

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 22-03-2015		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 21-Feb-2013 - 21-Feb-2015	
4. TITLE AND SUBTITLE Final Report: Interactive Acoustic Simulation in Urban and Complex Environments			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER W911NF-13-C-0037		
			5c. PROGRAM ELEMENT NUMBER 665502		
6. AUTHORS Dinesh Manocha, Anish Chandak			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES Impulsonic, Inc. 305 Brookside Drive Chapel Hill, NC 27516 -2905			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 62557-CS-ST2.11		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT Report developed under Topic #A11a-T006, contract W911NF-13-C-0037. Outdoor acoustics simulation plays a vital role in several military and commercial applications. State-of-the-art outdoor acoustics prediction methods are restricted, slow, or limited to simple environments. Specifically, no existing method can simultaneously model atmospheric acoustical phenomena, surface interactions due to complex objects, along with dynamic sound sources and listeners. We developed a multi-domain simulation approach that combines efficient non-linear ray tracing with accurate wave solver to model long range atmospheric effects as well as scattering due to complex objects.					
15. SUBJECT TERMS STTR Report, outdoor acoustics, equivalent source method, acoustic transfer operator, hybrid acoustics simulation, non-linear ray tracer, measurement and validation, numerical acoustic simulation, multi-core CPU					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Dinesh Manocha
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU			19b. TELEPHONE NUMBER 919-590-6049

Report Title

Final Report: Interactive Acoustic Simulation in Urban and Complex Environments

ABSTRACT

Report developed under Topic #A11a-T006, contract W911NF-13-C-0037. Outdoor acoustics simulation plays a vital role in several military and commercial applications. State-of-the-art outdoor acoustics prediction methods are restricted, slow, or limited to simple environments. Specifically, no existing method can simultaneously model atmospheric acoustical phenomena, surface interactions due to complex objects, along with dynamic sound sources and listeners. We developed a multi-domain simulation approach that combines efficient non-linear ray tracing with accurate wave solver to model long-range atmospheric effects as well as scattering due to complex objects. Accurate but computationally expensive wave solver models scattering effects near complex objects, while efficient non-linear ray tracer handles far-field interactions. The results from these solvers are combined using a novel two-way coupling based on equivalent source representation. Our technique is compute and memory efficient, requiring at least 100 times less memory and compute time than prior approaches. We validated our technique by comparing simulation results against analytical solutions and measurement data for outdoor scenes. The resulting acoustic simulation system can be used by the Army and DoD for battlefield planning, training, and simulation systems. It can also be used by federal agencies, acoustical consultants, and architects for urban noise modeling and reduction.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received

Paper

- | | | |
|------------|-------|---|
| 03/21/2015 | 1.00 | Ravish Mehra, Nikunj Raghuvanshi, Lakulish Antani, Anish Chandak, Sean Curtis, Dinesh Manocha. Wave-based sound propagation in large open scenes using an equivalent source formulation, ACM Transactions on Graphics, (04 2013): 0. doi: 10.1145/2451236.2451245 |
| 03/21/2015 | 2.00 | Hengchin Yeh, Ravish Mehra, Zhimin Ren, Lakulish Antani, Dinesh Manocha, Ming Lin. Wave-ray coupling for interactive sound propagation in large complex scenes, ACM Transactions on Graphics, (11 2013): 0. doi: 10.1145/2508363.2508420 |
| 03/21/2015 | 3.00 | Ravish Mehra, Lakulish Antani, Sujeong Kim, Dinesh Manocha. Source and Listener Directivity for Interactive Wave-Based Sound Propagation, IEEE Transactions on Visualization and Computer Graphics, (04 2014): 0. doi: 10.1109/TVCG.2014.38 |
| 03/21/2015 | 4.00 | Ravish Mehra, Atul Rungta, Abhinav Golas, Ming Lin, Dinesh Manocha. WAVE: Interactive Wave-based Sound Propagation for Virtual Environments, IEEE Transactions on Visualization and Computer Graphics, (04 2015): 0. doi: 10.1109/TVCG.2015.2391858 |
| 03/21/2015 | 5.00 | Carl Schissler, Ravish Mehra, Dinesh Manocha. High-order diffraction and diffuse reflections for interactive sound propagation in large environments, ACM Transactions on Graphics, (07 2014): 0. doi: 10.1145/2601097.2601216 |
| 03/21/2015 | 10.00 | Nikunj Raghuvanshi, Ravish Mehra, Anish Chandak, Donald G. Albert, D. Keith Wilson, Dinesh Manocha. Acoustic pulse propagation in an urban environment using a three-dimensional numerical simulation, The Journal of the Acoustical Society of America, (06 2014): 0. doi: 10.1121/1.4874495 |

TOTAL: 6

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

<u>Received</u>	<u>Paper</u>
03/21/2015	6.00 Qi Mo, Hengchin Yeh, Ming Lin, Dinesh Manocha. Analytic Ray Curve Tracing for Outdoor Sound Propagation, Applied Acoustics (03 2015)
03/21/2015	7.00 Qi Mo, Hengchin Yeh, Ming Lin, Dinesh Manocha. Outdoor Sound Propagation with Analytic Ray Curve Tracer and Gaussian Beam, JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA (03 2015)
TOTAL:	2

Number of Manuscripts:

Books

<u>Received</u>	<u>Book</u>
-----------------	-------------

TOTAL:

<u>Received</u>	<u>Book Chapter</u>
-----------------	---------------------

TOTAL:

Patents Submitted

Methods, systems, and computer readable media for simulating sound propagation in large scenes using equivalent sources

Sound Propagation in Large Complex Environments Using Wave-Ray Coupling

Dynamic Data Drive Source Directivity for Interactive Wave-based Sound Propagation

High-Order Diffraction and Diffuse Reflections For Interactive Sound Propagation in Large Environments

Interactive Sound Propagation and Rendering for Scenes with High Aural Complexity

Patents Awarded

Awards

Royster Student Scholarship Award, Acoustical Society of America, 2013

Link Foundation Fellowship in Advanced Simulation and Training, Link Foundation, 2013

Graduate Students

<u>NAME</u>	<u>PERCENT SUPPORTED</u>	Discipline
Hengchen Yeh	0.50	
Carl Schissler	0.25	
Qi Mo	0.50	
Ravish Mehra	0.50	
Alok Mesharam	0.50	
FTE Equivalent:	2.25	
Total Number:	5	

Names of Post Doctorates

<u>NAME</u>	<u>PERCENT SUPPORTED</u>
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

<u>NAME</u>	<u>PERCENT SUPPORTED</u>
FTE Equivalent:	
Total Number:	

Names of Under Graduate students supported

<u>NAME</u>	<u>PERCENT SUPPORTED</u>
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:..... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):..... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:..... 0.00

Names of Personnel receiving masters degrees

NAME

Alok Mesharam

Total Number: 1

Names of personnel receiving PHDs

NAME

Hengchin Yeh

Ravish Mehra

Total Number: 2

Names of other research staff

NAME

PERCENT SUPPORTED

FTE Equivalent:

Total Number:

Sub Contractors (DD882)

1 a. University of North Carolina - Chapel Hill

1 b. 104 Airport Drive, CB 1350

Suite 2200

Chapel Hill NC 275991350

Sub Contractor Numbers (c):

Patent Clause Number (d-1):

Patent Date (d-2):

Work Description (e):

Sub Contract Award Date (f-1): 2/21/13 12:00AM

Sub Contract Est Completion Date(f-2): 2/21/15 12:00AM

1 a. University of North Carolina - Chapel Hill

1 b. 440 W. Franklin Street

Campus Box 1350

Chapel Hill NC 275991350

Sub Contractor Numbers (c):

Patent Clause Number (d-1):

Patent Date (d-2):

Work Description (e):

Sub Contract Award Date (f-1): 2/21/13 12:00AM

Sub Contract Est Completion Date(f-2): 2/21/15 12:00AM

1 a. University of North Carolina - Chapel Hill

1 b. 104 Airport Drive, Cb 1350

Chapel Hill NC 275991350

Sub Contractor Numbers (c):

Patent Clause Number (d-1):

Patent Date (d-2):

Work Description (e):

Sub Contract Award Date (f-1): 2/21/13 12:00AM

Sub Contract Est Completion Date(f-2): 2/21/15 12:00AM

Inventions (DD882)

5 Dynamic Data Drive Source Directivity for Interactive Wave-based Sound Propagation

Patent Filed in US? (5d-1) Y

Patent Filed in Foreign Countries? (5d-2) N

Was the assignment forwarded to the contracting officer? (5e) N

Foreign Countries of application (5g-2):

5a: Lakulish Antani

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Ravish Mehra

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Dinesh Manocha

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5 High-Order Diffraction and Diffuse Reflections For Interactive Sound Propagation in Large Environments

Patent Filed in US? (5d-1) Y

Patent Filed in Foreign Countries? (5d-2) N

Was the assignment forwarded to the contracting officer? (5e) N

Foreign Countries of application (5g-2):

5a: Carl Schissler

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Dinesh Manocha

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5 Interactive Sound Propagation and Rendering for Scenes with High Aural Complexity

Patent Filed in US? (5d-1) Y

Patent Filed in Foreign Countries? (5d-2) N

Was the assignment forwarded to the contracting officer? (5e) N

Foreign Countries of application (5g-2):

5a: Dinesh Manocha

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Carl Schissler

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5 Methods, systems, and computer readable media for simulating sound propagation in large scenes using equivalent sources

Patent Filed in US? (5d-1) Y

Patent Filed in Foreign Countries? (5d-2) N

Was the assignment forwarded to the contracting officer? (5e) N

Foreign Countries of application (5g-2):

5a: Ravish Mehra

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Dinesh Manocha

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5 Sound Propagation in Large Complex Environments Using Wave-Ray Coupling

Patent Filed in US? (5d-1) Y

Patent Filed in Foreign Countries? (5d-2) N

Was the assignment forwarded to the contracting officer? (5e) N

Foreign Countries of application (5g-2):

5a: Ming C. Lin

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Dinesh Manocha

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Lakulish Antani

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Zhimin Ren

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Hengchin Yeh

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

5a: Ravish Mehra

5f-1a: University of North Carolina at Chapel Hill

5f-c: The Department of Computer Science Brooks Compute
Chapel Hill NC 27599

Scientific Progress

See Attachment.

Technology Transfer

The patents generated through the research work performed under the contract is licensed exclusively to Impulsonic, Inc. under the Carolina Express License.

Collaboration with Donal G. Albert and D. Keith Wilson from ERDC-CRREL-NH on "Acoustic Pulse Propagation in an Urban Environment using an three-dimensional Numerical Simulation"

Final Report: Interactive Acoustic Simulation in Urban and
Complex Environments

Dinesh Manocha, Anish Chandak

March 21, 2015

Technical Abstract

Report developed under Topic #A11a-T006, contract W911NF-13-C-0037. Outdoor acoustics simulation plays a vital role in several military and commercial applications. State-of-the-art outdoor acoustics prediction methods are restricted, slow, or limited to simple environments. Specifically, no existing method can simultaneously model atmospheric acoustical phenomena, surface interactions due to complex objects, along with dynamic sound sources and listeners. We developed a multi-domain simulation approach that combines efficient non-linear ray tracing with accurate wave solver to model long-range atmospheric effects as well as scattering due to complex objects. Accurate but computationally expensive wave solver models scattering effects near complex objects, while efficient non-linear ray tracer handles far-field interactions. The results from these solvers are combined using a novel two-way coupling based on equivalent source representation. Our technique is compute and memory efficient, requiring at least 100 times less memory and compute time than prior approaches. We validated our technique by comparing simulation results against analytical solutions and measurement data for outdoor scenes. The resulting acoustic simulation system can be used by the Army and DoD for battlefield planning, training, and simulation systems. It can also be used by federal agencies, acoustical consultants, and architects for urban noise modeling and reduction.

Key Words: STTR Report, outdoor acoustics, equivalent source method, acoustic transfer operator, hybrid acoustics simulation, non-linear ray tracer, measurement and validation, numerical acoustic simulation, multi-core CPU

Contents

1	Introduction	19
1.1	Phase II Key Results	20
1.2	Organization	23
2	Wave-Based Sound Propagation in Large Open Scenes using an Equivalent Source Formulation	24
2.1	Overview	24
2.2	Introduction	25
2.3	Related work	27
2.3.1	Numerical Wave Solvers	27
2.3.2	Interactive Geometric Techniques	29
2.3.3	Interactive Wave-simulation Techniques	29
2.4	The Equivalent Source Method	30
2.5	Sound propagation using ESM	32
2.5.1	Our Approach	33
2.5.2	Offset Surface Calculation	34
2.5.3	Per-object Transfer Function	34
2.5.4	Inter-object Transfer Function	36
2.5.5	Computing Equivalent Source Positions	39

2.5.6	Global Solve	40
2.5.7	Runtime Computation	41
2.6	Implementation	42
2.7	Results and Analysis	46
2.7.1	Scenarios	46
2.7.2	Error Analysis	48
2.7.3	Computational Complexity	48
2.7.4	Comparison with Prior Interactive Techniques	50
2.8	Conclusions and Discussion	51
2.9	Appendix	53
2.9.1	Two-object Steady State Field	53
2.9.2	Multiple Objects Steady State Field	55
2.9.3	Computational Complexity	55
3	Wave-Ray Coupling for Interactive Sound Propagation in Large Complex Scenes	59
3.1	Overview	59
3.2	Introduction	60
3.3	Related Work	62
3.3.1	Numerical Acoustic Techniques	62
3.3.2	Geometric Acoustic Techniques	63
3.3.3	Hybrid Techniques	63
3.3.4	Acoustic Kernel-Based Interactive Techniques	64
3.4	Overview	65
3.4.1	Sound Propagation	65
3.4.2	Acoustic Transfer Function	66
3.4.3	Hybrid Sound Propagation	66
3.5	Two-Way Wave-Ray Coupling	69

3.5.1	Geometric $\rightarrow$ Numerical	69
3.5.2	Numerical $\rightarrow$ Geometric	70
3.5.3	Fundamental solutions	72
3.5.4	Precomputed Transfer Functions	74
3.6	Implementation	75
3.6.1	Implementation details	77
3.6.2	Collocated equivalent sources	78
3.6.3	Auralization	78
3.7	Results and Analysis	79
3.7.1	Scenarios	79
3.7.2	Error Analysis	80
3.7.3	Complexity	81
3.7.4	Comparison with Prior Techniques	83
3.8	Limitations, Conclusion, and Future Work	85
4	Analytic Ray Curve Tracing for Outdoor Sound Propagation	88
4.1	Overview	88
4.2	Introduction	89
4.3	Prior Work	90
4.4	Analytic Ray Curve	92
4.4.1	Ray formulation	92
4.4.2	Ray properties	93
4.5	Ray Tracing Algorithm	96
4.5.1	Tracing Analytic Segments	96
4.5.2	Adaptive Media Traversal	97
4.5.3	Handling Boundary Surfaces	97
4.6	Results	98

4.6.1	Performance Comparison of Ray Models	98
4.6.2	Outdoor Applications of Analytic Ray Tracer	103
4.7	Discussions	105
4.8	Conclusion	106
5	Outdoor Sound Propagation with Analytic Ray Curve Tracer and Gaussian Beam	108
5.1	Overview	108
5.2	Introduction	109
5.3	Prior Work	110
5.3.1	Wave-based Methods	110
5.3.2	Geometric Acoustics Methods	111
5.4	Algorithm	113
5.4.1	Analytic Ray Curve Tracer	113
5.4.2	Dynamic Ray Tracing	114
5.4.3	Field Computation with Gaussian Beam	115
5.5	Validation	117
5.5.1	Benchmark A	117
5.5.2	Benchmark B	119
5.6	Application on General Outdoor Scene	119
5.6.1	Scene configuration	119
5.6.2	Results	120
5.7	Discussion	123
5.8	Conclusion	123
5.9	Derivations	124
5.9.1	Analytic evolution of ray trajectories	124
5.9.2	Analytic evolution of Cartesian P and Q	125
5.9.3	Analytic evolution of tranformation matrix	125

5.9.4	Evolution of ray-centered P and Q	126
5.9.5	Cartesian paraxial travel time	126
5.9.6	Gaussian beam summation	127
6	Acoustic pulse propagation in an urban environment using a three-dimensional numerical simulation	128
6.1	Overview	128
6.2	Introduction	129
6.3	Previous work	131
6.4	Measurements	132
6.5	Adaptive rectangular decomposition numerical modeling	134
6.5.1	The adaptive rectangular decomposition method	134
6.5.2	Validation	137
6.5.3	Source waveform	138
6.5.4	3D model	138
6.5.5	Simulation parameters	139
6.5.6	3D vs 2D wave simulation	139
6.6	Results	139
6.6.1	Simulation	140
6.6.2	Varying propagation speed	141
6.6.3	Error metrics	141
6.6.4	Comparison with measurements in time domain	143
6.6.5	Comparison with 2D FDTD	145
6.6.6	Comparison with measurements in spatial domain	147
6.7	Conclusion, Limitations and Future Work	148

List of Figures

1.1	<i>An overview of our multi-domain acoustic simulation approach. The circles denote source (S) and receiver (R), and the obstacles are denoted by the numbers 1 and 2. The scene is decomposed into sub-domains, i.e. boundary regions around the obstacles and the free region. Sound propagation from the source to the boundary regions is computed, followed by boundary-region computations, free-region computations, and interface-handling operations. Finally, the sound is propagated from the boundary regions to the receiver to compute the final output response.</i>	20
2.1	Our algorithm accurately models realistic acoustic effects, such as diffraction, scattering, focusing, and echoes, in large, open scenes at real-time rates. We reduce the runtime memory usage by orders of magnitude compared to state-of-the-art wave solvers, enabling real-time, wave-based sound propagation in scenes spanning hundreds of meters: a) reservoir scene (from Half-Life 2), b) Christmas scene, and c) desert scene.	27
2.2	a) A diagram illustrating a radiating object A , its corresponding boundary ∂A , exterior region A^+ , interior region A^- , and the set of equivalent sources (denoted by star shapes). b) Classification of objects in a scene. The triangle and rectangle constitute a single object, as their offset surfaces overlap. On the other hand, L-shaped shapes are classified as separate objects.	32

2.3	Overview of our wave-based sound propagation technique based on equivalent sources on a simple scene composed of two objects and a sound source (shown with a red dot). The magnitudes of pressure fields are visualized using the color scheme shown.	33
2.4	Magnitude of the pressure field (in Pa) at 170 Hz in a simple scene with a single object (rocks) and a single sound source (red dot). The difference between total and incident fields is the scattered field (scaled eight times for visualization). Note the high amplitude of the scattered field between the rocks representing the large difference in incident and total field that results from diffracted occlusion.	37
2.5	We show the scattering behavior of a wall ($2.3m \times 4.5m \times 3.7m$) and a concave reflector (diameter $8m$, thickness $1.2m$) at $160Hz$ using our technique (top row) and BEM (bottom row). The sound source is shown with a red dot.	57
2.6	Convergence: We show the variation of error $\ P_{ref} - P_{ESM}\ ^2 / \ P_{ref}\ ^2$ between the reference wave solver and our technique for varying values of scattering matrix error threshold σ for the two parallel walls scene (fixed $\eta = 1\%$). P_{ref} and P_{ESM} are vectors consisting of (complex) pressure values at all the receiver locations as given by the reference wave solver and our technique, respectively. The receivers are placed on a XY grid for this scene.	57
2.7	Variation of the number of outgoing equivalent sources with frequency, for four different objects. As the frequency increases (wavelength decreases), surface details of the size of the wavelength increase the complexity of the sound field. This results in a larger number of equivalent sources. When all the details of the object are captured, increasing the frequency has little effect and the number of equivalent sources begin to stabilize. Error thresholds are $\sigma = 15\%$ and $\eta = 1\%$	58

2.8	Comparison between the magnitude of the total pressure field (in Sound Pressure Level SPL, units dB) computed by our ESM, the reference wave-solver ARD, BEM and FMM-BEM techniques for the two parallel walls scene on a XY cutview grid of listeners. The red point denotes the position of the sound source. The error (defined in Figure 2.6 caption) between the ARD-ESM fields is $< 3\%$, BEM-ESM fields is $< 5\%$ and FMM BEM-ESM fields is $< 5\%$ for the frequencies shown.	58
3.1	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.	62
3.2	Overview of spatial decomposition in our hybrid sound propagation technique: In the pre-computation 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.	65
3.3	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.	67

3.4	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	69
3.5	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").	84
3.6	Error $\ P_{\text{ref}} - P_{\text{ref}}\ ^2 / \ P_{\text{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 3.5, last column).	85
3.7	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).	86

4.1	(a) Parabolic ray curve: shown in r - z plane, z -axis being the direction of ∇V^{-2} . Rays with two different initial directions (marked by arrows) are shown here. (x_f, z_f) defining the vertex of the parabola is computed from Eq. 4.6. (b) Efficiency of curve tracing: with constant $\ \nabla V^{-2}\ $, one segment of analytic curve (blue line) suffices as compared to many steps (magenta circles) taken by numerical ray integration. The trajectory traced out by RK-4 visibly diverges from the parabolic curve, which can be corrected by further reducing step sizes. (c) Adaptive segments of curves: with logarithmic sound speed profile, our ray tracer proceeds in segments of parabolic curves with adaptive sizes, evident from the visualization of the spheres of validity (Sec. 4.5.2).	93
4.2	Left steps of tracing one segment of analytic curve, among which the green-colored blocks have analytic solutions due to the parabolic ray formulation, leading to large savings of ray tracing cost. Right Besides the analytic formulation, other sources of performance improvements lead to our ray tracer tracing fewer segments with fewer intersection tests, and cheaper intersection tests as well. Those algorithmic improvements are connected to the steps in the flow chart where they happen.	95
4.3	Downward refracting atmosphere: ray plot of 201 rays computed by the analytic ray curve tracer.	100
4.4	Left: Cost-accuracy trade-off of RK 4, to achieve lower error the step size of RK4 needs to decrease, which leads to exponential growth in computation time. Right: Cost with increasing media gradient grows as both RK 4 and ray curve tracer adopt smaller steps to keep the error below 2×10^{-4} , but RK 4 shows both higher cost and lower scalability with media gradient. .	101
4.5	The range of validity for each segment of ray curve is controlled by the parameter δ and adapts to local media gradient. (a)Munk profile sound speed. (b,c,d) Segment sizes at spatial locations within the Munk profile visualized, with $\delta = 0.002, 0.001, 0.0005$ respectively.	102
4.6	Munk profile with 200km range, ray tracing cost-accuracy trade-off of Left: Runge-Kutta 4 and Right: analytic ray curve.	102

4.7	Munk profile characteristic ray plot computed by analytic ray tracer.	103
4.8	Outdoor benchmarks Desert and Christmas village. Example ray paths are shown with upward and downward refraction. The media gradient is exaggerated and the ray path number is kept very small for illustration purpose. See Table 4.3 for scene stats and actual performance numbers of tracing 10k rays up to 3 orders of reflections.	104
4.9	Spatial visualization of ray tracing cost: brighter color indicates higher intersection costs. It is shown that the intersection computation concentrates in areas of dense surface geometry and around silhouette of obstacles, which illustrates that the spatial culling of ray-surface intersections using BVH is effective. Left: Desert Right: Christmas village.	104
5.1	Algorithm overview.	112
5.2	(a) the analytic ray curve in ray-plane. Rays of two different θ_0 are drawn in blue and magenta, respectively. (x_f, z_f) is the vertex of the parabola. Assuming a locally constant ∇V^{-2} , of which the range of validity determines the extent of the ray segment. The two red dots mark the end points of ray segments, where the rays exit from this sphere of validity (Sec. A). (b,c) analytic evolution of P , Q are performed for each segment by transforming into and evolving in Cartesian coordinates before transforming back to ray-centered coordinates (Sec. B).	114
5.3	Benchmark A Range-TL plot: Source height $h_s = 5m$, receiver height $h_r = 1m$, range $10km$. Columns contain results for media profiles case 1, 2, and 3, while rows contain results for frequency 10, 100, and 1000 Hz. Comparison with Fig. 12-14[12].	118
5.4	Benchmark A 2D field: Source height $h_s = 5m$, receiver height $h_r = 1m$, frequency 10Hz. 2D vertical field of height up to $1km$ and range up to $10km$ is visualized on the left, and the corresponding contour plot is shown on the right for comparison with Fig. 15[12].	118
5.5	Benchmark B 2D field: Source depth 1000m, frequency 50Hz. 2D vertical field of TL is visualized for depth up to $5km$ and range up to $200km$, compared to Fig. 9[93].	120

5.6	Outdoor scene and field results for two source locations: (a) Wireframe rendering of the Reservoir scene, green dot represents a sound source located on the slope, and the red dot represents a sound source in the valley. (b) Slices of sound TL field visualized for the green source. (c) Slices of sound TL level visualization for the red source. Frequency 10Hz.	121
5.7	Vector wind for source in the valley: (a) Difference in horizontal field of sound pressure level between a north and a sound wind. (b) Difference in horizontal field of sound pressure level between an east and a west wind. (c) Difference in vertical field of sound pressure level between an east and a west wind. Frequency 10Hz.	121
5.8	Source on the slope: (a) Relative sound pressure level upward vs downward refraction, vertical field, (b) horizontal field. (c) Relative sound pressure level up wind vs down wind propagation, vertical field, (d) horizontal field. Frequency 10Hz.	122
5.9	Source in the valley: Relative sound pressure level upward vs downward refraction, (a) horizontal and (b) vertical field that passes the source location. Frequency 10Hz.	122
6.1	a) Top view of the urban scene used in the experimental study. Reproduced with permission from Liu and Albert (2010). b) An approximate 3D model of the scene constructed based on the 2D layout, photographs of the scene and heights of the buildings corners and roof tops. (Color online.)	133
6.2	Different stages of the Adaptive Rectangular Decomposition simulator. (Color online.) . . .	135
6.3	Validation of the ARD simulation results (dots) with analytical expressions (curves) for the scattering of a spherical wave by a rigid sphere. Normalized pressure is plotted in the radial axis. The radius of the sphere $a = 1m$ and wave numbers k considered are 0.7, 1.3 and $2 m^{-1}$, respectively. (Color online.)	149
6.4	Time and frequency responses produced by the Biot-Tolstoy-Medwin diffraction model (reference) and the ARD simulation for a right-angled rigid wall. (Color online.)	150
6.5	The source pulse used for modeling the blast signal produced in the experiment as calculated from Eq.(5) (Liu and Albert, 2006).	151

6.6	Calculated acoustic response for the source position SP1 in the artificial village scene using the ARD technique. Simulated wavefields are shown at times $t=75, 150, 225$, and 300 ms.	152
6.7	Waveforms calculated by the 3D ARD (upper) and 2D FDTD (lower) simulations and measured (middle) in the artificial village for source position 1 at 14 receiver positions. All the responses have been individually normalized and low-passed to the maximum frequency of 200 Hz. The error between the 2D FDTD simulation and measurement (dashed line) and 3D ARD simulation and measurement (solid line) has been calculated using the DTW-based SDM and ADM metric.	153
6.8	Waveforms calculated by the 3D ARD simulator (upper) and measured (lower) in the artificial village for source position 2 at 14 receiver positions. All the waveforms have been individually normalized and low-passed to the maximum frequency of 450 Hz. The error between the 3D ARD simulation and measurement has been calculated using the DTW-based SDM and ADM metric.	154
6.9	Waveforms calculated by the 3D ARD simulator (upper) and measured (lower) in the artificial village for source position 3 at 14 receiver positions. All the responses have been individually normalized and low-passed to the maximum frequency of 450 Hz. The error between the 3D ARD simulation and measurement has been calculated using the DTW-based SDM and ADM metric.	155
6.10	Comparison between the calculated and measured waveforms for the source position SP1 and receiver R09 behind building A. The upper trace (solid line) corresponds to the 3D ARD waveform whereas the lower trace (solid line) corresponds to the 2D FDTD waveform. The measured waveform is drawn as dotted line. Note that the 2D FDTD simulation cannot model the diffraction path from the building's rooftop resulting in the missing first arrival at 60 ms.	156

6.11	Waveforms calculated by the 3D ARD (upper) and 2D FDTD (lower) simulations and measured (middle) in the artificial village for source position 4 at 14 receiver positions. All the responses have been individually normalized and low-passed to the maximum frequency of 200 Hz. The error between the 2D FDTD simulation and measurement (dashed line) and 3D ARD simulation and measurement (solid line) has been calculated using the DTW-based SDM and ADM metric.	157
6.12	Rooftop diffraction Calculated acoustic response for the source position SP1 in the artificial village scene using the ARD technique. Receiver position R09 is marked by a white circle on the right of the first building (from right). Simulated wavefields are shown at times $t = 20, 30, 40, 50, 61, 83, 104, 161$ and 177 ms corresponding to marked points in Figure 6.10.	158
6.13	Variation of error with distance between the source-receiver positions.	159

List of Tables

2.1	Table of commonly used symbols.	34
2.2	Parameters used in our system.	44
2.3	Performance statistics. Abbreviations are as follows : “#objs.” denotes the number of objects in the scene, “#freq.” is the number of frequency samples in the range [0-1 kHz] and “#srcs” is the number of sound sources. For the precomputation stage, the term “wave sim.” is the total simulation time of the numerical wave solver for all frequencies, “per-object” denote the compute time for the per-object transfer function for all unique objects and “inter-object” denote the compute time for inter-object transfer functions for all object pairs, “source-field” is time to express each sound source in terms of incoming multipoles for all objects, and “global-solve” is time to compute equivalent source strengths for all objects. The “wave sim.” step is parallelized over all unique objects whereas the remaining precomputation steps are parallelized over all frequencies. The term “wall-clk time” is the total wall-clock time computed by uniformly distributing all the parallel processes over all the cores of the 64-node cluster with 8 cores per node (512 cores in total). At runtime, the total number of equivalent sources “# eq. srcs” (in million M), performance “eval.” and storage requirement “storage” (fixed and per source cost) for all objects for all frequencies are also specified. For column “#objs.”, notation $a^* + b^*$ denotes that first object has been instanced a times and second object instanced b times, but their per-object transfer functions are computed only once for each unique object.	46

2.4	Runtime memory requirements per source, for FDTD [145], ARD [124], BEM/FMM-BEM [92], and our ESM technique with error thresholds $\sigma = 15\%$, $\eta = 1\%$ at maximum simulation frequency $\nu_{\max} = 1018Hz$. Refer to Section 2.7.3 and Appendix 2.9.3 for more details.	50
3.1	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).	76

3.2	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.	79
3.3	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 performance Table. Compared to standard numerical techniques, our method provides 3 to 7 orders of magnitude of memory saving on the benchmark scenes.	84
4.1	Same quality comparison. A fan of rays are traced under a downward refracting atmosphere to a range of 10km. At each level of relative error in the ray tracing results, the average number of ray curve segments per propagation path is compared with the average number of Runge-Kutta 4 steps. Analytic ray curve is able to achieve same level of accuracy with much less tracing cost.	100
4.2	Same speed comparison. A fan of rays are traced under a downward refracting atmosphere to a range of 10km. With comparable number of ray curve segments and Runge-Kutta 4 steps, our ray tracer (RT) is able to achieve lower levels of relative error in ray tracing results (both hit points and path length) than RK4.	101
4.3	Breakdown of ray tracing time: the computation of closed-form ray trajectory and closed-form evaluations of ray properties for each segment takes very little computation (less than 0.001% of total frame time) and thus is omitted here. The percentage that each kind of computation takes relative to the total frame time is included in parenthesis, together with the number of ray-BV and ray-surface intersections.	105
6.1	Parameters for the ARD technique.	140

Chapter 1

Introduction

Outdoor acoustics simulation, i.e., computer modeling of outdoor sound propagation, plays a vital role in military, government, and civilian applications. Outdoor acoustics simulation can be used to determine the vulnerability of battlefield resources in urban and outdoor warfare scenarios. For example, acoustics prediction can be used to plan the path of a manned or unmanned aerial vehicle over an urban environment while taking into account its noise profile, so as to minimize its audible range and probability of detection. It is also useful for ground troops to be able to pinpoint the origin of gunshots based on their acoustic signatures. In all these examples, rapidly changing battlefield situations would necessitate rapid prediction using acoustics simulation. Furthermore, these applications require the ability to handle complex and dynamic environments, in order to be deployed in the field.

As part of our STTR Phase II, we developed outdoor acoustics simulation technology that can meet these stringent requirements. Figure 1.1 provides an overview of our system. In addition to battlefield and security applications such as tactical planning and sensor optimization, efficient outdoor acoustics prediction is also needed for many commercial applications: it can be used to reduce the noise profile of aircraft, aid architects in reducing the impact of outdoor noise on their constructions, and help city planners and industrial operators in complying with noise level regulations.

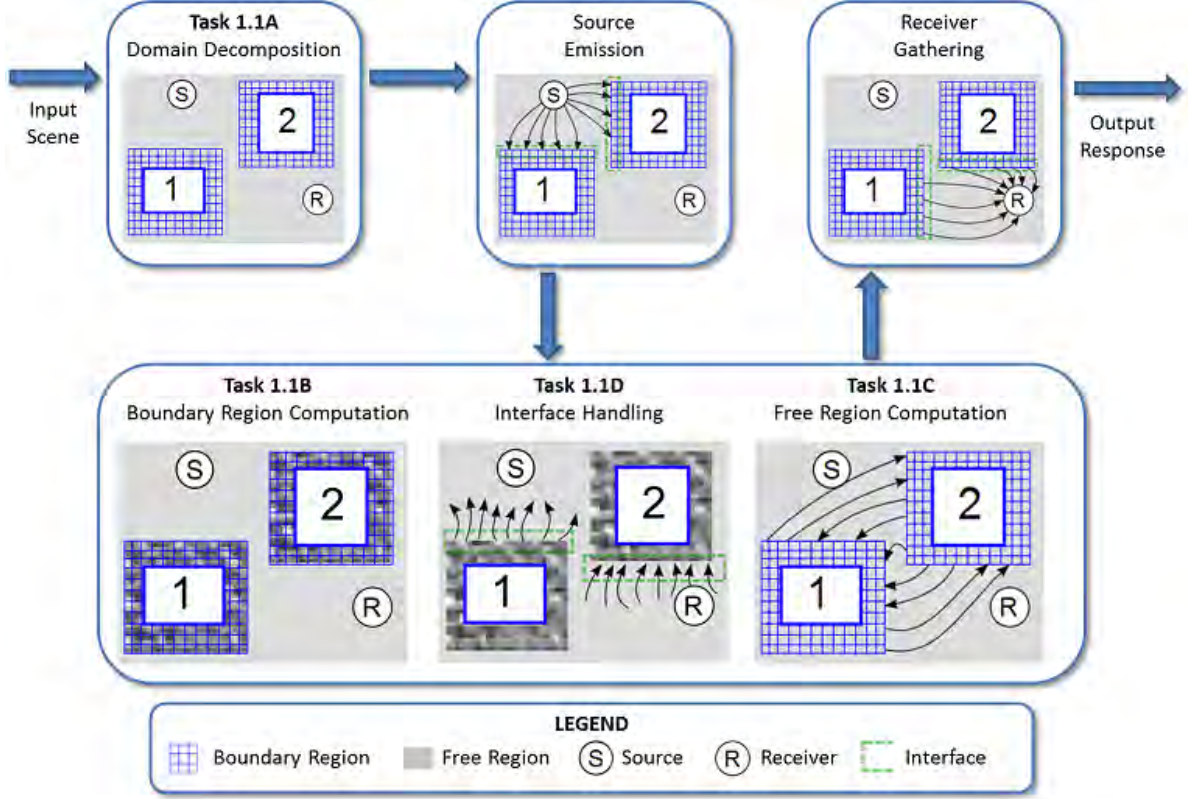


Figure 1.1: An overview of our multi-domain acoustic simulation approach. The circles denote source (S) and receiver (R), and the obstacles are denoted by the numbers 1 and 2. The scene is decomposed into sub-domains, i.e. boundary regions around the obstacles and the free region. Sound propagation from the source to the boundary regions is computed, followed by boundary-region computations, free-region computations, and interface-handling operations. Finally, the sound is propagated from the boundary regions to the receiver to compute the final output response.

1.1 Phase II Key Results

We developed a novel hybrid sound propagation approach suitable for modeling acoustics of large, open spaces spanning hundreds of meters, with a small memory footprint efficiently. At a high-level, the scene is decomposed into disjoint rigid objects. Compute-intensive wave-based methods are applied only around rigid objects, efficient non-linear ray tracing methods are used to model acoustics between the rigid objects while taking into account inhomogeneous medium properties, and finally combining the results from different techniques by coupling them at the boundary interfaces.

The acoustic behavior of each object is captured by a compact per-object transfer-function relating the amplitudes of a set of incoming equivalent sources to outgoing equivalent sources. Accurate wave-based

methods can be used to compute per-object transfer function as the size of these objects is much smaller compared to the overall scene. Pairwise acoustic interactions between objects can be computed either analytically for homogenous medium or using non-linear ray tracing for non-homogenous medium to yield compact inter-object transfer functions. The global sound field accounting for all orders of interaction is computed using these transfer functions.

In this section, we summarize our key Phase II accomplishments:

- *Pre-computed Outdoor Acoustics Transfer Operators:* We developed novel wave-based sound propagation algorithm that captures acoustic effects such as high-order diffraction and scattering, using an equivalent source formulation. Our technique can perform accurate sound propagation on large, open scenes in real-time, has a small memory footprint, and allows flexible efficiency-to-accuracy tradeoffs. Compared to directly storing and convolving wave-solver solutions for auralization, we reduce the memory usage more than 100 times.
- *Multi-domain Outdoor Acoustics Simulation:* We developed a novel hybrid technique for sound propagation that combines the strengths of numerical and geometric acoustic techniques for the different parts of the domain: the more accurate and costly numerical 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.
- *Efficient Non-Linear Ray Tracer:* We developed an efficient non-linear ray tracer for inhomogeneous media by tracing analytic ray curves based on local media gradients as primitives. Segments of ray curves are computed by sampling the media gradient on-the-fly, accounting for both inhomogeneous and moving media without the need to pre-compute explicit cell structures. Acceleration based on BVH is readily adapted to speed up surface intersections of the ray curves, enabling logarithmic scaling with scene complexity and allowing dynamic scenes.
- *Accuracy Analysis and Validation:* We compared the accuracy of our efficient pre-computed transfer

operators, multi-domain outdoor acoustics simulation, and non-linear ray tracer against analytical solutions and measurement data. Our results show good agreement between simulation and analytical or measurement data.

- *Performance Analysis:* We have performed performance analysis of our implementation of pre-computed transfer operators, multi-domain outdoor acoustics simulation, and efficient non-linear ray tracer.

Technical Submissions and Patents The work performed during Phase II has resulted in the following papers, which have either been accepted for publication or are under review:

1. Wave-based sound propagation in large open scenes using an equivalent source formulation. Ravish Mehra, Nikunj Raghuvanshi, Lakulish Antani, Anish Chandak, Sean Curtis, and Dinesh Manocha. 2013. ACM Trans. Graph. 32, 2, Article 19 (April 2013), 13 pages.
2. Wave-ray coupling for interactive sound propagation in large complex scenes. Hengchin Yeh, Ravish Mehra, Zhimin Ren, Lakulish Antani, Dinesh Manocha, and Ming Lin. 2013. ACM Trans. Graph. 32, 6, Article 165 (November 2013), 11 pages.
3. Source and Listener Directivity for Interactive Wave-Based Sound Propagation. Ravish Mehra, Lakulish Antani, Sujeong Kim, and Dinesh Manocha. 2014. IEEE Transactions on Visualization and Computer Graphics 20, 4 (April 2014), 495-503.
4. WAVE: Interactive Wave-based Sound Propagation for Virtual Environments. Ravish Mehra, Atul Rungta, Abhinav Golas, Ming C Lin, Dinesh Manocha. Visualization and Computer Graphics, IEEE Transactions on , vol.21, no.4, pp.434,442, April 18 2015.
5. Analytic Ray Curve Tracing for Outdoor Sound Propagation. Qi Mo, Hengchin Yeh, Ming Lin, Dinesh Manocha. Under review (Applied Acoustics)
6. Outdoor Sound Propagation with Analytic Ray Curve Tracer and Gaussian Beam. Qi Mo, Hengchin Yeh, Ming Lin, and Dinesh Manocha. Under review (The Journal of the Acoustical Society of America)

7. High-order diffraction and diffuse reflections for interactive sound propagation in large environments. Carl Schissler, Ravish Mehra, and Dinesh Manocha. 2014. ACM Trans. Graph. 33, 4, Article 39 (July 2014), 12 pages.
8. Acoustic pulse propagation in an urban environment using a three-dimensional numerical simulation. Ravish Mehra, Nikunj Raghuvanshi, Anish Chandak, Donald G. Albert, D. Keith Wilson, Dinesh Manocha. The Journal of the Acoustical Society of America, 135, 3231-3242 (2014)
9. Adaptive modeling of details for physically-based sound synthesis and propagation. Hengchin Yeh. PhD Thesis. August, 2014.
10. Efficient Techniques for Wave-Based Sound Propagation in Interactive Applications. Ravish Mehra. PhD Thesis. Under review.

We have also applied for five patents based on the work performed during our STTR Phase II.

1.2 Organization

As part of our Phase II efforts, we have developed the key components of a hybrid, multi-domain outdoor acoustics simulation algorithm, significantly advancing the state-of-the-art. The rest of the report is organized as follows: Chapter 2 discusses our wave-based sound propagation techniques for homeogenous medium based on pre-computed acoustic transfer operators. Chapter 3 discusses our novel hybrid approach that couples geometric acoustics/non-linear ray tracer and numerical acoustic techniques to for a multi-domain outdoor acoustics simulator. Chapter 4 presents our non-linear ray tracing method which uses analytic ray curve as tracing primitives to improve the efficiency of outdoor sound propagation in fully general settings. Chapter 5 and Chapter 6 analyze the accuracy of our algorithms.

Chapter 2

Wave-Based Sound Propagation in Large Open Scenes using an Equivalent Source Formulation

2.1 Overview

We present a novel approach for wave-based sound propagation suitable for large, open spaces spanning hundreds of meters, with a small memory footprint. The scene is decomposed into disjoint rigid objects. The free-field acoustic behavior of each object is captured by a compact per-object transfer-function relating the amplitudes of a set of incoming equivalent sources to outgoing equivalent sources. Pairwise acoustic interactions between objects are computed analytically to yield compact inter-object transfer functions. The global sound field accounting for all orders of interaction is computed using these transfer functions. The runtime system uses fast summation over the outgoing equivalent source amplitudes for all objects to auralize the sound field for a moving listener in real-time. We demonstrate realistic acoustic effects such as diffraction, low-passed sound behind obstructions, focusing, scattering, high-order reflections, and echoes, on a variety

of scenes.

2.2 Introduction

Interactive sound propagation has emerged as a powerful tool in computer graphics to enhance the realism of virtual worlds by predicting the behavior of sound as it interacts with the environment [146, 45, 94]. In order to accurately capture important acoustic phenomena in general scenarios, including interference, diffraction, scattering, sound focusing (caustics), and higher-order wave effects resulting from their combination, it is important to develop techniques that can directly solve the acoustic wave equation. There is extensive work in scientific computing and acoustics on numerical methods to solve the wave equation. Furthermore, there has been considerable interest in developing interactive wave-based techniques to model free-space sound radiation [71], first-order scattering from surfaces [156], and sound propagation for indoor scenes [135, 125].

Large, open scenes, which arise in many applications ranging from games to training or simulation systems, present a significant challenge for interactive, wave-based sound propagation techniques. State-of-the-art wave simulation methods can take hours of computation and gigabytes of memory for performing sound propagation in indoor scenes such as concert halls [131, 124]. For large, open scenes spanning hundreds of meters, it is challenging to run these techniques in real-time. On the other hand, geometric (ray-based) acoustic techniques can provide real-time performance for such environments. However, geometric techniques are better suited for higher frequencies due to the inherent assumption of rectilinear propagation of sound waves. Therefore, accurately modeling diffraction and higher-order wave effects with these techniques remains a significant challenge, especially at low frequencies.

In this paper, we present a novel approach for precomputed, wave-based sound propagation that is applicable to large, open scenes. It is based on the *equivalent source method*, which has been widely studied for radiation and scattering problems in acoustics and electromagnetics [40] and more recently introduced to computer graphics [71]. Our approach consists of two main stages: preprocessing and runtime. During preprocessing, we decompose the scene into disjoint, well-separated rigid objects. The acoustic behavior of

each object, taken independently, is characterized by its *per-object transfer function* that maps an arbitrary incident field on the object to the resulting scattered field. We propose an equivalent source formulation to express this transfer function as a compact *scattering matrix*. Pairwise acoustic coupling between objects is then modeled by computing *inter-object transfer functions* between all pairs of objects that maps the outgoing scattered field from one object to the incoming field on another object. These transfer functions are represented compactly by using the same equivalent source framework to yield *interaction matrices*. Acoustic transfer between multiple objects can therefore be represented using chained multiplication of their scattering and interaction matrices. Finally, the acoustic response of the scene to a static source distribution is computed by solving a global linear system that accounts for all orders of inter-object wave propagation.

At runtime, fast summation over all outgoing equivalent sources for all objects is performed at the listener location. The computed response is used for real-time sound rendering for a moving listener. Multiple moving sources, with a static listener, are handled by exploiting acoustic reciprocity. The runtime memory and computational requirements are proportional to the number of objects and their outgoing scattered field complexity (usually a few thousand equivalent sources per frequency for a few percent error), instead of the volume or surface area of the scene. Thus, our technique takes an object-centric approach to wave-based sound propagation. The key contributions of our work include:

- *Object-based sound field decomposition* using per-object and inter-object acoustic transfer functions for enabling real-time, wave-based sound propagation on large, open scenes.
- *Compact per-object transfer using equivalent sources* to model the scattering behavior of an object mapping arbitrary incident fields to the resultant scattered fields.
- *Compact analytical coupling of objects* is achieved by expressing inter-object transfer functions in the same, compact equivalent source basis as used for per-object transfer.
- *A fast, memory-efficient run-time* enables real-time sound rendering, while requiring only a few tens of megabytes of memory.

Our approach is well-suited for quick iterations while authoring scenes. Per-object transfer functions,

which take a significant portion of the precomputation time of our method, are independent of the scene and can thus be stored in a lookup table. Therefore, adding, deleting or moving a few objects in an existing scene has low precomputation overhead, linear in the number of objects.

We have tested our technique on a variety of scenarios (see Figure 2.1) and integrated our system with the Valve’s Source™ game engine from Half-Life 2. Our technique generates realistic acoustic effects and takes orders of magnitude less runtime memory compared to state-of-the-art wave solvers, enabling interactive performance. To the best of our knowledge, it is the first real-time technique for accurate, wave-based sound propagation in large, open scenes.



Figure 2.1: Our algorithm accurately models realistic acoustic effects, such as diffraction, scattering, focusing, and echoes, in large, open scenes at real-time rates. We reduce the runtime memory usage by orders of magnitude compared to state-of-the-art wave solvers, enabling real-time, wave-based sound propagation in scenes spanning hundreds of meters: a) reservoir scene (from Half-Life 2), b) Christmas scene, and c) desert scene.

2.3 Related work

Our technique has close theoretical parallels with prior numerical wave solvers. We first explore these connections, followed by related work on interactive geometric and wave-based techniques.

2.3.1 Numerical Wave Solvers

Research in wave-based acoustic simulation techniques spans a broad range of areas such as noise control, automotive design, urban architectural planning, and concert hall design. Wave solvers can be classified into *volumetric* and *surface-based* approaches. The most common among volumetric techniques are the finite element method (FEM) [183, 152] and finite difference time domain (FDTD) method [178, 145, 131],

which require a discretization of the entire volume of the 3D scene. The compute and memory usages of these methods scale linearly with the volume of the scene. Faster methods like pseudospectral time domain (PSTD) [91] and adaptive rectangular decomposition (ARD) [124] have been proposed and achieve good accuracy with a much coarser spatial discretization. Volumetric techniques are well-suited for scenes with high surface area and low air volume, which makes them highly applicable to indoor spaces.

Surface-based techniques are better suited for open scenes, for which scattering geometry is sparse with large regions of air with uniform wave-propagation speed. The most common approach here is the boundary element method (BEM) [35] that expresses the global acoustic field as the sum of elementary radiating fields from monopole and dipole sources placed on a uniform, sub-wavelength sampling of the scene’s surface. Traditional BEM scales as the square of the surface area but recent research on the fast multipole method for BEM (FMM-BEM)[92, 54] has improved the complexity to linear in surface area by creating a hierarchical clustering of BEM monopoles and dipoles using an octree, and approximating their interactions compactly using high-order multipole Green’s functions. Offline FMM-BEM solutions are infeasible for interactive applications because of the large, dense number of monopole and dipole sources in the final solution that need to be stored and summed on the fly.

For acoustic radiation and scattering problems, an efficient and powerful surface-based technique is the equivalent source method (ESM) [44, 84, 106, 113] that forms the basis of our formulation. Instead of relying on a boundary-integral formulation, as BEM does, ESM exploits the uniqueness of solutions to the acoustic boundary value problem. Equivalent multipole sources, Green’s functions, are placed at variable locations in space with the intent of making the total generated field match boundary conditions on the object’s surface, since uniqueness guarantees the correctness of the solution (Section 3 in [104]). The flexibility of location results in fewer multipole sources. The ESM can yield large gains in performance and memory-efficiency for scattering and radiation problems in large spaces, and has been used widely in both acoustic and electromagnetic applications [40]. Equivalent sources were introduced to computer graphics in the seminal work of [71] on sound generation from vibrating objects. ESM is an attractive starting point for such precomputation-based approaches, and our method, because it allows very flexible performance-

to-accuracy tradeoffs. More importantly, the compactness of the solutions reduces runtime memory and compute requirements by a large factor, making them amenable to real-time evaluation.

A related technique, called the transition-matrix method, has been used extensively for electromagnetic scattering, and also developed for acoustics [170]. The method relates the incoming and outgoing fields in the scattering process in terms of the coefficients of a complete system of vector basis functions that are not necessarily Green’s functions, unlike BEM or ESM.

2.3.2 Interactive Geometric Techniques

Most current interactive sound propagation systems are based on geometric acoustics, which applies the high-frequency Eikonal (ray) approximation to sound propagation. The image source method [6] is the most commonly used geometric technique, and there has been much research on improving its performance [45]. However, the image source method can only model purely specular reflections. Other techniques based on ray tracing [83, 166, 88] or radiosity [159] have been developed for modeling diffuse reflections, but these energy-based formulations may not model phase accurately. Techniques based on acoustic radiance transfer [138, 139] can model arbitrary surface interactions for wide-band signals, but cannot accurately model wave phenomena such as diffraction. The two main approaches for modeling diffraction in a geometric acoustics framework are the uniform theory of diffraction (UTD) [157] and the Biot-Tolstoy-Medwin (BTM) formulation [143]. UTD is an approximate formulation, while the BTM yields accurate results with a significant performance cost. Methods based on image source gradients [155] and acoustic radiance transfer operators [9] have been developed to interactively model higher-order propagation effects. Recent developments in fast ray tracing have enabled interactive geometric propagation in dynamic scenes, but these methods only model first-order edge diffraction based on UTD [151].

2.3.3 Interactive Wave-simulation Techniques

In recent years, we have seen increasing interest in developing interactive wave-simulation techniques for sound propagation in indoor and outdoor spaces. Sound radiation from a single vibrating object in free

space can be efficiently modeled using precomputed acoustic transfer [71]. These acoustic transfer functions approximate the radiation behavior of a complicated geometry by expressing it in terms of equivalent sources, which can be quickly evaluated at runtime to enable real-time performance. Tsingos et al. [156] solves the boundary integral formulation of the Helmholtz equation subject to the Kirchhoff approximation in real-time. Raghuvanshi et al. [125] relies on a volumetric sampling of acoustic responses on a spatial grid and perceptual encoding based on the acoustic properties of indoor spaces. Recent work [135] has shown that FDTD simulations can run in real-time on the GPU, but only for very small spaces that span a few meters across. We compare our method in more detail with these closely related interactive wave-simulation techniques in later sections.

2.4 The Equivalent Source Method

In this section, we give a brief review of the Equivalent Source Method. Consider the exterior scattering problem [154], a solid three-dimensional object A immersed in an unbounded air volume (see Figure 2.2(a)). Considering only time-harmonic vibrations, with angular frequency ω and a homogeneous medium with constant speed of sound c , acoustic wave propagation can be expressed as a boundary value problem for the Helmholtz equation:

$$\nabla^2 p + \frac{w^2}{c^2} p = 0 \text{ in } A^+, \quad (2.1)$$

where p is the (complex-valued) pressure field, A^+ is the domain exterior to the object, and ∇^2 is the Laplacian operator. At the boundary of the domain, ∂A , the pressure is specified using a Dirichlet boundary condition:

$$p = f(\mathbf{x}) \text{ on } \partial A. \quad (2.2)$$

To complete the problem specification, the behavior of p at infinity must be specified, usually by the *Sommerfeld radiation condition* [116]:

$$\lim_{r \rightarrow \infty} \left[\frac{\partial p}{\partial r} + j \frac{w}{c} p \right] = 0, \quad (2.3)$$

where $r = \|\mathbf{x}\|$ is the distance of point $\mathbf{x}$ from the origin and $\hat{j} = \sqrt{-1}$. The equivalent source method [104, 106, 113] relies on the existence of fundamental solutions also called Green's functions or equivalent sources $q(\mathbf{x}, \mathbf{y})$, of the Helmholtz equation (2.1) subject to the Sommerfeld radiation condition (2.3) for all $\mathbf{x} \neq \mathbf{y}$. An equivalent source $q(\mathbf{x}, \mathbf{y}_i)$ is the solution field induced at any point $\mathbf{x}$ due to a point source located at $\mathbf{y}_i$, and can be expressed as the sum:

$$q(\mathbf{x}, \mathbf{y}_i) = \sum_{l=0}^{L-1} \sum_{m=-l}^l d_{ilm} \varphi_{ilm}(\mathbf{x}) = \sum_{k=1}^{L^2} d_{ik} \varphi_{ik}(\mathbf{x}), \quad (2.4)$$

where k is a generalized index for (l, m) . The fundamental solution $\varphi_{ilm}(\mathbf{x})$ is the field due to a *multipole* source located at $\mathbf{y}_i$, d_{ilm} is its strength, and L is the order of the multipole ($L = 1$ is just a monopole, $L = 2$ includes dipole terms as well, and so on). The field due to a multipole located at point $\mathbf{y}_i$ is defined as

$$\varphi_{ilm}(\mathbf{x}) = \Gamma_{lm} h_l^{(2)}(wr_i/c) \psi_{lm}(\theta_i, \phi_i) \quad (2.5)$$

where (r_i, θ_i, ϕ_i) is the vector $(\mathbf{x} - \mathbf{y}_i)$ expressed in spherical coordinates, $h_l^{(2)}(wr_i/c)$ are the (complex-valued) spherical Hankel functions of the second kind [2], $\psi_{lm}(\theta_i, \phi_i)$ are the (complex-valued) spherical harmonic functions [63], and Γ_{lm} is the (real-valued) normalizing factor that makes the spherical harmonics orthonormal.

The fundamental solutions $\varphi_{ilm}(\mathbf{x})$ (or $\varphi_{ik}(\mathbf{x})$) are used to solve the Helmholtz equation. Consider the outgoing scattered field due to an object, and the associated Dirichlet boundary value problem on ∂A . Consider a discrete set of R source locations $\{\mathbf{y}_i\}_{i=1}^R$, all contained in the interior region A^- . The total field due to these sources at any $\mathbf{x} \in A^+$ is

$$p(\mathbf{x}) = \sum_{i=1}^R c_i q(\mathbf{x}, \mathbf{y}_i) = \sum_{i=1}^R \sum_{k=1}^{L^2} c_{ik} \varphi_{ik}(\mathbf{x}), \quad (2.6)$$

where $c_{ik} = c_i d_{ik}$ are the corresponding strengths of the equivalent sources. The main idea of the ESM is that if the equivalent source strengths c_{ik} and positions $\mathbf{y}_i$ are chosen to match the Dirichlet boundary

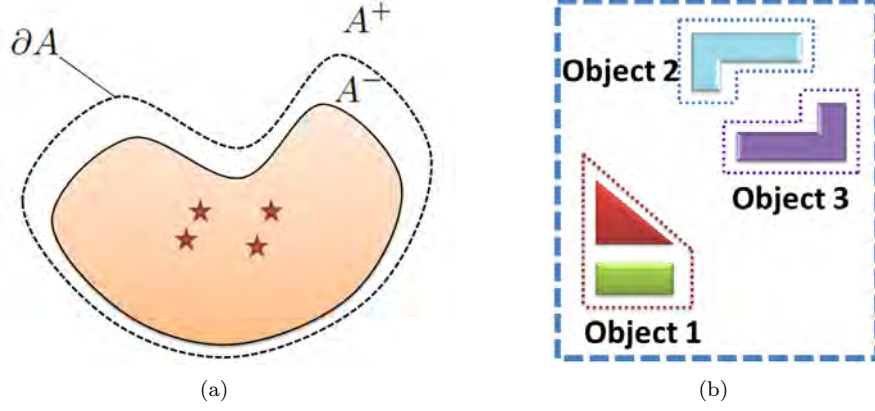


Figure 2.2: a) A diagram illustrating a radiating object A , its corresponding boundary ∂A , exterior region A^+ , interior region A^- , and the set of equivalent sources (denoted by star shapes). b) Classification of objects in a scene. The triangle and rectangle constitute a single object, as their offset surfaces overlap. On the other hand, L-shaped shapes are classified as separate objects.

condition on ∂A ,

$$p(\mathbf{x}) = \sum_{i=1}^R \sum_{k=1}^{L^2} c_{ik} \varphi_{ik}(\mathbf{x}) = f(\mathbf{x}); \quad \mathbf{x} \in \partial A, \quad (2.7)$$

then $p(\mathbf{x})$ is the correct solution over all A^+ .

This process can also be used to represent the incident field of an object, the only difference being that the equivalent sources are now placed in the exterior region A^+ . Again, by matching the boundary condition (2.7), we get the correct solution $p(\mathbf{x})$ for all $\mathbf{x}$ in the interior region A^- .

In practice, the boundary conditions (2.7) can only be satisfied approximately for a finite value of R , and the degree of approximation can be controlled by changing R . Since the strengths of multipoles of each source must be stored and its contribution evaluated at runtime, R is the main parameter for trading accuracy for runtime performance and memory requirements. This flexibility makes ESM highly suitable for interactive applications.

2.5 Sound propagation using ESM

We give a brief overview of the precomputation and runtime stages of our technique (see Figure 2.3). Our formulation is in the frequency domain. We construct a complex frequency response (containing magnitudes

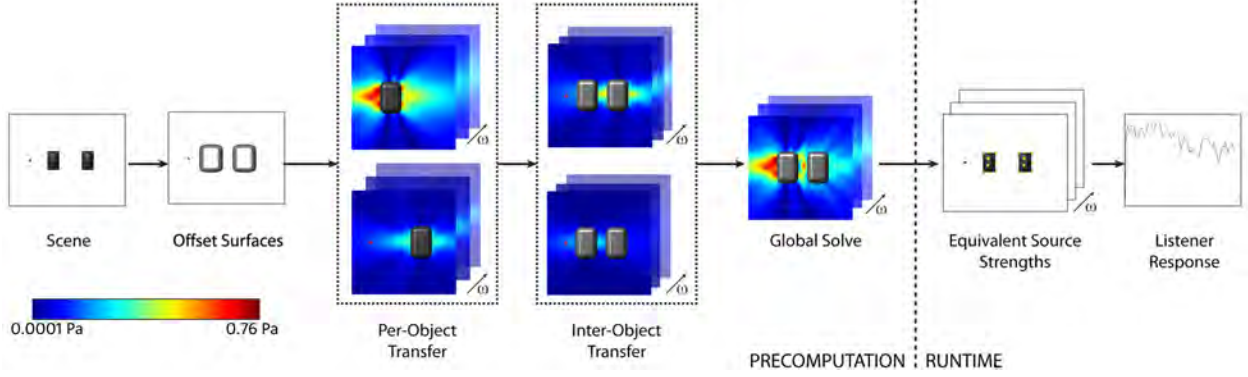


Figure 2.3: Overview of our wave-based sound propagation technique based on equivalent sources on a simple scene composed of two objects and a sound source (shown with a red dot). The magnitudes of pressure fields are visualized using the color scheme shown.

and phases), at regularly sampled frequencies, to model the delay information in the propagated sound. Thus, the steps outlined in this section, except the offset surface calculation, need to be performed for a regularly sampled set of frequencies in the range $[0, \nu_{\max}]$, where $\nu_{\max}$ is the maximum simulated frequency. We assume that the scene is composed of static objects. Table 2.1 provides a list of commonly used symbols.

2.5.1 Our Approach

Offset surface calculation: In the preprocessing stage, we classify objects in the scene and calculate the offset surface for each object.

Per-object transfer function: For each object, we compute a *per-object transfer function* that maps the incoming field incident on the object to the outgoing scattered field.

Inter-object transfer function: For each object pair, we precompute an *inter-object transfer function* that encodes how the outgoing scattered field of one object becomes the incoming field for the other object.

Global solve: Based on the per-object and inter-object transfer functions and a sound source configuration, we model acoustic interactions between the objects in the scene and solve for the global sound field. Thereby, we compute the strengths of all the outgoing scattered field equivalent sources of all objects.

Run-time pressure evaluation: At runtime, we add the pressure produced at the listener position by all outgoing field equivalent sources, for each frequency. This is an extremely fast computation, and can be

performed for a moving listener in real-time.

2.5.2 Offset Surface Calculation

The first step is to decompose the input scene into *well-separated* objects. To decide if two objects are well-separated, we use the notion of an *offset surface*. The offset surface is defined by taking the constant offset along the normal direction at each point on the boundary of the object. Two objects are considered disjoint if and only if their offset surfaces do not intersect. Otherwise, we combine them and treat them as a single object (see Figure 2.2(b)). We compute the offset surface of an object using distance field and the marching cubes algorithm similar to James et al. [71]. Typical values of distance field voxel resolution h and offset distance δ are specified in Table 2.2. The offset surface serves as the boundary of the domain ∂A . After decomposing the scene into well-separated objects, we compute the scattering properties for each object independently.

Symbols	Meaning
q_i^{in}, q_j^{out}	i^{th} & j^{th} eq. src for incoming, outgoing field resp.
$\varphi_{ik}^{in}, \varphi_{jh}^{out}$	k^{th} & h^{th} multipole term of eq. src. q_i^{in} & q_j^{out} resp.
Q, P	number of incoming, outgoing eq. srcs resp.
M, N	order of incoming, outgoing field multipoles resp.

Table 2.1: Table of commonly used symbols.

2.5.3 Per-object Transfer Function

In order to capture an object’s scattering behavior, we define the *per-object transfer function* f , a function which maps an arbitrary incoming field reaching the object to the corresponding outgoing scattered field after reflection, scattering and diffraction due to the object itself. This function is linear owing to the linearity of the wave equation and depends only on the shape and material properties of the object.

The incoming and outgoing fields for an object A are both expressed using equivalent sources. The outgoing field is represented by placing equivalent sources $\{q_1^{out}, q_2^{out}, q_3^{out}, \dots\}$ in the interior region A^- of the object. Similarly, the incoming field is represented by placing equivalent sources $\{q_1^{in}, q_2^{in}, q_3^{in}, \dots\}$ in the

exterior region A^+ . The transfer function f maps the basis of the incoming field (multipoles φ_{ik}^{in}) to the corresponding outgoing field expressed as a linear combination of its basis functions (multipoles φ_{jh}^{out}):

$$f(\varphi_{ik}^{in}) = \sum_{(j,h)=(1,1)}^{(P,N^2)} \alpha_{jh}^{ik} \varphi_{jh}^{out}; \quad (2.8)$$

$$\begin{aligned} \begin{bmatrix} f(\varphi_{11}^{in}) \\ f(\varphi_{12}^{in}) \\ \vdots \\ f(\varphi_{QM^2}^{in}) \end{bmatrix} &= \begin{bmatrix} \alpha_{11}^{11} & \alpha_{12}^{11} & \dots & \alpha_{PN^2}^{11} \\ \alpha_{11}^{12} & \alpha_{12}^{12} & \dots & \alpha_{PN^2}^{12} \\ \vdots & \vdots & \dots & \vdots \\ \alpha_{11}^{QM^2} & \alpha_{12}^{QM^2} & \dots & \alpha_{PN^2}^{QM^2} \end{bmatrix} \begin{bmatrix} \varphi_{11}^{out} \\ \varphi_{12}^{out} \\ \vdots \\ \varphi_{PN^2}^{out} \end{bmatrix} \\ &= T_A \Phi_A^{out}, \end{aligned} \quad (2.9)$$

where $\alpha_{jh}^{ik} \equiv T_A(ik, jh)$ is the (complex) amplitude for the outgoing multipole φ_{jh}^{out} induced by a unit-amplitude incoming multipole φ_{ik}^{in} . The per-object sound transfer function for object A is encoded in the coefficient matrix T_A , which we call the *scattering matrix*. We now explain how to compute the (complex) amplitudes α_{jh}^{ik} of the outgoing field multipoles. Details on choosing the number and positions of incoming and outgoing equivalent sources are given in Section 2.5.5.

Computing the Scattering Matrix For each incoming field multipole φ_{ik}^{in} in turn, we place a unit-amplitude sound source and use a numerical wave solver to compute the total pressure field at n uniformly-sampled locations $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ on ∂A . We subtract the incident field from the total pressure field to compute the outgoing scattered field at these sampled locations (see Figure 2.4), denoted by $\bar{p}_{ik} = \{p(\mathbf{x}_1), p(\mathbf{x}_2), \dots, p(\mathbf{x}_n)\}$.

We fit the outgoing field multipole expansion to the sampled scattered field, in a least-squares sense, by solving an over-determined linear system ($n > PN^2$) subject to a pre-specified error threshold σ for all

incoming field multipoles:

$$\sum_{(j,h)=(1,1)}^{(P,N^2)} \varphi_{jh}^{out}(\mathbf{x}_t) \alpha_{jh}^{ik} = p(\mathbf{x}_t), \text{ for } t = 1, \dots, n; \quad (2.10)$$

$$\mathbf{V} \boldsymbol{\alpha}^{ik} = \bar{\mathbf{p}}_{ik}. \quad (2.11)$$

The least-squares solution yields the coefficients $\boldsymbol{\alpha}^{ik}$ corresponding to the ik^{th} row of the scattering matrix T . This process is repeated for all incoming field multipoles to compute the scattering matrix. The solution can be computed efficiently using a single combined linear system

$$\mathbf{V} T_A^{tr} = \begin{bmatrix} \bar{\mathbf{p}}_{11} & \dots & \bar{\mathbf{p}}_{QM^2} \end{bmatrix}, \quad (2.12)$$

where T_A^{tr} is the transpose of T_A . The per-object transfer function is computed for all objects at sampled frequencies. The error threshold σ is used while deciding the number and placement of equivalent sources (Section 2.5.5) such that the above linear system gives error less than σ .

2.5.4 Inter-object Transfer Function

Scenes with multiple objects exhibit object-to-object interactions, where the outgoing field from one object serves as the incoming field for the other objects. For example, with two objects A and B , source s and listener l , the possible interactions that can occur from s to l are: direct sound (0^{th} order) $s \rightarrow l$, 1^{st} order $s \rightarrow A \rightarrow l$; $s \rightarrow B \rightarrow l$, 2^{nd} order $s \rightarrow A \rightarrow B \rightarrow l$; $s \rightarrow B \rightarrow A \rightarrow l$, and so on. We model these interactions by formulating an *inter-object transfer function*. For two objects A and B , the inter-object transfer function g_A^B expresses the outgoing field of A in terms of the basis of the incoming field of B . Like the per-object transfer function, the inter-object transfer function is also a linear function. The inter-object transfer function g_A^B maps each basis function of the outgoing field of A (multipoles φ_{jh}^{out}) to the corresponding incoming field

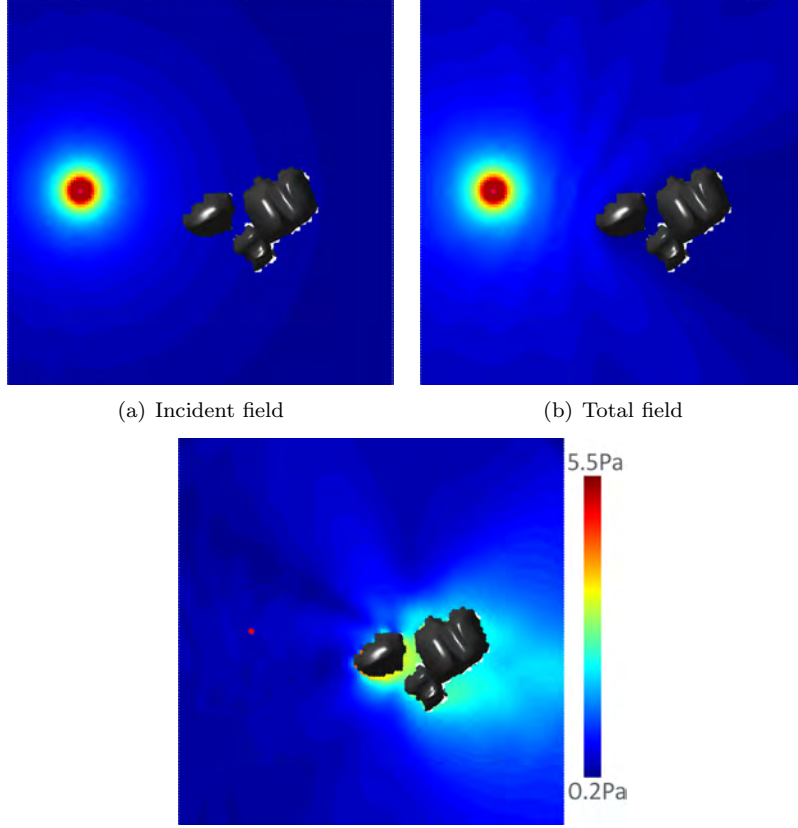


Figure 2.4: Magnitude of the pressure field (in Pa) at 170 Hz in a simple scene with a single object (rocks) and a single sound source (red dot). The difference between total and incident fields is the scattered field (scaled eight times for visualization). Note the high amplitude of the scattered field between the rocks representing the large difference in incident and total field that results from diffracted occlusion.

of B expressed as a linear combination of its basis functions (multipoles φ_{ik}^{in}):

$$g_A^B(\varphi_{jh}^{out}) = \sum_{(i,k)=(1,1)}^{(Q,M^2)} \beta_{ik}^{jh} \varphi_{ik}^{in}; \quad (2.13)$$

$$\begin{aligned}
\begin{bmatrix} g_A^B(\varphi_{11}^{out}) \\ g_A^B(\varphi_{12}^{out}) \\ \cdot \\ \cdot \\ g_A^B(\varphi_{PN^2}^{out}) \end{bmatrix} &= \begin{bmatrix} \beta_{11}^{11} & \beta_{12}^{11} & \dots & \beta_{QM^2}^{11} \\ \beta_{11}^{12} & \beta_{12}^{12} & \dots & \beta_{QM^2}^{12} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \beta_{11}^{PN^2} & \beta_{12}^{PN^2} & \dots & \beta_{QM^2}^{PN^2} \end{bmatrix} \begin{bmatrix} \varphi_{11}^{in} \\ \varphi_{12}^{in} \\ \cdot \\ \cdot \\ \varphi_{QM^2}^{in} \end{bmatrix} \\
&= G_A^B \Phi_B^{in},
\end{aligned} \tag{2.14}$$

where $\beta_{ik}^{jh} \equiv G_A^B(jh, ik)$ is the (complex) amplitude of the incoming multipole φ_{ik}^{in} of B induced by a unit-amplitude outgoing multipole φ_{jh}^{out} of A . The inter-object transfer function from A to B is thus encoded as G_A^B , which we call the *interaction matrix*. Generally, the interaction matrix is *not* symmetric, i.e., $G_A^B \neq G_B^A$. Since the outgoing field of an object is not defined in its interior region, G_A^A and G_B^B are zero matrices. We now explain how to compute the (complex) amplitudes β_{ik}^{jh} of the incoming field multipoles.

Computing the Interaction Matrix The interaction matrix G_A^B can be computed using a least-squares formulation similar to the one used for computing scattering matrices. However, the pressure values at the offset surface samples of B , $\bar{p}_{jh} = \{p(\mathbf{x}_1), p(\mathbf{x}_2), \dots, p(\mathbf{x}_n)\}$ are simpler to compute. In a homogenous medium, the outgoing field due to a multipole is the same as the free space field, for which analytical expressions exist (Equation 2.5). Therefore, we simply evaluate the analytical expressions of the outgoing field multipoles φ_{jh}^{out} of A at the sample points on the offset surface of B . The resulting linear system is solved subject to a separate error threshold, η :

$$\sum_{(i,k)=(1,1)}^{(Q,M^2)} \varphi_{ik}^{in}(\mathbf{x}_t) \beta_{ik}^{jh} = p(\mathbf{x}_t), \text{ for } t = 1, \dots, n. \tag{2.15}$$

Again, this process is repeated for each outgoing multipole of B , and solved efficiently as a single combined linear system:

$$U G_A^{Btr} = \begin{bmatrix} \bar{p}_{11} & \dots & \bar{p}_{PN^2} \end{bmatrix}. \tag{2.16}$$

The inter-object transfer functions are computed for all object pairs, independently for each frequency.

2.5.5 Computing Equivalent Source Positions

Choosing Offset Surface Samples Solving equations (2.12) and (2.16) at frequency ν involves computing the pressure at sampled locations $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ on the offset surface of each object. The number of sampled locations n depends on the spatial variation of the pressure field, which in turn, depends directly on its frequency ν or inversely on its wavelength λ since $\nu = c/\lambda$. As per the Nyquist Theorem, representing a signal of frequency ν with a finite number of samples requires a sampling rate of 2ν . The spatially-varying pressure field defined on the 2D offset surface must be sampled at a rate of 2ν in both dimensions. The distance between samples become $2\nu/c = 2/\lambda$. Therefore, we place $n \propto (2\nu/c)^2 \times \text{surface area} = (2/\lambda)^2 \times \text{surface area}$ samples uniformly on the offset surface.

Choosing Incoming Equivalent Sources Since the nature of the incoming field is not known *a priori*, it is difficult to optimize the number and position of incoming equivalent sources. We resolve this problem by generating another offset surface at distance $\Delta > \delta$ from the object, where δ is the original offset surface's distance, and placing incoming equivalent sources on this new surface (see Table 2.2 for the value of Δ). The number of incoming equivalent sources Q depends on the spatial variation of the incoming pressure field. As before, $Q \propto (2/\lambda)^2 \times \text{surface area}$ equivalent sources are uniformly placed. This allows us to represent the incoming field on the inner offset surface to good accuracy.

Choosing Outgoing Equivalent Sources The number of outgoing equivalent sources P and their positions are decided based on a multi-level source placement algorithm similar to James et al. [71]. The previous algorithm was designed to satisfy a single radiating field $\bar{\mathbf{p}}$ of an object at each frequency. It places equivalent sources in a greedy manner, where at each step a set of candidate positions χ are ranked based on their ability to reduce the pressure residual vector $\bar{\mathbf{r}} = \bar{\mathbf{p}}/\|\bar{\mathbf{p}}\|_2$ on the offset surface. The best candidate position $\mathbf{x}^*$ is chosen via the largest projection, i.e., $\mathbf{x}^* = \arg \max_{x \in \chi} u$, where projection $u = \|(U_{\mathbf{x}})^H \bar{\mathbf{r}}\|_2$. The unitary matrix corresponding to the subspace spanning all the previously selected positions is updated.

The residual vector is updated by removing its component in that subspace. The process is repeated until the value of the residual $\|\bar{\mathbf{r}}\|_2$ falls below the error tolerance. The set of best candidate positions selected in the process is the set of outgoing equivalent sources and its size gives us the value of P .

Our algorithm is designed to satisfy multiple outgoing radiating fields at each frequency simultaneously. In our case, at each frequency, we have as many outgoing radiating fields $\begin{bmatrix} \bar{\mathbf{p}}_{11} \dots \bar{\mathbf{p}}_{QM^2} \end{bmatrix}$ as the number of incoming multipoles QM^2 . This gives us a vector of pressure residual vectors $\mathbf{r} = [\bar{\mathbf{r}}_{11} \dots \bar{\mathbf{r}}_{QM^2}]$ and a corresponding vector of projections $\mathbf{u} = [u_{11} \dots u_{QM^2}]$ where $u_{ik} = \|(U_{\mathbf{x}})^H \bar{\mathbf{r}}_{ik}\|_2$. We choose the best candidate as the one that minimizes the pressure residual of all outgoing fields simultaneously via a modified largest projection $\mathbf{x}^* = \arg \max_{x \in \chi} \|\mathbf{u}\|_2$. We update the unitary matrix and for each residual vector we remove its component in the chosen subspace. We then compute the value of the modified residual $\|\mathbf{d}\|_2$, where $\mathbf{d} = [d_{11} \dots d_{QM^2}]$ and $d_{ik} = \|\bar{\mathbf{r}}_{ik}\|_2$. We repeat this process until the relative value of the modified residual falls below the error tolerance (σ in our case). Similar to the number of incoming equivalent sources Q , the number of outgoing equivalent sources P also increases with frequency. But it strongly depends on the shape of the object and the complexity of the outgoing scattered field that the object generates. We fit as many equivalent sources as necessary to satisfy the error threshold. As the frequency increases, more equivalent sources are needed but the accuracy of our technique is maintained. The candidate positions χ are chosen randomly on the surface of the object in the same manner as the previous algorithm. However, a minimum distance between any two equivalent sources is enforced to improve the condition number of the system; extremely close equivalent sources dominate the eigenvalues of the resulting system, adversely affecting its condition number. We choose a minimum distance of half the wavelength at any given frequency.

2.5.6 Global Solve

Once the scattering and interaction matrices are computed, and the sound source position has been decided, we solve for the global sound field and compute the outgoing equivalent source strengths of all the objects in the scene. The sound source can be a point source or a complex directional source (represented as a set of multipoles). We give an intuitive explanation here for a simple two-object scene and the detailed derivation

can be found in Appendix 2.9.1. For a scene composed of multiple objects, we derive the same equation with the symbols having analogous meanings, as described in detail in Appendix 2.9.2.

Assume the outgoing field in the scene is $\mathbf{C}$. This field when propagated through the scene, transferred via all possible object pairs using interaction matrix $\mathbf{G}$, generates an incoming field $\mathbf{GC}$ that, in addition to the source field $\mathbf{S}$, generates the total incoming field $(\mathbf{GC} + \mathbf{S})$ on the objects. This incoming field is then scattered by the object, via scattering matrix $\mathbf{T}$, to produce an outgoing field $\mathbf{T}(\mathbf{GC} + \mathbf{S})$. Under steady state, this outgoing field must equal $\mathbf{C}$. Mathematically, this can be written as

$$\mathbf{C} = \mathbf{T}(\mathbf{GC} + \mathbf{S}) \quad (2.17)$$

This yields a linear system for the outgoing source strengths for all objects:

$$(\mathbf{I} - \mathbf{TG})\mathbf{C} = \mathbf{TS}. \quad (2.18)$$

This linear system is solved for $\mathbf{C}$ at a regularly-sampled set of frequencies. This step has to be repeated for every sound source generating a distinct source field $\mathbf{S}$. In the absence of a source, the solution is identically zero.

2.5.7 Runtime Computation

At the end of the preprocessing stage, we obtain the outgoing equivalent source strengths for all objects at a regularly sampled set of frequencies corresponding to each sound source. During run-time, we use these strengths to compute the pressure field at any listener position $\mathbf{x}$:

$$p(\mathbf{x}) = \sum_{j=1}^{\kappa} C_{A_j}^{tr} \Phi_{A_j}^{out}(\mathbf{x}) + s(\mathbf{x}), \quad (2.19)$$

where κ is the number of objects in the scene, $C_{A_j}^{tr}$ and $\Phi_{A_j}^{out}$ are the strengths and multipoles of the outgoing equivalent sources for object A_j respectively, and $s(\mathbf{x})$ is the field generated by the sound source. This

computation is performed at a regularly-sampled set of frequencies and repeated for each source to compute a band-limited frequency response per source. Evaluating equation 2.19 for a new value of $\mathbf{x}$ is very efficient, allowing a moving listener to be handled naturally in real-time. Since the analytical expressions for multipoles of equivalent sources are used, the pressure can be evaluated at any position $\mathbf{x}$ in space and not necessarily at grid positions. Therefore, no spatial interpolation is required with our technique. Unlike grid-based approaches (such as FDTD), our equivalent source method is independent of the spatial discretization, resulting in a much smoother auralization for a moving listener.

Our technique allows auralization in a scene with multiple static sources and a moving listener. We can also handle the case of multiple moving sources and a static listener. First, we start with a scene with a static source and compute acoustic responses at multiple moving listeners using our runtime system. The principle of acoustic reciprocity states that we can reverse the sense of source and listener without changing the acoustic response [116, p. 195-199]. Using this principle, we now switch the roles of source and listeners while keeping the acoustic responses the same. This gives us acoustic response for the case of multiple moving sources with a static listener.

2.6 Implementation

In this section, we describe the implementation details of our technique. Typical parameter values used in our experiments are specified in Table 2.2.

Implementation Details The offset surface generation code is written in C++. When computing per-object transfer functions, outgoing scattered fields are computed on the offset surface (see Section 2.5.3) using an efficient GPU-based implementation of the ARD wave-solver [124, 100]. The solver treats the scattering objects as rigid (with no transmission) and handles the material properties using perfectly matched layer interfaces. The remaining parts of the preprocessing stage, solving the linear system for per-object transfer functions, inter-object transfer functions, and equivalent source strengths, are implemented in MATLAB. The runtime code is implemented in C++, and has also been integrated with Valve’s SourceTM engine, as

demonstrated in the supplementary video.

The timing results for offset surface generation, the ARD solver, and runtime code are measured on a single core of a 4-core 2.80 GHz Xeon X5560 desktop with 4 GB of RAM and NVIDIA GeForce GTX 480 GPU with 1.5 GB memory. Offset surface generation code takes < 1 sec for each object. The timing results for the MATLAB-based precomputation are measured on a 64-node CPU cluster (Xeon X5560 processor nodes, 8 cores, 2.80 GHz, 48 GB). Detailed statistics are provided in Table 2.3. Precomputation for each frequency is performed in parallel over all the nodes (and individual cores) of the CPU cluster. Given more nodes on the cluster, the per-object, inter-object, and source-field computations can be further parallelized over all unique objects, all object-pairs, and all objects, respectively.

Due to the computational overhead of the precomputation stage, we treat band-limited sources that emit sound whose frequency range is bounded by maximum frequency $\nu_{\max}$ (see Table 2.2), for the purpose of wave simulations (see Section 2.5.3). The pressure is computed at regularly sampled set of frequencies in the range $[0, \nu_{\max}]$ with a step size of $\Delta\nu$. The value of parameter $\Delta\nu$ is 4.08 Hz for concave, wall, rock, and parallel walls scenes and 2.04 Hz for desert, reservoir, and Christmas scenes.

Handling Ground Reflections To handle ground reflections, we assume the ground to be an infinite plane. Similar to the image source method [6], we reflect our equivalent sources about the ground plane and multiply their source strengths by the (complex) reflection coefficient of the ground. Since sound waves traveling in air maintain their phase upon reflection from a hard surface, we do not need to invert the strengths of the equivalent sources. To incorporate last ground reflection, this step is performed after the “Global solve” step (see Section 2.5.6). In order to handle all levels of ground reflections before that, this step needs to be performed while computing the interaction matrices as well (see Section 2.5.4). More accurate physical models of ground reflection coefficient based on Darcy’s law can also be used [147]. The assumption of infinite flat plane works very well for cases where the size of the ground perturbations is smaller than the minimum wavelength simulated (34cms for 1kHz). For cases where the ground contains terrain features that are much larger, like hillocks, these can be handled as separate objects in our ESM framework. Due

to the increased number of objects, the precomputation time and runtime memory would increase, but the accuracy of our technique would be maintained.

Parameter	Value	Description
c	340 m/s	speed of sound
$\nu_{\max}$	1 kHz	highest frequency simulated
h	$c/2\nu_{\max} = 0.17$ m	voxel resolution of distance field
δ	$5h = 0.85$ m	inner offset distance
Δ	$8h = 1.36$ m	outer offset distance
σ	15%	error threshold for scattering matrix
η	1%	error threshold for interaction matrix
M, N	2	order of incoming, outgoing multipoles resp.

Table 2.2: Parameters used in our system.

Spectral Extrapolation The band-limited nature of the frequency responses of our technique necessitates a plausible extrapolation to higher frequencies at runtime. Prior work on interactive wave-based methods has shown that spectral extrapolation techniques can be used to produce plausible results for higher frequencies [125]. However, using this method with our technique would incur an extra inverse FFT cost at every audio frame for time-domain processing. Therefore, we implemented a simple, fast extrapolation technique based on the edge-diffraction spectra [143]. As observed, the typical edge diffraction spectra are roughly linear on a log-log scale. Hence, we first estimate a trend-line by a least-squares fit to the maximas of the log magnitude spectrum till $\nu_{\max}$. We then adjust for the trend, to create a flat response, by multiplying with the inverse of the trend on a log frequency scale. This adjusted response is replicated to higher frequencies and then multiplied by the trend again for the entire frequency range, yielding the final wide-band spectrum. If the trend-line has positive slope, indicating a high-pass response, we flatten the trend-line for frequencies beyond $\nu_{\max}$. This extrapolation technique does not change the spectrum up to $\nu_{\max}$.

We evaluate our spectral extrapolation technique by comparing the audio quality of the results with the wide-band spectrum produced by the Biot-Tolstoy-Medwin (BTM) technique (0-22 kHz) as the ground truth for the single, finite-edge scenario created in the right-angled wall scene. In the BTM method, edge diffraction impulse responses are computed by evaluating a time-domain line integral over the finite length of the edge. This is essentially based on Huygens theory, where a diffracting sound wave is modeled as a superposition of an infinite number of secondary point sources situated along the diffracting edge, each with different

strengths and directivities. BTM has been shown to converge to the exact analytical solution for a simple scene like this [143]. We use the MATLAB-based edge diffraction toolbox (<http://www.iet.ntnu.no/svensson/software/index.html>) to generate the BTM results. As shown in the supplementary video, the final auralized audio generated by both the techniques on this scene sound similar.

This spectral extrapolation technique is approximate, and becomes exact in a specific, single-edge diffraction configuration. It does not guarantee accuracy on general scenes at high frequencies. While single-edge diffraction arises frequently in outdoor scenes, many other complex configurations also occur, such as double-diffraction and diffracted-reflection. Our extrapolation approach would be accurate in such cases only if the acoustic response is dominated by diffraction from a single edge. In other situations, we have observed that our extrapolation approach generates plausible results. A general spectral extrapolation approach for band-limited acoustic responses with guarantees on extrapolation error for arbitrary scenes, is an important area for future research.

Real-Time Auralization The sound sources used in our implementation play a pre-recorded audio clip. Audio is rendered using FMOD, and is processed in frames of 1024 audio samples, at a sampling rate of 44.1 kHz. In-game (“dry”) audio clips are pre-processed by computing a windowed short-time Fourier transform (STFT) on each frame (Blackman window). The STFTs are computed on audio frames after zero-padding by the maximum impulse response length to prevent aliasing artifacts. Real-time auralization is performed using overlap-add STFT convolutions. In each rendered frame, the dry audio frame for each source is multiplied in the frequency-domain with the corresponding frequency response. The results are then mixed, and an inverse FFT performed on the mixed audio. Finally, *overlap* from previous frames is added in, and overlap from the current frame is cached in a ring buffer. Frequency responses are updated asynchronously from the actual convolution processing. Spatialization is achieved by using a simplified spherical head model with two listeners, one for each ear. Richer spatialization that uses geometry information of an individual listener’s ears, head, and shoulders can be modeled using *head related transfer functions* (HRTFs), and can be easily integrated in our approach, but is computationally more expensive.

Scene	#objs.	# freq.	# srcs	wave sim. (total, per obj.)	per-object (per freq)	inter-object (per freq)	source field (per freq, per src)	global solve (per freq, per src)	wall clk time	# eq. srcs (total, per src)	eval. (total, per src)	storage (total, fixed + per src)
Concave	1	250	1	80 min	51 min	NA	1 min	0.1 min	132 min	0.1 M	3 ms	(1 + 4) MB
Wall	1	250	1	50 min	101 min	NA	3 min	0.1 min	154 min	0.1 M	4 ms	(2 + 5) MB
Rock	1	250	1	80 min	87 min	NA	1 min	0.1 min	168 min	0.4 M	10 ms	(4 + 11) MB
Parallel	2*	250	1	50 min	101 min	13 min	6 min	1 min	171 min	0.2 M	8 ms	(4 + 10) MB
Desert	4*+2*	500	3	180 min	196 min	98 min	9 min	26 min	509 min	1.1 M	26 ms	(12 + 33) MB
Reservoir	4*+1	500	2	146 min	224 min	63 min	7 min	15 min	455 min	1.3 M	33 ms	(15 + 41) MB
Christmas	2*+2*+1	500	2	297 min	301 min	71 min	7 min	18 min	694 min	1.5 M	38 ms	(18 + 47) MB

Table 2.3: **Performance statistics.** Abbreviations are as follows : “#objs.” denotes the number of objects in the scene, “#freq.” is the number of frequency samples in the range [0-1 kHz] and “#srcs” is the number of sound sources. For the precomputation stage, the term “wave sim.” is the total simulation time of the numerical wave solver for all frequencies, “per-object” denote the compute time for the per-object transfer function for all unique objects and “inter-object” denote the compute time for inter-object transfer functions for all object pairs, “source-field” is time to express each sound source in terms of incoming multipoles for all objects, and “global-solve” is time to compute equivalent source strengths for all objects. The “wave sim.” step is parallelized over all unique objects whereas the remaining precomputation steps are parallelized over all frequencies. The term “wall-clk time” is the total wall-clock time computed by uniformly distributing all the parallel processes over all the cores of the 64-node cluster with 8 cores per node (512 cores in total). At runtime, the total number of equivalent sources “# eq. srcs” (in million M), performance “eval.” and storage requirement “storage” (fixed and per source cost) for all objects for all frequencies are also specified. For column “#objs.”, notation $a^* + b^*$ denotes that first object has been instanced a times and second object instanced b times, but their per-object transfer functions are computed only once for each unique object.

Air Absorption High-frequency sounds are absorbed more aggressively by the atmosphere than low frequencies. This frequency dependent air absorption is currently not modeled by our technique. However, it can be included as a post-processing step to our auralization pipeline. Since we compute complex frequency responses containing phase information, propagation delays are modeled. These delays yield the propagation distances which can be used to calculate and apply a per-frequency attenuation filter in frequency domain to model atmospheric absorption.

2.7 Results and Analysis

In this section, we present the results of our technique on different scenarios, provide error analysis and compare it with prior work.

2.7.1 Scenarios

We have considered a variety of scenes for testing our technique. For auralizations corresponding to the scenes discussed below, refer to the supplementary video.

Single Object We considered various objects having different scattering characteristics: rocks, a wall, and a concave reflector. The rocks scatter sound in all directions (see Figure 2.4). We show magnitude of the scattered sound field for the wall generated by our technique and by BEM in Figure 2.5. As shown in the figure, the wall strongly scatters sound in the direction perpendicular to itself. As a more challenging scene, the magnitude of scattered sound field for a concave reflector is also shown. The reflector generates significant interference effects, resulting in caustic formation in the focal region. This is clearly captured by our technique, as the high amplitude (red) region in the figure, showing that our technique is able to approximate the phase of the scattered wave field with reasonable accuracy. The relative error, defined in Figure 2.6 caption, between the total pressure fields generated by our technique and by the BEM technique is less than 2% for the wall and 5% for the concave reflector.

Parallel Buildings This scene consists of two buildings situated parallel to one another. We show two walkthroughs of this scene, with a flying helicopter, and a person speaking, respectively. As the helicopter moves behind a building, diffraction leads to a distinct low-pass occlusion effect. The two walls trap sound between them, producing high-order reflections, so that the volume of someone talking between the buildings is markedly louder than someone standing even slightly to the side.

Desert This is a large scene with three sound sources spread throughout the scene: a jeep, a bird, and a radio. As the listener walks through the scene, the sound received from the various sources changes significantly depending on whether or not the listener is in the line-of-sight of the source(s). We also specifically demonstrate the effect of second-order diffracted occlusion of the jeep sound around two buildings.

Christmas Town This scene demonstrates sound propagation in a village with many houses, a church, a bell tower and large buildings. It shows reflection from buildings, diffraction around houses, sound propagation over large distances from the bell tower, and reflections between two parallel buildings, for multiple sources.

Reservoir We show that our technique can be integrated with an existing game (*Half-Life 2*) to generate realistic wave acoustic effects in a large, outdoor game map. Our method is the first wave-based sound propagation technique that can accurately model wave phenomena such as diffraction behind the rocks and scattering around buildings over large distances on such a scene in real-time.

In our single-object examples, helicopter behind rock, the occluder is placed in isolation without any surrounding objects. Due to the lack of reflections from surrounding objects, and the fact that high frequencies do exhibit quite sharp shadows, our diffraction and occlusion effects may sound exaggerated, compared to real life.

2.7.2 Error Analysis

Figure 2.6 shows the convergence of our method as the error threshold σ decreases. Since the number of outgoing equivalent sources is inversely proportional to σ , it also shows convergence of the technique with increasing number of outgoing equivalent sources. We also plot the variation in the number of outgoing equivalent sources with frequency to achieve given error thresholds (see Figure 2.7). In Figure 2.8, we compare the results of our ESM technique with the reference wave solver ARD [124], BEM [35], and FMM-BEM [92, 54] techniques on a spatial grid of listeners at different frequencies for the two parallel walls scene. We used the state-of-the-art FastBEM simulator (<http://www.fastbem.com/>) for generating BEM and FMM-BEM results up to the maximum frequency possible (358 Hz). This scene is acoustically very complex, even though individual objects seem simple, since there are multiple orders of interaction happening between these two walls. Our approach handles such effects accurately and produces good approximations, up to the user-specified error thresholds, while reducing memory usage by orders of magnitude (see Table 2.4).

2.7.3 Computational Complexity

Consider a scene with κ objects. To perform analysis for frequency ν , let the number of offset surface samples, incoming equivalent sources and outgoing equivalent sources at this frequency be n , Q and P respectively. We assume that all objects have equal volume u .

Pre-processing

Scattering Matrix: For each of the QM^2 incoming multipoles of an object, wave simulations are performed and a dense linear system of size $n \times PN^2$ is solved to find the object's scattering matrix. The cost for each simulation is $u \log u$, and the cost of solving the linear system¹ is nP^2N^4 . Hence, the total cost is $O(\kappa QM^2(nP^2N^4 + u \log u))$.

Interaction Matrix: For every pair of objects, PN^2 linear systems of size $n \times QM^2$ need to be solved to find the interaction matrix. In total, we have κ^2 object pairs. The cost of evaluating analytical expressions for multipole pressure is $O(1)$ each, and is dominated by the cost of solving the linear systems. Hence the total cost is $O(\kappa^2 PN^2 n Q^2 M^4)$.

The size of these linear systems vary linearly with n , which in turn varies quadratically with frequency (see Section 2.5.5). Thus, ensuring a few hours precomputation time on a small computational cluster (see Table 3) limits our technique to 1-2 kHz on typical outdoor scenes.

Computing Strengths: The incoming field produced by each sound source is represented in terms of the incoming equivalent sources of the objects. This requires solving κ linear system of size $n \times QM^2$ resulting in cost $O(\kappa n Q^2 M^4)$. The size of the final linear system for finding outgoing equivalent source strengths for all objects in response to a sound source is $\kappa PN^2 \times \kappa PN^2$. Solving it takes $O(\kappa^3 P^3 N^6)$ time.

It follows that the total pre-processing cost at frequency ν thus scales as

$$O(\kappa QM^2(nP^2N^4 + u \log u + \kappa PN^2 n QM^2 + n QM^2) + \kappa^3 P^3 N^6)$$

Runtime At runtime, we evaluate equation (2.19), which takes $O(\kappa PN^2)$ at frequency ν . The runtime memory requirement consists of positions (3 floats) and (complex-valued) strengths (2 floats) of equivalent sources, which comes out to be $\kappa(3P + 2PN^2)$ at frequency ν .

The precomputation and runtime complexity and memory requirement depend on the number of equivalent sources P , which scales quadratically with frequency, in an asymptotic sense. However, for practical

¹To solve a dense linear system of size $m \times n$ ($m > n$), the cost is mn^2

objects and frequencies up to 1 kHz, we observed linear scaling of equivalent sources with frequency, as shown in Figure 2.7.

We have to compute the pressure at the listener position over a regularly sampled set of frequencies in the range $[0, \nu_{max}]$ with a step of $\Delta\nu$. The total number of frequency samples becomes $\nu_{max}/\Delta\nu$. Thus, the above expressions are summed for all frequency samples in this range to give the total computational complexity and memory requirement. Since the computational cost and runtime memory scales with the multipole order, we limit equivalent sources to monopoles and dipoles, i.e., $N=M=2$. Low multipole orders (N, M) result in a larger number of equivalent sources for satisfying the same error thresholds. However, since we place only as many equivalent sources as required, low multipole order does not effect the quality of the final result. The theoretical runtime memory requirements for other wave-solvers are discussed in Appendix 2.9.3. We also compare the runtime memory requirements of these solvers with our technique on a variety of scenes (see Table 2.4).

Scene	air. vol.	surf. area	FDTD	ARD	BEM/ FMM	Ours
Concave	$(85m)^3$	$107\ m^2$	33 TB	0.9 TB	0.5 GB	5 MB
Wall	$(85m)^3$	$71\ m^2$	33 TB	0.9 TB	0.3 GB	7 MB
Rock	$(85m)^3$	$159\ m^2$	33 TB	0.9 TB	0.8 GB	15 MB
Parallel	$(85m)^3$	$142\ m^2$	33 TB	0.9 TB	0.7 GB	14 MB
Desert	$(180m)^3$	$1626\ m^2$	625 TB	17 TB	15 GB	45 MB
Reservoir	$(180m)^3$	$950\ m^2$	625 TB	17 TB	9 GB	56 MB
Christmas	$(180m)^3$	$2953\ m^2$	625 TB	17 TB	27 GB	65 MB

Table 2.4: Runtime memory requirements per source, for FDTD [145], ARD [124], BEM/FMM-BEM [92], and our ESM technique with error thresholds $\sigma = 15\%$, $\eta = 1\%$ at maximum simulation frequency $\nu_{max} = 1018Hz$. Refer to Section 2.7.3 and Appendix 2.9.3 for more details.

Our precomputation step is computationally heavy and takes a few hours to run on a CPU-cluster (Table 2.3). But this step is trivially parallel and could be performed easily on cheap and widely available cloud computing resources, such as Amazon EC2. Moreover, our current implementation is in MATLAB and we expect 10x improvement with an optimized C++ implementation.

2.7.4 Comparison with Prior Interactive Techniques

Our usage of equivalent sources for sound propagation is in a similar vein to prior work [71], where the authors represent arbitrary outgoing radiation field from a single, geometrically complex object. Our work

differs primarily in three regards: First, we model *mutual interactions* between objects in arbitrary scenes using inter-object transfer functions, accounting for high-order interactions, such as echoes and multiple diffraction. Secondly, we model acoustic scattering from objects (as opposed to radiation), which requires an approximation of both the *incoming and outgoing* pressure fields for an object. Finally, our outgoing equivalent sources are chosen to satisfy multiple outgoing scattered fields as opposed to a single radiation field.

The problem of real-time acoustic scattering has been previously addressed using GPUs [156]. First-order scattering effects are incorporated, but acoustic interactions between objects are not modeled. In contrast, our work can handle all orders of interactions between the objects using inter-object transfer functions.

A recent technique for interactive acoustics based on wave simulation was proposed in Raghuvanshi et al. [125], which relies on sampling the volume of the scene, and uses a perceptual compression specific to indoor scenes. The runtime memory requirement of their technique (per source) on our scenes (assuming a spatial sampling of $1m$) is 187 MB for the parallel walls and 1.8 GB for the reservoir scene. This technique is complimentary to our approach; it works best in indoor spaces with a lot of geometric clutter but limited volume, while our technique is better suited to outdoor spaces with well-separated objects. In fact, it would be quite natural to integrate this method with ours, with the indoor and outdoor propagation models coupled through transport operators defined on doors and windows.

2.8 Conclusions and Discussion

We have presented a novel wave-based sound propagation algorithm that captures acoustic effects such as high-order diffraction and scattering, using an equivalent source formulation. As a result, our technique can perform accurate sound propagation on large, open scenes in real-time, has a small memory footprint, and allows flexible efficiency-to-accuracy tradeoffs. Compared to directly storing and convolving wave-solver solutions for auralization, we reduce the memory usage more than 100 times.

Our approach is currently limited to static scenes, due to the computational cost of recomputing inter-

object transfers as objects move. We would like to combine our approach with Fast Multipole Method (FMM)[92, 54] to accelerate inter-object transfer evaluations using progressive far-field approximations. Moreover, real-time performance could be achieved by further using GPU-based dense linear solvers. The incoming field strength computation for a moving source is similar to inter-object transfer. Thus, the combination of FMM and GPU-based computations could enable dynamic sources along with a moving listener. Also, our current runtime system does not model Doppler effect, which we would like to address in future work.

Currently, our sound sources emit a pre-recorded audio clip. An interesting direction of future research would be to integrate acoustic radiators based on mechanical physical models [181, 31] as potential sound sources, thus enabling physically-based real-time sound synthesis and propagation.

The computational complexity and runtime memory requirement of our technique scale linearly with number of frequency samples which in turn scales linearly with the scene size (number of frequency samples $\propto$ length of impulse response $\propto$ scene size). Thus, our technique can easily handle scenes that are hundreds of meters wide. However, for massive outdoor scenes that span kilometers, the runtime memory requirement would become too high (GBs per source). We plan to address this in future by using FMM-based far field approximations to reduce the number of equivalent sources.

Our precomputation depends heavily on the maximum simulation frequency thereby limiting it to 1-2 kHz. This behavior is consistent with other wave-based techniques like BEM and FDTD, which are also computationally limited to a few kHz. Geometric approximations become quite accurate for outdoor scenes at higher frequencies because buildings and terrain have much larger dimensions than the wavelength of 17cm at 2kHz. Thus, hybridization of our technique with geometric methods could lead to accurate wide-band propagation techniques for open scenes. Hybridization is an active area of research in acoustics [140].

2.9 Appendix

2.9.1 Two-object Steady State Field

We describe in detail the way we compute the equivalent source strengths for a scene composed of two objects. Consider a scene with objects A and B and a sound source s . Let the incoming field multipoles for A and B be Φ_A^{in} and Φ_B^{in} , respectively. Similarly, let the multipoles for the outgoing field for A and B be Φ_A^{out} and Φ_B^{out} , respectively. The scattering matrices for A and B are T_A and T_B , respectively. Let the interaction matrices for the objects be G_A^B and G_B^A , respectively. First of all, we express the incoming field produced by sound source s on objects A and B in terms of their incoming field multipoles :

$$\begin{aligned} s_A^{in} &= \sum_{i=1}^Q \sum_{k=1}^{M^2} a_{ik} \varphi_{ik}^{in} = S_A^{tr} \Phi_A^{in}; \\ s_B^{in} &= \sum_{i=1}^Q \sum_{k=1}^{M^2} b_{ik} \varphi_{ik}^{in} = S_B^{tr} \Phi_B^{in}. \end{aligned}$$

Now assume that the steady state outgoing field of object A and B is P_A^{out} and P_B^{out} respectively.

$$P_A^{out} = \sum_{j=1}^P \sum_{h=1}^{N^2} c_{jh}^A \varphi_{jh}^{out} = C_A^{tr} \Phi_A^{out}; \quad (2.20)$$

$$P_B^{out} = \sum_{j=1}^P \sum_{h=1}^{N^2} c_{jh}^B \varphi_{jh}^{out} = C_B^{tr} \Phi_B^{out}. \quad (2.21)$$

The outgoing field of one object becomes the incoming field for the other object. Exploiting the linearity of the inter-object transfer function and (2.14), we find the incoming field for B produced by the outgoing field of A as

$$\hat{P}_B^{in} = g_A^B(C_A^{tr} \Phi_A^{out}) = C_A^{tr} G_A^B \Phi_B^{in}.$$

Similarly, we find the incoming field for A produced by the outgoing field of B as

$$\hat{P}_A^{in} = g_B^A(C_B^{tr} \Phi_B^{out}) = C_B^{tr} G_B^A \Phi_A^{in}.$$

The total incoming fields on objects A and B are given by

$$P_A^{in} = s_A^{in} + \hat{P}_A^{in} = S_A^{tr} \Phi_A^{in} + C_B^{tr} G_B^A \Phi_A^{in};$$

$$P_B^{in} = s_B^{in} + \hat{P}_B^{in} = S_B^{tr} \Phi_B^{in} + C_A^{tr} G_A^B \Phi_B^{in}.$$

Applying the linearity of per-object transfer function f and using (2.9), we get outgoing pressure P_{out}^A and P_{out}^B due to the scattering of incoming fields by the objects as

$$P_A^{out} = f(P_A^{in}) = (S_A^{tr} T_A + C_B^{tr} G_B^A T_A) \Phi_A^{out}, \quad (2.22)$$

$$P_B^{out} = f(P_B^{in}) = (S_B^{tr} T_B + C_A^{tr} G_A^B T_B) \Phi_B^{out}. \quad (2.23)$$

In steady state, this outgoing pressure should match the outgoing pressure we started with. Equating (2.22) with (2.20), and (2.23) with (2.21), we get

$$C_A^{tr} = S_A^{tr} T_A + C_B^{tr} G_B^A T_A;$$

$$C_B^{tr} = S_B^{tr} T_B + C_A^{tr} G_A^B T_B.$$

Combining the above two equations, and rearranging, we obtain

$$\begin{bmatrix} C_A \\ C_B \end{bmatrix} = \begin{bmatrix} T_A^{tr} & 0 \\ 0 & T_B^{tr} \end{bmatrix} \left(\begin{bmatrix} 0 & (G_B^A)^{tr} \\ (G_A^B)^{tr} & 0 \end{bmatrix} \begin{bmatrix} C_A \\ C_B \end{bmatrix} + \begin{bmatrix} S_A \\ S_B \end{bmatrix} \right).$$

In other words,

$$\mathbf{C} = \mathbf{T}(\mathbf{G}\mathbf{C} + \mathbf{S}), \quad (2.24)$$

or,

$$(\mathbf{I} - \mathbf{T}\mathbf{G})\mathbf{C} = \mathbf{T}\mathbf{S}, \quad (2.25)$$

which is a linear system $Ax = b$. We solve this linear system to get the outgoing equivalent source strengths

C. At runtime, the outgoing scattered field at any listener position $\mathbf{x}$ is given by

$$p(\mathbf{x}) = C_A^{tr} \Phi_A^{out}(\mathbf{x}) + C_B^{tr} \Phi_B^{out}(\mathbf{x}). \quad (2.26)$$

The total pressure field becomes

$$p(\mathbf{x}) = C_A^{tr} \Phi_A^{out}(\mathbf{x}) + C_B^{tr} \Phi_B^{out}(\mathbf{x}) + s(\mathbf{x}). \quad (2.27)$$

2.9.2 Multiple Objects Steady State Field

For a scene with κ objects, $A_1, A_2, \dots, A_\kappa$, equation (2.25) remains the same except the vectors and matrices are generalized for κ objects.

The total pressure field becomes

$$p(\mathbf{x}) = \sum_{j=1}^{\kappa} C_{A_j}^{tr} \Phi_{A_j}^{out}(\mathbf{x}) + s(\mathbf{x}). \quad (2.28)$$

2.9.3 Computational Complexity

BEM The storage requirements of BEM depends on the total surface area S of the objects in the scene and the number of frequency samples $\nu_{\max}/\Delta\nu$. Assuming BEM places τ samples per wavelength (usually $\tau = 12$), the number of BEM elements placed on the object's surface at frequency sample ν_i is equal to $S\tau^2\nu_i^2/c^2$. The total number of BEM elements for all the frequency samples is equal to $S\tau^2\nu_{\max}^3/(3c^2\Delta\nu)$, where each element is specified by its position (3 floats) and four complex amplitudes corresponding to pressure and its gradient (2 floats each). Total memory requirement of storing the simulation results becomes

$$11S\tau^2\nu_{\max}^3/(3c^2\Delta\nu).$$

ARD and FDTD The runtime memory requirements of ARD and FDTD are equal to the number of grid cells in the spatial discretization of the entire volume of the scene and the number of timesteps in the simulation. Assuming volume of the scene to be V , the grid size h , the maximum frequency $\nu_{\max}$, the speed of sound c , and the number of samples per wavelength τ (equal to 3 for ARD and 10 for FDTD), the number of grid cells are $(\tau\nu_{\max}/c)^3V$. The total number of time samples to store is at least twice the number of samples in the frequency domain. The total memory requirement of storing the simulation results for these techniques is thus

$$2\tau^3\nu_{\max}^4V/(c^3\Delta\nu).$$

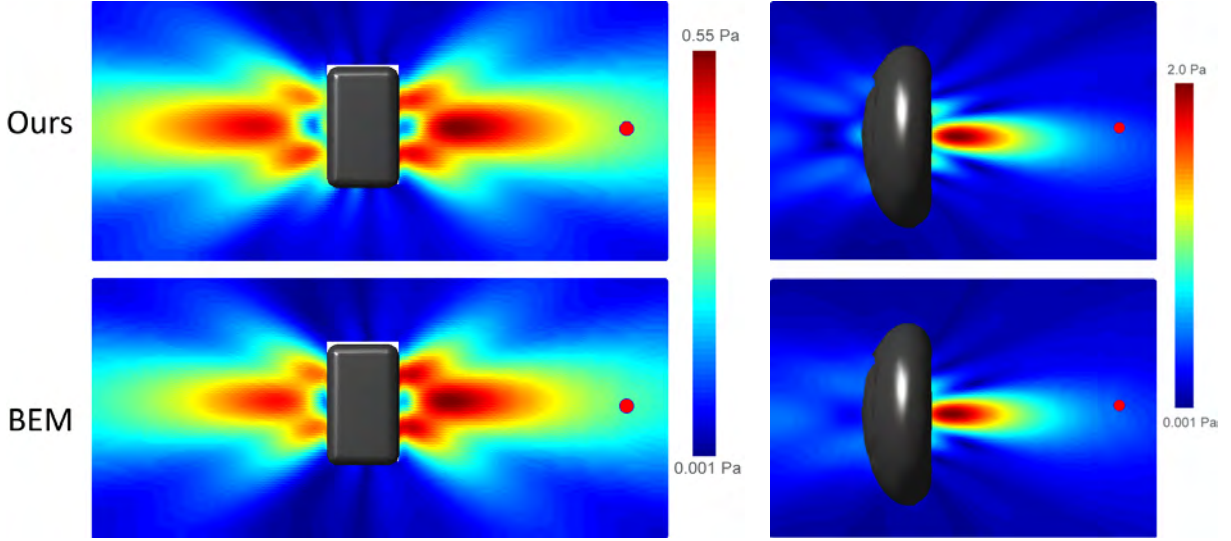


Figure 2.5: We show the scattering behavior of a wall (2.3m x 4.5m x 3.7m) and a concave reflector (diameter 8m, thickness 1.2m) at 160Hz using our technique (top row) and BEM (bottom row). The sound source is shown with a red dot.

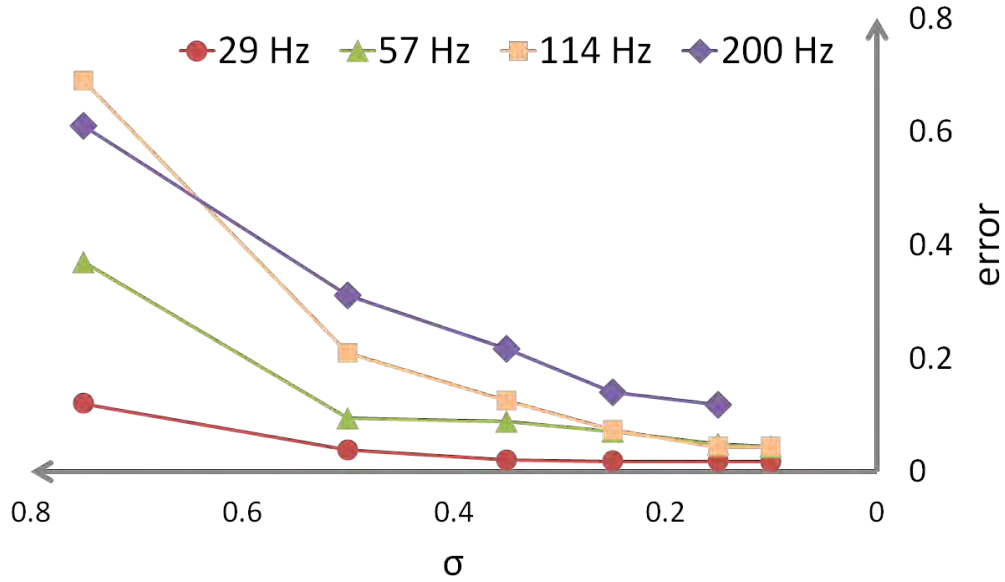


Figure 2.6: **Convergence:** We show the variation of error $\|P_{ref} - P_{ESM}\|^2 / \|P_{ref}\|^2$ between the reference wave solver and our technique for varying values of scattering matrix error threshold σ for the two parallel walls scene (fixed $\eta = 1\%$). P_{ref} and P_{ESM} are vectors consisting of (complex) pressure values at all the receiver locations as given by the reference wave solver and our technique, respectively. The receivers are placed on a XY grid for this scene.

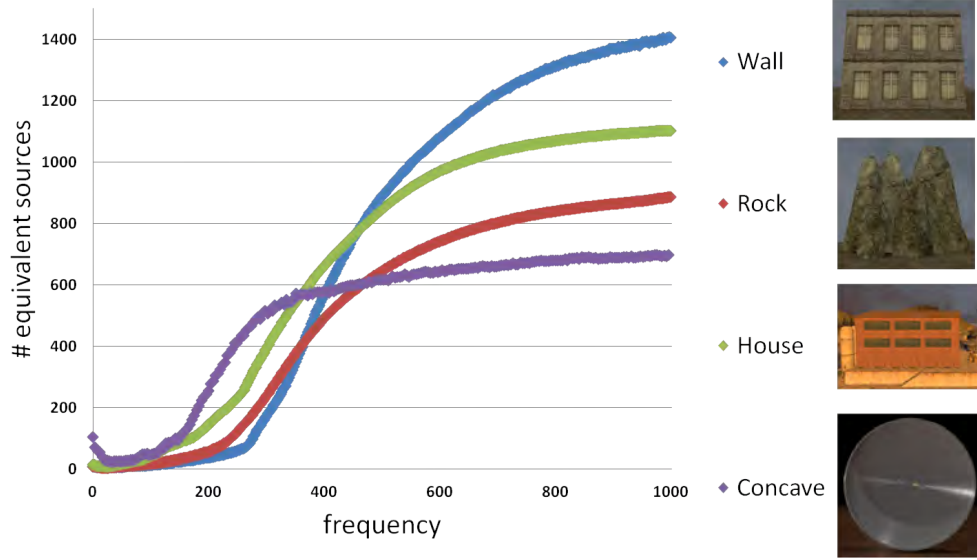


Figure 2.7: Variation of the number of outgoing equivalent sources with frequency, for four different objects. As the frequency increases (wavelength decreases), surface details of the size of the wavelength increase the complexity of the sound field. This results in a larger number of equivalent sources. When all the details of the object are captured, increasing the frequency has little effect and the number of equivalent sources begin to stabilize. Error thresholds are $\sigma = 15\%$ and $\eta = 1\%$.

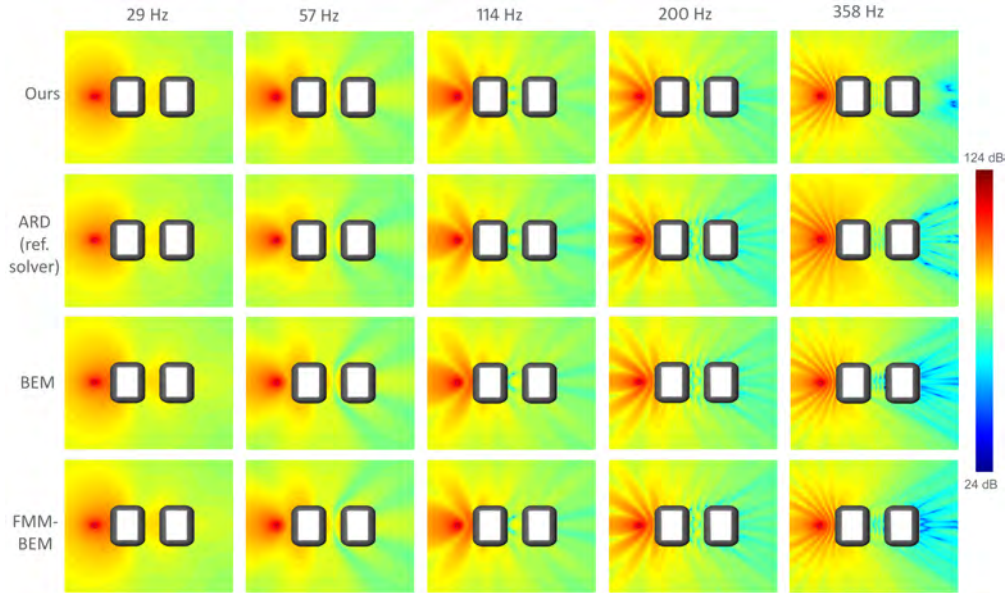


Figure 2.8: Comparison between the magnitude of the total pressure field (in Sound Pressure Level SPL, units dB) computed by our ESM, the reference wave-solver ARD, BEM and FMM-BEM techniques for the two parallel walls scene on a XY cutview grid of listeners. The red point denotes the position of the sound source. The error (defined in Figure 2.6 caption) between the ARD-ESM fields is $< 3\%$, BEM-ESM fields is $< 5\%$ and FMM BEM-ESM fields is $< 5\%$ for the frequencies shown.

Chapter 3

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

3.1 Overview

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.

3.2 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 [20].

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) [124, 98], 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

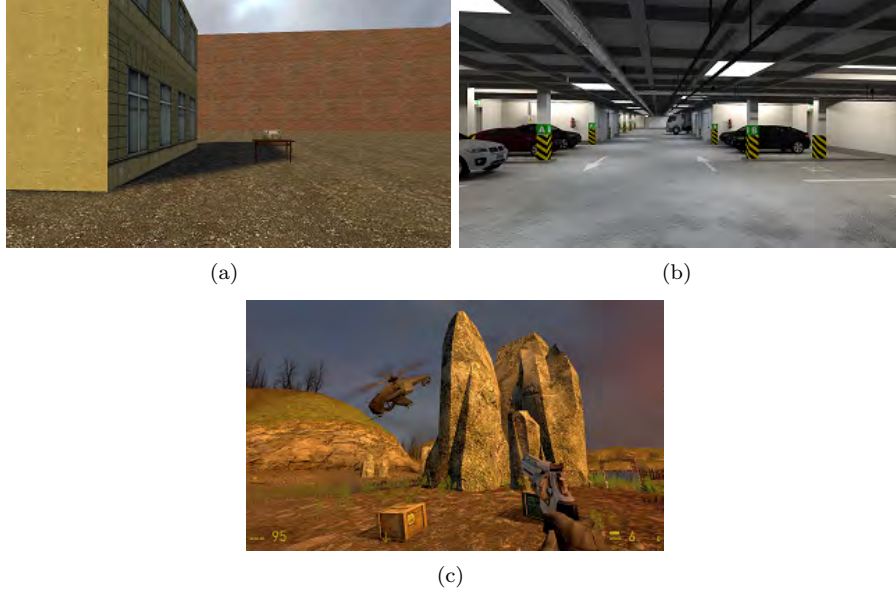


Figure 3.1: Our hybrid technique is able to model high-fidelity acoustic effects for large, complex indoor or outdoor scenes at interactive rates: (a) building surrounded by walls, (b) underground parking garage, and (c) reservoir scene in Half-Life 2.

SourceTM 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.

3.3 Related Work

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

3.3.1 Numerical Acoustic Techniques

Accurate, numerical acoustic simulations typically solve the acoustic wave equation using numerical methods, such as finite differences [23], finite elements [152], boundary elements, or adaptive rectangular decomposition [124]. 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 [104], 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. [71] solved a related *sound radiation* problem, using equivalent sources to represent the radiation field generated by a vibrating object.

3.3.2 Geometric Acoustic Techniques

Most acoustics simulation software and commercial systems are based on geometric techniques [45, 166] that assume sound travels along linear rays [47]. 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 [143, 158]. These ray-based techniques can interactively model early reflections and first order edge-diffraction [150]; 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 [88], the number of rays traced has to be limited for scenes with moving listeners in order to maintain real-time performance.

3.3.3 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), the digital waveguide mesh method (DWM), and the finite element method (FEM). 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 sim-

ulate wave effects, while far-field regions are handled by geometric acoustic techniques. Hampel et al. [56] 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. [169] 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. [15] propose a two-way coupling that combines the acoustic field generated using ray-tracing and FEM. Jean et al. [72] 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.

3.3.4 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 runtime for interactive propagation in static environments. Raghuvanshi et al. [125] 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. [98] 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 [155, 8] 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.

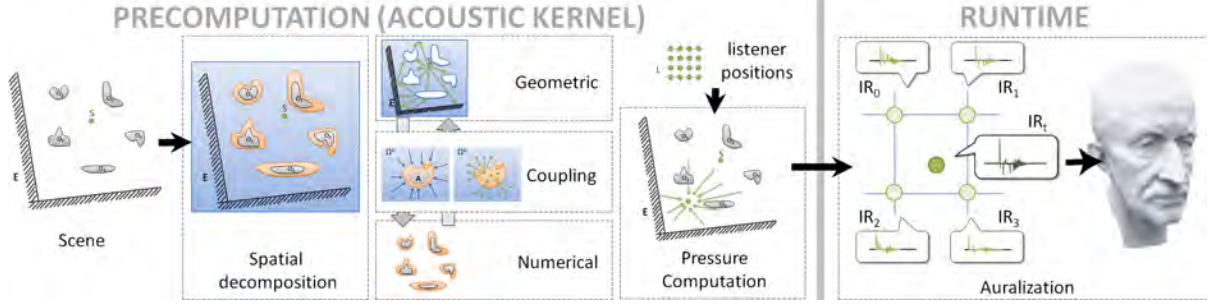


Figure 3.2: Overview of spatial decomposition in our hybrid sound propagation technique: In the precomputation phase, a scene is classified into objects and environment features. This includes near-object regions (shown in orange) and far-field regions (shown in blue). The sound field in near-object 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.

3.4 Overview

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

3.4.1 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 :

$$\nabla^2 p + \frac{\omega^2}{c^2} p = f; \quad \mathbf{x} \in \Omega, \quad (3.1)$$

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 an 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* [116],

$\lim_{||\mathbf{x}|| \rightarrow \infty} \left[\frac{\partial p}{\partial ||\mathbf{x}||} + \hat{j}\omega c p \right] = 0$, 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 (3.1) 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.

3.4.2 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 [171] and maps the incoming and outgoing fields in terms of the coefficients of a complete system of vector basis functions. Antani et al. [8] compute an acoustic radiance transfer operator that maps incident sound to diffusely reflected sound in a scene. Mehra et al. [98] model the free-field acoustic behavior of an object, as well as pairwise interactions between objects. In sound radiation problems, James et al. [71] map the vibration mode of an object to the radiated sound pressure field.

3.4.3 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 3.3. Since frequency decomposition is a standard technique [51], we mostly focus on spatial decomposition and our novel two-way coupling algorithm (see Figure 3.2).

Frequency Decomposition: We divide the modeled frequencies to low and high frequencies, with a crossover frequency $\nu_{\max}$. 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 $\nu_{\max}$ are 0.5-2 kHz, and a simple low-pass-high-pass filter

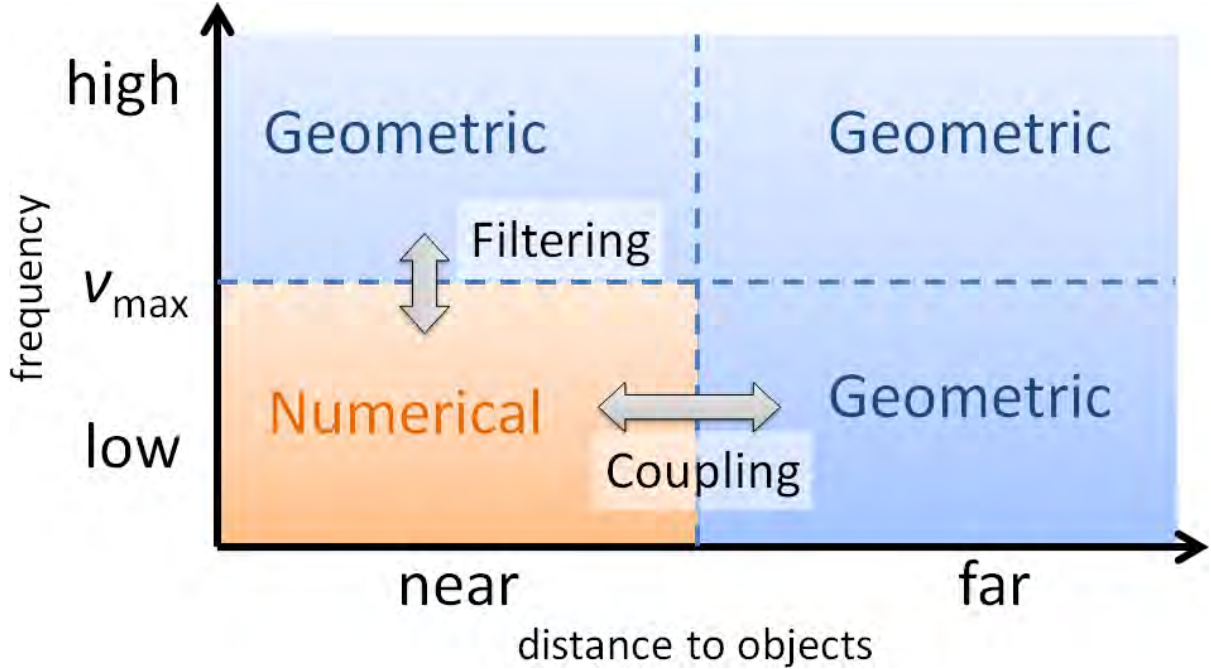


Figure 3.3: Frequency and spatial decomposition. High frequencies are simulated using geometric techniques, while low frequencies are simulated using a combination of numerical and geometric techniques based on a spatial decomposition.

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 3.2) 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 3.2).

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. [71]. 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 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 $p^G(\mathbf{x})$,

$$p^G(\mathbf{x}) = \sum_{r \in R} p_r(\mathbf{x}), \quad (3.2)$$

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 3.4). 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 3.5.

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

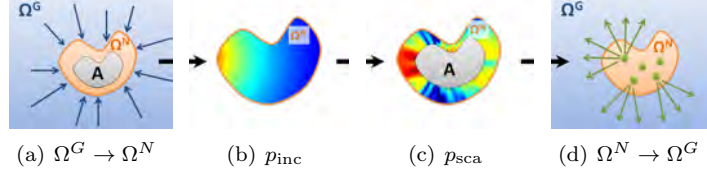


Figure 3.4: Two-way coupling of pressure values computed by geometric and numerical acoustic techniques. (a) The rays are collected at the boundary and the pressure evaluated. (b) The pressure on the boundary defines the incident pressure field p_{inc} in Ω^N , which serves as the input to the numerical 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 .

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 [125], and the output sound is auralized by convoluting the input sound with the IRs in real time.

3.5 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 (3.1) in both domains Ω^G and Ω^N .

3.5.1 Geometric $\rightarrow$ 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 (3.1) with forcing term $f = 0$.

Since there is no object in domain Ω^G ,

$$p_{\text{inc}}(\mathbf{x}) = p^G(\mathbf{x}); \quad \mathbf{x} \in \Omega^G. \quad (3.3)$$

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

$$p = p^G(\mathbf{x}); \quad \mathbf{x} \in \partial A^+, \quad (3.4)$$

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,

$$p_{\text{inc}}(\mathbf{x}) = \sum_i c_i \varphi_i(\mathbf{x}) \quad \mathbf{x} \in \Omega^N, \quad (3.5)$$

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 (3.4) is satisfied, then $p_{\text{inc}}(\mathbf{x})$ is the required *unique* solution to the boundary value problem (Section 3 in Ochmann [104]). 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 3.4(a) and 3.4(b) illustrate the process.

3.5.2 Numerical $\rightarrow$ 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

¹A 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 [165].

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,

$$p_{\text{sca}}(\mathbf{x}) = p^N(\mathbf{x}); \quad \mathbf{x} \in \Omega^N. \quad (3.6)$$

We need to find the scattered field outside of Ω^N , and model it as a set of rays. As before, Equation (3.6) defines a Dirichlet boundary condition on the interface ∂A^+ ,

$$p = p^N(\mathbf{x}); \quad \mathbf{x} \in \partial A^+. \quad (3.7)$$

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 :

$$p_{\text{sca}}(\mathbf{x}) = \sum_j c_j \varphi_j(\mathbf{x}); \quad \mathbf{x} \in \Omega^G, \quad (3.8)$$

and then find the coefficients c_j by satisfying the boundary condition (3.7). 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

$$\varphi_j(\mathbf{x}) = \sum_{r \in R_j^{\text{out}}} p_r(\mathbf{x}), \quad \mathbf{x} \in \Omega^G. \quad (3.9)$$

These rays correctly represent the outgoing scattered field in Ω^G . Figure 3.4(c) and 3.4(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 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 [150]). 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 (3.5)) and outgoing scattered field (Equation (3.8)) are *predetermined*, then the mapping from φ_i to φ_j can be precomputed. This precomputation process will be discussed in section 3.5.4.

3.5.3 Fundamental solutions

The requirement for the choice of fundamental solution φ_j is that it must satisfy the Helmholtz Equation (3.1) and the Sommerfeld radiation condition.

Equivalent Sources: One choice of fundamental solutions is based on *equivalent sources* [104]. 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 $\mathbf{y}_j$:

$$\varphi_j(\mathbf{x}) = \varphi_{jlm}(\mathbf{x}), \quad (3.10)$$

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

$$\varphi_{jlm} = \Gamma_{lm} h_l^{(2)}(\omega \rho_j / c) \psi_{lm}(\theta_j, \phi_j), \quad (3.11)$$

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 [10].

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 (3.9)) 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 (3.8) 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:

$$p(\mathbf{x}) = \int_{\partial\Omega} \left(G(\mathbf{y}, \mathbf{x}) \frac{\partial p(\mathbf{y})}{\partial n} - \frac{\partial G(\mathbf{y}, \mathbf{x})}{\partial n} p(\mathbf{y}) \right) d(\partial\Omega(\mathbf{y})), \quad (3.12)$$

where $\mathbf{y}$ is the approximated position of the triangle and G is the Green's Function $G(\mathbf{y}, \mathbf{x}) = \exp(j\omega|\mathbf{x} - \mathbf{y}|/c)/4\pi|\mathbf{x} - \mathbf{y}|$ [55].

Note that the discretization of Equation (3.12) also takes the form of Equation (3.8) as a linear combination of fundamental solutions:

$$p(\mathbf{x}) = \sum_j (c_j^1 \varphi_j^1(\mathbf{x}) + c_j^2 \varphi_j^2(\mathbf{x})), \quad (3.13)$$

where the two kinds of fundamental solutions are

$$\varphi_j^1(\mathbf{x}) = G(\mathbf{y}_j, \mathbf{x}) \frac{\partial p(\mathbf{y}_j)}{\partial n}; \quad \varphi_j^2(\mathbf{x}) = -\frac{\partial G(\mathbf{y}_j, \mathbf{x})}{\partial n} p(\mathbf{y}_j). \quad (3.14)$$

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 (3.12).

3.5.4 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 :

$$R^{\text{in}} \xrightarrow{\mathcal{M}} R^{\text{out}}, \quad (3.15)$$

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 [17]. 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 (3.4));
2. Express the incident field as a set of fundamental solutions (Equation (3.5));
3. Compute the outgoing scattered field using the numerical acoustic technique;
4. Express the outgoing scattered field as a set of fundamental solutions (Equation (3.8)); and finally,
5. Find a set of rays R^{out} that model these functions (Equation (3.9)).

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 transferred 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 [15, 72], 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.

3.6 Implementation

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

Scene	#src	#freq.	#eq. srcs	Hybrid Pressure Solving								Pressure
				Numerical				Geometric				Evaluation
				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 3.1: **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).

3.6.1 Implementation details

The geometric acoustics code is written in C++, based on the Impulsonic sound propagation SDK, 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 [124]. Per-object transfer functions, inter-object transfer functions, and equivalent source strengths are computed using a MATLAB implementation based on [98].

Table 3.1 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-CPU (8 dual-core 2.8GHz Opterons, 32GB RAM each) and forty-five 4-CPU (2 dual-core 2.6GHz Opterons, 8GB RAM each).

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 $\nu_{\max}$, $h = c/2\nu_{\max}$, 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.

3.6.2 Collocated equivalent sources

The positions of outgoing equivalent sources can be generated by a greedy algorithm that selects the best candidate positions randomly [71]. 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 $\nu_{\max}$, 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 reflected 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.

3.6.3 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 $\nu_{\max}$, we use a ray tracing solution, with diffractions approximated by the Uniform Theory of Diffraction (UTD) model [82]. 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-

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 3.2: **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.

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. [125]. 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 [116]. 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 3.2. The parking garage scene is rendered off-line and not included in this table.

3.7 Results and Analysis

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

3.7.1 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 [71, 124, 98] or geometric methods with precomputed high-order reverberation [155, 8]. 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 [150], which models specular reflection as well as edge diffraction through UTD; a pure numerical technique, the ARD wave-solver [124]. 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.

3.7.2 Error Analysis

In Figure 3.5 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 [15, 72, 56]. BEM results are generated by the FastBEM simulator. 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 3.6 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 $\|P_{\text{ref}} - P_{\text{ref}}\|^2 / \|P_{\text{ref}}\|$, where P_{ref} and P_{ref} 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 3.5). Higher reflection orders lead to more accurate results but require more rays to be traced.

3.7.3 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 [98]. In our system, the equivalent sources are limited to monopoles and dipoles, and the complexity follows:

$$O(\kappa n Q P^2 + \kappa^2 n P Q^2 + \kappa(u \log u) + \kappa^3 P^3), \quad (3.16)$$

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

$$O(\kappa^2 P n T^d) \quad (3.17)$$

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 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 [45], 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 3.7. 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 [124].

3.7.4 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 3.1) to achieve interactive runtime performance (see Table 3.2). 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 3.3, 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. [98] 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

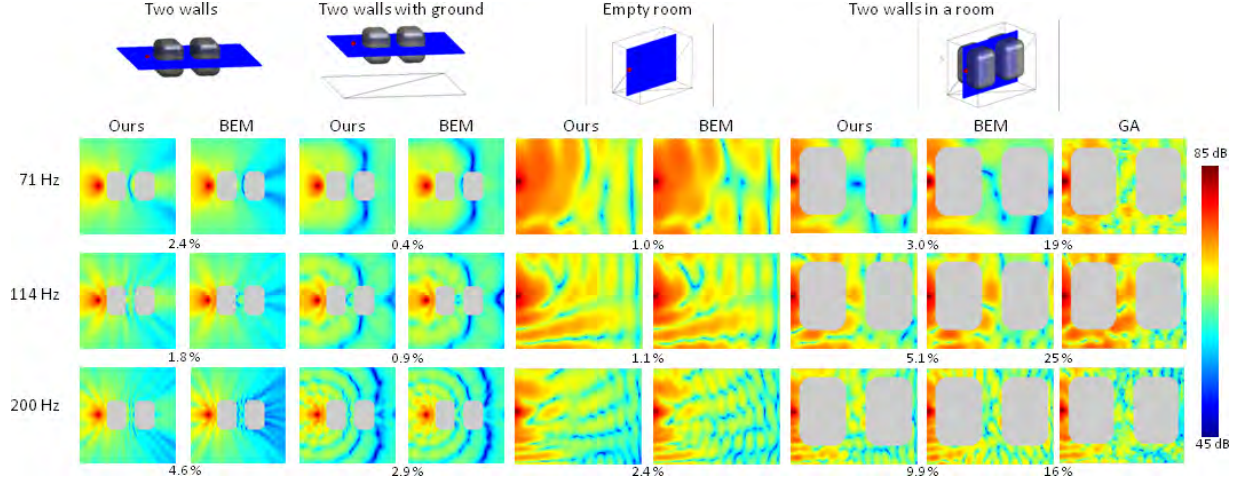


Figure 3.5: Comparison between the magnitude of the total pressure field computed by our hybrid technique and BEM for various scenes. In the top row, the red dot is the sound source, and the blue plane is a grid of listeners. Errors between our method and BEM for each frequency are shown in each row. 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”).

with uneven terrains. If an outdoor scene has a large object, the algorithm proposed in [98] 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 3.3: **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 performance Table. Compared to standard numerical techniques, our method provides **3 to 7 orders of magnitude** of memory saving on the benchmark scenes.

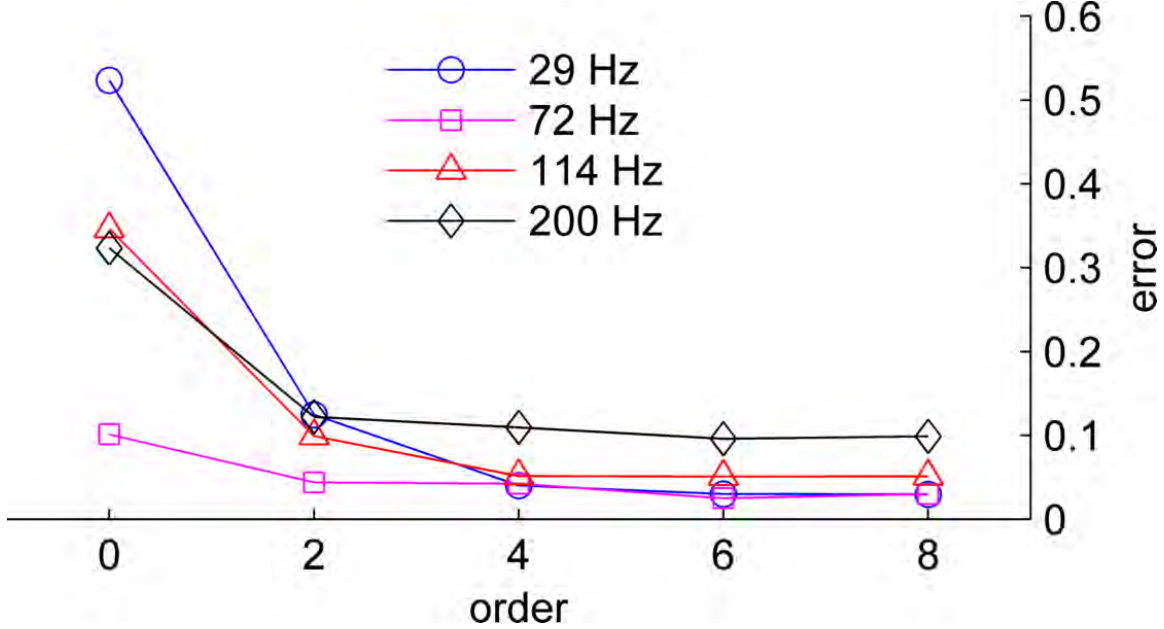


Figure 3.6: Error $\|P_{\text{ref}} - P_{\text{ref}}\|^2 / \|P_{\text{ref}}\|$ between the reference wave solver (BEM) and our hybrid technique for varying maximum order of reflections modeled. The tested scene is the "Two walls in a room" (see also Figure 3.5, last column).

3.8 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 numerical 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 its size is

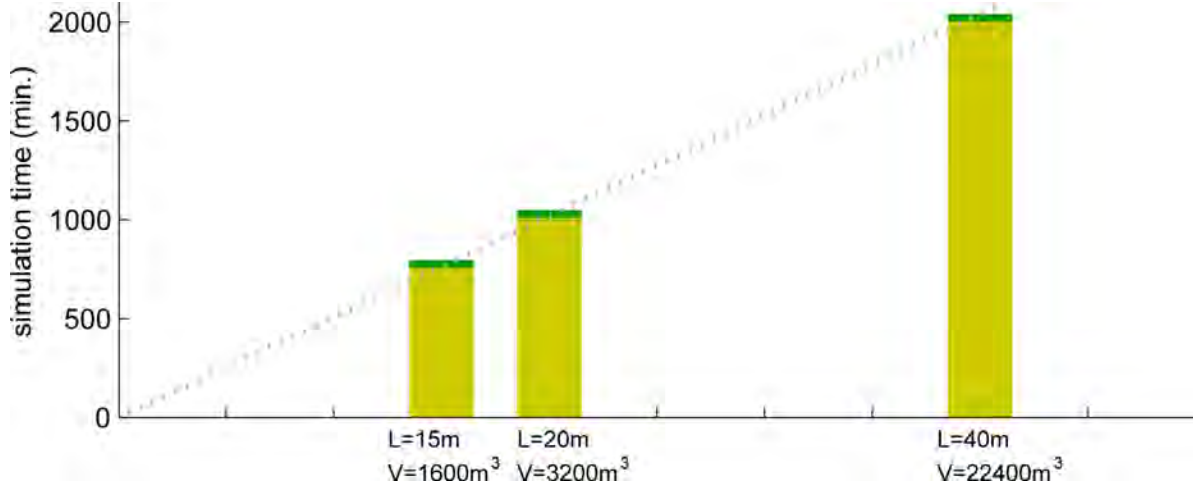


Figure 3.7: **Breakdown of Precomputation Time.** For a building placed in terrains of increasing volumes (small, medium, and large walls), the yellow part is the simulation time for the numerical method, and the green part is for the geometric method. The numerical simulation time scales linearly to the largest dimension (L) of the scene instead of the total volume (V).

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) [53] 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.

Chapter 4

Analytic Ray Curve Tracing for Outdoor Sound Propagation

4.1 Overview

Outdoor sound propagation presents challenging scenarios for computational simulation with inhomogeneous, moving media and complex obstacles. Existing methods rely on simplified models of media conditions or obstacles with limited complexity to keep the computational cost low. Numerical models that handle general media and arbitrary media boundaries do not scale well with the size and complexity of the scenes, making them prohibitively expensive for large three-dimensional outdoor scenes. In this paper, a ray tracing method using analytic ray curve as tracing primitives is presented to improve the efficiency of outdoor sound propagation in fully general settings. This ray curve tracer inherits the efficiency and flexibility of rectilinear ray tracers in handling boundary surfaces, and it overcomes the performance limitations imposed by approximating the curved propagation paths in inhomogeneous media with rectilinear rays. Adaptive media traversal as well as acceleration structures for surfaces intersection leads to further savings in computation. Speedup up to an order of magnitude is demonstrated on outdoor benchmark scenes in comparison with

existing ray models. This ray tracer serves as a stand-alone tool for fast outdoor sound propagation, and it is complementary to existing geometric and numerical methods and extendable in multiple ways.

4.2 Introduction

Sound propagation in outdoor environments[133, 11, 73] and room acoustics[85] are traditionally investigated separately. Room acoustics focus on the high order interactions that sound wave makes with surfaces within an enclosed space, assuming homogeneous media. On the other hand, outdoor sound propagation (atmospheric and underwater) deals with spatially varying as well as moving media, and the curved propagation trajectories resulted from refraction is the most prominent departure from indoor models. In the outdoor context, obstacles that constitute media boundaries are generally assumed to be sparse, and the complexity of the boundary surfaces is limited in the investigations.

However, many outdoor applications can benefit from simultaneous modeling of fully general media and complex boundaries, examples include large expanse of complex terrains and sprawling urban area that often has its own micro-climate. As the data on the spatially and temporally varying atmosphere and the ocean becomes increasingly available, and geographic information of terrains and man-made structures come in ever richer details, it would be ideal that the sound propagation can utilize the full scale of such data. Existing numerical methods such as FDTDs suffer from performance and scalability issues with large domain volume, high frequency, or complex boundaries. Geometric acoustics (GA)[115] methods like ray models are efficient at handling boundary surfaces, but with inhomogeneous media the rectilinear rays are forced to advance in small steps.

One way to speed up ray models for inhomogeneous media involves subdividing the media into spatial cells, so that within each cell the media assumes simple profile that leads to an analytic curved trajectory. The *cell method* was proposed as exactly such a model based on a triangular subdivision of two-dimensional media, and it has been implemented in some acoustic simulation software[130, 172]. A later work[101] for both light and sound propagation followed the idea of cell method and decomposed three-dimensional

media profile into tetrahedral cells. Mesh connectivity is utilized to trace rays across neighboring cells, and boundary surfaces are embedded into the mesh so that boundary intersection is solved in the same mesh traversing process.

This paper develops the idea of cell method further, and improves upon the tetrahedral cell method[101] in three important aspects of the algorithm:

1. **Simple ray formulation** We select the parabolic ray curve to be the ray tracing primitives, giving the simplest analytic form of trajectory, intersection, and ray properties (Sec.4.4).
2. **Implicit cell** We use a mesh-less approach for media traversal, tracing ray curve segments of adaptive sizes based on on-the-fly sampling of the media profile. This approach of implicit cells avoids the costly mesh construction, and it supports moving media as well as dynamic media (Sec.4.5.2).
3. **Acceleration of surface intersection** we adapt hierarchical acceleration structures used in rectilinear ray tracer for the ray curve tracer. Further acceleration is achieved by spatial bounding of ray curves based on their geometric properties for higher culling efficiency (Sec.4.5.3).

Overall this analytic ray curve tracer is designed to be efficient for both general media profile and complex surfaces. Its performance is demonstrated on outdoor benchmarks(Sec.4.6), showing significant improvements over previous ray models and software tools. It is complementary to a set of numerical and geometric methods, and can be extended in multiple ways as discussed in Sec. 4.7.

4.3 Prior Work

Outdoor sound propagation has been studied extensively in the context of underwater[73] as well as atmospheric acoustics[133]. Early methods include Fast Field Program (FFP)[39] and Parabolic Equation (PE)[102, 50]. Both methods provide frequency domain full wave solutions with simplifying assumptions about the media. More general numerical methods have been proposed to handle arbitrary media and obstacles, the most popular one being the Finite Difference Time Domain(FDTD) method that solves the

linearized Euler equation[110, 175]. The disadvantage of the FDTD is mainly its high computational costs, in particular, it does not scale well with the volume of the simulation domain or the frequency. Other frequency domain methods including FEM[103, 164], BEM[26], Pseudo Spectral Time Domain[67, 66], and Transmission Line Matrix (TLM)[64, 52, 13] face similar challenges in scalability, making them prohibitively expensive for large-scale broadband simulation.

Geometric acoustics (GA)[115] is a family of methods that are widely used in room acoustics[85]. They rely on the high frequency approximation of wave propagation with rays, which are perpendicular to the wave fronts and can be intersected efficiently with scene surfaces. Various GA techniques were developed, including image source method[7, 22], ray tracing[167, 149], frustum tracing[148], beam tracing[46, 32], sonal mapping[19, 38, 79], path tracing[136], and radiosity[160]. The sound speed is generally assumed to be constant so rectilinear rays are used to trace out straight-line paths. Complex interactions are modeled at ray-surface intersections, such as specular reflection, Snell’s Law refraction, diffuse reflection [149, 136], and Bidirectional Reflectance Distribution Function (BRDF)[79, 137].

The ray models have also been applied to inhomogeneous outdoor media by numerical integration of the differential ray equation, effectively tracing segments of rectilinear rays (see survey [133, 73]). The segment size is limited by the range within which the media can be assumed to be homogeneous, and intersection with surfaces need to be tested for each segment. With complex media the great number of small ray segments dominate the propagation computation. Higher order integration like the fourth-order Runge-Kutta method is adopted to improve efficiency, but with those methods the segment between each advancement of the ray integration is no longer rectilinear, complicating intersection tests with surfaces. Several widely used software are based on ray models, including HARPO[75], BELLHOP[118], WaveQ3D[126] in underwater acoustics, HARPA[74] in atmospheric acoustics (see [107] for the collection). Although performance of those ray-based software are orders of magnitude better than wave-based alternatives (a few seconds vs. thousands of seconds on 2D propagation in layered media, reported recently[168, 43]), they quickly becomes prohibitively expensive with general media and three-dimensional complex scenes.

4.4 Analytic Ray Curve

Ray models simulate the wave propagation with rays that are perpendicular to the wavefront. For certain media profiles, the ray trajectories as well as the pressure on the ray have analytic form that can be evaluated at any point along the ray with constant cost. Our ray tracer utilizes one of such analytic formulations, chosen for its particularly simple form, and the resulting ray curve lends itself to easy intersections with boundary surfaces. As can be seen from the algorithm outline in Fig. 4.2, this particular ray curve enables analytic evaluations of constant per-ray cost in key steps of the ray tracing (green-colored blocks). This ray formulation is a special case of the more general ray theory discussed in depth in Červený's comprehensive work[161].

4.4.1 Ray formulation

Within a media with spatially varying sound speed, we denote $V(x_i) = c$ as the sound speed at location x_i . The *arrival time* or *travel time* field T is defined as the time that a sound wave takes to travel from its source to a field location, and the spatial derivative of T : $\vec{p} = \nabla T$ is called the *slowness vector*. Within such a media, the ray trajectories can be solved from the general *Hamiltonian* form of the Eikonal equation, which in Cartesian coordinates reads:

$$\mathcal{H}(x_i, p_i) = n^{-1}(p_i p_i)^{n/2} - \frac{1}{V^n}, i = 1, 2, 3, \quad (4.1)$$

and n is a real number. Equation 4.1 can be solved in terms of *characteristics*: 3-D trajectories $x_i = x_i(u)$ along which 4.1 is satisfied, parameterized by u . The characteristic system of Equation 4.1 is:

$$\begin{aligned} \frac{dx_i}{du} &= (p_k p_k)^{n/2-1} p_i, & \frac{dp_i}{du} &= \frac{1}{n} \frac{\partial}{\partial x_i} \left(\frac{1}{V^n} \right), \\ \frac{dT}{du} &= (p_k p_k)^{n/2} = V^{-n}. \end{aligned} \quad (4.2)$$

When the gradient of V^{-n} is constant, 4.2 can generally be solved by quadrature. When we assume the

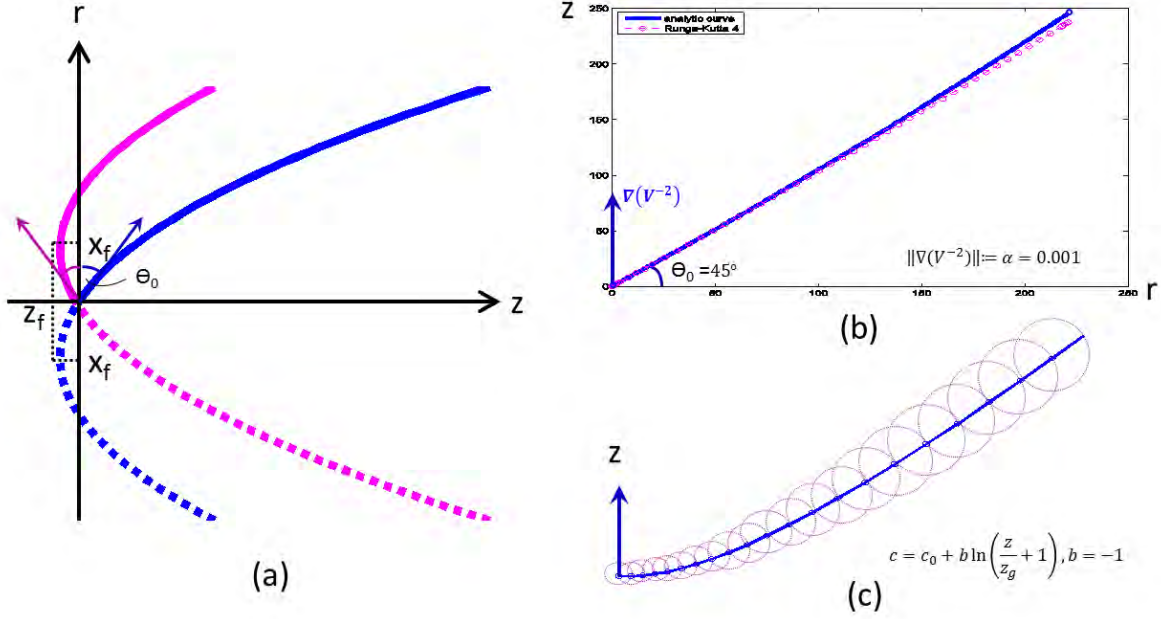


Figure 4.1: **(a) Parabolic ray curve:** shown in r - z plane, z -axis being the direction of ∇V^{-2} . Rays with two different initial directions (marked by arrows) are shown here. (x_f, z_f) defining the vertex of the parabola is computed from Eq. 4.6. **(b) Efficiency of curve tracing:** with constant $\|\nabla V^{-2}\|$, one segment of analytic curve (blue line) suffices as compared to many steps (magenta circles) taken by numerical ray integration. The trajectory traced out by RK-4 visibly diverges from the parabolic curve, which can be corrected by further reducing step sizes. **(c) Adaptive segments of curves:** with logarithmic sound speed profile, our ray tracer proceeds in segments of parabolic curves with adaptive sizes, evident from the visualization of the spheres of validity (Sec. 4.5.2).

gradient of V^{-2} is constant (i.e. $n = 2$): $V(\mathbf{x})^{-2} = A_0 + \vec{A} \cdot \mathbf{x}$, we obtain the simplest analytic form for ray trajectory and travel time, in terms of the parameter $u = \sigma$:

$$x_i(\sigma) = x_{i0} + p_{i0}(\sigma - \sigma_0) + \frac{1}{4}A_i(\sigma - \sigma_0)^2, \quad (4.3)$$

$$p_i(\sigma) = p_{i0} + \frac{1}{2}A_i(\sigma - \sigma_0), \quad (4.4)$$

$$T(\sigma) = T(\sigma_0) + V_0^{-2}(\sigma - \sigma_0) + \frac{1}{2}A_i p_{i0}(\sigma - \sigma_0)^2 + \frac{1}{12}A_i A_i(\sigma - \sigma_0)^3, \quad (4.5)$$

and σ is related to the travel time T and to arclength s by $d\sigma = V^2 dT = V ds$.

4.4.2 Ray properties

We take advantage of two key properties of the ray curve given in Sec. 4.4.1:

1. **Intersection** the ray is a parabolic plane curve and can be intersected analytically with planar surfaces to solve for intersection point.
2. **Evaluation** properties including but not limited to position, tangent direction, slowness vector, arrival time, and arclength can be evaluated analytically for arbitrary parameter σ along the ray, at constant cost.

It is evident from equation 4.3 that the analytic ray lies in a plane with the normal of $\vec{A} \times \vec{p}_0$. We denote the direction of the media gradient $\vec{A}$ as the z -axis, and $(\vec{A} \times \vec{p}_0) \times \vec{A}$ as the r -axis, the r - z plane is then called the *ray plane*. When the origin of the coordinate system is placed at the ray origin $\mathbf{x}_0$, the ray trajectory can be written as:

$$z(r) = \frac{\alpha}{4\xi_0^2}(r - r_f)^2 - z_f, \quad r_f = -\frac{2\xi_0\sqrt{-\xi_0^2 + n_0^2}}{\alpha}, z_f = \frac{n_0^2 - \xi_0^2}{\alpha} \quad (4.6)$$

where $\alpha = \|\vec{A}\|$ is the magnitude of media gradient, $\xi_0 = V(0)^{-1}\cos\theta_0$, $V(0)$ is the sound speed at ray origin, and θ_0 is the angle between the ray direction at origin and the r axis (see Figure 4.1). The parabolic plane curve given by Eq. 4.6 has closed-form solutions in terms of intersection with planar surfaces.

For an arbitrary point on the ray curve with coordinates (r_p, z_p) in the ray plane, the corresponding parameter σ_p can be solved from Equation 4.3 as $\sigma_p = r_p/\xi_0 + \sigma_0$, where σ_0 is the σ at ray origin. Given σ_p , the position $\vec{x}(\sigma_p)$, slowness vector $\vec{p}(\sigma_p)$, and travel time $T(\sigma_p)$ are given by Eq. 4.3, 4.4, 4.5, respectively. The tangent direction $\vec{t}$ coincides with the direction of the slowness vector $\vec{p}$ and can be evaluated by: $\vec{t}(\sigma_p) = \vec{p}(\sigma_p)/\|\vec{p}(\sigma_p)\|$. The arclength s of the parabolic ray curve can be evaluated at (r_p, z_p) by

$$s(r_p) = \left. \frac{\sqrt{(ar_p + b)^2 + 1}(ar_p + b) + \operatorname{arcsinh}(ar_p + b)}{2a} \right|_0^r, \quad (4.7)$$

where $a = \frac{1}{2} \frac{\alpha}{\xi_0^2}, \quad b = \frac{\sqrt{-\xi_0^2 + V(0)^{-2}}}{\xi_0'}.$

Similar to rectilinear rays, the parabolic ray curve assuming constant ∇V^{-2} comes with efficient operations for finding intersection and updating the key variables (position, direction, arclength, travel time,

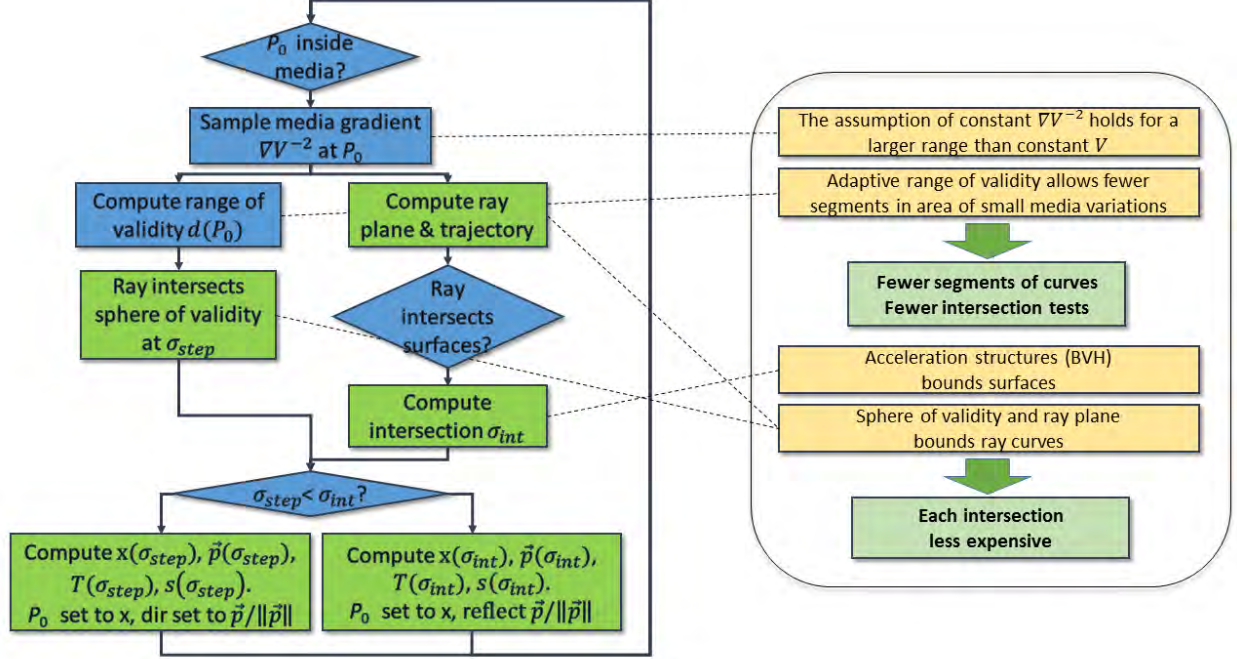


Figure 4.2: **Left** steps of tracing one segment of analytic curve, among which the green-colored blocks have analytic solutions due to the parabolic ray formulation, leading to large savings of ray tracing cost. **Right** Besides the analytic formulation, other sources of performance improvements lead to our ray tracer tracing fewer segments with fewer intersection tests, and cheaper intersection tests as well. Those algorithmic improvements are connected to the steps in the flow chart where they happen.

etc.) at intersection points, making it amenable to be used as a tracing primitive. The ray curve serves as a building block for a ray tracer that handles general media, by computing the propagation paths consisting of consecutive segments of such ray curves. The flow chart in Fig. 4.2 illustrates the steps involved in tracing one such segment, with the computational savings highlighted by green-colored blocks. The sources of performance improvements are listed on the right side of Fig. 4.2, and the ray tracing algorithm is detailed next.

4.5 Ray Tracing Algorithm

4.5.1 Tracing Analytic Segments

Our ray tracer computes the propagation paths in a segment-by-segment fashion, each segment being a parabolic ray curve formulated in Sec. 4.4. There are two criteria for terminating a segment and starting the next one:

1. **Media** being traversed no longer satisfy the assumption of constant ∇V^{-2} ,
2. **Surfaces** of media boundary or obstacles in the scene are encountered.

For the former we proposed an adaptive segment size computed from on-the-fly sampling of the media, discussed in details in Subsec. 4.5.2, and for the latter we used acceleration data structures and methods adapted from rectilinear ray tracer, discussed in Subsec. 4.5.3. Both aspects of the ray tracing algorithm contributes to its efficiency and performance improvements over previous methods.

Based on those criteria, we find out how far a particular segment goes, by computing the closest intersection point for surface encounter, and the point that the ray curve reach the edge of the range of validity for the constant media gradient assumption. The closest point between the two cases are taken as the end point of the current segment, at which σ , $\mathbf{x}(\sigma)$, $\vec{t}(\sigma)$, $T(\sigma)$, $s(\sigma)$ are evaluated. The next segment starts at $\mathbf{x}(\sigma)$, the trajectory of which is computed from the local media gradient $\nabla V^{-2}(\mathbf{x}(\sigma))$. σ , $T(\sigma)$, $s(\sigma)$ are continued over the next segment, and the ray direction $\vec{t}(\sigma)$ is also continued except for the surface encounter case. At surface encounter a reflection direction $\vec{t}$ is computed based on the surface reflection model as the initial direction of the next segment.

The ray traversal continue in this segment-by-segment way until it reaches outside the media volume, it interacts with surfaces a pre-determined number of times, or the pressure amplitude along the ray drops below a pre-set threshold. The ray tracer finds the set of propagation paths starting from the sound source, and the ray segments making up those paths can be queried for any field location to compute the pressure field.

4.5.2 Adaptive Media Traversal

Within a smoothly varying media with local coherence, the media gradient ∇V^{-2} can be assumed to be constant within a certain range of validity. This is a valid assumption for naturally occurring media such as the atmosphere or the ocean under stable conditions. Furthermore, our ray tracer uses the heuristics that the range of validity for a locally constant ∇V^{-2} should be proportional to $\|\nabla V^{-2}\|$, so that the media gradient remains valid for a relatively small spatial range in areas of great variations. Such adaptive sizing facilitates accurate traversal of the media without the cost of uniformly small segment size.

Based on this heuristics, we compute the range of validity $d(\mathbf{x})$ by $\delta = \frac{1}{8}(\nabla V^{-2})d^3(\mathbf{x})$, which is controlled by a global parameter δ . Given the local $\nabla V^{-2}(\mathbf{x})$, the ray curve given by Eq. 4.6 is computed, and continued until the border of a sphere centered at the ray origin $\mathbf{x}$ and with the radius $d(\mathbf{x})$. The parabolic ray curve has closed-form intersection with its validity sphere, and the exit point $\mathbf{x}'$ from the validity sphere is used to sample the media again for the next segment.

Given an input media profile $V(\mathbf{x})$, on-the-fly sampling of $\nabla V^{-2}(\mathbf{x})$ can be computed either analytically if the input profile is available as an analytic function, or by differentiation techniques such as central differences. Moving media can be approximated by the standard *effective sound speed* approach, adding to or subtracting the wind velocity $w(\mathbf{x})$ from $V(\mathbf{x})$, for upwind or downwind propagation respectively. Particularly for our ray tracer, the on-the-fly media gradient sampling enables handling of vector wind field by adding the wind speed projected onto the ray direction $V(\mathbf{x}) + \vec{w}(\mathbf{x}) \cdot \vec{t}(\mathbf{x})$.

4.5.3 Handling Boundary Surfaces

For existing rectilinear ray tracers, ray intersections with complex surfaces in the scene is the most expensive part of the computation, and many acceleration structures and techniques have been developed to speed it up. The parabolic ray curve we use as tracing primitives can be used with the hierarchical acceleration structures adapted from rectilinear ray tracing that bound the scene surfaces. The ray curves themselves can also be bounded spatially for further culling to save intersection costs.

For scene surfaces culling we built a BVH (Bounding Volume Hierarchy)[58] to group boundary surfaces

and bound them spatially. A BVH is a hierarchical tree structure whose internal nodes are bounding volumes (BVs), and we choose AABBs (axis-aligned bounding boxes) for the shape of the nodes. The parabolic ray curve can be intersected with an AABB much faster than with all its enclosed surfaces, and when the ray curve and the AABB are not intersecting, all those surfaces can be culled from further computations. To find the closest intersection with the scene surfaces, the BVH is traversed by each ray curve in a top-down fashion. Logarithmic intersection time on average is achieved with regard to the number of surfaces in the scene. In addition, the BVH can be re-fitted or rebuilt efficiently to accommodate dynamic scenes.

Besides spatial culling of the scene surfaces with a BVH, we also perform spatial bounding and culling of ray curves. As we traverse the media adaptively as described in Sec. 4.5.2, we compute a bounding sphere (the range of validity sphere) for each segment as a by-product. When intersecting such a segment with the scene BVH, the culling efficiency is improved by intersecting the bounding sphere of the segment with the BVs first, which is a simpler computation than intersecting the ray curve directly. Furthermore, as the parabolic ray is a plane curve, the ray plane is used to reduce the dimension of the intersection test which achieves further culling.

At the intersection point with a boundary surface, the interaction can be modeled by specular reflection, Snell’s law refraction, or BRDF-based reflection, depending on the surface properties, and the direction of the secondary ray segment can be computed accordingly. A surface impedance model can be used to modify the amplitude and the phase of the pressure along the ray, in addition to direction change.

4.6 Results

4.6.1 Performance Comparison of Ray Models

Ray models are the foundation of many software tools that are widely used in underwater acoustic simulation [130, 172, 118, 126], and to a lesser degree in atmospheric acoustics[74]. The majority of the available tools assume either range-independent media or profiles that vary in two dimensions, with the exception of HARPO and HARPA, which computes three-dimensional Hamiltonian ray tracing on clusters. The existing

software is written in different languages with different performance characteristics (e.g. MATLAB, Fortran), the available performance data are scarce and some of them rather dated. We reviewed the data we could find in Sec. 4.3, and here we compare the performance between our method and numerical ray integration that we implemented in our own code base.

Existing ray models use direct numerical integration of the ray equations with Euler method or higher order such as fourth order Runge-Kutta (RK4). The size of the integration timestep is an important factor that controls the trade-off between performance and accuracy. Our ray tracer is similarly controlled by the global parameter δ (Sec. 4.5.2), which also leads to smaller steps producing more accurate results yet longer computation time. To highlight this trade-off, We conduct two types of comparisons: same quality comparisons, for which ray models are tuned to achieve the same level of errors and their performances are compared; and same speed comparisons, for which ray models are run at the same level of performance, and the resulted errors are compared.

We measure the accuracy of ray tracer by the two variables that are crucial for pressure computation: hit location of the ray at a particular range and the arclength of the entire trajectory. To quantify the difference between two sets of results rt and $\tilde{rt}$ tracing the same set of initial rays, we compute the relative error: $E_{rel} = \frac{\|rt - \tilde{rt}\|}{\|rt\|}$.

We select two media profiles as our test cases: downward-refracting atmosphere with sound speed gradient of 0.1, and the Munk profile used in ocean acoustics. Both profiles have ∇V^{-2} and ∇V in analytic form, which decouple the measurements of ray tracing accuracy from the accuracy of gradient estimation. We also limit our test to 2D cases with a flat ground as obstacle, for fair comparison with the majority of existing ray-based software. The performance improvements for 3D cases is expected to be greater, which is demonstrated in Section V.B.4.6.2.

Downward refracting atmosphere

The first media profile we tested is a linear sound speed profile that leads to a downward refracting atmosphere: $c(z) = c(0) + 0.1m/s$, where $c(z)$ is the effective sound speed at the height z , and $c(0)$ is the effective

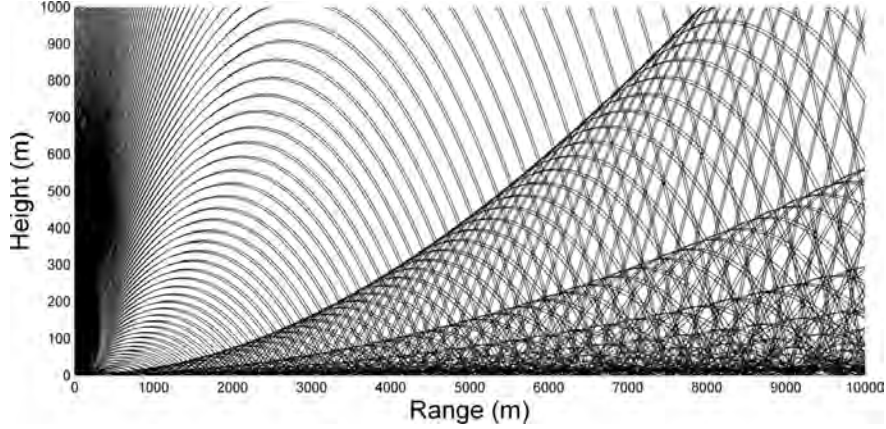


Figure 4.3: **Downward refracting atmosphere:** ray plot of 201 rays computed by the analytic ray curve tracer.

sound speed on the ground. We compute the ray tracing results for the range $0 - 10\text{km}$ and the height $1 - 1\text{km}$. The ray tracing results are gathered and compared at the 10km mark, at which the rays have been reflected up to 25 times. An example of the computed ray trajectories can be seen in Fig. 4.3, with a fan of 201 rays between the elevation angle -90° and 90° .

The same quality and the same speed comparison between our ray tracer and the RK4 ray integrator are listed in Table 4.1 and 4.2. To achieve the same error in ray tracing results, our ray tracer runs an order of magnitude faster than the numerical integrator. On the other hand, with the same running time the numerical integrator produce ray tracing results with higher levels of error. For this comparison only the computation of ray trajectories and arclength are included, while the pressure computation is excluded. However, the higher number of ray steps used by the numerical integrator indicate that with additional pressure-related computations at each step the performance improvements of our ray tracer will be even higher.

Rel. Error	1e-2	5e-3	1e-3	5e-4	2e-4	
Ray Curves	219	892	2,810	3,972	8,881	(segments/path)
RK4 Steps	467	2,330	4,881	24,588	49,158	(steps/path)

Table 4.1: **Same quality comparison.** A fan of rays are traced under a downward refracting atmosphere to a range of 10km . At each level of relative error in the ray tracing results, the average number of ray curve segments per propagation path is compared with the average number of Runge-Kutta 4 steps. Analytic ray curve is able to achieve same level of accuracy with much less tracing cost.

We also test the scalability of the ray models with increasing magnitude of media gradient in Fig. 4.4

# Ray steps/ RK4 segments	1×10^2	5×10^2	1×10^3	5×10^3	1×10^4
RT rel. error	1.31e-2	2.76e-3	8.04e-4	1.98e-4	4.33e-5
RK4 rel. error	4.35e-2	9.44e-3	4.92e-3	2.14e-3	9.12e-4

Table 4.2: **Same speed comparison.** A fan of rays are traced under a downward refracting atmosphere to a range of 10km. With comparable number of ray curve segments and Runge-Kutta 4 steps, our ray tracer (RT) is able to achieve lower levels of relative error in ray tracing results (both hit points and path length) than RK4.

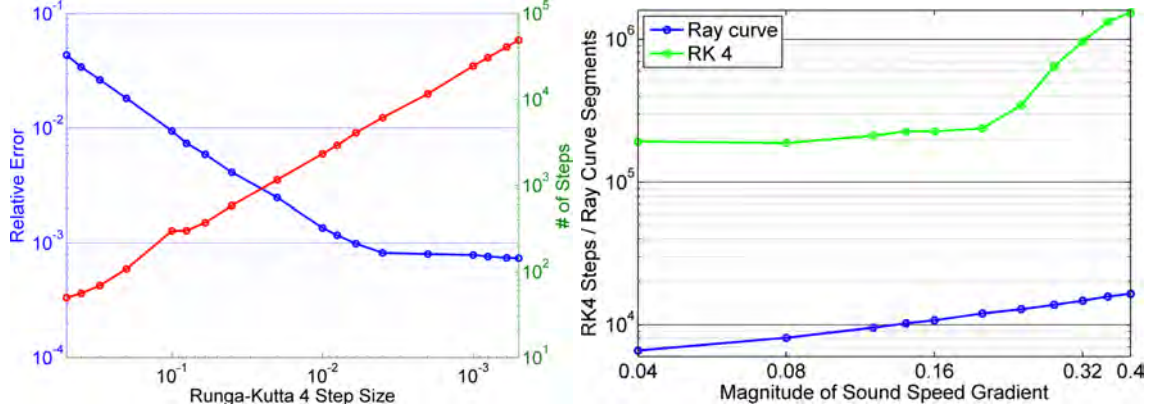


Figure 4.4: **Left:** Cost-accuracy trade-off of RK 4, to achieve lower error the step size of RK4 needs to decrease, which leads to exponential growth in computation time. **Right:** Cost with increasing media gradient grows as both RK 4 and ray curve tracer adopt smaller steps to keep the error below 2×10^{-4} , but RK 4 shows both higher cost and lower scalability with media gradient.

right. The media gradient is increased gradually from 0.04 to 0.4 with 0.04 increments, and the achieved relative error is kept at 2×10^{-4} by tuning the step size for RK 4 and the parameter δ for our ray curve tracer. Our ray tracer is shown to scale better with increased media variations, which we attribute to the adaptive segment size we used.

Munk profile

The second media profile we tested is the Munk profile commonly used in ocean acoustics:

$$c(z) = 15001.0 + 0.00737[x - 1 + \exp(-x)], \text{ where } c(z) \text{ is the effective sound speed at the depth } z, z \leq 5000m.$$

We compute the ray tracing results for the range $0 - 100km$ and the depth $0 - 5km$. An example of the computed ray trajectories can be seen in Fig. 4.7, with a fan of 21 rays between the elevation angle -13° and 13° . The relative errors of the numerical integrator results with decreasing step sizes is plotted in Fig. 4.6, contrasted with the same cost-accuracy analysis for the analytic ray tracer. Again, for this comparison only

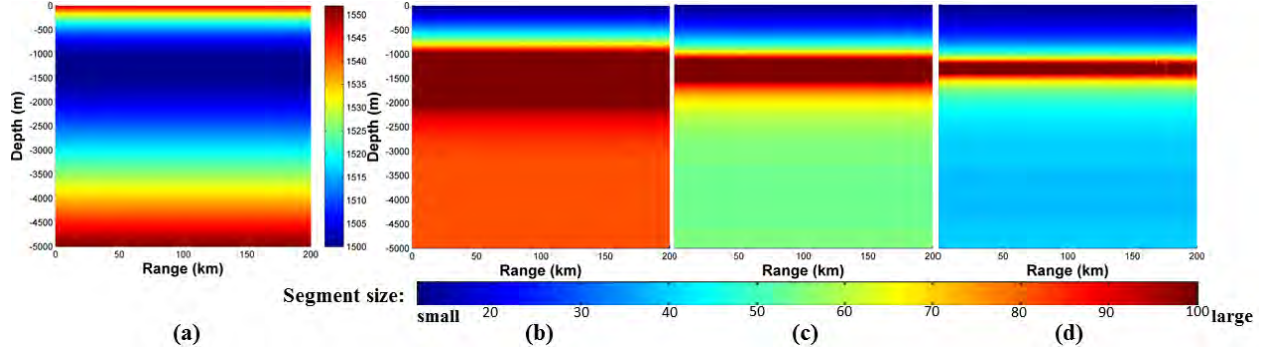


Figure 4.5: The range of validity for each segment of ray curve is controlled by the parameter δ and adapts to local media gradient. (a) Munk profile sound speed. (b,c,d) Segment sizes at spatial locations within the Munk profile visualized, with $\delta = 0.002, 0.001, 0.0005$ respectively.

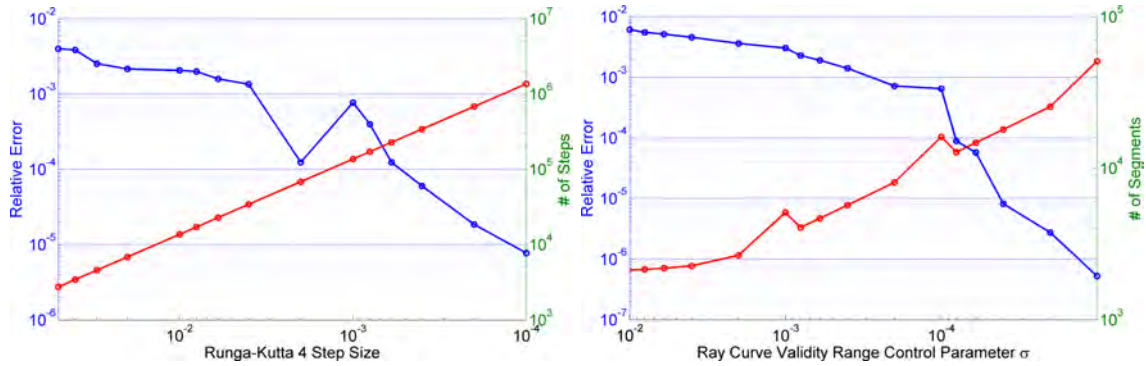


Figure 4.6: Munk profile with 200km range, ray tracing cost-accuracy trade-off of **Left:** Runga-Kutta 4 and **Right:** analytic ray curve.

the computation of ray trajectories and arclength are included, while the pressure computation is excluded.

For this profile we also show the adaptive segment size adopted by our ray tracer by visualizing in false color (Fig. 4.5) the size distribution in the domain. We trace rays spanning the launch angles between -90° and 90° for this visualization, and we contrast it with a false color visualization of the magnitude of gradient for Munk profile. The local coherence of media properties in the Munk profile is taken advantage of by our adaptive ray tracer to focus computation and improve efficiency. It is evident from the visualization that with a uniform step size the performance will reduce significantly even with the same analytic ray formulation.

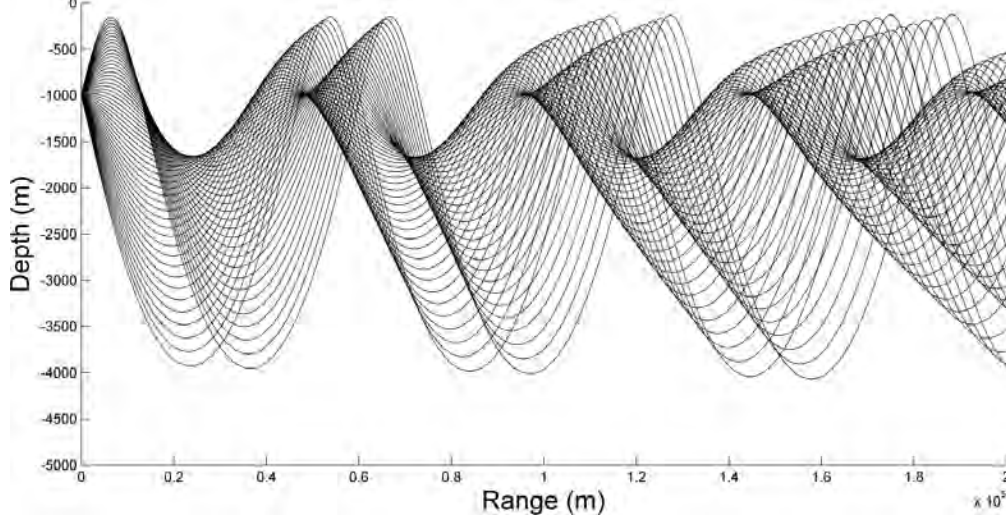


Figure 4.7: **Munk profile** characteristic ray plot computed by analytic ray tracer.

4.6.2 Outdoor Applications of Analytic Ray Tracer

To show the application of our ray tracer on fully general outdoor scenes, we generate a representative media profile from the stratified-plus-fluctuation model that is widely used in atmospheric acoustics[133]. The acoustic index of refraction in the atmosphere $n = c_0/c$, where c_0 is the reference sound speed, is modeled with a stratified component n_{str} and a fluctuation component n_{flu} , so that $n = n_{str} + n_{flu}$. The stratified component follows a logarithmic profile of the altitude z : $n_{str}(z) = c_0/(c_0 + b \ln(\frac{z}{z_g} + 1))$, with parameters n_0 , b , and z_g . c_0 is the sound speed at the ground surface, taken as the reference sound speed, and z_g is the roughness length of the ground surface. Typical values for parameter b are 1 m/s for a downward-refracting atmosphere and -1 m/s for an upward-refracting atmosphere. The fluctuation component models the three-dimensional random temperature and wind speed turbulence in the atmosphere. The value at position $\mathbf{x}$ can be computed as $n_{flu}(\mathbf{x}) = \sum_i G(\mathbf{k}_i) \cos(\mathbf{k}_i \cdot \mathbf{x} + \varphi_i)$, where $\mathbf{k}_i$ is the wave vector describing the *spatial* frequency of the fluctuation, φ_i is a random angle between $[0, 2\pi]$, and $G(\mathbf{k}_i)$ is a normalization factor.

The Desert and Christmas benchmarks represent large-volume outdoor acoustic scenes that have complex surface geometry (e.g. varying terrains and buildings). The surface primitive count and the volumetric expanse of the two scenes are listed in Table 4.3. Both scenes are visualized in Fig. 4.8, and we include illustrative 3D ray trace for both upward refractive and downward refractive media conditions.



Figure 4.8: **Outdoor benchmarks** Desert and Christmas village. Example ray paths are shown with upward and downward refraction. The media gradient is exaggerated and the ray path number is kept very small for illustration purpose. See Table 4.3 for scene stats and actual performance numbers of tracing 10k rays up to 3 orders of reflections.

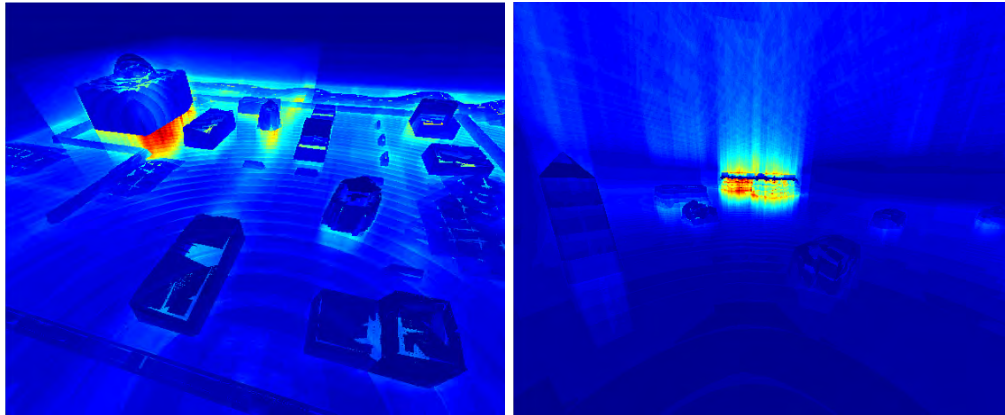


Figure 4.9: **Spatial visualization of ray tracing cost:** brighter color indicates higher intersection costs. It is shown that the intersection computation concentrates in areas of dense surface geometry and around silhouette of obstacles, which illustrates that the spatial culling of ray-surface intersections using BVH is effective. **Left:** Desert **Right:** Christmas village.

Scenes	# surf.	# seg per path	# BV int. per seg	# surf. int. per seg	frame time
Desert ($200 \times 200 \times 50$ m ³)	5,000	26.73	31.88 (94%)	2.32 (1.47%)	446 ms
Christmas ($120 \times 80 \times 50$ m ³)	16,000	44.45	34.31 (96.16%)	2.08 (1.22%)	917 ms

Table 4.3: **Breakdown of ray tracing time:** the computation of closed-form ray trajectory and closed-form evaluations of ray properties for each segment takes very little computation (less than 0.001% of total frame time) and thus is omitted here. The percentage that each kind of computation takes relative to the total frame time is included in parenthesis, together with the number of ray-BV and ray-surface intersections.

For each benchmark we trace 10K rays for up to three orders of reflections, and the overall ray tracing performance as well as a breakdown of the running time are reported in Table 4.3. We can see that the ray intersection computation dominates the total ray tracing cost, and the majority of the intersections are computed with Bounding Volumes (BVs). The intersection numbers with surface primitives are kept very low for each path, which shows that the spatial culling is effective by traversing the BVH. A detailed visualization in Fig. 4.9 demonstrates this point further that the intersection cost is concentrated in the vicinity of obstacles, whereas for the majority of the scenes the ray intersections are resolved with relatively low cost. The analytic ray tracer achieved close to interactive performance for both benchmarks, and it scales well with the complexity in the media and the boundaries.

4.7 Discussions

Our ray tracer can be extended in ways including but not limited to:

- coupling with flow models and physically realistic media profiles as inputs;
- augmenting GA methods with capability to handle inhomogeneous media;
- combining with other methods to form hybrid propagation algorithms.

The analytic ray curve tracer takes general media profile as input, as long as the profile can be sampled for sound speed and gradient at any spatial locations. The input can be an analytic function (e.g. Monin-

Obukhov similarity model[112, 141]) or comes on a set of discrete sample points (e.g. a grid input). Numerical methods have been coupled with a background flow model (e.g. linearized Euler equations[21, 182]), so that detailed media conditions output by the flow model serves as input to propagation simulation. Our ray tracer is also compatible with such a coupling.

GA methods such as path tracing[136] and sonal mapping[79] model complex surface interactions including diffuse reflections and surface BRDFs, but they assume homogeneous media and straight line propagation in between surfaces. Our method can substitute the trajectory computation in between surfaces to account for inhomogeneous media, while working seamlessly with the surface handling part. Such augmented GA solutions can be applied on simulating large indoor space such as auditoriums or dense outdoor urban areas.

The efficiency of our method makes it a good candidate to form hybrid schemes with wave-based numerical models. One such possibility is to couple the ray curve tracer with ESM (Equivalent Source Method) with a spatial and frequency division, similar to Yeh et al.[179]. The ray curve tracer is focused on high frequency and long range propagation in the sparse space in between complex scatterers. Alternatively, our ray tracer can be used for fast initial evaluations of the sound field for a wide area, and costly numerical methods such as FDTD can be applied economically to areas of significant field complexity[128, 129, 18, 163].

The analytic ray curve tracer has several limitations. The foundation of this method is geometric ray model, inherently a high frequency approximation as opposed to a full wave solution. The efficiency improvement from tracing analytic ray curves depend on media coherence, and it will degenerate to the performance of numerical ray integration if the media is chaotic without much spatial coherence.

4.8 Conclusion

This paper addresses the challenge of efficient sound propagation in inhomogeneous, moving media and large outdoor scenes with complex boundary surfaces. Our method adapts the geometric ray models to inhomogeneous media, and the key performance improvement upon rectilinear rays is achieved by tracing analytic ray curves based on local media gradients as primitives. Segments of ray curves are computed

by sampling the media gradient on-the-fly, accounting for both inhomogeneous and moving media without the need to pre-compute explicit cell structures. Acceleration based on BVH is readily adapted to speed up surface intersections of the ray curves, enabling logarithmic scaling with scene complexity and allowing dynamic scenes. The improved performance of this method is demonstrated in comparison to rectilinear ray models, as well as on scenarios that are prohibitively expensive with existing methods. In addition to being an efficient stand-alone tool, this method also has the potential of complementing other GA and numerical methods.

Chapter 5

Outdoor Sound Propagation with Analytic Ray Curve Tracer and Gaussian Beam

5.1 Overview

Outdoor sound propagation benefits from algorithms that are computationally efficient in handling inhomogeneous media, complex boundary surfaces, and large spatial expanse. One recent work [101] proposed a ray tracing method using analytic ray curves as tracing primitives, which retains the efficiency of the ray model in handling surface interactions, while overcoming the performance penalty incurred by inhomogeneous media on rectilinear rays. In this paper, we develop an algorithm for sound field computation of outdoor propagation that combines this fast ray tracer and the Gaussian beam model. Validation of the algorithm is performed by comparison to published results on benchmarks in atmospheric and ocean acoustics. The application of this algorithm is demonstrated on a scene with terrains and buildings of realistic complexity and a set of media profiles. Our algorithm is able to compute characteristic sound field for fully general

media profiles and complex 3D scenes at close-to-interactive speed.

5.2 Introduction

Sound propagation in outdoor environment[133, 11, 73], including atmospheric and underwater acoustics, deals with spatially varying as well as moving media. Empirical models, real-world measurements, and computational flow simulation can all supply the initial media profiles as input to sound propagation. Obstacles' shape and material properties play an important role as well, especially for scenes with complex terrains or area with dense man-made structures. As data that describe the media conditions and the scene obstacles become increasingly available with ever richer details, methods for outdoor sound propagation needs to be able to account for the full scale of those data.

Existing methods face many challenges in handling such complexity, they either make assumptions that preclude a fully general scene, or they become prohibitively expensive with large and complex scenes. Models such as Fast Field Program (FFP), Parabolic Equation (PE), and normal modes [133, 11, 73] fall in the former category, while general numerical methods like Finite Difference Time Domain (FDTD) [24], Finite Elements Method (FEM)[153], and Boundary Element Method (BEM)[26] belong to the latter. For outdoor settings many methods are focused on modeling the media complexity, while obstacles are generally assumed to be sparse and their complexity limited in the investigations.

Geometric acoustics (GA)[115] methods like ray models are known for their efficiency in handling boundary surfaces under the assumption of homogeneous media and rectilinear paths. A recent work[101] has also attempted to accommodate inhomogeneous media by tracing parabolic ray curves as primitives, which significantly accelerates path computation. On the other hand ray models suffer from artifacts in caustics and shadow zones when computing fields, and models such as Gaussian beam[161] performs better in this regard. However, the performance of Gaussian beam can be hindered by the underlying numerical ray integration that remains slow for inhomogeneous media.

We combines the performance of the ray curve tracer[101] and the accuracy of the Gaussian beam[161]

into an algorithm for outdoor sound propagation. In particular:

1. We compute analytic solutions to on-ray pressure as well as paraxial fields based on the parabolic ray formulation (Sec. III), which leads to efficient field computation that matches the efficiency of the path computation.
2. We combine Gaussian beam with the analytic ray formulation, and we validate the approach on simple benchmarks by comparing to published results generated by established techniques in both atmospheric and oceanic settings (Sec. IV).
3. We apply the algorithm on a complex outdoor scene with realistic configurations, which demonstrate its efficiency in generating characteristic sound field (Sec. V).

Overall we provide a validated solution to outdoor sound propagation that augments a fast analytic ray tracer with equally fast analytic field computations. This algorithm takes general media and scene input and computes the full 3D sound field at close-to-interactive speed, which can be useful for a wide range of outdoor sound applications (Sec. VI5.7).

5.3 Prior Work

Outdoor sound propagation has been studied extensively in the context of underwater[73] and atmospheric acoustics[133]. Here we first review numerical methods that provide full wave solutions, including various hybrid schemes aiming at reducing the high computation cost of those wave-based methods. Secondly ray-based methods are reviewed with their advantages and limitations for outdoor scenarios.

5.3.1 Wave-based Methods

Early methods include Fast Field Program (FFP)[39] and Parabolic Equation (PE)[102, 50], both provide frequency domain full wave solutions that account for the inhomogeneous media, but with simplifying assumptions. Given scenarios that meet those assumptions, these early models have been thoroughly validated [65, 180, 127, 142, 14] and often serve as reference solutions to test other models.

Among the more general numerical methods that handle arbitrary media and obstacles, the Finite Difference Time Domain (FDTD) method is widely used that solves the linearized Euler equation[110, 175]. FDTD have been coupled with background flow simulation[21, 182], applied with turbulence[176, 34], various ground conditions[36], terrains[61, 37, 41], and complex obstacles[109, 3, 42]. As a time domain method it is also used for pulse propagation[90, 3]. The disadvantage of FDTD is mainly its limited scalability with the domain volume or the frequency, making it prohibitively expensive for large-scale broadband simulation. FDTD has been limited to low frequency for it to be practical for wide area assessment[60]. Other frequency domain methods such as FEM[103, 164] face similar challenges in scaling with volume and frequency.

Hybrid methods were developed that use FDTD in a confined area and apply PE to propagate over long range with relatively sparse space[128, 129, 18]. Alternatively, methods such as BEM[26] or Equivalent Source Method (ESM)[105] were employed to limit the computation either to boundary surfaces, or to volumes that bound scatterers tightly. However, BEM needs to be coupled with specialized Green function for refractive media[119, 177], and it scales poorly with surface area and frequency. ESM was coupled with ray models[179] to handle large domains, but this hybrid method does not scale well with the number or complexity of scatterer objects. Methods such as Pseudo Spectral Time Domain[67, 66] and Transmission Line Matrix (TLM)[64, 52, 13] are more efficient, but they are still fundamentally limited by the cost of discretizing a large domain. A more recent method, Adaptive Rectangular Decomposition (ARD)[123, 99], took advantage of the analytic solution of wave equation in a rectangular domain, but it requires constant sound speed within each spatial subdivision, which is not easily adapted to a general media profile.

5.3.2 Geometric Acoustics Methods

Geometric acoustics (GA) methods[115] are widely used in room acoustics[85], to handle the high order surface interactions under the valid assumption of a homogeneous media. Examples include the image source method[7, 22], ray tracing[167, 149], frustum tracing[148], beam tracing[46, 32], sonal mapping[19, 38, 79], path tracing[136], and radiosity[160].

The ray models have also been applied to inhomogeneous media[28, 73] by numerically integrating the

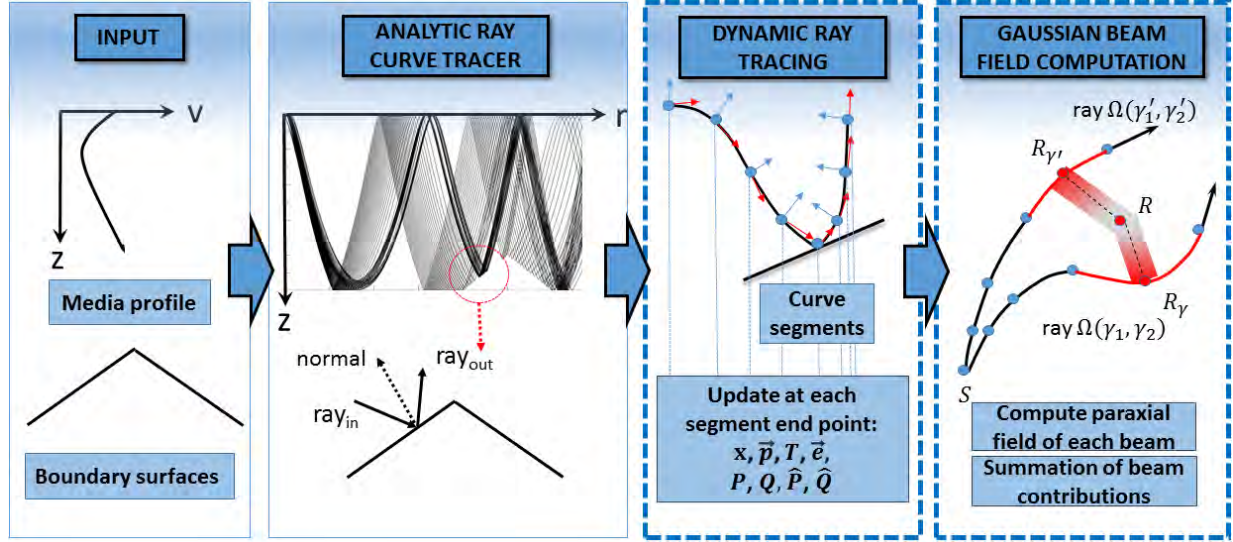


Figure 5.1: **Algorithm overview.**

ray equations. A sparse set of rays can be efficiently traced to plot out the sound propagation paths, but long range propagation in this manner becomes expensive. When the ray models are used to compute the pressure field in addition to the propagation paths, they are known to have issues in the caustic zones and in the shadow zones. With some ray models another source of inefficiency in inhomogeneous media is the need to locate eigen rays. The Gaussian beam approach[161] was developed in seismology and applied on underwater[16] and atmospheric[48] acoustics, which improves the accuracy in caustics and shadow zones and eliminates the need to locate eigen rays. However, when the underlying path still consists of rectilinear ray segments The performance of the Gaussian beam method is limited by the numerical integration step sizes. A recent work[101] achieved significant performance improvement by replacing the rectilinear ray with parabolic ray curve. Adaptive segment sizes based on on-the-fly media sampling as well as acceleration structure that bounds surfaces and ray curves leads to further speedup. We give an overview of this ray tracer in the next section, before we introduce our algorithm that combines this ray tracer with the Gaussian beam.

5.4 Algorithm

An overview of our outdoor sound propagation algorithm is illustrated in Figure 5.1. We built upon a very efficient ray tracer that computes propagation paths made up of segments of analytic curves (Sec. A5.4.1). Given the output from the ray tracer, we compute a set of additional variables per ray segment that are subsequently used for pressure computation both on the ray paths and in the paraxial regions. Those variables are computed by analytic evaluations of constant cost, which extends the efficiency of paths computation to pressure computation. The algorithm for on-ray pressures and ray paraxial fields are given in Sec. 5.4.2 and 5.4.3, respectively. The mathematical derivation that leads to this algorithm is a special case of the more general ray theory framework discussed in depth in Červený's comprehensive work[161]. They can be found in the Appendix to this paper for completeness of the presentation.

5.4.1 Analytic Ray Curve Tracer

Given a media profile with spatially varying sound speed $V(x_i)$, the *slowness vector* is the spatial derivative of the travel time field T , $\vec{p} = \nabla T$. Assuming a constant gradient of V^{-2} : $V(\mathbf{x})^{-2} = A_0 + \vec{A} \cdot \mathbf{x}$ within a certain range, the trajectory is a parabolic curve that lies in the plane with the normal of $\vec{A} \times \vec{p}_0$ (the *ray plane*, see Figure 5.2). The intersection between the parabolic ray curve and planar surfaces can be solved analytically, and key properties such as position $\mathbf{x}$, tangent direction $\vec{t}$ which coincides the direction of $\vec{p}$, $\vec{p}$ and T can be computed for any point along the ray by analytic evaluations of constant cost, as long as those properties are known at one point along the ray (e.g. the source).

This analytic ray formulation enables a ray tracer that computes propagation paths in a general media consisting of consecutive segments of parabolic curves. The ray properties are computed only at the end points and are continued across segments. In spatially coherent media, the assumption of constant ∇V^{-2} generally holds for a range larger than the assumption of constant V , enabling the ray curve tracer to advance in longer segments than rectilinear ray tracer, which is one of the key sources of performance improvements.

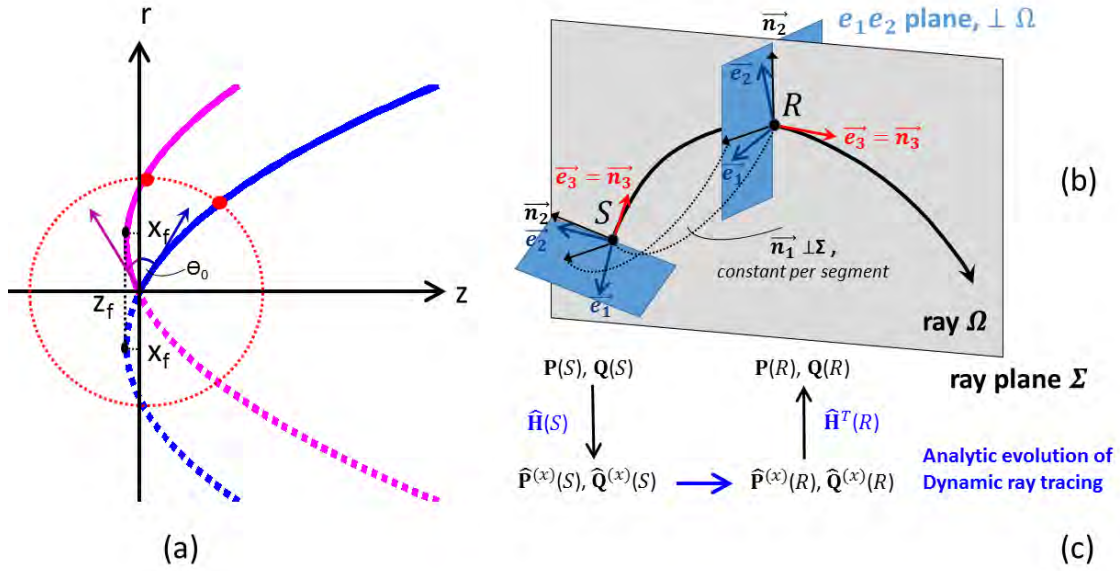


Figure 5.2: **(a)** the analytic ray curve in ray-plane. Rays of two different θ_0 are drawn in blue and magenta, respectively. (x_f, z_f) is the vertex of the parabola. Assuming a locally constant ∇V^{-2} , of which the range of validity determines the extent of the ray segment. The two red dots mark the end points of ray segments, where the rays exit from this sphere of validity (Sec. A). **(b,c)** analytic evolution of $\mathbf{P}$, $\mathbf{Q}$ are performed for each segment by transforming into and evolving in Cartesian coordinates before transforming back to ray-centered coordinates (Sec. B).

5.4.2 Dynamic Ray Tracing

Dynamic ray tracing keeps track of how a set of derivatives in ray-centered coordinates progress among propagating rays, and the derivatives subsequently are used in computation of pressure and travel time along the ray. For the parabolic ray curve, the derivatives in Cartesian coordinates can be evolved analytically between any two points within a ray curve, and the transformation between Cartesian coordinates and ray-centered coordinates can also be evolved analytically, thereby we achieve analytic evolution of the derivatives in ray-centered coordinates (See Fig. 5.2(b,c)).

First we define the coordinates involved. At any point along a particular ray Ω , the *ray-centered coordinates* q_1, q_2, q_3 is defined with origin at that point. Ray Ω is the q_3 -axis of the system, the q_1 -axis and q_2 -axis are taken to be perpendicular to q_3 -axis, and mutually perpendicular. The unit basis of the ray-centered coordinates are denoted by $\vec{e}_1, \vec{e}_2, \vec{e}_3$. The 3×3 transformation matrix from the ray-centered coordinates q_k to the Cartesian coordinates x_i is denoted by $\hat{H}$, $H_{ik} = \partial x_i / \partial q_k = \partial q_k / \partial x_i = e_{ki}$, $i, k = 1, 2, 3$.

The transformation matrix from the Cartesian coordinates to ray-centered coordinates is thus $\hat{H}^{-1} = \hat{H}^T$.

The derivatives we seek in dynamic ray tracing captures how a set of rays' spatial relationship with each other changes as they travel through a media profile. Consider an orthonormal system of rays starting from a sound source, parameterized by two ray parameters γ_1, γ_2 , taken here as the take-off azimuth ϕ_0 and elevation i_0 angles. The two 2×2 matrices of derivatives in the ray-centered coordinates $\mathbf{Q}$ and $\mathbf{P}$ are defined with elements $Q_{IJ} = (\partial q_I / \partial \gamma_J)_{T=const}$, $P_{IJ} = (\partial p_I^{(q)} / \partial \gamma_J)_{T=const}$, $I, J = 1, 2$. Correspondingly, $\hat{\mathbf{Q}}^{(x)}$ and $\hat{\mathbf{P}}^{(x)}$ in Cartesian coordinates are defined with elements $Q_{iJ}^{(x)} = (\partial x_i / \partial \gamma_J)_{\sigma=const}$, $P_{iJ}^{(x)} = (\partial p_i^{(x)} / \partial \gamma_J)_{\sigma=const}$, $i = 1, 2, 3$, $J = 1, 2$. We have $\hat{\mathbf{Q}}^{(x)} = \hat{\mathbf{H}}\mathbf{Q}$, $\hat{\mathbf{P}}^{(x)} = \hat{\mathbf{H}}\mathbf{P}$.

Under the assumption of constant ∇V^{-2} , the $\mathbf{P}$, $\mathbf{Q}$ are evolved between two points within the same ray curve segment by analytically evolving $\hat{\mathbf{Q}}^{(x)}$, $\hat{\mathbf{P}}^{(x)}$, and $\hat{H}$, in the steps shown in Fig. 5.2(b,c). $\mathbf{J} = \det \mathbf{Q}$ is the *ray Jacobian*, used in the geometric spreading term $|\mathbf{J}|^{1/2}$, which in turn relates the pressure amplitude at point s to the pressure amplitude at the source s_0 by:

$$P^{ray}(s) = [\frac{\rho(s)V(s)\mathbf{J}(s_0)}{\rho(s_0)V(s_0)\mathbf{J}(s)}]^{1/2} P(s_0), \quad (5.1)$$

where ρ is the density of the media. Both $\mathbf{P}$ and $\mathbf{Q}$ in ray-centered coordinates are also used in computing the paraxial field, in Sec. C.

5.4.3 Field Computation with Gaussian Beam

Given the dynamic ray tracing results for points along the ray paths, the Gaussian beam provides a model to approximate paraxial fields in the vicinity of the ray paths, which involves computing both paraxial travel time and pressure amplitude.

Paraxial Travel Time

Looking at the definition of $\mathbf{P}$ and $\mathbf{Q}$ which we compute analytically by dynamic ray tracing, we introduce the 2×2 matrix

$$\mathbf{M} = \mathbf{P}\mathbf{Q}^{-1}, \quad M_{IJ} = (\partial p_I^{(q)} / \partial q_J)_{T=\text{const}}, \quad I, J = 1, 2. \quad (5.2)$$

Recall the slowness vector $\vec{p}$ is the first derivative of T , M is therefore the second derivative of T with respect to ray-centered coordinates. For a point R' in the vicinity of a ray Ω , the paraxial travel time at R' can be computed given the T at a point R on Ω :

$$T(R', R) = T(R) + \frac{1}{2} \mathbf{q}^T(R') \mathbf{M}(R) \mathbf{q}(R'), \quad \mathbf{q} = (q_1, q_2)^T, \quad (5.3)$$

when $\Omega^\perp$ is the plane perpendicular to Ω that passes R' , and point R is the intersection of the ray Ω and $\Omega^\perp$. The derivatives of T can also be approximated in Cartesian coordinates, in which case any point R_γ on the ray that is close to R' can be selected, and computing $\Omega^\perp$ and R are not necessary (see Appendix for details).

Paraxial Pressure Amplitudes

There are multiple ways compute paraxial pressure amplitudes and the Gaussian beam model is one of them. It computes a paraxial amplitude centered on the ray with a Gaussian drop-off, which is achieved by allowing the matrix $\mathbf{M}$ to be complex: $\mathbf{M} = \text{Re}(\mathbf{M}) + \text{Im}(\mathbf{M})$. $\text{Im}(\mathbf{M})$ is chosen to be positive definite, so that

$$p^{\text{beam}}(R') = P^{\text{ray}}(R) \exp[-i\omega(-T(R) - \frac{1}{2} \mathbf{q}^T(R') \mathbf{M}(R) \mathbf{q}(R'))] \quad (5.4)$$

$$= P^{\text{ray}}(R) \exp[-i\omega(-T(R) - \frac{1}{2} \mathbf{q}^T(R') \text{Re}(\mathbf{M}(R)) \mathbf{q}(R'))] \quad (5.5)$$

$$\times \exp[-\frac{1}{2} \omega \mathbf{q}^T(R') \text{Im}(\mathbf{M}(R)) \mathbf{q}(R')]. \quad (5.6)$$

Matrices with suffix a ($\mathbf{M}^a$, $\mathbf{P}^a$, $\mathbf{Q}^a$) represents the matrices of the *actual* field (Eq. 5.2).

The contributions of Gaussian beams are then summed up by integral superposition. As shown in the rightmost block of Fig. 5.1, we gather all segments of ray curves that pass in the vicinity of a given field point, and we compute the paraxial pressure amplitude and travel time from each segment and sum those

up. For field computation that involves large volumes of points, we spread some of the costs of locating vicinity ray segments by reversing the process and distribute pressure contribution from each beam to the field points it covers.

5.5 Validation

We validated our algorithm on two benchmark scenes of atmospheric and oceanic sound propagation. The benchmarks are selected from literature with published results computed by alternative methods. Simple scenarios are modeled in those benchmarks, of which the sound field characteristics can be easily interpreted, and for which validated results computed by multiple methods are available for direct comparisons. After the validity of our method is established with these benchmarks, application on a real-world oriented scene with a more complex media conditions is presented in Section V. We compute and visualize the sound field in terms of the transmission loss (TL) for all benchmarks, which is defined as:

$$TL = 20 \log \frac{(Pressure \text{ at a field point})}{(Pressure \text{ of free field at 1m from source})}.$$

5.5.1 Benchmark A

Inhomogeneous atmosphere, flat ground with impedance

In Attenborough et al.[12], a set of benchmark cases for outdoor sound propagation is proposed, and results generated by a range of methods show good agreements, including FFP, PE, normal modes, ray and beam tracing. The boundary surface in the scene is a flat ground with impedance, while the media is inhomogeneous with three different profiles.

Media profile: sound speed $c(x)$ at spatial location x with height $z(x)$ is given by

case 1: downward refractive $c(x) = 343 + 0.1 * z(x)(m/s)$;

case 2: upward refractive $c(x) = 343 - 0.1 * z(x)(m/s)$;

case 3: duct condition: case 1 for $z(x) < 100m$, case 2 for $100m < z(x) < 300m$,

and constant c for $z(x) > 300m$.

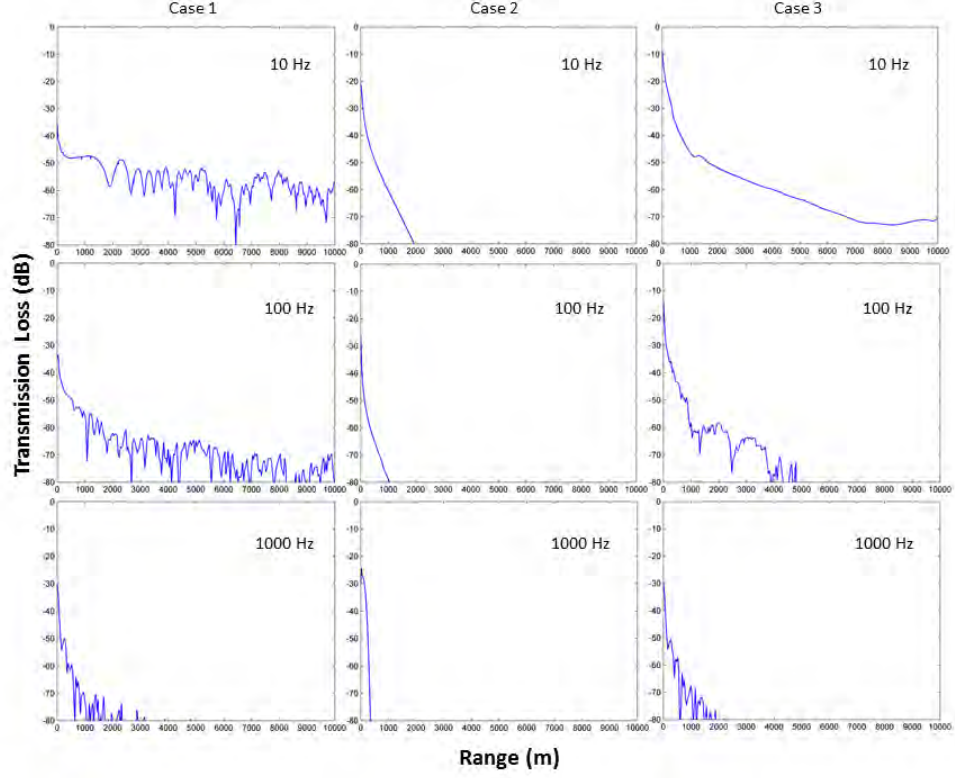


Figure 5.3: **Benchmark A Range-TL plot:** Source height $h_s = 5m$, receiver height $h_r = 1m$, range $10km$. Columns contain results for media profiles case 1, 2, and 3, while rows contain results for frequency 10, 100, and 1000 Hz. Comparison with Fig. 12-14[12].

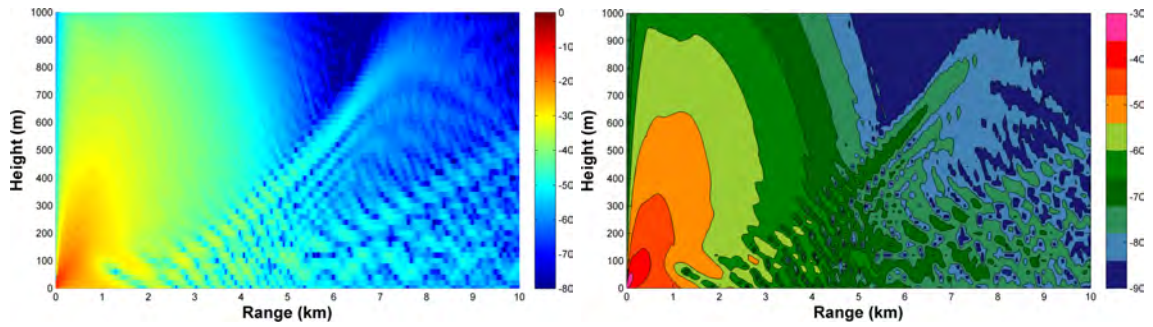


Figure 5.4: **Benchmark A 2D field:** Source height $h_s = 5m$, receiver height $h_r = 1m$, frequency 10Hz. 2D vertical field of height up to $1km$ and range up to $10km$ is visualized on the left, and the corresponding contour plot is shown on the right for comparison with Fig. 15[12].

Ground impedance: A four parameter model is used to compute the impedance of the flat ground, and the same parameters from Attenborough et al.[12] are used.

Results: As shown in Fig. 5.3 and 5.4, our algorithm is able to replicate the interference pattern for this set of media conditions at three different frequencies. 1D range plot and 2D vertical field of resulting TL is included for direct comparison with figures in [12]. The number of rays required is as low as 21 rays to achieve the results, more rays can be traced to compute asymptotically more accurate pressure fields.

5.5.2 Benchmark B

Munk profile with conical seamount

We validate our algorithm on an underwater benchmark with the standard Munk profile and a conical seamount as bathymetry. Published results computed by normal modes can be found in prior work[93].

Media profile: Munk profile with depth $0 - 5000m$.

Bathymetry: conical seamount located $100km$ from the source, radius of the base $20km$, and two different heights $1000m$ and $3800m$.

Bottom impedance: fluid half space with compressional speed of $2000m/s$, density of $1g/cm^3$, and attenuation of $0.1dB/\lambda$.

Results: The vertical 2D TL field is visualized for a range up to $200km$ and a depth up $5km$ in Fig. 5.5. Our ray tracer successfully replicate the characteristics of the field for this underwater benchmark with extensive volume, as compared with the normal mode results available in prior work[93].

5.6 Application on General Outdoor Scene

5.6.1 Scene configuration

We generate a general media profile based on an empirical model of the atmosphere [133]. The acoustic index of refraction in the atmosphere $n = c_0/c$, where c_0 is the reference sound speed, is modeled with a stratified component n_{str} and a fluctuation component n_{flu} , so that $n = n_{str} + n_{flu}$. The stratified

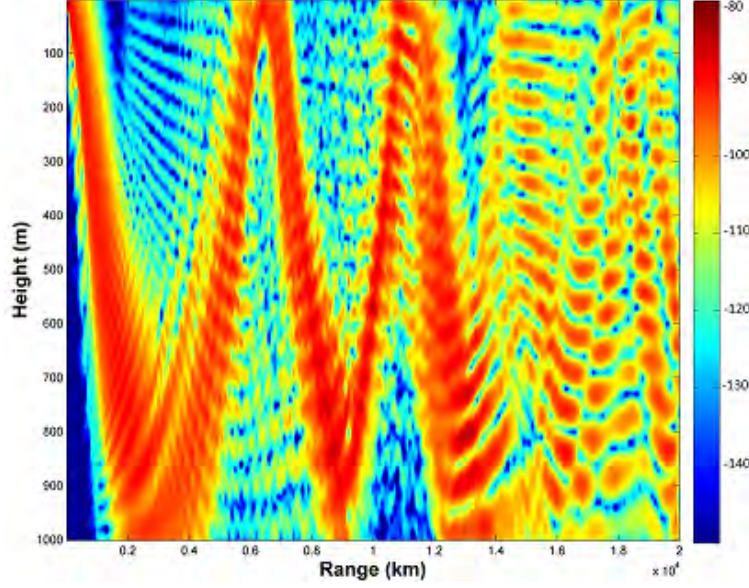


Figure 5.5: **Benchmark B 2D field:** Source depth 1000m, frequency 50Hz. 2D vertical field of TL is visualized for depth up to 5km and range up to 200km, compared to Fig. 9[93].

component follows a logarithmic profile of the altitude z : $n_{str}(z) = c_0 / (c_0 + b \ln(\frac{z}{z_g} + 1))$, the parameter z_g is the roughness length of the ground surface, and typical value for b are 1 m/s for a downward-refracting and -1 m/s for upward-refracting atmosphere. The fluctuation component at position $\mathbf{x}$ can be computed as $n_{flu}(\mathbf{x}) = \sum_i G(\mathbf{k}_i) \cos(\mathbf{k}_i \cdot \mathbf{x} + \varphi_i)$, where $\mathbf{k}_i$ is the wave vector describing the *spatial* frequency of the fluctuation, φ_i is a random angle between $[0, 2\pi]$, and $G(\mathbf{k}_i)$ is a normalization factor.

We use an artificial scene consisting of undulating terrains that depicts a reservoir and buildings. A wireframe rendering of the scene, and the two sound source locations marked by green and red dots can be found in Fig. 5.6(a). The scene has a physical dimension of $220m \times 150m \times 50m$, and is represented by 4,000 triangular surface primitives. Our algorithm can simulate propagation for any scenes that can be modeled or scanned into similar surface representations as demonstrated here.

5.6.2 Results

The diurnal changes in the atmosphere typically lead to an upward refractive condition during the day, and a downward refractive condition at night. We generate the sound field for the source on the slope of the

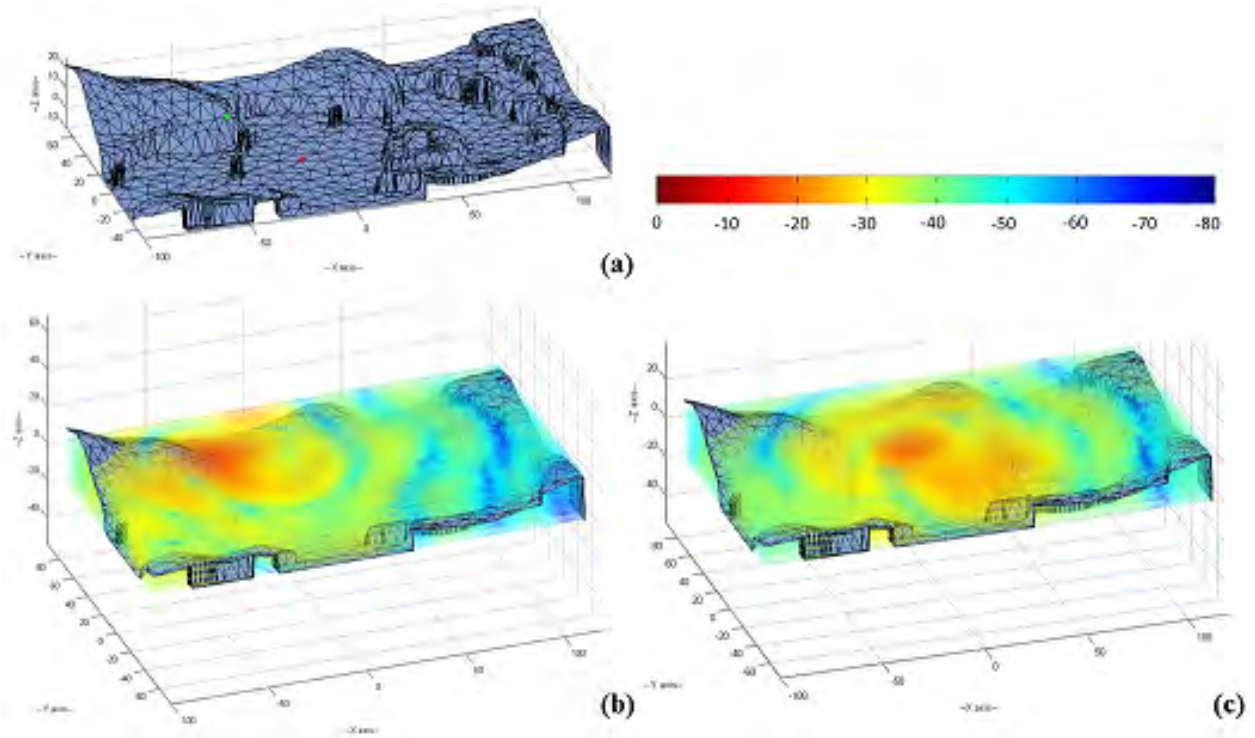


Figure 5.6: **Outdoor scene and field results for two source locations:** (a) Wireframe rendering of the Reservoir scene, green dot represents a sound source located on the slope, and the red dot represents a sound source in the valley. (b) Slices of sound TL field visualized for the green source. (c) Slices of sound TL level visualization for the red source. Frequency 10Hz.

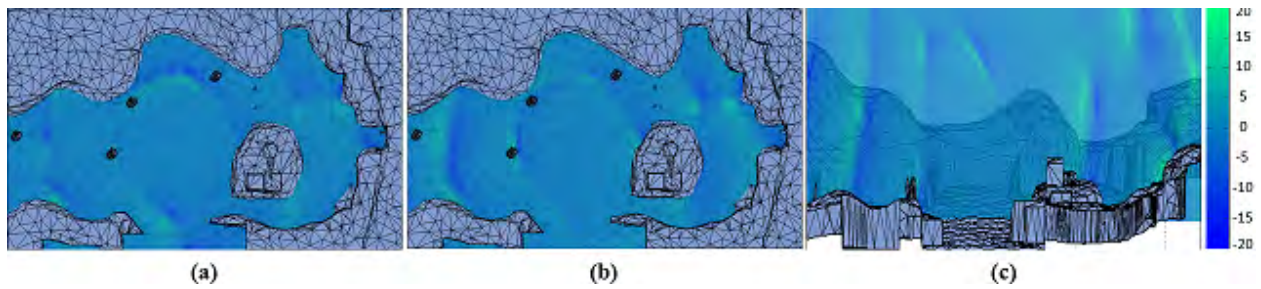


Figure 5.7: **Vector wind for source in the valley:** (a) Difference in horizontal field of sound pressure level between a north and a sound wind. (b) Difference in horizontal field of sound pressure level between an east and a west wind. (c) Difference in vertical field of sound pressure level between an east and a west wind. Frequency 10Hz.

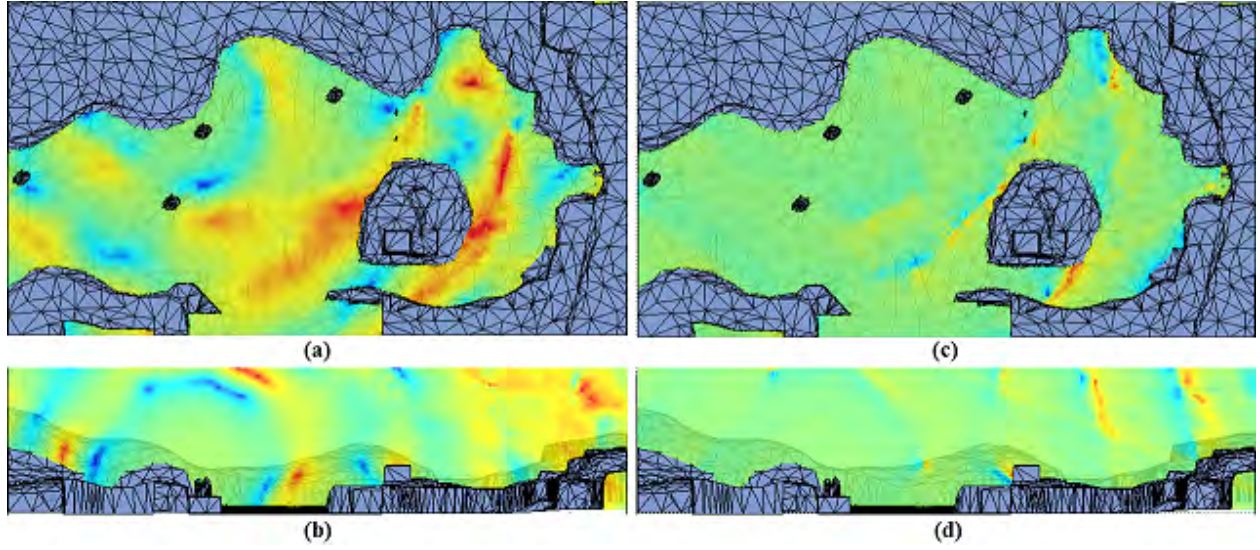


Figure 5.8: **Source on the slope:** (a) Relative sound pressure level upward vs downward refraction, vertical field, (b) horizontal field. (c) Relative sound pressure level up wind vs down wind propagation, vertical field, (d) horizontal field. Frequency 10Hz.

reservoir, and a source in the center of the valley, for upward and downward refractive conditions respectively, and visualize the difference (Fig. 5.8(a,b) and Fig. 5.9).

Wind plays an important role in atmospheric sound propagation, creating extra variations in the sound speed profile, and interacts with physical obstacles that further complicates the sound field. For the sound source on the slope, we also simulate the sound field for up-wind and down-wind conditions (Fig. 5.8(c,d)), which yields similar patterns as the relative difference between upward and downward refractive media.

With our ray tracer we also account for moving medium with vector wind field, and we show this

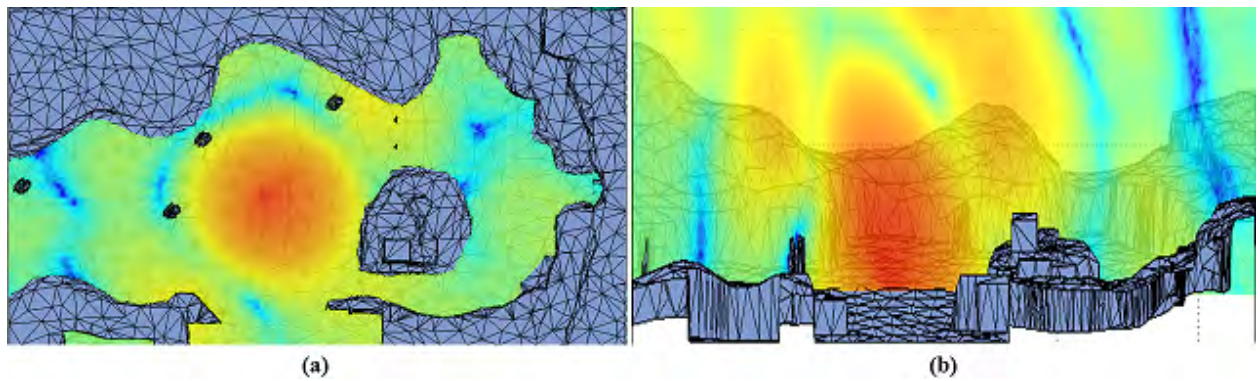


Figure 5.9: **Source in the valley:** Relative sound pressure level upward vs downward refraction, (a) horizontal and (b) vertical field that passes the source location. Frequency 10Hz.

capability with the sound source in the valley, and visualization of the difference in sound field between a west and a east wind, and between a north and a south wind, respectively (Fig. 5.7).

For the benchmark scene tested here the sound field displays complex characteristics resulting from the interaction of sound wave with the media and the obstacles. Our propagation algorithm can produce this complex sound field close to interactively, so that observations can be made while varying scene geometry, media profiles, or frequency. We show a stack of slices of the sound field in Fig. 5.6 for visualization purposes, while the full 3D sound field is computed by our algorithm. Details on performance of our method is can be found by referring to [101].

5.7 Discussion

This algorithm is complementary to many existing sound propagation techniques and can be extended or combined in multiple ways. With the analytic ray curve tracer[101] as a component, our method inherits its potential extensions, including augmenting GA methods, hybrid method based on frequency and spatial subdivision similar to Yeh et al.[179], and fast wide area assessment using the ray tracer with concentrate numerical methods in areas of interests. Furthermore, extensions[89, 121] can be built similar to other ray models to account for turbulence. It is also possible to accommodate sound sources other than a point source, using the techniques of Gaussian beam expansion of complex or sources with directivity[134, 81, 59].

As a ray model and Gaussian beam model, this algorithm inherits the limitations being a high frequency approximation, not a full wave solution. The analytic ray tracer relies on existence of spatial coherence in the media to perform efficiently. The Gaussian beam model that is used to compute the sound field relies on carefully chosen parameters that control the beam width[16], and it is best determined on a per-scene basis.

5.8 Conclusion

This paper combines the analytic ray curve tracer and Gaussian beam model to form an efficient solution for outdoor sound field computation. Based on the parabolic ray formulation[101], analytic solutions that

computes on-ray pressure as well as paraxial field are used with a Gaussian beam model. The efficiency of the paths computation by the analytic ray tracer is matched with the efficiency of the pressure computation, and the combined algorithm can efficiently simulate the propagated sound field for large outdoor scenes with general input media and arbitrary obstacles. This algorithm is validated on benchmarks with published results generated by validated methods including FDTD, PE, and normal modes. Simple benchmarks are chosen to facilitate the verification of the results, but they cover scenarios that span inhomogeneous media and obstacles with atmospheric and underwater propagation. The capability of this algorithm is further demonstrated with a complex artificial scene and a variety of media conditions that would present scalability challenges to existing methods. Results that reflect the characteristics of the scene and media are generated at close-to-interactive speed. As future work we hope to fully extend and apply this algorithm on more outdoor scenarios, and to obtain measured data or to run large scale numerical simulation for further validation of the algorithm.

5.9 Derivations

5.9.1 Analytic evolution of ray trajectories

With ray parameter σ defined by $d\sigma = V^2 dT = V ds$, the ray trajectories x_i , slowness p_i , and travel time T can be evolved analytically from σ_0 to any σ along the ray:

$$x_i(\sigma) = x_{i0} + p_{i0}(\sigma - \sigma_0) + \frac{1}{4}A_i(\sigma - \sigma_0)^2, \quad (5.7)$$

$$p_i(\sigma) = p_{i0} + \frac{1}{2}A_i(\sigma - \sigma_0), \quad (5.8)$$

$$T(\sigma) = T(\sigma_0) + V_0^{-2}(\sigma - \sigma_0) + \frac{1}{2}A_i p_{i0}(\sigma - \sigma_0)^2 + \frac{1}{12}A_i A_i(\sigma - \sigma_0)^3. \quad (5.9)$$

5.9.2 Analytic evolution of Cartesian P and Q

The characteristic system of the Hamiltonian for of the Eikonal equation gives:

$$\frac{dp_i}{d\sigma} = \frac{1}{2} \frac{\partial}{\partial x_i} \left(\frac{1}{V^2} \right), \quad \frac{dT}{d\sigma} = p_k p_k = V^{-2}. \quad (5.10)$$

From equation 5.10 and because partial derivative $\partial/\partial\gamma$ commutes with $d/d\sigma$, a simple *dynamic ray tracing system* can be derived as follows:

$$\frac{d}{d\sigma} Q_i^{(x)} = P_i^{(x)}, \quad \frac{d}{d\sigma} P_i^{(x)} = \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j} \left(\frac{1}{V^2} \right) Q_j^{(x)}, \quad (5.11)$$

For constant media gradient of V^{-2} , 5.11 can be solved analytically for any point R at σ along the ray Ω if $\hat{\mathbf{Q}}^{(x)}$ and $\hat{\mathbf{P}}^{(x)}$ are known at any other point S at σ_0 along the ray Ω :

$$P_{iJ}^{(x)}(x_i) = P_{iJ}^{(x)}(S), \quad Q_{iJ}^{(x)}(R) = Q_{iJ}^{(x)}(S) + (\sigma - \sigma_0) P_{iJ}^{(x)}(S). \quad (5.12)$$

5.9.3 Analytic evolution of tranformation matrix

For constant gradient of V^{-2} , $\hat{H}$ can be solved analytically for any point R from any other point S along Ω . This is achieved by computing the ray-centered coordinates unit basis $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$ that constitutes $\hat{H}$. Consider a set of orthonormal unit vectors $\vec{n}_1$, $\vec{n}_2$, $\vec{n}_3$ defined along ray Ω , parameterized by σ . Let $\vec{n}_3(\sigma) = V(\sigma)\vec{p}(\sigma)$ follow the tangent of the ray, $\vec{n}_2(\sigma)$ is selected to be perpendicular to the ray plane, $\vec{n}_1$ is then defined by $\vec{n}_1 = \vec{n}_2 \times \vec{n}_3$. Because the ray is a planar curve for constant gradient V^{-2} , $\vec{n}_1(\sigma) = \vec{n}_1(\sigma_0)$. Given Equations 5.7-5.9,

$$\begin{aligned} \vec{n}_2(\sigma) &= \vec{n}_1(\sigma) \times \vec{n}_3(\sigma) = \vec{n}_1(\sigma) \times V(\sigma)\vec{p}(\sigma) \\ &= \vec{n}_1(\sigma_0) \times V(\sigma)(\vec{p}(\sigma_0) + \frac{1}{2}\vec{A}(\sigma - \sigma_0)), \end{aligned} \quad (5.13)$$

As e_3 coincides with $\vec{n}_3$, $\vec{e}_1(\sigma)$, $\vec{e}_2(\sigma)$ can be determined from $\vec{e}_1(\sigma_0)$, $\vec{e}_2(\sigma_0)$ and the evolution of $\vec{n}_1$, $\vec{n}_2$ from σ_0 to σ is:

$$\begin{aligned}\vec{e}_1(\sigma) &= [\vec{e}_1(\sigma_0) \cdot \vec{n}_1(\sigma_0)]\vec{n}_1(\sigma) + [\vec{e}_1(\sigma_0) \cdot \vec{n}_2(\sigma_0)]\vec{n}_2(\sigma), \\ \vec{e}_2(\sigma) &= [\vec{e}_2(\sigma_0) \cdot \vec{n}_1(\sigma_0)]\vec{n}_1(\sigma) + [\vec{e}_2(\sigma_0) \cdot \vec{n}_2(\sigma_0)]\vec{n}_2(\sigma).\end{aligned}\tag{5.14}$$

5.9.4 Evolution of ray-centered $\mathbf{P}$ and $\mathbf{Q}$

1. Take initial condition for $\mathbf{P}$, $\mathbf{Q}$. Assuming a point source S and ϕ_0 and i_0 as the ray parameters γ_1 , γ_2 : $\mathbf{Q}(S) = 0$, $\mathbf{P}(S) = \frac{1}{V(S)} \begin{pmatrix} 1 & 0 \\ 0 & \sin i_0 \end{pmatrix}$.
2. Transform $\mathbf{P}$, $\mathbf{Q}$ into $\hat{\mathbf{P}}^{(x)}$ and $\hat{\mathbf{Q}}^{(x)}$ with $\hat{H}$,
3. Analytically evolve $\hat{\mathbf{P}}^{(x)}$ and $\hat{\mathbf{Q}}^{(x)}$ by Equation 5.12,
4. Analytically evolve $\hat{\mathbf{H}}$ by Equation 5.14,
5. Transform the evolved $\hat{P}^{(x)}$ and $\hat{Q}^{(x)}$ back to $\mathbf{P}$, $\mathbf{Q}$ with the evolved $\hat{H}^T$.

5.9.5 Cartesian paraxial travel time

Denote the Cartesian coordinates of R' and R_γ by $x_i(R')$ and $x_i(R_\gamma)$, and $x_i(R, R_\gamma) = x_i(R) - x_i(R_\gamma)$, the quadratic expansion of T from $T(R_\gamma)$ is:

$$T(R, R_\gamma) = T(R_\gamma) + x_i(R, R_\gamma)p_i^{(x)}(R_\gamma) + \frac{1}{2}x_i(R, R_\gamma)x_j(R, R_\gamma)M_{ij}^{(x)}(R_\gamma),\tag{5.15}$$

where M_{ij} are the elements of the 3×3 matrix $\hat{\mathbf{M}}^{(x)}$:

$$\hat{\mathbf{M}}(R_\gamma) = \hat{\mathbf{H}}(R_\gamma) \begin{pmatrix} \mathbf{M}(R_\gamma) & M_{13}(R_\gamma) \\ & M_{23}(R_\gamma) \\ M_{13}(R_\gamma) & M_{23}(R_\gamma) & M_{33}(R_\gamma) \end{pmatrix} \hat{\mathbf{H}}^T(R_\gamma).\tag{5.16}$$

Here $\mathbf{M}(R_\gamma)$ is defined in Eq. 5.2. The remaining elements can be derived [161] to be:

$$M_{13}(R_\gamma) = -(v^{-2}v_{,1})_{R_\gamma}, \quad M_{23}(R_\gamma) = -(v^{-2}v_{,2})_{R_\gamma}, \quad M_{33}(R_\gamma) = -(v^{-2}v_{,3})_{R_\gamma}, \quad (5.17)$$

$$v = [V(q_1, q_2, s)]_{q_1=q_2=0, s=s(R_\gamma)}, \quad v_{,i} = [\partial V(q_1, q_2, s)/\partial q_i]_{q_1=q_2=0, s=s(R_\gamma)}. \quad (5.18)$$

$v_{,i}$ can be solved by transforming to Cartesian coordinates first: $v_{,i} = \partial V/\partial q_i = H_{ki}\partial V/\partial x_k$, and $\partial V/\partial x_k$ can be solved analytically for constant gradient of V^{-2} by:

$$\partial V^{-2}/\partial x_k = -2V^{-3}\partial V/\partial x_k = A_k \implies \partial V/\partial x_k = -\frac{1}{2}V^3 A_k. \quad (5.19)$$

5.9.6 Gaussian beam summation

The contributions of Gaussian beams are then summed up by integral superposition:

$$p(R, \omega) = \iint_{\mathcal{D}} \Phi(\gamma_1, \gamma_2) P^{ray}(R_\gamma) \exp[i\omega T(R, R_\gamma)] d\gamma_1 d\gamma_2, \quad (5.20)$$

where R is the field point and R_γ is a point on the ray γ of the ray parameter γ_1, γ_2 . The weighting function Φ is derived to be:

$$\Phi(\gamma_1, \gamma_2) = (\omega/2\pi)[- \det(\mathbf{M}(R_\gamma) - \mathbf{M}^a(R_\gamma))]^{1/2} |\det \mathbf{Q}^a(R_\gamma)| \quad (5.21)$$

$$= (\omega/2\pi)[- \det(\mathbf{Q}^{aT}(\mathbf{M} - \mathbf{M}^a)\mathbf{Q}^a)]^{1/2}. \quad (5.22)$$

The choice of $\text{Re}(\mathbf{M})$ is related to the curvatures of the wavefront and the choice of $\text{Im}(\mathbf{M})$ is related to the width of the amplitude profile. They can be specified at R_γ or any other point along the central ray γ to control the shape of the beam.

Chapter 6

Acoustic pulse propagation in an urban environment using a three-dimensional numerical simulation

6.1 Overview

Acoustic pulse propagation in outdoor urban environments is a physically complex phenomenon due to the predominance of reflection, diffraction, and scattering. This is especially true in non-line-of-sight cases, where edge diffraction and high-order scattering are major components of acoustic energy transport. Past work [3] has shown that many of these effects can be captured using a two-dimensional finite-difference time-domain method, which was compared to the measured data recorded in an army training village. In this paper, a full three-dimensional analysis of acoustic pulse propagation is presented. This analysis is enabled by the

adaptive rectangular decomposition method [123], which models sound propagation in the same scene in three dimensions. The simulation is run at a much higher usable bandwidth (nearly 450 Hz), and took only a few minutes on a desktop computer. It is shown that a three-dimensional solution provides better agreement with measured data than two-dimensional modeling, especially in cases where propagation over rooftops is important. In general, the predicted acoustic responses match well with measured results for the source/sensor locations.

6.2 Introduction

Acoustic propagation in urban environments is a physically complex problem that has many practical applications. In urban planning and city design, acoustic propagation models can inform decisions on the location of noise-sensitive buildings like hospitals and schools [78]. Accurate computational modeling is also useful in designing baffles near areas of high traffic to control noise levels in residential neighborhoods [96, 69]. Acoustic modeling is also useful in sound-source localization: numerous sensors are placed in an urban environment to detect sound events and calculate the sound source’s position using the peak arrival times of the sound waves. This computation of sound source’s position can be used for gunshot localization for crime control in urban areas [86] and in many military applications [4].

Acoustic propagation modeling for urban areas is a challenging computational problem because of the complex building geometry and large domain size. High-order diffraction and scattering play a significant role in acoustic energy transport in urban areas, especially in cases when the source and receiver are not in line-of-sight. Previous work in the field has mainly focused on continuous noise sources to determine statistical quantities like reverberation time and noise levels [173, 29, 70]. However, these are gross acoustic parameters and do not give a detailed view of the actual propagation. Geometric techniques have been used to evaluate noise levels and calculate sound propagation in urban streets [62, 77]. However, due to inherent assumption of rectilinear propagation of sound waves, modeling wave effects such as diffraction and interference remains a significant challenge with these techniques.

Recent work in numerical techniques has focused on the use of acoustic pulse propagation techniques in time domain to get detailed characteristics of the complex propagation effects in urban scenes. Time-domain pulse propagation is preferred in urban acoustic modeling as it gives direct insight into the propagation by producing animations of pressure wavefronts. This allows one to quickly inspect the propagation path corresponding to dominant peaks in the response at a given sensor location. Recent studies used a finite-difference time-domain simulation to model acoustic pulse propagation and compared the results with real-world measurements [90, 5, 3]. Those studies were limited to 2D modeling due to the high computational cost and memory requirement of the finite-difference technique for this large domain size.

In this paper, a full 3D analysis of acoustic pulse propagation in time domain is presented. Our analysis is made on a virtual 3D model of the same scene as the prior 2D investigation [3]; this 3D analysis is made computationally feasible by using Adaptive rectangular decomposition (ARD) [123], an efficient time-domain numerical solver, allowing us to model propagation in this scene in three dimensions. ARD is much more compute- and memory-efficient for homogenous media than the finite-difference time-domain technique. The improved efficiency allows the simulations to have a much higher usable bandwidth, up to mid-range frequencies of 450 Hz, compared to 200Hz in prior work [3], while taking just a few minutes on a desktop computer. A detailed analysis of errors between measured data and 3D simulated data is performed, showing that 3D simulations provide better agreement with measured data than 2D simulations. The agreement is markedly better in cases where propagation over rooftops is important, a case which the 2D modeling cannot capture at all. In general, the predicted acoustic responses match well with measured results for most source/sensor locations, with typical errors being on the order of 3 dB. Visualizations of the time-domain simulation show that a rooftop-diffracted path provides important energy contributions at certain locations in the scene.

6.3 Previous work

Over the years, many techniques have been developed to study acoustic propagation in urban environments [78]. Analytical solutions are available for simple scenarios involving building edges and noise barriers [116]. Theoretical predictions have been used to predict the noise levels in urban street complexes [87]. Statistical analysis [49, 114] has been performed on measured data to analyze the reverberation time and sound levels in streets and to study the relationship between different noise descriptors in urban areas. Many ray-tracing based approaches [62, 70] have been proposed that evaluate the increase in traffic noise for street canyons due to the presence of buildings. Kang et al.[77] used a radiosity-based model to calculate sound propagation in interconnected urban streets. And the radiosity-based model has been combined with the image-source method to handle diffuse and geometrical boundaries, respectively, for street canyon scenarios [76].

Typical numerical approaches utilized to study urban acoustic propagation are Finite-Difference Time-Domain [23], Finite Element Method [152], Boundary Element Method [27], Equivalent Source Method [104] and Pseudo-Spectral Time-Domain [25]. The boundary element method has been applied to acoustic propagation in areas with noise barriers [120] and in outdoor scenes [33]. To model sound propagation in city canyons, Ogren and Kropp (2004) used the equivalent source method and Van Renterghem et al. (2006) used a coupled finite-difference-parabolic equation method [108, 162]. Ovenden et al. (2009) coupled the analytical calculation to a parabolic equation method for modeling noise propagation in urban freeways [111]. To model atmospheric sound propagation, a pseudo-spectral time-domain (PSTD) approach [68] was proposed. The finite-difference approach has been used in recent years to model acoustic pulse propagation in urban environments and compare the results with measured waveforms recorded at the physical site. This includes propagation for a right-angled wall [90], a single building [5], and a training village with multiple buildings [3]. However, due to computational limitations, all these approaches have been limited to propagation in two dimensions.

Some recent studies have modeled three-dimensional sound propagation. Ketcham et al. (2008) used a finite-difference approach for modeling the effect of urban infrastructure on sound scattering in three dimen-

sions [80], but the modeling required a computer cluster with hundreds of processors. Pollès et al. (2004) proposed a diffusion-equation-based approach to model 3D sound propagation in urban areas with multiple buildings [117]. Recently, a fast and efficient time-domain approach was proposed; this technique called adaptive rectangular decomposition, solves the wave equation in three dimensions for spatially invariant speed of sound [124, 100]. It is related to the PSTD technique but avoids the discrete integration in time by using analytical solutions of the wave equation for rectangular domains. For a more detailed survey of outdoor sound propagation techniques, the survey paper [174] is recommended.

6.4 Measurements

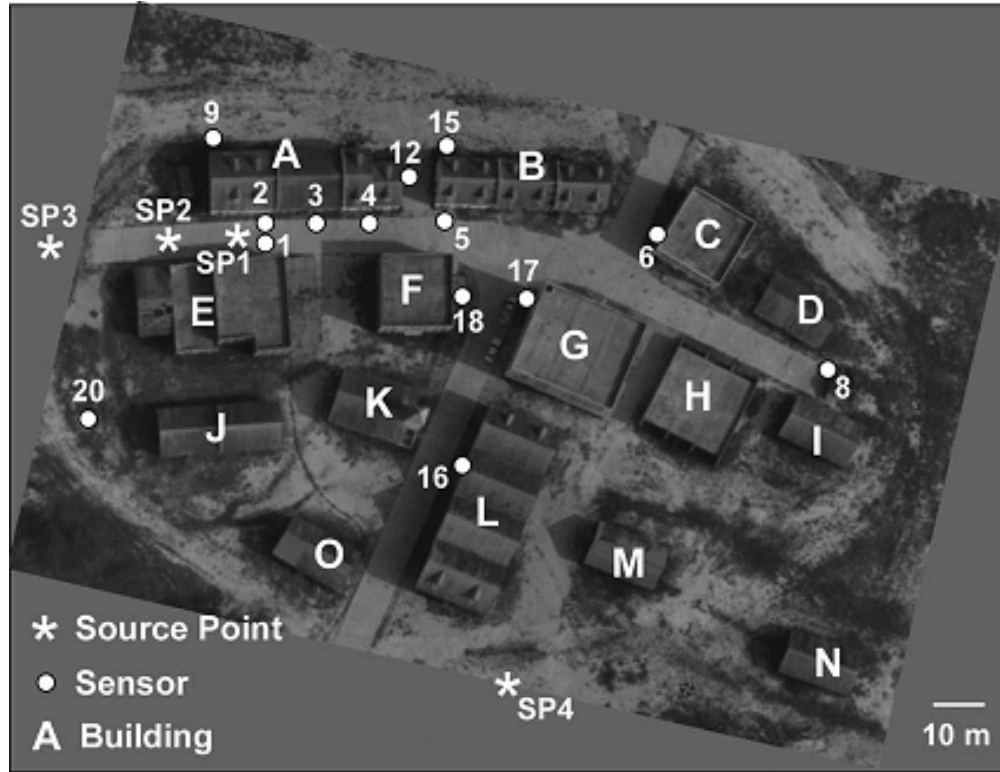
In this section, we discuss the real-world measurements used in the validation of the numerical simulation. This dataset was presented in the work of [3]. We provide a brief discussion here, but more details can be found in their previous work.

Scene Layout The experiment was conducted in an artificial village spanning a $150 \times 150 \text{ m}^2$ area with 15 buildings and two cross-streets: “Main street” running nearly horizontal, and “Church street” running in an approximately vertical direction. Figure 6.1(a) shows the 2D layout (top view) of the urban scene. The buildings in the village were two or three stories tall and made up of concrete blocks. The ground areas consisted of streets, grass areas, and hard-packed soil.

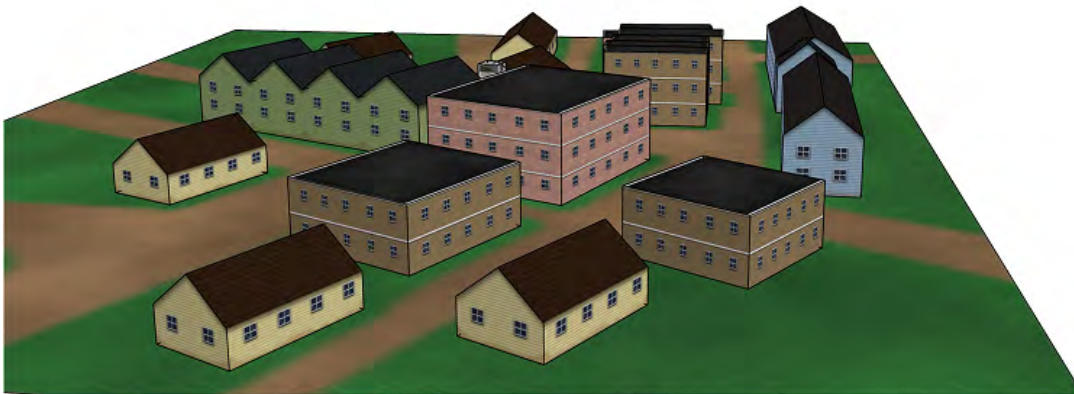
Weather Conditions: The experiment was conducted over two sunny days with temperature, wind and relative humidity variation between 8 to 19° C, 2 to 5 m/s and 30%-50% respectively.

Sources: Acoustic pulses were produced by using small explosives of 0.57 kg of C4 suspended at a height of 1.5 m from the ground. The measurements were recorded for four source positions SP1-SP4.

Receivers: Sensors were placed at 14 different receiver positions spread throughout the scene, in both line-of-sight(LOS) and non-line-of-sight(NLOS) positions. These sensors were connected to digital seismographs that recorded the pressure signal at a sampling rate of 5 or 8 kHz.



(a)



(b)

Figure 6.1: a) Top view of the urban scene used in the experimental study. Reproduced with permission from Liu and Albert (2010). b) An approximate 3D model of the scene constructed based on the 2D layout, photographs of the scene and heights of the buildings corners and roof tops. (Color online.)

6.5 Adaptive rectangular decomposition numerical modeling

In this section, we give an overview of the adaptive rectangular decomposition (ARD) simulation technique for modeling acoustic pulse propagation [124, 100, 122].

6.5.1 The adaptive rectangular decomposition method

Our starting point is the wave equation for constant sound speed,

$$\frac{\partial^2 p}{\partial t^2} - c^2 \nabla^2 p = f(\mathbf{x}, t), \quad (6.1)$$

where $p(\mathbf{x}, t)$ is the time-varying pressure field, $f(\mathbf{x}, t)$ is the force term corresponding to the volume sound sources, c is the speed of sound and ∇^2 is the Laplacian operator. The speed of sound in the medium is assumed to be spatially invariant.

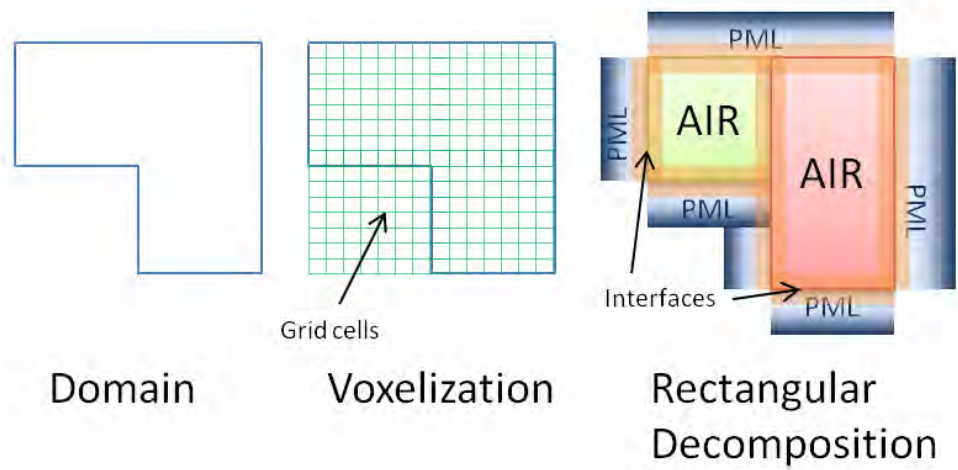
Figure 6.2 demonstrates the main stages of the ARD pipeline. The technique starts with a 3D model of the scene, voxelizes the air volume, and then decomposes the voxelization into rectangular partitions. The wave equation has a known analytical solution for rectangular domains for spatially-invariant speed of sound. Consider a rectangle in 3D of size (l_x, l_y, l_z) with perfectly reflecting walls. The analytical solution of the wave equation in this case can be written as:

$$p(\mathbf{x}, t) = \sum_{i=(i_x, i_y, i_z)} m_i(t) \Phi_i(\mathbf{x}), \quad (6.2)$$

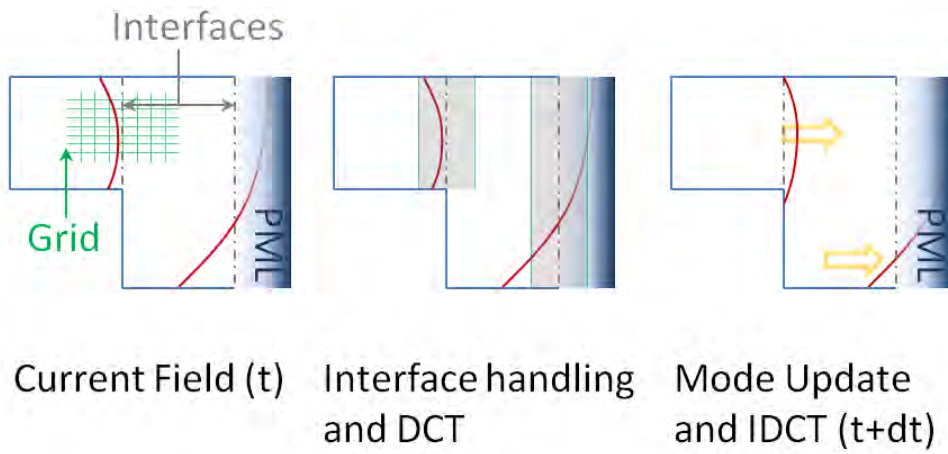
where i_x are x-indices in the range $[1 - l_x]$, and, i_y and i_z are y-and z-indices, respectively. Here $i = (i_x, i_y, i_z)$ is a generalized index over three dimensions, $m_i(t)$ are time-varying mode coefficients, and $\Phi_i(\mathbf{x})$ are eigenfunctions of the Laplacian, given by

$$\Phi_i(\mathbf{x}) = \cos\left(\frac{\Pi i_x}{l_x} x\right) \cos\left(\frac{\Pi i_y}{l_y} y\right) \cos\left(\frac{\Pi i_z}{l_z} z\right), \quad (6.3)$$

for a perfectly-reflecting boundary condition. In discrete interpretation, pressure can be transformed into



(a) Preprocessing



(b) Simulation

Figure 6.2: Different stages of the Adaptive Rectangular Decomposition simulator. (Color online.)

mode coefficients M_i at each time step by a Discrete Cosine Transform (DCT), since the eigenfunctions are cosines. The update rule for mode coefficients can be derived by taking a DCT of the wave equation and solving the resultant simple harmonic oscillator system, giving

$$M_i^{n+1} = 2M_i^n \cos(w_i \Delta t) - M_i^{n-1} + \frac{2F_i^n}{w_i^2} (1 - \cos(w_i \Delta t)), \quad (6.4)$$

where the superscript indicates the number of time steps, $w_i = c\pi \sqrt{\left(\frac{i_x^2}{l_x^2} + \frac{i_y^2}{l_y^2} + \frac{i_z^2}{l_z^2}\right)}$, F^n is the DCT of force $f(\mathbf{x}, t)$ at n^{th} time step and Δt is the size of the time step. Mode coefficients are then transformed back into pressure by an inverse DCT. This gives the pressure inside each rectangular partition. The pressure is propagated across neighbouring partitions by performing interface handling using a sixth-order finite-difference stencil. To incorporate sound absorption at the partition boundaries, Perfectly Matched Layer (PML) absorbing boundary conditions are used. Currently, the ARD simulator can handle absorption and reflection of sound while ignoring transmission through objects. For more details, please refer to the original texts [124, 100].

The ARD technique is more efficient than the FDTD technique because of its larger grid spacing and time steps. The grid spacing for the ARD technique is $h = c/(\nu_{\max}s)$, where $\nu_{\max}$ is the maximum simulation frequency and $s = 2.6$ is number of samples per wavelength. The ARD's simulation time step is restricted by the Courant-Friedrichs-Lewy (CFL) condition $\Delta t \leq h/(c\sqrt{3})$. In contrast, the FDTD technique requires a much higher value of samples per wavelength ($s = 10$ used in Taflov et. al [145] or $s = 20$ used in Albert et. al [3]), resulting in much denser grid and smaller time steps. Therefore, the ARD technique is computationally more efficient and requires less memory than the FDTD technique. This efficiency enables the ARD technique to perform 3D wave-simulations on large, complex scenes at a higher simulation frequency than FDTD can, all on a desktop computer.

The ARD technique has few intrinsic limitations, the primary one being the spatially invariant speed of sound. However, sound speed can change spatially due to many factors such as temperature gradient, humidity, or wind. The ARD technique does not model the effect of these factors on sound propagation.

Another limitation is that atmospheric absorption is currently not modeled in the simulation. Also, the simulation does not model sound transmission through the objects (walls, buildings, etc). In a general scenario, these limitations can have an effect on the quality of the simulation results. However, in the present study, these limitations have negligible effect on the prediction quality. The sensors are placed at a height of 1.5m and most of the energy recorded at these sensors is due to the acoustic propagation happening close to the ground ($< 15m$). For such small elevations, the temperature gradient of the atmosphere is negligible and speed of sound remains unaffected due to temperature gradient. The wind speed is also very low (2-5 m/s) to have any significant effect on arrival times of the acoustic pulses. Also, atmospheric absorption can be safely ignored in this case as the simulation frequency is less than 500Hz and the propagation distances are of the order of hundreds of meters for which the intrinsic absorption of the atmosphere is negligible. As for sound transmission, due to the very high impedance contrast between the air and the concrete buildings, very little acoustic energy can get transmitted through the buildings.

6.5.2 Validation

We provide validation results of the ARD technique on two benchmark test-cases: (a) spherical wave scattering by a rigid sphere and (b) edge-diffraction from a right-angled rigid wall. In the first case, the acoustic wave equation has known analytical solution [57]. The scene setup is as follows: a sphere of radius $a = 1m$, surrounded by air with speed of sound $343m/s$ and mean density of $1.21kg/m^3$, is centered at origin $(0, 0, 0)$. A spherical sound source (monopole source) is placed at position $(0, 0, -3)$. The spherical wave emitted by the source is scattered by the rigid sphere. The total field (incident+scattered field) is computed using the analytical solution of the wave equation at an angular distribution of listener positions at distance of 1.5m. The analytical solutions are compared against the simulation results at different wave numbers k as shown in Figure 6.3. The results are plotted versus the polar angle θ where $\theta = 180^\circ$ corresponds to the front end of sphere with respect to the incoming spherical wave. The comparison between the analytical expressions and the ARD simulation results show very good agreement.

In the second case, we perform validation of the ARD technique by comparing it against the edge

diffraction model proposed by Svensson et al. [144]. This model is an extension of Biot-Tolstoy-Medwin solution [97] to finite edges. The scene setup is as follows: a right-angled rigid wall of dimension $8m \times 12m$ is considered with the longer edge being the diffraction edge. Source and receiver are placed at symmetric positions with respect to the wall at $(-1.8, -0.9, -6.0)$ and $(0.9, 1.8, -6.0)$, respectively. The time- and frequency-domain responses are computed using the BTM finite-edge diffraction model and compared against the results of the ARD simulation. As shown in Figure 6.4, the agreement between the two responses is very good.

6.5.3 Source waveform

The source function used to model the explosive blast signal for calculations in the ARD simulator is described in [90]. Figure 6.5 shows the corresponding source function with peak pressure normalized to 1.

6.5.4 3D model

In order to run a 3D numerical simulation, a virtual 3D model of the scene is required. However, a detailed 3D model cannot be constructed due to the lack of availability of architectural blueprints or a laser-scanned point cloud of the site. Therefore, we construct a simplified 3D model of the scene using a 2D layout of the village, photographs, and the heights of corners and rooftops of buildings. This 3D model is an approximation to the actual geometry of the scene, since it lacks particular geometric details (peaked roofs, door/window locations, facade details, and extraneous geometry such as cars and a fountain). The dimensions of the simulation domain are $175m \times 140m \times 14m$. The heights of the building are between $6 - 9m$. Therefore, depending on the building, a vertical space of $5 - 8m$ is above the roof to the top of the simulation domain allowing correct simulation of rooftop diffraction. Figure 6.1(b) is a textured rendering of the 3D model. Based on the type of material present (e.g. concrete, grass, soil, etc.), we assign the appropriate absorption coefficients to the surfaces of the 3D model.

6.5.5 Simulation parameters

The ARD simulation was run with an acoustic wave velocity of $375ms^{-1}$ and an air density of $1.2kgm^{-3}$. The high value for the acoustic wave velocity comes from the propagation of the high-amplitude acoustic pulse generated by the C4 explosive used as the sound source. In case of concrete, the acoustic wave velocity is $2950ms^{-1}$ and density is $2300kgm^{-3}$, which results in a reflection coefficient of 0.99 for concrete. These parameters correspond to the values used in Albert and Liu (2010) for the finite-difference simulation.

We run the ARD simulator to propagate the acoustic pulse from each given sound source position, one by one. We record the response at the specified receiver positions and compare the results with the recorded measurement data.

6.5.6 3D vs 2D wave simulation

We discuss the advantages of the 3D wave simulator over the 2D wave simulator for acoustic pulse propagation. Firstly, a 3D wave simulation incorporates propagation paths over the top of walls or buildings, as well as wave diffraction from the upper edges, both of which are completely ignored by a 2D simulation. Secondly, in 3D simulation, the sound reflection from the ground terrain is handled accurately for all frequencies. For a 2D simulation, the pressure is simply doubled to approximate ground reflections, which is accurate only for frequencies up to $600Hz$, as discussed in Liu et al.(2006). Lastly, the results of a 2D simulation must be renormalized by an additional factor of $1/\sqrt{r}$ to account for 3D geometric spreading. This normalization is valid only for large kr (where k is the wave number and r is distance to the source). A 3D simulation requires no such normalization.

6.6 Results

In this section we compare the waveforms calculated by the ARD simulator with the measurements recorded in the village, and with the 2D FDTD simulation proposed by Liu and Albert (2010).

Table 6.1: Parameters for the ARD technique.

Parameters	Values
simulation frequency	450Hz
grid size	175m x 140m x 14m
grid spacing	0.31 m
# grid points	11 million
time step size	385 μ s
# time steps	2000
simulation length	0.77 s

6.6.1 Simulation

The ARD simulations were run for the four source positions SP1 to SP4 shown in Figure 6.1(a). The parameters used in the simulator are given in Table 6.1. The total simulation time for each source was 20 mins for the CPU-based ARD implementation [124] and 1–2 mins for the GPU-based ARD implementation [100]. The CPU-based implementation is C++ and the GPU-based implementation uses NVIDIA’s CUDA. These timing results were measured on a single core of a 4-core 2.80 GHz Xeon X5560 desktop machine with 4 GB of RAM and on a NVIDIA GeForce GTX 480 GPU with 448 CUDA cores and 1.5 GB memory.

The ARD responses were computed for four source positions, up to a maximum frequency of 450 Hz. The measured waveforms were low-passed to 450 Hz for source positions 2 and 3 to compare with the calculated ARD waveforms. For source position 1 and 4, both the ARD and the measured responses were low-passed to 200 Hz, since the 2D FDTD waveforms are available only up to 200 Hz. In this scene, the simulation frequency is less than 500Hz and the propagation distances are of the order of hundreds of meters, for which the intrinsic absorption of the atmosphere is negligible. Therefore, we have ignored atmospheric absorption during the simulation.

Figure 6.6 shows the visualization of the time-domain ARD simulation at specified time steps. These visualizations show the propagation of wave fronts in the scene and how they are modified by the multiple reflections and diffractions from the buildings, and reflections from the ground. These can be helpful in guiding engineering modifications to the scene.

6.6.2 Varying propagation speed

In the training village scene considered in this study, the sound source used is a C4 explosive. The high-amplitude explosion caused by this source generates a varying sound-speed profile. Figure 5 in Liu and Albert (2010) shows the measured values: over 400 m/s a meter or two away and 375 m/s at about 20 m distance. The speed varies with the amplitude of the traveling pulse, reducing with distance. To make things more complicated, the amplitude of a wave that diffracts around a building corner is reduced, and that diffracted wave travels slower than waves at the same propagation distance that are in line-of-sight paths. This generates a complicated wave-speed profile that is harder to model. The peak arrival times can get shifted in the measured waveform as compared to the calculated waveforms, which assume a constant speed of sound. Also, the individual peaks can get stretched due to decreasing wave speed. Therefore, there are complicated changes in acoustic wave speed even for different waves at the same location.

Similar to the 2D finite difference simulation of Albert and Liu (2010), our 3D ARD simulation also did not allow us to track these changes with amplitude, one would have to follow the individual wavefronts to do that correctly, or modify the code to explicitly include non-linear effects. Instead, we chose a constant wave speed to get the best waveform fits when compared with the measured data. Also, similar to Albert and Liu (2010), we perform time-shifting to the response of linear simulation to align the first arrival peak with the measured waveform. In addition to that, we try to account for these kinematic errors due to the varying sound speed with the “robust” error metric discussed later.

6.6.3 Error metrics

Basic error metric

In order to perform a quantitative evaluation between the measured and the calculated waveforms, two types of error metrics are used in the comparison: the spectrogram difference metric and the average dB error metric. The spectrogram difference metric (SDM) is computed on the spectrograms *SPEC* of the

time-domain pressure signals $\{a\}_{i=1}^N$ and $\{b\}_{i=1}^N$:

$$SDM(a, b) = \frac{\sum_{j=1}^M \sum_{k=1}^T \|SPEC(a)_{jk} - SPEC(b)_{jk}\|^2}{\sum_{j=1}^M \sum_{k=1}^T \|SPEC(a)_{jk}\|^2}, \quad (6.5)$$

where $M = \{N/2 \text{ if } N \text{ is even or } N/2 + 1 \text{ if odd}\}$ and T is the number of time segments in spectrogram.

The average dB error metric (ADM) is computed on the pressure signals in the dB scale $\{dB(a)\}_{i=1}^N$ and $\{dB(b)\}_{i=1}^N$ as follows:

$$ADM(a, b) = \sum_{i=1}^N \|dB(a)_i - dB(b)_i\|/N. \quad (6.6)$$

Robust error metric

The error metrics defined above are very sensitive to the time of arrival in the waveforms. In scenarios where the speed of sound is constant, these metrics perform well. However, in this case (as discussed in Section 6.6.2) where the speed of sound varies, the arrival times can be off by few milliseconds. Even though this small time-shift might not cause a big difference in individual waveform characteristics such as shape, frequency content, etc, it can generate a large error with these error metrics. One can think of using regular cross correlation to remedy this situation. However, regular cross correlation will match the peak arrival but will misalign the rest of the measured waveform.

We propose changes to the above error metrics that will make them more robust to the small time-shifts and peak stretching caused by varying sound speeds. Our proposed change is based on the idea of Dynamic Time Warping (DTW), which is a standard tool in signal-processing community to handle non-linear transformations in the time axis. In the DTW technique, the two input time signals are allowed to shift, stretch, or contract, in a limited manner to generate the most optimal matching between the signals. This technique has been widely used in the areas of speech processing [132], acoustics [95], bioinformatics [1], and medicine [30].

In our proposed solution, we align the first arrival peaks of the calculated and measured waveforms. This removes any time-shift that happens before the arrival of the first peak. Next, we take these signals and apply the DTW technique to align the remaining part, thus taking into account small time shifts

and stretching. Finally, the above error metrics are applied on these aligned signals to give a quantitative measure of the error; we also define at the same time the *confidence* we have in the resulting measure. The confidence measure is based on the intuition that the less DTW warping required to align the signals, the more confidence we have in the similarity of the signals.

$$\text{Confidence} = (1 - d_w/l_o) \times 100 \quad (\text{in percent}),$$

where $d_w = |l_w - l_o|$ is the difference in the length of the warped signal l_w and the original signal l_o . For all the results shown in the paper, we allow a warping length change of only 5 – 10%, resulting in a confidence measure of 90 – 95%.

6.6.4 Comparison with measurements in time domain

In this section, we compare the waveforms calculated using the ARD simulation with the measured waveforms for different source and receiver positions. Note that waveform modeling requires a strong agreement between the amplitudes and phases of the calculated and measured waveforms, making it a stringent test for any acoustic pulse propagation technique.

Figure 6.8 compares the ARD waveforms and the measurements for the source SP2 and the receiver positions. The upper traces in each panel correspond to the calculated ARD waveforms, and the lower traces correspond to the measurements. The source is located to the left of the narrow street canyon formed between buildings A and E (see Figure 6.1(a)). For all the line-of-sight (LOS) positions (R01 thru R06), we get an excellent match between the calculated and measured waveforms for the direct sound (first arrival) and the subsequent reflections. The main characteristic of LOS responses is the strong first arrival which dominates the signal, followed by (comparatively) weaker reflections. For receivers very close to the source (R01 thru R04), the high amplitude of direct sound completely dominates the later reflections, whereas for receivers far away (R05, R06), the reflections have considerable energy.

For non-line-of-sight (NLOS) positions, the waveform characteristic is position-dependent and more

complex. In the case of receiver R20, the propagation paths mainly consist of diffraction around and from the top of building E. For receiver R08, the diffracted first arrival is followed by high-order reflections from buildings C, H, D and I. In the case of receiver R15, diffracted arrivals around and from the top of buildings A and B are followed by high-order reflections and diffractions from A, F and B. Similarly, for R09, the first arrival corresponds to diffraction from the rooftop of building A followed by high-order reflections and diffractions from buildings A and E. For receivers R17 and R18 the first arrival is via diffraction, followed by multiple reflections trapped between buildings F and G. In the case of R17, the diffraction angle is 10-15 degrees, resulting in a high amplitude diffraction peak; in R18, by contrast, the diffraction angle is 90 degrees, which results in a low amplitude diffraction peak.

As can be seen in Figure 6.8, the calculated waveforms incorporate all the features and match with the measured waveforms to a high degree of accuracy. The biggest mismatch between the waveforms is for sensor R12, which is a NLOS position around the corner of the building A on Main Street. In this case, the ground floor room at the corner had two open windows facing Main Street, and one open window around the corner between Main Street and the sensor at R12. This resulted in a shorter path through that room to the sensor for sound coming from source positions SP2. This is probably what can be seen before the large arrival, which presumably diffracted around the building corner itself. The open windows had no glass at all and were about 2 x 4 feet in area, so this small size (compared to a wavelength of about 7 meters at the source) would reduce the amount of energy traveling on that shorter “indoor” path. Some of the additional high frequency arrivals later on in the measured waveform may be caused by reverberation inside that room. Due to lack of availability of window positions data, these were not included in the virtual 3D model constructed for the ARD simulation. Therefore, these early arrivals are not modeled in the simulation results. Same behavior is observed for source positions SP1 and SP3.

In terms of basic error metric (not shown in the figure), the highest value for this source simulation occurs at receiver R18 (basic SDM error=1.67) due to a decrease in the wave speed after the first diffraction from building F. This causes subsequent strong reflection from the opposite building G to arrive much later in time than the calculated waveform (which assumes a constant sound speed $375ms^{-1}$). This time stretching

cannot be modeled by a constant time shift applied at the beginning of the signal. Thus, the measured waveforms for R18 appear to be the stretched equivalents of the calculated waveforms; this results in a higher error using the basic metric. The DTW-based robust error metric takes this stretching into account, correctly predicting a low error as shown in the figure. Similar behavior can be observed for receivers R16 and R17.

In Figure 6.9, we perform the same comparison between calculated and measured waveforms for the source position SP3. All the LOS positions (R01 to R06, and R20) exhibit excellent matches between the calculated and measured waveforms. For NLOS positions involving first-order diffraction (i.e. R09), the calculated responses incorporate diffraction paths from both around and on top of the buildings to generate the correct waveform. For R08, the first diffraction arrival is weaker than the later reflection; this behavior is corrected modeled by the calculated ARD response. As described before, the waveforms of receivers R16, R17, R18 get stretched in time by the variable acoustic speed; this stretching results in high error with the unwarped metric and low error with DTW-based metric. Aside from sensor R12, the biggest mismatch between the waveforms is for receiver R15 for which the calculated waveform shows two distinct peaks as compared to only one peak for the measured waveform. One possible explanation is that the arrival times between the two diffraction peaks shown in the calculated response is smaller in the real scene (due to varying speed of sound), resulting in a constructive interference and peak merging in the measured waveform.

6.6.5 Comparison with 2D FDTD

In this section, we compare the calculated ARD and FDTD waveforms to the measured waveforms for two source positions, SP1 and SP4. The ARD waveforms are calculated by running a 3D simulation on the 3D model. The FDTD waveforms are calculated by running a 2D finite-difference simulation on a 2D grid as described in Liu and Albert (2006).

Figure 6.7 shows the calculated and measured waveforms for the source position SP1 and its receiver positions. The upper traces in each panel correspond to the calculated ARD response, the middle trace to the measured waveform, and the lower traces to the calculated FDTD responses. In case of LOS positions

(R01 to R06), the dominant propagation happens in the XY plane containing the sources and receivers. Therefore, the waveforms calculated using the 3D ARD simulation and the 2D FDTD simulation match equally well to the measured waveforms. The main difference between a fully 3D and a 2D simulation arises in cases where the sound waves diffract from the rooftops of the buildings, resulting in shorter propagation paths and higher energy (as illustrated in Figure 6.10). For receiver R09, the diffraction path from the top of building A is the shortest path and corresponds to the first arrival. This shortest path is modeled correctly by the 3D ARD simulation. In the case of receiver R20, the secondary arrival peak corresponds to the rooftop diffraction path, which is missing in the 2D simulation waveform. For receiver R15, the energy from rooftop diffraction paths is missing from the 2D simulation but not from the 3D simulation. Therefore, for these NLOS cases, the 3D simulation waveforms match far better with the measured waveforms than the waveforms from the 2D simulation.

In Figure 6.11, we do a similar comparison for source position SP4. This source is positioned outside the main village compound, and most of the receiver positions are non-line-of-sight. The only LOS position is receiver R20, where both the calculated waveforms match well with the measured waveforms. For NLOS positions, the measured waveforms are again stretched as compared to the calculated waveforms, as discussed before. The speed of sound reduces significantly after diffraction, resulting in high-order propagation peaks to arrive later in the measured waveforms than in ARD and FDTD waveforms (which assume constant speed of sound). In the case of receivers R04 to R06 and R15 to R18, the correct modeling of rooftop diffraction with a 3D simulation (ARD) results in a better match with the measured waveforms. The 2D FDTD simulation ignores these paths, resulting in a lower first arrival energy than in the measured data. The measurement data for receiver position R1 is not available for the source position SP4. As described in Liu and Albert (2010), this location was still fitted with the high-pressure blast sensor from the previous measurement, when it was measuring the pressure from the nearby explosive charge at SP1. This high-pressure sensor was unable to detect the low pressure waveform produced by the distant source SP4.

6.6.6 Comparison with measurements in spatial domain

One of the primary advantages of a time-domain wave simulation (FDTD or ARD) is the ability to save snapshots of the pressure field at any time step in the simulation. These snapshots can be assembled into a movie to elucidate wave-field evolution in time as the acoustic pulse travels through the environment. This movie can serve as a useful tool for studying in detail the complex wave-interactions involved in the acoustic pulse propagation.

As an example, wave-field snapshots validate the presence of rooftop diffraction paths in both measured and 3D ARD waveforms for sensor R09 and source SP1 (see Figure 6.12). In Figure 6.10 (upper trace), the first arrival for 3D ARD and the measured waveforms happens at $t = 61$ ms. The wave-field snapshot in Figure 6.12 at $t = 61$ ms shows that the corresponding propagation path is a rooftop diffraction path from the top of building A, followed by a diffraction from the side of the building at $t = 69$ ms. The 2D FDTD simulation cannot model 3D rooftop propagation paths and misses the energy corresponding to this path. Late arrivals in the waveform correspond to high-order reflections between building A and E that get diffracted around the side or the rooftop of building A before reaching the sensor. These correspond to peaks at $t = 83, 104$ and 177 ms, which are marked by circles on top the waveforms in Figure 6.10. These peaks are correctly modeled by the 3D ARD simulation, as shown by the wave-field snapshots for these times in Figure 6.12. The wave-field snapshots and movies can thus provide a more thorough understanding of acoustic pulse propagation.

Figure 6.13 shows the variation of error with distance between the source-receiver positions. The error seems to be independent of the source-receiver distance. Note that error values are much higher for source position SP4 than others. This is due to the presence of more NLOS positions in SP4, which typically have more complex propagation characteristics than LOS positions. The top three sensors with consistently high errors across all sources are R12, R16 and R15. As discussed before, R12 has high errors due to open windows in building A, resulting in shorter propagation paths that are not modeled in the simulation. As for sensor R16 and R15, these are NLOS positions for all sources, which typically have higher errors.

6.7 Conclusion, Limitations and Future Work

In this paper, acoustic pulse propagation results are presented for a large urban environment in three dimensions. The results of the 3D simulation provide better agreement with the measured data than the 2D simulation, especially in cases where rooftop diffraction is involved. This technique enables acoustic propagation in a three-dimensional scene with large size and broad frequency range on a desktop computer.

In future, we would like to explore the acoustic pulse propagation in three dimensions in the kHz range. This would require a very accurate geometric description of the scene (sub-meter accuracy) and parallelization of the ARD technique on a computer cluster. Secondly, we plan to investigate the benefit of a full 3D simulation for time reversal processing [4] to compute the source location given the recorded waveforms at sensor positions. Lastly, we also plan to study the transmission of sound through buildings to determine the noise levels inside the buildings from exterior sources.

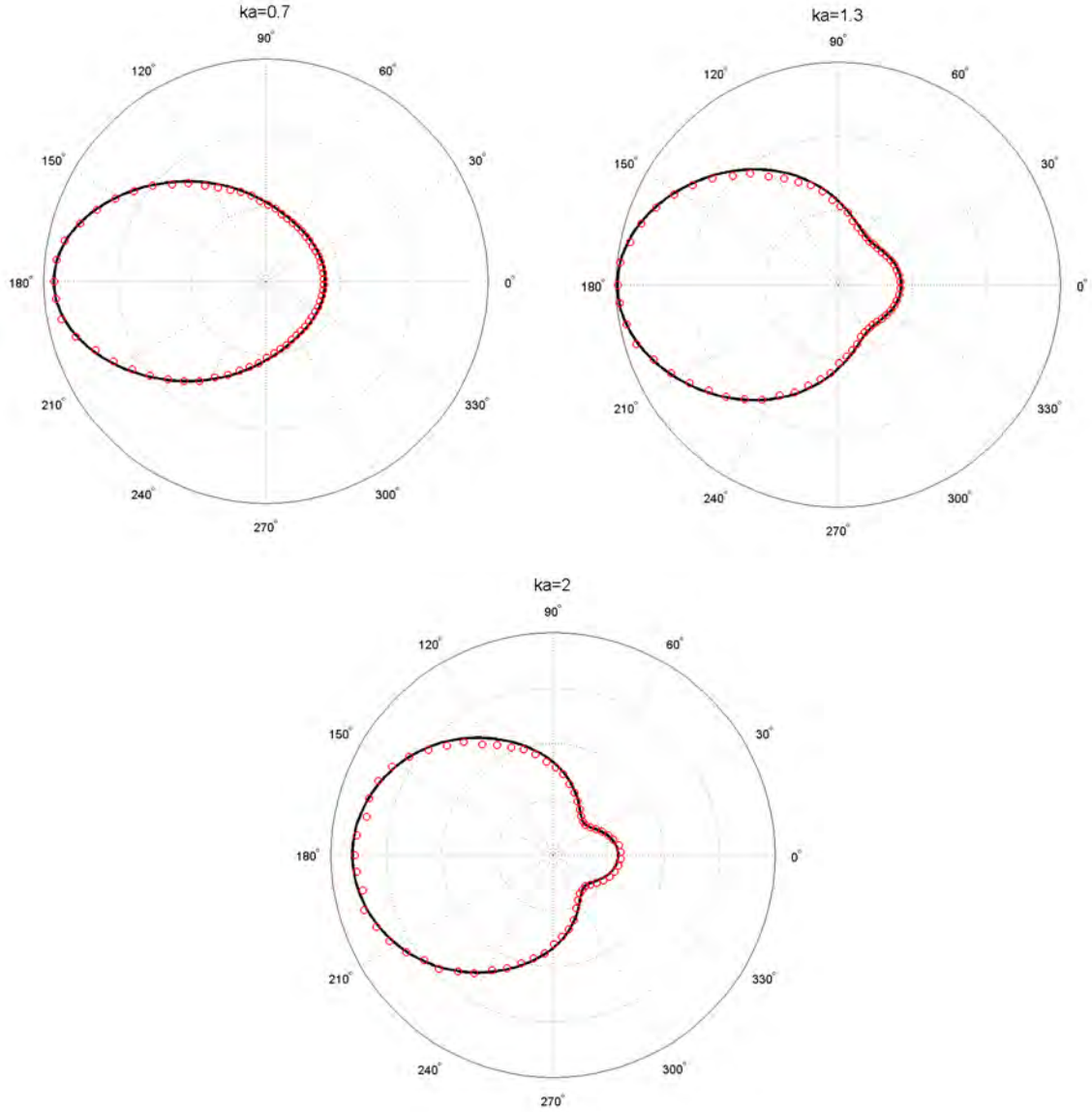
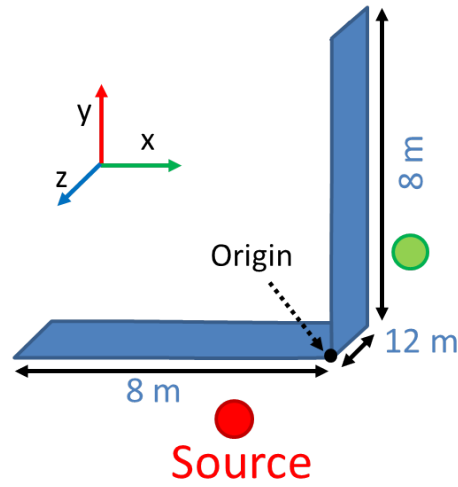
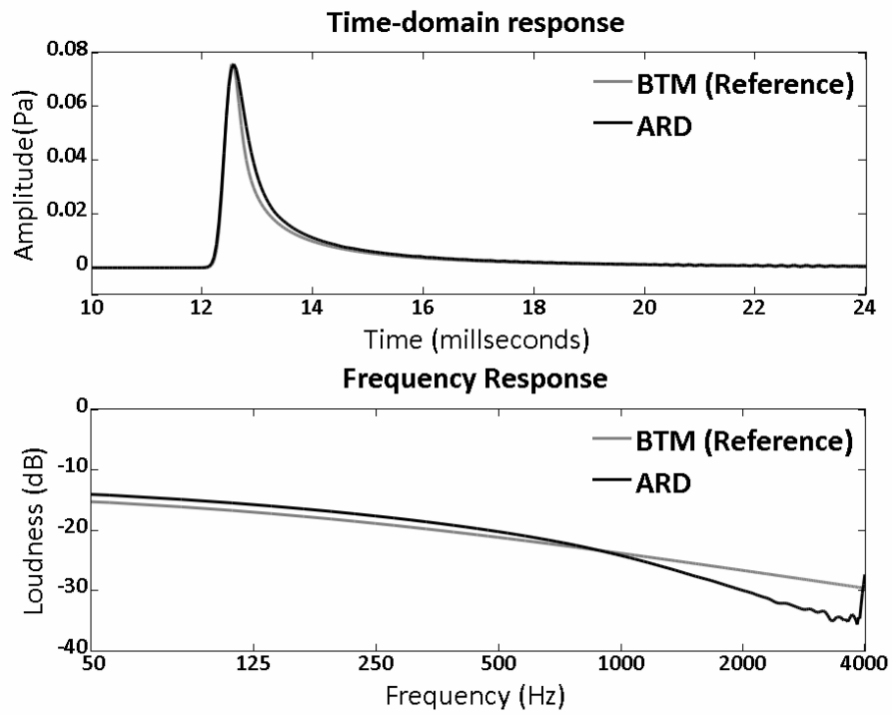


Figure 6.3: Validation of the ARD simulation results (dots) with analytical expressions (curves) for the scattering of a spherical wave by a rigid sphere. Normalized pressure is plotted in the radial axis. The radius of the sphere $a = 1m$ and wave numbers k considered are 0.7, 1.3 and $2 m^{-1}$, respectively. (Color online.)



(a)



(b)

Figure 6.4: Time and frequency responses produced by the Biot-Tolstoy-Medwin diffraction model (reference) and the ARD simulation for a right-angled rigid wall. (Color online.)

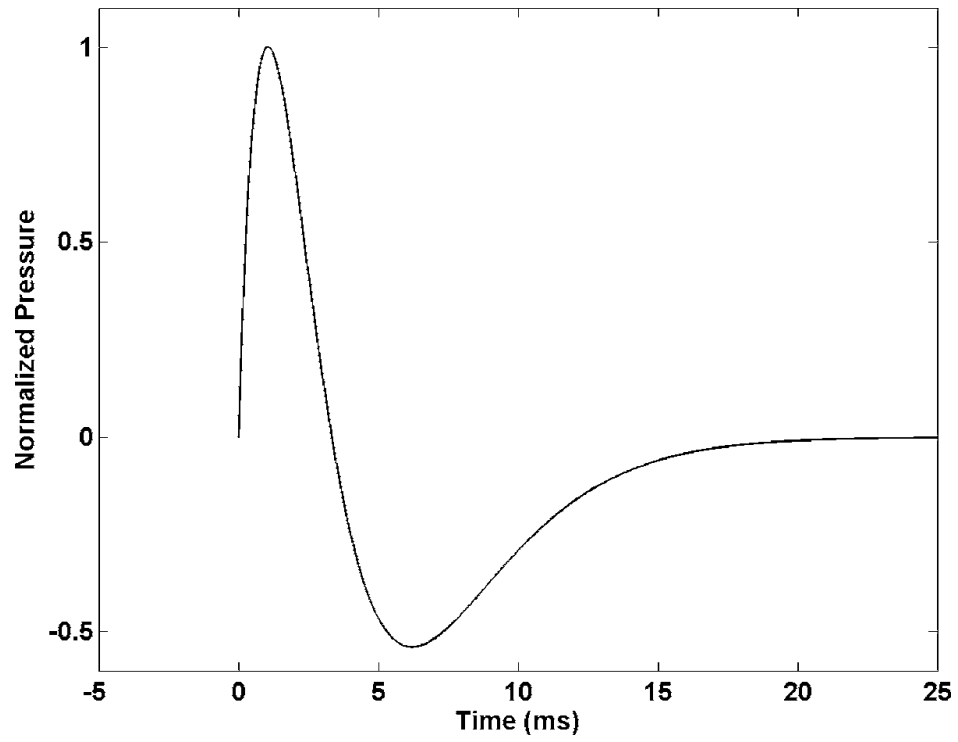


Figure 6.5: The source pulse used for modeling the blast signal produced in the experiment as calculated from Eq.(5) (Liu and Albert, 2006).

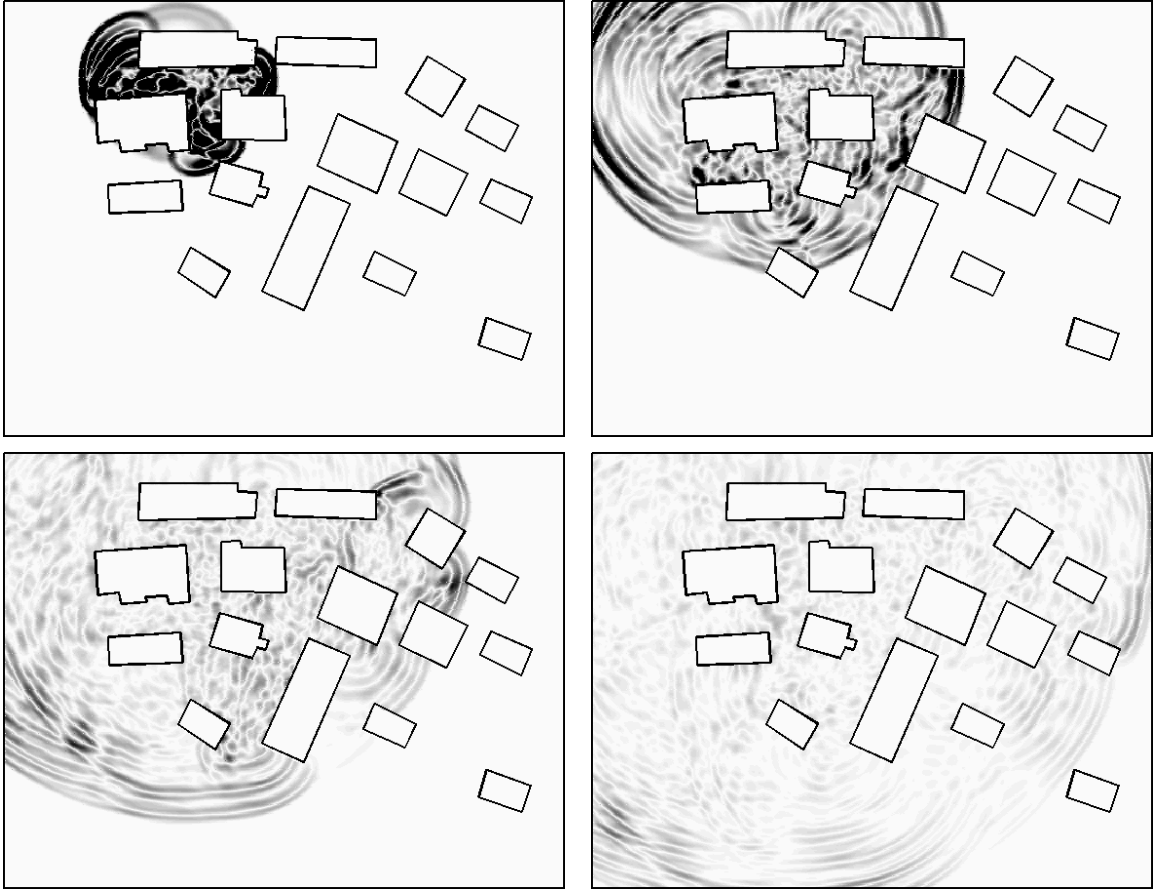


Figure 6.6: Calculated acoustic response for the source position SP1 in the artificial village scene using the ARD technique. Simulated wavefields are shown at times $t=75$, 150, 225, and 300 ms.

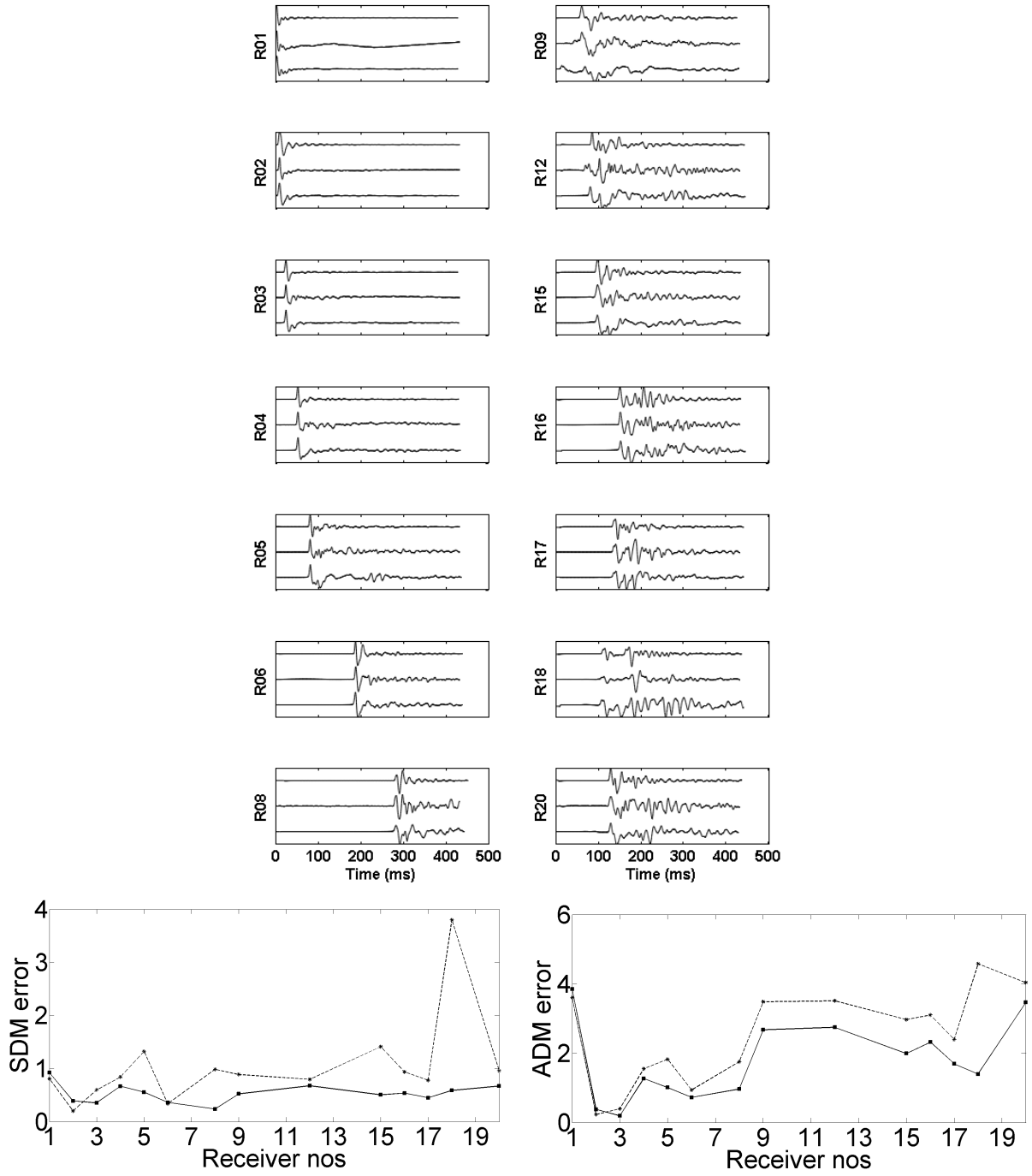


Figure 6.7: Waveforms calculated by the 3D ARD (upper) and 2D FDTD (lower) simulations and measured (middle) in the artificial village for source position 1 at 14 receiver positions. All the responses have been individually normalized and low-passed to the maximum frequency of 200 Hz. The error between the 2D FDTD simulation and measurement (dashed line) and 3D ARD simulation and measurement (solid line) has been calculated using the DTW-based SDM and ADM metric.

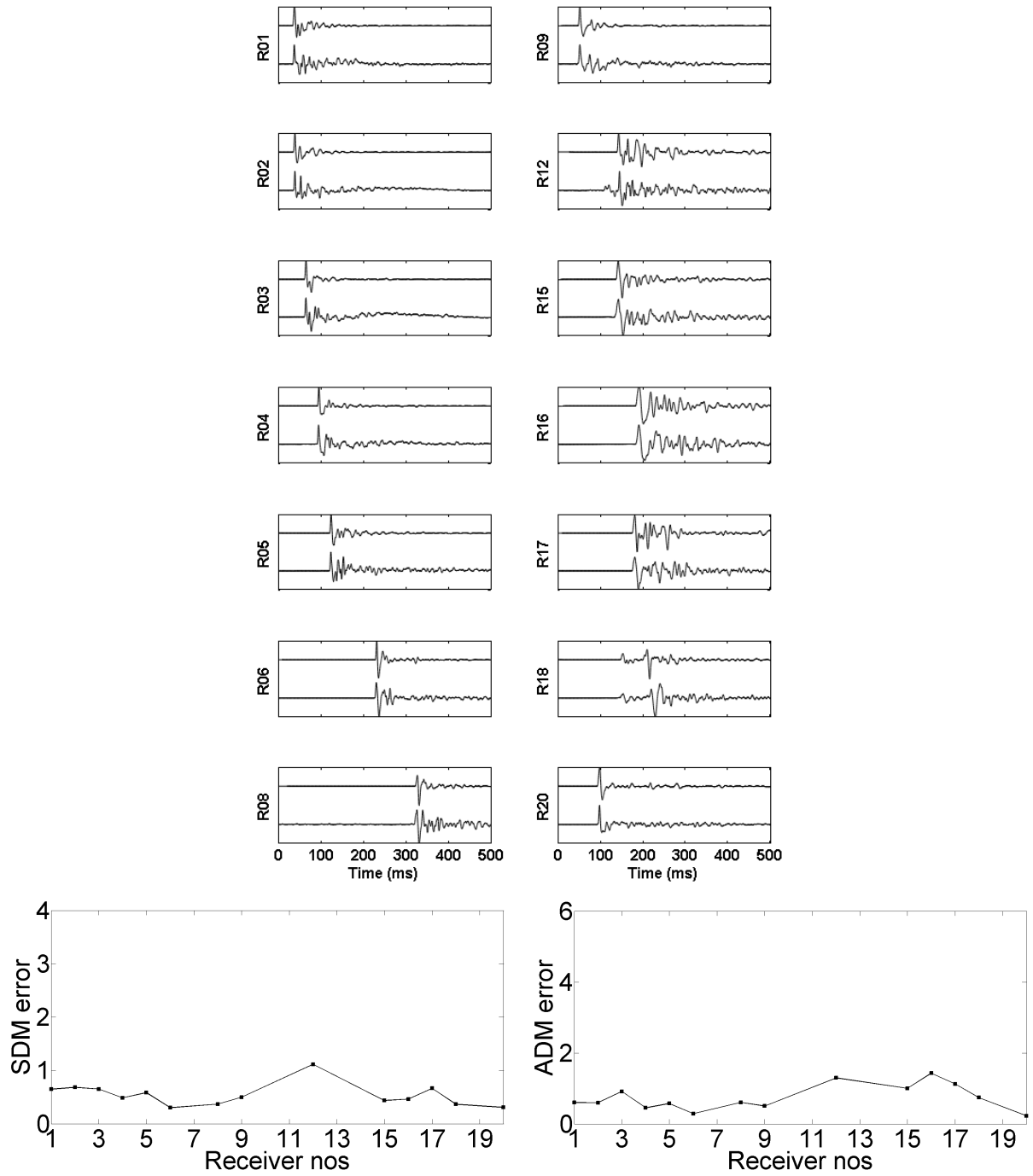


Figure 6.8: Waveforms calculated by the 3D ARD simulator (upper) and measured (lower) in the artificial village for source position 2 at 14 receiver positions. All the waveforms have been individually normalized and low-passed to the maximum frequency of 450 Hz. The error between the 3D ARD simulation and measurement has been calculated using the DTW-based SDM and ADM metric.

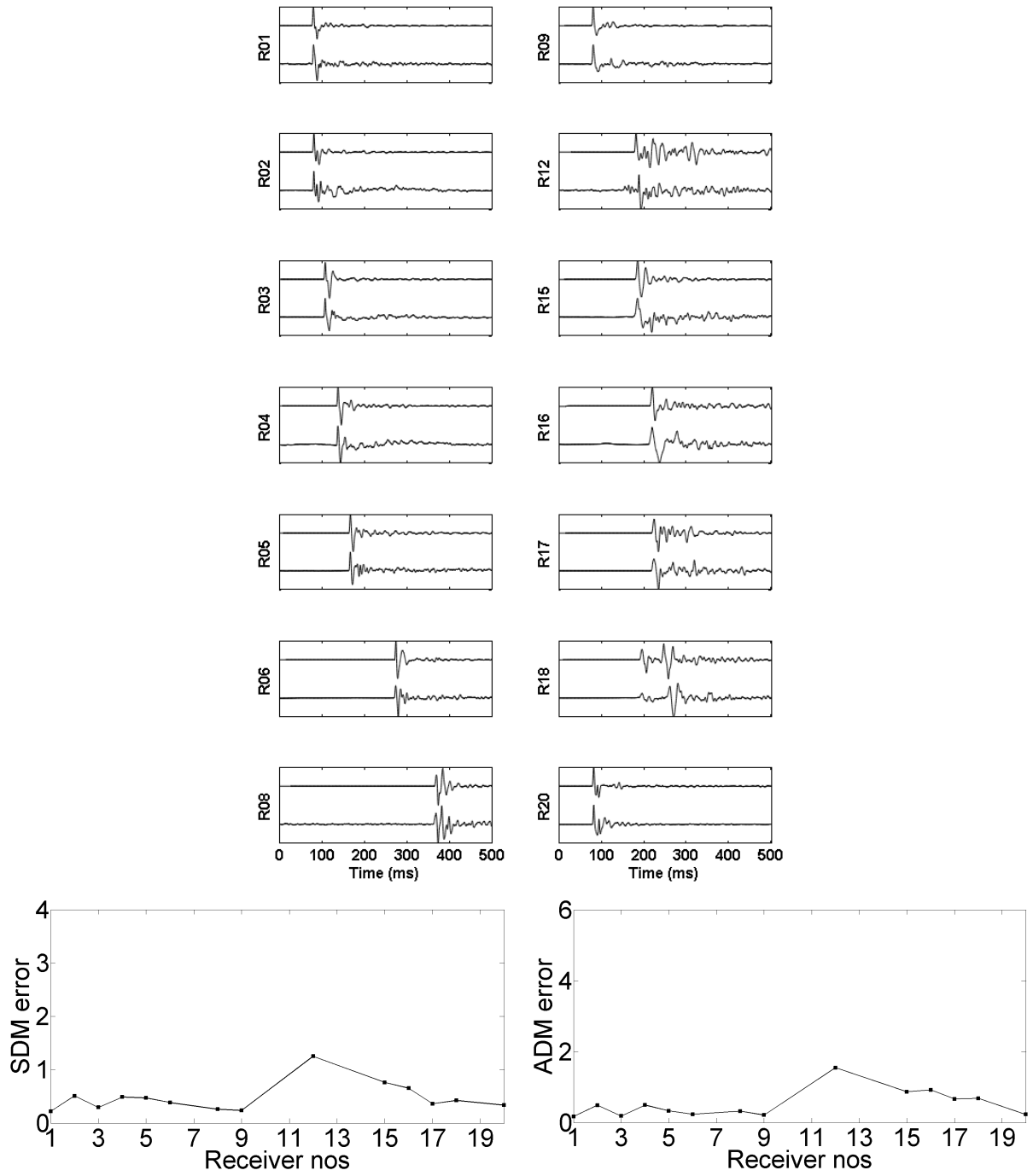


Figure 6.9: Waveforms calculated by the 3D ARD simulator (upper) and measured (lower) in the artificial village for source position 3 at 14 receiver positions. All the responses have been individually normalized and low-passed to the maximum frequency of 450 Hz. The error between the 3D ARD simulation and measurement has been calculated using the DTW-based SDM and ADM metric.

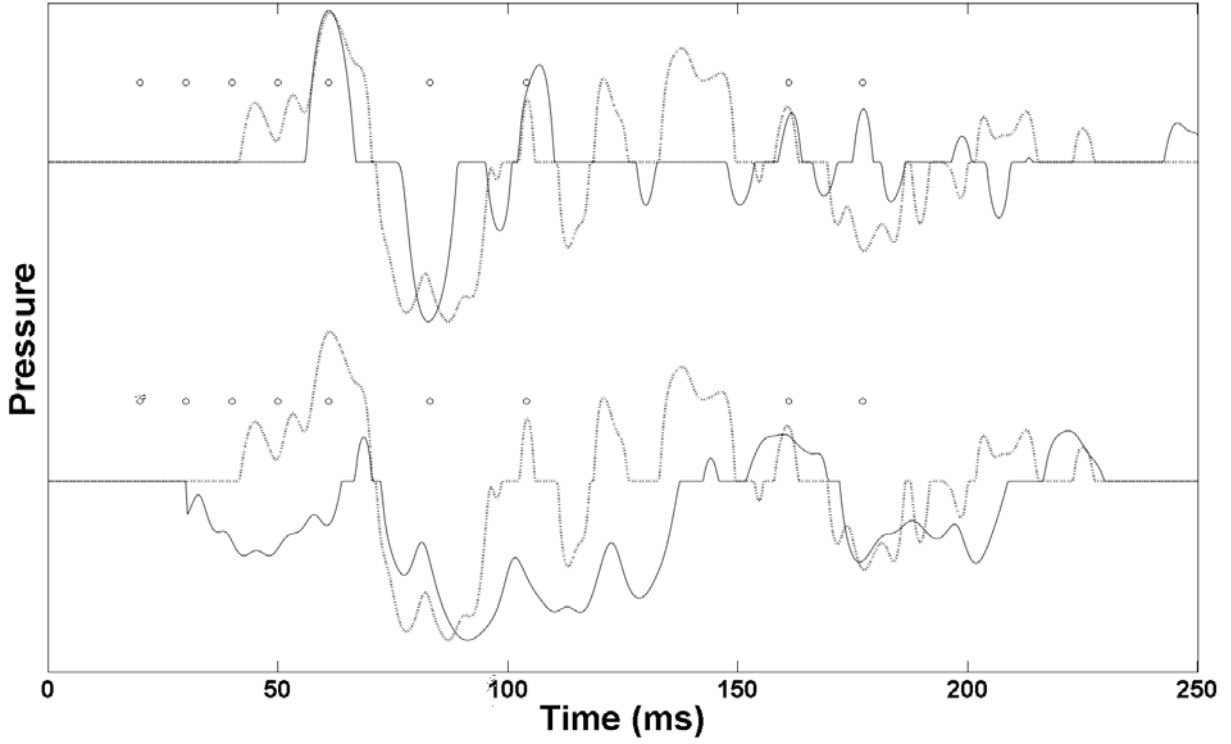


Figure 6.10: Comparison between the calculated and measured waveforms for the source position SP1 and receiver R09 behind building A. The upper trace (solid line) corresponds to the 3D ARD waveform whereas the lower trace (solid line) corresponds to the 2D FDTD waveform. The measured waveform is drawn as dotted line. Note that the 2D FDTD simulation cannot model the diffraction path from the building's rooftop resulting in the missing first arrival at 60 ms.

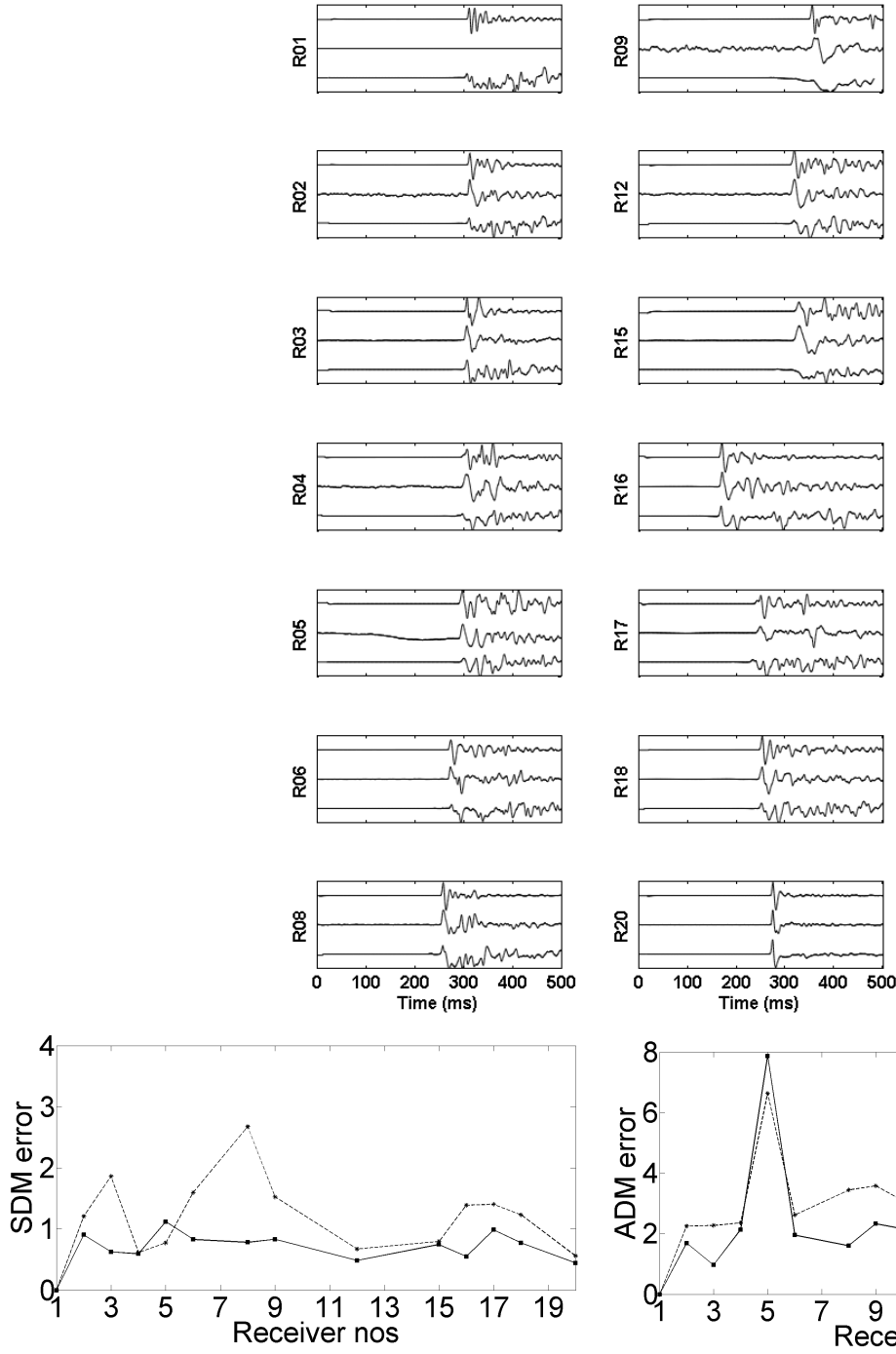


Figure 6.11: Waveforms calculated by the 3D ARD (upper) and 2D FDTD (lower) simulations and measured (middle) in the artificial village for source position 4 at 14 receiver positions. All the responses have been individually normalized and low-passed to the maximum frequency of 200 Hz. The error between the 2D FDTD simulation and measurement (dashed line) and 3D ARD simulation and measurement (solid line) has been calculated using the DTW-based SDM and ADM metric.



Figure 6.12: **Rooftop diffraction** Calculated acoustic response for the source position SP1 in the artificial village scene using the ARD technique. Receiver position R09 is marked by a white circle on the right of the first building (from right). Simulated wavefields are shown at times $t = 20, 30, 40, 50, 61, 83, 104, 161$ and 177 ms corresponding to marked points in Figure 6.10.

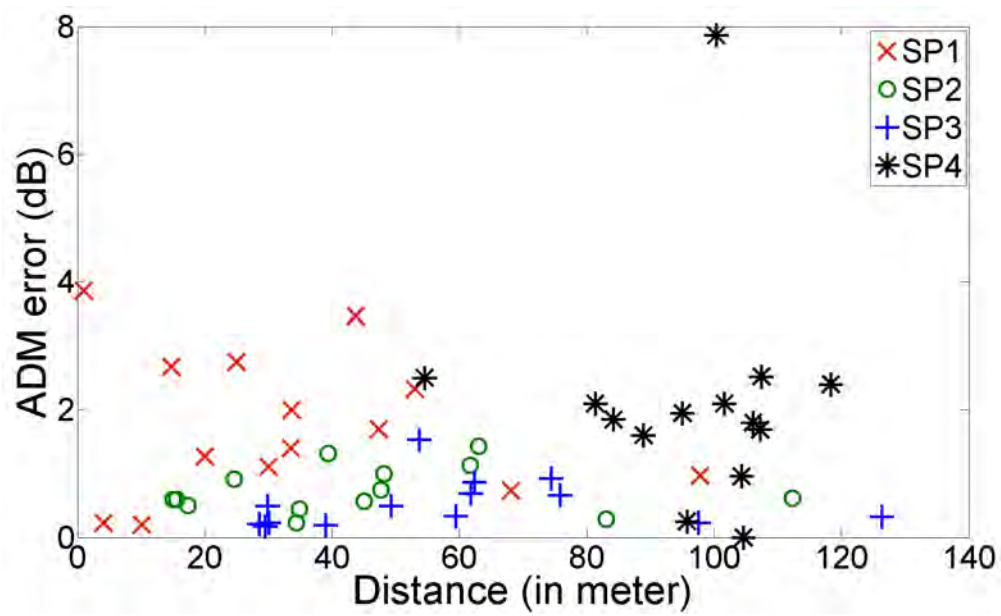


Figure 6.13: Variation of error with distance between the source-receiver positions.

Bibliography

- [1] AACH, J., AND CHURCH, G. M. Aligning gene expression time series with time warping algorithms. *Bioinformatics* 17, 6 (2001), 495–508.
- [2] ABRAMOWITZ, M., AND STEGUN, I. *Handbook of Mathematical Functions*, fifth ed. Dover, New York, 1964.
- [3] ALBERT, D. G., AND LIU, L. The effect of buildings on acoustic pulse propagation in an urban environment. *Journal of Acoustical Society of America* 127, 3 (2010), 1335–1346.
- [4] ALBERT, D. G., LIU, L., AND MORAN, M. L. Time reversal processing for source location in an urban environment. *The Journal of the Acoustical Society of America* 118, 2 (2005), 616–619.
- [5] ALBERTS, W. C. K., II, NOBLE, J. M., AND COLEMAN, M. A. Sound propagation in the vicinity of an isolated building: An experimental investigation. *The Journal of the Acoustical Society of America* 124, 2 (2008), 733–742.
- [6] ALLEN, J. B., AND BERKLEY, D. A. Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America* 65, 4 (1979), 943–950.
- [7] ALLEN, J. B., AND BERKLEY, D. A. Image method for efficiently simulating smallroom acoustics. *Journal of Acoustical Society of America* 65, 4 (1979), 943–950.
- [8] ANTANI, L., CHANDAK, A., SAVIOJA, L., AND MANOCHA, D. Interactive sound propagation using compact acoustic transfer operators. *ACM Trans. Graph.* 31, 1 (Feb. 2012), 7:1–7:12.

- [9] ANTANI, L., CHANDAK, A., TAYLOR, M., AND MANOCHA, D. Direct-to-indirect acoustic radiance transfer. *IEEE Transactions on Visualization and Computer Graphics* 18, 2 (2012), 261–269.
- [10] ARFKEN, G. B., WEBER, H. J., AND RUBY, L. *Mathematical methods for physicists*, vol. 3. Academic press San Diego, 1985.
- [11] ATTENBOROUGH, K., LI, K. M., AND HOROSHENKOV, K. *Predicting outdoor sound*. CRC Press, 2006.
- [12] ATTENBOROUGH, K., TAHERZADEH, S., BASS, H. E., DI, X., RASPET, R., BECKER, G. R., AND G A. Benchmark cases for outdoor sound propagation models. *Journal of Acoustical Society of America* 97, 1 (1995), 173–191.
- [13] AUMOND, P., GUILLAUME, G., GAUVREAU, B., LAC, C., MASSON, V., AND BÉRENGIER, M. Application of the transmission line matrix method for outdoor sound propagation modellingpart 2: Experimental validation using meteorological data derived from the meso-scale model Meso-NH. *Applied Acoustics* 76 (2014), 107–112.
- [14] BALLARD, M. S. Modeling three-dimensional propagation in a continental shelf environment. *Journal of Acoustical Society of America* 131, 3 (2012), 1969–1977.
- [15] BARBONE, P. E., MONTGOMERY, J. M., MICHAEL, O., AND HARARI, I. Scattering by a hybrid asymptotic/finite element method. *Computer methods in applied mechanics and engineering* 164, 1 (1998), 141–156.
- [16] BAXLEY, P. A., BUCKER, H., AND PORTER, M. B. Comparison of beam tracing algorithms. In *Proceedings of 5th European Conference on Underwater Acoustics* (2000).
- [17] BEN-ARTZI, A., EGAN, K., DURAND, F., AND RAMAMOORTHY, R. A precomputed polynomial representation for interactive BRDF editing with global illumination. *ACM Transactions on Graphics (TOG)* 27, 2 (2008), 13.

- [18] BERGEN, B., GENECHTEN, B. V., RENTERGHEM, T. V., PLUYMERS, B., VANDEPITTE, D., BOTTELDOOREN, D., AND DESMET, W. Comparison of numerical prediction techniques for sound propagation in complex outdoor environments. In *Proceedings of 15th International Congress on Sound and Vibration* (2008).
- [19] BERTRAM, M., DEINES, E., MOHRING, J., JEGOROV, J., AND HAGEN, H. Phonon tracing for auralization and visualization of sound. *IEEE Visualization* (2005).
- [20] BLAUERT, J. *Spatial hearing: The psychophysics of human sound localization*. MIT Press (Cambridge, Mass.), 1983.
- [21] BLUMRICH, R., AND HEIMANN, D. A linearized Eulerian sound propagation model for studies of complex meteorological effects. *Journal of Acoustical Society of America* 112, 2 (2002), 446–455.
- [22] BORISH, J. Extension of the image model to arbitrary polyhedra. *Journal of Acoustical Society of America* 75, 6 (1984), 1827–1836.
- [23] BOTTELDOOREN, D. Acoustical finite-difference time-domain simulation in a quasi-Cartesian grid. *The Journal of the Acoustical Society of America* 95, 5 (1994), 2313–2319.
- [24] BOTTELDOOREN, D. Acoustical finite-difference time domain simulation in a quasi-cartesian grid. *The Journal of the Acoustical Society of America* 95, 5 (1994), 2313–2319.
- [25] BOYD, J. P. Chebyshev and Fourier Spectral Methods. *Dover Publications, NY, USA*, (December 2001), 1–668.
- [26] BREBBIA, C. A. *Boundary element methods in acoustics*. Computational Mechanics Publications, 1991.
- [27] BREBBIA, C. A. Boundary Element Methods in Acoustics. *Springer, NY, USA*, (October 1991), 1–312.

- [28] BROWN, D. J., AND GARCÉS, M. A. Ray tracing in an inhomogeneous atmosphere with winds. In *Handbook of Signal Processing in Acoustics*. Springer, 2009, pp. 1437–1460.
- [29] BULLEN, R., AND FRICKE, F. Sound propagation in a street. *Journal of Sound and Vibration* 46, 1 (1976), 33 – 42.
- [30] CAIANI, E., PORTA, A., BASELLI, G., TURIEL, M., MUZZUPAPPA, S., PIERUZZI, F., CREMA, C., MALLIANI, A., AND CERUTTI, S. Warped-average template technique to track on a cycle-by-cycle basis the cardiac filling phases on left ventricular volume. In *Computers in Cardiology* (1998), pp. 73–76.
- [31] CHADWICK, J. N., AN, S. S., AND JAMES, D. L. Harmonic shells: a practical nonlinear sound model for near-rigid thin shells. In *SIGGRAPH Asia '09: ACM SIGGRAPH Asia 2009 papers* (New York, NY, USA, 2009), ACM, pp. 1–10.
- [32] CHANDAK, A., ANTANI, L., TAYLOR, M., AND MANOCHA, D. FastV: From-point visibility culling on complex models. *Computer Graphics Forum* 28, 4 (2009), 1237–1246.
- [33] CHANDLER-WILDE, S. The boundary element method in outdoor noise propagation. *Proc. Inst. Acoustics* 90, 2 (1997), 27–50.
- [34] CHEINET, S., EHRHARDT, L., JUVÉ, D., , AND BLANC-BENON, P. Unified modeling of turbulence effects on sound propagation. *Journal of Acoustical Society of America* 132, 4 (2012), 2198–2209.
- [35] CHENG, A., AND CHENG, D. Heritage and early history of the boundary element method. *Engineering Analysis with Boundary Elements* 29, 3 (Mar. 2005), 268–302.
- [36] COTTÉ, B., AND BLANC-BENON, P. Time-domain simulations of sound propagation in a stratified atmosphere over an impedance ground. *Journal of Acoustical Society of America* 125, 5 (2009), EL202–EL207.
- [37] DE GREVE, B., RENTERGHEM, T. V., AND BOTTELDOOREN, D. Long range fdtd over undulating terrain. In *Proceedings of Forum Acusticum* (2005), pp. 1141–1146.

- [38] DEINES, E., BERTRAM, M., MOHRING, J., JEGOROV, J., MICHEL, F., HAGEN, H., AND NIELSON, G. M. Comparative visualization for wave-based and geometric acoustics. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (2006), 1173–1180.
- [39] DINAPOLI, F. R., AND DEAVENPORT, R. L. Theoretical and numerical Green’s function field solution in a plane multilayered medium. *Journal of Acoustical Society of America* 67, 1 (1980), 92–105.
- [40] DOICU, A., EREMIN, Y. A., AND WRIEDT, T. *Acoustic and Electromagnetic Scattering Analysis Using Discrete Sources*, 1st ed. Academic Press, July 2000.
- [41] DRAGNA, D., BLANC-BENON, P., , AND POISSON, F. Time-domain solver in curvilinear coordinates for outdoor sound propagation over complex terrain. *Journal of Acoustical Society of America* 133, 6 (2013), 3751–3763.
- [42] DRAGNA, D., BLANC-BENON, P., AND POISSON, F. Time-domain simulations of outdoor sound propagation: experimental validation on a complex site.
- [43] DUSHAW, B. D., AND COLOSI, J. A. The eigenray acoustic ray tracing code, 2013.
- [44] FAIRWEATHER, G. The method of fundamental solutions for scattering and radiation problems. *Engineering Analysis with Boundary Elements* 27, 7 (July 2003), 759–769.
- [45] FUNKHOUSER, T., CARLBOM, I., ELKO, G., PINGALI, G., SONDH, M., AND WEST, J. A beam tracing approach to acoustic modeling for interactive virtual environments. In *Proc. of ACM SIGGRAPH* (1998), pp. 21–32.
- [46] FUNKHOUSER, T., CARLBOM, I., ELKO, G., PINGALI, G., SONDH, M., AND WEST, J. A beam tracing approach to acoustic modeling for interactive virtual environments. In *Proceedings of 25th Conference on Computer Graphics and Interactive Techniques* (1998).
- [47] FUNKHOUSER, T., TSINGOS, N., CARLBOM, I., ELKO, G., SONDH, M., WEST, J., PINGALI, G., MIN, P., AND NGAN, A. A beam tracing method for interactive architectural acoustics. *Journal of the Acoustical Society of America* 115, 2 (February 2004), 739–756.

- [48] GABILLET, Y., SCHROEDER, H., DAIGLE, G. A., AND LESPÉRANCE, A. Application of the gaussian beam approach to sound propagation in the atmosphere: Theory and experiments. *Journal of Acoustical Society of America* 93, 6 (1993), 3105–3116.
- [49] GARCIA, A., AND FAUS, L. Statistical analysis of noise levels in urban areas. *Applied Acoustics* 34, 4 (1991), 227 – 247.
- [50] GILBERT, K. E., AND WHITE, M. J. Application of the parabolic equation to sound propagation in a refracting atmosphere. *Journal of Acoustical Society of America* 85, 2 (1989), 630–637.
- [51] GRANIER, E., KLEINER, M., DALENBCK, B.-I., AND SVENSSON, P. Experimental auralization of car audio installations. *Journal of the Audio Engineering Society* 44, 10 (1996), 835–849.
- [52] GUILLAUME, G., AUMOND, P., GAUVREAU, B., AND DUTILLEUX, G. Application of the transmission line matrix method for outdoor sound propagation modellingpart 1: Model presentation and evaluation. *Applied Acoustics* 76 (2014), 113–118.
- [53] GUMEROV, N. A., AND DURAISWAMI, R. *Fast multipole methods for the Helmholtz equation in three dimensions*. Elsevier Science, 2004.
- [54] GUMEROV, N. A., AND DURAISWAMI, R. A broadband fast multipole accelerated boundary element method for the three dimensional Helmholtz equation. *The Journal of the Acoustical Society of America* 125, 1 (2009), 191–205.
- [55] GUMEROV, N. A., AND DURAISWAMI, R. Wideband fast multipole accelerated boundary element methods for the threedimensional helmholtz equation. *The Journal of the Acoustical Society of America* 125, 4 (2009).
- [56] HAMPEL, S., LANGER, S., AND CISILINO, A. P. Coupling boundary elements to a raytracing procedure. *International journal for numerical methods in engineering* 73, 3 (2008), 427–445.
- [57] HASEGAWA, T., OCHI, M., AND MATSUZAWA, K. Acoustic radiation pressure on a rigid sphere in a spherical wave field. *The Journal of the Acoustical Society of America* 67, 3 (1980), 770–773.

- [58] HAVRAN, V. *Heuristic Ray Shooting Algorithms*. PhD thesis, Faculty of Electrical Engineering, Czech Technical University, Prague, 2000.
- [59] HEILPERN, T., HEYMAN, E., AND TIMCHENKO, V. A beam summation algorithm for wave radiation and guidance in stratified media. *Journal of Acoustical Society of America* 121, 4 (2007), 1856–1864.
- [60] HEIMANN, D. Wide-area assessment of topographical and meteorological effects on sound propagation by time-domain modeling. *Journal of Acoustical Society of America* 133, 5 (2013), EL419–EL425.
- [61] HEIMANN, D., AND KARLE, R. A linearized euler finite-difference time-domain sound propagation model with terrain-following coordinates. *Journal of Acoustical Society of America* 119, 6 (2006), 3813–3821.
- [62] HEUTSCHI, K. A simple method to evaluate the increase of traffic noise emission level due to buildings, for a long straight street. *Applied Acoustics* 44, 3 (1995), 259 – 274.
- [63] HOBSON, E. W. *The Theory of Spherical and Ellipsoidal Harmonics*. Cambridge University Press, New York, NY, USA, 1955.
- [64] HOFMANN, J., AND HEUTSCHI, K. Simulation of outdoor sound propagation with a transmission line matrix method. *Applied Acoustics* 68, 2 (2007), 158–172.
- [65] HOLE, L. R. An experimental and theoretical study of propagation of acoustic pulses in a strongly refracting atmosphere. *Applied Acoustics* 53, 1 (1998), 77–94.
- [66] HORNIKX, M., AND WAXLER, R. The extended Fourier pseudospectral time-domain (pstd) method for fluid media with discontinuous properties. *Journal of Acoustical Society of America* 18, 04 (2010), 297–319.
- [67] HORNIKX, M., WAXLER, R., AND FORSSÉN, J. The extended Fourier pseudospectral time-domain method for atmospheric sound propagation. *Journal of Acoustical Society of America* 128, 4 (2010), 1632–1646.

- [68] HORNIKX, M., WAXLER, R., AND FORSSEN, J. The extended fourier pseudospectral time-domain method for atmospheric sound propagation. *The Journal of the Acoustical Society of America* 128, 4 (2010), 1632–1646.
- [69] HOTHERSALL, D., CHANDLER-WILDE, S., AND HAJMIRZAE, M. Efficiency of single noise barriers. *Journal of Sound and Vibration* 146, 2 (1991), 303 – 322.
- [70] IU, K. K., AND LI, K. M. The propagation of sound in narrow street canyons. *The Journal of the Acoustical Society of America* 112, 2 (2002), 537–550.
- [71] JAMES, D. L., BARBIČ, J., AND PAI, D. K. Precomputed acoustic transfer: output-sensitive, accurate sound generation for geometrically complex vibration sources. In *ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), SIGGRAPH '06, ACM, pp. 987–995.
- [72] JEAN, P., NOE, N., AND GAUDAIRE, F. Calculation of tyre noise radiation with a mixed approach. *Acta Acustica united with Acustica* 94, 1 (2008), 91–103.
- [73] JENSEN, F. B., KUPERMAN, W. A., PORTER, M. B., AND SCHMIDT, H. *Computational ocean acoustics*. Springer, 2011.
- [74] JONES, R. M., RILEY, J. P., AND GEORGES, T. M. HARPA: A versatile three-dimensional hamiltonian ray-tracing program for acoustic waves in the atmosphere above irregular terrain, 1986.
- [75] JONES, R. M., RILEY, J. P., AND GEORGES, T. M. HARPO: A versatile three-dimensional hamiltonian ray-tracing program for acoustic waves in an ocean with irregular bottom, 1986.
- [76] KANG, J. Sound propagation in street canyons: Comparison between diffusely and geometrically reflecting boundaries. *The Journal of the Acoustical Society of America* 107, 3 (2000), 1394–1404.
- [77] KANG, J. Sound propagation in interconnected urban streets: a parametric study. *Environment and Planning B: Planning and Design* 28, 2 (2001), 281–294.
- [78] KANG, J. Urban sound environment. *Taylor & Francis, NY, USA*, (2006), 1–304.

- [79] KAPRALOS, B., JENKIN, M., AND MILIOS, E. Sonel mapping: a probabilistic acoustical modeling method. *Building Acoustics* 15, 4 (2008), 289–313.
- [80] KETCHAM, S. A., WILSON, D. K., PARKER, M. W., AND CUDNEY, H. H. Signal fading curves from computed urban acoustic wave fields. *Proc. SPIE 6963* (2008), 69630K–69630K–8.
- [81] KLOSNER, J. M., FELSEN, L. B., LU, I. T., AND GROSSFELD, H. Threedimensional source field modeling by selfconsistent gaussian beam superposition. *Journal of Acoustical Society of America* 91, 4 (1992), 1809–1822.
- [82] KOUYOUMJIAN, R. G., AND PATHAK, P. H. A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface. *Proceedings of the IEEE* 62, 11 (1974), 1448–1461.
- [83] KROKSTAD, A., STROM, S., AND SORSDAL, S. Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibration* 8, 1 (July 1968), 118–125.
- [84] KROPP, W., AND SVENSSON, P. U. Application of the time domain formulation of the method of equivalent sources to radiation and scattering problems. *Acta Acustica united with Acustica* 81, 6 (1995), 528–543.
- [85] KUTTRUFF, H. *Room acoustics*. CRC Press, 2009.
- [86] LEDECZI, A., VOLGYESI, P., MAROTI, M., SIMON, G., BALOGH, G., NADAS, A., KUSY, B., DORA, S., AND PAP, G. Multiple simultaneous acoustic source localization in urban terrain. In *Fourth International Symposium on Information Processing in Sensor Networks* (2005), pp. 491–496.
- [87] LEE, K. P., AND DAVIES, H. G. Nomogram for estimating noise propagation in urban areas. *The Journal of the Acoustical Society of America* 57, 6 (1975), 1477–1480.
- [88] LENTZ, T., SCHRÖDER, D., VORLÄNDER, M., AND ASSENMACHER, I. Virtual reality system with integrated sound field simulation and reproduction. *EURASIP J. Appl. Signal Process.* 2007, 1 (2007), 187.

- [89] LI, K. M., TAHERZADEH, S., AND ATTENBOROUGH, K. An improved ray-tracing algorithm for predicting sound propagation outdoors. *Journal of Acoustical Society of America* 104, 4 (1998), 2077–2083.
- [90] LIU, L., AND ALBERT, D. G. Acoustic pulse propagation near a right-angle wall. *Journal of Acoustical Society of America* 119, 4 (2006), 2073–2083.
- [91] LIU, Q. H. The PSTD algorithm: A time-domain method combining the pseudospectral technique and perfectly matched layers. *The Journal of the Acoustical Society of America* 101, 5 (1997), 3182.
- [92] LIU, Y., SHEN, L., AND BAPAT, M. Development of the fast multipole boundary element method for acoustic wave problems. In *Recent Advances in Boundary Element Methods*. Springer Netherlands, 2009, pp. 287–303.
- [93] LUO, W., AND HENRIK, S. Three-dimensional propagation and scattering around a conical seamount. *Journal of Acoustical Society of America* 125, 1 (2009), 52–65.
- [94] MANOCHA, D., CALAMIA, P., LIN, M. C., MANOCHA, D., SAVIOJA, L., AND TSINGOS, N. Interactive sound rendering. In *ACM SIGGRAPH 2009 Courses* (New York, NY, USA, 2009), SIGGRAPH '09, ACM, pp. 15:1–15:338.
- [95] MASTERSON, C., KEARNEY, G., AND BOLAND, F. Acoustic impulse response interpolation for multichannel systems using dynamic time warping. In *35th AES Conference on Audio for Games* (2009), p. 34.
- [96] MAY, D., AND OSMAN, N. Highway noise barriers: new shapes. *Journal of Sound and Vibration* 71, 1 (1980), 73 – 101.
- [97] MEDWIN, H., CHILDS, E., AND JEBSEN, G. M. Impulse studies of double diffraction: A discrete huygens interpretation. *The Journal of the Acoustical Society of America* 72, 3 (1982), 1005–1013.

- [98] MEHRA, R., RAGHUVANSHI, N., ANTANI, L., CHANDAK, A., CURTIS, S., AND MANOCHA, D. Wave-based sound propagation in large open scenes using an equivalent source formulation. *ACM Trans. Graph.* 32, 2 (Apr. 2013), 19:1–19:13.
- [99] MEHRA, R., RAGHUVANSHI, N., CHANDAK, A., ALBERT, D. G., WILSON, D. K., AND MANOCHA, D. Acoustic pulse propagation in an urban environment using a three-dimensional numerical simulation. *Journal of Acoustical Society of America* 135, 6 (2014), 3231–3242.
- [100] MEHRA, R., RAGHUVANSHI, N., SAVIOJA, L., LIN, M. C., AND MANOCHA, D. An efficient gpu-based time domain solver for the acoustic wave equation. *Applied Acoustics* 73, 2 (2012), 83 – 94.
- [101] MO, Q., YEH, H., AND MANOCHA, D. Tracing analytic ray curves for light and sound propagation in non-linear media, arxiv preprint arxiv:1409.2235, 2014.
- [102] MYERS, M. K., AND MCANINCH, G. L. Parabolic approximation for sound propagation in the atmosphere. *Journal of AIAA* 16, 8 (1978), 836–842.
- [103] NOMURA, T., TAKAGI, K., AND SATO, S. Finite element simulation of sound propagation concerning meteorological conditions. *International Journal for Numerical Methods in Fluids* 64, 10-12 (2010), 1296–1318.
- [104] OCHMANN, M. The source simulation technique for acoustic radiation problems. *Acta Acustica united with Acustica* 81, 6 (1995), 512–527.
- [105] OCHMANN, M. The source simulation technique for acoustic radiation problems. *Acta Acustica united with Acustica* 81, 6 (1995), 512–527.
- [106] OCHMANN, M. The full-field equations for acoustic radiation and scattering. *The Journal of the Acoustical Society of America* 105, 5 (1999), 2574–2584.
- [107] OF NAVAL RESEARCH, U. O. Ocean acoustics library: Ray models/software, 2014.

- [108] OGREN, M., AND KROPP, W. Road traffic noise propagation between two dimensional city canyons using an equivalent sources approach. *Acta Acustica united with Acustica* 90, 2 (2004), 293–300.
- [109] OSHIMA, T., IMANO, M., HIRAGURI, Y., AND KAMOSHIDA, Y. Linearized euler simulations of sound propagation with wind effects over a reconstructed urban terrain using digital geographic information. *Applied Acoustics* 74, 12 (2013), 1354–1366.
- [110] OSTASHEV, V. E., WILSON, D. K., LIU, L., ALDRIDGE, D. F., AND SYMONS, N. P. Equations for finite-difference, time-domain simulation of sound propagation in moving inhomogeneous media and numerical implementation. *Journal of Acoustical Society of America* 117, 2 (2005), 503–517.
- [111] OVENDEN, N. C., SHAFFER, S. R., AND FERNANDO, H. J. S. Impact of meteorological conditions on noise propagation from freeway corridors. *The Journal of the Acoustical Society of America* 126, 1 (2009), 25–35.
- [112] PANOFSKY, H. A., AND DUTTON, J. A. *Atmospheric turbulence. Models and methods for engineering applications*. Wiley, 1984.
- [113] PAVIC, G. A technique for the computation of sound radiation by vibrating bodies using multipole substitute sources. *Acta Acustica united with Acustica* 92 (2006), 112–126(15).
- [114] PICAUT, J., POLLS, T. L., LHERMITE, P., AND GARY, V. Experimental study of sound propagation in a street. *Applied Acoustics* 66, 2 (2005), 149 – 173.
- [115] PIERCE, A. D. *Acoustics: an introduction to its physical principles and applications*. McGraw-Hill, 1981.
- [116] PIERCE, A. D. *Acoustics: An Introduction to Its Physical Principles and Applications*. Acoustical Society of America, 1989.
- [117] POLLÈS, T. L., PICAUT, J., BÉRENGIER, M., AND BARDOS, C. Sound field modeling in a street canyon with partially diffusely reflecting boundaries by the transport theory. *The Journal of the Acoustical Society of America* 116, 5 (2004), 2969–2983.

- [118] PORTER, M. B. The BELLHOP manual and users guide: Preliminary draft, 2011.
- [119] PREMAT, E., AND GABILLET, Y. A new boundary-element method for predicting outdoor sound propagation and application to the case of a sound barrier in the presence of downward refraction. *Journal of Acoustical Society of America* 108, 6 (2000), 2775–2783.
- [120] PREMAT, E., AND GABILLET, Y. A new boundary-element method for predicting outdoor sound propagation and application to the case of a sound barrier in the presence of downward refraction. *The Journal of the Acoustical Society of America* 108, 6 (2000), 2775–2783.
- [121] PROSPATHOPOULOS, J. M., AND VOUTSINAS, S. G. Determination of equivalent sound speed profiles for ray tracing in near-ground sound propagation. *Journal of Acoustical Society of America* 122, 3 (2007), 1391–1403.
- [122] RAGHUVANSHI, N., MEHRA, R., MANOCHA, D., AND LIN, M. C. Adaptive rectangular decomposition: A spectral, domain-decomposition approach for fast wave solution on complex scenes. *The Journal of the Acoustical Society of America* 132, 3 (2012), 1890–1890.
- [123] RAGHUVANSHI, N., NARAIN, R., , AND LIN, M. C. Efficient and accurate sound propagation using adaptive rectangular decomposition. *IEEE Transactions on Visualization and Computer Graphics* 15, 5 (2009), 789–801.
- [124] RAGHUVANSHI, N., NARAIN, R., AND LIN, M. C. Efficient and Accurate Sound Propagation Using Adaptive Rectangular Decomposition. *IEEE Transactions on Visualization and Computer Graphics* 15, 5 (2009), 789–801.
- [125] RAGHUVANSHI, N., SNYDER, J., MEHRA, R., LIN, M. C., AND GOVINDARAJU, N. K. Precomputed Wave Simulation for Real-Time Sound Propagation of Dynamic Sources in Complex Scenes. *SIGGRAPH 2010* 29, 3 (July 2010).
- [126] REILLY, S. M., POTTY, G., AND GOODRICH, M. WaveQ3D: Fast and accurate acoustic transmission loss (TL) eigenrays, in littoral environments, for active sonar simulation/stimulation systems, 2014.

- [127] RENTERGHEM, T. V., BOTTELDOOREN, D., AND LERCHER, P. Comparison of measurements and predictions of sound propagation in a valley-slope configuration in an inhomogeneous atmosphere. *Journal of Acoustical Society of America* 121, 5 (2007), 2522–2533.
- [128] RENTERGHEM, T. V., SALOMONS, E. M., AND BOTTELDOOREN, D. Efficient FDTD-PE model for sound propagation in situations with complex obstacles and wind profiles. *Acta Acustica united with acustica* 91, 4 (2005), 671–679.
- [129] RENTERGHEM, T. V., SALOMONS, E. M., AND BOTTELDOOREN, D. Parameter study of sound propagation between city canyons with a coupled FDTD-PE model. *Applied Acoustics* 67, 6 (2006), 487–510.
- [130] ROBERTS, B. G. Horizontal-gradient acoustical ray-trace program TRIMAIN, 1974.
- [131] SAKAMOTO, S., USHIYAMA, A., AND NAGATOMO, H. Numerical analysis of sound propagation in rooms using the finite difference time domain method. *The Journal of the Acoustical Society of America* 120, 5 (2006), 3008.
- [132] SAKOE, H., AND CHIBA, S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing* 26, 1 (1978), 43–49.
- [133] SALOMONS, E. M. *Computational atmospheric acoustics*. Kluwer Academic Publishers, 2001.
- [134] SAMPLE, J. D. Gaussian models for complex sound sources in the paraxial region. *Journal of Acoustical Society of America* 84, 6 (1988), 2252–2261.
- [135] SAVIOJA, L. Real-Time 3D Finite-Difference Time-Domain Simulation of Mid-Frequency Room Acoustics. In *13th International Conference on Digital Audio Effects (DAFx-10)* (Sept. 2010).
- [136] SCHISLER, C., MEHRA, R., AND MANOCHA, D. High-order diffraction and diffuse reflections for interactive sound propagation in large environments. *ACM Transactions on Graphics* 32, 4 (2014), 39.

- [137] SILTANEN, S., LOKKI, T., KIMINKI, S., AND SAVIOJA, L. The room acoustic rendering equation. *Journal of Acoustical Society of America* 122, 3 (2007), 1624–1635.
- [138] SILTANEN, S., LOKKI, T., KIMINKI, S., AND SAVIOJA, L. The room acoustic rendering equation. *The Journal of the Acoustical Society of America* 122, 3 (Sept. 2007), 1624–1635.
- [139] SILTANEN, S., LOKKI, T., AND SAVIOJA, L. Frequency domain acoustic radiance transfer for real-time auralization. *Acta Acustica united with Acustica* 95, 1 (2009), 106–117.
- [140] SOUTHERN, A., SILTANEN, S., AND SAVIOJA, L. Spatial room impulse responses with a hybrid modeling method. In *Audio Engineering Society Convention 130* (5 2011).
- [141] STULL, R. B. *An introduction to boundary layer meteorology*, vol. 13. Springer, 1988.
- [142] STURM, F., AND KORAKAS, A. Comparisons of laboratory scale measurements of three-dimensional acoustic propagation with solutions by a parabolic equation model. *Journal of Acoustical Society of America* 133, 1 (2013), 108–118.
- [143] SVENSSON, U. P., FRED, R. I., AND VANDERKOOY, J. An analytic secondary source model of edge diffraction impulse responses . *Acoustical Society of America Journal* 106 (Nov. 1999), 2331–2344.
- [144] SVENSSON, U. P., FRED, R. I., AND VANDERKOOY, J. An analytic secondary source model of edge diffraction impulse responses. *The Journal of the Acoustical Society of America* 106, 5 (1999), 2331–2344.
- [145] TAFLOVE, A., AND HAGNESS, S. C. Computational Electrodynamics: The Finite-Difference Time-Domain Method. *Artech House Publishers, London, UK and Boston, USA*, (June 2005), 1–1006.
- [146] TAKALA, T., AND HAHN, J. Sound rendering. *SIGGRAPH Comput. Graph.* 26, 2 (July 1992), 211–220.
- [147] TARALDSEN, G., AND JONASSON, H. Aspects of ground effect modeling. *The Journal of the Acoustical Society of America* 129, 1 (2011), 47–53.

- [148] TAYLOR, M., CHANDAK, A., ANTANI, L., AND MANOCHA, D. Resound: interactive sound rendering for dynamic virtual environments. In *Proceedings of 17th ACM international conference on Multimedia* (2009).
- [149] TAYLOR, M., CHANDAK, A., MO, Q., LAUTERBACH, C., SCHISLER, C., AND MANOCHA, D. Guided multiview ray tracing for fast auralization. *IEEE Transactions on Visualization and Computer Graphics* 18, 11 (2012), 1797–1810.
- [150] TAYLOR, M., CHANDAK, A., QI MO, LAUTERBACH, C., SCHISLER, C., AND MANOCHA, D. Guided multiview ray tracing for fast auralization. *Visualization and Computer Graphics, IEEE Transactions on* 18, 11 (Nov. 2012), 1797–1810.
- [151] TAYLOR, M. T., CHANDAK, A., ANTANI, L., AND MANOCHA, D. RESound: interactive sound rendering for dynamic virtual environments. In *MM '09: ACM Multimedia* (New York, NY, USA, 2009), ACM, pp. 271–280.
- [152] THOMPSON, L. L. A review of finite-element methods for time-harmonic acoustics. *The Journal of the Acoustical Society of America* 119, 3 (2006), 1315–1330.
- [153] THOMPSON, L. L. A review of finite-element methods for time-harmonic acoustics. *The Journal of the Acoustical Society of America* 119, 3 (2006), 1315–1330.
- [154] THOMPSON, L. L., AND PINSKY, P. M. *Acoustics*. John Wiley & Sons, Ltd, 2004.
- [155] TSINGOS, N. Pre-computing geometry-based reverberation effects for games. In *35th AES Conference on Audio for Games* (2009).
- [156] TSINGOS, N., DACHSBACHER, C., LEFEBVRE, S., AND DELLEPIANE, M. Instant Sound Scattering. In *Rendering Techniques (Proceedings of the Eurographics Symposium on Rendering)* (2007).
- [157] TSINGOS, N., FUNKHOUSER, T., NGAN, A., AND CARLBOM, I. Modeling acoustics in virtual environments using the uniform theory of diffraction. In *SIGGRAPH '01* (New York, NY, USA, 2001), ACM, pp. 545–552.

- [158] TSINGOS, N., FUNKHOUSER, T., NGAN, A., AND CARLBOM, I. Modeling acoustics in virtual environments using the uniform theory of diffraction. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 2001), SIGGRAPH '01, ACM, pp. 545–552.
- [159] TSINGOS, N., AND GASCUEL, J. D. A general model for the simulation of room acoustics based on hierarchical radiosity. In *ACM SIGGRAPH 97* (New York, NY, USA, 1997), SIGGRAPH '97, ACM.
- [160] TSINGOS, N., AND GASCUEL, J. D. A general model for the simulation of room acoustics based on hierarchical radiosity. *ACM SIGGRAPH technical sketch* (1997).
- [161] ČERVENÝ, V. *Seismic ray theory*. Cambridge University Press, 2005.
- [162] VAN RENTERGHEM, T., SALOMONS, E., AND BOTTELDOOREN, D. Parameter study of sound propagation between city canyons with a coupled FDTD-PE model. *Applied Acoustics* 67, 6 (2006), 487–510.
- [163] VECHERIN, S. N., WILSON, D. K., AND OSTASHEV, V. E. Incorporating source directionality into outdoor sound propagation calculations. *Journal of Acoustical Society of America* 130, 6 (2011), 3608–3622.
- [164] VENDHAN, C. P., DIWAN, G. C., AND BHATTACHARYYA, S. K. Finite-element modeling of depth and range dependent acoustic propagation in oceanic waveguides. *Journal of Acoustical Society of America* 127, 6 (2010), 3319–3326.
- [165] VLADIMIROV, V. S. Generalized functions in mathematical physics. *Moscow Izdatel Nauka* 1 (1976).
- [166] VORLANDER, M. Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *The Journal of the Acoustical Society of America* 86, 1 (1989), 172–178.
- [167] VORLÄNDER, M. Simulation of the transient and steady-state sound propagation in rooms using a new combined raytracing/imagesource algorithm. *Journal of Acoustical Society of America* 86, 1 (1989), 172–178.

- [168] WANG, P., ZHANG, L., AND LI, V. O. K. A stratified acoustic model accounting for phase shifts for underwater acoustic networks. *Sensors* 13, 5 (2013), 6183–6203.
- [169] WANG, Y., SAFAVI-NAEINI, S., AND CHAUDHURI, S. 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 2000), 743 –754.
- [170] WATERMAN, P. C. T-matrix methods in acoustic scattering. *The Journal of the Acoustical Society of America* 125, 1 (2009), 42–51.
- [171] WATERMAN, P. C. T-matrix methods in acoustic scattering. *The Journal of the Acoustical Society of America* 125, 1 (2009), 42–51.
- [172] WATSON, W. H., AND MCGIRR, R. W. RAYWAVE II: A propagation loss model for the analysis of complex ocean environments, 1975.
- [173] WIENER, F. M., MALME, C. I., AND GOGOS, C. M. Sound propagation in urban areas. *The Journal of the Acoustical Society of America* 37, 4 (1965), 738–747.
- [174] WILSON, D. K. Outdoor sound propagation : Recent modeling developments and applications to noise control. *Internoise*, (2011), K–4.
- [175] WILSON, D. K., AND LIU, L. Finite-difference, time-domain simulation of sound propagation in a dynamic atmosphere. *Cold Regions Research and Engineering Lab, Hanover NH*, No. ERDC/CRREL-TR-04-12. (2004).
- [176] WILSON, D. K., SYMONS, N. P., PATTON, E. G., SULLIVAN, P. P., MARLIN, D. H., ALDRIDGE, D. F., OSTASHEV, V. E., KETCHAM, S. A., ANDREAS, E. L., AND COLLIER, S. L. Simulation of sound propagation through high-resolution atmospheric boundary layer turbulence fields. In *Proceedings of 16th Symposium on Boundary Layers and Turbulence* (2004).

- [177] Y ALPERA, S. Q., JACOBSEN, F., LLER JUHL, P. M., AND HENRÍQUEZ, V. C. A bem approach to validate a model for predicting sound propagation over non-flat terrain. *Applied Acoustics* 64, 8 (2003), 781–791.
- [178] YEE, K. 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.
- [179] YEH, H., MEHRA, R., REN, Z., ANTANI, L., MANOCHA, D., AND LIN, M. C. Wave-ray coupling for interactive sound propagation in large complex scenes. *ACM Transactions on Graphics* 32, 6 (2013), 165.
- [180] ZAMAN, F. D., AND AL-MUHIAMEED, Z. I. Acoustic waves in a layered inhomogeneous ocean. *Applied Acoustics* 61, 4 (2000), 427–440.
- [181] ZHENG, C., AND JAMES, D. L. Rigid-body fracture sound with precomputed soundbanks. In *SIGGRAPH ’10: ACM SIGGRAPH 2010 papers* (New York, NY, USA, 2010), ACM, pp. 1–13.
- [182] ZHENG, Z. C., AND LI, W. Numerical stabilities and boundary conditions in time-domain eulerian simulations of acoustic wave propagations with and without background flow. *Applied Mathematics and Computation* 202, 1 (2008), 146–161.
- [183] ZIENKIEWICZ, O. C., TAYLOR, R. L., AND NITHIARASU, P. *The finite element method for fluid dynamics*, 6 ed. Butterworth-Heinemann, Jan. 2006.