2013 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)

Robust Source Localization and Tracking using MUSIC-Group Delay Spectrum over Spherical Arrays

Lalan Kumar, Kushagra Singhal, and Rajesh M Hegde
Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, India 208016
Email: {lalank}@iitk.ac.in

***Abstract*—In this paper, a novel method of robust source localization using MUSIC-Group delay (MUSIC-GD) spectrum computed over spherical harmonic components is described. Our earlier work on the MUSIC-GD spectrum has focused on uniform linear array (ULA) and uniform circular array (UCA) for resolving closely spaced speech sources using minimal number of sensors under reverberant environments. However, this work tries to utilize the advantages of the MUSIC-GD spectrum in a spherical harmonics framework that is computationally simple and more accurate. The MUSIC-GD spectrum for spherical harmonic components is first defined. Its advantages in high resolution DOA estimation are also discussed. Several experiments are conducted for 3-D source localization in reverberant environments and the performance of the MUSIC-GD is compared to other conventional methods. Additional experiments on source tracking are also conducted. The results obtained from the MUSIC-GD computed over spherical arrays are motivating enough to further investigate the method for multiple source tracking in reverberant environments.**

I. Introduction

Spherical microphone arrays have been a very active area of research in recent years [1], [2]. This is primarily because of the relative ease with which array processing can be performed in the spherical harmonics (SH) domain without any spatial ambiguity [3]. Due to the similarity in the formulation of various problems in spatial domain and spherical harmonics domain, the results of the spatial domain can directly be applied in the spherical harmonics domain.

One of the primary application of microphone array processing is the direction of arrival (DOA) estimation of the sources. Various DOA estimation methods have been proposed like beamforming based, maximum likelihood (ML) based [4] and subspace-based. The high resolution of subspace-based methods is due to subspace decomposition. MUltiple SIgnal Classification (MUSIC) [5] is the most popular subspace-based method due to its simplicity and high spatial resolution. The SH-MUSIC [6] utilizes spherical harmonics in conventional MUSIC for DOA estimation. However, for closely spaced sources it gives many spurious peaks, making DOA estimation challenging. Conventionally, DOA estimation utilizes the spectral magnitude of MUSIC to compute the DOA of multiple sources incident on an array of sensors. The phase information of the MUSIC spectrum has been studied in [7] for DOA estimation over a uniform linear array (ULA). In this work, we define and discuss the use of the negative differential of the unwrapped phase spectrum (group delay) of MUSIC for DOA estimation over spherical array of microphones. The primary contribution of this work is the utilization of the MUSIC-Group delay (MUSIC-GD) spectrum in spherical harmonics domain to localize closely spaced sources over spherical arrays under reverberant environments.

The rest of the paper is organized as follows. Section II presents the proposed MUSIC-GD spectrum over spherical array of microphones. The method is presented after brief discussion on spherical harmonics domain. The proposed method is evaluated in Section III. Section IV concludes the paper.

II. MUSIC-Group delay spectrum for source localization over Spherical array

Subspace-based method utilizing magnitude spectrum of MUSIC called MUSIC-Magnitude (MM) spectrum, gives large number of spurious peaks for closely spaced sources. Hence it requires comprehensive search algorithm for deciding the candidate peaks. The MM method is more prone to error under reverberant conditions. In [8], a high resolution source localization based on the MUSIC-GD spectrum over ULA has been proposed. The method is non-trivially extended for planar arrays in [9], [10]. In this work, a robust source localization method using MUSIC-GD spectrum is proposed in spherical harmonics domain. In the following Section, the spherical harmonics model for MUSIC-GD analysis is outlined.

A. The spherical harmonics model for MUSIC-GD analysis

We consider a spherical microphone array of order N with radius r and number of sensors I. A sound field of L plane-waves is incident on the array with wavenumber k. The l^{th} source location is denoted by $\mathbf{\Psi}_l = (\theta_l, \phi_l)$. The elevation angle θ is measured down from positive z axis, while the azimuthal angle ϕ is measured counterclockwise from positive x axis.

In spatial domain, the sound pressure at I microphones, $\mathbf{p}(k) = [p_1(k), p_2(k), \ldots, p_I(k)]^T$, is written as,

$$\mathbf{p}(k) = \mathbf{V}(k)\mathbf{s}(k) + \mathbf{n}(k), \tag{1}$$

where $\mathbf{V}(k)$ is $I \times L$ steering matrix, $\mathbf{s}(k)$ is $L \times 1$ vector of signal amplitudes, $\mathbf{n}(k)$ is $I \times 1$ vector of zero mean, uncorrelated sensor noise and $(.)^T$ denotes the transpose. The steering matrix $\mathbf{V}(k)$ is expressed as

$$\mathbf{V}(k) = [\mathbf{v_1}(k), \mathbf{v_2}(k), \ldots, \mathbf{v_L}(k)], \text{ where} \tag{2}$$
$$\mathbf{v}_l(k) = [e^{-j\mathbf{k}_l^T\mathbf{r}_1}, e^{-j\mathbf{k}_l^T\mathbf{r}_2}, \ldots, e^{-j\mathbf{k}_l^T\mathbf{r}_I}]^T \tag{3}$$
$$\mathbf{k}_l = -(k\sin\theta_l\cos\phi_l, k\sin\theta_l\sin\phi_l, k\cos\theta_l)^T \tag{4}$$
$$\mathbf{r}_i = (r\sin\theta_i\cos\phi_i, r\sin\theta_i\sin\phi_i, r\cos\theta_i)^T \tag{5}$$

Writing Equation 1 in spherical harmonics domain [11],

$$\mathbf{p_{nm}}(k,r) = \mathbf{B}(kr)\mathbf{Y}^H(\mathbf{\Psi})\mathbf{s}(k) + \mathbf{n_{nm}}(k) \tag{6}$$

where $\mathbf{p_{nm}}(k,r)$ is the vector of spherical Fourier co-efficients, $\mathbf{B}(kr)$ and $\mathbf{Y}(\mathbf{\Psi})$ are defined in Equation 11 and Equation 13 respectively. The spherical Fourier co-efficients are given by $\mathbf{p_{nm}}(k,r) = [p_{00}, p_{1(-1)}, p_{10}, p_{11}, \ldots, p_{NN}]^T$. Each p_{nm} is Spherical Fourier Transform (SFT) of received pressure on the surface of sphere, $p(k,r,\theta,\phi)$. The SFT is defined as [12]

$$p_{nm}(k,r) = \int_0^{2\pi}\int_0^{\pi} p(k,r,\theta,\phi)[Y_n^m(\theta,\phi)]^* \sin(\theta)d\theta d\phi, \tag{7}$$

The spherical harmonic of order n and degree m, Y_n^m is defined in Equation 14.

The pressure on the sphere is spatially sampled by the microphones. Hence, the Equation 7 can be re-written as

$$p_{nm}(k,r) \cong \sum_{i=1}^{I} a_i p(k,r,\mathbf{\Phi}_i)[Y_{nm}(\mathbf{\Phi}_i)]^* \tag{8}$$
$$\forall n \le N, -n \le m \le n$$

where a_i are the sampling weights [13], $\mathbf{\Phi}_i = (\theta_i, \phi_i)$ is the microphone angular position and $(.)^*$ denotes complex conjugate. The inverse SFT is defined as

$$p(k,r,\theta,\phi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} p_{nm}(k,r)Y_n^m(\theta,\phi), \tag{9}$$

For order limited pressure function, Equation (9) can be written as

$$p(k,r,\mathbf{\Phi}) \cong \sum_{n=0}^{N}\sum_{m=-n}^{n} p_{nm}(k,r)Y_n^m(\mathbf{\Phi}). \tag{10}$$

The $(N+1)^2 \times (N+1)^2$ matrix $\mathbf{B}(kr)$ is given by

$$\mathbf{B}(kr) = diag(b_0(kr), b_1(kr), b_1(kr), b_1(kr), \ldots, b_N(kr)) \tag{11}$$

where $b_n(kr)$ is called mode strength and for open sphere, it is given by

$$b_n(kr) = 4\pi j^n j_n(kr) \tag{12}$$

where $j_n(kr)$ is spherical Bessel function and j is the unit imaginary number. Figure 1 illustrates mode strength b_n as a function of kr and n for an open sphere. For kr=0.2, zeroth order mode amplitude is 0 dB, while the first order has amplitude -26.45 dB. Hence for order greater than kr, the mode strength b_n decreases significantly. For the order $N \ge kr$, the Equation 9 can be truncated at N. This justifies the finite order Equation 10.

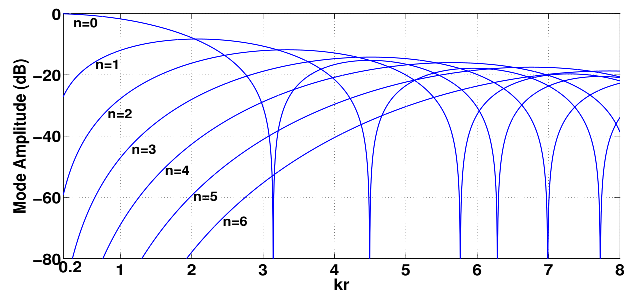


Fig. 1. Mode amplitude b_n plot for open sphere as a function of kr and n

$\mathbf{Y}(\mathbf{\Psi})$ is $L \times (N+1)^2$ steering matrix. A particular l^{th} steering vector is written as

$$\mathbf{y}_l = [Y_0^0(\mathbf{\Psi_1}), Y_1^{-1}(\mathbf{\Psi_1}), Y_1^0(\mathbf{\Psi_1}), Y_1^1(\mathbf{\Psi_1}), \ldots, Y_N^N(\mathbf{\Psi_1})] \tag{13}$$

The spherical harmonic of order n and degree m, $Y_n^m(\mathbf{\Psi})$ is given by

$$Y_n^m(\mathbf{\Psi}) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(cos\theta)e^{jm\phi} \tag{14}$$

with $\mathbf{\Psi} = (\theta, \phi)$ is direction of arrival of the plane wave. Y_n^m are solution to the Helmholtz equation [14] and P_n^m are associated Legendre function. For negative m, $Y_n^m(\mathbf{\Psi}) = (-1)^{|m|}Y_n^{|m|*}(\mathbf{\Psi})$. Figure 2 shows three spherical harmonics plot. The radius shows the magnitude and color shows the phase. It is to be noted that Y_0^0 is isotropic while Y_1^0 and Y_1^1 have directional characteristics.

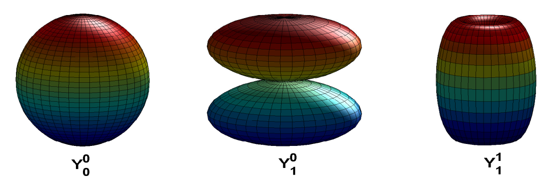


Fig. 2. Spherical harmonics plot, Y_0^0, Y_1^0, Y_1^1

Multiplying both side of Equation 6 by $\mathbf{B}^{-1}(kr)$, we get the final spherical harmonics model as follows

$$\mathbf{a_{nm}}(k) = \mathbf{Y}^H(\mathbf{\Psi})\mathbf{s}(k) + \mathbf{z_{nm}}(k), \tag{15}$$

where

$$\mathbf{z_{nm}}(k) = \mathbf{B}^{-1}(kr)\mathbf{n_{nm}}(k). \tag{16}$$

B. The MUSIC-Group delay spectrum computed for spherical harmonics components

The MUSIC-Magnitude spectrum for spherical array of microphones is given by

$$P_{MUSIC}(\mathbf{\Psi}) = \frac{1}{\mathbf{y}(\mathbf{\Psi})\mathbf{S_{a_{nm}}^{NS}}[\mathbf{S_{a_{nm}}^{NS}}]^H\mathbf{y}^H(\mathbf{\Psi})} \tag{17}$$

where $\mathbf{y}(\mathbf{\Psi})$ is a steering vector defined in Equation 13 and $\mathbf{S_{a_{nm}}^{NS}}$ is noise subspace obtained from eigenvalue decomposition of autocorrelation matrix, $\mathbf{S_{a_{nm}}} = E[\mathbf{a_{nm}}(k)\mathbf{a_{nm}}(k)^H]$.

The denominator takes zero when $\mathbf{\Psi}$ corresponds to DOA owing to orthogonality between noise eigenvector and steering vector. Hence, we get a peak in MUSIC-Magnitude spectrum. However, when sources are closely spaced, MUSIC-Magnitude spectrum is unable to resolve them accurately giving many spurious peaks. Figure 3 illustrates the MUSIC-Magnitude spectrum for an Eigenmike system [15]. The simulation was done considering open sphere for the sources at (20°,50°) and (15°,60°) with SNR = 10 dB. Frequency smoothing and whitening of noise was applied as in [11]. The MUSIC-Magnitude spectrum gives many spurious peaks for closely spaced sources and hence determining the candidate peak becomes challenging.

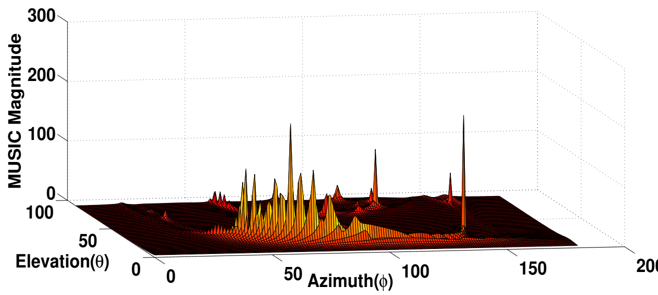


Fig. 3. MUSIC-magnitude spectrum for sources at (20°,50°) and (15°,60°), SNR=10 dB

Hence, we propose the use of group delay spectrum for resolving closely spaced sources. There is sharp change in phase spectrum of MUSIC at DOA as illustrated in [8]. Hence the differential magnitude of phase, called group delay, results in a peak at DOA. In practice abrupt changes can occur in the phase due to small variations in the signal caused by microphone calibration errors. This leads to many spurious peaks in group delay spectrum [10]. The proposed MUSIC-Group delay spectrum for spherical array of microphones is given by

$$P_{MGD}(\mathbf{\Psi}) = (\sum_{u=1}^{U} |\nabla arg(\mathbf{y}(\mathbf{\Psi}).q_u)|^2).P_{MUSIC}(\mathbf{\Psi}) \quad (18)$$

where $U = (N+1)^2 - L$, ∇ is the gradient operator, $arg(.)$ indicates unwrapped phase, and q_u represents the u^{th} eigenvector of the noise subspace, $\mathbf{S^{NS}_{a_{nm}}}$. The MUSIC-GD spectrum is product of MUSIC-Magnitude and group delay spectra. Hence the spurious peaks in MUSIC-Magnitude and group delay spectrum are removed and prominent peaks corresponding to DOAs are retained as illustrated in Figure 4. In addition, the phase spectrum follows additive property, that enables the group delay spectrum to better preserve the peak than magnitude spectrum following multiplicative property [9].

III. Performance Evaluation

Experiments on source localization and source tracking [16] are performed to evaluate the proposed method. The experiments on source localization are presented as root mean square error (RMSE) at various reverberation time T_{60} [17]. The proposed method is compared with MUSIC-Magnitude and minimum variance distortionless response (MVDR) [18].

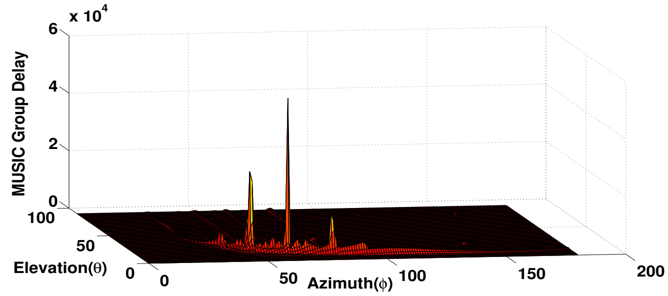


Fig. 4. MUSIC-Group delay spectrum for sources at (20°,50°) and (15°,60°), SNR=10 dB

Narrowband source tracking results are presented as two dimensional trajectory of the elevation angle for a fixed azimuth.

A. Experimental Conditions

The proposed algorithm was tested in a room with dimensions, $7.3m \times 6.2m \times 3.4m$. An Eigenmike microphone array [15] was used for the simulation. It consists of 32 microphones embedded in a rigid sphere of radius 4.2 cm. The order of the array was taken to be $N = 3$. The source localization experiments are done at various reverberation times (T_{60}). The room impulse response for spherical microphone array was generated as in [19].

B. Experiments on source localization

To evaluate the resolving power of the proposed method, two sources at $\mathbf{\Psi_1} = (30°, 60°)$ and $\mathbf{\Psi_2} = (35°, 50°)$ were considered. Localization experiments were conducted for 300 iterations at three different reverberation times, 150 ms, 200 ms and 250 ms. The experiment was repeated for three methods, MUSIC-GD, MUSIC-Magnitude and MVDR. The results are presented as RMSE values in Table I for source 1. MUSIC-GD has reasonably lower RMSE than other conventional methods proving it to be more robust.

TABLE I. Comparison of RMSE of various methods at different reverberation time, T_{60}.

Angle	Method	T_{60} (150ms)	T_{60} (200ms)	T_{60} (250ms)
θ	**MUSIC-GD**	0.6403	0.6419	0.6475
	MM	0.6688	0.8144	0.7989
	MVDR	1.1034	1.1579	1.1738
ϕ	**MUSIC-GD**	1.4387	1.4665	1.4866
	MM	1.7866	1.9127	1.6484
	MVDR	2.276	2.3481	2.4927

C. Experiments on narrowband source Tracking

To demonstrate the effectiveness of MUSIC-GD over MUSIC-Magnitude, elevation angle of a moving source is tracked in this section. The source continuously emits narrowband signal impinging on the array of microphone. The azimuthal angle (ϕ) of the source is fixed at 45° and the elevation angle is varied as the trajectory given in Figure 5. The elevation angle is tracked at fixed azimuth with MUSIC-GD and MUSIC-magnitude. Figure 6(a) illustrates the tracked trajectory by MM spectrum. It can be noted that the trajectory is zig-zag because of spurious peaks that arises in MM

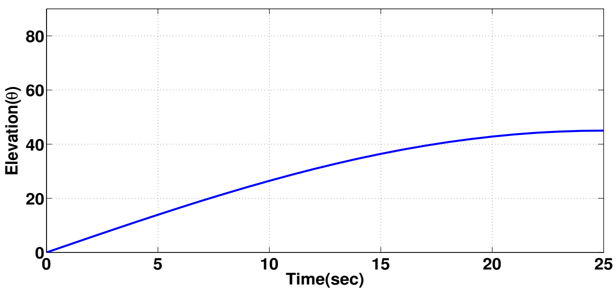
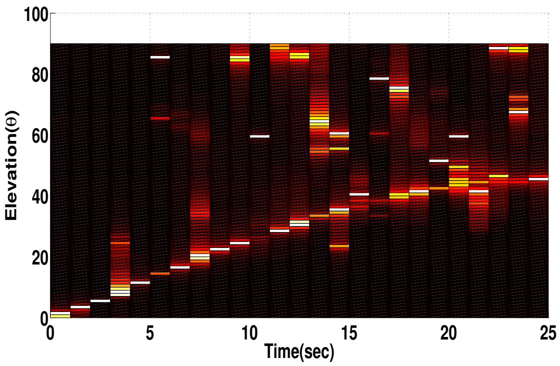
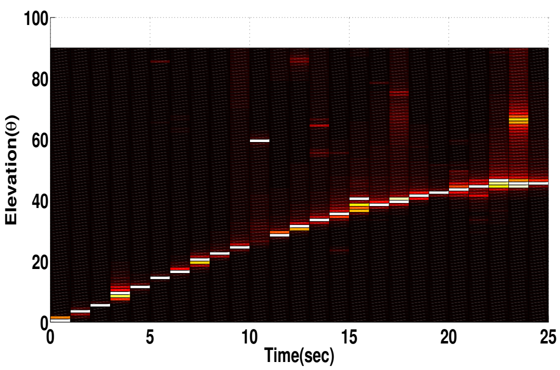


Fig. 5. Trajectory of elevation angle (θ) followed by the moving source with time for a fixed azimuth $\phi = 45°$.



(a)



(b)

Fig. 6. Tracking result for elevation (a)MUSIC-Magnitude and (b) MUSIC-Group delay. The azimuth is fixed at $45°$.

spectrum. MUSIC-GD tracking result is shown in Figure 6(b). This is very close to the actual trajectory leading to efficient tracking.

IV. Conclusion

In this paper, a high resolution source localization method for spherical array has been proposed using MUSIC-GD spectrum. Experimental results on multi-source localization show the robustness of the method. Experiments on tracking a single source are motivating enough to extend this approach to track multiple sources that are closely spaced in a kalman filter framework. The proof of additive property of group delay in the spherical harmonics domain is non trivial and is currently being developed. The question of utilizing the diffraction and scattering information provided by the spherical harmonics in the MUSIC-GD framework for reverberant environments is also being investigated.

Acknowledgment

This work was funded in part by TCS Research Scholarship Program under project number TCS/CS/20110191.

References

[1] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, vol. 2. IEEE, 2002, pp. II–1781.

[2] J. McDonough, K. Kumatani, T. Arakawa, K. Yamamoto, and B. Raj, "Speaker tracking with spherical microphone arrays," in *Acoustics Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013.

[3] I. Cohen and J. Benesty, *Speech processing in modern communication: challenges and perspectives*. Springer, 2010, vol. 3.

[4] J. Bohme, "Estimation of source parameters by maximum likelihood and nonlinear regression," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84.*, vol. 9. IEEE, 1984, pp. 271–274.

[5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation,," *IEEE Transactions on Antenna and Propagation,*, vol. AP-34, pp. 276–280, 1986.

[6] X. Li, S. Yan, X. Ma, and C. Hou, "Spherical harmonics music versus conventional music," *Applied Acoustics*, vol. 72, no. 9, pp. 646–652, 2011.

[7] K. Ichige, K. Saito, and H. Arai, "High resolution doa estimation using unwrapped phase information of music-based noise subspace," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. E91-A, pp. 1990–1999, August 2008.

[8] M. Shukla and R. M. Hegde, "Significance of the music-group delay spectrum in speech acquisition from distant microphones," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 2738–2741.

[9] A. Tripathy, L. Kumar, and R. M. Hegde, "Group delay based methods for speech source localization over circular arrays," in *Hands-free Speech Communication and Microphone Arrays (HSCMA), 2011 Joint Workshop on*. IEEE, 2011, pp. 64–69.

[10] L. Kumar, A. Tripathy, and R. M. Hegde, "Significance of the music-group delay spectrum in robust multi-source localization over planar arrays," *IEEE Trans. on Signal Processing, Under review*, 2013.

[11] D. Khaykin and B. Rafaely, "Acoustic analysis by spherical microphone array processing of room impulse responses," *The Journal of the Acoustical Society of America*, vol. 132, p. 261, 2012.

[12] J. R. Driscoll and D. M. Healy, "Computing fourier transforms and convolutions on the 2-sphere," *Advances in applied mathematics*, vol. 15, no. 2, pp. 202–250, 1994.

[13] B. Rafaely, "Analysis and design of spherical microphone arrays," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 1, pp. 135–143, 2005.

[14] E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*. Access Online via Elsevier, 1999.

[15] *The Eigenmike Microphone Array*, http://www.mhacoustics.com/.

[16] H.-C. Song and B.-w. Yoon, "Direction finding of wideband sources in sparse arrays," in *Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2002*. IEEE, 2002, pp. 518–522.

[17] P. Zahorik, in *Direct-to-reverberant energy ratio sensitivity*, vol. 112. Acoustical Society Of America, November, 2002, pp. 2110–2117.

[18] B. Rafaely, Y. Peled, M. Agmon, D. Khaykin, and E. Fisher, "Spherical microphone array beamforming," in *Speech Processing in Modern Communication*. Springer, 2010, pp. 281–305.

[19] D. P. Jarrett, E. A. Habets, M. R. Thomas, and P. A. Naylor, "Simulating room impulse responses for spherical microphone arrays," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 129–132.