Flexible and Optimal Design of Spherical Microphone Arrays for Beamforming

Zhiyun Li, *Member, IEEE*, and Ramani Duraiswami, *Member, IEEE*

***Abstract*—This paper describes a methodology for designing a flexible and optimal spherical microphone array for beamforming. Using the approach presented, a spherical microphone array can have very flexible layouts of microphones on the spherical surface, yet optimally approximate a desired beampattern of higher order within a specified robustness constraint. Depending on the specified beampattern order, our approach automatically achieves optimal performances in two cases: when the specified beampattern order is reachable within the robustness constraint we achieve a beamformer with optimal approximation of the desired beampattern; otherwise we achieve a beamformer with maximum directivity, both robustly. For efficient implementation, we also developed an adaptive algorithm for computing the beamformer weights. It converges to the optimal performance quickly while exactly satisfying the specified frequency response and robustness constraint in each step. One application of the method is to allow the building of a real-world system, where microphones may not be placeable on regions, such as near cable outlets and/or a mounting base, while having a minimal effect on the performance. Simulation results are presented.**

***Index Terms*—Beamforming, beampattern, directivity index (DI), optimization, quadrature, spherical microphone array, white noise gain (WNG).**

List of Symbols

c	Speed of sound. $c \approx 343$ m/s in air.
∇^2	Laplacian operator.
f	Frequency of wave.
$\boldsymbol{\theta} = (\theta, \varphi)$	Angular position.
$\boldsymbol{\theta}_0$	Look direction.
$\boldsymbol{\theta}_k$	Wave incident direction.
$\boldsymbol{\theta}_s$	Observation direction.
$\boldsymbol{r}_s = (\boldsymbol{\theta}_s, r_s)$	Observation point.
Ω_s	Surface of a unit sphere.
$p(\boldsymbol{r}, t)$	Pressure at the spatial location $\boldsymbol{r} = (\boldsymbol{\theta}, r)$ and the time t.
$\boldsymbol{k}$	Plane wave incident from the direction $\boldsymbol{\theta}_k$ with the wavenumber $k = \|\boldsymbol{k}\| = 2\pi f/c$.
$\psi(\boldsymbol{r}, \boldsymbol{k})$	Complex pressure in frequency domain at the location $\boldsymbol{r}$ for plane wave $\boldsymbol{k}$.
a	Radius of the spherical microphone array.
$j_n(x)$	Spherical Bessel function of order n. $j_n'(x) = \frac{d}{dx} j_n(x)$.
$h_n(x)$	nth order spherical Hankel function of the first kind. $h_n'(x) = \frac{d}{dx} h_n(x)$.
$Y_n^m(\boldsymbol{\theta})$	Spherical harmonics of order n and degree m.
$*$	Complex conjugation.
δ_{xy}	Kronecker delta.
$\delta(x)$	Delta function.
$C_{n'}^{m'}(\boldsymbol{\theta}_s)$	Quadrature coefficient for $Y_{n'}^{m'}$ at $\boldsymbol{\theta}_s$.
N_{eff}	Band limit of spatial frequency in terms of spherical harmonics orders.
N	Order of beamformer.
$N_{\max}$	Maximum order of a robust beamformer.
F_N	Regular beampattern of order N.
WNG	White noise gain.
DI	Directivity index.
$H(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$	Actual beampattern looking at $\boldsymbol{\theta}_0$.
$\boldsymbol{d}$	Vector of complex pressure at each microphone position produced by the plane wave of unit magnitude from the desired beamforming direction $\boldsymbol{\theta}_0$.
$\boldsymbol{W}$:	Vector of complex weights for each microphone.
$\psi_n^m(\boldsymbol{\theta}_s, \boldsymbol{\theta}_k, ka)$	One component of ψ at order n and degree m.
$\Pi_n^m(\boldsymbol{\theta}_k, ka)$	Decomposition result of $\psi_n^m(\boldsymbol{\theta}_s, \boldsymbol{\theta}_k, ka)$.
$\epsilon_{nn'}^{mm'}$	Othonormality errors caused by discreteness.
$B_{n'}^n(ka)$	Frequency-dependent scale factor in solving quadrature of orthonormalities.
$\boldsymbol{A}$	Coefficient matrix of soundfield expansion.
$\boldsymbol{B}_N$	Coefficient vector of beampattern.
c_N	Normalizing coefficient to satisfy the specified frequency response.
δ^2	Specified minimum WNG.
λ	Parameter in Tikhonov regularization.
μ	Step size in adaptive implementation.

Manuscript received June 7, 2005; revised November 30, 2005. This work was supported in part by National Science Foundation Award 0205271. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Futoshi Asano.

Z. Li was with the Perceptual Interfaces and Reality Lab, UMIACS, University of Maryland, College Park, MD 20742 USA. He is now with Leica, San Ramon, CA 94583 USA (e-mail: zli@cs.umd.edu).

Ramani Duraiswami is with the Perceptual Interfaces and Reality Lab, UMIACS, University of Maryland, College Park, MD 20742 USA (e-mail: ramani@umiacs.umd.edu).

Digital Object Identifier 10.1109/TASL.2006.876764

I. Introduction

SPHERICAL arrays of microphones are recently becoming a subject of interest as they allow three dimensional sampling of the soundfield, and may have applications in soundfield capture [14]. The paper [15] presented a first analysis of such arrays, and showed how sound can be analyzed using them.

This paper performed an elegant separation of the analysis and beamforming parts by using a modal beamformer structure. One implicit aspect of the analysis is that the distribution of microphones on the surface of the sphere seems to be redundant considering the results achieved. This is because the beamforming relies on numerical integration ("quadrature") of spherical harmonics. A quadrature scheme for spherical harmonics usually includes carefully chosen quadrature points on spherical surface and a set of optimal quadrature weights. In [15], this is done using a specified semi-regular distribution of points, which has two issues.

1) For practical arrays, it may not be possible to place microphones precisely at all the quadrature locations. Moving even one microphone slightly destroys the quadrature.
2) If higher order beamformers are necessary, quadrature points may be unavailable.

We discuss these issues further in the paper. Here, we propose an approach that allows flexible microphone placements. Then we show how the array can achieve optimal performance.[1]

This paper is organized into four sections. In Section II, we present the basic principle of beamforming using a spherical microphone array. In Section III, we give a theoretical analysis of the discrete system. This part includes a summary of previous work and an analysis of the orthonormality error: how it appears, how it gets amplified, and how it affects array performance. To cancel the error noise optimally, we propose an improved solution and compare several design examples including a practical one. In Section IV, we formulate our optimization problem into a linear system. We simplify the optimization by using reduced *degrees of freedom* (DOFs) for a specified beamforming direction. The resulting beamformer then is checked against the robustness constraint. The upper bound of the beampattern order is derived theoretically. We again use the example from Section III to demonstrate our simplified optimization. However, for especially ill-conditioned layouts, the solution can lack robustness. This limitation is then addressed in Section V by a controlled trade-off between the accuracy of approximation and a specified robustness criterion. We formulate this trade-off as a constrained optimization problem and develop an adaptive implementation. Our algorithm automatically optimizes in two different situations: the beampattern with maximum directivity or the desired beampattern of pre-specified order. Our adaptive implementation inherits the advantages of the classical ones in [7] and [3].

II. Background

The basic principle of a spherical beamformer is to make use of the orthonormality of spherical harmonics to decompose the soundfield arriving at a spherical array. The orthogonal components of the soundfield are then linearly combined to approximate a desired beampattern [15].

A. Scattering Theory

Acoustic wave propagation in a homogeneous medium is described by the wave equation

$$\nabla^2 p(\mathbf{r},t) = \frac{1}{c^2}\frac{\partial^2 p(\mathbf{r},t)}{\partial t^2} \tag{1}$$

[1]Some of the results were reported in [13].

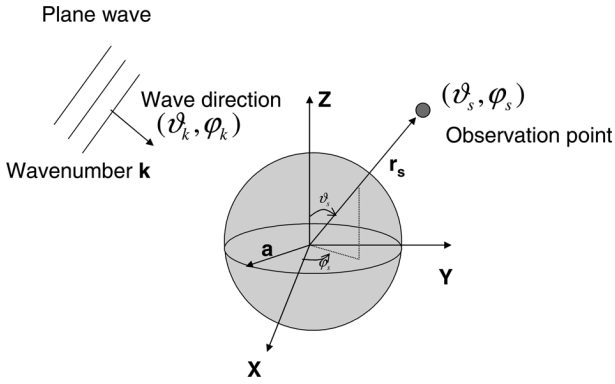

Fig. 1. Plane wave incident on a rigid sphere.

where p is the pressure, $\mathbf{r}$ and t are the location and time of the field point, ∇^2 is the Laplacian, and c is the speed of sound in the medium. Upon taking the Fourier transform, in the frequency domain, the wave equation becomes the Helmholtz equation

$$(\nabla^2 + k^2)\psi(\mathbf{r},\mathbf{k}) = 0, \quad k = 2\pi f/c \tag{2}$$

where f is the frequency and $k = \|\mathbf{k}\|$ is the wavenumber.

As denoted in Fig. 1, for a unit magnitude plane wave with wavenumber k incident from direction $\boldsymbol{\theta}_k = (\theta_k, \varphi_k)$, the incident field at an observation point $\mathbf{r}_s = (\boldsymbol{\theta}_s, r_s)$ can be expanded as

$$\psi_{in}(\mathbf{r}_s,\mathbf{k}) = 4\pi \sum_{n=0}^{\infty} i^n j_n(kr_s) \sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}_k) Y_n^{m*}(\boldsymbol{\theta}_s) \tag{3}$$

where j_n is the spherical Bessel function of order n, Y_n^m is the spherical harmonics of order n and degree m. Here, superscript $*$ denotes the complex conjugation. At the same point, the field scattered by the rigid sphere of radius a is [8]

$$\psi_{scat}(\mathbf{r}_s,\mathbf{k}) = -4\pi \sum_{n=0}^{\infty} i^n \frac{j_n'(ka)}{h_n'(ka)} h_n(kr_s) \times \sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}_k) Y_n^{m*}(\boldsymbol{\theta}_s). \tag{4}$$

The total field on the surface ($r_s = a$) of the rigid sphere is

$$\psi(\boldsymbol{\theta}_s,\boldsymbol{\theta}_k,ka) = \left[\psi_{in}(\mathbf{r}_s,\mathbf{k}) + \psi_{scat}(\mathbf{r}_s,\mathbf{k})\right]\big|_{r_s=a} = 4\pi \sum_{n=0}^{\infty} i^n b_n(ka) \sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}_k) Y_n^{m*}(\boldsymbol{\theta}_s), \tag{5}$$

$$b_n(ka) = j_n(ka) - \frac{j_n'(ka)}{h_n'(ka)} h_n(ka) \tag{6}$$

where h_n are the spherical Hankel functions of the first kind.

B. Soundfield Decomposition and Beamforming

Following [15], let us assume that the pressure recorded at each point $\boldsymbol{\theta}_s$ on the surface of the sphere Ω_s, is weighted by

$$W_{n'}^{m'}(\boldsymbol{\theta}_s, ka) = \frac{Y_{n'}^{m'}(\boldsymbol{\theta}_s)}{4\pi\, i^{n'} b_{n'}(ka)}. \tag{7}$$

Using the orthonormality of spherical harmonics

$$\int_{\Omega_s} Y_n^{m*}(\boldsymbol{\theta}_s) Y_{n'}^{m'}(\boldsymbol{\theta}_s) d\Omega_s = \delta_{nn'}\delta_{mm'} \tag{8}$$

the total output from a pressure-sensitive spherical surface weighted according to (7) is

$$\int_{\Omega_s} \psi(\boldsymbol{\theta}_s, \boldsymbol{\theta}_k, ka) W_{n'}^{m'}(\boldsymbol{\theta}_s, ka) d\Omega_s = Y_{n'}^{m'}(\boldsymbol{\theta}_k). \tag{9}$$

This shows the spatial response of the plane wave incident from $\boldsymbol{\theta}_k$, for a continuous pressure-sensitive spherical microphone, is $Y_{n'}^{m'}(\boldsymbol{\theta}_k)$. Since any square integrable function $F(\boldsymbol{\theta})$ on the unit sphere can be expanded in terms of complex spherical harmonics, we can implement arbitrary beampatterns from this class of functions. For example, an ideal beampattern looking at the direction $\boldsymbol{\theta}_0$ can be modeled as a delta function

$$F(\boldsymbol{\theta},\boldsymbol{\theta}_0) = \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0) \tag{10}$$

which can be expanded into an infinite series of spherical harmonics

$$F(\boldsymbol{\theta},\boldsymbol{\theta}_0) = 2\pi \sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_n^{m*}(\boldsymbol{\theta}_0) Y_n^m(\boldsymbol{\theta}). \tag{11}$$

So the weight at each point $\boldsymbol{\theta}_s$ to achieve this beampattern is

$$w(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s, ka) = \sum_{n=0}^{\infty} \frac{1}{2i^n b_n(ka)} \sum_{m=-n}^{n} Y_n^{m*}(\boldsymbol{\theta}_0) Y_n^m(\boldsymbol{\theta}_s). \tag{12}$$

The advantage of this system is that it can be steered into any three-dimensional (3-D) direction *digitally* with the same beampattern. This is of course for the ideal case of an ideal continuous microphone array on a spherical surface, and to achieve the ideal beam pattern we need to perform infinite summations.

III. Discrete Spherical Array Analysis

This section also follows [15], but we make the band limit restrictions explicit. For a discretely sampled array with S microphones mounted at $\boldsymbol{\theta}_s, s = 1, 2, \ldots, S$, the continuous integrals are approximated by weighted summations, or quadrature[2] [19, p. 71], as follows:

$$\frac{4\pi}{S} \sum_{s=1}^{S} Y_n^{m*}(\boldsymbol{\theta}_s) Y_{n'}^{m'}(\boldsymbol{\theta}_s) C_{n'}^{m'}(\boldsymbol{\theta}_s) = \delta_{nn'}\delta_{mm'} \tag{13}$$
$$(n = 0, \ldots, N_{\text{eff}}; m = -n, \ldots, n;$$
$$n' = 0, \ldots, N; m' = -n', \ldots, n')$$

where $C_{n'}^{m'}(\boldsymbol{\theta}_s)$ is the quadrature coefficient for $Y_{n'}^{m'}$ at $\boldsymbol{\theta}_s$. N_{eff} is the lesser of the following two values: the first is the maximum spatial order of spherical harmonics that can be resolved by a given array; the second is the maximum order at which the incoming sound signal shows spatial variability, which in turn is related to the temporal frequency of the signal as this specifies the wavenumber. As will be seen from the analysis in Section III-B, the precise choice of N_{eff} is not critical. N is the order of beamformer. (13) can be solved in the *least-squares* sense to minimize the 2-norm of the residues for (13).

Therefore, to approximate the truncated expansion of the ideal beampattern in (11) to order N, which we call the *regular beampattern*[3] of order N, defined as

$$F_N(\boldsymbol{\theta},\boldsymbol{\theta}_0) = 2\pi \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^{m*}(\boldsymbol{\theta}_0) Y_n^m(\boldsymbol{\theta}). \tag{14}$$

The weights to achieve the beampattern (14) are

$$w_N(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s, ka) = \sum_{n=0}^{N} \frac{1}{2i^n b_n(ka)} \sum_{m=-n}^{n} Y_n^{m*}(\boldsymbol{\theta}_0) Y_n^m(\boldsymbol{\theta}_s) C_n^m(\boldsymbol{\theta}_s). \tag{15}$$

To evaluate the robustness of a beamformer, we use the *white noise gain* (WNG) [2], usually in decibel scale

$$\text{WNG}(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s, ka) = 10\log_{10}\left(\frac{\left|\mathbf{d}^T\mathbf{W}\right|^2}{\mathbf{W}^H\mathbf{W}}\right) \tag{16}$$

where $\mathbf{d}$ is the column vector of complex pressure at each microphone position produced by the plane wave of unit magnitude from the desired beamforming direction $\boldsymbol{\theta}_0$ and $\mathbf{W}$ is the column vector of complex weights for each microphone. WNG defines the sensitivity on the white noise including the device noise and implicitly, the microphone position mismatches among other spatially uncorrelated perturbations. Positive WNG means an attenuation of white noise, whereas negative means an amplification.

To evaluate the directivity of a beampattern, we use the *directivity index* (DI) [2], also in decibels, as follows:

$$DI(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s, ka) = 10\log_{10}\left(\frac{4\pi|H(\boldsymbol{\theta}_0,\boldsymbol{\theta}_0)|^2}{\int_{\Omega_s}|H(\boldsymbol{\theta},\boldsymbol{\theta}_0)|^2 d\Omega_s}\right) \tag{17}$$

where $H(\boldsymbol{\theta},\boldsymbol{\theta}_0)$ is the actual beampattern looking at $\boldsymbol{\theta}_0$ and $H(\boldsymbol{\theta}_0,\boldsymbol{\theta}_0)$ is the value in that direction. The DI represents the ability of the array to suppress a diffuse noise field. It is the ratio of the gain for the look direction $\boldsymbol{\theta}_0$ to the average gain over all directions. If a spherical microphone array can precisely achieve the regular beampattern of order N as in (14), its theoretical DI is $20\log_{10}(N+1)$. We will show later that a spherical microphone array doesn't necessarily always achieve regular beampatterns of certain orders, it can also be optimized to achieve the maximum DI under a specified WNG constraint. In that case, the resulting beampattern may have irregular shapes other than the regular beampatterns defined in (14).

[2]While cubature is sometimes used for representing nodes and weights in two dimensions, we prefer to use the word quadrature, which is in any case the term used for still higher dimensions.

[3]It is also called the plane-wave decomposition pattern in [17], [16].

A. Previous Work

The previous work can be summarized according to the choices made for $C_n^m(\boldsymbol{\theta}_s)$ and the optimization method used, and the constraints enforced.

In [15], $C_n^m(\boldsymbol{\theta}_s)$ are chosen to be unity to provide relative accuracy for some "uniform" layouts such as the 32 nodes defined by a truncated icosahedron. This straightforward choice simplifies the computation; however, unity weights do not hold for "non-uniform" layouts and their use does not leave any other freedom for optimization of the array beam pattern, such as imposing the WNG constraint. In addition, we will see that even small errors can destroy the beampattern. In [1], several options are mentioned including equiangular grid layout [10] and an intuitive equidistant layout [6]. The common limitation of those schemes is that they are inflexible. If a patch of the spherical surface is inappropriate for mounting microphones, the orthonormality error may be large, thereby destroying the beampattern as the quadrature relation will not hold.

The approach in [11, ch. 3, sec. VI] is equivalent to choosing $C_n^m(\boldsymbol{\theta}_s)$ to be independent of $\boldsymbol{\theta}_s$. The remaining $(N+1)^2$ DOFs are not used to satisfy (13) but maximize the directivity within a specified WNG constraint.[4] The optimization is performed by using an undetermined Lagrangian multiplier. Since there is no simple relation between the multiplier and the resulting WNG, the implementation uses a straightforward trial-and-error strategy.

B. Orthonormality Error Noise Analysis

Unfortunately, (13) cannot be satisfied exactly for over-determined or rank-deficient systems in general, which is usually the case. In addition, the number of equations in (13) for each pair n' and m' depends on N_{eff}. Then, for any choice of $C_{n'}^{m'}(\boldsymbol{\theta}_s)$, we always have

$$\frac{4\pi}{S}\sum_{s=1}^{S} Y_n^{m*}(\boldsymbol{\theta}_s)Y_{n'}^{m'}(\boldsymbol{\theta}_s)C_{n'}^{m'}(\boldsymbol{\theta}_s) = \delta_{nn'}\delta_{mm'} + \epsilon_{nn'}^{mm'} \tag{18}$$

where $\epsilon_{nn'}^{mm'}$ is the nonzero error caused by discreteness.

Now, we will see how this error could degrade the performance of soundfield decomposition. To extract the component of order n' and degree m' from the soundfield (5), we consider the quadrature in (9) for $\psi_n^m(\boldsymbol{\theta}_s, \boldsymbol{\theta}_k, ka)$, denoting one component of ψ at order n and degree m

$$\Pi_n^m(\boldsymbol{\theta}_k, ka) = \int_{\Omega_s} \psi_n^m(\boldsymbol{\theta}_s, \boldsymbol{\theta}_k, ka)\frac{Y_{n'}^{m'}(\boldsymbol{\theta}_s)C_{n'}^{m'}(\boldsymbol{\theta}_s)}{4\pi i^{n'} b_{n'}(ka)} d\Omega_s \tag{19}$$

where

$$\psi_n^m(\boldsymbol{\theta}_s, \boldsymbol{\theta}_k, ka) = 4\pi i^n b_n(ka) Y_n^m(\boldsymbol{\theta}_k) Y_n^{m*}(\boldsymbol{\theta}_s). \tag{20}$$

Using S discrete points, we have

$$\Pi_n^m(\boldsymbol{\theta}_k, ka) = Y_n^m(\boldsymbol{\theta}_k)\left[\frac{i^n b_n(ka)}{i^{n'} b_{n'}(ka)}\right](\delta_{nn'}\delta_{mm'} + \epsilon_{nn'}^{mm'}). \tag{21}$$

[4]In [11, ch. 3, sec. VI], the beampattern is simplified by assuming it to be φ-independent and therefore m is dropped.

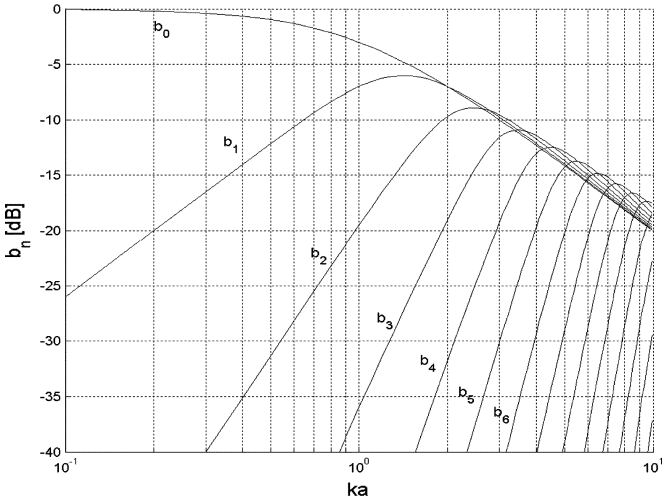

Fig. 2. $b_n(ka)$ for the first few orders. Given ka, $b_n(ka)$ decays very quickly with respect to n.

We notice that

$$\left[\frac{i^n b_n(ka)}{i^{n'} b_{n'}(ka)}\right]\delta_{nn'}\delta_{mm'} = \delta_{nn'}\delta_{mm'}. \tag{22}$$

So, (21) can be rewritten as

$$\Pi_n^m(\boldsymbol{\theta}_k, ka) = Y_n^m(\boldsymbol{\theta}_k)\{\delta_{nn'}\delta_{mm'} + \left[\frac{i^n b_n(ka)}{i^{n'} b_{n'}(ka)}\right]\epsilon_{nn'}^{mm'}\}. \tag{23}$$

The second term is the noise caused by the scaled orthonormality error $\epsilon_{nn'}^{mm'}$. We call it the *orthonormality error noise* (OEN) which is possible with any discrete microphone array layout. To prevent it from damaging the orthonormality we require to achieve beam patterns of various orders, we require

$$\left|\frac{i^n b_n(ka)}{i^{n'} b_{n'}(ka)}\epsilon_{nn'}^{mm'}\right| \ll 1. \tag{24}$$

So, we get:

$$|\epsilon_{nn'}^{mm'}| \ll \left|\frac{b_{n'}(ka)}{b_n(ka)}\right|, \quad \forall n, n', m, m'. \tag{25}$$

Since b_n decays very quickly with respect to n, as shown in Fig. 2, for a given number of microphones and layout, we cannot decompose the high-order component of soundfield if the condition (25) fails. In addition, we can see (25) is independent of the magnitude of the incoming sound wave. This means that even if the microphones have recorded the high-order components, the system may be unable to decompose them.

To prevent errors from being amplified, we include the frequency-dependent scale factor

$$B_{n'}^n(ka) = \frac{b_n(ka)}{b_{n'}(ka)} \tag{26}$$

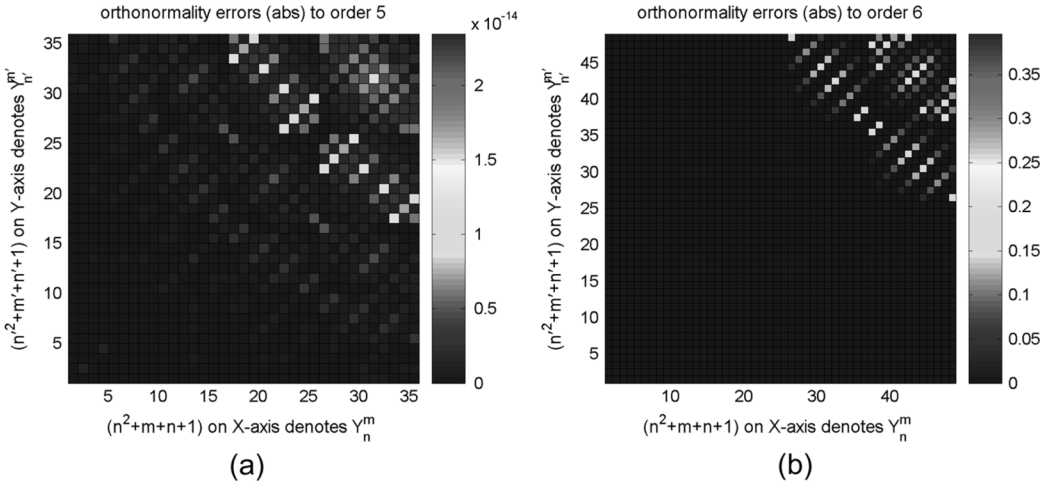

Fig. 3. Precision of the quadrature rules provided by the spherical t-design is only guaranteed in the specified band limit.

into (13). The linear system for quadrature coefficients then becomes a *weighted least-squares* problem

$$\min_{C_{n'}^{m'}(\boldsymbol{\theta}_s,ka)} \left\| \frac{4\pi}{S} \sum_{s=1}^{S} Y_n^{m*}(\boldsymbol{\theta}_s) Y_{n'}^{m'}(\boldsymbol{\theta}_s) C_{n'}^{m'}(\boldsymbol{\theta}_s, ka) B_{n'}^{n}(ka) - \delta_{nn'}\delta_{mm'} \right\|_2^2 \quad (27)$$
$$(n = 0, \ldots, N_{\text{eff}}; m = -n, \ldots, n$$
$$n' = 0, \ldots, N; m' = -n', \ldots, n').$$

The discrete orthonormalities in (27) that are weighted with smaller $|B_{n'}^{n}(ka)|$ are less important than those with larger $|B_{n'}^{n}(ka)|$. Since $B_{n'}^{n}(ka)$ converges to zero when n increases for given n', compared with (13), (27) suppresses the orthonormalities in the higher orders and adds more accuracy to the lower orders. In addition, since $B_{n'}^{n}(ka)$ converges to zero rapidly with respect to n, this weighting also significantly lessens the sensitivity of the solutions to the choice of N_{eff} only if $|b_{N_{\text{eff}}}(ka)|$ is small enough compared with $|b_N(ka)|$, e.g., for $ka = 1.83$, $N_{\text{eff}} = N + 3$ will make $|b_{N_{\text{eff}}}(ka)|$ at least 40dB below $|b_N(ka)|$ for $N = 3, 4, 5$.

C. Design Examples

A quadrature formula provides locations at which we evaluate the function and weights to multiply and sum up to obtain the integral. It was proven that any quadrature formula of order N over the sphere should have more than $S = (N+1)^2$ quadrature nodes [18], [9]. If the quadrature function is the multiplication of two band-limited functions up to order N, to achieve the exact quadrature using equiangular layout, we need $S = 4(N+1)^2$ nodes [4]. This is too large and redundant for our application. For special layouts, S can be made much smaller. For example, for a spherical grid which is a Cartesian product of equispaced grid in φ and the grid in θ in which the nodes distributed as zeros of the Legendre polynomial of degree $N+1$ with respect to $\cos\theta$, we need $S = 2(N+1)^2$ [17]. However, these points are quite inconveniently distributed. Another special design with equal quadrature coefficients is called the *spherical t-design* [9]. While this design achieves exact quadratures, it applies strictly to band-limited functions and S is still large considering our quadrature function is the multiplication of *two* band-limited functions. As shown in Fig. 3, using the 64 node t-design, orthonormality errors are extremely small up to order five, but increase significantly above order five.

If we use an approximate quadrature formula, then it may be possible to reduce S. Intuitively, we want the microphones distributed "uniformly" on the spherical surface. Unfortunately, it has been proven that only five regular polyhedrons (also called "Platonic solids") exist: the tetrahedron, cube, dodecahedron, octahedron, and icosahedron [5]. Semi-regular polyhedrons can be used also such as the truncated icosahedron used in [15] to layout 32 microphones. The general problem to distribute arbitrary number of points approximately "uniformly" on a spherical surface is numerically solved by Fliege in [6] by minimizing the potential energy of a distribution of movable electrons on a perfect conducting sphere. Then, a set of quadrature coefficients (which are unequal) are obtained by an optimization procedure. This is in contrast to spherical t-designs, where the quadrature coefficients are equal. In our experience, and as can be seen from a comparison of Fig. 3(b) and Fig. 4(b), the Fliege nodes are more robust to functions that have higher spatial frequency components. Fig. 4(a) shows Fliege's 64-node layout in [6]. Fig. 4(b) shows the orthonormality errors of spherical harmonics using those optimal quadrature coefficient for each nodes.

However, those optimal quadrature coefficients are not generally available for other flexible layouts. More importantly, for a given layout, especially for a "nonuniform" one, we will explain in the next section that there is no single optimal set of coefficients to satisfy all orthonormalities. For example, Fig. 5(a) is the layout of our array using the angular positions of those 64 nodes with four nodes at the bottom removed because of the cable outlet. Fig. 5(b) shows the orthonormality errors using the 60 nodes and quadrature coefficients. It is less accurate compared with Fig. 4(b). The top row in Fig. 6 shows the beampatterns from order three to five using this configuration. At order three, the beampattern is distorted. At order four, the orthonormality errors significantly damage the beampattern. At order five, the beampattern is almost completely destroyed. The bottom row in Fig. 6 shows the beampatterns using the quadrature coefficients solved by (27), which optimally approximate the regular beampatterns.

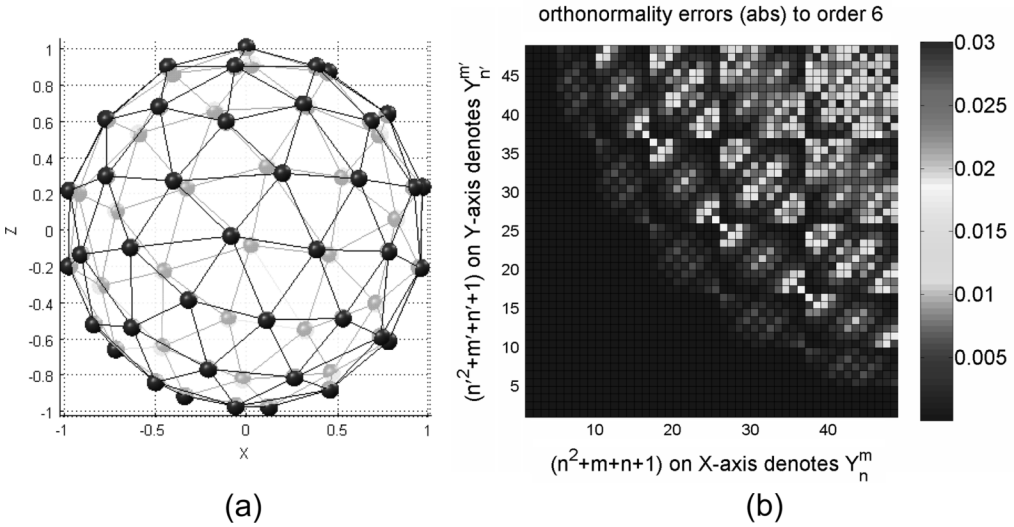

Fig. 4. (a) Fliege's 64 nodes. (b) Orthonormality errors.

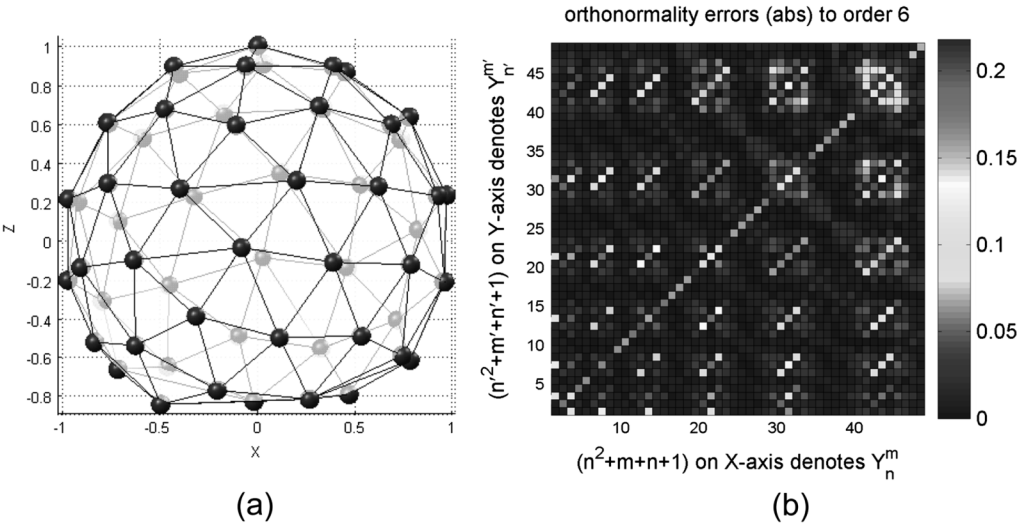

Fig. 5. (a) Same as in Fig. 4(a) except the bottom four nodes are removed. (b) Orthonormality errors.

IV. Simplified Optimization of Desired Beampattern for Discrete Array

In total, we have $S \times (N+1)^2$ quadrature coefficients for each frequency, which can be viewed as degrees of freedom. However, those coefficients are not directly related to the WNG constraint, and we cannot find an explicit constrained optimization easily. In addition, the $S \times (N+1)^2$ DOFs are intuitively redundant for satisfying the WNG constraint, specifically since we have only S microphones, and each microphone will be assigned to one final weight no matter how many intermediate coefficients are achieved. That means $C_{n'}^{m'}(\boldsymbol{\theta}_s, ka)$ should somehow be independent of n' and m'. Plausibly, we have the following least-squares problem with only S variables:

$$\frac{4\pi}{S}\sum_{s=1}^{S} Y_n^{m*}(\boldsymbol{\theta}_s) Y_{n'}^{m'}(\boldsymbol{\theta}_s) C(\boldsymbol{\theta}_s, ka) \frac{b_n(ka)}{b_{n'}(ka)} = \delta_{nn'}\delta_{mm'}. \tag{28}$$

This method, however, is infeasible especially for ill-conditioned layouts since in practice we aim to use as few microphones as possible, so the number of linear equations will largely outnumber the number of microphones which makes the optimization less meaningful. Instead of satisfying (28), we can use $C(\boldsymbol{\theta}_s, ka)$ to maximize the directivity subject to the WNG constraint, which simply returns to a classical solved problem of designing a superdirective and robust beamformer [3] in a white noise field. This method does not aim to optimally approximate a desired regular beampattern of a specified order N.

We pay the price that the beampattern coefficients are no longer the same in all steering directions. Because the optimization on (28) is independent of the beamforming direction, it is too ambitious to find such a single set of optimal weights for every beamforming directions. This, in turn, also implies that the optimization using S DOFs should be performed for each beamforming direction. In another words, to directly control the WNG for a flexible microphone layout, we have to sacrifice the modal beamformer structure described in [15]. To make the steering easier, we design an adaptive implementation in Section V-B.

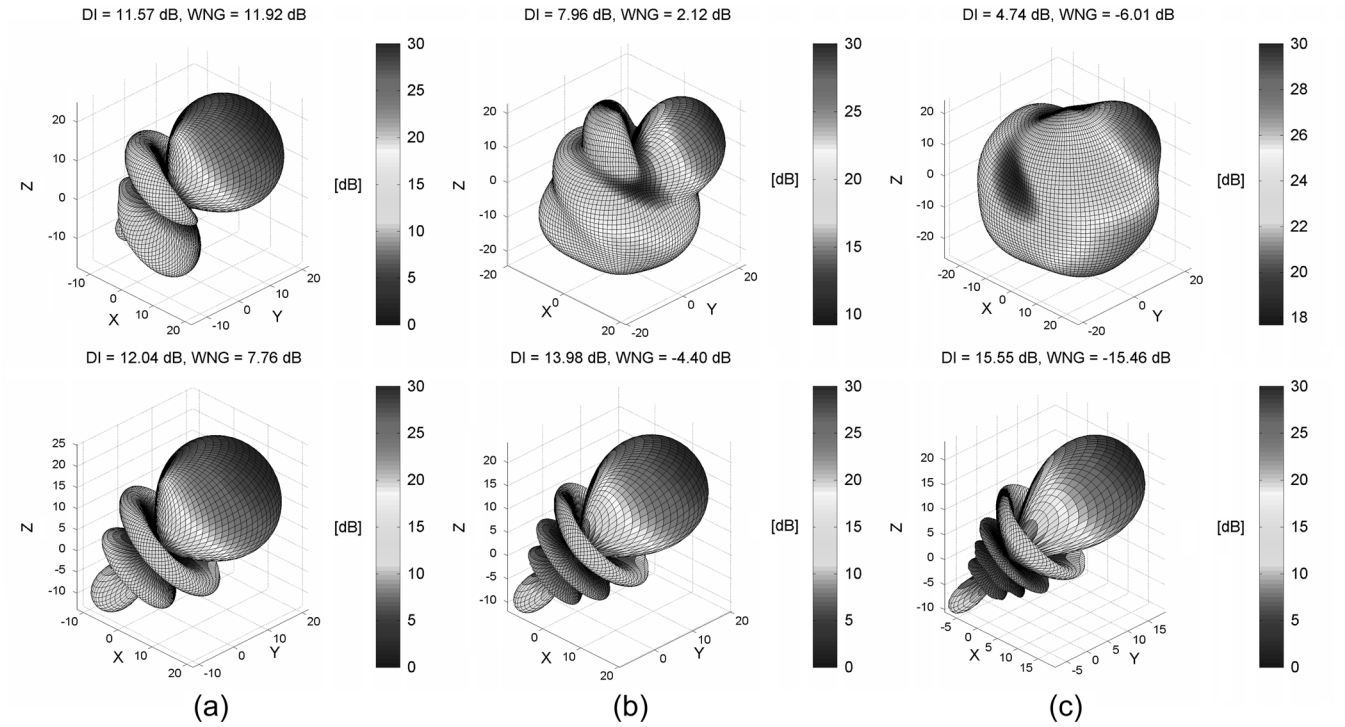

Fig. 6. (a)–(c) Beampatterns from order 3 to 5 using the 60-node array in Fig. 5. The top row uses Fliege's quadrature coefficients. The bottom row uses the coefficients solved by (27). Simulated at 1 kHz with $a = 10$ cm.

We will follow the above argument to simplify the optimization: instead of fixing individual orthonormalities of spherical harmonics as we did before, we will optimally fit the desired beampattern directly using reduced DOFs. In this case, we formulate the discrete spherical beamformer problem in to a linear system with respect to the S weights for the S microphones, respectively, so that the WNG can be directly controlled. The linear system says that if these S weights are solved under the WNG constraint, they will optimally generate the desired beampatterns. Then we will present straightforward solutions. In Section IV-B, we will derive the upper bound of beamforming order of a robust beamformer. Design examples are provided to demonstrate our approach.

A. Discrete Spherical Beamformer as A Constrained Linear System

To achieve a regular beampattern of order N (14), a discrete spherical beamformer with S microphones can be formulated as a finite linear system

$$\mathbf{AW} = c_N \mathbf{B}_N \tag{29}$$

$$\mathbf{d}^T \mathbf{W} = 1 \tag{30}$$

where (29) defines the beampattern, and (30) means that the plane wave incident from the beamforming direction will be faithfully reconstructed. In (29), $\mathbf{A}$ are the coefficients of the spherical harmonics expansion of the soundfield in (5)

$$\mathbf{A} = [\mathbf{A}_1\, \mathbf{A}_2 \cdots \mathbf{A}_S], \tag{31}$$

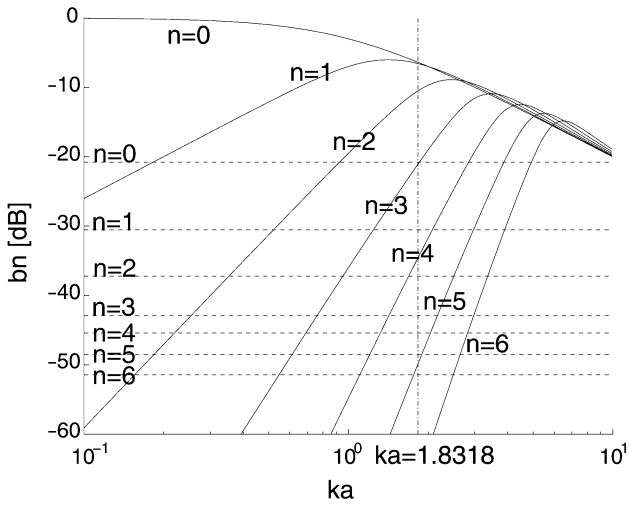

Fig. 7. Using the 60-nodes array in Fig. 5, the horizontal lines show the mode bounds for robust beamforming at given ka.

$$\mathbf{A}_s = 4\pi \begin{bmatrix} i^0 b_0(ka) Y_0^{0*}(\boldsymbol{\theta}_s) \\ i^1 b_1(ka) Y_1^{-1*}(\boldsymbol{\theta}_s) \\ \vdots \\ i^N b_N(ka) Y_N^{N*}(\boldsymbol{\theta}_s) \\ i^{(N+1)} b_{(N+1)}(ka) Y_{(N+1)}^{-(N+1)*}(\boldsymbol{\theta}_s) \\ \vdots \\ i^{N_{\text{eff}}} b_{N_{\text{eff}}}(ka) Y_{N_{\text{eff}}}^{N_{\text{eff}}*}(\boldsymbol{\theta}_s) \end{bmatrix}. \tag{32}$$

$(s = 1, \ldots, S.)$

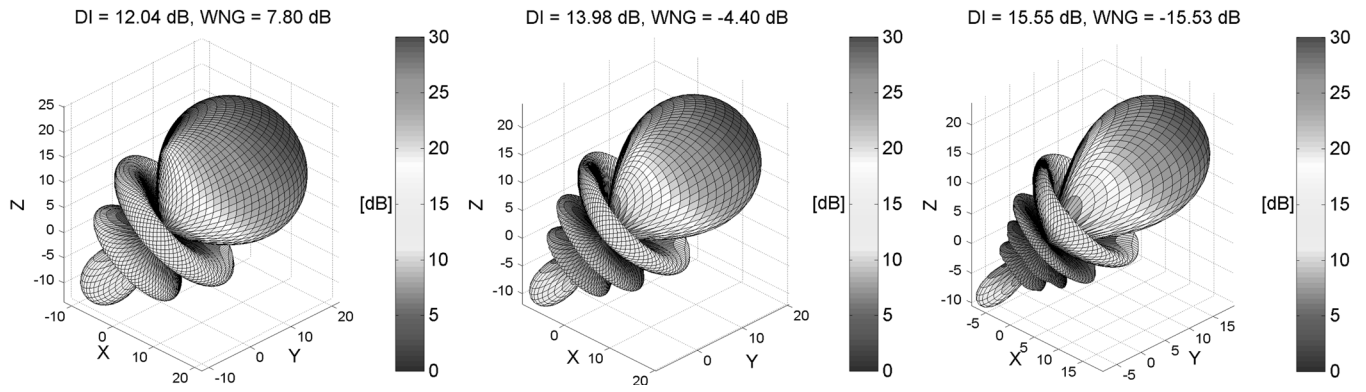

Fig. 8. Beampatterns of order 3 to 5 for the 60-nodes array in Fig. 5 using simplified optimization.

$\mathbf{W}$ is the vector of complex weights to be assigned to each microphone at $(\boldsymbol{\theta}_s, a)$

$$\mathbf{W} = \begin{bmatrix} W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, ka) \\ W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_2, ka) \\ \vdots \\ W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_S, ka) \end{bmatrix}. \tag{33}$$

As explained earlier, $\mathbf{W}$ is dependent on $\boldsymbol{\theta}_0$. $\mathbf{B}_N$ is the vector of coefficients of the beampattern of order N steered to $\boldsymbol{\theta}_0$ in (14)

$$\mathbf{B}_N = 2\pi \begin{bmatrix} Y_0^{0*}(\boldsymbol{\theta}_0) \\ Y_1^{-1*}(\boldsymbol{\theta}_0) \\ \vdots \\ Y_N^{N*}(\boldsymbol{\theta}_0) \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{34}$$

In (30), $\mathbf{d}$ is the row vector of the complex pressure (the Fourier transform of the signal recorded at each microphone position) produced by a plane wave of unit magnitude from the desired beamforming direction $\boldsymbol{\theta}_0$

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_S \end{bmatrix} = \begin{bmatrix} \psi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_0, ka) \\ \psi(\boldsymbol{\theta}_2, \boldsymbol{\theta}_0, ka) \\ \vdots \\ \psi(\boldsymbol{\theta}_S, \boldsymbol{\theta}_0, ka) \end{bmatrix}. \tag{35}$$

In (29), c_N is a normalizing coefficient to satisfy the all-pass frequency response constraint (30). The least-squares solution of (29) is

$$\mathbf{W} = \left[(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\right] c_N \mathbf{B}_N. \tag{36}$$

Then, c_N can be determined using (30). If we assume (29) has small residues, from (14), the *a priori* estimate of c_N is

$$c_N \approx \frac{1}{2\pi \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}_0) Y_n^{m*}(\boldsymbol{\theta}_0)}. \tag{37}$$

According to the spherical harmonic addition theorem, (37) can be simplified easily, and we see that c_N is independent of $\boldsymbol{\theta}_0$

$$c_N \approx \frac{1}{\sum_{n=0}^{N} \frac{2n+1}{2} P_n(\cos 0)} = \frac{2}{(N+1)^2}. \tag{38}$$

We will use it to predict the maximum order in the next subsection.

Note that because we have absorbed all frequency dependence (ka) into the linear system, it must be solved for each frequency.

B. Maximum Beampattern Order for Robust Beamformer

A robust beamformer requires a minimum WNG of δ^2 (such as the -6 decibel value used in [2])

$$\frac{\left|\mathbf{d}^T\mathbf{W}\right|^2}{\mathbf{W}^H\mathbf{W}} \geq \delta^2. \tag{39}$$

Substituting (30) into (39), we have a spherical constraint on $\mathbf{W}$

$$\mathbf{W}^H\mathbf{W} \leq \delta^{-2}. \tag{40}$$

Assume the maximum order we can possibly decompose robustly is $N_{\max}$, then the linear system (29) becomes

$$\mathbf{A}\mathbf{W} = c_{N_{\max}} \mathbf{B}_{N_{\max}} \tag{41}$$

where

$$c_{N_{\max}} \approx \frac{2}{(N_{\max}+1)^2}. \tag{42}$$

Suppose we have a least-squares solution of $\mathbf{W}$ to (41), considering the following equations of order $N_{\max}$:

$$\begin{aligned} &4\pi \begin{bmatrix} i^{N_{\max}} b_{N_{\max}}(ka) Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_1) \\ i^{N_{\max}} b_{N_{\max}}(ka) Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_2) \\ \cdots \\ i^{N_{\max}} b_{N_{\max}}(ka) Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_S) \end{bmatrix}^T \begin{bmatrix} W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, ka) \\ W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_2, ka) \\ \vdots \\ W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_S, ka) \end{bmatrix} \\ &\approx 2\pi c_{N_{\max}} Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_0) \\ &(M = -N_{\max}, \ldots, N_{\max}). \end{aligned} \tag{43}$$

Using Cauchy's inequality, we have

$$\begin{aligned}&\left|\sum_{s=1}^{S} i^{N_{\max}} b_{N_{\max}}(ka) Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_s) W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s, ka)\right|^2 \\ &\le \left(\sum_{s=1}^{S} \left|i^{N_{\max}} b_{N_{\max}}(ka) Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_s)\right| |W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s, ka)|\right)^2 \\ &\le \sum_{s=1}^{S} \left|i^{N_{\max}} b_{N_{\max}}(ka) Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_s)\right|^2 \sum_{s=1}^{S} |W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s, ka)|^2 \\ &= |b_{N_{\max}}(ka)|^2 \sum_{s=1}^{S} |Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_s)|^2 \sum_{s=1}^{S} |W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s, ka)|^2.\end{aligned}$$

So

$$\begin{aligned}\mathbf{W}^H\mathbf{W} &= \sum_{s=1}^{S} |W(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s, ka)|^2 \\ &\gtrsim \frac{|c_{N_{\max}} Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_0)|^2}{2|b_{N_{\max}}(ka)|^2 \sum_{s=1}^{S} \left|Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_s)\right|^2}. \end{aligned} \quad (44)$$

From (40), we have

$$\delta^{-2} \gtrsim \frac{\left|c_{N_{\max}} Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_0)\right|^2}{2|b_{N_{\max}}(ka)|^2 \sum_{s=1}^{S} |Y_{N_{\max}}^{M*}(\boldsymbol{\theta}_s)|^2} \quad (45)$$

which is

$$|b_{N_{\max}}(ka)| \gtrsim \frac{\sqrt{2}|Y_{N_{\max}}^{M}(\boldsymbol{\theta}_0)|\delta}{(N_{\max}+1)^2\sqrt{\sum_{s=1}^{S} |Y_{N_{\max}}^{M}(\boldsymbol{\theta}_s)|^2}} \quad (46)$$

for all $M = -N_{\max}, \ldots, N_{\max}$. Therefore, given a spherical microphone array and the beamforming direction, for each frequency, we know the upper bound order $N_{\max}$ of a robust beamformer.

C. Design Examples

We still use the 60 node layout at 1 kHz as an example. We set $\delta = 10^{-0.3}$, i.e., the minimum WNG is $-$6dB. Fig. 7 shows the bounds for $\boldsymbol{\theta}_0 = (\pi/4, \pi/4)$, where the dashed lines show the bounds. For example, the dash-dotted line shows 1 kHz ($ka = 1.8318$) in our case. Since the intersection of the black line and the $n = 4$ mode is above the $n = 4$ bound, while its intersection with $n = 5$ mode is below the $n = 5$ bound, we predict the maximum order of a robust beamformer is four. Fig. 8 shows the resulting beampatterns of order from three to five using the simplified optimization. The WNG of an order-5 beamformer falls below the minimum of -6 dB as predicted. These results are almost identical to the bottom row in Fig. 6

As a general example, Fig. 9 shows a random layout of 64 nodes. Fig. 10 shows the beampatterns from order three to five at 1 kHz. Their WNGs are all below -6 dB. It seems to be contradictory to the bounds shown in Fig. 11, however, we will show in the next section that the robust beamformers up to order four are still achievable, with constrained relaxation on the least-squares solution (36).

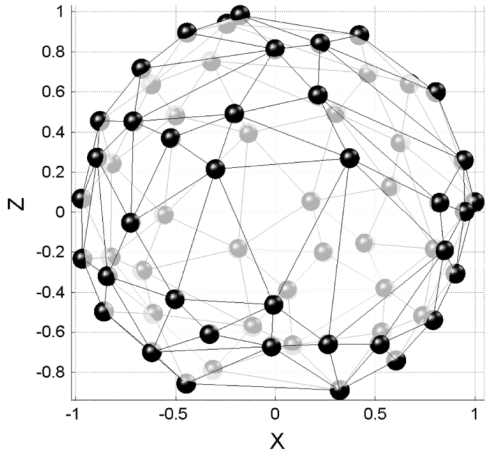

Fig. 9. Random layout of 64 microphones on a sphere of radius 10 cm.

V. Optimal Approximation Subject to the WNG Constraint

In the previous section, we minimize the residue of a finite linear system and check the resulting beamformer against the WNG constraint. In this section, we will extend the algorithm further to address the following two aspects:

1) we need to relax the approximations to stay within the constraint such as in Fig. 9;
2) we need a robust beamformer with maximum directivity index.

The two problems are closely related to each other and can be formulated as a unified constrained optimization problem.

A. Constrained Optimization

To design a robust spherical beamformer with finite microphones, yet optimally approximate the desired beampattern to a certain order (e.g., the ideal beampattern in our case), we need to optimize the following 2-norm function

$$\min_{\mathbf{W}} \|\mathbf{A}\mathbf{W} - c_N\mathbf{B}_N\|_2^2 \quad (47)$$

subject to

$$\mathbf{d}^T\mathbf{W} = 1 \quad (48)$$

$$\mathbf{W}^H\mathbf{W} \le \delta^{-2}. \quad (49)$$

This optimization can be numerically solved by some blackbox software packages, such as the MATLAB function `fmincon`, etc. Another way is to use Tikhonov regularization. Specifically, we place a 2-norm constraint on $\mathbf{W}$ by appending a damping matrix with the regularization parameter λ

$$\begin{bmatrix}\mathbf{A} \\ \lambda\mathbf{I}\end{bmatrix}\mathbf{W} = \begin{bmatrix}c_N\mathbf{B}_N \\ \mathbf{0}\end{bmatrix}. \quad (50)$$

The solution is

$$\mathbf{W} = [(\mathbf{A}^H\mathbf{A} + \lambda^2\mathbf{I})^{-1}\mathbf{A}^H]c_N\mathbf{B}_N.$$

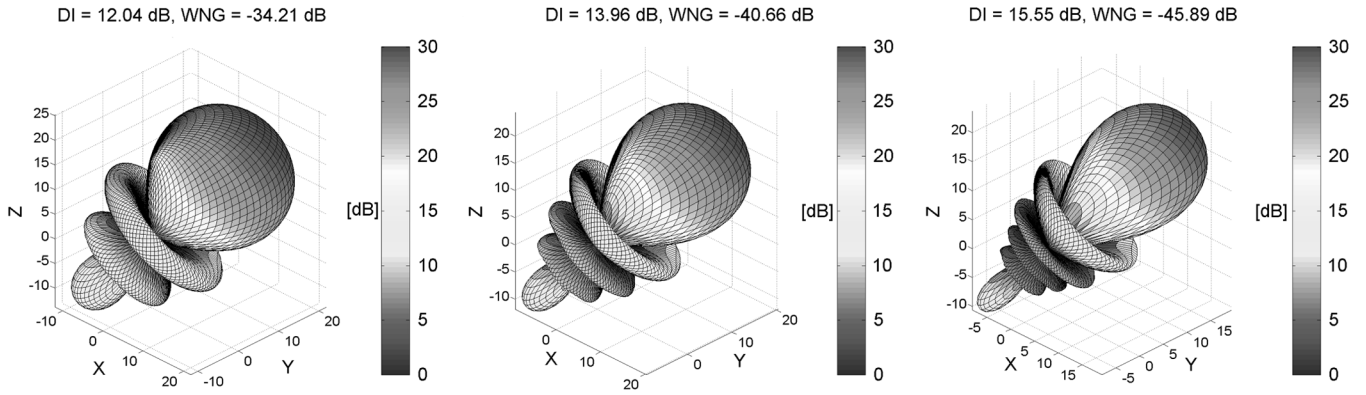

Fig. 10. Unconstrained beampatterns from order 3 to 5 for the array in Fig. 9.

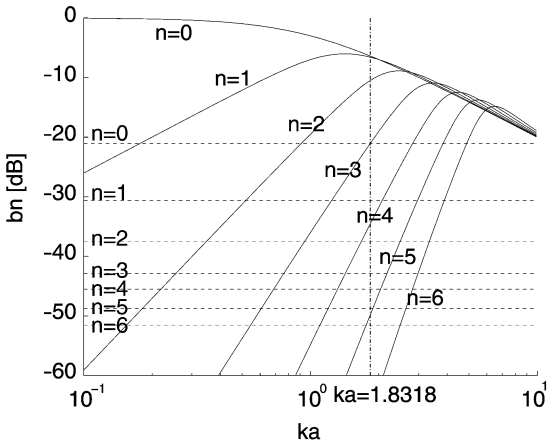

Fig. 11. Mode bounds for the array in Fig. 9.

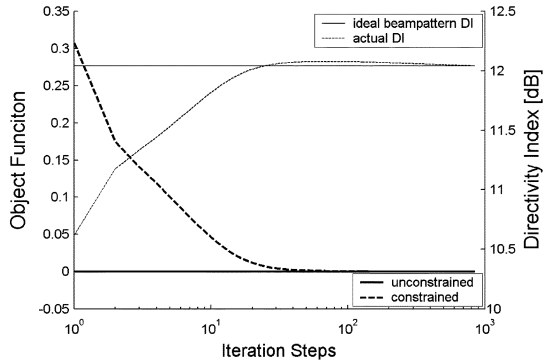

Fig. 12. Iteration process for beampattern of order 3. The thick curves use the left scale, thin curves right scale.

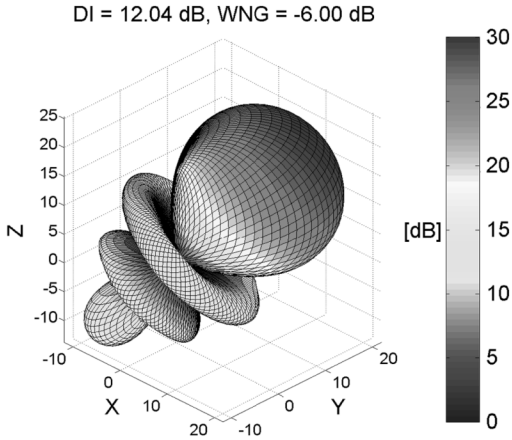

Fig. 13. Constrained optimal beampattern of order 3.

This regularization parameter λ, however, is not directly related to the WNG constraint. A trial-and-error strategy can be used in implementation.

B. Adaptive Implementation

The most straightforward way to implement this system is to precompute all the weights for each pre-defined 3-D direction and store them in a lookup table. This method, however, is not very efficient because of the obvious trade-off between the spatial resolution and the cost of storage. In this subsection, we reformulate our problem so that we can parallel the method presented in [3] to design an adaptive implementation which automatically and robustly converges to the desired beamformer of a specified order in any steering directions.

We rewrite the objective function into an ellipsoidal form, as follows:

$$\min_{\mathbf{W}} \|\mathbf{A}\mathbf{W} - c_N\mathbf{B}_N\|_2^2 = \min_{\tilde{\mathbf{W}}} \tilde{\mathbf{W}}^H\mathbf{R}\tilde{\mathbf{W}} \tag{51}$$

subject to

$$\mathbf{C}^H\tilde{\mathbf{W}} = \mathbf{g} \tag{52}$$

$$\tilde{\mathbf{W}}^H\tilde{\mathbf{W}} \le \delta^{-2} + 1 \tag{53}$$

where

$$\tilde{\mathbf{W}} = \begin{bmatrix} \mathbf{W} \\ W_0 \end{bmatrix}, \mathbf{R} = \begin{bmatrix} \mathbf{A}^H \\ c_N\mathbf{B}_N^H \end{bmatrix} [\mathbf{A} \quad c_N\mathbf{B}_N]$$

$$\mathbf{C} = \begin{bmatrix} d_1^* & 0 \\ d_2^* & 0 \\ \vdots & \vdots \\ d_S^* & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{g} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

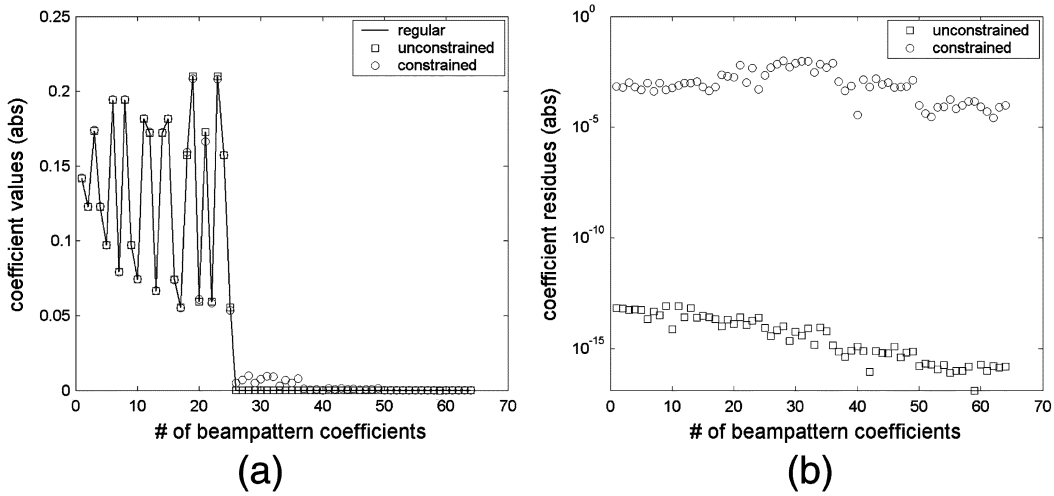

Fig. 14. Precision comparisons for order 4 beamforming. (a) Comparison of unconstrained and constrained beampattern coefficients with regular order four beampattern coefficients $c_4\mathbf{B}_4$. (b) Residue comparison between unconstrained and constrained beampattern coefficients. Both plots show the absolute values.

We know $W_0 = -1$ from (51), however, we include it as an extra variable into $\tilde{\mathbf{W}}$ and its actual value is automatically determined by the constraint (52) in the process of optimization.

To solve this optimization, we first decompose $\tilde{\mathbf{W}}$ into its orthogonal components:

$$\begin{aligned}\tilde{\mathbf{W}} &= \mathbf{W}_c + \mathbf{V}, \\ \mathbf{W}_c &= \mathbf{C}[\mathbf{C}^H\mathbf{C}]^{-1}\mathbf{g}.\end{aligned} \tag{54}$$

$\mathbf{W}_c$ is the least-squares solution to satisfy the linear constraint (52). The residue is expected to be zero since usually (52) is a highly under-determined system. Substituting (54) into (53), we have

$$\mathbf{V}^H\mathbf{V} \le \delta^{-2} + 1 - \mathbf{g}^H[\mathbf{C}^H\mathbf{C}]^{-1}\mathbf{g} = b^2. \tag{55}$$

Thus, the WNG constraint becomes a spherical constraint on $\mathbf{V}$. Since $\mathbf{R}\tilde{\mathbf{W}}(t)$ is the gradient of the object function (51) at step t, the tentative update vector is

$$\tilde{\mathbf{V}}(t+1) = \tilde{\mathbf{P}}_c[\mathbf{V}(t) - \mu\mathbf{R}\tilde{\mathbf{W}}(t)]. \tag{56}$$

$\mathbf{V}(t)$ is the scaled projection of $\tilde{\mathbf{V}}(t)$ into the sphere surface of radius b

$$\mathbf{V}(t) = \begin{cases} \tilde{\mathbf{V}}(t), & \text{for } |\tilde{\mathbf{V}}|^2 \le b^2 \\ b\dfrac{\tilde{\mathbf{V}}(t)}{|\tilde{\mathbf{V}}(t)|}, & \text{for } |\tilde{\mathbf{V}}|^2 > b^2 \end{cases}. \tag{57}$$

μ is the step size, and $\tilde{\mathbf{P}}_c$ is the null space of $\mathbf{C}^H$

$$\tilde{\mathbf{P}}_c = \mathbf{I} - \mathbf{C}[C^H\mathbf{C}]^{-1}\mathbf{C}^H. \tag{58}$$

The weights are updated as

$$\tilde{\mathbf{W}}(t+1) = \mathbf{W}_c + \begin{cases} \tilde{\mathbf{V}}(t+1) & \text{for } |\tilde{\mathbf{V}}|^2 \le b^2 \\ b\dfrac{\tilde{\mathbf{V}}(t+1)}{|\tilde{\mathbf{V}}(t+1)|} & \text{for } |\tilde{\mathbf{V}}|^2 > b^2. \end{cases} \tag{59}$$

$$= \mathbf{W}_c + \mathbf{V}(t+1). \tag{60}$$

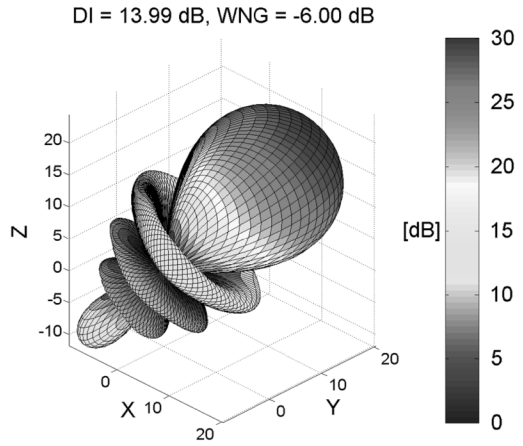

Fig. 15. Constrained optimal beampattern of order 4.

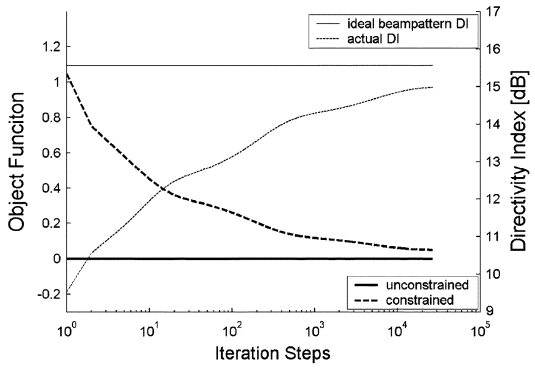

Fig. 16. Iteration process optimally approximates the DI of the regular beampattern of order 5. The thick curves use the left scale, thin curves right scale.

We set the initial guess as

$$\tilde{\mathbf{W}}(0) = \mathbf{R}^{-1}\mathbf{C}[\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}]^{-1}\mathbf{g} \tag{61}$$

$$\tilde{\mathbf{V}}(0) = \tilde{\mathbf{W}}(0) - \mathbf{W}_c \tag{62}$$

which is equivalent to the solution we obtained in Section IV. If the resulting WNG satisfies the constraint, the iteration will

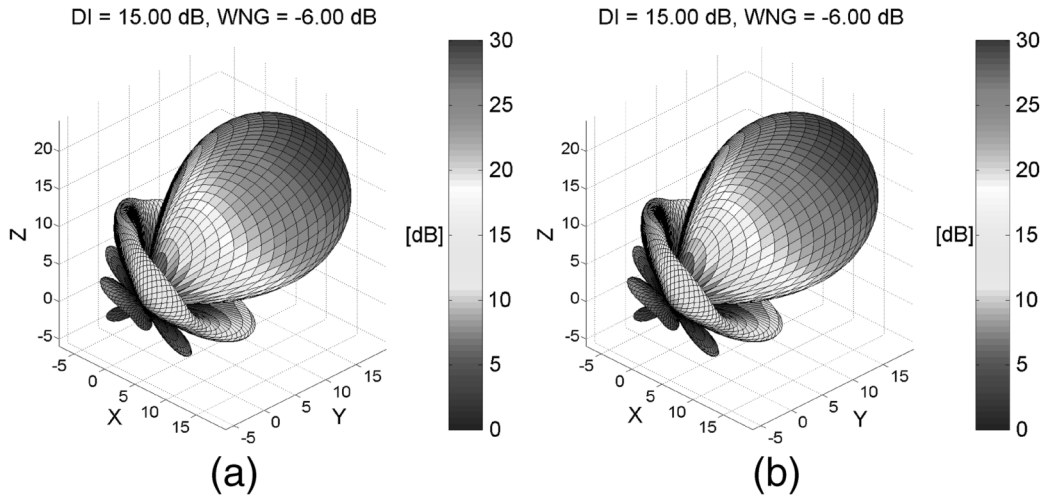

Fig. 17. (a) Constrained optimal beampattern of order 5. It is actually a superdirective beampattern in an ambient white noise field. (b) Beampattern of the regular implementation of a superdirective beamformer in an ambient white noise field.

stay with this solution, otherwise, it will start the constrained optimization process, both automatically. At each step, the constraints (52) and (53) are satisfied exactly. In addition, similar to the methods in [3] and [7], round-off errors do not accumulate. This iteration is independent of the actual signal processing rate, so it may be implemented more efficiently as a parallel unit with other processors.

C. Convergence and Optimal Step Size

From (56) and (59), we have

$$\begin{aligned}\tilde{\mathbf{V}}(t+1) &= \tilde{\mathbf{P}}_c[\mathbf{V}(t) - \mu\mathbf{R}[\mathbf{W}_c + \mathbf{V}(t)]] \\ &= \tilde{\mathbf{P}}_c[\mathbf{I}-\mu\mathbf{R}]\mathbf{V}(t) - \mu\tilde{\mathbf{P}}_c\mathbf{R}\mathbf{W}_c. \end{aligned} \tag{63}$$

Let

$$\tilde{\mathbf{V}}(t+1) = \gamma(t+1)\mathbf{V}(t+1) \tag{64}$$

$$\gamma(t+1) = \begin{cases} 1 & \text{for } |\tilde{\mathbf{V}}|^2 \le b^2 \\ \frac{|\tilde{\mathbf{V}}(t+1)|}{b} & \text{for } |\tilde{\mathbf{V}}|^2 > b^2 \end{cases} \ge 1. \tag{65}$$

We then have

$$\begin{aligned}&\mathbf{V}(t+1) \\ &= \frac{1}{\gamma(t+1)}\tilde{\mathbf{P}}_c[\mathbf{I}-\mu\mathbf{R}]\mathbf{V}(t) - \frac{1}{\gamma(t+1)}\mu\tilde{\mathbf{P}}_c\mathbf{R}\mathbf{W}_c \\ &= \frac{\tilde{\mathbf{P}}_c[\mathbf{I}-\mu\mathbf{R}]^{t+1}}{\prod_{i=0}^{t}\gamma(i+1)}\mathbf{V}(0) - \left[\sum_{k=0}^{t}\frac{\tilde{\mathbf{P}}_c^k[\mathbf{I}-\mu\mathbf{R}]^k}{\prod_{i=0}^{k}\gamma(i+1)}\right]\mu\tilde{\mathbf{P}}_c\mathbf{R}\mathbf{W}_c. \end{aligned} \tag{66}$$

To guarantee convergence, we need

$$0 < \mu < \frac{2}{\sigma_{\max}} \tag{67}$$

where $\sigma_{\max}$ is the maximum eigenvalue of $\mathbf{R}$. Although it is difficult to precisely model and control γ if possible, it will not cause divergence from (65) and will be very close to one in practice. Thus, the optimal step size μ_{opt} can be roughly estimated as the least-squares solution of

$$\mu_{\text{opt}}\boldsymbol{\sigma} = \mathbf{1} \tag{68}$$

which is

$$\mu_{\text{opt}} = [(\boldsymbol{\sigma}^T\boldsymbol{\sigma})^{-1}\boldsymbol{\sigma}^T]\mathbf{1}, \tag{69}$$

where $\boldsymbol{\sigma}$ is the column vector of eigenvalues of $\mathbf{R}$, and $\mathbf{1}^T \triangleq [1, \ldots, 1]$ is the vector of ones with the length of $\boldsymbol{\sigma}$.

D. Simulation Results

We use our algorithm to solve the two cases we mentioned at the beginning of this section. We first go back to the example in Figs. 9 and 10. We consider a beamformer of order three. Fig. 12 shows the iteration process using the optimal step size. As can be seen from this figure, the optimization goal is not to maximize the DI, instead it converges to the regular beampattern of order three. The resulting beampattern is shown in Fig. 13. Fig. 15 shows the optimal approximations of the regular beampattern of order four subject to the WNG constraint. There is minimal difference between the beampatterns in Figs. 15 and 9(b). The comparisons of residues are shown in Fig. 14.

If we desire optimal directivity, we can approximate the ideal beampattern as (11). In practice, we just need to approximate an order above the theoretical upper bound derived in (46), such as order 5 in this case. It is best to demonstrate this via simulations. Fig. 16 clearly shows the actual DI is approaching the regular DI of order five. Fig. 17(a) shows the resulted beampattern. Fig. 17(b) shows the regular implementation of superdirective beamformer in an ambient white noise field, which results in the nearly identical beampattern as Fig. 17(a). These simulations also demonstrates our algorithm can robustly reconfigure itself after microphone reorganization if the new positions are known.

VI. Conclusion

This paper describes a flexible and optimal design of spherical microphone arrays for beamforming. We analyze the effects

of discrete orthonormality errors of spherical harmonics on resulting spherical beamformers, especially for nonuniform microphone layouts. We first design a weighted least-squares approach to correct the effects of the individual orthonormality errors. However, since the solved coefficients cannot be used to control the resulting beamformer's robustness directly, we then design a constrained optimization algorithm. Instead of fixing individual orthonormality errors, this algorithm optimally approximates the desired beampatterns using reduced DOFs. To obtain easier steering, we develop an adaptive implementation, which optimally converges to the desired beampatterns under specified robustness constraint. The resulting beamformer has either regular or maximum directivity beampatterns. Various design examples are presented to demonstrate our algorithms. In current work [12], these algorithms have been practically implemented and found to exhibit performance close to the expected theoretical one, and these are being prepared for submission.

REFERENCES

[1] T. D. Abhayapala and D. B. Ward, "Theory and design of high order sound field microphones using spherical microphone array," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'02)*, May 2002, vol. 2, pp. 1949–1952.

[2] M. Brandstein and D. Ward, Eds., *Microphone Arrays*. New York: Springer-Verlag, 2001.

[3] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 10, pp. 1365–1376, Oct. 1987.

[4] J. R. Driscoll and J. D. M. Healy, "Computing fourier transforms and convolutions on the 2-sphere," *Adv. Appl. Math.*, vol. 15, pp. 202–250, 1994.

[5] Euclid, *The Thirteen Books of the Elements: Books X-XIII*. New York: Dover, 1956, vol. 3.

[6] J. Fliege and U. Maier, "The distribution of points on the sphere and corresponding cubature formulae," *IMA J. Numer. Anal.*, vol. 19, pp. 317–334, 1999.

[7] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.

[8] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. New Providence, NJ: Elsevier Science, 2005.

[9] R. H. Hardin and N. J. A. Sloane, "McLaren's improved snub cube and other new spherical designs in three dimensions," *Discrete and Comput. Geom.*, vol. 15, pp. 429–441, 1996.

[10] D. Healy, D. Rockmore, and S. Moor, "An FFT for the 2-sphere and applications," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'96)*, May 1996, vol. 3, pp. 1323–1326.

[11] Y. Huang and J. Benesty, Eds., *Audio Signal Processing For Next-Generation Multimedia Communication Systems*. Norwell, MA: Kluwer, 2004.

[12] Z. Li, "The Capture and Recreation of 3D Auditory Scenes," Ph.D. dissertation, Univ. Maryland, College Park, 2005.

[13] Z. Li and R. Duraiswami, "A robust and self-reconfigurable design of spherical microphone array for multi-resolution beamforming," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'05)*, Mar. 2005, vol. IV, pp. 1137–1140.

[14] Z. Li, R. Duraiswami, and N. A. Gumerov, "Capture and recreation of higher order 3D sound fields via reciprocity," in *Proc. 10th Int. Conf. Auditory Display (ICAD2004)*, Sydney, Australia, Jul. 6–9, 2004.

[15] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'02)*, May 2002, vol. 2, pp. 1781–1784.

[16] B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution," *J. Acoust. Soc. Amer.*, vol. 116, no. 4, pp. 2149–2157, 2004.

[17] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 1, pp. 135–143, Jan. 2005.

[18] M. Taylor, "Cubature for the sphere and the discrete spherical harmonic transform," *SIAM J. Numer. Anal.*, vol. 32, no. 2, pp. 667–670, 1995.

[19] C. W. Ueberhuber, *Numerical Computation 2: Methods, Software, and Analysis*. Berlin, Germany: Springer-Verlag, 1997.

Zhiyun Li (S'04–M'06) received the Ph.D. degree in computer science from the University of Maryland, College Park, in 2005.

His research interests include spatial audio, microphone arrays, fast multipole methods, and computer graphics. He is currently a Senior Algorithms Engineer at Leica, San Ramon, CA.

Ramani Duraiswami (M'99) received the B.Tech. degree from the Indian Institute of Technology, Bombay, India, in 1985 and the Ph.D. degree from The Johns Hopkins University, Baltimore, MD, in 1991.

He is a Faculty Member with the Department of Computer Science and the Institute for Advanced Computer Studies, University of Maryland, College Park, where he directs research at the Perceptual Interfaces and Reality Laboratory. His current research interests are in the areas of audio for virtual reality and human computer interaction, scientific computing (with a recent intense focus on the fast multipole method), and machine learning.

Dr. Duraiswami is an Associate Editor of the *ACM Transactions on Applied Perception* and a Member of the Audio and Electroacoustics Technical Committee of the IEEE Signal Processing Society.