HIGH RESOLUTION SPHERICAL QUANTIZATION OF SINUSOIDS WITH HARMONICALLY RELATED FREQUENCIES

Pim Korten, Jesper Jensen and Richard Heusdens

Information and Communication Theory Group, Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands
phone:+31 (0)15 27 82188, fax:+31 (0)15 27 81843
email: {p.e.l.korten, j.jensen, r.heusdens}@ewi.tudelft.nl

ABSTRACT

Sinusoidal coding is an essential tool in low-rate audio coding, and developing an efficient quantization scheme for the sinusoidal parameters is therefore crucial. In this work we derive optimal entropy constrained amplitude, phase and frequency quantizers for sinusoids whose frequencies are harmonically related, with respect to the ℓ_2 distortion measure. This scheme exploits the harmonic structure of many speech and audio signals in the sense that besides amplitudes and phases, only fundamental frequencies need to be quantized, resulting in a significant decrease in the number of bits assigned to frequency parameters. The asymptotically optimal quantizers minimize a high-resolution approximation of the expected ℓ_2 distortion while the corresponding quantization indices satisfy an entropy constraint. The quantizers turn out to be flexible and of low complexity, in the sense that they can be determined easily for varying bit rate requirements, without any sort of retraining or iterative procedures. In an objective rate-distortion comparison, the proposed scheme is shown to outperform two variants of a recently proposed scheme, in which all frequency parameters are quantized separately, either directly or differentially.

1. INTRODUCTION

Parametric coding is an often employed and efficient technique for representing audio signals at low bit rates [1, 2]. A typical procedure in parametric coding is to decompose an audio signal into three components: a sinusoidal component, a noise component and a transient component, each of which are coded by different subcoders. In most low-rate audio coders the main part of the bit budget will be assigned to the sinusoidal component [2], represented by amplitude, phase and frequency parameters. Typically, the total available bit rate is distributed over the different subcoders by a rate-distortion control algorithm, giving the bit budget available for encoding sinusoids a priori. Therefore we aim at developing simple and flexible quantizers which can adapt to changing bit rate requirements without any sort of retraining or iterative procedures. In this work, we focus on deriving such efficient quantizers for the sinusoidal parameters.

It is well-known that voiced speech signals [3] and many musical instruments [4] are approximately harmonically structured, i.e. they consist of sinusoidal components which are roughly located at integer multiples of a certain *fundamental frequency*. These sinusoidal components are also called *harmonics* or *partials*. Typically, speech is composed of one *layer* of harmonics, while in general audio may consist of several layers of harmonics, each layer having a different fundamental frequency depending on the underlying instrument. A speech or audio signal x is then modelled by

$$x(n) \approx \sum_{k}^{K} \sum_{l}^{L_k} a_{k,l} \cos\left(l\nu_k^0 n + \phi_{k,l}\right), \tag{1}$$

where K is the number of harmonic layers in the signal and L_k the number of harmonics in layer k. Furthermore, $a_{k,l}$, $\phi_{k,l}$ and ν_k^0 denote the amplitude, phase and fundamental frequency corresponding to harmonic $\{k, l\}$, respectively. Note that a model in which a signal is approximated by a sum of sinusoids whose frequencies are not necessarily harmonically related is included in (1), by choosing $L_k = 1$ for all layers, i.e. every component defines one layer. Several methods are developed for fundamental frequency estimation, e.g. [5, 6, 7]. Since we focus on quantization of the sinusoidal parameters in this work, it is assumed that the estimation of harmonics from the signal is already carried out by one of the existing algorithms. Hence the fundamental frequency and the number of harmonics in each layer, the amplitudes and phases for each harmonic, and the number of layers is assumed known.

The quantization scheme proposed in this work is called entropy constrained unrestricted[1] spherical quantization for harmonics (ECUSQ$_h$), and is a generalization of ECUSQ [10]. While the ECUSQ scheme quantizes all frequency parameters individually, the proposed ECUSQ$_h$ scheme only quantizes the fundamental frequencies; all the other quantized frequencies are then located at integer multiples of the quantized fundamental frequency. In this work we follow the procedure of [9, 10], and derive the optimal quantizers under high-resolution assumptions, i.e. a large number of quantization cells, which implies that the probability density functions of the input variables can assumed to be constant in each quantization cell.

In this work we derive scalar ECUSQ$_h$ quantizers, using the ℓ_2 error distortion measure. More specifically, under high-resolution assumptions, optimal amplitude, phase and fundamental frequency quantizers are derived which minimize the expected distortion, while satisfying an entropy constraint. This is done for multiple layers of harmonically related sinusoids. Furthermore, the distribution of the target entropy between the three parameter types is determined. Since besides amplitudes and phases, only fundamental frequencies are quantized in ECUSQ$_h$, the number of bits assigned to frequency is expected to be smaller than in ECUSQ. To see how this affects distortions, the rate distortion performances of the proposed scheme and two variants of the ECUSQ schemes are compared, using synthesized input data with imposed harmonic structure.

The research is supported by STW, applied science division of NWO and the technology programme of the Dutch ministry of Economic Affairs.

[1]The term unrestricted was introduced in [8] and refers to the fact that the sinusoidal parameters are quantized dependently.

2. ENTROPY CONSTRAINED UNRESTRICTED SPHERICAL QUANTIZATION OF HARMONIC SINUSOIDS

2.1. High-resolution expression for the expected distortion

In this section we derive a high-resolution approximation for the total expected ℓ_2 distortion. We assume the input signal consists of one segment with length N and is of the form (1). The generalization of the results in this work to multiple segments is straightforward. Since real sinusoids as in (1) can be expressed as a sum of complex exponentials, we consider quantization of complex exponentials. The quantization error signal for harmonic $\{k, l\}$ is then given by $\varepsilon_{k,l}(n) = a_{k,l}e^{j(l\nu_k^0 n+\phi_{k,l})} - \tilde{a}_{k,l}e^{j(l\tilde{\nu}_k^0 n+\tilde{\phi}_{k,l})}$. Let w denote an analysis window sequence, and let $\varepsilon_{k,l}^w$ denote the windowed error signal. The ℓ_2 error of $\varepsilon_{k,l}^w$ can be evaluated to

$$\|\varepsilon_{k,l}^w\|_2^2 = \|w\|^2(a_{k,l}^2 + \tilde{a}_{k,l}^2) - 2a_{k,l}\tilde{a}_{k,l} \times \sum_n w(n)^2 \cos\left(l(\nu_k^0 - \tilde{\nu}_k^0)n + \phi_{k,l} - \tilde{\phi}_{k,l}\right). \quad (2)$$

We assume that fundamental frequencies are sufficiently large such that the spectral overlap of harmonics within one layer is negligible. Furthermore, we assume that harmonics in different layers are statistically independent. In this case, it is valid to approximate the total expected distortion D by

$$D \approx \sum_{k=1}^{K}\sum_{l=1}^{L_k} E\left(\|\varepsilon_{k,l}^w\|_2^2\right), \quad (3)$$

where

$$E\left(\|\varepsilon_{k,l}^w\|_2^2\right) = \iiint f_{\boldsymbol{A}_k,\boldsymbol{\Phi}_k,\Upsilon_k^0}(\boldsymbol{a}_k, \boldsymbol{\phi}_k, \nu_k^0)\|\varepsilon_{k,l}^w\|_2^2 d\nu_k^0 d\boldsymbol{\phi}_k d\boldsymbol{a}_k. \quad (4)$$

Here $f_{\boldsymbol{A}_k,\boldsymbol{\Phi}_k,F_k^0}$ is the joint $(2L_k+1)$-dimensional probability density function of all random variables in layer k, consisting of amplitudes $\boldsymbol{A}_k \in \mathbb{R}^{L_k}$, phases $\boldsymbol{\Phi}_k \in \mathbb{R}^{L_k}$ and fundamental frequency $\Upsilon_k^0 \in \mathbb{R}$. We see that although $E\left(\|\varepsilon_{k,l}^w\|_2^2\right)$ defines the expected distortion for a single harmonic, the integration is performed over all variables within the corresponding layer, it will become clear later why this is necessary. As in [10], the integral in (4) can be evaluated by summing over all possible quantization cells, and using the fact that the probability density function can be considered constant in each cell, due to high-resolution assumptions. Substituting (2) in (4) and using Taylor expansions, we then obtain the following high-resolution approximation

$$E\left(\|\varepsilon_{k,l}^w\|_2^2\right) \approx \frac{\|w\|^2}{12}\iiint f_{\boldsymbol{A}_k,\boldsymbol{\Phi}_k,\Upsilon_k^0}(\boldsymbol{a}_k, \boldsymbol{\phi}_k, \nu_k^0) \times \left(g_{A_{k,l}}^{-2} + a_{k,l}^2\left(g_{\Phi_{k,l}}^{-2} + l^2\sigma^2 g_{\Upsilon_k^0}^{-2}\right)\right) d\nu_k^0 d\boldsymbol{\phi}_k d\boldsymbol{a}_k, \quad (5)$$

where $\sigma^2 = \frac{1}{\|w\|^2}\sum_n w(n)^2 n^2$. In this derivation quantization step sizes were replaced by so-called quantization point density functions g_A, g_Φ and g_{Υ^0} [11, 12], which when integrated over a region S gives the total number of quantization levels within S. In the case of scalar quantizers, this means that the quantizer step sizes are just given by the reciprocal values of the point densities, that is, $g = \Delta^{-1}$. In high-resolution theory, quantizers are described by these density functions, without exactly specifying the location of the quantization points. Note that since we consider unrestricted quantization, the quantization point density functions can depend on all variables in layer k. This dependence is omitted in (5) for notational convenience.

2.2. Entropy-constrained minimization of the expected distortion

In this section we determine the quantization point density functions that solve

$$\min_{g_{A_{k,l}}, g_{\Phi_{k,l}}, g_{\Upsilon_k^0}} D \text{ subject to } H \le H_t, \quad (6)$$

where H_t is the total target entropy, and $H = \sum_{k=1}^K H(\tilde{\boldsymbol{A}}_k, \tilde{\boldsymbol{\Phi}}_k, \tilde{\Upsilon}_k^0)$ is the total entropy of quantization indices, where we assume that entropies are additive over harmonic layers. Here $H(\tilde{\boldsymbol{A}}_k, \tilde{\boldsymbol{\Phi}}_k, \tilde{\Upsilon}_k^0)$ is the joint entropy of all amplitude, phase and fundamental frequency indices in layer k, where $\tilde{\boldsymbol{A}}_k$, $\tilde{\boldsymbol{\Phi}}_k$ and $\tilde{\Upsilon}_k^0$ denote their corresponding alphabets, respectively. Using high resolution assumptions, we approximate

$$\begin{aligned} H(\tilde{\boldsymbol{A}}_k, \tilde{\boldsymbol{\Phi}}_k, \tilde{\Upsilon}_k^0) \approx{}& h(\boldsymbol{A}_k, \boldsymbol{\Phi}_k, \Upsilon_k^0) \\ &+ \sum_{l=1}^{L_k}\iiint f_{\boldsymbol{A}_k,\boldsymbol{\Phi}_k,\Upsilon_k^0}\log_2(g_{A_{k,l}}) d\nu_k^0 d\boldsymbol{\phi}_k d\boldsymbol{a}_k \\ &+ \sum_{l=1}^{L_k}\iiint f_{\boldsymbol{A}_k,\boldsymbol{\Phi}_k,\Upsilon_k^0}\log_2(g_{\Phi_{k,l}}) d\nu_k^0 d\boldsymbol{\phi}_k d\boldsymbol{a}_k \\ &+ \iiint f_{\boldsymbol{A}_k,\boldsymbol{\Phi}_k,\Upsilon_k^0}\log_2(g_{\Upsilon_k^0}) d\nu_k^0 d\boldsymbol{\phi}_k d\boldsymbol{a}_k, \end{aligned}$$

where $h(\boldsymbol{A}_k, \boldsymbol{\Phi}_k, \Upsilon_k^0)$ is the joint differential entropy of amplitude, phase and fundamental frequency parameters corresponding to layer k. The constrained minimization problem in (6) is solved using the method of Lagrange multipliers. We compute a Lagrangian cost function $J = D + \lambda H$, where λ is the Lagrangian multiplier. Minimizing this cost function by evaluating the Euler-Lagrange equations for all three quantization point densities, yield analytical expressions for the optimal high-resolution ECUSQ$_h$ quantizers:

$$g_{A_{k,l}} = 2^{-\xi^{-1}\left(\sum_{m=1}^K\sum_{n=1}^{L_m} b(A_{m,n}) + \frac{1}{2}\sum_{m=1}^K \bar{b}(\boldsymbol{A}_m) + K\log_2(\sigma)\right)} \times 2^{\xi^{-1}\left(H_t - \sum_{m=1}^K h(\boldsymbol{A}_m, \boldsymbol{\Phi}_m, \Upsilon_m^0)\right)}, \quad (7)$$

$$g_{\Phi_{k,l}} = a_{k,l}g_{A_{k,l}}, \quad (8)$$

$$g_{\Upsilon_k^0} = \sigma\left(\sum_{l=1}^{L_k} l^2 a_{k,l}^2\right)^{\frac{1}{2}} g_{A_{k,l}}, \quad (9)$$

where $\xi = 2\sum_{k=1}^K L_k + K$ is the number of sinusoidal parameters in the signal model (1). Furthermore $b(A_{k,l}) = E\left(\log_2(A_{k,l})\right)$ and $\bar{b}(\boldsymbol{A}_k) = E\left(\log_2\left(\sum_{l=1}^{L_k} l^2 A_{k,l}^2\right)\right)$ are introduced for notational convenience. Note that the optimal amplitude quantizer is uniform and is the same for all amplitudes, i.e. $g_{A_{k,l}} \triangleq g_A$. Furthermore, phase quantization depends linearly on the corresponding amplitude, as is the case in the ordinary ECUSQ scheme for the ℓ_2 measure [10]. However, fundamental frequency quantization depends non-linearly on all amplitudes in the corresponding layer. Note from (9) that to compensate for the linearly growing frequency quantization errors as we increase l, the amplitudes corresponding to higher l are given more weight in determining the optimal fundamental frequency quantizer. Secondly, since the fundamental frequency quantizer depends on all amplitudes in the concerning layer, it is necessary to compute the expectation in (4) by integrating over all variables within the layer; the phase variables are integrated out. The high-resolution distortion-rate relation for ECUSQ$_h$ is now obtained by substituting (7), (8) and (9) in (5) and (3), giving

$$D_{ECUSQ_h} = \frac{\xi\|w\|^2}{12}g_A^{-2}, \quad \xi = 2\sum_{l=1}^{L_k} L_k + K. \quad (10)$$

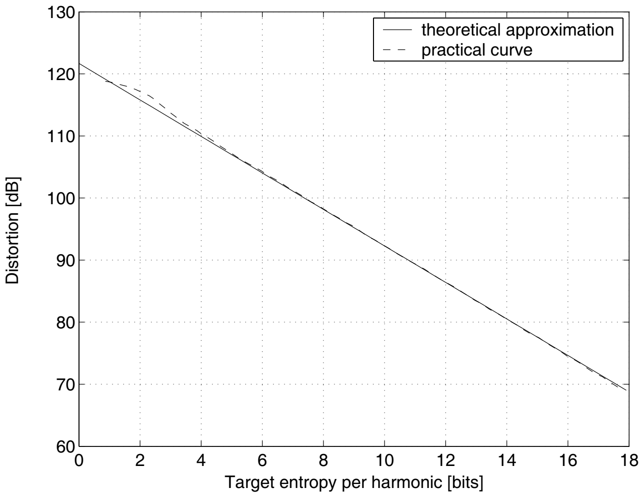

Fig. 1. Theoretical versus practical distortion-rate performance for $N = 1024$.

We can decompose the total distortion D_{ECUSQ_h} into contributions from individual parameter quantization errors as follows. From (3) we observe that D_{ECUSQ_h} is a sum of expected quantization errors $E\left(\|\varepsilon_{k,l}^{w}\|_2^2\right)$, related to individual complex exponentials. Furthermore, note that in (5) these individual quantization errors are evaluated by integrating over three terms, which correspond to the contribution of the amplitude, phase and frequency parameters, respectively, to these expected errors. Consequently, substituting the optimal point density functions in (5), and summing over all L_k harmonics in a layer k, we obtain an expression of the form $C_A+C_\Phi+C_{\Upsilon^0}$, in which C denotes the contribution of each parameter to the expected distortion in layer k. It is now easy to verify that for the optimal quantizers $C_A = C_\Phi = L_k \frac{\|w\|^2}{12} g_A^{-2}$ and $C_{\Upsilon^0} = \frac{\|w\|^2}{12} g_A^{-2}$, i.e. in fact each individual amplitude, phase and fundamental frequency parameter gives exactly the same contribution $\frac{\|w\|^2}{12} g_A^{-2}$ to the expected distortion in one layer, and hence to the total expected distortion (10).

2.3. Simulation example

In this section we compare the theoretical high-resolution distortion-rate relation derived in (10) to a practically obtained distortion-rate curve which is constructed by synthesizing a harmonically structured input signal according to (1), quantizing the corresponding parameters with the derived optimal ECUSQ$_h$ quantizers for different target entropies H_t, and measuring the resulting ℓ_2 distortion and entropy of quantization indices. The signal is synthesized by generating amplitudes, phases and fundamental frequencies from the same distribution for every parameter, i.e. $A_{k,l} = A$, $\Phi_{k,l} = \Phi$ and $\Upsilon_k^0 = \Upsilon^0$, where A is Rayleigh distributed ($\sigma_A = 1000$), Φ and Υ^0 are uniformly distributed on intervals $[0, 2\pi]$ and $[\frac{\pi}{400}, \frac{\pi}{45}]$ respectively. These three random variables are assumed independent. The chosen distributions are close to the ones measured in experiments with real audio data. The fundamental frequency interval approximately corresponds to the range 50-500 Hz which includes most speech and many musical instruments. We consider scalar coding of all quantization indices, so the joint differential entropy in the optimal quantizer formulas is replaced by the sum of all marginal differential entropies in the corresponding layer. Furthermore, the signal consists of 100 segments, having one layer each, where the number of harmonics in each layer is randomly chosen between 5 and 40. The frame length is set at $N = 1024$ and the analysis window w was a Hanning window. Note that an input signal consisting of 50 segments with equal frame lengths and two layers per segment or any other distribution of the 100 layers would produce the same results. Figure 1 shows the results, where the rate is averaged over all harmonics. Clearly, the two curves converge, which verifies that (10) is indeed a valid high-resolution approximation.

3. COMPARISON TO ECUSQ

In the *direct* ECUSQ scheme, as introduced in [10], optimal quantizers are derived for every individual frequency parameter, i.e. every frequency value, as well as every amplitude and phase value, is quantized and encoded separately. This scheme can be considered as a special case of the ECUSQ$_h$ scheme, if all layers consist of one harmonic. In *differential* ECUSQ the differences between consecutive frequencies are quantized and encoded separately, while amplitude and phase parameters are still treated directly, i.e. without using differential techniques. In [13] it is shown that differential ECUSQ yields better rate-distortion performance as compared to direct ECUSQ. In this section we will compare the proposed scheme with both versions of the ECUSQ-scheme in terms of rate-distortion performance, and the distribution of entropy between the three parameters. All results in this section are derived using the same settings as in the simulation example in the previous section.

The distribution of entropy between amplitude, phase and frequency in the optimal ECUSQ$_h$ scheme for these settings can be found by applying the entropy chain rule for every harmonic layer

$$H(\tilde{\boldsymbol{A}}_k, \tilde{\boldsymbol{\Phi}}_k, \tilde{\Upsilon}_k^0) = H(\tilde{\boldsymbol{A}}_k) + H(\tilde{\boldsymbol{\Phi}}_k|\tilde{\boldsymbol{A}}_k) + H(\tilde{\Upsilon}_k^0|\tilde{\boldsymbol{A}}_k, \tilde{\boldsymbol{\Phi}}_k).$$

Using high-resolution assumptions, we obtain:

$$\begin{aligned} H(\tilde{\boldsymbol{A}}_k) &\approx L_k\left(h(A) + \log_2(g_A)\right), \\ H(\tilde{\boldsymbol{\Phi}}_k|\tilde{\boldsymbol{A}}_k) &\approx L_k\left(h(\Phi) + b(A) + \log_2(g_A)\right), \\ H(\tilde{\Upsilon}_k^0|\tilde{\boldsymbol{A}}_k, \tilde{\boldsymbol{\Phi}}_k) &\approx h(\Upsilon^0) + b(A) + \log_2\left(\sum_{l=1}^{L_k} l^2\right) \\ &\quad + \log_2(\sigma) + \log_2(g_A). \end{aligned}$$

In [10] these entropies were derived in a similar way for the direct ECUSQ scheme. The entropy expressions in the differential case are exactly the same as in the direct case, however, for a harmonically structured input signal the differential entropy for frequency parameters in layer k is given by $\sum_{l=1}^{L_k} h(l\Upsilon^0) = L_k h(\Upsilon^0) + \sum_{l=1}^{L_k} \log_2(l)$ for direct ECUSQ, and is given by $L_k h(\Upsilon^0)$ for differential ECUSQ, resulting in different discrete entropies for both schemes. Summing over all layers, we then obtain the bit distribution in the entire signal, which is plotted in Figure 2 per parameter for all three schemes, as a function of the target entropy H_t. Given a value of H_t, the number of bits assigned to amplitude in a certain scheme can be looked up in the top plot. The middle and bottom plot give the number of bits assigned to phase and frequency, respectively. The range of H_t in the plots corresponds to 10 to 30 bits per harmonic. Clearly, the number of bits assigned to frequency parameters is much higher for both ECUSQ schemes, which is to be expected since in both versions of the ECUSQ scheme, each frequency parameter is encoded separately, while in ECUSQ$_h$ only fundamental frequency parameters are encoded. Consequently, in ECUSQ$_h$ more bits are left for amplitude and phase parameters, as can be seen in the upper two plots. Furthermore, the differential ECUSQ scheme assigns less bits to frequency than its direct form, which is due to the fact that differential entropy of frequency parameters is smaller in the differential case.

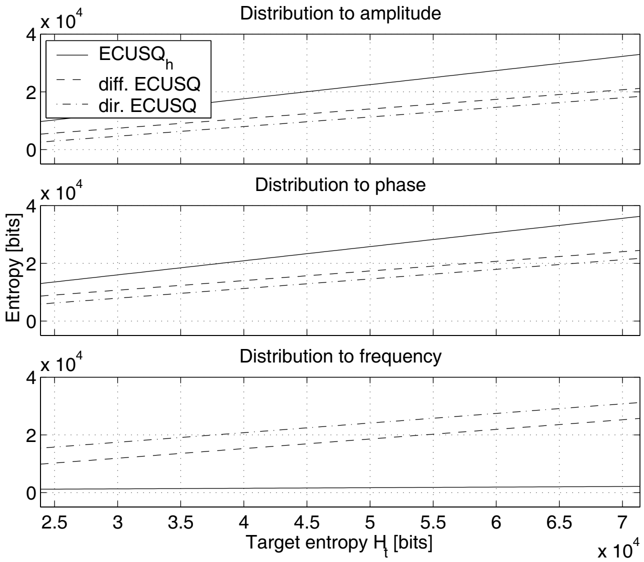

Fig. 2. Distribution of entropy between amplitude, phase and frequency for three schemes.

In Figure 3 the high-resolution rate-distortion relations for all three schemes are plotted, as derived in (10) for ECUSQ$_h$ and in [10] for direct ECUSQ. The rate-distortion relation for differential ECUSQ is equal to the direct relation, but as mentioned earlier, the differential entropies of frequency parameters differ for the ECUSQ schemes, resulting in different distortions for the same target entropy. The range for the target entropy H_t is the same as in Figure 2. It is clear that ECUSQ$_h$ significantly outperforms both ECUSQ schemes. The following observation can be made here. It is easy to verify that the distortion-rate relation for all three schemes is of the form

$$D = C 2^{-2\xi^{-1} H_t}. \tag{11}$$

For fixed input data C is a constant, which is different for each scheme and depends on source distributions A, Φ, and Υ^0, the window w, and the number of layers K and harmonics L_k. For the given input settings, C is smallest for the ECUSQ$_h$ scheme. Furthermore ξ is the total number of sinusoidal parameters needed for representing the signal in a given scheme. As we noted before, ξ in the ECUSQ$_h$ scheme is smaller than in the other two schemes. From (11), we then conclude that the slope of the distortion-rate curve for ECUSQ$_h$ is more steep, as is clear from Figure 3. It might be possible to find input settings, such that C would be lower for the two ECUSQ schemes. However, taking the target entropy H_t sufficiently large, the proposed scheme would still outperform both ECUSQ schemes.

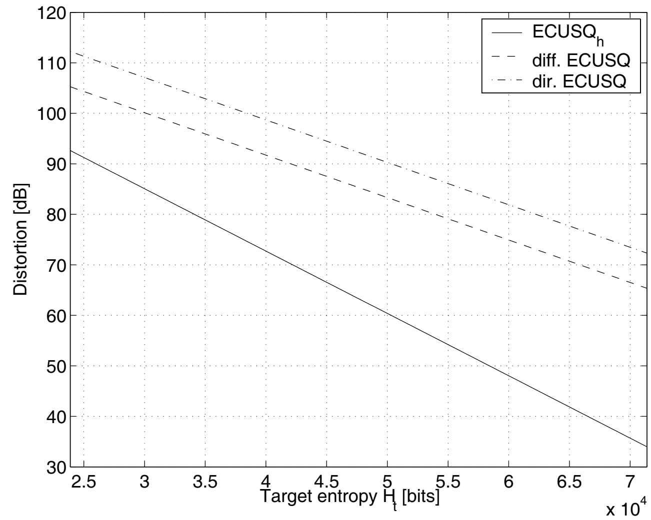

Fig. 3. Rate-distortion performance for ECUSQ$_h$ and ECUSQ

4. REFERENCES

[1] H. Purnhagen, "Advances in parametric audio coding," in *Proc. 1999 IEEE Workshop on Applications of Signal Proc. to Audio and Acoustics,*, (New Paltz, New York, USA), pp. W99-1-W99-4, 1999.

[2] T.S. Verma and T.H.Y. Meng, "A 6 kbps to 85 kbps scalable audio coder," in *Proc. IEEE Int. Conf. Acoust. Speech, and Signal Proc.*, vol. II, (Istanbul, Turkey), pp. 887-880, 2000.

[3] J. R. Deller, Jr., J.H.L. Hansen and J.G. Proakis, "Discrete-Time Processing of Speech Signals", IEEE Press, 2000.

[4] H. Purnhagen and N. Meine, "HILN – The MPEG-4 Parametric Audio Coding Tools", in *Proc. IEEE Int. Symp. on Circuits and Systems*, Geneva, Switzerland, May 2000.

[5] M. Zeytinoglu and K.M. Wong, "Detection of harmonic sets," in *IEEE Trans. Signal Processing*, vol. 43, pp. 2618-2630, Nov. 1995.

[6] R. Gribonval and E. Bacry, "Harmonic decomposition of audio signals with matching pursuit," in *IEEE Trans. Signal Processing*, vol. 51, pp. 101-111, Jan. 2003.

[7] A. de Cheveigné and A. Baskind, "F_0 estimation of one or several voices", in *Proc. EUROSPEECH 2003*, Geneva, Switzerland, pp. 833-836, 2003.

[8] G. Wilson, "Magnitude/phase quantization of independent Gaussian variates," *IEEE Trans. Commun.*, vol. 28, pp. 1924-1929, November 1980.

[9] R. Vafin and W. B. Kleijn, "Entropy-constrained polar quantization and its application to audio coding," *IEEE Trans. Speech Audio Processing*, vol. 13, no. 2, pp. 220-232, March 2005.

[10] P.E.L. Korten, J.Jensen and R.Heusdens "High rate spherical quantization of sinusoidal parameters," in *Proc. 12th European Signal Processing Conf.*, Vienna , Austria, pp. 1757-1760, Sept. 2004.

[11] R. M. Gray and D. L. Neuhoff, "Quantization," in *IEEE Trans. Information Theory*, vol. 44(6), pp. 2325-2383, Oct. 1998.

[12] S. P. Lloyd, Least squares quantization in PCM. *IEEE Trans. Information Theory*, 28:129–137, 1982.

[13] R. Heusdens, J. Jensen, P.E.L. Korten and R. Vafin, "Rate-distortion optimal high-resolution differential quantisation for sinusoidal coding of audio and speech," in *Proc. 2005 IEEE Workshop on Applications of Signal Proc. to Audio and Acoustics,*, New Paltz, New York, USA, pp. 243-246, 2005.