulation essentially coincides with max-log whereas at high rates it approaches soft ZF performance.

Comparison of Fig. 12(b) and Fig. 2(a) even suggests that there is a connection between the diversity of the demodulators in the quasi-static scenario and their SNR gap to optimum demodulation in the ergodic scenario. For all rates (SNRs), the max-log and hard ML demodulator both achieve constant (full) diversity in the quasi-static regime and maintain a roughly constant gap to soft MAP in the ergodic scenario. On the other hand, the diversity of soft MMSE demodulation in the quasi-static case and its SNR gap to soft MAP in the ergodic scenario both deteriorate with increasing rate/SNR. Note that also here soft ZF demodulation performs worst, maintaining a diversity equal to 1 for all rates.

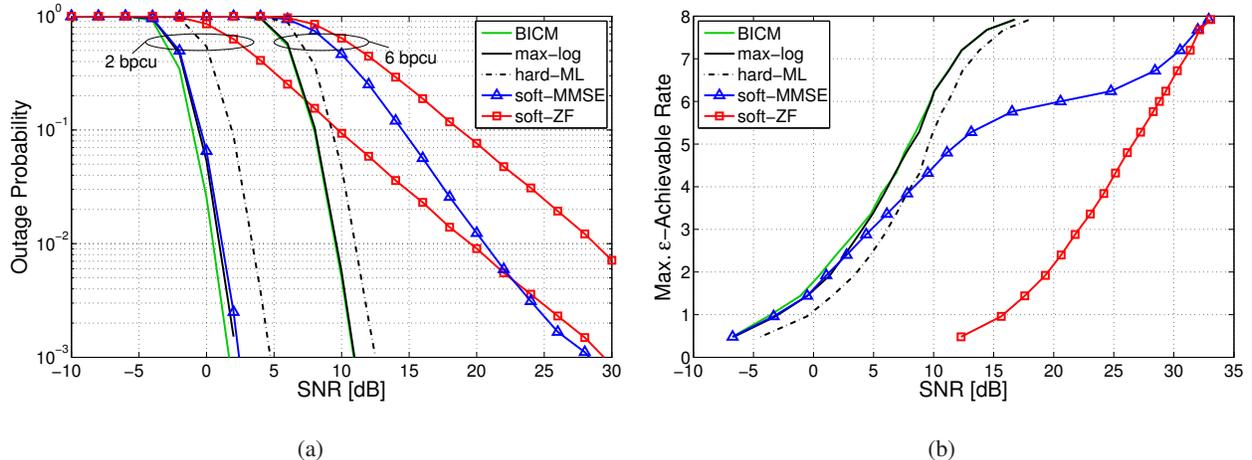


Fig. 12. Demodulator performance in quasi-static fading: (a) outage probability versus SNR for  $R=2$  bpcu and  $R=6$  bpcu, and (b)  $\epsilon$ -capacity versus SNR for  $\epsilon=10^{-2}$  ( $4 \times 4$  MIMO, 4-QAM, Gray labeling).

### VIII. KEY OBSERVATIONS AND DESIGN GUIDELINES

Based on the results of the previous sections, we summarize some key observations and provide practical guidelines for system design.

Soft MMSE demodulation was seen to approach max-log performance for low rates, both in the ergodic and the quasi-static regime and for various system configurations (see also [29]). Moreover, soft MMSE demodulation is very attractive since it has the lowest complexity among all the suboptimal demodulators discussed in this work. Therefore, soft MMSE demodulation is arguably the demodulator of choice when designing MIMO-BICM systems with outer codes of rate less than  $1/2$ . For the case of imperfect CSI, however, care should be taken to ensure a sufficient amount of training since otherwise soft MMSE may suffer severely from inaccurate estimates of the channel matrix and the noise variance. We note that soft ZF demodulation performed consistently poorer than soft MMSE at the same computational cost; thus, there appear to be no reasons to prefer soft ZF in practical implementations.

Soft MMSE has an even stronger case against its competitors for asymmetric MIMO systems (i.e.,  $M_T < M_R$ ), where it performs close to BICM capacity for *all* rates. Fig. 13 compares a  $4 \times 4$  MIMO system using 4-QAM (system I) and a  $2 \times 4$  system using 16-QAM (system II), both using Gray labeling and achieving  $R_0 = 8$ . Whereas at low rates soft MMSE demodulation performs better with system I than with system II, it is the other way around for high rates. For example, at 6 bpcu system II requires 1.25 dB less SNR than system I in spite of using fewer active transmit antennas. This observation is of interest when designing MIMO-BICM systems with adaptive modulation and coding. Specifically, with

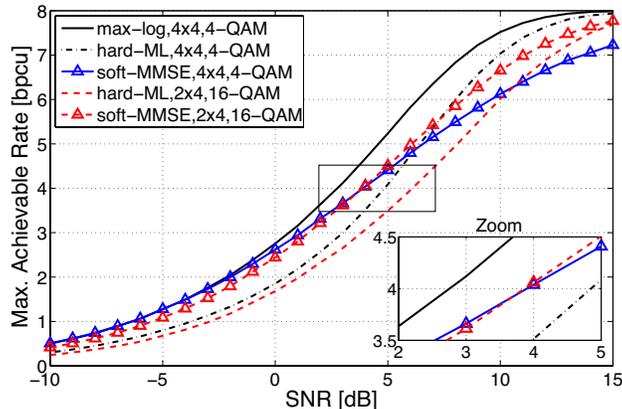


Fig. 13. System capacity of baseline demodulators for a  $4 \times 4$  MIMO system using 4-QAM and a  $2 \times 4$  MIMO system using 16-QAM (both with Gray labeling).

soft MMSE demodulation below 4 bpcu system I is preferable, whereas above 4 bpcu it is advantageous to deactivate two transmit antennas and switch to the 16-QAM constellation (system II). We note that with max-log and hard ML demodulation, system II performs consistently worse than system I.

The only regime where soft MMSE suffers from a noticeable performance loss is symmetric systems at high rates, where no low-complexity demodulation scheme is able to achieve max-log performance. These observations apply to the ergodic and quasi-static scenario. However, with perfect CSI the LSD and soft  $\ell^\infty$ -norm demodulation come reasonably close to max-log. The LSD has the additional advantages of being able to trade off performance for complexity reduction. Furthermore, note that for system I hard ML (i.e., the LSD with list size 1) outperforms most suboptimal soft demodulators for rates larger than 6 bpcu.

As a general rule of thumb, we conclude that at low rates linear soft demodulation is sufficient and generally preferable due to its low computational complexity. At high rates non-linear demodulator structures provide better performance, even when they deliver hard rather than soft outputs. If complexity is no issue, max-log demodulation is mostly superior to all other demodulators since it yields the highest rates over a wide range of system parameters and SNRs. Notable exceptions are SoftIC (which is a low complexity demodulator) and LSD which have the potential to outperform max-log at low rates and for imperfect CSI, respectively.

## IX. CONCLUSION

We provided a comprehensive performance assessment and comparison of soft and hard demodulators for non-iterative MIMO-BICM systems. Our comparison is based on the information-theoretic notion of *system capacity*, which can be interpreted as the maximum achievable rate of the equivalent “modulation” channel that comprises modulator, physical channel, and demodulator. As a performance measure, system capacity has the main advantage of being independent of any outer code. Extensive simulation results show that a universal demodulator performance ranking is *not* possible and that the demodulator performance can depend strongly on the rate (or equivalently the SNR) at which a system operates. In addition to ergodic capacity results, we investigated the non-ergodic fading scenario in terms of outage probability and  $\epsilon$ -capacity and analyzed the robustness of certain demodulators under imperfect channel state information. Our observations provide new insights into the design of MIMO-BICM systems (i.e., choice of demodulator, number of antennas, and symbol constellation). Moreover, our approach sheds light on issues that have not been apparent in the previously prevailing BER performance comparisons for specific outer codes. For example, an important key observation is that with low-rate outer codes soft MMSE demodulation is preferable over other demodulators since it has low complexity but nonetheless comes very close to max-log performance.

## ACKNOWLEDGMENT

The authors thank Gottfried Lechner for kindly providing his LDPC decoder implementation.

## REFERENCES

- [1] P. Fertl, J. Jaldén, and G. Matz, “Capacity-based performance comparison of MIMO-BICM demodulators,” in *IEEE 9th Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2008)*, Recife, Brazil, Jul. 2008, pp. 166–170.
- [2] G. Caire, G. Taricco, and E. Biglieri, “Bit-interleaved coded modulation,” *IEEE Trans. Inf. Theory*, vol. 44, no. 5, pp. 927–945, May 1998.
- [3] A. Guillén i Fàbregas, A. Martinez, and G. Caire, “Bit-Interleaved Coded Modulation,” *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 1–2, pp. 1–153, 2008.
- [4] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, “Multilevel codes: Theoretical concepts and practical design rules,” *European Trans. Telecomm.*, vol. 6, no. 5, pp. 557–567, Sep./Oct. 1995.
- [5] 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, Mar. 2000, pp. WA 4.1–4.6.
- [6] A. Stefanov and T. M. Duman, “Turbo-coded modulation for systems with transmit and receive antenna diversity over block fading channels: System model, decoding approaches, and practical considerations,” *IEEE J. Sel. Areas Comm.*, vol. 19, no. 5, pp. 958–968, May 2001.

- [7] S. H. Müller-Weinfurter, “Coding approaches for multiple antenna transmission in fast fading and OFDM,” *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2442–2450, Oct. 2002.
- [8] J. Jaldén and B. Ottersten, “Parallel implementation of a soft output sphere decoder,” in *Proc. 39th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, 2005, pp. 581–585.
- [9] —, “On the complexity of sphere decoding in digital communications,” *IEEE Trans. Signal Processing*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.
- [10] C. Studer, A. Burg, and H. Bölcskei, “Soft-output sphere decoding: Algorithms and VLSI implementation,” *IEEE J. Sel. Areas Comm.*, vol. 26, no. 2, pp. 290–300, Feb. 2008.
- [11] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bölcskei, “VLSI implementation of MIMO detection using sphere decoding algorithm,” *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, Jul. 2005.
- [12] D. Seethaler and H. Bölcskei, “Infinity-norm sphere-decoding,” in *IEEE Int. Symposium on Information Theory (ISIT)*, Toronto, Canada, Jul. 2008, pp. 2002–2006.
- [13] 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.
- [14] 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, Mar. 2003.
- [15] H. Yao and G. W. Wornell, “Lattice-reduction-aided detectors for MIMO communication systems,” in *Proc. IEEE GLOBECOM-2002*, vol. 1, Taipei, Taiwan, Nov. 2002, pp. 424–428.
- [16] D. Wübben, R. Böhnke, V. Kühn, and K. Kammeyer, “MMSE-based lattice-reduction for near-ML detection of MIMO systems,” in *Proc. ITG Workshop on Smart Antennas 2004*, Munich, Germany, Mar. 2004, pp. 106–113.
- [17] C. Windpassinger, L. H.-J. Lampe, and R. F. H. Fischer, “From lattice-reduction-aided detection towards maximum-likelihood detection in MIMO systems,” in *Proc. Wireless, Optical Communication Conference*, Banff, AB, Canada, Jul. 2003.
- [18] P. Silvola, K. Hooli, and M. Juntti, “Suboptimal soft-output MAP detector with lattice reduction,” *IEEE Signal Processing Letters*, vol. 13, no. 6, pp. 321–324, Jun. 2006.
- [19] V. Ponnampalam, D. McNamara, A. Lillie, and M. Sandell, “On generating soft outputs for lattice-reduction-aided MIMO detection,” in *Proc. IEEE ICC-07*, Glasgow, UK, Jun. 2007, pp. 4144–4149.
- [20] R. Wang and G. B. Giannakis, “Approaching MIMO channel capacity with soft detection based on hard sphere decoding,” *IEEE Trans. Comm.*, vol. 54, no. 4, pp. 587–590, Apr. 2006.
- [21] M. Butler and I. Collings, “A zero-forcing approximate log-likelihood receiver for MIMO bit-interleaved coded modulation,” *IEEE Commun. Letters*, vol. 8, no. 2, pp. 105–107, Feb. 2004.
- [22] 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.
- [23] D. Seethaler, G. Matz, and F. Hlawatsch, “An efficient MMSE-based demodulator for MIMO bit-interleaved coded modulation,” in *Proc. IEEE GLOBECOM-2004*, vol. IV, Dallas, Texas, Dec. 2004, pp. 2455–2459.
- [24] W.-J. Choi, K.-W. Cheong, and J. Cioffi, “Iterative soft interference cancellation for multiple antenna systems,” in *Proc. IEEE WCNC-2000*, vol. 1, Chicago, IL, Sep. 2000, pp. 304–309.
- [25] A. Paulraj, R. U. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge (UK): Cambridge Univ. Press, 2003.
- [26] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, “V-BLAST: An architecture for realizing very high

- data rates over the rich-scattering wireless channel,” in *Proc. URSI Int. Symp. on Signals, Systems and Electronics*, Pisa, Italy, Sep. 1998, pp. 295–300.
- [27] B. Hassibi, “A fast square-root implementation for BLAST,” in *Proc. 34th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov./Dec. 2000, pp. 1255–1259.
- [28] R. Böhnke, D. Wübben, V. Kühn, and K. D. Kammeyer, “Reduced complexity MMSE detection for BLAST architectures,” in *Proc. IEEE Globecom 2003*, vol. 4, San Francisco, CA, Dec. 2003, pp. 2258–2262.
- [29] P. Fertl, J. Jaldén, and G. Matz, “Performance assessment of MIMO-BICM demodulators based on system capacity: Further results,” Vienna University of Technology, Austria, Technical Report #09-1, Mar. 2009. [Online]. Available: [http://publik.tuwien.ac.at/files/PubDat\\_174303.pdf](http://publik.tuwien.ac.at/files/PubDat_174303.pdf)
- [30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [31] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Boston (MA): Cambridge University Press, 2005.
- [32] A. Lapidoth and P. Naryan, “Reliable communication under channel uncertainty,” *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2148–2177, Oct. 1998.
- [33] L. Paninski, “Estimation of entropy and mutual information,” *Neural Comput.*, vol. 15, no. 6, pp. 1191–1253, 2003.
- [34] 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.
- [35] 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, Sep. 2006.
- [36] M. Damen, H. El Gamal, and G. Caire, “On maximum-likelihood detection and the search for the closest lattice point,” *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.
- [37] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, “Factoring polynomials with rational coefficients,” *Math. Ann.*, vol. 261, pp. 515–534, 1982.
- [38] J. Jaldén, D. Seethaler, and G. Matz, “Worst- and average-case complexity of LLL lattice reduction in MIMO wireless systems,” in *Proc. IEEE ICASSP 2008*, Las Vegas, NV, Apr. 2008, pp. 2685–2688.
- [39] D. Wübben and D. Seethaler, “On the performance of lattice reduction schemes for MIMO data detection,” in *Proc. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, USA, Nov. 2007, pp. 1534–1538.
- [40] P. H. Tan and L. K. Rasmussen, “The application of semidefinite programming for detection in CDMA,” *IEEE J. Sel. Areas Comm.*, vol. 19, no. 8, pp. 1442–1449, Aug. 2001.
- [41] W. K. Ma, T. N. Davidson, K. M. Wong, Z. Q. Luo, and P. C. Ching, “Quasi-maximum-likelihood multiuser detection using semidefinite relaxation with application to synchronous CDMA,” *IEEE Trans. Signal Processing*, vol. 50, no. 4, pp. 912–922, Apr. 2002.
- [42] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge (UK): Cambridge Univ. Press, 1999.
- [43] M. Biguesh and A. B. Gershman, “Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals,” *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 884–893, Mar. 2006.
- [44] G. Taricco and E. Biglieri, “Space-time decoding with imperfect channel estimation,” *IEEE Trans. Wireless Comm.*, vol. 4, no. 4, pp. 1874–1888, Jul. 2005.