SPARSITY-BASED SOUND FIELD SEPARATION IN THE SPHERICAL HARMONICS DOMAIN

Mirco Pezzoli, Maximo Cobos[†*]*, Fabio Antonacci, Augusto Sarti*

Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy
[†]Departamento de Informática, Universitat de València, Valencia, Spain

ABSTRACT

Sound field analysis and reconstruction has been a topic of intense research in the last decades for its multiple applications in spatial audio processing tasks. In this context, the identification of the direct and reverberant sound field components is a problem of great interest, where several solutions exploiting spherical harmonics representations have already been proposed. However, the available techniques demand a large number of high-order microphones (HOMs) and high computational power in order to fulfill the necessary spatial sampling requirements, which can only be reduced by prior information obtained through acoustic measurements. Inspired by compressed sensing approaches, this paper proposes an alternative sparse formulation for estimating the exterior and interior sound field components in the spherical harmonics domain that allows to reduce hardware requirements without the need for additional acoustic measurements. The results show that a considerable reduction in the number of HOMs can be achieved while improving the estimation of the sound field components.

***Index Terms*—** sound field separation, spherical harmonics, sparse representations, sound field reconstruction, compressed sensing

1. INTRODUCTION

In the sound field processing literature, the identification of the direct sound emitted by target acoustic sources and the reverberation or interferences is a well-known challenging problem. This operation has inherent application in different tasks such as dereverberation [1], source separation [2], sound field reconstruction and navigation [3] for augmented/extended reality applications. The available solutions typically rely on multichannel recordings acquired by distributed microphones, possibly organized in arrays. In general, different classes of solutions can be identified for the estimation of the direct and reverberant sound field components. The so-called parametric approaches [3–7] describe the acoustic field through a compact signal model. In fact, the goal of these techniques is to provide a flexible and parsimonious system, at the cost of introducing some approximation in the estimated sound field components. A second class of solutions, known as non-parametric [1, 8–11] rely on the solutions of the wave equation to accurately describe the acoustic field captured by the sensors. Recently, a third class emerged with deep learning approaches to sound field analysis [12–16].

Among non-parametric solutions, spherical-harmonics-based techniques [1, 10, 11, 17] provide a convenient representation of the acoustic field. As a matter of fact, when all the target sources are enclosed in a region, the emitted direct sound is represented as an exterior field. Conversely, reverberation and interferences originating outside the source region can be expressed as an interior field, by suitably choosing the reference origin [18]. Therefore, the two sound field components can be compactly represented though the spherical harmonics coefficients associated to the exterior and the interior fields. In [1], the authors adopted a set of distributed higher order microphone arrays (HOMs) for the separation of the sound field into its exterior and interior components. The exterior and interior coefficients are estimated exploiting the spherical harmonics addition theorem [17, 19], which links the "local" spherical harmonics coefficients of the HOMs with the "global" coefficients of the sound field (exterior and interior coefficients). The main limitation of the technique in [1] concerns the spatial sampling required in order to accurately estimate the sound field components of a region e.g., 676 sensors for a region of radius $1\,\mathrm{m}$ at $1\,\mathrm{kHz}$. In order to reduce the huge hardware and computational requirements of [1], in [11] the authors propose to exploit a set of measured room impulse responses (RIRs) to describe the interior field produced by the reverberation as a function of the exterior sound field. Although effective, the technique in [11] requires the acquisition of a large number of RIRs, therefore, in [10] the authors propose to augment the available measurements with synthetic free field impulse responses.

A different class of non-parametric methods adopts compressed sensing principles [8, 9] in order to represent the sound field by means of sparse components. In [9], the authors exploit the spatial sparsity of the sources in order to represent the direct sound field through a sparse dictionary of Green's functions that constitutes a grid of sources generat-

[*]This work is partially supported by ERDF and the Spanish Ministry of Science, Innovation and Universities under Grant RTI2018-097045-B-C21, as well as grants AICO/2020/154 and AEST/2020/012 from Generalitat Valenciana.

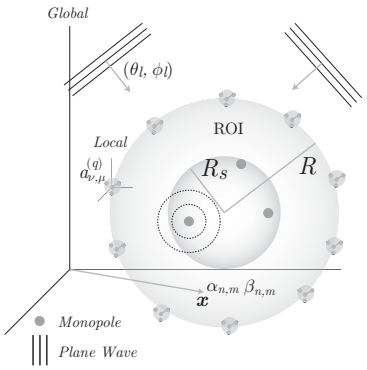

Fig. 1. Reference setup for the proposed method.

ing the sound field. Conversely, the reverberant component is modelled as a sparse distribution of plane waves with an additional *low-rank* term. Recently, in [8], this sound field model has been extended by including near-field early reflections. In particular, the authors in [8] model the early reflections in the reverberant component through image sources that augment the sparse dictionary of Green's functions. The computation of the image source terms requires prior information on the room geometry, in fact, they are modelled reflecting the grid of sources on the walls of the room.

In this paper, we propose a solution for the estimation of the exterior (direct sound) and interior (reverberation and interferences) sound field components. In particular, we exploit a sparse representation of the sound field in order to estimate the spherical harmonics sound field coefficients. Inspired by [9], we redefine the exterior sound field coefficients as a sparse dictionary of translated monopoles in the spherical harmonics domain. At the same time, the interior field is modelled through the spherical harmonics expansion of sparse plane waves. The use of sparse representations allows us to reduce the hardware requirements with respect to typical spherical-harmonics-based sound field decomposition techniques [1, 17]. In addition, differently from previous solutions [10, 11], the proposed approach does not require any additional information for the estimation of the sound field coefficients. The results show that the proposed solution is able to estimate the sound field coefficients using a reduced number of HOMs while providing more accuracy with respect to previous approaches.

2. DATA MODEL AND PROBLEM FORMULATION

Let us consider a set of Q distributed Vth order HOMs defining a region of interest ROI of radius R as depicted in Fig. 1. The HOMs surround J acoustic sources located inside a spherical region with radius R_s. Each qth HOM is composed of Q' sensors whose signals can be encoded in the spherical domain as

$$a^{(q)}_{\nu,\mu} = \frac{1}{b_\nu(kr_M)} \sum_{q'=1}^{Q'} P(k, \boldsymbol{x}_{q'}) Y^*_{\nu\mu}(\theta_{q'}, \phi_{q'}) \qquad (1)$$

where $k = 2\pi f/c$ is the wave number at frequency f and speed of sound c, $P(k, \boldsymbol{x}_{q'})$ is the sound pressure at the sensor located in $\boldsymbol{x}_{q'} = [r_M, \theta_{q'}, \phi_{q'}]^T$ (expressed with respect to the local origin of the HOM) and $Y_{\nu\mu}(\cdot)$ defines the spherical harmonic of order ν and degree μ with $\nu = 0, \dots, V$ and $\mu = -\nu, \dots, \nu$. The term $b_\nu(\cdot)$ in (1) is defined accordingly to the array type as [18]

$$b_v\left(kr_{\mathrm{HOM}}\right) = \begin{cases} j_v\left(kr_M\right) & \text{for open array} \\ j_v\left(kr_M\right) - \frac{j'_v(kr_M)}{h'_v(kr_M)} h_v\left(kr_M\right) & \text{for rigid array,} \end{cases} \qquad (2)$$

where $h_\nu(\cdot)$ and $j_\nu(\cdot)$ are the νth order spherical Hankel and Bessel functions of the first kind, respectively.

The sound pressure at the HOM sensors and, in general, at any point $\boldsymbol{x} = [r, \theta, \phi]^T$ inside the ROI can be expressed as the superposition of the exterior and interior field

$$P(\boldsymbol{x}, k) = P_E(\boldsymbol{x}, k) + P_I(\boldsymbol{x}, k), \qquad (3)$$

where $P_E(\boldsymbol{x}, k)$ is the exterior field generated from the source region and $P_I(\boldsymbol{x}, k)$ is the interior field entering the ROI and produced by reflections or other interfering sources outside the ROI. The two sound field components in (3) can be expressed through their spherical harmonics expansion as [18]

$$P_E(\boldsymbol{x}, k) = \sum_{n=0}^{N_E} \sum_{m=-n}^{n} \beta_{nm}(k) h_n(kr) Y_{nm}(\theta, \phi), \qquad (4)$$

$$P_I(\boldsymbol{x}, k) = \sum_{n=0}^{N_I} \sum_{m=-n}^{n} \alpha_{nm}(k) j_n(kr) Y_{nm}(\theta, \phi), \qquad (5)$$

where $\beta_{nm}(k)$ and $\alpha_{nm}(k)$ are the exterior and interior sound field coefficients, respectively. The order of the spherical harmonics expansion is given by the limits $N_E = \lceil keR_s/2 \rceil$ and $N_I = \lceil keR/2 \rceil$ [20]. The coefficients β in (4) and α in (5) are known as "global", in fact, by knowing their value we can completely characterize the sound field inside the ROI.

Let us collect the "local" spherical harmonics coefficients (1) of the Q HOMs in the vector $\mathbf{a} \in \mathbb{C}^{Q(V+1)^2 \times 1}$. We express the local coefficients as the superposition of the exterior and interior fields

$$\mathbf{a} = \mathbf{E}\mathbf{P}_E \boldsymbol{\beta} + \mathbf{E}\mathbf{P}_I \boldsymbol{\alpha}, \qquad (6)$$

where $\boldsymbol{\beta} \in \mathbb{C}^{(N_E+1)^2 \times 1}$ and $\boldsymbol{\alpha} \in \mathbb{C}^{(N_I+1)^2 \times 1}$ represent the vector containing the exterior and interior coefficients, respectively, $\mathbf{P}_E \in \mathbb{C}^{Q(V+1)^2 \times (N_E+1)^2}$ is the matrix defining the propagation of the exterior coefficients to the sensors (4), while $\mathbf{P}_I \in \mathbb{C}^{Q(V+1)^2 \times (N_I+1)^2}$ defines the propagation of the interior coefficients (5). The term $\mathbf{E}$ in (6) defines a complex $Q(V+1)^2 \times Q(V+1)^2$ block matrix that performs the local spherical harmonics encoding of the HOMs signals accordingly to (1). Note that dependency on the wave number k in (6) is omitted for simplicity. Our goal is the estimation of the global coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ from the HOMs in order to perform the separation of the sound field.

3. PROPOSED SPARSITY-BASED SPHERICAL HARMONICS MODEL (S-SH)

In this section, we formulate the estimation of the global sound field coefficients as the solution of a sparse optimization problem. We express the local spherical harmonics coefficients as the sum of the exterior and interior coefficients that are modelled through a sparse set of monopoles and plane waves in the spherical harmonics domain.

Following the approach in [9], we discretize the source region in a set of G points that constitute a grid of omnidirectional sources (monopoles). It follows that the exterior field coefficients β_{nm} in $\boldsymbol{\beta}$ (6) are approximated as a superposition of translated monopoles [21]

$$\beta_{nm}(k) \approx \sum_{g=1}^{G} c_{0,0}(k)\sqrt{4\pi} j_n(kr'_g) Y^*_{nm}(\theta'_g, \phi'_g), \quad (7)$$

where $c_{0,0}(k) = -ik$ is the $zero$th order coefficient of a unitary amplitude monopole, while $j_n(kr_g)Y^*_{nm}(\theta_g, \phi_g)$ expresses the translation [21] of the gth source from $\boldsymbol{x}'_g = [r'_g, \theta'_g, \phi'_g]^T$ to the reference origin.

As far as the interior field is concerned, we assume that the sound field entering the ROI can be modelled as a linear combination of L plane wave components expanded with respect to the reference origin, leading to the expression of the interior coefficients α_{nm} in $\boldsymbol{\alpha}$ (6) as

$$\alpha_{nm} \approx \sum_{l=1}^{L} \sqrt{4\pi}(-i)^n Y^*_{nm}(\theta_l, \phi_l), \quad (8)$$

where the inclination θ_l and the azimuth ϕ_l defines the direction of the lth plane wave.

Let us introduce the dictionary of monopoles derived from the G grid points translated in the origin as $\mathbf{B} \in \mathbb{C}^{(N_E+1)^2 \times G}$ whose elements are

$$[\mathbf{B}]_{\iota,g} = c_{0,0}(k)\sqrt{4\pi} j_n(kr'_g) Y^*_{nm}(\theta'_g, \phi'_g), \quad (9)$$

where $\iota = 1, \ldots, (N_E+1)^2$ is the row index for each nth order and mth degree. In practice, each column of (9) provides the expression of the exterior coefficients generated by the gth monopole. Similarly, we introduce the dictionary of the plane wave components with their spherical harmonics expansion $\mathbf{W} \in \mathbb{C}^{(N_I+1)^2 \times L}$ with elements

$$[\mathbf{W}]_{\iota,l} = \sqrt{4\pi}(-i)^n Y^*_{nm}(\theta_l, \phi_l), \quad (10)$$

where $\iota = 1, \ldots, (N_I+1)^2$ is the row index for each nth order and mth degree.

Given the signal model in (6) and the models of the exterior (7) and interior (8) coefficients, we rewrite the local coefficients as

$$\mathbf{a} = \mathbf{E}\mathbf{P}_E\mathbf{B}\mathbf{y} + \mathbf{E}\mathbf{P}_I\mathbf{W}\mathbf{u} \quad (11)$$

where $\mathbf{y} \in \mathbb{C}^{G\times 1}$ and $\mathbf{u} \in \mathbb{C}^{L\times 1}$ are the weight vectors of the translated monopoles and plane waves, respectively. Similarly to [9], the solution of the undetermined system in (11) is found through sparse optimization

$$\begin{aligned} \underset{\mathbf{y},\mathbf{u}}{\text{argmin}} \quad & \|\mathbf{y}\|_1 + \gamma\|\mathbf{u}\|_1 \\ \text{s.t.} \quad & \mathbf{a} = \mathbf{E}\mathbf{P}_E\mathbf{B}\mathbf{y} + \mathbf{E}\mathbf{P}_I\mathbf{W}\mathbf{u}, \end{aligned} \quad (12)$$

where γ is a regularization term. The solution to the optimization problem in (12) can be found using ADMM [22] as shown in [9]. In practice, the optimal weights $\mathbf{y}^\star$ and $\mathbf{u}^\star$ found solving (12) identify a limited "sparse" set of translated monopoles and plane wave components in (11) characterizing the sound field. Therefore, an estimate of the global sound field coefficients can be obtained as

$$\hat{\boldsymbol{\beta}} = \mathbf{B}\mathbf{y}^\star, \quad \hat{\boldsymbol{\alpha}} = \mathbf{W}\mathbf{u}^\star, \quad (13)$$

for the exterior and interior coefficients, respectively. The estimated coefficients can then be employed for reproducing the sound field at a target point inside the ROI using (4) and (5).

4. SIMULATION RESULTS

We evaluate the proposed technique (S-SH) through a simulation campaign in order to show the effectiveness of the sound field components estimation in different scenarios. We simulated a shoe-box room of dimensions $5\,\mathrm{m} \times 8\,\mathrm{m} \times 3\,\mathrm{m}$ with a variable number of 1st order HOMs ($V = 1$). The HOMs are placed around a ROI of $R = 1\,\mathrm{m}$ following the spiral sampling [23]. We consider $J = 2$ sources randomly placed inside a source region of radius $R_s = 0.1\,\mathrm{m}$. The grid of monopoles (9) is defined with $G = 300$ points placed inside the source region, while $L = 360$ directions sampled using [23] are employed for the plane wave components (10). The impulse responses between each source and sensor of the HOMs is computed using the RIR generator toolbox [24] with reflection order 10 and sampling frequency $8\,\mathrm{kHz}$. Gaussian white noise has been added to model a signal-to-noise ratio equal to $60\,\mathrm{dB}$.

We compare the estimation of the exterior field obtained by the proposed approach with the solution in [10], here referred to as VL. In particular, since S-SH does not require any prior measurement, we adopted [10] considering only ideal virtual loudspeakers (Green's functions) corresponding to the grid of G monopoles. As stated by the authors in [10], when no RIRs are availble VL performs comparably to [1]. It is worth noting that for the adopted ROI and frequency range (up to $4\,\mathrm{kHz}$), the techinque in [1] would require $Q > 2000$ HOMs. In addition, we compare the estimation of the sound field components provided by the sparsity-based technique in [9], referred to as SS, that differently from the optimization problem of (12) does not consider the spherical harmonics representation of the sound field. The estimated sound field is evaluated for a set of $M = 25$ test points randomly selected inside the ROI.

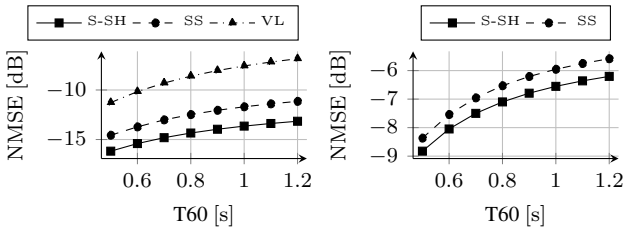

Fig. 2. Average values of NMSE at the test points as a function of the T60 for the exterior field only (a) and the full sound field (exterior and interior) (b).

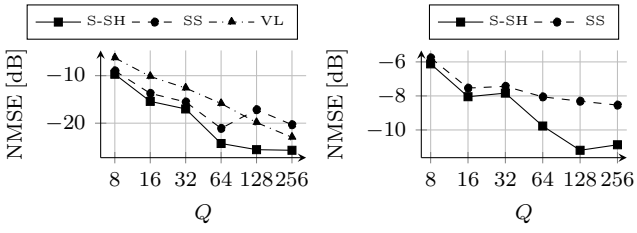

Fig. 3. Average values of NMSE at the test points as a function of the HOM number Q for the exterior field only (a) and the full sound field (exterior and interior) (b).

4.1. Metrics

In order to determine the performance of the sound field reconstruction at the test points, we adopt the NMSE defined as

$$\mathrm{NMSE}(\boldsymbol{x}_t) = 10\log_{10}\left(\frac{1}{\mathcal{T}}\frac{\sum_{\tau=1}^{\mathcal{T}}\left(\hat{p}\left(\boldsymbol{x}_t,\tau\right)-p\left(\boldsymbol{x}_t,\tau\right)\right)^2}{\sum_{\tau=1}^{\mathcal{T}}p\left(\boldsymbol{x}_t,\tau\right)^2}\right), \tag{14}$$

where $\mathcal{T}$ is the signal duration time, $p\left(\boldsymbol{x}_t,\tau\right)$ is the actual sound field at point $\boldsymbol{x}_t$ in time domain (discrete-time τ) and $\hat{p}\left(\boldsymbol{x}_t,\tau\right)$ represents its estimate.

As far as the sound field separation is concerned, we evaluate the amount of energy associated to the reverberation leaking in the exteior field estimation using the direct-to-reverberant ratio

$$\mathrm{DRR}(\boldsymbol{x}_t) = 10\log_{10}\left(\frac{1}{\mathcal{T}}\frac{\sum_{\tau=\tau_0-C}^{\tau_0+C}\hat{p}_E\left(\boldsymbol{x}_t,\tau\right)^2}{\sum_{\tau=\tau_0+C}^{\mathcal{T}}\hat{p}_E\left(\boldsymbol{x}_t,\tau\right)^2}\right), \tag{15}$$

where τ_0 is the time of arrival of the direct signal, and $C = 2.5\,\mathrm{ms}$ [10, 25]. Note that a higher value of DRR relates to a better exterior field estimate.

4.2. Discussion

As a first evaluation, we varied the T60 of the simulated room in the range from $0.5\,\mathrm{s}$ to $1.2\,\mathrm{s}$. We considered a fixed number of HOMs $Q = 16$ that corresponds to a total number of 64 channels. The main objective is to test the robustness of the sound field components estimation at different levels of reverberation. In Fig. 2, NMSE averaged for all the test points is reported for both the exterior field estimation and the complete field (exterior + interior components). Inspecting Fig. 2, we can observe that S-SH outperformed the estimation of VL. In addition, the proposed S-SH consistently provides more accurate results with respect to SS (see Fig. 2(b)). The sound field separation performance is evaluated through the DRR in Fig. 4(a). Similarly to the NMSE results, the proposed S-SH provides the highest average DRR. As expected, the DRR decreases at high T60, in fact, for higher reverberation times, the energy of the interior field component increases making the separation more challenging.

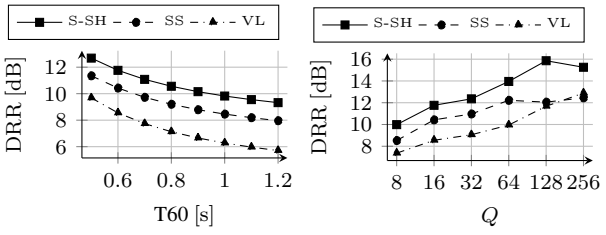

Fig. 4. Average values of DRR at the test points as a function of the T60 (a) and number of HOMs Q (b).

As a second evaluation, we analyze the sound field varying the number of HOMs from $Q = 8$ (32 capsules) up to $Q = 256$ (512 capsules), while the T60 is fixed at $0.6\,\mathrm{s}$. In Fig. 3, NMSE averaged for all the M test points is reported for both the exterior field (Fig. 3(a)) and the full sound field (Fig. 3(b)). Inspecting Fig. 3, we can note that the proposed technique consistently provides a more accurate estimation with respect to both SS and VL. In particular, S-SH shows an accurate estimation of the interior field, leading to lower NMSE in Fig. 3(b). This result can be also noted inspecting the DRR in Fig. 4(b) in which for $Q = [128, 256]$ SS performs comparably to VL. However, when more HOMs are available, the accuracy of exterior field estimation increases leading to higher DRR. Hence, for example, we can improve the separation in highly reverberant environments increasing the number of HOMs.

5. CONCLUSION

In this paper we presented a sparsity-based sound field model in the spherical harmonics domain that targets the separation of the acoustic field. Differently from previous approaches to sound field separation and reconstruction in the spherical harmonics domain, the proposed technique is able to work with a reduced number of HOMs without the need for prior measurement, showing, in addition, a greater accuracy. Moreover, the results suggest that working in the spherical harmonics domain increases the performance of the sparse model with respect to its counterpart that directly work on sensor signals. In the future, we foresee the extension of the proposed approach considering real measurements and different applications such as the separation of acoustic sources.

6. REFERENCES

[1] A. Fahim, P. N. Samarasinghe, and T. D. Abhayapala, "Sound field separation in a mixed acoustic environment using a sparse array of higher order spherical microphones," in *Hands-free Speech Commun. and Microphone Arrays*. IEEE, 2017, pp. 151–155.

[2] M. Pezzoli, J. J. Carabias-Orti, M. Cobos, F. Antonacci, and A. Sarti, "Ray-space-based multichannel nonnegative matrix factorization for audio source separation," *IEEE Signal Process. Lett.*, vol. 28, pp. 369–373, 2021.

[3] M. Pezzoli, F. Borra, F. Antonacci, S. Tubaro, and A. Sarti, "A parametric approach to virtual miking for sources of arbitrary directivity," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, pp. 2333–2348, 2020.

[4] O. Thiergart, G. Del Galdo, M. Taseska, and E. A. P. Habets, "Geometry-based spatial sound acquisition using distributed microphone arrays," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 21, no. 12, pp. 2583–2594, 2013.

[5] V. Pulkki, S. Delikaris-Manias, and A. Politis, *Parametric time-frequency domain spatial audio*. Wiley Online Library, 2018.

[6] M. Pezzoli, F. Borra, F. Antonacci, A. Sarti, and S. Tubaro, "Estimation of the sound field at arbitrary positions in distributed microphone networks based on distributed ray space transform," in *Int. Conf. Acoust. Speech Signal Process.* IEEE, 2018, pp. 186–190.

[7] M. Pezzoli, F. Borra, F. Antonacci, A. Sarti, and S. Tubaro, "Reconstruction of the virtual microphone signal based on the distributed ray space transform," in *26th Eur. Signal Process. Conf.* IEEE, 2018, pp. 1537–1541.

[8] S. Damiano, F. Borra, A. Bernardini, F. Antonacci, and A. Sarti, "Soundfield reconstruction in reverberant rooms based on compressive sensing and image-source models of early reflections," in *Workshop Appl. Signal Process. Audio Acoust.* IEEE, 2021, pp. 366–370.

[9] S. Koyama and L. Daudet, "Sparse representation of a spatial sound field in a reverberant environment," *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 1, pp. 172–184, 2019.

[10] F. Borra, S. Krenn, I. D. Gebru, and D. Marković, "1st-order microphone array system for large area sound field recording and reconstruction: Discussion and preliminary results," in *Workshop Appl. Signal Process. Audio Acoust.* IEEE, 2019, pp. 378–382.

[11] F. Borra, I. D. Gebru, and D. Markovic, "Soundfield reconstruction in reverberant environments using higher-order microphones and impulse response measurements," in *Int. Conf. Acoust. Speech Signal Process.* IEEE, 2019, pp. 281–285.

[12] M. J. Bianco, P. Gerstoft, J. Traer, E. Ozanich, M. A. Roch, S. Gannot, and C.-A. Deledalle, "Machine learning in acoustics: Theory and applications," *J. Acoust. Soc. Am.*, vol. 146, no. 5, pp. 3590–3628, 2019.

[13] M. Olivieri, R. Malvermi, M. Pezzoli, M. Zanoni, S. Gonzalez, F. Antonacci, and A. Sarti, "Audio information retrieval and musical acoustics," *IEEE Instrum. Meas. Mag.*, vol. 24, no. 7, pp. 10–20, 2021.

[14] M. Olivieri, M. Pezzoli, F. Antonacci, and A. Sarti, "Near field acoustic holography on arbitrary shapes using convolutional neural network," in *29th Eur. Signal Process. Conf.* IEEE, 2021, pp. 121–125.

[15] ——, "A physics-informed neural network approach for nearfield acoustic holography," *Sensors*, vol. 21, no. 23, 2021.

[16] F. Lluís, P. Martínez-Nuevo, M. Bo Møller, and S. Ewan Shepstone, "Sound field reconstruction in rooms: Inpainting meets super-resolution," *J. Acoust. Soc. Am.*, vol. 148, no. 2, pp. 649–659, 2020.

[17] P. N. Samarasinghe, T. D. Abhayapala, and M. A. Poletti, "3d spatial soundfield recording over large regions," in *Int. Workshop Acoust. Signal Enhanc.* VDE, 2012, pp. 1–4.

[18] E. G. Williams, *Fourier Acoustics*. London, UK: Academic Press, 1999.

[19] P. A. Martin, *Multiple scattering: interaction of time-harmonic waves with N obstacles*. Cambridge University Press, 2006, no. 107.

[20] H. M. Jones, R. A. Kennedy, and T. D. Abhayapala, "On dimensionality of multipath fields: Spatial extent and richness," in *Int. Conf. Acoust. Speech Signal Process*, vol. 3. IEEE, 2002, pp. III–2837–III–2840.

[21] I. Ben Hagai, M. Pollow, M. Vorländer, and B. Rafaely, "Acoustic centering of sources measured by surrounding spherical microphone arrays," *J. Acoust. Soc. Am.*, vol. 130, no. 4, pp. 2003–2015, 2011.

[22] S. Boyd, N. Parikh, and E. Chu, *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc, 2011.

[23] A. Semechko, "S2 sampling toolbox," https://github.com/AntonSemechko/S2-Sampling-Toolbox, 2020.

[24] E. A. P. Habets, "Room impulse response generator," *Technische Universiteit Eindhoven, Tech. Rep*, vol. 2, no. 2.4, p. 1, 2006.

[25] P. Zahorik, "Direct-to-reverberant energy ratio sensitivity," *J. Acoust. Soc. Am.*, vol. 112, no. 5, pp. 2110–2117, 2002.