Mutistage Integer-to-Integer Multichannel Prediction for Scalable Lossless Coding

D. Mary, D .T. M. Slock
Département de Communications Mobiles
Institut EURÉCOM*, 2229, Route des Crêtes,
06904, Sophia Antipolis Cedex, FRANCE

Abstract

We present a lossless coding procedure based on a recently introduced decorrelating scheme [1, 2], where both intra- and interchannel redundancies are removed by lossless prediction. The resulting signals are scalar entropy coded. We show that for continuous-amplitude Gaussian sources discretized with uniform scalar quantizers, no suboptimality occurs in the proposed lossless coding scheme by using scalar instead of vector entropy coders, except from a term caused by the lossless constraint, which vanishes in the limit of small distortions. This lossless constraint is described in terms of excess bit rate. The proposed coder may be used either as a compressor, or as a scalable lossless coder. In this case, a multistage version of the lossless coder based on ADPCM-like lossless prediction loops allows one to transmit the data by means of substreams, which represent different "resolution" levels. We show that this multiresolution approach is slightly suboptimal in comparison with a single global compression because of the noise feedback created in the ADPCM-like loops, but not of the "space filling" loss of the scalar quantizers. We propose a strategy to fix the stepsizes of these quantizers so that the delivered rates approach some predetermined target rates.

1 Introduction

1.1 Lossless Coding

A general framework for lossless coding can be depicted by Figure 1. Assume one disposes of a continuous-amplitude vectorial source $\{\underline{x}^c\}$. In a first step, this source is quantized. The box Q may for example represent the discretization realized by A/D converters, which may be followed by lossy source coders. The effect of Q is to provide a binary representation of $\{\underline{x}^c\}$ by means of a discrete amplitude

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

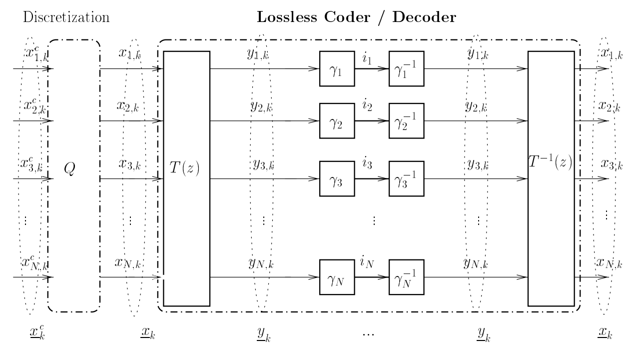

Figure 1: *General framework of a lossless coding scheme.*

source $\underline{x}$, whose samples are $\underline{x}_k = [x_{1,k}...x_{N,k}]^T$, where k refers to time. ($\underline{x}$ may in general, as $\underline{x}^c$, present both temporal and spatial dependencies, that is, dependencies between the $x_{i,k}$ and $x_{j,l}$ $\forall i, j, k$ and l.) The price paid for this digitization is indeed the introduction of some distortion, since some information about $\underline{x}^c$ has been lost. The aim of lossy coding is to optimize Q so that for a given distortion, the bitrate required to represent $\underline{x}$ will be minimal. Once some rate-distortion trade-off has been chosen, both rate and distortion are fixed to, say, some values r_0 and d_0. By the noiseless coding theorem of Shannon, r_0 is the entropy rate of the discrete-amplitude source $\underline{x}$. The aim of lossless coding is then nothing else that designing a coding procedure with the goal that the actual bitrate required to code $\underline{x}$ will be as small as possible, and, if possible, will reach r_0. Indeed, it is known that entropy coders which assigns adequate codewords to blocks of samples $\underline{x}_k$, according to the joint probability of these vectors, can reach r_0. The complexity of these vector entropy coders may, however, be prohibitive. Thus, an interesting question is that of designing a coding procedure which is performant in terms of rates, though maintaining a reasonnable complexity, by using scalar entropy coders. Assume in a first scenario that the components $\{x_i\}$ of the vectors $\underline{x}_k$ in Figure (1) are directly scalar entropy coded (entropy coders γ_i), resulting in a bitrate $r_{scal}(\underline{x})$. Assume in a second scenario that a reversible transformation $T(z)$, which tends to remove the intra- and inter- channel dependencies, is applied to $\underline{x}_k$ before scalar entropy coding, resulting in a bitrate $r_{scal}(\underline{y})$. Then one may define for this transform a lossless coding gain as $G_{T(z)} = r_{scal}(\underline{x}) - r_{scal}(\underline{y})$ bits per sample. One may indeed expect that $G_{T(z)}$ is upper bounded by some $G_{Max} = r_{scal}(\underline{x}) - r_0$.

1.2 Scalable lossless audio coding

Both temporal and spatial redundancies can be found in audio signals : starting from the monophonic and stereophonic technologies, new systems such as quadraphonic, 5.1, and up to 10.2 channels are now available. Thus, the field of audio coding appears as a natural space of application for lossless coding techniques, though these could be applied to the wide class of the vectorial sources.
However, most of the state of the art lossless audio coders do not take into account interchannel redundancy at all, or in a basic way only [3]. As seen in the previous subsection, this paper proposes a coding procedure where both types of redundancy are removed by lossless prediction.
As many audio data transfers are completed through the internet, scalable audio coders are particularily demanded (due to bandwidth constraints, browsing applications,...). These coders allow one to transmit in a first step a low resolution (lossy, low bitrate) version of the data, and to send the complement in a later stage. Such coders are proposed in Sections 4 and 5.

1.3 Overview

The paper is organized as follows : in Section 2, we derive the expression of G_{Max} for Gaussian sources with memory. We show in Section 3 that G_{Max} can almost by reached by an integer-to-integer implementation of a totally decorrelating transform, generalizing previous results on integer-to-integer transforms (causal LDU transform [4], unitary Karhunen-Loève transform [5]). This structure will be refered to as "one shot" multichannel prediction. The fourth part describes a lossless coding scheme allowing to progressively transmit the data by means of two complementary substreams. Finally, this structure is generalized in the last part, resulting in a multistage structure for the integer-to-integer prediction.

2 Minimum Birates and Maximum Lossless Coding Gain

2.1 Memoryless Gaussian sources

In the case where $\{\underline{x}\}$ is a uniformly quantized version of a memoryless Gaussian source $\{\underline{x}^c\}$, it was shown in [4] that the maximum lossless coding gain, that is, the number of bits per sample which can be saved by taking advantage of the redundancy of the source $\{\underline{x}\}$ is

$$G_{Max} = \frac{1}{2N} \log_2 \frac{\det diag\{R_{\underline{x}^c\underline{x}^c}\}}{\det R_{\underline{x}^c\underline{x}^c}}. \qquad (1)$$

2.2 Gaussian sources with memory

We first derive the minimal rate $r_0(\underline{x})$ required to represent the discrete-amplitude source $\{\underline{x}\}$, obtained from $\underline{x}^c$ by some discretization (or quantization) process, see Figure (1). Since $\underline{x}$ is a source with memory this rate corresponds, by the noiseless coding theorem of Shannon, to the entropy rate of the source

$$r_0(\underline{x}) = H_\infty(\underline{x}) + \epsilon \text{ bits per sample}, \qquad (2)$$

where H_∞ denotes the entropy rate, and ϵ is a positive value which can be made arbitrarily close to zero by means of vector entropy coders. We assume now that $\underline{x}$ is an uniformly quantized version of $\underline{x}^c$ with stepsizes Δ_i. Let the samples of $\{\underline{x}\}$ be collected in a vector $\underline{X}_k = [\underline{x}_1 ... \underline{x}_k]^T$ and denote by $\underline{X}_k^c$ the corresponding vector of samples of $\{\underline{x}^c\}$. The entropy rate $r_0(\underline{x})$ may then be written as

$$r_0(\underline{x}) = \lim_{k\to\infty} \frac{1}{Nk} H(\underline{X}_k). \qquad (3)$$

Now, for any continuous-amplitude source x_i^c uniformly quantized with stepsize Δ_i, the differential entropy $h(x_i^c)$ can be related to the discrete entropy $H(x_i)$ by [6]

$$H(x_i) + \log_2 \Delta_i \to h(x_i^c) \;\; as \;\; \Delta_i \to 0. \qquad (4)$$

This result can be extended to the NK-vector $\underline{X}_k$ (see [7] for a proof), leading, for Gaussian sources, to

$$\begin{aligned} r_0(\underline{x}) &= \lim_{k\to\infty} \tfrac{1}{Nk} h(\underline{X}_k^c) + \tfrac{1}{N}\sum_{i=1}^{N} \log_2 \Delta_i \\ &\approx \lim_{k\to\infty} \tfrac{1}{2}\log_2 2\pi e \left(\det R_{\underline{X}_k^c\underline{X}_k^c}\right)^{\frac{1}{N}} + \tfrac{1}{N}\sum_{i=1}^{N} \log_2 \Delta_i . \end{aligned} \qquad (5)$$

The Szego formula expresses the limit of the previous determinant as

$$\lim_{k\to\infty} \left(\det R_{\underline{X}_k^c\underline{X}_k^c}\right)^{\frac{1}{k}} = e^{\int_{-\frac{1}{2}}^{\frac{1}{2}} \det S_{\underline{x}^c\underline{x}^c}(f)\,df}, \qquad (6)$$

where $S_{\underline{x}^c\underline{x}^c}(f)$ is the power spectral density of the vectorial process $\{\underline{x}^c\}$. Thus, the minimum bitrate required to code the source $\{\underline{x}\}$ is

$$r_0(\underline{x}) = \frac{1}{2}\log_2 2\pi e\left(e^{\int_{-\frac{1}{2}}^{\frac{1}{2}} \det S_{\underline{x}^c\underline{x}^c}(f)\, df}\right)^{\frac{1}{N}} + \frac{1}{N}\sum_{i=1}^{N}\log_2\Delta_i. \quad (7)$$

As mentionned previously, this bitrate can be attained by vector entropy coding. If now we use scalar entropy coders to code the x_i, the bitrate becomes

$$\begin{aligned} & r_{scal}(\underline{x}) = \tfrac{1}{N}\sum_{i=1}^{N} H(x_i) \\ & \approx \tfrac{1}{N}\sum_{i=1}^{N} h(x_i) - \tfrac{1}{N}\sum_{i=1}^{N} \log_2 \Delta_i \\ & \approx \tfrac{1}{2}\log_2 2\pi e\left(\det(diag\{R_{\underline{x}^c\underline{x}^c}\})\right)^{\frac{1}{N}} - \tfrac{1}{N}\sum_{i=1}^{N} \log_2 \Delta_i, \end{aligned} \quad (8)$$

where $diag\{.\}$ denotes the diagonal matrix made with the diagonal elements of $\{.\}$. Finally, the maximum lossless coding gain corresponding to the bitrate saved by using vector instead of scalar entropy coders is

$$G_{Max} = r_{scal}(\underline{x}) - r_0(\underline{x}) \approx \frac{1}{2N}\log_2\frac{\det diag\{R_{\underline{x}^c\underline{x}^c}\}}{e^{\int_{-1/2}^{1/2}\ln\det S_{\underline{x}^c\underline{x}^c}(f)df}}. \quad (9)$$

Note that G_{Max} does not depend on the quantization stage but only on the spatial and temporal dependencies of the continuous amplitude sources x_i^c. Also, (1) is indeed a special case of (9), since in the case of memoryless sources, $S_{\underline{x}^c\underline{x}^c}(f)$ becomes $R_{\underline{x}^c\underline{x}^c}$, and $e^{\int_{-1/2}^{1/2}\ln\det S_{\underline{x}^c\underline{x}^c}(f)df}$ reduces to $\det R_{\underline{x}^c\underline{x}^c}$.

Vector entropy coding, though optimal, suffers however from a complexity which may prohibitive. In the next section, we show that G_{Max} can almost be reached by using a lossless transformation followed by scalar entropy coders.

3 "One Shot" Integer-to-Integer Multichannel Prediction

3.1 Triangular MIMO prediction

Consider the coding scheme of figure (2), described for $N = 2$. $L_{ij}(z)$, $i \neq j$ are Wiener filters, and $L_{ii}(z)$ are optimal causal linear prediction filters. The rounding operations denoted by Δ_i (high resolution is assumed) ensure the losslessness of the structure: each $\hat{x}_i$ is quantized to the same multiple of Δ_i as x_i, and x_i is losslessly recovered, at the decoder, by $y_i + \hat{x}_i$. The y_i are then scalar entropy coded. This scheme is a straightforward application of the triangular MIMO (Multi-Input Multi-Output) prediction to lossless coding [2]. The entries of the lower triangular MIMO prediction matrix $L(z)$ (which may be written as $I - \overline{L}(z)$, where I is the identity matrix) are $L_{ij}(z)$.

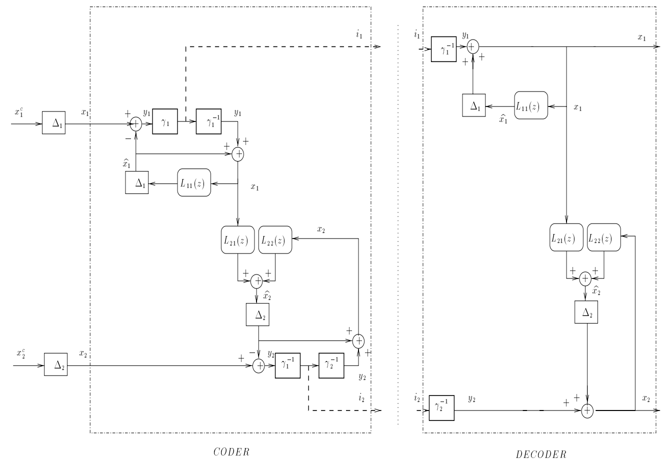

Figure 2: *"One shot" integer-to-integer multichannel prediction for $N = 2$.*

The triangular MIMO predictor can be seen, as the classical MIMO prediction, as a particular case of the generalized MIMO prediction [2]. The triangular case presents however over its classical counterpart several attractive coding advantages [8]. An application of the classical MIMO prediction was recently applied to lossless coding [9]. Concerning the decorrelating performances of the triangular transform, the following result was shown in [2].

Result 1: Suppose that we apply $L(z)$ to decorrelate some real valued vectorial source $\{\underline{w}\}$. The resulting vectorial process is $\{\underline{v}\}$. Then each value $v_{i,k}$ is the optimal prediction $w_{i,k}$ based on the past samples of w_i only, and on all the samples of the sources w_j, for all $j < i$. The vectorial process $\{\underline{v}\}$ is then totally decorrelated with diagonal covariance matrix $R_{\underline{v}\underline{v}}$:

$$R_{\underline{v}\underline{v}} = S_{\underline{v}\underline{v}}(f) = L(f)S_{\underline{w}\underline{w}}(f)L^T(f) = e^{\int_{-1/2}^{1/2}\ln\det S_{\underline{w}\underline{w}}df}. \quad (10)$$

3.2 Coding gain

Coming back to Figure (2), and assuming that the $\{y_i\}$ are scalar entropy coded, we can define the gain for the lossless implementation of the transform $L(z)$, $G_{L(z)}$, as the difference $r_{scal}(\underline{x}) - r_{scal,L(z)}(\underline{y})$, where $r_{scal}(\underline{x})$ is as in (8), and $r_{scal,L(z)}(\underline{y})$ is the bitrate required to scalar entropy code the decorrelated transform components y_i. We shall thus compute

$$\begin{aligned} G_{L(z)} &= r_{scal}(\underline{x}) - r_{scal,L(z)}(\underline{y}) \\ &= \tfrac{1}{N}\sum_{i=1}^{N} H(x_i) - \tfrac{1}{N}\sum_{i=1}^{N} H(y_i). \end{aligned} \quad (11)$$

Let us denote by $y_{i,k}^0$ the optimal prediction error obtained by applying $L(z)$ to $\underline{x}$ (that is, without the rounding operations ensuring the losslessness). Then the $y_{i,k}$ can be related to the $y^0_{i,k}$ by

$$\begin{aligned} \underline{y}_k &= \underline{x}_k - [\overline{L}(q)\underline{x}_k]_{\Delta_i} \\ &= [\underline{x}_k - \overline{L}(q)\underline{x}_k]_{\Delta_i} = [\underline{y}^0_k]_{\Delta_i}, \end{aligned} \qquad (12)$$

where $[\underline{y}^0]_{\Delta_i}$ denotes quantization with stepsize Δ_i of the ith component of $\underline{y}^0$, and the notation (q) denotes the unit delay operator. Thus, $y_{i,k}$ may be seen as the optimal prediction error $y^0_{i,k}$ quantized with the same stepsize as $x_{i,k}$. Now, applying the result (10) to the decorrelation of the process $\{\underline{x}\}$, the bitrate $r_{scal,L(z)}(\underline{y})$ may be written as

$$\begin{aligned} r_{scal,L(z)}(\underline{y}) &= \tfrac{1}{N}\sum_{j=1}^{N} i_j = \tfrac{1}{N}\sum_{i=1}^{N} h(x_i) - \tfrac{1}{N}\sum_{i=1}^{N}\log_2\Delta_i \\ &\approx \tfrac{1}{2}\log_2 2\pi e e^{(\int_{-1/2}^{1/2}\ln\det S_{\underline{x}\underline{x}}(f)df)^{\frac{1}{N}}} - \tfrac{1}{N}\sum_{i=1}^{N}\log_2\Delta_i. \end{aligned} \qquad (13)$$

Using (8), (11) and (13), we get the following expression of the gain :

$$G_{L(z)} = \frac{1}{2N}\log_2\frac{\det diag\{R_{\underline{x}^c\underline{x}^c}\}}{e^{\int_{-1/2}^{1/2}\ln\det S_{\underline{x}\underline{x}}(f)df}}. \qquad (14)$$

In the case of equal $\Delta_i = \Delta_{VHR}$, expression (14) may be approximated as

$$G_{L(z)} \approx \underbrace{\frac{1}{2N}\log_2\frac{\det diag\{R_{\underline{x}^c\underline{x}^c}\}}{e^{\int_{-1/2}^{1/2}\ln\det S_{\underline{x}^c\underline{x}^c}df}}}_{G_{Max}} - \underbrace{\frac{\Delta^2_{VHR}}{24N\ln 2}\left(\int_{-1/2}^{1/2} tr\, S^{-1}_{\underline{x}^c\underline{x}^c}(f)df\right)}_{Excess\ bit\ rate\ due\ to\ the\ lossless\ constraint}, \qquad (15)$$

where tr stands for the trace operator.
Thus, in the case of very high resolution, vector entropy coders performance can be approached by a lossless transformation followed by scalar entropy coders. In other words, vector entropy coders can be efficiently replaced by scalar entropy coders without (almost) any degradation to the overall rate-distortion function of the Gaussian source with memory $\underline{x}^c$, which is fixed by the preliminary quantization stage. Similar result have been reported in the case of memoryless sources in [5] (without excess bit rate analysis), and in [4].

4 Multiresolution Approach

4.1 Structure of the first stage

Consider now the coding scheme of figure (3), represented for $N = 2$. A uniform quantizer Q_1 is introduced in the ADPCM-like prediction loops, whose effect is to reduce the entropy of the transform signals y^q_i. These signals represent low resolution versions of the transform signals y_i described in the previous section. The error signals e_k, $k = 1, ..., N$, are

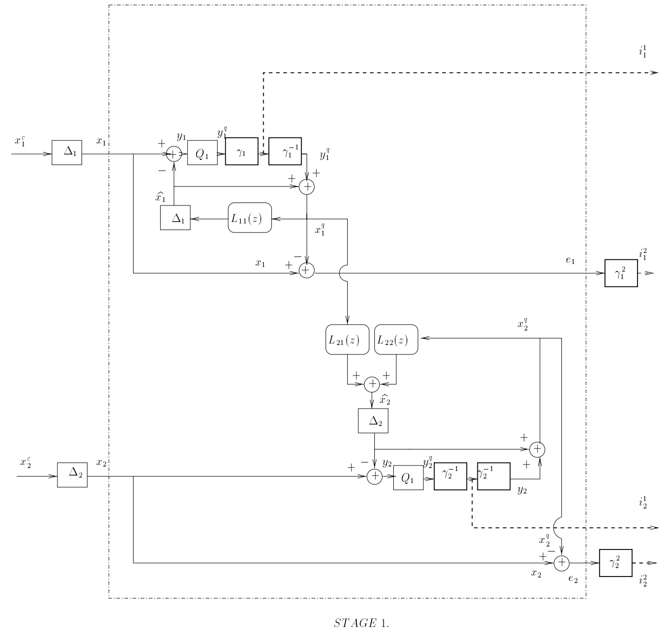

Figure 3: *First stage of the lossless multichannel prediction scalable encoder for $N = 2$. The bit rates for $\{i^1_{1:N}\}$ and $\{i^2_{1:N}\}$ are determined by the quantizer Q_1.*

then generated by substraction and separately entropy coded. Note that the transform signals are computed by substracting the optimal estimate of x_i based on the past *quantized* samples x^q_i, and by quantizing with stepsize Δ_i the resulting error prediction. The reason for computing the prediction by means of quantized data is that we are interested in a low resolution signal which can be computed *independently* of the error signals. Thus, only the available x^q_i at the decoder should be used to compute the remaining x^q_j, $j > i$. The total bit rate is thus the mean r_{LR} of the bit rates corresponding to the low resolution substreams $\{i^1_k\}$, $k = 1, ..., N$, plus the mean $\bar{r}$ of the rates corresponding to $\{i^2_k\}$, $k = 1, ..., N$ (substreams dedicated to the error signals). In order to simplify the derivations, we assume in this section that the Δ_i corresponding to the preliminary quantization stage are all equal, $\Delta_i = \Delta_{VHR}$. Moreover, we assume w.l.g. that the $\sigma^2_{x_i} >> 1$, and that $\Delta_{VHR} = 1$. Thus, x_i are integer valued, and $H(x_i) \approx h(x^c_i) - \log_2\Delta_i \approx h(x^c_i)$, and $S_{\underline{x}\underline{x}}(f) \approx S_{\underline{x}^c\underline{x}^c}(f)$. Note that quantizers after the predictors, with stepsizes $\Delta_i = \Delta_{VHR}$, are necessary to keep the structure lossless. Their effects on the several entropies are however small in comparison with the effects of the quantizer Q_1. (The stepsize Δ_{Q_1} is generally >> 1: for example, a Gaussian source with variance $\sigma^2_{y_i} = 10^4$ quantized with stepsize $\Delta_{VHR} = 1$ requires at least ≈ 8.7 bits to be entropy coded. Suppose we wish the bitrate corresponding to a low resolution version of this source to be $\approx 8.7/2$ bits, then the corresponding quantization stepsize should be $\Delta_{Q_1} \approx 20 >> \Delta_{VHR} = 1$.)

We shall now analyze the bitrate dedicated to the low resolution version $r_{LR} = \frac{1}{N}\sum_{i=1}^{N} H(y_i^q) = \frac{1}{N}\sum_{j=1} i_j^1$. Similarly to the previous section, each $y_{i,k}$ is now the optimal prediction of $x_{i,k}$ based on the past *quantized* value of x_i, and on all the *quantized* components of x_j, for all $j < i$. Assuming that the y_i are Gaussian, we have

$$\begin{aligned} r_{LR} &= \tfrac{1}{N}\sum_{i=1}^{N} H(y_i^q) \\ &\approx \tfrac{1}{2N}\log_2 (2\pi e)^N \prod_{i=1}^{N} \sigma^2_{y_i} - \tfrac{1}{N}\sum_{i=1}^{N} \log_2 \Delta_{Q_1}. \end{aligned} \tag{16}$$

We now use the following result from [2].

Result 2: Suppose that we apply $L(z)$ to decorrelate some real valued vectorial source $\{\underline{w}\}$ in closed loop around a quantizer of stepsize Δ, that is, by computing the predictions by means of quantized data of $\{\underline{w}\}$. The resulting vectorial process is $\{\underline{v}'\}$. Then the variances of the process $\{\underline{v}'\}$ can be related to the variances $\sigma^2_{v'_i}$ of $\{\underline{v}\}$, and to $S_{\underline{w}\underline{w}}(f)$ of Result 1 (10), by

$$\prod_{i=1}^{N} \sigma^2_{v'_i} \approx \prod_{i=1}^{N} \sigma^2_{v_i} \left(1 + \frac{\Delta_Q^2}{12}\left[\int_{-\frac{1}{2}}^{\frac{1}{2}} tr S^{-1}_{\underline{w}\underline{w}}(f) df - \sum_{i=1}^{N} \frac{1}{\sigma^2_{v_i}}\right]\right). \tag{17}$$

Applying Result 2 to (16) yields

$$\begin{aligned} & r_{LR} \approx \tfrac{1}{2}\log_2 2\pi e \left(\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln\det S_{\underline{x}^c\underline{x}^c}\right)^{1/N} \\ & \left(1 + \tfrac{\Delta^2_{Q_1}}{24N\ln 2}\left[\int_{-1/2}^{1/2} tr S^{-1}_{\underline{x}^c\underline{x}^c}(f) df - \sum_{i=1}^{N} \tfrac{1}{\sigma^2_{y_i^0}}\right]\right) - \log_2 \Delta_{Q_1} \\ & \approx r_{scal,L(z)}(\underline{y}) \times \\ & \left(1 + \underbrace{\frac{\Delta^2_{Q_1}}{24N\ln 2}\left[\int_{-1/2}^{1/2} tr\ S^{-1}_{\underline{x}^c\underline{x}^c}(f) df - \sum_{i=1}^{N} \frac{1}{\sigma^2_{y_i^0}}\right]}_{Excess\ bit\ rate\ due\ to\ noise\ feedback}\right) \\ & \underbrace{-\log_2 \Delta_{Q_1}}_{Bit\ rate\ reduction\ due\ to\ Q_1}, \end{aligned} \tag{18}$$

One can show from expression (18) that minimizing this excess bit rate entails maximizing $\sum_i \frac{1}{\sigma^2_{y_i^0}}$, which in turns entails processing the signals in order of decreasing variance [8].

Now, the bitrate dedicated to the error signals, $\overline{r} = \frac{1}{N}\sum_{j=1}^{N} i_j^2$, corresponds to the entropies of the r.v.s e_i, which are uniformly distributed over the interval $[-\frac{\Delta_{Q_1}}{2}, \frac{\Delta_{Q_1}}{2}]$, and whose entropies are consequently

$$H(e_i) \approx \int_{-\frac{\Delta_{Q_1}}{2}}^{\frac{\Delta_{Q_1}}{2}} -\frac{1}{\Delta_{Q_1}} \log_2 \frac{1}{\Delta_{Q_1}} dx = \log_2 \Delta_{Q_1}. \tag{19}$$

Thus

$$\overline{r} = \frac{1}{N}\sum_{j=1}^{N} i_j^2 \approx \log_2 \Delta_{Q_1}. \tag{20}$$

4.2 Multiresolution approach and space filling loss

Since the DPCM-like loops used in the multiresolution approach exposed above use entropy constrained uniform scalar quantizers (ECUQs), one may wonder if a suboptimality does not arise. ECUQs are known to be suboptmal in the rate-distortion sense, since for a given distortion (irrespectively of the p.d.f of the source), the rate of an ECUQ is $\frac{1}{2}\log_2 \frac{\pi e}{6} \approx 0.25$ bits above the rate-distortion bound. So indeed (neglecting the term corresponding to quantization noise feedback in (18)), the rate r_{LR} will be ≈ 0.25 bits more than that of an ideal "Shannon quantizer" for a given distortion. In this ideal coder, the quantization error would be Gaussian instead of uniform, which yields the factor $\frac{1}{2}\log 2\frac{\pi e}{6}$. Concerning the error signal however, Figure (3) shows that we have to code uniformly distributed r.v.s. In the ideal case of a "Shannon quantizer", the rate dedicated to code the error signal would then be ≈ 0.25 bits more than $\overline{r}$. Thus, neglecting the quantization noise feedback in (18), the birates r_{LR} and $\overline{r}$ sums in any case up to $r_{scal,L(z)}(\underline{y})$. As a conclusion, the price paid for using this multiresolution approach is the excess term in (18), which causes a slight suboptimality coming from quantization noise feedback, but not from space filling loss due to scalar quantizers. Similar considerations can be found in the case of multiresolution lossless transform coding in [10].

5 Mutistage Integer-to-Integer Multichannel Prediction

Finally, the previous scheme may be generalized to M stages, see Figure (4). Suppose we dispose of partially- or uncompressed data $\{\underline{x}\}$. In a first step, the minimum bit rate required to losslessly code these data is given by compressing the data with the "one shot" lossless coder. The size in bits of the compressed signal is then $r_{scal,L(z)}(\underline{y})$ (13). Suppose now we wish to transmit the data $\underline{x}$ by means of $M+1$ substreams corresponding to different resolution levels with imposed rates R_i ($\sum_{i=1}^{M+1} R_i \approx r_{scal,L(z)}(\underline{y})$).

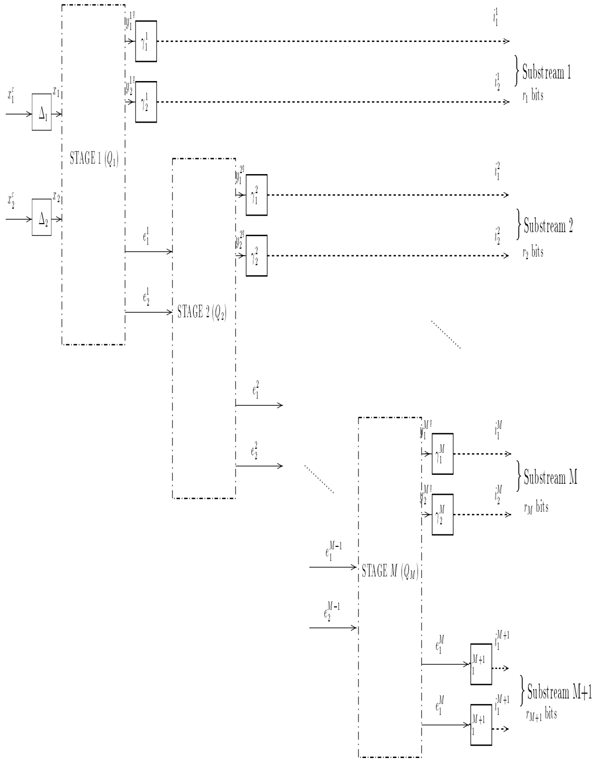

Figure 4: *Multistage structure of the lossless multichannel prediction scalable encoder for $N = 2$. The bit rates of the substreams are determined by the quantizers Q_i.*

How should we choose the stepsizes Δ_{Q_k} of the M uniform quantizers ?
As seen in the previous section, the one-stage structure of Figure (3) will, neglecting the term corresponding to the noise feedback in (18), yield a first substream with rate

$$r_1 = r_{LR} = \frac{1}{N}\sum_{k=1}^{N} i_k^1 \approx r_{scal,L(z)}(\underline{y}) - log_2 \Delta_{Q_1}, \tag{21}$$

and a complementary susbstream with rate $\overline{r} \approx \log_2 \Delta_{Q_1}$. If now we use a second stage, the previous error signal with rate $\overline{r}$ will be divided into two substreams with rates $r_2 \approx \log_2 \frac{\Delta_{Q_1}}{\Delta_{Q_2}}$, and $\overline{r_2} \approx \log_2 \Delta_{Q_2}$. (Note that for the stages $i > 2$, the prediction stucture becomes useless if the error signals are white (that is, if the decorrelation performed in the stage is efficient). Thus, a structure using M stages will yield a first substream with rate r_1 given by (21), $M-1$ complementary substreams with rates

$$r_j = \frac{1}{N}\sum_{k=1}^{N} i_k^j \approx \log_2 \frac{\Delta_{Q_{j-1}}}{\Delta_{Q_j}},\; j = 2, 3 \ldots M, \tag{22}$$

and a last substream with rate $r_{M+1} \approx \log_2 \Delta_{Q_M}$.
It can easily be checked that the constraint $r_1 \approx R_1$ imposes $\Delta_{Q_1} \approx 2^{r_{scal,L(z)}(\underline{y}) - R_1}$. Similarly, the constraints $r_k \approx R_k$ impose $\Delta_{Q_k} \approx \Delta_{Q_{k-1}} 2^{-R_k}$, for $k = 2, \ldots M$. Thus the stepsizes Δ_k of the M uniform quantizers should be determined by the simple rule of thumb

$$\Delta_{Q_k} = 2^{r_{scal,L(z)}(\underline{y}) - \sum_{i=1}^{k} R_i},\;\; k = 1, \ldots, M. \tag{23}$$

References

[1] S.-M. Phoong and Y.-P. Lin, "Prediction-based lower triangular transform," *IEEE trans. on Sig. Proc.*, vol. 48, July 2000.

[2] D. Mary and D. Slock, "Vectorial dpcm coding and application to wideband coding of speech," in *ICASSP 2001*, (Salt lake City), May 2001.

[3] M. Hans and R. Schafer, "Lossless Compression of Digital Audio," *IEEE Sig. Proc. Mag.*, vol. 18, Jul. 2001.

[4] D. Mary and D. T. M. Slock, "A performance analysis of integer-to-integer transforms for lossless coding of vectorial signals," in *Proc. SMMSP'02*, (Toulouse, France), September 2002.

[5] V. Goyal, "Transform coding with integer-to-integer transforms," *IEEE trans. on Inf. Theory*, vol. 46, March 2000.

[6] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley Series in Telecommunications, 1991.

[7] D. Mary and D. Slock, "A performance Analysis of Integer-to-Integer Transforms for Lossless Coding of Audio Signals," tech. rep., Institut Eurécom, March 2002. Available at http://www.eurecom.fr/~mary/publications.html.

[8] D. Mary and D. Slock, "Causal transform coding, generalized mimo prediction and application to vectorial dpcm coding of multichannel audio," in *WASPAA01*, (New York, USA), October 2001.

[9] T. Liebchen, "Lossless audio coding using adaptive multichannel prediction," in *113th AES Convention, Los Angeles*, October 2002.

[10] D. Mary and D. T. M. Slock, "On the suboptimality of orthogonal transforms for single- or multi-stage lossless transform coding," in *DCC*, 2003. Submitted Paper.