

ABSTRACT

Title of Dissertation: **NUMERICAL ACOUSTICS FOR PHYSICAL
AND SIMULATED ENVIRONMENTS**

Shoken Kaneko
Doctor of Philosophy, 2023

Dissertation Directed by: **Professor Ramani Duraiswami**
Department of Computer Science

Computer modeling and numerical analysis of acoustical phenomena have important applications including manufacturing, audio technologies in immersive multimedia, and machine learning systems involving audio. The focus of the present dissertation is the exploration of numerical methods for modeling, simulating, synthesizing, estimating, processing, controlling, and analyzing acoustical phenomena in the physical world as well as its applications to the virtual world, i.e. immersive technologies for creating virtual, augmented, and extended realities.

The dissertation is structured as follows. In chapter [1](#), I introduce some fundamentals and basic concepts of numerical acoustics and discuss existing practical problems in acoustics. In chapter [2](#) and chapter [3](#), I propose two novel techniques for three-dimensional sound field capturing and encoding for immersive audio applications, which are both based on (semi-)analytical cancellation of scattering caused by microphone arrays mounted on acoustic scatterers. In chapter [4](#) and chapter [5](#), I introduce a fast algorithm for synthesizing

acoustic impulse responses in large-scale forests, and use it to predict the performance of acoustic wildlife monitoring systems based on large-scale distributed microphone arrays. In chapter 6, I propose a novel general-purpose individual-agnostic binaural localizer which supports sound source localization from arbitrary directions without a priori knowledge of the process generating the binaural signal. In chapter 7 and chapter 8, I develop frameworks for regularized active sound control, using either point- or mode-control and using either distributed or local worn loudspeaker and microphone arrays with applications including speech privacy, personal active noise control, and local crosstalk cancellation with limited noise injection into the environment. In chapter 9, chapter 10 and chapter 11, three numerical methods for evaluating integrals arising in the (fast multipole accelerated) boundary element method are introduced. In chapter 9, a recursive algorithm is developed which allows efficient analytical evaluation of singular and nearly singular layer potential integrals arising in the boundary element method using flat high-order elements for Helmholtz and Laplace equations. In chapter 10, a differential geometry-based quadrature algorithm is developed which allows accurate evaluation of singular and nearly singular layer potential integrals arising in the boundary element method using smooth manifold boundary elements with constant densities for Helmholtz and Laplace equations. In chapter 11, an algorithm for efficient exact evaluation of integrals of regular solid harmonics over high-order boundary elements with simplex geometries is developed. In chapter 12, I discuss future research directions and conclude the dissertation.

NUMERICAL ACOUSTICS FOR PHYSICAL
AND SIMULATED ENVIRONMENTS

by

Shoken Kaneko

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:

Professor Ramani Duraiswami, Chair/Advisor
Professor Howard Elman
Professor Dinesh Manocha
Professor Maria Cameron, Dean's Representative
Doctor Hannes Gamper, Special Member

© Copyright by
Shoken Kaneko
2023

Preface

Computer modeling and numerical analysis of physical phenomena have been driving forces in many fields in science and technology. In manufacturing, computer simulation has become a common tool to predict the physical characteristics of products before building actual physical prototypes, and has served to shorten the development time, to reduce the cost of development, and to improve the quality of products. Immersive multimedia, such as 3D gaming and virtual reality, are other fields where physical simulation is an essential component; physical modeling of the real world is used for the creation of stimuli that fools human's perception in order to deliver an immersive experience of a virtual or augmented world. Also, machine learning, another rapidly growing branch of technology, consumes data to learn nontrivial mappings from inputs to responses. Machine learning models for real world applications require training data from the real physical world. Computer simulation of the physical world is therefore increasingly important as it is one of the most efficient methods to generate unlimited amounts of data that approximates data captured from the physical world.

The primary focus of the present dissertation is the exploration of numerical methods for modeling, simulating, capturing, processing, analyzing and synthesizing acoustical phenomena in the physical world as well as its applications to the virtual world.

In chapter 1, I introduce some fundamentals and basic concepts of numerical acoustics which are going to be used throughout the dissertation. I will also review some of the

existing practical problems in acoustics.

In chapter 2, I introduce multiple scattering ambisonics, a spatial sound field capturing technique for extended reproduction-time sweet-spots in XR audio, based on semi-analytical cancellation of multiple scattering arising in a system of interacting microphone arrays implemented on spherical scattering bodies.

In chapter 3, I introduce spheroidal ambisonics, a spatial sound field capturing technique for the creation of non-spherical reproduction-time sweet-spots in XR audio, based on the use of microphone arrays implemented on spheroid-shaped scattering bodies and spheroidal wave functions as the expansion basis.

In chapter 4, a fast and effective algorithm for the synthesis of acoustic impulse responses of large-scale forests is proposed. This algorithm models trees as single scattering cylinders and hence achieves linear complexity with respect to the number of trees in the system.

In chapter 5, simulation-based evaluation of sound source localization of birds using distributed microphone arrays in large-scale forests is carried out. This study is based on the forest acoustics simulation method developed in chapter 4 and aims for audio-based monitoring systems for wildlife conservation.

In chapter 6, a general-purpose individual-agnostic full-sphere binaural localizer model is introduced. This binaural localizer is designed for sound source localization from binaural audio sequences with unknown generation process, i.e. unknown recording / rendering-time head-related impulse response (HRIR), in noisy and reverberant conditions.

In chapter 7, a regularization strategy is developed for spherical harmonics-domain active noise control in reverberant rooms.

In chapter 8, a personal active sound field control framework with applications to speech privacy, local active noise control, and crosstalk cancellation for XR audio presentation is introduced.

In chapter 9, a recursive algorithm is developed which allows efficient analytical evaluation of singular and nearly singular layer potential integrals arising in the boundary element method using flat high-order elements for Helmholtz and Laplace equations.

In chapter 10, a differential geometry-based quadrature algorithm is developed which allows accurate evaluation of singular and nearly singular layer potential integrals arising in the boundary element method using smooth manifold boundary elements with constant densities for Helmholtz and Laplace equations.

In chapter 11, an algorithm for efficient exact evaluation of integrals of regular solid harmonics over high-order boundary elements with simplex geometries is developed.

In chapter 12, I discuss future research topics and conclude the dissertation.

The research presented in chapter 4, chapter 5, and chapter 6 was conducted while the author was a Research Intern at Microsoft Research.

Dedication to Mayu, Shima, Yulia, Kenshin, and Eugene.

Acknowledgments

I owe my gratitude to all the people who have made this dissertation possible.

First and foremost, I would like to deeply thank my advisor, Professor Ramani Duraiswami for his kindness and generosity for allowing me to work on the most exciting and challenging problems, for sharing his invaluable knowledge, expertise and insights, and for all sorts of support throughout the years. It has been not only a great opportunity for my growth, but also a great pleasure to work with such an excellent scientist and innovator.

I also express my deepest gratitude to my co-advisor, Dr. Nail Gumerov, who, tragically passed away in the midst of my graduate studies. Dr. Gumerov's works have been inspiring me already before I came to UMD - when I was a research engineer at Yamaha Corporation. Dr. Gumerov is partly responsible for me quitting the established company position and knocking on the doors of UMD as a student. It was an incredible privilege to be able to work with and learn from such an extraordinary scientist, whose works have not only strongly influenced my research, but will also be embraced by humanity forever.

I would like to thank Professor Nirupam Roy and Irtaza Shahid for all the support they provided in our collaboration.

I would like to thank my external collaborators and mentors in the industry; Dr.

Hannes Gamper and Dr. Ivan Tashev at Microsoft Research, Dr. Owen Brimijoin and Dr. Antje Ihlefeld at Meta Reality Labs, Dr. Abhaya Parthy and Dr. Deep Sen at Apple for having me as an intern on truly exciting research projects.

Thanks are due to Professor Howard Elman, Professor Dinesh Manocha, Professor Maria Cameron, and Dr. Hannes Gamper for agreeing to serve on my dissertation committee and for sparing their invaluable time reviewing the manuscript.

I would like to thank Professor Howard Elman, Professor Abhinav Bhatele, Professor David Mount, Professor Uzi Vishkin, Professor Eitan Tadmor, Professor Tobias Von Petersdorff, Professor Ricardo Nochetto, and Professor Dan Cristofaro-Gardiner, whose incredible lectures on computer science, applied and pure mathematics helped me deepening and expanding my scientific knowledge and skills.

I would like to thank my peer lab-mates and students colleagues at UMD, Dr. Sam Potter, Bowen Zhi, Monte Hoover, Siminfar Samakoush Galougah, James Robertson, John Baur, Meenakshi Krishnan, Armin Gerami, and Shuhong Chen for enriching my graduate life.

I would like to express my gratitude to assistant director Tom Hurst, who has been supporting me throughout the years with his extraordinary administrative work capabilities. His existence has been a relief so many times.

I would like to deeply acknowledge financial support from Japan Student Services Organization (JASSO), Watanabe Foundation, Roberta M. Ma Fund, Ann G. Wylie Fellowship, and all TA/RA sponsors throughout the years. I would also like to thank Microsoft Research, Meta Reality Labs, and Apple again for their paid internship opportunities which were essential for funding my studies. My graduate studies while having a family

was only possible with the generous financial support from these organizations, labs, and companies.

I owe my deepest thanks to my wife Mayu Kaneko who has always stood by me and supported me. Without her tremendous dedication, obtaining a Ph.D. degree while raising our four children in a foreign country during a pandemic was simply not possible. I would like to thank our four children - Shima, Yulia, Kenshin, and Eugene - for being part of my journey. The joy and distractions provided by them have been essential for pursuing my goals without burning out.

I thank my parents who are responsible for shipping me and for raising me to be who I am today.

Table of Contents

Preface	ii
Foreword	v
Dedication	vi
Acknowledgements	vii
Table of Contents	x
List of Tables	xv
List of Figures	xvi
List of Abbreviations	xxii
Chapter 1: Introduction	1
1.1 Introduction to acoustics	1
1.1.1 From continuum mechanics to the acoustic wave equation	1
1.1.2 The Helmholtz equation	4
1.2 Practical problems in acoustics	5
1.2.1 Capturing sound	5
1.2.2 Sound synthesis of sources and environments	5
1.2.3 Sound localization and binaural audition	6
1.2.4 Controlling sound	7
1.2.5 Numerical modeling and simulation of sound	7
1.3 Organization of the dissertation	11
1.3.1 Capture	11
1.3.2 Synthesis of an acoustic environment	11
1.3.3 Localization	12
1.3.4 Control	12
1.3.5 Boundary element analysis	12
1.4 Special functions used in this dissertation	13
1.4.1 Legendre polynomials	13
1.4.2 Associated Legendre polynomials	13
1.4.3 Spherical harmonics	14

Chapter 2:	Multiple scattering ambisonics: three-dimensional sound field estimation using interacting spheres	15
2.1	Introduction	15
2.2	Conventional ambisonics encoding using a single RSMA	18
2.3	Proposed method	21
2.3.1	The forward problem: multiple scattering due to an arbitrary incident field	21
2.3.2	The inverse problem: MS-HOA encoding	23
2.4	Numerical experiments	24
2.5	Related work and discussion	25
Chapter 3:	Spheroidal ambisonics: spatial audio in spheroidal bases	29
3.1	Introduction	29
3.2	Background: spherical ambisonics	31
3.2.1	Encoding in spherical ambisonics	32
3.3	Formulation of spheroidal ambisonics	35
3.3.1	Spheroidal coordinates	35
3.3.2	Scattering of an arbitrary incident wave by a sound-hard prolate spheroid	36
3.3.3	Spheroidal ambisonics encoding	38
3.4	Transcoding from spheroidal to spherical ambisonics	39
3.4.1	The transcoding formula	39
3.4.2	Mixed-order transcoding	40
3.5	Experimental evaluation	41
3.6	Conclusion	45
Chapter 4:	Fast forest reverberation synthesis using single scattering cylinders	46
4.1	Introduction	46
4.2	Proposed Approach	49
4.2.1	Algorithm overview	49
4.2.2	Implementation and optimization details	52
4.3	Experimental Evaluation	53
4.3.1	Echo density of synthetic and real forest IRs	53
4.3.2	Energy decay curves of synthetic and real forest IRs	55
4.3.3	Reverberation time and clarity	58
4.4	Summary and discussion	58
Chapter 5:	Localizing birds in forests with distributed microphone arrays	62
5.1	Introduction	62
5.2	DMA recording simulation	64
5.2.1	Bird sounds and ambient noise	65
5.2.2	Acoustic simulation	65
5.2.3	Multi-channel sound scene generation with measurement error	66
5.3	Sound source localization (SSL) and evaluation metrics	66
5.4	Problem formulation and the experimental setup	68

5.4.1	Forest impulse responses	69
5.4.2	Sound scene synthesis	69
5.4.3	Localization parameters	70
5.5	Results	70
5.5.1	The effect of each explanatory variable	70
5.5.2	The effect of microphone spacing, noise type, and modifications to improve localization accuracy	72
5.5.3	The effect of the SNR	74
5.6	Conclusion and future work	74
Chapter 6: A general-purpose individual-agnostic full-sphere binaural localizer		77
6.1	Introduction	78
6.2	Proposed method	80
6.3	Experimental evaluation	81
6.3.1	Model training and validation	83
6.3.2	Single static source localization	85
6.3.3	Multiple or moving source localization	88
6.3.4	Response to panned stereo audio	89
6.4	Conclusion and discussion	90
Chapter 7: Regularized spherical harmonics-domain spatial active noise cancellation in a reverberant room		92
7.1	Introduction	93
7.2	Problem formulation	94
7.3	Spherical harmonics-domain ANC method	97
7.3.1	Scenario <i>known</i> : known monopole noise source in a known room	98
7.3.2	Scenario <i>unknown</i> : arbitrary noise source in a known room	98
7.3.3	Scenario <i>blind</i> : arbitrary noise source and estimated room transfer functions	101
7.3.4	The global cost function for hyperparameter optimization	101
7.4	Numerical Experiments	102
7.4.1	Room condition	102
7.4.2	Room transfer function simulation	103
7.4.3	Cancellation loudspeaker array and estimation microphone array configurations	103
7.4.4	Optimization procedure	104
7.5	Results	104
7.6	Conclusion	109
Chapter 8: Personal active soundfield control (PASCAL) framework		110
8.1	Introduction	110
8.2	Problem Formulation	112
8.3	Tasks and objectives	113
8.3.1	Objective functions and optimal driving signals	114
8.3.2	Cost function for regularization parameter optimization	116

8.4	Numerical Experiments	117
8.4.1	Loudspeaker configurations and FMMBEM simulation	117
8.4.2	Radiation cancellation simulation	119
8.4.3	Personal ANC simulation	120
8.5	Summary and related work	123
Chapter 9: Recursive analytical quadrature for the close evaluation of Laplace and Helmholtz layer potentials over flat boundary elements in \mathbb{R}^3		126
9.1	Introduction	127
9.2	Boundary element method and layer potentials	130
9.3	Problem statement	131
9.4	Method: Recursive Integrals for Polynomial Elements (RIPE)	134
9.4.1	Dimension reduction via divergence theorem	134
9.4.2	Laplace kernel with polynomials of arbitrary order	136
9.4.3	Helmholtz kernel with polynomials of arbitrary order	145
9.4.4	Complexity	154
9.5	Numerical evaluation	155
9.5.1	Accuracy	155
9.5.2	Stability	158
9.5.3	Computation time	160
9.6	Conclusion	160
Chapter 10: Layer potential quadrature on manifold boundary elements with constant densities for Laplace and Helmholtz kernels in \mathbb{R}^3		163
10.1	Introduction	164
10.2	Boundary element method and layer potentials	166
10.3	Differential geometry preliminaries	168
10.3.1	Curvature of regular surfaces	168
10.3.2	Stokes' theorem on smooth manifolds	169
10.4	Problem statement	169
10.5	Proposed method	170
10.6	Numerical evaluation	179
10.6.1	Element-level tests	179
10.6.2	Integrated BEM test: thin spherical cavity problem	183
10.7	Conclusion	186
Chapter 11: Efficient exact quadrature of regular solid harmonics times polynomials over simplices in \mathbb{R}^3		188
11.1	Introduction	189
11.2	Potential integrals in BEM and vortex methods	190
11.2.1	Multipole expansion of potentials	192
11.3	Problem statement	194
11.4	Q2XP: Q2X for Polynomial elements	195
11.4.1	Recursions	196
11.4.2	Starting values for the recursions	200

11.4.3 Algorithm	200
11.5 Numerical experiment	202
11.6 Conclusion	203
Chapter 12: Conclusion and outlook	205
Appendix A: Derivation of the spheroidal ambisonics transcoding formula	208
Appendix B: Elementary integral computations arising in the RIPE method	214
B.1 Computation of integrals i_m	214
B.2 Computation of integrals k_m	215
Appendix C: Proof of relation (10.15) used in the manifold element integrals	217
Bibliography	220

List of Tables

5.1	List of variables and their values used in the experiments.	71
6.1	Model hyperparameters	82
6.2	BRIR simulation parameters and their ranges from which the parameters are uniformly sampled.	85
6.3	Mean angular error (degrees) for various models.	86
7.1	Microphone / loudspeaker array configuration used in the numerical experiments.	104
8.1	Loudspeaker channels used in each configuration.	118
9.1	Complexity for evaluating layer potentials for p_s of $O(1)$ and general order, respectively. Note that the complexities for Gauss-Legendre are for one layer potential evaluation, while for RIPE they are the complexities to computes all terms with the monomial factor $x^b y^c$ for $0 \leq b + c \leq$ p_s in a single execution of the recursive algorithm. These terms can be reused when evaluating multiple layer potentials with different linear combinations of these monomial terms.	155
9.2	Single layer potentials for the Laplace kernel evaluated over the element with second-order shape functions of a conforming Lagrange triangle element. Numbers for RIPE (P_{RIPE}), Guiggiani's method (P_{Gui}), and their relative differences are shown.	157
9.3	Hypersingular potentials for the Laplace kernel evaluated over the second- order element.	158
9.4	Single layer potentials for the Helmholtz kernel ($kD = 1$) over the second- order element.	158
9.5	Hypersingular potentials for the Helmholtz kernel ($kD = 1$) over the second-order element.	158
10.1	Summary of related quadrature methods for layer potentials or its multiple expansions based on analytical or dimensionality-reduction based evaluation. Methods for high order elements can be applied to constant elements, and methods for curved elements can be applied to flat elements. A summary of methods based on other approaches e.g. singularity subtraction/cancellation can be found in [1].	165
10.2	Coefficients introduced in (10.12) for Laplace and Helmholtz, single and double layer potentials.	174

List of Figures

2.1	Left: examples of reconstruction sweet-spots using a 252-channel RSMA for estimating incident fields with various frequencies. Details of the RSMA are described in Section 8.4. The truncation degree of HOA which provides the largest sweet-spot is chosen for each frequency independently. Right: illustration of a sound field recording setup using MS-HOA (top) and its reproduction setup over headphones allowing translation of the listener (bottom).	17
2.2	The scheme of the forward problem, i.e. computation of the total sound field at the microphone positions given the incident field, and the inverse problem, i.e. estimation of the incident field given the sound pressure at the microphone positions.	24
2.3	Results for the linear 6-sphere RSMA grid for an incident field of 4 kHz. Top row from left to right: real part of the sound pressure for the ground truth incident field, incident field estimated by HOA , Single , and MS-HOA , respectively. Bottom row from left to right: SDR map of the estimated fields with respect to the ground truth field using HOA , Single , and MS-HOA , respectively. The blue circles represent the positions and sizes of the RSMAs. The truncation number is $N_c = 14$ in HOA , $N_c^{(\text{in})} = 55$ and $N_c^{(\text{fwd})} = 20$ in Single and MS-HOA .	26
2.4	Results for the regular Cartesian 9-sphere RSMA grid. The estimated incident field (top row) and SDR map (bottom row) for Single (left column) and MS-HOA (right column), respectively. The blue circles represent the positions and sizes of the RSMAs. The same incident field as Fig. 2.3 is used. The truncation numbers are $N_c^{(\text{in})} = 45$ and $N_c^{(\text{fwd})} = 16$.	27
3.1	The overview of the proposed spheroidal ambisonics. The microphone capsule positions used in the numerical experiments presented in this paper are shown as the red dots in the images.	31
3.2	Left: prolate spheroidal coordinates ξ and η for $a = 1$. The lines of equal values of ξ and η are shown on the xz -plane, i.e. for $\varphi = \{0, \pi\}$. Right: truncation error $ e^{i\mathbf{k}\cdot\mathbf{r}} - p_{\text{in}}^{\text{pw}(N)} $ of the series expansion of a plane wave (3.16) truncated at $n = N$ as a function of the spherical coordinate $(r, \theta, \varphi = 0)$ with $a = 1$, $\mathbf{k} = (2\pi f/v, 0, 0)$, $f = 541.8$ Hz, and $v = 340$ m/s.	36

3.3	Results for an incident plane wave travelling along the long axis of the spheroidal array, which is set parallel to the x -axis. The first row from left to right: the real part of the sound pressure of the ground truth incident field, the field reconstructed from spherical ambisonics coefficients with $N' = 12$ (HOA), the field reconstructed from the prolate spheroidal ambisonics coefficients with $N = 12$ transcoded to spherical ambisonics coefficients with $N' = 12$ (ps-HOA), and the field reconstructed from the prolate spheroidal ambisonics coefficients with $N = 12$ transcoded to spherical ambisonics coefficients with $N' = 16$ using mixed-order transcoding (ps-HOA-mo). The second row presents the SDR of the reconstructed fields for HOA (left), ps-HOA (center), and ps-HOA-mo (right). The region with SDR higher than 30 dB was considered as the sweet-spot and is colored in red.	43
3.4	Results for an incident plane wave travelling along the short axis of the spheroidal array which is set parallel to the y -axis (upper two rows), and for an incident plane wave with a normalized wave vector $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0)$ (lower two rows). The definition of each of the subplot is identical to Fig. 3.3.	44
4.1	The amplitude spectrum of the angle-dependent tree scattering filter H_{angle} [2].	52
4.2	The waveforms (left) and frequency responses (right) of various forest IRs. From top to bottom: Koli national park in winter and summer [3–5], treeverb algorithm [6, 7], waveguide web algorithm [7, 8], proposed method with $N_c = 100\text{k}, 200\text{k}, 500\text{k}$	56
4.3	The echo density profiles of forest IRs. Comparison with reference forest IRs (top), and comparison among forest IRs generated by the proposed method for a square-shaped forest of $1 \text{ km} \times 1 \text{ km}$, with various D and N_c (bottom). Each subplot in the bottom shows results for $N_c = 100\text{k}$ (bottom left), 200k (bottom center), and 500k (bottom right), respectively.	57
4.4	Top: EDCs for various short distance forest IRs. The set of forest IRs is the same as in Section 4.3.1. Bottom: EDCs for synthetic and real forest IRs with varying D . N_c was set to 500k for the proposed method. All EDCs are computed from the IRs after applying a band-pass filter with a passband of $500\text{Hz}-8\text{kHz}$	59
4.5	T60 as a function of octave-band frequency (left) and C50 as a function of D (right). The set of shown forest IRs is the same as in Section 4.3.1 and Section 4.3.2 for the left and right sub-plot, respectively. IRs from the Koli forest and previous algorithms are not included in the C50 plot since the available dataset only contained IRs for a fixed distance.	60
5.1	Diagram of a possible bird monitoring system. D denotes the microphone spacing, assuming a regular Cartesian grid.	64
5.2	Left: histograms of localization errors. Right: the standard deviations with respect to the LSR along each dimension of the variable space. Note that the noise type dimension is split into noise category and noise level.	72

5.3	The LSRs for the cases with or without BL and NS. The results are for the realistic scenario (directional sources, with synchronization error, perturbed microphone grid with positional error, and GCC-PHAT algorithm), averaged over source positions and tree density, using the Xeno-canto dataset. The sub-plots represent the baseline (left-top), only BL (right-top), only NS (left-bottom), and both BL and NS (right-bottom).	73
5.4	The LSR as function of average SNR using various NS algorithms. This results are for the realistic setup using directional sources, <i>pert-pert</i> grid with synchronization error, and the GCC-PHAT algorithm. Band-limiting is applied except for the baseline case. Localization results from all noise types except for the clean case, all source positions, forests with non-zero trees, and the Xeno-canto dataset is used here.	75
6.1	The network architecture of the proposed model.	82
6.2	Typical output activation maps taken from LOCATA-task1 test clips. The right hemisphere of the map is shown with the L^∞ norm of the post-sigmoid output vector normalized after taking the maximum over time frames.	86
6.3	Output maps shown in the lateral-polar grid, for LOCATA-task2 test clips with two to four sources. The ovals denote the true source directions.	88
6.4	The estimated lateral (left column) and polar (right column) angle as function of ITD (top row) or ILD (bottom row) of the panned stereo signal.	90
7.1	An illustration of the SHANC system. The microphones and loudspeakers in the present work are placed on spheres concentric with the target sphere. The target sphere, the microphone array, and the loudspeaker array have radii r_{rec} , r_{mic} , and r_{sp} , respectively, with $r_{\text{rec}} < r_{\text{mic}} < r_{\text{sp}}$	94
7.2	The resulting fields on the listening plane for the case of a noise source of 500 Hz and a SHANC system consisting of 12 estimation microphones and 12 cancellation loudspeakers, in the <i>blind</i> scenario. From left to right: sound pressure level of the noise field, anti-noise field, total residual field, the RoS/RoS map, and the contour of the ZoS, respectively. The upper and lower row shows the result for the proposed regularized SHANC method and the reference case without regularization, respectively.	105
7.3	The resulting fields on the listening plane for the case of a noise source of 2 kHz and a SHANC system consisting of 128 estimation microphones and 128 cancellation loudspeakers, in the <i>blind</i> scenario. The meaning of each of the subplot is identical to Fig. 7.2.	105
7.4	Summary of results for the <i>known</i> scenario. Noise reduction level on the target sphere (left) and the radius of the SoS (right), for various device configurations and frequencies.	106
7.5	Summary of results for the <i>unknown</i> scenario. Noise reduction level on the target sphere (left) and the radius of the SoS (right), for various device configurations and frequencies.	107

7.6	Summary of results for the <i>blind</i> scenario. Noise reduction level on the target sphere (left) and the radius of the SoS (right), for various device configurations and frequencies.	108
8.1	An example system implementing PASCAL. A set of loudspeakers and microphones is used to present the transaural signal and to suppress the external field by emitting a cancellation field.	111
8.2	Left: the shape model used for the BEM simulation. The red, purple, blue, and cyan dots represent the set of loudspeakers, probe points, test points, and optimization points respectively. The probe points, test points, and optimization points are defined on both sides of the head. Right: loudspeaker channel assignment for the neckband array (top) and the laptop array (bottom). In the neckband array, channel 4 and 10 correspond to front and back, respectively.	118
8.3	The radiation field (top left) and residual field (top right) at 500 Hz using the 7ch system, reduction level (middle) of the fully-simulated ARC experiment, and the test point amplitude spectrum of the ARC simulation using real IRs (bottom).	120
8.4	The reduction levels for the neckband system (top) and laptop system (bottom) using two or seven loudspeakers.	122
8.5	Top: the cancellation signal power spectra. Bottom: amplitudes of the cancellation filters $H^{(\text{can})}$ for the four paths connecting the 2ch probe microphones and the 2ch loudspeakers. The solid lines and dotted lines represent filter spectra for the neckband and laptop setup, respectively.	123
8.6	The external field (top row), cancellation field (middle row), and residual field (bottom row) at 500 Hz in the x - y plane on the height of the listener's head for a monopole noise source located at (2m, 1m, 0) and the head located at the origin, for the stereo neckband (left column) and stereo laptop (right column), respectively.	124
9.1	Types of layer potential integrals categorized by element-evaluation point distance. (1) singular or hypersingular case where the evaluation point is on the element surface, (2) nearly singular case where the evaluation point is off the element surface but too close for simple Gauss-Legendre quadrature being accurate, (3) intermediate distance where direct evaluation is still applied and Gauss-Legendre quadrature is effective, and (4) far-field, where far interactions are clustered using the fast multipole method. This paper provides efficient and accurate computation of integrals of types (1) and (2) for all four layer potential kernels (Eqs. (9.3) and (9.5)), for both the Laplace and Helmholtz kernels.	127
9.2	A boundary element and its associated coordinate frames. We have the <i>element coordinate frame</i> $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ and <i>edge coordinate frame</i> $(\hat{\mathbf{i}}^j, \hat{\mathbf{j}}^j, \hat{\mathbf{k}}^j)$ for each edge with index j . The origin of the element frame is set to \mathbf{v}_1	132

9.3	Left: the surface integral over element S of a function which can be written as the surface divergence of a vector field \mathbf{m} . Right: after the application of the divergence theorem, the surface integral over S is converted into a contour integral over the contour $C = \partial S$ and further into line integrals due to the straightness of the edges.	134
9.4	Conventional Gauss-Legendre quadrature requires evaluation of the integrand at multiple nodes on the element S , which may be near or coincide with the evaluation point, or collocation node, \mathbf{r}_p . After application of the divergence theorem, the surface integral over S is converted into an integral over the contour $C = \partial S$ and further into line integrals, which in the end results in the evaluation of analytical expressions at the endpoints of the edges, i.e. at the vertices.	136
9.5	Computing $\xi_{b,c,a}$ using algorithm 2. The blue arrows indicate terms in the recurrence relations used in each step, here for the Laplace single layer potential with $p_s = 5$. Step 1 (top left): compute seed term $\xi_{0,0,0}$, step 2 (top center): compute $\xi_{0,0,a}$ using recurrences along the a -axis, step 3 (top right): compute $\xi_{0,c,a}$ on the $b = 0$ plane using recurrences in the c -directions, step 4-6 (bottom left to right): compute $\xi_{b,c,a}$ for $b > 0$ using recurrences in the b -direction.	142
9.6	Relative difference of the layer potentials between RIPE or 12-th order Gauss-Legendre quadrature and the reference adaptive Gauss-Kronrod quadrature, for the shape function $N(x, y) = x^3$. Results for single layer potential (top left), double layer potential (top right), adjoint double layer potential (bottom left), and hypersingular potential (bottom right).	157
9.7	Relative difference of the layer potentials between RIPE and the adaptive Gauss-Kronrod quadrature, for various shape function orders. Laplace (left column) and Helmholtz (right column), for single layer, double layer, adjoint double layer, and hypersingular potentials (top to bottom).	159
9.8	The total wall-clock time for computing all four layer potentials for a 5th order element, for RIPE and for adaptive Gauss-Kronrod quadrature. Note that the times for RIPE are the times spent to compute integrals weighted by $x^b y^c$ for all $((p_s + 1)(p_s + 2))/2 = 21$ terms in $0 \leq b + c \leq 5$ in one run of the recursion, while the times for Gauss-Kronrod are for computing only the integral with shape function x^5	161
10.1	Left: The normal vectors of a regular surface S and a regular curve c at point p . Right: the surface S and its contour ∂S	169
10.2	The manifold boundary element S and the tangent plane $T_q(S)$ at \mathbf{r}_q . Red, blue, and green lines indicate edges of the element, vectors parallel to the tangent plane $T_q(S)$, and vectors orthogonal to the tangent plane, respectively. \mathbf{n}_q , $\hat{\boldsymbol{\rho}} \equiv \boldsymbol{\rho}/ \boldsymbol{\rho} $ and $\tilde{\boldsymbol{\rho}} \equiv \mathbf{n}_q \times \hat{\boldsymbol{\rho}}$ can be used to construct a local orthogonal coordinate frame centered at \mathbf{r}_q	170
10.3	The coordinate transform used for the quadrature of the curvature term.	176

10.4	A 2D illustration of curved elements and their normal bundles. The element, their outward and inward normal bundles are drawn by black, blue, and red lines, respectively. Left: outward and inward normal bundles do not intersect. In this case, evaluation points in the inward normal bundle can avoid the singularity by flipping the parametrization. Right: outward and inward normal bundles intersect and if \mathbf{r}_p is in the intersection, the element has to be subdivided so that \mathbf{r}_p does not belong to an intersection of normal bundles.	178
10.5	The manifold boundary elements with positive and negative curvature used in the numerical tests. <i>Element 1</i> (left, $\sigma = -0.6$) and <i>element 2</i> (right, $\sigma = 0.6$). The colors are for visual aid.	180
10.6	Relative difference of the layer potentials on <i>element 1</i> for various methods and the reference adaptive Gauss-Kronrod quadrature. Results for single layer potential with Laplace kernel (top left), double layer potential with Laplace kernel (top right), single layer potential with Helmholtz kernel (bottom left), and double layer potential with Helmholtz kernel (bottom right). All results are for the case with quadrature order set to 20.	181
10.7	Relative difference of the layer potentials on <i>element 2</i> for various methods and the reference adaptive Gauss-Kronrod quadrature. Results for single layer potential with Laplace kernel (top left), double layer potential with Laplace kernel (top right), single layer potential with Helmholtz kernel (bottom left), and double layer potential with Helmholtz kernel (bottom right). All results are for the case with quadrature order set to 20.	182
10.8	Relative difference $ P_{\text{prop}} - P_{\text{Gui}} / P_{\text{Gui}} $ of the singular case layer potentials on <i>element 1</i> for the proposed method with various quadrature orders and the Guiggiani's reference method.	183
10.9	The cross section of the spherical cavity setup used as a benchmark problem. The white region in between the two spheres is the domain $\Omega = \{r a \leq r \leq b\}$ subject to analysis. The vibrating surface is the upper hemisphere of the interior sphere and is indicated as the red arc.	184
10.10	Results of the spherical cavity BEM test problem. Relative L_2 -norm of the error at the collocation points for two geometry conditions (flat or non-flat manifold elements) and two algorithms used for nearly singular integrals (Gauss-Legendre quadrature or the proposed method).	186
11.1	2- and 1-simplex in a cell of an octree data structure commonly used in the FMM-BEM. The evaluation point \mathbf{r}_p is located in a well-separated position outside the cell the elements reside.	190
11.2	Wall clock times for running Q2XP up to $p_s = p_d = 20$	202
11.3	Maximum relative differences between the integrals computed by Q2XP and exact Gauss-Legendre quadrature for various p_d and p_s values. Left: L integrals and right: M integrals. Good accuracy is observed in all conditions.	203

List of Abbreviations

ANC	Active Noise Control
AoN	Area of Noise
AR	Augmented Reality
ARC	Active Radiation Cancellation
BAD	Bird Activity Detector
BCE	Binary Cross Entropy
BEM	Boundary Element Method
BL	Binaural Localizer
BRIR	Binaural Room Impulse Response
CAD	Computer-Aided Design
CAE	Computer-Aided Engineering
CCE	Categorical Cross Entropy
CNN	Convolutional Neural Network
DMA	Distributed Microphone Array
DNN	Deep Neural Network
DoF	Degrees-of-Freedom
EDC	Energy Decay Curve
FC	Fully-Connected
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FMM	Fast Multipole Method
FMM-BEM	Fast Multipole-accelerated Boundary Element Method
GCC	Generalized Cross Correlation
GCC-PHAT	Generalized Cross Correlation - Phase Transform
HOA	Higher Order Ambisonics
HRIR	Head-Related Impulse Response
HRTF	Hear-Related Transfer Function
ILD	Interaural Level Difference
IPD	Interaural Phase Difference
ITD	Interaural Time Difference
IR	Impulse Response
LSR	Localization Success Rate
MAE	Mean Angular Error
ML	Maximum Likelihood

MMSE-STSA	Minimum Mean-Square Estimator Short-Term
Q2X	Quadrature to Expansion
Q2XP	Quadrature to Expansion for Polynomial elements
MAE	Mean Angular Error
MAP	Maximum A-Posteriori
MOI	Method of Images
MR	Mixed Reality
MS-HOA	Multiple Scattering Higher Order Ambisonics
NN	Neural Network
NS	Noise Suppression
PASCAL	Personal Active Soundfield Control Algorithm
PDE	Partial Differential Equation
RIPE	Recursive Integrals for Polynomial Elements
RIR	Room Impulse Response
RoN	Region of Noise
RoS	Region of Silence
RSMA	Rigid Spherical Microphone Array
RTF	Room Transfer Function
SDR	Signal-to-Distortion Ratio
SH	Spherical Harmonics
SHANC	Spherical Harmonics-domain Active Noise Control
SMA	Spherical Microphone Array
SNAD	Silent-to-Noisy Area Difference
SNR	Signal-to-Noise Ratio
SoS	Sphere of Silence
SPL	Sound Pressure Level
SS	Spectral Subtraction
SSC	Single Scattering Cylinder
SSL	Sound Source Localization
STFT	Short-Time Fourier Transform
TDOA	Time-Difference of Arrival
TF	Transfer Function
VAR	Voice Activity Detector
VR	Virtual Reality
XTC	Crosstalk Cancellation
ZoS	Zone of Silence

Chapter 1: Introduction

In this chapter some fundamentals of theoretical and numerical acoustics are reviewed in order to introduce some of the basic concepts which are repeatedly used throughout the dissertation. I then discuss types of practical problems arising in acoustics and audio engineering.

1.1 Introduction to acoustics

1.1.1 From continuum mechanics to the acoustic wave equation

Sound is the excitation of the sound field, where the excitation propagates in form of a longitudinal wave in a medium. The amplitude of excitation, which is called the *sound pressure*, is a function of position \mathbf{r} and time t , hence it can be expressed as $p(\mathbf{r}, t)$. A monochromatic plane wave is a form of wave which is characterized by the direction of propagation \mathbf{k} , which is called the *wave vector*, an angular frequency ω , and an amplitude A , and is mathematically expressed as:

$$p(\mathbf{r}, t, \mathbf{k}, \omega, A) = Ae^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}. \quad (1.1)$$

As any sound field can be mathematically expressed as an integral of plane waves over all possible directions of propagation and angular frequencies, it is beneficial to understand the properties of a plane wave. The dynamics of a plane wave propagating towards the positive direction of the x -axis can be expressed by the displacement $d(x, t)$ which is a function of position x and time t . The ratio of volume increase of a small volume, which is sufficiently larger than the scale of atoms and sufficiently smaller than the wavelength, is given by:

$$\Delta = \frac{\partial d}{\partial x}. \quad (1.2)$$

In an adiabatic process with linearity approximation, the sound pressure p , which is the difference of the pressure and the equilibrium pressure of the medium, can be written as:

$$p = -K\Delta, \quad (1.3)$$

where the proportionality coefficient K is called the *bulk modulus*. Using the particle velocity $v = \partial d / \partial t$, this equation can be rewritten as:

$$\frac{\partial v}{\partial x} = -\frac{1}{K} \frac{\partial p}{\partial t}. \quad (1.4)$$

This equation is called the continuity equation. On the other hand, the equation of motion for a unit volume is given as:

$$\rho \frac{\partial v}{\partial t} = -\frac{\partial p}{\partial x}, \quad (1.5)$$

with ρ the density of the medium. By differentiating the continuity equation by time t and the equation of motion by position x , we obtain the acoustic wave equation for the sound pressure p :

$$\frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} - \frac{\partial^2 p}{\partial x^2} = 0, \quad (1.6)$$

with c the sound velocity which is defined as:

$$c \equiv \sqrt{\frac{K}{\rho}}. \quad (1.7)$$

For three-dimensional sound fields, this one-dimensional wave equation can be written as follows:

$$\frac{1}{c^2} \frac{\partial^2 p(\mathbf{r}, t)}{\partial t^2} - \nabla^2 p(\mathbf{r}, t) = 0, \quad (1.8)$$

where $\mathbf{r} = (x, y, z)$ is the position in three-dimensional space. In the presence of sound sources, this homogeneous wave equation is generalized as:

$$\frac{1}{c^2} \frac{\partial^2 p(\mathbf{r}, t)}{\partial t^2} - \nabla^2 p(\mathbf{r}, t) = f(\mathbf{r}, t), \quad (1.9)$$

where $f(\mathbf{r}, t)$ is the forcing term (source term). This equation is called the inhomogeneous wave equation.

For the sake of completeness, I present the definitions of *hyperbolic*, *parabolic*, and *elliptic* partial differential equations (PDEs). Any second-order linear PDE with two

variables can be written in the following form:

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu + G = 0, \quad (1.10)$$

where A, B, C, D, E, F, G are functions of x and y . If the conditions $B^2 - AC > 0$, $B^2 - AC = 0$, or $B^2 - AC < 0$ are satisfied, the PDE is called *hyperbolic*, *parabolic*, or *elliptic*, respectively [9]. The wave equation is categorized as a hyperbolic PDE.

1.1.2 The Helmholtz equation

Given the inhomogeneous wave equation for a field $U(\mathbf{r}, t)$ with a source term $F(\mathbf{r}, t)$:

$$\frac{1}{c^2} \frac{\partial^2 U(\mathbf{r}, t)}{\partial t^2} - \nabla^2 U(\mathbf{r}, t) = F(\mathbf{r}, t), \quad (1.11)$$

and assuming a time-harmonic field $U(\mathbf{r}, t) = u(\mathbf{r})e^{-i\omega t}$ and a source $F(\mathbf{r}, t) = f(\mathbf{r})e^{-i\omega t}$, we obtain the Helmholtz equation:

$$-k^2 u(\mathbf{r}) - \nabla^2 u(\mathbf{r}) = f(\mathbf{r}), \quad (1.12)$$

where $k = \omega/c$ is the wavenumber. The Helmholtz equation, as well as its static case, the Laplace equation, are categorized as elliptic PDEs.

1.2 Practical problems in acoustics

1.2.1 Capturing sound

Capturing sound is one of the central tasks involved in various applications in audio and acoustics. It is particularly important in the context of spatial audio, which is a area of active research stimulated by the rise of interactive multimedia including VR/AR. Among various approaches which include sparse microphone arrays based on empirical audio engineering for surround sound recording [10–13] and wave field synthesis [14], a common approach for first-person spatial audio is the use of dense microphone arrays which are used to convert the raw microphone array signals to a spatial audio representation which is ideally independent from both the recording system and the playback system [15–18]. The spatial audio paradigm known as *ambisonics* [15,16] achieves this decoupling of recording, transmission, and rendering, and has become one of the most widely accepted spatial audio description format.

1.2.2 Sound synthesis of sources and environments

Capturing sound is not the only way to obtain audio for playback. Sound can be simulated or synthesized. While the difference between simulation and synthesis is not always clear, in this dissertation I use the term *simulation* for computation which rely on first-principles or governing equations with physical and mathematical justification, while *synthesis* is used for computation which may involve phenomenological or empirical models or approximations which may not necessarily have physical justification. In

audio engineering, the requirement of physical correctness can be sometimes relaxed or is not required at all, which is particularly true for artistic or gaming applications. For example, synthesized sound scenes may not precisely reflect the acoustics of the real world, but may be sufficient to be used in a VR gaming environment. The target to be simulated or synthesized can either be a sound source [19, 20], or an impulse response, e.g., reverberation of indoor [21] or outdoor environments [8, 22]. Simulating or synthesizing room impulse responses has been a subject of extensive study [23–27], as it finds important applications in e.g., concert hall or architectural design or rendering virtual environments in VR. Physically rigorous room acoustics simulation tend to be computationally too expensive for real-time applications, hence challenges remain for achieving a good balance between perceptual quality and computation cost.

1.2.3 Sound localization and binaural audition

Sound source localization is a task to localize the position or direction of one or multiple sound sources using one or more sensors, typically microphones. Due to its practical importance, it is an active field of research [28]. Humans and various animals are able to localize sound with surprisingly high precision with only two sensors - the left and the right ear. This mechanism is called binaural audition or binaural hearing and has a long history of study [29]. An important aspect of binaural hearing is that it is largely individual-dependent due to the highly varying geometry of the ears among individuals. Hence, individualization of the signal processing mimicking spatial hearing is an active field of research [30–38].

1.2.4 Controlling sound

Passive or active control of sound is an important application of acoustics and digital signal processing. Headphones with active noise cancellation (ANC) [39, 40] have become common. Yet, challenging problems remain, such as noise control in open spaces or spatial ANC of sound fields [41–43].

1.2.5 Numerical modeling and simulation of sound

1.2.5.1 Preliminaries: the weak form and boundary integral form of the Helmholtz equation

By multiplying an arbitrary weighing function $w(\mathbf{r})$ with the Helmholtz equation (1.12) and integrating over the domain Ω , a volumetric integral equation is obtained:

$$-\int_{\Omega} w(\mathbf{r})(\nabla^2 u(\mathbf{r}) + k^2 u(\mathbf{r}))d\Omega = \int_{\Omega} w(\mathbf{r})f(\mathbf{r})d\Omega. \quad (1.13)$$

This volumetric integral equation is called the weak form of the Helmholtz equation.

Using Green's theorem, we obtain

$$-\int_{\Omega} u(\mathbf{r})(\nabla^2 w(\mathbf{r}) + k^2 w(\mathbf{r}))d\Omega - \int_{\Gamma} \left(w(\mathbf{r}) \frac{\partial u(\mathbf{r})}{\partial \mathbf{n}^{(-)}} - u(\mathbf{r}) \frac{\partial w(\mathbf{r})}{\partial \mathbf{n}^{(-)}} \right) d\Gamma = \int_{\Omega} w(\mathbf{r})f(\mathbf{r})d\Omega, \quad (1.14)$$

where Γ is the boundary of Ω and $\partial/\partial \mathbf{n}^{(-)}$ is the outward normal derivative. By substituting

$w(\mathbf{r}) = G(\mathbf{r}_p, \mathbf{r})$, where $G(\mathbf{r}_p, \mathbf{r}_q) = \frac{\exp(ik|\mathbf{r}_p - \mathbf{r}_q|)}{4\pi|\mathbf{r}_p - \mathbf{r}_q|}$ is the free-field Green function of the

Helmholtz equation in 3D, we obtain:

$$\begin{aligned} \int_{\Omega} u(\mathbf{r})\delta(\mathbf{r}_p, \mathbf{r})d\Omega - \int_{\Gamma} \left(G(\mathbf{r}_p, \mathbf{r})\frac{\partial u(\mathbf{r})}{\partial \mathbf{n}^{(-)}} - u(\mathbf{r})\frac{\partial G(\mathbf{r}_p, \mathbf{r})}{\partial \mathbf{n}^{(-)}} \right) d\Gamma &= \int_{\Omega} G(\mathbf{r}_p, \mathbf{r})f(\mathbf{r})d\Omega \\ \Rightarrow c_p u(\mathbf{r}_p) &= \int_{\Gamma} \left(G(\mathbf{r}_p, \mathbf{r})\frac{\partial u(\mathbf{r})}{\partial \mathbf{n}^{(-)}} - u(\mathbf{r})\frac{\partial G(\mathbf{r}_p, \mathbf{r})}{\partial \mathbf{n}^{(-)}} \right) d\Gamma + \int_{\Omega} G(\mathbf{r}_p, \mathbf{r})f(\mathbf{r})d\Omega. \end{aligned} \quad (1.15)$$

Here, c_p is a constant depending on the solid angle of the boundary at \mathbf{r}_p . At a smooth surface point, $c_p = \frac{1}{2}$. This equation is called the Kirchhoff-Helmholtz boundary integral equation (BIE). The Green function satisfies the inhomogeneous Helmholtz equation:

$$-\nabla^2 G(\mathbf{r}_p, \mathbf{r}_q) - k^2 G(\mathbf{r}_p, \mathbf{r}_q) = \delta(|\mathbf{r}_p - \mathbf{r}_q|) \quad (1.16)$$

with $\delta(x)$ the Dirac's delta function.

1.2.5.2 Numerical methods for acoustic simulation

Modeling and simulating sound has tremendous applications in the manufacturing industry, where simulation, or *computer-aided engineering* (CAE), can be used to study the properties of the products before or without making real physical prototypes. This helps to improve the quality, reduce the cost, and shorten the delivery time of the products. Various numerical methods are employed in the analysis of acoustical phenomena. Different methods have their own pros and cons. The Finite Element Method (FEM) discretizes the weak form (1.13) of the PDE by splitting the domain into volumetric finite elements in the case of a three-dimensional domain. The FEM boils down to solving a sparse linear system whose size scales as $O(D^3)$ due to the volumetric nature, where D is

the domain size of the problem. Another popular method in numerical acoustics is the Finite Difference Time-Domain (FDTD) method, which is a finite difference method discretizing both time and space using Cartesian grids, also known as voxels. The FDTD method is another volumetric method which may be suited for time-domain problems due to its formulation in the time-domain. The Boundary Element Method (BEM) discretizes the BIE (1.15) which was obtained from the weak form (1.13) after the dimensionality reduction via Green's theorem. This dimensionality reduction allows the method to scale as $O(D^2)$ for three-dimensional domains. The BEM, too, boils down to solving a linear system, which, unlike the FEM, is dense. Fortunately, the solution of this dense linear system can be accelerated by iterative solvers [44] and the Fast Multipole Method (FMM) [45]. The BEM is particularly suited for problems which involve infinite domains, which frequently arise in scattering and radiation problems in acoustics. Infinite domains are naturally treated in the BEM using the Sommerfeld radiation condition [46]:

$$\lim_{r \rightarrow \infty} r^{\frac{n-1}{2}} \left(\frac{\partial}{\partial r} - ik \right) u(\mathbf{r}) = 0, \quad (1.17)$$

with $r = |\mathbf{r}|$. However, its treatment is nontrivial in the FEM or FDTD methods as they require designing special boundary conditions to compensate for their formulation or limitation to finite domains [47]. In this dissertation I focus on the BEM as the method for solving the Helmholtz equation numerically. This is also motivated by its wide use in audio and acoustics engineering. For example, the BEM has been heavily used for the analysis of head-related transfer functions [48–51], which is one of the most important applications of numerical acoustics for modern spatial audio technologies

including Virtual Reality (VR) and Augmented Reality (AR).

1.2.5.3 Layer potential integrals arising in the boundary element method

We define four boundary integral operators arising in the BEM for solving the Helmholtz equation:

$$\begin{aligned}
\{V\psi\}(\mathbf{r}_p) &\equiv \int_{\mathbf{r}_q \in S} G(\mathbf{r}_p, \mathbf{r}_q) \psi(\mathbf{r}_q) dS^{(q)}, \\
\{K\phi\}(\mathbf{r}_p) &\equiv \int_{\mathbf{r}_q \in S} \frac{\partial G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_q} \phi(\mathbf{r}_q) dS^{(q)}, \\
\{K'\psi\}(\mathbf{r}_p) &\equiv \int_{\mathbf{r}_q \in S} \frac{\partial G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_p} \psi(\mathbf{r}_q) dS^{(q)}, \\
\{D\phi\}(\mathbf{r}_p) &\equiv - \int_{\mathbf{r}_q \in S} \frac{\partial^2 G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_p \partial \mathbf{n}_q} \phi(\mathbf{r}_q) dS^{(q)},
\end{aligned} \tag{1.18}$$

where \mathbf{n}_q is the outward unit normal vector at point \mathbf{r}_q , and ψ, ϕ are arbitrary density functions. The normal derivative is defined as $\frac{\partial}{\partial \mathbf{n}_q} \equiv \mathbf{n}_q \cdot \nabla_{\mathbf{r}_q}$. The boundary integrals associated with operators V, K, K' , and D are referred to as the single layer potential, double layer potential, adjoint double layer potential, and hypersingular potential, respectively.

The entries of the system matrix arising in the BEM after discretizing the BIE are typically linear combinations of these four layer potential integrals evaluated over boundary elements.

We define the boundary trace operator $\gamma_{0,q}$ and the outward normal derivative operator $\gamma_{1,q}$ with respect to \mathbf{r}_q as :

$$\{\gamma_{0,q}u\}(\mathbf{r}_q) \equiv \lim_{\hat{\mathbf{r}}_q \in \Omega \rightarrow \mathbf{r}_q \in \Gamma} u(\hat{\mathbf{r}}_q), \quad \{\gamma_{1,q}u\}(\mathbf{r}_q) \equiv \mathbf{n}_q \cdot \nabla_q u(\mathbf{r}_q), \quad \mathbf{r}_q \in S = \partial\Omega, \tag{1.19}$$

The Kirchhoff-Helmholtz boundary integral equation in (1.15) can be written concisely using these boundary integral operators:

$$\{(c_p\gamma_{0,p} + K\gamma_{0,q} - V\gamma_{1,q})u\}(\mathbf{r}_p) = \{N_0f\}(\mathbf{r}_p). \quad (1.20)$$

Here, $u(\mathbf{r})$ is the total sound pressure at position \mathbf{r} , \mathbf{r}_p is the observation point. N_0 is the volume potential (Newton potential) integral operator over the domain Ω defined as:

$$\{N_0f\}(\mathbf{r}_p) = \int_{\mathbf{r} \in \Omega} G(\mathbf{r}_p, \mathbf{r})f(\mathbf{r})d\Omega, \quad (1.21)$$

with f the source distribution.

1.3 Organization of the dissertation

The dissertation is organized as follows.

1.3.1 Capture

In chapter 2 and chapter 3, two novel techniques for three-dimensional sound field capturing and encoding are introduced. Both are based on (semi-)analytical cancellation of scattering caused by microphone arrays mounted on acoustic scatterers.

1.3.2 Synthesis of an acoustic environment

In chapter 4, I introduce a fast algorithm for synthesizing acoustic impulse responses in large-scale forests, which is used in chapter 5 to study the performance of acoustic

wildlife monitoring systems based on large-scale distributed microphone arrays.

1.3.3 Localization

In chapter 5, I present an empirical study for understanding the performance of acoustic wildlife monitoring systems based on large-scale distributed microphone arrays. This study involves a fast synthesis method for forest acoustics, which is introduced in chapter 4. In chapter 6, I propose a novel general-purpose individual-agnostic binaural localizer which supports sound source localization from arbitrary directions without a priori knowledge of the process generating the binaural signal.

1.3.4 Control

In chapter 7 and chapter 8, I develop frameworks for regularized active sound control, using either point- or mode-control and using either distributed or local worn loudspeaker and microphone arrays with applications including speech privacy, personal active noise control, and local crosstalk cancellation with limited noise injection into the environment.

1.3.5 Boundary element analysis

In chapter 9, chapter 10 and chapter 11, three numerical methods for evaluating integrals arising in the (fast multipole accelerated) boundary element method are introduced. In chapter 9, a recursive algorithm is developed which allows efficient analytical evaluation of singular and nearly singular layer potential integrals arising in the boundary element

method using flat high-order elements for Helmholtz and Laplace equations. In chapter 10, a differential geometry-based quadrature algorithm is developed which allows accurate evaluation of singular and nearly singular layer potential integrals arising in the boundary element method using smooth manifold boundary elements with constant densities for Helmholtz and Laplace equations. In chapter 11, an algorithm for efficient exact evaluation of integrals of regular solid harmonics over high-order boundary elements with simplex geometries is developed.

1.4 Special functions used in this dissertation

Here, some special functions are defined, which serve as our basic tools for the numerical analysis. It is particularly important to clearly state the definitions of these functions, since some of them have many different conventions.

1.4.1 Legendre polynomials

Definition (Rodrigues' formula):

$$P_n(x) \equiv \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \quad (n \geq 0) \quad (1.22)$$

1.4.2 Associated Legendre polynomials

Definition for non-negative orders:

$$P_n^m(x) \equiv (-1)^m (1 - x^2)^{m/2} \frac{d^m}{dx^m} (P_n(x)) \quad (1.23)$$

Definition for negative orders ($-m < 0$):

$$P_n^{-m}(x) \equiv (-1)^m \frac{(n-m)!}{(n+m)!} P_n^m(x) \quad (1.24)$$

1.4.3 Spherical harmonics

$$Y_n^m(\theta, \varphi) \equiv \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos \theta) e^{im\varphi}, \quad (1.25)$$

where θ and φ are the polar and azimuthal angular variables, respectively, of a *physical* spherical coordinate system defined in [52].

Chapter 2: Multiple scattering ambisonics: three-dimensional sound field estimation using interacting spheres

Rigid spherical microphone arrays (RSMAs) have been widely used in *ambisonics* [15] sound field recording. While it is desired to combine the information captured by a grid of densely arranged RSMAs for expanding the area of accurate reconstruction, or *sweet-spots*, this is not trivial due to inter-array interference. Here we propose *multiple scattering ambisonics*, a method for three-dimensional ambisonics sound field recording using multiple acoustically interacting RSMAs. Numerical experiments demonstrate the sweet-spot expansion realized by the proposed method. The proposed method can be used with existing RSMAs as building blocks and opens possibilities including higher degrees-of-freedom spatial audio.¹

2.1 Introduction

Audio is indispensable in immersive technologies such as mixed reality (MR) and virtual reality (VR), which are receiving much attention. For these applications, it is essential to develop technologies to capture, process, and render spatial sound fields with high precision for the presentation of truly realistic and immersive MR/VR experiences.

¹The research presented in this chapter has been published in [53].

Ambisonics [15] as well as **higher-order ambisonics** (HOA) [16], which are established spatial audio frameworks to capture, process and reproduce spatial sound fields based on its representation in the spherical harmonics domain, are receiving much attention due to the popularization of MR/VR platforms [54, 55], and its high compatibility with first-person view MR/VR. Ambisonics spatial audio capturing and processing consists of a microphone array and signal processing that is used to encode the raw microphone array signal to the spherical harmonics-domain spatial description format, which is referred to the ambisonics signal. This ambisonics signal is decoded to the signal which is fed to loudspeaker arrays to render the spatial sound field. Such loudspeaker arrays are often virtualized by means of binaural technologies [56–58] and played back via headphones. Hence the high compatibility of ambisonics with MR/VR applications that usually use headphones for audio playback. Due to its formulation in the spherical harmonics-domain, a typical implementation of an ambisonics recording device is employing a **spherical microphone array** (SMA) [15, 16, 18, 59, 60]. Often, SMAs are mounted on sound-hard spherical scatterers in order to avoid the instability arising in encoding filters for hollow microphone arrays due to singularities originating from the roots of the spherical Bessel function [16], and for its mechanical stability as hardware. This form of a SMA is referred to as a **rigid SMA** (RSMA). Despite its success in first-person immersive audio with only three **degrees-of-freedom** (DoF) which are associated with the rotation of the listener, ambisonics suffers from the diminishing size of the accurate reconstruction area, referred to as the *sweet-spot*, as the frequency increases, hence limiting its efficacy in higher DoF spatial audio reproduction allowing translation of the listener. This is visualized in Fig. 2.1 (left), showing the resulting reconstruction sweet-spots for incident

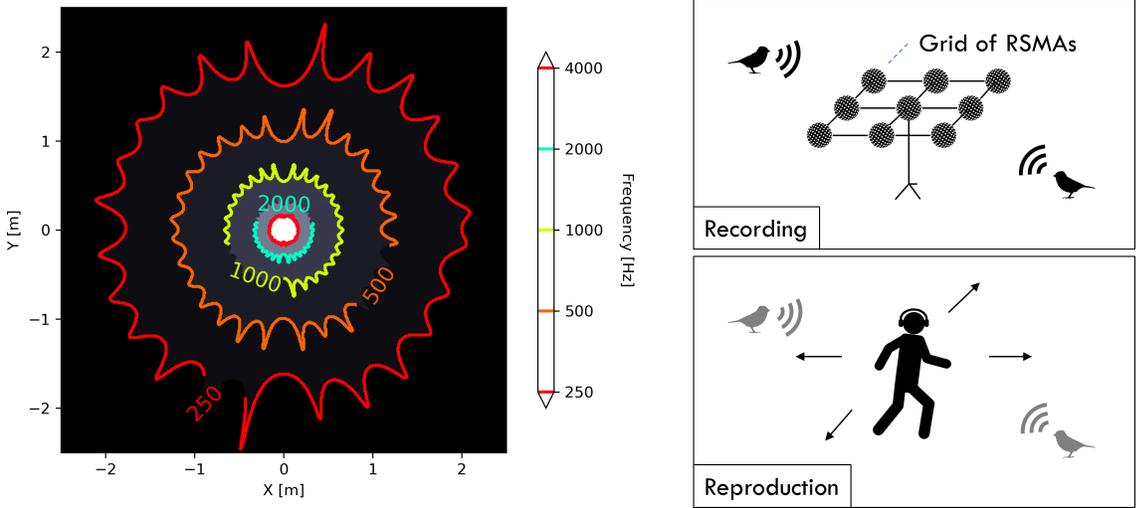


Figure 2.1: Left: examples of reconstruction sweet-spots using a 252-channel RSMA for estimating incident fields with various frequencies. Details of the RSMA are described in Section 8.4. The truncation degree of HOA which provides the largest sweet-spot is chosen for each frequency independently. Right: illustration of a sound field recording setup using MS-HOA (top) and its reproduction setup over headphones allowing translation of the listener (bottom).

plane waves with various frequencies. Here, the sweet-spot is defined as the region where the signal-to-distortion ratio (SDR) of the estimated field with respect to the ground truth incident field is above 30 dB. In order to expand the sweet-spot of ambisonics reproduction, the simplest way is to develop RSMA's with larger number of microphones. Although this is an effective approach, it comes with a significant development and device cost. An alternative approach is to combine multiple existing RSMA's and integrate the captured information. However, this is not a trivial task due to the inter-array interference. Here, *multiple scattering higher-order ambisonics* (MS-HOA), a three-dimensional (3D) sound field capturing scheme using multiple RSMA's with fully considering inter-array interaction due to multiple scattering [61] is proposed. Numerical experiments show that MS-HOA successfully creates sound field representations with expanded sweet-spots even when the RSMA's are densely arranged with small spacing, which is not achieved

without the consideration of inter-array interaction. An example sound field recording and reproduction setup allowing translation of the listener is illustrated in Fig. 2.1 (right).

2.2 Conventional ambisonics encoding using a single RSMA

The conventional framework of ambisonics encoding using a single RSMA is briefly reviewed. Ambisonics encoding and decoding can be performed by either relying on solving a linear system using least squares [16] or relying on spherical harmonic transformation using numerical integration [62]. Since the first approach allows more flexibility of the microphone array configuration, this approach is adopted here. In the present work, all formulations are presented in the frequency-domain, which can be converted to time-domain representations by inverse Fourier transform. All individual microphone capsules are assumed to be omnidirectional. The spherical harmonics used are defined as

$$Y_n^m(\theta, \varphi) \equiv \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos\theta) e^{im\varphi}, \quad (2.1)$$

with θ and φ the polar and azimuthal angle, respectively, and $P_n^m(x)$ and $P_n(x)$ respectively the associated and regular Legendre polynomials:

$$P_n^m(x) \equiv (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} (P_n(x)), \quad P_n(x) \equiv \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2-1)^n. \quad (2.2)$$

The above definition of spherical harmonics provides an orthonormal basis:

$$\int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} Y_n^m(\theta, \varphi) Y_{n'}^{m'}(\theta, \varphi)^* d\Omega = \delta_{nn'} \delta_{mm'}, \quad (2.3)$$

with δ_{ij} the Kronecker delta.

The process of obtaining the ambisonics signal $A_n^m(k)$, the weights of the spherical basis functions of the three dimensional sound field representing an arbitrary incident field of wavenumber k , from the signal captured by the microphone array is referred to as ambisonics *encoding*. An arbitrary incident field can be expanded in terms of the regular spherical basis functions $j_n(kr)Y_n^m(\theta, \varphi)$ of the three-dimensional Helmholtz equation in the spherical coordinate system (r, θ, φ) :

$$p_{\text{in}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m(k) j_n(kr) Y_n^m(\theta, \varphi), \quad (2.4)$$

with $j_n(x)$ the spherical Bessel function of degree n . The total field p_{tot} , which is the sum of the incident field and the field scattered by a rigid sphere with radius R located at O , the origin, is given by:

$$p_{\text{tot}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m(k) \left\{ j_n(kr) - h_n(kr) \frac{j_n'(kR)}{h_n'(kR)} \right\} Y_n^m(\theta, \varphi), \quad (2.5)$$

with $h_n(x)$ the spherical Hankel function of the first kind with degree n [2]. On the surface of the rigid sphere, i.e. $r = R$, this total field is evaluated as:

$$p_{\text{tot}}|_{r=R} = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m(k) \frac{i}{(kR)^2 h_n'(kR)} Y_n^m(\theta, \varphi). \quad (2.6)$$

The total field captured by the q -th microphone on the surface of the RSMA located at

(R, θ_q, φ_q) is therefore given by:

$$p_{\text{tot}}^{(q)} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \frac{i}{(kR)^2 h'_n(kR)} Y_n^m(\theta_q, \varphi_q) A_n^m(k). \quad (2.7)$$

By truncating the infinite series with $n = N_c$, this result can be represented in the following vector form:

$$\mathbf{p}_{\text{tot}} = \Lambda \mathbf{A}, \quad (2.8)$$

where \mathbf{p}_{tot} is a vector holding $p_{\text{tot}}^{(q)}$ in its q -th entry, \mathbf{A} is a vector holding $A_n^m(k)$ in its $(n^2 + n + m + 1)$ -th entry, and Λ is the matrix holding $\frac{i}{(kR)^2 h'_n(kR)} Y_n^m(\theta_q, \varphi_q)$ in its $(q, n^2 + n + m + 1)$ entry. The goal of ambisonics encoding is to obtain $A_n^m(k)$ for all ns and ms up to the truncation degree $n = N_c^{(\text{in})}$, i.e. $0 \leq n \leq N_c$ and $|m| \leq n$, from the observation \mathbf{p}_{tot} . This problem can be solved by regularized least squares with a minimization objective:

$$L_{\text{enc}} = \|\mathbf{p}_{\text{tot}} - \Lambda \mathbf{A}\|_2^2 + \sigma \|\mathbf{A}\|_2^2, \quad (2.9)$$

with σ a regularization parameter, and the solution given by:

$$\mathbf{A}^{(\text{est})} = \underset{\mathbf{A}}{\text{argmin}} L_{\text{enc}} = (\Lambda^H \Lambda + \sigma I)^{-1} \Lambda^H \mathbf{p}_{\text{tot}} = E \mathbf{p}_{\text{tot}}, \quad (2.10)$$

where $E \equiv (\Lambda^H \Lambda + \sigma I)^{-1} \Lambda^H$ is the regularized encoding matrix.

2.3 Proposed method

In the proposed method, a grid of multiple RSMAs is used to estimate p_{in} (3.4). The goal of ambisonics encoding in MS-HOA is to estimate $A_n^m(k)$ (3.4), for $0 \leq n \leq N_c$ and $|m| \leq n$ from observations of the sound pressure at discrete microphone capsule positions mounted on the surfaces of multiple RSMAs. In the following, a system of $N_s \geq 2$ RSMAs where each RSMA has a radius a_s is considered. Here, s is the index of the RSMA. Hereafter, the argument k is omitted from $A_n^m(k)$.

2.3.1 The forward problem: multiple scattering due to an arbitrary incident field

It is known that the problem of multiple scattering in a system of multiple spherical scatterers, i.e. computing the scattered field p_{scat} given p_{in} and the configuration of the scattering spheres, can be solved analytically [61, 63]. This problem is referred to as the *forward problem*. The procedure of solving the forward problem is briefly described here. First, A_n^m , the expansion coefficients at O (3.4) truncated at degree $n = N_c^{(\text{in})}$, are translated to the positions of the RSMAs using the translation operators $T_{R|R}^{(s,O)}$ resulting in the expansions $\mathbf{A}^{(s)} = T_{R|R}^{(s,O)} \mathbf{A}^{(\text{in})}$, where $\mathbf{A}^{(s)}$ and $\mathbf{A}^{(\text{in})}$ are vectors of length $L^{(\text{in})} \equiv (N_c^{(\text{in})} + 1)^2$ holding $A_n^{m(s)}$ and A_n^m in its $(n^2 + n + m + 1)$ -th entry, respectively. The $\mathbf{A}^{(s)}$ coefficients are then further truncated at degree $n = N_c^{(\text{fwd})}$. Two distinct truncation numbers $N_c^{(\text{in})}$ and $N_c^{(\text{fwd})}$ are introduced here in order to achieve sufficient accuracy of the translation operation while limiting numerical error in the computation of the scattered

field. Given the set of expansions $A_n^{m(s)}$ at each RSMA position, the contribution of each RSMA to the scattered field $B_n^{m(s)}$ can be computed by solving the linear system:

$$\mathbf{A}' = S\mathbf{B}', \quad (2.11)$$

where \mathbf{A}' and \mathbf{B}' are concatenations of N_S vectors $\{\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(N_S)}\}$ and $\{\mathbf{B}^{(1)}, \mathbf{B}^{(2)}, \dots, \mathbf{B}^{(N_S)}\}$, where $\mathbf{B}^{(s)}$ are vectors of length $L^{(\text{fwd})} \equiv (N_c^{(\text{fwd})} + 1)^2$ holding $B_n^{m(s)}$ in its $(n^2 + n + m + 1)$ -th entry. S is referred to as the *system matrix*, which is a block matrix holding the inter-sphere translation operator $T_{S|R}^{(s,t)}$ from the t -th sphere to the s -th sphere in its off-diagonal (s, t) -block and the “single scattering matrix” $\Lambda^{(s)}$ in its diagonal blocks:

$$S = \begin{pmatrix} \Lambda^{(1)} & -T_{S|R}^{(1,2)} & \dots & -T_{S|R}^{(1,N_S)} \\ -T_{S|R}^{(2,1)} & \Lambda^{(2)} & \dots & -T_{S|R}^{(2,N_S)} \\ \vdots & \vdots & \ddots & \vdots \\ -T_{S|R}^{(N_S,1)} & -T_{S|R}^{(N_S,2)} & \dots & \Lambda^{(N_S)} \end{pmatrix}, \quad (2.12)$$

where $\Lambda^{(s)}$ is a diagonal matrix holding $-\frac{h'_n(ka_s)}{j'_n(ka_s)}$ in its (l, l) entry with $l = n^2 + n + m + 1$.

The translation operators $T_{R|R}^{(s,O)}$ and $T_{S|R}^{(s,t)}$ can be computed by various methods, including explicit expressions based on Clebsch-Gordan coefficients or Wigner 3-j symbols [64], or methods based on recurrence relations [65]. The total field p_{tot} evaluated at $\mathbf{r}_q^{(s)}$, the q -th microphone position belonging to the s -th RSMA, is the sum of the scattered field

contributions from all the RSMAs and the incident field p_{in} :

$$p_{\text{tot}}(\mathbf{r}_q^{(s)}) = p_{\text{scat}}(\mathbf{r}_q^{(s)}) + p_{\text{in}}(\mathbf{r}_q^{(s)}) = \sum_{n=0}^{N_c^{(\text{fwd})}} \sum_{m=-n}^n \left(\sum_{t=1}^{N_S} B_n^{m(t)} S_n^{m(t)}(\mathbf{r}_q^{(s)}) + A_n^{m(s)} R_n^{m(s)}(\mathbf{r}_q^{(s)}) \right), \quad (2.13)$$

where $R_n^{m(s)}(\mathbf{r}_q^{(s)})$ and $S_n^{m(t)}(\mathbf{r}_q^{(s)})$ are the regular and singular spherical basis functions expanded at the location of the s -th and t -th sphere, respectively, and evaluated at the position of the q -th microphone capsule belonging to the s -th RSMA:

$$R_n^{m(s)}(\mathbf{r}_q^{(s)}) = j_n(ka_s) Y_n^m(\mathbf{r}_q^{(s)} - \mathbf{r}_s), \quad S_n^{m(t)}(\mathbf{r}_q^{(s)}) = h_n(k|\mathbf{r}_q^{(s)} - \mathbf{r}_t|) Y_n^m(\mathbf{r}_q^{(s)} - \mathbf{r}_t). \quad (2.14)$$

Alternatively, $p_{\text{in}}(\mathbf{r}_q^{(s)})$ could be evaluated directly using A_n^m instead of the translated $A_n^{m(s)}$ coefficients. The whole procedure of the forward problem can be expressed by a linear operator T_F which is referred to as the forward operator:

$$\mathbf{p}_{\text{tot}} = T_F \mathbf{A}^{(\text{in})}, \quad (2.15)$$

where \mathbf{p}_{tot} is a vector holding the values of $p_{\text{tot}}(\mathbf{r}_q^{(s)})$.

2.3.2 The inverse problem: MS-HOA encoding

The matrix representing T_F can be constructed by applying the operator to all bases up to $n \leq N_c^{(\text{in})}$. The estimate of the incident field can then be obtained via regularized least squares:

$$\mathbf{A}^{(\text{est})} = (T_F^H T_F + \sigma I)^{-1} T_F^H \mathbf{p}_{\text{tot}} = T_I \mathbf{p}_{\text{tot}}, \quad (2.16)$$

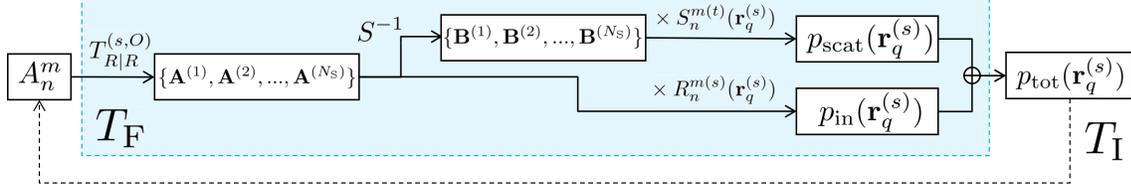


Figure 2.2: The scheme of the forward problem, i.e. computation of the total sound field at the microphone positions given the incident field, and the inverse problem, i.e. estimation of the incident field given the sound pressure at the microphone positions.

where $\mathbf{A}^{(\text{est})}$ is a vector holding the estimated coefficients $A_n^{m(\text{est})}$ in its $(n^2 + n + m + 1)$ -th entry up to $n \leq N_c^{(\text{in})}$ and $T_I \equiv (T_F^H T_F + \sigma I)^{-1} T_F^H$ is the encoding matrix for MS-HOA with σ a regularization parameter. The scheme of the forward and inverse problem is summarized in Fig. 2.2.

2.4 Numerical experiments

MS-HOA recording and encoding into HOA coefficients was validated by numerical experiments. Grids of RSMAs where each individual RSMA is a 252-channel SMA mounted on a rigid spherical scatterer with a radius of 8 cm are considered. The spherical Fibonacci grid [18, 66] of 252 points was used for the microphone capsule positions. A real-world implementation of a 252-channel RSMA with a similar size has been demonstrated in the past [60]. As the RSMA grid, a linear grid of 6 RSMAs and a regular Cartesian grid of 9 RSMAs was used in the experiments. The spacing between the nearest neighbour RSMA was set to 25 cm. The sound field generated by a monopole source located at $\mathbf{r}_s = (10\text{m}, 10\text{m}, 10\text{m})$ was used as the incident field. The signal captured by the grid of RSMAs was encoded into the HOA coefficients $\mathbf{A}^{(\text{est})}$ with the proposed method (**MS-HOA**). While prior works on the forward problem report heuristics for choosing the

parameter $N_c^{(\text{fwd})}$, e.g. $N_c^{(\text{fwd})} = \lfloor eka \rfloor$ [63], here $N_c^{(\text{fwd})}$ was treated as a free hyperparameter. The case where inter-sphere interaction is switched off, i.e. the method which only considers single scattering (**Single**), and the case of conventional HOA encoding using only one building block RSMA (**HOA**) are also computed as baselines. The analytical reconstruction of the estimated incident field was computed by (3.4) and was compared to the ground truth incident field p_{in} in terms of the SDR and the size of the reconstruction **sweet-spot area** (SSA) measured in the xy -plane or the yz -plane depending on the configuration of the RSMA grid. The SSA is defined here as the total area where the SDR surpasses 30 dB, which is measured using the regular Cartesian grid points on a plane which correspond to the pixels in Fig. 2.3-Fig. 2.4. The regularization hyperparameter was optimized by grid search independently for both the **Single** baseline and the proposed **MS-HOA**. Regularization was not applied to the single sphere **HOA** baseline due to its minor effect for this case, while the truncation number N_c was chosen as the one providing the largest SSA for the given RSMA. The results for the linear 6-sphere RSMA grid with a incident field frequency of 4kHz are shown in Fig. 2.3. The sweet-spot of reconstruction is successfully expanded with MS-HOA while the SSA and SDR is significantly degraded if only single scattering is considered. The results for the regular Cartesian 9-sphere grid is shown in Fig. 2.4, demonstrating planar expansion of the sweet-spot.

2.5 Related work and discussion

Multiple scattering ambisonics, a method to capture 3D sound fields using multiple acoustically interacting RSMAs, was proposed. MS-HOA allows to integrate the information

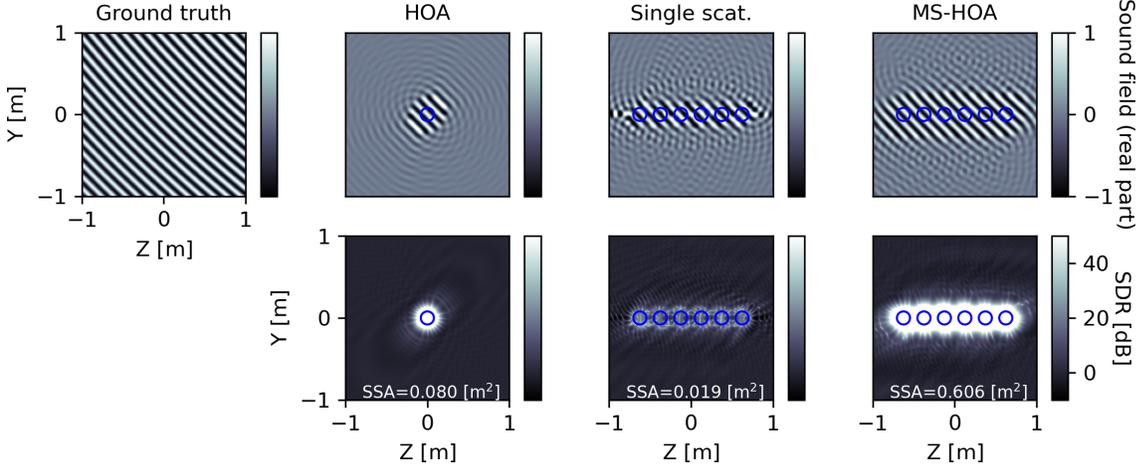


Figure 2.3: Results for the linear 6-sphere RSMA grid for an incident field of 4 kHz. Top row from left to right: real part of the sound pressure for the ground truth incident field, incident field estimated by **HOA**, **Single**, and **MS-HOA**, respectively. Bottom row from left to right: SDR map of the estimated fields with respect to the ground truth field using **HOA**, **Single**, and **MS-HOA**, respectively. The blue circles represent the positions and sizes of the RSMAs. The truncation number is $N_c = 14$ in **HOA**, $N_c^{(\text{in})} = 55$ and $N_c^{(\text{fwd})} = 20$ in **Single** and **MS-HOA**.

captured by multiple densely arranged RSMAs and can be used to expand the reconstruction sweet-spots in 3D sound field reproduction. The numerical experiments demonstrated that the proposed method successfully captures spatial sound fields with expanded reconstruction sweet-spots which was not possible without the consideration of inter-array interaction due to multiple scattering.

A related method using the translation of multipoles was introduced in [67]. This method was based on the assumption that the SMAs do not physically interact with each other, i.e. the SMAs do not cause scattering that affect other SMAs. This assumption is violated if the SMAs are densely arranged RSMAs, which scatter the incident field and interact with each other by multiple scattering. As shown in the numerical experiments, the approach without considering inter-array interaction becomes inaccurate if the RSMAs are arranged with small spacing. Recently, the consideration of inter-array multiple

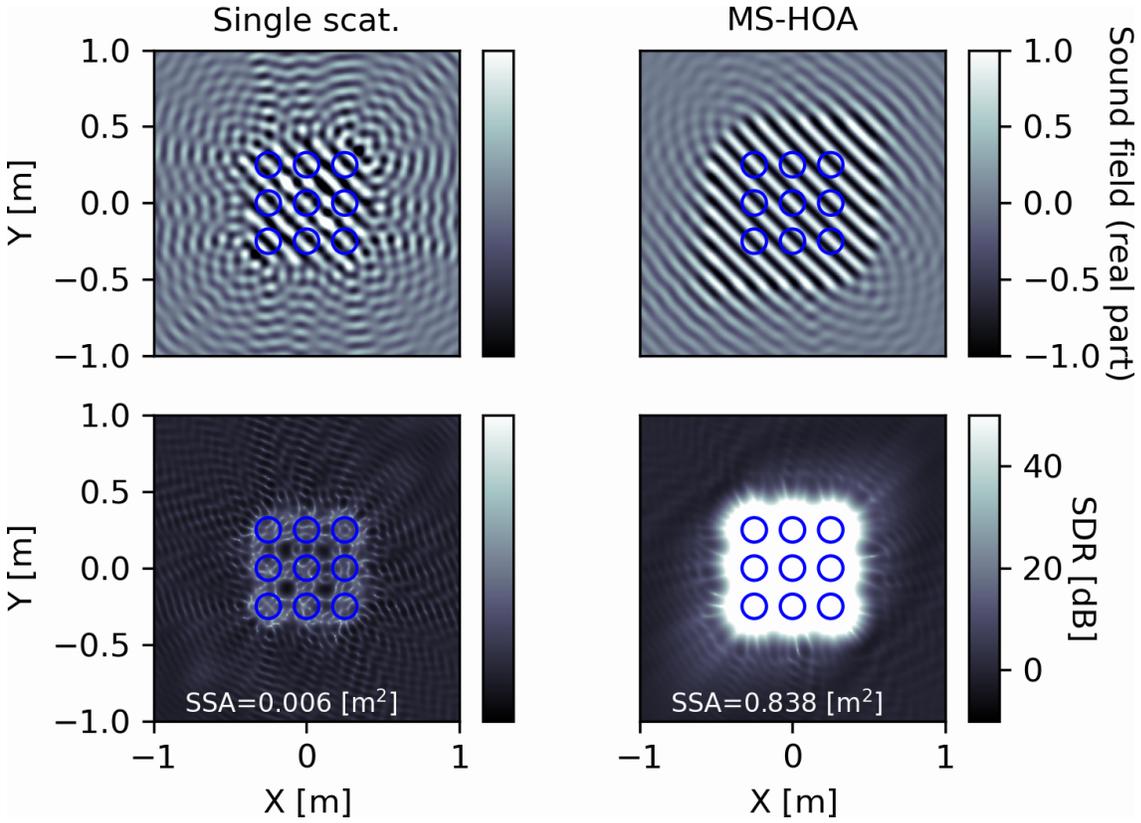


Figure 2.4: Results for the regular Cartesian 9-sphere RSMA grid. The estimated incident field (top row) and SDR map (bottom row) for **Single** (left column) and **MS-HOA** (right column), respectively. The blue circles represent the positions and sizes of the RSMAs. The same incident field as Fig. 2.3 is used. The truncation numbers are $N_c^{(\text{in})} = 45$ and $N_c^{(\text{fwd})} = 16$.

scattering has been demonstrated to improve the reconstruction accuracy in a two-dimensional sound field reconstruction problem using multiple cylindrical microphone arrays [68]. Two-dimensional modeling, however, is insufficient for modern spatial audio applications such as MR/VR where 3D audio representation and rendering is essential. Our work enables the use of interacting rigid microphone arrays for 3D spatial audio.

The expanded reconstruction sweet-spots with linear or planar spreads realized by the proposed method could be useful in applications including sound field reproduction in theaters or in meeting rooms where the sweet-spot should cover multiple listeners sitting next to each other, or higher DoF MR/VR where the translation of the listener needs to be supported. Developing techniques to reduce the cost of MS-HOA recording in terms of hardware, computation, and bandwidth is important for practical applications and are subjects of future research.

Chapter 3: Spheroidal ambisonics: spatial audio in spheroidal bases

Ambisonics is an established framework to capture and reproduce spatial sound fields based on the spherical harmonics representation [15]. A generalization – *spheroidal ambisonics* – based on spheroidal wave functions is proposed for use with spheroidal microphone arrays. Analytical conversion from spheroidal ambisonics to spherical ambisonics are derived to ensure compatibility with the existing ambisonics ecosystem. Numerical experiments verify spheroidal ambisonics encoding and transcoding for spatial sound field recording. The sound field reconstructed from the transcoded coefficients has a zone of accurate reconstruction which is prolonged towards the long axis of a prolate spheroidal microphone array.¹

3.1 Introduction

Immersive multimedia technologies such as augmented reality (AR) and virtual reality (VR) are receiving much attention. Audio is indispensable in these, and it is essential to be able to capture, process, and render spatial sound fields with high precision for presentation of plausible AR/VR and the creation of immersive experiences. The spatial audio framework of ambisonics [15] as well as higher-order ambisonics (HOA) [16]

¹The research presented in this chapter has been published in [69].

is receiving much attention due to the popularization of AR/VR, as well as the ability to stream this representation using standard platforms [54, 55], and its compatibility with first-person view AR/VR. Ambisonics spatial audio capturing and processing consists of a microphone array and signal processing algorithms that are used to encode the raw microphone array signals to the spherical harmonics-domain spatial description format, which is referred to as the ambisonics signal. This ambisonics signal is decoded to the signals which is fed to loudspeaker arrays to render the spatial sound field. Such loudspeaker arrays can also be virtualized by means of binaural technologies [56–58] and played over headphones. Hence the high compatibility of ambisonics with AR/VR applications that usually use headphones for audio playback. Due to its formulation in the spherical harmonics-domain, the most natural implementation of ambisonics recording devices is employing spherical microphone arrays [15, 16, 18, 59]. In this work, we generalize the framework of ambisonics into spheroidal coordinates and define *spheroidal ambisonics*, which uses spheroidal wave functions for the representation of spatial sound fields. While the use of spheroidal microphone arrays for sound field recording was claimed in a patent [17], its description was limited to the case of a spherical embodiment. We describe a formulation for the case of prolate spheroidal ambisonics, allowing the use of prolate spheroidal microphone arrays in an analytical manner in contrast to a recently proposed approach which allows arbitrary shaped microphone arrays but relies on numerical simulation to encode the captured field [70]. In addition to the basic formulation, an analytical conversion formula from spheroidal ambisonics to spherical ambisonics is derived. This allows the utilization of the existing ecosystem around spherical ambisonics after recording the spatial audio with a spheroidal microphone array. The overview of

the proposed schemes of spheroidal ambisonics encoding and transcoding is shown in Fig. 3.1. Numerical experiments are performed to validate and demonstrate spheroidal ambisonics encoding and transcoding when used for spatial sound field recording.

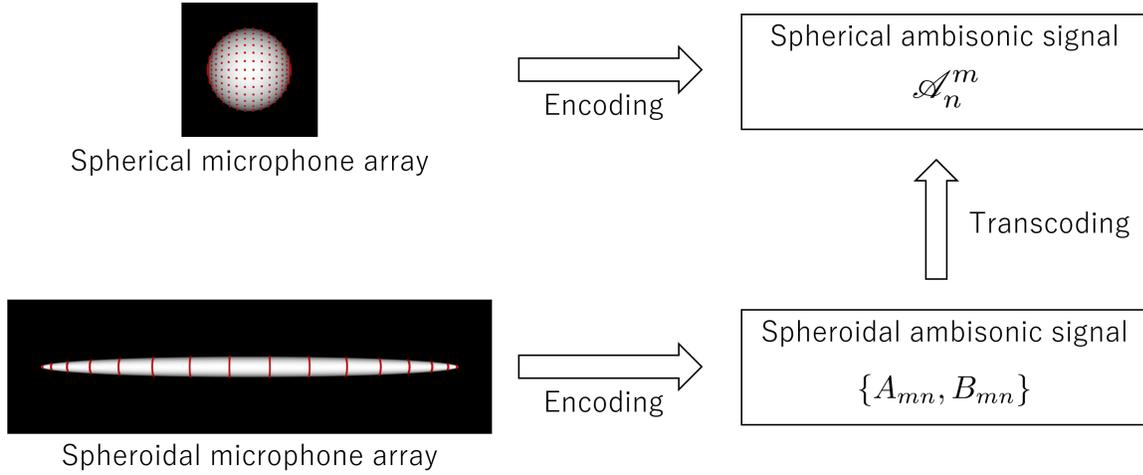


Figure 3.1: The overview of the proposed spheroidal ambisonics. The microphone capsule positions used in the numerical experiments presented in this paper are shown as the red dots in the images.

3.2 Background: spherical ambisonics

The conventional framework of ambisonics using spherical basis functions is reviewed here. Ambisonics encoding and decoding can be performed by either relying on solving a linear system using least squares [16] or relying on spherical harmonic transformations using numerical integration [62]. Since the first approach allows more flexibility of the microphone array configuration, this approach is adopted here. Throughout the paper, only microphone arrays mounted on surfaces of rigid scatterers are considered. This avoids the instability arising in encoding filters for hollow microphone arrays due to singularities originating from the roots of the spherical Bessel function [16]. All formulations are presented in the frequency-domain, which can be converted into the time-domain

representations by inverse Fourier transform. The spherical harmonics used are defined as

$$Y_n^m(\theta, \varphi) \equiv \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos\theta) e^{im\varphi}, \quad (3.1)$$

with θ and φ the polar and azimuthal angle, respectively, and $P_n^m(x)$ and $P_n(x)$ respectively the associated and regular Legendre polynomials:

$$P_n^m(x) \equiv (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} (P_n(x)), \quad P_n(x) \equiv \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2-1)^n. \quad (3.2)$$

The above definition of spherical harmonics provides an orthonormal basis:

$$\int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} Y_n^m(\theta, \varphi) Y_{n'}^{m'}(\theta, \varphi)^* d\Omega = \delta_{nn'} \delta_{mm'}, \quad (3.3)$$

with δ_{ij} the Kronecker delta.

3.2.1 Encoding in spherical ambisonics

The process of obtaining the ambisonics signal \mathcal{A}_n^m , the weights of the spherical basis functions of the three dimensional sound field representing an arbitrary incident field to the microphone array, from the signal captured by the microphone array is referred to as ambisonics *encoding*. An arbitrary incident field to the spherical microphone array mounted on a rigid sphere with radius R and located at O , the origin of the spherical coordinate system (r, θ, φ) , can be expanded in terms of the regular spherical basis functions

$j_n(kr)Y_n^m(\theta, \varphi)$ of the three-dimensional Helmholtz equation:

$$p_{\text{in}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \mathcal{A}_n^m(k) j_n(kr) Y_n^m(\theta, \varphi), \quad (3.4)$$

with $j_n(x)$ the spherical Bessel function of degree n and k the wavenumber. The total field p_{tot} , which is the sum of the incident field and the scattered field is given by:

$$p_{\text{tot}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \mathcal{A}_n^m(k) \left\{ j_n(kr) - h_n(kr) \frac{j'_n(kR)}{h'_n(kR)} \right\} Y_n^m(\theta, \varphi), \quad (3.5)$$

with $h_n(x)$ the spherical Hankel function of the first kind with degree n . On the surface of the rigid sphere, i.e. $r = R$, this total field is evaluated as:

$$p_{\text{tot}}|_{r=R} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \mathcal{A}_n^m(k) \frac{i}{(kR)^2 h'_n(kR)} Y_n^m(\theta, \varphi) \quad (3.6)$$

The total field captured by the q -th microphone located at (R, θ_q, φ_q) is therefore given by:

$$p_{\text{tot}}^{(q)} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \frac{i}{(kR)^2 h'_n(kR)} Y_n^m(\theta_q, \varphi_q) \mathcal{A}_n^m(k). \quad (3.7)$$

Truncating the infinite series at $n = N$, the equation can be written in vector form as

$$\mathbf{p}_{\text{tot}} = \Lambda^\bullet \mathbf{A}^\bullet \quad (3.8)$$

where \mathbf{p}_{tot} is a vector holding $p_{\text{tot}}^{(q)}$ in its q -th entry (in this paper, indices are 0-based), \mathbf{A}^\bullet is a vector holding $\mathcal{A}_n^m(k)$ in its $(n^2 + n + m)$ -th entry, and Λ^\bullet is the ‘‘inverse’’ encoding

matrix for rigid sphere microphone arrays which is a matrix holding $\frac{i}{(kR)^2 h'_n(kR)} Y_n^m(\theta_q, \varphi_q)$ in its $(q, n^2 + n + m)$ entry. The dot (\bullet) in the superscripts is used here to indicate variables associated to the spherical case, in order to distinguish them from the spheroidal case introduced later. The goal of ambisonics encoding is to obtain $\mathcal{A}_n^m(k)$ from the observation \mathbf{p}_{tot} . Typically, this is solved via regularized least squares minimization of:

$$L_{\text{enc}} = \|\mathbf{p}_{\text{tot}} - \Lambda^\bullet \mathbf{A}^\bullet\|_2^2 + \sigma \|\mathbf{A}^\bullet\|_2^2, \quad (3.9)$$

with σ a regularization parameter, and the solution given by:

$$\mathbf{A}_{\text{est}}^\bullet = \underset{\mathbf{A}^\bullet}{\text{argmin}} L_{\text{enc}} = (\Lambda^{\bullet H} \Lambda^\bullet + \sigma I)^{-1} \Lambda^{\bullet H} \mathbf{p}_{\text{tot}} = E \mathbf{p}_{\text{tot}}, \quad (3.10)$$

where $E \equiv (\Lambda^{\bullet H} \Lambda^\bullet + \sigma I)^{-1} \Lambda^{\bullet H}$ is the regularized encoding matrix. The regularization parameter σ is used to prevent over-fitting via an output signal of excessive energy and can be determined by optimizing a cost metric of the user's choice.

A plane wave $p_{\text{in}}^{\text{pw}} = e^{i\mathbf{k}\cdot\mathbf{r}}$ with a wave vector in spherical coordinates $\mathbf{k} = (k, \theta_i, \varphi_i)$

can be written as below with ambisonics coefficients \mathcal{A}_n^m :

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n 4\pi i^n Y_n^m(\theta_i, \varphi_i)^* j_n(kr) Y_n^m(\theta, \varphi), \quad \mathcal{A}_n^{m(\text{pw})}(k) = 4\pi i^n Y_n^m(\theta_i, \varphi_i)^*. \quad (3.11)$$

3.3 Formulation of spheroidal ambisonics

The fact that the three-dimensional Helmholtz equation is separable in the spheroidal coordinate system allows us to formulate spheroidal ambisonics, and details are presented here.

3.3.1 Spheroidal coordinates

While there are two types of spheroidal coordinates - the prolate and the oblate - details for only the prolate spheroidal ambisonics are presented here. The case for oblate spheroidal coordinates can be derived in a similar fashion. The definition of prolate spheroidal coordinates itself has some variations [71]. The definition also used in [72] is employed here. The prolate spheroidal coordinate system has three coordinates ξ , η , and φ , which is also characterized by the parameter a , where $2a$ is the distance between the two foci of the prolate spheroid. The domain of ξ and η is $\xi \geq 1$ and $|\eta| \leq 1$, respectively. The conversion with the Cartesian coordinates (x, y, z) is given by:

$$x = a\sqrt{1 - \eta^2}\sqrt{\xi^2 - 1} \cos(\varphi), \quad y = a\sqrt{1 - \eta^2}\sqrt{\xi^2 - 1} \sin(\varphi), \quad z = a\eta\xi. \quad (3.12)$$

$$\Leftrightarrow \begin{cases} \xi = \frac{1}{2a}(\sqrt{x^2 + y^2 + (z + a)^2} + \sqrt{x^2 + y^2 + (z - a)^2}) \\ \eta = \frac{1}{2a}(\sqrt{x^2 + y^2 + (z + a)^2} - \sqrt{x^2 + y^2 + (z - a)^2}) \\ \varphi = \arctan\left(\frac{y}{x}\right) \end{cases} \quad (3.13)$$

The long radius r_{long} and short radius r_{short} of a prolate spheroid is related with a and ξ_1 by:

$$r_{\text{long}} = a\xi_1, \quad r_{\text{short}} = a\sqrt{\xi_1^2 - 1}; \quad a = \sqrt{r_{\text{long}}^2 - r_{\text{short}}^2}, \quad \xi_1 = \frac{r_{\text{long}}}{\sqrt{r_{\text{long}}^2 - r_{\text{short}}^2}} = \frac{r_{\text{long}}}{a} \quad (3.14)$$

The prolate spheroidal coordinates ξ and η are illustrated in Fig. 3.2 (left) for $a = 1$.

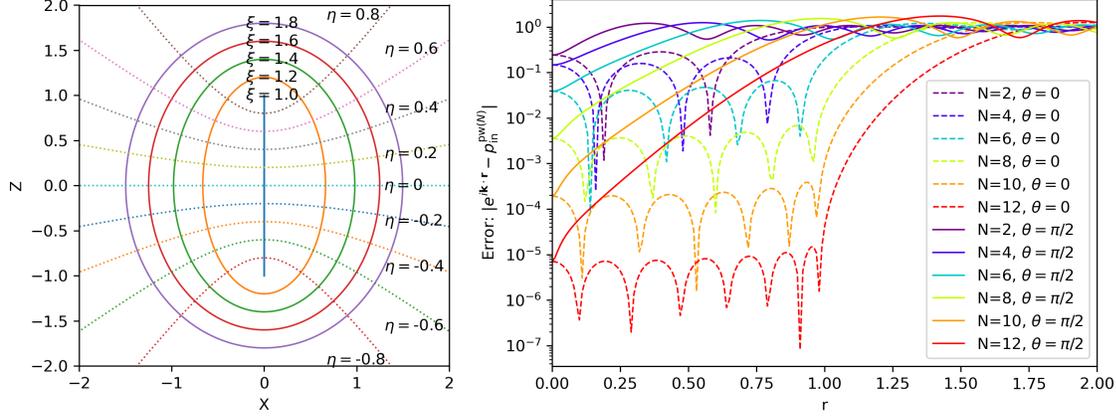


Figure 3.2: Left: prolate spheroidal coordinates ξ and η for $a = 1$. The lines of equal values of ξ and η are shown on the xz -plane, i.e. for $\varphi = \{0, \pi\}$. Right: truncation error $|e^{i\mathbf{k}\cdot\mathbf{r}} - p_{\text{in}}^{\text{pw}(N)}|$ of the series expansion of a plane wave (3.16) truncated at $n = N$ as a function of the spherical coordinate $(r, \theta, \varphi = 0)$ with $a = 1$, $\mathbf{k} = (2\pi f/v, 0, 0)$, $f = 541.8$ Hz, and $v = 340$ m/s.

3.3.2 Scattering of an arbitrary incident wave by a sound-hard prolate spheroid

An arbitrary incident wave can be expanded using radial spheroidal wave functions

$R_{mn}^{(1)}$ and angular spheroidal wave functions S_{mn} [71]:

$$p_{\text{in}} = \sum_{n=0}^{\infty} \sum_{m=0}^n R_{mn}^{(1)}(c, \xi) S_{mn}(c, \eta) (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi), \quad (3.15)$$

where $c = ka$ with k the wave number. The *spheroidal ambisonics coefficients* are defined as the collection of the $\{A_{mn}, B_{mn}\}$ coefficients. A canonical example of an incident wave is a plane wave $p_{\text{in}}^{\text{pw}} = e^{i\mathbf{k}\cdot\mathbf{r}}$ with $\mathbf{k} = k(\sin\theta_0 \cos\varphi_0, \sin\theta_0 \sin\varphi_0, \cos\theta_0)$ the wave vector represented in the Cartesian coordinates. The incident plane wave can be expanded as:

$$p_{\text{in}}^{\text{pw}} = \sum_{n=0}^{\infty} \sum_{m=0}^n \frac{2i^n \varepsilon_m}{N_{mn}(c)} R_{mn}^{(1)}(c, \xi) S_{mn}(c, \eta) S_{mn}(c, \cos\theta_0) \cos(m(\varphi - \varphi_0)), \quad (3.16)$$

which yields

$$A_{mn}^{\text{pw}} = \frac{2i^n \varepsilon_m}{N_{mn}(c)} S_{mn}(c, \cos\theta_0) \cos m\varphi_0, \quad B_{mn}^{\text{pw}} = \frac{2i^n \varepsilon_m}{N_{mn}(c)} S_{mn}(c, \cos\theta_0) \sin m\varphi_0. \quad (3.17)$$

The truncation error of the expression in (3.16), given by $|e^{i\mathbf{k}\cdot\mathbf{r}} - p_{\text{in}}^{\text{pw}(N)}|$ where $p_{\text{in}}^{\text{pw}(N)}$ is the series truncated at $n = N$, is shown for an example configuration in Fig. 3.2 (right).

The total field after scattering an arbitrary incident field characterized by $\{A_{mn}, B_{mn}\}$ is then given by:

$$p_{\text{tot}} = \sum_{n=0}^{\infty} \sum_{m=0}^n \left\{ R_{mn}^{(1)}(c, \xi) - \frac{R_{mn}^{(1)'}(c, \xi_1)}{R_{mn}^{(3)'}(c, \xi_1)} R_{mn}^{(3)}(c, \xi) \right\} S_{mn}(c, \eta) (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi). \quad (3.18)$$

On the surface of the spheroid, i.e. $\xi = \xi_1$, by using the Wronskian relation $W^{(1,3)} = R^{(1)}(c, \xi)R^{(3)'}(c, \xi) - R^{(1)'}(c, \xi)R^{(3)}(c, \xi) = \frac{i}{c(\xi^2-1)} = iW^{(1,2)}$, the total field can be

written as:

$$p_{\text{tot}}|_{\xi=\xi_1} = \sum_{n=0}^{\infty} \sum_{m=0}^n \frac{iS_{mn}(c, \eta)}{c(\xi_1^2 - 1)R_{mn}^{(3)'}(c, \xi_1)} (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi) \quad (3.19)$$

3.3.3 Spheroidal ambisonics encoding

The goal of spheroidal ambisonics encoding is to estimate the spheroidal ambisonics coefficients from observations by a limited number of microphones mounted on the surface of a spheroid-shaped baffle. As mentioned earlier, it is assumed here that the baffle is a sound-hard prolate spheroid.

By truncating the expansion order by $N > 0$, (3.19) can be rewritten in vector form:

$$\mathbf{p}_{\text{tot}} = (S^{(A)}, S^{(B)}) \begin{pmatrix} R^{(A)} & 0 \\ 0 & R^{(B)} \end{pmatrix} \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \equiv \Lambda^{(P)} \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}. \quad (3.20)$$

Here, \mathbf{A} and \mathbf{B} are vectors holding A_{mn} and B_{mn} in their $\tilde{l}^{(A)} = ((n^2 + n)/2 + m)$ -th and $\tilde{l}^{(B)} = ((n^2 - n)/2 + m - 1)$ -th entry, respectively. The lengths of these vectors are $L^{(A)} = \frac{(N+1)(N+2)}{2}$ and $L^{(B)} = \frac{N(N+1)}{2}$, respectively. $R^{(A)}$ and $R^{(B)}$ are diagonal matrices holding $\frac{i}{c(\xi_1^2 - 1)R_{mn}^{(3)'}(c, \xi_1)}$ in their $\tilde{l}^{(A)}$ -th and $\tilde{l}^{(B)}$ -th diagonal entries, respectively. $S^{(A)}$ and $S^{(B)}$ are matrices with entries:

$$S^{(A)}_{q, \tilde{l}^A(m,n)} = S_{mn}(c, \eta_q) \cos m\varphi_q, \quad S^{(B)}_{q, \tilde{l}^B(m,n)} = S_{mn}(c, \eta_q) \sin m\varphi_q, \quad (3.21)$$

respectively, where q the sensor index. \mathbf{p}_{tot} is a vector holding $p_{\text{tot}}^{(q)}$, the observed sound

pressure at the q -th microphone, in its q -th entry. \mathbf{p}_{tot} , $S^{(A)}$, and $S^{(B)}$ have shapes of $[Q]$, $[Q \times L^{(A)}]$, and $[Q \times L^{(B)}]$, respectively, where Q is the number of microphones. For a truncation order N , the total number of coefficients in \mathbf{A} and \mathbf{B} is $L = L^{(A)} + L^{(B)} = (N + 1)^2$, which is the same as the total number of spherical ambisonics coefficients $\{\mathcal{A}_n^m\}$ with maximum order N . $\Lambda^{(P)}$ is referred to as the “inverse” encoding matrix for sound-hard prolate spheroidal ambisonics.

The unknowns A_{mn} and B_{mn} can be estimated from observations of the sound field with multiple sensors mounted on the spheroidal baffle, by solving (3.20) with least squares. This process is referred to as *spheroidal ambisonics encoding*. The regularized least squares solution is given by:

$$(\mathbf{A}^T, \mathbf{B}^T)^T = (\Lambda^{(P)H} \Lambda^{(P)} + \sigma I)^{-1} \Lambda^{(P)H} \mathbf{p}_{\text{tot}} = E^{(P)} \mathbf{p}_{\text{tot}}, \quad (3.22)$$

with σ a regularization constant and $E^{(P)} \equiv (\Lambda^{(P)H} \Lambda^{(P)} + \sigma I)^{-1} \Lambda^{(P)H}$ the encoding matrix for sound-hard prolate spheroidal ambisonics.

3.4 Transcoding from spheroidal to spherical ambisonics

3.4.1 The transcoding formula

The sound field encoded as a spheroidal ambisonics signal can be converted into a conventional spherical ambisonics representation. This process is referred to as *transcoding*. The following relation between spheroidal wave functions and spherical Bessel functions

and associated Legendre polynomials [71]:

$$S_{mn}(c, \eta) R_{mn}^{(1)}(c, \xi) = \sum_{r=0}^{\infty} \delta_{(n-m)\%2, r\%2} i^{m-n+r} d_r^{mn}(c) j_{m+r}(kr) P_{m+r}^m(\cos \theta), \quad (3.23)$$

can be utilized for the derivation of the transcoding formula, where $\%_0$ is the modulo operator and $d_r^{mn}(c)$ are the expansion coefficients:

$$S_{mn}(c, \eta) = \sum_{r=0}^{\infty} \delta_{(n-m)\%2, r\%2} d_r^{mn}(c) P_{m+r}^m(\eta). \quad (3.24)$$

It can be shown that the analytical transcoding formula from spheroidal ambisonics coefficients $\{A_{mn}, B_{mn}\}$ to spherical ambisonics coefficients $\mathcal{A}_n^{m'}$ is given as the following:

$$\mathcal{A}_n^{m'} = \alpha_{m'} \sqrt{\frac{\pi(n' + |m'|)!}{(2n' + 1)(n' - |m'|)!}} \sum_{n=|m'|}^{\infty} \delta_{(n-n')\%2, 0} (-1)^{\frac{n'-n}{2}} d_{n'-|m'|}^{|m'|n}(c) (A_{|m'|n} - i \operatorname{sgn}(m') B_{|m'|n}) \quad (3.25)$$

where $\alpha_{m'} = (-1)^{m'}$ for negative m' and $\alpha_{m'} = 1 + \delta_{m', 0}$ otherwise. The derivation of this transcoding formula can be found in appendix A. The transcoded signal \mathcal{A}_n^m can be then stored / transmitted / processed with any existing signal processing pipeline for spherical ambisonics signals, e.g. rotation / filtering / decoding, and any technique or knowledge established for spherical ambisonics can be applied here.

3.4.2 Mixed-order transcoding

It can be noticed from (3.25) that the truncation number of the transcoded spherical ambisonics signal, which is hereafter referred to as N' , does not need to be the same

as the truncation number N of the spheroidal ambisonics signal. In fact, a truncated approximation of $\mathcal{A}_n^{m'}$ with $|m'| \leq N$ can be computed for any n' independently of N , as long as the truncation number R with respect to r in the table d_r^{mn} satisfies $n' - |m'| \leq R$. This consideration leads to the notion of *mixed-order transcoding*, which computes all spherical ambisonics coefficients $\mathcal{A}_n^{m'}$ with a truncation number $N' > N$, but only holding those coefficients that satisfy $|m'| \leq N$ and regarding $\mathcal{A}_n^{m'} = 0$ otherwise. The resulting transcoded ambisonics signal $\mathcal{A}_n^{m'}$ can be seen as a special form of the mixed-order ambisonics scheme introduced for conventional ambisonics [73]. Compression of the transcoded signal by discarding a subset of the $\mathcal{A}_n^{m'}$ coefficients similarly to conventional mixed-order ambisonics based on the direction dependence of the perceptual sensitivity [74] may be explored as well, which is out of the scope of the present work.

3.5 Experimental evaluation

Prolate spheroidal ambisonics encoding as well as its transcoding into spherical ambisonics was validated by numerical experiments. Encoding and transcoding of a plane wave with three different incident angles was performed with a sound-hard spherical microphone array as well as a sound-hard prolate spheroidal microphone array. The spherical array had a radius of 0.198 m. The prolate spheroidal microphone array had $r_{\text{short}} = 0.05$ m and $r_{\text{long}} = 1$ m. The arrays were designed to have the same surface area and both had 512 microphone capsules located on a 16-point grid of Gauss-Legendre quadrature nodes for θ and η and on a 32-point equispaced grid for φ . The long axis of

the prolate spheroidal array was set parallel to the x -axis. Fig. 3.1 shows the experimental procedure and the two microphone arrays used for the experiments. Spherical and spheroidal ambisonics encoding was performed using (3.10) and (3.22), respectively. Computation of the coefficient tables of spheroidal wave functions were performed using the software library *Spheroidal* [72], which is relying on arbitrary precision arithmetic using GNU MPFR [75] for accurate computation of the spheroidal wave functions. The truncation number was set to $N' = N = 12$ for the baseline spherical ambisonics and spheroidal ambisonics. The regularization parameter σ was set to zero for both spherical and spheroidal encoding, i.e. no regularization was applied. Transcoding from spheroidal ambisonics to spherical ambisonics was performed using (3.25), truncated at $n \leq N' = 12$, while $N' = 16$ was used for mixed-order transcoding. The estimated incident field for the encoded spherical ambisonics coefficients was reconstructed and compared to the ground truth incident field. The reconstruction of the estimated incident fields was performed using (3.4) truncated at $n \leq N'$. The signal-to-distortion ratio (SDR) of the reconstructed fields was computed for evaluation points in the x - y plane. The region with SDR higher than 30 dB was considered as the sweet-spot of accurate reconstruction.

Fig. 3.3 and Fig. 3.4 shows the results for incident waves with normalized wave vectors, expressed in the Cartesian coordinates, of $(1, 0, 0)$, $(0, 1, 0)$, and $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0)$, respectively. The frequency of the incident wave was set to 541.8 Hz. It can be observed that the width of the sweet-spot of precise reconstruction in spheroidal ambisonics is shorter in the shorter axis of the spheroid, but longer in the longer axis of the spheroid, compared to the width in the baseline spherical ambisonics case. With mixed-order transcoding, this prolongation is even more notable. This asymmetry of the sweet-spot

shape could be useful in some applications, in which a non-spherical sweet-spot is desired. An example application is sound field reproduction for multi-person home-theater systems in which the sweet-spot should cover multiple listeners sitting next to each other.

Note that the presented reconstruction is theoretical. If the reconstruction is performed via playback of the decoded spherical ambisonics signal using a limited number of loudspeakers, additional accuracy limitations apply. This is a problem of the decoding stage of spherical ambisonics and is out of the scope of the present work which is focused on recording, encoding, and transcoding. Any established technique for spherical ambisonics decoding and playback can be applied here.

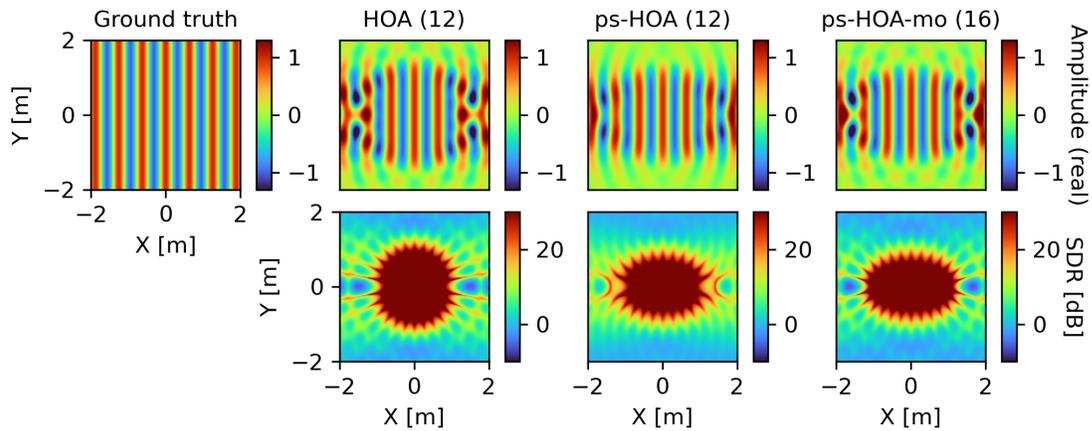


Figure 3.3: Results for an incident plane wave travelling along the long axis of the spheroidal array, which is set parallel to the x -axis. The first row from left to right: the real part of the sound pressure of the ground truth incident field, the field reconstructed from spherical ambisonics coefficients with $N' = 12$ (HOA), the field reconstructed from the prolate spheroidal ambisonics coefficients with $N = 12$ transcoded to spherical ambisonics coefficients with $N' = 12$ (ps-HOA), and the field reconstructed from the prolate spheroidal ambisonics coefficients with $N = 12$ transcoded to spherical ambisonics coefficients with $N' = 16$ using mixed-order transcoding (ps-HOA-mo). The second row presents the SDR of the reconstructed fields for HOA (left), ps-HOA (center), and ps-HOA-mo (right). The region with SDR higher than 30 dB was considered as the sweet-spot and is colored in red.

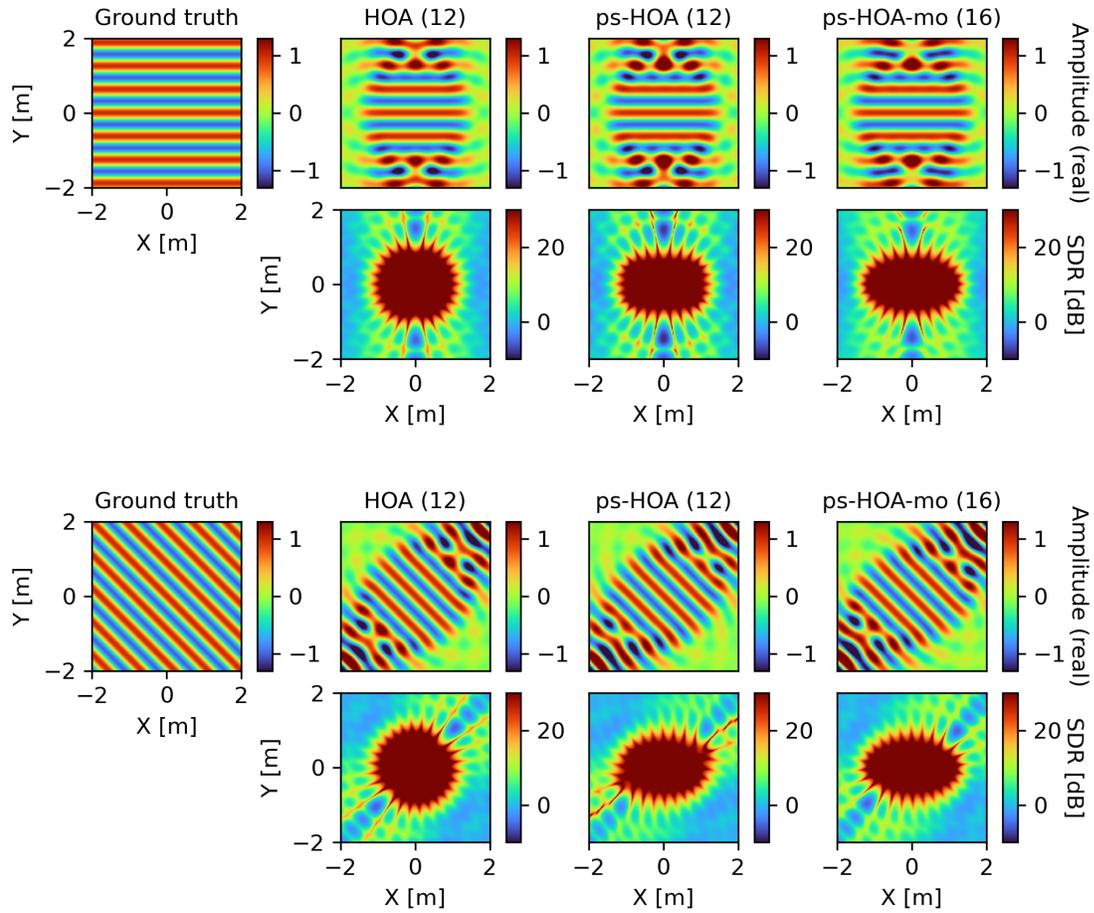


Figure 3.4: Results for an incident plane wave travelling along the short axis of the spheroidal array which is set parallel to the y -axis (upper two rows), and for an incident plane wave with a normalized wave vector $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0)$ (lower two rows). The definition of each of the subplot is identical to Fig. 3.3.

3.6 Conclusion

The framework of spheroidal ambisonics, a natural extension of ambisonics into spheroidal coordinates, was proposed. Spheroidal ambisonics enables analytical encoding of the spatial sound field into spheroidal ambisonics coefficients using spheroidal microphone arrays. An analytical transcoding formula from spheroidal ambisonics into conventional spherical ambisonics was derived, in order to ensure compatibility with the existing software ecosystem around spherical ambisonics. The concept of mixed-order transcoding which allows transcoding to spherical ambisonics of higher truncation numbers was introduced. The numerical experiments demonstrated that the sweet-spot of reconstruction in spheroidal ambisonics has an asymmetric shape which is prolonged towards the longer axis of the prolate spheroidal microphone array, realizing non-spherical sweet-spots in ambisonics reconstruction, which could be useful in some applications. The case of oblate spheroidal microphone arrays can be derived in a similar fashion and will be published elsewhere. A recently proposed microphone array for three-dimensional ambisonics recording, which uses a sound-hard circular disc as the scattering body [76], can be seen as a special case of an oblate spheroidal ambisonics microphone array. Another future research topic is the optimization of the microphone capsule configuration on the spheroid. In a practical setup, care must be taken for spatial aliasing [77] and a careful design of the microphone array configuration is important. While the subject of optimizing the microphone array configuration for spherical arrays has been studied extensively in the past [78,79], optimization of the array configuration in the case of spheroidal microphone arrays requires further research.

Chapter 4: Fast forest reverberation synthesis using single scattering cylinders

Simulating forest acoustics has important applications for rendering forest sound scenes in mixed and virtual reality, developing wildlife monitoring systems that use microphone arrays distributed in a forest, or as an artistic sound effect. Previously proposed methods for forest impulse response (IR) synthesis are limited to small or sparse forests because of their cubic asymptotic complexity with respect to the number of trees. Here we propose a simple and efficient parametric forest IR generation algorithm that relies on a multitude of single scattering cylinders to approximate scattering caused by tree trunks. The proposed method was compared to measured forest IRs in terms of the IR echo density, energy decay, reverberation time (T60), and clarity (C50). Experimental results indicate that the proposed algorithm generates forest reverb with acoustic characteristics similar to real forest IRs at a low computational cost.¹²

4.1 Introduction

Modeling forest acoustic impulse responses enables applications including designing or training forest wildlife monitoring systems based on microphone arrays and audio

¹This work was done while Shoken Kaneko was an intern at Microsoft Research Labs in Redmond, WA, USA, and was part of Microsoft's *AI for Earth* program.

²The research presented in this chapter has been published in [80].

signal processing, synthesizing artificial forest sound scenes to generate or augment training data for machine learning systems for sound event detection, creating reverberation sound effects for artistic expression and multimedia, or enabling plausible auditory rendering of virtual forests in mixed reality (MR) and immersive gaming. Studies related to forest acoustics date back to the 1940s [81, 82]. Prior work on forest acoustics or systems of multiple scatterers focused on computing the attenuation of a sound wave propagating through a forest, which can be represented by the bulk effective wavenumber [83–85]. Under the assumption that trees can be approximated by ideal cylinders, previous works considered multiple scattering by trees [83–85]. However, the bulk effective wavenumber is a result of statistical averaging and does not take into account geometric details including tree positions. Therefore, a bulk effective wavenumber alone is not sufficient to reproduce the echo patterns due to individual scattering events that depend on source, receiver, and tree positions.

Another line of research related to forest acoustics is recent work in the context of audio effects and outdoor reverberation synthesis. Spratt et al. devised a forest reverb synthesis algorithm based on a digital waveguide approach, referred to as the treeverb algorithm [6]. Stevens et al. introduced a more general scheme to model outdoor reverberation which is referred to as the waveguide web algorithm [8]. It too is based on waveguide techniques but not limited to forests. Although both of these algorithms consider multiple scattering, their asymptotic complexity is cubic with respect to the number of trees. This computational cost severely limits the size and density of the forest that can be modeled. Only results for small or sparse forests consisting of 25 to 30 trees have been reported using these algorithms [6, 8]. As shown later in this paper, experimental results indicate

that 30 or fewer trees, even if multiple scattering is considered, may not be sufficient to generate acoustically plausible forest IRs.

Here we propose a simple, efficient, and scalable algorithm for parametric forest impulse response synthesis. Source code and sample audio of the proposed method are available online.³ The proposed algorithm models the effect of tree scattering analytically via scattering from rigid cylinders, and approximates multiple scattering in a forest by single scattering from a multitude of virtual *single scattering cylinders* (SSCs). Since the SSCs do not interact with each other, this approximation results in an algorithm which can be easily parallelized and which has linear asymptotic complexity with respect to the number of SSCs, allowing the algorithm to generate forest IRs using hundreds of thousands of SSCs within a few seconds. Rather than aiming at a physically accurate simulation of a particular forest scene with given tree positions, the proposed algorithm produces acoustically plausible forest IRs. For many practical applications, including sound effects, audio rendering in MR, or data augmentation for training machine learning systems, physical accuracy may be of lower importance than subjective plausibility, computational cost, and parametric tunability of the reverberator. Experimental results indicate that the proposed algorithm produces synthetic forest IRs that exhibit acoustic and sonic qualities similar to real measured forest IRs [3], given a sufficiently large number of SSCs. This is achieved at a significantly lower computational cost compared to previous cubic order methods that try to model multiple scattering accurately. Using recorded real forest IRs as a baseline, the synthetic IRs produced by the proposed algorithm compared favorably to IRs synthesized by previously proposed algorithms in terms of energy decay

³https://github.com/microsoft/Forest_IR_synthesis

curves, reverberation time (T60), clarity (C50) [86], and a recently proposed echo density measure [87].

4.2 Proposed Approach

4.2.1 Algorithm overview

Assuming linearity and time-invariance, the acoustic path from a sound source to a receiver position is described by an impulse response (IR). A forest IR h_{forest} is modeled here as:

$$h_{\text{forest}} = h_{\text{direct}} + h_{\text{ground}} + h_{\text{tree}}, \quad (4.1)$$

where h_{direct} , h_{ground} , and h_{tree} are IRs describing the direct path, ground reflection, and tree scattering components, respectively. The direct-path component is given as:

$$h_{\text{direct}} = h_{\text{air}}(|\mathbf{r}_r - \mathbf{r}_s|) * h_{\text{dist}}^{(1)}(\mathbf{r}_r, \mathbf{r}_s), \quad (4.2)$$

where $h_{\text{air}}(x) = \mathcal{F}^{-1}(e^{-bx})$ denotes the air dissipation filter given the frequency-dependent sound attenuation factor b and the propagation distance x . \mathcal{F}^{-1} denotes the inverse discrete Fourier transform (DFT). Note that the IRs are functions of time t while the variable t is omitted for clarity. The source and receiver positions are denoted by \mathbf{r}_s and \mathbf{r}_r , respectively, and $*$ denotes the convolution operator. $h_{\text{dist}}^{(\alpha)}$ describes a generalized distance attenuation and delay filter defined as:

$$h_{\text{dist}}^{(\alpha)}(\mathbf{r}_1, \mathbf{r}_2) = \delta(t - |\mathbf{r}_1 - \mathbf{r}_2|/c) |\mathbf{r}_1 - \mathbf{r}_2|^{-\alpha}, \quad (4.3)$$

where δ is Dirac's delta function, c is the speed of sound, and α a distance decay parameter. This parameter allows the spatial spreading of the scattered wave to be interpolated between ideal cylindrical spreading ($\alpha = 0.5$) and spherical spreading ($\alpha = 1$). In the proposed model, the sound wave emitted from the source is modeled as a spherical wave with $\alpha = 1$, while the scattered waves traveling from the tree trunks to the receiver are modeled with a parameter $0.5 \leq \alpha_c \leq 1$. This is to account for the fact that real trees are not ideal cylinders with infinite length. The ground reflection component is modeled as:

$$h_{\text{ground}} = g(\mathbf{r}_r, \mathbf{r}_s) \left(h_{\text{air}}(|\mathbf{r}_r - \check{\mathbf{r}}_s|) * h_{\text{dist}}^{(1)}(\mathbf{r}_r, \check{\mathbf{r}}_s) \right), \quad (4.4)$$

where $\check{\mathbf{r}}_s$ denotes the source position reflected by the ground plane and g the ground reflection coefficient which is implemented as a frequency-independent reflection constant. If the frequency dependence of a forest ground reflection coefficient is known, that can be used here instead of a constant coefficient. The frequency spectrum of the reflection filter could be used to model different boundary conditions, e.g., the presence or absence of snow on the ground, which we leave for future studies. The tree scattering component is defined as:

$$h_{\text{tree}} = \sum_{m=1}^{N_c} h_{\text{air}}(r_{r,c_m,s}) * h_{\text{scat}}(\mathbf{r}_r, \mathbf{r}_{c_m}, \mathbf{r}_s) * h_{\text{dist}}^{(1)}(\mathbf{r}_{c_m}, \mathbf{r}_s), \quad (4.5)$$

where m denotes the SSC index, N_c the total number of SSCs, \mathbf{r}_{c_m} the position of the m th SSC, $r_{r,c_m,s} = |\mathbf{r}_r - \mathbf{r}_{c_m}| + |\mathbf{r}_{c_m} - \mathbf{r}_s|$ the total path length of the wave scattered by the m th SSC, and h_{scat} the scattering-angle dependent tree-scattering filter. The path lengths

$|\mathbf{r}_r - \mathbf{r}_{c_m}|$ and $|\mathbf{r}_{c_m} - \mathbf{r}_s|$ are evaluated in two dimensions, while the lengths of the direct path and the ground reflection path is evaluated in three-dimensional coordinates. The effect of tree scattering is modeled as the scattering of a plane wave by a rigid cylinder [2]:

$$h_{\text{scat}}(\mathbf{r}_r, \mathbf{r}_{c_m}, \mathbf{r}_s) = \beta h_{\text{dist}}^{(\alpha_c)}(\mathbf{r}_r, \mathbf{r}_{c_m}) * h_{\text{angle}}(\mathbf{r}_r, \mathbf{r}_{c_m}, \mathbf{r}_s), \quad (4.6)$$

with β denoting a parameter controlling the scattering amplitude. The scattering-angle dependent component h_{angle} is defined as:

$$h_{\text{angle}}(\mathbf{r}_r, \mathbf{r}_{c_m}, \mathbf{r}_s) = \mathcal{F}^{-1}(H_{\text{angle}}), \quad (4.7)$$

where H_{angle} denotes a frequency-domain filter modeling the frequency and scattering-angle dependent attenuation to an incoming wave scattered by a rigid cylinder:

$$H_{\text{angle}} = e^{i\frac{\pi}{4}} \sqrt{\frac{2}{\pi k}} \sum_{n=0}^{N_{\text{max}}} (2 - \delta_{n,0}) \sin(\gamma_n) \cos(n\varphi) e^{i\gamma_n}, \quad (4.8)$$

$$\gamma_n = \arctan \frac{J_{n-1}(ka) - J_{n+1}(ka)}{N_{n-1}(ka) - N_{n+1}(ka)}, \quad (4.9)$$

and

$$\gamma_0 = \arctan \frac{J_1(ka)}{N_1(ka)}, \quad (4.10)$$

with $\varphi = \arccos\left(\frac{(\mathbf{r}_r - \mathbf{r}_{c_m}) \cdot (\mathbf{r}_{c_m} - \mathbf{r}_s)}{|\mathbf{r}_r - \mathbf{r}_{c_m}| |\mathbf{r}_{c_m} - \mathbf{r}_s|}\right)$ denoting the scattering angle, k the wave number, a the tree radius, and N_{max} the truncation number of the series expansion. J_n and N_n are the Bessel functions of the first and second kind, respectively. To remove the tree-receiver distance dependency on the spectrum, the far-field approximation is used here [2].

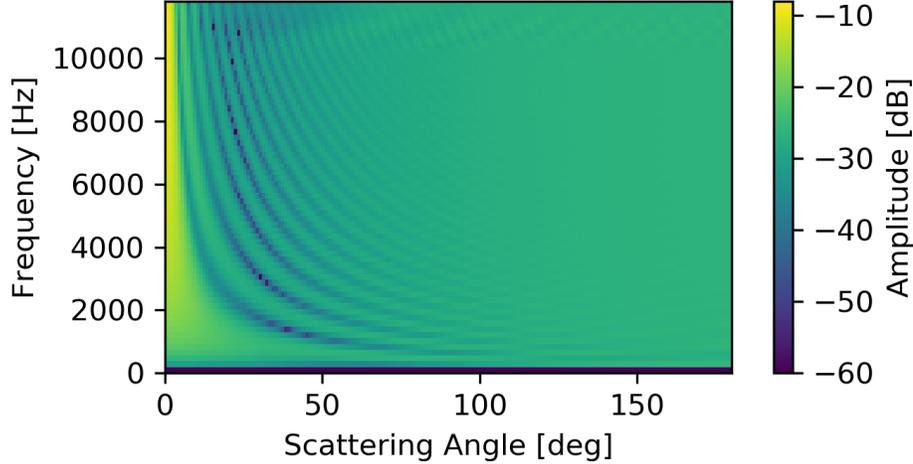


Figure 4.1: The amplitude spectrum of the angle-dependent tree scattering filter H_{angle} [2].

As can be observed from (4.1), (4.2), (4.4) and (4.5), the asymptotic complexity of the proposed forest IR generation algorithm is linear with respect to the number of SSCs. In addition, since all scattering paths can be evaluated in parallel, the proposed algorithm allows efficient synthesis of forest IRs with dense scattering patterns. In our experiments, we compute IRs for a forest of $1 \text{ km} \times 1 \text{ km}$ size with up to 500 000 SSCs.

4.2.2 Implementation and optimization details

The tree scattering filter h_{scat} is implemented as a 128-tap finite impulse response (FIR) filter. The scattering-angle dependent tree scattering amplitude spectra are shown in Fig. 4.1. The scattered wave has an amplitude drop of approximately 30 dB relative to the incident wave for scattering angles larger than about 20 degrees, implying that the majority of the multiple scattering paths will experience a quick decay of amplitude while traveling from the source to the receiver.

The air dissipation filter h_{air} is modeled based on the ISO9613-1 standard [88], and

was implemented as a 128-tap FIR filter. We found the incorporation of air dissipation to play an important role for obtaining a realistic timbre in the late reverberation. The ground reflection coefficient g was set to 0.8. For faster computation, all tree scattering and air dissipation filter coefficients are pre-computed for predefined discrete angles and distances. The processing cost is dominated by the computation of the tree scattering component h_{tree} . For better performance, the convolution is performed in the frequency-domain. Note that the frequency-domain filter banks for air dissipation and cylindrical scattering can be pre-computed. A Python implementation of the proposed algorithm was able to compute IRs of a forest with $1 \text{ km} \times 1 \text{ km}$ size and 100 000 SSCs at a sampling rate of 24 kHz in about 1.2 seconds on a laptop PC with an Intel Core i7-8665U CPU. In contrast, a Matlab implementation [22] of the algorithm by Spratt et al. [6] took about 31 minutes on the same machine to compute a forest IR considering multiple scattering for a forest with 50 trees. Note that this significant difference of computation cost, which is a consequence of the cubic complexity of the previously proposed method, prevents direct comparison of the methods using the same number of scatterers since the proposed method uses large numbers of scatterers by design.

4.3 Experimental Evaluation

4.3.1 Echo density of synthetic and real forest IRs

The echo density measure [87] of forest IRs for two existing algorithms [6,8], forest IRs synthesized by the synthesis algorithm proposed here, and real forest IRs measured in the Koli national park in Finland [3], was evaluated. All reference forest IRs were

downloaded from Openair [4, 5, 7, 89]. The IRs using the treeverb and waveguide web algorithms were computed for a forest of 25 trees and source-receiver distance D of 5.8 m. These IRs generated by previous algorithms, which are computed for a forest of a significantly smaller number of scatterers compared to the proposed method, should be considered only as references and not as a direct comparison since their cubic complexity only allows computation of small forests within a realistic amount of time. The real forest IRs were measured with $D = 9$ m. For the proposed algorithm, we set $\alpha_c = 0.7$, $a = 0.25$ m, source and receiver height to 1.5 m, $D = 10$ m, and $\beta = 2$. The SSC positions were randomly sampled using a uniform distribution within a square-shaped forest of size $1 \text{ km} \times 1 \text{ km}$. A constant tree radius was used here to allow reusing a single scattering coefficient table for all SSCs for the sake of computational efficiency, although this too can be randomized with an additional computational cost linear with respect to the number of SSCs. It was found by informal listening that randomness of the SSC positions is important to synthesize perceptually natural IRs. Configurations with regularity, e.g. regular Cartesian grids or two-dimensional quasi-random sequences with approximate regularity, tend to result in IRs with noticeable unnatural sound.

The waveforms and frequency responses of each forest IR are shown in Fig. 4.2. The comb-filter effect visible in the spectrum of the synthetic IRs obtained with the proposed method is a result of the interference of the direct path and the ground reflection. This effect can be controlled by tuning the amplitude, or, if frequency-dependence is considered, the spectrum, of the ground reflection coefficient g . While this comb-filtering is noticeable in the spectra, it did not result in unnatural sound in informal listening.

Fig. 4.3 shows the echo density profiles for all evaluated IRs. It is notable that

real forest IRs have nearly constant echo density profiles. We can see that this property is present in forest IRs synthesized with the proposed algorithm using a large number of SSCs ($N_c = 500k$), especially at low source–receiver distances D . However, this is not the case when using fewer SSCs ($N_c \leq 200k$) or for synthetic IRs obtained with previously proposed algorithms. While the reason for this near-constant echo density profile in real forest IRs is unclear, the results suggest that having a large number of SSCs may be necessary to produce comparable synthetic IRs.

Informal listening indicates that for $N_c \leq 200k$, the reduced echo density in the first 200 ms compared to real forest IRs is audible. It should be noted that $N_c = 500k$ corresponds to an SSC density about five to ten times higher than tree densities observed in typical forests [90]. We hypothesize that the high SSC density used here leads to denser echo profiles similar to real forest IRs by compensating for the fact that the SSCs are modeled as simple cylinders rather than complex geometric objects and that multiple scattering is not considered.

4.3.2 Energy decay curves of synthetic and real forest IRs

Fig. 4.4 shows the energy decay curves (EDCs) of the reference and synthetic forest IRs with parameters described in Section 4.3.1. As can be seen, the synthetic forest IRs obtained with the proposed method exhibit energy decay characteristics similar to real forest IRs. On the other hand, the EDCs of the previous algorithms exhibit characteristics which are far from real forest IRs. This difference was clearly audible in informal listening. To study the source–receiver distance (D) dependence of the EDCs,

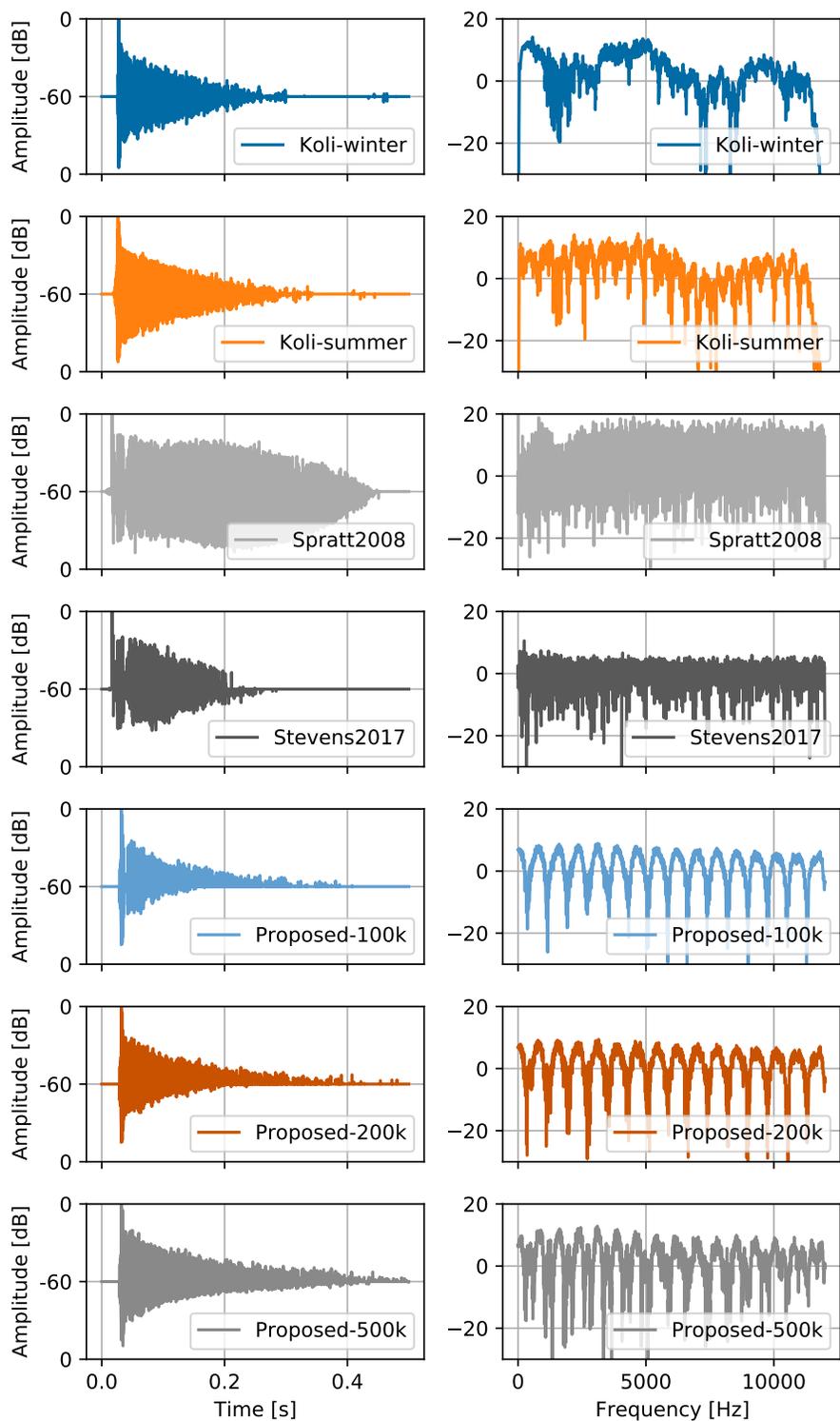


Figure 4.2: The waveforms (left) and frequency responses (right) of various forest IRs. From top to bottom: Koli national park in winter and summer [3–5], treeverb algorithm [6, 7], waveguide web algorithm [7, 8], proposed method with $N_c = 100k, 200k, 500k$.

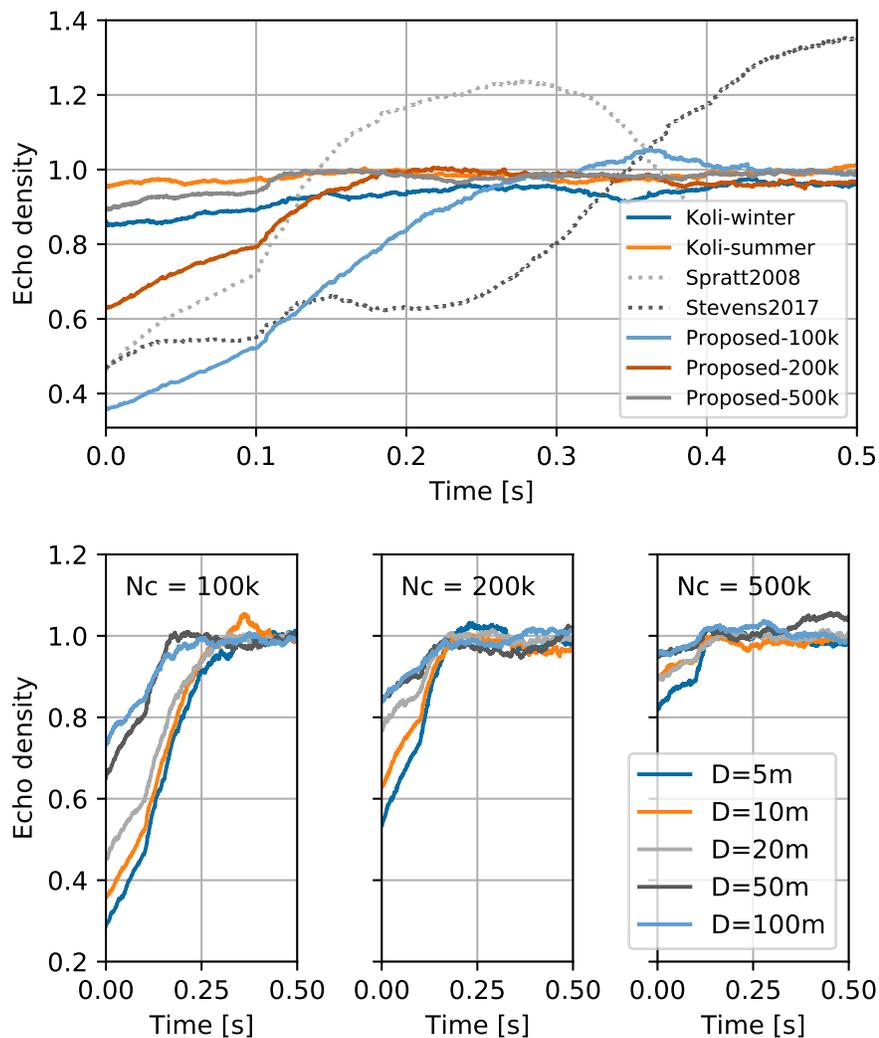


Figure 4.3: The echo density profiles of forest IRs. Comparison with reference forest IRs (top), and comparison among forest IRs generated by the proposed method for a square-shaped forest of $1 \text{ km} \times 1 \text{ km}$, with various D and N_c (bottom). Each subplot in the bottom shows results for $N_c = 100k$ (bottom left), $200k$ (bottom center), and $500k$ (bottom right), respectively.

we measured forest IRs in a real forest (Volunteer Park, Seattle, USA) at various distances D . As can be seen in Fig. 4.4 (bottom), the energy decays more slowly with increasing D , a property present in both the synthetic and real measured forest IRs.

4.3.3 Reverberation time and clarity

Fig. 4.5 shows the T60 and C50 of the synthetic forest IRs and the reference forest IRs introduced in Section 4.3.1 and Section 4.3.2. In the proposed method, T60 was able to be controlled over a wide range by varying N_c . Clear frequency dependence of T60 was only observed in real forest IRs and in synthesized forest IRs by the proposed algorithm. The synthetic forest IRs obtained with the proposed method exhibit a decrease of C50 as a function of D , which was also observed in measured forest IRs.

4.4 Summary and discussion

We proposed a simple and efficient parametric forest reverberator that approximates multiple scattering in real forests by scattering from a multitude of single scattering cylinders (SSCs). Experimental results show that the proposed algorithm synthesizes forest IRs with plausible acoustic and sonic properties at a low computational cost. The synthetic forest IRs were compared to real forest IRs, captured at two forests with varying seasons or source-receiver distances, in terms of the echo density, energy decay curves, reverberation time, and clarity. When modeling 0.5 million SSCs distributed over one square kilometer, the proposed algorithm produces IRs with an echo density profile similar to real forest IRs. The experimental results also indicate that the proposed algorithm

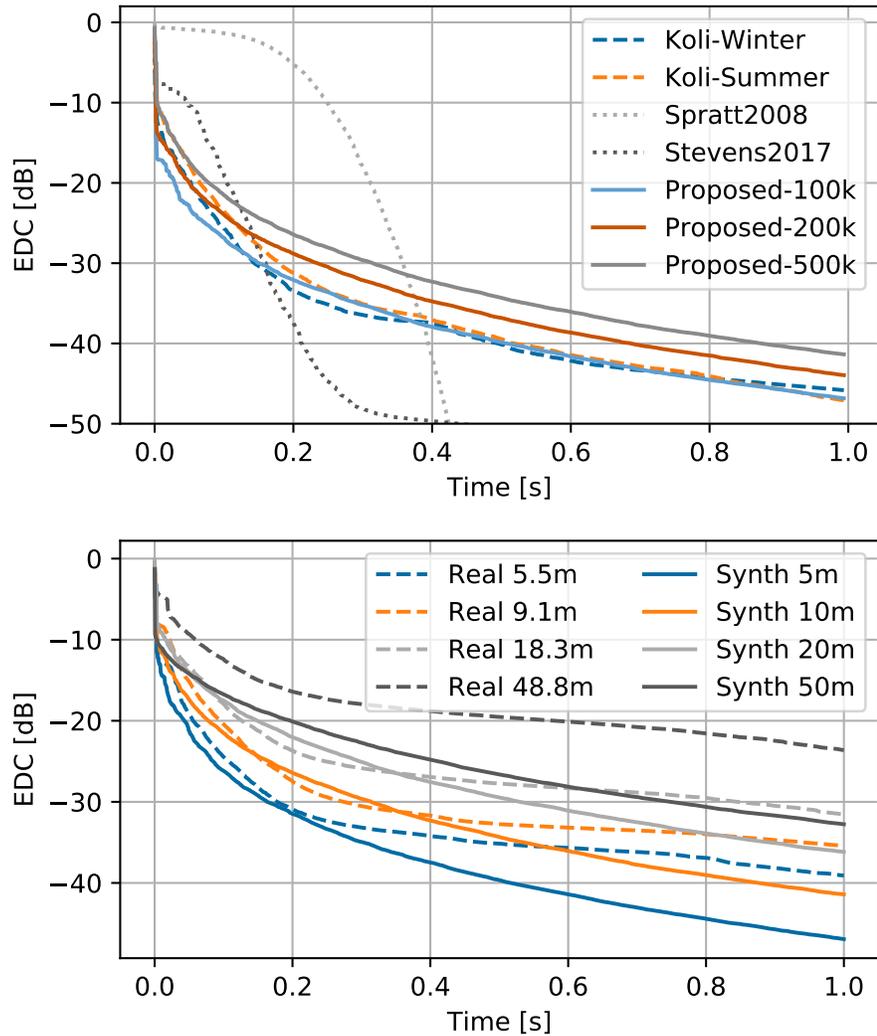


Figure 4.4: Top: EDCs for various short distance forest IRs. The set of forest IRs is the same as in Section 4.3.1. Bottom: EDCs for synthetic and real forest IRs with varying D . N_c was set to 500k for the proposed method. All EDCs are computed from the IRs after applying a band-pass filter with a passband of 500Hz-8kHz.

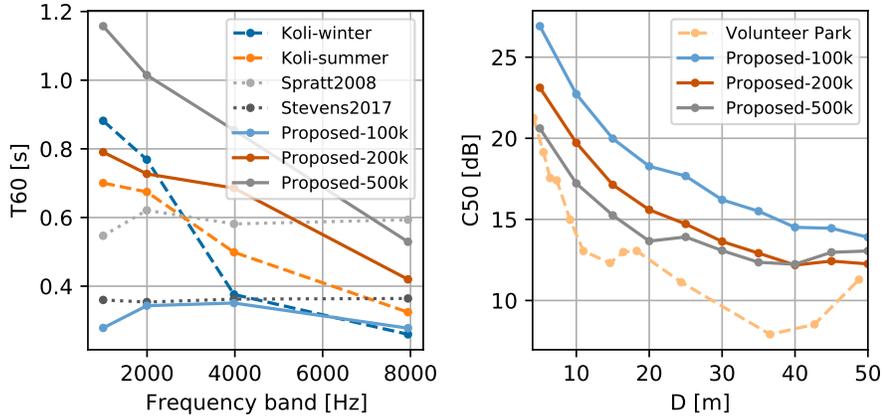


Figure 4.5: T60 as a function of octave-band frequency (left) and C50 as a function of D (right). The set of shown forest IRs is the same as in Section 4.3.1 and Section 4.3.2 for the left and right sub-plot, respectively. IRs from the Koli forest and previous algorithms are not included in the C50 plot since the available dataset only contained IRs for a fixed distance.

produces plausible energy decay profiles and characteristics similar to real forest IRs, such as the frequency-dependence of T60 and the distance-dependence of C50. The cubic arithmetic complexity of previously proposed algorithms may be too high to model forests with large tree numbers, resulting in synthetic IRs with audibly different echo density profiles compared to real forest IRs.

Note that the exact relation between the model parameter N_c determining the number of simulated SSCs and the number of actual trees in a real forest is unknown. While this relation could potentially be determined by additional measurements and simulation, this was outside the scope of the present work. This parameter can be subjectively tuned by a sound designer to match the desired acoustic characteristics. A practical recommended default value is between $N_c = 200$ k and $N_c = 500$ k for a forest of 1 km^2 , which results in a good balance between echo density and energy decay characteristics, as shown in the experiments. A limitation of both previous methods as well as the method proposed

here is that scattering from fine structures, e.g., leaves and branches, is not considered. It is possible that the high density of SSCs used in the experiments here is necessary to account for this limitation of the models. Efficient incorporation of multiple scattering and the consideration of scattering from fine structures, which may reduce the number N_c that delivers realistic acoustic profiles, is a subject of future research.

Chapter 5: Localizing birds in forests with distributed microphone arrays

Acoustic wildlife monitoring systems are important tools for capturing information about animal habitation in ecosystems. Previous work has demonstrated the effectiveness of audio-based bird localization techniques. However, few studies have investigated the performance and robustness of large-scale distributed systems. Here, we examine the performance of distributed microphone arrays for localizing birds in simulated forest scenes with added reverberation, ambient noise, and measurement errors. The experimental results may guide the design of practical large-scale wildlife monitoring systems and suggest promising directions for further improvements.¹²

5.1 Introduction

Autonomous monitoring methods are useful for studying animal habitats [92, 93]. Common examples include systems based on computer vision [94], RFID tags [95], or drones [96]. Audio-based approaches have the particular advantage of being able to detect and localize sound-producing animals, including birds and insects, in environments with visual obstacles, such as trees in dense forests, while theoretically completely covering

¹This work was done while Shoken Kaneko was an intern at Microsoft Research Labs in Redmond, WA, USA.

²The research presented in this chapter has been published in [91].

large areas, as the individual acoustic sensors are not restricted by a “field of view”.

Various systems have been proposed previously that use of microphone arrays to perform sound source localization (SSL) [97] for determining the density or location of birds. Dawson et al. used a four-channel distributed microphone array (DMA) with capsules located at the vertices of a square with 21 m sides. This array was manually moved over a grid of 75 locations in a 14 ha area to estimate the bird population density [98]. Mennil et al. used four stereo microphones placed at the vertices of a square with side length of 25 m or 50 m. The microphones were synchronized using the GPS clock to within a few milliseconds, achieving an average localization accuracy of 1.87 ± 0.13 m via time-difference of arrival (TDOA) estimation [99]. Stepanian et al. used an array consisting of six microphones placed at the vertices of an equilateral triangular prism with a height of 10 m and triangle sides of 20 m for three-dimensional localization [100]. Although these previous works have demonstrated that DMAs can be successfully applied for the localization of birds, the area covered by a single measurement using these monitoring systems has been relatively limited. Moreover, most prior works focus on real measurements with a limited set of variables. The effect of the DMA configuration as well as measurement errors, background noise, and reverberation on SSL in forests remains unclear, and is the focus of this work.

Fig. 5.1 illustrates a forest bird localization system. To evaluate the performance of such a system, SSL was performed on DMA signals derived from synthetic forest recordings. The recordings consisted of clean bird samples reverberated with a recently proposed forest impulse response (IR) model [80] and combined with ambient noise recordings. The system performance was evaluated with respect to various sources of

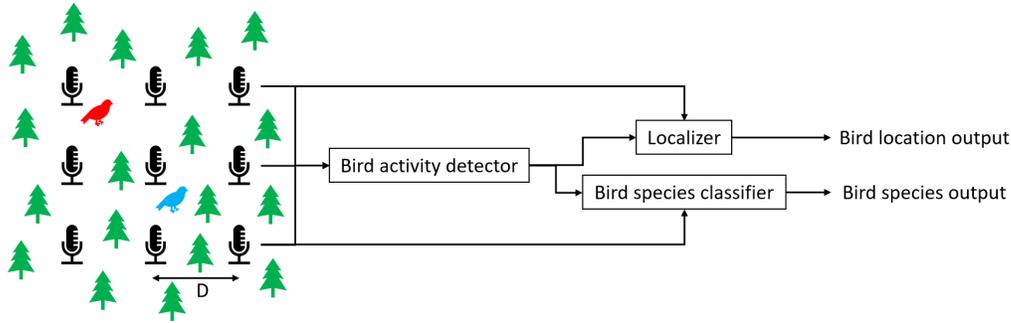


Figure 5.1: Diagram of a possible bird monitoring system. D denotes the microphone spacing, assuming a regular Cartesian grid.

noise and error, including the level of ambient noise, positional ambiguity of the recording devices, and temporal misalignment of the microphone signals due to synchronization issues encountered in distributed wireless systems. Furthermore, the effect of design parameters including the choice of microphone spacing and the SSL algorithm were studied. The experimental results may guide the design of practical large-scale wildlife monitoring systems and suggest promising directions for further improvements.

5.2 DMA recording simulation

To simulate the forest bird SSL, we synthesized DMA recordings using a dataset of clean bird sound clips and background noise clips, a bird sound directivity model, a forest reverberation model, and a measurement error model. The DMA recordings were synthesized by convolving simulated forest IRs that include a bird directivity model with clean bird sound clips extracted from bird sound databases. Various sources of error and noise were modeled, including ambient noise and microphone positioning and synchronization errors. Details of this simulation process are described in the following.

5.2.1 Bird sounds and ambient noise

Clean bird sounds and background noise clips were randomly sampled from *The Cornell Guide to Bird Sounds: Master Set for North America (version 2020)* [101] and the *Xeno-canto* [102] database. A general-purpose voice activity detector (VAD) [103] was used to separate segments of bird vocalizations and background noise in each recording. An approximate signal-to-noise ratio (SNR), defined as the ratio between the average level of the segments classified as signal and noise, was computed for each recording. Bird sound clips with an approximate SNR above 30 dB were treated as clean. The resulting collection contained 1036 clean bird sound clips and 2301 background noise clips from the Cornell Lab dataset, and 1189 clean bird clips and 43 265 background noise clips from the Xeno-canto database. In addition to recorded background noise clips, we used stationary noise with spectra shaped according to recorded clips, as well as simulated wind noise [104].

5.2.2 Acoustic simulation

To simulate multi-channel recordings in large scale forests, the forest IR simulation algorithm that models trees as single scattering cylinders (SSCs), as presented in chapter 4, was employed. A bird vocalization directivity model was incorporated by approximating the radiation from a bird as the radiation from a point source on a small spherical baffle. The far-field approximation of this radiated field is the following [2, 105]:

$$p_{\text{bird}}(r, k, \theta) = p_{\text{free}}(r)g(k, \theta) = p_{\text{free}}(r) \sum_{n=0}^{\infty} \frac{(-1)^{n-1}(2n+1)}{(ka)^2 h'_n(ka)} P_n(\cos \theta) \quad (5.1)$$

where r denotes the distance from the sphere to the observation point, k the wave number, θ the angle between the direction of the point source and the observation point seen from the center of the spherical baffle, $p_{\text{free}}(r)$ the pressure radiated from a point source in free space without the spherical baffle, $g(k, \theta)$ the angle-dependent filter due to the spherical baffle, a the radius of the spherical baffle, $h_n(x)$ the spherical Hankel functions of the first kind, and $P_n(x)$ the Legendre polynomials. The filter $g(k, \theta)$ is applied to the direct path, ground reflection path, and the tree scattering paths that compose the forest IR model in chapter 4.

5.2.3 Multi-channel sound scene generation with measurement error

A simulated forest sound scene is generated by randomly picking a sound source position, bird sound clip, and noise clips. The selected bird sound clip is convolved with the simulated multi-channel forest IR, where each channel represents the IR from the sound source to each microphone position including positional error. A random noise clip is added to each microphone channel at a specified A-weighted sound pressure level (SPL) [106]. The inter-channel synchronization error was modeled by applying a random temporal shift to each of the channels.

5.3 Sound source localization (SSL) and evaluation metrics

For the SSL, time-difference of arrival (TDOA) estimation methods were employed. TDOA estimation is based on the computation of the generalized cross correlation (GCC)

function [97],

$$\mathbf{R}_{1,2}(t) = \text{iFFT}(\Psi_{1,2}\mathbf{X}_1\mathbf{X}_2^H), \quad (5.2)$$

with $\Psi_{1,2}$ a spectral weighting function, \mathbf{X}_1 and \mathbf{X}_2 the Fourier transform of the microphone signals of the first and second microphone, respectively. Among many variants of the weighting function $\Psi_{1,2}$, the Roth weighting [107] and the Phase Transform (PHAT) weighting:

$$\Psi_{1,2}^{\text{Roth}} = \frac{1}{\mathbf{X}_1\mathbf{X}_1^H}, \quad \Psi_{1,2}^{\text{PHAT}} = \frac{1}{|\mathbf{X}_1\mathbf{X}_2^H|}. \quad (5.3)$$

was used in the experiments. A symmetric form, $\Psi_{1,2}^{\text{Roth(sym)}} = \Psi_{1,2}^{\text{Roth}} + \Psi_{2,1}^{\text{Roth}}$, was used for the Roth weighting. The information from all microphone pairs was integrated by constructing a spatial likelihood function [108]:

$$L(\mathbf{x}) = \sum_{k=1}^{M-1} \sum_{l=k+1}^M \mathbf{R}_{k,l}(t = \tau_{k,l}(\mathbf{x})), \quad (5.4)$$

where M is the number of microphones and $\tau_{k,l}(\mathbf{x})$ is the TDOA of the microphone pair (k, l) for a sound source located at position \mathbf{x} . Under the assumption that only a single source is active, the estimate of the source position \mathbf{x}_{est} can be obtained by:

$$\mathbf{x}_{\text{est}} = \underset{\mathbf{x}}{\text{argmax}}(L(\mathbf{x})). \quad (5.5)$$

The case of multiple simultaneously active sources is left for future work.

The localization success rate (LSR) was used as the evaluation metric, where the localization was considered successful when the localization error, defined as the Euclidean

distance between the true and estimated source position, was smaller than 5 m.

5.4 Problem formulation and the experimental setup

The subject is to localize the position of a bird in a large forest using a DMA of hundreds of nodes, each equipped with a single omnidirectional microphone. An approximately regular Cartesian grid is assumed as the DMA configuration. Localizing sound sources with TDOA methods with this large microphone numbers and target area becomes computationally expensive, therefore a two-step approach utilizing a machine-learning based bird activity detector (BAD) [109] is considered. The BAD classifies the presence / absence of a bird sound in a time window of fixed-length. In the first step, a reliable BAD detects bird activation on each microphone channel independently and an $N \times N$ sub-grid of microphones that is most likely to enclose the active bird is estimated. As the second step, the SSL algorithm is applied on the signal captured by the $N \times N$ sub-grid. In this work, the first step is assumed to be sufficiently robust and hence the focus is solely on the performance of the second step. A holistic simulation, integrating the first step as well, is left for future work. For simplicity all microphones are assumed to be sufficiently far from the boundary of the forest, and hence edge-effects are ignored. The size of the DMA sub-grid N was set to three. Two microphone grid types are considered. The first is an ideal regular Cartesian grid, as is depicted in Fig. 5.1, which is referred to as *regular*. The second is a regular Cartesian grid with random perturbation, which is referred to as *perturbed*. This grid is used to simulate microphone placements in real forests where the placement to precise grid points is impossible or impractical. The

microphone nodes are assumed to be synchronized using the GPS clock in each recording device.

5.4.1 Forest impulse responses

In the forest IR simulation, the bird positions were uniformly sampled from the area covered by the 3×3 DMA sub-grid. One hundred bird positions were considered for each experimental setup. The azimuth angle of the bird's head direction was uniformly sampled from $[0, 2\pi]$ independently for each bird position, and the elevation angle was fixed to zero. The microphone positions for the perturbed grid were uniformly sampled from discs with a radius of 1 m and centered at the regular grid points. Three different SSC densities, namely $\{0, 10 \text{ k/km}^2, 100 \text{ k/km}^2\}$, were considered for the forest IR simulation. SSCs that are included in a $1 \text{ km} \times 1 \text{ km}$ square concentric with the 3×3 sub-grid were used in the IR simulation. The positions of the SSCs were uniformly sampled within this square region. The height of the birds and microphones were both set to 1.5 m from the ground.

5.4.2 Sound scene synthesis

Recordings with a duration of 5 seconds each including a single bird vocalization were simulated. The following noise categories were considered: clean (no additive noise), shaped stationary noise, clips extracted from the field recordings, or simulated wind noise. The bird SPL at 1 m from the bird position was set to 80 dB_A as a reference, and the noise SPL at the microphone positions was set to 40, 50, or 60 dB_A . Here,

dB_A denotes the A-weighted SPL [106]. This resulted in an average SNR, averaged over microphone channels, ranging from about -60 dB to 10 dB, with a median of -18 dB. To simulate synchronization errors, random temporal shifts were applied to each microphone channel, sampled uniformly from the interval [-1ms, 1ms].

5.4.3 Localization parameters

The resolution of the spatial likelihood function was set to 0.2 m. In reality, it can be difficult to obtain the exact positions of the distributed microphones. To simulate this positional error, the microphone positions used in the localization algorithm were decoupled from the ones used at the recording stage and were chosen from $\{\textit{regular}, \textit{perturbed}, \textit{perturbed}'\}$, where *regular* is the regular Cartesian grid as depicted in Fig. 5.1, *perturbed* is the perturbed Cartesian grid which was also used for simulation of the forest DMA recordings, and *perturbed'* is a perturbed Cartesian grid generated with a different random seed resulting in another perturbed grid that is not identical with the one used for the recording simulation.

5.5 Results

5.5.1 The effect of each explanatory variable

Table 5.1 summarizes all experimental conditions. Five different combinations of microphone grid selections were used. *Reg-reg* and *Reg-pert* uses *regular* for the recording stage and *regular* or *perturbed* at localization stage, respectively. *Pert-reg*, *Pert-pert*, and *Pert-pert'* uses *perturbed* for the recording stage and *regular*, *perturbed*, or

perturbed’ at localization stage, respectively.

As noise types, the symbols “s”, “c” and “w” are used as abbreviations of shaped stationary noise, clips extracted from the field recordings, and simulated wind noise. The number following the noise type symbol is the noise level in dB_A . The set of explanatory

Table 5.1: List of variables and their values used in the experiments.

Variable	Values	#choices
Source position	100 positions within the 3×3 sub-grid	100
Microphone spacing	{25m, 50m, 100m, 200m}	4
Microphone grid type	{Reg-reg, Reg-pert, Pert-reg, Pert-pert, Pert-pert’}	5
Tree density	{0, 10k/ km^2 , 100k/ km^2 }	3
Directivity	{On, Off}	2
Noise type	{clean, s40, c40, w40, s50, c50, w50, s60, c60, w60}	10
Synchronization error	{On, Off}	2
GCC algorithm	{PHAT, Roth}	2
Dataset	{Cornell, Xeno-canto}	2

variables, as listed in Table 5.1, forms a 9-dimensional variable space. All combinations of variables in this space were evaluated in an exhaustive manner, resulting in 0.96 million localization experiments. The histograms of localization error is shown in Fig. 5.2 (left). The standard deviation with respect to the LSR along each of the dimension in the variable space is shown in Fig. 5.2 (right). It can be observed that the microphone spacing, noise category, the choice of the spectral weighting function in the GCC algorithm, and the noise level have the largest impact on the LSR. The inclusion of source directivity, synchronization error, and microphone position error reduced the average LSR by 9.3%, 1.8%, and 7.8%, respectively. GCC-PHAT was found to outperform GCC-Roth by 23.2% in average. These results implies that the SNR and the spectral characteristics of the signal and noise strongly affects the localization accuracy, and lead us to the experiments in the next section.

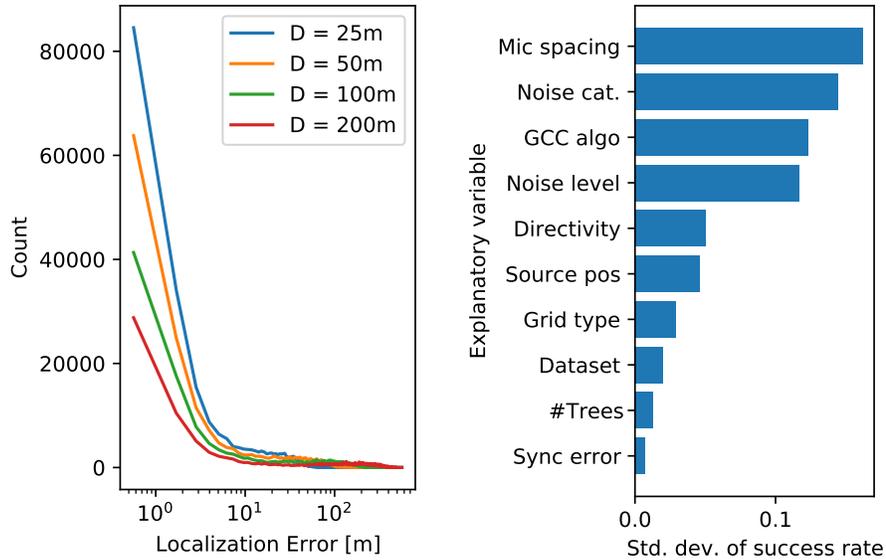


Figure 5.2: Left: histograms of localization errors. Right: the standard deviations with respect to the LSR along each dimension of the variable space. Note that the noise type dimension is split into noise category and noise level.

5.5.2 The effect of microphone spacing, noise type, and modifications to improve localization accuracy

Here, the variable space was reduced to two dimensions by selecting the realistic scenario (with source directivity, synchronization error, and the *pert-pert*' grid for microphone positions which contains positional error), by averaging over the dimensions that had relatively small variance (the number of trees, source positions, and dataset), and by choosing the GCC-PHAT algorithm for its superior performance. The results are shown in Fig. 5.3 (top left). It can be observed that the LSR decreases with increased microphone spacing and noise level.

Among the variables that strongly affect the localization performance, the microphones spacing and noise level is related to the SNR, whereas the noise category and choice

of the GCC weighting function is related to the spectral characteristics of the signal or noise. Therefore, one strategy to improve the system’s performance is to increase the SNR and to modify the spectrum of the recorded signal. As such attempts, two modifications were investigated, namely band-limiting (BL) in the GCC algorithm and applying noise suppression (NS) as preprocessing. For the BL, a frequency-domain rectangular window is multiplied with the GCC spectrum which nullifies the frequency components below 1 kHz and above 8 kHz before the inverse Fourier transform. This frequency range covers the vocalization frequency spectra of a wide variety of bird species. For the NS, the MMSE-STSA algorithm [110] was applied to the individual microphone channels independently before feeding it to the localizer. The resulting LSRs as function of microphone spacing and noise type with and without these modifications are shown in Fig. 5.3. It was observed that the NS without BL, BL without NS, and BL with NS improves the LSR by 7.9%, 9.0%, and 14.4% in average over the baseline. These improvements are important

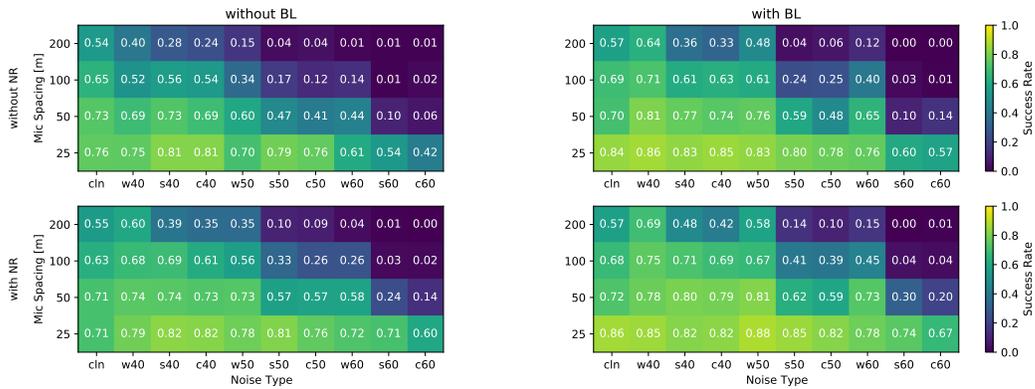


Figure 5.3: The LSRs for the cases with or without BL and NS. The results are for the realistic scenario (directional sources, with synchronization error, perturbed microphone grid with positional error, and GCC-PHAT algorithm), averaged over source positions and tree density, using the Xeno-canto dataset. The sub-plots represent the baseline (left-top), only BL (right-top), only NS (left-bottom), and both BL and NS (right-bottom).

for practical forest bird localization systems, since they allow the use of a sparser grid of

microphones implying a larger coverage area or a reduced cost of the system.

5.5.3 The effect of the SNR

Since the results in the previous sections implied the importance of SNR with respect to the localization performance, the relation between the SNR and the LSR was studied. The SNR values used here are the ground truth values computed from the signal and noise before mixing, averaged over the nine microphone channels. Different noise suppression rules were compared how they affect the localization performance. The suppression rules that were compared are: Wiener filtering [111], maximum-likelihood (ML) [112], spectral subtraction (SS) [113], minimum mean-square estimator short-term spectral amplitude (MMSE-STSA) [110], maximum a-posteriori spectral amplitude (MAP) [114], and Laplace-Gauss rule (LG) [115, 116]. The results are shown in Fig. 5.4. It can be observed that the application of BL and NS helps the localizer to improve the LSR by about 0.28 in the best case, or, to maintain the same LSR with larger noise level of about 9 dB, in challenging settings with SNRs of around -30 dB to -20 dB. The optimal suppression rule was found to be dependent on the average SNR.

5.6 Conclusion and future work

We studied the performance of a bird sound localization system using a distributed microphone array in the presence of reverberation, background noise, and measurement errors. To enable numerical experiments of bird localization in forests, a synthetic forest sound scene simulator for DMA recordings that utilizes an efficient forest IR simulation

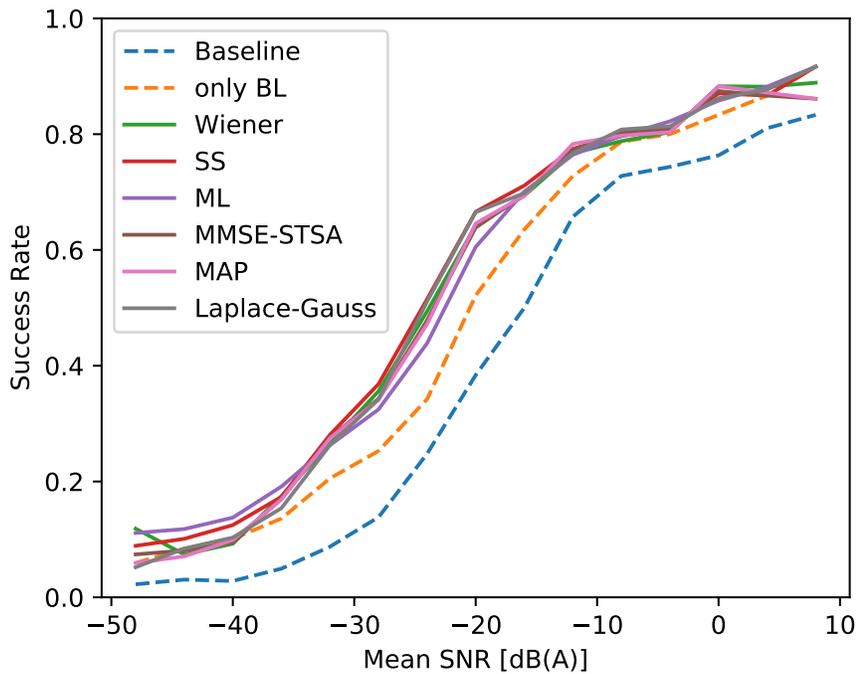


Figure 5.4: The LSR as function of average SNR using various NS algorithms. This results are for the realistic setup using directional sources, *pert-pert*' grid with synchronization error, and the GCC-PHAT algorithm. Band-limiting is applied except for the baseline case. Localization results from all noise types except for the clean case, all source positions, forests with non-zero trees, and the Xeno-canto dataset is used here.

algorithm was developed. Through the experiments, the variables that significantly affect localization performance were identified. It was also shown that spectral weighting in the GCC algorithm and NS significantly improves the localization accuracy. We hope that these results will help design and improve practical large-scale forest bird monitoring systems.

For future work, we suggest to develop and integrate a machine learning-based bird audio extractor/suppressor, which is trained on real world data of bird sound. This could be an effective way to improve the SNR. The forest sound scene synthesizer developed in this work can be useful as a data generation/augmentation engine when training such machine-learning models. It might be also beneficial to use GCC weighting schemes that take into account the spectral characteristics of the signal and noise [117]. Particularly, we suggest to utilize statistical knowledge about bird species-dependent vocalization frequency spectra. By using the estimated species label provided by the bird species classifier (as shown in Fig. 5.1), it is possible to apply a bird species-dependent spectral weighting in the GCC algorithm, which could significantly improve the localization accuracy in low-SNR situations. Additionally, multi-channel NS schemes [118,119] could potentially improve the NS preprocessing and overall localization performance. Since such multi-channel NS may require the estimation of the TDOA, it is implied that the problem of multi-channel NS and SSL should be solved jointly [120–122]. Using dense microphone arrays as the individual nodes in the sparsely distributed grid [123] is another promising approach to improve the localization performance. Experimental studies on the performance of large-scale forest bird localization systems deployed in real forests are also left for future research.

Chapter 6: A general-purpose individual-agnostic full-sphere binaural localizer

Sound source localization from binaural signals has important applications ranging from machine listening to psychoacoustics, yet challenges including generalization and robustness under noisy and reverberant conditions remain. Here we propose a binaural localizer (BL) framework that produces full-sphere spatial activity maps. The framework enables individual-agnostic training of a convolutional neural network using head-related impulse response (HRIR) sets with arbitrary measurement grids and is shown to perform well on unseen HRIRs and binaural recordings. Unlike BLs trained with the HRIRs of a specific known subject or dummy head, the proposed individual-agnostic BL is intended to perform robustly without any a priori knowledge about the process creating the binaural signals. Localization tests using binaural renderings and recordings show that the proposed BL is robust to noisy and reverberant conditions and compares favorably to individual-specific BLs. Furthermore, preliminary results indicate that the proposed BL is applicable to multiple simultaneous and moving sources.¹²

¹This work was done while Shoken Kaneko was an intern at Microsoft Research Labs in Redmond, WA, USA.

²The research presented in this chapter has been published in [124].

6.1 Introduction

Humans localize sound by learning to map features embedded into the two ear input signals to a source's location. A machine able to localize sound from binaural audio has important applications including robotics, sound scene analysis, as well as psychoacoustics and binaural audio evaluation. However, challenges remain for existing binaural localizers (BLs) including generalization across individuals and robustness under realistic noisy and reverberant conditions.

Machine learning-based BLs date back to the early 90s [125–127]. Neti et al. created a neural network (NN) that learns the mapping from the HRIRs to the sound source direction using HRIRs of a cat [127]. Jin et al. reported reasonable agreement in terms of the localization characteristics between a human subject and a NN-based BL trained with the same subject's HRIR set using band noise sources [128]. Jiang et al. developed a deep NN (DNN) for binaural sound source separation of speech based on time-frequency bin-wise classification [129]. Ma et al. used a DNN for binaural localization of multiple sources in the horizontal plane, which also incorporated active head movements [130]. Thuillier et al. studied saliency maps of a convolutional NN (CNN)-based BL for median plane localization trained with an individual-agnostic setup [131]. Wu et al. developed a random forest-based BL trained on a single subject's HRIR set for localization of both azimuthal and elevation angle and reported mean angular errors of about 10 degrees on binaural signals recorded in a laboratory. Wang et al. studied DNN-based BLs in the mismatched HRIR condition where the HRIR set used at test time is different from the set used for training, and proposed a method for clustering

HRIR sets based on the similarity of the localization performance of BLs [132]. Yang et al. developed a multi-task CNN-based BL for lateral and polar angle classification [133]. Franci et al. studied a massive CNN-based BL trained on dummy head HRIRs and reported various similarities between human spatial hearing and model behaviour, e.g., the emergence of the precedence effect, sensitivity to spatial cues, and the challenges arising in localization of concurrent sources [134]. Their results also indicated that an individual-specific BL may have limited generalization ability across individuals in elevation localization.

Most existing BLs are either individual-specific or limited to dataset-dependent directions on a subset of the sphere. While an individual-specific BL may be useful, training such a BL requires individual HRIRs, which may not be available in practice. This motivates the development of a general-purpose individual-agnostic BL which can provide localization estimates of binaurally spatialized audio without a priori knowledge of the process or HRIRs creating the binaural audio. This would allow localizing sounds in a broad variety of binaural media (games, music, movies, video conferencing calls), based on the sole assumption that the spatialized sound is intended for binaural playback to a human listener. Another potential use case for a general-purpose BL is to provide an estimate for how an *average* listener might localize a certain binaural rendering. Here, we propose a general-purpose individual-agnostic BL using a CNN that outputs a spatial activation map covering the entire sphere for each processed audio frame, allowing extension to multiple/moving sources. The novelty and contributions of the present work can be summarized as: a) design of model output format (HRIR dataset-independent, full-sphere coverage), b) design of training scheme (individual-agnostic, augmentation by simulated noise/reverb, use of soft targets rather than hard binary targets), and c) a general-purpose

BL with robust localization performance on unseen binaural recordings under unknown HRIRs, noise, and reverberation, which opens new applications (localization from arbitrary binaural audio with unknown generation process).

6.2 Proposed method

The proposed BL takes a pair of binaural audio signals as input and produces directional activation maps (see Fig. 6.2) for each frame. A short-time Fourier transform (STFT) converts both channels into a $2 \times F \times T$ log-magnitude spectrogram as well as a $1 \times F \times T$ spectrogram of interaural phase differences (IPDs), where T and F are the number of time frames and frequency bins, respectively. The log-magnitude spectrogram and the IPD spectrogram are fed into a “two-legged” CNN, inspired by [133]. The network architecture and hyperparameters are shown in Fig. 6.1 and Table 6.1, respectively. Each leg of the CNN has six convolution blocks, where each block consists of convolution, batch normalization, max-pooling, and the nonlinearity. Max-pooling was applied with a pooling factor of two in the frequency axis and no pooling was applied in the time axis. The outputs of the two CNNs are concatenated and fed into a fully-connected (FC) network with one hidden layer and an output layer. The leaky-ReLU activation function was used for all layers except for the output layer which has the sigmoid activation function. The output layer forms the directional activation map, where each output neuron is assigned to a specific direction on the sphere. The output of the network is a $D \times T$ matrix where D is the number of bins in the output direction map. Here, the 2048-point spherical Fibonacci grid [135] was used for the output direction

map. This novel dataset-independent output format was motivated by the fact that most available public HRIR datasets do not have a common angular grid which makes the mixing of datasets difficult and hence restricts the amount of training data. The proposed output format allows a mixture of datasets with arbitrary angular grids, scaling up the amount of available training data. This output format is also source-number independent and facilitates the application of the model to multiple source localization. The network is trained using the AdaMod optimizer [136] with the multivariate binary cross-entropy (BCE) loss between the network output and the ground truth target as the minimization objective. Rather than treating the problem as hard binary classification, dynamic soft targets based on the von Mises-Fisher probability density function are used as the target for a source with ground truth direction $\boldsymbol{\mu}$:

$$\boldsymbol{f}(\mathbf{x}; \boldsymbol{\mu}, \kappa) = C(\kappa) \exp(\kappa \mathbf{x} \boldsymbol{\mu}) \in \mathbb{R}^D, \quad (6.1)$$

where \mathbf{x} is the $D \times 3$ matrix of output directions, $C(\kappa)$ is a coefficient normalizing the L^∞ norm of the vector $\boldsymbol{f}(\mathbf{x}; \boldsymbol{\mu}, \kappa)$, and \exp is the element-wise exponential function. The concentration parameter κ is initialized with 2, and is doubled every 100 iterations until reaching 512, to gradually sharpen the soft targets. The use of soft targets was inspired by the success of “fuzzy” targets in musical onset detection [137].

6.3 Experimental evaluation

The proposed method was evaluated in terms of the localization error angle of a single static source, with preliminary results for multiple or moving sources. For the

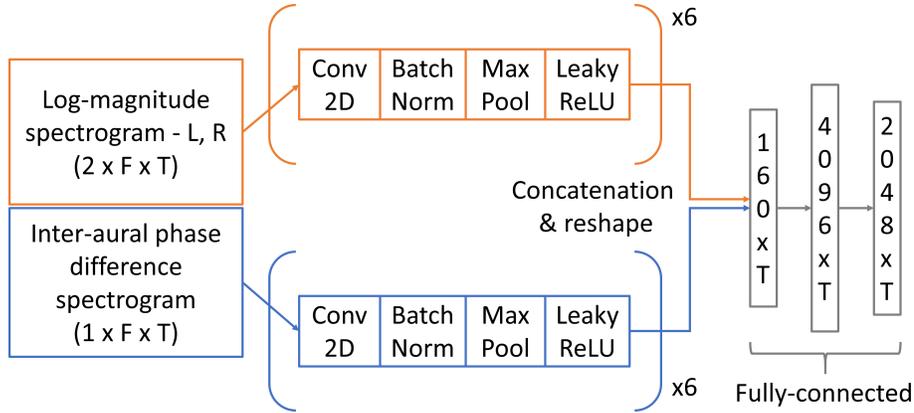


Figure 6.1: The network architecture of the proposed model.

Table 6.1: Model hyperparameters

Sampling rate	32kHz
FFT size	512 taps
STFT hop size	120 taps
STFT window	Kaiser ($\beta = 4$)
Kernel sizes (frequency-axis)	(5, 5, 5, 3, 3, 3)
Kernel sizes (time-axis)	(5, 5, 5, 3, 3, 3)
# Feature maps in each conv. layer	(5, 10, 20, 40, 80, 80)
# Hidden units in FC layer	4096
# Output units	2048

single static case, three different test tasks were used. The first two tasks are localization given binaural audio synthesized by convolving speech with public BRIR datasets, namely a dataset of horizontal plane BRIRs for rooms with various reverberation times (*IoSR Rooms*) [138], and a dataset of 22.2 channel BRIRs (*IoSR 22.2ch*) [139]. The third test task is localization of static sound sources present in binaural recordings; LOCATA challenge corpus - evaluation set - task 1 (*LOCATA Task1*) [140].

6.3.1 Model training and validation

The input binaural signal for training was generated by convolving HRIRs from public datasets with clean speech signals. The CIPIC [141], ARI [142], RIEC [143], ITA [144], Viking [145], and CHEDAR [146] HRIR datasets were used for training, resulting in 1680 subjects in total. The HRIRs captured using the KEMAR manikin were excluded from the CIPIC and Viking datasets. From the CHEDAR dataset, the HRIRs with source distances of 0.5 m, 1 m, and 2 m for the first 1240 shapes were used for training. The HRIRs associated with the first thirteen shapes and source distance of 2 m from the CHEDAR dataset were used to evaluate the model on seen HRIR data (*Seen*). Acoustically measured HRIRs from the last thirteen subjects excluding dummy heads from the HUTUBS dataset [147] were used for validating the model on unseen HRIR data (*Unseen*). The GRID corpus [148] was used as the speech source. The set of speakers was split into 90%, 5%, and 5% and the first two subsets were used for training and validation, respectively. It is known that noise and room reverberation degrades the performance of a BL [130, 133, 134, 149]. To emulate challenging real-world conditions, stereo white noise and BRIRs simulated by the image source method [21] were added to the input binaural signal. During training, the signal-to-noise ratio (SNR) was uniformly sampled from L_{Noise}^{\max} dB_A to 40 dB_A and the noise was normalized accordingly. L_{Noise}^{\max} is a hyperparameter corresponding to the maximum additive noise level which was chosen from {20, 25, 30, 40, No noise}. An ensemble of the proposed models was formed from five models trained with these five different maximum noise level conditions. The BRIRs were precomputed using HRIRs from the train and validation set, excluding the

direct path to allow addition with the direct path signal which is prepared separately on-the-fly during training. The room simulation parameters were randomly sampled from predefined ranges of room dimensions and impedance of the walls, as shown in Table 6.2. Four different sets of synthesized binaural signals were used to monitor the performance of the model during training, which includes binaural signals synthesized using a subset of the seen training set HRIRs and the unseen validation set HRIRs, with or without additive stereo noise and binaural reverberation. These four datasets are denoted in the following as *Seen clean* (same HRIRs in training and validation), *Seen N+R* (additive noise and reverberation), *Unseen clean*, and *Unseen N+R*, respectively. 64 directions based on the spherical Fibonacci grid were used to sample HRIR directions for the validation runs. During validation, the SNR was set to 30 dB_A for the *N+R* cases. The direction with maximum activation in the output direction map is considered as the estimated sound source direction in the static single source case. The direction maps were accumulated by taking the maximum over all time frames before making the decision about the sound source direction. Analyses based on lateral and polar angle error have revealed that the error is dominated by the polar angle error while the lateral angle localization can be highly precise. The polar angle alone, however, is problematic as an evaluation metric since it has singularities at the left and right pole. Hence, following prior works [128,149], the models were evaluated by the total mean angular error (MAE) where the angular error is the angle between the estimated and ground truth source direction. The models were trained for 200 k iterations and the model with the best total MAE averaged over *Unseen clean* and *Unseen N+R* was chosen as the model to evaluate the performance on the separate test sets.

Table 6.2: BRIR simulation parameters and their ranges from which the parameters are uniformly sampled.

Room width, depth	(3m, 12m)
Room height	(3m, 10m)
Source/receiver position	At least 1 m from the walls
Source/receiver height	(1m, 2m)
Reflection order	20 (fixed)
Wall impedance ratio	(5, 19)
Receiver’s sight direction	Yaw angle in $(0, 2\pi)$

6.3.2 Single static source localization

Visualizations of the output direction maps of the proposed model on the LOCATA-task1 test clips are shown in Fig. 6.2. It is interesting to observe the emergence of cones of confusion, i.e., rings of constant lateral angle on the spherical map, even though the proposed model does not explicitly estimate lateral or polar angles. Similarly, human listeners tend to have worse polar angle localization than lateral angle localization, especially when listening to non-individual HRIRs [150]. The localization results are summarized in Table 6.3.

The proposed model or its training condition was modified by not adding white noise and/or BRIR at training time, by band-limiting the input speech via applying a low-pass filter with cutoff frequency of 8 kHz, by using white noise as the source signal, or by using hard targets instead of the soft targets for supervision. Modifications to the sources, i.e., band-limiting or using white noise, were applied to both the *Seen* and *Unseen* evaluations. In the case of hard targets, the direction bin closest to the ground truth direction was set to one and all other bins were set to zero in the target signal. It can be observed from the results that the addition of BRIRs, the use of speech as the source

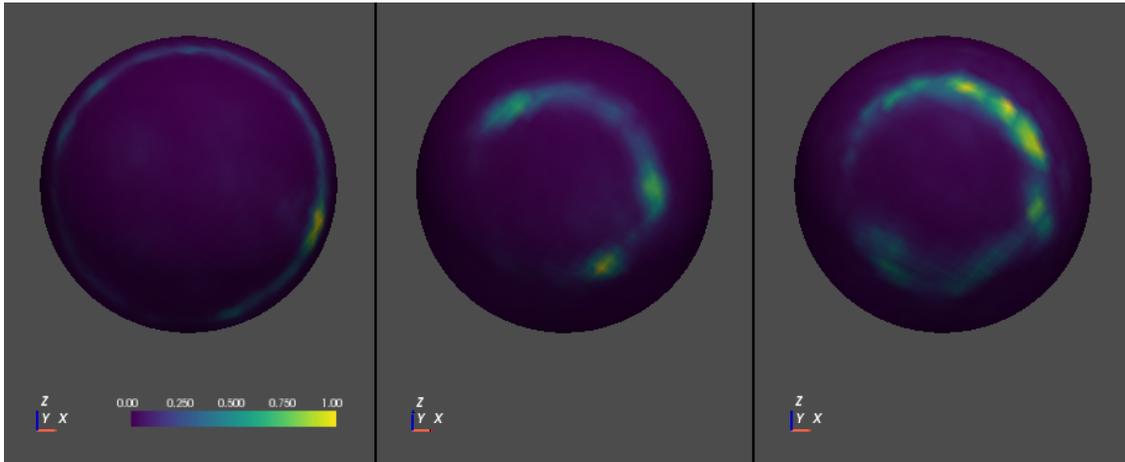


Figure 6.2: Typical output activation maps taken from LOCATA-task1 test clips. The right hemisphere of the map is shown with the L^∞ norm of the post-sigmoid output vector normalized after taking the maximum over time frames.

Table 6.3: Mean angular error (degrees) for various models.

Model	Seen Clean	Seen N+R	Unseen Clean	Unseen N+R	IoSR Rooms	IoSR 22.2ch	LOCATA Task1	Test Avg.	Test Max.
Proposed (BCE + ST ³)	37.8	48.0	43.6	50.0	16.4	28.6	27.9	24.3	28.6
w/o Noise	31.3	47.6	43.5	50.5	30.7	27.7	24.9	27.8	30.7
w/o Reverb	6.3	53.6	30.5	56.5	35.1	33.2	71.1	46.4	71.1
w/o Noise & Reverb	5.4	69.8	30.4	69.7	38.7	42.9	60.3	47.3	60.3
Band-limited speech ⁴	35.9	47.1	39.3	47.0	31.7	30.0	26.4	29.4	31.7
White noise source ⁵	20.1	38.5	31.6	49.9	49.7	57.1	65.7	57.5	65.7
BCE + hard targets	46.2	55.0	48.9	51.7	24.9	30.3	49.9	35.0	49.9
CCE loss	49.8	52.6	46.0	52.5	36.9	34.7	38.3	36.6	38.3
Multi-task loss [133]	46.4	51.0	43.7	50.3	46.4	35.8	41.3	41.2	46.4
Proposed ensemble	34.3	45.0	40.9	47.9	14.4	25.0	26.0	21.8	26.0
CHEDAR (IS-average)	27.2	39.2	58.3	59.8	28.6	49.6	23.4	33.9	49.6
CHEDAR (IS-ensemble)	-	-	-	-	23.6	47.2	17.4	29.4	47.2
HUTUBS-sim (IS-avg.)	1.3	7.1	40.7	40.3	43.4	34.5	64.7	47.5	64.7
HUTUBS-sim (IS-ens.)	-	-	-	-	39.2	26.5	58.3	41.4	58.3
HUTUBS-meas (IS-avg.)	0.3	2.6	38.4	35.5	56.8	43.2	70.9	57.0	70.9
HUTUBS-meas (IS-ens.)	-	-	-	-	53.7	35.5	66.3	51.8	66.3

signal, and the use of soft targets substantially improve localization performance.

The proposed model was compared with two reference models which use either the categorical cross entropy (CCE) or a multi-task CCE loss [133] which is the mean of two CCE losses of two classifiers estimating the lateral and polar angle separately. The CCE loss has been commonly used in previous DNN-based BLs [130–132, 134]. The CCE loss model here uses the same 2048 direction bins as the direction classes. The multi-task model uses 37 lateral and 72 polar angle classes with 5 degrees step, resulting in 2522 different directions. Except for the output layer, the network architecture and training conditions of the reference models are identical to the proposed model. From Table 6.3 it can be observed that the proposed model outperforms these reference models.

In order to study the localization performance of individual-specific (IS) BLs, 13 subjects each from the CHEDAR, simulated HUTUBS, and measured HUTUBS dataset were used to train IS models. An ensemble of IS models was also formed for each HRIR dataset. It was observed that some IS models from the CHEDAR dataset outperform the proposed model in LOCATA-task1. However, as can be seen in the MAE on the other test sets, the maximum MAE for the IS models tend to be larger, indicating that the individual-agnostic model operates more robustly. The results also imply that for a IS model to be effective, an individual which delivers high expected localization performance needs to be empirically sought.

6.3.3 Multiple or moving source localization

To test multiple source localization, binaural recordings from the evaluation set of LOCATA-task2 were used. Example network output activation maps, projected onto the lateral-polar grid, for input audio containing up to four sound sources are shown in Fig. 6.3. The number of feature maps in the l -th convolution layer was set to 2^{l+2} in this experiment and the model was trained for 1 million weight updates using training data with two sound sources. It can be observed that the model output has multiple vertical lines of high activation in the lateral-polar projection which correspond to the rings of confusion observed in the spherical output map. The proposed model can be

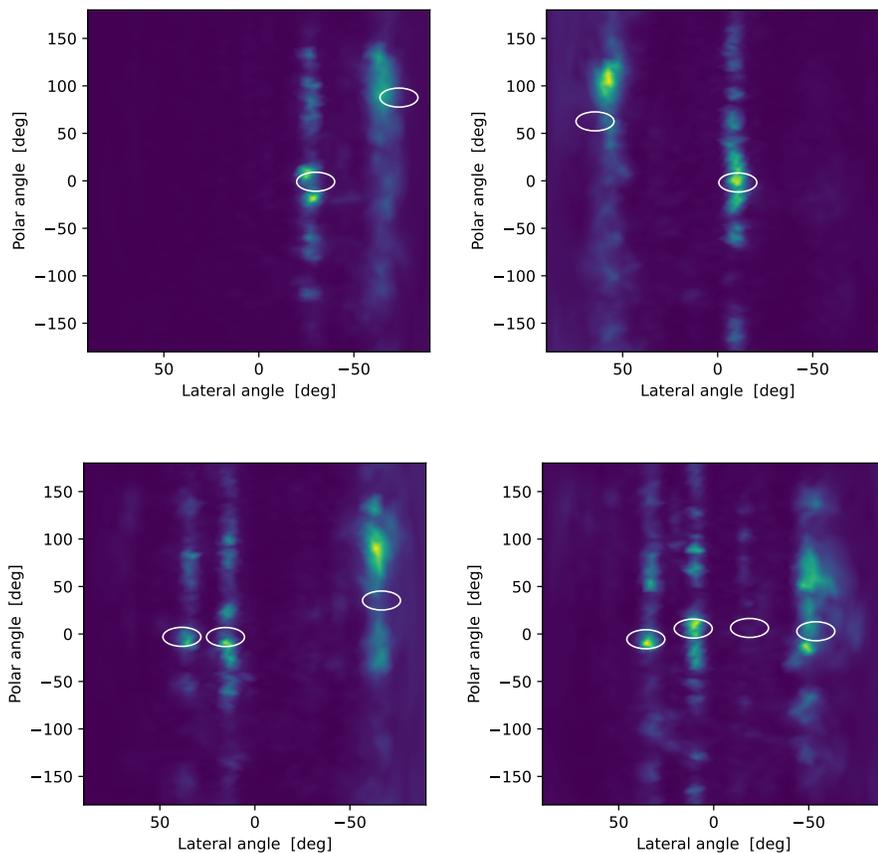


Figure 6.3: Output maps shown in the lateral-polar grid, for LOCATA-task2 test clips with two to four sources. The ovals denote the true source directions.

used in a sliding window fashion to produce a sequence of directional activation maps. To demonstrate moving source localization, the binaural recordings from the LOCATA-task3 were processed with the proposed general-purpose BL and the sequences of output activation maps were recorded. Animated visualizations of the model output maps are available online⁶.

6.3.4 Response to panned stereo audio

It is known that the interaural time difference (ITD) and the interaural level difference (ILD) are important acoustic cues that affect binaural localization. Amplitude- or delay panning are established techniques to create sound images in stereophonic sound by artificially modifying the ILD or the ITD. Stereo signals created by panning can be seen as a simple approximation of binaural audio. We use panned stereo to test whether the model generalizes to binaural-like signals not seen during training. Fig. 6.4 shows the resulting lateral and polar angle estimates made by the proposed model given panned stereo input. The results for delay panning suggest that the model has learned to associate the ITD with the lateral angle. Amplitude panning also affected the estimated lateral angle, however it did not exhibit a smooth curve as in the case of delay panning, but resulted in a noisy step-like profile.

⁶<https://github.com/microsoft/DIABLo-demos>

¹Band-limited speech was used for training, seen and unseen evaluation

²White noise source was used for training, seen and unseen evaluation

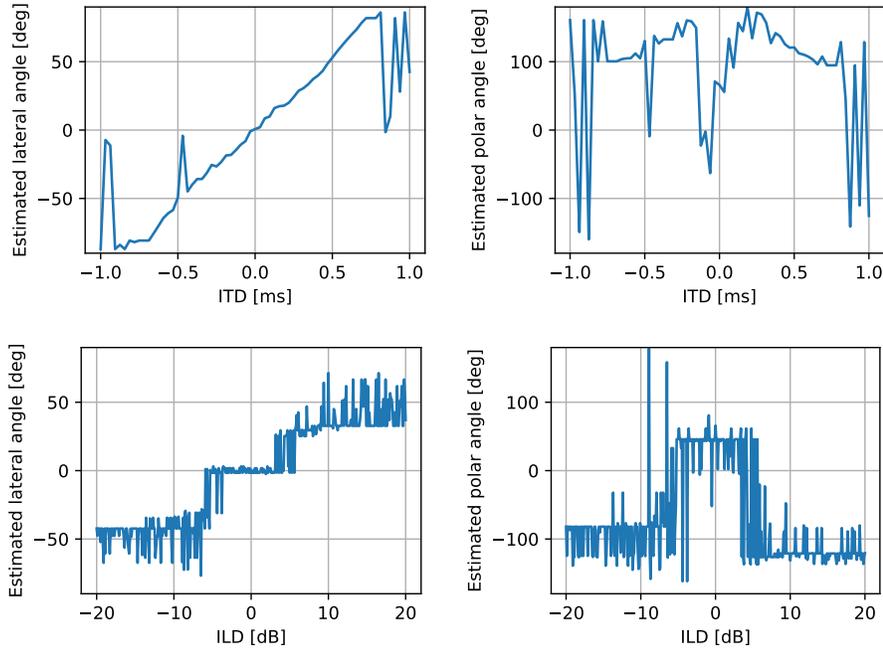


Figure 6.4: The estimated lateral (left column) and polar (right column) angle as function of ITD (top row) or ILD (bottom row) of the panned stereo signal.

6.4 Conclusion and discussion

We proposed a general-purpose individual-agnostic binaural localizer that does not require a priori knowledge about the generation process of the binaural audio and covers sound sources located anywhere on the 4π sphere. The novel model output format allows the mixture of HRIR datasets with arbitrary angular grids which realizes the expansion of available training data, facilitates individual-agnostic training, and together with the carefully designed training scheme, resulted in a robust binaural localizer which theoretically generalizes to alternative spatialization methods and multiple/moving sources. We have empirically shown the benefits of individual-agnostic training, data augmentation by noise and reverberation, and the use of soft targets. The proposed method was tested with real BRIRs as well as binaural recordings which contain noise and reverberation and

exhibited superior localization performance compared to individual-specific BLs or BLs trained with CCE or multi-task losses or hard targets. While some single-subject models performed well on one of the test sets, the results indicate that the individual-agnostic scheme may be more robust and generalize better.

Future work may include further refinement of the proposed model for multiple/moving sources. The use of recurrent layers or attention mechanisms in the network architecture and/or a probabilistic post-process, e.g., Bayesian filtering might improve the performance for moving sources. Finally, a comparison with humans' subjective localization tests may indicate whether the proposed model could be used to predict subjective localization performance in psychoacoustic studies.

Chapter 7: Regularized spherical harmonics-domain spatial active noise cancellation in a reverberant room

Active Noise Cancellation (ANC) at a target area in an open space, as opposed to cancellation in the ears through headphones, can lead to future applications. For instance, a personal acoustic environment in an airplane seat or inside a car, or a quiet zone in a noisy shared workspace can be possible using such open-space ANC without any uncomfortable on-body audio equipment. Recent advancements reinforce the practicality of such systems. However, regularization of the cancellation signal has been a crucial challenge in open-space ANC as it causes amplification of noise at locations away from the target area. This work presents a spherical harmonics-domain feed-forward spatial ANC method with a room-wide global cost function to address this issue. This room-wide global cost function is used for optimizing the set of regularization hyperparameters, while at run time only local information captured by a microphone array surrounding the target listening zone is required. Numerical experiments applying the proposed method in a simulated reverberant room show the effectiveness of the proposed method in creating a specific zone of silence with low to moderate noise amplification in the rest of the room.¹

¹The research presented in this chapter has been published in [151].

7.1 Introduction

Active noise cancellation (ANC) is a well studied problem in acoustics and audio signal processing, aiming at the reduction of the acoustic noise around the listening position using loudspeakers emitting a cancellation signal, which is sometimes referred to as the *anti-noise* [39,40]. While the traditional approach to this problem is to define a collection of control points in the vicinity of the listening position and generate a cancellation field from the loudspeakers which would reduce the total sound pressure at those control points, recently, *spherical harmonics* (SH)-domain signal processing has been applied to this problem in order to realize *spatial* ANC, where the cancellation problem is formulated in the SH-domain [152–154]. A potential problem of an ANC system is that the anti-noise emitted from multiple loudspeakers can constructively interfere with themselves or with the original noise field, which would turn the anti-noise into extra noise in some locations in the environment. A careful regularization methodology for synthesizing the cancellation signal is therefore a vital component in the design of such systems. Here, a regularization strategy, consisting of a local and a room-wide global cost function, is proposed to reduce such side-effects of a *spherical harmonics-domain ANC* (SHANC) system. The proposed global cost function is based on a metric which is referred to as the *silent-to-noisy area difference* (SNAD), and allows the regularized SHANC to create a *zone of silence* (ZoS) around the target position while limiting noise amplification outside the ZoS.

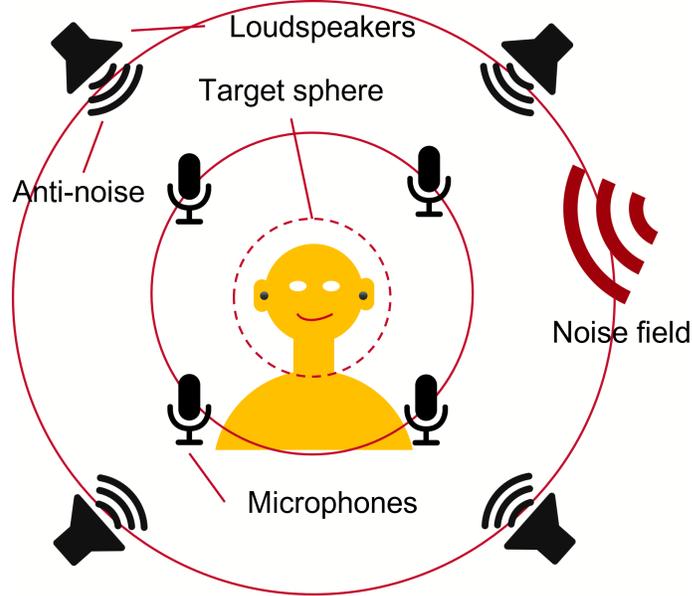


Figure 7.1: An illustration of the SHANC system. The microphones and loudspeakers in the present work are placed on spheres concentric with the target sphere. The target sphere, the microphone array, and the loudspeaker array have radii r_{rec} , r_{mic} , and r_{sp} , respectively, with $r_{\text{rec}} < r_{\text{mic}} < r_{\text{sp}}$.

7.2 Problem formulation

The goal of the SHANC system is to cancel the acoustic noise at a listening position in a reverberant room using a loudspeaker array and a microphone array surrounding the listening position. Examples of real world applications include spatial ANC in offices or car cabins. To achieve this, the algorithm needs to find an appropriate set of driving signals for the loudspeaker array given the observations by the microphone array. In the following, the formulation is presented in the frequency domain for a single frequency, while a time-domain representation can be obtained via inverse Fourier transform.

Fig. 8.1 shows the setup of the SHANC system presented in this paper. The origin O of the coordinate system is set to the listening position. The *target sphere* is defined as a sphere with radius r_{rec} centered at O . The complex amplitude of the signal emitted by

each loudspeaker is denoted as c_i with i the index of the individual loudspeaker. These amplitudes are referred to as the cancellation amplitudes. The objective is to reduce the band-limited spherical harmonics (SH)-domain total signal on the target sphere: $p_n^m(\mathbf{c}, r = r_{\text{rec}}) \equiv p_l$, while keeping the amount of extra noise added to the remaining area of the room minimal. Here, $0 \leq n \leq N_{\text{SH}}$, $-n \leq m \leq n$, $l \equiv n^2 + n + m$, and $\mathbf{c} = (c_0, c_1, \dots, c_{N_{\text{sp}}-1})^T$, with N_{SH} and N_{sp} the truncation order of the SH expansion and the number of loudspeakers, respectively. The SH-domain signal is defined as the following SH transform of a space-domain signal $p(r, \theta, \varphi)$:

$$p_n^m = \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} p(r, \theta, \varphi) Y_n^m(\theta, \varphi)^* d\Omega, \quad (7.1)$$

with Y_n^m the spherical harmonics of degree n and order m . Hereafter, $\mathbf{p}^{(\text{SH})} = (p_{l=0}, p_{l=1}, \dots, p_{l=(N_{\text{SH}}+1)^2-1})$ denotes a vector containing the $L_{\text{SH}} = (N_{\text{SH}} + 1)^2$ SH expansion coefficients of the field at $r = r_{\text{rec}}$. We distinguish three fields, namely the noise field $\mathbf{p}_{\text{noise}}^{(\text{SH})}$, the cancellation field $\mathbf{p}_{\text{cancel}}^{(\text{SH})}$ created by the cancellation loudspeaker array, and the total field $\mathbf{p}_{\text{total}}^{(\text{SH})}$. The total field $\mathbf{p}_{\text{total}}^{(\text{SH})}$ can be decomposed into the noise field $\mathbf{p}_{\text{noise}}^{(\text{SH})}$ and the generated cancellation field $\mathbf{p}_{\text{cancel}}^{(\text{SH})}$:

$$\mathbf{p}_{\text{total}}^{(\text{SH})} = \mathbf{p}_{\text{noise}}^{(\text{SH})} + \mathbf{p}_{\text{cancel}}^{(\text{SH})} = \mathbf{p}_{\text{noise}}^{(\text{SH})} + \sum_{i=0}^{N_{\text{sp}}-1} \mathbf{h}_{\text{sp} \rightarrow \text{rec}}^{(i)} c_i, \quad (7.2)$$

where the *room transfer function* (RTF) from the i -th loudspeaker (sp_i) position $\mathbf{r}_{\text{sp}}^{(i)}$ to the SH-domain representation of the signal at the listener position \mathbf{r}_{rec} is denoted as $\mathbf{h}_{\text{sp} \rightarrow \text{rec}}^{(i)} = \mathbf{h}(\mathbf{r}_{\text{sp}}^{(i)} \rightarrow \mathbf{r}_{\text{rec}})$. Again, $\mathbf{h}_{\text{sp} \rightarrow \text{rec}}$ is a vector representation of all the $L_{\text{SH}} = (N_{\text{SH}} + 1)^2$ SH expansion coefficients. The SH-domain RTFs h_n^m can be computed as SH transforms of

point-to-point RTFs $h(r_{\text{rec}}, \theta, \varphi)$ sampled on a sphere with radius r_{rec} ,

$$h_n^m = \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} h(r_{\text{rec}}, \theta, \varphi) Y_n^m(\theta, \varphi)^* d\Omega, \quad (7.3)$$

with $Y_n^m(\theta, \varphi)$ the spherical harmonics of the sampling points (θ, φ) in the spherical coordinates. The point-to-point RTFs $h(r_{\text{rec}}, \theta, \varphi)$ can be either measured in the real world or can be computed using simulation methods such as the boundary element method [155, 156]. For rooms with simple shapes, the method of images [21] can be employed for fast computation.

In the following, three scenarios, with different degrees of knowledge about the noise and room that is subject to cancellation, are defined.

In the first scenario, which is referred to as the *known* scenario, it is assumed that the noise source is a point source and the RTF from the noise source position \mathbf{r}_{src} to the SH-domain signal $\mathbf{p}_{\text{noise}}^{(\text{SH})}$ at the target sphere located at the listener position \mathbf{r}_{rec} , denoted as $\mathbf{h}_{\text{src} \rightarrow \text{rec}} = \mathbf{h}(\mathbf{r}_{\text{src}} \rightarrow \mathbf{r}_{\text{rec}})$, is known. It is also assumed that the true RTF from the cancellation loudspeakers to the target sphere is known. In this case, $\mathbf{p}_{\text{noise}}^{(\text{SH})} = \mathbf{h}_{\text{src} \rightarrow \text{rec}} s$, with s the complex amplitude of the signal emitted from the noise source. This scenario serves as an ideal baseline case.

In the second scenario, which is referred to as the *unknown* scenario, the noise field can be an arbitrary field, however it is assumed that the true RTF from the cancellation loudspeakers to the target sphere is known. In this scenario, the SH representation of the noise field $\mathbf{p}_{\text{noise}}^{(\text{SH})}$ needs to be estimated using a set of microphones. This set of microphones is referred to as the *estimation microphone array*.

In the third scenario, which is referred to as the *blind* scenario, the noise field can

be an arbitrary field, and also the true RTFs from the cancellation loudspeakers to the target sphere are not assumed to be known, but need to be estimated using the estimation microphone array.

7.3 Spherical harmonics-domain ANC method

A *local* cost function L which is a sum of the norm of the SH-domain signal and a regularization term is defined as:

$$L_{\lambda}^{(\text{loc})}(\mathbf{c}) = \|\mathbf{p}_{\text{total}}^{(\text{SH})}\|_2^2 + \lambda \|\mathbf{c}\|_2^2, \quad (7.4)$$

where $\lambda > 0$ is a regularization parameter. This regularization parameter controls the trade-off between the cancellation level on the target sphere and the amount of extra noise added to the room.

The problem is to find the vector \mathbf{c} minimizing the local cost function $L_{\lambda}^{(\text{loc})}(\mathbf{c})$. In the case where both of the norms in the cost function $L_{\lambda}^{(\text{loc})}(\mathbf{c})$ are l^2 -norms, the optimal \mathbf{c} is given by the least-square solution:

$$\mathbf{c}_{\text{opt}} = -(H_{\text{sp} \rightarrow \text{rec}}^H H_{\text{sp} \rightarrow \text{rec}} + \lambda I)^{-1} H_{\text{sp} \rightarrow \text{rec}}^H \mathbf{p}_{\text{noise}}^{(\text{SH})}, \quad (7.5)$$

with $H_{\text{sp} \rightarrow \text{rec}} = (\mathbf{h}_{\text{sp} \rightarrow \text{rec}}^{(0)}, \mathbf{h}_{\text{sp} \rightarrow \text{rec}}^{(1)}, \dots, \mathbf{h}_{\text{sp} \rightarrow \text{rec}}^{(N_{\text{sp}}-1)})$ an $L_{\text{SH}} \times N_{\text{sp}}$ matrix of RTFs.

Alternatively, the l^1 -norm could be used as the regularization term in L :

$$L_{\lambda}^{(\text{loc})}(\mathbf{c}) = \frac{1}{2L_{\text{SH}}} \|\mathbf{p}_{\text{total}}^{(\text{SH})}\|_2^2 + \lambda \|\mathbf{c}\|_1, \quad (7.6)$$

This problem is known as LASSO optimization. The optimal \mathbf{c} can be obtained by algorithms such as coordinate descent [157] or least-angle regression [158]. Preliminary

experiments using l^1 -norm regularization (7.6) did not indicate improvement over the standard l^2 -norm based regularization (7.4) in terms of the performance metrics that will be defined below, hence this type of regularization will not be discussed further in this paper.

7.3.1 Scenario *known*: known monopole noise source in a known room

In this scenario, the cost function in the case of l^2 -regularization becomes:

$$L_{\lambda}^{(\text{loc})}(\mathbf{c}) = \|\mathbf{h}_{\text{src} \rightarrow \text{rec}} s + \mathbf{p}_{\text{cancel}}^{(\text{SH})}\|_2^2 + \lambda \|\mathbf{c}\|_2^2, \quad (7.7)$$

and the optimal \mathbf{c}_{opt} is given by:

$$\mathbf{c}_{\text{opt}} = -(H_{\text{sp} \rightarrow \text{rec}}^H H_{\text{sp} \rightarrow \text{rec}} + \lambda I)^{-1} H_{\text{sp} \rightarrow \text{rec}}^H \mathbf{h}_{\text{src} \rightarrow \text{rec}} s. \quad (7.8)$$

7.3.2 Scenario *unknown*: arbitrary noise source in a known room

In the case of arbitrary noise sources, the SH representation of the noise field $\mathbf{p}_{\text{noise}}^{(\text{SH})}$ needs to be estimated using the estimation microphone array. The microphones are assumed to be located on a surface of a hollow sphere with radius r_{mic} centered at O . The position of the j -th microphone is denoted as $\mathbf{r}_{\text{mic}}^{(j)}$.

An arbitrary incident noise field p_{noise} can be expressed using the regular spherical basis functions of the Helmholtz equation:

$$p_{\text{noise}}(r, \theta, \phi, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m(k) j_n(kr) Y_n^m(\theta, \phi). \quad (7.9)$$

For processing in reality, the infinite series is truncated and in (7.9) can be written in

vector form:

$$\mathbf{p}_{\text{noise}}(\mathbf{r}_{\text{mic}}) = \Lambda(r_{\text{mic}})\mathbf{A}, \quad (7.10)$$

where $\mathbf{p}_{\text{noise}}(\mathbf{r}_{\text{mic}})$ is a vector holding $p_{\text{noise}}(r_j, \theta_j, \phi_j, k)$ in its j -th entry, \mathbf{A} is a vector holding $A_n^m(k)$ in its $(n^2 + n + m + 1)$ -th entry, and Λ is a matrix which is holding $j_n(kr_{\text{mic}})Y_n^m(\theta_j, \phi_j)$ in its $(j, n^2 + n + m + 1)$ entry.

$\mathbf{p}_{\text{noise}}^{(\text{SH})}$, which is the representation at $r = r_{\text{rec}}$, need to be estimated from p_{noise} sampled at a finite number of microphones located at $r = r_{\text{mic}}$. This can be solved by least-squares:

$$\begin{aligned} \mathbf{p}_{\text{noise}}^{(\text{SH})}(\mathbf{r}_{\text{rec}})|_n^m &= j_n(kr_{\text{rec}}) \cdot [(\Lambda^H \Lambda + \sigma I)^{-1} \Lambda^H \mathbf{p}_{\text{noise}}(\mathbf{r}_{\text{mic}})]_n^m \\ &= j_n(kr_{\text{rec}}) \cdot [H^{(\text{enc})}(\sigma) \mathbf{p}_{\text{noise}}(\mathbf{r}_{\text{mic}})]_n^m, \end{aligned} \quad (7.11)$$

with $H^{(\text{enc})}(\sigma) \equiv (\Lambda^H \Lambda + \sigma I)^{-1} \Lambda^H$.

In vector form, this can be written as:

$$\mathbf{p}_{\text{noise}}^{(\text{SH})}(\mathbf{r}_{\text{rec}}) = J(kr_{\text{rec}})H^{(\text{enc})}(\sigma)\mathbf{p}_{\text{noise}}(\mathbf{r}_{\text{mic}}), \quad (7.12)$$

where $J(kr_{\text{rec}})$ is a diagonal matrix containing $j_n(kr_{\text{rec}})$ in its $(l + 1, l + 1)$ entry for $l \in [n^2, n^2 + 2n]$, $0 \leq n \leq N_{\text{SH}}$.

The noise field at the microphone position j , $p_{\text{noise}}^{(j)} = p_{\text{noise}}(\mathbf{r}_{\text{mic}}^{(j)})$ can be computed by subtracting the component generated by the cancelling loudspeakers from the total observed amplitude:

$$p_{\text{noise}}^{(j)} = p_{\text{total}}^{(j)} - \sum_{i=0}^{N_{\text{sp}}-1} h_{s_i \rightarrow m_j} c_i, \quad (7.13)$$

or in vector form:

$$\mathbf{p}_{\text{noise}}(\mathbf{r}_{\text{mic}}) = \mathbf{p}_{\text{total}}(\mathbf{r}_{\text{mic}}) - H_{\text{sp} \rightarrow \text{mic}} \mathbf{c}. \quad (7.14)$$

Substituting this into (7.12) yields:

$$\mathbf{p}_{\text{noise}}^{(\text{SH})}(\mathbf{r}_{\text{rec}}) = J(kr_{\text{rec}})H^{(\text{enc})}(\sigma)(\mathbf{p}_{\text{total}}(\mathbf{r}_{\text{mic}}) - H_{\text{sp} \rightarrow \text{mic}}\mathbf{c}). \quad (7.15)$$

The SH-representation of the total field on the target sphere is therefore given as:

$$\begin{aligned} \mathbf{p}_{\text{total}}^{(\text{SH})} &= \mathbf{p}_{\text{noise}}^{(\text{SH})} + \mathbf{p}_{\text{cancel}}^{(\text{SH})} \\ &= \mathbf{p}_{\text{noise}}^{(\text{SH})} + \sum_{i=0}^{N_{\text{sp}}-1} \mathbf{h}_{\text{sp} \rightarrow \text{rec}}^{(i)} c_i \\ &= \mathbf{p}_{\text{noise}}^{(\text{SH})} + H_{\text{sp} \rightarrow \text{rec}}\mathbf{c} \\ &= J(kr_{\text{rec}})H^{(\text{enc})}(\sigma)\mathbf{p}_{\text{noise}}(\mathbf{r}_{\text{mic}}) + H_{\text{sp} \rightarrow \text{r}}\mathbf{c}. \end{aligned} \quad (7.16)$$

Here it is assumed that $H_{\text{sp} \rightarrow \text{mic}}$, the room impulse responses from the cancellation loudspeakers to the estimation microphones, is known and therefore $\mathbf{p}_{\text{noise}}(r_{\text{mic}})$ can be obtained from the signal recorded by the estimation microphones.

Hence, the optimal cancellation amplitudes \mathbf{c}_{opt} can be obtained by:

$$\begin{aligned} \mathbf{c}_{\text{opt}} &= -(H_{\text{sp} \rightarrow \text{rec}}^H H_{\text{sp} \rightarrow \text{rec}} + \lambda I)^{-1} H_{\text{sp} \rightarrow \text{rec}}^H \mathbf{p}_{\text{noise}}^{(\text{SH})}(\sigma) \\ &= -(H_{\text{sp} \rightarrow \text{rec}}^H H_{\text{sp} \rightarrow \text{rec}} + \lambda I)^{-1} H_{\text{sp} \rightarrow \text{rec}}^H J(kr_{\text{rec}})H^{(\text{enc})}(\sigma)\mathbf{p}_{\text{noise}}(\mathbf{r}_{\text{mic}}), \end{aligned} \quad (7.17)$$

Note that this solution diverges at frequencies where $j_n(kr_{\text{mic}}) = 0$. This property would limit the controllable frequencies. The first positive frequency where this happens is where $kr_{\text{mic}} = \pi$. In $0 < kr_{\text{mic}} < \pi$, there are no such poles. For an estimation microphone array of $r_{\text{mic}} = 20$ cm, the maximum frequency of this stable range is about 850 Hz. In addition, the microphone spacing also limits the frequency range due to spatial aliasing. Note that a second hyperparameter σ was introduced for the estimation procedure, which makes the local cost function $L_{\lambda, \sigma}^{(\text{loc})}(\mathbf{c})$ conditioned by two

hyperparameters λ and σ .

7.3.3 Scenario *blind*: arbitrary noise source and estimated room transfer functions

The estimation of the RTFs from the cancellation loudspeakers to the target sphere is performed in the same way as presented in the *unknown* scenario, using the same microphone array.

7.3.4 The global cost function for hyperparameter optimization

It is left to choose a set of hyperparameters that are suited for the objective of the regularized SH-domain ANC problem. For this, the following notions are defined. *Control zone*: a 2- or 3-dimensional subset of the room in which the amount of extra added noise due to the ANC system should be limited. *Region of silence (RoS)*: the region in the control zone which has cancellation above 10 dB. *Region of noise (RoN)*: the region in the control zone which has more than 10 dB of extra added noise due to the ANC system. *Zone of silence (ZoS)*: the connected domain of the RoS which contains the target listening position \mathbf{r}_{rec} . *Sphere of silence (SoS)*: the inscribing sphere of the ZoS centered at \mathbf{r}_{rec} . The *global* cost function, which is based on a metric that is referred to as the silent-to-noisy area difference (SNAD), is defined as the following:

$$L_w^{(\text{glob})}(\lambda, \sigma) = -\text{SNAD}(\lambda, \sigma) = -\{wA_{\text{SoS}}(\lambda, \sigma) - (1 - w)A_{\text{RoN}}(\lambda, \sigma)\}, \quad (7.18)$$

where A_{SoS} and A_{RoN} are the areas (or volumes, if the control zone is a 3-dimensional domain) of the SoS and RoN, respectively, and w is a weighting hyper-hyperparameter. This definition of the global cost function is assuming that the noise field in the control

zone and the RTFs from the cancellation loudspeakers to points in the control zone can be measured during the development phase of the ANC system for a specific noise and room, in order to determine an appropriate set of hyperparameters. Note that this global control zone-wide information is only used for the optimization of the global cost function and not for the local optimization problem which is, in the *blind* case, only relying on the signal captured by the estimation microphone array. This means that after identifying an appropriate set of hyperparameters in the development phase, at run time, the SHANC system can operate using only the local information captured by the estimation microphone array.

7.4 Numerical Experiments

Numerical experiments were performed to study the performance of the regularized SHANC system. Here, the experimental conditions are described.

7.4.1 Room condition

A cuboid shaped room was used for the numerical experiments. The size of the room in the x , y , and z axis was set to 4, 3, and 2.5 m, respectively. Hence, the room is defined as the interior of the cuboid $\{(x, y, z) | 0 \leq x \leq L_x = 4, 0 \leq y \leq L_y = 3, 0 \leq z \leq L_z = 2.5\}$, where the coordinates here are in the room frame, which is different than the listener frame used in the previous section for the spherical harmonics expansion. The specific acoustic impedance of the walls of the room was set to 19, and the same value was used for all six walls. The position of the listener \mathbf{r}_{rec} was set to $(1, 1, 1)$ in the room coordinate frame which has its origin at one corner of the room. The noise source was a

monopole source with its position set to $(3, 2, 1)$ in the room coordinate frame.

7.4.2 Room transfer function simulation

The RTFs required for the numerical experiments were computed using the method of images [21], for 200 Hz, 500 Hz, 1kHz, and 2kHz, with maximum reflection order of 30. The control zone was chosen as the set of points that are lying in a plane which passes through the target listener position \mathbf{r}_{rec} and which is parallel to the ground. This plane is referred to as the listening plane. The RTFs from sound sources to points on this listening plane were computed for a regular Cartesian grid with resolution of 2 cm.

7.4.3 Cancellation loudspeaker array and estimation microphone array configurations

In the numerical experiments, the configuration and the number of estimation microphones were fixed to the same as the cancellation loudspeakers. The loudspeakers and microphones were placed on concentric hollow spheres with radii of $r_{\text{sp}} = 0.5$ m and $r_{\text{mic}} = 0.2$ m, respectively. The configuration types used in the experiments are listed in Table 8.1. In addition to the point sets based on the vertices of regular and semi-regular polyhedra, spherical Fibonacci grids [66] were used in the experiments. The spherical Fibonacci grids is a point set defined on a sphere for any positive number of total points. Given an arbitrary number of points, it has an explicit expression of point coordinates. Due to this convenient property and its fairly uniform point distribution, it has served in various applications including quasi-Monte Carlo integration in computer graphics [159] and as a microphone array configuration for spatial sound field capturing [18].

Table 7.1: Microphone / loudspeaker array configuration used in the numerical experiments.

Configuration type	#microphones / loudspeakers
Tetrahedron	4
Octahedron	6
Cube	8
Icosahedron	12
Dodecahedron	20
Pentakis dodecahedron	32
Truncated icosahedron	60
Spherical Fibonacci grid	16, 24, 48, 64, 96, 128

7.4.4 Optimization procedure

The optimization procedure consists of two steps. For a given set of hyperparameters $\{\lambda, \sigma\}$, the local optimization problem $\mathbf{c}_{\text{opt}} = \operatorname{argmin}_{\mathbf{c}} L_{\lambda, \sigma}^{(\text{loc})}(\mathbf{c})$ is solved by least squares. The optimizer of the global cost $L_w^{(\text{glob})}(\lambda, \sigma)$ is sought with grid-search over the two-dimensional hyperparameter space. The maximum order of the spherical harmonics expansion N_{SH} was set to the smallest integer that satisfies $(N_{\text{SH}} + 1)^2 > N_{\text{sp}}$. The weighting hyper-hyperparameter w was set to 0.95.

7.5 Results

The resulting sound fields on the listening plane for the case of a noise source of 500 Hz and a SHANC system of 12 channels are shown in Fig. 7.2 (top row) for the *blind* scenario. It can be observed that the system successfully creates a SoS with a radius of 19.1 cm. On the target sphere, an average noise reduction level of 29.8 dB was observed while preserving the AoN to 0.279 m², which is only 2.3 % of the area of the room. As reference, SHANC was applied with the regularization terms set to zero (Fig. 7.2, bottom row). In this case, the resulting SoS radius was only 13.2 cm while the injected anti-noise

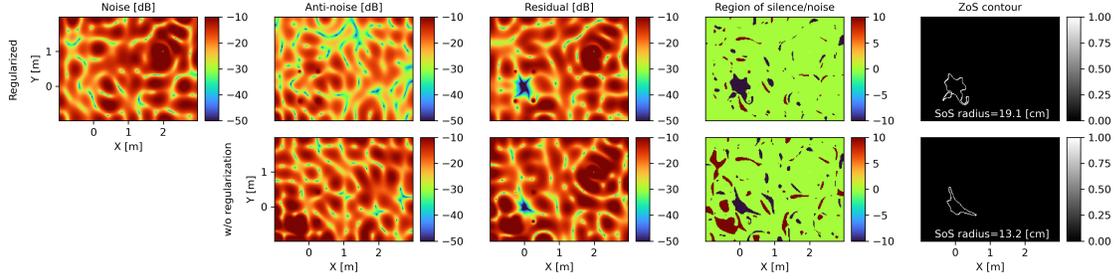


Figure 7.2: The resulting fields on the listening plane for the case of a noise source of 500 Hz and a SHANC system consisting of 12 estimation microphones and 12 cancellation loudspeakers, in the *blind* scenario. From left to right: sound pressure level of the noise field, anti-noise field, total residual field, the RoS/RoN map, and the contour of the ZoS, respectively. The upper and lower row shows the result for the proposed regularized SHANC method and the reference case without regularization, respectively.

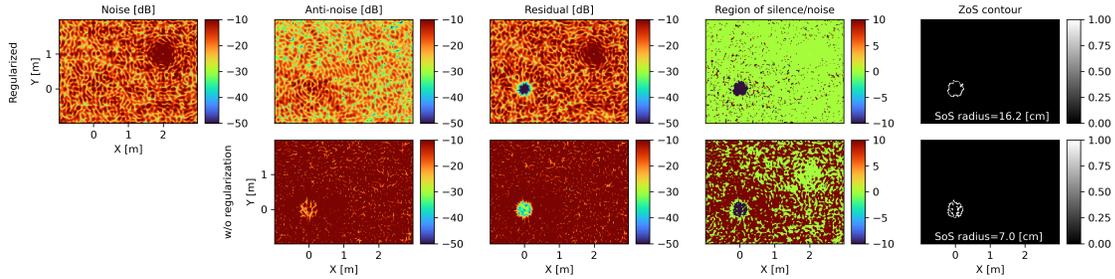


Figure 7.3: The resulting fields on the listening plane for the case of a noise source of 2 kHz and a SHANC system consisting of 128 estimation microphones and 128 cancellation loudspeakers, in the *blind* scenario. The meaning of each of the subplot is identical to Fig. 7.2.

was about 5 dB higher in level. Results for 2kHz with $N_{sp} = 128$ are shown in Fig. 7.3, where the average noise reduction level on the target sphere was 37 dB with an AoN of only 0.423 m² using the regularized SHANC method.

The level of noise reduction on the target sphere, as well as the radius of the SoS for various device configurations and target frequencies are summarized in Fig. 7.4, Fig. 7.5, and Fig. 7.6 for the *known*, *unknown*, and *blind* scenario, respectively. The number of estimation microphones required for a successful creation of a sufficiently large SoS was found to be strongly dependent on the noise frequency.

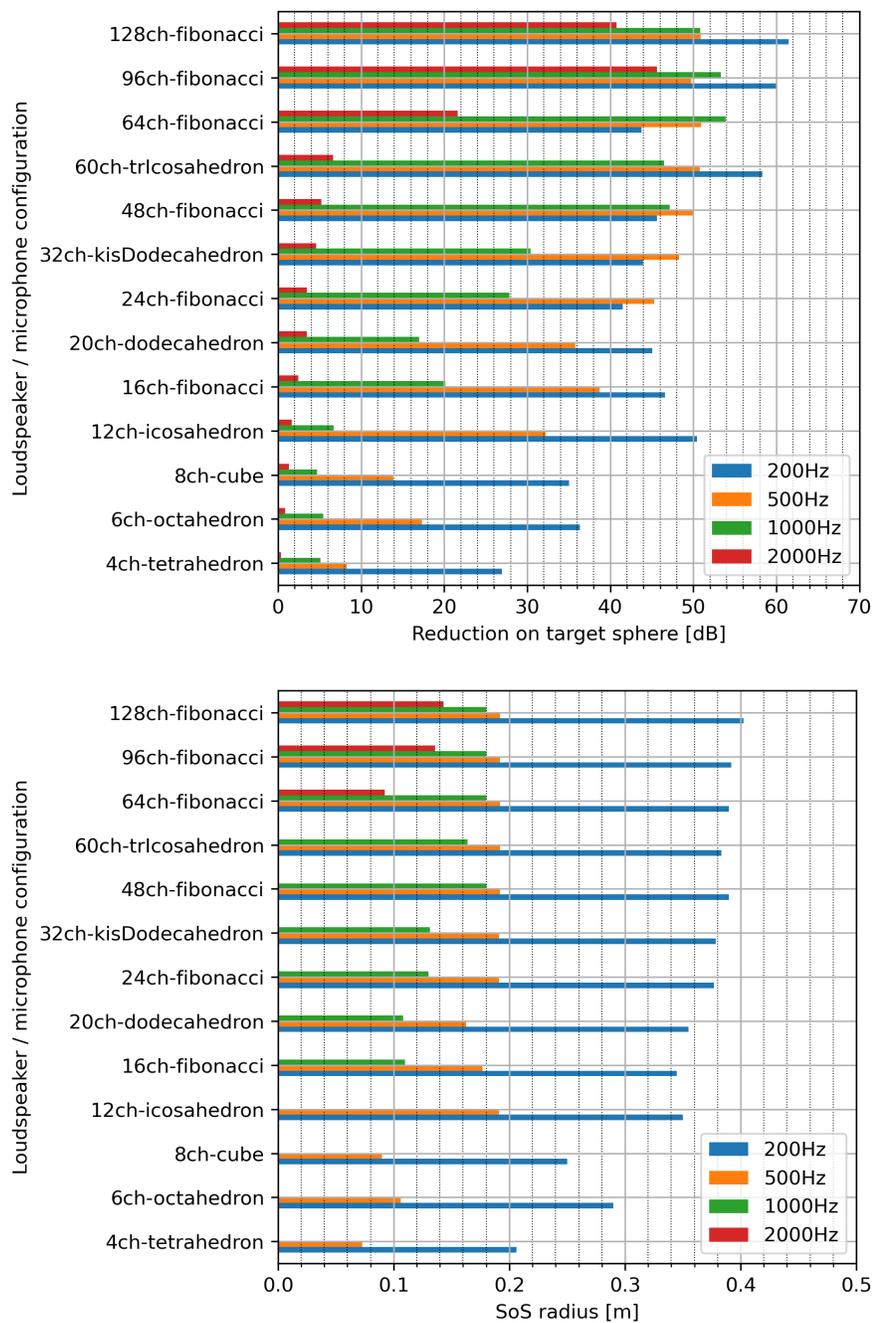


Figure 7.4: Summary of results for the *known* scenario. Noise reduction level on the target sphere (left) and the radius of the SoS (right), for various device configurations and frequencies.

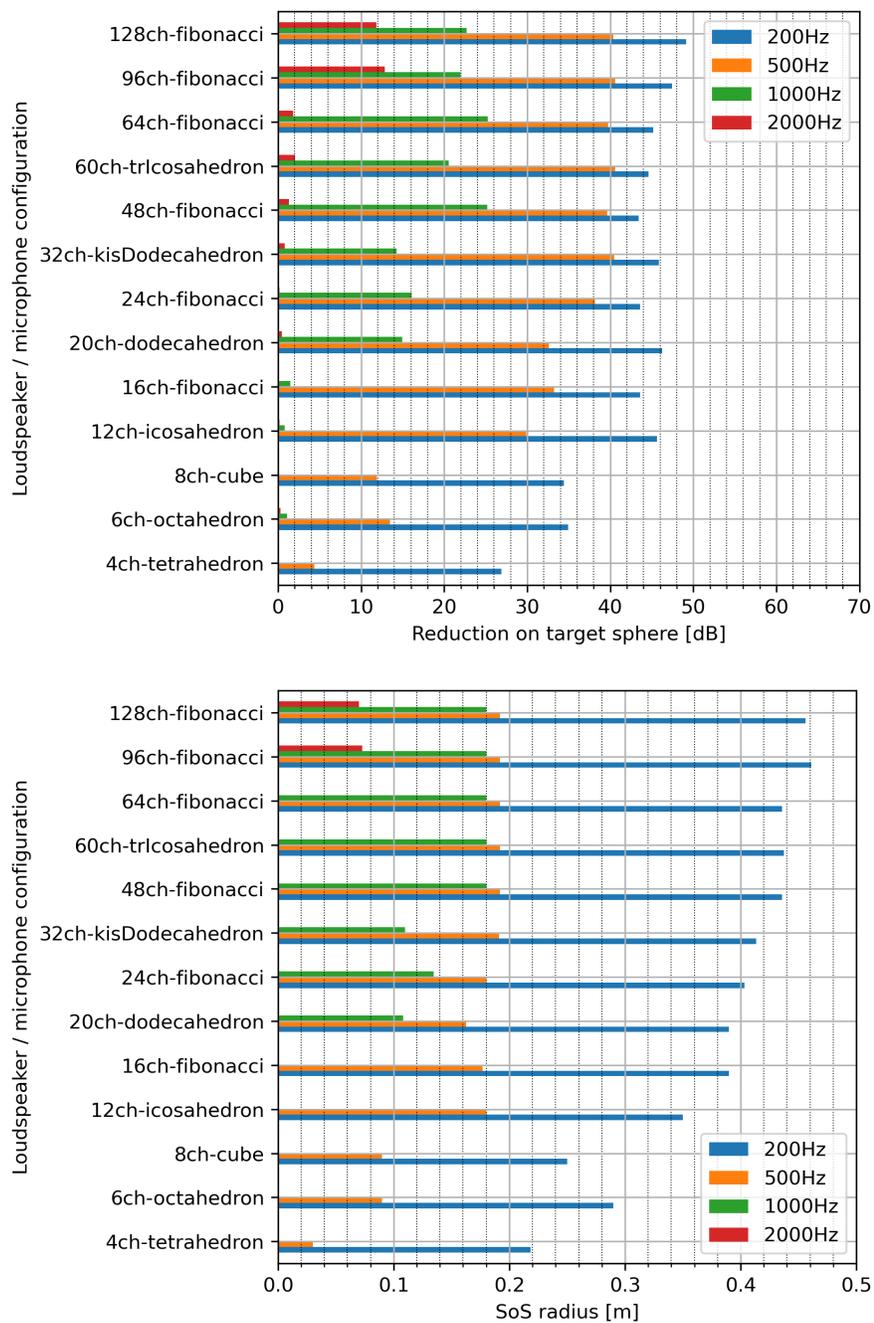


Figure 7.5: Summary of results for the *unknown* scenario. Noise reduction level on the target sphere (left) and the radius of the SoS (right), for various device configurations and frequencies.

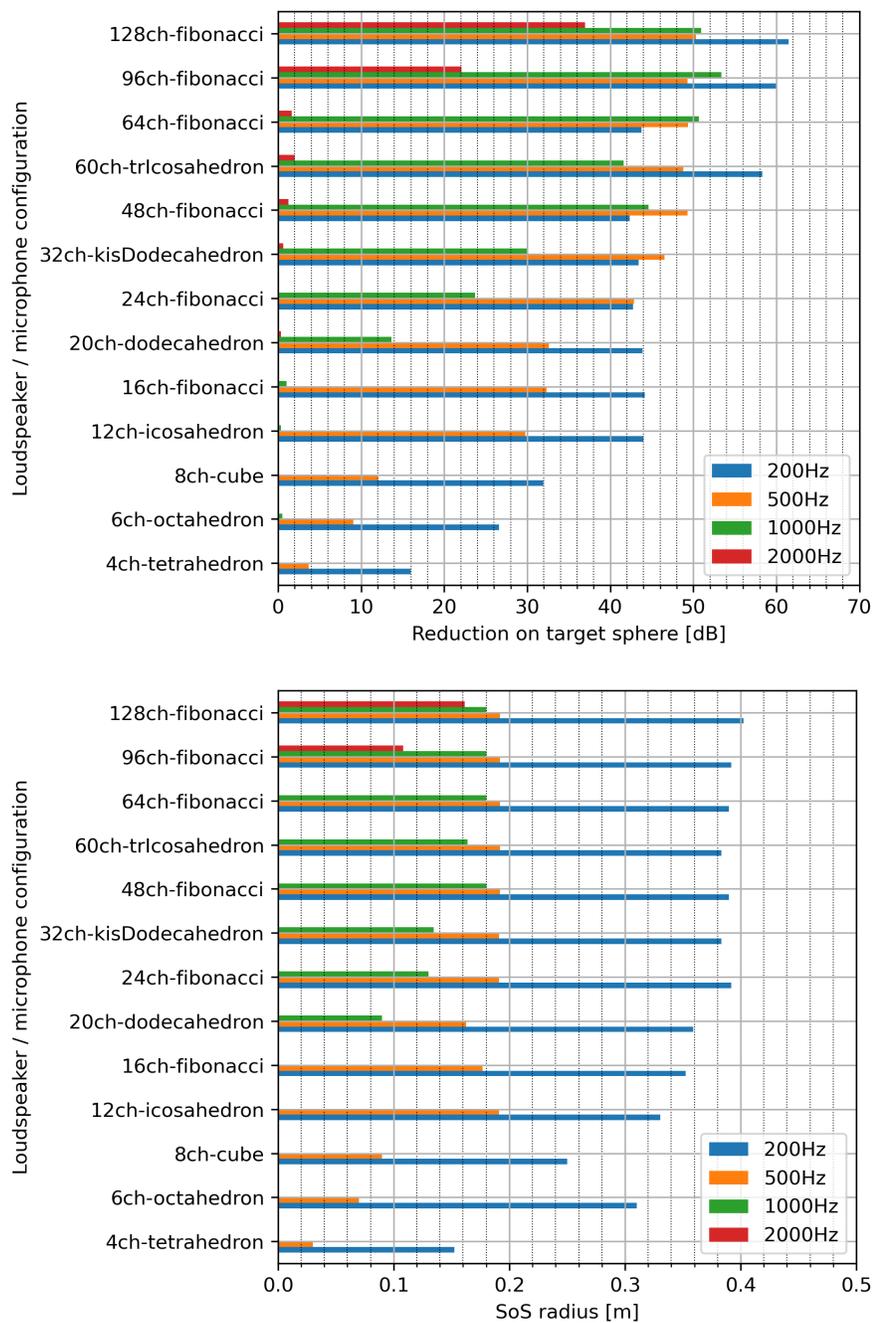


Figure 7.6: Summary of results for the *blind* scenario. Noise reduction level on the target sphere (left) and the radius of the SoS (right), for various device configurations and frequencies.

7.6 Conclusion

A regularized variant of spherical harmonics-domain spatial ANC, with a particular focus on limiting the noise amplification outside the target listening zone, was presented. Results of numerical experiments applying the proposed method to a SHANC problem in a reverberant room indicated successful creation of a zone of silence while maintaining noise amplification in the rest of the room moderate. The comparison with the reference method without regularization clearly shows that the proposed regularized SHANC method shrinks noisy regions and expands the silent region while reducing the anti-noise power. Future works include studying other microphone/loudspeaker array configurations for the reduction of non-periodic noise fields.

Chapter 8: Personal active soundfield control (PASCAL) framework

A framework for actively controlling radiated, incident, or local personal sound fields is presented. It relies on loudspeakers and microphones either worn by the user or surrounding the user. The framework aims to address tasks such as speech privacy, personal active noise cancellation, and immersive audio presentation with limited amplification / injection of noise or leakage of private speech into the environment. The formulation relies on modeling and simulation of the sound field using a fast multipole accelerated boundary element method, spectral or point measurements of the sound field, and regularized optimization of the field created by actively controlled speakers. The use of acoustic simulation enables the utilization of transfer functions associated with a large number of points distributed in space resulting in effective regularization. Radiation cancellation of up to 20 dB was observed in low frequencies below 1 kHz in a numerical experiment using real-world impulse responses of a wearable loudspeaker setup.¹

8.1 Introduction

Environmental and body worn loudspeakers and microphones open up the possibility of measuring, predicting and controlling the sound field reaching the ears of a listener,

¹The research presented in this chapter has been published in [160].

or radiated from their mouths. We propose methods to control the sound field to provide various possibilities including (i) **active radiation cancellation** (ARC) for speech privacy, (ii) personal **active noise cancellation** (ANC), and (iii) immersive audio presentation via transaural playback.

We develop a formulation in Section 8.2 that accounts for the scattering that is inherent in the applications envisaged – off the room, off room and furniture surfaces, and the person – via solutions of the Helmholtz equation via the **fast multipole accelerated boundary element method** (FMM-BEM) [155, 161, 162], and measurements of the field using microphones distributed in the environment and on the person, and that can be processed and collected in an approximately common clock by taking advantage of internet connectivity and the slow speed of sound [163, 164]. Section 8.3 discusses the control algorithms that are to be solved for each of the three problems identified, and the possible issues and their mitigation via regularization strategies. Section 8.4 presents results for the ARC and ANC task. Section 8.5 concludes the paper with discussions on future developments.

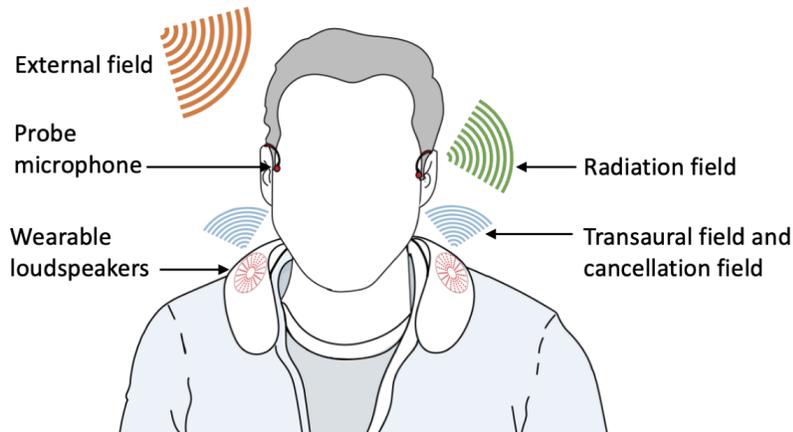


Figure 8.1: An example system implementing PASCAL. A set of loudspeakers and microphones is used to present the transaural signal and to suppress the external field by emitting a cancellation field.

8.2 Problem Formulation

The proposed framework aims to control the radiated, incident, or local sound field associated with an individual user who can either be a speaker or a listener. A key to achieve this is modeling of the acoustic **transfer functions** (TFs) which include all acoustic scattering involved in the environment. The system consists of a user, an arbitrary set of scatterers, a set of N_s loudspeakers and N_m microphones which can be either body-worn or distributed in space, N_t target listening points, and optionally a room enclosing all these objects. The microphone positions are referred to as the *probe points*. The target listening points are referred to as *test points*, and correspond to the listener's ear positions. Given a specific geometry and boundary conditions in the system, the TFs from sound sources to observation points can be evaluated by solving the Helmholtz equation numerically using the FMMBEM. We decompose the total sound field in the system $p^{(\text{tot})}$ as:

$$p^{(\text{tot})} = p^{(\text{rad})} + p^{(\text{ext})} + p^{(\text{can})} + p^{(\text{tra})}, \quad (8.1)$$

where $p^{(\text{rad})}$ is the field generated by the user's vocal radiation, $p^{(\text{ext})}$ is the ambient sound arriving from the environment, $p^{(\text{can})}$ is the field generated by the loudspeakers in order to cancel $p^{(\text{ext})}$ and/or $p^{(\text{rad})}$, and $p^{(\text{tra})}$ is the sound field generated by the loudspeakers in order to present the immersive audio to the user, respectively. We refer to $p^{(\text{rad})}$, $p^{(\text{ext})}$, $p^{(\text{can})}$, and $p^{(\text{tra})}$ as the *radiation field*, *external field*, *cancellation field*, and *transaural field*, respectively. Sound field control can be based either on a multi-point pressure-matching scheme or on a band-limited **spherical harmonics** (SH)-domain mode-matching

scheme. The SH-domain representation can be computed by SH transform defined as:

$$p_n^m = \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi} p(r, \theta, \varphi) Y_n^m(\theta, \varphi)^* \sin \theta d\theta d\varphi, \quad (8.2)$$

with p the sound pressure at point (r, θ, φ) in the spherical coordinates and Y_n^m the spherical harmonic of degree n and order m . Fig. 8.1 illustrates an example setup of the **personal active soundfield control algorithm** (PASCAL) framework.

8.3 Tasks and objectives

Our first task is (i) **active radiation cancellation** (ARC) of $p^{(\text{rad})}$ generated by the user. The objective can be the minimization of the residual field $p^{(\text{res})} = p^{(\text{rad})} + p^{(\text{can})}$ at target points, the SH-domain signal on a sphere enclosing the user, or an objective related to intelligibility.

The second task is (ii) **personal ANC** [39, 40] of the external field. The goal is to generate a cancellation field which minimizes the amplitude of the residual field $p^{(\text{res})}$, defined as $p^{(\text{res})} = p^{(\text{ext})} + p^{(\text{can})}$, at the test points. Alternatively, the objective can be the minimization of the band-limited SH-domain representation of $p^{(\text{res})}$ [152–154]. This can be done by capturing the SH-domain external field using a spherical microphone array enclosing the user, or using a microphone array worn by the user who is considered as a scatterer [70].

The third task is (iii) **crosstalk cancellation** (XTC) [165] for immersive spatial audio presentation. The goal is to produce sound at the user’s ears which matches an input binaural signal. Alternatively, as in ambisonics decoding in spatial audio reproduction [16],

the SH-domain incident field can be the reproduction target.

These tasks are instances of a constrained multi-point pressure-matching problem, or a SH mode-matching problem, and solved by computing the required TFs, represented either in space-domain or SH-domain, using FMMBEM and computing the optimal filters as described below. The following challenges must be addressed: (i) the radiation cancellation field can constructively interfere with the radiated field in some locations, resulting in leakage of private speech (e.g., in a coffee shop or airport setting). (ii) the noise cancellation field may constructively interfere with the external field at some locations and create extra noise there. (iii) the transaural field is radiated not only to the listener's ears but also the environment as distracting sound. It is therefore crucial to avoid fields $p^{(\text{can})}$ and $p^{(\text{tra})}$ with excessive amplitudes which cause these side-effects, and is done via a regularization scheme described later.

8.3.1 Objective functions and optimal driving signals

The formulation is presented in the frequency domain and arguments denoting frequency are omitted. In the following, the cancellation signal is denoted by a vector $\mathbf{c}^{(\text{can})} = (c_1, c_2, \dots, c_{N_s})^T \in \mathbb{C}^{N_s}$. Here, c_i is the frequency dependent complex amplitude representing the signal for the i -th loudspeaker. For the ARC task (i), the cost function is defined as:

$$L_\lambda^{(\text{ARC})} = \|\mathbf{p}_{\text{test}}^{(\text{rad})} + H_{\text{ts}}\mathbf{c}^{(\text{can})}\|_2^2 + \lambda\|\mathbf{c}^{(\text{can})}\|_2^2, \quad (8.3)$$

with $\mathbf{p}_{\text{test}}^{(\text{rad})} \in \mathbb{C}^{N_t}$ the radiation field at the test points, H_{ts} the $N_t \times N_s$ matrix holding the TFs from the loudspeakers to the test points, and $\lambda > 0$ a regularization parameter. The

optimal driving signal is obtained by least squares:

$$\mathbf{c}_{\text{opt}}^{(\text{can})} = -(H_{\text{ts}}^H H_{\text{ts}} + \lambda I)^{-1} H_{\text{ts}}^H \mathbf{p}_{\text{test}}^{(\text{rad})}, \quad (8.4)$$

where $\mathbf{p}_{\text{test}}^{(\text{rad})}$ is computed from the signal captured via a microphone near the mouth and TFs from the mouth to the test points.

In the ANC task (ii), $\mathbf{c}^{(\text{can})}$ is produced using the total field measured at the probe points $\mathbf{p}_{\text{probe}}^{(\text{tot})} \in \mathbb{C}^{N_m}$. $\mathbf{c}^{(\text{can})}$ should minimize the residual field at the test points $\mathbf{p}_{\text{test}}^{(\text{res})}$ and at the same time should have as small energy as possible. The cost function $L_{\lambda}^{(\text{ANC})}$ is defined as:

$$L_{\lambda}^{(\text{ANC})} = \|\mathbf{p}_{\text{probe}}^{(\text{res})}\|_2^2 + \lambda \|\mathbf{c}^{(\text{can})}\|_2^2. \quad (8.5)$$

The vector $\mathbf{c}^{(\text{can})}$ minimizing the cost $L_{\lambda}^{(\text{ANC})}$ is:

$$\mathbf{c}_{\text{opt}}^{(\text{can})} = -(H_{\text{ms}}^H H_{\text{ms}} + \lambda I)^{-1} H_{\text{ms}}^H \mathbf{p}_{\text{probe}}^{(\text{ext})} = H^{(\text{can})} \mathbf{p}_{\text{probe}}^{(\text{ext})}, \quad (8.6)$$

with H_{ms} a $N_m \times N_s$ matrix holding the TFs from the loudspeakers to the probe points and $\mathbf{p}_{\text{probe}}^{(\text{ext})} = \mathbf{p}_{\text{probe}}^{(\text{res})} - H_{\text{ms}} \mathbf{c}^{(\text{can})}$ the signal component due to the external field. $H^{(\text{can})} \equiv -(H_{\text{ms}}^H H_{\text{ms}} + \lambda I)^{-1} H_{\text{ms}}^H$ is referred to as the cancellation filter matrix.

For the XTC task (iii), the objective function is:

$$L_{\lambda}^{(\text{XTC})} = \|\mathbf{s}^{(\text{bin})} - H_{\text{rtf}} H_{\text{xtc}} \mathbf{s}^{(\text{bin})}\|_2^2 + \lambda \|H_{\text{xtc}} \mathbf{s}^{(\text{bin})}\|_2^2, \quad (8.7)$$

with H_{xtc} a $N_s \times 2$ matrix holding the XTC filters which is multiplied with the input

binaural signal $\mathbf{s}^{(\text{bin})}$ for generating $p^{(\text{tra})}$, and H_{rtf} a $2 \times N_s$ matrix holding the HRTFs from the loudspeakers to the listener's ears. The optimal XTC filters are obtained similarly:

$$H_{\text{xtc}} = (H_{\text{rtf}}^H H_{\text{rtf}} + \lambda I)^{-1} H_{\text{rtf}}^H. \quad (8.8)$$

The SH-domain formulation for each task replaces the signal vectors and TFs by their SH-domain counterparts. The L_1 norm could be alternatively used for regularization to enforce sparsity of loudspeaker activations, but losing the closed form results [158, 166]. The TFs required for the framework are assumed to be pre-computed and dynamically updated via motion and head tracking of the user. HRTF individualization techniques [30, 167, 168] including head/ear shape modeling techniques [58], allow adaptation to the user's individual head and ear shapes and can be utilized in PASCAL, as HRTFs are highly personal.

8.3.2 Cost function for regularization parameter optimization

The sound pressure at an additional set of N_o points, the *optimization points*, $\mathbf{p}_{\text{opt}} \in \mathbb{C}^{N_o}$ is used for the optimization of the regularization parameter λ in the development stage of the system. For task (i), the reduction level $R(\mathbf{p}_Q)$ on a set of points Q is defined as

$$R(\mathbf{p}_Q) = 20(\log_{10} \|\mathbf{p}_Q^{(\text{rad})}\|_2 - \log_{10} \|\mathbf{p}_Q^{(\text{res})}\|_2), \quad (8.9)$$

where $\mathbf{p}_Q^{(\text{rad})}$ and $\mathbf{p}_Q^{(\text{res})}$ are vectors holding the sound pressure of the external and the residual field at points Q , respectively. For the optimization of λ , the global cost function

is defined as

$$L^{(\text{glob})}(\lambda) = -R(\mathbf{p}_{\text{opt}}). \quad (8.10)$$

The sound pressure at the optimization points \mathbf{p}_{opt} is only used at the development stage for optimizing the hyperparameter λ and is not required at run time. Similar global cost functions can be designed for (ii) and (iii). Replacing $\mathbf{p}_Q^{(\text{rad})}$ by $\mathbf{p}_Q^{(\text{ext})}$ gives a global cost function for task (ii). Note that the number of test and optimization points is practically unlimited since the proposed framework relies on FMMBEM computation of the sound field, allowing the design of effective regularization. This is hardly achieved with expensive real-world multi-point IR measurements, potentially limiting the efficacy of regularization.

8.4 Numerical Experiments

We consider *wearable neckband loudspeakers* which are loudspeakers closely located to the user's ears. The performance of PASCAL using two to seven loudspeakers is evaluated via simulation.

8.4.1 Loudspeaker configurations and FMMBEM simulation

The neckband and laptop arrays are simulated. To include scattering from the head, neck, torso, laptop screen, and desk, a simplified mesh was used for the FMMBEM simulation of the TFs, as shown in Fig. 8.2 (left). A Robin boundary condition with specific admittance of 5.3×10^{-2} was used for all boundaries. The loudspeakers were modelled as monopole sources for the purpose of this exercise, though more complex radiation patterns can be simulated. Their positions are indicated with red dots in Fig. 8.2

(left). Twelve equispaced points on the circle of radius 10 cm were defined as the neckband loudspeaker positions. The laptop loudspeakers were located on seven equispaced points on a line segment with a length of 29 cm. Arrays consisting of two or seven loudspeakers were used in the experiments, by choosing a subset. The channels used in each configuration are shown in Fig. 8.2 (right) and Table 8.1, respectively. The BEM simulation of the TFs in the given system took around 30 seconds for each sound source at 500 Hz using a laptop.

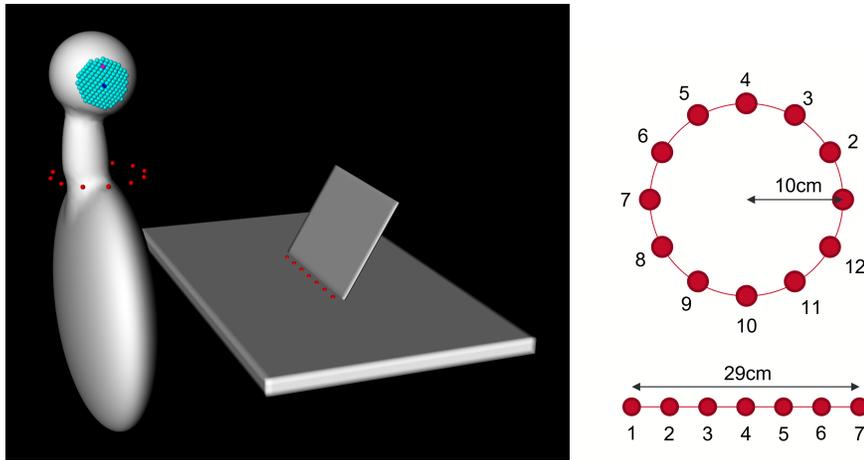


Figure 8.2: Left: the shape model used for the BEM simulation. The red, purple, blue, and cyan dots represent the set of loudspeakers, probe points, test points, and optimization points respectively. The probe points, test points, and optimization points are defined on both sides of the head. Right: loudspeaker channel assignment for the neckband array (top) and the laptop array (bottom). In the neckband array, channel 4 and 10 correspond to front and back, respectively.

Table 8.1: Loudspeaker channels used in each configuration.

# Loudspeakers	Neckband	Laptop
2	{1, 7}	{1, 7}
7	{1, 3, 5, 7, 9, 10, 11}	{1, 2, 3, 4, 5, 6, 7}

8.4.2 Radiation cancellation simulation

The ARC task (i) was studied for the neckband system. The vocal radiation was approximated by a monopole source located at the mouth. The test and optimization points were 2048 spherical Fibonacci grid points on spheres of radii 0.5 m and 0.7 m centered at the origin. The fields before and after cancellation for the 7ch system are shown in Fig. 8.3 (top). The reduction level for the 2ch and 7ch arrays are shown in Fig. 8.3 (middle). The ARC simulation was also conducted using IRs measured in a real anechoic room. A 48-channel planar microphone array (VisiSonics Digital Array Microphones) forming a 8×6 mesh grid of size 35 cm \times 25 cm was used for measuring the IRs from a source loudspeaker (YAMAHA VXS1MLB) and a cancellation stereo neckband loudspeaker (SONY SRS-WS1). The IR measurement points at $d = 76$ cm and $d = 96$ cm were used as the optimization points and test points, respectively, with d the distance from the source loudspeaker to the microphone array plane. The ARC filters were designed as time domain FIR filters and were convolved with a female human speech source signal to create the cancellation signal. The amplitudes at the microphone positions were then computed by convolving the source and cancellation signals with the measured IRs from source and cancellation loudspeakers, respectively. The amplitude before and after ARC at a single test point is shown in Fig. 8.3 (bottom). A reduction of up to 20 dB can be observed at frequencies below 1000 Hz.

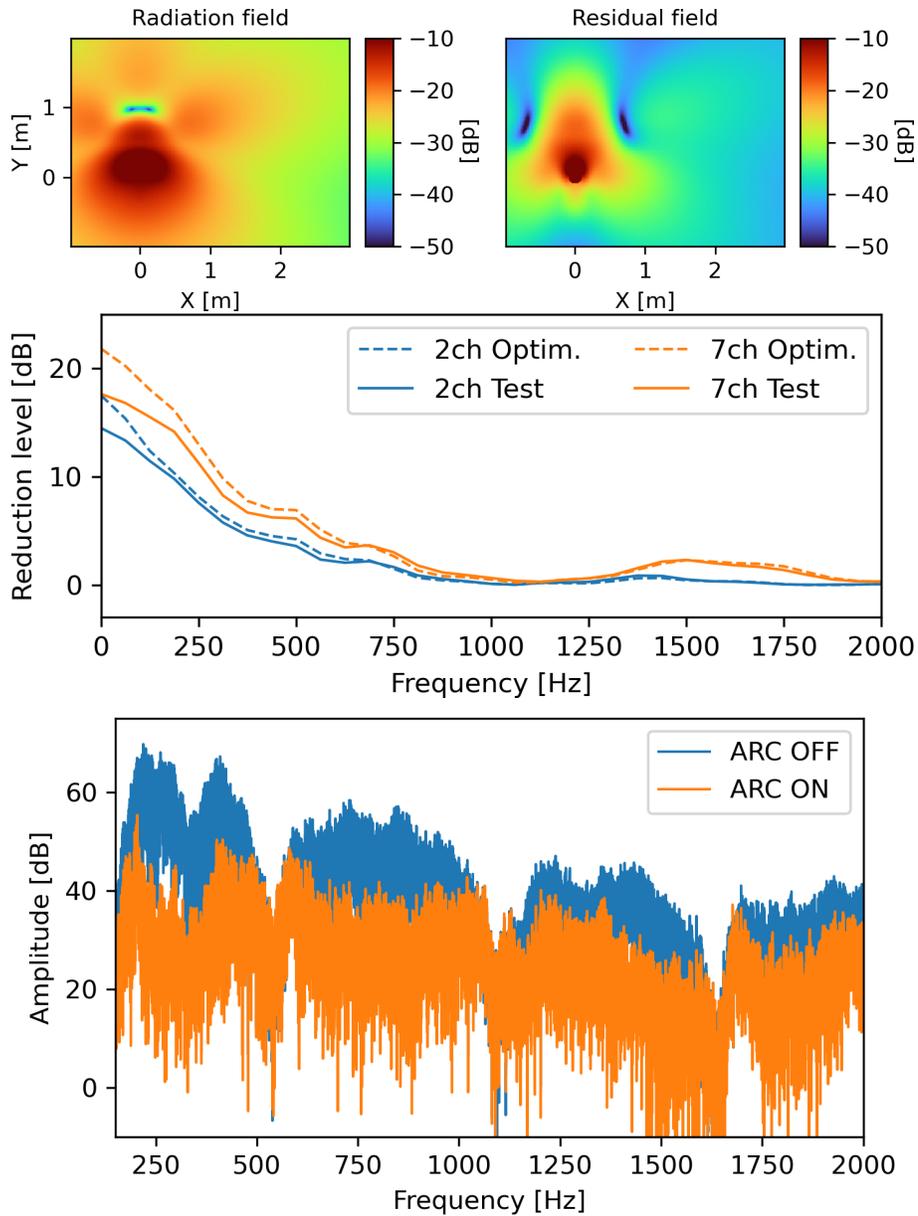


Figure 8.3: The radiation field (top left) and residual field (top right) at 500 Hz using the 7ch system, reduction level (middle) of the fully-simulated ARC experiment, and the test point amplitude spectrum of the ARC simulation using real IRs (bottom).

8.4.3 Personal ANC simulation

A monopole noise source generating a signal with amplitude -22 dB at 1 m distance was placed at point $\mathbf{r}_{\text{noise}} = (2\text{m}, 1\text{m}, 0)$ in Cartesian coordinates where the origin is the

center of sphere approximating the listener’s head and the x - and y -axes are pointing to the right and to the front from the origin. The test points corresponding to the entrance of the ear canals were placed 1 cm from the surface of the sphere with a radius of 9 cm approximating the head of the listener. The microphones, i.e. probe points are placed in the vicinity of the test points approximating the microphones installed in a smart glass or headset device. The distance between the test points and the nearest probe points was 38 mm and 37 mm for the left and right sides, respectively. The set of optimization points consists of $N_o = 271$ points in the vicinity of the assumed ear position including the probe points and test points and is indicated as the cyan dots in Fig. 8.2 (left). A subset of a spherical Fibonacci grid [66] inside of a cone with solid angle 0.842 for both the left and right direction were used as the optimization points. The regularization parameter λ minimizing $L^{(\text{glob})}(\lambda)$ was sought by grid search on a logarithmic grid in $[10^{-15}, 1]$. The reduction level on the probe points, optimization points, and test points was computed using the simulated TFs and is shown in Fig. 8.4 for four different loudspeaker configurations. The reduction level $R(\mathbf{p}_{\text{test}})$ at the test points at 500 Hz is about 10 dB and 11 dB for the neckband and laptop case, respectively. The increased number of loudspeakers did not improve $R(\mathbf{p}_{\text{test}})$ in the current setup which has only two probe points. The total power of the cancellation signal defined as $P_c = 20 \log_{10} \|\mathbf{c}^{(\text{can})}\|_2$ is shown in Fig. 8.5 (top). The power of the cancellation signal produced by the neckband is considerably lower than the laptop array in the lower frequencies where the ANC is effective, with a difference of about 14 dB at 500 Hz in the stereo loudspeaker case. Increasing the number of loudspeakers helps reduce the total power of the cancellation signal. Fig. 8.5 (bottom) shows the cancellation filters for

the stereo loudspeaker setups. A significant boost in amplitude can be observed at low frequencies for the laptop setup - a common phenomenon in XTC systems using frontal loudspeakers with small spacing relative to the wavelength. This low frequency boost is avoided in the neckband setup.

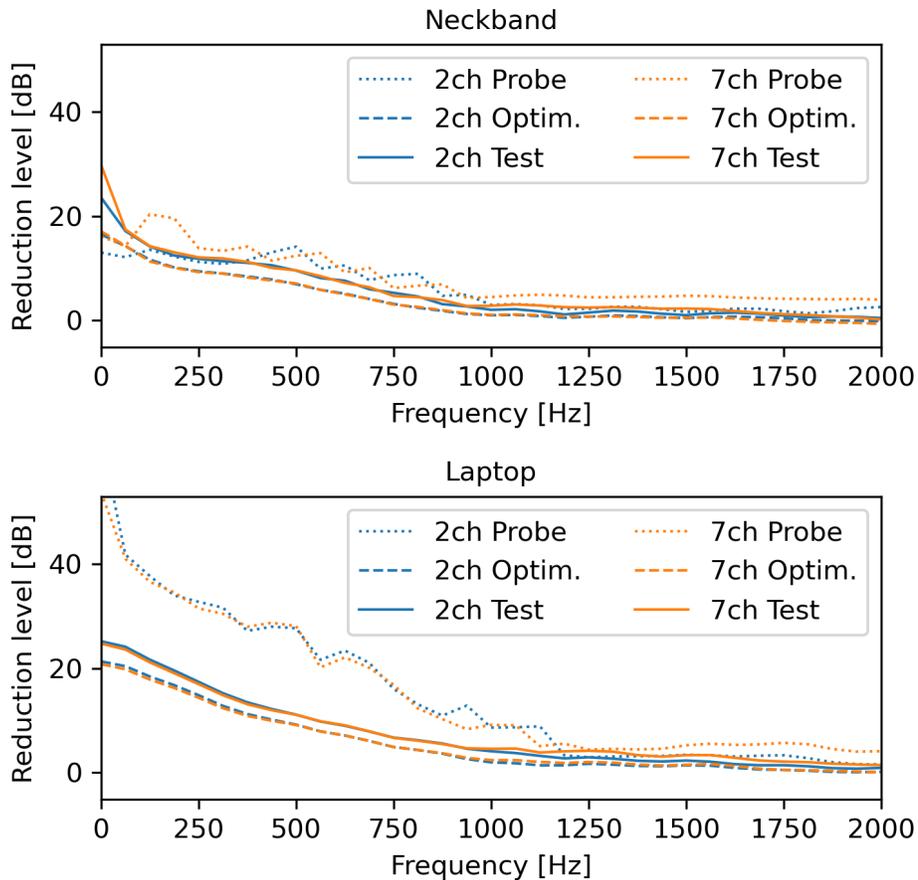


Figure 8.4: The reduction levels for the neckband system (top) and laptop system (bottom) using two or seven loudspeakers.

Fig. 8.6 illustrates the sound fields in the horizontal plane at the height of the listener’s head. The external, cancellation, and the residual fields at 500 Hz for the example problem with a single monopole noise source are shown. Both the neckband and laptop systems are seen to create silence zones in the vicinity of the test points. The laptop system needs to inject more anti-noise to do so, which results in the creation of

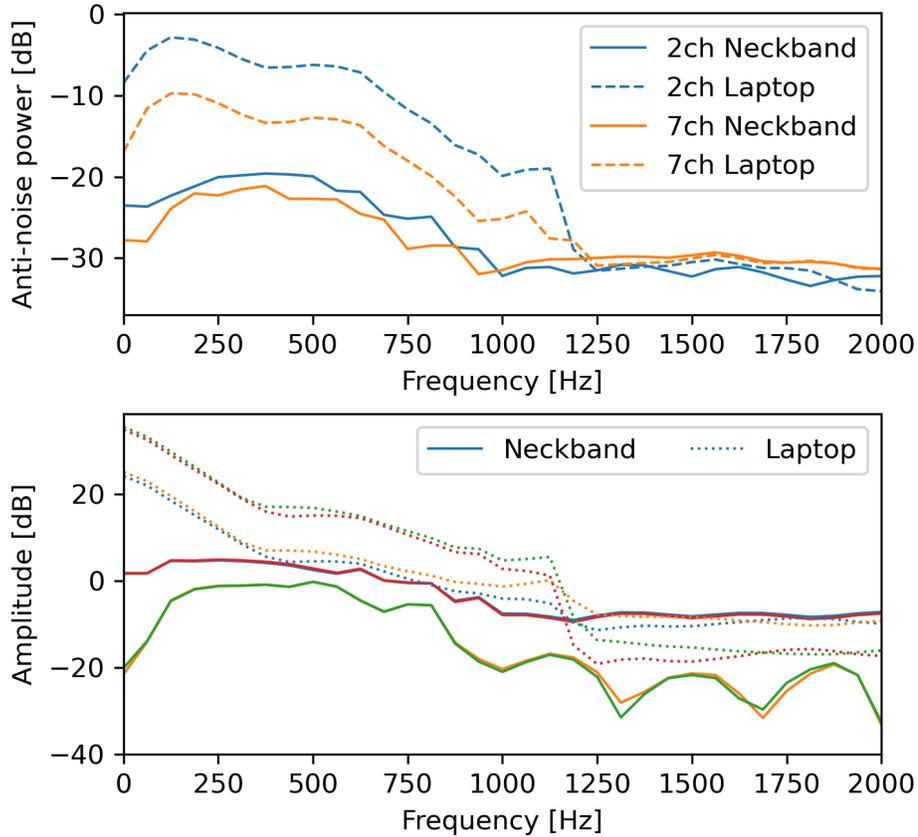


Figure 8.5: Top: the cancellation signal power spectra. Bottom: amplitudes of the cancellation filters $H^{(\text{can})}$ for the four paths connecting the 2ch probe microphones and the 2ch loudspeakers. The solid lines and dotted lines represent filter spectra for the neckband and laptop setup, respectively.

additional noisy zones, as seen in the residual field. Audio examples of the PASCAL simulations are available online.² Note that these examples are generated in an ideal setup for demonstration where the incident noise can be perfectly separated from the signal to be delivered to the listener and hence ANC is applied to the incident noise only.

8.5 Summary and related work

PASCAL, a framework for personal active sound field control on radiated, incident, or local personal sound fields, using a single setup of loudspeakers and microphones was

²<https://github.com/kaneko60/PASCAL-demos>

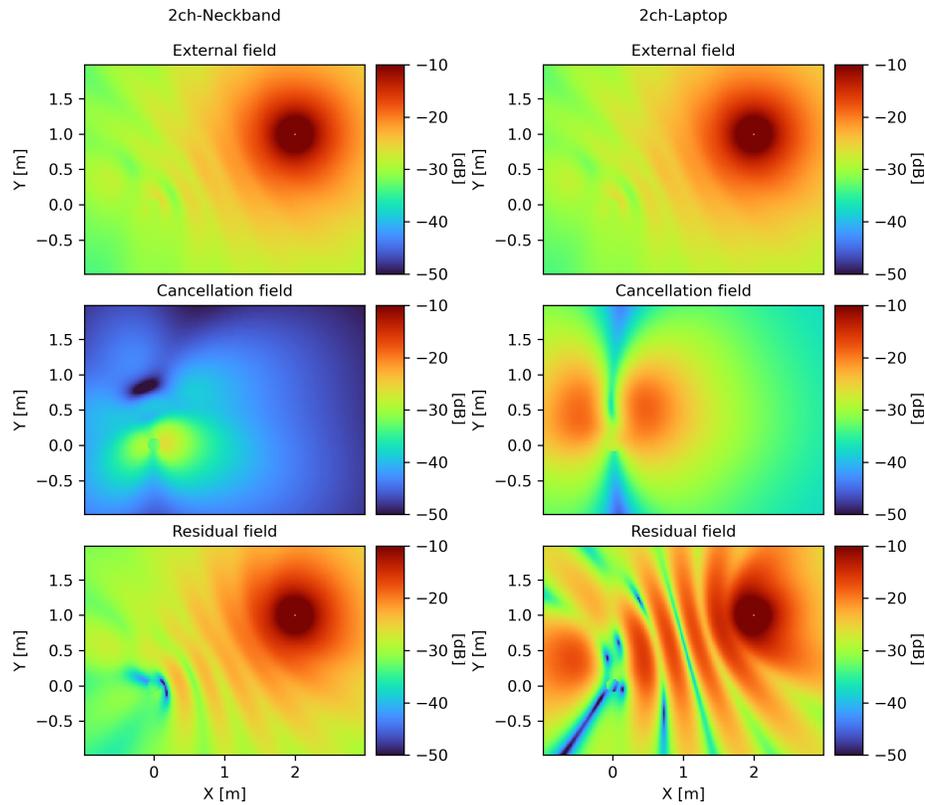


Figure 8.6: The external field (top row), cancellation field (middle row), and residual field (bottom row) at 500 Hz in the x - y plane on the height of the listener’s head for a monopole noise source located at (2m, 1m, 0) and the head located at the origin, for the stereo neckband (left column) and stereo laptop (right column), respectively.

presented. The framework relies on FMMBEM simulation for the TF preparation, which allows the use of a large number of test and optimization points which are utilized to design effective regularization. A system using wearable loudspeakers was simulated for the ARC and ANC tasks. A reduction level of 10-20 dB was observed at frequencies below 500 Hz. In the ANC task using a stereo neckband loudspeaker, the noise injected into the environment was found to be significantly less compared to the case of stereo loudspeakers of a laptop with comparable noise reduction level. A known issue of small loudspeakers used in real-world wearable audio devices is the limited producible frequency range which leaves active control of frequencies below 100 Hz challenging. Fortunately,

human speech has a higher frequency range and therefore this limitation may not be a serious problem in the ARC task.

Related systems have been extensively studied for ANC in vehicles [169] and ANC for headrests [170, 171]. These include time-domain variants e.g. the filtered-X LMS algorithm [172]. The utilization of these algorithms may be beneficial for future versions of the present work.

The performance of audio classification, source separation, and signal enhancement methods has significantly improved in the recent years [173–175]. As future developments, such methods could be utilized to realize PASCAL which recognizes different signal and noise types in the environment and selectively suppresses/enhances certain types of radiated/incident sounds.

Chapter 9: Recursive analytical quadrature for the close evaluation of Laplace and Helmholtz layer potentials over flat boundary elements in \mathbb{R}^3

A method for the analytical evaluation of singular and nearly singular layer potentials arising in the collocation boundary element method for the Laplace and Helmholtz equation is developed for flat boundary elements with polynomial shape functions. The method is based on dimension-reduction via the divergence theorem and a Recursive scheme for evaluating the resulting line Integrals for Polynomial Elements (RIPE). It is used to evaluate single layer, double layer, adjoint double layer, and hypersingular potentials, for both the Laplace and the Helmholtz kernels. It naturally supports nearly singular, singular, and hypersingular integrals under a single framework. The developed recursive algorithm allows accurate evaluation of layer potentials associated with $O(p^2)$ density functions used in a p -th order boundary element in $O(p^3)$ time for the Laplace case.¹

¹The research presented in this chapter is currently under review for publication. Its preprint can be found in [176].

9.1 Introduction

Accurate, efficient, and error-controlled numerical evaluation of layer potentials is needed to build the linear systems that must be solved in boundary element methods (BEM). This is nontrivial because the integrands are not always simple analytic functions but are singular or nearly singular. Fig. 9.1 illustrates the different types of element-level integrals that need to be evaluated in the fast multipole accelerated BEM. Quadrature

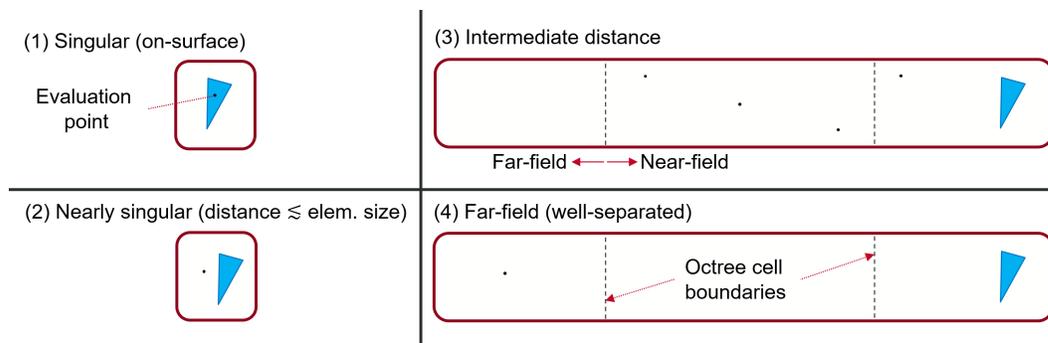


Figure 9.1: Types of layer potential integrals categorized by element-evaluation point distance. (1) singular or hypersingular case where the evaluation point is on the element surface, (2) nearly singular case where the evaluation point is off the element surface but too close for simple Gauss-Legendre quadrature being accurate, (3) intermediate distance where direct evaluation is still applied and Gauss-Legendre quadrature is effective, and (4) far-field, where far interactions are clustered using the fast multipole method. This paper provides efficient and accurate computation of integrals of types (1) and (2) for all four layer potential kernels (Eqs. (9.3) and (9.5)), for both the Laplace and Helmholtz kernels.

schemes which are effective for integrating polynomials of limited degree, e.g. Gauss-Legendre quadrature, are known to produce inaccurate results when the evaluation point is close to the element. Many techniques have been developed over the years to accurately evaluate boundary integrals in such nearly singular cases [1, 177–187]. Ref. [1] provides a recent extensive survey on this subject. The approaches developed include singularity cancellation using coordinate transforms [177, 179, 180], singularity subtraction [178],

continuation approach [188], dimension reduction [181–183], adaptive subdivision [184], and quadrature by expansion [185, 186, 186, 187]. The PART method [177] applies coordinate transforms which reduce the effect of the diverging integrand via the Jacobian of the transform. [180] introduced another coordinate transform approach, using sinh transforms. Singularity subtraction methods [178] split the integrand into a singular part and a regular part, where the former is evaluated analytically and the latter via Gauss-Legendre quadrature. Another approach for evaluating the layer potentials is to derive analytical expressions of the integrals. Newman developed a method to evaluate layer potentials for the Laplace kernel on quadrilateral elements for shape functions of arbitrary order [189]. Lenoir and Salles developed a semi-analytical method using dimension-reduction via the divergence theorem for constant and linear elements in Galerkin BEM [181, 190], where the layer potential associated with the singular part of the Helmholtz kernel is evaluated analytically, but the regular part is evaluated via conventional quadrature. Recently, Zhu and Veerapaneni [183] introduced a method for Laplace layer potentials on high-order curved elements using dimension reduction via Stokes’ theorem and quaternion algebra. As for the far-field case arising in the fast multipole method (FMM) [45] accelerated BEM, multipole expansions of layer potentials need to be evaluated where the integrands are the spherical basis functions and their normal derivatives instead of the Green functions themselves. The authors have recently developed efficient analytical quadrature schemes for these integrands for both constant [191] and high-order simplex elements [192].

In this paper, we propose a method to analytically evaluate all four layer potentials targeting singular and nearly singular cases arising in collocation BEM for the Helmholtz and Laplace kernels over flat boundary elements with high-order polynomial shape functions.

Flat boundary elements find important applications in problems where the boundary is inherently made of piecewise flat surfaces. This includes, but is not limited to, room acoustics [162]. In room acoustics, the boundary is made of millions or more flat boundary elements and hence efficient algorithms tailor-made for flat elements are highly desired. The proposed method is based on reduction of the surface integrals to line integrals via the divergence theorem in the element plane and efficient computation of the resulting line integrals using recursions which are derived via the introduction of auxiliary vector fields. In comparison to [181] we treat both the Laplace and Helmholtz kernels within a single framework, and are able to analytically evaluate integrals with general order polynomial shape functions via recursions. The method in [183] is most related to ours in that it is based on dimension reduction and supports polynomial elements. However our work differs from [183] in a number of ways: (1) it is analytical and solely uses recursions, in contrast to the use of Gauss-Legendre quadrature for the line integrals, (2) it is derived for both the Laplace and Helmholtz kernels and for all four basic layer potentials, i.e. single layer, double layer, adjoint double layer, and hypersingular potential, (3) it is restricted to flat elements by relying on the divergence theorem, whereas [183] is based on Stokes' theorem for manifolds and can handle curved elements, and (4) where [183] involves an $O(p^6)$ coordinate transform step, our method has an overall complexity of $O(p^3)$ for the Laplace case, with p the polynomial order. We remark that the divergence theorem in the flat element plane can be derived from Stokes' theorem, so the method could have been equivalently presented in terms of Stokes' theorem.

One of the benefits of our recursion-based approach is error control, as some of the recursions may be truncated as soon as a prescribed threshold is achieved, a feature

that is missing in fixed-order Gauss-Legendre quadrature. Error control is important when quadrature is part of a more complicated solver with multiple components, e.g. the FMM-BEM solved via an iterative method, where the numerical error arising from different parts should all be kept consistent to avoid redundant computation. While adaptive error control can be achieved by classical methods e.g. Gauss-Kronrod quadrature, such methods are not as efficient as our method, as seen in §5. The simplest case of our method for piecewise constant elements was initially developed by Gumerov and Duraiswami [193] and used in a production collocation BEM solver for the Helmholtz equation [162].

9.2 Boundary element method and layer potentials

The boundary element method is extensively used for numerical solution of partial differential equations, e.g. the Helmholtz equation and the Laplace equation, respectively given by

$$-k^2 u(\mathbf{r}) - \nabla^2 u(\mathbf{r}) = f(\mathbf{r}), \quad -\nabla^2 u(\mathbf{r}) = f(\mathbf{r}), \quad (9.1)$$

with wavenumber k , field u , and source f . The weak form of (11.1) can be written in terms of single- and double layer potentials V , K [194]:

$$\{(c_p \gamma_{0,p} + K \gamma_{0,q} - V \gamma_{1,q})u\}(\mathbf{r}_p) = \{N_0 f\}(\mathbf{r}_p), \quad (9.2)$$

with $c_p = 1/2$, γ_0 and γ_1 the boundary trace and normal derivative operators, N_0 the Newton potential operator, and layer potentials defined over the boundary Γ as:

$$\{V\psi\}(\mathbf{r}_p) \equiv \int_{\mathbf{r}_q \in \Gamma} G(\mathbf{r}_p, \mathbf{r}_q) \psi(\mathbf{r}_q) d\Gamma, \quad \{K\phi\}(\mathbf{r}_p) \equiv \int_{\mathbf{r}_q \in \Gamma} \frac{\partial G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_q} \phi(\mathbf{r}_q) d\Gamma, \quad (9.3)$$

where $G(\mathbf{r}_p, \mathbf{r}_q)$ is the Helmholtz or Laplace Green function:

$$G_H(\mathbf{r}_p, \mathbf{r}_q) = \frac{e^{ikr}}{4\pi r}, \quad G_L(\mathbf{r}_p, \mathbf{r}_q) = \frac{1}{4\pi r}, \quad r \equiv |\mathbf{r}_q - \mathbf{r}_p|. \quad (9.4)$$

Some BEM formulations [195, 196] use the normal derivative form of the boundary integral equation with the adjoint double layer potential K' and hypersingular potential D .

$$\{K'\psi\}(\mathbf{r}_p) \equiv \int_{\mathbf{r}_q \in \Gamma} \frac{\partial G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_p} \psi(\mathbf{r}_q) d\Gamma, \quad \{D\phi\}(\mathbf{r}_p) \equiv \int_{\mathbf{r}_q \in \Gamma} \frac{-\partial^2 G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_p \partial \mathbf{n}_q} \phi(\mathbf{r}_q) d\Gamma. \quad (9.5)$$

In the collocation BEM the boundary Γ is discretized into elements, typically triangular, and the layer potential integrals over these elements are evaluated. The densities ψ , ϕ are approximated via local polynomial functions (also called *shape functions*).

9.3 Problem statement

Assume a flat triangle element S with vertices \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 in \mathbb{R}^3 , where the *element coordinate frame* $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ has its origin in the element plane, $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ are unit

vectors lying in the element plane, and $\hat{\mathbf{k}} = \hat{\mathbf{i}} \times \hat{\mathbf{j}} = \mathbf{n}_{\text{face}}$, where \mathbf{n}_{face} is the unit normal to the element. We also define the *edge coordinate frame* $(\hat{\mathbf{i}}^j, \hat{\mathbf{j}}^j, \hat{\mathbf{k}}^j)$ for edge j as: $\hat{\mathbf{i}}^j = (\mathbf{v}_{(j+1)\%3} - \mathbf{v}_j) / |\mathbf{v}_{(j+1)\%3} - \mathbf{v}_j|$, $\hat{\mathbf{k}}^j = \mathbf{n}_{\text{face}}$, and $\hat{\mathbf{j}}^j = \hat{\mathbf{k}}^j \times \hat{\mathbf{i}}^j \equiv -\mathbf{n}^j$. $\mathbf{n}^1, \mathbf{n}^2$, and \mathbf{n}^3 are the unit outward edge normal vectors in the element plane for edges $(\mathbf{v}_1, \mathbf{v}_2)$, $(\mathbf{v}_2, \mathbf{v}_3)$, and $(\mathbf{v}_3, \mathbf{v}_1)$, respectively. The coordinate frames associated with the boundary element are shown in Fig. 10.2. In the following, the superscript j refers to the edge index and is not an exponent. \mathbf{r}_p and $\check{\mathbf{r}}_p$ are respectively defined as the observation point and its orthogonal projection on the element plane. The points \mathbf{r}_q and \mathbf{r}_p are represented as $\mathbf{r}_q = (x, y, z)^T$ and $\mathbf{r}_p = (x_p, y_p, z_p)^T$ in the element frame. Also define $r \equiv |\mathbf{r}_q - \mathbf{r}_p|$, $\boldsymbol{\rho} \equiv \mathbf{r}_q - \check{\mathbf{r}}_p$, $\rho \equiv |\boldsymbol{\rho}|$, $h \equiv \hat{\mathbf{k}} \cdot (\mathbf{r}_p - \check{\mathbf{r}}_p)$, $x_d \equiv x - x_p$, $y_d \equiv y - y_p$, and $z_d \equiv z - z_p$.

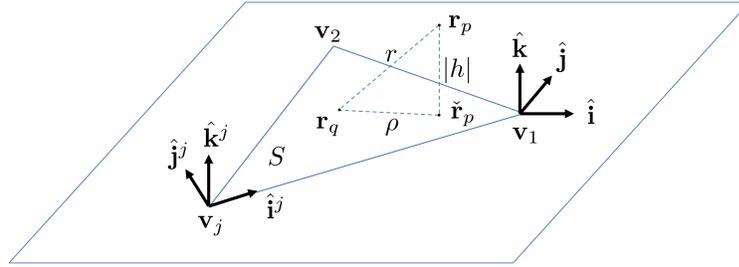


Figure 9.2: A boundary element and its associated coordinate frames. We have the *element coordinate frame* $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ and *edge coordinate frame* $(\hat{\mathbf{i}}^j, \hat{\mathbf{j}}^j, \hat{\mathbf{k}}^j)$ for each edge with index j . The origin of the element frame is set to \mathbf{v}_1 .

In this setup, a point in space can be expressed as:

$$\mathbf{r} = \mathbf{v}_1 + x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} = \mathbf{v}_1 + u(\mathbf{v}_2 - \mathbf{v}_1) + v(\mathbf{v}_3 - \mathbf{v}_1) + w\mathbf{n}_{\text{face}}, \quad (9.6)$$

where (u, v) are the coordinates in the two-dimensional reference frame with $0 \leq u, v \leq 1$ and $u + v \leq 1$, and (x, y, z) are the coordinates in the element frame. Projection from

the reference frame to the element frame is a linear transform:

$$(x, y)^T = \Xi(u, v)^T, \quad (9.7)$$

with Ξ the (2×2) transformation matrix. Points in the element frame are expressed by their coordinates in the edge frame as:

$$(x, y, z)^T = B(x^j, y^j, z^j)^T + \mathbf{b}, \quad B \equiv \begin{pmatrix} \alpha_x & \beta_x & 0 \\ \alpha_y & \beta_y & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{b} \equiv \begin{pmatrix} \gamma_x \\ \gamma_y \\ 0 \end{pmatrix}. \quad (9.8)$$

Identifying the entries of Ξ , B , and \mathbf{b} for a given element is straightforward.

Shape functions $N(u, v)$ for p_s -th order elements have the following form:

$$N_{p_s}(u, v) = \sum_{b'=0}^{p_s} \sum_{c'=0}^{p_s-b'} a_{b',c'} u^{b'} v^{c'} = \sum_{b=0}^{p_s} \sum_{c=0}^{p_s-b} A_{b,c} x^b y^c. \quad (9.9)$$

Shape functions are specified by the $a_{b',c'}$ coefficients in the (u, v) reference frame. The coefficients $A_{b,c}$ for the element frame expression can be obtained by substituting (9.7) into (9.9) and solving a linear system. Since the coefficients $a_{b',c'}$ and $A_{b,c}$ only interact if $b + c = b' + c'$ due to (9.7), this can be achieved by solving $p_s + 1$ small linear systems, and each of these can be solved via efficient stabilized algorithms for Vandermonde systems [197], resulting in a complexity of $O(p_s^3)$ for this stage.

9.4 Method: Recursive Integrals for Polynomial Elements (RIPE)

9.4.1 Dimension reduction via divergence theorem

We assume flat triangular elements and reduce the surface integral to a contour integral via the divergence theorem, and discretize the contour integral into a sum of line integrals over the straight edges. We seek vector fields \mathbf{m}_i whose surface divergence in the element plane gives the integrand in question, i.e. a polynomial shape function multiplied with the Green function:

$$\sum_i A_i \nabla_s \cdot \mathbf{m}_i = N(u(x, y), v(x, y)) G(\mathbf{r}_p, \mathbf{r}_q), \quad (9.10)$$

then we can apply the divergence theorem for the evaluation of the element-wise layer potentials of interest. This is depicted in Fig. 9.3. The single layer potential V_e , double

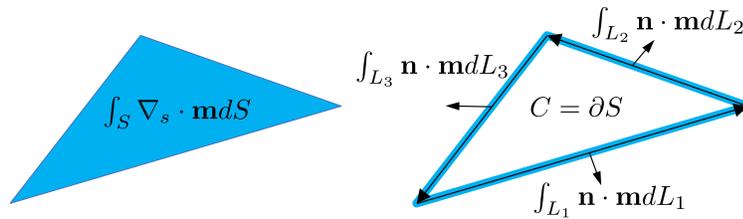


Figure 9.3: Left: the surface integral over element S of a function which can be written as the surface divergence of a vector field \mathbf{m} . Right: after the application of the divergence theorem, the surface integral over S is converted into a contour integral over the contour $C = \partial S$ and further into line integrals due to the straightness of the edges.

layer potential K_e , adjoint double layer potential K'_e , and hypersingular potential D_e can

be all then written as:

$$\begin{aligned}
V_e &= \int_{S_q} G(\mathbf{r}_p, \mathbf{r}_q) N(\mathbf{r}_q) dS_q = \int_{S_q} \sum_i A_i \nabla_s \cdot \mathbf{m}_i dS_q = \sum_i A_i \int_{S_q} \nabla_s \cdot \mathbf{m}_i dS_q \\
&= \sum_i A_i \left(\oint_C \mathbf{n}_C \cdot \mathbf{m}_i dC + T_{\text{sing}}^{(i)} \right) = \sum_i A_i \left(\sum_j \int_{L_j} \mathbf{n}_j \cdot \mathbf{m}_i dL_j + T_{\text{sing}}^{(i)} \right), \\
K_e &= \int_{S_q} \frac{\partial G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_q} N(\mathbf{r}_q) dS_q = \sum_i A_i \left(\sum_j \frac{\partial}{\partial z} \int_{L_j} \mathbf{n}_j \cdot \mathbf{m}_i dL_j + \frac{\partial T_{\text{sing}}^{(i)}}{\partial z} \right), \\
K'_e &= \int_{S_q} \frac{\partial G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_p} N(\mathbf{r}_q) dS_q = \sum_i A_i \left(\sum_j \frac{\partial}{\partial \mathbf{n}_p} \int_{L_j} \mathbf{n}_j \cdot \mathbf{m}_i dL_j + \frac{\partial T_{\text{sing}}^{(i)}}{\partial \mathbf{n}_p} \right), \\
-D_e &= \int_{S_q} \frac{\partial^2 G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_p \partial \mathbf{n}_q} N(\mathbf{r}_q) dS_q = \sum_i A_i \left(\sum_j \frac{\partial^2}{\partial \mathbf{n}_p \partial z} \int_{L_j} \mathbf{n}_j \cdot \mathbf{m}_i dL_j + \frac{\partial^2 T_{\text{sing}}^{(i)}}{\partial \mathbf{n}_p \partial z} \right),
\end{aligned} \tag{9.11}$$

with $\nabla_s = \nabla - \hat{\mathbf{k}}(\hat{\mathbf{k}} \cdot \nabla)$ the surface divergence operator in the element plane and $C = \partial S_q$ the contour of the element surface S_q . Note that we have to add the singularity term $T_{\text{sing}}^{(i)}$ associated with \mathbf{m}_i to the contour integral if $\check{\mathbf{r}}_p$, the projection of \mathbf{r}_p onto the element plane, falls onto the element and if the integrand $\mathbf{n}_C \cdot \mathbf{m}_i$ has a singularity at $\mathbf{r}_q = \check{\mathbf{r}}_p$ on the element. In this case, the singularity has to be removed when applying the divergence theorem, i.e. the contour integral over an infinitesimal circle C_ε centered at the singularity on the element has to be subtracted from the contour integral over S : $T_{\text{sing}}^{(i)} \equiv \lim_{\varepsilon \rightarrow 0} \oint_{C_\varepsilon} \mathbf{n}_{C_\varepsilon} \cdot \mathbf{m}_i dC_\varepsilon$. Thus, the problem reduces to the evaluation of line integrals $\int_{L_j} \mathbf{n}_j \cdot \mathbf{m}_i dL_j$ associated with each edge and singularity terms associated with the element. One way to evaluate these line integrals is to rely on Gauss-Legendre quadrature, resulting in semi-analytical methods [183]. If these line integrals can be computed analytically, they can be further reduced to the evaluation at the end-points of the edges due to the fundamental theorem of calculus. Fig. 9.4 depicts the contrast of

the conventional approach of surface integrals based on Gauss-Legendre quadrature and the proposed RIPE method.

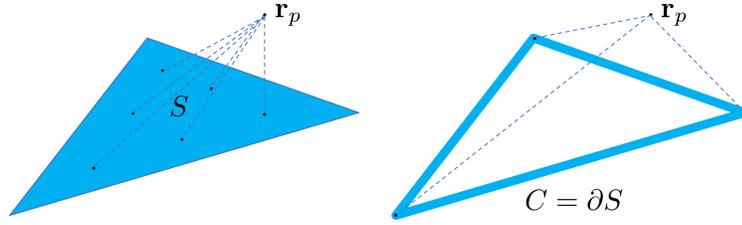


Figure 9.4: Conventional Gauss-Legendre quadrature requires evaluation of the integrand at multiple nodes on the element S , which may be near or coincide with the evaluation point, or collocation node, \mathbf{r}_p . After application of the divergence theorem, the surface integral over S is converted into an integral over the contour $C = \partial S$ and further into line integrals, which in the end results in the evaluation of analytical expressions at the endpoints of the edges, i.e. at the vertices.

To execute this strategy we need to find expressions for the vector fields \mathbf{m}_i , whose surface divergence equals the integrand $G(\mathbf{r}_p, \mathbf{r}_q)N(\mathbf{r}_q)$, or, at least, the terms $\int_{L_j} \mathbf{n}_j \cdot \mathbf{m}_i dL_j$ and $T_{\text{sing}}^{(i)}$. While the proposed method can be naturally generalized to polygonal elements, here we only consider triangular elements for their simplicity and efficiency. In the rest of the paper, $\partial f / \partial x = \partial_x f = f^{(x)}$ all denote partial derivatives of a function f with respect to variable x .

9.4.2 Laplace kernel with polynomials of arbitrary order

9.4.2.1 Single layer potential

Our aim is to find vector fields whose surface divergence equals the integrand in question, i.e., a polynomial of x and y with degree up to p_s multiplied with the Green function:

$$\nabla_s \cdot \mathbf{m}_{b,c} = \frac{x^b y^c}{4\pi r}, \quad b, c \geq 0, \quad b + c \leq p_s. \quad (9.12)$$

We introduce auxiliary vector fields $\mathbf{m}_{b,c,a}$ which satisfy the following relation:

$$\nabla_s \cdot \mathbf{m}_{b,c,a} = \frac{x^b y^c r^a}{4\pi r}, \quad a, b, c \geq 0, \quad a + b + c \leq p_s. \quad (9.13)$$

Lemma 1 (Recurrence for auxiliary vector fields - Laplace kernel). *If fields $\mathbf{m}_{b-1,c,a+1}$ and $\mathbf{m}_{b,c,a-1}$ satisfy (9.13), then the field $\mathbf{m}_{b+1,c,a-1}$ given by:*

$$a\mathbf{m}_{b+1,c,a-1} = -b\mathbf{m}_{b-1,c,a+1} + ax_p \mathbf{m}_{b,c,a-1} + \left(\frac{x^b y^c r^a}{4\pi} \hat{\mathbf{i}} \right), \quad a \geq 1, \quad (9.14)$$

also satisfies (9.13) for the index tuple $(b + 1, c, a - 1)$.

Proof. If vector fields $\mathbf{m}_{b-1,c,a+1}$ and $\mathbf{m}_{b,c,a-1}$ satisfy (9.13), the surface divergence of the right hand side of (9.14) yields:

$$\begin{aligned} & -b\nabla_s \cdot \mathbf{m}_{b-1,c,a+1} + ax_p \nabla_s \cdot \mathbf{m}_{b,c,a-1} + \nabla_s \cdot \left(\frac{x^b y^c r^a}{4\pi} \hat{\mathbf{i}} \right) \\ &= -b \frac{x^{b-1} y^c r^a}{4\pi r} + ax_p \frac{x^b y^c r^{a-1}}{4\pi r} + y^c \frac{ax^{b+1} r^{a-1} - ax_p x^b r^{a-1} + bx^{b-1} r^{a+1}}{4\pi r} = a \frac{x^{b+1} y^c r^{a-1}}{4\pi r}, \end{aligned} \quad (9.15)$$

hence, $\nabla_s \cdot \mathbf{m}_{b+1,c,a-1} = x^{b+1} y^c r^{a-1} / (4\pi r)$, which means that $\mathbf{m}_{b+1,c,a-1}$ satisfies (9.13)

for the index tuple $(b + 1, c, a - 1)$. □

Let us define:

$$\xi_{b,c,a} \equiv \int \mathbf{m}_{b,c,a} \cdot \mathbf{n}^j dx^j, \quad i_m \equiv \int r^m dx^j, \quad k_m \equiv -y_p^j \int \frac{r^m}{\rho^2} dx^j. \quad (9.16)$$

From (9.14) it follows:

$$\xi_{b+1,c,a} = -\frac{b}{a+1}\xi_{b-1,c,a+2} + x_p \xi_{b,c,a} + \frac{\hat{\mathbf{i}} \cdot \mathbf{n}^j}{(a+1)4\pi} \int x^b y^c r^{a+1} dx^j, \quad (9.17)$$

and by symmetry, the following recurrence also holds:

$$\xi_{b,c+1,a} = -\frac{c}{a+1}\xi_{b,c-1,a+2} + y_p \xi_{b,c,a} + \frac{\hat{\mathbf{j}} \cdot \mathbf{n}^j}{(a+1)4\pi} \int x^b y^c r^{a+1} dx^j. \quad (9.18)$$

These recursions hold for $b \geq 0$ and $c \geq 0$. For $b = c = 0$, the following field $\mathbf{m}_{0,0,a}$:

$$\mathbf{m}_{0,0,a} = \frac{r^{a+1}}{(a+1)4\pi\rho^2} \boldsymbol{\rho}, \quad (9.19)$$

satisfies (9.13) and can be used to initiate the recursion. This yields:

$$\xi_{0,0,a} = \int \frac{r^{a+1}}{(a+1)4\pi\rho^2} \boldsymbol{\rho} \cdot \mathbf{n}^j dx^j = \frac{y_p^j}{(a+1)4\pi} \int \frac{r^{a+1}}{\rho^2} dx^j = \frac{-k_{a+1}}{(a+1)4\pi}. \quad (9.20)$$

The integral k_m and i_m can be computed efficiently via the following recurrence:

$$k_{m+2} = (z_p^j)^2 k_m - y_p^j i_m, \quad i_{m+2} = ((y_p^j)^2 + (z_p^j)^2) \frac{m+2}{m+3} i_m + \frac{x_d^j r^{m+2}}{m+3}, \quad (9.21)$$

with closed form initial terms in appendix B. The following are used later:

$$\begin{aligned} \partial_{x_p^j} k_l &= y_p^j \frac{r^l}{\rho^2}, & \partial_{y_p^j} k_l &= x_d^j \frac{r^l}{\rho^2} - l i_{l-2}, & \partial_{z_p^j} k_l &= -l z_d^j k_{l-2}, \\ \partial_{x_p^j} i_l &= -r^l, & \partial_{y_p^j} i_l &= l y_p^j i_{l-2}, & \partial_{z_p^j} i_l &= -l z_d^j i_{l-2}. \end{aligned} \quad (9.22)$$

It remains to compute the last integral in (9.17) and (9.18) with form $\int x^b y^c r^a dx^j$. Recall $x = \alpha_x x^j + \beta_x y^j + \gamma_x$ and $y = \alpha_y x^j + \beta_y y^j + \gamma_y$. Also, $y^j = z^j = 0$ for points on edge j . Following a coordinate transform from x^j to x_d^j this can be written as:

$$\begin{aligned} \kappa_{b,c,a} &\equiv \int x^b y^c r^a dx^j = \int (\alpha_x x_d^j - \hat{\alpha})^b (\alpha_y x_d^j - \hat{\beta})^c r^a dx_d^j, \\ \hat{\alpha} &\equiv -(\beta_x y^j + \gamma_x) - \alpha_x x_p^j, \quad \hat{\beta} \equiv -(\beta_y y^j + \gamma_y) - \alpha_y x_p^j, \quad \gamma^2 \equiv r^2 - (x_d^j)^2. \end{aligned} \quad (9.23)$$

Later we use:

$$\hat{\alpha}^{(x_p^j)} = -\alpha_x, \quad \hat{\beta}^{(x_p^j)} = -\alpha_y, \quad \hat{\alpha}^{(y_p^j)} = \hat{\alpha}^{(z_p^j)} = \hat{\beta}^{(y_p^j)} = \hat{\beta}^{(z_p^j)} = 0. \quad (9.24)$$

The function $\kappa_{b,c,a}$ satisfies the following recurrence relations:

$$\begin{aligned} \kappa_{b+1,c+1,a} &= \alpha_x \alpha_y \kappa_{b,c,a+2} - \hat{\beta} \kappa_{b+1,c,a} - \hat{\alpha} \kappa_{b,c+1,a} - \left(\hat{\alpha} \hat{\beta} + \alpha_x \alpha_y \gamma^2 \right) \kappa_{b,c,a}, \\ \kappa_{b+2,c,a} &= \alpha_x^2 \kappa_{b,c,a+2} - 2\hat{\alpha} \kappa_{b+1,c,a} - \left(\hat{\alpha}^2 + \alpha_x^2 \gamma^2 \right) \kappa_{b,c,a}, \\ \kappa_{b,c+2,a} &= \alpha_y^2 \kappa_{b,c,a+2} - 2\hat{\beta} \kappa_{b,c+1,a} - \left(\hat{\beta}^2 + \alpha_y^2 \gamma^2 \right) \kappa_{b,c,a}. \end{aligned} \quad (9.25)$$

The following special cases are used to initiate the recursion:

$$\begin{aligned} \kappa_{0,0,a} &= i_a, \quad \kappa_{1,0,a} = \alpha_x \int x_d^j r^a dx_d^j - \hat{\alpha} i_a = \alpha_x \frac{r^{a+2}}{a+2} - \hat{\alpha} \kappa_{0,0,a}, \\ \kappa_{0,1,a} &= \alpha_y \int x_d^j r^a dx_d^j - \hat{\beta} i_a = \alpha_y \frac{r^{a+2}}{a+2} - \hat{\beta} \kappa_{0,0,a}. \end{aligned} \quad (9.26)$$

Later, we will use the following relations:

$$\begin{aligned}
\partial_{x_p^j} \kappa_{b,c,a} &= \partial_{x_p^j} \int x^b y^c r^a dx^j = \int x^b y^c \partial_{x_p^j} (r^a) dx^j = - \int x^b y^c \partial_{x^j} (r^a) dx^j \\
&= -x^b y^c r^a + \alpha_x b \kappa_{b-1,c,a} + \alpha_y c \kappa_{b,c-1,a}, \\
\partial_{y_p^j} \kappa_{b,c,a} &= -a y_d \kappa_{b,c,a-2}, \quad \partial_z \kappa_{b,c,a} = a z_d \kappa_{b,c,a-2}, \\
\partial_{z,x_p^j}^2 \kappa_{b,c,a} &= -x^b y^c a z_d r^{a-2} + \alpha_x b \kappa_{b-1,c,a}^{(z)} + \alpha_y c \kappa_{b,c-1,a}^{(z)} \\
&= a z_d (-x^b y^c r^{a-2} + \alpha_x b \kappa_{b-1,c,a-2} + \alpha_y c \kappa_{b,c-1,a-2}).
\end{aligned} \tag{9.27}$$

Note that we also need to compute the singularity term associated with $\mathbf{m}_{b,c,a}$:

$$T_{b,c,a} \equiv \lim_{\varepsilon \rightarrow 0} \oint_{C_\varepsilon} \mathbf{m}_{b,c,a} \cdot \mathbf{n}_{C_\varepsilon} dC_\varepsilon, \tag{9.28}$$

if $\check{\mathbf{r}}_p$ falls onto the element and if the integrand has a singularity on the element, with C_ε a circle of radius ε centered at the singularity and $\mathbf{n}_{C_\varepsilon}$ the inward unit normal vector along the circle. This singularity term also obeys the recursions:

$$T_{b+1,c,a} = \frac{-b}{a+1} T_{b-1,c,a+2} + x_p T_{b,c,a}, \quad T_{b,c+1,a} = \frac{-c}{a+1} T_{b,c-1,a+2} + y_p T_{b,c,a}, \tag{9.29}$$

and, if the fields (9.19) are used for $b = c = 0$, the following are initial values:

$$T_{0,0,a} = \lim_{\varepsilon \rightarrow 0} \oint_{C_\varepsilon} \mathbf{m}_{0,0,a} \cdot \mathbf{n}_{C_\varepsilon} dC_\varepsilon = \lim_{\varepsilon \rightarrow 0} \oint_{C_\varepsilon} \frac{r^{a+1} \boldsymbol{\rho} \cdot \mathbf{n}_{C_\varepsilon}}{(a+1)4\pi\rho^2} dC_\varepsilon = -\frac{|h|^{a+1}}{2(a+1)}. \tag{9.30}$$

Thus we can evaluate all terms needed for the single layer potential with an arbitrary polynomial shape function of order p_s by running the above recursions. The procedures

for computing $\kappa_{b,c,a}$ and $\xi_{b,c,a}$ are given in algorithm 1 and algorithm 2. The values $\xi_{b,c,0}$ are the line integrals needed to compute the layer potential. The steps for computing the $\xi_{b,c,a}$ coefficients using algorithm 2 are illustrated in Fig. 9.5 in the (b, c, a) index space for the Laplace single layer potential. By taking derivatives and second derivatives with respect to z and/or \mathbf{r}_p , similar recurrences are derived for the double layer, adjoint double layer, and hypersingular potentials. The same scheme applies for the evaluation of these layer potentials, also for the Helmholtz kernel as shown later.

Algorithm 1 Compute $\kappa_{b,c,a}$ for $b + c + a \leq p_s$

1. Compute all $\kappa_{0,0,a}$ terms by running the recursion (9.26) on the a -axis.
 2. Compute all $\kappa_{0,1,a}$ terms using (9.26).
 3. Compute all $\kappa_{0,c,a}$ terms by running recursion (9.25) on the $b = 0$ plane for $c = 2, 3, \dots, p_s$.
 4. Compute all $\kappa_{1,0,a}$ terms using (9.26).
 5. Compute all $\kappa_{b,0,a}$ terms by running recursion (9.25) on the $c = 0$ plane for $b = 2, 3, \dots, p_s$.
 6. Compute all remaining $\kappa_{b+1,c+1,a}$ terms for each c -plane with $c = 0, 1, \dots, p_s - 1$, for $b = 0, 1, \dots, p_s - c - 1$, using recursion (9.25).
-

Algorithm 2 Compute $\xi_{b,c,a}$ for $b + c + a \leq p_s$

1. Compute the seed term $\xi_{0,0,0}$ using the expression (9.20).
 2. Compute all $\xi_{0,0,a}$ terms along the a -axis using the expression (9.20).
 3. Compute all $\xi_{0,c,a}$ terms in the $b = 0$ plane for $c = 1, 2, \dots, p_s$, using (9.18).
 4. Compute all remaining $\xi_{b,c,a}$ terms for each c -plane with $c = 0, 1, \dots, p_s$, for $b = 1, 2, \dots, p_s - c$, using recursion (9.17).
-

Here is our main theorem:

Theorem 2 (Layer potential integral by recursion via auxiliary vector fields - Laplace kernel). *algorithm 1 and algorithm 2 allow evaluation of the indefinite integral $\xi_{b,c,0} = \int \mathbf{m}_{b,c} \cdot \mathbf{n}^j dx^j$ needed for the computation of the single layer potential weighted by shape functions $x^b y^c$ based on (9.11).*

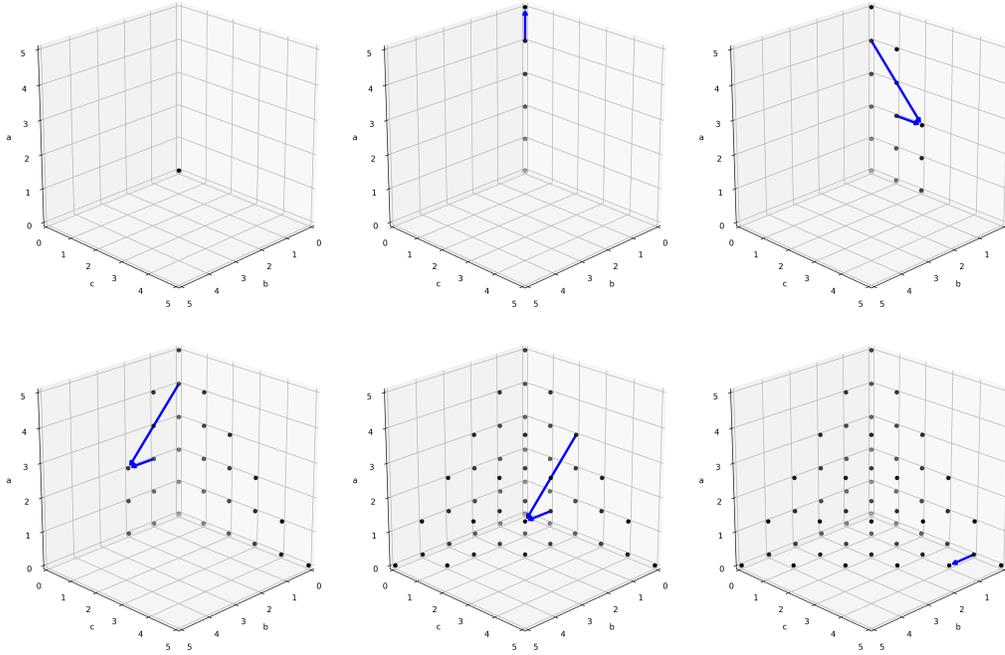


Figure 9.5: Computing $\xi_{b,c,a}$ using algorithm 2. The blue arrows indicate terms in the recurrence relations used in each step, here for the Laplace single layer potential with $p_s = 5$. Step 1 (top left): compute seed term $\xi_{0,0,0}$, step 2 (top center): compute $\xi_{0,0,a}$ using recurrences along the a -axis, step 3 (top right): compute $\xi_{0,c,a}$ on the $b = 0$ plane using recurrences in the c -directions, step 4-6 (bottom left to right): compute $\xi_{b,c,a}$ for $b > 0$ using recurrences in the b -direction.

Proof. The theorem follows from the initial conditions (9.26), (9.20), the recurrence relations (9.25), (9.17) and (9.18) due to lemma 1, and induction. □

Alternatively, the recursions can be executed such that $\xi_{b,c,0}$ are obtained in ascending order of $b + c$. This may be useful if application using RIPE has a mechanism to evaluate and accept results with dynamically truncated orders.

9.4.2.2 Double layer potential

By taking the derivative of (9.17) and (9.18) with respect to z , we obtain recursions for computing the double layer potential:

$$\begin{aligned}\xi_{b+1,c,a}^{(z)} &= \frac{-b}{a+1}\xi_{b-1,c,a+2}^{(z)} + x_p \xi_{b,c,a}^{(z)} + \frac{\hat{\mathbf{i}} \cdot \mathbf{n}^j}{4\pi} z_d \int x^b y^c r^{a-1} dx^j, \\ \xi_{b,c+1,a}^{(z)} &= \frac{-c}{a+1}\xi_{b,c-1,a+2}^{(z)} + y_p \xi_{b,c,a}^{(z)} + \frac{\hat{\mathbf{j}} \cdot \mathbf{n}^j}{4\pi} z_d \int x^b y^c r^{a-1} dx^j, \quad \xi_{0,0,a}^{(z)} = \frac{z_p^j k_{a-1}}{4\pi},\end{aligned}\tag{9.31}$$

with $\xi_{b,c,a}^{(z)} \equiv \partial_z \xi_{b,c,a}$. For the singularity terms $T_{b,c,a}^{(z)} \equiv \partial_z T_{b,c,a}$, from (9.29) we have:

$$\begin{aligned}T_{b+1,c,a}^{(z)} &= x_p T_{b,c,a}^{(z)} - \frac{b}{a+1} T_{b-1,c,a+2}^{(z)}, \quad T_{b,c+1,a}^{(z)} = y_p T_{b,c,a}^{(z)} - \frac{c}{a+1} T_{b,c-1,a+2}^{(z)}, \\ T_{0,0,a}^{(z)} &= \frac{z_p^j |h|^{a-1}}{2} \text{ if } |h| \neq 0 \text{ else } 0.\end{aligned}\tag{9.32}$$

9.4.2.3 Adjoint double layer potential recursions

$$\xi_{b+1,c,a}^{(x_p^j)} = -\frac{b}{a+1}\xi_{b-1,c,a+2}^{(x_p^j)} + \alpha_x \xi_{b,c,a} + x_p \xi_{b,c,a}^{(x_p^j)} + \frac{\hat{\mathbf{i}} \cdot \mathbf{n}^j}{(a+1)4\pi} \partial_{x_p^j} \kappa_{b,c,a+1},\tag{9.33}$$

$$\xi_{b+1,c,a}^{(y_p^j)} = -\frac{b}{a+1}\xi_{b-1,c,a+2}^{(y_p^j)} + \beta_x \xi_{b,c,a} + x_p \xi_{b,c,a}^{(y_p^j)} - \frac{\hat{\mathbf{i}} \cdot \mathbf{n}^j}{4\pi} y_d^j \kappa_{b,c,a-1},$$

$$\xi_{b,c+1,a}^{(x_p^j)} = -\frac{c}{a+1}\xi_{b,c-1,a+2}^{(x_p^j)} + \alpha_y \xi_{b,c,a} + y_p \xi_{b,c,a}^{(x_p^j)} + \frac{\hat{\mathbf{j}} \cdot \mathbf{n}^j}{(a+1)4\pi} \partial_{x_p^j} \kappa_{b,c,a+1},\tag{9.34}$$

$$\xi_{b,c+1,a}^{(y_p^j)} = -\frac{c}{a+1}\xi_{b,c-1,a+2}^{(y_p^j)} + \beta_y \xi_{b,c,a} + y_p \xi_{b,c,a}^{(y_p^j)} - \frac{\hat{\mathbf{j}} \cdot \mathbf{n}^j}{4\pi} y_d^j \kappa_{b,c,a-1},$$

$$\xi_{0,0,a}^{(x_p^j)} = \frac{-y_p^j}{(a+1)4\pi} \frac{r^{a+1}}{\rho^2}, \quad \xi_{0,0,a}^{(y_p^j)} = \frac{1}{4\pi} \left(i_{a-1} - \frac{x_d^j r^{a+1}}{(a+1)\rho^2} \right), \quad \xi_{b,c,a}^{(z_p^j)} = -\xi_{b,c,a}^{(z)},\tag{9.35}$$

with $\xi_{b,c,a}^{(x_p^j)} \equiv \partial_{x_p^j} \xi_{b,c,a}$ and $\xi_{b,c,a}^{(y_p^j)} \equiv \partial_{y_p^j} \xi_{b,c,a}$. For the singularity terms $T_{b,c,a}^{(x_p)} \equiv \partial_{x_p} T_{b,c,a}$ and $T_{b,c,a}^{(y_p)} \equiv \partial_{y_p} T_{b,c,a}$ we have:

$$T_{b+1,c,a}^{(x_p)} = x_p T_{b,c,a}^{(x_p)} - \frac{b}{a+1} T_{b-1,c,a+2}^{(x_p)} + T_{b,c,a}, \quad T_{b+1,c,a}^{(y_p)} = x_p T_{b,c,a}^{(y_p)} - \frac{b}{a+1} T_{b-1,c,a+2}^{(y_p)}, \quad (9.36)$$

$$T_{b,c+1,a}^{(x_p)} = y_p T_{b,c,a}^{(x_p)} - \frac{c}{a+1} T_{b,c-1,a+2}^{(x_p)}, \quad T_{b,c+1,a}^{(y_p)} = y_p T_{b,c,a}^{(y_p)} - \frac{c}{a+1} T_{b,c-1,a+2}^{(y_p)} + T_{b,c,a}, \quad (9.37)$$

$$T_{0,0,a}^{(x_p)} = T_{0,0,a}^{(y_p)} = 0, \quad T_{b,c,a}^{(z_p^j)} = -T_{b,c,a}^{(z)}. \quad (9.38)$$

9.4.2.4 Hypersingular Laplace potential

For this potential we obtain:

$$\begin{aligned} \xi_{b+1,c,a}^{(z,x_p^j)} &= -\frac{b}{a+1} \xi_{b-1,c,a+2}^{(z,x_p^j)} + \alpha_x \xi_{b,c,a}^{(z)} + x_p \xi_{b,c,a}^{(z,x_p^j)} + \frac{\hat{\mathbf{i}} \cdot \mathbf{n}^j}{(a+1)4\pi} \kappa_{b,c,a+1}^{(z,x_p^j)}, \\ \xi_{b+1,c,a}^{(z,y_p^j)} &= -\frac{b}{a+1} \xi_{b-1,c,a+2}^{(z,y_p^j)} + \beta_x \xi_{b,c,a}^{(z)} + x_p \xi_{b,c,a}^{(z,y_p^j)} - \frac{\hat{\mathbf{i}} \cdot \mathbf{n}^j (a-1)}{4\pi} y_d^j z_d \int x^b y^c r^{a-3} dx^j, \\ \xi_{b+1,c,a}^{(z,z_p^j)} &= -\frac{b}{a+1} \xi_{b-1,c,a+2}^{(z,z_p^j)} + x_p \xi_{b,c,a}^{(z,z_p^j)} - \frac{\hat{\mathbf{i}} \cdot \mathbf{n}^j}{4\pi} (\kappa_{b,c,a-1} + z_d^2 (a-1) \kappa_{b,c,a-3}), \end{aligned} \quad (9.39)$$

$$\begin{aligned} \xi_{b,c+1,a}^{(z,x_p^j)} &= -\frac{c}{a+1} \xi_{b,c-1,a+2}^{(z,x_p^j)} + \alpha_y \xi_{b,c,a}^{(z)} + y_p \xi_{b,c,a}^{(z,x_p^j)} + \frac{\hat{\mathbf{j}} \cdot \mathbf{n}^j}{(a+1)4\pi} \kappa_{b,c,a+1}^{(z,x_p^j)}, \\ \xi_{b,c+1,a}^{(z,y_p^j)} &= -\frac{c}{a+1} \xi_{b,c-1,a+2}^{(z,y_p^j)} + \beta_y \xi_{b,c,a}^{(z)} + y_p \xi_{b,c,a}^{(z,y_p^j)} - \frac{\hat{\mathbf{j}} \cdot \mathbf{n}^j (a-1)}{4\pi} y_d^j z_d \int x^b y^c r^{a-3} dx^j, \\ \xi_{b,c+1,a}^{(z,z_p^j)} &= -\frac{c}{a+1} \xi_{b,c-1,a+2}^{(z,z_p^j)} + y_p \xi_{b,c,a}^{(z,z_p^j)} - \frac{\hat{\mathbf{j}} \cdot \mathbf{n}^j}{4\pi} (\kappa_{b,c,a-1} + z_d^2 (a-1) \kappa_{b,c,a-3}), \end{aligned} \quad (9.40)$$

$$\begin{aligned}
\xi_{0,0,a}^{(z,x_p^j)} &= \frac{1}{4\pi} \partial_{x_p^j} (z_p^j k_{a-1}) = \frac{1}{4\pi} z_p^j \partial_{x_p^j} k_{a-1} = \frac{1}{4\pi} z_p^j y_p^j \frac{r^{a-1}}{\rho^2} \\
\xi_{0,0,a}^{(z,y_p^j)} &= \frac{1}{4\pi} \partial_{y_p^j} (z_p^j k_{a-1}) = \frac{1}{4\pi} z_p^j \partial_{y_p^j} k_{a-1} = \frac{1}{4\pi} z_p^j (x_d^j \frac{r^{a-1}}{\rho^2} - (a-1)i_{a-3}) \\
\xi_{0,0,a}^{(z,z_p^j)} &= \frac{1}{4\pi} \partial_{z_p^j} (z_p^j k_{a-1}) = \frac{k_{a-1}}{4\pi} + \frac{a-1}{4\pi} (z_p^j)^2 k_{a-3} = \frac{1}{4\pi} (ak_{a-1} + (a-1)y_p^j i_{a-3}),
\end{aligned} \tag{9.41}$$

with $\xi_{b,c,a}^{(z,x_p^j)} \equiv \partial_{z,x_p^j}^2 \xi_{b,c,a}$ and $\xi_{b,c,a}^{(z,y_p^j)} \equiv \partial_{z,y_p^j}^2 \xi_{b,c,a}$.

For the singularity terms $T_{b,c,a}^{(z,x_p)} \equiv \partial_{z,x_p}^2 T_{b,c,a}$ and $T_{b,c,a}^{(z,y_p)} \equiv \partial_{z,y_p}^2 T_{b,c,a}$ we have:

$$\begin{aligned}
T_{b+1,c,a}^{(z,x_p)} &= -\frac{b}{a+1} T_{b-1,c,a+2}^{(z,x_p)} + x_p T_{b,c,a}^{(z,x_p)} + T_{b,c,a}^{(z)}, \\
T_{b+1,c,a}^{(z,y_p)} &= -\frac{b}{a+1} T_{b-1,c,a+2}^{(z,y_p)} + x_p T_{b,c,a}^{(z,y_p)}, \quad T_{b+1,c,a}^{(z,z_p)} = -\frac{b}{a+1} T_{b-1,c,a+2}^{(z,z_p)} + x_p T_{b,c,a}^{(z,z_p)},
\end{aligned} \tag{9.42}$$

$$\begin{aligned}
T_{b,c+1,a}^{(z,y_p)} &= -\frac{c}{a+1} T_{b,c-1,a+2}^{(z,y_p)} + y_p T_{b,c,a}^{(z,y_p)} + T_{b,c,a}^{(z)}, \\
T_{b,c+1,a}^{(z,x_p)} &= -\frac{c}{a+1} T_{b,c-1,a+2}^{(z,x_p)} + y_p T_{b,c,a}^{(z,x_p)}, \quad T_{b,c+1,a}^{(z,z_p)} = -\frac{c}{a+1} T_{b,c-1,a+2}^{(z,z_p)} + y_p T_{b,c,a}^{(z,z_p)},
\end{aligned} \tag{9.43}$$

$$T_{0,0,a}^{(z,x_p^j)} = T_{0,0,a}^{(z,y_p^j)} = 0, \quad T_{0,0,a}^{(z,z_p^j)} = \frac{|h|^{a-1} + (z_p^j)^2 (a-1) |h|^{a-3}}{2} \text{ if } |h| \neq 0 \text{ else } 0. \tag{9.44}$$

9.4.3 Helmholtz kernel with polynomials of arbitrary order

9.4.3.1 Single layer potential

We want to find vector fields in the form:

$$\nabla_s \cdot \mathbf{m}_{b,c} \equiv \frac{x^b y^c e^{ikr}}{4\pi r}. \tag{9.45}$$

For this, the following auxiliary vector fields are introduced:

$$\nabla_s \cdot \mathbf{m}_{b,c,a} \equiv \frac{x^b y^c r^a e^{ikr}}{4\pi r}. \quad (9.46)$$

Lemma 3 (Recurrence relation for auxiliary vector fields - Helmholtz kernel). *If vector fields $\mathbf{m}_{b+1,c,a}$, $\mathbf{m}_{b,c,a}$, $\mathbf{m}_{b-1,c,a+1}$, and $\mathbf{m}_{b,c,a-1}$ satisfy (9.46) and $a \geq 1$, then the vector field $\mathbf{m}_{b+1,c,a-1}$ given by the following equation:*

$$\frac{a}{ik} \mathbf{m}_{b+1,c,a-1} = x_p \mathbf{m}_{b,c,a} - \mathbf{m}_{b+1,c,a} - \frac{b}{ik} \mathbf{m}_{b-1,c,a+1} + \frac{ax_p}{ik} \mathbf{m}_{b,c,a-1} + \frac{x^b y^c r^a e^{ikr}}{4\pi ik} \hat{\mathbf{i}} \quad (9.47)$$

also satisfies (9.46) for the index tuple $(b+1, c, a-1)$.

Proof. If vector fields $\mathbf{m}_{b,c,a}$, $\mathbf{m}_{b+1,c,a}$, $\mathbf{m}_{b-1,c,a+1}$ and $\mathbf{m}_{b,c,a-1}$ satisfy (9.46), the surface divergence of the right hand side of (9.47) yields: $\frac{ax^{b+1}y^c r^{a-1} e^{ikr}}{4\pi ik r}$, hence, $\nabla_s \cdot \mathbf{m}_{b+1,c,a-1} = \frac{x^{b+1}y^c r^{a-1} e^{ikr}}{4\pi r}$, which means that $\mathbf{m}_{b+1,c,a-1}$ satisfies (9.46) for the index tuple $(b+1, c, a-1)$. \square

This yields the following recurrence relation and its counterpart due to symmetry:

$$\begin{aligned} \xi_{b+1,c,a} &= x_p \xi_{b,c,a} - \frac{a}{ik} \xi_{b+1,c,a-1} - \frac{b}{ik} \xi_{b-1,c,a+1} + \frac{ax_p}{ik} \xi_{b,c,a-1} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{i}}}{4\pi ik} \theta_{b,c,a}, \\ \xi_{b,c+1,a} &= y_p \xi_{b,c,a} - \frac{a}{ik} \xi_{b,c+1,a-1} - \frac{c}{ik} \xi_{b,c-1,a+1} + \frac{ay_p}{ik} \xi_{b,c,a-1} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{j}}}{4\pi ik} \theta_{b,c,a}, \end{aligned} \quad (9.48)$$

where we have defined:

$$\xi_{b,c,a} \equiv \int \mathbf{m}_{b,c,a} \cdot \mathbf{n}^j dx^j, \quad \theta_{b,c,a} \equiv \int x^b y^c r^a e^{ikr} dx^j. \quad (9.49)$$

The challenge is to evaluate $\theta_{b,c,a}$. Assuming $k|r - r_0| \leq 1$ with $r_0 \equiv |\mathbf{r}_p - \mathbf{v}^j|$, which is usually satisfied in the BEM as the mesh size is typically set to be smaller than $\lambda/6$ with λ the wavelength, we expand the e^{ikr} term as Taylor series:

$$\begin{aligned}
\theta_{b,c,a} &= \int x^b y^c r^a e^{ikr} dx^j = e^{ikr_0} \int x^b y^c r^a e^{ik(r-r_0)} dx^j \\
&\approx e^{ikr_0} \sum_{l=0}^{p_e-1} \frac{(ik)^l}{l!} \int x^b y^c r^a (r - r_0)^l dx^j = e^{ikr_0} \sum_{l=0}^{p_e-1} A_l^{(p_e)} \int x^b y^c r^{a+l} dx^j \quad (9.50) \\
&= e^{ikr_0} \sum_{l=0}^{p_e-1} A_l^{(p_e)} \kappa_{b,c,a+l}, \quad A_l^{(p)} \equiv \frac{(ik)^l}{l!} \sum_{m=0}^{p-l-1} \frac{(-ikr_0)^m}{m!}.
\end{aligned}$$

For later use, we also define:

$$S_s \equiv \frac{e^{ikr_0}}{4\pi} \sum_{l=0}^{p_e-1} A_l^{(p_e)} k_{l+s}, \quad U_s \equiv \frac{e^{ikr_0}}{4\pi} \sum_{l=0}^{p_e-1} A_l^{(p_e)} i_{l+s}. \quad (9.51)$$

The following derivatives of these functions are also used later for the computation of derivatives of the single layer potential:

$$\partial_{x_p^j} S_{a+1} = \frac{e^{ikr_0}}{4\pi} \sum_{l=0}^{p_e-1} A_l^{(p_e)} \partial_{x_p^j} k_{l+a+1} = \frac{e^{ikr_0}}{4\pi} \sum_{l=0}^{p_e-1} A_l^{(p_e)} y_p^j \frac{r^{l+a+1}}{\rho^2} \approx \frac{y_p^j r^{a+1} e^{ikr}}{4\pi \rho^2}, \quad (9.52)$$

$$\begin{aligned}
\partial_{y_p^j} S_{a+1} &= \frac{e^{ikr_0}}{4\pi} \sum_{l=0}^{p_e-1} A_l^{(p_e)} \left(x_d^j \frac{r^{l+a+1}}{\rho^2} - (l+a+1) i_{l+a-1} \right) \\
&\approx \frac{x_d^j r^{a+1} e^{ikr}}{4\pi \rho^2} - \frac{e^{ikr_0}}{4\pi} \left(\sum_{l=0}^{p_e-1} A_l^{(p_e)} (l i_{l+a-1} + (a+1) i_{l+a-1}) \right) \\
&\approx \frac{x_d^j r^{a+1} e^{ikr}}{4\pi \rho^2} - \frac{e^{ikr_0}}{4\pi} \sum_{l=0}^{p_e-1} A_l^{(p_e)} ik i_{l+a} - (a+1) U_{a-1} \\
&= \frac{x_d^j r^{a+1} e^{ikr}}{4\pi \rho^2} - (ik U_a + (a+1) U_{a-1}),
\end{aligned} \quad (9.53)$$

$$\partial_z S_{a+1} = \frac{e^{ikr_0}}{4\pi} \sum_{l=0}^{p_e-1} A_l^{(p_e)} (l+a+1) z_d k_{l+a-1} \approx z_d (ik S_a + (a+1) S_{a-1}), \quad (9.54)$$

$$\partial_z U_a = \frac{e^{ikr_0}}{4\pi} \sum_{l=0}^{p_e-1} A_l^{(p_e)} (l+a) z_d i_{l+a-2} \approx z_d (ik U_{a-1} + a U_{a-2}). \quad (9.55)$$

Thus, by computing the values of $\kappa_{b,c,a}$ with maximum a index of $p_s + p_e - 1$, the approximation of $\theta_{b,c,a}$ can be computed. Note that (9.50) is a discrete convolution along the a -axis in the coefficient grid, which may be accelerated via fast Fourier transform (FFT) when both p_s and p_e are large.

The following recurrence relation can be used to generate vector fields $\mathbf{m}_{0,0,a}$ which are needed to initiate the recursions in the b - and c -axis (9.48):

$$\mathbf{m}_{0,0,a+1} = -\frac{a+1}{ik} \mathbf{m}_{0,0,a} + \frac{1}{ik} \boldsymbol{\chi}_{a+1}, \quad \mathbf{m}_{0,0,0} = \frac{1}{ik} \boldsymbol{\chi}_0, \quad \boldsymbol{\chi}_a \equiv \frac{r^a e^{ikr}}{4\pi \rho^2} \boldsymbol{\rho}. \quad (9.56)$$

This recurrence yields the recurrence for the $\xi_{b,c,a}$ coefficients:

$$\begin{aligned} \xi_{0,0,a+1} &= -\frac{a+1}{ik} \xi_{0,0,a} + \frac{1}{ik} \int \boldsymbol{\chi}_{a+1} \cdot \mathbf{n}^j dx^j = -\frac{a+1}{ik} \xi_{0,0,a} + \frac{y_p^j}{ik} \int \frac{r^{a+1} e^{ikr}}{4\pi \rho^2} dx^j \\ &\approx -\frac{a+1}{ik} \xi_{0,0,a} - \frac{e^{ikr_0}}{4\pi ik} \sum_{l=0}^{p_e-1} A_l^{(p_e)} k_{l+a+1} = -\frac{a+1}{ik} \xi_{0,0,a} - \frac{1}{ik} S_{a+1}, \\ \xi_{0,0,0} &= \frac{y_p^j}{ik} \int \frac{e^{ikr}}{4\pi \rho^2} dx^j \approx \frac{e^{ikr_0}}{4\pi ik} \sum_{l=0}^{p_e-1} A_l^{(p_e)} y_p^j \int \frac{r^l}{\rho^2} dx^j = -\frac{1}{ik} S_0. \end{aligned} \quad (9.57)$$

The singularity term also obeys the recursions:

$$\begin{aligned} T_{b+1,c,a} &= x_p T_{b,c,a} - \frac{a}{ik} T_{b+1,c,a-1} - \frac{b}{ik} T_{b-1,c,a+1} + \frac{ax_p}{ik} T_{b,c,a-1}, \\ T_{b,c+1,a} &= y_p T_{b,c,a} - \frac{a}{ik} T_{b,c+1,a-1} - \frac{c}{ik} T_{b,c-1,a+1} + \frac{ay_p}{ik} T_{b,c,a-1}. \end{aligned} \quad (9.58)$$

If the vector fields $\mathbf{m}_{0,0,a}$ are chosen using the recurrence (9.56), we have the following recurrence along the a -axis:

$$T_{0,0,a+1} = -\frac{a+1}{ik}T_{0,0,a} + \frac{1}{ik} \lim_{\varepsilon \rightarrow 0} \oint_{C_\varepsilon} \boldsymbol{\chi}_{a+1} \cdot \mathbf{n}_{C_\varepsilon} dC_\varepsilon = -\frac{a+1}{ik}T_{0,0,a} - \frac{|h|^{a+1}e^{ik|h|}}{2ik}. \quad (9.59)$$

Note that this equation holds for $a = -1$ and gives the seed term $T_{0,0,0} = ie^{ik|h|}/(2k)$.

9.4.3.2 Double layer potential

The derivatives of the $\xi_{0,0,a}$ are given by:

$$\xi_{0,0,0}^{(z)} \approx -z_d^j S_{-1}, \quad \xi_{0,0,a+1}^{(z)} \approx -\frac{a+1}{ik} \xi_{0,0,a}^{(z)} - z_d S_a - \frac{(a+1)z_d}{ik} S_{a-1}. \quad (9.60)$$

Recurrences along b - and c -axis:

$$\begin{aligned} \xi_{b+1,c,a}^{(z)} &= x_p \xi_{b,c,a}^{(z)} - \frac{a}{ik} \xi_{b+1,c,a-1}^{(z)} - \frac{b}{ik} \xi_{b-1,c,a+1}^{(z)} + \frac{ax_p}{ik} \xi_{b,c,a-1}^{(z)} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{i}}}{4\pi ik} \theta_{b,c,a}^{(z)}, \\ \xi_{b,c+1,a}^{(z)} &= y_p \xi_{b,c,a}^{(z)} - \frac{a}{ik} \xi_{b,c+1,a-1}^{(z)} - \frac{c}{ik} \xi_{b,c-1,a+1}^{(z)} + \frac{ay_p}{ik} \xi_{b,c,a-1}^{(z)} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{j}}}{4\pi ik} \theta_{b,c,a}^{(z)}. \end{aligned} \quad (9.61)$$

Here, the derivative $\theta_{b,c,a}^{(z)} \equiv \partial_z \theta_{b,c,a}$ can be computed as:

$$\begin{aligned}
\theta_{b,c,a}^{(z)} &\approx e^{ikr_0} \sum_{l=0}^{p_e-1} A_l^{(p_e)} (a+l) z_d \int x^b y^c r^{a+l-2} dx^j \\
&= e^{ikr_0} a z_d \sum_{l=0}^{p_e-1} A_l^{(p_e)} \int x^b y^c r^{a+l-2} dx^j + e^{ikr_0} \sum_{l=1}^{p_e-1} A_l^{(p_e)} l z_d \int x^b y^c r^{a+l-2} dx^j \\
&= e^{ikr_0} a z_d \sum_{l=0}^{p_e-1} A_l^{(p_e)} \int x^b y^c r^{a+l-2} dx^j + e^{ikr_0} i k z_d \sum_{l=1}^{p_e-1} A_{l-1}^{(p_e-1)} \int x^b y^c r^{a+l-2} dx^j \\
&\approx e^{ikr_0} a z_d \sum_{l=0}^{p_e-1} A_l^{(p_e)} \int x^b y^c r^{a+l-2} dx^j + e^{ikr_0} i k z_d \sum_{l=0}^{p_e-1} A_l^{(p_e)} \int x^b y^c r^{a+l-1} dx^j \\
&= z_d (a \theta_{b,c,a-2} + i k \theta_{b,c,a-1}).
\end{aligned} \tag{9.62}$$

If the vector fields $\mathbf{m}_{0,0,a}$ are chosen using the recurrence (9.56), we have the following recurrence along the a -axis for the singularity term:

$$\begin{aligned}
T_{0,0,0}^{(z)} &= -\partial_z \frac{e^{ik|h|}}{2ik} = \frac{-z_d e^{ik|h|}}{2|h|} \quad \text{if } z \neq z_p \text{ else } 0, \\
T_{0,0,a+1}^{(z)} &= -\frac{a+1}{ik} T_{0,0,a}^{(z)} + \frac{z_d (i(a+1) - k|h|) e^{ik|h|} |h|^{a-1}}{2k}.
\end{aligned} \tag{9.63}$$

The recursion for the singularity term along the b - and c - axis is given by:

$$\begin{aligned}
T_{b+1,c,a}^{(z)} &= x_p T_{b,c,a}^{(z)} - \frac{a}{ik} T_{b+1,c,a-1}^{(z)} - \frac{b}{ik} T_{b-1,c,a+1}^{(z)} + \frac{ax_p}{ik} T_{b,c,a-1}^{(z)}, \\
T_{b,c+1,a}^{(z)} &= y_p T_{b,c,a}^{(z)} - \frac{a}{ik} T_{b,c+1,a-1}^{(z)} - \frac{c}{ik} T_{b,c-1,a+1}^{(z)} + \frac{ay_p}{ik} T_{b,c,a-1}^{(z)}.
\end{aligned} \tag{9.64}$$

9.4.3.3 Adjoint double layer potential

Seed term:

$$\xi_{0,0,0}^{(x_p^j)} = -\frac{y_p^j e^{ikr}}{4\pi ik\rho^2}, \quad \xi_{0,0,0}^{(y_p^j)} = U_{-1} - \frac{x_d^j e^{ikr}}{4\pi ik\rho^2}. \quad (9.65)$$

Recurrence along a -axis is given by:

$$\begin{aligned} \xi_{0,0,a+1}^{(x_p^j)} &= -\frac{a+1}{ik} \xi_{0,0,a}^{(x_p^j)} - \frac{1}{ik} S_{a+1}^{(x_p^j)} = -\frac{a+1}{ik} \xi_{0,0,a}^{(x_p^j)} - \frac{y_p^j r^{a+1} e^{ikr}}{4\pi ik\rho^2}, \\ \xi_{0,0,a+1}^{(y_p^j)} &= -\frac{a+1}{ik} \xi_{0,0,a}^{(y_p^j)} + U_a + \frac{a+1}{ik} U_{a-1} - \frac{x_d^j r^{a+1} e^{ikr}}{4\pi ik\rho^2}. \end{aligned} \quad (9.66)$$

Recurrence along b - and c -axis are given by:

$$\begin{aligned} \xi_{b+1,c,a}^{(x_p^j)} &= x_p \xi_{b,c,a}^{(x_p^j)} + \alpha_x \xi_{b,c,a} - \frac{a}{ik} \xi_{b+1,c,a-1}^{(x_p^j)} - \frac{b}{ik} \xi_{b-1,c,a+1}^{(x_p^j)} + \frac{ax_p}{ik} \xi_{b,c,a-1}^{(x_p^j)} + \frac{a\alpha_x}{ik} \xi_{b,c,a-1} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{i}}}{4\pi ik} \theta_{b,c,a}^{(x_p^j)}, \\ \xi_{b+1,c,a}^{(y_p^j)} &= x_p \xi_{b,c,a}^{(y_p^j)} + \beta_x \xi_{b,c,a} - \frac{a}{ik} \xi_{b+1,c,a-1}^{(y_p^j)} - \frac{b}{ik} \xi_{b-1,c,a+1}^{(y_p^j)} + \frac{ax_p}{ik} \xi_{b,c,a-1}^{(y_p^j)} + \frac{a\beta_x}{ik} \xi_{b,c,a-1} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{i}}}{4\pi ik} \theta_{b,c,a}^{(y_p^j)}, \\ \xi_{b,c+1,a}^{(x_p^j)} &= y_p \xi_{b,c,a}^{(x_p^j)} + \alpha_y \xi_{b,c,a} - \frac{a}{ik} \xi_{b,c+1,a-1}^{(x_p^j)} - \frac{c}{ik} \xi_{b,c-1,a+1}^{(x_p^j)} + \frac{ay_p}{ik} \xi_{b,c,a-1}^{(x_p^j)} + \frac{a\alpha_y}{ik} \xi_{b,c,a-1} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{j}}}{4\pi ik} \theta_{b,c,a}^{(x_p^j)}, \\ \xi_{b,c+1,a}^{(y_p^j)} &= y_p \xi_{b,c,a}^{(y_p^j)} + \beta_y \xi_{b,c,a} - \frac{a}{ik} \xi_{b,c+1,a-1}^{(y_p^j)} - \frac{c}{ik} \xi_{b,c-1,a+1}^{(y_p^j)} + \frac{ay_p}{ik} \xi_{b,c,a-1}^{(y_p^j)} + \frac{a\beta_y}{ik} \xi_{b,c,a-1} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{j}}}{4\pi ik} \theta_{b,c,a}^{(y_p^j)}, \end{aligned} \quad (9.67)$$

where we used:

$$\begin{aligned} \theta_{b,c,a}^{(x_p^j)} &\approx -x^b y^c r^a e^{ikr} + \alpha_x b \theta_{b-1,c,a} + \alpha_y c \theta_{b,c-1,a}, \\ \theta_{b,c,a}^{(y_p^j)} &\approx -y_d^j (a \theta_{b,c,a-2} + ik \theta_{b,c,a-1}). \end{aligned} \quad (9.69)$$

Recursions for the singularity term are given by:

$$\begin{aligned}
T_{b+1,c,a}^{(x_p)} &= x_p T_{b,c,a}^{(x_p)} + T_{b,c,a} - \frac{a}{ik} T_{b+1,c,a-1}^{(x_p)} - \frac{b}{ik} T_{b-1,c,a+1}^{(x_p)} + \frac{ax_p}{ik} T_{b,c,a-1}^{(x_p)} + \frac{a}{ik} T_{b,c,a-1}, \\
T_{b+1,c,a}^{(y_p)} &= x_p T_{b,c,a}^{(y_p)} - \frac{a}{ik} T_{b+1,c,a-1}^{(y_p)} - \frac{b}{ik} T_{b-1,c,a+1}^{(y_p)} + \frac{ax_p}{ik} T_{b,c,a-1}^{(y_p)},
\end{aligned} \tag{9.70}$$

$$\begin{aligned}
T_{b,c+1,a}^{(x_p)} &= y_p T_{b,c,a}^{(x_p)} - \frac{a}{ik} T_{b,c+1,a-1}^{(x_p)} - \frac{c}{ik} T_{b,c-1,a+1}^{(x_p)} + \frac{ay_p}{ik} T_{b,c,a-1}^{(x_p)}, \\
T_{b,c+1,a}^{(y_p)} &= y_p T_{b,c,a}^{(y_p)} + T_{b,c,a} - \frac{a}{ik} T_{b,c+1,a-1}^{(y_p)} - \frac{c}{ik} T_{b,c-1,a+1}^{(y_p)} + \frac{ay_p}{ik} T_{b,c,a-1}^{(y_p)} + \frac{a}{ik} T_{b,c,a-1}.
\end{aligned} \tag{9.71}$$

Note that $T_{0,0,a}^{(x_p)} = T_{0,0,a}^{(y_p)} = 0$, $\xi_{b,c,a}^{(z_p^j)} = -\xi_{b,c,a}^{(z)}$, and $T_{b,c,a}^{(z_p^j)} = -T_{b,c,a}^{(z)}$.

9.4.3.4 Hypersingular potential

Seed terms:

$$\begin{aligned}
\xi_{0,0,0}^{(z,x_p^j)} &= y_p^j z_p^j \frac{e^{ikr}}{4\pi\rho^2 r}, \quad \xi_{0,0,0}^{(z,y_p^j)} = z_d(ikU_{-2} - U_{-3}) + z_p^j \frac{x_d^j e^{ikr}}{4\pi\rho^2 r}, \\
\xi_{0,0,0}^{(z,z_p^j)} &= (z_d^j S_{-1})^{(z)} = S_{-1} + z_d^j S_{-1}^{(z)} = S_{-1} + (z_p^j)^2 (ikS_{-2} - S_{-3}).
\end{aligned} \tag{9.72}$$

Recurrence along a -axis:

$$\begin{aligned}
\xi_{0,0,a+1}^{(z,x_p^j)} &= -\frac{a+1}{ik} \xi_{0,0,a}^{(z,x_p^j)} + y_p^j z_p^j \frac{(kr - i(a+1)) r^{a-1} e^{ikr}}{4\pi k \rho^2}, \\
\xi_{0,0,a+1}^{(z,y_p^j)} &= -\frac{a+1}{ik} \xi_{0,0,a}^{(z,y_p^j)} + z_d(ikU_{a-1} + (2a+1)U_{a-2} + \frac{a^2-1}{ik}U_{a-3}) + z_p^j \frac{x_d^j (kr - i(a+1)) r^{a-1} e^{ikr}}{4\pi k \rho^2}, \\
\xi_{0,0,a+1}^{(z,z_p^j)} &= -\frac{a+1}{ik} \xi_{0,0,a}^{(z,z_p^j)} + S_a + \left(\frac{a+1}{ik} + z_d^2 ik\right) S_{a-1} + z_d^2 ((2a+1)S_{a-2} + \frac{a^2-1}{ik} S_{a-3}).
\end{aligned} \tag{9.73}$$

Recurrence along b - and c -axis:

$$\begin{aligned}
\xi_{b+1,c,a}^{(z,x_p^j)} &= x_p \xi_{b,c,a}^{(z,x_p^j)} + \alpha_x \xi_{b,c,a}^{(z)} - \frac{a}{ik} \xi_{b+1,c,a-1}^{(z,x_p^j)} - \frac{b}{ik} \xi_{b-1,c,a+1}^{(z,x_p^j)} + \frac{ax_p}{ik} \xi_{b,c,a-1}^{(z,x_p^j)} + \frac{a\alpha_x}{ik} \xi_{b,c,a-1}^{(z)} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{i}}}{4\pi ik} \theta_{b,c,a}^{(z,x_p^j)}, \\
\xi_{b+1,c,a}^{(z,y_p^j)} &= x_p \xi_{b,c,a}^{(z,y_p^j)} + \beta_x \xi_{b,c,a}^{(z)} - \frac{a}{ik} \xi_{b+1,c,a-1}^{(z,y_p^j)} - \frac{b}{ik} \xi_{b-1,c,a+1}^{(z,y_p^j)} + \frac{ax_p}{ik} \xi_{b,c,a-1}^{(z,y_p^j)} + \frac{a\beta_x}{ik} \xi_{b,c,a-1}^{(z)} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{i}}}{4\pi ik} \theta_{b,c,a}^{(z,y_p^j)}, \\
\xi_{b+1,c,a}^{(z,z_p^j)} &= x_p \xi_{b,c,a}^{(z,z_p^j)} - \frac{a}{ik} \xi_{b+1,c,a-1}^{(z,z_p^j)} - \frac{b}{ik} \xi_{b-1,c,a+1}^{(z,z_p^j)} + \frac{ax_p}{ik} \xi_{b,c,a-1}^{(z,z_p^j)} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{i}}}{4\pi ik} \theta_{b,c,a}^{(z,z_p^j)},
\end{aligned} \tag{9.74}$$

$$\begin{aligned}
\xi_{b,c+1,a}^{(z,x_p^j)} &= y_p \xi_{b,c,a}^{(z,x_p^j)} + \alpha_y \xi_{b,c,a}^{(z)} - \frac{a}{ik} \xi_{b,c+1,a-1}^{(z,x_p^j)} - \frac{c}{ik} \xi_{b,c-1,a+1}^{(z,x_p^j)} + \frac{ay_p}{ik} \xi_{b,c,a-1}^{(z,x_p^j)} + \frac{a\alpha_y}{ik} \xi_{b,c,a-1}^{(z)} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{j}}}{4\pi ik} \theta_{b,c,a}^{(z,x_p^j)}, \\
\xi_{b,c+1,a}^{(z,y_p^j)} &= y_p \xi_{b,c,a}^{(z,y_p^j)} + \beta_y \xi_{b,c,a}^{(z)} - \frac{a}{ik} \xi_{b,c+1,a-1}^{(z,y_p^j)} - \frac{c}{ik} \xi_{b,c-1,a+1}^{(z,y_p^j)} + \frac{ay_p}{ik} \xi_{b,c,a-1}^{(z,y_p^j)} + \frac{a\beta_y}{ik} \xi_{b,c,a-1}^{(z)} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{j}}}{4\pi ik} \theta_{b,c,a}^{(z,y_p^j)}, \\
\xi_{b,c+1,a}^{(z,z_p^j)} &= y_p \xi_{b,c,a}^{(z,z_p^j)} - \frac{a}{ik} \xi_{b,c+1,a-1}^{(z,z_p^j)} - \frac{c}{ik} \xi_{b,c-1,a+1}^{(z,z_p^j)} + \frac{ay_p}{ik} \xi_{b,c,a-1}^{(z,z_p^j)} + \frac{\mathbf{n}^j \cdot \hat{\mathbf{j}}}{4\pi ik} \theta_{b,c,a}^{(z,z_p^j)}.
\end{aligned} \tag{9.75}$$

Here we used:

$$\begin{aligned}
\theta_{b,c,a}^{(z,x_p^j)} &\approx -x^b y^c (a + ikr) r^{a-2} z_d e^{ikr} + \alpha_x b \theta_{b-1,c,a}^{(z)} + \alpha_y c \theta_{b,c-1,a}^{(z)}, \\
\theta_{b,c,a}^{(z,y_p^j)} &\approx -y_d^j (a \theta_{b,c,a-2}^{(z)} + ik \theta_{b,c,a-1}^{(z)}), \quad \theta_{b,c,a}^{(z)} \approx z_d (a \theta_{b,c,a-2} + ik \theta_{b,c,a-1}), \\
\theta_{b,c,a}^{(z,z_p^j)} &\approx -z_d^j (a \theta_{b,c,a-2}^{(z)} + ik \theta_{b,c,a-1}^{(z)}) - (a \theta_{b,c,a-2} + ik \theta_{b,c,a-1}).
\end{aligned} \tag{9.76}$$

Recursions for the singularity term:

$$\begin{aligned}
T_{0,0,a}^{(z,x_p)} &= T_{0,0,a}^{(z,y_p)} = 0, \quad T_{0,0,0}^{(z,x_p)} = \frac{ike^{ik|h|}}{2}, \\
T_{0,0,a+1}^{(z,x_p)} &= -\frac{a+1}{ik} T_{0,0,a}^{(z,x_p)} - \frac{e^{ik|h|} |h|^{a-1}}{2k} (-2(a+1)k|h| + i(a(a+1) - k^2|h|^2)),
\end{aligned} \tag{9.77}$$

$$\begin{aligned}
T_{b+1,c,a}^{(z,x_p)} &= x_p T_{b,c,a}^{(z,x_p)} + T_{b,c,a}^{(z)} - \frac{b}{ik} T_{b-1,c,a+1}^{(z,x_p)} + \frac{a}{ik} (x_p T_{b,c,a-1}^{(z,x_p)} + T_{b,c,a-1}^{(z)} - T_{b+1,c,a-1}^{(z,x_p)}), \\
T_{b+1,c,a}^{(z,y_p)} &= x_p T_{b,c,a}^{(z,y_p)} - \frac{a}{ik} T_{b+1,c,a-1}^{(z,y_p)} - \frac{b}{ik} T_{b-1,c,a+1}^{(z,y_p)} + \frac{ax_p}{ik} T_{b,c,a-1}^{(z,y_p)}, \\
T_{b+1,c,a}^{(z,z_p)} &= x_p T_{b,c,a}^{(z,z_p)} - \frac{a}{ik} T_{b+1,c,a-1}^{(z,z_p)} - \frac{b}{ik} T_{b-1,c,a+1}^{(z,z_p)} + \frac{ax_p}{ik} T_{b,c,a-1}^{(z,z_p)},
\end{aligned} \tag{9.78}$$

$$\begin{aligned}
T_{b,c+1,a}^{(z,x_p)} &= y_p T_{b,c,a}^{(z,x_p)} - \frac{a}{ik} T_{b,c+1,a-1}^{(z,x_p)} - \frac{c}{ik} T_{b,c-1,a+1}^{(z,x_p)} + \frac{ay_p}{ik} T_{b,c,a-1}^{(z,x_p)}, \\
T_{b,c+1,a}^{(z,y_p)} &= y_p T_{b,c,a}^{(z,y_p)} + T_{b,c,a}^{(z)} - \frac{c}{ik} T_{b,c-1,a+1}^{(z,y_p)} + \frac{a}{ik} (y_p T_{b,c,a-1}^{(z,y_p)} + T_{b,c,a-1}^{(z)} - T_{b,c+1,a-1}^{(z,y_p)}), \\
T_{b,c+1,a}^{(z,z_p)} &= y_p T_{b,c,a}^{(z,z_p)} - \frac{a}{ik} T_{b,c+1,a-1}^{(z,z_p)} - \frac{c}{ik} T_{b,c-1,a+1}^{(z,z_p)} + \frac{ay_p}{ik} T_{b,c,a-1}^{(z,z_p)}.
\end{aligned} \tag{9.79}$$

9.4.4 Complexity

For the Laplace kernel, evaluating the recursions in algorithm 1 and algorithm 2 has a complexity of $O(p_s^3)$. Note that this is the complexity to obtain all terms with the factor $x^b y^c$ for $b + c \leq p_s$. The layer potential for one shape function can then be computed as a linear combination of $O(p_s^2)$ terms. If the layer potential was to be evaluated by Gauss-Legendre quadrature of order $b + c + p'_e$ with p'_e the degrees to accommodate the non-polynomial part, i.e., the Green function part of the integrand, the complexity to evaluate one layer potential would be $O((p_s + p'_e)^2)$, assuming the use of $O((p_s + p'_e)^2)$ quadrature nodes. For the Helmholtz case, if we use p_e degrees for the Taylor series expansion of the oscillatory term e^{ikr} , the complexity of the recursions in the RIPE method is $O(p_s^3 p_e)$ with a simple implementation and $O(p_s^2 (p_s + p_e) \log(p_s + p_e))$ if FFT is used for the convolution. For Gauss-Legendre, the complexity is $O((p_s + p'_e)^2)$, where p'_e degrees are used to accommodate the non-polynomial part. Note that p_e and p'_e , although both parameters specify the precision of the integral evaluation, cannot be directly associated; Gauss-Legendre would require larger p'_e for evaluation points in the near-field since near the singularity the integrand cannot be well approximated by low-order polynomials. On the other hand, p_e in the RIPE method can be smaller in the near-field as it is the

expansion order of the Taylor series. This implies that the appropriate method depends on the distance of the evaluation point from the element. If low-order elements with $p_s = O(1)$ are used, the complexity of the RIPE method for the Laplace and Helmholtz kernel reduces to $O(1)$ and $O(p_e)$, respectively. The complexities for the low-order case and general case are summarized in Table 9.1.

	Laplace		Helmholtz	
	$p_s \approx 1$	General	$p_s \approx 1$	General
RIPE	$O(1)$	$O(p_s^3)$	$O(p_e)$	$O(p_s^3 p_e)$
RIPE (+FFT)	-	-	-	$O(p_s^2(p_s + p_e) \log(p_s + p_e))$
Gauss-Legendre	$O(p_e'^2)$	$O((p_s + p_e')^2)$	$O(p_e'^2)$	$O((p_s + p_e')^2)$

Table 9.1: Complexity for evaluating layer potentials for p_s of $O(1)$ and general order, respectively. Note that the complexities for Gauss-Legendre are for one layer potential evaluation, while for RIPE they are the complexities to compute all terms with the monomial factor $x^b y^c$ for $0 \leq b + c \leq p_s$ in a single execution of the recursive algorithm. These terms can be reused when evaluating multiple layer potentials with different linear combinations of these monomial terms.

9.5 Numerical evaluation

9.5.1 Accuracy

9.5.1.1 Nearly-singular case

The method was tested for all four layer potentials, for both the Laplace and Helmholtz kernels. Adaptive Gauss-Kronrod quadrature, implemented in QUADPACK [198], was used to compute the layer potentials P_{GK} over a boundary element. The result of RIPE, P_{RIPE} was compared against this reference result via the relative error $|P_{\text{RIPE}} - P_{\text{GK}}| / |P_{\text{GK}}|$. The maximum expansion order p_e in RIPE was set to 32, and the error tolerance of

Gauss-Kronrod was set to 10^{-12} . The layer potentials were also evaluated using 12th order Gauss-Legendre quadrature which has 33 quadrature nodes on a triangle element. The triangle $\{(0, 0, 0), (1, 0, 0), (0, 1, 0)\}$ and point $\mathbf{r}_p = (1/3, 1/3, |h|)$ were used as the element and observation points, respectively. $N(x, y) = x^3$ was used as the shape function. Results are shown in Fig. 9.6. Our RIPE method delivers remarkable accuracy particularly in the challenging nearly singular regime $|h|/D < 1$, with D the maximum edge length of the element. Gauss-Legendre quadrature is not reliable in this domain. Note that the condition $kD = 1$ approximately corresponds to six wavelengths per element, which is typically used as the maximum mesh size in boundary element analysis. We also observe that the RIPE method has worse performance for larger $|h|/D$, where Gauss-Legendre, on the other hand, delivers good accuracy. Given this complementary character, one can switch the integration routine based on the distance of the evaluation point from the element.

9.5.1.2 Singular and hypersingular case

Cases with the observation point \mathbf{r}_p on the element (singular and hypersingular cases) are handled without any modifications to the framework. For these the method was compared with results obtained by Guiggiani's method [178], internally using Gauss-Legendre quadrature of 20th order. The same element as in the previous section with \mathbf{r}_p on the center of the element was used to compute the single layer and the hypersingular potentials. The six shape functions of a conforming second-order Lagrange triangle element were used. Results for both Laplace and Helmholtz kernels are shown in 9.2–9.5.

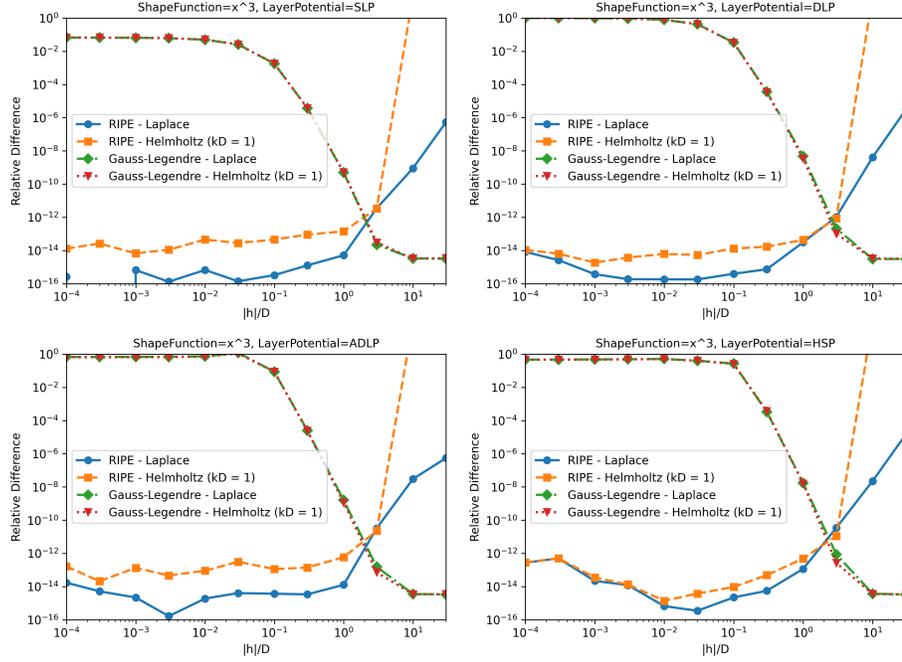


Figure 9.6: Relative difference of the layer potentials between RIPE or 12-th order Gauss-Legendre quadrature and the reference adaptive Gauss-Kronrod quadrature, for the shape function $N(x, y) = x^3$. Results for single layer potential (top left), double layer potential (top right), adjoint double layer potential (bottom left), and hypersingular potential (bottom right).

A relative difference $|P_{\text{RIPE}} - P_{\text{Gui}}|/|P_{\text{Gui}}|$ from 10^{-10} to 10^{-13} was observed.

	P_{RIPE}	P_{Gui}	$ P_{\text{RIPE}} - P_{\text{Gui}} / P_{\text{Gui}} $
N_1	-0.0096108650741614	-0.0096108650753968	$1.28540284 \times 10^{-10}$
N_2	-0.0096108650741614	-0.0096108650753968	$1.28539201 \times 10^{-10}$
N_3	-0.0059161308348599	-0.0059161308348537	$1.05940117 \times 10^{-12}$
N_4	0.0733163156462961	0.0733163156487268	$3.31537174 \times 10^{-11}$
N_5	0.0716914080260122	0.0716914080259944	$2.48939665 \times 10^{-13}$
N_6	0.0716914080260122	0.0716914080259944	$2.48552512 \times 10^{-13}$

Table 9.2: Single layer potentials for the Laplace kernel evaluated over the element with second-order shape functions of a conforming Lagrange triangle element. Numbers for RIPE (P_{RIPE}), Guiggiani's method (P_{Gui}), and their relative differences are shown.

	P_{RIPE}	P_{Gui}	$ P_{\text{RIPE}} - P_{\text{Gui}} / P_{\text{Gui}} $
N_1	0.3411586129005689	0.3411586129009882	$1.22913865 \times 10^{-12}$
N_2	0.3411586129005690	0.3411586129009879	$1.22799965 \times 10^{-12}$
N_3	0.5031187119584526	0.5031187119589535	$9.95654936 \times 10^{-13}$
N_4	-0.7261344637586460	-0.7261344637604444	$2.47674552 \times 10^{-12}$
N_5	-0.9322819538428125	-0.9322819538448199	$2.15320509 \times 10^{-12}$
N_6	-0.9322819538428122	-0.9322819538448149	$2.14820345 \times 10^{-12}$

Table 9.3: Hypersingular potentials for the Laplace kernel evaluated over the second-order element.

	P_{RIPE}	P_{Gui}	$ P_{\text{RIPE}} - P_{\text{Gui}} / P_{\text{Gui}} $
N_1	-0.0097575874677327	-0.0097575874673111	$1.65408707 \times 10^{-10}$
N_2	-0.0097575874677327	-0.0097575874673111	$1.65411332 \times 10^{-10}$
N_3	-0.0059358291069226	-0.0059358291069324	$2.15796857 \times 10^{-12}$
N_4	0.0724350497721009	0.0724350497713923	$4.33315252 \times 10^{-11}$
N_5	0.0707990955161933	0.0707990955162131	$4.29383222 \times 10^{-13}$
N_6	0.0707990955161934	0.0707990955162131	$4.30321106 \times 10^{-13}$

Table 9.4: Single layer potentials for the Helmholtz kernel ($kD = 1$) over the second-order element.

	P_{RIPE}	P_{Gui}	$ P_{\text{RIPE}} - P_{\text{Gui}} / P_{\text{Gui}} $
N_1	0.3387374371700406	0.3387374371703727	$1.34783671 \times 10^{-12}$
N_2	0.3387374371700407	0.3387374371703721	$1.34571877 \times 10^{-12}$
N_3	0.5016372264001558	0.5016372264006584	$1.00193060 \times 10^{-12}$
N_4	-0.7079157406214216	-0.7079157406230446	$2.45050302 \times 10^{-12}$
N_5	-0.9144708322784949	-0.9144708322805092	$2.20267780 \times 10^{-12}$
N_6	-0.9144708322784947	-0.9144708322805069	$2.20036963 \times 10^{-12}$

Table 9.5: Hypersingular potentials for the Helmholtz kernel ($kD = 1$) over the second-order element.

9.5.2 Stability

The RIPE method was evaluated by testing its accuracy with high order elements up to $p_s = 9$. Fig. 9.7 shows results for both Laplace and Helmholtz layer potentials. The relative difference with the reference Gauss-Kronrod increases with element order, and this divergence is more for the Helmholtz kernel.

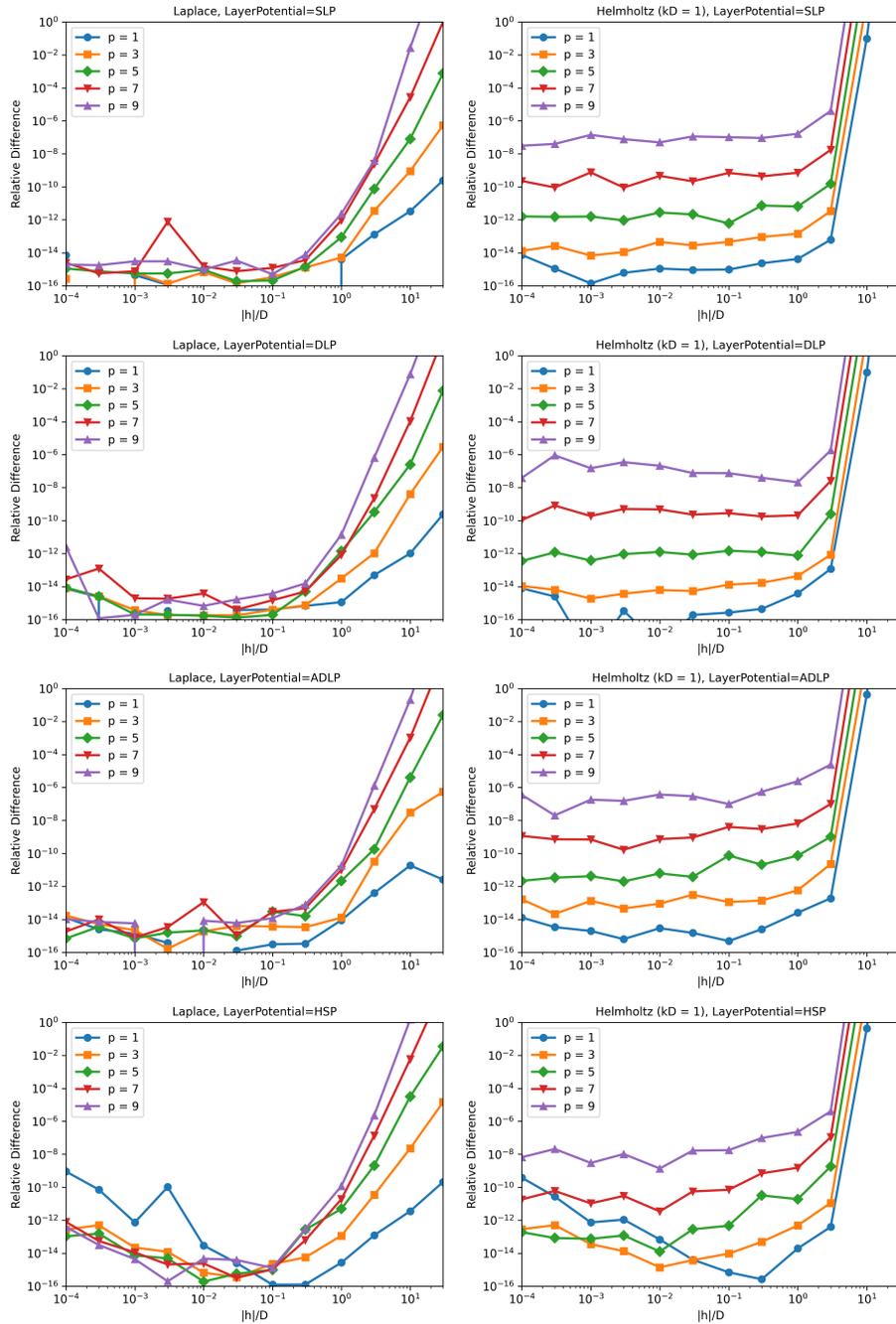


Figure 9.7: Relative difference of the layer potentials between RIPE and the adaptive Gauss-Kronrod quadrature, for various shape function orders. Laplace (left column) and Helmholtz (right column), for single layer, double layer, adjoint double layer, and hypersingular potentials (top to bottom).

9.5.3 Computation time

Computation time of the layer potentials with 5th order shape functions over the element $\{(0, 0, 0), (1, 0, 0), (0, 1, 0)\}$ and observation point $\mathbf{r}_p = (1/3, 1/3, |h|)$ were measured and compared against the adaptive Gauss-Kronrod baseline. The expansion order for the Helmholtz kernel in RIPE was set to $p_e = 32$, and the error tolerance of Gauss-Kronrod was set to 10^{-12} . The result is shown in Fig. 11.2. The times for RIPE are those spent to compute integrals weighted by the shape function $N(x, y) = x^b y^c$ for all terms in $0 \leq b + c \leq p_s$ in one run of the recursion, while for Gauss-Kronrod these are for the computation of a *single* term with the shape function $N(x, y) = x^5$. Gauss-Kronrod requires long computation time in the near-singular domain with small $|h|/D$, while computation time for RIPE does not depend on $|h|/D$ and can be orders of magnitudes faster. The measurements for RIPE were performed using a Python prototype without performance optimization, and Gauss-Kronrod was measured using Scipy's interface to QUADPACK. The advantage of the RIPE method should be more significant in a production collocation BEM setup where the $x^b y^c$ terms for all the $0 \leq b + c \leq p_s$ combinations are needed.

9.6 Conclusion

A recursive algorithm (RIPE) to evaluate layer potentials arising in the collocation BEM for the Laplace and Helmholtz equation, tailored for piecewise flat boundary elements with polynomial shape functions of arbitrary orders was proposed. Numerical tests showed that RIPE exhibits remarkable accuracy in the nearly-singular regime, where Gauss-

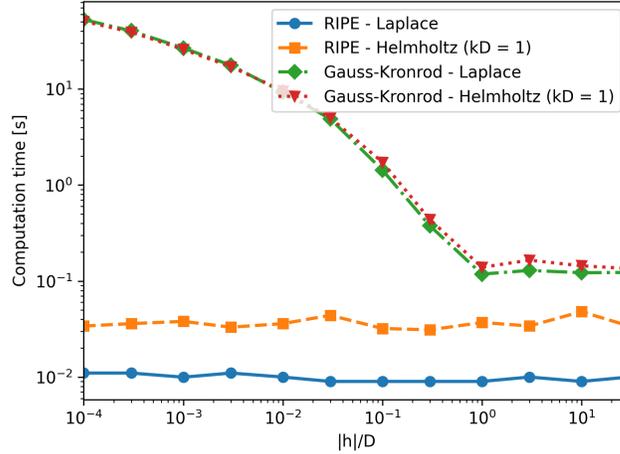


Figure 9.8: The total wall-clock time for computing all four layer potentials for a 5th order element, for RIPE and for adaptive Gauss-Kronrod quadrature. Note that the times for RIPE are the times spent to compute integrals weighted by $x^b y^c$ for all $((p_s + 1)(p_s + 2))/2 = 21$ terms in $0 \leq b + c \leq 5$ in one run of the recursion, while the times for Gauss-Kronrod are for computing only the integral with shape function x^5 .

Legendre quadrature is not effective. Experiments indicate that RIPE is several orders of magnitude faster than Gauss-Kronrod quadrature, a general-purpose adaptive quadrature method. RIPE would hence serve as an efficient routine for nearly singular, singular, and hypersingular integrals for Laplace and Helmholtz layer potentials. The benefits of the RIPE method are that it offers: (1) analytical integration for polynomial elements achieved by a simple formulation using auxiliary vector fields, (2) easier error control for the Helmholtz case compared to methods based on Gauss-Legendre quadrature due to the series expansion-based formulation, (3) accuracy and efficiency for nearly singular, singular, and hypersingular integrals, achieved under a single framework naturally supporting all of these cases without separate modifications, for all four standard layer potentials, for both Laplace and Helmholtz kernels.

The RIPE method has multiple frontiers for further development. Supporting manifold surfaces as shown in [183] is an important direction for development since the reduction

of geometrical errors via higher-order geometry representations can be beneficial for curved objects [199]. While [183] demonstrates an inspiring scheme which appears to be effective for the double layer potential of the Laplace kernel, room for improvement seem to remain particularly in the treatment of the single layer potential which does not offer an exact differential 2-form allowing straightforward application of Stokes' theorem. The authors have recently developed a layer potential quadrature scheme for manifold boundary elements with constant densities which is effective for the close evaluation of the single layer potential [200]. The development of a layer potential quadrature scheme which is both efficient and accurate for general boundary elements with high-order expressions for both the density and geometry is still an open problem. Although the RIPE method is restricted to flat elements, we note that many practical BEM software are written for flat elements, and many problems have meshes from architecture or CAD that are truly flat, or where plane triangular meshes are more easily available. Extending RIPE to other kernels, Galerkin BEM and integrating it into FMM-BEM solvers are other directions for future work.

Chapter 10: Layer potential quadrature on manifold boundary elements
with constant densities for Laplace and Helmholtz kernels
in \mathbb{R}^3

A method is proposed for evaluation of single and double layer potentials of the Laplace and Helmholtz equations on piecewise smooth manifold boundary elements with constant densities. The method is based on a novel two-term decomposition of the layer potentials, derived by means of differential geometry. The first term is an integral of a differential 2-form which can be reduced to contour integrals using Stokes' theorem, while the second term is related to the element curvature. This decomposition reduces the degree of singularity and the curvature term can be further regularized by a polar coordinate transform. The method can handle singular and nearly singular integrals. Numerical results validating the accuracy of the method are presented for all combinations of single and double layer potentials, for the Laplace and Helmholtz kernels, and for singular and nearly singular integrals.¹

¹The research presented in this chapter is currently under review for publication. Its preprint can be found in [201].

10.1 Introduction

Boundary element methods (BEM) are widely used for solving partial differential equations arising in science and engineering. In the classical BEM, the boundary of the problem domain is typically represented using polygon meshes composed of piecewise flat boundary elements. This simple representation of the geometry allowed the development of efficient analytical methods tailored for flat boundary elements, e.g., [176, 181, 182, 191]. Methods capable of solving problems with geometries represented by piecewise manifold surfaces which are not necessarily piecewise flat are receiving attention due to their ability to represent the geometry of a wide variety of problems accurately or exactly, and thereby eliminating a source of discretization error [199, 202].

Practical BEM solvers are composed of multiple building blocks including iterative linear system solvers, fast matrix-vector product evaluation routines using fast multipole methods, etc. One of the essential computation routines in the BEM is the numerical evaluation of layer potentials integrals required for computing the near field interactions. This task is nontrivial because the integrands can be singular or nearly singular. Standard quadrature schemes which are effective for integrating polynomials of limited degrees, e.g. Gauss-Legendre quadrature, are known to produce inaccurate results when the evaluation point is close to the element. Many techniques have been developed over the years to accurately evaluate boundary integrals in such cases [1, 177–181, 183–187]. Ref. [1] provides a recent extensive survey on this subject. The approaches developed include singularity cancellation using coordinate transforms [177, 179, 180], singularity subtraction [178], continuation approach [188], dimension reduction [176, 181–183], adaptive subdivision [184],

and quadrature by expansion [185–187]. The authors have recently proposed analytical methods based on dimensionality reduction for both collocation [176] and Galerkin BEM [182], tailored for flat boundary elements. Zhu and Veerapaneni [183] recently introduced a method for Laplace layer potentials on high-order curved elements using dimensionality reduction via Stokes’ theorem and quaternion algebra. This method exploits the fact that an exact differential form is available for the Laplace double layer potential. While the application of this method to the evaluation of the Laplace single layer potential was discussed in passing in [183], numerical results were only presented for the double layer potential case. A summary of related quadrature methods for layer potentials or their multipole expansions is shown in Table 10.1.

Distance → Element type ↓	Singular / nearly singular	Far-field expansions
Flat, constant	Lenoir & Salles [181], Gumerov & Duraiswami [193]	Gumerov, Kaneko & Duraiswami [191]
Flat, high order	Newman [189], Kaneko, Gumerov & Duraiswami [176]	Newman [189], Kaneko & Duraiswami [192]
Curved, constant	Present work	
Curved, high order	Zhu & Veerapaneni [183], Klöckner et al. [185], Rosen & Cormack [188]	

Table 10.1: Summary of related quadrature methods for layer potentials or its multipole expansions based on analytical or dimensionality-reduction based evaluation. Methods for high order elements can be applied to constant elements, and methods for curved elements can be applied to flat elements. A summary of methods based on other approaches e.g. singularity subtraction/cancellation can be found in [1].

In this work, we focus on nearly singular and singular layer potential evaluation and propose a method which supports both Laplace and Helmholtz kernels for both single and double layer potentials on manifold boundary elements for the special case of constant

densities. The method is based on a decomposition of the layer potentials into two terms. The first term is an integral of a differential 2-form, which can be evaluated via one-dimensional contour integrals after applying Stokes' theorem on manifolds, while the second term with reduced singularity which is related with the curvature of the element. The singularity in the second term can be further reduced by the classical technique of polar coordinate transform, used in e.g. [177–179]. Layer potentials with higher order densities are important but require further development which may involve a redesign of the set of basis functions to obtain convenient exact differential forms, as shown for the Laplace double layer case [183]. In this work we instead focus on the constant element case and provide formulations and numerical results for both Laplace and Helmholtz kernels for both the single and double layer potentials. The accuracy of the proposed method was confirmed via element-level tests and also using an example benchmark problem for which an analytical solution is available.

10.2 Boundary element method and layer potentials

The boundary element method is extensively used for numerical solution of partial differential equations, e.g. the Helmholtz equation and the Laplace equation, respectively given by

$$-k^2 u(\mathbf{r}) - \nabla^2 u(\mathbf{r}) = f(\mathbf{r}), \quad -\nabla^2 u(\mathbf{r}) = f(\mathbf{r}), \quad \mathbf{r} \in \Omega \subset \mathbb{R}^3, \quad (10.1)$$

with wavenumber k , field u in domain $\Omega \subset \mathbb{R}^3$, and source f . The weak form of (11.1) can be written in terms of single- and double layer potentials V , K [194]:

$$\begin{aligned} \{(c_p \gamma_{0,p} + K \gamma_{0,q} - V \gamma_{1,q})u\}(\mathbf{r}_p) &= \{N_0 f\}(\mathbf{r}_p), \\ \{V \psi\}(\mathbf{r}_p) &\equiv \int_{\mathbf{r}_q \in \Gamma} G(\mathbf{r}_p, \mathbf{r}_q) \psi(\mathbf{r}_q) d\Gamma, \quad \{K \phi\}(\mathbf{r}_p) \equiv \int_{\mathbf{r}_q \in \Gamma} \frac{\partial G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_q} \phi(\mathbf{r}_q) d\Gamma, \end{aligned} \quad (10.2)$$

with $c_p = 1/2$ on a smooth boundary, γ_0 and γ_1 the boundary trace and normal derivative operators, and N_0 the Newton potential operator, defined as:

$$\begin{aligned} \{\gamma_{0,q} u\}(\mathbf{r}_q) &\equiv \lim_{\hat{\mathbf{r}}_q \in \Omega \rightarrow \mathbf{r}_q \in \Gamma} u(\hat{\mathbf{r}}_q), \quad \{\gamma_{1,q} u\}(\mathbf{r}_q) \equiv \mathbf{n}_q \cdot \nabla_q u(\mathbf{r}_q), \quad \mathbf{r}_q \in \Gamma = \partial\Omega, \\ \{N_0 f\}(\mathbf{r}_p) &= \int_{\mathbf{r}_q \in \Omega} G(\mathbf{r}_p, \mathbf{r}_q) f(\mathbf{r}_q) d\Omega, \quad \mathbf{r}_p \in \mathbb{R}^3, \end{aligned} \quad (10.3)$$

where $G(\mathbf{r}_p, \mathbf{r}_q)$ is the respective Laplace or Helmholtz Green function:

$$G_L(\mathbf{r}_p, \mathbf{r}_q) = \frac{1}{4\pi r}, \quad G_H(\mathbf{r}_p, \mathbf{r}_q) = \frac{e^{ikr}}{4\pi r}, \quad r \equiv |\mathbf{r}_q - \mathbf{r}_p|. \quad (10.4)$$

In the BEM the boundary Γ is discretized into surface boundary elements which can be either flat or curved, and which may exactly discretize the original geometry when the closed-form representation of the geometry is available. This applies to e.g. surfaces generated using computer-aided design (CAD) software. The layer potential integrals over these elements are evaluated to form the linear system of equations. The densities ψ , ϕ are approximated via local, typically polynomial, functions (also called *shape functions*) with unknown coefficients which must be determined. In the present work we assume that the boundary $\Gamma = \partial\Omega$ is a union of boundary elements $\Gamma = \bigcup_i S_i$,

where each S_i is a smooth oriented Riemannian submanifolds with a boundary [203] and has constant density. Geometrical singularities e.g. wedges or corners need to be removed by subdividing the surface before applying the proposed method.

10.3 Differential geometry preliminaries

10.3.1 Curvature of regular surfaces

In the differential geometry of curves and surfaces, various types of curvatures are defined. Here we briefly review the definition of the normal curvature, as it is central to the proposed method. The normal curvature $\kappa_N(p, c)$ of a regular curve c on a regular surface S at point $p \in c \subset S$ is defined as:

$$\kappa_N(p, c) \equiv \kappa(p, c) \mathbf{n}_p(c) \cdot \mathbf{n}_p(S), \quad \kappa(p, c) \equiv \frac{\|\mathbf{r}''(t) \times \mathbf{r}'(t)\|}{\|\mathbf{r}'(t)\|^3}, \quad \mathbf{n}_p(c) \equiv \frac{\hat{\mathbf{t}}'(t)}{\|\hat{\mathbf{t}}'(t)\|}, \quad (10.5)$$

where $\kappa(p, c)$ is the curvature of curve c at p , $p = \mathbf{r}(t) \in \mathbb{R}^3$ the parametrization of c , $\mathbf{n}_p(c)$ the unit normal vector of c at p , $\hat{\mathbf{t}} = \mathbf{r}'(t)/\|\mathbf{r}'(t)\|$ the unit tangent vector of c at p , and $\mathbf{n}_p(S)$ the unit normal vector of S at p [204]. This setup is illustrated in Fig. 10.1 (left). While c can be any regular curve on S passing through p , it is useful to consider the c which is the cross section of S with respect to a normal plane of S at p which is spanned by $\mathbf{n}_p(S)$ and a direction vector $\hat{\rho}$, which is a tangent vector of S at p . We refer to the normal curvature associated with direction $\hat{\rho}$ as $\kappa_N(\hat{\rho})$.

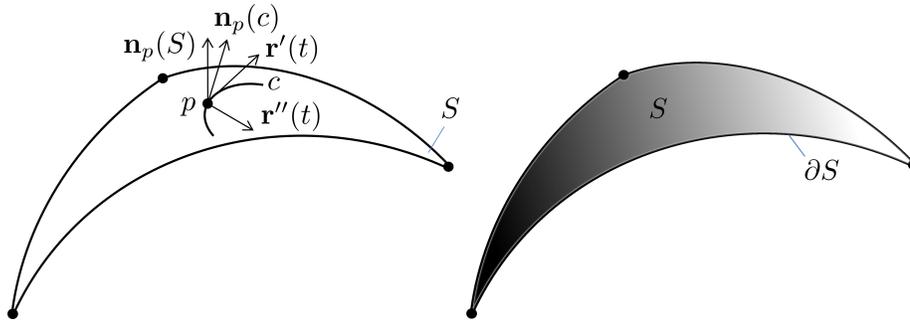


Figure 10.1: Left: The normal vectors of a regular surface S and a regular curve c at point p . Right: the surface S and its contour ∂S .

10.3.2 Stokes' theorem on smooth manifolds

Another important tool available for the calculus on manifolds is Stokes' theorem [205]:

$$\int_S d\omega = \oint_{\partial S} \omega, \quad (10.6)$$

where S is an oriented smooth submanifold with a boundary ∂S as illustrated in Fig. 10.1 (right), ω a differential form and $d\omega$ the exterior derivative of ω . Stokes' theorem allows the reduction of a surface integral into a contour integral, and can be used to evaluate integrals of exterior derivatives of exact differential forms. This applies to the Laplace double layer potential as demonstrated in [183].

10.4 Problem statement

The setup under consideration is shown in Fig. 10.2. Let our boundary element S be an oriented two-dimensional smooth Riemannian submanifold with a boundary [203] in \mathbb{R}^3 , which is parametrized via the function $\mathbf{r}_q(u, v) \in \mathbb{R}^3$ with variables $u, v \in \mathbb{R}$ defined on a reference triangle $\{u, v | 0 \leq u, 0 \leq v, u + v \leq 1\}$. In the rest of the paper we

will refer to this element simply as *manifold element*. Let us denote the vertices of S as \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 , and the unit normal vector and tangent plane at point \mathbf{r}_q as \mathbf{n}_q and $T_q(S)$, respectively. Let us denote the normal plane at point \mathbf{r}_q spanned by \mathbf{n}_q and a tangent vector \mathbf{t} as $N_q(\mathbf{t})$. The projection of an evaluation point \mathbf{r}_p onto a given tangent plane is denoted as $\tilde{\mathbf{r}}_p$. We also define $r \equiv |\mathbf{r}_q - \mathbf{r}_p|$, $\boldsymbol{\rho} \equiv \mathbf{r}_q - \tilde{\mathbf{r}}_p$, $\rho \equiv |\boldsymbol{\rho}|$, $h \equiv \mathbf{n}_q \cdot (\mathbf{r}_p - \mathbf{r}_q)$. The goal is to develop a method to evaluate the single and double layer potentials over a

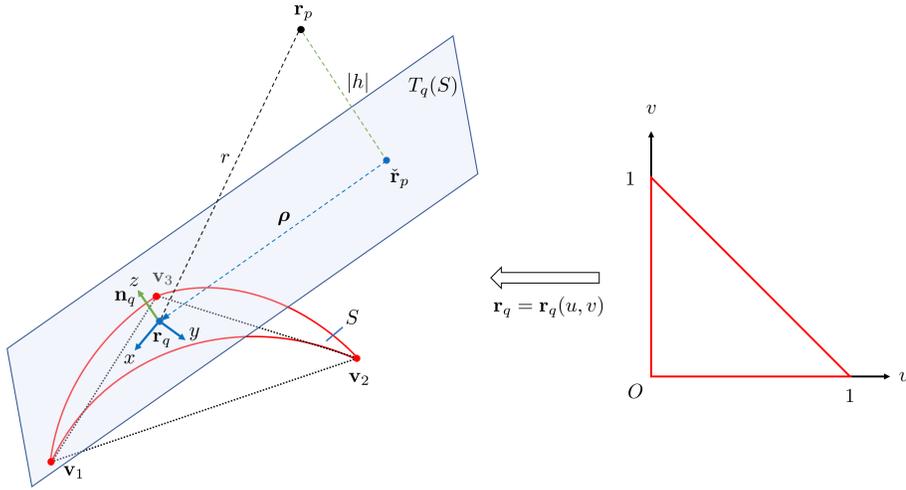


Figure 10.2: The manifold boundary element S and the tangent plane $T_q(S)$ at \mathbf{r}_q . Red, blue, and green lines indicate edges of the element, vectors parallel to the tangent plane $T_q(S)$, and vectors orthogonal to the tangent plane, respectively. \mathbf{n}_q , $\hat{\boldsymbol{\rho}} \equiv \boldsymbol{\rho}/|\boldsymbol{\rho}|$ and $\tilde{\boldsymbol{\rho}} \equiv \mathbf{n}_q \times \hat{\boldsymbol{\rho}}$ can be used to construct a local orthogonal coordinate frame centered at \mathbf{r}_q .

given manifold element S which is accurate in the nearly singular ($0 < \min_{\mathbf{r}_q} r \ll 1$) and singular ($\min_{\mathbf{r}_q} r = 0$) cases.

10.5 Proposed method

The following proposition provides a decomposition of the Green functions and their normal derivatives which allows the application of Stokes' theorem. It is a main result of our paper.

Proposition 4 (Decomposition of Green functions and their normal derivatives). *Green functions and their normal derivatives for the Laplace and Helmholtz equation can be decomposed into two terms as follows:*

$$\begin{aligned} G_K(\mathbf{r}_p, \mathbf{r}_q) &= (\nabla_{\mathbf{r}_q} \times \mathbf{f}_K) \cdot \mathbf{n}_q + \frac{1}{4\pi} (C_K \kappa_N(\tilde{\boldsymbol{\rho}}) + D_K \kappa_N(\hat{\boldsymbol{\rho}})), \\ \frac{\partial G_K(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_q} &= (\nabla_{\mathbf{r}_q} \times \mathbf{f}'_K) \cdot \mathbf{n}_q + \frac{1}{4\pi} (C'_K \kappa_N(\tilde{\boldsymbol{\rho}}) + D'_K \kappa_N(\hat{\boldsymbol{\rho}})), \end{aligned} \quad (10.7)$$

with $K = \{L, H\}$ the identifier whether the kernel is Laplace or Helmholtz, \mathbf{n}_q the unit normal vector at point \mathbf{r}_q , $\kappa_N(\hat{\boldsymbol{\rho}})$ and $\kappa_N(\tilde{\boldsymbol{\rho}})$ the normal curvature of the element at point \mathbf{r}_q for the normal planes spanned by \mathbf{n}_q and the tangent vectors $\hat{\boldsymbol{\rho}} \equiv \boldsymbol{\rho}/|\boldsymbol{\rho}|$ and $\tilde{\boldsymbol{\rho}} \equiv \mathbf{n}_q \times \hat{\boldsymbol{\rho}}$, respectively, \mathbf{f}_K and \mathbf{f}'_K the pseudo potential fields, C_K , C'_K , D_K and D'_K the weights of the curvatures defined as follows:

$$\begin{aligned} \mathbf{f}_L &\equiv \frac{\rho \tilde{\boldsymbol{\rho}}}{4\pi(r+h)}, \quad C_L \equiv \frac{h}{r+h}, \quad D_L \equiv \frac{r}{r+h}, \quad \mathbf{f}'_L \equiv \frac{\mathbf{f}_L}{r}, \quad C'_L \equiv \frac{C_L}{r}, \quad D'_L \equiv \frac{D_L}{r}, \\ \mathbf{f}_H &\equiv \frac{e^{ikr} - e^{ikh}}{4\pi ik\rho} \tilde{\boldsymbol{\rho}}, \quad C_H \equiv \frac{h(e^{ikr} - e^{ikh})}{ik\rho^2}, \quad D_H \equiv e^{ikh} - \frac{h(e^{ikr} - e^{ikh})}{ik\rho^2}, \\ \mathbf{f}'_H &\equiv \frac{re^{ikh} - he^{ikr}}{4\pi r\rho} \tilde{\boldsymbol{\rho}}, \quad C'_H \equiv \frac{h(re^{ikh} - he^{ikr})}{r\rho^2}, \quad D'_H \equiv \frac{re^{ikr} - he^{ikh}}{\rho^2} - ike^{ikh}. \end{aligned} \quad (10.8)$$

Proof. Let us define the following pseudo potential fields (see Section 10.4 and Fig. 10.2 for the definition of r , h and $\boldsymbol{\rho}$):

$$\mathbf{m}_L \equiv \frac{(r-h)\boldsymbol{\rho}}{4\pi\rho^2}, \quad \mathbf{m}'_L \equiv \frac{(r-h)\boldsymbol{\rho}}{4\pi r\rho^2}, \quad \mathbf{m}_H \equiv \frac{e^{ikr} - e^{ikh}}{4\pi ik\rho^2} \boldsymbol{\rho}, \quad \mathbf{m}'_H \equiv \frac{re^{ikr} - he^{ikh}}{4\pi r\rho^2} \boldsymbol{\rho}, \quad (10.9)$$

where L and H indicate the Laplace and Helmholtz kernels respectively and the prime

denotes corresponding expressions for the double layer potential. In the following we will use the notation:

$$\begin{aligned}
\mathbf{r}_u &\equiv \frac{\partial \mathbf{r}_q}{\partial u}, & \mathbf{r}_v &\equiv \frac{\partial \mathbf{r}_q}{\partial v}, & \mathbf{n}_u &\equiv \frac{\partial \mathbf{n}_q}{\partial u}, & \mathbf{n}_v &\equiv \frac{\partial \mathbf{n}_q}{\partial v}, \\
\mathbf{r}_{uu} &\equiv \frac{\partial^2 \mathbf{r}_q}{\partial u^2}, & \mathbf{r}_{uv} &\equiv \frac{\partial^2 \mathbf{r}_q}{\partial u \partial v}, & \mathbf{r}_{vv} &\equiv \frac{\partial^2 \mathbf{r}_q}{\partial v^2}, \\
r_{u\parallel} &\equiv \mathbf{r}_u \cdot \hat{\boldsymbol{\rho}}, & \mathbf{r}_{u\parallel} &\equiv (\mathbf{r}_u \cdot \hat{\boldsymbol{\rho}}) \hat{\boldsymbol{\rho}}, & r_{u\perp} &\equiv \mathbf{r}_u \cdot \tilde{\boldsymbol{\rho}}, & \mathbf{r}_{u\perp} &\equiv (\mathbf{r}_u \cdot \tilde{\boldsymbol{\rho}}) \tilde{\boldsymbol{\rho}}, \\
r_{v\parallel} &\equiv \mathbf{r}_v \cdot \hat{\boldsymbol{\rho}}, & \mathbf{r}_{v\parallel} &\equiv (\mathbf{r}_v \cdot \hat{\boldsymbol{\rho}}) \hat{\boldsymbol{\rho}}, & r_{v\perp} &\equiv \mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}}, & \mathbf{r}_{v\perp} &\equiv (\mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}}) \tilde{\boldsymbol{\rho}}, \\
n_{u\parallel} &\equiv \mathbf{n}_u \cdot \hat{\boldsymbol{\rho}}, & \mathbf{n}_{u\parallel} &\equiv (\mathbf{n}_u \cdot \hat{\boldsymbol{\rho}}) \hat{\boldsymbol{\rho}}, & n_{u\perp} &\equiv \mathbf{n}_u \cdot \tilde{\boldsymbol{\rho}}, & \mathbf{n}_{u\perp} &\equiv (\mathbf{n}_u \cdot \tilde{\boldsymbol{\rho}}) \tilde{\boldsymbol{\rho}}, \\
n_{v\parallel} &\equiv \mathbf{n}_v \cdot \hat{\boldsymbol{\rho}}, & \mathbf{n}_{v\parallel} &\equiv (\mathbf{n}_v \cdot \hat{\boldsymbol{\rho}}) \hat{\boldsymbol{\rho}}, & n_{v\perp} &\equiv \mathbf{n}_v \cdot \tilde{\boldsymbol{\rho}}, & \mathbf{n}_{v\perp} &\equiv (\mathbf{n}_v \cdot \tilde{\boldsymbol{\rho}}) \tilde{\boldsymbol{\rho}}.
\end{aligned} \tag{10.10}$$

We consider local Cartesian coordinate frames with origin at \mathbf{r}_q , x - and y - axes lying in the tangent plane $T_q(S)$, and the z -axis pointing towards the normal direction \mathbf{n}_q (see Fig. 10.2). With $\boldsymbol{\rho} \equiv \mathbf{r}_q - \mathbf{r}_p + (\mathbf{n}_q \cdot (\mathbf{r}_p - \mathbf{r}_q)) \mathbf{n}_q = \mathbf{r}_q - \mathbf{r}_p + h \mathbf{n}_q$, we have:

$$\begin{aligned}
\partial_u \boldsymbol{\rho} &= \mathbf{r}_u + h \mathbf{n}_u + c \mathbf{n}_q, & \partial_u \rho &= \hat{\boldsymbol{\rho}} \cdot \partial_u \boldsymbol{\rho} = \hat{\boldsymbol{\rho}} \cdot (\mathbf{r}_u + h \mathbf{n}_u), & \partial_u r &= \frac{\mathbf{r}_u \cdot \boldsymbol{\rho}}{r}, \\
\partial_u \mathbf{n}_q &\equiv \mathbf{n}_u = \frac{1}{J} (\partial_u \mathbf{C} - \mathbf{n}_q (\mathbf{n}_q \cdot \partial_u \mathbf{C})), & \mathbf{C} &\equiv \mathbf{r}_u \times \mathbf{r}_v, & J &= |\mathbf{C}|, \\
\partial_u h &= \mathbf{n}_u \cdot (\mathbf{r}_p - \mathbf{r}_q) = -\mathbf{n}_u \cdot \boldsymbol{\rho}, & \partial_z h &= -1, & \partial_z \boldsymbol{\rho} &= \mathbf{0}, & \partial_z r &= \frac{-h}{r},
\end{aligned} \tag{10.11}$$

where c is some real number. It can be found that the partial derivative of the pseudo potential fields in (10.9) with respect to u and v have the general form:

$$\begin{aligned}
4\pi \partial_u \mathbf{m} &= A \mathbf{r}_{u\perp} + B \mathbf{r}_{u\parallel} + C \mathbf{n}_{u\perp} + D \mathbf{n}_{u\parallel} + c \mathbf{n}_q, \\
4\pi \partial_v \mathbf{m} &= A \mathbf{r}_{v\perp} + B \mathbf{r}_{v\parallel} + C \mathbf{n}_{v\perp} + D \mathbf{n}_{v\parallel} + c \mathbf{n}_q,
\end{aligned} \tag{10.12}$$

where \mathbf{m} is the appropriately chosen pseudo potential field in (10.9). Since we have

$$4\pi J \partial_x \mathbf{m}_x = (\mathbf{r}_v)_y 4\pi \partial_u \mathbf{m}_x - (\mathbf{r}_u)_y 4\pi \partial_v \mathbf{m}_x, \quad 4\pi J \partial_y \mathbf{m}_y = -(\mathbf{r}_v)_x 4\pi \partial_u \mathbf{m}_y + (\mathbf{r}_u)_x 4\pi \partial_v \mathbf{m}_y, \quad (10.13)$$

it follows that the surface divergence $\nabla_s = \nabla - \mathbf{n}_q(\mathbf{n}_q \cdot \nabla)$ of \mathbf{m} multiplied by $4\pi J$ can be computed as:

$$\begin{aligned} & 4\pi J(\partial_x \mathbf{m}_x + \partial_y \mathbf{m}_y) \\ &= -(-r_{v\perp} \hat{\boldsymbol{\rho}} + r_{v\parallel} \tilde{\boldsymbol{\rho}}) \cdot 4\pi \partial_u \mathbf{m} + (-r_{u\perp} \hat{\boldsymbol{\rho}} + r_{u\parallel} \tilde{\boldsymbol{\rho}}) \cdot 4\pi \partial_v \mathbf{m} \\ &= -(-r_{v\perp} \hat{\boldsymbol{\rho}} + r_{v\parallel} \tilde{\boldsymbol{\rho}}) \cdot (A \mathbf{r}_{u\perp} + B \mathbf{r}_{u\parallel} + C \mathbf{n}_{u\perp} + D \mathbf{n}_{u\parallel}) \\ &\quad + (-r_{u\perp} \hat{\boldsymbol{\rho}} + r_{u\parallel} \tilde{\boldsymbol{\rho}}) \cdot (A \mathbf{r}_{v\perp} + B \mathbf{r}_{v\parallel} + C \mathbf{n}_{v\perp} + D \mathbf{n}_{v\parallel}) \\ &= +r_{v\perp} (B r_{u\parallel} + D n_{u\parallel}) - r_{v\parallel} (A r_{u\perp} + C n_{u\perp}) - r_{u\perp} (B r_{v\parallel} + D n_{v\parallel}) + r_{u\parallel} (A r_{v\perp} + C n_{v\perp}) \\ &= (A + B)(r_{u\parallel} r_{v\perp} - r_{u\perp} r_{v\parallel}) + C(r_{u\parallel} n_{v\perp} - r_{v\parallel} n_{u\perp}) + D(r_{v\perp} n_{u\parallel} - r_{u\perp} n_{v\parallel}) \\ &= (A + B)J - C J \kappa_N(\tilde{\boldsymbol{\rho}}) - D J \kappa_N(\hat{\boldsymbol{\rho}}), \end{aligned} \quad (10.14)$$

where $\kappa_N(\tilde{\boldsymbol{\rho}})$ and $\kappa_N(\hat{\boldsymbol{\rho}})$ are the normal curvatures at point \mathbf{r}_q on the element with respect to normal planes spanned by \mathbf{n}_q and tangent vectors $\tilde{\boldsymbol{\rho}}$ and $\hat{\boldsymbol{\rho}}$, respectively. In the last step, we have used the following lemma:

Lemma 5.

$$C(r_{v\parallel} n_{u\perp} - r_{u\parallel} n_{v\perp}) + D(r_{u\perp} n_{v\parallel} - r_{v\perp} n_{u\parallel}) = J(C \kappa_N(\tilde{\boldsymbol{\rho}}) + D \kappa_N(\hat{\boldsymbol{\rho}})), \quad (10.15)$$

Proof. See appendix C. □

Table 10.2: Coefficients introduced in (10.12) for Laplace and Helmholtz, single and double layer potentials.

	$A + B$	C	D
\mathbf{m}_L	$\frac{1}{r}$	$\frac{h}{r+h}$	$\frac{r}{r+h}$
\mathbf{m}'_L	$\frac{h}{r^3}$	$\frac{h}{r(r+h)}$	$\frac{1}{r+h}$
\mathbf{m}_H	$\frac{e^{ikr}}{r}$	$\frac{h(e^{ikr}-e^{ikh})}{ik\rho^2}$	$e^{ikh} - \frac{h(e^{ikr}-e^{ikh})}{ik\rho^2}$
\mathbf{m}'_H	$\frac{he^{ikr}(1-ikr)}{r^3}$	$\frac{h(re^{ikh}-he^{ikr})}{r\rho^2}$	$\frac{re^{ikr}-he^{ikh}}{\rho^2} - ike^{ikh}$

(10.17)

From (10.14) it follows:

$$\frac{A+B}{4\pi} = \nabla_s \cdot \mathbf{m} + \frac{C\kappa_N(\tilde{\boldsymbol{\rho}}) + D\kappa_N(\hat{\boldsymbol{\rho}})}{4\pi}. \quad (10.16)$$

The coefficients $A+B$, C , and D for each pseudo potential field in (10.9) are summarized in (10.17). It turns out that for all cases, the left hand side of (10.16) is nothing but the Green function or its normal derivative. Lastly, by using $\mathbf{f}_K = \mathbf{n}_q \times \mathbf{m}_K$ and $\mathbf{f}'_K = \mathbf{n}_q \times \mathbf{m}'_K$, we finally obtain decomposition (10.7):

$$\begin{aligned} G_K(\mathbf{r}_p, \mathbf{r}_q) &= (\nabla_{\mathbf{r}_q} \times \mathbf{f}_K) \cdot \mathbf{n}_q + \frac{1}{4\pi} (C_K \kappa_N(\tilde{\boldsymbol{\rho}}) + D_K \kappa_N(\hat{\boldsymbol{\rho}})), \\ \frac{\partial G_K(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_q} &= (\nabla_{\mathbf{r}_q} \times \mathbf{f}'_K) \cdot \mathbf{n}_q + \frac{1}{4\pi} (C'_K \kappa_N(\tilde{\boldsymbol{\rho}}) + D'_K \kappa_N(\hat{\boldsymbol{\rho}})). \end{aligned} \quad (10.18)$$

□

We refer to the first and second term in decomposition (10.7) as the *Stokes term* and the *Curvature term*, respectively. The Stokes term offers a differential 2-form: $d\omega = (\nabla_{\mathbf{r}_q} \times \mathbf{f}_K) \cdot \mathbf{n}_q dS$, whose integral can be reduced to a contour integral of a differential 1-form $\omega = \mathbf{f}_K \cdot d\mathbf{l}$ due to Stokes' theorem with $d\mathbf{l}$ the line element vector along the contour

∂S . Hence, the layer potentials now can be expressed as follows:

$$\begin{aligned} \{V_K\}(\mathbf{r}_p) &= \oint_{\partial S} \mathbf{f}_K \cdot d\mathbf{l} + \frac{1}{4\pi} \int_S (C_K \kappa_N(\tilde{\boldsymbol{\rho}}) + D_K \kappa_N(\hat{\boldsymbol{\rho}})) dS, \\ \{K_K\}(\mathbf{r}_p) &= \oint_{\partial S} \mathbf{f}'_K \cdot d\mathbf{l} + \frac{1}{4\pi} \int_S (C'_K \kappa_N(\tilde{\boldsymbol{\rho}}) + D'_K \kappa_N(\hat{\boldsymbol{\rho}})) dS, \end{aligned} \quad (10.19)$$

with subscript $K = \{L, H\}$ indicating the type of the kernel.

Remark. *The double layer potential of the Laplace kernel can be expressed as an integral of an exact form only and [183] utilized this fact. On the other hand, the single layer potential was approximated in [183] by a double layer potential with a modified density function. The efficacy of this approach was not discussed in [183] and is unclear. We present numerical results for both the present method based on decomposition (10.7) and the method presented in [183] in Section 10.6 for a comparison.*

Remark. *It is interesting that the proposed method, which is derived from the perspective of differential geometry, resembles a feature of the continuation approach [188] which also results in a decomposition of the integrand into two parts where one of the terms absorbs the “worst part” of the singularity. The decomposition in the continuation approach is based on Taylor series expansions and the geometric meanings of the decomposed terms are not clear. In contrast, decomposition (10.7) offers two terms with clear geometric meanings associated with differential forms and the curvature of the element.*

Remark. *If the element is flat, the curvature term vanishes and the layer potential evaluation (10.19) reduces to the evaluation of just contour integrals. Furthermore, this case can be evaluated analytically using the RIPE method [176].*

As a consequence, the one-dimensional integral of the Stokes term over the curvilinear boundary can be evaluated using Gauss-Legendre quadrature, and methods for two-dimensional quadrature can be applied to the curvature term. For non-negative h , as $h \rightarrow 0$ with $r \rightarrow 0$, the curvature term has a regularity of $1/r^{n-1}$ as opposed to $1/r^n$ of the original integrand before the decomposition, where $n = 1$ for the single layer potential and $n = 2$ for the double layer potential. This means that the integral still contains a (near-) singularity in the double layer potential case. To further regularize the singularity, we employ the classical technique of polar coordinate transform around the singularity [206], which is illustrated in Fig. 10.3. A polar coordinate system (R, θ) is set up on a flat surrogate element \bar{S} whose vertices are identical to those of the original manifold element. Points on this surrogate element \bar{S} are mapped to the reference triangle via an affine mapping, which are then mapped to the manifold element S via the parametrization $\mathbf{r}_q = \mathbf{r}_q(u, v)$. Point $R = 0$, i.e. the origin of the polar coordinates on \bar{S} , is chosen to be the point which

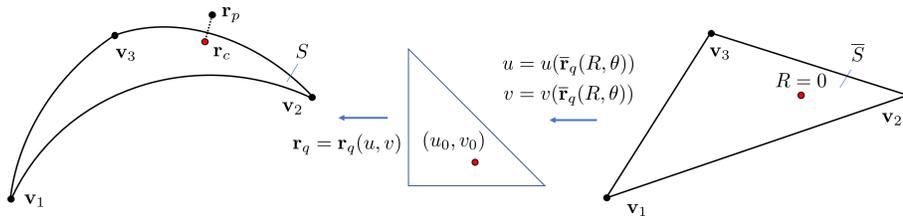


Figure 10.3: The coordinate transform used for the quadrature of the curvature term.

maps onto $\mathbf{r}_c = \mathbf{r}_q(u_0, v_0)$, the point on S nearest to the evaluation point \mathbf{r}_p . This mapping was also used in [206], though [206] applies subsequent coordinate transformations. Here we only use the basic polar coordinate transform.

Remark. *The Stokes term can be still nearly singular if \mathbf{r}_p is close to an edge of the element. This can be avoided by using nonconforming boundary elements where the*

collocation points are in the interior of the element. This is naturally satisfied in constant elements with center-panel collocation.

Remark. For negative values of h , the functions in (10.8) have a singularity at $h = -r$ which hinders the evaluation of the integrals using the presented approach. Such singularities arise when \mathbf{r}_p is in the inward normal bundle of the element. This can be resolved by utilizing respectively the symmetry and anti-symmetry of the layer potentials with respect to the exchange of variables u and v in the parametrization $\mathbf{r}_q(u, v)$. For the single layer potential we have the symmetry:

$$\int_{u=0}^1 \int_{v=0}^{1-u} G(\mathbf{r}_p, \mathbf{r}_q(u, v)) J(u, v) dv du = \int_{u=0}^1 \int_{v=0}^{1-u} G(\mathbf{r}_p, \mathbf{r}_q(v, u)) J(v, u) dv du \quad (10.20)$$

with J the Jacobian of the transform from \mathbf{r}_q to (u, v) and for the double layer potential we have the anti-symmetry where this exchange results in a sign flip. This exchange of variables also flips the direction of the normal vectors and the sign of h . For a given evaluation point \mathbf{r}_p , therefore, we can evaluate the same layer potentials using this symmetry property to avoid the singularity $h = -r$. This may not be always possible, since the outward and inward normal bundles of the element can have a non-empty intersection and evaluation points in this intersection cannot avoid the singularity. This can happen if the element is too curved. In such cases, the element can be subdivided until the evaluation point \mathbf{r}_p can avoid the inward normal bundle in one of the parametrizations and the proposed method can be applied to the subdivided elements. The geometry is illustrated in Fig. 10.4. An example pseudo-code implementing the procedure to avoid singularities is listed in algorithm 3. Various optimization methods could be used for

step 1 in algorithm 3. The Newton-Raphson method was used in our implementation. While this step adds additional computation cost, similar computation is needed in other methods for nearly singular integrals (e.g. [206]), where the first step is to find the point on the element closest to \mathbf{r}_p . As Newton's method converges quickly its computational overhead is limited and can be practically considered constant per element.

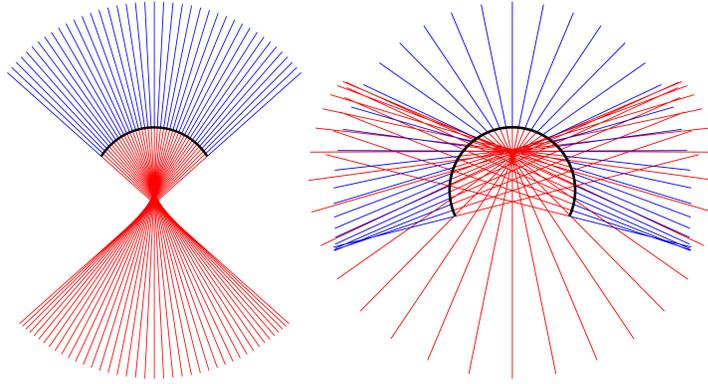


Figure 10.4: A 2D illustration of curved elements and their normal bundles. The element, their outward and inward normal bundles are drawn by black, blue, and red lines, respectively. Left: outward and inward normal bundles do not intersect. In this case, evaluation points in the inward normal bundle can avoid the singularity by flipping the parametrization. Right: outward and inward normal bundles intersect and if \mathbf{r}_p is in the intersection, the element has to be subdivided so that \mathbf{r}_p does not belong to an intersection of normal bundles.

Algorithm 3 EvaluateLayerPotential(S : element, \mathbf{r}_p : evaluation point)

1. Find $[\hat{h}_{\min}, \hat{h}_{\max}]$, the range of h/r over all points on the element.
 2. If $\hat{h}_{\min} = -1$ and $\hat{h}_{\max} = 1$, subdivide the element such that $S = \bigcup_i s_i$.
 $I := \sum_i \text{EvaluateLayerPotential}(s_i, \mathbf{r}_p)$. Return I .
 3. If $|\hat{h}_{\min} + 1| < |\hat{h}_{\max} - 1|$, flip the parametrization of S from $\mathbf{r}_q(u, v)$ to $\mathbf{r}_q(v, u)$.
 4. Evaluate the integral I using (10.19).
 5. If the target is the double layer potential and the parametrization was flipped in step 3., $I := -I$.
 6. Return I .
-

10.6 Numerical evaluation

10.6.1 Element-level tests

10.6.1.1 Nearly singular case

The method was tested for both the single and double layer potentials, for both the Laplace and Helmholtz kernels. Adaptive Gauss-Kronrod quadrature, implemented in QUADPACK [198], was used to compute the reference values of the layer potentials P_{GK} over a boundary element. The error tolerance of Gauss-Kronrod was set to 10^{-12} . The layer potentials were computed using two standard techniques: (1) two-dimensional Gauss-Legendre quadrature [207] (*GL2D*), (2) two-dimensional quadrature using the polar coordinate transform, i.e. nested one-dimensional Gauss-Legendre quadrature over R and θ (*GL2D(Polar)*), and with two methods using the proposed decomposition (10.7): (3) evaluating the Stokes term via one-dimensional Gauss-Legendre quadrature and the curvature term via two-dimensional Gauss-Legendre quadrature [207] (*Stokes+GL2D*), and (4) evaluating the Stokes term via one-dimensional Gauss-Legendre quadrature and the curvature term via polar coordinate transform i.e. nested one-dimensional Gauss-Legendre quadrature (*Stokes+GL2D(Polar)*). For the Laplace layer potentials, we also computed the integrals using the method described in [183] (*Stokes*) using 20th degree polynomials as the basis for the density function approximation. The computed potentials P were compared against the reference result in terms of the relative error $|P - P_{\text{GK}}|/|P_{\text{GK}}|$. 20th order Gauss-Legendre quadrature was used for all integrals. Curved triangles parametrized via $\mathbf{r}_q(u, v) = (u, v, f(u, v))^T$ with $f(u, v) = \sigma((u - 1/4)^2 + (v - 1/4)^2)$ were used as test

cases where $\sigma = \{-0.6, 0.6\}$. The test elements are referred to as *element 1* and *element 2* and are shown in Fig. 10.5. The evaluation point was placed at $(u_p, v_p, f(u_p, v_p) + |h|\mathbf{n}_p)^T$,

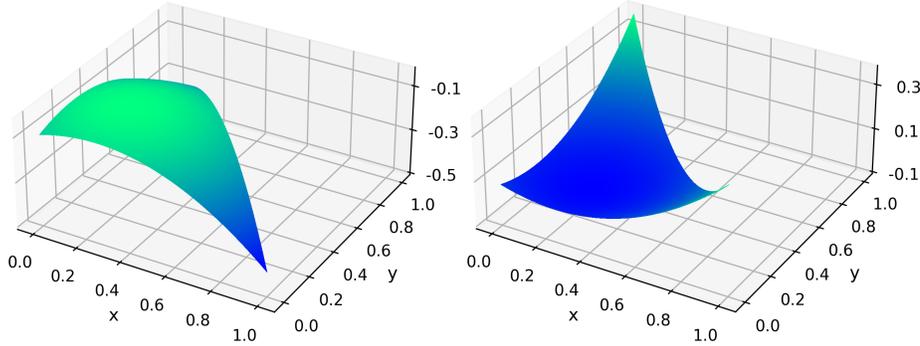


Figure 10.5: The manifold boundary elements with positive and negative curvature used in the numerical tests. *Element 1* (left, $\sigma = -0.6$) and *element 2* (right, $\sigma = 0.6$). The colors are for visual aid.

where \mathbf{n}_p is the unit normal vector at point $\mathbf{r}_q(u_p, v_p)$ with $(u_p, v_p) = (0.2, 0.3)$. Results are shown in Figures 10.6 and 10.7 for *elements 1* and *2*, respectively. It was found that the proposed decomposition improves the accuracy of the numerical results in the nearly singular regime $|h|/d < 1$ compared to baseline methods without the decomposition, with d the maximum inter-vertex distance of the element. In the Laplace double layer potential case, the Stokes-only approach by [183] (*Stokes*) was found to deliver better accuracy than the proposed method. However, larger error was observed for the single layer potential case using this method. This could be because of the approximation introduced in this method where the single layer potential is considered a double layer potential with a modified density function. See [183] for details on how this approximation is constructed. The method in [183], therefore, appears to be preferable for the double layer potential, while the present method may be preferable for the single layer potential as it is free from the type of approximation error introduced by the approach taken in [183]. Note that the condition $kd = 1$ for the Helmholtz case approximately corresponds to six wavelengths

per element, which is typically used as the maximum mesh size in boundary element analysis. A Python implementation of the proposed method exhibited computation times comparable to the baseline $GL2D(Polar)$ method, although optimized implementations via compiled languages should be used for more accurate performance evaluations.

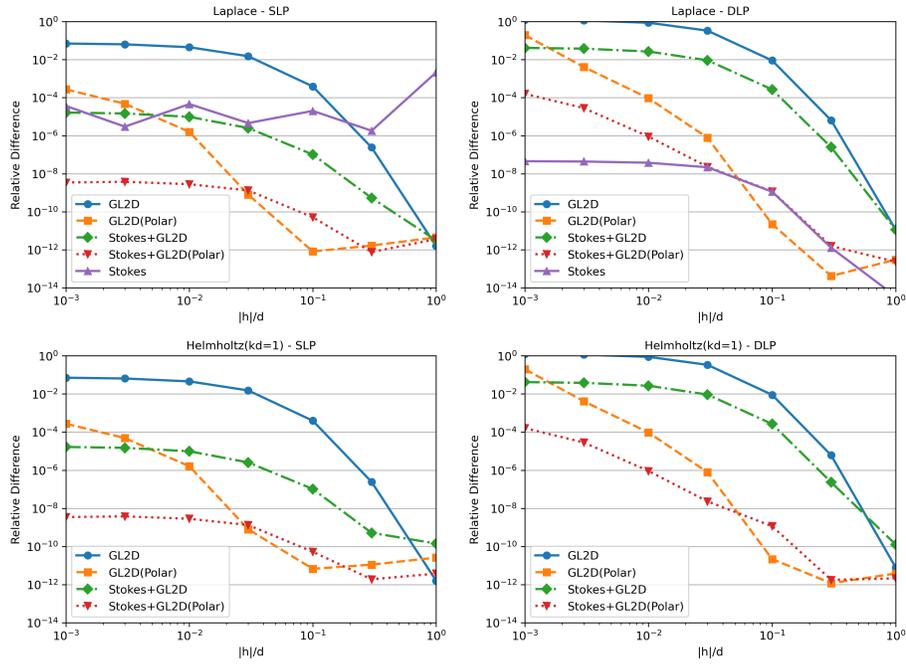


Figure 10.6: Relative difference of the layer potentials on *element 1* for various methods and the reference adaptive Gauss-Kronrod quadrature. Results for single layer potential with Laplace kernel (top left), double layer potential with Laplace kernel (top right), single layer potential with Helmholtz kernel (bottom left), and double layer potential with Helmholtz kernel (bottom right). All results are for the case with quadrature order set to 20.

10.6.1.2 Singular case

Singular cases with the evaluation point \mathbf{r}_p on the element can be handled with the proposed method. The only change to be made is that the singularity has to be excluded from the integration domain when applying Stokes' theorem in the double layer potential case. Technically, this results in the subtraction of the contribution of the singularity,

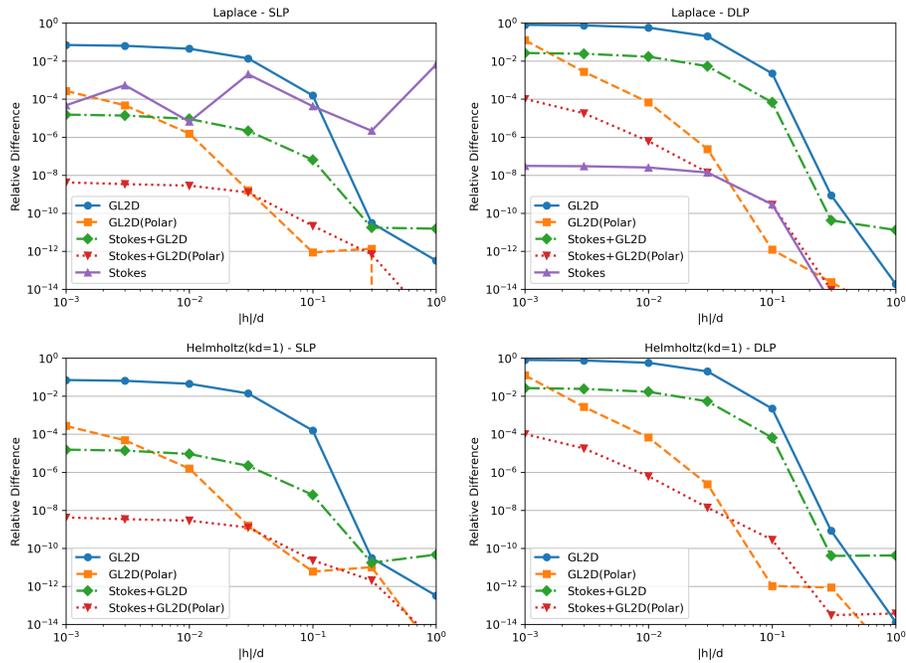


Figure 10.7: Relative difference of the layer potentials on *element 2* for various methods and the reference adaptive Gauss-Kronrod quadrature. Results for single layer potential with Laplace kernel (top left), double layer potential with Laplace kernel (top right), single layer potential with Helmholtz kernel (bottom left), and double layer potential with Helmholtz kernel (bottom right). All results are for the case with quadrature order set to 20.

which is a constant of $1/2$. The layer potentials evaluated by the proposed method (*Stokes+GL2D(Polar)* in the previous section) P_{prop} were compared with reference results P_{Gui} obtained by Guiggiani’s method [178], internally using Gauss-Legendre quadrature of 50th order. The quadrature order in the proposed method was varied from 10 to 40. The same quadrature order was used for the Stokes term and the curvature term. Element 1 from the previous section with \mathbf{r}_p on point $\mathbf{r}_q(0.2, 0.3)$ was used to compute the single and double layer potentials for the Laplace and Helmholtz kernels. Results in Fig. 10.8 show p -convergence and a good agreement with the reference at sufficiently high quadrature orders.

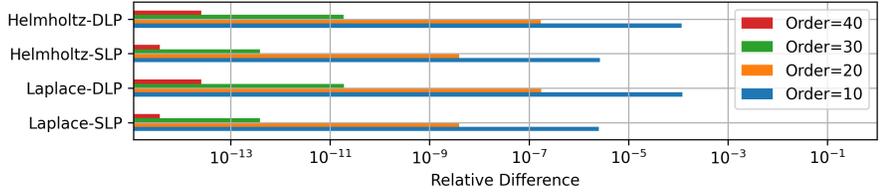


Figure 10.8: Relative difference $|P_{\text{prop}} - P_{\text{Gui}}|/|P_{\text{Gui}}|$ of the singular case layer potentials on *element 1* for the proposed method with various quadrature orders and the Guiggiani’s reference method.

10.6.2 Integrated BEM test: thin spherical cavity problem

The method was integrated in a prototype BEM solver and was evaluated by solving a benchmark problem where we consider an interior Helmholtz problem in a spherical cavity, which is illustrated in Fig. 10.9. In this setup, a rigid sphere of radius a is placed concentrically in a spherical room of radius b with rigid boundaries, and the upper half ($\theta < \pi/2$) of the internal sphere is vibrating at velocity v_0 . This problem can be solved analytically and the solution expressed in spherical coordinates (r, θ, ϕ) is given

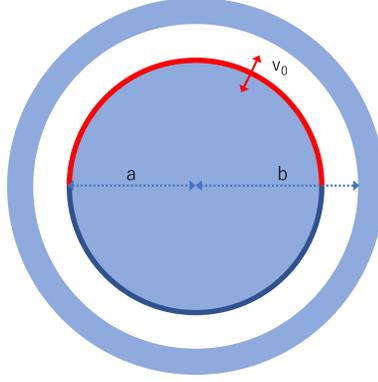


Figure 10.9: The cross section of the spherical cavity setup used as a benchmark problem. The white region in between the two spheres is the domain $\Omega = \{r|a \leq r \leq b\}$ subject to analysis. The vibrating surface is the upper hemisphere of the interior sphere and is indicated as the red arc.

as follows (see e.g. [156, 208]):

$$p(r, \theta, \phi) = \sum_{n=0}^{\infty} \alpha_n R_n^0(r, \theta, \phi) + \beta_n S_n^0(r, \theta, \phi),$$

$$R_n^m(r, \theta, \phi) = j_n(kr)Y_n^m(\theta, \phi), \quad S_n^m(r, \theta, \phi) = h_n(kr)Y_n^m(\theta, \phi), \quad (10.21)$$

$$\alpha_n = iv_0 c_s q \sqrt{(2n+1)\pi} \left(j_n'(ka) - \frac{j_n'(kb)}{h_n'(kb)} h_n'(ka) \right)^{-1}, \quad \beta_n = -\frac{j_n'(kb)}{h_n'(kb)} \alpha_n,$$

with j_n and h_n the spherical Bessel and Hankel function of the first kind, respectively, j_n' and h_n' the derivative of j_n and h_n with respect to the argument, respectively, Y_n^m the spherical harmonics, c_s the speed of sound, and q the density of the medium. We fixed $b = v_0 = 1$ and $k = 2$. The numerical solution to this problem was computed using BEM where the singular integrals, i.e. layer potentials with the evaluation point on the same element, were computed using the proposed *Stokes+GL2D(Polar)* method. Layer potentials with the evaluation point \mathbf{r}_p which satisfy $\|\mathbf{r}_p - \mathbf{r}_e\| < l_e$ with $\mathbf{r}_e = \mathbf{r}_q(1/3, 1/3)$ and l_e the maximum length of the straight line segments connecting the vertices of the element, were considered nearly singular and were evaluated using either Gauss-Legendre

quadrature (*GL2D*) [207] or the proposed *Stokes+GL2D(Polar)* method. The quadrature order was set to 20 for all singular and nearly singular integrals. The manifold elements were parametrized as:

$$\mathbf{r}_q(u, v) = s \frac{\tilde{\mathbf{r}}(u, v)}{|\tilde{\mathbf{r}}(u, v)|}, \quad \tilde{\mathbf{r}} = \mathbf{v}_1 + u(\mathbf{v}_2 - \mathbf{v}_1) + v(\mathbf{v}_3 - \mathbf{v}_1), \quad (10.22)$$

with s the radius of the spherical triangle and $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 the vertices of the spherical triangle. Note that this parametrization represents the surface exactly, hence the discretization error in the numerical solution for the manifold element case is solely due to the discretization of the function space. On the other hand, the flat polygon mesh introduces geometric approximation error. The numerical solution at the collocation points \mathbf{p}_{BEM} was compared against the analytical solution $\mathbf{p}_{\text{exact}}$ given by (10.21) in terms of the relative L_2 -norm of the difference vector $\|\mathbf{p}_{\text{BEM}} - \mathbf{p}_{\text{exact}}\|_2 / \|\mathbf{p}_{\text{exact}}\|_2$. We ran the experiments for two geometry representation conditions, where in one case the elements were represented as exact spherical triangles and in the other case polygon meshes with flat triangle elements were used to approximate the geometry of the spherical surfaces. In both representations, the boundary was represented by 3668 elements in total. The result is shown in Fig. 10.10. It was found that *GL2D* diverges from the analytical solution in the nearly singular regime, and that the numerical results using the exact spherical surface representation delivers up to about one order of magnitude smaller error compared to the flat element counterpart. The error in the curved element case is bound by the constant density approximation; expanding the surface density using higher order basis functions would further improve the accuracy.

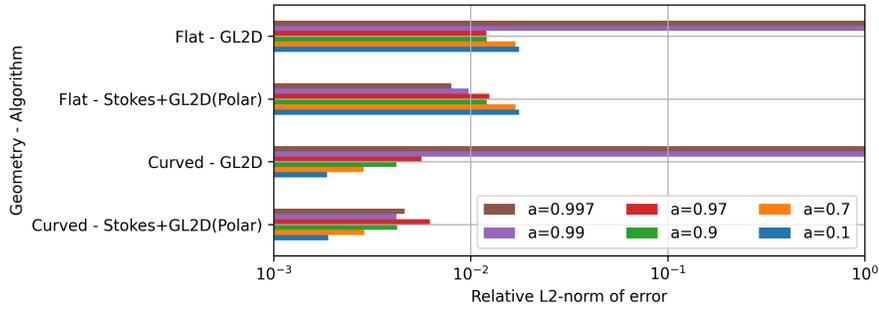


Figure 10.10: Results of the spherical cavity BEM test problem. Relative L_2 -norm of the error at the collocation points for two geometry conditions (flat or non-flat manifold elements) and two algorithms used for nearly singular integrals (Gauss-Legendre quadrature or the proposed method).

10.7 Conclusion

A method for the evaluation of nearly singular and singular integrals for single and double layer potentials over manifold boundary elements for Laplace and Helmholtz kernels was proposed. The method uses a novel decomposition of the layer potentials into an integral of a differential 2-form which can be reduced to a contour integral via Stokes' theorem and a second term related to the curvature of the element, which can be further regularized via a polar coordinate transform and integrated via existing quadrature methods. Numerical tests showed that the proposed method delivers accurate results in the nearly singular and singular regime where a naïve use of Gauss-Legendre quadrature is not effective. One of the benefits of the proposed method, which is shared with some modern techniques (e.g. [176, 185, 209]), is that it supports both nearly singular and singular layer potential integrals within the same framework. The proposed method covers both single and double layer potentials for both the Laplace and Helmholtz equations. While we introduced the method for constant densities in the present work, supporting higher order density functions is indeed desired for a boundary element method with

higher order accuracy and this is a natural next step. Nevertheless, constant elements are still useful in solving large scale problems with millions or billions of unknowns under limited compute resources and in applications where accuracy requirements are relatively relaxed. Generalizing the present method to other layer potentials, other kernels, and integrating it into a FMM-BEM solver are other directions for future work.

Chapter 11: Efficient exact quadrature of regular solid harmonics times polynomials over simplices in \mathbb{R}^3

A generalization of a recently introduced recursive numerical method [210] for the exact evaluation of integrals of regular solid harmonics and their normal derivatives over simplex elements in \mathbb{R}^3 is presented. The original *Quadrature to Expansion* (Q2X) method [210] achieves optimal per-element asymptotic complexity for computing $O(p_s^2)$ integrals of all regular solid harmonics bases with truncation degree p_s by exploiting recurrence and differential relations of the regular solid harmonics as well as the flatness and straightness of the faces and edges, respectively, of simplex elements. However, it considered only constant density functions over the elements. Here, we generalize this method to support arbitrary degree polynomial density functions, which is achieved in an extended recursive framework while maintaining the optimality of the per-element complexity for evaluating all regular solid harmonics and monomial density functions. The method is derived for 1- and 2- simplex elements in \mathbb{R}^3 and can be used for the boundary element method and vortex methods coupled with the fast multipole method.¹

¹The research presented in this chapter is currently under review for publication. Its preprint can be found in [192].

11.1 Introduction

Recently, Gumerov et al. introduced *Quadrature to Expansion (Q2X)*, a recursive method for the analytical evaluation of integrals of spherical basis functions for the Laplace equation in \mathbb{R}^3 , aka *solid harmonics*, over d -simplex elements for $d \in \{1, 2, 3\}$ with constant densities [210]. This method achieves exact integration of all bases up to truncation degree p_s with optimal complexity $O(p_s^2)$ per element for any d in $\{1, 2, 3\}$, by exploiting both (i) the flatness of the faces and/or the straightness of the edges and (ii) recurrence and differential relations of the special functions in the integrand. This is useful for the Boundary Element Method (BEM) [194] coupled with the Fast Multipole Method (FMM) [45], particularly for problems where the boundary can be accurately expressed via piecewise flat surfaces or straight lines. The conventional FMM performs an approximate summation of monopoles and dipoles centered at points \mathbf{r}_q distributed in space, by expanding them into truncated multipole expansions centered at points \mathbf{r}_p . Many such expansions are consolidated into one, in a hierarchical manner, to achieve efficiency. In the Q2X approach, recently introduced by Gumerov et al. [210], layer potential integrals over a surface or a line discretized via simplices were represented as such expansions, with the expansion coefficients, obtained by quadrature over the simplex, evaluated analytically via an optimal recursive procedure. Fig. 11.1 shows the geometrical relation of the elements with respect to the expansion center \mathbf{r}_* of the solid harmonics, which is typically the centroid of a cell in an octree data structure to which the elements belong. Recursive analytical methods for high order flat surface elements have been developed for the close evaluation of layer potential integrals [176], and such a method is also desired for the

integration of regular solid harmonics commonly arising in the FMM-BEM for the Laplace equation or vortex methods, typically in the first step of the far-field potential evaluation which is sometimes referred to as the “particle-to-multipole” step. [189] discusses a related method for quadrilateral elements, however a full algorithm for computing integrals for all harmonics and all monomial densities with finite degrees has not been described. Here, we extend the Q2X method to elements with polynomial densities which allows the evaluation of integrals for all regular solid harmonics up to degree p_s and all monomial density functions up to degree p_d with optimal complexity $O(p_s^2 p_d^d)$ per element. All formulation is done for simplex elements in \mathbb{R}^3 .

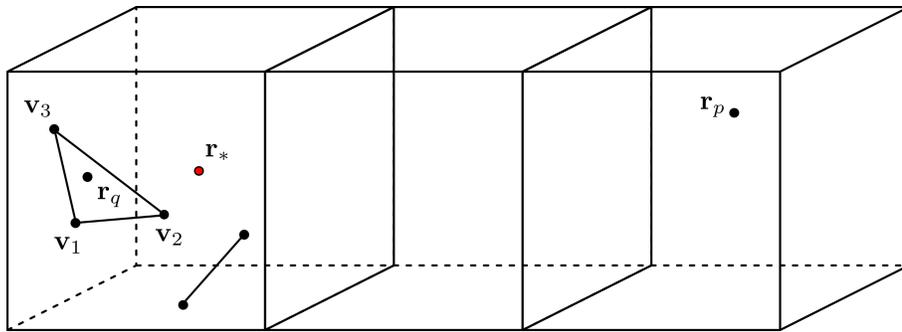


Figure 11.1: 2- and 1-simplex in a cell of an octree data structure commonly used in the FMM-BEM. The evaluation point r_p is located in a well-separated position outside the cell the elements reside.

11.2 Potential integrals in BEM and vortex methods

The BEM is widely used for numerical solution of partial differential equations, e.g. the Poisson equation:

$$-\nabla^2 u(\mathbf{r}) = f(\mathbf{r}), \quad \mathbf{r} \in \Omega \subset \mathbb{R}^3, \quad (11.1)$$

with field u in domain $\Omega \subset \mathbb{R}^3$, and source f ($= 0$ for Laplace). The weak form of (11.1) can be written in terms of single- and double layer potentials L, M [194]:

$$\begin{aligned} \{(c_p \gamma_{0,p} + M \gamma_{0,q} - L \gamma_{1,q})u\}(\mathbf{r}_p) &= \{N_0 f\}(\mathbf{r}_p), \\ \{L\sigma\}(\mathbf{r}_p) &\equiv \int_{\mathbf{r}_q \in \Gamma} G(\mathbf{r}_p, \mathbf{r}_q) \sigma(\mathbf{r}_q) d\Gamma, \quad \{M\sigma\}(\mathbf{r}_p) \equiv \int_{\mathbf{r}_q \in \Gamma} \frac{\partial G(\mathbf{r}_p, \mathbf{r}_q)}{\partial \mathbf{n}_q} \sigma(\mathbf{r}_q) d\Gamma, \end{aligned} \quad (11.2)$$

with $c_p = 1/2$ on a smooth boundary, \mathbf{n}_q the outward unit normal, γ_0 and γ_1 the boundary trace and normal derivative operators, respectively, N_0 the Newton potential operator, and $G(\mathbf{r}_p, \mathbf{r}_q)$ the Green function for the Laplace equation:

$$\begin{aligned} \{\gamma_{0,q}u\}(\mathbf{r}_q) &\equiv \lim_{\hat{\mathbf{r}}_q \in \Omega \rightarrow \mathbf{r}_q \in \Gamma} u(\hat{\mathbf{r}}_q), \quad \{\gamma_{1,q}u\}(\mathbf{r}_q) \equiv \mathbf{n}_q \cdot \nabla_q u(\mathbf{r}_q), \quad \mathbf{r}_q \in \Gamma = \partial\Omega, \\ \{N_0 f\}(\mathbf{r}_p) &= \int_{\mathbf{r}_q \in \Omega} G(\mathbf{r}_p, \mathbf{r}_q) f(\mathbf{r}_q) d\Omega, \quad G(\mathbf{r}_p, \mathbf{r}_q) = \frac{1}{4\pi r}, \quad r \equiv |\mathbf{r}_q - \mathbf{r}_p|, \quad \mathbf{r}_p \in \mathbb{R}^3. \end{aligned} \quad (11.3)$$

In the BEM the boundary Γ is discretized into boundary elements which can be either flat or curved surfaces. The layer potential integrals over these elements are evaluated to form the linear system of equations. The densities σ are approximated via local, typically polynomial, functions with unknown coefficients which must be determined. In the present work we assume the boundary $\Gamma = \partial\Omega$ is a union of boundary elements $\Gamma = \bigcup_i S_i$, where each S_i is a flat triangular element with polynomial density functions.

Similarly, in vortex methods the Bio-Savart integral is used to compute potentials

due to line elements Λ_q :

$$\begin{aligned} \mathbf{H}(\mathbf{r}_p) &= \int_{\mathbf{r}_q \in \Lambda_q} \frac{(\mathbf{I}_q \sigma(\mathbf{r}_q)) \times (\mathbf{r}_p - \mathbf{r}_q)}{4\pi |\mathbf{r}_p - \mathbf{r}_q|^3} d\Lambda_q = \nabla \times (\mathbf{I}_q K_q(\mathbf{r}_p)), \\ \{K_q \sigma\}(\mathbf{r}_p) &\equiv \int_{\mathbf{r}_q \in \Lambda_q} G(\mathbf{r}_p, \mathbf{r}_q) \sigma(\mathbf{r}_q) d\Lambda_q, \end{aligned} \quad (11.4)$$

where \mathbf{I}_q is the circulation of the q -th line element.

11.2.1 Multipole expansion of potentials

We summarize the formulation also used in [210]. The following definition of regular and singular spherical basis functions, $R_n^m(\mathbf{r})$ and $S_n^m(\mathbf{r})$ is accepted:

$$\begin{aligned} R_n^m(\mathbf{r}) &= \frac{(-1)^n i^{|m|}}{(n + |m|)!} r^n P_n^{|m|}(\cos \theta) e^{im\varphi}, \quad S_n^m(\mathbf{r}) = i^{-|m|} (n - |m|)! r^{-n-1} P_n^{|m|}(\cos \theta) e^{im\varphi}, \\ n &= 0, 1, 2, \dots, \quad m = -n, \dots, n, \quad \mathbf{r} = (x, y, z)^T = r (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)^T, \end{aligned} \quad (11.5)$$

respectively, where (r, θ, φ) are the spherical coordinates of \mathbf{r} and P_n^m are the associated Legendre functions [211],

$$P_n^m(\mu) = \frac{(-1)^m (1 - \mu^2)^{m/2}}{2^n n!} \frac{d^{m+n}}{d\mu^{m+n}} (\mu^2 - 1)^n, \quad (11.6)$$

for nonnegative m . R_n^m and S_n^m obey the symmetry:

$$R_n^{-m}(\mathbf{r}) = (-1)^m \overline{R_n^m(\mathbf{r})}, \quad S_n^{-m}(\mathbf{r}) = (-1)^m \overline{S_n^m(\mathbf{r})}, \quad (11.7)$$

where the bar indicates the complex conjugate. In these bases the Green's function can be expanded as

$$G(\mathbf{r}_p, \mathbf{r}_q) = \sum_{n=0}^{\infty} \frac{(-1)^n}{4\pi} \sum_{m=-n}^n R_n^{-m}(\mathbf{r}_q - \mathbf{r}_*) S_n^m(\mathbf{r}_p - \mathbf{r}_*), \quad (11.8)$$

given $|\mathbf{r}_p - \mathbf{r}_*| > |\mathbf{r}_q - \mathbf{r}_*|$ where \mathbf{r}_* is the center of expansion. The integrals of the spherical basis functions over element S_q, Λ_q are defined as:

$$\begin{aligned} L_n^m(\mathbf{r}_*) &\equiv \frac{(-1)^n}{4\pi} \int_{\mathbf{r}_q \in S_q} R_n^{-m}(\mathbf{r}_q - \mathbf{r}_*) \sigma(\mathbf{r}_q) dS_q, \\ M_n^m(\mathbf{r}_*) &\equiv \frac{(-1)^n}{4\pi} \mathbf{n}_q \cdot \int_{\mathbf{r}_q \in S_q} (\nabla R_n^{-m}(\mathbf{r}_q - \mathbf{r}_*)) \sigma(\mathbf{r}_q) dS_q, \\ K_n^m(\mathbf{r}_*) &\equiv \frac{(-1)^n}{4\pi} \int_{\mathbf{r}_q \in \Lambda_q} R_n^{-m}(\mathbf{r}_q - \mathbf{r}_*) \sigma(\mathbf{r}_q) d\Lambda_q, \end{aligned} \quad (11.9)$$

and are used to expand the layer potentials for S_q as:

$$\{F\sigma\}(\mathbf{r}_p) = \sum_{n=0}^{\infty} \sum_{m=-n}^n S_n^m(\mathbf{r}_p - \mathbf{r}_*) F_n^m(\mathbf{r}_*), \quad (11.10)$$

where $F = \{L, M, K\}$. R_n^m satisfies (see e.g., [212]):

$$\begin{aligned} \frac{\partial}{\partial \eta} R_n^m(\mathbf{r}) &= iR_{n-1}^{m+1}(\mathbf{r}), & \frac{\partial}{\partial \xi} R_n^m(\mathbf{r}) &= iR_{n-1}^{m-1}(\mathbf{r}), & \frac{\partial}{\partial z} R_n^m(\mathbf{r}) &= -R_{n-1}^m(\mathbf{r}), \\ \xi &\equiv \frac{x+iy}{2}, \quad \eta \equiv \frac{x-iy}{2}, & \partial_\eta &\equiv \frac{\partial}{\partial \eta} = \frac{\partial}{\partial x} + i\frac{\partial}{\partial y}, & \partial_\xi &\equiv \frac{\partial}{\partial \xi} = \frac{\partial}{\partial x} - i\frac{\partial}{\partial y}. \end{aligned} \quad (11.11)$$

Thus, with $\mathbf{n}_q = (n_x, n_y, n_z)^T$ we have:

$$\mathbf{n}_q \cdot \nabla R_n^m(\mathbf{r}) = i \frac{n_x}{2} [R_{n-1}^{m+1}(\mathbf{r}) + R_{n-1}^{m-1}(\mathbf{r})] + \frac{n_y}{2} [R_{n-1}^{m+1}(\mathbf{r}) - R_{n-1}^{m-1}(\mathbf{r})] - n_z R_{n-1}^m(\mathbf{r}). \quad (11.12)$$

11.3 Problem statement

For $d = 2$, we denote the vertices of the 2-simplex element as $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and use the parametrization:

$$\begin{aligned} \mathbf{r}_q(u, v) &= \mathbf{r}_u u + \mathbf{r}_v v + \mathbf{v}_1 = A(u, v, 0)^T + \mathbf{v}_1 = (x, y, z)^T, \\ \mathbf{r}_u &\equiv \frac{\partial \mathbf{r}_q}{\partial u}, \quad \mathbf{r}_v \equiv \frac{\partial \mathbf{r}_q}{\partial v}, \quad A \equiv (\mathbf{r}_u, \mathbf{r}_v, \mathbf{n})^T. \end{aligned} \quad (11.13)$$

In contrast to [210] which used constant parametrization, we extend the method to density functions σ defined over the elements which are polynomials of the form:

$$N(u, v) = \sum_{b=0, c=0}^{b+c \leq p_d} B_b^c N_b^c(u, v), \quad N_b^c(u, v) \equiv u^b v^c. \quad (11.14)$$

The integrand of the integrals L_n^m, M_n^m have the form:

$$Q_{n,b}^{m,c}(u, v) \equiv R_n^m(\mathbf{r}_q(u, v) - \mathbf{r}_*) N_b^c(u, v). \quad (11.15)$$

We set the origin of the coordinate system so that $\mathbf{r}_* = \mathbf{0}$. The goal is to derive an algorithm for efficient analytical evaluation of integrals (11.9) for $\sigma = N_b^c(u, v)$:

$$\begin{aligned} L_{n,b}^{m,c}(\mathbf{r}_*) &\equiv \frac{(-1)^n}{4\pi} \int_{\mathbf{r}_q \in S_q} R_n^{-m}(\mathbf{r}_q - \mathbf{r}_*) N_b^c(u, v) dS_q, \\ M_{n,b}^{m,c}(\mathbf{r}_*) &\equiv \frac{(-1)^n}{4\pi} \mathbf{n}_q \cdot \int_{\mathbf{r}_q \in S_q} (\nabla R_n^{-m}(\mathbf{r}_q - \mathbf{r}_*)) N_b^c(u, v) dS_q, \end{aligned} \quad (11.16)$$

for all index tuples (n, m, b, c) satisfying $0 \leq n \leq p_s$, $|m| \leq n$, $0 \leq b$, $0 \leq c$, and $b + c \leq p_d$. The coefficients $K_{n,b}^m$ for the case $d = 1$ can be expressed by setting $c = 0$, $\mathbf{r}_q = \mathbf{r}_u u + \mathbf{v}_1$, and replacing S_q by Λ_q in $L_{n,b}^{m,c}$.

11.4 Q2XP: Q2X for Polynomial elements

We first describe the case $d = 2$. The Q2X method [210] was derived by utilizing the fact that the regular spherical basis functions R_n^m are homogeneous polynomials of degree n of arguments (x, y, z) . Similarly, N_b^c are also homogeneous polynomials of (u, v) with degree $b + c$. Hence, by considering \mathbf{r} and (u, v) as functions of (ξ, η, z) , Euler's homogeneous function theorem gives:

$$\begin{aligned} nR_n^m(\mathbf{r}) &= (\xi\partial_\xi + \eta\partial_\eta + z\partial_z)R_n^m(\mathbf{r}), \quad (b+c)N_b^c(u, v) = (\xi\partial_\xi + \eta\partial_\eta + z\partial_z)N_b^c(u, v), \\ (n+b+c)Q_{n,b}^{m,c}(u, v) &= (\xi\partial_\xi + \eta\partial_\eta + z\partial_z)Q_{n,b}^{m,c}(u, v). \end{aligned} \quad (11.17)$$

With a_{st} the (s, t) entry of A^{-1} , $\alpha_s^{(\pm)} \equiv a_{s1} \pm ia_{s2}$, \mathbf{a}_s the s -th row of A^{-1} , and $\beta_s \equiv \mathbf{a}_s \cdot \mathbf{v}_1$

we obtain:

$$(n + b + c)Q_{n,b}^{m,c} = -b\beta_1 Q_{n,b-1}^{m,c} - c\beta_2 Q_{n,b}^{m,c-1} + \xi(iQ_{n-1,b}^{m-1,c} + bQ_{n,b-1}^{m,c} \alpha_1^{(-)} + cQ_{n,b}^{m,c-1} \alpha_2^{(-)}) \\ + \eta(iQ_{n-1,b}^{m+1,c} + bQ_{n,b-1}^{m,c} \alpha_1^{(+)} + cQ_{n,b}^{m,c-1} \alpha_2^{(+)}) + z(-Q_{n-1,b}^{m,c} + ba_{13}Q_{n,b-1}^{m,c} + ca_{23}Q_{n,b}^{m,c-1}). \quad (11.18)$$

We define

$$\xi \equiv \xi_u u + \xi_v v + \xi_0, \quad \eta \equiv \eta_u u + \eta_v v + \eta_0, \quad z \equiv z_u u + z_v v + z_0, \quad (11.19)$$

where the subscripts u and v denote partial derivatives with respect to these variables, and

$$\psi_{n,b}^{m,c} \equiv \int_0^1 \int_0^{1-u} Q_{n,b}^{m,c}(u, v) dv du, \quad q_{n,b}^{m,c} \equiv Q_{n,b}^{m,c}(1, 0), \quad (11.20) \\ h_{n,b}^{m,c}(u) \equiv Q_{n,b}^{m,c}(u, 1-u), \quad j_{n,b}^{m,c} \equiv \int_0^1 h_{n,b}^{m,c}(u) du.$$

11.4.1 Recursions

The Q2XP method is constructed by utilizing nested recurrence relations which we present in the following three lemmas. Using these recurrence relations we can build the integrals for the higher order harmonics and monomial basis functions using the lower order ones.

Lemma 6. *The coefficients $\psi_{n,b}^{m,c}$ satisfy the recursion:*

$$(n + b + c + 2)\psi_{n,b}^{m,c} = \xi_0 i \psi_{n-1,b}^{m-1,c} + \eta_0 i \psi_{n-1,b}^{m+1,c} - z_0 \psi_{n-1,b}^{m,c} + j_{n,b}^{m,c}. \quad (11.21)$$

Proof.

$$\begin{aligned}
u \frac{\partial Q_{n,b}^{m,c}}{\partial u} + v \frac{\partial Q_{n,b}^{m,c}}{\partial v} &= u (\xi_u \partial_\xi + \eta_u \partial_\eta + z_u \partial_z) Q_{n,b}^{m,c} + v (\xi_v \partial_\xi + \eta_v \partial_\eta + z_v \partial_z) Q_{n,b}^{m,c} \\
&= (u \xi_u + v \xi_v) \partial_\xi Q_{n,b}^{m,c} + (u \eta_u + v \eta_v) \partial_\eta Q_{n,b}^{m,c} + (u z_u + v z_v) \partial_z Q_{n,b}^{m,c} \\
&= (\xi - \xi_0) \partial_\xi Q_{n,b}^{m,c} + (\eta - \eta_0) \partial_\eta Q_{n,b}^{m,c} + (z - z_0) \partial_z Q_{n,b}^{m,c} \\
&= \xi \frac{\partial Q_{n,b}^{m,c}}{\partial \xi} + \eta \frac{\partial Q_{n,b}^{m,c}}{\partial \eta} + z \frac{\partial Q_{n,b}^{m,c}}{\partial z} - \left(\xi_0 \frac{\partial Q_{n,b}^{m,c}}{\partial \xi} + \eta_0 \frac{\partial Q_{n,b}^{m,c}}{\partial \eta} + z_0 \frac{\partial Q_{n,b}^{m,c}}{\partial z} \right) \\
&= (n + b + c) Q_{n,b}^{m,c} + b \beta_1 Q_{n,b-1}^{m,c} + c \beta_2 Q_{n,b}^{m,c-1} - \xi_0 (i Q_{n-1,b}^{m-1,c} + b Q_{n,b-1}^{m,c} \alpha_1^{(-)} + c Q_{n,b}^{m,c-1} \alpha_2^{(-)}) \\
&\quad - \eta_0 (i Q_{n-1,b}^{m+1,c} + b Q_{n,b-1}^{m,c} \alpha_1^{(+)} + c Q_{n,b}^{m,c-1} \alpha_2^{(+)}) - z_0 (-Q_{n-1,b}^{m,c} + b a_{13} Q_{n,b-1}^{m,c} + c a_{23} Q_{n,b}^{m,c-1}).
\end{aligned} \tag{11.22}$$

We integrate both sides over the unit triangle $T = \{(u, v) | 0 \leq u, 0 \leq v, u + v \leq 1\}$ and applying integration by parts. The left hand side gives:

$$\begin{aligned}
&\int_{v=0}^1 \int_{u=0}^{1-v} u \frac{\partial Q_{n,b}^{m,c}}{\partial u} + v \frac{\partial Q_{n,b}^{m,c}}{\partial v} dudv = \int_{v=0}^1 \int_{u=0}^{1-v} u \frac{\partial Q_{n,b}^{m,c}}{\partial u} dudv + \int_{v=0}^1 \int_{u=0}^{1-v} v \frac{\partial Q_{n,b}^{m,c}}{\partial v} dudv \\
&= \int_{v=0}^1 \int_{u=0}^{1-v} u \frac{\partial Q_{n,b}^{m,c}}{\partial u} dudv + \int_{u=0}^1 \int_{v=0}^{1-u} v \frac{\partial Q_{n,b}^{m,c}}{\partial v} dvdu \\
&= \int_{v=0}^1 \left\{ [u Q_{n,b}^{m,c}]_{u=0}^{1-v} - \int_{u=0}^{1-v} Q_{n,b}^{m,c} du \right\} dv + \int_{u=0}^1 \left\{ [v Q_{n,b}^{m,c}]_{v=0}^{1-u} - \int_{v=0}^{1-u} Q_{n,b}^{m,c} dv \right\} du \\
&= \int_{v=0}^1 (1-v) Q_{n,b}^{m,c}(1-v, v) dv + \int_{u=0}^1 (1-u) Q_{n,b}^{m,c}(u, 1-u) du - 2 \int_T Q_{n,b}^{m,c} dT \\
&= \int_{u=0}^1 u Q_{n,b}^{m,c}(u, 1-u) du + \int_{u=0}^1 (1-u) Q_{n,b}^{m,c}(u, 1-u) du - 2 \psi_{n,b}^{m,c} \\
&= \int_{u=0}^1 h_{n,b}^{m,c}(u) du - 2 \psi_{n,b}^{m,c} = j_{n,b}^{m,c} - 2 \psi_{n,b}^{m,c}.
\end{aligned} \tag{11.23}$$

The right hand side gives:

$$\begin{aligned}
& (n + b + c)\psi_{n,b}^{m,c} + b\beta_1\psi_{n,b-1}^{m,c} + c\beta_2\psi_{n,b}^{m,c-1} - \xi_0(i\psi_{n-1,b}^{m-1,c} + b\psi_{n,b-1}^{m,c}\alpha_1^{(-)} + c\psi_{n,b}^{m,c-1}\alpha_2^{(-)}) \\
& - \eta_0(i\psi_{n-1,b}^{m+1,c} + b\psi_{n,b-1}^{m,c}\alpha_1^{(+)} + c\psi_{n,b}^{m,c-1}\alpha_2^{(+)}) - z_0(-\psi_{n-1,b}^{m,c} + ba_{13}\psi_{n,b-1}^{m,c} + ca_{23}\psi_{n,b}^{m,c-1}).
\end{aligned} \tag{11.24}$$

Equating the two sides yields:

$$\begin{aligned}
& (n + b + c + 2)\psi_{n,b}^{m,c} = \\
& \xi_0(i\psi_{n-1,b}^{m-1,c} + b\psi_{n,b-1}^{m,c}\alpha_1^{(-)} + c\psi_{n,b}^{m,c-1}\alpha_2^{(-)}) + \eta_0(i\psi_{n-1,b}^{m+1,c} + b\psi_{n,b-1}^{m,c}\alpha_1^{(+)} + c\psi_{n,b}^{m,c-1}\alpha_2^{(+)}) \\
& + z_0(-\psi_{n-1,b}^{m,c} + ba_{13}\psi_{n,b-1}^{m,c} + ca_{23}\psi_{n,b}^{m,c-1}) - b\beta_1\psi_{n,b-1}^{m,c} - c\beta_2\psi_{n,b}^{m,c-1} + j_{n,b}^{m,c} \\
& = \xi_0i\psi_{n-1,b}^{m-1,c} + \eta_0i\psi_{n-1,b}^{m+1,c} - z_0\psi_{n-1,b}^{m,c} + j_{n,b}^{m,c},
\end{aligned} \tag{11.25}$$

where we have used the following relations in the last line:

$$\begin{aligned}
& \xi_0 = \frac{v_{1x} + iv_{1y}}{2}, \quad \eta_0 = \frac{v_{1x} - iv_{1y}}{2}, \quad z_0 = v_{1z}, \quad \mathbf{v}_1 \equiv (v_{1x}, v_{1y}, v_{1z})^T, \\
& \xi_0(a_{s1} - ia_{s2}) + \eta_0(a_{s1} + ia_{s2}) + z_0a_{s3} - \mathbf{a}_s \cdot \mathbf{v}_1 \\
& = v_{1x}a_{s1} + v_{1y}a_{s2} + v_{1z}a_{s3} - \mathbf{a}_s \cdot \mathbf{v}_1 = 0 \quad (s = 1, 2)
\end{aligned} \tag{11.26}$$

□

Lemma 7. *The coefficients $j_{n,b}^{m,c}$ satisfy the recursion:*

$$(n + b + c + 1)j_{n,b}^{m,c} = (\xi_0 + \xi_v)ij_{n-1,b}^{m-1,c} + (\eta_0 + \eta_v)ij_{n-1,b}^{m+1,c} - (z_0 + z_v)j_{n-1,b}^{m,c} + cj_{n,b}^{m,c-1} + q_{n,b}^{m,c}. \tag{11.27}$$

Proof. Replacing $Q_{n,b}^{m,c}(u, v)$ in (11.22) by $h_{n,b}^{m,c}(u)$, integrating both sides over domain

$u \in [0, 1]$, and applying integration by parts yields:

$$\begin{aligned}
& (n + b + c + 1)j_{n,b}^{m,c} = \\
& (\xi_0 + \xi_v)(ij_{n-1,b}^{m-1,c} + bj_{n,b-1}^{m,c}\alpha_1^{(-)} + cj_{n,b}^{m,c-1}\alpha_2^{(-)}) + (\eta_0 + \eta_v)(ij_{n-1,b}^{m+1,c} + bj_{n,b-1}^{m,c}\alpha_1^{(+)} + cj_{n,b}^{m,c-1}\alpha_2^{(+)}) \\
& + (z_0 + z_v)(-j_{n-1,b}^{m,c} + ba_{13}j_{n,b-1}^{m,c} + ca_{23}j_{n,b}^{m,c-1}) - b\beta_1j_{n,b-1}^{m,c} - c\beta_2j_{n,b}^{m,c-1} + q_{n,b}^{m,c}.
\end{aligned} \tag{11.28}$$

Using (11.26) and

$$\xi_v(a_{s1} - ia_{s2}) + \eta_v(a_{s1} + ia_{s2}) + z_v a_{s3} = \mathbf{a}_s \cdot \mathbf{b}_2 = \delta_{s,2} \quad (s = 1, 2), \tag{11.29}$$

where \mathbf{b}_t is the t -th column of matrix A and $\delta_{s,t}$ is Kronecker's delta, yields (11.27). \square

Lemma 8. *The coefficients $q_{n,b}^{m,c}$ satisfy the recursion:*

$$(n + b + c)q_{n,b}^{m,c} = (\xi_0 + \xi_u)iq_{n-1,b}^{m-1,c} + (\eta_0 + \eta_u)iq_{n-1,b}^{m+1,c} - (z_0 + z_u)q_{n-1,b}^{m,c} + bq_{n,b-1}^{m,c} \tag{11.30}$$

Proof. Substituting $u = 1$ and $v = 0$ to (11.18) yields:

$$\begin{aligned}
& (n + b + c)q_{n,b}^{m,c} = \\
& (\xi_0 + \xi_u)(iq_{n-1,b}^{m-1,c} + bq_{n,b-1}^{m,c}\alpha_1^{(-)} + cq_{n,b}^{m,c-1}\alpha_2^{(-)}) + (\eta_0 + \eta_u)(iq_{n-1,b}^{m+1,c} + bq_{n,b-1}^{m,c}\alpha_1^{(+)} + cq_{n,b}^{m,c-1}\alpha_2^{(+)}) \\
& + (z_0 + z_u)(-q_{n-1,b}^{m,c} + ba_{13}q_{n,b-1}^{m,c} + ca_{23}q_{n,b}^{m,c-1}) - b\beta_1q_{n,b-1}^{m,c} - c\beta_2q_{n,b}^{m,c-1}.
\end{aligned} \tag{11.31}$$

Using (11.26) and

$$\xi_u(a_{s1} - ia_{s2}) + \eta_u(a_{s1} + ia_{s2}) + z_u a_{s3} = \mathbf{a}_s \cdot \mathbf{b}_1 = \delta_{s,1} \quad (s = 1, 2), \tag{11.32}$$

yields (11.30). □

11.4.2 Starting values for the recursions

Note that in all recursions $\psi_{n,b}^{m,c}$, $j_{n,b}^{m,c}$, and $q_{n,b}^{m,c}$ should be set to zero for $|m| > n$.

This follows from the fact that $R_n^m(\mathbf{r}) = 0$ for $|m| > n$. The starting values are:

$$q_{0,b}^{0,c} = \delta_{c,0}, \quad j_{0,b}^{0,c} = \int_0^1 u^b (1-u)^c du \equiv \kappa_{b,c} \quad (11.33)$$

where table $\kappa_{b,c}$ can be computed using recursion:

$$\kappa_{b,c} = \frac{c}{b+1} \kappa_{b+1,c-1} \quad (c > 0), \quad \kappa_{b,0} = \frac{1}{b+1} \quad (c = 0). \quad (11.34)$$

Lastly, the starting values for $\psi_{0,b}^{0,c}$ are given by:

$$\psi_{0,b}^{0,c} = \int_0^1 \int_0^{1-u} u^b v^c dv du = \int_0^1 u^b \left[\frac{v^{c+1}}{c+1} \right]_0^{1-u} du = \frac{1}{c+1} \int_0^1 u^b (1-u)^{c+1} du = \frac{\kappa_{b,c+1}}{c+1}. \quad (11.35)$$

11.4.3 Algorithm

The Q2XP method follows from the nested recurrence relations (11.30), (11.27), and (11.21), together with the fact that the expansion coefficients $L_{n,b}^{m,c}$ and $M_{n,b}^{m,c}$ can be

computed from the $\psi_{n,b}^{m,c}$ coefficients using $J = |\mathbf{r}_u \times \mathbf{r}_v|$:

$$\begin{aligned}
L_{n,b}^{m,c} &= \frac{J}{4\pi} (-1)^n \psi_{n,b}^{-m,c}, & M_{n,b}^{m,c} &= \frac{J}{4\pi} (-1)^n l_{n,b}^{-m,c}, \\
l_{n,b}^{-m,c} &= i \frac{n_x}{2} [\psi_{n-1,b}^{-m+1,c} + \psi_{n-1,b}^{-m-1,c}] + \frac{n_y}{2} [\psi_{n-1,b}^{-m+1,c} - \psi_{n-1,b}^{-m-1,c}] - n_z \psi_{n-1,b}^{-m,c}.
\end{aligned} \tag{11.36}$$

The pseudocode of the method is listed in algorithm 4.

Algorithm 4 Evaluate multipole expansion coefficients $L_{n,b}^{m,c}$ and $M_{n,b}^{m,c}$ for $0 \leq n \leq p_s$, $|m| \leq n$, $0 \leq b + c \leq p_d$

1. Compute the starting values $q_{0,b}^{0,c}$, $j_{0,b}^{0,c}$, and $\psi_{0,b}^{0,c}$ using (11.33) and (11.35).
 2. Compute coefficients $q_{n,b}^{m,c}$ using recursion (11.30).
 3. Compute coefficients $j_{n,b}^{m,c}$ using recursion (11.27) and the $q_{n,b}^{m,c}$ coefficients computed in the previous step.
 4. Compute coefficients $\psi_{n,b}^{m,c}$ using recursion (11.21) and the $j_{n,b}^{m,c}$ coefficients computed in the previous step.
 5. Compute coefficients $L_{n,b}^{m,c}$ and $M_{n,b}^{m,c}$ using (11.36).
-

Remark. *Q2XP achieves optimal per-element complexity $O(p_s^2 p_d^d)$ for evaluating all index tuples in question $((n, m, b, c)$ for $d = 2$ and (n, m, b) for $d = 1$) which contains all the integrals associated with $O(p_d^d)$ monomials. An arbitrary set of $O(p_d^d)$ density polynomials can be computed as a linear combination of the complete set of p_d -th degree monomials, therefore the total cost for computing all the integrals needed for a boundary element with p_d -th degree polynomials density functions is $O(p_s^2 p_d^{2d})$. A naive application of Gauss-Legendre quadrature requires $O(p_s^2 p_d^d (p_s + p_d)^d)$ for exact evaluation, indicating that the proposed method has lower asymptotic complexity.*

Remark. *The case $d = 1$ can be easily obtained by minor modifications to the case $d = 2$, i.e. setting $c = 0$, $\xi_v = \eta_v = z_v = 0$, $J = |\mathbf{r}_u|$, modifying the recursions accordingly,*

skipping step 4 and 5 in algorithm 4, and computing the result by:

$$K_{n,b}^m = \frac{J}{4\pi} (-1)^n j_{n,b}^{-m,0}. \quad (11.37)$$

11.5 Numerical experiment

Q2XP was tested for the case $d = 2$ on a configuration given by: $\mathbf{r}_c = (\sqrt{3}/2, 0, 0)^T$, $\mathbf{v}_1 = \mathbf{r}_c + r_t (1, 0, 0)^T$, $\mathbf{v}_{2,3} = \mathbf{r}_c + r_t (-1/2, \pm\sqrt{3}/2, 0)^T$, and $r_t = 0.1$. The wall clock times for running a Python implementation of Q2XP for polynomial degrees of up to $p_s = p_d = 20$ are shown in Fig. 11.2. Given a complexity expression $t = Cp_s^\alpha p_d^\beta$ with t the computation time and C a constant, the exponents are extracted as $\alpha = 1.75$ and $\beta = 1.83$ from the numerical experiment.

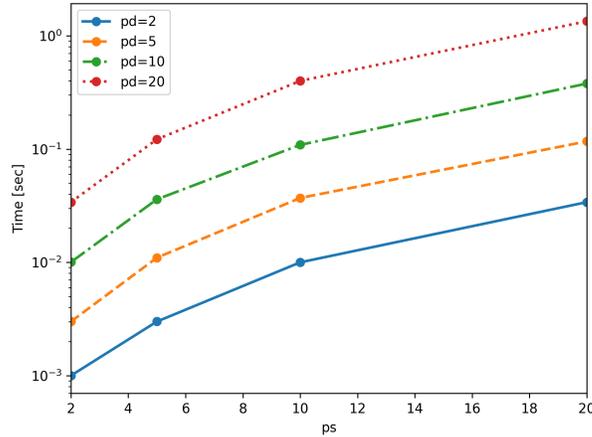


Figure 11.2: Wall clock times for running Q2XP up to $p_s = p_d = 20$.

Fig. 11.3 shows the maximum relative difference between the $L_{n,b}^{m,c}$ and $M_{n,b}^{m,c}$ coefficients for truncation degrees up to $p_s = p_d = 10$ computed by Q2XP and exact Gauss-Legendre quadrature. The largest maximum relative difference over all computed $L_{n,b}^{m,c}$ and $M_{n,b}^{m,c}$ coefficients was 3.5×10^{-14} , indicating good accuracy.

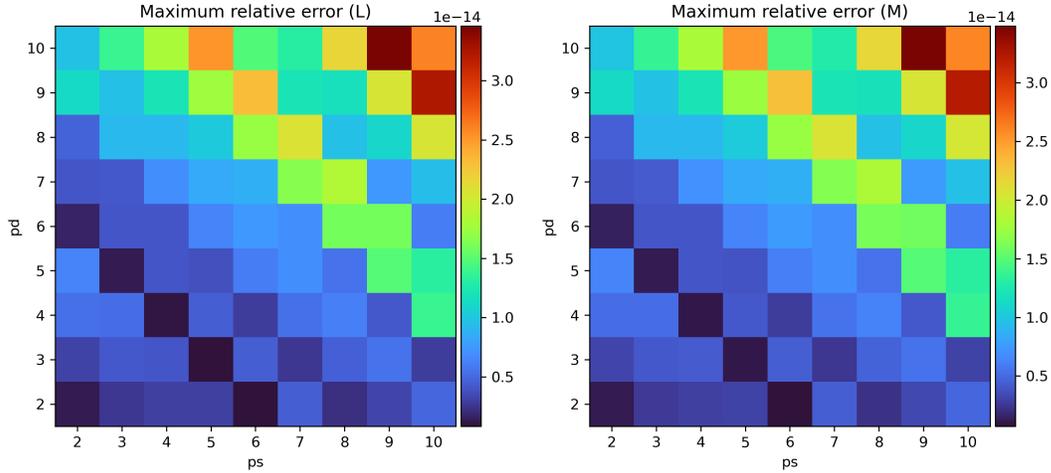


Figure 11.3: Maximum relative differences between the integrals computed by Q2XP and exact Gauss-Legendre quadrature for various p_d and p_s values. Left: L integrals and right: M integrals. Good accuracy is observed in all conditions.

11.6 Conclusion

We have extended the recursive algorithm presented in [210] for the analytical evaluation of integrals of spherical basis functions over d -simplex elements arising in the FMM-BEM for the Laplace equation in \mathbb{R}^3 to the case of high order polynomial density functions. All the integrals are evaluated analytically up to spherical basis degree of p_s and the density's monomial degree of p_d with optimal complexity $O(p_s^2 p_d^d)$. This complexity was confirmed via numerical experiments. While we limited the discussion to $d \in \{1, 2\}$, the case $d = 3$ could also be supported by following the formulation presented here and in [210]. The proposed method, as well as the previous Q2X method, has been introduced by exploiting the flatness and straightness of the faces and edges, respectively, of simplex elements. Quadrature methods over general curved elements have been developed for the close evaluation of layer potentials [183, 200]. Developing efficient methods for integrating spherical basis functions over general curved elements

with high-order densities is a direction for future development. Generalizing the present method to other kernels and integrating it into a FMM-BEM solver are other directions for future work.

Chapter 12: Conclusion and outlook

In this dissertation I tackled problems ranging from fundamental numerical methods to applied algorithms arising in the field of audio and acoustics.

Two methods generalizing ambisonics-based spatial sound field capture have been developed for expanding and changing the shape of reconstruction sweet spots. A fast method for realistic forest reverberation synthesis has been developed and was used for simulating distributed microphone array recordings of forest sound scenes for studying bird sound localization systems deployed in large reverberant forests. An individual-agnostic neural network model for binaural localization was developed, which revealed that such neural network-based binaural localizers experience the same “cone of confusion” as humans do. Sound field control methods for open personal sound fields have been developed. This dissertation was focused on the development of algorithms and their method of evaluation was limited to numerical experiments. As some of the developed methods are meant for audio applications for human listeners, it is desired to evaluate these methods via formal subjective listening experiments. For that, it is also desired to implement these methods as audio devices operating in the real world. These topics are left for future research.

Three numerical algorithms for evaluating integrals arising in the fast multipole-

accelerated BEM for the Helmholtz and Laplace equation have been developed for flat high order or curved constant boundary elements. These numerical quadrature methods are not limited to acoustics but can be also applied to problems in other domains, e.g. electrostatics or fluid dynamics. These methods, however, also have some limitations in their current form which suggest opportunities for future research. Currently, the RIPE method (chapter 9) and the Q2XP method (chapter 11) support arbitrarily high order polynomial densities over the elements, however these methods are limited to flat boundary elements. Conversely, the method for manifold elements discussed in chapter 10 supports curved boundary elements with high generality, however it is limited to constant densities and also to single and double layer potentials in its current form. Therefore, generalizing the RIPE method and Q2XP method to curved elements, or, the method for manifold elements to high-order densities may be possible approaches towards an algorithm which supports high-order curved elements. While the algorithm developed in [183] in principle supports high-order curved elements, in its current form it only supports Laplace kernels and its empirical accuracy on single layer potentials, as observed in chapter 10, may not be as satisfying as that for the double layer potential, implying that room for improvement remains. Indeed, classical methods for layer potential quadrature, e.g. Gauss-Legendre quadrature, various methods for singular, hypersingular, or nearly singular do exist and can be used for certain cases, however they lack a mechanism which allows efficient evaluation of high-order layer potentials via exploiting recurrence relations. Hence, the development of a numerical quadrature algorithm for the evaluation of general layer potentials arising in the BEM with high-order densities over curved elements which offers both efficiency and accuracy is still an open problem. We hope

that the algorithms presented in this dissertation serve as steps towards such an algorithm.

Appendix A: Derivation of the spheroidal ambisonics transcoding formula

The following relation is utilized for the derivation of the transcoding formula:

$$S_{mn}(c, \eta) R_{mn}^{(1)}(c, \xi) = \sum_{r=0}^{\infty} \delta_{(n-m)\%2, r\%2} i^{m-n+r} d_r^{mn}(c) j_{m+r}(kr) P_{m+r}^m(\cos \theta), \quad (\text{A.1})$$

where $d_r^{mn}(c)$ are the expansion coefficients:

$$S_{mn}(c, \eta) = \sum_{r=0}^{\infty} \delta_{(n-m)\%2, r\%2} d_r^{mn}(c) P_{m+r}^m(\eta), \quad (\text{A.2})$$

By substituting (A.1) into (3.15), the expression of p_{in} , we obtain:

$$\begin{aligned} p_{\text{in}} &= \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{r=0}^{\infty} d_r^{mn}(c) P_{m+r}^m(\cos \theta) j_{m+r}(kr) \delta_{(n-m)\%2, r\%2} i^{m-n+r} \quad (\text{A.3}) \\ &\quad \times (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi) \\ &= \sum_{n=0}^{\infty} \sum_{m=0}^n (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi) \\ &\quad \times \sum_{r=0}^{\infty} d_r^{mn}(c) P_{m+r}^m(\cos \theta) j_{m+r}(kr) \delta_{(n-m)\%2, r\%2} i^{m-n+r} \\ &= \sum_{n=0}^{\infty} \sum_{m=0}^n (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi) \\ &\quad \times \sum_{r=0}^{\infty} d_r^{mn}(c) P_{m+r}^m(\cos \theta) j_{m+r}(kr) \delta_{n\%2, (m+r)\%2} i^{m-n+r} \end{aligned}$$

$$\begin{aligned}
&\approx \sum_{n=0}^N \sum_{m=0}^n (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi) \\
&\quad \times \sum_{r=0}^{N-m} d_r^{mn}(c) \delta_{n\%2, (m+r)\%2} j_{m+r}(kr) P_{m+r}^m(\cos \theta) i^{m-n+r}
\end{aligned}$$

We can also write:

$$\begin{aligned}
p_{\text{in}} &= \sum_{n=0}^{\infty} \sum_{m=-n}^n \mathcal{A}_n^m j_n(kr) Y_n^m(\theta, \varphi) \\
&= \sum_{n=0}^{\infty} \sum_{m=-n}^n \mathcal{A}_n^m(k) j_n(kr) N_n^m P_n^m(\cos \theta) e^{im\varphi} \\
&= \sum_{n=0}^{\infty} \sum_{m=-n}^n \mathcal{A}_n^m(k) j_n(kr) N_n^m P_n^m(\cos \theta) (\cos(m\varphi) + i \sin(m\varphi)) \\
&\approx \sum_{n=0}^N j_n(kr) \sum_{m=-n}^n \mathcal{A}_n^m(k) (\cos(m\varphi) + i \sin(m\varphi)) N_n^m P_n^m(\cos \theta)
\end{aligned} \tag{A.4}$$

with

$$N_n^m \equiv \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}}. \tag{A.5}$$

Multiplying $Y_{n'}^{m'}(\theta, \varphi)^*$ with (A.4) and integrating over the unit sphere yields:

$$\begin{aligned}
\mathcal{A}_{n'}^{m'}(k) &= \frac{1}{j_{n'}(kr)} \int p_i Y_{n'}^{m'}(\theta, \varphi)^* d\Omega \\
&= \frac{1}{j_{n'}(kr)} \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} p_i Y_{n'}^{m'}(\theta, \varphi)^* \sin \theta d\theta d\varphi \\
&= \frac{1}{j_{n'}(kr)} \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{r=0}^{\infty} d_r^{mn}(c) P_{m+r}^m(\cos \theta) j_{m+r}(kr) \\
&\quad \times \delta_{(n-m)\%2, r\%2} i^{m-n+r} (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi) Y_{n'}^{m'}(\theta, \varphi)^* \sin \theta d\theta d\varphi \\
&= \frac{1}{j_{n'}(kr)} \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{r=0}^{\infty} d_r^{mn}(c) \delta_{(n-m)\%2, r\%2} i^{m-n+r} j_{m+r}(kr) N_{n'}^{m'} \\
&\quad \times \int_{\theta=0}^{\pi} P_{m+r}^m(\cos \theta) P_{n'}^{m'}(\cos \theta) \sin \theta d\theta
\end{aligned} \tag{A.6}$$

$$\times \int_{\varphi=0}^{2\pi} (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi) e^{-im'\varphi} d\varphi$$

Now, we define:

$$\begin{aligned}
F(m, n, m') &\equiv \int_{\varphi=0}^{2\pi} (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi) e^{-im'\varphi} d\varphi \\
&= \begin{cases} 2\pi A_{mn} & \text{if } m = m' = 0 \\ \pi (A_{mn} + iB_{mn}) & \text{else if } m = -m' \\ \pi (A_{mn} - iB_{mn}) & \text{else if } m = m' \\ 0 & \text{else} \end{cases} \quad (\text{A.7}) \\
&= F(m, n, m') \delta(m, |m'|) \\
&= F(|m'|, n, m') \delta(m, |m'|)
\end{aligned}$$

In the case $\delta(m, |m'|) = 1$, we also define:

$$\begin{aligned}
G(m, m', n', r) &\equiv \int_{\theta=0}^{\pi} P_{m+r}^m(\cos \theta) P_{n'}^{m'}(\cos \theta) \sin \theta d\theta \\
&= \begin{cases} \frac{2(n'+m)!}{(2n'+1)(n'-m)!} \delta_{m+r, n'} & \text{for } m = m' \\ \frac{2(-1)^m}{(2n'+1)} \delta_{m+r, n'} & \text{for } m = -m' \end{cases} \quad (\text{A.8}) \\
&= G(m, m', n', r) \delta(m + r, n') \\
&= G(m, m', n', n' - m) \delta(m + r, n')
\end{aligned}$$

This yields:

$$\begin{aligned}
\mathcal{A}_{n'}^{m'}(k) &= \frac{1}{j_{n'}(kr)} \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{r=0}^{\infty} d_r^{mn}(c) \delta_{(n-m)\%2, r\%2} i^{m-n+r} j_{m+r}(kr) N_{n'}^{m'} \\
&\quad \times \int_{\theta=0}^{\pi} P_{m+r}^m(\cos \theta) P_{n'}^{m'}(\cos \theta) \sin \theta d\theta \times \int_{\varphi=0}^{2\pi} (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi) e^{-im'\varphi} d\varphi \\
&= \frac{1}{j_{n'}(kr)} \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{r=0}^{\infty} d_r^{mn}(c) \delta_{(n-m)\%2, r\%2} i^{m-n+r} j_{m+r}(kr) N_{n'}^{m'} \\
&\quad \times G(m, m', n', n' - m) \delta(r, n' - m) F(|m'|, n, m') \delta(m, |m'|) \\
&= \frac{1}{j_{n'}(kr)} \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{r=0}^{\infty} d_{n'-m}^{mn}(c) \delta_{n\%2, n'\%2} i^{m-n+r} j_{n'}(kr) N_{n'}^{m'} \\
&\quad \times G(m, m', n', n' - m) \delta(r, n' - m) F(|m'|, n, m') \delta(m, |m'|) \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^n d_{n'-m}^{mn}(c) \delta_{n\%2, n'\%2} N_{n'}^{m'} \\
&\quad \times G(m, m', n', n' - m) F(|m'|, n, m') \delta(m, |m'|) \sum_{r=0}^{\infty} \delta(r, n' - m) i^{m-n+r} \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^n d_{n'-m}^{mn}(c) \delta_{n\%2, n'\%2} N_{n'}^{m'} \\
&\quad \times G(m, m', n', n' - m) F(|m'|, n, m') \delta(m, |m'|) H(n' \geq m) i^{n'-n} \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^n d_{n'-|m'|}^{|m'|n}(c) \delta_{n\%2, n'\%2} N_{n'}^{m'} \\
&\quad \times G(|m'|, m', n', n' - |m'|) H(n' \geq |m'|) F(|m'|, n, m') \delta(m, |m'|) i^{n'-n} \\
&= \sum_{n=0}^{\infty} d_{n'-|m'|}^{|m'|n}(c) \delta_{n\%2, n'\%2} N_{n'}^{m'} \\
&\quad \times G(|m'|, m', n', n' - |m'|) F(|m'|, n, m') H(|m'| \leq n') i^{n'-n} \\
&= G(|m'|, m', n', n' - |m'|) N_{n'}^{m'} H(|m'| \leq n') \\
&\quad \times \sum_{n=0}^{\infty} d_{n'-|m'|}^{|m'|n}(c) \delta_{n\%2, n'\%2} F(|m'|, n, m') i^{n'-n} \\
&= G(|m'|, m', n', n' - |m'|) N_{n'}^{m'} \times \sum_{n=0}^{\infty} d_{n'-|m'|}^{|m'|n}(c) \delta_{n\%2, n'\%2} F(|m'|, n, m') i^{n'-n}
\end{aligned} \tag{A.9}$$

In the above, $H(n \geq m)$ is the Heaviside step function:

$$H(n \geq m) = \begin{cases} 1 & \text{for } n \geq m \\ 0 & \text{for } n < m \end{cases} \quad (\text{A.10})$$

From (A.8), it follows:

$$G(|m'|, m', n', n' - |m'|) = \begin{cases} \frac{2(n'+m')!}{(2n'+1)(n'-m')!} & \text{for } m' \geq 0 \\ \frac{2(-1)^{m'}}{(2n'+1)} & \text{for } m' \leq 0 \end{cases} \quad (\text{A.11})$$

Also from (A.7) it follows:

$$\begin{aligned} F(|m'|, n, m') &= \begin{cases} 2\pi A_{|m'|n} & \text{if } m' = 0 \\ \pi (A_{|m'|n} + iB_{|m'|n}) & \text{else if } m' < 0 \\ \pi (A_{|m'|n} - iB_{|m'|n}) & \text{else if } m' > 0 \end{cases} \\ &= \begin{cases} 2\pi A_{0n} & \text{if } m' = 0 \\ \pi (A_{|m'|n} - i\text{sgn}(m')B_{|m'|n}) & \text{else} \end{cases} \quad (\text{A.12}) \\ &= \pi (A_{|m'|n} - i\text{sgn}(m')B_{|m'|n}) \times \begin{cases} 2 & \text{if } m' = 0 \\ 1 & \text{else} \end{cases} \end{aligned}$$

The remaining quantity that is needed to evaluate $\mathcal{A}_{n'}^{m'}(k)$ is the table of $d_{n'-|m'|}^{|m'|n}$ coefficients. This table can be calculated by the method described in [72].

To summarize,

$$\begin{aligned}
\mathcal{A}_{n'}^{m'}(k) &= G(|m'|, m', n', n' - |m'|) N_{n'}^{m'} \\
&\times \sum_{n=0}^{\infty} d_{n'-|m'|}^{|m'|n}(c) \delta_{n\%2, n'\%2} F(|m'|, n, m') i^{n'-n} \\
&= \sqrt{\frac{(n' + |m'|)!}{\pi(2n' + 1)(n' - |m'|)!}} J(m') \\
&\times \sum_{n=0}^{\infty} d_{n'-|m'|}^{|m'|n}(c) \delta_{n\%2, n'\%2} F(|m'|, n, m') i^{n'-n}
\end{aligned} \tag{A.13}$$

where

$$J(m') = \begin{cases} (-1)^{m'} & \text{for } m' \leq 0 \\ 1 & \text{for } m' \geq 0 \end{cases} \tag{A.14}$$

Finally, the following transcoding formula is obtained.

Prolate spheroidal ambisonics to spherical ambisonics transcoding formula

$$\begin{aligned}
\mathcal{A}_{n'}^{m'} &= \alpha(m') \sqrt{\frac{\pi(n' + |m'|)!}{(2n' + 1)(n' - |m'|)!}} \\
&\times \sum_{n=|m'|}^{\infty} \delta_{(n-n')\%2, 0} (-1)^{\frac{n'-n}{2}} d_{n'-|m'|}^{|m'|n}(c) (A_{|m'|n} - i \operatorname{sgn}(m') B_{|m'|n})
\end{aligned} \tag{A.15}$$

where

$$\alpha(m') = \begin{cases} (-1)^{m'} & \text{for } m' < 0 \\ 2 & \text{for } m' = 0 \\ 1 & \text{for } m' > 0 \end{cases} \tag{A.16}$$

Appendix B: Elementary integral computations arising in the RIPE method

B.1 Computation of integrals i_m

We show analytical computation of

$$i_m(x; a) = \int r^m dx, \quad r = \sqrt{x^2 + a^2}, \quad a^2 = y'^2 + z'^2, \quad m = 0, \pm 1, \pm 2, \dots \quad (\text{B.1})$$

For small values of even and odd m these integrals can be computed

$$i_0(x; a) = x, \quad i_{-1}(x; a) = \ln |r + x|, \quad i_{-2}(x; a) = \frac{1}{a} \arctan \frac{x}{a}, \quad i_{-3}(x; a) = \frac{x}{a^2 r}.$$

We have the recurrence:

$$i_{m+2} = \int r^{m+2} dx = x r^{m+2} - (m+2) \int x^2 r^m dx = x r^{m+2} - (m+2) i_{m+2} + (m+2) a^2 i_m.$$
$$i_{m+2} = \frac{x r^{m+2}}{m+3} + \frac{m+2}{m+3} a^2 i_m, \quad m \neq -3. \quad (\text{B.2})$$

We need integral values for $m \geq -3$, and have explicit expressions for all required non-positive m and can recursively find all positive m starting the recurrence from $m = 0$ for even m and from $m = -1$ for odd m . Note that primitives $i_m(x; a)$ may have singularities

if $a = 0$. For these we have

$$i_m(x; 0) = \frac{x|x|^m}{m+1}, \quad m \neq -1, \quad i_{-1}(x; 0) = \text{sgn}(x) \ln|x|. \quad (\text{B.3})$$

B.2 Computation of integrals k_m

In this appendix we show how integrals

$$k_m(x; y', z') = z' \int \frac{r^m}{x^2 + z'^2} dx, \quad r = \sqrt{x^2 + y'^2 + z'^2}, \quad m = 0, \pm 1, \pm 2, \dots \quad (\text{B.4})$$

can be computed analytically. Note, that for $y' = 0$ these reduce to the integrals i_m :

$$k_m(x; 0, z') = z' \int \frac{r^m}{r^2} dx = z' i_{m-2}(x; |z'|). \quad (\text{B.5})$$

So we only need consider the case $y' \neq 0$. We derive recurrences for all needed k_m and find initial values to use them. We have

$$k_{m+2} = z' \int \frac{r^{m+2}}{x^2 + z'^2} dx = z' \int \frac{(x^2 + z'^2) + y'^2}{x^2 + z'^2} \rho^m dx = z' i_m(x; \sqrt{y'^2 + z'^2}) + y'^2 k_m.$$

With known expressions for the initial values (odd and even), all integrals can be computed recursively. For this we have

$$\begin{aligned} k_{-1}(x; y', z') &= \frac{z'}{y'|z'|} \arctan \frac{y'x}{|z'|r}, & k_0(x; y', z') &= \frac{z'}{|z'|} \arctan \frac{x}{|z'|}, \\ k_1(x; y', z') &= \frac{y'z'}{|z'|} \arctan \frac{y'x}{|z'|r} + z' \ln|r+x|. \end{aligned} \quad (\text{B.6})$$

Note that $y'^{k-1}(x; y', z')$ entering the primitive expressions are not singular even when z' and y' approaches zero, while depend on the path (ratio y'/z').

Appendix C: Proof of relation (10.15) used in the manifold element integrals

We use the following definitions of first and second fundamental forms [204]:

$$E \equiv \mathbf{r}_u^2, \quad F \equiv \mathbf{r}_u \cdot \mathbf{r}_v, \quad G \equiv \mathbf{r}_v^2, \quad e \equiv \mathbf{r}_{uu} \cdot \mathbf{n}_q, \quad f \equiv \mathbf{r}_{uv} \cdot \mathbf{n}_q, \quad g \equiv \mathbf{r}_{vv} \cdot \mathbf{n}_q. \quad (\text{C.1})$$

With $\mathbf{C}_u \equiv \partial \mathbf{C} / \partial u$, $\mathbf{C}_v \equiv \partial \mathbf{C} / \partial v$, θ_u the angle of $\hat{\boldsymbol{\rho}}$ from \mathbf{r}_u and θ_v the angle of $\hat{\boldsymbol{\rho}}$ from \mathbf{r}_v measured in the tangent plane of S at \mathbf{r}_q ,

$$\begin{aligned} K &\equiv -Dr_{u\perp n_{v\parallel}} + Cr_{u\parallel n_{v\perp}} + Dr_{v\perp n_{u\parallel}} - Cr_{v\parallel n_{u\perp}} & (\text{C.2}) \\ &= -D(\mathbf{r}_u \cdot \tilde{\boldsymbol{\rho}})(\mathbf{n}_v \cdot \hat{\boldsymbol{\rho}}) + C(\mathbf{r}_u \cdot \hat{\boldsymbol{\rho}})(\mathbf{n}_v \cdot \tilde{\boldsymbol{\rho}}) + D(\mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}})(\mathbf{n}_u \cdot \hat{\boldsymbol{\rho}}) - C(\mathbf{r}_v \cdot \hat{\boldsymbol{\rho}})(\mathbf{n}_u \cdot \tilde{\boldsymbol{\rho}}) \\ &= -D(\mathbf{r}_u \cdot \tilde{\boldsymbol{\rho}})(\mathbf{n}_v \cdot \hat{\boldsymbol{\rho}}) + D(\mathbf{r}_u \cdot \hat{\boldsymbol{\rho}})(\mathbf{n}_v \cdot \tilde{\boldsymbol{\rho}}) + C(\mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}})(\mathbf{n}_u \cdot \hat{\boldsymbol{\rho}}) - C(\mathbf{r}_v \cdot \hat{\boldsymbol{\rho}})(\mathbf{n}_u \cdot \tilde{\boldsymbol{\rho}}) \\ &\quad + (C - D)(\mathbf{r}_u \cdot \hat{\boldsymbol{\rho}})(\mathbf{n}_v \cdot \tilde{\boldsymbol{\rho}}) - (C - D)(\mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}})(\mathbf{n}_u \cdot \hat{\boldsymbol{\rho}}) \\ &= C(\mathbf{r}_v \times \mathbf{n}_u) \cdot (\tilde{\boldsymbol{\rho}} \times \hat{\boldsymbol{\rho}}) - D(\mathbf{r}_u \times \mathbf{n}_v) \cdot (\tilde{\boldsymbol{\rho}} \times \hat{\boldsymbol{\rho}}) + (C - D) ((\mathbf{r}_u \cdot \hat{\boldsymbol{\rho}})(\mathbf{n}_v \cdot \tilde{\boldsymbol{\rho}}) - (\mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}})(\mathbf{n}_u \cdot \hat{\boldsymbol{\rho}})) \\ &= -\mathbf{n}_q \cdot (-D(\mathbf{r}_u \times \mathbf{n}_v) + C(\mathbf{r}_v \times \mathbf{n}_u)) + (C - D) ((\mathbf{r}_u \cdot \hat{\boldsymbol{\rho}})(\mathbf{n}_v \cdot \tilde{\boldsymbol{\rho}}) - (\mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}})(\mathbf{n}_u \cdot \hat{\boldsymbol{\rho}})) \\ &= \frac{1}{J} \mathbf{n}_q \cdot (D(\mathbf{r}_u \times \mathbf{C}_v) - C(\mathbf{r}_v \times \mathbf{C}_u)) + (C - D) ((\mathbf{r}_u \cdot \hat{\boldsymbol{\rho}})(\mathbf{n}_v \cdot \tilde{\boldsymbol{\rho}}) - (\mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}})(\mathbf{n}_u \cdot \hat{\boldsymbol{\rho}})) \\ &= -\frac{1}{J} (C|\mathbf{r}_v|^2(\mathbf{r}_{uu} \cdot \mathbf{n}_q) + D|\mathbf{r}_u|^2(\mathbf{r}_{vv} \cdot \mathbf{n}_q) - (C + D)(\mathbf{r}_u \cdot \mathbf{r}_v)(\mathbf{r}_{uv} \cdot \mathbf{n}_q)) \\ &\quad + (C - D) ((\mathbf{r}_u \cdot \hat{\boldsymbol{\rho}})(\mathbf{n}_v \cdot \tilde{\boldsymbol{\rho}}) - (\mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}})(\mathbf{n}_u \cdot \hat{\boldsymbol{\rho}})) \\ &= -\frac{1}{J} (C|\mathbf{r}_v|^2(\mathbf{r}_{uu} \cdot \mathbf{n}_q) + D|\mathbf{r}_u|^2(\mathbf{r}_{vv} \cdot \mathbf{n}_q) - (C + D)(\mathbf{r}_u \cdot \mathbf{r}_v)(\mathbf{r}_{uv} \cdot \mathbf{n}_q)) \end{aligned}$$

$$\begin{aligned}
& + \frac{C-D}{J} |\mathbf{r}_u| \cos \theta_u (|\mathbf{r}_v| \cos \theta_v \mathbf{r}_{uv} \cdot \mathbf{n}_q - |\mathbf{r}_u| \cos \theta_u \mathbf{r}_{vv} \cdot \mathbf{n}) \\
& - \frac{C-D}{J} |\mathbf{r}_v| \sin \theta_v (-|\mathbf{r}_v| \sin \theta_v \mathbf{r}_{uu} \cdot \mathbf{n}_q + |\mathbf{r}_u| \sin \theta_u \mathbf{r}_{uv} \cdot \mathbf{n}_q).
\end{aligned}$$

Here we used:

$$\mathbf{r}_u \cdot \hat{\boldsymbol{\rho}} = |\mathbf{r}_u| \cos \theta_u, \quad \mathbf{r}_u \cdot \tilde{\boldsymbol{\rho}} = |\mathbf{r}_u| \sin \theta_u, \quad \mathbf{r}_v \cdot \hat{\boldsymbol{\rho}} = |\mathbf{r}_v| \cos \theta_v, \quad \mathbf{r}_v \cdot \tilde{\boldsymbol{\rho}} = |\mathbf{r}_v| \sin \theta_v. \tag{C.3}$$

?? continues as:

$$\begin{aligned}
K &= -\frac{1}{J} (C |\mathbf{r}_v|^2 (\mathbf{r}_{uu} \cdot \mathbf{n}_q) + D |\mathbf{r}_u|^2 (\mathbf{r}_{vv} \cdot \mathbf{n}_q) - (C+D) (\mathbf{r}_u \cdot \mathbf{r}_v) (\mathbf{r}_{uv} \cdot \mathbf{n}_q)) \\
& + \frac{C-D}{J} (\sin^2 \theta_v |\mathbf{r}_v|^2 (\mathbf{r}_{uu} \cdot \mathbf{n}_q) - \cos^2 \theta_u |\mathbf{r}_u|^2 (\mathbf{r}_{vv} \cdot \mathbf{n}_q)) + \frac{C-D}{J} |\mathbf{r}_u| |\mathbf{r}_v| \cos(\theta_u + \theta_v) (\mathbf{r}_{uv} \cdot \mathbf{n}_q) \\
& = \frac{1}{J} (((C-D) \sin^2 \theta_v - C) |\mathbf{r}_v|^2 (\mathbf{r}_{uu} \cdot \mathbf{n}_q)) + \frac{1}{J} ((-D - (C-D) \cos^2 \theta_u) |\mathbf{r}_u|^2 (\mathbf{r}_{vv} \cdot \mathbf{n}_q)) \\
& + 2 \frac{1}{J} (C \cos \theta_u \cos \theta_v - D \sin \theta_u \sin \theta_v) |\mathbf{r}_u| |\mathbf{r}_v| (\mathbf{r}_{uv} \cdot \mathbf{n}_q) \\
& = \frac{-1}{J} ((D \sin^2 \theta_v + C \cos^2 \theta_v) |\mathbf{r}_v|^2 (\mathbf{r}_{uu} \cdot \mathbf{n}_q)) + \frac{-1}{J} ((D \sin^2 \theta_u + C \cos^2 \theta_u) |\mathbf{r}_u|^2 (\mathbf{r}_{vv} \cdot \mathbf{n}_q)) \\
& + 2 \frac{1}{J} (C \cos \theta_u \cos \theta_v - D \sin \theta_u \sin \theta_v) |\mathbf{r}_u| |\mathbf{r}_v| (\mathbf{r}_{uv} \cdot \mathbf{n}_q) \\
& = -\frac{D}{J} (\sin^2 \theta_v \mathbf{r}_v^2 e + \sin^2 \theta_u \mathbf{r}_u^2 g - 2 \sin \theta_u \sin \theta_v |\mathbf{r}_u| |\mathbf{r}_v| f) \\
& - \frac{C}{J} (\cos^2 \theta_v \mathbf{r}_v^2 e + \cos^2 \theta_u \mathbf{r}_u^2 g - 2 \cos \theta_u \cos \theta_v |\mathbf{r}_u| |\mathbf{r}_v| f) \\
& = -JD \frac{(\sin^2 \theta_v) Ge + (\sin^2 \theta_u) Eg - 2(\sin \theta_u \sin \theta_v) |\mathbf{r}_u| |\mathbf{r}_v| f}{EG - F^2} \\
& - JC \frac{(\cos^2 \theta_v) Ge + (\cos^2 \theta_u) Eg - 2(\cos \theta_u \cos \theta_v) |\mathbf{r}_u| |\mathbf{r}_v| f}{EG - F^2} \\
& = -J(C\kappa_N(\tilde{\boldsymbol{\rho}}) + D\kappa_N(\hat{\boldsymbol{\rho}})). \tag{C.4}
\end{aligned}$$

Remark. From the definition of the normal curvature, it follows that the quantities $\kappa_N(\hat{\boldsymbol{\rho}})$

and $\kappa_N(\tilde{\boldsymbol{\rho}})$ given by:

$$\begin{aligned}\kappa_N(\hat{\boldsymbol{\rho}}) &= \frac{Ge \sin^2 \theta_v + Eg \sin^2 \theta_u - 2f|\mathbf{r}_u||\mathbf{r}_v| \sin \theta_u \sin \theta_v}{EG - F^2}, \\ \kappa_N(\tilde{\boldsymbol{\rho}}) &= \frac{Ge \cos^2 \theta_v + Eg \cos^2 \theta_u - 2f|\mathbf{r}_u||\mathbf{r}_v| \cos \theta_u \cos \theta_v}{EG - F^2},\end{aligned}\tag{C.5}$$

are nothing but the normal curvature of the surface at \mathbf{r}_q in direction $\hat{\boldsymbol{\rho}}$ and $\tilde{\boldsymbol{\rho}}$, respectively.

Bibliography

- [1] Hadrien Montanelli, Matthieu Aussal, and Housseem Haddar. Computing weakly singular and near-singular integrals over curved boundary elements. *SIAM Journal on Scientific Computing*, 44(6):A3728–A3753, 2022.
- [2] Philip McCord Morse and K Uno Ingard. *Theoretical acoustics*. Princeton university press, 1986.
- [3] Simon Benedict Shelley, Damian Thomas Murphy, and Andrew James Chadwick. B-format acoustic impulse response measurement and analysis in the forest at koli national park, finland. In *Proceedings of the 16th International Conference on Digital Audio Effects (DAFx13)*, pages 351–355. York, 2013.
- [4] Koli national park – winter – Openair. https://openairlib.net/?page_id=584.
- [5] Koli national park – summer – Openair. https://openairlib.net/?page_id=577.
- [6] Kyle Spratt and Jonathan S Abel. A digital reverberator modeled after the scattering of acoustic waves by trees in a forest. In *Audio Engineering Society Convention 125*. Audio Engineering Society, 2008.
- [7] Waveguide web example audio – Openair. https://openairlib.net/?page_id=782.
- [8] Francis Stevens, Damian T Murphy, Lauri Savioja, and Vesa Välimäki. Modeling sparsely reflecting outdoor acoustic scenes using the waveguide web. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25:1566–1578, 2017.
- [9] Walter A Strauss. *Partial differential equations: An introduction*. John Wiley & Sons, 2007.

- [10] Timo Hiekkänen, Tero Lempinen, Martti Mattila, Ville Pulkki, and Ville Veijanen. Reproduction of virtual reality with multichannel microphone techniques. In *Audio Engineering Society Convention 122*. Audio Engineering Society, 2007.
- [11] Akira Fukada. A challenge in multichannel music recording. In *Audio Engineering Society Conference: 19th International Conference: Surround Sound-Techniques, Technology, and Perception*. Audio Engineering Society, 2001.
- [12] Rafael Kassier, Hyun-Kook Lee, Tim Brookes, and Francis Rumsey. An informal comparison between surround-sound microphone techniques. In *118th AES Convention*, page 17, 2005.
- [13] Toru Kamekawa, Atsushi Marui, and Hideo Irimajiri. Correspondence relationship between physical factors and psychological impressions of microphone arrays for orchestra recording. In *Audio Engineering Society Convention 123*. Audio Engineering Society, 2007.
- [14] Augustinus J Berkhout, Diemer de Vries, and Peter Vogel. Acoustic control by wave field synthesis. *The Journal of the Acoustical Society of America*, 93(5):2764–2778, 1993.
- [15] Michael A Gerzon. Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society*, 21:2–10, 1973.
- [16] Jérôme Daniel, Sebastien Moreau, and Rozenn Nicol. Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging. In *Audio Engineering Society Convention 114*. Audio Engineering Society, 2003.
- [17] Gary Elko, Robert Kubli, and Jens Meyer. Audio system based on at least second-order eigenbeams, August 7 2003. US Patent App. 10/315,502.
- [18] Shoken Kaneko, Tsukasa Suenaga, Hitoshi Akiyama, Yoshiro Miyake, Satoshi Tominaga, Futoshi Shirakihara, and Hiraku Okumura. Development of a 64-channel spherical microphone array and a 122-channel loudspeaker array system for 3D sound field capturing and reproduction technology research. In *Audio Engineering Society Convention 144*. Audio Engineering Society, 2018.
- [19] John M Chowning. The synthesis of complex audio spectra by means of frequency modulation. *Journal of the audio engineering society*, 21(7):526–534, 1973.
- [20] Xavier Serra and Julius Smith. Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition. *Computer Music Journal*, 14(4):12–24, 1990.
- [21] Jont B Allen and David A Berkley. Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65:943–950, 1979.
- [22] Travis Wiens. Forest reverb generator – matlab central file exchange, 2008.

- [23] Dick Botteldooren. Finite-difference time-domain simulation of low-frequency room acoustic problems. *The Journal of the Acoustical Society of America*, 98(6):3302–3308, 1995.
- [24] Lauri Savioja. Real-time 3D finite-difference time-domain simulation of low- and mid-frequency room acoustics. In *13th Int. Conf on Digital Audio Effects*, volume 1, page 75, 2010.
- [25] Andrzej Pietrzyk. Computer modeling of the sound field in small rooms. In *Audio Engineering Society Conference: 15th International Conference: Audio, Acoustics & Small Spaces*. Audio Engineering Society, 1998.
- [26] Tetsuya Sakuma, Shinichi Sakamoto, and Toru Otsuru. *Computational simulation in architectural and environmental acoustics*. Springer, 2014.
- [27] Nikunj Raghuvanshi, Rahul Narain, and Ming C Lin. Efficient and accurate sound propagation using adaptive rectangular decomposition. *IEEE Transactions on Visualization and Computer Graphics*, 15(5):789–801, 2009.
- [28] Christine Evers, Heinrich W Löllmann, Heinrich Mellmann, Alexander Schmidt, Hendrik Barfuss, Patrick A Naylor, and Walter Kellermann. The LOCATA challenge: Acoustic source localization and tracking. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28:1620–1643, 2020.
- [29] Jens Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.
- [30] Shoken Kaneko, Tsukasa Suenaga, and Satoshi Sekine. Deepearnet: individualizing spatial audio with photography, ear shape modeling, and neural networks. In *Audio Engineering Society Conference: 2016 AES International Conference on Audio for Virtual and Augmented Reality*. Audio Engineering Society, 2016.
- [31] Kazuhiko Yamamoto and Takeo Igarashi. Fully perceptual-based 3d spatial sound individualization with an adaptive variational autoencoder. *ACM Transactions on Graphics (TOG)*, 36(6):1–13, 2017.
- [32] Faiyadh Shahid, Nikhil Javeri, Kapil Jain, and Shruti Badhwar. AI DevOps for Large-Scale HRTF Prediction and Evaluation: An End to End Pipeline. In *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*. Audio Engineering Society, 2018.
- [33] Corentin Guezenoc and Renaud Segulier. HRTF individualization: A survey. *arXiv preprint arXiv:2003.06183*, 2020.
- [34] Robert Pelzer, Manoj Dinakaran, Fabian Brinkmann, Steffen Lepa, Peter Grosche, and Stefan Weinzierl. Head-related transfer function recommendation based on perceptual similarities and anthropometric features. *The Journal of the Acoustical Society of America*, 148(6):3809–3817, 2020.

- [35] Bowen Zhi, Dmitry N Zotkin, and Ramani Duraiswami. Towards fast and convenient end-to-end HRTF personalization. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 441–445. IEEE, 2022.
- [36] Navid H Zandi, Awny M El-Mohandes, and Rong Zheng. Individualizing Head-Related Transfer Functions for Binaural Acoustic Applications. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 105–117. IEEE, 2022.
- [37] Simone Angelucci, Claudia Rinaldi, Fabio Franchi, and Fabio Graziosi. Comparison of ML Solutions for HRIR Individualization Design in Binaural Audio. In *International Conference on Advanced Information Networking and Applications*, pages 271–278. Springer, 2023.
- [38] Ziran Jiang, Jinqiu Sang, Chengshi Zheng, Andong Li, and Xiaodong Li. Modeling individual head-related transfer functions from sparse measurements using a convolutional neural network. *The Journal of the Acoustical Society of America*, 153(1):248–259, 2023.
- [39] Stephen J Elliott and Philip A Nelson. Active noise control. *IEEE signal processing magazine*, 10:12–35, 1993.
- [40] Sen M Kuo and Dennis R Morgan. Active noise control: a tutorial review. *Proceedings of the IEEE*, 87:943–973, 1999.
- [41] Prasanga N Samarasinghe, Wen Zhang, and Thushara D Abhayapala. Recent advances in active noise control inside automobile cabins: Toward quieter cars. *IEEE Signal Processing Magazine*, 33(6):61–73, 2016.
- [42] Jihui Zhang, Thushara D Abhayapala, Wen Zhang, Prasanga N Samarasinghe, and Shouda Jiang. Active noise control over space: A wave domain approach. *IEEE/ACM Transactions on audio, speech, and language processing*, 26(4):774–786, 2018.
- [43] Yuchen Dong, Jie Chen, and Wen Zhang. Distributed wave-domain active noise control based on the diffusion strategy. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8429–8433. IEEE, 2020.
- [44] Yousef Saad. *Iterative methods for sparse linear systems*. SIAM, 2003.
- [45] Leslie Greengard and Vladimir Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys.*, 73(2):325–348, 1987.
- [46] Arnold Sommerfeld. Die Greensche Funktion der Schwingungsgleichung. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 21:309–352, 1912.

- [47] Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 114(2):185–200, 1994.
- [48] Brian FG Katz. Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation. *The Journal of the Acoustical Society of America*, 110(5):2440–2448, 2001.
- [49] Nail A Gumerov, Ramani Duraiswami, and Dmitry N Zotkin. Fast multipole accelerated boundary elements for numerical computation of the head related transfer function. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, volume 1, pages I–165. IEEE, 2007.
- [50] Shoken Kaneko, Tsukasa Suenaga, Mai Fujiwara, Kazuya Kumehara, Futoshi Shirakihara, and Satoshi Sekine. Ear Shape Modeling for 3D Audio and Acoustic Virtual Reality: The Shape-Based Average HRTF. In *Audio Engineering Society Conference: 61st International Conference: Audio for Games*. Audio Engineering Society, 2016.
- [51] Shoken Kaneko, Tsukasa Suenaga, and Satoshi Sekine. DeepEarNet: individualizing spatial audio with photography, ear shape modeling, and neural networks. In *Audio Engineering Society Conference: 2016 AES International Conference on Audio for Virtual and Augmented Reality*. Audio Engineering Society, 2016.
- [52] ISO/TC 12 Committee et al. Iso 80000-2: 2019 Quantities and units—part 2: Mathematics, Aug. 2019.
- [53] Shoken Kaneko and Ramani Duraiswami. Multiple scattering ambisonics: Three-dimensional sound field estimation using interacting spheres. *JASA Express Letters*, 1(8), 2021.
- [54] YouTube. Use spatial audio in 360-degree and VR videos. <https://support.google.com/youtube/answer/6395969>.
- [55] Facebook. Creating Videos with Spatial Audio for Facebook 360. <https://facebookincubator.github.io/facebook-360-spatial-workstation/KB/CreatingVideosSpatialAudioFacebook360.html>.
- [56] Markus Noisternig, Thomas Musil, Alois Sontacchi, and Robert Holdrich. 3D binaural sound reproduction using a virtual ambisonic approach. In *IEEE International Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2003. VECIMS'03. 2003*, pages 174–178. IEEE, 2003.
- [57] Dmitry N Zotkin, Ramani Duraiswami, and Larry S Davis. Rendering localized spatial audio in a virtual auditory space. *IEEE Transactions on multimedia*, 6:553–564, 2004.

- [58] Shoken Kaneko, Tsukasa Suenaga, Mai Fujiwara, Kazuya Kumehara, Futoshi Shirakihara, and Satoshi Sekine. Ear shape modeling for 3D audio and acoustic virtual reality: The shape-based average HRTF. In *Audio Engineering Society Conference: 61st International Conference: Audio for Games*. Audio Engineering Society, 2016.
- [59] Jens Meyer and Gary Elko. A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages II–1781. IEEE, 2002.
- [60] Shuichi Sakamoto, Satoshi Hongo, Takuma Okamoto, Yukio Iwaya, and Yōiti Suzuki. Sound-space recording and binaural presentation system based on a 252-channel microphone array. *Acoustical Science and Technology*, 36:516–526, 2015.
- [61] Paul A Martin. *Multiple scattering: interaction of time-harmonic waves with N obstacles*. Cambridge University Press, 2006.
- [62] Mark A Poletti. Three-dimensional surround sound systems based on spherical harmonics. *Journal of the Audio Engineering Society*, 53:1004–1025, 2005.
- [63] Nail A Gumerov and Ramani Duraiswami. Computation of scattering from n spheres using multipole reexpansion. *Journal of the Acoustical Society of America*, 112:2688–2701, 2002.
- [64] Michael A Epton and Benjamin Dembart. Multipole translation theory for the three-dimensional Laplace and Helmholtz equations. *SIAM Journal on Scientific Computing*, 16:865–897, 1995.
- [65] WC Chew. Recurrence relations for three-dimensional scalar addition theorem. *Journal of Electromagnetic Waves and Applications*, 6:133–142, 1992.
- [66] Richard Swinbank and R James Purser. Fibonacci grids: A novel approach to global modelling. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 132:1769–1793, 2006.
- [67] Prasanga Samarasinghe, Thushara Abhayapala, and Mark Poletti. Wavefield analysis over large areas using distributed higher order microphones. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22:647–658, 2014.
- [68] Masahiro Nakanishi, Natsuki Ueno, Shoichi Koyama, and Hiroshi Saruwatari. Two-dimensional sound field recording with multiple circular microphone arrays considering multiple scattering. In *2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 368–372. IEEE, 2019.
- [69] Shoken Kaneko and Ramani Duraiswami. Spheroidal ambisonics: Spatial audio in spheroidal bases. *JASA Express Letters*, 1(8), 2021.

- [70] Dmitry N Zotkin, Nail A Gumerov, and Ramani Duraiswami. Incident field recovery for an arbitrary-shaped scatterer. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 451–455. IEEE, 2017.
- [71] Carson Flammer. *Spheroidal wave functions*. Courier Corporation, 2014.
- [72] Ross Adelman, Nail A Gumerov, and Ramani Duraiswami. Software for computing the spheroidal wave functions using arbitrary precision arithmetic. *arXiv preprint arXiv:1408.0074*, 2014.
- [73] Chris Travis. A new mixed-order scheme for ambisonic signals. In *Proc. of the Ambisonics Symposium, Graz, Austria, 2009*.
- [74] Virginia Ann Best. *Spatial hearing with simultaneous sound sources: A psychophysical investigation*. PhD thesis, University of Sydney. Medicine, 2004.
- [75] Laurent Fousse, Guillaume Hanrot, Vincent Lefèvre, Patrick Pélissier, and Paul Zimmermann. MPFR: A multiple-precision binary floating-point library with correct rounding. *ACM Transactions on Mathematical Software (TOMS)*, 33(2):13:1–13:15, 2007.
- [76] Svein Berge. Acoustically hard 2D arrays for 3D HOA. In *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*. Audio Engineering Society, 2019.
- [77] Boaz Rafaely, Barak Weiss, and Eitan Bachmat. Spatial aliasing in spherical microphone arrays. *IEEE Transactions on Signal Processing*, 55:1003–1010, 2007.
- [78] Zhiyun Li, Ramani Duraiswami, Elena Grassi, and Larry S Davis. Flexible layout and optimal cancellation of the orthonormality error for spherical microphone arrays. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 4, pages iv–iv. IEEE, 2004.
- [79] Zhiyun Li and Ramani Duraiswami. Flexible and optimal design of spherical microphone arrays for beamforming. *IEEE Transactions on Audio, Speech, and Language Processing*, 15:702–714, 2007.
- [80] Shoken Kaneko and Hannes Gamper. A fast forest reverberator using single scattering cylinders. In *2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–5. IEEE, 2021.
- [81] Carl F Eyring. Jungle acoustics. *Journal of the Acoustical Society of America*, 18:257–270, 1946.
- [82] Leslie L Foldy. The multiple scattering of waves. i. general theory of isotropic scattering by randomly distributed scatterers. *Physical review*, 67:107, 1945.

- [83] Victor Twersky. On scattering of waves by random distributions. i. free-space scatterer formalism. *Journal of Mathematical Physics*, 3:700–715, 1962.
- [84] SK Bose and AK Mal. Longitudinal shear waves in a fiber-reinforced composite. *International Journal of Solids and Structures*, 9:1075–1085, 1973.
- [85] CM Linton and PA Martin. Multiple scattering by random configurations of circular cylinders: Second-order corrections for the effective wavenumber. *Journal of the Acoustical Society of America*, 117:3413–3423, 2005.
- [86] Gilbert A Soulodre and John S Bradley. Subjective evaluation of new room acoustic measures. *The Journal of the Acoustical Society of America*, 98(1):294–301, 1995.
- [87] Helena Peić Tukuljac, Ville Pulkki, Hannes Gamper, Keith Godin, Ivan J Tashev, and Nikunj Raghuvanshi. A sparsity measure for echo density growth in general environments. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2019.
- [88] ISO Central Secretary. Acoustics — Attenuation of sound during propagation outdoors — Part 1: Calculation of the absorption of sound by the atmosphere. Standard ISO 9613-1:1993, International Organization for Standardization, June 1993.
- [89] Damian T Murphy and Simon Shelley. Openair: An interactive auralization web resource and database. In *Audio Engineering Society Convention 129*. Audio Engineering Society, 2010.
- [90] T. W. Crowther, H. B. Glick, K. R. Covey, et al. Mapping tree density at a global scale. *Nature*, 525:201—205, 2015.
- [91] Shoken Kaneko and Hannes Gamper. Large-scale simulation of bird localization systems in forests with distributed microphone arrays. *JASA Express Letters*, 2(10), 2022.
- [92] Kevin Darras, Péter Batáry, Brett J Furnas, Ingo Grass, Yeni A Mulyani, and Teja Tschardtke. Autonomous sound recording outperforms human observation for sampling birds: a systematic map and user guide. *Ecological Applications*, 29:e01954, 2019.
- [93] Tessa A Rhinehart, Lauren M Chronister, Trieste Devlin, and Justin Kitzes. Acoustic localization of terrestrial wildlife: Current practices and future opportunities. *Ecology and Evolution*.
- [94] Luis F Gonzalez, Glen A Montes, Eduard Puig, Sandra Johnson, Kerrie Mengersen, and Kevin J Gaston. Unmanned aerial vehicles (UAVs) and artificial intelligence revolutionizing wildlife monitoring and conservation. *Sensors*, 16(1):97, 2016.

- [95] Vladimir Dyo, Stephen A Ellwood, David W Macdonald, Andrew Markham, Cecilia Mascolo, Bence Pásztor, Salvatore Scellato, Niki Trigoni, Ricklef Wohlers, and Kharsim Yousef. Evolution and sustainability of a wildlife monitoring sensor network. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, pages 127–140, 2010.
- [96] Jarrod C Hodgson, Rowan Mott, Shane M Baylis, Trung T Pham, Simon Wotherspoon, Adam D Kilpatrick, Ramesh Raja Segaran, Ian Reid, Aleks Terauds, and Lian Pin Koh. Drones count wildlife more accurately and precisely than humans. *Methods in Ecology and Evolution*, 9(5):1160–1167, 2018.
- [97] Charles Knapp and Glifford Carter. The generalized correlation method for estimation of time delay. *IEEE transactions on acoustics, speech, and signal processing*, 24:320–327, 1976.
- [98] Deanna K Dawson and Murray G Efford. Bird population density estimated from acoustic signals. *Journal of Applied Ecology*, 46:1201–1209, 2009.
- [99] Daniel J Mennill, Matthew Battiston, David R Wilson, Jennifer R Foote, and Stephanie M Doucet. Field test of an affordable, portable, wireless microphone array for spatial monitoring of animal ecology and behaviour. *Methods in Ecology and Evolution*, 3:704–712, 2012.
- [100] Phillip M Stepanian, Kyle G Horton, David C Hille, Charlotte E Wainwright, Phillip B Chilson, and Jeffrey F Kelly. Extending bioacoustic monitoring of birds aloft through flight call localization with a three-dimensional microphone array. *Ecology and evolution*, 6:7039–7046, 2016.
- [101] The Cornell Lab. The Cornell Guide to Bird Sounds: Master Set for North America. <https://www.macaulaylibrary.org>.
- [102] xeno canto. xeno-canto. <https://www.xeno-canto.org>.
- [103] Jongseo Sohn, Nam Soo Kim, and Wonyong Sung. A statistical model-based voice activity detection. *IEEE signal processing letters*, 6(1):1–3, 1999.
- [104] Daniele Mirabilli and Emanuël AP Habets. Simulating multi-channel wind noise based on the corcos model. In *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*, pages 560–564. IEEE, 2018.
- [105] Earl G Williams. *Fourier acoustics: sound radiation and nearfield acoustical holography*. Elsevier, 1999.
- [106] International Electrotechnical Commission. Electroacoustics - Sound level meters - Part 1: Specifications. International Standard IEC 61672-1:2013, IEC, 2013.
- [107] Peter R Roth. Effective measurements using digital signal analysis. *IEEE spectrum*, 8:62–70, 1971.

- [108] Parham Aarabi. The fusion of distributed microphone arrays for sound localization. *EURASIP Journal on Advances in Signal Processing*, 2003:860465, 2003.
- [109] Mario Lasseck. Acoustic bird detection with deep convolutional neural networks. Technical report, DCASE2018 Challenge, September 2018.
- [110] Yariv Ephraim and David Malah. Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. *IEEE Transactions on acoustics, speech, and signal processing*, 32:1109–1121, 1984.
- [111] Norbert Wiener. Extrapolation, interpolation and smoothing of stationary. *Time Series, with Engineering Applications*, 1949.
- [112] Robert McAulay and Marilyn Malpass. Speech enhancement using a soft-decision noise suppression filter. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28:137–145, 1980.
- [113] Steven Boll. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on acoustics, speech, and signal processing*, 27:113–120, 1979.
- [114] Patrick J Wolfe and Simon J Godsill. Simple alternatives to the ephraim and malah suppression rule for speech enhancement. In *Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing (Cat. No. 01TH8563)*, pages 496–499. IEEE, 2001.
- [115] Rainer Martin. Speech enhancement using mmse short time spectral estimation with gamma distributed speech priors. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages I–253. IEEE, 2002.
- [116] Ivan Jelev Tashev. *Sound capture and processing: practical approaches*. John Wiley & Sons, 2009.
- [117] C Eckhart. Optimal rectifier systems for detection of steady signals, scripps. *Inst Oceanography, Marine Physical Lab. Univ. California, Rep. SIO*, 12692:52–11, 1952.
- [118] Marek B Trawicki and Michael T Johnson. Distributed multichannel speech enhancement with minimum mean-square error short-time spectral amplitude, log-spectral amplitude, and spectral phase estimation. *Signal Processing*, 92:345–356, 2012.
- [119] Thomas Lotter, Christian Benien, and Peter Vary. Multichannel direction-independent speech enhancement using spectral amplitude estimation. *EURASIP Journal on Advances in Signal Processing*, 2003:705085, 2003.
- [120] Amin Hassani, Alexander Bertrand, and Marc Moonen. Cooperative integrated noise reduction and node-specific direction-of-arrival estimation in a fully connected wireless acoustic sensor network. *Signal Processing*, 107:68–81, 2015.

- [121] Alexander Bertrand and Marc Moonen. Distributed adaptive node-specific signal estimation in fully connected sensor networks—part i: Sequential node updating. *IEEE Transactions on Signal Processing*, 58:5277–5291, 2010.
- [122] Alexander Bertrand and Marc Moonen. Distributed adaptive node-specific signal estimation in fully connected sensor networks—part ii: Simultaneous and asynchronous node updating. *IEEE Transactions on Signal Processing*, 58:5292–5306, 2010.
- [123] Daniel Gabriel, Ryosuke Kojima, Kotaro Hoshiba, Katsutoshi Itoyama, Kenji Nishida, and Kazuhiro Nakadai. 2D sound source position estimation using microphone arrays and its application to a vr-based bird song analysis system. *Advanced Robotics*, 33:403–414, 2019.
- [124] Shoken Kaneko and Hannes Gamper. Towards all-purpose full-sphere binaural localization. In *2022 30th European Signal Processing Conference (EUSIPCO)*, pages 334–338. IEEE, 2022.
- [125] F Palmieri, M Datum, A Shah, and A Moiseff. Learning binaural sound localization through a neural network. In *Proceedings of the 1991 IEEE Seventeenth Annual Northeast Bioengineering Conference*, pages 13–14. IEEE, 1991.
- [126] A Moiseff, F Palmieri, M Datum, and A Shah. An artificial neural network for studying binaural sound localization. In *Proceedings of the 1991 IEEE Seventeenth Annual Northeast Bioengineering Conference*, pages 1–2. IEEE, 1991.
- [127] Chalapathy Neti, Eric D Young, and Michael H Schneider. Neural network models of sound localization based on directional filtering by the pinna. *The Journal of the Acoustical Society of America*, 92(6):3140–3156, 1992.
- [128] Craig Jin, Markus Schenkel, and Simon Carlile. Neural system identification model of human sound localization. *The Journal of the Acoustical Society of America*, 108(3):1215–1235, 2000.
- [129] Yi Jiang, DeLiang Wang, RunSheng Liu, and ZhenMing Feng. Binaural classification for reverberant speech segregation using deep neural networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(12):2112–2121, 2014.
- [130] Ning Ma, Tobias May, and Guy J Brown. Exploiting deep neural networks and head movements for robust binaural localization of multiple sources in reverberant environments. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(12):2444–2453, 2017.
- [131] Etienne Thuillier, Hannes Gamper, and Ivan J Tashev. Spatial audio feature discovery with convolutional neural networks. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6797–6801. IEEE, 2018.

- [132] Jing Wang, Jin Wang, Kai Qian, Xiang Xie, and Jingming Kuang. Binaural sound localization based on deep neural network and affinity propagation clustering in mismatched hrtf condition. *EURASIP Journal on Audio, Speech, and Music Processing*, 2020(1):1–16, 2020.
- [133] Yichen Yang, Jingwei Xi, Wen Zhang, and Lijun Zhang. Full-Sphere Binaural Sound Source Localization Using Multi-task Neural Network. In *2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 432–436. IEEE, 2020.
- [134] Andrew Francis and Josh H McDermott. Deep neural network models of sound localization reveal how perception is adapted to real-world environments. *Nature Human Behaviour*, 6(1):111–133, 2022.
- [135] Álvaro González. Measurement of areas on a sphere using Fibonacci and latitude–longitude lattices. *Mathematical Geosciences*, 42(1):49–64, 2010.
- [136] Jianbang Ding, Xuancheng Ren, Ruixuan Luo, and Xu Sun. An adaptive and momental bound method for stochastic learning. *arXiv preprint arXiv:1910.12249*, 2019.
- [137] Jan Schlüter and Sebastian Böck. Improved musical onset detection with convolutional neural networks. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6979–6983. IEEE, 2014.
- [138] C. Hummersone. Binaural Room Impulse Response Measurements. <https://github.com/IoSR-Surrey/RealRoomBRIRs>, 2011.
- [139] Jon Francombe. IoSR Listening Room Multichannel BRIR dataset’. https://github.com/IoSR-Surrey/IoSR_ListeningRoom_BRIRs, 2017.
- [140] Heinrich W Löllmann, Christine Evers, Alexander Schmidt, Heinrich Mellmann, Hendrik Barfuss, Patrick A Naylor, and Walter Kellermann. The LOCATA challenge data corpus for acoustic source localization and tracking. In *2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 410–414. IEEE, 2018.
- [141] V Ralph Algazi, Richard O Duda, Dennis M Thompson, and Carlos Avendano. The CIPIC HRTF database. In *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575)*, pages 99–102. IEEE, 2001.
- [142] P. Majdak. ARI HRTF database. <https://www.oeaw.ac.at/isf/das-institut/software/hrtf-database>, 2017.
- [143] Kanji Watanabe, Yukio Iwaya, Yôiti Suzuki, Shouichi Takane, and Sojun Sato. Dataset of head-related transfer functions measured with a circular loudspeaker array. *Acoustical science and technology*, 35(3):159–165, 2014.

- [144] Ramona Bomhardt, Matias de la Fuente Klein, and Janina Fels. A high-resolution head-related transfer function and three-dimensional ear model database. In *Proceedings of Meetings on Acoustics 172ASA*, volume 29, page 050002. Acoustical Society of America, 2016.
- [145] Simone Spagnol, Kristján Bjarki Purkhús, Rúnar Unnthórsson, and Sverrir Karl Björnsson. The Viking HRTF dataset. In *16th Sound and music computing conference*, pages 55–60. Sound and Music Computing Network, 2019.
- [146] Slim Ghorbal, Xavier Bonjour, and Renaud Séguier. Computed hrirs and ears database for acoustic research. In *Audio Engineering Society Convention 148*. Audio Engineering Society, 2020.
- [147] Brinkmann Fabian, Dinakaran Manoj, Pelzer Robert, Wohlgemuth Jan Joschka, Seipel Fabian, Voss Daniel, Grosche Peter, and Weinzierl Stefan. The HUTUBS head-related transfer function (HRTF) database. 2019.
- [148] Martin Cooke, Jon Barker, Stuart Cunningham, and Xu Shao. An audio-visual corpus for speech perception and automatic speech recognition. *The Journal of the Acoustical Society of America*, 120(5):2421–2424, 2006.
- [149] Xiang Wu, Dumidu S Talagala, Wen Zhang, and Thushara D Abhayapala. Individualized interaural feature learning and personalized binaural localization model. *Applied Sciences*, 9(13):2682, 2019.
- [150] Elizabeth M Wenzel, Marianne Arruda, Doris J Kistler, and Frederic L Wightman. Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America*, 94(1):111–123, 1993.
- [151] Shoken Kaneko, Nirupam Roy, Nail Gumerov, and Ramani Duraiswami. Regularized spherical harmonics-domain spatial active noise cancellation in a reverberant room. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 263, pages 4733–4742. Institute of Noise Control Engineering, 2021.
- [152] Hanchi Chen, Prasanga Samarasinghe, Thushara D Abhayapala, and Wen Zhang. Spatial noise cancellation inside cars: Performance analysis and experimental results. In *2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 1–5. IEEE, 2015.
- [153] Hanchi Chen, Jihui Zhang, Prasanga N Samarasinghe, and Thushara D Abhayapala. Evaluation of spatial active noise cancellation performance using spherical harmonic analysis. In *2016 IEEE International Workshop on Acoustic Signal Enhancement (IWAENC)*, pages 1–5. IEEE, 2016.
- [154] Huiyuan Sun, Thushara D Abhayapala, and Prasanga N Samarasinghe. Time domain spherical harmonic analysis for adaptive noise cancellation over a spatial region. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 516–520. IEEE, 2019.

- [155] Prasanta Kumar Banerjee and Roy Butterfield. *Boundary element methods in engineering science*, volume 17. McGraw-Hill London, 1981.
- [156] Nail A Gumerov and Ramani Duraiswami. *Fast multipole methods for the Helmholtz equation in three dimensions*. Elsevier, 2005.
- [157] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33:1, 2010.
- [158] Bradley Efron, Trevor Hastie, Iain Johnstone, Robert Tibshirani, et al. Least angle regression. *The Annals of statistics*, 32:407–499, 2004.
- [159] Ricardo Marques, Christian Bouville, Mickaël Ribardière, Luis Paulo Santos, and Kadi Bouatouch. Spherical fibonacci point sets for illumination integrals. In *Computer Graphics Forum*, volume 32, pages 134–143. Wiley Online Library, 2013.
- [160] Shoken Kaneko, Irtaza Shahid, Nirupam Roy, Nail A Gumerov, and Ramani Duraiswami. Personal active soundfield control (PASCAL) framework. In *Audio Engineering Society Conference: AES 2022 International Audio for Virtual and Augmented Reality Conference*. Audio Engineering Society, 2022.
- [161] Nail A Gumerov and Ramani Duraiswami. A broadband fast multipole accelerated boundary element method for the three dimensional Helmholtz equation. *Journal of the Acoustical Society of America*, 125:191–205, 2009.
- [162] Nail A Gumerov and Ramani Duraiswami. Fast multipole accelerated boundary element methods for room acoustics. *arXiv preprint arXiv:2103.16073*, 2021.
- [163] Sheng Shen, Nirupam Roy, Junfeng Guan, Haitham Hassanieh, and Romit Roy Choudhury. MUTE: Bringing IoT to noise cancellation. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, pages 282–296, 2018.
- [164] Sheng Shen, Nirupam Roy, Junfeng Guan, Haitham Hassanieh, and Romit Roy Choudhury. Networked acoustics around human ears. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, pages 747–749, New York, NY, USA, October 2018. Association for Computing Machinery.
- [165] Peter Damaske. Head-related two-channel stereophony with loudspeaker reproduction. *Journal of the Acoustical Society of America*, 50:1109–1115, 1971.
- [166] Craig T Jin, Nicolas Epain, Tahereh Noohi, V Pulkki, S Delikaris-Manias, and A Politis. Sound field analysis using sparse recovery. In *Parametric Time-Frequency-Domain Spatial Audio*. John Wiley & Sons, 2017.

- [167] DYN Zotkin, Jane Hwang, R Duraiswaini, and Larry S Davis. HRTF personalization using anthropometric measurements. In *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 157–160. IEEE, 2003.
- [168] Song Xu, Zhizhong Li, and Gavriel Salvendy. Individualization of head-related transfer function for three-dimensional virtual auditory display: a review. In *International Conference on Virtual Reality*, pages 397–407. Springer, 2007.
- [169] Hisashi Sano, Toshio Inoue, Akira Takahashi, Kenichi Terai, and Yoshio Nakamura. Active control system for low-frequency road noise combined with an audio system. *IEEE Transactions on speech and audio processing*, 9:755–763, 2001.
- [170] Stephen J Elliott and Matthew Jones. An active headrest for personal audio. *Journal of the Acoustical Society of America*, 119:2702–2709, 2006.
- [171] Alessandro Opinto, Marco Martalò, Carlo Tripodi, Alessandro Costalunga, Luca Cattani, and Riccardo Raheli. Performance Analysis of Feedback MIMO ANC in Experimental Automotive Environment. In *2020 14th International Conference on Signal Processing and Communication Systems (ICSPCS)*, pages 1–6. IEEE, 2020.
- [172] Elias Bjarnason. Analysis of the filtered-X LMS algorithm. *IEEE Transactions on Speech and Audio Processing*, 3:504–514, 1995.
- [173] Shawn Hershey, Sourish Chaudhuri, Daniel PW Ellis, Jort F Gemmeke, Aren Jansen, R Channing Moore, Manoj Plakal, Devin Platt, Rif A Saurous, Bryan Seybold, et al. CNN architectures for large-scale audio classification. In *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 131–135. IEEE, 2017.
- [174] DeLiang Wang and Jitong Chen. Supervised speech separation based on deep learning: An overview. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26:1702–1726, 2018.
- [175] Yong Xu, Jun Du, Li-Rong Dai, and Chin-Hui Lee. A regression approach to speech enhancement based on deep neural networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23:7–19, 2014.
- [176] Shoken Kaneko, Nail A Gumerov, and Ramani Duraiswami. Recursive Analytical Quadrature of Laplace and Helmholtz Layer Potentials in \mathbb{R}^3 . *arXiv preprint arXiv:2302.02196*, 2023.
- [177] K Hayami and CA Brebbia. Quadrature methods for singular and nearly singular integrals in 3-D boundary element method. *Boundary elements X*, 1:237–264, 1988.

- [178] M. Guiggiani, G. Krishnasamy, T. J. Rudolphi, and F. J. Rizzo. A General Algorithm for the Numerical Solution of Hypersingular Boundary Integral Equations. *J. Appl. Mech.*, 59(3):604–614, 09 1992.
- [179] Wolfgang Hackbusch and Stefan A Sauter. On numerical cubatures of nearly singular surface integrals arising in BEM collocation. *Computing*, 52(2):139–159, 1994.
- [180] Barbara M Johnston, Peter R Johnston, and David Elliott. A sinh transformation for evaluating two-dimensional nearly singular boundary element integrals. *Int. J. Numer. Methods Eng.*, 69(7):1460–1479, 2007.
- [181] Marc Lenoir and Nicolas Salles. Evaluation of 3-D singular and nearly singular integrals in Galerkin BEM for thin layers. *SIAM J. Sci. Comput.*, 34(6):A3057–A3078, 2012.
- [182] Nail A Gumerov, Shoken Kaneko, and Ramani Duraiswami. Analytical Galerkin boundary integrals of Laplace kernel layer potentials in \mathbb{R}^3 . *arXiv preprint arXiv:2302.03247*, 2023.
- [183] Hai Zhu and Shravan Veerapaneni. High-order close evaluation of Laplace layer potentials: A differential geometric approach. *SIAM J. Sci. Comput.*, 44(3):A1381–A1404, 2022.
- [184] Ross Adelman, Nail A Gumerov, and Ramani Duraiswami. Computation of Galerkin double surface integrals in the 3-D boundary element method. *IEEE Trans. Antennas Propag.*, 64(6):2389–2400, 2016.
- [185] Andreas Klöckner, Alexander Barnett, Leslie Greengard, and Michael O’Neil. Quadrature by expansion: A new method for the evaluation of layer potentials. *J. Comput. Phys.*, 252:332–349, 2013.
- [186] Matt Wala and Andreas Klöckner. A fast algorithm for quadrature by expansion in three dimensions. *J. Comput. Phys.*, 388:655–689, 2019.
- [187] Matt Wala and Andreas Klöckner. Optimization of fast algorithms for global Quadrature by Expansion using target-specific expansions. *J. Comput. Phys.*, 403:108976, 2020.
- [188] Dan Rosen and Donald E Cormack. The continuation approach: A general framework for the analysis and evaluation of singular and near-singular integrals. *SIAM Journal on Applied Mathematics*, 55(3):723–762, 1995.
- [189] John Nicholas Newman. Distributions of sources and normal dipoles over a quadrilateral panel. *Journal of Engineering Mathematics*, 20(2):113–126, 1986.
- [190] Marc Lenoir and Nicolas Salles. Exact evaluation of singular and near-singular integrals in Galerkin BEM. *Proceedings of ECCOMAS2012*, pages 1–20, 2012.

- [191] Nail A Gumerov, Shoken Kaneko, and Ramani Duraiswami. Recursive computation of the multipole expansions of layer potential integrals over simplices for efficient fast multipole accelerated boundary elements. *Journal of Computational Physics*, 486:112118, 2023.
- [192] Shoken Kaneko and Ramani Duraiswami. Efficient exact quadrature of regular solid harmonics times polynomials over simplices in \mathbb{R}^3 . *arXiv preprint arXiv:2307.12202*, 2023.
- [193] Nail A Gumerov and Ramani Duraiswami. Analytical computation of boundary integrals for the Helmholtz equation in three dimensions. *arXiv:2103.17196*, 2021.
- [194] Stefan A Sauter and Christoph Schwab. Boundary element methods. In *Boundary Element Methods*, pages 183–287. Springer, 2010.
- [195] AJ Burton and GF Miller. The application of integral equation methods to the numerical solution of some exterior boundary-value problems. *Proc. R. Soc. A*, 323(1553):201–210, 1971.
- [196] Nail Gumerov, Ross Adelman, and Ramani Duraiswami. Fast multipole accelerated indirect boundary elements for the Helmholtz equation. In *Proc. Meet. Acoust. ICA2013*, volume 19, page 015097. Acoustical Society of America, 2013.
- [197] Ake Björck and Victor Pereyra. Solution of Vandermonde systems of equations. *Mathematics of computation*, 24(112):893–903, 1970.
- [198] Robert Piessens, Elise de Doncker-Kapenga, Christoph W Überhuber, and David K Kahaner. *Quadpack: a subroutine package for automatic integration*, volume 1. Springer, 2012.
- [199] Leslie Greengard, Michael O’Neil, Manas Rachh, and Felipe Vico. Fast multipole methods for the evaluation of layer potentials with locally-corrected quadratures. *J. Comput. Phys.: X*, 10:100092, 2021.
- [200] Shoken Kaneko and Ramani Duraiswami. Layer potential quadrature on manifold boundary elements with constant densities for Laplace and Helmholtz kernels in \mathbb{R}^3 . *arXiv preprint arXiv:2309.07802*, 2023.
- [201] Shoken Kaneko and Ramani Duraiswami. Layer potential quadrature on manifold boundary elements with constant densities for Laplace and Helmholtz kernels in \mathbb{R}^3 . *arXiv preprint arXiv:2309.07802*, 2023.
- [202] Gernot Beer, Benjamin Marussig, and Christian Duenser. *The isogeometric boundary element method*. Springer, 2020.
- [203] John M Lee. *Introduction to Riemannian manifolds*, volume 2. Springer, 2018.
- [204] Manfredo P Do Carmo. *Differential geometry of curves and surfaces: revised and updated second edition*. Courier Dover Publications, 2016.

- [205] Michael Spivak. *Calculus on manifolds: a modern approach to classical theorems of advanced calculus*. CRC press, 2018.
- [206] Ken Hayami. Variable transformations for nearly singular integrals in the boundary element method. *Publications of the Research Institute for Mathematical Sciences*, 41(4):821–842, 2005.
- [207] D.A. Dunavant. High degree efficient symmetrical gaussian quadrature rules for the triangle. *International journal for numerical methods in engineering*, 21(6):1129–1148, 1985.
- [208] Earl G Williams and J Adin Mann III. *Fourier acoustics: sound radiation and nearfield acoustical holography*, 2000.
- [209] Matthew J Morse, Abtin Rahimian, and Denis Zorin. A robust solver for elliptic PDEs in 3D complex geometries. *Journal of Computational Physics*, 442:110511, 2021.
- [210] Nail A Gumerov, Shoken Kaneko, and Ramani Duraiswami. Recursive Computation of the Multipole Expansions of Layer Potential Integrals over Simplices for Efficient Fast Multipole Accelerated Boundary Elements. *Journal of Computational Physics*, 486(1):112118, 2023.
- [211] Milton Abramowitz, Irene A Stegun, and Robert H Romer. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 56. American Journal of Physics, 1988.
- [212] Nail A. Gumerov and Ramani Duraiswami. Fast multipole method for the biharmonic equation in three dimensions. *Journal of Computational Physics*, 215(1):363–383, 2006.