Derivation of Subband Coding Gain: The Most General Case

Anamitra Makur

School of EEE

Nanyang Technological University

Singapore

1 Reconstruction Error Variance

Consider the subband coder of figure 1, consisting of $k = 0, 1, ..., M-1$ channels. Each channel has an analysis filter $H_k(z)$, a synthesis filter $F_k(z)$, down/up-sampling ratio m_k, and a b_k bit quantizer. The input is $x[n]$, after passing through the analysis filters it becomes $x'_k[n]$, and the subband signals are $x_k[n]$. The subband coder is maximally decimated,

$$\sum_{k=0}^{M-1} m_k = 1. \tag{1}$$

Assumption 1: Assume the input $x[n]$ to be wide sense stationary (WSS) [1, page 806].

Assumption 2: Assume the input $x[n]$ to be zero mean. It follows that all subsequent signals (including quantization noise) become zero mean. (This assumption does not cause any difficulty in applying subband coding to nonzero mean signals such as image, since that system is equivalent to a zero mean system obtained by subtracting the mean from the input.)

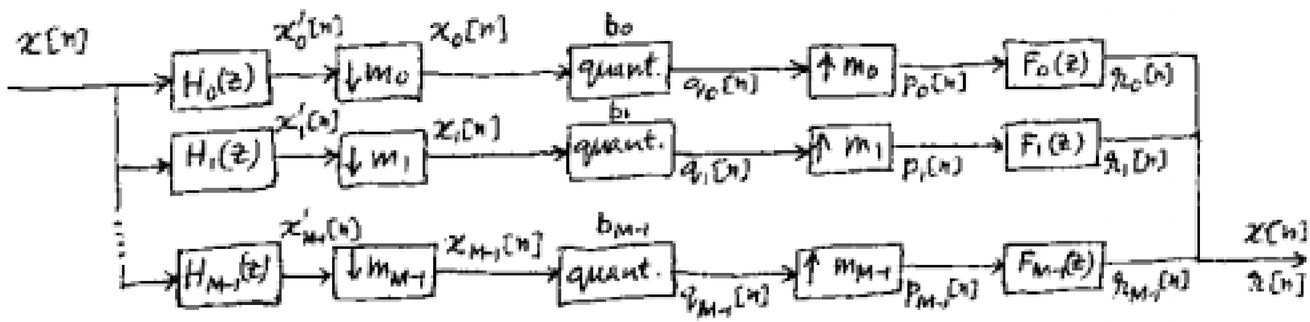

Figure 1: Subband coder

From assumption 1, $S_{xx}(e^{j\omega})$ is the power spectral density (PSD) [1, page 807] of $x[n]$. From assumption 2, variance of $x[n]$ is

$$\sigma_x^2 = \int_{-\pi}^{\pi} S_{xx}(e^{j\omega}) \frac{d\omega}{2\pi}. \tag{2}$$

After passing through the analysis filter $H_i(z)$, variance of $x'_k[n]$ is [1, page 810]

$$\sigma_k^2 = \int_{-\pi}^{\pi} S_{xx}(e^{j\omega}) |H_k(e^{j\omega})|^2 \frac{d\omega}{2\pi} \tag{3}$$

for $k = 0, 1, ..., M-1$. From assumption 1, variance does not change after downsampling [1, page 406] [2, page 136], so variance of $x_k[n]$, the k-th subband, is still σ_k^2.

A quantizer is modelled as additive noise model [3, chapter 4]. So, the output from a quantizer is $x_k[n] + q_k[n]$, where $x_k[n]$ is its input and $q_k[n]$ is the quantization noise. For both uniform quantizers and probability density function (PDF) optimized quantizers, the quantization noise variance is [3, chapter 4] [1, page 818]

$$\sigma_{q_k}^2 = \epsilon_k 2^{-2b_k} \sigma_k^2 \tag{4}$$

where ϵ_k is a constant that depends on the PDF of the k-th subband signal.

1.1 General Filter Bank

Consider figure 1 after quantization. $x_k[n] + q_k[n]$ enters each channel of the synthesis bank. If only subband signals $x_k[n]$ enter, since the filter bank is perfect reconstruction [1, page 196], the output will be equal to $x[n]$, the input (neglecting delay). If only noises $q_k[n]$ enter, the corresponding signals are shown in figure 1 below the signal arrows. $q_k[n]$ after upsampling becomes $p_k[n]$, which after the filters becomes $r_k[n]$, and the output will be $r[n]$. Due to linearity of the synthesis bank, the output with both signal and noise entering will be equal to the sum of these two outputs, $x[n] + r[n]$. Therefore, the reconstruction error (output minus input) is $r[n]$, entirely obtained from $q_k[n]$ passing through the synthesis filter bank.

Assumption 3a: Assume uniform (also, high bit rate) quantizers.

From assumption 3a, $q_k[n]$ of different channels are uncorrelated among themselves and white [3, chapter 4] [1, page 818]. This model of uniform quantizers is for bit rate more than 3 bits/sample [4], but is no longer appropriate for lower rate uniform quantizers. Since these are white and zero mean, the PSD of $q_k[n]$ is just a constant $\sigma_{q_k}^2$. $p_k[n]$, obtained after upsampling $q_k[n]$, is

$$p_k[n] = \begin{cases} q_k[n/m_k] & \text{if } n \bmod m_k = 0 \\ 0 & \text{else} \end{cases} \tag{5}$$

which is no longer WSS but is cyclo wide sense stationary with period m_k [2, page 136]. In other words, the PSD of $p_k[n]$ if $n \bmod m_k = 0$ is $\sigma^2_{q_k}$, and its PSD for the remaining samples is 0. $p_k[n]$ passes through the synthesis filter $F_k(z)$ to give $r_k[n]$ which is also cyclo-WSS with period m_k. The average variance of $r_k[n]$, averaged over the period m_k, is

$$\frac{1}{m_k}\left[\int_{-\pi}^{\pi} \sigma^2_{q_k}|F_k(e^{j\omega})|^2 \frac{d\omega}{2\pi} + 0 + \cdots + 0\right] \tag{6}$$

where $m_k - 1$ zeros come from the zero noise samples. The reconstruction error $r[n]$ is the sum of all $r_k[n]$. Since $r_k[n]$ across the channels are uncorrelated from assumption 3a, the variance of their sum is the sum of their variances [5, page 153],

$$\sigma_r^2 = \sum_{k=0}^{M-1} \frac{\sigma^2_{q_k}}{m_k} \int_{-\pi}^{\pi} |F_k(e^{j\omega})|^2 \frac{d\omega}{2\pi} \tag{7}$$

where σ_r^2 is the reconstruction error variance. Let $n_k = \int_{-\pi}^{\pi} |F_k(e^{j\omega})|^2 d\omega/2\pi$ denote the synthesis filter norms (note that for an FIR filter with impulse response coefficients $[h_0, h_1, \ldots, h_L]$ the norm is $h_0^2 + h_1^2 + \cdots + h_L^2$). Then using equation (4), the above equation becomes

$$\sigma_r^2 = \sum_{k=0}^{M-1} \frac{\epsilon_k 2^{-2b_k} \sigma_k^2 n_k}{m_k}. \tag{8}$$

1.2 Orthogonal Filter Bank

Assumption 3b: Assume the filter bank to be orthogonal.

A filter bank is orthogonal/paraunitary when the (analysis and) synthesis filters and their appropriately shifted versions are orthogonal to each other [6, page 1832]. If such a filter bank has unequal resampling ratios (m_k are not equal as in figure 1), it may be transformed into a filter bank with equal resampling ratios (all m_k are equal) [6, page 1833] as described below.

1.2.1 Unequal to Equal Resampling Ratios

Let $N = \text{lcm}(m_0, m_1, \ldots, m_{M-1})$. A channel with down/up-sampling ratio m_k is shown in figure 2. Introduce a delay chain system with resampling ratio $\frac{N}{m_k}$ between the analysis and the synthesis side, as shown in figure 3. Since the delay chain system introduced is

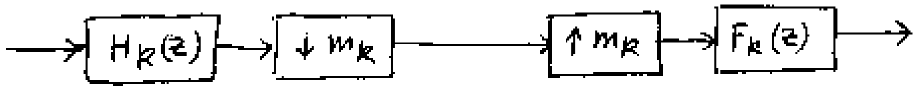

Figure 2: k-th channel

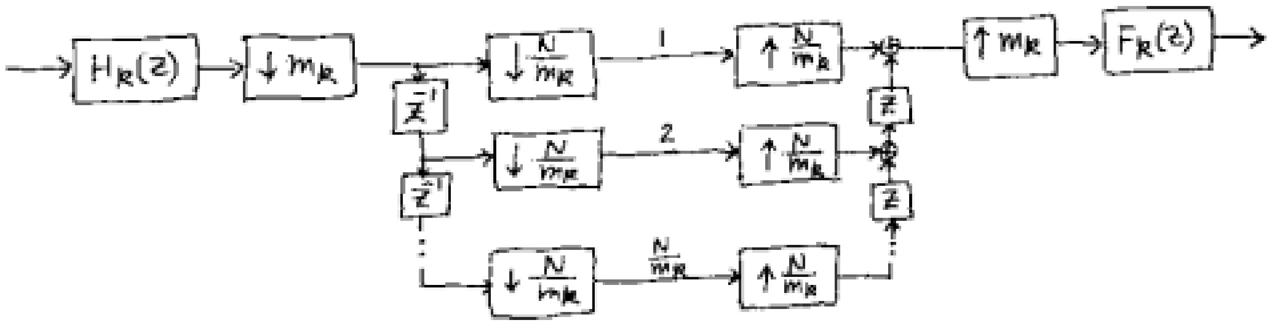

Figure 3: k-th channel with a delay chain introduced

perfect reconstruction [1, page 235], and has zero delay (since positive power of z is used in the synthesis side), both systems of figures 2 and 3 are identical. Moving the delays of the introduced system before the m_k-downsampler and after the m_k-upsampler results in the system of figure 4. Effectively, one channel with resampling ratio m_k is replaced by $\frac{N}{m_k}$ channels with resampling ratio N.

In a similar fashion, all channels $0 \leq k \leq M-1$ may be converted to resampling ratio N systems. The total number of channels will be

$$\frac{N}{m_0} + \frac{N}{m_1} + \cdots + \frac{N}{m_{M-1}} = N \tag{9}$$

using equation (1). Thus, the filter bank with unequal resampling ratios is expressed as a filter bank with equal resampling ratios having N channels.

It is known that for an orthogonal filter bank with equal resampling ratios, the analysis and

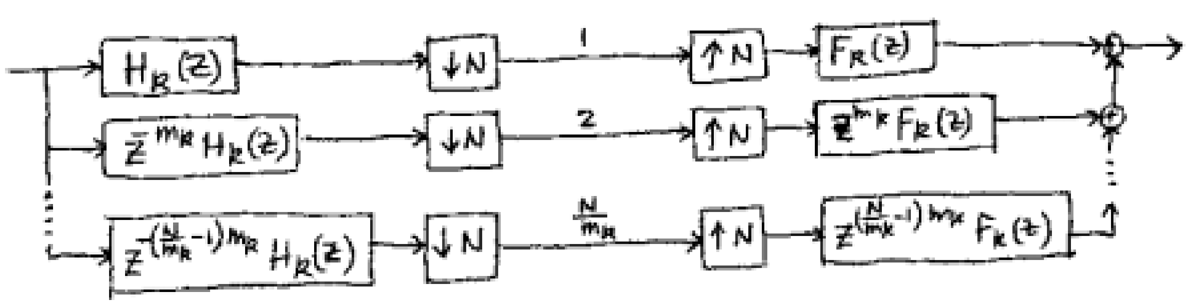

Figure 4: Equivalent system to the k-th channel

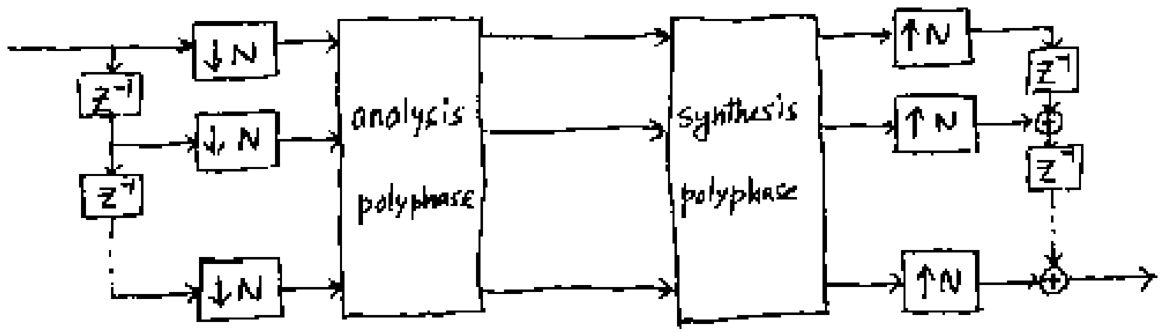

Figure 5: Subband coder in polyphase form

synthesis polyphase matrices are paraunitary [1, page 294]. The polyphase matrix, defined for filter banks with equal resampling ratio, may be extended to unequal resampling ratios of figure 1, since from figure 4 its equivalent $N \times N$ polyphase matrix may be constructed. These polyphase matrices turn out to be also paraunitary [6, page 1833]. (Such orthogonal filter bank with unequal resampling ratio may be constructed using few orthogonal filter banks with equal resampling ratio in a tree structure.)

1.2.2 Error Variance

The analysis and synthesis sides of figure 1 may therefore be expressed in polyphase forms, as shown in figure 5. For a paraunitary system, the lossless (or, energy balance) property says that the output energy is equal to the input energy [1, page 725] [also shown in Appendix 2]. Consider only quantization noise entering the synthesis bank. Let ${\sigma'}^2_{q_i}$ be the quantization noise variance of the i-th channel for $i = 0, \ldots, N-1$. Then the input variance (energy) averaged over all channels is $\frac{1}{N}\sum_{i=0}^{N-1} {\sigma'}^2_{q_i}$. This is equal to the average output variance after the synthesis polyphase. The upsample, delay and add chain simply converts the output vector sequence into a scalar sequence. Therefore, this is equal to the average output variance, or the reconstruction error variance.

$$\sigma_r^2 = \frac{1}{N}\sum_{i=0}^{N-1} {\sigma'}^2_{q_i} \tag{10}$$

Note that N channels of figure 5 are obtained from M channels of figure 1. Therefore, first $\frac{N}{m_0}$ channels of figure 5 have the same quantizer as channel 0 of figure 1, so ${\sigma'}^2_{q_i}$ is equal to $\sigma^2_{q_0}$ for $i = 0, \ldots, \frac{N}{m_0} - 1$. Similarly other variances may be equated, resulting in

$$\sigma_r^2 = \frac{1}{N}\sum_{k=0}^{M-1} \frac{N}{m_k}\sigma^2_{q_k} = \sum_{k=0}^{M-1} \frac{\sigma^2_{q_k}}{m_k}. \tag{11}$$

Using equation (4), it becomes

$$\sigma_r^2 = \sum_{k=0}^{M-1} \frac{\epsilon_k 2^{-2b_k} \sigma_k^2}{m_k}. \tag{12}$$

Comparing equations (8) and (12), for orthogonal case the synthesis filter norms n_k do not appear. This is because lossless property implies unit energy filters or $n_k = 1$ for all k [1, page 297]. Also note that assumption 3b replaces assumption 3a, or the assumption of uniform high rate quantizers is not necessary to derive this result. In the quantizer sense, therefore, the result is more general for orthogonal case.

2 Optimal Bit Allocation

Referring to figure 1, b_k is the k-th quantizer bit rate per quantized sample. But due to a downsampling ratio of m_k, one sample is quantized by this quantizer for every m_k input samples. Therefore, its bit rate is b_k/m_k bits per input sample. The average bit rate of the subband system is, therefore,

$$b = \sum_{k=0}^{M-1} \frac{b_k}{m_k} \tag{13}$$

bits per input sample. The optimal bit allocation problem is to find $b_0, b_1, \ldots, b_{M-1}$ that minimizes the reconstruction error variance σ_r^2 of equation (8) subject to the constraint of equation (13).

This constrained minimization may be solved using the Lagrange multiplier method (an alternative method is given in Appendix 3). Define the cost

$$C = \sum_{k=0}^{M-1} \frac{1}{m_k} \epsilon_k 2^{-2b_k} \sigma_k^2 n_k + \lambda \left(\sum_{k=0}^{M-1} \frac{b_k}{m_k} - b \right) \tag{14}$$

where λ is the Lagrange multiplier. Differentiating C with respect to b_k and equating to zero, $\frac{\partial C}{\partial b_k} = 0$ or

$$\frac{1}{m_k} \epsilon_k 2^{-2b_k} \sigma_k^2 n_k (-2 \ln 2) + \lambda \frac{1}{m_k} = 0 \tag{15}$$

which follows from $\frac{\partial}{\partial b_k} 2^{-2b_k} = \frac{\partial}{\partial b_k} e^{-2 \ln 2 b_k} = -2 \ln 2 e^{-2 \ln 2 b_k}$. Therefore

$$2 \ln 2 \epsilon_k 2^{-2b_k} \sigma_k^2 n_k = \lambda \tag{16}$$

or

$$2^{-2b_k} = \frac{\lambda}{2 \ln 2 \epsilon_k \sigma_k^2 n_k}. \tag{17}$$

Taking logarithm base 2 of both sides,

$$-2b_k = \log_2 \frac{\lambda}{2\ln 2\epsilon_k \sigma_k^2 n_k} \tag{18}$$

or

$$b_k = \frac{1}{2}\log_2 \frac{2\ln 2\epsilon_k \sigma_k^2 n_k}{\lambda} = \frac{1}{2}\log_2 \frac{2\ln 2}{\lambda} + \frac{1}{2}\log_2(\epsilon_k \sigma_k^2 n_k). \tag{19}$$

The above equation is valid for all values of k from 0 to $M-1$. Substituting equation (19) in the constraint, or equation (13), we obtain

$$\begin{aligned} b &= \sum_{k=0}^{M-1} \frac{1}{m_k}\left[\frac{1}{2}\log_2 \frac{2\ln 2}{\lambda} + \frac{1}{2}\log_2(\epsilon_k \sigma_k^2 n_k)\right] && (20) \\ &= \frac{1}{2}\log_2 \frac{2\ln 2}{\lambda} \sum_{k=0}^{M-1} \frac{1}{m_k} + \frac{1}{2}\sum_{k=0}^{M-1} \log_2(\epsilon_k \sigma_k^2 n_k)^{1/m_k}. && (21) \end{aligned}$$

The first sum is 1 from equation (1). The second sum of logarithms may be written as a logarithm of product. Using these simplifications,

$$b = \frac{1}{2}\log_2 \frac{2\ln 2}{\lambda} + \frac{1}{2}\log_2 \prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2 n_i)^{1/m_i} \tag{22}$$

or

$$\frac{1}{2}\log_2 \frac{2\ln 2}{\lambda} = b - \frac{1}{2}\log_2 \prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2 n_i)^{1/m_i}. \tag{23}$$

Substituting the above result in equation (19), we obtain

$$\begin{aligned} b_k &= b - \frac{1}{2}\log_2 \prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2 n_i)^{1/m_i} + \frac{1}{2}\log_2(\epsilon_k \sigma_k^2 n_k) && (24) \\ &= b + \frac{1}{2}\log_2 \frac{\epsilon_k \sigma_k^2 n_k}{\prod_{i=0}^{M-1}(\epsilon_i \sigma_i^2 n_i)^{1/m_i}} && (25) \end{aligned}$$

for all values of k from 0 to $M-1$. This is the optimal bit allocation. Note that for orthogonal case, the optimal bit allocation is obtained by substituting $n_k = 1$ for all k. Note also that the resulting bit rates are real and may be negative.

3 Subband Coding Gain

We now find the minimum reconstruction error variance using the optimal bit allocation. From equation (25),

$$2^{-2b_k} = 2^{-2b} \cdot \frac{\prod_{i=0}^{M-1}(\epsilon_i \sigma_i^2 n_i)^{1/m_i}}{\epsilon_k \sigma_k^2 n_k} \tag{26}$$

for all k from 0 to $M-1$. Substituting this into equation (8),

$$\sigma_r^2 = \sum_{k=0}^{M-1} \frac{1}{m_k} 2^{-2b} \prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2 n_i)^{1/m_i} \tag{27}$$

$$= 2^{-2b} \prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2 n_i)^{1/m_i} \sum_{k=0}^{M-1} \frac{1}{m_k} \tag{28}$$

$$= 2^{-2b} \prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2 n_i)^{1/m_i} \tag{29}$$

using equation (1) in the last step.

Subband coding gain is defined as the ratio of the reconstruction error variances of the fullband coder (or, PCM) to the subband coder for same bit rate. The fullband coder simply quantizes input $x[n]$ using a b bit quantizer. Therefore, from equation (4), its quantization error variance (which is also its reconstruction error variance) is

$$\sigma_q^2 = \epsilon 2^{-2b} \sigma_x^2 \tag{30}$$

where ϵ is as before. Therefore the subband coding gain is

$$G = \frac{\epsilon 2^{-2b} \sigma_x^2}{2^{-2b} \prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2 n_i)^{1/m_i}} = \frac{\epsilon \sigma_x^2}{\prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2 n_i)^{1/m_i}}. \tag{31}$$

Note that the assumptions used are 1, 2, and 3a.

Some special cases of subband coding gain are provided in the following subsections. Subband coding gain for two-dimensional filter bank is given in Appendix 1. Optimal bit allocation and subband coding gain derivations may be found for Gaussian, equal resampling ratio, orthogonal case in [1, page 821,823]; for equal resampling ratio orthogonal, tree structured orthogonal, and unequal resampling ratio orthogonal case in [6]; for equal resampling ratio general case in [7]; and for two-dimensional, Gaussian, unequal resampling ratio, general case in [8]. Extensions for more general quantizer or more appropriate bit allocation have been done. A discussion of additive model and gain-plus-noise model of quantizer is found in [4]. Optimal bit allocation and subband coding gain results for gain-plus-noise model of quantizer is found in [4,9]. Optimal non-negative bit allocation may be found in [10].

3.1 Gaussian Input

Assumption 4a: Assume the input $x[n]$ to be Gaussian (samples have Gaussian PDF).

Then it is known that output of a linear system with Gaussian input is also Gaussian [5, page 243]. Thus all subband signals will be Gaussian, resulting in $\epsilon = \epsilon_0 = \epsilon_1 = \cdots = \epsilon_{M-1}$

(assumption 3a is also necessary). Therefore

$$G = \frac{\epsilon \sigma_x^2}{\epsilon^{\frac{1}{m_0}+\cdots+\frac{1}{m_{M-1}}} \prod_{i=0}^{M-1} (\sigma_i^2 n_i)^{1/m_i}} = \frac{\sigma_x^2}{\prod_{i=0}^{M-1} (\sigma_i^2 n_i)^{1/m_i}}. \tag{32}$$

Note that the assumptions used are 1, 2, 3a, and 4a.

3.2 Equal Resampling Ratio Filter Bank

Assumption 4b: Assume the resampling ratios to be equal, $m_0 = m_1 = \cdots = m_{M-1}$.

Then from equation (1) each ratio is equal to $1/M$, and the subband coding gain becomes

$$G = \frac{\epsilon \sigma_x^2}{\left(\prod_{i=0}^{M-1} \epsilon_i \sigma_i^2 n_i \right)^{\frac{1}{M}}}. \tag{33}$$

Note that the assumptions used are 1, 2, 3a, and 4b.

3.3 Orthogonal Filter Bank

Let us now consider assumption 3b of orthogonal filter bank instead of assumption 3a. Since $n_k = 1$ as stated earlier, the subband coding gain is

$$G = \frac{\epsilon \sigma_x^2}{\prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2)^{1/m_i}}. \tag{34}$$

Further, the lossless property stated earlier may be applied to the analysis polyphase matrix of figure 5 in a way similar to section 1.2.2. Since the input to the analysis polyphase is the input $x[n]$ (in a vector or parallel form), its average input variance is σ_x^2. Let ${\sigma'}_i^2$ denote the output variances of $i = 0, \ldots, N-1$ subbands. Then from lossless property

$$\sigma_x^2 = \frac{1}{N} \sum_{i=0}^{N-1} {\sigma'}_i^2. \tag{35}$$

Since N channels of figure 5 are obtained from M channels of figure 1, first $\frac{N}{m_0}$ channels of figure 5 have the same subband variance as channel 0 of figure 1, so ${\sigma'}_i^2$ is equal to σ_0^2 for $i = 0, \ldots, \frac{N}{m_0} - 1$. Similarly other subband variances may be equated, resulting in

$$\sigma_x^2 = \frac{1}{N} \sum_{k=0}^{M-1} \frac{N}{m_k} \sigma_k^2 = \sum_{k=0}^{M-1} \frac{\sigma_k^2}{m_k}. \tag{36}$$

Using this in equation (34), the subband coding gain becomes

$$G = \frac{\epsilon \sum_{i=0}^{M-1} \sigma_i^2 / m_i}{\prod_{i=0}^{M-1} (\epsilon_i \sigma_i^2)^{1/m_i}}. \tag{37}$$

Note that the assumptions used are 1, 2, and 3b.

3.4 Gaussian, Equal Resampling Ratio, and Orthogonal

It is possible to obtain other cases by considering more than one of assumptions 4a, 4b and 3b. The most simplified expression for subband coding gain is obtained by considering assumptions 4a, 4b and 3b simultaneously, and is

$$G = \frac{\frac{1}{M}\sum_{i=0}^{M-1}\sigma_i^2}{\left(\prod_{i=0}^{M-1}\sigma_i^2\right)^{\frac{1}{M}}} \tag{38}$$

which becomes the ratio of arithmetic mean to geometric mean of non-negative quantities σ_i^2. Since arithmetic mean is more than or equal to geometric mean [1, page 820], it therefore follows that $G \geq 1$, or the subband coder performs equal or better to the fullband coder for any input.

References

[1] P. P. Vaidyanathan, "*Multirate systems and filter banks*," Prentice Hall: Englewood Cliffs, 1993.

[2] V. P. Sathe, P. P. Vaidyanathan, "Effects of multirate systems on the statistical properties of random signals," *IEEE Transactions Signal Processing*, 41(1), January 1993, pp. 131–146.

[3] N. S. Jayant, P. Noll, "*Digital coding of waveforms*," Prentice Hall: Englewood Cliffs, 1984.

[4] N. Uzun, R. A. Haddad, "Cyclostationary modeling, analysis, and optimal compensation of quantization errors in subband codecs," *IEEE Transactions Signal Processing*, 43(9), September 1995, pp. 2109–2119.

[5] A. Papoulis, "*Probability, random variables, and stochastic processes*," McGraw-Hill: Tokyo, 1984.

[6] A. K. Soman, P. P. Vaidyanathan, "Coding gain in paraunitary analysis/synthesis systems," *IEEE Transactions Signal Processing*, 41(5), May 1993, pp. 1824–1835.

[7] I. Djokovic, P. P. Vaidyanathan, "On optimal analysis/synthesis filters for coding gain maximization," *IEEE Transactions Signal Processing*, 44(5), May 1996, pp. 1276–1279.

[8] G. Calvagno, G. A. Mian, R. Rinaldo, "Computation of the coding gain for subband coders," *IEEE Transactions Communications*, 44(4), April 1996, pp. 475–487.

[9] C. W. Kok, T. Q. Nguyen, "Multirate filter banks and transform coding gain," *IEEE Transactions Signal Processing*, 46(7), July 1998, pp. 2041–2044.

[10] K. Gosse, P. Duhamel, "Perfect reconstruction versus MMSE filter banks in source coding," *IEEE Transactions Signal Processing*, 45(9), September 1997, pp. 2188–2202.

Appendix 1: Two-Dimensional Filter Bank

For image applications, typically a one-dimensional filter bank is used to construct a two-dimensional separable filter bank [1, page 640]. For example, figure 6 shows one channel of a two-dimensional synthesis filter bank using a 2-channel one-dimensional filter bank. The subband image is upsampled by a factor of 2 row-wise (add all-zero column between every pair of columns) and by a factor of 2 column-wise (add all-zero row between every pair of rows). The one-dimensional low-pass filter F_0 is applied row-wise on the upsampled image (denoted by $F_0(z_1)$ where z_1 is the horizontal z-transform variable). Then the one-dimensional high-pass filter F_1 is applied column-wise on the resulting output (denoted by $F_1(z_2)$). (The sequence of row-wise before column-wise may be reversed.) This is called the reconstruction of the low-high or LH band in image compression literature. The upsampling ratio in two dimensions is a matrix [1, page 572], and m_k in the subband coding gain expressions is the determinant of this matrix. In the example, the upsampling ratio is $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, and

$$m_k = \det \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 4. \tag{39}$$

The effective filter is the product of filters. In the example, the filter is $F_0(z_1)F_1(z_2)$. The effective synthesis filter norm in the subband coding gain expressions is the product of each one-dimensional filter norm. If F_0 has a norm n_0, and F_1 has a norm n_1 (note that filter norm does not change whether the filter is applied row-wise or column-wise), then the effective

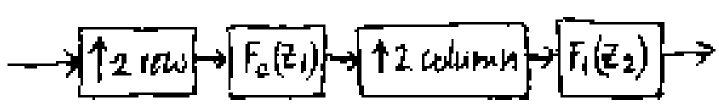

Figure 6: one channel of a two-dimensional synthesis filter bank

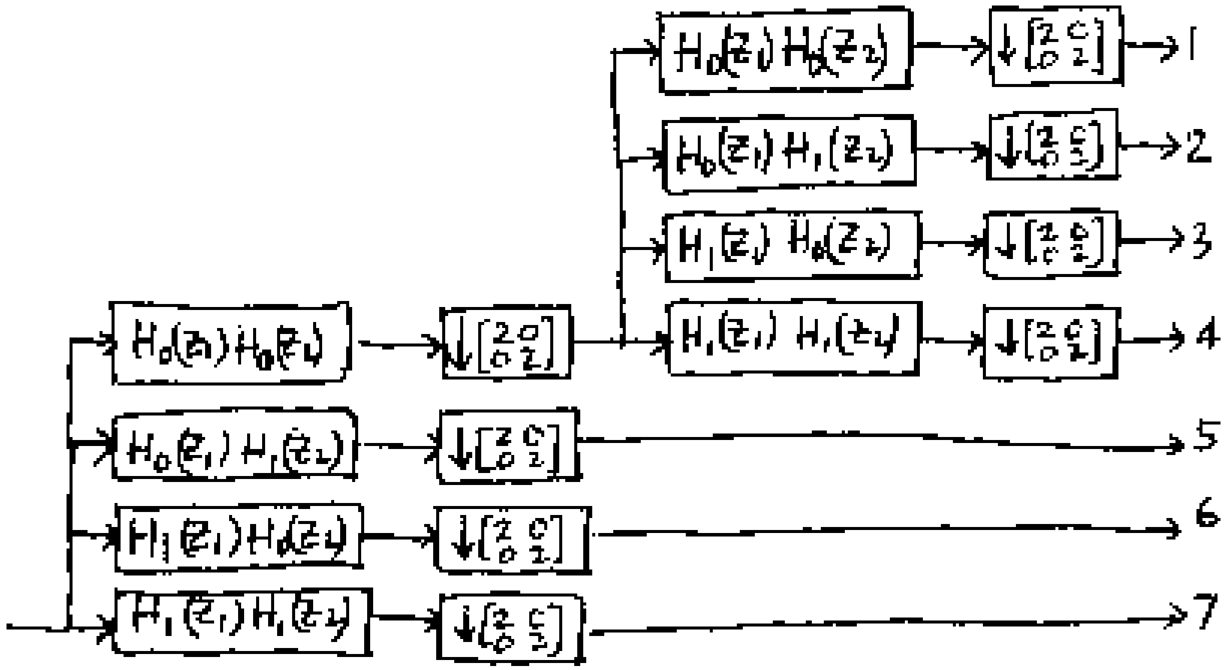

Figure 7: 2 level wavelet decomposition (analysis only)

norm is $n_0 n_1$. (If non-separable filter is used, then the filter norm may be computed by squaring and summing its impulse response coefficients.)

To take an example of a complete two-dimensional filter bank, consider the 2 level wavelet decomposition shown in figure 7 that has 7 subbands. Let n_0 be the synthesis filter norm of the one-dimensional low-pass F_0, and n_1 be the synthesis filter norm of the high-pass F_1. The subband coding gain expression for this coder will use resampling ratios $m_k = 16, 16, 16, 16, 4, 4, 4$ and synthesis filter norms $n_k = n_0^4, n_0^3 n_1, n_0^3 n_1, n_0^2 n_1^2, n_0 n_1, n_0 n_1, n_1^2$ respectively.

Appendix 2: Lossless Property of a Paraunitary System

Let $\mathbf{x}[n] = \begin{bmatrix} x_0[n] \\ \vdots \\ x_{N-1}[n] \end{bmatrix}$ be the input vector and $\mathbf{y}[n] = \begin{bmatrix} y_0[n] \\ \vdots \\ y_{N-1}[n] \end{bmatrix}$ be the corresponding output vector through an $N \times N$ paraunitary transfer matrix $\mathbf{A}(z)$. Let $\mathbf{S_{xx}}(e^{j\omega})$ be the $N \times N$ PSD matrix of the input vector $\mathbf{x}(n)$. Note that the PSD of the i-th input component, $x_i[n]$, is the ii-th element of $\mathbf{S}_{\mathbf{xx}}(e^{j\omega})$. Therefore the variance of $x_i[n]$ is

$$\int_{-\pi}^{\pi} \left[\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\right]_{ii} \frac{d\omega}{2\pi}. \tag{40}$$

The average input variance is

$$\frac{1}{N}\sum_{i=0}^{N-1} \text{variance of } x_i[n] = \frac{1}{N}\int_{-\pi}^{\pi} \text{trace}\left(\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\right) \frac{d\omega}{2\pi}. \tag{41}$$

The PSD of the output vector is

$$\mathbf{S}_{\mathbf{yy}}(e^{j\omega}) = \mathbf{A}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{A}^H(e^{j\omega}) \tag{42}$$

where H denotes Hermitian. Therefore, the average output variance is

$$\frac{1}{N}\int_{-\pi}^{\pi} \text{trace}\left(\mathbf{S}_{\mathbf{yy}}(e^{j\omega})\right) \frac{d\omega}{2\pi} = \frac{1}{N}\int_{-\pi}^{\pi} \text{trace}\left(\mathbf{A}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{A}^H(e^{j\omega})\right) \frac{d\omega}{2\pi}. \tag{43}$$

Since trace(AB) =trace(BA), it simplifies to

$$\frac{1}{N}\int_{-\pi}^{\pi} \text{trace}\left(\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{A}^H(e^{j\omega})\mathbf{A}(e^{j\omega})\right) \frac{d\omega}{2\pi}. \tag{44}$$

Since $\mathbf{A}(z)$ is paraunitary, $\mathbf{A}^H(e^{j\omega})\mathbf{A}(e^{j\omega}) = \mathbf{I}$, and it becomes

$$\frac{1}{N}\int_{-\pi}^{\pi} \text{trace}\left(\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\right) \frac{d\omega}{2\pi} \tag{45}$$

or the average output variance is equal to the average input variance. This shows that the energy is preserved through a paraunitary system.

Appendix 3: Another Method to Optimal Bit Allocation

The optimal bit allocation problem is to find $b_0, \ldots, b_{M-1}$ subject to $b = \sum_k b_k/m_k$ (constraint) that minimizes $\sum_k \epsilon_k 2^{-2b_k}\sigma_k^2 n_k/m_k$ (cost).

The constraint may be written as

$$2^{-2b} = \prod_k \left(2^{-2b_k}\right)^{1/m_k}. \tag{46}$$

Multiply both sides by $\prod_k \left(\epsilon_k\sigma_k^2 n_k\right)^{1/m_k}$ to get

$$2^{-2b}\prod_k \left(\epsilon_k\sigma_k^2 n_k\right)^{1/m_k} = \prod_k \left(2^{-2b_k}\epsilon_k\sigma_k^2 n_k\right)^{1/m_k}. \tag{47}$$

The left hand side is a constant since there is no b_k. Express the right hand side using $X_k = \epsilon_k 2^{-2b_k}\sigma_k^2 n_k$ (where X_k are variables since they are functions of b_k) as

$$\text{constant} = \prod_k X_k^{1/m_k}. \tag{48}$$

The cost may also be simplified using X_k. So, the problem becomes to find $X_0, \ldots, X_{M-1}$ subject to constant$= \prod_k X_k^{1/m_k}$ that minimizes $\sum_k X_k/m_k$.

The constraint may further be expressed as constant$= \prod_{k=0}^{M-1}(X_k^{N/m_k})^{\frac{1}{N}}$. Define N variables Y_i such that first $\frac{N}{m_0}$ Y_i are equal to X_0, and so forth. Then the constraint becomes constant$= \prod_{i=0}^{N-1}(Y_i)^{\frac{1}{N}}$. Taking both sides to the power of N, it is constant$= \prod_i Y_i$. Similarly the cost becomes $\sum_{k=0}^{M-1} \frac{1}{N} \cdot \frac{N}{m_k} X_k = \sum_{i=0}^{N-1} \frac{1}{N} Y_i$ where the constant factor $\frac{1}{N}$ may be neglected. So, the problem becomes to find $Y_0, \ldots, Y_{N-1}$ subject to constant$= \prod_i Y_i$ that minimizes $\sum_i Y_i$.

Since $X_k = \epsilon_k 2^{-2b_k}\sigma_k^2 n_k \geq 0$, the variables Y_i are non-negative. The solution to minimizing the sum of a few non-negative variables given that their product is a constant, is known to be when all variables are equal. Therefore the solution is $Y_0 = Y_1 = \cdots = Y_{N-1}$, or $X_0 = X_1 = \cdots = X_{M-1} = X$ for some value X. The physical interpretation of this is as follows. All $\epsilon_k 2^{-2b_k}\sigma_k^2 n_k$ are equal, which means from equation (4) all $\sigma_{q_k}^2 n_k$ are equal, or the quantization noise energy observed after the synthesis bank is equal in each channel.

The value of X may be found from the contraint. Since $X_k = X$, we have $2^{-2b_k} = X/(\epsilon_k \sigma_k^2 n_k)$. Putting this in equation (46),

$$2^{-2b} = \prod_k \left(\frac{X}{\epsilon_k \sigma_k^2 n_k}\right)^{1/m_k} = \frac{X^{\frac{1}{m_0}+\cdots+\frac{1}{m_{M-1}}}}{\prod_k(\epsilon_k \sigma_k^2 n_k)^{1/m_k}} \tag{49}$$

or using equation (1)

$$X = 2^{-2b}\prod_k(\epsilon_k \sigma_k^2 n_k)^{1/m_k}. \tag{50}$$

The optimal bit allocation of equation (25) readily follows.