

# ROBUST SPHERICAL HARMONIC DOMAIN INTERPOLATION OF SPATIALLY SAMPLED ARRAY MANIFOLDS

*Alastair H. Moore, Mike Brookes and Patrick A. Naylor*

Imperial College London, Dept. of Electrical and Electronic Engineering, UK.

## ABSTRACT

Accurate interpolation of the array manifold is an important first step for the acoustic simulation of rapidly moving microphone arrays. Spherical harmonic domain interpolation has been proposed and well studied in the context of head-related transfer functions but has focussed on perceptual, rather than numerical, accuracy. In this paper we analyze the effect of measurement noise on spatial aliasing. Based on this analysis we propose a method for selecting the truncation orders for the forward and reverse spherical Fourier transforms given only the noisy samples in such a way that the interpolation error is minimized. The proposed method achieves up to 1.7 dB improvement over the baseline approach.

**Index Terms**— interpolation, HRTF, array manifold, spherical harmonics, microphone array

## 1. INTRODUCTION

For many years multichannel acoustic signal processing has targeted scenarios in which the microphone arrays are static or move sufficiently slowly that a short-term stationary assumption can be used. However, for applications such as robot audition and hearing aids an assumption of stationarity is not normally satisfied in practice. To validate existing algorithms and to aid the development of new algorithms which do not assume stationarity, requires simulation tools that can accurately predict the signals received by moving microphone arrays.

Assuming the incident sound field can be expressed as a weighted sum of plane waves impinging on the array from different Directions-of-Arrival (DOAs), we require the response of the array to a unit amplitude plane wave from each DOA, that is the array manifold or steering vector. Under idealized conditions, the array manifold for some very simple array geometries has an analytical solution. In practice, array manifolds must be measured for a finite number of DOAs, a process known as sampling on the sphere, and interpolated to the required DOAs. Generalizing the weighted sum of plane waves model of a sound field to a plane-wave density [1, 2], it becomes more convenient to express both the array manifold and the sound field as a Spherical Harmonic (SH) expansion [3, 4]. In this way the Wigner-D rotation matrices can be used to rotate the array manifold with respect to the sound field [5], and such rotation can vary rapidly with time. The interpolation of the array manifold is then achieved implicitly.

The problem we seek to address is how to obtain the SH domain representation of a measured array manifold such that it can be most accurately interpolated. The SH domain representation of

Head-related Transfer Functions (HRTFs) in particular has been studied quite extensively [6–12]. However, most studies have analyzed the interpolation accuracy from the perspective of human sound localization cues, and have consequently focussed only on the magnitude spectrum or indeed the perceived position. Some studies have proposed methods to make interpolation robust to missing data in the HRTF measurements [8, 9]. This offers some insight into our scenario but cannot be compared directly since we assume that the sampling grid is well distributed over the full sphere. In [13], the SH expansion included regularization, as proposed in [7], but the regularization parameter needed to be ‘carefully chosen’ and no guidance was provided as to how this should be set. An analysis of the effect of SH order and number of sample points on the reconstruction error was presented and a theoretical law for selecting the SH order proposed. In [11] this law, which is based on the wavenumber and the radius of the array, which in the case of HRTFs is the radius of the head, was shown to provide a lower bound on the required SH order. As such the authors are not aware of a systematic approach to selecting the truncation order of spherically sampled data.

The novel contributions of this paper are: an analysis of the aliasing structure of sampling on the sphere in the presence of measurement noise (Sec. 3); a systematic approach to estimating the SH coefficients of the array manifold (Sec. 4); and an evaluation of the proposed approach using measured data (Sec. 5).

## 2. SPHERICAL FOURIER TRANSFORM

The Spherical Fourier Transform (SFT) and its inverse are well described in, for example, [14–16]. Here we introduce the key equations using our notation which makes a clear distinction between the truncation orders used for the forward and reverse transforms, which are generally assumed to be the same, but need not be.

Let  $A(\theta, \phi)$  be a square-integrable function defined on the surface of a sphere, where  $\theta$  and  $\phi$  denote the inclination and azimuth angles, respectively. We can express  $A(\theta, \phi)$  as a series expansion

$$A(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{l,m} Y_l^m(\theta, \phi) \quad (1)$$

where

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\phi} \quad (2)$$

are the SH basis functions of order  $l \in \mathbb{N}$  and degree  $m \in \{-l \dots l\}$  and  $P_l^m(\cdot)$  is the associated Legendre function. The SH coefficients are obtained from the forward SFT of  $A(\theta, \phi)$  as

$$A_{l,m} = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} A(\theta, \phi) [Y_l^m(\theta, \phi)]^* \sin \theta d\theta d\phi. \quad (3)$$

The research leading to these results has received funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 609465. This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/M026698/1].

If the spatial bandwidth of  $A(\theta, \phi)$  is limited such that  $A_{l,m} = 0 \forall l > L_A$  then the infinite summation in (1) can be truncated without error

$$A(\theta, \phi) = \sum_{l=0}^{L_{\text{rev}}} \sum_{m=-l}^l A_{l,m} Y_l^m(\theta, \phi) \quad (4)$$

where  $L_{\text{rev}} \geq L_A$ . Note that given the set of SH coefficients up to order  $L_{\text{rev}}$ ,  $A(\theta, \phi)$  can be evaluated for any direction. Interpolation is therefore possible if these coefficients can be obtained from discrete samples of  $A(\theta, \phi)$ .

Provided that  $A(\theta, \phi)$  is order-limited, (3) can be discretized

$$A_{l,m} = \sum_p w_p A(\theta_p, \phi_p) [Y_l^m(\theta_p, \phi_p)]^*, \quad l \leq L_{\text{fwd}} \quad (5)$$

where  $\{w_p\}_{p=1 \dots P}$  are the quadrature weights of the sampling scheme. The number of sample points required depends on their spatial distribution with a lower bound of  $P \geq (L_{\text{fwd}} + 1)^2$ , where  $L_{\text{fwd}}$  is the maximum order of the SH coefficients to be estimated and  $L_{\text{fwd}} \geq L_A$ . Rewriting (5) in matrix form gives

$$\mathbf{a} = \mathbf{Y}_{\text{fwd}}^H \mathbf{W} \mathbf{a} \quad (6)$$

where  $\mathbf{W} = \text{diag}\{w_p\}_{p=1 \dots P}$  and

$$\begin{aligned} \mathbf{a} &= [A(\theta_1, \phi_1) \quad A(\theta_2, \phi_2) \quad \cdots \quad A(\theta_P, \phi_P)]^T \\ \mathbf{y}_{l,m} &= [Y_l^m(\theta_1, \phi_1) \quad Y_l^m(\theta_2, \phi_2) \quad \cdots \quad Y_l^m(\theta_P, \phi_P)]^T \\ \mathbf{Y}_{\text{fwd}} &= [\mathbf{y}_{0,0} \quad \mathbf{y}_{1,-1} \quad \mathbf{y}_{1,0} \quad \mathbf{y}_{1,1} \quad \cdots \quad \mathbf{y}_{L_{\text{fwd}}, L_{\text{fwd}}}] \\ \mathbf{a} &= [A_{0,0} \quad A_{1,-1} \quad A_{1,0} \quad A_{1,1} \quad \cdots \quad A_{L_{\text{fwd}}, L_{\text{fwd}}}]^T. \end{aligned}$$

The quadrature weights are selected such that the orthonormality properties of the SFT are ensured, i.e.  $\mathbf{Y}_{\text{fwd}}^H \mathbf{W} \mathbf{Y}_{\text{fwd}} = \mathbf{I}_{(L_{\text{fwd}}+1)^2}$ .

In general the sampling schemes used for array manifold measurements do not have a closed-form solution for the quadrature weights. In this case, the forward discrete SFT is approximated as

$$\mathbf{a} = \mathbf{Y}_{\text{fwd}}^\dagger \mathbf{a}, \quad L_{\text{fwd}} \geq L_A \quad (7)$$

where  $\mathbf{Y}_{\text{fwd}}^\dagger = (\mathbf{Y}_{\text{fwd}}^H \mathbf{Y}_{\text{fwd}})^{-1} \mathbf{Y}_{\text{fwd}}^H$  is the Moore-Penrose pseudo-inverse.

In summary, using (7),  $(L_{\text{fwd}} + 1)^2$  SH coefficients are obtained from  $P$  spatial samples of  $A(\theta, \phi)$  and interpolation to  $(\theta_q, \phi_q)$  is achieved using (4) with  $L_A \leq L_{\text{rev}} \leq L_{\text{fwd}} \leq L_{\text{max}}$ , where  $L_{\text{max}} \leq \sqrt{P}-1$  is the highest SH order coefficient which can be obtained without spatial aliasing for the chosen sampling scheme.

### 3. ALIASING AND THE EFFECT OF MEASUREMENT NOISE

In the case where  $L_A$  exceeds  $L_{\text{fwd}}$  the forward SFT is

$$\hat{\mathbf{a}} = \mathbf{Y}_{\text{fwd}}^\dagger \mathbf{a}, \quad L_A > L_{\text{fwd}} \quad (8)$$

where  $\epsilon = \hat{\mathbf{a}} - \mathbf{a}$  is the error introduced by spatial aliasing. The aliasing projection matrix [17, 18] provides a means of analyzing the extent to which spatial frequency components which exceed  $L_{\text{fwd}}$  will be aliased into the estimated SH coefficients. It is defined as

$$\mathbf{D} = \mathbf{Y}_{\text{fwd}}^\dagger \mathbf{Y}_\Upsilon \quad (9)$$

where  $\mathbf{Y}_\Upsilon = [\mathbf{y}_{0,0} \quad \mathbf{y}_{1,-1} \quad \mathbf{y}_{1,0} \quad \mathbf{y}_{1,1} \quad \cdots \quad \mathbf{y}_{L_\Upsilon, L_\Upsilon}]$  and  $L_\Upsilon$  is arbitrarily large. For each row of  $\mathbf{D}$ , indexed as  $l^2 + l + m + 1$ , the entry in the column indexed by  $(l')^2 + l' + m' + 1$  contains the projection operator from  $A_{l',m'}$  onto  $A_{l,m}$ . Provided the distribution of spatial samples is adequate for the  $L_{\text{fwd}}$ -th order forward SFT, the aliasing projection matrix is structured as  $\mathbf{D} = [\mathbf{I}_{(L_{\text{fwd}}+1)^2} \quad \mathbf{\Delta}_\epsilon]$  where  $\mathbf{\Delta}_\epsilon$  denotes the extent of the aliasing for SH coefficients with  $L_{\text{fwd}} < l' \leq L_\Upsilon$ .

In practical measurements the measured impulse response from a source on a sphere of radius  $r$  from the  $p$ -th direction,  $(\theta_p, \phi_p)$ , to the  $\nu$ -th microphone is given by

$$\tilde{h}_\nu(\theta_p, \phi_p, t) = h_\nu(\theta_p, \phi_p, t) + n_{\nu,p}(t) \quad (10)$$

where  $h_\nu$  is true impulse response,  $n_{\nu,p}$  is the measurement noise and  $t$  is the time sample index. In the frequency domain (10) becomes

$$\tilde{H}_\nu(\theta_p, \phi_p, \omega) = H_\nu(\theta_p, \phi_p, \omega) + N_{\nu,p}(\omega).$$

By working in the frequency domain we assume that the noise is sampled from a zero-mean, Gaussian distribution with frequency-dependent variance  $\sigma_\nu^2(\omega)$ . Since  $N_{\nu,p}(\omega)$  for each  $p$  is an independent realization, the noise in each  $\tilde{H}_\nu(\theta_p, \phi_p, \omega)$  is uncorrelated. In other words each  $N_{\nu,p}(\omega)$  can be seen as a spatial sample of an underlying function  $N_\nu(\theta, \phi, \omega)$  which is spatially white and therefore has infinite spatial bandwidth. The SFT of  $\tilde{H}_\nu(\theta_p, \phi_p, \omega)$  is therefore

$$\tilde{\mathbf{h}}(\omega) = \mathbf{Y}_{\text{fwd}}^\dagger \tilde{\mathbf{h}}(\omega) \quad (11)$$

$$= \mathbf{Y}_{\text{fwd}}^\dagger \mathbf{h}(\omega) + \mathbf{Y}_{\text{fwd}}^\dagger \mathbf{n}(\omega) \quad (12)$$

$$= \mathbf{h}(\omega) + \hat{\mathbf{n}}(\omega), \quad L_H(\omega) < L_{\text{fwd}} \quad (13)$$

where  $L_H(\omega)$  is defined similarly to  $L_A$  and the remaining terms are defined according to the corresponding term in (6). Since  $N_\nu(\theta, \phi, \omega)$  is not bandlimited, it is inevitable that, regardless of how high  $L_{\text{fwd}}$  is chosen, the noise will be aliased. Consider as an example the aliasing projection matrices for  $L_{\text{fwd}} \in [4, 6]$  shown in Fig. 1 for a sampling scheme with  $L_{\text{max}} = 5$ . With  $L_{\text{rev}} = L_{\text{fwd}} = 3$ , all SH coefficients with  $l > 3$  will be aliased and so lead to inaccuracies in the interpolation. On the other hand, selecting a higher value of  $L_{\text{fwd}}$  one can prevent some of the aliasing, since the higher order SHs are explicitly estimated. By selecting  $L_{\text{rev}} < L_{\text{fwd}}$ , the SHs with  $L_{\text{rev}} < l \leq L_{\text{fwd}}$ , which we assume contain only noise, do not contribute to the interpolated response. However, for a given sampling scheme, increasing  $L_{\text{fwd}}$  also increases the condition number of  $\mathbf{Y}_{\text{fwd}}$  [3], making the pseudo-inverse more sensitive to noise. There is, therefore, a compromise required between avoiding spatial aliasing of noise and making the SFT more sensitive to noise in the SHs with  $l < L_{\text{rev}}$ . For comparison, Fig. 1(d) shows the effect of under-sampling, i.e. attempting to perform the forward SFT with truncation order higher than the sampling scheme is capable of; aliasing occurs to some extent for all  $l < L_{\text{fwd}}$ .

### 4. PROPOSED METHOD

We have so far assumed that  $L_H(\omega) \leq L_{\text{rev}}$ . In practice, whilst we might expect the array manifold to be band-limited, we do not know the value of  $L_H(\omega)$ . For a microphone mounted on a rigid sphere,  $H_\nu(\theta, \phi, \omega)$  depends on the radius of the sphere,  $r_a$ , and  $H_{l,m} \rightarrow 0$  for  $l > \omega r_a / c$  where  $c$  is the speed of sound [19]. A rigid sphere has often been used as a simple approximation of the

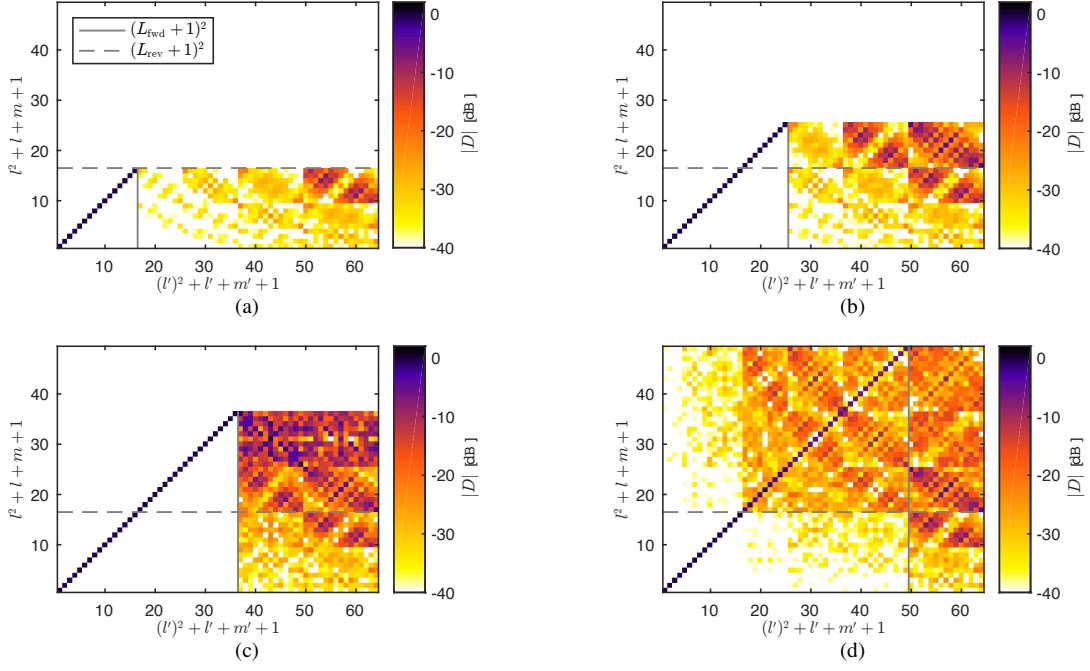


**Fig. 1.** Aliasing projection matrix for (a)  $L_{\text{fwd}} = 3$ , (b)  $L_{\text{fwd}} = 4$ , (c)  $L_{\text{fwd}} = 5$  and (d)  $L_{\text{fwd}} = 6$  for sampling scheme with  $L_{\text{max}} = 5$ . Only SH coefficients below the dashed horizontal line are included in the reverse SFT when  $L_{\text{rev}} = 3$ . Off-diagonal elements indicate aliasing.

human head to describe the directional properties of HRTFs and indeed at low frequencies accounts for the main wave phenomena [20]. However, at higher spatial bandwidth, resonances within the pinna lead to a much higher spatial bandwidth. For an arbitrary geometry, though the dimensions of the baffle may give some indication, the correct choice of  $L_{\text{rev}}$  is ultimately unknown. We therefore now propose a systematic method to select both  $L_{\text{fwd}}$  and  $L_{\text{rev}}$ .

Let the continuous function we wish to estimate be  $H_\nu(\theta, \phi, \omega)$  and the available data be noisy samples,  $\tilde{H}_\nu(\theta_\gamma, \phi_\gamma, \omega)$ , where  $(\theta_\gamma, \phi_\gamma)$ ,  $\gamma \in 1 \dots \Gamma$  are the sample directions. These directions are first partitioned into a training set,  $\mathcal{P}$ , of size  $P$ , a development set,  $\mathcal{Q}$ , of size  $Q$ , and a test set,  $\mathcal{R}$ , of size  $R$ . The  $L_{\text{fwd}}$ -order forward SFT (11) is applied to the training set measurements  $\tilde{H}_\nu(\theta_\gamma, \phi_\gamma, \omega) \forall \gamma \in \mathcal{P}$  to obtain  $\tilde{H}_{\nu,l,m}(\omega, L_{\text{fwd}}) \forall l \leq L_{\text{fwd}}$ . Applying (4) and dropping the dependence on  $\omega$  for brevity gives

$$\tilde{H}_\nu(\theta, \phi, L_{\text{fwd}}, L_{\text{rev}}) = \sum_{l=0}^{L_{\text{rev}}} \sum_{m=-l}^l \tilde{H}_{\nu,l,m}(L_{\text{fwd}}) Y_l^m(\theta, \phi)$$

such that the relative absolute error between the noisy measurements and their reconstructed approximations is [13]

$$E_\nu(\theta_\gamma, \phi_\gamma, L_{\text{fwd}}, L_{\text{rev}}) = \frac{\left\| \tilde{H}_\nu(\theta_\gamma, \phi_\gamma) - \tilde{H}_\nu(\theta, \phi, L_{\text{fwd}}, L_{\text{rev}}) \right\|}{\left\| \tilde{H}_\nu(\theta_\gamma, \phi_\gamma) \right\|}$$

where  $\|\cdot\|$  denotes the Euclidean norm. Note that, since the arguments are complex-valued, this measure inherently takes account of the phase accuracy.

It has been previously observed that overfitting can make the error for the training set arbitrarily low, but this increases the error for

interpolation points. We therefore consider only the average interpolation error, which for the development set is defined as

$$E_{\nu,\mathcal{Q}}(L_{\text{fwd}}, L_{\text{rev}}) = \frac{1}{Q} \sum_{(\theta_\gamma, \phi_\gamma) \in \mathcal{Q}} E_\nu(\theta_\gamma, \phi_\gamma, L_{\text{fwd}}, L_{\text{rev}}) \quad (14)$$

and for the test set is  $E_{\nu,\mathcal{R}}(L_{\text{fwd}}, L_{\text{rev}})$ , which is obtained as in (14) but over the test set,  $\mathcal{R}$ .

Since the conditioning of  $\mathbf{Y}_{\text{fwd}}$  depends on the directions in the training set, cross-validation is used to avoid any dependence on a particular allocation of directions to the training set. Averaging  $E_{\nu,\mathcal{Q}}(L_{\text{fwd}}, L_{\text{rev}})$  and  $E_{\nu,\mathcal{R}}(L_{\text{fwd}}, L_{\text{rev}})$  over all folds of the cross-validation gives,  $\bar{E}_{\nu,\mathcal{Q}}(L_{\text{fwd}}, L_{\text{rev}})$  and  $\bar{E}_{\nu,\mathcal{R}}(L_{\text{fwd}}, L_{\text{rev}})$ , respectively. The proposed method is then to solve the optimization problem

$$\arg \min_{L_{\text{fwd}}, L_{\text{rev}}} \bar{E}_{\nu,\mathcal{Q}}(L_{\text{fwd}}, L_{\text{rev}}), \quad L_{\text{rev}} \leq L_{\text{fwd}} \leq L_{\text{max}}. \quad (15)$$

Since the solution space is discrete and bounded it is quite practical to perform an exhaustive search. Finally, the effectiveness of the optimization is assessed using  $\bar{E}_{\nu,\mathcal{R}}(L_{\text{fwd}}, L_{\text{rev}})$ .

As discussed, the value of  $L_{\text{max}}$  depends on the sampling scheme. The directions for which HRTF or array manifold measurements are made are generally chosen to obtain a reasonably uniform distribution or to satisfy particular sampling criteria [11]. In our method, by assigning a random subset of measurement directions to the development and test partitions, the remaining directions are unlikely to be optimally distributed and so  $L_{\text{max}}$  will be somewhat lower than the upper bound suggested by  $L_{\text{max}} \leq \sqrt{P}-1$ . Noting the structure of the aliasing projection matrix in (9), the aliasing of SH coefficients with  $l \leq L_{\text{fwd}}$  can be quantified as

$$\varepsilon(L_{\text{fwd}}) = \text{tr} \left( (\mathbf{D}_{\text{fwd}} - \mathbf{I}_{(L_{\text{fwd}}+1)^2})^H (\mathbf{D}_{\text{fwd}} - \mathbf{I}_{(L_{\text{fwd}}+1)^2}) \right)$$

---

**Algorithm 1** Proposed algorithm to select  $L_{\text{fwd}}$  and  $L_{\text{rev}}$ 


---

Partition measurements into  $G$  folds containing training, development and test sets

**for each fold do**

    Determine  $L_{\text{max}}$  for training set

**for**  $L_{\text{fwd}} = 1 \dots L_{\text{max}}$  **do**

        Calculate  $\bar{H}_{\nu,l,m}(\omega, L_{\text{fwd}}) \forall l \leq L_{\text{fwd}}$  using (11)

**for**  $L_{\text{rev}} = 1 \dots L_{\text{fwd}}$  **do**

            Calculate  $E_{\nu,Q}(L_{\text{fwd}}, L_{\text{rev}})$  using (14)

    Calculate average interpolation error  $\bar{E}_{\nu,Q}(L_{\text{fwd}}, L_{\text{rev}})$

    Select optimum  $L_{\text{fwd}}$  and  $L_{\text{rev}}$  using (15)

---

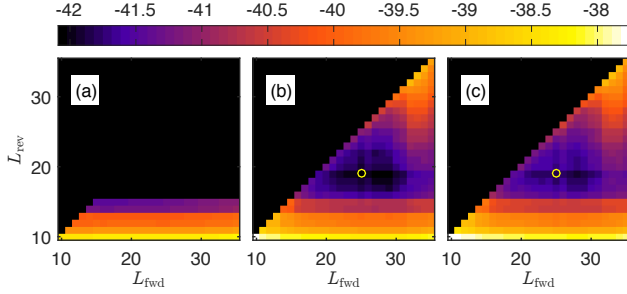


**Fig. 2.** Mean error in dB for (a) training, (b) development and (c) test points averaged over 100 trials at 685 Hz. Optimum combination of  $L_{\text{fwd}}$  and  $L_{\text{rev}}$  determined from development points shown as circle.

where  $\mathbf{D}_{\text{fwd}} = \mathbf{Y}_{\text{fwd}}^\dagger \mathbf{Y}_{\text{fwd}}$ . Defining  $\delta$  as a small positive constant,  $L_{\text{max}}$  is the largest value of  $L_{\text{fwd}}$  which satisfies  $\varepsilon(L_{\text{fwd}}) < \delta$ .

The complete method is summarized in Algorithm 1.

## 5. EVALUATION

The proposed method was evaluated on the left HRTF data from [21] which includes measurements for 1784 directions. These directions lie on contours of equal inclination, spaced  $5^\circ$  apart, with the azimuth spacing also  $5^\circ$  at the equator and increasing towards the poles to maintain an approximately uniform coverage. In each of 100 folds, the directions were randomly assigned to  $\mathcal{P}$ ,  $\mathcal{Q}$  or  $\mathcal{R}$ , with  $P = 1521$ ,  $Q = 132$  and  $R = 131$ . With this partitioning the theoretical maximum value of  $L_{\text{max}}$  is 38, but due to the random allocations one would not expect this best-case condition to be met.

Algorithm 1 was used to determine  $L_{\text{max}}$  and the optimal values of  $L_{\text{fwd}}$  and  $L_{\text{rev}}$ . Figure 2 shows the interpolation error,  $\bar{E}_{\nu,Q}(L_{\text{fwd}}, L_{\text{rev}})$  and  $\bar{E}_{\nu,R}(L_{\text{fwd}}, L_{\text{rev}})$ , at a representative frequency. Also shown is the reconstruction error which is defined as in (14) but for the training points,  $\mathcal{P}$ . It can be seen that, for all  $L_{\text{fwd}}$  the reconstruction error reduces with increasing  $L_{\text{rev}}$ . On the other hand, the interpolation error is not monotonically decreasing, indicating the overfitting to noise. It can be seen that for all values of  $L_{\text{fwd}}$  at least a marginal reduction in the interpolation error can be achieved by setting  $L_{\text{rev}} < L_{\text{fwd}}$ . The optimum combination of  $L_{\text{fwd}}$  and  $L_{\text{rev}}$  determined from the development set is shown as a circle in Fig. 2(b) and also in Fig. 2(c).

Figure 3(a) shows the selected values of  $L_{\text{fwd}}$  and  $L_{\text{rev}}$  over all frequencies tested using the proposed method. As a baseline, it also shows the selected truncation order when the optimization was constrained such that  $L_{\text{fwd}} = L_{\text{rev}}$ . In general, the proposed method selects a higher value of  $L_{\text{fwd}}$  and lower value of  $L_{\text{rev}}$  than the base-

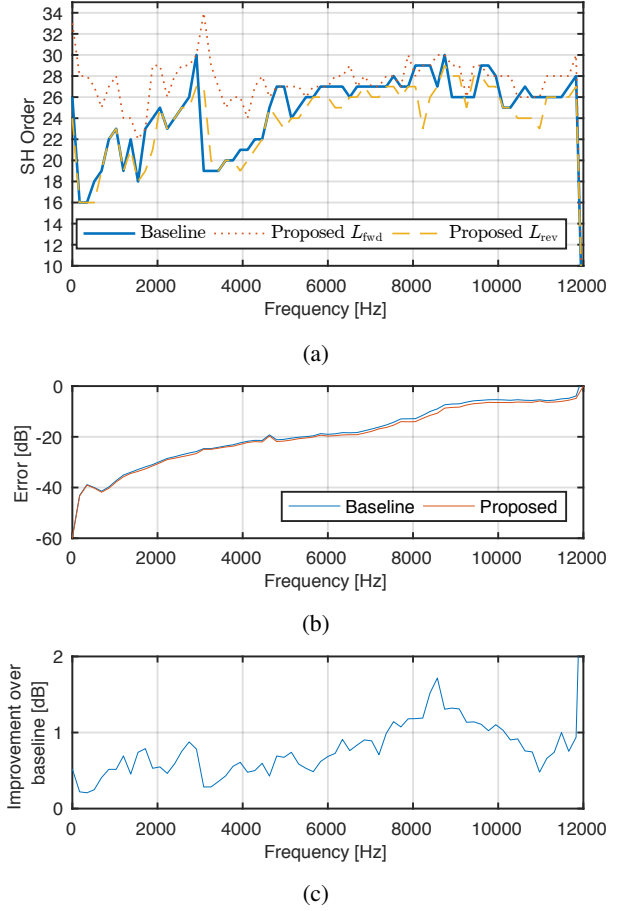


**Fig. 3.** (a) Optimum SH truncation order for baseline ( $L_{\text{fwd}} = L_{\text{rev}}$ ) and for proposed method. (b) Average interpolation error for test points using optimum values determined using development points. (c) Benefit of using the proposed method over the baseline.

line, which is consistent with the analysis in Sec. 3 that it is preferable to use a higher truncation order for the forward transform to avoid aliasing of noise, but to select a lower truncation order for the reverse transform to exclude this noise from the interpolated estimates. The error obtained for the test set with both the proposed and the baseline method is shown in Fig. 3(b). Looking at just the difference between the two, shown in Fig. 3(c), the proposed method always increases the interpolation accuracy with the improvement varying between 0.2 and 1.7 dB. Since the most computationally expensive operation in the proposed method is the calculation of  $\mathbf{Y}_{\text{fwd}}^\dagger$  which is required for both the baseline and the proposed method, the proposed method has minimal additional cost and so is preferred.

## 6. CONCLUSIONS

A new approach to spherical harmonic interpolation of array manifolds has been proposed based on an analysis of the impact of measurement noise on spatial aliasing. The proposed method consistently outperformed the baseline approach by between 0.2 and 1.7 dB allowing for more accurate interpolation of array manifolds.

## 7. REFERENCES

- [1] B. Rafaely, "Plane-wave decomposition of the pressure on a sphere by spherical convolution," *J. Acoust. Soc. Am.*, vol. 116, no. 4, pp. 2149–2157, Oct. 2004.
- [2] D. Menzies and M. Al-Akaidi, "Nearfield binaural synthesis and ambisonics," *J. Acoust. Soc. Am.*, vol. 121, no. 3, pp. 1559–1563, 2007.
- [3] B. Rafaely and A. Avni, "Interaural cross correlation in a sound field represented by spherical harmonics," *J. Acoust. Soc. Am.*, vol. 127, pp. 823, 2010.
- [4] J. Sheaffer and B. Rafaely, "Equalization strategies for binaural room impulse response rendering using spherical arrays," in *Proc. IEEE Convention of Electrical & Electronics Engineers in Israel (IEEEI)*, Eilat, Israel, 2014, IEEE.
- [5] B. Rafaely and M. Kleider, "Spherical microphone array beam steering using Wigner-D weighting," *IEEE Signal Process. Lett.*, vol. 15, pp. 417–420, 2008.
- [6] M. J. Evans, J. A. S. Angus, and A. I. Tew, "Analyzing head-related transfer function measurements using surface spherical harmonics," *J. Acoust. Soc. Am.*, vol. 104, no. 4, pp. 2400–2411, June 1998.
- [7] R. Duraiswami, D. N. Zotkin, and N. A. Gumerov, "Interpolation and range extrapolation of HRTFs," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2004, vol. 4, pp. iv–45.
- [8] D. N. Zotkin and R. Duraiswami, "Plane-wave decomposition of a sound scene using a cylindrical microphone array," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2009, pp. 85–88.
- [9] J. Ahrens, M. R. Thomas, and I. Tashev, "HRTF magnitude modeling using a non-regularized least-squares fit of spherical harmonics coefficients on incomplete data," in *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*. IEEE, 2012, pp. 1–5.
- [10] W. Zhang, M. Zhang, R. A. Kennedy, and T. D. Abhayapala, "On high-resolution head-related transfer function measurements: An efficient sampling scheme," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 2, pp. 575–584, 2012.
- [11] A. P. Bates, Z. Khalid, and R. A. Kennedy, "Novel sampling scheme on the sphere for head-related transfer function measurements," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 6, pp. 1068–1081, 2015.
- [12] G. D. Romigh, D. S. Brungart, R. M. Stern, and B. D. Simpson, "Efficient real spherical harmonic representations of head-related transfer functions," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 5, pp. 921–930, 2015.
- [13] M. Aussal, F. Alouges, and B. Katz, "A study of spherical harmonics interpolation for HRTF exchange," in *Proceedings of Meetings on Acoustics*. Acoustical Society of America, 2013, vol. 19, p. 050010.
- [14] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, first edition, 1999.
- [15] B. Rafaely, *Fundamentals of Spherical Array Processing*, Springer Topics in Signal Processing. Springer, Berlin Heidelberg, 2015.
- [16] D. P. Jarrett, E. A. Habets, and P. A. Naylor, *Theory and Applications of Spherical Microphone Array Processing*. Springer-Verlag, Berlin, Germany, 2016.
- [17] D. L. Alon and B. Rafaely, "Beamforming with optimal aliasing cancellation in spherical microphone arrays," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 1, pp. 196–210, 2016.
- [18] P. Plessas, *Rigid sphere microphone arrays for spatial recording and holography*, Ph.D. thesis, Graz University of Technology, Austria, Nov. 2009.
- [19] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 1, pp. 135–143, Jan. 2005.
- [20] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Am.*, vol. 104, no. 5, pp. 3048–3058, Jan. 1998.
- [21] C. Oreinos and J. M. Buchholz, "Measurement of a full 3D set of HRTFs for in-ear and hearing aid microphones on a head and torso simulator (HATS)," *Acta Acustica united with Acustica*, vol. 99, no. 5, pp. 836–844, 2013.