Multilevel Coding: Use of Hard Decisions in Multistage Decoding

Udo Wachsmann, Robert Fischer, Johannes Huber

Lehrstuhl für Nachrichtentechnik II (Digital Transmission and Mobile Communications)
Universität Erlangen–Nürnberg, Cauerstrasse 7, D-91058 Erlangen, Germany
Phone: +49-9131-857667, Fax: +49-9131-858849,
E-Mail: wachsman@nt.e-technik.uni-erlangen.de

Abstract — **Multilevel coding (MLC) schemes with powerful softly decodable component codes like turbo codes allow very power and bandwidth efficient digital communication [8, 2, 6]. But the use of the non–quantized channel output in each stage of the multistage decoding procedure leads to highly complex receivers. In this paper, the employment of hard decision decoding at several coding levels is proposed as a very efficient way to save complexity. It turns out that, for Ungerböck's labeling of signal points, the performance loss compared to soft decision decoding (in terms of capacity of the equivalent binary channel as well as in simulations) is only about 0.13 dB for ASK transmission over the AWGN channel, when hard decision decoding is employed at all but the lowest level. If error–correction techniques are exclusively employed, an MLC approach with ℓ individual binary error–correcting codes is recommended.**

1 Introduction

Multilevel coding (MLC) together with the relatively low complex multistage decoding (MSD) is well–known to be a theoretically optimum approach to coded modulation, cf. [6, 8]. Key point is the proper assignment of rates to the individual component codes. In [6, 2, 8] it is shown that properly designed MLC schemes with powerful softly decodable component codes lead to very power and bandwidth efficient digital communication schemes close to limits from information theory. The drawback is the relatively high complexity of the receiver due to soft decision maximum likelihood (ML) decoding in each stage. Complexity and decoding delay can be substantially reduced by using simple FEC coding at some levels instead. But hard decision decoding usually causes a substantial performance loss when compared to soft decision decoding. For example, the loss for binary antipodal signaling is about 2 dB due to hard decision decoding and an additional 1 dB due to bounded minimum distance (BMD) instead of ML decoding. In previous work, MLC schemes are mostly treated, where all component codes are either exclusively error–correcting codes like algebraic block codes, cf. e.g. [5] or softly decodable codes like convolutional or turbo codes [4, 8]. In this paper, a performance analysis of MLC schemes using hard decisions at some levels in MSD is presented, cf. also [7].

In Section 2 the underlying system model and the concept of equivalent channels at the individual levels of a MLC scheme is introduced. In Section 3 the capacity of coded modulation schemes with a combination of hard and soft decoded levels is derived. For the example of a three–level coded 8ASK modulation, the capacity results are discussed in Section 4. In order to confirm the theoretical results, simulation results for selected schemes are presented in Section 5.

2 System Model

Digital transmission by M–ary pulse amplitude modulation, $M = 2^\ell$, $\ell > 1$, with ℓ–level coding over the AWGN channel is considered. A binary address vector or label $\mathbf{x} = (x^0 x^1 \dots x^{\ell-1})$ is assigned to each signal point a_m out of the D–dimensional signal set $\mathbf{A} = \{a_0, \dots, a_{M-1}\}$. This mapping is usually derived from successive partitioning of the set $\mathbf{A}$ into subsets. Each subset at partitioning level i is uniquely labeled by the path $(x^0 \dots x^{i-1})$ in the partitioning tree from the root to the subset:

$$\begin{aligned}&\mathbf{A}(x^0 \dots x^{i-1}) = \\ &\{a | a \leftrightarrow \mathbf{x} = (x^0 \dots x^{i-1} b^i \dots b^{\ell-1}), \\ &\quad (b^i \dots b^{\ell-1}) \in \{0,1\}^{\ell-i}\}. \qquad (1)\end{aligned}$$

A similar paper was presented at the 35th Annual Allerton Conference on Communication, Control, and Computing, Urbana, Illinois, Sept./Oct., 1997

Here, the discussion is restricted to schemes with regular partitioning, where the subsets at one partitioning level differ only by translation and rotation, i.e. they provide equal capacities.

The core of the MLC approach is to partition a block of data symbols into ℓ levels and to encode them individually by component codes of rates R^i and equal lengths N into the binary address symbols x^i. The total rate of the scheme is $R = \sum_{i=0}^{\ell-1} R^i$. Key point to create powerful MLC schemes is the proper assignment of code rates R^i to the component codes. Since overall maximum likelihood decoding is not feasible in practice, the component codes are successively decoded starting with the level 0 code. Thereby, not only the channel output $y \in \mathbb{R}^D$ but also decisions $\hat{x}^j$, $j = 0, \ldots, i-1$, from previous decoding stages j, $j < i$, are used for decoding at level i.

Due to the individual encoding at each level, the transmission of address vector $\mathbf{x}$ with binary digits x^i, $i = 0, \ldots, \ell - 1$, over the physical channel can be virtually separated into parallel transmission of each digit x^i over ℓ equivalent channels [8, 6]. When hard decision decoding is used at decoding stage i, the continuous channel output y is quantized to a binary symbol $\tilde{x}^i$.

For example, let us consider the equivalent channel for transmission of symbol x^1 for an 8ASK scheme when hard decision decoding is used, see Fig. 1. With the restriction to equiprobable si-

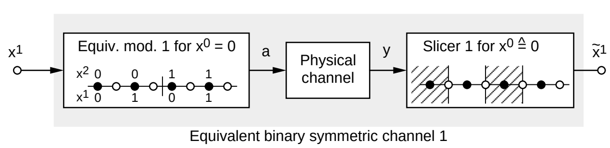

Figure 1: Equivalent BSC 1 for the transmission of symbol x^1 for an 8ASK scheme when hard decision decoding is used and symbol $x^0 = 0$ is assumed at the lowest level.

gnal points, the optimum detection thresholds are half–way between two adjacent points. Due to the multiple representation of one binary symbol by signal points, the decision region of binary symbol $\tilde{x}^i$ is the union of the corresponding regions $\mathcal{R}_{\mathbf{A}(x^0 \ldots x^{i-1})}(a)$, see Fig. 1:

$$\mathcal{R}(\tilde{x}^i | x^0 \ldots x^{i-1}) = \bigcup_{a \in \mathbf{A}(x^0 \ldots x^{i-1} \tilde{x}^i)} \mathcal{R}_{\mathbf{A}(x^0 \ldots x^{i-1})}(a), \tag{2}$$

where the decision region $\mathcal{R}_{\mathbf{A}}(a)$ of signal point $a \in \mathbf{A}$ is defined by

$$\mathcal{R}_{\mathbf{A}}(a) = \{ y | d_E(y, a) \leq d_E(y, a'), a' \in \mathbf{A}, a \neq a' \}. \tag{3}$$

$d_E(y, a)$ denotes the Euclidean distance between y and a.

For equiprobable signal points and regular partitioning, the resulting equivalent channel is a *binary symmetric* channel (BSC). To distinguish between equivalent channel i with hard and soft decision decoding, we refer to the former one as equivalent BSC i. Capacity of the equivalent channel i without quantization is denoted by C^i, whereas capacity of the equivalent BSC i is denoted by C_H^i (index H indicates hard decision decoding).

3 Capacity

Since we consider capacities, the digits at lower levels can be assumed to be estimated without errors, as long as rates R^j are chosen not to be greater than capacities C^j or C_H^j, $j < i$, resp. Hence, the capacity of the equivalent BSC i is simply given by

$$C_H^i = 1 - \mathrm{H}_2(\varepsilon_i), \tag{4}$$

where ε_i denotes the bit error probability of the equivalent BSC i and $\mathrm{H}_2(\cdot)$ is the binary entropy function. The only task for calculating C_H^i is to determine the transition probability $\Pr\{\tilde{x}^i \neq x^i | x^i, x^0 \ldots x^{i-1}\} = \varepsilon_i$. The particular choice of lower level digits $(x^0 \ldots x^{i-1})$ has no influence on the transition probability since regular partitioning and equiprobable signal points are considered.

First, let us assume that $x^i = 0$ is transmitted and represented by a special signal point a out of the subset $\mathbf{A}(x^0 \ldots x^{i-1} 0)$. The error event of detecting $\tilde{x}^i = 1$ occurs if the channel output y falls in the decision region $\mathcal{R}_{\mathbf{A}(x^0 \ldots x^{i-1})}(\tilde{a})$ of any signal point $\tilde{a}$ out of the subset $\mathbf{A}(x^0 \ldots x^{i-1} 1)$ representing the binary symbol $x^i = 1$. Thus, the probability of detection error is given by

$$\begin{aligned} &\Pr\{\tilde{x}^i = 1 | a, x^0 \ldots x^{i-1}\} \\ &= \Pr\{\tilde{a} \in \mathbf{A}(x^0 \ldots x^{i-1} 1) | a, x^0 \ldots x^{i-1}\} \\ &= \sum_{\tilde{a} \in \mathbf{A}(x^0 \ldots x^{i-1} 1)} \Pr\{\tilde{a} | a, x^0 \ldots x^{i-1}\}; \\ &\quad \text{for} \quad a \in \mathbf{A}(x^0 \ldots x^{i-1} 0). \end{aligned} \tag{5}$$

Here, $\Pr\{\tilde{a} | a, x^0 \ldots x^{i-1}\}$ denotes the probability that y falls into $\mathcal{R}_{\mathbf{A}(x^0 \ldots x^{i-1})}(\tilde{a})$, when signal point

a is transmitted. Detecting any signal point $\tilde{a} \in \mathbf{A}(x^0 \dots x^{i-1}0)$, $\tilde{a} \neq a$, which also represents the symbol $x^i = 0$, does not lead to an error. The probability of detection error is upper bounded by the conditioned symbol error probability of the underlying signal subset for transmitting $a \in \mathbf{A}(x^0 \dots x^{i-1}0)$, namely

$$\begin{aligned} &\Pr\{\tilde{x}^i = 1 | a, x^0 \dots x^{i-1}\} \\ &\leq \Pr\{\tilde{a} \in \mathbf{A}(x^0 \dots x^{i-1}), \tilde{a} \neq a | a, x^0 \dots x^{i-1}\} \\ &= \sum_{\tilde{a} \in \mathbf{A}(x^0 \dots x^{i-1}), \tilde{a} \neq a} \Pr\{\tilde{a} | a, x^0 \dots x^{i-1}\}, \\ &\quad \text{for} \quad a \in \mathbf{A}(x^0 \dots x^{i-1}0). \end{aligned} \tag{6}$$

For Ungerböck's set partitioning, the most relevant case in practice, all signal points closest to the transmitted one represent the inverse binary symbol. Here, the detection error probability is mainly determined by the nearest neighbor signal points, and, hence, approximation by the symbol error probability (6) is quite tight. For example, the bit error probabilities ε_i, $i = 0, \dots, \ell - 1$, for a 2^ℓ–ary ASK scheme with $a \in \mathbf{A} = \{\pm 1, \dots, \pm(2^\ell - 1)\}$ are simply approximated by the well–known formulae for the symbol error probabilities for $2^{\ell - i}$–ary ASK.

With Eq. (5) the probability of detecting symbol $\tilde{x}^i = 1$ for transmission of symbol $x^i = 0$ is given by the expected value of $\Pr\{\tilde{x}^i | a, x^0 \dots x^{i-1}\}$ over all possible signal points a representing symbol $x^i = 0$:

$$\begin{aligned} &\Pr\{\tilde{x}^i = 1 | x^i = 0, x^0 \dots x^{i-1}\} \\ &= \mathrm{E}_{a \in \mathbf{A}(x^0 \dots x^{i-1}0)} \left\{ \Pr\{\tilde{x}^i | a, x^0 \dots x^{i-1}\} \right\}. \end{aligned} \tag{7}$$

The results derived above can simply be generalized to the case of an arbitrary probability distribution of signal points and arbitrary labeling, cf. [7].

The capacity of a coded modulation scheme with a mix of hard and soft decision decoding at the individual levels is obtained by the sum over the capacities of the corresponding equivalent channels i, $i = 0, \dots, \ell - 1$. For example, consider a 3–level scheme with soft decision decoding at level 0 and hard decision decoding at levels 1 and 2:

$$C_{SHH} = C^0 + C_H^1 + C_H^2. \tag{8}$$

The index of the capacity C_{SHH} denotes the decoding method (soft or hard) at the individual levels. If the index is omitted soft decision decoding at all levels is assumed.

It was shown in [8, 6] that, in the case of soft decision decoding at all levels, MLC together with MSD is an asymptotically optimum coding approach. However, when hard decision decoding at several levels is used, this statement does not hold any longer, and hence, the MLC approach is not necessarily optimum. Thus, the MLC/MSD transmission scheme with hard decision decoding at all levels operating on the AWGN channel is compared to a coded modulation scheme operating on a M–ary *discrete memoryless channel* (DMC). The transition probability $\Pr\{a_k | a_j\}$ for receiving symbol a_k is given by the probability that the output y of the underlying AWGN channel falls into the decision region $\mathcal{R}_{\mathbf{A}}(a_k)$ if symbol a_j is transmitted. The capacity of this scheme is given by

$$C_{DMC} = \mathrm{E}_{(a_k \in \mathbf{A}, a_j \in \mathbf{A})} \left\{ \log_2 \frac{\Pr\{a_k | a_j\}}{\Pr\{a_k\}} \right\}. \tag{9}$$

Finally, we compare all these methods of coded modulation with a traditional FEC approach where the 2^ℓ–ary modulation scheme is actually ignored. Hard decision decoding together with short bit interleaving provides a memoryless BSC. In order to minimize the bit error probability ε of the resulting BSC, Gray mapping, $a = \mathcal{M}_G(x^0, \dots, x^{\ell-1})$ is applied. The capacity reads

$$C_{FEC} = \ell \cdot (1 - \mathrm{H}_2(\varepsilon)). \tag{10}$$

For the bit error probability ε we obtain

$$\begin{aligned} \varepsilon &= \frac{1}{\ell} \mathrm{E}_{(a_k \in \mathbf{A}, a_j \in \mathbf{A})} \left\{ \Pr\{a_j\} \Pr\{a_k | a_j\} \cdot \right. \\ &\quad \left. d_H(\mathcal{M}_G^{-1}(a_k), \mathcal{M}_G^{-1}(a_j)) \right\}. \end{aligned} \tag{11}$$

where $d_H(\cdot)$ denotes the Hamming distance.

4 Examples and Discussion

For example, 3–level coded 8ASK modulation with natural labeling is investigated. In Fig. 2 the

- capacity C, i.e. soft decision decoding at all levels (case SSS, reference),
- capacity C_{SHH}, i.e. soft decision decoding at level 0 and hard decision decoding at levels 1 and 2 (case SHH),
- capacity C_{HHH}, i.e. hard decision decoding at all levels (case HHH),

- capacity C_{DMC} of the 8–ary DMC (dashed line),
- capacity C_{FEC} (dashed–dotted line),
- capacities of the equivalent channels and the equivalent BSCs

are depicted.

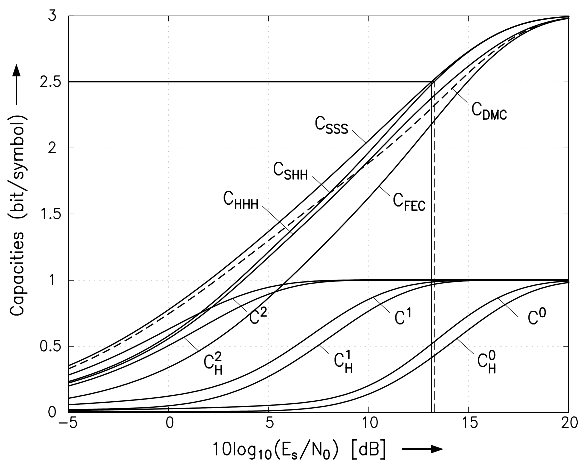

Figure 2: Capacities of 8–ary ASK over the AWGN channel for decoding strategies SSS, SHH, and HHH (see text) and corresponding capacities C^i and C^i_H for the equivalent and equivalent binary symmetric channel i, $i = 0, 1, 2$.

Firstly, we look at the capacities C^i of the equivalent channels i and C^i_H, $i = 0, 1, 2$, of the equivalent BSCs i, respectively. At level 2, the underlying signal set for the transmission of symbol x^2 is the ordinary BPSK signal set, i.e. symbol x^2 is represented uniquely. In this case, the well–known gap of about 1.7 dB between the capacities $C^2 = C^2_H = 0.5$ can be observed. At level 1, symbol x^1 is represented double in the underlying 4ASK signal set. Here, the gap between the capacities $C^1 = C^1_H = 0.5$ reduces to 1.1 dB. This gap is further reduced to 0.9 dB at level 0. Thus, for transmission where symbols are multiply represented by signal points the gap between soft and hard decision decoding becomes smaller. We interpret this different behavior by the following observation: For binary antipodal signaling the gain of soft decision decoding mainly results from those highly reliable symbols which are received far away from the (hard) decision boundary. But for multiamplitude/multiphase modulation the existence of boundaries of the decision regions on all sides of inner points rather excludes such points and thus reduces the soft decision advantage. Additionally, as usual the gap between capacities C^i and C^i_H decreases for increasing SNR.

Secondly, the cases SSS and SHH are compared. It is apparent from Fig. 2 that the gap between C and C_{SHH} is negligible for rates above 2.5 bit/symbol. In particular, we focus on the rate 2.5 bit/symbol, since 0.5 bit redundancy per dimension is sufficient to achieve near optimum performance as long as equiprobable signal points are used, cf. e.g. [2]. For $R = 2.5$ bit/symbol the loss for the case SHH vs. the optimum case SSS is 0.13 dB. Even for rates down to 2.0 bit/symbol only a loss of max. 0.6 dB is visible. Hence, the performance loss for a coded modulation scheme due to hard decision decoding at higher levels is dramatically reduced when compared to the case of BPSK. The explanation is as follows. For set partitioning according to Ungerböck's criterion, the individual rates increase from lowest level to highest level. Thus, the performance loss due to hard decision decoding decreases. Hence, if hard decision decoding is only used at higher levels, where high–rate codes are active, the loss remains small.

An additional loss occurs if maximum likelihood decoding is replaced by bounded minimum distance decoding. This loss cannot be assessed exactly, but binary codes with rate very close to 1 are not far away from the sphere–packing bound and hence the difference between bounded minimum distance decoding and hard decision maximum likelihood decoding is not essential. In the case SHH, where only high–rate codes are used for hard decision decoding, there is only a little additional loss for practical bounded minimum distance decoding algorithms.

If we design the system for a total rate $R = 2.5$ bit/symbol, the rate distribution according to capacities in the case SSS (solid vertical line) is

$$R^0/R^1/R^2 = 0.516/0.984/1, \qquad (12)$$

whereas in the case SHH (dashed vertical line)

$$R^0/R^1/R^2 = 0.532/0.968/1 \qquad (13)$$

results. For SHH rate at level 0 with soft decision decoding is slightly increased while rate at level 1 with hard decision decoding is decreased by the same amount when compared to SSS. This can be interpreted as the increased rate at level 0 compensates for the loss due to hard decision decoding at level 1. On the other hand, if the

SNR of 13.13 dB (capacity limit of SSS for total rate 2.5 bit/dimension) is fixed, the rate design

$$R^0/R^1/R^2 = 0.516/0.964/1 \qquad (14)$$

in the case SHH results. Thus, SHH decoding corresponds to a very small rate decrease of 0.02 bit/dimension when compared to SSS.

Next, we look at hard decision decoding at all levels (HHH). For a total rate $R = 2.5$ bit/symbol, the gap between C and C_{HHH} is about 0.9 dB (see Fig. 2). Here, the loss due to exclusively hard decision decoding in a coded modulation scheme is substantially reduced compared to BPSK. The reason is that for Ungerböck's set partitioning, the lowest rate is transmitted at level 0 and, hence, the capacity loss at this level dominates. But, as shown above, at level 0 the loss due to hard decision decoding is moderate because of the multiple representation of the symbol x^0 in the underlying signal set.

Concluding we see that for 2^ℓ–ary ASK and $R = \ell - 0.5$ bit/symbol, it is sufficient to employ soft decision decoding only at level 0. Hard decision decoding at higher levels can be done without any remarkable performance loss while offering a reasonable reduction in complexity. Since in general the effort for hard decision decoding is substantially lower than for soft decision decoding, we can state that power efficient coded 2^ℓ–ary, $\ell > 1$, modulation requires only slightly additional complexity when compared to coded binary transmission. Thus, bandwidth efficient digital transmission with a multiamplitude/multiphase constellation close to capacity limit requires much less decoding complexity per bit than binary antipodal signaling. Even if hard decision decoding is used at all levels the capacity loss compared to soft decision decoding is lower than 1 dB. Notice that in the interesting region near $R = \ell - 0.5$ bit/dimension the curve C_{FEC} shows exactly the same gap of 1.8 dB to the capacity C of the AWGN channel as in the case of binary antipodal signaling (i.e. $\ell = 1$). The difference between C_{DMC} and C_{FEC} is due to the memory destroying process of interleaving.

Finally, the differences in capacities C_{DMC} (8–ary one–level coding for the 8–ary DMC), C_{FEC} (binary one–level coding for 8ASK, Gray mapping, and bit interleaving) and C_{HHH} (MLC/MSD with hard decision decoding at each level) are discussed. It can be observed from Fig. 2 that the MLC approach with hard decision decoding at each level substantially outperforms the one–level coding schemes. The difference between the schemes is the manner how the soft output of the underlying AWGN channel is exploited. For the 8–ary DMC as well as for one–level FEC the soft channel output is quantized once. In the staged decoding of a MLC scheme, the quantization of the soft value at each level in order to perform the respective binary detection is controlled by the decoding results at lower levels. This leads to the important result that, in the case of hard decision decoding, the most efficient way to exploit the channel information is to split up the coding and decoding procedure in multiple, ideally binary, levels and, thus, to use the soft channel output multiply without the need of complex soft decoding algorithms. That means, 1 dB gain is simply possible by quantizing the channel output step–by–step without any additional complexity. Therefore, dealing with design proposals for coded modulation schemes, the two cases soft and hard decision decoding have to be carefully distinguished. As shown in [1, 7], in the case of soft decision decoding and a very large block length it is possible to amalgamate several coding levels to a single one without remarkable performance loss when Gray mapping is used. In the case of hard decision decoding, this is not true since combining multiple levels to a single one results in an unavoidable performance loss due to an suboptimum exploitation of the soft information. It is well–known that the adoption of a single binary error–correcting code to a M–ary transmission scheme by Gray mapping does not lead to satisfying results. In contrast, the MLC approach with individual error correcting codes at each level promises quite better performance.

5 Simulation Results

In order to verify the presented capacity results, simulations for 8ASK transmission with a MLC/MSD scheme over the AWGN channel were performed. In particular, we are interested in the loss of a MLC scheme using hard decision decoding at the 2 higher levels when compared to a scheme using entirely soft decision decoding. The individual rates are chosen equal to the capacities of the respective equivalent channels, cf. [3, 8, 6]. For reference, the MLC scheme with the individual rates $R^0/R^1/R^2 = 0.52/0.98/1.0$ is used, where turbo codes over 16–state convolutional codes and code word length $N = 4095$

are employed as component codes. Flexible rates of turbo codes are achieved via puncturing, see [8]. For the competing scheme, the same turbo code with $R^0 = 0.52$ is employed at level 0 and a primitive BCH code of length $N = 4095$ at level 1. Level 2 remains uncoded. The error correcting capability of the BCH code is adapted in such a way that the individual performances at level 0 and level 1 are similar. As a result, the required error correcting capability for the BCH code is $t = 14$ errors and, hence, $R^1 = 0.96$, which is exactly equal to the capacity C^1 in (14). The simulation results are depicted in Fig. 3. For the

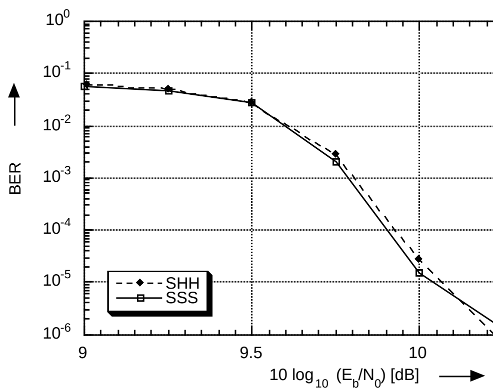

Figure 3: BER of 8ASK transmission with MLC/MSD over the AWGN channel. Code word length $N = 4095$. Solid line (case SSS): Rate distribution $R^0/R^1/R^2 = 0.52/0.98/1.0$, total rate $R = 2.5$ bit/symbol. Turbo codes are employed as component codes. Dashed line (case SHH): Rate distribution $R^0/R^1/R^2 = 0.52/0.96/1.0$, total rate $R = 2.48$ bit/symbol. Component codes: Turbo code at level 0, 14–errors–correcting BCH code at level 1. Simulation results.

reference scheme $10\log_{10}(E_b/N_0) = 10.04$ dB is required to achieve BER $\leq 10^{-5}$ with a total rate $R = 2.5$ bit/symbol. Since the capacity $C = 2.5$ bit/symbol is reached for $10\log_{10}(E_b/N_0) = 9.15$ dB, this scheme works about 0.89 dB above capacity. From Fig. 3, it can be seen that the competing scheme with the BCH code at level 1 achieves BER $\leq 10^{-5}$ at $10\log_{10}(E_b/N_0) = 10.06$ dB with a total rate $R = 2.48$ bit/symbol. Here, the capacity $C = 2.48$ bit/symbol is reached for $10\log_{10}(E_b/N_0) = 9.04$ dB. Hence, the MLC scheme using a BCH code at level 1 works about 1.02 dB above capacity, resulting in a loss of about 0.13 dB vs. the MLC scheme using turbo codes at both coded levels. Thus, the loss of 0.13 dB obtained by simulations coincides with the loss predicted from capacity argument.

Acknowledgment

The authors are greatly indebted to *Deutsche Forschungsgemeinschaft* for supporting this work. Additionally, the authors like to thank Rolf Weber [9] for performing some of the calculations in this paper.

References

[1] G. Caire, G. Taricco, and E. Biglieri. Capacity of bit–interleaved channels. *Electronics Letters*, vol.32:pp.1060–1061, June 1996.

[2] R. Fischer, J. Huber, and U. Wachsmann. Multilevel coding: Aspects from information theory. In *Proc. CTMC at IEEE Global Telecommun. Conf. (GLOBECOM'96)*, pages 26–30, London, UK, Nov. 1996.

[3] J. Huber and U. Wachsmann. Capacities of Equivalent Channels in Multilevel Coding Schemes. *Electronics Letters*, vol. 30:pp. 557–558, March 1994.

[4] Y. Kofman, E. Zehavi, and S. Shamai (Shitz). Performance Analysis of a Multilevel Coded Modulation System. *IEEE Trans. Commun.*, vol.42:pp.299–312, 1994.

[5] G.J. Pottie and D.P. Taylor. Multilevel Codes based on Partitioning. *IEEE Trans. Inf. Theory*, vol.35:pp.87–98, Jan. 1989.

[6] U. Wachsmann, R. Fischer, and J. Huber. Multilevel Coding: Basic Concepts, Capacity, and Rate Design. Submitted to IEEE Trans. Inf. Theory, 1997.

[7] U. Wachsmann, R. Fischer, and J. Huber. Multilevel Coding: Dimensionality of the Constituent Signal Set, Labeling, and Hard Decision Decoding. Submitted to IEEE Trans. Inf. Theory, 1997.

[8] U. Wachsmann and J. Huber. Power and bandwidth efficient digital communication using turbo codes in multilevel codes. *Europ. Trans. Telecommun. (ETT)*, vol. 6:pp. 557–567, Sept.-Oct. 1995.

[9] R. Weber. *Einsatz von fehlerkorrigierenden Block–Codes in der codierten Modulation.* Diploma thesis, Lehrstuhl für Nachrichtentechnik, Universität Erlangen-Nürnberg, Germany, 1997. In German.