Article

Time Domain Spherical Harmonic Processing with Open Spherical Microphones Recording

Huiyuan Sun [1,*,†], **Thushara D. Abhayapala** [1,†] **and Prasanga N. Samarasinghe** [1,†]

Audio & Acoustic Signal Processing Group, College of Engineering and Computer Science, Australian National University, Canberra 2601, Australia; thushara.abhayapala@anu.edu.au (T.D.A.); prasanga.samarasinghe@anu.edu.au (P.N.S.)
* Correspondence: huiyuan.sun@anu.edu.au
† These authors contributed equally to this work.

Featured Application: Spatial Active Noise control.

Abstract: Spherical harmonic analysis has been a widely used approach for spatial audio processing in recent years. Among all applications that benefit from spatial processing, spatial Active Noise Control (ANC) remains unique with its requirement for open spherical microphone arrays to record the residual sound field throughout the continuous region. Ideally, a low delay spherical harmonic recording algorithm for open spherical microphone arrays is desired for real-time spatial ANC systems. Currently, frequency domain algorithms for spherical harmonic decomposition of microphone array recordings are applied in a spatial ANC system. However, a Short Time Fourier Transform is required, which introduces undesirable system delay for ANC systems. In this paper, we develop a time domain spherical harmonic decomposition algorithm for the application of spatial audio recording mainly with benefit to ANC with an open spherical microphone array. Microphone signals are processed by a series of pre-designed finite impulse response (FIR) filters to obtain a set of time domain spherical harmonic coefficients. The time domain coefficients contain the continuous spatial information of the residual sound field. We corroborate the time domain algorithm with a numerical simulation of a fourth order system, and show the proposed method to have lower delay than existing approaches.

Keywords: spatial audio recording; spherical harmonic; time domain signal processing

1. Introduction

Spherical harmonic analysis has been widely used for spatial acoustic signal processing for years [1]. Sound field recordings can be decomposed into a set of orthogonal spatial basis functions and respective coefficients when an appropriately designed spherical microphone array is used [2,3]. The spherical harmonic decomposition has the advantage that a given sound field can be analyzed over a continuous spatial region rather than a set of distributed points [4]. This has embraced a wide range of algorithms in three-dimensional (3D) audio signal processing such as: sound intensity analysis [5], sound field diffusive analysis [6], beamforming [7,8], source localization [9,10], and spatial Active Noise Control (ANC) [11,12].

A spatial ANC system aims to reduce the unwanted acoustic noise [13] over a space in order to create a silent zone for people. Multiple microphones are used to record the residual noise, and multiple loudspeakers are used to generate the anti-noise field. The recording's accuracy of the residual sound field can highly influence the performance of an ANC system. Furthermore, recording efficiency is also important, as ANC usually focuses on low frequency and time-variant noise. As a result, an accurate and low latency algorithm for residual sound field recording is desired [14].

The sound field recording step in a spatial ANC system focuses on obtaining the location independent spherical harmonic coefficients that represent the residual sound field inside a region of interest. This is different to real time spherical harmonic beamforming or directivity analysis which focuses on extracting source location information from the spatial recording. Moreover, spatial ANC mainly focuses on reducing the sound field inside the spherical microphone array (the region of interest). While other spatial recording applications may focus on analysing the sound field exterior to the array. Additionally, although most of the spatial audio applications utilize a rigid spherical array [15–17] for its convenience to build and use, an open spherical array is considered to be more suitable for a spatial ANC system. This is because users should be able to enter and move within the ANC region of interest that is surrounded by the spherical microphone array [12,18]. Furthermore, there exists previous work focusing on optimising the open array for spherical harmonic recording [19,20], and for spatial ANC systems [21]. However, we consider the optimisation of the open microphone array design to be outside of the scope of this paper, and instead focus on a time-domain recording algorithm.

Real-time spatial beamforming systems illustrate that applications with strict delay requirements can highly benefit from the small latency and efficient computation of time domain processing [22,23]. By posing the signal processing algorithm in the time domain, system performance can be optimized with real-valued lower order filters [24], and lower modeling delays [25]. Specifically, for a spatial ANC system, the system delay which includes the filter group delay (signal processing algorithm), the A/D and D/A converter, and the data processing delay, should be less than the acoustic delay from the reference microphones to the secondary loudspeakers in order to satisfy causality [26]. Furthermore, a longer signal processing delay slows down the convergence speed of the adaptive filtering and may lead to an unstable system [27,28]. Therefore, it is worthwhile to consider a time domain spherical harmonic decomposition method to achieve sound field recording with an open spherical array for the application of spatial ANC.

Frequency domain spherical harmonic recording has been well developed with various optimised filters [29–31]. One benefit of developing the method in the frequency domain is that the influence of the spherical Bessel zeros can be easily removed by avoiding the estimation of the coefficients at these erroneous frequency bins [19,32–34]. However, when we consider a time domain method, we can not simply avoid the Bessel zeros because we do not apply a Fourier Transformation to separate the Bessel zero frequency components from the others.

Meanwhile, there are also several works relate to time domain spatial audio signal processing. In [35], Poletti and Abhayapala give a time domain description of the free-space Green's function in the spherical harmonic domain. This provides a solution to decompose the free-space channel between a loudspeaker and microphone into the time-space domain. This work only targets the free-space Green's function, and as a result, the method is highly limited to the application of free space sound field reproduction. In [36], a time domain wave field synthesis method is presented. Although an IFFT is applied to derive the time domain solution, the work still demonstrates that time-domain wave field synthesis can be beneficial to time-varying spatial acoustic applications. In [37], Hahn and Spors offer a time domain representation of the spherical harmonic equation. They relate the time domain spherical harmonic coefficients to the sound pressure, but do not include the method of obtaining the time domain coefficients from a given recording. Time domain beamformers are designed in [38,39] with the IFFT of spherical harmonics. These papers show certain advantages for finite impulse response (FIR) filtering based signal processing systems. Overall, these time domain approaches illustrate the advantages of time domain signal processing, however, they remain unable to obtain location-independent spherical harmonic coefficients. This makes them ill-suited for spatial ANC systems, as these location-independent coefficients provide necessary information about the continuous residual sound field inside the region of interest.

In this paper, we propose a FIR filter based time domain spherical harmonic analysis method to accurately record spatial sound fields with an open spherical microphone array for the purpose of spatial ANC. We note that this work focuses solely on the problem of sound field recording, and that the spatial ANC application acts purely as motivation to our problem. Therefore, with spatial ANC in mind, the recording method prioritizes a minimum processing delay, a bandwidth of interest (low frequencies for typical noise scenarios), and a practical array geometry (open sphere surrounding a quite zone). Employing the recording method in an actual ANC system, and its evaluation, is out of the scope of this paper. The novelty of the presented work is the investigation of time domain spherical harmonic coefficients. These time domain coefficients match the properties of conventional frequency domain spherical harmonic coefficients. That is, the coefficients are location independent within the region of interest, and they represent the continuous sound field over the space. Additionally, these coefficients are obtained in time domain, which relieves the block processing constraint (and can do sample-by-sample processing) and results in lower system delay. Hence, the proposed method is considered to be highly beneficial to spatial ANC systems.

We organise the main body of this paper as follows: In Section 2 we detail the background of the frequency domain spherical harmonic algorithm for spatial sound field recording. Additionally, we introduce the time domain equation of spherical harmonic decomposition, while addressing the challenges of recording time domain spherical harmonic coefficients. The filter's design and implementation to obtain time domain spherical harmonic coefficients is presented in Section 3, along with error analysis. Effects of truncation and filter length are shown in Section 4 via initial simulations of filter performance. Section 5 presents simulation results for the proposed method's estimation of spherical harmonic coefficients, as well as sound field reconstruction performance at a point and over space, verifying the effectiveness of the proposed theory and design. We conclude the findings and insights gained from this work in Section 6.

2. Problem Formulation

We begin this section by reviewing the well-known frequency domain spherical harmonic decomposition method. We then introduce the corresponding time domain formulation, and detail the Fourier Transform relationship between the components in the frequency domain equation and the time domain equivalent. Finally, we show the difficulties in obtaining spherical harmonic coefficients in the time domain.

2.1. Spherical Harmonic Decomposition of Sound Field in Frequency-Space Domain

An incident sound field at any arbitrary point $x = (r, \theta, \phi)$ inside a source free 3D spherical region Ω, where r refers to the distance between the point x and the origin, θ and ϕ denote elevation and azimuth angles, respectively [40], can be expressed in the frequency domain as [1,41]

$$S(x, k) = \sum_{n=0}^{N} \sum_{m=-n}^{n} \alpha_{nm}(k) j_n(kr) Y_{nm}(\theta, \phi), \tag{1}$$

where order n ($n \geq 0$) and mode m are integers, $N = \lceil kR \rceil$ [1], $k = 2\pi f/c$ is the wave number, f is frequency, c is the speed of sound, R is the radius of Ω, $\alpha_{nm}(k)$ is a set of spherical harmonic coefficients representing the sound field inside Ω, $j_n(kr)$ is the nth order spherical Bessel function of the first kind, $Y_{nm}(\theta, \phi)$ are the spherical harmonic functions. For convenience, we use real spherical harmonics in this paper, given by [42]

$$Y_{nm}(\theta, \phi) = (-1)^{|m|} \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}}$$

$$\times \begin{cases} P_{nm}(\cos\theta)\cos(m\phi) & m \geq 0 \\ P_{nm}(\cos\theta)\sin(m\phi) & m < 0 \end{cases}, \tag{2}$$

where $P_{nm}(\cdot)$ is the associated Legendre function. Real spherical harmonics have the orthogonality property of

$$\int_0^{2\pi} \int_0^{\pi} Y_{nm}(\theta, \phi) Y_{n'm'}(\theta, \phi) \sin\theta d\theta d\phi = \delta_{nn'}\delta_{mm'}. \tag{3}$$

If the spherical harmonic coefficients $\alpha_{nm}(k)$ are available for a sound field, then these coefficients can fully describe the sound field over the continuous spatial region of interest. Traditionally, when spatial harmonic processing is used to record a spatial sound field $S(\boldsymbol{x}, k)$, it is recorded over a spherical surface of radius R_Q ($R_Q \geq r$). The corresponding $\alpha_{nm}(k)$ are extracted by integrating (1) over the spherical surface while exploiting the orthogonality property of $Y_{nm}(\cdot)$ in (3), which gives [2]

$$\alpha_{nm}(k) = \frac{1}{j_n(kr)} \int_0^{2\pi} \int_0^{\pi} S(r, \theta, \phi, k) Y_{nm}(\theta, \phi) \sin\theta d\theta d\phi. \tag{4}$$

In practice, this integration is realized using an equivalent discrete summation of spatial samples over the sphere.

2.2. Equivalent Spherical Harmonic Decomposition of a Sound Field in Time-Space Domain

While the frequency domain spatial sound field capture is well established as explained in Section 2.1, in this paper, our objective is to investigate the possibility of an analogous spherical harmonic analysis in time domain. In a similar fashion to (1) and (4), we now consider the relationship between sound pressure $s(\boldsymbol{x}, t)$ recorded by a spherical microphone array and the time domain spherical harmonic coefficients, denoted as $a_{nm}(t)$. It is desirable to have these time domain coefficients $a_{nm}(t)$ independent of the measurement radius. Thus, we only need to record $a_{nm}(t)$ to obtain the sound field over the entire region of interest Ω. A time domain method can directly extract $a_{nm}(t)$, thus avoiding the Fourier transformation of signals.

As a time domain analysis is usually with real-valued components, we rewrite (1) in the form of

$$S(\boldsymbol{x}, k) = \sum_{n=0}^{N} \sum_{m=-n}^{n} i^n \alpha_{nm}(k) \frac{j_n(kr)}{i^n} Y_{nm}(\theta, \phi), \tag{5}$$

where $i = \sqrt{-1}$, in order to make the inverse Fourier transform of all terms to be real. Taking the inverse Fourier transformation of (5), we obtain

$$s(\boldsymbol{x}, t) = \sum_{n=0}^{N} \sum_{m=-n}^{n} a_{nm}(t) * p_n(t, r) Y_{nm}(\theta, \phi), \tag{6}$$

where $*$ denotes the convolution operation,

$$a_{nm}(t) \xrightarrow{\mathcal{F}} i^n \alpha_{nm}(k), \tag{7}$$

where $\xrightarrow{\mathcal{F}}$ denotes the Fourier transform operator,

$$p_n(t, r) \xrightarrow{\mathcal{F}} \frac{j_n(kr)}{i^n}, \tag{8}$$

which is given by

$$p_n(t, r) = \begin{cases} \frac{c}{2r} P_n(\frac{tc}{r}) & -\frac{r}{c} \leq t \leq \frac{r}{c} \\ 0 & |t| > \frac{r}{c} \end{cases}, \tag{9}$$

where $P_n(\cdot)$ is the Legendre function. The proof of (9) is given in Appendix A. We note that every component in (6) is real valued.

Equation (6) shows how to reconstruct the sound pressure at $\boldsymbol{x} = (r, \theta, \phi)$ with the recorded time domain spherical harmonic coefficients $a_{nm}(t)$. We consider an alternative

time domain filter to obtain $a_{nm}(t)$ from the recorded signals rather than taking the inverse Fourier transform of (4) since $1/j_n(kr)$ is unbounded when $j_n(kr) = 0$. Note that $j_n(kr)$ as a filter has order dependent zeros when $j_n(kr) = 0$. As a result, $1/j_n(kr)$ approaches infinity at these frequencies, making it unstable to have an inverse Fourier transform. In other words, the z-transform of $p_n(t,r)$ given in (9), has zeros on the unit circle because of Bessel zeros, refers to a non-minimum phase system. In this case, the inverse system of $p_n(t,r)$, with the frequency response of $1/j_n(kr)$ is not stable. As a result, we first define

$$b_{nm}(t,r) \triangleq a_{nm}(t) * p_n(t,r), \tag{10}$$

which has a frequency response of

$$b_{nm}(t,r) \xrightarrow{\mathcal{F}} i^n \alpha_{nm}(k) \frac{j_n(kr)}{i^n} = \alpha_{nm}(k) j_n(kr). \tag{11}$$

Since $Y_{nm}(\theta, \phi)$ is independent to both frequency and time, $b_{nm}(t,r)$ can be obtained by integrating (6) over a sphere of radius r such that

$$b_{nm}(t,r) = \int_0^{2\pi} \int_0^{\pi} S(r, \theta, \phi, t) Y_{nm}(\theta, \phi) \sin(\theta) d\theta d\phi. \tag{12}$$

If we regularly place $Q \geq (N+1)^2$ omni-directional microphones on a sphere of radius R_Q, we can estimate the integration in (12) with a finite summation such that

$$b_{nm}(t, R_Q) \approx \sum_{q=1}^{Q} S(\mathbf{x}_q, t) Y_{nm}(\theta_q, \phi_q). \tag{13}$$

To simplify the implementation, we sample the signals with sampling time T such that $t = \nu T = \nu/F_s$, where ν is the time index and F_s is the sampling frequency. We rewrite (10) as

$$
b_{nm}(\nu, R_Q) = a_{nm}(\nu) * p_n(\nu, R_Q)
$$
$$
= \sum_{\mu=-L_p}^{L_p} p_n(\mu, R_Q) a_{nm}(\nu - \mu), \tag{14}
$$

where

$$p_n(\nu, R_Q) = \begin{cases} \frac{c}{2R_Q} P_n\left(\frac{\nu c}{R_Q F_s}\right) & -\frac{R_Q F_s}{c} \leq \nu \leq \frac{R_Q F_s}{c} \\ 0 & |\nu| > \frac{R_Q F_s}{c} \end{cases}, \tag{15}$$

is a time limited function with $p_n(\nu, r) \neq 0$ when $-L_p \leq \nu \leq L_p$, $L_p = \lceil R_Q F_s / c \rceil$ such that the length of $p_n(\nu, R_Q)$ is $2L_p + 1$.

With (14) in hand, our problem reduces to obtaining $a_{nm}(\nu)$ from the measured $b_{nm}(\nu, R_Q)$. This is not achievable since it is an under-determined problem. We always have $2L_p + 1$ more unknowns ($a_{nm}(\nu)$) than knowns ($b_{nm}(\nu, R_Q)$). Moreover, this is not practically feasible because the z-transform of $p_n(\nu, R_Q)$ has zeros on the unit circle, resulting in poles on the unit circle in its direct inverse, making the system unstable. Alternatively, $a_{nm}(\nu)$ can be extracted from $b_{nm}(\nu, R_Q)$ using an appropriately designed filter.

In this paper, we attempt to design a filtering solution while overcoming the above challenges. It is important to note that the Fourier transform relationship discussed in this section were solely used to formulate the definition of the time-domain spherical harmonic decomposition of a sound field. From this point onward, we will focus on signal processing of the captured sound field only in the time domain.

3. Filter Design for Obtaining Time Domain Spherical Coefficients

In Section 2, we have presented a method to obtain $b_{nm}(\nu, R_Q)$ from recorded sound

pressure $S(x_q, \nu)$ with a spherical microphone array. In this section, we design a series of FIR filters to obtain $a_{nm}(\nu)$ from given $b_{nm}(\nu, R_Q)$.

3.1. Stability of Ideal Inverse Filter

Due to the challenges mentioned in Section 2, rather than directly using (14), we pre-design a series of filters $\rho_n(\nu, r)$ such that

$$
\begin{aligned}
b_{nm}(\nu, R_Q) * \rho_n(\nu, R_Q) \\
= a_{nm}(\nu) * p_n(\nu, R_Q) * \rho_n(\nu, R_Q) = a_{nm}(\nu),
\end{aligned}
\tag{16}
$$

where

$$
p_n(\nu, R_Q) * \rho_n(\nu, R_Q) \approx \delta(\nu).
\tag{17}
$$

We note here that $\rho_n(\nu, r)$ should be order n dependent but mode m independent, as is the same property with $p_n(\nu, r)$.

However, we can never achieve a precise $\delta(\nu)$ in (17), as the energy of measured sound pressure at the frequency bins of Bessel zeros has been filtered to zero by $p_n(\nu, R_Q)$. Therefore, we refrain from designing the inverse filter at these zero positions. Instead, we modify $\delta(\nu)$ to $\hat{z}_n(\nu)$ such that its frequency response $\hat{Z}_n(f)$ is given by

$$
\hat{Z}_n(f) = \begin{cases} 1 & |j_n(kr)| \geq \epsilon \\ 0 & |j_n(kr)| < \epsilon \end{cases},
\tag{18}
$$

where ϵ is a small positive constant threshold which satisfies $j_n(kr) \approx 0$ when $|j_n(kr)| < \epsilon$. For a fixed R_Q, both $j_n(2\pi f R_Q/c)$ and $\hat{Z}_n(f)$ can be seen as a function of f. Figure 1 shows $j_n(2\pi f R_Q/c)$ and $\hat{Z}_n(f)$ with $\epsilon = 1/40$ of the first four orders of n.
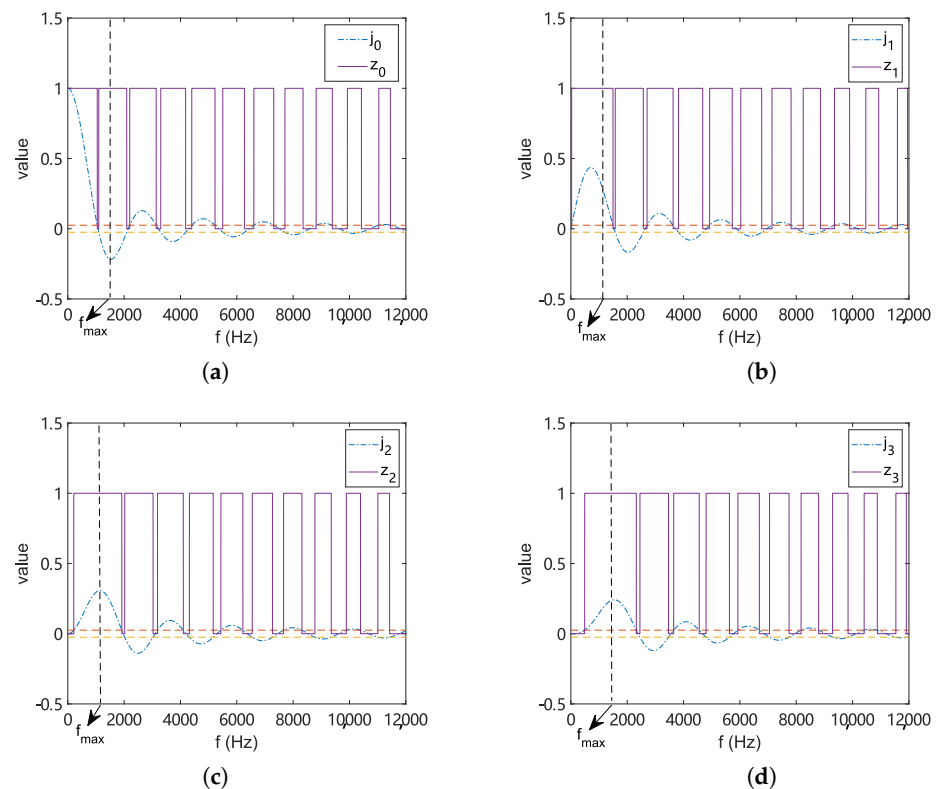

Figure 1. The spherical Bessel function $j_n(2\pi f R_Q/c)$ and $\hat{Z}_n(f)$ of order (**a**) $n = 0$, (**b**) $n = 1$, (**c**) $n = 2$, (**d**) $n = 3$ with $f_{max} \approx 1360$ Hz, $\epsilon = 1/40$, $R_Q = 0.16$ m and $c = 343$ m/s.

From Figure 1, we can see that $\hat{Z}_n(f)$ is a superposition of a series of rectangular windows, meaning its inverse Fourier transformation, $\hat{z}_n(\nu)$, should be a superposition

of sinc functions. In practice, due to inherent properties of $j_n(2\pi f R_Q/c)$, for a given maximum frequency $f_{\max}$, the number of active spherical harmonic orders is up to $N \approx \lceil kR_Q \rceil$ [1]. We use the same truncation limit when designing $\hat{Z}_n(f)$, resulting in $\hat{z}_n(\nu)$ to be a superposition with a finite number of sinc functions. The necessity and the influence of this truncation on frequency $f_{\max}$ will be further discussed in Section 4.1.

Let us define $w^{(n)}$ in radian (rad), such that $j_n(w^{(n)}F_sR_Q/c) = \epsilon$, where ϵ is the positive threshold we explained in the last paragraph. Therefore, $w^{(n)}$ can be considered as the edges of window in $\hat{Z}_n(f)$ (see Figure 1). Given the vector of $[w_1^{(n)}, w_2^{(n)}, w_3^{(n)}, \cdots]$, we can write $\hat{z}_n$ as

$$
\begin{aligned}
\hat{z}_n(\nu) = \sum_{s=1}^{S} (w_{2s}^{(n)} - w_{2s-1}^{(n)}) \mathrm{sinc}\Big(\frac{w_{2s}^{(n)} - w_{2s-1}^{(n)}}{2}\nu\Big) \\
\times \cos\Big(\frac{w_{2s}^{(n)} + w_{2s-1}^{(n)}}{2}\nu\Big),
\end{aligned}
\tag{19}
$$

where S is the number of rectangular windows in $\hat{Z}_n(f)$ for $-f_{\max} \le f \le f_{\max}$. Furthermore, $w^{(n)}$s are dependent on the radius of the microphone array R_Q, sampling frequency F_s and the speed of sound c, but the value of $w^{(n)}F_sR_Q/c$ remains constant for each order n such that $|j_n(w^{(n)}F_sR_Q/c)| = \epsilon$. The first four order of $w^{(n)}$ is given in Table 1 with the highest frequency limit of $f_{\max} = 2047$ Hz and sampling frequency $F_s = 48,000$ Hz. Note that for the zero-th order, we set $w_1 = 8.9 \times 10^{-4}$ to block DC component in practice.

If we have a series of concentric spherical microphone arrays with the radii of $r_1, r_2, \cdots$, the value of $w^{(n)}F_sr_q/c$ would be different from a single sphere model, which can be calculated by $|j_n(w^{(n)}r_1/F_sc) + j_n(w^{(n)}r_2/F_sc) + \cdots| = \epsilon$.

Table 1. The first four order of $w^{(n)}$ to derive $\hat{z}_n(\nu)$ with $f_{\max} \approx 2047$ Hz, $\epsilon = 1/40$, $R_Q = 0.16$ m and $c = 343$ m/s.

n (Order)	$\mathbf{w}_1^{(n)}$	$\mathbf{w}_2^{(n)}$	$\mathbf{w}_3^{(n)}$	$\mathbf{w}_4^{(n)}$
0	0.0009	0.1369	0.1439	0.2680
1	0.0033	0.1957	0.2059	0.2680
2	0.0276	0.2509	0.3643	0.2680
3	0.0640	0.2680	-	-

3.2. Modified Inverse Filter

Now that the design for $\hat{z}_n(\nu)$ is established, our next step is to design filters $\rho_n(\nu, R_Q)$ which satisfies

$$
p_n(\nu, R_Q) * \rho_n(\nu, R_Q) = \hat{z}_n(\nu).
\tag{20}
$$

We notice in (20) that $p_n(\nu, R_Q)$ is a finite length vector and we would like $\rho_n(\nu, R_Q)$ also to be a finite length vector. However, $\hat{z}_n(\nu)$ is infinitely long with a series of sinc functions. If we perform linear convolution of $p_n(\nu, R_Q)$ with $\rho_n(\nu, R_Q)$, we would obtain a vector with the length of $2(L + L_p) + 1$ samples, where $2L + 1$ is the filter length of ρ_n, such that $\rho_n(\nu, R_Q)$ has none-zero values for $-L \le \nu \le L$. Thus, we need to truncate the infinite length $\hat{z}_n(\nu)$ to $2(L + L_p) + 1$ samples for every order of n where

$$
z_n(\nu) \triangleq \begin{cases} \hat{z}_n(\nu) & -(L + L_p) \le \nu \le L + L_p \\ 0 & \text{otherwise} \end{cases}.
\tag{21}
$$

We can then write (20) in a finite summation form as

$$
\begin{aligned}
z_n(\nu) &= p_n(\nu, R_Q) * \rho_n(\nu, R_Q) \\
&= \sum_{\mu=-L}^{L} p_n(\nu - \mu, R_Q)\rho_n(\mu, R_Q).
\end{aligned}
\tag{22}
$$

We rewrite (22) into matrix form

$$z_n = P_n \rho_n, \tag{23}$$

where $z_n = [z_n(-(L + L_p)), z_n(-(L + L_p) + 1), \cdots, z_n((L + L_p))]^T$, $\rho_n = [\rho_n(-L, R_Q), \rho_n(-L + 1, R_Q), \cdots, \rho_n(L, R_Q)]^T$, and P_n is the convolution matrix based on the Toeplitz structure of $p_n(\nu, R_Q)$, given in (24).

$$P_n = \begin{bmatrix} p_n(-L_p, R_Q) & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ p_n(-L_p + 1, R_Q) & p_n(-L_p, R_Q) & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & & \ddots & & & & \vdots \\ p_n(L_p, R_Q) & \cdots & p_n(-L_p, R_Q) & 0 & \cdots & \cdots & 0 \\ 0 & p_n(L_p, R_Q) & \cdots & p_n(-L_p, R_Q) & 0 & \cdots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \cdots & 0 & p_n(L_p, R_Q) & \cdots & p_n(-L_p + 1, R_Q) & p_n(-L_p, R_Q) \\ 0 & \cdots & \cdots & 0 & p_n(L_p, R_Q) & \cdots & p_n(-L_p + 1, R_Q) \\ \vdots & & & & \ddots & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & p_n(L_p, R_Q) & p_n(L_p - 1, R_Q) \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & p_n(L_p, R_Q) \end{bmatrix}. \tag{24}$$

The size of matrix P_n is $[2(L + L_p) + 1, 2L + 1]$, where we choose the filter length $2L + 1$ of $\rho_n(\nu, R_Q)$ to be significantly larger than both $2L_p + 1$ and the main lobe width of function $z_n(\nu)$, to avoid P_n being ill-conditioned and minimize the error of truncating $z_n(\nu)$ into a finite length signal. The influence of choosing L will be detailed in Section 4.2.

Since (23) is an over-determined system of equations, we apply LMS method to (23) to obtain

$$\rho_n = P_n^+ z_n, \tag{25}$$

where P_n^+ refers to the Moore-Penrose inverse of P_n. As a result, with (16) and (25), $a_{nm}(\nu)$ can be estimated by

$$\begin{aligned} a_{nm}(\nu) &\approx b_{nm}(\nu, R_Q) * \rho_n(\nu, R_Q) \\ &= \left(\sum_{q=1}^{Q} S(x_q, t) Y_{nm}(\theta_q, \phi_q) \right) * \rho_n(\nu, R_Q). \end{aligned} \tag{26}$$

In this way we obtain $a_{nm}(\nu)$ while overcoming the challenges listed in Section 2.

3.3. Practical Considerations of Filter Implementation

In (26), $a_{nm}(\nu)$ is obtained by filtering $b_{nm}(\nu, R_Q)$ with $\rho_n(\nu, R_Q)$, where we get

$$b_{nm}(\nu, R_Q) * \rho_n(\nu, R_Q) = a_{nm}(\nu) * z_n(\nu) \approx a_{nm}(\nu). \tag{27}$$

Naturally $a_{nm}(\nu)$ at time index ν is only influenced by $[b_{nm}(\nu - L_p, R_Q), b_{nm}(\nu - L_p + 1, R_Q), \cdots, b_{nm}(\nu + L_p, R_Q)]$ because of the Legendre function in $p_n(\nu, R_Q)$. However, with the influence of sinc functions in $z_n(\nu)$ in our proposed filters $\rho_n(\nu, R_Q)$, we now need the past L samples and the future L samples of $b_{nm}(\nu, R_Q)$ to obtain $a_{nm}(\nu)$ at time index ν. For offline signal processing, L samples of zeros should be added both in the beginning and the end of the vector of b_{nm} before filtering it with pre-designed $\rho_n(\nu, R_Q)$. Moreover, an overlap of $2L + 1$ samples is needed for frame based signal processing. For on-line real time signal processing, we cannot obtain future samples of $b_{nm}(\nu, R_Q)$. As a result, we add L samples of zeros in front of the filter $\rho_n(\nu, R_Q)$, and create a buffer of the past $2L + 1$ samples of $b_{nm}(\nu, R_Q)$. At time index ν, we obtain $a_{nm}(\nu - L)$ with the buffer of $[b_{nm}(\nu - 2L, R_Q), \cdots, b_{nm}(\nu, R_Q)]$. Thus, there is a L samples of group delay of the

system. We further discuss and compare the group delay with frequency domain method in Section 5.5.

3.4. Error Analysis

We define the error $e_{nm}(v)$ as the difference between the desired time domain spherical harmonic and the coefficients we obtained by the proposed method, which can be decomposed to:

$$e_{nm}(v) = e_{\text{filter}}(n, m, v) + e_{\text{position}}(n, m) + e_{\text{truncation}}(n), \tag{28}$$

where $e_{\text{filter}}(n, m, v)$ is filtering error introduced by $\rho_n(v, R_Q)$, $e_{\text{truncation}}(n)$ is the truncation error of order N, and $e_{\text{position}}(n, m)$ is due to the microphones position error. The qualitative analysis of $e_{\text{truncation}}(n)$ and $e_{\text{position}}(n, m)$ based on the frequency domain method are addressed in [3], where we draw a similar conclusion in time domain that with increasing number of microphones and fixed N, $e_{\text{truncation}}(n)$ decreases. Meanwhile, $e_{\text{position}}(n, m)$ depends on the nature of inaccurate microphone positioning, referring to the distance between the desired point and microphone location. We mainly focus on $e_{\text{filter}}(n, m, v)$ here as it is the main error contribution due to the proposed filtering approach.

According to (27), $e_{\text{filter}}(n, m, v)$ at a specific order n and mode m can be expressed as

$$\begin{aligned} e_{\text{filter}}(n, m, v) &= |a_{nm}(v) * z_n(v) - a_{nm}(v)| \\ &= |a_{nm}(v) * e_n(v)|, \end{aligned} \tag{29}$$

where

$$e_n(v) \triangleq \delta(v) - z_n(v). \tag{30}$$

Using (18) and (21), the Fourier transform of $e_n(v)$ is

$$e_n(v) \xrightarrow{\mathcal{F}} E_n(f) = \begin{cases} 1 & |j_n(kr)| \le \epsilon, kr < N \\ 0 & \text{otherwise} \end{cases}, \tag{31}$$

with the same truncation in frequency as $\hat{Z}_n(f)$. Thus, $e_n(v)$ can be expressed as

$$\begin{aligned} e_n(v) = \sum_{s=0}^{S-1} \left(w_{2s+1}^{(n)} - w_{2s}^{(n)} \right) \text{sinc} \left(\frac{w_{2s+1}^{(n)} - w_{2s}^{(n)}}{2} v \right) \\ \times \cos \left(\frac{w_{2s+1}^{(n)} + w_{2s}^{(n)}}{2} v \right), \end{aligned} \tag{32}$$

where S and $w^{(n)}$ have the same definition as in (19) and $w_0^{(n)} = 0$.

With (29) and (32) we can quantitatively calculate the filtering error $e_{\text{filter}}(n, m, v)$ introduced by $\rho_n(v, R_Q)$. The total error caused by filtering can be calculated by a summation of $e_{\text{filter}}(n, m, v)$ over every order of n and mode of m. As this filtering error is mainly due to Bessel zeros, it can be reduced by limiting the highest order N of the system, where a smaller N results in lower Bessel zeros hence a smaller $e_{\text{filter}}(n, m, v)$. Also, N depends on the highest wave number k and the radius of the microphone array R_Q. By choosing N with a pre-knowledge of the frequency limit of the input signals and R_Q also helps to minimize the filtering error $e_{\text{filter}}(n, m, v)$.

4. A Filter Design Example

To provide a further understanding of the filter design process, we present a design example of a fourth ($N = 4$) order spherical microphone array of $R_Q = 0.16$ m, designed to record the time domain spherical harmonic coefficients within the spatial region enclosed by the array with a desired frequency band of $[20, 1360]$ Hz. Let $F_s = 48,000$ Hz and $c = 343$ m/s. Before we apply the proposed method to recording signals, we first analyze the influence of several steps in designing the proposed filter $\rho_n(v, R_Q)$.

4.1. Effect of Frequency Truncation of $Z_n(\nu)$

As audio signals are often band limited in ANC applications [14], we can have a finite truncation on spherical harmonic decomposition with order $N = \lceil kR_Q \rceil$. In other words, if we have a fixed N-th order system, the highest frequency that the system can successfully capture is given by $f_{\max} = Nc/(2\pi R_Q) \approx 1360$ Hz. Figure 2 shows the frequency response of $\rho_n(\nu, R_Q)$, refers to $\Phi_n(f, R_Q)$, which is designed using (25) with $z_n(\nu)$ truncated at $f_1 = 1023.6$ Hz (Figure 2a), $f_2 = 1364.8$ Hz (Figure 2b), $f_3 = 2047.1$ Hz (Figure 2c), respectively. The filter length is set to be 500. To obtain the frequency response of $\rho_n(\nu, R_Q)$, a FFT of $I = 4096$ points is applied with zero padding to $\rho_n(\nu, R_Q)$. We remind here that $z_n(\nu)$ is given by (19) in time domain, which does not rely on any frequency domain processing.
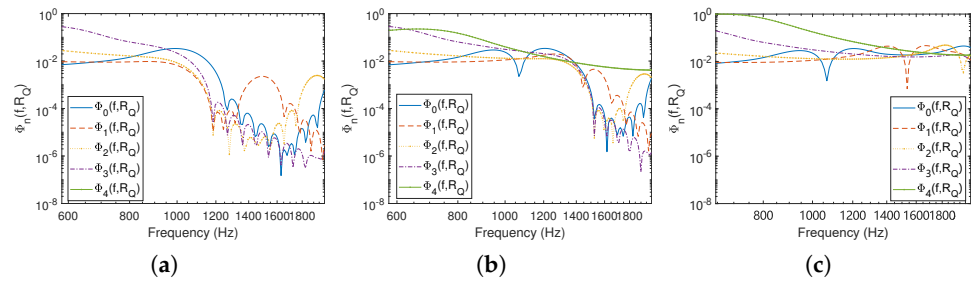

Figure 2. Frequency response of up to 4-th order of the pre-designed order dependent FIR filter $\rho_n(\nu, R_Q)$ with $z_n(\nu)$ frequency truncated at (**a**) f_1, (**b**) f_2 (**c**) f_3.

We observe that for a $N = $ 4th order system, the truncation at f_1 is not enough to get an accurate frequency response of $\rho_n(\nu, R_Q)$, as the frequency response $\Phi_n(f, R_Q)$ begins to decline at f_1. In this case, $\rho_n(\nu, R_Q)$ can not provide an acceptable filtering result with signals containing higher frequency components. Truncation at both f_2 and f_3 can give a satisfied frequency response when $f < f_{\max}$. As the frequency range of the system is also limited by $N = \lceil kR_Q \rceil$, it is not necessary to look at the frequency response when $f > f_{\max}$. So in both cases $\rho_n(\nu, R_Q)$ can give an acceptable filtering output. As a result, we choose to truncate $z_n(\nu)$ at $f_{\max}$, where $2\pi f_{\max} R_Q/c = N$. If the recorded signal is known as a band limited signal where its highest frequency component is less than $f_{\max}$, an alternative choice of the frequency truncation of $z_n(\nu)$ is at this highest frequency to reduce the computation complexity. Meanwhile, if $z_n(\nu)$ has been designed with a higher frequency truncation, it can also be used in a lower order system with a lower requirement of frequency truncation.

4.2. Choice of Filter Length of $\rho_n(\nu, R_Q)$

Intuitively, a longer filter often brings us less error and better performance. Figure 3 supports this idea by showing the result of $\rho_n(\nu, R_Q) * p_n(\nu, R_Q) - z_n(\nu)$ with different choices of L, which refers to the error introduced into the system by the filtering processing. We observe that the filtering error decreases across all of the orders with a higher L. This is due to the time truncation of $z_n(\nu)$ (length of vector z_n in (25)), being related to L. Thus, a higher L leads to less information loss in the time truncation of $z_n(\nu)$, hence smaller error in $\rho_n(\nu, R_Q)$. However, Figure 4 shows the time domain filter $\rho_n(\nu, R_Q)$ with different lengths. We observe that a longer filter results in a higher group delay of filtering. This is not desirable because it leads to a higher system delay of our proposed method, while lowering the system delay is one of the most important motivations that we develop the proposed time domain method.
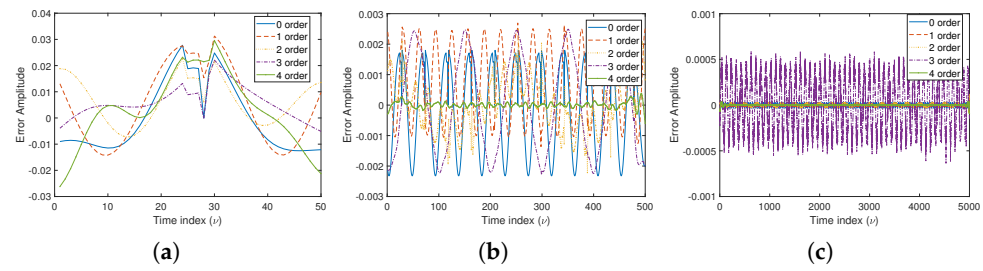
Figure 3. Error of $\rho_n(\nu, R_Q) * p_n(\nu, R_Q)$ with length $L =$ (**a**) 25, (**b**) 250 (**c**) 2500.
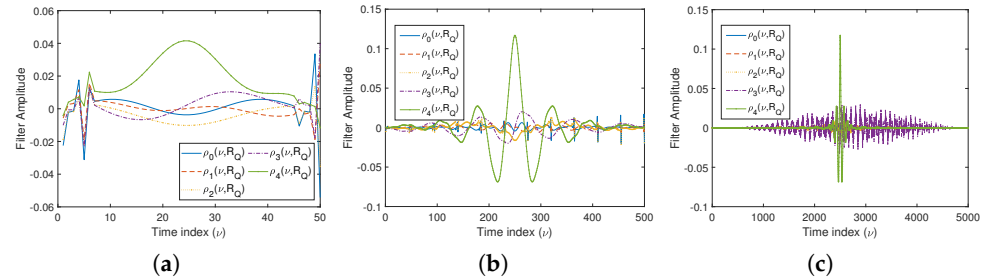

Figure 4. Time representation of the pre-designed order dependent FIR filter $\rho_n(\nu, R_Q)$ with length $L =$ (**a**) 25, (**b**) 250 (**c**) 2500.

As a result, we need to balance the noise tolerance, group delay, and the filtering error when we choose L. We suggest that filter length $2L + 1$ should be significantly larger than the main lobe width of $z_n(\nu)$ and $2L_P + 1$, the length of $p_n(\nu, R_Q)$, but no more than 50 times of $2L_P + 1$. Additionally, L should be less than the maximum tolerance of the delay of the system. Based on these guidelines, for the current example, we choose $2L + 1 = 501$.

5. Simulation Results and Analysis

In this section, we evaluate the result of the proposed algorithm for time domain spherical harmonic analysis using a fourth order ($N = 4$) system. We consider 32 microphones regularly placed on an open spherical array of $R_Q = 0.16$ m, where the analysis region of interest is inside the array. A point source is placed at $[1, 2, 1]$ m with respect to the origin which coincides with the origin of the microphone array. The sampling frequency is 48,000 Hz, and the filter length $2L + 1$ is 501. A noise signal at 40 dB SNR is added to each microphone to reflect thermal noise. Considering the application of the proposed method to be spatial ANC, we construct the desired frequency band to cover the target noise band, and construct the radius of the region to be wide enough to fit one human head.

It is difficult to validate our method in time domain directly because the coefficients are time dependent and no ground truth has been given. Therefore, we first validate our proposed time domain spherical harmonic coefficients in the frequency domain. Thus, we compare the Fourier transformation of the time domain coefficients to the theoretical frequency domain coefficients given in (4). Next, to clarify that our proposed method has the ability to record a sound field in the region of interest in the time domain, we reconstruct sound pressure at an arbitrary point as well as over a plane inside the region of interest with the captured time domain spherical harmonic coefficients by (6). Finally, the time delay of the proposed method is given.

5.1. Comparison between the Time Domain and the Frequency Domain Spherical Harmonic Coefficients

We use a narrow band signal at 1200 Hz to test if our proposed method can obtain time domain spherical harmonic coefficients $a_{nm}(\nu)$ correctly with (26). In (11), we give the relationship between $a_{nm}(\nu)$ and $\alpha_{nm}(k)$. We compare the Fourier transformation result of our obtained time domain spherical harmonic coefficients $\mathcal{FT}\{a_{nm}(\nu)\}$ with the desired frequency domain spherical harmonic coefficients $\alpha_{nm}(k)$, obtained by Equation (4) in

frequency domain. Fourier transformations use $J = 1024$ points. We do not compare the phase of these coefficients since the group delay of the time domain method and the frequency domain method is different. Instead, we compare the phase difference, given by $\alpha_{nm}(k) - \alpha_{n(m-1)}(k)$. The results of both amplitude and phase difference are shown in Figure 5.
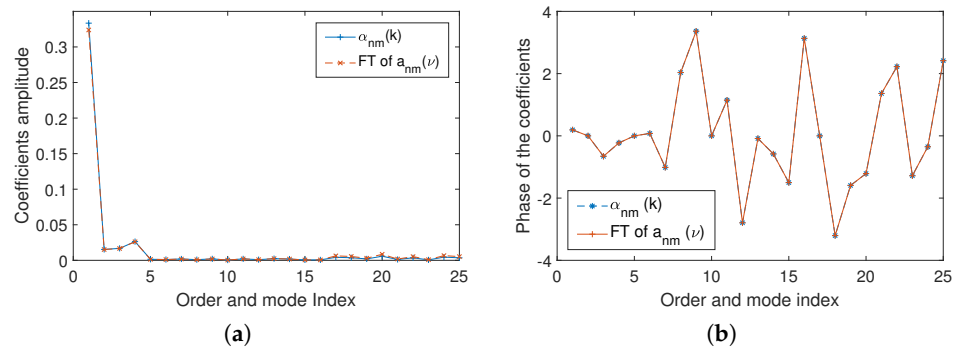

(a)

(b)

Figure 5. (**a**) Amplitude and (**b**) phase difference comparison between the Fourier Transform of the time domain spherical harmonic coefficients $a_{nm}(\nu)$ and frequency domain spherical harmonic coefficients $\alpha_{nm}(k)$ at a single frequency $f = 1200$ Hz.

In Figure 5 we see that there is little to no difference on both amplitude and phase difference between the Fourier Transformed time domain coefficients and the frequency domain coefficients over all the order and modes. Thus, our proposed time domain method successfully obtained the time domain spherical harmonic coefficients, which can be related to the frequency domain coefficients by Fourier transformation.

Next, we compare the coefficients over different frequencies with a wide band test signal within the frequency limited of $[20, 1300]$ Hz. In Figure 6, we show the comparison of amplitude at $FT\{a_{00}(\nu)\}$ and $\alpha_{00}(k)$, $FT\{a_{11}(\nu)\}$ and $\alpha_{11}(k)$, and $FT\{a_{31}(\nu)\}$ and $\alpha_{31}(k)$ over frequencies respectively while Figure 7 shows the phase difference.
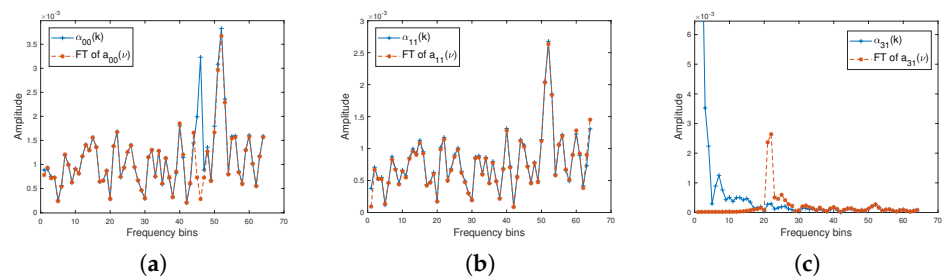

(a)

(b)

(c)

Figure 6. Amplitude comparison between the Fourier Transform of the time domain spherical harmonic coefficients $a_{nm}(\nu)$ and frequency domain spherical harmonic coefficients $\alpha_{nm}(k)$ at mode (**a**) 00, (**b**) 11 and (**c**) 31 with a white Gaussian noise.

A huge error is observed in Figure 6a at the 46th frequency bin. This error is due to that there is a Bessel zero of the zeroth order at this frequency bin (around 1072 Hz). We see the frequency domain spherical harmonic coefficients $\alpha_{00}(k)$ has a much higher amplitude, while our proposed method suppressed the amplitude at this certain frequency bin. Meanwhile, we can see in Figures 6 and 7 that the error at $a_{31}(\nu)$ is higher compared to the other two modes. As order increases, the error increases. This error can be decreased by applying more microphones on the array. We also obtain a non-negligible error before the 30th frequency bin of the coefficients amplitude for $(n, m) = (3, 1)$ in Figures 6c and 7c. This error is because our time domain proposed method and conventional frequency domain method have different processing on suppressing Bessel zeros. During the reconstruction process, the high pass property of spherical Bessel function removes the information at this frequency bin. Thus, this error will not influence the reconstruction of the sound field.
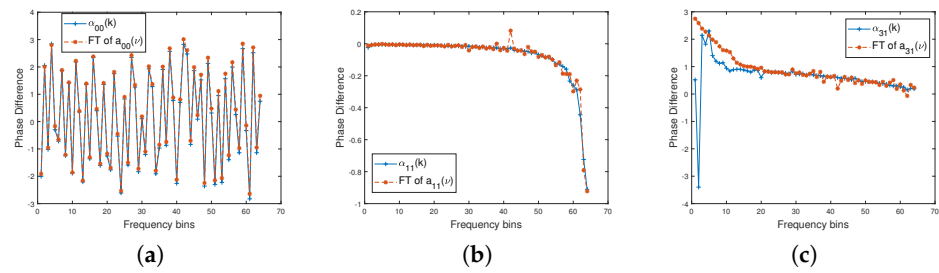
Figure 7. Phase difference comparison between the Fourier Transform of the time domain spherical harmonic coefficients $a_{nm}(\nu)$ and frequency domain spherical harmonic coefficients $\alpha_{nm}(k)$ at mode (**a**) 00, (**b**) 11 and (**b**) 31 with a white Gaussian noise.

5.2. Sound Pressure Comparison at a Point Of Interest

In this section, we reconstruct the sound field with the captured time domain spherical harmonic coefficients at a point in the region of interest, and compare it with the desired sound field at the same point of interest. We use a signal containing three frequency components of 600 Hz, 850 Hz, and 1300 Hz. Figure 8 shows the desired sound pressure and the reconstructed sound pressure calculated by $a_{nm}(\nu)$ at the point $[-0.13, 0.07, 0.02]$ m and $[-0.03, 0.01, 0.1]$ m inside the region of interest in time domain. The desired sound field has been manually delayed for 272 samples to match the group delay of the reconstructed sound field, where the details of this delay will be shown in Section 5.5.
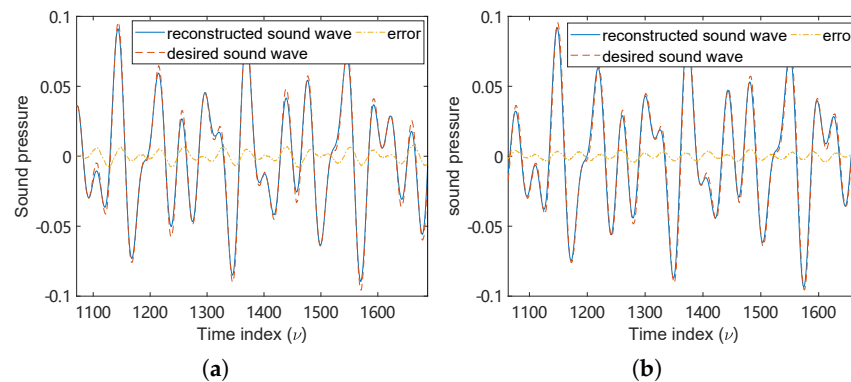

Figure 8. Comparison between reconstructed sound pressure and desired sound pressure at the point (**a**) $(-0.13, 0.07, 0.02)$ m and (**b**) $(-0.03, 0.01, 0.1)$ m.

We note here that when reconstructing the sound-field with (6), we face the problem that at a point $x = (r, \theta, \phi)$ where the radius r is very small, the filter $p_n(\nu, r)$, whose filter length dependents on rF_s/c, is too short to perform efficient filtering. To overcome this problem, we up-sample the obtained $a_{nm}(\nu)$ with a rate of R_Q/r and construct corresponding $p_n(\nu, r)$ with the same length of $L_p = 2 * R_Q F_s/c + 1$. We then down-sample the resulting $b_{nm}(\nu, r)$ with a rate of r/R_Q to keep the sampling frequency consistent with F_s. We can see that the obtained $a_{nm}(\nu)$ by our proposed method can successfully reconstruct the sound pressure at a point inside the region of interest with a tolerable error. This supports that our time domain coefficients contain certain spatial information of the sound field that the sound pressure at an arbitrary point inside the region of interest can be properly calculated with the measurements only being taken on the boundary of the region.

5.3. Sound Field Comparison over a Plane

To further evaluate our method on reconstructing sound field over space, we now reconstruct the sound field by $a_{nm}(\nu)$ over a plane. We use a narrow band signal of 1200Hz here that the sound field in the region of interest is simple and clearly understood. Although the sound field is reconstructed over time, a 2D plot can only show the result

of one time index. Figure 9 shows the reconstructed sound field and the desired sound field over the plane parallel to the x-y plane, with $z = 0.02$ m at $t = 0.3$ s. The 272 samples group delay is manually fixed and will be discussed later in the next subsection.
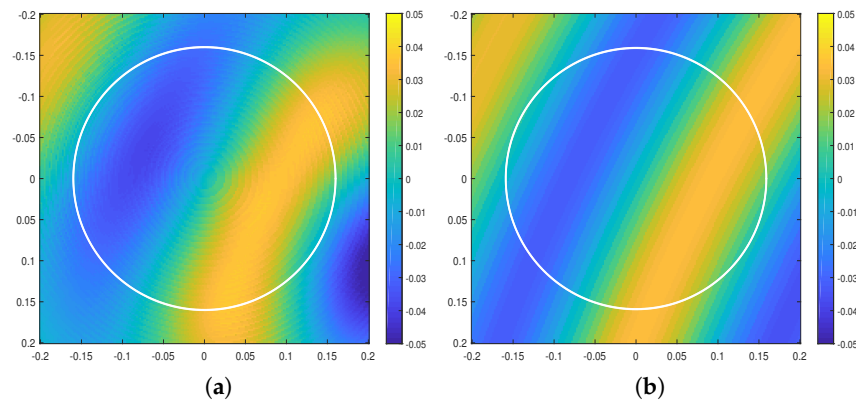

Figure 9. Comparison between (**a**) reconstructed sound field and (**b**) desired sound field at the horizontal plane with a height of z = 0.02 m.

The white line in Figure 9 bounds the region of interest. We can see that the reconstructed sound field inside this region in Figure 9a is roughly the same as the desired sound field in Figure 9b. This confirms that the coefficients recorded by our proposed method are able to capture the sound field inside the region of interest.

5.4. Sound Field Error Estimation over The Region

To evaluate the reconstructed sound field over time, we calculate the instantaneous average squared spatial error over time, which is defined by

$$e(\nu) \triangleq \frac{\sum_\Omega \|S_r(\boldsymbol{x}, \nu) - S_d(\boldsymbol{x}, \nu)\|^2}{\sum_\Omega}. \tag{33}$$

Figure 10 shows how the error fluctuates with time in a tolerable range (no more than 5×10^{-4}) with a 900 Hz tone and a 1072 Hz tone. We have already observed in Figure 8 that the error of the sound pressure at a point of interest is proportional to the desired sound pressure. We observe the same trend when we evaluate the error over the region that the error increases when the sound field inside the region of interest is at peak amplitude. We also observe in Figure 10 that the error with 1072 Hz signal is higher than 900 Hz signal. This is due to that there is a Bessel zero of the zeroth order ($j_0(kr)$) at 1072 Hz in the proposed spatial ANC system. Hence, the amplitude of $a_{00}(\nu)$ is suppressed by the proposed method, leading to a higher error in reconstructing the sound field.
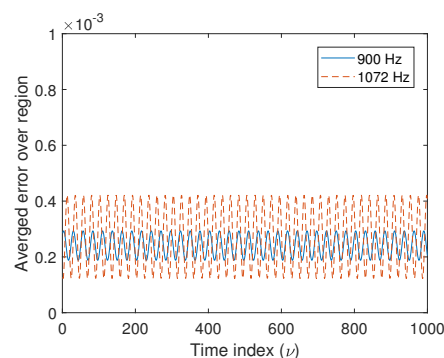

Figure 10. Instantaneous region averaged squared spatial error of the proposed method for sound field reconstruction over space at 900 Hz and 1072 Hz (Bessel zero).

5.5. Processing Delay Analysis

In this section, we indicate the group delay of our method. Figure 11 shows the desired sound pressure and the reconstructed sound pressure of a signal containing three frequency components of $[600, 850, 1300]$ Hz at the point $[-0.13, 0.07, 0.02]$ m. We can obtain from Figure 11 that the processing delay of the system is $1046 - 774 = 272$ samples, which equals to $L + \lceil R_Q F_s / c \rceil$. The L samples of the delay is from the group delay of the proposed filter $\rho_n(\nu, R_Q)$, while $R_Q F_s / c$ is the delay introducing by the Legendre function within filter $p_n(\nu, r)$ to reconstruct the sound pressure at a point with the time domain spherical harmonic coefficients. Comparing to a conventional frequency domain scenario with 512 frame-size and 75% of overlap Short Time Fourier transformation, which refers to a 2048 samples of delay [43], our proposed method can significantly reduce the processing delay.

Comparing to one of the start-of-art frequency domain spherical harmonic filter designs [44], which states a 75 ms delay with a 900 sample long filter, our method can achieve a 972 samples (20.25 ms with 48k Hz sampling frequency) delay with the same length of filter. Meanwhile, as our method is processed in time domain, there is nothing to stop us from doing a sample by sample signal processing instead of frame based signal processing. This sample based processing considerably extends the application of spherical harmonic analysis.
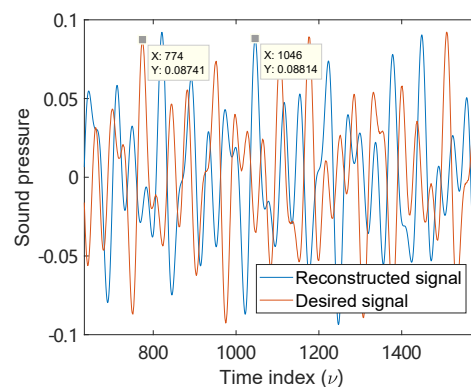

Figure 11. Delay analysis between the desired signal and the reconstructed signal at a point inside the region of interest.

6. Conclusions

In this paper, a time domain spherical harmonic analysis method for spatial sound field recording over 3D space has been developed with the goal to minimize processing delay. This favours the specific application of spatial ANC. With the proposed FIR filter design, the time domain spherical harmonic coefficients can be obtained from the sound pressure measurements of an open spherical microphone array. The filters are designed based on the inverse of the Legendre function. Additionally, the filters are modified with considerations of stability and practical implementation. We have provided simulation results proving the validity of the proposed method.

We note that by obtaining the proposed time domain spherical harmonic coefficients, the desired sound field can be efficiently captured and reconstructed over space. The proposed time domain spherical harmonic coefficients can be related to the conventional frequency domain coefficients, where both have the same location independent property. The proposed method has the prominent advantage of lower delay since it is developed in the time domain without the introduction of a Fourier transformation or inverse Fourier transformation. Furthermore, the proposed time domain filtering method can support sample based signal processing instead of frame based, which indicates that the frame size can be one sample if necessary. As a result, we consider the proposed time domain spherical harmonic analysis method to be highly suitable for a spatial ANC system where accurate spatial recording with low delay is desired.

The most important future work is practically introducing the proposed spatial recording method to a spatial ANC system. Currently the proposed method utilizes open spherical microphone arrays, where the difficulties of constructing this array limit the potential applications. Hence, applying the proposed method to alternative optimised open microphone arrays is another direction for future work.

Author Contributions: Conceptualization, H.S., T.D.A. and P.N.S.; Funding acquisition, T.D.A. and P.N.S.; Investigation, H.S.; Methodology, H.S., T.D.A. and P.N.S.; Project administration, T.D.A.; Supervision, T.D.A. and P.N.S.; Validation, H.S.; Writing—original draft, H.S.; Writing—review and editing, T.D.A. and P.N.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANC	Active noise control
3D	Three-dimensional
IFFT	Inverse fast fourier transform
FIR	Finite impulse response
SNR	Signal to noise ratio

Appendix A. Proof of Equation (9)

We have the Fourier relationship between the spherical Bessel function $j_n(kr)$ and the Legendre function $P_n(t)$ given by [45]

$$\int_{-\infty}^{\infty} e^{ikrt} j_n(kr) dk = \pi i^n P_n(t). \tag{A1}$$

With (A1), $p_n(t,r)$ in (8) can be express as

$$\begin{aligned} p_n(t,r) &= \frac{c}{i^n 2\pi r} \int_{-\infty}^{\infty} j_n(kr) e^{\frac{itckr}{r}} dkr \\ &= \begin{cases} \frac{c}{2r} P_n\left(\frac{tc}{r}\right) & -\frac{r}{c} \le t \le \frac{r}{c} \\ 0 & \pm t > \frac{r}{c} \end{cases}. \end{aligned} \tag{A2}$$

This completes the prove of (9).

References

1. Ward, D.B.; Abhayapala, T.D. Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Trans. Speech Audio Process.* **2001**, *9*, 697–707. [CrossRef]
2. Abhayapala, T.D.; Ward, D.B. Theory and design of high order sound field microphones using spherical microphone array. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, USA, 13–17 May 2002; Volume 2, pp. 1949–1952.
3. Rafaely, B. Analysis and design of spherical microphone arrays. *IEEE Trans. Speech Audio Process.* **2004**, *13*, 135–143. [CrossRef]
4. Park, M.; Rafaely, B. Sound-field analysis by plane-wave decomposition using spherical microphone array. *J. Acoust. Soc. Am.* **2005**, *118*, 3094–3103. [CrossRef]
5. Zuo, H.; Samarasinghe, P.N.; Abhayapala, T.D.; Dickins, G. Spatial sound intensity vectors in spherical harmonic domain. *J. Acoust. Soc. Am.* **2019**, *145*, EL149–EL155. [CrossRef] [PubMed]
6. Epain, N.; Jin, C.T. Spherical harmonic signal covariance and sound field diffuseness. *IEEE/ACM Trans. Audio Speech Lang. Process. (TASLP)* **2016**, *24*, 1796–1807. [CrossRef]

7. Koretz, A.; Rafaely, B. Dolph–Chebyshev beampattern design for spherical arrays. *IEEE Trans. Signal Process.* **2009**, *57*, 2417–2420. [CrossRef]

8. Gover, B.N.; Ryan, J.G.; Stinson, M.R. Microphone array measurement system for analysis of directional and spatial variations of sound fields. *J. Acoust. Soc. Am.* **2002**, *112*, 1980–1991. [CrossRef]

9. Argentieri, S.; Danes, P.; Soueres, P. Modal analysis based beamforming for nearfield or farfield speaker localization in robotics. In Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 9–15 October 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 866–871.

10. Birnie, L.; Abhayapala, T.D.; Chen, H.; Samarasinghe, P.N. Sound Source Localization in a Reverberant Room Using Harmonic Based Music. In Proceedings of the ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 651–655.

11. Spors, S.; Buchner, H. An approach to massive multichannel broadband feedforward active noise control using wave-domain adaptive filtering. In Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, USA, 21–24 October 2007; pp. 171–174.

12. Zhang, J.; Abhayapala, T.D.; Zhang, W.; Samarasinghe, P.N.; Jiang, S. Active Noise Control Over Space: A Wave Domain Approach. *IEEE/ACM Trans. Audio Speech Lang. Process. (TASLP)* **2018**, *26*, 774–786. [CrossRef]

13. Cassina, L.; Fredianelli, L.; Menichini, I.; Chiari, C.; Licitra, G. Audio-visual preferences and tranquillity ratings in urban areas. *Environments* **2018**, *5*, 1. [CrossRef]

14. Kuo, S.M.; Morgan, D.R. *Active Noise Control Systems: Algorithms and DSP Implementations*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1995.

15. Baumgartner, R.; Pomberger, H.; Frank, M. Practical implementation of radial filters for ambisonic recordings. In Proceedings of the ICSA 2011, Detmold, Germany, 10–13 November 2011.

16. Moreau, S.; Daniel, J.; Bertet, S. 3D sound field recording with higher order ambisonics–Objective measurements and validation of a 4th order spherical microphone. In Proceedings of the 120th Convention of the AES, Paris, France, 20–23 May 2006; pp. 20–23.

17. Zotter, F. A linear-phase filter-bank approach to process rigid spherical microphone array recordings. In Proceedings of the 5th International Conference on Electrical, Electronic and Computing Engineering, Palić, Serbia, 11–14 June 2018.

18. Sun, H.; Abhayapala, T.D.; Samarasinghe, P.N. Time Domain Spherical Harmonic Analysis for Adaptive Noise Cancellation over a Spatial Region. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 516–520.

19. Chen, H.; Abhayapala, T.D.; Zhang, W. Theory and design of compact hybrid microphone arrays on two-dimensional planes for three-dimensional soundfield analysis. *J. Acoust. Soc. Am.* **2015**, *138*, 3081–3092. [CrossRef]

20. Abhayapala, T.D.; Gupta, A. Spherical harmonic analysis of wavefields using multiple circular sensor arrays. *IEEE Trans. Audio Speech Lang. Process.* **2010**, *18*, 1655–1666. [CrossRef]

21. Sun, H.; Abhayapala, T.D.; Samarasinghe, P.N. Active Noise Control Over 3D Space with Multiple Circular Arrays. In Proceedings of the 2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, USA, 20–23 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 135–139.

22. Bilbao, S.; Ahrens, J.; Hamilton, B. Incorporating source directivity in wave-based virtual acoustics: Time-domain models and fitting to measured data. *J. Acoust. Soc. Am.* **2019**, *146*, 2692–2703. [CrossRef] [PubMed]

23. Farina, A.; Capra, A.; Chiesi, L.; Scopece, L. A spherical microphone array for synthesizing virtual directive microphones in live broadcasting and in post production. In Proceedings of the Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space, Tokyo, Japan, 8–10 October 2010; Audio Engineering Society: New York, NY, USA, 2010.

24. Mabande, E.; Schad, A.; Kellermann, W. A time-domain implementation of data-independent robust broadband beamformers with lowfilter order. In Proceedings of the 2011 Joint Workshop on Hands-free Speech Communication and Microphone Arrays, Edinburgh, UK, 30 May–1 June 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 81–85.

25. Simón Gálvez, M.F.; Elliott, S.J.; Cheer, J. Time domain optimization of filters used in a loudspeaker array for personal audio. *IEEE/ACM Trans. Audio Speech Lang. Process. (TASLP)* **2015**, *23*, 1869–1878. [CrossRef]

26. Nelson, P.A.; Elliott, S.J. *Active Control of Sound*; Academic Press: Cambridge, MA, USA, 1991.

27. Long, G.; Ling, F.; Proakis, J.G. The LMS algorithm with delayed coefficient adaptation. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 1397–1405. [CrossRef]

28. Long, G.; Ling, F.; Proakis, J.G. Corrections to 'The LMS algorithm with delayed coefficient adaptation'. *IEEE Trans. Signal Process.* **1992**, *40*, 230–232. [CrossRef]

29. Lösler, S.; Zotter, F. Comprehensive radial filter design for practical higher-order Ambisonic recording. In Proceedings of the Fortschritte der Akustik DAGA, Nuremberg, Germany, 16–19 March 2015; pp. 452–455.

30. Jin, C.T.; Epain, N.; Parthy, A. Design, optimization and evaluation of a dual-radius spherical microphone array. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2013**, *22*, 193–204. [CrossRef]

31. Politis, A.; Gamper, H. Comparing modeled and measurement-based spherical harmonic encoding filters for spherical microphone arrays. In Proceedings of the 2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 15–18 October 2017, New Paltz, NY, USA; IEEE: Piscataway, NJ, USA, 2017; pp. 224–228.

32. Balmages, I.; Rafaely, B. Open-sphere designs for spherical microphone arrays. *IEEE Trans. Audio Speech Lang. Process.* **2007**, *15*, 727–732. [CrossRef]

33.   Chardon, G.; Kreuzer, W.; Noisternig, M. Design of spatial microphone arrays for sound field interpolation. *IEEE J. Sel. Top. Signal Process.* **2015**, *9*, 780–790. [CrossRef]

34.   Ueno, N.; Koyama, S.; Saruwatari, H. Sound field recording using distributed microphones based on harmonic analysis of infinite order. *IEEE Signal Process. Lett.* **2017**, *25*, 135–139. [CrossRef]

35.   Poletti, M.; Abhayapala, T.D.; Teal, P.D. Time domain description of spatial modes of 2D and 3D free-space greens functions. In Proceedings of the Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control, Guildford, UK, 18–20 July 2016; Audio Engineering Society: New York, NY, USA, 2016.

36.   Winter, F.; Hahn, N.; Spors, S. Time-domain realisation of model-based rendering for 2.5 D local wave field synthesis using spatial bandwidth-limitation. In Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO), Kos, Greece, 28 August–2 September 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 688–692.

37.   Hahn, N.; Spors, S. Time-Domain Representations of a Plane Wave with Spatial Band-Limitation in the Spherical Harmonics Domain. In Proceedings of the Meeting of the German Acoustical Society (DAGA), Rostock, Germany, 18–21 March 2019.

38.   Yan, S.; Sun, H.; Ma, X.; Svensson, U.P.; Hou, C. Time-domain implementation of broadband beamformer in spherical harmonics domain. *IEEE Trans. Audio Speech Lang. Process.* **2010**, *19*, 1221–1230.

39.   Ren, W.; Chen, H.; Gao, W. On the design of time-domain implementation structure for steerable spherical modal beamformers with arbitrary beampatterns. *Appl. Acoust.* **2017**, *122*, 146–151. [CrossRef]

40.   Arfken, G.B.; Weber, H.J. *Mathematical Methods for Physicists*; Academic Press: San Diego, CA, USA, 1999.

41.   Williams, E.G. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*; Elsevier: Amsterdam, The Netherlands, 1999.

42.   Homeier, H.H.H.; Steinborn, E.O. Some properties of the coupling coefficients of real spherical harmonics and their relation to Gaunt coefficients. *J. Mol. Struct. THEOCHEM* **1996**, *368*, 31–37. [CrossRef]

43.   Oppenheim, A.V. *Discrete-Time Signal Processing*; Pearson Education India: Tamil Nadu, India, 1999.

44.   Langrenne, C.; Bavu, E.; Garcia, A. A linear phase IIR filterbank for the radial filters of ambisonic recordings. In Proceedings of the EAA Spatial Audio Signal Processing Symposium, Paris, France, 6–7 September 2019.

45.   Chang, H.P.; Sarkar, T.K.; Pereira-Filho, O.M. Antenna pattern synthesis utilizing spherical Bessel functions. *IEEE Trans. Antennas Propag.* **2000**, *48*, 853–859.