forum acusticum 2026
8-12 September · Graz, Austria

TIME-DOMAIN BEHAVIOUR OF SPATIAL ALIASING ARTEFACTS IN SPHERICAL MICROPHONE ARRAY PROCESSING

Yueheng Li[1], Filippo Maria Fazi[1], Nara Hahn[1]

[1] *Institute of Sound and Vibration Research, University of Southampton*

This contribution has been accepted based on the submitted abstract. The submitted manuscript was not reviewed and is reproduced here without edits or corrections by the conference board. Neither the EAA nor the AAA take responsibility for its contents.

Abstract

Spatial aliasing in spherical microphone arrays (SMAs) arises from insufficient spatial sampling to capture high-order components, which are folded back into lower orders. This phenomenon is well understood: the frequency-independent spatial aliasing matrix characterises the mapping by which higher-order components are aliased into lower orders, as determined by the array's sampling geometry. In addition, the mathematical expression of spatial aliasing also involves a frequency-dependent term determined by radial functions of different orders. This can be rigorously analysed by poles and zeros in the Laplace domain. Upon this, a closed-form expression for the corresponding time-domain spatial aliasing artefact is derived, which demonstrates that spatial aliasing introduces temporal spreading of higher-order Ambisonic signals measured with SMAs.

Keywords: *spherical harmonics, spherical microphone array, ambisonics, spherical radial function*

1. Introduction

Spherical microphone arrays (SMAs) sample the sound field with a finite number of microphone capsules distributed over a spherical surface according to a prescribed sampling scheme, such as Gaussian, equiangular sampling or spherical t-design. The measured sound pressure signals are thereafter encoded to Ambisonic format, in which the captured sound field is represented as coefficients of spherical harmonics up to a certain order. Ambisonic encoding comprises two steps: a weighted summation of the microphone signals with the spherical harmonics evaluated at the capsule directions, and an equalisation that removes the effect of the array sphere, which is described by the radial function. Spatial aliasing takes place during the encoding. The pressure on the array sphere is an infinite-order physical function, whereas a finite number of microphones can only support up to a limited order of spherical harmonics. The surface pressure is therefore spatially undersampled, and the components of an order higher than the array can resolve are aliased onto the lower orders. This was well studied in [1], [2], [3], where the aliasing is effectively understood in terms of the spatial aliasing matrix, which describes the projections between high and low orders. Methods for mitigating spatial aliasing have also received considerable research interests, see [1], [4], [5], [6].

However, because the radial functions introduce a frequency dependence, spatial aliasing also produces a frequency-dependent artefact and, correspondingly, a time-domain one. It is shown below that each aliased component is high-pass filtered by this cross-order term, and therefore the encoded output spreads temporally. In this study, this artefact is investigated with Laplace domain analysis where the pole-zero expression of the cross-order term enables: a continuous closed-form expression of the time-domain behaviour, an analytical interpretation of the time-domain behaviour, and the characterisation of the frequency response with poles and zeros. The Laplace-domain representation of radial functions was introduced in [7] in the context of regularised inverse filtering. The present work applies that representation to the ratio of two radial functions of different order, which is the quantity that governs the frequency dependence of spatial aliasing.

2. Spatial Aliasing

Consider an SMA whose L microphones are positioned at the directions $\Omega_l = (\theta_l, \phi_l), l = 1, \cdots, L$ on a sphere with radius r, with $\theta_l \in [0, \pi]$ the colatitude and $\phi_l \in [0, 2\pi)$ the azimuth. Using a set of quadrature weights w_l, the Ambisonic signal of order n and degree

Corresponding author: *Yueheng.Li@soton.ac.uk.*

m is estimated as

$$\hat{a}_{nm}(k) = \frac{1}{b_n(kr)} \sum_{l=1}^{L} w_l p(k, \Omega_l, r) [Y_n^m(\Omega_l)]^*, \quad (1)$$

where $[\cdot]^*$ denotes complex conjugate $k = \omega/c$ is the wavenumber with ω being the angular frequency and c the speed of sound. w_l is the quadrature sampling weight corresponding to the sampling scheme of SMA. $\hat{a}_{nm}(k)$ is the encoded Ambisonic signal, and $(N+1)^2 \leq L$ is the maximum spherical harmonic order the SMA with L microphones can analyse. This expression assumes a direct inverse without using a regularised inverse radial filter for an intuitive analysis.

The term $b_n(kr)$ is the radial function which models how much each spherical harmonic order is present in the measured pressure given the sampling sphere. For a rigid sphere array, the radial function can be expressed in a compact form via the Wronskian relation,

$$b_n(kr) = \frac{-4\pi i^{n+1}}{(kr)^2 h_n^{(2)\prime}(kr)}, \quad (2)$$

where $h_n^{(2)}(kr)$ is the spherical Hankel function of the second kind, and $(\cdot)'$ is the first derivative operator. Assuming the sound field is composed of plane wave(s), spherical harmonic expansion allows us to express the sound pressure measured on the array surface as

$$p(k, \Omega_l, r) = \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} b_{n'}(kr) a_{n'm'}(k) Y_{n'}^{m'}(\Omega_l), \quad (3)$$

where $a_{n'm'}$ and $Y_{n'}^{m'}(\Omega_l)$ denote the true Ambisonic signal and the corresponding spherical harmonics, respectively, with the order n' going up to infinity since the expansion is not band-limited in practice. In the following the infinite summation over n' is truncated and analysed up to a high order N'. Substituting Eq. (3) into Eq. (1) yields

$$\hat{a}_{nm}(k) = \sum_{n'=0}^{N'} \sum_{m'=-n'}^{n'} \frac{b_{n'}(kr)}{b_n(kr)} a_{n'm'}(k) \varepsilon_{nm}^{n'm'} \quad (4)$$

$$= a_{nm}(k) + \sum_{n'>N}^{N'} \sum_{m'=-n'}^{n'} \frac{b_{n'}(kr)}{b_n(kr)} a_{n'm'}(k) \varepsilon_{nm}^{n'm'}. \quad (5)$$

Eq. (5) expresses the encoded Ambisonic signal $\hat{a}_{nm}(k)$ as the sum of two terms: the desired signal $a_{nm}(k)$ of the same order and degree, and a spatial aliasing term. The latter is a superposition of the unresolvable high orders, each of which is weighted by the gain in $\varepsilon_{nm}^{n'm'}$ and filtered by the ratio of the radial functions of orders n' and n.

$\varepsilon_{nm}^{n'm'}$ (see Eq. (6)) shows that spherical harmonic orthogonality holds approximately up to order N with the approximation of quadrature sampling weights. Spatial aliasing occurs for orders greater than N,

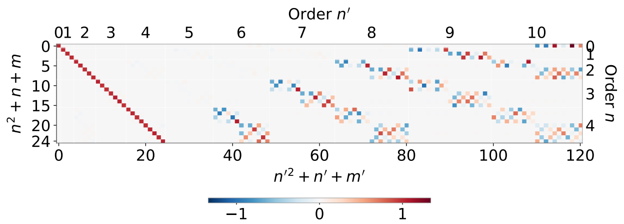

Figure 1: Spatial aliasing matrix with the Eigenmike em32 configuration with $L = 32$ microphones in a nearly-uniform spherical sampling scheme. The axes indicate the ACN indices. Maximum analysable spherical harmonic order $N = 4$ and the aliased high orders are plotted up to $N' = 10$.

where the orthogonality is no longer satisfied. The second equality is under the assumption that the exact orthogonality is preserved for order $n \leq N$. The entry $\varepsilon_{nm}^{n'm'}$ of the spatial aliasing matrix is shown in Fig. 1,

$$\varepsilon_{nm}^{n'm'} = \sum_{l=1}^{L} [Y_n^m(\Omega_l)]^* w_l Y_{n'}^{m'}(\Omega_l). \quad (6)$$

Note that $\varepsilon_{nm}^{n'm'}$ depends on the spherical sampling scheme alone, whereas the ratio of radial functions is frequency dependent. The former is the spatial aliasing matrix described in the literature, and the latter is analysed in the following section.

3. Laplace-Domain Analysis of the Cross-Order Term

From Eq. (5), other than the constant spatial aliasing gain given by $\varepsilon_{nm}^{n'm'}$, the contribution of spatial aliasing is also determined by the cross-order *transfer function*,

$$\tilde{H}_{n,n'}(\omega) = \frac{H_{n'}(\omega)}{H_n(\omega)} = \frac{b_{n'}(\frac{\omega}{c}r)}{b_n(\frac{\omega}{c}r)}, \quad (7)$$

where the dependency on radius is omitted for $H_n(\omega)$ since it is fixed for a given single-surface SMA. The Laplace variable is defined as $s = \sigma + i\omega$, where σ and ω denote its real and imaginary parts, respectively, and i is the imaginary unit. Evaluating a Laplace-domain function on the imaginary axis $s = i\omega$ recovers the corresponding frequency response. Eq. (7) are continued analytically into the s plane such that $H_n(s)|_{s=i\omega} = H_n(\omega)$. Following [7, Eq. (31)(32)], the expression of the radial function $H_n(s)$ in the pole-zero form is given as:

$$H_n(s) = 4\pi e^{\frac{r}{c}s} \frac{-\frac{c}{r}s^n}{\prod_{p=1}^{n+1}(s - s_{n,p})}, \quad (8)$$

where $s_{n,p}, p = 1, \cdots, n+1$ denote the poles of the order-n radial function, all of which are distinct poles lying in the left half-plane in the s plane with $\Re(s_{n,p}) < 0$. Since the numerator always has a lower order than the denominator, i.e. $n < n+1$, Eq. (8)

shows a strictly proper transfer function. For a stable radial function, its region of convergence (ROC) is the open half-plane to the right of the rightmost pole, $\Re(s) > \max_p(\Re(s_{n,p}))$, which encloses the imaginary axis and confirms the existence of the frequency response. Substituting Eq. (8) into Eq. (7),

$$\tilde{H}_{n,n'}(s) = s^{n'-n}\frac{\prod_{p=1}^{n+1}(s-s_{n,p})}{\prod_{p=1}^{n'+1}(s-s_{n',p})}, \quad (9)$$

$\tilde{H}(s)$ becomes the rational transfer function, with $n'+1$ off-origin poles, $n'-n>0$ zeros at the origin and another $n+1$ zeros lying in the left half-plane, in total $n'+1$ zeros, the same as the number of poles, implying a bi-proper transfer function. The behaviour of $\tilde{H}_{n,n'}$ at the two ends of the frequency range follows directly from Eq. (9). For $n \geq 0$, both radial functions in Eq. (7) have a zero at the origin, so the ratio takes the form $0/0$ at $s=0$. Eq. (9) provides the analytic continuation through the origin. Since the fraction $\prod_{p=1}^{n+1}(s-s_{n,p})/\prod_{p=1}^{n'+1}(s-s_{n',p})$ in Eq. (9) is finite and non-zero as $s \to 0$, the behaviour at the origin is therefore governed entirely by $s^{n'-n}$. Provided that $n' > n$, the limit of this term is zero as $\omega \to 0$, so the low-frequency limit of $\tilde{H}_{n,n'}(i\omega)$ in Eq. (9) is zero. On the other hand, since both the numerator and denominator of Eq. (9) are monic polynomials of degree $n'+1$, the high-frequency limit is unity. Evaluating on the imaginary axis,

$$\lim_{\omega\to 0}\tilde{H}_{n,n'}(i\omega) = 0, \qquad \lim_{|\omega|\to\infty}\tilde{H}_{n,n'}(i\omega) = 1. \quad (10)$$

Since $\tilde{H}_{n,n'}(s)$ is bi-proper, extracting the unit feed-through term yields the partial fraction expansion,

$$\tilde{H}_{n,n'}(s) = 1 + \sum_{p=1}^{n'+1}\frac{r_p}{s-s_{n',p}}, \quad (11)$$

where r_p is the residual corresponding to the pole $s_{n',p}$, given as

$$r_p = \frac{s_{n',p}^{n'-n}\prod_{q=1}^{n+1}(s_{n',p}-s_{n,q})}{\prod_{q\neq p}(s_{n',p}-s_{n',q})}. \quad (12)$$

Since all $s_{n',p}$ are located in the open left half-plane, operating the single-sided inverse Laplace transform with the ROC to the right of the rightmost pole gives the causal and continuous time-domain response as a closed-form expression of the spatial aliasing time-domain artefact,

$$\mathcal{L}^{-1}\{\tilde{H}_{n,n'}(s)\} = \delta(t) + \sum_{p=1}^{n'+1} r_p e^{s_{n',p}t}u(t), \quad (13)$$

where $\delta(t)$ is the Dirac Delta function and $u(t)$ is the Heaviside unit step function. The time-domain behaviour of spatial aliasing therefore consists of a Dirac impulse at $t=0$ followed by a sum of $n'+1$ exponentials, whose decay rates are given by $\Re(s_{n',p})$ and whose oscillation frequencies are given by $\Im(s_{n',p})$.

4. Evaluations

To numerically evaluate the expressions given in the previous section, the Eigenmike em32 configuration is adopted, with the radius $r = 0.042$ m. The sound speed $c = 343$ m/s. The cross-order transfer function is examined for $n \in \{0, \cdots, N\}$ with $N=4$, and $n' \in \{N+1, \cdots, N'\}$ with $N'=10$.

Fig. 3 shows the inverse DFT of Eq. (2) together with the expression in Eq. (13). Eq. (2) is numerically evaluated at $2^{10}+1$ single-sided frequency bins from DC to $fs/2$, where $fs = 4.8 \times 10^5$ Hz, corresponding to a DFT length of 2^{11}. At the DC bin, the limit value of zero is assigned in accordance with Eq. (10). The time-domain responses are therefore obtained by their inverse DFT followed by a circular shift of half the DFT length. For the comparison, decaying tails in (13) are evaluated with the same sampling rate, and the amplitude is rescaled by the sampling interval $1/fs$ for a fair comparison in the discretised scale. Also note that the inverse DFT evaluation is of finite bandwidth, whereas Eq. (13) is a continuous-time expression of infinite bandwidth.

The pole-zero loci corresponding to different n, n' pairs are displayed in Fig. 2 to examine the time-domain behaviour in terms of the pole-zero distribution. It is observed that for a fixed n and increasing n', the poles migrate outwards to the left half-plane and the number of poles also increases, while the $n+1$ off-origin zeros remain fixed. The shape of the tail is determined by the order n', since by Eq. 13 the decaying rate and oscillation frequencies of $n'+1$ exponentials are determined by the real and imaginary part of poles $\{s_{n',p}\}|_{p=1}^{n'+1}$, and the number of exponentials is solely given by $n'+1$. Increasing n' introduces more poles, some of which have greater imaginary parts but almost the same real parts, therefore creating more ringing as shown along rows of Fig. 3.

Eq. (12) expresses the residual of the pole $s_{n',p}$ as a product of distances in the s plane: in the numerator, the distance from the pole to the origin, and to the $n+1$ off-origin zeros; in the denominator, the distances from the pole to all remaining n' poles. Increasing n with a fixed n' relocates off-origin zeros $\{s_{n,p}\}_{p=1}^{n+1}$ outwards to the left half-plane, while leaving poles unchanged. So the exponent $n'-n$ decreases and additional off-origin zeros enter the numerator. The effect of the exponent in the numerator dominates, so the residual reduces, which is the reduction in amplitudes of the response observed down a column in Fig, 3.

Fig. 4 plots the frequency responses of the cross term in the Fourier domain. The upper plot compares the frequency response of the first column of the cross-order term, i.e. $n'=6$ and $n=0,\cdots,4$, whereas the lower plot shows a row where $n=0$ and $n'=5,\cdots,10$.

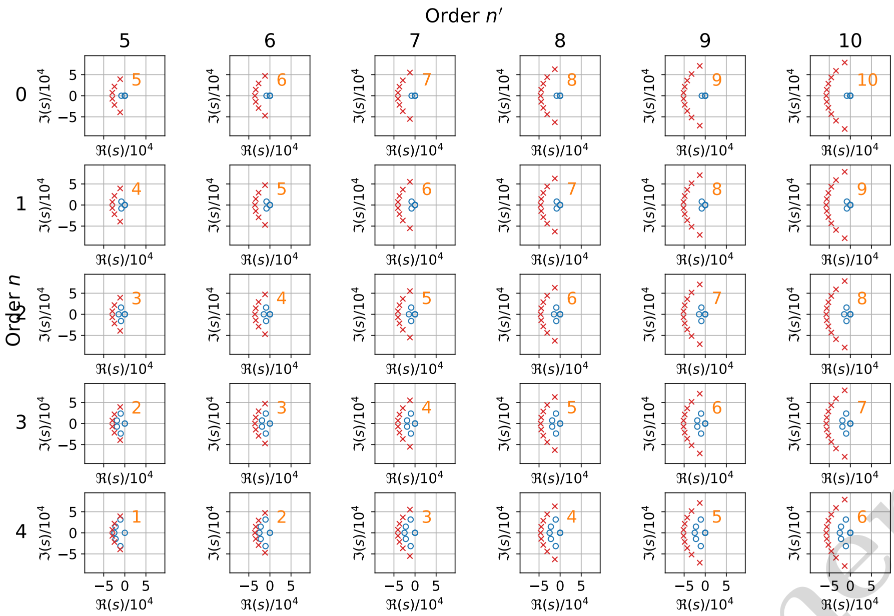

Figure 2: The pole-zero plot for the cross-order $\tilde{H}_{n,n'}(s)$ for n (rows) and n' (columns). Poles are shown as crosses and zeros as circles. The multiplicity of $n' - n$ zeros at the origin is annotated with orange numbers.

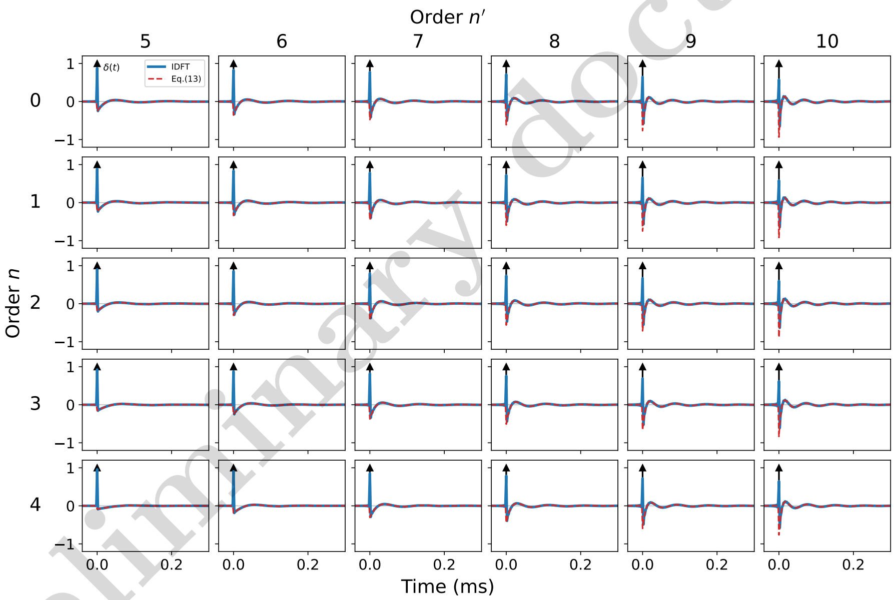

Figure 3: Time-domain cross-order response of the term $b_{n'}(kr)/b_n(kr)$. The blue line is evaluated numerically with the inverse DFT at the sampling rate $f_s = 4.8 \times 10^5$ Hz, while the red dashed line is the implementation of the exponential tails in Eq. (13) (without the Dirac Delta), with amplitudes rescaled by the sampling interval $1/fs$. The black arrows at $t = 0$ represent the continuous Dirac Delta.

The full frequency response can be interpreted as three bands divided by two transition zones: below the knee frequency ($\omega_{knee} \approx nc/r$), between the knee frequency and cut-on frequency ($\omega_c \approx n'c/r$), and above the cut-on frequency. Below the knee frequency, the response slope of $(n' - n) \cdot 20$ dB/decade is attributed to the $n' - n$ zeros at the origin. Between the knee frequency and the cut-on frequency, all $n' + 1$ zeros contribute to the $(n' + 1) \cdot 20$dB/decade slope steepening as a small hump. Aliased components are not merely passed but slightly amplified around the cut-on frequency. Above the cut-on frequency, the response converges to a flat unit asymptote since the number of poles and zeros are equal and both numerator and denominator of Eq. (9) are monic.

A case study is also presented to illustrate the overall spatial aliasing effect, i.e. the combined action of the frequency-independent aliasing gains $\varepsilon_{nm}^{n'm'}$ and

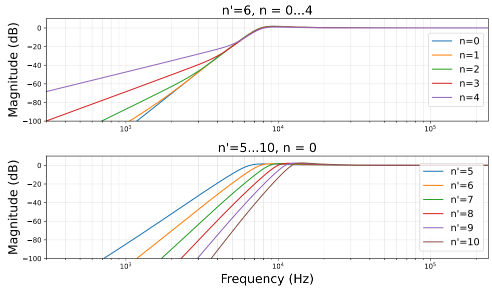

Figure 4: Magnitude response of the cross-order term $b_{n'}(kr)/b_n(kr)$ evaluated with the radial function in Eq. (2) for fixed $n' = 6$ with varying n (the upper plot), and for fixed $n = 0$ with varying n' (the lower plot). The array radius $r = 0.042$ m.

the frequency-dependent cross-order terms. Consider a unit-amplitude broadband impulse incident as a plane wave from the positive z-axis in a free field, i.e. from the north pole $\theta_0 = 0°$. The array signals are simulated for an Eigenmike em32 at a sampling rate of $f_s' = 4.8 \times 10^4$ Hz, and the spatial aliasing is evaluated up to order $N' = 18$, which satisfies $\pi f_s' r/c < N'$. Rewriting Eq. (5) in the time domain,

$$\hat{a}_{nm}(t) = a_{nm}(t) + \sum_{n'=N+1}^{N'} h_{n,n'}(t) * \left(\sum_{m'=-n'}^{n'} \varepsilon_{nm}^{n'm'} a_{n'm'}(t) \right), \tag{14}$$

where $*$ denotes the time-domain convolution and $h_{n,n'}(t)$ is the time-domain response of the cross-order term. For a unit-amplitude plane wave incident from the north pole, the true Ambisonic signals are,

$$a_{nm}(t) = x(t)[Y_n^m(0, \phi)]^*, \tag{15}$$

where $x(t)$ time signal of the incident plane wave. With the orthonormal normalisation, the spherical harmonics evaluated at the north pole are

$$Y_n^m(0, \phi) = 0, \text{for } m \neq 0 \tag{16}$$

$$Y_n^0(0, \phi) = \sqrt{\frac{2n+1}{4\pi}}. \tag{17}$$

independently of the azimuth ϕ. Only the channels with $m = 0$ are therefore excited, each containing a scaled copy of the source impulse. Therefore, Fig. 5 compares the encoded channels $\hat{a}_{n0}(t)$ with the true Ambisonic signals $a_{n0}(t)$. The true signals are impulses of amplitude $\sqrt{(2n+1)/4\pi}$ located at the arrival time indicated by the dashed line. The encoded channels, in contrast, are spread in time, and the aliasing artefacts are causal in the sense that they appear only after the arrival of the direct wavefront.

5. Conclusion

This paper has analysed the frequency-dependent component of the spatial aliasing artefact in spherical microphone arrays. This component is governed by the ratio of two radial functions of different orders, for which a pole–zero representation has been derived in the Laplace domain. The corresponding magnitude response has been analysed accordingly and characterised in three bands partitioned by the modal cut-on frequencies of the two orders. A closed-form expression of the time-domain artefact has been derived and analysed in terms of zeros and poles. The expression was also validated with the numerical evaluation using inverse DFT, and close agreement between the two evaluations was shown. It was found that the decay rates and oscillation frequencies of the exponential tails are determined solely by n', whereas the amplitudes depend on both n' and n through the residuals and grow with the order difference $n' - n$. A complementary case study illustrates that this temporal spreading can substantially distort the shape of the encoded time-domain Ambisonic signals.
Since the cross-order term does not suffer from ill-conditioned problems at low frequencies as the inverse filtering in the encoder does, this paper assumes an ideal, unregularised radial filter. In practice, implementing the regularisation will create a circle of poles around the origin, and the impulse responses are acausal.[7].

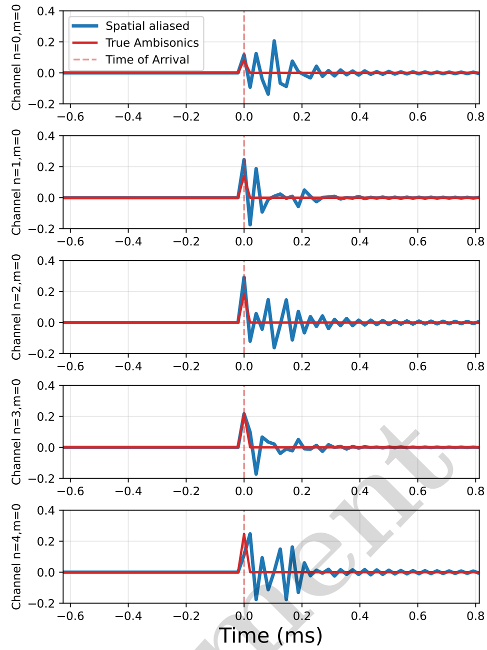

Figure 5: The comparison between the encoded Ambisonic signals contaminated by spatial aliasing and the true Ambisonic signals for channel $m = 0$ of different orders n. The dashed line shows the time of arrival of the impulse.

6. Acknowledgment

The authors would like to thank the Sony Research Award Program for funding and supporting this work.

References

[1] 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.

[2] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Transactions on speech and audio processing*, vol. 13, no. 1, pp. 135–143, 2005.

[3] B. Rafaely, *Fundamentals of spherical array processing*. Springer, 2015, vol. 8.

[4] J. Meyer and G. W. Elko, "Handling spatial aliasing in spherical array applications," *2008 Hands-Free Speech Communication and Microphone Arrays*, pp. 1–4, 2008.

[5] D. L. Alon and B. Rafaely, "Beamforming with optimal aliasing cancellation in spherical microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 1, pp. 196–210, 2015.

[6] J. Lin, X. Wu, and T. Qu, "Anti spatial aliasing hoa encoding method based on aliasing projection matrix," in *2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP)*, IEEE, 2020, pp. 321–325.

[7] N. Hahn and F. M. Fazi, "Regularized inverse filter design for rigid spherical microphone array processing: Laplace-and time-domain representations," in *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2026, pp. 14 637–14 641.