

---

**A PERFORMANCE ANALYSIS  
OF INTEGER-TO-INTEGERS TRANSFORMS  
FOR LOSSLESS CODING OF VECTORIAL SIGNALS**

---

*David Mary, Dirk T.M. Slock*

March 20, 2002

Institut EURÉCOM  
2229, Route des Crêtes, BP 193, 06904 Sophia Antipolis, France  
Tel: +33 4 9300 2916/2606; Fax: +33 4 9300 2627  
e-mail: {David.Mary,Dirk.Slock}@eurecom.fr

## ABSTRACT

An efficient lossless coding procedure should take advantage of the multichannel aspect of some data standards (such as audio standards for example). In [6], the instantaneous decorrelation of several quantized scalar signals is shown to be efficiently realized by a lossless (integer-to-integer) implementation of the Karhunen-Loeve Transform (KLT, unitary transform). The implementation of this integer-to-integer transform involves a cascade of triangular matrices and truncations. We present in this paper a lossless coding procedure based on a recently introduced decorrelating scheme (Lower Diagonal Upper factorization, LDU, causal transform). We define the lossless coding gain for a transformation as the number of bits which are saved by using the corresponding lossless coding scheme. In a first step, we analyze and compare the effects of the truncations on the coding gains for the two transformations. In a second step, we analyze the effects of estimation noise upon the coding gains : in this case, the transforms are based on an estimate  $\widehat{R_{\underline{x}^q \underline{x}^q}}$  of the covariance matrix of the quantized signals  $R_{\underline{x}^q \underline{x}^q}$ . We find that for stationary Gaussian i.i.d. signals, the coding gains are close to their maxima after a few tens of decoded vectors. Moreover, the LDU based approach is shown to yield the highest coding gain. Theoretical assertions are confronted with simulations results.

### Keywords

Lossless, Integer-to-integer, adaptive transform coding, compression, entropy, prediction, estimation, quantization, causal, unitary transformation, LDU, KLT.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Lossless Audio Coding . . . . .	4
1.2	Framework of this study . . . . .	4
1.3	Multichannel Audio . . . . .	4
1.4	Overview of this work . . . . .	5
<b>2</b>	<b>Ideal Coding Gain and Mutual Information</b>	<b>5</b>
2.1	Ideal Lossless Coding Gain . . . . .	5
2.2	Lossless Coding Gain and Mutual Information . . . . .	7
<b>3</b>	<b>Integer-to-integer transforms</b>	<b>7</b>
3.1	Integer-to-Integer implementation of the LDU . . . . .	8
3.1.1	LDU Transform Coding . . . . .	8
3.1.2	Integer-to-Integer Implementation of the LDU Transform . . . . .	8
3.2	Integer-to-Integer implementation of the KLT . . . . .	11
<b>4</b>	<b>Effects of the Estimation Noise on the Lossless Coding Gains</b>	<b>12</b>
4.1	Coding Gain for the integer-to-integer LDU . . . . .	13
4.2	Lossless Coding Gain for the integer-to-integer KLT . . . . .	14
<b>5</b>	<b>Simulations</b>	<b>14</b>
5.1	Lossless Coding Gains without estimation noise . . . . .	15
5.1.1	Results for $N = 2$ . . . . .	15
5.1.2	Position of the first signal . . . . .	18
5.1.3	Results for $N > 2$ . . . . .	18
5.2	Coding Gains with estimation noise . . . . .	20
<b>6</b>	<b>Appendix</b>	<b>22</b>
<b>7</b>	<b>References</b>	<b>22</b>

## List of Figures

1	Transform based lossless coding scheme embedded in a lossy codec. . . . .	4
2	Two entropy coding schemes. (1) : Direct entropy coding of the $x_i^q$ . (2) : Introduction of a lossless transform. . . . .	6
3	LDU transform as a vectorial DPCM coding scheme. . . . .	8
4	Lossless implementation of the LDU transform. . . . .	9
5	Lossless coding gains for integer-to-integer implementations of the LDU and KLT vs quantization stepsize. $N = 2, \rho = 0.9$ . . . . .	15
6	Percentage of the total bit rate saved by using integer-to-integer transform. $N = 2, \rho = 0.9$ . . . . .	16
7	Correlation coefficient of quantization noises vs Correlation coefficient of the variables $x_1$ and $x_2$ for several quantization stepsize. The variables $x_1$ and $x_2$ have variance 1. . . . .	17
8	Actual minus theoretic entropy for uniformly quantized Gaussian signal vs quantization stepsize. . . . .	17
9	Importance of the coarseness of the quantization of the first signal. . . . .	18
10	Lossless Coding Gain for integer-to-integer LDU with $N = 5, \Delta = 0.51$ . . . . .	19
11	Lossless Coding Gain for integer-to-integer LDU with $N = 5, \Delta = 0.21$ . . . . .	19
12	Coding Gains with estimation noise versus K for $N=2, \frac{\Delta}{\sigma_x} = 0.51$ . . . . .	20
13	Coding Gains with estimation noise versus K for $N=5, \frac{\Delta}{\sigma_x} = 0.51$ . . . . .	21
14	Coding Gains with estimation noise versus K for $N=5, \Delta = 0.21$ . . . . .	21

# 1. INTRODUCTION

## 1.1. Lossless Audio Coding

While relatively little attention has been paid in the past to lossless audio coding (mainly because it provides lower compression ratios than lossy coding), many modern important applications suggest or demand the use of a powerful lossless audio coding technique. When the coding procedure is not subject to a strong constraint of bitrate saving, as this is for example the case for Digital Versatile Disks, lossless coding obviously appears as the best technique. Besides, some applications of very high fidelity music distribution over the internet could provide lossy compressed audio clips in a first step (allowing the music lover to browse and select the desired clip in a reasonable time), and then provide a losslessly compressed version of the original. For archiving and mixing applications, lossless compression avoids the degradation of the signal when successively coded/decoded with lossy coders [2]. It can also be observed that an increasing number of companies now provide products for lossless audio compression [14]. A survey complementary to that of [2] reviewing free competitive lossless coders can be found in [13].

## 1.2. Framework of this study

An useful application of a lossless coding scheme is described in figure 1. The lossless coder is embedded in the core of

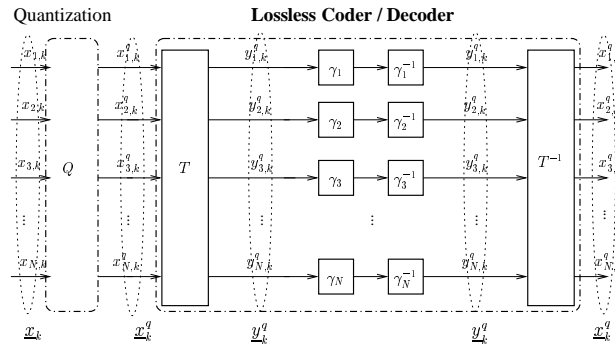


Fig. 1. Transform based lossless coding scheme embedded in a lossy codec.

a lossy encoder, whose performance may thereby be improved. In a first step, a very high resolution vectorial source  $\underline{x}$  is quantized using a lossy source codec, represented by the box  $Q$  ( $Q$  may represent the discretization realized by any lossy coder/decoder, e.g. independent uniform scalar quantizers, independent ADPCM or MPEG Audio Codecs...). The set of quantized values  $\{x_i^q\}$  obtained in the lossy coder is classically directly entropy coded using independent entropy coders  $\{\gamma_i\}$ . However, since these coefficients are *independently* entropy coded, it may be worth applying, after the quantization stage, a lossless transformation  $T$ .  $T$  aims to reduce the intra- and inter-signal dependencies, and hence, by further entropy coding the discrete transformed signals  $y_i^q$ , the total bitrate. Indeed, no additional degradation of the signals should occur thereby. Integer-to-integer implementations of optimal linear transforms in the context of transform coding can reduce intersignal dependencies, and may therefore be used in such a scheme. Two of them are reviewed in this work, the optimal causal transform (LDU) and the optimal unitary transform (KLT). The set of operations  $\{Q, T, \gamma, \gamma^{-1}, T^{-1}\}$ , which represents an "enhanced" lossy codec, constitutes the framework of this work. The field of audio coding appears as a natural space of application for these techniques, which may however be applied to the wide class of the vectorial sources. In the particular case of MPEG-4, MPEG members are now discussing issues in considering lossless audio coding as an extension to the MPEG-4 standard [14, 15].

## 1.3. Multichannel Audio

An important issue which should be taken into account in the lossless audio coding procedure is the multichannel aspect of recent audio technologies. Starting from the monophonic and stereophonic technologies, new systems (mainly due to the film industry and home entertainments systems) such as quadrasonic, 5.1 and 10.2 channels are now available. An efficient coding procedure aiming storage or transmission of these signals should, as much as possible, take advantage of

the correlations between these signals. Multichannel audio sources can be roughly classified into three categories : signals used for broadcasting, where the channels can be totally different from one to another (e.g. different audio programs in each channel, or the same program in different languages), films soundtracks (typically the format of 5.1 channels) which present a high correlation between certain channels, and finally multichannel audio sources resulting from a recording of the same scene by multiple microphones (in this case, there is indeed a great advantage to be taken from the structure of the multichannel audio signal) [4]. However, the correlations between the different channels are in most of the state of the art lossless audio coders not taken into account at all, or in a basic way only, by computing sums and differences.

#### 1.4. Overview of this work

This work only considers transformations attempting to reduce instantaneous correlations between the scalar signals  $x_i^q$ . In the next section, we derive the expression of the ideal lossless coding gain, that is, the maximum coding gain one can expect by a transform making the transformed signal independent, though preserving the whole information about the vectorial source  $\underline{x}^q$ . The third part compares the gains obtained with two approaches for lossless coding based on approximation of linear transforms, a unitary (KLT) and a causal (LDU) transform. The fourth section is dedicated to estimation noise and derives the coding gains of the two approaches when the transformations are based on an estimate of the covariance matrix. The fifth section exposes and discusses several simulations results.

## 2. IDEAL CODING GAIN AND MUTUAL INFORMATION

Consider the two coding schemes of figure (2), simplified from figure (1). In both cases, the sources  $x_i$  are quantized using stepsizes  $\Delta_i$ . In the first scheme, the resulting discrete valued scalar sources  $x_i^q$  are directly entropy coded using a set of independent scalar entropy coders  $\gamma_i$  (codewords  $i_i$  are transmitted). As stated in introduction however, the sources  $x_i$  are generally not independent, and neither are indeed their quantized versions. Thus, in order to avoid to code any redundancy, one may apply a transform  $T$  before entropy coding. The resulting discrete scalar sources  $y_i$  are further entropy coded (codewords  $i'_i$  are transmitted to the decoder). The transform  $T$  is chosen to be invertible so that the decoder can losslessly recover the data  $\{x_i^q\}$ . The lossless coding gain obtained for the transform, expressed in bits, may then be written as

$$G_T = \sum_i^N [H(x_i^q) - H(y_i^q)], \quad (1)$$

where  $H$  denotes (zeroth order) discrete entropy.

### 2.1. Ideal Lossless Coding Gain

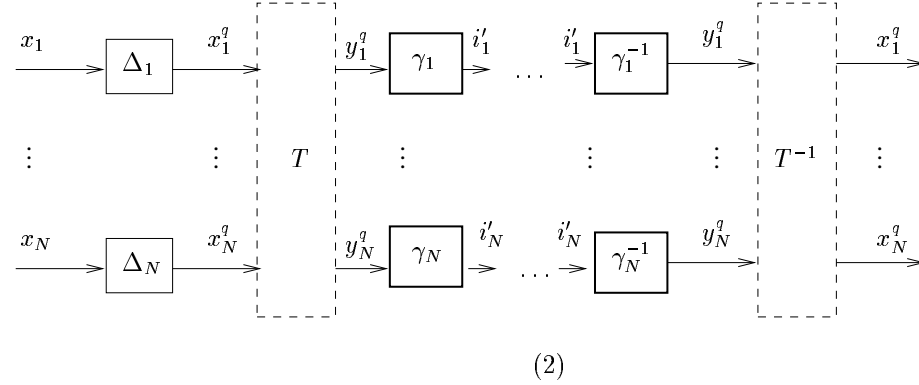
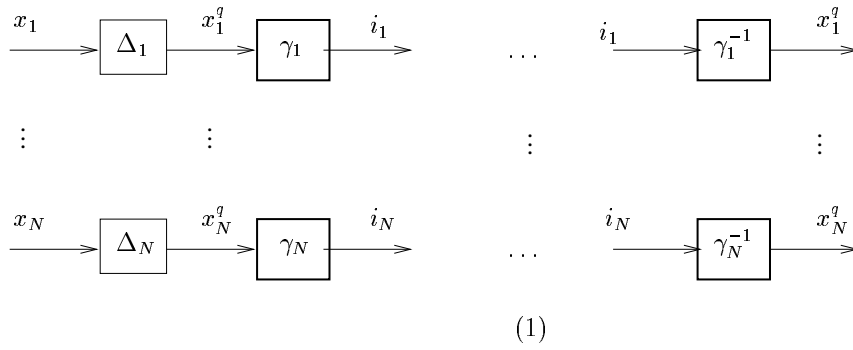
The question of the maximum coding gain  $G_T$  now arises, in other words: how many bits can we expect to save by making the transformed signals independent ?

Since the coding scheme must be lossless, the amount of information about the vectorial source  $\underline{x}^q$  conveyed to the decoder must be at least  $H(\underline{x}^q)$ . Now, ideally, the several signals  $\{y_i^q\}$  will be made independent by the transform, that is, the coding scheme will take advantage from a non redundant repartition of the whole information  $H(\underline{x}^q)$ . In this case, since an invertible transform does not change entropy, the bit rate required to independently code the  $y_i^q$  is  $\sum_{i=1}^N H(y_i^q) = H(\underline{y}^q) = H(\underline{x}^q)$ , which is also the minimum bit rate required to losslessly code the vectorial source  $\underline{x}^q$ . Now, the relation of differential to discrete entropy of the uniformly quantized sources with stepsize  $\Delta_i$  is [1]

$$H(x_i^q) + \log_2 \Delta_i \rightarrow h(x_i) \text{ as } \Delta_i \rightarrow 0. \quad (2)$$

For the N-vectorial source  $\underline{x}$ , a similar relation holds (see [6], and Appendix for a proof)

$$H(\underline{x}^q) + \sum_{i=1}^N \log_2 \Delta_i \rightarrow h(\underline{x}) \text{ as } \Delta_i \rightarrow 0, i = 1, \dots, N. \quad (3)$$



**Fig. 2.** Two entropy coding schemes. (1) : Direct entropy coding of the  $x_i^q$ . (2) : Introduction of a lossless transform.

For Gaussian random variables (r.v.s)  $x_i$ , the differential entropy  $h(x_i)$  equals  $\frac{1}{2} \log_2 2\pi e \sigma_{x_i}^2$ . It can then easily be shown that for sufficiently small quantization stepsizes  $\Delta_i$ , we obtain

$$H(\underline{x}^q) \approx \frac{1}{2} \log_2 \frac{(2\pi e)^N \det R_{\underline{x}\underline{x}}}{\prod_{i=1}^N \Delta_i^2}, \quad (4)$$

where  $R_{\underline{x}\underline{x}} = E \underline{x}\underline{x}^T$ . The maximum coding gain is then

$$\begin{aligned} G_{max} &= \sum_{i=1}^N H(x_i^q) - H(\underline{x}^q) \\ &= \sum_{i=1}^N h(x_i) - h(\underline{x}) \\ &= \frac{1}{2} \log_2 \frac{\det \text{diag}\{R_{\underline{x}\underline{x}}\}}{\det R_{\underline{x}\underline{x}}}, \end{aligned} \quad (5)$$

where  $\text{diag}\{\cdot\}$  denotes the diagonal matrix made with the diagonal elements of  $\{\cdot\}$ .

It is now shown that  $G_{max}$  is ideal because it corresponds in the Gaussian case to the gain obtained with a linear decorrelating transform placed before the quantizers. By writing  $\det R_{\underline{x}\underline{x}} = \prod_{i=1}^N \sigma_{y_i}^2 = \prod_{i=1}^N \lambda_i$ , where  $\{\sigma_{y_i}^2\}$  and  $\{\lambda_i\}$  are respectively the optimal prediction error variance of  $x_i$  based on  $\underline{x}_{1:i-1}$ , and the eigenvalues of  $R_{\underline{x}\underline{x}}$ , we can write equation (4) as

$$\begin{aligned} H(\underline{x}^q) &\approx \sum_{i=1}^N \frac{1}{2} \log_2 2\pi e \sigma_{y_i}^2 - \log_2 \Delta_i \\ &\approx \sum_{i=1}^N \frac{1}{2} \log_2 2\pi e \lambda_i - \log_2 \Delta_i. \end{aligned} \quad (6)$$

Equation (6) shows that the entropy of the vector  $\underline{x}^q$  may be written as the sum of the entropies of  $N$  independent r.v.s of variances  $\{\sigma_{y_i}^2\}$  (or  $\{\lambda_i\}$ ), quantized with quantization stepsizes  $\Delta_i$ . The required bitrate for independently coding these quantized variables is  $H(\underline{x}^q)$ : if we apply first a KLT or an LDU to the unquantized source  $\underline{x}$ , and then quantize the transformed signals with stepsizes  $\Delta_i$ , then the minimum bit rate required to entropy code these transformed signals is given by (6). Hence, the gain (5) would be obtained by a classical transform coding scheme by inverting the transformation and

quantization operations. As will be illustrated in the next section however, the performance of realizable lossless coding schemes based on approximations of linear transforms must be expected to be lower than the expression (5): indeed, since the transform is placed after the quantizers and just before scalar entropy coders, its output should be discrete valued, which is not the case for optimal linear decorrelating transforms. Thus, truncations are necessary which increase the entropy of the transform signals :  $\sum_i H(y_i^q)$  is generally greater than  $H(\underline{x}^q)$ .

## 2.2. Lossless Coding Gain and Mutual Information

We now show that the ideal lossless coding gain (5) can be easily related to the mutual information between quantized r.v.s  $\{x_i^q\}$  in the Gaussian case.

Suppose we dispose of a set of  $i - 1$  quantized scalar sources  $x_j^q, j = 1 \dots i - 1$ , and that we wish to code an  $i - th$  source  $x_i^q$ , which is not independent from the  $i - 1$  others. Intuitively, the best strategy would be to code the only information contained in the  $i - th$  r.v. which is not shared with the  $i - 1$  previous variables. The mutual information  $I(x_i^q; \underline{x}_{1:i-1}^q)$  allows one to quantize this idea : it represents the amount of information that the r.v.  $x_i^q$  shares with the  $i - 1$  others (vector  $\underline{x}_{1:i-1}^q$ ), and is defined by

$$I(x_i^q; \underline{x}_{1:i-1}^q) = H(x_i^q) + H(\underline{x}_{1:i-1}^q) - H(x_i^q, \underline{x}_{1:i-1}^q) = H(x_i^q) + H(\underline{x}_{1:i-1}^q) - H(\underline{x}_{1:i}^q). \quad (7)$$

By writing the expressions of the mutual information between  $x_i^q$  and  $\underline{x}_{1:i-1}^q$  for  $i = 2, \dots, N$ , we obtain

$$\begin{aligned} I(x_2^q; x_1^q) &= H(x_2^q) + H(x_1^q) - H(\underline{x}_{1:2}^q) \\ I(x_3^q; \underline{x}_{1:2}^q) &= H(x_3^q) + H(\underline{x}_{1:2}^q) - H(\underline{x}_{1:3}^q) \\ &\vdots \\ I(x_{N-1}^q; \underline{x}_{1:N-2}^q) &= H(x_{N-1}^q) + H(\underline{x}_{1:N-2}^q) - H(\underline{x}_{1:N-1}^q) \\ I(x_N^q; \underline{x}_{1:N-1}^q) &= H(x_N^q) + H(\underline{x}_{1:N-1}^q) - H(\underline{x}_{1:N}^q). \end{aligned} \quad (8)$$

Then by summing the previous expressions, we get

$$\begin{aligned} \sum_{i=2}^N I(x_i^q; \underline{x}_{1:i-1}^q) &= \sum_{i=1}^N H(x_i^q) - H(\underline{x}^q) \\ &= \sum_{i=1}^N h(x) - h(\underline{x}) \\ &= G_{max}. \end{aligned} \quad (9)$$

Thus, the maximum bitrate that can be saved using a lossless coding scheme corresponds to the sum of the mutual information shared between each new random variable and the previous ones.

Note also from (9) that high resolution quantizing does not change mutual information, which comes from the whiteness and independence of the quantization noises.

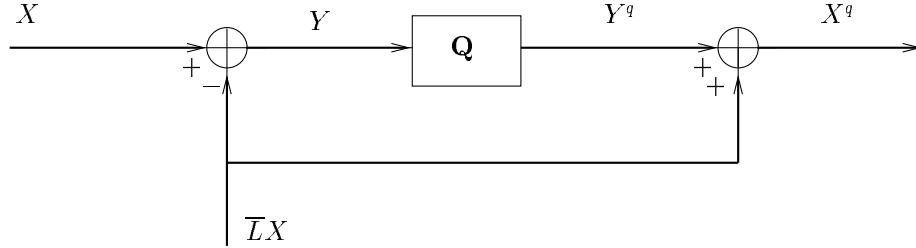
## 3. INTEGER-TO-INTEGERS TRANSFORMS

Suppose one disposes of  $N$  quantized scalar signals  $x_i^q$ . Each one of this source is a quantized version of  $x_i$  to the nearest multiple of  $\Delta_i$  (denoted by  $[\cdot]_{\Delta_i}$ ), and takes values in the set  $\Delta_i \mathbb{Z}$ :  $\underline{x}_{i,k}^q = [x_{1,k}^q, x_{2,k}^q, \dots, x_{n,k}^q]^T = \left[ [x_{1,k}]_{\Delta_1}, [x_{2,k}^q]_{\Delta_2}, \dots, [x_{n,k}^q]_{\Delta_N} \right]^T$ . An integer-to-integer transform  $T: \Delta_1 \mathbb{Z} \times \Delta_2 \mathbb{Z} \dots \times \Delta_N \mathbb{Z} \rightarrow \Delta_1 \mathbb{Z} \times \Delta_2 \mathbb{Z} \dots \times \Delta_N \mathbb{Z}$  associates to each quantized  $N$ -vector  $\underline{x}_{i,k}^q$  an  $N$ -vector  $\underline{y}_{i,k}^q = T \underline{x}_{i,k}^q$  whose components  $y_i^q$  are quantized to the same resolution  $\Delta_i$  as the corresponding  $x_i^q$ . The transformation is chosen to be invertible so that the decoder can losslessly compute the original data by  $\underline{x}_{i,k}^q = T^{-1} \underline{y}_{i,k}^q$ . Since the aim of the transform  $T$  is to make the transform signals independent, the problem of its design is very similar to that of designing the best transformation in a transform coding framework. The transform  $T$  can be chosen to approximate linear decorrelating transforms such as the LDU or the KLT, which are optimal for Gaussian signals in the classical transform coding case [10, 9]. However, since the transform must be integer-to-integer,  $T$  is only an approximation of the chosen linear decorrelating transform. Although both integer-to-integer implementations tend to the maximum gain of expression (5) for quantization stepsizes arbitrarily small, a quantifiable loss in performance occurs in practical coding situations whose effects are analyzed in the following.

### 3.1. Integer-to-Integer implementation of the LDU

#### 3.1.1. LDU Transform Coding

We first recall the optimal causal transform in the case of transform coding. Consider a stationary Gaussian vectorial source  $\{\underline{x}\}$ . This source may be composed of any scalar sources  $\{x_i\}$ . In the classical transform coding framework, a linear transformation  $T$  is applied to each N-vector  $\underline{x}_k$  to produce an N-vector  $\underline{y}_k = T\underline{x}_k$  whose components are independently quantized using scalar quantizers  $Q_i$ . A number of bits  $r_i$  is attributed to each  $Q_i$  under the constraint  $\sum_i r_i = Nr$ . In the case of the LDU transform, the transform vector  $\underline{y}_k$  is chosen to be  $\underline{y}_k = \underline{x}_k - \widehat{\underline{x}}_k = L\underline{x}_k = \underline{x}_k - \overline{L}\underline{x}_k$ , where  $\overline{L}\underline{x}_k$  is the reference vector. The output  $\underline{x}_k^q$  is  $\underline{y}_k^q + \overline{L}\underline{x}_k$ , see Figure 3. This scheme appears as a generalization to the vectorial case of the classical scalar DPCM coding scheme. As detailed in [9, 10, 11], the optimal  $L$  in terms of transform coding



**Fig. 3.** LDU transform as a vectorial DPCM coding scheme.

gain is such that  $LR_{\underline{x}\underline{x}}L^T = \text{diag}\{\sigma_{y_1}^2, \dots, \sigma_{y_N}^2\}$ , where  $\text{diag}\{\dots\}$  represents a diagonal matrix whose elements are  $\sigma_{y_i}^2$ . In other words, the components  $y_{i,k}$  are the prediction errors of  $x_{i,k}$  with respect to the past values of  $\underline{x}_k$ , the  $\underline{x}_{1:i-1,k}$ , and the optimal coefficients  $-L_{i,1:i-1}$  are the optimal prediction coefficients. Since each prediction error  $y_{i,k}$  is orthogonal to the subspaces generated by the  $\underline{x}_{1:i-1,k}$ , the  $y_{i,k}$  are orthogonal. It follows that  $R_{\underline{x}\underline{x}} = L^{-1}R_{\underline{y}\underline{y}}L^{-T}$ , which represents the LDU factorization of  $R_{\underline{x}\underline{x}}$ . Moreover, it can be shown [3] that the decorrelating and unimodularity properties of  $L$  lead to the same coding gain as the Karhunen-Loeve Transform.

#### 3.1.2. Integer-to-Integer Implementation of the LDU Transform

In a first step, the linear transform  $L^q = I - \overline{L}^q$  is optimized to decorrelate the quantized data  $x_i^q$ . Thus, we look for

$$\min_{L_{i,1:i-1}^q} L_i^q (R_{\underline{x}^q \underline{x}^q}) L_i^{qT}, \quad (10)$$

which leads to the normal equations

$$\begin{bmatrix} R_{\underline{x}^q \underline{x}^q} \end{bmatrix}_{1:i,1:i} \begin{bmatrix} L_{i,i-1}^q \\ \vdots \\ L_{i,1}^q \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \sigma_{y_i}^2 \end{bmatrix},$$

where  $\sigma_{y_i}^2$  is the optimal prediction error variance corresponding to the optimal (continuous valued) prediction error  $y'_{i,k} = x_{i,k}^q - L_{i,1:i-1}^q \underline{x}_{1:i-1,k}^q = x_{i,k}^q - \widehat{x}_{i,k}^q$ . The optimal transform vector is then  $\underline{y}'_k = \underline{x}_k^q - \overline{L}^q \underline{x}_k^q$ , and the optimal transform  $L^q$  corresponds in this case to the LDU factorization of the covariance matrix of quantized data  $R_{\underline{x}^q \underline{x}^q}$

$$R_{\underline{x}^q \underline{x}^q} = L^{q-1} R_{\underline{y}' \underline{y}'} L^{q-T}. \quad (11)$$

The second step is now to design an approximation  $\widehat{L}_{int}^q$  of  $L^q$  which allows one to keep the transform structure lossless. This can easily be realized by truncating each estimate  $\widehat{x}_{i,q}^q$  of  $x_{i,k}^q$ . Each transform coefficient is computed by

$$y_{i,k}^q = x_{i,k}^q - [\widehat{x}_{i,k}^q]_{\Delta_i} = x_{i,k}^q - [L_{i,1:i-1}^q \underline{x}_{1:i-1,k}^q]_{\Delta_i}, \quad (12)$$





of the variables  $y_i^q$ . Since the source  $x_i^q$  is discrete, we have

$$\begin{aligned} y_{i,k}^q &= x_{i,k}^q - \left[ \widehat{x}_{i,k}^q \right]_{\Delta_i} \\ &= \left[ x_{i,k}^q - \widehat{x}_{i,k}^q \right]_{\Delta_i} \\ &= \left[ y'_{i,k} \right]_{\Delta_i}. \end{aligned} \quad (16)$$

Thus, the entropy  $H(y_i^q)$  may be written as

$$H(y_i^q) \approx h(y'_i) - \log_2 \Delta_i, \quad (17)$$

which assumes a quantization noise uniformly distributed over  $[-\frac{\Delta_i}{2}, \frac{\Delta_i}{2}[$  (small quantization stepsizes). The continuous r.v.s  $y'_i$  are not strictly Gaussian since each  $y'_i$  is a linear combination of  $i$  Gaussian r.v.s and  $i - 1$  uniform r.v.s. However, since the probability density function of a sum of uniform r.v.s tends quickly to a Gaussian p.d.f., we assume that this is the case, and

$$H(y_i^q) \approx \frac{1}{2} \log_2 2\pi e \sigma_{y'_i}^2 - \log_2 \Delta_i. \quad (18)$$

Note that in the integer-to-integer implementation of the LDU, the first scalar signal remains unchanged, and only  $N - 1$  rounding operations are involved in the lossless transformation. The bit rate required to entropy code the discrete r.v.s  $y_i^q$  is then

$$E \sum_{i=1}^N H(y_i^q) \approx \frac{1}{2} \log_2 (2\pi e) \sigma_{x_1}^2 - \log_2 \Delta_1 + \sum_{i=2}^N \frac{1}{2} \log_2 (2\pi e)^{N-1} \sigma_{y'_i}^2 - \log_2 \Delta_i. \quad (19)$$

The lossless coding gain for the integer-to-integer LDU may then be written as

$$\begin{aligned} G_{L_{int}^q} &\approx \sum_{i=1}^N H(x_i^q) - H(y_i^q) \\ &= \frac{1}{2} \log_2 \frac{\prod_{i=2}^N \sigma_{x_i}^2}{\prod_{i=2}^N \sigma_{y'_i}^2} \\ G_{L_{int}^q} &\approx \frac{1}{2} \log_2 \frac{\det \text{diag}\{R_{\underline{x}\underline{x}}\}}{\sigma_{x_1}^2 \prod_{i=2}^N \sigma_{y'_i}^2}, \end{aligned} \quad (20)$$

where subscript  $L_{int}^q$  refers to the integer-to-integer implementation of  $L^q$ . The last equality shows that  $G_{L_{int}^q}$  is indeed inferior to  $G_{max}$  since the denominator involves the optimal prediction error variances obtained from  $R_{\underline{x}^q \underline{x}^q} = R_{\underline{x}\underline{x}} + D$  (where  $D$  is the diagonal of the variances of the quantization noises whose  $i$ -th entry is  $D(i, i) = \Delta_i^2/12$ ) instead of  $R_{\underline{x}\underline{x}}$ . Moreover, since  $L^q$  diagonalizes  $R_{\underline{x}^q \underline{x}^q}$ , we have  $\prod_{i=1}^N \sigma_{y'_i}^2 = \det R_{\underline{x}^q \underline{x}^q}$ , where  $\sigma_{y'_1}^2 = \sigma_{x_1}^2$ . Thus the coding gain  $G_{L_{int}^q}$  may alternatively be approximated as

$$\begin{aligned} G_{L_{int}^q} &\approx \frac{1}{2} \log_2 \frac{\det \text{diag}\{R_{\underline{x}\underline{x}}\}}{\left(\sigma_{x_1}^2 - \frac{\Delta_1^2}{12}\right) \prod_{i=2}^N \sigma_{y'_i}^2} \\ &= \frac{1}{2} \log_2 \frac{\det \text{diag}\{R_{\underline{x}\underline{x}}\}}{\det R_{\underline{x}^q \underline{x}^q}} + \frac{1}{2} \log_2 \left(1 + \frac{\Delta_1^2}{12\sigma_{x_1}^2}\right) \\ G_{L_{int}^q} &\approx \frac{1}{2} \log_2 \frac{\det \text{diag}\{R_{\underline{x}\underline{x}}\}}{\det R_{\underline{x}^q \underline{x}^q}} + \frac{\Delta_1^2}{24 \ln 2 \sigma_{x_1}^2}. \end{aligned} \quad (21)$$

This last expression shows that we should position the most coarsely quantized signal (strongest  $\frac{\Delta_i}{\sigma_{x_i}}$ ) in first position in order to maximize  $G_{L_{int}^q}$ . Moreover, one can check that  $G_{L_{int}^q}$  indeed tends to  $G_{max}$  as  $\Delta_i \rightarrow 0$ ,  $i = 1, \dots, N$ , which means that the transform is optimal in terms of lossless coding gains in the case of negligible rounding effects.

Another expression of (21) involving  $G_{max}$  may be found by writing

$$\begin{aligned} \det R_{\underline{x}^q \underline{x}^q} &= \det(R_{\underline{x}\underline{x}} + D) \\ &\approx \det R_{\underline{x}\underline{x}} \left(1 + \text{tr}\{R_{\underline{x}\underline{x}}^{-1} D\}\right), \end{aligned} \quad (22)$$

wich leads to

$$\begin{aligned}
G_{L_{int}^q} &\approx \frac{1}{2} \log_2 \frac{\det \text{diag}\{R_{\underline{x}\underline{x}}\}}{\det R_{\underline{x}\underline{x}}} - \frac{1}{2} \log_2 \left( 1 + \text{tr}\{R_{\underline{x}\underline{x}}^{-1}D\} \right) - \frac{1}{2} \log_2 \left( 1 - \frac{\Delta_1^2}{12\sigma_{x_1}^2} \right) \\
&\approx G_{max} - \frac{1}{2 \ln 2} \left[ \text{tr}\{R_{\underline{x}\underline{x}}^{-1}D\} - \frac{\Delta_1^2}{12\sigma_{x_1}^2} \right].
\end{aligned} \tag{23}$$

We should here underline the similarity between the integer-to-integer implementation of the LDU and the lossless matrixing described in [12], in which however the diagonalizing aspect of the transform (and thus its optimality for Gaussian signals in the case of negligible perturbation effects) was not established. Moreover, the pertubation effects due to truncations and estimation noise are not, to our knowledge, analyzed in their published related work.

### 3.2. Integer-to-Integer implementation of the KLT

Concerning the KLT (unitary case), the integer-to-integer approximation is based on the factorization of a unimodular matrix cascaded with rounding operations ensuring the invertibility of the global transform. In [6], this transform was shown to be equivalent to the original KLT for arbitrarily small  $\Delta_i$ . The loss in the bitrate saving which is due to the rounding operations occuring in actual coding situations was however neglected in [6], and is analyzed here for  $N = 2$ .

Let us denote by  $V^q$  the KLT computed on  $R_{\underline{x}^q \underline{x}^q}$ . We then have

$$\Lambda' = V^q R_{\underline{x}^q \underline{x}^q} V^{qT}. \tag{24}$$

We denote by  $\lambda_i^q$  the variances of the (continuous) transform signals.

We now briefly recall the construction of the integer-to-integer transform based on  $V^q$ . As any  $2 \times 2$  unimodular transform,  $V^q$  can be factored into at most three lower- and upper-triangular matrices with unit diagonal as

$$\begin{aligned}
V^q &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} = V_1^q V_2^q V_3^q, \\
V_1^q &= \begin{bmatrix} 1 & \frac{a-1}{c} \\ 0 & 1 \end{bmatrix}, V_2^q = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}, V_3^q = \begin{bmatrix} 1 & \frac{d-1}{c} \\ 0 & 1 \end{bmatrix}.
\end{aligned} \tag{25}$$

The transform vector  $\underline{y}_k$  is then losslessly obtained by using the integer-to-integer transform  $V_{int}^q$

$$\underline{y}_k^q = V_{int}^q \underline{x}_k^q = \left[ \begin{array}{c} \left[ \begin{array}{c} \left[ \begin{array}{c} \underbrace{\left[ \begin{array}{c} V_3^q \underline{x}_k^q \\ \underline{y}_k^1 \end{array} \right]}_{\Delta_1} \\ \underline{y}_k^2 \end{array} \right]}_{\Delta_2} \\ \underline{y}_k^3 \end{array} \right] \end{array} \right]_{\Delta_1}. \tag{26}$$

Since the matrices are triangular, their inverses are simply computed by changing the signs of the off-diagonal elements.

In the  $N = 2$  case, one can analyze the effects of the truncations at each step of the cascade (26). Denoting by  $\delta_{i,j}$  the error due to truncation of the  $i$ -th component of the vector  $\underline{y}_k^j$ , it can easily be shown that the final (discrete valued) transform vector  $\underline{y}_k^q$  is obtained by

$$\underline{y}_k^q = \begin{bmatrix} y_{1,k} = \left[ x_1^q + \frac{d-1}{c} x_2^q + \delta_{1,1} + \frac{a-1}{c} (c x_1^q + c \delta_{1,1} + d x_2^q + \delta_{2,2}) \right]_{\Delta_1} \\ y_{2,k} = \left[ c x_1^q + c \delta_{1,1} + d x_2^q \right]_{\Delta_2} \end{bmatrix}. \tag{27}$$

Assuming small quantization stepsizes (i.e. the independence of the quantization noises  $\delta_{i,j}$ , and the Gaussianity of the transformed signals), the discrete entropy of each transformed random variable may be approximated as

$$\begin{aligned}
H(y_1^q) &\approx \frac{1}{2} \log_2 2\pi e \underbrace{\left( \lambda_1' + \frac{a^2 \Delta_1^2}{12} + \frac{(a-1)^2 \Delta_1^2}{c} \frac{\Delta_1^2}{12} \right)}_{\lambda_1'' = \lambda_1' + e_1} - \log_2 \Delta_1 \\
H(y_2^q) &\approx \frac{1}{2} \log_2 2\pi e \underbrace{\left( \lambda_2' + \frac{c^2 \Delta_1^2}{12} \right)}_{\lambda_2'' = \lambda_2' + e_2} - \log_2 \Delta_2.
\end{aligned} \tag{28}$$

Thus,  $y_i^q$  may be seen as a continuous r.v. of variance  $\lambda_i'' = \lambda_i' + e_i$ , quantized with stepsize  $\Delta_i$ . The terms  $e_i$  are the increase in the variance of the transform signals due to the truncations. The corresponding expression for the lossless coding gain in the  $N = 2$  case is then

$$\begin{aligned}
G_{V_{int}^q} &= \sum_{i=1}^2 H(x_i^q) - H(y_i^q) \\
&\approx \frac{1}{2} \log_2 \frac{\prod_{i=1}^2 \sigma_{x_i}^2}{\prod_{i=1}^2 \lambda_i''},
\end{aligned} \tag{29}$$

where subscript  $V_{int}^q$  refers to the integer-to-integer implementation of  $V^q$ . Comparing with the gain obtained for the lossless implementation of the LDU (20) we have  $G_{V_{int}^q} \leq G_{L^q, int}$  (this follows from the following series of inequalities  $\prod_{i=1}^2 \lambda_i'' \geq \prod_{i=1}^2 \lambda_i' = \prod_{i=1}^2 \sigma_{y_i}^2 > \sigma_{x_1}^2 \sigma_{y_2}^2$ ). Thus the gain for the integer-to-integer KLT is clearly inferior to that of the integer-to-integer LDU for the  $N = 2$  case. Indeed, only one triangular transform/truncation is involved in the LDU case whereas three triangular transforms/truncations are generally necessary to losslessly implement the KLT. In the general  $N$ -case, the triangular structure of the prediction matrix allows one to implement the lossless causal transform using  $N - 1$  truncations (see (14)), which is most probably less than the number required in the unitary case, where the transform matrix has not a triangular structure.

An alternative expression of  $G_{V_{int}^q}$  may be obtained by approximating the following product under high resolution assumption

$$\prod_{i=1}^2 \lambda_i'' = \prod_{i=1}^2 \lambda_i' \left( 1 + \frac{e_i}{\lambda_i'} \right) \approx \prod_{i=1}^2 \lambda_i' \left( 1 + \sum_{i=1}^2 \frac{e_i}{\lambda_i'} \right). \tag{30}$$

We get

$$\begin{aligned}
G_{V_{int}^q} &\approx \frac{1}{2} \log_2 \frac{\det \text{diag}\{R_{xx}\}}{\det R_{xx^q}} - \frac{1}{2} \log_2 \left( 1 + \sum_{i=1}^2 \frac{e_i}{\lambda_i'} \right) \\
&\approx \frac{1}{2} \log_2 \frac{\det \text{diag}\{R_{xx}\}}{\det R_{xx^q}} - \frac{1}{2 \ln 2} \sum_{i=1}^2 \frac{e_i}{\lambda_i'}.
\end{aligned} \tag{31}$$

As (29), this expression holds for  $N = 2$ , since the perturbation terms on the variances  $e_i$  in (28) have been analytically derived in this case only. However, the truncation effects can be similarly analyzed for a general  $N$ , and the expressions (29), (31) would hold more generally by plugging in the corresponding  $e_i$ .

As in the previous subsection, (31) may be rewritten by expressing  $G_{max}$  as

$$G_{V_{int}^q} \approx G_{max} - \frac{1}{2 \ln 2} \left[ \text{tr}\{R_{xx}^{-1} D\} + \sum_{i=1}^2 \frac{e_i}{\lambda_i'} \right]. \tag{32}$$

Finally, as expected,  $G_{V_{int}^q}$  tends to  $G_{max}$  as  $\Delta_i$  tends to 0,  $i = 1, \dots, N$ .

#### 4. EFFECTS OF THE ESTIMATION NOISE ON THE LOSSLESS CODING GAINS

We analyze in this section the coding gains of an adaptive scheme based on an estimate of the covariance matrix  $\widehat{R_{xx^q}} = R_{xx^q} + \Delta R = \frac{1}{K} \sum_{i=1}^K \underline{x}_i^q \underline{x}_i^{qT}$ , where  $K$  is the number of previously decoded vectors available at the decoder. We suppose independent identically distributed Gaussian real vectors  $\underline{x}_i^q$ , which is for example the case if the sampling period of the scalar signals is high in comparison with their typical correlation time. (Again, the r.v.s are not strictly Gaussian because of the contribution of the uniform quantization noise. This contribution is however small for a high resolution quantization). Thus, the first and second order statistics of  $\Delta R$  are known [16]:  $(\Delta R)_{ii}$  is, for sufficiently high  $K$ , a zero mean Gaussian random variable with covariance matrix such that  $\text{Evec}(\Delta R) (\text{vec}(\Delta R))^T = \frac{2}{K} R_{xx^q} \otimes R_{xx^q}$ , where  $\otimes$  denotes the Kronecker product. For each realization of  $\Delta R$ , the coder computes in a first step the linear transformation  $\widehat{T}$  ( $\widehat{T} = \widehat{L}^q$  or  $\widehat{V}^q$ )

which diagonalizes  $\widehat{R_{\underline{x}^q \underline{x}^q}} : \widehat{T} \widehat{R_{\underline{x}^q \underline{x}^q}} \widehat{T}^T = \widehat{\Sigma}$ . Then, by using the previously exposed factorizations, the coder computes the corresponding integer-to-integer transform  $\widehat{T}_{int}$ .

In order to derive the coding gains for the two approaches in presence of estimation noise, we need the following result ([7]). **Result** Suppose that the transformation  $\widehat{T}$  ( $\widehat{T} = \widehat{L}^q$  or  $\widehat{V}^q$ ) is based on an estimate of the covariance matrix  $\frac{1}{K} \sum_{i=1}^K \underline{x}_i^q \underline{x}_i^{qT}$ . Without estimation noise, the variance of the transform signals obtained by applying the transform  $T$  to the quantized signals would be  $(TR_{\underline{x}^q \underline{x}^q}T^T)_{ii} = (\Sigma)_{ii}$ . (For the LDU,  $(\Sigma)_{ii} = R_{\underline{y}' \underline{y}'}$ , and  $(\Sigma)_{ii} = \Lambda'$  for the KLT, as in equations (11) and (24)). Now, the actual variances of the signals obtained by applying  $\widehat{T}$  to  $\underline{x}^q$  are  $E(\widehat{T} R_{\underline{x}^q \underline{x}^q} \widehat{T}^T)_{ii} = (\Sigma + \Delta \Sigma)_{ii}$ . Then for  $\widehat{T} = \widehat{L}^q$ ,  $\widehat{V}^q$ , it can be shown that for sufficiently high  $K$

$$E \sum_{i=1}^N \frac{(\Delta \Sigma)_{ii}}{(\Sigma)_{ii}} \approx \frac{N(N-1)}{2K}, \quad (33)$$

where  $E$  denotes the expectation. We can now derive the gains obtained when the transformations are based on an estimate of the covariance matrix by means of  $K$  vectors.

#### 4.1. Coding Gain for the integer-to-integer LDU

One has to compute the difference

$$G_{\widehat{L}_{int}^q}(K) = E \sum_{i=1}^N H(x_i^q) - H(y_i^q, K), \quad (34)$$

where only the entropies  $H(y_i^q, K)$  of the discrete variables  $y_i^q$ , obtained by applying  $\widehat{L}_{int}^q$  to  $\underline{x}^q$ , depend on  $K$ . Since the variance of the first variable  $y_1^q$  is not affected by the transformation, we have

$$E H(y_1^q, K) = H(x_1^q) \approx \frac{1}{2} \log_2(2\pi e) \sigma_{x_1}^2 - \log_2 \Delta_1. \quad (35)$$

Concerning the  $N-1$  remaining r.v.s  $y_i^q$ , they may be seen as r.v.s obtained by applying  $L^q$  to  $\underline{x}^q$ , and then by quantizing the resulting continuous valued result with stepsize  $\Delta_i$ . Thus, by denoting  $(\widehat{L}^q R_{\underline{x}^q \underline{x}^q} \widehat{L}^q)_{ii} = (R_{\underline{y}' \underline{y}'})_{ii} + \Delta(R_{\underline{y}' \underline{y}'})_{ii}$ , we have

$$\begin{aligned} E H(y_i^q, K) &\approx E \frac{1}{2} \log_2(2\pi e) (\widehat{L}^q R_{\underline{x}^q \underline{x}^q} \widehat{L}^q)_{ii} - \log_2 \Delta_i (2\pi e) \\ &= \frac{1}{2} \log_2 \left( (2\pi e (R_{\underline{y}' \underline{y}'})_{ii} \left( 1 + \frac{\Delta (R_{\underline{y}' \underline{y}'})_{ii}}{(R_{\underline{y}' \underline{y}'})_{ii}} \right) \right) - \log_2 \Delta_i \\ &\approx \frac{1}{2} \log_2 2\pi e (R_{\underline{y}' \underline{y}'})_{ii} - \log_2 \Delta_i + \frac{1}{2 \ln 2} E \frac{\Delta (R_{\underline{y}' \underline{y}'})_{ii}}{(R_{\underline{y}' \underline{y}'})_{ii}}. \end{aligned} \quad (36)$$

Thus, we have

$$E \sum_{i=1}^N H(y_i^q) \approx \frac{1}{2} \log_2(2\pi e) \sigma_{x_1}^2 - \log_2 \Delta_1 + \sum_{i=2}^N \frac{1}{2} \log_2(2\pi e)^{N-1} \sigma_{y_i'}^2 - \log_2 \Delta_i + \sum_{i=2}^N \frac{1}{2 \ln 2} E \frac{\Delta (R_{\underline{y}' \underline{y}'})_{ii}}{(R_{\underline{y}' \underline{y}'})_{ii}}. \quad (37)$$

Comparing with the bit rate required to code the  $y_i^q$  when the transformation is not perturbed (19), the last term corresponds to an excess bit rate due to estimation noise. Using the result (33), and the fact that  $E \Delta (R_{\underline{y}' \underline{y}'})_{11} = 0$ , this term may be written as

$$\sum_{i=2}^N \frac{1}{2 \ln 2} E \frac{\Delta (R_{\underline{y}' \underline{y}'})_{ii}}{(R_{\underline{y}' \underline{y}'})_{ii}} = \frac{1}{2 \ln 2} E \sum_{i=1}^N \frac{\Delta (R_{\underline{y}' \underline{y}'})_{ii}}{(R_{\underline{y}' \underline{y}'})_{ii}} \approx \frac{N(N-1)}{4 \ln 2K}. \quad (38)$$

Finally, the lossless coding gain for an integer-to-integer implementation of the LDU when the transform is based on  $K$  observed vectors may be approximated as

$$\begin{aligned} G_{\widehat{L}_{int}^q}(K) &= E \sum_{i=1}^N H(x_i^q) - H(y_i^q, K) \\ &\approx G_{L_{int}^q} - \frac{N(N-1)}{4 \ln 2K}. \end{aligned} \quad (39)$$

for high  $K$  and under high resolution assumption.

## 4.2. Lossless Coding Gain for the integer-to-integer KLT

In this case, one has to compute the difference

$$G_{\widehat{V}_{int}^q}(K) = E \sum_{i=1}^N H(x_i^q) - H(y_i^q, K), \quad (40)$$

where only the entropies  $H(y_i^q, K)$  of the discrete variables  $y_i^q$ , obtained by applying  $\widehat{V}_{int}^q$  to  $\underline{x}^q$ , depend on  $K$ .

In a first step, the variances of the continuous transform signals obtained by applying  $\widehat{V}^q$  to the quantized signals  $x_i^q$  may, as in the previous section, be written as

$$E(\widehat{V}^q R_{\underline{x}^q \underline{x}^q} \widehat{V}^q)_{ii} = (\Lambda')_{ii} + (\Delta \Lambda')_{ii} = \lambda'_i + (\Delta \Lambda')_{ii} \quad (41)$$

Let us now analyze the effects of the estimation noise when the integer-to-integer KLT is implemented : the transform  $\widehat{V}_{int}^q$  is for this purpose computed using a factorization similar to (25), involving  $\widehat{V}^q$ . Because of the rounding operations necessary to losslessly implement the transform  $\widehat{V}_{int}^q$ , the r.v.s  $y_i^q$  may be seen as continuous r.v.s of variances  $\lambda'_i + (\Delta \Lambda')_{ii} + \hat{e}_i$ , where the perturbing terms  $\hat{e}_i = e_i + \delta e_i$  indeed depend on  $K$  since they involve the coefficients of the estimated transform (see (28)). However, the dependence in  $K$  is a perturbation occurring on terms which are themselves small. Thus, we assume that  $e_i(K) \approx e_i$  for high  $K$ . The bit rate required to entropy code the discrete transform signals is hence

$$\begin{aligned} E \sum_{i=1}^N H(y_i^q, K) &\approx E \sum_{i=1}^N \frac{1}{2} \log_2 2\pi e \left( \lambda'_i + (\Delta \Lambda')_{ii} + e_i \right) - \log_2 \Delta_i \\ &\approx E \frac{1}{2} \log_2 \frac{(2\pi e)^N}{\prod_{i=1}^N \Delta_i^2} \left( \prod_{i=1}^N \lambda'_i + (\Delta \Lambda')_{ii} + e_i \right) \\ &\approx \frac{1}{2} \log_2 \frac{(2\pi e)^N}{\prod_{i=1}^N \Delta_i^2} \left( \prod_{i=1}^N \lambda'_i + (\Delta \Lambda')_{ii} \right) \left( 1 + \sum_{i=1}^N \frac{e_i}{\lambda'_i + (\Delta \Lambda')_{ii}} \right) \\ E \sum_{i=1}^N H(y_i^q, K) &\approx E \frac{1}{2} \log_2 \frac{(2\pi e)^N}{\prod_{i=1}^N \Delta_i^2} \left( \prod_{i=1}^N \lambda'_i \right) \left( 1 + \sum_{i=1}^N \frac{(\Delta \Lambda')_{ii}}{\lambda'_i} \right) \left( 1 + \sum_{i=1}^N \frac{e_i}{\lambda'_i + (\Delta \Lambda')_{ii}} \right) \\ &= E \frac{1}{2} \log_2 \frac{(2\pi e)^N}{\prod_{i=1}^N \Delta_i^2} \det R_{\underline{x}^q \underline{x}^q} + \frac{1}{2} \log_2 \left( 1 + \sum_{i=1}^N \frac{(\Delta \Lambda')_{ii}}{\lambda'_i} \right) + \frac{1}{2} \log_2 \left( 1 + \sum_{i=1}^N \frac{e_i}{\lambda'_i + (\Delta \Lambda')_{ii}} \right). \end{aligned} \quad (42)$$

By result (33), the second term in (42) may, for sufficiently high  $K$ , be approximated as

$$E \frac{1}{2} \log_2 \left( 1 + \sum_{i=1}^N \frac{(\Delta \Lambda')_{ii}}{\lambda'_i} \right) \approx \frac{N(N-1)}{4 \ln 2K}. \quad (43)$$

The last term in (42) may also be, for high  $K$  and under high resolution assumptions, be approximated as

$$\frac{1}{2} \log_2 \left( 1 + \sum_{i=1}^N \frac{e_i}{\lambda'_i + (\Delta \Lambda')_{ii}} \right) \approx \frac{1}{2 \ln 2} \sum_{i=1}^N \frac{e_i}{\lambda'_i + (\Delta \Lambda')_{ii}} \approx \frac{1}{2 \ln 2} \sum_{i=1}^N \frac{e_i}{\lambda'_i}. \quad (44)$$

Finally, the lossless coding gain with estimation noise for the integer-to-integer KLT may be approximated as

$$\begin{aligned} G_{\widehat{V}_{int}^q}(K) &\approx \frac{1}{2} \log_2 \frac{\det \text{diag}\{R_{\underline{x}^q \underline{x}^q}\}}{\det R_{\underline{x}^q \underline{x}^q}} - \frac{1}{2 \ln 2} \sum_{i=1}^N \frac{e_i}{\lambda'_i} - \frac{N(N-1)}{4 \ln 2K} \\ &\approx G_{V_{int}^q} - \frac{N(N-1)}{4 \ln 2K}, \end{aligned} \quad (45)$$

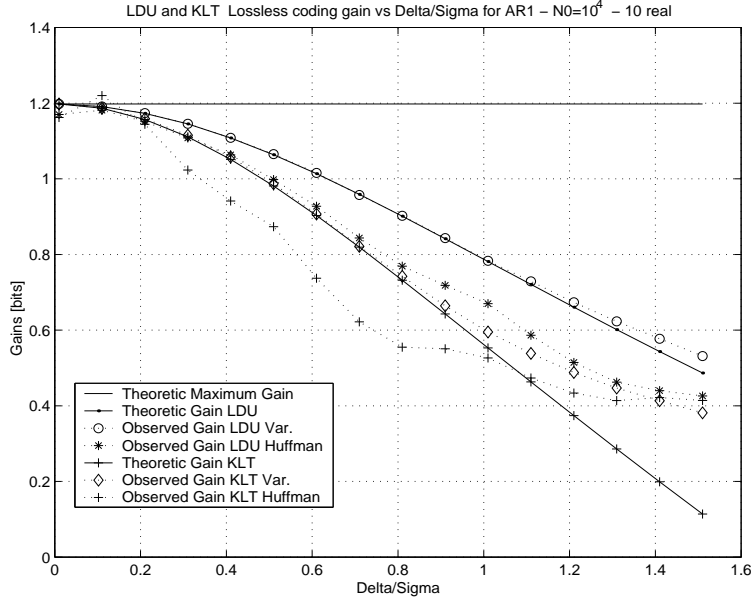
under high resolution assumption and for sufficiently high  $K$ . As in section 3.2, this expression holds for  $N = 2$  (in which case we have derived analytically the gain  $G_{V_{int}^q}$ ), but would hold more generally with the corresponding  $G_{V_{int}^q}$ .

## 5. SIMULATIONS

For the simulations, we used entropy constrained uniform scalar quantizers and real Gaussian i.i.d. vectors. In the first part of this section, we compare the lossless coding gains obtained for the integer-to-integer implementations of the LDU and the KLT for  $N=2$ . Then simulation results for higher values of  $N$  are presented in the case of the LDU. The second part of this section describes the effects of estimation noise on the coding gains in the case  $N = 2$ .

## 5.1. Lossless Coding Gains without estimation noise

### 5.1.1. Results for $N = 2$ .



**Fig. 5.** Lossless coding gains for integer-to-integer implementations of the LDU and KLT vs quantization stepsize.  $N = 2$ ,  $\rho = 0.9$ .

In order to check the theoretical results we generated real Gaussian vectors of covariance matrix  $R_{\underline{x}\underline{x}}$  (covariance matrix of a first order autoregressive process with normalized correlation coefficient  $\rho$ ). The number of vectors was  $N_0 = 10^4$ . The vectors were then quantized using the same normalized quantization stepsize for all  $i$ :  $frac{\Delta_i \sigma_{x_i}}{\sigma_x} = \frac{\Delta}{\sigma_x}$ . For several values of  $\frac{\Delta}{\sigma_x}$ , the optimal decorrelating transformations  $L^q$  and  $V^q$  were computed using the covariance matrix  $R_{\underline{x}^q \underline{x}^q}$ . The integer-to-integer transforms  $L_{int}^q$  and  $V_{int}^q$ , based on the transforms  $L^q$  and  $V^q$  were then implemented and used to compute the transformed data  $\underline{y}_i^q$ . We realized this experience ten times and averaged the different resulting gains. These gains are plotted in figure (5) versus  $\frac{\Delta}{\sigma_x}$ .

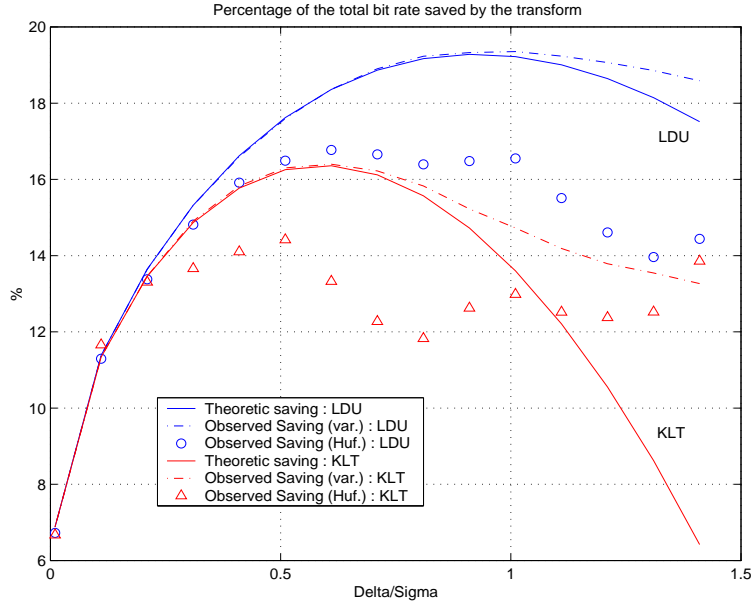
The theoretic maximum coding gain is related to the mutual information between the variables as defined in (9). The theoretic gains for LDU and KLT are then given by (20) and (29) respectively. The observed lossless coding gains were then computed in two different ways.

Since under high resolution assumption there is a one to one correspondence between the discrete entropy and the variances of the transform signals through the relations (17) and (28)), the first way was to estimate the variances of the transform signals before quantization ( $\sigma_{y_i}^2$  (17) for the LDU and  $\lambda_i''$  for the KLT (28)). The observed gains are then obtained by plugging in the relation (20) and (29) the measured variance of the test signals. These gains are referred to by "Observed Gain  $\{KLT, LDU\}$  Var." in figure (5).

The second way is obviously to design Huffman codes for the quantized signals  $x_i^q$ , and then other Huffman codes for the transformed discrete signals  $y_i^q$ . Since the average length of the codewords will be close to the entropy of each scalar source, the difference of the average codelengths gives a precise insight of how many bits are saved by using integer-to-integer transform. These gains are referred to by "Observed Gain  $\{KLT, LDU\}$  Huffman" in figure (5).

It first can be seen that for high resolution (small values of  $\frac{\Delta}{\sigma_x}$ ), the observed gains correspond well to the predicted ones. Moreover, the bit rate that can be saved by using integer-to-integer transforms is rather high in this case. For example, 11% of the total bit rate can be saved for  $\frac{\Delta}{\sigma_x} = 0.1$  by using any of the two integer-to-integer transformations analyzed in this work. For  $\frac{\Delta}{\sigma_x} = 0.51$ , 16% of the total bit rate can be saved by using the integer-to-integer implementation of the LDU, and 15% for the integer-to-integer implementation of the KLT. Figure (6) illustrates the percentage of the total bit rate which can be saved by using the two analyzed integer-to-integer transforms.

Coming back to figure (5), the rounding effects due to the lossless implementation of the transforms indeed can be seen to



**Fig. 6.** Percentage of the total bit rate saved by using integer-to-integer transform.  $N = 2$   $\rho = 0.9$

increase as the quantization gets coarser. The observed coding gains based on the estimates of the variances of the transformed signals correspond well to the predicted ones until a ratio  $\frac{\Delta}{\sigma_x} \approx 1$ . When the quantization becomes even coarser, the quantization noises are not independent anymore, and the mutual information between the quantized variables  $\{x_i^q\}$  becomes superior to the theoretical one. Figure (7) shows the normalized correlation coefficients of the quantization noise versus the normalized correlation coefficient of the variables  $x_1$  and  $x_2$  for several quantization stepsizes. It indicates that for many usual coding situations, the hypothesis of independence of the quantization noises is very reasonable. When the correlation is not negligible, the transforms take more advantage of the information shared between the quantized variables, and the gains become slightly superior to the predicted ones.

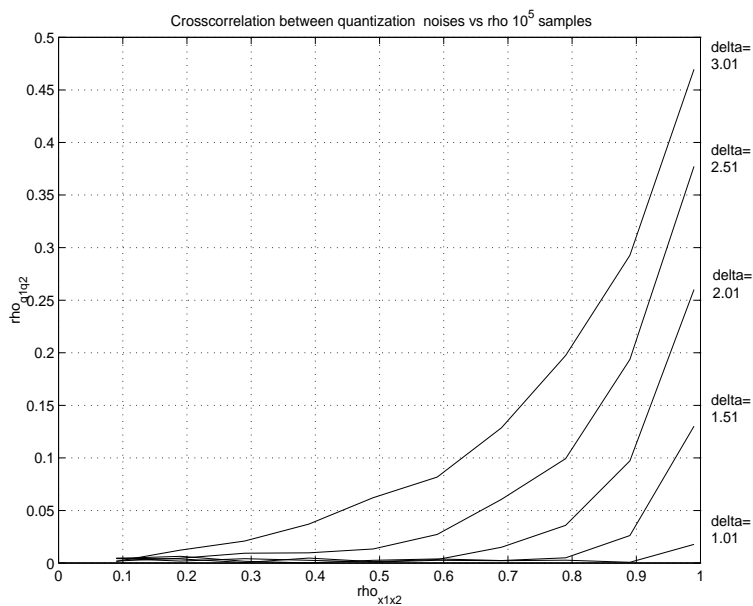
Concerning the gains based on a Huffman coding of the losslessly transformed signals, they are lower than the predicted ones though the difference is weak (typically  $\approx 0.1$  bit/sample for "reasonable" values of  $\frac{\Delta}{\sigma_x}$ ). For example, this mismatch is almost  $10^{-1}$  bit for the integer-to-integer LDU at  $\frac{\Delta}{\sigma_x} = 0.51$ . A possible explanation is as follows. The gains based on the average lengths,  $l_{Huffman}$ , of the codewords obtained by Huffman coding the discrete variables can be written as

$$\begin{aligned}
 G_{Huffman}(T) &= E l_{Huffman}(x_1^q) + l_{Huffman}(x_2^q) - l_{Huffman}(y_1^q) - l_{Huffman}(y_2^q) \\
 &= E l_{Huffman}(x_2^q) - l_{Huffman}(y_2^q) \\
 &= E \widehat{H}(x_2^q) - \widehat{H}(y_2^q), \\
 &\approx H(x_2^q) - H(y_2^q)
 \end{aligned} \tag{46}$$

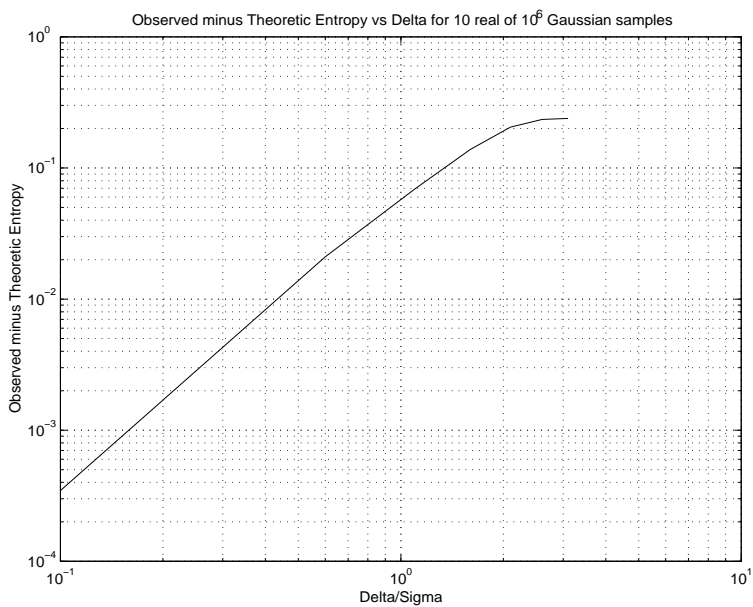
where  $\widehat{H}$  denotes an estimate of the entropy. The approximation in (46) comes from the fact that the averaged length of the Huffman codewords is indeed not guaranteed to correspond to entropies of the discrete r.v.s since the Shannon-Fano integer constraint<sup>1</sup> is not imposed. This may lead to a slight mismatch between actual entropies and observed gains. Moreover, we supposed theoretically that the relation of differential to discrete entropy is given by (2). This assumes a constant p.d.f. within the quantization bins, which is only asymptotically true. Figure (8) plots the difference of the actual discrete entropy of a uniformly quantized Gaussian scalar source  $\{x\}$  minus the theoretic entropy (given by (2)), versus  $\frac{\Delta}{\sigma_x}$ . This shows that the actual entropy is greater than the theoretical one, and that the difference indeed depends on  $\frac{\Delta}{\sigma_x}$ . Now, our theory predicts the same relationship between discrete and differential entropies for the variables  $x_2^q$  and  $y_2^q$ . The latter, however, may be seen (see Section 3) as the optimal prediction  $y_2^q$  of  $x_2^q$  quantized with stepsize  $\Delta_2$ . Thus, relatively to its own standard deviation, it is coarsely quantized than  $x_2^q$ , and the actual entropy of the quantized r.v.  $y_2^q$  is greater than the predicted one. This mismatch is greater than the mismatch between predicted and observed entropies for  $x_2^q$  (which can be seen from figure (8) to be roughly

<sup>1</sup>which states that the code will be optimal if the probabilities  $p_i$  of the symbols  $i$  are related to the corresponding codeword lengths  $l(i)$  by  $p_i = 2^{-l(i)}$





**Fig. 7.** Correlation coefficient of quantization noises vs Correlation coefficient of the variables  $x_1$  and  $x_2$  for several quantization stepsize. The variables  $x_1$  and  $x_2$  have variance 1.



**Fig. 8.** Actual minus theoretic entropy for uniformly quantized Gaussian signal vs quantization stepsize.

$10^{-2}$  bit. For  $y'_2$ , whose standard deviation for  $\rho = 0.9$  is about  $\sqrt{5}$  times smaller than  $x_2$ , the relative quantization stepsize is  $\sqrt{5}$  times bigger,  $\approx 2$ , which from figure (8) correspond to an overestimation of  $\approx 10^{-1}$  bit. Hence, the gain obtained by Huffman coding is globally lower than predicted from roughly 0.1 bit. A similar analysis can be made for the transform signals obtained with a lossless implementation of the KLT (where the mismatch is even greater since the lowest variance is generally lower than  $\sigma_{y'_2}^2$ ).

As a conclusion, the curves obtained for  $N = 2$  correspond well to the predicted results for a wide range of  $\frac{\Delta}{\sigma_x}$ , roughly  $\frac{\Delta}{\sigma_x} \approx 1$ . They show that a non negligible part of the bit rate may be saved by using an integer-to-integer transform, and finally that the lossless implementation of the LDU is superior to that of the KLT.

### 5.1.2. Position of the first signal

Figure (9) shows the codings gains obtained for the integer-to-integer LDU applied to scalar sources of unit variance, versus their correlation coefficient  $\rho$ . In the first case, denoted by "1" in the legend, the first signal  $\{x_1\}$  is quantized with stepsize  $\Delta_1 = 0.1$  and the second signal  $\{x_2\}$  with stepsize  $\Delta_2 = 1$ . In the second case, denoted by "2" in the legend, the stepsizes are 1 for  $\{x_1\}$  and 0.1 for  $\{x_2\}$ . The curves show as expected that the most coarsely quantized signal must be placed in first place in order to maximize the lossless coding gain.

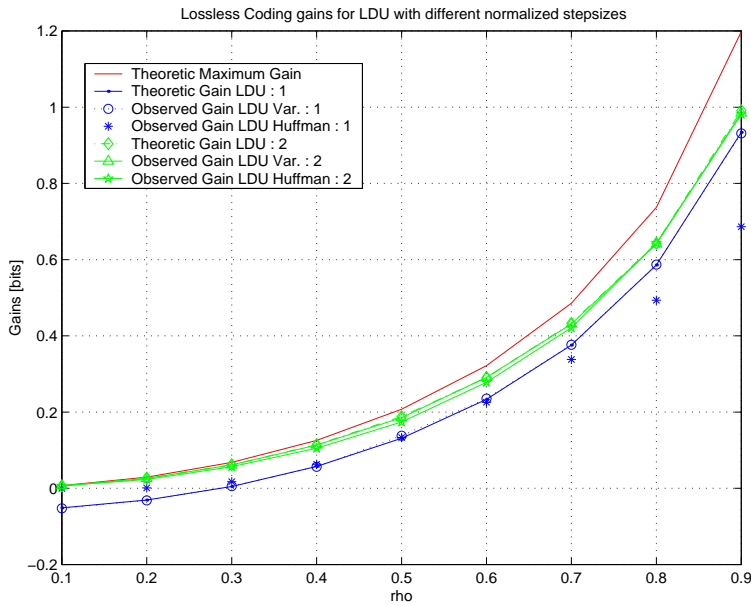


Fig. 9. Importance of the coarseness of the quantization of the first signal.

### 5.1.3. Results for $N > 2$ .

The coding gains obtained for the integer-to-integer LDU with  $N = 5$ ,  $\Delta = 0.51$  and  $N = 5$ ,  $\Delta = 0.21$  are presented in figures (10) and (11) respectively. In this case, the data are real Gaussian i.i.d. vectors with covariance matrix  $R = H R_{\underline{x}\underline{x}} H^T$ .  $R_{\underline{x}\underline{x}}$  is the covariance matrix of a first order autoregressive process with normalized correlation coefficient  $\rho$ .  $H$  is a diagonal matrix whose  $i$ th entry is  $(i)^{2/3}$  (decreasing variances, ranging from 1 to 8.56). Hence, the coarseness of the quantization decreases as  $i$  increases. It can be seen that the observed gains match well the predicted ones. In particular, the previously exposed mismatch between the predicted gains and the observed gains based on Huffman coding concerns only the few first  $y_i^q$  because the stepsize  $\Delta$  becomes, as  $i$  grows, relatively smaller comparatively to the standard deviation of the prediction error  $y_i'$ . Thus, this mismatch becomes negligible relatively to the total gain.

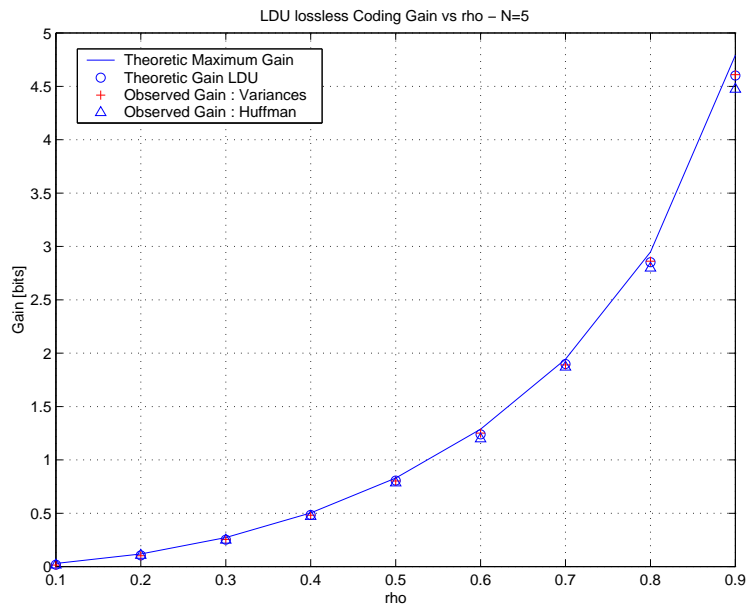


Fig. 10. Lossless Coding Gain for integer-to-integer LDU with  $N = 5$ ,  $\Delta = 0.51$ .

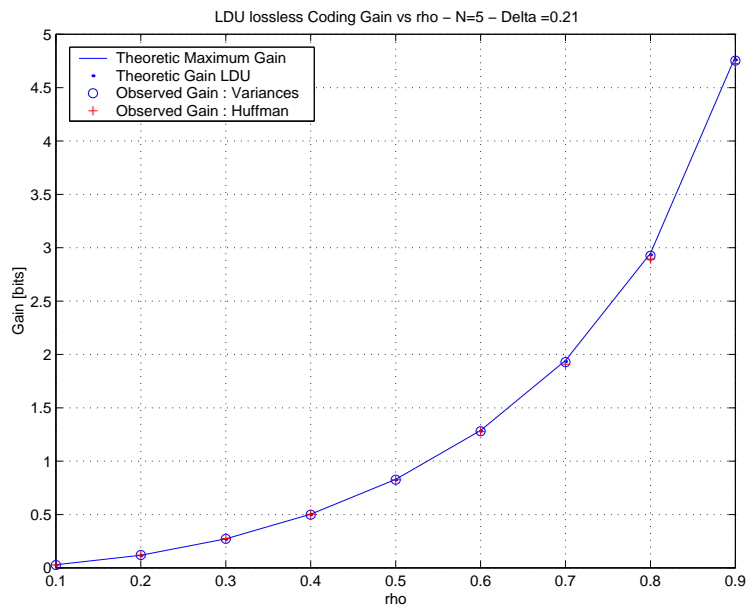
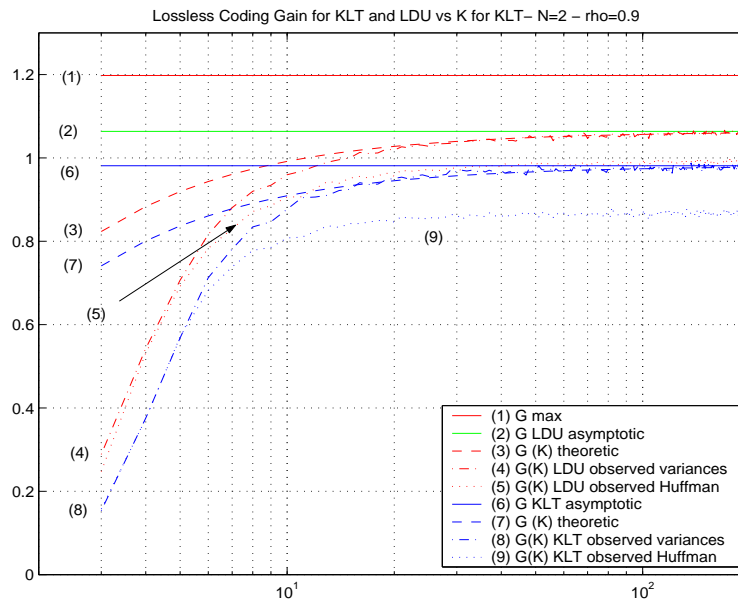


Fig. 11. Lossless Coding Gain for integer-to-integer LDU with  $N = 5$ ,  $\Delta = 0.21$ .

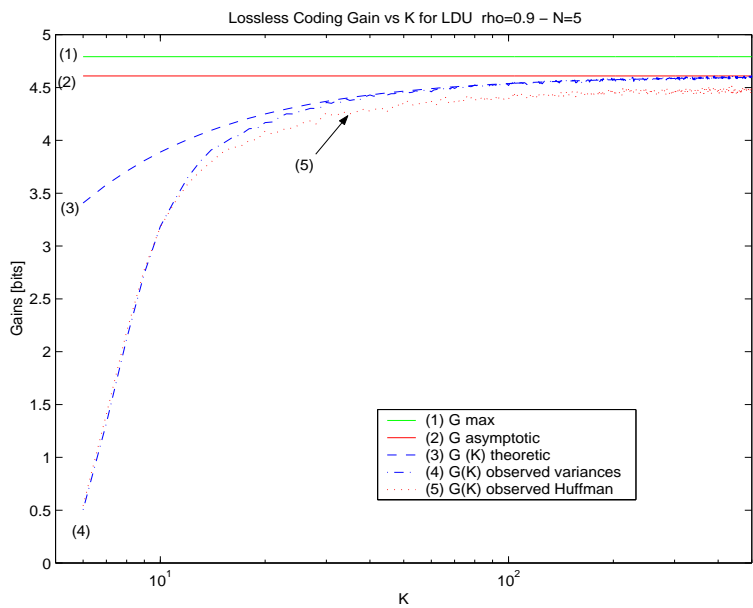
## 5.2. Coding Gains with estimation noise

In the first experiment, the data were generated as in section 5.1.1, for  $N = 2$ . The coding gains with estimation noise are plotted in figure (12). The normalized quantization stepsize is  $\frac{\Delta}{\sigma_x} = 0.51$ . The coding gain  $G_{max}$  refers to the mutual information given by (9). The theoretic gains for LDU and KLT are given by (39) and (45) respectively (gains referred to as "G(K) {KLT, LDU} Asymptotic"). The observed coding gains are, as in the previous section, whether based on the estimates of the variances of the transform signals (gains referred to as "G(K) {KLT, LDU} observed variances"), whether based on the actual gain computed by Huffman coding. In this case, a Huffman code is designed for the signals obtained with integer-to-integer transforms based on an estimate of the covariance matrix of quantized data  $\widehat{R_{x^q x^q}}$  with  $K$  vectors. The theoretic curves correspond well to the observed ones for the observed gains based on variances estimates for  $K \approx$  a few tens. The slight mismatch between Huffman based and variance based observed gains may be due to the overestimation explained in the previous subsection. Huffman based and variance based observed gains reach 90% of their maximal value for  $K \approx 10$  decoded vectors. That is, basing our conclusions on averaged codewords lengths obtained with a Huffman code, 90% of 16% of the total bit rate can be saved for entropy coding each 2-vector as soon as  $K \approx 10$  in the case of the integer-to-integer LDU. In the case of the integer-to-integer KLT, 90% of 15% of the total bit rate can be saved for a comparable estimation noise.

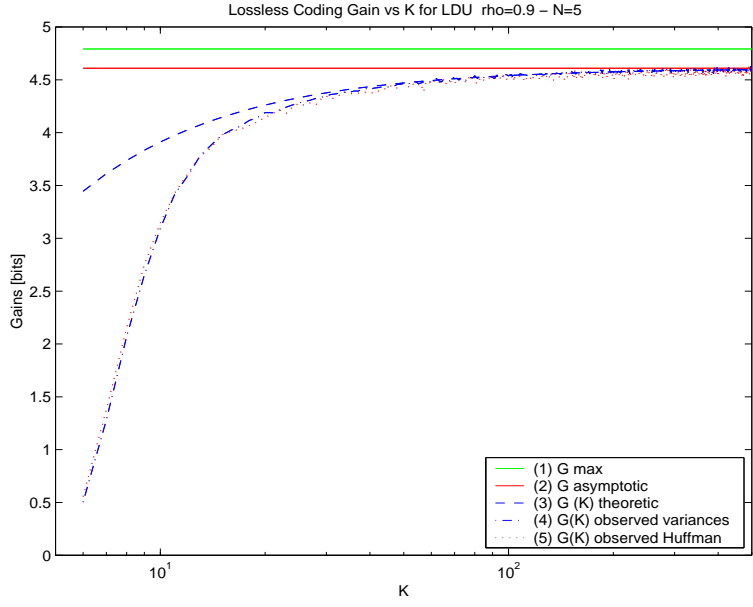
Figure (13) and (14) plot the lossless coding gain with estimation noise versus  $K$  for the same type of data as in the previous subsection 5.1.3 :  $N = 5$ ,  $\Delta = 0.51$ , and  $\Delta = 0.21$ . Theoretic and observed gains correspond well for  $K \approx$  a few tens.



**Fig. 12.** Coding Gains with estimation noise versus  $K$  for  $N=2$ .  $\frac{\Delta}{\sigma_x} = 0.51$ .



**Fig. 13.** Coding Gains with estimation noise versus K for N=5.  $\Delta = 0.51$ .



**Fig. 14.** Coding Gains with estimation noise versus K for N=5.  $\Delta = 0.21$ .

## 6. APPENDIX

The proof of expression (3) is a straightforward extension to  $N$  r.v.s and different quantization stepsizes of Theorem 9.3.1, [1].

Consider a set of random variables  $x_1, x_2, \dots, x_N$  with density  $f(x_1, x_2, \dots, x_N)$ . Suppose we divide the range of each  $x_i$  into bins (quantization cells) of length  $\Delta_i$ . The density is assumed to be continuous within the bins. The proof includes three steps : firstly we find the mean of each quantization cell. Then we express the probability that a vector is quantized into a particular cell, and finally we estimate the entropy of the quantized source.

By the mean value theorem, there exists a set of values  $\underline{x}_j = [x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N]^T$  within each bin such that

$$\Delta_1 \Delta_2 \dots \Delta_N f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) = \int_{j_1 \Delta_1}^{(j_1+1)\Delta_1} \dots \int_{j_N \Delta_N}^{(j_N+1)\Delta_N} f(x_1, x_2, \dots, x_N) dx_1 \dots dx_N. \quad (47)$$

Consider now the set of quantized r.v.  $\underline{x}^q = [x_{1,j_1}^q \dots x_{N,j_N}^q]^T$ , defined by

$$x_{i,j_i}^q = x_{j_i}^i, \quad \text{if } j_i \Delta_i \leq x_i < (j_i + 1) \Delta_i. \quad (48)$$

The probability that a certain quantized vector  $\underline{x}^q$  is chosen to be represented by  $\underline{x}_j$  is

$$p_{j_1, \dots, j_N} = \int_{j_1 \Delta_1}^{(j_1+1)\Delta_1} \dots \int_{j_N \Delta_N}^{(j_N+1)\Delta_N} f(x_1, x_2, \dots, x_N) dx_1 \dots dx_N = f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) \Delta_1 \Delta_2 \dots \Delta_N. \quad (49)$$

The entropy of the vectorial source quantized using the rule (48) is

$$\begin{aligned} H(\underline{x}^q) &= - \sum_{j_1=-\infty}^{\infty} \dots \sum_{j_N=-\infty}^{\infty} p_{j_1, \dots, j_N} \log_2 p_{j_1, \dots, j_N} \\ &= - \sum_{j_1=-\infty}^{\infty} \dots \sum_{j_N=-\infty}^{\infty} f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) \Delta_1 \Delta_2 \dots \Delta_N \log_2 (f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) \Delta_1 \Delta_2 \dots \Delta_N) \\ &= - \sum_{j_1=-\infty}^{\infty} \dots \sum_{j_N=-\infty}^{\infty} f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) \Delta_1 \Delta_2 \dots \Delta_N \log_2 f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) \\ &\quad - \sum_{j_1=-\infty}^{\infty} \dots \sum_{j_N=-\infty}^{\infty} \Delta_1 \Delta_2 \dots \Delta_N f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) \log_2 (\Delta_1 \Delta_2 \dots \Delta_N). \end{aligned} \quad (50)$$

If  $f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) \log_2 f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N)$  is Riemann integrable, the first approaches the integral of  $f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) \log_2 f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) = h(\underline{x})$  by definition of Riemann integrability. The second term in (50) may be written as

$$\begin{aligned} & - \Delta_1 \Delta_2 \dots \Delta_N \sum_{j_1=-\infty}^{\infty} \dots \sum_{j_N=-\infty}^{\infty} f(x_{j_1}^1, x_{j_2}^2, \dots, x_{j_N}^N) \log_2 \Delta_1 \Delta_2 \dots \Delta_N \\ &= - \underbrace{\left( \Delta_1 \Delta_2 \dots \Delta_N \sum_{j_1=-\infty}^{\infty} \dots \sum_{j_N=-\infty}^{\infty} \int_{j_1 \Delta_1}^{(j_1+1)\Delta_1} \dots \int_{j_N \Delta_N}^{(j_N+1)\Delta_N} f(x_1, x_2, \dots, x_N) dx_1 \dots dx_N \right)}_{\rightarrow 1} \log_2 \Delta_1 \Delta_2 \dots \Delta_N \\ &= \log_2 \Delta_1 \Delta_2 \dots \Delta_N \\ &= \sum_{i=1}^N \log_2 \Delta_i \end{aligned} \quad (51)$$

Hence, we get

$$H(\underline{x}^q) + \sum_{i=1}^N \log_2 \Delta_i \rightarrow h(\underline{x}) \quad \text{as } \Delta_i \rightarrow 0, \quad i = 1 \dots N. \quad (52)$$

## 7. REFERENCES

- [1] T. M. Cover, J.A. Thomas ,“Elements of Information Theory,” *Wiley Series in Telecommunications*, 1991.
- [2] M. Hans, R.W. Schafer ,“Lossless Compression of Digital Audio,” *IEEE Sig. Proc. Mag.*, vol. 18, no. 4, July 2001.
- [3] D. Mary, D. T.M. Slock, “Comparison between Unitary and Causal Approaches to Backward Adaptive Transform Coding of Vectorial Signals,” *ICASSP 2002*, Orlando, USA, May 2002.

- [4] D. Yang, H. Ai, C. Kyriakakis, C.-C. Jay Kuo, "An Inter-Channel Redundancy Removal Approach for High-Quality Multichannel Audio Compression," *AES 109th Convention*, Los Angeles, USA, September 2000.
- [5] D. Mary, D. T.M. Slock, "Backward Adaptive Transform Coding of Vectorial Signals : A Comparison between Unitary and Causal Approaches," Accepted paper, *EUSIPCO-2002*, Toulouse, France, September 2002.
- [6] V. K. Goyal, "Transform Coding with Integer-to-Integer Transforms" *IEEE trans. on Inf. Theory*, vol. 46, no. 2, March 2000.
- [7] D. Mary, D. T.M. Slock, "Comparison between Unitary and Causal Approaches to Transform Coding of Vectorial Signals," Corrected Version of *GRETSI01*, Toulouse, France, September 2001. Available at <http://www.eurecom.fr/mary/publications.html>
- [8] D. Mary, D. T.M. Slock, "Causal Transform Coding, Generalized MIMO Prediction and Application to Vectorial DPCM Coding of Multichannel Audio," *WASPAA01*, New York, USA, October 2001.
- [9] S.-M. Phoong, Y.-P. Lin, "Prediction-Based Lower Triangular Transform," *IEEE trans. on Sig. Proc.*, vol. 48, no. 7, July 2000.
- [10] D. T.M. Slock, K. Maouche, D. Mary, "Codage DPCM Vectoriel et Application au Codage de la Parole en Bande Élargie,"
- [11] D. Mary, D. T.M. Slock, "Vectorial DPCM Coding and Application to Wideband Coding of Speech," *CORESA 2000*, Poitiers, France, October 2000.
- [12] J. R. Stuart, P. G. Craven, M. A. Gerzon, M. J. Law, R. J. Wilson "MLP Lossless Compression" *AES 9th Convention*, Tokyo.
- [13] R. Whittle, "Lossless Audio Compression" URL : <http://www.firstpr.com.au/audiocomp/lossless/>.
- [14] T. Moriya, "Report of AHG on Issues in Lossless Audio Coding", *Draft for Discussion*, ISO/IEC JTC1/SC29/WG11 M7678, December 2001.
- [15] *mpeg – audiolossless@research.att.com* : Reflector for AHG on Issues in Lossless Audio Coding.
- [16] B. Ottersten, M. Viberg, *Sensor Array Signal Processing*, Technical Report, Royal Institute of Tech., Chalmers Technical Inst., Sweden, January 1994.