JOINT SOURCE-CHANNEL CODING FOR LATTICE WATERMARKING

Abdellatif ZAIDI and Pierre DUHAMEL

Laboratoire des Signaux et Systèmes LSS/CNRS, SUPELEC
Plateau de Moulon, 3 rue Joliot-Curie - 91190 Gif sur Yvette - FRANCE
phone: + (33) 1 69 85 17 61, fax: + (33) 1 69 85 17 65, email: {zaidi,pierre.duhamel}@lss.supelec.fr
web: www.lss.supelec.fr

ABSTRACT

Coset-based codes are often proposed as an alternative to the theoretical probabilistic random binning in network coding. In this paper nested lattice codes recently proposed for multiterminal binning are used to devise a structured high dimensional Costa scheme for data hiding. The resulting embedding scheme overcomes both the famous Scalar Costa Scheme (SCS) and regular Quantization Index Modulation (QIM). Performances are studied within the context of a Modulo Lattice Additive Noise (MLAN) channel. We first show that the gap to the full AWGN capacity can be partially bridged up using some finite-dimension lattices with good packing properties. Next we use a binning interpretation to argue that information embedding can also be understood as a source-channel coding problem and that nested lattices provide means of constructing efficient low complexity, good source-channel codes. The resulting paradigm connects information embedding to the two rich area of source and channel coding and gives insights -through an example- into the construction of fine/coarse lattices. For illustrations purposes, Monte-Carlo integration-based capacity and simulation-based bit-error rate (BER) curves are provided.

1. INTRODUCTION

Consider the channel $\mathbf{Y} = \mathbf{X} + \mathbf{S} + \mathbf{V}$, where $\mathbf{X}$ and $\mathbf{Y}$ are the channel input and output respectively, $\mathbf{V}$ is an unknown additive noise and $\mathbf{S}$ is an interference known to the transmitter but not the receiver. Coding for such a channel is commonly known as "channel coding with state information available only at the transmitter" and it dates back to Gel'fand and Pinsker [1] and Heegard and El Gamal [2]. They showed that in case of a random state vector $\mathbf{S} = (S_1, \ldots, S_n)$ with independent and identically distributed (i.i.d) components and when the encoder chooses the entire state vector before choosing the channel input $\mathbf{X}$, capacity writes

$$C = \sup_{p(\mathbf{u},\mathbf{x}|\mathbf{s})} \big\{ I(\mathbf{U};\mathbf{Y}) - I(\mathbf{U};\mathbf{S}) \big\}. \quad (1)$$

$\mathbf{U}$ is an auxiliary random variable (codebook) chosen so that $\mathbf{U} \to (\mathbf{X},\mathbf{S}) \to \mathbf{Y}$ form a Markov Chain and $p(\mathbf{u},\mathbf{x}|\mathbf{s}) = \delta\big(\mathbf{x} - f(\mathbf{u},\mathbf{s})\big)p(\mathbf{u}|\mathbf{s})$. Costa, in his "Writing on Dirty Paper" [3], adhering to Gelfand-Pinsker setting [1], considered the special case of Gaussian signals[1] and showed that if $\mathbf{S}$ and $\mathbf{V}$ are statistically independent Gaussian variables with $\frac{1}{n}\mathbb{E}_{\mathbf{X}}[\mathbf{X}^2] = P$ and $\frac{1}{n}\mathbb{E}_{\mathbf{V}}[\mathbf{V}^2] = N$, then the capacity (1) is given by

$$C = \frac{1}{2}\log\left(1 + \frac{P}{N}\right). \quad (2)$$

Thus the effect of the interference $\mathbf{S}$ is canceled out completely, as if either it were zero or it were available also at the receiver. The proof is based on the random binning argument for general channels with state information [1]. The idea of "binning" is a key element in the solutions of information network problems. A binning scheme divides a set of codewords into random subsets or "bins" in such a way that the codewords in each subset are as far apart as possible. However, this probabilistic construction is convenient only for theoretical analysis, not for practical applications. For these, structured low-complexity codebooks have to be found. Recently, in a watermarking context, Chen and Wornell [4] and Eggers and al. [5] designed practical quantization-based schemes to achieve this side-information capacity. These two sample-wise schemes are referred to as "Quantization Index Modulation" (QIM) and "Scalar Costa Scheme" (SCS), respectively. Independently, in a unifying framework, Zamir and al. [6] proposed nested linear/lattice codes for algebraic coding schemes for symmetric/Gaussian multiterminal communication networks. Nested lattice codes are obtained through a layered lattice based construction: a fine lattice Λ_1 together with a coarse lattice Λ_2, with $\Lambda_2 \subset \Lambda_1$. In this paper we first show that nested lattices can readily be applied for high dimensional data embedding and derive corresponding performances, illustrating the achieved gain over scalar approaches SCS and QIM. Next we use a binning interpretation to argue that lattice-based watermarking can be understood as a source-channel coding problem and that nested lattices provide a good framework for designing low complexity good source-channel codes. We use the minimum distance criterion for the selection of the fine code for high rate transmissions. The use of the minimum-distance criterion is motivated by the fact that at high Signal-to-Noise Ratio channel codes performances depend almost only on their minimum-distance. Analysis is supported by an illustrative example using a fine Reed Solomon code together with Construction A [7]. Though non-optimal in a general setting, our main contribution- at this stage- is that of showing that careful design of joint source-channel codes enables reliable transmission at relatively high transmission rates in Costa-based watermarking

The paper is organized as follows: Section 2 presents the formal statement of the watermarking problem as coding with state information available at the transmitter together with the nested encoding and decoding functions. Also, a short review of fundamental lattice properties is given. In section 3, analogy with Modulo Lattice Additive Noise (MLAN) channel is stated and capacity is derived accordingly. In section 4 a binning interpretation is given and the watermarking problem is formulated as a source-channel coding problem. Nested lattices are then used as a good framework for theoretical analysis and the minimum-distance criterion is discussed together with an illustrative example. Finally we give some concluding remarks in section 5.

2. NESTED LATTICES FOR COSTA-BASED WATERMARKING

A lattice-based transmission diagram for the Costa problem (used here for watermarking) is depicted in Fig. 1. All signals are Gaussian. An index $m \in \mathscr{M}$ with $\mathscr{M} = \{1, \ldots, M\}$ has to be sent to a receiver in n uses of some channel denoted by the *watermark channel*. m is encoded into a code vector $\mathbf{x}$ called the watermark which is added to the cover signal $\mathbf{s} \in \mathbb{R}^n$ to form the watermarked or composite signal $\mathbf{s} + \mathbf{x}$. Here, M is the greatest integer smaller than or equal to 2^{nR} and R is the transmission rate. We shall assume that the encoder and the decoder share *common*

The authors would like to thank the RNRT project SDMO for funding.

[1] In the following, this situation is referred to as the "Costa problem"

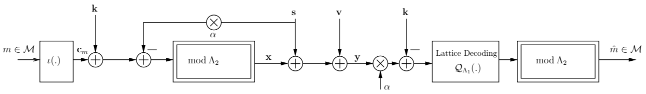

Figure 1: Nested coding scheme for watermarking. The coarse lattice Λ_2 should be a good source-code and the fine lattice Λ_1 should be a good channel code.

randomness, so that the key $\mathbf{k}$ is available to both of them. The one-dimensional SCS is obtained with the particular case of an integer lattice $\Lambda = \mathbb{Z}$ but with scaled signals $\mathbf{s}' = \mathbf{s}/\alpha$, $\mathbf{x}' = \mathbf{x}/\alpha$ and $\mathbf{v}' = \mathbf{v}/\alpha$. The inflating parameter α is chosen so as to maximize the system performances (see below). Lattices are extensively studied in [7]. In this paper, only the required ingredients are briefly reviewed. A lattice Λ of dimensionality n and with generator matrix $\mathbf{G}$ is defined as $\Lambda = \{\mathbf{G}.\mathbf{i} : \mathbf{i} \in \mathbb{Z}^n\}$. The nearest neighbor quantizer $\mathcal{Q}(.)$ associated with Λ is defined by $\mathcal{Q}(\mathbf{x}) = \mathbf{l} \in \Lambda$ if $||\mathbf{x} - \mathbf{l}|| \leq ||\mathbf{x} - \mathbf{l}'|| \, \forall \, \mathbf{l}' \in \Lambda$. The fundamental Voronoi cell of Λ is $\mathcal{V}(\Lambda) \triangleq \{\mathbf{x} : \mathcal{Q}(\mathbf{x}) = \mathbf{0}\}$. The modulo operation consists in a reduction modulo Λ: $\forall \mathbf{s} \in \mathbb{R}^n$, $\mathbf{s} \bmod \Lambda \triangleq \mathbf{s} - \mathcal{Q}_{\Lambda}(\mathbf{s}) \in \mathcal{V}(\Lambda)$.

The pair of n-dimensional lattices (fine, Λ_1, and coarse, Λ_2) of Fig. 1 is nested in the sense that each point of Λ_2 is also a point of Λ_1, i.e., $\Lambda_2 \subset \Lambda_1$. Their corresponding generator matrices satisfy $\mathbf{G}_2 = \mathbf{G}_1.\mathbf{J}$ with $\det\{\mathbf{J}\} \geq 1$. Also, the volumes V_1 of $\mathcal{V}(\Lambda_1)$ and V_2 of $\mathcal{V}(\Lambda_2)$ are such that $V_2 = det\{\mathbf{J}\}.V_1$. An important parameter is the *nesting ratio* defined as $\mu(\Lambda_1, \Lambda_2) = \sqrt[n]{V_2/V_1}$. The set of points of Λ_1 that are inside the Voronoi region of Λ_2

$$\mathscr{C}_m \triangleq \{\Lambda_1 \mathrm{mod} \Lambda_2\} = \{\Lambda_1 \cap \mathcal{V}(\Lambda_2)\}$$

forms the *coset leaders* of Λ_2 relative to Λ_1. For each $\mathbf{c} \in \mathscr{C}_m$, $\Lambda_2^{(\mathbf{c})} = \mathbf{c} + \Lambda_2$ is a coset of Λ_2 relative to Λ_1. There are $|\mathscr{C}_m| = \lfloor V_2/V_1 \rfloor$ different cosets whose union gives the fine lattice Λ_1,

$$\bigcup_{\mathbf{c} \in \mathscr{C}_m} \Lambda_2^{(\mathbf{c})} = \Lambda_1.$$

We may view these cosets as structured "bins" of the random codebook in [1]. To design a transmission scheme based on nested lattices, let $\iota(.)$ a certain indexing function that arbitrarily associates each message $m \in \mathscr{M}$ to a unique codeword $\mathbf{c}_m = \iota(m) \in \mathscr{C}_m$. Note that this implies $M \leq |\mathscr{C}_m|$, or equivalently that,

$$R \leq R^{\max} = \frac{1}{n}\log_2(\mu(\Lambda_1, \Lambda_2)). \tag{3}$$

Nested encoding and decoding are performed according to

$$\mathbf{x}(\mathbf{s}; m, \Lambda_1, \Lambda_2) = (\mathbf{c}_m + \mathbf{k} - \alpha\mathbf{s}) \bmod \Lambda_2, \tag{4a}$$

$$\hat{\mathbf{c}_m}(\mathbf{y}; m, \Lambda_1, \Lambda_2) = \mathcal{Q}_{\Lambda_1}(\alpha\mathbf{y} - \mathbf{k}) \bmod \Lambda_2. \tag{4b}$$

Eq. (4a) means that the transmitted signal is the error quantization between $\alpha\mathbf{s} - \mathbf{k}$ and the selected coset $\Lambda_2^{(\mathbf{c}_m)}$. Eq. (4b) means that the overall decoding is performed through successive (layered) decoding: first, use the fine lattice Λ_1 to find the quantizer representative $\mathcal{Q}_{\Lambda_1}(\alpha\mathbf{y} - \mathbf{k})$ of $\alpha\mathbf{y} - \mathbf{k}$. Next, use the coarse lattice Λ_2 to quantize $\mathcal{Q}_{\Lambda_1}(\alpha\mathbf{y} - \mathbf{k})$ and reconstruct the message as the index of the unique coset containing $\mathcal{Q}_{\Lambda_1}(\alpha\mathbf{y} - \mathbf{k})$. Hence equation (4b) is equivalent to $\hat{m} = \underset{m = 1, \ldots, M}{\operatorname{argmin}} \quad ||\mathcal{Q}_1(\alpha\mathbf{y} - \mathbf{k}) \bmod \Lambda_2^{\mathbf{c}_m}||$.

3. CAPACITY ANALYSIS

Consider the channel depicted in Fig.1. Using the distributive property of the modulo operation

$$((\mathbf{x} \bmod \Lambda) + \mathbf{y}) \bmod \Lambda = (\mathbf{x} + \mathbf{y}) \bmod \Lambda, \ \forall \, (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{2n}, \tag{5}$$

Eq. (4b) can be rewritten as $\hat{\mathbf{c}_m} = \mathcal{Q}_{\Lambda_1}(\mathbf{y}') \bmod \Lambda_2$ with,

$$\begin{aligned} \mathbf{y}' &= (\alpha\mathbf{y} - \mathbf{k}) \bmod \Lambda_2 \\ &= (\mathbf{y} - (1-\alpha)\mathbf{y} - \mathbf{k}) \bmod \Lambda_2 \\ &= ((\mathbf{c}_m + \mathbf{k} - \alpha\mathbf{s}) + \mathbf{s} + \mathbf{v} - (1-\alpha)(\mathbf{x} + \mathbf{s} + \mathbf{v}) - \mathbf{k}) \bmod \Lambda_2 \\ &= (\mathbf{c}_m + \alpha\mathbf{v} - (1-\alpha)\mathbf{x}) \bmod \Lambda_2. \end{aligned} \tag{6}$$

If the key $\mathbf{K}$ is uniformly distributed over $\mathcal{V}(\Lambda_2)$, it can be used as a dither vector. Dithering is a well known capacity maximizing technique. In this case, the "inflated lattice" Lemma in [8] ensures that the equivalent noise $\tilde{\mathbf{V}} = (\alpha\mathbf{V} - (1-\alpha)\mathbf{X}) \bmod \Lambda_2$ is independent of the the input $\mathbf{C}_m$ even if the high resolution quantization assumption is violated. This is due to the fact that dithering makes $\mathbf{X}$ (almost) uniform over $\mathcal{V}(\Lambda_2)$. Thus transmission over channel in Fig.1 is equivalent to that over an MLAN channel (modulo Λ_2) with input $\mathbf{C}_m$ and noise $\tilde{\mathbf{V}}$. Consequently capacity is attained with an uniform input and it calculates (in bits per dimension) to [9]

$$C(\Lambda_1/\Lambda_2) = \max_{\alpha} \frac{1}{n}\left(\log_2 V_2 - h(\tilde{\mathbf{v}})\right) < \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right), \tag{7}$$

where $h(.)$ denotes differential entropy and the right side hand term of (7) is the full capacity $C^{\max}$ of a channel AWGN with Signal-to-Noise Ratio P/N. Note that the noise $\tilde{\mathbf{V}}$ is given by the convolution of an uniform self noise $(1-\alpha)\mathbf{X}$ and a Gaussian scaled noise $\alpha\mathbf{V}$. If $\alpha = 1$ ("Zero-Forcing" approach or *regular* QIM), $\tilde{\mathbf{V}} = \mathbf{V} \bmod \Lambda_2$ is the restriction of a Gaussian PDF over $\mathcal{V}(\Lambda_2)$. Moreover, maximizing $C(\Lambda_1/\Lambda_2)$ in (7) amounts to minimizing the noise entropy $h(\tilde{\mathbf{v}})$. A quite tight approximation is obtained by minimizing the variance $\mathrm{VAR}(\alpha\mathbf{v} - (1-\alpha)\mathbf{x}) = \alpha^2 N + (1-\alpha)^2 P$. The optimal choice for the inflating parameter is the MMSE solution $\alpha = P/(P+N)$. This corresponds to DC-QIM. In general, no closed-form expression of (7) can be derived and numerical integration is needed to evaluate the differential entropy. However, two approximate expressions can be found (i) if $\alpha = 1$, $C^{(\text{approx})} = \max\{0, \frac{1}{2}\log_2\left(\frac{V_2^2}{(2\pi eN)^n}\right)\}$, and (ii) if $\alpha = P/(P+N)$, we have $h(\tilde{\mathbf{V}}) \leq h(\alpha\mathbf{V} - (1-\alpha)\mathbf{X}) \leq \log(2\pi e\alpha N)$. A lower bound on $C(\Lambda_1/\Lambda_2)$ follows

$$C(\Lambda_1/\Lambda_2 \geq \frac{1}{n}\left(\frac{1}{2}\log(1 + P/N) - \frac{1}{2}\log(2\pi e G(\Lambda_2))\right), \tag{8}$$

where $G(\Lambda)$ is the normalized second moment of the lattice Λ defined as $G(\Lambda) = \frac{1}{n}\int_{\mathcal{V}(\Lambda)} ||\mathbf{r}||^2 / V(\Lambda)^{(2/n+1)}$. The volume V_2 in (7) characterizes the average transmit power needed to transmit the set of indexes $m \in \mathscr{M}$. With respect to the baseline cubic lattice $\mathbb{Z}^n$, the reduction in this transmission power is given by the shaping gain $\gamma_s(\Lambda_2)$ of the lattice Λ_2 given by $\gamma_s(\Lambda_2) = 1/12G(\Lambda_2)$. Replacing V_2 in (7) by its expression as a function of $\gamma_s(\Lambda_2)$, the capacity (7) writes

$$C(\Lambda_1/\Lambda_2) = \max_{\alpha} \frac{1}{2}\log_2\left(12G(\Lambda_2)V_2^{2/n}\gamma_s(\Lambda_2)\right) - \frac{1}{n}h(\tilde{\mathbf{V}}). \tag{9}$$

The n-dimensional lattices considered in this paper (coarse lattice Λ_2) are summarized in table below, together with their important parameters. These lattices are used for numerical Monte-Carlo-based integration.

Lattice	Name	n	$G(\Lambda)$	$\gamma_s(\Lambda)$
$\mathbb{Z}$	Integer Lattice	1	$\frac{1}{12}$	0.000
A_2	Hexagonal Lattice	2	$\frac{5}{36\sqrt{3}}$	0.028
D_4	$4D$ Checkerboard L.	4	0.0766	0.061
E_8	Gosset Lattice	8	0.0717	0.108

Capacity curves in bits per dimension are plotted in Fig.2. We observe that:

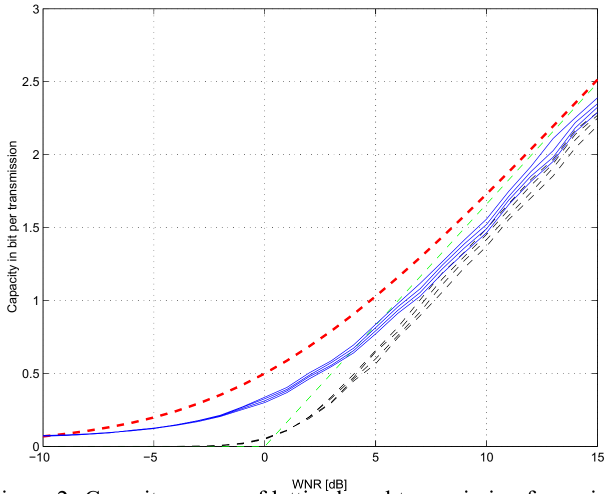

Figure 2: Capacity curves of lattice based transmission for various lattices over the Watermark-to-Noise Ratio $\text{WNR} = 10\log_{10}(P/N)$. Bottom to top: $\mathbb{Z}, A_2, D_4$ and E_8 lattices. Solid: Capacity curves of DC-QIM. Dashed: AWGN capacity and asymptotic-limit. Dashed-dotted: Capacity curves of the Zero-Forcing approach.

(i) Due to its small shaping gain, the integer lattice $\mathbb{Z}$ provides the lowest capacity. The gap to AWGN capacity is particularly large for low WNRs. At low rates (below 0.1 bit/dimension), a gap of about 4 dB is observed. At high WNR, this gap is already partially bridged up using lattices A_2, D_4 and E_8.

(ii) The improvement due to the shaping gain $\gamma_s(\Lambda)$ of the lattice is particularly visible at high rates where the shaping gain becomes significant. Convergence toward the full AWGN capacity C^{max} is such that

$$0 \leq C^{\text{max}} - C_{\Lambda} < \frac{1}{2}\log_2(2\pi e G(\Lambda)).$$

At low rates however, the shaping gain $\gamma_s(\Lambda) \approx \frac{\pi e}{6}(1-2^{-2R})$ is very small and increase in capacity is marginal.

(iii) As expected, DC-QIM with optimal lattice encoding/decoding outperforms the Zero-Forcing approach. Also, the higher the lattice dimension n, the tighter are both the lower bound (8) and the approximation C^{approx}.

4. JOINT SOURCE-CHANNEL CODING THROUGH NESTED LATTICES

From a strict functional viewpoint, the watermarking problem depicted in Fig.1 is primarily a channel-coding problem, that is, for transmitting messages. However, the "power constraint" of the input of the communication channel is the quantization error of the side information. Hence side information $\mathbf{S}$ necessitates a good source coding in order to satisfy efficiently this power constraint. In other terms, the encoding process (4a) together with the power constraint $\mathbb{E}_{\mathbf{X}}[\mathbf{X}^2] = P$ is basically a source coding problem. The only minor difference with respect to classic source coding quantization is that quantization is message-based (through a binning scheme). In addition a good quantizer would be one that, for the same transmission rate R, minimizes the quantization error (thus allowing more information at the channel input for the same input power). So, in the watermarking problem shown in Fig.1, and broadly in the more general "Costa problem", source coding is used to design channel codewords that have the appropriate energy at the input of the channel. This is ensured by grouping channel codewords into (appropriate) cosets of (appropriate) source codes.

4.1 Binning interpretation

The basic concept of combined source-channel coding in lattice-based watermarking is inherently implicit in the original random binning coding argument for channels with state information [1]. "Binning" consists in randomly dividing the codebook entries into subsets (cosets or bins) such that the codewords are far apart as possible. Hence, the set of codewords in all cosets may be viewed as a set of channel codewords. Moreover, to transmit a message $m \in \mathscr{M}$, a codeword that is distortion jointly typical with the state information $\mathbf{S}$ has to be found. This can be viewed as quantizing $\mathbf{s}$ to the nearest codeword in the bin identified by m. The set of codewords collapsed into the same bin m may then be viewed as a set of source codewords. The efficiency of this source code is measured by the distortion introduced in quantizing $\mathbf{s}$. Thus the channel coding problem of Costa-based watermarking can also be understood as a source-channel coding problem when considering that the watermark signal is obtained through quantization.

4.2 Design of practical good nested codes

Construction of good nested codes is still a challenging problem. In [6], authors tune the fine lattice Λ_1 so as to be an "exponentially good channel code". This approach is convenient for theoretical analysis but not for practical implementations. Here we use a less stringent, but more feasible approach. Namely, since the use of a simple cubic lattice $\mathbb{Z}^n$ for shaping leaves only 0.255 bit per dimension unexploited, we use $\Lambda_2 = \mathbb{Z}^n$ as a coarse lattice. Also, since the transmission rate and the error probability are obviously conflicting requirements (see [10]) and since high rates are targeted we use Construction A [7] to obtain the fine lattice Λ_1. However Construction A is a means of building a lattice from a linear code. The efficiency of this lattice naturally depends on that of the linear code. Here we ask the fine code to be "good" enough in a minimum-distance sense.

4.2.1 RS codes and minimum-distance criterion

The use of the *minimum distance* criterion is motivated by the fact that for high WNRs, channel codes performances depend almost only on their *minimum distance*. The remaining weight distribution does not much matter. So, we proceed as follows: (i) select a good fine code $\mathscr{C}(n,k)$ according to the *minimum distance* criterion, (ii) use Construction A [7] to obtain the corresponding fine lattice Λ_1 and finally (iii) use a cubic lattice as coarse lattice Λ_2. An important class of codes having good (large) *minimum-distance* is that of Reed-Solomon codes. A Reed-Solomon code $\text{RS}(N,K,D)$, $N = 2^m - 1$, is Maximum-Distance-Separable (MDS) meaning that it has the largest minimum-distance among all codes of the same dimensionality N. For instance the singleton bound is attained, i.e. $D = N - K + 1$. However the RS code being defined over a Galois-Field $GF(q)$ with $q = 2^m$, an equivalent binary representation (over $GF(2)$) should be found to use construction A. The RS code $\text{RS}(N,K,D)$ over $GF(q)$ translates to the binary code $\mathscr{C}(n,k,d)$=$\mathscr{C}(mN,mK,d)$. Note that the the minimum-distance d of the binary code $\mathscr{C}(n,k,d)$ is not explicitly related to D. A loss in the relative distance may occur when transforming an RS code into a binary code. But in most of the cases large minimum-distance D over $GF(q)$ leads to sufficiently large minimum-distance d over $GF(2)$. Thus, (binary versions of) RS codes are good candidates for building the fine lattice Λ_1 with Construction A. In addition, following Zamir et al. construction of family of codes that are asymptotically "good", RS codes represent a good starting point for a class of asymptotically (in dimension n) good channel codes. These are called *Justesen codes*. Justesen codes [11] satisfy both *Gilbert-Varshamov* lower-bound and *McEliece-Rodemich-Rumsey-Welch* upper bound. These bounds characterize channel codes for which both R and d/n remain bounded away from zero as n increases.

(a) *Gilbert-Varshamov lower-bound*: Let $\delta \in [0, \frac{1}{2}[$. There exist linear codes $\mathscr{C}(n,k,d)$ over $GF(q)$ with minimum distance d and rate k/n such that $d/n \geq \delta$ and $k/n \geq 1 - H_q(\delta) - \delta\log_q(q-1), \forall n$.

(b) *McEliece-Rodemich-Rumsey-Welch upper bound*: For each linear code $\mathscr{C}(n,k,d)$ of minimum distance d, the rate k/n is such that $k/n \leq H_2\left(\frac{1}{2}-\sqrt{\frac{d}{n}(1-\frac{d}{n})}\right)$ for n sufficiently large.

A Justesen code $\mathscr{C}(2N,2K)$ may be obtained from the RS code RS(N,K) as follows: let α be a primitive element of $GF(q)$, i.e. $\alpha^N = 1$. If $\mathbf{c} = (c_1,\ldots,c_N)$, $c_i \in GF(q)$ is an arbitrary codeword of RS(N,K), $\mathbf{c}' = (c_1,c_1,c_2,\alpha c_2,\ldots,c_N,\alpha^{N-1}c_N)$ and $\mathbf{c}''$ the corresponding binary m-tuple, the set of all codewordS $\mathbf{c}''$ for $\mathbf{c} \in$ RS(N,K) forms a Justesen code $\mathscr{C}(2nN,mK)$. The minimum distance of this Justesen code satisfies

$$\frac{d}{n} \gtrsim 0.11(1-2R). \qquad (10)$$

4.2.2 Example: RS(7,5,3)

We consider the RS code RS$(7,5,3)$ over $GF(8)$. We use the corresponding binary code $\mathscr{C}(21,15)$ for construction A. Decoding of construction A amounts to decoding the binary code $\mathscr{C}(21,15)$. We implemented a soft decision decoder based on the Euclidean distance. Given some received sequence $\mathbf{y}$, the soft decision decoder searches for the closest codeword to $\alpha\mathbf{y}-\mathbf{k}$- among the set of all $2^k = 2^{15}$ codewords of $\mathscr{C}(21,15)$. To obtain the list of all these codewords, the binary generator matrix G_{bin} of code $\mathscr{C}(21,15)$ is needed. G_{bin} is obtained as the dual of the binary parity check matrix H_{bin}, with H_{bin} the binary representation of the the parity check matrix H_q over $GF(8)$ given by

$$H_q = \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 & \alpha & \alpha^3 & \alpha^5 \end{pmatrix}.$$

In Fig.3, the per-dimension bit error probability reduction that results from the use of RS$(7,5,3)$ in the construction of the fine lattice Λ_1 is compared to that using the Gosset lattice E_8 obtained from Construction A and also transmission with deep holes of lattices $\mathbb{Z}$, A_2 and D_4. The use of lattice holes as a good channel codebook for low rate transmission is reported in [10]. It can be seen that the gain is particularly significant for low to medium WNR but may vanish for very high WNR. The reason is that the minimum distance of the fine lattice is bounded by $\min(2,\sqrt{2})$. Note that this design, corresponding to a transmission rate that is 3.75 times that of lattice D_4 and 5 times that of lattice A_2, using the RS code $(7,5,3)$, would also enable further reduction of bit error probabilities if one relaxes the transmission rate. Hence reliable transmission together with relatively high payloads are made possible.

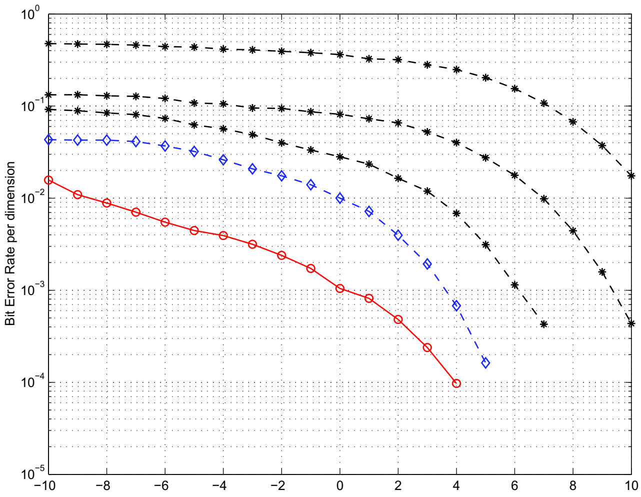

Figure 3: Bit Error Probability v.s. WNR for DC-QIM information embedding. Dashed: from bottom to top: lattices E_8 obtained from Construction A, D_4, A_2 and $\mathbb{Z}$. Solid: using the RS code $(7,5,3)$ for the design of the fine lattice with Construction A.

4.2.3 Discussion

In the example above we considered an RS code as a start point for the fine lattice Λ_1. This choice may be non-optimal, but it already shows the gain achieved when channel codewords (fine lattice points) are carefully designed. Of course more sophisticated linear/trellis codes can be used. Here, our main goal is to point out the source and channel coding problem in Watermarking and to give insights -through an example- into the proper design of the involved codes. The resulting construction has the advantage of enabling low error rates at relatively high payloads showing thus that the trade-off between error probability and transmission rate mentioned above may have good solutions. Note that the use of construction A may be non-optimal for very high embedding dimensions. Other constructions (constructions B, C, and D [7] for instance) may be used instead. The principle described here remains unchanged. These constructions have a greater complexity, however.

5. CONCLUSION

In this paper, we focused on nested lattices based information embedding techniques for data hiding. Theoretical performances have been derived through analogy with Modulo Lattice Additive Noise (MLAN) channel, illustrating the achieved gain over scalar approaches (SCS and regular QIM). Then, we use a binning interpretation to argue that lattice Costa-based watermarking can be understood as a source-channel coding problem. Interestingly, nested lattice structure turns to be useful (through decoupling of source and channel codes) for good source and channel codes design. We also, propose a simple *minimum distance*-based approach for selecting good fine channel codes. Both Monte Carlo based integration (for capacity) and simulation (for BER) are provided for illustrations.

REFERENCES

[1] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and IT.*, vol. 9, pp. 19–31, 1980.

[2] C. D. Heegard and A. A. E. Gamal, "On the capacity of computer memory with defects," *IEEE Transactions on Information Theory*, vol. IT-29, pp. 731–739, September 1983.

[3] M. H. M. Costa, "Writing on dirty papers," *IEEE Trans. on IT*, vol. IT-29, pp. 439–441, may 1983.

[4] B. Chen and G. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory*, vol. 47, pp. 1423–1443, may 2001.

[5] J. J. Eggers, R. Bäuml, R. Tzschoppe, and B. Girod, "Scalar costa scheme for information embedding," *IEEE Transactions on Signal Processing*, pp. 1–39, 2002.

[6] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Transactions on Information Theory*, vol. IT-48, pp. 1250–1276, June 2002.

[7] J. H. Conway and N. J. A. Sloane, *Sphere Packing, Lattices and Groups*, S. Verlag, Ed. New York: John Willey & Sons INC., 1988.

[8] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for cancelling known interference," in *Int. Symps. on IT and Its Applications, ISITA*, Honolulu, Hawaii, 2000, pp. 681–684.

[9] G. D. Forney, M. D. Trott, and S. Y. Chung, "Sphere-bound-achieving cosets codes and multilevel coset codes," *IEEE Trans. on IT*, vol. IT-46, pp. 820–850, 2000.

[10] A. Zaidi and P. Duhamel, "Modulo lattice additive noise channel for QIM watermarking," in *proc of Int. Conf. Image Processing ICIP*, Genova, Italy, september 2005.

[11] J. Justesen, "A class of constructive asymptotically good algebraic codes," in *PGIT*, vol. 18, 1972, pp. 652–656.