A Theoretical High Rate Analysis of Causal versus Unitary On-line Transform Coding †

David L. Mary and Dirk T.M. Slock *

Aryabhatta Research Institute of Observational Sciences
Manora Peak, Nainital-263 129
Uttaranchal, INDIA
Tel: 9105942-233727, extension: 228
Fax: 9105942-235136
email: dmary@upso.ernet.in

EURECOM Institute
Mobile Communications Department
2229 route des Crêtes, B.P. 193
06904 Sophia Antipolis Cdx, FRANCE
Tel: +33 4 9300 2606
Fax: +33 4 9300 2627
email: dirk.slock@eurecom.fr

Submitted to the IEEE Transactions on Signal Processing

EDICS: SP 3-CODC

Abstract

Backward adaptive or "on-line" transform coding (TC) of Gaussian sources is investigated. We compare in this context the Karhunen-Loève Transform (KLT, unitary approach) to the Causal Transform (CT, causal approach). When the covariance matrix $R_{\underline{x}}$ of the source is used in the TC scheme, KLT and CT present similar coding gains at high rates [1], [2], [3]. The aim of this study is to model analytically the behavior of these two coding structures when the ideal TC scheme gets perturbed, that is, when only a perturbed value $R_{\underline{x}} + \Delta R$ is known at the encoder. In the on-line TC schemes considered here, this estimate is used to compute both the transform and the bit assignment. ΔR is caused by two noise sources : estimation noise (finite set of available data at the encoder) and quantization noise (quantized data at the decoder). Furthermore, not only the transformation itself gets perturbed, but also the bit assignment. In this framework, theoretical expressions for the coding gains in both the unitary and the causal cases are derived under high rate assumption.

† Eurécom's research is partially supported by its industrial partners: Ascom, Swisscom, Thales Communications, ST Microelectronics, CEGETEL, France Télécom, Bouygues Télécom, Hitachi Europe Ltd., and Texas Instruments. This work was also supported in part by the french RNRT project COBASCA.

I. INTRODUCTION

A. Karhunen-Loève and Causal Transforms in classical TC

In the classical transform coding (TC) framework (high rate, optimal bit assignment [4], [5]) the Karhunen-Loève transform (KLT) has become a benchmark, since it has been proved optimal for Gaussian sources[1] [9], [10], [11], [12]. A transform is optimal in TC if the distortion between original and quantized data (usually the Mean Squared Error, MSE) is minimized for a given source and a given bitrate [4], [5].

Following initial work [13], others have demonstrated theoretically [13], [14], [1], [15], [16] and numerically [2] that the CT performs as well as the KLT for high bitrates (see Sec II for a quick overview of the CT). The causal transform being moreover less computationnally expensive than the KLT, this makes it very attractive for TC.

B. Backward Adaptive TC schemes

Classical analyses of TC (see *e.g.* [4], [5] and the references above) assume that the TC coding parameters (bit assignment and transform) are available at the encoder and decoder side. Equivalently, the covariance matrix of the data, from which the parameters can be computed, is assumed known at both sides of the coder. Most of the time however, TC schemes deal with non- or locally- stationary signals. In this case, sending the updates of the signal-dependent transformation and bit assignment as side information may cause a considerable overhead for the overall bitrate. Hence, one can seek to adapt these parameters on the basis of the data available at the decoder only. This backward adaptive framework may be related to the general problem of *universal lossy quantization*. Universality is meant here [2] as the ability of a system which has no *a priori* knowledge about the source, to achieve the same rate-distortion performance as a system designed with that knowledge. Very few works have investigated the feasibility of universal transform codes in the literature. Some techniques have been proposed [18], [19] which rely on so-called *two-stages* codes: the first stage codes the identity of the code that will be used to code the data; the second stage codes the data with the previously chosen code. Using one method [19] a pair (KLT; bit assignment) is chosen among a codebook of transformations

[1]For non Gaussian sources, different transforms may yield better compression results, see *e.g.* [6], [7], [8].

[2]Different kind of universality for lossy coding, or coding with a fidelity criterion, are defined in [17].

and bit assignment pairs; the index of the chosen pair is sent as side-information to the decoder. This type of technique is universal in the sense that it allows one to code with the best transform and bit assignment any source among a particular class. The methods investigated in the present work are different in the sense that they do not rely on "universal codebooks" of any kind. Instead of choosing among several precomputed transforms and bit assignments, we wish the encoder and the decoder to compute these parameters using previously decoded data only. This technique is computationally more expensive, but does not require any side-information. The approach of the proposed analysis is similar to that of [12], where backward adaptivity of the KLT is considered, using equal step size quantizers. It is proved in these works that such systems may produce the same coding performances than TC systems designed with the a priori knowledge of $R_{\underline{x}}$ when the number K of available quantized vectors becomes infinite. In the present works we propose to model and compare the coding performances as a functions of K for the KLT and the CT, when both the transform and the bit assignment are backward adaptive.

C. Formulation of the problem

Let us state more precisely the terms of the proposed evaluations. The backward adaptive systems considered here require that neither the transformation nor the parameters of the bit assignment be transmitted to the decoder. For the purpose of our analysis, we shall assume that the signal is a locally stationary Gaussian vectorial signals $\underline{x}$ with covariance matrix $R_{\underline{x}}$. Each source vector $\underline{x}_k$ [3] $= [x_{1,k} \;\; x_{2,k} \cdots x_{N,k}]^t$ may seen as the sample of a vector signal, whose components $x_{i,k}$ are the samples of N scalar signals $\{x_i\}$, $i = 1, \cdots, N$, taken at time k. The components of the corresponding transform vector $\underline{y}_k$ form a set of transform coefficients which are independently quantized using scalar quantizers.

In the classical TC framework, the KLT (denoted by V) or the CT (denoted by L) are computed so that $VR_{\underline{x}}V^t$, or $LR_{\underline{x}}L^t$ is diagonal. Let $R_{\underline{y}}$ denote the covariance matrix of the transformed signals. The variances $\sigma^2_{y_i}$ of the transform signals are $(R_{\underline{y}})_{ii}$, where $(.)_{ii}$ denotes the ith diagonal element of $(.)$. The number of bits b_i optimally allocated to each transform component is $b_i = b + \frac{1}{2}\log_2 \frac{\sigma^2_{y_i}}{(\prod_{i=1}^N \sigma^2_{y_i})^{\frac{1}{N}}}$. This

[3] Vectors will be denoted by underlined lowercase letters, and matrices by uppercase letters. The notation $L_{i,j}$ denotes the element on the ith row and jth column of L, and superscript ^{t} stands for transposition.

bit assignment algorithm is optimal in the sense that for a given set of $\{\sigma^2_{y_i}\}$ and a given per component bitrate b, the distortion is minimized [5]. This yields the same distortion $\mathbb{E}\,(y_i - y_i^q)^2 = \mathbb{E}\,\widetilde{y_i}^2$ on each component. The per component distortion may further may be expressed as $\mathbb{E}\,\widetilde{y}^2 = c\; 2^{-2b_i}\sigma^2_{y_i}$, where c is the quantizer performance factor w.r.t. the source [4]. When no transform is used (or equivalently, the Identity transform) the distortion becomes $\frac{1}{N}\,\mathbb{E}\,\|\underline{\widetilde{y}}\|^2_I = c2^{-2b}(\det \overline{\text{diag}}\,\{R_{\underline{x}}\})^{1/N}$, where $\overline{\text{diag}}\,\{\underline{a}\}$ represent the diagonal matrix with diagonal $\underline{a}$. For the KLT, the distortion becomes $\frac{1}{N}\,\mathbb{E}\,\|\underline{\widetilde{y}}\|^2_V = c2^{-2b}\det\{R_{\underline{x}}\}^{1/N}$. In the above distortions, the subscripts I and V refer to the transform. [4], [5]). The corresponding coding gain for KLT is then

$$G^0 = \frac{\mathbb{E}\,\|\underline{\widetilde{y}}\|^2_I}{\mathbb{E}\,\|\underline{\widetilde{y}}\|^2_V} = \left(\frac{\det\{\,\overline{\text{diag}}\,R_{\underline{x}}\}}{\det R_{\underline{x}}}\right)^{\frac{1}{N}}. \tag{1}$$

The backward adaptive TC systems considered here can only rely on the previously decoded data. These schemes are thus based on $\widehat{R}_{\underline{x}} = R_{\underline{x}} + \Delta R$ instead of $R_{\underline{x}}$, where $\widehat{R}_{\underline{x}}$ is an estimate of $R_{\underline{x}}$ available at both the encoder and the decoder. Hence the transformations ($\widehat{V}$ for the KLT and $\widehat{L}$ for the CT) will be such that $\widehat{V}\widehat{R}_{\underline{x}}\widehat{V}^t$ or $\widehat{L}\widehat{R}_{\underline{x}}\widehat{L}^t$ is diagonal. Let $\widehat{T}$ denote either $\widehat{V}$ or $\widehat{L}$. The per component distortion will be proportional to the variances of the signals transformed by means of $\widehat{T}$, say $\sigma'^2_{y_i}$. Regarding the bit assignment, the bits $\widehat{b}_i$ should be attributed on the basis of estimates of the variances available at both encoder and decoder also. With the notations above, these variance estimates are $(\widehat{T}\widehat{R}_{\underline{x}}\widehat{T})_{ii}$, which yields

$$\widehat{b}_i = b + \frac{1}{2}\log_2 \frac{(\widehat{T}\widehat{R}_{\underline{x}}\widehat{T}^t)_{ii}}{(\prod_{i=1}^N(\widehat{T}\widehat{R}_{\underline{x}}\widehat{T}^t)_{ii})^{\frac{1}{N}}}. \tag{2}$$

For most of transformations used in TC, the distortion in the transform domain $\mathbb{E}\,\|\underline{\tilde{y}}\|^2$ and in the signal domain $\mathbb{E}\,\|\underline{\tilde{x}}\|^2$ is the same. This property is sometimes referred to as “Unity Noise Gain Property” [1]. This is indeed true for orthogonal transforms (KLT, DCT, etc...) and for the causal transform [1], [2]. We obtain therefore the following measure of distortion for a system using a transformation $\widehat{T}$ based on $\widehat{R}_{\underline{x}}$:

$$\mathbb{E}\,\|\underline{\widetilde{y}}\|^2_{\widehat{T}} = \mathbb{E}\sum_{i=1}^N c2^{-2\widehat{b}_i}\sigma'^2_{y_i} = \mathbb{E}\sum_{i=1}^N c2^{-2[b+\frac{1}{2}\log_2\frac{(\widehat{T}\widehat{R}_{\underline{x}}\widehat{T}^t)_{ii}}{(\prod_{i=1}^N(\widehat{T}\widehat{R}_{\underline{x}}\widehat{T}^t)_{ii})^{\frac{1}{N}}}]}\sigma'^2_{y_i}. \tag{3}$$

where the expectation $\mathbb{E}$ is w.r.t. ΔR in case it is non-deterministic[4].

D. *High rate assumptions*

Several assumptions are implicitly or not made by the above description. Firstly, we assume a Gaussian source model. Secondly, the rate must be sufficiently high. The bit assignment mechanism (2) neglects the fact that $\widehat{b}_i$ can be non integer and negative. This would happen for low values of the average bitrate budget b, or even at higher values of b, for low values of some variances ${\sigma'}^2_{y_i}$. Thirdly, the expression (3) assumes that the quantizers' operational distortion-rate laws are of the form $c2^{-2b_i}\sigma^2_{y_i}$. This assumes, besides high rates (independence of c w.r.t. b_i) and significance of all the transform signals (they are assigned nonzero b_i), that these transform signals belong to the family of Gaussian probability density functions (p.d.f.s). For jointly Gaussian scalar sources x_i composing a vectorial source $\underline{x}$, this assumption is clearly true for the transform signals obtained by means of a KLT. In the case of a causal transform however, this is not rigorously true, because the prediction residuals $\{y_i\}, i = 2 \cdots N$, contain a quantized component through the closed loop prediction (see [2]). At high rates however, this perturbation is small and the shapes of the p.d.f. of the $\{y_i\}, i = 2 \cdots N$, are accurately approximated by Gaussian p.d.f. (see [20]). Additionally, we shall assume that the effects of quantization are to introduce on the data an uncorrelated white noise with variance $c2^{-2\widehat{b}_i}{\sigma'}^2_{y_i}$ which is a customary model in high rate TC, see *e.g.* [21], [1]. Finally, for estimation noise, the vectors to be coded will be assumed independent and identically distributed (i.i.d.). This may be the case if the sampling period of the scalar signals is high in comparison with their typical correlation time.

Hence, on the one hand, the proposed analysis (3) is indeed a modelization in the sense that high rate, Gaussian sources, etc ..., may not be verified in practice by any TC system. Also some practical TC algorithms may not provide the optimal, non integer bitrates bit assignment mentioned in (2) (e.g. greedy algorithm, etc...). On the other hand, these assumptions are quite customary in TC, and without these assumptions theoretical investigations of TC become very difficult.

[4]As in (3), the sign $=$ will be used along the derivations though this equality is correct only asymptotically (w.r.t. the rate); the sign $\approx$ will be used when the original expression (3) will be replaced by an approximation based on the dominant perturbation terms.

Paper Outline : The main characteristics of the CT are first outlined in Section II. The expressions of the distortion (3) and that of the corresponding coding gain are then compared for the KLT and the CT. This is done in three cases. In the first case (Section III), ΔR is caused by a quantization noise: the coding schemes are based on the statistics of the data corrupted by an additive white noise. In the second case, ΔR corresponds in Section IV to an estimation noise : the coding schemes are based the sample covariance matrix $\widehat{R}_{\underline{x}} = \frac{1}{K}\sum_{i=1}^{K} \underline{x}_i \underline{x}_i^t$. Finally, both influences of quantization and estimation noises are analyzed in Section V. Numerical simulations are presented in Section VI. The last Section summarizes the main results and draws some conclusions.

II. Classical Causal Transform Coding at a glance

The causal transform was first proposed in [13]. In the causal case, the transformed vector is obtained by subtracting the reference vector : $\underline{y}_k = \underline{x}_k - \overline{L}\underline{x}_k^q$, where $\overline{L}$ is a lower triangular matrix whose diagonal entries are zeros. The reference signal $\overline{L}\underline{x}_k^q$ is based on the past quantized samples [13]. The components $y_{i,k}$ appear as the prediction errors of $x_{i,k}$ with respect to the previous (whence the name of causality) quantized components, the $\{x_{1,k}^q \cdots x_{i-1,k}^q\}$. For optimal bit assignment, the optimal linear CT is unit diagonal and lower triangular. It may be written as $L = I - \overline{L}$ where I denotes the $(N \times N)$ Identity matrix. The non-zero coefficients $\{-L_{i,1} \cdots - L_{i,i-1}\}$ of $\overline{L}$ are the optimal linear prediction coefficients [22]. In other words, L is such that

$$LR_{\underline{x}}L^t = R_{\underline{y}} = \overline{\text{diag}}\,\{\sigma_{y_1}^2 \cdots \sigma_{y_N}^2\}. \tag{4}$$

It follows that $R_{\underline{x}} = L^{-1}R_{\underline{y}}L^{-t}$, which represents the LDU (Lower-Diagonal-Upper) factorization of $R_{\underline{xx}}$ [4]. Extensive details about the Causal Transform can be found in [13], [1], [2], [3]. If we neglect the fact that the prediction is based on quantized data, then $\mathbb{E}\,\|\widetilde{\underline{y}}\|_L^2 = \mathbb{E}\,\|\widetilde{\underline{y}}\|_V^2$: CT and KLT present the same coding gain G^0 of eq. (1). These distortions and G^0 shall be used as references in the sequel.

III. Quantization Effects on the Coding Gains

In this case, transformations and bit assignment are computed using quantized data . The statistics of the quantized data is assumed to be perfectly known in this section. In other words, we assume that an infinite number of quantized vectors $\underline{x}_i^q$ is available at the decoder, so that $R_{\underline{x}^q\underline{x}^q}$ is known.
Under the assumptions discussed in Sec. I-D, $\Delta R = E\tilde{\underline{x}}\tilde{\underline{x}}^t = \sigma_q^2 I$, where σ_q^2 denotes the variance of the quantization noise. Thus, the distortion (3) becomes

$$\mathbb{E}\,\|\tilde{\underline{y}}\|^2_{\widehat{T},q} = \sum_{i=1}^{N} c2^{-2[b+\frac{1}{2}\log_2 \frac{(\widehat{T}R_{\underline{x}^q}\widehat{T}^t)_{ii}}{(\prod_{i=1}^N (\widehat{T}R_{\underline{x}^q}\widehat{T}^t)_{ii})^{\frac{1}{N}}}]} {\sigma'}^2_{y_i}, \tag{5}$$

where $\widehat{T}$ refers to the transformation, and q refers to quantization. Expression (5) may now be evaluated for $\widehat{T} = I$, $\widehat{T} = \widehat{V}$ and $\widehat{T} = \widehat{L}$.

A. Identity Transformation

In this case, the number of bits attributed to the quantizer Q_i is

$$\widehat{b}_i = b + \frac{1}{2}\log_2 \frac{(R_{\underline{x}^q})_{ii}}{(\prod_{i=1}^N (R_{\underline{x}^q})_{ii})^{\frac{1}{N}}}, \tag{6}$$

and the variance ${\sigma'}^2_{y_i}$ are indeed $(R_{\underline{x}})_{ii}$. The distortion (5), where $\widehat{T}$ is replaced by I and ${\sigma'}^2_{y_i}$ by $(R_{\underline{x}})_{ii}$, becomes

$$\mathbb{E}\,\|\tilde{\underline{y}}\|^2_{I,q} \;=\; \sum_{i=1}^{N} c2^{-2r}\left(\det\ \mathrm{diag}\, R_{\underline{x}^q}\right)^{\frac{1}{N}} \frac{(R_{\underline{x}})_{ii}}{(R_{\underline{x}^q})_{ii}}, \tag{7}$$

where $\mathrm{diag}\, A$ denotes the diagonal matrix with same diagonal as A. This leads to[5]

$$\mathbb{E}\,\|\tilde{\underline{y}}\|^2_{I,q} \;=\; \mathbb{E}\,\|\tilde{\underline{y}}\|^2_I \tfrac{1}{N}(\det(I+\sigma_q^2(\mathrm{diag}\, R_{\underline{x}})^{-1}))^{\frac{1}{N}}\ \mathrm{tr}\left\{\left(I+\sigma_q^2(\mathrm{diag}\, R_{\underline{x}})^{-1}\right)^{-1}\right\}. \tag{8}$$

The distortion is increased (w.r.t. a scheme based on $R_{\underline{x}}$) because the bits allocated on the basis of the variances of the quantized signals are not the optimal ones. An approximation of (8) up to the second order of the perturbations gives

$$\begin{aligned} \mathbb{E}\,\|\tilde{\underline{y}}\|^2_{I,q} &= c2^{-2r}(\det\ \mathrm{diag}\,\{R_{\underline{x}}\})^{1/N}\left(\Pi_{i=1}^N(\tfrac{\sigma_q^2}{(R_{\underline{x}})_{ii}}))^{1/N}\sum_{i=1}^N(1+\frac{1}{(R_{\underline{x}})_{ii}}\right)^{-1} \\ &\approx \mathbb{E}\,\|\tilde{\underline{y}}\|^2_I\left[1+\tfrac{\sigma_q^4}{N^2}(\tfrac{N-1}{2}\sum_{i=1}^N\frac{1}{(R_{\underline{x}})_{ii}^2}-\sum_{i=1}^N\sum_{j>i}\frac{1}{(R_{\underline{x}})_{ii}(R_{\underline{x}})_{jj}})\right]. \end{aligned} \tag{9}$$

[5]The calculations for the present and the following subsections are omitted for lack of space but can be found in [3].

The perturbation effect w.r.t. the ideal case is only caused by the perturbation upon the bit assignment. These perturbation terms are of the form $(\frac{\sigma_q^2}{(R_{\underline{x}})_{ii}})^2$. High rate means that the quantization noise variance is small in comparison with that of the signal components. Hence we see from eq. (9) that this perturbation is a second order term.

B. KLT

As observed in [12] also, if V denotes a KLT of $R_{\underline{x}}$, then $V(R_{\underline{x}} + \sigma_q^2 I)V^t = \Lambda + \sigma_q^2 I = \Lambda^q$, and V is also a KLT of $R_{\underline{x}} + \sigma_q^2 I$. Thus, the perturbation term $\sigma_q^2 I$ on $R_{\underline{x}}$ does not change the backward adapted transformation: $\widehat{V} = V$. The variances of the transformed signals remain unchanged: ${\sigma'}_{y_i}^2 = (VR_{\underline{x}}V^t)_{ii} = \lambda_i$. However, the variance estimates at the decoder are $(VR_{\underline{x}^q}V^t)_{ii} = \lambda_i + \sigma_q^2$. These variances are used to assign the bits $\widehat{b}_i$. These are computed as in eq. (2), where $\widehat{V}$ replaces $\widehat{T}$ and $R_{\underline{x}^q}$ replaces $\widehat{R}_{\underline{x}}$. The actual distortion becomes

$$\begin{aligned} \mathbb{E}\,\|\underline{\tilde{y}}\|^2_{V,q} &= \sum_{i=1}^{N} c 2^{-2[b + \frac{1}{2}\log_2 \frac{(VR_{\underline{x}^q}V^t)_{ii}}{(\prod_{i=1}^{N}(VR_{\underline{x}^q}V^t)_{ii})^{\frac{1}{N}}}]} (VR_{\underline{x}}V^t)_{ii} \\ &= \sum_{i=1}^{N} c 2^{-2r} \left(\det \text{ diag}\,\{VR_{\underline{x}^q}V^t\}\right)^{\frac{1}{N}} \frac{(VR_{\underline{x}}V^t)_{ii}}{(VR_{\underline{x}^q}V^t)_{ii}}. \end{aligned} \tag{10}$$

Since $VR_{\underline{x}}V^t$ and $VR_{\underline{x}^q}V^t$ are diagonal, one can show that

$$\sum_{i=1}^{N} \frac{(VR_{\underline{x}}V^t)_{ii}}{(VR_{\underline{x}^q}V^t)_{ii}} = \text{tr}\,\{\left(I + \sigma_q^2(R_{\underline{x}}^{-1})\right)^{-1}\} = \text{tr}\,\{\left(I + \sigma_q^2(\Lambda)^{-1}\right)^{-1}\}. \tag{11}$$

Also,

$$\det\left(R_{\underline{x}^q}\right) = \det\left(R_{\underline{x}}\right)\det(I + \sigma_q^2(R_{\underline{x}}^{-1})). \tag{12}$$

Finally, the distortion for the KLT with quantization noise is

$$\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{V,q} = \mathbb{E}\,\|\underline{\tilde{y}}\|^2_{V} \frac{1}{N}(\det(I + \sigma_q^2(\Lambda^{-1})))^{\frac{1}{N}} \text{tr}\,\{\left(I + \sigma_q^2(\Lambda^{-1})\right)^{-1}\}. \tag{13}$$

Again, the increase in distortion comes from the perturbation occurring upon the bit assignment mechanism. An expression approximating this distortion may be obtained by

$$\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{V,q} = c2^{-2r}(\det \text{ diag}\,\{R_{\underline{x}}\})^{\frac{1}{N}} \frac{1}{N} \left(\prod_{i=1}^{N}(1 + \frac{\sigma_q^2}{\lambda_i})\right)^{\frac{1}{N}} \sum_{i=1}^{N} \left(1 + \frac{\sigma_q^2}{\lambda_i}\right)^{-1}. \tag{14}$$

By developing the product and the sum in (14), it can be checked that the terms proportional to σ_q^2 vanish, so that

$$\left(\prod_{i=1}^{N}(1+\frac{\sigma_q^2}{\lambda_i})\right)^{\frac{1}{N}} \sum_{i=1}^{N}\left(1+\frac{\sigma_q^2}{\lambda_i}\right)^{-1} \approx N + \frac{N-1}{2N}\sum_{i}\frac{\sigma_q^4}{\lambda_i} - \frac{1}{N}\sum_{i=1}^{N}\sum_{j>i}\frac{\sigma_q^4}{\lambda_i\lambda_j}. \quad (15)$$

This leads to the following approximated distortion

$$\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{V,q} \;\approx\; \mathbb{E}\,\|\underline{\tilde{y}}\|^2_{V}\left[1+\tfrac{\sigma_q^4}{N^2}\left(\tfrac{N-1}{2}\sum_{i=1}^{N}\frac{1}{\lambda_i^2} - \sum_{i=1}^{N}\sum_{j>i}\frac{1}{\lambda_i\lambda_j}\right)\right] \quad (16)$$

Using (8) and (13), the corresponding expression for the coding gain in the unitary case with quantization noise is

$$G_{V,q} = G^0\,\frac{(\det(I+\sigma_q^2(\operatorname{diag} R_{\underline{x}})^{-1}))^{\frac{1}{N}}\operatorname{tr}\{\left(I+\sigma_q^2(\operatorname{diag} R_{\underline{x}})^{-1}\right)^{-1}\}}{(\det(I+\sigma_q^2(\Lambda^{-1})))^{\frac{1}{N}}\operatorname{tr}\{\left(I+\sigma_q^2(\Lambda^{-1})\right)^{-1}\}}. \quad (17)$$

With (9) and (16), $G_{V,q}$ can be approximated as

$$G_{V,q} \approx G^0\left[1+\frac{\sigma_q^4}{N^2}\left(\frac{N-1}{2}\sum_{i=1}^{N}(\frac{1}{(R_{\underline{x}})_{ii}^2}-\frac{1}{(\lambda_i)^2}) - \sum_{i=1}^{N}\sum_{j>i}(\frac{1}{(R_{\underline{x}})_{ii}(R_{\underline{x}})_{jj}}-\frac{1}{\lambda_i\lambda_j})\right)\right]. \quad (18)$$

The perturbation effect w.r.t. the ideal case is only caused by the perturbation upon the bit assignment. As in the case of Identity transformation, the perturbation terms in eq. (18) are second order terms of the form $(\frac{\sigma_q^2}{(R_{\underline{x}})_{ii}})^2$ or $(\frac{\sigma_q^2}{\lambda_i})^2$.

C. Causal Transform (CT)

In the causal case, the encoder computes a transformation $\widehat{L} = L'$ such that $L'R_{\underline{x}^q}L'^T = R'_{\underline{y}}$. The causal transform corresponds to a LDU factorization of $R_{\underline{x}^q}$. $R'_{\underline{y}}$ is the diagonal matrix of the variances used for the bit assignment (L' and $R'_{\underline{y}}$ are both available to the decoder). In this case, the difference vector $\underline{y}$ is $\underline{x} - \overline{L'}\underline{x}^q$. By the analysis of [2], the quantization noise is filtered by the rows of $\overline{L'}$ (see Figure 1). Note that in this case $\mathbb{E}\,\|\underline{\tilde{x}}\|^2_{L',q}$ still equals $\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{L',q}$, since $\underline{\tilde{x}} = \underline{x}^q - \underline{x} = \underline{y}^q + \overline{L'}\underline{x}^q - \underline{x} = \underline{y}^q - \underline{y} = \underline{\tilde{y}}$.

Regarding the estimates of the rates, they are computed by eq. (2), where $\widehat{T}$ is replaced by $\widehat{L}'$, and $\widehat{R}_{\underline{x}}$ by $R_{\underline{x}^q}$. At high rates, it is shown in [2] that the actual variances of the signals y_i obtained by means of

L' may be approximated as $(L'R_{\underline{x}^q}L'^T - \sigma_q^2 I)_{ii}$. Using (5), the distortion $\mathbb{E}\,\|\tilde{\underline{y}}\|^2_{L',q}$ is then given by

$$\begin{aligned}\mathbb{E}\,\|\tilde{\underline{y}}\|^2_{L',q} &= c2^{-2[b+\frac{1}{2}\log_2\frac{(L'R_{\underline{x}^q}L'^T)_{ii}}{(\prod_{i=1}^N (L'R_{\underline{x}^q}L'^T)_{ii})^{\frac{1}{N}}}]}\sum_{i=1}^N (L'R_{\underline{x}^q}L'^T - \sigma_q^2 I)_{ii} \\ &\approx \sum_{i=1}^N c2^{-2r}\left(\det\ \mathrm{diag}\,\{L'R_{\underline{x}^q}L'^T\}\right)^{\frac{1}{N}}\left(1-\frac{(\sigma_q^2 I)_{ii}}{(L'R_{\underline{x}^q}L'^T)_{ii}}\right).\end{aligned} \tag{19}$$

Since the transformation L' is unimodular[6], the determinant in the previous expression equals the determinant in (12). The sum in (19) may be written as $\mathrm{tr}\,\{(I-\sigma_q^2(L'R_{\underline{x}^q}L'^T)^{-1})\} = \mathrm{tr}\,\{(I-\sigma_q^2 R'^{-1}_{\underline{y}})\}$. Thus (19) becomes

$$\mathbb{E}\,\|\tilde{\underline{y}}\|^2_{L',q} = \mathbb{E}\,\|\tilde{\underline{y}}\|^2_L \frac{1}{N}(\det(I+\sigma_q^2(\Lambda^{-1})))^{\frac{1}{N}}\,\mathrm{tr}\,\{\left(I-\sigma_q^2(R'^{-1}_{\underline{y}})\right)\}. \tag{20}$$

The excess in distortion comes not only from the perturbation occurring on the bit assignment mechanism but also from the filtering of the quantization noise. Up to the first order of perturbations, we obtain

$$\begin{aligned}\mathbb{E}\,\|\tilde{\underline{y}}\|^2_{(L',q} &= c2^{-2r}(\det\ \mathrm{diag}\,\{R_{\underline{x}}\})^{\frac{1}{N}}\left(\prod_{i=1}^N(1+\frac{\sigma_q^2}{\lambda_i})\right)^{\frac{1}{N}}\sum_{i=1}^N\left(1-\sigma_q^2\frac{1}{(R'_{\underline{yy}})_{ii}}\right) \\ &\approx \mathbb{E}\,\|\tilde{\underline{y}}\|^2_V\left[1+\frac{\sigma_q^2}{N}\sum_{i=1}^N\left(\frac{1}{\lambda_i}-\frac{1}{\sigma^2_{y_i}}\right)\right],\end{aligned} \tag{21}$$

where the $\sigma^2_{y_i}$ correspond to optimal prediction error variances in absence of quantization noise.

The corresponding exact expression for the coding gain is

$$G_{L',q} = G^0\frac{(\det(I+\sigma_q^2(\,\mathrm{diag}\,\{R_{\underline{x}}\})^{-1}))^{\frac{1}{N}}\,\mathrm{tr}\,\{\left(I+\sigma_q^2(\,\mathrm{diag}\,\{R_{\underline{x}}\})^{-1}\right)^{-1}\}}{(\det(I+\sigma_q^2(\Lambda^{-1})))^{\frac{1}{N}}\,\mathrm{tr}\,\{\left(I-\sigma_q^2(R'^{-1}_{\underline{y}})\right)\}}. \tag{22}$$

Up to the first order of perturbation we get,

$$G_{L',q} \approx G^0\left[1-\frac{\sigma_q^2}{N}\sum_{i=1}^N\left(\frac{1}{\lambda_i}-\frac{1}{\sigma^2_{y_i}}\right)\right]. \tag{23}$$

The approximated expression (23) shows that the perturbation effects of the bit assignment mechanism (2nd order terms) are in the causal case negligible in comparison with those of the noise feedback (1st order terms). This coding gain is similar to that obtained in [2], where only the noise feedback was accounted for (no perturbation on the bit assignment).

An interesting consequence of (23) is that the performance of the causal TC scheme depend on the

[6] L being unit diagonal and lower triangular, its determinant equals the product of its diagonal elements, which is one.

order in which the signals $\{x_i\}$ get decorrelated. As shown in [2], the signals x_i should be decorrelated by order of decreasing variance if we want $G_{L',q}$ to be maximized (see also Fig. 3 and 6 in Section VI). In other words, in the vector $\underline{x}_k = [x_{1,k}\, x_{2,k} \ldots x_{N,k}]^t$, the component x_1 should be that of largest variance, x_2 the component with second largest variance, etc..., if we want the noise feedback to be minimized.

IV. Estimation Noise

We analyze in this section the coding gains of a backward adaptive scheme based on an estimate of the covariance matrix $\widehat{R}_{\underline{x}} = \frac{1}{K}\sum_{i=1}^{K} \underline{x}_i \underline{x}_i^t = R + \Delta R$, where ΔR corresponds to the estimation noise. In the following, the subscript K refers to the estimation noise corresponding to K vectors. In this case, one can show that ΔR is a zero mean Gaussian random variable, with

$$\mathbb{E}\,\mathrm{vec}(\Delta R)\,(\mathrm{vec}(\Delta R))^t \approx \frac{2}{K} R_{\underline{x}} \otimes R_{\underline{x}}, \tag{24}$$

where $\otimes$ denotes the Kronecker product.

Using K data vectors, encoder and decoder compute a transformation $\widehat{T}$ which diagonalizes $\widehat{R}_{\underline{x}}$: $\widehat{T}\widehat{R}_{\underline{x}}\widehat{T} = \widehat{R}_{\underline{y}}$. The number of bits assigned to each component is as in eq. (2), with the definition of $\widehat{T}$ and $\widehat{R}_{\underline{x}}$ above.

Now, the actual variances of the signals obtained by applying $\widehat{T}$ to $\underline{x}$ are $(\widehat{T}R_{\underline{x}}\widehat{T}^t)_{ii}$. Note that in the causal case, $\underline{y} = I - \widehat{\overline{L}}\underline{x} = \widehat{L}\underline{x}$, so that $R'_{\underline{y}} = \widehat{L}R_{\underline{x}}\widehat{L}^t$. 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 resulting distortion for a sample covariance matrix based on K vectors is as in eq. (3), with $\sigma^{2\prime}_{y_i} = (\widehat{T}R_{\underline{x}}\widehat{T}^t)_{ii}$.

A. Identity Transformation

With $\widehat{T} = I$, and using a similar analysis as in the previous section, we obtain for the distortion

$$\begin{aligned}
\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{I,K} &= \mathbb{E}\, c2^{-2r}\left(\det\,\mathrm{diag}\,\{R_{\underline{x}}\}\right)^{\frac{1}{N}} \left(\prod_{i=1}^{N}(1+\frac{(\Delta R)_{ii}}{(R_{\underline{x}})_{ii}})\right)^{\frac{1}{N}} \sum_{i=1}^{N}\left(1+\frac{(\Delta R)_{ii}}{(R_{\underline{x}})_{ii}}\right)^{-1} \\
&\approx \mathbb{E}\,\|\underline{\tilde{y}}\|^2_{I}\left(1+\ \mathbb{E}\ \tfrac{N-1}{2N^2}\sum_{i=1}^{N}(\frac{(\Delta R)_{ii}}{(R_{\underline{x}})_{ii}})^2 -\ \mathbb{E}\ \frac{1}{N^2}\sum_{i}\sum_{j>i}\frac{(\Delta R)_{ii}}{(R_{\underline{x}})_{ii}}\frac{(\Delta R)_{jj}}{(R_{\underline{x}})_{jj}}\right)
\end{aligned} \tag{25}$$

With (24), the second expectation in (25) may be written as

$$\mathbb{E}\ \frac{N-1}{2N^2}\sum_{i=1}^{N}\left(\frac{(\Delta R)_{ii}}{(R_{\underline{x}})_{ii}}\right)^2 \approx \frac{N-1}{2N^2}\sum_{i=1}^{N}\frac{2(R_{\underline{x}})_{ii}^2}{K(R_{\underline{x}})_{ii}^2} = \frac{N-1}{2N^2}\frac{2N}{K} = \frac{N-1}{NK}, \tag{26}$$

and the third expectation leads to

$$\begin{aligned}\mathbb{E}\ \tfrac{1}{N^2}\sum_i\sum_{j>i}\frac{(\Delta R)_{ii}}{(R_{\underline{x}})_{ii}}\frac{(\Delta R)_{jj}}{(R_{\underline{x}})_{jj}} &\approx \tfrac{2}{KN^2}\sum_i\sum_{j>i}\frac{(R_{\underline{x}})_{ij}^2}{(R_{\underline{x}})_{ii}(R_{\underline{x}})_{jj}} \\ &\approx \tfrac{2}{KN^2}\|\triangleright\left((\text{ diag}\{R_{\underline{x}}\})^{1/2}R_{\underline{x}}(\text{ diag}\{R_{\underline{x}}\})^{1/2}\right)\|^2\end{aligned} \tag{27}$$

where $\triangleright(A)$ denotes the strictly lower triangular matrix made with the strictly lower triangular part of A, and $\|.\|^2$ denotes the Frobenius norm. If D denotes $\text{diag}\{R_{\underline{x}}\}$, we obtain

$$\begin{aligned}\mathbb{E}\ \tfrac{1}{N^2}\sum_i\sum_{j>i}\frac{(\Delta R)_{ii}}{(R_{\underline{x}})_{ii}}\frac{(\Delta R)_{jj}}{(R_{\underline{x}})_{jj}} &\approx \tfrac{1}{KN^2}\left(\|D^{-\frac{1}{2}}R_{\underline{x}}D^{-\frac{1}{2}}\|^2 - \underbrace{\|\text{ diag}\{D^{-\frac{1}{2}}R_{\underline{x}}D^{-\frac{1}{2}}\|^2}_{N}\right) \\ &\approx \tfrac{1}{KN^2}(\text{ tr}\{R_{\underline{x}}D^{-1}R_{\underline{x}}D^{-1}\} - N).\end{aligned} \tag{28}$$

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

$$\mathbb{E}\,\|\underline{\tilde{y}}\|_{I,K}^2 \approx \mathbb{E}\,\|\underline{\tilde{y}}\|_I^2\left(1+\frac{1}{K}[1-\frac{1}{N^2}\text{ tr}\{R_{\underline{x}}(\text{ diag}\{R_{\underline{x}}\})^{-1}R_{\underline{x}}(\text{ diag}\{R_{\underline{x}}\})^{-1}\}]\right). \tag{29}$$

B. KLT

In the unitary case, the expected distortion $\mathbb{E}\,\|\underline{\tilde{y}}\|_{\widehat{V},K}^2$ is as in eq. (3), with $\widehat{T}$ replaced by $\widehat{V}$, and ${\sigma^2}'_{y_i}$ by $(\widehat{V}R_{\underline{x}}\widehat{V}^t)_{ii}$. Using an analysis similar to the previous subsection, the expected distortion for the KLT when the transformation is based on K vectors becomes, for sufficiently large K

$$\begin{aligned}\mathbb{E}\,\|\underline{\tilde{y}}\|_{\widehat{V},K}^2 &= \mathbb{E}\,\|\underline{\tilde{y}}\|_V^2\left(\tfrac{1}{N}\ \mathbb{E}\ (\det(I+R_{\underline{x}}^{-1}\Delta R))^{\frac{1}{N}}\text{ tr}\{\left(I+R_{\underline{x}}^{-1}\Delta R\right)^{-1}\}\right). \\ &\approx \mathbb{E}\,\|\underline{\tilde{y}}\|_V^2\left(1+\tfrac{N-1}{K}\left[\tfrac{1}{2}+\tfrac{1}{N}\right]\right)\end{aligned} \tag{30}$$

The corresponding coding gain is

$$G_{\widehat{V},K} = \frac{\mathbb{E}\,\|\underline{\tilde{y}}\|_{I,K}^2}{\mathbb{E}\,\|\underline{\tilde{y}}\|_{\widehat{V},K}^2} \approx G^0\left(1-\frac{1}{K}\left[\frac{\text{tr}\{R(\text{ diag}\{R_{\underline{x}}\})^{-1}R(\text{ diag}\{R_{\underline{x}}\})^{-1}\}}{N^2}+\frac{N-1}{2}-\frac{1}{N}\right]\right). \tag{31}$$

C. Causal Transform (CT)

As commented in the introduction of this section, the expected distortion with $\widehat{L}$ computed with $\widehat{R}_{\underline{x}}$ is

$$\begin{aligned}\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{\widehat{L},K} &= \mathbb{E}\sum_{i=1}^{N} c2^{-2[b+\frac{1}{2}\log_2\frac{(\widehat{L}\widehat{R}_{\underline{x}}\widehat{L}^t)_{ii}}{(\prod_{i=1}^{N}(\widehat{L}\widehat{R}_{\underline{x}}\widehat{L}T)_{ii})^{\frac{1}{N}}}]}(\widehat{L}R_{\underline{x}}\widehat{L}^t)_{ii} \\ &= \mathbb{E}\,c2^{-2r}\left(\det\widehat{L}\widehat{R}_{\underline{x}}\widehat{L}^t\right)^{\frac{1}{N}}\sum_{i=1}^{N}\frac{(\widehat{L}R_{\underline{x}}\widehat{L}^t)_{ii}}{(\widehat{L}\widehat{R}_{\underline{x}}\widehat{L}^t)_{ii}},\end{aligned} \tag{32}$$

where we used a factorization similar to that used in (7). Now by the unimodularity property of $\widehat{L}$, we can write the determinant in (32) as

$$\left(\det\widehat{L}\widehat{R}_{\underline{x}}\widehat{L}^t\right)^{\frac{1}{N}} = \det\widehat{R}_{\underline{x}} = \det(R_{\underline{x}})\det(I+R_{\underline{x}}^{-1}\Delta R), \tag{33}$$

and since $\widehat{L}$ diagonalizes $\widehat{R}_{\underline{x}}$, we can write the sum in (32) as

$$\sum_{i=1}^{N}\frac{(\widehat{L}R_{\underline{x}}\widehat{L}^t)_{ii}}{(\widehat{L}\widehat{R}_{\underline{x}}\widehat{L}^t)_{ii}} = \text{tr}\,\{(I+R_{\underline{x}}^{-1}\Delta R)^{-1}\}. \tag{34}$$

The perturbation terms in eq. (33) and (34) are the same in the causal and the unitary case : the equality of the determinants in eq. (33) comes from the unimodularity of the transformations $\widehat{L}$ and $\widehat{V}$, and the equality of the traces in (34) comes from their decorrelating property. Hence, because both CT and unitary KLT are decorrelating and unimodular transforms, they yield the same distortion $\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{\widehat{L},K} = \mathbb{E}\,\|\underline{\tilde{y}}\|^2_{\widehat{V},K}$, as given by eq. (30). The coding gains with estimation noise are thus equal for KLT and CT and may be approximated by eq. (31).

V. Quantization and Estimation Noise

This Section deals with the most general case of this study. In presence of quantization and estimation noises, transforms and bit assignment should be computed using a number K of decoded vectors, or equivalently using $\widehat{R}_{\underline{x}^q} = \frac{1}{K}\sum_{i=1}^{K}\underline{x}_i^q\underline{x}_i^{q^t}$. The estimated transform $\widehat{T}$ is such that $\widehat{T}\widehat{R}_{\underline{x}^q}\widehat{T}^t$ is a diagonal matrix, which corresponds to the estimated variances of the transformed signals. We shall continue denoting by $\sigma'^2_{y_i}$ the actual variances of the transformed signals (obtained by applying $\widehat{T}$ to $\underline{x}_k$). The expected distortion $\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{\widehat{T},K,q}$ can be computed as in eq. (3), with $\widehat{R}_{\underline{x}}$ replaced by $\widehat{R}_{\underline{x}^q}$ (the subscripts $_q$ and $_K$ refer to the presence of quantization and estimation noise). This distortion must now be evaluated for Identity, KL and causal transforms.

A. Identity Transformation

With $\widehat{T} = I$, and by writing $R_{\underline{x}} = R_{\underline{x}^q} - \sigma_q^2 I$, we obtain

$$\begin{aligned}\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{I,K,q} = {} & \mathbb{E}\sum_{i=1}^{N} c2^{-2[b+\frac{1}{2}\log_2 \frac{(\widehat{R}_{\underline{x}^q})_{ii}}{(\prod_{i=1}^{N}(\widehat{R}_{\underline{x}^q})_{ii})^{\frac{1}{N}}}]}(R_{\underline{x}^q})_{ii} \\ & -\sigma_q^2\,\mathbb{E}\sum_{i=1}^{N} c2^{-2[b+\frac{1}{2}\log_2 \frac{(\widehat{R}_{\underline{x}^q})_{ii}}{(\prod_{i=1}^{N}(\widehat{R}_{\underline{x}^q})_{ii})^{\frac{1}{N}}}]}.\end{aligned} \tag{35}$$

For sufficiently high resolution and large K, the expected distortion for Identity transform with quantization and estimation noise leads to

$$\begin{aligned}\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{I,K,q} \approx {} & \mathbb{E}\,\|\underline{\tilde{y}}\|^2_{I}\left(\det(I+\sigma_q^2(\text{ diag }\{R_{\underline{x}}\})^{-1})\right)^{1/N} \\ & \times\left[1+\tfrac{1}{K}\left[1-\tfrac{1}{N^2}\text{ tr }\{R_{\underline{x}^q}(\text{ diag }R_{\underline{x}^q})^{-1}R_{\underline{x}^q}(\text{ diag }R_{\underline{x}^q})^{-1}\}\right]-\tfrac{\sigma_q^2}{N}\text{ tr }\{(\text{ diag }R_{\underline{x}^q})^{-1}\}\right].\end{aligned} \tag{36}$$

B. KLT

In the unitary case, $\sigma^{2\prime}_{y_i} = (\widehat{V}R_{\underline{x}}\widehat{V}^t)_{ii}$. After some computation we find for the expected distortion in the unitary case, when the transformation is based on K quantized vectors,

$$\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{\widehat{V},K,q} \approx \mathbb{E}\,\|\underline{\tilde{y}}\|^2_{V}\left(\det(I+\sigma_q^2(R_{\underline{x}})^{-1})\right)^{\frac{1}{N}}\left[1+\frac{N-1}{K}\left[\frac{1}{2}+\frac{1}{N}\right]-\frac{\sigma_q^2}{N}\text{ tr }\{(R_{\underline{x}^q})^{-1}\}\right], \tag{37}$$

for large K and under high resolution assumption. The corresponding expression for the coding gain is

$$\begin{aligned}G_{\widehat{V},K,q} = {} & \frac{\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{I,K,q}}{\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{\widehat{V},K,q}} \approx G^0\,\frac{\left(\det(I+\sigma_q^2(\text{ diag }\{R_{\underline{x}}\})^{-1})\right)^{1/N}}{\left(\det(I+\sigma_q^2(R_{\underline{x}})^{-1})\right)^{1/N}} \\ & \times\frac{\left[1+\frac{1}{K}(1-\frac{1}{N^2}\text{ tr }\{R_{\underline{x}^q}(\text{ diag }\{R_{\underline{x}^q}\})^{-1}R_{\underline{x}^q}(\text{ diag }\{R_{\underline{x}^q}\})^{-1}\})-\frac{\sigma_q^2}{N}\text{ tr }\{(\text{ diag }\{R_{\underline{x}^q}\})^{-1}\}\right]}{\left[1+\frac{N-1}{K}(\frac{1}{2}+\frac{1}{N})-\frac{\sigma_q^2}{N}\text{ tr }\{(R_{\underline{x}^q})^{-1}\}\right]}.\end{aligned} \tag{38}$$

The above expression exhibit three kinds of terms : those regarding estimation noise only (through K), those regarding quantization noise only (through σ_q^2), and cross influence terms.

C. Causal Transform (CT)

In the causal case, an estimate $\widehat{L}'$ is computed from $\widehat{R}_{\underline{x}^q}$, and the actual variances are $\sigma'^2_{y_i} = \mathbb{E}\,(\widehat{L}' R_{\underline{x}^q} \widehat{L}'^T - \sigma_q^2 I)_{ii}$. Thus, when the transformation is based on K quantized vectors (for high K and under high resolution assumption) the distortion becomes

$$\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{\widehat{L}',K,q} = \mathbb{E}\sum_{i=1}^{N} c2^{-2[b+\frac{1}{2}\log_2 \frac{(\widehat{L}'\widehat{R}_{\underline{x}^q}\widehat{L}'^T)_{ii}}{(\prod_{i=1}^{N}(\widehat{L}'\widehat{R}_{\underline{x}^q}\widehat{L}'^T)_{ii})^{\frac{1}{N}}}]} (\widehat{L}' R_{\underline{x}^q} \widehat{L}'^T - \sigma_q^2 I)_{ii}. \qquad (39)$$

The above expression leads to

$$\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{\widehat{L}',K,q} \approx \mathbb{E}\,\|\underline{\tilde{y}}\|^2_{L} \left(\det(I + \sigma_q^2 (R_{\underline{x}})^{-1})\right)^{1/N} \left[1 + \frac{N-1}{K}\left[\frac{1}{2} + \frac{1}{N}\right] - \frac{\sigma_q^2}{N} \operatorname{tr}\{(R'_{\underline{y}})^{-1}\}\right]. \qquad (40)$$

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

$$G_{\widehat{L}',K,q} = \frac{\mathbb{E}\,\|\underline{\tilde{x}}\|^2_{I,K,q}}{\mathbb{E}\,\|\underline{\tilde{y}}\|^2_{\widehat{L}',K,q}} \approx G^0 \frac{\left(\det(I + \sigma_q^2 (\operatorname{diag}\{R_{\underline{x}}\})^{-1})\right)^{1/N}}{\left(\det(I + \sigma_q^2 (R_{\underline{x}})^{-1})\right)^{1/N}}$$
$$\times \frac{\left[1 + \frac{1}{K}\left[1 - \frac{1}{N^2}\operatorname{tr}\{R_{\underline{x}^q}(\operatorname{diag}\{R_{\underline{x}^q}\})^{-1} R_{\underline{x}^q}(\operatorname{diag}\{R_{\underline{x}^q}\})^{-1}\}\right] - \frac{\sigma_q^2}{N}\operatorname{tr}\{(\operatorname{diag}\{R_{\underline{x}^q}\})^{-1}\}\right]}{\left[1 + \frac{N-1}{K}\left[\frac{1}{2} + \frac{1}{N}\right] - \frac{\sigma_q^2}{N}\operatorname{tr}\{(L' R_{\underline{x}^q} L'^T)^{-1}\}\right]}. \qquad (41)$$

Again, perturbation terms regarding the influence of quantization, estimation noise, and both can be identified.

It can be checked that the expressions (41) and (38) tend to (17) and (22) respectively as $K \to \infty$, and both to (31) as $\sigma_q^2 \to 0$. This means indeed that as $K \to \infty$, the estimation noise vanishes, and we face a quantization noise problem only, which leads to the results of Sec. III. As $\sigma_q^2 \to 0$ also, only estimation noise remains, which leads to the results of Sec. IV.

VI. SIMULATIONS

For the simulations, we generated real Gaussian i.i.d. vectors with covariance matrix $R_{\underline{x}_j} = H_j R_{AR1} H_j^t$, $j = 1, 2$. R_{AR1} denotes the covariance matrix of a first order autoregressive process with normalized cross correlation coefficient ρ. H_j is a diagonal matrix whose ith entry is $i^{1/3}$ for H_1 (increasing variances), and $(N-i+1)^{1/3}$ for H_2 (decreasing variances). The goal of these numerical evaluations is first to check whether the generic distortion as described in eq. (3) (and the corresponding coding gains) corresponds to their theoretical expressions derived in the three cases of quantization, estimation noise, and both. Also, these curves may give more visual insight on the actual behavior of the backward adaptive TC schemes than the mathematical expressions may do. In the curves, G^0 correspond the maximum gain in TC as defined in eq. (1). The following algorithms were therefore used to check our analytical results.

A. Quantization Noise

For several rates (from 2 to 6 b/s), bit assignments and transforms ($\widehat{T} = I, L'$ and V respectively) were computed using $\widehat{R}_{\underline{x}} = R_{\underline{x}_j} + \sigma_q^2 I$, where $\sigma_q^2 = c 2^{-2r} \det R_{\underline{x}_j}{}^{1/N}$ (that is, the distortion occurring in a high rate transform coding framework with optimal bit assignment). The choice of the constant is not relevant because (3) is very general; we chose $c = \frac{\pi e}{6}$ which correspond to entropy coded uniform quantization. The bits to be allocated were computed by (2), with the appropriate $\widehat{T}$ and ${\sigma^2}'_{y_i}$ for the three cases. In a similar manner, the corresponding distortions where computed using (3). These resulting distortions were then used to compute the coding gains, which were compared with the theoretical expressions.

- In Figure 2, G^0 is the upper straight line. The coding gain with quantization noise is plotted for the KLT (upper solid curves) and the CT (lower solid curves), for signals of decreasing variances, and with $\rho = 0.9$, $N = 4$. The theoretical exact expressions are given by (17) and (22), the corresponding curves are dotted. The theoretical approximated expressions are given by (18) and (23), and the corresponding curves are dashed.
- Figure 3 shows the influence of the variance ordering in the decorrelation process. The upper curves

(solid: observed and dots: theoretical) depict the gain obtained with the CT by decorrelating the signals by decreasing order of variance ($R_{\underline{x}_2}$), and the lower curves (solid and diamond) by increasing order ($R_{\underline{x}_1}$). The theoretical expression is eq. (22).

From these Figures, it is checked that the expressions (17) and (22) are actually exact. From Fig. 2, approximated expressions (18) and (23) match their exact counterparts as the rate increases. The performances of the CT are slightly inferior to those of the KLT (from a few percents) and vanishes at high rates. From Fig. 3, it appears that processing the signals by order of decreasing variance maximizes the coding gain, as discussed in Sec. III-C.

B. Estimation Noise

In this case, estimates of the covariance matrix of the data were computed using K vectors by $\frac{1}{K}\sum_{i=1}^{K}\underline{x}_i\underline{x}_i^t$, $K = N, N{+}1, \cdots, 10^3$. For each estimate $\widehat{R}_{\underline{x}}$, the transforms $\widehat{T} = \widehat{V}, \widehat{L}$ were computed so that $\widehat{T}\widehat{R}_{\underline{x}}\widehat{T}^t$ is diagonal, and the bit assignments were computed using estimates of the variances $(\widehat{T}\widehat{R}_{\underline{x}}\widehat{T}^t)_{ii}$. In order to evaluate the expected distortion (3), the sum in (3) was considered as a random variable, whose expectation was evaluated by Monte Carlo simulations. This was done for the Identity transform, in the causal and in the unitary case. The coding gains in presence of estimation noise are compared for $N = 4$ and $\rho = 0.9$. The ratio of the corresponding distortions are the "Observed G" in Figure 4. The corresponding theoretical expression ("Theoretical G") is given by (31) (it should be the same for the KLT and CT because both transforms are decorrelating and unimodular). G^0 is the upper straight line.

As expected, there is no difference between the unitary and the causal case. Our calculations assume small perturbations (large K). It can be observed that the model matches the actual coding gain after a few tens of vectors. Backward adaptive systems yield similar performances as systems designed with the knowledge of $R_{\underline{x}}$ after a few hundreds of decoded vectors. Note also that it is always useful to use backward adaptive TC schemes (the coding gain is superior to 1 for $K > N + 1$).

C. Quantization and Estimation Noise

In this case, the quantized vectors were obtained for each rate r by adding to the sets of i.i.d. Gaussian vectors uncorrelated white noise vectors with covariance matrix $\sigma_q^2 I = c2^{-2r}(\det R_{\underline{x}})^{\frac{1}{N}} I$. For each set of K quantized vectors, an estimate of the covariance matrix of the data was computed by $\frac{1}{K}\sum_{i=1}^{K} \underline{x}_i^q \underline{x}_i^{qt}$, $K = N, N+1, \cdots, 10^3$. Again, for each estimate $\widehat{R}_{\underline{x}^q}$, the transforms $\widehat{T} = \widehat{V}, \widehat{L}$ were computed so that $\widehat{T}\widehat{R}_{\underline{x}^q}\widehat{T}^t$ is diagonal, and the bit assignments were computed using estimates of the variances $(\widehat{T}\widehat{R}_{\underline{x}^q}\widehat{T}^t)_{ii}$. In order to evaluate the expected distortion for the three transformations, the sum was considered as a random variable, whose expectation was evaluated by Monte Carlo simulations. The ratio of the corresponding distortions are the "Observed Gains" of the following figures. The theoretical gains are given by (38) for KLT and (41) for LDU.

- The coding gains in presence of estimation and quantization noise are compared for KLT and CT (signals of decreasing variances) in Figure 5 for $N = 4$, $\rho = 0.9$ and a rate of 3 bits per sample. Upper straight line is G^0. The upper solid line curve is the theoretical coding gain for KLT, and the lower solid line curve the theoretical coding gain for CT. The upper dashed curve is the observed coding gain for KLT, and the lower dashed curve the observed coding gain for CT.

The observed behaviors of the transformation are relatively well matched by the theoretically predicted ones as K amounts to a few tens. As K amounts to a few hundreds, the performances of on-line systems approach those of systems designed with the optimal transforms and bit assignment. The performances of the CT are slightly inferior to those of the KLT. This difference vanishes at high rates (cf Fig. 2). In Fig. 5, the coding gains toward which both the KLT and the CT system converge can be read from Fig. 3, with $r = 3$ b/s.

- The influence of the ordering of the signals for the same parameters as above is plotted in Figure 6.

In the limit of large K, the actual gains converge to the results obtained in the case where quantization noise only is considered (the estimation noise vanishes). The proposed model matches the actual convergence behaviors in the causal and unitary cases after a few tens of decoded vectors. Finally, decorrelating the signals by order of decreasing variance appears the best strategy.

VII. Summary and Conclusions

We proposed an analytical model for the performances of causal and unitary on-line TC schemes. We described the effects of backward adaptation as perturbation effects : backward adaptation impacts the ideal high rate TC framework by perturbing both the transforms' design and the bit assignment mechanism.

It appears that as quantization noise only is considered, only the bit assignment mechanism is perturbed for the KLT (2nd order perturbation term), whereas the CT suffers additionally from quantization noise feedback (1st order term). As one accounts for estimation noise only, both transforms present the same performances because they are both decorrelating and unimodular. As both types of perturbations are accounted for, the CT remains slightly inferior from a rate-distortion point of view to its unitary counterpart because of the quantization noise feedback. This drawback vanishes at high rates. It can be minimized if the signals get decorrelated by order of decreasing variances.

As K amounts to a few hundreds, the performances of on-line TC systems approach those of systems designed with the optimal transforms and bit assignment. The on-line TC systems modeled by eq. (2) and (3) are advantageous w.r.t. a system using no transform for values of K larger than $\approx N+1$ vectors.

The results of simulation show that the analytical description of the considered systems is fairly accurate. We provided exact expressions for the coding gains as far as the quantization noise only is concerned. When estimation noise is accounted for, the proposed analysis reliably estimates the distortions and the corresponding coding gains after a few tens of decoded vectors.

As a follow-up of these works, we are currently investigating systems using different bit assignment mechanisms than that assumed in eq. (2).

REFERENCES

[1] S.-M. Phoong and Y.-P. Lin, "Prediction-based lower triangular transform," *IEEE trans. on Sig. Proc.*, July 2000.

[2] D. Mary and D.T.M. Slock, "On quantization noise feedback in causal tranform coding," 2004, Submitted to IEEE Trans. on Signal Processing.

[3] D. Mary, *Causal Lossy and Lossless Coding of Vectorial Signals*, Ph.D. thesis, ENST, Paris, March 2003.

[4] N.S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice Hall, 1984.

[5] A. Gersho and R.M. Gray, *Vector quantization and signal compression*, Kluwer Academic, 1992.

[6] V.K. Goyal, "Theoretical foundations of transform coding," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 9–21, Sept. 2001.

[7] M. Effros, H. Feng, and K. Zeger, "Suboptimality of Karhunen-Loeve transform for transform coding," *IEEE Trans. on Inf. Th.*, pp. 1605–1619, August 2004.

[8] M. Effros, "Rate-distortion bounds for fixed- and variable-rate multiresolution sources codes," Submitted to IEEE Trans. on Inf. Theory on March 26, 1998.

[9] K. Karhunen, "Uber lineare Methoden in der Wahrscheinlichkeitsrechnung," *Ann. Acad. Sci. Fenn., Ser. A1,: Math.-Phys.*, vol. 37, pp. 3–79, 1947.

[10] M. Loeve, *Processus stochastiques et mouvements Browniens*, chapter : Fonctions aleatoires de second ordre, P. Levy, Ed. Paris, France : Gauthier-Villars, 1948.

[11] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *J. Educ. Psychology*, vol. 24, pp. 417–441, 498–520, 1933.

[12] V. Goyal, J. Zhuang, and M. Vetterli, "Transform coding with backward adaptive updates," *IEEE Trans. on Inf. Th.*, vol. 46, no. 4, July 2000.

[13] Habibi A. and Hershell R.S., "A unified framework of differential pulse-code modulation (DPCM and transform coding systems," *IEEE Trans. on Com.*, pp. 692–696, May 1974.

[14] S.M. Phoong and Y.P. Lin, "PLT versus KLT," in *IEEE Int. Symp. Circ. Syst.*, May 1999.

[15] D. Mary and D. T. M. Slock, "Codage DPCM vectoriel et application au codage de la parole en bande elargie," in *CORESA 2000*, Poitiers, France, October 2000.

[16] F. Lahouti and A.K. Khandani, "Sequential vector decorrelation technique," Tech. Rep., Univ. of Waterloo, 2001.

[17] D. L. Neuhoff, R. M. Gray, and L. D. Davisson, "Fixed rate universal block source coding with a fidelity criterion," *IEEE Trans. Inf. Theory*, vol. IT-21, pp. 511–523, Sept. 1975.

[18] P. A. Chou, M. Effros, and R.M. Gray, "A vector quantization approach to universal noiseless coding and quantization," *IEEE Trans. on Inf. Theory*, vol. 42, no. 4, pp. 1109–1138, July 1996.

[19] M. Effros and P. A. Chou, "Weighted universal transform coding: Universal image compression with the Karhunen-Love transform," in *Proc. Int. Conf. Image Processing*, Oct. 1995, vol. II, pp. 61–64.

[20] D. Mary and D. T. M. Slock, "On the suboptimality of orthogonal transforms for single- or multi-stage lossless transform

coding," in *DCC*, 2003.

[21] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, NJ, 1993.

[22] J. Makhoul, "Linear prediction : a tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, April 1975.

List of Figures

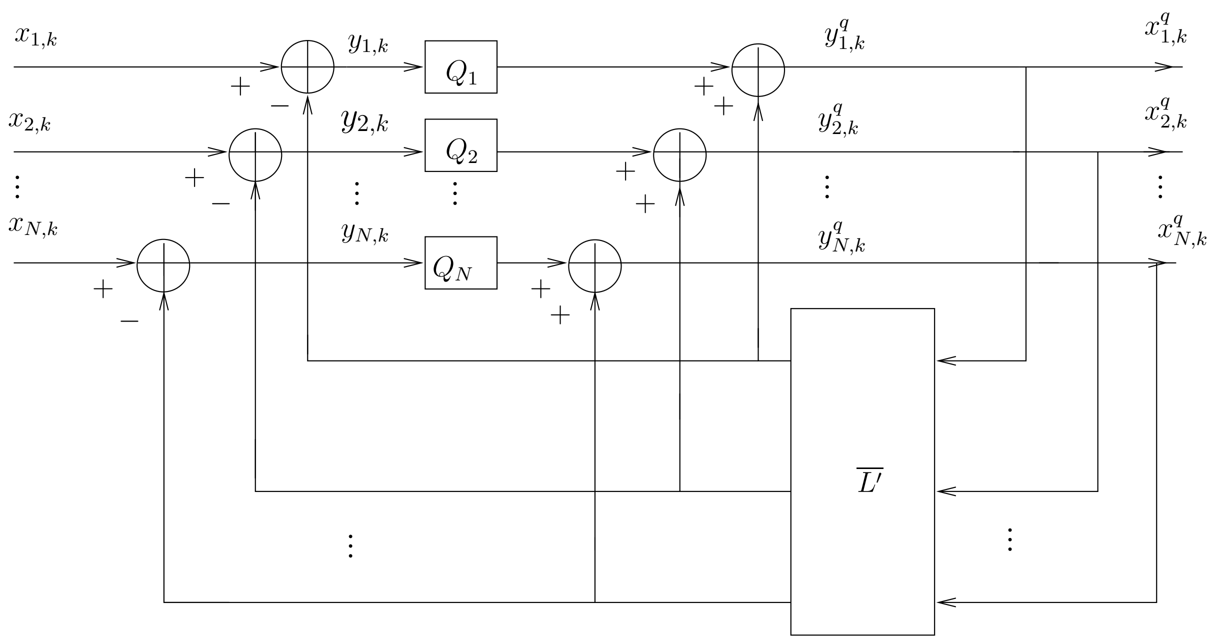
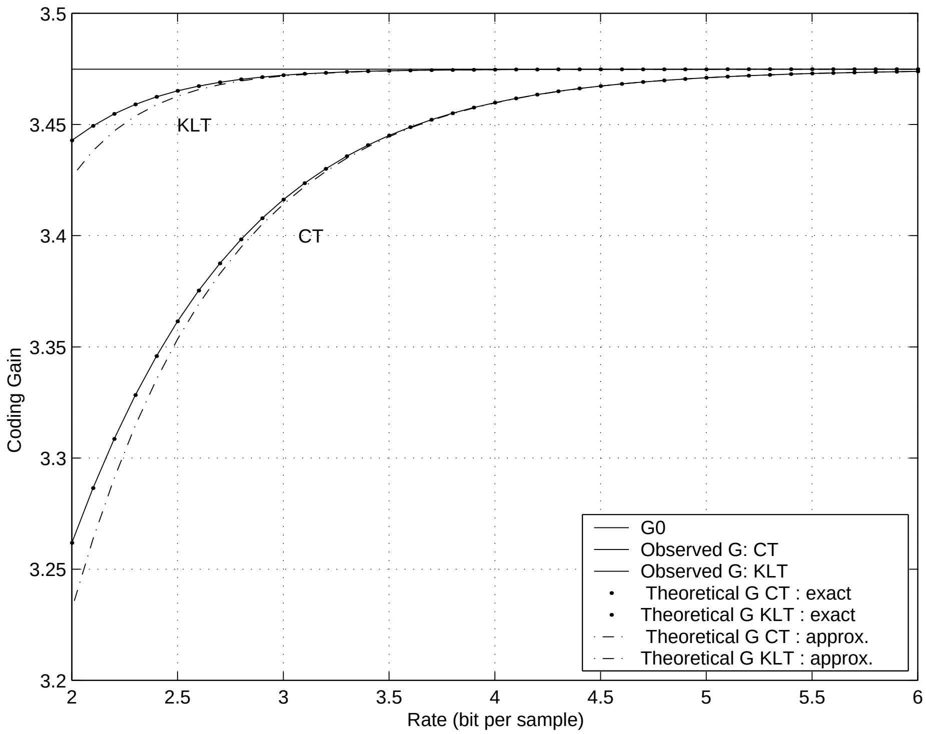
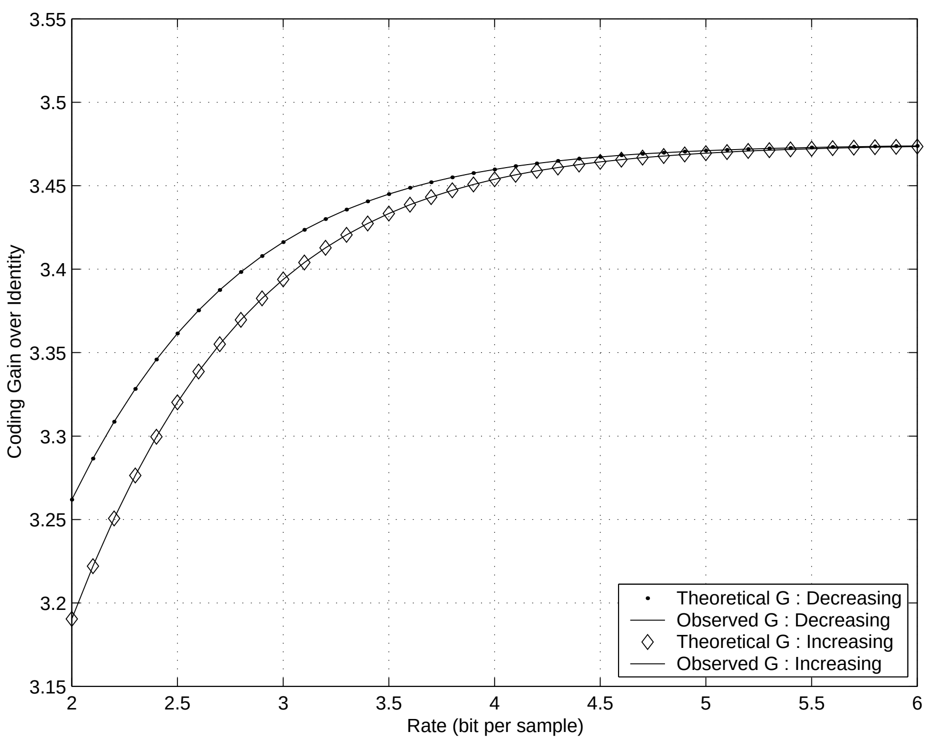


Fig. 1.



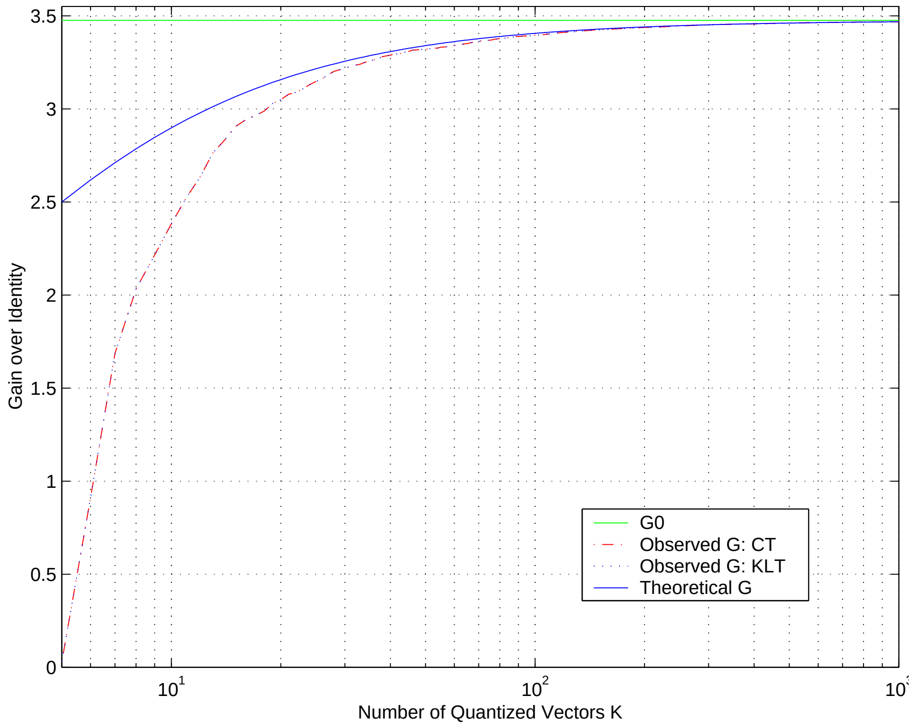
3.5
3.45
3.4
3.35
3.3
3.25
3.2
Coding Gain
KLT
CT
G0
Observed G: CT
Observed G: KLT
Theoretical G CT : exact
Theoretical G KLT : exact
Theoretical G CT : approx.
Theoretical G KLT : approx.
2
2.5
3
3.5
4
4.5
5
5.5
6
Rate (bit per sample)

Fig. 2.



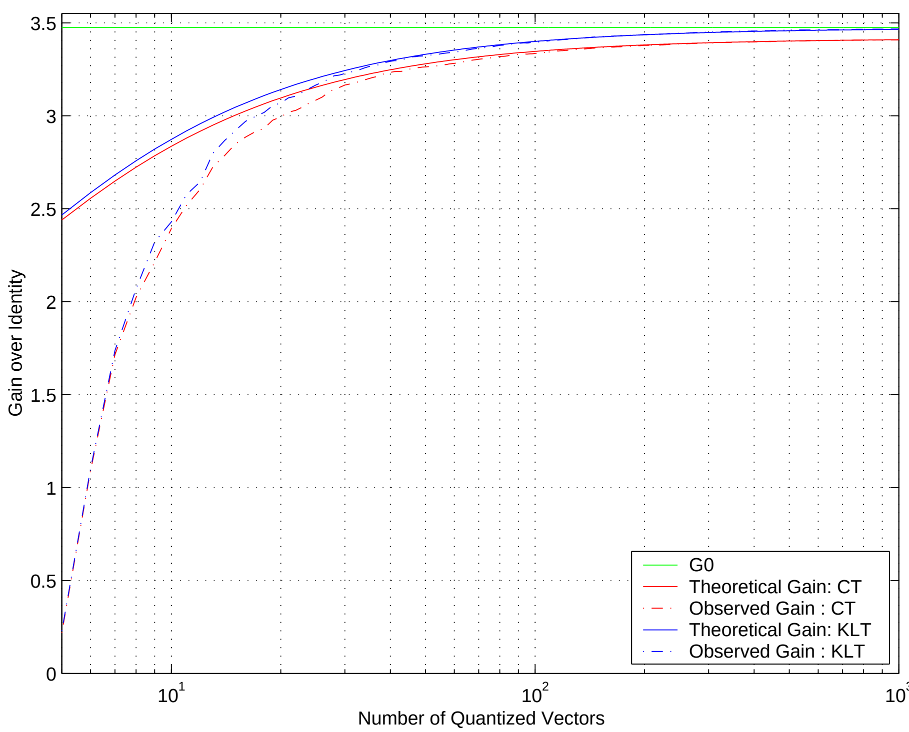
3.55
3.5
3.45
3.4
3.35
3.3
3.25
3.2
3.15
Coding Gain over Identity
Theoretical G : Decreasing
Observed G : Decreasing
Theoretical G : Increasing
Observed G : Increasing
2
2.5
3
3.5
4
4.5
5
5.5
6
Rate (bit per sample)

Fig. 3.



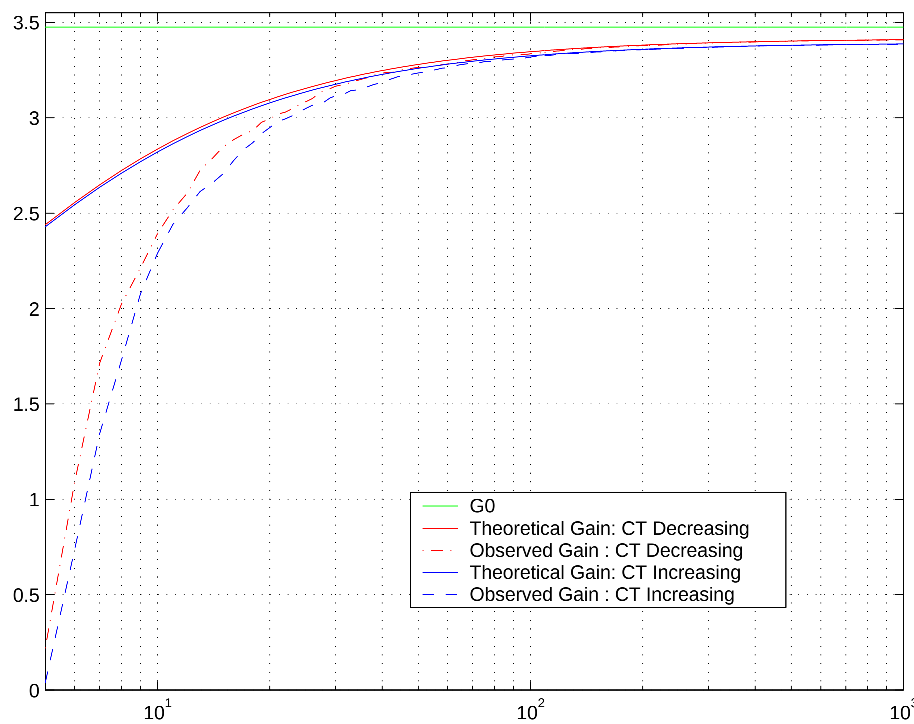
3.5
3
2.5
2
1.5
1
0.5
0
Gain over Identity
10^1
10^2
10^3
Number of Quantized Vectors K
G0
Observed G: CT
Observed G: KLT
Theoretical G

Fig. 4.



3.5
3
2.5
2
1.5
1
0.5
0
Gain over Identity
10^1
10^2
10^3
Number of Quantized Vectors
G0
Theoretical Gain: CT
Observed Gain : CT
Theoretical Gain: KLT
Observed Gain : KLT

Fig. 5.



3.5
3
2.5
2
1.5
1
0.5
0
10^1
10^2
10^3
G0
Theoretical Gain: CT Decreasing
Observed Gain : CT Decreasing
Theoretical Gain: CT Increasing
Observed Gain : CT Increasing

Fig. 6.