# Polar Quantization of Sinusoids from Speech Signal Blocks

*Harald Pobloth, Renat Vafin, and W. Bastiaan Kleijn*

Department of Signals, Sensors, and Systems
KTH (Royal Institute of Technology)
100 44 Stockholm, Sweden
{ Harald.Pobloth, Renat.Vafin, Bastiaan.Kleijn }@s3.kth.se

## Abstract

We introduce a block polar quantization (BPQ) procedure that minimizes a weighted distortion for a set of sinusoids representing one block of a signal. The minimization is done under a resolution constraint for the entire signal block. BPQ outperforms rectangular quantization, strictly polar quantization, and unrestricted polar quantization (UPQ) both when assuming the Cartesian coordinates of the sinusoidal components to be Gaussian and for sinusoids found from speech data. In the case of speech data we found a significant performance gain (about 4 dB) over the best performing polar quantization (UPQ).

## 1. Introduction

The quantization of the amplitude and phase of complex variables is a common task in speech, audio, and image processing. In speech coding, this task is relevant for sinusoidal coding, e.g., [1], and waveform interpolation coding, e.g., [2]. In this paper, we consider the scalar quantization of the amplitudes and phases of a block of $L$ sinusoids, representing a speech segment containing $N$ samples, with a weighted mean squared error criterion and a resolution constraint. In contrast to many conventional methods, our quantizers are adaptive. That is, the quantizers adapt to the particular realization of the signal. Only a small computational overhead is required for this adaptation. We refer to our new method as block polar quantization (BPQ).

BPQ operates with a given bit budget constraint of $B$ bits for all $L$ amplitudes and phases. If $B_A$ and $B_\Phi$ are the numbers of bits spent for the set of $L$ amplitudes and phases, respectively, then the overall bit budget is $B = B_A + B_\Phi$. Given this bit budget constraint, BPQ asymptotically (for high rates) minimizes the weighted distortion measure

$$D = \mathrm{E}\left[\frac{1}{L}\sum_{l=1}^{L} w_l \left|\left(a_l \exp(j\phi_l) - \hat{a}_l \exp(j\hat{\phi}_l)\right)\right|^2\right], \quad (1)$$

where $\mathrm{E}[\cdot]$ denotes expectation, $w_l$, $a_l$, and $\phi_l$ are the weight, amplitude, and phase of sinusoid $l$, respectively, and the symbol $\hat{}$ indicates that the variables are quantized. Weights that make the distortion measure perceptually meaningful can, e.g, be based on the inverse of the masking threshold. We note that the entropy-constrained case of this problem was earlier solved in [3]. (For a resolution constraint the rate for each signal block is fixed while for an entropy constraint the rate for the signal blocks can vary and the average rate is fixed.)

To evaluate the performance of our asymptotically optimal quantizers, we compare their performance with the performance of rectangular quantization (RQ), e.g, [4], strictly polar quantization (SPQ), e.g, [4], and unrestricted polar quantization (UPQ), e.g, [5] for the Gaussian case. Our method gives lower distortions than RQ, SPQ, and UPQ with only a marginal increase in complexity. We show that, despite the fact that the quantizers are only asymptotically optimal, they provide good performance at practical rates. More-over the asymptotic predictions of the distortion-rate relations are useful at practical rates. BPQ performance is particularly good for speech signals.

As was mentioned above, BPQ outperforms other polar quantization methods. Spherical quantizers, e.g., [5, 6], also outperform such polar quantizers while maintaining similar complexity for the unweighted mean squared error criterion. However, in contrast to BPQ, existing spherical quantizers do not incorporate a weighting and it is not trivial to include weighting in them. Since weighting is essential for efficient quantization of a speech signal, and since BPQ outperforms other quantizers that easily facilitate weighting, it forms an attractive method for quantizing sinusoids in speech.

## 2. Block polar quantization (BPQ)

We derive the optimal polar quantizers assuming independent, identically distributed (i.i.d.) amplitudes and uniform i.i.d. phases for a set of $L$ sinusoids representing each block of speech data. The $L$ amplitudes are quantized and transmitted first, and afterwards the phase quantizers are adapted for the observed amplitudes.

The distortion of eq. 1 can be rewritten as

$$D = \mathrm{E}\left[\frac{1}{L}\sum_{l=1}^{L} w_l \left(a_l^2 + \hat{a}_l^2 - 2a_l\hat{a}_l\cos(\phi_l - \hat{\phi}_l)\right)\right]. \quad (2)$$

Denoting $d(a_l, \phi_l, \hat{a}_l, \hat{\phi}_l) = a_l^2 + \hat{a}_l^2 - 2a_l\hat{a}_l\cos(\phi_l - \hat{\phi}_l)$, the average distortion in one given quantization cell, bounded between $a_{k_l}$ and $a_{k_l} + \Delta a_{k_l}$ and $\phi_{i_l}$ and $\phi_{i_l} + \Delta\phi_{i_l}$ is

$$D(\hat{a}_{k_l}, \hat{\phi}_{i_l}, \Delta a_{k_l}, \Delta\phi_{i_l}) = \frac{\int_{a=a_{k_l}}^{a_{k_l}+\Delta a_{k_l}} \int_{\phi=\phi_{i_l}}^{\phi_{i_l}+\Delta\phi_{i_l}} f_{A,\Phi}(a,\phi) d(a, \hat{a}_{k_l}, \phi, \hat{\phi}_{i_l}) \mathrm{d}a\mathrm{d}\phi}{\int_{a=a_{k_l}}^{a_{k_l}+\Delta a_{k_l}} \int_{\phi=\phi_{i_l}}^{\phi_{i_l}+\Delta\phi_{i_l}} f_{A,\Phi}(a,\phi) \mathrm{d}a\mathrm{d}\phi}, \quad (3)$$

where $f_{A,\Phi}(a, \phi)$ is the joint probability density function (PDF) of the amplitude and phase, $\hat{a}_{k_l}$ is the $k$-th amplitude reconstruction point and $\hat{\phi}_{i_l}$ is the $i$-th phase reconstruction point in dimension $l$.

In the high-rate case, we approximate $f_{A,\Phi}(a, \phi) \approx f_{A,\Phi}(\hat{a}_{k_l}, \hat{\phi}_{i_l})$ within amplitude cell $k_l$ and phase cell $i_l$. Further, we assume the reconstruction points $\hat{a}_{k_l}$ and $\hat{\phi}_{i_l}$ to be the mid-points of the quantization intervals $[a_{k_l}, a_{k_l} + \Delta a_{k_l})$ and $[\phi_{i_l}, \phi_{i_l} + \Delta\phi_{i_l})$, respectively. Thus, eq. 3 simplifies to

$$D(\hat{a}_{k_l}, \Delta a_{k_l}, \Delta\phi_{i_l}) = \frac{\Delta a_{k_l}^2}{12} + 2\hat{a}_{k_l}^2 - 4\hat{a}_{k_l}^2 \frac{\sin\left(\frac{\Delta\phi_{i_l}}{2}\right)}{\Delta\phi_{i_l}} . \quad (4)$$

In the next step, we introduce the reconstruction point densities $g_{A_l}(a)$ and $g_{\Phi_l}(\hat{\mathbf{a}})$, which are the inverse of the quantization cell sizes $\Delta a_{k_l}$ and $\Delta\phi_{i_l}$. Since $\hat{\mathbf{a}} = (\hat{a}_{k_1}, \ldots, \hat{a}_{k_L})^T$, the phase reconstruction point density is a function of all quantized amplitudes. The independence of $g_{\Phi_l}(\hat{\mathbf{a}})$ from $\phi_l$ results from the uniform PDF of $\phi_l$. The distortion in a given cell becomes

$$D(\hat{a}_{k_l}, g_{A_l}(\hat{a}_{k_l}), g_{\Phi_l}(\hat{\mathbf{a}})) = \frac{g_{A_l}^{-2}(\hat{a}_{k_l})}{12} + 2\hat{a}_{k_l}^2 - 4\hat{a}_{k_l}^2 \frac{\sin\left(\frac{g_{\Phi_l}^{-1}(\hat{\mathbf{a}})}{2}\right)}{g_{\Phi_l}^{-1}(\hat{\mathbf{a}})} . \quad (5)$$

For high rates, the total distortion can be approximated using eq. 5 and eq. 2:

$$D_{BPQ} \approx \frac{1}{L}\sum_{l=1}^{L} w_l \left[\int_{a=0}^{\infty} f_A(a)\left(\frac{g_{A_l}^{-2}(a)}{12} + 2a^2\right)\mathrm{d}a - 4\int_{\mathbf{a}=\mathbf{0}}^{\infty} f_{\mathbf{A}}(\mathbf{a}) a_l^2 \sin\left(\frac{g_{\Phi_l}^{-1}(\mathbf{a})}{2}\right) g_{\Phi_l}(\mathbf{a})\mathrm{d}\mathbf{a}\right] . \quad (6)$$

### 2.1. Amplitude reconstruction point densities

Observing that $\int_{a=0}^{\infty} g_{A_n}(a)\mathrm{d}a$ is equal to the number of amplitude reconstruction points in dimension $n$, the constraint for finding the $g_{A_l}(a)$ that minimizes the distortion of eq. 6 becomes

$$B_A = \sum_{n=1}^{L} \log_2\left(\int_{a=0}^{\infty} g_{A_n}(a)\mathrm{d}a\right) . \quad (7)$$

When considering only the part of eq. 6 that depends on $g_{A_l}(a)$ and rewriting the constraint of eq. 7 the function to minimize is

$$\eta = \sum_{l=1}^{L}\int_{\mathbf{a}}\left(w_l f_{\mathbf{A}}(\mathbf{a})\frac{g_{A_l}^{-2}(a_l)}{12} + \frac{\lambda}{L}\prod_{n=1}^{L} g_{A_n}(a_n)\right)\mathrm{d}\mathbf{a} , \quad (8)$$

where $\lambda$ is the Lagrange multiplier. Then $\eta$ is minimized by the solution to the Euler-Lagrange equation

$$\frac{\partial\eta}{\partial g_{A_l}(a_l)} = -f_A(a_l) w_l \frac{g_{A_l}^{-3}(a_l)}{6} + \frac{\lambda}{L}\int_{\mathbf{a}_{\forall n\neq l}} \frac{\prod_{n=1}^{L} g_{A_n}(a_n)}{g_{A_l}(a_l)}\mathrm{d}\mathbf{a}_{\forall n\neq l} \equiv 0 . \quad (9)$$

From eq. 9 it follows that

$$\frac{w_l}{6}\int_a f_A(a) g_{A_l}^{-2}(a)\mathrm{d}a = \frac{\lambda}{L}2^{B_A} . \quad (10)$$

Thus, $w_l \int_a f_A(a) g_{A_l}^{-2}(a)\mathrm{d}a$ is constant for all $l$ and a general $f_A(a)$. This can be achieved when

$$g_{A_l}(a) = \sqrt{w_l}\, g_A(a) . \quad (11)$$

Using $g_{A_l}(a)$ of eq. 11 in eq. 8 we obtain

$$\eta = \int_a \left(f_A(a)\frac{g_A^{-2}(a)}{12} + \lambda g_A(a)\right)\mathrm{d}a , \quad (12)$$

which leads to the solution

$$g_{A_l}(a) = \sqrt{\frac{w_l}{\bar{w}}}\,\frac{f_A^{1/3}(a)2^{B_A/L}}{\int_x f_A^{1/3}(x)\mathrm{d}x} , \quad (13)$$

where $\bar{w} = \prod_{n=1}^{L} w_n^{1/L}$ is the geometric mean of the weights.

### 2.2. Phase reconstruction point densities

Given that the $L$ amplitudes are quantized to $\hat{\mathbf{a}}$, the average distortion is given by

$$D(\hat{\mathbf{a}}) = \frac{1}{L}\sum_{l=1}^{L} w_l \left(\frac{g_{A_l}^{-2}(\hat{a}_{k_l})}{12} + 2\hat{a}_{k_l}^2 - 4\hat{a}_{k_l}^2 \frac{\sin\left(\frac{g_{\Phi_l}^{-1}(\hat{\mathbf{a}})}{2}\right)}{g_{\Phi_l}^{-1}(\hat{\mathbf{a}})}\right) , \quad (14)$$

where we used eq. 5. Since $2\pi g_{\Phi_l}(\hat{\mathbf{a}})$ is the number of phase reconstruction points for dimension $l$, $D(\hat{\mathbf{a}})$ is minimized under the constraint

$$B_\Phi = \sum_{l=1}^{L} \log_2\left(2\pi g_{\Phi_l}(\hat{\mathbf{a}})\right) . \quad (15)$$

Using the method of Lagrange multipliers we find

$$g_{\Phi_l}(\hat{\mathbf{a}}) = \sqrt{\frac{w_l}{\bar{w}}}\,\frac{\hat{a}_{k_l} 2^{B_\Phi/L}}{2\pi\prod_{n=1}^{L}\hat{a}_{k_n}^{1/L}} . \quad (16)$$

The phase quantizers $g_{\Phi_l}(\hat{\mathbf{a}})$ have to be found for every transmitted set of amplitudes. Thus, eq. 16 represents the (only) computational overhead in BPQ compared to conventional polar quantizers.

### 2.3. Distortion-rate relation

Using the results of eqs. 13 and 16 for amplitude and phase reconstruction point densities in the expression for the distortion given in eq. 6 we find

$$D_{BPQ} = \frac{b}{2^{2B_A/L}} + \frac{c}{2^{2B_\Phi/L}} , \quad (17)$$

where we defined

$$b = \frac{1}{12}\left(\int_a f_A^{1/3}(a)\mathrm{d}a\right)^3 \bar{w} , \quad (18)$$

$$c = \frac{\pi^2}{3}\left(\int_a f_A(a) a^{2/L}\mathrm{d}a\right)^L \bar{w} \quad (19)$$

for notational convenience, and where we used the approximation $\sin(x^{-1}/2)x \approx 1/2 - x^{-2}/48$ for large $x$.

To find the optimal distribution of bits between the amplitude quantizers $B_A$ and the phase quantizers $B_\Phi$ for a given total bit budget $B = B_A + B_\Phi$ we denote $B_A = \alpha B$ and $B_\Phi = (1-\alpha)B$. This reduces the problem of bit distribution to the problem of finding the $\alpha$ that minimizes the distortion

$$D_{BPQ} = \frac{b}{2^{2\alpha B/L}} + \frac{c}{2^{2(1-\alpha)B/L}} . \quad (20)$$

The $\alpha$ minimizing eq. 20 is

$$\alpha = \frac{1}{2} + \frac{L}{4B}\log_2\left(\frac{b}{c}\right) , \quad (21)$$

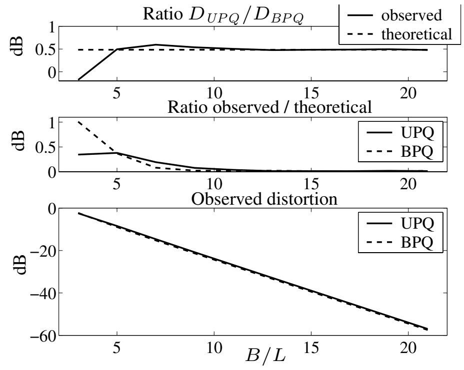


Figure 1: Comparison of predicted and observed performance of UPQ and BPQ with Gaussian data. The amplitude PDF is found using a histogram. The variance $\sigma^2$ is 1 and $L$ is 40.

which is independent of the weights. With this $\alpha$ the distortion-rate relation becomes

$$D_{BPQ} = \frac{2}{2^{B/L}}\sqrt{bc}\,. \qquad (22)$$

Since $D_{BPQ} \sim 2^{-B/L}$ the distortion reduces by 3 dB when $B/L$ is increased by 1 bit.

## 3. Gaussian data

In the literature the performance of polar quantization is often considered for the case that the real and imaginary parts of all complex components are i.i.d. Gaussian variables with variance $\sigma^2$. This does not only lead to simple expressions but also has practical significance as argued, e.g., in [7]. The Gaussian distribution of the complex components makes the PDF of the amplitude a Rayleigh distribution. The phase in this case is uniformly distributed.

The amplitude reconstruction point density becomes

$$g_{A_l}(a) = \sqrt{\frac{w_l}{\bar{w}}}\,\frac{2^{B_A/L}(2a)^{1/3}}{(9\sigma^4)^{1/3}\Gamma\left(\frac{2}{3}\right)\exp\left(\frac{a^2}{6\sigma^2}\right)}\,, \qquad (23)$$

the constants $b$ and $c$ become

$$b = \frac{3\sigma^2}{8}\Gamma^3\left(\frac{2}{3}\right)\bar{w}\,, \qquad (24)$$

$$c = \frac{2\sigma^2\pi^2}{3}\Gamma^L\left(\frac{1}{L}+1\right)\bar{w}\,, \qquad (25)$$

and the distortion-rate expression is

$$D_{BPQ} = \frac{\pi\sigma^2}{2^{B/L}}\Gamma^{3/2}\left(\frac{2}{3}\right)\Gamma^{L/2}\left(\frac{1}{L}+1\right)\bar{w}\,. \qquad (26)$$

To compare the results of BPQ to previously known polar quantizers we set all weights $w_l$ to 1. Figure 1 compares the predicted performance of the BPQ and the UPQ and shows the performance of actual quantizers found by averaging over the squared error from 50 000 data points. Since in resolution-constrained UPQ it is not possible to transmit the amplitude independent of the phase, it is not possible to optimize the phase density for the observed amplitudes. Thus, the UPQ is identical in all dimensions.

The quantizers were not designed exploiting the knowledge of the Rayleigh amplitude PDF. Instead, the amplitude PDF is found using a histogram over 50 000 data points (disjoint from the encoded ones) and numerical integration was used to find the reconstruction point densities. (For the Gaussian data used here this is not necessary since the PDF is known but for general data as the speech data of section 4 the PDF has to be approximated by, e.g., histograms.) Figure 1 shows that the approximation of the PDF does not influence the result significantly.

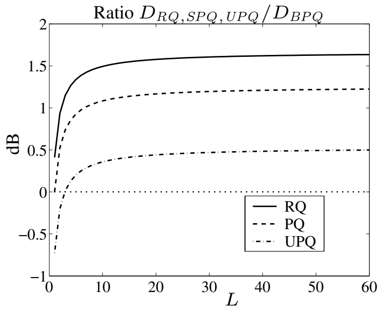


Figure 2: Difference between BPQ and RQ, SPQ, and UPQ as given in eqs. 27, 28, and 29.

Next, we give expressions for how the theoretical performance of BPQ compares to rectangular quantization (RQ), strictly polar quantization (SPQ), and unrestricted polar quantization (UPQ). For $L = 1$ BPQ reduces to SPQ and thus has equal performance.

The distortion for RQ can be found in [4, eq. 19 and eq. 20] and the ratio of RQ and BPQ distortion is

$$\frac{D_{RQ}}{D_{BPQ}} = \sqrt{3}\Gamma^{-3/2}\left(\frac{2}{3}\right)\Gamma^{-L/2}\left(\frac{1}{L}+1\right)\,. \qquad (27)$$

The distortion for SPQ is found by minimizing eq. 13 in [4] for a general variance $\sigma^2$ and the resulting SPQ/BPQ distortion ratio is

$$\frac{D_{SPQ}}{D_{BPQ}} = \Gamma^{-L/2}\left(\frac{1}{L}+1\right)\,. \qquad (28)$$

Finally we compare BPQ to UPQ. An expression for UPQ distortion is given in [5, eq. 1] and the distortion ratio is

$$\frac{D_{UPQ}}{D_{BPQ}} = \frac{4}{3}\Gamma^{-3/2}\left(\frac{2}{3}\right)\Gamma^{-L/2}\left(\frac{1}{L}+1\right)\,. \qquad (29)$$

Noting that $\lim_{L\to\infty}\Gamma^{-L/2}(1/L+1) = \exp(\gamma/2)$, where $\gamma \approx 0.5772157$ is Euler's constant, we find the values shown in table 1. From figure 2 we can conclude that BPQ outperforms all of the considered quantizers for $L \geq 3$. The fact that UPQ performs closest to BPQ motivated the comparison to UPQ only in figures 1, 3, and 4.

## 4. Speech data

In this section, we apply BPQ to the sinusoids found by using the analysis/synthesis system described in [8]. The input data is speech from the TIMIT database [9] sampled at 8 kHz. For each analysis block of 30 ms (240 samples) a set of $L = 40$ sinusoids is extracted. Before encoding the amplitudes, they are normalized by the maximum amplitude in the set, which has to

| | RQ/BPQ | SPQ/BPQ | UPQ/BPQ |
|---|---|---|---|
| $L = 1$ | 0.4108 dB | 0 dB | -0.7255 dB |
| $L \to \infty$ | 1.664 dB | 1.253 dB | 0.5280 dB |

Table 1: Performance ratios for Gaussian data according to eqs. 27, 28, and 29 in dB.

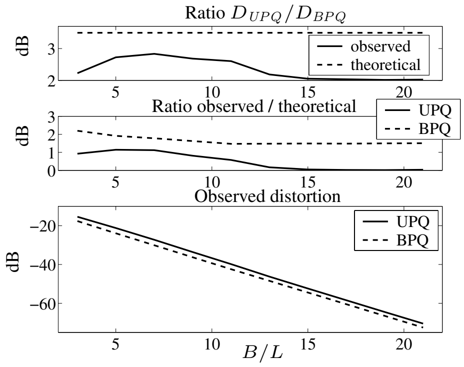


Figure 3: Distortion in the sinusoidal components (eq. 1) as a function of rate for speech data. Block length 30 ms (240 samples), $L = 40$.

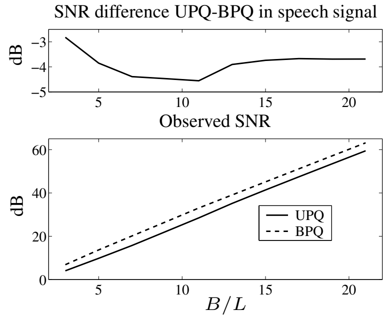


Figure 4: SNR in the speech signal caused by quantization of the sinusoidal components. Block length 30 ms (240 samples), $L = 40$.

be transmitted as side information. For easy comparison with UPQ all weights $w_l$ are set to 1. Figure 3 shows the performance of the actual BPQ and UPQ quantizers compared to the predictions for BPQ in eq. 22 and for UPQ in [5].

The PDF $f_A(a)$ was found via a histogram from 11 minutes of speech from the TIMIT database [9] while the distortion of the actual quantizers was found by averaging over another 12 minutes of speech from TIMIT. The gain from using BPQ instead of UPQ for the quantization of the normalized sinusoids is about 2 dB. This corresponds to about 0.67 bit per sinusoidal component and relates to a rate reduction of 1.8 kbit/s for the chosen block length and 50% overlap. The only increase in complexity to achieve this gain is the design of the phase quantizers according to eq. 16 each time the amplitudes are received. The increase in deviation from predicted and actual performance of BPQ is likely to be due to the fact that some phase quantizers are assigned very low rates.

Figure 4 shows the observed signal to noise ratio (SNR) in the reconstructed speech signal for BPQ and UPQ. The performance gain for BPQ is about 4 dB corresponding to about 3.4 kbit/s, which is higher than the gain for the sinusoidal components. The increased difference between BPQ and UPQ is due to the denormalization of all $L$ amplitudes in one set by the maximum amplitude in the set. We observed that sets with high maximum amplitudes often contained only few sinusoids with high amplitudes and many sinusoids with low amplitudes. BPQ performs very efficiently for such blocks since it adapts the phase quantizers for the different sinusoids. In addition the high maximum amplitude blocks are of increased relative importance for the reconstructed speech. In figure 4 only the distortion caused by the quantization of the sinusoidal components, not the distortion caused by the fact that the sinusoidal components do not represent the original signal without distortion is considered. In [8] it is argued that the part of the signal not captured by the sinusoids is described more efficiently by other coders, as, e.g., waveform coders.

An efficient way of finding weights $w_l$ is to exploit the masking found from the amplitudes $a_l$ as described in [3]. In this scenario the amplitude quantizers are optimized for all $w_l = 1$ and the phase quantizers are optimized for the $w_l$ found from the masking function. This way the weights are adapted to the signal and no additional side information has to be transmitted. In [3] the gain from omitting the side information overcompensates the suboptimal $g_{A_l}(a)$. The fact that phase often is considered of secondary importance compared to amplitude can be taken into account by adjusting $\alpha$.

# 5. Conclusions

Our results show that BPQ outperforms RQ, SPQ, and UPQ for practical rates and data distributions. Thus, the performance of scalar polar quantization with a perceptually meaningful distortion measure is getting closer to the performance of computationally complex vector quantization. The results also show that quantizers that adapt to the observed signal can outperform non-adaptive quantizers. For BPQ the adaptation requires only a marginal increase in computations and, for fixed weights or weights derived from the quantized amplitudes, no additional rate is required. We conclude that BPQ is an appealing alternative to be used in speech coding, audio coding, and image coding.

# 6. References

[1] R. J. McAulay and T. F. Quatieri, "Sinusoidal coding", in W. B. Kleijn and K. K. Paliwal, editors, *Speech Coding and Synthesis*, pp. 121–173. Elsevier Science, Amsterdam, 1995.

[2] W. B. Kleijn and J. Haagen, "Waveform interpolation for speech coding and synthesis", in W. B. Kleijn and K. K. Paliwal, editors, *Speech Coding and Synthesis*, pp. 175–208. Elsevier Science, Amsterdam, 1995.

[3] R. Vafin and W. B. Kleijn, "Entropy-constrained polar quantization and its application to audio coding", *submitted to IEEE Trans. on Speech and Audio Proc.*

[4] J. A. Bucklew and N. C. Gallagher Jr., "Two-dimensional quantization of bivariate circularly symmetric densities", *IEEE Trans. Inform. Theory*, vol. 25, pp. 667 – 671, 1979.

[5] P. F. Swaszek and T. W. Ku, "Asymptotic performance of unrestricted polar quantizers", *IEEE Trans. Inform. Theory*, vol. 32, pp. 330 – 333, March 1986.

[6] J. Hamkins and K. Zeger, "Gaussian source coding with spherical codes", *IEEE Trans. Inform. Theory*, vol. 48, pp. 2980 – 2989, 2002.

[7] W. A. Pearlman and R. M. Gray, "Source coding of the discrete Fourier transform", *IEEE Trans. Inform. Theory*, vol. 24, pp. 683 – 692, 1978.

[8] K. Vos, R. Vafin, R. Heusdens, and W. B. Kleijn, "High-quality consistent analysis-synthesis in sinusoidal coding", *in Proc. AES 17th Int. Conf., 'High-Quality Audio Coding'*, pp. 244 – 250, 1999.

[9] Darpa, "TIMIT", CD-ROM, October 1990, NIST Speech Disc 1-1.1.