Iteratively Decodable Codes for Bridging the Shaping Gap in Communication Channels

Nedeljko Varnica, Xiao Ma and Aleksandar Kavčić*
Division of Engineering and Applied Sciences, Harvard University
33 Oxford Street, Cambridge, MA 02138
e-mail: [varnica,xiaoma,kavcic]@hrl.harvard.edu

Abstract

We consider the power-constrained complex memoryless additive white Gaussian noise channel whose channel inputs are drawn from a finite alphabet. It is well known that if the probability mass function over the finite alphabet is uniform, a shaping gap is created that asymptotically approaches 1.53dB as the constellation cardinality approaches infinity. In a recent paper, we proposed a method to compute the shaping gap for a finite alphabet size and finite SNR. Here, we take advantage of constellations that can be represented as cross-products of the in-phase (real) and quadrature (imaginary) une dimensional constellations (e.g., a 16-QAM constellation). For a 256-QAM constellation, we construct separate simple in-phase and quadrature inner trellis codes whose combined information rate bridges (and nearly closes) the shaping gap. We then demonstrate that a judiciously constructed outer iteratively decodable low-density parity-check code performs inside the shaping gap, that is, very near the channel capacity.

1 Introduction

When a uniform input distribution is used over a finite-size alphabet at the input of a memoryless additive white Gaussian noise (AWGN) channel, an information loss known as the shaping gap incurs [1]. For quadrature amplitude modulation (QAM) input alphabets as the alphabet cardinality goes to infinity, this shaping gap asymptotically approaches 1.53dB at high signal-to-noise ratios (SNRs) [1]. In a recent paper [2], we developed a procedure to evaluate the shaping gap of a finite alphabet at a finite SNR. The same can be achieved by using the method proposed in [3], but with much higher (and often prohibitive) computational cost [2].

In this paper, we propose a code over a finite input alphabet, that bridges the shaping gap and approaches the capacity of the channel. The design strategy is to decompose the code into an inner trellis code and an outer iteratively decodable low-density parity-check (LDPC) code [4]. The inner trellis code is constructed as a *matched information rate* code [5], whose name comes from the property that the symbol occurrence probabilities match the information maximizing (i.e., capacity achieving) source distribution.

A simplification in the code design is achieved by using an input alphabet that can be decomposed as a cross-product of an in-phase (real) alphabet and a quadrature (imaginary) alphabet. This allows us to separate the trellis code into two independent (real and imaginary) trellis codes without loss of optimality. This separability into real and imaginary trellis codes is conceptually different from previously proposed trellis code constructions, both without constellation shaping [6] and with constellation shaping [7], [8]. The difference is a result of the fact we are using an *information rate* criterion to construct the trellis code, whereas prior work constructed trellis codes using an algebraic *minimum Euclidean distance* criterion. In our view, if the objective is to achieve the capacity, the construction that uses the information rate criterion is more appropriate.

2 The channel capacity

Let t denote discrete time, and let the channel input and output (complex) random variables be denoted by X_t and Y_t, respectively. The complex memoryless AWGN channel is given by

$$Y_t = X_t + W_t, \qquad (1)$$

where W_t is a circularly symmetric white Gaussian noise process with per-dimension variance σ^2. Let P be the channel input power constraint, i.e., $\mathrm{E}\left(|X_t|^2\right) \le P$. The capacity of the channel in (1) is [9]

$$C = \log_2\left(1 + \frac{P}{2\sigma^2}\right) = \log_2\left(1 + SNR\right), \qquad (2)$$

* This research was supported in part by NSF Grant No. CCR-0118701

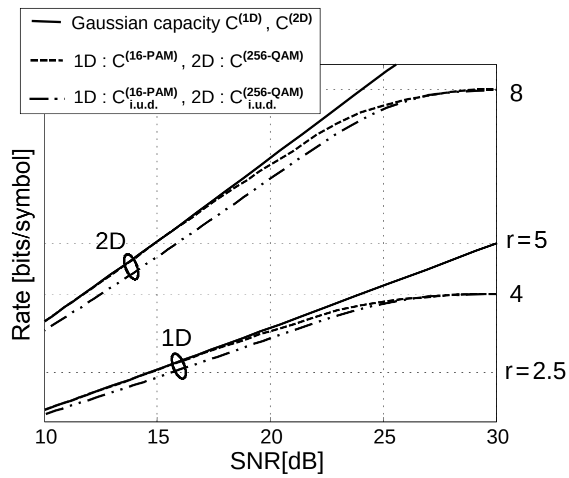

Figure 1. Capacity curves. $C^{(\mathrm{1D})}$ and $C^{(\mathrm{2D})}$ are the capacities of the 1-dimensional and 2-dimensional AWGN channels. $C^{(16-\mathrm{PAM})}$ and $C^{(16-\mathrm{PAM})}_{\mathrm{i.u.d}}$ are the capacity and i.u.d. capacity of a 16-PAM alphabet over an AWGN channel. $C^{(256-\mathrm{QAM})}$ and $C^{(256-\mathrm{QAM})}_{\mathrm{i.u.d}}$ are the capacity and i.u.d. capacity of a 256-QAM alphabet over an AWGN channel.

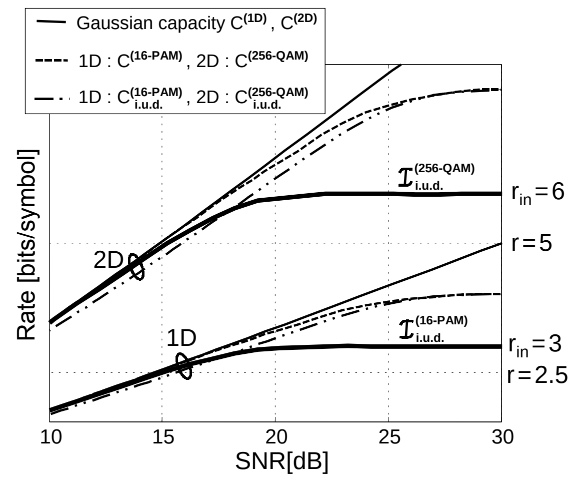

Figure 2. The i.u.d. superchannel information rate $\mathcal{I}^{(16-\mathrm{PAM})}_{\mathrm{i.u.d.}}$ of an inner matched information rate trellis code. By Theorem 1, by combining identical in-phase and quadrature trellis codes, we get a 256-QAM trellis code whose i.u.d. supercahnnel information rate is $\mathcal{I}^{(256-\mathrm{QAM})}_{\mathrm{i.u.d.}} = 2\mathcal{I}^{(16-\mathrm{PAM})}_{\mathrm{i.u.d.}}$.

and is achieved by a circularly symmetric white Gaussian input process with power $\mathrm{E}\left(|X_t|^2\right) = P$.

We now constrain the channel input random variable X_t to be drawn from a finite-size alphabet $\mathcal{X}$, where the cardinality of $\mathcal{X}$ is $|\mathcal{X}| < \infty$, and consider the channel

$$Y_t = \alpha \cdot X_t + W_t, \tag{3}$$

where α is a real constant chosen such that $\mathrm{E}\left(\alpha^2 |X_t|^2\right) = P$. For the channel with a finite input alphabet in (3), we may utilize the numeric procedure proposed in [2] to numerically evaluate the channel capacity under a power constraint P.

In this paper, we consider constellation alphabets that can be decomposed into a cross-product of identical in-phase (real) and quadrature (imaginary) constellation alphabets

$$\mathcal{X} = \mathcal{A} \times \mathcal{A}. \tag{4}$$

For example, if $\mathcal{A}$ is the 4-PAM (pulse amplitude modulated) input alphabet $\mathcal{A} = \{-3, -1, 1, 3\}$, then $\mathcal{X} = \mathcal{A} \times \mathcal{A}$ is the 16-QAM input alphabet.

Theorem 1: For the AWGN channel in (3), if the input alphabet is separable, i.e., $\mathcal{X} = \mathcal{A} \times \mathcal{A}$, and the power constraint is $\mathrm{E}\left(\alpha^2 |X_t|^2\right) \leq P$, the capacity is achieved by a memoryless random process whose in-phase (real) and quadrature (imaginary) components are independent and identically distributed over the alphabet $\mathcal{A}$, each with power $P/2$.

Armed with Theorem 1, we shall construct a (nearly) optimal trellis code for the 16-PAM input alphabet $\mathcal{A}^{(16-\mathrm{PAM})}$. The cross product of two such trellis codes is automatically a (nearly) optimal trellis code for the 256-QAM alphabet $\mathcal{X}^{(256-\mathrm{QAM})} = \mathcal{A}^{(16-\mathrm{PAM})} \times \mathcal{A}^{(16-\mathrm{PAM})}$.

3 Matched information rate inner trellis code

Figure 1 shows the capacities and information rates for both the one dimensional (16-PAM) and two dimensional (256-QAM) constellations, computed using the algorithm in [2]. Note that by Theorem 1, the two-dimensional information rates and capacities are twice those of the one-dimensional counterparts.

Our goal is to use a 256-QAM alphabet to communicate over the channel with a rate, say, $r = 5$bits/256-QAM-symbol. By Theorem 1, this is equivalent to using the 16-PAM alphabet to communicate over a one-dimensional (i.e., real) AWGN channel with a rate $r = 2.5$bits/16-PAM-symbol. Indeed, from Figure 1, we observe that $C^{(16-\mathrm{PAM})} = 2.5$bits/16-PAM-symbol

$\alpha = 0.19 \quad SNR = 14.93dB$	
16-PAM symbol	Symbol probability
± 1	0.145
± 3	0.126
± 5	0.097
± 7	0.064
± 9	0.037
± 11	0.019
± 13	0.008
± 15	0.004

TABLE I
Optimized transition probabilities and α for 16-PAM inputs to an AWGN channel for rate $r = 2.5$bits/16-PAM-symbol.

is achieved at SNR 14.93dB, and that $C^{(256-\mathrm{QAM})} =$ 5bits/256-QAM-symbol is achieved also at the same SNR 14.93dB. Table I gives the probability mass function over the 16-PAM alphabet that achieves the capacity $C^{(16-\mathrm{PAM})} = 2.5$bits/16-PAM-symbol, as computed by the algorithm in [2].

We shall construct a concatenated inner and outer code with rate $r = r_{\mathrm{in}} \cdot r_{\mathrm{out}} = 2.5$bits/16-PAM-symbol. The main feature of the inner *matched information rate* trellis code is that the symbol occurrence probabilities match those given in Table I. Rules for constructing matched information rate trellis codes for binary intersymbol interference channels have been presented in [5] (longer version is under review [10]). These rules readily apply to memoryless AWGN channels. Using these rules, we constructed a 40-state trellis code with rate $r_{\mathrm{in}} = 3$bits/16-PAM-symbol. (The trellis of this code is too complex to give here.) The constructed matched information rate trellis code has the feature that for every $k = 3$ binary input digits, an output symbol from the 16-PAM alphabet is released.

We define the superchannel as the concatenation of the inner trellis code and the AWGN channel. The i.u.d. superchannel information rate is the information rate achieved through the superchannel when the superchannel inputs (i.e., inner trellis code inputs) are independent and uniformly distributed (i.u.d.), and can be computed by the method in [11], [12]. The i.u.d. superchannel information rate (denoted by $\mathcal{I}_{\mathrm{i.u.d.}}^{16-\mathrm{PAM}}$) for the constructed 40-state matched information rate code is given in Figure 2. The importance of the i.u.d. superchannel information rate is underlined by the fact that $\mathcal{I}_{\mathrm{i.u.d.}}^{16-\mathrm{PAM}}$ can be achieved by an outer random linear block code, see Theorems 1 and 2 in [10]. Here, we approach $\mathcal{I}_{\mathrm{i.u.d.}}^{16-\mathrm{PAM}}$ using an outer iteratively decodable LDPC code.

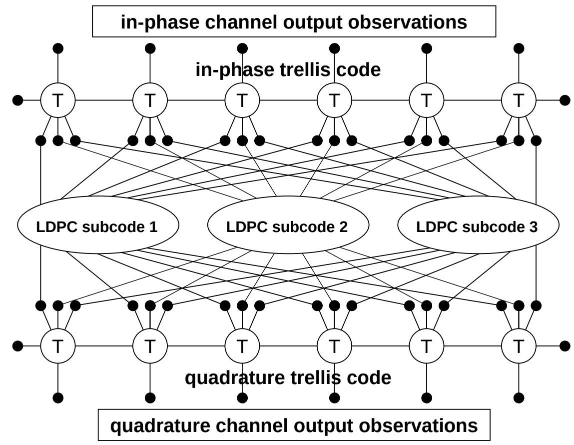

Figure 3. Joint graph of the in-phase (real) and quadrature (imaginary) inner matched information rate trellis codes, together with the outer LDPC code, where the LDPC code is separated into $k = 3$ subcodes.

4 Outer LDPC code

Our goal is to construct an outer linear code whose rate is $r_{\mathrm{out}} = r/r_{\mathrm{in}}$, where r is the overall target code rate, and r_{in} is the rate of the inner trellis code. For example, in Section 3, we had $r = 2.5$bits/16-PAM-symbol, and $r_{\mathrm{in}} = 3$bits/16-PAM-symbol. The resulting outer code rate is $r_{\mathrm{out}} = 2.5/3$. Since each stage of the inner trellis code converts a block of $k = 3$ binary symbols into a 16-PAM symbol, the outer code consists of $k = 3$ separate subcodes.

The i-th subcode encodes the i-th symbol of the block of $k = 3$ binary symbols. The rate $r^{(i)}$ of the i-th subcode is determined as the information rate between the binary digits of the i-th subsequence and the channel output sequence, under the assumption that all the symbols of subcodes 1 through $i-1$ are known to the receiver, and that all symbols are i.u.d. This information rate can be computed by appropriately modifying the Monte Carlo method for computing information rates on a trellis given in [11], [12].

For the purpose of designing the code (especially for designing the degree sequences of the constituent irregular LDPC subcodes), we may assume that the decoder employs the multistage decoding strategy [13]. The multistage strategy assumes that the subcodes 1 through i–1 are decoded without errors, and that these decoded symbols are available to the decoder of subcode i.

The multistage strategy is assumed only in the design process. The actual decoding is performed by message passing [14] on a joint graph of the inner trel-

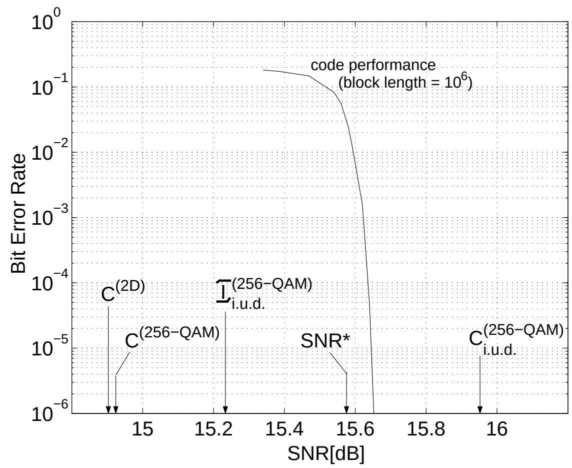

Figure 4. Simulation of an outer LDPC code when the inner code is the matched information rate trellis code.

lis code an the outer LDPC code. This joint graph is depicted in Figure 3 for the example of $k = 3$ outer subcodes, where the inner code maps a (k=3)-tuple into a 16-PAM symbol. To plot the graph in Figure 3, we have taken advantage of symmetry between the in-phase (real) and quadrature (imaginary) components of the 256-QAM constellation, such that the trellis code portion of the graph has separate in-phase and quadrature trellis codes (which is justified by Theorem 1).

5 Code example and simulation results

We revert to the example of a one-dimensional 16-PAM alphabet (or equivalently a 256-QAM alphabet) at rate $r = 2.5$bits/16-PAM-symbol with an inner code whose i.u.d. superchannel rate is given in Figure 2. For this example, we constructed an outer LDPC code with $k = 3$ constituent subcodes. Subcodes 1 and 2 are irregular LDPC codes whose degree coefficients are given in Table II (see [15], [16] for definitions of irregular LDPC codes and their degree coefficients). We use a regular LDPC code as Subcode 3 since regular codes perform very near the capacities at very high code rates [17]. In Figure 4, we show the performance of a finite-length LDPC code whose degree coefficients are given in Table II. Notice that the code performs inside the shaping gap (i.e., between $C^{(256-\mathrm{QAM})}_{\mathrm{i.u.d.}}$ and $C^{(256-\mathrm{QAM})}$). Also notice that the performance is only 0.74dB away from the AWGN capacity $C^{(2D)}$.

Subcode 1			Subcode 2			Subcode 3		
$r^{(1)} = 0.555$			$r^{(2)} = 0.955$			$r^{(3)} = 0.990$		
x	$\lambda_x^{(1)}$	$\rho_x^{(1)}$	x	$\lambda_x^{(2)}$	$\rho_x^{(2)}$	x	$\lambda_x^{(3)}$	$\rho_x^{(3)}$
2	0.2079		2	0.1844		3	1.0	
3	0.1328		3	0.3206		300		1.0
4	0.2054		7	0.1498				
5	0.0581	0.3943	10	0.0204				
6		0.0136	20	0.0231				
15		0.0041	35		0.0052			
16	0.1627		37	0.2214				
17	0.0029		49	0.0803				
26		0.0041	60		0.3215			
30		0.5839	74		0.0109			
41	0.0215		132		0.0005			
42	0.2087		140		0.6619			
threshold 1 $SNR_1^* = 15.51$dB			threshold 2 $SNR_2^* = 15.58$dB			threshold 3 $SNR_3^* = 15.56$dB		
$SNR^* = \max\left(SNR_1^*, SNR_2^*, SNR_3^*\right) = 15.58$ dB								

TABLE II
Degree sequences of constituent LDPC subcodes and their noise tolerance thresholds.

6 Conclusion

We have constructed an iteratively decodable code for bridging the shaping gap of AWGN channels. The code is a concatenated code with an inner matched information rate trellis code. The inner code was constructed over the 256-QAM symbol alphabet, which factors into a cross-product of two 16-PAM alphabets. We showed that by constructing two separate 16-PAM trellis codes (and combining them into a 256-QAM code) we loose nothing in terms of the achievability of the maximal information rate, but we gain in terms of simplicity of the code design. The outer code was constructed as an LDPC code, where the decoding was conducted by message-passing on a joint inner/outer code graph. We demonstrated that with this coding strategy, we were able to enter the shaping gap and approach the channel capacity of the AWGN channel to within a fraction of a decibel.

References

[1] G. D. Forney, Jr. and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Transactions on Information Theory*, vol. 44, pp. 2384–2415, October 1998.

[2] N. Varnica, X. Ma, and A. Kavčić, "Capacity of power constrained memoryless AWGN channels with fixed input constellations." To appear in Globecom2002, November 2002.

[3] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. London: Academic Press, 1981.

[4] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1962.

[5] X. Ma, N. Varnica, and A. Kavčić, "Matched information rate codes for binary ISI channels," in *International Symposium on Information Theory*, (Lausanne, Switzerland), 2002.

[6] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, pp. 55–67, January 1982.

[7] L. Wei, "Trellis-coded modulation with multidimensional

constellations," *IEEE Transactions on Information Theory*, vol. 33, pp. 483–501, July 1987.

[8] G. D. Forney, Jr. and L. Wei, "Multidimensional constellations-Part I:Introduction, figures of merit, and generalized cross constellations," *J. Select. Areas Commun.*, vol. SAC-7, pp. 877–892, August 1989.

[9] C. E. Shannon, "A mathematical theory of communications," *Bell Systems Technical Journal*, vol. 27, pp. 379–423 (part I) and 623–656 (part II), 1948.

[10] A. Kavčić, X. Ma, and N. Varnica, "Matched information rate codes for partial response channels." submitted for publication in *IEEE Trans. Inform. Theory*, June 2002.

[11] D. Arnold and H.-A. Loeliger, "On the information rate of binary-input channels with memory," in *Proceedings IEEE International Conference on Communications 2001*, (Helsinki, Finland), June 2001.

[12] H. D. Pfister, J. B. Soriaga, and P. H. Siegel, "On the achievable information rates of finite state ISI channels," in *Proceedings IEEE Global communications Conference 2001*, (San Antonio, Texas), pp. 2992–2996, November 2001.

[13] U. Wachsmann, R. F. Fischer, and J. B. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Transactions on Information Theory*, vol. 45, pp. 1361–1391, July 1999.

[14] N. Wiberg, H.-A. Loeliger, and R. Kötter, "Codes and iterative decoding on general graphs," *European Trans. on Commun.*, vol. 6, pp. 513–526, September 1995.

[15] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Improved low density parity check codes using irregular graphs and belief propagation," in *Proc. IEEE Int. Symp. Inform. Theory*, (Cambridge, MA), p. 117, Aug. 1998.

[16] T. Richardson and R. Urbanke, "The capacity of low-density parity check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 599–618, February 2001.

[17] A. Kavčić, X. Ma, and M. Mitzenmacher, "Binary intersymbol interference channels: Gallager codes, density evolution and code performance bounds." 44 pages, accepted for publication after revisions in *IEEE Trans. Inform. Theory*; available at http://hrl.harvard.edu/~kavcic/publications.html, February 2001.