

PRACTICAL CONCENTRIC OPEN SPHERE CARDIOD MICROPHONE ARRAY DESIGN FOR HIGHER ORDER SOUND FIELD CAPTURE

Mark R. P. Thomas

Dolby Laboratories
San Francisco, CA 94103

ABSTRACT

The problem of higher order sound field capture with spherical microphone arrays is considered. While A-format cardioid designs are commonplace for first order capture, interest remains in the increased spatial resolution delivered by higher order arrays. Spherical arrays typically use omnidirectional microphones mounted on a rigid baffle, from which higher order spatial components are estimated by accounting for radial mode strength. This produces a design trade-off between with small arrays for spatial aliasing performance and large arrays for reduced amplification of instrument noise at low frequencies. A practical open sphere design is proposed that contains cardioid microphones mounted at multiple radii to fulfill both criteria. A design example with a two spheres of 16-channel cardioids at 42 mm and 420 mm radius produces white noise gain above unity on third order components down to 200 Hz, a decade lower than a rigid 32-channel 42 mm sphere of omnidirectional microphones.

Index Terms— Fourier acoustics, spherical microphone arrays, sound field capture.

1. INTRODUCTION

A sound field is commonly defined as a spherical harmonic expansion of modal coefficients about a notional origin, in particular in the context of spherical microphone arrays and Higher Order Ambisonics [1]. Ideally, colocated microphones with directivity pattern equalling that of the spherical harmonics would yield the sound field coefficients up to a given order, as is the case for orders 0 and 1 with the ideal B-format microphone that contains an omnidirectional and two or three coincident pressure gradient microphones¹. Owing to the difficulty of manufacturing microphones with directivity patterns beyond the first order, higher order capture typically involves relatively large numbers of omnidirectional pressure microphones near-uniformly distributed on a sphere of fixed radius [2], using an analytic model to compensate for the sound field's radial mode strength and thereby inferring the sound field at the origin. Free-space spheres of

omnidirectional microphones exhibit spectral nulls in the mode strength due to periodic zero crossings in the spherical Bessel functions; as such it has become popular to mount microphones on a rigid spherical baffle, using the scattered sound field to 'fill in' the nulls [2]. Unfortunately, embedding microphones in the surface of a rigid scatterer produces a tradeoff between high frequency aliasing performance and low frequency noise gain depending upon the choice of scatterer radius. Dual-radius open spheres [3], reconfigurable multi-radius spheres [4] and open spheres surrounding a rigid sphere [5] have been proposed to address this issue. In the case of radial cardioid microphones, which can be viewed as an acoustic mixer between pressure and pressure gradient operation, periodic zero crossings in the Bessel and Bessel derivatives occur out of phase with one another so that term fills in spectra when the other term is deficient [3].

This paper proposes a set of design criteria for a practical dual-radius cardioid sphere. Using arbitrary numbers of microphones whose distributions are derived from the minimized potential energy of charged particles on a unit sphere, it is shown that 16 and 32-microphone solutions exhibit some attractive properties. Microphone capsules are mounted on the vertices of polyhedra for acoustic transparency, with an inner sphere attached to an outer sphere with the same vertex configuration by elasticated supports. Comparisons in white noise gain are drawn between an example case with 16-channel cardioid spheres of radius 42 mm and 420 mm and a 32-channel sphere of pressure microphones on a 42 mm sphere, demonstrating the efficacy of the proposed topology.

2. PROPOSED DESIGN

Let (r, Ω) be a location in 3D space, where $\Omega = (\theta, \phi)$, $\theta = [0, \pi]$, $\phi = [0, 2\pi)$ are colatitude and azimuth angles. Given a sound field in the spherical harmonic domain $\check{S}_l^m(\omega)$, a microphone at radius r receives a signal at frequency ω [2]

$$P(r, \Omega, \omega) = \sum_{l=0}^N \sum_{m=-l}^l b_l \left(\frac{\omega}{c} r \right) \check{S}_l^m(\omega) Y_l^m(\Omega), \quad (1)$$

Email: mark.r.thomas@ieee.org

¹A similar statement can be made for the ideal A-format microphone.

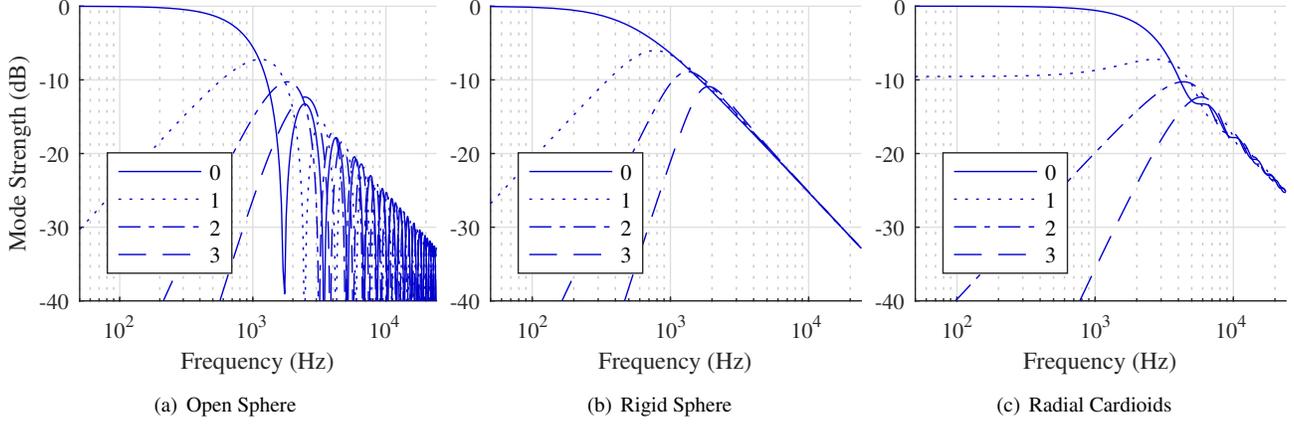


Fig. 1. Mode strengths for $r = 100$ mm, orders 0 to 3.

where l and m are degree and order and $Y_l^m(\Omega)$ are the spherical harmonics² up to order N . The mode strength $b_l(\frac{\omega}{c}r)$ is a function of radius and microphone directivity [3]

$$b_l\left(\frac{\omega}{c}r\right) = 4\pi i^l \begin{cases} j_l\left(\frac{\omega}{c}r\right) & \text{Open sphere} \\ j_l\left(\frac{\omega}{c}r\right) - \frac{j_l'\left(\frac{\omega}{c}a\right)}{h_l'\left(\frac{\omega}{c}a\right)} h_l\left(\frac{\omega}{c}r\right) & \text{Rigid sphere} \\ j_l\left(\frac{\omega}{c}r\right) - i j_l'\left(\frac{\omega}{c}r\right) & \text{Radial cardioid,} \end{cases} \quad (2)$$

where a is a rigid sphere radius, $j_l(x)$ are the spherical Bessel functions, $h_l(x)$ the spherical Hankel functions³ and $(\cdot)'$ is a derivative w.r.t. the argument. Expansion (1) may be simplified to

$$P(r, \theta, \phi, \omega) = \sum_{l=0}^N \sum_{m=-l}^l \check{S}_l^m(\omega) \Psi_l^m(r, \Omega, \omega) \quad (3)$$

by defining a new basis

$$\Psi_l^m(r, \Omega, \omega) = b_l\left(\frac{\omega}{c}r\right) Y_l^m(\Omega). \quad (4)$$

Further simplifying the notation as a vector product for P discrete microphone locations, (3) becomes

$$\mathbf{P}(\omega) = \mathbf{\Psi}(\omega) \check{\mathbf{S}}(\omega), \quad (5)$$

where $\mathbf{\Psi}(\omega)$ is a $P \times (N+1)^2$ basis matrix, $\check{\mathbf{S}}(\omega)$ an $(N+1)^2 \times 1$ vector of expansion coefficients and $\mathbf{P}(\omega)$ a $P \times 1$ vector of microphone signals. A practical solution for $\check{\mathbf{S}}(\omega)$ requires constraints to prevent excessive microphone capsule gain. In an exemplary unconstrained case,

$$\check{\mathbf{S}}(\omega) = \mathbf{\Psi}(\omega)^\dagger \mathbf{P}(\omega) = \boldsymbol{\xi}(\omega) \mathbf{P}(\omega) \quad (6)$$

is the least-squares solution where $(\cdot)^\dagger$ is a pseudo-inverse.

²The type of spherical harmonics, including complex/real, normalization conventions phase conventions, is unimportant.

³The type of Hankel function is a function of Fourier Transform convention and sign convention on direction of propagation.

Fig. 1 shows normalized mode strengths as a function of frequency for 100 mm radius open omni, rigid omni, and open cardioid configurations. The rapid low frequency mode strength attenuation, coupled with spatial aliasing that is limited by inter-microphone spacing, yields a design tradeoff in microphone radius between low frequency noise and high frequency aliasing performance for rigid designs. Fig. 1(c) shows that free-space cardioids offer two main benefits. Firstly, for a given radius they provide greater low frequency mode strength at all orders, in particular the first. Secondly, free-space operation allows the embedding of spheres at arbitrary radii that can be optimized for different frequencies. The remainder of this paper investigates the design of practical dual-radius cardioid spherical microphone arrays.

2.1. Design Criteria

In order to make a practical dual open-sphere solution, it is proposed to use a polyhedral design with which each microphone is mounted to a vertex. The array should aim to fulfill the following criteria:

1. Near-uniformity to aid the conditioning of (6) [2].
2. A small number of unique edge lengths and vertex pieces to simplify the BOM and aid manufacture.
3. Multiples of 8 microphones for compatibility with common microphone amplifiers.
4. $\geq (N+1)^2$ microphones, preferably $= (N+1)^2$ to minimize the number of microphones for order N .
5. Ease of mounting to a microphone stand.
6. Acoustic transparency in the audible frequency range.
7. Unoccluded backport for pressure-gradient operation.
8. Ability to suspend smaller spheres from larger outer spheres concentrically.

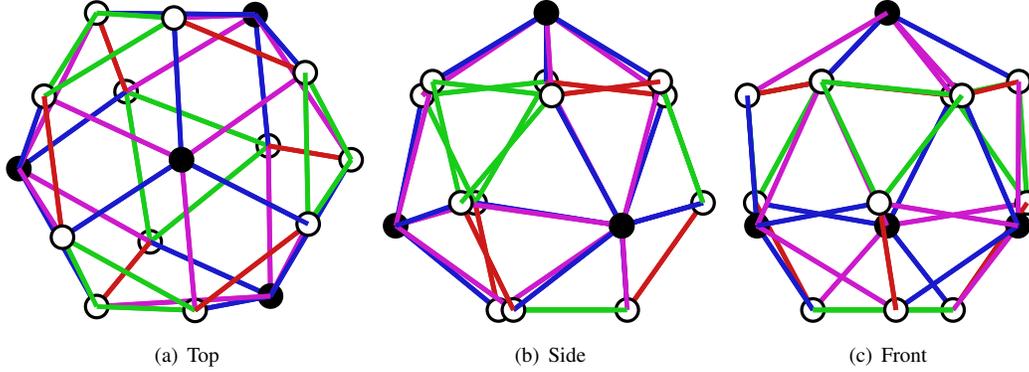


Fig. 2. Minimum energy solution. 16 vertices, 28 faces, 42 edges. Vertices ($12 \times v_5$, $4 \times v_6$) and edges ($\{0.83, 0.88, 0.93, 1.02\}r_0$) are color coded by type. The v_6 nodes (black) form a tetrahedron. A horizontal equilateral triangle (green) lies at the bottom.

2.2. Minimum Energy Distributions

Several solutions to Criterion 1 are available as standard polyhedra [6], minimum energy [7, 8], hyperinterpolation [9] and T-designs [10]. Minimum energy designs are a natural starting point for microphone distributions as they generalize to any number of nodes and are relatively straightforward to calculate. Here we explore some properties of minimum energy solutions for practical open-sphere microphone arrays.

Treating nodes as equally charged particles that mutually repel one another on the surface of a unit sphere ($r_0 = 1$), a cost function J measures total potential energy as the sum of Euclidian distances between nodes [7, 11],

$$J = \sum_{i=1}^P \sum_{j=i+1}^P \frac{1}{\|\mathbf{p}_i - \mathbf{p}_j\|_2}, \quad (7)$$

where $\mathbf{p}_i = [\theta_i \ \phi_i]^T$ are the spherical coordinates of the i th node. The optimization problem can then be formulated as

$$\begin{aligned} & \text{minimize} \quad \sum_{i=1}^P \sum_{j=i+1}^P d(\theta_i, \phi_i, \theta_j, \phi_j) \\ & \text{subject to} \quad 0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi, \end{aligned} \quad (8)$$

where

$$\begin{aligned} d(\theta_i, \phi_i, \theta_j, \phi_j) = & ((\cos(\phi_i) \sin(\theta_i) - \cos(\phi_j) \sin(\theta_j))^2 \\ & + (\sin(\phi_i) \sin(\theta_i) - \sin(\phi_j) \sin(\theta_j))^2 \\ & + (\cos(\theta_i) - \cos(\theta_j))^2)^{-1/2}. \end{aligned} \quad (9)$$

Table 1 lists all minimum energy configurations for square numbers, powers of 2 and multiples of 8 up to 64 nodes. The five rightmost columns are costs associated with Criteria 1–4. The condition number measures the invertibility of a $P \times (N + 1)^2$ matrix of spherical harmonics $Y_l^m(\Omega)$ as a ratio of largest to smallest singular values, measuring the impact of the angular configuration on (6) for order N . \mathcal{E} is

P	Poly. Name	J_{opt}	Cond. (N)	$ \mathcal{E} $	$ \mathcal{V} $	Excess
4	Tetrahedron	3.67	1 (1)	1	1	0
6	Octahedron	9.99	1 (1)	1	1	2
8	Square Antiprism	19.68	1.06 (1)	3	3	4
9	Triaugmented Triang. Prism	25.76	1.42 (2)	3	4	0
12	Icosahedron	49.17	1 (2)	1	1	3
16	–	92.91	1.58 (3)	4	2	0
24	Snub Cube	223.35	1.07 (3)	4	3	8
25	–	243.81	2.22 (4)	20	25	0
32	–	412.26	1.04 (4)	2	2	7
36	–	529.12	3.36 (5)	18	20	0
40	–	660.68	1.19 (5)	5	7	4
48	–	968.71	1.11 (5)	6	5	12
49	–	1011.56	3.32 (6)	20	49	0
56	–	1337.09	1.34 (6)	19	56	7
64	–	1765.81	6.28 (7)	13	20	0

Table 1. Properties of minimum energy sphere distributions.

the set of unique edge lengths (rounded to the nearest 1% of the radius) and \mathcal{V} the set of unique types of vertex with angles rounded to the nearest 0.1° . Excess is $P - (N + 1)^2$ for $N = \arg \max_N \{(N + 1)^2 \leq P\}$.

As expected, the uniform distributions $P = \{4, 6, 12\}$ yield unity for condition number, $|\mathcal{E}|$ and $|\mathcal{V}|$, with $P = 4$ being used extensively for A-format microphones. The icosahedron is less desirable due to its Excess cost. The $P = 25$ case, which would be ideal for harmonic order $N = 4$, is highly nonuniform with 20 unique edge lengths and 25 unique vertex pieces, making its fabrication impractical. All other cases, particularly square numbers, are similarly unattractive.

Cases $P = \{16, 32\}$ are close approximation to the hexakis truncated tetrahedron and pentakis dodecahedron respectively [6]. Both fulfill Criteria 1–3 but $P = 32$ does not fulfill Criterion 4 with an excess of 7. Only case $P = 16$ fulfills



Fig. 3. Full Assembly viewed from the front.

Criteria 1-4, yielding a maximum harmonic order $N = 3$, and was chosen for the forthcoming design.

2.3. Completing the Design

The 16-node design in Fig. 2 consists of 12 vertices joining 5 edges (v_5) and 4 vertices joining 6 edges (v_6) that lie in a tetrahedral distribution, with 42 edges and 28 faces in total. Edges and vertices are color coded to indicate similar types. The solution has been aligned so that a v_6 vertex lies on the positive z -axis. At the very bottom lies an equilateral triangle aligned parallel to the $x - y$ plane that helps fulfill Criterion 5.

Fig. 3 shows a completed assembly with an inner sphere of radius 42 mm and outer 420 mm. The design was assembled from machined carbon fiber edge rods that mate with an interference fit into sleeves on 3D printed vertices. Features were designed to be small for acoustic transparency (Criterion 6). Modular microphone ‘cages’, perforated with holes for the front and back ports, accept a 9.7mm electret cardioid microphone capsule (Criterion 7) whose cable exits through a hole in the rear of the vertex. The final Criterion is the ability to suspend spheres concentrically. An elasticated hook is mounted under tension to the edges of the lower equilateral triangle and three of the six hexagonal edges around the vertex lying on the z -axis.

3. EVALUATION

The example 42 mm and 420 mm radius 2x16-cardioid design is compared with a 32-microphone 42 mm rigid omnidirectional design. The unconstrained least-squares solution in (6) yields increasing white noise gain with increased mode strength, defined as [12]

$$\text{WNG}(\omega) = 1/\text{diag} \left\{ \boldsymbol{\xi}(\omega) \boldsymbol{\xi}^H(\omega) \right\}, \quad (10)$$

and is a measure of the gain applied to uncorrelated white noise in the capture apparatus. The curves in Fig. 4 show

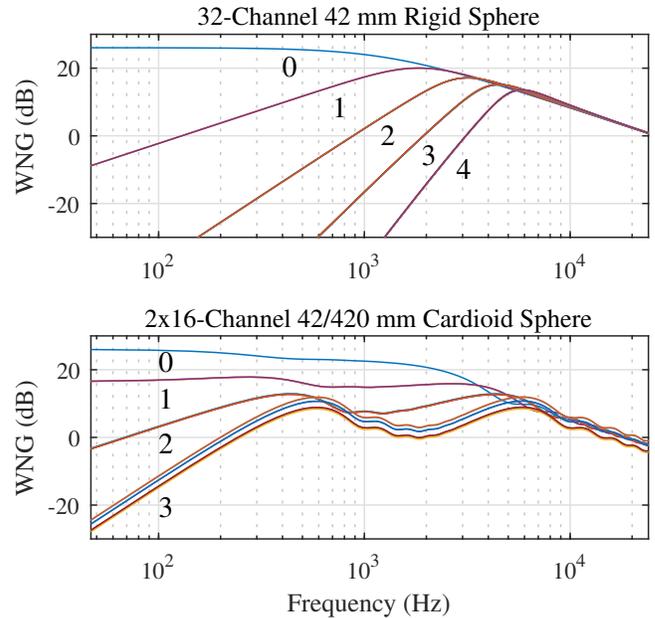


Fig. 4. WNG for rigid omni and dual-radius cardioid designs.

white noise gain above unity with the proposed design down to about 200 Hz. The effect of the outer and inner spheres can be seen as the two sets of curves intersect to produce peaks around 600 Hz and 6 kHz respectively. The 3rd order curves do not lie precisely on top of one another due to nonuniformity in the design, as reflected by the slightly elevated condition number compared to the 32-microphone distribution. The first order noise gain asymptotes to 18 dB due to the cardioid’s ability to measure 1st order modes by its sensitivity to pressure gradients. The 32-channel array carries the advantage of resolving 4th order components above ~ 3 kHz. At the third order, low frequency noise gains drop below unity below 2 kHz, and are approximately 58 dB at 200 Hz. Below these frequencies, only first order components are practical in the rigid 32-microphone case.

4. CONCLUSION

A set of design criteria was proposed for practical open sphere microphone array designs. It was shown that cardioid microphones placed on spheres of multiple radii provide improved low frequency white noise gain performance compared with the same number of omnidirectional microphones on a rigid sphere. Using near-uniform microphone distributions by minimizing the potential energy of a system of charged particles on the unit sphere, 16- and 32-point distributions were highlighted as having good practical properties. An exemplary case of a 42 mm and 420 mm dual radius open cardioid sphere shows unit WNG down to 200 Hz, compared with 2 kHz for a 32-channel 42 mm radius rigid omnidirectional configuration.

5. REFERENCES

- [1] B. Rafaely, *Fundamentals of Spherical Array Processing*, vol. 8, Springer, 2015.
- [2] B. Rafaely, “Analysis and design of spherical microphone arrays,” *IEEE/ACM Trans. on Audio, Speech and Lang. Process.*, vol. 13, no. 1, pp. 135–143, Jan. 2005.
- [3] I. Balmages and B. Rafaely, “Open-sphere design for spherical microphone arrays,” *IEEE/ACM Trans. on Audio, Speech and Lang. Process.*, vol. 15, no. 2, pp. 727–732, Feb. 2007.
- [4] R. González, J. Pearce, and T. Lokki, “Modular design for spherical microphone arrays,” in *AES International Conference on Audio for Virtual and Augmented Reality*, Aug. 2018.
- [5] C. Jin, N. Epain, and A. Parthy, “Design, optimization and evaluation of a dual-radius spherical microphone array,” *IEEE/ACM Trans. on Audio, Speech and Lang. Process.*, vol. 22, no. 1, pp. 193–204, Jan. 2014.
- [6] J. H. Conway and H. Burgiel, *The Symmetries of Things*, CRC Press, 2008.
- [7] J. Fliege and U. Maier, “A two-stage approach for computing cubature formulae on the sphere,” in *Mathematik 1937*. Fachbereich Mathematik, Universität Dortmund, 44221 Dortmund, Germany, 1996.
- [8] R. H. Frickel and B. V. Bronk, “Symmetries of configuration of charges on a sphere,” *Canadian Journal of Chemistry*, vol. 66, no. 9, pp. 2161–2165, 1988.
- [9] I. H. Sloan, “Polynomial interpolation and hyperinterpolation over general regions,” *Journal of Approximation Theory*, vol. 83, no. 2, pp. 238–254, Nov. 1995.
- [10] R. H. Hardin and N. J. A. Sloane, “McLaren’s improved snub cube and other new spherical designs in three dimensions,” *Discrete and Computational Geometry*, vol. 15, no. 4, pp. 429–441, Apr. 1996.
- [11] M. R. P. Thomas, “Fast computation of cubature formulae for the sphere,” in *Proc. Joint Workshop on Hands-Free Speech Communication and Microphone Arrays (HSCMA)*, San Francisco, CA, Mar. 2017.
- [12] A. Politis and H. Gamper, “Comparing modeled and measurement-based spherical harmonic encoding filters for spherical microphone arrays,” in *Proc. IEEE Workshop on Applications of Signal Proc. to Audio and Acoustics*, Oct. 2017.