Audio Engineering Society
Convention Paper 7370
Presented at the 124th Convention
2008 May 17–20 Amsterdam, The Netherlands



# Quantization of 2D Higher Order Ambisonics wave fields

Audun Solvang[1], U. Peter Svensson[1], Erik Hellerud[1]

[1]*Centre for Quantifiable Quality of Service in Communication Systems*, NTNU, O.S. Bragstads plass 2E, NO-7491 Trondheim, Norway*

Correspondence should be addressed to Audun Solvang (Audun.Solvang@q2s.ntnu.no)

## ABSTRACT

The spatial distribution of the quantization noise for a 2-D Higher Order Ambisonics (HOA) signal is investigated analytically. Uniformly distributed loudspeakers radiating plane waves in a non reverberant environment and frequency domain quantization are presumed. It is found that employing the same quantization interval for all orders leads to uniformly distributed quantization noise in space. Assigning a larger quantization interval (i.e. fewer bits) to higher orders leads to a radially increasing quantization noise. Matching the quantization error to the reproduction error at the near perfect reconstruction boundary suggests that as little as four bits per sample can be used for quantization. Furthermore, high-pass filtering the HOA components opens up for employing as little as three bits per sample. This quantization strategy seems very promising for reducing the rate of HOA.

## 1. INTRODUCTION

The transmission of audio signals over channels with various channel capacities, e.g. the Internet, requires an adaptation of the rate of the audio stream, and this implies distortion. Contemporary methods for lossy compression of audio signals are able to make the distortion perceptually less annoying by shaping it in time and frequency [1]. With the introduction of multiple channels and spatiality in audio, attempts to shape the distortion in space have also been made [2–6]. These are all sweet spot techniques since they are intended for formats like 5.1. However, the sound reproduction formats Wavefield synthesis (WFS) [7, 8] and Higher Order Ambison-

**Centre for Quantifiable Quality of Service in Communication Systems, Centre of Excellence* appointed by The Research Council of Norway, funded by the Research Council, NTNU and UNINETT. http://www.ntnu.no/Q2S/

ics (HOA) [9–11] offer reproduction over an extended area and this leads to new quantization effects. Here we look into the HOA format, which employs a decomposition of a sound field into cylindrical harmonics in 2D (CHD) and spherical harmonics in 3D (SHD) [9, 11, 12]

HOA reproduction with an error smaller than 4% is guaranteed as long as $kr < N$ [13, 14] where $N$ is the truncation order of the harmonic decomposition, $r$ is the radius of the reproduction area and $k$ is the wave number. Other limits for different errors have also been presented [10]. This relationship implies that CHD and SHD are layered transforms where the distortion can be shaped within the listening area or across the frequency spectrum by the number of bits assigned to each order. This is elucidated in this work where plane waves in a non reverberant environment and loudspeakers uniformly distributed on a circle are assumed for simplicity.

The paper is structured as follows. Section 2 reviews the cylindrical harmonics decomposition of a 2D sound field, and its relationship to HOA encoding and decoding. In section 3 the spatial distribution of quantization errors is studied for the quantization of the (complex) amplitude B-format HOA signals. The possibility to shape the quantization noise spatially is discussed in section 4 and numerical simulations are used to demonstrate these effects. Conclusions for the use of HOA as a transmission format are presented in section 5.

## 2. HIGHER ORDER AMBISONICS ENCODING OF A WAVE FIELD

The CHD of a monochromatic two-dimensional wave field can be expressed by the relation [11, 12]

$$\begin{aligned} p(r,k,\theta) &= B_0^{+1} J_0(kr) \\ &+ \sum_{n=1}^{\infty} \sqrt{2}(B_n^{+1}\cos(n\theta) + B_n^{-1}\sin(n\theta)) J_n(kr), \end{aligned} \tag{1}$$

where $J_n(z)$ is the Bessel function of order $n$ and first kind, $r$ is the distance from the point of origin and $k = \frac{2\pi}{\lambda} = \frac{2\pi f}{c}$ is the wave number. $B$ form the HOA B-format signals in the frequency domain.

Presuming that the sound field results from $S$ plane waves from different directions and using the Jacobi-Anger expansion (see eq. (30) in the appendix) leads to

$$\begin{aligned} p(r,k,\theta) &= \sum_{s=1}^{S} p(0,k,\theta_s) e^{jkr\cos(\theta-\theta_s)} \\ &= \sum_{s=1}^{S} p(0,k,\theta_s)[J_0(kr) \\ &+ 2\sum_{n=1}^{\infty} j^n J_n(kr)\cos(n(\theta-\theta_s))]. \end{aligned} \tag{2}$$

It can easily be shown by combining eq. (1) and eq. (2) that the HOA B-format signals can be formulated as

$$\begin{aligned} B_0^{+1} &= \sum_{s=1}^{S} p(0,k,\theta_s), \\ B_n^{+1} &= \sum_{s=1}^{S} p(0,k,\theta_s) j^n \sqrt{2}\cos(n\theta_s), \\ B_n^{-1} &= \sum_{s=1}^{S} p(0,k,\theta_s) j^n \sqrt{2}\sin(n\theta_s). \end{aligned} \tag{3}$$

In practice, the number of orders is truncated because of limitation in information capacity and the HOA encoding can be expressed in matrix notation;

$$\mathbf{b} = \mathbf{I_j E s}, \tag{4}$$

where $\mathbf{b}$ contains the $2N+1$ HOA B-format signals in eq. (3),

$$\mathbf{E} = \begin{bmatrix} 1 & \dots & 1 \\ \sqrt{2}j\cos(1\theta_1) & \dots & \sqrt{2}j\cos(1\theta_S) \\ \sqrt{2}j\sin(1\theta_1) & \dots & \sqrt{2}j\sin(1\theta_S) \\ \dots\dots & \dots & \dots\dots \\ \dots\dots & \dots & \dots\dots \\ \sqrt{2}j^N\cos(N\theta_1) & \dots & \sqrt{2}j^N\cos(N\theta_S) \\ \sqrt{2}j^N\sin(N\theta_1) & \dots & \sqrt{2}j^N\sin(N\theta_S) \end{bmatrix},$$

$$\mathbf{I_j} = \begin{bmatrix} 1 & 0 & \dots \\ 0 & j & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & j^N & 0 \\ \dots & 0 & j^N \end{bmatrix},$$

and $\mathbf{s}$ contains the $S$ source signals.

In order to reproduce the wave field by $M$ loudspeakers the "re-encoding principle" [11, 15] can be utilized. The loudspeaker signals are denoted the HOA D-format. The re-encoding principle leads to solving the relation

$$\mathbf{l} = \mathbf{D}\mathbf{I_j}^H \mathbf{b} = \mathbf{D}\mathbf{I_j}^H \mathbf{I_j}\mathbf{E}\mathbf{s} \tag{5}$$

where $\mathbf{l}$ contains the loudspeaker signals. This leads to the relation

$$\mathbf{D} = pinv(\mathbf{E_d}) = \mathbf{E_d}^H(\mathbf{E_d}.\mathbf{E_d}^H)^{-1} \tag{6}$$

and $\mathbf{E_d}$ is on the same form as $\mathbf{E}$ but with loudspeaker angles instead of source angles. This implies that $M \geq 2N+1$. $\mathbf{I_j}$ is often omitted in the encoding and decoding [9, 11, 14, 16].

The signal to reproduction error ($\text{SNR}_{\text{repro}}$)is given by [13]

$$\text{SNR}_{\text{repro}} = -10\log\left(1 - \sum_{n=-N}^{N} J_n(kr)^2\right) \tag{7}$$

and a rule of thumb is that if $kr < N$, the $\text{SNR}_{\text{repro}} >$ 15dB; near perfect reproduction.

## 3. QUANTIZATION ERROR IN THE SOUND FIELD

The quantization of signals is done in the frequency domain throughout this study, and the real part and the imaginary parts are quantized with the same quantization intervals (i.e., the same number of bits). The quantization of a signal $p_n$ ( implicitly dependent on frequency) is denoted as

$$p_{n,Q} = Q(p_n), \tag{8}$$

where $Q(\cdot)$ is the quantization operator. Thereby, the variance of the quantization noise $\sigma^2_{\varepsilon,p_n}$ can be derived by the using the expectation operator $E(\cdot)$;

$$\sigma^2_{\varepsilon,p_n} = E((Q(p_n) - p_n)(Q(p_n) - l_n)^*) = \frac{\Delta^2}{12}. \tag{9}$$

The last relation is valid if the quantization interval, $\Delta$ is small relative to the variance of $p_n$. The quantization noise can then be said to be uniformly distributed in amplitude and uncorrelated to the input [17] and thereby the resulting quantization noise from different input signals also uncorrelated.

The sound field in position $(r, \theta)$, in polar coordinates, resulting from $M$ plane wave radiating loudspeakers can be expressed as

$$p(r, k, \theta) = \sum_{m=1}^{M} e^{jkr\cos(\theta_m - \theta)} l_m(0, k, \theta_m) \tag{10}$$

where $l_m(0, k, \theta_m)$ is the sound pressure from the loudspeaker of index $m$.

The quantization of loudspeaker signals, the HOA D-format signals, leads to a uniform spatially distributed quantization noise by the relation

$$\begin{aligned}
&\sigma^2_{\varepsilon,lsp}(kr,\theta) = E[p(r,k,\theta)p(r,k,\theta)^*] \\
&= \sum_{n=1}^{M}\sum_{m=1}^{M} e^{jkr(\cos(\theta_m-\theta)-\cos(\theta_m-\theta))} \\
&\times E[(Q(l_m) - l_m)(Q(l_n) - l_n)^*] \\
&= M\frac{\Delta_l^2}{12}.
\end{aligned} \tag{11}$$

When quantizing the HOA B-format signals the quantization noise becomes dependent on the covariance matrix of the quantization noise.

$$\begin{aligned}
&\sigma^2_{\varepsilon,CHD}(kr,\theta) \\
&= \sum_{n=1}^{M}\sum_{m=1}^{M} e^{jkr(\cos(\theta_m-\theta)-\cos(\theta_m-\theta))} \\
&\times \sum_{v=1}^{2N+1}\sum_{u=1}^{2N+1} d_{m,v}d_{n,u} \\
&\times E[(Q(b_v) - b_v)(Q(b_u) - b_u)^*]
\end{aligned} \tag{12}$$

where $b_v$ is the v'th Ambisonics B-format signal in eq. (4) and $d_{m,v}$ is the element in column $m$, row $v$, in the decoding matrix in eq. (6).

Since $b_v$ is dependent on the encoding coefficients in $\mathbf{E}$ the dynamic range will be dependent on the number of sources, their angular extent and angles of incidence. The chance of he dynamic range of some of the $b_v$ being too low compared to the quantization interval increases if the number of sources is small, their angular extent is narrow and the HOA order is high. The quantization noise covariance matrix is therefore not necessarily a diagonal matrix, there can be some covariance. However, assuming that no angle of incidence is preferred leads to

$$\begin{aligned}
&E[(Q(b_v) - b_v)(Q(b_u) - b_u)^*] \\
&\approx \frac{\Delta_B^2(v)}{12}\delta(v-u) = \sigma^2_\varepsilon(v)\delta(v-u),
\end{aligned} \tag{13}$$

calculatign the expectation over all possible angles of incidence. Now, the covariance matrix is diagonal and only dependent on the quantization interval for each b-format signal, $\Delta_B(v)$, and this assumption is used throughout this paper for the theoretical derivations.

The expression for the spatial distribution of the quantizaion noise in eq. (12) can be explored further using the shifting property in eq. (13) and by assuming M uniformly distributed loudspeakers [11]. The second double-sum factor in eq. (12) can then be written

$$\begin{aligned}&\sum_{v=1}^{2N+1} d_{m,v}d_{n,v}\sigma_\varepsilon^2(v)\\ =&\frac{\sigma_\varepsilon^2(1)}{M^2}+\frac{2}{M^2}\sum_{v=1}^{N}\sigma_\varepsilon^2(2v)\\ &\times(\cos(\theta_m v)\cos(\theta_n v)+\sin(\theta_m v)\sin(\theta_n v))\\ =&\frac{1}{M^2}\sum_{v=-N}^{N}(\sigma_\varepsilon^2(2|v|))e^{j\frac{2\pi v(m-n)}{M}}.\end{aligned} \tag{14}$$

This last form can be identified as an inverse discrete Fourier transform (IDFT) (see eq. (27) in the Appendix) of a function

$$G(p)=\begin{cases}\sigma_\varepsilon^2(2|p|), & -N\le p\le N\\ 0, & otherwise\end{cases}, \tag{15}$$

so that

$$\sum_{v=1}^{2N+1} d_{m,v}d_{n,v}\sigma_\varepsilon^2(v)=\frac{g(m-n)}{M} \tag{16}$$

where $g$ is the IDFT of $G$ in eq. (15)

This result can be inserted in eq. (12) and using the definitions of Bessel function of first kind, $J_n(z)$, the DFT, a Fourier series and the relation between DFT and Fourier series, which all can be found in the Appendix, leads to the relation for the quantization noise

$$\begin{aligned}&\sigma_{\varepsilon,CHD}^2(kr,\theta)\\ =&\sum_{n=1}^{M}\sum_{m=1}^{M}e^{jkr(\cos(\frac{2\pi m}{M}-\theta)-\cos(\frac{2\pi n}{M}-\theta))}\frac{g(m-n)}{M}\\ =&\sum_{n=1}^{M}\sum_{m=1-n}^{M-n}\frac{g(m)}{M}e^{jkr(\cos(\frac{2\pi(m+n)}{M}-\theta)-\cos(\frac{2\pi n}{M}-\theta))}\\ =&\sum_{m=1}^{M}\frac{g(m)}{M}\sum_{n=1}^{M}e^{-j2kr\sin(\frac{\pi(m+2n)}{M}-\theta)\sin(\frac{\pi m}{M})}\\ =&\sum_{m=1}^{M}g(m)\sum_{n=-\infty}^{\infty}J_{nM}(2kr\sin(\frac{2\pi m}{M}))e^{jn(\pi m-M(\theta+\pi))}.\end{aligned} \tag{17}$$

The expression in eq. (17) is a Fourier series representation of a harmonic signal that has been sampled at a rate of $M$ samples/period and filtered with the discrete filter $G(p)$. The sampling theorem will be violated for $kr>\frac{M}{2}$ because Bessel terms of order other than 0 will contribute significantly and aliasing effects leading to dependencies in $\theta$ will occur.

## 4. SPATIAL NOISE SHAPING

Assuming that the listener is seated in the near perfect reproduction area $kr<N$ is fulfilled and since the magnitude of $J_n(z)$ are small for $z<n$ and the criterion $M>2N+1$ must be fulfilled [11] the quantization noise expressed in eq. (17) becomes

$$\begin{aligned}&\sigma_{\varepsilon,CHD}^2(kr,\theta)\\ &\approx\sum_{m=1}^{M}g(m)J_0(2kr\sin(\frac{2\pi m}{M})).\end{aligned} \tag{18}$$

This implies that there will be no angular dependency for the error in the near perfect reproduction area.

The expected radial error for all $kr$ is found by averaging eq. (17) over $\theta$

$$\begin{aligned}\sigma_{\varepsilon,CHD}^2(kr)&=\frac{1}{2\pi M}\sum_{m=1}^{M}g(m)\\ &\times\sum_{n=1}^{M}\int_{-\pi}^{\pi}e^{-j2kr\sin(\frac{\pi(m+2n)}{M}-\theta)\sin(\frac{\pi m}{M})}d\theta\\ &=\sum_{m=0}^{M}g(m)J_0[2kr\sin(\pi\frac{m}{M})]\end{aligned} \tag{19}$$

which obviously corresponds to the expression in eq. (18) for $kr<N$ where there is no angular dependency.

The quantization noise can therefore be shaped radially by altering $g(m)$ which corresponds to shaping the quantization noise over the HOA orders (see eqs. (15) and (27)).

### 4.1. Uniformly distributed number of bits

Matching the number of loudspeakers to the HOA order so that $M=2N+1$ and using uniform bit allocation so that $\sigma_{\varepsilon,p}^2=\frac{\Delta_B^2}{12}$ results in $g(n)=\frac{\Delta_B^2}{12}\delta(n)$ and $\sigma_{\varepsilon,CHD}^2(kr,\theta)=\frac{\Delta_B^2}{12}$ for all $kr$ and $\theta$

Using eq. (11) suggests a $\frac{M\Delta_l^2}{\Delta_B^2}$ signal to quantization noise gain factor relative to the quantization of loudspeaker signals. However, the real dynamic range is highly dependent on the sound field, and a rate/distortion discussion of frequency domain quantized B- and D-format signals without entropy encoding doesn't make sense. This is out of scope of this work but is addressed in an accompanying paper [18].

The $\text{SNR}_{\text{repro}}$ is about 15dB on the $kr = N$ boundary and this SNR value can be used as a lower boundary for the signal to total error in the near perfect reproduction area. The signal to quantization noise ratio (SQNR) is dependent on the statistical properties of the signal. Presuming Gaussian distribution and a range of $\pm 3\sigma^2 x$ which would lead to an overload for less than 0.3% of the samples leads to $S\text{QNR} = 6.02b + 1.25$ [19]. This implies that as long as $b > 2$ then SQNR $> 15$ which suggests that as few as 3 bits per sample can be employed. The total error is, however, the sum of the quantization noise and the noise from the truncation in HOA.

A higher total signal to noise ratio can be achieved by increasing the number of bits and HOA order according to eq. (7).

### 4.2. Decreasing number of bits by order

Using a linearly decreasing bit allocation across the HOA orders so that

$$\sigma_\varepsilon^2(2v) = \begin{cases} \frac{2^{2(v-\lfloor \frac{N}{2} \rfloor + 1 - b)}}{12}, & 0 \le v \le \lfloor \frac{N}{2} \rfloor \\ \frac{2^{2(v-\lceil \frac{N}{2} \rceil + 1 - b)}}{12}, & \lceil \frac{N}{2} \rceil \le v \le N \end{cases} \tag{20}$$

leads to the spatial distribution of quantization noise observed in figs. 1 and 2.

It can be seen in fig. 2 that the error level for the linearly decreasing bit allocation relative to the uniform distribution crosses 0 dB around $kr \approx \frac{N}{2}$, i.e., the quantization noise for a linearly decreasing bit allocation is lower than $\frac{\Delta_B^2}{12}$ for $kr \lessapprox \frac{N}{2}$. Furthermore, the relative error level increases linearly up to $kr = N$. Since the quantization is performed in the frequency domain this means that the quantization noise will increase up to $r \approx \frac{54N}{f}$ where $f$ is the quantized frequency.

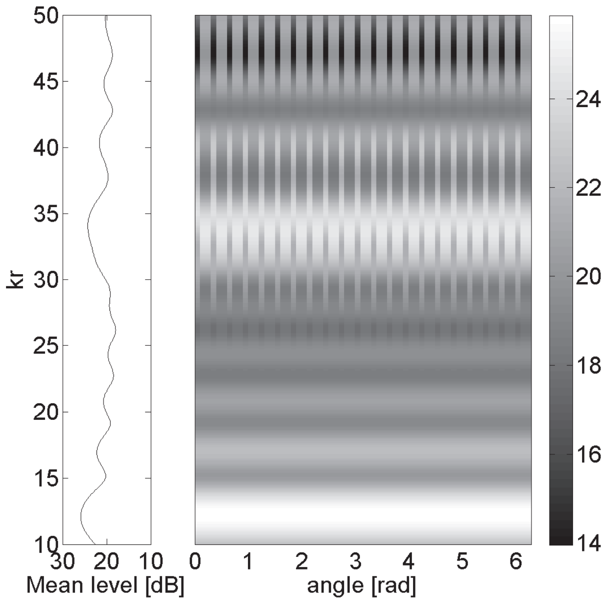


**Fig. 1:** Spatial distribution of the quantization error for a linearly decreasing bit allocation, given in eq. (20), relative to a uniform bit allocation. The HOA order N = 10.

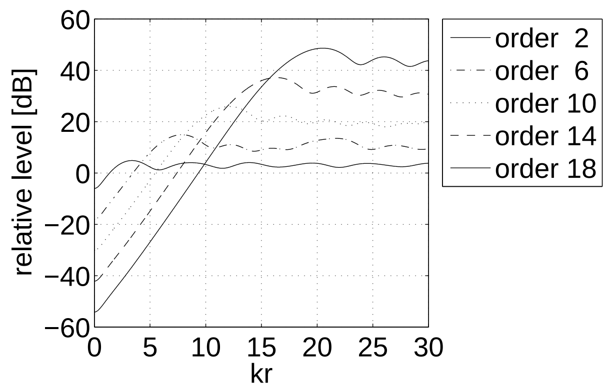


**Fig. 2:** Mean quantization error for a linearly decreasing bit allocation, given in eq. (20), relative to a uniform bit allocation.

Regarding the fact that $J_nM[2kr\sin(\frac{\pi m}{M})]$ in eq. (19) will have a non-zero value only for $m = 0$ when $kr \rightarrow \infty$ means that the noise will converge to

$$\lim_{kr\rightarrow\infty} \frac{\sigma^2_{\varepsilon,CHD,var}(kr)}{\sigma^2_{\varepsilon,CHD,uni}(kr)} = \begin{cases} \frac{2^{-(N)}(2^{2N+3}-5)}{3M}, & Neven \\ \frac{2^{-(N)}(2^{2(N+1)}-3\times 2^{(N+1)}+5)}{3M}, & Nodd \end{cases} \tag{21}$$

This implies that higher order leads to a higher quantization noise in the non-perfect reconstruction area. This could be compensated for by increasing $M$, i.e. using more loudspeakers will attenuate the quantization noise. However, the HOA is formulated from coherent summation of loudspeaker signals and the summation will become incoherent as $kr \rightarrow \infty$. This means that using more loudspeakers will also attenuate the signal level leading to the same SQNR.

### 4.3. Frequency dependent bit allocation

The fact that near perfect reproduction only can be achieved for $kr < N$ motivates for high-pass filtering the HOA signals [16] with filters having cut-off frequency at $f_c \approx \frac{54n}{r_{max}}$ . This is analog to assigning 0 bits to the HOA orders over $N_c = \lceil \frac{f_i r_{max}}{54} \rceil$ when encoding HOA. Here, $f_i$ is the highest frequency in frequency band $i$.

This can also lead to a lower reproduction error for low frequencies provided the same number of loudspeakers is used for the whole frequency range, i.e. if more loudspeakers than necessary is employed for low frequencies.

Consider the case of $M$ uniformly distributed loudspeakers on a circle reproducing a monochromatic wave of amplitude $A_1$ from direction $\theta_1$ represented in HOA order $N$. The sound pressure can be expressed as

$$p(r,k,\theta)_A = A_1 \sum_{l=-N}^{N} e^{jl(\theta_1-\theta)} \times \sum_{n=-\infty}^{\infty} j^{nM+l} J_{nM+l}(kr) e^{-jnM\theta}. \tag{22}$$

The Bessel terms which orders absolute value is larger or equal to $M - N$ will have a destructive effect, i.e. the phase will be altered by the $e^{-jnM\theta}$ term. This effect will be largest at the near perfect reconstruction boundary $kr \approx N$ and the problem is analogous to the

**Table 1:** High-pass cut-off frequencies

| order | $f_c$ [Hz] |
|---|---|
| **0** | 0.000 |
| **1** | 85.000 |
| **2** | 169.000 |
| **3** | 254.000 |
| **4** | 423.000 |
| **5** | 508.000 |
| **6** | 592.000 |
| **7** | 762.000 |
| **8** | 846.000 |
| **9** | 931.000 |
| **10** | 1100.000 |

sampling theorem and aliasing [19]. The effect will be smaller if the number of loudspeakers, $M$, relative the order, $N$, increases which follows from the property of the Bessel function of first kind

$$|J_n(x)| < |J_m(x)|, \; |x| < |m| < |n|, \tag{23}$$

which is analogous to over-sampling.

However, one could also match the number of loudspeakers to the order used to represent each frequency region. This would lead to a reduced number of required loudspeakers at lower frequencies. This suggests a trade-off between reproduction accuracy and the number of loudspeakers used in the low frequency region.

#### 4.3.1. Case: Quantizing a pulse

As a numerical example, the case of a plane wave pulse is studied. The pulse is generated at 2.2 kHz sampling frequency and truncated to 26 samples. The source direction chosen is indicated in fig. 3, which shows a snapshot of the soundfield over a circle of radius 1 m. The signal is transformed by a DFT and HOA encoded at order N = 10 which should give a near-perfect reproduction radius of r $\approx$ 0.5 m for the highest frequency of 1.1. kHz. Both the real and imaginary DFT coefficients are quantized by three or four bits resolution. High-pass filtering of the B-format orders as shown in table 1is employed as well. Decoding is done using 21 loudspeakers and fig. 4 shows a snapshot for basic HOA decoding and fig. 5 shows a snapshot for the high-pass filtered HOA. The angle-averaged errors are calculated and presented in figs. 6 and 7. For the non-quantized cases the error equals the reproduction error, while the quantized cases will display a sum of reproduction error and quantization

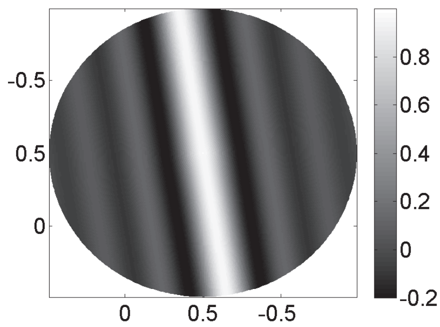


**Fig. 3:** Original wave.

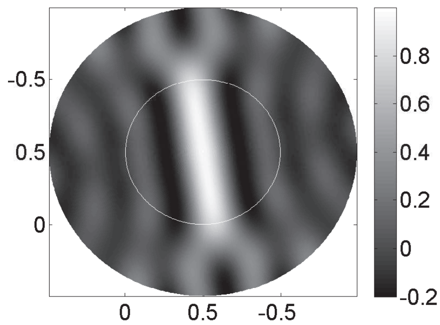


**Fig. 4:** Basic HOA order 10 highest frequency 1.1 kHz, reproduction error<-15 dB within white circle (r=0.5 m).

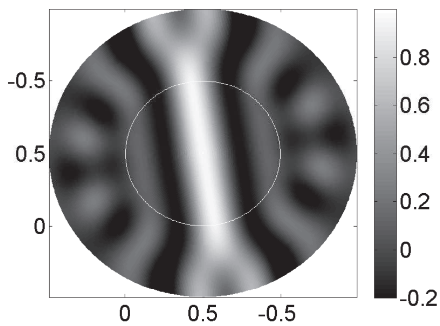


**Fig. 5:** High-pass filtered HOA order 10 highest frequency 1.1 kHz, reproduction error<-15 dB within white circle (r=0.5 m).

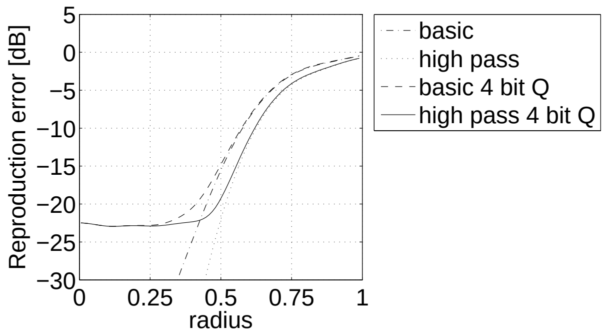


**Fig. 6:** Radial error for HOA order 10 at the frequency 1.1 kHz. The near-perfect reproduction radius is 0.5 m.

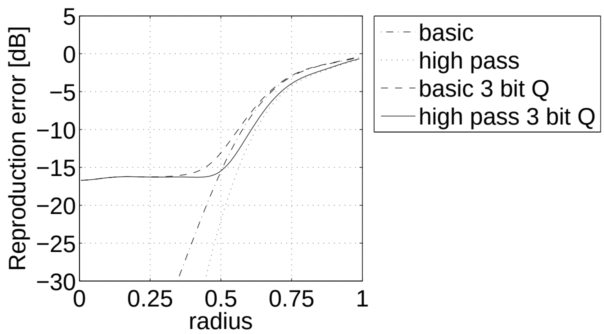


**Fig. 7:** Radial error for HOA order 10 at the frequency 1.1 kHz. The near-perfect reproduction radius is 0.5 m.

error. As mentioned in section 4.3, and as can be seen in figs. 6 and 7, the error at the kr = N boundary is lower for the high-pass filtered HOA.

From fig. 6 it can be seen that utilizing four bits per sample will lead to no larger error in the reproduction area than the error at the $kr = N$ boundary for both basic HOA and high-pass filtered HOA. Figure 7 shows that utilizing three bits leads to a larger reproduction noise for basic HOA but applied to the high-pass filtered HOA, the reproduction error stays -15 db below the signal level.

## 5. CONCLUSION

The reproduction error can be shaped radially in the near perfect reproduction area by assigning more bits to the lower orders than the higher ones. Uniformly distributed bits across order leads to uniformly distributed quantization noise. The number of bits needed in order to keep the error below 15 dB, the HOA error for $kr = N$, is dependent on the signal distribution and quantization range but can typically be about four bits. Furthermore, utilizing high-pass filtering of the HOA channels leads to a larger reproduction radius if utilizing the same number of loudspeakers for all frequencies. High-pass filtering opens therefore up for removing irrelevancy and further reducing the number of bits representing the remaining information to about three bits.

These results suggest that the large amount of bandwidth needed for representing wave fields by HOA can be highly reduced.

## 6. REFERENCES


[1] A. Spanias, T. Painter, and V. Atti. *Audio Signal Processing and Coding*. Wiley-Interscience, 2007.

[2] F. Baumgarte and C. Faller. Binaural cue coding-Part I: psychoacoustic fundamentals and design principles. *IEEE transactions on speech and audio processing*, 11(6):509–519, 2003.

[3] J. Merimaa and V. Pulkki. Spatial Impulse Response Rendering I: Analysis and Synthesis. *J. Aud. Eng. Soc.*, 53(12):1115–1127, December 2005.

[4] V. Pulkki and J. Merimaa. Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests. *J. Aud. Eng. Soc.*, 54(1/2):3–20, January/February 2006.

[5] V Pulkki and C Faller. Directional Audio Coding: Filterbank and STFT-Based Design. In Proc. of the 120th AES Convention, Paris, France, May 20-23 2006. Preprint 6658.

[6] J. Breebaart, J. Herre, C. Faller, J. Rödén, F. Myburg, S. Disch, H. Purnhagen, G. Hotho, M. Neusinger, K. Kjörling, and W. Oomen. MPEG Spatial Audio Coding / MPEG Surround: Overview and Current Status. In Proc. of the 119th AES Convention, New York, USA, October 7-10 2005. Preprint 6599.

[7] A. J. Berkhout. A holographic approach to acoustic control. *J. Audio Eng. Soc*, 36(12):977–995, December 1988.

[8] M. M. Boone and E. N. G. Verheijen. Multichannel sound reproduction based on wavefield synthesis. In *The Convention 95th AES Conv.*, October 1993. Preprint 3719.

[9] S. Moreau, J. Daniel, and S. Bertet. 3D Sound Field Recording with Higher Order Ambisonics - Objective Measurements and Validation of Spherical Microphone. In *Audio Engineering Society's 120th Convention 2006 May 20-23 Paris, France*, 2006. Preprint 6857.

[10] R. Nicol and M. Emerit. 3d-sound reproduction over an extensive listening area: A hybrid method derived from holophony and ambisonic. In *Proc. of the 16th AES International Conference*, March 1999. Paper Number 16-039.

[11] J. Daniel, S. Moreau, and R. Nicol. Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging. In *The 114th AES Conv.*, February 2003. Preprint 5788.

[12] M. A. Poletti. A unified theory of horizontal holographic sound systems. *J. Aud. Eng. Soc.*, 48(12):1155–1182, 2000.

[13] T.D. Ward, D.B.; Abhayapala. Reproduction of a plane-wave sound field using an array of loudspeakers. *Speech and Audio Processing, IEEE Transactions on*, 9(6):697–707, Sep 2001.

[14] B. Støfringsdal and U. P. Svensson. Conversion of discretely sampled sound field data to auralization formats. *J. Aud. Eng. Soc.*, 54(5):380–400, May 2006.

[15] M. A. Poletti. Three-dimensional surround sound systems based on spherical harmonics. *J. Aud. Eng. Soc.*, 53(11):1004–1025, November 2005.

[16] J. Daniel and S. Moreau. Further study of sound field coding with higher order ambisonics. In *The 116th AES Conv.*, May 2004. Preprint 6017.

[17] S. P. Lipshitz, R. A. Wannamaker, and J. Vanderkooy. Quantization and dither: A theoretical survey. *J. Audio Eng. Soc*, 40(5):355–375, May 1992.

[18] E. Hellerud, I. Burnett, A. Solvang, and U. P. Svensson. Encoding higher order ambisonics with aac. In *in Proc. of the 112th AES Conv.*, May 2008.

[19] J. G. Proakis and D. G. Manolakis. *Digital Signal Processing: Principles, Algorithms, and Applications*. Prentice Hall, Upper Saddle River, NJ, 3rd edition, 1996.


## APPENDIX

### The Fourier series and its inverse

$$W(n) = \mathcal{F}(w(\theta)) = \frac{1}{2\pi}\int_{-\pi}^{\pi} w(\theta)e^{-jn\theta}d\theta \qquad (24)$$

$$w(\theta) = \mathcal{F}^{-1}(W(n)) = \sum_{n=-\infty}^{\infty} W(n)e^{jn\theta} \qquad (25)$$

### Discrete Fourier transform

$$F(m) = \mathit{DFT}_M(f(n)) = \sum_{n=0}^{M-1} f(n)e^{-j\frac{2\pi mn}{M}} \qquad (26)$$

$$f(n) = \mathit{IDFT}_M(F(m)) = \frac{1}{M}\sum_{m=0}^{M-1} F(m)e^{j\frac{2\pi mn}{M}} \qquad (27)$$

Relation to Fourier series

$$F(m) = M\sum_{n=-\infty}^{\infty} W(m+nM) \qquad (28)$$

### Bessel function of first kind

$$J_n(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-jn\theta+jz\sin(\theta)}d\theta = \mathcal{F}(e^{jz\sin(\theta)}) \qquad (29)$$

Jacobi-Anger Expansion:

$$e^{jz\cos\theta} = \sum_{n=-\infty}^{\infty} j^n J_n(z)e^{jn\theta} = \mathcal{F}^{-1}(j^n J_n(z)) \qquad (30)$$