ERROR ANALYSIS OF SPHERICAL HARMONIC SOUNDFIELD REPRESENTATIONS IN TERMS OF TRUNCATION AND ALIASING ERRORS

Stefanie Brown[⋆] *and Deep Sen*[⋆†]

[⋆]Acoustics Laboratory, School of Electrical Engineering and Telecommunications
The University of New South Wales, Sydney, Australia
Email: {stefanie, d.sen}@unsw.edu.au

[†]Qualcomm Inc., Audio Research & Development, San Diego, CA, USA

ABSTRACT

The use of the spherical harmonic representation of a soundfield is useful when attempting to record, reproduce or manipulate the spatial qualities of the soundfield. However, the practical requirement of discrete sampling in the spatial domain brings errors to the system, namely those of truncation and spatial aliasing. The truncation error can be seen in the synthesized pressure, while spatial aliasing is apparent when looking at the spherical harmonic coefficients themselves. These errors are linked to each other through the number and position of the microphones in the array, as well as the method used to perform numerical integration on the sphere, but they can exist separately. This paper discusses the above topics and investigates two approaches to numerical integration in regards to sampling the soundfield using an em32 Eigenmike® microphone array.

Index Terms— Spherical microphone array, spatial aliasing error, truncation error, spherical harmonics, quadrature coefficients.

1. INTRODUCTION

Spherical microphone arrays have a wide variety of uses in spatial audio scenarios, such as beamforming [1], room acoustics measurement, soundfield recording and Higher Order Ambisonics [2] amongst others, due to the ease of their use with the spherical harmonic representation of a soundfield. This representation allows the soundfield to be described spatially in terms of radius, azimuth and inclination, and thus a soundfield can be decomposed into spatial components. The advantage of a spherical array over other shapes is that it will have high spatial resolution in all directions and the geometry allows the use of the orthonormality properties of spherical harmonics to aid in extracting the coefficients of the representation.

Several types of errors are involved with the spherical harmonic representation that include some that arise from mathematical approximations, such as spatial aliasing [2, 3, 4], errors due to the limitations of physical array set up (number of elements and positioning) e.g. truncation error [5, 6, 7, 8, 9] and those stemming from the non-idealities associated with the construction of the array and the microphones themselves [10]. Minimisation of the orthonormality error, a component of the spatial aliasing error, has been used to select optimal numerical integration, or quadrature, coefficients for non-uniformly spaced array designs [11].

Truncation and spatial aliasing errors are in a practical sense closely related to the array geometry and numerical integration coefficients used, however they are separable in theory. This paper explicitly defines the aliasing and truncation errors as separate entities. It goes on to investigate an example of the spherical array em32 Eigenmike® for two particular sets of quadrature coefficients and comments on the relative benefits of each.

2. BACKGROUND

We have found that conventions and notations used in this field vary widely between and even amongst research groups. As such, we find it prudent to specify the assumptions we have made to avoid confusion.

2.1. Conventions and Definitions

Spherical coordinates are defined by $\mathbf{r} = (r, \theta, \phi)$ as the radius, angle of inclination from the positive z-axis and angle of azimuth from the positive x-axis respectively.

We define the pressure function of a unit amplitude spherical wave in (1), where ω and k are the angular frequency and wave number and $\mathbf{r_s} = (r_s, \theta_s, \phi_s)$ is the source location.

$$p(t, k, \mathbf{r}) = \frac{exp\{i(\omega t - k|\mathbf{r} - \mathbf{r_s}|)\}}{|\mathbf{r} - \mathbf{r_s}|} \tag{1}$$

$P(k, \mathbf{r})$ is the Fourier transform of $p(t, k, \mathbf{r})$, taken using the engineering convention where $S(\omega) = \int_{-\infty}^{\infty} s(t) e^{-i\omega t} dt$. Expanding $P(k, \mathbf{r})$ in terms of r, θ and ϕ, we arrive at the general interior solution to the wave equation in a non-scattering environment, valid where $r < r_s$ [12]

$$P(k, \mathbf{r}) = \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^{n} A_n^m(k) Y_n^m(\theta, \phi) \tag{2}$$

where $j_n(kr)$ is the spherical Bessel function, $Y_n^m(\theta, \phi)$ is the spherical harmonic of order n and degree m, or mode (n, m) and $A_n^m(k)$ are source-dependent spherical harmonic coefficients. There are several definitions for spherical harmonics, but we choose to use the following, which includes the associated Legendre function $P_n^{|m|}(\cos\theta)$.

$$Y_n^m(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos\theta) e^{im\phi} \tag{3}$$

The theoretical A_n^m for a point source at $\mathbf{r_s}$ are defined as follows to conform with the engineering Fourier transform convention, and include the spherical Hankel function of the second kind $h_n^{(2)}(kr_s)$. $*$ indicates the complex conjugate.

$$A_n^m(k) = -4\pi ik h_n^{(2)}(kr_s) Y_n^{m*}(\theta_s, \phi_s) \tag{4}$$

2.2. Scattering from a rigid microphone array

An array with microphones placed on the surface of a rigid sphere of radius a acts as a scatterer. This results in a compound soundfield, where the $j_n(kr)$ term in (2) is replaced by $b_n(kr, ka)$ [12]

$$b_n(kr, ka) = j_n(kr) - \frac{j_n'(kr)}{h_n^{(2)'}(ka)} h_n^{(2)}(kr), \tag{5}$$

where the first term represents the incident field, the second term the scattered field and $'$ indicates the derivative of the function with respect to kr. This term can be simplified by the Wronskian relationships at $r = a$ to

$$b_n(ka) = \frac{-i}{(ka)^2 h_n^{(2)'}(ka)} \tag{6}$$

2.3. Orthonormal property of Spherical Harmonics

Spherical harmonic functions are orthonormal when multiplied and integrated over the sphere. This can be expressed for the continuous case by an integral over the sphere $\int d\Omega$, as in (7), and approximated for the discrete case by a sum over the microphone positions indicated by subscript s in (8).

$$\int Y_n^m(\theta, \phi) Y_{n'}^{m'*}(\theta, \phi) d\Omega = \delta_{nn'} \delta_{mm'} \tag{7}$$

$$\sum_{s=1}^{S} \alpha_s Y_n^m(\theta_s, \phi_s) Y_{n'}^{m'*}(\theta_s, \phi_s) = \delta_{nn'} \delta_{mm'} + \epsilon_{nn'}^{mm'} \tag{8}$$

where α_s are the numerical integration, or quadrature coefficients, for the chosen sampling points (θ_s, ϕ_s). Quadrature refers to performing a numerical integration of a function over a surface accurately up to a certain function order. $\epsilon_{nn'}^{mm'}$ refers to the orthonormality error between modes (n, m) and (n', m').

The α_s are chosen such that $\epsilon_{nn'}^{mm'} = 0$ for $n,\ n' \leq N$, the truncation order. It is therefore possible to extract the A_n^m's from pressure signals in the spatially continuous/discrete case using the following integral/sum:

$$A_n^m(k) = \frac{1}{b_n(ka)} \int Y_n^{m*}(\theta, \phi) P(k, a, \theta, \phi) d\Omega \tag{9}$$

$$A_n^m(k) = \frac{1}{b_n(ka)} \sum_{s=1}^{S} \alpha_s Y_n^{m*}(\theta_s, \phi_s) P(k, a, \theta_s, \phi_s) \tag{10}$$

3. SEPARATION OF THE TRUNCATION AND SPATIAL ALIASING ERRORS

A spherical wave has an infinite representation in the spherical harmonic domain. We often choose, or are forced by the physical constraints of an array, to limit that representation to a particular order N, often called the truncation or array order. This limitation leads to both truncation and spatial aliasing errors. Several authors have investigated truncation and spatial aliasing errors both theoretically and practically. They are closely related, but occur due to different causes. This section will discuss each type of error individually and how they both manifest in a practical scenario.

3.1. Truncation error

The truncation error arises when a limit is placed on the infinite sum in (2) such that n ranges from 0 to some order N, i.e. $P(k, \mathbf{r}) = P_N(k, \mathbf{r}) + P_{n>N}(k, \mathbf{r})$.

$$P_N(k, \mathbf{r}) = \sum_{n=0}^{N} j_n(kr) \sum_{m=-n}^{n} A_n^m(k) Y_n^m(\theta, \phi) \tag{11}$$

$$P_{n>N}(k, \mathbf{r}) = \sum_{n>N} j_n(kr) \sum_{m=-n}^{n} A_n^m(k) Y_n^m(\theta, \phi) \tag{12}$$

Hypothetically, if a continuously pressure-sensitive spehrical array existed it would be possible to extract all A_n^m's exactly, i.e. with no spatial aliasing, but for computational or data storage purposes we might still wish to truncate the series. In terms of resynthesizing a soundfield from the stored coefficients, since the $j_n(kr)$ term becomes smaller with increasing order n, high order modes contribute less information to the pressure signal while requiring many more coefficients. As such there are diminishing returns to increasing the truncation order N.

Several closed-form solutions for bounding functions of the truncation error have been derived for both plane and spherical waves [7, 8, 9], given certain assumptions. All are derived from (12), the tail end of the infinite summation over n. These authors used theoretical spherical harmonic coefficients in their calculations, and as such do not include any spatial aliasing. Other authors include a Gaussian noise term derived from the effects of the $N + 1^{\text{th}}$ order modes [5, 6], and thus include spatial aliasing errors within the truncation error definition.

3.2. Spatial aliasing errors

Spatial aliasing manifests as a difference between the theoretical (A_n^m) and the array-extracted $(\hat{A}_n^m)$ spherical harmonic coefficients of a signal, $\widetilde{A}_n^m$. It generally occurs when high order modal components of the signal are shifted to low order coefficients, due to discrete orthonormality errors. Spatial aliasing will not occur if the signal in question is spatially bandlimited, or essentially truncated, to an order lower than the order of the array.

The aliasing effect of all higher order modes into the mode (n, m), designated $\widetilde{A}_n^m(k)$, is related to the truncation error through $P_{n>N}$ and is dependent upon the number of sampling points S, their positions (θ_s, ϕ_s) and the quadrature coefficients α_s. As such, good choices for these parameters are important.

$$\widetilde{A}_n^m(k) = \frac{1}{b_n(ka)} \sum_{s=1}^{S} \alpha_s Y_n^{m*}(\theta_s, \phi_s) P_{n>N}(ka, \theta_s, \phi_s) \tag{13}$$

Spatial aliasing can occur when there is no truncation error. Equation (10) permits the extraction of all the soundfield coefficients, such that no truncation error exists, however the accuracy of the extracted coefficients is determined by how closely the chosen microphone positions and quadrature coefficients approximate orthonormality, which is generally not well for increasing order.

Various sampling schemes requiring greater or fewer points and their impact on spatial aliasing have been investigated in [3, 10]. The particular scheme of interest to us is that used by the em32 and this will be covered in greater detail in Section 4.

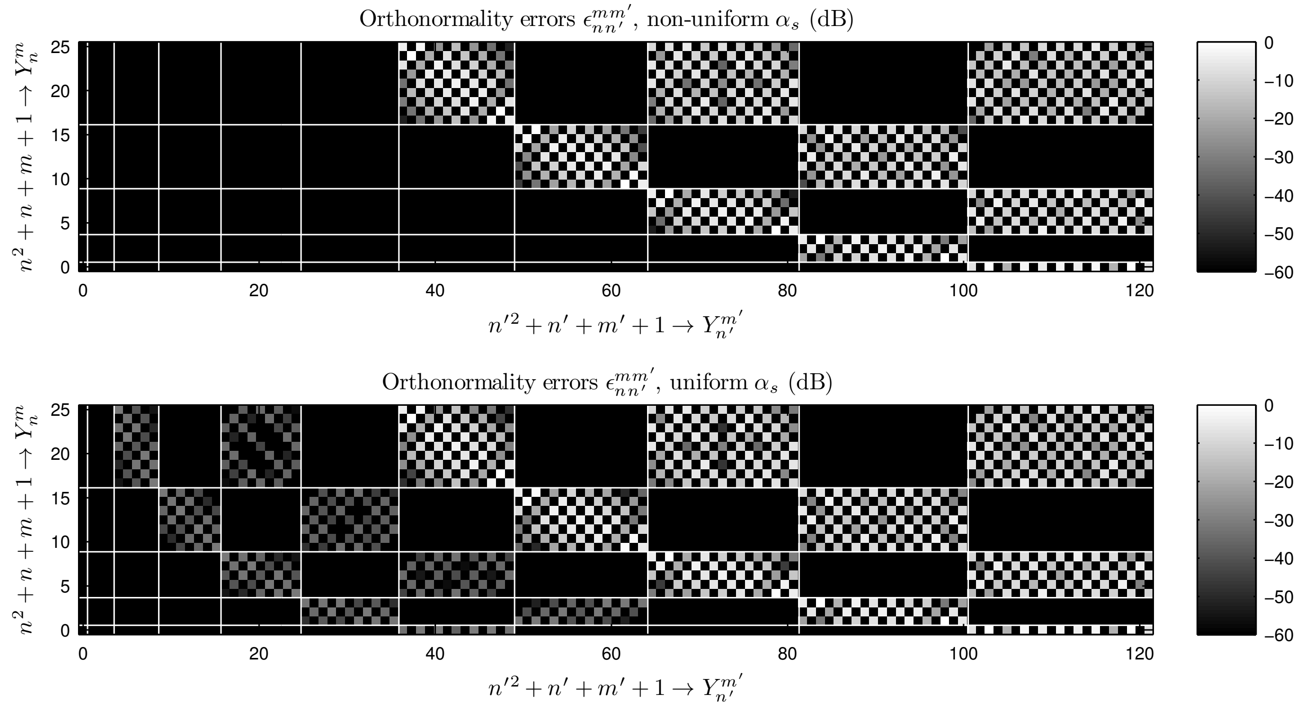

Fig. 1. Orthonormality errors $\epsilon_{nn'}^{mm'}$ for the em32 array using *'non-uniform'* (top) and *'uniform'* (bottom) α_s.

3.3. Choice of array/truncation order

The truncation order N is generally described as being limited by the number of elements in the array, $(N+1)^2 \leq S$ [12]. This lower bound is only met if the method used to extract A_n^m's does not make use of the orthonormality principle. A method to extract N^{th} order coefficients involving matrix inversion requires at least $(N+1)^2$ sampling points so that the system of equations to be solved is not underdetermined [8]. When exploiting the orthonormality property the number of elements required to achieve good extraction up to order N is often much higher and could be greater than the minimum number of points to extract $N+1$ orders [10, 13].

In practice, once a particular array is chosen the tradeoff between truncation error and spatial aliasing error is what really determines the chosen truncation order. The following section will discuss in detail the nearly-uniform sampling scheme of the em32 spherical array and two different choices of quadrature coefficients.

4. SPATIAL ALIASING ERRORS WITH THE EM32

The em32 Eigenmike®is a 4^{th}order array consisting of 32 omnidirectional electret microphones positioned on the surface of a rigid sphere. Their radii are 4.2 cm and (θ, ϕ) coordinates correspond to the face-centers of a truncated icosahedron. This sampling scheme is deemed nearly uniform and uses the least microphones necessary to obtain a accurate 4^{th}order extraction [10]. We have two options when choosing the quadrature coefficients α_s that give slightly different results. They can uniformly be set to $4\pi/32$ as in [4], or alternatively can vary with position (*'non-uniform'*) and be $9\pi/70$ for points corresponding to the twenty hexagonal faces and $5\pi/42$ for the twelve pentagonal faces [14].

Let us return to equation (13) which decribes the aliasing of all higher order modes $(n > N, m)$ into a low order mode (n, m). By looking at the aliasing of a single higher order mode $(\mathcal{N}, \mathcal{M})$ into a particular lower order mode (n, m), designated $\widetilde{A}_{n,\mathcal{N}}^{m,\mathcal{M}}(k)$, we can see that it is dependent on the orthonormality error for those two modes (and thus on the choice of α_s) and the ratio of the $b_n(ka)$ terms of both modes.

$$\begin{aligned}\widetilde{A}_{n,\mathcal{N}}^{m,\mathcal{M}}(k) &= \frac{1}{b_n(ka)} \sum_{s=1}^{S} \alpha_s Y_n^{m*}(\theta_s, \phi_s)\ A_{\mathcal{N}}^{\mathcal{M}}(k) b_{\mathcal{N}}(ka) Y_{\mathcal{N}}^{\mathcal{M}}(\theta_s, \phi_s) \\ &= A_{\mathcal{N}}^{\mathcal{M}}(k) \frac{b_{\mathcal{N}}(ka)}{b_n(ka)} \epsilon_{n\mathcal{N}}^{m\mathcal{M}} \qquad (14)\end{aligned}$$

where $\epsilon_{n\mathcal{N}}^{m\mathcal{M}}$ is the orthonormality error between modes (n, m) and $(\mathcal{N}, \mathcal{M})$. Thus we can rewrite (13) as

$$\widetilde{A}_n^m(k) = \sum_{\mathcal{N}>N} \frac{b_{\mathcal{N}}(ka)}{b_n(ka)} \sum_{\mathcal{M}=-\mathcal{N}}^{\mathcal{N}} A_{\mathcal{N}}^{\mathcal{M}}(k) \epsilon_{n\mathcal{N}}^{m\mathcal{M}} \qquad (15)$$

The following sections will investigate the impact of both choices of α_s on the spatial aliasing equation.

4.1. Orthonormality Errors

Figure 31 shows the patterns in the orthonormality errors $\epsilon_{nn'}^{mm'}$ from (8) that occur when using *'uniform'* or *'non-uniform'* quadrature coefficients. The aliasing patterns and magnitude of the orders greater than 5 are similar in both cases. The main difference between the two methods is that with *'uniform'* α_s and for $n \leq 4$, the orthonormality error is not vanishingly small at roughly -40 dB. However compared to the higher order errors, it seems that this could be insignificant.

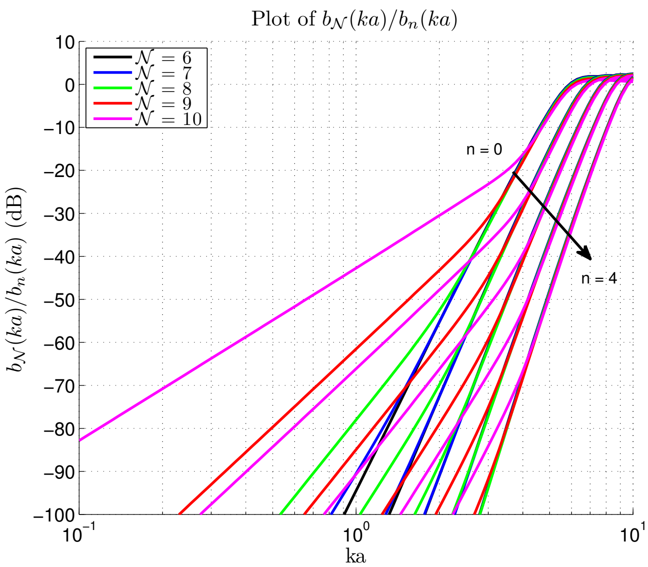

Fig. 2. $b_{\mathcal{N}/n}(ka)$ plotted for $n = 0:4$ and $\mathcal{N} = 6:10$ for varying ka. Please view the PDF to see in color.

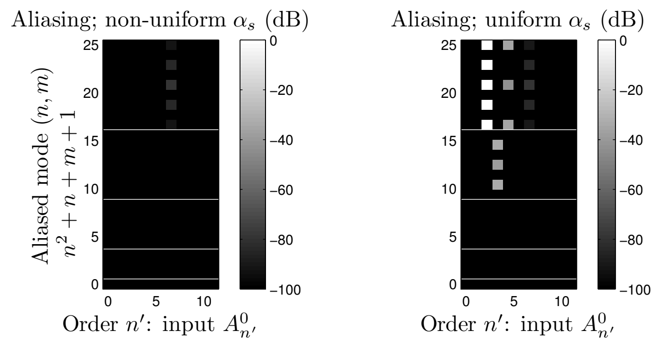

Fig. 3. Spatial aliasing for each mode $A^0_{n'}$ of a 2 kHz point source at $(1, 0, 0)$ when using '*non-uniform*' (left) and '*uniform*' (right) α_s.

In both cases, the stepped pattern of larger errors appearing with increasing order n' is caused by the fact that this combination of microphone positions and quadrature coefficients allow functions of up to 9^{th}order to be integrated accurately, i.e. when $n+n' \leq 9$ since the integrated function is the multiple of two spherical harmonics. It is also observed that larger orthonormality errors occur when both $n + n'$ and $m + m'$ are even. The quadrature coefficients perform better when either $n+n'$ or $m+m'$ are odd, even when beyond the accurate integration limit of the array.

4.2. $\mathbf{b_n(ka)}$ modification of the orthonormality error

We will now look at the magnitude of the portion of the aliasing term involving $b_n(ka)$. Using (6), this simplifies to the following, which we will refer to as $b_{\mathcal{N}/n}(ka)$ henceforth.

$$\frac{b_{\mathcal{N}}(ka)}{b_n(ka)} = \frac{h_n^{(2)'}(ka)}{h_{\mathcal{N}}^{(2)'}(ka)} \tag{16}$$

Figure 2 shows how the absolute value of this function behaves with $n = 0 : N$, $\mathcal{N} > N$ and ka. As both n and $\mathcal{N}$ increase, the magnitude decreases for a specific ka. This graph can be used to determine frequencies at which particular orders may begin to have significant aliasing, as long as there is a corresponding high orthonormality error.

4.3. Spatial aliasing of a spherical wave

To simplify testing the spatial aliasing errors due to the two different sets of quadrature coefficients, we will use a 2 kHz spherical wave positioned at $(1, 0, 0)$. This means for any mode where $m \neq 0$, the spherical harmonic coefficient will be zero, which allows us to observe the minimal number of coefficients aliasing to other modes.

Figure 3 was created by extracting all modes up to 4^{th} order from a signal comprised of a single modal component of the wave described above, i.e. implementing (14) for $\mathcal{N} = 0 : 10$, $\mathcal{M} = 0$, then subtracting the theoretical A_n^m from the appropriate modes to observe the error. It can be seen that the 6^{th} order mode aliases into the 4^{th} slightly (around -70 dB) in both cases, in a pattern following that shown in the Figure 3. In the case of '*uniform*' α_s, aliasing of both 3^{rd} and 4^{th} order modes into modes of the same order but different degrees, while none appears for the '*non-uniform*' case. The patterns we see here again match the error patterns shown in Figure 1, but with the added effect of the $b_{\mathcal{N}/n}(ka)$ term. In Figure 2, $\mathcal{N}$ was selected to be greater than the array order N and it was seen that with increasing order $\mathcal{N}$, the overall affect of aliasing on the array decreases for a particular frequency. However if we allow $\mathcal{N} < n$, the term becomes large. The -40 dB orthonormality errors for this region seem insignificant on their own, but when combined with the $b_{\mathcal{N}/n}(ka)$ term become significant. This shows how important it is that the orthonormality error should be zero for n and $n' < N$, as stated in Section 2.3 - for '*uniform*' α_s this is simply not the case.

In terms of the higher orders, we can see that the $b_{\mathcal{N}/n}(ka)$ term decays after 6^{th} order to counteract the higher orthonormality errors for those orders, making aliasing from higher orders negligible.

5. CONCLUSION

This paper has described how both truncation and spatial aliasing errors exist independently of each other and also explained how they are related to each other in practical spherical array scenarios. It has also looked at a specific array, the mh acoustics em32 Eigenmike®, and investigated the effect of using two different sets of quadrature coefficients, '*non-uniform*' and '*uniform*', on the spatial aliasing error. It is recommended that quadrature coefficients be chosen to minimise the orthonormality error for orders less than the truncation order N, as these errors are amplified in the spatial aliasing error, while orthonormality errors for $n > N$ are attenuated.

6. RELATION TO PRIOR WORK

The work presented here has focused on separating the definitions of truncation error and spatial aliasing error, as well as looking at how spatial aliasing error is affected by the choice of quadrature coefficients. Past studies have combined the two errors by treating higher order components as noise that is included in the truncation error [6] and combined capsule noise and spatial aliasing [2] which can confuse the issue. '*Uniform*' quadrature coefficients have been used in [4] but this work has shown that better '*non-uniform*' coefficients can be chosen to yield more accurate results.

7. REFERENCES

[1] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the sound-field." International Conference on Acoustics, Speech, and Signal Processing (ICASSP), May 2002, pp. II–1781–1784.

[2] S. Kordon, A. Kruger, J.-M. Batke, and H. Kropp, "Optimization of spherical microphone array recordings." International Conference on Spatial Audio, November 2011.

[3] B. Rafaely, B. Weiss, and E. Bachmat, "Spatial aliasing in spherical microphone arrays," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 1003–1010, 2007.

[4] J. Meyer and G. Elko, "Handling spatial aliasing in spherical array applications." Hands-free Speech Communication and Microphone Arrays (HSCMA), May 2008.

[5] E. G. Williams, N. Valdivia, P. C. Herdic, and J. Klos, "Volumetric acoustic vector intensity imager," *The Journal of the Acoustical Society of America*, vol. 120, no. 4, pp. 1887–1897, 2006.

[6] E. G. Williams and K. Takashima, "Vector intensity reconstructions in a volume surrounding a rigid spherical microphone array," *The Journal of the Acoustical Society of America*, vol. 127, no. 2, pp. 773–783, 2010.

[7] T. Abhayapala, T. Pollock, and R. Kennedy, "Characterization of 3d spatial wireless channels." IEEE 58th Vehicular Technology Conference, October 2003, pp. 123 – 127.

[8] S. Brown and D. Sen, "Analysis of truncation and sampling errors in the spherical harmonic representation of soundfields." 162nd Meeting of Acoustical Society of America, October 2011.

[9] S. Brown, S. Wang, and D. Sen, "Analysis of the spherical-wave truncation error for spherical harmonic soundfield expansions." International Conference on Acoustics, Speech, and Signal Processing (ICASSP), March 2012, pp. 5 –8.

[10] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 135–143, 2005.

[11] Z. Li and R. Duraiswami, "Flexible and optimal design of spherical microphone arrays for beamforming," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 2, pp. 702–714, 2007.

[12] E. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*. London, UK: Academic Press, 1999.

[13] R. Hardin and N. Sloane, "Mclaren's improved snub cube and other new spherical designs in three dimensions," *Discrete & Computational Geometry*, vol. 15, pp. 429–441, 1996.

[14] S. Sobolev, "On mechanical quadrature formulas on the sphere," *Siberian Mathematical Journal*, vol. 3, pp. 769–796, 1962, (Russian).