SOUNDFIELD RECONSTRUCTION IN REVERBERANT ENVIRONMENTS USING HIGHER-ORDER MICROPHONES AND IMPULSE RESPONSE MEASUREMENTS

Federico Borra[*]

Politecnico di Milano
Piazza Leonardo da Vinci 32
20133 Milano, Italy

Israel Dejene Gebru, Dejan Marković

Facebook Reality Labs, Pittsburgh
4420 Bayard Street, Suite 100
Pittsburgh, PA 15213

ABSTRACT

This paper addresses the problem of soundfield reconstruction over a large area using a distributed array of higher-order microphones. Given an area enclosed by the array, one can distinguish between two components of the soundfield: the interior soundfield generated by sources outside of the enclosed area and the exterior soundfield generated by sources inside the enclosed area. These components form an indistinguishable mixture and, despite the existence of theoretical solutions to separate them, practical implementation is challenging due to high number of microphones needed for large regions and high frequencies. In this work, we consider a scenario where the interior soundfield is characterized by reverberation and show how a set of RIR measurements can be used to parametrize the interior component as a function of the exterior component, effectively reducing the unknowns of the problem.

Index Terms— Ambisonics, microphone arrays, virtual microphone, soundfield recording, spherical harmonics.

1. INTRODUCTION

The problem of estimating the soundfield at points in space that differ from the points it was measured is gaining interest in emerging applications such as e.g., navigation of captured scenes in VR [1, 2] and analysis of the spatial properties of sound sources [3, 4]. It is an open and challenging problem in spatial audio processing. In the literature this problem is referred to as acoustic holography [5], wavefield analysis [6, 7], spatial sound recording [8–10], virtual microphone synthesis [10–12], etc. In this work, we adopt the term soundfield reconstruction to describe an algorithm that uses a distributed array of microphones to estimate the soundfield in arbitrary positions of the captured acoustic scene.

The problem of soundfield reconstruction has been recently addressed using two different approaches: parametric (model based) and non-parametric. The parametric methods, such as [10–13], are based on simplified assumptions, e.g. existence of a discrete number of point-like sources, and are less demanding from the hardware perspective. While they yield perceptually plausible results, they are less suitable for analysis purposes. On the other hand, the non-parametric methods require a huge number of spatial samples (microphones). At the expense of their individual complexity, the number of microphone units required can be reduced using higher-order microphones. The higher-order microphones, commonly used for 360 VR audio, can also be employed in arrays based on the spherical harmonics addition theorem [14]. The non-parametric reconstruction algorithms that use higher-order microphones were presented in [7, 8, 15] and [9, 16] for 2D and 3D soundfields, respectively.

In particular, [16] discusses the case of exterior/interior soundfield separation in the spherical harmonics domain [5]. The exterior field is defined as the soundfield outside a sphere that encompasses all the sources that generate it, and the interior field is defined as the soundfield inside a sphere that does not contain any sources. In a mixed field scenario the overall soundfield is a superposition of these two components. The authors in [16] proposed a separation/reconstruction algorithm that jointly estimates the exterior and interior soundfield coefficients. The number of spherical harmonic coefficients needed to describe the two components is proportional to the radius of the respective spheres. In comparison to the exterior field, estimating the interior field is more challenging due to hardware and computational requirements. A similar approach to the one presented in [16], but based on differential microphones has been presented in [17].

The work presented in this paper is inspired by [16] and considers a scenario in which the acoustic scene of interest is surrounded by an array of higher-order microphones with no independent sources outside. Despite being less general w.r.t to [16], this is a common configuration for both capture and analysis purposes. With only reverberation as a source, the interior soundfield can be considered a function of the exterior soundfield and the environment, whose response to acoustic events can be probed by measuring the room impulse response (RIR) for a set of points. In this paper, we show how the RIR measurements can be used to reduce the separation problem to the estimation of exterior coefficients only. Moreover, although the exterior soundfield reconstruction is enough for applications such as analysis of sound sources or capture of acoustic scenes without the environmental fingerprint, in some cases it may be desirable to estimate the reverberation component as well. We therefore propose an approach to estimate the interior field for points where RIR measurements are available.

The rest of the paper is organized as follows. Section 2 formulates the problem in the spherical harmonics domain. In Section 3 we describe how RIR measurements can be used to estimate both the exterior and interior components of the soundfield. Simulation results and comparisons with the algorithm in [16] are presented in Section 4. Moreover, we also present some preliminary real-world results obtained with 40 first-order microphones enclosing an area of radius 2.74 m.

2. PROBLEM FORMULATION

Let us consider two concentric spheres with the center in the origin of global reference system, and radii R_0 and R_s, where $R_0 > R_s$.

[*]Work done during an internship at Facebook Reality Labs, Pittsburgh.

Let all the sound sources be confined within the smaller sphere of radius R_s, and the only additional sound source is the room i.e. reverberation. Let Q units of Vth-order microphones be placed along the radius R_0. Since the image (wall-mirrored) sources causing reverberation are found outside of sphere R_0, the area between R_s and R_0 is source-free. Our goal is to reconstruct the soundfield in this region, i.e. estimate the signal at any point $\boldsymbol{x} = (r, \theta, \phi)$, $R_s < r < R_0$.

Let $k = 2\pi f/c$ be the wave number, f being the frequency and c indicating the speed of sound. The acoustic signal at point $\boldsymbol{x}$ is the superposition of exterior and interior soundfields [5], i.e. $p(\boldsymbol{x}, k) = p_E(\boldsymbol{x}, k) + p_I(\boldsymbol{x}, k)$, with

$$\begin{aligned} p_E(\boldsymbol{x}, k) &= \sum_{n=0}^{N_E} \sum_{m=-n}^{n} \beta_{nm}(k) h_n(kr) Y_{nm}(\theta, \phi) \,, \\ p_I(\boldsymbol{x}, k) &= \sum_{n=0}^{N_I} \sum_{m=-n}^{n} \alpha_{nm}(k) j_n(kr) Y_{nm}(\theta, \phi) \,, \end{aligned} \tag{1}$$

where $p_E(\boldsymbol{x}, k)$ and $p_I(\boldsymbol{x}, k)$ are, respectively, the exterior and interior field components, $Y_{nm}(\theta, \phi)$ represents the spherical harmonic of order n and degree m, $h_n(\cdot)$ and $j_n(\cdot)$ are, respectively, n-th order spherical Hankel and Bessel functions, $\beta_{nm}(k)$ and $\alpha_{nm}(k)$ denote the exterior and interior soundfield coefficients, respectively.

Let $\alpha_{\nu\mu}^{(q)}(k)$ indicate the soundfield coefficients captured by the qth microphone relative to the microphone's local coordinate system which is located at $\boldsymbol{x}_q = (r_q, \theta_q, \phi_q)$. The relationship between the local coefficients $\alpha_{\nu\mu}^{(q)}(k)$ and the global coefficients $\beta_{nm}(k)$ and $\alpha_{nm}(k)$ in (1) is given by the addition theorem [14, 16]

$$\begin{aligned} \alpha_{\nu\mu}^{(q)}(k) = &\sum_{n=0}^{N_E} \sum_{m=-n}^{n} \beta_{nm}(k) S_{n\nu}^{m\mu}(\boldsymbol{x}_q, k, h_{(\cdot)}) \\ &+ \sum_{n=0}^{N_I} \sum_{m=-n}^{n} \alpha_{nm}(k) S_{n\nu}^{m\mu}(\boldsymbol{x}_q, k, j_{(\cdot)}) \,, \end{aligned} \tag{2}$$

where

$$\begin{aligned} S_{n\nu}^{m\mu}(\boldsymbol{x}_q, k, f_{(\cdot)}) = 4\pi i^{\nu-n} \sum_{l=0}^{n+\nu} i^l (-1)^{2m-\mu} f_l(kr_q) Y_{l(\mu-m)}^*(\theta_q, \phi_q) \\ \sqrt{(2n+1)(2\nu+1)(2l+1)/4\pi} W_1 W_2 \,, \end{aligned} \tag{3}$$

encodes the global to local coefficient translation, i is the imaginary unit, * denotes the complex conjugate, and W_1 and W_2 are Wigner 3-j symbols defined as in [14].

The separation algorithm proposed in [16] forms a linear system using (2) for all local coefficients $\alpha_{\nu\mu}^{(q)}(k)$, $1 \le \nu \le V$, $-\nu \le \mu \le \nu$, and all microphones $1 \le q \le Q$. Solution of the inverse problem gives the estimates of the global coefficients $\hat{\beta}_{nm}(k)$ and $\hat{\alpha}_{nm}(k)$. By inserting these estimates in (1), one can reconstruct the soundfield at any arbitrary point $\boldsymbol{x}$ in the region of interest.

It is worth noting that the actual setup does not need to be concentric and the microphones are not required to sample a sphere. In particular, the microphone placement influences the conditioning number of the problem [7]. Nonetheless, the simplified setup comes in handy for theoretical analysis purposes. In practice, the summations in (1) are truncated at $N_E = \lceil keR_s/2 \rceil$ and $N_I = \lceil keR_0/2 \rceil$ [9], where e is Euler's number and $\lceil \cdot \rceil$ denotes the ceiling operator. It is straightforward to see that the order needed to accurately describe the soundfield components is directly proportional to the radius and frequency. On the other hand, the maximum frequency that can be reconstructed without aliasing artifact is inversely proportional to the radius [9], i.e. $f_E = \frac{c[\sqrt{Q}(V+1)-1]}{\pi e R_s}$ and $f_I = \frac{c[\sqrt{Q}(V+1)-1]}{\pi e R_0}$. Particularly, the complexity of the exterior field depends on the radius R_s and the sources confined within. However, one can reduce R_s with a suitable choice of origin [3]. The interior field instead depends on R_0, i.e. the dimension of the capture area, $R_0 \ge R_s$, thus the estimation of its soundfield coefficients is more demanding in both hardware and computational requirements.

3. SOUNDFIELD RECONSTRUCTION

We consider a scenario in which the interior field is due to the effect of reverberation; hence it can be expressed as a function of the exterior field and the environment. In theory, given detailed knowledge of geometry and acoustic properties of the environment, one should be able to compute the reverberation component given the external soundfield. In practice, the environment's response to acoustic events can be evaluated by measuring the RIRs [18,19], e.g. by moving a speaker in the capture area. One can think of this process as an assembly of a virtual loudspeaker array whose environment responses are known for all the microphones.

Assuming we have RIR measurements for a set of virtual loudspeakers, we propose to tackle the soundfield reconstruction problem in five steps: **(I)** use the virtual speakers to reproduce a desired (unknown) exterior field; **(II)** apply the known RIRs to the virtual speakers to obtain the microphone signals containing both soundfield components; **(III)** invert the problem and compute the exterior field using the microphone signals. Once the exterior field is estimated, by using the RIR source/receiver duality, we can: **(IV)** use microphones to "reproduce" the exterior field; and **(V)** compute the interior field in the given points by taking into account the RIRs. The following sections describe each step in detail.

3.1. Exterior soundfield

Step I In order to describe the proposed methodology for the estimation of the exterior soundfield, we start by outlining the forward problem. Let us assume that the exterior soundfield coefficients are known. We can use the virtual loudspeakers to reproduce this soundfield. Although a high number of speakers may be needed for correct reproduction, the array is purely virtual (obtained moving a single speaker) and, as long as the environment characteristics remain stationary over time, the RIR measurements do not need to be repeated. Let $\boldsymbol{x}_l$ be the position of the lth virtual loudspeaker, $l = 1, \ldots, L$. The relation between the virtual loudspeaker weights and the exterior soundfield coefficients is described by [20]

$$\sum_{l=1}^{L} \sum_{\hat{n}=0}^{\hat{N}} \sum_{\hat{m}=-\hat{n}}^{\hat{n}} w_{\hat{n}\hat{m}}^{(l)}(k) S_{\hat{n}n}^{\hat{m}m}(\boldsymbol{x}_l, k, j_{(\cdot)}) = \beta_{nm}(k) \,, \tag{4}$$

where $\hat{N}$ is the maximum loudspeaker order, $w_{\hat{n}\hat{m}}^{(l)}(k)$ is the lth loudspeaker weight, $S_{\hat{n}n}^{\hat{m}m}(\cdot)$ is defined in (3) and $\beta_{nm}(k)$ are the desired field coefficients in (1). Given all the coefficients $\boldsymbol{\beta}(k) = [\beta_{00}(k), \ldots, \beta_{N_E N_E}(k)]^T$, we can write (4) in matrix form as $\mathbf{J}(k)\mathbf{w}(k) = \boldsymbol{\beta}(k)$. Then the speaker weights $\mathbf{w}(k) = [w_{00}^{(1)}(k), \ldots, w_{\hat{N}\hat{N}}^{(L)}(k)]^T$ can be computed as

$$\mathbf{w}(k) = \mathbf{J}^\dagger(k)\boldsymbol{\beta}(k), \tag{5}$$

where $(\cdot)^\dagger$ denotes the matrix pseudo-inverse. The expressions in (4) and (5) are general. However, in practice the zero-order loudspeakers are commonly used for RIR measurements; hence $\hat{N} = 0$.

Step II The relationship between the lth zero-order loudspeaker weights, $w_{00}^{(l)}(k)$, and the soundfield coefficients acquired by qth microphone at order ν and mode μ, $\tilde{\alpha}_{\nu\mu}^{(q)}(k)$, can be expressed using the transfer function $\gamma_{\nu\mu}^{(l,q)}(k)$, obtained from the RIR measurement, as [21–23]

$$\tilde{\alpha}_{\nu\mu}^{(q)}(k) = \sum_{l=1}^{L} w_{00}^{(l)}(k) \frac{\gamma_{\nu\mu}^{(l,q)}(k)}{ik} \,, \tag{6}$$

where $\tilde{\alpha}_{\nu\mu}^{(q)}(k)$ is the real equivalent of complex coefficient $\alpha_{\nu\mu}^{(q)}(k)$ in (2) [24]. Using all Q Vth-order microphones, we can write (6) in matrix form as $\tilde{\boldsymbol{\alpha}}(k) = \boldsymbol{\Gamma}(k)\mathbf{w}(k)$. Using the speaker weights computed in (5), the forward problem is formulated as

$$\tilde{\boldsymbol{\alpha}}(k) = \boldsymbol{\Gamma}(k)\mathbf{J}^{\dagger}(k)\boldsymbol{\beta}(k) \,. \tag{7}$$

This expression relates the signals $\tilde{\boldsymbol{\alpha}}(k) = [\tilde{\alpha}_{00}^{(1)}(k), \ldots, \tilde{\alpha}_{VV}^{(Q)}(k)]^T$ acquired by the higher-order microphones with the exterior soundfield coefficients $\boldsymbol{\beta}(k)$.

Step III From (7), solving the inverse problem, the exterior soundfield coefficients can be estimated as

$$\hat{\boldsymbol{\beta}}(k) = \left[\boldsymbol{\Gamma}(k)\mathbf{J}^{\dagger}(k)\right]^{\dagger} \tilde{\boldsymbol{\alpha}}(k) \,. \tag{8}$$

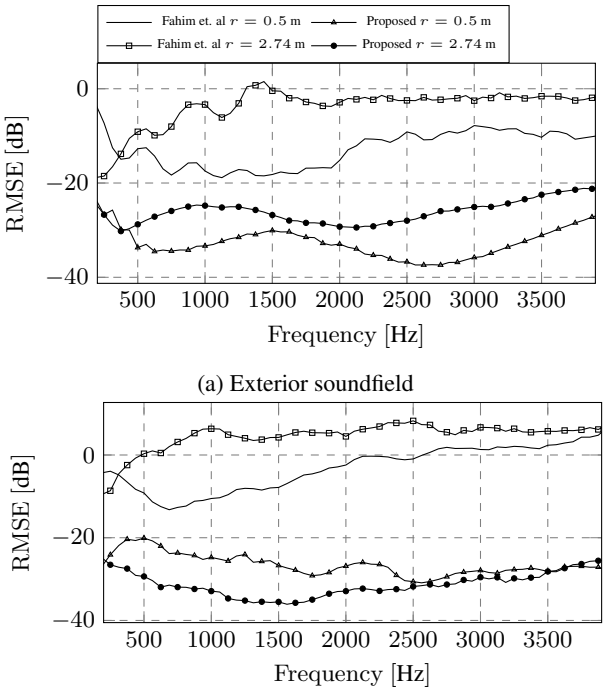
If $\boldsymbol{\Gamma}(k)$ is full column rank and $\mathbf{J}^{\dagger}(k)$ is full row rank, the expression in (8) can be simplified as $\hat{\boldsymbol{\beta}}(k) = \mathbf{J}(k)\boldsymbol{\Gamma}^{\dagger}(k)\tilde{\boldsymbol{\alpha}}(k)$. One can interpret this solution as dereverberation and a projection of the microphone coefficients into the speaker weights, followed by the translation of the weights to form global soundfield coefficients.

3.2. Interior soundfield

Given the exterior field estimate from (8) and using the duality between speakers and microphones, one can use Q Vth-order virtual speakers (microphones) to "reproduce" the exterior field outside of R_0, and the environment would respond by producing the desired interior field inside of R_0. However, in practice, we are only able to estimate the interior field in a discrete number of control points, i.e. the points in which the RIR measurements are available (RIRs can be extrapolated in other points using [21, 22]). Furthermore, since the number of microphones is limited, the frequency range in which we can correctly "reproduce" the exterior soundfield component without aliasing artifacts is restricted. To tackle this problem, we propose to estimate the interior field individually for each control point by requiring the correct "reproduction" only of a part of exterior field that contributes to that point. This is done as follows.

Step IV Let $\boldsymbol{x}_c = (r_c, \theta_c, \phi_c)$ be the position of the cth control point, $c = 1, \ldots, C$, and $\boldsymbol{x}_{cc'} = (r_{cc'}, \theta_{cc'}, \phi_{cc'})$, $c' = 1, \ldots, C'(c)$, the positions of its image sources [25] found through the RIR analysis, e.g. [26]. To estimate the interior field at $\boldsymbol{x}_c$, our strategy is to find a set of microphones weights $\eta_{\nu\mu}^{(c,q)}$ that correctly "reproduce" the exterior field at set of points $\boldsymbol{x}_{cc'}$ (reverberation paths) and cancel the soundfield at the control point $\boldsymbol{x}_c$ (direct path), i.e.

$$\begin{aligned}
&\sum_{q=1}^{Q}\sum_{\nu=0}^{V}\sum_{\mu=-\nu}^{\nu} \eta_{\nu\mu}^{(c,q)}(k) r_{\nu\mu}(\boldsymbol{x}_{cc'}|\boldsymbol{x}_q) G(k, \boldsymbol{x}_{cc'}|\boldsymbol{x}_q) = \\
&\quad = \sum_{n=0}^{N_E}\sum_{m=-n}^{n} \beta_{nm}(k) h_n(kr_{cc'}) Y_{nm}(\theta_{cc'}, \phi_{cc'}) \;,\; \forall c' \,, \\
&\sum_{q=1}^{Q}\sum_{\nu=0}^{V}\sum_{\mu=-\nu}^{\nu} \eta_{\nu\mu}^{(c,q)}(k) r_{\nu\mu}(\boldsymbol{x}_c|\boldsymbol{x}_q) G(k, \boldsymbol{x}_c|\boldsymbol{x}_q) = 0 \,,
\end{aligned} \tag{9}$$



(a) Exterior soundfield

(b) Interior soundfield

Fig. 1: Comparison in terms of RMSE of the method presented in [16] and the proposed method for both interior and exterior soundfields.

where $r_{\nu\mu}(\cdot)$ is the radiance pattern of the Vth-order microphone and $G(\cdot)$ is the well-known free-field Green's function [5]. In matrix form this writes $\mathbf{G}(k)\boldsymbol{\eta}^{(c)}(k) = \mathbf{D}(k)\boldsymbol{\beta}(k)$. Given the value of $\hat{\boldsymbol{\beta}}(k)$ estimated in (8), we can compute the weights $\boldsymbol{\eta}^{(c)}(k) = [\eta_{00}^{(c,1)}, \ldots, \eta_{VV}^{(c,Q)}]^T$ as

$$\hat{\boldsymbol{\eta}}^{(c)}(k) = \mathbf{G}^{\dagger}(k)\mathbf{D}(k)\hat{\boldsymbol{\beta}}(k) \,. \tag{10}$$

Step V Finally, given the transfer functions $\psi_{\nu\mu}^{(c,q)}(k)$ between the cth control point and the qth microphone, we can estimate the interior soundfield at the control point as

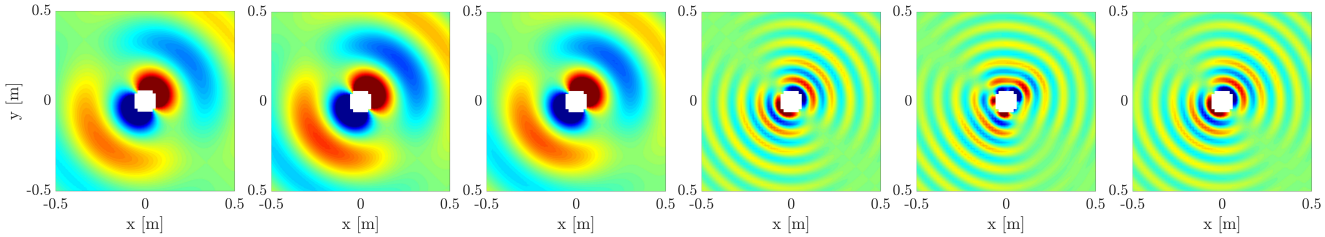
$$p_I(\boldsymbol{x}_c, k) = \sum_{q=1}^{Q}\sum_{\nu=0}^{V}\sum_{\mu=-\nu}^{\nu} \eta_{\nu\mu}^{(c,q)}(k)\psi_{\nu\mu}^{(c,q)}(k) \,. \tag{11}$$

4. EXPERIMENTS

This section describes simulations and real-world experiments performed in order to validate the proposed methodology.

4.1. Simulations

We simulated a distributed array of 1st-order microphones with $Q = 150$ elements. In particular, the 1st-order microphones are chosen since they are the most commercially available higher-order microphones. We simulated two scenarios. In the first, we placed the microphones on a surface of a sphere of radius $R_0 = 0.5\,\text{m}$ which is then placed at the center of $2{\times}2{\times}2\,\text{m}$ shoe-box room. For the second scenario, we set $R_0 = 2.74\,\text{m}$ and a shoe-box room of size $6{\times}6{\times}6\,\text{m}$. We simulated reverberation with the image source



(a) Desired at 750 Hz (b) Fahim et. al at 750 Hz (c) Proposed at 750 Hz (d) Desired at 3.4 kHz (e) Fahim et. al at 3.4 kHz (f) Proposed at 3.4 kHz

Fig. 2: The ground truth and estimated external soundfields on the horizontal plane with $z = 0$ (slice passing by the array center) for frequencies 750 Hz and 3.4 kHz.

method [25] using the implementation in [27]. Reflections are computed up to the second order assuming perfectly reflective walls. For the interior field estimate we assumed to know the positions of the image sources, which in a real-world scenario can only be estimated (using for example [26]). In both scenarios we randomly placed 50 virtual loudspeakers inside a sphere of radius 0.12 m.

We compare the proposed method with the approach proposed in [16] using a frequency-dependent root-mean-square error (RMSE) defined as follows

$$\mathrm{RMSE}(f) = \frac{\sum_{\forall \boldsymbol{x}} |p_{(\cdot)}(\boldsymbol{x}, f) - \hat{p}_{(\cdot)}(\boldsymbol{x}, f)|^2}{\sum_{\forall \boldsymbol{x}} |p_{(\cdot)}(\boldsymbol{x}, f)|^2}, \qquad (12)$$

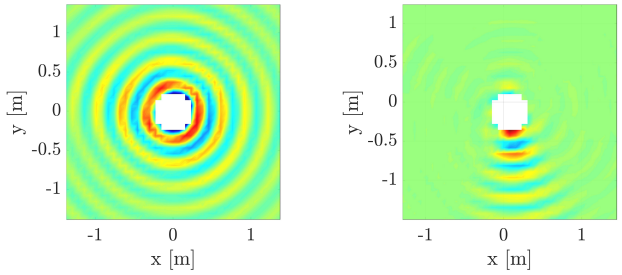
where $p_{(\cdot)}$ and $\hat{p}_{(\cdot)}$ represent, respectively, the ground truth and estimated soundfields and, depending on the evaluation, they will be properly substituted with exterior or interior components.

In the first test, we simulated an omnidirectional point source at $[0.03, 0.02, 0]$ m. Figure 1a reports the RMSE results for the exterior soundfield, and Figure 1b shows the results for the interior soundfield. In both cases the proposed method achieve better resuts. This is not surprising since the method in [16] estimates coefficients of both components simultaneously and, given the higher number of unknowns, it is less robust in the highly reverberant scenario under consideration. Furthermore, we can observe degradation around $\sim$1.8 kHz for $R_0 = 0.5$ m array and around $\sim$346 Hz for $R_0 = 2.74$ m array, which are the corresponding aliasing frequencies f_I for the interior field. On the other hand, the performance of the proposed method remains satisfactory across all frequencies since the aliasing is governed by the exterior field and the distribution of the virtual loudspeakers. However, we stress that the proposed method estimates the interior soundfield for each control point individually.

Figure 2 shows snapshots of the exterior soundfield generated by two omnidirectional point sources at $\pm\,[0.01, 0.01, 0]$ m, for the array configuration with $R_0 = 0.5$ m. In particular, snapshots for 750 Hz and 3.4 kHz are shown. As expected, both [16] and the proposed method perform well at the lower frequency. However, at the higher frequency the performance of [16] degrades while the soundfield estimated using the proposed method still strongly resembles the desired one. Finally, it is worth noting that the method is equally applicable to directional and volumetric sources, and point-like sources were used to simplify simulation of room acoustics.

4.2. Real-world Recording

We tested the reconstruction algorithm using 40 Core Sound TetraMics distributed along a radius $R_0 = 2.74$ m with a limited vertical variation. The Figures 3a and 3b show time snapshots of the exterior field estimates in the horizontal plane for an omnidirectional (B&K



(a) (b)

Fig. 3: Reconstruction of the exterior soundfield generated by an omnidirectional loudspeaker (a) and a directional loudspeaker (b).

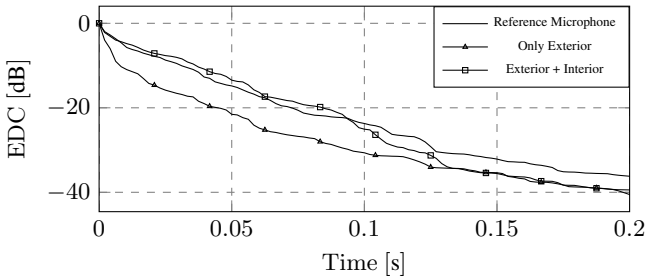


Fig. 4: Comparison in terms of Energy Decay Curve (EDC).

OmniSource Type 4295) and directional (Holosonics Audio Spotlight AS-16i) loudspeakers, respectively, both emitting a linear sine sweep. Finally, we compared the results of the reconstruction w.r.t a reference signal captured by an omnidirectional microphone present in the scene. In particular, we used the Energy Decay Curve (EDC) as a comparison metric. The EDC measures the energy decrease in the impulse response [28]. Figure 4 reports the EDC for the exterior field estimate and the superposition of the exterior and interior estimates relative to the scene with the omnidirectional speaker.

5. CONCLUSION

In this work we formulated a solution for soundfield reconstruction inside reverberant environments and showed its feasibility through simulations and real-world experiments. Although the first experiments have shown promising results, they are still limited and a lot of effort will be needed to make the method more reliable and robust for real-world capture of more complex acoustic scenes. However, to the best of the authors knowledge these experiments represent the first time a reconstruction algorithm of this kind has been employed in practice for such a large area and a non-anechoic environment.

6. REFERENCES

[1] J. G. Tylka and E. Choueiri, "Comparison of techniques for binaural navigation of higher-order ambisonic soundfields," in *Audio Engineering Society Convention 139*, Oct 2015.

[2] J. G. Tylka and E. Choueiri, "Soundfield navigation using an array of higher-order ambisonics microphones," in *Audio Engineering Society Conference: 2016 AES International Conference on Audio for Virtual and Augmented Reality*, Sep 2016.

[3] B. Rafaely, "Spatial alignment of acoustic sources based on spherical harmonics radiation analysis," in *Communications, Control and Signal Processing (ISCCSP), 2010 4th International Symposium on*. IEEE, 2010, pp. 1–5.

[4] A. Canclini, L. Mucci, F. Antonacci, A. Sarti, and S. Tubaro, "Estimation of the radiation pattern of a violin during the performance using plenacoustic methods," in *Proc. AES 138th Conv.*, 2015.

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

[6] A. Kuntz and R. Rabenstein, "Cardioid pattern optimization for a virtual circular microphone array," in *Proc. EAA Joint Symposium on Auralization and Ambisonics*, Espoo, FI, June 15-17 2009.

[7] P. N. Samarasinghe, T. D. Abhayapala, and M. Poletti, "Wavefield analysis over large areas using distributed higher order microphones," *IEEE/ACM Trans. Audio, Speech, Language Processing*, vol. 22, no. 3, pp. 647–658, 2014.

[8] P. N. Samarasinghe, T. D. Abhayapala, and M. Poletti, "Spatial soundfield recording over a large area using distributed higher order microphones," in *Proc. IEEE Workshop on Applications of Signal Process. to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, Oct. 16-19 2011.

[9] P. N. Samarasinghe and T. D. Abhayapala, "3D spatial soundfield recording over large regions," in *Proc. Int. Workshop on Acoustic Signal Enhancement (IWAENC)*, Aachen, DE, Sept. 4-6 2012.

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

[11] 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 *2018 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2018, Calgary, AB, Canada, April 15-20*, 2018, pp. 186–190.

[12] 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 *Proc. 26th European Signal Processing Conf. (EUSIPCO)*, 2018.

[13] Y. Takida, S. Koyama, and H. Saruwataril, "Exterior and interior sound field separation using convex optimization: Comparison of signal models," in *2018 26th European Signal Processing Conference (EUSIPCO)*, Sep. 2018, pp. 2549–2553.

[14] P. A. Martin, *Multiple Scattering: Interaction of Time-Harmonic Waves with N Obstacles*, Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2006.

[15] A. Fahim, P. N. Samarasinghe, and T. D. Abhayapala, "Extraction of exterior field from a mixed sound field for 2D height-invariant sound propagation," *2016 IEEE International Workshop on Acoustic Signal Enhancement IWAENC*, pp. 1–5, 2016.

[16] 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 *2017 Hands-free Speech Communications and Microphone Arrays (HSCMA)*, March 2017, pp. 151–155.

[17] H. Zuo, P. N. Samarasinghe, and T. D. Abhayapala, "Exterior - interior 3D sound field separation using a planar array of differential microphones," in *Proc. IEEE International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2018.

[18] A. Farina, A. Amendola, A. Capra, and C. Varani, "Spatial analysis of room impulse responses captured with a 32-capsules microphone array," in *Proc. AES 130th Conv.*, 2011.

[19] A. Farina, "Advancements in impulse response measurements by sine sweeps," in *Audio Engineering Society Convention 122*. Audio Engineering Society, 2007.

[20] P. N. Samarasinghe, M. A. Poletti, S. M. A. Salehin, T. D. Abhayapala, and F. M. Fazi, "3D soundfield reproduction using higher order loudspeakers," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, 2013.

[21] P. N. Samarasinghe, T. D. Abhayapala, and M. A. Poletti, "Synthesis of room transfer function over a region of space by multiple measurements using a higher-order directional microphone," in *2014 IEEE China Summit International Conference on Signal and Information Processing (ChinaSIP)*, July 2014, pp. 6–10.

[22] M. Poletti P. N. Samarasinghe, T. Abhayapala and T. Betlehem, "An efficient parameterization of the room transfer function," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 12, pp. 2217–2227, Dec 2015.

[23] P. N. Samarasinghe and T. D. Abhayapala, "Room transfer function measurement from a directional loudspeaker," in *2016 IEEE International Workshop on Acoustic Signal Enhancement (IWAENC)*, Sept 2016, pp. 1–5.

[24] M. Poletti, "Unified description of ambisonics using real and complex spherical harmonics," in *Proc. Ambisonics Symposium*, 2009.

[25] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.

[26] Y. E. Baba, A. Walther, and E. A. P. Habets, "3D room geometry inference based on room impulse response stacks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 5, pp. 857–872, May 2018.

[27] E. A. P. Habets, "RIR generator, version 2.1.20141124," Nov. 2014.

[28] "ISO 3382-3:2012, acoustics – measurement of room acoustic parameters," Standard, International Organization for Standardization, 2012.