Applying Erez and Ten Brink's Dirty Paper Codes to Data-Hiding *†

Pedro Comesaña[1], Fernando Pérez-González[1] and Frans M. J. Willems[2]

[1] Dept. Teoria de la Señal y Comunicaciones. ETSI Telecom., Universidad de Vigo, 36310 Vigo, Spain

[2] Dept. of Electrical Engineering, Technische Universiteit Eindhoven, Postbus 513, 5600 MB Eindhoven, The Netherlands

ABSTRACT

Structured codes are known to be necessary in practical implementations of capacity-approaching "dirty paper schemes". In this paper we study the performance of a recently proposed dirty paper technique, by Erez and ten Brink which, to the authors' knowledge, is firstly applied to data-hiding, and compare it with other existing approaches. Specifically, we compare it with conventional side-informed schemes previously used in data-hiding based on repetition and turbo coding. We show that a significant improvement can be achieved using Erez and ten Brink's proposal. We also study the considerations we have to take into account when these codes are used in data-hiding, mainly related with perceptual questions.

1. INTRODUCTION.

In the last years the usefulness of approaching watermarking as a communication problem with side information known at the encoder but not at the decoder has been proven. This model was shown by Costa[1] to achieve the same capacity as if the side information were also made available to the decoder. Nevertheless, the main problem with Costa's construction is that it relies on random codes, which require an exhaustive search strategy for selecting the codeword to be used, something that is largely impractical. Due to the importance of Costa's result, not only to watermarking, but also to many other applications in communications, a large number of papers dealing with the possibility of approaching the same result using structured codes have been written.[2,3]

Erez and Zamir have recently shown[3] that Costa's result can be achieved with nested lattices. In fact, they have proven a stronger result in which a modulo-lattice transformation of the received signal is considered at the decoder (*lattice decoding* will be dealt with in Section 3). This obviously implies a huge reduction in complexity, as well as the possibility of achieving capacity without explicitly knowing the probability density function (pdf) of the host signal. Nevertheless, the question of code construction is not completely solved: Erez and Zamir's result applies to lattices verifying quite strict conditions which require that the fundamental regions approach hyperspheres asymptotically as the number of dimensions is increased. Unfortunately, those conditions fall short of being met by the simplest (and mostly used) lattices, such as the cubic ones. Therefore, practical solutions demand the use of strategies whose complexity does not rely exclusively on lattices.

The usually followed solution is to encode the information bits with a near-Shannon-limit channel code and then take the output bits to index the sub-lattice used to quantize the host signal. Due to the redundancy introduced by the channel code, this lattice can be a very simple one, even allowing for scalar quantization.

*This work was partially funded by *Xunta de Galicia* under projects PGIDT04 TIC322013PR and PGIDT04 PXIC32202PM; MEC project DIPSTICK, reference TEC2004-02551/TCM; FIS project IM3, reference G03/185, and European Comission through the IST Programme under Contract IST-2002-507932 ECRYPT.

Further author information: (Corresponding author: P.C.)
P.C.: E-mail: pcomesan@gts.tsc.uvigo.es, Telephone: +34 986 812683
F.P-G.: E-mail: fperez@tsc.uvigo.es, Telephone: +34 986 812124
F.M.J.W.: E-mail: f.m.j.willems@tue.nl, Telephone: +31 40 247 3539

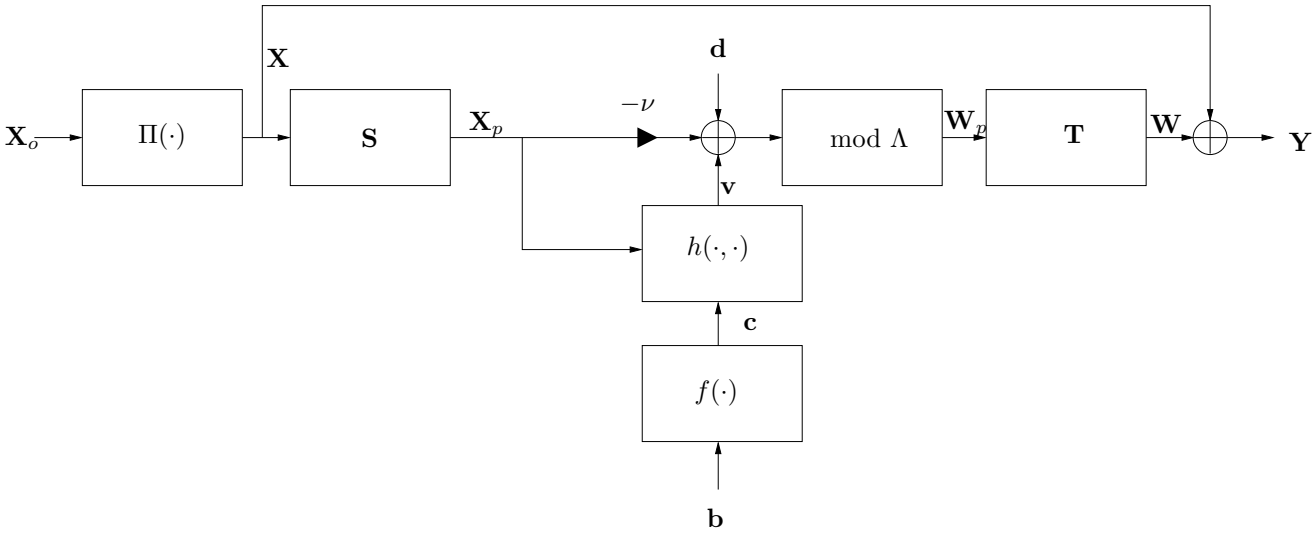


Figure 1. General structure of a dirty-paper encoder.

The good results obtained with this kind of schemes can be explained from the fact that the channel code concatenated with the simple lattice is equivalent to a better (and also more involved) lattice.

Summarizing, most of the practical schemes that use structured codes to approach Costa's result are composed of a good channel code concatenated with a quite simple lattice. The encoding and decoding with the channel code is usually relatively easy, and the same applies when a simple lattice is chosen, in such a way that the resulting dirty paper coding schemes fall very close to Shannon's limit, while keeping a reasonable computational cost.

The rest of the paper is organized as follows: In Sect. 2 the notation and a unified framework are introduced. In Sect. 3 the lattice decoding strategy is reviewed. The classical dirty paper coding schemes are studied following the unified framework in Sect. 4 whereas Erez and ten Brink's scheme constitutes the main subject of Sect. 5. Experimental results are shown in Sect. 6 and conclusions in Sect. 7.

2. NOTATION AND UNIFIED FRAMEWORK

We will denote scalar random variables with capital letters (e.g., X), and their outcomes with lowercase letters (e.g., x). The same notation criterion applies to random vectors and their outcomes, denoted in this case by bold letters (e.g., $\mathbf{X}$, $\mathbf{x}$).

The general diagram of the dirty-paper coding schemes studied in this paper is plotted in Fig. 1. We assume without loss of generality that the host signal is represented by a zero-mean random vector $\mathbf{X}^o = (X_1^o, \cdots, X_L^o)^T$. If necessary, these particulars can always be achieved by subtracting any non-zero mean from the host, and by using an arbitrary bijective transformation from the original arrangement of the host signal samples to a unidimensional one. Prior to embedding we apply a key-dependent pseudorandom permutation $\Pi(\cdot)$ to $\mathbf{X}^o$. The permuted host is $\mathbf{X} \triangleq \Pi(\mathbf{X}^o)$. Apart from the security increase due to the uncertainty that this permutation procedure causes to an attacker unaware of the key, an important advantage from an analytical point of view is the statistical independence between consecutive samples brought about by such key-dependent permutation. We will denote the average host signal power as

$$D_h = \frac{1}{L}\sum_{i=1}^{L}\sigma_{X_i}^2, \tag{1}$$

where $\sigma^2_{X_i} \triangleq \mathrm{Var}\{X_i\}$. The permuted host could be projected onto a P-dimensional space ($P < L$), in order to perform the embedding in the new domain (as, for instance, in the Spread Transform Dither Modulation proposed by Chen and Wornell[2]), yielding

$$\mathbf{X}_p = \mathbf{S} \cdot \mathbf{X}, \tag{2}$$

being $\mathbf{S}$ a $P \times L$ matrix.

In side-informed schemes (on which we will exclusively focus), the watermarked signal $\mathbf{Y}$ will be obtained from both the projected host signal $\mathbf{X}_p$ and the M_b-length information message $\mathbf{b}$ to be embedded. We will assume that $\mathbf{b} = (b_1, \cdots, b_{M_b})^T$ is a binary vector, $b_j \in \{0, 1\}$ for $j = 1, \cdots, M_b$. This message could go through a binary channel encoder $f(\cdot)$, so $\mathbf{c} = f(\mathbf{b})$ is the M_c-length channel-coded message.

The code in the vector quantizer $h(\cdot, \cdot)$ will transform the M_c-length channel-coded binary message $\mathbf{c}$ into a P-length vector $\mathbf{v}$, with elements in the alphabet $\mathcal{V}$. The vector $\mathbf{v}$ will depend on both $\mathbf{c}$ and $\mathbf{X}_p$. Therefore, $h(\cdot, \cdot)$ will only make sense when the vector quantizer is really use; for example, in Sect. 4, where cartesian products of scalar vectors are used, its output $\mathbf{v}$ will be just $\mathbf{c}$.

Let

$$\Lambda \triangleq |\mathcal{V}|\mathbb{Z}^P, \tag{3}$$

then, given $\mathbf{v}$, a shifted-lattice quantizer, $\mathbf{Q}_{\mathbf{v}}(\cdot)$, based on a minimum Euclidean distance criterion is defined as

$$\mathbf{Q}_{\mathbf{v}}(\mathbf{a}) = \mathbf{Q}_{\Lambda}\left(\mathbf{a} - \mathbf{t}(\mathbf{v})\right) + \mathbf{t}(\mathbf{v}), \quad \text{for any } \mathbf{a} \in \mathbb{R}^P \tag{4}$$

where $\mathbf{Q}_{\Lambda}(\cdot)$ is the minimum Euclidean distance quantizer induced by the lattice Λ, and $\mathbf{t}(\mathbf{v}) = \mathbf{v} + \mathbf{d}$. Vector $\mathbf{d}$ is a realization of a key-dependent pseudorandom dither vector $\mathbf{D}$, which is uniformly distributed over the Voronoi region of Λ, so in the j-th component $D_j \sim U[-|\mathcal{V}|/2, |\mathcal{V}|/2), 1 \leq j \leq P$. The purpose of vector $\mathbf{d}$ is to increase security by making it available only to encoder and decoder.

The watermark in the projected domain ($\mathbf{W}_p$) will be then computed as

$$\mathbf{W}_p \triangleq \mathbf{Q}_{\mathbf{v}}(\nu\mathbf{X}_p) - \nu\mathbf{X}_p, \tag{5}$$

which is nothing but the quantization error resulting when quantizing $\nu\mathbf{X}_p$ with the quantizer $\mathbf{Q}_{\mathbf{v}}(\cdot)$ corresponding to the message $\mathbf{v}$. Considering the structure of the lattice defined in (3), it is clear that the quantization in (5) can be implemented in sample-by-sample basis. The distortion-compensation parameter ν, $0 < \nu \leq 1$, is an optimizable variable akin to the one in Costa's paper.[‡]

The inverse projection will be given by the $L \times P$-matrix $\mathbf{T}$, so

$$\mathbf{W} = \mathbf{T}\mathbf{W}_p, \tag{6}$$

where $\mathbf{T}$ could be any matrix verifying

$$\mathbf{S} \cdot \mathbf{T} = \mathbf{I}_{P \times P}, \tag{7}$$

although the minimum-norm valid matrix is the pseudoinverse $\mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1}$.

In a similar way that the average host signal power, the *embedding distortion* is defined as

$$D_w \triangleq \frac{1}{L}\sum_{i=1}^{L} \mathrm{E}\{W_i^2\}. \tag{8}$$

The imperceptibility of the differences between $\mathbf{X}$ and $\mathbf{Y} \triangleq \mathbf{X} + \mathbf{W}$ has to be guaranteed by means of a perceptual analysis of the host signal previous to the embedding operation. This procedure is intrinsically

[‡]In fact ν should be a Wiener filter, as it is discussed by Yu et al.[4] Nevertheless, since this would require the embedder to know the noise power distribution we will assume it to be a scalar, even if it is not the optimal choice.

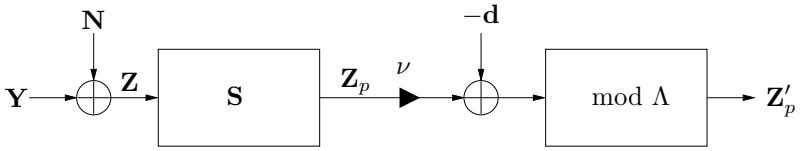


Figure 2. Scheme of the channel and predecoder.

dependent on the type of host signal in question. Due to this fact, we will consider that the host ($\mathbf{X}^o$) is a multimedia signal given in a certain domain of interest. The only requirement is that the domain chosen is suited to compute a *perceptual mask* $\boldsymbol{\alpha}$, taking into account human perceptual features. We assume in the following that the maximum unnoticeable amplitude modification of the corresponding host signal sample X_i is proportional to α_i; therefore, the quantization step in each dimension Δ_i will be made proportional to α_i.

Decoding is accomplished by the receiver after the watermarked signal $\mathbf{y}$ has undergone an attack channel. Throughout this paper we will focus on the case where this channel is a zero-mean additive probabilistic channel independent of $\mathbf{X}$ and $\mathbf{b}$, yielding a received signal $\mathbf{Z} = \mathbf{Y} + \mathbf{N}$. This type of channel model has been consistently used for benchmarking purposes in most relevant data hiding research. We will also assume that the samples in $\mathbf{N}$ are mutually independent, with diagonal covariance matrix $\Gamma = \text{diag}(\sigma^2_{N_1}, \cdots, \sigma^2_{N_L})$. The *channel distortion* D_c can be then defined in a similar fashion as the embedding distortion, i.e.,

$$D_c \triangleq \frac{1}{L}\sum_{i=1}^{L}\sigma^2_{N_i}. \tag{9}$$

We would like to remark that this kind of measurement would in principle allow to concentrate all the attacking distortion on a single sample of $\mathbf{Y}$ or spread it over all the vector. This freedom to distribute distortion hints at the poor connection existing between perceptual issues and this kind of mean square error (MSE) distortion measurements.

Nevertheless, we will undertake all subsequent analyses using MSE, as this criterion has been the most employed in the literature so far for the sake of tractability. Notice, for instance, that the hypotheses of Costa's result are stated for this type of restriction. In any case, an attacker may try to partially relieve the intrinsic inconveniences of MSE in order to comply with the usual requirement of minimal perceptual impact of the attack. Assuming the adequacy of the perceptual mask, it is clear that one way to meet this condition is to *perceptually shape* the added noise, such that its variance at each dimension is proportional to the corresponding allowable perceptual energy. Last, we will find it useful to introduce the *watermark-to-noise ratio* as

$$\text{WNR} \triangleq \frac{D_w}{D_c}, \tag{10}$$

that relates the power of the embedding and channel distortion, establishing a working point similar to the signal-to-noise ratio (SNR) in communications.

The received signal $\mathbf{Z}$ will be projected using $\mathbf{S}$ in order to obtain $\mathbf{Z}_p = \mathbf{S}\cdot\mathbf{Z}$, the projected signal which will be used in the decoding process. In a similar way, $\mathbf{N}_p = \mathbf{S}\cdot\mathbf{N}$ denotes the projected noise.

3. LATTICE DECODING

In this section we describe the lattice decoding technique proposed by Erez et al.,[3] and which is used by all the methods studied in this paper. Both the channel and predecoder are plotted in Fig. 2. Lattice decoding is solely based on the statistic

$$\begin{aligned}
\mathbf{Z}'_p &= (\nu\mathbf{Z}_p - \mathbf{d}) \mod \Lambda && (11)\\
&= (\mathbf{W}_p - \mathbf{d} + (\nu-1)\mathbf{W}_p + \nu\mathbf{X}_p + \nu\mathbf{N}_p) \mod \Lambda && (12)\\
&= [(\mathbf{v} + \mathbf{d} - \nu\mathbf{X}_p) \mod \Lambda - \mathbf{d} - (1-\nu)\mathbf{W}_p + \nu\mathbf{X}_p + \nu\mathbf{N}_p] \mod \Lambda && (13)\\
&= [\mathbf{v} - (1-\nu)\mathbf{W}_p + \nu\mathbf{N}_p] \mod \Lambda, && (14)
\end{aligned}$$

This means that not all the possible centroids associated to a symbol have to be considered, but just one. Moreover, a huge complexity reduction can be achieved when $\mathbf{W}_p = [\mathbf{d} + \mathbf{v} - \nu\mathbf{X}_p] \mod \Lambda$ is assumed to be uniform in the Voronoi region of Λ. This assumption can be justified by making two considerations:

- Although the decoder knows the value of $\mathbf{d}$, since he/she requires it to perform the predecoding (11), this knowledge is disregarded in the decoding stage in order to reduce its complexity. In this case, the distribution of $\mathbf{D}$ is assumed to be uniform over the fundamental Voronoi region of Λ, yielding an identical distribution for $\mathbf{W}_p$[§].
- The decoder knows the value of $\mathbf{d}$, and he/she tries to use this knowledge in both the predecoding and decoding stages. Nevertheless, the pdf of $\nu\mathbf{X}$ is so smooth, that $(\nu\mathbf{X}) \mod \Lambda$ is almost uniform in the fundamental Voronoi region of Λ. This gives again a uniform distribution for $\mathbf{W}_p$, and it is reasonable in in data hiding applications, where the power of the host signal is orders of magnitude larger than that of the watermark.

The component $(1-\nu)\mathbf{W}_p$ is usually termed as self-noise, since it is caused by the watermarking process itself due to the distortion compensation. As it is well-known, performance improvements are obtained by using $\nu < 1$, i.e., allowing a certain degree of self-noise.

Finally, the decoder decision will only rely on the observation of $\mathbf{Z}'_p$, using some criteria, as ML-lattice decoding or Euclidean (minimum distance) lattice decoding.[3] Therefore, the decoder has not to perform an exhaustive search but he/she has only to look for the most suitable symbol in the Voronoi region of Λ.

4. CLASSICAL APPROACHES.

Once the general framework for decoding has been introduced in the previous section, we will show how classical approaches fit in this framework. These methods are typically based on the use of scalar quantizers instead of a vectorial one, and the differences among them are given by their values of $\mathbf{S}$, $\mathbf{T}$, $f(\cdot)$ and $h(\cdot,\cdot)$.

4.1. Repetition coding with no projection.

In this case, the following values are taken:

$$\mathbf{S} = \text{diag}(1/\Delta_1, \cdots, 1/\Delta_L), \tag{15}$$
$$\mathbf{T} = \text{diag}(\Delta_1, \cdots, \Delta_L), \tag{16}$$
$$c_j = b_i, \quad (i-1)L/M_b < j \leq iL/M_b, \text{ and } 1 \leq i \leq M_b, \tag{17}$$
$$v_j = c_j, \quad 1 \leq j \leq L, \tag{18}$$

in such a way that any bit b_j is repeated L/M_b times[¶], so $M_c = L = P$. Note also that given (18), $\mathbf{v}$ will not depend on $\mathbf{X}_p$ but only on $\mathbf{c}$, since a scalar quantizer is being used.

Summarizing, the initial values of $\mathbf{X}$ are normalized by the corresponding quantization step $\mathbf{\Delta}$ in order to take into account the perceptual constraints in the embedding, and the input bits are repeated L/M_b times.

4.2. Repetition coding with projection.

For this method, we have:

$$s_{ij} = \begin{cases} \frac{M_b}{L}\frac{q_j}{\Delta_j}, & \text{if } (i-1)L/M_b < j \leq iL/M_b, \text{ and } 1 \leq i \leq M_b \\ 0, & \text{otherwise} \end{cases}, \tag{19}$$

$$t_{ij} = \begin{cases} q_i\Delta_i, & \text{if } (j-1)L/M_b < i \leq jL/M_b, \text{ and } 1 \leq j \leq M_b \\ 0, & \text{otherwise} \end{cases}, \tag{20}$$

$$c_j = b_j, \quad 1 \leq j \leq L, \tag{21}$$
$$v_j = c_j, \quad 1 \leq j \leq L, \tag{22}$$

[§]In the paper by Erez and Zamir[3] this simplification does not imply any loss in performance, since capacity can be achieved even when $\mathbf{W}_p$ is assumed to be uniform. Nevertheless, this is only possible for asymptotically spherical lattices.

[¶]We will assume that L/M_b is an integer.

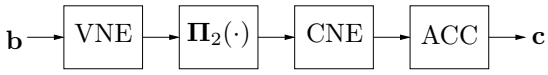


Figure 3. Structure of $f(\cdot)$ for Erez and Ten Brink's scheme.

with $q_j \in \{-1, +1\}$ a pseudorandomly generated spreading sequence, known to both encoder and decoder. Note that $\mathbf{q}$ could follow any other zero-mean unit-variance distribution (e.g., a Gaussian). The definition of $\mathbf{T}$ is also based on perceptual constraints. The fact of not having a vector quantizer, but a scalar one is again reflected in (22).

Be aware that both of these methods could be seen as extreme cases of a general one, where the repetition rate L/M_b is achieved by a first step which projects from L dimensions to L' and then a repetition channel code which transforms the M_b bits into L'. Nevertheless, the optimal value for L' when all the samples are idependent and identically distributed is $L' = M_b$, as was shown in.[5]

4.3. Channel coding with no projection

In this case, we have:

$$\begin{aligned} \mathbf{S} &= \text{diag}(1/\Delta_1, \cdots, 1/\Delta_L), && (23)\\ \mathbf{T} &= \text{diag}(\Delta_1, \cdots, \Delta_L), && (24)\\ \mathbf{c} &= f(\mathbf{b}), && (25)\\ v_j &= c_j. \quad 1 \le j \le L, && (26) \end{aligned}$$

As it can be clearly seen, repetition coding without projection is just a particular case of the previous methods. Nevertheless, it is interesting to address it separately due to its practical importance. In practical situations $f(\cdot)$ could be any kind of channel code: turbo,[6,7] serially concatenated,[8] block,[9] LDPC,[10] etc.

4.4. Channel coding with projection

Finally a last alternative could be:

$$\begin{aligned} s_{ij} &= \begin{cases} \frac{M_c}{L}\frac{q_j}{\Delta_j}, & \text{if } (i-1)L/M_c < j \le iL/M_c, \text{ and } 1 \le i \le M_c \\ 0, & \text{otherwise} \end{cases}, && (27)\\ t_{ij} &= \begin{cases} q_i\Delta_i, & \text{if } (j-1)L/M_c < i \le jL/M_c, \text{ and } 1 \le j \le M_c \\ 0, & \text{otherwise} \end{cases}, && (28)\\ \mathbf{c} &= f(\mathbf{b}), && (29)\\ v_j &= c_j, \quad 1 \le j \le L, && (30) \end{aligned}$$

where the same comments made in Sect.4.3 are still valid.

5. EREZ AND TEN BRINK'S APPROACH.

Erez and ten Brink's scheme[11] can be regarded to as one of the foremost existing dirty paper codes, which to the best of our knowledge has not been applied yet in data hiding scenarios.

It consists of a check-biregular, repeat-irregular nonsystematic repeat-accumulate code concatenated with a vector quantizer. In other words, the encoder is composed of a variable node encoder (VNE), which is nothing but a variable-rate repetition encoder, whose output is permuted using $\mathbf{\Pi}_2(\cdot)$ to become the input of a check node encoder (CNE), which is a single parity check encoder. The variable node encoder has 64.36% of the nodes of degree 3, 31.24% degree 10 and 4.4% of degree 76. 80% of the check nodes have degree 1 and 20% degree 3. The concatenation of both of them, yields a total rate 1/6. The bits in the output of this check node encoder go through a recursive accumulator (ACC). All the variable node encoder, the permuter, the check node encoder

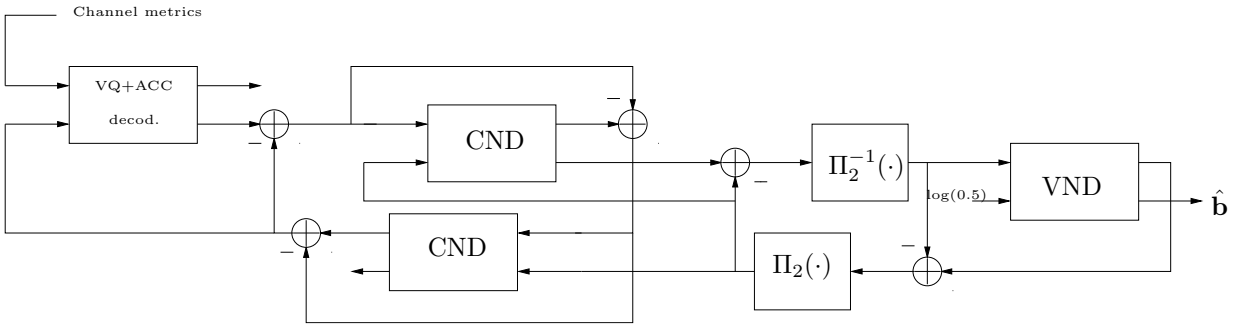


Figure 4. Structure of Erez and ten Brink's decoder. The lower input and output to the decoding blocks are the *a priori* and *a posteriori* log-probabilities of the input of the corresponding encoding block. The upper ones are the same probabilities, but corresponding to the output of the encoding block.

and the recursive accumulator can be seen as a channel code $f(\cdot)$ and its output $\mathbf{c}$ constitutes the input of a vector quantizer, that will be explained in Sect. 5.1. The structure of $f(\cdot)$ for this scheme is plotted in Fig. 3.

This quantizer finds that centroid of a lattice (which depends on the input bits) which minimizes the distortion between the side information ($\mathbf{X}$) and the output signal ($\mathbf{Y}$). This distortion measure can be changed depending on the requirements of our system, although for Erez and ten Brink's paper the Euclidean distance between both signals is employed. The search of this centroid implies using a Viterbi algorithm, so the embedding process is computationally much more expensive than for turbo-codes. In the data-hiding problem, a typical choice for the distortion measure could be a perceptual measure, which will obviously depend on the nature of the host signal. For example, when $\mathbf{X}$ is the 8×8 block-wise DCT of an image, the perceptual measure by Watson could be used.[12] Other alternative could be a weighted Euclidean distance, which normalizes the distortion in each dimension by the perceptual mask (α). In our implementation, we have followed the last strategy for the sake of simplicity.

Another problem to be solved is how to increase the redundancy for a fixed structure (which implies a fixed rate) of the channel code and vector quantizer. The solution we have adopted is based on projecting the initial vector $\mathbf{X}$ onto a lower-dimensional space (using $\mathbf{S}$). In this way the SNR per dimension will be increased in average by L/P.

As a consequence of the previous discussion, we can write

$$s_{ij} = \begin{cases} \frac{P}{L}\frac{q_j}{\Delta_j}, & \text{if } (i-1)L/P < j \leq iL/P, \text{ and } 1 \leq i \leq P \\ 0, & \text{otherwise} \end{cases}, \tag{31}$$

$$t_{ij} = \begin{cases} q_i\Delta_i, & \text{if } (j-1)L/P < i \leq jL/P, \text{ and } 1 \leq j \leq P \\ 0, & \text{otherwise} \end{cases}, \tag{32}$$

The decoding is carried out by the iterative decoding of three blocks: vector quantizer and accumulator (VQ + ACC), check node decoder (CND), and variable node decoder (VND). In exchange for this increase in complexity, significant performance gains can be achieved, as it is shown in Sect. 6. Fig. 4 shows the structure of the decoder.

5.1. Vector Quantizer.

The vector quantizer proposed by Erez and ten Brink[11] (see Fig. 5) groups the bits into triplets. One bit per triplet is duplicated and combined with the output of a non-systematic convolutional code with feedforward polynomials 07_8 and 05_8, whose input are the *virtual bits*. These *virtual bits* are not information bits but a tool to shape the quantization region of the vector quantizer; they can be arbitrarily flipped and give a degree of freedom to modify the watermark in such a way that a distortion measure between the original host signal and the watermarked one is minimized. The presence of these *virtual bits* is what accounts for the difference

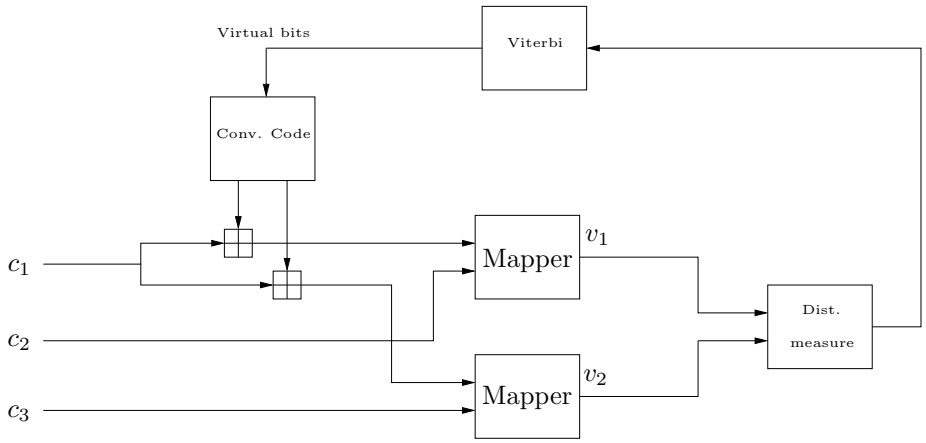


Figure 5. Structure of Erez and ten Brink's vector quantizer.

between a scalar quantizer and a vectorial one. The optimal sequence of virtual bits, i.e. that minimizing the target distortion measure, is computed using a Viterbi algorithm, and the resulting output is combined with the information coded bits, yielding 4 bits which are used to index two 4-PAM symbols (or, equivalently, a 16-QAM symbol) with alphabet $\mathcal{V} = \{-3/2, -1/2, 1/2, 3/2\}$, obtaining $\mathbf{v}$, wich is used in (5) to get $\mathbf{W}_p$. Moreover, $\mathbf{v}$ is taken into account to measure the distortion, which is used by a Viterbi algorithm to determine the optimal virtual bits sequence. Bearing this structure in mind, the total rate of the scheme is 1/4.

This vector quantizing resembles the method proposed by Miller et al.,[13] since both of them try to find a watermark which minimizes a distortion measure taking into account all the components of the watermark. Nevertheless, the differences are evident: Miller et al.'s method is based on trellis coding, while the search of the optimal watermark is more systematic in Erez and ten Brink's scheme.

6. EXPERIMENTAL RESULTS.

For the experimental part of this paper we have watermarked *Lena* 256×256 in the mid-frequencies of the 8×8-DCT domain,[14] using a perceptual mask based on Watson's distortion.[12] In all experiments each information bit was hidden in 20 coefficients, giving a total payload of $1,122$ bits. The channel-noise was chosen to be Gaussian with the same power in all the coefficients (i.i.d.). In order to address a real scenario, a value of ν was set for each experiment and hold constant for the entire range of WNR's.

First of all, we have compared the repetition coding schemes, both with and without projecting. The values of ν were 0.5 and 0.9 respectively. This difference is due to the different SNR per dimension in each scheme, since the optimal ν in the first case is computed by taking into account the SNR in the projected domain, which is increased by $10 \log_{10}$ of the projection factor. In Fig. 6 the improvement due to projecting is shown. Both of them were decoding using Maximum Likelihood (ML) lattice decoding.[3]

In order to compare dirty paper schemes which use repetition coding with those using channel coding, we have chosen a serially concatenated code proposed by Benedetto et al.[8] with outer code $G_o(D) = [1 + D, 1 + D + D^3]$ and inner $G_i(D) = [1, (1+D+D^3)/(1+D)]$, giving a total rate 1/4, which is used with projection of rate 1/5. In Fig. 7(a) the turbo-cliff of this code for dirty paper coding when all the components are i.i.d. is shown. Fig. 7(a) also shows the turbocliff of Erez and ten Brink's scheme for the same scenario. In the paper by Erez and ten Brink[11] the turbocliff was at WNR $= -1.9$ dB (1.93 dB from capacity limit) or, equivalently, $E_b/N_o = 1.1$ dB, so taking into account the increase in the WNR due to the projection, one would expect such turbocliff to show up at -8.9 dB (2.55 dB from capacity limit). Nevertheless, Fig. 7(a) (where we use $\nu = 0.42$) shows it around -7.8 dB. In fact, we can decompose the gap to the capacity limit (3.64 dB) in a gap due to the method itself (1.92 dB), other part due to projecting instead of using a more sophisticated code (0.63 dB), and finally the part

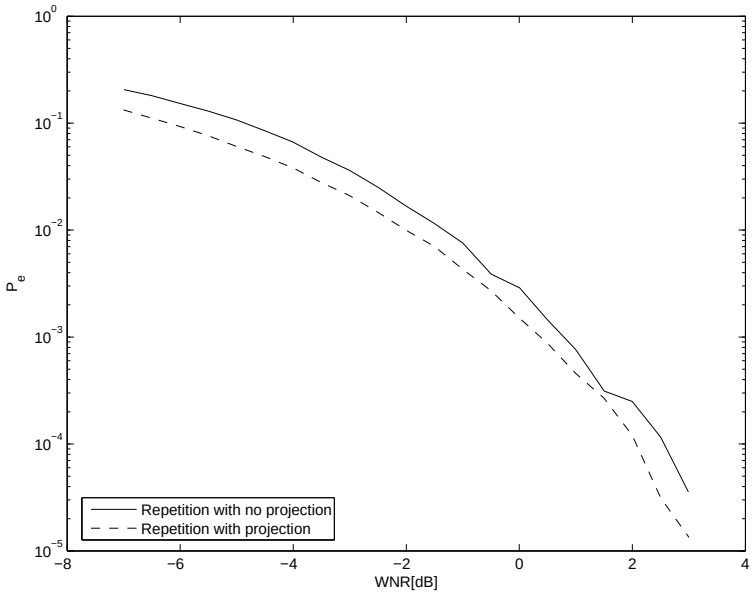
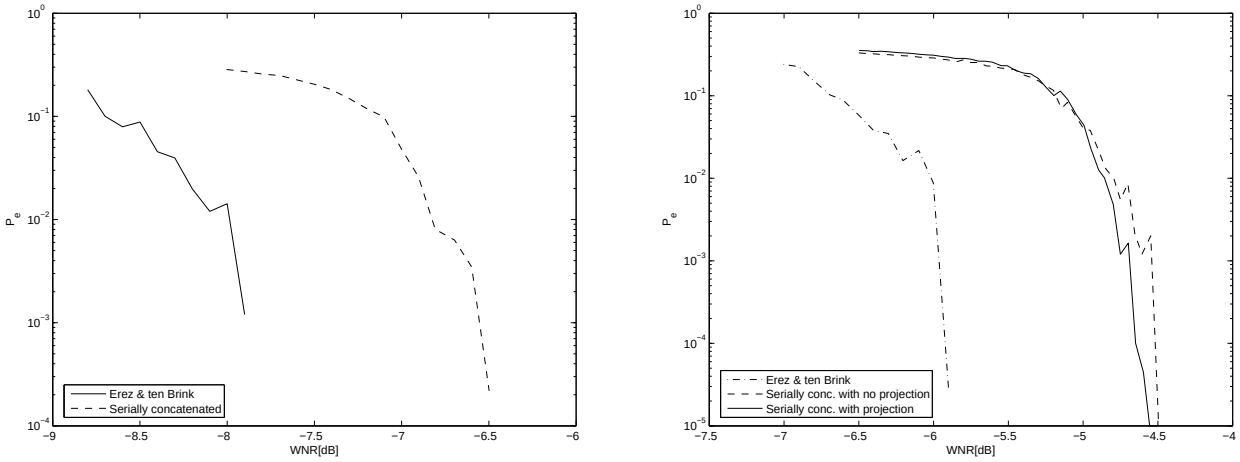


Figure 6. Comparison of repetition coding with and with no projection. $L/M_b = 20$.



(a) (b)

Figure 7. (a): Comparison between a serially-concatenated code concatenated ($\nu = 0.5$) with a scalar quantizer and Erez and ten Brink's scheme ($\nu = 0.42$) when the noise components are i.i.d.. $M_b/M_c = 1/4$ and $M_c/L = 1/5$ with projection.(b): Comparison between a serially-concatenated code concatenated with projection ($\nu = 0.6$) and with no projection ($\nu = 0.3$) with a scalar quantizer, and Erez and ten Brink's scheme ($\nu = 0.415$) when the noise components after normalizing by the perceptual mask are independent but not identically distributed. $M_b/M_c = 1/4$ and $M_c/L = 1/5$.

corresponding to the use of a limited-size permuter (1.09 dB). In any case, the gain achieved by using Erez and ten Brink's scheme compared with the serially concatenated codes is around 1.3 dB, see Fig. 7(a).

Fig. 7(b) shows the results when noise samples are Gaussian and independent but not identically distributed. The gain by using Erez and ten Brink's scheme is still around 1.5 dB, but both plots are now shifted almost 2 dB to the right, so the turbocliffs are found now at -5.8 dB and -4.5 dB. Finally, it is interesting to remark that the gain due to projecting when the serially concatenated codes are used, is almost negligible, as can be seen in Fig. 7(b).

7. CONCLUSIONS.

In this paper we have proposed a framework that encompasses many side-informed methods with coding for data-hiding, and reviewed state-of-the-art methods, specifying two possible ways to increase the operating SNR: repetition coding with and without projection. Moreover, we have introduced for what we believe is the first time in watermarking a capacity-approaching dirty-paper scheme by Erez and ten Brink. The gap to capacity of this scheme is measured for Gaussian i.i.d. noise, showing the different causes of this loss. Experimental results comparing the performance of that scheme with serially concatenated codes and repetition, with and without projection, are also introduced for non-i.i.d. noise, showing again a similar improvement when the new scheme is used.

Further research lines include a varying distortion compensating parameter ν, replacing it by its optimal value, i.e., the Wiener filter. This would improve the system performance in the case of Gaussian, independent but not identically distributed noise. Another challenge is the design of similar methods to those introduced here, by specifying different repetition and checking rates for the VNE and CNE repectively, in order to use them in scenarios with lower SNR's, instead of increasing the operating SNR through projection.

ACKNOWLEDGMENTS

Thanks are due to Dr. T. Kalker for some enlightening discussions about Erez and ten Brink's scheme.

REFERENCES

1. M. H. Costa, "Writing on dirty paper," *IEEE Trans. on Information Theory* **29**, pp. 439–441, May 1983.
2. B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. on Information Theory* **47**, pp. 1423–1443, May 2001.
3. U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log(1+\text{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Transactions on Information Theory* **50**, pp. 2293–2314, October 2004.
4. W. Yu, A. Sutivong, D. Julian, T. Cover, and M. Chiang, "Writing on colored paper," in *Proceedings IEEE International Symposium on Information Theory*, p. 302, IEEE, June 2001.
5. F. Pérez-González, F. Balado, and J. R. Hernández, "Performance analysis of existing and new methods for data hiding with known-host information in additive channels," *IEEE Trans. on Signal Processing* **51**, pp. 960–980, April 2003. Special Issue "Signal Processing for Data Hiding in Digital Media & Secure Content Delivery".
6. C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. on Communications*, pp. 1064–1070, May 1993.
7. J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory* **42**, pp. 429–445, March 1996.
8. S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Transactions on Information Theory* **44**, pp. 909–926, May 1998.
9. S. Lin and D. Costello, *Error Control Coding: Fundamentals and Applications*, Prentice-Hall, 2004.
10. R. G. Gallager, *Low-Density Parity-Check Codes*. PhD thesis, M.I.T., 1963.
11. U. Erez and S. ten Brink, "Approaching the dirty paper limit for cancelling known interference," October 2003. 41th Ann. Allerton Conf. on Commun., Control, and Computing.
12. A. B. Watson, "Dct quantization matrices visually optimized for individual images," in *Proceedings of the SPIE*, 1993. in Human Vision, Visual Processing and Digital Display III.
13. M. L. Miller, G. J. Dorr, and I. J. Cox, "Applying informed coding and embedding to design a robust high-capacity watermarking," *IEEE Transactions on Image Processing* **13**, pp. 792–807, June 2004.
14. J. R. Hernández, M. Amado, and F. Pérez-González, "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE Trans. on ImageProcessing* **9**, pp. 55–68, January 2000. Special Issue on Image and Video Processing for Digital Libraries.