Approaching the Dirty Paper Limit for Canceling Known Interference

Uri Erez
EECS, MIT
uri@allegro.mit.edu

Stephan ten Brink
Realtek, Irvine, CA
stenbrink@realtek-us.com

Presented at
41th Ann. Allerton Conf. on Commun., Control, and Computing
Oct. 1–3, 2003

Abstract

Costa's "writing on dirty paper"–channel model offers an information theoretic framework for precoding techniques for canceling arbitrary known interference. Using lattice strategies and MMSE scaling, lossless precoding is theoretically possible at any SNR. Following this approach, we report an end-to-end coding realization of a system materializing a significant portion of the promised gains. We employ iterative detection and decoding of capacity approaching codes, where code design is done using the EXIT chart technique.

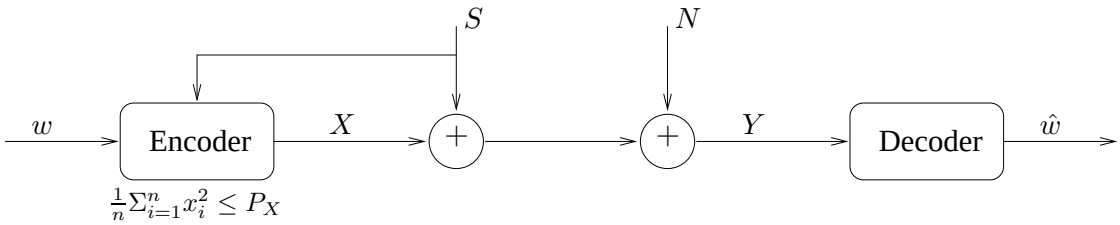

Figure 1: The generalized Costa channel

1 Introduction

It has recently been shown [10] that an information theoretic framework for the study of efficient known interference cancellation (precoding) techniques may be found in Costa's "Writing on dirty paper" [8]. The (generalized) dirty paper channel model is depicted in Fig. 1.

$$Y = X + S + N, \tag{1}$$

where S is arbitrary interference known at the transmitter (noncausally) and N is a statistically independent Gaussian random variable with variance P_N, and where the encoder has power P_X. This channel model was proposed by Cover with Gaussian S and N and in [8]. Costa showed that in this case the capacity is equal to $\frac{1}{2}\log(1 + P_X/P_N)$. Therefore the noise S does not incur any loss in capacity. We treat the generalized model where S can be an arbitrary signal, deterministic or random, for which this result holds as well [10].

Willems suggested schemes for coding for the dirty paper channel (for causally known interference) in [16]. In [10] it was shown that the full capacity may be achieved using a scheme based on lattices and MMSE scaling. Related schemes were developed in the context of information embedding in [7, 9]. In [15] a realization of the necessary lattice transmission scheme based on trellis shaping [12, 11] and "syndrome dilution" was proposed. In this work we extend this approach by employing capacity approaching codes and using iterative detection and decoding. We design a complete end-to-end dirty paper transmission system which attains a significant portion of the promised gains. The system incorporates a check–biregular, repeat–irregular nonsystematic repeat–accumulate (RA) code [14, 3] concatenated with a trellis shaping code. The variable node decoder (VND) of the RA code is designed for iterative quantization detection and decoding using the EXIT chart technique. We note that, in principle, a low–density parity–check code could be used in a similar set–up as well. However, RA codes exhibit a nice linear encoding complexity.

1.1 Lattice precoding

We review the lattice precoding approach proposed in [10]. Let Λ denote an n-dimensional lattice and let $\mathcal{V}$ denote its fundamental Voronoi region. Also let $\mathbf{U} \sim \text{Unif}(\mathcal{V})$, that is $\mathbf{U}$ is a random variable (dither) uniformly distributed over $\mathcal{V}$. The scheme is given by,

- *Transmitter:* The input alphabet is restricted to $\mathcal{V}$. For any $\mathbf{v} \in \mathcal{V}$, the encoder sends:

$$\mathbf{x} = [\mathbf{v} - \alpha\mathbf{s} - \mathbf{u}] \bmod \Lambda. \tag{2}$$

- *Receiver:* The receiver computes

$$\mathbf{y}' = [\alpha\mathbf{y} + \mathbf{u}] \mod \Lambda. \tag{3}$$

The resulting channel is a mod-Λ additive noise channel described by the following lemma:

Lemma 1 ([10]) *The mod-Λ channel defined by (1),(2) and (3) satisfies*

$$\mathbf{Y}' = \mathbf{v} + \mathbf{N}' \mod \Lambda \tag{4}$$

with

$$\mathbf{N}' = (1-\alpha)\mathbf{U} + \alpha\mathbf{N} \mod \Lambda. \tag{5}$$

The mutual information of the channel is maximized by a uniform input, giving

$$\frac{1}{n}I(\mathbf{V};\mathbf{Y}) = \frac{1}{n}h(\mathbf{Y}') - \frac{1}{n}h(\mathbf{N}') = \frac{1}{2}\log\frac{P_X}{G(\Lambda)} - \frac{1}{n}h(\mathbf{N}'), \tag{6}$$

where $G(\Lambda) = \frac{1}{n}\frac{\int_{\mathcal{V}}\|\mathbf{x}\|^2 d\mathbf{x}}{|V|^{1+2/n}}$ is the normalized second moment of Λ. Taking $\alpha = \frac{P_X}{P_X+P_N}$ we have

$$\mathrm{Var}((1-\alpha)\mathbf{U} + \alpha\mathbf{N}) = (1-\alpha)^2\frac{1}{n}\mathrm{Var}(\mathbf{U}) + \alpha^2\frac{1}{n}\mathrm{Var}(\mathbf{N}) = \alpha P_N. \tag{7}$$

Since for a given second moment a Gaussian random vector has the greatest entropy it follows that

$$\frac{1}{n}h(\mathbf{N}') \le \frac{1}{n}h((1-\alpha)\mathbf{U} + \alpha\mathbf{N}) \le \log\left(2\pi e\alpha P_N\right) \tag{8}$$

where the first inequality follows since the modulo operation can only decrease the entropy. Substituting (8) in (6) we obtain the following lower bound on the achievable rate as a function of $G(\Lambda)$,

$$I(\mathbf{V};\mathbf{Y}) \ge \frac{1}{2}\log(1+SNR) - \frac{1}{2}\log 2\pi e G(\Lambda). \tag{9}$$

Thus, in principle, for a given lattice Λ, the gap to capacity of a precoding system may be made smaller than $\log 2\pi e G(\Lambda)$. This is depicted in Fig. 2 from which the gap in dB may be inferred. For optimal lattices for quantization we have $G(\Lambda) \to 2\pi e$, and the gap goes to zero.

Note that when Λ is one-dimensional, the lattice precoding scheme is based simply on scalar quantization (SQ) and is an extension of Tomlinson-Harashima precoding. For this case, the achievable mutual information of the mod-Λ channel (4) may be easily computed and is depicted in the Fig. 2. Note that while the gap to capacity of a scalar system is 1.53dB at high SNR, the lowest possible E_b/N_0–operating point is at 2.4dB. This means that the gap to capacity approaches 4dB at zero spectral efficiency (see Fig 2). For this reason we concentrate our efforts on the low SNR regime.

In the next section we describe coding for the scalar case. This will serve as a baseline reference for our main results, reported in Section 3. See also [6].

2 One–Dimensional Quantization

We briefly describe a one-dimensional lattice transmission system. The effective noise channel (4) now takes the form

$$Y' = X + \underbrace{(1-\alpha)\,U}_{\substack{\text{uniform in interval}\\ [-A(1-\alpha),A(1-\alpha)]}} + \underbrace{\alpha N}_{\substack{\text{Gaussian}\\ N\left(0,\alpha^2 P_N\right)}} \mod [-A, A].$$

Coding for this channel is essentially not much different than for an AWGN channel. As our focus is on the low SNR regime we chose our target operating spectral efficiency (passband) to be 1bit/s/Hz. A SQ transmission system is depicted in Fig. 3.

We use an off–the–shelf parallel concatenated (turbo) code (PCC, [2]). Fig. 4 depicts a typical effective noise channel and the respective log–likelihood ratio values (L–values [13]).

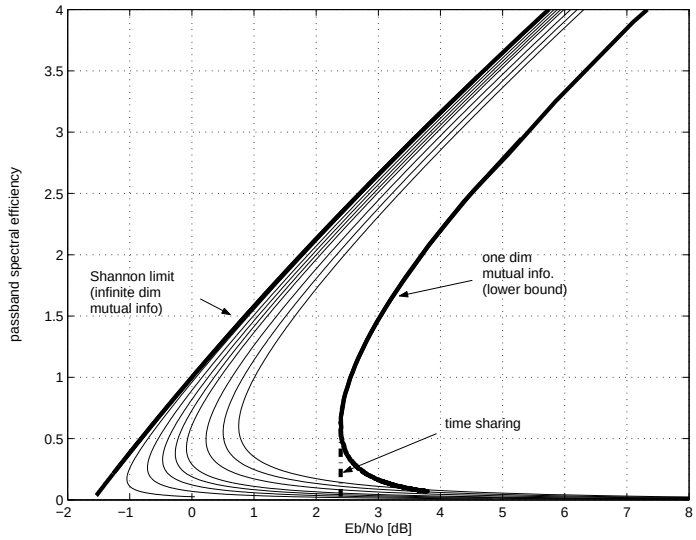

Figure 2: Lower bound (9) on achievable rates as a function of $G(\Lambda)$. From left to right: AWGN capacity limit; lower bound for shaping gains $1.5, 1.45, 1.4, 1.35, 1.3, 1.2, 1.1$ and 1.0dB; mutual information of one-dimensional scheme.

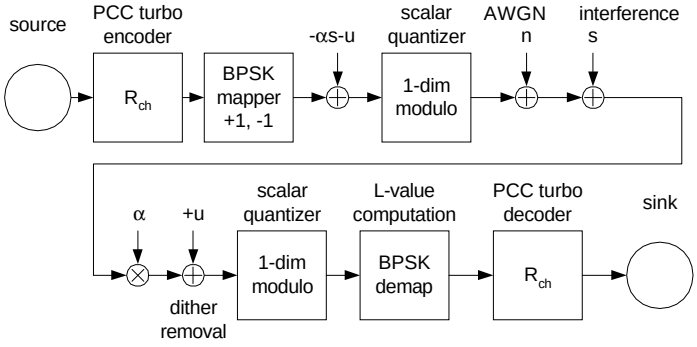

Figure 3: Dirty paper coding with turbo codes and a scalar quantizer (SQ).

The L–values are the input to the turbo decoder. For a spectral efficiency of 1bit/s/Hz we use a PCC of rate $R_{ch} = 1/2$ and BPSK modulation per dimension. The code is of memory 4, and has generator polynomials 037_8 (feedback) and 021_8 (feedforward). With $\alpha = 0.65$, we obtain a turbo cliff at about 3dB (length $K = 10^5$ systematic bits, 20 iterations) which is just about 0.4dB from the performance predicted by the mutual information limits of the scalar quantizer, and 3dB away from the AWGN capacity limit (see Fig. 9). Similarly, for a spectral efficiency of 0.667bit/s/Hz, we use a PCC of rate $R_{ch} = 1/3$, memory 4, and polynomials 025_8 (feedback), 037_8 (feedforward). Setting $\alpha = 0.55$, we get the turbo cliff at about 2.8dB which is 0.4dB from the mutual information limit, and 3.4dB away from the respective AWGN capacity limit.

3 Multidimensional Quantization

3.1 Background: Obtaining lattices from linear codes

The modulo operations performed at both transmission ends mean that we may equivalently view a message selection as specifying a *coset* $\mathbf{t} + \Lambda$. The actual transmitted signal is the difference between $\alpha\mathbf{S} + \mathbf{U}$ and the nearest point of the coset. As S is arbitrary (unbounded) this

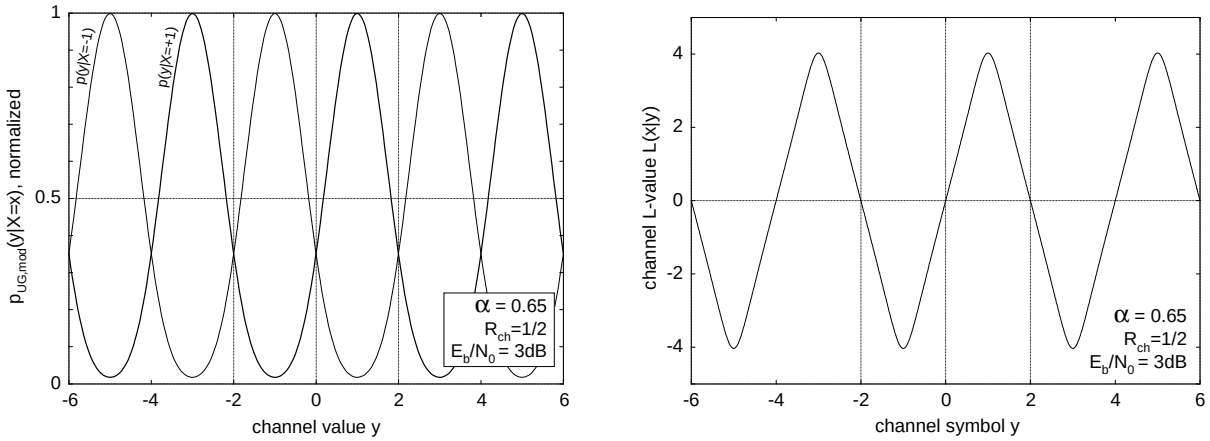

Figure 4: Left: Conditional probability density functions (convolution of Gaussian and uniform densities) after scalar quantizer (modulo). Right: Corresponding log–likelihood ratio values (L–values); modulo interval from $-2, ..., +2$; $\alpha = 0.65$; $E_b/N_0 = 3$dB at code rate $R_{ch} = 1/2$.

means that we have to search over the infinite lattice. Similarly, while we may closely approximate the "unfolded" effective noise $(1-\alpha)\mathbf{U}+\alpha\mathbf{N}$ as Gaussian noise of variance $\alpha P_N = \frac{P_X P_N}{P_X+P_N}$, the modulo operation (folding) at the receiver means that we have to compute the metrics $\sum_\lambda \exp\left\{-\|\mathbf{y}-\mathbf{v}+\lambda\|^2/\alpha P_N\right\}$. Again, this involves a summation over the infinite lattice. Fortunately, a standard method for constructing lattices from linear codes, i.e., Construction A, yields lattices that are also periodic in the cubic lattice $p\mathbb{Z}^n$ (p prime). Furthermore, lattices which are optimal for shaping, i.e., having $\log 2\pi e G(\Lambda)$ as small as desired, may be obtained having this structure (although one would have to use non–binary codes). This reduces the search to that of first performing *one* dimensional quantization, and then performing a search over the finite set of coset representatives of the quotient group $\Lambda/p\mathbb{Z}^n$. The separation of the search into these two stages is done in trellis precoding in [11] and in the context of dirty paper coding in [15]. Fig. 5 illustrates the construction by example. To visualize this construction the example is over the prime field $\mathbb{Z}_{11}$.

Example:

We take block length $n = 2$ and field $\mathbb{Z}_{11}$. We use a rate 1/2 block code ($k = 1$) given by the generating matrix (vector) $G = [2, 3]$ so that

$$\mathcal{C} = \{x \cdot [2, 3] \mod 11 \ : \ x \in \mathbb{Z}_{11}\}$$

We embed the code "as is" in Euclidean space as depicted in Fig. 5 (left). Then using this "finite lattice" we tessellate the whole of $\mathbb{R}^2$ giving the lattice

$$\Lambda = \mathcal{C} + 11\mathbb{Z}^2$$

The eleven points contained in the fundamental Voronoi region serve as coset representatives and correspond to the choice of the vector $\mathbf{v}$ in Section 1. The next step is to choose a code, i.e., use a subset of these coset leaders and map the messages onto them.

3.2 Designed system

Our system follows this line of construction, replacing lattices with trellis codes as is often done in practice. The system model is depicted in Fig. 6.

Transmitter: The transmitter is a concatenation of a nonsystematic repeat–accumulate (RA) code [3], performing the "coset dilution", and a trellis shaping code (i.e. the vector quantizer). The RA encoder is composed of an outer mixture of repetition codes of different rates (variable nodes), an edge interleaver, and an inner mixture of single parity check codes

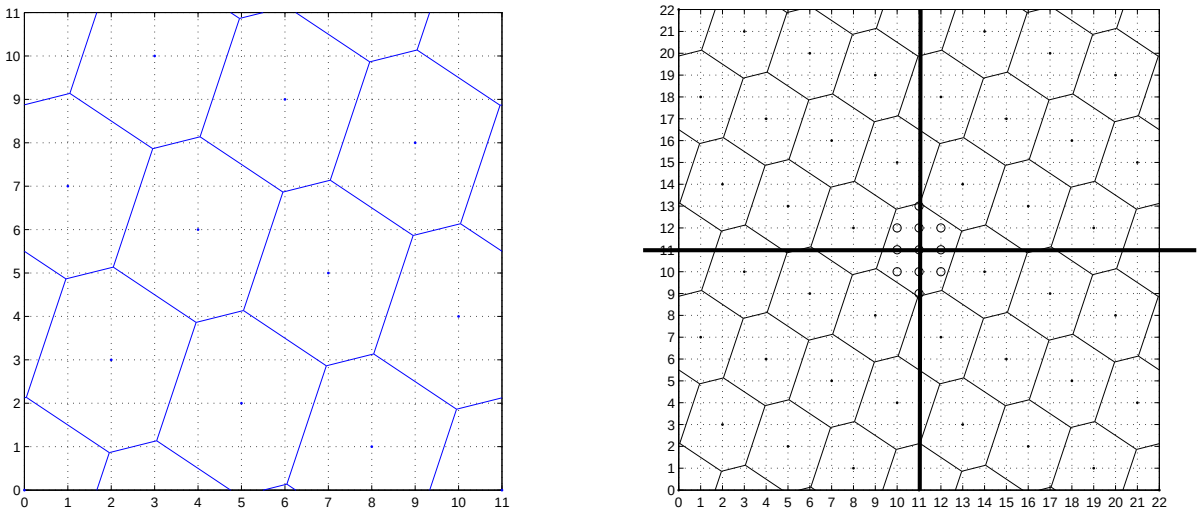

Figure 5: Construction A: Linear code is embedded in $\mathbb{R}^n$ (left), then space is tessellated (right).

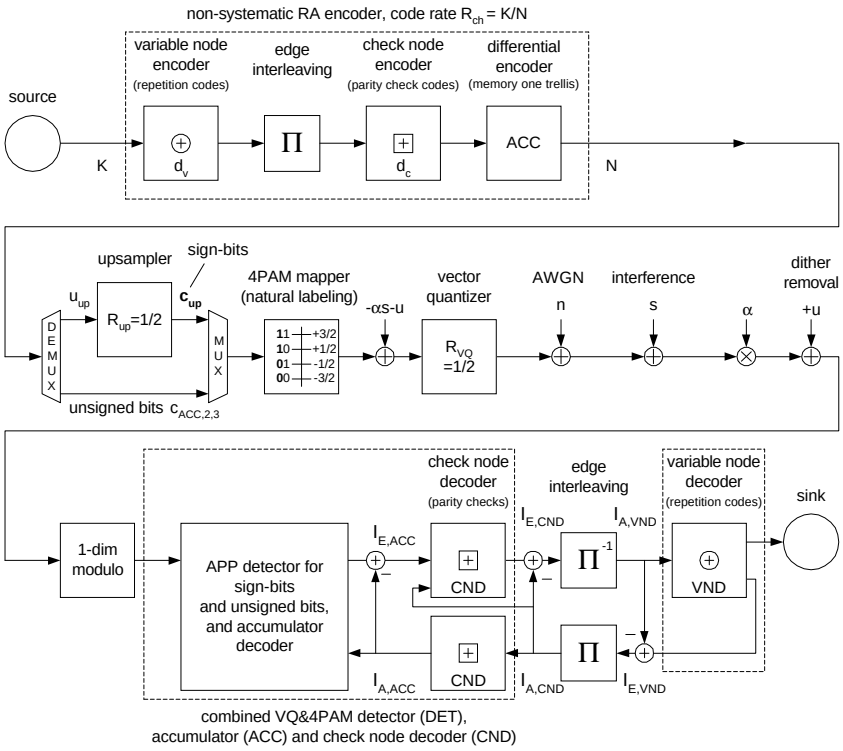

Figure 6: Dirty paper coding with nonsystematic repeat-accumulate codes and a vector quantizer (VQ); iterative quantization and decoding.

of different rates (check nodes), followed by a memory one differential encoder (accumulator, ACC). Code design is performed by appropriately choosing repetition and check node degree distributions. The encoded bits are grouped into triplets $(c_1, c_2, c_3)_{ACC}$ and demultiplexed into "upsampler" bits $u_{up} = c_{ACC,1}$ and unsigned bits $c_{ACC,2}, c_{ACC,3}$. The upsampler (replacing the inverse syndrome former in trellis shaping) has rate $R_{up} = 1 - R_{VQ} = 1/2$. The sign–bits $c_{up,1}, c_{up,2}$ generated by the upsampler, and the unsigned bits are mapped onto 4-PAM symbols using natural labeling. After adding the scaled interference and a uniformly distributed dither signal, the vector quantizer determines the minimum energy sequence, and the quantization error vector is transmitted over the communication channel.

Channel: On the channel, white Gaussian noise is added, with double–sided noise power spectral density $P_N = N_0/2$ and mean zero. Interference is added. For 16-QAM (4-PAM per dim.) and $R_{VQ} = 1/2$, we have $E_s/N_0 = 2(1 + 0.5)R_{ch}E_b/N_0$. Thus, for simulation we set $P_N = E_s/(3R_{ch}2E_b/N_0)$, whereby E_s is the average energy per *complex* output symbol measured *after* the VQ at the transmitter.

Receiver: At the receiver, MMSE α–scaling is applied, and the dither signal u is removed; a one-dimensional modulo is performed prior to putting the signal into a soft in/soft out vector quantizer which performs an *a posteriori* probability (APP) detection of the sign–bits and the unsigned bits respectively, using the BCJR algorithm [1] on an appropriately defined trellis structure. The vector quantizer, thus, can be viewed as an APP detector, computing extrinsic information on the sign–bits and unsigned bits respectively, which is forwarded to the RA decoder. The RA decoder is composed of an inner accumulator decoder (ACC), check node decoder (CND), and an outer variable node decoder (VND), which, in turn, provides *a priori* information for the APP VQ detector to improve the quantization result ("iterative quantization and RA decoding"). The structure is close to the scheme presented in [3]. As we merge the APP vector quantizer with the inner accumulator decoder of the RA code, we obtain a variant of "trellis detection", similar to [4].

Fig. 7 aids in understanding the structure of the joint trellis processing over accumulator trellis (memory $\nu_{ACC} = 1$), vector quantizer trellis (memory ν_{VQ}), upsampler, and modulo symbol metric based on two 4-PAM symbols per three hypothesized accumulator bits $(u_1, u_2, u_2)_{ACC}$. One *trellis column* comprises $2^{\nu_{ACC}} = 2$ states of the memory one accumulator, and $2^{\nu_{VQ}}$ vector quantizer states per accumulator state, i.e. in total $2^{\nu_{ACC}+\nu_{VQ}}$ states. One *state transition* is labeled by the three *input* bits to the accumulator $(u_1, u_2, u_3)_{ACC}$, the virtual input bit to the VQ u_{VQ}, and by the two 4-PAM *output* symbols s_1, s_2. The intermediate outputs are the hypothesized coded bits of the accumulator $(c_1, c_2, c_3)_{ACC}$, the output of the rate 1/2 upsampler $c_{up,1}, c_{up,2}$, the hypothesized coded bits of the VQ $c_{VQ,1}, c_{VQ,2}$, and the sign–bits $c_{s,1} = c_{up,1} + c_{VQ,1}|_{\mathrm{mod}\,2}, c_{s,2} = c_{up,2} + c_{VQ,2}|_{\mathrm{mod}\,2}$. Thus, with inputs $(u_1, u_2, u_3)_{ACC}$ and u_{VQ}, there are $2^4 = 16$ state transitions entering and leaving each state of the trellis. *A priori* information is provided by the outer variable node decoder on the inner information bits with respect to the accumulator, i.e., on $(u_1, u_2, u_3)_{ACC}$.

Note that the information bit u_{VQ} of the vector quantizer is "virtual": By keeping u_{VQ} undetermined ("floating"), all VQ–codewords are allowed. Of course, since the VQ at the transmitter has taken the liberty to change the sign–bits to its liking (according to its codebook), namely, to find/shape the minimum energy sequence, all VQ–codewords are equally likely and have to be "overlayed" in the trellis structure to perform appropriate detection of the sign–bits and unsigned bits respectively. This corresponds to the summation over the coset specified in Section 3.1.

3.3 Mutual information

With the chain–rule of mutual information [5], we can compute the mutual information of an "equivalent bit channel", i.e. the channel that, effectively, is experienced by the channel decoder after VQ APP detection. For this, we compute the mutual information transfer curve of the

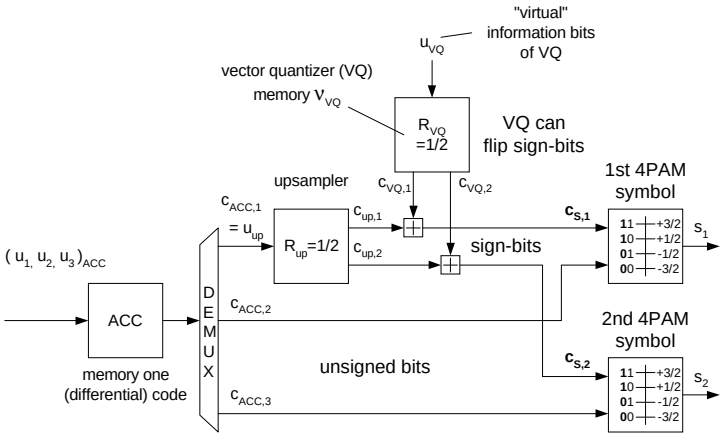

Figure 7: Illustration of joint accumulator, upsampler, vector quantizer and 4–PAM trellis processing.

VQ APP detector using *a priori* knowledge that is modeled as stemming from a binary erasure channel (BEC). Examples of such curves are depicted in Fig. 8 (left) for different VQ memory; observe that the S–shape is more pronounced for bigger memory. An integration over the area under these curves gives an estimate of the mutual information that is available to the channel decoder, provided that perfect iterative decoding over detector and decoder could be performed. In the figure, the area under the memory 4 and 6 VQ curves is a little greater (for fixed E_b/N_0) if compared to the memory 2 VQ, since the shaping gains for memory 4 and 6 are bigger.

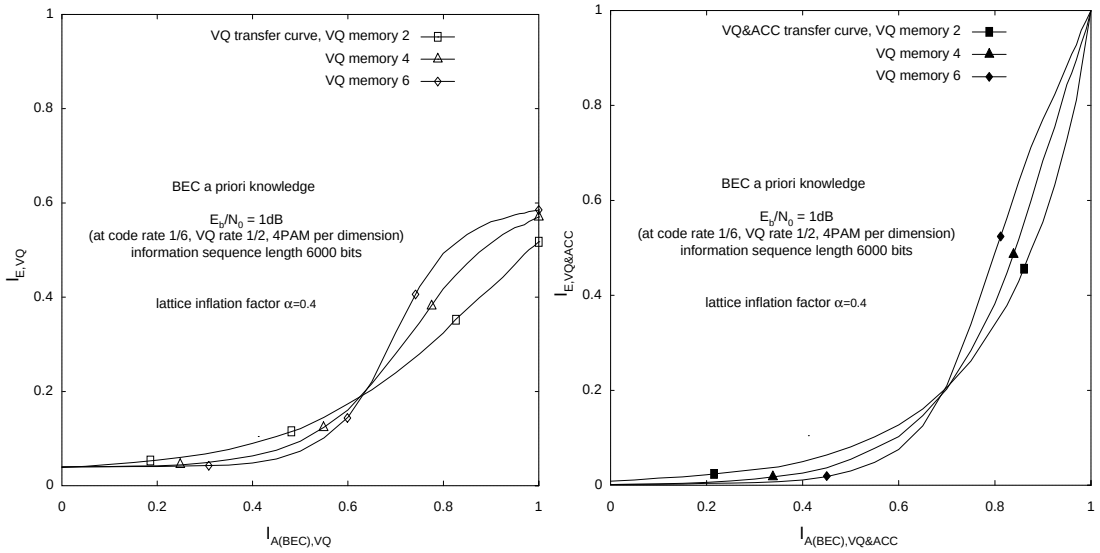

Figure 8: Left: Transfer curves of rate $R_{VQ} = 1/2$ vector quantizer and 4-PAM/dim.–metric for $\alpha = 0.4$ at $E_b/N_0 = 1$dB (code rate $R_{ch} = 1/6$ assumed); BEC *a priori* knowledge. Right: Curves from left, combined with a memory one accumulator decoder.

If we include the accumulator decoder into the inner detector, and use the joint trellis processing discussed in the previous section, we obtain the curves on the right, which, now, go up to $(1, 1)$. For the same VQ memory, the area under the curves stays about the same. Changing the upsampler has little influence on the shape of the VQ transfer curve as the vector quantizer "undos" any code protection offered by, e.g., a convolutional upsampler with memory, since it can freely flip the sign–bits at the transmitter. The upsampler is part of the VQ mapping

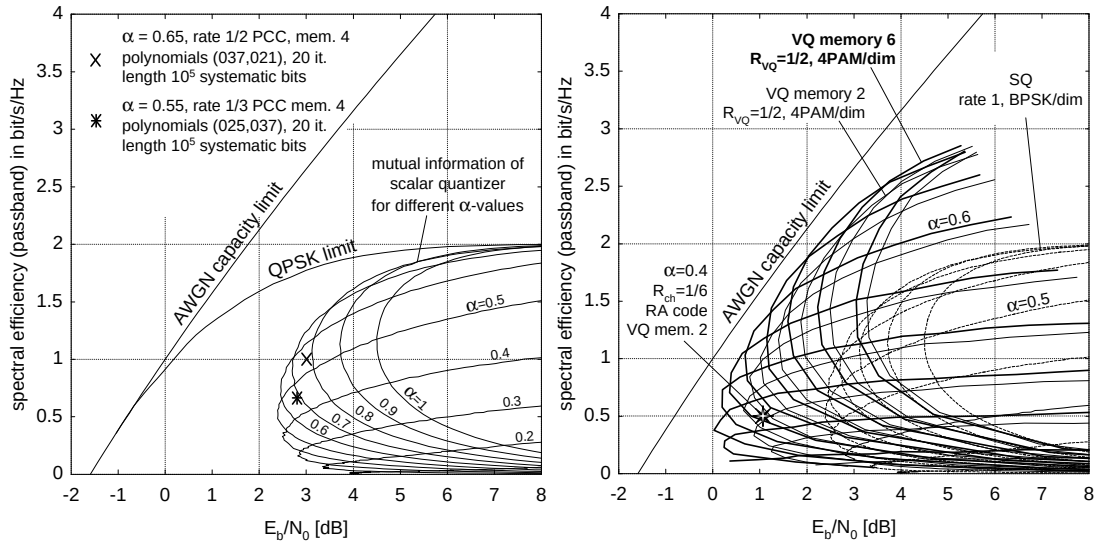

Figure 9: Left: Mutual information limits of SQ and BPSK/dim. for different α–values in the spectral efficiency plane; PCC turbo cliff positions at 3dB (1bit/s/Hz) and 2.8dB (0.667bit/s/Hz) respectively. Right: VQ and 4-PAM/dim., $R_{VQ} = 1/2$, VQ of memory 2 and memory 6; RA code turbo cliff position at 1.1dB (0.5bit/s/Hz).

process (mapping input bits to points in the Voronoi region), its role being to match the rates. Thus, we fixed it to the simplest possible code, i.e., a repetition code.

By computing transfer curves over different α– and E_b/N_0–values, and numerically evaluating the corresponding area, we obtain the mutual information limits given in Fig. 9 (right), plotted in the spectral efficiency chart.

As can be seen, the advantage of a memory 6 VQ over a memory 2 VQ is bigger for smaller spectral efficiencies. Note that these mutual information limits are obtained using area integration over EXIT curves. We still need to design an appropriate *iterative* decoding scheme to materialize these gains. This is discussed in the next section.

3.4 Code design example

We designed an RA code of rate $R_{ch} = 1/6$, to obtain an overall spectral efficiency of 0.5bit/s/Hz. The EXIT chart technique was used to find appropriate VND degree distributions. For this, the outer VND transfer curve is matched to the inner VQ&ACC&CND curve by means of curve fitting. For details on RA code design using the EXIT chart, see [3]. We chose a VQ of rate $R_{VQ} = 1/2$ and memory 2, with feedforward polynomials 07_8 and 05_8. A 4-PAM constellation was applied per dimension using natural labeling. We computed the inner detector curve (including VQ&4PAM, ACC and CND) by Monte–Carlo simulation, assuming a Gaussian model for the *a priori* information. We chose a biregular check node layer, with 80% of the check nodes being degree 1, and 20% being degree 3. Curve fitting at $E_b/N_0 = 1$dB yields a VND degree distribution of 64.36% variable nodes being degree 3, 31.24% degree 10, and 4.402% degree 76. We achieve convergence at 1.1dB (codeword length $K = 6 \cdot 10^4, N = 3.6 \cdot 10^5$ bits, 100 iterations, $\alpha = 0.4$). No error floor was observed for 40 blocks simulated, which can be attributed to the fact that there are no degree 2 variable nodes, and the lowest variable node degree is 3. Fig. 10 shows inner and outer transfer curves, and a simulated decoding trajectory at 1.2dB. The trajectory follows the individual transfer curves reasonably well.

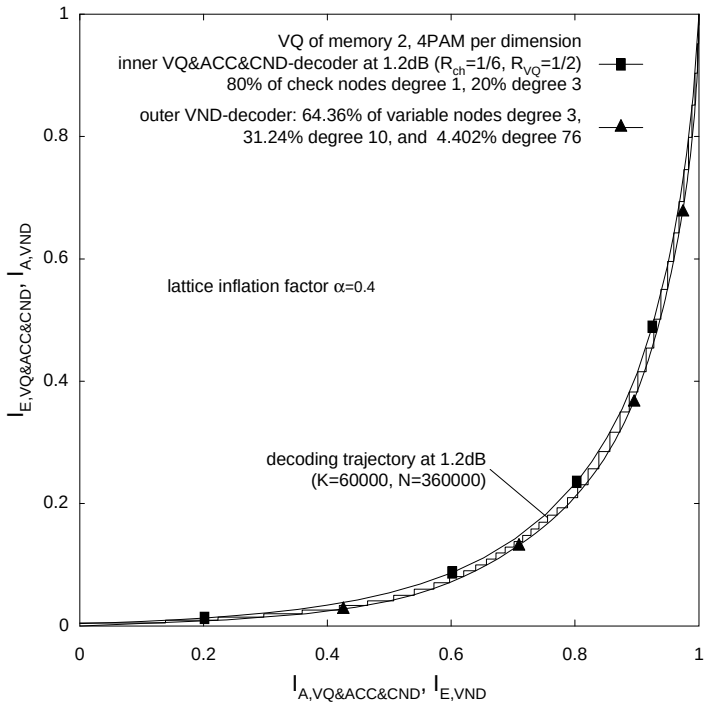

Figure 10: EXIT chart of rate $R_{ch} = 1/6$ nonsystematic RA code designed by curve fitting, with inner VQ&4PAM, ACC, CND–curve and outer VND–curve; $E_b/N_0 = 1.2$dB, codeword length $N = 3.6 \cdot 10^5$ coded bits; 100 iterations; $\alpha = 0.4$.

4 Summary

We presented a multidimensional dirty paper coding system that offers substantial gains over the performance of systems based on one-dimensional scalar quantization. While for a SQ system a simple AWGN turbo code together with a log–likelihood ratio modulo metric is sufficient to achieve reliable communication close to its mutual information limits, the gap to capacity of such systems is large at low SNR, and hence multidimensional quantization is required.

For the vector quantizer case, we showed how to perform iterative quantization and decoding, using a 4-PAM mapping, APP sign–bit and unsigned bit detection, together with a nonsystematic RA code. The combination of vector quantizer, upsampler and 4-PAM demapper was regarded as a detector, computing the respective L–values to be processed by a channel decoder. Moreover, merging this quantization detector with the inner memory one trellis decoder of an RA code allowed to obtain a joint inner curve which simplified code design. The design was exemplified for 0.5bit/s/Hz using an RA code of rate $R_{ch} = 1/6$. The improvement is more than 1.6dB over the best SQ–case. Still, a gap of about 1.8dB remains to the AWGN capacity at this spectral efficiency. However, the structure presented in this paper paves the way to even better performing VQ and decoding schemes for the dirty paper channel.

5 Acknowledgment

The authors would like to thank J. Salz and G. J. Foschini for providing the initial impetus for the endeavor reported in this work.

References

[1] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, *Optimal decoding of linear codes for minimizing symbol error rate*, *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, Mar. 1974.

[2] C. Berrou, A. Glavieux, and P. Thitimajshima, *Near Shannon limit error–correcting coding and decoding: Turbo–codes*, *Proc. ICC*, pp. 1064–1070, May 1993.

[3] S. ten Brink and G. Kramer, *Design of repeat-accumulate codes for iterative detection and decoding, to appear in IEEE Trans. Sign. Processing (Special Issue on MIMO Proc. Techn.)*, vol. 51, no. 11, Nov. 2003.

[4] S. ten Brink and G. Kramer, *Turbo processing for scalar and vector channels, Proc. 3rd Int. Symp. Turbo Codes and Related Topics*, Brest, France, pp. 23–30, Sept. 2003.

[5] S. ten Brink, *Exploiting the Chain Rule of Mutual Information for the Design of Iterative Decoding Schemes*, *39th Ann. Allerton Conf. on Commun., Control, and Computing*, Oct. 2001.

[6] G. Caire and S. Shamai, *Writing on Dirty Tape with LDPC codes. DIMACS Workshop on Signal Processing for Wireless Transmission, Rutgers University, NJ, USA*, Oct. 7-9, 2002.

[7] B. Chen and G. W. Wornell. *Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. IEEE Trans. Information Theory*, IT–47:1423–1443, May 2001.

[8] M. H. M. Costa. *Writing on dirty paper*, *IEEE Trans. Information Theory*, IT-29:439–441, May 1983.

[9] J. J. Eggers, J. K. Su, and B. Girod. *A blind watermarking scheme based on structured codebooks.* In *IEE Colloquium: Secure images and image authentication, London, UK*, Apr. 2000.

[10] U. Erez, S. Shamai (Shitz), and R. Zamir. *Capacity and lattice-strategies for cancelling known interference.* In *Proceedings of ISITA 2000, Honolulu, Hawaii*, pp. 681–684, Nov. 2000.

[11] M. V. Eyuboglu and G. D. Forney Jr. *Trellis precoding: combined coding, precoding and shaping for intersymbol interference channels*, *IEEE Trans. Information Theory*, IT-38:301–314, Mar. 1992.

[12] G. D. Forney. *Trellis shaping*, *IEEE Trans. Information Theory*, IT-38:281–300, Mar. 1992.

[13] J. Hagenauer, E. Offer, and L. Papke, *Iterative decoding of binary block and convolutional codes, IEEE Trans. Information Theory*, vol. 42, no. 2, pp. 429–445, Mar. 1996.

[14] H. Jin, A. Khandekar and R. McEliece, *Irregular repeat–accumulate codes,* In *Proc. 2nd Int. Conf. on Turbo Codes and Related Topics*, Brest, France, pp. 1–8, Sept. 2000.

[15] T. Philosof, U. Erez, and R. Zamir. *Combined shaping and precoding for interference cancellation at low SNR.* In *Proc. of Int. Symp. Inform. Theory (ISIT2003), Yokohama, Japan*, pp. 68, June 2003.

[16] F. M. J. Willems. *On Gaussian channels with side information at the transmitter.* In *Proc. of the Ninth Symposium on Information Theory in the Benelux*, Enschede, The Netherlands, 1988.