IEEE Transactions on Information Theory, Vol. 41, pp. 2094–2100, Nov. 1995

Information Theoretic Performance of Quadrature Mirror Filters *

Ajay Divakaran †and William A. Pearlman ‡

Abstract

Most existing QMFs closely match the derived closed form expression for an efficient class of Quadrature Mirror Filters. We use the closed form expressions to derive the relationship between information theoretic loss and the frequency selectivity of the QMF, by calculating first order entropy as well as rate distortion theoretic performance of a two band QMF system. We find that practical QMFs do not suffer a significant information theoretic loss with first order autoregressive Gaussian sources. With second order autoregressive sources we find that practical QMFs suffer a notable information theoretic loss when the bandwidth of the source is extremely narrow, but incur a small loss when the bandwidth is wider. We suggest that our results broadly apply to higher order autoregressive sources as well.

Index Terms - quadrature mirror filters, subband coding, source coding, rate-distortion theory.

1 Introduction

The discrete-time Fourier transforms of the impulse responses $h_l(\omega)$ (low-pass)and $h_u(\omega)$ (high-pass) of a Quadrature Mirror Filter (QMF) [6] pair satisfy the following conditions for perfect reconstruction:

$$H_u(\omega) = e^{j\omega} H_l(\pi - \omega) \tag{1}$$

$$|H_l(\omega)|^2 + |H_l(\pi - \omega)|^2 = 1 \tag{2}$$

The objective is, therefore, to design a linear phase FIR filter $H(\omega)$ that satisfies the power complementarity condition (2). Past approaches to this problem [11, 14, 16, 9] have not yielded a simple and general analytical framework, because they have relied on optimization by means of numerical search algorithms. In our previous work [4, 3, 5, 2], we derived simple closed form expressions for an efficient class of QMFs, and compared to existing QMF designs and developed a new implementation technique which exploits the closed form expressions derived in [4, 3, 5, 2], and yields FIR implementations that equal existing designs in performance.

*This work was performed at Rensselaer Polytechnic Institute and was supported in part by the National Science Foundation under Grant No. NCR-9004758. The U.S. Government has certain rights in this material.

†A. Divakaran is currently with Iterated Systems Inc., 3525 Piedmont Rd., Atlanta, GA 30305-1530, e-mail: ajayd@iterated.com.

‡W.A. Pearlman is with the Electrical, Computer and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY 12180-3590, e-mail: pearlman@ecse.rpi.edu.

Brickwall filters provide perfect frequency selectivity but their infinite roll-off cannot be realized in practice with a finite number of filter taps or coefficients. While we would like to use as few filter taps as possible, the fewer the filter taps, the lower the frequency selectivity. Reduction in frequency selectivity of QMFs has deleterious information theoretic consequences.

Rao and Pearlman [15] have derived an expression for the difference in first order entropy between the source and a subband decomposition thereof, and shown that it is always non-negative. This entropy difference is exactly the rate advantage of scalar encoding the subbands over scalar encoding the fullband, i.e., the source. Although the rate expression was derived under the assumption of ideal or brickwall filtering, it applies to the usual situation of independent coding of subbands. This motivates us to look into the relationship between first order entropy of feasible QMF pyramids and frequency selectivity. Furthermore, Fischer [7] has shown that realizable QMFs incur a non-zero rate distortion theoretic penalty that increases as the frequency selectivity goes down. He has derived a very convenient formula for calculating the rate distortion theoretic loss incurred by a particular QMF. However, in the absence of a general formula for QMFs, we have to proceed on a filter by filter basis by substituting practical filter tap values each time. This procedure is cumbersome and does not give a direct indication of the information theoretic consequences of changing frequency selectivity.

In this paper, we use our closed form expression for QMFs to derive the relationship between frequency selectivity and information theoretic performance of QMFs, by calculating the first order entropy as well as the rate distortion theoretic performance of a two band QMF system. We conclude that with as few as 9 filter taps, the information theoretic loss incurred by realizable QMFs with first order auto-regressive Gaussian sources is negligible. With second order auto-regressive sources with wider bandwidth we find that we need at least 64 filter taps to keep the deterioration of scalar coding performance below 10%. When the bandwidth of AR(2) sources is extremely narrow, the information theoretic penalty incurred by realizable QMFs is considerable.

2 An efficient class of QMFs

We now describe (see [4, 3, 5, 2]) a class of efficient QMFs which require the fewest taps for FIR implementation, for a given frequency tuning. We let

$$H_p(\omega) = |H(\omega)|^2 \tag{3}$$

Then we can write an expression for a generic $H_p(\omega)$, which has zero attenuation in the pass band and infinite attenuation in the stop band, bearing in mind that the ultimate aim is to design a low pass filter.

$$H_p(\omega) = \begin{cases} 1 & \text{if } |\omega| < \omega_p \\ \mathrm{f}(\omega) & \text{if } \omega_p < |\omega| < \pi - \omega_p \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

We need only $f(\omega)$ and ω_p to define $H_p(\omega)$. $f(\omega)$ has the same mirror symmetry as does $H_p(\omega)$, so as to maintain the QMF property (2). Since $H_p(\omega)$ is a low pass filter and should be as smooth as possible (to facilitate FIR implementation) we can place the following requirements on $f(\omega)$:

$$f'(\omega_p) = f'(\pi - \omega_p) = 0 \tag{5}$$

and

1. $f(\omega)$ is strictly monotonic decreasing.

2. $f^{'}(\omega)$ is strictly monotonic decreasing on $[\omega_p, \frac{\pi}{2}]$.

We can show that the $f(\omega)$ that yields the fastest converging Fourier series expansion of $H_p(\omega)$ and thus a $H_p(\omega)$ that requires the fewest taps for FIR approximation is given by:

$$f_{opt}(\omega) = (1 - W)^2(2W + 1) = 2W^3 - 3W^2 + 1 \tag{6}$$

where W is given by:

$$W = \frac{\omega - \omega_p}{\pi - 2\omega_p} \tag{7}$$

and hence :

$$\sqrt{f_{opt}(\omega)} = (1 - W)\sqrt{(2W + 1)} \tag{8}$$

The efficient class of QMFs $H_{eff}(\omega)$ is given by:

$$H_{eff}(\omega) = \begin{cases} 1 & \text{if } |\omega| < \omega_p \\ \sqrt{f_{opt}(\omega)} & \text{if } \omega_p < |\omega| < \pi - \omega_p \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

The impulse response $H_{eff}(z)$ is given by:

$$H_{eff}(z) = \sum_{n=-\infty}^{\infty} h_{eff}(n) z^{-n} \tag{10}$$

$$h_{eff}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} H_{eff}(\omega) cos(n\omega)\, d\omega \tag{11}$$

The function $\sqrt{f_{opt}(\omega)}$ can easily be expanded in a Taylor series. In practice, five Taylor series terms should suffice. We can then get closed form expressions for the impulse response $h_{eff}(n)$ as well.

3 Comparison with existing designs: Numerical and Analytical results

For a given number N of filter taps, we would like to achieve as high frequency selectivity, and hence as high ω_p as possible. However, increasing ω_p beyond a certain upper limit compromises the reconstruction fidelity.

We have a simple rule of thumb see [4, 3, 5, 2]) to calculate the maximum feasible ω_p for a given number N of filter taps - $\omega_{p,max}(N)$. Let for even N, $N = 2n$, and for odd N, $N = 2n + 1$. Then we can plot the $(n+1)^{th}$ and $(n+2)^{th}$ coefficients, i.e. $h(n+1)$ and $h(n+2)$ from (11), versus ω_p. It is reasonable to assume that as long as these two coefficients are negligibly low in magnitude, the $N = 2n + 1$ tap FIR approximation will be accurate, given of course that the coefficients of order $k > n+2$ are also negligible. As ω_p increases, these coefficients assume more and more significance because the frequency selectivity of the filter goes up. The increase in the magnitude of these coefficients is not monotonic but tends to go through a series of maxima and minima. We illustrate this in figure 1 in which we plot the fifth and sixth coefficients versus the passband width ω_p, to see how we do with 9 taps. For values of $\omega_p$that lie between the intersections of these plots with the ω_p axis, the magnitudes of these coefficients are negligible. At such values, the FIR approximation will hence be accurate, if the magnitudes of the coefficients of order

$k > 6$ are also negligible. This implies that the maximum feasible passband edge frequency $\omega_{p,max}(N)$ lies between one of the closely spaced pairs of intersections of the plots with the ω_p axis. Thus we only need to choose the pair of intersections which gives the highest value for the $\omega_{p,max}(N)$. In this case $N = 9$. As shown, we reject the first intersection pair because we get a value of $\omega_{p,max}(9)$ that is less than what we could get with just five taps. The second pair of intersections gives a value that improves upon the passband width of the five tap filter. We stop at this pair because subsequent intersections are at values of ω_p which are so high that the coefficients of order greater than 6 cease to be negligible. This turns out to be a recurring pattern. We can thus generalize this procedure for $N = 2n + 1$ taps and say that the maximum feasible passband edge frequency $\omega_{p,max}(N)$ will lie between the first non-trivial intersection of the amplitude of the $(n+1)^{th}$ coefficient with the ω_p axis and that of the $(n+2)^{th}$ coefficient with the $\omega_p$axis. We define a *non-trivial intersection* as an intersection ω_{int} such that ω_{int} is greater than the maximum feasible passband edge frequency for $N - 2$ taps. i.e.

$$\omega_{int} \geq \omega_{p,max}(N-2)$$

In this region, both coefficients have a suitably low magnitude. Only the first non-trivial intersections will yield correct values of $\omega_{p,max}(N)$ because the magnitude of the higher order coefficients will not be negligible at subsequent intersections. We can then state the following rule of thumb:

Rule of Thumb 3.1 *The maximum feasible passband edge frequency $\omega_{p,max}(N)$ lies between the first non-trivial intersections of the $(n+1)^{th}$ and the $(n+2)^{th}$ coefficients with the ω_p axis, and can be estimated from the plots, where $N = 2n + 1$ or $N = 2n$.*

Application of the rule of thumb shows that the convenient positioning of the intersections illustrated in figure 1 occurs only for every other pair $n + 1$ and $n + 2$ of coefficients. This is because the even and odd coefficient intersections with the ω_p axis are "staggered" in such a way that only every other pair of consecutive coefficients has conveniently located intersections with the ω_p axis. For instance, the intersections of the fourth and fifth coefficients are spaced so wide apart that the higher order coefficients are no longer negligible at the estimated $\omega_{p,max}(N)$ got from the rule of thumb. At the intersection of the fourth coefficient and the $\omega_p$axis , the fifth coefficient is not negligible. Hence, the $\omega_{p,max}(N)$ is unchanged, i.e. the seven tap filter is not able to improve upon the frequency selectivity of the five tap filter in spite of the increased taps. An analogous argument holds for even tap implementation. We can sum up this result [4, 3, 5, 2] as follows:

$$\omega_{p,max}(2n+1) = \omega_{p,max}(2n+3) \tag{12}$$

$$\omega_{p,max}(2n) = \omega_{p,max}(2n+2) \tag{13}$$

for even n.

For instance, $\omega_{p,max}(N)$ is the same for 5 tap and 7 tap filters, which implies that there is no gain in frequency selectivity in spite of the increased taps. Equations (12) and (13) imply that for even n, the frequency selectivity of a $2n + 1$ or $2n$ tap filter can be improved if and only if the number of taps is increased by at least 4. Existing designs verify these results.

Most existing QMFs closely match the analytical expression derived here when we substitute $\omega_p =$ effective $\omega_p \approx \omega_{p,max}(N)$, where $\omega_{p,max}(N)$ is obtained, in radians, as described earlier and N is the number of taps. Recall that the passband width ω_p is the only variable parameter in the analytical expression. We illustrate this in figure 1, in which we plot the frequency responses of Johnston's 32 tap C filter and the rectangular window implementation, as well as the analytical expression (9) with $\omega_p = 1.32 \approx \omega_{p,max}(32) = 1.31$. We have also developed a new implementation technique that uses the analytical expression to obtain designs that match existing designs in performance (see [5, 2, 4]).

Having established the relationship between achievable frequency selectivity and number of filter taps, we can now investigate the information theoretic consequences of using fewer filter taps by studying the effects of reducing frequency selectivity.

4 Information Theoretic Performance of QMFs

Having derived expressions for the class of efficient QMFs and Perfect Reconstruction filters (see [2]), we can now study the information theoretic consequences of varying the passband width ω_p from the brickwall value $\frac{\pi}{2}$ to zero, which are:

1. Loss in the entropy reduction.
2. A Rate-Distortion theoretic penalty.

Nanda and Pearlman [12] showed that $D(R)$ can be achieved by subband coding using brickwall filters. Fischer has shown that using realizable i.e. non-brickwall QMFs causes a distortion-rate theoretic penalty. Rao and Pearlman have shown that scalar coding of ideally (i.e. brickwall) filtered subbands gives gain over scalar coding of fullband. There is a reduction in the aforementioned gain when we use realizable QMFs because they are less frequency selective than are brickwall filters. Therefore we need to find out to what extent the loss incurred by using realizable QMFs counteracts the scalar coding rate gain offered by subband coding, and how it varies with respect to the frequency selectivity of the QMF. We do so by varying the passband width ω_p from the brickwall value of $\frac{\pi}{2}$ to zero. Note that the $D(R)$ penalty applies to the theoretical and difficult to implement case, while the scalar coding gain applies to practical coders.

Note that since only the squared magnitude of the filter frequency response is used in information theoretic calculations, the filter's linearity of phase, or lack thereof, is immaterial. As we saw in our previous work [2]), our formula provides an excellent approximation for linear phase QMFs and a fair one for non-linear phase QMFs.

5 First Order Entropy of Subband Pyramids

5.1 The Spectral Flatness Measure and the Spectral Roughness Measure

Consider a discrete time, stationary process X_n, with power spectral density $S_X(\omega)$. The spectral flatness measure of a source with power spectral density $S_X(\omega)$ is defined as:

$$\gamma_x^2 = \frac{exp\left[\frac{1}{2\pi}\int_{-\pi}^{+\pi}\log S_x(\omega)d\omega\right]}{\sigma_x^2} \tag{14}$$

We use natural logarithms throughout. The variance of the aforementioned source is:

$$\sigma_x^2 = \frac{1}{2\pi}\int_{-\pi}^{+\pi} S_X(\omega)d\omega \tag{15}$$

Rao and Pearlman [15] define a quantity called the Spectral Roughness Measure (SRM) Δ_∞^x as the difference between the first order entropy of the source and its entropy rate i.e.

$$\Delta_\infty^x = h_1(X) - h_\infty(X) \tag{16}$$

Furthermore, for small distortion D,

$$\Delta_\infty^x = R_1(D) - R(D) \tag{17}$$

where $R(D)$ and $R_1(D)$ are the rate-distortion functions of the source and a memoryless source with the same marginal probability density respectively. We express rate in nats throughout. If X is Gaussian, then

$$h_1(X) = \frac{1}{2}\log(2\pi e\sigma_x^2) \tag{18}$$

$$h_\infty(X) = \frac{1}{2}\log(2\pi e\gamma_x^2\sigma_x^2) \tag{19}$$

Which from (14) and (16) shows that for Gaussian sources, Δ_∞^x is related to the spectral flatness measure as follows:

$$\Delta_\infty^x = -\frac{1}{2}\log(\gamma_x^2) \tag{20}$$

It is evident that the spectral roughness measure increases as the spectral flatness measure decreases and vice versa. Hence the name "roughness measure" has some justification. It is a measure of the memory of the source. For i.i.d. sources γ_x^2 is equal to 1 and hence Δ_∞^x is equal to 0. For sources with memory, the spectrum is no longer smooth. As γ_x^2 approaches zero, Δ_∞^x approaches infinity.

5.2 Theorem [15]

The spectral roughness measure (SRM), Δ_∞, for an ideally filtered multiresolution pyramid representation of a Gaussian source is no greater than the Δ_∞ for fullband. i.e.

$$\Delta_\infty^{Fullband} - \Delta_\infty^{Subband} \geq 0 \tag{21}$$

Furthermore, the difference between the two increases as the number of levels in the pyramid is increased

5.3 Discussion

Equation (17) implies that the difference of the SRM of the fullband and the combined SRM of the subbands $\Delta\mathcal{H} = \Delta_\infty^{Fullband} - \Delta_\infty^{Subband}$ is the rate difference between optimal scalar coding of the fullband and the subbands. Hence the above result implies that we need fewer bits to scalar quantize the pyramid as compared to the original source if we use the same quantizer. We can explain this from a frequency domain point of view as well, because the bandpass filtering followed by subsampling stretches the spectrum in each subband. This implies that each subband has less inter-sample correlation and hence can be more efficiently scalar quantized than can the fullband.

Furthermore, it has been shown in [13] that:

$$\Delta\mathcal{H} = \Delta_\infty^{Fullband} - \Delta_\infty^{Subband} = \frac{1}{2}\log G_{SB/PCM}, \tag{22}$$

where [10]

$$G_{SB/PCM} = \frac{\frac{1}{M}\sum_{k=1}^{M}\sigma_{xk}^2}{\left[\prod_{k=1}^{M}\sigma_{xk}^2\right]^{\frac{1}{M}}} \tag{23}$$

is the well-known coding gain for PCM coding of M equal width subbands, with σ_{xk}^2 = the variance of the kth subband. This coding gain is defined to be the ratio between distortions of fullband and subband PCM coding. Equation (22) shows that the distortion gain ratio of subband PCM is a consequence of the reduction of the difference between first order entropy and entropy rate from that of the full band source.

5.4 Results on AR(1) sources

We investigate the behavior of the combined first order entropy of auto-regressive first order Markov (AR(1)) sources, with a two band split. We consider AR(1) sources because they provide a simple and powerful approximate model for natural signals such as images and speech. While natural signals cannot be perfectly modeled by such simple models, the AR(1) model is a good analytical tool for first order prediction statistics. Note that the spectral roughness measure is also a first order entropy measure. Since the subband pyramid is built by recursively splitting the lower band, the first split should establish a trend which the subsequent splits should follow.

The output $X(n)$ of an AR(1) source is given by [10]:

$$X(n) = \rho X(n-1) + Z(n) \tag{24}$$

Z is a zero-mean, white Gaussian noise process with $\sigma_n^2 = 1$.

Let the source spectral density be $S_X(\omega)$ and the filter characteristic be $H(\omega)$. The source spectral density of an AR(1) source is given by:

$$S_X(\omega) = \frac{1-\rho^2}{1+\rho^2 - 2\rho cos(\omega)}\sigma_x^2 \tag{25}$$

and its spectral flatness measure is given by:

$$\gamma^2 = 1 - \rho^2 \tag{26}$$

Let the lower subband spectral density be $S_l(\omega)$ and the upper subband spectral density be $S_u(\omega)$. Then taking into account the aliasing and the decimation by a factor of two we get:

$$S_l(\omega) = \frac{1}{2}\left[|H(\frac{\omega}{2})|^2 S_X(\frac{\omega}{2}) + |H(\frac{\omega}{2}+\pi)|^2 S_X(\frac{\omega}{2}+\pi)\right] \tag{27}$$

$$S_u(\omega) = \frac{1}{2}\left[|H(\frac{\omega}{2}+\pi)|^2 S_X(\frac{\omega}{2}) + |H(\frac{\omega}{2})|^2 S_X(\frac{\omega}{2}+\pi)\right] \tag{28}$$

The variance σ_x^2 of the source is related to the variances of the subbands as follows (see [7]):

$$\sigma_x^2 = \sigma_l^2 + \sigma_u^2 \tag{29}$$

Each subband is separately encoded and therefore the above expressions for the subband spectra can be substituted into the definition of the SRM Δ_∞, so as to obtain a valid expression for the combined SRM. Recall that the efficient class of QMFs is given by (8), (9) and (10). From (20), the SRM of the lower subband is given by

$$\Delta_\infty^l = -\frac{1}{2}\log \gamma_l^2 \tag{30}$$

γ_l is obtained by substituting $S_l(\omega)$ in (14).

We can similarly get expressions for the SRM Δ_∞^u of the upper subband. Since the subbands are encoded independently, we can add the SRM's of the two subbands to get the combined SRM i.e.

$$\Delta_\infty^{Subband} = \frac{1}{2}\left[\Delta_\infty^l + \Delta_\infty^u\right] \tag{31}$$

The weighting factor $\frac{1}{2}$ is introduced because of the decimation by a factor of two.

The combined SRM for a Gaussian source is therefore:

$$\Delta_{\infty}^{Subband} = -\frac{1}{2}\log\sqrt{\gamma_l^2\gamma_u^2} \tag{32}$$

In Figure 4, we plot the rate difference $\Delta\mathcal{H}$ i.e. the difference of the SRM of the fullband and the combined SRM of the subbands versus the passband edge frequency ω_p. Note that the maximum rate difference is achieved by brickwall filters i.e. when $\omega_p = \frac{\pi}{2}$. Note that this rate difference is very high for highly correlated sources but reduces as the source becomes less correlated, which is as expected.

When we reduce ω_p from its brickwall value, we find that there is a loss in the rate difference, i.e. a loss in first order entropy reduction. However, the loss is quite low even when we reduce ω_p to zero, for instance it is 3.3% for $\rho = 0.9$. Even at its peak it is about 3.8%. Furthermore, the loss is more pronounced when the source is highly correlated but goes down as ρ is reduced. The value of ω_p achieved by practical 9 tap filter designs, for instance, [16, 2, 5] is 0.895 radians, which is approximately equal to $0.57\frac{\pi}{2}$. Our rule of thumb for frequency selectivity estimates the maximum achievable ω_p with 9 taps to be 0.88 radians which is equal to $0.56\frac{\pi}{2}$, which is evidently close to the ω_p achieved in practice. Even with as few as 5 filter taps, when the effective ω_p is 0.48 radians, i.e. about $0.3\frac{\pi}{2}$, the loss is very low. Bearing this in mind, let us look once again at the plots. We find that rate difference decreases monotonically but negligibly across the range of values of ω_p.

Furthermore, the analytical formula yields results which are approximately the same as those obtained with existing filter designs. Let us take Simoncelli's 9 tap filter for example. It matches the analytical formula quite closely, but to a much lesser extent than do other existing designs. The value of the combined SRM with Simoncelli's 9 tap filter is 0.32717 nats, while the formula (with ω_p=0.895) yields 0.32715, i.e. a difference in the fifth significant place. With other designs the formula gives even greater accuracy. This is not very surprising, because the squared magnitude $|H_{eff}(\omega)|^2$ of the filter response, which is used for information theoretic purposes, has double the attenuation in the stop band compared to the QMF frequency response $H_{eff}(\omega)$. Hence, the squared expression is, for practical purposes, almost the same as the generic low pass filter we developed earlier, that had infinite attenuation in the stop band. Thus, our formula is very useful, in spite of the seemingly naive assumption of infinite attenuation in the stop band.

5.5 Results on AR(2) Sources

The AR(2) process of zero mean is given by [10]

$$X(n) = b_1X(n-1) + b_2X(n-2) + Z(n) \tag{33}$$

with Z a white noise process as defined earlier. The spectral flatness measure of an AR(2) source is given by:

$$\gamma^2 = \frac{(1+b_2)(1-b_1-b_2)(1+b_1-b_2)}{(1-b_2)} \tag{34}$$

We now carry out the rate difference calculations for a two band split for AR(2) sources, so as to extend our results to higher-order sources. Moreover, AR(2) sources provide a good fit to the long-time-averaged second-order statistics of speech. The source spectral density of an AR(2) source is given by:

$$S_X(\omega) = \frac{1}{1+b_1^2+b_2^2-2b_1(1-b_2)cos(\omega)-2b_2cos(2\omega)}\sigma_z^2 \tag{35}$$

where σ_z^2 is the variance of the input white noise process used to generate the AR(2) process. We can substitute this expression in the expressions for the combined SRM we obtained in the previous section, to get the combined SRM. Note that now we have two parameters b_1 and b_2. Unlike with an AR(1) process, it is now possible to produce spectra with a peak or trough at an intermediate angular frequency. Equation (35) implies that the power spectral density has an extremum at $\omega = arccos\,[b_1(b_2-1)/(4b_2)]$ For negative b_2 [10], the extremum is a peak and is at $\omega = \frac{\pi}{2}$, if $b_1 = 0$. As b_1 increases, the peak shifts away from $\omega = \frac{\pi}{2}$ towards $\omega = 0$, so that for higher values of b_1 the AR(2) source spectrum begins to resemble the AR(1) spectrum more and more closely. Since we have already obtained results on AR(1) sources, it is reasonable to begin with an AR(2) source spectrum that is least like that of an AR(1) source and perturb it until it begins to resemble a typical AR(1) spectrum. We do so by fixing $b_2 = -0.9$ and varying b_1 from zero to higher values. We choose a high value of b_2 so as to keep the bandwidth narrow and thus maximize the departure from the AR(1) spectrum.

In Figure 5, we plot the rate difference $\Delta\mathcal{H}$. Note that in this case, there is a sharp drop in rate gain as we reduce ω_p from its brickwall value of $\frac{\pi}{2}$. For very low values of b_1, even a slight deviation from the brickwall value leads to a high percentage drop in rate gain. For $b_1 > .4$, there is less than 10% loss in rate gain if we go down to $\omega_p = 0.9\frac{\pi}{2}$, and the loss reduces monotonically as we increase b_1. Note that $\omega_p = 0.9\frac{\pi}{2}$ is achievable with 64 taps, as per our rule of thumb.

Next we apply Fischer's analysis to investigate rate distortion theoretic losses incurred by non-brickwall filters.

6 Rate Distortion Performance of Feasible QMFs

6.1 Fischer's Result

Let us assume that $x(n)$ is real-valued, zero mean, wide sense stationary (wss), and Gaussian. For small distortion, the full band distortion-rate function is

$$D_x(R) = \sigma_x^2 \gamma_x^2 e^{-2R} \tag{36}$$

Let us split the signal into two subbands as described earlier. Fischer has shown that the subband coding distortion-rate function of $x(n)$ is given by:

$$\begin{aligned} D(R) &= e^{-2R} exp\left\{\frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} ln\left[\Delta(\omega) + S_x(\omega)S_x(\omega+\pi)\right] d\omega\right\} \\ &\geq e^{-2R} exp\left[\frac{1}{2\pi}\int_{-\pi/2}^{+\pi/2} ln[S_x(\omega)S_x(\omega+\pi)]d\omega\right] = D_x(R) \end{aligned} \tag{37}$$

where

$$\Delta(\omega) = |H(\omega)|^2 |H(\omega+\pi)|^2 [S_X(\omega) - S_X(\omega+\pi)]^2 \tag{38}$$

and $H(\omega)$ is a low pass filter response that satisfies the power complementarity condition. Furthermore, the inequality in (37) is strict if $\Delta(\omega) > 0$ on a subset of $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

Equations (36) and (37) imply that the encoding performance with subband coding is generally inferior to the rate distortion function of the source (fullband). In other words, subband coding generally incurs a rate distortion theoretic penalty ΔR. From (37) and (38) we get the fractional distortion increase

$$\eta = \frac{D(R)}{D_x(R)} \geq 1 \tag{39}$$

with equality when $\Delta(\omega) = 0$, which occurs when brickwall filters are used or if $S_X(\omega)$ is symmetric about $\omega = \frac{\pi}{2}$. Thus, the rate penalty corresponding to a given value of η is

$$\Delta R = \frac{1}{2} log(\eta) \tag{40}$$

The above result (i.e. equations (37) and (38)) provides a straightforward way of determining the rate distortion theoretic penalty incurred by the usage of realizable filters. The above expression is very easy to use since we now have a simple expression for $H(\omega)$. Fischer has suggested direct substitution of practical filter tap values to obtain $H(\omega)$. Clearly, this procedure will be very tedious for filters that have a large number of taps. Our expression for the efficient class circumvents this problem and provides a simple way to evaluate $\Delta(\omega)$. Note that the expression for $\Delta(\omega)$ will then depend only on ω_p.

6.2 Results on AR(1) sources

We can now get an expression for ΔR that has only one variable parameter viz. the passband width ω_p and plot ΔR versus ω_p. We illustrate our results in Figure 6. We can see that the rate distortion penalty is zero for brickwall filters, but rises slowly as we reduce the passband width ω_p from the brickwall value of $\frac{\pi}{2}$, i.e. we reduce ω_p to a value less than $\frac{\pi}{4}$. As ω_p falls below half of the brickwall value the penalty begins to rise at a faster rate, and is quite significant for $\omega_p = 0$. In other words, the rate distortion penalty rises monotonically as we reduce ω_p from its brickwall value of $\frac{\pi}{2}$ to zero, but the rate of increase becomes significantly high only after ω_p is reduced below $\frac{\pi}{4}(= 0.5\frac{\pi}{2})$. Also, as expected, the rate distortion penalty is highest for the highest value of the correlation coefficient ρ for a given value of ω_p. But the rate distortion penalty for even moderately correlated sources is significant for $\omega_p \leq \frac{\pi}{4}$.

We find that the rate distortion theoretic penalty is not very high. In fact, at $\omega_p = 0.57\frac{\pi}{2}$ i.e the value of $\omega_p$for Simoncelli's 9 tap filter, the ΔR for $\rho = 0.9$ is merely 0.00366 nats. This implies that for values of ω_p that we can achieve with as few as 9 taps, the rate distortion theoretic penalty is extremely low. This turns out to be true for as few as five taps as well.

Furthermore, the analytical formula yields results which are approximately the same as those obtained with existing filter designs. Let us take Simoncelli's 9 tap filter for example. It matches the analytical formula quite closely, but to a much lesser extent than do other existing designs. The value of η with Simoncelli's 9 tap filter is 1.010271, while the formula (with ω_p=0.895) yields 1.007362, i.e. a difference in the third significant place. This is equivalent to an error of 0.288%, which is very low. With other designs the formula gives even greater accuracy which, as we observed in the previous section, is as expected.

6.3 Results on AR(2) Sources

We now calculate the ratio ΔR for AR(2) sources. We vary the parameters b_1 and b_2 in exactly the same way as we did for the first order entropy calculations. If we set $b_1 = 0$, the resulting source spectrum is symmetric about $\omega = \frac{\pi}{2}$, which results in a zero rate distortion theoretic penalty irrespective of the frequency selectivity of the QMF. Hence, as we increase the value of b_1 we in fact perturb the symmetry about $\frac{\pi}{2}$ and thus make the rate distortion theoretic penalty non-zero for realizable QMFs. Note that at first the energy of the signal is almost entirely in the vicinity of $\omega = \frac{\pi}{2}$, but as the value of b_1 increases, the energy in the vicinity of $\omega = \frac{\pi}{2}$ reduces until for very high values of b_1 the low energy in the vicinity of $\omega = \frac{\pi}{2}$ has a much more significant effect than does the lack of symmetry of the source spectrum about $\omega = \frac{\pi}{2}$.

We illustrate our results in Figure 7. Note that for $\omega_p = 0.9\frac{\pi}{2}$ the rate distortion penalty is very low . For lower values however, the penalty goes up steadily. The "crossover" in the curves is due to the reduction in energy in the vicinity of $\omega = \frac{\pi}{2}$, as b_1 is increased. Hence for higher values of b_1 there is actually less of a loss at $\omega_p = 0.9\frac{\pi}{2}$ than there is for certain lower values of b_1. However it is clear that even at $\omega_p = 0.8\frac{\pi}{2}$ the loss is high. As per our rule of thumb, the minimum number of filter taps needed to achieve $\omega_p = 0.9\frac{\pi}{2}$ is 64. Note that the first order entropy results indicate a slightly greater loss than does the rate-distortion theoretic calculation which is not unexpected.

7 Recapitulation

Let us sum up this paper. We reviewed Rao and Pearlman's result that states that the combined first order entropy of an optimally filtered pyramid is less than is the first order entropy of the full band process. We also recalled Fischer's result on subband coding viz. with realizable filters, the subband coding of a wide sense stationary Gaussian source is rate distortion theoretically sub-optimal.

Rao and Pearlman assumed brickwall filtering. However, the combined formulas are valid when the subbands are coded independently even when we use realizable QMFs. Fischer's analysis assumed a generic QMF. Since we have closed form expressions for an efficient class of QMFs, we were able to look into entropy reduction properties of subband pyramids formed with realizable QMFs, as well as the rate distortion penalty incurred by using realizable (non-brickwall) QMFs. In both cases, we used auto-regressive Markov (AR(1) and AR(2)) sources as test sources, since such sources provide a simple and powerful model for natural signals such as images and speech. Our closed form expressions allowed us to express both the entropy reduction and the rate distortion in terms of just one parameter viz. the passband width of the QMF ω_p.

This allowed us to study the variation in information theoretic performance as the QMF is made less frequency selective i.e. ω_p is reduced from its brickwall value of $\frac{\pi}{2}$ to 0. Good information theoretic performance implies high entropy reduction as well as a low rate distortion theoretic penalty. Our formula closely approximates existing designs, and therefore matches information theoretic results obtained from practical filters. We found that the information theoretic performance with AR(1) sources does not deteriorate significantly for feasible values of ω_p. Such values of ω_p are attainable with as few as 5 taps.

Since AR(1) sources have a small fraction of their signal energy in the vicinity of $\omega = \frac{\pi}{2}$, we would expect the information theoretic penalty to be low. With AR(2) sources that have the bulk of their energy in the vicinity of $\omega = \frac{\pi}{2}$, realizable QMFs incur a high information theoretic penalty. Such sources are narrowband and thus do not lend themselves well to two band decompositions. For a large category of AR(2) sources, i.e. those that have a smaller fraction of total signal energy in the vicinity of $\omega = \frac{\pi}{2}$, the information theoretic penalty is negligibly low as we go down to $\omega_p = 0.9\frac{\pi}{2}$. Such a value of ω_p is achievable with 64 filter taps, as per our rule of thumb.

Our results suggest therefore that with sources such as images, which have a lowpass monotonically decreasing frequency spectrum, we can use as few as 5 filter taps without incurring a significant information theoretic penalty. Practical image coding results using short filters such as those in [1] support such an inference. Speech spectra [8] on the other hand are rarely monotonic decreasing, but typically do not have both a narrow bandwidth and the bulk of signal energy in the vicinity of $\omega = \frac{\pi}{2}$, either with reference to the original sampling frequency or to decimated sampling frequencies resulting from recursive two-band or M-band splittings. (Moreover, decimation tends to flatten or broaden continuous subband spectra as a function of normalized frequency $\omega, -\pi < \omega \leq \pi$.) Since AR(2) sources are known to fit speech sources well, our results suggest that with 64 filter taps or more, the information theoretic penalty with speech will not be significant. Unlike AR(2) and AR(1) sources, auto-regressive Gaussian sources of order higher than 2 can have multiple peaks and troughs. However we can see that our conclusions regarding the relationship between information theoretic loss and the fraction of signal energy in the vicinity of $\omega = \frac{\pi}{2}$ will still be largely applicable to such sources because we can think of roughly approximating them as sums of AR(2) processes. The only way to ensure negligible information theoretic loss with all possible sources is to use very long QMFs that achieve $\omega_p >= 0.99\frac{\pi}{2}$.

8 Acknowledgement

We would like to thank one of our anonymous reviewers for suggesting the perturbation strategy that we have used for the results on AR(2) sources, as well as for the valuable comments on the implications of signal energy in the vicinity of $\omega = \frac{\pi}{2}$.

REFERENCES

[1] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image Coding using Wavelet Transform," *IEEE Trans. on Image Processing,* Vol. 1, pp. 205–220, April 1992.

[2] Ajay Divakaran. *Quadrature Mirror Filters: FIR Implementation, Wavelet Analysis and Information Theoretic Analysis.* PhD thesis, Electrical, Computer and Systems Engineering Dept., Rensselaer Polytechnic Institute, August 1993.

[3] A. Divakaran and W. A. Pearlman, "A new approach to quadrature mirror filter design," in *Proceedings of 1992 Conf. on Information Science and Systems*, pp. 324–329, Princeton, NJ, March 1992.

[4] A. Divakaran and W. A. Pearlman, "A closed form expression for an efficient class of quadrature mirror filters and its FIR approximation," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, Vol. 43, pp. 207–219, Mar. 1996.*

[5] A. Divakaran and W. A. Pearlman, "A new quadrature mirror filter design technique," in *Proc. Conf. on Information Science and Systems*, Baltimore, MD, March 1993.

[6] D. Esteban and C. Galand, "Application of quadrature mirror filters to split band voice coding schemes," in *Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP)*, pp. 191–195, 1977.

[7] T. R. Fischer, "On the rate-distortion efficiency of subband coding," *IEEE Trans. on Information Theory*, Vol. 38, pp. 426 – 428, March 1992.

[8] J. L. Flanagan, M. R. Schroeder, B. S. Atal, R. E. Crochiere, N. S. Jayant, and J. M. Tribolet. "Speech coding," *IEEE Trans. on Communications*, pp. 710–737, May 1979.

[9] V. K. Jain and R. E. Crochiere, "Quadrature mirror filter design in the time domain," *IEEE Trans. on Acoust., Speech and Signal Processing*, Vol. ASSP-32, pp. 353–360, April 1984.

[10] N. S. Jayant and P. Noll, *Digital Coding of Waveforms,* Prentice-Hall, Englewood Cliffs, NJ, 1984.

[11] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in *Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP)*, pp. 291–294, 1980.

[12] S. Nanda and W. A. Pearlman, "Tree coding of image subbands," *IEEE Trans. on Image Processing,* Vol. 1, pp. 133–147, April 1992.

[13] W. A. Pearlman, "Performance bounds for subband coding," Chap. 1 in J. W. Woods, editor, *Subband Image Coding.* Kluwer Academic Publishers, 1991.

[14] G. Pirani and V. Zingarelli, "An analytical formula for the design of quadrature mirror filters," *IEEE Trans. on Acoust., Speech and Signal Processing* , Vol. ASSP-32, pp. 645–648, June 1984.

[15] R. P. Rao and W. A. Pearlman, "On entropy of pyramid structures," *IEEE Trans. on Information Theory*, Vol. IT-37, pp. 407–413, March 1991.

[16] E. P. Simoncelli and E. H. Adelson, "Subband transforms," in J. W. Woods, editor, *Subband Image Coding.* Kluwer Academic Publishers, 1991.

List of Figures

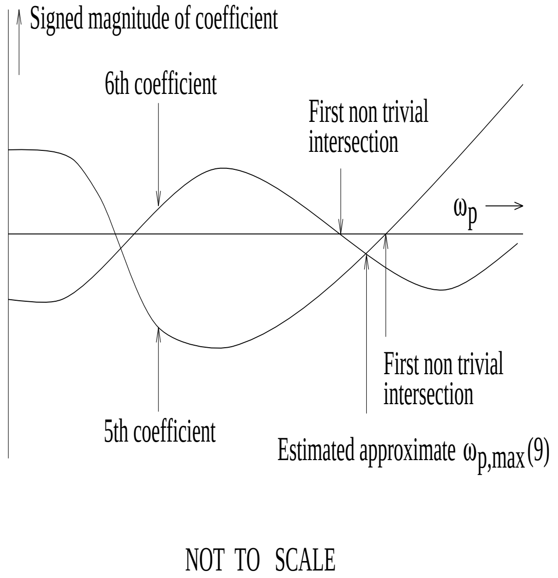

Figure 1: How to find $\omega_{p,max}(9)$

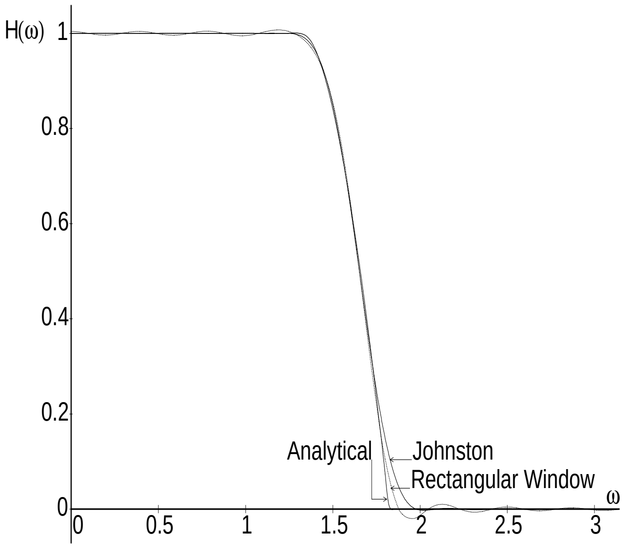

Figure 2: Comparison of Johnston's 32 tap "C" filter and the analytical expression with the Rectangular Window Implementation: $\omega_p = 1.32 \approx \omega_{p,max}(32) = 1.31$radians

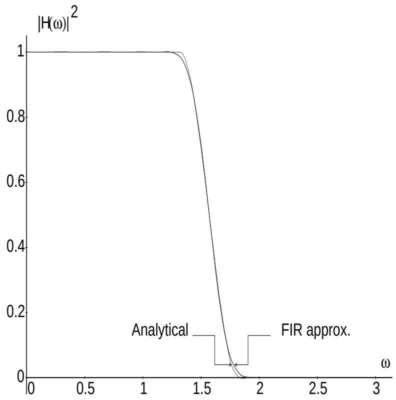

Figure 3: Comparison of the square of the analytical formula with the square of our 32 tap filter characteristic: $\omega_p = 1.32 radians \approx \omega_{p,max}(32) = 1.31 radians$

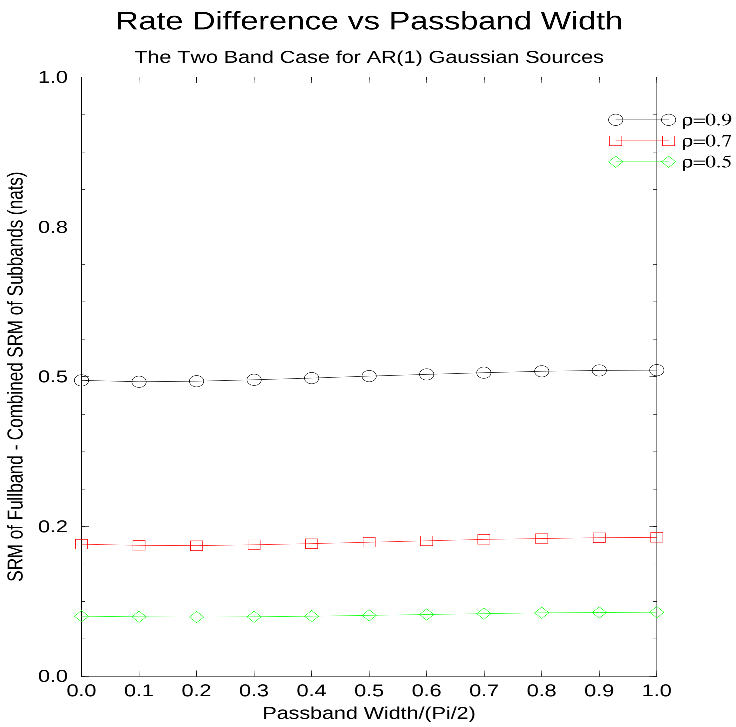

Figure 4: The variation of reduction in SRM with the passband edge frequency ω_p : The AR(1) case.

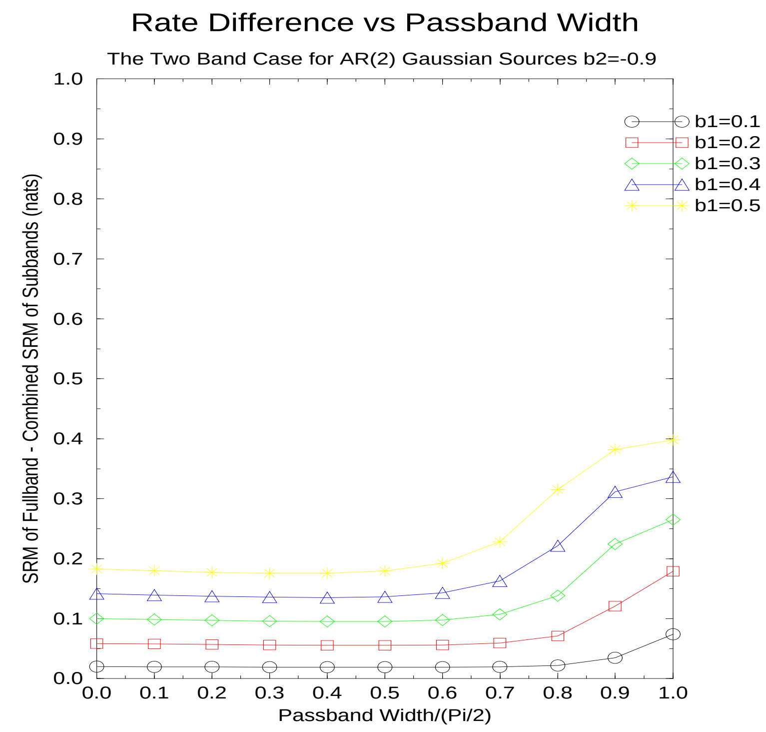

Figure 5: The variation of reduction in SRM with the passband edge frequency ω_p : The AR(2) case.

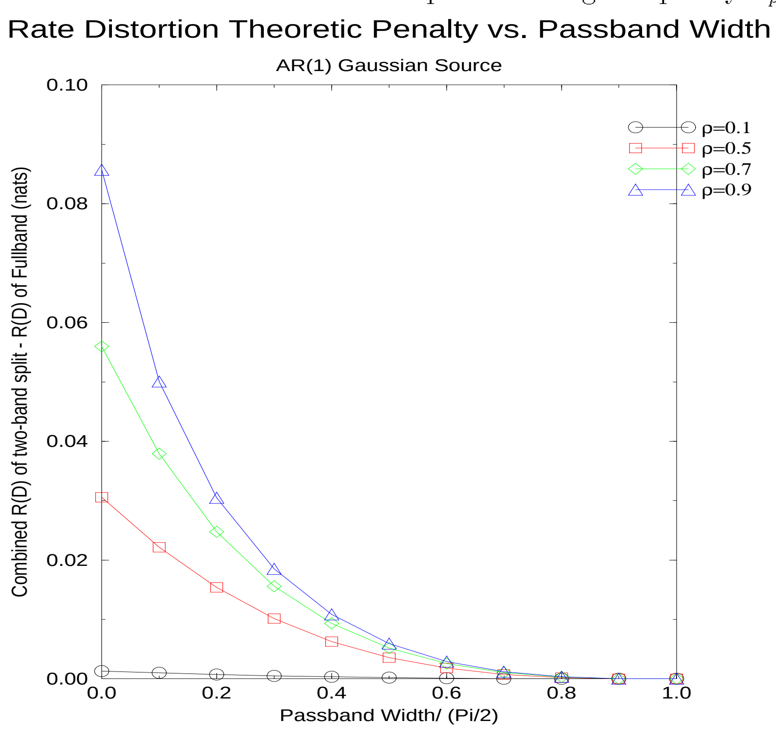

Figure 6: The variation of the Rate Distortion Penalty with the passband edge frequency ω_p : The AR(1) case.

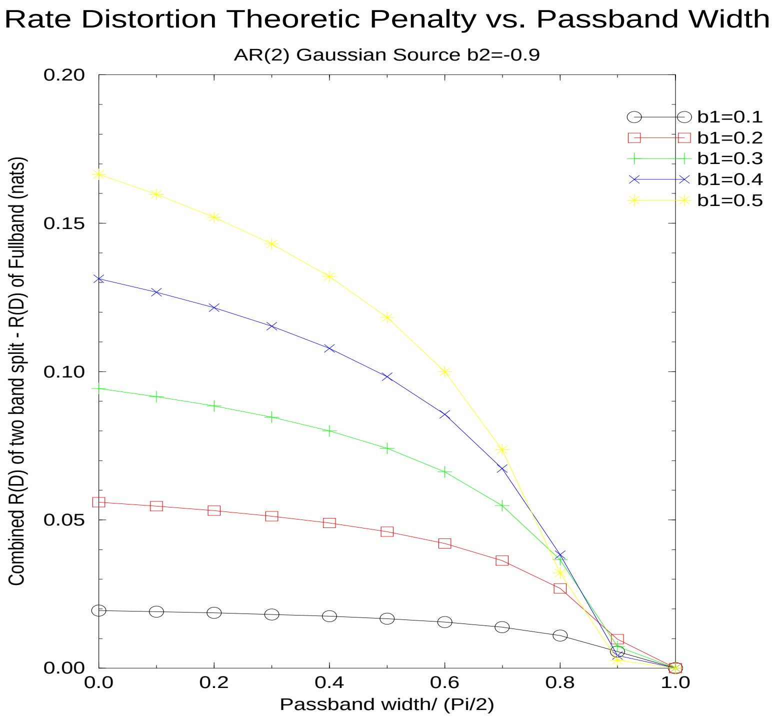

Figure 7: The variation of the Rate Distortion Penalty with the passband edge frequency ω_p: The AR(2) case.