Flexible Beampattern Design Algorithm for Spherical Microphone Arrays

LEI WANG, JIE ZHU
Department of Electronic Engineering, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China
Corresponding author: Jie Zhu (e-mail: zhujie@sjtu.edu.cn).

This work was supported by the National Key Research and Discovery Plan No. 2017YFF0210903 and the National Natural Science Foundation of China (grant 11433002).

ABSTRACT
Specific beampatterns of spherical microphone arrays (SMAs) are designed for different noise scenarios. The standard beampattern usually achieves optimum value in one of the performance indicators. Additionally, in acoustics scenarios where positions of several noise sources are known, the beampattern for SMA should be designed to reject signals from noise orientations. This paper presents a flexible beamformer design method for SMA. The beampattern expression is transformed from spherical harmonics expansion to the common polynomial expression, which facilitates the implementation of the standard beampattern and the novel pattern where the zero positions are assigned to reject multiple noise sources. Simulation results reveal that the proposed beamformer shows better flexibility over the high directivity beamformer and Dolph-Chebychev beamformer. The deep-null problem existing in the open-sphere SMA is also discussed and can be mitigated by reducing the radius or increasing the order of the beamformer.

INDEX TERMS
spherical microphone arrays, beampattern design, frequency invariance, white noise gain, directivity factor, deep nulls

I. INTRODUCTION

By exploiting the spatial diversity information in the sound field, the microphone arrays (MAs) are capable of enhancing the speech signal contaminated by noise, reverberation and interference [1]–[3]. According to the spacing among elements, MAs can be divided into compact MAs and distributed MAs. Considering the limited consumption of size, weight and power in many applications, MAs with compact apertures have drawn lots of attention recently, where the far-field acoustic model hypothesis holds and facilitates the follow-up algorithm [4]–[6].

The performance of the MA is affected by various factors like signal spectral range, acoustic surrounding, applied algorithm, element quantity and geometry. Among these, the geometry is thought to play a significant part in the algorithm formulation [3], [7], [8]. With simple acoustic modeling and estimation calculation, the linear MA (LMA) is proposed and many algorithms are developed [9], [10]. Since the beampattern of the LMA is dependent on the steering direction, the poor steering flexibility becomes a big drawback of the LMA. The circular MA (CMA), also with a regular geometry, stands out due to its flexible steering characteristic. The look direction of the CMA can be steered to arbitrary azimuth in the sensor plane with the constant array response [11], [12]. Compared with MAs of the two above geometries, the spherical MA (SMA) can capture the three-dimensional (3D) sound field and becomes an attracting research topic.

The beamforming algorithms and beampattern design methods have been widely investigated for multi-channel speech enhancement by using SMA [3], [13]–[16]. The delay-and-sum beamformer, realized by aligning signals from each channel with time of arrival or phase shifts, stands out due to its simplicity and robustness [14]. Besides, the extensive ranges of optimal beamforming algorithms are also studied for SMA, like minimum-variance distortionless response (MVDR), generalized side-lobe canceler (GSC) and linearly constrained minimum variance (LCMV). Some of them show a better performance in the spherical harmonics domain than the space domain [3]. But the frequency response of these beamformers is inconstant, which prevents them from the frequency-invariant beampattern design.

In practical noise scenarios, specific beampatterns are desired to capture the target speech and reject noise sources. Standard beampatterns are the most interesting theoretical beampatterns because they usually achieve optimum value in one of the performance indicators, e.g., hypercardioid beam-
pattern is the best in terms of DF and supercardioid beampattern has highest front-to-back ratio [2]. In other cases, the shape of the beampattern need be designed according to the direction of the target sound source and the noises. In beampattern design, Amitai Koretz et al. proposed the Dolph-Chebychev beampattern method to control the main-lobe width and maximum the side-lobe level directly by using the similarity between the Legendre polynomials and Chebyshev polynomials [15], where the width of every side lobe is fixed. Gongping Huang et al. developed a high directivity beamformer, which can only design the hypercardioid pattern of any order for SMA [16]. Hence, a more general design algorithm is desired to obtain arbitrary beampatterns.

In this paper, we present a beampattern design method for the SMA beamformer. Arbitrary beampatterns can be implemented easily, including standard beampatterns such as cardioid, hypercardioid, supercardioid and novel patterns where the zeros can be assigned to eliminate multiple noise sources. To obtain the desired beampattern, the processing algorithm is formulated in the spherical harmonics domain, where the coefficients of the beamformer filter are computed from a constrained optimization problem. A regularized beamforming filter is also proposed to improve the white noise amplification problem. The simulation experiments prove the proposed design algorithm obtains the frequency-invariant beampattern. Compared to the high directivity beamformer and Dolph-Chebychev beamformer, the proposed beamformer is capable of designing arbitrary beampattern and eliminating noise sources efficiently. Moreover, the array aperture restriction and the beamformer order increase are proved to mitigate the deep-null problem. We also prove the white noise amplification aggravated by the two methods can be alleviated by using the proposed regularized beamforming filter.

II. SIGNAL MODEL AND PROBLEM FORMULATION

A far-field source signal propagates as a plane wave in an anechoic environment at the speed of sound, i.e., $c = 340$ m/s, and impinges on the SMA. The SMA is composed of $M$ omnidirectional microphones located on a sphere with a radius of $r$. The center of the SMA is assumed to be consistent with the origin of the three-dimensional Cartesian coordinate system. The azimuth angles are measured anticlockwise from the x-axis and the elevation angles are measured downward from the z-axis. In this case, the location of the $m$th $(m = 1, 2, ..., M)$ microphone in the SMA is expressed as $r \times r_m$, where

\[ r_m = [\sin \theta_m \cos \psi_m \quad \sin \theta_m \sin \psi_m \quad \cos \theta_m]^T, \tag{1} \]

with the superscript $T$ being the transpose operator, $\theta_m$ being the elevation angle and $\psi_m$ being the azimuth angle of the $m$th microphone.

Steering the beam to the direction $(\theta, \psi)$, we represent the steering vector corresponding to the plane wave from the far-field sound source of length $M$ as

\[ d(\omega, \theta, \psi) = [e^{j\omega r \cos \theta_1} \quad e^{j\omega r \cos \theta_2} \quad \cdots \quad e^{j\omega r \cos \theta_M}]^T, \tag{2} \]

where $\theta \in [0, \pi]$, $\psi \in [0, 2\pi)$, $j$ is the imaginary unit with $j^2 = -1$,

\[ \omega = \frac{\omega r}{c} \tag{3} \]

with $\omega = 2\pi f$ being the angular frequency and $f$ being the temporal frequency and $p = [\sin \theta \cos \psi \quad \sin \theta \sin \psi \quad \cos \theta]^T$.

Considering that the target speech signal comes from the direction $(\theta_s, \psi_s)$, the corresponding propagation vector is $d(\omega, \theta_s, \psi_s)$. The signal received by microphones is defined in the frequency domain as

\[ x(\omega) = \begin{bmatrix} X_1(\omega) & X_2(\omega) & \cdots & X_M(\omega) \end{bmatrix}^T = d(\omega, \theta_s, \psi_s)S(\omega) + v(\omega), \tag{4} \]

where $X_m(\omega)$ is the received signal at $m$th microphone, $S(\omega)$ is the desired source signal with mean 0 and $v(\omega) = [V_1(\omega) \quad V_2(\omega) \quad \cdots \quad V_M(\omega)]^T$ is the noise vector with $V_m(\omega)$ being the additive noise signal at $m$th microphone in the frequency domain.

To recover the desired signal from the noisy observation vector, a complex weight, $H_m^*(\omega)$, is applied to the received signal of the $m$th sensor, where the superscript * means complex conjugation. Then the weighted outputs are summed up to obtain the beamformer’s output, that is,

\[ Z(\omega) = \sum_{m=1}^{M} H_m^*(\omega)X_m(\omega) \]

\[ = h^H(\omega)x(\omega) = h^H(\omega)d(\omega, \theta_s, \psi_s)S(\omega) + h^H(\omega)v(\omega), \tag{5} \]

where

\[ h(\omega) = [H_1(\omega) \quad H_2(\omega) \quad \cdots \quad H_M(\omega)]^T \tag{6} \]

is the spatial filter of length $M$ with $H_m(\omega)$ being the weight function of the $m$th microphone.
To make the source signal pass through the beamformer without distortion, the distortionless constraint in the desired direction is assumed to be true in most cases, i.e.,

$$h^H(\omega)d(\omega, \theta_s, \psi_s) = 1. \quad (7)$$

### III. PERFORMANCE MEASURES

Generally, several metrics are used to evaluate a beamformer. In this paper we introduce three of them: the beampattern (directivity pattern), the white noise gain (WNG) and the directivity factor (DF) [17].

The beampattern describes the sensitivity of the beamformer to a plane wave impinging on the SMA from the direction ($\theta, \psi$), which is defined as

$$B[h(\omega), \theta, \psi] = h^H(\omega)d(\omega, \theta, \psi)$$

$$= \sum_{m=1}^{M} H_m^*(\omega)e^{j\varphi m} r_m. \quad (8)$$

The signal-to-noise ratio (SNR) gain is an important criterion for beamformer performance. If the first microphone in SMA is chosen to be the reference, the input SNR is defined as

$$\text{iSNR}(\omega) = \frac{\phi_s(\omega)}{\phi_v(\omega)}, \quad (9)$$

where $\phi_s(\omega) = E|S(\omega)|^2$ and $\phi_v(\omega) = E|V_1(\omega)|^2$ are the variances of $S(\omega)$ and $V_1(\omega)$ respectively with $E(\cdot)$ being the mathematical expectation. According to (5), the output SNR can be derived as

$$\text{oSNR}[h(\omega)] = \frac{\phi_s(\omega)h^H(\omega)d(\omega, \theta_s, \psi_s)^2}{\phi_v(\omega)} = \frac{\phi_s(\omega)h^H(\omega)d(\omega, \theta_s, \psi_s)^2}{\phi_v(\omega)h^H(\omega)\Phi(\omega)h(\omega)}, \quad (10)$$

where $\Phi(\omega) = E[V(\omega)v^H(\omega)]$ and $\Phi(\omega)/\phi_v(\omega)$ are the correlation and pseudo-coherence matrices of $v(\omega)$ respectively. SNR gain is defined as oSNR to iSNR ratio. From (9) and (10), we obtain

$$G[h(\omega)] = \frac{\text{oSNR}[h(\omega)]}{\text{iSNR}(\omega)} = \frac{h^H(\omega)d(\omega, \theta_s, \psi_s)^2}{h^H(\omega)\Phi(\omega)h(\omega)}. \quad (11)$$

Generally, two kinds of noise are considered to evaluate the performance of beamformers: the white noise and the diffuse noise.

The temporally and spatially white noise with the same variance across all the microphones models the electronic and sensor noise as well as the mismatch among different microphones in a MA system. In this case, $\Phi(\omega) = \mathbf{I}_M$, where $\mathbf{I}_M$ is the $M \times M$ identity matrix. Hence (11) becomes

$$W[h(\omega)] = \frac{h^H(\omega)d(\omega, \theta_s, \psi_s)^2}{h^H(\omega)h(\omega)}, \quad (12)$$

which is called the white noise gain (WNG).

Another kind of noise is the diffuse noise, which corresponds to the spherically isotropic noise field. The $(i, j)$th element of the noise pseudo-coherence matrix $\Gamma_{\psi}(\omega)$ becomes

$$[\Gamma_{\psi}(\omega)]_{ij} = [\Gamma_{d,\psi}(\omega)]_{ij} = \text{sinc} \left(\frac{\omega d_{ij}}{c}\right), \quad (13)$$

where $i, j = 1, 2, ..., M$ and

$$\delta_{ij} = ||r_i - r_j||_2 \quad (14)$$

is the Euclidean distance between microphones $i$ and $j$ with $r_i, r_j$ being the coordinates of microphones, as defined in (1). Then the SNR gain in (11) becomes

$$D[h(\omega)] = \frac{|h^H(\omega)d(\omega, \theta_s, \psi_s)|^2}{h^H(\omega)\Gamma_d(\omega)h(\omega)}, \quad (15)$$

which is called the directivity factor (DF).

### IV. BEAMFORMING FOR SMA

Generally, in an open space, the unit-amplitude plane wave that comes from the direction $(\theta, \psi)$ can be decomposed in the spherical harmonic domain [3]

$$e^{j\varphi m} r_m = \sum_{n=0}^{\infty} \sum_{l=-n}^{n} b_n(\varpi) Y_n^l(\theta_m, \psi_m) Y_n^l(\theta, \psi)^* \quad (16)$$

where $b_n(\varpi) = 4\pi j^n J_n(\varpi)$ (17)

with $J_n(\varpi)$ being the $n$th-order spherical Bessel function of the first kind.

$$Y_n^l(\theta, \psi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-l)!}{(n+l)!}} P_n^l(\cos \theta) e^{jl\psi} \quad (18)$$

denotes the spherical harmonic of order $n$ ($n = 0, 1, 2, ..., +\infty$) and degree $l$ ($l = 0, \pm 1, \pm 2, ..., \pm n$), where $(\cdot)!$ denotes the factorial function and $P_n^l(\cos \theta)$ is the associated Legendre function of order $n$ and degree $l$ [18].

Substituting (16) into (8), we can derive the beampattern as

$$B[h(\omega), \theta, \psi] = \sum_{m=1}^{M} H_m^*(\omega) \sum_{l=-n}^{n} \sum_{l=-n}^{n} b_n(\varpi) Y_n^l(\theta_m, \psi_m) Y_n^l(\theta, \psi)^*$$

$$= \sum_{m=1}^{M} H_m^*(\omega) Y_n^l(\theta_m, \psi_m). \quad (19)$$

By limiting the expansion to the order $N$, the following equation is obtained.

$$B_N[h(\omega), \theta, \psi] = \sum_{n=0}^{N} b_n(\varpi) \sum_{l=-n}^{n} Y_n^l(\theta_m, \psi_m) Y_n^l(\theta, \psi)^*$$

$$\times \sum_{m=1}^{M} H_m^*(\omega) Y_n^l(\theta_m, \psi_m). \quad (20)$$
For small \( \varpi \), the values of high-order spherical harmonics are negligible and \( \varpi \ll N \) is also satisfied, so (20) truncates (19) to a finite order \( N \) with little truncation error and aliasing error. Whereas for large \( \varpi \), i.e., for high frequency and big array aperture, the effect of truncation and spatial aliasing cannot be ignored [19], [20].

In the plane-wave amplitude density domain, the axisymmetric beampattern can be formulated in terms of series expansion [3]:

\[
B[d_n, \theta, \psi] = \sum_{n=0}^{\infty} d_n \sum_{l=-n}^{n} Y_n^l(\theta, \psi_s)[Y_n^l(\theta, \psi)]^*, \tag{21}
\]

where \( d_n \) is a coefficient only depending on \( n \). If we limit (21) to the order of \( N \), the beampattern can be expressed as

\[
B_N[d_n, \theta, \psi] = \sum_{n=0}^{N} d_n \sum_{l=-n}^{n} Y_n^l(\theta, \psi_s)[Y_n^l(\theta, \psi)]^*. \tag{22}
\]

Equating beampatterns in (20) and (22), we get the following equation,

\[
b_n(\varpi) \sum_{m=1}^{M} H_n^*(\omega) Y_n^l(\theta, \psi_m) = d_n Y_n^l(\theta, \psi_s). \tag{23}
\]

A more compact expression of (23) is,

\[
\mathbf{Y} \mathbf{h}^*(\omega) = \mathbf{\zeta}(\varpi), \tag{24}
\]

where

\[
\mathbf{Y} = \begin{bmatrix}
Y_0^0(\theta, \psi_1) & Y_0^0(\theta, \psi_2) & \cdots & Y_0^0(\theta, \psi_M) \\
Y_1^{-1}(\theta, \psi_1) & Y_1^{-1}(\theta, \psi_2) & \cdots & Y_1^{-1}(\theta, \psi_M) \\
Y_1^1(\theta, \psi_1) & Y_1^1(\theta, \psi_2) & \cdots & Y_1^1(\theta, \psi_M) \\
Y_2^{-2}(\theta, \psi_1) & Y_2^{-2}(\theta, \psi_2) & \cdots & Y_2^{-2}(\theta, \psi_M) \\
\vdots & \vdots & \ddots & \vdots \\
Y_N^{-l}(\theta, \psi_1) & Y_N^{-l}(\theta, \psi_2) & \cdots & Y_N^{-l}(\theta, \psi_M)
\end{bmatrix}
\]

(25)

is a matrix of size \((N + 1)^2 \times M, M \geq (N + 1)^2\) and

\[
\mathbf{\zeta}(\varpi) = \begin{bmatrix}
d_0 Y_0^0(\theta, \psi_1) / b_0(\varpi) \\
d_1 Y_1^{-1}(\theta, \psi_1) / b_1(\varpi) \\
\vdots \\
d_N Y_N^{-l}(\theta, \psi_1) / b_N(\varpi)
\end{bmatrix}^T \tag{26}
\]

is a vector of length \((N + 1)^2\). Minimizing \( \mathbf{h}^H(\omega) \mathbf{h}(\omega) \) subject to the constraint in (24), an optimization problem is constructed [2]:

\[
\min_{\mathbf{h}(\omega)} \mathbf{h}^H(\omega) \mathbf{h}(\omega) \quad \text{s.t.} \quad \mathbf{Y} \mathbf{h}^*(\omega) = \mathbf{\zeta}(\varpi). \tag{27}
\]

The minimum-norm solution of (27) is the \( N \)th-order beamforming filter \( \mathbf{h}_N(\omega) \):

\[
\mathbf{h}_N(\omega) = \mathbf{Y}^T(\mathbf{Y} \mathbf{Y}^T)^{-1}\mathbf{\zeta}^*(\varpi). \tag{28}
\]

Then the array output is given in the space domain as,

\[
\mathbf{z}_N(\omega) = \mathbf{h}_N^H(\omega) \mathbf{x}(\omega). \tag{29}
\]

Note that the filter coefficient \( \mathbf{h}_N(\omega) \) in (28) can be considered as the weighted sum of the elements in \( \mathbf{\zeta}^*(\varpi) \), whose denominator is \( b_N(\varpi) \) (see (26)). \( b_n \) shares the same zeros with the spherical Bessel function \( J_n(\varpi) \), as indicated in (17). Obviously the zero of \( b_n \) lead to an infinite value of the filter coefficient \( \mathbf{h}_N(\omega) \), which further cause significant SNR gain (WNG and DF defined in (12) and (15) respectively) degradation at some frequency bands [3], [19], [21]. The zero locations in \( b_n(\varpi) \) are decided by the value of \( \varpi \) \( (\varpi = \varpi r/c) \), which is an increasing function of \( r \). Figure 2 displays the curves of \( b_n \) versus frequency \( f \) at \( 0\sim8kHz \). As shown in Fig. 2(a), the first zero \( \varpi_0 \) of \( b_n \) in the positive domain belongs to \( b_0 \). Therefore, to avoid deep nulls, we need \( \varpi < \varpi_0 \) holds at desired frequency band, that is,

\[
r < \varpi r/c. \tag{30}
\]

It can be seen that when \( r=5cm \), i.e., (30) is not always true, there are several zeros in observed frequency range. However, when \( r=2cm \), i.e., (30) is true in the range of \( 0\sim8kHz \) (0Hz is excluded), no zero exists in \( b_0(\varpi) \).

From Fig. 2 (a), we can also notice that all curves have a zero at 0Hz except the 0th-order term. At low frequency, the value of \( b_n \) approaches zero except \( b_0 \) and higher order \( n \) corresponds to smaller \( b_n \), which causes white noise amplification, especially in the cases of high-order beamformer and small aperture array. For example, (30) works well for 1st-order beamformer but it will suffer from severe white noise amplification for higher orders, which provides a low-frequency limit for the beamformer design. The beamforming filter in (28) can be regularized to alleviate the white noise amplification and improve the robustness of the SMA beamformer. By rearranging (23), the matrix (25) and vector (26) can be rewritten as

\[
\mathbf{Y}_r = \begin{bmatrix}
b_0(\varpi) \\
b_1(\varpi) \\
\vdots \\
b_N(\varpi)
\end{bmatrix}^T \begin{bmatrix}
y_0^0(\theta, \psi_s) \\
y_0^{-1}(\theta, \psi_s) \\
\vdots \\
y_N^{-l}(\theta, \psi_s)
\end{bmatrix} \tag{31}
\]

and

\[
\mathbf{\zeta}_r = \begin{bmatrix}
d_0 Y_0^0(\theta, \psi_s) \\
d_1 Y_1^{-1}(\theta, \psi_s) \\
\vdots \\
d_N Y_N^{-l}(\theta, \psi_s)
\end{bmatrix}, \tag{32}
\]
which make (24) and (27) hold as well. Then the regularized beamforming filter is expressed as

\[ \mathbf{h}_{N,\epsilon}(\omega) = \mathbf{Y}^T \mathbf{Y} + \epsilon \mathbf{I}_{(N+1)^2} \mathbf{\zeta}_{\epsilon}^{-1}(\omega), \] (33)

where \( \epsilon \geq 0 \) is the regularization parameter and \( \mathbf{I}_{(N+1)^2} \) is the identity matrix of size \((N + 1)^2 \times (N + 1)^2\). (33) is equivalent to (28) when \( \epsilon \) equals to zero.

Apart from restricting the radius of SMA to overcome deep-null problem, some other solutions are also studied. In [22], microphones are positioned on concentric spheres of different radii, which avoids calculating the inversion of the spherical Bessel function by increasing the physical complexity. The rigid-sphere SMA also provides a robust numerical condition to avoid the deep-null problem, but the accuracy of the sound field measurement is reduced as the sound scattering of a rigid sphere will interferes with the surrounding sound field [23]. Although the effect is significant for large radius and becomes negligible for small radius [19], the accuracy of the sound field measurement is more important in this paper to ensure the frequency-invariance and the accuracy of the beampattern design. Hence the open-sphere configuration is preferable in this paper. The elements of open-sphere SMA can be cardioid microphones or omnidirectional ones. Using cardioid microphones helps avoid the deep-null problem but is sensitive to the directional misalignment [23]. In this paper, the open-sphere SMA with omnidirectional elements is considered.

V. BEAMFORMER DESIGN ALGORITHM

A. FLEXIBLE BEAMFORMER DESIGN

Generally, desired beampatterns are various due to multiple possibilities in terms of the positions of target speech and noises. To obtain the optimal beampattern in different acoustic scenarios, the value of corresponding \( d_n \) needs to be known. However, the exact \( d_n \) is hard to be derived directly from (21) because of its complicated expression as a nested series expansion. According to the spherical harmonics addition theorem [3], (21) can be rewritten as

\[ B_N(d_n, \theta, \psi) = \sum_{n=0}^{N} d_n \cdot \frac{2n + 1}{4\pi} \mathbf{P}_n(\cos \Theta), \] (34)

where \( \mathbf{P}_n(\cdot) \) is the Legendre polynomial of degree \( n, \Theta \) is the angular difference between vectors \((\theta, \psi)\) and \((\theta_s, \psi_s)\), and \( \cos \Theta = \cos \theta \cos \theta_s + \sin \theta \sin \theta_s \cos (\psi - \psi_s) \). Substituting \( \beta_n = \frac{2n+1}{4\pi} d_n \) and \( x = \cos \Theta \in [\pm 1, 1]\) into (34), we obtain

\[ B_N(\beta_n, x) = \sum_{n=0}^{N} \beta_n \mathbf{P}_n(x) = \beta^T \mathbf{P}(\beta), \] (35)

where \( \beta = [\beta_0 \beta_1 ... \beta_N]^T \) is a vector consisting of constant coefficients and \( \mathbf{P}(x) = [\mathbf{P}_0(x) \mathbf{P}_1(x) ... \mathbf{P}_N(x)]^T \). By this means, the beampattern is expressed as a linear combination of the Legendre polynomial \( \mathbf{P}_n(x) \), which is one kind of orthogonal polynomials [24].

In the two-dimensional planar plane, any symmetric frequency-invariant beampattern can be expressed in form of a simple series expansion [25]:

\[ B_N(\alpha, x) = \sum_{n=0}^{N} \alpha_n x^n = \alpha^T \mathbf{x}_N, \] (36)

where \( \alpha = [\alpha_0 \alpha_1 ... \alpha_N]^T \), \( \mathbf{x}_N = [1 \ x ... x^N]^T \) with \( x = \cos \Theta \in [\pm 1, 1] \).

To extend it to the 3D space, the coefficient \( \beta \) in (35) should be expressed as a function of \( \alpha \), because \( \beta \) is involved in the beampattern design of SMA. Appendix B in reference [26] shows the derivation of the relationship between \( \beta \) and the Jacobi polynomials’ coefficients \( \beta_j \), which is a more general expression of \( \beta \). Firstly the \( n \)-th order Jacobi polynomial can be expressed as a simple series expansion,

\[ P_{n;i,j}^{(i,j)}(x) = \sum_{i=0}^{n} \xi_{n,i} x^i = \xi^T x_n, \] (37)

where \( \xi_n = [\xi_{n,0} \xi_{n,1} ... \xi_{n,n}]^T \) and \( \mathbf{x}_n = [1 \ x ... x^n]^T \) are vectors of length \( n + 1 \). Then the relationship between the coefficients \( \alpha \) and \( \beta_j \) is deduced as

\[ \beta_j = \Xi^{-1} \alpha, \] (38)

where the transfer matrix between \( \beta_j \) and \( \alpha \) is,

\[ \Xi = \begin{bmatrix} \xi_{0,0} & \xi_{1,0} & \xi_{2,0} & \cdots & \xi_{0,N} \\ 0 & \xi_{1,1} & \xi_{2,1} & \cdots & \xi_{1,N} \\ 0 & 0 & \xi_{2,2} & \cdots & \xi_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \xi_{N,N} \end{bmatrix}. \] (39)

Coefficients \( \xi_n \) can be recursively computed,

\[ \xi_{n+1} = \mathbf{R}^T \xi_n - \mathbf{R} \xi_{n-1}, n \geq 1, \] (40)

where \( \xi_0 = 1, \xi_1 = [(i - j)/2 \ (i + j + 2)/2]^T \), \( \mathbf{R}^+ \) and \( \mathbf{R}^- \) are recurrence coefficients.

In (35), \( \beta \) is coefficients of the Legendre polynomial \( \mathbf{P}_n(x) \), which is a special case of the Jacobi polynomial. In this case, \( i = 0, j = 0 \) and recurrence coefficients \( \mathbf{R}^+ = \frac{n+1}{n+2} \mathbf{I}_{(n+1) \times (n+1)} \), \( \mathbf{R}^- = -\frac{n}{n+1} \mathbf{I}_{n \times n} \)

Substituting them into (40), the coefficient for the Legendre polynomial is represented as,

\[ \xi_{n+1} = \frac{n+1}{n+2} \begin{bmatrix} \mathbf{I}_{(n+1) \times (n+1)} \\ \mathbf{I}_{(n+1) \times (n+1)} \end{bmatrix} \xi_n - \frac{n}{n+1} \begin{bmatrix} \mathbf{I}_{n \times n} \\ \mathbf{I}_{0 \times n} \end{bmatrix} \xi_{n-1} \] (41)

with \( \xi_0 = 1, \xi_1 = [0 \ 1]^T \), where \( n \geq 1 \), \( \mathbf{0}_{k \times l} \) is the zero matrix of size \( k \times l \) and \( \mathbf{I}_{k \times l} \) is the identity matrix of size \( k \times l \). \( \xi_n \) is a special case of \( \xi_n \), with the same length of \( n + 1 \). Substituting \( \xi_n \) into (39), the transfer matrix between \( \beta \) and \( \alpha \) is obtained as,

\[ \Xi = \begin{bmatrix} \xi_0 & \xi_1 & \xi_2 & \cdots & \xi_N \\ \mathbf{0}_{0 \times n} & \mathbf{0}_{(N-1) \times 1} & \mathbf{0}_{(N-2) \times 1} & \cdots & \mathbf{0}_{0 \times 1} \end{bmatrix}. \] (42)

which is an upper-triangular matrix of size \((N + 1) \times (N + 1)\).

Obviously \( \Xi \) is a special form of \( \Xi \) in the case of the Legendre polynomial. Hence we obtain the relationship between \( \beta \) and \( \alpha \),

\[ \beta = \Xi^{-1} \alpha. \] (43)
B. BEAMPATTERN DESIGN IN TWO APPLICATIONS

The value of α decides the beam shape. Here we introduce two kinds of practical beampatterns and how to derive α respectively:

- The standard beampatterns are the most widely used beampatterns as they usually achieve optimum value in one of the performance indicators. For instance, hypercardioid beampattern is the best in terms of DF and supercardioid beampattern has the highest front-to-back ratio [2]. The value of the coefficient α corresponding to multiple-order standard directivity patterns is calculated and listed in [9].

- Novel beampatterns for mitigating multiple noise sources by placing nulls. In the case of multiple noise sources, the magnitude of beamformer directional response is expected to be as low as possible in the directions of noises, i.e., zeros in beampattern should be located in these directions. We will present how to assign known zeros in noise directions to the directivity pattern in the following.

As (36) is a N-order polynomial, a N-order directivity pattern contains N zeros at most. Assume that N noise sources distribute in the 3D space and the direction of the nth noise is expressed as (θ0,ψ0), where n = 1, 2, ..., N, θ0 is the elevation angle and ψ0 is the azimuth angle of the nth noise source. Given that the target speech signal comes from the direction (θs,ψs), the angle difference between the nth noise source and the target speech source is derived as,

\[ μ_n = \arccos[\sinθ_0\sinθ_s\cos(ψ_s−ψ_0) + \cosθ_0\cosθ_s], \]

where \( n = 1, 2, ..., N \). Since the magnitude of the beampattern in the noise source directions should be zero, \( θ_2 ∈ \{ μ_1, μ_2, ..., μ_N \} \) makes the value of the beampattern expression (36) be zero. Obviously \( \{ \cos μ_1, \cos μ_2, ..., \cos μ_N \} \) are the roots of the polynomial in (36), hence the beampattern can be reconstructed as,

\[ B_N^\Theta(μ, x) = \prod_{n=1}^{N} (x − \cos μ_n), \]

where \( μ = [μ_1, μ_2, ..., μ_N]^T \) and \( x = \cos Θ (x ∈ [-1, 1]) \) with Θ being the angular difference between vectors (θ, ψ) and (θs, ψs). The value of α in (36) is obtained by unfolding (45). Then the SMA filter can be calculated through (43) and (28). After passing through the spatial filters, the known noises are eliminated.

VI. SIMULATIONS

In this section, we evaluate and present the performance of the proposed design algorithm for SMA beamformers. In our simulation, the open-sphere SMA is composed of 32 omnidirectional microphones distributing along equally spaced azimuth (8 steps) and elevation (4 steps), which performs the Gaussian sampling scheme [19]. The steering direction (θs, ψs) is set as (90°, 120°), which, of course, could be any direction in the 3D space.
FIGURE 3: Designed beampatterns (blue solid) and desired beampatterns (red dashed) of 1st-, 2nd-, 3rd-order cardioid, supercardioid and hypercardioid, where $M = 32$, $r = 2$ cm and $f = 1$ kHz.

locations, i.e., all the noise sources are depressed. As the frequency increases, the SNR gain rises gradually, which means the attenuation in noise directions gets weakened. There are two possible reasons: Firstly, the higher frequency diminishes the attenuation in zeros, as mentioned before. Secondly, as the frequency increases, the effect of aliasing error becomes more significant, which causes the slight change in beam shape, as shown in Fig. 5. As a result, zero positions minorly shift and deviate from noise direction.

B. PERFORMANCE COMPARISON
We compare the following beamformers to illuminate the improvement of the proposed design algorithm for SMA:

- Proposed flexible beamformer designing 3rd-order hypercardioid beampattern, which maximizes DF theoretically;
- The 3rd-order high directivity (HD) beamformer proposed in [16];
- The beamformer derived from the 3rd-order Dolph-Chebyshev (Chev) beampattern with the first zero located at $51.43^\circ$, which coincides with the position of the first zero in 3rd-order hypercardioid beampattern. But other zeros in the Chev beampattern cannot be assigned to certain positions [15];
- Traditional delay-and-sum (DS) beamformer.

Note that mentioned beampatterns are all axisymmetrical. The radius of the SMA is 2 cm, i.e., the deep-null problem does not exist in above beamformers. Figure 7 displays the WNG and DF comparison among the 4 beamformers at the frequency range from 0 kHz to 8 kHz. As shown in the two graphs, DS beamformer achieves high WNG, but the poor DF values in $0 \sim 3$ kHz prevent it from being a good speech enhancement method. The proposed beamformer obtains slightly higher WNG and DF than Chev beamformer as the beam shape of Chev beamformer is not the one achieving the highest directivity. The HD beamformer is almost coincident with the proposed beamformer in the two indicators, which indicates HD beamformer can be considered as a special case of the proposed beamformer. Overall, compared to the Chev beamformer, the proposed algorithm performs better at assigning zeros to the noise directions more accurately. Also, the proposed beamformer is more flexible than HD beamformer since all kinds of standard patterns as well as patterns with specific zeros can be designed.

C. DISCUSSION ABOUT THE DEEP-NULL PROBLEM
To evaluate the influence of the deep-null problem in the open-sphere SMA as mentioned in Sec.IV, the WNG and
FIGURE 4: Designed beampatterns (blue solid) and desired beampatterns (red dashed) of 3rd-order cardioid, supercardioid and hypercardioid at 1kHz, 3kHz, 5kHz and 7kHz, where $M = 32$, $r = 2\text{ cm}$.

FIGURE 5: Designed beampatterns (blue solid) and desired beampatterns (red dashed) in 2-noise and 3-noise scenarios at 1kHz, 3kHz, 5kHz and 7kHz, where $M = 32$, $r = 2\text{ cm}$.
FIGURE 6: Beampattern view in condition of known noise sources illustrated by (a) 2D plot in 2-noise scenario, (b) 2D plot in 3-noise scenario, (c) 3D plot in 2-noise scenario and (d) 3D plot in 3-noise scenario, where $M = 32$, $r = 2cm$.

TABLE 2: The SNR gain at noise directions in the designed beampattern with the target speech coming from $(90^{\circ}, 120^{\circ})$.

<table>
<thead>
<tr>
<th>Noise Location</th>
<th>$1kHz$</th>
<th>$3kHz$</th>
<th>$5kHz$</th>
<th>$7kHz$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Noises</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(270^{\circ}, 60^{\circ})$</td>
<td>-87.4</td>
<td>-68.8</td>
<td>-60.2</td>
<td>-48.8</td>
</tr>
<tr>
<td>$(30^{\circ}, 120^{\circ})$</td>
<td>-86.6</td>
<td>-66.6</td>
<td>-55.2</td>
<td>-42.7</td>
</tr>
<tr>
<td>3 Noises</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(30^{\circ}, 110^{\circ})$</td>
<td>-72.6</td>
<td>-53.3</td>
<td>-44.1</td>
<td>-38</td>
</tr>
<tr>
<td>$(130^{\circ}, 30^{\circ})$</td>
<td>-72.6</td>
<td>-53.6</td>
<td>-44.8</td>
<td>-38.8</td>
</tr>
<tr>
<td>$(300^{\circ}, 120^{\circ})$</td>
<td>-74.2</td>
<td>-55.1</td>
<td>-46.4</td>
<td>-42.7</td>
</tr>
</tbody>
</table>

DF performance in different configurations are displayed in Fig. 8 and Fig. 9. Figure 8 depicts the WNG and DF of the 1st-order hypercardioid with 3 radii in the frequency band over $0 \sim 8kHz$. Figure 9 depicts the WNG and DF of the 1st-, 2nd- and 3rd-order hypercardioid in the frequency band of $0 \sim 8kHz$. The radius of SMA is 3cm. According to Fig. 8, as the value of $r$ increases, the number of deep nulls in WNG and DF also increases because the range of $\varpi$ in the spherical Bessel function $J_n(\varpi)$ includes more zeros. When $r = 2cm$, (30) holds well at the frequency band of (0kHz, 8kHz], hence no deep null exists in WNG and DF. Figure 9 shows the deep null becomes less obvious as the order increases, meanwhile, the WNG deteriorates at low frequency.

To sum up, in open-sphere SMA beamformer design, restricting the radius to the limitation in (30) will avoid the deep-null problem in the SNR gains and increasing the order of the beampattern will mitigate the deep-null problem but worsen the WNG performance. Whereas it can be observed that smaller array aperture and higher order lead to more serious white noise amplification. To solve the problem, the regularized beamformer filter in (33) can be used to improve the WNG at low frequency. Figure 10 plots SNR gains at different values of the regularization parameter $\epsilon$. The SNR gains in the case without regularization by using the filter in (28) are also plotted as a comparison. When $\epsilon = 10^{-12}$, the figures are coincident with those without regularization, i.e.,
the regularization has no significant effect. As $\epsilon$ increases, WNG at low frequency is improved but DF descends, which means the beampattern shape becomes less satisfactory at low frequency. Overall, regularization is denoted to mitigate white noise amplification by sacrificing the DF indicator.

At the absence of the deep-null problem, as shown in Fig.4, the designed patterns perform well in terms of frequency-invariance. In addition, the concentric SMA is capable of deal with the deep-null problem without restricting the array aperture, which could be studied in our future research.

VII. CONCLUSION

This paper proposes an algorithm to design arbitrary beampatterns flexibly for the SMA beamformer in the spherical harmonics domain. The simulation results illustrate the proposed beamformer is nearly frequency-invariant at the absence of the deep-null problem and it achieves better flexibility compared to Chev and HD beamformers. By transforming the desired directivity pattern from the spherical harmonics expansion to the common polynomial expression, arbitrary beampatterns are implemented, including standard beampatterns achieving the optimum performance indicators or novel patterns derived from specified zero positions to reject noise source directions. The designed pattern is obtained by calculating the beamformer filter coefficients derived from a constrained optimization problem. The deep-null problem caused by the zeros of the spherical Bessel function is proved to be mitigated by restricting the array radius and increasing the beamformer order, which introduce the white noise amplification. To solve this problem, the beamforming filter can be regularized to compromise between white noise amplification and the directivity factor. Overall, the proposed beampattern design algorithm can apply SMA in various noise scenarios and has good prospect for applications and popularization.

VIII. REFERENCES

REFERENCES


***