EXIT CHARTS FOR ITERATIVELY DECODED MULTILEVEL MODULATION

Frank Schreckenbach

Institute for Communications Engineering (LNT)
Munich University of Technology (TUM)
80290 Munich, Germany
email: frank.schreckenbach@ei.tum.de

Gerhard Bauch

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

ABSTRACT

We investigate EXIT charts for the design and analysis of Bit-Interleaved Coded Modulation (BICM) with iterative decoding. Analytical expressions for EXIT functions of convolutional codes and mappings are derived. For the demapper, we show EXIT functions for different channel models and introduce bit-wise EXIT functions and a symbol-wise description. These functions are used in an optimization procedure to find new mappings.

1. INTRODUCTION

Several coded modulation schemes, where channel coding and high order modulation are combined for bandwidth efficient transmission have been proposed. We will focus on BICM [1][2], that is the concatenation of an encoder, an interleaver, and a mapper. The advantages of BICM are the simple and flexible implementation possibilities and the good performance in fading channels, resulting from the maximized diversity order obtained by bit-wise interleaving.

The performance of BICM can be greatly improved through iterative decoding. One strategy is to use a strong channel code which allows iterative decoding, e.g. a parallel concatenated turbo code [3] or a Low Density Parity Check (LDPC) code [4]. Since the mapping introduces dependencies between the bits, it can also be used as rate 1 code. Thus, an other strategy is to include the mapping in the iterative decoding process and to perform iterative decoding between the inner decoder, i.e. the demapper, and the outer channel decoder [5, 6]. This system is usually referred to as BICM with iterative decoding (BICM-ID) and is well suited for a combination with e.g. iterative equalization or MIMO detection [7]. Other strategies, not addressed in this paper, include the use of an additional precoder [8] or the combination of concatenated channel codes and iterative demapping, resulting in a threefold concatenated system.

It was soon recognized that the choice of the mapping of binary indeces to signal points is a crucial design parameter for BICM and BICM-ID. It is well known that Gray mapping is optimal for BICM [2]. A method to optimize mappings for BICM-ID with respect to the coding gain over the iterations was presented in [9].

A powerful tool to design and optimize iterative systems is the *Extrinsic Information Transfer (EXIT) chart* [10]. The EXIT chart visualizes the exchange of mutual information in an iterative system. The system components are characterized by their *EXIT functions*, which describe the output mutual information as a function of the input mutual information. For an introduction to EXIT charts, we refer to [11].

This paper aims on the following: First, we devise properties and new analytical results for EXIT functions of convolutional codes for a Binary Erasure Channel (BEC). Then, EXIT functions of mappings with different channel models are compared and analytical expressions are derived. Bit-wise EXIT functions and a symbol-wise description of mappings are introduced. Finally, we use these results and the optimization method based on a binary switching algorithm proposed in [9] to find new mappings with desired properties.

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

2. SYSTEM MODEL

We consider the BICM system depicted in Fig. 1.

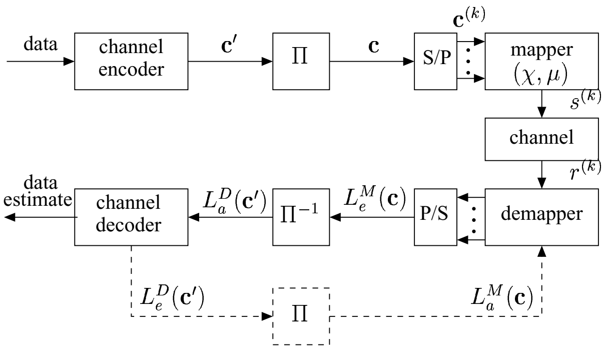

Figure 1: System model.

A block of data bits is encoded by a channel encoder and bit-interleaved by the random interleaver Π. m consecutive bits of the coded and interleaved sequence are grouped to form the subsequences $\mathbf{c}^{(k)} = (c_0^{(k)}, \ldots, c_i^{(k)}, \ldots, c_{m-1}^{(k)})$. 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 χ according to the labeling map μ. In the following, we will disregard the index (k).

The channel is described by $r = a \cdot s + n$, where a denotes the fading coefficient and n 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 and the corresponding a priori log-likelihood ratios (LLRs or L-values) $L_a^M(C_i) = \log(P(C_i = 0)/P(C_i = 1))$[1] of the coded bits and outputs the extrinsic LLRs:

$$L_e^M(C_i) = \log \frac{\sum_{s \in \chi_0^i} p(r|s) \cdot \prod_{j=1, j \neq i}^{m} \mathrm{e}^{-L_a^M(C_j) \cdot c_j}}{\sum_{s \in \chi_1^i} p(r|s) \cdot \prod_{j=1, j \neq i}^{m} \mathrm{e}^{-L_a^M(C_j) \cdot c_j}}, \qquad (1)$$

where $\chi_{c_i}^i$ denotes the subset of symbols $s \in \chi$ whose bit labels have the value $C_i = c_i \in \{0, 1\}$ in position $i \in \{1, \ldots, m\}$.

The extrinsic estimates $L_e^M(C_i)$ are deinterleaved and applied to the APP channel decoder. If no feedback from the channel decoder to the demapper is implemented or during the initial demapping step, the a priori LLRs $L_a^M(C_i)$ are equal to zero. Otherwise the estimates of the coded bits from the decoder are fed back and regarded as a priori information at the demapper.

[1] C_i denotes the binary random variable with realizations $c_i \in \{0, 1\}$

3. EXIT FUNCTION OF CHANNEL CODE

Fig. 2 depicts EXIT functions $T_D(i)^{-1}$ (inverted, with flipped axes) of several channel codes used as outer codes in a serial concatenated system.

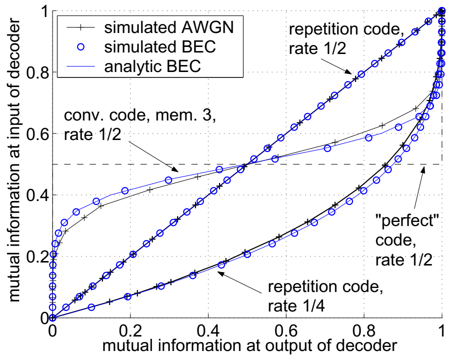

Figure 2: EXIT functions of different channel codes as outer codes in the serial concatenated system with AWGN (simulation results) and BEC (simulation and analytical results) input LLRs.

A weak channel code (e.g. a repetition code) gathers in general more extrinsic information at low input mutual information while at high input mutual information, a strong code (e.g. a convolutional code with high memory) outputs more extrinsic information. A "perfect" channel code allowing error free transmission at channel capacity would output ideal extrinsic information once the input mutual information exceeds the code rate, as illustrated in Fig. 2. A code with close to optimum performance could be e.g. an optimized irregular LDPC code with a high number of decoding iterations and with a very large block length [12].

An interesting property of the EXIT function $T_D(i)^{-1}$ of an outer decoder is that the area under $T_D(i)^{-1}$ is equal to the code rate [13] (compare EXIT functions of rate $1/2$ and rate $1/4$ codes in Fig. 2).

Analytical expressions for EXIT functions of several codes have been derived for BEC and BSC channels. As simple example, the EXIT function of the outer rate $1/n$ repetition code for BEC input information $i = 1 - p_c$ is given by $T_D(i) = 1 - p_c^{n-1}$ [13], where p_c is the probability of erasure of the input of the decoder. The analytical EXIT functions in Fig. 2 of a rate $1/2$ and $1/4$ repetition code match exactly the simulated curves for BEC input. The curves for BEC input are an upper bound for the curves with AWGN input.

Recently, polynomial expressions for the exact erasure probability of several convolutional codes on the BEC channel have been derived [14]. We will use these expressions to compute EXIT functions of outer convolutional codes for BEC input information. The main observation in [14] is that the possible state vectors in the forward and backward recursion steps in the APP decoding algorithm are finite and reasonably small in number. Using Markov chain theory, reasonably complex polynomials which describe the exact erasure probability are derived. The polynomial degree increases rapidly with the code memory, but the results are manageable up to memory 4 convolutional codes.

The extrinsic erasure probability p_e, required for EXIT charts, can be easily derived from the APP erasure probability p_{APP} and channel erasure probability p_c since with an APP algorithm, the output will be an erasure only if the estimate from the channel and the extrinsic estimate are erasures:

$$p_e = p_{APP}/p_c. \tag{2}$$

As example, the analytical EXIT function of a memory 3, rate 1/2 convolutional code is depicted in Fig. 2. The analytical results correspond exactly to the simulated curves with BEC input and differ only a little from the simulated curves with AWGN input.

4. EXIT FUNCTION OF MAPPING

The applied labeling map is a crucial design parameter for the considered BICM and BICM-ID systems. A mapping can be characterized basically in two ways:

The first approach is to use the Euclidean *distance spectrum* [15], that is the average number of bit errors if symbol errors occur at a specific Euclidean distance between the transmitted and the received signal. This approach is similar to the characterization of channel codes with the Hamming distance spectrum. Out of the Euclidean distance spectrum, characteristic values such as the *harmonic mean of the Euclidean distance* [2] can be extracted.

It is reasonable to characterize a mapping with two distance spectra: one for the case where *no* a priori information about the coded bits is available at the demapper and one for the case with *ideal* a priori information. With ideal a priori information, all the bits are perfectly known at the demapper, except for the bit to be detected, since only the extrinsic information is used. The signal constellation is reduced to a pair of signal points which differ only in the bit to be detected.

The second approach to characterize mappings is to use EXIT charts, i.e. the average mutual information between the coded bits $\mathbf{c}$ and the extrinsic LLRs after the demapper $L_e^M(\mathbf{c})$ as a function of the a priori information. For an APP demapper, this is equivalent to the mutual information between the coded bits $\mathbf{c}$ and the received symbol r as a function of the a priori information [13].

Fig. 3 depicts the EXIT functions $T_M(i)$ for an AWGN channel with Gray mapping, Ungerboecks Set Partitioning (SP) mapping [16] and $M16^a$ mapping [9]. The mappings are defined in Fig. 4.

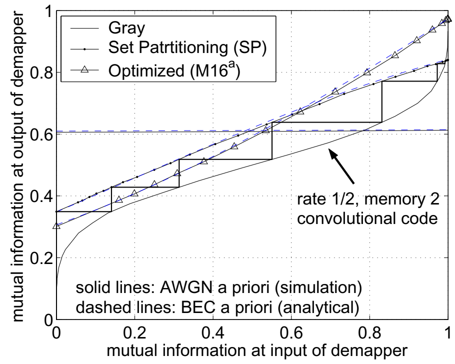

Figure 3: EXIT functions of 16QAM mappings, AWGN channel, $E_b/N_0 = 4$dB (with rate $1/2$ code).

map binary indices $\{0000,0001,0010,\ldots,1111\}$ to the corresponding signal points:
Gray: 1,2,4,3,5,6,8,7,13,14,16,15,9,10,12,11
SP: 10,11,13,16,12,9,15,14,4,1,7,6,2,3,5,8
$M16^a$: 13,6,7,16,3,12,14,5,8,15,9,2,10,1,4,11

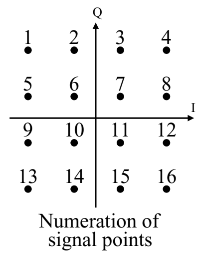

Figure 4: Definition of mappings.

As shown in [17], the channel capacity $C(\chi)$ of the 2^m-ary signal constellation χ (and of the BICM-ID system) is equal to $C(\chi) = A \cdot m$, where A denotes the area under the EXIT function $T_M(i)$. $C(\chi)$ is independent of the applied mapping. However, if no a priori information is used at the demapper, e.g. in the BICM system, it can be shown with the chain rule of mutual information that the "BICM capacity" is equal to $C_{BICM}(\chi) = m \cdot T_M(0)$.

This result has two important consequences: First, it restricts the shape of $T_M(i)$ since high $T_M(0)$ usually implies low $T_M(1)$ and vice versa. Second, recalling that the area under the EXIT function of the channel code corresponds to the rate (Section 3), the design of capacity approaching iterative systems reduces to a curve-fitting problem: If the two curves do not intersect and if the area between the two curves is minimized, the difference between the code rate and the capacity is minimized. Thus, we approach Shannon's capacity.

Concerning the mapping, high values of $T_M(0)$ (without a priori knowledge) and $T_M(1)$ (with ideal a priori knowledge) are desirable in order to avoid an early crossing with the EXIT function of the decoder, which would cause the iterative process to stop, and to reach low error rates respectively. As illustrated in Fig. 3, Gray mapping has the highest value $T_M(0)$, $M16^a$ mapping has an optimized value $T_M(1)$ and Set Partitioning (SP) mapping is a trade-off.

In order to match the EXIT functions of decoder (Fig. 2) and demapper (Fig. 3), we state that the steeper the EXIT function of the mapping is, the less powerful the channel code has to be and the more iterations between the decoder and demapper are required. With Gray mapping, iterations between decoder and demapper are not necessary but a powerful code is required. To further optimize the match of the EXIT functions, irregular codes may be used [8].

Fig. 5 depicts the EXIT functions of 16QAM with Gray and $M16^a$ mapping for a Rayleigh channel with independent fading coefficients for every symbol compared to an AWGN channel. The different E_b/N_0 are set to have a similar capacity of the Rayleigh and AWGN channel in order to obtain the same area under the EXIT function for a better comparison. We observe that with the $M16^a$ mapping, BICM-ID is not very robust against channel model variations: the slope of the function is lower in the Rayleigh than in the AWGN channel, requiring an other channel code for optimum iterative decoding.

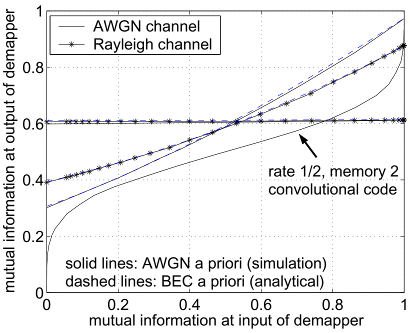

Figure 5: EXIT functions of Gray and $M16^a$ 16QAM mappings for a Rayleigh channel ($E_b/N_0 = 6.1$dB) and AWGN channel ($E_b/N_0 = 4$dB) with rate $1/2$ code.

In the following, we will derive an analytical expression for the demapper EXIT function $T_M(i)$ for an AWGN and Rayleigh channel and BEC a priori information. The approach is similar to the one presented in [17].

Let $\mathbf{C}_L$ denote a vector of L arbitrary a priori known bits C_j, $j \neq i$. The bit-wise mutual information $I(C_i;R|\mathbf{C}_L)$ describes the information that can be gained by the demapper about the transmitted bit C_i by observing the received symbol R and using the a priori known bits $\mathbf{C}_L$. For a specific selection of a priori known bits, $I(C_i;R|\mathbf{C}_L)$ is evaluated by numerical integration over the signal space $\mathbb{C}$ and by averaging over the possible values of C_i and $\mathbf{C}_L$:

$$I(C_i;R|\mathbf{C}_L) = \frac{1}{2^{L+1}} \cdot \sum_{c_i=0}^{1} \sum_{\forall \mathbf{c}_L} \int_{\mathbb{C}} p(a) \int_{\mathbb{C}} p(r|C_i = c_i, \mathbf{C}_L = \mathbf{c}_L) \cdot$$

$$\cdot \log_2 \frac{2 \cdot p(r|C_i = c_i, \mathbf{C}_L = \mathbf{c}_L)}{p(r|C_i = 0, \mathbf{C}_L = \mathbf{c}_L) + p(r|C_i = 1, \mathbf{C}_L = \mathbf{c}_L)} \, \mathrm{d}r \, \mathrm{d}a, \quad (3)$$

with

$$p(r|C_i = c_i, \mathbf{C}_L = \mathbf{c}_L)) = \frac{1}{2^{m-1-L}} \cdot \sum_{s \in \chi_{i,L}^{c_i,\mathbf{c}_L}} p(r|s), \quad (4)$$

where $\chi_{i,L}^{c_i,\mathbf{c}_L}$ denotes the subset of symbols $s \in \chi$ whose bit labels have the a priori known values $\mathbf{C}_L = \mathbf{c}_L$ in L positions and the value $C_i = c_i$ in position $i \in \{1,\ldots,m\}$. For an AWGN channel, the integration over the Rayleigh fading coefficient a can be omitted.

Averaging (3) over the $\binom{m-1}{L}$ possible a priori known bits and over the m bit positions gives

$$I_L = \frac{1}{m} \frac{1}{\binom{m-1}{L}} \sum_{m=0}^{m-1} \sum_{i=1}^{\binom{m-1}{L}} I(C_i;R|\mathbf{C}_L). \quad (5)$$

I_L is the average conditional mutual information given that L bits are a priori known at the demapper, $0 \leq L \leq m-1$. In the EXIT chart, we have the points $T_M(L/(m-1)) = I_L$ for the a priori mutual information $i = 1 - p_c = L/(m-1)$, where p_c is the erasure probability of the a priori information.

For a transition to continuous a priori knowledge, the mapping is repeated to obtain smaller fractions of a priori known bits [17]. The comparison of analytical and simulated EXIT functions with AWGN and Rayleigh channel is shown in Fig. 3 and 5.

To obtain *bit-wise* EXIT functions, we consider every bit position in the binary label separately. The analytical bit-wise EXIT functions are determined by omitting the averaging over the m bit positions in (5). Fig. 6a) and 6b) depict the simulated and analytical EXIT functions and the decoding trajectory split up in the four bit positions of the Gray and $M16^a$ 16QAM mappings. Those plots illustrate the different reliability of the bit positions.

A *symbol-wise* description of mappings is obtained if we fix the transmitted signal point and observe the L-values at the output of the demapper as a function of the a priori information. However, the transmitted bit c_i is then no more a variable and the mutual information between c_i and the received symbol R becomes zero. If we use the equation $I(C;L) = 1 - 1/N \cdot \sum_{n=1}^{N} \log_2(1 + \mathrm{e}^{-x_n \cdot L_n})$ [8] [11], where the bit $c_n \in \{0,1\}$ is mapped to $x_n \in \{+1,-1\}$ and N is the block length, the mutual information is a function of the distribution of the LLRs normalized by the sign of the transmitted bit, but no more a function of the specific bit sequence. Thus, we get a meaningful symbol-wise description, even though the specific bit sequence is not a variable.

In Fig. 6c) and 6d), we observe that the 16 symbols of the 16QAM Gray and $M16^a$ mappings are grouped in three and four groups respectively with similar characteristics.

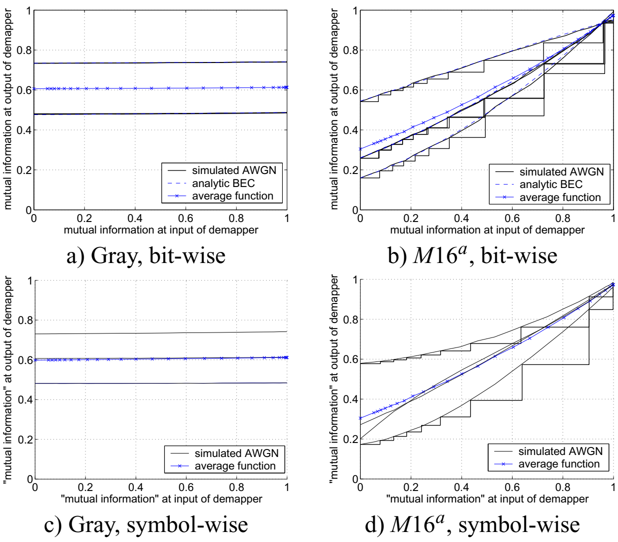

Figure 6: Bit-wise EXIT functions with trajectory and symbol-wise description.

5. NEW OPTIMIZED MAPPINGS

A method based on a binary switching algorithm is proposed in [9] to optimize the assignment of the binary labels to signal points.

Using the bit-wise mutual information and symbol-wise description derived in Section 4 in this optimization method, new mappings can be found for BICM-ID.

Only the value $T_M(1)$ of the EXIT function of the demapper has been optimized in [9], resulting in the $M16^a$ and $M16^r$ mappings for AWGN and Rayleigh channel respectively. To include $T_M(0)$ in the optimization, the binary switching algorithm requires a symbol-wise performance measure. With the symbol-wise description introduced in Section 4 we can now optimize the value of e.g. $T_M(0) + T_M(1)$ to avoid an early crossing of the EXIT functions *and* to reach low error rates. We found at $E_b/N_0 = 4$dB (with rate $1/2$ code) the optimized mapping $I16$, where the binary indices are mapped to the signal points $\{15,9,1,7,6,4,12,14,5,3,11,13,16,10,2,8\}$ according to the definition in Fig. 4.

If we want to have a different reliability of the bit positions in the binary label, we can optimize the bit-wise EXIT functions separately. The mapping $\{13,14,9,10,11,12,15,16,1,5,2,6,7,3,8,4\}$ has e.g. two bit positions with high reliability and two bit positions with low reliability.

6. SIMULATION RESULTS

The BER performance of several systems with iterative decoding is shown in Fig. 7 after 1 and 10 iterations, where the code rate is $1/2$, the interleaver length is 10000 bits, the signal constellation is 16QAM and the channel is an AWGN channel.

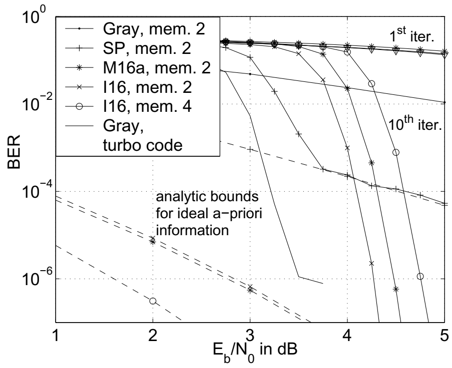

Figure 7: BER of BICM with iterative decoding for 16QAM, AWGN channel, $R = 1/2$ convolutional code and turbo code (2-bit/channel use).

The analytical error bounds with ideal a priori information are computed as described in [2], [6] using the Gauss-Chebyshev method. According to the prediction of the EXIT chart, a strong trade-off exists between the performance at low E_b/N_0 and the error bound at high E_b/N_0 for different mappings and different channel codes. The new $I16$ mapping with a memory 2 convolutional code is a good trade-off since it has a good performance at low E_b/N_0 and reaches similar error rates at high E_b/N_0 than the $M16^a$ mapping optimized for ideal a priori information. The combination of a turbo code with Gray mapping outperforms the other systems at low E_b/N_0 but has a higher error floor at high E_b/N_0. The advantage of BICM-ID systems is their flexibility and their good combination possibilities with e.g. iterative equalization and MIMO detection.

7. CONCLUSION

In this contribution, we use EXIT charts to illustrate the design possibilities of Bit-Interleaved Coded Modulation (BICM) with iterative decoding. Both the combination of a strong channel code with Gray mapping and the combination of a weak channel code with optimized mappings are promising approaches. EXIT functions of channel codes and mappings for different channels are analyzed and analytical expressions of these functions are derived. We use the bit-wise mutual information and a symbol-wise description to illustrate additional characteristics of mappings and to find new optimized mappings outperforming previously proposed ones.

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] 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.

[4] R. G. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, pp. 21–28, January 1962.

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

[6] 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.

[7] S. Bäro, "Turbo detection for MIMO systems: bit labeling and pre-coding," in *ITG Fachtagung*, Erlangen, January 2004.

[8] M. Tüchler, "Design of serially concatenated systems for short or long block lengths," accepted for *IEEE Transactions on Communications*, August 2003.

[9] F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Communications Letters*, vol. 7, no. 12, pp. 593–595, December 2003.

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

[11] J. Hagenauer, "The EXIT chart - introduction to the extrinsic information transfer in iterative processing," Submitted to EUSIPCO, 2004.

[12] S.-Y. Chung, D. Forney, T. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Communications Letters*, vol. 5, no. 2, pp. 58–60, February 2001.

[13] G. Kramer A. Ashikhmin and S. ten Brink, "Extrinsic information transfer functions: Model and erasure channel properties," submitted to *IEEE Transactions on Information Theory*.

[14] P. Siegel B. Kurkoski and J. Wolf, "Exact probability of erasure and a decoding algorithm for convolutional codes on the binary erasure channel," in *Proc. IEEE Globecom Conference*, San Francisco, December 2003.

[15] F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, "Optimized symbol mappings for bit-interleaved coded modulation with iterative decoding," in *Proc. IEEE Globecom Conference*, San Francisco, December 2003.

[16] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55–67, January 1982.

[17] 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.