

Acoustic Reliefs

JEREMY CHEW, ETH Zürich, Switzerland

MICHAL PIOVARČI, ETH Zürich, Switzerland

KANGRUI XUE, Stanford University, United States of America

DOUG L. JAMES, Stanford University, United States of America

BERND BICKEL, ETH Zürich, Switzerland

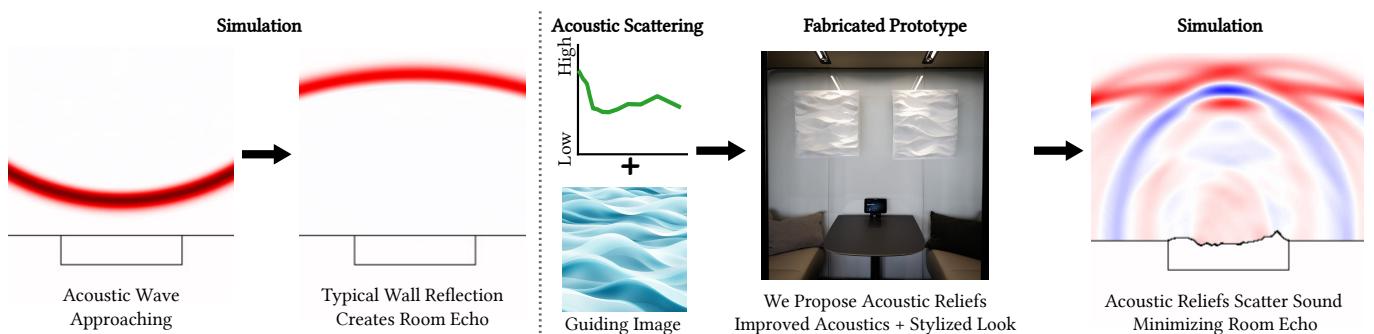


Fig. 1. When an acoustic wave approaches a flat wall it undergoes a mirror reflection. This induces a familiar boomy, harsh, and echoey sound. To improve the acoustic properties of the room and its visual appeal, we propose *Acoustic Reliefs*. *Acoustic Reliefs* combine the functionality of acoustic diffusers with appearance of bas-reliefs. *Acoustic Reliefs* break up sound waves, reduce echoes, promote balanced sound, and provide customized image-driven appearance.

We present a framework to optimize and generate *Acoustic Reliefs*: acoustic diffusers that not only perform well acoustically in scattering sound uniformly in all directions, but are also visually interesting and can approximate user-provided images. To this end, we develop a differentiable acoustics simulator based on the boundary element method, and integrate it with a differentiable renderer coupled with a vision model to jointly optimize for acoustics, appearance, and fabrication constraints at the same time. We generate various examples and fabricate two room-scale reliefs. The result is a validated simulation and optimization scheme for generating acoustic reliefs whose appearances can be guided by a provided image.

CCS Concepts: • Applied computing → Sound and music computing; Computer-aided design; • Computing methodologies → Modeling and simulation.

Additional Key Words and Phrases: Acoustics, Optimization, Fabrication, Diffuser design, Relief design

ACM Reference Format:

Jeremy Chew, Michal Piovarči, Kangrui Xue, Doug L. James, and Bernd Bickel. 2025. Acoustic Reliefs. *ACM Trans. Graph.* 44, 6, Article 241 (December 2025), 23 pages. <https://doi.org/10.1145/3763287>

Authors' addresses: Jeremy Chew, ETH Zürich, Zürich, Switzerland, jchew@ethz.ch; Michal Piovarči, ETH Zürich, Zürich, Switzerland, michal.piovarci@gmail.com; Kangrui Xue, Stanford University, Stanford, California, United States of America, kangruix@stanford.edu; Doug L. James, Stanford University, Stanford, California, United States of America, djames@cs.stanford.edu; Bernd Bickel, ETH Zürich, Zürich, Switzerland, bickelb@ethz.ch.

Please use nonacm option or ACM Engage class to enable CC licenses
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

© 2025 Copyright held by the owner/author(s).

ACM 0730-0301/2025/12-ART241

<https://doi.org/10.1145/3763287>



1 INTRODUCTION

The sound we hear from the environment consists of the direct sound from the source(s), and also the indirect sound that is reflected and scattered off the various surfaces around us. Acoustic engineers are usually concerned about the latter, and within a given space, the materials, geometry, and positions of various surfaces affect how sound is both reflected and absorbed into the environment. Most architectural acoustics is concerned with sound absorption, but managing how sound is reflected and diffused across surfaces is just as important. When applied correctly, good acoustic diffusers promote spaciousness, reduce unwanted echoes, and generally improve speech intelligibility while maintaining other aspects of acoustics such as sound reverberance and envelope [Cox and D'Antonio 2016].

Most modern diffuser designs (Figure 2) can be attributed to Schroeder [1975], who used a constrained set of design equations to produce acoustic diffusers that dispersed in a predictable manner, resulting in a family of so-called *Schroeder diffusers*. Five years later, de Jong and van den Berg [1980] proposed using an iterative method to further optimize these Schroeder diffusers. However, it was only after Cox [1995] and D'Antonio [1995] experimentally showed the marked improvement that optimization had over the initial Schroeder designs that the concept started seeing more widespread use and acoustic engineers began exploring the idea. Nevertheless, recent work on acoustic diffuser design is still fairly limited, often using simplified models and limiting the types of geometry that can be generated as diffusers [Li and Hamilton 2020; Rust 2021; Tan et al. 2022]. Furthermore, it is also notable that such optimization efforts often rely on blackbox or heuristic optimization methods, as the derivative is usually not known and difficult to compute [Cox



Fig. 2. Examples of Acoustic Diffusers. Left to right, top to bottom: Acoustic Diffuser Panels, IDL, ETH Zürich [Rust 2021]. City Wood by ALPHA Acoustiki. Primitive Root Diffuser by Music City Acoustics. Sound Diffusers by Acoustic Fields. Acoustic diffusing discs in the Royal Albert Hall (CC BY-SA 4.0 Colin).

and D’Antonio 2016]. In this regard, there is a distinct lack of a validated method that is able to efficiently optimize acoustic diffusers of arbitrary geometry.

Enabling a wider variety of geometries for acoustic optimization has the potential to yield improvements to more than the acoustic properties. For a diffuser to be effective, its physical size must be substantial with respect to the target room [D’Antonio and Cox 2000]. As a result, acoustic diffusers form large panels that cover entire walls, and thus their appearances directly influence the overall visual appeal of a room. Unfortunately, current optimization schemes focus solely on acoustics, leaving the visual appearance to follow the functional form. As a result, fitting a diffuser in a room can be challenging from an aesthetic point of view. Often suboptimal diffusers are used to achieve a more harmonious match with the atmosphere [Rust 2021].

In this paper, we introduce a system that integrates a differentiable acoustics simulator with a differentiable renderer to allow for visually guided acoustic diffuser optimization. We call our generated acoustic diffusers *Acoustic Reliefs*, as a reference to relief sculptures, whose figures are projected from a planar background. For reproducibility purposes, our code and generated *Acoustic Reliefs* can be found on Github¹.

Specifically, we first develop a differentiable Boundary Element Method (BEM) solver to solve the Helmholtz equation at various frequencies (Section 3) and compute the derivatives of the diffusion metric with respect to the vertices of a mesh (Section 4.1). Then, we integrate a differentiable renderer and combine it with pre-trained neural networks for image guidance (Section 4.2). We then introduce various additional losses for fabrication constraints and quality enhancement and use the complete pipeline to jointly optimize a mesh for both acoustic diffusion and appearance (Section 4.3). We validate our simulator and showcase various generated acoustic reliefs in Section 5. Using a GPU finite-difference time-domain (FDTD) room acoustics solver, we render audio previews in simple rooms before

¹https://github.com/mickey1356/acoustic_reliefs

and after being treated with our acoustic reliefs. We also physically fabricate several reliefs.

To our knowledge, this is the first system that is able to perform gradient-based optimization of an acoustic diffuser. Furthermore, our system is unique in that it is not limited to specific families of diffusers, but instead uses a heightfield parameterization which allows for much greater flexibility in the geometry of the diffuser. Additionally, the coupled appearance optimization allows for the complete customization of the visual aspects of an acoustic diffuser, and could thus be useful for both acousticians and designers alike, allowing either to generate both artistic and functional acoustic diffusers.

2 RELATED WORK

2.1 Sound Simulation

The simulation of acoustic waves can be classified into three main categories: (1) wave-based methods, (2) geometric methods, and (3) hybrid methods. We will provide a brief overview of each category to motivate our simulation choice. For more details on individual methods, we refer the reader to the survey conducted by Liu and Manocha [2021].

Wave-based methods directly simulate acoustic waves that pass through volumes or interact with surfaces [Mehra et al. 2013]. One of the most common volumetric methods is the finite-difference time-domain (FDTD) method [Botteldooren 1994; Savioja et al. 1994; Yee 1966]. Due to its generality and ease of parallelization, it is widely applied in acoustic design [Bilbao 2009, 2012; Wang et al. 2018; Xue et al. 2023, 2024]. Unfortunately, the method requires meshing the entire volume domain. This requirement is impractical for diffuser design, where large or open spaces are often considered and would require specialized treatments to admit general geometries [Berenger 1994; Nabizadeh et al. 2021]. As a result, the FDTD method significantly restricts the family of realizable designs [Li and Hamilton 2020; Pilch 2021]. One of the most common surface-based methods is the Boundary Element Method (BEM) [Ciszkowski and Brebbia 1991; Fairweather et al. 2003; Kirkup 2007; Marburg 2018]. Most BEM methods solve the Helmholtz equation in the frequency domain. This has been shown to produce more accurate simulations than alternative methods [Cox 1994; D’Antonio et al. 2022; Gumerov and Duraiswami 2009; Lock and Holloway 2016; Randrianoelina et al. 2012]. Moreover, the method is widely used in diffuser design with many specialized approximations that improve the time complexity [Kirchhoff 1883; Ogilvy 1986; Pignier et al. 2015; Tsingos et al. 2007].

An alternative to wave-based simulation methods are geometric methods. Geometric methods are analogous to ray tracing and follow acoustic rays through the scene to simulate sound propagation [Krokstad et al. 1968]. Such a formulation is more computationally efficient than wave-based methods and works well for high frequencies. Unfortunately, it is not well suited to modeling wave phenomena such as diffraction and interference [Bork 2005a,b]. Nevertheless, geometrical methods remain popular among acoustic engineers, and is often the method of choice used in industry sound simulation software such as Odeon [Odeon A/S 2024] and EASE [AFMG 2024]. Last but not least, hybrid methods seek to strike

the balance between accuracy and computational efficiency. Techniques range from hybrid FEM-BEM formulations [Granier et al. 1996], exploiting precomputations [Rungta et al. 2018], as well as taking advantage of machine learning techniques [Lee et al. 2023; Ratnarajah and Manocha 2024; Tang et al. 2021].

We base our choice of the simulation framework on the two physical phenomena that govern the behavior of acoustic diffusers: diffraction and scattering [Bork 2005a,b; Schroeder 1975]. In order to accurately model these effects, we opt for a wave-based formulation. During diffuser optimization, we employ surface-based discretization BEM which allows us to further simplify the boundary conditions while maintaining accuracy [Yuan et al. 1997]. Afterwards, we use the FDTD method to simulate time-domain audio previews of our diffusers.

2.2 Acoustic Diffusers

In this section, we highlight some key works in acoustic diffusers. We refer the reader to the textbook *Acoustic Absorbers and Diffusers* by Cox and D’Antonio [2016] which is widely considered the definitive work in the field.

Although acousticians have been designing diffusing surfaces back when Sabine [1922] formalized the field of building acoustics, it was only in the 1970s when Schroeder [1975] introduced Schroeder diffusers that the acoustic properties of such scattering surfaces could be easily defined and designed. A Schroeder diffuser is a patterned acoustic panel with a series of evenly spaced pillars that vary in height. After D’Antonio and Konnert [1984] presented a review on the theory and application of Schroeder diffusers, including experimental validation, they also commercialized and, therefore, popularized these diffusers, making them widely known and used, even to today. To evaluate the quality of a diffuser Hargreaves et al. [2000] introduced a new diffusion coefficient based on the autocorrelation function, which is now part of the ISO standard 17497-2:2012 [ISO 2012]. The most recent advances focus on optimization of the design parameters [Cox 1995; D’Antonio 1995; de Jong and van den Berg 1980; Li and Hamilton 2020; Tan et al. 2022], used materials [Dillon 2013; Redondo et al. 2016], inclusion of location-specific objectives [Pilch 2021], data-driven approximations [Rust 2021], and multi-diffuser optimizations [Quiles et al. 2024].

Most notably, the optimization schemes mentioned above rely on gradient-free frameworks. Cox and D’Antonio [2016] writes that the derivative is often unknown and its numerical approximations are not sufficiently precise for acoustic design. In this work, we revisit the problem of diffuser design through the lens of computer graphics. We propose a relaxation of the diffuser shape domain that is well suited for optimization of both acoustic and visual properties of the diffuser. This enables us to perform gradient-based optimization of both properties on room-scale-sized diffusers.

2.3 Relief Generation

The first algorithms for automatic bas-relief generation relied on converting 3-dimensional scenes to heightfields [Cignoni et al. 1997]. Later, these reliefs were generalized to ensure fabricability [Weyrich et al. 2007] or to follow an arbitrary surface [Schüller et al. 2014]. In this work, our aim is to convert arbitrary input images to reliefs. This

problem is complicated by bas-relief ambiguity, where it has been shown that for a single image, it is impossible to uniquely determine a single heightfield [Belhumeur et al. 1997]. This is similar to the shape-from-shading problem [Horn 1970] and similar heuristics were proposed for image-to-relief generation [Alexa and Matusik 2010; Bermano et al. 2012; Peng et al. 2019]. Alternatively additional user guidance can be used to solve the ambiguity [Dvorožňák et al. 2018; Kolomenkin et al. 2011; Zeng et al. 2014; Zhang et al. 2018a].

In this work, our aim is to extend the field of relief design by combining visual optimization with an acoustic objective. As our ultimate goal is to manufacture the diffusers, we opt to base our approach on that proposed by Weyrich et al. [2007] that explicitly constrains surface variations to a fabricable subset.

2.4 Differentiable Simulation

During modeling and design, it is often desirable to solve the so-called inverse problem, i.e., to calculate the optimal design configuration needed to achieve a desired functional outcome. In the past, this involved specialized solutions [Chen et al. 2014]. However, in recent years, we have seen rapid advancement of differentiable simulators [Hu et al. 2019b]. These methods provide forward predictions with derivative information, facilitating direct update of the optimization parameters. Such tools were used for rigid body design [Hu et al. 2019a], soft body design [Du et al. 2021], fluid simulation [Du et al. 2020], and even acoustics [Caeiro et al. 2017; Jin et al. 2024; Morales and Manocha 2016]. Unfortunately, current acoustic models cannot be applied to diffuser designs. Therefore, in this work, we propose the first differentiable numerical simulator capable of designing room-scale acoustic diffusers.

Differentiability is now widely considered even in the solution of the rendering equation. Thanks to the precise tracking of light transport, it is possible to compute the gradients of many rendering parameters, such as object location, scene materials, or environmental lighting [Jakob et al. 2022a]. However, due to the discrete nature of a geometric changes that must be considered, a computation of gradients for shape optimization stood as a significant challenge in the field [Kato et al. 2020]. Recently methods have been developed that are capable of processing heightfields [Tong et al. 2023], pillar-like structures [Perroni-Scharf and Rusinkiewicz 2023], or general geometries [Nicolet et al. 2021]. In this work, we rely on Mitsuba 3 [Jakob et al. 2022a] for differentiable rendering and integrate it with our own differentiable acoustic simulator.

3 ACOUSTIC DIFFUSERS

Numerically modeling acoustic diffusion is a well-established process [Ciskowski and Brebbia 1991; Fairweather et al. 2003; Kirkup 2007; Marburg 2018]. We first consider the Helmholtz equation, which represents a time-independent form of the acoustic wave equation:

$$\nabla^2 u(\mathbf{x}) + k^2 u(\mathbf{x}) = 0, \quad (1)$$

where $u(\mathbf{x}) \in \mathbb{C}$ is the **complex** acoustic pressure at some point \mathbf{x} in space, k is the wavenumber given by $k = \frac{2\pi f}{v}$, f represents the frequency (Hz) of the wave, and v represents the wave’s phase velocity (m/s). Additionally, we solve this equation in an unbounded domain, to mimic the standardized way of measuring the diffusion

coefficient where the surface to be measured is placed in an anechoic chamber, and only the surface's reflections are considered [ISO 2012].

We apply the Boundary Element Method to solve the equation. The core idea is to discretize the domain at the boundary. To this end, we mesh the boundary into **triangular** elements, and employ the collocation method. We take the collocation point to be the centroid of each element, treating the pressures and its derivatives to be constant over the element. In other words, to compute the boundary pressure \mathbf{u}_p for some boundary element e_p (with collocation point $\mathbf{p} \in e_p$), we have

$$\frac{1}{2}\mathbf{u}_p = \mathbf{u}^i(\mathbf{p}) + \sum_{e_q \in S} \int_{e_q} \mathbf{u}_q \frac{\partial G(\mathbf{p}, \mathbf{q})}{\partial n_{e_q}} d\mathbf{q}, \quad (2)$$

where $\mathbf{u}^i(\mathbf{p})$ is the incident pressure on element p , S is the boundary of the domain, n_{e_q} is the outward-pointing normal of element q , and $G(\mathbf{p}, \mathbf{q})$ is the associated Green's function for the Helmholtz equation.

Following the ISO standard, we use a point source as our excitation signal, and thus the incident pressure $\mathbf{u}^i(\mathbf{p})$ is given by

$$\mathbf{u}^i(\mathbf{p}) = G(\mathbf{p}_{src}, \mathbf{p}). \quad (3)$$

We note that the pressure derivative term has been neglected. In general, the ideal diffuser should reflect as much sound as possible [Cox and D'Antonio 2016]. To reflect this condition, we only consider sound-hard surfaces characterized by $\frac{\partial u}{\partial n} = 0$. We discuss this consideration in greater detail in Section 5.3

For a more in-depth derivation please refer to Appendix A.

We can then set up a system of N equations, leading to an $N \times N$ matrix, where N is the number of elements in the boundary mesh:

$$\mathbf{M}\mathbf{u} := \left(\frac{1}{2}\mathbf{I} - \mathbf{B}\right)\mathbf{u} = \mathbf{u}^i, \quad (4)$$

where each entry $B_{i,j}$ can be intuitively understood as treating element j as a point source, and computing its scattered pressure on element i .

Then, the external scattered pressures \mathbf{u}^e at each listener l can be computed using

$$\mathbf{u}_l^e = \sum_{e_q \in S} \int_{e_q} \mathbf{u}_q \frac{\partial G(\mathbf{p}_l, \mathbf{q})}{\partial n_{e_q}} d\mathbf{q}, \quad (5)$$

or, in matrix form:

$$\mathbf{u}^e = \mathbf{Q}\mathbf{u}, \quad (6)$$

where each entry $Q_{i,j}$ can be intuitively understood as treating element j as a point source with magnitude \mathbf{u}_j , and computing its scattered pressure on listener i .

In order to explicitly compute the entries $B_{i,j}$ and $Q_{i,j}$, we define the following function f :

$$f(k, \mathbf{x}, \mathbf{y}, n) = \frac{(ik|\mathbf{r}| + 1)}{4\pi|\mathbf{r}|^3} e^{ik|\mathbf{r}|} (\mathbf{r} \cdot n), \quad (7)$$

where $\mathbf{r} = \mathbf{x} - \mathbf{y}$. This is the integration kernel $\int_{e_q} \frac{\partial G(\mathbf{p}_l, \mathbf{q})}{\partial n_{e_q}} d\mathbf{q}$ in Equations 2 and 5.

We then denote \mathbf{c}_i to be the centroid of element i , n_j to be the normal of element j , and $\mathbf{q}_{j,\{1\dots K\}}$ to be the quadrature points for

element j , with quadrature weights $w_{q,\{1\dots K\}}$. We also denote \mathbf{l}_i to be the position of the i th listener. Then, we have

$$B_{i,j} = \frac{1}{2} \sum_t^K w_{q,t} f(k, \mathbf{c}_i, \mathbf{q}_{j,t}, n_j), \quad (8)$$

and

$$Q_{i,j} = \frac{1}{2} \sum_t^K w_{q,t} f(k, \mathbf{l}_i, \mathbf{q}_{j,t}, n_j). \quad (9)$$

Diffusion Coefficient. To evaluate acoustic diffusers, there is an ISO standard on measuring the “directional diffusion coefficient” of a given surface in the free field [ISO 2012]. This coefficient is a per-frequency measure that ranges from near 0 (for a concave surface focusing sound to a single receiver) to 1 (for a surface that evenly reflects the sound to all receivers). To measure the diffusion coefficient, we compute the external pressures \mathbf{u}^e at listeners evenly placed on a hemisphere centered on the diffuser. We then use the diffusion coefficient formula [ISO 2012, Eq. 5] to obtain the diffusion coefficient for a given frequency band. While there is yet to be a comprehensive user study that relates the coefficients to the subjective responses in human listeners [Cox and D'Antonio 2016], it nevertheless is an industry standard that provides us with an objective metric on which to optimize and evaluate our diffusers. Mathematically, for a given receiver l , we have

$$y_l = \left(\sum^F |u_l^e| \right)^2, \quad (10)$$

where F is the total number of frequencies sampled in the frequency band, and $|z|$ is the magnitude of the complex number z . The diffusion coefficient c is then defined as

$$c := \frac{(\sum_l y_l)^2 - \sum_l y_l^2}{(L-1) \sum_l y_l^2}, \quad (11)$$

where L is the number of listeners.

Following D'Antonio et al. [2022], we compute the diffusion coefficient for 17 one-third octave bands from 100 to 4000 Hz, sampling $F = 3$ frequencies per band on a logarithmic scale. We place the sound source and receivers at 100 m and 50 m, respectively, from the center of the diffuser, with the listeners placed at a spatial resolution of 5° in azimuth and elevation, totalling 1297 listeners. For each frequency band, we plot the calculated diffusion coefficient c against its respective frequency band, resulting in the diffuser's *diffusion curve*. Figure 3 shows two examples of a diffusion curve for a rectangular cuboid with dimensions $0.6 \times 0.15 \times 0.6$ m, as well as a commercial Schroeder diffuser, namely, the primitive root diffuser. The ideal diffusion curve is a horizontal line with $c = 1$, representing a diffuser that uniformly scatters sound in all directions at every frequency band. Where applicable, we also display the diffusion coefficient averaged across all frequency bands to provide a single, quantitative measure of a diffuser's performance.

Low-rank Approximation. The system in Equation 4 is nonsparse, and thus solving it requires $O(N^2)$ memory and has a runtime on the order of $O(N^3)$. As such, to make solving the system feasible for meshes with a large number of elements, we use \mathcal{H} -matrices and adaptive cross-approximation introduced by Hackbusch [1999]. We also follow Börm et al. [2003], and use a cluster tree (with parameter

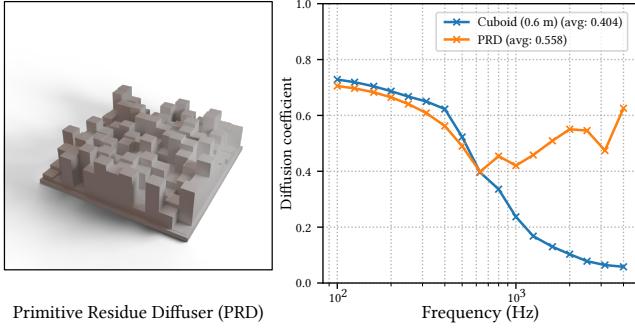


Fig. 3. **Diffusion Curves.** We plot the diffusion curves for a 0.6 m cuboid mesh (blue), and the primitive root diffuser (pictured on the left) (orange) for the one-third octave bands from 100 to 4000 Hz. For visualization purposes, the frequency bands are plotted on a log-scale.

ε) to group elements of the boundary mesh by their distances from one another. Elements that are close to each other correspond to blocks in the \mathcal{H} -matrix that have to be directly computed, while elements which are further away correspond to low-rank matrix blocks, which can be approximated instead. This reduces the memory cost of the system to $O(\varepsilon N \log N)$, and the runtime of the linear solve to $O(\varepsilon N^2 \log N)$ [Bebendorf 2000; Börm et al. 2003; Grasedyck and Hackbusch 2003; Hackbusch 1999]. The full forward pipeline is summarized in Algorithm 1.

ALGORITHM 1: Diffusion coefficient forward pipeline

Input: Input mesh (V, E) , number of listeners L , listener positions P_L , number of sampled frequencies F , and frequency band F_b
Output: Diffusion coefficient c
Build cluster tree \mathcal{T} using input mesh (V, E) ;
Sample F frequencies from F_b on a logarithmic scale;
Set $y_t = 0$;
for each sampled frequency **do**
 Approximate boundary matrix M using cluster tree \mathcal{T} ;
 Approximate listener matrix Q using input mesh (V, E) and
 listener positions P_L ;
 Solve for \mathbf{u} in $M\mathbf{u} = \mathbf{u}^i$;
 Compute $\mathbf{u}^e = Q\mathbf{u}$;
 $y_t \leftarrow y_t + |\mathbf{u}^e|$;
end
Compute $y = y_t^2$;
Compute $c = \frac{(\sum y)^2 - \sum y^2}{(L-1) \sum y^2}$;

4 METHOD

The input to our method is an image that guides the final appearance (Figure 4). The design space is a 2-dimensional heightmap representing the diffuser surface. To optimize the acoustic and visual properties we derive a differentiable acoustic simulator (Section 4.1) and couple it with a differentiable renderer (Section 4.2). To optimize *Acoustic Reliefs* we propose a set of loss terms that guarantee acoustic quality, visual match, and fabricability (Section 4.3). Our loss

terms iteratively update the heightmap with gradient information until the final design is optimized (Section 5).

4.1 Acoustics Gradient

Recently, automatic differentiation (AD) tools have become a viable option for deriving the gradient with respect to the design parameters. In the case of acoustic diffusers, we are interested in computing the system matrices M and Q , as well as the solution \mathbf{u} . To compute these quantities, we rely on iterative solvers. The iterative nature of such solvers inhibits direct application of standard AD tools. Repeated iterations often result in large, inefficient computational graphs (due to the exponential nature of the chain rule), or even inaccurate derivatives as the solver and its derivative might converge at different rates [Beck 1994; Fischer 1991; Hueckelheim et al. 2023].

Therefore, we must rely on traditional methods to compute the derivative. In prior work, the application of such methods was considered infeasible [Cox and D'Antonio 2016]. The chief reason for this conclusion is the considered shape domain of the diffusers. Traditionally, diffusers are modeled as meshes. During the simulation the mesh undergoes deformation which inhibits the calculation of the derivative [Cox and D'Antonio 2016]. In this work, we propose to reparametrize the diffusers with a heightfield. We start by applying the adjoint method [Johnson 2021] to manually compute the gradient of the diffusion coefficient c with respect to the vertex positions V of the mesh. Next, we associate the vertex position V with the heightfield through bilinear interpolation. Such a setup is robust with respect to the meshing and avoids self-intersections during the optimization process.

By letting g be the composition of the functions $c \circ y \circ y_t$, we can rewrite Algorithm 1 mathematically as

$$c = g(QM^{-1}\mathbf{u}^i), \quad (12)$$

where M and Q are approximated using \mathcal{H} -matrices as explained in Section 3.

We introduce $\mathbf{w} := \frac{dg}{d\mathbf{u}^e}^T$ and solve for the so-called *adjoint variable* λ in

$$M^T \lambda = Q^T \mathbf{w}. \quad (13)$$

This allows the derivative $\frac{dc}{dV}$ to be computed as

$$\frac{dc}{dV} = \mathbf{w}^T \frac{\partial Q}{\partial V} \mathbf{u} + \lambda^T \frac{d\mathbf{u}^i}{dV} - \lambda^T \frac{\partial M}{\partial V} \mathbf{u}. \quad (14)$$

The full derivation can be found in Appendix B.

While this method only requires one additional matrix solve (for λ), the two matrix derivatives $\frac{\partial Q}{\partial V}$ and $\frac{\partial M}{\partial V}$ are 3-dimensional tensors, and with a large N , are infeasible to store explicitly. Therefore, we also reformulate the vector-tensor-vector products as

$$\mathbf{w}^T \frac{\partial Q}{\partial V} \mathbf{u} = \sum_{i=1}^L \mathbf{w}_i \sum_{j=1}^N \mathbf{u}_j \frac{dQ_{ij}}{dV} \quad (15)$$

$$\lambda^T \frac{\partial M}{\partial V} \mathbf{u} = \sum_{i=1}^N \lambda_i \sum_{j=1}^N \mathbf{u}_j \frac{dM_{ij}}{dV} \quad (16)$$

respectively, where the derivatives $\frac{dQ_{ij}}{dV}$ and $\frac{dM_{ij}}{dV}$ can be found analytically using, for example, a symbolic mathematics library

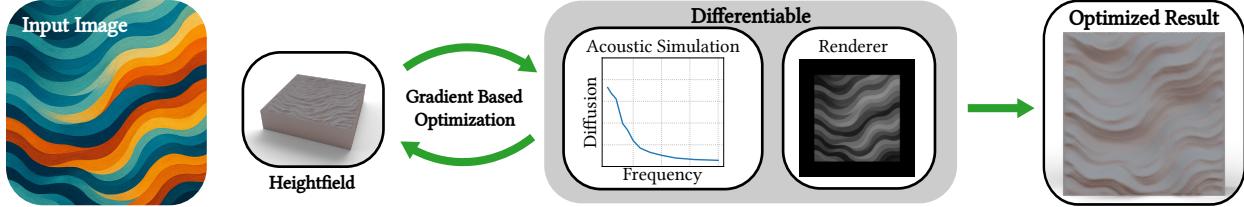


Fig. 4. Overview. Given an input image we construct an optimization heightfield. The heightfield is updated using the gradient information flowing from our proposed differentiable acoustic simulator and a differentiable renderer. The final result is a fabricable *Acoustic Relief*.

such as SymPy² [Meurer et al. 2017]. This avoids the memory limitations by computing one “column” of $\frac{\partial Q}{\partial V}$ and $\frac{\partial M}{\partial V}$ at a time.

4.2 Diffuser Appearance

We use Mitsuba 3 [Jakob et al. 2022b,a] to differentiably render the acoustic relief for the appearance optimization. In order to ensure that the appearance remains largely view-angle-independent, we render the mesh at various camera poses and jointly optimize the appearance from all the views at the same time. The input to guide the appearance is an image.

We project the provided image onto a flat plane and render it with the same camera settings and poses as the *Acoustic Relief*. A CLIP model [Radford et al. 2021] is used to obtain the geometric and semantic loss between the two images. The gradients of the losses can then be backpropagated through the CLIP guidance model and the differentiable renderer to optimize the vertex positions of the acoustic relief.

The CLIP model encodes an image into its semantic and geometric meaning, allowing for looser matches between the rendered image and the target appearance. This enables the optimization to accept “similar”-looking appearances instead of exact matches, and has been demonstrated to be effective in other uses such as sketch generation [Vinker et al. 2023, 2022], and 3D wire art [Tojo et al. 2024].

4.3 Optimization

We parameterize an acoustic relief using a heightfield $H \in \mathbb{R}^{D \times D}$ which is a suitable parametrization to jointly optimize both acoustics and appearance properties.

Starting with a rectangular cuboid of some pre-specified dimensions, we mesh its surface using GMSH [Geuzaine and Remacle 2009]. Then, we apply a heightfield texture to the top face of the box, where the y-displacements of the vertices are given by bilinearly interpolating the pixels of the heightfield texture. We use GMSH instead of constructing a fine, regular, rectangular cuboid mesh as GMSH attempts to mesh geometry into equilateral triangles of roughly equal area. As we will show in Section 5.3, this avoids the bias towards a given diagonal direction that would be present in a grid-like rectangular cuboid mesh, resulting in a generally better visual appearance.

We add a small border of $B = 2$ pixels around the heightfield texture. This prevents any y displacement of the edge vertices, which

may otherwise cause the mesh to self-intersect. As such, there would be $(D - 2B)^2$ pixels in the heightfield texture which can be freely optimized.

Losses. Our losses take into account the acoustic and visual properties, as well as the fabricability of the final result. Since we wish to **maximize** the diffusion coefficient, we define the acoustic loss as

$$E_a := 1 - c. \quad (17)$$

Our rendering loss uses a CLIP loss C which computes the L2 distance between intermediate layer activations, specifically layer 3, of the ViT-B/32 CLIP model. This represents low-level geometric information, such as pose and structure, of the image. Mathematically, this loss is defined as

$$C(I_1, I_2) := \|CLIP_l(I_1) - CLIP_l(I_2)\|_2^2 \quad (18)$$

where $CLIP_l$ refers to activations of layer l of the CLIP model, and I_1, I_2 are RGB images.

Hence, assuming a differentiable renderer \mathcal{R} along with camera poses p_{cam} , we compute the rendering loss as

$$E_r := C(\mathcal{R}(V, F; p_{cam}), \mathcal{R}(V_{cuboid}, F_{cuboid}; I, p_{cam})), \quad (19)$$

where I is the provided image, and V_{cuboid}, F_{cuboid} are the vertices and faces of a rectangular cuboid, respectively. In other words, we apply the provided image as a texture on the top face of a rectangular cuboid, and render it with the same settings as that of our reliefs.

To take into account the depth of the acoustic diffuser we further augment the visual loss with a CLIP term acting directly on the heightfield:

$$E_i := C(H, I) \quad (20)$$

To avoid sharp peaks and otherwise non-fabricable reliefs, we add a smoothness loss similar to Alexa and Matusik [2010] that minimizes the mean squared Laplacian of each pixel in the heightfield texture:

$$E_s := \frac{1}{D^2} \sum_{x=1}^{D-1} \sum_{y=1}^{D-1} (H_{x+1,y} + H_{x-1,y} + H_{x,y+1} + H_{x,y-1} - 4H_{x,y})^2 \quad (21)$$

²<https://www.sympy.org/en/index.html>

Note that this includes the border, so we avoid sharp changes from the edges of the mesh to the center.

To respect a limited milling tool height we also add a barrier loss that limits the maximum displacement of the heightfield to some

maximum height h_{\max} :

$$E_b := -\frac{1}{D^2} \sum_{x,y} \min(\log(H_{x,y} + h_{\max}), \log(h_{\max} - H_{x,y})), \quad (22)$$

and we include a loss that tries to avoid “sunken-in” reliefs, by penalizing negative values in the heightfield texture:

$$E_n := \frac{1}{D^2} \sum_{x,y} \max(0, -H_{x,y}) \quad (23)$$

We compute the gradient for each of the above losses, normalize them individually for better loss balancing, and take a weighted sum:

$$\nabla E = w_a s_a \nabla E_a + w_r s_r \nabla E_r + w_i s_i \nabla E_i + \\ w_s s_s \nabla E_s + w_b s_b \nabla E_b + w_n s_n \nabla E_n \quad (24)$$

where the w 's are weights that can be pre-set, and the s 's represent the scaling factors that normalizes each gradient vector ∇E .

The final gradient is then passed into the optimizer. For all examples in the paper, we use the *Adam* optimizer [Kingma and Ba 2017], as well as the same set of gradient weights.

Lastly, we manually clip the y-displacements after each optimization update to lie between the range, $[-h_{\max}, h_{\max}]$. This works with the barrier loss to ensure malleability while preventing self-intersections with the bottom face of the mesh.

Frequency Selection. The diffusion coefficient computed in Algorithm 1 works on a **per-frequency-band** basis. In order to target even diffusion across more than a single frequency band, during each optimization step, we randomly sample from a list of provided frequency bands with probability proportional to the inverse of its value. Specifically, given a list of frequency bands $F = \{f_1, f_2 \dots f_k\}$, we sample (with replacement) frequency f_i with probability proportional to $\frac{1}{f_i}$.

As we show in Section 5.3, this results in a more even diffusion across multiple frequency bands and a better overall performance for the diffuser. This weighting scheme intuitively works since diffuser geometries are often more tuned towards lower frequencies [Cox and D’Antonio 2016]. As such, we use this weighting scheme to focus optimization on the critical lower frequencies, instead of the higher frequencies where acoustic diffusion is less noticeable.

Post-processing. To further improve visual appearance, we take advantage of human perception of images and sound. More specifically, we are generally interested in optimizing the acoustic diffusers up to the frequencies commonly used during speech, that is, up to 4000 Hz. In contrast, our visual system is capable of processing information at a much finer resolution. To take advantage of this property, we perform a fine-tuning step. First, we subsample the heightfield and the top mesh by the Nyquist rate to ensure that we are not affecting the already optimized audio frequencies. Next, we run *only the appearance optimization portion* of the entire pipeline on the finer mesh. Finally, to ensure that the final mesh remains close to the original optimized heightfield, we add a regularization term:

$$E_{\text{reg}} := \frac{1}{D'^2} (H'_{x,y} - H'^{\text{init}}_{x,y})^2, \quad (25)$$

where $H' \in \mathbb{R}^{D' \times D'}$ is the up-sampled heightfield.

5 RESULTS

To model our acoustic reliefs, we follow Cox and D’Antonio [2016] which suggests using elements smaller than an eighth of the smallest wavelength (4000 Hz ~ 0.086 m). Thus, we mesh our diffusers to element sizes of roughly 0.01 m.

This section is split into five parts. We first evaluate the accuracy and stability of our simulations and their gradients (Section 5.1). Then, we optimize cuboid meshes for the generation of *Acoustic Reliefs*: diffusers that not only perform well acoustically, but are also visually interesting (Section 5.2). We also performed an ablation study on the various components and hyperparameters that make up our optimization pipeline (Section 5.3). Using a GPU FDTD room acoustics solver, we render audio previews illustrating the effects of diffusers in reducing echo and promoting spaciousness in a room (Section 5.4). Finally, we fabricate some of our *Acoustic Reliefs* and perform several impulse measurements, comparing them with the simulated results (Section 5.5).

5.1 Numerical Accuracy

Analytical Solution. We start by evaluating our solver against a problem with a known analytical solution. More specifically, we simulate the plane wave scattering off a rigid sphere [Hart 1951]. We mesh a sphere of radius 1 m and use elements of 0.06 m. We then set up the same plane wave scenario in our simulator and plot the exterior field pressures at a radial distance of 10 m on a polar graph (Figure 5). The pressures were computed with an incident plane wave frequency of 250 and 1000 Hz. In general, we observe an excellent agreement between the simulation (smooth curves) and the analytical solution (discrete points). The main discrepancy lies in the imaginary component of lower-frequency simulations. However, this does not significantly affect the computation of the diffusion coefficient. During the computation, we average over multiple frequencies per band, thus reducing the errors observed at individual frequencies [D’Antonio et al. 2022].

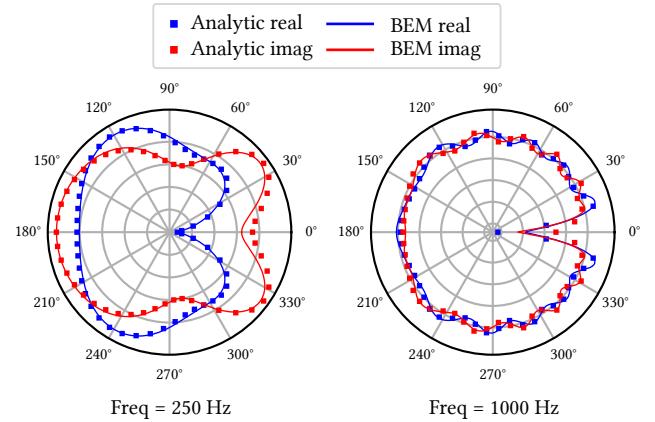


Fig. 5. **Analytical Comparisons.** Scattered pressure of plane wave at 250 and 1000 Hz off a rigid sphere with a radius of 1 m. For the BEM computation, 360 receiver points were evenly spaced placed on a circle with radius 10 m. The analytical solution was computed for 50 evenly spaced listener points.

Commercial Software. In addition, we compare our simulator with standard software. For the comparison, we picked COMSOL® [COMSOL 2024], a general-purpose physics simulation software commonly used by acoustic engineers [Dillon 2013; Khrystoslavenko and Grubliauskas 2017; Ramakrishnan 2010; Tan et al. 2022]. Following Tan et al. [2022], we use the *Pressure Acoustics, Boundary Elements* interface to simulate and measure the diffusion coefficient of our *Acoustic Reliefs*. Figure 6 shows a comparison between the coefficients computed by our differentiable simulator and the coefficients computed by COMSOL® for the *Cat* (0.6 m), and *Cat* (0.9 m) *Acoustic Reliefs*. Close matches suggest that any errors introduced by \mathcal{H} matrices are negligible when computing the full diffusion coefficient, thus confirming that our simulation is accurate in predicting the diffusion coefficients of our reliefs.

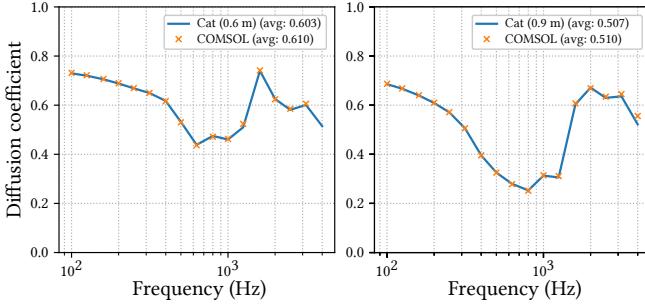


Fig. 6. **COMSOL® Comparisons.** We compare the coefficients computed by our simulator with those computed by COMSOL®. For better visualization, the frequency bands are plotted on a log-scale. We note that on the *Cat* (0.6 m) relief, COMSOL® fails to converge to a solution for the 4000 Hz frequency band. Nevertheless, for the other frequency bands, the two computed coefficients are extremely close, validating the accuracy of our simulator.

Gradients. Our use of \mathcal{H} -matrices can introduce approximation errors into the computed gradients. To evaluate their performance, we compare them against a complete analytical solution and finite differences (Figure 7). Gradients were computed on a height field of $D = 64$ pixels for a cuboid mesh with dimensions $0.6 \times 0.15 \times 0.6$ m at a frequency of 1000 Hz. The cuboid was meshed to an element size of 0.02 m, resulting in 6466 elements. The finite-difference gradients were computed using a step size of 10^{-6} . We can observe that the gradients computed with our adjoint formulation are robust with respect to the computation with either full or \mathcal{H} -approximated matrices. In contrast, the finite difference gradients experience stability issues. This corroborates the findings of prior work that identified derivative computation as a challenging problem [Cox and D’Antonio 2016] and further highlights the benefits of our adjoint formulation.

5.2 Acoustic and Optical Performance

We use the proposed optimization method to generate several acoustic reliefs. In all examples, we first mesh a cuboid to elements of 0.02 m, and then subdivide the top face to elements of 0.01 m. The final mesh elements are then processed using a cluster tree with a radius factor of 1.5 and a maximum cluster size of 128. We found

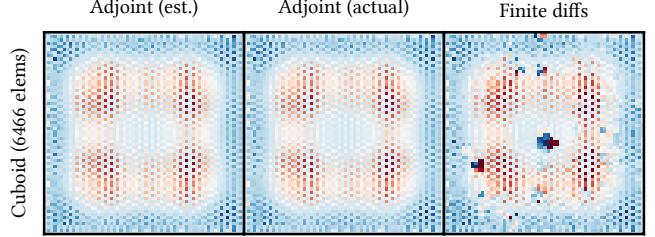


Fig. 7. **Gradient Comparisons.** Gradients for a cuboid of $0.6 \times 0.15 \times 0.6$ m meshed to 0.02 m (6466 elements). The gradients were computed either by approximating using \mathcal{H} -matrices (Left), computing the full boundary matrices (Middle), and finite-differencing (Right).

these values to offer a good tradeoff between matrix compression and computational efficiency.

In Mitsuba, we use the *direct_projective* integrator [Zhang et al. 2023] to render the scenes. The camera is placed 1 m away for 0.6 m reliefs, and 1.5 m for 0.9 m reliefs. During the optimization phase, we render and optimize our reliefs at 5 different camera angles: one at an elevation of 90° facing directly down, and 4 placed at an elevation angle of 45° with azimuthal angles of $0^\circ, 90^\circ, 180^\circ$, and 270° . For lighting, we use a single white point light source with an intensity value of 1.5 placed 1 m directly above the relief. The material of the relief is set to be a purely diffusing white material.

Additionally, we set $w_a = 7$, $w_i = 5$, $w_s = 15$, $w_b = 1$, and $w_n = 1$ for the various loss weights detailed in Section 4.3. For the rendered weights, we use a weight of $w_r = 7$ for the camera placed directly above, and a weight of $w_r = 2$ for the other cameras. Additionally, we set $h_{\max} = 0.1$ m. In the post-processing step, we keep the same weights as the initial optimization, turning off the acoustics optimization, and adding the regularization term (Equation 25) with a corresponding weight of $w_{\text{reg}} = 5$. The initial optimization uses a learning rate of 10^{-3} for 500 iterations. The post-process optimization uses the same learning rate for 50 iterations.

Finally, the frequency bands that we sample from (Section 4.3) are

$$\{1000, 1250, 1600, 2000, 2500, 3150\} \text{ Hz}. \quad (26)$$

These loss weights and parameters were selected via qualitative experimentation to yield visually and acoustically compelling results. Since the gradients are normalized prior to weighting, the weights provide an interpretable balance between objectives. While a full hyperparameter search is possible, the subjective nature (especially in the aesthetic sense) of the final loss limits its practical benefit.

Acoustic Performance. To evaluate the acoustic response of our diffusers, we investigate their behavior when interacting with an impulse. To this end, we can draw an analog with BRDF visualizations, where an incident ray is reflected from a surface and one observes the intensity of light reflected in all directions. Similar experiments can be conducted for sound waves. However, the light ray needs to be replaced by an audio wave (Figure 8 Top-left). The wave then approaches a surface and is reflected back (Figure 8 Top-middle). An ideal diffuser will break the wave and scatter it as much as possible (Figure 8 Top-right).

For easier visualization, we focus on 2-dimensional slices of the scattered sound (Figure 8 Bottom). We can observe that a flat plane has a significant specular component in the reflection. In contrast, our *Acoustic Reliefs* significantly reduce the amount of sound reflected in the mirror direction. Instead, the sound is more uniformly redirected to other locations of the room.

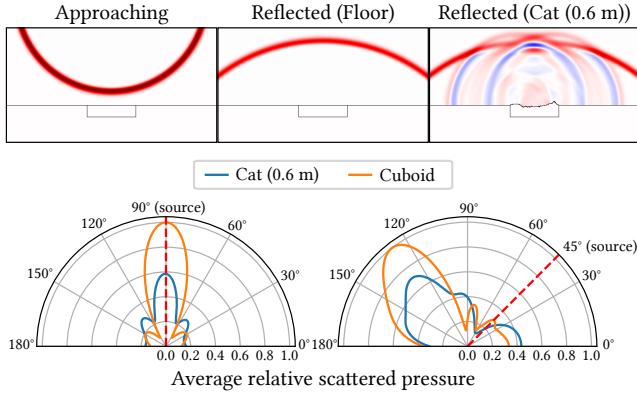


Fig. 8. FDTD and Polar Plots. We show how our *Acoustic Reliefs* significantly breaks up the waveform of an incoming wave compared to a flat surface (Top). Additionally, we plot the average relative scattered pressure for a source located at 90° (Bottom-left) and 45° (Bottom-right).

Optical Performance. To evaluate the optical quality of our diffusers, we render an exemplar diffuser from multiple novel viewpoints (Figure 9). We observe that the novel viewpoints have an appearance similar to the frontal view included in the optimization. With grazing angles the performance diminishes, however, the target image is still recognizable. While the decrease in optical fidelity might be more pronounced when compared to image-only reliefs, we consider the grazing angle performance of acoustic diffusers to be a reasonable trade-off. This compromise is justified by the acoustic properties they provide.

Generalizability. We optimize several acoustic reliefs using different guiding images (Figures 10 and 11). We observe that the guiding images are well represented even from novel viewpoints. The low frequencies of the overall acoustic relief surface remain consistent across different guiding images. However, we can observe that the higher frequencies are adapting to accommodate the images while preserving sound characteristics. We further evaluate the diffusion coefficients of the optimized samples (Figure 12).

As a reminder, the ideal diffusion curve is a horizontal line with $c = 1$ (Section 3). As such, we observe that while there are some discrepancies between the different *Acoustic Reliefs*, they still perform very well. We therefore conclude that the guiding image has a minimal influence on the acoustic properties.

Comparing with Commercial Diffusers. We further compare our optimized *Acoustic Reliefs* with commercial diffusers in Figure 13. Our frame of reference is a commercial Schroeder diffuser, specifically, the primitive root diffuser (PRD). We compare it to the optimized *Stripes (0.6 m) Acoustic Relief*. For additional diffuser surfaces,

please refer to Appendix C. We observe that even with appearance constraints, our *Acoustic Reliefs* are able to outperform the acoustic performance of the PRD in most frequency bands.

5.3 Ablation Experiments

We ablate various hyperparameters and design choices of our proposed optimization system. We use a cuboid as initial geometry and mesh the top face to elements of 0.01 m, and other faces to elements of 0.02 m.

Acoustics-only. By setting the weights of the **appearance-related losses** to zero, we can optimize our reliefs for pure acoustic diffusion (Figure 14). We contrast the results with our combined acoustic and visual optimization, specifically the *Waves (0.6 m)* and *Matterhorn (0.9 m) Acoustic Reliefs*. As expected, there is a trade-off between the two objectives for the 0.6 m diffuser. However, this effect is less prominent for the 0.9 m diffuser. Larger designs have significantly more freedom to accommodate the underlying image. This leads to an improved match in acoustical capabilities, marking only a minor loss with respect to the optimal diffuser.

Appearance-only. By setting the weights of the **acoustic loss** to zero, we can generate reliefs parameterized by a heightfield in the same vein as Alexa and Matusik [2010] and Schüller et al. [2014] (Figure 15). We can observe that the high-frequency information is captured more faithfully with pure reliefs. Unfortunately, as the generated reliefs are extremely flat, their acoustic performance has a negligible difference compared to a non-optimized flat surface. In contrast, our *Acoustic Reliefs* provide diffuser-level performance while maintaining a reasonably high definition of the final image to capture its salient features. An interesting note is that the acoustic diffuser might appear more visually pleasing. We attribute this to the more pronounced depth variations creating a 3D-pop effect. A regular relief does not exhibit this effect as the image metric primarily focuses on matching the high-frequencies rather than on producing an aesthetically pleasing image.

Acoustics-Appearance Trade-Off. We also investigate how the generated relief changes from an image-only relief to an acoustic-only relief. To do so, starting from the initial set of weights, we freeze either the acoustics or appearance weight, and set the other weight to equally spaced values between 0 and the initial weight. Figure 16 shows both the evolution of the generated relief, as well as their diffusion curves.

The diffusion curves and image quality generally behave as expected, with an increase in the acoustics/appearance weights corresponding to an increase in quality of the acoustics/appearance respectively. However, we note that the initial change from acoustics-only to the combined optimization results in slightly better acoustical performance. We attribute this to the fact that adding in the appearance optimization creates high-frequency details which may otherwise be neglected by the acoustics optimization. This helps break up the reflected sound, thus improving the diffusion coefficient metric.

Material Absorption. We previously made the assumption that the material of the diffuser was sound-hard, thus allowing us to ignore



Fig. 9. **Optical Performance.** We render an example *Acoustic Relief (Waves (0.6 m))* from multiple novel viewpoints. While most novel viewpoints have an appearance similar to the frontal view, the effect is diminished at grazing angles.

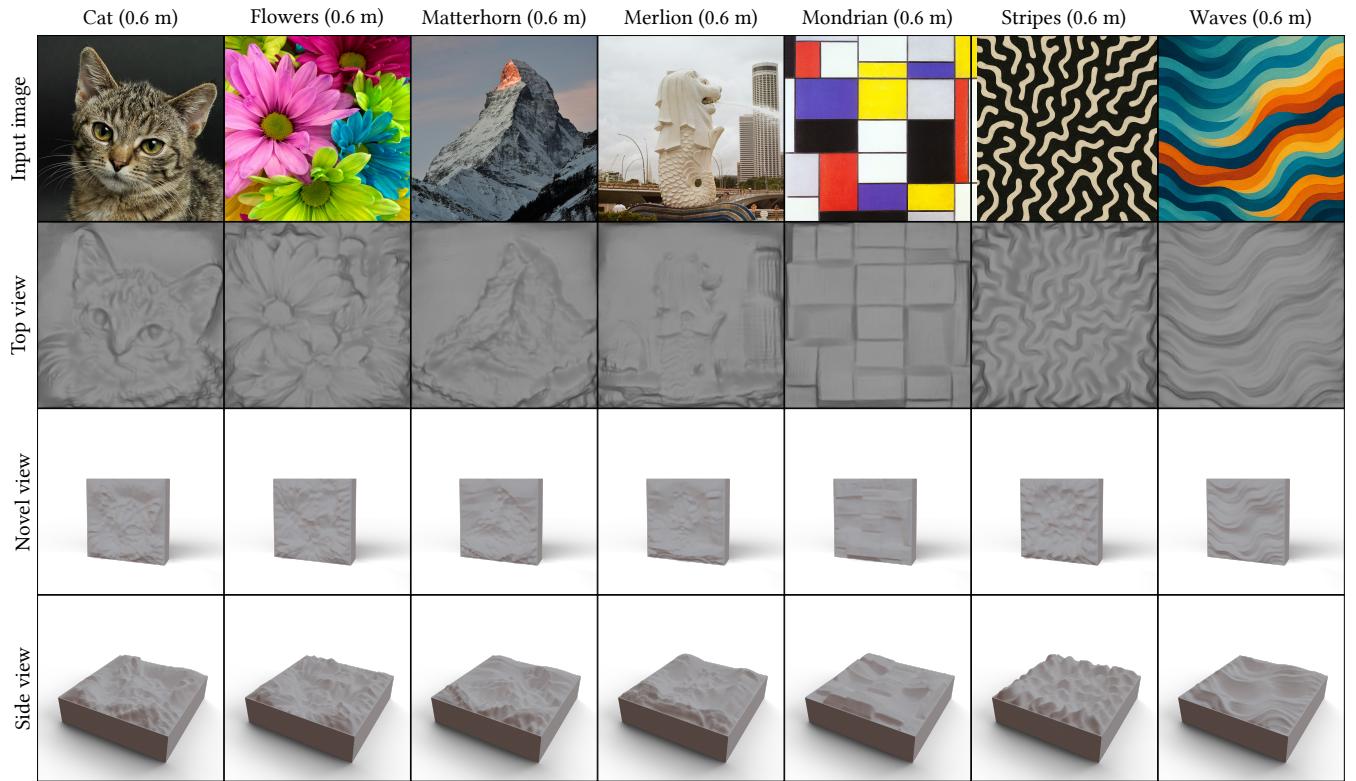


Fig. 10. **0.6 m Optimized Diffusers.** We showcase several **0.6 m Acoustic Reliefs**. From top to bottom, we show the guiding image, the top view which was used in optimization, a novel view, and a side view that shows the height profile of the relief. The guiding images are used with the courtesy of Kumpicki [2019] (*Cat*), Srirugsar [2016] (*Flowers*), Niedermann [2021] (*Matterhorn*), Vinh Lâm [2022] (*Merlion*), and Mondrian [1923] (*Mondrian*).

the pressure derivative term when deriving Equation 2. In Figure 17, we use COMSOL® to compute the diffusion coefficient for various material absorption values. As can be seen, at low levels of material absorption, which is valid for materials such as wood, the diffusion coefficient does not vary significantly with respect to the material absorption. Intuitively, the diffusion coefficient can be thought of as an auto-correlation coefficient computed across all listener positions. Since the material absorption uniformly reduces the amount of per-listener scattered sound energy, the overall diffusion coefficient remains relatively unchanged.

Source Positions. We follow the ISO standard [ISO 2012] for the diffusion coefficient which is the most widely recognized standard for evaluating diffuser performance [Cox and D’Antonio 2016]. As such, our acoustical metric E_a (Eq. 17) is computed under this standard setup, which involves using a source placed directly above the center of the diffuser (normal incidence). At first glance, this may lead to situations where our *Acoustic Reliefs* overfit on this specific configuration. In Figure 18, we vary the elevation angle of the source positions from 0° (normal incidence) to 75° with intervals of 15° , and compute the corresponding diffusion coefficients for the *Waves (0.6 m) Acoustic Relief*. The diffusion curves show fairly

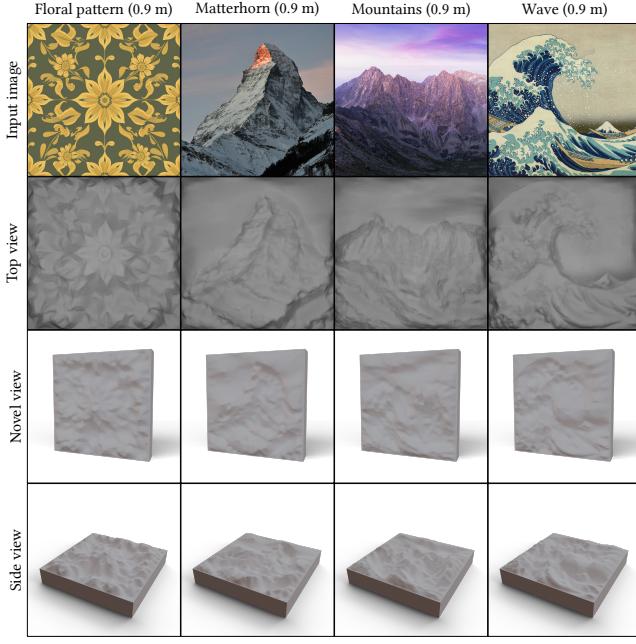


Fig. 11. 0.9 m Optimized Diffusers. We showcase 0.9 m examples of our *Acoustic Reliefs*. Again, from top to bottom, we show the guiding image, the top view which was used in optimization, a novel view, and a side view that shows the height profile of the relief. The guiding images are used with the courtesy of Stefancik [2016] (*Mountains*), and Hokusai [1831] (*Wave*).

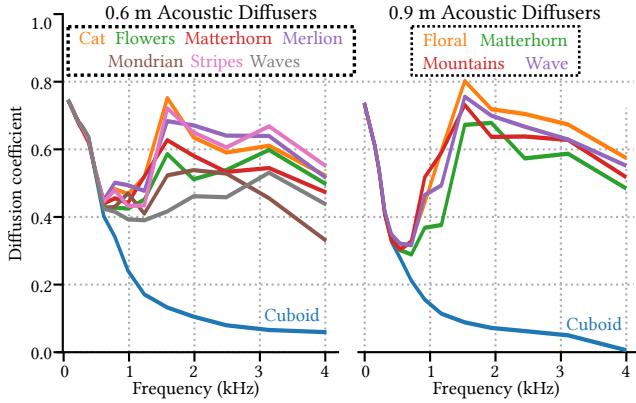


Fig. 12. Diffusion Coefficients. We plot the diffusion coefficients of the optimized diffusers in Figures 10 and 11, and compare their graphs to that of a rectangular cuboid of the corresponding dimensions (blue). For both the 0.6 m (left) and 0.9 m (right) diffusers, we observe that the acoustic performance remains fairly similar regardless of the choice of the guiding image.

consistent performance across different source positions, except at low grazing angles where the effect is slightly diminished. Despite this, we still show marked improvements over a flat surface. We also highlight the fact that our framework places no requirements on the source position. As such, it would be straightforward to optimize

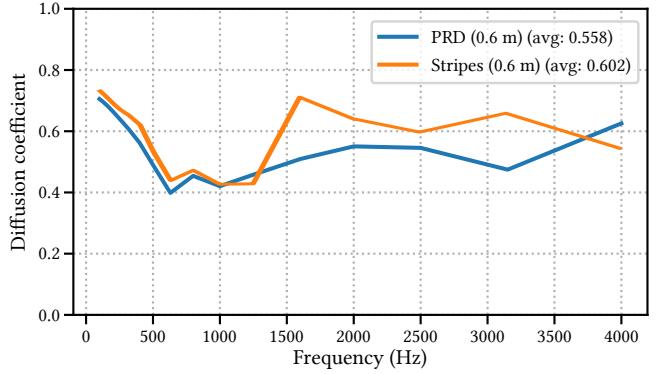


Fig. 13. PRD Comparison. We plot the diffusion curves of the PRD (blue) and the *Stripes* (0.6 m) Acoustic Relief (orange). On average, our optimized *Acoustic Reliefs* achieve greater acoustic diffusion than the PRD.

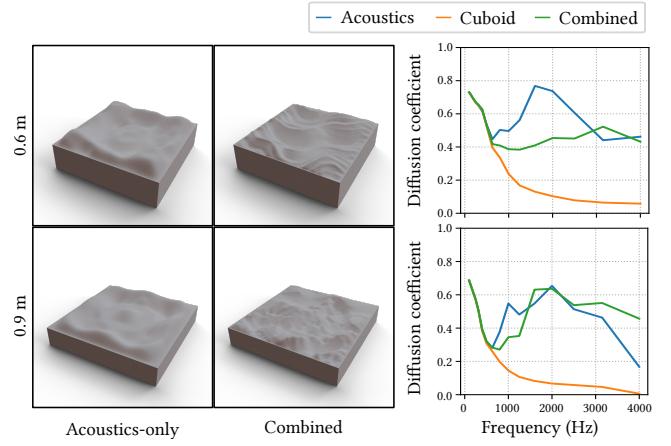


Fig. 14. Acoustics-only Optimization. We optimize a 0.6 m cuboid (Top) and a 0.9 m cuboid (Bottom) for pure acoustic diffusion, and compare their performance (blue) to a cuboid (orange), and our *Acoustic Reliefs* (green).

for source positions specifically tailored to the planned installation locations of our *Acoustic Reliefs*.

Frequency Sampling. During optimization we propose to change the simulated frequency of the diffuser at each iteration to promote a more uniform design across all frequencies. We evaluate the effect of this change in Figure 19. We can observe that an *Acoustic Relief* optimized for a single frequency performs significantly better at the specific frequency. However, to achieve this performance, both the acoustic and visual properties are sacrificed. In contrast, our proposed frequency-sampling optimization produces an overall more balanced diffuser. Moreover, the added freedom in frequency optimization allows for better balancing between acoustic and visual loss, yielding visually more pleasing results.

Fabrication Losses. To evaluate the influence of fabrication on the final design, we set the weights of the fabrication losses (w_s, w_b, w_n) to 0 in Figure 20. We can observe that the resulting acoustic diffusers

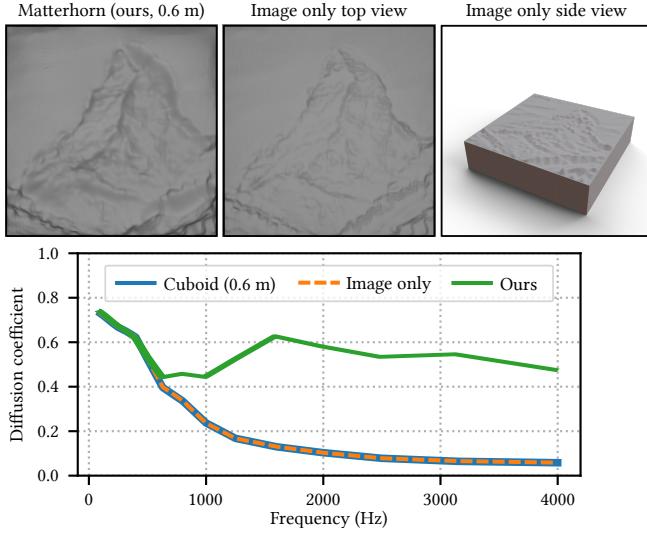


Fig. 15. **Appearance-only Optimization.** We generate heightfield-parameterized reliefs by ignoring the acoustic loss and only optimizing for appearance. The resultant meshes are extremely flat and from their diffusion curves (orange), their acoustical properties are essentially the same as a non-optimized cuboid mesh (blue, or see Fig. 3 for the 0.6 m curve).

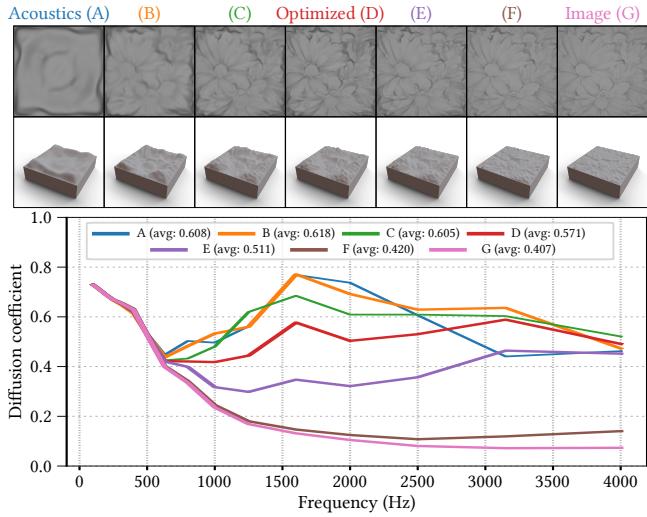


Fig. 16. **Acoustics-Appearance Weights.** We vary the weights of the acoustic and appearance losses, and plot the evolution of the resultant *Acoustic Relief* as well as their diffusion curves. Generally, the behavior is as expected, with increased weights corresponding to increasing the quality of one loss over the other.

have a very high acoustic and visual quality. Unfortunately, these gains are driven by a very spiky height field. Such a heightfield is challenging to produce even with modern fabrication devices. In contrast, acoustic diffusers optimized with our full pipeline demonstrate only a minor loss in acoustic and visual properties while being readily manufacturable with standard hardware.

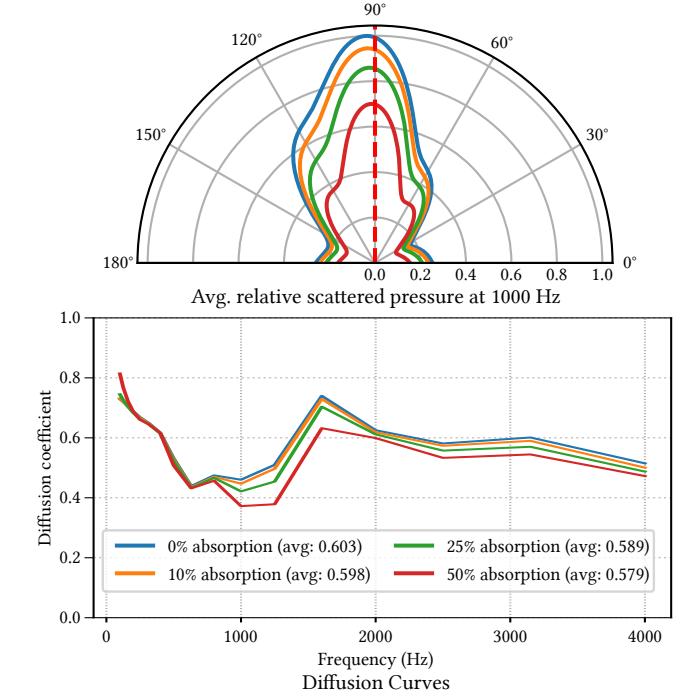


Fig. 17. **Material Absorption.** We use COMSOL® to test the effect of varying material absorption values on the diffusion coefficient. Changing the material absorption value uniformly changes the scattered pressure (Top), so the diffusion coefficient remains relatively unchanged (Bottom).

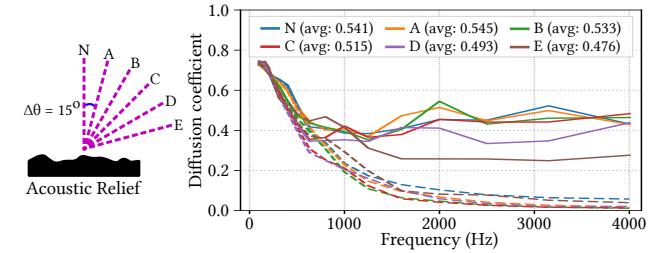


Fig. 18. **Source Positions.** We vary the positions of the source (Left) and plot their diffusion curves (Right, solid). For comparison, we also plot the diffusion curves of a flat surface (Right, dashed). Even though there is a slightly diminished effect at the grazing angle of 75°, our *Acoustic Reliefs* still outperform a flat surface even though they were only optimized under normal incidence.

Post-processing. To further optimize the visual quality of the *Acoustic Reliefs*, we implement a post-processing step (Figure 21). We observe that the visual quality is significantly improved. Due to the regularization term, the post-processing acts similarly to a sharpening filter and improves the fine-resolution details in the relief images. Furthermore, when investigating the acoustic performance, we observe that both the initial and post-processed *Acoustic Reliefs* have almost identical diffusion coefficient across all frequencies.

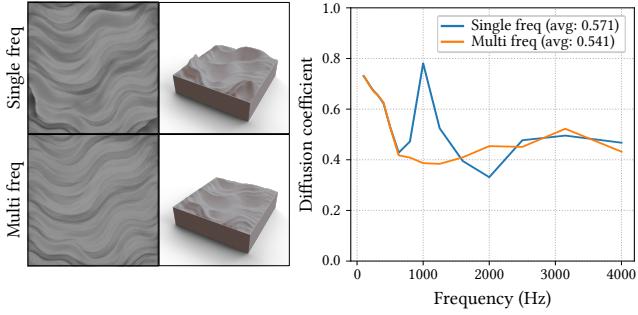


Fig. 19. **Frequency Sampling.** We optimize our diffusers for the 1000 Hz frequency band (Top, blue), and compare it to the diffuser optimized for multiple frequencies (Bottom, orange). The resultant diffuser performs much better at the 1000 Hz frequency band, but doesn't perform as well at other frequencies.

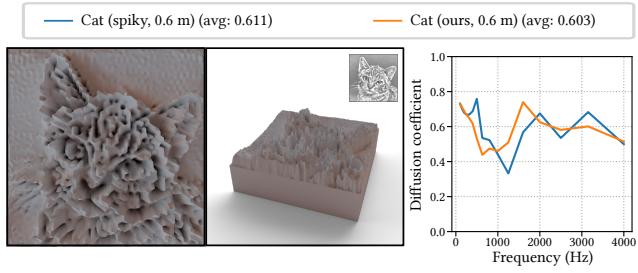


Fig. 20. **Fabrication Losses.** Setting the weights of the fabrication loss to zero results in extremely spiky meshes. The optimization is dominated by the texture loss, resulting in the heightfield texture (Top-right) almost exactly matching the input image.

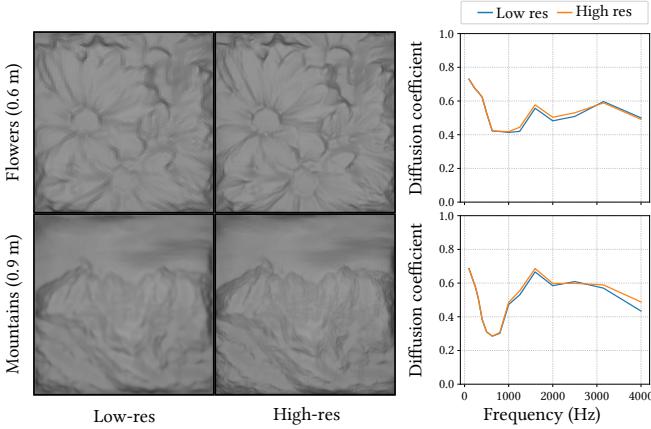


Fig. 21. **Post-processing Details.** The post-processing steps adds detail to the initial low-resolution relief (Left), improving its visual quality (Right). The diffusion curves for the low-resolution relief (blue) and the high-resolution relief (orange) remain relatively similar.

GMSH Visual Differences. As our BEM simulation uses triangular elements (Section 3), we use GMSH [Geuzaine and Remacle 2009]

to mesh a rectangular cuboid into triangles of roughly equal area. Compared to using a rectangular grid of right-angled triangles, not only does this result in a higher quality mesh [Shewchuk 2002], but it also results in subjectively improved appearances. As seen in Figure 22, using a rectangular grid of triangles results in visual artefacts, though admittedly the effect is less evident after the post-processing step previously discussed. Nevertheless, using GMSH enables us to use any arbitrary mesh as a starting base mesh (instead of a rectangular cuboid), allowing for a greater variety in terms of design choices.

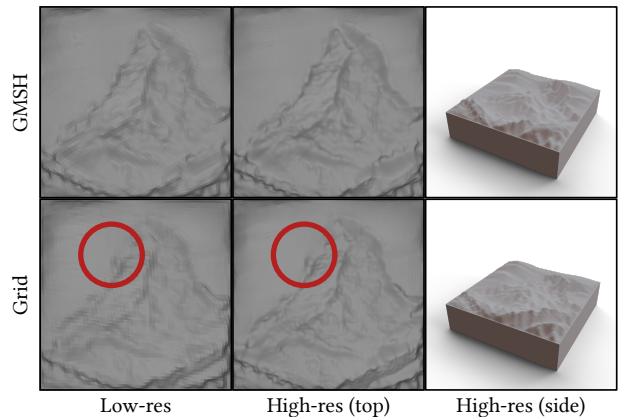


Fig. 22. **GMSH Visual Differences.** We compare the visual results of using GMSH to initialize the triangle mesh (Top) versus using a regular grid of triangles (Bottom). The triangle grid results in less robust mesh generation, though the effect is reduced after the post-processing step.

Lighting. We use a single point light as our lighting condition to simulate a controlled, indoor, environment such as a studio, where the lighting is fixed and does not vary much. However, as our appearance optimization is built upon the Mitsuba renderer [Jakob et al. 2022b], we are able to easily change the lighting conditions. We experiment with using an environment map³ and observe their performance in Figure 23. The lighting conditions have clear influences on the final shape. On one hand, the point light source casts darker shadows, resulting in reliefs that are more pronounced, but at the same time, are more sensitive to the lighting condition and view angle. On the other hand, the reliefs generated using the environment map have a slightly more washed-out appearance, but are less sensitive to the view angle and various lighting scenarios. In either case, both reliefs perform acoustically well, so the choice of the lighting for the renderer should be determined by where the relief would be placed in the real world.

Optimizing Bas Reliefs. Our method involves the joint optimization of both the acoustics and the appearance in the same step. An alternative approach might be to use existing relief generation methods (e.g., as presented in Weyrich et al. [2007]) to first generate a relief, and then use it as an initial guess for the optimization. Starting with an input image, we generate a depth map

³Distant Evening Sun (Hallstatt) by Bernhard Vogl (<http://dativ.at/lightprobes/>)

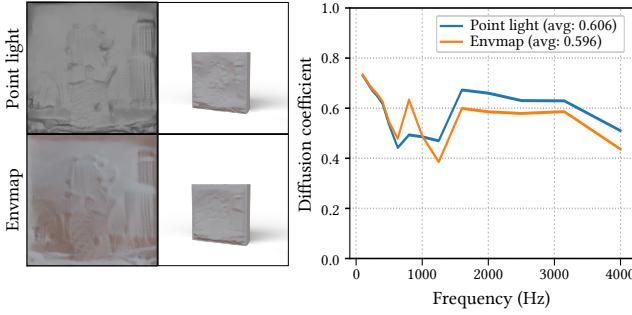


Fig. 23. Different Lighting Condition. We compare the reliefs optimized under a single point light source (blue), versus one optimized under an environment map (orange). The former results in reliefs with darker shadows at a slight cost of generalizability to other view angles. The latter has a more washed out look, but is less sensitive to novel views.

using DepthAnythingV2 [Yang et al. 2024], and use the algorithm presented in Weyrich et al. [2007] to generate the relief sculpture as a heightfield. This is then passed into the optimization routine, where we disregard the CLIP-based appearance loss in favor of a regularization loss that keeps that optimized relief close to the input heightfield. The results are presented in Figure 24.

The visual performance of the bas-reliefs offer significantly sharper contrast than our method. Additionally, the optimized bas-reliefs are also capable sound diffusers. However, our *Acoustic Reliefs* still offer superior acoustic performance around 1 kHz and 3 kHz bands which are associated with human speech and musical instruments, respectively. Moreover, even though our manufacturing losses are employed, bas-relief initialization results in a spiky geometry that is challenging to manufacture. We attribute this to the fact that the initial solution is a local minima where the acoustics and appearance losses dominate the fabrication loss. On the other hand, our method inherently couples all optimization objectives to each other. This helps to avoid such local minima, resulting in smooth surfaces, with superior acoustic properties, suitable for direct manufacturing.

Visual Models. An alternative to the CLIP loss for image-guided optimization is a perceptual similarity metric such as LPIPS [Zhang et al. 2018b]. In Figure 25, we replace the CLIP model with a LPIPS model that uses a VGG backend trained on the ImageNet classification tasks [Simonyan and Zisserman 2015; Zhang et al. 2018b], and only optimize for the appearance. We note that while the LPIPS loss is able to capture the overall shape of the image, the CLIP loss maintains greater detail and is more robust for the image-guided optimization in our pipeline.

Additionally, we compare both losses to a naive pixel-based distance metric. However, due to the underlying differences between a projected image and the rendered relief, the optimization is unable to reach the targeted result, instead resulting in unrecognizable blobs. This shows the need for a flexible loss metric which both the CLIP and LPIPS model can achieve.

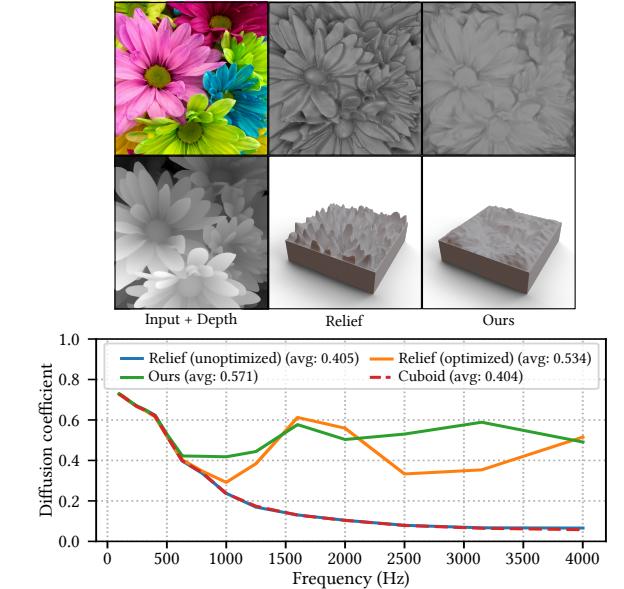


Fig. 24. Bas Reliefs. We use the relief generation method presented in Weyrich et al. [2007] to generate an initial heightmap (blue), optimize its acoustics (orange), and then compare it to our *Acoustic Reliefs* (green).

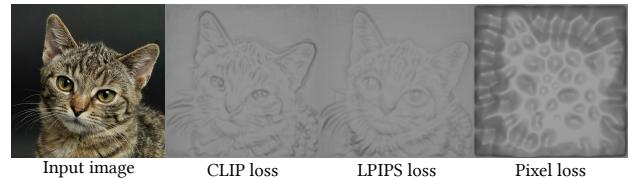


Fig. 25. Visual Models. We show the replacement of the CLIP loss with the LPIPS similarity loss, as well as a naive pixel-based distance loss. The visual results from using the CLIP model is the most robust for our purposes.

5.4 Simulated Audio Previews

To preview the qualitative performance of our acoustic reliefs, we employ a GPU FDTD room acoustics solver (see Appendix D for simulator details and validation) to render room impulse responses (RIRs) in various acoustic treatment settings. We consider two scenarios: (1) reducing *flutter echoes* in a simple corridor, and (2) reducing coloration in a small listening room.

Corridor. For our untreated corridor, we use a rectangular $12 \times 2.7 \times 3.5$ m room and model all surfaces as locally reacting, lightly absorptive material with a frequency-independent normal-incidence absorption coefficient of 0.09. We compare between two treatment settings. In the first setting (“Reliefs”), we place three of our *Acoustic Reliefs* (*Wave Left* (0.6 m), *Cat* (0.9 m), and *Wave Right* (0.6 m)) on the left wall (see Figure 26). In the second setting (“PRDs”), we replicate the setup but place three commercial primitive root diffusers instead. In both settings, we use sound-hard surfaces for the diffusers, while the room walls remain unchanged.

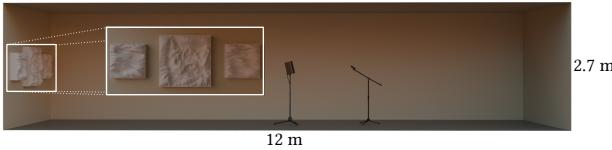


Fig. 26. **Corridor Layout.** Three diffusers are aligned on the left wall of a $12 \times 2.7 \times 3.5$ m room. A close-up of the reliefs is shown.

We run our FDTD simulations on an NVIDIA RTX 4090 (24 GB) GPU, with a cell size of 4 mm and a sampling frequency of 150 kHz. We use a point source placed at the room's center and a point receiver placed 1.5 m directly to the right. To filter out high frequencies with an excess numerical dispersion error, we applied a fourth-order Butterworth low-pass filter with a 10 kHz cutoff frequency. The final audio previews are obtained by convolving the processed RIRs with an anechoic audio recording.

Figure 27 shows spectrograms of a simulated hand clap test. Due to elongated room dimensions and parallel walls, prominent *flutter echoes* form along the 12 m length of the corridor [Cox and D'Antonio 2016]. Treating the left wall with diffusers removes these echoes while maintaining the overall spaciousness of the room, and we observe minimal perceptual difference between our optimized *Acoustic Reliefs* and the commercial PRDs. See the supplemental video for audio comparisons.

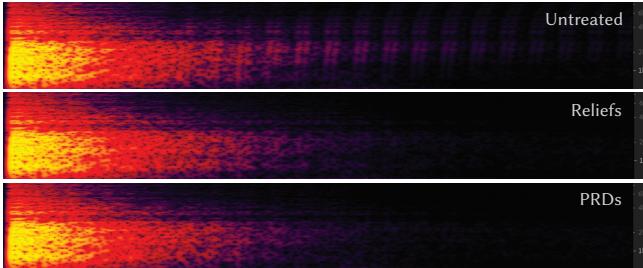


Fig. 27. **Simulated Clap Test Spectrograms.** *Flutter echoes* are prominent in the Untreated spectrogram but disappear after the room is treated with either our *Acoustic Reliefs* or with commercial PRDs.

Listening Room. For our second scenario, we consider a more subjective treatment of a small listening room. We start with an untreated rectangular $3.5 \times 2.7 \times 5$ m room with a desk, stereo speakers, and chair placed near the front wall (see Figure 28). To reduce coloration from early reflections, we place absorbers at the first reflection points (side walls and ceiling), creating a *reflection free zone* [Cox and D'Antonio 2016]. However, without treating the rear wall, strong parallel reflections may still form between the rear and the front, so we then place three of our *Acoustic Reliefs* on the rear wall to break up echoes while maintaining spaciousness.

We simulate audio previews at each stage of treatment (untreated, absorbers-only, fully treated) using identical settings as before. Surfaces are idealized as locally reacting, frequency-independent materials with varying absorption coefficients. We use two point sources



Fig. 28. **Listening Room Layout.** Our fully treated $3.5 \times 2.7 \times 5$ m listening room features absorbers placed at the first reflection points (side walls and ceiling) along with diffusers placed on the rear wall.

placed at the front of each loudspeaker, along with a point receiver located directly above the chair.

Perceptually significant differences can be heard at each stage of treatment, with the fully treated room having reduced coloration and echo. See the supplemental video for audio comparisons. We note that due to perfect rectangular room symmetries and idealized materials, simulated echoes may be stronger than they would be in analogous real-world settings.

5.5 Fabricated Prototypes

We fabricate full-scale versions of the *WaveLeft* (0.6 m) and *WaveRight* (0.6 m) *Acoustic Reliefs*. We also fabricate scaled-down (0.2 m) versions of the *Cat* (0.6 m), *Flowers* (0.6 m), *Mondrian* (0.6 m), *Floral Pattern* (0.9 m), *Matterhorn* (0.9 m) and *Wave* (0.9 m) *Acoustic Reliefs*. We use a Bamboo Lab X1 Carbon 3D printer to print the reliefs. Due to the limited print size, we split the full-scale reliefs into 9 pieces, and print them separately before gluing the pieces together. We also fabricate *EnvCat* (0.6 m), an *Acoustic Relief* optimized under the environment map used in Figure 23, and use it to physically validate our simulations (Figure 29).

Acoustic Properties. To evaluate the acoustic properties, we measure a Gaussian impulse signal reflected from the *EnvCat* (0.6 m) *Acoustic Relief*. The measurement setup consists of the acoustic diffuser, a speaker, and a microphone (Figure 29 left). The speaker is placed 2m away from the diffuser and the microphones are placed at the same distance facing the diffuser from 0° , 45° , and 60° . We utilized a laser positioning tool to ensure proper placement of all elements. To avoid other sources of reflections the setup was built on tall stands and we ensured that the ground and other walls were at least 2.5m away. We then played the impulse and captured the reflected sound either from a reference flat wall or from our diffuser. To validate the results, we isolate the first reflections, and compare them against FDTD simulations of the same setup (Figure 29 right). Since the simulation and real world have different sound magnitudes, we rescale the measurements to match the simulated amplitudes.

Overall, we note an excellent match of our measurements with the simulation. The direct reflection at 0° shows a reduced impulse

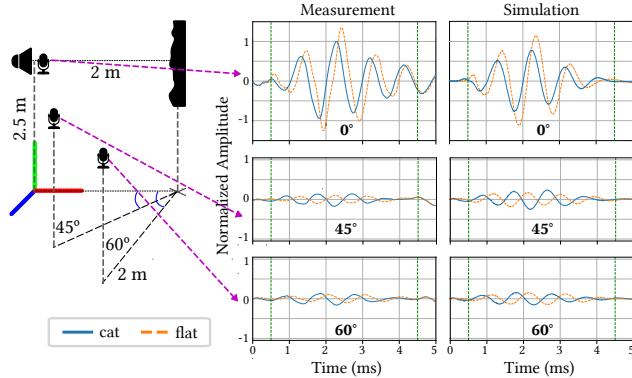


Fig. 29. Measurements. We measure *EnvCat* (0.6 m) using the experimental setup detailed above (Left). We play a Gaussian pulse centered around 1000 Hz and record the resultant waveforms using microphones placed at 0° (Top-right), 45° (Middle-right), and 60° (Bottom-right). We isolate the first reflections. To match the amplitudes between simulation and measurements we apply a uniform scaling factor that matches the peak-to-peak response of the simulated and measured flat plane reflections at 60°.

amplitude that follows our prediction. In contrast, at 45° we observe that the diffuser has a stronger reflection than a flat plane. Lastly, at 60° the angle is too steep and both the diffuser and flat plane provide similar reflections. We conclude that the diffuser design was effective in redistributing the acoustic pressure away from the specular direction. We also highlight the excellent match between simulated and measured results that further confirms the accuracy of our numerical modeling.

Reduced Diffusion Coefficient. We additionally conduct another experiment to further validate the effectiveness of our *Acoustic Reliefs*, comparing it with other diffuser geometries. Specifically, we compare *WaveLeft* (0.6 m) with a 3D printed PRD, as well as a flat surface obtained by flipping the diffuser. We measure a reduced version of the diffusion coefficient presented in the ISO standard [ISO 2012], as measuring the full hemispherical diffusion coefficient is infeasible without a full industrial setup. The measurement setup (Figure 30 left) consists of the three diffusers, a speaker, and a microphone. The speaker is fixed at a position of 3 m away from and directly facing the measured diffuser. The microphones are moved at angular intervals of 11.25° with a radius of 1.7 m, ranging from -90° to 90°.

The impulse responses are computed using a logarithmic sine sweep [Farina 2000] from 100 to 4000 Hz with Room EQ Wizard⁴. The resultant sound pressure levels (SPLs) are then averaged into one-third octave bands [Cox and D'Antonio 2016; ISO 2012], and the diffusion coefficient formula [ISO 2012, Eq. 5] is used to compute the coefficient values. The resultant curves are shown on the right in Figure 30.

In accordance to our expectations, we observe a significant improvement in diffusion with both the PRD and our *Acoustic Reliefs* as compared to a flat, rectangular surface. Additionally, we note that the measurements also show that our *Acoustic Reliefs* outperforms

⁴<https://www.roomeqwizard.com/>

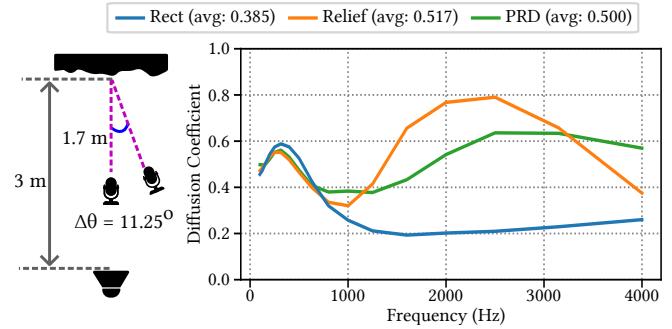


Fig. 30. Reduced Diffusion Coefficient. We measure *WaveLeft* (0.6 m), a PRD, and a flat, rectangular surface using the experimental setup detailed above (Left). We play a logarithmic sine sweep from 100 to 4000 Hz and record the resultant waveforms using microphones placed 11.25° apart from -90° to 90° (Right). We isolate the first reflections, compute the SPLs, and plot the obtained diffusion curves.

the acoustic performance of the PRD is most frequency bands, which is in-line with the simulation results presented in Figure 13.

Visual Properties. We evaluate the full-sized diffusers, *WaveLeft* (0.6 m) and *WaveRight* (0.6 m), by mounting them on an office wall (Figure 31). The smaller *Acoustic Reliefs* are captured under a light-box to mimic the controlled point light source used during the optimization. The resultant images are shown in Figure 32. We observe that the reliefs closely follow the input images and are readily recognizable.



Fig. 31. Office Diffusers. We fabricate two of our 0.6 m *Acoustic Reliefs*: *WaveLeft* (0.6 m) and *WaveRight* (0.6 m), and mount them on an office wall. For scale, the length of the office wall is 2.1m.

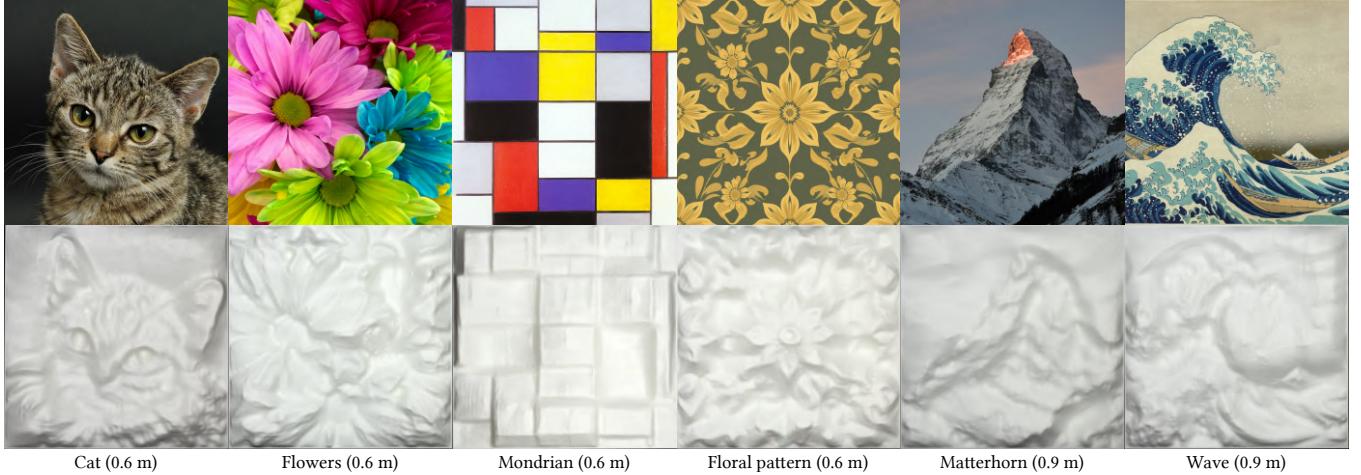


Fig. 32. **Fabricated Prototypes.** In order to ease fabrication, we scale our *Acoustic Reliefs* down to 0.2 m, and compare the 3D printed versions (Bottom) to their guiding images (Top).

6 LIMITATIONS AND FUTURE WORK

In this section we discuss the limitations of our system and propose several promising directions of future work.

Kirchoff Approximation. Though sufficiently accurate, the full BEM pipeline is fairly slow, especially for larger meshes, requiring the computation of a large, non-sparse matrix, as well as a matrix solve (or two, in the backward pass). However, if accuracy is not required, the Kirchoff approximation [Kirchhoff 1883; Ogilvy 1986; Pignier et al. 2015; Tsingos et al. 2007] could be used instead to estimate first-order sound reflections. This would only require a single matrix multiplication (Equation 6), and would be much faster to compute compared to the full BEM. We propose using this approximation as an initialization step, computing a result which can be used as the initial guess for the main optimization pipeline, thus improving the main pipeline’s convergence and reducing the number of iterations needed to obtain a satisfactory diffuser.

General Acoustics. Our system is inherently tuned for the optimization of acoustic diffusers, and as such several assumptions made in the simulation domain reflect this focus. Firstly, we use the vanilla formulation of the BEM. However, when used to solve the Helmholtz equation in exterior domains, this formulation is susceptible to so-called *fictitious frequencies* [Chen et al. 2009]. While Cox and D’Antonio [2016] mentions that these fictitious frequencies are usually a nonissue for diffusers due to their geometry, it may be interesting to explore other formulations of the BEM that are able to deal with the fictitious frequency problem. Alternative formulations include the CHIEF method [Schenck 1968], or the Burton-Miller formulation [Burton and Miller 1971], although the key problem would be to ensure differentiability.

Secondly, we only consider acoustically hard surfaces (i.e., $\frac{\partial u}{\partial n} = 0$) since it does not significantly affect the diffusion coefficient, but it might be interesting to expand the allowed boundary conditions to locally reacting or even absorbtive surfaces. If properly coupled with

the gradient computation, this could open avenues for gradient-based material optimization of acoustic diffusers and absorbers.

Our FDTD audio previews were simulated without frequency-dependent air absorption. While important for accurate attenuation of high-frequencies, modeling air absorption in FDTD incurs significant memory and computational costs, although such effects could be approximated as a post-process [Hamilton 2021]. Furthermore, our simulations assume locally reacting frequency-independent materials, but general impedance models [Bilbao et al. 2016] should be used when simulating more complex settings.

Last but not least, it is also possible to generalize the heightfield parameterization to any arbitrary mesh representation. Coupled with the re-parameterization technique introduced by Nicolet et al. [2021], one could perform acoustics-based mesh optimization while avoiding self-intersections and maintaining a high-quality mesh.

Visual Appearance. Generally, we want to use images with distinct features. This allows the acoustics optimization to have more flexibility in introducing the peaks and troughs that result in better diffusion. For instance, in Figure 12, *Mondrian (0.6m)* underperforms relative to *Cat (0.6m)* due to its large flat color regions.

Currently, we are focusing on reliefs with self-shadowing effects. Extending this to multicolor or stylistically guided reliefs is a promising direction for future work.

7 CONCLUSION

In this paper, we have introduced a system for optimizing and generating *Acoustic Reliefs*: sound diffusers that not only perform well acoustically but whose appearances can also be guided by user-provided images. We developed a differentiable boundary element simulator that solves the Helmholtz equation, allowing gradient-based optimization of acoustic diffusers. We also coupled the solver to a differentiable renderer and jointly optimized the appearance of the diffusers at the same time. We validated our differentiable simulator, generated various image-guided reliefs, and also physically

fabricated several examples of our acoustic reliefs. To our knowledge, our system is the first to perform gradient-based optimization for acoustic diffusers.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their comments and suggestions. We would also like to thank Desmond Liu for their feedback on the manuscript, Achilleas Xydis for the discussion on acoustic diffusers, Bharath Seshadri, Anna Maria Eggler, and Henry Welch for their help in the fabrication of the *Acoustic Reliefs*, as well as Rafael Bischof, Pengbin Tang, and Yuhang Zhang for their help in the measurement of the *Acoustic Reliefs*.

The guiding images for the *Stripes (0.6 m)*, *Waves (0.6 m)*, *WaveLeft (0.6 m)*, *WaveRight (0.6 m)*, and *Floral pattern (0.9 m)* *Acoustic Reliefs* shown in Figures 1, 4, 10, 11, and 31 were generated using Generative AI.

Partial support for this work was provided by the Office of Naval Research (Grant N000142512024) for Doug L. James and Kangrui Xue.

REFERENCES

- AFMG. 2024. EASE 5 Third Edition. <https://www.afmg.eu/en/ease>. Accessed: 2024-12-08.
- Tomas Akenine-Möller. 2005. Fast 3D Triangle-Box Overlap Testing. In *ACM SIGGRAPH 2005 Courses* (Los Angeles, California) (*SIGGRAPH '05*). Association for Computing Machinery, New York, NY, USA, 8–es. <https://doi.org/10.1145/1198555.1198747>
- Marc Alexa and Wojciech Matusik. 2010. Reliefs as Images. *ACM Transactions on Graphics (TOG)* 29, 4 (July 2010), 60:1–60:7. <https://doi.org/10.1145/1778765.1778797>
- Mario Bebendorf. 2000. Approximation of Boundary Element Matrices. *Numer. Math.* 86, 4 (Oct. 2000), 565–589. <https://doi.org/10.1007/PL00005410>
- Thomas Beck. 1994. Automatic Differentiation of Iterative Processes. *J. Comput. Appl. Math.* 50, 1 (May 1994), 109–118. [https://doi.org/10.1016/0377-0427\(94\)90293-3](https://doi.org/10.1016/0377-0427(94)90293-3)
- P.N. Belhumeur, D.J. Kriegman, and A.L. Yuille. 1997. The Bas-Relief Ambiguity. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Comput. Soc, San Juan, Puerto Rico, 1060–1066. <https://doi.org/10.1109/CVPR.1997.609461>
- Jean-Pierre Berenger. 1994. A Perfectly Matched Layer for the Absorption of Electromagnetic Waves. *J. Comput. Phys.* 114, 2 (Oct. 1994), 185–200. <https://doi.org/10.1006/jcph.1994.1159>
- Amit Bermano, Ilya Baran, Marc Alexa, and Wojciech Matusik. 2012. ShadowPix: Multiple Images from Self Shadowing. *Computer Graphics Forum* 31, 2pt3 (May 2012), 593–602. <https://doi.org/10.1111/j.1467-8659.2012.03038.x>
- Stefan Bilbao. 2009. *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics*. Wiley Publishing.
- Stefan Bilbao. 2012. Optimized FDTD Schemes for 3-D Acoustic Wave Propagation. *IEEE Transactions on Audio, Speech, and Language Processing* 20, 5 (July 2012), 1658–1663. <https://doi.org/10.1109/TASL.2012.2186806>
- Stefan Bilbao, Brian Hamilton, Jonathan Botts, and Lauri Savioja. 2016. Finite Volume Time Domain Room Acoustics Simulation under General Impedance Boundary Conditions. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 24, 1 (2016), 161–173. <https://doi.org/10.1109/TASLP.2015.2500018>
- Ingolf Bork. 2005a. Report on the 3rd Round Robin on Room Acoustical Computer Simulation – Part I: Measurements. *ACTA Acustica united with Acustica* 91 (2005), 740 – 752.
- Ingolf Bork. 2005b. Report on the 3rd Round Robin on Room Acoustical Computer Simulation – Part II: Calculations. *ACTA Acustica united with Acustica* 91 (2005), 753 – 763.
- Steffen Börn, Lars Grasedyck, and Wolfgang Hackbusch. 2003. Introduction to Hierarchical Matrices with Applications. *Engineering Analysis with Boundary Elements* 27, 5 (May 2003), 405–422. [https://doi.org/10.1016/S0955-7997\(02\)00152-2](https://doi.org/10.1016/S0955-7997(02)00152-2)
- D. Botteldooren. 1994. Acoustical Finite-difference Time-domain Simulation in a Quasi-Cartesian Grid. *The Journal of the Acoustical Society of America* 95, 5 (May 1994), 2313–2319. <https://doi.org/10.1121/1.409866>
- A. J. Burton and G. F. Miller. 1971. The Application of Integral Equation Methods to the Numerical Solution of Some Exterior Boundary-Value Problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 323, 1553 (1971), 201–210. [jstor:78022](https://doi.org/10.1088/0305-4470/323/1553/201)
- Faisal Caeiro, Carlo Sovardi, Kilian Förner, and Wolfgang Polifke. 2017. Shape Optimization of a Helmholtz Resonator Using an Adjoint Method. *International Journal of Spray and Combustion Dynamics* 9 (May 2017), 175682771770357. <https://doi.org/10.1177/1756827717703576>
- J. T. Chen, S. R. Lin, and J. J. Tsai. 2009. Fictitious Frequency Revisited. *Engineering Analysis with Boundary Elements* 33, 11 (Nov. 2009), 1289–1301. <https://doi.org/10.1016/j.enganabound.2009.06.001>
- Xiang Chen, Changxi Zheng, Weiwei Xu, and Kun Zhou. 2014. An Asymptotic Numerical Method for Inverse Elastic Shape Design. *ACM Transactions on Graphics (TOG)* 33, 4, Article 95 (July 2014), 11 pages. <https://doi.org/10.1145/2601097.2601189>
- P. Cignoni, C. Montani, and R. Scopigno. 1997. Computer-Assisted Generation of Bas-and High-Reliefs. *Journal of Graphics Tools* 2, 3 (Jan. 1997), 15–28. <https://doi.org/10.1080/10867651.1997.10487476>
- Robert D Ciskowski and Carlos Alberto Brebbia. 1991. *Boundary Element Methods in Acoustics*. Vol. 20. Springer.
- COMSOL. 2024. COMSOL Multiphysics®v. 6.2. <https://www.comsol.com/>. Accessed: 2024-12-09.
- Trevor Cox. 1995. The Optimization of Profiled Diffusers. *Journal of The Acoustical Society of America - J ACoust SOC AMER* 97 (May 1995), 2928–2936. <https://doi.org/10.1121/1.412972>
- Trevor Cox and Peter D'Antonio. 2016. *Acoustic Absorbers and Diffusers, Third Edition: Theory, Design and Application*. CRC Press.
- Trevor J. Cox. 1994. Predicting the Scattering from Reflectors and Diffusers Using Two-Dimensional Boundary Element Methods. *The Journal of the Acoustical Society of America* 96, 2 (Aug. 1994), 874–878. <https://doi.org/10.1121/1.410262>
- Peter D'Antonio. 1995. Performance Evaluation of Optimized Diffusors. *The Journal of the Acoustical Society of America* 97, 5 (May 1995), 2937–2941. <https://doi.org/10.1121/1.413102>
- Peter D'Antonio and Trevor J. Cox. 2000. Diffusor Application in Rooms. *Applied Acoustics* 60, 2 (June 2000), 113–142. [https://doi.org/10.1016/S0003-682X\(99\)00054-7](https://doi.org/10.1016/S0003-682X(99)00054-7)
- Peter D'Antonio, Luiz Ferraz Alvim, Rinaldi Petroloti, Eric Brandao, and Paulo Mareze. 2022. A Cloud-Based 3D BEM Framework for Predicting the Diffusion Coefficient. In *Proceedings of Meetings on Acoustics*, Vol. 46, 015003. <https://doi.org/10.1121/2.0001645>
- P. D'Antonio and J. Konnert. 1984. The Reflection Phase Grating Diffusor: Design Theory and Application. *Journal of The Audio Engineering Society* (April 1984).
- B. A. de Jong and P. M. van den Berg. 1980. Theoretical Design of Optimum Planar Sound Diffusers. *The Journal of the Acoustical Society of America* 68, 4 (Oct. 1980), 1154–1159. <https://doi.org/10.1121/1.385000>
- Christina Dillon. 2013. Designing, Modeling, Constructing, and Testing a Flat Panel Speaker and Sound Diffuser for a Simulator.
- Tao Du, Kui Wu, Pingchuan Ma, Sebastian Wah, Andrew Spielberg, Daniela Rus, and Wojciech Matusik. 2021. DiffPD: Differentiable Projective Dynamics. *ACM Transactions on Graphics (TOG)* 41, 2, Article 13 (Nov. 2021), 21 pages. <https://doi.org/10.1145/3490168>
- Tao Du, Kui Wu, Andrew Spielberg, Wojciech Matusik, Bo Zhu, and Eftychios Sifakis. 2020. Functional Optimization of Fluidic Devices with Differentiable Stokes Flow. *ACM Transactions on Graphics (TOG)* 39, 6 (2020), 1–15.
- Marek Dvorožník, Saman Sepehri Nejad, Ondřej Jamriška, Alec Jacobson, Ladislav Kavan, and Daniel Sýkora. 2018. Seamless Reconstruction of Part-Based High-Relief Models from Hand-Drawn Images. In *Proceedings of the Joint Symposium on Computational Aesthetics and Sketch-Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering (Expressive '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3229147.3229153>
- Graeme Fairweather, Andreas Karageorghis, and P. A. Martin. 2003. The Method of Fundamental Solutions for Scattering and Radiation Problems. *Engineering Analysis with Boundary Elements* 27, 7 (July 2003), 759–769. [https://doi.org/10.1016/S0955-7997\(03\)00017-1](https://doi.org/10.1016/S0955-7997(03)00017-1)
- A. Farina. 2000. Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. *Journal of The Audio Engineering Society* (Feb. 2000).
- Herbert Fischer. 1991. Special Problems in Automatic Differentiation. *Automatic Differentiation of Algorithms: Theory, Implementation, and Application* (1991), 43–50.
- Christophe Geuzaine and Jean-François Remacle. 2009. Gmsh: A 3-D Finite Element Mesh Generator with Built-in Pre- and Post-Processing Facilities. *Internat. J. Numer. Methods Engrg.* 79, 11 (2009), 1309–1331. <https://doi.org/10.1002/nme.2579>
- E. Gravier, M. Kleiner, Bengt-Inge Dalenbäck, and P. Svensson. 1996. Experimental Auralization of Car Audio Installations. *Journal of The Audio Engineering Society* (Oct. 1996).
- Lars Grasedyck and Wolfgang Hackbusch. 2003. Construction and Arithmetics of \mathcal{H} -Matrices. *Computing* 70, 4 (Aug. 2003), 295–334. <https://doi.org/10.1007/s00607-003-0019-1>
- Nail A. Gumerov and Ramani Duraiswami. 2009. A Broadband Fast Multipole Accelerated Boundary Element Method for the Three Dimensional Helmholtz Equation. *The Journal of the Acoustical Society of America* 125, 1 (Jan. 2009), 191–205. <https://doi.org/10.1121/1.3021297>

- Wolfgang Hackbusch. 1999. A Sparse Matrix Arithmetic Based on \mathcal{H} -Matrices. Part I: Introduction to \mathcal{H} -Matrices. *Computing* 62, 2 (April 1999), 89–108. <https://doi.org/10.1007/s006070050015>
- Brian Hamilton. 2021. Adding Air Attenuation to Simulated Room Impulse Responses: A Modal Approach. In *2021 Immersive and 3D Audio: from Architecture to Automotive (3DA)*. 1–10. <https://doi.org/10.1109/3DA48870.2021.9610871>
- Tristan Hargreaves, Trevor Cox, Yiu Lam, and Peter D'Antonio. 2000. Surface Diffusion Coefficients for Room Acoustics: Free-Field Measures. *The Journal of the Acoustical Society of America* 108 (Nov 2000), 1710–20. <https://doi.org/10.1121/1.1310192>
- Robert W. Hart. 1951. Sound Scattering of a Plane Wave from a Nonabsorbing Sphere. *The Journal of the Acoustical Society of America* 23, 3 (May 1951), 323–329. <https://doi.org/10.1121/1.1906766>
- Katsushika Hokusai. 1831. The Great Wave off Kanagawa.
- Berthold K. P. Horn. 1970. Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View. (Nov 1970).
- Yuanming Hu, Luke Anderson, Tzu-Mao Li, Qi Sun, Nathan Carr, Jonathan Ragan-Kelley, and Frédéric Durand. 2019a. DiffTaichi: Differentiable Programming for Physical Simulation. *arXiv preprint arXiv:1910.00935* (2019).
- Yuanming Hu, Jiancheng Liu, Andrew Spielberg, Joshua B Tenenbaum, William T Freeman, Jiajun Wu, Daniela Rus, and Wojciech Matusik. 2019b. Chainqueen: A Real-Time Differentiable Physical Simulator for Soft Robotics. In *2019 International conference on robotics and automation (ICRA)*. IEEE, 6265–6271.
- Jan Hueckelheim, Harshitha Menon, William S. Moses, Bruce Christianson, Paul Hovland, and Laurent Hascoet. 2023. A Short Review of Automatic Differentiation Pitfalls in Scientific Computing. In *ICML 2023 Workshop on Differentiable Almost Everything: Differentiable Relaxations, Algorithms, Operators, and Simulators*.
- ISO. 2012. ISO 17497-2:2012: *Acoustics – Sound-scattering Properties of Surfaces Part 2: Measurement of the Directional Diffusion Coefficient in a Free Field*. Technical Report. International Organization for Standardization.
- Wenzel Jakob, Sébastien Speirer, Nicolas Roussel, Merlin Nimier-David, Delio Vicini, Tizian Zeltner, Baptiste Nicolet, Miguel Crespo, Vincent Leroy, and Ziyi Zhang. 2022b. *Mitsuba 3 Renderer*. <https://mitsuba-renderer.org>.
- Wenzel Jakob, Sébastien Speirer, Nicolas Roussel, and Delio Vicini. 2022a. DRJIT: A Just-in-Time Compiler for Differentiable Rendering. *ACM Transactions on Graphics (TOG)* 41, 4 (July 2022), 124:1–124:19. <https://doi.org/10.1145/3528223.3530099>
- Xutong Jin, Chenxi Xu, Ruohan Gao, Jiajun Wu, Guoping Wang, and Sheng Li. 2024. DiffSound: Differentiable Modal Sound Rendering and Inverse Rendering for Diverse Inference Tasks. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers '24*. ACM, Denver CO USA, 1–12. <https://doi.org/10.1145/3641519.3657493>
- Steven G Johnson. 2021. Notes on Adjoint Methods for 18.335. *Introduction to Numerical Methods* (2021).
- Hiroharu Kato, Deniz Beker, Mihai Morariu, Takahiro Ando, Toru Matsuoka, Wadim Kehl, and Adrien Gaidon. 2020. Differentiable Rendering: A Survey. <https://doi.org/10.48550/arXiv.2006.12057> [cs]
- Olga Khrystolavlenko and Raimondas Gribliauskas. 2017. Simulation of Room Acoustics using COMSOL Multiphysics. <https://doi.org/10.3846/aainz.2017.06>
- Diederik P. Kingma and Jimmy Ba. 2017. Adam: A Method for Stochastic Optimization. <https://doi.org/10.48550/arXiv.1412.6980> [cs]
- G. Kirchhoff. 1883. Zur Theorie Der Lichtstrahlen. *Annalen der Physik* 254, 4 (1883), 663–695. <https://doi.org/10.1002/andp.18832540409>
- Stephen Kirkup. 2007. *The Boundary Element Method in Acoustics*. Vol. 8. Integrated Sound Software.
- Michael Kolomenkin, George Leifman, Ilan Shimshoni, and Ayallet Tal. 2011. Reconstruction of Relief Objects from Line Drawings. In *CVPR 2011*. 993–1000. <https://doi.org/10.1109/CVPR.2011.5995643>
- A. Krokstad, S. Strom, and S. Sørsdal. 1968. Calculating the Acoustical Room Response by the Use of a Ray Tracing Technique. *Journal of Sound and Vibration* 8, 1 (July 1968), 118–125. [https://doi.org/10.1016/0022-460X\(68\)90198-3](https://doi.org/10.1016/0022-460X(68)90198-3)
- Wojciech Kupnicki. 2019. Brown Tabby Cat. <https://www.pexels.com/photo/brown-tabby-cat-2071882/>
- Soo Young Lee, Choon-Su Park, Keonhyeok Park, Hyung Jin Lee, and Seungchul Lee. 2023. A Physics-informed and Data-Driven Deep Learning Approach for Wave Propagation and Its Scattering Characteristics. *Engineering with Computers* 39, 4 (Aug. 2023), 2609–2625. <https://doi.org/10.1007/s00366-022-01640-7>
- Yuqing Li and Brian Hamilton. 2020. Modelling and Optimizing Acoustic Diffusers Using Finite-Difference Time-Domain Method. In *Proceedings of DAGA2020*.
- Qing-Huo Liu and Jianping Tao. 1997. The Perfectly Matched Layer for Acoustic Waves in Absorptive Media. *The Journal of the Acoustical Society of America* 102, 4 (10 1997), 2072–2082. <https://doi.org/10.1121/1.419657> arXiv:https://pubs.aip.org/asa/jasa/article-pdf/102/4/2072/8083331/2072_1_online.pdf
- Shiguang Liu and Dinesh Manocha. 2021. Sound Synthesis, Propagation, and Rendering: A Survey. <https://doi.org/10.48550/arXiv.2011.05538> arXiv:2011.05538 [cs]
- Andrew Lock and Damien Holloway. 2016. Boundary Element Modelling of a Novel Simple Enhanced Bandwidth Schroeder Diffuser Offering Comparable Performance to a Fractal Design. *Acoustics Australia* 44 (March 2016). <https://doi.org/10.1007/s40857-016-0049-4>
- Steffen Marburg. 2018. Boundary Element Method for Time-Harmonic Acoustic Problems. In *Computational Acoustics*, Manfred Kaltenbacher (Ed.). Springer International Publishing, Cham, 69–158. https://doi.org/10.1007/978-3-319-59038-7_3
- Ravish Mehra, Nikunj Raghuvanshi, Lakulish Antani, Anish Chandak, Sean Curtis, and Dinesh Manocha. 2013. Wave-Based Sound Propagation in Large Open Scenes Using an Equivalent Source Formulation. *ACM Transactions on Graphics (TOG)* 32, 2 (April 2013), 19:1–19:13. <https://doi.org/10.1145/2451236.2451245>
- Aaron Meurer, Christopher P. Smith, Mateusz Paprocki, Ondřej Čertík, Sergey B. Kirpichev, Matthew Rocklin, AMIT Kumar, Sergiu Ivanov, Jason K. Moore, Sartaj Singh, Thilina Rathnayake, Sean Vig, Brian E. Granger, Richard P. Muller, Francesco Bonazzi, Harsh Gupta, Shivam Vats, Fredrik Johansson, Fabian Pedregosa, Matthew J. Curry, Andy R. Terrel, Štěpán Roučka, Ashutosh Saboo, Isuru Fernando, Sumith Kulal, Robert Cimrman, and Anthony Scopatz. 2017. SymPy: Symbolic Computing in Python. *PeerJ Computer Science* 3 (Jan. 2017), e103. <https://doi.org/10.7717/peerj.cs.103>
- Piet Mondrian. 1923. Composition A.
- Nicolas Morales and Dinesh Manocha. 2016. Efficient Wave-Based Acoustic Material Design Optimization. *Computer-Aided Design* 78 (Sept. 2016), 83–92. <https://doi.org/10.1016/j.cad.2016.05.011>
- Mohammad Sina Nabizadeh, Ravi Ramamoorthi, and Albert Chern. 2021. Kelvin Transformations for Simulations on Infinite Domains. *ACM Transactions on Graphics (TOG)* 40, 4 (July 2021), 97:1–97:15. <https://doi.org/10.1145/3450626.3459809>
- Baptiste Nicolet, Alec Jacobson, and Wenzel Jakob. 2021. Large Steps in Inverse Rendering of Geometry. *ACM Transactions on Graphics (TOG)* 40, 6 (Dec. 2021), 248:1–248:13. <https://doi.org/10.1145/3478513.3480501>
- Timo Niedermann. 2021. Scenic View of Matterhorn, Switzerland. <https://www.pexels.com/photo/scenic-view-of-matterhorn-switzerland-10570243/>
- Odeon A/S. 2024. Odeon Room Acoustics Software. <https://odeon.dk/>. Accessed: 2024-12-08.
- J. A. Ogilvy. 1986. An Estimate of the Accuracy of the Kirchhoff Approximation in Acoustic Wave Scattering from Rough Surfaces. *Journal of Physics D: Applied Physics* 19, 11 (Nov. 1986), 2085. <https://doi.org/10.1088/0022-3727/19/11/008>
- Hao Peng, Lin Lu, Lin Liu, Andrei Sharf, and Baoquan Chen. 2019. Fabricating QR Codes on 3D Objects Using Self-Shadows. *Computer-Aided Design* 114, C (Sept. 2019), 91–100. <https://doi.org/10.1016/j.cad.2019.05.029>
- Ken Perlin. 2002. Improving Noise. In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '02)*. Association for Computing Machinery, New York, NY, USA, 681–682. <https://doi.org/10.1145/566570.566636>
- Maxine Perroni-Scharf and Szymon Rusinkiewicz. 2023. Constructing Printable Surfaces with View-Dependent Appearance. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Proceedings*. ACM, Los Angeles CA USA, 1–10. <https://doi.org/10.1145/3588432.3591526>
- Nicolas J. Pignier, Ciarán J. O'Reilly, and Susann Boij. 2015. A Kirchhoff Approximation-Based Numerical Method to Compute Multiple Acoustic Scattering of a Moving Source. *Applied Acoustics* 96 (Sept. 2015), 108–117. <https://doi.org/10.1016/j.apacoust.2015.03.016>
- Adam Pilch. 2021. Optimized Diffusers for Shoe-Box Shaped Performance Halls. *Applied Acoustics* 178 (July 2021), 108019. <https://doi.org/10.1016/j.apacoust.2021.108019>
- Domingo Quiles, Ignacio Rodríguez-Rodríguez, and José-Victor Rodríguez. 2024. Design of Optimal Sound Absorbers Using Acoustic Diffusers for Multipurpose Auditoriums. *Acoustics* 6 (March 2024), 219–239. <https://doi.org/10.3390/acad6010012>
- Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning Transferable Visual Models From Natural Language Supervision. <https://arxiv.org/abs/2103.00020v1>
- Ramani Ramakrishnan. 2010. Validation of COMSOL Multiphysics and Acoustical Performance of Splitter-Silencers. *Canadian Acoustics* 38, 3 (Sept. 2010), 178–179.
- Ando Randrianoelina, Maarten Hornikx, Renz van Luxemburg, Stef van Eijndhoven, and Azar Dastouri. 2012. BEM for the Prediction of Standardized Scattering and Diffusion Coefficients of Diffusers. In *Acoustics 2012 Nantes*, Société Française d'Acoustique (Ed.), Nantes, France.
- Anton Ratnarajah and Dinesh Manocha. 2024. Listen2Scene: Interactive Material-Aware Binaural Sound Propagation for Reconstructed 3D Scenes. <https://doi.org/10.48550/arXiv.2302.02809> [eess]
- J. Redondo, J. V. Sánchez-Pérez, X. Blasco, J. M. Herrero, and M. Vorländer. 2016. Optimized Sound Diffusers Based on Sonic Crystals Using a Multiobjective Evolutionary Algorithm. *The Journal of the Acoustical Society of America* 139, 5 (May 2016), 2807–2814. <https://doi.org/10.1121/1.4948580>
- Atul Rungta, Carl Schissler, Nicholas Rewkowski, Ravish Mehra, and Dinesh Manocha. 2018. Diffraction Kernels for Interactive Sound Propagation in Dynamic Environments. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1613–1622. <https://doi.org/10.1109/TVCG.2018.2794098>

- Romana Rust. 2021. Computational Design and Evaluation of Acoustic Diffusion Panels for the Immersive Design Lab - An Acoustic Design Case Study. In *Stojakovic, V and Tepavcevic, B (Eds.), Towards a New, Configurable Architecture - Proceedings of the 39th eCAADe Conference - Volume 1, University of Novi Sad, Novi Sad, Serbia, 8-10 September 2021, Pp. 515-524*. CUMINCAD.
- Wallace C. Sabine. 1922. *Collected Papers on Acoustics*. Harvard University Press.
- Lauri Sarioja, T. Rinne, and T. Takala. 1994. Simulation of Room Acoustics with a 3-D Finite Difference Mesh. In *International Conference on Mathematics and Computing*.
- Harry A. Schenck. 1968. Improved Integral Formulation for Acoustic Radiation Problems. *The Journal of the Acoustical Society of America* 44, 1 (July 1968), 41–58. <https://doi.org/10.1121/1.1911085>
- M. R. Schroeder. 1975. Diffuse Sound Reflection by Maximum-length Sequences. *The Journal of the Acoustical Society of America* 57, 1 (Jan. 1975), 149–150. <https://doi.org/10.1121/1.380425>
- Christian Schüller, Daniele Panozzo, and Olga Sorkine-Hornung. 2014. Appearance-Mimicking Surfaces. *ACM Transactions on Graphics (TOG)* 33, 6 (Nov. 2014), 1–10. <https://doi.org/10.1145/2661229.2661267>
- Jonathan Richard Shewchuk. 2002. What Is a Good Linear Element? Interpolation, Conditioning, and Quality Measures. In *11th International Meshing Roundtable, [IMR] 2002*. Ithaca, United States.
- Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. [https://doi.org/10.48550/arXiv.1409.1556 \[cs\]](https://doi.org/10.48550/arXiv.1409.1556)
- Arnold Sommerfeld. 1912. Die Greensche Funktion der Schwingungsgleichung. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 21 (1912), 309–352. <http://eudml.org/doc/145344>
- Unchalee Siruragsar. 2016. Pink Red Yellow Petaled Flower in Close Up Shot. <https://www.pexels.com/photo/pink-red-yellow-petaled-flower-in-close-up-shot-85773/>.
- Stefan Stefancik. 2016. Brown Black and White Mountain Under White and Blue Cloudy Sky. <https://www.pexels.com/photo/brown-black-and-white-mountain-under-white-and-blue-cloudy-sky-109243/>.
- Lingge Tan, Jieun Yang, Jian Kang, and Hongpeng Xu. 2022. Sound-Scattering Properties of Sierpinski Triangle Fractal Structures in the near Field. *Applied Acoustics* 196 (July 2022), 108892. <https://doi.org/10.1016/j.apacoust.2022.108892>
- Zhenyu Tang, Hsien-Yu Meng, and Dinesh Manocha. 2021. Learning Acoustic Scattering Fields for Dynamic Interactive Sound Propagation. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, 835–844. [https://doi.org/10.1109/VR50410.2021.00111 \[cs\]](https://doi.org/10.1109/VR50410.2021.00111)
- Kenji Tojo, Ariel Shamir, Bernd Bickel, and Nobuyuki Umetani. 2024. Fabricable 3D Wire Art. In *ACM SIGGRAPH 2024 Conference Papers (SIGGRAPH '24)*. Association for Computing Machinery, New York, NY, USA, 1–11. <https://doi.org/10.1145/3641519.3657453>
- Xiaochun Tong, Hsueh-Ti Derek Liu, Yotam Gingold, and Alec Jacobson. 2023. Differentiable Heightfield Path Tracing with Accelerated Discontinuities. In *ACM SIGGRAPH 2023 Conference Proceedings (SIGGRAPH '23)*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3588432.3591530>
- Nicolas Tsingos, Carsten Dachsbarer, Sylvain Lefebvre, and Matteo Dellepiane. 2007. *Instant Sound Scattering*. The Eurographics Association.
- Vinh Läm. 2022. White Concrete Statue Near Body of Water. <https://www.pexels.com/photo/white-concrete-statue-near-body-of-water-12657483/>.
- Yael Vinker, Yuval Alaluf, Daniel Cohen-Or, and Ariel Shamir. 2023. CLIPascene: Scene Sketching with Different Types and Levels of Abstraction. [https://doi.org/10.48550/arXiv.2211.17256 \[cs\]](https://doi.org/10.48550/arXiv.2211.17256)
- Yael Vinker, Ehsan Pajouheshgar, Jessica Y. Bo, Roman Christian Bachmann, Amit Hain Bermano, Daniel Cohen-Or, Amir Zamir, and Ariel Shamir. 2022. CLIPasso: Semantically-Aware Object Sketching. *ACM Transactions on Graphics (TOG)* 41, 4 (July 2022), 86:1–86:11. <https://doi.org/10.1145/3528223.3530068>
- Jui-Hsien Wang, Ante Qu, Timothy R. Langlois, and Doug L. James. 2018. Toward Wave-Based Sound Synthesis for Computer Animation. *ACM Transactions on Graphics (TOG)* 37, 4 (July 2018), 109:1–109:16. <https://doi.org/10.1145/3197517.3201318>
- Craig J. Webb and Stefan Bilbao. 2011. Computing Room Acoustics with CUDA - 3D FDTD schemes with Boundary Losses and Viscosity. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 317–320. <https://doi.org/10.1109/ICASSP.2011.5946404>
- Tim Weyrich, Jia Deng, Connelly Barnes, Szymon Rusinkiewicz, and Adam Finkelstein. 2007. Digital Bas-Relief from 3D Scenes. *ACM Transactions on Graphics (TOG)* 26, 3 (July 2007), 32-es. <https://doi.org/10.1145/1276377.1276417>
- Kangrui Xue, Ryan M. Aronson, Jui-Hsien Wang, Timothy R. Langlois, and Doug L. James. 2023. Improved Water Sound Synthesis Using Coupled Bubbles. *ACM Transactions on Graphics (TOG)* 42, 4 (July 2023), 127:1–127:13. <https://doi.org/10.1145/3592424>
- Kangrui Xue, Jui-Hsien Wang, Timothy Langlois, and Doug James. 2024. WaveBlender: Practical Sound-Source Animation in Blended Domains. In *SIGGRAPH Asia 2024 Conference Papers (SA '24)*. Association for Computing Machinery, New York, NY, USA, 1–10. <https://doi.org/10.1145/3680528.3687696>
- Lihe Yang, Bingyi Kang, Zilong Huang, Zhen Zhao, Xiaogang Xu, Jiashi Feng, and Hengshuang Zhao. 2024. Depth Anything V2. [https://doi.org/10.48550/arXiv.2406.09414 \[cs\]](https://doi.org/10.48550/arXiv.2406.09414)
- Kane Yee. 1966. Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media. *IEEE Transactions on Antennas and Propagation* 14, 3 (May 1966), 302–307. <https://doi.org/10.1109/TAP.1966.1138693>
- Xiaojuen Yuan, D. Borup, J.W. Wiskin, M. Berggren, R. Eidens, and S.A. Johnson. 1997. Formulation and Validation of Berenger's PML Absorbing Boundary for the FDTD Simulation of Acoustic Scattering. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 44, 4 (July 1997), 816–822. <https://doi.org/10.1109/58.655197>
- Qiong Zeng, Ralph R. Martin, Lu Wang, Jonathan A. Quinn, Yuhong Sun, and Changhe Tu. 2014. Region-Based Bas-Relief Generation from a Single Image. *Graphical Models* 76, 3 (May 2014), 140–151. <https://doi.org/10.1016/j.gmod.2013.10.001>
- Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. 2018b. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. [https://doi.org/10.48550/arXiv.1801.03924 \[cs\]](https://doi.org/10.48550/arXiv.1801.03924)
- Yu-Wei Zhang, Yanzhai Chen, Hui Liu, Zhongping Ji, and Caiming Zhang. 2018a. Modeling Chinese Calligraphy Reliefs from One Image. *Computers & Graphics* 70 (Feb. 2018), 300–306. <https://doi.org/10.1016/j.cag.2017.07.022>
- Ziyi Zhang, Nicolas Roussel, and Wenzel Jakob. 2023. Projective Sampling for Differentiable Rendering of Geometry. *ACM Transactions on Graphics (TOG)* 42, 6 (Dec. 2023), 212:1–212:14. <https://doi.org/10.1145/3618385>

A BACKGROUND ON BOUNDARY ELEMENT ANALYSIS

We first consider the Helmholtz equation, which represents a time-independent form of the acoustic wave equation:

$$\nabla^2 u(\mathbf{x}) + k^2 u(\mathbf{x}) = 0 \quad (27)$$

where $u(\mathbf{x}) \in \mathbb{C}$ is the **complex** acoustic pressure at some point \mathbf{x} in space, k is the wavenumber given by $k = \frac{2\pi f}{v}$, f represents the frequency (Hz) of the wave, and v represents the wave's phase velocity (m/s).

The Helmholtz equation has an infinite number of solutions. Thus, an additional boundary condition known as the Sommerfeld radiation condition [Sommerfeld 1912] is often imposed, limiting solutions to waves that radiate outwards from sources, rather than allowing arbitrary inbound waves from the infinite domain. For the three-dimensional case, it can be written as

$$\lim_{|\mathbf{x}| \rightarrow \infty} |\mathbf{x}| \left(\frac{\partial}{\partial |\mathbf{x}|} - ik \right) u(\mathbf{x}) = 0, \quad (28)$$

where i is the imaginary unit and $|\cdot|$ is the Euclidean norm.

To solve Equation 27 we employ the Boundary Element Method (BEM). BEMs reduce the dimension of the problem by one, and their solutions innately satisfy the Sommerfeld radiation condition. More specifically, we use Green's second identity to transform the Helmholtz equation (Equation 27) into the Kirchoff-Helmholtz integral (Equation 29, below), which formulates the pressure at any point in terms of the values of the pressure and its derivative at all points on a given closed surface:

$$C(\mathbf{x})u(\mathbf{x}) = u^i(\mathbf{x}) + \int_{S_y} (u(y) \frac{\partial G(\mathbf{x}, y)}{\partial n_y} - G(\mathbf{x}, y) \frac{\partial u(y)}{\partial n_y}) dy \quad (29)$$

where S is the boundary of the domain, u^i is the incident pressure, n_y is the outward-pointing normal at y , and $G(\mathbf{x}, y)$ is the associated Green's function for the Helmholtz equation, which, in three dimensions, is given by

$$G(\mathbf{x}, y) = \frac{e^{-ikr}}{4\pi r} \quad (30)$$

with $r = |\mathbf{x} - \mathbf{y}|$. $C(\mathbf{x})$ is also given by

$$C(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in D \\ \frac{1}{2} & \mathbf{x} \in S \\ 1 & \mathbf{x} \in E \end{cases} \quad (31)$$

where D is the interior domain, S is the boundary, and E is the exterior (infinite) domain.

B DERIVATION OF ADJOINT GRADIENT

Starting from Equation 12, the chain rule gives the expression for the gradient as

$$\frac{dc}{dV} = \frac{dg}{d\mathbf{u}^e} \left(\frac{\partial Q}{\partial V} \mathbf{u} + QM^{-1} \left[\frac{d\mathbf{u}^i}{dV} - \frac{\partial M}{\partial V} \mathbf{u} \right] \right). \quad (32)$$

We introduce $\mathbf{v} = \frac{dg}{d\mathbf{u}^e}^T$ and can rewrite Equation 32 as

$$\frac{dc}{dV} = \mathbf{v}^T \frac{\partial Q}{\partial V} \mathbf{u} + \mathbf{v}^T QM^{-1} \left(\frac{d\mathbf{u}^i}{dV} - \frac{\partial M}{\partial V} \mathbf{u} \right). \quad (33)$$

The problematic term is $M^{-1} \left(\frac{d\mathbf{u}^i}{dV} - \frac{\partial M}{\partial V} \mathbf{u} \right)$. It either involves a matrix inversion, or $|V|$ additional matrix solves.

However, if we introduce the adjoint variable $\lambda^T = \mathbf{v}^T QM^{-1}$, we have

$$\begin{aligned} \lambda^T &= \mathbf{v}^T QM^{-1} \iff \lambda = M^{-T} Q^T \mathbf{v} \\ &\iff M^T \lambda = Q^T \mathbf{v} \end{aligned} \quad (34)$$

and hence we can compute λ using only one additional matrix solve.

By substituting λ^T back into Equation 33, we obtain the following expression for the gradient:

$$\frac{dc}{dV} = \mathbf{v}^T \frac{\partial Q}{\partial V} \mathbf{u} + \lambda^T \frac{d\mathbf{u}^i}{dV} - \lambda^T \frac{\partial M}{\partial V} \mathbf{u} \quad (35)$$

C ADDITIONAL DIFFUSER COMPARISONS

Our main frame of reference is the primitive root diffuser (PRD) shown in Figure 3. This represents one of the most common industry-standard diffusers on the market, with easy access to online generators⁵. We also test other diffuser geometries in Figure 33, including the quadratic residue diffuser (QRD) [Schroeder 1975], a noisy surface [Tsingos et al. 2007] generated using Perlin noise [Perlin 2002], and tiled cylinders and hemispheres [Cox and D'Antonio 2016]. The results clearly indicate the strong performance of the PRD, which we use as a baseline for our *Acoustic Reliefs*.

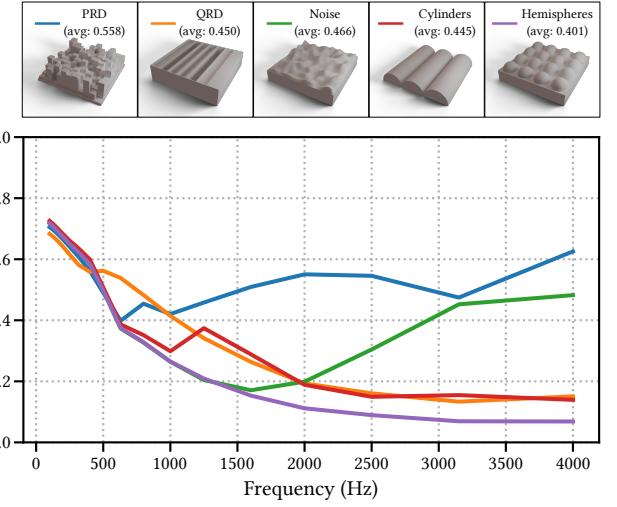


Fig. 33. **Other Diffusers.** We evaluate and plot the diffusion curves for various diffuser geometries suggested by literature [Cox and D'Antonio 2016; Schroeder 1975; Tsingos et al. 2007].

D FDTD SIMULATOR DETAILS

Our CUDA-based FDTD room acoustics solver closely follows Webb and Bilbao [2011]. In short, it uses the standard six-neighbor stencil, with support for frequency-independent boundary losses (but no air viscosity). To rasterize solid boundaries onto our FDTD grid, we use a conservative rasterizer based on the triangle-box overlap test from Akenine-Möller [2005]. For FDTD simulations in open spaces,

⁵<https://actools.tunetown.de/prd/>

we also implemented the split-field PML from Liu and Tao [1997] to absorb outgoing waves at the domain boundary, allowing us to truncate our simulation domains; in all of our experiments, we use a PML width of eight cells.

We validated our FDTD solver against BEM by comparing the diffusion coefficients (11) computed by both methods in Figure 34. Following the procedure outlined in Section 3, we place the source 1 m above the diffuser and evaluate the scattered pressure on a 1 m radius hemisphere centered on the diffuser. Note that unlike the setup used for optimizing our diffusers, the distances used here are considerably shorter to accommodate for small FDTD cell sizes.

To isolate the external pressure scattered by the diffuser from the total pressure field, we first run an FDTD simulation without any geometry. This procedure yields the direct path pressure, $p_{\text{direct}}(\mathbf{x}, t)$ (and is performed once for each FDTD resolution). We then incorporate the diffuser into the simulation to obtain $p_{\text{total}}(\mathbf{x}, t)$. Finally, the pressure scattered by the diffuser is computed as:

$$p_{\text{diffuse}}(\mathbf{x}, t) = p_{\text{total}}(\mathbf{x}, t) - p_{\text{direct}}(\mathbf{x}, t) \quad (36)$$

To compute the diffusion coefficient across multiple frequency bands, we use a delta function as our source and take an FFT to extract the scattered pressure per frequency band, interpolating between frequencies as needed.

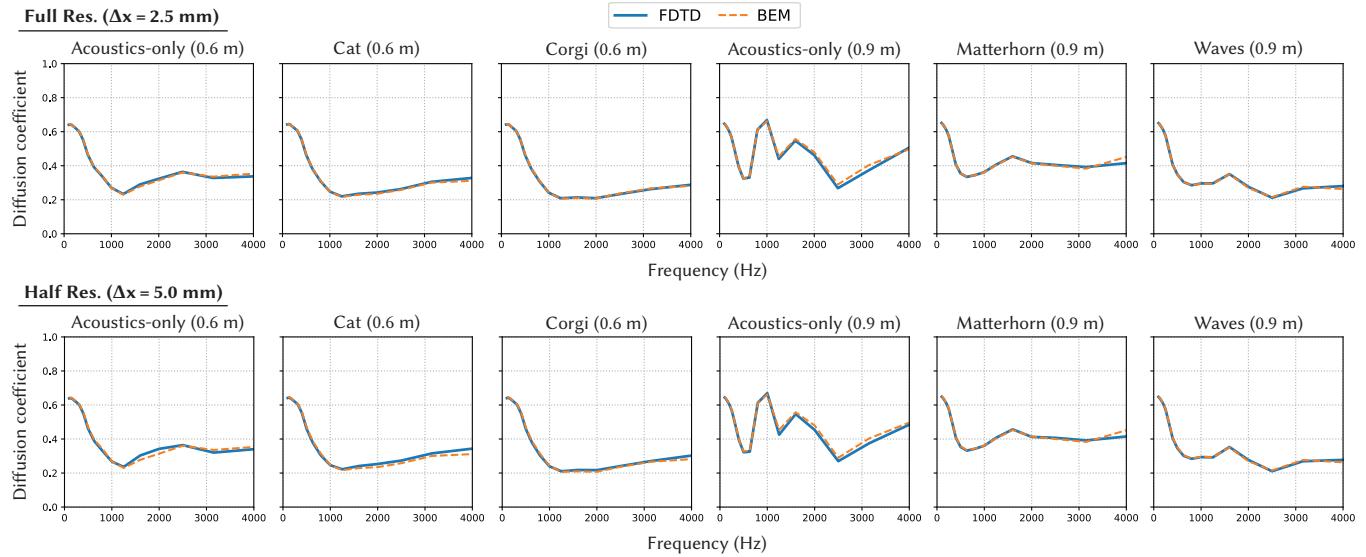


Fig. 34. FDTD vs. BEM comparison of diffusion coefficients shows strong agreement across (left-to-right) six different diffuser geometries and (top-to-bottom) two different FDTD resolutions. Here, Δx denotes the FDTD cell size, with $\Delta x = 2.5$ mm corresponding to a 240 kHz sampling rate and $\Delta x = 5$ mm corresponding to 120 kHz. Simulation times (for a hemispherical array of 0.025 s impulse responses) are 1.6 min. for each mesh at $\Delta x = 2.5$ and 19 s for each mesh at $\Delta x = 5.0$.