

NOISE FIELD CONTROL USING ACTIVE SOUND PROPAGATION AND OPTIMIZATION

Zhenyu Tang

University of North Carolina-Chapel Hill
Department of Computer Science

Dinesh Manocha*

University of Maryland
Department of Computer Science
Department of Electrical & Computer Engineering

ABSTRACT

We present a novel algorithm that uses active loudspeakers to reduce the noise level in an acoustic environment. We utilize a state-of-the-art sound propagation technique to compute a room's impulse responses accurately and then use it to perform acoustic optimization. We minimize the overall noise level in a target region by solving the driving signals of active loudspeakers, which interfere with and cancel out the noise from a linear system. Our algorithm is evaluated on complex indoor benchmarks and shows an overall noise reduction of up to 30.0dB with noise frequency up to 8kHz.

Index Terms— Active noise control, sound propagation

1. INTRODUCTION

Excessive acoustic noise in human living environments has long been a major concern for human health [1]. Noise can be generated by various sources that are common in daily life: electric fans, HVACs, engines, etc. To deal with the noise challenge, different measures have been proposed in material science, acoustic engineering, and computer aided design (CAD). Active noise control (ANC) systems [2, 3, 4] are more effective at dealing with a wide range of noise, but they also require more sophisticated instrumentation for digital devices. With modern advances in digital signal processing (DSP) software and hardware, ANC has been successfully applied to several venues [5].

Despite their success in some scenarios, earlier ANC systems are constrained by certain factors, including the number and distance of noise sources, reverberation time, physical contact with the user, etc. However, recent developments in modeling acoustic environment information can facilitate even more effective noise control with fewer constraints. For example, time difference of arrival (TDOA) based techniques can be used to robustly determine both source and receiver locations [6]. Acoustic classification and optimization methods can be combined to estimate real world acoustic material properties automatically [7], along with 3D model reconstruc-

tion using RGB-D cameras [8]. Using this environmental information, it is possible to use accurate sound simulation as part of the ANC pipeline.

Main Results: We present a novel algorithm to control the noise field in a target region actively using sound propagation and optimization. Given a static environment with known non-stationary noise sources, our approach dynamically optimizes the driving signals for a set of active loudspeakers to control the noise field effectively. Our formulation is general and makes no assumptions about the complexity of the noise field and does not constrain loudspeaker placement.

2. RELATED WORK

2.1. Active Noise Control

ANC refers to methods that introduce new sound sources in the environment to reduce the original noise. ANC relies on the principle of superposition for sound waves. The basic system for ANC requires a reference microphone to record the noise digitally and feed the signal to a loudspeaker to generate sound that has equal amplitudes but opposite phases to that of the recorded one [9]. However, when the noise is non-stationary, adaptive methods [3] are used to compensate for time-varying errors. Conventional implementation of ANC tries to match sound pressure in the frequency domain [10], while the recent wave domain approach [11] performs ANC with approximate room modes. Depending on the location of the reference microphones, ANC can be categorized into *feedforward* control, where the reference noise is recorded before the noise propagates past our active loudspeakers, and *feedback* control, where the combined noise and control signal are sensed at the control region [4]. In this paper we adopt a feedforward scheme and the frequency domain formulation.

2.2. Sound Propagation

When modeling acoustic environments, it is crucial to account for the propagation of sound and noise waves, through which the original signal experiences attenuation and delay. The simplest models, which only consider direct sound in the free field, tend to ignore the reflection, reverberation,

*Thanks to ARO grants W911NF14-1-0437 and W911NF-18-1-0313, NSF grant 1320644, and Intel for funding.

and diffraction of sound, which are important real-world phenomena. Prior methods for modeling sound propagation can be classified into either *wave-based* or *geometric* techniques. Wave-based methods solve the acoustic wave equation directly [12] using finite element methods [13], boundary element methods [14], finite difference time domain methods [15], or adaptive rectangular decomposition methods [16]. These methods provide the most accurate solutions to wave equations, but suffer from an expensive computation load, becoming especially slow for large scenes and high frequency (e.g. $> 500\text{Hz}$) waves. In contrast, geometric methods regard sound waves as geometric rays, which have much lower computational complexities than wave-based methods. Common geometric methods include image source methods [17], ray-tracing methods [18], and beam/frustrum tracing methods [19]. In this paper, we utilize the state-of-the-art geometric propagation technique, as we are dealing with noise frequencies as high as 8kHz , beyond which human ears become less sensitive to noise [20].

3. NOISE FIELD CONTROL

Table 1: Notation and symbols used throughout the paper.

| | |
|--|--|
| $P(\mathbf{x}, t)$ | Sound pressure at point \mathbf{x} at time t |
| $S(\mathbf{x}, \omega)$ | Sound pressure at point \mathbf{x} of frequency ω |
| $D(\mathbf{x}, \omega)$ | Driving signal at point \mathbf{x} of frequency ω |
| $\text{IR}(\mathbf{x}_s, \mathbf{x}_r, t)$ | Impulse response at time t from point source at location \mathbf{x}_s to the listener at location \mathbf{x}_r |
| $G(\mathbf{x}_s, \mathbf{x}_r, \omega)$ | Frequency response at ω from point source at location \mathbf{x}_s to the listener at location \mathbf{x}_r |

In this section, we present the noise field control problem in its specific form and as a linear system. As in conventional active noise cancellation methods [9], we use loudspeakers to emit destructive sound signals to cancel out the existing noise field in a target region. The notation used in the paper is summarized in Table 1.

3.1. Modeling Sound Pressure Field

Consider a sound receiver at location $\mathbf{x}_r \in V$ and a sound source placed at $\mathbf{x}_s \in V$, where V represents the 3D space. The sound pressure from \mathbf{x}_s will undergo both phase and strength changes when propagating to \mathbf{x}_r . Suppose the temporal sound pressure sequence at source location \mathbf{x}_s is $P(\mathbf{x}_s, t)$; we need to study what the induced sound pressure $P(\mathbf{x}_r, t)$ at each receiver location \mathbf{x}_r will be. A conventional way to model the transfer of acoustic waves is using homogeneous and in-homogeneous Helmholtz equations. In free-field conditions, the transfer function can be described analytically with the free-field Green’s function. However, the free-field condition does not hold when significant reverberations occur in the environment. A proposed sound field

decomposition method based on the assumption of sparse and low-rank source signals has been proven to work better than traditional models [21]. In our work, however, we make no assumptions about the structure or characteristics of the unknown sound field.

We use impulse responses (IRs) to model the sound propagation. Let $\text{IR}(\mathbf{x}_s, \mathbf{x}_r, t)$ denote the time-varying IR between a source-receiver pair \mathbf{x}_s and \mathbf{x}_r , then it follows that

$$P(\mathbf{x}_r, t) = P(\mathbf{x}_s, t) * \text{IR}(\mathbf{x}_s, \mathbf{x}_r, t), \quad (1)$$

where $(*)$ means convolution between two sequences. By taking fast Fourier transform (FFT) of Equation (1) we obtain its complex frequency domain representation:

$$S(\mathbf{x}_r, \omega) = D(\mathbf{x}_s, \omega)G(\mathbf{x}_s, \mathbf{x}_r, \omega), \quad (2)$$

which applies to all angular frequencies $\omega = 2\pi f$. We also use the notation of $D(\mathbf{x}_s, \omega)$, which is the driving signal of the sound source at \mathbf{x}_s , to differentiate between signals of sources and receivers. When dealing with multiple sound sources, the resulting sound pressure at \mathbf{x}_r is the superposition of propagated sound from all sources:

$$S(\mathbf{x}_r, \omega) = \sum_s D(\mathbf{x}_s, \omega)G(\mathbf{x}_s, \mathbf{x}_r, \omega). \quad (3)$$

3.2. Our Optimization Algorithm

Our goal is to control the noise field in a local target region $V \subset \mathbb{R}^3$. Assume that there are N_S known noise sources at $\mathbf{x}_s \notin V, s \in \{1, \dots, N_S\}$ in the environment. To negate the noise, we add N_L active loudspeakers at $\mathbf{y}_l \notin V, l \in \{1, \dots, N_L\}$. Then we need to sample N_M monitor points $\mathbf{p}_m \in V, m \in \{1, \dots, N_M\}$ from the target region for optimization. As in Equation (3), the corresponding noise field consists of the two types of sources:

$$S(\mathbf{p}_m) = \sum_{s=1}^{N_S} D(\mathbf{x}_s)G(\mathbf{x}_s, \mathbf{p}_m) + \sum_{l=1}^{N_L} D(\mathbf{y}_l)G(\mathbf{y}_l, \mathbf{p}_m). \quad (4)$$

We omit ω for brevity. We also assume that the signal $D(\mathbf{x}_s)$ at each noise source is known by close measurement.

Since we want to eliminate noise as much as possible, the desired noise field would be $S(\mathbf{p}_m) = 0$ for all monitor points in Equation (4). To formulate our linear system, we define

$$\begin{aligned} \mathbf{g}_S(\mathbf{p}_m) &= [G(\mathbf{x}_1, \mathbf{p}_m), \dots, G(\mathbf{x}_{N_S}, \mathbf{p}_m)]^T, \\ \mathbf{D}(\mathbf{x}) &= [D(\mathbf{x}_1), \dots, D(\mathbf{x}_{N_S})]^T; \end{aligned} \quad (5)$$

for noise sources. Similarly, for active loudspeakers the system is:

$$\begin{aligned} \mathbf{g}_L(\mathbf{p}_m) &= [G(\mathbf{y}_1, \mathbf{p}_m), \dots, G(\mathbf{y}_{N_L}, \mathbf{p}_m)]^T, \\ \mathbf{D}(\mathbf{y}) &= [D(\mathbf{y}_1), \dots, D(\mathbf{y}_{N_L})]^T. \end{aligned} \quad (6)$$

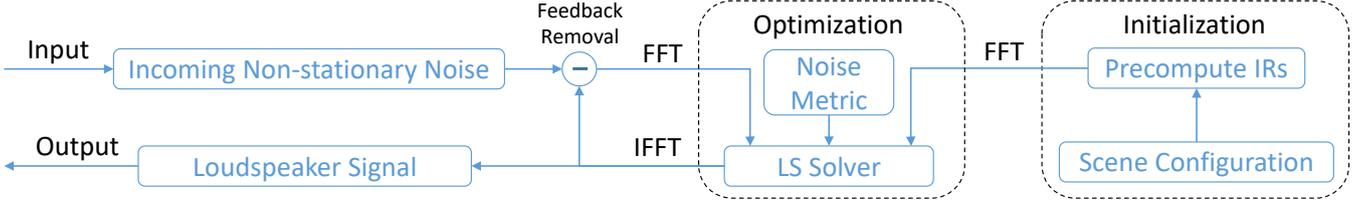


Fig. 1: The stages and flow of our algorithm. The noise is measured by noise level (see Equation (14)). The driving functions for the loudspeakers are solved from linear systems using complex regularized least-squares (LS).

We can compactly rewrite Equation (4) in terms of the desired field for all monitor points as

$$\mathbf{0} = \mathbf{g}_S^T(\mathbf{p}_m)\mathbf{D}(\mathbf{x}) + \mathbf{g}_L^T(\mathbf{p}_m)\mathbf{D}(\mathbf{y}). \quad (7)$$

The underlying optimization problem can be formulated as:

$$\begin{aligned} C_m &= -\mathbf{g}_S^T(\mathbf{p}_m)\mathbf{D}(\mathbf{x}), \\ f(X_m, \mathbf{D}(\mathbf{y})) &= \mathbf{g}_L^T(\mathbf{p}_m)\mathbf{D}(\mathbf{y}); \end{aligned} \quad (8)$$

with our objective function being

$$\arg \min_{\mathbf{D}(\mathbf{y})} \sum_{m=1}^{N_M} [C_m - f(X_m, \mathbf{D}(\mathbf{y}))]^2. \quad (9)$$

To solve the above linear system, we further define the acoustic transfer matrix $\mathbf{Q} = [\mathbf{g}_L(\mathbf{p}_1), \dots, \mathbf{g}_L(\mathbf{p}_{N_L})]^T$, which has a dimension $N_M \times N_L$ and the vector $\mathbf{C} = [C_1, \dots, C_{N_M}]^T$. Then we can obtain the solution in the least-square sense:

$$\mathbf{D}(\mathbf{y}) = (\mathbf{Q}^H \mathbf{Q} + \lambda \mathbf{I}_{N_L})^{-1} \mathbf{Q}^H \mathbf{C}, \quad (10)$$

where $(\cdot)^H$ denotes the Hermitian transpose of a complex matrix, \mathbf{I}_{N_L} is the identity matrix of order N_L , and λ is the regularization term, which is set through experiments to 0.01 to prevent unreasonably high loudspeaker efforts.

3.3. Feedback Removal

Frequency domain noise processing requires segmenting the input signal in the time domain and performing a short time Fourier transform. Consequently, Equation (10) is evaluated for each segment to obtain the loudspeaker driving function. Our derivation in Section 3.2 assumes reliable sensing of $D(\mathbf{x}_s, \omega)$ for all noise sources. However, the loudspeaker signal at one time segment can affect the sensing in the next segment, a common issue in many ANC systems.

We address this issue by identifying active signals generated from previous segments. In the time domain view, the signal propagating to the source at \mathbf{x}_s from segment i is

$$P^{(i)}(\mathbf{x}_s, t) = \sum_{l=1}^{N_L} P^{(i)}(\mathbf{y}_l, t) * \text{IR}(\mathbf{y}_l, \mathbf{x}_s, t). \quad (11)$$

Then, for a following segment j , the interfering signal from segment i is

$$P^{(j,i)}(\mathbf{x}_s, t) = P^{(i)}(\mathbf{x}_s, t)W((j-i)T, (j-i+1)T), \quad (12)$$

where $W(t_1, t_2)$ is a window function truncating signals outside $[t_1, t_2]$ and T is the segment time length. Let the raw noise sensor signal affected by interference at \mathbf{x}_s be $P_c^{(j)}(\mathbf{x}_s, t)$, we recover the true noise signal $P^{(j)}(\mathbf{x}_s, t)$ as

$$P^{(j)}(\mathbf{x}_s, t) = P_c^{(j)}(\mathbf{x}_s, t) - \sum_{i < j} P^{(j,i)}(\mathbf{x}_s, t). \quad (13)$$

Then the noise signal can be transformed to $D(\mathbf{x}_s, \omega)$ as the input to our algorithm pipeline, as shown in Fig. 1.

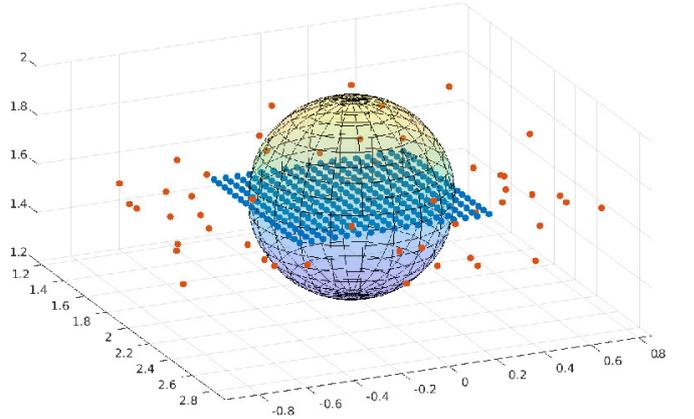


Fig. 2: Distribution of sound sources and receivers for scenes in Table 2. The sphere mesh denotes the target control region, which has a radius of 0.4 meters; the blue dots are dense monitor points, but only a subset of them is used in our optimization scheme; and the red dots are active loudspeakers we added to the scene, which are all outside of the control region.

4. EXPERIMENTS

4.1. Computing Noise Metric

When we have computed the driving signals from Equation (10), we can substitute $\mathbf{D}(\mathbf{y})$ into Equation (4) to obtain the combined noise field $S(\mathbf{p}_m, \omega)$. With Parseval's theorem, the noise level can be calculated as

$$L(\omega) = 20 \log_{10} \frac{|S(\mathbf{p}_m, \omega)|}{p_0} + A\left(\frac{\omega}{2\pi}\right), \quad (14)$$

where $p_0 = 2 \times 10^{-5}$ Pa is the reference pressure for the human hearing threshold, and $A(f)$ is the most commonly used

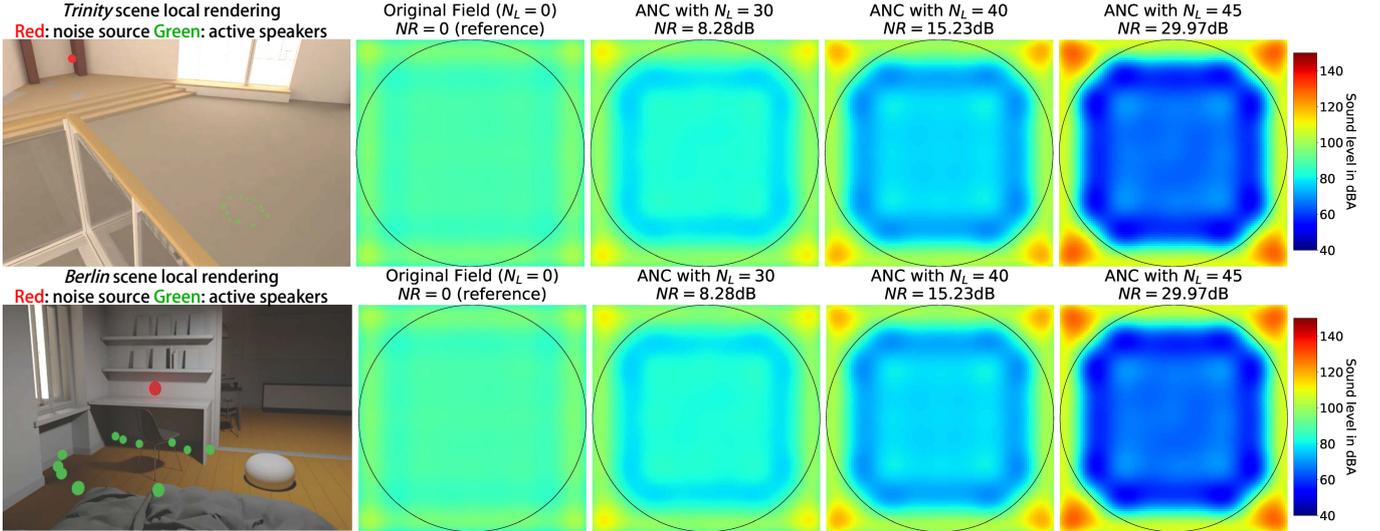


Fig. 3: Noise field and noise reduction (NR) values by varying the number of loudspeakers (N_L) in the *Trinity* (top row) and *Berlin* (bottom row) scenes. In each field plot, the circle with a radius of $0.4m$ encloses the target control region. $N_M = 45$ monitor points are uniformly sampled in the circle. The rendering in the leftmost column shows two possible noise source and loudspeaker configurations, while the actual number of sources and loudspeakers varies in the experiment. Note that with more loudspeakers, the noise field is better controlled with max NR values of $30.0dB$ for *Trinity* and $24.1dB$ for *Berlin*. Reductions are achieved for overall target areas with noise frequency up to $8kHz$ due to accurate sound propagation and optimization.

A-weighting noise level adjustment for human hearing [20]. The resulting noise level will have a corrected decibel unit dBA. Based on this metric, we calculate the average noise reduction (NR) value in dB with respect to the original noise field for each synthesized field to indicate ANC performance.

Table 2: Specifications of our experiments on the CAD models of two large real-world scenes. We highlight the noise reduction value before and after ANC.

| Scene | Size | N_S | N_M | Noise Level Before | Noise Level After |
|---------|----------|-------|-------|--------------------|-------------------|
| Trinity | $450m^3$ | 9 | 45 | 88.37dBA | 58.40dBA |
| Berlin | $370m^3$ | 5 | 45 | 96.46dBA | 72.36dBA |

4.2. Setup and Results

We use the *Trinity* and *Berlin* 3D models shown in Fig. 3, which are captured from real-world buildings, as our benchmarks. We highlight various details of the models in Table 2, and the sample point distribution of the two scenes with different centers is shown in Fig. 2. Although our formulation is general for 3D space, in this paper we visualize noise fields for a representative 2D plane. IRs are precomputed using geometric propagation that traces specular and diffuse rays and performs up to 200 bounces for accurate reverberation.

To evaluate the impact of the number of active loudspeakers in use, we vary parameter N_L and visualize the noise field before and after our optimization algorithm in Fig. 3. In both scenes, with an insufficient number of active loudspeakers

(e.g. $N_L = 30$ or $N_L = 40$), we only have weak control over the target regions because we are solving an under-determined system. However, as we increase the number of loudspeakers, we gain more control over the target region. The noise power can be significantly reduced with $N_L = 45$ loudspeakers, achieving $NR = 30.0dB$ and $NR = 24.1dB$ for *Trinity* and *Berlin*, respectively, for the whole region. Further increasing N_L does not give significantly more noise reduction.

In addition, our results show consistency with the multiple-input multiple-output inverse theorem (MINT) [22]. Because our IRs are not derived from simple principle approximations, given the acoustic complexity of common rooms, our system will not suffer from the sensitivity problem at some seemingly symmetrical positions of sources and receivers.

5. CONCLUSION AND FUTURE WORK

In this paper, we proposed a novel method to dynamically compute the optimal driving signals for a set of active loudspeakers based on accurate IRs to minimize the overall noise level over a target region. We evaluated the performance on two complex scenes and observed considerable reduction in noise levels. Despite our flexibility in loudspeaker placement, our method requires more loudspeakers to achieve similar levels of noise reduction compared to mode-matching based methods [11]. Future integration of mode analysis into our algorithm may relax this requirement. Moreover, our implementation can be extended to use hybrid sound propagation to handle low-frequency noise more accurately.

6. REFERENCES

- [1] Mathias Basner, Wolfgang Babisch, Adrian Davis, Mark Brink, Charlotte Clark, Sabine Janssen, and Stephen Stansfeld, "Auditory and non-auditory effects of noise on health," *The Lancet*, vol. 383, no. 9925, pp. 1325–1332, 2014.
- [2] Michael L Honig and David G Messerschmitt, "Adaptive filters: structures, algorithms and applications," 1984.
- [3] Peter M Clarkson, *Optimal and adaptive signal processing*, Routledge, 2017.
- [4] Stephen J Elliott and Trevor J Sutton, "Performance of feedforward and feedback systems for active control," *IEEE Transactions on Speech and Audio Processing*, vol. 4, no. 3, pp. 214–223, 1996.
- [5] LJ Eriksson, "Computer-aided silencing emerging technology," *Sound Vib*, vol. 24, no. 7, pp. 42–45, 1990.
- [6] Trung-Kien Le and Nobutaka Ono, "Robust tdoa-based joint source and microphone localization in a reverberant environment using medians of acceptable recovered toas," in *Acoustic Signal Enhancement (IWAENC), 2016 IEEE International Workshop on*. IEEE, 2016, pp. 1–5.
- [7] Carl Schissler, Christian Loftin, and Dinesh Manocha, "Acoustic classification and optimization for multimodal rendering of real-world scenes," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 3, pp. 1246–1259, 2018.
- [8] Mingsong Dou, Li Guan, Jan-Michael Frahm, and Henry Fuchs, "Exploring high-level plane primitives for indoor 3d reconstruction with a hand-held rgb-d camera," in *Asian Conference on Computer Vision*. Springer, 2012, pp. 94–108.
- [9] Lueg Paul, "Process of silencing sound oscillations," June 9 1936, US Patent 2,043,416.
- [10] Jacob Benesty and Dennis R Morgan, "Frequency-domain adaptive filtering revisited, generalization to the multi-channel case, and application to acoustic echo cancellation," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*. IEEE, 2000, vol. 2, pp. II789–II792.
- [11] Jihui Zhang, Thushara D Abhayapala, Wen Zhang, Prasanga N Samarasinghe, and Shouda Jiang, "Active noise control over space: A wave domain approach," *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 26, no. 4, pp. 774–786, 2018.
- [12] David T Blackstock, "Fundamentals of physical acoustics," 2001.
- [13] Philippe Langlet, Anne-Christine Hladky-Hennion, and Jean-Noël Decarpigny, "Analysis of the propagation of plane acoustic waves in passive periodic materials using the finite element method," *The Journal of the Acoustical Society of America*, vol. 98, no. 5, pp. 2792–2800, 1995.
- [14] Mingsian R Bai, "Application of bem (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries," *The Journal of the Acoustical Society of America*, vol. 92, no. 1, pp. 533–549, 1992.
- [15] Shinichi Sakamoto, Ayumi Ushiyama, and Hiroshi Nagatomo, "Numerical analysis of sound propagation in rooms using the finite difference time domain method," *The Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 3008–3008, 2006.
- [16] Nikunj Raghuvanshi, Rahul Narain, and Ming C Lin, "Efficient and accurate sound propagation using adaptive rectangular decomposition," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 15, no. 5, pp. 789–801, 2009.
- [17] Jont B Allen and David A Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.
- [18] Carl Schissler and Dinesh Manocha, "Interactive sound propagation and rendering for large multi-source scenes," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 1, pp. 2, 2016.
- [19] Anish Chandak, Christian Lauterbach, Micah Taylor, Zhimin Ren, and Dinesh Manocha, "Ad-frustum: Adaptive frustum tracing for interactive sound propagation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1707–1722, 2008.
- [20] Electroacoustics-Sound level meters Part, "1: Specifications, international standard iec 61672-1: 2013," Tech. Rep., Tech. Rep, 2013.
- [21] Shoichi Koyama and Hiroshi Saruwatari, "Sound field decomposition in reverberant environment using sparse and low-rank signal models," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 395–399.
- [22] Masato Miyoshi and Yutaka Kaneda, "Inverse filtering of room acoustics," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 36, no. 2, pp. 145–152, 1988.