PSK versus QAM for Iterative Decoding of Bit-Interleaved Coded Modulation

Thorsten Clevorn, Susanne Godtmann, and Peter Vary
Institute of Communication Systems and Data Processing (ind)
RWTH Aachen University, Germany
{clevorn,vary}@ind.rwth-aachen.de

***Abstract*— Bit-interleaved coded modulation with iterative decoding (BICM-ID) is a transmission scheme with excellent performance on fading channels. The mapping of the bit patterns to the elements of the signal constellation set (SCS) is the key design element of BICM-ID, especially regarding the asymptotic performance of the BER. In this paper we will demonstrate that SCSs of PSK (phase shift keying) with optimum bit pattern assignment are significantly superior in the asymptotic error-floor region to the respective square QAM (quadrature amplitude modulation) SCSs. When modifying the PSK SCSs to non-regular SCSs performance gains more than 1 dB are realizable in case of 4 bits per channel symbol. The achievable gains are theoretically determined and analyzed, and verified by simulations.**

I. Introduction

In digital communications the transmitted signal is often affected by fast fading due to, e.g., multi-path propagation. *Bit-interleaved coded modulation* (BICM) [1],[2] is a band-width efficient coded modulation scheme which increases the time-diversity. The key element of BICM is the serial concatenation of channel coding, bit-interleaving, and multilevel modulations, a typical design of todays communication systems [3]. In order to increase the performance of BICM, in [4],[5],[6],[7] a feedback loop is added to the decoder, which results in a turbo-like decoding process [8],[9]. This new scheme is known as BICM with *iterative decoding* (BICM-ID). When using non-Gray mappings in the modulator significant improvements by the iterations can be achieved. Several square 16QAM (quadrature amplitude modulation) mappings have been proposed for BICM-ID, e.g., in [10],[11],[12]. In [11],[12] an optimization algorithm is presented. By *mapping* we denote the whole design process of locating the possible positions of the channel symbols in the signal space, the *signal constellation set* (SCS), and assigning the possible bit patterns to these symbols.

For further improvement of BICM-ID we propose to use PSK (phase shift keying) mappings instead of square QAM mappings. We will show that the SCSs of PSK are much more suited for the optimization process of the mapping due to the fact that the channel symbols are located all on a circle. Additionally, PSK mappings allow an easy application of the *non-regular* SCSs technique proposed in [13],[14] which provides further gains for BICM-ID. Optimum mappings are presented for a 16PSK SCS. The advantages of PSK SCSs are examined and visualized by error bounds and the *mutual information* measure. EXIT charts [14],[15],[16],[17] confirm the convergence behavior and simulations demonstrate the achievable performance gains.

II. The BICM-ID System

Fig. 1 depicts the baseband model of the BICM-ID system considered in this paper. A block of data bits u is encoded by a standard non-systematic feed-forward convolutional encoder. The resulting encoded bits x are permuted by a pseudo-random bit-interleaver π to $\tilde{x}$ and grouped consecutively into bit patterns $\underline{\tilde{x}}_t = [\tilde{x}_t^{(1)}, \ldots \tilde{x}_t^{(I)}]$, where $\tilde{x}_t^{(i)}$ denotes the i^{th} bit in the bit pattern at time index t, $t = 1, \ldots T$. I is the number of bits that will be mapped to one channel symbol later on, e.g., $I = 4$ in case of 16QAM.

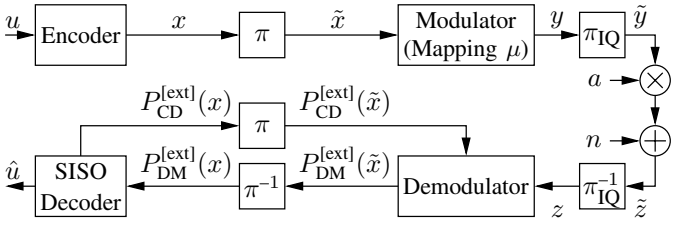


Fig. 1. Baseband model of the BICM-ID system.

The modulator maps an interleaved bit pattern $\underline{\tilde{x}}_t$ according to a mapping rule μ to a complex channel symbol y_t out of the signal constellation set (SCS) $\mathcal{Y}$

$$y_t = \mu(\underline{\tilde{x}}_t) \quad . \tag{1}$$

The respective inverse relation is denoted by μ^{-1}, with

$$\underline{\tilde{x}}_t = \mu^{-1}(y_t) = [\mu^{-1}(y_t)^{(1)}, \ldots \mu^{-1}(y_t)^{(I)}] \quad . \tag{2}$$

The channel symbols are normalized to an average energy of $E_s = E\{\|y_t\|^2\} = 1$. In case of Rayleigh fading an IQ interleaver π_{IQ} [10] can be applied to the modulated symbols. With an IQ interleaver the in-phase (I) and quadrature (Q) component of a modulated symbol shall be made to fade independently. For a memoryless fading channel a simple wraparound shift by one symbol of the Q component is sufficient [10],

$$\tilde{y}_t = \begin{cases} \text{Re}\{y_t\} + j \cdot \text{Im}\{y_{t-1}\} & 2 \le t \le T \\ \text{Re}\{y_t\} + j \cdot \text{Im}\{y_T\} & t = 1 \end{cases} \quad . \tag{3}$$

The transmitted symbols $\tilde{y}_t$ are faded by the Rayleigh distributed coefficients a_t with $E\{\|a_t\|^2\} = 1$. In this paper we assume that these coefficients are known at the receiver, i.e., perfect knowledge of the channel state information (CSI). The effects of imperfect CSI on BICM-ID have been studied,

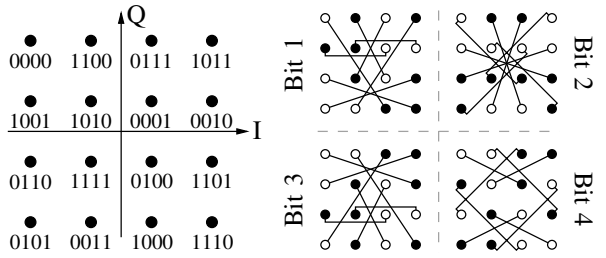


Fig. 2. 16QAM-R mapping with signal-pair distances $\|y-\check{y}\|$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

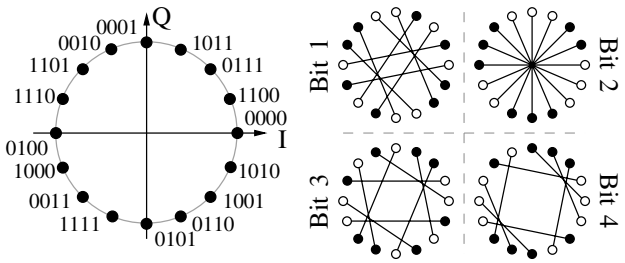


Fig. 3. 16PSK-R mapping with signal-pair distances $\|y-\check{y}\|$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

e.g., in [18]. Next, complex zero-mean white Gaussian noise $n_t = n_t' + jn_t''$ with a known power spectral density of $\sigma_n^2 = N_0$ ($\sigma_{n'}^2 = \sigma_{n''}^2 = N_0/2$) is added. Thus, the received channel symbols $\tilde{z}_t$ can be written as

$$\tilde{z}_t = a_t \tilde{y}_t + n_t \quad . \tag{4}$$

After the IQ deinterleaver π_{IQ}^{-1} is applied the demodulator (DM) computes *extrinsic* probabilities $P_{\text{DM}}^{[\text{ext}]}(\tilde{x})$ for each bit $\tilde{x}_t^{(i)}$ being $b \in \{0,1\}$ according to [6]

$$P_{\text{DM}}^{[\text{ext}]}(\tilde{x}_t^{(i)}=b) \sim \sum_{\hat{y}\in\mathcal{Y}_b^i} P(z_t|\hat{y}) P_{\text{CD}}^{[\text{ext,i}]}(\hat{y}) \quad , \tag{5}$$

$$\text{with} \quad P_{\text{CD}}^{[\text{ext,i}]}(\hat{y}) \triangleq \prod_{j=1, j\neq i}^{I} P_{\text{CD}}^{[\text{ext}]}\left(\tilde{x}_t^{(j)} = \mu^{-1}(\hat{y})^{(j)}\right) \quad . \tag{6}$$

Each $P_{\text{DM}}^{[\text{ext}]}(\tilde{x})$ consists of the sum over all possible channel symbols $\hat{y}$ for which the i^{th} bit of the corresponding bit pattern $\underline{\tilde{x}} = \mu^{-1}(\hat{y})$ is b. These channel symbols form the subset $\mathcal{Y}_b^i$ with $\mathcal{Y}_b^i = \{\mu([\tilde{x}^{(1)}, \ldots \tilde{x}^{(I)}]) | \tilde{x}^{(i)} = b\}$. In the first iteration the feedback probabilities $P_{\text{CD}}^{[\text{ext}]}(\tilde{x})$ are initialized as equiprobable, i.e., $P_{\text{CD}}^{[\text{ext}]}(\tilde{x}) = 0.5$. In case of IQ interleaving the I and Q component of z are faded independently. Thus, the conditional probability density $P(z_t|\hat{y})$ describing the channel is given by

$$P(z_t|\hat{y}) = \frac{1}{\pi\sigma_n^2} \exp\left(-\frac{d_{z\hat{y}}^2}{\sigma_n^2}\right) \quad \text{, with} \tag{7}$$

$$d_{z\hat{y}}^2 = |\text{Re}\{z_t\} - a_t \text{Re}\{\hat{y}\}|^2 + |\text{Im}\{z_t\} - a_{t-1}\text{Im}\{\hat{y}\}|^2. \tag{8}$$

Without IQ interleaving (8) simplifies to $d_{z\hat{y}}^2 = \|z_t - a_t\hat{y}\|^2$.

After appropriately deinterleaving the $P_{\text{DM}}^{[\text{ext}]}(\tilde{x})$ to $P_{\text{DM}}^{[\text{ext}]}(x)$, the $P_{\text{DM}}^{[\text{ext}]}(x)$ are fed into a Soft-Input Soft-Output (SISO) channel decoder (CD) [19], which computes *extrinsic* probabilities $P_{\text{CD}}^{[\text{ext}]}(x_t^{(i)})$ of the encoded bits $x_t^{(i)} = \{0,1\}$ in addition to the preliminary estimated decoded data bits $\hat{u}$. For the next iteration the $P_{\text{CD}}^{[\text{ext}]}(x)$ are interleaved again to $P_{\text{CD}}^{[\text{ext}]}(\tilde{x})$ in order to be fed into the Demodulator.

III. Mappings for Rayleigh channels

Since for non-iterative BICM Gray mappings are optimal [1],[2], we concentrate on the analysis of the performance in the iterative case with BICM-ID. If the channel is sufficiently good and enough iterations are carried out we can assume *error-free feedback* (EFF) [5],[6], i.e., we assume that the feedback $P_{\text{CD}}^{[\text{ext}]}(\hat{x})$ for all bits except the one currently considered is correct with perfect reliability (best case). Using the perfectly reliable *extrinsic* information in the EFF results in a BPSK decision between the two channel symbols y and $\check{y}$, which posses identical bit patterns except for the position i of the currently considered bit. Here, the bit $\mu^{-1}(\check{y})^{(i)}$ is inverted with respect to y. On the right side of Fig. 2 the resulting BPSK decisions are depicted for the mapping we denote by 16QAM-R (see Section III-B, "-R" indicating Rayleigh).

A. Performance bound

For the case of an AWGN channel with Rayleigh fading and no IQ interleaving the performance bound for the bit-error rate (BER) P_b of BICM-ID is derived, e.g., in [6]. The asymptotic behavior of P_b can be described by

$$\log_{10} P_b \approx \frac{-d_{\text{Ham}}(\mathcal{C})}{10} \left[\left(R \cdot \check{d}_h^2(\mu)\right)_{\text{dB}} + \left(\frac{E_b}{N_0}\right)_{\text{dB}} \right] + const. \tag{9}$$

The minimum Hamming distance $d_{\text{Ham}}(\mathcal{C})$ of the channel code $\mathcal{C}$ with rate r defines the slope of the bound. Whereas $\check{d}_h^2(\mu)$ is the *harmonic mean* of the *squared Euclidean distances* of the BPSK decisions with EFF. The product of the information rate $R = rI$ and the *harmonic mean* $\check{d}_h^2(\mu)$ determines an E_b/N_0 offset of the bound, i.e., a horizontal offset in a BER vs. E_b/N_0 plot. $E_b = E_s/R$ is the energy per information bit. For a mapping with I bits the *harmonic mean* $\check{d}_h^2(\mu)$ is given by

$$\check{d}_h^2(\mu) = \left(\frac{1}{I2^I} \sum_{i=1}^{I} \sum_{b=0}^{1} \sum_{y\in\mathcal{Y}_b^i} \frac{1}{\|y-\check{y}\|^2} \right)^{-1}. \tag{10}$$

In case of non-iterative BICM, i.e., no feedback, $\check{d}_h^2(\mu)$ has to be replaced by the harmonic mean $\bar{d}_h^2(\mu)$. $\bar{d}_h^2(\mu)$ is computed similar to $\check{d}_h^2(\mu)$, except that $\check{y}$ is replaced by $\bar{y}$, where $\bar{y}$ denotes the nearest neighbor y with an inverted bit at position i, whatever the rest of the bit pattern contains.

The ratio $\mathcal{G}(\mu) = (\check{d}_h^2(\mu)/\bar{d}_h^2(\mu)^{[\max]})_{\text{dB}}$ is called the *offset gain* (OG) [6] and describes the possible gain of BICM-ID with a certain mapping with respect to the optimum mapped non-iterative BICM in case of Rayleigh fading. The optimum mapping for the non-iterative BICM is usually a Gray mapping, i.e., the 16QAM-Gray mapping (e.g. [10]) with $\bar{d}_h^2(\mu)^{[\max]} = \bar{d}_h^2(\text{16QAM-Gray}) = 0.492$. The *offset gain* $\mathcal{G}(\mu)$ and the corresponding $\check{d}_h^2(\mu)$ serve as key indicators for the evaluation of the theoretically achievable performance of mappings with BICM-ID [6],[13]. In [11],[12] the inverse of $\check{d}_h^2(\mu)$ denoted by $D_1^r(\mu) = 1/\check{d}_h^2(\mu)$ is used in a similar way.

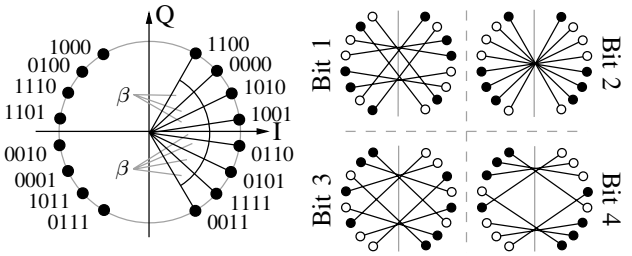


Fig. 4. β-16PSK-R mapping with signal-pair distances $\|y-\check{y}\|$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

TABLE I

$\check{d}_h^2$, D_1^r AND OFFSET GAINS $\mathcal{G}$ FOR MAPPINGS WITH $I=4$ BITS.

mapping μ	D_1^r	$\check{d}_h^2$	offset gain $\mathcal{G}$
16QAM-Gray [10]	1.944	0.514	0.19 dB
16QAM-R (Fig. 2) $\widehat{=}$ *M16^r* [11],[12]	0.368	2.719	7.42 dB
16PSK-R (Fig. 3)	0.321	3.114	8.01 dB
17°-16PSK-R (Fig. 4, $\beta=17°$)	0.285	3.509	8.53 dB
BPSK bound [13] (Fig. 4, $\beta\to 0°$)	0.25	4	9.10 dB

B. Optimization of 16QAM mappings

To optimize the asymptotic behavior of the bit pattern assignment for a certain SCS, i.e., increase the obtainable *offset gain* by the mapping, $\check{d}_h^2(\mu)$ has to be maximized or respectively $D_1^r(\mu)$ has to be minimized. There exist $(2^I)!$ different bit pattern assignments. In [11],[12] a binary switching algorithm (BSA) has been presented for this task, finding a (possibly local) optimum. For $I\leq 4$ we developed a recursive exhaustive search algorithm, which finds the global optimum. It successively places a bit pattern in each recursion and immediately adds the available costs of new fixed signal-pair distances. Recursions with no chance of being optimal anymore are adaptively omitted, i.e., search tree pruning. The algorithm requires only several seconds or a few minutes on a regular PC for $I=4$. Fig. 2 depicts the optimum mapping for a 16QAM SCS. We denote this mapping by 16QAM-R. It is similar to the *M16^r* mapping in [11],[12] and provides an *offset gain* of $\mathcal{G}(\text{16QAM-R})=7.42$ dB.

C. 16PSK signal constellation set

However, the drawback of the 16QAM SCS are the four channel symbols in the center of the SCS. They imply small signal-pair distances $\|y-\check{y}\|$ which dominate the optimization with $\check{d}_h^2(\mu)$ or $D_1^r(\mu)$. Thus, if interested in superior asymptotic performance, we propose the usage of the 16PSK SCS. Using the optimization algorithm described above, we obtained the 16PSK-R mapping presented in Fig. 3. As visible, with all channel symbols on the unit circle the minimum signal-pair distances $\|y-\check{y}\|$ are significantly increased. With an *offset gain* of $\mathcal{G}(\text{16PSK-R})=8.01$ dB it outperforms the optimum 16QAM mapping (16QAM-R) by ≈ 0.6 dB.

Another nice feature of PSK SCSs is the easy application of the technique of *non-regular signal constellation sets* [13],[14] for further improvement of the asymptotic performance. As an example for a *non-regular* SCS we consider the β-16PSK-R mapping depicted in Fig. 4. It splits the regular SCS of the 16PSK-R mapping into two parts and adjusts the phase angle between the channel symbols of each part to β. For $\beta=22.5°$ the β-16PSK-R mapping is identical to a rotated 16PSK-R mapping. However, as visible on the right side of Fig. 4 for a smaller β all signal-pair distances $\|y-\check{y}\|$ are increased since they all cross the split axis between the two parts. For $\beta=17°$ and $\mathcal{G}(17°\text{-16PSK-R})=8.53$ dB, e.g., we can obtain ≈ 0.5 dB additional improvement of the *offset gain*. With $\beta\to 0°$ we would reach the BPSK bound derived in [13],[14]. Also other *non-regular* SCSs can be constructed. When using the method of assigning several bit patterns to a single channel symbol [13] possible SCSs are, e.g., 12QAM (not using the center points of 16QAM) or 12PSK. But we restrict the discussion in this paper to the β-16PSK SCS with unequal distributed channel symbols as *non-regular* feature.

Table I lists the *offset gains* and the values of D_1^r and $\check{d}_h^2$ for the mappings with $I=4$ bits considered for Rayleigh fading. In the asymptotic error-floor region the 16PSK mappings significantly outperform the 16QAM mappings. With 17°-16PSK-R mapping we approach the BPSK bound by 0.57 dB.

D. Mutual information analysis

The asymptotic performance of different mappings can also be analyzed by the *mutual information* measure, either using computational expensive numerical integration (described, e.g., in [11],[12] as alternative to the efficient D_1^r cost function) or simply by the histogram method of EXIT charts [14]. The EFF corresponds to perfect a-priori *mutual information* for the demodulator, i.e., $\mathcal{I}_{\text{DM}}^{[\text{apri}]}=1$, for the EXIT characteristics of a mapping. The obtained extrinsic *mutual information* $\mathcal{I}_{\text{DM}}^{[\text{ext}]}(\mathcal{I}_{\text{DM}}^{[\text{apri}]}=1)$ converges towards 1 for high E_s/N_0. For a better visualization we plot $1-\mathcal{I}_{\text{DM}}^{[\text{ext}]}(\mathcal{I}_{\text{DM}}^{[\text{apri}]}=1)$ on a logarithmic scale versus E_s/N_0 in Fig. 5. The horizontal offsets between the curves for no IQ interleaving match the predicted *offset gains* given in Table I. The curves with the solid markers depict the results when IQ interleaving is applied. The relative ordering of the mappings is similar to the case without IQ interleaving. Note, to benefit from IQ interleaving the SCS of 16QAM-Gray mapping needs to be rotated [10], e.g., by $45°$.

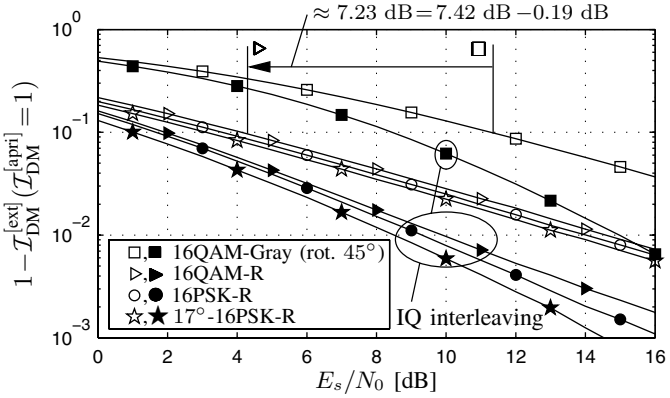


Fig. 5. $\mathcal{I}_{\text{DM}}^{[\text{ext}]}$ for mappings with EFF, i.e., $\mathcal{I}_{\text{DM}}^{[\text{apri}]}=1$, Rayleigh channel (solid markers when IQ interleaving is used).

E. Simulation results

The results of a simulation of the BER performance for data bits u are depicted in Fig. 6. The block size is set to 12000 information bits per frame and 30 iterations are carried out. As channel code serves a 4-state, rate-1/2 convolutional code with generator polynomials $\{7,5\}_8$ and $d_{\text{Ham}}(\mathcal{C}) = 5$. The solid markers indicate the usage of IQ interleaving. As expected the proposed 16PSK mappings noticeably outperform the optimal 16QAM mapping in the asymptotic error-floor region. In this region the horizontal offsets between the curves for no IQ interleaving coincide with predicted *offset gains* given in Table I and the offsets visible in Fig. 5. Using IQ interleaving further improves the performance of all mappings significantly. The convergence behavior is confirmed by the EXIT chart in Fig. 7. The depicted trajectory for the 16PSK-R mapping shows that at $E_b/N_0 = 8.5$ dB only approximately 7 iterations are required to reach the error-floor.

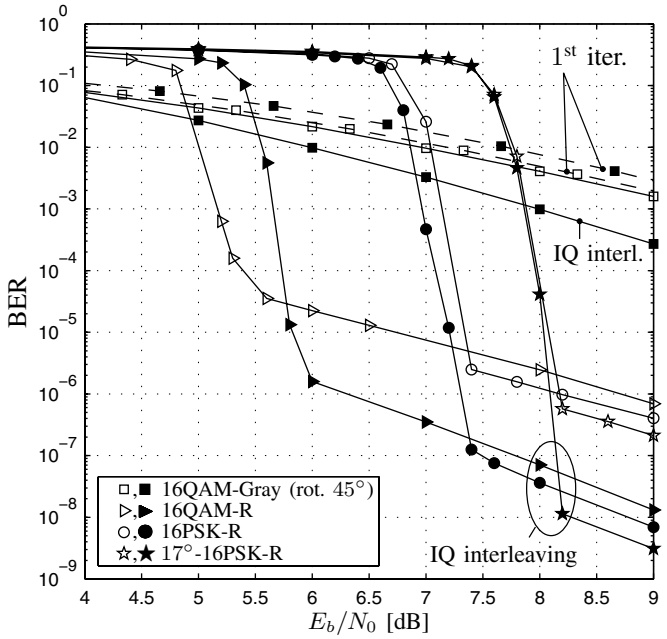


Fig. 6. BER for $I = 4$, 4-state, rate-1/2 code, 30 iterations, Rayleigh channel (solid markers when IQ interleaving is used).
Shannon limits $(E_b/N_0)_{\min}$ for a rate-1/2 code and Rayleigh fading: 3.92 dB (16QAM), 5.00 dB (16PSK), 5.76 dB (17°-16PSK)

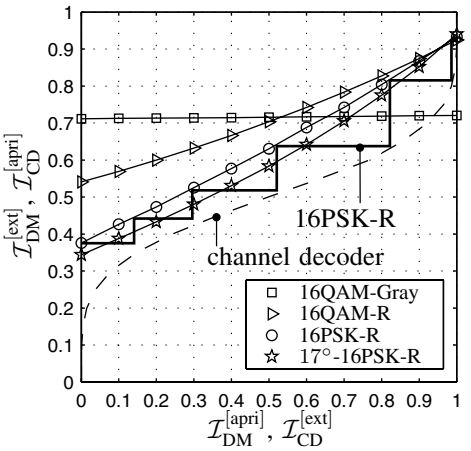
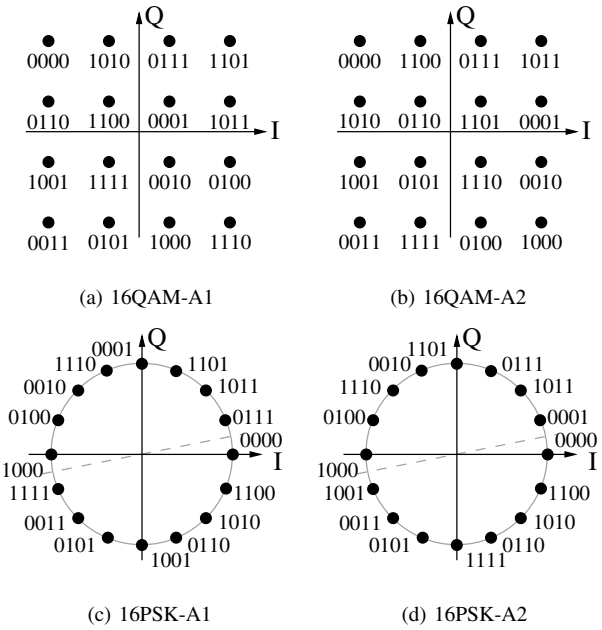


Fig. 7. EXIT chart for Fig. 6 at $E_b/N_0 = 8.5$ dB.



(a) 16QAM-A1 (b) 16QAM-A2

(c) 16PSK-A1 (d) 16PSK-A2

Fig. 8. Optimized mappings for 16QAM and 16PSK SCSs.

IV. Mappings for AWGN channels

For the case of a non-fading AWGN channel the cost function corresponding to D_1^r for the optimization of a mapping by minimization is [11],[12]

$$D_1^a = \frac{1}{I2^I} \sum_{i=1}^{I} \sum_{b=0}^{1} \sum_{y \in \mathcal{Y}_b^i} \exp\left(-\frac{E_s}{4N_0}\|y - \tilde{y}\|^2\right) \quad . \quad (11)$$

Thus, the optimum mapping now depends on E_s/N_0. For a 16QAM SCS and $E_s/N_0 = 0$ dB the optimization routine returns the 16QAM-A1 mapping ("-A" indicating AWGN) depicted in Fig. 8(a). This mapping is similar to the *M16*^{a} mapping in [11],[12]. However, for $E_s/N_0 < -0.4$ dB the 16QAM-A2 mapping (Fig. 8(b)) and for 0.2 dB $< E_s/N_0 < 7.6$ dB the 16QAM-R mapping (Fig. 2) are optimal for a 16QAM SCS, though the differences, especially for the latter one, are small. Using the 16QAM-A1 mapping as reference the difference $D_1^a(\text{16QAM-A1}) - D_1^a(\mu)$ is plotted in Fig. 9. Thus, a positive value of this difference indicates a superior asymptotic performance of the mapping μ.

Also depicted in Fig. 9 are the values for several mappings with a 16PSK SCS, which all significantly outperform the 16QAM mappings for almost all E_s/N_0. The 16PSK-R mapping (Fig. 3) has been presented in Section III. The 16PSK-A1 mapping and the 16PSK-A2 mapping are depicted in Figs. 8(c) and 8(d), with the dashed line indicating the split axis if *non-regular* SCSs as presented in Section III-C shall be used. The 17°-16PSK-R mapping with a *non-regular* SCS given in Fig. 4 surpasses the 16QAM mappings and the *regular* 16PSK mappings noticeably. For Fig. 10 the results of Fig. 9 are normalized again, this time by $D_1^a(\text{16QAM-A1}) - D_1^a(\text{16PSK-A1})$, the difference between the two optimum mappings at $E_s/N_0 = 0$ dB and regular

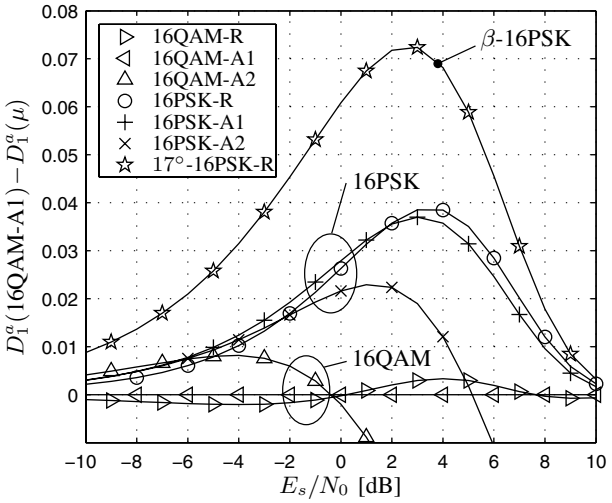


Fig. 9. Mappings with $I = 4$ bits and a non-fading AWGN channel.

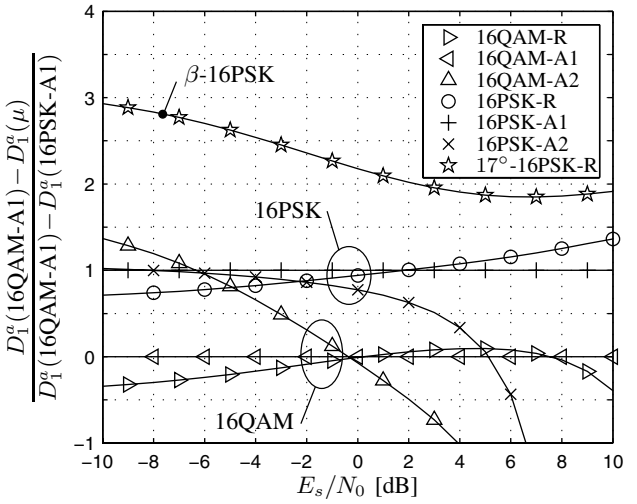


Fig. 10. Mappings with $I = 4$ bits and a non-fading AWGN channel.

SCSs, to avoid the scale problem for different E_s/N_0 and better illustrate the differences at low and high E_s/N_0.

In Fig. 11 the *mutual information* $\mathcal{I}_{\text{DM}}^{[\text{ext}]}(\mathcal{I}_{\text{DM}}^{[\text{apri}]} = 1)$ is plotted for an AWGN channel in a similar way as in Fig. 5 for the Rayleigh channel. The depicted curves confirm the results presented in Figs. 9 and 10. The obtainable gains depend on E_s/N_0 as implied by (11).

V. Conclusion

In this paper we showed that the capabilities of BICM-ID in the asymptotic error-floor region can be significantly improved when PSK signal constellation sets are used instead of QAM signal constellation sets. The possible performance gains are analyzed and presented by several means using, e.g., error bounds and the *mutual information* measure. Optimum 16PSK mappings for different channel scenarios, obtained by a search algorithm, are presented. Furthermore, the PSK mappings can be easily modified to *non-regular* signal constellation sets for additional enhancement. The theoretically predicted gains of more than 1 dB in E_b/N_0 are verified by EXIT charts and BER simulations.

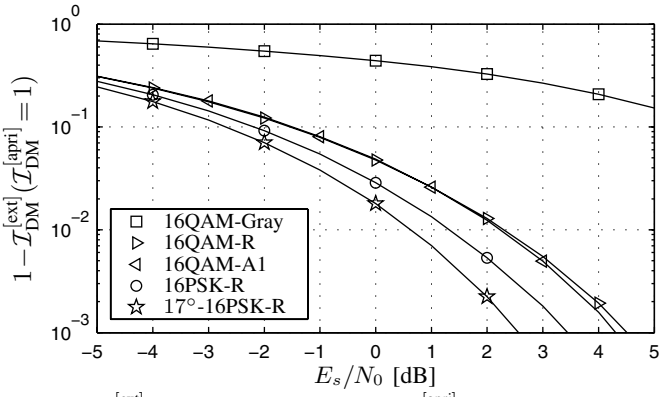


Fig. 11. $\mathcal{I}_{\text{DM}}^{[\text{ext}]}$ for mappings with EFF, i.e., $\mathcal{I}_{\text{DM}}^{[\text{apri}]} = 1$, AWGN channel.

References

[1] E. Zehavi, "8-PSK Trellis Codes for a Rayleigh Channel," *IEEE Trans. Comm.*, pp. 873–884, May 1992.

[2] G. Caire, G. Taricco, and E. Biglieri, "Bit-Interleaved Coded Modulation," *IEEE Trans. Inform. Theory*, pp. 927–946, May 1998.

[3] J. Huber, U. Wachsmann, and R. Fischer, "Coded Modulation by Multilevel-Codes: Overview and State of the Art," in *2nd Intern. ITG Conf. on Source and Channel Coding (SCC)*, Aachen, Germany, Mar. 1998.

[4] X. Li and J. A. Ritcey, "Bit-Interleaved Coded Modulation with Iterative Decoding," *IEEE Comm. Lett.*, pp. 169–171, Nov. 1997.

[5] X. Li and J. A. Ritcey, "Trellis-Coded Modulation with Bit Interleaving and Iterative Decoding," *IEEE J. Select. Areas Commun.*, pp. 715–724, Apr. 1999.

[6] X. Li, A. Chindapol, and J. A. Ritcey, "Bit-Interleaved Coded Modulation With Iterative Decoding and 8PSK Signaling," *IEEE Trans. Comm.*, pp. 1250–1257, Aug. 2002.

[7] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping for QPSK modulation," *IEEE Electr. Lett.*, pp. 1459–1460, July 1998.

[8] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: Turbo-Codes," *IEEE Trans. Comm.*, pp. 1261–1271, Oct. 1996.

[9] J. Hagenauer, E. Offer, and L. Papke, "Iterative Decoding of Binary Convolutional Codes," *IEEE Trans. Comm.*, pp. 429–445, Mar. 1996.

[10] A. Chindapol and J. A. Ritcey, "Design, Analysis, and Performance Evaluation for BICM-ID with Square QAM Constellations in Rayleigh Fading Channels," *IEEE J. Select. Areas Commun.*, pp. 944–957, May 2001.

[11] F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, "Optimized Symbol Mappings for Bit-Interleaved Coded Modulation with Iterative Decoding," in *Globecom 2003*, San Francisco, Dec. 2003.

[12] F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, "Optimization of Symbol Mappings for Bit-Interleaved Coded Modulation With Iterative Decoding," *IEEE Comm. Lett.*, pp. 593–595, Dec. 2003.

[13] T. Clevorn and P. Vary, "Iterative Decoding of BICM with Non-Regular Signal Constellation Sets," in *5th Intern. ITG Conf. on Source and Channel Coding (SCC)*, Erlangen, Germany, Jan. 2004.

[14] T. Clevorn, S. Godtmann, and P. Vary, "EXIT Chart Analysis of Non-Regular Signal Constellation Sets for BICM-ID," in *International Symposium on Information Theory and its Applications (ISITA 2004)*, Parma, Italy, Oct. 2004.

[15] S. ten Brink, "Convergence of Iterative Decoding," *IEEE Electr. Lett.*, pp. 806–808, May 1999.

[16] S. ten Brink, "Designing Iterative Decoding Schemes with the Extrinsic Information Transfer Chart," *International Journal of Electronics and Communications (AEÜ)*, pp. 389–398, Dec. 2000.

[17] S. ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes," *IEEE Trans. Comm.*, pp. 1727–1737, Oct. 2001.

[18] Y. Huang and J. A. Ritcey, "16-QAM BICM-ID in Fading Channels With Imperfect Channel State Information," *IEEE Trans. Wireless Comm.*, pp. 1000–1007, Sep. 2003.

[19] S. Benedetto, G. Montorsi, D. Divsalar, and F. Pollara, "Soft-Input Soft-Output Modules for the Construction and Distributed Iterative Decoding of Code Networks," *European Transactions on Telecommunications (ETT)*, pp. 155–172, Mar. 1998.