Iterative Decoding of BICM with Non-Regular Signal Constellation Sets

Thorsten Clevorn and Peter Vary

Institute of Communication Systems and Data Processing (ind)
Aachen University (RWTH), Germany
Email: {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. A key design element, especially regarding the asymptotic performance of the BER, is the mapping of the bit patterns to the elements of the signal constellation set. In this paper we present mappings and new non-regular signal constellation sets with superior asymptotic performance. These proposed mappings use signal constellation sets with less elements than the number of bit patterns to be transmitted and an appropriate assignment of the bit patterns to them. The puncturing introduced by the not one-to-one relation between assigned bit patterns and the elements of the signal constellation set is compensated by the usage of extrinsic information obtained in the iterations. The convergence behavior is confirmed by EXIT charts. We further show the theoretical limits for the improvement of BICM-ID by this technique.**

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 and consequently improves the performance on memoryless Rayleigh fading channels.

The key element of BICM is the serial concatenation of channel encoding, bit-interleaving, and multilevel modulations at the transmitter. The respective receiver comprises demapping by metric computation, deinterleaving, and channel decoding. The bit-interleaving is the major difference to *trellis-coded modulation* (TCM) [3], which uses symbol-wise interleaving and gives better results on non-fading Gaussian channels. BICM shows superior performance on fading channels since the bit-wise deinterleaver separates the coded bits of a channel symbol, which is distorted by fading.

To further increase the performance of BICM in [4], [5], [6] a feedback loop is added to the decoder, which results in a turbo-like decoding process [7], [8]. This new scheme is known as BICM with *iterative decoding* (BICM-ID). A similar approach is presented in [9]. Another key element of BICM-ID is the usage of non-Gray mappings in the modulator, since only with these a significant improvement by the iterations can be achieved. By *mapping* we will denote the whole design process of locating the possible positions of the channel symbols in the signal space, the *signal constellation set*, and assigning the possible bit patterns to these symbols.

In this paper we will develop fundamentally new mappings with *non-regular signal constellation set* to improve the capabilities of BICM-ID. These mappings owe their superior performance to a relation between channel symbols and assigned bit-patterns which is not one-to-one and unequally distributed positions of the channel symbols in the signal space. The achievable performance gains are demonstrated by simulations. The convergence behavior is verified by EXIT charts [10], [11], [12]. Additionally a theoretical limit for the improvement of BICM-ID by the proposed technique is derived.

II. The BICM-ID System

Fig. 1 depicts the baseband model of the BICM-ID system used in this paper. The data bits u are encoded by a standard non-systematic feed-forward convolutional encoder. The resulting encoded bits are grouped consecutively into bit patterns $\underline{x}_t = [x_t^{(1)}, \dots x_t^{(I)}]$, where $x_t^{(i)}$ denotes the ith bit in the bit pattern at time index t, $t = 1, \dots T$. I is the number of bits that will be mapped to one channel symbol later on, e.g., $I = 3$ in case of 8PSK (phase shift keying).

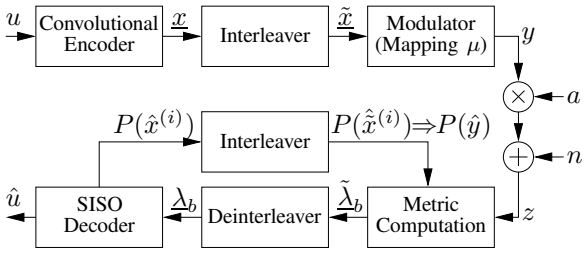

Fig. 1. Baseband model of the BICM-ID system.

Afterwards the bit patterns $\underline{x}_t$ are permuted by a pseudo-random block interleaver to $\underline{\tilde{x}}_t = [\tilde{x}_t^{(1)}, \dots \tilde{x}_t^{(I)}]$. The interleaving is done bit-wise in the time domain t, meaning that each bit $x_t^{(i)}$ of $\underline{x}_t$ is interleaved separately to different symbols $\underline{\tilde{x}}_t$, which are far apart in time. Note that the position i of a bit in a bit pattern is not changed. Therefore, the interleaver practically consists

of I separate different sub-interleavers, one for each of the I bit positions [1], [6].

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* $\mathcal{Y}$

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

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 $E\{\|y_t\|^2\}=1$.

On the channel the transmitted symbols 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 effect of imperfect CSI on BICM-ID has been studied in [13], [14]. Next, complex 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 z_t can be written as

$$z_t = a_t y_t + n_t \quad . \tag{3}$$

At the receiver in each iteration for each bit $\tilde{x}_t^{(i)}$ of a transmitted bit pattern $\underline{\tilde{x}}_t$ the two metrics $\tilde{\lambda}_0(\tilde{x}_t^{(i)})$ and $\tilde{\lambda}_1(\tilde{x}_t^{(i)})$ are generated [6]

$$\begin{aligned} \tilde{\lambda}_b\left(\tilde{x}_t^{(i)}\right) &= P\left(\tilde{x}_t^{(i)} = b | z_t\right) \\ &\sim \sum_{\hat{y}\in\mathcal{Y}_b^i} P(\hat{y}|z_t) \\ &\sim \sum_{\hat{y}\in\mathcal{Y}_b^i} P(z_t|\hat{y})P(\hat{y}) \quad , \end{aligned} \tag{4}$$

with $b = 0, 1$ for each $i = 1, \ldots I$. A metric $\tilde{\lambda}_b(\tilde{x}_t^{(i)})$ consists of the sum over all possible channel symbols $\hat{y}$ for which the ith 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 $\mathcal{Y}_b^i = \{\mu([\tilde{x}^{(1)}, \ldots \tilde{x}^{(I)}]) | \tilde{x}^{(i)} = b\}$. The conditional probability density $P(z_t|\hat{y})$ is given by

$$P(z_t|\hat{y}) = \frac{1}{\pi\sigma_n^2} \exp\left(-\frac{\|z_t - a_t\hat{y}\|^2}{\sigma_n^2}\right) \quad . \tag{5}$$

The probability $P(\hat{y})$ in (4) is the feedback computed later on in the decoding process. It contains the obtained a priori knowledge on the transmitted channel symbols. In the first iteration, when no feedback is available, the $P(\hat{y})$ are set as equiprobable.

Afterwards the metrics vectors $\underline{\tilde{\lambda}}_b$ are appropriately deinterleaved to $\underline{\lambda}_b$. Finally the $\underline{\lambda}_b$ are fed into a Soft-Input Soft-Output (SISO) channel decoder [15], which computes the probabilities $P(\hat{x}_t^{(i)} = b) \triangleq P(\hat{x}_t^{(i)} = b | \underline{\lambda}_0, \underline{\lambda}_1)$ of the encoded bits $x_t^{(i)}$ in addition to the estimated decoded data bits $\hat{u}$.

For the next iteration the $P(\hat{x}_t^{(i)})$ are interleaved again to $P(\hat{\tilde{x}}_t^{(i)})$. A large block size T and an appropriately designed interleaver ensure the independence of the $P(\hat{\tilde{x}}_t^{(i)})$ of one symbol $\underline{\hat{\tilde{x}}}_t$. Thus, the feedback input $P(\hat{y})$ of the metric computation can be determined as

$$P(\hat{y}) = P\left(\mu(\underline{\hat{\tilde{x}}})\right) = \prod_{j=1}^{I} P\left(\hat{\tilde{x}}^{(j)} = \mu^{-1}(\hat{y})^{(j)}\right) \quad , \tag{6}$$

which results in the new *extrinsic* metric [6]

$$\tilde{\lambda}_b^E\left(\tilde{x}_t^{(i)}\right) \sim \sum_{\hat{y}\in\mathcal{Y}_b^i} P(z_t|\hat{y}) \prod_{j=1, j\neq i}^{I} P\left(\hat{\tilde{x}}_t^{(j)} = \mu^{-1}(\hat{y})^{(j)}\right). \tag{7}$$

The metric is *extrinsic* since the feedback probability of the bit $\tilde{x}_t^{(i)}$ itself is excluded, $j \neq i$.

III. Performance Analysis

A performance bound of the bit-error rate (BER) P_b of BICM on Rayleigh fading channels was derived in [2]. The asymptotic behavior of P_b can be described by

$$\log_{10} P_b \approx \frac{-d_H(C)}{10}\left[\left(R\cdot\bar{d}_h^2(\mu)\right)_{\text{dB}} + \left(\frac{E_b}{N_0}\right)_{\text{dB}}\right] + const. \tag{8}$$

The minimum Hamming distance $d_H(C)$ of the code defines the slope of the bound. $\bar{d}_h^2(\mu)$ is the *harmonic mean* of the *minimum squared Euclidean distances* of the mapping μ. The product of the information rate R and the *harmonic mean* $\bar{d}_h^2(\mu)$ determines an E_b/N_0 offset of the bound, i.e., a horizontal offset in a BER vs. E_b/N_0 plot. For a mapping with I bits the *harmonic mean* $\bar{d}_h^2(\mu)$ is given by

$$\bar{d}_h^2(\mu) = \left(\frac{1}{I2^I}\sum_{i=1}^{I}\sum_{b=0}^{1}\sum_{y\in\mathcal{Y}_b^i}\frac{1}{\|y-\bar{y}\|^2}\right)^{-1} \quad , \tag{9}$$

where $\bar{y}\in\overline{\mathcal{Y}_b^i}$ denotes the nearest neighbor of $\underline{y}$ in the set $\overline{\mathcal{Y}_b^i}$, which is the complementary set to $\mathcal{Y}_b^i$, $\overline{\mathcal{Y}_b^i} = \mathcal{Y}\backslash\mathcal{Y}_b^i$.

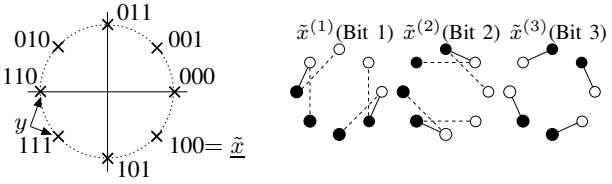

Fig. 2. 8PSK-Gray mapping [6] with signal-pair distances $\|y-\bar{y}\|^2$ for $\bar{d}_h^2(\mu)$. $\circ \leftrightarrow \tilde{x}^{(i)} = 0$ and $\bullet \leftrightarrow \tilde{x}^{(i)} = 1$

On the right side of Fig. 2 the *minimum squared Euclidean distances* $\|y-\bar{y}\|^2$ of Gray mapped 8PSK used in (9) are depicted. Black dots denote $\tilde{x}^{(i)} = \mu^{-1}(y)^{(i)} = 1$ and white dots $\tilde{x}^{(i)} = \mu^{-1}(y)^{(i)} = 0$. If we assume, e.g., $i=1$ and $b=0$ in the two outer sums of (9) we get for the inner sum $y\in\mathcal{Y}_0^1$. The subset $\mathcal{Y}_0^1$ contains the channel symbols with the bit patterns $\underline{\tilde{x}}\in\{[000], [001], [010], [011]\}$. Thus, the inner summation comprises all channel symbols with

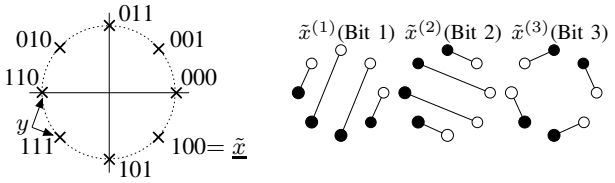

Fig. 3. 8PSK-Gray mapping [6] with signal-pair distances $\|y-\check{y}\|^2$ for $\check{d}_h^2(\mu)$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

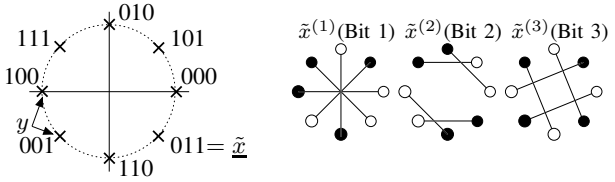

Fig. 4. 8PSK-SSP mapping [6] with signal-pair distances $\|y-\check{y}\|^2$ for $\check{d}_h^2(\mu)$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

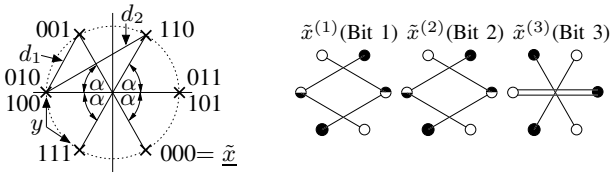

Fig. 5. α-6PSK$_{I=3}$ mapping with signal-pair distances $\|y-\check{y}\|^2$ for $\check{d}_h^2(\mu)$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

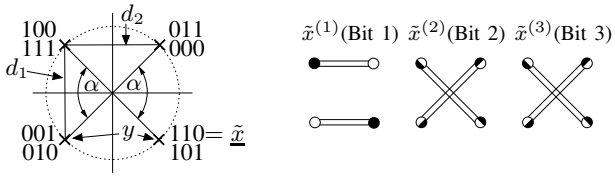

Fig. 6. α-4PSK$_{I=3}$ mapping with signal-pair distances $\|y-\check{y}\|^2$ for $\check{d}_h^2(\mu)$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

$\tilde{x}^{(1)}=0$, i.e., all white dots in the left of the three signal spaces. For each of these channel symbols y (or white dots) we search for the nearest channel symbol $\bar{y}\in\overline{\mathcal{Y}_b^i}$, i.e., the black dot with the *minimum squared Euclidean distance*. For, e.g., $y\in\{\mu([000]),\mu([001])\}$ this is $\bar{y}=\mu([100])$ and for $y\in\{\mu([010]),\mu([011])\}$ we get $\bar{y}=\mu([110])$. Thus, we can calculate the *minimum squared Euclidean distances* $\|y-\bar{y}\|^2$. The solid lines in the three signal spaces on the right side of Fig. 2 shall indicate that the respective *Euclidean distances* are used twice in the computation of the *harmonic mean* $\bar{d}_h^2(\mu)$ in (9), while *Euclidean distances* with dashed lines are used only once.

For the bound of the iterative case with BICM-ID we assume *error-free feedback* (EFF) [6], i.e., we assume that the feedback $P(\hat{y})$ for all bits except the one currently considered is correct with perfect reliability (best case). In a real system this assumption is obviously not fulfilled for a low E_b/N_0, but, as simulations will show, the bound can be reached at a high E_b/N_0. When using the *extrinsic* information in the EFF we do not use the nearest neighbor $\bar{y}$ anymore, but we use the possibly different channel symbol $\check{y}$, for which the rest of the corresponding bit pattern $\underline{\tilde{x}}=\mu^{-1}(\check{y})$ matches the EFF, i.e., $\check{y}$ possesses the identical bit pattern as y except for an inverted bit at position i. For example, if $i=1$ and $y=\mu([001])$ the assumed feedback is $\tilde{x}^{(2)}=0$ and $\tilde{x}^{(3)}=1$, resulting in $\check{y}=\mu([101])$. In some cases $\check{y}$ and $\bar{y}$ can be identical. When replacing $\bar{y}$ by $\check{y}$ in (9) we get a new *harmonic mean* [6]

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

which has to be used in (8) instead of $\bar{d}_h^2(\mu)$ to obtain the bound for the iterative case [6]. The right side of Fig. 3 depicts the *squared Euclidean distances* $\|y-\check{y}\|^2$ of Gray mapped 8PSK. Note that some *Euclidean distances* have been slightly increased with respect to Fig. 2, e.g., for $i=1$ and $y=\mu([001])$ the corresponding other channel symbol changed from $\bar{y}=\mu([100])$ to $\check{y}=\mu([101])$.

Since for a constant information rate R the *harmonic means* $\bar{d}_h^2(\mu)$ and $\check{d}_h^2(\mu)$ define the horizontal offset in the computation of the bound in (8) a high $\bar{d}_h^2(\mu)$ or $\check{d}_h^2(\mu)$ is an indicator for a good asymptotic performance with non-iterative resp. iterative BICM. In the non-iterative BICM case Gray mappings have the largest $\bar{d}_h^2(\mu)$. But for BICM-ID other mappings with a significantly larger $\check{d}_h^2(\mu)$ than the one of the respective Gray mapping exist. They can outperform Gray mapping in the iterations at high E_b/N_0. The ratio $(\check{d}_h^2(\mu)/\bar{d}_h^2(\mu)^{[\max]})_{\text{dB}}$ is called the *offset gain* [6] and describes the possible gain of BICM-ID with a certain mapping with respect to optimum Gray-mapped BICM. Therefore, in this paper $\check{d}_h^2(\mu)$ and the corresponding *offset gain* will serve as the key indicators for the evaluation of the theoretically achievable performance of the new and some known mappings with BICM-ID.

Mappings with *regular* 8PSK *signal constellation sets* were investigated in [6]. Semi Set-Partitioning (SSP) mapping depicted in Fig. 4 has the largest $\check{d}_h^2(\text{8PSK-SSP})=2.877$, which yields an offset gain of 5.74 dB in relation to Gray mapped 8PSK BICM ($\bar{d}_h^2(\mu)_{I=3}^{[\max]}=0.766$) in Fig. 2.

IV. Non-Regular Signal Constellation Sets

To further increase the *offset gain*, respectively improve the asymptotic performance, we propose the usage of new mappings μ with *signal constellation sets* with less than the usual 2^I different channel symbols y for bit patterns $\underline{\tilde{x}}$ with I bits and an unequal distribution

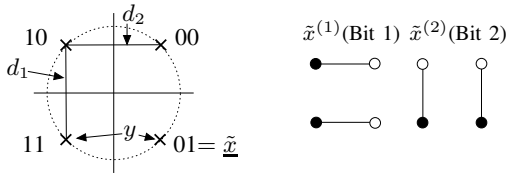

Fig. 7. 4PSK-Gray mapping with signal-pair distances $\|y-\check{y}\|^2$ for $\check{d}_h^2(\mu)$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

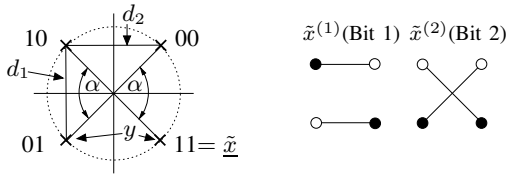

Fig. 8. α-4PSK$_{I=2}$ mapping with signal-pair distances $\|y-\check{y}\|^2$ for $\check{d}_h^2(\mu)$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

of the channel symbols in the signal space. Reducing the number of channel symbols y is a contrary strategy as in TCM, which expands the *signal constellation sets* by increasing the number of channel symbols [3]. In the following we discuss different examples for these *non-regular signal constellation sets*. The obtained numerical values for $\check{d}_h^2(\mu)$ and the corresponding *offset gains* are summarized in Tables I and II. To simplify notations the time index t is omitted throughout this section.

A. α-6PSK mapping for $I=3$

First, we consider the mapping depicted in Fig. 5, which maps $2^{I=3}=8$ different bit patterns $\underline{\tilde{x}}$ to 6 channel symbols y and which will be denoted as α-6PSK$_{I=3}$ mapping. The angle α defines the angle between certain channel symbols on the unit circle with respect to the origin. Thus, by changing α different mappings can be obtained. With this α-6PSK$_{I=3}$ mapping twice two bit patterns $\underline{\tilde{x}}$ are mapped to the same channel symbol y, i.e., $\underline{\tilde{x}}_{1,2}=[010],[100]\Rightarrow y=-1$ and $\underline{\tilde{x}}_{1,2}=[011],[101]\Rightarrow y=+1$. In the first iteration this results in an automatic partial puncturing of the terms $P(z|\hat{y})P(\hat{y})$ in the metric computation in (4). At the bit positions i where the bit patterns differ, $\tilde{x}_1^{(i)}\neq\tilde{x}_2^{(i)}$, the terms $P(z|\hat{y})P(\hat{y})$ with $\hat{y}_{1,2}=\mu(\underline{\tilde{x}}_{1,2})$ are distributed equally weighted to $\tilde{\lambda}_0(\tilde{x}^{(i)})$ and $\tilde{\lambda}_1(\tilde{x}^{(i)})$, since $P(\hat{y}_1)=P(\hat{y}_2)$. This partial puncturing of the metrics implies a severe degradation of the performance in the initial iteration.

However, if the E_b/N_0 is sufficiently high the decoder can cope the introduced puncturing and generates improving feedback $P(\hat{y})$. This feedback will be different for $\underline{\tilde{x}}_1$ and $\underline{\tilde{x}}_2$, $P(\mu(\underline{\tilde{x}}_1))\neq P(\mu(\underline{\tilde{x}}_2))$, and thus it improves the metric computation by compensating (or reducing) the partial puncturing.

At high E_b/N_0 and a sufficient number of iterations the feedback may finally converge to EFF, the necessary

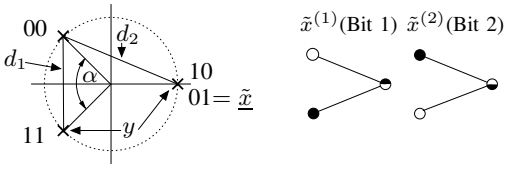

Fig. 9. α-3PSK$_{I=2}$ mapping with signal-pair distances $\|y-\check{y}\|^2$ for $\check{d}_h^2(\mu)$. $\circ \leftrightarrow \tilde{x}^{(i)}=0$ and $\bullet \leftrightarrow \tilde{x}^{(i)}=1$.

assumption for the computation of $\check{d}_h^2(\mu)$ in (10). If the two bit patterns assigned to a channel symbol differ at a certain bit i, $\tilde{x}_1^{(i)}\neq\tilde{x}_2^{(i)}$, this is designated by the partial black, partial white dots as in the left and middle signal spaces in Fig. 5. Also some *Euclidean distances* are used four times in (10), indicated by two parallel solid lines as in the right signal space in Fig. 5.

For the angles $\alpha=60°$ and $\alpha=45°$ we get with (10) $\check{d}_h^2(60°\text{-6PSK}_{I=3})=3.273$ and $\check{d}_h^2(45°\text{-6PSK}_{I=3})=3.589$, translating to *offset gains* of 6.30 dB resp. 6.71 dB, which are higher than the one of 8PSK-SSP mapping. Thus, the two α-6PSK$_{I=3}$ mappings will outperform 8PSK-SSP mapping by 0.56 dB resp. 0.97 dB when the E_b/N_0 is such high that in both cases the bound is reached. Note, in case a relatively high target BER is given, the mapping with the largest *offset gain* may not be the best choice, since it might only be superior at a BER below the target BER. Further note that the (in this case two) $\underline{\tilde{x}}$ mapped to one y must differ in at least two bits, since otherwise for a single different bit i, we would have $\check{y}=y$ at this bit, resulting in $\check{d}_h^2(\mu)=0$. The *extrinsic* feedback information for the single different bit being 0 or 1 would be always identical.

B. α-4PSK mapping for $I=3$

Next, we will still consider a mapping for $2^{I=3}=8$ different bit patterns, but further reduce the number of channel symbols y to 4. A possible layout for the signal constellation set and the bit assignment is the α-4PSK$_{I=3}$ mapping depicted in Fig. 6. In case of $\alpha=90°$ the *signal constellation set* is identical to the one of regular QPSK. With this mapping two bit patterns $\underline{\tilde{x}}$ are assigned to every channel symbol y. Thus, all channel symbols are subject to the automatic partial puncturing in the first iteration introduced by the *non-regular signal constellation sets*. Since for all channel symbols y the only common bit of the two corresponding bit patterns $\underline{\tilde{x}}$ is the first bit $\tilde{x}^{(1)}$, the two remaining bits $\tilde{x}^{(2)}$ and $\tilde{x}^{(3)}$ are always initially punctured with this mapping. Thus, in the first decoding cycle two out three bits are punctured, which yields a very poor performance in this first iteration.

Nevertheless, at high E_b/N_0 strong channel codes can compensate this puncturing in the succeeding iterations as described in Section IV-A. For a 90°-4PSK$_{I=3}$ mapping, i.e., using the channel symbols of regular QPSK,

TABLE I

$\check{d}_h^2$ AND OFFSET GAINS FOR MAPPINGS WITH $I=3$ BITS.

mapping μ	$\check{d}_h^2(\mu)$	offset gain
8PSK-Gray [6] (Fig. 3)	0.810	0.24 dB
8PSK-SSP [6] (Fig. 4)	2.877	5.74 dB
90°-4PSK$_{I=3}$ (Fig. 6, $\alpha=90°$)	3	5.93 dB
60°-6PSK$_{I=3}$ (Fig. 5, $\alpha=60°$)	3.273	6.30 dB
45°-6PSK$_{I=3}$ (Fig. 5, $\alpha=45°$)	3.589	6.71 dB
45°-4PSK$_{I=3}$ (Fig. 6, $\alpha=45°$)	3.784	6.93 dB
BPSK bound (Figs. 5 and 6, $\alpha\to0°$)	4	7.18 dB

TABLE II

$\check{d}_h^2$ AND OFFSET GAINS FOR MAPPINGS WITH $I=2$ BITS.

mapping μ	$\check{d}_h^2(\mu)$	offset gain
4PSK-Gray (Fig. 7)	2	0 dB
90°-4PSK$_{I=2}$ (Fig. 8, $\alpha=90°$)	2.667	1.25 dB
120°-3PSK$_{I=2}$ (Fig. 9, $\alpha=120°$)	3	1.76 dB
45°-4PSK$_{I=2}$ (Fig. 8, $\alpha=45°$)	3.684	2.65 dB
65°-3PSK$_{I=2}$ (Fig. 9, $\alpha=65°$)	3.687	2.66 dB
45°-3PSK$_{I=2}$ (Fig. 9, $\alpha=45°$)	3.922	2.93 dB
BPSK bound (Figs. 8 and 9, $\alpha\to0°$)	4	3.01 dB

we obtain with (10) a $\check{d}_h^2(90°\text{-4PSK}_{I=3})=3$ and an *offset gain* of 5.93 dB. Thus, the 90°-4PSK$_{I=3}$ mapping will not surpass the two exemplary α-6PSK$_{I=3}$ mappings presented above. The relatively small *Euclidean distances* d_2 of the first bit dominate the computation of $\check{d}_h^2$. But if α is similarly reduced to $\alpha = 45°$, the *offset gain* increases to 6.93 dB, with $\check{d}_h^2(45°\text{-4PSK}_{I=3})=3.784$, due to the larger d_2. Therefore, the 45°-4PSK$_{I=3}$ mapping outperforms the 45°-6PSK$_{I=3}$ mapping by approximately 0.2 dB, at least asymptotically. In the initial iterations the situation is of course vice-versa, since in case of 45°-4PSK$_{I=3}$ mapping more bits are automatically punctured.

C. α-4PSK mapping for $I=2$

For the case of transmitting $I=2$ bits per channel symbol the baseline system uses Gray mapping with a *regular* 4PSK (or QPSK) *signal constellation set* as presented in Fig. 7. With this mapping the maximum $\bar{d}_h^2(\mu)$ for $I=2$ is obtained, $\bar{d}_h^2(\mu)_{I=2}^{[\max]}=2$. Thus, this mapping gives the best performance for the case of non-iterative BICM. But, since the Euclidean distances of the signal-pairs are also the minimum distances as visible in the right part of Fig. 7, no improvement is possible by adding iterations. Analytically this fact is substantiated by $\check{d}_h^2(\text{4PSK-Gray})=2$, resulting in an *offset gain* of 0 dB. Also by introducing an angle α as above this mapping cannot be improved. A larger d_2, e.g., would imply a smaller d_1, leading to a decreasing harmonic mean $\check{d}_h^2(\mu)$.

An iterative gain can be obtained, e.g., by changing the labeling of the channel symbols to Anti-Gray as in the α-4PSK$_{I=2}$ mapping in Fig. 8. Now, the Euclidean distances of the signal-pairs for bit $\tilde{x}^{(2)}$ are larger than the minimum distance. Thus, this difference can be explored in the iterations. If the position of the channel symbols remains unchanged with respect to 4PSK-Gray mapping, i.e., $\alpha = 90°$, we get $\check{d}_h^2(90°\text{-4PSK}_{I=2})=2.667$ and a corresponding *offset gain* of 1.25 dB.

This α-4PSK$_{I=2}$ mapping can also be additionally improved in terms of *offset gain* by reducing α. In contrast to 4PSK-Gray mapping an increasing d_2 does not affect the other distances in the *harmonic mean* $\check{d}_h^2(\mu)$, since in the *harmonic mean* computation of (10) the distance d_1 is not used anymore. As an example we consider $\alpha = 45°$. With $\check{d}_h^2(45°\text{-4PSK}_{I=2})=3.684$ the resulting *offset gain* of 2.65 dB of 45°-4PSK$_{I=2}$ mapping is more than twice the one of 90°-4PSK$_{I=2}$ mapping.

Since with any α-4PSK$_{I=2}$ mapping still only one bit pattern $\underline{\tilde{x}}$ is assigned to each channel symbol y there is no automatic puncturing in the first iteration for this mapping. Thus, only a medium performance degradation in the first iteration can be expected, implying that relatively few iterations are necessary to explore the achievable *offset gain*.

D. α-3PSK mapping for $I=2$

For a reduced number of three allowed channel symbols for the case of $I = 2$ we propose the α-3PSK$_{I=2}$ mapping depicted in Fig. 9. Note, with this α-3PSK$_{I=2}$ mapping the channel symbol $y=+1$ is completely punctured in the first iteration since the assigned bit patterns are entirely different. Equally distributed channel symbols, i.e., $\alpha=120°$, result in $\check{d}_h^2(120°\text{-3PSK}_{I=2})=3$ and an *offset gain* of 1.76 dB, surpassing the 90°-4PSK$_{I=2}$ mapping with also equally distributed, but not multiple used channel symbols by about 0.5 dB.

If an identical angle of $\alpha=45°$ is used for α-3PSK$_{I=2}$ mapping and α-4PSK$_{I=2}$ mapping this gap reduces to approximately 0.3 dB since for 45°-3PSK$_{I=2}$ mapping the *offset gain* computes as 2.93 dB with $\check{d}_h^2(45°\text{-3PSK}_{I=2})=3.922$.

For comparison of a α-3PSK$_{I=2}$ mapping and a α-4PSK$_{I=2}$ mapping with identical *offset gains* a 65°-3PSK$_{I=2}$ mapping can be used, which has with $\check{d}_h^2(65°\text{-3PSK}_{I=2})=3.687$ and an corresponding *offset gain* of 2.66 dB similar values as the 45°-4PSK$_{I=2}$ mapping.

E. BPSK mapping for $I\geq 2$

Reducing the number of possible channel symbols y to 2, i.e., using a *regular* BPSK *signal constellation set*, is not feasible since no bit pattern assignment for $I\geq2$ exists which does not violate one of the

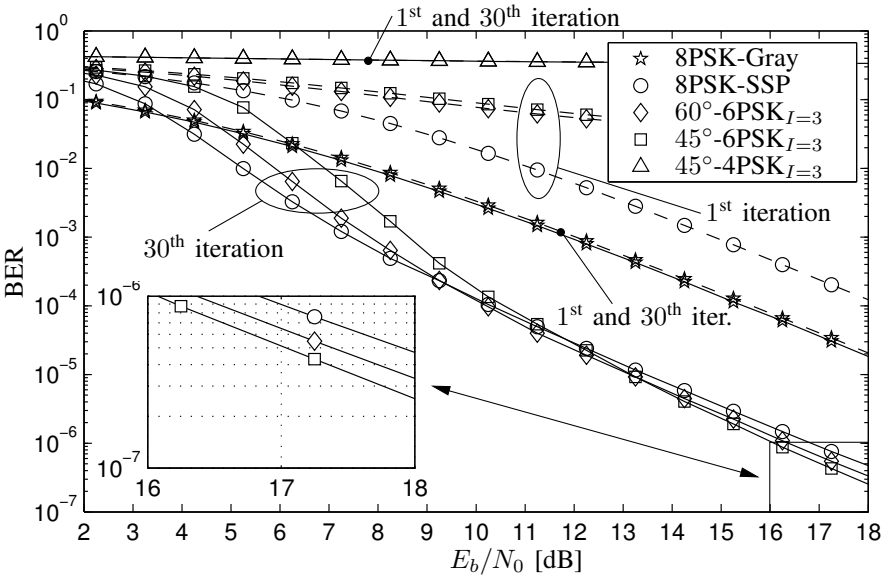

Fig. 10. BER for $I=3$ with a 2-state, rate-1/2 code and 12000 information bits per frame.

two necessary conditions for the bit patterns $\underline{\tilde{x}}$ of the channel symbols y. On the one hand the bit patterns assigned to one channel symbols must differ in at least 2 positions, because otherwise $\check{d}_h^2(\mu)=0$, as described in Section IV-A. On the other hand at least for some channel symbols the assigned bit patterns must have common bits, since otherwise all bits are automatically punctured in the first iteration and no usable feedback can be generated.

V. Theoretical Limit

In this section we derive the theoretical maximum *offset gain*. First a system with $I=3$ bits per channel symbol is considered. For the extreme case of $\alpha\to 0$ in Figs. 5 and 6, this results on the one hand in $d_1\to 0$, meaning that $E_b/N_0\to\infty$ is required to have improving feedback. On the other hand, looking at the asymptotic performance with EFF, we have $d_2\to 2$ which yields in $\check{d}_h^2(\mu)\to 4$. The corresponding *offset gain* is 7.18 dB. With EFF all signal-pairs resemble now a BPSK *signal constellation set* with $y=\pm 1$. This is obviously the best possible case for a symbol energy of $E_s=1$ and therefore a bound for the obtainable *offset gain* with BICM-ID and $I=3$. Due to the similarity of the *signal constellation sets* we will denote this bound in the following as BPSK bound.

Obviously, the value $\check{d}_h^2(\mu)=4$ marks an upper limit also for systems with $I\neq 3$. The corresponding maximum achievable *offset gain*, however, additionally depends on $\bar{d}_h^2(\mu)^{[\max]}$, which differs for different values of I and is usually achieved with Gray or near-Gray mappings. For $I=2$ a theoretically maximum possible *offset gain* of 3.01 dB can be obtained with $\bar{d}_h^2(\mu)_{I=2}^{[\max]}=\bar{d}_h^2(\text{4PSK-Gray})=2$.

VI. Numerical and Simulation Results

Table I lists the $\check{d}_h^2(\mu)$ and the respective *offset gains* for the mappings with $I=3$ considered in this paper. The calculation of the *offset gains* is based on $\bar{d}_h^2(\mu)_{I=3}^{[\max]}=0.766$ for Gray-mapped 8PSK. Note that an increase in $\check{d}_h^2(\mu)$ and especially the usage of *non-regular signal constellation sets* usually inflicts a performance degradation in the first iteration and a higher E_b/N_0 at which the asymptotic bound is reached.

Fig. 10 depicts the results of a simulation comparing 8PSK-SSP mapping with the α-6PSK$_{I=3}$ mappings. Additionally as a reference the performance of BICM with 8PSK-Gray mapping, i.e., the optimum mapping for the non-iterative case, is included. The block size is set to 12000 information bits per frame and 30 iterations are performed. As channel code serves a 2-state, rate-1/2 convolutional code with generator polynomials $\{1,3\}$ and $d_H(C)=3$. This weak code was used since with a low $d_H(C)$ the slope of the bound given in (8) is relatively flat. Thus, the bound is reached at a higher BER, which simplifies simulations. In the case of Fig. 10 all mappings seem to have reached the bound at a BER of $\sim 10^{-6}$. As visible in the enlargement the horizontal offsets between the different mappings approximately match the differences in the calculated *offset gains* given in Table I, where $60°$-6PSK$_{I=3}$ and $45°$-6PSK$_{I=3}$ surpass 8PSK-SSP by 0.56 dB resp. 0.97 dB. Also the expected significant degradation in performance of the α-6PSK$_{I=3}$ mappings in the first iteration can be observed. The bad performance of $45°$-4PSK$_{I=3}$ mapping, the mapping with largest *offset gain* in Table I, is due to the weak, rate-1/2 code used for the simulations. The code is not strong enough to generate usable feedback. Too many bits, two out of three, are automatically punctured in the first iteration, resulting in an effective code rate larger than 1.

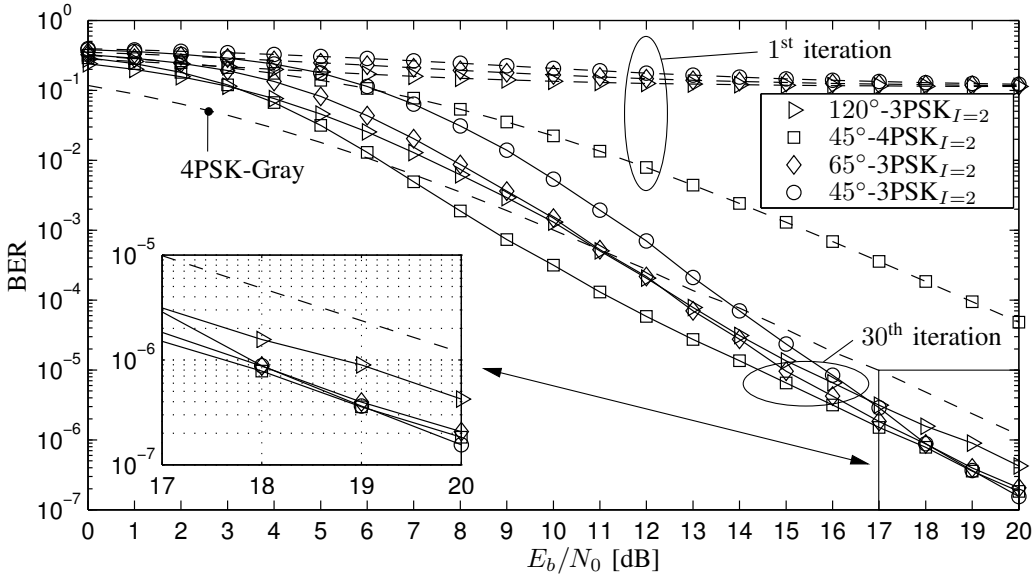

Fig. 11. BER for $I=2$ with a 2-state, rate-1/2 code and 12000 information bits per frame.

Table II contains a summary of the *harmonic means* $\check{d}_h^2(\mu)$ and the respective *offset gains* for the mappings with $I = 2$ discussed in this paper. The *offset gains* are computed with $\bar{d}_h^2(\mu)_{I=2}^{[\max]} = 2$. Fig. 11 depicts simulation results for several of these mappings. Again, the reference for the non-iterative case, i.e., BICM with 4PSK-Gray mapping, is included. The simulation parameters are identical to the ones given above for Fig. 10, except that the number of iterations is limited to 15. As visible, considerable improvements can be achieved by *non-regular signal constellation sets* also for $I=2$. $65°$-3PSK$_{I=2}$ mapping does not reach exactly the performance of $45°$-4PSK$_{I=2}$ mapping, which has an approximately equal *offset gain*. More iterations are required for $65°$-3PSK$_{I=2}$ mapping to reach its bound.

When comparing Figs. 10 and 11 it can be observed that for similar simulation parameters, since the performance bounds are reached the number of iterations has no influence anymore, the mappings with $I = 3$ outperform the mappings with $I=2$ in terms of BER versus E_b/N_0. This is a result of the significantly larger *offset gains* that can be achieved with $I=3$ and implies that, in contrast to BICM, for BICM-ID transmitting fewer symbols with mappings with large I is superior to the transmission of a lot of symbols with a small I.

VII. EXIT Chart Analysis

EXIT (*extrinsic* information transfer) charts [10], [11], [12] are a powerful tool to analyze the convergence behavior of iterative systems which utilize the Turbo principle, i.e., exchange and refine *extrinsic* information [7], [8]. The capabilities of the components, in our case the demodulator (DM) and the SISO channel decoder (CD), are analyzed separately by determining the *mutual information* $\mathcal{I}^{[\text{ext}]}$ between the *extrinsic* information and the data, which is obtained by a component for a certain *mutual information* $\mathcal{I}^{[\text{apri}]}$ of input *a priori* information. Since the *extrinsic* information of one component serves as input *a priori* information for the other component, the two resulting EXIT characteristics are plotted in a single graph with swapped axes. The EXIT characteristic of the demodulator, i.e., of the considered mapping, depends on the E_b/N_0 of the channel, while the EXIT characteristic of the channel decoder is independent of the E_b/N_0 since it has no access to the received channel symbols z.

In Fig. 12 EXIT characteristics for the mappings used in Fig. 10 are presented for $E_b/N_0 = 5.25$ dB. As visible the EXIT characteristic of 8PSK-Gray mapping is almost horizontal, which prohibits gains by iterations. The EXIT characteristics of the other mappings exhibit noticeable slopes, allowing improvement by numerous iterations except for the $45°$-4PSK$_{I=3}$ mapping. For the latter case the EXIT characteristics of demodulator and channel decoder intersect at very low values of *mutual information* $\mathcal{I}_{\text{DM}}^{[\text{ext}]}$ and $\mathcal{I}_{\text{CD}}^{[\text{ext}]}$. For the 2-state, rate 1/2 case the chosen generator polynomials $\{1,3\}$ are the only reasonable option. But for channel codes with more states the generators polynomials can be optimized by the EXIT chart to exhibit a similar (with swapped axes) EXIT characteristic as the chosen mapping.

The convergence behavior of the complete system is demonstrated in EXIT charts by plotting the EXIT trajectory (step-curve) of the exchanged *extrinsic* information. In Fig. 12 this EXIT trajectory is depicted for the $60°$-6PSK$_{I=3}$ mapping at $E_b/N_0 = 5.25$ dB, proving that despite the introduced *non-regular signal constellation sets* with fewer channel symbols y than possible bit patterns $\underline{\tilde{x}}$ the proposed mappings feature the usual iterative convergence behavior.

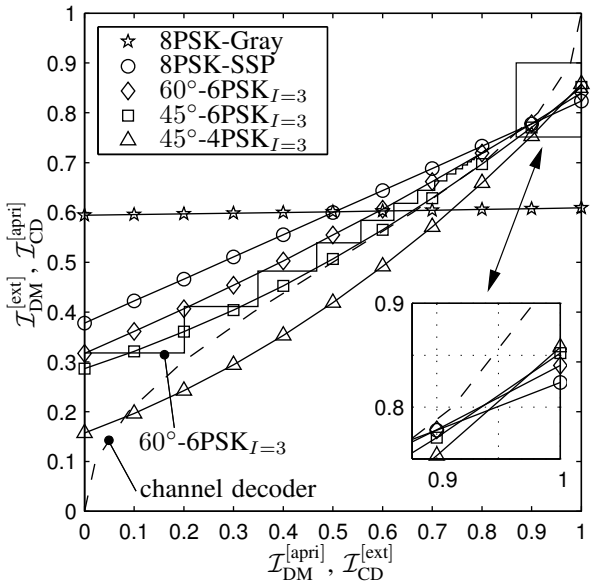

Fig. 12. EXIT charts for Fig. 10 at $E_b/N_0 = 5.25$ dB ($I = 3$; 2-state, rate-1/2 code; 12000 information bits per frame).

The performance analysis in Section III and the development of the *non-regular signal constellation sets* in Section IV are not based on the convergence behavior but the asymptotic behavior. To compute the achievable *offset gains* the assumption of *error-free feedback* (EFF) is made. In the EXIT charts this EFF corresponds to $\mathcal{I}_{\mathrm{DM}}^{[\mathrm{apri}]} = 1$ for the EXIT characteristics of the different mappings. Thus, information on the asymptotic behavior should be available on the right vertical axis, i.e., $\mathcal{I}_{\mathrm{DM}}^{[\mathrm{apri}]} = 1$. However, to generate EFF a channel code usually requires also $\mathcal{I}_{\mathrm{CD}}^{[\mathrm{apri}]} = 1$. Consequently, the asymptotic behavior cannot be satisfactorily analyzed since the required EFF exists only in the point $\mathcal{I}_{\mathrm{DM}}^{[\mathrm{apri}]} = \mathcal{I}_{\mathrm{CD}}^{[\mathrm{apri}]} = 1$, the upper right corner of the EXIT chart. Nevertheless, an initial statement on the asymptotic behavior can be made by considering the EXIT characteristics of the different mappings at $\mathcal{I}_{\mathrm{DM}}^{[\mathrm{apri}]} = 1$. As depicted in the enlargement in Fig. 12, already for $E_b/N_0 = 5.25$ dB the values for $\mathcal{I}_{\mathrm{DM}}^{[\mathrm{ext}]}$ at $\mathcal{I}_{\mathrm{DM}}^{[\mathrm{apri}]} = 1$ are in the same order as the achievable *offset gains* listed in Table I.

VIII. Conclusion

In this paper we extended the capabilities of iteratively decoded BICM by using fundamentally new mappings with *non-regular signal constellation sets*. These mappings have less channel symbols than possible bit patterns. Thus, more than one bit pattern may be assigned to one channel symbol. Also the position of the channel symbols are not necessarily equally distributed anymore. By this technique larger *offset gains* are achievable in comparison to previously known mappings. Numerous mappings for systems with two or three bits per channel symbol, i.e., for improvement of QPSK and 8PSK baseline systems, are presented and the theoretically obtainable performance gains demonstrated by simulations. The convergence behavior is verified by EXIT charts. Furthermore a bound for the achievable improvement by this technique is derived.

IX. Acknowledgment

The authors would like to acknowledge Susanne Godtmann for providing the simulation results of Fig. 12.

References

[1] E. Zehavi, "8-PSK Trellis Codes for a Rayleigh Channel," *IEEE Trans. Comm.*, vol. 40, no. 5, pp. 873–884, May 1992.

[2] G. Caire, G. Taricco, and E. Biglieri, "Bit-Interleaved Coded Modulation," *IEEE J. Select. Areas Commun.*, vol. 44, no. 3, pp. 927–946, May 1998.

[3] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. 28, no. 1, pp. 56–67, Jan. 1982.

[4] X. Li and J. A. Ritcey, "Bit-Interleaved Coded Modulation with Iterative Decoding," *IEEE Comm. Lett.*, vol. 1, no. 6, pp. 169–171, Nov. 1997.

[5] ——, "Trellis-Coded Modulation with Bit Interleaving and Iterative Decoding," *IEEE J. Select. Areas Commun.*, vol. 17, no. 4, 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.*, vol. 50, no. 8, pp. 1250–1257, Aug. 2002.

[7] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: Turbo-Codes," *IEEE Trans. Comm.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.

[8] J. Hagenauer, E. Offer, and L. Papke, "Iterative Decoding of Binary Convolutional Codes," *IEEE Trans. Comm.*, vol. 42, no. 2, pp. 429–445, Mar. 1996.

[9] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping for QPSK modulation," *IEEE Electronics Lett.*, vol. 34, no. 15, pp. 1459–1460, 1998.

[10] S. ten Brink, "Convergence of Iterative Decoding," *IEEE Electronics Lett.*, vol. 35, no. 10, pp. 806–808, May 1999.

[11] ——, "Designing Iterative Decoding Schemes with the Extrinsic Information Transfer Chart," *International Journal of Electronics and Communications (AEÜ)*, vol. 54, no. 6, pp. 389–398, Dec. 2000.

[12] ——, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes," *IEEE Trans. Comm.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[13] Y. Huang and J. A. Ritcey, "16-QAM BICM-ID in Fading Channels With Imperfect Channel State Information," *IEEE Trans. Wireless Comm.*, vol. 2, no. 5, pp. 1000–1007, Sep. 2003.

[14] ——, "EXIT Chart Analysis of BICM-ID With Imperfect Channel State Information," *IEEE Comm. Lett.*, vol. 7, no. 9, pp. 434–436, Sep. 2003.

[15] 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*, vol. 9, no. 2, pp. 155–172, Mar. 1998.