Wave-Ray Coupling for Interactive Sound Propagation in Large Complex Scenes

Hengchin Yeh*   Ravish Mehra†   Zhimin Ren*   Lakulish Antani*   Dinesh Manocha*
Ming Lin*
University of North Carolina at Chapel Hill

Figure 1: *Our hybrid technique is able to model high-fidelity acoustic effects for large, complex indoor or outdoor scenes at interactive rates: (a) building surrounded by walls, (b) underground parking garage, and (c) reservoir scene in Half-Life 2.*

Abstract

We present a novel hybrid approach that couples geometric and numerical acoustic techniques for interactive sound propagation in complex environments. Our formulation is based on a combination of spatial and frequency decomposition of the sound field. We use numerical wave-based techniques to precompute the pressure field in the near-object regions and geometric propagation techniques in the far-field regions to model sound propagation. We present a novel two-way pressure coupling technique at the interface of near-object and far-field regions. At runtime, the impulse response at the listener position is computed at interactive rates based on the stored pressure field and interpolation techniques. Our system is able to simulate high-fidelity acoustic effects such as diffraction, scattering, low-pass filtering behind obstruction, reverberation, and high-order reflections in large, complex indoor and outdoor environments and Half-Life 2 game engine. The pressure computation requires orders of magnitude lower memory than standard wave-based numerical techniques.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Physically based modeling; I.3.5 [Computer Graphics]: Applications—Sound rendering; I.6.8 [Simulation and Modeling]: Types of simulation—Animation;

Keywords: sound propagation, interactive applications, source simulation, scattering

Links: ◆DL 🅿PDF 🌐WEB

*{hyeh,zren,lakulish,dm,lin}@cs.unc.edu
†ravish.mehra07@gmail.com

1 Introduction

Sound propagation techniques are used to model how sound waves travel in the space and interact with various objects in the environment. Sound propagation algorithms are used in many interactive applications, such as computer games or virtual environments, and offline applications, such as noise prediction in urban scenes, architectural acoustics, virtual prototyping, etc.. Realistic sound propagation that can model different acoustic effects, including diffraction, interference, scattering, and late reverberation, can considerably improve a user's immersion in an interactive system and provides spatial localization [Blauert 1983].

The acoustic effects can be accurately simulated by numerically solving the acoustic wave equation. Some of the well-known solvers are based on the boundary-element method, the finite-element method, the finite-difference time-domain method, etc. However, the time and space complexity of these solvers increases linearly with the volume of the acoustic space and is a cubic (or higher) function of the source frequency. As a result, these techniques are limited to interactive sound propagation at low frequencies (e.g. 1-2KHz) [Raghuvanshi et al. 2010; Mehra et al. 2013], and may not scale to large environments.

Many interactive applications use geometric sound propagation techniques, which assume that sound waves travels like rays. This is a valid assumption when the sound wave travels in free space or when the size of intersecting objects is much larger than the wavelength. As a result, these geometric techniques are unable to simulate many acoustic effects at low frequencies, including diffraction, interference, and higher-order wave effects. Many hybrid combinations of numeric and geometric techniques have been proposed, but they are limited to small scenes or offline applications.

Main Results: In this paper, we present a novel hybrid approach that couples geometric and numerical acoustic techniques to perform interactive and accurate sound propagation in complex scenes. Our approach uses a combination of spatial decomposition and frequency decomposition, along with a novel two-way wave-ray coupling algorithm. The entire simulation domain is decomposed into different regions, and the sound field is computed separately by geometric and numerical techniques for each region. In the vicinity of objects whose sizes are comparable to the simulated wavelength (near-object regions), we use numerical wave-based methods to simulate all wave effects. In regions away from objects (far-field

regions), including the free space and regions containing objects that are much larger than the wavelength, we use a geometric ray-tracing algorithm to model sound propagation. We restrict the use of numeric propagation techniques to small regions of the environment and precompute the pressure field at low frequencies. The rest of the pressure field is precomputed using ray tracing.

At the interface between near-object and far-field regions, we need to couple the pressures computed by the two different (one numerical and one geometric) acoustic techniques. Rays entering a near-object region define the incident pressure field that serves as the input to the numerical acoustic solver. The numerical solver computes the outgoing scattered pressure field, which in turn has to be represented by rays exiting the near-object region. At the core of our hybrid method is a two-way coupling procedure that handles these cases. We present a scheme that represents two-way coupling using *transfer functions* and computes all orders of interaction.

The key results of this work include:

1. *An efficient hybrid approach* that decomposes the scene into regions that are more suitable for either geometric or numerical acoustic techniques, exploiting the strengths of both.

2. *Novel two-way coupling between wave-based and ray-based acoustic simulation* based on fundamental solutions at the interface that ensures the consistency and validity of the solution given by the two methods. Transfer functions are used to model two-way couplings to allow multiple orders of acoustic interactions.

3. *Fast, memory-efficient interactive audio rendering* that only uses tens to hundreds of megabytes of memory.

We have tested our technique on a variety of scenarios and integrated our system with the Valve's Source™ game engine. Our technique is able to handle both large indoor and outdoor scenes (similar to geometric techniques) as well as generate realistic acoustic effects (similar to numeric wave solvers), including late reverberation, high-order reflections, reverberation coloration, sound focusing, and diffraction low-pass filtering around obstructions. Furthermore, our pressure evaluation takes orders of magnitude less memory compared to state-of-the-art wave equation solvers.

2 Related Work

In this section, we present a brief overview of related work on sound propagation and reverberation.

2.1 Numerical Acoustic Techniques

Accurate, numerical acoustic simulations typically solve the acoustic wave equation using numerical methods, such as finite differences [Botteldooren 1994], finite elements [Thompson 2006], boundary elements [Gumerov and Duraiswami 2009], or adaptive rectangular decomposition [Raghuvanshi et al. 2009]. However, their time and space complexity increases as a third or fourth power of frequencies. Despite recent advances, they remain impractical for many large scenes.

The equivalent source [Ochmann 1995], expresses the solution fields of the wave equation in terms of a linear combination of points sources of various order (monopoles, dipoles, etc). James et al. [2006] solved a related *sound radiation* problem, using equivalent sources to represent the radiation field generated by a vibrating object.

2.2 Geometric Acoustic Techniques

Most acoustics simulation software and commercial systems are based on geometric techniques [Funkhouser et al. 1998; Vorlander 1989] that assume sound travels along linear rays [Funkhouser et al. 2004]. The simplified assumption of rays limits these methods to accurately capture specular and diffuse reflections only at high frequencies. Diffraction is typically modeled by identifying individual diffracting edges [Svensson et al. 1999; Tsingos et al. 2001]. These ray-based techniques can interactively model early reflections and first order edge-diffraction [Taylor et al. 2012]; however, they cannot interactively model the reverberation of the impulse response explicitly, since that would require high-order reflections and wave effects such as scattering, interference, and diffraction. While ray-tracing has been successfully used in many interactive acoustics systems [Lentz et al. 2007], the number of rays traced has to be limited for scenes with moving listeners in order to maintain real-time performance.

2.3 Hybrid Techniques

Several methods for combining geometric and numerical acoustic techniques have been proposed. One line of work is based on *frequency decomposition*: dividing the frequencies to be modeled into low and high frequencies. Low frequencies are modeled by numerical acoustic techniques, and high frequencies are treated by geometric methods, including the finite difference time domain method (FDTD) [Southern et al. 2011; Lokki et al. 2011], the digital waveguide mesh method (DWM) [Murphy et al. 2008], and the finite element method (FEM) [Granier et al. 1996; Aretz 2012]. However, these methods use numerical methods at lower frequencies over the entire domain. As a result, they are limited to offline applications and may not scale to very large scenes.

Another method of hybridization is based on *spatial decomposition*. The entire simulation domain is decomposed to different regions: near-object regions are handled by numerical acoustic techniques to simulate wave effects, while far-field regions are handled by geometric acoustic techniques. Hampel et al. [2008] combine the boundary element method (BEM) and geometric acoustics using a spatial decomposition. Their method provides a one-way coupling from BEM to ray tracing, converting pressures in the near-object region (computed by BEM) to rays that enter the far-field region containing the listener. In electromagnetic wave propagation, Wang et al. [2000] propose a hybrid technique combining ray tracing and FDTD. Their technique is also based on a one-way coupling, where rays are traced in the far-field region and collected at the boundaries of the near-object regions. The pressures are then evaluated and serve as the boundary condition for the FDTD method. These one-way coupling methods do not allow rays to enter and exit the near-object regions of an object, and therefore acoustic effects of that object will not be propagated to the far-field regions. Barbone et al. [1998] propose a two-way coupling that combines the acoustic field generated using ray-tracing and FEM. Jean et al. [2008] present a hybrid BEM/beam tracing approach to compute the radiation of tyre noise. However, these methods do not describe how multiple entrance of rays into near-object regions of different objects is handled, which is crucial when simulating interaction between multiple objects.

2.4 Acoustic Kernel-Based Interactive Techniques

There has been work in enabling interactive auralization for acoustic simulations through precomputation. At a high level, these techniques tend to precompute an acoustic kernel, which is used at run-time for interactive propagation in static environments. Raghuvan-

shi et al.[2010] precompute acoustic responses on a sampled spatial grid using a numerical solver. They then encode perceptually salient information to perform interactive sound rendering. Mehra et al. [2013] proposed an interactive sound propagation technique for large outdoor scenes based on equivalent sources. Other techniques use geometric methods to precompute high-order reflections or reverberation [Tsingos 2009; Antani et al. 2012] and compactly store the results for interactive sound propagation at runtime. Our method can be integrated into any of these systems as an acoustic kernel that can efficiently capture wave effects in a large scene.

3 Overview

In this section we give an overview of sound propagation and our proposed approach.

3.1 Sound Propagation

For a sound pressure wave with angular frequency ω, speed of sound c, the problem of sound propagation in domain Ω in the space can be expressed as a boundary value problem for the Helmholtz equation :

$$\nabla^2 p + \frac{\omega^2}{c^2} p = f; \quad \mathbf{x} \in \Omega, \tag{1}$$

where $p(\mathbf{x})$ is the complex valued pressure field, ∇^2 is the Laplacian operator, and $f(\mathbf{x})$ is the *source term*, (e.g. $= 0$ in free space and $\delta(\mathbf{x}')$ for a point source located at $\mathbf{x}'$). Boundary conditions are specified on the boundary $\partial \Omega$ of the domain (which can be the surface of a solid object, the interface between different media, or an arbitrarily defined surface) by a Dirichlet boundary condition that specifies pressure, $p(\mathbf{x}) = 0; \mathbf{x} \in \partial \Omega$, a Neumann boundary condition that specifies the velocity of medium, $\frac{\partial p(\mathbf{x})}{\partial n} = 0; \mathbf{x} \in \partial \Omega$, or a mixed boundary condition that specifies a complex-valued constant Z, so that $Z \frac{\partial p(\mathbf{x})}{\partial n} + p(\mathbf{x}) = 0; \mathbf{x} \in \partial \Omega$.

The pressure p at infinity must also be specified, usually by the *Sommerfeld radiation condition* [Pierce 1989], $\lim_{||\mathbf{x}|| \to \infty} \left[\frac{\partial p}{\partial ||\mathbf{x}||} + \hat{j} \omega c p \right] = 0$, where $||\mathbf{x}||$ is the distance of point $\mathbf{x}$ from the origin and $\hat{j} = \sqrt{-1}$.

Different acoustic techniques aim to solve the above equations with different formulations. Numerical acoustic techniques discretize Equation (1) and solve for p numerically with boundary conditions. Geometric acoustic techniques model p as a discrete set of rays emitted from sound sources which interact with the environment and propagate the pressure.

3.2 Acoustic Transfer Function

When modeling the acoustic effects due to objects or surfaces in a scene, it is often useful to define the *acoustic transfer function*. Many different acoustic transfer functions have been proposed to simulate different acoustic effects. In sound propagation problems, the acoustic transfer function maps an incoming sound field to an outgoing sound field. For example, Waterman developed a *transition-matrix method* for acoustic scattering [Waterman 2009] and maps the incoming and outgoing fields in terms of the coefficients of a complete system of vector basis functions. Antani et al. [2012] compute an acoustic radiance transfer operator that maps incident sound to diffusely reflected sound in a scene. Mehra et al. [2013] model the free-field acoustic behavior of an object, as well as pairwise interactions between objects. In sound radiation problems, James et al. [2006] map the vibration mode of an object to the radiated sound pressure field.

3.3 Hybrid Sound Propagation

We describe the various components of our hybrid sound propagation technique. Our approach uses a combination of frequency decomposition and spatial decomposition, as shown schematically in Figure 3. Since frequency decomposition is a standard technique [Granier et al. 1996], we mostly focus on spatial decomposition and our novel two-way coupling algorithm (see Figure 2).
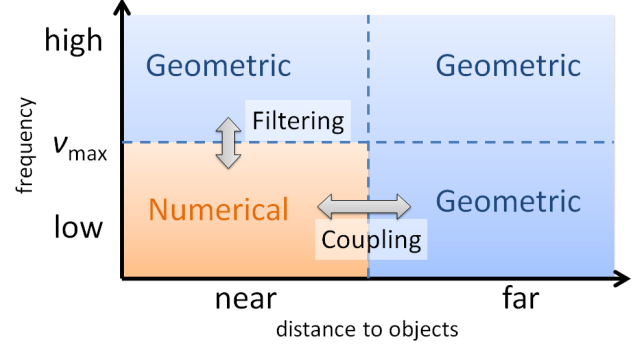

Figure 3: *Frequency and spatial decomposition. High frequencies are simulated using geometric techniques, while low frequencies are simulated using a combination of numerical and geometric techniques based on a spatial decomposition.*

Frequency Decomposition: We divide the modeled frequencies to low and high frequencies, with a crossover frequency ν_{max}. For high frequencies, geometric techniques are used throughout the entire domain. For low frequencies, a combination of numerical and geometric techniques is used based on a spatial decomposition described below. Typical values for ν_{max} are 0.5-2 kHz, and a simple low-pass–high-pass filter combination is usually used to join the results at the crossover frequency region.

Spatial decomposition: Given a scene we first classify it into *small objects* and *environment features*. The small objects, or simply *objects*, are of size comparable to or smaller than the wavelength of the sound pressure wave being simulated. The environment features represent objects much larger than the wavelength (like terrain). The wavelength that is used as the criterion for distinguishing small or large objects is a user-controlled parameter. One possible choice is the maximum audible wavelength (17 m), corresponding to the lowest audible frequency for human (20 Hz). When sound interacts with objects, wave phenomena are prominent only when the objects are small relative to the wavelength. Therefore we only need to compute accurate wave propagation in the local neighborhood of small objects. We call this neighborhood the *near-object region* (orange region in Figure 2) of an object, and numerical acoustic techniques are used to compute the sound pressure field in this region. The region of space away from small objects is called the *far-field region* and is handled by geometric acoustic techniques (blue region in Figure 2).

The spatial decomposition is performed as follows: For a small object A, we compute the offset surface ∂A^+ and define the near-object region, denoted as Ω^N, as the space inside the offset surface. The offset surface of an object is computed using discretized distance fields and the marching cubes algorithm similar to James et al. [2006]. If the offset surfaces of two objects intersect then they are treated as a single object and are enclosed in one Ω^N. The space complementary to the near-object region is defined as the far-field region, and is denoted as Ω^G.

Geometric acoustics: The pressure waves constituting the sound field in Ω^G are modeled as a discrete set of rays. Their propagation
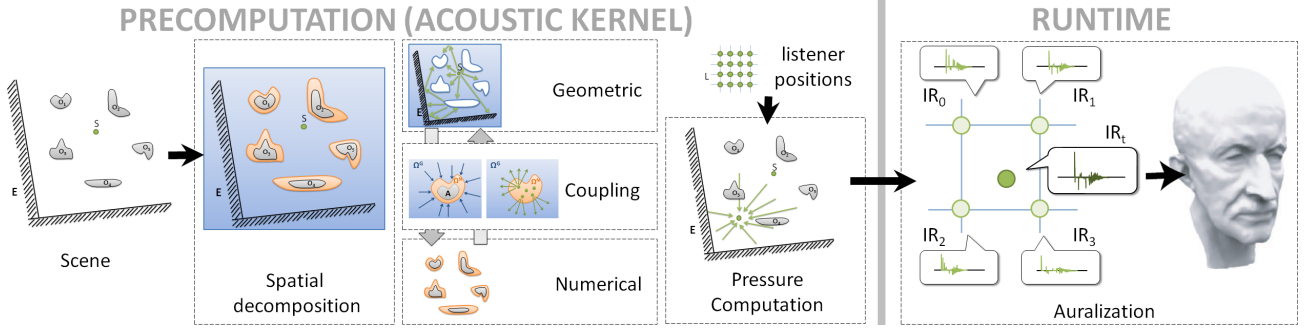
Figure 2: *Overview of spatial decomposition in our hybrid sound propagation technique: In the precomputation phase, a scene is classified into objects and environment features. This includes near-object regions (shown in orange) and far-field regions (shown in blue). The sound field in near-object regions is computed using a numerical wave simulation, while the sound field in far-field region is computed using geometric acoustic techniques. A two-way coupling procedure couples the results computed by geometric and numerical methods. The sound pressures are computed at different listener positions to generate the impulse responses. At runtime, the precomputed impulse responses (IR_0-IR_3) are retrieved and interpolated for the specific listener position (IR_t) at interactive rates, and final sound is rendered.*

in space and interaction with environment features (e.g. reflection from walls) are governed by geometric acoustic principles. We denote the pressure value defined collectively by the rays at position $\mathbf{x}$ as $p^G(\mathbf{x})$,

$$p^G(\mathbf{x}) = \sum_{r \in R} p_r(\mathbf{x}), \qquad (2)$$

where p_r is the contribution from one ray r in a set of rays R.

Numerical acoustic techniques: The sound pressure field scattered by objects in Ω^N is treated by wave-based numerical techniques for lower frequencies, in which the wave phenomena such as diffraction and interference are inherently modeled. We denote the pressure value at position $\mathbf{x}$ computed using numerical techniques as $p^N(\mathbf{x})$.

Coupling: At the interface between near-object and far-field regions, the pressures computed by the two different acoustic techniques need to be coupled (Figure 4). Rays entering a near-object region define the incident pressure field that serves as the input to the numerical solver. Similarly, the outgoing scattered pressure field computed by the numerical solver must be converted to a set of rays. The two-way coupling are modeled as transfer functions between incoming and outgoing rays. The process is detailed in Section 4.

Pressure computation: At each frequency lower than $\nu_{\max}$, the coupled geometric and numerical methods are used to solve the global sound pressure field. All frequencies higher than $\nu_{\max}$ are handled by geometric techniques throughout the entire domain.

Acoustic kernel: The previous stages serve as an *acoustic kernel*, which computes the impulse responses (IRs) for a given source-listener position pair. For each sound source, the pressure value at each listener position is evaluated for all simulated frequencies to give a complete acoustic *frequency response* (FR), which can in turn be converted to an impulse response (IR) through Fourier transform. IR's for predefined source-listener positions (usually on a grid) are precomputed and stored.

Auralization: At runtime, the IR for a general listener position is obtained by interpolating the neighboring precomputed IR's [Raghuvanshi et al. 2010], and the output sound is auralized by convoluting the input sound with the IRs in real time.

(a) $\Omega^G \rightarrow \Omega^N$ **(b)** p_{inc} **(c)** p_{sca} **(d)** $\Omega^N \rightarrow \Omega^G$
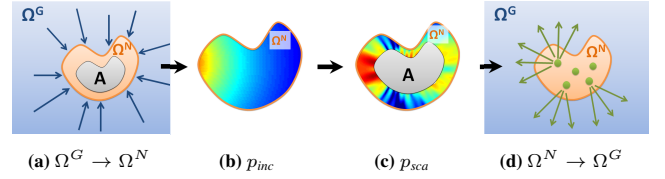
Figure 4: *Two-way coupling of pressure values computed by geometric and numerical acoustic techniques. (a) The rays are collected at the boundary and the pressure evaluated. (b) The pressure on the boundary defines the incident pressure field p_{inc} in Ω^N, which serves as the input to the numerical solver. (c) The numerical solver computes the scattered field p_{sca}, which is the effect of object A to the pressure field. (d) p_{sca} is expressed as fundamental solutions and represented as rays emitted to Ω^G.*

4 Two-Way Wave-Ray Coupling

In this section, we present the details of our two-way coupling procedure. We also highlight the precomputation and runtime phases. The coupling procedure ensures the consistency between p^G and p^N, the pressures computed by the geometric and numerical acoustic techniques, respectively. Any exchange of information at the interface between Ω^G and Ω^N must result in valid solutions to the Helmholtz equation (1) in both domains Ω^G and Ω^N.

4.1 Geometric → Numerical

From the pressure field p^G, we want to find the *incident pressure field* p_{inc}, which serves as the input to the numerical solver inside Ω^N. The incident pressure field is defined as the pressure field that corresponds to the solution of the wave equation if there were no objects in Ω^N.

Mathematically p_{inc} is the solution of the free-space Helmholtz Equation (1) with forcing term $f = 0$. Since there is no object in domain Ω^G,

$$p_{inc}(\mathbf{x}) = p^G(\mathbf{x}); \quad \mathbf{x} \in \Omega^G. \qquad (3)$$

This equation defines a Dirichlet boundary condition on the interface ∂A^+:

$$p = p^G(\mathbf{x}); \quad \mathbf{x} \in \partial A^+, \qquad (4)$$

The uniqueness of the acoustic boundary value problem guarantees that the solution of the free-space Helmholtz Equation, along with the specified boundary condition, is unique inside Ω^N. The unique solution $p_{\text{inc}}(\mathbf{x})$ can be found by expressing it as a linear combination of *fundamental solutions*. [1] If $\varphi_i(\mathbf{x})$ is a fundamental solution, and $p_{\text{inc}}(\mathbf{x})$ is expressed as a linear combination,

$$p_{\text{inc}}(\mathbf{x}) = \sum_i c_i \varphi_i(\mathbf{x}) \quad \mathbf{x} \in \Omega^N, \tag{5}$$

then the linearity of the wave equation implies that $p_{\text{inc}}(\mathbf{x})$ is also a solution. Furthermore, if the coefficients c_i are such that the boundary condition (4) is satisfied, then $p_{\text{inc}}(\mathbf{x})$ is the required *unique* solution to the boundary value problem (Section 3 in Ochmann [1995]). Therefore, the resultant pressure field is a valid incoming field in the numerical domain. The numerical solver takes the incident pressure field, considers the effect of the object inside Ω^N, and computes the outgoing scattered field. Figures 4(a) and 4(b) illustrate the process.

4.2 Numerical → Geometric

In order to transfer information from Ω^N to Ω^G, a discrete set of rays must be determined to represent the computed pressure p^N. These outgoing rays may be emitted from some starting points located in Ω^N and carry different information related to the modeled pressure waves (strength, phase, frequency, spatial derivatives of pressure, etc.) The coupling procedure thus needs to compute the appropriate outgoing rays, given the numerically computed p^N.

The scattered field in the numerical domain due to the object can be simply written as,

$$p_{\text{sca}}(\mathbf{x}) = p^N(\mathbf{x}); \quad \mathbf{x} \in \Omega^N. \tag{6}$$

We need to find the scattered field outside of Ω^N, and model it as a set of rays. As before, Equation (6) defines a Dirichlet boundary condition on the interface ∂A^+,

$$p = p^N(\mathbf{x}); \quad \mathbf{x} \in \partial A^+. \tag{7}$$

The free space Helmholtz Equation, along with this boundary condition, uniquely defines the scattered field p_{sca} outside Ω^N. We again express p_{sca} as a linear combination of fundamental solutions φ_j:

$$p_{\text{sca}}(\mathbf{x}) = \sum_j c_j \varphi_j(\mathbf{x}); \quad \mathbf{x} \in \Omega^G, \tag{8}$$

and then find the coefficients c_j by satisfying the boundary condition (7). This gives us a unique solution for scattered field $p_{\text{sca}}(\mathbf{x})$ outside Ω^N. We then use a set of rays R_j^{out} to model the fundamental solutions $\varphi_j(\mathbf{x})$ such that

$$\varphi_j(\mathbf{x}) = \sum_{r \in R_j^{\text{out}}} p_r(\mathbf{x}), \quad \mathbf{x} \in \Omega^G. \tag{9}$$

These rays correctly represent the outgoing scattered field in Ω^G. Figure 4(c) and 4(d) illustrate the process.

The coupling process described above is a general formulation and is independent of the underlying numerical solver (BEM, FEM, etc.) that is used to compute p^N as long as the pressure on the interface ∂A^+ can be evaluated and expressed as a set of fundamental

---

[1]A fundamental solution F for a linear operator L (in this case the Helmholtz operator $L = \nabla^2 + \frac{\omega^2}{c^2}$) is defined as the solution to the equation $LF = \delta(\mathbf{x})$, where δ is the Dirac delta function [Vladimirov 1976].

---

solutions. Depending on the mathematical formulation of the selected set of fundamental solutions $\varphi_j(\mathbf{x})$, different rays (starting points, directions, information carried, etc.) can be defined. However, a general principle is that if $\varphi_j(\mathbf{x})$ has a singularity at $\mathbf{y}_j$, then $\mathbf{y}_j$ is a natural starting point from which rays are emitted. The directions of rays sample a unit sphere uniformly or with some distribution function (e.g. guided sampling [Taylor et al. 2012]). The choice of fundamental solutions will be discussed in the next section.

Note that if the fundamental solutions φ_i and φ_j used to express the incident field (Equation (5)) and outgoing scattered field (Equation (8)) are *predetermined*, then the mapping from φ_i to φ_j can be precomputed. This precomputation process will be discussed in section 4.4.

4.3 Fundamental solutions

The requirement for the choice of fundamental solution φ_j is that it must satisfy the Helmholtz Equation (1) and the Sommerfeld radiation condition.

Equivalent Sources: One choice of fundamental solutions is based on *equivalent sources* [Ochmann 1995]. Each fundamental solution is chosen to correspond to the field due to *multipole sources* of order L ($L = 1$ is a monopole, $L = 2$ is a dipole, etc.) located at $\mathbf{y}_j$:

$$\varphi_j(\mathbf{x}) = \varphi_{jlm}(\mathbf{x}), \tag{10}$$

for $l \leq L - 1$ and $-l \leq m \leq l$, and

$$\varphi_{jlm} = \Gamma_{lm} h_l^{(2)}(\omega \rho_j / c) \psi_{lm}(\theta_j, \phi_j), \tag{11}$$

where $(\rho_j, \theta_j, \phi_j)$ corresponds to the vector $(\mathbf{x} - \mathbf{y}_j)$ expressed in spherical coordinates, $h_l^{(2)}(\cdot)$ is the complex-valued spherical Hankel function of the second kind, $\psi_{lm}(\theta_j, \phi_j)$ is the complex-valued spherical harmonic function, and Γ_{lm} is the real-valued normalizing factor that makes the spherical harmonics orthonormal [Arfken et al. 1985]. We use a shorthand generalized index h for (l, m), such that $\varphi_{jh}(\mathbf{x}) \equiv \varphi_{jlm}(\mathbf{x})$.

For pressure fields outside of ∂A^+ (i.e. in Ω^G), these equivalent sources are placed inside of ∂A^+ (i.e. in Ω^N). In a similar fashion, for pressure fields inside Ω^N, the equivalent sources must be placed outside Ω^N.

We model the outgoing pressure field from these equivalent sources using rays (Equation (9)) as follows. Rays are emitted from the source location $\mathbf{y}_j$. For a ray of direction (θ, ϕ) that has traveled a distance ρ, its pressure is scaled by $\psi_{lm}(\theta, \phi)$ and $h_l^{(2)}(\omega \rho / c)$.

Note that we can use equivalent sources to express a pressure field independently of how the pressure field was computed. For a computed p^N, we only need to find the locations $\mathbf{y}_j$ and coefficients c_j of the equivalent sources. This is performed by satisfying the boundary condition (8) in a least squared sense.

Boundary Elements: If the underlying numerical acoustic technique of choice is the boundary element method (BEM), then another set of fundamental solutions which is directly based on the BEM formulation is possible. For a domain with boundary $\partial \Omega$, the boundary element method solves the boundary integral equation of the Helmholtz equation. The boundary $\partial \Omega$ is discretized into triangular surface elements, and the equation is solved numerically for two variables; the pressure p and its normal derivative $\frac{\partial p}{\partial n}$ on the boundary. Once the boundary solutions p and $\frac{\partial p}{\partial n}$ are known, the sound pressure in the domain can be found for any point $\mathbf{x}$ by summing

all the contributions from the surface triangles:

$$p(\mathbf{x}) = \int_{\partial\Omega} \left(G(\mathbf{y}, \mathbf{x}) \frac{\partial p(\mathbf{y})}{\partial n} - \frac{\partial G(\mathbf{y}, \mathbf{x})}{\partial n} p(\mathbf{y}) \right) d\left(\partial\Omega(\mathbf{y})\right),$$
(12)

where $\mathbf{y}$ is the approximated position of the triangle and G is the Green's Function $G(\mathbf{y}, \mathbf{x}) = \exp(\hat{j}\omega|\mathbf{x} - \mathbf{y}|/c)/4\pi|\mathbf{x} - \mathbf{y}|$ [Gumerov and Duraiswami 2009].

Note that the discretization of Equation (12) also takes the form of Equation (8) as a linear combination of fundamental solutions:

$$p(\mathbf{x}) = \sum_j \left(c_j^1 \varphi_j^1(\mathbf{x}) + c_j^2 \varphi_j^2(\mathbf{x}) \right),$$
(13)

where the two kinds of fundamental solutions are

$$\varphi_j^1(\mathbf{x}) = G(\mathbf{y_j}, \mathbf{x}) \frac{\partial p(\mathbf{y_j})}{\partial n}; \; \varphi_j^2(\mathbf{x}) = -\frac{\partial G(\mathbf{y_j}, \mathbf{x})}{\partial n} p(\mathbf{y_j}).$$
(14)

Under this formulation, we can represent the pressure field as two kinds of rays emitted from each triangle location $\mathbf{y_j}$, each modeling $\varphi_j^1(\mathbf{x})$ and $\varphi_j^2(\mathbf{x})$ respectively. Then for a point in Ω^G the pressure field defined by the rays is computed according to Equation (12).

4.4 Precomputed Transfer Functions

If we consider what happens in Ω^N as a black box, the net result of the coupling and the numerical solver is that a set of rays enter Ω^N and then another set of rays exit Ω^N:

$$R^{\text{in}} \xrightarrow{\mathcal{M}} R^{\text{out}},$$
(15)

where R^{in} is the set of incoming rays entering Ω^N, R^{out} is the set of outgoing rays, and $\mathcal{M}$ is the ray transfer function. In this case, the function $\mathcal{M}$ is similar to the bidirectional reflectance distribution function (BRDF) for light [Ben-Artzi et al. 2008]. In our formulation, $\mathcal{M}$ encodes all the operations for the following computations:

1. Collect pressures defined by R^{in} to form the incident field on the interface (Equation (4));
2. Express the incident field as a set of fundamental solutions (Equation (5));
3. Compute the outgoing scattered field using the numerical acoustic technique;
4. Express the outgoing scattered field as a set of fundamental solutions (Equation (8)); and finally,
5. Find a set of rays R^{out} that model these functions (Equation (9).

A straightforward realization of hybrid sound propagation technique is possible: from each sound source rays are traced, interacting with the environment features, entering and exiting the near-object regions transfered by different $\mathcal{M}$'s, and finally reaching a listener. However, as the first step of $\mathcal{M}$ depends on the incoming rays R^{in}, a different $\mathcal{M}$ must be computed each time the rays enter the same near-object region. Moreover, the process must be repeated until the solution converges to a steady state, which may be too time-consuming for a scene (e.g. an indoor scene) with multiple ray reflections causing multiple entrances to near-object regions.

While previous two-way hybrid techniques do not consider this problem [Barbone et al. 1998; Jean et al. 2008], we address this problem by observing that if the fundamental solutions in Step 2 (denoted as φ_i^{in}) and Step 4 (denoted as φ_j^{out}) are predefined, then we can precompute the results of Step 2-Step 5 for an object. Similar to the BRDF for light, one can define the BRDF for sound. The mapping of φ_i^{in} to φ_j^{out} for an object is called the *per-object transfer*

function. For different R^{in} that define an incident field p_{inc} on the interface, we only need to compute the expansion coefficients d_i of the fundamental solutions φ_i^{in}; the outgoing rays are computed by applying the precomputed per-object transfer function.

The outgoing scattered field that is modeled as outgoing rays from an object A may, after propagating in space and interacting with the environment, enter as incoming rays into the near-object region of another object B. For a scene where the environment and relative positions of various objects are fixed, we can precompute all the propagation paths for rays that correspond to A's outgoing basis functions $\varphi_{j,A}^{\text{out}}$ and that reach B's near-object region. These rays determine the incident pressure field arriving at object B, which can again be expressed as a linear combination of a set of basis functions $\varphi_{i,B}^{\text{in}}$. The mapping from $\varphi_{j,A}^{\text{out}}$ to $\varphi_{i,B}^{\text{in}}$, called the *inter-object transfer function*, which is a fixed function and can also be precomputed. Interactions between multiple objects can therefore be found by a series of applications of the inter-object transfer functions.

Based on the per-object and inter-object transfer functions, all orders of acoustic interaction (corresponding to multiple entrance of rays to near-object regions) in the scene can be found for the total sound field by solving a global linear system, which is much faster than the straightforward hybridization, where the underlying numerical solver is invoked multiple times for each order of interactions. The trade-off is that the transfer functions have to be precomputed. However, the pre-object transfer functions can be reused even when the objects are moved. This characteristic is beneficial for quick iterations when authoring scenes, and can potentially be a cornerstone for developing sound propagation systems that supports fully dynamic scenes.

5 Implementation

In this section we discuss the implementation aspect for our technique.

5.1 Implementation details

The geometric acoustics code is written in C++, based on the Impulsonic Acoustect SDK[2], which implements a ray-tracing based image source method. For the numerical acoustic technique we use a GPU-based implementation of the ARD wave-solver [Raghuvanshi et al. 2009]. Per-object transfer functions, inter-object transfer functions, and equivalent source strengths are computed using a MATLAB implementation based on [Mehra et al. 2013].

Table 1 provides the detailed timing results for the precomputation stage. The timings are divided into two groups. The first group, labeled as "Hybrid Pressure Solving," consists of all the steps required to compute the final equivalent source strengths. These computations are performed once for a given scene. The second group, labeled as "Pressure Evaluation," involves the computation of the pressures contributed by all equivalent sources at a listener position. This computation is performed once for each sampled listener position.

The timing results for "wave sim." (simulation time of the ARD wave solver), and "Pressure Evaluation" are measured on a single core of a 4-core 2.80 GHz Xeon X5560 desktop with 4GB of RAM and NVIDIA GeForce GTX 480 GPU with 1.5 GB of RAM. All the other results are measured on a cluster containing a total of 436 cores, with sixteen 16-CPUs (8 dual-core 2.8GHz Opterons, 32GB RAM each) and forty-five 4-CPU (2 dual-core 2.6GHz Opterons, 8GB RAM each).

				Hybrid Pressure Solving								Pressure Evaluation
				Numerical				Geometric				
Scene	#src	#freq.	#eq. srcs	wave sim.	per-object	inter-object	source + global field	#tris	order	#rays	propagation time	
Building+small	1	300	220 K	163 min	552 min	22 min	19 min	60	3	4096	41 min	81 sec
Building+medium	1	400	290 K	217 min	736 min	33 min	23 min	60	3	4096	39 min	81 sec
Building+large	1	800	580 K	435 min	1472 min	54 min	40 min	60	3	4096	39 min	81 sec
Reservoir	1	500	500 K	254 min	252 min	4 min	2.6 min	16505	2	262144	1.9 min	10 sec
Parking	2	500	123 K	55 min	40 min	3 min	0.9 min	5786	2	4096	6.6 min	24 sec

Table 1: Precomputation Performance Statistics. *The rows "Building+small", "Building+medium", and "Building+large" correspond to scenes with a building surrounded by small, medium, and large walls, respectively. "Reservoir" and "Parking" denote the reservoir and underground parking garage scene respectively. For a scene, "#src" denotes the number of sound sources in the scene, "#freq." is the number of frequency samples, and "#eq. srcs" denotes the number of equivalent sources. The first part, "Hybrid Pressure Solving". includes all the steps required to compute the final equivalent source strengths, and is performed once for a given sound source and scene geometries. The second part, "Pressure Evaluation", corresponds to the cost of evaluating the contributions from all equivalent sources at a listener position and is performed once for each listener position. For the numerical technique, "wave sim." refers the total simulation time of the numerical wave solver for all frequencies; "per-object" denotes the computation time of for per-object transfer functions; "inter-object" is the inter-object transfer functions for each pair of objects (including self-inter-object transfer functions, where the pressure wave leaves a near-object region and reflected back to the same object); "source + global solve" is the time to solve the linear system to determine the strengths of incoming and outgoing equivalent sources. For the geometric technique, "# tris" is the number of triangles in the scene; "order" denotes the order of reflections modeled; "# rays" is the number of rays emitted from a source (sound source or equivalent source). The column "propagation time" includes the time of finding valid propagation paths and computing pressures for any intermediate step (e.g. from one object to another object's offset surface).*

We assume the scene is given as a collection of objects and terrains. In the spatial decomposition step, the offset surface is computed using distance fields. One important parameter is the spatial Nyquist distance h, corresponding to the highest frequency simulated ν_{max}, $h = c/2\nu_{max}$, where c is the speed of sound. To ensure enough spatial sampling on the offset surface, we choose the voxel resolution of distance field to be h, and the sample points are the vertices of the surface given by the marching cubes algorithm. The offset distance is chosen to be $8h$. In general, a larger offset distance means a larger spatial domain for the numerical solver and is therefore more expensive. On the other hand, a larger offset distance results in a pressure field with less detail (i.e. reduced spatial variation) on the offset surface, and fewer outgoing equivalent sources are required to achieve the same error threshold.

5.2 Collocated equivalent sources

The positions of outgoing equivalent sources can be generated by a greedy algorithm that selects the best candidate positions randomly [James et al. 2006]. However, if each frequency is considered independently, a total of $1M$ or outgoing equivalent sources may arise across all simulated frequencies. Because we must trace N_r rays, (typically thousands or more) from each equivalent source, this computation becomes a major bottleneck in our hybrid framework. This may cause a computation bottleneck in our hybrid framework, because we need to trace N_r rays (typically thousands or more) from each equivalent source.

We resolve this issue by reusing equivalent sources positions across different frequencies as much as possible. First, the equivalent sources for the highest frequency ν_{max}, which requires the highest number of equivalent sources, P_{max}, are computed using the greedy algorithm. For lower frequencies, the candidate positions are drawn from the P_{max} existing positions, which guarantees that a total of P_{max} collocated positions is occupied. Indeed, when the path is frequency-independent, rays emitted from collocated sources will travel the same path, which reduces the overall ray-tracing cost. The frequency-independent path assumption holds for paths containing only specular reflections, in which case the incident and re-

flected directions are determined. We observe a $60-100X$ speedup while maintaining the same error bounds over methods without the collocation scheme. All the timings results in this section are based on this optimization.

5.3 Auralization

We compute the frequency responses using our spatial decomposition approach up to $\nu_{max} = 1$ kHz with a sampling step size of 2.04 Hz. For frequencies higher than ν_{max}, we use a ray tracing solution, with diffractions approximated by the Uniform Theory of Diffraction (UTD) model [Kouyoumjian and Pathak 1974]. We join the low- and high-frequency responses in the region $[800, 1000]$ Hz using a low-pass–high-pass filter combination.

The sound sources in our system are recorded audio clips. The auralization is performed using overlap-add STFT convolutions. A "dry" input audio clip is first segmented into overlapping frames, and a windowed (Blackman window) Short-Time Fourier transform (STFT) is performed. The transformed frames are multiplied by the frequency responses corresponding to the current listener position. The resulting frequency-domain frames are then transformed back to time-domain frames using inverse FFT, and the final audio is obtained by overlap-adding the frames. For spatialization we use a simplified spherical head model with one listener position for each ear. Richer spatialization can be modeled using *head related transfer functions* (HRTFs), which are easily integrated in our approach.

For the interactive auralization we implemented a simplified version of the system proposed by Raghuvanshi et al. [2010]. Only the listener positions are sampled on a grid (of 0.5m-1m grid size), and the sound sources are kept static. The case of moving sound sources and a static listener is handled using the principle of acoustic reciprocity [Pierce 1989]. The interactive auralization is demonstrated through integration with Valve's Source™ game engine. Audio processing is performed using FMOD at a sampling rate of 44.1 kHz; the audio buffer length is 4096 samples, and the FFTs are computed using the Intel MKL library. The runtime performance statistics are summarized in Table 2. The parking garage scene is rendered offline and not included in this table.

Scene	#IR samples	Memory	Time
Building+small	960	19 MB	3.5 ms
Building+med	1600	32 MB	3.5 ms
Building+large	6400	128 MB	3.5 ms
Reservoir	17600	352 MB	1.8 ms

Table 2: Runtime Performance on a Single Core. *For each scene, "#IR samples" denotes the number of IR's sampled in the scene to support moving listeners or sources; "Memory" shows the memory to store the IR's; "Time" is the total running time needed to process and render each audio buffer.*

6 Results and Analysis

In this section we present the results of our hybrid technique in different scenarios and error analysis.

6.1 Scenarios

We demonstrate the effectiveness of our technique in a variety of scenes. The scenes are at least as complex as those shown in previous wave-based sound simulation techniques [James et al. 2006; Raghuvanshi et al. 2009; Mehra et al. 2013] or geometric methods with precomputed high-order reverberation [Tsingos 2009; Antani et al. 2012]. Please refer to the supplementary video for the auralizations. In each scene, we compare the audio generated by our method with existing sound propagation methods: a pure geometric technique is used for comparison [Taylor et al. 2012], which models specular reflection as well as edge diffraction through UTD; a pure numerical technique, the ARD wave-solver [Raghuvanshi et al. 2009]. Comparisons with ARD are done only in a limited selection of scenes (Building), while the other scenes (Underground Parking Garage and Reservoir) are too large to fit in the memory using ARD.

Building. As the listener walks behind the building, we observe the low-pass occlusion effect with smooth transition as a result of diffraction. We also observe the reflection effects due to the surrounding walls. We show how sound changes as the distance from the listener to the walls and the height of the walls vary.

Underground Parking Garage. This is a large indoor scene with two sound sources, a human and a car, as well as vehicles that scatter and diffract sound. As the listener walks through the scene, we observe the characteristic reverberation of a parking garage, as well as the variation of sound received from various sources depending on whether the listener is in the line-of-sight of the sources.

Reservoir. We demonstrate our system in a large outdoor scene from the game Half-Life 2, with a helicopter as the sound source. This scene shows diffraction and scattering due to a rock; it also shows high-order interactions between the scattered pressure and the surrounding terrain, which is most pronounced when the user walks through a passage between the rock and the terrain. Interactive auralization is achieved by precomputing the IRs at a grid of predefined listener positions. We also make the helicopter fly and thereby demonstrate the ability to handle moving sound sources and high-order diffractions.

6.2 Error Analysis

In Figure 5 we compare the results of our hybrid technique with BEM on a spatial grid of listener locations at different frequencies for several scenes: two parallel walls, two walls with a ground, an empty room, and two walls in a room. BEM is one of the most accurate wave-based simulators available, and comparing with high-accuracy simulated data is a widely adopted practice [Barbone et al. 1998; Jean et al. 2008; Hampel et al. 2008]. BEM results are generated by the FastBEM simulator[3]. A comparison with a geometric technique for the last scene is also provided. The geometric technique models 8 orders of reflection and 2 orders of diffraction through UTD.

We also compute the difference in pressure field (i.e. the error) between our hybrid technique with varying reflection orders and BEM, as shown in Figure 6 for the "Two Walls in a Room" scene. The error between the pressure fields generated by the reference wave solver and by our hybrid method , is computed as $||P_{\text{ref}} - P_{\text{hybrid}}||^2/||P_{\text{ref}}||$, where P_{ref} and P_{hybrid} are vectors consisting of complex pressure values at all the listener positions and $||\cdot||$ denotes the two-norm of complex values, summed over all positions $\mathbf{x}$ (the grid of listeners as shown in Figure 5). Higher reflection orders lead to more accurate results but require more rays to be traced.

6.3 Complexity

Consider a scene with κ objects. We perform the complexity analysis for frequency ν and discuss the cost of numerical and geometric techniques used.

Numerical Simulation and Pre-Processing: The pre-processing involves several steps: (1) performing the wave simulation using numerical techniques, (2) computing per-object and inter-object transform matrix, and (3) solving linear systems to determine the strengths of incoming and outgoing equivalent sources [Mehra et al. 2013]. In our system, the equivalent sources are limited to monopoles and dipoles, and the complexity follows:

$$O(\kappa n Q P^2 + \kappa^2 n P Q^2 + \kappa(u \log u) + \kappa^3 P^3), \qquad (16)$$

where Q, P are the number of incoming and outgoing equivalent sources respectively, n is the number of offset surface samples, and u is the volume of an object. The number of equivalent sources P and Q scale quadratically with frequency.

Ray Tracing: Assume the scene has T triangles, and from each source we trace N_r rays to the scene. The cost for one bounce of tracing from a source is $O(N_r \log T)$ on average and $O(N_r T)$ in the worst case. If the order of reflections modeled is d, then the (worst case) cost of ray-tracing is $O(N_r T^d)$. This cost is multiplied by the number of sources (sound sources and equivalent sources) and the number of points where the pressure values need to be evaluated. The total cost is dominated by computing inter-object transfer functions, where the pressure from P outgoing equivalent sources from an object needs to be evaluated at n sample positions on the offset surface of another object. This results in

$$O(\kappa^2 P n T^d) \qquad (17)$$

for a total of κ^2 pairs of objects in the scene.

In our collocated equivalent source scheme, however, the P outgoing sources for different frequencies share a total of P_{col} positions. The rays traced from a shared position can be reused, so for all frequencies ν, we only need to trace rays from P_{col} positions instead of $\sum_\nu P(\nu)$ positions .

The choice of N_r is scene-dependent. In theory, in order to discover all possible reflections from all scene triangles without missing a propagation path, the ray density along every direction should be high enough so that the triangle spanning the smallest solid angle

---

[3]http://www.fastbem.com/

viewed from the source can be hit by at least one ray. The problem of missing propagation paths is intrinsic to all ray-tracing methods. It can be overcome by using beam-tracing methods [Funkhouser et al. 1998], but they are considerably more expensive and are only practical for simple scenes.

The order of reflection d also depends on the scene configuration. For an outdoor scene where most reflections come from the ground, a few reflections are sufficient. In enclosed or semi-enclosed spaces more reflections are needed. In practice it is common to stop tracing rays when a given bound of reflection is reached, or when the reflected energy is less than a threshold.

Scalability Although the computation domain of the numerical solver, Ω^N, is smaller than the entire scene, the size of the entire scene still matters. Larger scenes require longer IR responses and therefore more frequency samples, which affect the cost of both numerical and geometric components of our hybrid approach. Larger scenes in general require more triangles, assuming the terrain has the same *feature density*. For a scene whose longest dimension is L, the number of IR samples (and therefore frequency samples) scales as $O(L)$, and the number of triangles scales as $O(L^2)$, - giving overall numerical and ray-tracing complexities of - $O(L)$ and $O(L^3 \log L)$ respectively. This is better than most numerical methods; for example, the time complexity of ARD are $O(L^4 \log L)$ and FDTD scale $O(L^4)$.

We tested the scalability of our method with the building scene by increasing the size of the scene and measuring the performance. The results are shown in Figure 7. Since the open space is handled by geometric methods, whose complexity of the geometric method is not a direct function of the total volume, it is not necessary to divide the open space into several connected smaller domains, as some previous methods did [Raghuvanshi et al. 2009].

## 6.4 Comparison with Prior Techniques

Compared with geometric techniques, our approach is able to capture wave effects such as scattering and high-order diffraction, thereby generate sound of higher quality. When compared with performing numerical wave-based techniques such as ARD and BEM, over the entire domain, our approach is much faster as we use a numerical solver only in near-object regions, as opposed to the entire volume. We do not have a parallel BEM implementation, but extrapolating from the data in Figure 6, FastBEM would take $100+$ hours for Underground Parking Garage and $1000+$ hours for Reservoir on a 500-core cluster to simulate sound up to 1 kHz, assuming full parallelization. In comparison, our method can perform all (numeric, geometric, and coupling) precomputations in a few hours for these two scenes (as shown in Table 1) to achieve interactive runtime performance (see Table 2). Moreover, numerical techniques typically require memory proportional to the third or fourth power of frequency to evaluate pressures and compute I's at different listener positions. As shown in Table 3, our method requires orders of magnitude less memory than several standard numerical techniques. We have also highlighted the relative benefits of our two-way coupling algorithms with other hybrid methods used in acoustic and electromagnetic simulation (see Section 2.3). In many ways, our coupling algorithm ensures continuity and consistency of the field computed by numeric and geometric techniques at the artificial boundary between their computational domains.

The method proposed by Mehra et al. [2013] is also able to simulate the acoustic effects of objects in large outdoor scenes. Their formulation, however, only allows objects to be situated in an empty space or on an infinite flat ground, and therefore cannot model large indoor scenes (e.g. parking lot) or outdoor scenes with uneven terrains. If an outdoor scene has a large object, the algorithm pro-

posed in [Mehra et al. 2013] would slow down considerably. The coupling with geometric propagation algorithm, on the other hand, enables us to model acoustic interactions with all kinds of environment features. It is relatively easier to extend our hybrid approach to inhomogeneous environments by using curved ray tracing. Furthermore, geometric ray tracing is also used to perform frequency decomposition and this results in improved sound rendering.

Scene	air vol. (m^3)	surf. area (m^2)	FDTD	ARD	BEM/ FMM	Ours
Bldg+small	1800	660	0.2 TB	5 GB	6 GB	**12 MB**
Bldg+med	3200	1040	0.3 TB	9 GB	9 GB	**12 MB**
Bldg+large	22400	3840	2.2 TB	60 GB	34 GB	**12 MB**
Reservoir	5832000	32400	578 TB	16 TB	307 GB	**42 MB**
Parking	9000	2010	0.9 TB	24 GB	2 GB	**9 MB**

Table 3: Memory Cost Saving. *The memory required to evaluate pressures at a given point of space. This corresponds to the same operation shown in the rightmost column of Table 1. Compared to standard numerical techniques, our method provides* **3 to 7 orders of magnitude** *of memory saving on the benchmark scenes.*
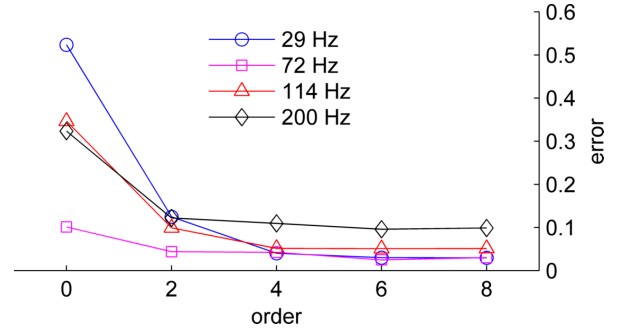

Figure 6: *Error $||P_{ref} - P_{hybrid}||^2/||P_{ref}||$ between the reference wave solver (BEM) and our hybrid technique for varying maximum order of reflections modeled. The tested scene is the "Two walls in a room" (see also Figure 5, last column).*
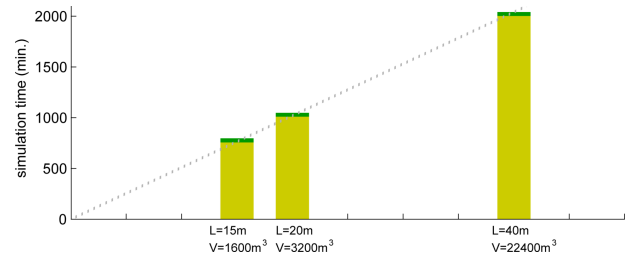
Figure 7: Breakdown of Precomputation Time. *For a building placed in terrains of increasing volumes (small, medium, and large walls), the yellow part is the simulation time for the numerical method, and the green part is for the geometric method. The numerical simulation time scales linearly to the largest dimension (L) of the scene instead of the total volume (V).*

7 Limitations, Conclusion, and Future Work

We have presented a novel hybrid technique for sound propagation in large indoor and outdoor scenes. The hybrid technique combines the strengths of numerical and geometric acoustic techniques for the different parts of the domain: the more accurate and costly numer-
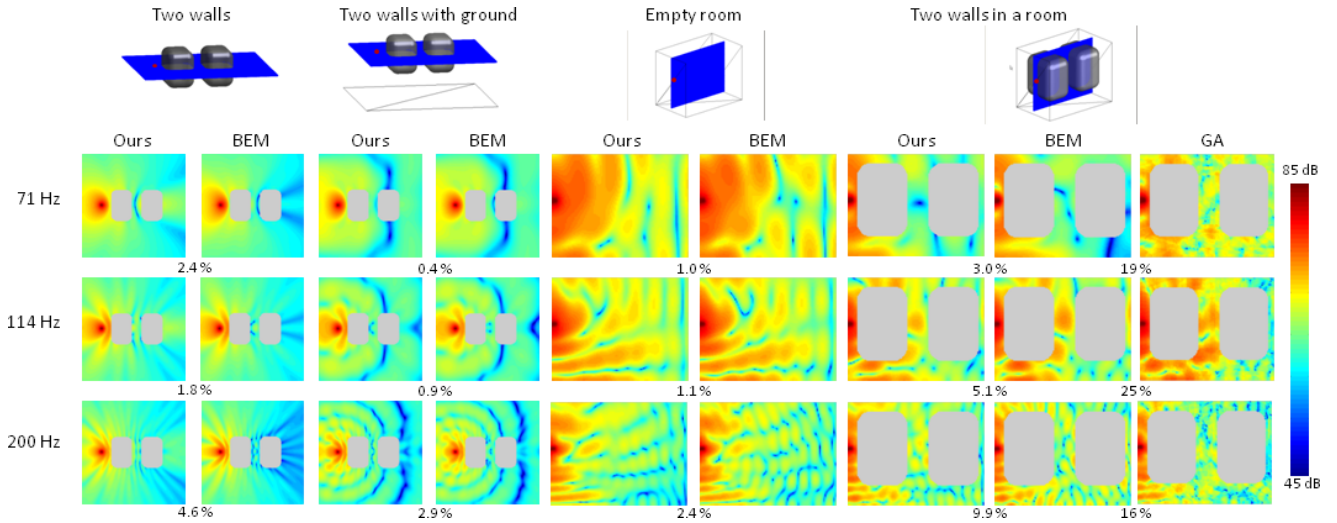
Figure 5: *Comparison between the magnitude of the total pressure field computed by our hybrid technique and BEM for various scenes. In the top row, the red dot is the sound source, and the blue plane is a grid of listeners. Errors between our method and BEM for each frequency are shown in each row. For our hybrid technique, the effect of the two walls are simulated by numerical acoustic techniques, and the interaction between the ground or the room is handled by geometric acoustic techniques. For BEM, the entire scene (including the walls, ground, and room) is simulated together. The last column also shows comparison with a pure geometric technique (marked as "GA").*

ical technique is used to model wave phenomena in near-object regions, while the more efficient geometric technique is used to handle propagation in far-field regions and interaction with the environment. The sound pressure field generated by the two techniques is coupled using a novel two-way coupling procedure. The method is successfully applied to different scenarios to generate realistic acoustic effects.

Our approach has a few limitations. Currently our geometric technique assumes homogeneous medium and traces straight ray paths. However, in the case of inhomogeneous medium, where the speed of sound is not constant and the rays may travel in curved paths, a nonlinear ray-tracing module can be integrated into our hybrid system instead.

The performance of our spatial decomposition depends greatly on the size of Ω^N. Although it size is smaller than the entire simulation domain, an individual Ω^N may still be too large, especially when the wave effects near a large object need to be computed and this increases the complexity of our algorithm. One interesting topic to investigate is the possibility of not enclosing the whole object, but only parts of it (e.g. small features) in Ω^N.

We currently compare our simulation results with simulated data from a high-accuracy BEM solver. It would be an important future work to validate these results with recorded audio measurements, when accurate measurements with binaural sound recordings and spatial sampling in complex environments are available.

Additionally our approach and system implementation is currently limited to mostly static scenes with moving sound sources and/or listeners. Nonetheless the use of transfer functions lays the foundation for future extension to fully dynamic scenes, as the per-object transfer functions of an object can be reused even when the object is moved. In order to recompute inter-object transfers as multiple objects move in a dynamic scene, a large number of rays (the number of outgoing sources for all frequency samples multiplied by thousands of rays emitted per source) need to be retraced. We would like to explore the use of the Fast Multipole Method (FMM) [Gumerov and Duraiswami 2004] to reduce the number of outgoing sources for far-field approximations. The computation of transfer function is currently implemented with unoptimized MATLAB code, and using high-performance linear solvers (CPU- or GPU-based) can greatly improve the performance.

Acknowledgements

References

ANTANI, L., CHANDAK, A., SAVIOJA, L., AND MANOCHA, D. 2012. Interactive sound propagation using compact acoustic transfer operators. *ACM Trans. Graph. 31*, 1 (Feb.), 7:1–7:12.

ARETZ, M. 2012. *Combined wave and ray based room acoustic simulations of small rooms: challenges and limitations on the way to realistic simulation results*. PhD thesis, Aachener Beitrge zur Technischen Akustik.

ARFKEN, G. B., WEBER, H. J., AND RUBY, L. 1985. *Mathematical methods for physicists*, vol. 3. Academic press San Diego.

BARBONE, P. E., MONTGOMERY, J. M., MICHAEL, O., AND HARARI, I. 1998. Scattering by a hybrid asymptotic/finite element method. *Computer methods in applied mechanics and engineering 164*, 1, 141–156.

BEN-ARTZI, A., EGAN, K., DURAND, F., AND RAMAMOORTHI, R. 2008. A precomputed polynomial representation for interactive BRDF editing with global illumination. *ACM Transactions on Graphics (TOG) 27*, 2, 13.

BLAUERT, J. 1983. *Spatial hearing: The psychophysics of human sound localization*. MIT Press (Cambridge, Mass.).

BOTTELDOOREN, D. 1994. Acoustical finite-difference time-domain simulation in a quasi-Cartesian grid. *The Journal of the Acoustical Society of America 95*, 2313.

FUNKHOUSER, T., CARLBOM, I., ELKO, G., PINGALI, G., SONDHI, M., AND WEST, J. 1998. A beam tracing approach to acoustic modeling for interactive virtual environments. In *Proc. SIGGRAPH 1998*, 21–32.

FUNKHOUSER, T., TSINGOS, N., AND JOT, J.-M. 2004. Survey of methods for modeling sound propagation in interactive virtual environment systems. *Presence*.

GRANIER, E., KLEINER, M., DALENBCK, B.-I., AND SVENSSON, P. 1996. Experimental auralization of car audio installations. *Journal of the Audio Engineering Society 44*, 10, 835–849.

GUMEROV, N. A., AND DURAISWAMI, R. 2004. *Fast multipole methods for the Helmholtz equation in three dimensions*. Elsevier Science.

GUMEROV, N. A., AND DURAISWAMI, R. 2009. A broadband fast multipole accelerated boundary element method for the three-dimensional helmholtz equation. *J. Acoustical Society of America 125*, 1, 191–205.

HAMPEL, S., LANGER, S., AND CISILINO, A. P. 2008. Coupling boundary elements to a raytracing procedure. *International journal for numerical methods in engineering 73*, 3, 427–445.

JAMES, D. L., BARBIC, J., AND PAI, D. K. 2006. Precomputed acoustic transfer: output-sensitive, accurate sound generation for geometrically complex vibration sources. In *Proc. of ACM SIGGRAPH*, 987–995.

JEAN, P., NOE, N., AND GAUDAIRE, F. 2008. Calculation of tyre noise radiation with a mixed approach. *Acta Acustica united with Acustica 94*, 1, 91–103.

KOUYOUMJIAN, R. G., AND PATHAK, P. H. 1974. A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface. *Proc. of IEEE 62* (Nov.), 1448–1461.

LENTZ, T., SCHROEDER, D., VORLANDER, M., AND ASSENMACHER, I. 2007. Virtual reality system with integrated sound field simulation and reproduction. *EURASIP J. Applied Signal Processing*.

LOKKI, T., SOUTHERN, A., SILTANEN, S., AND SAVIOJA, L. 2011. Studies of epidaurus with a hybrid room acoustics modeling method. In *Acoustics of Ancient Theaters Patras, Greece*.

MEHRA, R., RAGHUVANSHI, N., ANTANI, L., CHANDAK, A., CURTIS, S., AND MANOCHA, D. 2013. Wave-based sound propagation in large open scenes using an equivalent source formulation. *ACM Trans. Graph. 32*, 2 (Apr.), 19:1–19:13.

MURPHY, D., SHELLEY, S., BEESON, M., MOORE, A., AND SOUTHERN, A. 2008. Hybrid room impulse response synthesis in digital waveguide mesh based room acoustics simulations. In *Proc. of the Int. Conference on Digital Audio Effects (DAFx-08)*.

OCHMANN, M. 1995. The source simulation technique for acoustic radiation problems. *Acta Acustica united with Acustica 81*, 6, 512–527.

PIERCE, A. D. 1989. *Acoustics: an introduction to its physical principles and applications*. Acoustical Society of America.

RAGHUVANSHI, N., NARAIN, R., AND LIN, M. C. 2009. Efficient and Accurate Sound Propagation Using Adaptive Rectangular Decomposition. *IEEE Transactions on Visualization and Computer Graphics 15*, 5, 789–801.

RAGHUVANSHI, N., SNYDER, J., MEHRA, R., LIN, M., AND GOVINDARAJU, N. 2010. Precomputed wave simualtion for real-time sound propagation of dynamic sources in complex scenes. *ACM Trans. on Graphics (Proc. of ACM SIGGRAPH) 29*, 3.

SOUTHERN, A., SILTANEN, S., AND SAVIOJA, L. 2011. Spatial room impulse responses with a hybrid modeling method. In *Audio Engineering Society Convention 130*.

SVENSSON, U. P., FRED, R. I., AND VANDERKOOY, J. 1999. An analytic secondary source model of edge diffraction impulse responses. *J. Acoustical Society of America 106*, 5, 2331–2344.

TAYLOR, M., CHANDAK, A., QI MO, LAUTERBACH, C., SCHISSLER, C., AND MANOCHA, D. 2012. Guided multiview ray tracing for fast auralization. *Visualization and Computer Graphics, IEEE Transactions on 18*, 11 (Nov.), 1797–1810.

THOMPSON, L. L. 2006. A review of finite-element methods for time-harmonic acoustics. *J. Acoustical Society of America 119*, 3, 1315–1330.

TSINGOS, N., FUNKHOUSER, T., NGAN, A., AND CARLBOM, I. 2001. Modeling acoustics in virtual environments using the uniform theory of diffraction. In *Proc. SIGGRAPH 2001*, 545–552.

TSINGOS, N. 2009. Pre-computing geometry-based reverberation effects for games. *35th AES Conference on Audio for Games*.

VLADIMIROV, V. S. 1976. Generalized functions in mathematical physics. *Moscow Izdatel Nauka 1*.

VORLANDER, M. 1989. Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *J. Acoustical Society of America 86*, 1, 172–178.

WANG, Y., SAFAVI-NAEINI, S., AND CHAUDHURI, S. 2000. A hybrid technique based on combining ray tracing and FDTD methods for site-specific modeling of indoor radio wave propagation. *Antennas and Propagation, IEEE Transactions on 48*, 5 (May), 743–754.

WATERMAN, P. C. 2009. T-matrix methods in acoustic scattering. *The Journal of the Acoustical Society of America 125*, 1, 42–51.