RESEARCH REPORT N- RR-01-057

COMPARISON BETWEEN UNITARY AND CAUSAL APPROACHES TO BACKWARD ADAPTIVE TRANSFORM CODING OF VECTORIAL SIGNALS

David Mary, Dirk T.M. Slock

December 13, 2001

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

Eurécom's research is partially supported by its industrial partners: Ascom, Swisscom, Thales Communications, STMicroelectron ics, CEGETEL, Motorola, France Télécom, Bouygues Telecom, Hitachi Europe Ltd. and Texas Instruments.

.

ABSTRACT

In a transform coding framework we compare the optimal causal approach (LDU, Lower-Diagonal-Upper) to the optimal unitary approach (Karhunen-Loeve Transform, KLT). The criterion of merit used for this comparison is the coding gain, defined for a transformation T as the ratio of the average distortion obtained with identity transformation over the average distortion obtained with T. In absence of perturbation, both transforms have been recenlty shown to yield the same gain [8, 7]. The purpose of this report is to compare the behavior of these two transformations when the ideal transform coding scheme gets perturbed, that is, when only an estimate $R_{XX} + \Delta R$ of R_{XX} is known. In this case, not only the transformation itself will be perturbated, but also the bit allocation mechanism. We compare the two approaches in three cases. Firstly, ΔR is caused by a quantization noise : the coding scheme is based on the statistics of the quantized data. We find that the coding gain in the unitary case is higher than in the causal case. In a second case, ΔR corresponds to an estimation noise : the coding scheme is based on an estimate of R_{XX} based on a finite amount of available data. In this case, both causal and unitary approaches are strictly equivalent, because of the unimodularity and decorrelating properties of the transformations. Finally, the influence of both perturbations is considered, as this is the case in a real backward adaptive transform coding scheme. Simulations results confirming the predicted behavior of the coding gains with perturbations are reported. The results of this work have been submitted in [2] and [3]. This report is available at http://www.eurecom.fr/˜mary/publications.html .

Keywords

Transform coding, backward adaptive, prediction, estimation, quantization, causal, unitary transformation, coding gain.

Contents

List of Figures

1. INTRODUCTION

Consider a stationary Gaussian vectorial source $\{X\}$. This source may be composed of any scalar source $\{x_i\}$. In the classical transform coding framework, a linear transformation is applied to each N-vector X to produce an N-vector Y 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 $\sum_i r_i = Nr$, see Figure 1. For an entropy constrained scalar quantizer of a Gaussian source, the high resolution

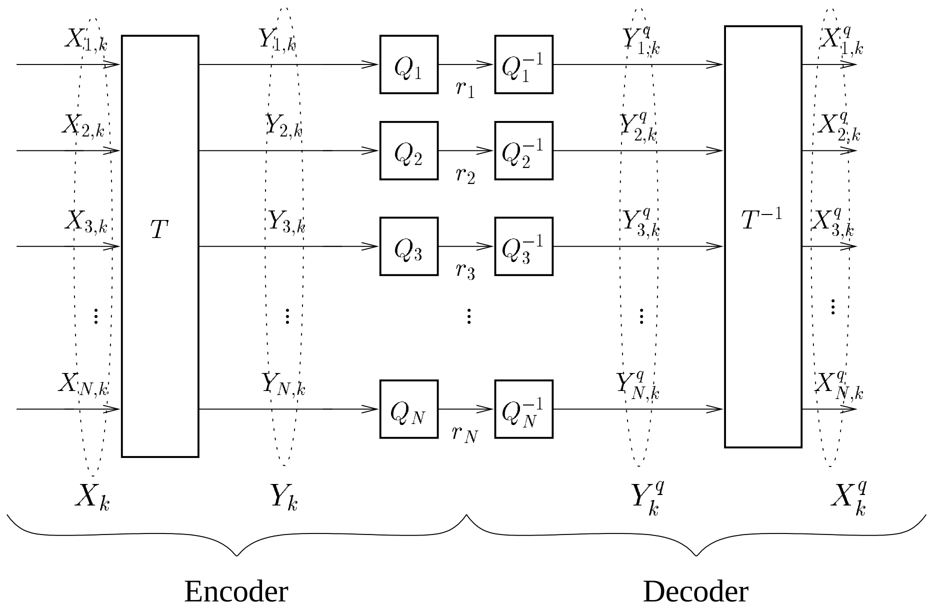

Fig. 1. Classical Transform Coding scheme. $\{Q_i\}$ denotes entropy constrained uniform scalar quantization.

distortion is $\mathrm{E}(y_i^q(k) - y_i(k))^2 = \sigma^2_{q_i} = \frac{\pi e}{6} 2^{-2r_i}\sigma^2_{y_i}$. The constant $c = \frac{\pi e}{6}$ will denote in this report the proportionality coefficient between variance and distortion in the Gaussian case.

An important property of commonly used transformations is that, if a noise (for example quantization noise) is added to the signal in the transformed domain, then its power will be the same in the transform and in the signal domains. This property is sometimes referred to as "unity noise gain" property [8]. The coding gain for such transformations T is defined as

$$G_T = \frac{E||\tilde{X}||^2_{(I)}}{E||\tilde{X}||^2_{(T)}} = \frac{E||\tilde{X}||^2_{(I)}}{E||\tilde{Y}||^2_{(T)}}, \tag{1}$$

where I is the identity matrix, and the notation $||\tilde{X}||^2_{(T)}$ denotes the variance of the quantization error on the vector X, obtained for a transformation T. The optimal bit allocation yields the well known distortion for the vectorial signal $\{Y\}$: $\mathrm{E}||\tilde{Y}||^2_T = \frac{1}{N}\sum_{i=1}^N \sigma^2_{q_i} = N\sigma^2_q = Nc2^{-2r}\left(\prod_{i=1}^N \sigma^2_{y_i}\right)^{\frac{1}{N}}$. The optimal variance of the quantization noise for each component is independent of i, and the number of bits assigned to the i$-th$ component is $r + \frac{1}{2}log_2\frac{\sigma^2_{y_i}}{(\prod_{i=1}^N \sigma^2_{y_i})^{\frac{1}{N}}}$. In the next section, we recall the main characteristics of the optimal causal approach (LDU) when optimized on R_{XX}, and summarize the reasons why its performance is the same as the best unitary approach (KLT).

However, a backward adaptive coding scheme generally deals with non- or locally- stationary signals. Sending the updates of the signal-dependent transformation and bit assignment as side information would cause a considerable overhead for the total bit rate. Thus, the backward adaptive coding scheme should require that neither the transformation nor the parameters of the bit assignment are transmitted to the decoder. So suppose now that the coding scheme is based on $\hat{R}_{XX} = R_{XX} + \Delta R$ instead of R_{XX}, where $\hat{R}_{XX}$ is available at both encoder and decoder. Then the computed transformation will be $\hat{T} = T + \Delta T$, and the distortion will be proportional to the variances of the signals transformed by means of $\hat{T}$ instead of T, say $\sigma^{2'}_{y_i}$. Moreover, the bits r_i should be attributed on the basis of estimates of the variances available at both encoder and decoder also, that is, $(\hat{T}\hat{R}_{XX}\hat{T})_{ii}$, where $(.)_{ii}$ denotes the i$-th$ diagonal element of $(.)$. Hence, under the assumption of Gaussianity for the transformed signals, we get the following measure of distortion for a transformation $\hat{T}$ based on $\hat{R}_{XX}$:

$$E||\tilde{Y}||^2_{(\hat{T})} = E\sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{T}\hat{R}_{XX}\hat{T}^T)_{ii}}{(\Pi_{i=1}^{N}(\hat{T}\hat{R}_{XX}\hat{T}^T)_{ii})^{\frac{1}{N}}}]}\sigma^{2'}_{y_i}. \tag{2}$$

where the expectation is w.r.t. ΔR in case it is non-deterministic. We compare this distortion and the corresponding gain for the KLT and the LDU in three cases. In the third part, ΔR is caused by a quantization noise : the coding scheme is based on the statistics of the quantized data, under high resolution assumption. In the fourth part, ΔR corresponds to an estimation noise : the coding scheme is based on an estimate of R_{XX} due to a finite amount of K vectors : $\hat{R}_{XX} = \frac{1}{k}\sum_{i=1}^{K} X_i X_i^T$.

2. OPTIMAL CAUSAL AND UNITARY APPROACHES WITHOUT PERTURBATION

In the causal case, $Y = LX = X - \overline{L}X$, where $\overline{L}X$ is the reference vector. The output X^q is $Y^q + \overline{L}X$. Note that the reconstruction error $\tilde{X}$ equals the quantization error $\tilde{Y}$:

$$\tilde{X} = X - X^q = X - (Y^q + \overline{L}X) = X - \overline{L}X - Y^q = Y - Y^q = \tilde{Y}, \tag{3}$$

so that the causal transform possesses the unity noise gain property. This is also true in the unitary case, even though $\tilde{Y}$ is in general different from $\tilde{X}$. As detailed in [6, 7], the optimal L in terms of coding gain is such that

$$LR_{XX}L^T = R_{YY} = D = diag\{\sigma^2_{y_1}, ...\sigma^2_{y_N}\}, \tag{4}$$

where $diag\{...\}$ represents a diagonal matrix whose elements are $\sigma^2_{y_i}$. In other words, the components y_i are the prediction errors of x_i with respect to the past values of X, the $X_{1:i-1}$, and the optimal coefficients $-L_{i,1:i-1}$ are the optimal prediction coefficients. Since each prediction error y_i is orthogonal to the subspaces generated by the $X_{1:i-1}$, the y_i are orthogonal, and D is diagonal. It follows that

$$R_{XX} = L^{-1}R_{YY}L^{-T}, \tag{5}$$

which represents the LDU factorization of R_{XX}. Referring to (1), the coding gain without perturbation for the optimal causal transform can be written as

$$G_L^{(0)} = \frac{E||\tilde{Y}||^2_I}{E||\tilde{X}||^2_L} = \frac{E||\tilde{X}||^2_I}{E||\tilde{Y}||^2_L} = \left(\frac{\det[diag(R_{XX})]}{\det[diag(LR_{XX}L^T)]}\right)^{\frac{1}{N}}, \tag{6}$$

where $diag(R)$ denotes here the diagonal matrix that corresponds to the diagonal of the matrix R. Now, since the diagonalizing transformation matrix L is unimodular, $\det(diag(R_{YY})) = \det(R_{XX}) = \det\Lambda$, where Λ is the eigenvalue matrix of R_{XX}. The coding gain is

$$\begin{aligned} G_L^{(0)} &= \left(\frac{\det[diag(R_{XX})]}{\det[diag(LR_{XX}L^T)]}\right)^{\frac{1}{N}} \\ &= \left(\frac{\det[diag(R_{XX})]}{\det\Lambda}\right)^{\frac{1}{N}} = G_K^{(0)}, \end{aligned} \tag{7}$$

where K denotes a KLT of R_{XX}. Thus, for an optimal bit allocation, the coding gains of the KLT and the LDU are the same without perturbation for three reasons : both transformations ensure that the power of the quantization error is the same in the transform and in the signal domains, they are totally decorrelating transforms, and finally they are unimodular. Moreover, this is the best coding gain achievable among all unimodular transforms (a proof based on Hadamard's inequality for symmetric positive semidefinite matrices can be found in [4]).

3. QUANTIZATION EFFECTS ON THE CODING GAINS

Suppose we compute the transformation on the basis of quantized data. The statistics of the quantized data is assumed to be perfectly known in this section. In other words, we assume that the decoder disposes of an infinite number of quantized vectors X_i^q, and we know $R_{X^qX^q}$. (As stated in introduction, the perturbing effects of partially known statistics are analyzed in section 4 and 5.) Under the assumptions of high resolution (uncorrelated white noise), optimal bit assignment and unity noise gain property of the transformation, $\Delta R = E\tilde{X}\tilde{X}^T = \sigma_q^2 I$, where $\sigma_q^2 = c2^{-2r}\left(\prod_{i=1}^N \sigma_{y_i}^2\right)^{\frac{1}{N}}$ for Gaussian transform signals. Thus, one should compute

$$E||\tilde{Y}||^2_{(\hat{T},q)} = \sum_{i=1}^N c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{T}R_{X^qX^q}\hat{T}^T)_{ii}}{(\Pi_{i=1}^N(\hat{T}R_{X^qX^q}\hat{T}^T)_{ii})^{\frac{1}{N}}}]}\sigma_{y_i}^{2'}, \tag{8}$$

where q refers to quantization. Evaluating (8) for $\hat{T} = I, \hat{V}$ and $\hat{L}$ gives the following results.

3.1. Identity Transformation

In this case, the number of bits attributed to the quantizer Q_i is $r + \frac{1}{2}log_2\frac{(R_{X^qX^q})_{ii}}{(\prod_{i=1}^N(R_{X^qX^q})_{ii})^{\frac{1}{N}}}$, and the variance $\sigma_{y_i}^{2'}$ are indeed $(R_{XX})_{ii}$. Since the signals to be quantized are Gaussian, we have

$$E||\tilde{Y}||^2_{(I,q)} = \sum_{i=1}^N c2^{-2[r+\frac{1}{2}log_2\frac{(R_{X^qX^q})_{ii}}{(\Pi_{i=1}^N(R_{X^qX^q})_{ii})^{\frac{1}{N}}}]}(R_{XX})_{ii} \tag{9}$$

$$= \sum_{i=1}^N c2^{-2r}\left(\det diag\{R_{X^qX^q}\}\right)^{\frac{1}{N}}\frac{(R_{XX})_{ii}}{(R_{X^qX^q})_{ii}}. \tag{10}$$

The second equality comes from the fact that optimal bit assignment produces equal distortion on each component : suppose we compute the optimal bit assignment for the signal with covariance matrix $R_{X^qX^q}$. Then we can write

$$\sum_{i=1}^N c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{T}R_{X^qX^q}\hat{T}^T)_{ii}}{(\Pi_{i=1}^N(\hat{T}R_{X^qX^q}\hat{T}^T)_{ii})^{\frac{1}{N}}}]}(R_{X^qX^q})_{ii} = c2^{-2r}\left(\det diag\{R_{X^qX^q}\}\right)^{\frac{1}{N}} \tag{11}$$

where all $c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{T}R_{X^qX^q}\hat{T}^T)_{ii}}{(\Pi_{i=1}^N(\hat{T}R_{X^qX^q}\hat{T}^T)_{ii})^{\frac{1}{N}}}]}(R_{X^qX^q})_{ii}$ are equal (each term equals the geometric mean of all the terms). Factorizing by the geometric mean gives (10).
Now, by writing $R_{X^qX^q} = R_{XX} + \sigma_q^2 I$ in (10), one shows that

$$\sum_{i=1}^N \frac{(R_{XX})_{ii}}{(R_{X^qX^q})_{ii}} = tr\{\left(I + \sigma_q^2(diagR_{XX})^{-1}\right)^{-1}\},$$

where tr denotes the trace operator, and

$$\det\left(diagR_{X^qX^q}\right) = \det\left(diagR_{XX}\right)\det(I + \sigma_q^2(diagR_{XX})^{-1})$$

and we find

$$\begin{aligned} E||\tilde{Y}||^2_{(I,q)} \quad &= \quad E||\tilde{Y}||^2_{(I)}\frac{1}{N}(\det(I + \sigma_q^2(diagR_{XX})^{-1}))^{\frac{1}{N}} \\ &\quad tr\{\left(I + \sigma_q^2(diagR_{XX})^{-1}\right)^{-1}\}. \end{aligned} \tag{12}$$

The distortion is slightly increased because the bits allocated on the basis of variances of quantized signals are not the optimal ones (the variance of the quantization noise is not equal in each branch i). An approximation of (12) up to the second order of the perturbation gives

$$\begin{aligned} &E||\tilde{Y}||^2_{(I,q)} \\ &= c2^(-2r)(\det diag\{R_{XX}\})^{1/N}(\Pi_{i=1}^N(\frac{\sigma_q^2}{(R_{XX})_{ii}}))^{1/N}\sum_{i=1}^N(1+\frac{1}{(R_{XX})_{ii}})^{-1} \\ &\approx E||\tilde{Y}||^2_{(I)}\left[1+\frac{\sigma_q^4}{N^2}(\frac{N-1}{2}\sum_{i=1}^N\frac{1}{(R_{XX})_{ii}^2} - \sum_{i=1}^N\sum_{j>i}\frac{1}{(R_{XX})_{ii}(R_{XX})_{jj}})\right] \end{aligned} \tag{13}$$

3.2. KLT

As observed in [9] also, if V denotes a KLT of R_{XX}, then $V(R_{XX}+\sigma_q^2 I)V^T = \Lambda + \sigma_q^2 I = \Lambda^q$, and V is also a KLT of $R_{XX}+\sigma_q^2 I$. Thus, the perturbation term $\sigma_q^2 I$ on R_{XX} does not change the backward adapted transformation, and the variances of the transformed signals remain unchanged : $\sigma_{y_i}^{2'} = (VR_{XX}V^T)_{ii} = \lambda_i$. However, the decoder can only estimate the variances $(VR_{X^qX^q}V^T)_{ii} = \lambda_i + \sigma_q^2$, on the basis of which the coder has to assign the bits r_i. The transformed signals are Gaussian ($Y = TX$), thus the actual distortion is

$$E\|\tilde{Y}\|^2_{(V,q)} = \sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2 \frac{(VR_{X^qX^q}V^T)_{ii}}{(\Pi_{i=1}^N (VR_{X^qX^q}V^T)_{ii})^{\frac{1}{N}}}]} (VR_{XX}V^T)_{ii}. \tag{14}$$

$$= \sum_{i=1}^{N} c2^{-2r} \left(\det diag\{VR_{X^qX^q}V^T\}\right)^{\frac{1}{N}} \frac{(VR_{XX}V^T)_{ii}}{(VR_{X^qX^q}V^T)_{ii}}. \tag{15}$$

Since $VR_{XX}V^T$ and $VR_{X^qX^q}V^T$ are diagonal, one shows that

$$\sum_{i=1}^{N} \frac{(VR_{XX}V^T)_{ii}}{(VR_{X^qX^q}V^T)_{ii}} = tr\{\left(I+\sigma_q^2(R_{XX}^{-1})\right)^{-1}\} = tr\{\left(I+\sigma_q^2(\Lambda)^{-1}\right)^{-1}\}. \tag{16}$$

Also,

$$\det(R_{X^qX^q}) = \det(R_{XX}) \det(I+\sigma_q^2(R_{XX}^{-1})). \tag{17}$$

Finally, the distortion for the KLT with quantization noise is

$$\begin{aligned} E\|\tilde{Y}\|^2_{(V,q)} \quad &= \quad E\|\tilde{Y}\|^2_{(V)} \tfrac{1}{N} (\det(I+\sigma_q^2(\Lambda^{-1})))^{\frac{1}{N}} \\ & \qquad tr\{\left(I+\sigma_q^2(\Lambda^{-1})\right)^{-1}\}. \end{aligned} \tag{18}$$

Again, the increase in distortion comes from the perturbation occuring on the bit allocation mechanism. An expression approximating this distortion is

$$\begin{aligned} E\|\tilde{Y}\|^2_{(V,q)} \quad &= \quad c2^(-2r)(\det diag\{R_{XX}\})^{1/N} (\Pi_{i=1}^N (\tfrac{\sigma_q^2}{\lambda_i}))^{1/N} \textstyle\sum_{i=1}^N (1+\frac{1}{\lambda_i})^{-1} \\ &\approx \quad E\|\tilde{Y}\|^2_{(V)} \left[1 + \tfrac{\sigma_q^4}{N^2}(\tfrac{N-1}{2} \textstyle\sum_{i=1}^N \frac{1}{(\lambda_i)^2} - \sum_{i=1}^N \sum_{j>i} \frac{1}{\lambda_i\lambda_j})\right] \end{aligned} \tag{19}$$

Using (12) and (18) the corresponding expression for the coding gain in the unitary case with quantization noise is

$$G_{V,q} = G^0 \frac{(\det(I+\sigma_q^2(diagR_{XX})^{-1}))^{\frac{1}{N}} tr\{\left(I+\sigma_q^2(diagR_{XX})^{-1}\right)^{-1}\}}{(\det(I+\sigma_q^2(\Lambda^{-1})))^{\frac{1}{N}} tr\{\left(I+\sigma_q^2(\Lambda^{-1})\right)^{-1}\}}, \tag{20}$$

which, with (13) and (19), can be approximated as

$$G_{V,q} \approx G^0 \left[1 + \frac{\sigma_q^4}{N^2}\left(\frac{N-1}{2}\sum_{i=1}^{N}(\frac{1}{(R_{XX})_{ii}^2} - \frac{1}{(\lambda_i)^2}) - \sum_{i=1}^{N}\sum_{j>i}(\frac{1}{(R_{XX})_{ii}(R_{XX})_{jj}} - \frac{1}{\lambda_i\lambda_j})\right)\right]. \tag{21}$$

3.3. LDU

In the causal case, the coder uses a transformation L' such that $L'R_{X^qX^q}L^{'T} = R'_{YY}$. R'_{YY} is the diagonal matrix of the estimated variances involved in the bit allocation (L' and R'_{YY} are both available to the decoder). In this case, the difference vector $Y = X - \overline{L'}X^q$, the quantization noise is filtered by the rows of $\overline{L'}$, see figure 2. Note that in this case $E\|\tilde{X}\|^2_{L',q}$ still equals $E\|\tilde{Y}\|^2_{L',q}$, since $\tilde{X} = X^q - X = Y^q + \overline{L'}X^q - X = Y^q - (X - \overline{L'}X^q) = Y^q - Y = \tilde{Y}$. It can be shown that the actual variances of the signals y_i obtained with L' are $(L'R_{X^qX^q}L^{'T} - \sigma_q^2 I)_{ii}$ [7].

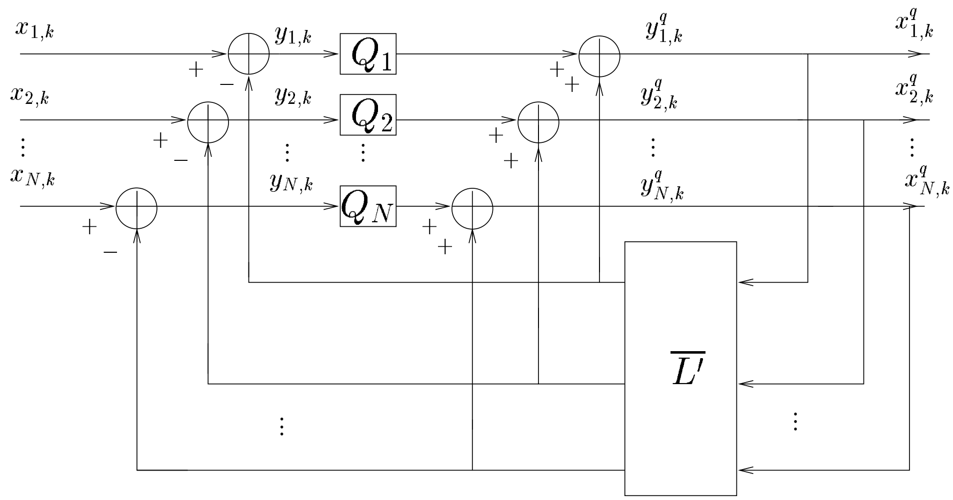

Fig. 2. Vectorial DPCM coding scheme.

We now show that evaluating (8) with $\hat{T} = L'$ yields an upper bound for $E||\tilde{Y}||^2_{(L',q)}$, and thus an lower bound for the coding gain $G_{L',q}$ in the causal case. Consider the signal in the transform domain :

$$y_{i,k} = x_{i,k} - \sum_{j=1}^{i-1} \overline{L'}_{ij} x_{i-j,k} - \sum_{j=1}^{i-1} \overline{L'}_{ij} q_{i-j,k}. \tag{22}$$

The random variable (r.v.) $\{y_i\}$ is the sum of a Gaussian r.v. $\{z_i\}$ ($z_{i,k} = x_{i,k} - \sum_{j=1}^{i-1} \overline{L}'_{ij} x_{i-j,k}$) and of $i-1$ uniform r.v.s $\{u_{ij}\}$ ($u_{ij,k} = \overline{L}'_{ij} q_{i-j,k}$). Hence, for i greater than 1, $\{y_i\}$ is not Gaussian. In order to compare the actual rate-distorsion function of the $\{y_i\}$ with that of a Gaussian r.v., denote by $H(y_i^q)$ the discrete entropy of the quantized variable y_i^q, $h(y_i)$ the differential entropy of y_i, r_i the minimum number of bits per sample necessary to losslessly code y_i^q, and Δ_i the step size of the uniform quantizer Q_i. Then we have under high resolution assumption $H(y_i^q) = r_i \approx h(y_i) - log_2\Delta_i$, where $log_2\Delta_i$ corresponds to the differential entropy of the uniformly distributed r.v. $\{q_i\}$: $log_2\Delta_i = h(q_i) = \frac{1}{2}log_2(12\sigma^2_{q_i}) = \frac{1}{2}log_2(12E||\tilde{y}_i||^2)$. Thus $r_i \approx h(y_i) - \frac{1}{2}log_2(12E||\tilde{y}_i||^2)$. Since for a given variance the Gaussian probability density function (p.d.f) maximizes the differential entropy $h(y_i)$, an upper bound for r_i may be found as $r_i < \frac{1}{2}log_2\frac{2\pi e\sigma^2_{y'_i}}{12E||\tilde{y}_i||^2}$, which gives $E||\tilde{y}_i||^2 < \frac{\pi e}{6}2^{-2r_i}\sigma^2_{y'_i}$. The distortion $E||\tilde{Y}||^2_{(L',q)}$ is then upper bounded by (8), that is,

$$\begin{aligned} E||\tilde{Y}||^2_{(L',q)} &= \textstyle\sum_{i=1}^N E||\tilde{y}_i||^2 \\ &< \textstyle\sum_{i=1}^N c2^{-2r_i}\sigma^2_{y'_i} = \sum_{i=1}^N c2^{-2[r+\frac{1}{2}log_2\frac{(R'_{YY})_{ii}}{(\Pi_{i=1}^N (R'_{YY})_{ii})^{\frac{1}{N}}}]}\sigma^{2'}_{y_i}. \end{aligned} \tag{23}$$

Let us compute this upper bound. The distortion (8) may be written as

$$E||\tilde{Y}||^{2^{up}}_{(L',q)} = \sum_{i=1}^N c2^{-2[r+\frac{1}{2}log_2\frac{(L'R_{X^qX^q}L'^T)_{ii}}{(\Pi_{i=1}^N (L'R_{X^qX^q}L'^T)_{ii})^{\frac{1}{N}}}]}(L'R_{X^qX^q}L'^T - \sigma_q^2 I)_{ii} \tag{24}$$

$$= \sum_{i=1}^N c2^{-2r}\left(\det diag\{L'R_{X^qX^q}L'^T\}\right)^{\frac{1}{N}}\left(1 - \frac{(\sigma_q^2 I)_{ii}}{(L'R_{X^qX^q}L'^T)_{ii}}\right). \tag{25}$$

Since the transformation L' is unimodular, the determinant in the previous expression equals the determinant in (17). The sum in (25) may be written as $tr\{(I - \sigma_i^2 L'R_{X^qX^q}L'^T)\} = tr\{(I - \sigma_i^2 R'_{YY})\}$. Thus this bound becomes

$$E||\tilde{Y}||^{2^{up}}_{(L',q)} = E||\tilde{Y}||^2_{(L)}\frac{1}{N}(\det(I + \sigma_q^2(\Lambda^{-1})))^{\frac{1}{N}}tr\{\left(I + \sigma_q^2(R'^{-1}_{YY})\right)\}. \tag{26}$$

The increase in distortion comes not only from the perturbation occuring on the bit allocation mechanism but also from the filtering of the quantization noise. Up to the first order of perturbation,we obtain

$$
\begin{aligned}
E\|\tilde{Y}\|^{2^{up}}_{(L',q)} &= c2^{(}-2r)(\det diag\{R_{XX}\})^{1/N}(\Pi_{i=1}^{N}(\frac{\sigma_q^2}{\lambda_i}))^{1/N}\sum_{i=1}^{N}(1+\frac{1}{R'_{Y_iY_i}})^{-1} \\
&\approx E\|\tilde{Y}\|^2_{(K)}\left[1+\frac{\sigma_q^2}{N}(\sum_{i=1}^{N}\frac{1}{\lambda_i}-\frac{1}{\sigma_{y_i}^2})\right],
\end{aligned}
\tag{27}
$$

where the $\sigma_{y_i}^2$ correspond to optimal prediction error variance in absence of quantization noise.
Now, the gain in the causal case $G_{L',q}=\frac{E\|\tilde{Y}\|^2_{(I,q)}}{E\|\tilde{Y}\|^2_{(L',q)}}$ becomes lower bounded by $\frac{E\|\tilde{Y}\|^2_{(I,q)}}{\sum_{i=1}^{N}c2^{-2r_i}\sigma^2_{y'_i}}=G^{low}_{L',q}$. The corresponding exact expression for the bound on the coding gain is

$$
G^{low}_{L',q}=G^0\frac{(\det(I+\sigma_q^2(diagR_{XX})^{-1}))^{\frac{1}{N}}tr\{\left(I+\sigma_q^2(diagR_{XX})^{-1}\right)^{-1}\}}{(\det(I+\sigma_q^2(\Lambda^{-1})))^{\frac{1}{N}}tr\{\left(I+\sigma_q^2(R'^{-1}_{YY})\right)\}}.
\tag{28}
$$

Up to the first order of perturbation we get,

$$
G^{low}_{L',q}\approx G^0\left[1-\frac{\sigma_q^2}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i}-\frac{1}{\sigma_{y_i}^2}\right].
\tag{29}
$$

Now how close is the actual coding gain in the causal case to this bound ? First recall that the r.v. $\{y_1\}$ is strictly Gaussian. Moreover, the convolution of $i-1$ Uniform p.d.f. is known to tend quickly, as i grows, to a Gaussian p.d.f.. Thus, $\{y_i\}$ tends to be Gaussian and for reasonably high N, this bound is a fairly precise measure of the actual coding gain : $G_{L',q}\approx G^{low}_{L',q}$. As developped in [5], an interesting consequence of (29) is that we should decorrelate the signals $\{x_i\}$ in order of decreasing variance if we want $G_{L',q}$ to be maximized (see the simulations in the last section).

4. ESTIMATION NOISE

We analyze in this section the coding gains of a backward adaptive scheme based on an estimate of the covariance matrix $\hat{R}_{XX}=\frac{1}{K}\sum_{i=1}^{N}X_iX_i^T$. We assume independent identically distributed (i.i.d.) real vectors X_i, which is for example the case if the sampling period of the scalar signals is high in comparison with their typical correlation time. Thus, the first and second order statistics of ΔR are known : $(\Delta R)_{ii}$ is, for sufficiently high K, a zero mean Gaussian random variable with covariance matrix such that

$$
Evec(\Delta R)\left(vec(\Delta R)\right)^T=\frac{2}{K}R_{XX}\otimes R_{XX},
\tag{30}
$$

where $\otimes$ denotes the Kronecker product. For each realization of ΔR, the coder computes a transformation $\hat{T}$ which diagonalizes $\hat{R}_{XX}:\hat{T}\hat{R}_{XX}\hat{T}=\hat{R}_{YY}$. The number of bits assigned to each component is $r_i=r+\frac{1}{2}log_2\frac{(\hat{T}\hat{R}_{XX}\hat{T}^T)_{ii}}{(\prod_{i=1}^{N}(\hat{T}\hat{R}_{XX}\hat{T}^T)_{ii})^{\frac{1}{N}}}$. Now, the true variances of the signals obtained by applying $\hat{T}$ to X are $(\hat{T}R_{XX}\hat{T}^T)_{ii}$. Note that in the causal case, $Y=I-\hat{\overline{L}}X=\hat{L}X$, so that $R'_{YY}=\hat{L}R_{XX}\hat{L}^T$. In the causal case, there is a qualitative difference with the previous section, where the quantization noise was filtered by the predictors of $\overline{L'}$. Here, the estimation noise does not perturb signals, but only transformations and bit assignments. The transformed signals are Gaussian for the three transformations, and the resulting distortion by estimating T and R_{XX} by means of K vectors is

$$
E\|\tilde{Y}\|^2_{(\hat{T},K)}=E\sum_{i=1}^{N}c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{T}\hat{R}_{XX}\hat{T}^T)_{ii}}{(\Pi_{i=1}^{N}(\hat{T}\hat{R}_{XX}\hat{T}^T)_{ii})^{\frac{1}{N}}}]}(\hat{T}R_{XX}\hat{T}^T)_{ii}.
\tag{31}
$$

4.1. Identity Transformation

With $\hat{T}=I$, we obtain

$$
E\|\tilde{Y}\|^2_{(I,K)}=E\sum_{i=1}^{N}c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{R}_{XX})_{ii}}{(\Pi_{i=1}^{N}(\hat{R}_{XX})_{ii})^{\frac{1}{N}}}]}(R_{XX})_{ii}.
\tag{32}
$$

Using a similar analysis as in section 3, we obtain

$$\begin{aligned} E\|\tilde{Y}\|^2_{(I,k)} &= Ec2^{-2r}\left(\det diag\{R_{XX}\}\right)^{\frac{1}{N}}\left(\prod_{i=1}^N(1+\tfrac{(\Delta R)_{ii}}{(R_{XX})_{ii}})\right)^{\frac{1}{N}}\sum_{i=1}^N\left(1+\tfrac{(\Delta R)_{ii}}{(R_{XX})_{ii}}\right)^{-1} \\ &\approx E\|\tilde{Y}\|^2_{(I)}\left(1+\ E\ \tfrac{N-1}{2N^2}\sum_{i=1}^N(\tfrac{(\Delta R)_{ii}}{(R_{XX})_{ii}})^2-\ E\ \tfrac{1}{N^2}\sum_i\sum_{j>i}\tfrac{(\Delta R)_{ii}}{(R_{XX})_{ii}}\tfrac{(\Delta R)_{jj}}{(R_{XX})_{jj}}\right) \end{aligned} \tag{33}$$

With (30), the second expectation in (33) may be written as

$$E\ \frac{N-1}{2N^2}\sum_{i=1}^N\left(\frac{(\Delta R)_{ii}}{(R_{XX})_{ii}}\right)^2=\frac{N-1}{2N^2}\sum_{i=1}^N\frac{2(R_{XX})^2_{ii}}{K(R_{XX})^2_{ii}}=\frac{N-1}{2N^2}\frac{2N}{K}=\frac{N-1}{NK}. \tag{34}$$

The third expectation is

$$\begin{aligned} E\ \tfrac{1}{N^2}\sum_i\sum_{j>i}\tfrac{(\Delta R)_{ii}}{(R_{XX})_{ii}}\tfrac{(\Delta R)_{jj}}{(R_{XX})_{jj}} &\approx \tfrac{2}{K}\sum_i\sum_{j>i}\tfrac{(R_{XX})^2_{ij}}{(R_{XX})_{ii}(R_{XX})_{jj}} \\ &\approx \tfrac{2}{K}\|\rhd\left((diag\{R_{XX}\})^{1/2}R_{XX}(diag\{R_{XX}\})^{1/2}\right)\|^2_F \end{aligned} \tag{35}$$

where $\rhd(A)$ denotes the stritcly lower triangular matrix made with the strictly lower triangular part of A, and $\|A\|^2_F$ the squared Frobenius norm of A. If D denotes $diag\{R_{XX}\}$, we obtain

$$E\ \frac{1}{N^2}\sum_i\sum_{j>i}\frac{(\Delta R)_{ii}}{(R_{XX})_{ii}}\frac{(\Delta R)_{jj}}{(R_{XX})_{jj}}\approx\frac{1}{K}\left(\|D^{-\frac{1}{2}}R_{XX}D^{-\frac{1}{2}}\|^2_F-\|diag\{D^{-\frac{1}{2}}R_{XX}D^{-\frac{1}{2}}\|^2_F\right)=\frac{1}{K}(tr\{R_{XX}D^{-1}R_{XX}D^{-1}\}). \tag{36}$$

Finally, the expected distortion for Identity with estimation noise is, for sufficiently high K,

$$E\|\tilde{Y}\|^2_{(I,K)}\approx E\|\tilde{Y}\|^2_{(I)}\left(1+\frac{1}{K}[1-\frac{1}{N^2}tr\{R_{XX}D^{-1}R_{XX}D^{-1}\}]\right). \tag{37}$$

4.2. KLT

In the unitary case, the expected distortion is

$$E\|\tilde{Y}\|^2_{(\hat{V},K)}=E\sum_{i=1}^N c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{V}\hat{R}_{XX}\hat{V}^T)_{ii}}{(\Pi_{i=1}^N(\hat{V}\hat{R}_{XX}\hat{V}^T)_{ii})^{\frac{1}{N}}}]}(\hat{V}R_{XX}\hat{V}^T)_{ii}. \tag{38}$$

Using the fact $\hat{V}\hat{R}_{XX}\hat{V}^T$ is diagonal, we can write (38) as

$$E\|\tilde{Y}\|^2_{(\hat{V},K)} = Ec2^{-2r}\left(\det\hat{V}\hat{R}_{XX}\hat{V}^T\right)^{\frac{1}{N}}\sum_{i=1}^N\tfrac{(\hat{V}R_{XX}\hat{V}^T)_{ii}}{(\hat{V}\hat{R}_{XX}\hat{V}^T)_{ii}}. \tag{39}$$

Because of the unimodularity of $\hat{V}$, the determinant in (39) may be written as

$$\begin{aligned} \det\hat{V}\hat{R}_{XX}\hat{V}^T &= \det(R_{XX}+\Delta R) \\ &= (\det R_{XX})\det(I+R^{-1}_{XX}\Delta R). \end{aligned} \tag{40}$$

The sum in (39) may be written as

$$\begin{aligned} \sum_{i=1}^N\tfrac{(\hat{V}R_{XX}\hat{V}^T)_{ii}}{(\hat{V}\hat{R}_{XX}\hat{V}^T)_{ii}} &= tr\{(\hat{V}\hat{R}_{XX}\hat{V}^T)^{-1/2}\hat{V}R_{XX}\hat{V}^T(\hat{V}\hat{R}_{XX}\hat{V}^T)^{-1/2}\} \\ &= tr\{\hat{V}R_{XX}\hat{V}^T(\hat{V}\hat{R}_{XX}\hat{V}^T)^{-1}\}=tr\{\hat{V}R_{XX}\hat{R}^{-1}_{XX}\hat{V}^{-1}\} \\ &= tr\{R_{XX}\hat{R}^{-1}_{XX})\}=tr\{R_{XX}(R_{XX}+\Delta R)^{-1}\} \\ &= tr\{(I+R^{-1}_{XX}\Delta R)^{-1}\}. \end{aligned} \tag{41}$$

Thus, (39) is equivalent to

$$E\|\tilde{Y}\|^2_{(\hat{V},K)} = E\|\tilde{Y}\|^2_{(V)}\tfrac{1}{N}\ E\ (\det(I+R^{-1}_{XX}\Delta R))^{\frac{1}{N}}tr\{\left(I+R^{-1}_{XX}\Delta R\right)^{-1}\}. \tag{42}$$

In order to find the expectation of $E||\tilde{Y}||^2_{(\hat{V},K)}$, let us develop (38) as

$$E||\tilde{Y}||^2_{(\hat{V},K)} = Ec2^{-2r}\left(\det \hat{V}R_{XX}\hat{V}^T\right)^{\frac{1}{N}}\left(\prod_{i=1}^{N}(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{XX}\hat{V}^T)_{ii}})\right)^{\frac{1}{N}}\sum_{i=1}^{N}\left(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{XX}\hat{V}^T)_{ii}}\right)^{-1}. \quad (43)$$

The determinant in (43) may also be written as

$$\begin{aligned}\det \hat{V}R_{XX}\hat{V}^T &= \prod_{i=1}^{N}(\lambda_i+\delta_{\lambda_i}) \\ &= (\prod_{i=1}^{N}\lambda_i)(\prod_{i=1}^{N}1+\frac{\delta_{\lambda_i}}{\lambda_i}) \\ &\approx (\prod_{i=1}^{N}\lambda_i)(1+\sum_{i=1}^{N}\frac{\delta_{\lambda_i}}{\lambda_i}),\end{aligned} \quad (44)$$

where λ_i ans δ_{λ_i} are the diagonal elements of Λ and $\Delta\Lambda$, which is defined by

$$\hat{V}R_{XX}\hat{V}^T = \Lambda + \Delta\Lambda. \quad (45)$$

Now (43) may then be approximated as

$$\begin{aligned}E||\tilde{Y}||^2_{(\hat{V},K)} &= E||\tilde{Y}||^2_{(K)}\left(1+\frac{1}{N}\sum_{i=1}^{N}\frac{\delta_{\lambda_i}}{\lambda_i}\right)\left(\prod_{i=1}^{N}(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{XX}\hat{V}^T)_{ii}})\right)^{\frac{1}{N}}\sum_{i=1}^{N}\left(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{XX}\hat{V}^T)_{ii}}\right)^{-1} \\ &\approx E||\tilde{Y}||^2_{(K)}\left(1+\ E\frac{1}{N}\sum_{i=1}^{N}\frac{\delta_{\lambda_i}}{\lambda_i}+\ E\ \frac{N-1}{2N^2}\sum_{i=1}^{N}(\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{XX}\hat{V}^T)_{ii}})^2-\ E\ \frac{1}{N^2}\sum_i\sum_{j>i}\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{XX}\hat{V}^T)_{ii}}\frac{(\hat{V}\Delta R\hat{V}^T)_{jj}}{(\hat{V}R_{XX}\hat{V}^T)_{jj}}\right).\end{aligned} \quad (46)$$

Computation of the second expectation in (46).
Using the unitarity of $\hat{V} = V + \Delta V$, we have

$$diag\{\Delta\Lambda\} = diag\{\Delta V^T R_{XX}\Delta V - \Delta V^T\Delta V\Lambda\}. \quad (47)$$

The expectation of the diagonal elements of the first term in (47) is

$$E\ (\Delta V^T R_{XX}\Delta V)_{ii} = tr\ E\ R_{XX}\Delta V_i\Delta V_i^T = \frac{\lambda_i}{K}\sum_{j\neq i}\frac{\lambda_j^2}{(\lambda_j-\lambda_i)^2}, \quad (48)$$

where we have used the following classical result in perturbation theory of matrices [1], for sufficiently high K

$$E\ \Delta V_i\Delta V_i^T = \frac{\lambda_i}{K}\sum_{j\neq i}\frac{\lambda_j}{(\lambda_j-\lambda_i)^2}T_jT_j^T. \quad (49)$$

The expectation of the diagonal elements of the second term in (47) is

$$E\ (\Delta V^T\Delta V\Lambda)_{ii} = \lambda_i E(\Delta V_i^T\Delta V_i) = \frac{\lambda_i^2}{K}\sum_{j\neq i}\frac{\lambda_j}{(\lambda_j-\lambda_i)^2}. \quad (50)$$

Hence, we get from (47,48), and (50)

$$E\delta\lambda_i = \frac{\lambda_i}{K}\sum_{j\neq i}\frac{\lambda_j^2-\lambda_i\lambda_j}{(\lambda_j-\lambda_i)^2} = \frac{\lambda_i}{K}\sum_{j\neq i}\frac{\lambda_j}{\lambda_j-\lambda_i}, \quad (51)$$

from which it is easy to show that

$$E\ \sum_{i=1}^{N}\frac{\delta\lambda_i}{\lambda_i} = \frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}\frac{\lambda_j}{\lambda_j-\lambda_i} = \frac{1}{K}\frac{N(N-1)}{2}. \quad (52)$$

Computation of the second and third expectations in (46).

The perturbing term $\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{XX}\hat{V}^T)_{ii}}$ may also be approximated as

$$\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{XX}\hat{V}^T)_{ii}} = \frac{V\Delta RV^T + V\Delta R\Delta V^T + \Delta V\Delta RV^T + \Delta V\Delta R\Delta V^T}{VR_{XX}V^T + VR_{XX}\Delta V^T + \Delta VR_{XX}V^T + \Delta VR_{XX}\Delta V^T} \approx \frac{V\Delta RV^T}{VR_{XX}V^T} \tag{53}$$

Let $\hat{\Lambda}$ be

$$\begin{aligned} \hat{\Lambda} &= V\hat{R}_{XX}V^T = VR_{XX}V^T + V\Delta RV^T \\ &= \tfrac{1}{K}\textstyle\sum_{i=1}^{N} VX_iX_i^TV^T \\ &= \tfrac{1}{K}\textstyle\sum_{i=1}^{N} Y_iY_i^T. \end{aligned} \tag{54}$$

Thus, $V\Delta RV^T = \Lambda - \frac{1}{K}\sum_{i=1}^{N} Y_iY_i^T$: $(V\Delta RV^T)_{ii}$ is a real zero mean Gaussian random variable, corresponding to the estimation error of Λ obtained with a covariance matrix computed with K vectors. Hence, we have $Evec(V\Delta RV^T)(vecV\Delta RV^T)^T = \frac{2\Lambda\otimes\Lambda}{K}$, whence

$$\sum_{i=1}^{N}\left(\frac{(V\Delta RV^T)_{ii}}{(VR_{XX}V^T)_{ii}}\right)^2 = \frac{2N}{K}, \tag{55}$$

and

$$\sum_{i}\sum_{j>i}\frac{(V\Delta RV^T)_{ii}}{(VR_{XX}V^T)_{ii}}\frac{(V\Delta RV^T)_{jj}}{(VR_{XX}V^T)_{jj}} = \sum_{i}\sum_{j>i}\frac{2}{K}\frac{\Lambda_{ij}^2}{\lambda_i\lambda_j} = 0. \tag{56}$$

Finally, the expected distortion for the KLT when the transformation is based on K vectors is, under high resolution assumption

$$E||\tilde{Y}||^2_{(\hat{V},K)} \approx E||\tilde{Y}||^2_{(V)}\left(1 + \frac{N-1}{K}\left[\frac{1}{2} + \frac{1}{N}\right]\right). \tag{57}$$

The associated coding gain is

$$G_{\hat{V},K} = \frac{E||\tilde{Y}||^2_{(I,K)}}{E||\tilde{Y}||^2_{(\hat{V},K)}} \approx G^0\left(1 - \frac{1}{K}\left[\frac{tr\{RD^{-1}RD^{-1}\}}{N^2} + \frac{N-1}{2} - \frac{1}{N}\right]\right), \tag{58}$$

where $D = diag\{R_{XX}\}$.

4.3. LDU

As stated in the introduction of this section, the expected distortion with $\hat{L}$ computed on $\hat{R}_{XX}$ is

$$\begin{aligned} E||\tilde{Y}||^2_{(\hat{L},K)} &= E\textstyle\sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{L}\hat{R}_{XX}\hat{L}^T)_{ii}}{(\Pi_{i=1}^{N}(\hat{L}\hat{R}_{XX}\hat{L}^T)_{ii})^{\frac{1}{N}}}]}(\hat{L}R_{XX}\hat{L}^T)_{ii} \\ &= Ec2^{-2r}\left(\det\hat{V}\hat{R}_{XX}\hat{V}\right)^{\frac{1}{N}}\textstyle\sum_{i=1}^{N}\frac{(\hat{L}R_{XX}\hat{L}^T)_{ii}}{(\hat{L}\hat{R}_{XX}\hat{L}^T)_{ii}}, \end{aligned} \tag{59}$$

where we used a factorization similar to (10). Now using the unimodularity of $\hat{L}$, we can write the determinant as

$$\left(\det\hat{V}\hat{R}_{XX}\hat{V}\right)^{\frac{1}{N}} = \det\hat{R}_{XX} = \det(R_{XX})\det(I + R_{XX}^{-1}\Delta R), \tag{60}$$

and using the decorrelating property of $\hat{L}$, we can write the sum as

$$\sum_{i=1}^{N}\frac{(\hat{L}R_{XX}\hat{L}^T)_{ii}}{(\hat{L}\hat{R}_{XX}\hat{L}^T)_{ii}} = tr\{(I + R_{XX}^{-1}\Delta R)^{-1}\}. \tag{61}$$

Thus, comparing with (38), we have $E||\tilde{Y}||^2_{(\hat{L},K)} = E||\tilde{Y}||^2_{(\hat{V},K)}$: the distortion and coding gain with quantization noise are the same in the causal and the unitary cases, and are given for high K by (38) and (58) respectively.

5. QUANTIZATION AND ESTIMATION NOISE

We arrive now to the most general case of this comparison between causal and unitary approaches. As stated in Introduction, in a real backward adaptive scheme, the coder should attribute the bits on the basis of $\hat{R}_{X^qX^q} = R_{X^qX^q} + \Delta R = \frac{1}{K}\sum_{i=1}^{K} X_i^q X_i^{qT}$. As in the previous section, we assume independent identically distributed real vectors X_i. The estimated transform is $\hat{T}$, such that $\hat{T}\hat{R}_{X^qX^q}\hat{T}^T$ is a diagonal matrix, which corresponds to the estimated variances of the transformed signals. If we continue denoting by $\sigma_{y_i}^{2'}$ the actual variances of the transformed signals (obtained by applying $\hat{T}$ to X), the expected distortion obtained with $\hat{T}$ using K quantized vectors is

$$E\|\tilde{Y}\|^2_{(\hat{T},K,q)} = E\sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{T}\hat{R}_{X^qX^q}\hat{T}^T)_{ii}}{(\Pi_{i=1}^{N}(\hat{T}\hat{R}_{X^qX^q}\hat{T}^T)_{ii})^{\frac{1}{N}}}]}(\hat{T}R_{XX}\hat{T}^T)_{ii}, \tag{62}$$

where the subscripts q and K refer to the presence of quantization and estimation noise, and the constant c assumes Gaussianity of the transformed signals. Equation (62) must be evaluated for Identity, KLT and LDU transforms. As in Section 3, the computation of (62) in the causal case will provide an upper bound for the distortion because of the uniform quantization noise feedback uppon the $\{y_i\}$.

5.1. Identity Transformation

In this case the transformed signals y_i are indeed still Gaussian. With $\hat{T} = I$ we obtain for (62), by writing $R_{XX} = R_{X^qX^q} - \sigma_q^2 I$,

$$E\|\tilde{Y}\|^2_{(I,K,q)} = E\sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{R}_{X^qX^q})_{ii}}{(\Pi_{i=1}^{N}(\hat{R}_{X^qX^q})_{ii})^{\frac{1}{N}}}]}(R_{X^qX^q})_{ii} - \sigma_q^2 E\sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2\frac{(\hat{R}_{X^qX^q})_{ii}}{(\Pi_{i=1}^{N}(\hat{R}_{X^qX^q})_{ii})^{\frac{1}{N}}}]}. \tag{63}$$

The first term may be written as

$$\begin{aligned}&c2^{-2[r+\frac{1}{2}log_2\frac{(R_{X^qX^q})_{ii}}{(\Pi_{i=1}^{N}(R_{X^qX^q})_{ii})^{\frac{1}{N}}}]}(R_{X^qX^q})_{ii}E\left(\prod_{i=1}^{N}1+\frac{(\Delta R)_{ii}}{(R_{X^qX^q})_{ii}}\right)^{\frac{1}{N}}\sum_{i=1}^{N}\left(1+\frac{(\Delta R)_{ii}}{(R_{X^qX^q})_{ii}}\right)^{-1}\\ &\approx E\|\tilde{Y}\|^2_{(I)}c2^{-2r}(\det(diag\{R_{X^qX^q}\}))^{\frac{1}{N}}\left(1+\ E\ \frac{N-1}{2N^2}\sum_{i=1}^{N}(\frac{(\Delta R)_{ii}}{(R_{X^qX^q})_{ii}})^2-\ E\ \frac{1}{N^2}\sum_i\sum_{j>i}\frac{(\Delta R)_{ii}}{(R_{X^qX^q})_{ii}}\frac{(\Delta R)_{jj}}{(R_{X^qX^q})_{jj}}\right).\end{aligned} \tag{64}$$

The equality concerning the determinant comes from a factorization similar to (10). The expectations in (64) are computed in the same manner as in section 4.1. Note however that in this case, the r.v. ΔR corresponds to the estimation error of $R_{X^qX^q}$, which is not the covariance matrix of jointly Gaussian i.i.d. r.v.s because of the uniformly distributed quantization noise q_i perturbing the x_i. Since this perturbation is small (and the vectors X_i^q are still i.i.d.), we assume that ΔR can be considered as a zero mean r.v. with covariance matrix

$$Evec(\Delta R)\,(vec(\Delta R))^T \approx \frac{2}{K}R_{X^qX^q}\otimes R_{X^qX^q}. \tag{65}$$

With (65), the second expectation in (64) may be approximated as

$$E\ \frac{N-1}{2N^2}\sum_{i=1}^{N}\left(\frac{(\Delta R)_{ii}}{(R_{X^qX^q})_{ii}}\right)^2 \approx \frac{N-1}{2N^2}\sum_{i=1}^{N}\frac{2(R_{X^qX^q})_{ii}^2}{K(R_{X^qX^q})_{ii}^2} = \frac{N-1}{2N^2}\frac{2N}{K} = \frac{N-1}{NK}, \tag{66}$$

and the third expectation as

$$\begin{aligned}E\ \frac{1}{N^2}\sum_i\sum_{j>i}\frac{(\Delta R)_{ii}}{(R_{X^qX^q})_{ii}}\frac{(\Delta R)_{jj}}{(R_{X^qX^q})_{jj}} &\approx \frac{2}{K}\sum_i\sum_{j>i}\frac{(R_{X^qX^q})_{ij}^2}{(R_{X^qX^q})_{ii}(R_{X^qX^q})_{jj}}\\ &\approx \frac{2}{K}\|\triangleright\left((diag\{R_{X^qX^q}\})^{1/2}R_{X^qX^q}(diag\{R_{X^qX^q}\})^{1/2}\right)\|_F^2.\end{aligned} \tag{67}$$

If D^q denotes $diag\{R_{X^qX^q}\}$, we obtain

$$\begin{aligned}E\ \frac{1}{N^2}\sum_i\sum_{j>i}\frac{(\Delta R)_{ii}}{(R_{X^qX^q})_{ii}}\frac{(\Delta R)_{jj}}{(R_{X^qX^q})_{jj}} &\approx \frac{1}{K}\left(\|(D^q)^{-\frac{1}{2}}R_{XX}(D^q)^{-\frac{1}{2}}\|_F^2 - \|diag\{(D^q)^{-\frac{1}{2}}R_{XX}(D^q)^{-\frac{1}{2}}\}\|_F^2\right)\\ &\approx \frac{1}{K}(tr\{R_{XX}(D^q)^{-1}R_{XX}(D^q)^{-1}\}).\end{aligned} \tag{68}$$

The second term in (63) is small because of σ_q^2, and we neglect the estimation errors in this term (estimation errors being itself small), so that we make the approximation

$$\sigma_q^2 E \sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2 \frac{(\hat{R}_{X^qX^q})_{ii}}{(\Pi_{i=1}^{N}(\hat{R}_{X^qX^q})_{ii})^{\frac{1}{N}}}]} \approx \sigma_q^2 c2^{-2r} \left(\det(diag\{R_{X^qX^q}\})\right)^{\frac{1}{N}} \left[-\sum_{i=1}^{N} \frac{1}{(R_{X^qX^q})_{ii}}\right]. \tag{69}$$

Finally, using (64),(68) and (69), the expected distortion for Identity transform with quantization and estimation noise may, for sufficiently high resolution and K, be written as

$$E\|\tilde{Y}\|^2_{(I,K,q)} \approx E\|\tilde{Y}\|^2_{(I)} \left(\det(I+\sigma_q^2(diag\{R_{XX}\})^{-1})\right)^{1/N} \left[1+\frac{1}{K}\left[1-\frac{tr\{R_{X^qX^q}D^{q-1}R_{X^qX^q}D^{q-1}\}}{N^2}\right]-\frac{\sigma_q^2}{N}tr\{(diagR_{X^qX^q})^{-1}\}\right]. \tag{70}$$

5.2. KLT

In the unitary case also, expression (62) with $\hat{T} = \hat{V}$ gives the exact expression of the distortion since each transform coefficient is Gaussian (linear combination of Gaussian r.v.s.). The expected distortion with quantization and estimation noise is

$$E\|\tilde{Y}\|^2_{(\hat{V},K,q)} = E\sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2 \frac{(\hat{V}\hat{R}_{X^qX^q}\hat{V}^T)_{ii}}{(\Pi_{i=1}^{N}(\hat{V}\hat{R}_{X^qX^q}\hat{V}^T)_{ii})^{\frac{1}{N}}}]} (\hat{V}R_{XX}\hat{V}^T)_{ii}. \tag{71}$$

By writing $\hat{V}R_{XX}\hat{V}^T = \hat{V}R_{X^qX^q}\hat{V}^T - \sigma_q^2\hat{V}\hat{V}^T = \hat{V}R_{X^qX^q}\hat{V}^T - \sigma_q^2 I$, we get

$$\begin{aligned} E\|\tilde{Y}\|^2_{(\hat{V},K,q)} &= E\sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2 \frac{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}}{(\Pi_{i=1}^{N}(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii})^{\frac{1}{N}}}]} \left(\prod_{i=1}^{N}(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}})\right)^{\frac{1}{N}} \\ &\quad \times \left(\frac{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}} - \frac{\sigma_q^2}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}}\right)\left(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}}\right) \\ &= Ec2^{-2r}\left(\det \hat{V}R_{X^qX^q}\hat{V}^T\right)^{\frac{1}{N}} \left(\prod_{i=1}^{N}(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}})\right)^{\frac{1}{N}} \\ &\quad \times \left[\sum_{i=1}^{N}\left(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}}\right)^{-1} - \sigma_q^2\sum_{i=1}^{N}((\hat{V}R_{X^qX^q}\hat{V}^T)_{ii})^{-1}\right] \end{aligned} \tag{72}$$

Now, let $\widehat{\Lambda^q} = \hat{V}R_{X^qX^q}\hat{V}^T = \Lambda^q + \Delta\Lambda^q = \Lambda + \sigma_q^2 I + \Delta\Lambda^q$, and let $\delta^q_{\lambda_i}$ be the diagonal elements of $\Delta\Lambda^q$. Then, the first term of (72) may be approximated as

$$\begin{aligned} &E\|\tilde{Y}\|^2_{(K)}\left(1+\frac{1}{N}\sum_{i=1}^{N}\frac{\delta^q_{\lambda_i}}{(\Lambda^q)_{ii}}\right)\left(\prod_{i=1}^{N}(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}})\right)^{\frac{1}{N}}\sum_{i=1}^{N}\left(1+\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}}\right)^{-1} \\ &\approx E\|\tilde{Y}\|^2_{(K)}\left(1+ E\frac{1}{N}\sum_{i=1}^{N}\frac{\delta^q_{\lambda_i}}{(\Lambda^q)_{ii}} + E\,\frac{N-1}{2N^2}\sum_{i=1}^{N}(\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}})^2 - E\,\frac{1}{N^2}\sum_i\sum_{j>i}\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}}\frac{(\hat{V}\Delta R\hat{V}^T)_{jj}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{jj}}\right). \end{aligned} \tag{73}$$

Using an analysis similar to (46) and using the same classical result in pertubation theory of matrices as in the previous section [1], one can show that

$$E\sum_{i=1}^{N}\frac{\delta\lambda_i^q}{(\Lambda^q)_{ii}} = \frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}\frac{\lambda_j^q}{\lambda_j^q - \lambda_i^q} = \frac{1}{K}\frac{N(N-1)}{2}. \tag{74}$$

Also, the expectation of the second term in (73) may be computed as in (46). By using the approximation

$$\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}} \approx \frac{(V\Delta RV^T)_{ii}}{(VR_{X^qX^q}V^T)_{ii}}, \tag{75}$$

(that is, by keeping only linear terms in ΔR) and writing $V\Delta RV^T$ as $\Lambda^q - \frac{1}{K}\sum_{i=1}^N Y_i^q Y_i^{qT}$, the random variable $(V\Delta RV^T)_{ii}$ corresponds now to the estimation error of Λ^q obtained with a covariance matrix computed with K quantized vectors. As in (65) however, this is an approximation since the y_i^q are not Gaussian. Thus we assume $Evec(V\Delta RV^T)(vec(V\Delta RV^T))^T \approx \frac{2\Lambda^q \otimes \Lambda^q}{K}$, whence

$$\sum_{i=1}^N \left(\frac{(V\Delta RV^T)_{ii}}{(VR_{XX}V^T)_{ii}}\right)^2 \approx \frac{2N}{K}, \tag{76}$$

and

$$\sum_i \sum_{j>i} \frac{(V\Delta RV^T)_{ii}}{(VR_{X^qX^q}V^T)_{ii}} \frac{(V\Delta RV^T)_{jj}}{(VR_{X^qX^q}V^T)_{jj}} = \sum_i \sum_{j>i} \frac{2}{K} \frac{(\Lambda^q)_{ij}^2}{\lambda_i^q \lambda_j^q} \approx 0. \tag{77}$$

Thus, using the unimodularity of $\hat{T}$, the first term becomes

$$\begin{aligned} & c2^{-2r}(\det R_{X^qX^q})^{1/N}\left[1 + \tfrac{N(N-1)}{2K} + \tfrac{2N(N-1)}{2N^2K}\right] \\ & \approx c2^{-2r}(\det R_{XX})^{1/N}\left(\det(I + \sigma_q^2(R_{XX})^{-1})\right)^{1/N}\left[1 + \tfrac{1}{K}(\tfrac{N-1}{2} + \tfrac{N-1}{N})\right]. \end{aligned} \tag{78}$$

The second term in (72) may approximated under the assumptions of high resolution and high K as

$$\begin{aligned} & \sigma_q^2 E c^{-2r}(\det R_{X^qX^q})^{\frac{1}{N}} \\ & \underbrace{\left(1 + \frac{1}{N}\sum_{i=1}^N \frac{\delta_{\lambda_i}^q}{\lambda_i^q}\right)\left(\prod_{i=1}^N (1 + \frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}})\right)^{\frac{1}{N}} \sum_{i=1}^N \frac{1}{(\hat{V}R_{X^qX^q}\hat{V})_{ii}}(1 - \frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}} + (\frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}})^2)\left(1 + \frac{(\hat{V}\Delta R\hat{V}^T)_{ii}}{(\hat{V}R_{X^qX^q}\hat{V}^T)_{ii}}\right)^{-1}}_{\approx \sum_{i=1}^N \frac{1}{(VR_{X^qX^q}V^T)_{ii}}} \\ & \approx \sigma_q^2 E c^{-2r}(\det R_{X^qX^q})^{\frac{1}{N}} tr\{(\Lambda^q)^{-1}\} \\ & = \sigma_q^2 c 2^{-2r}(\det R_{XX})^{1/N}\left(\det(I + \sigma_q^2(R_{XX})^{-1})\right)^{1/N} tr\{(\Lambda^q)^{-1}\}. \end{aligned} \tag{79}$$

Finally, the expected distortion for the KLT when the transformation is based on K quantized vectors is for high K and under high resolution assumption

$$E\|\tilde{Y}\|^2_{(\hat{V},K,q)} \approx E\|\tilde{Y}\|^2_{(K)}\left(\det(I + \sigma_q^2(R_{XX})^{-1})\right)^{1/N} \times \left[1 + \frac{N-1}{K}\left[\frac{1}{2} + \frac{1}{N}\right] - \frac{\sigma_q^2}{N} tr\{(\Lambda^q)^{-1}\}\right]. \tag{80}$$

The corresponding expression for the coding gain is

$$\begin{aligned} G_{(\hat{V},K,q)} &= \frac{E\|\tilde{I}\|^2_{(\hat{V},K,q)}}{E\|\tilde{Y}\|^2_{(\hat{V},K,q)}} \\ &\approx G^0 \frac{\left(\det(I + \sigma_q^2(diag\{R_{XX}\})^{-1})\right)^{1/N}}{\left(\det(I + \sigma_q^2(R_{XX})^{-1})\right)^{1/N}} \frac{\left[1 + \frac{1}{K}(1 - \frac{tr\{R_{X^qX^q}D^{q-1}R_{X^qX^q}D^{q-1}\}}{N^2}) - \frac{\sigma_q^2}{N} tr\{(diag R_{X^qX^q})^{-1}\}\right]}{\left[1 + \frac{N-1}{K}(\frac{1}{2} + \frac{1}{N}) - \frac{\sigma_q^2}{N} tr\{(\Lambda^q)^{-1}\}\right]} \end{aligned} \tag{81}$$

where $D^q = diag\{R_{XX}\}$ and $\Lambda^q = VR_{X^qX^q}V^T$.

5.3. LDU

In the causal case, and estimate $\hat{L}'$ of L is computed, and the r.v.s in the transform domain are not Gaussian :

$$y_{i,k} = x_{i,k} - \sum_{j=1}^{i-1} \hat{L'}_{ij} x_{i-j,k} - \sum_{j=1}^{i-1} \hat{L'}_{ij} q_{i-j,k}. \tag{82}$$

As in (23), the expected distortion for the LDU $E||\tilde{Y}||^2_{(\hat{L}',K,q)}$ is upper bounded by the rate-distortion function of a set of Gaussian r.v.s of same variances $\sigma^2_{y'_i}$, that is

$$E||\tilde{Y}||^2_{(\hat{L}',K,q)} < \sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2 \frac{(\hat{L}'\hat{R}_{X^qX^q}\hat{L}'^T)_{ii}}{(\Pi_{i=1}^N(\hat{L}'\hat{R}_{X^qX^q}\hat{L}'^T)_{ii})^{\frac{1}{N}}}]}\sigma^{2'}_{y_i}, \tag{83}$$

with $\sigma^{2'}_{y_i} = (\hat{L}'R_{X^qX^q}\hat{L}'^T - \sigma_q^2 I)_{ii}$. Thus, computing (62) when the transformation is based on K quantized vectors (for high K and under high resolution assumption) gives an upper bound, which is

$$E||\tilde{Y}||^{2^{low}}_{(\hat{L}',K,q)} = E\sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2 \frac{(\hat{L}'\hat{R}_{X^qX^q}\hat{L}'^T)_{ii}}{(\Pi_{i=1}^N(\hat{L}'\hat{R}_{XX}\hat{L}'^T)_{ii})^{\frac{1}{N}}}]}(\hat{L}'R_{X^qX^q}\hat{L}'^T - \sigma_q^2 I)_{ii}, \tag{84}$$

This bound (84) can be developped as

$$\begin{aligned}
&E||\tilde{Y}||^{2^{low}}_{(\hat{L}',K,q)} = E\sum_{i=1}^{N} c2^{-2[r+\frac{1}{2}log_2 \frac{(\hat{L}'R_{X^qX^q}\hat{L}'^T)_{ii}}{(\Pi_{i=1}^N(\hat{L}'R_{X^qX^q}\hat{L}'^T)_{ii})^{\frac{1}{N}}}]}\left(\prod_{i=1}^N(1+\frac{(\hat{L}'\Delta R\hat{L}'^T)_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}'^T)_{ii}})\right)^{\frac{1}{N}}\\
&\times\left(\frac{(\hat{L}'R_{X^qX^q}\hat{L}'^T)_{ii}}{(\hat{L}'R_{X^qX^q}\hat{L}'^T)_{ii}} - \frac{\sigma_q^2}{(\hat{L}'R_{X^qX^q}\hat{L}'^T)_{ii}}\right)\left(1+\frac{(\hat{L}'\Delta R\hat{L}'^T)_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}'^T)_{ii}}\right)\\
&=Ec2^{-2r}\left(\det\hat{L}'R_{X^qX^q}\hat{L}'^T\right)^{\frac{1}{N}}\left(\prod_{i=1}^N(1+\frac{(\hat{L}'\Delta R\hat{L}'^T)_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}'^T)_{ii}})\right)^{\frac{1}{N}}\left[\sum_{i=1}^N\left(1+\frac{(\hat{L}'\Delta R\hat{L}'^T)_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}'^T)_{ii}}\right)^{-1} - \sigma_q^2\sum_{i=1}^N((\hat{L}'R_{X^qX^q}\hat{L}'^T)_{ii})^{-1}\right].
\end{aligned} \tag{85}$$

Now, let $\hat{R}^q_{YY}$ be $\hat{L}'R_{X^qX^q}\hat{L}' = R^q_{YY} + \Delta R^q_{YY}$, where $R^q_{YY} = L'R_{X^qX^q}L'^T$, and let δ_{y^q} be the diagonal elements of ΔR^q_{YY}. Then, the first term of (85) may be written as

$$\begin{aligned}
&E||\tilde{Y}||^{2^{low}}_{(\hat{L}',K,q)} = E||\tilde{Y}||^2_{(L)}\left(1+\frac{1}{N}\sum_{i=1}^N\frac{\delta_{y^q}}{(R^q_{YY})_{ii}}\right)\left(\prod_{i=1}^N(1+\frac{(\hat{L}'\Delta R\hat{L}'^T)_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}'^T)_{ii}})\right)^{\frac{1}{N}}\sum_{i=1}^N\left(1+\frac{(\hat{L}'\Delta R\hat{L}'^T)_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}'^T)_{ii}}\right)^{-1}\\
&\approx E||\tilde{Y}||^2_{(L)}\left(1+\ E\frac{1}{N}\sum_{i=1}^N\frac{\delta_{y^q}}{(R^q_{YY})_{ii}} +\ E\ \frac{N-1}{2N^2}\sum_{i=1}^N(\frac{(\hat{L}'\Delta R\hat{L}'^T)_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}'^T)_{ii}})^2 -\ E\ \frac{1}{N^2}\sum_i\sum_{j>i}\frac{(\hat{L}'\Delta R\hat{L}'^T)_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}'^T)_{ii}}\frac{(\hat{L}'\Delta R\hat{L}'^T)_{jj}}{(\hat{L}'R_{X^qX^q}\hat{L}'^T)_{jj}}\right)
\end{aligned} \tag{86}$$

Using a similar analysis as in (46), one can show that

$$E\sum_{i=1}^N\frac{\delta_{y^q}}{(R^q_{YY})_{ii}} = \frac{1}{K}\sum_{i=1}^N\sum_{j\neq i}\frac{(R^q_{YY})_{jj}}{(R^q_{YY})_{jj}-(R^q_{YY})_{ii}} = \frac{1}{K}\frac{N(N-1)}{2}. \tag{87}$$

Thus, using the unimodularity of $\hat{L}'$, the first term may be approximated as

$$\begin{aligned}
&c2^{-2r}(\det R_{X^qX^q})^{1/N}\left[1+\frac{N(N-1)}{2K}+\frac{2N(N-1)}{2N^2K}\right]\\
&= c2^{-2r}(\det R_{XX})^{1/N}\left(\det(I+\sigma_q^2(R_{XX})^{-1})\right)^{1/N}\left[1+\frac{1}{K}\left(\frac{N-1}{2}+\frac{N-1}{K}\right)\right].
\end{aligned} \tag{88}$$

The expectation of the second term in (85) can be computed as in (46) also. By using the approximation

$$\frac{(\hat{L}'\Delta R\hat{L}'^T)_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}'^T)_{ii}} \approx \frac{(L'\Delta RL'^T)_{ii}}{(L'R_{X^qX^q}L'^T)_{ii}} \tag{89}$$

We can write $L'\Delta RL'^T = R^q_{YY} - \frac{1}{K}\sum_{i=1}^N Y_i^q Y_i^{qT}$: the random variable $(L'\Delta RL'^T)_{ii}$ corresponds now to the estimation error of R^q_{YY} obtained with a covariance matrix computed with K quantized vectors. Again, we make the approximation of

Gaussianity for y_i^q (this approximation is more justified than in the unitary case since now each y_i^q is the sum of one Gaussian r.v. and of i uniform r.v.s). Thus we assume $Evec(L'\Delta RL^{'T})(vec(L'\Delta RL^{'T}))^T \approx \frac{2R_{YY}^q \otimes R_{YY}^q}{K}$, whence

$$\sum_{i=1}^{N} \left(\frac{(L'\Delta RL^{'T})_{ii}}{(L'R_{XX}L^{'T})_{ii}} \right)^2 \approx \frac{2N}{K}, \tag{90}$$

and

$$\sum_i \sum_{j>i} \frac{(L'\Delta RL^{'T})_{ii}}{(L'R_{X^qX^q}L^{'T})_{ii}} \frac{(L'\Delta RL^{'T})_{jj}}{(L'R_{X^qX^q}L^{'T})_{jj}} = \sum_i \sum_{j>i} \frac{2}{K} \frac{(R_{YY}^q)_{ij}^2}{(R_{YY}^q)_{ii}(R_{YY}^q)_j} \approx 0. \tag{91}$$

The second term in (85) may approximated under the assumptions of high resolution and high K as

$$\begin{aligned}
&\sigma_q^2 E c^{-2r} (\det R_{X^qX^q})^{\frac{1}{N}} \\
&\underbrace{\left(1+\frac{1}{N}\sum_{i=1}^{N} \frac{\delta_{y^q}}{(R_{YY}^q)_{ii}}\right)\left(\prod_{i=1}^{N}(1+\frac{(\hat{L}'\Delta R\hat{L}^{'T})_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}^{'T})_{ii}}\right)^{\frac{1}{N}} \sum_{i=1}^{N}\frac{1}{(\hat{L}'R_{X^qX^q}\hat{L}')_{ii}}\left(1-\frac{(\hat{L}'\Delta R\hat{L}^{'T})_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}^{'T})_{ii}}+(\frac{(\hat{L}'\Delta R\hat{L}^{'T})_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}^{'T})_{ii}})^2\right)\left(1+\frac{(\hat{L}'\Delta R\hat{L}^{'T})_{ii}}{(\hat{L}R_{X^qX^q}\hat{L}^{'T})_{ii}}\right)^{-1}}_{\approx \sum_{i=1}^{N} \frac{1}{(L'R_{X^qX^q}L^{'T})_{ii}}} \\
&\approx \sigma_q^2 E c^{-2r} (\det R_{X^qX^q})^{\frac{1}{N}} tr\{(R_{YY}^q)^{-1}\} \\
&= \sigma_q^2 c 2^{-2r} (\det R_{XX})^{1/N} \left(\det(I+\sigma_q^2(R_{XX})^{-1})\right)^{1/N} tr\{(R'_{YY})^{-1}\}
\end{aligned} \tag{92}$$

Finally, using (83), the expected distortion for the LDU when the transformation is based on K quantized vectors is for high K and under high resolution assumption can be upper bounded as

$$E\|\tilde{Y}\|^2_{(\hat{L}',K,q)} < E\|\tilde{Y}\|^{2^{low}}_{(\hat{L}',K,q)} = E\|\tilde{Y}\|^2_{(L)} \left(\det(I+\sigma_q^2(R_{XX})^{-1})\right)^{1/N} \left[1+\frac{N-1}{K}\left[\frac{1}{2}+\frac{1}{N}\right] - \frac{\sigma_q^2}{N} tr\{(R'_{YY})^{-1}\}\right], \tag{93}$$

The corresponding expression for the coding gain in the causal case can then be lower bounded as

$$\begin{aligned}
G_{(\hat{L}',K,q)} &= \frac{E\|\tilde{X}\|^2_{(I,K,q)}}{E\|\tilde{Y}\|^2_{(\hat{L}',K,q)}} \\
&> \frac{E\|\tilde{X}\|^2_{(I,K,q)}}{E\|\tilde{Y}\|^{2^{low}}_{(\hat{L}',K,q)}} = G^0 \frac{\left(\det(I+\sigma_q^2(diag\{R_{XX}\})^{-1})\right)^{1/N}}{\left(\det(I+\sigma_q^2(R_{XX})^{-1})\right)^{1/N}} \frac{\left[1+\frac{1}{K}\left[1-\frac{tr\{R_{X^qX^q}D^{q-1}R_{X^qX^q}D^{q-1}\}}{N^2}\right]-\frac{\sigma_q^2}{N}tr\{(diagR_{X^qX^q})^{-1}\}\right]}{\left[1+\frac{N-1}{K}\left[\frac{1}{2}+\frac{1}{N}\right]-\frac{\sigma_q^2}{N}tr\{(R'_{YY})^{-1}\}\right]}.
\end{aligned} \tag{94}$$

where $D^q = diag\{R_{XX}\}$. As in Section 3, since the r.v. $\{y_i\}$ tends to be Gaussian very quickly as i grows, the bound in (94) is a fairly precise measure of the actual coding gain $G_{(\hat{L}',K,q)}$ for reasonably high N.
Indeed, it can be checked that the expression (94) and (81) tend to (20) and (28) respectively as $K \to \infty$, and both to (58) as $\sigma_q^2 \to 0$.

6. SIMULATIONS

For the simulations, we generated real Gaussian i.id. vectors with covariance matrix $R_{xx_j} = H_j R_{AR1} H_j^T$, $j = 1, 2$. R_{AR1} is the covariance matrix of a first order autoregressive process with normalized correlation coefficient ρ: R_{AR1} =Toeplitz ($[1\ \rho ... \rho^{N-1}]$), where Toeplitz denotes the Toeplitz matrix made with its first row $[1\ \rho ... \rho^{N-1}]$. H_j is a diagonal matrix whose i-th entry is $i^{1/3}$ for H_1 (increasing variances), and $(N-i+1)^{1/3}$ (decreasing variances). We assumed entropy constrained scalar quantizers Q_i with high resolution rate-distortion function $\sigma_{q_i}^2 = \frac{\pi e}{6} 2^{-2r_i} \sigma_{y_i}^2$. Thus, when computing the distortion, the Gaussianity of the transform signals was assumed in the causal case.
In Figure 3, the coding gain with quantization noise is plotted for KLT (upper curves) and LDU (lower curves), for signals of decreasing variances, and with $\rho = 0.9$, $N = 4$. The theoretical exact expressions are given by (20) and (28), and the

approximated expressions by (21) and (29). In Figure 4, the influence of the ordering of the variances is shown. The upper curve depicts the gain obtained with the causal approach by decorrelating the signals by decreasing order of variance (R_{xx_2}), and the lower curve by increasing order (R_{xx_1}).
The coding gains in presence of estimation noise are compared for LDU and KLT in Figure 5, for $N = 4$ and $\rho = 0.9$ (mean over 100 realizations).
The coding gains in presence of estimation noise and quantization are compared for KLT and LDU (signals of decreasing variances) in Figure 6, for $N = 8$, $\rho = 0.9$ and a rate of 3 bits per sample (mean over 100 realizations). The theoretical gains are given by (81) and (94). The observed behaviors of the transformation corresponds quite well to the theoretically predicted ones for $K \approx$ a few tens.

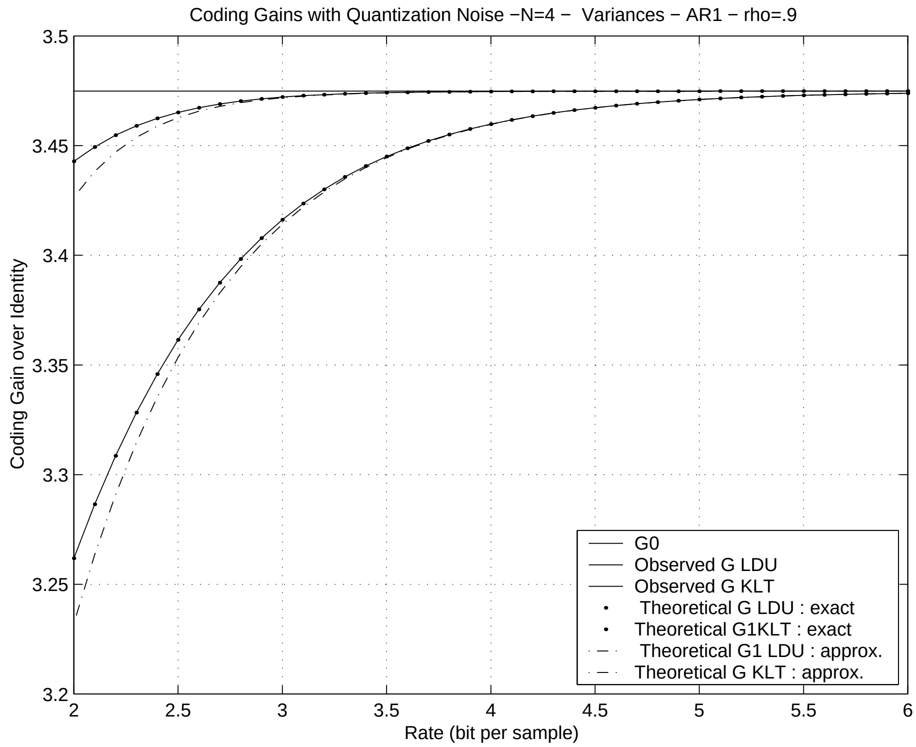

Fig. 3. Coding Gains vs rate in bit/sample.

7. REFERENCES

[1] B. Ottersten, M. Viberg, *Sensor Array Signal Processing*, Technical Report, Royal Institute of Tech., Chalmers Technical Inst., Sweden, January 1994.

[2] D. Mary, Dirk T.M. Slock, “Comparison between Unitary and Causal Approaches to Backward Adaptive Transform Coding of Vectorial Signals,” Submitted to *ICASSP 2002*, Orlando, USA, May 2002.

[3] D. Mary, Dirk T.M. Slock, “Backward Adaptive Transform Coding of Vectorial Signals : A Comparison between Unitary and Causal Approaches,” Submitted to *EUSIPCO-2002*, Toulouse, France, September 2002.

[4] V.K. Goyal, “ Transform Coding with Integer-to-Integer Transforms” *IEEE trans. on Inf. Theory*, vol. 46, no. 2, March 2000.

[5] D. Mary, Dirk T.M. Slock, “Causal Transform Coding, Generalized MIMO Prediction and Application to Vectorial DPCM Coding of Multichannel Audio,” *WASPAA01*, New York, USA, October 2001.

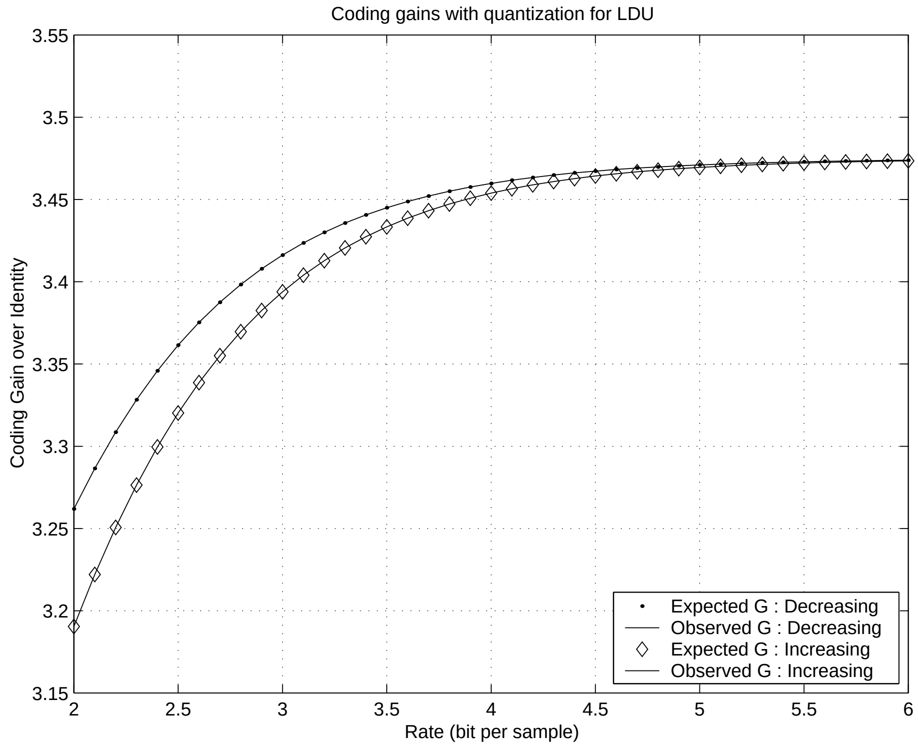

Fig. 4. Influence of the ordering of the signals $\{x_i\}$

[6] Dirk T.M. Slock, K. Maouche, D. Mary "Codage DPCM Vectoriel et Application au Codage de la Parole en Bande Élargie," *CORESA 2000*, Poitiers, France, October 2000.

[7] D. Mary, Dirk T.M. Slock, "Vectorial DPCM Coding and Application to Wideband Coding of Speech," *ICASSP 2001*, Salt Lake City, USA, May 2001.

[8] S.-M. Phoong, Y.-P. Lin, "Prediction-Based Lower Triangular Transform," *IEEE trans. on Sig. Proc.*, vol. 48, no. 7, July 2000.

[9] V. Goyal, J. Zhuang, M. Vetterli, " Transform Coding with Backward Adaptive Updates," *IEEE trans. on Inf. Theory*, vol. 46, no. 4, July 2000.

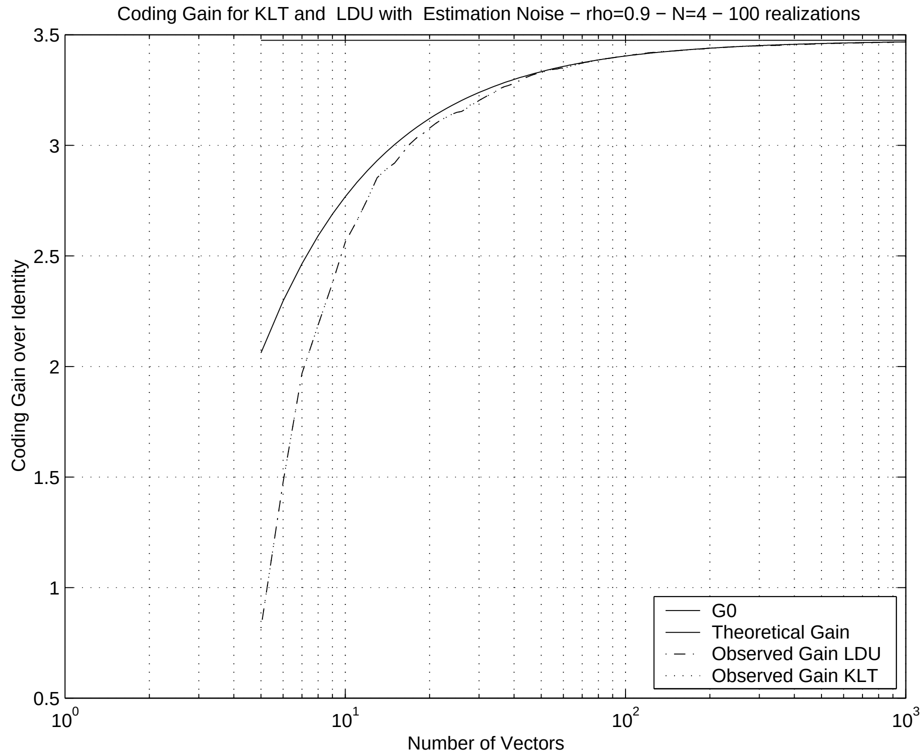

Fig. 5. Gains for KLT and LDU with estimation noise

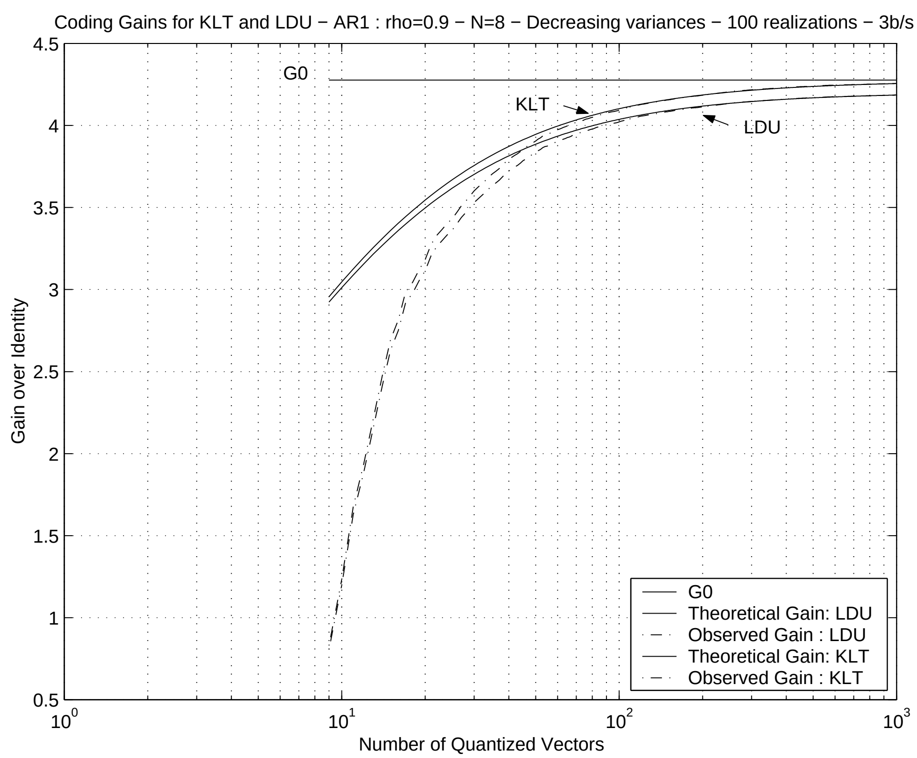

Fig. 6. Gains for KLT and LDU.