# Optimized Symbol Mappings for Bit-Interleaved Coded Modulation with Iterative Decoding

F. Schreckenbach, N. Görtz, J. Hagenauer
Institute for Communications Engineering (LNT)
Munich University of Technology (TUM)
80290 Munich, Germany
Email: frank.schreckenbach@ei.tum.de

G. Bauch
DoCoMo Communications Laboratories Europe GmbH
Landsbergerstr. 312
80687 Munich, Germany
Email: bauch@docomolab-euro.com

***Abstract*— We investigate bit-interleaved coded modulation with iterative decoding (BICM-ID) for bandwidth efficient transmission, where the bit error rate is reduced through iterations between a multilevel demapper and a simple channel decoder. In order to achieve a significant turbo-gain, the assignment strategy of the binary indices to signal points is crucial. We address the problem of finding the most suitable index assignments to arbitrary, high order signal constellations. A new method based on the *binary switching algorithm* is proposed that finds optimized mappings outperforming previously known ones.**

## I. Introduction

Bit-interleaved coded modulation (BICM) [1], [2] is the concatenation of an encoder, an interleaver and a symbol mapper, and is well suited for bandwidth efficient transmission over fading channels. The performance of BICM can be greatly improved through iterative information exchange between the *inner* decoder, i.e. the demapper, and the *outer* channel decoder, similar to iterative decoding of serial concatenated codes (SCCC) [3]. This system, introduced in [4], [5], is usually referred to as BICM with iterative decoding (BICM-ID). It is a promising low complexity alternative to turbo codes [6] well suited for a combination with e.g. iterative equalization or MIMO detection. It was soon recognized that the choice of the mapping (labeling map) is the crucial design parameter to achieve a high coding gain over the iterations. With increasing constellation order, an exhaustive computer search to find suitable mappings becomes intractable due to high complexity. Several mappings for BICM-ID were introduced in [7]−[10], but to the best of our knowledge, no general design algorithm has been proposed.

In this paper, we introduce a low complexity method to find labeling maps with desired characteristics for arbitrary signal constellations. The optimization scheme is based on the *binary switching algorithm (BSA)* previously used for index optimization in vector quantization [11]. The labeling indexes are assigned to constellation points so that an average cost is optimized. We investigate several cost functions for BICM-ID based on mutual information and error bounds.

Section II introduces the system model. Then, in Section III, a *distance spectrum* for mappings and the EXIT charts [12] are used to characterize mappings. The BSA and appropriate cost functions are investigated in Section IV. The analysis and simulation results of the new mappings show the performance gains of the optimization with the BSA in Section V.

This work was supported by DoCoMo Communications Laboratories Europe GmbH, Munich, Germany

## II. System model

We consider the BICM system with iterative decoding (BICM-ID) depicted in Fig. 1.

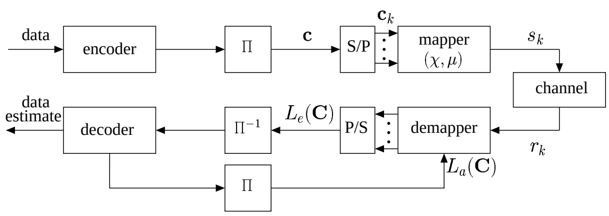



Fig. 1. BICM-ID system model.

A block of data bits is encoded by a convolutional encoder and bit-interleaved by the random interleaver $\Pi$. The coded and interleaved sequence is denoted by $\mathbf{c}$; $m$ consecutive bits of the sequence $\mathbf{c}$ are grouped to form the subsequences $\mathbf{c}_k = (c_k(1), \ldots, c_k(m))$. Each subsequence $\mathbf{c}_k$ is mapped to a complex symbol $s_k = \mu(\mathbf{c}_k)$ chosen from the $2^m$-ary signal constellation $\chi$ according to the labeling map $\mu$.

The channel is described by $r_k = a_k \cdot s_k + n_k$, where $a_k$ denotes the fading coefficient, and $n_k$ is the complex zero-mean Gaussian noise with variance $\sigma_n^2 = N_0/2E_s$ in each real dimension.

At the receiver, the demapper processes the received complex symbols $r_k$ and the corresponding a priori log-likelihood ratios (LLRs) $L_a(C_k(i)) = \log(\frac{P(C_k(i)=0)}{P(C_k(i)=1)})$[1] of the coded bits and outputs the extrinsic LLRs [5]:

$$L_e(C_k(i)) = \log \frac{P(c_k(i) = 0 | r_k, L_a(\mathbf{C}_k))}{P(C_k(i) = 1 | r_k, L_a(\mathbf{C}_k))} - L_a(C_k(i)). \tag{1}$$

Let $\chi_b^i$ denote the subset of symbols $s_k \in \chi$ whose bit labels have the value $b \in \{0,1\}$ in position $i \in \{1, \ldots, m\}$, i.e. $\chi_b^i = \{s_k = \mu(\mathbf{c}_k), \forall \mathbf{c}_k \in \{0,1\}^m | c_k(i) = b\}$. Using Bayes' rule and taking the expectation of $p(r_k|s_k)$ over $P(s_k|C_k(i) = b)$, $s_k \in \chi_b^i$, yields

$$L_e(C_k(i)) = \log \frac{\sum_{s_k \in \chi_0^i} p(r_k|s_k) \cdot P(s_k|C_k(i) = 0)}{\sum_{s_k \in \chi_1^i} p(r_k|s_k) \cdot P(s_k|C_k(i) = 1)}. \tag{2}$$

[1] $C_k(i)$ denotes the binary random variable with realizations $c_k(i) \in \{0,1\}$

The first term $p(r_k|s_k)$ is computed according to the channel model out of the Gaussian distribution:

$$p(r_k|s_k) = \frac{1}{\sqrt{2\pi}\sigma_n} \mathrm{e}^{-\frac{|r_k - a_k \cdot s_k|^2}{2\sigma_n^2}}. \tag{3}$$

The second term $P(s_k|C_k(i) = b)$ is computed from the a priori information of the individual bits:

$$P(s_k|C_k(i) = b) = \prod_{j=1, j\neq i}^{m} \frac{1}{1 + \mathrm{e}^{-L_a(C_k(j))}} \mathrm{e}^{-L_a(C_k(j))\cdot c_k(j)}. \tag{4}$$

The extrinsic estimates $L_e(C_k(i))$ are deinterleaved and applied to the APP channel decoder. Performing iterative decoding, extrinsic information about the coded bits from the decoder is fed back and regarded as a priori information $L_a(C_k(i))$ at the demapper. During the initial demapping step, the a priori LLRs are set to zero.

## III. Characteristics of Labeling Maps

The applied labeling map is the crucial design parameter for the considered BICM-ID system. The labeling map can be either characterized through a *distance spectrum* or through EXIT charts.

### A. Distance spectrum

Similar to the distance spectrum of Hamming weights for a channel code, we introduce a distance spectrum of Euclidean distances for mappings with the values $\bar{N}(d_E)$, which are defined as the number $N_N(d_E, s_k)$ of symbols $\hat{s}_k \in \chi_{\bar{b}}^i$ [2] at Euclidean distance $d_E$ from the symbol $s_k \in \chi_b^i$, averaged over all bits $i = 1, \ldots, m$ and $2^m$ symbols $s_k$:

$$\bar{N}(d_E) = \frac{1}{m2^m} \sum_{i=1}^{m} \sum_{b=0}^{1} \sum_{s_k \in \chi_b^i} N_N(d_E, s_k). \tag{5}$$

In other words, the distance spectrum is the average number of bit errors made at a specific Euclidean distance $d_E$:

$$\bar{N}(d_E) = \frac{1}{m2^m} \sum_{s_k \in \chi} N_H(d_E, s_k), \tag{6}$$

where $N_H(d_E, s_k)$ is the total Hamming distance between the symbol $s_k$ and the symbols at Euclidean distance $d_E$ from $s_k$.

As an example, we will consider three characteristic 16-QAM mappings proposed in the literature and depicted in Fig. 2: Gray, where symbols at minimum Euclidean distance differ in one bit, Modified Set Partitioning (MSP) [8] and Maximum Squared Euclidean Weight (MSEW) mapping [10].

We consider first the case where *no a priori* information about the coded bits is available at the demapper (e.g. during the initial demapping step). The shaded regions in Fig. 2 correspond to the decision regions for bit $i$ having the value 1 (subset $\chi_1^i$), the unshaded regions correspond to the decision regions for bit $i = 1$ having the value 0 (subset $\chi_0^i$). If a symbol error occurs within one decision region, no error will be made on the corresponding bit. Large decision regions provide a high protection for the corresponding bit, since the number of *nearest neighbors*, i.e. the number of symbols $\hat{s}_k \in \chi_{\bar{b}}^i$ at minimum Euclidean distance from $s_k \in \chi_b^i$ is minimized. In other words, the Hamming distances between symbols at small Euclidean distance should be minimized in order to minimize the number of bit errors for one symbol error.

[2] $\chi_{\bar{b}}^i$ denotes the complementary subset to $\chi_b^i$

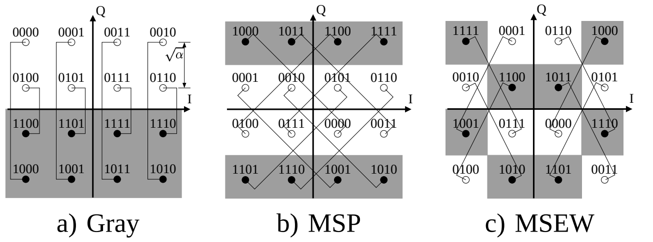



Fig. 2. For bit $i$=1: 16-QAM mappings with decision regions without a priori information and binary signal constellations selected by the ideal a priori information of the last 3 bits.

Table I shows the distance spectrum as defined above for the considered signal constellations. $\alpha$, shown in Fig. 2, denotes the minimum squared Euclidean distance ($d_E^2$) of 16-QAM. The performance is expected to be the best for Gray labeling, where the average number of nearest neighbors is minimized, followed by MSP and MSEW mapping.

| $d_E^2$ | $\alpha$ | $2\alpha$ | $4\alpha$ | $5\alpha$ | $8\alpha$ | $9\alpha$ | $10\alpha$ | $13\alpha$ | $18\alpha$ |
|---|---|---|---|---|---|---|---|---|---|
| Gray | 0.75 | 1.12 | 1 | 2.25 | 1 | 0.25 | 0.75 | 0.75 | 0.12 |
| MSP | 1.62 | 1.19 | 0.75 | 2.25 | 0.25 | 0.62 | 0.75 | 0.5 | 0.06 |
| MSEW | 2.25 | 1.12 | 1 | 0.75 | 1 | 0.75 | 0.75 | 0.25 | 0.12 |

TABLE I

Distance spectrum with no a priori information.

The performance with *ideal a priori* information at the demapper represents the achievable gain over the iterations. In this case, all the bits are perfectly known at the demapper, except for the bit to be detected, since only the *extrinsic* information is used. The a priori known bits select a pair of symbols which differ only in the bit $i$ to be detected. Possible symbol pairs for bit $i = 1$ are shown in Fig. 2. The symbol pairs consist of the subsets $\chi_1^i$ and $\chi_0^i$ which are reduced through ideal a priori knowledge from the shaded and unshaded regions to one symbol, denoted by the filled and unfilled signal points respectively. The distance spectrum, shown in Table II, can be obtained similar to the case without a priori information. With Gray mapping, the number of distances at minimum Euclidean distance is not reduced through a priori knowledge. Thus, only very small performance improvement is expected over the iterations. With MSP, designed in [8] to provide a good trade-off between performance with and without a priori information, the number of small distances is reduced and with MSEW mapping, the minimum squared Euclidean distance between the symbol pairs is maximized. It is interesting that, if the binary bit labels are converted to decimal numbers, the MSEW mapping is a perfect magic square, since the sum of all rows, columns, diagonals, $2 \times 2$ sub-squares and $2 \times 2$ cyclic sub-squares is equal to 30. We also observed that every perfect magic square has the distance spectrum of the MSEW mapping.

| $d_E^2$ | $\alpha$ | $2\alpha$ | $4\alpha$ | $5\alpha$ | $8\alpha$ | $9\alpha$ | $10\alpha$ | $13\alpha$ | $18\alpha$ |
|---|---|---|---|---|---|---|---|---|---|
| Gray | 0.75 | 0 | 0 | 0 | 0 | 0.25 | 0 | 0 | 0 |
| MSP | 0 | 0.06 | 0.25 | 0.12 | 0.25 | 0 | 0.12 | 0.12 | 0.06 |
| MSEW | 0 | 0 | 0 | 0.75 | 0 | 0 | 0 | 0.25 | 0 |

TABLE II

DISTANCE SPECTRUM WITH IDEAL A PRIORI INFORMATION.

### B. Analysis with mutual information: EXIT chart

Additional insight can be gained through the analysis of labeling maps based on the average mutual information between the coded bits and the output of the demapper as a function of the a priori input. The iterative exchange of this mutual information between the demapper and the decoder can be visualized by EXIT charts. For an introduction to EXIT charts we refer to [12].

Fig. 3 depicts the transfer functions in the EXIT chart for an AWGN channel of the three 16-QAM mappings presented in Section III-A, together with the transfer function of a 4-state, rate-$1/2$ convolutional code and a simulated trajectory which passes through the tunnel. Let us denote the values of the demapper functions with no a priori knowledge and ideal a priori knowledge by $I_0$ and $I_1$ respectively. High values of $I_0$ and $I_1$ are desirable in order to avoid an early crossing of the transfer functions, which would cause the iterative process to stop, and to reach low error rates respectively. As expected, Gray mapping has the highest $I_0$, MSEW mapping the highest $I_1$, while MSP is a good trade off, which can be explained as follows: the area $A$ under the transfer function of the demapper chart corresponds approximately to $A \approx C(\chi)/m$ [13], where $C(\chi)$ denotes the channel capacity of the signal constellation $\chi$. Since the capacity is independent of the applied mapping, high $I_0$ usually corresponds to low $I_1$ and vice versa. MSP mapping is particularly interesting because of its slightly curved shape, which allows both high $I_0$ and $I_1$.

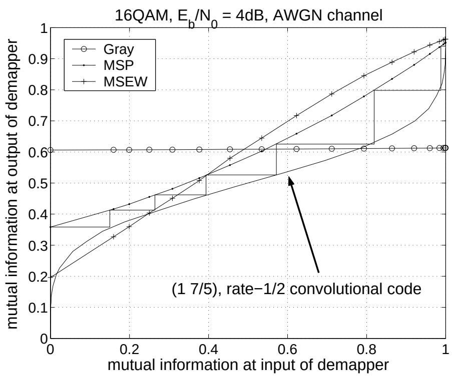



Fig. 3. Transfer functions of different mappings in the EXIT chart.

## IV. OPTIMIZATION OF THE INDEX ASSIGNMENTS

The goal is to have a method, where weights for the performance without and with ideal a priori knowledge can be set and the algorithm finds an optimal mapping for the selected weight distribution. Those mappings could be found through exhaustive search, which becomes intractable for higher order constellations since $2^m!$ different possibilities have to be checked. We propose a binary switching algorithm (BSA) to overcome the complexity problems of the brute-force approach. This algorithm finds a local optimum on a given cost function. First, the BSA for mappings is presented. Then, we investigate several possible cost functions.

### A. Binary switching algorithm

The binary switching algorithm (BSA) [11] is started with an initial mapping. Using one of the cost functions defined in the next sections, the cost of each symbol and the total cost are calculated. An ordered list of symbols, sorted by decreasing costs, is generated. The idea is to pick the symbol with the highest cost in the list (which has the strongest contribution to a "bad" performance), and to try to switch the index of this symbol with the index of another symbol. The latter is selected such, that the decrease of the total cost due to the switch is as large as possible. If no switch partner can be found for the symbol with the highest cost, the symbol with the second-highest cost will be tried to switch next. This process continues for symbols in the list with decreasing costs until a symbol is found that allows a switch that lowers the total cost. After an accepted switch, a new ordered list of symbols is generated, and the algorithm continues as described above until no further reduction of the total cost is possible. Several algorithm executions with random initial mappings yield to the presumed global optimum, as the BSA finds a local optimum.

### B. Cost function based on mutual information

We first consider the mutual information $I(C;R) \in [0,1]$ between the coded bits and the received channel output as a cost function for the BSA. This mutual information, also used in the EXIT charts [12], can be evaluated by numerical integration over the signal space $\mathbb{C}$:

$$I(C;R) = \frac{1}{2m} \cdot \sum_{i=1}^{m} \sum_{b=0}^{1} \int_{\mathbb{C}} p(a) \int_{\mathbb{C}} p(r|C(i)=b) \cdot \\ \cdot \log_2 \frac{2 \cdot p(r|C(i)=b)}{p(r|C(i)=0) + p(r|C(i)=1)} \, \mathrm{d}r \, \mathrm{d}a, \quad (7)$$

with

$$p(r|C(i)=b) = \frac{1}{2^{m-1}} \cdot \sum_{s \in \chi_b^i} p(r|s), \quad (8)$$

for independent and uniformly distributed code bits. For AWGN channels, the integration over the probability density function $p(a)$ of the fading coefficient $a$ can be omitted as $a = 1$. $p(r|s)$ is given by the Gaussian distribution 3.

The mutual information is a robust performance measure. However, without a priori information and without approximations, we can only give a separate performance measure for each bit position rather than for each symbol as required for the BSA. Thus, we will use the mutual information only for the case of ideal a priori information, where the subset $\chi_b^i$ in the summation of (8) is reduced to one symbol:

$$D^{Ir} = 1 - I(C;R) \text{ and } D^{Ia} = 1 - I(C;R)|_{a=1}\,, \quad (9)$$

for the Rayleigh and AWGN channel respectively.

*C. Cost functions based on error bounds*

Instead of using the mutual information computed by expensive numerical integration in (7) as cost function, we can use simple terms which characterize the influence of the mapping in error bounds. Our main goal is not to use tight error bounds, but to have a *qualitative* measure to define costs for mappings.

Let $P(s_k \rightarrow \hat{s}_k)$ denote the probability of choosing the symbol $\hat{s}_k$ instead of the transmitted symbol $s_k$. In the general case of a Rician fading channel with Rice-Factor $K$, the Chernoff upper bound of $P(s_k \rightarrow \hat{s}_k)$ is given by [14]:

$$P(s_k \rightarrow \hat{s}_k) \leq \frac{1+K}{1+K+\frac{E_s}{4N_0}|s_k-\hat{s}_k|^2} \times \exp\left(-\frac{K\frac{E_s}{4N_0}|s_k-\hat{s}_k|^2}{1+K+\frac{E_s}{4N_0}|s_k-\hat{s}_k|^2}\right). \quad (10)$$

Let $\mathbf{c}$ and $\hat{\mathbf{c}}$ denote two coded bit sequences which differ in $d$ consecutive positions. The bits $\mathbf{c}$ are transmitted within distinct symbols. $P(\mathbf{c} \rightarrow \hat{\mathbf{c}})$ is the *pairwise error probability* (PEP), i.e. the probability of choosing the sequence $\hat{\mathbf{c}}$ instead of the transmitted sequence $\mathbf{c}$. Assuming perfect interleaving and averaging over all symbols and bit positions, the PEP of the two sequences $\mathbf{c}$ and $\hat{\mathbf{c}}$ is given by [2]:

$$P(\mathbf{c} \rightarrow \hat{\mathbf{c}}) = \left(\frac{1}{m2^m}\sum_{i=1}^{m}\sum_{b=0}^{1}\sum_{s_k\in\chi_b^i}\sum_{\hat{s}_k\in\chi_{\bar{b}}^i} P(s_k \rightarrow \hat{s}_k)\right)^d. \quad (11)$$

For small values of $K$ (e.g. $K$=0, Rayleigh fading) and high SNR ($(1+K) \ll \frac{E_s}{4N_0}|s_k-\hat{s}_k|^2$), using (10) and (11), the influence of the mapping on the PEP is described by

$$D^r = \frac{1}{m2^m}\sum_{i=1}^{m}\sum_{b=0}^{1}\sum_{s_k\in\chi_b^i}\sum_{\hat{s}_k\in\chi_{\bar{b}}^i}\frac{1}{|s_k-\hat{s}_k|^2}. \quad (12)$$

The inverse of $D^r$ is interpreted in [2] as the *harmonic mean* of the Euclidean distance.

For large values of $K$ (e.g. $K \rightarrow \infty$, AWGN channel) or low SNR ($(1+K) \gg \frac{E_s}{4N_0}|s_k-\hat{s}_k|^2$), the influence of the mapping on the PEP is described by

$$D^a = \frac{1}{m2^m}\sum_{i=1}^{m}\sum_{b=0}^{1}\sum_{s_k\in\chi_b^i}\sum_{\hat{s}_k\in\chi_{\bar{b}}^i}\exp\left(-\frac{E_s}{4N_0}|s_k-\hat{s}_k|^2\right). \quad (13)$$

Depending on the channel, $D^r$ or $D^a$ can be used as cost functions for the BSA. Separate costs for each symbol, required by the BSA, can be easily determined by considering only the selected symbol $s_k$ in the summation in (12) and (13). We distinguish between the case without a priori knowledge ($D_0^r$, $D_0^a$) and with perfect a priori knowledge ($D_1^r$, $D_1^a$). In the latter case, the signal subset $\chi_{\bar{b}}^i$ in the summations (12) and (13) is reduced to one symbol. We propose to use the weighted combination

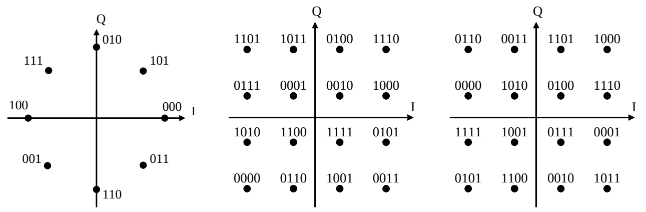
$$D^{Er} = \lambda_0 \cdot D_0^r + \lambda_1 \cdot D_1^r \text{ and } D^{Ea} = \lambda_0 \cdot D_0^a + \lambda_1 \cdot D_1^a \quad (14)$$

as a cost function for the Rayleigh and AWGN channel respectively; $\lambda_0$ and $\lambda_1$ denote the weights for the performance without and with perfect a priori knowledge.

## V. Results

The cost functions $D^{Ir}$, $D^{Ia}$ (9), $D^{Er}$ and $D^{Ea}$ (14) can be determined for arbitrary signal constellations. The BSA can output for different initial mappings and cost functions different optimized mappings with the same distance spectrum.

Setting $\lambda_0 = 1$ and $\lambda_1 = 0$ in (14) maximizes the performance without a priori knowledge. As expected, the BSA outputs Gray or quasi-Gray mappings for all signal constellations. By setting $\lambda_0 = 0$ and $\lambda_1 = 1$ in (14) or using (9), the performance with ideal a priori information is maximized. For 8-PSK, all cost functions result in the same optimum mapping *M8* shown in Fig. 4a), which is identical to the one proposed in [9] found by exhaustive search. Our optimization, however, took only a few seconds on a conventional PC, which shows the efficiency and the principal functionality of our algorithm. For 16-QAM, the BSA finds two new mappings shown in Fig. 4b) and 4c). *M16*$^a$ results from the cost functions $D^{Ia}$ and $D^{Ea}$ at high SNR. Due to the exponential cost decrease with the Euclidean distance and the SNR in AWGN channels, the minimum of the squared Euclidean distance is maximized. *M16*$^r$ results from the cost function $D^{Ir}$, $D^{Er}$ as well as from $D^{Ia}$ and $D^{Ea}$ at low SNR.




a) *M8* b) *M16*$^a$ c) *M16*$^r$

Fig. 4. Optimized mappings for ideal a priori information.

| $d_E^2$ | $\alpha$ | $2\alpha$ | $4\alpha$ | $5\alpha$ | $8\alpha$ | $9\alpha$ | $10\alpha$ | $13\alpha$ | $18\alpha$ |
|---|---|---|---|---|---|---|---|---|---|
| $M16^a$ | 1.75 | 1.31 | 1.25 | 1.25 | 0.5 | 0.75 | 0.75 | 0.25 | 0.19 |
| $M16^r$ | 1.75 | 1.31 | 1 | 1.75 | 0.25 | 0.75 | 0.5 | 0.5 | 0.19 |

TABLE III

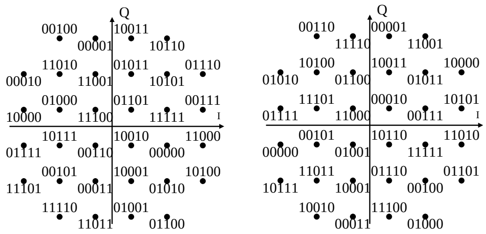
Distance spectrum with no a priori information.

| $d_E^2$ | $\alpha$ | $2\alpha$ | $4\alpha$ | $5\alpha$ | $8\alpha$ | $9\alpha$ | $10\alpha$ | $13\alpha$ | $18\alpha$ |
|---|---|---|---|---|---|---|---|---|---|
| $M16^a$ | 0 | 0 | 0 | 0.5 | 0.12 | 0 | 0.12 | 0.25 | 0 |
| $M16^r$ | 0 | 0 | 0.12 | 0.25 | 0.25 | 0 | 0.25 | 0.12 | 0 |

TABLE IV

Distance spectrum with ideal a priori information.

For 32-QAM, the BSA finds two new mappings, shown in Fig. 5. Similar to the mappings *M16*$^a$ and *M16*$^r$ for 16-QAM, the mapping *M32*$^a$ is optimal for AWGN channels at high SNR, *M32*$^r$ is optimal for Rayleigh channel or AWGN channel at low SNR.




a) *M32*$^a$ b) *M32*$^r$

Fig. 5. Optimized mappings for ideal a priori information.

Fig. 6 depicts the transfer functions in the EXIT chart for an AWGN channel of the mappings investigated in this section. *M16*$^a$ and *M32*$^a$ are expected to have a better performance at higher SNR than *M16*$^r$ and *M32*$^r$, but for the considered SNR, their performance is similar.

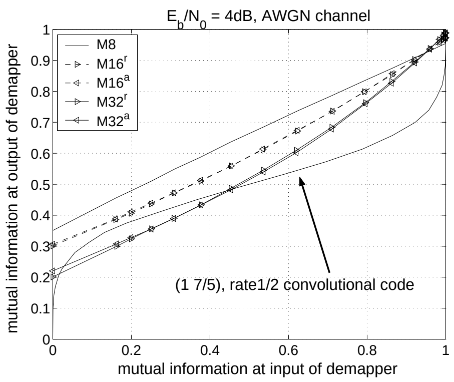



Fig. 6. Transfer functions of new mappings in the EXIT chart.

The BER performance of the investigated 16-QAM mappings is shown in Fig. 7 after 1 and 10 iterations, where the convolutional code is a 4-state, rate-1/2 code, the channel is an AWGN channel, and the interleaver length is 10000 bits. The analytic error bounds with ideal a priori information are computed as described in [2], [8] using the Gauss-Chebyshev method. The two optimized mappings *M16*$^a$ and *M16*$^r$ clearly outperform the other mappings once the a priori feedback information has a certain reliability. Since the new mappings are only optimized for ideal a priori information, other mappings can converge at lower SNR. The MSEW mapping is outperformed in the whole SNR range by the optimized mappings. Even though the *M16*$^r$ mapping is optimized for the fading channel, its performance is similar to the *M16*$^a$ mapping in the considered SNR range. All those relations are similar for a fading channel.

## VI. Conclusions

In this paper we have proposed a new method with low complexity which finds optimized mappings for BICM-ID. The method is based on a binary switching algorithm. New 16-QAM and 32-QAM mappings are given that outperform previously proposed ones at high SNR. The proposed method

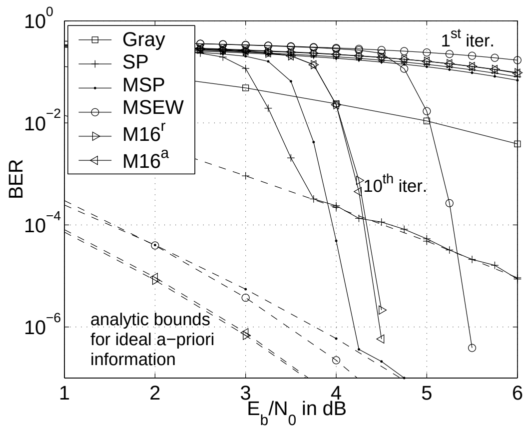



Fig. 7. BER of BICM-ID with 16-QAM mappings, AWGN channel, R=1/2, 4-state convolutional code (2-bit/channel use).

can be applied to any arbitrary signal constellation. Simulation results show that mappings optimized for fading channels are also well suited for AWGN channels.

## References

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

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

[3] G. Montorsi S. Benedetto, D. Divsalar and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 909–926, May 1998.

[4] X. Li and J. Ritcey, "Bit-interleaved coded modulation with iterative decoding using soft feedback," *Electronic Letters*, vol. 34, no. 10, pp. 942–943, May 1998.

[5] J. Speidel S. ten Brink and R. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. IEEE Globcom Conference*, Sydney, November 1998, pp. 579–584.

[6] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error–correcting coding and decoding: Turbo-codes," in *IEEE International Conference on Communications (ICC)*, Geneva, Switzerland, May 1993, pp. 1064–1070.

[7] S. ten Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," *AEÜ International Journal of Electronics and Communications*, vol. 54, no. 6, pp. 389–398, November 2000.

[8] A. Chindapol and J. Ritcey, "Design, analysis and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 5, pp. 944–957, May 2001.

[9] A. Chindapol and J. Ritcey, "Bit-interleaved coded modulation with iterative decoding and 8PSK signaling," *IEEE Transactions on Communications*, vol. 50, no. 8, pp. 1250–1257, August 2002.

[10] J. Tan and G.L. Stüber, "Analysis and design of interleaver mappings for iteratively decoded BICM," in *IEEE International Conference on Communications (ICC)*, New York, May 2002, vol. 3, pp. 1403–1407.

[11] K. Zeger and A. Gersho, "Pseudo-Gray coding," *IEEE Transactions on Communications*, vol. 38, pp. 2147–2158, December 1990.

[12] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, October 2001.

[13] S. ten Brink, "Exploiting the chain rule of mutual information for the design of iterative decoding schemes," in *Proc. 39th Allerton Conf. on Communications, Control and Computing*, Monticello, Urbana-Champaign, October 2001.

[14] S. Benedetto and E. Biglieri, *Principles of Digital Transmission*, Kluwer Academic / Plenum publishers, 1999.