

# LOG-LIKELIHOOD RATIO CLIPPING IN MIMO-BICM SYSTEMS: INFORMATION GEOMETRIC ANALYSIS AND IMPACT ON SYSTEM CAPACITY

Stefan Schwandter, Peter Fertl, Clemens Novak, and Gerald Matz

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology  
Gusshausstrasse 25/389, 1040 Vienna, Austria; Email: {sschwand,pfertl,cnovak,gmatz}@nt.tuwien.ac.at

## ABSTRACT

The clipping of log-likelihood ratios (LLRs) in soft demodulators for multiple-input multiple-output (MIMO) systems with bit-interleaved coded modulation (BICM) was recently observed to allow for enormous complexity savings. In this paper we first provide an information-geometric interpretation of LLR clipping as information projection onto a log-convex manifold. Then we study the system capacity of MIMO-BICM systems that use LLR clipping. Our results show that strong LLR clipping is possible without significant capacity loss. We finally propose an LLR transformation scheme which is necessary for approaching capacity in case of strong clipping. The usefulness of this LLR transformation is illustrated by numerical simulations for MIMO-BICM systems employing low-density parity check (LDPC) codes.

**Index Terms**— log-likelihood ratio, soft demodulation, information geometry, BICM capacity

## 1. INTRODUCTION

Bit-interleaved coded modulation (BICM) [1] is a promising transmission scheme that has recently received increasing interest in the context of multiple-input multiple-output (MIMO) systems. BICM receivers comprise a demapper (demodulator) that provides soft information about the code bits to the channel decoder. Usually, this soft information consists of bit-wise log-likelihood ratios (LLRs). Since optimum demapping is exponentially complex in the number of transmit antennas, a large number of low-complexity detectors have been proposed. Specifically, soft sphere decoding with single tree search and LLR clipping was recently observed to provide an excellent performance-complexity tradeoff [2]. This provides the motivation for the present paper.

In this paper, we first provide an interpretation of LLR clipping in terms of notions from information geometry [3]. In particular, we show that clipping LLRs amounts to a projection onto a log-convex submanifold. We then study the capacity of a MIMO-BICM system with LLR clipping using the performance measure proposed in [4], which reveals that strong clipping is possible without noticeable capacity reduction. However, the theoretical capacity limits cannot be approached by plain application of low-density parity check (LDPC) codes with sum-product decoding. We argue that this is due to the fact that clipped LLRs tend to be overly pessimistic and we propose a simple LLR remapping that avoids this problem. All of our claims are substantiated via numerical simulations. We note that a similar

This work was supported by the FWF project “Information Networks” (S10603-N13) within the National Research Network SISE and by the STREP project MASCOT (IST-026905) within the Sixth Framework Programme of the European Commission.

problem in the context of iterative tree search has been investigated in [5].

The rest of this paper is organized as follows. Section 2 presents the MIMO-BICM system model. In Section 3, we provide an information geometric interpretation of LLR clipping and Section 4 studies the capacity of MIMO-BICM systems with LLR clipping. A correction of the clipped LLR values is introduced in Section 5. Conclusions are provided in Section 6.

## 2. SYSTEM MODEL

In our MIMO-BICM system, a block of  $L$  information bits  $\mathbf{b}$  is encoded with a rate  $R$  code, yielding a block of  $N = L/R$  code bits. The code bits are passed through an interleaver  $\Pi$  and groups of  $R_0 = mM_T$  interleaved bits<sup>1</sup>  $c_l[n]$ ,  $l = 1, \dots, R_0$ , are mapped to complex-valued transmit vectors  $\mathbf{s}[n] \triangleq (s_1[n] \dots s_{M_T}[n])^T$  with symbols  $s_k$  from an alphabet  $\mathcal{X}$  of size  $|\mathcal{X}| = 2^m$ . Assuming  $M_R$  receive antennas and fast Rayleigh fading, the output of the MIMO channel is

$$\mathbf{y}[n] = \mathbf{H}[n]\mathbf{s}[n] + \mathbf{w}[n],$$

where  $\mathbf{H}[n]$  is the  $M_R \times M_T$  channel matrix and  $\mathbf{w}[n]$  is the noise vector with i.i.d. circularly symmetric complex Gaussian components of variance  $\sigma_w^2$ . The transmit vector is subject to the power constraint  $\mathbb{E}\{\|\mathbf{s}[n]\|^2\} = E_s$ . Given normalized entries of the channel matrix, the receive SNR therefore equals  $\rho = E_s/\sigma_w^2$ .

Assuming perfect knowledge of the channel state  $\mathbf{H}[n]$  and the noise variance  $\sigma_w^2$  at the receiver, the demapper calculates LLRs for the code bits:

$$\Lambda_l[n] \triangleq \log \frac{\mathbb{P}\{c_l[n]=1|\mathbf{y}[n]\}}{\mathbb{P}\{c_l[n]=0|\mathbf{y}[n]\}}. \quad (1)$$

With the max-log approximation [1] there is

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

Here,  $\mathcal{X}_l^d$  denotes the subset of transmit vectors  $\mathbf{s}$  for which the  $l$ th bit equals  $d$ . These LLRs are de-interleaved ( $\Pi^{-1}$ ) and passed to a soft-in channel decoder that exploits the code structure to deliver final estimates of the information bits.

Soft sphere decoding with LLR clipping uses a threshold  $\hat{\Lambda}$  to terminate the tree search as soon as it turns out that  $\tilde{\Lambda}_l[n]$  is larger than  $\hat{\Lambda}$  or smaller than  $-\hat{\Lambda}$ , in which case  $\pm\hat{\Lambda}$  is passed to the channel decoder. This can equivalently be formulated as actual clipping:

$$\tilde{\Lambda}_l[n] = \begin{cases} \hat{\Lambda}, & \text{for } \Lambda_l[n] > \hat{\Lambda}, \\ \Lambda_l[n], & \text{for } |\Lambda_l[n]| \leq \hat{\Lambda}, \\ -\hat{\Lambda}, & \text{for } \Lambda_l[n] < -\hat{\Lambda}. \end{cases} \quad (2)$$

<sup>1</sup>Here,  $n = 1, \dots, N/R_0$  denotes discrete symbol time.

### 3. INFORMATION-GEOMETRIC INTERPRETATION OF LLR CLIPPING

We next provide some basics on information geometry which views probability distributions as points on a manifold. For further details we refer to [3] and [6].

**Fundamentals.** In this paper, we are mostly concerned with length- $N$  sequences  $\mathbf{c} = (c_1 \dots c_N)$  of code bits  $c_{(n-1)R_0+l} = c_l[n]$ . The  $2^N$  possible sequences can be characterized via their (sequence) *probability mass functions (pmf)*  $p(\mathbf{c}_k) = \mathbb{P}\{\mathbf{c} = \mathbf{c}_k\}$ ,  $k = 0, \dots, K$ , where  $K = 2^N - 1$  and  $\mathbb{P}\{E\}$  denotes the probability of the event  $E$ . Since  $\sum_{k=0}^K p(\mathbf{c}_k) = 1$ , i.e.,  $p(\mathbf{c}_0) = 1 - \sum_{k=1}^K p(\mathbf{c}_k)$ , the collection of all possible pmfs of  $\mathbf{c}$  constitutes an exponential manifold of dimension  $K = 2^N - 1$ . Without loss of generality, we assume that the  $2^N$  possible sequences are indexed such that  $\mathbf{c}_k$  equals the length- $N$  binary representation of  $k$ . Equivalent coordinates for pmfs on this manifold are given by  $\tilde{\mathbf{p}} = (\tilde{p}^1 \dots \tilde{p}^K)^T$  with  $\tilde{p}^k = p(\mathbf{c}_k)$  (“expectation parameters”) or  $\tilde{\boldsymbol{\eta}} = (\tilde{\eta}^1 \dots \tilde{\eta}^K)^T$  with  $\tilde{\eta}^k = \log \frac{p(\mathbf{c}_k)}{p(\mathbf{c}_0)}$  (“natural parameters”).

Distances on the manifold of pmfs can be measured in terms of the *Kullback-Leibler (KL) divergence* [7] between two distributions  $p(\mathbf{c})$  and  $q(\mathbf{c})$ ,

$$D(p||q) = \sum_{k=0}^K p(\mathbf{c}_k) \log \frac{p(\mathbf{c}_k)}{q(\mathbf{c}_k)}.$$

There is  $D(p||q) \geq 0$  with equality if and only if  $p(\mathbf{c}) = q(\mathbf{c})$ . However, the KL divergence is non-symmetric and hence no metric.

The orthogonal projection of Euclidean geometry has an analog in information geometry which is termed *I-projection* [6]. Given a distribution  $q$ , the I-projection onto a convex set  $\mathcal{S}$  is defined as the distribution  $p^* \in \mathcal{S}$  that is “closest” to  $q$  in terms of KL divergence:

$$p^* = \arg \min_{p \in \mathcal{S}} D(p||q). \quad (3)$$

The asymmetry of the KLD further allows to define a so-called *rI-Projection* (the ‘r’ stands for ‘reverse’):

$$p^* = \arg \min_{p \in \mathcal{S}} D(q||p), \quad (4)$$

where  $\mathcal{S}$  now is a log-convex set.

**Demapper Output PMF.** A special case important for our analysis is the set of bit sequence distributions that are factorizable:

$$\mathcal{F} = \left\{ p: p(\mathbf{c}) = \prod_{i=1}^N p_i(c_i) \right\}, \quad (5)$$

Here,  $p_i(c_i) = \sum_{\sim c_i} p(\mathbf{c})$  where  $\sim c_i$  means summation with respect to all bits except  $c_i$ . It can easily be verified that  $\mathcal{F}$  is an  $N$ -dimensional log-convex submanifold with coordinates  $(p_1 \dots p_N)^T$  where  $p_i = \mathbb{P}\{c_i = 1\}$  or  $\boldsymbol{\eta} = (\eta_1 \dots \eta_N)^T$  with  $\eta_i = \log \frac{\mathbb{P}\{c_i=1\}}{\mathbb{P}\{c_i=0\}}$ . Specifically, using the factorization property in (5) allows to rewrite the natural sequence parameters as

$$\tilde{\boldsymbol{\eta}} = \mathbf{B}\boldsymbol{\eta}, \quad (6)$$

where  $\mathbf{B}$  is a  $K \times N$  matrix whose  $k$ th row contains the length- $N$  binary representation of  $k$ . Thus, in the log-domain  $\mathcal{F}$  is an  $N$ -dimensional linear submanifold.

Note that the natural coordinates are just the bit LLRs. It can be shown (cf. [8] for the single-antenna case) that the demapper output corresponds to a factorizable sequence pmf

$$p^{\text{dem}}(\mathbf{c}) = \prod_{i=1}^N p_i^{\text{dem}}(c_i), \quad p_i^{\text{dem}}(c_i) = \frac{1}{1 + e^{-(2c_i-1)\Lambda_i}},$$

(where  $\Lambda_{(n-1)R_0+l} = \Lambda_l[n]$ ), which is the factorizable pmf closest to the “observed” distribution  $p^{\text{obs}}(\mathbf{c}) \propto p(\{\mathbf{y}[1], \dots, \mathbf{y}[N/R_0]\}|\mathbf{c})$ . More precisely,

$$p^{\text{dem}}(\mathbf{c}) = \min_{p \in \mathcal{F}} D(p^{\text{obs}}||p).$$

**Clipping Manifold.** In the following, we will show that bit-wise LLR clipping can be interpreted as a reverse I-projection of the observed pmf onto a log-convex submanifold (termed “clipping manifold”). To begin with, we consider a single bit  $c_i$  with pmf  $p_i(c_i)$ , i.e., with expectation parameter  $p_i$  and natural parameter  $\eta_i$ . Clipping the LLR  $\Lambda_i = \eta_i$  of this bit using a threshold  $\hat{\Lambda}$  is equivalent to restricting  $p_i$  to the symmetric interval

$$\hat{p} \leq p_i \leq 1 - \hat{p}, \quad (7)$$

with  $\hat{p} = 1/(1 + e^{\hat{\Lambda}})$ . An equivalent formulation of this restriction can be given in terms of the KL divergence between  $p_i(c_i)$  and the uniform distribution  $\bar{p}(0) = \bar{p}(1) = 1/2$ , or equivalently, the entropy of  $p_i(c_i)$ , i.e., the clipped distributions  $p_i$  satisfy

$$D(p_i||\bar{p}) = 1 - h_2(p_i) \leq \hat{D},$$

with  $\hat{D} = 1 - h_2(\hat{p})$  and  $h_2(p)$  denoting the binary entropy function.

Applying the above conditions to all bit LLRs in a sequence defines the “clipping manifold”

$$\mathcal{M} = \{p \in \mathcal{F}: |\eta_i| \leq \hat{\Lambda}, \quad i = 1, \dots, N\}$$

Clearly,  $\mathcal{M} \subset \mathcal{F}$ . In view of (6), it is straightforward to show that this is a log-convex manifold, i.e., for two distributions  $p_0(\mathbf{c}) \in \mathcal{M}$  and  $p_1(\mathbf{c}) \in \mathcal{M}$  with natural coordinates  $\tilde{\boldsymbol{\eta}}^0$  and  $\tilde{\boldsymbol{\eta}}^1$ , respectively, the log-domain convex combination  $\tilde{\boldsymbol{\eta}}^\alpha = \alpha \tilde{\boldsymbol{\eta}}^1 + (1 - \alpha) \tilde{\boldsymbol{\eta}}^0$  corresponds to a distribution again in  $\mathcal{M}$ . Fig. 1 illustrates the clipping manifold for  $N = \hat{\Lambda} = 2$ .

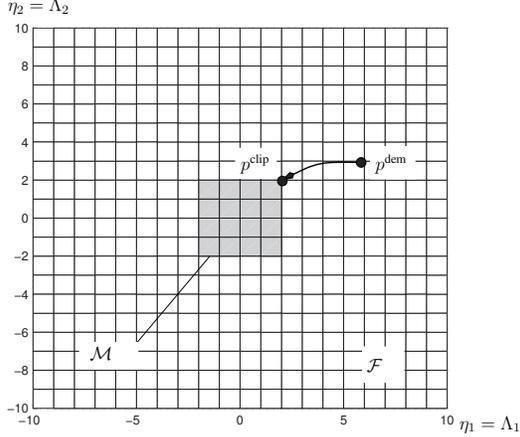
**Projection Interpretation.** We next provide an interpretation of LLR clipping in terms of information projections. Consider the distribution  $p^{\text{clip}}(\mathbf{c}) \in \mathcal{F}$  with natural coordinates  $\eta_i = \hat{\Lambda}_i$ , which corresponds to the output of the demodulator after LLR clipping. In the following, we will argue that

$$p^{\text{clip}}(\mathbf{c}) = \arg \min_{q \in \mathcal{M}} D(p^{\text{obs}}||q), \quad (8)$$

i.e.,  $p^{\text{clip}}(\mathbf{c})$  is the unique distribution that is factorizable, has LLRs with magnitude not larger than  $\hat{\Lambda}$ , and is closest to the observed distribution  $p^{\text{obs}}(\mathbf{c})$ .

We provide the main arguments underlying (8). First we recall that  $\mathcal{M} \subset \mathcal{F}$  and hence the projection in (8) can be performed by first projecting onto  $\mathcal{F}$ , which yields  $p^{\text{dem}}(\mathbf{c})$ , and subsequently projecting onto the clipping manifold (see Fig. 1):

$$p^{\text{clip}}(\mathbf{c}) = \arg \min_{q \in \mathcal{M}} D(p^{\text{dem}}||q). \quad (9)$$



**Fig. 1.** Bit LLR clipping: interpretation as projection of  $p^{\text{dem}}$  onto the clipping manifold  $\mathcal{M}$ .

Since both  $p^{\text{dem}}(\mathbf{c})$  and  $q(\mathbf{c}) \in \mathcal{M}$  are factorizable, the sequence KL divergence can be split into a sum of bit-wise KL divergences,

$$D(p^{\text{dem}} \| q) = \sum_{i=1}^N D(p_i^{\text{dem}} \| q_i),$$

where each term can now be individually minimized subject to the side constraint  $q \in \mathcal{M}$ . On the bit-level, this constraint reduces to (7). Clearly if  $p_i^{\text{dem}}(c_i)$  satisfies (7), we can achieve  $D(p_i^{\text{dem}} \| q_i) = 0$  by setting  $q_i(c_i) = p_i^{\text{dem}}(c_i)$ . Otherwise, it can be shown by direct verification that the minimum of  $D(p_i^{\text{dem}} \| q_i)$  is achieved by choosing  $q_i(c_i = 1) = \hat{p}$  or  $q_i(c_i = 1) = 1 - \hat{p}$ , depending on which one is closer to  $p^{\text{dem}}(c_i)$ .

#### 4. CAPACITY ANALYSIS

**Performance Measure.** In the following, we investigate the impact of LLR clipping on the rates achievable in a non-iterative MIMO-BICM system by using the code-independent performance measure described in [4]. Specifically, it was proposed to quantify the capacity that can be achieved with a specific MIMO-BICM demodulator in terms of the mutual information between the input and output of an equivalent “modulation” channel. This modulation channel comprises the modulator, the wireless channel, and the demodulator, i.e., the channel inputs are given by<sup>2</sup>  $c_l$  and the channel outputs are given by the LLRs. In our case, where the demodulator performs LLR clipping (cf. (2)), the channel is completely characterized by the conditional probability density function (pdf)

$$f(\tilde{\Lambda}_l | c_l) = f_0(\tilde{\Lambda}_l | c_l) \text{rect}(\Lambda_l; \hat{\Lambda}) + P_{c_l} \delta(\tilde{\Lambda}_l - \hat{\Lambda}_l) + P_{1-c_l} \delta(\tilde{\Lambda}_l + \hat{\Lambda}_l), \quad (10)$$

where  $f_0(\Lambda_l | c_l)$  is the conditional pdf of the true LLRs (assumed to satisfy  $f_0(\Lambda_l | c_l) = f_0(-\Lambda_l | 1 - c_l)$ ),  $P_{c_l} = \text{P}\{\Lambda_l > \hat{\Lambda} | c_l\}$  denotes the conditional clipping probabilities, and  $\text{rect}(\tilde{\Lambda}_l; \hat{\Lambda})$  is 1 if  $\tilde{\Lambda}_l \in ] - \hat{\Lambda}; \hat{\Lambda} [$  and zero otherwise.

Following [4], the achievable rate of MIMO-BICM with LLR clipping at the demodulator is given by

$$R = R_0 - \sum_{l=1}^{R_0} \sum_{c_l=0}^1 \int_{-\infty}^{\infty} \frac{1}{2} f(\tilde{\Lambda}_l | c_l) \log_2 \frac{f(\tilde{\Lambda}_l)}{\frac{1}{2} f(\tilde{\Lambda}_l | c_l)} d\tilde{\Lambda}_l, \quad (11)$$

<sup>2</sup>For convenience, we suppress the symbol time index  $n$  in the following.

where  $f(\tilde{\Lambda}_l) = \frac{1}{2} \sum_{b=0}^1 f(\tilde{\Lambda}_l | c_l = b)$ ; here, we assumed that the code bits  $c_l$  are uniformly distributed and statistically independent (as guaranteed by an ideal interleaver). Note that the measure  $R$  constitutes a code-independent performance measure. Moreover, the data processing inequality [7] implies that  $R \leq C_{\text{BICM}}$ , where  $C_{\text{BICM}}$  denotes the capacity delivered by the MIMO-BICM system employing an optimal soft MAP demodulator [9].

**Capacity Results.** In what follows, we evaluate the capacity in (11) using various clipping thresholds under the assumption of ergodic i.i.d. fast Rayleigh fading. The required pdfs were estimated by means of Monte-Carlo simulations. Fig. 2(a) shows the results obtained for clipped max-log demodulation in bit per channel use (bpcu) versus clipping threshold  $\hat{\Lambda}$  for a  $2 \times 2$  MIMO system with 16-QAM (here,  $R_0 = 8$ ) and Gray labeling for various SNRs (shown as curve labels). As a reference we also plot the performance curves for a hard ML detector that corresponds to infinitesimally small clipping threshold. It is seen that above a certain threshold (about  $\hat{\Lambda} > 5$ ), the capacity at all SNRs is virtually the same as without clipping. In contrast, for  $\hat{\Lambda} < 0.5$  performance rapidly approaches that of hard ML detection. Further simulations (not shown due to lack of space) revealed that a similar behavior holds true for other labeling strategies, constellation sizes, and antenna configurations. Moreover, it can be observed that the relative capacity loss incurred by strong clipping is lower at higher SNR, i.e., the impact of clipping is more pronounced at low SNR.

Complementing the plots of maximum rate at fixed SNR, Fig. 2(b) shows the minimum SNR required to achieve a fixed target rate versus the clipping threshold. Again, the required SNR increases only if the clipping threshold is below a certain value. Very strong clipping results in an SNR penalty of about 3 dB at rate  $1/4$  and about 2 dB at rate  $3/4$ . This confirms that LLR clipping incurs performance loss particularly at low rates/low SNR.

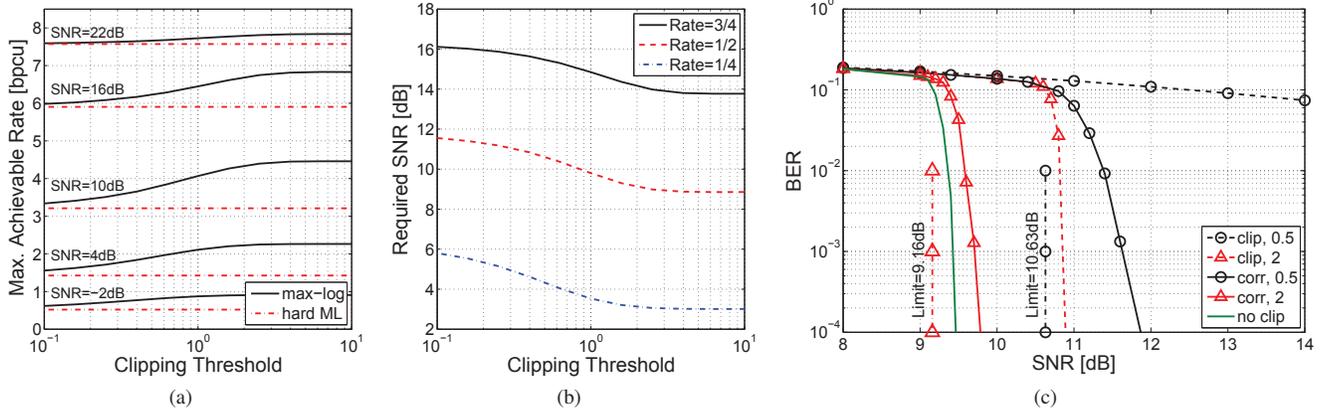
#### 5. LLR CORRECTION

The LLR magnitude provides reliability information about the corresponding code bit, which can be exploited by the channel decoder. LLR clipping always renders the corresponding bit less reliable than it actually is, i.e., the clipping amounts to an overly pessimistic worst-case representation of all LLRs whose magnitude is larger than the clipping threshold. This suggests that clipped LLRs should be mapped to a different value that on average better represents the actual reliabilities. This view can be supported by the equivalent modulation channel given by (10). Clearly, there are two discrete channel outputs ( $\pm \hat{\Lambda}$ ) with positive probabilities and the decoder should actually use the LLR of these discrete channel outputs. We note that belief propagation [10], predominantly used for LDPC decoding, is particularly sensitive to incorrect reliability information. We remark that the capacity results of Section 4 do not depend on the actual representative of clipped LLRs since mutual information does not depend on the channel inputs and outputs but only on their probabilities; however, incorrect reliability information passed to the channel decoder may prevent the system from approaching capacity.

Based on (10), we obtain the LLRs for the equivalent channel as

$$\Lambda_l^* = \log \frac{f(\tilde{\Lambda}_l | c_l = 1)}{f(\tilde{\Lambda}_l | c_l = 0)} = \begin{cases} \hat{\Lambda}^*, & \text{for } \Lambda_l > \hat{\Lambda}, \\ \Lambda_l, & \text{for } |\Lambda_l| \leq \hat{\Lambda}, \\ -\hat{\Lambda}^*, & \text{for } \Lambda_l < -\hat{\Lambda}. \end{cases} \quad (12)$$

Here,  $\hat{\Lambda}^*$  denotes the corrected value of the clipped LLRs that is



**Fig. 2.** Simulation results for a  $2 \times 2$  MIMO-BICM system with 16-QAM: (a) Capacity vs. clipping threshold for different SNRs, (b) required SNR vs. clipping value for different code rates, (c) BER vs. SNR for different clipping levels with and without LLR correction.

given by

$$\hat{\Lambda}^* = \log \frac{P_1}{P_0} = \log \frac{P\{\Lambda_l > \hat{\Lambda} | c_l = 1\}}{P\{\Lambda_l > \hat{\Lambda} | c_l = 0\}}. \quad (13)$$

Since closed-form expressions for the distribution of the optimum LLRs  $\Lambda_l$  are not available, Monte Carlo simulations are required to determine the clipping probabilities  $P_0$  and  $P_1$ . In a practical implementation, the corrected LLRs  $\Lambda_i^*$  may be precomputed and stored for the system and channel parameters of interest.

**Simulation results.** To illustrate the effect of LLR correction, we simulated a  $2 \times 2$  MIMO-BICM system with 16-QAM signal constellation, Gray labeling, and a regular rate  $1/2$  LDPC code<sup>3</sup> of block length 64000. Fig.2(c) shows BER versus SNR for conventional LLR clipping according to (2) (labeled “clip”) and for LLR correction according to (12) (labeled “corr”), both for clipping thresholds of  $\hat{\Lambda} = 0.5$  and  $\hat{\Lambda} = 2$ . The capacity limits in terms of minimum SNR required to support rate  $1/2$  (cf. Fig. 2(b)) are also indicated. With clipping threshold  $\hat{\Lambda} = 2$ , LLR correction is within 0.6 dB of capacity, outperforming plain clipping by roughly 1 dB and approaching the performance of an idealized system without LLR clipping. In contrast, with clipping threshold  $\hat{\Lambda} = 0.5$ , plain clipping is practically useless while LLR correction allows coming within roughly 1.5 dB of the capacity limit of 10.63 dB. We conclude that LLR correction is more important for small clipping thresholds and can provide significant performance improvements.

## 6. CONCLUSIONS

We have investigated the problem of LLR clipping in demodulators for MIMO-BICM systems. An information geometric analysis revealed that bit-wise LLR clipping can be interpreted in terms of an rI-projection on a log-convex manifold, thereby providing a first step towards a better theoretical understanding of LLR clipping. We then used a capacity-based performance measure to investigate the effect of LLR clipping onto the ultimate performance of a MIMO-BICM demodulator. It turned out that it is possible to use quite small clipping thresholds without noticeable capacity loss. Finally, we proposed an LLR correction scheme for the clipped LLRs which is necessary in order for the decoder to indeed approach capacity.

<sup>3</sup>The code design was performed using the EPFL web-application at <http://lthcwww.epfl.ch/research/ldpcopt>.

## ACKNOWLEDGEMENT

The authors thank Prof. Helmut Bölcskei for suggesting to study the effect of LLR clipping in MIMO systems.

## 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] C. Studer, A. Burg, and H. Bölcskei, “Soft-output sphere decoding: Algorithms and VLSI implementation,” submitted to *IEEE J. Sel. Areas Comm.*, Apr. 2007.
- [3] S. Amari and H. Nagaoka, *Methods of Information Geometry*, American Mathematical Society and Oxford University Press, New York, 2000.
- [4] Peter Fertl, Joakim Jaldén, and Gerald Matz, “Capacity-based performance comparison of MIMO-BICM demodulators,” in *Proc. IEEE SPAWC-08 Workshop on Signal Processing Advances in Wireless Communications*, Recife, Brazil, July 2008, pp. 166–170.
- [5] Y. L. C. de Jong and T. J. Willink, “Iterative tree search detection for MIMO wireless systems,” *IEEE Trans. Communications*, vol. 53, no. 6, pp. 930–935, June 2005.
- [6] I. Csiszár and P. Shields, “Information theory and statistics: A tutorial,” *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 4, pp. 420–527, 2004.
- [7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.
- [8] B. Muquet, *Novel receiver and decoding schemes for wireless OFDM systems with cyclic prefix or zero padding*, Ph.D. thesis, ENST, Paris (France), June 2001.
- [9] S. H. Müller-Weinfurter, “Coding approaches for multiple antenna transmission in fast fading and OFDM,” *IEEE Trans. Signal Processing*, vol. 50, pp. 2442–2450, Oct. 2002.
- [10] 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.