Optimal beamformer design in spherical sector harmonics domain ☆

Deepika Kumari [a], Lalan Kumar [b],*

[a] *Department of Electrical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India*
[b] *Department of Electrical Engineering and Bharti School of Telecommunication, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India*

ARTICLE INFO

ABSTRACT

Beamforming using spherical microphone arrays (SMAs) with processing in spherical harmonics (SH) domain is widely being studied. This is due to the ease of array processing in SH domain with no spatial ambiguity. Despite the widespread applications of SMA in beamforming, its size makes it inconvenient for practical or commercial use. Building of SMA over a rigid sphere is also a challenging task. Additionally, use of the entire sphere comes at the cost of more microphones and signals to process. Hence, it is uneconomic to use full SMA when sources are only present in restricted regions of the environment. Attempts have been made in literature to use hemispherical microphone array for beamforming. Acoustic image principle was utilized to enable application of SH but with greater computational complexity. In this paper, the use of a spherical sector microphone array is proposed for beamforming. An orthonormal spherical sector harmonics (S^2H) basis function is used for data model development. An ideal direction-invariant beampattern using S^2H basis function is utilized. Subsequently, to make the beamformer robust against sensor noise with a sharp directivity pattern at the direction of arrival and zero elsewhere, a new maximum directivity and white noise gain (MD-WNG) beamformer is designed for a spherical sector microphone array. Performance analysis of the proposed beamformer is presented using various experiments on directivity pattern and array gain. The proposed method is additionally compared with several existing beamforming techniques. The comparison reveals that the S^2H based beamforming provides reasonably better results compared to the existing methods.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Beamforming using a sensor array has been an active area of research because of its application in the noise reduction, radar, sonar, speech processing, room reflections analysis and bio-medicine [1–4]. In particular, spherical microphone arrays (SMAs) have been widely used for beamforming [5–9]. The widespread application of SMAs is due to ease of array signal processing in spherical harmonics (SH) domain without any spatial ambiguity [10]. Owing to the similarity in the formulation of various problems in the spatial and SH domains, results in the spatial domain can be directly applied in SH domain. Additionally, SMA has the benefit of providing directional independent beampattern [11].

A rigid SMA was utilized in [11] for beamforming. The approach was based on a spherical harmonic decomposition of the sound field over SMA. A delay and sum beamforming was proposed in the spherical harmonics domain in [12]. The designed beamformer in the SH domain was more robust towards the isotropic noise sources compared to the conventional delay and sum beamformer but with poor white noise gain (WNG) at a certain frequency. A WNG constrain was imposed in [13] for optimal beampattern design. The method optimally converges to the desired beampatterns under specified robustness constraints. Frequency-invariant maximum directivity (MD) beamformer was discussed in [14]. The beamformer reduces the effects of reverberation and spread noise sources but suffers from poor WNG at low frequencies. To improve the WNG, a flexible high directivity (HD) beamformer was proposed in [8]. The beamforming filter coefficients were regularized in [15] to alleviate the white noise amplification and improve the robustness of the beamformer.

Despite the widespread applications of SMA in beamforming, its size makes it inconvenient for practical or commercial use. Building of SMA over a rigid sphere is also a challenging task. Additionally, use of the entire sphere comes at the cost of more microphones and signals to process. Hence, it is uneconomic to use full SMA when sources are only present in restricted regions of the environment. A hemispherical microphone array (HMA) was utilized in [16] for sound acquisition and beamforming. The sources were placed on one side of a rigid plane. As the data was

---

present only on the upper half of a sphere, the acoustic image principle (AIP) was utilized therein to construct pressure over the entire sphere and thus enabling the application of spherical harmonics. The proposed configuration and the principle were applicable only when a rigid plane is attached to the bottom of the hemispherical array resulting in limited applications. This also adds to the computational complexity as it utilizes imaginary microphones and sources. It also requires to maintain uniformity and symmetry across the boundary of a real and imaginary hemisphere. Microphone array placed over one-eighth, quarter and half-spaces was utilized in [17] for beamforming. A new basis function is extracted from the SHs basis by only keeping those that have a null derivative with respect to θ and ϕ over the boundaries. The rigid boundary also introduces an image beampattern and degrades the array performance.

In general, spherical harmonics are utilized to represent function defined over the entire sphere. Accurate representation of data over the hemisphere by SH requires more number of SH coefficients due to discontinuities at the boundary of the hemisphere [18]. Hence, hemispherical harmonics (HSH) basis functions were utilized for representing the bidirectional reflectance distribution function (BRDF) hemispherical function in [19], brain source localization in [20], and acoustic source localization in [21]. An orthonormal spherical sector harmonics (S²H) basis function over a general spherical sector microphone array is derived in [22] for sound-field representation and is subsequently utilized for source localization and beamforming in [23]. An ideal directional-invariant beampattern with unit gain at the desired direction and zeroes elsewhere is modeled using S²H basis function.

In this work, an optimal S²H based beamforming is presented using microphone array over spherical sector. The novel contributions are listed herein.

- A data model for maximum directivity and white noise gain (MD-WNG) based robust beamformer is developed in S²H domain.
- The proposed robust MD-WNG beamformer additionally achieves the maximum possible DF when compared with the existing methods.
- The inherent rigid boundary assumption in state-of-the-art method results in larger side lobe with shifted main lobe. Spherical sector harmonics based approach has no such assumptions and limitation.

2. Signal model

A microphone array with I identical and omnidirectional microphones is uniformly distributed over the surface of a spherical sector. The sector is defined by elevation $[\theta_1, \theta_2]$, azimuth $[\phi_1, \phi_2]$ and radius r as shown in Fig. 1. The elevation angle θ is measured downward from positive Z axis. The azimuthal angle ϕ is measured counter clockwise from positive X axis. For the plausible solution to wave equation over the sector, ϕ_1 is taken to be zero and $\phi_2 = 2\pi/u$, where $u \geqslant 1$ [22]. $\Phi_i = (\theta_i, \phi_i)$ denotes the angular position of the ith microphone. A plane wave is incident on the array from a far-field source with wave number k and direction of arrival (DOA) as $\Psi = (\theta_l, \phi_l)$. The signal received at I microphones can be written in the spatio-frequency domain under free field assumption as [3]

$$\begin{aligned} \mathbf{P}(k) &= [P_1(k), P_2(k), \ldots, P_I(k)]^T, \\ &= \mathbf{a}(k)S(k) + \mathbf{V}(k), \end{aligned} \tag{1}$$

where $S(k)$ is the target signal, $\mathbf{V}(k) = [V_1(k), V_2(k), \ldots, V_I(k)]^T$ is uncorrelated sensor noise, and $[.]^T$ denotes the transpose. The steering vector $\mathbf{a}(k)$ for the plane wave incident from (θ_l, ϕ_l) is expressed as
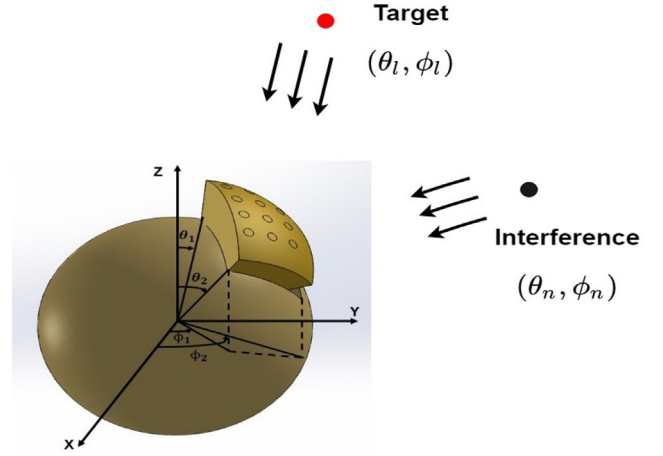

Fig. 1. Illustration of S²MA system. The sector assumed with $\theta \in [\theta_1, \theta_2]$ and $\phi \in [\phi_1, \phi_2]$.

$$\mathbf{a}(k) = \left[e^{-j\mathbf{k}_l^T \mathbf{r}_1}, \quad e^{-j\mathbf{k}_l^T \mathbf{r}_2}, \quad \ldots, \quad e^{-j\mathbf{k}_l^T \mathbf{r}_I} \right]^T, \tag{2}$$

where $\mathbf{k}_l = -[k\sin\theta_l\cos\phi_l, \quad k\sin\theta_l\sin\phi_l, \quad k\cos\theta_l]^T, j = \sqrt{-1}$, and $\mathbf{r}_i$ represents the position vector of the ith microphone, given by

$$\mathbf{r}_i = [r\sin\theta_i\cos\phi_i, \quad r\sin\theta_i\sin\phi_i, \quad r\cos\theta_i]^T. \tag{3}$$

For the spherical sector microphone array (S²MA), the microphones are present only over the sector. Hence, S²H basis function can more appropriately be utilized to represent the pressure over the sector. Consequently, the unit amplitude plane wave solution (given by the ith term in (2)) over the sector can be rewritten in spherical co-ordinate system as [22,23]

$$e^{-j\mathbf{k}_l^T \mathbf{r}_i} = \sum_{n=0}^{N} \sum_{m=-n}^{n} \tilde{b}_n(kr) \left[T_n^m(\theta_l, \phi_l) \right]^\star T_n^m(\theta_i, \phi_i), \tag{4}$$

where $T_n^m(\theta, \phi)$ is the spherical sector harmonics, given by

$$\begin{aligned} T_n^m(\Phi) &= K_n^m \tilde{\mathscr{P}}_n^m(\cos\theta)\tilde{e}^{jm\phi}, \forall 0 \leqslant n \leqslant N, 0 \leqslant m \leqslant n, \\ &= (-1)^{|m|} T_n^{|m|\star}(\Phi), \forall -n \leqslant m < 0, \end{aligned} \tag{5}$$

where $\tilde{\mathscr{P}}_n^m(x)$ is shifted associated Legendre polynomial. $\tilde{\mathscr{P}}_n^m(x)$ is related to the associated Legendre polynomial $(\mathscr{P}_n^m(x))$ as $\tilde{\mathscr{P}}_n^m(x) = \mathscr{P}_n^m(q_1 x + q_2)$ with $x = \cos(\theta)$. $\tilde{e}^{jm\phi} = e^{jmu\phi}$ is scaled exponential function, and K_n^m is the normalization constant given by

$$K_n^m = \sqrt{\frac{(2n+1)(n-m)!(q_1 \times u)}{4\pi(n+m)!}}. \tag{6}$$

The value of q_1 and q_2 are given by the solutions of the equations [23]

$$\begin{pmatrix} \cos(\theta_2) & 1 \\ \cos(\theta_1) & 1 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \tag{7}$$

resulting from orthogonality of shifted associated Legendre polynomials. The order N in (4) is chosen based on the inequality [24]

$$kr < N \leqslant \sqrt{I} - 1. \tag{8}$$

The far-field sectoral mode strength $\tilde{b}_n(kr)$ for rigid and open (free field) sector is given by

$$\tilde{b}_n(kr) = \begin{cases} \frac{4\pi j^n}{q_1^{(n+1)} u} \left(j_n(kr) - \frac{j_n'(kr)}{h_n^{(2)'}(kr)} h_n^{(2)}(kr) \right) & \\ & \text{(rigid spherical sector)}, \\ \frac{4\pi j^n}{q_1^{(n+1)} u} j_n(kr) & \text{(open spherical sector)}. \end{cases} \tag{9}$$

Here, $j_n(kr)$ denotes the spherical Bessel function of first kind. $h_n^{(2)}(kr)$ represents spherical Hankel function of second kind, and $(\prime)$ refers to the first derivative. The sectoral mode strength variation with frequency and radius r is presented in Fig. 2. An open spherical sector with $\theta \in [0°, 90°]$ and $\phi \in [0, 360°]$ is considered. It may be noted that as the range r increases, $\tilde{b}_n(kr)$ assumes the value 0 more frequently. Utilizing (4) in (2) the steering vector can be re-written as

$$\mathbf{a}(k) = \mathbf{T}(\Phi)\mathbf{B}(kr)\mathbf{t}^H(\Psi), \tag{10}$$

where $[.]^H$ stands for conjugate transpose operator, and $\mathbf{T}(\Phi)$ is a $I \times (N+1)^2$ matrix whose ith row is given by

$$\mathbf{t}(\Phi_i) = \begin{bmatrix} T_0^0(\Phi_i), & T_1^{-1}(\Phi_i), & T_1^0(\Phi_i), & \ldots, & T_N^N(\Phi_i) \end{bmatrix}. \tag{11}$$

$\mathbf{t}(\Psi)$ is a $1 \times (N+1)^2$ vector can be obtained by replacing Φ_i with Ψ in (11). $\mathbf{B}(kr)$ is $(N+1)^2 \times (N+1)^2$ matrix given by

$$\mathbf{B}(kr) = \text{diag}\{\tilde{b}_0(kr), \quad \tilde{b}_1(kr), \quad \tilde{b}_1(kr), \quad \tilde{b}_1(kr), \ldots, \tilde{b}_n(kr)\}. \tag{12}$$

The spherical sector harmonics decomposition of the received pressure $P(k)$ over the sector, is given by

$$\begin{aligned} P_{nm}(k) &= \int_{\phi_1}^{\phi_2} \int_{\theta_1}^{\theta_1} P(k, \theta, \phi) \left[T_n^m(\theta, \phi) \right]^\star \sin(\theta) d\theta d\phi \\ &\cong \sum_{i=1}^{I} a_i P_i(k, \Phi_i) \left[T_n^m(\Phi_i) \right]^\star, \end{aligned} \tag{13}$$

where the sampling weight of ith microphone (a_i) is chosen based on the sampling scheme [14]. Eq. (13) for $n \in [0, N]$ and $m \in [-n, n]$, can be expressed in matrix form as

$$\mathbf{P}_{nm}(k) = \mathbf{T}^H(\Phi)\mathbf{\Gamma P}(k), \tag{14}$$

where $\mathbf{P}_{nm}(k) = [P_{00}, P_{1-1}, P_{10}, P_{11}, \ldots, P_{NN}]^T$ and $\mathbf{\Gamma} = \text{diag}\{a_1, a_2, a_3, \ldots a_I\}$. The orthonormal relation of the spherical sector harmonics suggests [23]

$$\mathbf{T}^H(\Phi)\mathbf{\Gamma T}(\Phi) \approx \mathbf{I}, \tag{15}$$

where $\mathbf{I}$ is an identity matrix of dimension $(N+1)^2 \times (N+1)^2$. Substituting (10) in (1), multiplying $\mathbf{T}^H(\Phi)\mathbf{\Gamma}$ on both side, and utilizing (14) and (15), transform the data in the spherical sector harmonics domain, given by

$$\mathbf{P}_{nm}(k) = \mathbf{B}(kr)\mathbf{t}^H(\Psi)\mathbf{S}(k) + \mathbf{V}_{nm}(k), \tag{16}$$

where $\mathbf{V}_{nm}(k) = \mathbf{T}^H(\Phi)\mathbf{\Gamma}V(k)$. Based on the comparison of the data model in (1) and (16), the new steering matrix can be written in S²H domain as $\mathbf{a}_{nm}(k) = \mathbf{B}(kr)\mathbf{t}^H(\Psi)$.

3. Maximum Directivity-WNG Beamformer for S²MA

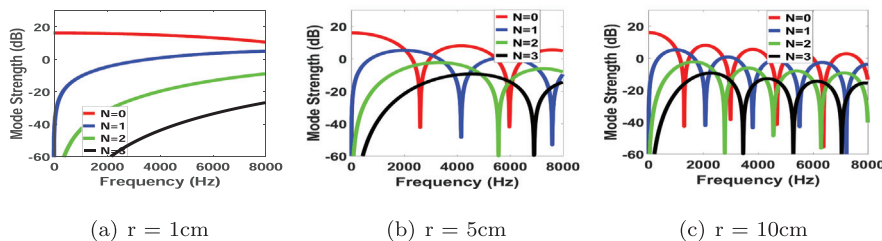Having formulated the data model in spherical sector harmonics domain, maximum directivity-WNG beamformer for S²MA is

presented in this section. The beamformed output in S²H domain is given by

$$\mathbf{Z}_{nm}(k) = \mathbf{w}_{nm}^H(k)\mathbf{P}_{nm}(k), \tag{17}$$

where $\mathbf{w}_{nm}(k)$ is the spherical sector harmonics decomposition of $\mathbf{w}(k)$. Here, $\mathbf{w}(k)$ represents the beamforming coefficient/weight vector, expressed as

$$\mathbf{w}(k) = [W_1(k), W_2(k), \ldots, W_I(k)]^T, \tag{18}$$

where $W_i(k)$ is the beamforming weight at the ith microphone. The sensitivity of the beamformer is evaluated using beampattern. The beampattern due to a unit amplitude plane wave incident on the array from an angle (θ_l, ϕ_l) is expressed as [14]

$$B[\mathbf{w}(k), \theta_l, \phi_l] = \sum_{i=1}^{I} W_i^*(k) e^{-j\mathbf{k}_l^T \mathbf{r}_i}, \tag{19}$$

Substituting (4) in (19), we have

$$B[\mathbf{w}(k), \theta_l, \phi_l] = \sum_{n=0}^{N} \sum_{m=-n}^{n} \tilde{b}_n(kr) \left[T_n^m(\theta_l, \phi_l) \right]^\star \sum_{i=1}^{I} W_i^*(k) T_n^m(\theta_i, \phi_i) \tag{20}$$

3.1. Ideal Beampattern

Beampattern provides response of an array steered at direction (θ_s, ϕ_s) for source incident from various directions (θ_l, ϕ_l). In particular, the beampattern with large gain at steered direction (θ_s, ϕ_s), and no gain at all other directions (θ_l, ϕ_l) is modeled using a delta function (ideal beampattern) [8], given by

$$B[\mathbf{w}(k), \theta_l, \phi_l] = \delta(\cos(\theta_l) - \cos(\theta_s))\delta(\phi_l - \phi_s). \tag{21}$$

The completeness relation of spherical sector harmonics suggests [23]

$$\sum_{n=0}^{\infty} \sum_{m=-n}^{n} T_n^m(\theta_s, \phi_s) \left[T_m^m(\theta_l, \phi_l) \right]^\star = \frac{1}{|q_1||u|} \delta(\cos(\theta_l) - \cos(\theta_s)) \times \delta(\phi_l - \phi_s). \tag{22}$$

Substituting the completeness relation in (21) under limited order assumption (N), we have

$$B[\mathbf{w}(k), \theta_l, \phi_l] = |q_1||u| \sum_{n=0}^{N} \sum_{m=-n}^{n} T_n^m(\theta_s, \phi_s) \left[T_m^m(\theta_l, \phi_l) \right]^\star. \tag{23}$$

It is to be noted from (20) and orthonormal relation of S²H that the beampattern expression in (23) can be achieved by using the weight as

$$W_i^\star(k) = \sum_{n=0}^{N} \sum_{m=-n}^{n} \frac{|q_1||u|}{\tilde{b}_n(kr)} \left[T_n^m(\theta_i, \phi_i) \right]^\star T_n^m(\theta_s, \phi_s). \tag{24}$$

Fig. 3 illustrates the 3D and 2D spectra of an ideal S²H based beampattern. A narrowband source with frequency 2 kHz was taken in the far-field region at $(45°, -90°)$. 20 microphones are

(a) r = 1cm   (b) r = 5cm   (c) r = 10cm

Fig. 2. Sectoral mode strength variation with radius $r = 1cm$ in (a), $r = 5cm$ in (b), and $r = 10cm$ in (c). The open spherical sector is considered with elevation $\theta \in [0°, 90°]$, and azimuth $\phi \in [0°, 360°]$.

considered to be distributed uniformly on the surface of open spherical sector. The sector is defined by $r = 2cm, \theta \in [0°, 90°]$, and $\phi \in [0°, 360°]$. High resolution directivity pattern is observed in the array steered direction with negligible side-lobes.

3.2. Maximizing the Directivity and WNG

The maximum directivity (MD) beamforming design expects the beampattern as sharp as possible in the direction of arrival and zero elsewhere. The MD condition is mathematically obtained by equating the expression in (23) with (20), given by

$$\sum_{n=0}^{N}\sum_{m=-n}^{n}\tilde{b}_n(kr)\sum_{i=1}^{I}W_i^*(k)T_n^m(\theta_i,\phi_i) = \qquad (25)$$

$$|q_1||u|\sum_{n=0}^{N}\sum_{m=-n}^{n}T_n^m(\theta_s,\phi_s)$$

For (25) to hold true, we have

$$\sum_{i=1}^{I}W_i^*(k)T_n^m(\theta_i,\phi_i) = \frac{|q_1||u|}{\tilde{b}_n(kr)}T_n^m(\theta_s,\phi_s) \qquad (26)$$

For order N and degree m, (26) can be rewritten in matrix form as

$$\widetilde{\mathbf{T}}\mathbf{w}^\star(k) = \boldsymbol{\eta}(k) \qquad (27)$$

where $\widetilde{\mathbf{T}}$ is a matrix of dimension $(N+1)^2 \times I$, given by

$$\begin{bmatrix} T_0^0(\theta_1,\phi_1) & T_0^0(\theta_2,\phi_2) & T_0^0(\theta_3,\phi_3) & \dots & T_0^0(\theta_I,\phi_I) \\ T_1^{-1}(\theta_1,\phi_1) & T_1^{-1}(\theta_2,\phi_2) & T_1^{-1}(\theta_3,\phi_3) & \dots & T_1^{-1}(\theta_I,\phi_I) \\ T_1^0(\theta_1,\phi_1) & T_1^0(\theta_2,\phi_2) & T_1^0(\theta_3,\phi_3) & \dots & T_1^0(\theta_I,\phi_I) \\ T_1^1(\theta_1,\phi_1) & T_1^1(\theta_2,\phi_2) & T_1^1(\theta_3,\phi_3) & \dots & T_1^1(\theta_I,\phi_I) \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ T_N^N(\theta_1,\phi_1) & T_N^N(\theta_2,\phi_2) & T_N^N(\theta_3,\phi_3) & \dots & T_N^N(\theta_I,\phi_I) \end{bmatrix}$$

and $\boldsymbol{\eta}(k)$ represents a vector of length $(N+1)^2$ expressed as

$$\boldsymbol{\eta}(k) = \left[\frac{\zeta T_0^0(\theta_s,\phi_s)}{b_0(k)}, \frac{\zeta T_1^{-1}(\theta_s,\phi_s)}{b_1(k)}, \right. \\ \left. \frac{\zeta T_1^0(\theta_s,\phi_s)}{b_1(k)}, \dots, \frac{\zeta T_N^N(\theta_s,\phi_s)}{b_N(k)} \right]^T \qquad (28)$$

with $\zeta = |q_1||u|$. To make the beamformer robust against sensor noise, white noise gain (WNG) is maximized by $min\mathbf{w}^H(k)\mathbf{w}(k)$ [14].

Utilizing the concept of sharp directivity pattern and maximum WNG, the maximum directivity-WNG (MD-WNG) beamformer can now be designed by solving the optimization problem given by

$$min_{\mathbf{w}(k)}\mathbf{w}^H(k)\mathbf{w}(k) \text{ s.t } \widetilde{\mathbf{T}}\mathbf{w}^\star(k) = \boldsymbol{\eta}(k) \qquad (29)$$

The weight vector obtained from the solution of (29) can be written as

$$\mathbf{w}(k) = \widetilde{\mathbf{T}}^T\left(\widetilde{\mathbf{T}}^\star\widetilde{\mathbf{T}}^T\right)^{-1}\boldsymbol{\eta}^\star(k) \qquad (30)$$

Utilizing (14), the weight vector can now be formulated in spherical sector harmonics domain as

$$\mathbf{w}_{nm}(k) = \mathbf{T}^H\boldsymbol{\Gamma}\left[\widetilde{\mathbf{T}}^T\left(\widetilde{\mathbf{T}}^\star\widetilde{\mathbf{T}}^T\right)^{-1}\boldsymbol{\eta}^\star(k)\right] \qquad (31)$$

Eqs. (31) and (28) suggest that the beamforming weight $\mathbf{w}_{nm}(k)$ includes a division by $\tilde{b}_n(kr)$. It may also be noted from Fig. 2 that $\tilde{b}_n(kr)$ assumes the value 0 more frequently with increase in range r. Hence, the weights cannot be determined for these values of frequency. To overcome this a compact array, limited frequency or a rigid S^2MA may be utilized.

By substituting the weight vector given by (31) in (19), the MD-WNG beamformer output in S^2H domain is given by

$$\mathbf{Z}_{nm}(k) = \left[\mathbf{T}^H\boldsymbol{\Gamma}\left(\widetilde{\mathbf{T}}^T\left(\widetilde{\mathbf{T}}^\star\widetilde{\mathbf{T}}^T\right)^{-1}\boldsymbol{\eta}^\star(k)\right)\right]^H\mathbf{P}_{nm}(k) \qquad (32)$$

The MD-WNG beampattern in S^2H domain can be written as

$$B[\mathbf{w}_{nm}(k),\theta_l,\phi_l] = \left[\mathbf{T}^H\boldsymbol{\Gamma}\left(\widetilde{\mathbf{T}}^T\left(\widetilde{\mathbf{T}}^\star\widetilde{\mathbf{T}}^T\right)^{-1}\boldsymbol{\eta}^\star(k)\right)\right]^H\mathbf{a}_{nm}(k) \qquad (33)$$

Fig. 4 compares the beampatterns of the S^2H based ideal beamformer with the MD-WNG beamformer. A narrowband source of frequency 4 kHz was taken in the far-field region at (30°, 0°). The experimental conditions remain same as detailed in the previous section. It can be seen that the main-lobe of the MD-WNG beampattern almost coincides with the ideal beampattern, but the MD-WNG beamformer has a comparably larger side-lobe. This is mainly due to the presence of two design objectives (sharp directivity pattern along with maximum WNG) in MD-WNG beamformer.

4. Performance Evaluation

In this section, numerous simulation results associated with the proposed S^2H based beamforming are presented. A narrowband source with frequency 2 kHz is taken in the far-field region at (45°, −90°). 20 microphones are considered to be distributed uniformly on the surface of open spherical sector. The sector is defined by $r = 2cm, \theta \in [0°, 90°]$, and $\phi \in [0°, 360°]$ under the free-field assumption. Various performance metrics are presented to evaluate the proposed method.

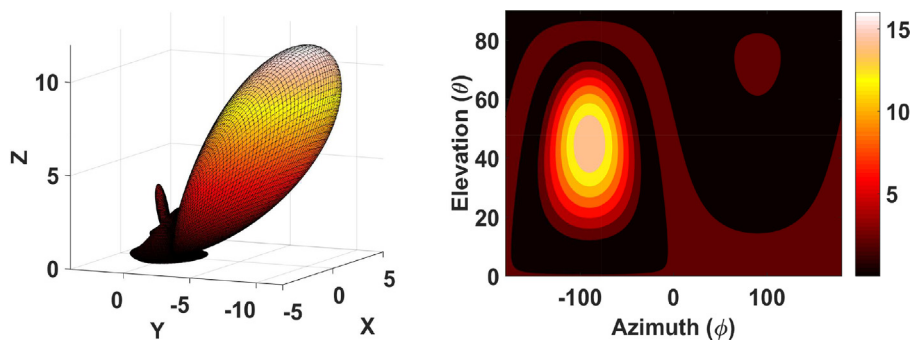**Fig. 3.** 3D and 2D spectra of an ideal beampattern using S^2MA. The array of order $N = 3$ is steered at (45°, −90°).
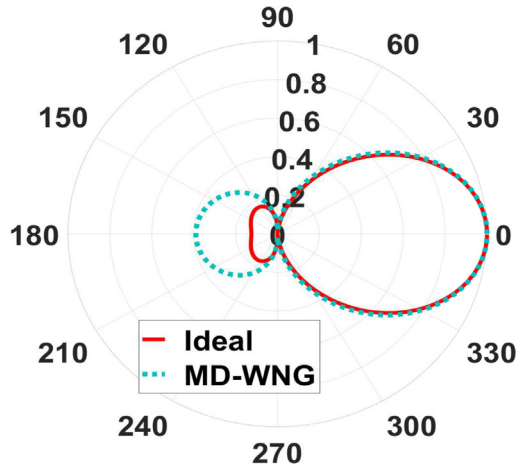
Fig. 4. Ideal (red solid) and MD-WNG (blue dashed) beampatterns using S²MA of order $N = 3$ and $r = 2cm$.

4.1. Beampattern/Directivity pattern analysis

The proposed beamformer performance is presented herein using the directivity pattern. Particularly, the effect of increasing order N on an ideal and MD-WNG beampattern is explored. Additionally, the proposed S²H beampattern is also compared with the existing acoustic image principle (AIP) based AIP-SH beampattern [16].

4.1.1. Effect of array order on ideal S²H beampattern

Fig. 5 illustrates the effect of increasing order N on the ideal beampattern. The simulation was performed by considering the signal with frequency 2 kHz. The array steering direction (θ_s, ϕ_s) is set to $(45°, -90°)$. It can be seen that the beampattern has higher resolution with increase in the order of an array. This is due to increase in the number of spherical sector harmonics coefficients with increasing N in (23). An array with high order N results in better approximation of beampattern [14].

4.1.2. Effect of array order on S²H based MD-WNG beampattern

Fig. 6 illustrates the effect of array order N on MD-WNG beampattern. Particularly, 3D, 2D and polar plots are shown in details. A signal with 6 kHz frequency is taken with steering direction at $(30°, 280°)$. As expected, the beampattern has higher resolution with increase in the order of the array. This is due to the increase in the number of S²H coefficients in (33).

4.1.3. Beampattern comparison with AIP-SH beamformer

The AIP and S²H based beampatterns are illustrated in Fig. 7. An extra competing lobe may be seen in Fig. 7(a), resulting from inherent assumption of imaginary source. The additional image source

results in two correlated sources which degrades the beamformer performance. However, spherical sector harmonics based approach utilizes S²H basis for the beamforming. This results in providing the narrower main lobe only. Hence, S²H basis function gives more accurate representation of beampattern over spherical sector than usual spherical harmonics.

4.1.4. Beampattern comparison with spherical fraction harmonics based beamformer

The spherical fraction harmonics (SFHs) and S²H based beampatterns are illustrated in Fig. 8 for $N = 3$. The array steering direction is set to $(45°, 45°)$. The SFHs based beampatterns in Fig. 8(a) has comparably larger side-lobe heights, resulting from inherent assumption of imaginary source. The superposition and interference of the beampatterns due to real and imaginary source influence the resultant beampattern shape [17]. However, spherical sector harmonics based approach provides the main lobe with lower side-lobes because of absence of any imaginary source. It may also be noted that the position of main lobe is shifted in SFHs based approach. This may be due to the presence of rigid boundaries condition in the SFHs based approach when compared to free-field boundaries in S²H based beampattern.

4.2. White noise gain and directivity factor

To evaluate the sensitivity of the proposed beamformer design towards sensor and diffused noise sources, the parameters like WNG and directivity factor (DF) are utilized in this section. The simulations are done considering the steering direction $(\theta_s, \phi_s) = (30°, 280°)$. As WNG and DF utilize array gain and SNR gains hence a brief discussion about each measure is presented first. The input signal to noise ratio (iSNR) with reference to the first microphone is given by

$$iSNR(k) = \frac{\phi_s(k)}{\phi_{v_1}(k)}, \tag{34}$$

where $\phi_s(k) = E\left[|S(k)|^2\right]$ is the signal variance and $\phi_{v_1}(k) = E\left[|V_1(k)|^2\right]$ is the noise variance at the first microphone. $E[x]$ denotes the mathematical expectation of x and $V_1(k)$ is the first element of $\mathbf{V}(k)$. Similarly, the output signal to noise ratio (oSNR) of the beamformer can be written as

$$oSNR[\mathbf{w}(k)] = \frac{\phi_s(k)|\mathbf{w}^H(k)\mathbf{a}(k)|^2}{\mathbf{w}^H(k)\boldsymbol{\Phi}_v(k)\mathbf{w}(k)},$$
$$= \frac{\phi_s(k)}{\phi_{v_1}(k)} \times \frac{|\mathbf{w}^H(k)\mathbf{a}(k)|^2}{\mathbf{w}^H(k)\boldsymbol{\Gamma}_v(k)\mathbf{w}(k)}, \tag{35}$$

where $\boldsymbol{\Gamma}_v(k)$ is the pseudo-coherence matrix of $\mathbf{V}(k)$ given by $\boldsymbol{\Gamma}_v(k) = \frac{\boldsymbol{\Phi}_v(k)}{\phi_{v_1}(k)}$ with $\boldsymbol{\Phi}_v(k) = E\left[\mathbf{V}(k)\mathbf{V}^H(k)\right]$. Using (34) and (35) the array gain is obtained as [8]

(a) N=1

(b) N=2

(c) N=3
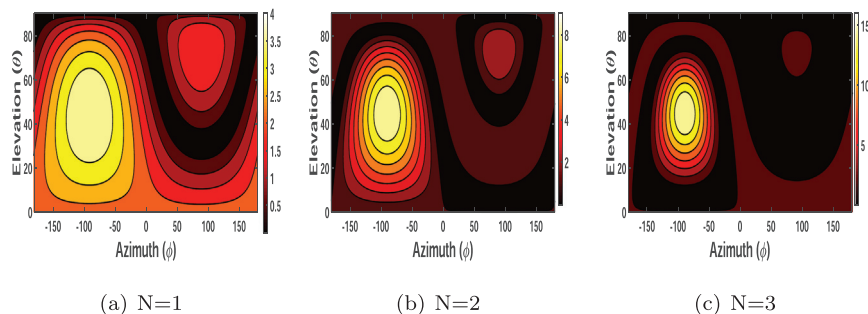
Fig. 5. Variation of S²H based ideal beampattern with array order N. The steering direction (θ_s, ϕ_s) is set to $(45°, -90°)$.

(a) N=1

(b) N=1
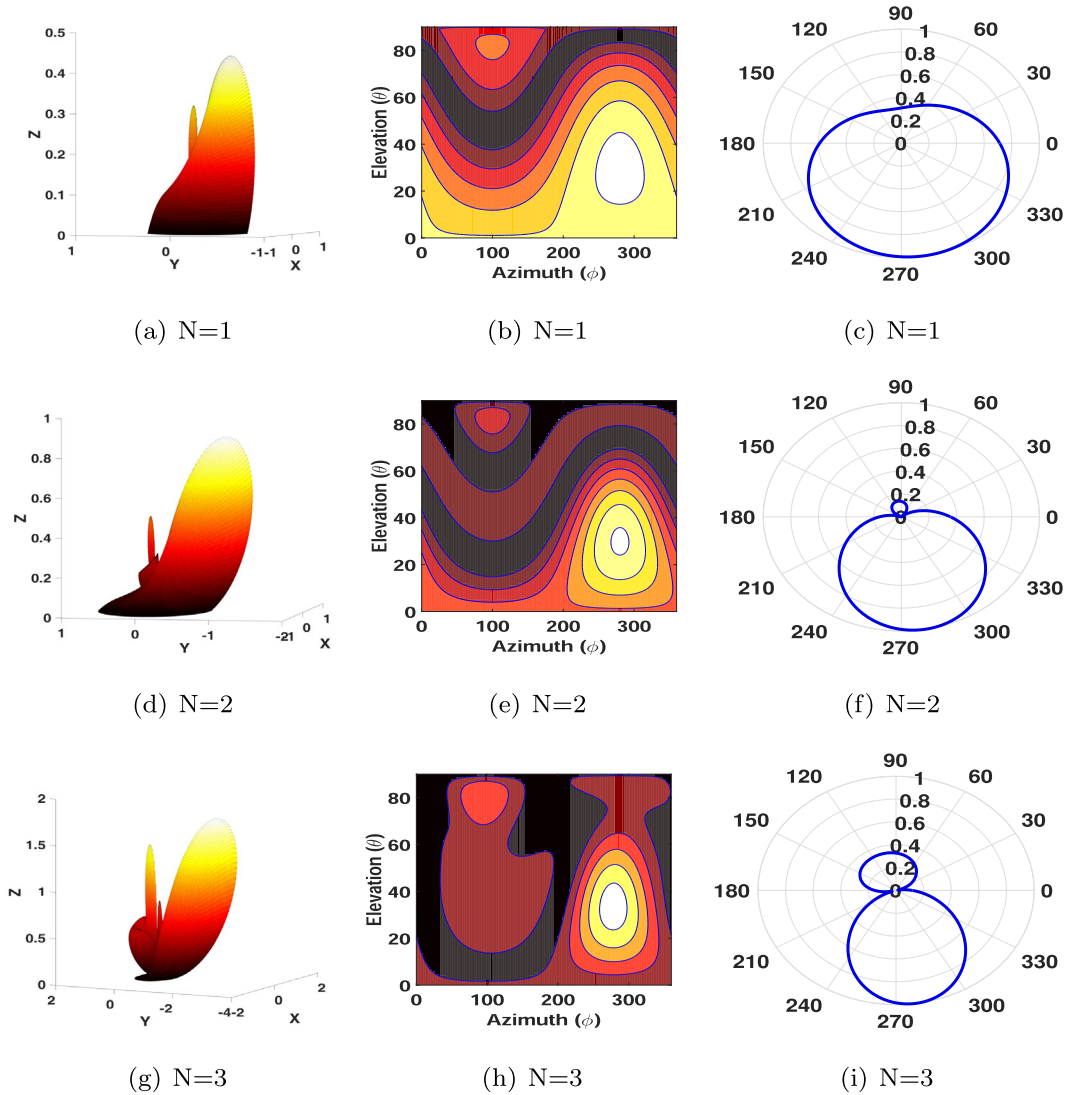
(c) N=1

(d) N=2

(e) N=2

(f) N=2

(g) N=3

(h) N=3

(i) N=3

Fig. 6. S²H based MD-WNG Beampattern variation with order N. The array is steered in the direction $(30°, 280°)$. 3D spectra in (a), (d) and (g), 2D spectra in (b), (e) and (h), and polar directivity plot in (c), (f) and (i).

$$\mathscr{G}[\mathbf{w}(k)] = \frac{oSNR[\mathbf{w}(k)]}{iSNR(k)}, = \frac{|\mathbf{w}^H(k)\mathbf{a}(k)|^2}{\mathbf{w}^H(k)\mathbf{\Gamma}_v(k)\mathbf{w}(k)}. \tag{36}$$

The beamformer is evaluated in the presence of different noisy conditions. In general, white noise and diffuse noise sources are considered. Based on the type of noise sources, the array gain is termed as white noise gain (WNG) and directivity factor (DF) as detailed in the following section. A white noise gain is used to measure the sensitivity of the beamformer to it's imperfections such as sensor noise, electronic noise and mismatch among the different sensors of microphone array. As the noise is temporally and spatially white in nature and assumes to have same variance at all microphones, the pseudo-coherence matrix is replaced by the identity matrix of dimension $I \times I$ i.e $\mathbf{\Gamma}_v(k) = \mathbf{I}_I$. Substituting in (36), WNG can be expressed as

$$\mathscr{W}[\mathbf{w}(k)] = \frac{|\mathbf{w}^H(k)\mathbf{a}(k)|^2}{\mathbf{w}^H(k)\mathbf{w}(k)}. \tag{37}$$

WNG in spherical sector harmonics domain can be written as

$$\mathscr{W}[\mathbf{w}_{nm}(k)] = \frac{|\mathbf{w}_{nm}^H(k)\mathbf{a}_{nm}(k)|^2}{\mathbf{w}_{nm}^H(k)\mathbf{w}_{nm}(k)}. \tag{38}$$

Directivity factor is formulated by assuming noise sources to be uniformly distributed over the spherical surface. Here, the noise fields are assumed to be spherically isotropic or 3D diffused in nature. The pseudo-coherence matrix of $\mathbf{V}(\mathbf{k})$, for spherically isotropic noise is given by [25]

$$\begin{aligned}[\mathbf{\Gamma}_v(k)]_{ij} &= [\mathbf{\Gamma}_d(k)]_{ij}, \\ &= sinc(kd_{ij}), \end{aligned} \tag{39}$$
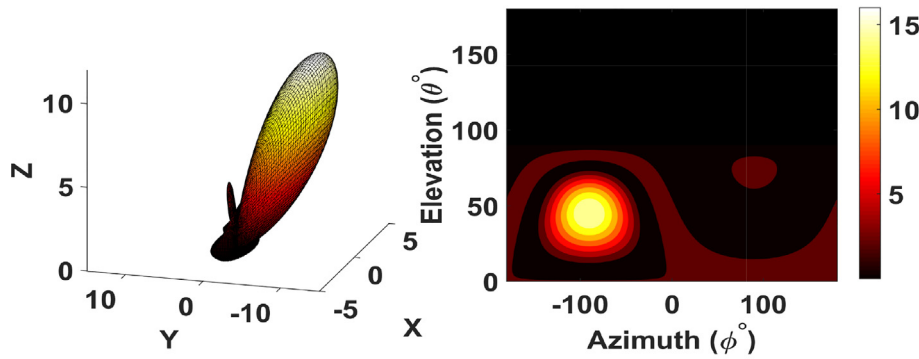
where $sinc(kd_{ij}) = \frac{\sin(kd_{ij})}{kd_{ij}}$ with $i, j = 1, 2, 3, \ldots, I$ and d_{ij} is the distance between the ith and jth microphones. Utilizing (39) in (36), the array gain becomes the directivity factor and can be expressed as

$$\mathscr{D}[\mathbf{w}(k)] = \frac{|\mathbf{w}^H(k)\mathbf{a}(k)|^2}{\mathbf{w}^H(k)\mathbf{\Gamma}_d(k)\mathbf{w}(k)}. \tag{40}$$

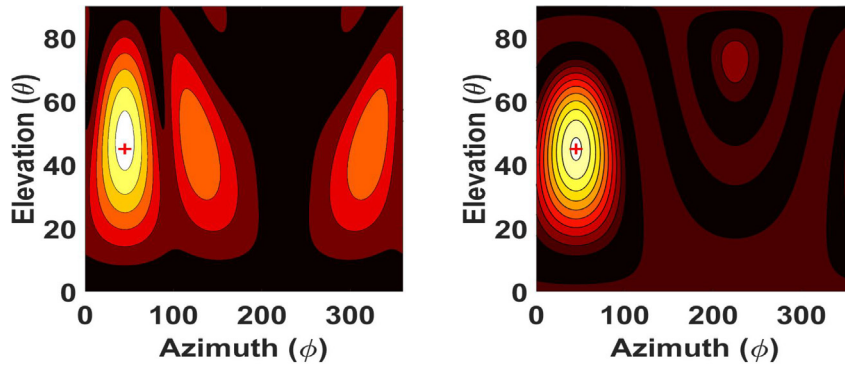Hence, the DF in spherical sector harmonics domain can be written as
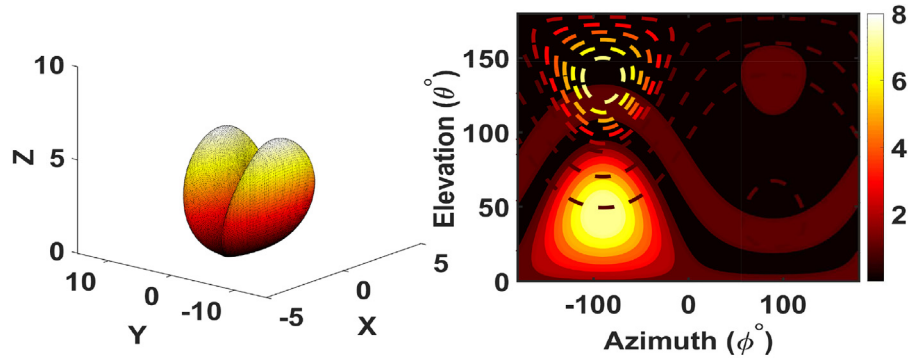
(a)

(b)

Fig. 7. 3D and 2D spectra of an ideal beampattern using AIP-SH in (a) and S^{2}H in (b). The array of order $N = 3$ is steered in the direction $(45°, -90°)$.

(a)

(b)

Fig. 8. 2D spectra of an ideal beampattern using SFHs in (a) and S^{2}H in (b). The array of order $N = 3$ is steered in the direction $(45°, 45°)$.

$$\mathscr{D}[\mathbf{w}_{nm}(k)] = \frac{|\mathbf{w}_{nm}^H(k)\mathbf{a}_{nm}(k)|^2}{\mathbf{w}_{nm}^H(k)\mathbf{\Gamma}_{nm}(k)\mathbf{w}_{nm}(k)}. \tag{41}$$

The WNG and DF are utilized in the ensuring section for evaluating the proposed beamformer.

4.2.1. Effect of deep nulls on WNG and DF

Fig. 9 plots the WNG and DF of the proposed S^{2}H MD-WNG beamformer. The WNG and DF are shown as a function of frequency with varying radius r. The array order is fixed at $N = 3$. It can be noted that at $r = 1cm$, the WNG and DF has no deep nulls. With the increase of r, the deep nulls get introduced and increase

as a function of r. Hence, for an open spherical sector microphone array, restriction on r is necessary to avoid the deep null problems [15].

4.2.2. Comparison of the proposed S^{2}H beamformer with the AIP-SH based beamformer

Fig. 10 illustrates the comparison of the proposed S^{2}H MD-WNG beamformer with the AIP-SH based beamformer. The S^{2}H based beamformer outperforms the AIP-SH based beamformer in terms of WNG and DF. The low resolution of AIP-SH based method is mainly due to the presence of correlated imaginary source.
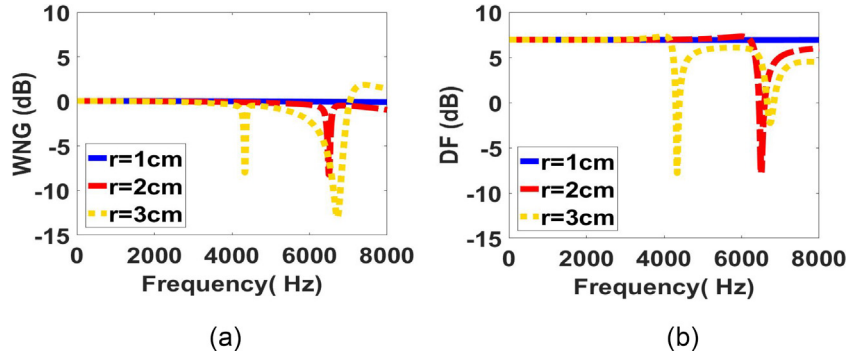
Fig. 9. (a) WNG and (b) DF of the MD-WNG beamformer for S²MA as a function of the frequency with $I = 20$ and $N = 3$, for three different radius r.
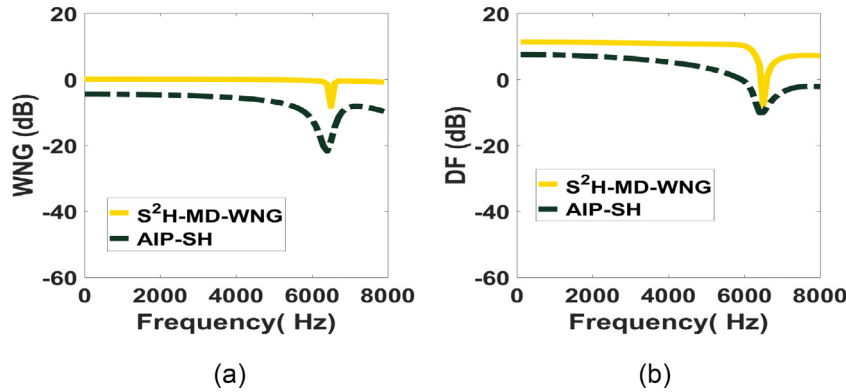

Fig. 10. (a) WNG and (b) DF comparison of the MD-WNG beamformer for S²MA and AIP-SH based beamformer for HMA as a function of the frequency with $I = 20$ and $r = 2cm$.

4.2.3. Comparison of the proposed S²H-MD-WNG beamformer with three different beamformers for S²MA

In this subsection, an additional comparison is made between MD-WNG beamformer with the three different beamformers for S²MA. In particular, the WNG and DF of the proposed beamformer is compared with minimum variance distortionless response (MVDR) beamformer, delay-and-sum (DS) beamformer and maximum directivity (MD) beamformer, all computed in S²H domain. The plots are shown in Fig. 11. For a given $N = 3$, the DS beamformer obtains high WNG but poor directivity factor. This makes the DS beamformer a poor speech enhancement method. On the other hand, the WNG of the proposed MD-WNG beamformer is higher than the MVDR and MD beamformers but less than DS

beamformer. The DF of the MD-WNG beamformer is higher than the MVDR, DS, and MD beamformers. Hence, the proposed beamformer is relatively more robust in presence of white noise as well as spherically isotropic noise.

4.3. Computational Complexity Analysis

The computational complexity of the proposed method for beamforming is presented in this subsection. A system with Intel® i7-6700 processor, RAM 16 GB, instruction set 64-bit and clock speed 3.4 GHz is utilized for the evaluation.

The computation time for S²H based ideal beampattern is compared with AIP based counterpart in Table 1. The array is steered at
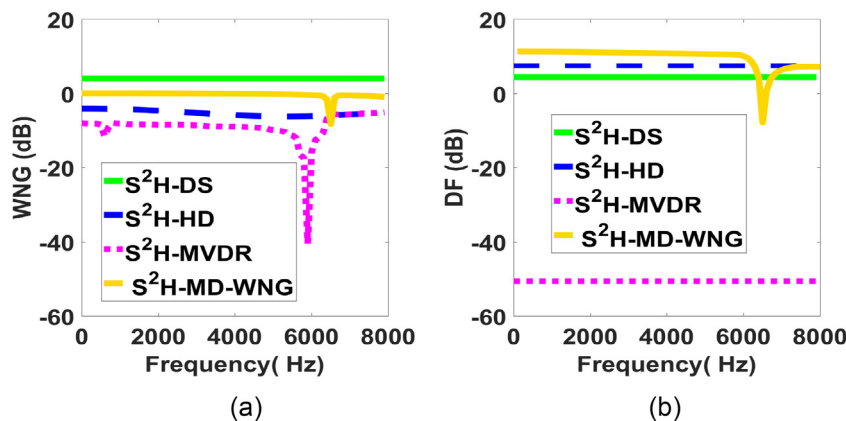

Fig. 11. (a) WNG and (b) DF comparison of the different beamformers for S²MA as a function of the frequency with $N = 3$ and $r = 2cm$.

Table 1

Computation time comparison for beampattern synthesis, using S^2H and AIP based approach.

Functions	Computational time (sec)
S^2H based Ideal Beampattern	16.290120
AIP-SH based Beampattern	48.058153

Table 2

Computational time for MD-WNG beampattern in spatial and S^2H domains.

Functions	Computational time (sec)
Spatial domain	28.7340
S^2H domain	22.6013

$(45°, 90°)$ for array order $N = 3$. It is to note that S^2H based ideal beampattern is computationally more efficient than the AIP based method. Hence, the beampattern synthesis is performed efficiently using spherical sector harmonics.

The computation time for the MD-WNG beampattern in the spatial and S^2H domains is presented in Table 2. The array is steered at $(30°, 280°)$ for array order $N = 3$. It may be noted that S^2H based beampattern exhibits greater computational efficiency when compared to the spatial counterpart. Hence, the beampattern synthesis is more efficient using S^2H.

5. Conclusion

This paper describes an optimal design of beamformers for spherical sector microphone arrays. An ideal directivity pattern with maximum response in the desired direction and zero elsewhere is detailed. To make the beamformer robust against sensor noise along with a maximum directivity pattern in the desired direction, MD-WNG beamformer is designed in S^2H domain. The designed beamformer weights are estimated by solving the constrained optimization problem. The performance evaluation of the proposed S^2H MD-WNG beamformers is presented. A spherical sector array prototype is currently being developed.

CRediT authorship contribution statement

Deepika Kumari: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing. **Lalan Kumar:** Visualization, Investigation, Supervision, Resources, Funding acquisition, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Chiariotti P, Martarelli M, Castellini P. Acoustic beamforming for noise source localization–reviews, methodology and applications. Mech Syst Signal Process 2019;120:422–48.

[2] Haykin SS, Liu KR. Handbook on array processing and sensor networks. Wiley Online Library; 2009.

[3] Van Trees HL. Optimum array processing: Part IV of detection, estimation, and modulation theory. John Wiley & Sons; 2004.

[4] de Araujo FH, Pinto FAdNC, Torres JCB, Room reflections analysis with the use of spherical beamforming and wavelets, Applied Acoustics 131 (2018) 192–202.

[5] Wang L, Zhu J. Regularized beamformer for the spherical microphone array to cope with the white noise amplification. In: ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2020. p. 4657–61.

[6] Salvati D, Drioli C, Foresti GL. Diagonal unloading beamforming in the spherical harmonic domain for acoustic source localization in reverberant environments. IEEE/ACM Trans Audio, Speech, Language Processing 2020.

[7] Ping G, Chu Z, Yang Y. Compressive spherical beamforming for acoustic source identification. Acta Acustica united with Acustica 2019;105(6):1000–14.

[8] Huang G, Chen J, Benesty J. A flexible high directivity beamformer with spherical microphone arrays. J Acoust Soc Am 2018;143(5):3024–35.

[9] Chu Z, Yang Y, Yang Y. A new insight and improvement on deconvolution beamforming in spherical harmonics domain. Appl Acoust 2021;177:107900.

[10] Cohen I, Benesty J, Gannot S. Speech processing in modern communication: Challenges and perspectives, Vol. 3. Springer Science & Business Media; 2009.

[11] Meyer J, Elko G. A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 2. IEEE; 2002. pp. II–1781.

[12] Rafaely B. Phase-mode versus delay-and-sum spherical microphone array processing. IEEE signal processing Letters 2005;12(10):713–6.

[13] Li Z, Duraiswami R. Flexible and optimal design of spherical microphone arrays for beamforming. IEEE Trans Audio, Speech, Language Process 2007;15(2):702–14.

[14] Rafaely B. Fundamentals of spherical array processing, Vol. 8. Springer; 2015.

[15] Wang L, Zhu J. Flexible beampattern design algorithm for spherical microphone arrays. IEEE Access 2019;7:139488–98.

[16] Li Z, Duraiswami R. Hemispherical microphone arrays for sound capture and beamforming. In: Applications of Signal Processing to Audio and Acoustics, 2005. IEEE Workshop on. IEEE; 2005. p. 106–9.

[17] Lecomte P, Melon M, Simon L. Spherical fraction beamforming. IEEE/ACM Trans Audio, Speech, Language Processing 2020;28:2996–3009.

[18] Cabral B, Max N, Springmeyer R, Bidirectional reflection functions from surface bump maps, in: ACM Siggraph Computer Graphics, Vol. 21, ACM, 1987, pp. 273–281.

[19] Gautron P, Krivanek J, Pattanaik SN, Bouatouch K. A novel hemispherical basis for accurate and efficient rendering. Rendering Techniques 2004;2004:321–30.

[20] Giri A, Kumar L, Gandhi T, Eeg dipole source localization in hemispherical harmonics domain, in: 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), IEEE, 2018, pp. 679–684.

[21] Kumari D, Kumar L. Acoustic source localization in hemispherical harmonics domain. In: Proc. EuroNoise. p. 2575–80.

[22] Kumari D, Kumar L. Spherical sector harmonics representation of sound fields using a microphone array over spherical sector. J Acoust Soc Am 2021;149(1):145–57.

[23] Kumari D, Kumar L. S^2H domain processing for acoustic source localization and beamforming using microphone array on spherical sector. IEEE Trans Signal Process 2021;69:1983–94.

[24] Meyer J, Elko G, A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield, in: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 2, 2002, pp. II–1781–II–1784.

[25] Habets EA, Gannot S. Generating sensor signals in isotropic noise fields. J Acoust Soc Am 2007;122(6):3464–70.