Intralevel Interleaving for BICM in OFDM Scenarios

Clemens Stierstorfer and Robert F.H. Fischer

Institute for Information Transmission, University of Erlangen–Nuremberg

Cauerstraße 7/LIT, 91058 Erlangen, Germany, Email: {clemens,fischer}@LNT.de

Abstract—Interleaver design for bit-interleaved coded modulation in highly bandwidth-efficient transmission is investigated. Based on an equivalent channel model with parallel binary input channels, the variation of the so-called bit level capacities is identified as an additional source of "fading" that has to be taken into account for interleaver design. We assess the more appropriate approach of intralevel interleaving and present numerical results that support the superiority of the suggested design over global interleavers.

I. INTRODUCTION

We consider very high data rate transmission over frequency-selective channels. The occurring intersymbol interference (ISI) is treated by using <u>orthogonal frequencydivision multiplexing</u> (OFDM) [2]. To support the requested high data rates large signal constellations have to be employed in each carrier. Additionally, the introduction of channel coding into the system is inevitable to enhance the reliability. Aiming at low latencies, the codeword size is limited to an OFDM symbol, i.e., a codeword spans all carriers. The channel "experienced" by a codeword, i.e., the different channel conditions of the carriers, can be well modeled by a Rayleigh fading channel.

A well-known coding scheme tailored to fading channels is <u>bit-interleaved coded modulation</u> (BICM)¹ [3]. The employed bit interleaver is crucial for the performance of BICM and its superiority over other coded modulation schemes (e.g., <u>trellis coded modulation</u> (TCM) [10]) in fading environments. In the initial publication on BICM [13], Zehavi used three independent bit interleavers for transmission with 8-PSK (<u>phase shift keying</u>). Regarding a triple of binary symbols mapped onto an 8-PSK symbol, each of the three binary symbols has undergone independent interleaving, so-called *intralevel interleaving*. Caire et al. rejected this approach in [3], arguing that there were no reasons justifying it, flexibility would be limited, analysis complicated, and unequal error protection introduced. Instead they proposed the application of a global interleaver.

In this work we investigate interleavers for BICM using large signal cosntellations which justify the initial attempt in [13] and disprove the superiority of global interleavers in terms of the achievable <u>bit-error ratio</u> (BER). Furthermore, we extend Zehavi's approach and drop the natural ordering of the interleaved bits in the mapping onto channel symbols. The introduction of this additional degree of freedom compared to [13] leads to significant performance improvements. Numerical results reveal gains up to 2 dB for a 64-ASK (amplitude shift keying).

Section II presents the channel and BICM system model. In Section III the decoding and the impact of the interleaver are analyzed. Section IV gives design rules for intralevel interleavers, followed by numerical results in Section V. Section VI concludes the paper.

II. CHANNEL AND BICM SYSTEM MODEL

A. Channel Model

Transmission over a single-input/single-output frequencyselective channel employing OFDM [2] is studied. Considering frequency domain (D (used) carriers), the received signal of carrier d reads

$$y_d = h_d a_d + n_d, \quad d = 1, \dots, D.$$
 (1)

Here, $a_d \in \mathcal{A} \subseteq \mathbb{C}$ is the transmitted symbol drawn from the signal constellation \mathcal{A} ($M = |\mathcal{A}|$, $m = \log_2(M)$, $m \in$ \mathbb{N} , $m \ge 2$), h_d denotes the channel coefficient, and n_d the additive white Gaussian noise (AWGN). The variance of the channel symbols is given as $\sigma_a^2 = E_s/T_s$ (E_s : average energy per symbol; $1/T_s$: symbol rate), that of the noise as $\sigma_n^2 = N_0/T_s$ (N_0 : one-sided noise power spectral density).

The channel description given in (1) indicates a similarity to fading channels. Indeed, it can be shown [7] that for channels of interest the coefficients h_d are complex Gaussian with zero mean and unit variance $(h_d \sim C\mathcal{N}(0, 1))$, i.e., a Rayleigh fading channel is present. Hence, in the following we investigate block-based transmission over a Rayleigh fading channel; codewords shall be restricted to a block, i.e., they comprise D symbols. Since our investigations on interleaver design are based on bit level capacities, the channel coefficients are assumed to be statistically independent. Furthermore, for sake of simplicity we resort to M-ary amplitude shift keying (ASK) constellations ($\mathcal{A} = \mathcal{A}_{ASK} = \{\pm 1, \pm 3, \dots, \pm (M-1)\}$). The translation of the obtained insights to M^2 -QAM constellations is immediate.

B. Equivalent Channel Model

Considering signal constellations with $M \ge 4$, a binary sequence of information bearing symbols $x_d = (x_d^{(1)}, x_d^{(2)}, \ldots, x_d^{(m)})$ has to be mapped onto the channel symbols a_d . In the following we include the mapping \mathcal{M} : $x_d \in \mathbb{F}_2^m \to a_d \in \mathcal{A}$ into the channel model. According to [11], the combination of mapping and channel can be equivalently represented by a set of m parallel subchannels with binary inputs and continuous output. The binary mtuples x_d (aka. binary labels of the a_d), can be identified as the respective inputs; the received signal y_d constitutes

This work was supported by Deutsche Forschungsgemeinschaft (DFG) under grant HU 634/5–2 within the framework TakeOFDM.

¹In the context of this paper and in order to keep the receiver's complexity low, we restrict ourselves to the non-iteratively decoded variant of BICM. The term "bit" is used for binary digits in this context. To denote the binary information unit we employ "information bit".

the continuous output. Apparently, the μ -th label bit $x_d^{(\mu)}$ is transmitted over the μ -th subchannel (aka. μ -th bit level), cf. Fig. 1. We define the capacity of the μ -th subchannel



Fig. 1. Block diagram of equivalent channel model with m binary input channels and scalar continuous output.

in carrier d under the assumption of known channel states h_d and equally distributed input signals, as the mutual information between the receive signal y_d and the binary input signal $x_d^{(\mu)}$ (upper case letters denote the respective random variables)

$$C_d^{(\mu)}(E_s/N_0, h_d) = I(Y_d; X_d^{(\mu)}|h_d, E_s/N_0).$$
(2)

For a fixed state h_d (i.e., AWGN channel), numerical computations reveal rather varying $C_d^{(\mu)}$'s depending on μ , cf. Fig. 2 and [5].



Fig. 2. Exemplary subchannel capacities of equivalent channel model over $10\log_{10}(E_{\rm s}/N_0)$. AWGN channel. Left: m = 4 respective curves for 16-ASK. Right: m = 6 respective curves for 64-ASK.

Referring to [3] for a detailed description of the computation of a lossless bit metric, we state that there is a bijective mapping of the received value y_d onto an *m*-tuple of bit metrics $\Lambda_d = (\lambda_d^{(1)}, \lambda_d^{(2)}, \dots, \lambda_d^{(m)})$. The output of the μ -th level is represented by $\lambda_d^{(\mu)} = [\lambda_{d,0}^{(\mu)}, \lambda_{d,1}^{(\mu)}]^{\mathsf{T}}$, i.e., by a pair of metrics for both hypotheses (transmitted zero/transmitted one). Serializing both, input and output sequence yields a system with a symbol rate of $1/T_c = m/T_s$, cf. Fig. 3. The resulting channel exhibits fading characteristic due to the h_d 's and the subchannel selected for transmission is changed periodically with a period of length *m*.



Fig. 3. Block diagram of serialized channel model with m periodically chosen binary input channels.

C. Bit-Interleaved Coded Modulation

1) System Model: The transmitter of the employed coded modulation scheme is sketched in Fig. 4(a). We use a ratek/n convolutional encoder to encode binary k-tuples $q_{\delta} =$ $(q_{\delta}^{(1)}, q_{\delta}^{(2)}, \dots, q_{\delta}^{(k)})$ of information bits into binary *n*-tuples $c_{\delta} = (c_{\delta}^{(1)}, c_{\delta}^{(2)}, \dots, c_{\delta}^{(n)})$ of encoded bits. Here, the symbol rate of the information bits is denoted by $1/T_{\rm b}$, that of the encoded bits by $1/T_{\rm c}$, and the discrete index of the tuples by $\delta, \delta = 1, \dots, \Delta = \frac{Dm}{n}$. The sequence of encoded tuples is



Fig. 4. Block diagram of employed coded modulation transmitter (a) and detailed diagrams of interleaver concepts. (b): global interleaving. (c): intralevel interleaving, natural ordering of levels acc. to [13]. (d): intralevel interleaving, optimized ordering of levels (two exemplary orderings depicted). Conversion from *n*-tuples c_{δ} to *m*-tuples x_d .

then passed through a bit interleaver Π which also converts the stream of *n*-tuples with index δ into a sequence of *m*tuples $\boldsymbol{x}_d = (x_d^{(1)}, x_d^{(2)}, \dots, x_d^{(m)})$ with index *d*, cf. Fig. 5. These tuples \boldsymbol{x}_d are then mapped onto a sequence of *D* channel symbols a_d which are finally transmitted over the channel as described in Section II-A.



Fig. 5. Exemplary (global) interleaving for 16-ASK with R = 1/2 $(n = 2, m = 4, D = 4, \Delta = 8)$. Conversion of pairs c_{δ} into quadruples \boldsymbol{x}_d . Color of bit levels acc. to Fig. 2.

2) Capacity: Regarding the capacity of BICM, we have to take its parallel decoding character into account, cf. [11]. Contrary to joint decoding of the outputs of m subchannels, any knowledge originating from other subchannels is discarded. As a result the capacity of BICM is slightly inferior to that of joint decoding² and determined by the sum over the level capacities as given in (2). The capacity per symbol a_d of the entire BICM scheme thus reads

$$C(E_{\rm s}/N_0) = \frac{1}{D} \sum_{d=1}^{D} \sum_{\mu=1}^{m} C_d^{(\mu)}(E_{\rm s}/N_0, h_d) \,.$$
(3)

Neglecting the scaling factor 1/D, (3) can be interpreted as the computation of the ergodic capacity of a binary input channel with time-varying channel conditions. In addition

²Here, we use binary reflected Gray mappings to minimize this loss in capacity, cf. [8].

to the fading states h_d , we can recognize the changing capacities of the subchannels as another source of "fading" in the system—the subchannel capacities depicted in Fig. 2 may be interpreted as capacity curves resulting from the same general binary input channel affected by different channel attenuations, i.e., fading states. Noteworthy, interleaving changes the order of summation in (3). Hence, due to the commutativity of summation, the capacity of BICM is not affected by interleaving; only the arrangement of the subchannel capacities within a transmitted block can be influenced.

III. DECODING AND THE IMPACT OF INTERLEAVING

As we focus on the application of non-iteratively decoded BICM the employed decoder uses the Viterbi algorithm [6]. The path metric in the Viterbi algorithm is obtained from the deinterleaved bit metrics $\tilde{\Lambda}_{\delta} = (\tilde{\lambda}_{\delta}^{(1)}, \tilde{\lambda}_{\delta}^{(2)}, \dots, \tilde{\lambda}_{\delta}^{(n)})$ with $\tilde{\lambda}_{\delta}^{(\nu)} = [\tilde{\lambda}_{\delta 0}^{(\nu)}, \tilde{\lambda}_{\delta,1}^{(\nu)}]^{\mathsf{T}}$. The conversion from *m*-tuples Λ to *n*-tuples $\tilde{\Lambda}$ is also performed by the deinterleaver. Denoting the decoder's path metric with λ^{VA} , the total path metric for the hypothesis $(\bar{c}_{1}^{(1)}, \bar{c}_{1}^{(2)}, \dots, \bar{c}_{\Delta}^{(n)})$ is given as

$$\lambda^{\mathrm{VA}}(\bar{c}_{1}^{(1)}, \bar{c}_{1}^{(2)}, \dots, \bar{c}_{\Delta}^{(n)}) = \sum_{\delta=1}^{\Delta} \sum_{\nu=1}^{n} \tilde{\lambda}_{\delta, \bar{c}_{\delta}^{(\nu)}}^{(\nu)}.$$
(4)

The decoder returns an estimate $(\hat{q}_1^{(1)}, \hat{q}_1^{(2)}, \dots, \hat{q}_{\Delta}^{(k)})$ on the transmitted information sequence based on the hypothesis $(\bar{c}_1^{(1)}, \bar{c}_1^{(2)}, \dots, \bar{c}_{\Delta}^{(n)})$ with minimal total path metric λ^{VA} .

Obviously, the metric (4) for an entire codeword is not affected by interleaving. However, the crucial role of the interleavers in BICM schemes for fading channels is wellknown. The great impact of the interleaver on the performance of BICM results from the sliding window characteristic and the immediate decisions of the Viterbi algorithm. According to a rule of thumb, the Viterbi algorithm returns decoding results on symbols after processing approximately five times the constraint length of the convolutional code [6]. The aim of the interleavers is to break up statistical dependencies between neighboring receive bits and to avoid an aggregation of unreliable metrics within the decoding window of the Viterbi algorithm. Ideal interleavers as they are employed in the theoretical analysis of BICM in [3] would completely remove statistical dependencies between any bits and prevent the clustering of unreliable metrics.

Present attempts on the design of implementable interleavers, like, e.g., simple block interleavers (spread adjacently received bits over the entire length of a block) or, more sophisticated, an s-random interleaver (introduce a minimum spacing), e.g., [4], [1], recognize the fading channel as the reason for more or less reliable bit metrics. The intention is to place bit metrics affected by the same fading states as far apart as possible in order to avoid the occurrence of clusters of unreliable bit metrics resulting from deep fades. An inherent conclusion from this goal is that the aggregation of reliable bit metrics in the trellis is not rewarding as well.

IV. (OPTIMIZED) INTRALEVEL INTERLEAVING

According to Section II-C the fading radio channel is not the only source of unreliable bit metrics, but the varying subchannel capacities can be interpreted as another "fading" process. To demonstrate the impact of the level capacities onto the performance of the decoder we first give a counterexample to intralevel interleaving.

A. Counterexample

Consider a system employing 16-ASK, a non-recursive, non-systematic encoder of code rate R = 1/2 and 1024 states [12], and blocks of length D = 1024 with uncorrelated, complex Gaussian channel coefficients h_d . The interleaver may sort the bits according to their levels (cf. Fig. 6) and creates four subsequences each comprising the D bit metrics of a level at the receiver (natural ordering within a subsequence). Thereby, the bit metrics affected by the same fading state h_d are spaced at maximum distance D. The decoder however, sees DR = 512 trellis segments with a path metric originating from a particular level at a time. Fig. 7 shows the resulting BERs over the transmitted block



Fig. 6. Bit level sorting interleaver for 16-ASK with R = 1/2 $(n = 2, m = 4, D = 4, \Delta = 8)$. Color of bit levels acc. to Fig. 2.

(length 2048) of bits. Obviously, the levels have a dramatic impact on the achievable performance. Subchannels with a low capacity lead to a large number of errors, whereas over subchannels with high capacities we can communicate almost without any errors.



Fig. 7. BER over position l within block for bit level sorting (solid lines); Rayleigh fading channel. Blue: $10\log_{10}(E_b/N_0) = 8 \text{ dB}$. Green: $10\log_{10}(E_b/N_0) = 10 \text{ dB}$. Red: $10\log_{10}(E_b/N_0) = 12 \text{ dB}$. Bit level: $\mu = 1$ for $l = 1, \ldots, 512$; $\mu = 2$ for $l = 513, \ldots, 1024$; $\mu = 3$ for $l = 1025, \ldots, 1536$; $\mu = 4$ for $l = 1537, \ldots, 2048$. Level μ acc. to Fig. 2. BER of global random interleaver (dashed) and optimized intralevel interleaving (dotted) for comparison.

B. Design of Intralevel Interleavers

This counterexample uses a particular implementation of a global interleaver, cf. Fig. 4(b). Surely, we selected a very unfavorable version, but the constraints on the spacing of bit metrics affected by correlated fading states were fulfilled. The obtained results suggest an examination of the interleaver design with respect to the bit level capacities. Following Section III, the objective of interleaving is to separate similar or equal fading states. Transfered to the bit levels this implies a design that prevents aggregations of bit metrics in the decoding originating from identical bit levels. Ideally, the interleaver would ensure some kind of minimum spacing between identical or similar levels, i.e., a strategy equivalent to the s-random interleaver has to be found. Though, the limited number of m different level "states" facilitates the design of respective interleavers, there is still a large number of possible concepts. Here, we focus on so-called intralevel interleaving.

1) Conventional Intralevel Interleaving: We drop the currently favored concept of global interleaving (cf. Fig. 4(b)) and return to the initial scheme of BICM as introduced in [13]. There, the idea of intralevel interleaving is employed for transmission with 8-PSK. In order to cope with the fading channel m = 3 independent interleavers are introduced into the three parallel bit streams before the mapping. Hence, a bit is transmitted over the same level as it would be without interleaving, cf. Fig. 4(c).

Regarding the distribution of the levels in the resulting path metric, we can observe a regular structure. The natural ordering of the levels is preserved and hence, the occurrence of clusters of metrics originating from equal subchannels is prevented. Even a spacing at a minimum distance of metrics from equal levels is implemented. For the decoding this interleaving approach provides almost equally distributed levels within the sliding window of the Viterbi algorithm which cannot be guaranteed by global interleaving.

2) Optimized Intralevel Interleaving: For an optimization of the bit level arrangement, we neglect the fading channel and focus on the bit level capacities. Considering the path metrics of isolated trellis segments, it can be noticed that in a single segment n bit metrics are combined into a path metric to obtain k estimates of binary information symbols $\hat{q}_{\delta}^{(\kappa)}$. Apparently, the natural ordering of the bit levels can lead to a periodically changing reliability of the path metric of a trellis segment. Consider again the example of a 16-ASK with a code rate of R = 1/2. Here, a segment's path metric is computed from two bit metrics at a time. Presuming a natural ordering of the bit levels it is either the sum of metrics originating from levels one and two or from levels three and four. As the first two subchannels exhibit significantly better capacities than the last two levels (cf. Fig. 2), the reliability of the respective path metric is most likely larger than that of a path metric based on the two weak levels. The result is a periodically varying reliability of consecutive segmental path metrics.

In our previous considerations on the distribution of fading states and bit levels over a codeword, averaging of "conditions" has been identified as the interleaver's main function. Consequently, an averaged reliability of the metrics of trellis segments could improve performance: the interleaver should combine weak and strong levels for a path metric. In our example of a 16-ASK with a code rate R = 1/2 this could be implemented by combining the bit metrics of level one and four into a path metric and levels two and three in another one, cf. Fig. 9. Compared to conventional intralevel interleaving [13] we introduce an additional degree of freedom which can be used for optimization.



Fig. 9. Exemplary intralevel interleaving for 16-ASK with R = 1/2 ($n = 2, m = 4, D = 4, \Delta = 8$). Color of bit levels acc. to Fig. 2. Ordering of bit levels $(c_{\delta}^{(1)}, c_{\delta}^{(2)}, c_{\delta+1}^{(1)}, c_{\delta+1}^{(2)}) \rightarrow (x_{d_1}^{(1)}, x_{d_2}^{(4)}, x_{d_3}^{(2)}, x_{d_4}^{(3)}), \delta \mod 2 = 1, d_i \in \{1, \dots, D\}, i = 1, 2, 3, 4.$

In some settings, the intralevel interleaver inherently already averages over the level conditions just by the conversion of *n*-tuples c_{δ} on *m*-tuples x_d (e.g., 8-ASK with code rate R = 1/2). Nevertheless, the arrangement of the bit levels is still worth an optimization. The latter might require an exhaustive search over all possible permutations of the bit levels. Though, this can be done offline and the results can be tabulated. The preparation of a comprehensive survey of optimum bit level arrangements is subject of current work. Results of this optimization employed for the numerical examples in Section V can be found in [9].

V. NUMERICAL EXAMPLES

Numerical simulations support our theoretical considerations. We implemented several intralevel interleavers based on the structure depicted in Fig. 4(d). Within each level random interleaving was employed. We show BER curves for several settings, comparing intralevel interleavers to global interleaving. The best arrangement of the bit levels in the intralevel interleaving has been determined by an exhaustive search over all m! possible orderings, cf. [9]. For all settings we employed the best known (wrt. free distance) convolutional codes [12, Tab. 11.3-11.7]³ of code rates R = 1/3, R = 1/2, and R = 2/3 with a non-recursive, non-systematic encoder. Information sequences were zeropadded to ensure terminated trellises; block length was chosen to D = 1024 and 10^5 blocks were simulated.

8-ASK: Here, we compared various codes with encoder states 2^l , l = 2, ..., 10, cf. top of Fig. 8. Apparently, gains of about 0.2 dB can be achieved due to intralevel interleaving for any of the presented settings (for code rates R = 1/3 and R = 2/3 and four encoder states even larger gains occur). Intralevel interleaving can in some cases even halve the number of trellis states compared to global interleaving. Obviously, the advantage of intralevel interleaving grows with decreasing code rate.

16-ASK, 32-ASK, and 64 ASK: On bottom of Fig. 8 the respective BER curves for 16-ASK, 32-ASK, and 64-ASK are depicted. Here, we employed an encoder with $2^{10} = 1024$ states. At BER = 10^{-5} 16-ASK (left) exhibits

³Indices κ and ν are shifted by 1 compared to [12] $(c_{\delta}^{(0)} \rightarrow c_{\delta}^{(1)}$ etc.).



Fig. 8. BER over $10\log_{10}(E_{\rm b}/N_0)$ for several ASK constellations. Rayleigh fading channel. Global random interleaving (blue) and best intralevel interleaving (red). Top: 8-ASK; encoder with 2^{ℓ} states ($\ell = 2, ..., 10$). Left: R = 2/3. Center: R = 1/2. Right: R = 1/3. Bottom: several settings with 1024 states and code rates R = 1/3 (dashed), R = 1/2 (solid), R = 2/3 (dash-dotted). Left: 16-ASK. Center: 32-ASK. Right: 64-ASK.

gains of 0.3 dB for R = 2/3, 0.6 dB for R = 1/2, and 0.6 dB for R = 1/3; for 32-ASK (center) we can see gains of 0.4 dB for R = 2/3, 0.8 dB for R = 1/2, and 1.2 dB for R = 1/3, and 64-ASK (right) shows gains of 1 dB for R = 2/3, 1.5 dB for R = 1/2, and 2 dB for R = 1/3. Again, we can recognize the superiority of intralevel interleaving over global interleaving and the growing gains for decreasing code rate. As we can see from the comparison of all four applied signal constellations, the gains increase with growing constellation sizes, which is a result of the rising variance of the bit level capacities.

VI. CONCLUSION

Interleaving for non-iteratively decoded BICM in combination with large signal constellations has been investigated. Based on an equivalent channel model and the capacities of the employed subchannels, the decoding of BICM and the impact of the interleaver have been studied. We have identified the varying subchannel capacities as another source of fading and suggested an according interleaver design. Numerical results have been presented that support the superiority of intralevel interleaving compared to global interleaving. The obtained gains grow with a decreasing code rate and an increasing size of the applied signal constellation. The advantage of intralevel interleaving comes with no additional cost in complexity or delay compared to global interleaving.

REFERENCES

- A. S. Barbulescu and S. S. Pietrobon. Terminating the trellis of turbo-codes in the same state. *Elect. Let.*, 31:22–23, Jan. 1995.
- [2] J. A. C. Bingham. Multicarrier modulation for data transmission: An idea whose time has come. *IEEE Communications Magazine*, pages 5–14, May 1990.
- [3] G. Caire, G. Taricco, and E. Biglieri. Bit-Interleaved Coded Modulation. *IEEE Transactions on Information Theory*, 44(3):927–946, 1998.
- [4] E. Dunscombe and F. C. Piper. Optimal interleaving scheme for convolutional coding. *Elect. Let.*, 25(22):1517–1518, Oct. 1989.
- [5] R. F. H. Fischer and C. Stierstorfer. Comparison of code design requirements for single- and multicarrier transmission over frequencyselective MIMO channels. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, Adelaide, Australia, Sept. 2005.
- [6] G. D. Forney Jr. The viterbi algorithm. Proceedings of the IEEE, 61(3):268–278, Mar. 1973.
- [7] S. Müller-Weinfurtner. OFDM for Wireless Communications: Nyquist Windowing, Peak-Power Reduction, and Synchronization. PhD thesis, Universität Erlangen-Nürnberg, 2000.
- [8] C. Stierstorfer and R. F. H. Fischer. (Gray) Mappings for Bit-Interleaved Coded Modulation. In Proc. IEEE Vehicular Technology Conference Spring (VTC Spring), Dublin, Irleand, Apr. 2007.
- [9] C. Stierstorfer and R. F. H. Fischer. Optimum Intralevel Interleaving for BICM. http://www.LNT.de/lit/People/CStierstorfer.shtml, 2007. in preparation.
- [10] G. Ungerböck. Channel coding with multilevel/phase signals. *IEEE Transactions on Information Theory*, 28(1):55–67, 1982.
- [11] 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.
- [12] S. B. Wicker. Error Control Systems for Digital Communications and Storage. Prentice–Hall, Inc., Upper Saddle River, NJ, USA, 1 edition, 1995.
- [13] E. Zehavi. 8-PSK Trellis Codes for a Rayleigh Channel. IEEE Transactions on Communications, 40(5):873–884, 1992.