Optimizing BICM with Convolutional Codes for Transmission over the AWGN Channel

Clemens Stierstorfer, Robert F.H. Fischer, and Johannes B. Huber
Lehrstuhl für Informationsübertragung, Friedrich–Alexander–Universität Erlangen–Nürnberg
Cauerstraße 7/LIT, 91058 Erlangen, Germany, Email: {clemens,fischer,jbhuber}@LNT.de

***Abstract*—The usual comparison of trellis coded modulation and bit-interleaved coded modulation, both using convolutional codes and Viterbi decoding, leads to the well-known result that for the AWGN channel TCM clearly outperforms BICM. In fading scenarios BICM shows superior results. Based on recent results on optimized bit mappings and bit-interleaver designs for BICM, we demonstrate that the BER of BICM on AWGN channels can be significantly lowered at no additional cost. Depending on the signal constellation size and the constraint length of the convolutional code gains up to 7 dB can be achieved over BICM with random bit interleavers.**

I. Introduction

The comparison of trellis coded modulation (TCM) [8] and bit-interleaved coded modulation [11] as for example performed in [2], leads to the well-known result that on the additive white Gaussian noise (AWGN) channel TCM outperforms BICM in terms of capacity and bit error ratio (BER). In fading scenarios, on the contrary, BICM shows superior results. The provided numerical results are mostly based on non-iteratively decoded implementations using convolutional codes and Viterbi decoders. This classic variant of BICM recently was identified as optimum solution for low-delay applications [4]. Here, we show that in part the performance of conventional BICM on AWGN channels has been considerably underestimated so far.

For the analysis of BICM an equivalent channel model initially introduced in the context of multilevel codes [5], [9] has proven to be helpful. Based on this model we recently studied bit mappings for BICM transmission [6] and optimized the design of the bit interleaver [7]. In particular the latter offers a large potential for optimizations at no additional costs.

In this contribution, our recent insights are combined with the knowledge that in non-fading scenarios bit interleaving is not necessarily beneficial, but may be even counter-productive [9]. We propose some slight modifications for BICM transmission over AWGN channels which at higher spectral efficiency significantly lower the resulting BER.

II. General System Model

We investigate block-based coded transmission over an AWGN channel (see Fig. 1). A rate-$R_\mathrm{c} = k/n$ convolutional encoder (ENC) is used to encode a sequence $\boldsymbol{q}$ of K_bs binary source bits q_κ, $\kappa = 1, \ldots, K_\mathrm{bs}$, originating from a discrete i.i.d. memoryless source, into a binary sequence $\boldsymbol{c}$ of $N_\mathrm{bs} = K_\mathrm{bs}/R_\mathrm{c}$ encoded symbols c_ν, $\nu = 1, \ldots, N_\mathrm{bs}$. The symbol rate of the source bits is denoted by $1/T_\mathrm{b}$, that of the encoded bits by $1/T_\mathrm{c}$. The sequence of encoded bits is passed on to a block bit interleaver Π which permutes the encoded binary symbols and generates an output stream of R_M-tuples $\boldsymbol{x} = [x_\delta^{(1)}, x_\delta^{(2)}, \ldots, x_\delta^{(R_M)}]$ with index $\delta = 1, \ldots, \Delta$; $\Delta = N_\mathrm{bs}/R_M$. These binary R_M-tuples $\boldsymbol{x}$ are then bijectively mapped onto channel symbols a

$$\mathcal{M} : \boldsymbol{x}_\delta \in \mathbb{F}_2^{R_M} \mapsto a_\delta \in \mathcal{A} \subset \mathbb{C}\,. \tag{1}$$

Here, $\mathcal{A}$ denotes an M-ary signal constellation and $M = 2^{R_M} = |\mathcal{A}|$ holds. For simplicity, we restrict our considerations to M-ary amplitude-shift keying (ASK) constellations ($\mathcal{A} = \{\pm 1, \pm 3, \ldots, \pm(M-1)\}$), i.e., we focus on one of the quadrature components of an M^2-QAM constellation. If $\mathcal{M}$ is a Gray mapping the two quadrature components of a QAM constellation are independent an can be processed successively.

The received and AWGN-corrupted signal reads

$$y_\delta = a_\delta + n_\delta\,. \tag{2}$$

The variance of the channel symbols[1] is given as $\sigma_\mathrm{A}^2 = E_\mathrm{s}/T_\mathrm{s}$ with the average energy per symbol E_s and the channel symbol rate $1/T_\mathrm{s}$. The variance of the noise samples per quadrature component reads $\sigma_\mathrm{N}^2 = N_0/(2T_\mathrm{s})$ with N_0 denoting the one-sided noise power spectral density. We define a signal-to-noise ratio (SNR) as $E_\mathrm{s}/N_0 = \sigma_\mathrm{A}^2/(2\sigma_\mathrm{N}^2)$.

At the receiver R_M-tuples $\boldsymbol{\Lambda}_\delta$ of pairs of bit metrics $\boldsymbol{\lambda}_\delta^{(\mu)}$, $\mu = 1, \ldots, R_M$, are determined by the bit-metric calculator $\mathcal{L}$ from the y_δ. After deinterleaving the sequence of N_bs pairs of deinterleaved bit metrics $\boldsymbol{\lambda}_\nu$ is processed by the Viterbi decoder (DEC) which finally returns a sequence $\hat{\boldsymbol{q}}$ of K_bs estimates $\hat{q}_\kappa$ on the initial source symbols q_κ.

A. Equivalent Channel Model

Due to the bijection between the R_M-tuples $\boldsymbol{x}$ and the channel symbols[2] a, the combination of $\mathcal{M}$ and AWGN channel can be equivalently represented by a set of R_M parallel subchannels (aka. bit levels) with binary inputs and continuous output, cf. [9] and Fig. 1. The binary labels $\boldsymbol{x}$ are the discrete channel input; the received signal y is the continuous output. The μ-th label bit $x^{(\mu)}$ is transmitted via the μ-th bit level.

B. Bit-Level Capacity and Parallel-Decoding Capacity

The bit levels are characterized by the respective bit-level capacity ${}^{\mathsf{bl}}C$. For the μ-th bit level ${}^{\mathsf{bl}}C^{(\mu)}$ is defined as the mutual information between y and $x^{(\mu)}$ [9], i.e.,

$${}^{\mathsf{bl}}C^{(\mu)} \triangleq I(Y; X^{(\mu)})\,, \tag{3}$$

which for AWGN and fading channels can only be evaluated numerically. Exemplary results for 16-ASK ($R_M = 4$) and 64-ASK ($R_M = 6$) are depicted in Fig. 3. Obviously, there is a

[1] Upper case letters denote the respective random variables to scalars. Vectors and matrices are set in lower resp. upper case boldface letters.

[2] The discrete time index δ is dropped for convenience.

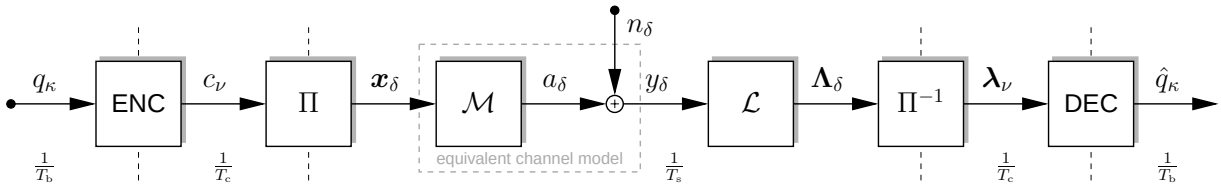

Fig. 1. System model of BICM transmission over AWGN channel.

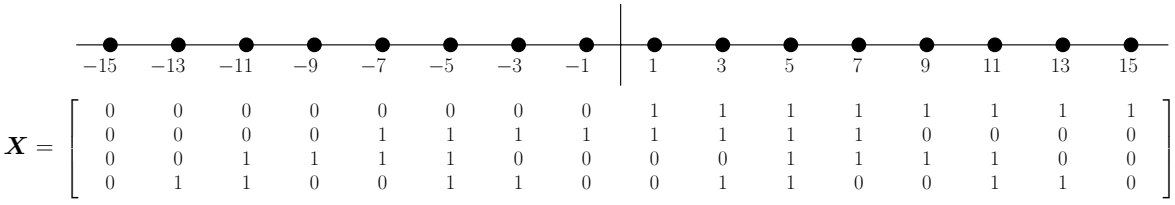

Fig. 2. Illustration of exemplary bit mapping for 16-ASK constellation. $\boldsymbol{X}$ realizes a binary-reflected Gray mapping.

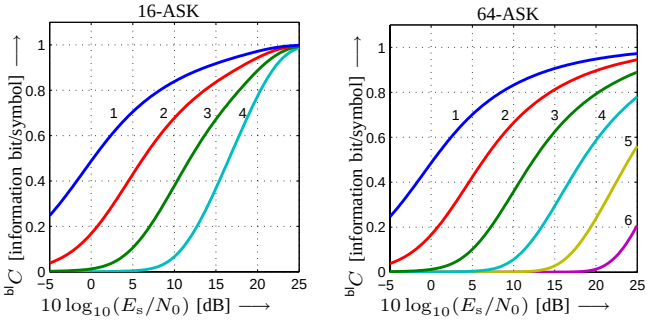

Fig. 3. Exemplary bit-level capacities over $10\log_{10}(E_\mathrm{s}/N_0)$ for AWGN channel. Left: $R_M = 4$ curves for 16-ASK. Right: $R_M = 6$ curves for 64-ASK. Respective level indices μ for binary-reflected Gray mapping given in plots.

strong relation between the $^{\mathsf{bl}}C$'s and μ resp. M. Furthermore, there is also a significant dependency of the average reliability of the bit metrics $\boldsymbol{\lambda}^{(\mu)}$ on the related $^{\mathsf{bl}}C^{(\mu)}$, cf. [7].

The sum over the R_M individual bit-level capacities yields the parallel-decoding capacity (PDC) [9]

$$^{\mathsf{pd}}C = \sum_{\mu=1}^{R_M} {}^{\mathsf{bl}}C^{(\mu)} \,. \tag{4}$$

In contrast to successive or joint decoding, where knowledge of preceeding or even all bit levels is exploited at the decoder, parallel decoding neglects any knowledge originating from other subchannels. The decoding of BICM realizes a parallel-decoding approach, cf. [2], [9], and thus the PDC usually is referred to as 'BICM capacity' in the literature, e.g., [3]. Due to the omitted knowledge, the PDC is lower than the constellation-constraint capacity. The gap between these two variables depends on the bit mapping $\mathcal{M}$. In [6] binary-reflected Gray mappings (BRGM) were shown to minimize the loss in terms of the PDC for medium to high SNRs and thus are used in the remainder. In [1] BRGM were proven to be optimal for uncoded transmission.

C. Bit Mapping

The binary labeling rules $\mathcal{M}$ can be described by using an $(R_M \times M)$-matrix $\boldsymbol{X}$ and a function $\mathrm{ColNr}_{\boldsymbol{X}}\{\boldsymbol{x}\}$. The columns of $\boldsymbol{X}$ contain the M potential binary R_M-tuples, i.e., $\boldsymbol{X} = [\boldsymbol{x}_1 \ldots \boldsymbol{x}_M]$. $\mathrm{ColNr}_{\boldsymbol{X}}\{\boldsymbol{x}\}$ returns the column index (from 1 to M) of $\boldsymbol{x}$ in $\boldsymbol{X}$. We can then write the mapping as

$$\mathcal{M} : \boldsymbol{x} \mapsto a = 2 \cdot \underset{\boldsymbol{X}}{\mathrm{ColNr}}\{\boldsymbol{x}\} - M - 1 \,. \tag{5}$$

Fig. 2 shows an exemplary matrix for 16-ASK and a BRGM.

Obviously, the $^{\mathsf{bl}}C^{(\mu)}$ is entirely determined by the structure of the μ-th row of $\boldsymbol{X}$ denoted by $\boldsymbol{\chi}^{(\mu)}$. Respective $^{\mathsf{bl}}C$'s for the exemplary $\mathcal{M}$ given in Fig. 2 are depicted in Fig. 3. The first row of $\boldsymbol{X}$ leads to the left-most curve and so forth.

The matrix $\boldsymbol{X}$ given in Fig. 2 describes only one of many potential $\mathcal{M}$'s applicable to 16-ASK. Some of these mappings perform entirely identical in terms of, e.g., the BER of uncoded transmission or the PDC. In [1] *trivial operations* on $\boldsymbol{X}$ were defined which do not affect the BER of uncoded tranmission or the PDC. These operations comprise complementations of rows of $\boldsymbol{X}$, interchanging rows in $\boldsymbol{X}$, and reflection of rows wrt. to the symmetry axes of the signal constellation. Note, in the BICM scheme the interchanging of rows of $\boldsymbol{X}$ could be performed implicitly by a respectively designed bit interleaver Π.

III. The Viterbi Decoder and Bit Interleaving

The optimization of BICM for AWGN channels is based on an analysis of the Viterbi decoder and its interaction with the bit interleaver Π. Formally, the latter implements a mapping described as

$$\Pi : \boldsymbol{c} \mapsto [\boldsymbol{x}_1 \ldots \boldsymbol{x}_\Delta] \,. \tag{6}$$

A sequence of N_bs encoded binary symbols c is mapped onto a sequence of Δ binary R_M-tuples $\boldsymbol{x}$.

The inverse operation Π^{-1} at the receiver reads

$$\Pi^{-1} : [\boldsymbol{\Lambda}_1 \ldots \boldsymbol{\Lambda}_\Delta] \mapsto [\boldsymbol{\lambda}_1 \ldots \boldsymbol{\lambda}_{N_\mathrm{bs}}] \,, \tag{7}$$

i.e., a sequence of Δ R_M-tuples $\boldsymbol{\Lambda}_\delta$ of pairs of bit metrics is converted into a sequence of N_bs pairs of bit metrics $\boldsymbol{\lambda}_\nu$.

The problem finally to be solved by the decoder (DEC) is

$$\hat{\boldsymbol{c}} = \underset{\boldsymbol{c}\in\mathcal{C}}{\mathrm{argmin}} \left\{ \sum_{\nu=1}^{N_\mathrm{bs}} \lambda_{\nu,c_\nu} \right\} . \tag{8}$$

The estimated sequence $\hat{\boldsymbol{c}}$ minimizes the overall path metric. Here, $\lambda_{\nu,b}$ denotes the bit metric (Euclidean distance) for the ν-th encoded bit to be $b \in \{0,1\}$ and $\boldsymbol{\lambda}_\nu = [\lambda_{\nu,0}, \lambda_{\nu,1}]^\mathsf{T}$.

The importance of bit interleaving is determined by the 'sliding window' characteristic of the Viterbi decoder. Decoding results strongly rely on localized arrangements of bit metrics within the sequence. In turn, aggregations of unreliable bit metrics most likely lead to errors; the processing window covers only a very small fraction of the trellis. Bit interleavers are mostly designed to compensate for the effects of fading scenarios and to (randomly) spread initially neighbored bit metrics—usually affected by similar fading states and thus equally reliable—over the entire codeword.

In [7] the individual bit levels were identified as an inherent source of 'fading'. The varying bit-level capacities induce varying average bit-metric reliabilities. In contrast to the 'real' fading process, the 'bit-level' fading is known prior to transmission. An approach taking advantage of this knowledge for BICM is, e.g., *adaptive bit interleaving* [7] which leads to significant gains compared to conventional designs.

IV. Optimization of BICM for AWGN Channels

A. Bit Interleaver

For the optimization of BICM for transmission over AWGN channels the usually block-based bit interleaver Π is reduced to simply provide binary R_M-tuples to the mapper $\mathcal{M}$; no 'real' bit interleaving is performed. This approach was already vaguely discussed in some early publications on coded modulation, e.g., [9]. The optimization is entirely shifted to the bit mapping $\mathcal{M}$.

B. Bit Mapping

The optimization of the bit mapping is motivated by the bit interleaver designs presented in [7]. For simplicity, we introduce the idea for rate-$1/2$ convolutional codes; an extension to other code rates is straightforward.

Consider the decoding of a rate-$1/2$ convolutional code as illustrated by the trellis diagram in Fig. 4. In each trellis segment two bit metrics $\lambda_{\nu,b}$ are combined into a segmental path metric. If $R_M = 2$ (4-ASK) is assumed, bit metrics originating from the only two existing bit levels are combined in each segment. The virtual sliding processing window of the decoder comprises several trellis segments and we can easily see that a shift of the window does not affect the average reliability of the bit metrics within its span. Regarding in contrast $R_M = 4$, i.e., 16-ASK transmission and the mapping $\mathcal{M}$ specified by $\boldsymbol{X}$ given in Fig. 2, distinct variations of the average reliability of the segmental path metrics, can be observed, cf. Fig. 4. Segments combining bit metrics of the stronger bit levels (1st/2nd) are succeeded by segments where the bit metrics of the weaker bit levels (3rd/4th) are combined. Shifting the window may affect the average bit-metrics reliability within its span. For larger signal constellations, e.g., 64-ASK, this effect is even more pronounced.

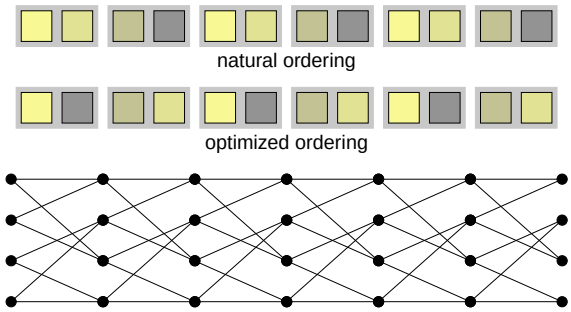

Fig. 4. Illustration of ordering of bit metrics in trellis for rate-$1/2$ code ($\nu_\mathrm{c} = 2$) and 16-ASK transmission ($R_M = 4$). Squares represent bit metrics $\lambda_{\nu,b}$. Brightness grows with average bit metric reliability.

To avoid such disadvantageous arrangements of the bit metrics, we slightly modify $\mathcal{M}$. Consider again 16-ASK transmission using the matrix $\boldsymbol{X}$ given in Fig. 2. By interchanging $\boldsymbol{\chi}_2$ and $\boldsymbol{\chi}_4$ of $\boldsymbol{X}$ we obtain a new matrix $\tilde{\boldsymbol{X}}$, cf. (9), and the average reliability of the segmental path metrics in the trellis is equalized, cf. Fig. 4. In terms of the BER of uncoded transmission and the PDC this modification is a trivial operation, i.e., irrelevant. With regard to the BER of coded transmission, however, $\tilde{\boldsymbol{X}}$ represents a new mapping. For larger signal constellation sizes M the procedure follows the same line: weaker levels are combined with stronger ones.

V. Numerical Results

A. Simulation Settings

Numerical results for the BER of coded transmission are provided for three scenarios: $R_M = 2$, 4, and 6. We used non-recursive, non-systematic convolutional rate-$1/2$ encoders (best known wrt. free distance [10]) with $\nu_\mathrm{c} = 2$, 10, and 13. $\Delta = 1000$ channel symbols were transmitted per block and the initial binary data sequences were zero-padded to ensure terminated trellises. We assessed three different configurations:

(C1) random bit interleaving,
(C2) no interleaving, standard BRGM,
(C3) no interleaving, optimized BRGM acc. to Sec. IV.

The analytically derived BERs of uncoded ASK transmission with equal spectral efficiencies are given for comparison.

B. Results

On the left of Fig. 5 the results of coded 4-ASK transmission and the BER of uncoded 2-ASK transmission are shown. Here, (C2) and (C3) coincide; an optimization is not feasible.

$$\tilde{\boldsymbol{X}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{9}$$

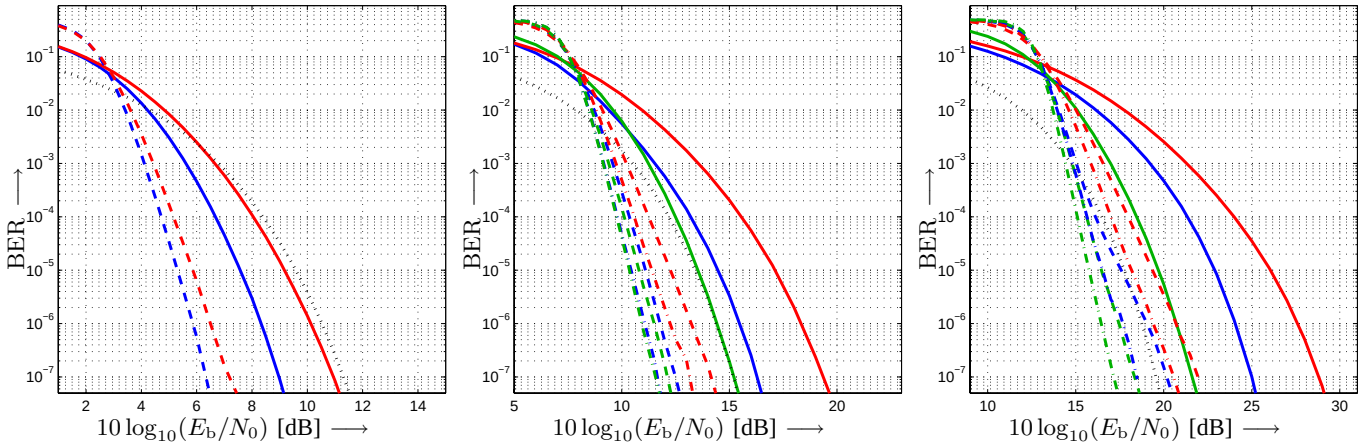

Fig. 5. Bit error ratio of rate-$1/2$ coded 4- (left), 16- (center), and 64-ASK (right) transmission over $10\log_{10}(E_b/N_0)$ [dB]. $\nu_c = 2$ (solid lines), $\nu_c = 10$ (dashed lines), and $\nu_c = 13$ (dash-dotted lines). Random bit interleaving (C1) (red), no bit interleaving (C2) (blue), optimized bit mapping (C3) (green). BER of uncoded 2- (left), 4- (center), and 8-ASK (right) transmission given for comparison (dotted).

Nevertheless, the curves illustrate the advantage of not interleaving on AWGN channels. Especially for $\nu_c = 2$ a careful arrangement of the bit metrics is beneficial. At $\mathrm{BER} = 10^{-6}$ a gain of about $1.8\,\mathrm{dB}$ is achieved at no additional cost compared to the standard scheme. Actually, we can even reduce latency as the block interleaver is removed and such show to advantage a convenient feature convolutional codes exhibit over block-based coding like, e.g., LDPC codes [4].

In the center of Fig. 5 the BERs of coded 16-ASK transmission ($R_M = 4$) and uncoded 4-ASK are given. The plot reveals a significant advantage of (C2) over (C1). Simply dropping the bit interleaver leads to a gain of $3.2\,\mathrm{dB}$ at $\mathrm{BER} = 10^{-6}$ for $\nu_c = 2$. Optimizing the bit mapping (C3)adds another decibel resulting in $4.2\,\mathrm{dB}$ at $\mathrm{BER} = 10^{-6}$. For larger ν_c the gains decrease a little, but are still remarkable. For $\nu_c = 10$ we obtain $1.4\,\mathrm{dB}$ gain of (C2) over (C1) at $\mathrm{BER} = 10^{-6}$. The optimized $\mathcal{M}$ leads to an additional $0.3\,\mathrm{dB}$. The improvements for $\nu_c = 13$ are only slightly smaller.

The results for $R_M = 6$ (64-ASK) are depicted on the right of Fig. 5. Here, the gaps between (C1), (C2), and (C3) are tremendous. For $\nu_c = 2$ the difference between random bit interleaving and no bit interleaving amounts to more than $3.5\,\mathrm{dB}$ at $\mathrm{BER} = 10^{-6}$. Optimizing the bit mapping yields almost another $3.5\,\mathrm{dB}$ adding up to a total of over $7\,\mathrm{dB}$ at $\mathrm{BER} = 10^{-6}$. The weak code with four encoder states achieves a performance similar to the code with $\nu_c = 10$ and random bit interleaving. For $\nu_c = 10$ the gap between no bit interleaving and the optimized bit mapping is $1.5\,\mathrm{dB}$ at $\mathrm{BER} = 10^{-6}$; we gain $3\,\mathrm{dB}$ over random bit interleaving. For $\nu_c = 13$ the optimization yields comparable gains.

C. Discussion

The presented results emphasize the need for a sensible choice of bit mapping and bit interleaver for BICM transmission over AWGN channels. The gains due to the suggested modifications over conventional BICM with convolutional codes are tremendous. This holds in particular for shorter constraint lengths. Stronger codes better mitigate the 'fading' effect of the individual bit levels and the gains due to optimized bit mappings decrease. Nevertheless, the introduced improvements induce no additional complexity and latency is even reduced.

VI. Conclusion

We proposed modifications for BICM transmission over AWGN channels using convolutional codes and Viterbi decoders. Based on a brief analysis of the Viterbi decoder and some recent results on the design of bit interleavers tailored to fading channels, we suggested to abandon bit interleaving on AWGN channels and rearrange the usually employed bit mapping. Such, we achieve a more equalized average bit metric reliability within the relevant processing range of the decoder. The BER of BICM can be significantly lowered at no additional cost and latency is reduced. Consequently, the gap of BICM to TCM in non-fading scenarios is not as large as usually stated in the literature.

References

[1] E. Agrell, J. Lassing, E. G. Ström, and T. Ottosson. On the optimality of the binary reflected Gray code. *IEEE Transactions on Information Theory*, 50(12):3170–3182, 2004.

[2] G. Caire, G. Taricco, and E. Biglieri. Bit-Interleaved Coded Modulation. *IEEE Transactions on Information Theory*, 44(3):927–946, 1998.

[3] A. Guillén i Fàbregas, A. Martinez, and G. Caire. Bit-Interleaved Coded Modulation. *Foundations and Trends in Communications and Information Theory*, 5(1/2):1–153, 2009.

[4] T. Hehn and J. B. Huber. LDPC codes and convolutional codes with equal structural delay: a comparison. *IEEE Transactions on Communications*, 57(6):1683–1692, June 2009.

[5] H. Imai and S. Hirakawa. A new multilevel coding method using error-correcting codes. *IEEE Transactions on Information Theory*, 23(3):371–377, May 1977.

[6] C. Stierstorfer and R. F. H. Fischer. (Gray) mappings for bit-interleaved coded modulation. In *Proceedings IEEE Vehicular Technology Conference Spring (VTC Spring)*, Dublin, Ireland, Apr. 2007.

[7] C. Stierstorfer and R. F. H. Fischer. Adaptive interleaving for bit-interleaved coded modulation. In *Proceedings 7th International ITG Conference on Source and Channel Coding (SCC)*, Ulm, Germany, Jan. 2008.

[8] G. Ungerböck. Channel coding with multilevel/phase signals. *IEEE Transactions on Information Theory*, 28(1):55–67, 1982.

[9] U. Wachsmann, R. F. H. Fischer, and J. B. Huber. Multilevel Codes: Theoretical Concepts and Practical Design Rules. *IEEE Transactions on Information Theory*, 45(5):1361–1391, July 1999.

[10] S. B. Wicker. *Error Control Systems for Digital Communications and Storage*. Prentice-Hall, Upper Saddle River, NJ, USA, 1 edition, 1995.

[11] E. Zehavi. 8-PSK Trellis Codes for a Rayleigh Channel. *IEEE Transactions on Communications*, 40(5):873–884, 1992.