EXTENDED MAPPINGS FOR BIT-INTERLEAVED CODED MODULATION

Patrick Henkel

Institute of Communications and Navigation (NAV)
Munich University of Technology (TUM)
80290 Munich, Germany
patrick.henkel@tum.de

ABSTRACT

The performance of Bit-Interleaved Coded Modulation with Iterative Decoding (BICM-ID) depends strongly on the applied mapping. We analyse the use of extended mapping schemes with more than one label per signal point: Assuming a fix signal constellation and a constant number of information bits per symbol, the increase in label length results in an ambiguous mapping scheme but also provides additional parity bits which allow the use of a stronger channel code.

We derive the optimum MAP detector and give some optimized mapping schemes for 16-QAM. Simulation results include an EXIT-Chart analysis and show the considerable reduction of the asymptotic bit error probability.

As a consequence of the more powerful mapping, we can obtain a very low bit error floor with extremely simple channel codes.

I. INTRODUCTION

Bit-Interleaved Coded Modulation is a bandwidth efficient transmission scheme built up of a serially concatenated encoder, bit interleaver and mapper [1] [2].

Several approaches of combinatorial mathematics, e.g. the Binary Switching Algorithm, the Reactive Tabu Search Algorithm and the Quadratic Integer Programming [3] have been investigated in [4] [5] for the optimization of the mapping. A distance spectrum for the mapping has been introduced in [4] to characterize the mapping. The free distance of the mapping is given by the first non-zero entry in the spectrum.

Irregular PSK signal constellations with a non-uniform angle between the signal points enable a considerable "offset gain" for high SNR [6]. Moreover, the use of larger label alphabets with more than one label per signal point is suggested for PSK signal constellations.

In this paper, we present the optimum MAP detector for BICM-ID with mappings that are based on extended label alphabets. The complexity of the MAP detector is analyzed by counting the number of additive operations and comparing them to a traditional mapping scheme.

We describe the fundamental trade-off underlying extended mapping schemes: Assuming a fixed signal constellation and transmission with a constant number of information bits per symbol, the increase in label length means an ambiguous mapping but also additional parity bits and a stronger channel code. The second effect predominates for a sufficiently high number of iterations and enables a considerable E_b/N_0 gain at moderate SNR.

In Section II, we introduce the system model and derive the optimum MAP detector for BICM-ID with extended mappings. The trade-off between the ambiguity of the mapping and a more powerful channel code is described in Section III. Section IV is devoted to the optimization of the extended mappings and includes some results for 16-QAM. The obtained mappings are characterized by a distance spectrum. Section V contains some simulation results that show the performance gains due to the increase of the label length. The paper is summarized in Section VI.

II. SYSTEM MODEL WITH EXTENDED MAPPINGS

We describe an extended system model of Bit-Interleaved Coded Modulation with Iterative Decoding (BICM-ID) depicted in Fig. 1. A block of l_I information bits $u_i \in \{\pm 1\}$ is encoded by a convolutional code of rate R_C and memory M. After random interleaving, a multiplexer partitions the bit stream into blocks of length l which are mapped to symbols s_t according to the labeling map μ and an N-ary signal constellation χ.

Note that the label length does not necessarily fulfill $l = \log_2(N)$. The use of a larger label alphabet does not change the channel capacity as long as an equal number of labels is assigned to each signal point. The received signal $y_t = s_t + n_t$
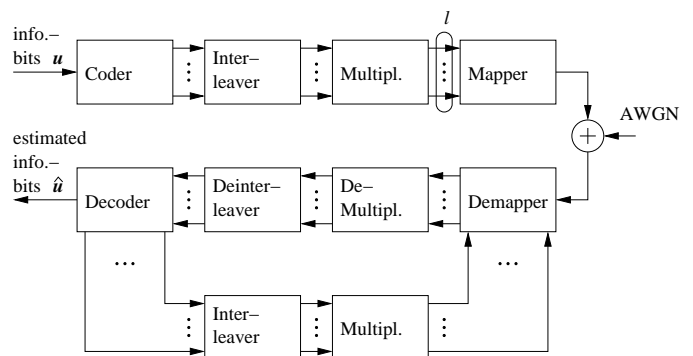

Figure 1: System model for Generalized Signal Shaping

is distorted by AWGN of variance $\sigma_n^2 = \frac{N_0}{2E_S}$ for each real dimension. For simplicity, we do not write any longer the time index t. The MAP detector computes the extrinsic reliability

information [7]:

$$L(x_k|y) = \ln\left(\frac{P(x_k = +1|y)}{P(x_k = -1|y)}\right)$$

$$\ln\left(\frac{\sum\limits_{\boldsymbol{x}|x_k=+1} \exp\left(-\frac{|y-s(\boldsymbol{x})|^2}{2\sigma_n^2} + \sum\limits_{j=1,j\neq k}^{l}\frac{x_j L(x_j)}{2}\right)}{\sum\limits_{\boldsymbol{x}|x_k=-1} \exp\left(-\frac{|y-s(\boldsymbol{x})|^2}{2\sigma_n^2} + \sum\limits_{j=1,j\neq k}^{l}\frac{x_j L(x_j)}{2}\right)}\right), \quad (1)$$

where the first addend of the argument of the exp function describes the geometric Maximum Likelihood (ML) decision metric and the second addend considers the a priori information from previous iterations.

Let us consider the complexity of equation (1) by counting the number of additions per L-value. There exist 2^{l-1} different a priori information vectors $\boldsymbol{x}|x_k = +1$ and each exponential function implies $l + 1$ summations leading to totally $(l+1)2^{l-1} + 2^{l-1} - 1$ additions in the numerator of equation (1). Consequently, the price of the extended labels is an increase in complexity of the demapper.

We can simplify the MAP detector of equation (1) by summation of symbols instead of labels such that the ML metric is computed only N times, i.e.

$$L(x_k|y) =$$

$$\ln\frac{\sum\limits_{s}\left[\exp\left(-\frac{|y-s(\boldsymbol{x})|^2}{2\sigma_n^2}\right)\sum\limits_{\substack{\boldsymbol{x}|\mu(\boldsymbol{x})=s \wedge \\ x_k=+1}}\exp\left(\sum\limits_{\substack{j=1 \\ j\neq k}}^{l}\frac{x_j L(x_j)}{2}\right)\right]}{\sum\limits_{s}\left[\exp\left(-\frac{|y-s(\boldsymbol{x})|^2}{2\sigma_n^2}\right)\sum\limits_{\substack{\boldsymbol{x}|\mu(\boldsymbol{x})=s \wedge \\ x_k=-1}}\exp\left(\sum\limits_{\substack{j=1 \\ j\neq k}}^{l}\frac{x_j L(x_j)}{2}\right)\right]}. \quad (2)$$

The numerator of equation (2) requires $2^l - N$ additions less than equation (1) and the ML metric can be reused in the denominator.

III. Trade-off between Ambiguity of Mapping and an Improved Channel Code

Traditional mapping schemes use an unambiguous mapping which assigns one bitlabel to one signal point. Consequently, the label length fulfills $l = \log_2(N)$.

Increasing the label length l results in a larger label alphabet and an ambiguous mapping with $2^l/N$ labels per signal point. The advantage of the increased label length is the option to use a more powerful channel code. Let us compare two different label alphabets of length l_1 and l_2 with $l_2 > l_1$: In both cases, we want to obtain the same spectral efficiency $\eta = R_{C,i} \cdot l_i$, $i \in \{1, 2\}$ whereas we assumed the same signal constellation χ. A larger label length l_2 means additional parity bits and, thus, a lower coding rate $R_{C,2}$.

Clearly, the unambiguous mapping ($l = \log_2(N)$) is optimum for non-iterative BICM where any ambiguity can not be removed. For BICM-ID, the ambiguity can be solved through the a priori information obtained over the iterations. After a sufficiently high number of iterations, the gain due to the more

powerful channel code predominates the loss due to the ambiguous mapping.

IV. Optimization of the Extended Mapping

We optimize the mapping with the aid of the Chernoff upper bound [2] on the symbol error probability $P(s \rightarrow \hat{s})$.

The cost function for the mapping without a priori information is obtained by summation over all possibly detected bits $k = \{1, \ldots, l\}$, all transmitted and received labels $\boldsymbol{x}$ and $\boldsymbol{y}$ which differ in the detected bit:

$$Z_{\mathrm{no}} = \frac{1}{l \cdot 2^{2(l-1)}} \sum_{k=1}^{l} \sum_{\boldsymbol{x}|\boldsymbol{x}_k=0} \sum_{\boldsymbol{y}|\boldsymbol{y}_k=1} \exp\left(-\frac{E_S}{4N_0}|\mu(\boldsymbol{x}) - \mu(\boldsymbol{y})|^2\right), \quad (3)$$

where we used the euclidian metric of [1] [4] for the AWGN channel. For ideal a priori information, the third summation of equation (3) is reduced to a single vector. We add up over all possibly a priori information vectors $\boldsymbol{x}_{\mathrm{ap}}$ and obtain

$$Z_{\mathrm{id}} = \frac{1}{l \cdot 2^{l-1}} \sum_{k=1}^{l} \sum_{\theta(\boldsymbol{x}_{\mathrm{ap}})=0}^{2^l-1} \exp\left(-\frac{E_S}{4N_0} \underbrace{|\mu(\boldsymbol{x}) - \mu(\boldsymbol{y})|^2}_{x_j \overset{!}{=} y_j \overset{!}{=} x_{\mathrm{ap},j} \ \forall j \neq k}\right), \quad (4)$$

where θ denotes the binary-to-decimal conversion operator and the labels $\boldsymbol{x}$ and $\boldsymbol{y}$ differ only in the detected bit ($x_k \neq y_k$).

A comparison of different optimization algorithms for the mapping is given in [5]. We investigate the Binary Switching Algoritm (BSA) which is initialized by a random assignment μ of the labels to the symbols. In each iteration, the cost of each label and the total cost are calculated with equations (3,4). The label with the largest cost is selected and a switch partner is searched such that the decrease in total cost is maximized. If no switch partner is found, the label with the second highest cost will be considered. This process continues until two switch partners are found. The realization of a switch completes an iteration. The BSA terminates when an iteration brings no further reduction of the total cost. 100 random initializations of the BSA are sufficient to find the global minimum of the cost function with high reliability [5].

Fig. 2 and 3 show optimized extended mappings for 16-QAM, $E_s/N_0 = 5$ dB, $l = \{5, 6\}$ and ideal a priori information.

We can derive a distance spectrum for the mapping from equations (3,4) by replacing the exp function with the $\delta[m]$ function which is defined by $\delta[0] = 1$ and $\delta[m] = 0$ for $m \neq 0$, i.e. we are interested only in the number of label pairs at certain euclidian distances d_E.

The distance spectrum with no a priori information can be written as

$$\Psi_{\mathrm{no}} = \frac{1}{l \cdot 2^{2(l-1)}} \sum_{k=1}^{l} \sum_{\boldsymbol{x}|\boldsymbol{x}_k=0} \sum_{\boldsymbol{y}|\boldsymbol{y}_k=1} \delta\left[m - d_E^2(\mu(\boldsymbol{x}), \mu(\boldsymbol{y}))\right]. \quad (5)$$
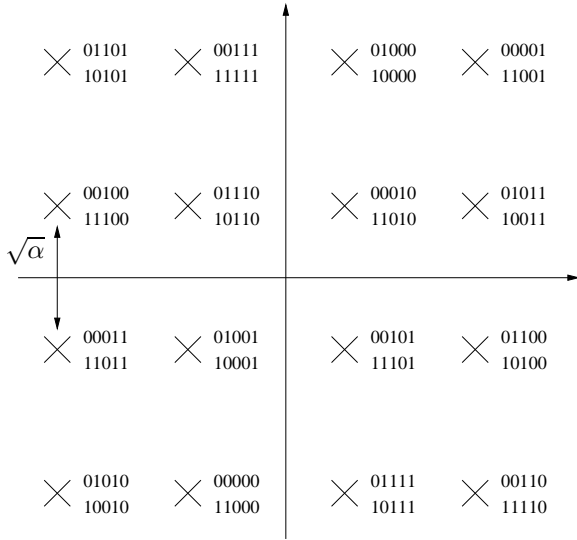
Figure 2: Extended mapping for 16-QAM with $l = 5$ and ideal a priori information
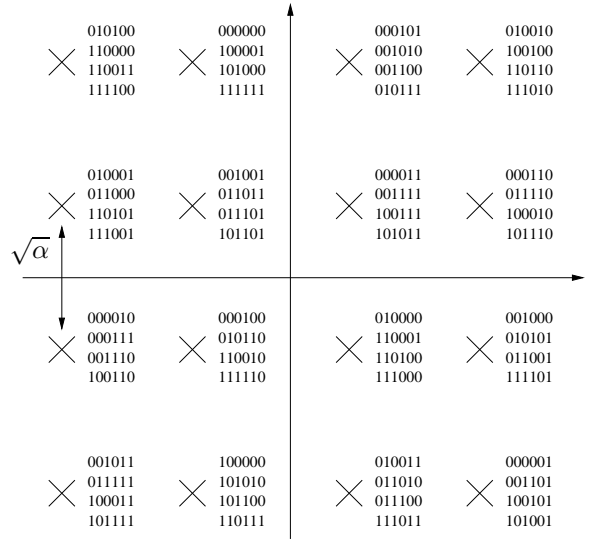

Figure 3: Extended mapping for 16-QAM with $l = 6$ and ideal a priori information

For ideal a priori information, we obtain from equation (4):

$$\Psi_{\mathrm{id}} = \frac{1}{l \cdot 2^{l-1}} \sum_{k=1}^{l} \sum_{\substack{\theta(\boldsymbol{x}_{\mathrm{ap}})=0}}^{2^l-1} \delta \left[\underbrace{m - d_E^2(\mu(\boldsymbol{x}), \mu(\boldsymbol{y}))}_{x_j \overset{!}{=} y_j \overset{!}{=} x_{\mathrm{ap},j} \ \forall j \neq k} \right]. \quad (6)$$

Tab. 1 and 2 compare the distance spectra of different label lengths for 16-QAM. We restrict to mappings which are optimized for ideal a priori information. The unambiguous mapping ($l = 4$) is characterized by a larger free distance d_f^{M} which is defined as the number of zeros entries until the first non-zero entry of the distance spectrum. The extended mapping of Fig. 3 benefits from a significantly lower $\Psi_{\mathrm{id}}(d_{\mathrm{E}}^2 = 5\alpha)$.

Table 1: Distance spectrum Ψ_{no} with no a priori inform.

d_{E}^2	0α	1α	2α	4α	5α	$\cdots$	18α
$l = 6$	0.04	0.19	0.15	0.15	0.19	$\cdots$	0.02
$l = 4$	0	0.20	0.16	0.16	0.19	$\cdots$	0.02

Table 2: Distance spectrum Ψ_{id} with ideal a priori inform.

d_{E}^2	0α	1α	2α	4α	5α	$\cdots$	18α
$l = 6$	0	0	0	0.03	0.34	$\cdots$	0
$l = 4$	0	0	0	0	0.50	$\cdots$	0

V. SIMULATION RESULTS

The iterative information transfer between the demapper and the decoder is analysed by EXtrinsic Information Transfer (EXIT) Charts [8]. The abscissa represents the mutual information I_{i} between the coded and interleaved information bits and the L-values at the input of the demapper whereas the ordinate shows the mutual information I_{o} between the coded and interleaved information bits and the L-values at the output of the demapper. For a more detailled introduction to EXIT-Charts we refer to [8].
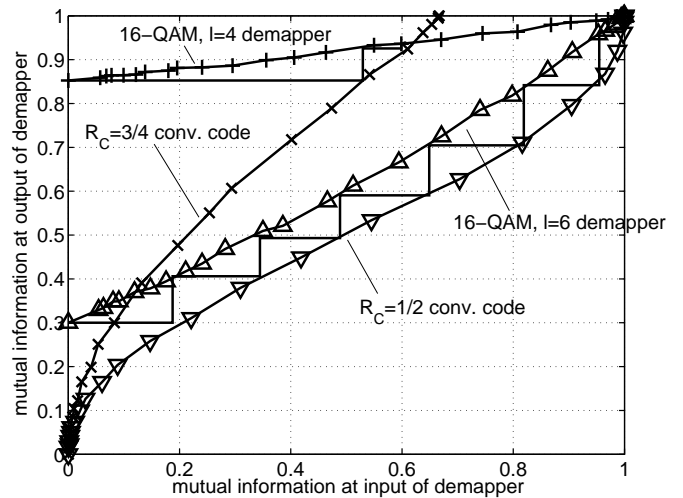

Figure 4: EXIT-Charts for $\eta = 3$, convolutional coding ($M = 1$), $10\log(E_b/N_0) = 8$ dB and label length $l = \{4, 6\}$

Fig. 4 shows the EXIT-Charts for 16-QAM and $E_b/N_0 = 8$ dB with $M = 1$ convolutional coding and label length $l \in \{4, 6\}$. The curve of the demapper with $l = 6$ starts with a considerably lower $I_{\mathrm{E}}(0)$ due to the ambiguity of the mapping. Moreover, the curves of the $M = 1$ convolutional codes differ significantly in form and shape: The rate $R_{\mathrm{C}} = 1/2$ code is characterized by an almost constant moderate slope such that the maximum $I_{\mathrm{E}}(1) = 1$ is obtained after nearly 10 iterations. On the contrary, the poorer $R_{\mathrm{C}} = 3/4$ code is decribed by a large slope for low I_{i} leading to an early intersection with the

curve of the corresponding demapper. The EXIT-Chart also visualizes the channel capacity which is independant of the applied label length, i.e.

$$C = l \cdot \int_0^1 I_o(I_i, l) dI_i = \text{const} \quad \forall \, l. \tag{7}$$

We compare the BER of two scenarios with common $\eta = 3$ and $M = 1$ but different $l = \{4, 6\}$ and $R_C = \eta/l$ in Fig. 5. A traditional label length means a very poor channel code which leads to a relatively small iterative gain even if the mapping is optimized for ideal a priori information. The extended mapping is characterized by a strong bit error cliff which results in an E_b/N_0 gain of 4 dB at an $\text{BER} = 10^{-3}$ and 6 dB at 10^{-6} over traditional schemes.
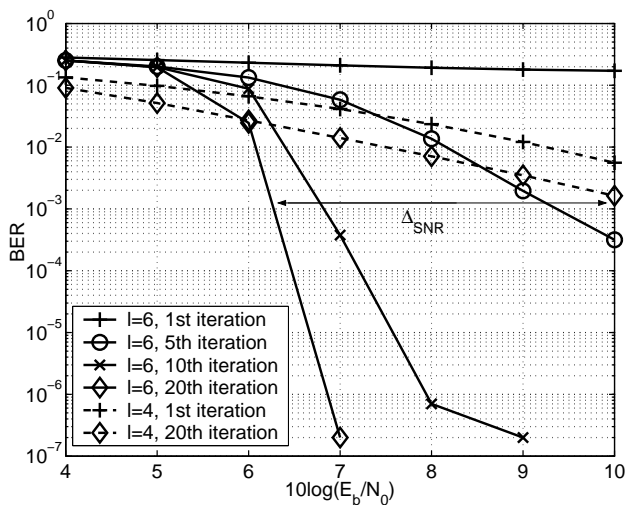

Figure 5: Comparison of BER for $\eta = 3$, convolutional coding ($M = 1$) and label length $l = \{4, 6\}$

One might think of increasing the memory length of traditional schemes because an $M = 1$ code is rarely used in practice. Nevertheless, an important advantage of extended mappings is the sufficiently low bit error floor even for $M = 1$. Fig. 6 depicts a comparison of BER for $\eta = 3$ where we combined $l = 4$ with $M = 2$ versus $l = 6$ with $M = 1$. A saturation appears for the first iteration of the $l = 6$ scenario due to the ambiguity of the mapping. The $l = 4$, $M = 2$ scheme benefits from an earlier bit error cliff but the $l = 6$, $M = 1$ scheme outperforms the traditional one in the bit error floor area.

When we assume the same constellation χ, the same spectral efficiency η (number of information bits per channel use) and the same code memory M, the extended mapping always shows a lower bit error floor than the traditional $l = \log_2(N)$ scheme.

VI. CONCLUSION

In this paper, we have introduced overdetermined mapping schemes with more than one label per signal point. The extended label alphabet results in an ambiguous mapping but also allows the use of a channel code with lower coding rate. The
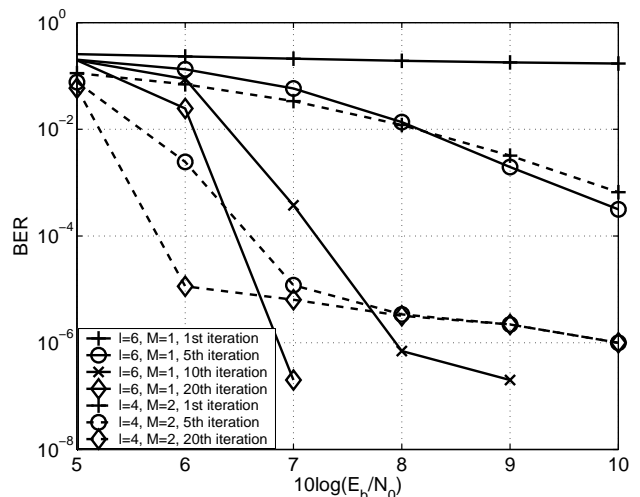

Figure 6: Comparison of BER for $\eta = 3$, convolutional coding with $l = 6$, $M = 1$ and $l = 4$, $M = 2$

effect of the stronger channel code predominates for high SNR and leads to a considerably reduced asymptotic bit error rate.

REFERENCES

[1] G. Caire, G. Taricco, and E. Biglieri, "Bit-Interleaved Coded Modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[2] S. Benedetto and E. Biglieri, "Principles of Digital Transmission," *Kluwer Academic/ Plenum Publisher*, 1999.

[3] E. Cela, *The Quadratic Assignment Problem*, Kluwer Academics, Norwell, MA, 1998.

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

[5] F. Schreckenbach, P. Henkel, N. Goertz and G. Bauch, "Analysis and Design of Mappings for Iterative Decoding of BICM," *XI National Symposium of Radio Sciences*, Poznan, Apr. 2005.

[6] T. Clevron and P. Vary, "Iterative Decoding of BICM with Non-Regular Signal Constellation Sets," *5th International ITG Conference on Source and Channel Coding (SCC 2004)*, Erlangen, Germany, Jan. 2004.

[7] J. Hagenauer, E. Offer, and L. Papke, "Iterative Decoding of Binary Block and Convolutional Codes," *IEEE Transactions on Information Theory*, vol. 42, pp. 429–445, Mar. 1996.

[8] S. ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[9] T. Cover and J. Thomas, *Elements of Information Theory*, J. Wiley and Sons, N.Y., 1991.