



Rate-adaptive Constellation Shaping for Near-capacity Achieving Turbo Coded BICM

Yankov, Metodi Plamenov; Forchhammer, Søren; Larsen, Knud J.; Christensen, Lars P. B.

Published in:
Proceedings of IEEE ICC 2014

Link to article, DOI:
[10.1109/icc.2014.6883635](https://doi.org/10.1109/icc.2014.6883635)

Publication date:
2014

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Yankov, M. P., Forchhammer, S., Larsen, K. J., & Christensen, L. P. B. (2014). Rate-adaptive Constellation Shaping for Near-capacity Achieving Turbo Coded BICM. In Proceedings of IEEE ICC 2014 (pp. 2112-2117). IEEE. DOI: 10.1109/icc.2014.6883635

DTU Library

Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Rate-adaptive Constellation Shaping for Near-capacity Achieving Turbo Coded BICM

Metodi Yankov, Søren Forchhammer and Knud J. Larsen

Department of Photonics Engineering
Technical University of Denmark
2800 Kgs Lyngby, Denmark
Email: {meya, sofo, knj}@fotonik.dtu.dk

Lars P. B. Christensen
Wireless Algorithm Design
Renesas Mobile Europe
2450 Copenhagen, Denmark

Email: lars.christensen@renesasmobile.com

Abstract—In this paper the problem of constellation shaping is considered. Mapping functions are designed for a many-to-one signal shaping strategy, combined with a turbo coded *Bit-interleaved Coded Modulation* (BICM), based on symmetric Huffman codes with binary reflected Gray-like properties. An algorithm is derived for finding the Huffman code with such properties for a variety of alphabet sizes, and near-capacity performance is achieved for a wide SNR region by dynamically choosing the optimal code rate, constellation size and mapping function based on the operating SNR point and assuming perfect channel quality estimation. Gains of more than 1dB are observed for high SNR compared to conventional turbo coded BICM, and it is shown that the mapping functions designed here significantly outperform current state of the art Turbo-Trellis Coded Modulation and other existing constellation shaping methods.

I. INTRODUCTION

It is well known that at high SNR, uniformly distributed constellation input to a Gaussian channel achieves a rate that is 1.53dB away from channel capacity, as given by Shannon. In order to close that gap, a continuous Gaussian distributed input is required. For discrete modulation used in any practical digital communication system, that gap can be closed for very high order constellations (high number of coded bits per constellation symbol), which have a sampled Gaussian distribution [1]. *Signal shaping* is the process of achieving non-uniform input *Probability Mass Function* (PMF), from the uniformly distributed data. Since this usually comes at the cost of additional redundancy, the channel code rate has to be increased in order to compensate and keep the same data rate, or the constellation size has to be increased to accommodate the extra redundancy bits and retain the high spectral efficiency. These in turn degrade the performance and/or lead to increased complexity at the receiver. However, the energy savings resulting from signal shaping can compensate for the performance degradation. These energy savings at a given data rate are referred to as *Shaping gain*.

In [2] an algorithm is derived for finding the optimal PMF on an AWGN channel for a certain SNR, which was later extended to cover fading scenarios in [3]. The exact optimal PMF is however very hard to achieve in a practical communication system, transmitting coded binary data. In [4], the authors use a shaping code to explicitly select points

with low energy more often. This method results in a PMF with only a few levels, and thus suffers a significant loss in *Mutual Information* (MI) between the input and the output of the channel w.r.t. a continuous Gaussian distribution. In [5] superposition modulation is used together with an optimized irregular convolutional code to combat this problem and achieve a near-capacity performance. However, the decoding complexity grows exponentially with the number of levels. Furthermore, different transmission rates would require different codes, which is impractical in a rate adaptive system, such as e.g. LTE. In [6] the authors approximate the optimal PMF with a *dyadic* one (to be defined in Section II). Dyadic PMFs are attractive because they can be achieved with binary data in a number of ways, e.g. with *Variable Length Codes* (VLC) [6] or a *many-to-one* strategy [7]. In [7] Raphaeli proposed a simple method for achieving a dyadic PMF, using the notion of pragmatic turbo codes [8]. However, the PMF is fixed and is designed for a certain SNR region, making it sub-optimal if a change in the spectral efficiency is required.

In this work a rate-adaptive system is designed, by tuning the transmit PMF depending on the SNR value and the desired spectral efficiency. The desired PMF is achieved by a many-to-one mapping function, and near-capacity performance is reported for a wide SNR region when the mapping function is combined with a turbo coded *Bit Interleaved Coded Modulation* (BICM) [9].

II. SYSTEM OVERVIEW AND MANY-TO-ONE SHAPING

In [10] it is shown how the optimal discrete PMF can be approximated with a dyadic one. A dyadic PMF has the form $p(x_i) = 2^{-l_i}$, where l_i is a positive integer, and x_i is the i 'th symbol of the alphabet. The authors in [10] prove that the dyadic PMF which minimizes the $KLD = D(p_{approx} || p_{optimum})$ is found by *Geometric Huffman Codes* (GHC), where KLD is the Kullback-Leibler Divergence, $p_{optimum}$ is the capacity achieving PMF, and p_{approx} is any dyadic PMF. The MI between the input and the output of a real-valued system (i.e. PAM), when signaling with the optimal input PMF is given in Fig. 1, together with the Shannon capacity, the MI with an uniform input PMF, and the MI with GHC approximation. We see that even though the dyadic PMF cannot achieve capacity, significant gains can still be expected

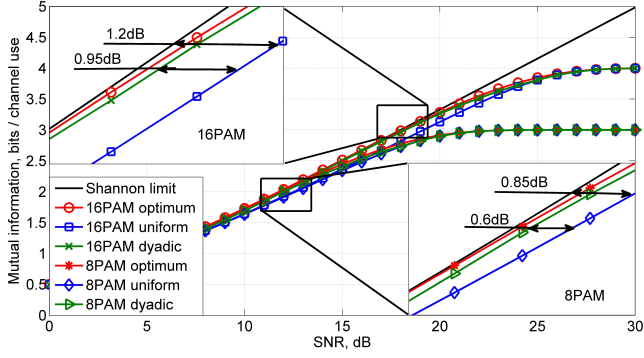


Fig. 1. MI for different input PMFs

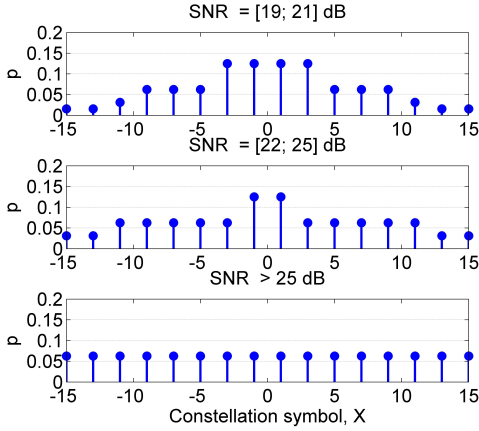


Fig. 2. Dyadic PMFs, approximating the optimal PMF for selected SNR regions

at the examined SNR regions and modulation orders. We also see that after a certain SNR, increase in the modulation size is needed (around 15dB for e.g. 8PAM, or equivalently 64QAM) if Gaussian capacity is targeted. This is due to the limited entropy of the small constellations. Examples of dyadic approximations to the optimal PMF at different SNRs are given in Fig. 2. The PMF for $SNR \in [19; 21]dB$ is the one the authors used in [7]. Higher or lower SNR will have different optimal PMFs and dyadic GHC approximations.

The transmitter from [7], which we also use here, is given in Fig. 3. The data are encoded, and then serial to parallel converted. Puncturing is then applied to the parity streams, so that the remaining parity and data bits can be rearranged into m streams, where the size of the modulation format is 2^m . Each stream is then interleaved, and the signal is modulated. The receiver is given in Fig. 4. The signal is demodulated and the data are deinterleaved. The parity streams are de-punctured and sent for decoding together with the data bits. In case of iterative demapping, the decoder sends extrinsic information back to the demapper, after proper puncturing and interleaving.

Dyadic PMF of the output symbols can be achieved by decoding the binary stream into the constellation symbols using an entropy rate source code, e.g. VLC [11]. The VLC needs to be an entropy code, because otherwise there will exist forbidden sequences in the code, and thus a part of

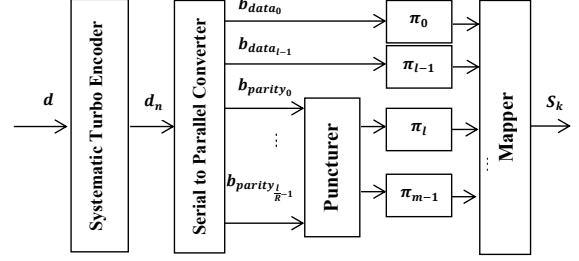


Fig. 3. Turbo coded BICM transmitter

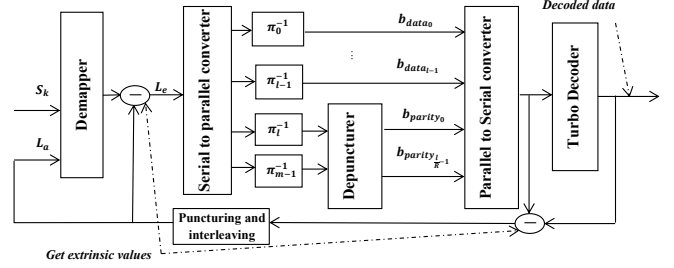


Fig. 4. Turbo coded BICM system receiver

the data will be impossible to decode to any symbol. The probabilities of the output symbols depend on the length of the bit labels assigned to them. For dyadic PMFs, the prefix-free Huffman code is an entropy rate code [11][12]. However, the lack of any structure in the Huffman code results in a severe synchronization problem, especially when there is no error protection before it (as is the case if the VLC is used as an inner code, or in this case - a mapping function).

In [7], the authors tackle this problem by appending ambiguous bits (taking a value of both '1' and a '0') to each bit sequence, after a constellation symbol is chosen. In this way the dyadic PMF is retained, but the label length is constant across symbols. This process can also be viewed as a many-to-one mapping, where the ambiguous bits have the effect of assigning multiple bit sequences with the same prefix (defined by the Huffman code) to the same output symbol. An example of such a mapping function is given in Table I, where the ambiguous bits are labeled as 'X'. For example, constellation symbol '-5' will carry all sequences of length 5, that start with '0010', namely '00100' and '00101'. If the data are i.i.d., the probability of encountering this prefix sequence, and thus this constellation symbol, is clearly 2^{-4} . Compared to

TABLE I
EXAMPLE OF 8PAM MAPPING FUNCTION WITH OPTIMAL DYADIC PMF

Constellation symbol X	-7	-5	-3	-1	1	3	5	7
Bit Label	0	0	0	1	1	0	0	0
	0	0	1	0	1	1	0	0
	1	1	0	X	X	1	0	1
	1	0	X	X	X	X	X	1
	1	X	X	X	X	X	X	0
$p(X)$	2^{-5}	2^{-4}	2^{-3}	2^{-2}	2^{-2}	2^{-3}	2^{-3}	2^{-5}

conventional turbo coded BICM, the only difference in the receiver is the demapper, which needs to take the many-to-one ambiguities into account. Since the complexity of the turbo-decoder dominates over that of the de-mapper, the complexity of the receiver overall is not significantly increased.

A. Mapping design criteria for iterative and non-iterative receivers

In a general BICM, when no iterations are allowed between the de-mapper and decoder, binary-reflected Gray coding is chosen since it provides maximum MI at the output of the demapper with no a-priori knowledge at its input [13]. A popular assumption is that the MI after the de-mapper in that case does not improve with iterative de-mapping (or that estimation of the bit at a certain position in the labeling scheme does not benefit from knowing the rest of the bits), i.e. [14]:

$$I(Y; X) = \sum_m I(Y; b_m | b_1, \dots, b_{m-1}) \approx \sum_m I(Y; b_m) \quad (1)$$

where X is the channel input, Y is the channel output, and b_m is the m 'th bit in the labeling scheme. However, (1) does not hold in the general case. Having a channel code will introduce dependencies between the bits, and (1) therefore has the form:

$$I(Y; X) = \sum_m I(Y; b_m | b_1, \dots, b_{m-1}) \geq \sum_m I(Y; b_m) \quad (2)$$

Where the last inequality turns into equality for binary-reflected Gray mapping *only in perfect conditions*, such as ideal interleaving, perfectly white noise, perfectly white data, optimal de-mapper etc. The usual methods for generating an EXIT chart, e.g. [15], do not take this into account, and the EXIT function of a Gray mapper therefore appears completely flat. In a practical system it will be slightly inclined. In Fig. 5, the bit-wise ($\sum_m I(Y; b_m)$) and symbol-wise ($I(Y; X)$) MI are given for a 16PAM binary-reflected Gray mapping, with the optimal maximum a-posteriori probability de-mapper. We see that at low-to-medium SNR, gains can be expected from iterative de-mapping even with Gray mappings. As the SNR increases and the MI curve starts to level out, this gain disappears. The many-to-one ambiguities will have the effect of further inclining the EXIT function of the de-mapper, resulting in a very good match with the EXIT function of the turbo code, as can be seen from Fig. 6. The shaped system will thus benefit from iterative de-mapping and decoding. Moreover, it is well known that convolutional and turbo codes perform best, when the systematic bits are not punctured. Adding ambiguities on the bit positions can be considered as puncturing, because at the receiver those bits will have equal probability for a '1' and a '0'. We are therefore designing the mapper in a way, so as the unique bit positions resemble a binary-reflected Gray code, and the systematic bits are mapped to them. These criteria also ensure very good performance, when iterations between the decoder and the de-mapper are not allowed.

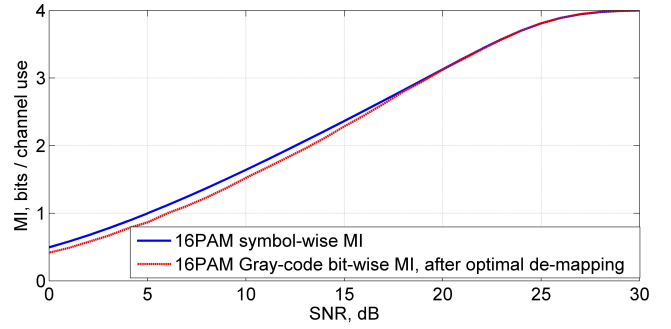


Fig. 5. Bit-wise vs. symbol-wise MI for binary-reflected Gray coded 16 PAM

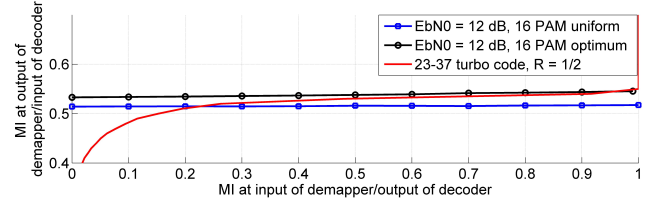


Fig. 6. EXIT chart of 16PAM de-mappers

III. OBTAINING BINARY REFLECTED GRAY-LIKE MANY-TO-ONE MAPPING FUNCTION

In [7], the authors use a single 16PAM mapping function with label length of 6 bits, including the ambiguous bits. As shown in Section I, that PMF of the symbols is optimal only for a limited SNR range (Fig. 2, optimum PMF for $SNR \in [19; 21]dB$). A straight-forward way of obtaining a mapping function for any dyadic PMF is to use a Huffman tree [12]. However, this method will generally produce a mapping function without any structure, and very far from the target binary reflected Gray code. Here we propose a method for designing the mapping function for a 1D constellation, with the additional constraint that the dyadic PMF is symmetric (we note that the GHC approximation from [10] does not necessarily yield a symmetric dyadic PMF, even though the optimal PMF is always symmetric).

A. Obtaining the optimal symmetric dyadic PMF

Let \mathbf{p} be the optimum PMF of the constellation points and the constellation have N points, $\mathbf{p}_{1:N}$ is the PMF of the points from 1 to N , and p_i is the probability of point i . Let \mathbf{p}^{half} be the normalized version of $\mathbf{p}_{1:N/2}$ (we assume even number of constellation points). For symmetric PMFs, it is clear that $\mathbf{p}^{half} = 2\mathbf{p}_{1:N/2}$.

Theorem 1: The symmetric dyadic PMF \mathbf{q} which minimizes $D(\mathbf{q}||\mathbf{p})$ is the one, for which \mathbf{q}^{half} is the GHC approximation of \mathbf{p}^{half}

Proof: $D(\mathbf{q}||\mathbf{p}) = \sum_{i=1}^N q_i \log_2 \frac{q_i}{p_i} = \sum_{i=1}^{N/2} 2q_i \log_2 \frac{2q_i}{2p_i} = D(\mathbf{q}^{half}||\mathbf{p}^{half})$, where the last two equalities follow from the fact that \mathbf{q} and \mathbf{p} are symmetric. As proven in [6], $D(\mathbf{q}^{half}||\mathbf{p}^{half})$ is minimized by GHC, which proves the theorem. ■

The loss in MI from constraining the dyadic PMF to a symmetric one is given in Fig. 7 for a 16PAM constellation. We only see loss in the low SNR region, up to 14 dB. Surprisingly, Raphaeli's PMF [7], even though optimal for around 20dB, performs very well also for lower SNRs. We note that the GHC dyadic approximation from [6], and the symmetric constraint we apply are considered optimal in the sense that they minimize $D(p_{approx}||p_{optimum})$. This optimality does not guarantee that these approximations result in the minimum loss in MI between input and output (unless $D(p_{approx}||p_{optimum}) \rightarrow 0$), which is also what we can see in Fig. 7. Finding such an approximation is currently an NP hard problem [6]. This figure also shows that the PMF from [7] suffers significant loss at high SNR.

The proposed algorithm for finding the binary reflected Gray-like Huffman code is as follows:

- **Consider the PMF of half the constellation points** (e.g. negative points in a symmetric PAM constellation)

- Assign an all-ones string to the point with shortest label (symbol number $N/2$, where N is the number of constellation points)
- **for** $i = N/2 - 1$ **down to** 1 **do**
 - 1) Copy label $i + 1$ to label i
 - 2) For label i , flip the right-most possible bit, making sure the resulting table is prefix free (bit is highlighted in the example in Table II)
 - 3) For label i , set bits $(l_{i+1}; l_i]$ to 1, where l_i is the length of label i (bits are underlined in the example in Table II)

- Once the final label is reached, mirror the table to the other half of the PMF, while flipping the left-most bit, ensuring the labels are prefix free

An example of the obtained Huffman code is given in Table II. The PMF we used there is an example 16PAM dyadic, namely, optimized for $SNR \in [22; 25]dB$ (see also Fig. 2). As discussed above, the mapping function is obtained by appending ambiguous bits to the short labels. Shaped QAM PMFs are obtained as product of PAM PMFs, preserving the Gray-code property.

IV. RESULTS

In this section we present results for the mapping functions designed in Section III, and compare them with the system, originally designed by Raphaeli in [7], and also other near-capacity achieving schemes. The systems are compared in terms of *Bit Error Rate* (BER) and *Block Error Rate* (BLER). The BER analysis is chosen for easier comparison with other systems and capacity. However, the BLER gives a better idea of the achievable throughput in a real communication system over a wide SNR range. Consequently, we report the achieved throughput T in bits per channel use. It is calculated as $T = (1 - BLER)\eta$, where η is the spectral efficiency in bits

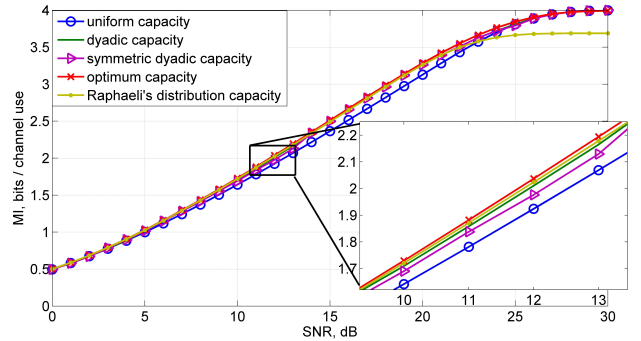


Fig. 7. MI loss from symmetric dyadic approximation

TABLE II
ALGORITHM FOR FINDING THE BINARY REFLECTED GRAY-LIKE HUFFMAN CODE FOR A 16PAM DYADIC PMF

X	i	Bit label						l_i
-15	1	1	0	1	0	<u>0</u>	5	
-13	2	1	0	1	<u>0</u>	<u>1</u>	5	
-11	3	1	0	<u>1</u>	1		4	
-9	4	1	0	0	<u>1</u>		4	
-7	5	1	<u>0</u>	0	0		4	
-5	6	1	1	0	<u>0</u>		4	
-3	7	1	1	<u>0</u>	<u>1</u>		4	
-1	8	1	1	1			3	
1	9	0	1	1			3	
3	10	0	1	0	1		4	
5	11	0	1	0	0		4	
7	12	0	0	0	0		4	
9	13	0	0	0	1		4	
11	14	0	0	1	1		4	
13	15	0	0	1	0	1	5	
15	16	0	0	1	0	0	5	

per channel use $[b/cu]$. Table III summarizes the analyzed modulation schemes. QAM constellations are designed as a product of two identical PAM constellations. Therefore, $l_{max}^{QAM} = 2l_{max}^{PAM}$, where l_{max} is the length of the longest label in the QAM/PAM modulation table.

The system parameters are 5000 information bits block length, 20 turbo iterations, 5 demapping iterations, two convolutional component codes with polynomials (23, 37) in standard octal notation and uniformly distributed puncturing

TABLE III
SUMMARY OF USED MAPPING FUNCTIONS

Constellation	l_{max}	Type
256QAM	12	Raphaeli's [7]
64QAM	8	Designed in Section III
256QAM	10	
1024QAM	14	
1024QAM	16	
64QAM	6	Binary-reflected Gray mappings with uniform PMF
256QAM	8	
1024QAM	10	

pattern. Simulated data rates (modulation spectral efficiency) are between 3 and 9, with a step of $0.5 b/cu$. We note, that not all formats from Table III support all possible rates. This is because the lowest possible code rate of the turbo code in question is $1/3$, and therefore constellation with bit label length of e.g. 16 bits cannot support spectral efficiency of less than $5.5 b/cu$. We simulated 10000 blocks for each code rate at each relevant SNR.

A. AWGN channel

As we saw in Fig. 7, Raphaeli's PMF suffers from a severe loss above a certain SNR. In Fig. 8, we see the performance of the mapping functions at $\eta = 7b/cu$. The best performing 256QAM PMF is the one designed based on the Huffman code from Table II. However, at this point the ambiguities have a significant effect on the waterfall properties of the turbo code. We see that Raphaeli's mapping function (which employs more ambiguities) performs worse than the conventional system with uniform 256QAM, if BER below 10^{-3} is targeted. When we increase the modulation order (equivalently decrease the code rate), the gains are significantly improved to about 0.9 dB. We see that even for these rather short block lengths and very high spectral efficiency, we are very close to uniform capacity (within 0.6dB at $BER \approx 10^{-5}$).

In Fig. 9 the throughput envelope curves are given. The envelope is obtained by taking the throughput of the best-performing combination code-rate/mapping function/modulation size at that SNR operating point. We see that the performance is within 1dB from uniform capacity (0.3dB at $8b/cu$), and that steady shaping gains of 1dB and above are achieved for high spectral efficiency ($\eta > 5b/cu$). Raphaeli's system, even though applicable for a wide SNR range, cannot be used for $\eta > 6.5b/cu$, and at $\eta > 6b/cu$ is already outperformed by the mapping functions designed here. The near-capacity *Turbo Trellis Coded Modulation* (TTCM) from [16] at $\eta = 5b/cu$ is also outperformed by about 1dB (for that system we only report the error-free transmission SNR point, or the SNR, for which the $BER < 10^{-5}$). We note that TTCM is generally limited by the constellation size (namely increasing the constellation size increases the complexity exponentially). For comparison we also plot the performance of the sub-constellation shaping system from [4] at $\eta = 3b/cu$ with block length of 64800 bits, which employs an LDPC code and 32APSK modulation. This system is also significantly outperformed. We also note, that the explicit shaping code in such a system significantly increases both the receiver and transmitter complexity, whereas the mapping functions designed here (as discussed in Section II) only add complexity at the de-mapper, which is not so crucial overall.

B. Rayleigh fading channel

Our preliminary results suggested that fading channels will require stronger codes, and therefore increase in modulation size, if the spectral efficiency is to be kept high. This can also be related to the design criteria of trellis codes. Namely increasing the Euclidian distance is not the main design

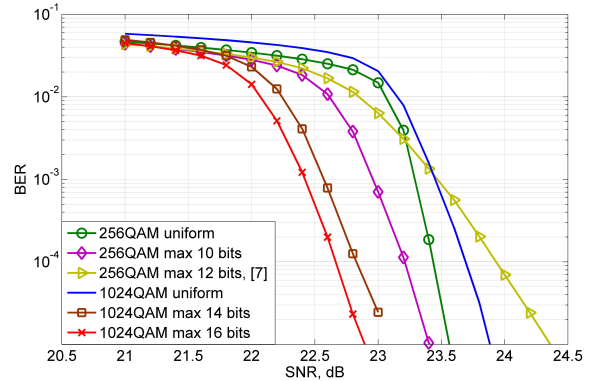


Fig. 8. BER at 7 bits/symbol. Shannon capacity at 21.04dB, 1024QAM uniform capacity at 22.28dB, 256QAM uniform capacity at 22.48dB

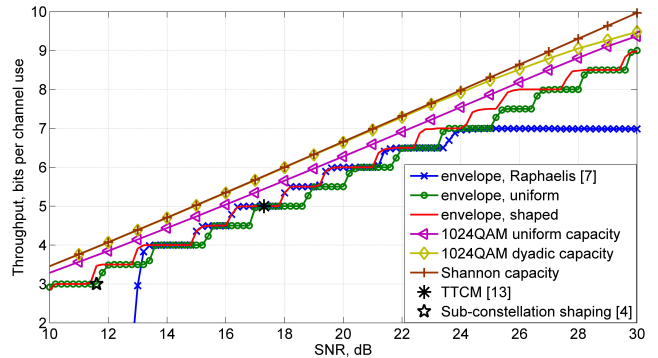


Fig. 9. Throughput envelope, AWGN channel

criterion, but maximizing the product of the metrics on the shortest error-event path and maximizing the length of that path [17]. Decreasing the code rate (reducing the puncturing) directly increases this product for trellis (convolutional) codes. In Fig. 10 the BER performance is given at $\eta = 5b/cu$. We see that already here Raphaeli's system is outperformed. This is due to the required increase in modulation size. The same follows for the TTCM from [18], designed specifically for fading channels, where a 64QAM is used at this spectral efficiency, and the 32 APSK from [4] at $\eta = 3b/cu$. In Fig. 11 a throughput envelope curve is given. We see that our mappings perform close to capacity also on the fading channel (within 1 – 1.5dB, depending on the modulation size and SNR). The gain over the uniform system vanishes at high SNR due to the limited size of the analyzed constellations (in this case - 1024QAM). As seen from Fig. 10, large constellations with low code rate experience an error-floor at around 10^{-4} . This may be partly contributed to the rather short block lengths considered here. We note that the reported performance of both TTCM systems from [13] (for AWGN channel) and [14] (for fading channels) is achieved for much longer codes - 15000 information bits in [16] and 10000η information bits in [18], whereas our system employs 5000 information bits only.

C. Iterative vs. non-iterative de-mapping

As discussed in Section II, even though usually employed in non-iterative receivers, binary-reflected Gray mapping can

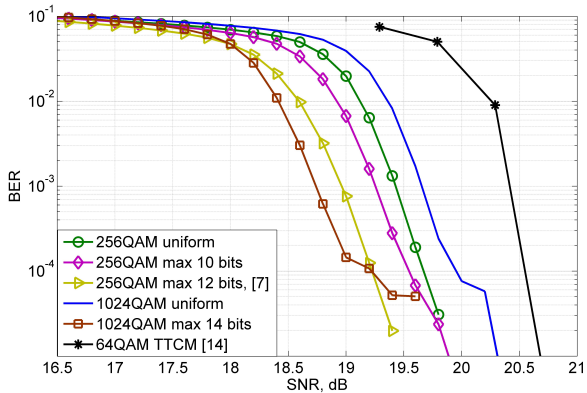


Fig. 10. BER at 5 bits/symbol. Ergodic capacity at 17.15dB, 64QAM uniform capacity at 19.70dB, 256QAM uniform capacity at 18.18dB, 1024QAM uniform capacity at 18.12dB

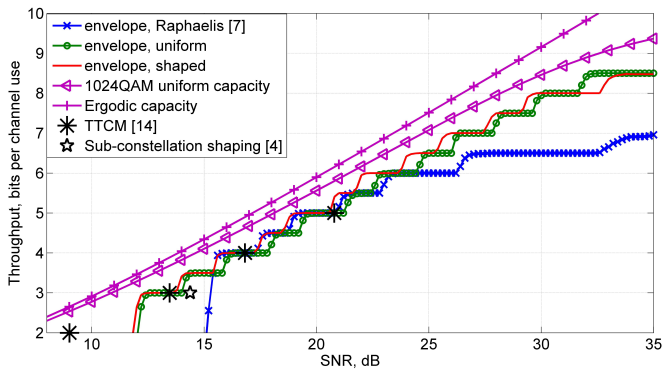


Fig. 11. Throughput envelope, i.i.d. Rayleigh fading channel

still be expected to benefit from iterations when used together with many-to-one shaping and a turbo code. On Fig. 12 we see this effect over a wide range of SNR (i.e. the throughput curves) for AWGN and i.i.d. fading channels. The iterative gain is kept for a wide region of SNR, due to the fact that as the SNR increases, the envelope follows the performance of the higher order modulation formats. i.e. the operating point always falls in the region, where the MI curve appears linear in the log-SNR domain, and the bit-wise MI is smaller than the symbol-wise (see Fig. 5). Towards the very high SNR this gain is not so pronounced because we are approaching the point, where the MI for the largest constellation considered (here 1024QAM), starts to level out.

V. CONCLUSION

In this paper a family of mapping functions were designed for a turbo coded BICM system employing many-to-one constellation shaping. The mappings were designed based on a Huffman code with binary-reflected Gray code properties. Performance of the mappings was analyzed on AWGN and i.i.d. Rayleigh fading channels over a wide range of SNR and spectral efficiencies. It was shown that they outperform current state of the art TTCM, as well as the original many-to-one shaping from [7], further reducing the gap to capacity. It was also shown that iterative de-mapping improves the

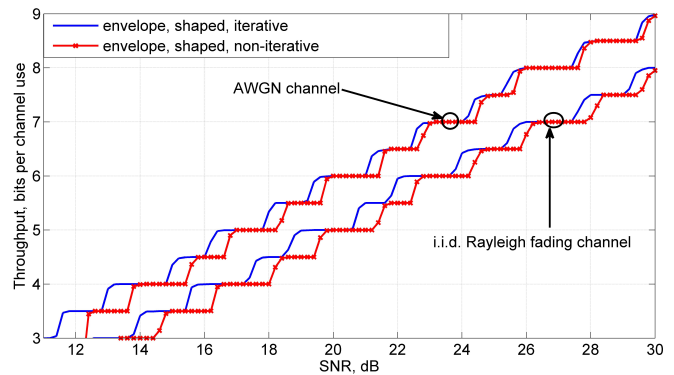


Fig. 12. Illustration of the iterative processing gain

performance of binary-reflected Gray coded mappings, when combined with a turbo code.

REFERENCES

- [1] R. F. H. Fischer, "Precoding and Signal Shaping for Digital Transmission", New York, John Wiley & Sons, 2002
- [2] N. Varnica, X. Ma and A. Kavcic, "Capacity of power constrained memoryless AWGN channels with fixed input constellations," GLOBECOM, pp. 1339-1343 vol.2, 2002.
- [3] J. Bellorado, S. Ghassemzadeh and A. Kavcic, "Approaching the Capacity of the MIMO Rayleigh Flat-Fading Channel with QAM Constellations, Independent across Antennas and Dimensions", *IEEE Trans. Commun.*, vol. 5, no. 6, pp. 1322-1332, 2006
- [4] M. C. Valenti, X. Xiang, "Improving DVB-S2 performance through constellation shaping and iterative demapping," *IEEE Military Communications Conference*, pp. 549-554, 2011.
- [5] M. Noemm, A. Mourad and P. A. Hoeher, "Superposition Modulation with Irregular Convolutional Coding", GLOBECOM, pp. 2346-2350, 2012.
- [6] G. Bocherer, F. Altenbach, and R. Mathar, "Capacity Achieving Modulation for Fixed Constellations with Average Power Constraint", *IEEE International Conference on Communications*, pp. 1-5, 2011.
- [7] D. Raphaeli and A. Gurevitz, "Constellation shaping for pragmatic turbo-coded modulation with high spectral efficiency," *IEEE Trans. Commun.*, vol. 52, no. 3, pp. 341-345, 2004.
- [8] S. Le Goff, A. Glavieux, and C. Berrou, "Turbo-Codes and High Spectral Efficiency Modulation" *IEEE International Conference on Communications*, vol. 2, pp. 645-649, 1994.
- [9] G. Caire, G. Taricco, and E. Biglieri, "Bit-Interleaved Coded Modulation" *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927-946, 1998.
- [10] G. Bocherer and R. Mathar, "Matching Dyadic Distributions to Channels", *Data Compression Conference (DCC)*, pp. 23-32, 2011.
- [11] L. Hanzo, "Near-capacity variable-length coding : regular and exit-chart-aided irregular designs". Chichester, West Sussex, U.K, John Wiley & Sons, 2011
- [12] D. A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes," *Proc. IRE*, vol. 40, no. 9, pp. 1098-1101, 1952.
- [13] C. Stierstorfer and R. F. H. Fischer, "(Gray) Mappings for bit-interleaved coded modulation," *IEEE Vehicular Technology Conference*, pp. 1703-1707, Dublin, Ireland, Apr. 2007.
- [14] F. Schreckenbach, *Iterative Decoding of Bit-Interleaved Coded Modulation*, Ph.D. dissertation, Technical University of Munich, 2007.
- [15] J. Hagenauer, "The EXIT Chart", *Signal Processing - European Conference*, vol. 2, no. 12, pp. 1541-1548, 2004.
- [16] P. Robertson and T. Wörz, "Bandwidth-Efficient Turbo Trellis-Coded Modulation Using Punctured Component Codes", *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 206-218, 1998
- [17] L. Hanzo, T. H. Liew, B. L. Yeap, R. Y. S. Tee and S. X. Ng "Turbo Coding, Turbo Equalisation and SpaceTime Coding". Chichester, West Sussex, U.K, John Wiley & Sons, 2011
- [18] S. X. Ng, O. R. Alamri, Y. Li, J. Klierer and L. Hanzo, "Near-Capacity Turbo Trellis Coded Modulation Design Based on EXIT Charts and Union Bounds," *IEEE Trans. Commun.*, vol. 56, no. 12, pp. 2030-2039, 2008.