Perfect Reconstruction versus MMSE Filter Banks in Source Coding

Karine Gosse, *Member, IEEE,* and Pierre Duhamel, *Senior Member, IEEE*

***Abstract*— Classically, the filter banks (FB's) used in source coding schemes have been chosen to possess the perfect reconstruction (PR) property or to be maximally selective quadrature mirror filters (QMF's). This paper puts this choice back into question and solves the problem of minimizing the reconstruction distortion, which, in the most general case, is the sum of two terms: a first one due to the non-PR property of the FB and the other being due to signal quantization in the subbands. The resulting filter banks are called minimum mean square error (MMSE) filter banks.**

In this paper, several quantization noise models are considered. First, under the classical white noise assumption, the optimal positive bit rate allocation in any filter bank (possibly nonorthogonal) is expressed analytically, and an efficient optimization method of the MMSE filter banks is derived. Then, it is shown that while in a PR FB, the improvement brought by an accurate noise model over the classical white noise one is noticeable, it is not the case for MMSE FB. The optimization of the synthesis filters is also performed for two measures of the bit rate: the classical one, which is defined for uniform scalar quantization, and the order-one entropy measure. Finally, the comparison of rate-distortion curves (where the distortion is minimized for a given bit rate budget) enables us to quantify the SNR improvement brought by MMSE solutions.

I. Introduction

CLASSICALLY, in transform coding schemes, the signal to be encoded is split into several decorrelated subband components prior to quantization. This encoding process (transform and quantization) performs lossy compression, and a classical problem is to choose the quantization steps so that the original signal is reconstructed with minimum distortion for a given bit rate budget. Usually, the signal reconstruction is done by means of the inverse transform applied to the quantized coefficients, and in the case of filter banks, it thus relies on their perfect reconstruction (PR) property. Since the reconstruction error is cancelled, this is the best possible choice in the absence of quantization in the subbands. However, the presence of quantization puts this optimality back into question.

In fact, taking the effects of noise into account in the optimization of the system has been known to be successful elsewhere. In the communication area, the analysis FB is equivalent to sending the input signal into a multichannel. The quantization noise is equivalent to the channel noise, and the synthesis PR FB thus corresponds to a zero-forcing equalizer (ZFE), which is the equalizer that exactly inverts the channel filters when no noise is present. Yet, in noisy channels, ZFE are largely overcome by minimum mean squared error (MMSE) equalizers. As another example, in the signal processing area, minimizing the distortion due to additive noise has led to the so-called Wiener filters [1]. Recently, such Wiener filters have even been introduced in each subband of an orthonormal wavelet filter bank (i.e., lossless PR FB) for the restoration of $1/f$ fractal signals distorted by a transmission channel and additive noise [2].

Focusing on M-band FB-based systems, this paper proposes solutions for obtaining the reconstruction FB's that minimize the mean squared (MS) distortion introduced by the quantizers. These results are compared with the PR case in a situation where the distortion is minimized for a given bit rate budget.

Note that previous works were related to such a problem. The closest one can be found in [3], in which an optimization of the synthesis window of an analysis/synthesis system using the weighted overlap-add synthesis method is performed. Such a scheme is a modulated filter bank (MFB). When quantization noise is present, the authors also use a statistical model for designing the optimal synthesis filter in the MSE sense. This approach is further generalized by a study of matrix Wiener filters for subband coders in [4], by a multirate Kalman filtering formalism in [5], and by an MMSE design of two-dimensonal (2-D) filters using a recursive pseudo-adaptive algorithm [6]. Our work can be seen as an extension of these studies since it addresses the tuning of both the synthesis filters and the subband quantization steps in the general case of filter banks, including orthogonal transforms and MFB's. The performances of various other analysis/synthesis systems (including the discrete Fourier transform (DFT), quadrature mirror filter (QMF), and pseudo-QMF filter banks) in the presence of quantization have also been measured in [7].

Expressing the MSE to be minimized requires the choice of a quantization noise model. Here, for the sake of simplicity, we consider uniform scalar quantization, and this results in a simple additive noise model. Uniform quantization is widely used in subband coding because of its simplicity, but other choices, such as pdf-optimized scalar quantization, are also compatible with our approach. Previous work somewhat related to ours was undertaken in this case. Westerink [8] uses

Manuscript received April 13, 1995; revised November 11, 1996. This work was supported by a grant of the GdR "Traitement du signal et des images" of the CNRS. The associate editor coordinating the review of this paper and approving it for publication was Dr. Truong Q. Nguyen.

K. Gosse is with the Centre de Recherche de Motorola, Paris, France (e-mail: gosse@crm.mot.com).

P. Duhamel is with Département Signal, ENST, Paris, France (e-mail: duhamel@sig.enst.fr).

Publisher Item Identifier S 1053-587X(97)06450-7.

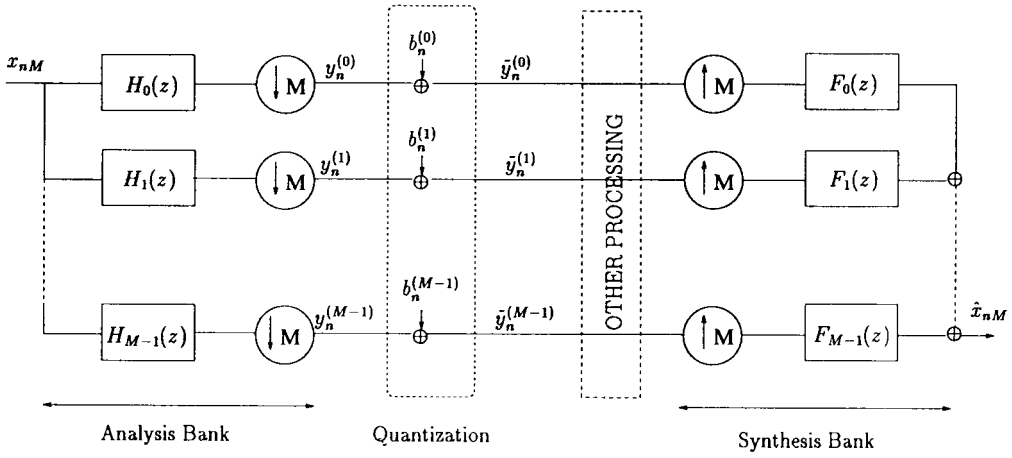

Fig. 1. Compression system including an FB and a quantization stage.

a gain-plus-additive noise model [9] of the quantizer in order to evaluate the amount of quantization error introduced by a QMF filter bank in an image coding context. This work was then extended by Uzun [10], Haddad [11], [12], and Kovačević [13], [14]. They use the decomposition of the mean squared distortion into a signal term and a random term emphasized by Westerink. The signal distortion part can be cancelled by the introduction of a compensation matrix in the synthesis part of the bank. Then, the remaining random term is minimized over the set of paraunitary or biorthogonal FB's. Note that this requirement sets heavy constraints on the filters. This approach amounts to restoring the PR property by considering the amount of noise correlated to the signal as part of the signal. In our approach, the PR property is broken in order to optimize the whole synthesis part for a given set of quantizers.

In practice, entropy coding is widely used in transform coding schemes to perform a final lossless compression in the subbands. When it is the case, uniform and optimized quantization give similar performances to the overall system. In theory, uniform quantization is even the optimal solution of the entropy-constraint optimization of the quantization steps [9]. Therefore, uniform scalar quantization is chosen here, and synthesis filter banks that minimize the MSE under entropy constraint (thus minimizing the ultimate performance of the coder) are also presented.

This paper intends to remain as general as possible; thus, the plain MSE criterion is chosen to optimize the filter banks. No treatment corresponding to perceptual characteristics is considered because it would depend on the desired application. However, we provide simulations with various input signals (synthetic AR processes, music samples) in order to illustrate that this approach is not linked to specific properties of the signal. Clearly, a practical use of these results would require, for example, the use of other distortion measures; it would also introducing perceptual criteria, as was done in a classical subband coding context by Vandendorpe [15]. He expresses the distortion as the weighted sum of the noise power in the subbands, where the weights are function of the eye sensitivity. Nevertheless, he restricted the optimization to the bit rate allocation for a given set of filters, whereas our results highlight the gain obtained by optimizing these filters. Other criteria to be used when designing filters in a subband coding context including quantization are proposed in [16].

Finally, we point out that our approach is compatible with an optimization of the analysis FB, such as what was done in [17] using orthonormal wavelet bases. The comparison between this technique and an on-line optimization of both the MMSE filter banks and the quantization steps is beyond the scope of this work, but we give some results that should be useful for on-line MMSE optimization purposes.

The paper is organized as follows. Section II states formally the problem at hand in the most general case. Then, Section III describes the bit rate constrained optimization of both PR and MMSE FB's under white and uniform quantization noise assumption. This optimization is further improved by the use of an accurate quantization noise model in Section IV. Section V addresses the entropy-constrained optimization of the filter bank. Finally, Section VI gives an estimation of the SNR improvement obtained by optimizing the synthesis filters over classical approaches on both synthetic and music signals.

II. Analytical Formulation of the Problem in the General Case

In a subband coding frame, as depicted on Fig. 1, the quantizers introduce some distortion in the subbands, whose amount depends on the bit rate allocation. A relevant problem for a system designer is to minimize the reconstruction error under bit rate constraint. Here, the optimization criterion is the MS distortion $\mathcal{E}[|\hat{x}_n - x_n|^2]$, where

- $\mathcal{E}$ mathematical expectation;
- x_n input signal;
- $\hat{x}_n$ reconstructed samples.

This paper intends mainly to compare the classical PR FB's having their quantization steps tuned to MMSE FB's in which *both* the quantization steps and the synthesis filters are tuned. Note that the analysis filter bank is unchanged in our procedure so that the comparisons remain reliable. The improvement in terms of SNR achievable with MMSE FB's is evaluated by comparing the rate-distortion curves of both kinds of schemes. Obtaining such curves in the PR FB case has been treated elsewhere, usually in the lossless case, which simplifies the

computations. In the MMSE case, this requires a solution of the following problems.

1) When the analysis filters are known and for a given choice of quantizers, what are the synthesis filters minimizing the distortion? The resulting filter bank depends clearly on the signal statistics. It does not keep its PR property, but it is optimal in the sense of minimum distortion, given the amount of quantization noise added in the subband.
2) Once Problem 1 above has allowed to obtain the *optimal* distortion as a function of the quantization steps, what set of quantizers minimizes this optimal distortion subject to a channel bit rate constraint?

This section aims at presenting a formal statement of the optimization problem by means of expressions of the MSE criterion and of the bit-budget constraint. Note that the matrix formulation of the input–output relation in a FB has been reported elsewhere (see [18], for example), but here, we use a polyphase description of the filtering process in presence of quantization. The underlying assumptions are that the quantization noise is additive and that the input signal is stationary.

A. The Criterion to Be Minimized

The general structure of the subband coding scheme is shown in Fig. 1. The M analysis and synthesis filters are, respectively, denoted by $H_k(z) = \sum_{j=0}^{L-1} h_k(j) z^{-j}$ and $F_k(z) = \sum_{j=0}^{L-1} f_k(j) z^{-j}$ for $0 \le k \le M-1$. They are assumed to have the same length L, and a reconstruction delay $T = L-1$ is introduced (it would be the delay of a lossless PR FB). In the following, $L = KM$ without loss of generality for the proposed solutions. Choosing other values for L would only lead to consider polyphase components of the filters with various lengths.

The synthesis operation is not time invariant but has period M. The vector of M consecutive output samples $\hat{X}_n^T = (\hat{x}_{nM} \hat{x}_{nM+1} \cdots \hat{x}_{nM+M-1})$ is the filtering of the quantized subband samples vector $\tilde{Y}_n^T = (\tilde{y}_n^{(0)}, \tilde{y}_n^{(1)} \cdots \tilde{y}_n^{(M-1)}, \tilde{y}_{n-1}^{(0)} \cdots \tilde{y}_{n-K+1}^{(M-1)})$ by a matrix F of synthesis coefficients (with these notations, $\tilde{y}_n^{(k)}$ is the nth quantized sample in subband k; all other subband signal vectors are then defined in the same way as $\tilde{Y}_n$). Therefore, F is built as in (1), shown at the bottom of the page, and $\tilde{Y}_n$ results from the quantization of $Y_n^T = (y_n^{(0)}, y_n^{(1)} \cdots y_n^{(M-1)}, y_{n-1}^{(0)} \cdots y_{n-K+1}^{(M-1)})$, where this quantization is modeled by an additive noise $B_n = Y_n - \tilde{Y}_n$. Thus, $\hat{X}_n = F\tilde{Y}_n = F(Y_n - B_n)$.

The vector of the subband signals Y_n is provided by the multiplication of the input signal vector $X_n^T = (x_{nM} x_{nM-1} \cdots x_{nM-2L+M+1})$ by H, which is the analysis filtering matrix. This matrix is block Toeplitz and contains K identical blocks $\mathcal{H}$ of analysis coefficients (of size $M \times L$), each of them being located M columns to the right of the one above

$$\mathcal{H} = \begin{bmatrix} h_0(0) & h_0(1) & \cdots & h_0(L-1) \\ \vdots & & & \cdots \\ h_{M-1}(0) & h_{M-1}(1) & \cdots & h_{M-1}(L-1) \end{bmatrix}$$

and

$$H = \begin{pmatrix} \ddots & & & 0 \\ & \mathcal{H} & & \\ & & \mathcal{H} & \\ 0 & & & \ddots \end{pmatrix}. \tag{2}$$

The MSE is obtained by comparing the reconstructed signal $\hat{X}_n^T$ to the vector of corresponding input samples $X_n'^T = (x_{nM-L+1} x_{nM-L+2} \cdots x_{nM-L+M})$ and is given by

$$\begin{aligned} D &= \frac{1}{M}\mathcal{E}[|\hat{X}_n - X_n'|^2] = \frac{1}{M}\mathcal{E}[(F\tilde{Y}_n - X_n')^T (F\tilde{Y}_n - X_n')] \\ &= \frac{1}{M}\mathcal{E}\left[\tilde{Y}_n^T F^T F \tilde{Y}_n - 2\tilde{Y}_n^T F^T X_n'\right] + \sigma_x^2 \\ &= \frac{1}{M}\mathcal{E}\left[\mathrm{Tr}\left(F\tilde{Y}_n\tilde{Y}_n^T F^T - 2F\tilde{Y}_n X_n'^T\right)\right] + \sigma_x^2 \end{aligned} \tag{3}$$

where σ_x^2 is the input signal variance. By swapping the mathematical expectation and the Trace operator and by using $\tilde{Y}_n = HX_n - B_n$ as well as F_i^T, which is the ith line of F (and the ith polyphase components of the synthesis filters), we have

$$D = \overbrace{\sigma_x^2 + \frac{1}{M}\sum_{i=0}^{M-1}\left(F_i^T H R_{xx} H^T F_i - 2F_i^T H \mathcal{E}[X_n x_{nM-L+1+i}]\right)}^{D_f} + \underbrace{\frac{1}{M}\sum_{i=0}^{M-1}\left(F_i^T (R_{bb} - 2R_{by}) F_i + 2F_i^T \mathcal{E}[B_n x_{nM-L+1+i}]\right)}_{D_b} \tag{4}$$

where R_{bb} and R_{xx} are, respectively, the autocorrelation matrices of the quantization noise (vector B_n) and the input signal (vector X_n). As for R_{by}, it is the intercorrelation matrix of the noise and the subband signals. These correlation matrices contain cross-correlation terms between the noise and signals in various subbands.

The MSE is thus divided in two terms; the first error term D_f is due to the non-PR property of the FB. It vanishes if the synthesis filters, which are represented here by vectors F_i, form with the analysis a PR filter bank. The second term D_b

$$F = \begin{bmatrix} f_0(0) & \cdots & f_{M-1}(0) & f_0(M) & \cdots & f_{M-1}(M) & \cdots & f_{M-1}((K-1)M)) \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ f_0(M-1) & \cdots & f_{M-1}(M-1) & f_0(2M-1) & \cdots & f_{M-1}(2M-1) & \cdots & f_{M-1}(KM-1)) \end{bmatrix} \tag{1}$$

is due to additive quantization noise. At very high bit rates as well as in absence of quantization, $\tilde{Y}_n$ approaches Y_n, and D_b tends to zero. Therefore, at high bit rates, the optimized filter bank converges to a PRFB.

Equation (4) can also be written in terms of the quantized signals and their autocorrelation matrix $R_{\tilde{y}\tilde{y}}$ as

$$D = \sigma_x^2 + \frac{1}{M}\sum_{i=0}^{M-1}\left(F_i^T R_{\tilde{y}\tilde{y}} F_i - 2F_i^T \mathcal{E}[\tilde{Y}_n x_{nM-L+1+i}]\right) = \frac{1}{M}\sum_{i=0}^{M-1} D(i) \tag{5}$$

where $D(i)$ is the reconstruction error of the phase $x_{nM-L+1+i}$ of the input.

Transform-Based Schemes: Transform coding is a particular case of subband coding, and a square transform is a specific filter bank of length M. Hence, the polyphase components of the FB are constants, and F is a square synthesis matrix. Equation (4) still holds, and the MMSE optimization can be interpreted as follows: If the coding stage of a system involves the DCT, using the inverse DCT on the decoding side does not minimize the distortion in presence of quantization noise.

B. The Constraint

The optimization is undertaken under the constraint that the sum of the bit rates in each subband is equal to some given value R_T. Classically, given $d(k)$, the dynamic range in subband k, which is defined as $d(k) = \max_y(y^{(k)}) - \min_y(y^{(k)})$, the bit rate R_k and the quantization step q_k are related by

$$R_k = \log_2 \frac{d(k)}{q_k} \tag{6}$$

C. The MMSE Solutions

It is clear from (4) that computing the MMSE FB requires the choice of a quantization noise model, thus enabling the criterion to be written in terms of the variables to be tuned: the quantization steps.

Then, for a given set of quantizers, the MSE is a quadratic form in terms of the polyphase components of the optimal synthesis filters. They are thus obtained by setting the derivative of D with respect to F_i to zero: $\forall i\ 0 \le i \le M-1$ $\frac{\partial D}{\partial F_i} = \frac{\partial D(i)}{\partial F_i} = 0$, which amounts to solving a set of linear equations.

Once the optimal filters are obtained, the total distortion is minimized with respect to the quantization steps. This second optimization has no analytical solution. Depending on the quantization noise model chosen, algorithms dedicated to this problem are proposed below.

III. Bit-Rate Constrained Optimization under White Noise Assumption

The purpose of this section is to make explicit the optimization of both PR and MMSE FB's under the classical assumption of uniform and white additive input-independent quantization noise (i.e., high-resolution assumption). In the following, they are, respectively, denoted as PR-WN and MMSE-WN FB's. Optimization procedures providing the optimum bit rate allocation are developed for both systems. In the PR case, assuming, in addition, that the FB is lossless leads, classically, to a very simple form of the MSE, and we shall stick with it.

A. Optimization of a PR-WN Filter Bank

In a lossless PRFB, defining D_k, which is the distortion introduced by the quantization in subband k, and under the classical assumption of uniform quantization noise, we have [19]

$$D = \frac{1}{M}\sum_{k=0}^{M-1} D_k \tag{7}$$

$$D_k = \frac{q_k^2}{12} = \frac{d(k)^2}{12} 2^{-2R_k} = c_k 2^{-2R_k} \tag{8}$$

with $d(k)$ defined as in Section II-B.

The optimal rate in subband k is a well-known result (see, for example, [20]), which is found by Lagrangian techniques, using the functional

$$J(\{R_k\}_{0\le k<M}, \lambda) = D(\{R_k\}_{0\le k<M}) + \lambda \sum_{k=0}^{M-1} R_k \tag{9}$$

the solution of which is given by

$$R_k = \frac{R_T}{M} + \frac{1}{2}\log_2 \frac{c_k}{\prod_{k=0}^{M-1} c_k^{\frac{1}{M}}}. \tag{10}$$

However, (10) does not ensure that the bit rates will be nonnegative since no such constraint was considered.

Thus, we now establish the analytical expression of the *positive* optimal bit rates R_k. By expressing the bit rates as the square of some quantity $R_k = \rho_k^2$ and minimizing the new Lagrangian functional

$$J(\{\rho_k^2\}_{0\le k<M}, \lambda) = \sum_{k=0}^{M-1} c_k 2^{-2\rho_k^2} + \lambda \sum_{k=0}^{M-1} \rho_k^2 = \sum_{k=0}^{M-1} J_k(\rho_k, \lambda) \tag{11}$$

with the constraint $R_T = \sum_{k=0}^{M-1} \rho_k^2$. The optimal bit rates R_k are then obtained classically in two steps. First, express the ρ_k minimizing J for a given λ. Let $J^*(\lambda) = \sum_{k=0}^{M-1} J_k^*(\lambda)$ denote the corresponding minimum. Second, maximize $W(\lambda) = J^*(\lambda) - \lambda R_T$ over λ.

1) First Step: $J^*(\lambda)$ is first found by canceling the derivative of J with respect to ρ_k, $0 \le k \le M-1$, leading to

$$2\lambda\rho_k - 4\ln(2) c_k \rho_k 2^{-2\rho_k^2} = 0, \quad \text{for } 0 \le k \le M-1. \tag{12}$$

Depending on the value of λ, the variables ρ_k minimizing J are given by the first of the two expressions given at the bottom of the page. In the second case ($\lambda \leq 2\ln(2)c_k$), note that for any value of c_k, $J^*_{k,1} > J^*_{k,2}$. Thus, $J^*_{k,1}$ is a maximum of J_k, and $J^*_{k,2}$ is our solution. It can be seen that this minimization amounts to distributing the bit rate budget among certain subbands only so that no bit rate is allocated to subbands corresponding to a small c_k. With the classical optimization method, a negative bit rate would be attributed to these very subbands.

2) Second Step: Given $J^*(\lambda)$, the maximum of $W(\lambda) = J^*(\lambda) - \lambda R_T$ occurs when its derivative with respect to λ vanishes, and it is unique because W is concave. The corresponding λ^* thus determines precisely the subbands in which some bit rate should be allocated. First, rearrange the constants c_k in decreasing order in vector $\underline{\gamma} = (\gamma_0, \ldots \gamma_{M-1})$. γ_k is thus a permutation of c_k. Note that γ_k is constant in our optimization problem, depending only from the input signal and the (given) analysis filters. If $\lambda < 2\ln(2)c_k$, the derivative of $J^*_k(\lambda)$ reads

$$[J^*_k(\lambda)]' = \log_2 \sqrt{\frac{2\ln(2)c_k}{\lambda}} \tag{13}$$

leading to the following possible forms for $[W(\lambda)]'$:

- For $\lambda > 2\ln(2)\gamma_0$, $[W(\lambda)]' = -R_T < 0$.
- For $2\ln(2)\gamma_k < \lambda < 2\ln(2)\gamma_{k-1}$, $[W(\lambda)]' = \sum_{i=0}^{k-1} \log_2 \sqrt{\frac{2\ln(2)\gamma_i}{\lambda}} - R_T$.
- For $\lambda < 2\ln(2)\gamma_{M-1}$, $[W(\lambda)]' = \sum_{i=0}^{M-1} \log_2 \sqrt{\frac{2\ln(2)\gamma_i}{\lambda}} - R_T$.

Due to the concavity of W, the index K characterizing the interval $[\gamma_{K+1}; \gamma_K]$ where $[W(\lambda)]'$ vanishes is easily obtained by an evaluation of $[W(\lambda)]'$ at the points $\lambda = 2\ln(2)\gamma_k$, $0 \leq k < M$. Let N be the lowest subscript k so that $[W(2\ln(2)\gamma_k)]' > 0$. If $[W(2\ln(2)\gamma_{M-1})]' < 0$, N is set to M. Hence, $2\ln(2)\gamma_{N-1} > \lambda^* > 2\ln(2)\gamma_N$, and the whole bit rate budget is distributed among the N subbands corresponding to constants $\gamma_0, \gamma_1, \ldots, \gamma_{N-1}$. Using the constraint $R_T = \sum_{k=0}^{M-1} \rho_k^2$ and the expression of ρ_k^2 enables us to find λ^* and the analytical expression of the optimal bit rates in (14), shown at the bottom of the page. This equation implies that the distortions are equal in all subbands with nonzero bit rate.

3) Practical Use of the Algorithm:

- First, rearrange the c_k into vector $\underline{\gamma}$. In the case of audio or image signals that are lowpass, the variance σ_k^2 of the signal in subband k decreases as k increases. Since c_k is closely related to σ_k^2, this will often lead in these cases to $\gamma_k = c_k$ $\forall k, 0 \leq k < M$.
- Second, while $1 \leq k \leq M-1$, compare $T = \frac{1}{2}\log_2 \frac{\prod_{i=0}^{k-1}\gamma_i}{\gamma_k}$ and R_T. Stop as soon as $T > R_T$ for a value of k denoted as N.
- Finally, apply (14) in order to get the optimal bit rates.

In comparison to the classical iterative "greedy bit allocation algorithm" [20], our procedure has the advantage of giving the *optimal* nonnegative bit rate allocation in the subbands instead of being a heuristic. MMSE FB optimizations got trapped in local minima with the iterative method. However, our algorithm does not provide an integer bit allocation. In terms of complexity and speed, finding the subband with maximum demand in the "greedy" algorithm may be computationally expensive for a large number M of subands (it requires M comparisons each time one bit is allocated).

B. Optimization of an MMSE-WN FB

Concerning the MMSE-WN FB, the whiteness of the quantization error yields simplifications of (4). In fact, under this assumption, R_{bb} is diagonal with nonzero terms given by (8), and all crosscorrelation terms either between noise and subband signals or noise and input signals vanish. $D(i)$ thus reads

$$D(i) = F_i^T[R_{yy} + R_{bb}]F_i - 2F_i^T H\mathcal{E}[X_n x_{nM-L+i+1}] + \sigma_x^2 \tag{15}$$

Hence, the ith polyphase component of the synthesis filters minimizing the distortion (15) in the white noise model are

$$F_i = [R_{yy} + R_{bb}]^{-1} H\mathcal{E}[X_n x_{nM-L+i+1}] \tag{16}$$

and the expression of the MSE on phase i of the reconstructed signal reads

$$D(i) = \sigma_x^2 - \mathcal{E}\left[X_n^T x_{nM-L+i+1}\right] H^T [R_{yy} + R_{bb}]^{-1} \times H\mathcal{E}[X_n x_{nM-L+i+1}]. \tag{17}$$

Numerical problems might occur while inverting matrix $[R_{yy} + R_{bb}]$ in case of highly lowpass signals (this never happened in our simulations). However, in this case, no bit rate would be allocated to highpass subbands, and the remaining synthesis coefficients could be obtained by extracting the corresponding submatrix of $[R_{yy} + R_{bb}]$.

In an MMSE FB, the output distortion cannot be directly written as a sum of independent subband contributions as in the lossless PR case. This makes the minimization of (17) over the quantizers rather intricate. Generic nonlinear optimization

$$\begin{cases} \text{if } \lambda \geq 2\ln(2)c_k \quad \rho_k = 0 \quad \text{and} \quad J^*_k(\lambda) = c_k \\ \text{otherwise} \end{cases} \qquad \begin{cases} \rho_k = 0 & \text{and} \quad J^*_{k,1}(\lambda) = c_k \\ \text{or} \quad \rho_k^2 = \frac{1}{2}\log_2 \frac{2\ln(2)c_k}{\lambda} & \text{and} \quad J^*_{k,2}(\lambda) = \frac{\lambda}{2\ln(2)} + \frac{\lambda}{2}\log_2 \frac{2\ln(2)c_k}{\lambda} \end{cases}$$

$$\begin{cases} R_\kappa = \frac{R_T}{N} + \frac{1}{2}\log_2 \frac{\gamma_k}{\prod_{j=0}^{N-1}\gamma_j^{\frac{1}{N}}} & \text{for } 0 \leq k < N \text{ and } \kappa \text{ verifying } \gamma_k = c_\kappa \\ R_\kappa = 0 & \text{otherwise} \end{cases} \tag{14}$$

methods should be implemented. A straightforward choice for such an optimization procedure, using, for example, the Matlab Optimization Toolbox, revealed difficulties of convergence. This is why we propose here to proceed as a sequence of independent optimizations: For a given set of quantizers, first find the optimum filters minimizing the distortion. The solution to this problem is given by (16). Then, given the set of synthesis and analysis filters, find the quantizers minimizing the distortion (15) for a given bit rate budget. The obtained solution serves as an initialization to the first step, and the whole procedure is iterated.

It is clear that this procedure provides a sequence of decreasing distortions. If the underlying function is convex, the procedure would converge to the global optimum. Despite the fact that we could not prove such a property on the cost function, we never obtained misconvergence of this procedure.

We show below that the second step of this procedure simplifies to a problem of the same form as the one described for the PR-WN case: The distortion follows (8) with another set of constants a_k instead of c_k. In fact, while fixing the synthesis filters, some components of D in (15), such as σ_x^2, $2F_i^T H\mathcal{E}[X_n x_{nM-L+i+1}]$ or $F_i^T R_{yy} F_i$, become constants. As a consequence, the bit rates minimizing D also minimize the reduced criterion:

$$\tilde{D} = \sum_{i=0}^{M-1} F_i^T R_{bb} F_i = \sum_{k=0}^{M-1} D_k \sum_{n=0}^{L-1} f_k(n)^2 = \sum_{k=0}^{M-1} \left(c_k \sum_{n=0}^{L-1} f_k(n)^2 \right) 2^{-2R_k} \tag{18}$$

where D_k is the quantized signal variance in subband k of (8). It turns out that $\tilde{D}$ is of the form $\tilde{D} = \sum_{k=0}^{M-1} a_k 2^{-2R_k}$, exactly like the distortion in a PR-WN scheme. Therefore, for a given set of synthesis filters, the optimal bit rate allocation in an MMSE-WN FB verifies (14).

Cascading both steps of the algorithm results in an efficient optimization method of the MMSE-WN filter banks under bit rate constraint, and the corresponding algorithm is summed up as follows:

1) *Initialization:* The procedure is initialized by a PR-WN system: The optimized synthesis filters, which are denoted as $F_{k,\mathrm{opt}}(z)$, $0 \leq k < M$, are set to PR-WN FB's, the subband bit rates in the MMSE FB are set to the optimal bit rate vector of the PR-WN system, and the variable MSE_1, which is the current predicted distortion in the MMSE system, is thus the corresponding minimum distortion.
2) Set $MSE_2 = MSE_1$. According to (16), compute the optimum synthesis filters $F_{k,\mathrm{opt}}(z)$, $0 \leq k < M$, minimizing the distortion for the given bit rate allocation.
3) Find the allocation of the bit rate budget in the subbands for minimum distortion, given $F_{k,\mathrm{opt}}(z)$, using the Lagrangian method detailed in the PR-WN case and (18).
4) MSE_1 is the corresponding MSE at the output of an MMSE-WN system using the synthesis filters $F_{k,\mathrm{opt}}(z)$ computed in Step 2 and the bit rate allocation computed in Step 3.
5) Go back to Step 2 unless $MSE_2 - MSE_1 < \epsilon$, where ϵ depends on the desired accuracy.

On-Line Optimization, Nonorthogonality, and Additivity: The previous section emphasizes the following property: The subband distortions additivity is a well-known property of orthogonal filter banks. In MMSE FB's, which are not orthogonal, subband distortions are no longer additive. Nevertheless, for a given set of synthesis filters, the overall distortion as a function of the subband bit rates differs from a sum of subband contributions by a constant. Therefore, minimizing the additive criterion (18) leads to the same solutions as would be obtained by minimizing the plain criterion *as long as* the analysis filters remain fixed. Hence, the above approach is valid for any bit rate estimation (entropy, Huffman, scalar or vector quantization, uniform or optimized quantizer) using techniques described, e.g., in [17]. This would require an estimation of individual rate-distortion curves for each subband for each synthesis filter bank in the iterations. However, note that minimizing the additive criterion (18) leads to the optimal bit rates but that its value does not give a correct estimation of the distortion.

Such a procedure also emphasizes that if an MMSE FB has been tuned beforehand on a large set of signals, the tuning of the sole quantization steps remains a simple procedure (as simple as in the classical PR case). This suggests that on-line optimization could be performed by updating the filters very infrequently, whereas the quantization steps could be varied more often. We believe that such procedures could be very close to optimum.

IV. Bit-Rate Constrained Optimization Under Colored Noise Assumption

For medium compression rates, and especially for lowpass signals, quantizers minimizing the MSE allocate a small, nonzero bit rate to high subbands in which the signal power is also very small. Thus, the white and uniform noise model does not accurately fit the quantization error since the correlation between subband signal and quantization noise is not negligible. Moreover, with this model, the signal variance is sometimes smaller than the estimated noise variance, which is clearly impossible.

A. Accurate Model of the Quantization Noise

It is therefore expected that more accurate noise models could produce synthesis filters that more carefully match the actual quantization noise. This section recalls an accurate model of the quantization error based on the results presented in [21]. It details the correlation terms needed to compute (4):

i) $\mathcal{E}[b_n^{(k)} b_m^{(k)}]$ $\forall n, m$, $0 \leq k < M$;
ii) $\mathcal{E}[b_n^{(k)} b_m^{(l)}]$ $k \neq l$ and $\forall n, m$;
iii) $\mathcal{E}[b_n^{(k)} y_m^{(k)}]$ $\forall n, m$, $0 \leq k < M$;
iv) $\mathcal{E}[b_n^{(k)} y_m^{(l)}]$ $k \neq l$ and $\forall n, m$ (v) $\mathcal{E}[b_n^{(k)} x_{nM-L+i+1}]$ $\forall n$, $0 \leq k, i < M$.

These correlations are estimated [21] by making use of the probability density and of the joint probability density of the quantization noise, both of them being expressed in terms of Φ_x, which is the characteristic function of the input signal. For example, the probability density of the quantization error is expressed as

$$f_b(B) = \begin{cases} \frac{1}{q} + \frac{1}{q}\sum_{k\neq 0} \Phi_x\left(\frac{2\pi k}{q}\right) \exp\left(\frac{-j2\pi kB}{q}\right) & \text{if } -\frac{q}{2} \leq B < \frac{q}{2} \\ 0 & \text{otherwise.} \end{cases}$$

As an example of computation, according to [21], the correlation between noise $b_n^{(k)}$ and subband signal $y_n^{(k)}$ [see iii)] can be written as

$$\mathcal{E}\left[y_n^{(k)} b_n^{(k)}\right] = \frac{q_k}{2\pi} \sum_{r\neq 0} \frac{(-1)^r}{r} \Phi'_x\left(\frac{2\pi r}{q_k}\right) \tag{19}$$

where Φ'_x is the derivative of Φ_x. Similar expressions provide the second-order statistics required for computing the MSE at the output of an MMSE FB.

Since the characteristic function of the input signal is difficult to estimate for real signals, we decided to approximate these various correlation matrices by the ones that would be obtained with a Gaussian input signal having the same correlation matrix R_{xx} as the real signal. Simulations show that this approximation provides very accurate results, even on real (music) signals (see Section VI).

Under this assumption, the expressions providing estimates of the various correlation matrices are summarized below. First, our noise model disregards the correlation between signals and noises belonging to different subbands [see ii) and iv)] because they are second-order terms if the frequency bands of the filter bank overlap only reasonably. Otherwise, simulations confirm that this approximation holds for FB's with overlapping between adjacent subbands (such as the ELT [22]) for subband bit rates above 0.18 bits/sample/subband. Therefore, this colored noise model improves significantly the white noise one in most cases.

Hence, the corresponding coefficients of R_{bb} and R_{yb} are set to zero. The remaining terms are obtained as

$$\mathcal{E}\left[b_n^{(k)^2}\right] = \frac{q_k^2}{12}\left[1 + \frac{12}{\pi^2}\sum_{r=1}^{\infty} \frac{(-1)^r}{r^2} \exp\left(-\frac{2\pi^2 r^2 \sigma^2_{y^{(k)}}}{q_k^2}\right)\right] \tag{20}$$

$$\mathcal{E}\left[b_n^{(k)} b_m^{(k)}\right] = \frac{q_k^2}{\pi^2}\sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \frac{(-1)^{r+s}}{rs} \times \exp\left(-\frac{2\pi^2}{q_k^2}\sigma^2_{y^{(k)}}(r^2+s^2)\right) \times \sinh\left(\frac{4\pi^2 rs\mathcal{E}\left[y_n^{(k)} y_m^{(k)}\right]}{q_k^2}\right) \tag{21}$$

$$\mathcal{E}\left[y_n^{(k)} b_n^{(k)}\right] = -2\sigma^2_{y^{(k)}} \sum_{r=1}^{\infty} (-1)^r \exp\left(-\frac{2\pi^2 r^2 \sigma^2_{y^{(k)}}}{q_k^2}\right) \tag{22}$$

$$\mathcal{E}\left[y_n^{(k)} b_m^{(k)}\right] = -2\mathcal{E}\left[y_n^{(k)} y_m^{(k)}\right] \sum_{r=1}^{\infty} (-1)^r \exp\left(-\frac{2\pi^2 r^2 \sigma^2_{y^{(k)}}}{q_k^2}\right). \tag{23}$$

Equation (23) is obtained by techniques similar to those in [21] by using the joint characteristic function of $y^{(k)}$. Finally, the last correlation term $\mathcal{E}[B_n x_{nM-L+i+1}]$ with $0 \leq i < M$ can be rewritten as

$$\mathcal{E}[B_n x_{nM-L+i+1}] = \mathcal{E}[\hat{x}_{nM+i} B_n] = \mathcal{E}[B_n Y_n G_i] = R_{by} G_i \tag{24}$$

if G_i denotes the vector of the ith polyphase coefficients of the synthesis providing PR when associated with the analysis of the considered MMSE filter bank.

B. Optimization of PR and MMSE Schemes with Accurate Noise Model

Computing optimum quantizers and filters with the colored noise model leads to a comparison of two other compression schemes: a PR one (which is referred to as PR-CN FB) and a MMSE one (which is known as MMSE-CN FB). They are compared in Section VI with PR-WN and MMSE-WN systems since only simulations can indicate the pertinence of either one model or the other.

1) MSE Expression for a PR-CN FB: The purpose of introducing an error model valid at low bit rates in a PR FB is the handling of a better prediction of the distortion level during the optimization in order to find the quantizers that will be optimum in practice. The MSE in a PR-CN filter bank is still given by (7) (subband distortions additivity), and only D_k is now given by (20).

2) MSE Expression for an MMSE-CN Filter Bank: The MSE at the output of MMSE-CN FB is given by (4), when the correlation matrices are estimated using the colored noise model. The optimal synthesis filters are obtained analytically by $\forall i\ 0 \leq i \leq M-1\ \frac{\partial D(i)}{\partial F_i} = 0$:

$$F_i = R^{-1}_{\tilde{y}\tilde{y}}(H\mathcal{E}[X_n x_{nM-L+i+1}] - \mathcal{E}[B_n x_{nM-L+i+1}]) \tag{25}$$

where $\tilde{y}$ denotes the quantized subband signals.

3) General Optimization Method of the PR-CN and MMSE-CN Cases: For both schemes using the accurate noise model, the expression of the correlation terms is too complex for using the optimization methods established in the WN case. Finding the optimal bit rates requires us to set a general procedure based on standard algorithms performing nonlinear optimization.

Since unconstrained nonlinear optimization methods are much more reliable, the constraints have been included in the function to be minimized. First, we carry out the variable change: $R_k = \rho_k^2$ that forces R_k to be positive. Then, the linear constraint relating the subband bit rates becomes, in terms of ρ_k, $\sum_{k=0}^{M-1} \rho_k^2 = R_T$ and describes a hypersphere. Its parameterization can be done with

$$\begin{cases} \rho_0 = \cos\theta_0 \\ \rho_k = \sin\theta_0 \sin\theta_1 \cdots \sin\theta_{k-1} \cos\theta_k \\ \rho_{M-1} = \sin\theta_0 \cdots \sin\theta_{M-2} \end{cases}$$

By means of this correspondence, we have transformed our constrained minimization problem of a function of M variables into an unconstrained minimization of a function of $M-1$ parameters θ_k.

Among unconstrained optimization methods, gradient methods are generally more efficient than simple search methods when the function to be minimized is continuous in its first derivative. They require an analytical expression of the gradient, which can easily be obtained from (4). We have used the implementation of the quasi-Newton algorithm found in the Matlab Optimization Toolbox. Unfortunately, we cannot ensure that the general algorithm reaches the global minimum, especially since the optimized functional is highly nonlinear with respect to the θ_k. Only the comparison with the white noise case and the shape of the obtained rate/distortion curve (regular or not) give some confidence that we are close to global convergence. Comments on the relative efficiencies of the various methods are provided in Section VI.

V. Entropy-Constrained Optimization of the Filter Bank

A main concern with the approach described above is the precise definition used to estimate the bit rate for a given quantization step: Equation (6) does not take into account any entropy coding of the subband signals. The improvement brought by this (simple) procedure is noticeable, as is shown in Section VI, but the use of a more realistic bit rate evaluation is certainly more convincing. A procedure for making use of an actual coding such as Huffman has already been outlined in Section III-B. This section provides further analytical results for the optimization of the filters and the quantization steps in an MSE sense under entropy constraint. It thus aims at giving an upper bound to the SNR(dB) reachable by a filter bank (with optimized or PR filters) followed by uniform quantization and entropy coding. First, following the various steps of Section II, we express the criterion to be minimized and the constraint as a function of order-one entropies in the subbands; then, an optimization solution dedicated to the problem is chosen among the previously proposed ones.

The order-one entropy H_k of the signal in subband k, which is defined below, is used as a bit rate measure

$$H_k = -\sum_{j\geq 1} p_j^{(k)} \log_2 p_j^{(k)} \tag{26}$$

where $p_j^{(k)}$ denotes the occurrence probability of the jth quantizer output in subband k. The constraint is $\sum_{k=0}^{M-1} H_k = H_T$.

If the high-resolution assumption is met, i.e., if the probability density of the quantization noise is supposed to be uniform over a quantization step, the entropies are easily introduced as parameters of the MSE criterion of relation (4) since the noise variance σ_b^2 is related to the order-one entropy H of the quantizer input X [9] by

$$\sigma_b^2 = \frac{1}{12} 2^{2h(X)} 2^{-2H} \tag{27}$$

where $h(X)$ is the differential entropy of X, depending on $p_X(x)$, which is the probability density of X.

The situation would be much more complex in the colored noise case; hence, only two systems will be optimized under entropy constraint (EC), the PR-WN, and the MMSE-WN schemes, which are denoted as EC PR-WN FB and EC MMSE-WN FB. Since the simulations shown in Section VI indicate that the performances of the MMSE-WN and MMSE-CN systems are equivalent, we are confident that we are close to the optimal situation. However, this is not the only required assumption: Equation (27) shows that further assumption on the probability density of the subband signals is required. This probability is assumed to be Gaussian to remain consistent with Section IV. Moreover, simulations show that entropy evaluations made on actual subband signals of real signals are very close to the estimate obtained with the above hypothesis. Equation (27) applied to a Gaussian signal $y^{(k)}$ in subband k gives

$$D_k = \sigma_{b^{(k)}}^2 = \frac{\pi e}{6} \sigma_{y^{(k)}}^2 2^{-2H_k}. \tag{28}$$

A similar expression describes the distortion for subband signals having a Laplacian probability density [9]; $\frac{\pi e}{6}$ should only be changed into $\frac{e^2}{6}$.

At this point, the choice of optimization methods is straightforward since both expressions of the MSE as a function of the subband entropies (28) or as a function of the subband bit rates (8) are similar. All optimization procedures given in Section III-A are valid after substituting the constant $\frac{\pi e}{6}\sigma_{y^{(k)}}^2$ for $\frac{d(k)^2}{12}$ and the entropy H_k for the bit rate R_k.

A full procedure also requires an expression relating the quantization steps to the subband entropy. It is derived from the probability density of the signals. Given the quantities e_i defined by

$$e_i = \operatorname{erf}\left(\frac{(2i+1)q_k}{2\sqrt{2}\sigma_{y^{(k)}}}\right)$$

with

$$\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_0^x \exp(-t^2)\, dt \tag{29}$$

the quantized signal entropy in the kth subband is provided by

$$H\left(\frac{q_k}{\sigma_{y^{(k)}}}\right) = -e_0 \log_2(e_0) - \sum_{i=1}^{\infty} (e_i - e_{i-1}) \log_2\left(\frac{e_i - e_{i-1}}{2}\right). \tag{30}$$

The optimal quantizers corresponding to a given subband entropy are then computed with MATLAB by evaluating the $\frac{q_k}{\sigma_{y^{(k)}}}$ rate for a given value of H (interpolation of the reciprocal function of $H(\frac{q_k}{\sigma_{y^{(k)}}})$ or table lookup).

However, when simulating the encoding process with the analytically computed optimal subband quantizers, the measured order 1 entropies may differ slightly from the predicted ones: In fact, our method does not take into account either granular or overload noise and requires assumptions on subband signal statistics. More details on the entropy estimation error can be found in Section VI. Moreover, a general method for designing the quantizers under entropy constraint on given

training sequences but without any analytical calculation is given in [23].

VI. Simulations

To summarize the various schemes that are compared, recall that two different bit rate measures are used: the classical measure relating quantization step and bit rate [which is shown in (6)] and the order-one entropy of the signals. The first one leads to the optimization of four schemes, consisting only of a FB and a uniform quantization stage: the PR-WN, PR-CN, MMSE-WN, and MMSE-CN filter banks. The simulations corresponding to these four cases aim at quantifying the SNR improvement brought by MMSE FB (compared with PR FB), thus enabling the more relevant choice between the four schemes, depending on the desired compression ratio.

As for the second bit rate measure, it is aimed at estimating the asymptotic gains and checking that improvements with MMSE FB's are still there if entropy coding is performed. In this case, two kinds of systems including uniform quantization and entropy coding are considered: the EC PR-WN filter bank and the EC MMSE-WN filter bank.

A. The Simulation Context

1) The Signals Tested: These schemes were run on synthetic signals (order 1 AR processes, with correlation coefficient ranging from 0.1 to 0.9) as well as on an audio signal (the beginning of Vivaldi's *The Four Seasons: The Spring*, having CD quality, i.e., sampled at 44.1 kHz with 16 bits/sample).

2) The Filter Bank: The number of subbands M varies between 2 and 32 (Layer I and II of MPEG-1 involve a 32-band filter bank), and the total channel rate lies between 0.5 and 8 bits/sample, corresponding to a compression ratio ranging between 2 and 32 for CD quality signals. For comparison purposes, transparent audio compression of CD quality signals at 64 kbits/s would correspond to a compression ratio of 11, but the use of masking characteristics of the human hearing process is to be taken into account. Furthermore, a Huffmann-like encoder would certainly be applied on the subband signals, thus improving the compression ratio. A curve involving Huffman coding is provided in Fig. 10.

The PR FB chosen as a reference is an extended lapped transform (ELT) taken from [22] (lossless case) with filter length $L = 4M$. The analysis filters coefficients, for $0 \leq k < M$ and $0 \leq n < L$, are given by

$$h_k(n) = h(n)\sqrt{\frac{2}{M}}\cos\left[\frac{\pi}{4M}(2n - M + 1)(2k+1)\right]$$

with

$$h(n) = -\frac{1}{2\sqrt{2}} + \frac{1}{2}\cos\left[\frac{\pi}{2M}(2n+1)\right]. \tag{31}$$

The filters composing such a four-band filter bank are shown on Fig. 2. However, the frequency selectivity of the analysis filters can be improved by increasing the overlapping factor, and we also consider 16 modulated filters of length 256 designed according to [24] in order to show the influence of the reference analysis filter selectivity on the resulting MMSE performances (the corresponding prototype has stopband attenuation of 65 dB).

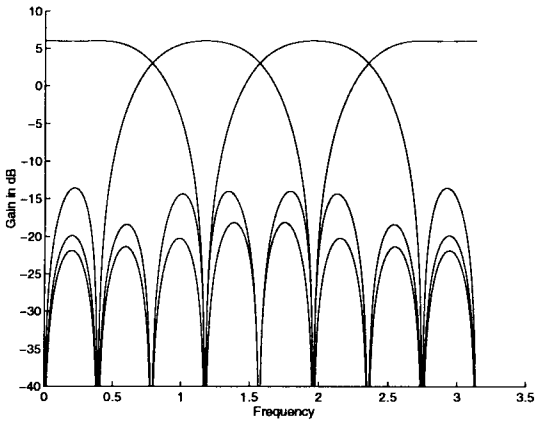

Fig. 2. Magnitude frequency response of Malvar's ELT with $M = 4$ and $L = 16$.

3) Autocorrelation Matrix of Input Samples: The MMSE FB was computed from an estimate of R_{xx} over a large number of samples $(N = 2^{17})$. Depending on the simulations, its performance in terms of distortion was estimated either on these N samples or on a much longer signal; this is specified in the text. In order to estimate possible effects due to a poor approximation of R_{xx}, MMSE FB's were also computed while modelizing the Vivaldi signal by an AR(2) process.

B. Rate-Distortion Curves Obtained with Uniform Quantization

This section aims at providing the first analysis and conclusions on the improvement brought by MMSE filter banks over PR banks with the classical bit rate measure. The various schemes are optimized according to Sections III-A and IV-B. In the case of the general optmization method of Section IV-B, choosing, as a starting point, the optimal solution found in the white noise case proved to be useful. Then, the optimal quantizers are computed using (8) and the choice $d(k) = 12\sigma_{y^{(k)}}\forall k$ such as $0 \leq k < M$. It turns out that 99.99% of Vivaldi signal samples belong to this interval.

When the quantizers are found, the synthesis filters are set, and the SNR(dB) is estimated using $\sum_{i=0}^{N-1} x_i^2 / \sum_{i=0}^{N-1} |x_i - \hat{x}_i|^2$. The segmental SNR measure was not elected in our study simply because it is not the criterion minimized over the set of filters and quantizers. Nevertheless, it has been checked that segmental SNR curves have the same shape.

1) Comparison of Predicted and Observed Distortions: Figs. 4–6 provide a comparison of the predicted and observed distortions at the output of, respectively, a 16-band PR-WN, PR-CN, and MMSE-CN system. Globally, they fit astonishingly well, but this requires further explanation. First, in the PR-WN case, the predicted curve is (at most) 4 dB under the curve of observed distortion at low bit rates because of the poor estimation of the quantization noise variance in this interval. The plot corresponding to the PR-CN filter bank illustrates the estimation improvement allowed by the colored noise model

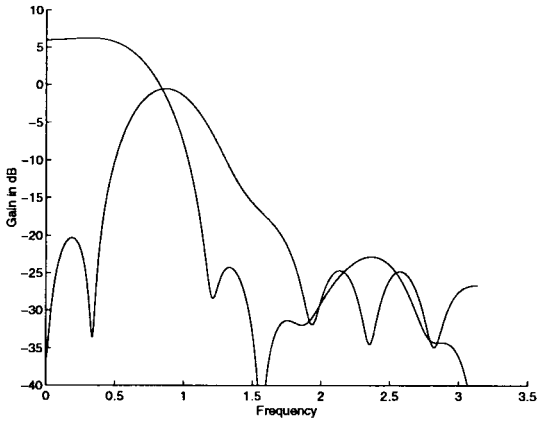

Fig. 3. Magnitude frequency response of optimized synthesis filters for $M = 4$, Vivaldi samples, and $\underline{R} = [6.3; 1.7; 0; 0]$.

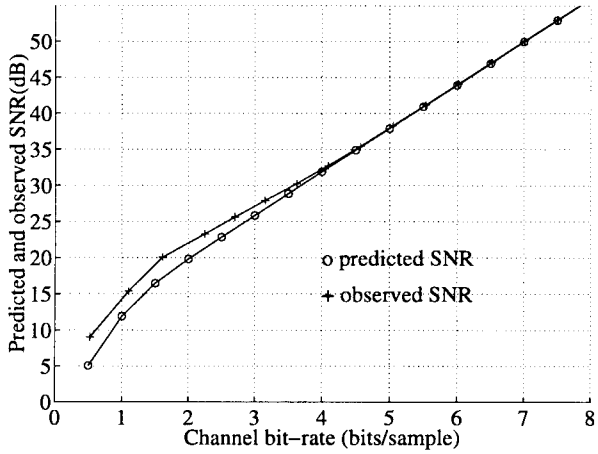

Fig. 4. Predicted and observed SNR versus channel bit rate for Vivaldi input samples (16 bits/sample), $M = 16$, and PR-WN synthesis.

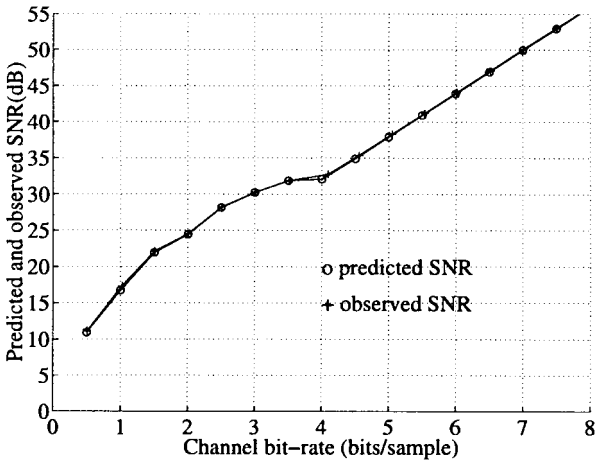

Fig. 5. Predicted and observed SNR versus channel bit rate for Vivaldi input samples (16 bits/sample), $M = 16$, and PR-CN synthesis.

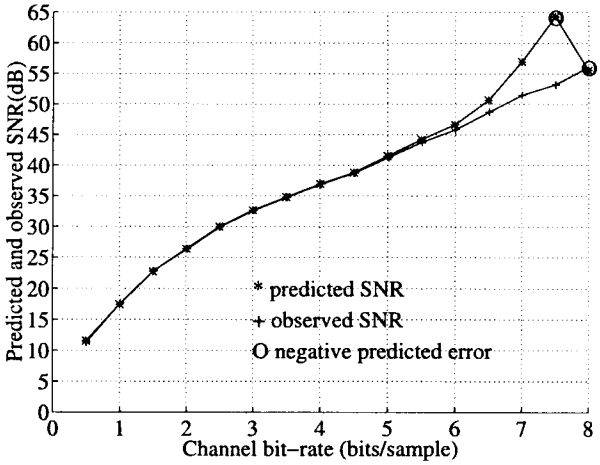

Fig. 6. Predicted and observed SNR versus channel bit rate for Vivaldi input samples (16 bits/sample), $M = 16$, and MMSE-CN synthesis.

introduced in Section IV. Both estimated and measured curves fit perfectly, and the approximation of the subband samples being Gaussian seems accurate.

Concerning MMSE FB's, only one out of both cases (WN and CN) is presented since they behave exactly alike. This time, the predicted curve fits the observed curve at low bit rates. In fact, an error on the MMSE filters coefficients results from the estimation error on the calculated quantization noise statistics. For rates over 5 bits/sample, the predicted output distortion is small compared with the obtained accuracy (due to the error on the filters) and even becomes negative (the circled last two points) over 7 bits/sample. Nevertheless, the noises being very small, this does not influence the synthesis FB optimization, whose result is close to a PR-WN filter bank in this bit rate range.

2) Comments About Rate-Distortion Curves Estimated Over 2^{17} Signal Samples: We give the rate-distortion curves for the signal Vivaldi compressed by a four-band, a 16-band, and a 32-band filter bank, respectively, in Figs. 7–9. Whatever the signals, the plots have these common characteristics:

1) The four curves merge above some bit rate (very high quality coding) for the following reasons: The error becomes too small to allow noticeable gains through the optimization, and furthermore, the MMSE solution tends to the PR solution as the bit rate increases [this was shown in (4)]. Moreover, the white noise model becomes accurate, explaining that the PR-WN and the PR-CN curves merge before the other ones.
2) The PR-WN scheme show poor performances at low rates due to an inaccurate estimation of the distortion level, resulting in an unappropriate choice of the quantization steps (emphasized in Fig. 4). Moreover, if the classical equation (10) giving the optimal R_k was valid at all bit rates, the PR-WN curve would be a line of slope 6 dB per bit. Since this equation gives rise to negative bit rates at low bit rates, (14) was substituted for it, but this breaks the linearity of the rate-distortion curve in the bit rate range where both solutions differ.
3) The MMSE-WN and MMSE-CN curves are almost identical whatever the signals and the number of subbands. This is explained as follows: In all schemes, when the bit-budget increases, some subbands in which the bit rate was null now obtain a certain bit rate amount. In MMSE schemes, this amount is never small

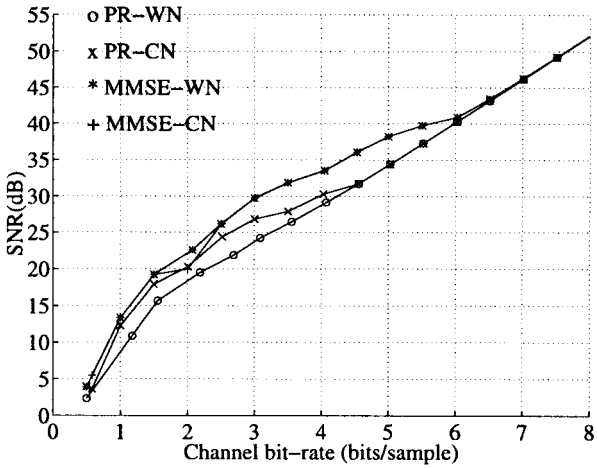

Fig. 7. Measured SNR versus channel bit rate for Vivaldi input samples (16 bits/sample) and $M = 4$.

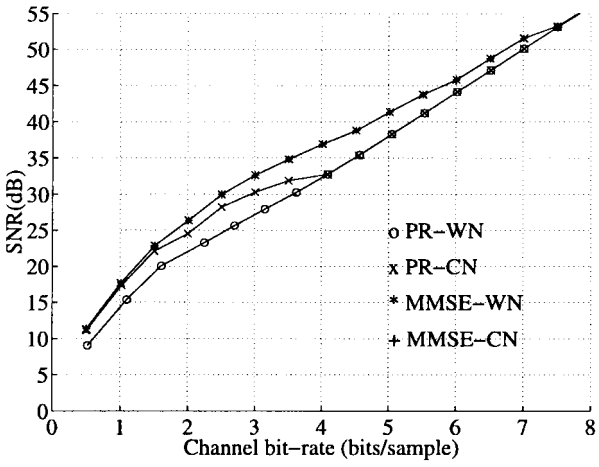

Fig. 8. Measured SNR versus channel bit rate for Vivaldi input samples (16 bits/sample) and $M = 16$.

(this is a secondary effect of the filters optimization). As a consequence, both white and elaborated noise models give approximately the same estimation of the noise variance. To illustrate this, consider the MMSE-WN scheme with four subbands applied to the Vivaldi signal. The bit rate allocation vector is $[8.6; 5.4; 0; 0]$ for a budget of $R_T = 14$ bits/block and becomes $[8.4; 5.1; 2.5; 0]$ for $R_T = 16$ bits/block. In the third subband, for a bit rate of 2.5 bits/block, the variance of the signal $\sigma^2_{y^{(3)}} = 0.0015$ for both models.

4) Below 2 bits/sample, MMSE curves and PR-CN curves are almost identical because only a few subband bit rates are not set to zero, and thus, only a few filters of the synthesis bank are optimized (the others ones are set to zero!).

Observation of the curves for AR process and music samples show that the improvement brought by MMSE FB compared with PR-WN FB can reach 2 to 5 dB in a large range of bit rates for all numbers of subbands. On the music signal presented in Fig. 7 with $M = 4$, the improvement is greater than 2.5 dB from 1 to 5.5 bits/sample; it reaches 5.5 dB for 3 bits/sample. For the 16- and the 32-band filter banks, the improvement obtained with MMSE FB's is slightly smaller (the maximum is, respectively, 4.6 and 4.3 dB at 3 bits/sample). The main difference is that the four curves merge later, respectively, about 7.5 bits/sample and over 8 bits/sample versus 6 bits/sample for the $M = 4$ case. This can be explained by a better allocation of the available bit rate in the subbands that really need it when the frequency band is split in smaller intervals and the budget is increased. On the other hand, in the four-band filter bank, the four solutions have the same performances at very low bit rate (i.e., 0.5 bit/sample) because the optimal quantizers found simply verify that "the whole bit-budget is given to the low-pass subband." Increasing the number of subbands yields an improvement of the SNR of approximately 4 dB from 4 to 16 subbands but is only 0.9 dB from 16 to 32 subbands. The relation between M, which is the number of subbands, and this improvement depends on the signal spectrum.

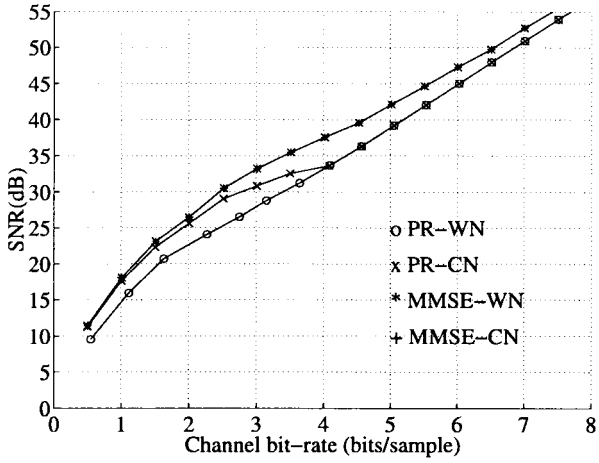

Fig. 9. Measured SNR versus channel bit rate for Vivaldi input samples (16 bits/sample) and $M = 32$.

An Huffman encoding stage after quantization has been simulated for $M = 32$, and the resulting rate-distortion curves are reported in Fig. 10. They give a more realistic estimation of the bit rate budget needed in a source coding system using MMSE filters. Their shapes are quite similar to the ones in Fig. 9 but with a shrunk bit rate scale. Therefore, the improvement brought by MMSE FB's in this case is kept but on a smaller bit rate range.

The tests on various AR processes show that the improvement brought by MMSE FB is all the more important as the input signal is correlated (see Fig. 11 for a Markov process $\rho = 0.9$). Astonishingly, the best results have been obtained on the real signals (Vivaldi music samples).

3) About the Optimized Synthesis Filters: Another point of interest is the shape of the optimized synthesis filters. The ones depicted in Fig. 3 are optimal in a four-band filter bank with a bit rate budget of $R_T = 8$. The optimal bit rates in the subbands are $[6.3; 1.7; 0; 0]$, and the optimal filters in subbands 3 and 4 are null. When compared with the ones of Fig. 2, the filter corresponding to subband 1 has smaller sidelobes, and filter 2 even presents a bandpass attenuation. The contribution

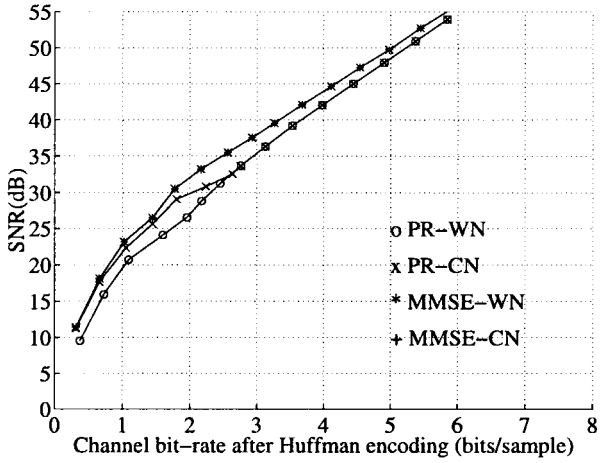

Fig. 10. Measured SNR versus channel bit rate for Vivaldi input samples (16 bits/sample), where $M = 32$ after Huffman encoding.

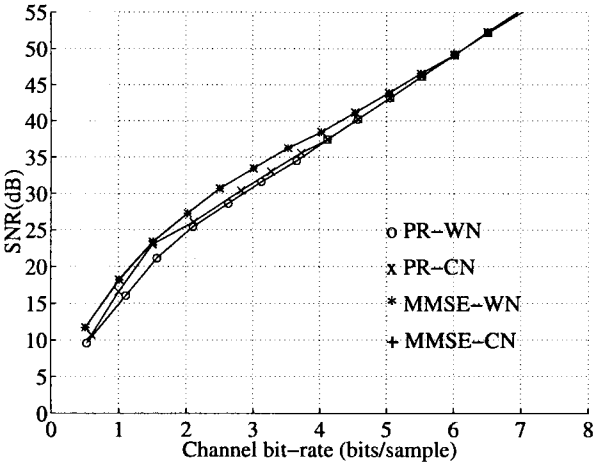

Fig. 12. Measured SNR versus channel bit rate for Vivaldi input samples (16 bits/sample) and $M = 16$ filters of length 256.

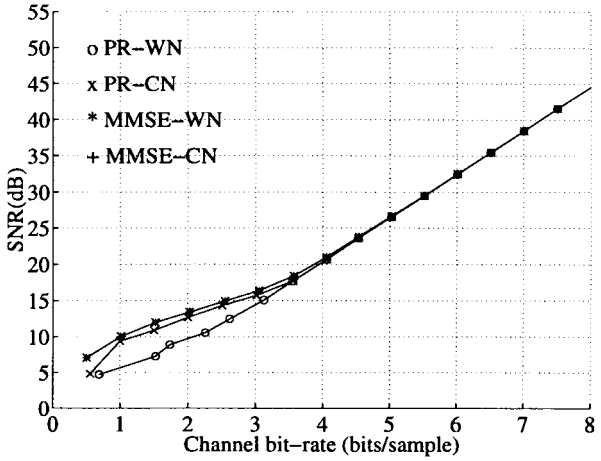

Fig. 11. Measured SNR versus channel bit rate for AR1 ($\rho = 0.9$) input samples, where $M = 16$.

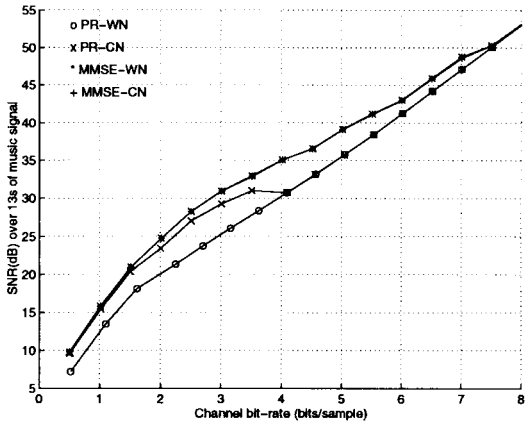

Fig. 13. Measured SNR versus channel bit rate for 6.610^4 Vivaldi input samples (16 bits/sample) $M = 16$.

of the second subband in the final reconstruction of the signals is thus reduced because the bit rate in this subband is small, and the corresponding noise is large.

A related question is about how much distortion comes from the quantization noise, compared with the reconstruction error in the MMSE-FB. In our examples, it was found that more than 40% the total distortion was due to the non-PR property of the synthesis FB in the bit rate range [0.5–4] bits/sample. This shows how the noise amplification can be lowered by relaxing the PR constraint since the total distortion remains smaller than the one only due to the noise in an optimized PR-FB.

4) Concerning the Influence of Analysis Filter Selectivity on MMSE Performances: Here, we consider a modulated filter bank of $M = 16$ subbands and filters of length $L = 256$ taken from [24]. The increased length improves the stopband attenuation of the filters, which is now about 65 dB. Corresponding performances are plotted in Fig. 12. Part of the gain brought by MMSE solutions is due to a improvement of the frequency selectivity in certain subbands (see Fig. 3). Thus, if the analysis bank has better stopband attenuation, the gain of MMSE solutions over PR systems is reduced but still present. At low bit rates, performances of PR-CN and MMSE schemes are the same.

5) Variants for the Estimation of R_{xx}: In a first set of simulations, the optimal synthesis filters are calculated using R_{xx} estimated over 2^{17} signal samples, and the SNR(dB) is obtained by filtering six times more samples (13 s of signal) with them. It is particularly promising that the improvement brought by MMSE FB's be kept under these new conditions. Indeed, it allows us to think of an optimization method involving on-line quantizer optimization thanks to (14) and a much less frequent optimization of the filters, which remain efficient because they correspond to an average situation. Fig. 13 supports these comments.

In a second step, R_{xx} is computed as the autocorrelation matrix of the AR(2) signal modelizing Vivaldi samples. Fig. 14 is obtained by processing 2^{17} Vivaldi samples by the resulting MMSE filters and quantizers. The improvement brought by approximated MMSE filters remains unchanged in the range [1–3] bits/sample and is still noticeable until 5 bits/sample, although R_{xx} estimation is simplified. Of course, a modelization using an AR signal of higher order would yield

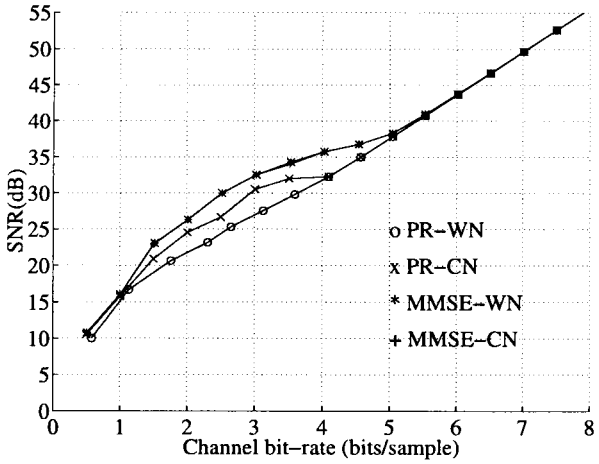

Fig. 14. Measured SNR versus channel bit rate for Vivaldi input samples. MMSE filters ($M = 16$) and quantizers are obtained with an AR(2) model of the input.

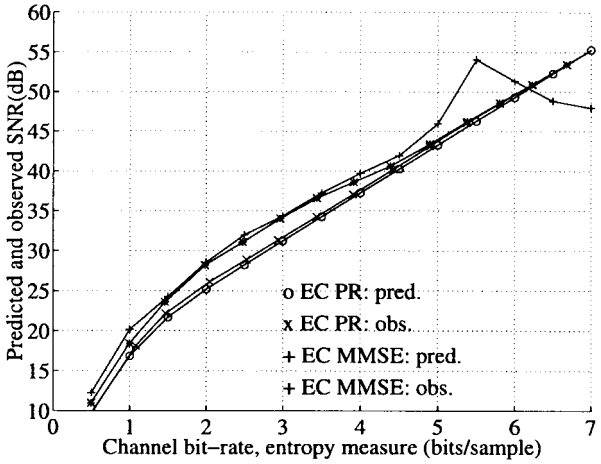

Fig. 15. Predicted and observed SNR versus entropy bit rate for Vivaldi input samples (16 bits/sample), $M = 4$, and with EC PR-WN or EC MMSE-WN synthesis.

better results. A practical use of such a technique would require letting the coefficients of the AR(2) signal vary with time.

Results of same quality can also be obtained while calculating optimal filters with the AR(1) R_{xx} matrix but with the optimal bit rate allocation with the real subband signal variances of "Vivaldi." MMSE filters are less sensible to poor R_{xx} estimation than MMSE quantizers. Indeed, optimal quantizers obtained with the AR(1) approximation associated with optimal filters for the "Vivaldi" signal yield poor performances. Therefore, a fruitful strategy should be to calculate MMSE filters over a certain number of "Vivaldi" samples modeled by an AR process and then to update the quantizers often by using actual information on subband signal variances.

C. Entropy-Distortion Curves Obtained with Uniform Quantization

Here, the SNR is estimated over the 2^{17} samples used for the estimation of R_{xx}. As in the previous subsection, the efficient optimization methods developed in Section III were used for both systems studied. In the EC PR-WN case, the optimal entropy allocation corresponds to the optimal bit rates of the PR-WN schemes for a same budget. Of course, this comes from the similarity of the distortion expressions since the sum of subband contributions in 2^{-2R_k} differ only by the multiplicative coefficients a_k. However, in the MMSE-WN case, both optimal allocations are not similar because the role played by the coefficients a_k is more complex: They appear in the autocorrelation matrix of the quantized subband signals, which has to be inverted for the estimation of the theoretical distortion.

1) Comparison of Predicted and Observed Distortions: In order to verify the accuracy of the assumptions needed to set the MSE expression as a function of the subband entropies, the predicted and observed entropy-distortion curves corresponding to both optimized schemes with $M = 4$ are plotted in Fig. 15. The Gaussian hypothesis concerning the subband signals seems to be fairly correct since the difference between prediction and observation is approximately of 0.5 dB. At low rates, however, the prediction error is closer to 1 dB because the signals in the lowpass subbands are further from Gaussianity than the other ones.

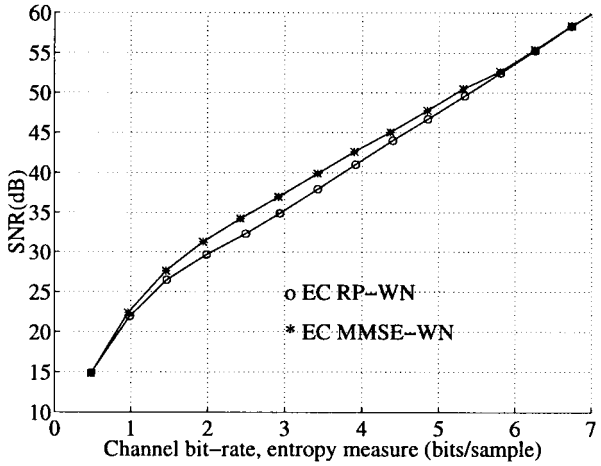

Fig. 16. Measured SNR versus entropy bit rate for Vivaldi input samples (16 bits/sample) $M = 16$.

2) Comments About the Observed Entropy-Distortion Curves: The improvement brought by the EC MMSE-WN filter bank, when the bit rate measured is the order-one entropy, can be observed in Figs. 15–17 for, respectively, a 4-, 16-, and 32-band filter bank. Yet it has been reduced to a maximum of 2.7 dB with $M = 4$, 2.1 dB with $M = 16$, and to 1.8 dB with $M = 32$. This is not really surprising since we consider an asymptotic bound of the performances.

The bit rate range in which the performances of the two systems differ still widens when M increases, starting from $[0.5-5]$, whereas $M = 4$ and reaching [1]–[7] while $M = 32$. Here, the curves merge at very low bit rates for $M = 16$ and $M = 32$, which was not the case with the classical bit rate definition.

3) First Conclusions: The results obtained with entropy-constrained systems confirm that SNR improvements can be obtained while using optimized filter banks instead of PR FB's in a source coding scheme including quantization and entropy coding. Moreover, we present three schemes allowing SNR

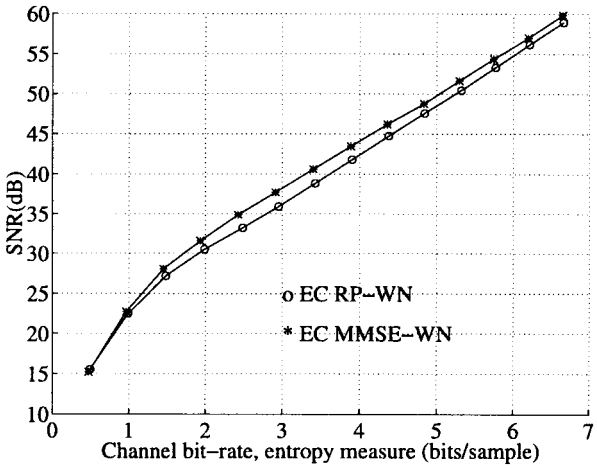

Fig. 17. Measured SNR versus entropy bit rate for Vivaldi input samples (16 bits/sample) $M = 32$.

improvements with respect to the classical PR solution under white noise assumption. The choice among them should rely on the following considerations. The PR-CN system allows us to keep the interesting structure of lossless filter banks, but its optimization is complex (general nonlinear procedure), and the resulting improvement is the smallest in the most narrow bit rate range. As for the MMSE-WN and the MMSE-CN schemes, they have exactly the same performances, but efficient optimization methods are established in the white noise case. It thus seems useless to introduce a colored noise model in the MMSE FB's.

VII. Conclusion

This paper emphasizes the usefulness of relaxing the perfect reconstruction property of the synthesis filter bank; the improvement that has been obtained with a SNR criterion is noticeable, at least in a specific compression rate range. Moreover, this improvement was observed for a source coding scheme, including a uniform scalar quantization stage, in two different situations: When the bit rate allocation is optimized, and when the order-one entropy allocation is optimized (i.e., entropy-constrained optimization of the filter bank). Thus, MMSE filter banks could increase the output SNR of any source coding scheme involving scalar quantization and entropy coding.

This paper also provides side results useful to designers of source coding schemes: It gives an analytical expression of the *positive* optimal bit rate allocation in any filter bank, under high-resolution assumption, whereas the classical optimal solutions in the lossless PR case can become negative. Based on this result, an efficient method opening the way to online optimization of the bit rates and fast optimization of the synthesis filters in a MMSE FB is given.

A further work under consideration is the optimization of the synthesis filters according to a specific source coding application, taking into account perceptual characteristics such as the variations of sensitivity of the ear at various frequencies. Linear masking effects, such as the inverse absolute hearing threshold, can be taken into account easily by using a frequency-weighted psychoacoustic criterion. Corresponding results are reported in [25]. Then, in a second additional step, nonlinear frequency masking effects could also be introduced in the criterion, and this is the subject of further studies.

References

[1] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1986.
[2] B.-S. Chen and C.-W. Lin, "Multiscale Wiener filter for the restoration of fractal signals: Wavelet filter bank approach," *IEEE Trans. Signal Processing*, vol. 42, pp. 2972–2982, Nov. 1994.
[3] A. Dembo and D. Malah, "Statistical design of analysis/synthesis systems with quantization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 328–341, Mar. 1988.
[4] P. P. Vaidyanathan and T. Chen, "Statistically optimal synthesis banks for subband coders," in *Proc. IEEE Asilomar Conf.*, Monterey, CA, Nov. 1994.
[5] B.-S. Chen, C.-W. Lin, and Y.-L. Chen, "Optimal signal reconstruction in noisy filter bank systems: Multirate Kalman synthesis filtering approach," *IEEE Trans. Signal Processing*, vol. 43, pp. 2496–2504, Nov. 1995.
[6] A. Delopoulos and S. Kollias, "Optimal filter banks for signal reconstruction from noisy subband components," *IEEE Trans. Signal Processing*, vol. 44, pp. 212–224, Feb. 1996.
[7] T. Lookabaugh, M. Perkins, and C. Cadwell, "Analysis/synthesis systems in presence of quantization," in *Proc. Int. Conf. Audio Speech Signal Process.*, 1989, pp. 1341–1344.
[8] P. H. Westerink, J. Biemond, and D. E. Boekee, "Scalar quantization error analysis for image subband coding using QMF's," *IEEE Trans. Signal Processing*, vol. 40, pp. 421–428, Apr. 1992.
[9] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[10] N. Uzun and R. A. Haddad, "Modeling and analysis of quantization errors in two channel subband filter structure," in *SPIE Visual Commun. Image Process.*, pp. 1446–1457, Nov. 1992.
[11] N. Uzun and R. A. Haddad, "Cyclostationary modeling, analysis, and optimal compensation of quantization errors in subband codecs," *IEEE Trans. Signal Processing*, vol. 43, pp. 2109–2119, Sept. 1995.
[12] R. Haddad and K. Park, "Modeling, analysis, and optimum design of quantized M-band filter banks," *IEEE Trans. Signal Processing*, vol. 43, pp. 2540–2549, Nov. 1995.
[13] J. Kovacević, "Subband coding systems incorporating quantizer models," in *Proc. Data Compression Conference*, Snowbird, UT, Mar. 1993, p. 486.
[14] ———, "Subband coding systems incorporating quantizer models," *IEEE Trans. Image Processing*, vol. 4, pp. 543–553, May 1995.
[15] L. Vandendorpe, "Optimized quantization for image subband coding," *Signal Process.: Image Commun.*, vol. 4, pp. 65–78, Nov. 1991.
[16] T. Kronander, "New criteria for optimization of QMF banks to be used in an image coding system," in *Proc. Int. Symp. Circ. Syst.*, Portland, OR, 1989, pp. 1354–1357.
[17] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. Image Processing*, vol. 2, pp. 160–175, Apr. 1993.
[18] K. Nayebi, T. P. Barnwell III, and M. J. T. Smith, "Time-domain filter bank analysis: A new design theory," *IEEE Trans. Signal Processing*, vol. 40, pp. 1412–1429, June 1992.
[19] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[20] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.
[21] A. B. Sripad and D. L. Snyder, "A necessary and sufficient condition for quantization errors to be uniform and white,"*IEEE Trans. Acoust., Speech, Signal Processing*, vol. 25, pp. 442–448, Oct. 1977.
[22] H. S. Malvar, *Signal Processing with Lapped Transforms*. Norwood, MA: Artech House, 1992.
[23] P. Chou, T. Lookabaugh, and R. Gray, "Entropy-constrained vector quantization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 31–42, Jan. 1989.
[24] T. Q. Nguyen, "Digital filter bank design: Quadratic-constrained formulation," *IEEE Trans. Signal Processing*, vol. 43, pp. 2103–2108, Sept. 1995.
[25] K. Gosse, O. Pothier, and P. Duhamel, "Optimizing the synthesis filter bank in audio coding for minimum distortion using a frequency weighted psychoacoustic criterion," in *IEEE ASSP Workshop Applications Signal Process. Audio Acoust.*, New Paltz, NY, Oct. 1995.

Karine Gosse (M'97) was born in Grenoble, France, on May 18, 1971. She graduated from the Ecole Nationale Superieure des Telecommunications (ENST), Paris, France, in 1993 and received the Ph.D. degree, also from ENST, in 1996.

Her research interests include multirate filtering and source coding, especially audio coding in subbands. She is currently working for the Centre de Recherche de Motorola, Paris, France, as a research engineer.

Pierre Duhamel (SM'87) was born in France in 1953. He received the Ing. degree in electrical engineering from the National Institute for Applied Sciences (INSA), Rennes, France, in 1975, the Dr. Ing. Degree in 1978, and the Doctorat és sciences degree in 1986, both from Orsay University, Orsay, France.

From 1975 to 1980, he was with Thomson-CSF, Paris, France, where his research interests were in circuit theory and signal processing, including digital filtering and analog fault diagnosis. In 1980, he joined the National Research Center in Telecommunications (CNET), Issy les Moulineaux, France, where his research activities were first concerned with the design of recursive CCD filters. Later, he worked on fast Fourier transforms and convolution algorithms and applied similar techniques to adaptive filtering, spectral analysis, and wavelet transforms. He is now developing studies in channel equalization (including multicarrier systems) and source coding (including joint source/channel coding). Since June 1993, he has been professor at Télécom Paris (ENST) with research activities in the same areas. He was recently appointed Head of the Signal Processing Department.

Dr. Duhamel is chairman of the IEEE DSP Committee, was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1989 to 1991, and was an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He was a guest editor for the special issue of the IEEE TRANSACTIONS ON SIGNAL PROCESSING on wavelets.