DOI: 10.61782/fa.2025.0815

FORUM ACUSTICUM EURONOISE 2025

# SPATIALLY EXTENDED RECONSTRUCTION OF ROOM IMPULSE RESPONSES USING A GENERALIZABLE WAVE MODEL

**Efren Fernandez-Grande**[1] **Antonio Figueroa-Duran**[2] **Samuel A. Verburg**[2]
[1] Universidad Politecnica de Madrid (UPM), Madrid, Spain
[2] Technical University of Denmark (DTU), Kgs. Lyngby, Denmark

## ABSTRACT

In this study we present an experimental method to capture sound fields over extended spatial apertures based on sparse measurements with an array of microphones. The methodology builds upon recent approaches based on elementary wave models and kernel methods, that account for the general spatio-temporal properties of sound fields in enclosed spaces. The model observes the general structure of reflections and reflection density, the second-order statistics of the late reverberation field, and other properties such as decay rates, frequency dependency of reflections, etc. These properties are incorporated in the model through fitting the measured data, i.e., the wave coefficients are estimated to fit the observations of the sound field. We examine the prediction ability in real rooms, with a particular interest in auditoria and concert halls.

**Keywords:** *Sound field Analysis, Microphone arrays, Room impulse response, Sound field reconstruction, Signal Processing*

## 1. INTRODUCTION

Sound field reconstruction methods enable the estimation of the sound field at positions that have not been directly measured. This capability is essential in various applications, including active noise control [1–4], audio reproduction technology [5–8], and architectural acoustics [9–11]. In this study we examine a novel approach to reconstruct the sound field over large spatial apertures, such as an entire audience area in a concert hall, using a very limited number of measurements. Specifically, we explore the use of a single compact spherical microphone array to estimate the sound field across an extended volume of space. The key lies in incorporating meaningful and generalizable properties of the sound field in a large room into the wave model, in order to successfully predict spatial variations using very limited measurement data and faraway from the measurement aperture [12–14]. To achieve this, the proposed method employs elementary wave models and statistical processing methods (kernel-based techniques) to account for the fundamental spatiotemporal properties of enclosed sound fields. By incorporating these generalizable sound field properties, which are commonly found in any room, the model can extrapolate beyond the specific data measured and lead to a reconstruction of the sound field in the homogeneous source-free domain. The model used in this study is based on Ref. [14].

**Corresponding author*: *efren.fernandez@upm.es.*


## 2. METHODOLOGY

A wave propagation model is formulated, which can be fitted from measurements with an array of microphones. The analytical model is expressed as [12]

$$p(t,\mathbf{r}) = \sum_{d=1}^{D} Q_d(t) * \frac{\delta(t - \|\mathbf{r}-\mathbf{r}_d\|/c)}{4\pi\|\mathbf{r}-\mathbf{r}_d\|} + \sum_{i=1}^{I}\sum_{s=1}^{S} A_{s,i}(t) * \frac{\delta(t - \|\mathbf{r}-\mathbf{r}_{s,i}\|/c)}{4\pi\|\mathbf{r}-\mathbf{r}_{s,i}\|} + \sum_{j=1}^{J} B_j(t) * \delta\left(t - \langle \mathbf{r},\mathbf{r}_j\rangle/c\right). \tag{1}$$

The left hand side of Eq. (1) corresponds to the sound pressure in the domain, measured with a compact array of microphones. The first two terms on the right hand side of Eq. (1) represent the direct sound and early reflections, which are modeled by a cluster of point sources (index $d$ for the direct sound and $s$ for the cluster of each reflection $i$). The third term consists of a plane wave expansion that models the remaining part of the response, namely the less energetic components that are not modeled by the other two terms, as well as the late reverberation . This last term runs throughout the entire response , overlapping in time with the early reflections and extending until the end of the response. The objective of the problem is to estimate the unknown coefficients $Q_d$, $\Lambda_{s,i}$ and $B_j$ to fit the array measurements. This model then can be employed to predict the sound field at other positions that were not measured, far from the array. Further details of this type of model can found in Refs. [12, 14, 15].

The key of the proposed model is to account for generalizable properties of reverberant fields, specifically: 1) the general time structure of the RIR, comprised of the direct sound, followed by early reflections and late reverberation, 2) the variable nature of the different wavefronts reflected by scatterers and diffusers, and 3) the ensemble spatial statistics of the field, where the acoustic pressure is assumed to be an ergodic, normally-distributed stochastic process with spatially uniform energy density. Additionally, the spatial correlation of the field is assumed to exhibit a *sinc*-like behavior, in accordance to the random wave theory [16–18].

The methodology introduced in this work is an extension of [14] for single-array measurements. From the observations, $p(t, \mathbf{r})$ in Eq. (1), a regularized inversion is used to estimate the model coefficients $\{Q_d, A_{s,i}, B_j\}$. First, the method models the direct sound and salient reflections individually. A salient reflection is defined as a peak above an energy threshold $\zeta_{\mathrm{dB}}$ defined in relation (below) the level of the direct sound. The model iteratively identifies the peaks, localizes its direction-of-arrival (DOA) via SRP-PHAT [19], estimates its range based on the propagation time (time-of-flight), and estimates the coefficients $Q_d$ and $A_{s,i}$ that fit the observations. After each reflection estimate, the model subtracts the reconstructed pressure from the measured data, obtaining a residual signal that contains the less energetic components (cf. third term in Eq. (1)). The residual coefficients are obtained via kernel ridge regression [20], which assumes that the sound field can be described as a linear combination of correlation (kernel) functions [21]. Since the spatial correlation is modeled via the *sinc* function, often referred to as the Bessel kernel, there exists an inherent energetic decay when extrapolating far from the measurement array. To circumvent this energy decay away from the measurements, additional synthesized pressure points are distributed along the reconstruction domain that serve as additional virtual measurements. These points follow the observed statistics of the field, preserving a homogeneous energy density even at large distances and the defined spatial correlation among them.

## 3. RESULTS

The model was tested experimentally in a historical Opera Hall to demonstrate the effectiveness of the proposed reconstruction method in real-world scenarios. The hall is a horseshoe-shaped opera hall of 1400 seats including 18 rows on the stalls (approx. 500 seats) and four levels of balconies. A dodecahedral source was placed in the center of the stage, and measurements were performed with a compact 32-channel spherical array of 4.2 cm radius (Eigenmike EM32). The array was positioned in the central area of the stalls/parterre, in row 7, 2.2 m to the left of the center axis, and at distance 16 m from the source. Regarding implementation parameters, the reflection threshold was set to $\zeta_{\mathrm{dB}} = -30\,\mathrm{dB}$, and wavefronts were isolated using a Tukey window of 128 samples and cosine fraction of 0.5. The coefficients for individual wavefronts, $\{Q_d, A_{s,i}\}$, were iteratively obtained using least squares regularization, and the kernel ridge regression problem is solved via Tikhonov regularization, with a regularization parameter of 30 dB. The kernel coefficients are estimated using time-frequency analysis to capture the temporal energy decay, with a FFT window size of 4096 samples.

Figure 1 shows the reconstruction of the sound field over a large spatial aperture of 8 m length, which spans the width of the audience area in the stalls. The figure shows the reconstructed room impulse responses in a space-time representation, where the $x$-axis represents the spatial coordinate of the reconstruction along the line, and the $y$-axis of the figure shows time, covering the first 110 ms of the response - with zero time corresponding to the onset of the source. The position of the microphone array (4.2 cm radius) is indicated by the dashed line at 4.5 m. The reconstruction in Fig. 1 shows first the direct sound arriving at approximately 47 ms at the array. The wave front is clearly reconstructed and the curvature corresponds to the approximately spherical wavefront generated, with a slight tilt towards a side, as the array is not in the cen-

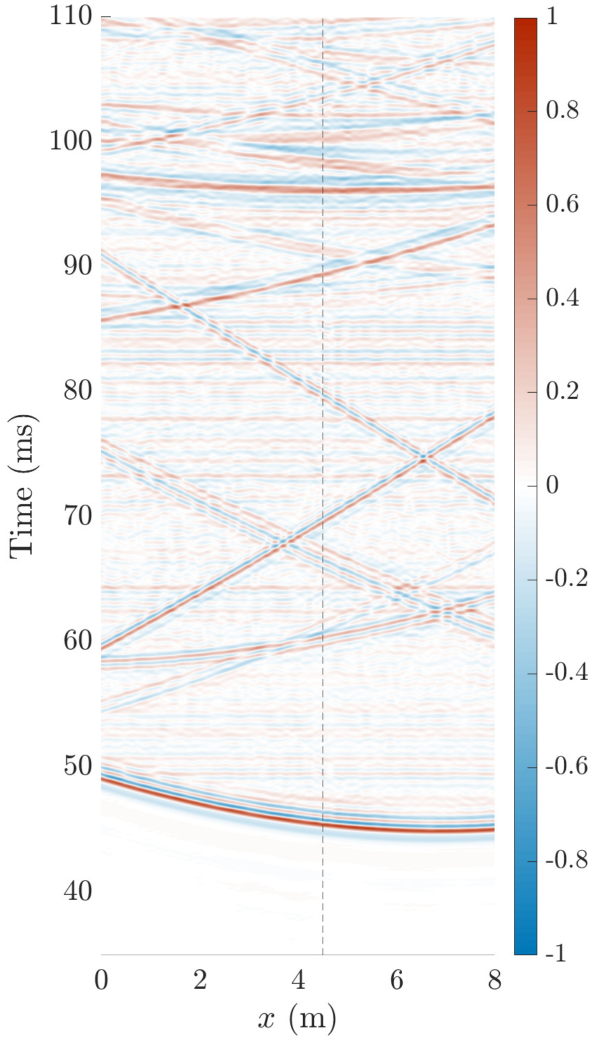


**Figure 1**. Spatio-temporal visualization of the RIR reconstruction from a single spherical microphone array of radius 4.2 cm. The dashed line indicates the position of the array.

tral axis of the stalls, but displaced 2.2 m to the side. Soon after the direct sound arrival, a few lateral reflections are present - the earliest ones are one from each side of the proscenium, where two large reflection walls are present. Each of them is soon followed by another reflection, more oblique than the previous, which are from the sidewalls of the hall (audience area). Soon after that, the reflections arriving at approximately 90 m are reflections from the top edges of the proscenium, shortly followed by reflections from the face of the first balcony and the ceiling (96 ms), and afterwards from the back-wall of the stalls area. The results were compared with measurements on site performed along the same row, and a good agreement is found, showing that the methodology can recover successfully the directions of the incoming reflections. It is also noticeable in these results that the smaller components of the sound field due to scattering diffraction and less energetic reflections have a marked directionality along the $x$-axis, which is not physical. These are a result of the statistical modeling, as relatively few experimental observations are present, and the spatial structure is biased by the temporal one. This can be easily overcome by introducing a few additional observations, or by modeling explicitly the distribution of the sound pressure.

## 4. CONCLUSIONS

The study examines a novel approach to sound field reconstruction using a generalizable wave model. The effectiveness of the method is demonstrated through experimental validation in a historical opera hall. The results show that exploiting a single measurement with a compact microphone array, enables the reconstruction of room impulse responses over extended spatial regions. We present a reconstruction based on measurements with a 4.2 cm radius spherical array, and a reconstruction over a 8 m aperture, spanning the entire audience area. The outcomes of the study are potentially significant for application in room acoustics analysis, spatial audio rendering, and real-time acoustic environment modeling.

## 5. ACKNOWLEDGMENTS

This work has been supported by the Beatriz Galindo Senior Fellowship under grant BG22-00177 and VILLUM Foundation under the grant 19179. The authors thank Anders Christian Gade, Xenofon Karakonstantis and Henrik Hvidberg, who also participated in the experiments, and HBK for lending material for the experiments.

## 6. REFERENCES


[1] P. A. Nelson, "Active control of acoustic fields and the reproduction of sound," *J Sound and Vib*, vol. 177, no. 4, pp. 447–477, 1994.

[2] D. Caviedes-Nozal, F. M. Heuchel, J. Brunskog, N. A. B. Riis, and E. Fernandez-Grande, "A Bayesian spherical harmonics source radiation model for sound field control," *J Acoust Soc Am*, vol. 146, no. 5, pp. 3425–3435, 2019.

[3] S. Koyama, J. Brunnström, H. Ito, N. Ueno, and H. Saruwatari, "Spatial Active Noise Control Based on Kernel Interpolation of Sound Field," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 29, pp. 3052–3063, 2021.

[4] F. M. Heuchel, D. Caviedes-Nozal, J. Brunskog, F. T. Agerkvist, and E. Fernandez-Grande, "Large-scale outdoor sound field control," *J Acoust Soc Am*, vol. 148, pp. 2392–2402, 2020.

[5] K. Müller and F. Zotter, "Auralization based on multi-perspective ambisonic room impulse responses," *Act Acust*, vol. 4, no. 6, 2020.

[6] H. Zuo, L. I. Birnie, P. N. Samarasinghe, T. D. Abhayapala, and V. Tourbabin, "Particle-Velocity-Based Mixed-Source Sound Field Translation for Binaural Reproduction," *Appl. Sci.*, vol. 13, no. 11, 2023.

[7] F. M. Heuchel, E. Fernandez-Grande, F. T. Agerkvist, and E. Shabalina, "Active room compensation for sound reinforcement using sound field separation techniques," *J Acoust Soc Am*, vol. 143, no. 3, pp. 1346–1354, 2018.

[8] M. B. Møller, J. K. Nielsen, E. Fernandez-Grande, and S. K. Olesen, "On the influence of transfer function noise on sound zone control in a room," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 9, pp. 1405–1418, 2019.

[9] I. B. Witew, M. Vorländer, and N. Xiang, "Sampling the sound field in auditoria using large natural-scale array measurements," *J Acoust Soc Am*, vol. 141, no. 3, pp. EL300–EL306, 2017.

[10] M. Nolan, S. A. Verburg, J. Brunskog, and E. Fernandez-Grande, "Experimental characterization of the sound field in a reverberation room," *J Acoust Soc Am*, vol. 145, no. 4, pp. 2237–2246, 2019.

[11] E. Fernandez-Grande, X. Karakonstantis, D. Caviedes-Nozal, and P. Gerstoft, "Generative models for sound field reconstruction," *J Acoust Soc Am*, vol. 153, no. 2, pp. 1179–1190, 2023.

[12] E. Fernandez-Grande, D. Caviedes-Nozal, M. Hahmann, X. Karakonstantis, and S. A. Verburg, "Reconstruction of room impulse responses over extended domains for navigable sound field reproduction," in *I3DA*, 2021.

[13] D. Caviedes-Nozal and E. Fernandez-Grande, "Spatio-Temporal Bayesian Regression for Room Impulse Response Reconstruction With Spherical Waves," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 31, pp. 3263–3277, 2023.

[14] A. Figueroa-Duran and E. Fernandez-Grande, "Reconstruction of reverberant sound fields over large spatial domains," *J Acoust Soc Am*, vol. 157, no. 1, pp. 180–190, 2025.

[15] A. Figueroa-Duran and E. Fernandez-Grande, "Navigable Reconstruction of Reverberant Sound Fields Using Distributed Microphone Arrays," in *I3DA*, 2023.

[16] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thompson, "Measurement of Correlation Coefficients in Reverberant Sound Fields," *J Acoust Soc Am*, vol. 27, no. 6, pp. 1072–1077, 1955.

[17] F. Jacobsen and A. Rodríguez Molares, "The ensemble variance of pure-tone measurements in reverberation rooms," *J Acoust Soc Am*, vol. 127, no. 1, pp. 233–237, 2010.

[18] R. Badeau, "Common mathematical framework for stochastic reverberation models," *J Acoust Soc Am*, vol. 145, no. 4, pp. 2733–2745, 2019.

[19] M. Cobos, A. Marti, and J. J. Lopez, "A Modified SRP-PHAT Functional for Robust Real-Time Sound Source Localization With Scalable Spatial Sampling," *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 71–74, 2010.

[20] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[21] N. Aronszajn, "Theory of reproducing kernels," *Transactions of the American Mathematical Society*, vol. 68, no. 3, pp. 337–404, 1950.