Combined Shaping and Precoding for Interference Cancellation at Low SNR

Tal Philosof, Uri Erez and Ram Zamir
Dept. of Elect. Eng. - Systems, Tel Aviv University, ISRAEL

Abstract

We investigate the application of the nested coding framework for cancelling known interference at low SNR, that is $P_X/P_Z \approx 1$ and below. We consider multi-dimensional precoding, with anticipation of $N_s > 1$ "future" interference samples. Unlike non-precoded transmission, where capacity can be achieved at low SNR without shaping, optimum precoding at low SNR *does* require shaping. Eyuboglu and Forney's trellis precoding scheme combines Tomlinson-Harashima precoding and trellis shaping, to achieve both coding and shaping gains in transmission over inter-symbol interference channels with Euclidean distance decoder. However, the standard configuration of this scheme does not support precoding at low SNR, where the capacity is less than $1\frac{\text{bit}}{\text{dim}}$. We propose a low rate precoding scheme which combines MMSE estimation, dithering and a variant of nested codes, based on concatenation of a "syndrome dilution" code and a "syndrome-to-coset" modulation code. We provide simulation results for several configurations of nested trellis codes, which reduce the gap to capacity.

1 Introduction

Consider the interference channel

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

where X is the channel input with power constraint $EX^2 \leq P_X$ and Y is the channel output, Z is additive i.i.d Gaussian noise $Z \sim N(0, P_Z)$ and S is an arbitrary interference that is known at the encoder. Let the SNR be $\frac{P_X}{P_Z}$. The specific channel model where the interference S is i.i.d. Gaussian and known non-causally at the transmitter was considered in [1], where it was shown that in this case $C = \frac{1}{2}\log_2(1 + SNR)$. Erez, Shamai and Zamir [3, 10] developed structured coding scheme for cancelling known interference using nested lattice strategies/codes ("lattice precoding"). This scheme uses a fine lattice - Λ_c in which a coarse lattice - Λ_s is nested, where the basic cell of the coarse lattice, $\mathcal{V}_s$, defines the region of the code and the codewords are points of the fine lattice. Thus the coarse lattice determines the shaping gain while the fine lattice the coding gain. Although Λ_c and Λ_s belong to the same multi-dimensional space, their *effective* dimensions N_c and N_s may differ.

The lattice precoding scheme also incorporates common randomness ("dither") and minimum mean squared error (MMSE) estimation, α. Specifically, the transmitted signal is given by $\mathbf{x} = [\mathbf{v} - \alpha\mathbf{s} - \mathbf{d}] \bmod \Lambda_s$ where $\mathbf{v}$ is the information bearing signal and $\mathbf{d} \sim U(\mathcal{V}_s)$ is the dither. Moreover, the receiver front end is $\mathbf{y}' = [\alpha\mathbf{y} + \mathbf{d}] \bmod \Lambda_s$. One of the main results of [3] is that this scheme induces an equivalent modulo-Λ_s additive noise channel

$$\mathbf{Y}' = \Big[\mathbf{V} + \mathbf{Z}'\Big] \bmod \Lambda_s, \tag{2}$$

where $\mathbf{Z}' = [(1-\alpha)\mathbf{U} + \alpha\mathbf{Z}] \bmod \Lambda_s$ is the equivalent noise. That is, $\mathbf{Z}'$ is a mixture of the Gaussian noise $\mathbf{Z}$ and "self noise", $\mathbf{U}$, which is uniformly distributed over the basic Voronoi region of Λ_s. For $\alpha = P_X/(P_X + P_Z)$ and high dimensional Λ_s ($N_s \to \infty$) with normalized second moment $G(\Lambda_s) \approx 1/2\pi e$, the self noise component approaches i.i.d Gaussian, consequently $\mathbf{Z}'$ is nearly i.i.d Gaussian with $E\mathbf{Z}'^2 = P_X P_Z/(P_X + P_Z)$. In this case, the full capacity $C = \frac{1}{2}\log(1 + SNR)$ is achieved.

The lattice precoding scheme may be viewed as a generalization of Tomlinson-Harashima (TH) precoding [9, 5], where S plays the role of inter symbol interference (ISI), and the scalar modulo operation of TH amounts to the special case of a $\mathbb{Z}^n$ coarse lattice, implying effectively one dimensional ($N_s = 1$) shaping, i.e., no shaping gain. A well known improvement to the TH precoding is the combined shaping and precoding scheme of Eyuboglu and Forney (EF), trellis precoding [4]. In this scheme EF combines between TH precoding and trellis shaping [6]. Trellis precoding achieves shaping gain by enforcing the multi-dimensional transmitted signal to be in a Voronoi like region.

In this work we assume that N_s is large enough so that the self noise $\mathbf{U}$ is "Gaussian enough", and thus an Euclidean decoder is nearly the optimum decoder. However in a related study [8], for $N_s = 1$ we showed that a decoder matched to the equivalent noise $\mathbf{Z}'$ has improved performance compared to an Euclidean decoder.

2 Precoding "shaping-gap"

The ultimate precoding "shaping-gap" is the power (or SNR) gap between precoding using one dimensional coarse lattice and precoding using high dimensional coarse lattice with $G(\Lambda_s) \approx 1/2\pi e$ (which achieves the capacity), for a fixed mutual information. Unlike the interference-free AWGN channel, this gap is particularly significant at the low SNR regime for the interference channel. Figure 1.a shows the mutual information of a one dimensional lattice strategy achieved by (2), the mutual information achieved by a uniform input over an interference-free AWGN channel, and the capacity $C = \frac{1}{2}\log(1 + SNR)$. We refer to the SNR gap between the first two and the latter as "shaping-gap", at high SNR the interference-free AWGN and the interference channels have an identical shaping gap of 1.53 dB. At low SNR, the interference-free AWGN channel has no shaping gap, on the other hand the interference channel has over 1.53 dB precoding shaping-gap. From Figure 1.a, at $SNR = 0$ dB the precoding shaping-gap is 3.1 dB, while the shaping-gap of interference-free AWGN channel is close to zero.

The rate loss between one dimensional precoding scheme and N_s dimensional precoding scheme is bounded by $\frac{1}{2}\log(2\pi e G_{N_s}(\Lambda_s)) \leq 0.254$ bit/dim, [3], for *any* SNR. Nevertheless, the capacity curve with respect to the SNR (in logarithmic scale) is more sensitive at low SNR (the "6 dB per bit" is not valid at low SNR), i.e, the loss in dB increases.

The shaping gap at the low SNR regime is well understood by drawing the mutual information with respect to $\frac{SNR}{2I(\mathbf{V};\mathbf{Y}')}$ (or $\frac{SNR}{2R}$), which is equivalent to E_b/N_0. Figure 1.b shows the capacity, the mutual information of one dimensional lattice strategy and one dimensional lattice strategy with time sharing at the interval $SNR \in [0, 1]$ (where $SNR_c = 1$ is critical SNR for optimal transmission with time sharing at the low SNR regime). For $R \to 0$, the Shannon limit is at $E_b/N_0 = -1.59$ dB, the gap between the lattice strategy with time sharing and the capacity is bounded by 4 dB.

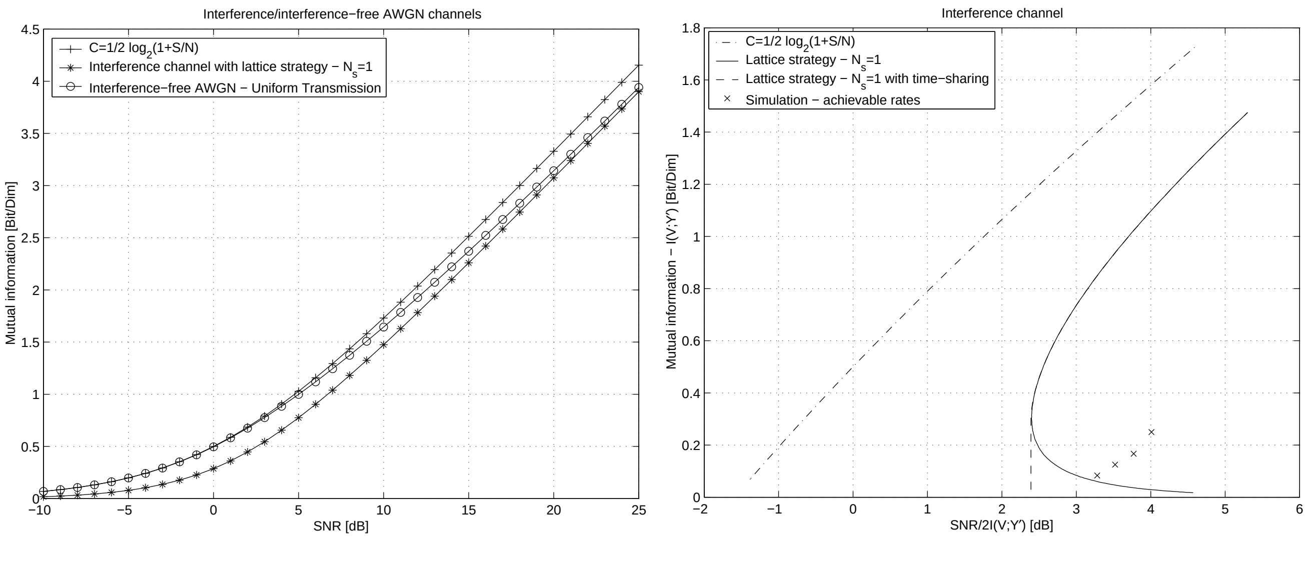

Figure 1: a. Precoding shaping-gap vs SNR. b. Precodeing shaping-gap vs $\frac{SNR}{2I(\mathbf{V};\mathbf{Y}')}$.

3 The nested lattices construction

"Construction A" is a fundamental method to construct a lattice $\Lambda \in \mathbb{R}^n$ using a linear code $\mathcal{C}(n, k) \in \mathbb{Z}_p$, where p is prime number. Since $\mathcal{C}$ is in the region of n-cube $[0, p)^n$, the lattice Λ is obtained by tessellating

$\mathbb{R}^n$ with translations of $\mathcal{C}$, i.e., $\Lambda = \mathcal{C} + p\mathbb{Z}^n$, thus Λ is a sub-lattice of $\mathbb{Z}^n$. Furthermore, "construction A" of random ensembles of lattices achieve the interference-free AWGN channel capacity [7, 2], therefore good lattices can be generated by "construction A". Generally, a sub-lattice Λ_2 of Λ_1 induces a nested lattice partition Λ_1/Λ_2 of Λ_1 into $|\{\Lambda_1 \bmod \Lambda_2\}|$ cosets of Λ_2.

We present the nested lattices construction for Λ_c and Λ_s so that $\mathbb{Z}^N/\Lambda_c/\Lambda_s/p\mathbb{Z}^N$, as shown in [8]. Initially, we construct Λ'_s by "construction A" with linear code $\mathcal{C}_s(n_s, k_s)$ over $\mathbb{Z}_p$, meaning $\Lambda'_s = \mathcal{C}_s + p\mathbb{Z}^{n_s}$, thus Λ'_s is a sub-lattice of $\mathbb{Z}^{n_s}$. The code $\mathcal{C}_s$ contains all the codewords $\{\mathbf{c}_s\}$ so that $\mathbf{c}_s \in \mathbb{Z}_p^{n_s}$ and whose syndrome, $\mathbf{s_y} \triangleq H_s\mathbf{c}_s$, is equal to zero where $\mathbf{s_y} \in \mathbb{Z}_p^{n_s-k_s}$. The coset leader group is defined as $\{\mathbb{Z}^{n_s} \bmod \Lambda'_s\}$, while a coset representative group includes a unique representative of each coset, but not necessarily the coset leaders. The coset representative and coset leader groups have an equal number of elements and they are equivalent modulo Λ'_s. There are $p^{n_s}/|\mathcal{C}_s| = p^{n_s}/p^{k_s} = p^{n_s-k_s}$ coset representatives, which are equal to number of $\mathcal{C}_s$ syndromes. Each syndrome $\mathbf{s_y}$ corresponds to a unique coset representative. In order to obtain a $\mathcal{C}_s$ coset from a specific syndrome, we can apply the "pseudo" right inverse parity check matrix on the syndrome [4], $\mathbf{t} = H_s^{-1}\mathbf{s_y}$ so that $\mathbf{s_y} = H_s\mathbf{t}$, that is, $\mathbf{t}$ represents a coset of $\mathcal{C}_s$. Finally, the sub-lattice Λ_c of $\mathbb{Z}^{n_s}$ is obtained by dilution of $\mathbb{Z}^{n_s}$ codewords, which is accomplished by dilution of $\mathcal{C}_s$ syndromes. Specifically, it is done by linear code $\mathcal{C}_c(n_c, k_c)$ over $\mathbb{Z}_p$, which produces a sub-group of syndromes. The nested lattices construction is shown schematically in Figure 2.

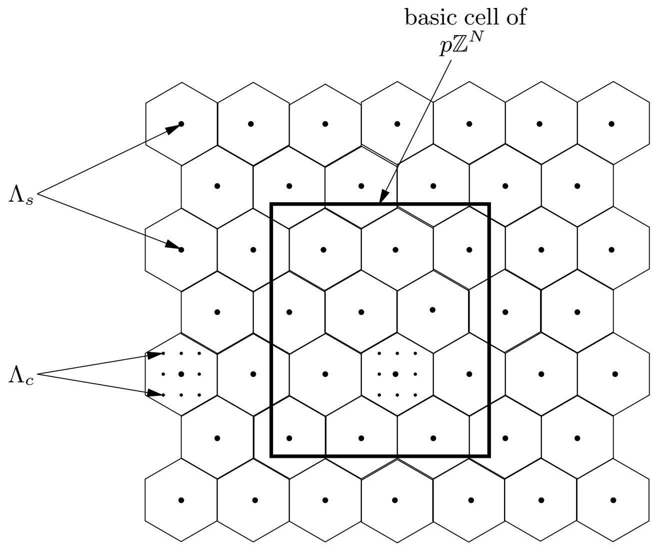

Figure 2: Nested lattices construction

The above construction implies effective rate of $\frac{1}{N}\log_2(V_s/V_c)$ [bit/dim], where V_c and V_s are the cells volume of Λ_c and Λ_s, respectively. Assume n_c is a multiple of $n_s - k_s$, then the construction $\mathbb{Z}^N/\Lambda_c/\Lambda_s/p\mathbb{Z}^N$ has $N = n_s\frac{n_c}{n_s-k_s}$ and Λ_s is $\frac{n_c}{n_s-k_s}$ times a Cartesian product of Λ'_s, where the rate is

$$R = \frac{1}{N}\log_2(V_s/V_c) = \frac{1}{N}\log_2(p^{k_c}) = \frac{k_c(n_s - k_s)}{n_c n_s}\log_2 p \text{ [bit/dim]}. \tag{3}$$

Furthermore, for any n_c the construction should be extended by Cartesian product to dimension $N = n_s\frac{LCM(n_c, n_s-k_s)}{n_s-k_s}$, nevertheless the rate is unchanged.

4 Transmission scheme

The transmission scheme is based on the following generalization of the "inflated lattice lemma" [3].

Lemma 1. ***(Generalized inflated lattice lemma)*** *For the nested lattices chain $\Lambda_G/\Lambda_c/\Lambda_s/\Lambda_A$ and the channel defined by (1).*
Encoder:

$$\mathbf{x} = [\mathbf{v}' - \alpha\mathbf{s} - \tilde{\mathbf{d}}] \; mod \, \Lambda_s, \tag{4}$$

where $\mathbf{v}' \in \{\Lambda_c \cap \mathcal{V}_A\}$, $\tilde{\mathbf{d}} \sim U(\mathcal{V}_A)$ and $\mathcal{V}_A$ is the basic cell Λ_A.
Decoder:

$$\mathbf{y}' = [\alpha\mathbf{y} + \tilde{\mathbf{d}}] \; mod \, \Lambda_A. \tag{5}$$

For $\mathbf{v} \in \{\Lambda_c \cap \mathcal{V}_s\}$ the equivalent channel satisfies

$$\mathbf{Y}' = [\mathbf{v} + \mathbf{B} + \mathbf{Z}'] \; mod \, \Lambda_A, \tag{6}$$

with

$$\mathbf{Z}' = [(1-\alpha)\mathbf{U} + \alpha\mathbf{Z}] \; mod \, \Lambda_A, \tag{7}$$

where $\mathbf{B}$ is some point in Λ_s which is a function of $\mathbf{v}'$, and $\mathbf{U} \sim U(\mathcal{V}_s)$.

Applying the lemma to the nested lattices construction, where $\Lambda_G = \mathbb{Z}^N$ and $\Lambda_A = p\mathbb{Z}^N$, the equivalent channel till $\mathbf{y}'$ is given by $\mathbf{Y}' = [\mathbf{v} + \mathbf{B} + \mathbf{Z}']$ mod $p\mathbb{Z}^N$ and $\mathbf{B} \in \Lambda_s$. The effect of $\mathbf{B}$ can be cancelled by maximum likelihood (ML) decoding of $\mathbf{y}'$ for the nearest Λ_c codeword, and by syndrome detection of the corresponding coset of Λ_s in Λ_c.

The transmission scheme which incorporates the lattice precoding and the nested lattices construction is presented in Figure 3. The Λ_s second moment is given by $\sigma_s^2 = P_X$.

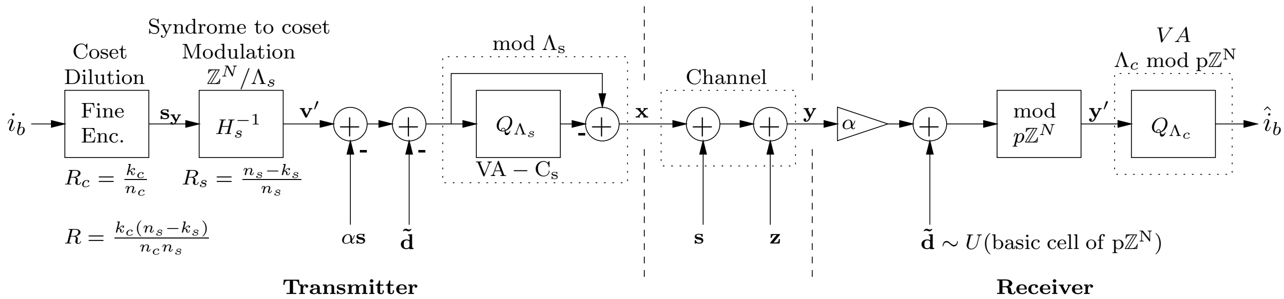

Figure 3: Transmission scheme

Transmitter - the $\mathcal{C}_s$ syndromes dilution is done by a fine code encoder, where its output is a $\mathcal{C}_s$ syndrome, $\mathbf{s_y} \in \mathbb{Z}_p^{n_s-k_s}$. The shaping encoder, H_s^{-1}, transforms a $\mathcal{C}_s$ syndrome to $\mathcal{C}_s$ coset representative $\mathbf{v}'$, of Λ_s in Λ_c over the basic cell of $p\mathbb{Z}^N$, specifically $\mathbf{v}' = H_s^{-1}\mathbf{s_y}$ is a discrete point in Λ_c. The dither signal $\tilde{\mathbf{d}}$ is uniformly distributed over the basic cell of $p\mathbb{Z}^N$. The transmitted vector is $\mathbf{x} = [\mathbf{v}' - \alpha\mathbf{s} - \tilde{\mathbf{d}}]$ mod Λ_s, it can be written as $\mathbf{x} = [\mathbf{v} - \alpha\mathbf{s} - \mathbf{d}]$ mod Λ_s where $\mathbf{v} \in \{\Lambda_c \text{ mod } \Lambda_s\}$ is coset leader and $\mathbf{d} \sim U(\mathcal{V}_s)$, as in [3]. Since $\mathbf{d}$ is uniform over $\mathcal{V}_s$, the transmitted vector $\mathbf{x}$ is also uniform over $\mathcal{V}_s$ by the property of dithered quantization. Therefore, the power constraint is $\frac{1}{N}E\{||\mathbf{X}||^2\} = \sigma_s^2 = P_X$.
Receiver - for information bits reconstruction, the receiver has to estimate the transmitted coset. Initially, the receiver calculates $\mathbf{y}' = [\alpha\mathbf{y} + \tilde{\mathbf{d}}]$ mod $p\mathbb{Z}^N$. Finally, since the equivalent channel till $\mathbf{y}'$ is given by $\mathbf{Y}' = [\mathbf{v}+\mathbf{B}+\mathbf{Z}']$ mod $p\mathbb{Z}^N$ where $\mathbf{B} \in \Lambda_s$, therefore $\mathbf{V} + \mathbf{B}$ is in the same coset of $\mathbf{V}$. The coset estimation can be computed as $\widehat{(\mathbf{v} + \mathbf{b})} = Q_{\Lambda_c}(\mathbf{y}')$ where Q_{Λ_c} is performed over the basic cell of $p\mathbb{Z}^N$.

Although the transmission scheme has been constructed from nested lattice codes, it can be interpreted as structure of concatenated codes. This structure is composed of an inner code - performs syndrome-to-coset-modulation for constellation shaping, and an outer code - adds redundancy for error correction to enhance the noise immunity. Furthermore, the outer code is responsible for the coding gain while the inner code for the shaping gain.

A detailed "rate flow diagram" of the encoder is illustrated in Figure 4, where the m_s, m_c integers are determined by $LCM(n_c, n_s - k_s) = m_s(n_s - k_s) = m_c n_c$. A K-tuples of p-ary symbols is segmented into m_c vectors of length k_c, each of the m_c vector is encoded by $R_c = k_c/n_c$ fine encoder. Therefore, at the fine encoder output there are m_c vectors of length n_c. The $m_c n_c$ p-ary symbols are segmented into m_s vectors of $n_s - k_s$ symbols, each of the m_s vector is encoded by $R_s = \frac{n_s-k_s}{n_s}$ shaping encoder, thus at the shaping encoder output there are m_s vectors of length n_s symbols. The encoder rate is given by

$$R = \frac{K}{N} = \frac{m_c k_c}{m_s n_s} = \frac{k_c(n_s - k_s)}{n_c n_s} \log_2 p \tag{8}$$

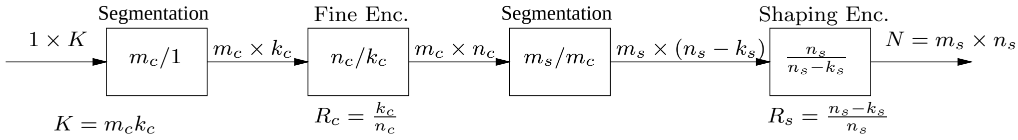

Figure 4: Encoder rate flow diagram

5 Nested trellis codes

A straightforward implementation of the above scheme, for $p = 2$, uses binary convolutional codes $\mathcal{C}_c(n_c, k_c)$, $\mathcal{C}_s(n_s, k_s)$ to construct the nested binary lattice construction $\mathbb{Z}^N/\Lambda_c/\Lambda_s/2\mathbb{Z}^N$. The loss of using $p = 2$ with shaping code rate $R_s = 1/2$ is 0.25 dB from the entire 3.1 dB shaping gap at SNR = 0 dB, see [8]. Furthermore, as long as the SNR decreases the loss increases, therefore larger p is needed to achieve the capacity at low SNR.

The coset representative, $\mathbf{v}'$, is a series of one dimensional symbols $\{0, 1\}$. The dither vector $\tilde{\mathbf{d}} \sim U(\text{basic cell of } 2\mathbb{Z}^{\mathrm{N}})$ consists of continuous values random variables $\tilde{d}_i \sim U(-1, 1)$. The modulo Λ_s operation can be written as $\mathbf{x} \bmod \Lambda_s = \mathbf{x} - Q_{\Lambda_s}(\mathbf{x})$, where Q_{Λ_s} can be implemented by Viterbi algorithm (VA) decoder for the nearest $\mathcal{C}_s$ codeword, since Λ_s has Voronoi partition. The receiver performs modulo $2\mathbb{Z}^N$, or equivalently one dimensional modulo to the interval $[-1, 1)$. The quantizer Q_{Λ_c} over the region of the cube $[-1, 1)^N$ is implemented by VA decoder for the nearest Λ_c codeword in the region of $[-1, 1)^N$. The lattices effective dimensions N_c, N_s are mainly affected by the constraint length ν_c, ν_s of $\mathcal{C}_c$ and $\mathcal{C}_s$, respectively.

5.1 Performance results

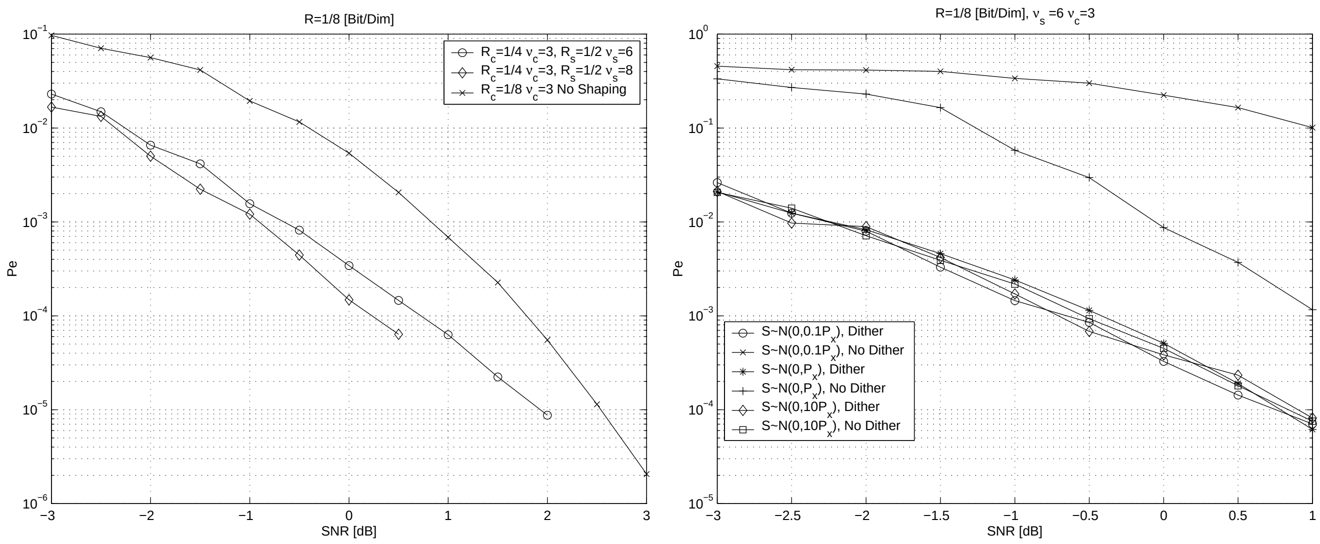

Figure 5: a. Shaped versus un-shaped system. b. Gaussian interference.

Figure 5.a shows 1.8 dB precoding shaping gain of the shaped system with $\nu_s = 8$, compared to the un-shaped system at $P_e = 1 \cdot 10^{-3}$ (about $SNR = 0\ dB$). However, the precoding shaping-gap is bounded by 3.1 dB at $SNR = 0\ dB$. The precoding system immunity to the interference signal is shown in Figure 5.b for Gaussian interference. As long as the dither signal is used, the system performance is the same for any interference power level and any interference distribution. For non-dithered system the performance is interference dependent.

Figure 1.b illustrates the shaped systems performances with transmission rate $\frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{12}$ bit/dim at $P_e = 1 \cdot 10^{-3}$. These systems operate about 4.5 dB from the capacity, although modest convolutional codes have been used as fine code.

6 Discussion

The above scheme can incorporate different codes, especially Turbo codes or LDPC codes as a fine code with some decoder changes. Using these codes in our scheme enables to approach the interference channel capacity. Theoretically, for any good fine code we use, without a shaping code the capacity can not be achieved. Furthermore, at low SNR the precoding shaping-gap increases, therefore the shaping code requires specific consideration.

Acknowledgments

The authors wish to thank G. D. Forney for his helpful comments.

References

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

[2] U. Erez, S. Litsyn, and R. Zamir, *Lattices which are good for (almost) everything*, IEEE Trans. Information Theory (in preparation).

[3] U. Erez, S. Shamai (Shitz), and R. Zamir, *Capacity and lattice-strategies for cancelling known interference*, IEEE Trans. Information Theory (submitted), Also presented at the ISITA 2000, Honolulu, Hawaii, pages 681–684 (Nov. 2000).

[4] 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** (March, 1992), 301–314.

[5] H. Harashima and H. Miyakawa, *Matched-transmission technique for channels with intersymbol interference*, IEEE Trans. Communications **COM-20** (Aug. 1972), 774–780.

[6] G. D. Forney Jr., *Trellis shaping*, IEEE Trans. Information Theory **IT-38** (Mar. 1992), 281–300.

[7] H. A. Loeliger, *Averaging bounds for lattices and linear codes*, IEEE Trans. Information Theory **43** (Nov. 1997), 1767–1773.

[8] T. Philosof, *Precoding for interference cancellation at low SNR using nested codes*, Master's thesis, Tel-Aviv University, (in preparation).

[9] M. Tomlinson, *New automatic equalizer employing modulo arithmetic*, Electronic Lett. **vol. 7** (Mar. 1971), 138–139.

[10] R. Zamir, S. Shamai (Shitz), and U. Erez, *Nested linear/lattice codes for structured multiterminal binning*, IEEE Trans. Information Theory **IT-48** (June 2002), 1250–1276.