



# THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Source-Receiver Wavefield Interferometry  
in Scattering Media

**Katrin Lör**

Doctor of Philosophy



THE UNIVERSITY  
*of* EDINBURGH

2015



Alles Wissen und alles Vermehren unseres Wissens  
endet nicht mit einem Schlusspunkt,  
sondern mit einem Fragezeichen.

*Hermann Hesse*

All knowledge and all growth of our knowledge  
does not end with a full stop  
but with a question mark.

*Hermann Hesse*



# Lay Summary

Seismic waves are waves of energy generated by seismic sources such as earthquakes or explosions. They can propagate through the subsurface over large distances and down to many kilometers of depth before they are recorded at so-called seismometers (or simply receivers) at the surface. Seismic wave records can be analysed to determine the origin of an earthquake, for example, but they also contain information about the medium the waves have travelled through. Hence, seismologists can use seismic waves to obtain an image of the Earth's interior, just as doctors use ultrasonic waves to examine the interior of a human body.

Conventional seismic tools have recently been complemented by a new set of methods, commonly referred to as seismic interferometry, which provides a means to construct synthetic measurements of seismic waves that have not been recorded. Interferometry uses combinations of other wavefields caused either by active seismic sources or by the ambient seismic noise wavefield, which is continuously generated, for example by wind, traffic, or the ocean swell. The first and revolutionary application of interferometry was the construction of a seismic signal propagating between two receivers as if one of them had been a seismic source, using recordings of ambient noise only.

This thesis focuses on an advanced version of interferometry called source-receiver interferometry (SRI). It constructs the signal between a source and a receiver that has not been recorded using wavefields propagating from and to surrounding sources and receivers. Using synthetic data generated by a wavefield modeling code on a computer, I investigate how the method performs in a medium containing one or more scatterers or reflecting interfaces. A scatterer is a small object that interacts with the waves and scatters the incoming energy in every direction. In the context of the Earth, a scatterer can be a small void or a sharp edge of a geological layer; in medical applications it would correspond to a gall or a kidney stone, for example. Large objects or layered interfaces, on the other hand, cause specular reflections. In my thesis I gradually increase the complexity of the numerical models used, starting with a single scatterer experiment, over multiple scatterers, and finishing with a multiple layer scenario.

I find that SRI provides valuable information about the scattered wavefield in all three cases even if theoretical requirements of the method cannot be met (for example if surrounding sources and receivers do not form a closed boundary). For the single scatterer case I show that under conditions resembling a real seismic experiment a novel type of energy, called pseudo-physical energy, is constructed, which emulates the scattered wavefield. This type of energy also comes into play when predicting the arrival time of multiply scattered waves. Further, it can also be used to identify multiply reflected waves in a layered medium. These so-called internal multiples are a known source of error in conventional imaging schemes and need to be detected and removed from the data in order to avoid artefacts in the seismic image. Finally, the internal-multiple equation derived from SRI provides a direct link to an existing algorithm for internal-multiple prediction; however, it is computationally much more efficient.

Overall, this thesis contributes to a better understanding of source-receiver interferometry in a scattering medium, it demonstrates how the method can be used to analyse multiply scattered wavefields, and how it relates to other methods of wavefield analysis.

# Abstract

Seismic or wavefield interferometry refers to a set of methods that synthesize wavefields between pairs of receivers, pairs of sources, or a source and a receiver, using wavefields propagating from and to surrounding boundaries of sources and/or receivers. Starting from cross-correlations of ambient seismic noise recordings, which provide the signal between two receivers as if one of them had been an active source, interferometric methods developed rapidly within the last decade, revolutionizing the way in which seismic, acoustic, elastic, or electromagnetic waves are used to image and monitor the interior of a medium. Only recently, an explicit link was found between the methods of source-receiver interferometry (SRI) and seismic imaging, a technique widely used in seismic exploration to map diffractors and reflectors in the subsurface, but also in more academic studies investigating, for example, deep crustal processes. This link is particularly interesting because SRI, in contrast to classical imaging schemes, does not rely on the single-scattering assumption but accounts for all multiple-scattering effects in the medium. While first non-linear imaging schemes based on SRI have been proposed, the full potential of the method remains to be explored and a number of open questions concerning, for example, the role of non-physical energy in interferometric wavefield estimates, require further investigation.

The aim of this thesis is to gain more insight into the method of source-receiver interferometry in the context of wavefield construction and analysis in multiply scattering media, especially when theoretical requirements of the method (such as complete boundaries of sources and receivers, surrounding the medium of interest) are not met. First I analyse the single diffractor case using partial surface boundaries only. I find that only two out of eight terms of the SRI equation are required to construct a robust estimate of the scattered wavefield, and that one of these two terms is also used in seismic imaging. The other term provides a pseudo-physical estimate of the scattered wave; this is a new type of non-physical energy that emulates the kinematics of a physically scattered wave. I then proceed to a multiple scattering scenario, using the pseudo-physical term to predict the travel times and exact scattering paths of multiply diffracted waves. The presented algorithm is purely data-driven and fully automated and, as a by-product,

provides a new tool to isolate primary diffracted waves from a complex multiply diffracted wavefield. Finally, the concept is expanded to multiply reflecting media. In reflection seismic data, multiply reflected waves should be removed prior to migration in order to avoid artefacts in the seismic image. I demonstrate how internal multiples can be estimated and attenuated using pseudo-physical energy constructed from SRI. Moreover, an explicit link is derived between the internal-multiple equation based on SRI and the internal-multiple equation derived from the inverse-scattering series (ISS), currently the most capable algorithm for internal-multiple attenuation. Using the insight provided by the SRI approach, I suggest an alternative equation that estimates internal multiples more efficiently compared to the current method.

Overall, this thesis improves our understanding of how physical, non-physical, and pseudo-physical wavefields are constructed in SRI, how new information about multiply scattered wavefields can be inferred, and how SRI relates to other methods of wavefield analysis, in particular seismic imaging and the ISS.

# Declaration

I confirm that the work submitted is my own, except where work which has formed part of jointly-authored publications has been included. My contribution and the contribution of other authors to this work has been explicitly indicated below. I confirm that appropriate credit has been given within the thesis where reference has been made to the work of others. I declare that this thesis has not been submitted, either in whole or in part, in any previous application for a degree.

This thesis contains three chapters that have been published in peer-reviewed journals or are intended for submission. I, as lead author, have done the writing of the manuscripts and performed numerical experiments, mathematical derivations, and analyses as described below. Co-authors gave advice and support on the scope of the project, provided background knowledge, contributed computer codes, and helped editing manuscripts.

Chapter 2 is published as

- K. Löer, G.A. Meles, A. Curtis, and I. Vasconcelos. Diffracted and pseudo-physical waves from spatially limited arrays using source-receiver interferometry (SRI). *Geophysical Journal International*, 196(2): 1043–1059, 2014. DOI: 10.1093/gji/ggt435

The numerical wavefield simulation for the construction of synthetic data as shown in this chapter was carried out by me using and modifying an analytic wavefield modeling code (Foldy, 1945; Galetti et al., 2013). The stationary phase analyses for the case of incomplete boundaries in source-receiver interferometry (SRI) that explains the non-physical energy in the interferometric wavefield estimate and its potential benefit for the localization of the scatterer is directly attributable to me. Further, benefitting from insightful discussions with my co-authors, I derived the mathematical proof of so-called pseudo-physical energy in interferometric wavefields.

Chapter 3 is published as

- K. Löer, G.A. Meles, and A. Curtis. Automatic identification of multiply

diffracted waves and their ordered scattering paths. *Journal of the Acoustical Society of America*, 137(4): 1834–1845, 2015. DOI: 10.1121/1.4906839

The automation of the process that identifies multiply scattered waves and their ordered scattering path is my work. In particular, I developed and implemented a new method to estimate individual primary diffraction move-outs by cross-correlation of multiple data gathers. I created the synthetic data sets used to demonstrate the applicability of the algorithm using an analytic wavefield modeling code (Foldy, 1945; Galetti et al., 2013) and a finite difference acoustic modeling code written by Matteo Ravasi (PhD student, University of Edinburgh).

Chapter 5 has been submitted as

- K. Lör, G.A. Meles, and A. Curtis. Relating source-receiver interferometry to the inverse-scattering series provides a new method to estimate internal multiples. *Geophysics*.

The derivation of the internal multiples equation based on source-receiver interferometry has been developed by me. While the alternative representation that allows more efficient computation of internal multiples was originally proposed by my co-author Giovanni A. Meles, the link between this equation and the equation by Weglein et al. (1997) based on the inverse-scattering series has been formulated by me. I also wrote the Matlab codes that perform internal multiples prediction. The synthetic data set the method was tested on was provided by Total S.A.

Katrin Lör  
September 2015

# Acknowledgments

It was probably at a Geophysics seminar in Münster, where I was a Bachelor student at the time, that Andrew Curtis's enthusiastic lecture about something I had never heard of before became the origin of the vague but enthralling idea of doing a PhD in seismic interferometry. More than five years later it still seems unreal to me that I should have done it, and first of all I would like to thank Tine Thomas who encouraged me to go to Edinburgh in the first place. Today, I owe my sincere gratitude to my supervisor Andrew Curtis for giving me the opportunity to spend three fantastic years working with him and an excellent team of fellow PhD students and researchers. Thanks for your profound feedback and encouragement during this time. And thank you for challenging us both as researchers and as human beings.

*Wer fertig ist, dem ist nichts recht zu machen; ein Werdender wird immer dankbar sein* (Johann Wolfgang von Goethe, in Faust)—I am particularly thankful for the guidance and support provided by my second supervisor Giovanni Meles. I will gladly remember our fruitful discussions on beauty, trees and, of course, interferometry. They've always been a source of inspiration to me.

Thank you Carlos, Claire, Erica, Lizzie, Matteo and Zara, for your help with codes and maths and with all the small things in a PhD student's life that become so much more endurable when you can share them with fellow-sufferers... It's been great fun working with you, all the best for the rest of your PhDs and the life beyond!

I highly appreciate the stimulating diversity of research within the School of Geo-Sciences and the University of Edinburgh as a whole. Numerous seminars, talks, discussions and conferences continuously helped me to challenge and acknowledge the bigger picture of my own work. Thanks to everyone in the School who made this a great place to work and learn in. Special thanks to all the PhD students in the Grant Institute (in particular Axel and Johannes), who shared with me the delights and desperation of a PhD (and occasional coffee breaks), and to Iain Woodhouse for an enriching exchange of views and ideas.

I would like to thank Alexander Kritski, Ali Tura, Ed Kragh, Emilien Bruthiaux

and Ivan Vasconcelos from the Edinburgh Interferometry Project (EIP) sponsors (ConocoPhillips, Schlumberger Cambridge Research, Statoil and Total) for supporting my research and for insightful discussions and constructive comments at our regular meetings, which I have always enjoyed. I also thank the people from Total in Pau (France), in particular Emilien Bruthiaux, Peigen Xie and Jean-Luc Boelle, who welcomed and supported me during my three-month internship, and Philippe Roux, who provided me with the opportunity to use the laboratory facilities in Grenoble, and who helped me a lot in interpreting the tricky bits of the data.

Andy, Em, Julia, Liv, Mark, Matt, Stu, Viv—I will surely miss our Sunday brunches, weekends in the hills, barbecues in the Meadows, nights in the pub... But first and foremost I feel extremely lucky that I have you as friends! Your help and reassurance helped me through the tough times. Thanks to Nic, the best flatmate and landlady in the world. I had the best home I could wish for, a perfect place to relax and unwind. I will always remember your Saturday morning jazz tunes. Massive credits to Calum, Claire, Dougal, Ian, Jamie, Jim, Liv and Paul of Edinburgh's most fabulous funk/soul/R&B band Cat Named Jack—your good vibes helped me forget about seismic waves for a while when I needed it. I had a helluva time with you! Everyone back home (I'm sure I'd miss someone if I started listing names), I am really thankful for your numerous visits, hour-long skype dates, your (pretended) interest in my work and mental support. You made me feel like I was never really far away.

*Can you explain in a few sentences and in simple terms what you do?*—Despite all the thesis writing workshops I've visited, which are supposed to prepare you for this most difficult of all questions, I am not convinced I've ever managed to provide you, my parents, with a satisfying answer. All the more astonishing that during all my studies you granted me your unconditional trust and support. Knowing I have a place to come back to, no matter what, gave me the strength to carry on. I cannot but be extremely grateful.

In amazement and disbelief I look back on these three most intense years I spent in Scotland and I am extremely grateful that I could share this adventure with you, Martin. Thank you for accepting the challenge. Thank you for pushing my strengths and allowing my weaknesses. Thanks for your faith and your patience. I could never have made it without you.



# Contents

<b>Lay Summary</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Declaration</b>	<b>vii</b>
<b>Acknowledgments</b>	<b>ix</b>
<b>Contents</b>	<b>xi</b>
<b>List of Tables</b>	<b>xiv</b>
<b>List of Figures</b>	<b>xviii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The principle of interferometry . . . . .	4
1.2 Source-receiver interferometry . . . . .	9
1.3 Outline of the thesis . . . . .	12
1.4 Computer Codes . . . . .	13
1.5 Publications . . . . .	14
<b>2 Diffracted and pseudo-physical waves in source-receiver interferometry</b>	<b>16</b>
2.1 Introduction . . . . .	17
2.2 Source-receiver interferometry in a scattering medium . . . . .	20
2.3 Non-physical and pseudo-physical energy . . . . .	25
2.4 Numerical examples . . . . .	29
2.5 Discussion . . . . .	37
2.6 Conclusions . . . . .	46
<b>3 Automatic identification of multiply diffracted waves and their ordered scattering paths</b>	<b>47</b>
3.1 Introduction . . . . .	48
3.2 Theory: Fingerprinting diffractors . . . . .	50
3.3 The automated scheme . . . . .	54

3.3.1	Isolating primary move-outs using cross-correlation of gathers	54
3.3.2	Detecting multiples using semblance analysis . . . . .	60
3.3.3	Predicting travel times of higher-order multiples . . . . .	61
3.4	Numerical examples . . . . .	65
3.5	Discussion . . . . .	69
<b>4</b>	<b>Automatic identification of diffracted waves</b>	
	<b>in acoustic laboratory data</b>	<b>75</b>
4.1	Acoustic finite-difference modeling tests . . . . .	75
4.2	Geometry and acquisition of the laboratory experiment . . . . .	83
4.3	Analysis of raw data and interpretation of artefacts . . . . .	86
4.4	Automatic identification of primaries and secondaries . . . . .	89
4.5	Discussion . . . . .	99
4.6	Conclusions . . . . .	102
<b>5</b>	<b>Computationally efficient internal multiple prediction</b>	
	<b>based on source-receiver interferometry</b>	<b>103</b>
5.1	Introduction . . . . .	104
5.2	An equation for internal multiples derived from the inverse-scattering series . . . . .	106
5.2.1	Forward scattering theory . . . . .	106
5.2.2	Inverse scattering and the internal multiples generator . . . . .	108
5.2.3	A migration-demigration process . . . . .	109
5.2.4	An equation for internal multiples derived from the ISS . . . . .	110
5.3	An equation for internal multiples derived from source-receiver interferometry . . . . .	113
5.3.1	Introduction to SRI . . . . .	113
5.3.2	The multiple condition . . . . .	115
5.3.3	The equation for internal multiples from SRI . . . . .	116
5.4	Alternative representation . . . . .	118
5.4.1	Correlation- and convolution-type representation of internal multiples . . . . .	118
5.4.1.1	Cross-correlation and convolution uncoupled . . . . .	120
5.5	Examples from a synthetic data set . . . . .	122
5.6	Computational cost . . . . .	126
5.7	Discussion . . . . .	128
5.8	Conclusions . . . . .	130

<b>6 Discussion</b>	<b>131</b>
6.1 Cancellation of pseudo-physical energy in SRI and the optical theorem . . . . .	131
6.2 Alternative forms of SRI . . . . .	134
6.3 SRI and Marchenko autofocusing . . . . .	135
6.4 Primary move-out estimation and interferometry . . . . .	137
6.5 Open issues in internal multiple prediction from SRI . . . . .	138
<b>7 Conclusions</b>	<b>139</b>
<b>Bibliography</b>	<b>142</b>
<b>Appendices</b>	<b>154</b>
<b>A The method of stationary phase</b>	<b>154</b>
<b>B Pseudo-physical energy</b>	<b>155</b>
<b>C Supporting Information</b>	<b>158</b>

# List of Tables

2.1	Stationary point pairs for different terms in the SRI equation. . .	24
2.2	Key to symbols used to denote parameter constellations employed in each numerical example. . . . .	32
3.1	Estimated and true travel times of primary arrivals on the mutual trace in the 3-diffractors example. . . . .	63
3.2	Estimated and true travel times of secondary arrivals on the mutual trace in the 3-diffractors example. . . . .	63
3.3	Estimated and true travel times of primary arrivals on the mutual trace in the 3-diffractors example (heterogeneous background). . .	70
3.4	Estimated and true travel times of secondary arrivals on the mutual trace in the 3-diffractors example (heterogeneous background). . .	70
4.1	Material parameters tested in numerical scattering experiments. .	77
4.2	Estimated and true travel times of secondary arrivals on the mutual trace (1 mm-scatterer, aluminium). . . . .	81
4.3	Comparison of measured and computed scatterer locations in the laboratory experiment. . . . .	85
4.4	Estimated and true travel times of primary arrivals on the mutual trace in the 3-diffractors example (laboratory data). . . . .	98
4.5	Estimated and true travel times of secondary arrivals on the mutual trace in the 3-diffractors example (laboratory data). . . . .	98
6.1	Travel times and amplitudes of physical, non-physical and pseudo-physical events provided by the scattered terms in the SRI equation.	132
6.2	Overview of complex conjugated forms of the SRI equation. . . .	135

# List of Figures

1.1	Example seismogram recorded on a broadband seismometer. . . .	2
1.2	Geometry of a numerical inter-receiver interferometry experiment.	5
1.3	Correlograms showing the contribution of each source on a circular boundary to the interferometric estimate. . . . .	7
1.4	Conceptual geometry of source-receiver interferometry (SRI). . . .	10
2.1	Three example canonical configurations used in wavefield interferometry. . . . .	18
2.2	Geometry for SRI with complete boundaries. . . . .	21
2.3	Stationary points for SRI in a homogeneous medium containing a single point scatterer. . . . .	24
2.4	Generation of non-physical and pseudo-physical energy from different terms in the SRI equation. . . . .	26
2.5	Pseudo-physical event constructed from the term $G_S G_S G_S^*$ using 2D Green's functions. . . . .	28
2.6	Pseudo-physical event constructed from the term $G_S G_S G_S^*$ using 3D Green's functions. . . . .	29
2.7	Geometry of SRI used for numerical examples. . . . .	30
2.8	SRI estimate of the Green's function between a source and a receiver compared to the true Green's function. . . . .	32
2.9	Result of integrating each term of the SRI equation separately compared to the true Green's function between a source and a receiver. . . . .	33
2.10	Correlation gathers for each term of the SRI equation. . . . .	35
2.11	SRI estimate after eliminating $G_0^*(\mathbf{x}', \mathbf{x})$ from the data. . . . .	38
2.12	Pseudo-physical and non-physical events for different horizontal scatterer positions . . . . .	39
2.13	Correlation gathers of $G_0 G_0 G_S^*$ for different parameter constellations.	41
2.14	Correlation gathers of $G_0 G_0 G_0^*$ for different parameter constellations.	41
2.15	(Pseudo-)Physical scattered wave energy constructed using the terms $G_0 G_0 G_S^*$ and $G_S G_S G_S^*$ only compared to the true scattered wave. . . . .	42

2.16	Geometry used in multiple diffractors example. . . . .	45
2.17	Pseudo-physical events in multiple diffractors example constructed using $G_S G_S G_S^*$ . . . . .	45
3.1	Example geometry for a 3-diffractors model. . . . .	51
3.2	Synthetic CSG, CRG and sketch of the data cube. . . . .	52
3.3	Schematic construction of multiply diffracted events from primaries and secondaries. . . . .	53
3.4	CSGs with different source locations and result of cross-correlation between the two gathers. . . . .	55
3.5	Cross gathers for different time lags. . . . .	58
3.6	Result of semblance analysis for a CSG. . . . .	60
3.7	Mutual trace of a CSG and a CRG with identified primaries, sec- ondaries and higher-order multiples. . . . .	64
3.8	Percentage of primaries and secondaries identified correctly . . . .	66
3.9	Density and velocity distribution in a 3-diffractors model with a heterogeneous background structure. . . . .	67
3.10	CSG and corresponding cross-gathers obtained from the 3-diffractors example with heterogeneous background. . . . .	68
3.11	Mutual trace of the CSG and the CRG obtained from the 3- diffractors example with heterogeneous background. . . . .	71
4.1	Geometry of the velocity and density model used in FD modeling. . . .	76
4.2	Results from the single scattering modeling experiment testing scatterers with different diameters and densities. . . . .	76
4.3	Velocity and density model used in the 3-scatterers case with scat- terers of diameter $d = 1$ mm. . . . .	79
4.4	Synthetic CSG and CRG generated from the 3-scatterers model with 1 mm-scatterers. . . . .	79
4.5	Primaries estimated from the 3-scatterers case with 1 mm-scatterers. . . .	80
4.6	Estimated primaries and secondaries on the mutual trace compared to the true arrivals. . . . .	81
4.7	Velocity and density model used in the 3-scatterers case with scat- terers of diameter $d = 6$ mm. . . . .	82
4.8	Synthetic CSG and CRG generated from the 3-scatterers model with 6 mm-scatterers. . . . .	82
4.9	Plan view of acquisition set up for an acoustic multiple scattering laboratory experiment. . . . .	84

4.10	Spurious events in single-scattering laboratory data. . . . .	87
4.11	Spurious scattered wave observed at three different CSGs. . . . .	88
4.12	Data example for the 2-scatterers case. . . . .	90
4.13	Primaries estimated from the 2-scatterers case. . . . .	91
4.14	Data example for the 3-scatterers case. . . . .	92
4.15	Primaries estimated from the 3-scatterers case. . . . .	93
4.16	Data example for the 4-scatterers case. . . . .	94
4.17	Primaries estimated from the 4-scatterers case. . . . .	95
4.18	Data example for the 5-scatterers case. . . . .	96
4.19	Primaries estimated from the 5-scatterers case. . . . .	97
4.20	Tertiaries on the mutual trace of the 3-scatterers case. . . . .	100
5.1	Example geometry for a third-order scattering event. . . . .	108
5.2	Conceptual geometry for SRI using complete or partial boundaries. . . . .	114
5.3	Raypaths of primaries used in the multiple condition and raypath of internal multiple. . . . .	115
5.4	Velocity and density profile of the acoustic model. . . . .	123
5.5	Internal multiples estimated from the ISS for a collocated source and receiver. . . . .	124
5.6	Internal multiples estimated from SRI for a collocated source and receiver. . . . .	125
5.7	Estimate of internal multiples for a common-shot gather. . . . .	127
5.8	Demultiplied scattered wavefield for a common-shot gather. . . . .	127
B.1	Sketch of ray paths constructing $G_S(\mathbf{x}_2, \mathbf{x}_1)$ from inter-source in- terferometry. . . . .	156
B.2	Sketch of ray paths constructing $G_0^*(\mathbf{x}', \mathbf{x}_2)$ from interreceiver in- terferometry. . . . .	157
B.3	Sketch of ray paths constructing $\hat{G}_0^*(\mathbf{x}', \mathbf{x}_2)$ from interreceiver in- terferometry. . . . .	157
C.1	Interferometric result of the full integral $GGG^*$ . . . . .	159
C.2	Example set of raypaths used in $G_0G_0G_S^*$ . . . . .	160
C.3	Interferometric result of the cross term $G_0G_0G_S^*$ . . . . .	160
C.4	Example set of raypaths used in $G_SG_SG_S^*$ . . . . .	162
C.5	Interferometric result of the cross term $G_SG_SG_S^*$ . . . . .	162
C.6	$G_0G_SG_S^*$ for different apertures of source and receiver boundaries. . . . .	163
C.7	Example set of raypaths used in $G_0G_SG_S^*$ . . . . .	164
C.8	Interferometric result of the cross term $G_0G_SG_S^*$ . . . . .	164

C.9	Example set of raypaths used in $G_S G_0 G_S^*$ .	165
C.10	Interferometric result of the cross term $G_S G_0 G_S^*$ .	165
C.11	$G_0 G_0 G_0^*$ for different depths of the source-receiver pair.	166
C.12	Example set of raypaths used in $G_0 G_0 G_0^*$ .	167
C.13	Interferometric result of the cross term $G_0 G_0 G_0^*$ .	167
C.14	Example set of raypaths used in $G_S G_S G_0^*$ .	169
C.15	Interferometric result of the cross term $G_S G_S G_0^*$ .	169
C.16	Example set of raypaths used in $G_0 G_S G_0^*$ .	170
C.17	Interferometric result of the cross term $G_0 G_S G_0^*$ .	170
C.18	Example set of raypaths used in $G_S G_0 G_0^*$ .	172
C.19	Interferometric result of the cross term $G_S G_0 G_0^*$ .	172



# Chapter 1

## Introduction

Seismic wavefields, generated by earthquakes or artificial sources of energy, provide a means to illuminate the Earth's interior and create an image of the subsurface: just as doctors use ultrasonic waves to examine the interior of the human body, seismologists analyse seismic waves to examine the Earth's internal structures and processes. While serving a general scientific interest in the origin, history and development of our planet, knowing what the Earth looks like beneath our feet also plays a key role in a number of societal issues, ranging from natural hazard assessment (earthquakes, volcanoes, landslides) through natural resource exploration and production (water, minerals, ores, oil, gas, geothermal energy) to subsurface waste storage (nuclear waste, CO<sub>2</sub>). In all of these cases information about the subsurface is vital.

Traditional methods of seismic exploration use active, natural or artificial, seismic sources and one or more receivers (seismometers) to record the energy that propagates from the source through the medium (the subsurface) to the receiver. These recordings, also called seismograms or seismic traces, provide the displacement, velocity, or acceleration of the medium caused by a passing wave as a function of time (Fig. 1.1). The distinct arrivals (packets of energy) on a trace can be associated with different components of the wavefield (surface waves, body waves, reflected and diffracted waves) and provide information, for example, about the seismic velocity of the medium or the location of geological structures.

A relatively novel method called *seismic interferometry* uses the *passive* seismic wavefield, the so-called seismic noise, to infer information about the subsurface. Seismic noise refers to the permanent vibrations of the ground generated by wind, rain, traffic, human and industrial activity, and the oceanic microseism, but is also used for the multiply scattered energy in the seismic coda (the “tail” of the signal, Fig. 1.1), which until recently has been too complex to use or to interpret. The origins of seismic or wavefield interferometry can be traced back to Claerbout

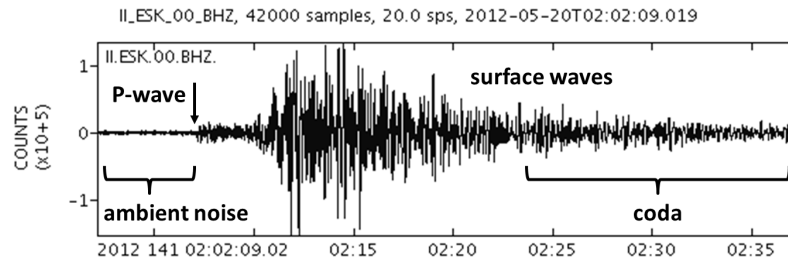


Figure 1.1: Seismogram of an earthquake in Northern Italy recorded on the vertical component of a broadband seismometer (station ESK) in Scotland. Indicated on the seismic trace are the ambient noise wavefield preceding the first arriving body wave (P-wave), the large amplitude surface waves, and the seismic coda. The data is provided by the IRIS Data Management Center<sup>1</sup>.

(1968) who showed that the surface reflection response of a layered medium can be constructed from the autocorrelation of a transmitted wave emitted by a source located below the bottom layer. From his findings Claerbout inferred that it should be possible to construct the wavefield between *two receivers* as if one of them had in fact been a source, by cross-correlating the background noise recorded at the two receiver locations. However, almost three decades passed before the idea was successfully proven in helioseismology: Duvall et al. (1993) were able to extract time-distance information from cross-correlations of intensity fluctuations (“solar noise”) measured on the sun’s surface. The first laboratory demonstrations followed a few years later (Lobkis and Weaver, 2001; Derode et al., 2003a; Larose et al., 2005) supported by the mathematical proof for acoustic (Wapenaar, 2003; Van Manen et al., 2005) and elastic media (Wapenaar, 2004; Van Manen et al., 2006).

Campillo and Paul (2003) were the first to demonstrate the method in a seismological context using the late seismic coda of local earthquakes to construct inter-receiver surface waves between all the stations of a seismic network. Sabra et al. (2005a,b) and Shapiro et al. (2005) constructed inter-receiver surface waves from ambient seismic noise caused by the oceanic microseisms. The constructed wavefields were then used in a standard tomographic inversion scheme to produce the first seismic velocity maps from ambient noise. Since then ambient noise surface wave tomography has become a popular tool to seismologically map regions with low seismic activity that were previously “invisible” to the seismologists (e.g. Yang et al., 2007; Bensen et al., 2008; Stehly et al., 2009; Nicolson et al., 2012).

<sup>1</sup>Incorporated Research Institutions for Seismology, <http://www.iris.edu>

Not only did seismic interferometry revolutionize the field of seismic tomography, but it transformed the way in which seismologists think about both passive and active wavefields and how they use them to infer information about the subsurface. Over the last decade interferometric methods developed rapidly within the seismology community and found application, for example, in seismic exploration (Bakulin and Calvert, 2006; Halliday et al., 2007; Halliday et al., 2010), velocity analysis (King and Curtis, 2011; King et al., 2011), near-surface geophysics (Galetti and Curtis, 2012; Harmankaya et al., 2013; Kaslilar et al., 2014), seismic monitoring (Snieder, 2002; Sens-Schönfelder and Wegler, 2006; Draganov et al., 2012) and imaging (Vasconcelos, 2008; Sava and Vasconcelos, 2011; Fleury and Vasconcelos, 2012). Other fields of physics such as electromagnetics (e.g. Slob and Wapenaar, 2007; Hunziker et al., 2012) or biophysics (Sabra et al., 2007) also benefitted from the advances in interferometry.

The link between interferometry and seismic imaging was made explicit by Halliday and Curtis (2010) after a new type of interferometry called *source-receiver interferometry* (SRI) had been introduced (Curtis and Halliday, 2010). SRI constructs the wavefield between a real source and a real receiver using only energy that has travelled from and to surrounding boundaries of sources and receivers. The SRI wavefield representation is particularly interesting because, in contrast to most imaging schemes, it does not rely on the single-scattering Born approximation and hence accounts for all multiple scattering in the medium. The term “scattering” generally refers to the interactions of the wavefield with heterogeneities inside the medium that influence the wave’s propagation. In general, one distinguishes between *diffracted* energy which originates from structures of small spatial extent compared to the wavelength (e.g. angular boundaries, voids, faults, or fractures), and *reflected* energy which originates from structures such as interstrata interfaces.

This thesis investigates some interesting issues that arise from the theoretical link between the two domains of interferometry (SRI in particular) and imaging, and addresses open questions regarding the use of SRI in the construction and interpretation of scattered wavefields: How is the scattered wave construction from SRI affected if theoretical requirements of interferometry cannot be met in practice? What is the role of non-physical energy that is introduced by not meeting theoretical requirements and how can it either be used or suppressed? What are the implications for a multiple scattering scenario? Can SRI help to interpret a complex multiply scattered wavefield? And finally, can the theory

developed for diffracted waves be extended to the case of reflection data?

In section 1.2, I provide further detail on the theory and current state of SRI, and give a brief introduction to seismic imaging and the link between the two concepts. In section 1.3 I summarise how the different chapters of this thesis address the questions raised above. Thereafter, I briefly describe the programming tools that I used and provide a list of chapters of this thesis that have been published as jointly authored papers.

However, first let us begin with an introduction to standard interferometry.

## 1.1 The principle of interferometry

*The term interferometry generally refers to the study of interference phenomena between pairs of signals in order to obtain information from the differences between them.*

(Curtis et al., 2006). Interference phenomena are better known in the field of optics, with Young's double slit experiment being the most prominent example (a demonstration of the principle of wave-particle duality): a beam of light is split into two waves when passing through a double slit. A screen is located behind the slit that displays the interference pattern of the two superimposing waves resulting from constructive and destructive summation of the waves' crests and troughs.

Seismic interferometry, as the name suggests, refers to interference phenomena of seismic waves propagating in the Earth's interior and along its surface. In fact, interference in the context of seismic interferometry does not refer to the physical superposition of waves but rather to a mathematical operation: a wavefield is recorded at two different locations and the two recordings are combined in a way that removes what is similar and highlights what is different between them. The mathematical tool used to do this is an operation called cross-correlation. Differences in the recordings are due to the different locations of the two receivers and allow us to infer the wavefield between the two locations.

The canonical geometry of standard interferometry comprises a pair of receivers surrounded by a closed boundary of sources, which generate the wavefield (Fig. 1.2). The sources could either represent individual active sources (active source interferometry) or ambient noise sources (passive or ambient noise interferometry).

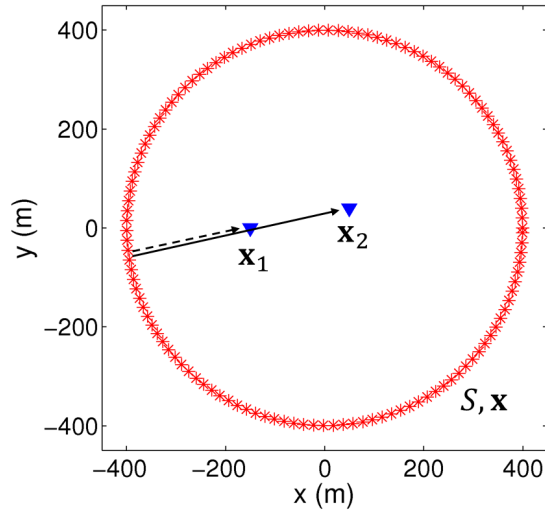


Figure 1.2: Geometry of a numerical inter-receiver interferometry experiment. Stars are sources, triangles are receivers. Cross-correlation and summation of the wavefields recorded at the two receivers provides the homogeneous Green’s function between the two receiver as if one receiver had been an impulsive source. Arrows indicate raypaths (direction of wave propagation) between one boundary source and the two receivers. Travel times of dashed and solid raypaths that run parallel cancel each other in the cross-correlation process. For the rays shown, this produces a result with the travel time of the ray between  $\mathbf{x}_1$  and  $\mathbf{x}_2$ .

Active source interferometry consists of only two steps: (a) the cross-correlation of signals from each source recorded at the two receivers, and (b) the summation (integration) of the cross-correlation results from all sources. Note that in passive interferometry the summation over sources is replaced by temporal stacking of long recording times since the sources are already (randomly) superposed or stacked. The following derivations apply to active source interferometry.

Provided a few assumptions are met, which I define later, performing these two steps gives the signal between the two receivers as if one receiver had in fact been an impulsive source. Therefore, the method is also referred to as *inter-receiver* interferometry. The constructed signal is the homogeneous *Green’s function*. A Green’s function denotes the wavefield measured at a receiver if the source is a delta impulse (a “peak” in space and time). It is also known as the medium’s response to an impulsive source, as it is entirely defined by the properties of the medium and independent from the source function. The Green’s function comprises all possible wave types propagating between source and receiver (surface waves, direct wave, scattered waves, etc.) and allows us to derive information about the medium between the two locations. The *homogeneous* Green’s function refers to the superposition of the causal and acausal (time-reversed) Green’s

functions, which both solve the homogeneous wave equation.

Mathematically, the interferometric construction of a homogeneous Green's function  $G(\mathbf{x}_2, \mathbf{x}_1, \omega) + G^*(\mathbf{x}_2, \mathbf{x}_1, \omega)$  between two receivers at  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , respectively, can be expressed in the frequency domain as

$$G(\mathbf{x}_2, \mathbf{x}_1, \omega) + G^*(\mathbf{x}_2, \mathbf{x}_1, \omega) \approx \frac{2}{\rho c} \int_S G^*(\mathbf{x}_1, \mathbf{x}, \omega) G(\mathbf{x}_2, \mathbf{x}, \omega) d^2\mathbf{x} \quad (1.1)$$

(Wapenaar and Fokkema, 2006), where  $\omega$  is the angular frequency,  $\mathbf{x}$  denotes the source locations on boundary  $S$ ,  $\rho$  and  $c$  are the density and the velocity of the medium, respectively. The integration performs summation over source locations  $\mathbf{x}$ .  $G(\mathbf{x}_1, \mathbf{x}, \omega)$  denotes the Green's function between a source at  $\mathbf{x}$  and a receiver at  $\mathbf{x}_1$  in the frequency domain. The star  $*$  denotes complex conjugation, which is equivalent to time reversal in the time domain. In the frequency domain, cross-correlation is defined as the product of two functions, one of which has been complex conjugated, which is exactly what we see in the integrand. Note that when multiplying a complex conjugated function  $A_1^* e^{-i\omega t}$  with a function  $A_2 e^{i\omega t'}$  the travel times in the phase term are subtracted according to  $A_1^* A_2 e^{i\omega(t'-t)}$ , which is illustrated in Fig. 1.2 as the subtraction of raypaths. Eq. 1.1 is an approximation of the actual interferometry equation; it is quite accurate if the boundary has a very large radius such that all rays travel approximately perpendicular to the boundary.

To understand how Eq. 1.1 works, let us consider the wavefield generated by a single source on boundary  $S$  located in line with the two receivers at  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , respectively (Fig. 1.2). At the first receiver ( $\mathbf{x}_1$ ) we record the energy that has propagated from the source directly to that receiver (dashed arrow). At the second receiver ( $\mathbf{x}_2$ ) we record the energy that has propagated from the source via the first receiver to the second receiver (solid arrow). Comparing the corresponding raypaths we find that they have a common part, namely the bit from the source to the first receiver. Cross-correlation cancels the travel time of the common part in the phase term, leaving the travel time of the bit from the first receiver to the second receiver intact. After this operation has been integrated around the boundary the result is as if the raypath originated at the first receiver, hence, as if the first receiver had actually been a source: the receiver at  $\mathbf{x}_1$  has become a *virtual source* (an imagined source).

If the raypaths between any other boundary source and the two receivers are considered, it becomes clear that they do not necessarily have a common part, hence

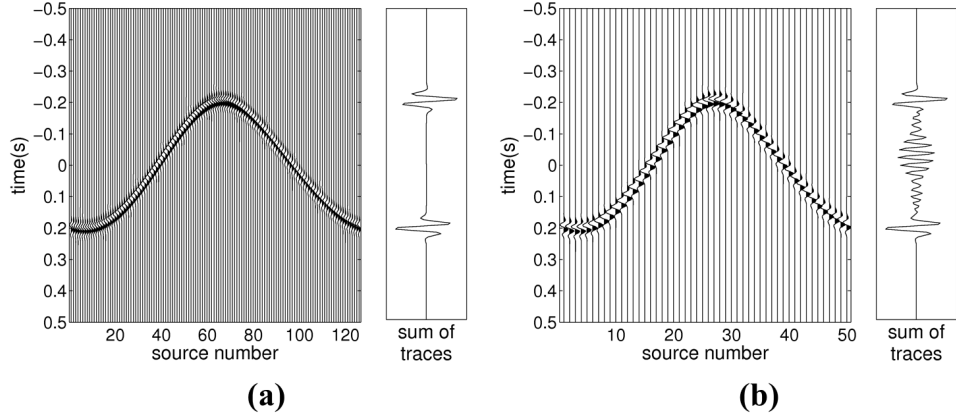


Figure 1.3: **(a)** Correlogram showing the contribution of each source on boundary  $S$  (Eq. 1.1, Fig. 1.2) to the interferometric estimate. Each trace in the left diagram is the result of the cross-correlation of signals generated by one source and recorded at the two receivers. For clarity, only every second trace is plotted. The plot in the right-hand side represents the sum of all the traces on the left, i.e. the homogeneous Green's function between the two receivers (Eq. 1.1). **(b)** As in (a), but with fewer sources on the boundary (hence, fewer traces in the correlogram) so that the Nyquist criterion is not satisfied. This introduces artefacts in the Green's function estimate (right plot) where non-stationary energy has not completely cancelled.

cross-correlation would not have the above effect. Summing over all sources, however, cancels the contribution from unfavourable source locations and preserves the contribution from the in-line sources. This process is illustrated in a so-called correlogram (Fig. 1.3), which displays the cross-correlation results from all sources on the boundary (left-hand plot in both a and b). The trace on the right is the sum of all traces on the left (the integrand in Eq. 1.1). It turns out that where the energy in the correlogram is stationary, i.e., where there is a maximum or minimum, the traces are in phase and sum constructively. All other traces cancel each other as they are out of phase, i.e., their crests and troughs sum destructively. For the direct wave, the two stationary points for which the integrand is stationary with respect to the sources on the boundary are located on the inter-receiver line: one of them accounts for the construction of the causal direct wave (as described above), the other for the construction of the acausal direct wave. A mathematical derivation of the stationary points for both direct and scattered waves in inter-receiver interferometry is given in Snieder (2004a).

From the correlogram it is clear that also the sources in the immediate neighbourhood of the inter-receiver line contribute to the interferometric construction. This is related to the Fresnel zone of any wave propagating at a finite frequency. The size of the Fresnel zone depends on the wavelength and thus the frequency

of the signal. All waves originating from within the Fresnel zone arrive in phase at the receiver and interfere constructively. What is more, if the medium contains heterogeneities, an enclosing source boundary is required to account for the construction of scattered waves. The stationary points of scattered waves depend on the locations of the scatterers and are not known in advance if the medium is unknown; this is why theory requires the source boundary to completely surround the medium of interest.

The correlogram also illustrates another requirement of interferometry: for the non-stationary contributions to cancel out completely, the inter-source spacing (spatial sampling) along the boundary has to be sufficiently small, otherwise spurious energy remains in the interferometric estimate (Fig. 1.3b). The spatial sampling is controlled by the Nyquist criterion, which states that the distance between two sources should be smaller than or equal to the Nyquist wavelength  $\lambda_{Nyq} = \frac{c}{f_{Nyq}}$ , where the Nyquist frequency  $f_{Nyq}$  is half the sampling frequency, and  $c$  is the velocity of the medium.

In practical applications of interferometry the theoretical requirements discussed above often cannot be satisfied. Typically, this introduces so-called non-physical energy into the Green's function estimate. While this type of energy does not relate directly to physical wave propagation it can still be used to infer physical information about the medium (e.g. Mikesell et al., 2009; 2012; Harmankaya et al., 2013; Meles and Curtis, 2013), provided that we understand the origin of non-physical energy, and how it relates to physical properties.

Eq. 1.1 is valid given the particular geometry shown in Fig. 1.2, however, with a few modifications other geometries are possible. Firstly, by invoking the theorem of source-receiver reciprocity, which states that the wavefield propagating between a source at  $\mathbf{x}$  and a receiver at  $\mathbf{x}_1$  is identical to the wavefield propagating between a source at  $\mathbf{x}_1$  and a receiver at  $\mathbf{x}$ , source and receiver locations in Fig. 1.2 can be interchanged, so that a pair of sources is surrounded by a boundary of receivers. Performing the same steps as above (cross-correlation and summation) provides the Green's function between the two source locations at  $\mathbf{x}_1$  and  $\mathbf{x}_2$ ; hence, this second type of interferometry is referred to as *inter-source* interferometry and turns one source into a *virtual receiver*. Secondly, if instead of cross-correlation, convolution or deconvolution is applied to the wavefields then one receiver could be located outside of the source boundary. Convolution is similar to cross-correlation, except that both functions are causal (no complex conjugation), and hence travel times in the phase term are added rather than subtracted.



Different geometries and the corresponding interferometric equations for a variety of cases are discussed, for example, in Slob and Wapenaar (2007), Vasconcelos and Snieder (2008), Wapenaar et al. (2010c) and Galetti et al. (2013).

Reciprocity and geometrical flexibility allow us to derive a third type of interferometry that is basically a combination of inter-source and inter-receiver interferometry and is referred to as *source-receiver interferometry* (SRI). While less understood, SRI further expands the range of applications of interferometric methods and provide a new perspective on problems in wavefield analysis, as will be discussed in the following section.

## 1.2 Source-receiver interferometry

Curtis and Halliday (2010) show that the homogeneous Green's function  $G(\mathbf{x}_2, \mathbf{x}_1) + G^*(\mathbf{x}_2, \mathbf{x}_1)$  between a source at  $\mathbf{x}_1$  and a receiver at  $\mathbf{x}_2$  can be represented by a double surface integral of the form

$$G(\mathbf{x}_2, \mathbf{x}_1) + G^*(\mathbf{x}_2, \mathbf{x}_1) \approx \frac{4}{(\rho c)^2} \int_S \int_{S'} G(\mathbf{x}', \mathbf{x}_1) G(\mathbf{x}_2, \mathbf{x}) G^*(\mathbf{x}', \mathbf{x}) d^2\mathbf{x}' d^2\mathbf{x} \quad (1.2)$$

where  $\mathbf{x}$  denotes a source on boundary  $S$  and  $\mathbf{x}'$  denotes a receiver on boundary  $S'$  (Fig. 1.4),  $\rho$  and  $c$  are the medium's density and velocity, respectively. All Green's functions are in the frequency domain (the dependency on  $\omega$  has been dropped for brevity). Eq. 1.2 is the source-receiver interferometry (SRI) equation in the monopole approximation. That is, it is assumed that the two boundaries have large radii such that all rays travel perpendicular to the boundaries. While at first sight the interferometric construction of a Green's function between a real source and a real receiver may seem unnecessary, I will show below that the SRI Green's function representation offers a new perspective on wavefield construction and analysis, which has proven useful in a number of different problems.

For example, in SRI the source at  $\mathbf{x}_1$  and the receiver at  $\mathbf{x}_2$  do not have to be deployed at the same time in order to obtain the Green's function between them, which makes it possible to retrospectively observe seismograms from old earthquakes on newly installed sensors (Curtis et al., 2012; Entwistle et al., 2015): the Green's function between an earthquake and a receiver deployed in the vicinity of the earthquake *after* it has happened is constructed using recordings of the

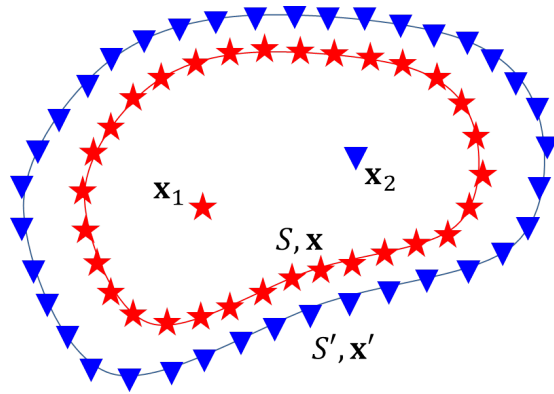


Figure 1.4: Conceptual geometry of source-receiver interferometry (SRI). Symbols as in Fig. 1.2. SRI constructs the Green’s function between the source at  $\mathbf{x}_1$  and the receiver at  $\mathbf{x}_2$  using only energy that has travelled from and to surrounding boundaries of sources and receivers  $S$  and  $S'$ , respectively.

earthquake at a backbone array of receivers and long-term recordings of ambient noise, measured at the new receiver and the backbone array. Comparisons with real recordings suggest that SRI provides a good estimate of the surface wave component of the wavefield generated by the earthquake.

Surface wave construction also plays a role in exploration seismology. Halliday et al. (2007) showed how inter-receiver interferometry provides a means to estimate and remove scattered surface waves, the so-called ground-roll in land seismic data, provided that a receiver is collocated with each source location. Duguid et al. (2011) showed that using SRI instead provides an even better estimate of the ground-roll and that SRI does not require a receiver to be collocated with the seismic source used in the survey, which could be difficult in practice. In a recent work, Halliday et al. (2015) showed how the need for an additional receiver boundary in ground-roll removal by SRI can be circumvented using a model-driven approach.

Poliannikov (2011) and Poliannikov et al. (2012) use the benefits of SRI for the construction of body wave reflections between two borehole receivers or between two earthquakes, respectively. Their technique combines inter-receiver or inter-source interferometry with SRI and was tested successfully on numerical models.

The reason why SRI seems to outperform inter-receiver interferometry in practical applications is explained by Meles and Curtis (2013). In inter-receiver interferometry, the construction of scattered waves require wavefields to be generated around specific stationary points on the boundary. In contrast, by applying sta-

tionary phase analysis to the SRI integral, Meles and Curtis (2013) show that SRI produces kinematic information about the scattered field using wavefields propagating from and to *any* point on the boundary. This property is particularly useful when boundaries are only partially available or sparsely sampled.

The theoretical analysis of scattered wave SRI also provided the explicit link between SRI and imaging. Korneev and Bakulin (2006) were the first to demonstrate the equivalence of the virtual-source method (a variety of inter-receiver interferometry used in exploration seismics) and a version of the Kirchhoff-Helmholtz integral that uses experimentally measured Green's functions. The Kirchhoff-Helmholtz integral in the representation of Korneev and Bakulin (2006) reads

$$I_{\Sigma}(M) = \frac{1}{4\pi} \int_{\Sigma} (G(S, M) \frac{\partial \tilde{u}}{\partial n}(S) - \tilde{u}(S) \frac{\partial G(S, M)}{\partial n}) dS \quad (1.3)$$

where  $\Sigma$  is a closed surface boundary,  $G(S, M)$  is the causal Green's function between a point  $S$  on the surface and an internal point  $M$  inside  $\Sigma$ , and  $\tilde{u}$  is the Fourier transform of an acoustic field  $u$ . It is the fundamental basis for seismic imaging or migration, a common technique used in exploration geophysics to map reflecting and diffracting discontinuities in the subsurface (Claerbout, 1985). The basic idea behind seismic imaging methods is the correlation of a forward propagated and a backpropagated wavefield at a point in the subsurface. Using a smooth model of the subsurface, the wavefield generated at the source is forward propagated to a point in the model space, the image point. The scattered wavefield recorded at a receiver is backpropagated to the same point. Only if the image point coincides with a true scatterer or reflector is the result of the correlation of the two wavefields (the imaging condition) non-zero at zero time.

The relationship between interferometry and imaging was further explored by Thorbecke and Wapenaar (2007) and Vasconcelos (2008). However, it was not until the development of SRI (Halliday and Curtis, 2010) that the link between the two domains was made explicit. Using new representation theorems for scattered waves (Vasconcelos, 2008) they showed that when the central source and the receiver in the SRI geometry are collocated, an SRI representation for scattered waves can be derived that is equivalent to the generalized imaging condition given by Oristaglio (1989). They also showed that while the single-scattering Born approximation is inherent to Oristaglio's formula, the interferometric version of the theory naturally accounts for non-linearities, in this case caused by multiple scat-

tering in the medium. This insight encouraged further research in the field of SRI and led to new advances in seismic imaging methods by introducing non-linear imaging conditions that account for multiply scattered and multiply converted waves (e.g. Vasconcelos et al., 2010; Fleury and Vasconcelos, 2012; Ravasi and Curtis, 2013; Ravasi et al., 2014).

A common problem in standard (linear) seismic imaging are surface-related or internal multiply reflected waves, so-called “multiples”. Similar to Oristaglio’s formula, most standard migration algorithms are linear. The linearity comes about by assuming that all waves in the data have scattered only once (are so-called “primaries”). The fact that some of the data consists of multiples introduces spurious reflectors in the seismic image. This is why multiples should be identified and removed from the data prior to linear migration. Different authors (Wapenaar et al., 2010c; Behura and Forghani, 2012) have implied a relationship between interferometry, particularly SRI, and the construction of internal multiples in seismic reflection data, but so far an explicit link has remained illusive.

### 1.3 Outline of the thesis

In this thesis the method of SRI and its potential to both construct and analyse scattered wavefields is tested in increasingly complex scenarios, starting with the conceptual single-scattering case, through multiply diffracted wavefields to multiply reflected wavefields.

Chapter 2 evaluates the performance of SRI in a medium containing a single scatterer when theoretical requirements are not met. In particular, the robustness of the method is tested when only sparsely populated surface boundaries of sources and receivers are available, as typically used in seismic exploration. Separating the homogeneous Green’s function into an unperturbed and a perturbed (scattered) part expands the SRI equation into the sum of eight terms. Each of these terms is analysed individually for its stationary points and for its contribution to the final interferometric wavefield estimate given the limited source and receiver distribution along the boundaries. I find that only two out of eight terms are required to obtain a robust estimate of the scattered part of the wavefield, and show for the first time that one of these terms provides a new type of non-physical energy that we call *pseudo*-physical: the energy is non-physical but emulates the kinematics of a physically scattered wave.

In chapter 3 an acoustic medium with multiple scatterers is considered and an automatic algorithm is presented that both identifies singly and twice scattered waves in the data, and predicts the arrival times and scattering paths of all other multiply scattered events of any order. The prediction process is based on SRI and provides a pseudo-physical estimate (as defined in chapter 2) of higher-order scattered waves by cross-correlating and convolving lower-order events. The method is entirely data driven and works also in media with a heterogeneous background velocity distribution where the travel-time curves of diffracted waves are not hyperbolic. Numerical examples are presented to support the theory.

Chapter 4 tests the applicability of the automatic algorithm introduced in chapter 3 to an acoustic data set obtained from a laboratory experiment. I present results from different scattering scenarios, considering up to five scatterers in a homogeneous background medium, and comment on the difficulties arising when using real data as compared to synthetic data.

In chapter 5 the multiples prediction process presented in chapter 3 is extended to media with reflecting interfaces in order to identify internal multiples in reflection data. Internal multiples cause artefacts in seismic imaging and are therefore required to be removed prior to linear migration. Here, a connection is established between the internal multiples equation derived from SRI and an existing formula based on the inverse-scattering series. It is demonstrated that both derivations rely on pseudo-physical energy of the kind presented in chapter 2 to obtain an estimate of internal multiples. Moreover, it is shown explicitly how the SRI perspective leads to an alternative representation of the same equation that allows one to compute internal multiples far more efficiently. This is significant in practical reflection seismology where the computational cost of applying the existing methods is extremely high, often prohibitively so. Synthetic data examples are provided to compare the existing and the alternative formula both qualitatively and quantitatively in terms of computational cost.

## 1.4 Computer Codes

I used computer codes written in Matlab<sup>2</sup> and Python<sup>3</sup> to model synthetic scattered wavefields and to analyse these wavefields using interferometric methods.

---

<sup>2</sup><http://www.mathworks.com>

<sup>3</sup><http://www.python.org>

Synthetic wavefields from scattering media with homogeneous background velocity structure were obtained from an acoustic wavefield modeling code, which is an implementation of Foldy’s method (Foldy, 1945; Galetti et al., 2013). Synthetic wavefields from scattering media with heterogeneous background velocity structure were modeled using an acoustic finite-difference (FD) code provided by Matteo Ravasi (PhD student, University of Edinburgh). The code is written using Madagascar<sup>4</sup>, an open-source software package based on Python.

I modified the Matlab code provided by Galetti et al. (2013) to compute individual terms of the SRI equation separately (chapter 2) using partial surface boundaries of sources and receivers. For chapter 3, I wrote a set of Matlab codes to perform the different steps of the automatic algorithm that analyses multiply diffracted wavefields. Acoustic experiments performed in the laboratory (chapter 4) were controlled using a Matlab code by Philippe Roux (ISTerre, University Joseph Fourier, Grenoble, France). In chapter 5, I used the synthetic wavefields obtained from a one-dimensional horizontally layered velocity and density model provided by Total S.A. The wavefields were computed from SISMOS, an analytic wavefield modeling code for relatively simple layered media, which allows one to compute the primary wavefield only. The standard and the alternative equation for internal multiple prediction were solved using a set of Matlab codes partially based on a code provided by Giovanni A. Meles (University of Edinburgh).

## 1.5 Publications

This thesis contains three chapters that have been published in or are intended for submission to peer-reviewed journals.

Chapter 2 is published as

- K. Lörer, G.A. Meles, A. Curtis, and I. Vasconcelos. Diffracted and pseudo-physical waves from spatially limited arrays using source-receiver interferometry (SRI). *Geophysical Journal International*, 196(2): 1043–1059, 2014. DOI: 10.1093/gji/ggt435

Chapter 3 is published as

- K. Lörer, G.A. Meles, and A. Curtis. Automatic identification of multiply

---

<sup>4</sup><http://www.ahay.org>

diffracted waves and their ordered scattering paths. *Journal of the Acoustical Society of America*, 137(4): 1834–1845, 2015. DOI: 10.1121/1.4906839

Chapter 5 has been submitted as

- K. Löer, G.A. Meles, and A. Curtis. Relating source-receiver interferometry to the inverse-scattering series provides a new method to estimate internal multiples. *Geophysics*.

In addition to these publications I have also contributed to the following publication:

- G.A. Meles, K. Löer, M. Ravasi, A. Curtis, and C. A. da Costa Filho. Internal multiple prediction and removal using Marchenko autofocusing and seismic interferometry. *Geophysics*, 80(1): A7–A11, 2015.  
DOI: 10.1190/GEO2014-0408.1

# Chapter 2

## Diffracted and pseudo-physical waves in source-receiver interferometry

I start the analysis of source-receiver interferometry (SRI) in scattering media considering the most simple scenario of a single point diffractor embedded in a homogeneous background medium. Source and receiver boundaries are linear arrays located above the scatterer resembling the acquisition geometry typically used in real seismic experiments. I analyse the Green's function estimate constructed from this configuration focusing on the contribution of individual terms of the SRI equation. Further, I examine the effect of partial boundaries and comment on the origin and the value of non-physical and pseudo-physical energy in the interferometric estimate. While the scenario considered in this chapter may not have an immediate application, it sheds light on the physics that underly scattered wave SRI and provides a sound basis for the analysis of more complex systems.

This chapter has been published as a jointly-authored paper<sup>1</sup>. I, as lead author, have done the writing of the paper, performed the numerical experiments, and analysed the results. Co-authors gave advice and support on the scope of the project, provided background knowledge, and helped editing the manuscript.

### Abstract

Source-receiver interferometry (SRI) refers to a technique to construct the Green's function between a source and a receiver using only energy that has travelled from and to surrounding boundaries of sources and receivers. If a background medium

---

<sup>1</sup>K. L er, G.A. Meles, A. Curtis, and I. Vasconcelos. Diffracted and pseudo-physical waves from spatially limited arrays using source-receiver interferometry (SRI). *Geophysical Journal International*, 196:1043–1059, 2014. DOI: 10.1093/gji/ggt435.



is perturbed, the corresponding interferometric equation can be expressed as the sum of eight terms, which result from the separation of the total wavefield into an unperturbed background field and the perturbed scattered field. In this chapter, the contribution of each individual term is identified for singly diffracted waves using the methods of stationary phase analysis and waveform modelling. When the data acquisition boundary requirements for seismic interferometry are violated, non-physical energy is introduced into Green’s function estimates. Our results show that four terms produce purely non-physical, non-stationary energy and that these can be suppressed, and that a combination of only two terms can be used to estimate diffracted wavefields robustly. One of the two terms is precisely that used in geophysical imaging schemes. A key result is that this term also produces non-physical energy, except when the integration boundaries are truncated to span only part of the medium’s free surface: we thus show that in this sense, partial boundaries can be seen as a positive advantage for migration or imaging methods. The other term produces non-physical energy which nevertheless emulates physical energy; such energy is therefore called pseudo-physical. We present for the first time a complete mathematical derivation of this new category of energy complemented with illustrative examples. Overall, this work significantly enhances our understanding of how scattered wave SRI works.

## 2.1 Introduction

Seismic or wavefield interferometry commonly refers to the use of wavefields from a boundary of sources recorded at two receivers to construct the signal that would have been obtained at one of the two receivers if the other receiver had instead been an impulsive source (Lobkis and Weaver, 2001; Campillo and Paul, 2003; Derode et al., 2003b; Wapenaar, 2004; van Manen et al., 2005, 2006; Wapenaar et al., 2005). This signal is referred to as the Green’s function of the medium and is estimated by cross-correlation, convolution, or deconvolution of the wavefields measured at the two receiver positions. This technique is known as inter-receiver interferometry since it yields the Green’s function between two receivers, turning one receiver into a so-called “virtual” (imagined) source. Further types of interferometry are referred to as either inter-source (Hong and Menke, 2006; Curtis et al., 2009) or source-receiver interferometry, also referred to as SRI (Curtis and Halliday, 2010). These construct either the Green’s function between two sources from the wavefield recorded at an enclosing boundary of receivers, or the Green’s

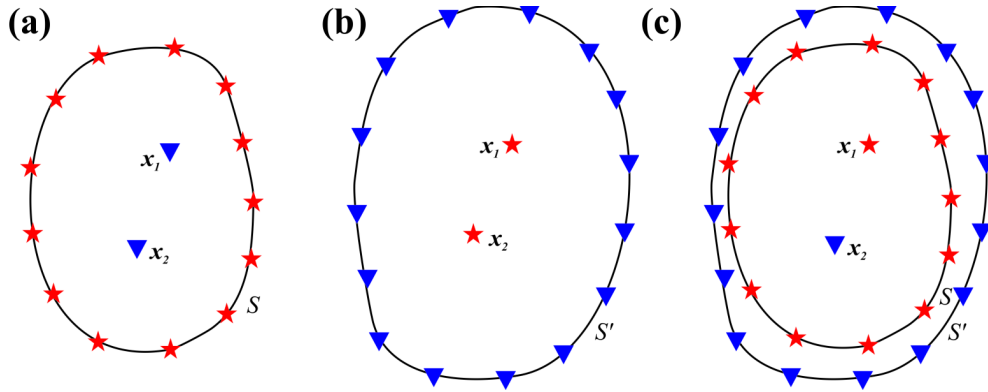


Figure 2.1: Three example canonical configurations used in wavefield interferometry to construct the Green’s function between locations  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . Triangles indicate receivers, stars indicate sources, and  $S$  and  $S'$  denote a source and a receiver boundary, respectively. (a) inter-receiver interferometry: the receiver at either  $\mathbf{x}_1$  or  $\mathbf{x}_2$  is turned into a virtual source using the wavefields generated by the surrounding sources on  $S$ . (b) inter-source interferometry: the source at either  $\mathbf{x}_1$  or  $\mathbf{x}_2$  is turned into a virtual receiver using the wavefield response recorded on the receiver boundary  $S'$ . (c) Source-receiver interferometry (SRI): the Green’s function between a real source and a real receiver can be constructed using the energy travelling from and to the surrounding boundaries  $S$  and  $S'$  of sources and receivers, respectively.

function between a source and a receiver using only the energy that has propagated to surrounding receivers or from surrounding sources. Example canonical geometries for these three types of interferometry are shown in Fig. 2.1, and reviews and tutorials on the various methods are given in Curtis et al. (2006), Wapenaar et al. (2010a,c) and Galetti and Curtis (2012).

Recently, Halliday and Curtis (2010) derived an explicit link between scattered-wave SRI and seismic imaging or migration, a common technique used in exploration geophysics to map reflecting and diffracting discontinuities in the subsurface (Claerbout, 1985; Oristaglio, 1989). As has been shown by Halliday and Curtis (2010), Vasconcelos et al. (2010), and Ravasi and Curtis (2013), seismic interferometry has the potential to improve current migration schemes since it can circumvent the need for the single-scattering Born approximation and is in principle able to account for all possible nonlinearities, such as those due to multiply scattered waves. Other potential applications of SRI are ground-roll removal (Duguid et al., 2011) or reflection imaging from below or above the reflector (Poliannikov, 2011; Poliannikov et al., 2012).

Scattering occurs in all regimes of energy propagation (acoustic, elastic, electromagnetic, etc.) when the propagating wavefield interacts with perturbations inside the medium. In general, one distinguishes between reflected energy which

originates from structures such as interstrata interfaces, and diffracted energy which originates from structures of small spatial extent compared to the wavelength, such as angular boundaries, voids, faults, or fractures. Although standard industrial seismic imaging procedures were originally developed based on diffracted energy (Miller et al., 1987), they are principally designed to image reflecting interfaces rather than diffracting structures. However, understanding and imaging diffractions is a topic of ongoing research (Khaidukov et al., 2004; Berkovitch et al., 2009; Faccipieri et al., 2013), as it helps to interpret recorded data and to enhance the resolution of seismic images. In this study, we use SRI to construct the diffracted wavefield associated with a single isotropic point diffractor, also referred to here as scatterer, in an otherwise homogeneous medium. The simplicity of this medium allows us to illuminate the internal workings of SRI for diffracted energy.

The principles of seismic interferometry can be illustrated using the method of stationary phase (see appendix A). Although mainly applicable for relatively simple media it provides a means to understand the underlying physics and the generation of so-called spurious or non-physical energy in estimated Green's functions (Snieder et al., 2006, 2008; Halliday and Curtis, 2009; Mikesell et al., 2009; King and Curtis, 2012). Using stationary phase analysis, Snieder (2004a,b) showed that the main contribution to the constructed Green's function in coda-wave interferometry comes from so-called stationary points (appendix A). Waves radiated from sources in regions near these points interfere constructively, whereas waves coming from sources in non-stationary regions destructively cancel each other when summing or integrating over a complete source boundary.

In theory, interferometry requires complete, closed boundaries of sources or receivers that surround a portion of the medium of interest. In practice, however, this can seldom be realized (for example, it is usually impossible to place a complete boundary of sources in the interior of solid bodies) so often only partial boundaries, usually spatially limited arrays on the Earth's surface, are available. In addition the spatial sampling density of sources or receivers along the surface may not fulfil the usual Nyquist requirements of wavefield sampling. In such cases, spurious or non-physical energy appears in the interferometric results giving errors in the Green's function estimates due, for example, to incomplete destructive interference of energy from non-stationary regions or to the omission of sources or receivers at or around stationary points. These non-physical events are not per se unfavourable: it has been shown (Mikesell et al., 2009; King and

Curtis, 2011, 2012; King et al., 2011; Harmankaya et al., 2013; Meles and Curtis, 2013) that non-physical energy can in fact be used to extract physical information about the medium. However, our ability to use such energy originates from our understanding of how it relates to physical properties and recording geometries.

In this work we contribute to a deeper understanding of SRI in a scattering medium. We expand the kinematic analysis of Meles and Curtis (2013) by considering also dynamic waveforms of events constructed from SRI, and analyse the effect of limited integration boundaries represented by linear source and receiver arrays on one side of the medium. We examine the origin of non-physical diffracted energy by invoking the method of stationary phase, and focus on a new category of non-physical energy which emulates physical energy and is therefore referred to as *pseudo-physical*. We present a new mathematical derivation that explains the origin of pseudo-physical energy, and use a numerical model to demonstrate the construction of physical, non-physical, and pseudo-physical energy in the Green's function estimates. In appendix C, we provide a detailed parametric study of each term of the interferometric equation in SRI, illustrating the sensitivity of different terms to changes in the model parameters and in the data processing, and showing how this can be used to suppress undesired non-physical and non-stationary arrivals. As a result, this paper shows how non-physical energy can be suppressed, how pseudo-physical energy is related to physical energy, and how all these types of energy can be used to interrogate the interior of a solid medium such as the Earth.

## 2.2 Source-receiver interferometry in a scattering medium

SRI constructs the signal between a source and a receiver using the energy travelling from and to surrounding boundaries of sources and receivers. It can be thought of as a combination of inter-receiver and inter-source interferometry, performed sequentially. In the following we focus on the specific geometry shown in 2.2, where the outer boundary,  $S$ , is the source boundary and the inner boundary,  $S'$ , is a receiver boundary (these may be interchanged without loss of generality). The first step of SRI comprises the construction of the so-called homogeneous Green's function  $G_h(\mathbf{x}', \mathbf{x}_2) = G(\mathbf{x}', \mathbf{x}_2) + G^*(\mathbf{x}', \mathbf{x}_2)$  between the receiver at  $\mathbf{x}_2$

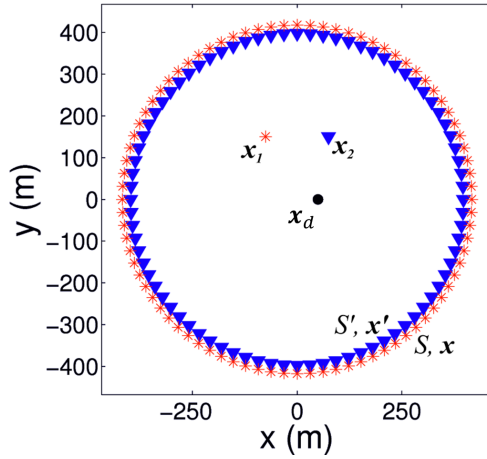


Figure 2.2: Geometry for SRI with complete boundaries. Stars are sources, triangles are receivers, and the black dot marks the location of an isotropic point scatterer.  $\mathbf{x}$  denotes any source on source boundary  $S$ ,  $\mathbf{x}'$  denotes any receiver on receiver boundary  $S'$ .

and any receiver at  $\mathbf{x}'$  on boundary  $S'$  using

$$G(\mathbf{x}', \mathbf{x}_2) + G^*(\mathbf{x}', \mathbf{x}_2) = \frac{-1}{i\omega\rho} \int_{S'} \{G(\mathbf{x}_2, \mathbf{x})\partial_i G^*(\mathbf{x}', \mathbf{x}) - [\partial_i G(\mathbf{x}_2, \mathbf{x})]G^*(\mathbf{x}', \mathbf{x})\} n_i d^2\mathbf{x} \quad (2.1)$$

(Wapenaar and Fokkema, 2006) where  $i = \sqrt{-1}$  is the imaginary unit,  $\omega$  denotes the angular frequency,  $\rho$  denotes the density of the medium (assumed to be constant herein),  $G(\mathbf{x}', \mathbf{x})$  is the Green's function between a source at  $\mathbf{x}$  and a receiver at  $\mathbf{x}'$ ,  $n_i \partial_i G$  is a Green's function's derivative in direction  $i$ , the star  $*$  denotes complex conjugation, and the integration is over variable  $\mathbf{x}$ . The Einstein summation convention applies to repeated indices. Note that the explicit dependency on frequency in the Green's functions has been dropped for notational convenience only, and all expressions herein are in the frequency domain. For each receiver pair this is equivalent to standard inter-receiver interferometry: the wavefields from each source on  $\mathbf{x}$  recorded at one receiver pair at  $\mathbf{x}'$  and  $\mathbf{x}_2$  are cross-correlated and summed (integrated) over source positions. This turns one of the receivers, here the central receiver at  $\mathbf{x}_2$ , into a virtual source, as can be seen in terms  $G(\mathbf{x}', \mathbf{x}_2)$  on the left-hand side. In the second step we construct the homogeneous Green's function  $G_h(\mathbf{x}_2, \mathbf{x}_1) = G(\mathbf{x}_2, \mathbf{x}_1) + G^*(\mathbf{x}_2, \mathbf{x}_1)$  between the real source at  $\mathbf{x}_1$  and the virtual source at  $\mathbf{x}_2$ , using the recorded wavefields between  $\mathbf{x}_1$  and points on  $S'$ , and the inter-receiver wavefields between  $\mathbf{x}_2$  and points on  $S'$  obtained in the first step. This corresponds to inter-source

interferometry and is given by

$$G(\mathbf{x}_2, \mathbf{x}_1) + G^*(\mathbf{x}_2, \mathbf{x}_1) = \frac{-1}{i\omega\rho} \int_{S'} \{G^*(\mathbf{x}', \mathbf{x}_2)\partial'_i G(\mathbf{x}', \mathbf{x}_1) - [\partial'_i G^*(\mathbf{x}', \mathbf{x}_2)]G(\mathbf{x}', \mathbf{x}_1)\} n'_i d^2\mathbf{x}' \quad (2.2)$$

(Hong and Menke, 2006). The Green's function  $G^*(\mathbf{x}', \mathbf{x}_2)$  required in 2.2 can be obtained from the homogeneous Green's function  $G_h(\mathbf{x}', \mathbf{x}_2)$  (Eq. 2.1) by windowing the acausal part of the time-domain signal then transforming back to the frequency domain. Assuming that the Sommerfeld radiation conditions (Born and Wolf, 1999) apply, we can write the Green's functions' derivatives in Eqs. 2.1 and 2.2 as  $n_i \partial_i G = \pm i(\frac{\omega}{c}G)$ , where  $c$  is velocity and  $'-'$  and  $'+'$  indicate outgoing or incoming waves, respectively (Wapenaar and Fokkema, 2006). This reduces the integrals to the simpler forms

$$G(\mathbf{x}', \mathbf{x}_2) + G^*(\mathbf{x}', \mathbf{x}_2) \approx \frac{2}{\rho c} \int_S G(\mathbf{x}_2, \mathbf{x}) G^*(\mathbf{x}', \mathbf{x}) d^2\mathbf{x} \quad (2.3)$$

$$G(\mathbf{x}_2, \mathbf{x}_1) + G^*(\mathbf{x}_2, \mathbf{x}_1) \approx \frac{2}{\rho c} \int_{S'} G(\mathbf{x}', \mathbf{x}_2) G^*(\mathbf{x}', \mathbf{x}_1) d^2\mathbf{x}' \quad (2.4)$$

(Curtis et al., 2012). We will refer to Eqs. 2.3 and 2.4 as the 'monopole approximation' because the dipole sources and receivers indicated by the Green's functions' derivatives in Eqs. 2.1 and 2.2, are approximated by monopole sources and receivers in Eqs. 2.3 and 2.4. Sommerfeld's radiation conditions assume that all ray paths are normal to the boundaries. Eqs. 2.3 and 2.4 can be combined to a more applicable form,

$$\begin{aligned} G(\mathbf{x}_2, \mathbf{x}_1) + G^*(\mathbf{x}_2, \mathbf{x}_1) & \\ & \approx \frac{4}{(\rho c)^2} \int_S \int_{S'} G(\mathbf{x}', \mathbf{x}_1) G(\mathbf{x}_2, \mathbf{x}) G^*(\mathbf{x}', \mathbf{x}) dS' d^2\mathbf{x} \\ & \quad - \frac{2}{\rho c} \int_{S'} G(\mathbf{x}', \mathbf{x}_1) G(\mathbf{x}', \mathbf{x} - 2) d^2\mathbf{x}' \end{aligned} \quad (2.5a)$$

$$\approx \frac{4}{(\rho c)^2} \int_S \int_{S'} G(\mathbf{x}', \mathbf{x}_1) G(\mathbf{x}_2, \mathbf{x}) G^*(\mathbf{x}', \mathbf{x}) d^2\mathbf{x}' d^2\mathbf{x} \quad (2.5b)$$

(Curtis and Halliday, 2010), where the second integral in Eq. 2.5a goes to zero if the radiation condition applies, that is, if the boundaries are in the far-field and are perpendicular to the outgoing wavefield. This expression allows a clear understanding of the problem that is presented in the following. As shown in Fig. 2.2 we assume a scattering medium and herein we consider only the case of a single scatterer or diffractor. In this case it is useful to separate the full

wavefield  $G$  into the unperturbed background field  $G_0$ , which would be obtained if the scatterer was not present, and the perturbed or scattered field  $G_S$  defined according to

$$G = G_0 + G_S \quad (2.6)$$

where  $G$  is any Green's function measured in the perturbed medium. Substituting Eq. 2.6 for each Green's function in Eq. 2.5b gives a double-surface integral over the sum of eight terms:

$$\begin{aligned} G(\mathbf{x}_2, \mathbf{x}_1) + G^*(\mathbf{x}_2, \mathbf{x}_1) &= [G_0(\mathbf{x}_2, \mathbf{x}_1) + G_S(\mathbf{x}_2, \mathbf{x}_1)] \\ &\quad + [G_0(\mathbf{x}_2, \mathbf{x}_1) + G_S(\mathbf{x}_2, \mathbf{x}_1)]^* \\ &\approx \frac{4}{(\rho c)^2} \int_S \int_{S'} [G_0(\mathbf{x}', \mathbf{x}_1) G_0(\mathbf{x}_2, \mathbf{x}) G_S^*(\mathbf{x}', \mathbf{x}) \\ &\quad + G_S(\mathbf{x}', \mathbf{x}_1) G_S(\mathbf{x}_2, \mathbf{x}) G_S^*(\mathbf{x}', \mathbf{x}) \\ &\quad + G_0(\mathbf{x}', \mathbf{x}_1) G_S(\mathbf{x}_2, \mathbf{x}) G_S^*(\mathbf{x}', \mathbf{x}) \\ &\quad + G_S(\mathbf{x}', \mathbf{x}_1) G_0(\mathbf{x}_2, \mathbf{x}) G_S^*(\mathbf{x}', \mathbf{x}) \\ &\quad + G_0(\mathbf{x}', \mathbf{x}_1) G_0(\mathbf{x}_2, \mathbf{x}) G_0^*(\mathbf{x}', \mathbf{x}) \\ &\quad + G_S(\mathbf{x}', \mathbf{x}_1) G_S(\mathbf{x}_2, \mathbf{x}) G_0^*(\mathbf{x}', \mathbf{x}) \\ &\quad + G_0(\mathbf{x}', \mathbf{x}_1) G_S(\mathbf{x}_2, \mathbf{x}) G_0^*(\mathbf{x}', \mathbf{x}) \\ &\quad + G_S(\mathbf{x}', \mathbf{x}_1) G_0(\mathbf{x}_2, \mathbf{x}) G_0^*(\mathbf{x}', \mathbf{x})] d^2 \mathbf{x}' d^2 \mathbf{x} \end{aligned} \quad (2.7)$$

where subscript 0 refers to a direct wave and subscript  $S$  refers to a scattered wave (Vasconcelos et al., 2009). A similar decomposition is presented by Vasconcelos (2013), which includes all of the gradient terms at both source and receiver locations similarly to Eqs. 2.1 and 2.2, rather than invoking the monopole approximation in Eqs. 2.3 and 2.4 as above.

The double integral on the right-hand side of Eq. 2.7 can be evaluated using stationary phase analysis. This method has been used in standard interferometry (Snieder et al., 2006; Snieder et al., 2008; Halliday and Curtis, 2009; Snieder and Fleury, 2010) to analyse interferometric integrals assuming that the main contribution to the integrand comes from so-called stationary points (appendix A). Recently, Meles and Curtis (2013) have performed stationary phase analysis for the kinematics of SRI. Due to the double boundary the stationary points of inter-receiver interferometry become stationary point-pairs (or just pairs) consisting of a stationary point  $\mathbf{x}$  on source boundary  $S$  and a stationary point  $\mathbf{x}'$  on receiver boundary  $S'$  (Fig. 2.3, Table 2.1). We invoke this method of analysis below and in appendix C.

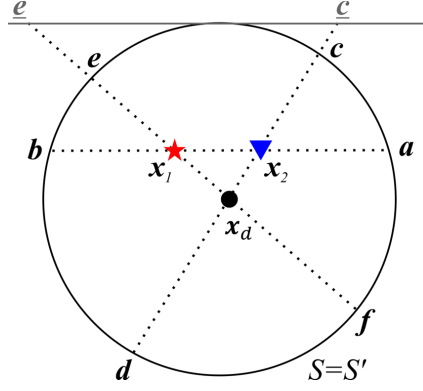


Figure 2.3: Stationary points (**a** to **f**) for source–receiver interferometry in a homogeneous medium containing a single point scatterer. Symbol key as in Fig. 2.2. The source and receiver boundary coincide and are represented by the circular solid line. A stationary pair consists of a boundary source and receiver at a pair of stationary points. The combination of stationary points in this pair varies for each term. Points **e** and **c** are projections of points **e** and **c** on to a horizontal surface (such as the Earth’s surface in seismic interferometry in a vertical plane).

Terms in the integrand of Equation (8)	Corresponding stationary points	
$G_0(\mathbf{x}', \mathbf{x}_1) G_0(\mathbf{x}_2, \mathbf{x}) G_0^*(\mathbf{x}', \mathbf{x})$	$\mathbf{x} = \mathbf{a}; \mathbf{x}' = \mathbf{b};$	$\underline{\mathbf{x} = \mathbf{b}; \mathbf{x}' = \mathbf{a};}$
$G_0(\mathbf{x}', \mathbf{x}_1) G_s(\mathbf{x}_2, \mathbf{x}) G_0^*(\mathbf{x}', \mathbf{x})$	$\mathbf{x} = \mathbf{e}; \mathbf{x}' = \mathbf{f};$	$\mathbf{x} = \mathbf{f}; \mathbf{x}' = \mathbf{e};$
$G_s(\mathbf{x}', \mathbf{x}_1) G_0(\mathbf{x}_2, \mathbf{x}) G_0^*(\mathbf{x}', \mathbf{x})$	$\mathbf{x} = \mathbf{d}; \mathbf{x}' = \mathbf{c};$	$\mathbf{x} = \mathbf{c}; \mathbf{x}' = \mathbf{d};$
$G_s(\mathbf{x}', \mathbf{x}_1) G_s(\mathbf{x}_2, \mathbf{x}) G_0^*(\mathbf{x}', \mathbf{x})$	any pair connected by a straight line passing through the scatterer	
$G_0(\mathbf{x}', \mathbf{x}_1) G_0(\mathbf{x}_2, \mathbf{x}) G_s^*(\mathbf{x}', \mathbf{x})$	$\underline{\mathbf{x} = \mathbf{e}; \mathbf{x}' = \mathbf{c};}$ $\mathbf{x} = \mathbf{f}; \mathbf{x}' = \mathbf{d};$	$\mathbf{x} = \mathbf{d}; \mathbf{x}' = \mathbf{e};$ $\mathbf{x} = \mathbf{c}; \mathbf{x}' = \mathbf{f};$
$G_0(\mathbf{x}', \mathbf{x}_1) G_s(\mathbf{x}_2, \mathbf{x}) G_s^*(\mathbf{x}', \mathbf{x})$	$\mathbf{x} = \mathbf{e}; \mathbf{x}' = \mathbf{f};$	$\mathbf{x} = \mathbf{f}; \mathbf{x}' = \mathbf{e};$
$G_s(\mathbf{x}', \mathbf{x}_1) G_0(\mathbf{x}_2, \mathbf{x}) G_s^*(\mathbf{x}', \mathbf{x})$	$\mathbf{x} = \mathbf{d}; \text{any } \mathbf{x}';$	$\mathbf{x} = \mathbf{c}; \text{any } \mathbf{x}';$
$G_s(\mathbf{x}', \mathbf{x}_1) G_s(\mathbf{x}_2, \mathbf{x}) G_s^*(\mathbf{x}', \mathbf{x})$	any pair	

Table 2.1: Stationary point pairs for different terms in Eq. 2.7 (see Fig. 2.3; Meles and Curtis, 2013).  $\mathbf{x}$  refers to a source and  $\mathbf{x}'$  to a receiver location. The events constructed occur at travel times associated with the causal direct wave (blue), the acausal direct wave (blue, underlined), the causal scattered wave (green), and the acausal scattered wave (green, underlined). Events constructed from stationary points marked red do not relate to physical arrivals.



## 2.3 Non-physical and pseudo-physical energy

An event constructed from interferometry is referred to as “non-physical” if it does not correspond to a physical wave that would propagate between the source and receiver locations  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , such as a direct or scattered wave. Fig. 2.4a illustrates geometrically how one such event is generated as an example. Snieder et al. (2008) analysed the properties of non-physical events in inter-receiver interferometry in a scattering medium. They showed the contributions from different terms (in their case four terms rather than eight) towards physical and non-physical energy, and how non-physical energy is cancelled out after the summation of all terms and integration over a closed boundary. If, however, the different terms of the integrand are used separately, or if the boundary is not complete, the non-physical energy does not cancel out but instead gives spurious contributions to the interferometric estimate.

Geophysical seismic imaging as well requires integration over boundaries. It is well known that artefacts in the image occur due to limited boundaries: sources and receivers can usually only be placed on the Earth’s surface and hence cannot be said to surround any portion of the medium through which the energy propagates and which we hope to image (the subsurface). From interferometry we now understand that these artefacts correspond to non-physical energy that is not cancelled out due to missing sources and receivers in the subsurface.

In SRI, unlike in standard interferometry, we find that some non-physical energy arrives at exactly the travel time of the expected scattered waves, thus approximately (or exactly) emulating physical energy. Fig. 2.4b illustrates how such a pseudo-physical event is constructed from SRI when using diffracted waves  $G_S$  only, that is, within the term  $G_S G_S G_S^*$ . For this term, any source-receiver pair on the boundaries is stationary and is sufficient to construct the event (Meles and Curtis, 2013; Fig. 2.3 and Table 2.1). This property makes the term  $G_S G_S G_S^*$  particularly useful when boundaries are only partially available or are strongly decimated, as is often the case in practical experiments. Note that the stationarity properties of this term only apply to diffracted waves; for reflecting media the behaviour is substantially different. As has been shown by Meles and Curtis (2013), the travel time of the constructed event equals the travel time of the causal scattered wave. They point out, however, that the interferometric event is only *proportional* to the causal scattered wave: its amplitude also depends on

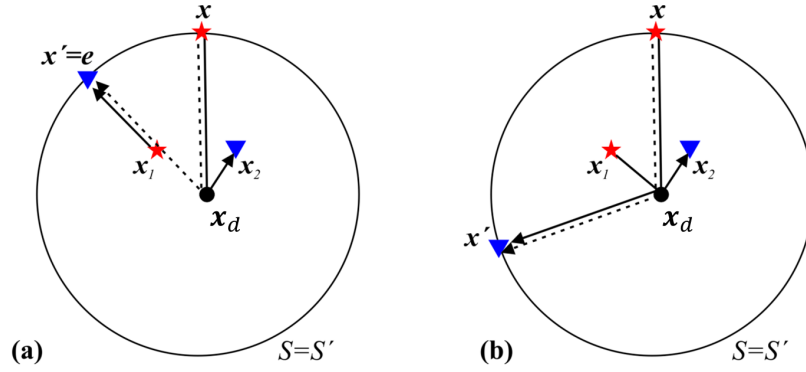


Figure 2.4: Generation of non-physical and pseudo-physical energy from different terms in Eq. 2.7 and Table 2.1; symbol key as in Fig. 2.2. Solid ray paths indicate travel times that are added to the phase of the result of Eq. 2.7, the dashed line indicates a travel time that is subtracted from the phase: hence, portions of solid and dashed lines that span the same path give phase contributions that exactly cancel. **(a)**  $G_0 G_S G_S^*$ : the stationary point  $\mathbf{e}$  on  $S'$  and any point  $\mathbf{x}$  on  $S$  give rise to a stationary, but non-physical event with a travel time equal to  $t(\mathbf{x}_2, \mathbf{x}_S) - t(\mathbf{x}_1, \mathbf{x}_S)$ , where  $t(\mathbf{x}_2, \mathbf{x}_S)$  is the travel time from point  $\mathbf{x}_S$  to  $\mathbf{x}_2$ . **(b)**  $G_S G_S G_S^*$ : any source-receiver pair on the boundary is stationary and gives rise to a non-physical event with a travel time equal to that of the physical diffracted wave. The kinematics of this energy thus emulate that of physical energy, and hence, the event is called *pseudo-physical*

a real factor  $\lambda$ . Expanding their analysis, we provide in the following a detailed mathematical derivation explaining the origin of pseudo-physical energy in SRI and the properties of  $\lambda$  for 2D and 3D Green's functions. The derivation below is complemented by an alternative derivation based on the scattered wave representation theorems of Vasconcelos et al. (2009) given in appendix B.

We start by rewriting the diffracted wavefield as a concatenation of direct waves according to

$$G_S(\mathbf{x}_2, \mathbf{x}_1) = G_0(\mathbf{x}_d, \mathbf{x}_1) A(\mathbf{k}_2, -\mathbf{k}_1) \tilde{G}_0(\mathbf{x}_2, \mathbf{x}_d) \quad (2.8)$$

(Snieder et al., 2008; Wapenaar et al., 2010b) where  $\mathbf{x}_d$  is the location of a point diffractor,  $A(\mathbf{k}_2, -\mathbf{k}_1)$  is the complex-valued scattering matrix,  $-\mathbf{k}_1$  is the direction of the incident wavefield, and  $\mathbf{k}_2$  the direction of the scattered wavefield. In the far field, the Green's function  $G_0(\mathbf{x}_d, \mathbf{x}_1)$  in 2D is defined as

$$G_0(\mathbf{x}_d, \mathbf{x}_1) = -\frac{\rho\omega}{4} e^{-i(k|\mathbf{x}_d - \mathbf{x}_1|)} \sqrt{\frac{2}{\pi k |\mathbf{x}_d - \mathbf{x}_1|}} \quad (2.9)$$

and the volume injection Green's function  $\tilde{G}_0(\mathbf{x}_2, \mathbf{x}_d)$  is given by

$$\begin{aligned}\tilde{G}_0(\mathbf{x}_2, \mathbf{x}_d) &= \frac{1}{\rho\omega} G_0(\mathbf{x}_2, \mathbf{x}_d) \\ &= -\frac{1}{4} e^{-ik|\mathbf{x}_2 - \mathbf{x}_d|} \sqrt{\frac{2}{\pi k|\mathbf{x}_2 - \mathbf{x}_d|}}\end{aligned}\quad (2.10)$$

For isotropic point scatterers the scattering matrix  $A(\mathbf{k}_2, -\mathbf{k}_1)$  in Eq. 2.8 does not depend on the direction of the incident wavefield, and an equal amount of energy is scattered in any direction; hence the scattering matrix can be abbreviated as a scalar  $A$ . Substituting each Green's function in  $G_S G_S G_S^*$  for Eq. 2.8, rearranging the terms, and using Eqs. 2.9 and 2.10 gives

$$\begin{aligned}& \frac{4}{(\rho c)^2} \int_S \int_{S'} G_S(\mathbf{x}', \mathbf{x}_1) G_S(\mathbf{x}_2, \mathbf{x}) G_S^*(\mathbf{x}', \mathbf{x}) dS' dS \\ &= G_0(\mathbf{x}_2, \mathbf{x}_d) A \tilde{G}_0(\mathbf{x}_d, \mathbf{x}_1) \times A A^* \frac{4}{(\rho c)^2} \int_S G_0(\mathbf{x}_d, \mathbf{x}) G_0^*(\mathbf{x}_d, \mathbf{x}) dS \\ &\quad \times \int_{S'} \tilde{G}_0(\mathbf{x}', \mathbf{x}_d) \tilde{G}_0^*(\mathbf{x}', \mathbf{x}_d) dS' \\ &= G_S(\mathbf{x}_2, \mathbf{x}_1) \times A A^* \frac{1}{(4\pi)^2} \int_S \frac{1}{|\mathbf{x}_d - \mathbf{x}|} dS \int_{S'} \frac{1}{|\mathbf{x}' - \mathbf{x}_d|} dS'\end{aligned}\quad (2.11)$$

Without loss of generality we assume that the scatterer is located at the origin  $\mathbf{x}_d = [0, 0]$ , so that the terms in the integrands simplify to  $\frac{1}{|\mathbf{x}|}$  and  $\frac{1}{|\mathbf{x}'|}$ , respectively. In the case of circular boundaries it is convenient to move to a polar coordinate system, such that  $|\mathbf{x}| = r$ ,  $|\mathbf{x}'| = r'$ ,  $dS = r d\phi$  and  $dS' = r' d\phi'$ . Integration over  $\phi$  and  $\phi'$  from 0 to  $2\pi$  shows that each surface integral reduces to a factor  $2\pi$  (note that when the boundaries are only partially available each surface integral will give a fraction of  $2\pi$  depending on the portion of the circle included in the boundaries). From the relationship between real and imaginary parts of the scattering amplitude (optical theorem) it follows that  $A A^* = -4\Im(A)$ , with  $0 \geq \Im(A) \geq -4$ , where  $\Im(A)$  is the imaginary part of  $A$  (Groenenboom and Snieder, 1995; Snieder, 1999; Galetti et al., 2013). Consequently, 2.11 becomes

$$\begin{aligned}& \frac{4}{(\rho c)^2} \int_S \int_{S'} G_S(\mathbf{x}', \mathbf{x}_1) G_S(\mathbf{x}_2, \mathbf{x}) G_S^*(\mathbf{x}', \mathbf{x}) dS' dS \\ &= G_S(\mathbf{x}_2, \mathbf{x}_1) \times [-\Im(A)]\end{aligned}\quad (2.12)$$

Eq. 2.12 shows that the contribution of the term  $G_S G_S G_S^*$  from complete boundaries equals the causal scattered wave  $G_S(\mathbf{x}_2, \mathbf{x}_1)$  multiplied by a real-valued positive constant that is proportional to the imaginary part of the scattering matrix.

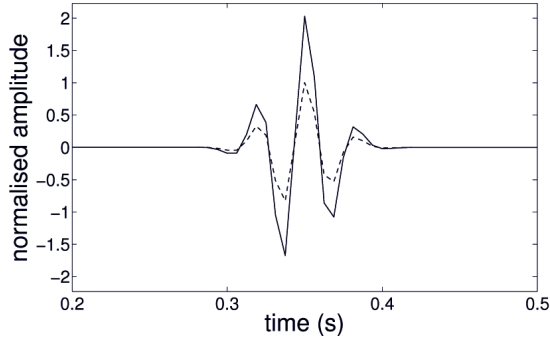


Figure 2.5: Interferometric result of the cross-term  $G_S G_S G_S^*$  in Eq. 2.7 (solid line) compared to the modelled causal scattered wave  $G_S(\mathbf{x}_2, \mathbf{x}_1)$  (dashed line) using 2D Green’s functions and complete circular boundaries. Amplitudes have been normalized with respect to the maximum of the modelled trace. The imaginary part of the scattering matrix  $A$  has been set to  $-2$ ; hence, according to Eq. 2.12, the interferometric result equals  $2 \cdot G_S(\mathbf{x}_2, \mathbf{x}_1)$ .

This causes an amplitude change but no shift in phase or travel time compared to the modelled arrival  $G_S(\mathbf{x}_2, \mathbf{x}_1)$ . For partial boundaries the overall amplitude is reduced according to the portion of the circle included in the boundaries. Thus we show that although the term  $G_S G_S G_S^*$  is non-physical, in 2D it provides the correct travel time and waveform of the causal scattered wave, hence, it is *pseudo-physical*. If the scatterer is non-isotropic the scattering matrix  $A(\mathbf{k}_2, -\mathbf{k}_1)$  cannot be reduced to the scalar  $A$  and the analysis is more complicated. The amplitude of the pseudo-physical arrival will then be a function of the source and receiver positions relative to the scatterer. Nevertheless, the kinematic analysis applies just as well for non-isotropic scatterers, which allows us to estimate the travel time of the causal scattered wave.

The above results are only valid for the 2D case where the 2D Green’s functions defined in Eqs. 2.9 and 2.10 are used. Using 3D Green’s functions and the corresponding relationship  $AA^* = -\frac{1}{k}\Im(A)$  (Wapenaar et al., 2010b) Eq. 2.12 changes to

$$\begin{aligned} & \frac{4}{(\rho c)^2} \int_S \int_{S'} G_S(\mathbf{x}', \mathbf{x}_1) G_S(\mathbf{x}_2, \mathbf{x}) G_S^*(\mathbf{x}', \mathbf{x}) dS' dS \\ & = G_S(\mathbf{x}_2, \mathbf{x}_1) \times (-4k) [-\Im(A)] \end{aligned} \quad (2.13)$$

Since  $k = \frac{\omega}{c}$  the result now depends on the frequency content and therefore distorts the waveform (Fig 2.6a). However, this effect can be removed by dividing the result by  $k$  (in the frequency domain), and thus the correct travel time and waveform information can also be obtained from the contribution of  $G_S G_S G_S^*$  in the 3D case (Fig. 2.6b).

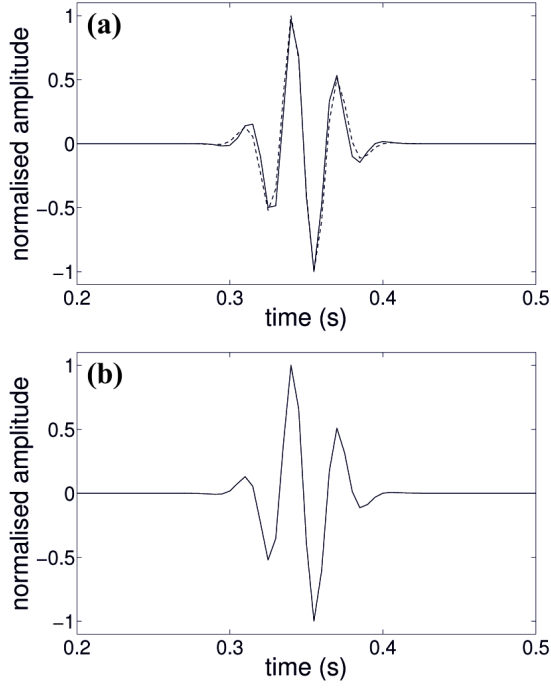


Figure 2.6: **(a)** Interferometric result of the term  $G_S G_S G_S^*$  in Eq. 2.7 (solid line) compared to the modelled causal scattered wave  $G_S(\mathbf{x}_2, \mathbf{x}_1)$  (dashed line) using 3D Green's functions and the geometry shown in Fig 2.7. Amplitudes have been normalized with respect to the maximum of each trace. **(b)** As in (a) but divided by the wavenumber  $k$  in the frequency domain (Eq. 2.13): the waveforms of the two curves are now identical.

Note that the non-physical energy provided by  $G_S G_S G_S^*$  would be destructively cancelled out within an integration over complete boundaries that included the summation over all terms in Eq. 2.7. Destructive cancellation occurs on account of other terms that provide non-physical energy at the same travel time but with different amplitude and phase. Following similar arguments as for  $G_S G_S G_S^*$ , it can be shown that this cancelling energy is provided by the terms  $G_0 G_0 G_S^*$ ,  $G_0 G_S G_0^*$ ,  $G_0 G_S G_S^*$ ,  $G_S G_0 G_S^*$ , and  $G_S G_S G_0^*$  (cf. Table 2.1). Thus we can also show that only the term  $G_S G_0 G_0^*$  constructs the physical causal scattered wave (see Eqs. B.1 and B.2 in appendix B), given that the integration boundaries span the stationary point pair associated with that term ( $\mathbf{x} = \mathbf{d}, \mathbf{x}' = \mathbf{c}$  in Fig. 2.3).

## 2.4 Numerical examples

In the numerical examples that follow, the integration boundaries  $S$  and  $S'$  are reduced to finite linear arrays above the scatterer, as illustrated in Fig. 2.7. Compared to the ideal geometry in Figs. 2.2 and 2.3, this omits some of the

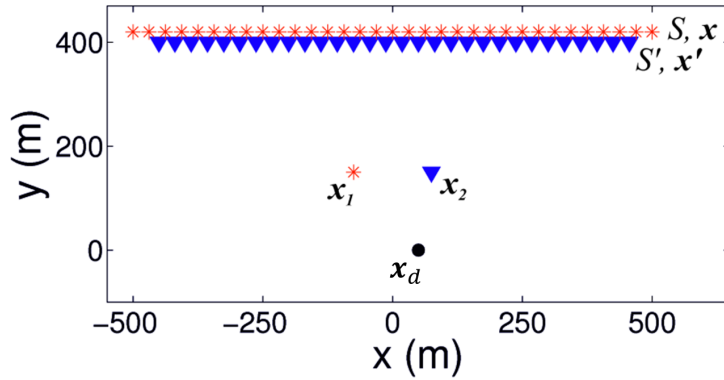


Figure 2.7: Geometry used for numerical examples with incomplete boundaries represented by linear source and receiver arrays; symbol key as in Fig. 2.2. Only every fifth source and receiver is plotted for clarity.

stationary points: in fact, only two stationary points,  $\underline{e}$  and  $\underline{c}$ , equivalent to  $\mathbf{e}$  and  $\mathbf{c}$  (Table 2.1, Fig. 2.3), are populated by sources and receivers. To model acoustic wavefields in a scattering medium we use a direct scattering matrix-based scheme that is a variant of Foldy’s method (Foldy, 1945; Groenenboom and Snieder, 1995; Galetti et al., 2013). This method yields the full, non-linear scattering response of multiple isotropic point scatterers embedded in an otherwise homogeneous medium. In the modelling code of Galetti et al. (2013), inter-receiver and inter-source interferometry are performed sequentially using Eq. 2.1 and 2.2 for the full wavefield (monopoles and dipoles), or Eq. 2.3 and 2.4 for the monopole approximation. The background velocity and density of the model are here taken to be  $v = 1000$  m/s and  $\rho = 1000$  kg/m<sup>3</sup>, respectively. The scatterer at  $\mathbf{x}_d = [50 \text{ m}, 0 \text{ m}]$  is a point diffractor, the imaginary part of the scattering amplitude is chosen to be  $-2$  in accordance with the conditions of the acoustic optical theorem (Groenenboom and Snieder, 1995). The maximum frequency is  $f_{max} = 80$  Hz and the central frequency of the applied Ricker wavelet is  $f_c = 30$  Hz. The spatial sampling, that is, the inter-source and inter-receiver distance within the arrays, is controlled by the Nyquist wavelength  $\lambda_{Nyq}$  and is given by multiples of this value.  $\lambda_{Nyq}$  describes half the minimum wavelength defined by the velocity  $v$  and the temporal Nyquist frequency  $f_{Nyq}$  as

$$\lambda_{Nyq} = \frac{v}{f_{Nyq}} = \frac{v}{2f_{max}} = 6.25 \text{ m} \quad (2.14)$$

given the signal’s maximum frequency  $f_{max}$ . The length of the source array is set to 1000 m and the receiver array is 900 m long, which gives a maximum of 161 sources and 144 receivers, respectively. The receiver array is located 20 m below the source array and 400 m above the diffractor. The single source is located at

$\mathbf{x}_1 = [-75 \text{ m}, 150 \text{ m}]$  and the receiver at  $\mathbf{x}_2 = [75 \text{ m}, 150 \text{ m}]$ . The interferometric results are studied in the time domain in a window between  $-0.8 \text{ s}$  and  $0.8 \text{ s}$ . Negative times are referred to as the acausal part (in the frequency domain, the complex conjugate) of the Green's function and positive times represent the causal part.

In the examples shown in Figs. 2.8 and 2.9 a spatial tapering function has been applied to the cross-correlated traces prior to the summation over sources and receivers. This means that the contributions associated with sources or receivers towards the endpoints of the arrays have been down weighted using half-cosine windowing functions.

The solid trace in Fig. 2.8 gives the interferometric estimate of the Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  using the geometry in Fig. 2.7 and the full wavefield (i.e. all of Eq. 2.7, but with incomplete integration boundaries) and the dashed trace represents the true Green's function modelled directly between  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . Table 2.2 provides the key to symbols used to denote parameter constellations employed in Fig. 2.8 and other figures. Since amplitudes of the constructed trace are expected to be incorrect due to the limited number of sources and receivers along the boundaries, the maximum amplitude of each trace has been normalized to one. Note that the normalization does not change the phase or the waveform shapes and therefore does not affect our analysis. Despite the incomplete boundaries, it appears that both the causal and the acausal scattered wave are constructed surprisingly well from SRI. What is not apparent, however, is that the arrival that looks like the causal scattered wave is in fact a non-physical arrival, which has a physical travel time, hence, is a pseudo-physical arrival. Moreover, note that non-physical events appear with differing amplitudes between zero time and the scattered wave arrivals for both positive and negative times. The direct wave is not recovered at all.

Fig. 2.9 displays the contribution of each individual term (solid lines) of Eq. 2.7, compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  (dashed line). The acausal scattered wave is solely constructed by the term  $G_0 G_0 G_S^*$  (Fig. 2.9a) and the term  $G_S G_S G_S^*$  contributes the above mentioned pseudo-physical arrival at the travel time of the causal scattered wave (Fig. 2.9b). As has been demonstrated in the previous section, the pseudo-physical energy associated with the term  $G_S G_S G_S^*$  can be used to estimate the waveforms of arriving physical energy.

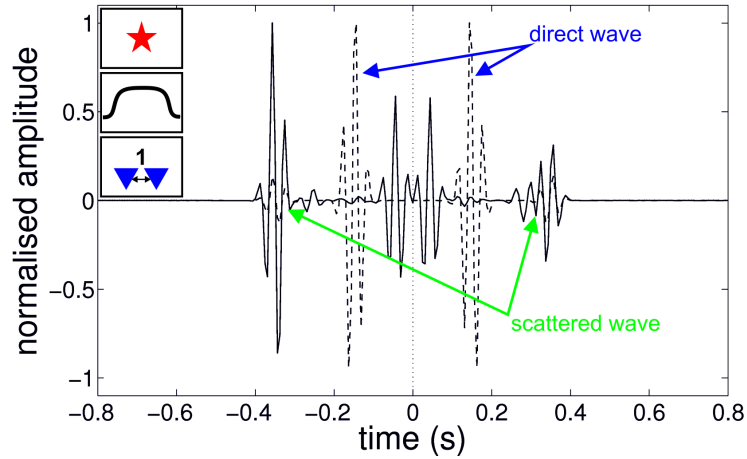


Figure 2.8: Source-receiver interferometric estimate of the Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  using the full wavefield and the incomplete boundaries in Fig. 2.7 (solid line), compared to the true Green's function (dashed line). Amplitudes of each waveform have normalized maximum values. For legend key see Table 2.2.





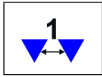
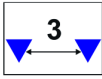
Symbol	Legend
	only monopole sources/receivers are used on the boundary
	monopole and dipole sources/receivers are used on the boundary
	a spatial half-cosine tapering weight is applied to the integrand towards the ends of the boundaries
	no spatial half-cosine tapering weight is applied
	the distance between two sources/receivers on a boundary equals $1 * \lambda_{Nyq}$
	the distance between two sources/receivers on a boundary equals $3 * \lambda_{Nyq}$

Table 2.2: Key to symbols used to denote parameter constellations employed in each numerical example.  $\lambda_{Nyq}$  indicates the spatial wavelength of the Nyquist frequency of the wavefield.



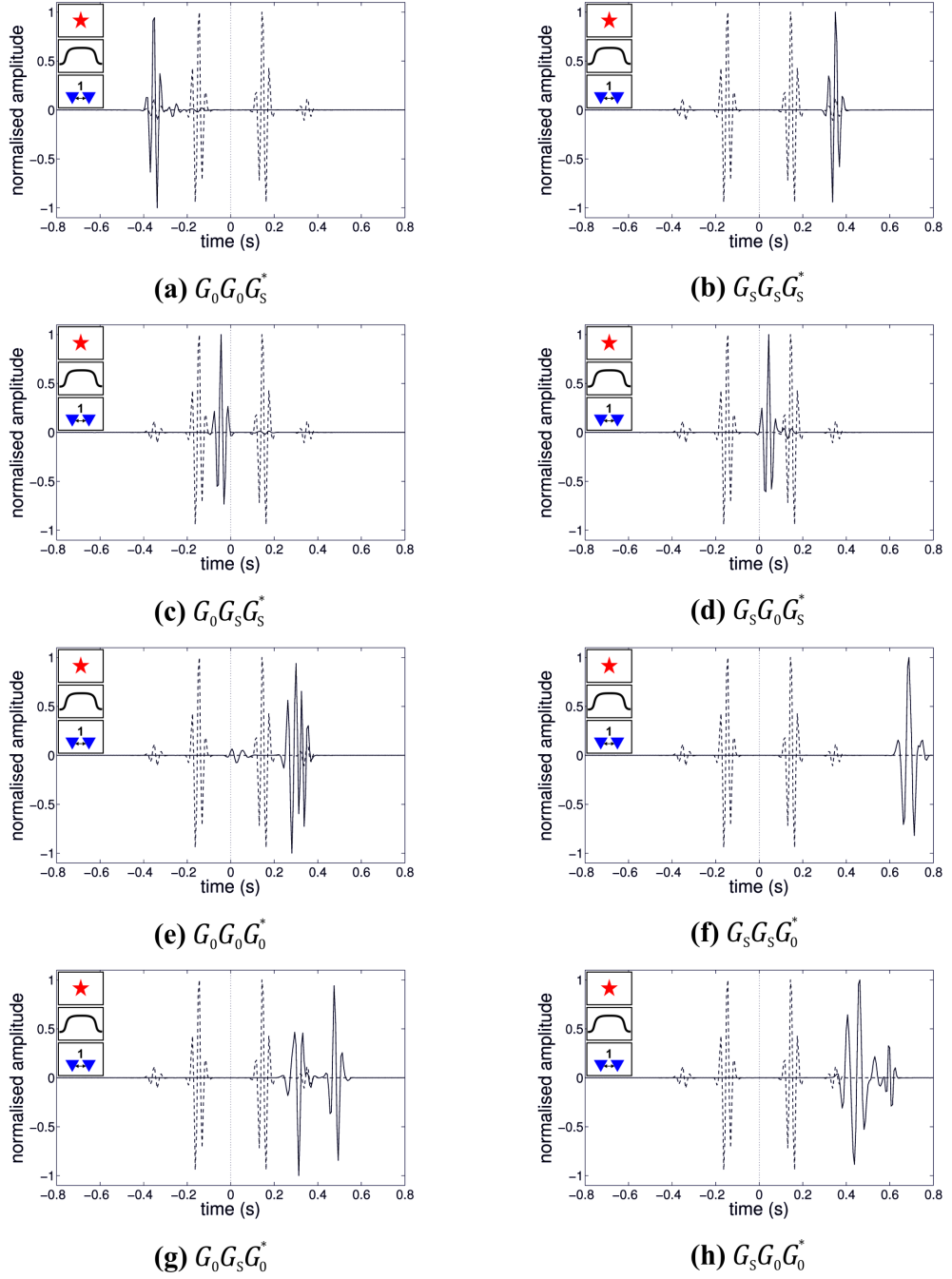


Figure 2.9: Interferometric result of integrating each term in Eq. 2.7 separately (solid lines) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  (dashed line). The term used in each case is noted beneath the plot. Model and processing parameters are defined according to the key symbols (Table 2.2). All maximum amplitudes have been normalized to one.

All other terms generate events that cannot be associated with the expected Green’s function and therefore count as non-physical events. The maximum amplitude in each trace has again been normalized to one; this means that the spurious events from the endpoints of the boundaries have been magnified in Fig. 2.9e–h due to the normalization since they are the largest events on the trace. In fact, they have very low amplitudes—for example, the event in Fig. 2.9f does not show up at all in Fig. 2.8 because its amplitude is too small to see compared to the maximum amplitude in the full trace. In appendix C we show systematically how different events are effected by variations in the model parameters, and by variations in the data processing.

To determine the origin of the constructed signals it is useful to display the so-called correlation gathers (van Manen et al., 2005; Mehta et al., 2008). In standard inter-receiver interferometry the correlation gather is simply the set of integrands that are integrated in the interferometric equation. It provides the contribution of each source on the boundary to the interferometric estimate between two receivers, prior to the summation over sources. Zero-slope areas (i.e. flat areas) in the correlation gather indicate stationary points: the stationary phase approach assumes that the contributions from the Fresnel zone around such points sum constructively, while the contributions from all other source locations cancel each other out. In inter-source interferometry the correlation gather displays the contribution for each specific receiver location. In source–receiver interferometry, however, we must consider both one correlation gather for each receiver pair in the first step (inter-receiver interferometry), and the correlation gather of the (virtual) source pair in the second step (inter-source interferometry).

In Fig. 2.10 we show the correlation gather of one specific receiver pair, namely the receiver at  $\mathbf{x}_2$  and the leftmost receiver on boundary  $S'$  (location  $\mathbf{x}'_l$ ), and the resulting correlation gather of the sources (one virtual, one real) located at  $\mathbf{x}_1$  and  $\mathbf{x}_2$  for each individual term. Some of the inter-source correlation gathers exhibit zero traces (“gaps”) over a range of receiver locations (e.g. Fig. 2.10j). These gaps occur when the inter-receiver energy constructed in the first step (Eq. 2.3) has positive arrival times only. According to Eq. 2.4 only the acausal component  $G^*(\mathbf{x}', \mathbf{x}_2)$ , which corresponds to negative travel times, should be used in the second step. If this is zero, the cross-correlation  $G^*(\mathbf{x}', \mathbf{x}_2)G(\mathbf{x}', \mathbf{x}_1)$  (Eq. 2.4) yields a zero trace and thus a gap in the correlation gather.

The correlation gathers reveal that both physical and non-physical events constructed from the terms  $G_0G_0G_S^*$ ,  $G_SG_SG_S^*$ ,  $G_0G_SG_S^*$ , and  $G_SG_0G_S^*$  (Figs. 2.9a–d)

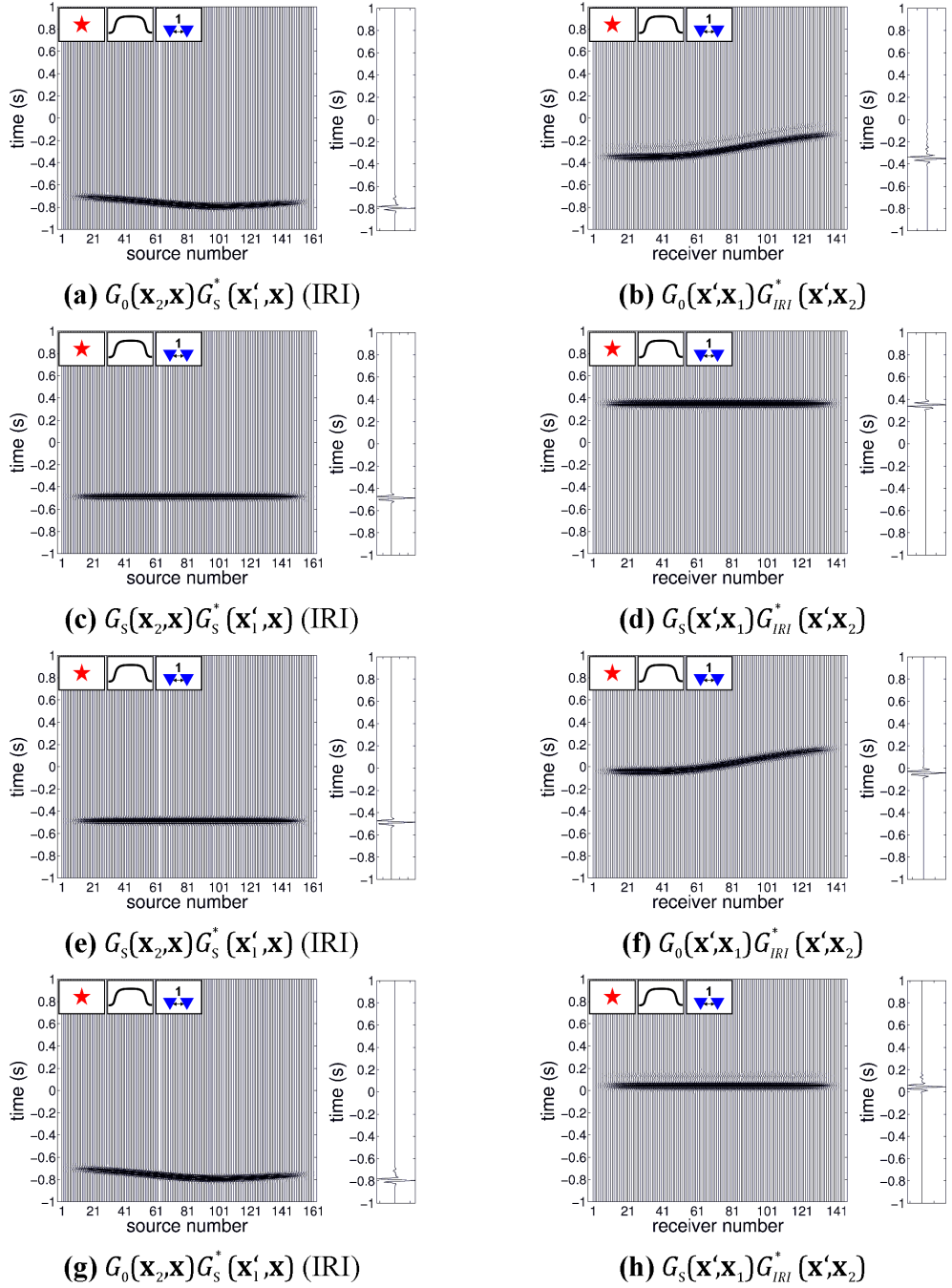


Figure 2.10: Correlation gathers for the first four terms of Eq. 2.7 (continues on next page). The left-hand column displays the inter-receiver interferometry results between the receiver at  $\mathbf{x}_2$  and the leftmost receiver  $\mathbf{x}'_l$  on boundary  $S'$  for each source on boundary  $S$ ; the right-hand plot in the left-hand column gives the sum over all sources. The acausal part of this trace  $G_{IRI}^*(\mathbf{x}'_l, \mathbf{x}_2)$  is then cross-correlated with  $G_S(\mathbf{x}'_l, \mathbf{x}_1)$ , which gives the leftmost trace in the gather in the right-hand column. The right-hand column displays the inter-source interferometry results between the source at  $\mathbf{x}_1$  and the virtual source at  $\mathbf{x}_2$  for each receiver on boundary  $S'$ ; the right-hand plot in the right-hand column gives the sum over all receivers showing how results in Fig. 2.9 are constructed.  $G_{IRI}^*(\mathbf{x}'_l, \mathbf{x}_2)$  refers to the result of inter-receiver interferometry (IRI) carried out in the first step.

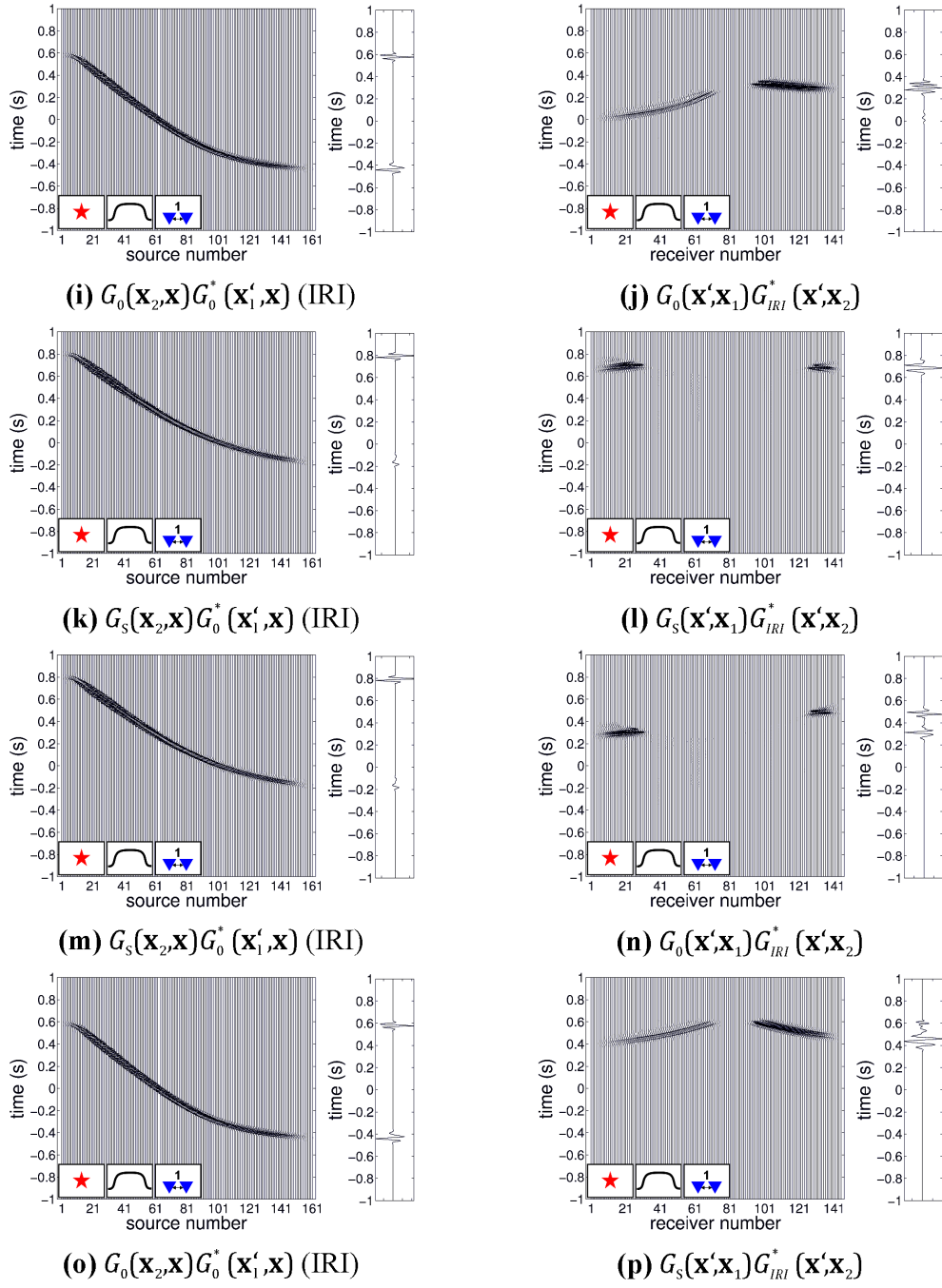


Figure 2.10: Correlation gathers for the last four terms of Eq. 2.7 (continued from previous page).

originate from stationary points. Non-stationary contributions from the endpoints of the arrays are down-weighted by a cosine taper. The events constructed from  $G_0G_0G_0^*$ ,  $G_SG_SG_0^*S$ ,  $G_0G_SG_0^*$ , and  $G_SG_0G_0^*$  (Figs. 2.9e-h) are non-physical and non-stationary without exception as they all originate from the endpoints of the arrays. As before, the traces related to the endpoint sources and receivers, respectively, are down weighted by a taper, however here the summed traces have been normalized to one, which especially magnifies these non-physical, non-stationary events.

## 2.5 Discussion

In interferometry, using incomplete boundaries of sources and receivers, such as linear arrays, causes non-physical arrivals in Green’s function estimates (Fig. 2.8) due to both inadequate sampling of stationary points and abrupt truncation of the boundaries. In appendix C, we analyse the origin of physical, non-physical and pseudo-physical energy from each term in Eq. 2.7 in detail. In this discussion we draw together the principal findings from above and from appendix C.

We first distinguish between stationary and non-stationary non-physical events. Non-stationary events are associated with the contributions from sources and receivers at the endpoints of the arrays. They occur in every term, except for  $G_SG_SG_S^*$  where every source–receiver pair is stationary and gives a pseudo-physical contribution (appendix B; also Meles and Curtis, 2013). As has been demonstrated in previous papers (e.g. Snieder et al., 2006) and throughout this study, the amplitudes of such non-stationary events can all be suppressed by down-weighting the contributions from the endpoints of surface arrays with a taper.

We also find that non-physical energy associated with the direct wave (non-scattered) Green’s function  $G_0^*(\mathbf{x}', \mathbf{x})$  can be reduced in amplitude by using the exact interferometric representation (Eqs. 2.1 and 2.2) rather than the monopole approximation (Eqs. 2.3 and 2.4). This is because  $G_0^*(\mathbf{x}', \mathbf{x})$  does not fulfil the far-field assumptions when boundaries  $S$  and  $S'$  are close to one-another (Fig. 2.7): the assumption that all ray paths (including those between the source and receiver boundaries) are normal to the boundaries is not valid in this case. (Note that when using  $G_S^*(\mathbf{x}', \mathbf{x})$  instead of  $G_0^*(\mathbf{x}', \mathbf{x})$ , the assumption is more reasonable because the energy generated at the source boundary travels towards the scatterer

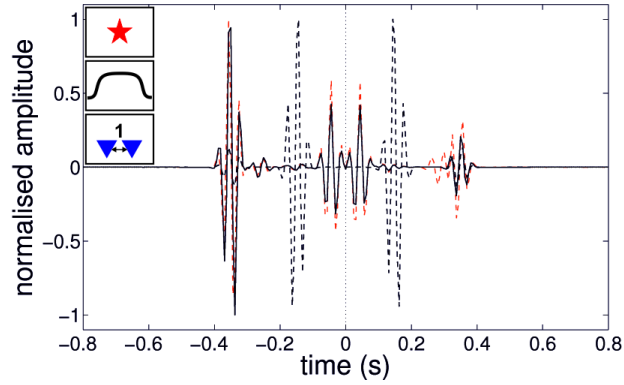


Figure 2.11: Source–receiver estimate of the Green’s function between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  (solid line) constructed using the partial boundaries in Fig. 2.7, when  $G_0^*(\mathbf{x}', \mathbf{x})$  is eliminated from the data prior to the cross-correlation. The dashed black line represents the true Green’s function; the dashed red line corresponds to the interferometric trace in Fig. 2.8. Maximum amplitudes in each trace are normalized to one. The non-physical arrival before the pseudo-physical causal scattered wave in Fig. 2.8 is completely suppressed, resulting in a good estimate of the scattered wavefield.

first before being recorded at the receiver boundary: provided the boundaries are far from the scatterer this leads to an ultimate propagation direction that is closer to the normal to the boundary.) In fact, terms  $G_0G_0G_0^*$ ,  $G_SG_SG_0^*$ ,  $G_0G_SG_0^*$ , and  $G_SG_0G_0^*$  only contribute non-physical non-stationary energy. Since all of these terms, and only these terms, contain the direct wave arrival  $G_0^*(\mathbf{x}', \mathbf{x})$  between boundaries  $S$  and  $S'$ , eliminating this direct wave component altogether from the interferometry (i.e. setting it to zero prior to cross-correlating wavefields) reduces the amount of non-physical energy without losing physical information about the scattered wavefield (Fig. 2.11). In this way, the monopole or far-field approximation can be used without causing significant negative effects, even if the boundaries are close together or even if they are collocated.

Stationary but non-physical events are constructed from the terms  $G_0G_SG_S^*$  and  $G_SG_0G_S^*$  on account of the stationary pairs  $\mathbf{x}' = \underline{\mathbf{c}}, \forall \mathbf{x}$  and  $\mathbf{x} = \underline{\mathbf{e}}, \forall \mathbf{x}'$  (Fig. 2.3). Usually these events would be cancelled out by other non-physical events associated with the stationary points at  $\mathbf{d}$  and  $\mathbf{f}$ . Using only partial boundaries, however, the linear arrays omit the corresponding stationary points and thus preserve the non-physical energy in the constructed trace. They are identified as the first arrivals, with travel times corresponding to  $t(\mathbf{x}_2, \mathbf{x}_d) - t(\mathbf{x}_1, \mathbf{x}_d)$  and  $t(\mathbf{x}_1, \mathbf{x}_d) - t(\mathbf{x}_2, \mathbf{x}_d)$ , respectively, where  $t(\mathbf{x}_1, \mathbf{x}_d)$  is the travel time from the diffractor at  $\mathbf{x}_d$  to  $\mathbf{x}_1$ . Although their travel times do not relate to physical ray paths they still contain information about the medium, especially about the location of the scatterer. For example, if the scatterer was located at the midpoint

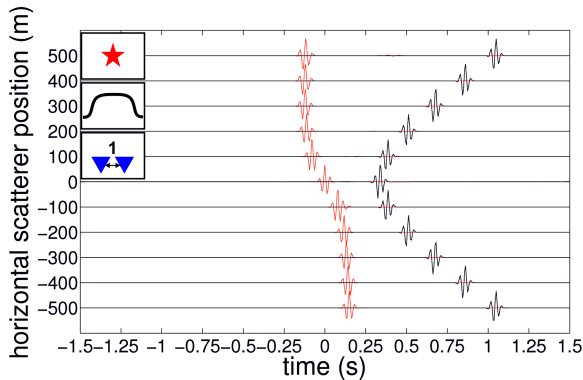


Figure 2.12: Pseudo-physical (black lines) and non-physical (red lines) events constructed from  $G_S G_S G_S^*$  and  $G_0 G_S G_S^*$ , respectively, for different horizontal scatterer positions—moving the scatterer in Fig. 2.7 relative to  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , while keeping the vertical location at 0 m. For any fixed scatterer position relative to  $\mathbf{x}_1$  and  $\mathbf{x}_2$  the position of the scatterer can be estimated using the combined travel time information from both events.

between source and receiver the travel times of both events would be zero and they would coincide at zero lag-time. We can thus use the information from the travel times of non-physical events to constrain the position of the scatterer. Combining this information with the travel time of a physical scattered wave, causal or acausal, given by  $t(\mathbf{x}_d, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_d)$  or  $-[t(\mathbf{x}_d, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_d)]$ , respectively, the position of a scatterer located below the source-receiver pair is uniquely defined (Fig. 2.12). Non-physical energy from standard interferometry has been well studied and shown to be useful in velocity analysis (King and Curtis, 2011) and locating near-surface scatterers (Harmankaya et al., 2013; Kaslilar et al., 2014). Similar applications appear feasible for non-physical energy constructed from SRI. Further research could examine the potential of using non-physical energy to constrain the scattering amplitude.

In appendix C, we provide an analysis of different parameterizations of the numerical model using, for example, a larger spatial sampling interval while the lateral extents of  $S$  and  $S'$  are held constant. When the sampling interval is increased the contributions from neighbouring traces at non-stationary points may not cancel out and may thus introduce non-physical energy (see Figs. 2.13 and 2.14). Analysing each term of Eq. 2.7 individually we find that some terms contribute energy that is relatively robust to changes of the sampling interval: the terms  $G_0 G_0 G_S^*$  and  $G_S G_S G_S^*$  still provide good estimates of the acausal and causal scattered wave, respectively, when the sampling interval equals six times the spatial Nyquist wavelength  $\lambda_{Nyq}$  (Fig. 2.15), which corresponds to a spacing of 38 m. This is also true for the stationary non-physical events in  $G_0 G_S G_S^*$  and

$G_S G_0 G_S^*$ . Thus, the additional non-physical energy introduced by depopulating the boundaries appears solely on account of the terms  $G_0 G_0 G_0^*$ ,  $G_S G_S G_0^*$ ,  $G_0 G_S G_0^*$ , and  $G_S G_0 G_0^*$ , all of which contain the direct wave arrival between the two boundaries,  $G_0(\mathbf{x}', \mathbf{x})$ . Again, by eliminating this component prior to the cross-correlation of wavefields we can therefore reduce the amount of non-physical energy and apply a coarser source and receiver spacing without loss of resolution of the scattered waves.

The differing behaviour with respect to the spacing on each boundary can be understood by considering the correlation gathers: for each term, the maximum allowable spacing is determined by the slope of the travel time curve in the correlation gather, which depends on the choice of the Green's functions in the cross-correlation. When  $G_0(\mathbf{x}', \mathbf{x})$  is used rather than  $G_S(\mathbf{x}', \mathbf{x})$  the travel time curve in the first correlation gather (corresponding to inter-receiver interferometry) has a much steeper slope (compare Figs. 2.13 and 2.14, for example) and therefore causes incomplete cancellations even for a small increase of the sampling interval above the Nyquist wavelength. In general, the travel time slope depends on the velocity of the medium, the depth of the source–receiver pair, and the depth of the scatterer (Mehta et al., 2008). Further, the behaviour of the travel time curves may be different for multiply scattered or reflected waves. The maximum allowable spacing is thus defined by the geometry and material properties of the problem at hand.

For the geometry used, only the term  $G_0 G_0 G_S^*$  gives the acausal scattered wave (Figs. 2.9a, 2.10a and b) on account of the stationary pair  $\mathbf{x} = \underline{\mathbf{e}}, \mathbf{x}' = \underline{\mathbf{c}}$  (Fig. 2.3). This term has also been used by Poliannikov (2011) to recover the reflection response of a layered medium using SRI. Moreover, an estimate of the causal scattered wave is obtained from the pseudo-physical event constructed from  $G_S G_S G_S^*$  (Fig. 2.9b), for which every source-receiver pair is stationary (Fig. 2.10c and d). Note that in the geometry used herein,  $G_S G_S G_S^*$  is the only term that contains information about the causal scattered wave. If we therefore use only terms  $G_0 G_0 G_S^*$  and  $G_S G_S G_S^*$ , we obtain a good estimate of the causal and acausal scattered field (Figs. 2.15a and c) even if the boundary source and receiver sampling is depleted (Figs. 2.15b and d).

Considering the applicability of these results in an imaging context, when the aim is to image the scatterer we have to consider the following limitations: compared to the geometry used in SRI (Fig. 2.7), in a seismic experiment the subsurface source at  $\mathbf{x}_1$  and the receiver at  $\mathbf{x}_2$  are physically not available, so the wavefields



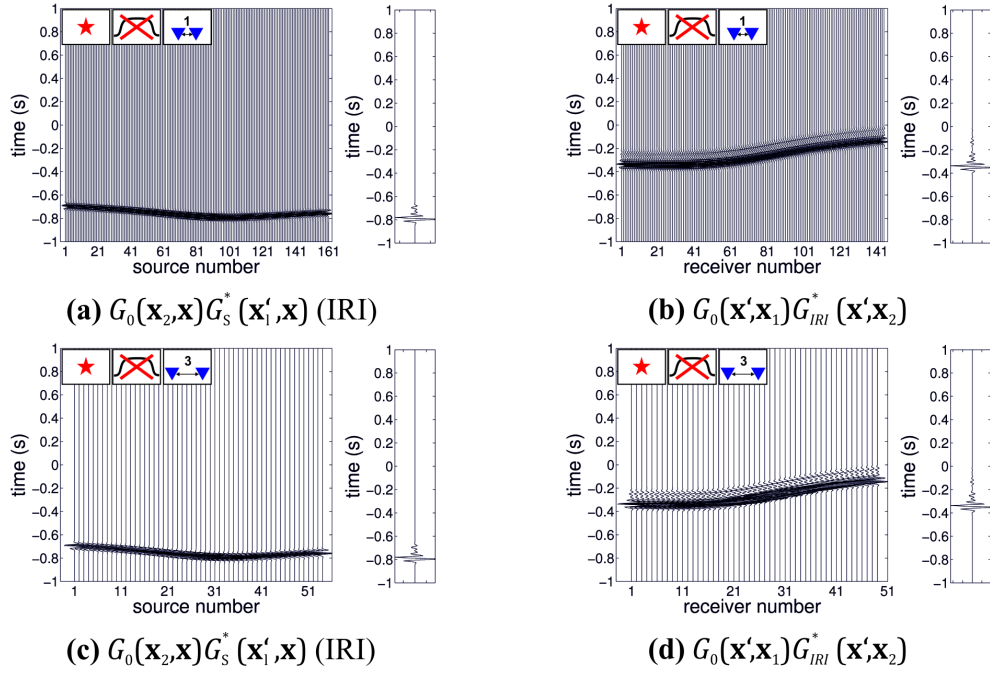


Figure 2.13: Correlation gathers of  $G_0 G_0 G_S^*$  for different parameter constellations (for symbol key see Table 2.2). Panels (a) and (b) correspond to (a) and (b) in Fig. 2.10 (c) As in (a) but with the source interval equal to  $3 \cdot \lambda_{Nyq}$ . (d) As in (b) but with the receiver interval equal to  $3 \cdot \lambda_{Nyq}$ . Figure layout as in Fig. 2.10. In this example the coarser spatial sampling does not affect the interferometric result.

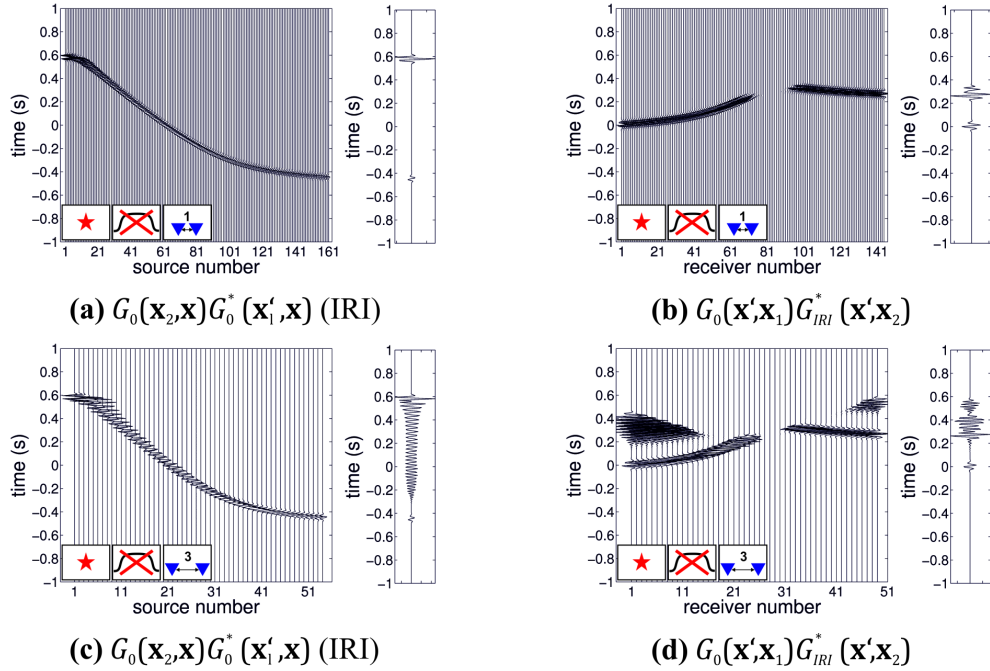


Figure 2.14: As in Fig. 2.13 but for  $G_0 G_0 G_0^*$ . (a) and (b) above correspond to (i) and (j) in Fig. 2.10. When a coarser spatial sampling is applied (c and d) additional non-physical energy is introduced in the interferometric estimate.

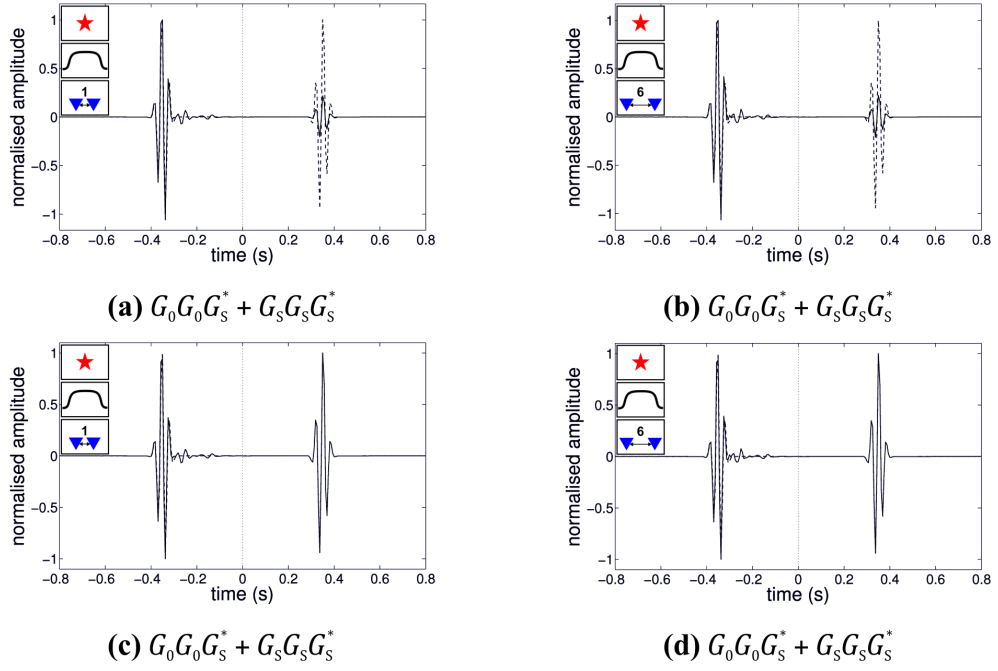


Figure 2.15: (Pseudo-)Physical scattered wave energy constructed using the terms  $G_0 G_0 G_S^*$  and  $G_S G_S G_S^*$  only (solid line) compared to the true scattered wave (dashed line). The maximum amplitudes have been normalized to one. **(a)** Source and receiver spacing is equal to  $1 \cdot \lambda_{Nyq}$ . **(b)** Source and receiver spacing is equal to  $6 \cdot \lambda_{Nyq}$ . **(c)** and **(d)** show the same results as **(a)** and **(b)**, but causal and acausal (positive and negative time) sides have been normalized independently to better illustrate the quality of the interferometric result. Despite the strongly depleted boundaries in **(b)** and **(d)** the scattered wavefield is well constructed and no additional spurious energy is introduced.

$G(\mathbf{x}', \mathbf{x}_1)$  and  $G(\mathbf{x}_2, \mathbf{x})$  are not recorded. In imaging methods these wavefields are modelled using a smooth background model, usually obtained from velocity analysis and waveform inversion (Pratt, 1999; Yilmaz, 2001). Under these conditions we find that  $G_0 G_0 G_S^*$  has a crucial advantage over  $G_S G_S G_S^*$ : while modelling of  $G_S(\mathbf{x}', \mathbf{x}_1)$  and  $G_S(\mathbf{x}_2, \mathbf{x})$  requires information about the scatterer (which is usually not available from the background model),  $G_0(\mathbf{x}', \mathbf{x}_1)$  and  $G_0(\mathbf{x}_2, \mathbf{x})$  are solely defined by the background model. In fact,  $G_0 G_0 G_S^*$  can be compared directly to the imaging condition for a migrated image (Claerbout, 1985) given that source and receiver coincide on the image point  $\mathbf{x}_1$ , since the explicit link between imaging and SRI was provided by Halliday and Curtis (2010). They derive the scattered wave components of SRI from reciprocity relations for perturbed media, and show that under the Born approximation the scattering potential  $f$  at a point  $\mathbf{x}_1$  is given by

$$f(\mathbf{x}_1) = \frac{-4}{j c_0} \int_{-\infty}^{\infty} d\omega (-j\omega) \times \int_S [\partial_i \Phi(\mathbf{x}_1, \mathbf{x}) G_0^*(\mathbf{x}_1, \mathbf{x}) - \Phi(\mathbf{x}_1, \mathbf{x}) \partial_i G_0^*(\mathbf{x}_1, \mathbf{x})] n_i dS \quad (2.15)$$

where  $\Phi(\mathbf{x}_1, \mathbf{x})$  represents the back propagated wavefield at  $\mathbf{x}_1$

$$\Phi(\mathbf{x}_1, \mathbf{x}) = \frac{-1}{j\omega\rho} \int_{S'} [\partial'_i G_S(\mathbf{x}', \mathbf{x}) G_0^*(\mathbf{x}_1, \mathbf{x}') - G_S(\mathbf{x}', \mathbf{x}) \partial'_i G_0^*(\mathbf{x}_1, \mathbf{x}')] n'_i dS' \quad (2.16)$$

Note that  $G_0(\mathbf{x}_1, \mathbf{x})$  and  $G_0(\mathbf{x}_1, \mathbf{x}')$  are not measured quantities but synthetic forward-propagating (from sources at  $\mathbf{x}$ ) and backpropagating (from receivers at  $\mathbf{x}'$ ) Green's functions, respectively, calculated using the background model. As Halliday and Curtis (2010) explain, Eq. 2.15 is directly related to the imaging condition that Oristaglio (1989) derived using a double-focusing algorithm. We find that Eqs. 2.15 and 2.16 show striking similarities to Eq. 2.2 and 2.1, respectively, assuming that  $\mathbf{x}_2 = \mathbf{x}_1$ , and setting  $G(\mathbf{x}', \mathbf{x}_1) = G_0(\mathbf{x}', \mathbf{x}_1)$ ,  $G(\mathbf{x}_2, \mathbf{x}) = G_0(\mathbf{x}_2, \mathbf{x})$ , as if using the term  $G_0 G_0 G_S^*$  only and invoking source–receiver reciprocity.

Note that Halliday and Curtis (2010) use a complete circular boundary of sources and receivers. In their derivation they find that not only the scattered wave is constructed but also its time reverse, as well as two events similar to stationary but non-physical arrivals in interferometry. This is consistent with the events expected from the term  $G_0 G_0 G_S^*$  by stationary phase analysis when using a full boundary (Table 2.1). When restricting the boundary to the surface, however,

only the acausal scattered wave is constructed due to the lack of stationary points associated with the causal scattered wave and the two stationary, non-physical events. Hence an imaging condition can be derived from SRI even if the boundaries are only partially available on top of the scattering medium. Indeed the suppression of the two non-physical events when using partial boundaries can be seen as a positive advantage of using incomplete boundaries, since that non-physical energy will not disturb the image.

As the interferometric approach does not make use of the Born approximation it is in principle able to account for non-linearities associated with multiple scatterers, which are currently not considered by standard migration schemes. When an initial estimate of the scattered wavefield is included in the reference wavefield  $G_0$  it also becomes possible to use additional interferometric terms for non-linear imaging, and some work has already been done in this area (Fleury and Vasconcelos, 2012; Ravasi and Curtis, 2013; Vasconcelos, 2013). The question of how our specific results generalize to the case of a multiply scattering medium, and how this could be used to enhance resolution in seismic images, will be addressed in future research. An example for pseudo-physical energy constructed in a multiple scattering case is provided in Fig. 2.17. A full boundary has been used (Fig. 2.16) to highlight the effect of the scattering amplitude on the amplitude of the constructed events. For first-order scattering the analysis provided for a single scatterer applies just as well in the multiple-scattering case: the amplitude of the constructed event is proportional to the imaginary part of the scattering amplitude of the corresponding scatterer. Note that the scatterers have different scattering amplitudes, which results in different amplitudes of the primary events on the constructed trace. In principle, scattering events of any order are constructed using  $G_S G_S G_S^*$  only. For example, a secondary event can be seen at around 0.7s. For a kinematic analysis of higher-order scattering see Meles and Curtis (2013). How the amplitude (and phase) is affected by the scattering amplitudes of the individual diffractors has to be clarified in future research. Moreover, non-physical events are introduced from the correlation of cross-terms (e.g. at 0.2s). Note that those may superimpose pseudo-physical arrivals and affect their amplitudes and waveforms. Nevertheless, this example shows that the single-scattering analysis presented herein is useful and applicable, at least in relatively simple multiple-scattering scenarios.

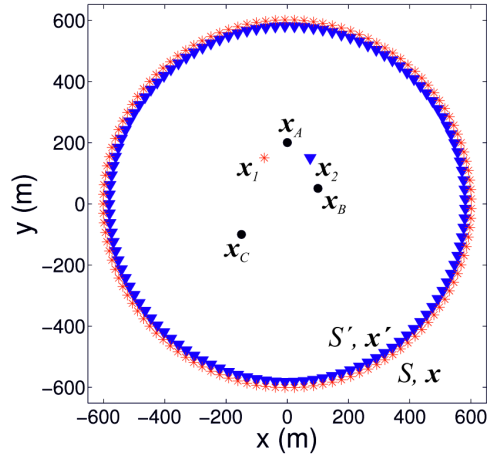


Figure 2.16: Geometry used in multiple diffractors example. Symbol key as in Fig. 2.2. The imaginary part of the scattering amplitude of the diffractors located at  $\mathbf{x}_A$  and  $\mathbf{x}_B$  is set to  $-1$ , for the scatterer at  $\mathbf{x}_C$  it is set to  $-2$ .

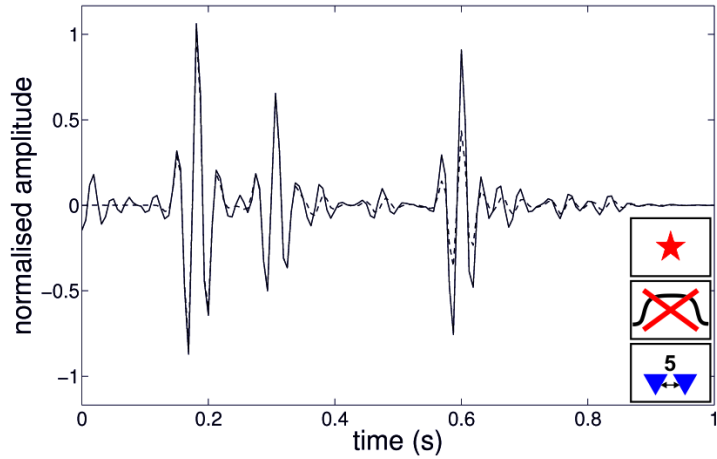


Figure 2.17: Pseudo-physical events (solid line) constructed using  $G_S G_S G_S^*$  only and the geometry shown in Fig. 2.16 containing three diffractors. The dashed line represents the true scattered Green's function. All first-order scattering events (primaries) are constructed with the correct travel time and waveform. According to Eq. 2.12 the primaries around 0.2s and 0.3s have exactly the amplitude of the true events, since the imaginary part of the scattering amplitude of the corresponding scatterers ( $\Im(A_A)$  and  $\Im(A_B)$ , respectively) equals  $-1$ . Analogously, the primary at 0.6s is constructed with twice the correct amplitude, since  $\Im(A_C) = -2$ . Moreover, good estimates of higher-order scattering events have been constructed, for example, between 0.4s and 0.6s and around 0.8s. Note, however, that they appear to be shifted in phase. The small amplitude events before 0.15s are non-physical.

## 2.6 Conclusions

Using synthetic acoustic scattered waves we have illustrated the ability of SRI to provide information about scatterers embedded in a smooth background medium using a limited geometry of source and receiver boundaries representing linear arrays used in industrial geophysics. By separating the wavefield into a background component and a scattered wave component and analysing individual cross-terms of the interferometric equation using the method of stationary phase, we determine the origin of both physical and non-physical energy in the resulting Green's function estimates. We identify a new category of non-physical energy, referred to as pseudo-physical energy, which can be used to estimate physical energy directly. We show that the scattered wave is constructed by only one term of the equation, referred to as  $G_0G_0G_S^*$ , which is directly linked to the imaging condition used in standard seismic migration schemes. We showed that for this term the partial boundary may be a positive advantage as this suppresses non-physical energy in resulting images. The term  $G_SG_SG_S^*$  provides a pseudo-physical event, which is naturally non-physical but can be used as an estimate of the causal scattered wave. For the first time a complete mathematical derivation for the generation of pseudo-physical energy is provided. Two other terms contain stationary non-physical energy that is not cancelled out when using incomplete boundaries; however, this energy was shown to provide novel information about the location of scatterers. Non-stationary, non-physical energy associated with the abrupt truncation of the boundaries and the monopole approximation can be reduced by using a spatial taper, the use of dipole sources, or the elimination of the direct wave component between the boundaries,  $G_0(\mathbf{x}', \mathbf{x})$ , prior to the cross-correlation of wavefields. Considering source and receiver coverage along the boundaries, our studies reveal four terms that permit deviations from theoretical sampling requirements while still providing a reliable estimate of stationary energy, and fortunately for practical applications, these include the two terms  $G_0G_0G_S^*$  and  $G_SG_SG_S^*$  that construct the (pseudo-)physical scattered waves.

# Chapter 3

## Automatic identification of multiply diffracted waves and their ordered scattering paths

In the previous chapter, I studied a single point diffractor and determined the contribution of individual terms of the SRI equation to the scattered wavefield estimate when source and receiver boundaries are only partially available at the surface. From this analysis, the term  $G_S G_S G_S^*$ , which has only scattered wave arguments, stands out because of its ability to construct a pseudo-physical estimate of the causal scattered wave from *any* source receiver pair on the boundary and thus provides useful results even for very limited boundaries.

In this chapter, I consider multiple isotropic point scatterers and focus on the analysis rather than on the construction of a multiply scattered wavefield  $G_S$ .  $G_S$  consists of singly scattered and multiply scattered waves and I show how these different components can be distinguished and the exact scattering path of any event be identified using the term  $G_S G_S G_S^*$  of the SRI equation. However, I do not apply full wavefield cross-correlation and convolution as in SRI but focus on the summation and subtraction of travel times in the phase term, which is sufficient to determine the arrival times of multiply scattered waves. The algorithm presented in this chapter is a modified version of a method described in detail in Meles and Curtis (2014a) and predicts rather than retraces multiply scattered arrivals. Special emphasis is placed on the automation of the different steps of the algorithm involving isolation of primaries (singly-scattered waves), identification of secondaries (twice-scattered waves), and prediction of higher-order multiples.

This chapter has been published as a jointly-authored paper<sup>1</sup>. I, as lead author,

---

<sup>1</sup>K. Lör, G.A. Meles, and A. Curtis. Automatic identification of multiply diffracted waves and their ordered scattering paths. *Journal of the Acoustical Society of America*, 137(4): 1834–

have done the writing of the paper, performed the numerical experiments, and analysed the results. Co-authors gave advice and support on the scope of the project, provided background knowledge, and helped editing the manuscript. For the construction of synthetic data sets I used an analytic wavefield modeling code (Foldy, 1945; Galetti et al., 2013) and a finite difference acoustic modeling code written by Matteo Ravasi (PhD student, University of Edinburgh).

## Abstract

An automated algorithm is presented that uses recordings of acoustic energy across a spatially-distributed array to derive information about multiply scattered acoustic waves in heterogeneous media. The arrival time and scattering-order of each recorded diffracted acoustic wave, and the exact sequence of diffractors encountered by that wave, are estimated without requiring an explicit model of the medium through which the wave propagated. Individual diffractors are identified on the basis of their unique single-scattering relative travel-time curves (move-outs) across the array, and secondary (twice-scattered) waves are detected using semblance analysis along temporally offset primary move-outs. This information is sufficient to estimate travel times and scattering paths of all multiply diffracted waves of any order using a process based on source-receiver interferometry. These events can then be identified in recorded data. The algorithm is applied to synthetic acoustic data sets from a variety of media, including different numbers of point diffractors and a medium with strong heterogeneity and non-hyperbolic move-outs.

## 3.1 Introduction

The phenomenon of diffraction describes the interaction of propagating energy (e.g., acoustic and elastic waves, electromagnetic radiation, or moving particles) with sub-wavelength heterogeneities within the medium that are generally referred to as scatterers or diffractors (Born and Wolf, 1999). Diffractions play an important role in many fields of theoretical and applied acoustics, including in medical imaging (Insana et al., 1990; Tadayyon et al., 2014), localization and destruction of kidney stones using medical ultrasound and lithotripsy (Fink et al.,

---

1845, 2015. DOI: 10.1121/1.4906839.



2003), ocean acoustics for the detection of marine organisms (Brekhovskikh and Lysanov, 2003; Foote, 2008), but also in other fields of physics including quantum mechanics (Friedrich, 2006), non-destructive testing (Prada et al., 2002), remote sensing (Ferretti et al., 2001), ground-penetrating radar (Papziner and Nick, 1998), near-surface geophysics (Harmankaya et al., 2013; Kaslilar et al., 2014), seismic exploration and monitoring (Landa et al., 1987; Khaidukov et al., 2004; Pacheco and Snieder, 2006; Halliday and Curtis, 2009; Halliday et al., 2010; Jixiang et al., 2014), and global seismology (Wu and Aki, 1988). In all such cases being able to predict or interpret diffracted energy is crucial.

Although in most applications the medium of interest contains multiple diffractors, it is often assumed that significant recorded wave energy has only scattered once, in order to simplify the wave theory considered. In that case, only energy that has interacted with a single diffractor is correctly taken into account. Apart from the fact that this assumption neglects many of the data which contain additional information about the medium, it may also lead to misinterpretation of the data and hence to incorrect conclusions. As examples in various fields of ongoing research show (Stanton, 1982; Gao et al., 1983; Bordier et al., 1991; De Rosny and Roux, 2001), taking multiple scattering into account often leads to improved results. Some authors (Larose et al., 2006; Aubry and Derode, 2010) also address the problem of separating singly from multiply scattered wavefields and analyzing the information content in different parts of the wavefield separately.

Recently, Meles and Curtis (2014a) presented a new method to identify multiply diffracted waves in acoustic data gathers. Moreover, it identifies all individual diffractors involved in the corresponding scattering path, and the sequential order in which they were encountered. It relies on fingerprinting individual diffractors in common-source and common-receiver gathers (data subsets) by means of their unique move-out (travel-time variation across arrays of receivers or sources, respectively). The method has a range of possible applications such as improved localization of diffractors or estimation of inter-scatterer medium properties, or discriminating physical from non-physical energy in wavefield interferometry (Meles and Curtis, 2014b); however, until now the method required substantial manual intervention. This is possible, though time-consuming, only in a medium of low complexity (with very few diffractors) and for data with very low noise levels.

We present an algorithm that automatically identifies primary and secondary waves (waves that have diffracted exactly once or twice, respectively), and pre-

dicts arrival times of higher-order multiply diffracted waves (those that have diffracted three times or more). The latter multiply diffracted wave arrivals may then be identified in recorded data, and associated with an exact scattering path, despite the method requiring no explicit model of the medium or of the diffractor locations. In this paper we briefly revise the theoretical concept of the method of Meles and Curtis (2014a) and introduce the new automated algorithm that predicts the arrival times of multiply scattered events in three steps: (i) automatic extraction of primary travel-time curves (fingerprints) from common-source and common-receiver data gathers, (ii) identification of secondaries on the mutual trace common to both gathers, and (iii) prediction of travel times of all multiples up to any specified order by summing and subtracting primary and secondary travel times. We test the algorithm on a range of synthetic data sets involving different numbers of scatterers, varying noise-to-signal ratios, and non-hyperbolic move-outs. The results are compared to the true arrival times and scattering paths computed from the numerical models, and limitations of the method are addressed in the Discussion.

### 3.2 Theory: Fingerprinting diffractors

The fingerprint of an individual point diffractor corresponds to a unique travel-time curve across a common-receiver gather (CRG) or a common-source gather (CSG) (Meles and Curtis, 2014a). A common-source gather is the set of time series recorded at an array of receivers when the recorded energy has been generated by a single (common) source. Similarly, a common-receiver gather is the set of time series recorded between a single (common) receiver and an array of sources. Fig. 3.1 shows the 3-diffractors model that was used to generate an example of a synthetic common-source gather (Fig. 3.2a) and a common-receiver gather (Fig. 3.2b) using an implementation of Foldy’s method (Foldy, 1945; Galetti et al., 2013). When both a source array and a receiver array are used, a cube of recorded data is generated defined by the sources along one axis, the receivers along another axis, and the recording time along the third axis (Fig. 3.2c). Two crossing slices within this cube correspond to a common-receiver gather and a common-source gather, and have a mutual trace which is common to both gathers and is the record of the wavefield generated from the common source recorded at the common receiver (vertical bold line in Fig. 3.2a, b and c).

We refer to any distinctly observed arriving packet of energy as an event, and to

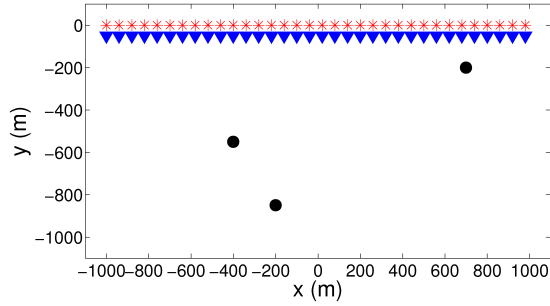


Figure 3.1: Example geometry for a 3-diffractors model. Stars are sources, triangles are receivers, and black solid circles denote the locations of point diffractors. For clarity, only every third source and receiver is plotted. The velocity of the background medium is 1000 m/s

the recorded time series at a single receiver as a trace. The travel-time variations of a singly scattered wave (a primary) prescribe a so-called move-out along the traces of the common-source or common-receiver gather. The move-out gives the relative time of arrival of the diffracted wave as a function of receiver or source position respectively, and is solely determined by the properties of the diffractor and the background medium, and the location of the diffractor with respect to the receiver array or the source array. Each diffractor can be related to a unique move-out, its so-called fingerprint. Meles and Curtis (2014a) show that this uniqueness also holds for inhomogeneous media, other than for pathological cases.

The number of distinct primary move-outs in the data (the common-source or common-receiver gathers) corresponds to the number of diffractors in the illuminated part of the medium. Further, multiple occurrences of the same move-out arriving at different times indicate multiply scattered events. In a common-source gather, multiply scattered waves can be classified by the last diffractor they have visited: the move-out curve of any multiply scattered wave with last diffractor  $L$  is equal to that of the primary of  $L$ , with an additional constant travel-time shift that accounts for the longer raypath before visiting diffractor  $L$  (Figs. 3.3a and b). Hence, all events with the same last diffractor have the same move-outs in a common-source gather. In a common-receiver gather, all events can be classified similarly according to the first diffractor along the scattering path (Fig. 3.3c and d). Combining the information from both gathers on any chosen mutual trace (that is common to both gathers; Fig. 3.3e) we can identify the first and last diffractor along the scattering path for any event observed on the trace: thus, every event observed on any chosen trace can be associated with a move-out pair, namely the fingerprints of the first and last diffractor. This is important because

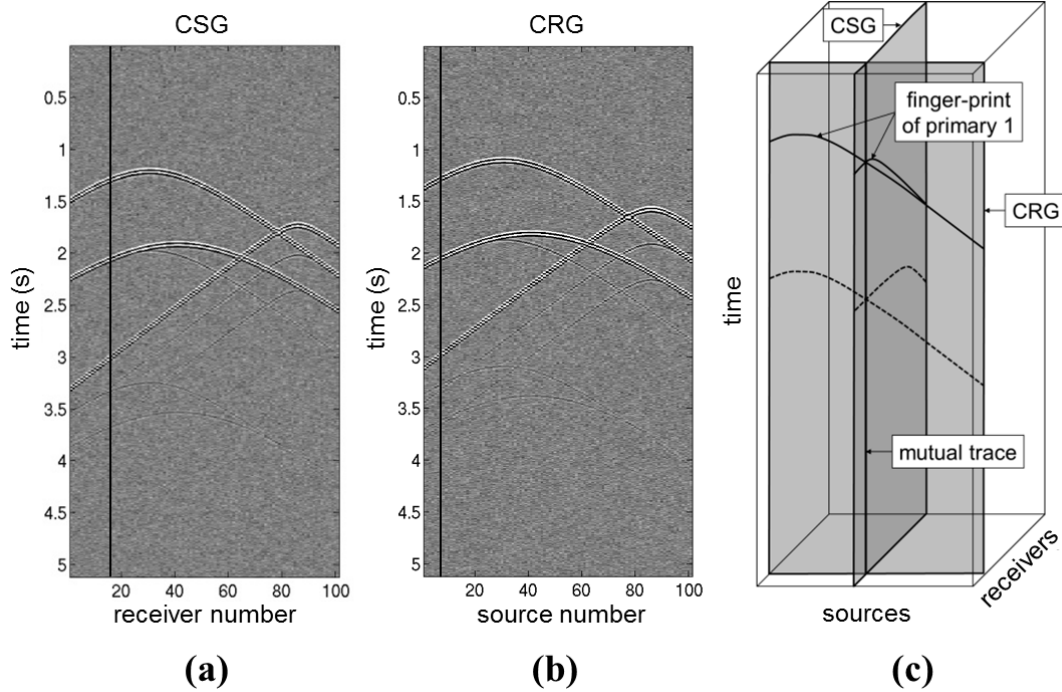


Figure 3.2: (a) Synthetic common-source gather (CSG) and (b) common-receiver gather (CRG) obtained using the 3-diffractors model shown in Fig. 3.1 with the common source located at  $[-880\text{ m}, 0\text{ m}]$  and the common receiver located at  $[-680\text{ m}, -50\text{ m}]$ , respectively. The noise-to-signal ratio is 0.5, where a ratio of 1 means that the root-mean square noise is equal to the typical amplitude of a secondary (a twice-scattered wave). (c) The data-cube showing a CSG and a CRG as two orthogonal (schematic) sections. The vertical bold line in both panels of (a) marks the mutual trace that is common to both of the gathers, and this is denoted also in the schematic gathers in (c).

primaries and secondaries can all be identified uniquely as the first events on the recorded trace associated with a particular move-out pair: for primaries both move-outs match the fingerprints of the same (single) diffractor, while for secondaries the source and receiver move-outs match the fingerprints of different first and last diffractors.

Meles and Curtis (2014a) then choose an arbitrary event on any trace and analyse the sequential order of diffractors involved in its scattering path. The last diffractor  $L$  in the scattering path is classified according to its move-out as shown above. The penultimate diffractor,  $L - 1$ , is identified by using a particular combination of cross-correlation and convolution of primaries and secondaries (or by a combination of additions and subtractions of their travel times). By induction, the diffractors  $L - 2$ ,  $L - 3$ , and so on are identified in the same way until the full scattering path of any chosen event is recovered.

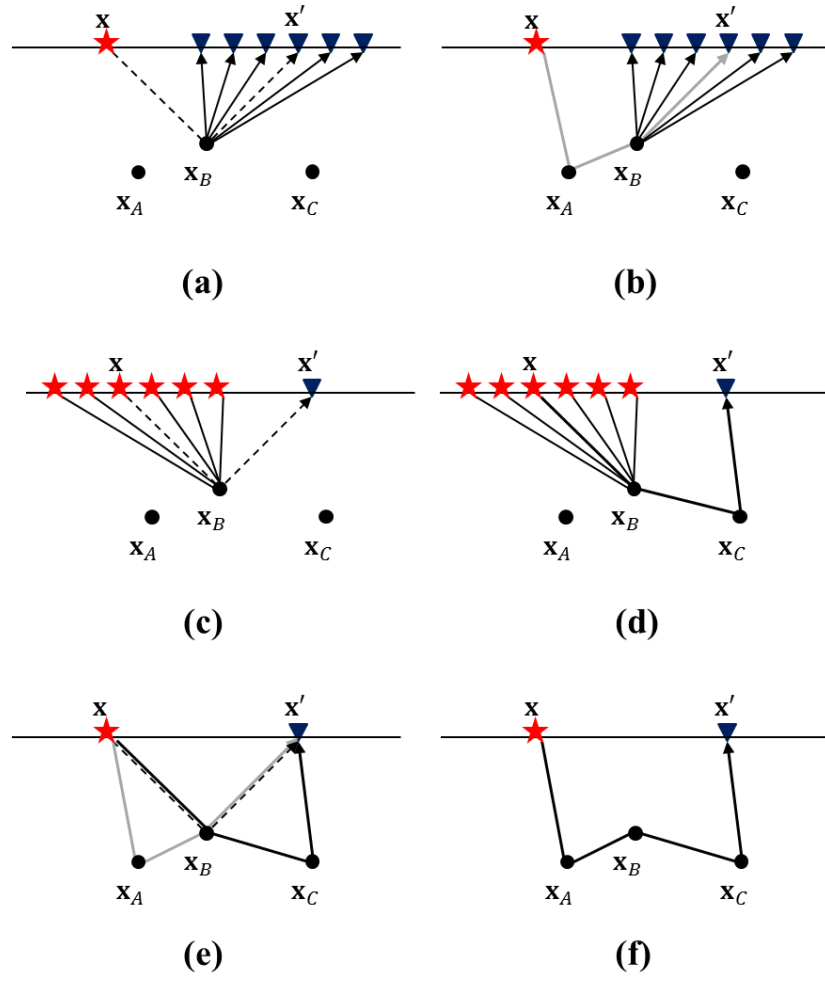


Figure 3.3: Schematic construction of multiply diffracted events from primaries and secondaries. Symbol key as in Fig. 3.1. (a) and (b) CSG: raypaths of the primary scattered at  $\mathbf{x}_B$  and of the secondary with first diffractor  $\mathbf{x}_A$  and last diffractor  $\mathbf{x}_B$ . Note that the receiver-side move-outs are the same. (c) and (d) CRG: raypaths of the primary scattered at  $\mathbf{x}_B$  and of the secondary with first diffractor  $\mathbf{x}_B$  and last diffractor  $\mathbf{x}_C$ . Note that the source-side move-outs are the same. (e) and (f) Events on the mutual trace that are used to estimate the travel time of a tertiary scattered event: two secondaries (grey and bold) are convolved and the result is cross-correlated with a primary (dashed). Travel times along dashed and solid raypath components that follow the same paths cancel each other as a result of cross-correlation. This operation gives the travel time of the tertiary event scattered sequentially at diffractors  $\mathbf{x}_A$ ,  $\mathbf{x}_B$ , and  $\mathbf{x}_C$ .

The method requires that the move-out of the event can be classified uniquely on both the common-source and the common-receiver gather. This can be difficult especially for late arriving waves that undergo higher-order scattering, have low amplitudes, and are embedded in a complex wavefield where individual move-outs may not be distinguishable. Nevertheless, the method has been demonstrated successfully for scattering events up to fourth order in a noise-free synthetic acoustic data set; that is, the scattering path of any observed event could be interpreted provided the energy in that event had not scattered more than four times.

In this work we implement, automate, and demonstrate the reversed process to that of Meles and Curtis (2014a): rather than analyzing scattering paths of events observed on the recorded trace, our method predicts travel times of all multiply diffracted waves (Fig. Fig. 3.3e and f), which can then be identified in recorded data. In fact, this method was proposed schematically by Meles and Curtis (2014a), but was neither implemented nor tested.

### **3.3 The automated scheme**

Automation of the new algorithm requires a method to detect move-outs of different diffractors, and repetitions of the same move-out across each gather. It should be applicable to either simple or complex wavefields where the human eye may not be able to recognize individual move-outs. We now describe three stages in an algorithm that achieves this.

#### **3.3.1 Isolating primary move-outs using cross-correlation of gathers**

We first isolate and identify the primary move-outs, which provide the basis for all further analyses. To do so, we exploit the fact that the shape of a primary diffracted move-out in a common-source gather is invariant with respect to the source position apart from a constant shift in time. That is, if we compare a common-source gather to a second such gather with a different source position, we will find the same primary move-out curves in the second gather, but each curve will be shifted by a different amount along the time axis (Figs. 3.4a and b). This assumption applies equally to common-receiver gathers with respect to location of the receiver.

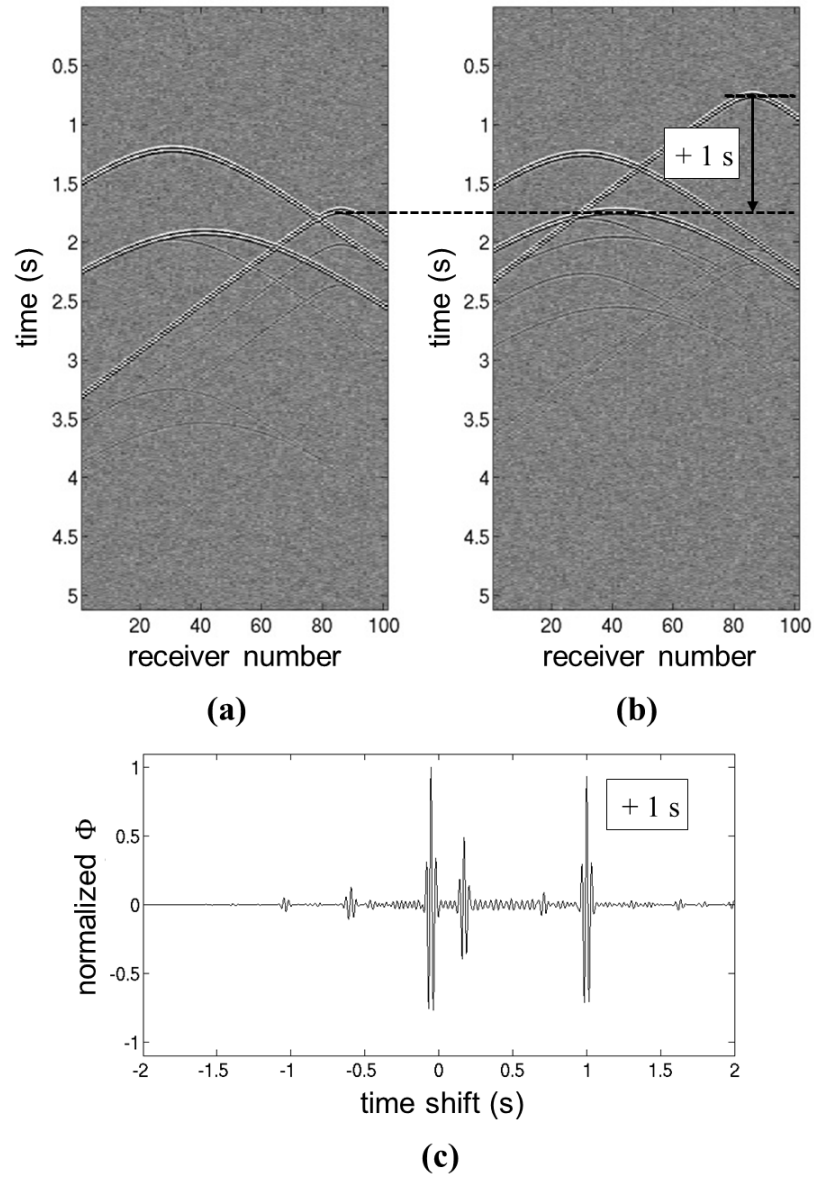


Figure 3.4: **(a)** Common-source gather with the source located at  $[-880\text{m}, 0\text{m}]$  and **(b)** common-source gather with the source located at  $[-380\text{m}, 0\text{m}]$  (see Fig. 3.1). The noise-to-signal ratio is 0.5. Both gathers contain the same three primary move-outs (fingerprints of each of the three diffractors in Fig. 3.1), but shifted in time (i.e., vertically up or down the gather) due to the change in source position. As an example the dashed lines indicate the shift by 1 s of the rightmost move-out. **(c)** Result of cross-correlation ( $\Phi_1$  in Eq. 3.1) between the two gathers shown in (a) and (b). Time lag  $i$  in Eq. 3.1 has been converted to seconds. The peaks indicate the three time shifts under which both gathers are most alike with the peak at 1 s corresponding to the shift of the rightmost move-out in both gathers. These time shifts are used to isolate the individual move-outs from the left gather in (a).

The invariance of diffraction move-outs across multiple common-source gathers allows us to isolate them from the rest of the data and estimate their arrival-time curves using cross-correlation. Standard cross-correlation allows one to estimate the time shift between two traces under which they are most alike with respect to a squared norm. We wish to find similarities between two gathers ( $A$  and  $B_1$ ) rather than two individual traces, so we first perform cross-correlation between the pair of traces at each receiver in the two gathers, then sum the results over receivers. This can be written in the time domain for discrete time samples  $i$  as

$$\Phi_1(i) = \sum_{m=1}^{\min(M_A, M_B)} \sum_{n=1}^N A(m, n) B_1(m + i, n) \quad (3.1)$$

where  $\Phi_1(i)$  denotes the correlation coefficient at time shift  $i$ ,  $M_A$  and  $M_B$  are the record lengths in gathers  $A$  and  $B_1$  respectively, and  $N$  is the number of traces in each gather. In gather  $A(m, n)$ , for example,  $m$  is the time index and  $n$  is the receiver index.

$\Phi$  has its maximum recorded amplitude at time shift  $i = i_1$  say, for which  $A$  and  $B_1$  are most alike.  $i_1$  is usually the time by which one of the recorded primary move-outs (typically the one with the largest amplitude) is shifted due to the relative shift in source position between gathers  $A$  and  $B_1$  (Fig. 3.4c). To identify this particular move-out, gather  $B_1$  is shifted in time by  $i_1$  and multiplied element-wise with gather  $A$  according to

$$C_1(m, n | i = i_1) = A(m, n) | B_1(m + i_1, n) | \quad (3.2)$$

Ideally, the resulting cross-gather  $C_1$  has maximum values along the primary move-out associated with time shift  $i_1$ : all other elements should have close to zero amplitudes (Figs. 3.5a, b and c). However, when the data are noisy or contain many intersecting move-outs, residual energy that is not related to that move-out remains in the cross-gather and affects the accuracy of the estimated move-out. To attenuate this energy, the obtained cross-gather is correlated with a new common-source gather,  $B_2$  say. The maximum amplitude of that correlation function indicates the time-shift between the move-out in the cross-gather and the equivalent primary move-out in gather  $B_2$ , and therefore allows one to compute an updated cross-gather. This process is repeated iteratively, each iteration consisting of three steps:

- (i) Cross-correlation of two gathers (one of which is the previously determined



cross-gather) according to

$$\Phi_{j+1}(i) = \sum_{m=1}^{\min(M_A, M_B)} \sum_{n=1}^N C_j(m, n|i = i_j) B_{j+1}(m + i, n) \quad (3.3)$$

and identification of the time shift  $i_{j+1}$  with the largest correlation coefficient.

- (ii) Element-wise multiplication of the two gathers—one of them shifted by the time shift  $i_{j+1}$  determined in step (i)—to produce a new cross-gather according to

$$C_{j+1}(m, n|i = i_{j+1}) = C_j(m, n|i = i_j) |B_{j+1}(m + i_{j+1}, n)| \quad (3.4)$$

The cross-gather should have maximum values along the primary move-out associated with time shift  $i_{j+1}$ ; all other elements should have close to zero amplitudes.

- (iii) The corresponding travel-time curve is estimated as a function of receiver position by picking the maximum amplitude arrival on each trace of the cross-gather (Fig. 3.5d, e and f):

$$t_j(n|i = i_j) = \arg \max(|C_j(m, n|i = i_j)|) \quad (3.5)$$

The energy along the primary move-out in the cross-gather in step (ii) should be amplified by multiple iterations, allowing a better estimate of the primary travel-time curve. Iterations cease when the extracted travel-time curve does not exhibit significant changes compared to that in the previous iteration.

So far the algorithm identifies one move-out, namely that which exhibits the largest correlation coefficient  $\Phi_1$ . To find other diffraction move-outs we use an additional iterative loop: at the beginning of each iteration recorded energy with previously detected move-outs is removed from the original common-source gather (gather  $A$ ) by muting all traces around the identified arrival-time curves. The muted gather is then used in the initial correlation step (Eq. 3.1). However, when computing the first cross-gather using Eq. 3.2, the original (unmuted) common-source gather is used in order to avoid gaps in the move-out energy due to muting, which would result in discontinuous arrival-time estimates. Thus we iteratively find successive move-out curves.

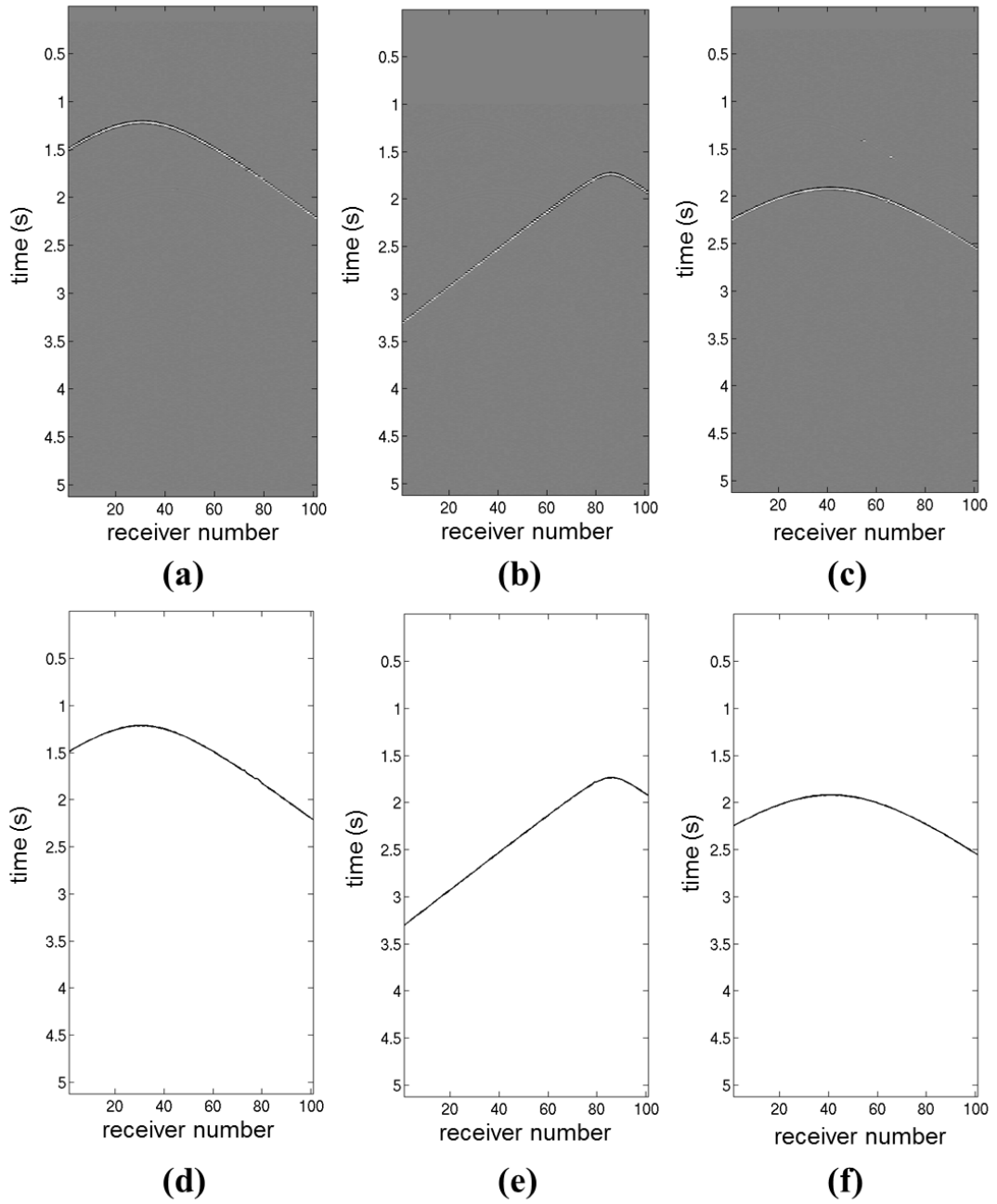


Figure 3.5: (a), (b) and (c) Cross gathers  $C$  for different time lags (Eq. 3.2) each showing a single move-out isolated from the common-source gather in Fig. 3.4(a). (d), (e) and (f) Travel time curve as a function of receiver position extracted from the cross-gathers using Eq. 3.5.

The search for additional move-outs ceases when the maximum of  $\Phi$  falls below a certain threshold, here defined as  $0.1 * \max(\Phi_1) * 0.9^l$ , where  $\max(\Phi_1)$  is the maximum of the first correlation coefficient and  $l$  is the number of move-outs already detected and removed from the gather. This dynamic threshold accounts for the fact that in each iteration the total amount of energy in the gather is reduced due to the muting. The two variable parameters (0.1 and 0.9) are chosen empirically and depend on the relative locations and scattering amplitudes of the diffractors, as well as on the noise level. The influence of this threshold on the results is explained in more detail in the Discussion.

Note that the initial gather  $A$  can be chosen arbitrarily, although certain source locations may provide a better illumination angle (resulting in fewer intersections or larger amplitudes of diffracted energy) depending on the distribution of diffractors in the medium with respect to the source and the receiver array. Nevertheless the identified move-outs should be identical and the extracted arrival-time curves only vary in absolute travel times, so that a comparison of the move-outs obtained from different initial gathers could be used to check the accuracy of results.

The method is in principle able to extract travel-time curves of arbitrary complexity. Below we will show synthetic data examples containing non-hyperbolic move-outs obtained from a numerical model with a highly heterogeneous background velocity distribution. Note that the background medium has to be sufficiently smooth in order not to generate diffracted or reflected energy, i.e., velocity or density variations occur only gradually at length scales larger than the typical wavelength.

The identification of primaries is carried out on both the common-source and the common-receiver gather. We then check if any two (or more) primaries arrive close to simultaneously on the mutual trace that is common to both gathers. If this is the case, a new trace is picked and the process is repeated until a suitable trace is found for which all primary arrivals are separated by at least the length of a wavelet. This reduces ambiguity about which common-receiver and common-source move-outs correspond to the same event. The mutual trace is chosen quasi randomly from different sub-sets (“bins”) of the data cube, each containing 25 sources and 25 receivers in the examples herein. By testing traces from different bins we make sure that a variety of illumination angles are considered, which accelerates the search for a suitable mutual trace.

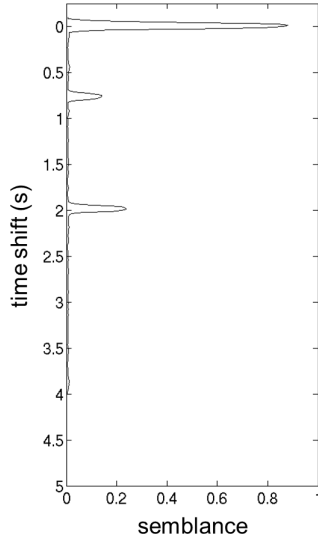


Figure 3.6: Result of semblance analysis ( $S(t_{shift})$  in Eq. 3.6) for a common-source gather (Fig. 3.4a) using the travel-time curve shown in Fig. 3.5d at different time shifts. Time shifts are relative to the arrival time of the primary, represented by the peak at 0s. The two later peaks indicate arrivals of multiples with the same move-out and hence the same last scatterer.

### 3.3.2 Detecting multiples using semblance analysis

The travel-time curves extracted from the cross-gathers are used to detect multiples with the same fingerprint, hence the same last (first) diffractor, arriving at later times in the common-source (common-receiver) gather. This is achieved using semblance analysis, a technique commonly used in seismic velocity analysis (Thorson and Claerbout, 1985; Sheriff and Geldart, 1995). The semblance denotes the ratio of the total energy of the stack (sum) of traces along a travel-time curve to the sum of the energy of the individual traces (Sheriff and Geldart, 1995). In our case the shape of the travel-time curve is already known and the only unknown parameter is a constant time shift  $t_{shift}$  between a primary and a multiple. The primary travel-time curve  $t(x)$  is shifted across the gather by  $t_{shift}$ , and the semblance  $S$  along the curve  $t(x) + t_{shift}$  is computed according to

$$S(t_{shift}) = \frac{1}{N} \frac{\sum_{t_w = -\frac{m}{2}\Delta}^{\frac{m}{2}\Delta} [\sum_{x=1}^N g(t(x) + t_{shift} + t_w; x)]^2}{\sum_{t_w = -\frac{m}{2}\Delta}^{\frac{m}{2}\Delta} \sum_{x=1}^N [g(x) + t_{shift} + t_w; x]^2} \quad (3.6)$$

where  $N$  is the number of traces, and  $g(t(x) + t_{shift} + t_w; x)$  is the value of trace  $x$  at time  $t(x) + t_{shift} + t_w$ , where  $t_w$  runs over the temporal length of the semblance window according to the outer summation in Eq. 3.6. Stacking over a time

window of width  $m\Delta$ , where  $m$  is a positive scalar and  $\Delta$  is the inverse of the temporal sampling rate, accounts for the fact that a coherent arrival on each trace extends over a finite time interval, namely roughly the length of the source wavelet. Due to the normalization factor  $1/N$  values for  $S$  range between 0 and 1. When  $S$  is close to 1, all amplitudes sum coherently and the proposed travel-time curve  $t(x) + t_{shift}$  fits a move-out that exists in the gather. The time shifts  $t_{shift}$  for which  $S$  is large thus provide the arrival times of multiples with the same last (or first) scatterer as the primary described by the travel-time curve  $t(x)$  (i.e., all arrival times are relative to the corresponding primary arrival). An example is shown in Fig. 3.6.

In order to increase the amplitudes of weak arrivals (especially secondaries) and equalize amplitudes across the whole move-out, a temporally adaptive gain was applied to the data before performing semblance analysis. This maintains the relative amplitudes of all arrivals in a specified time window (gain window) but normalizes the maxima of all time windows. If the gain window is sufficiently small, this process sets all arrivals on a trace to the same amplitude including primaries and multiples, but also noise. For the semblance analysis, however, the amplitude of the noise is not important, assuming that it is incoherent in phase and therefore should not sum constructively. In our examples a gain window equivalent to the length of the wavelet provided the best results for noisy data.

Having identified the first diffractor of each event on the common-receiver gather, and the last diffractor of each event on the common-source gather, we combine these pieces of information on the mutual trace, which allows us to identify all secondary events: these are the first arriving events associated with each pair of different common-source and common-receiver move-out.

### 3.3.3 Predicting travel times of higher-order multiples

Once all primary and secondary events have been identified they can be used to construct any tertiary multiply scattered arrival by combining three events:

- (1) any primary (e.g. Fig. 3.3a)
- (2) a secondary associated with the same last diffractor as the primary (Fig. 3.3b)

- (3) a secondary associated with the same first diffractor as the primary (Fig. 3.3d)

The travel times of the two secondaries are added, and the travel time of the primary is subtracted, which yields the travel time of a tertiary arrival that should be able to be observed on the trace. The events that were used in the construction process also define the scattering path of the tertiary arrival: this is the concatenation of the scattering paths of the two secondaries (minus one instance of the common scatterer, Fig. 3.3e and f). We can go through this example using the vector scattering paths of three events travelling between a source at  $\mathbf{x}$  and a receiver at  $\mathbf{x}'$  (Fig. 3.3):

- (1) a primary scattered at diffractor  $\mathbf{x}_B$  (scattering path  $\overline{\mathbf{x}\mathbf{x}_B} + \overline{\mathbf{x}_B\mathbf{x}'}$ ),
- (2) a secondary scattered at diffractors  $\mathbf{x}_A$  and  $\mathbf{x}_B$  (scattering path  $\overline{\mathbf{x}\mathbf{x}_A} + \overline{\mathbf{x}_A\mathbf{x}_B} + \overline{\mathbf{x}_B\mathbf{x}'}$ ),
- (3) and a secondary scattered at diffractors  $\mathbf{x}_B$  and  $\mathbf{x}_C$  (scattering path  $\overline{\mathbf{x}\mathbf{x}_B} + \overline{\mathbf{x}_B\mathbf{x}_C} + \overline{\mathbf{x}_C\mathbf{x}'}$ )

Adding and subtracting the travel times as described informally then relates to adding and subtracting vector scattering path components:

$$\begin{aligned} & (\overline{\mathbf{x}\mathbf{x}_A} + \overline{\mathbf{x}_A\mathbf{x}_B} + \overline{\mathbf{x}_B\mathbf{x}'}) + (\overline{\mathbf{x}\mathbf{x}_B} + \overline{\mathbf{x}_B\mathbf{x}_C} + \overline{\mathbf{x}_C\mathbf{x}'}) - (\overline{\mathbf{x}\mathbf{x}_B} + \overline{\mathbf{x}_B\mathbf{x}'}) \\ & = \overline{\mathbf{x}\mathbf{x}_A} + \overline{\mathbf{x}_A\mathbf{x}_B} + \overline{\mathbf{x}_B\mathbf{x}_C} + \overline{\mathbf{x}_C\mathbf{x}'} \end{aligned} \quad (3.7)$$

Hence, the predicted tertiary has encountered the three diffractors in the sequential order  $\mathbf{x}_A$ ,  $\mathbf{x}_B$ , and  $\mathbf{x}_C$ .

Subsequently, the travel times of higher-order multiples can be computed by replacing one of the secondaries with a tertiary in the construction process. This process allows the travel times of multiply scattered waves of any order to be predicted, each being verified by direct observation of that multiple's move-out in the data. We will show that the time of arrivals can be predicted correctly even if the arrivals themselves are buried in a complex noise field.

Fig. 3.7 displays the mutual trace of the 3-diffractors example obtained from the model in Fig. 3.1 with white noise added with a noise-to-signal ratio of 0.5 (where a ratio of 1 means the root-mean square noise is equivalent to the typical amplitude of a secondary arrival): here, all primary and secondary arrivals have been identified correctly by the algorithm (see Tables 3.1 and 3.2) and the

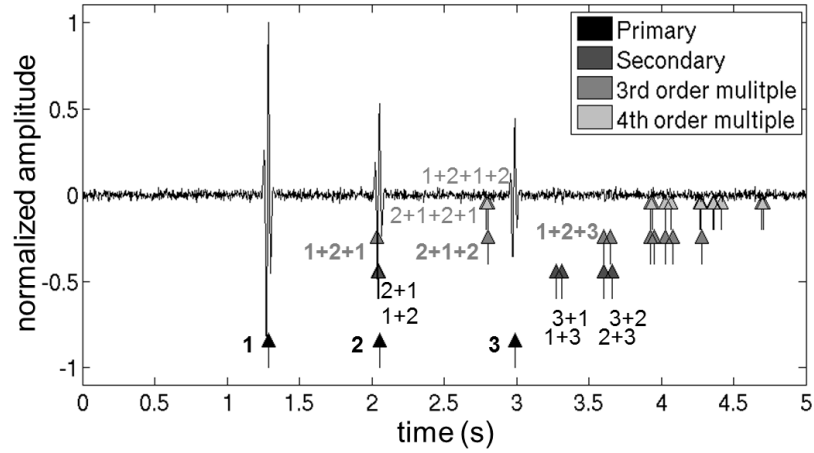
Primary	$t_{est}$	$t_{true}$	$ t_{est} - t_{true} $
1	1.29 s	1.28 s	0.01 s
2	2.05 s	2.05 s	0.00 s
3	2.99 s	2.98 s	0.01 s

Table 3.1: Comparison of estimated ( $t_{est}$ ) and true travel times ( $t_{true}$ ) of primary arrivals on the mutual trace in the 3-diffractors example (Fig. 3.7). All estimated travel times lie within the permitted deviation of half the length of a wavelet (0.05 s).

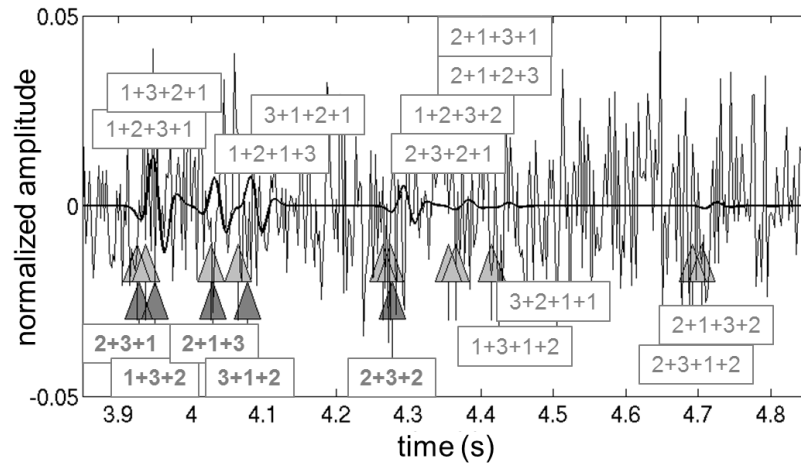
Secondary	$t_{est}$	$t_{true}$	$ t_{est} - t_{true} $
2+1	2.04 s	2.05 s	0.01 s
1+2	2.05 s	2.05 s	0.00 s
1+3	3.28 s	3.27 s	0.00 s
3+1	3.31 s	3.32 s	0.01 s
2+3	3.60 s	3.61 s	0.01 s
3+2	3.66 s	3.66 s	0.00 s

Table 3.2: Comparison of estimated ( $t_{est}$ ) and true travel times ( $t_{true}$ ) of secondary arrivals on the mutual trace in the 3-diffractors example (Fig. 3.7). All estimated travel times lie within the permitted deviation of half the length of a wavelet (0.05 s).

resulting prediction for third and fourth order multiples are also marked in the figure. The computing time required to analyse one such trace in this example is of the order of one minute on a standard desktop computer, though finding a suitable mutual trace in the data cube requires additional time (up to one minute for each trace that is tested). For increasing numbers of primary move-outs (i.e., numbers of diffractors in the model) the computing time also increases.



(a)



(b)

Figure 3.7: (a) Mutual trace of the common-source and common-receiver gather shown in Fig. 3.2. The maximum amplitude has been normalized to one. Arrows mark the arrivals of estimated primary, secondary, third order, and fourth order events. In this example the algorithm detected all primaries and secondaries correctly (Tables 3.1 and 3.2), i.e., the deviation from the true arrival time was below half the length of a wavelet (0.05 s); hence, all higher-order events are estimated within permitted accuracy. Numbers indicate the scattering path of individual events (up to 3.6 s), e.g.,  $\mathbf{3} + \mathbf{1}$  means a secondary scattered first at diffractor  $\mathbf{3}$  and then diffractor  $\mathbf{1}$ . (b) Zoom on the trace in (a) between 3.85 s and 4.85 s superimposed by the noise-free trace (bold) amplified by a gain of four to highlight small amplitude third and fourth order scattered arrivals embedded in the noise but predicted correctly by the algorithm.



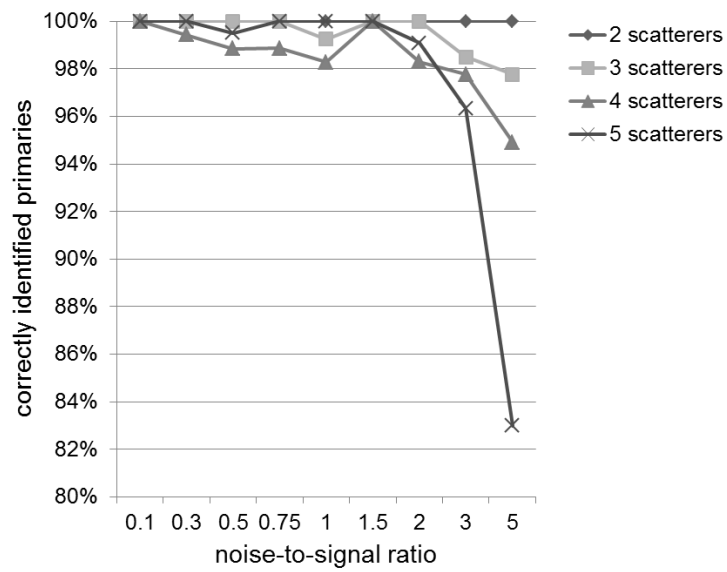
### 3.4 Numerical examples

We test the automated algorithm on a range of synthetic data sets obtained from different two-dimensional acoustic models containing multiple isotropic point scatterers embedded in a homogeneous background medium (density  $\rho = 1000 \text{ kg/m}^3$ , velocity  $v = 1000 \text{ m/s}$ ). A variety of Foldy's method embodied within a freely available wavefield modelling code (Foldy, 1945; Galetti et al., 2013) is used to compute the diffracted wavefields including all orders of multiple scattering. Direct wave arrivals are not modelled since they have no interactions with the diffractors. In our example, source and receiver arrays consist of 101 sources and 101 receivers, respectively, have the same lateral extension and the same horizontal spacing of 20 m (however, this is not a requirement of the method).

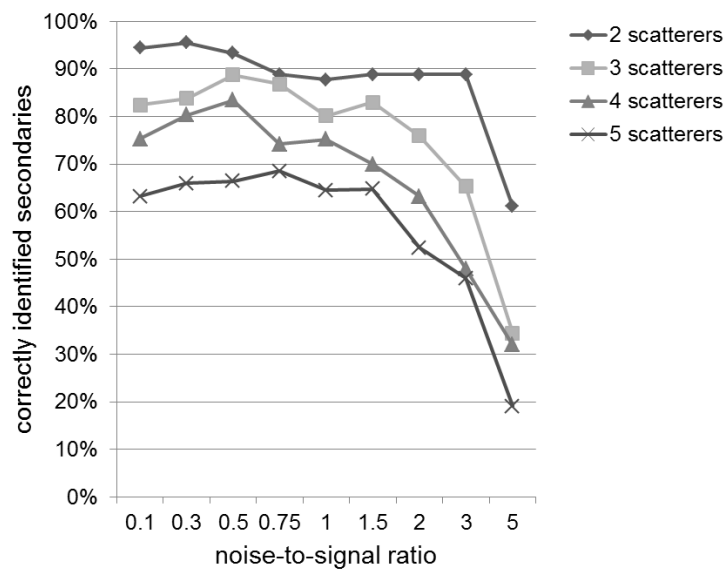
Our aim is to demonstrate the performance of the algorithm in detecting primary and secondary arrivals correctly for (i) models containing varying numbers of diffractors, and (ii) different noise-to-signal ratios. In our algorithm the prediction of higher-order multiples depends solely on the correct identification of primaries and secondaries. The number of diffractors in the model (between two and five) is pre-defined in each example; the locations of the diffractors are chosen randomly within a box of  $2000 \text{ m} \times 1000 \text{ m}$  beneath the source array and the receiver array (roughly the area shown in Fig. 3.1), such that the apexes of all move-outs are well-defined by the recorded data. The only other constraint is that the distance between any two diffractors is larger than twice the typical wavelength  $\lambda \approx 33 \text{ m}$  so that each produces identifiable scattered wave energy.

For each set number of diffractors, we test the algorithm on 15 different models, all containing the same number of diffractors but in different random locations. For each model we count the number of secondaries detected by the algorithm, and determine how many of them are predicted correctly (at the correct travel time on the mutual trace) and how many are incorrect. For comparison we use the true travel times of primary and secondary waves computed directly from the model. An event is classified as incorrect if the estimated travel time deviates by more than half the length of a wavelet from the true travel time. To test the influence of the choice of the mutual trace, we examine three different such traces for each model.

Next, we contaminate each data set with increasing noise-to-signal ratios (between



(a)



(b)

Figure 3.8: Percentage of (a) primaries and (b) secondaries identified correctly (i.e., with the correct travel time, to within an accuracy limit of half of the source wavelet length) for different numbers of diffractors in the model and varying noise-to-signal ratios.

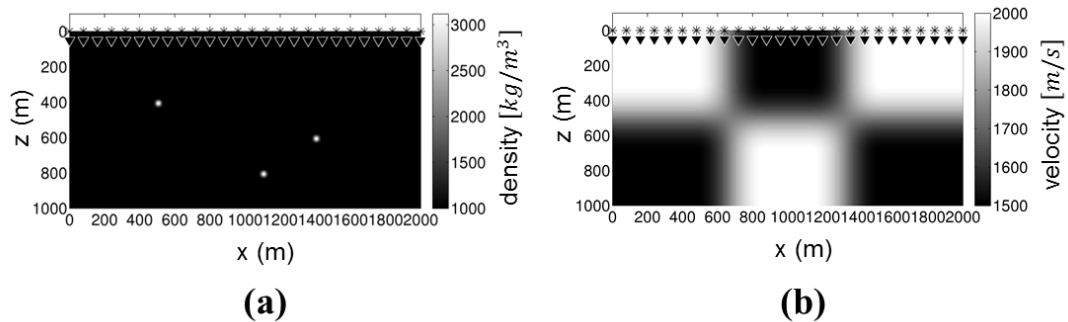


Figure 3.9: **(a)** Density and **(b)** velocity distribution in a 3-diffractors model with a heterogeneous background structure. Symbol key as in Fig. 3.1. Only every fourth receiver and source is plotted for clarity. Scatterers are modelled as sharp density contrasts in a homogeneous background, whereas smooth checkerboard-like heterogeneities comprise the non-diffracting velocity distribution.

0.1 and 5, where a ratio of 1 means that the root-mean square noise is equal to the typical amplitude of a secondary), and for each we again count the number of primaries and secondaries detected correctly and those detected incorrectly. This process is repeated for different numbers of diffractors and the results tabulated. A summary for primary and secondary events is given in Fig. 3.8, since travel times of all subsequent multiples depend only on these results as shown above and in Meles and Curtis (2014a).

To demonstrate that neither the isolation of primaries nor the detection of secondaries is restricted to media with homogeneous background velocities and hence data with hyperbolic move-outs, we include an illustrative example for data that exhibit clearly non-hyperbolic move-outs. This was created using a finite-difference modelling code with absorbing boundaries and a Ricker wavelet with central frequency of 30 Hz. Three randomly placed diffractors are represented as locations of high density contrast ( $3000 \text{ kg/m}^3$  compared to  $1000 \text{ kg/m}^3$  in the background) and the velocity varies according to a smooth checkerboard pattern between  $1500 \text{ m/s}$  and  $2000 \text{ m/s}$  (Fig. 3.9). The acquisition geometry is identical to that in previous examples and no noise is added to the data. Fig. 3.10 shows a typical common-source gather from this experiment and the identified primary move-outs. Primary and secondary travel times as determined by our algorithm (Fig. 3.11) are compared to the true travel times calculated from the finite-difference model in Tables 3.3 and 3.4.

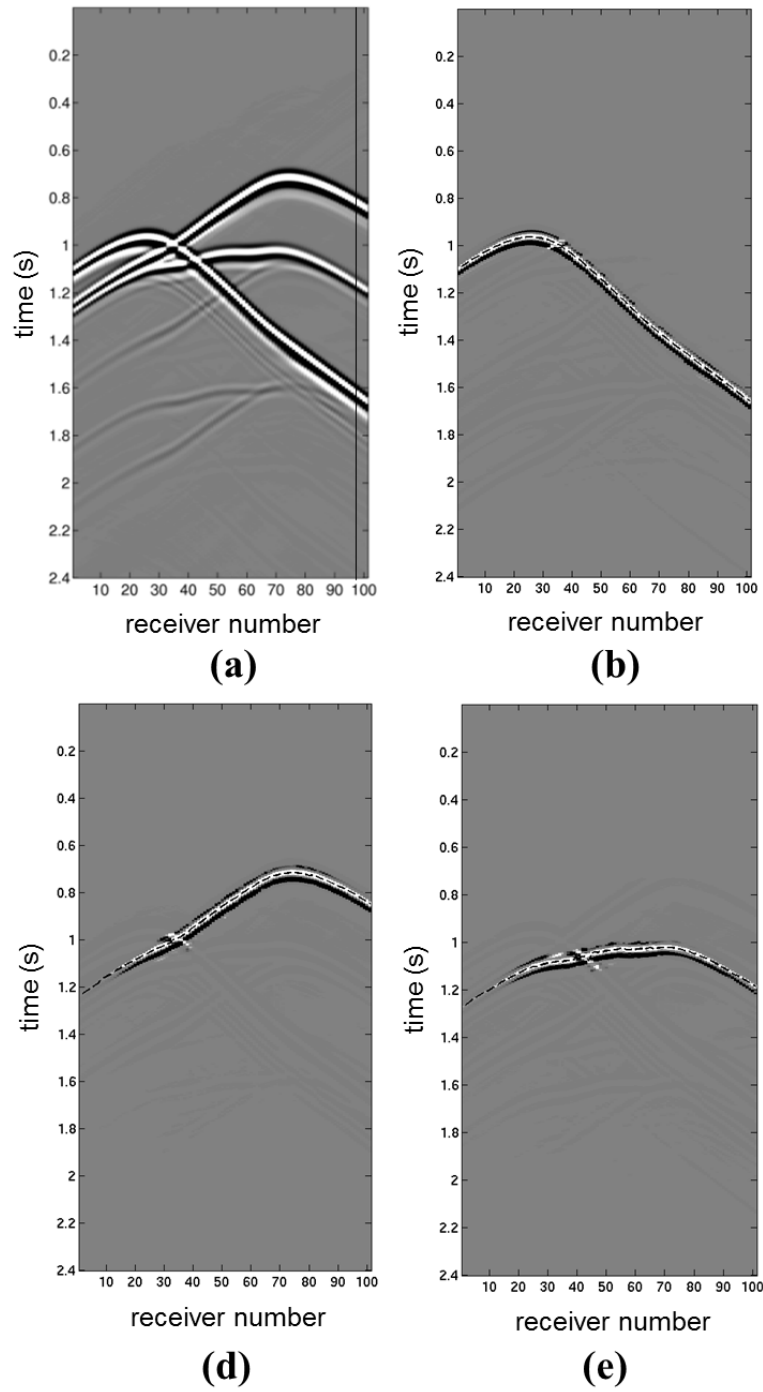


Figure 3.10: **(a)** Common-source gather obtained from the model in Fig. 3.9 with the common source located at [600 m, 0 m]. The vertical bold line marks the mutual trace. **(b)**, **(c)** and **(d)**: Cross-gathers showing primary move-outs extracted from the common-source gather in **(a)**. Estimated travel-time curves are superimposed as dashed lines.

### 3.5 Discussion

The identification of primaries by cross-correlation of two gathers is successfully carried out in over 90% of the cases even for several diffractors and relatively high noise-to-signal ratios (Fig. 3.8a). One point of weakness in the algorithm is that a suitable detection threshold must be found that discriminates cross-correlations of primaries from those of multiply diffracted waves, based on the amplitude of the cross-correlation coefficient (e.g.,  $\Phi$  in Fig. 3.4c). If diffractors are distributed widely across the medium, the amplitudes of primary diffractions can vary significantly (the order of a factor of ten in our study). The threshold needs to be set sufficiently low to enable detection of the weaker primaries. On the other hand, if diffractors are located close together, multiply diffracted waves can have relatively large amplitudes, and the threshold needs to be high enough so as not to mistake them for primaries. However, if the move-out of a multiply diffracted event is erroneously identified as a primary, it must in fact be identical to one of the primary move-outs but with a larger absolute arrival time. Comparing the move-outs thus allows one to reject those related to multiply diffracted waves that are simply later repetitions of other detected earlier-arriving move-outs. This can therefore be used as a criterion to set the detection threshold automatically: it should be chosen low enough that setting it to lower values yields no new and distinct move-outs, and large enough that no multiples are mistaken for primaries.

To our knowledge, the only other techniques that identify and isolate individual diffraction move-outs in a multiply scattering medium are methods based on time-reversal such as those proposed by Prada and Fink (1994) and Montaldo et al. (2004). These methods, however, require a transducer array (collocated sources and receivers) that records and, in the iterative process described by Montaldo et al. (2004), also re-injects the scattered wavefield. In our method, while to predict the full scattering path both a source and a receiver array is needed (which do not have to be collocated), primary move-outs can in principle be identified and isolated using only a receiver array that records the response from two separate sources. More sources can be used to improve the identified travel-time curves, but a densely sampled source array is not required.

The crucial step in the algorithm is the identification of secondaries. This step becomes more challenging the more diffractors are involved. The number of events that have been scattered  $m$  times in a medium containing  $n$  diffractors

Primary	$t_{est}$	$t_{true}$	$ t_{est} - t_{true} $
1	0.82 s	0.82 s	0.00 s
2	1.16 s	1.16 s	0.00 s
3	1.63 s	1.62 s	0.01 s

Table 3.3: Comparison of estimated ( $t_{est}$ ) and true travel times ( $t_{true}$ ) of primary arrivals on the mutual trace in the 3-diffractors example with a heterogeneous background velocity structure (Fig. 3.11). All estimated travel times lie within the permitted deviation of half the length of a wavelet (0.072 s).

Secondary	$t_{est}$	$t_{true}$	$ t_{est} - t_{true} $
2+1	1.18 s	1.22 s	0.03 s
1+2	1.20 s	1.24 s	0.04 s
3+1	1.75 s	1.74 s	0.02 s
3+2	1.75 s	1.79 s	0.04 s
1+3	1.78 s	1.82 s	0.04 s
2+3	1.79 s	1.85 s	0.05 s

Table 3.4: Comparison of estimated ( $t_{est}$ ) and true travel times ( $t_{true}$ ) of secondary arrivals on the mutual trace in the 3-diffractors example with a heterogeneous background velocity structure (Fig. 3.11). All estimated travel times lie within the permitted deviation of half the length of a wavelet (0.072 s).

is  $n \cdot (n - 1)^{(m-1)}$ , hence, we expect  $n$  primaries (since  $m = 1$ ) and  $n \cdot (n - 1)$  secondaries ( $m = 2$ ). This means that for  $n = 3$ , as in the examples shown earlier, the number of secondaries is only 6, but for  $n = 5$  the number of secondaries is already 20.

As an aside, note that in an elastic medium, where both compressional (P) waves and shear (S) waves propagate, waves are converted every time they scatter. In theory, the assumption of a unique fingerprint for each diffractor still holds, the difference being that the elastic fingerprint consists of two travel-time curves in each data gather, one for the P-wave primary and one for the S-wave primary. Thus, assuming that the source generates only P-waves, the number of primaries doubles while the number of secondaries becomes four times as large compared to the acoustic case. We have not tested the elastic case but assume that performance would deteriorate.

Fig. 3.8b shows that when the number of diffractors increases, the detected arrival times of secondaries are indeed increasingly error-prone. The sources of this error are discussed as follows. A prerequisite for the identification of secondary move-outs is the correct estimation of primary travel-time curves along which the

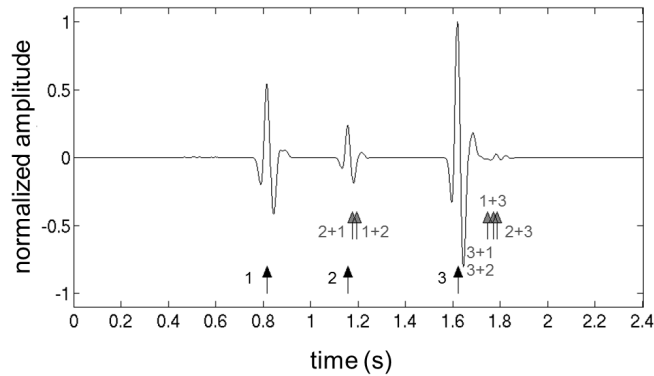


Figure 3.11: Mutual trace of the common-source and common-receiver gather obtained from the 3-diffractors example with heterogeneous background velocity structure (Figs. 3.9, bold vertical line in 3.10a). The maximum amplitude has been normalized to one. Arrows and numbers mark the arrivals of estimated primaries (black text) and secondaries (grey text). In this example the algorithm detected all primaries and secondaries correctly (Tables 3.3 and 3.4), i.e., the deviation from the true arrival time was below half the length of a wavelet ( $0.072s$ ).

semblance analysis is performed. If the extracted curve deviates from the true travel-time curve, computed semblance values are lower and may fall below the threshold for secondary detection. The results for primaries (Fig. 3.8a) suggest that almost all primaries are identified correctly on the mutual trace. Note, however, that the estimated travel time was compared to the true travel time on the mutual trace only, which does not mean that the travel times are correct everywhere along the estimated curve. Errors in primary travel-time curves can be produced by an unfortunate combination of gathers, in which the travel-time shift of a move-out is not unique, i.e., two or more move-outs experience the same travel-time shift due to the spatial shift of the common source or the common receiver. This is more likely to happen the more primaries there are in the data, hence, the more diffractors there are in the medium.

The accuracy with which travel times of tertiaries and higher-order scattered waves are predicted depends on the errors in the estimated primary and secondary travel times. Since these errors propagate through the algorithm, travel-time prediction gets less accurate with increasing scattering order: let  $F_{t(B)}$  be the absolute error (uncertainty) of the travel time of primary  $B$ ,  $F_{t(AB)}$  the absolute error of the travel time of secondary  $AB$ , and  $F_{t(BC)}$  the absolute error of the travel time of secondary  $BC$ , then the absolute error of the predicted tertiary  $ABC$  is given by  $F_{t(ABC)} = F_{t(AB)} + F_{t(B)} + F_{t(BC)}$ . Equivalently, the errors of tertiaries contribute to the errors of fourth order scattered events and so on. In this study the prediction was limited to fourth order scattering where most events

were still recognizable as individual, distinct arrivals. Just as in the algorithm of Meles and Curtis (2014a), this means that in principle the propagation of error could be corrected at each iteration if a clearly identifiable arrival on the trace can be associated with the predicted higher-order scattering event: the predicted arrival time may simply be replaced by the observed arrival time. This assumes that the initial errors in primaries and secondaries are sufficiently small that the correct higher-order event is identified.

As we move to higher and higher-order scattering, more events become superimposed or merged into a continuous coda, and the predicted travel times can no longer be related to individual arrivals. Also, for most applications (e.g., diffraction imaging) considering low-order scattering will be sufficient, especially as our method demonstrates explicitly that theoretically all kinematic information is contained in primaries and secondaries.

In models with three or more diffractors the superposition of primaries and secondaries on the mutual trace can lead to ambiguity and hence misinterpretation. When, for example, primary  $A$  coincides with secondary  $B + C$ , two different move-outs intersect in each gather: move-outs  $A$  and  $B$  in the common-receiver gather (fingerprints of the first diffractors), and move-outs  $A$  and  $C$  in the common-source gather (fingerprints of the last diffractors). After the algorithm has detected all move-outs, it then identifies four possible combinations of first and last fingerprints:  $A + A$ , which corresponds to primary  $A$ ;  $B + C$ , which corresponds to the true secondary; and two false events, namely the secondaries  $A + C$  and  $B + A$ . This ambiguity is more likely to occur the more diffractors there are in the model, and hence the more fingerprints overlap in the data gathers. Note that the superposition of primary arrivals, which is another potential source of misinterpretation, is avoided automatically by the algorithm through the appropriate selection of the gathers to be analysed.

Fig. 3.8b also shows that, as we would expect, the secondary detection in general deteriorates when the noise-to-signal ratio is increased. A considerable drop is observed for noise-to-signal ratios of 3 and higher, which means that the average noise amplitude is at least three times as large as the typical amplitude of a secondary arrival. We interpret this as the point where the coherency of the multiply scattered arrivals is completely extinguished by the incoherent noise field. Before this point, larger noise levels may affect the accuracy of the extracted primary travel-time curves, which results in minor errors in the semblance analysis. Also, the threshold used to distinguish multiply scattered waves from background



noise (defined as the mean of the semblance vector) may not be suitable for all noise-to-signal ratios.

Fig. 3.8 implies that in some cases the algorithm returns better results when the noise level is higher. This unexpected pattern of performance (improvement with noisier data) can be explained by a problem that mainly affects low noise data, where move-outs of higher-order multiples may also be observed in the data with similar amplitudes to secondaries. Superposition of these arrivals with primary or secondary arrivals on the mutual trace causes similar ambiguities as described above. For noisier data these higher-order multiples are not detected due to their relatively small amplitudes compared to the noise level and the source of error is avoided, which results in more accurate detection of secondaries.

Inherent ambiguities due to the simultaneous arrival of different events cannot be solved on the basis of a single trace. However, we have the option to consult new (perhaps neighboring) mutual traces in the data cube, analyse the corresponding common-receiver and common-source gather, and compare the results. In this case source-receiver interferometry (SRI) operations can be used to redatum the estimated arrival times to the old source and receiver positions (Curtis and Halliday, 2010; Halliday and Curtis, 2010; Curtis et al., 2012) and thus to compare arrivals on two different traces directly. If the results are inconsistent, more traces must be analysed in order to decide between them on a statistical basis. Experimental results suggest that the choice of the trace has a considerable impact on results, which can in fact be more significant than the effect of the noise level.

The problems described above occur especially for larger numbers of diffractors since with increasing complexity of the wavefield the identification of secondaries becomes more difficult. Nevertheless, for a small number of strong diffractors we have demonstrated that a development of the method by Meles and Curtis (2014a) can be automated, and have successfully applied it to synthetic acoustic noisy data. Future research should test its applicability to real data sets, for example from an acoustic laboratory experiment or field data. The method could eventually find application in improved diffractor localization by using multiples, or to provide new information about inter-diffractor paths. One important transferable learning that both Meles and Curtis (2014a) and this work demonstrate is that in theory all travel time information in the multiply diffracted wavefield is in fact included in only the primary and secondary arrivals. This implies, for example, that since the travel times of higher order scattered waves are explicitly

related to order one and two waves, observations of higher order travel times might be used to improve estimates of lower order travel times.

Moreover, the identification and isolation of primary events by computing cross gathers and evaluating the energy distribution in these gathers could be a useful method on its own right. Since it is entirely non-parametric (and can therefore identify entirely non-hyperbolic move-outs) it can also be applied to more generally heterogeneous velocity structures, as demonstrated in Fig. 3.10. This may be of interest for a variety of data processing operations, for example in travel-time tomography or event-based filtering.

Finally, while all of our tests have been in 2D models, there is an obvious extension of the method to 3D using planes rather than lines of receivers by simply allowing parameter  $n$  in the above equations to index over the plane rather than only a line. In fact, this may improve rather than diminish performance for any particular number of diffractors, because there would usually then be far more data to discriminate between different move-outs, and because planes of receivers allow move-outs to be discriminated in two spatial directions rather than only one. Thus the method might be usefully applied in applications where one can only access one side of a  $3D$  medium.

# Chapter 4

## Automatic identification of diffracted waves in acoustic laboratory data

Chapter 3 introduced an automatic algorithm that isolates primary move-outs from a multiply diffracted wavefield and predicts travel times of higher-order multiply diffracted waves. It was demonstrated that this algorithm is applicable to a range of different synthetic data sets including different numbers of point diffractors, different noise-to-signal ratios and different (heterogeneous) background media. In this chapter, I test the algorithm on real data sets obtained from an acoustic laboratory experiment, which I conducted at Joseph Fourier University in Grenoble, France, in cooperation with Philippe Roux. I describe the numerical simulations performed prior to designing the laboratory experiment, provide details on the experimental setup and the data acquisition, comment on artefacts in the data and interpret the results provided by the automatic algorithm.

### 4.1 Acoustic finite-difference modeling tests

Before designing the final experiment in the laboratory, we run numerical simulations of different experimental configurations using a two-dimensional acoustic finite-difference (FD) modeling code (provided by Matteo Ravasi, PhD student, University of Edinburgh). In the numerical model, we vary the size and the properties (velocity and density) of the scatterers in order to simulate different materials, and evaluate the diffraction properties for the different cases. This allows us to define the type of scatterers suitable for the real experiment, which should be more or less isotropic and should not produce internal scattering or ringing that interferes with or superimposes multiply scattered waves in the data. Internal ringing refers to the reverberation of waves *within* the scatterer. This generates a ringing coda after the first arriving scattered wave. Anisotropy in the context of scattering means that the scattering amplitude depends on the

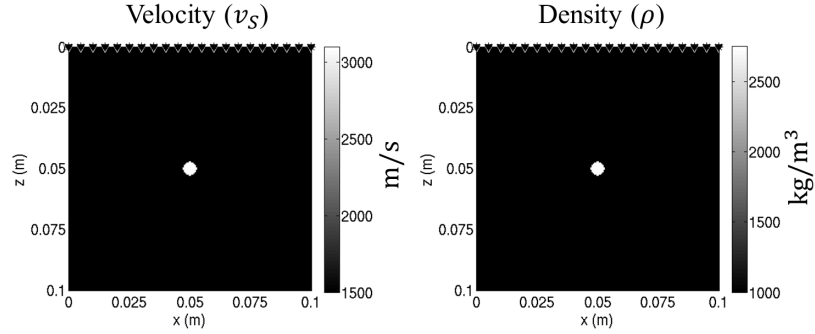


Figure 4.1: Example geometry for the velocity model (left) and the density model (right) used in an acoustic finite-difference modeling code with a scatterer of diameter  $d = 6$  mm, velocity  $v_S^{scatt} = 3100$  m/s and density  $\rho^{scatt} = 2700$  kg/m<sup>3</sup> (values of  $v_S^{scatt}$  and  $\rho^{scatt}$  correspond to aluminium). We chose to use the S-wave velocity of aluminium to better capture the effects of converted waves in the acoustic model. Velocity and density of the background correspond to water ( $v_P^0 = 1500$  m/s and  $\rho^0 = 1000$  kg/m<sup>3</sup>, respectively).

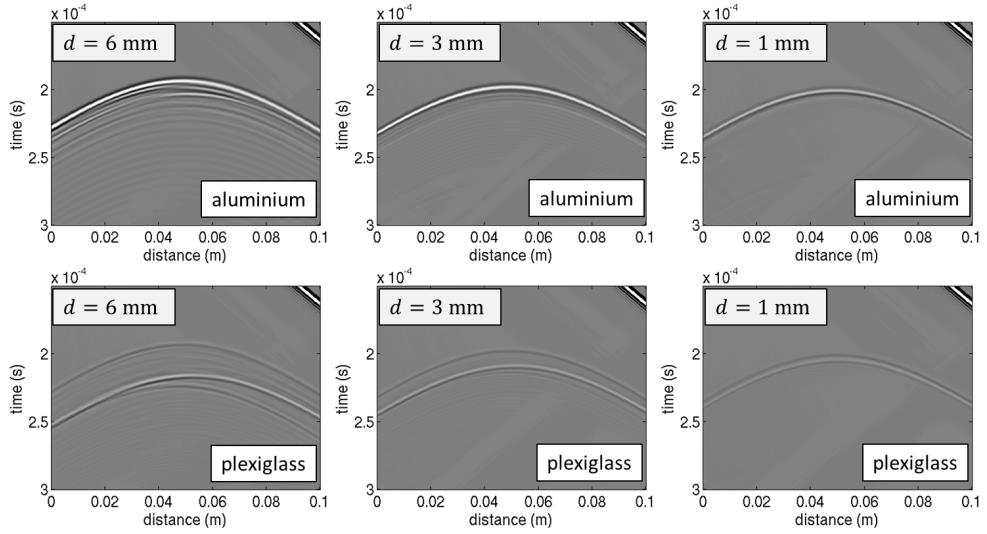


Figure 4.2: Results from the single scattering modeling experiment testing scatterers with different diameters and with different velocities and densities corresponding to aluminium ( $v_S^{alu} = 3100$  m/s and  $\rho^{alu} = 2700$  kg/m<sup>3</sup>) and plexiglass ( $v_S^{plexi} = 1100$  m/s and  $\rho^{plexi} = 1150$  kg/m<sup>3</sup>).

Material	$\rho$	$v_P$	$v_S$
Aluminium	2700 kg/m <sup>3</sup>	6300 m/s	3100 m/s
Plexiglass	1150 kg/m <sup>3</sup>	2700 m/s	1100 m/s
Teflon	2150 kg/m <sup>3</sup>	1400 m/s	-
Neoprene	1300 kg/m <sup>3</sup>	1600 m/s	-
Glass	3600 kg/m <sup>3</sup>	5300 m/s	3000 m/s

Table 4.1: Velocities  $v_P$  and  $v_S$  and densities  $\rho$  of materials tested in numerical scattering experiments. Where applicable, we use the S-wave velocities to better capture the effects of converted waves in the acoustic model.

direction of incoming and outgoing energy at the scatterer. This property has a negative influence when estimating primaries by cross-correlation and will be discussed in more detail later in this section and in section 4.5.

The numerical model is scaled in the order of the geometry envisaged for the real experiment with a modeling space of  $0.1 \text{ m} \times 0.1 \text{ m}$  (Fig. 4.1). The background medium has velocity  $v_S^{water} = 1500 \text{ m/s}$  and density  $\rho^{water} = 1000 \text{ kg/m}^3$ , as corresponding to water. We model the scatterers as discretized circles from elements of size  $0.25 \text{ mm} \times 0.25 \text{ mm}$  and test different combinations of velocities and densities corresponding to different materials listed in Table 4.1. Note that, where applicable, we use S-wave velocities of scatterers to better capture the effects of converted waves (P-S and S-P) in the acoustic model. The central frequency of the excited wavefield is  $f_c = 500 \text{ kHz}$ , which gives a typical wavelength of  $\lambda = 3 \text{ mm}$  in water. We test scatterers with diameter  $\lambda < 6 \text{ mm}$ ,  $\lambda \approx 3 \text{ mm}$  and  $\lambda > 1 \text{ mm}$ .

We start with an experiment that contains a single scatterer and compare the diffraction curves produced by different parameter combinations. Some examples are shown in Fig. 4.2. Comparing the results from the 6 mm-scatterer for aluminium and plexiglass (left column in Fig. 4.2), it becomes clear that different materials generate different anisotropy patterns. For the case of plexiglass, one can clearly distinguish the reflection from the top and the bottom of the scatterer, while for aluminium these two events almost superimpose. Further, the amplitude of the top reflection appears much stronger for the aluminium rod. In both cases one can see a low amplitude ringing, which originates from waves being reflected multiple times between the bottom and the top of the scatterer. These features are weaker but still present when using a 3 mm-scatterer (central column). However, when reducing the diameter of the scatterer to 1 mm (right column), which is smaller than the typical wavelength of the signal, both cases

produce an isotropic diffraction curve and the internal ringing is eliminated. In fact, for a 1 mm-scatterer all materials tested (Table 4.1) produce an isotropic diffraction curve and no internal ringing. Fig. 4.2 shows that the amplitudes of the scattered wave differ for different materials. Note, however, that these amplitudes are likely to be incorrect since we did not consider the P-wave velocities of aluminium and plexiglass in the modeling. Nevertheless the results suggest that materials with strong velocity and density contrasts to the background medium (such as aluminium or glass) produce stronger amplitudes, which is crucial when multiply scattered waves are supposed to be generated.

Next, we model three circular scatterers with a diameter of 1 mm and the properties of aluminium, generate a synthetic data cube and apply the algorithm introduced in chapter 3 to estimate primary diffraction curves and identify secondary scattered waves in the data. Fig. 4.4 shows a common-source gather (CSG) and a common-receiver gather (CRG) from the data cube. Fig. 4.5 shows the estimated primary move-outs and Fig. 4.6 shows the mutual trace that is common to both gathers in Fig. 4.4 with true and estimated travel times of primaries and secondaries indicated by arrows. Both Fig. 4.5 and Fig. 4.4 confirm that primary travel times and diffraction curves are estimated correctly and that most secondaries are identified. Table 4.2 compares the estimated secondary travel times to the true travel times on the mutual trace, which have been computed from the (known) scatterer locations and the background velocity.

We repeat the numerical experiment using larger scatterers (diameter  $d = 6$  mm) to examine the influence of anisotropy and internal ringing on the performance of the algorithm (Figs. 4.7 and 4.8). Using the same CSG and CRG and the same parameters in the algorithm as before (namely the number of cross-correlations,  $n_c$ , and the threshold for primary detection, see chapter 3), only one out of three primaries is detected in the CSG and multiple spurious primaries are detected in the CRG. As a result, secondaries cannot be identified correctly and higher-order multiples cannot be predicted. We attribute these problems to the anisotropy of the diffraction curves and the presence of internal ringing that generates additional arrivals interfering with primary and secondary scattered waves. The method used to estimate primary move-outs (see section 3.3.1) compares these move-outs in different CSGs (or CRGs) using cross-correlation. It is based on the assumption that changing the position of the common source (or the common receiver) only affects the absolute travel times of primaries, so that identical diffraction curves can be found at different travel times in different gathers.

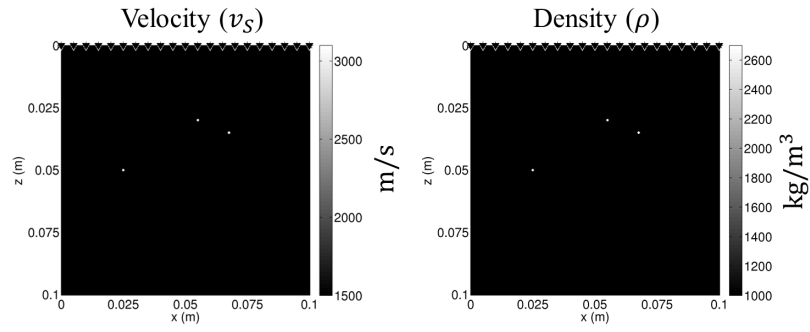


Figure 4.3: Velocity model (left) and density model (right) used in the 3-scatterers case with scatterers of diameter  $d = 1$  mm, velocity  $v_S^{scatt} = 3100$  m/s and density  $\rho^{scatt} = 2700$  kg/m<sup>3</sup> (values of  $v_S^{scatt}$  and  $\rho^{scatt}$  correspond to aluminium). Velocity and density of the background correspond to water ( $v_P^0 = 1500$  m/s and  $\rho^0 = 1000$  kg/m<sup>3</sup>, respectively).

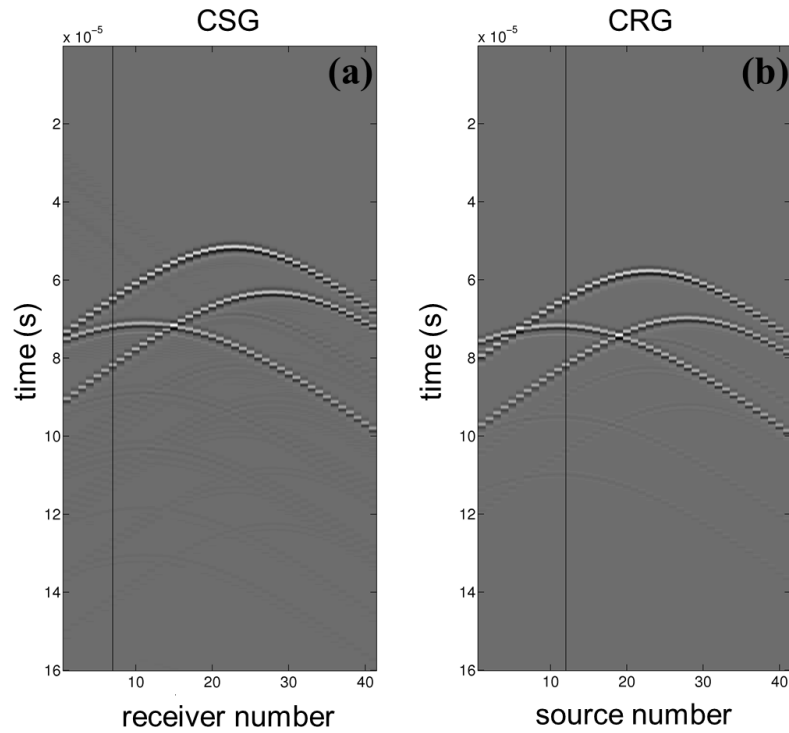


Figure 4.4: (a) Synthetic common-source gather (CSG) and (b) common-receiver gather (CRG) generated from the 3-scatterers model shown in Fig. 4.3.

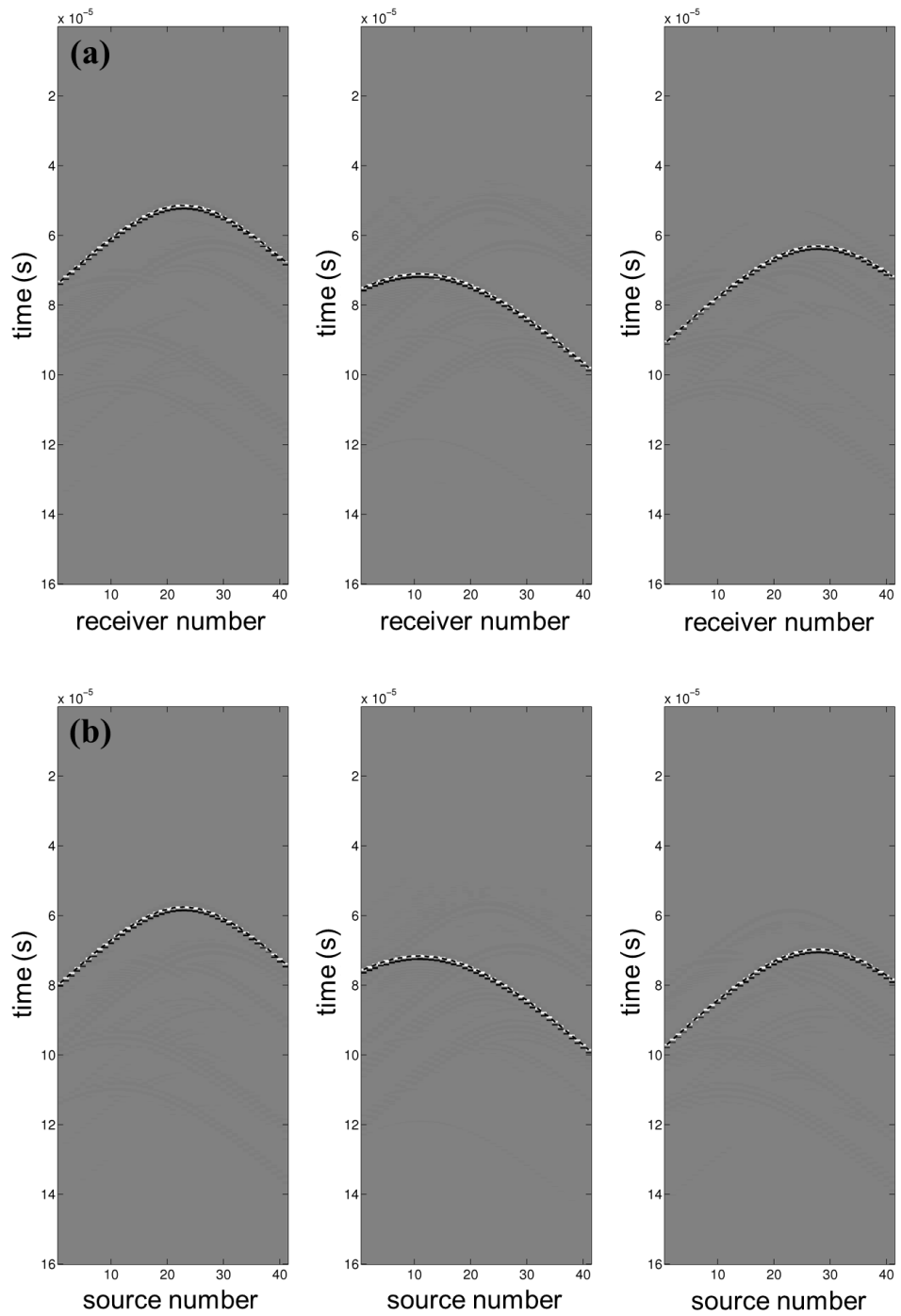


Figure 4.5: Left, central and right plots show estimated primaries (a) from the CSG in Fig. 4.3a and (b) from the CRG in Fig. 4.3b.



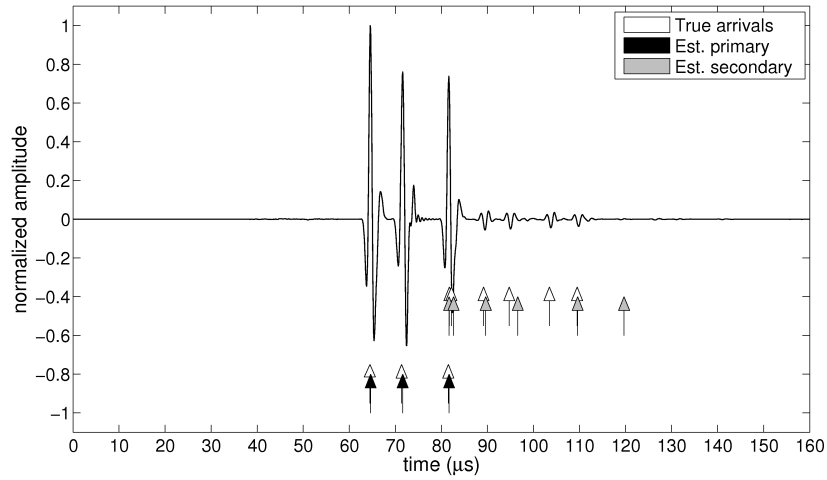


Figure 4.6: Estimated primaries and secondaries (black and grey arrows, respectively) on the mutual trace compared to the true arrivals (white arrows). All primaries and four out of six secondaries have been predicted correctly (see true and estimated travel times in Table 4.2).

Secondary	$t_{est}$	$t_{true}$	$ t_{est} - t_{true} $
3+1	81.7 $\mu s$	81.7 $\mu s$	0 $\mu s$
1+3	82.2 $\mu s$	82.6 $\mu s$	0.4 $\mu s$
1+2	89.2 $\mu s$	89.6 $\mu s$	0.4 $\mu s$
2+1	94.5 $\mu s$	96.6 $\mu s$	2.1 $\mu s$
3+2	103.5 $\mu s$	199.7 $\mu s$	16.2 $\mu s$
2+3	109.5 $\mu s$	109.6 $\mu s$	0.1 $\mu s$

Table 4.2: Comparison of estimated ( $t_{est}$ ) and true travel times ( $t_{true}$ ) of secondary arrivals on the mutual trace (Fig. 4.6). Most estimated travel times lie within the permitted deviation of half a wavelet ( $2 \mu s$ ), only secondaries 2 + 1 and 3 + 2 have not been identified correctly.

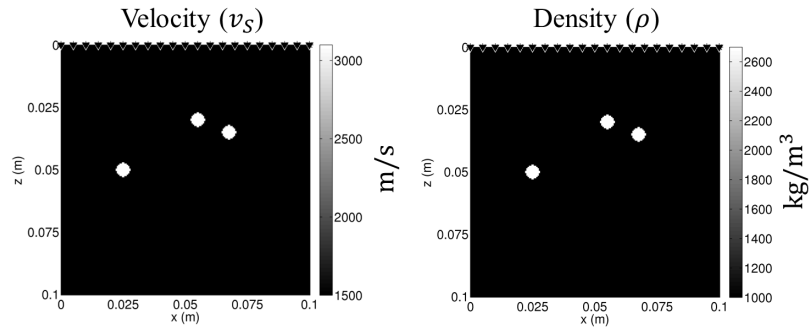


Figure 4.7: Velocity model (left) and a density model (right) used in the 3-scatterers case with scatterers of diameter  $d = 6$  mm, velocity  $v_S^{scatt} = 3100$  m/s and density  $\rho^{scatt} = 2700$  kg/m<sup>3</sup> (values of  $v_S^{scatt}$  and  $\rho^{scatt}$  correspond to aluminium). Velocity and density of the background correspond to water ( $v_P^0 = 1500$  m/s and  $\rho^0 = 1000$  kg/m<sup>3</sup>, respectively).

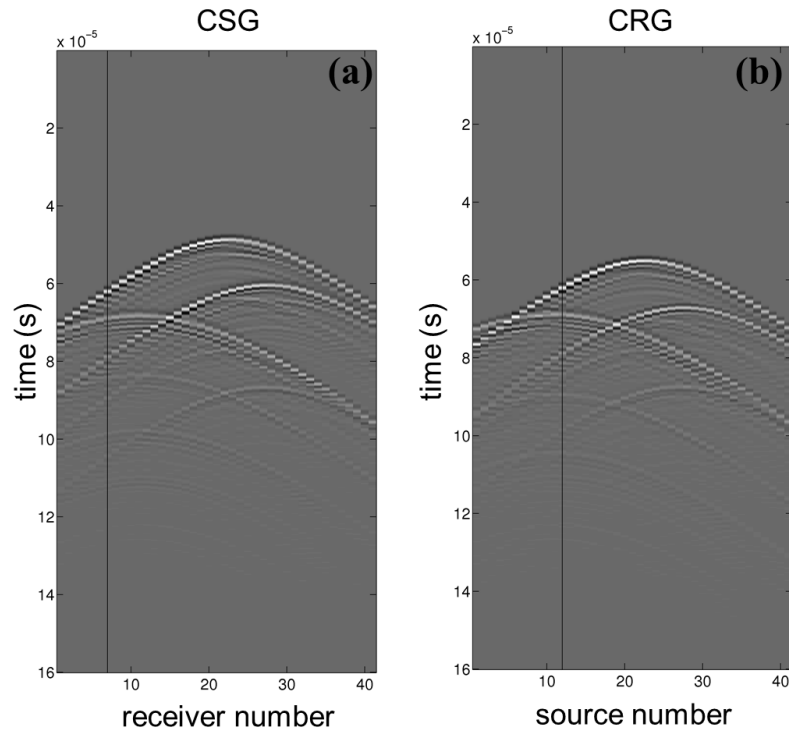


Figure 4.8: (a) Synthetic common-source gather (CSG) and (b) common-receiver gather (CRG) generated from the 3-scatterers model shown in Fig. 4.7.

Cross-correlating two different gathers provides the time shifts between identical diffraction curves that can be used to isolate individual move-outs from the total wavefield. If a scatterer is anisotropic, however, changing the location of the source (the receiver) not only affects the absolute travel times but also changes the wavelets of the primary arrivals on each trace (because the scattering amplitude depends on the angles of incoming and outgoing energy at the scatterer). The cross-correlation of two different CSGs (CRGs) will thus have lower amplitudes compared to the isotropic case and the time shifts between two primary move-outs are more difficult to detect. What is more, spurious correlations of additional arrivals introduced by internal ringing generate noise in the correlation result and make it even more difficult to detect the correlation of actual primaries. Also the detection of secondaries using semblance analysis deteriorates for anisotropic diffraction curves, since the wavelets along the move-out in this case do not necessarily sum constructively.

Based on the numerical simulations, we choose rods with a diameter of  $d = 1$  mm as point scatterers in our real experiment. We used rods made from stainless steel as they were readily available at the correct diameter. Unfortunately, the acoustic code we use to simulate scattered wavefields becomes unstable for high density contrasts ( $|\rho^{steel} - \rho^{water}| \approx 6800 \text{ kg/m}^3$ ), so we could not simulate the diffraction behaviour for the case of steel rods. However, the numerical experiments performed for different material parameters suggest that if the scatterers are small enough (smaller than the wavelength) any material would produce a mostly isotropic scattered wavefield without internal ringing. Further, the high density and velocities of steel ( $v_P \approx 5800 \text{ m/s}$ ,  $v_S \approx 3100 \text{ m/s}$ , and  $\rho^{steel} = 7800 \text{ kg/m}^3$ ) should produce a large enough scattering amplitude to generate multiply scattered waves.

## 4.2 Geometry and acquisition of the laboratory experiment

A multiply scattered wavefield is excited and recorded by a transducer array located at one side of a homogeneous background medium containing multiple point scatterers. The background medium is a gel made from polyvinyl alcohol (PVA) with P-wave velocity  $v_P \approx 1500 \text{ m/s}$  and S-wave velocity  $v_S \approx 4 \text{ m/s}$ . Because of the low S-wave velocity, shear waves are not relevant within the time scale of the experiment (in the range of  $\mu\text{s}$ ) and thus the medium can be considered to be acoustic. Steel rods of diameter  $d = 1$  mm, length  $l \approx 11$  cm and velocities

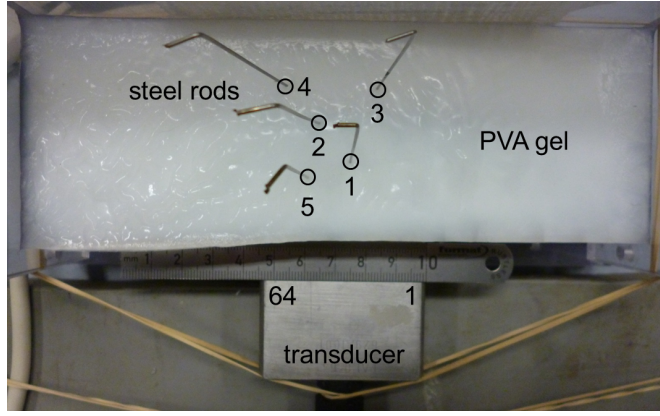


Figure 4.9: Plan view of acquisition set up for an acoustic multiple scattering experiment in a PVA gel. Five vertically oriented steel rods with diameter  $d = 1$  mm are used as point scatterers. Scatterer number 1 is used in the single-scattering case, all other scatterers are added one after the other according to the numbers given in the picture. The transducer array is attached to the front side of the gel with channel number 1 on the right-hand side and channel number 64 on the left-hand side.

$v_P \approx 5800$  m/s and  $v_S \approx 3100$  m/s that can be stuck into the gel are used as point scatterers (Fig. 4.9).

The transducer array consists of 64 channels evenly spaced at intervals of 0.75 mm. Each channel can both emit and receive a signal, hence we can think of one channel as a collocated source-receiver pair and the recorded data set corresponds to a data cube of intersecting CSGs and CRGs. A Gaussian pulse, designed in the controlling computer, is sent into the medium via a digital-to-analogue converter. The wavelet has a central frequency of  $f_0 = 0.5$  MHz, a peak amplitude of  $A = 50$  V and a total length of  $T = 8.1$   $\mu$ s. Given the P-wave velocity of the gel and the central frequency of the signal, the typical wavelength of the signal is  $\lambda = v_P/f_0 = 3$  mm, hence, three times as large as the diameter of the scatterers. Each source is fired separately 100 times with a temporal delay of approximately 2 ms between two subsequent shots. The corresponding wavefields are recorded at the receiver array and an average is computed for each source. Recording starts 20  $\mu$ s after each shot in order to minimize the effect of surface wave energy at shorter travel times, and ends after 85  $\mu$ s just before the waves reflected off the boundaries of the gel arrive. The sampling rate at the receivers is  $f_s = 10$  MHz.

The gel is installed in a plexiglass frame that is open on the top and the front sides, and the transducer is firmly attached to the front side of the gel. If desired, the whole system can be placed in a water tank, the advantages of which are discussed below. When everything is set up, it is recommended to wait for about 1 h for the

Measured		Computed		Difference	
x	y	x	y	x	y
28 mm	24 mm	22.5 mm	22.6 mm	5.5 mm	1.4 mm
35 mm	35 mm	30.7 mm	37.2 mm	4.3 mm	2.3 mm
17 mm	45 mm	15.0 mm	45.8 mm	2.0 mm	0.8 mm
44 mm	46 mm	39.8 mm	47.4 mm	4.2 mm	1.4 mm
38 mm	19 mm	35.3 mm	19.1 mm	2.7 mm	0.1 mm

**Table 4.3:** Comparison of measured and computed scatterer locations. x-values are with respect to channel number 1 on the right-hand side of the transducer (measured in the leftward direction) and y-values are with respect to the transducer-gel boundary (Fig. 4.9). Deviations are likely to occur due to scatterers not being exactly vertical in the gel, hence their true position at the transducer level differs from the location measured at the surface. Computed values are inferred from the apices of primary move-outs using a background velocity of  $v_P = 1475$  m/s.

system to stabilize before adding the first scatterer to the system. During this time the gel adjusts to the frame, deforming due to the pressure caused by the attached transducers. To evaluate the end of the stabilization process, monitoring acquisitions are carried out at regular intervals. If the difference between the wavefields of two subsequent acquisitions becomes negligible (below 1%) the gel is stable enough for the experiment to start. After adding new scatterers to the gel we allow the system to stabilize for approximately 10 min before making the next measurement. This is to ensure that the wavefield does not change between two acquisitions due to changes of the background medium, and that all changes observed are solely related to the additional scatterers.

Scatterer positions are measured manually with a scale at the surface of the gel (see Table 4.3). x-values are with respect to the location of channel 1 of the transducer array and y-values are with respect to the interface between gel and transducer. The locations are required in order to compute theoretical travel times of scattered waves and to compare them to the travel times estimated by the algorithm. However, we notice significant deviations between the theoretical travel times and real arrivals in the data, which are likely to be due to inaccurate measurements of scatterer locations: note that while scatterer positions were determined at the surface of the gel, the transducer was located about 2 cm below the surface. If the scattering rods have not been oriented exactly vertical, the position of the scatterers at the level of the transducer might deviate from the surface position. We refine the locations using the spatial and temporal apex of the diffraction curves and a background velocity of  $v_P = 1475$  m/s, which provides the best fit for most primary travel-time curves. Table 4.3 lists the measured and

computed scatterer locations and the differences between the two.

### 4.3 Analysis of raw data and interpretation of artefacts

For a first quality check of the data, we consider the singly-scattered wavefield obtained after sticking the first rod into the gel. In this case we expect to see only the direct wave (surface wave) and the primary scattered wave associated with scatterer 1. We find that the data are dominated by high amplitude surface waves propagating between the transducer array and the surface of the gel with a velocity of approximately 1000 m/s. Due to their high energy and low velocity surface waves are likely to superimpose low amplitude scattered waves. In order to avoid this effect, a reference field is recorded using only the background medium without any scatterers. The reference wavefield contains only surface waves; subtracting these recordings from any scattered wavefield removes only the surface wave component. For this procedure to work, it is important that the gel has stabilized before the reference field is recorded, otherwise the surface waves will not subtract perfectly from the scattered wavefields and surface wave energy will remain in the data. Note that the reference field needs to be recorded before the first rod is put into the gel, as each rod leaves a hole in the medium when removed, which also acts as a scatterer. From hereon all data will be shown post surface wave removed.

At the outermost channels we observe spurious linear arrivals that seem to be travelling backwards along the array (Fig. 4.10, solid arrow). Because these waves travel with the velocity of surface waves but arrive only after the scattered wave, we assume that these waves are secondary surface waves excited by scattered waves that have propagated back to the surface. In particular, these waves seem to start at the end points of the transducer array, hence we assume that the end points of the array act as scatterers that transform the backscattered body waves into surface waves.

Another spurious event arriving just after the scattered wave has the same move-out as the scattered wave itself (Figs. 4.10, dashed arrow, and 4.11). Again, this artefact can be attributed to the end points of the transducer array acting as spurious scatterers; in fact, here we observe the reverse process to that described above: at the end points of the transducer, surface waves propagating along the array are scattered into the medium where they scatter again at the steel rods

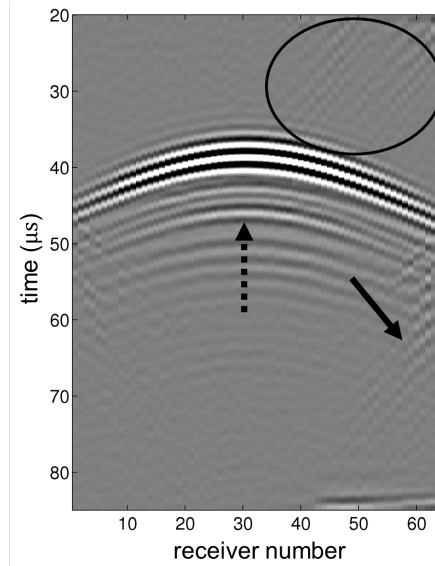


Figure 4.10: Linear converted body/surface wave (solid arrow) and scattered body wave (dashed arrow) spurious events in the single-scattering data. Encircled are remnants of primary surface wave energy that did not cancel out completely after subtraction of the reference field.

before being recorded at the receivers. The raypath of such an event is thus given by the path from the source to the end of the transducer, from there to the scatterer and from the scatterer back to the receiver array. Because this path is slightly longer compared to that of the primary scattered waves due to the additional surface wave part, the spurious wave reaches the receivers slightly later than the actual scattered wave, hence a constant time shift is observed between the two arrivals. This explanation is confirmed by the fact that the travel time difference of the spurious event in two neighbouring CSGs is equal to the travel time of a surface wave propagating between two channels ( $\approx 0.8 \mu\text{s}$ , see Fig. 4.11). Also, the time shift between spurious and actual arrival decreases when the source moves towards one of the end points of the array, since the part of the raypath travelled as surface wave becomes smaller.

In order to avoid spurious scattering at the transducer's end points, we lowered the central frequency of the signal and used different wavelets, testing both broader and narrower bandwidths. Lower frequency waves have larger wavelengths and are thus less sensitive to small scale scatterers. However, if the frequency is too low, the wavefield is not sensitive for the real scatterers in the medium anymore. By changing the bandwidth we might be able to filter out the frequencies particularly sensitive to end point scattering. Unfortunately, we did not succeed in finding a suitable trade-off that allows us to reduce the spurious scattering while

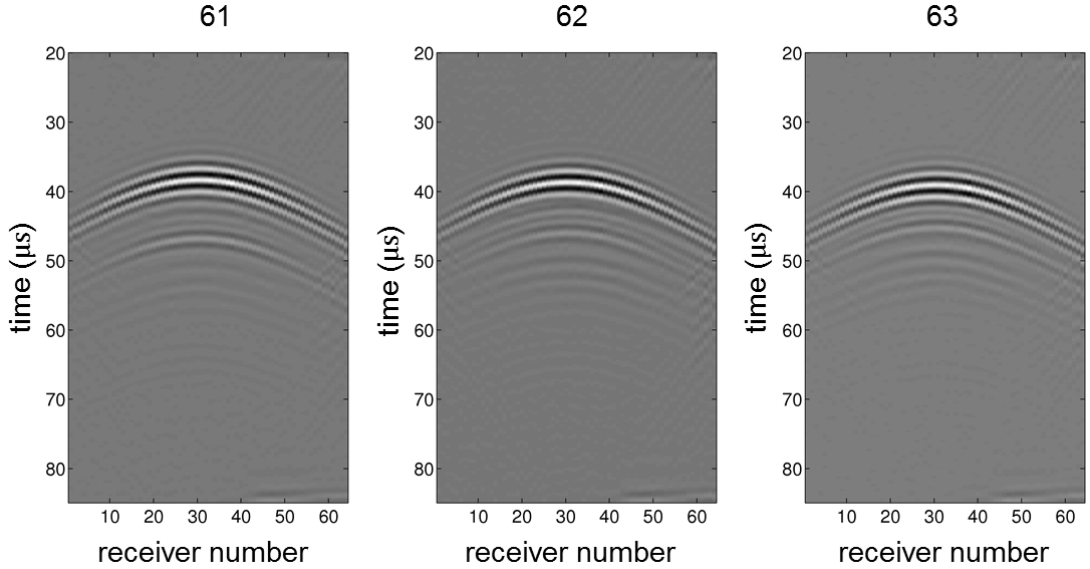


Figure 4.11: Spurious scattered wave observed at three CSGs (source numbers 61, 62 and 63). The difference in travel time of the spurious arrival between two neighbouring gathers ( $\approx 0.8 \mu\text{s}$ ) corresponds to the travel time of a surface wave between two neighbouring channels.

preserving the “real” scattering. By moving the whole experimental setup into a water tank, we tried to reduce the contrast between the gel, the transducer and the surrounding medium (air) in order to reduce the amplitude of the scattered surface waves. Finally, we used different types of transducer arrays with different coupling properties to see whether any type would create lower amplitude artefacts. However, none of these approaches noticeably reduced the amount of spurious energy.

Linear spurious events are unlikely to affect the performance of the automated algorithm; hence, they can remain in the data. Scattered spurious events, on the other hand, are more likely to distort the results since their arrival times, move-outs and amplitudes resemble secondary arrivals. Note, however, that unlike a true secondary, they have the same move-out in both a common-source and a common-receiver gather, since they originate from only one scatterer. In this sense they are more similar to primaries, however their amplitudes are much lower. Nevertheless, we will show below that despite these problems we observe relatively few errors directly related to spurious scattered events, which might be due to these unique properties.



## 4.4 Automatic identification of primaries and secondaries

Before running the automatic algorithm, the following parameters need to be specified: the number of cross-correlations that provide the final cross-gather ( $n_c$ ) and the thresholds that control the detection of primaries and multiples. We expect the values of these parameters to differ from those used in numerical tests, since they depend on the amount of noise in the data as well as on the distribution and the scattering amplitude of the scatterers. For two and three scatterers we obtain best results for  $ng = 12$ , a primary threshold of  $0.2 * \max(\Phi_1) * 0.9^n$  (where  $\max(\Phi_1)$  is the maximum of the first cross-correlation and  $n$  the number of primaries already detected), and a multiple threshold of half the mean of the semblance (for details see chapter 3, section 3.3.1). For four and five scatterers we need to amend the primary threshold to  $0.1 * \max(\Phi_1) * 0.9^n$ , because the scatterers are more widely distributed and as a result the difference in amplitude between the strongest and the weakest primary is more pronounced due to geometrical spreading.

Other than in numerical tests we use muted gathers (where detected primaries have been subtracted) not only in the cross-correlation of two gathers, but also when computing cross-gathers. This avoids errors caused by non-unique time shifts of primaries between two gathers, which play a larger role in real data due to the presence of spurious events, and is possible without diminishing the quality of the estimated travel-time curves when few move-outs are crossing.

Figs. 4.12 to 4.19 show the results for two to five scatterers on one mutual trace each and the corresponding CSGs and CRGs. Displayed are the estimated primary and secondary travel times on the mutual trace compared to the “true” arrival times computed from the measured scatterer locations (e.g. Fig. 4.12c), and the primary move-outs estimated from both the CSG and the CRG (e.g. Fig 4.13). In each case the mutual trace was picked randomly from the data cube. The algorithm verifies if the number of move-outs detected in the corresponding CSG and the CRG are equal, and if this is not the case a new mutual trace is evaluated. The examples shown represent cases in which the correct number of move-outs has been detected in both gathers.

For up to four scatterers the estimated primary move-outs are in good agreement with the data, and the few errors that occur can be attributed to the presence of spurious scattered waves (e.g. primary number 2 in Fig. 4.17a). For the 3-

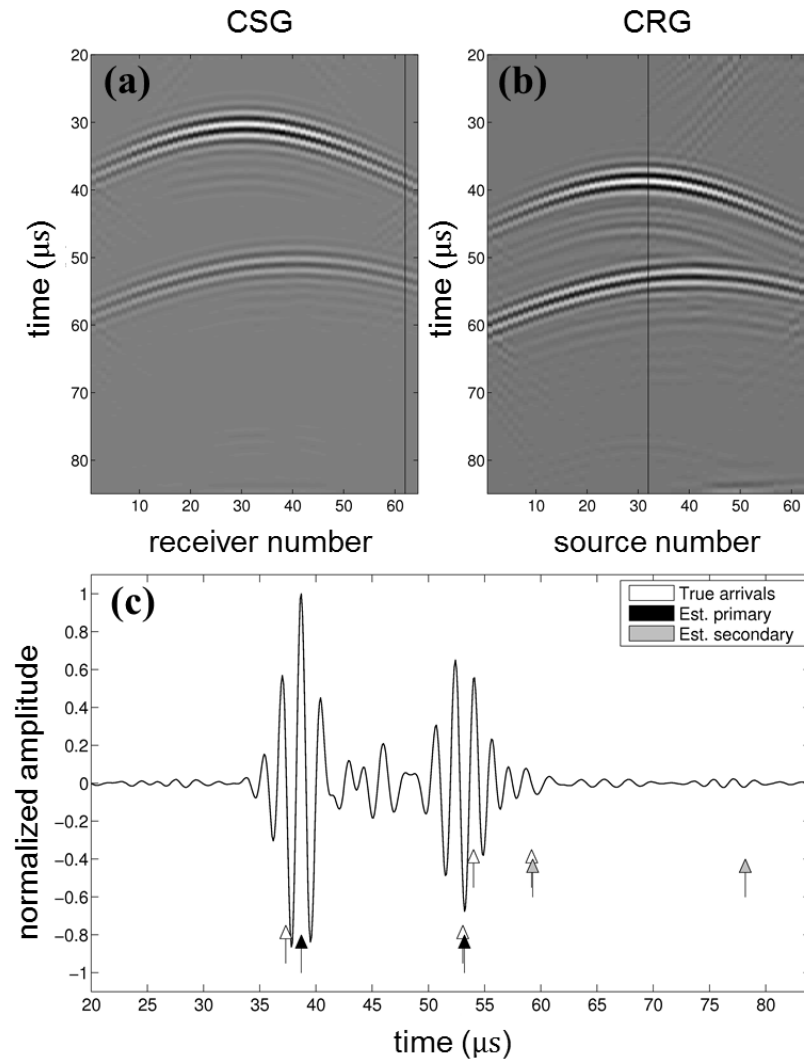


Figure 4.12: Data example for the 2-scatterers case. (a) Common-source gather (CSG) from source number 32 and (b) common-receiver gather (CRG) from receiver number 62. The solid black line marks the mutual trace that is common to both gathers. (c) Estimated primaries and secondaries (black and grey arrows, respectively) on the mutual trace compared to the true arrivals (white arrows).

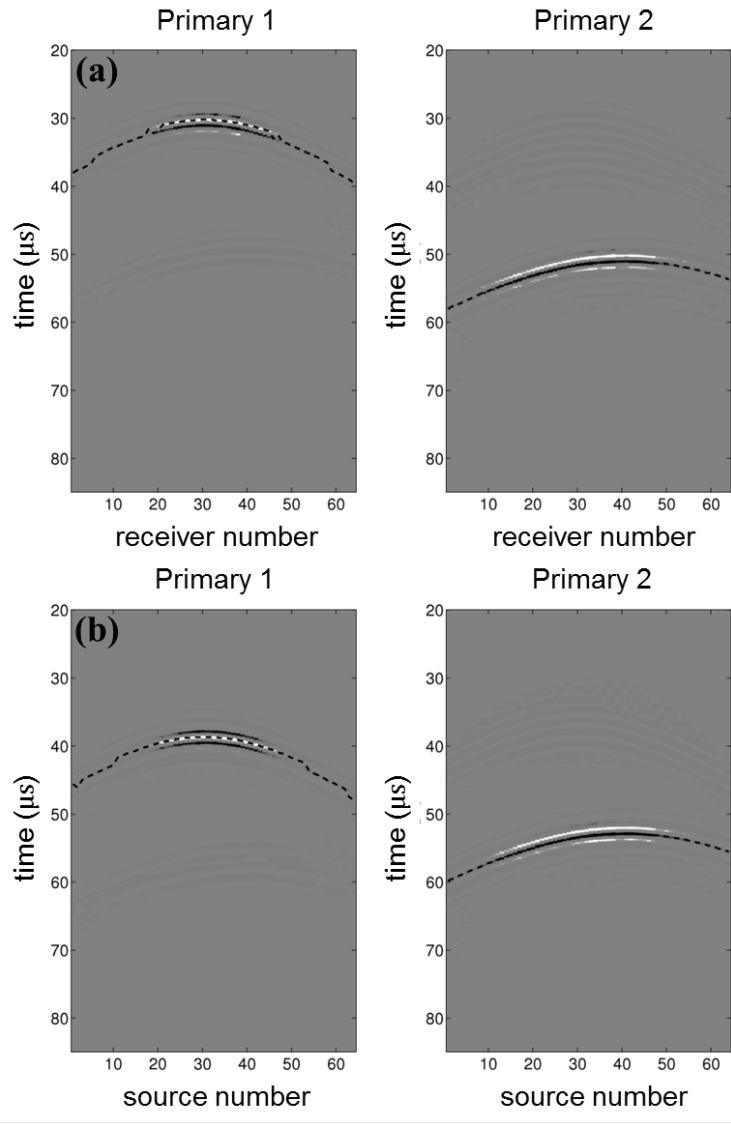


Figure 4.13: Left and right plots show estimated primaries (a) from the CSG in Fig. 4.12a and (b) from the CRG in Fig. 4.12b.

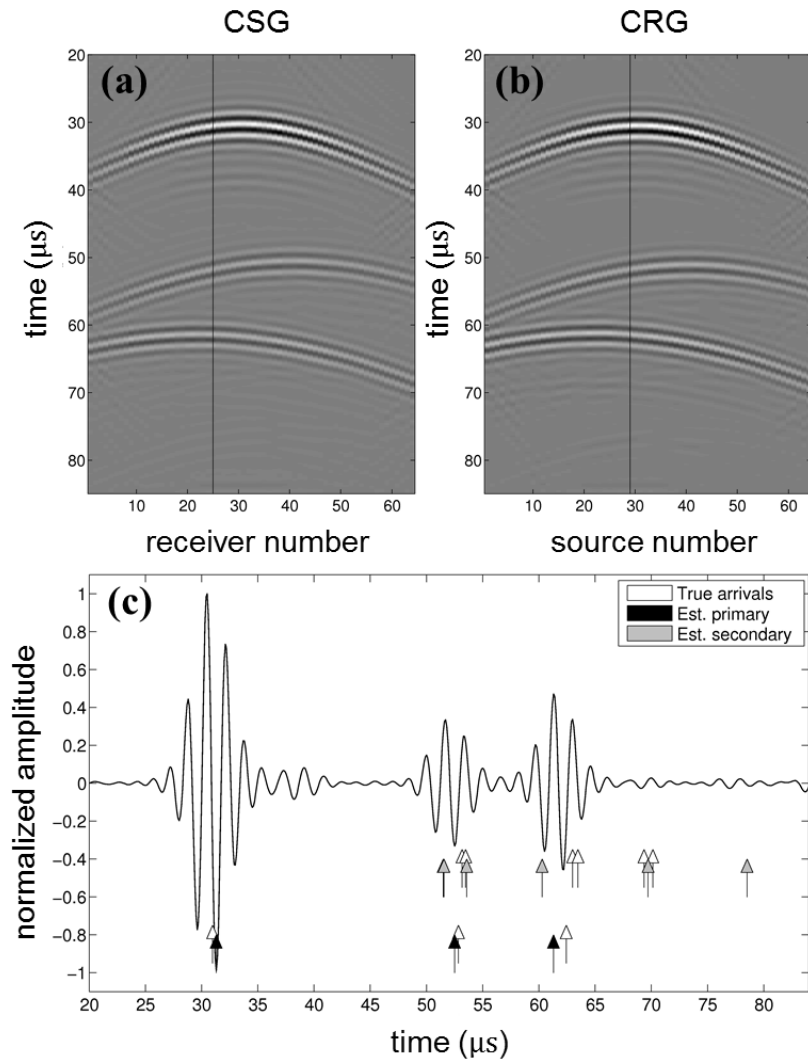


Figure 4.14: Data example for the 3-scatterers case. (a) CSG from source number 29 and (b) CRG from receiver number 25. The solid black line marks the mutual trace that is common to both gathers. (c) Estimated primaries and secondaries (black and grey arrows, respectively) on the mutual trace compared to the true arrivals (white arrows).

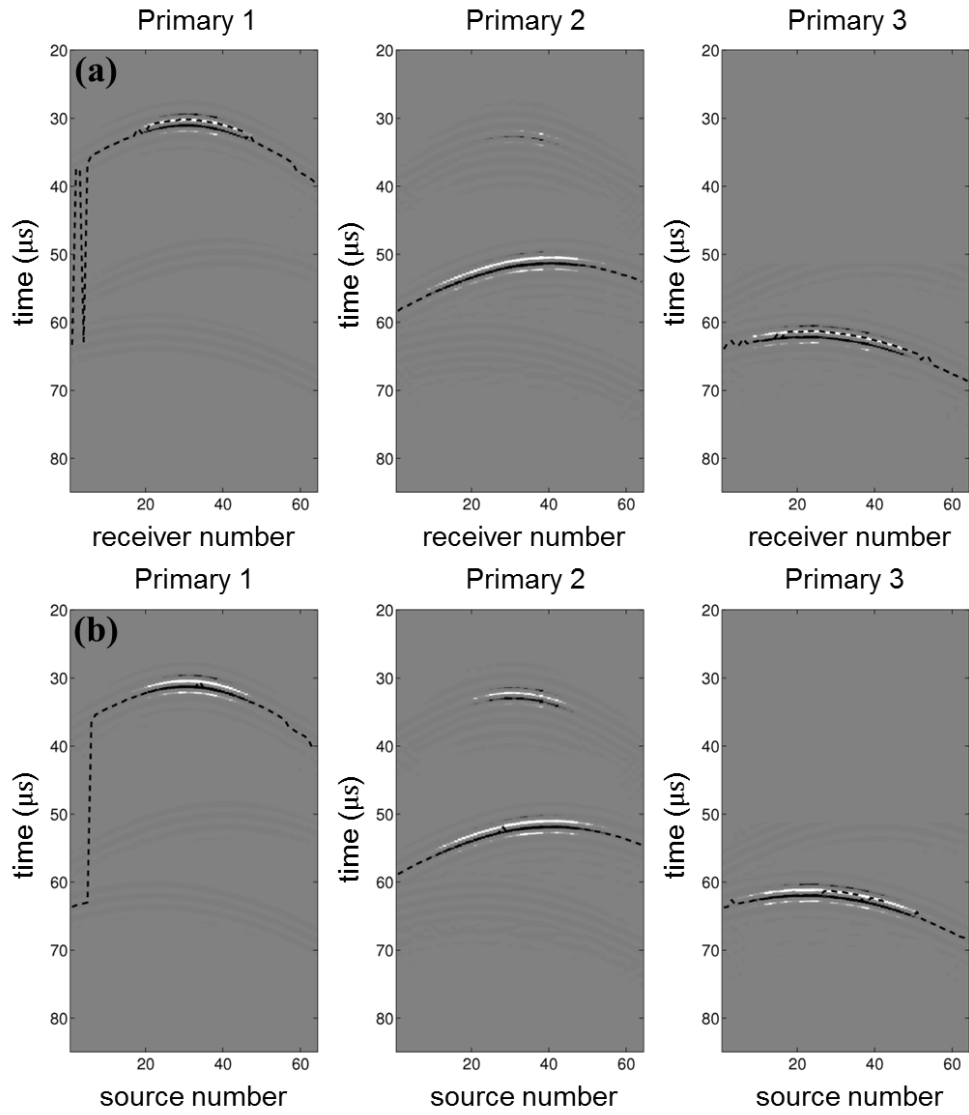


Figure 4.15: Left, central and right plots show estimated primaries (a) from the CSG in Fig. 4.14a and (b) from the CRG in Fig. 4.14b.

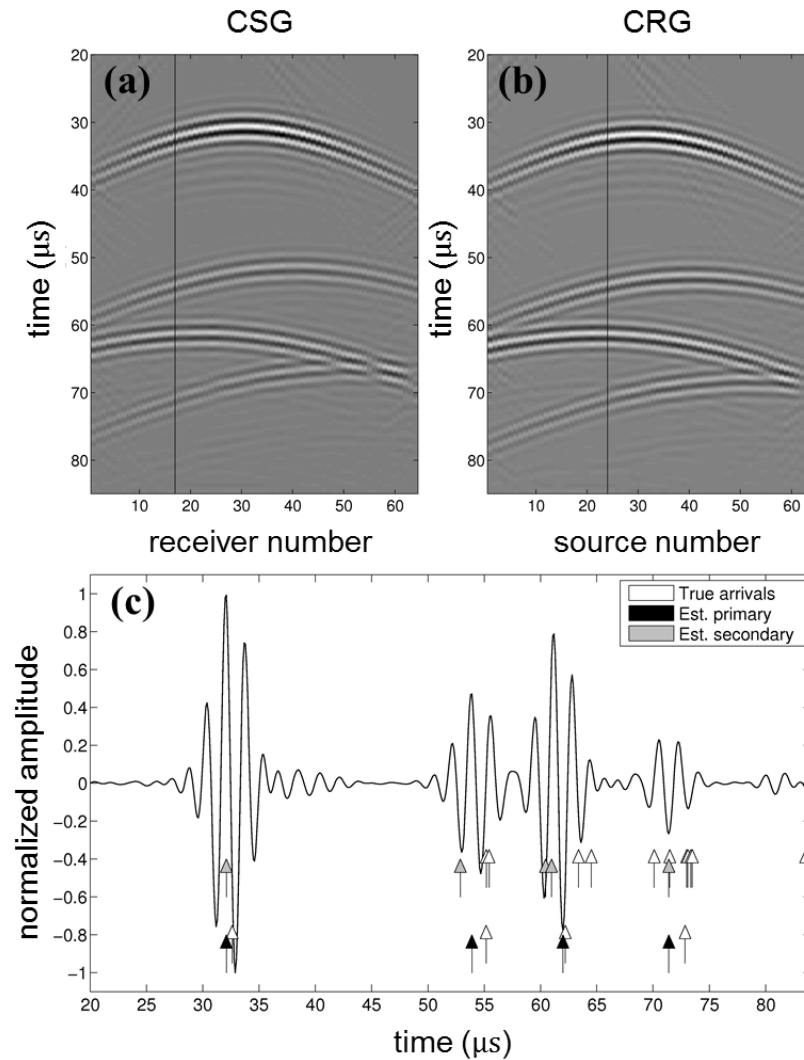


Figure 4.16: Data example for the 4-scatterers case. (a) CSG from source number 24 and (b) CRG from receiver number 17. The solid black line marks the mutual trace that is common to both gather. (c) Estimated primaries and secondaries (black and grey arrows, respectively) on the mutual trace compared to the true arrivals (white arrows).

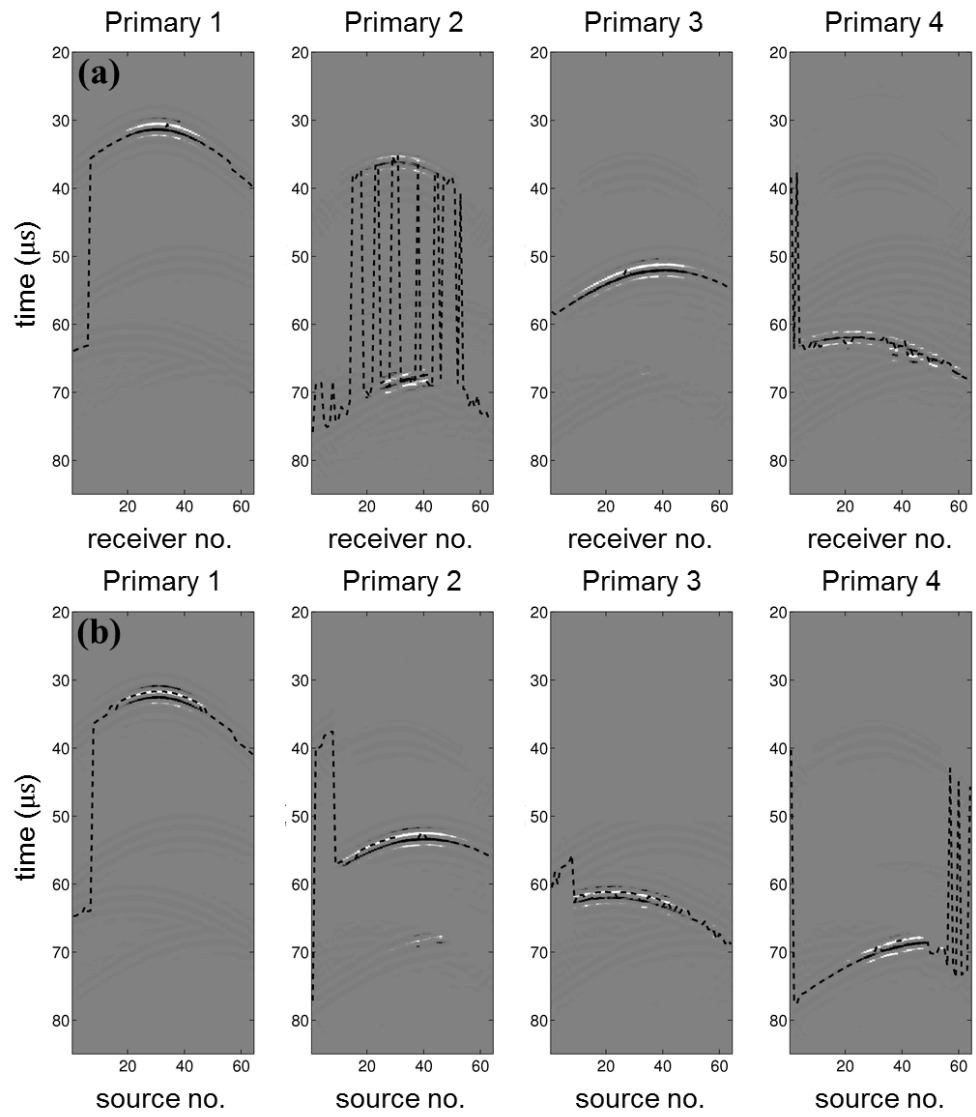


Figure 4.17: Plots across show estimated primaries **(a)** from the CSG in Fig. 4.16a and **(b)** from the CRG in Fig. 4.16b. Primary number 2 in (a) is strongly distorted by the spurious arrival after the first primary.

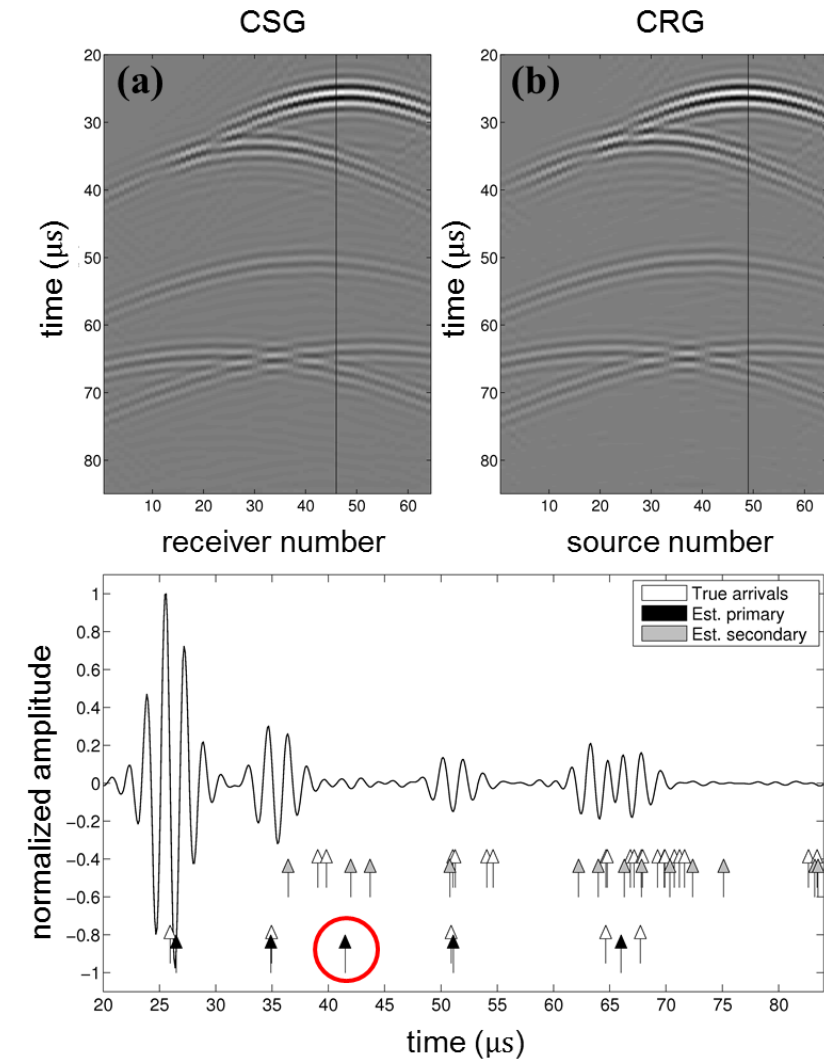


Figure 4.18: Data example for the 5-scatterers case. (a) CSG from source number 49 and (b) CRG from receiver number 48. The solid black line marks the mutual trace that is common to both gathers. (c) Estimated primaries and secondaries (black and grey arrows, respectively) on the mutual trace compared to the true arrivals (white arrows). The circle indicates an erroneously estimated primary arrival.



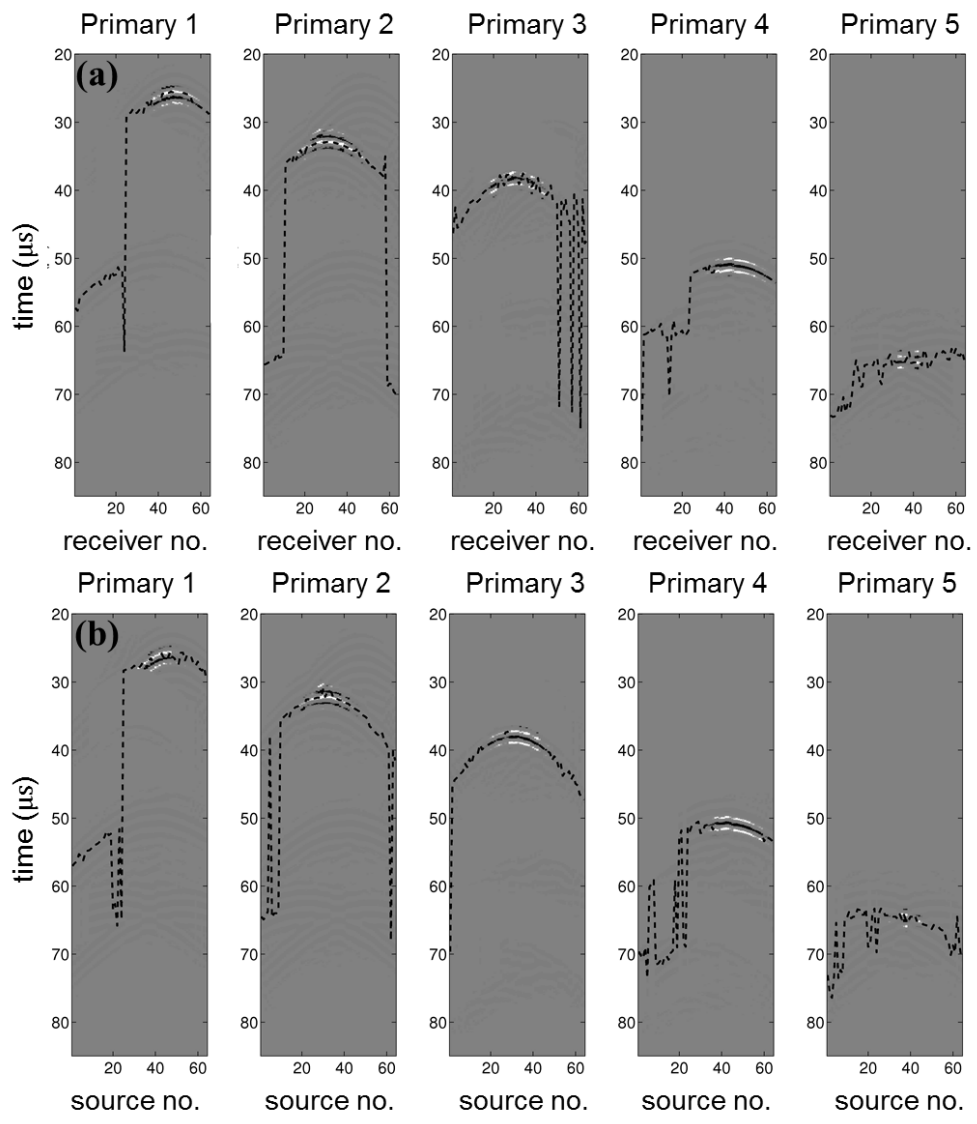


Figure 4.19: Plots across show estimated primaries (a) from the CSG in Fig. 4.18a and (b) from the CRG in Fig. 4.18b. Primary number 3 is not a real primary but a spurious event. The very last event in both gathers in Fig. 4.18 is not detected.

Primary	$t_{est}$	$t_{true}$	$ t_{est} - t_{true} $
1	31.3 $\mu s$	31.0 $\mu s$	0.3 $\mu s$
2	52.5 $\mu s$	52.9 $\mu s$	0.4 $\mu s$
3	61.3 $\mu s$	62.4 $\mu s$	1.1 $\mu s$

Table 4.4: Comparison of estimated ( $t_{est}$ ) and true travel times ( $t_{true}$ ) of primary arrivals on the mutual trace in the 3-diffractors example (Fig. 4.14c). All estimated travel times lie within the permitted deviation of 6.1  $\mu s$ .

Secondary	$t_{est}$	$t_{true}$	$ t_{est} - t_{true} $
2+1	51.5 $\mu s$	53.2 $\mu s$	1.7 $\mu s$
1+2	51.6 $\mu s$	53.5 $\mu s$	1.9 $\mu s$
1+3	53.6 $\mu s$	63.0 $\mu s$	9.7 $\mu s$
3+1	61.3 $\mu s$	63.5 $\mu s$	2.2 $\mu s$
2+3	69.7 $\mu s$	69.4 $\mu s$	0.3 $\mu s$
3+2	78.5 $\mu s$	70.1 $\mu s$	8.4 $\mu s$

Table 4.5: Comparison of estimated ( $t_{est}$ ) and true travel times ( $t_{true}$ ) of secondary arrivals on the mutual trace in the 3-diffractors example (Fig. 4.14c). Most estimated travel times lie within the permitted deviation of 7.1  $\mu s$ , only secondaries 1 + 3 and 3 + 2 are not identified correctly.

scatterers case the true and estimated primary travel times on the mutual trace are compared in Table 4.4. We need to take into account that also the “true” travel times are errorprone, since the locations of the scatterers could only be determined at the surface of the gel, and a comparison with locations estimated from the data (using the apex and curvature of the respective diffraction move-out) suggests that the uncertainty of the locations is about 3.7 mm in the x-direction and 1.2 mm in the y-direction on average (see Table 4.3). According to the propagation of uncertainties, this results in a travel time uncertainty of about 2  $\mu s$  for primaries. Adding this to the usual threshold for primary detection of half a wavelet yields a permitted deviation of 6.1  $\mu s$ ; hence, according to this threshold, all primaries in table 4.4 are estimated correctly. For “true” travel times of secondaries, the estimated uncertainty is 3  $\mu s$  on average, which gives a permitted deviation of 7.1  $\mu s$ . In most cases, the algorithm identifies only few secondaries correctly, rather it does not detect them at all. Both errors are likely to be related to the occurrence of spurious events that mask secondary arrivals.

Table 4.5 gives an example in which the difference between the estimated and the true secondary travel time is smaller than 7.1  $\mu s$  for most secondaries. Using

only these correct secondaries (and primaries) we predict related tertiaries and demonstrate that these are also visible in the data (Fig. 4.20). Note that due to the propagation of travel time errors when summing and subtracting primary and secondary travel times, predicted tertiary travel times may deviate by more than the above thresholds from the true arrival.

In the 5-scatterers example not all primary move-outs are estimated correctly. While in the example shown in Fig. 4.19 the estimated number of different primaries is correct, one detected primary is in fact a spurious event and the very last primary in the data (Fig. 4.18) is not identified at all. Testing different mutual traces and varying parameter  $n_c$  and the primary threshold in the algorithm did not enhance the results. We assume that the cross-correlation coefficient associated with the last primary is too small to be detected, because this primary, which has a low amplitude anyway, is partially removed when the previous, crossing move-out is muted. This example underlines one of the major problems in primary detection when the number of scatterers increases and multiple move-outs are crossing, and encourages the search for alternative procedures to remove individual primary energy without affecting other waves, as such procedures would complement the current method. Note that the other steps of the algorithm are independent of the method used for move-out estimation and energy removal.

## 4.5 Discussion

Compared to the synthetic data sets analysed in chapter 3, we observe that multiples have much lower amplitudes and are more difficult to distinguish because of spurious scattered energy in the data. It is likely that this problem is further enhanced by the fact that the steel rods used to emulate two-dimensional scatterers were not placed exactly vertically and thus allowed the scattered energy to radiate in three dimensions, which significantly reduces their amplitudes. Also, the setup of the experiment did not allow us to record the later part of the wavefield, which contains most of the multiply scattered energy, without interfering with reflected waves from the boundaries of the gel. These problems could be addressed by repeating the experiment using, for example, a larger block of gel with boundaries further away from the scatterers, a different background medium (e.g. water) that grants better control over the orientation of scatterers, and a transducer array that causes less scattering at its end points.

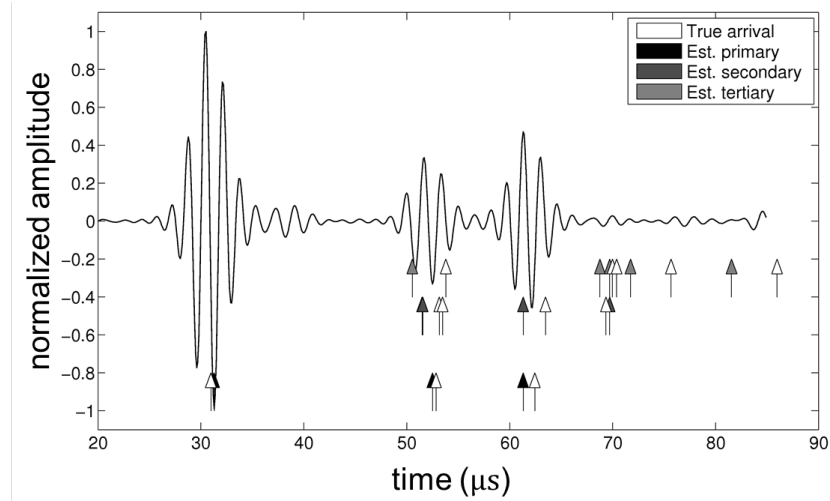


Figure 4.20: Mutual trace for the 3-scatterers case in Fig. 4.14. Black and grey arrows indicate correctly estimated travel times of primaries, secondaries and tertiaries; white arrows indicate the true travel times for comparison. For primaries and secondaries, travel times count as “correct” when they differ by less than  $6.1 \mu\text{s}$  or  $7.1 \mu\text{s}$ , respectively, from the true travel times (see Tables 4.4 and 4.5). Tertiaries may deviate by more than that due to the propagation of errors when summing and subtracting primary and secondary travel times.

We tried to attenuate spurious energy in the post-processing by deconvolving the data with an average wavelet that comprises both the scattered wave arrival and the spurious arrival. For the single-scatterer case one can define an average wavelet for each CSG, since the arrival time of the spurious wave with respect to the true scattered wave depends only on the scatterer position and the source position, but is constant along the receiver array. Deconvolving each trace in the CSG with the average wavelet removes the spurious arrival and sharpens the scattered wave arrival. However, this approach does not work with multiple scatterers in the medium, since the travel time difference between true and spurious arrival is different for each scatterer, hence the average wavelet computed from the single-scatterer case does not fit for other scatterers. In that case, deconvolution can produce even more spurious energy due to an unfavourable choice of wavelet.

The method used to estimate individual primary move-outs in a multiply diffracted wavefield relies on the invariance of primary move-outs in different common-source or common-receiver gathers. Further, it requires the waveforms along two equivalent move-outs to be similar in order to detect them. This assumption, however, is only valid for isotropic point diffractors that scatter the incoming energy equally in all directions. Testing different rod sizes and materials in numerical experi-

ments we noticed that anisotropic and internal scattering negatively affects the performance of the algorithm.

Anisotropy in the context of scattering means that the scattering amplitude  $A$  is directionally dependent. That is,  $A$  is not a constant but a function of the angle of incoming and outgoing energy,  $A(\mathbf{k}_1, \mathbf{k}_2)$ , where  $\mathbf{k}_1$  and  $\mathbf{k}_2$  are the wave vectors of the incoming and outgoing waves, respectively. This is the case when the diameter of the rod is similar to or larger than the typical wavelength. Moreover, not only amplitudes but also waveforms can be affected by this phenomenon. The angle of incoming energy at the scatterer is determined by the source location with respect to the scatterer location. This means that a diffraction move-out associated with a particular scatterer observed in different common-source gathers will have different amplitude and phase properties due to the difference in source location.

Our numerical simulations show that anisotropic behaviour changes the requirements and the performance of the automated algorithm and particularly affects the methods of primary and secondary detection, which look for *identical* waveforms in different gathers or at different travel times, respectively. Further, it raises the question about the definition of a move-out curve in media with anisotropic scatterers: by considering only kinematic information, i.e., the travel-time curves, a lot of information about the properties of the scatterers is neglected and the scattered wavefield is not sufficiently described. It remains the subject of future research to show how the algorithm described here or other interferometric methods can be used to extract and interpret information about anisotropic properties from scattered wavefields.

The method used for tertiary and higher-order multiple prediction implies that, in theory, all *kinematic* information about a multiply scattered wavefield is contained in its primary and secondary components. *Dynamic* information, however, is not preserved by our method, since the prediction of multiples is based on a term from the SRI equation that provides pseudo-physical energy with correct travel times but erroneous amplitude and phase information (see Eq. 2.12 in chapter 2). (Note that in this study we use a simplified version of SRI that does not perform convolution and cross-correlation of actual waveforms but simply sums and subtracts travel times; Meles and Curtis (2013) show examples where also the waveforms of multiples are reconstructed). For anisotropic scattering this problem is complicated by the fact that amplitudes and waveforms of diffracted waves depend on the direction of incoming and outgoing energy with respect

to the scatterer. Whether this information is preserved in cross-correlation and convolution processes performed in SRI could be a topic of future research.

## 4.6 Conclusions

In this section, we demonstrated that the automatic algorithm that estimates primary move-outs and predicts higher order multiples, is in principle applicable to data sets obtained from laboratory experiments in acoustic media containing multiple point scatterers. For up to four scatterers primary move-outs were estimated reliably and some secondaries were identified correctly. In most cases only few secondaries were identified incorrectly, i.e., with travel times estimated outside of a half-wavelet limit of the true travel time. Rather the algorithm did not find a secondary at all. The performance was negatively influenced by spurious energy in the data that resembled scattered wave energy and could not be sufficiently attenuated by changing the experimental setup or post-processing the data. It would be desirable to repeat the experiments with an improved setup to reduce the amount of spurious energy in the data and to further investigate the shortcomings of the method in real data applications.

# Chapter 5

## Computationally efficient internal multiple prediction based on source-receiver interferometry

In previous chapters I demonstrated that the term  $G_S G_S G_S^*$  of the SRI equation provides pseudo-physical scattered energy that can be used to predict arrival times of multiply diffracted waves and their ordered scattering paths in multiply diffracted wavefields. I now consider a scattering medium that contains reflecting interfaces rather than individual point scatterers. Using the kinematic connection between multiply diffracted waves and multiply *reflected* waves I show that the term  $G_S G_S G_S^*$  can also be used to estimate internal multiples in reflection seismic data. Internal multiples cause artefacts in seismic imaging and are therefore required to be removed prior to linear migration. I demonstrate that the internal multiples equation derived from SRI is equivalent to an existing equation derived from the inverse-scattering series (ISS) and yields an explicit link between the two concepts of SRI and the ISS. Further, I show that from the SRI perspective an alternative representation can be inferred that is based on cross-correlation and convolution—two operations that are computationally cheap and routinely used and therefore well understood in interferometric methods. The alternative representation therefore provides an efficient way to estimate and attenuate internal multiples, which can present a crucial advantage particularly in 3D applications.

This chapter has been submitted as a jointly-authored publication<sup>1</sup>. I, as lead author, have done the writing of the manuscript, performed the mathematical derivations and provided major contributions to the writing of the programming codes. Co-authors gave advice and support on the scope of the project, pro-

---

<sup>1</sup>K. Lör, G.A. Meles, and A. Curtis. Relating source-receiver interferometry to the inverse-scattering series provides a new method to estimate internal multiples. *Geophysics*.

vided background knowledge and draft computer codes, and helped editing the manuscript.

## Abstract

An explicit relationship between the representations of internal multiples by source-receiver interferometry (SRI) and the inverse-scattering series (ISS) is presented. This provides new insight into the role of pseudo-physical energy in multiply reflected wavefield estimates and leads to an alternative, computationally more efficient way to predict internal multiples.

## 5.1 Introduction

Interferometry refers to a set of methods that allow us to synthesize Green's functions between pairs of receivers (inter-receiver interferometry—Wapenaar, 2004; van Manen et al., 2005, 2006; Wapenaar and Fokkema, 2006), pairs of sources (inter-source interferometry—Hong and Menke, 2006; Curtis et al., 2009), or a source and a receiver (source-receiver interferometry or SRI—Curtis and Halliday, 2010) by means of cross-correlation, convolution or deconvolution (e.g. Vasconcelos and Snieder, 2008) of wavefields. The latter of the three methods, SRI, has been subject to increasing interest due to its close relationship to seismic imaging methods (Halliday and Curtis, 2010) and the new perspective it provides on non-linear imaging schemes and so-called extended images (Vasconcelos et al., 2010; Fleury and Vasconcelos, 2012; Ravasi and Curtis, 2013; Ravasi et al., 2014). Other applications of SRI include ground-roll removal in land-based exploration seismology (Duguid et al., 2011), construction of underside reflections from borehole recordings (Poliannikov, 2011), retrospectively observing seismograms from old earthquakes in seismology (Curtis et al., 2012; Entwistle et al., 2015), suppression of non-physical reflections in standard interferometry (King and Curtis, 2012), and prediction of multiply diffracted events and identification of scattering paths (Meles and Curtis, 2014a; L er et al., 2015). We focus on this last application and show that by considering multiply reflected scattering paths a new method is obtained to predict internal multiples in reflection seismic data.

While surface-related multiples cause major problems in marine seismic data,



internal multiples (i.e., interbed multiples generated between subsurface stratal interfaces) affect both marine and land data in the presence of strong reflectors, such as the water bottom, and the top and bottom of salt or basalt layers. Though there have been attempts to use multiply scattered waves in seismic imaging (e.g., Jiang et al., 2007; Malcolm et al., 2009; Fleury, 2013), most migration schemes rely on a single-scattering assumption and therefore require both surface-related and internal multiples to be removed from the data prior to migration.

Suppression of surface-related multiples has been addressed successfully by a number of methods (for a review see Verschuur, 2013), whereas relatively few methods exist that identify and attenuate internal multiples. Methods that rely on move-out discrimination, for example in the radon domain (Hampson, 1986), tend to fail for internal multiples as their move-out velocities are often similar to those of primaries. Berkhout and Verschuur (1997) propose a layer-related internal multiple elimination scheme (based on surface-related multiple elimination, SRME; Verschuur et al., 1992) that downward extrapolates shot records to a virtual acquisition surface and eliminates all multiples generated by that surface. This method, however, requires a velocity model to create the redatumed data. Jakubowicz (1998) suggests a data-driven approach based on the work of Keydar et al. (1997) that combines three primary reflections to predict a first order multiple. However, the primary reflection from the interface generating the interbed multiple needs to be identified and isolated from the recorded data, which can be difficult. Other schemes based on the same idea were proposed by Hung and Wang (2012) and Behura and Forghani (2012). Recently, Meles et al. (2014) proposed a scheme to estimate internal multiples based on Marchenko imaging and interferometry that requires autofocusing of wavefields, a relatively novel technique in seismics. The method has been applied successfully to synthetic data sets but still needs to be tested on real data. Another data-driven algorithm that predicts all internal multiples at once was first presented in Araújo et al. (1994) and is described in detail in Weglein et al. (1997, 2003). Travel times of internal multiples are predicted using a sub-series of the inverse-scattering series (ISS), which is derived from the Lippmann-Schwinger equation (Lippmann and Schwinger, 1950). This method is promising but has the downside that it is computationally expensive.

In this paper we show explicitly for wave propagation in a 1D medium (a medium that varies only in one dimension) and a collocated source and receiver that the SRI equation to estimate internal multiples is in fact equivalent to the internal-

multiple attenuation formula derived from the ISS (Weglein et al., 1997). We provide a concise derivation of both equations and demonstrate their equivalence by making use of a representation of Weglein’s formula provided by Ten Kroode (2002). Finally, an alternative representation of the same equation based on cross-correlation and convolution is presented, which provides a more efficient way to compute travel times of internal multiples and which decreases computational cost by many orders of magnitude.

## 5.2 An equation for internal multiples derived from the inverse-scattering series

We begin by deriving the internal-multiple equation based on the ISS (Weglein et al., 1997), starting with an introduction to (forward) scattering theory.

### 5.2.1 Forward scattering theory

In a scattering medium, the response to an impulsive source, the so-called Green’s function  $G$ , can be written as the sum of an unperturbed component  $G_0$  that propagates in a background or reference medium, and a perturbed component  $G_S$  that interacts with added scattering perturbations to the medium:

$$\mathbf{G} = \mathbf{G}_0 + \mathbf{G}_S \quad (5.1)$$

where the Green’s functions  $\mathbf{G}$ ,  $\mathbf{G}_0$  and  $\mathbf{G}_S$  are matrices in which the first two indices represent different spatial coordinates and the last index represents different temporal frequencies. Their elements are  $G(\mathbf{x}_j, \mathbf{x}_i, \omega)$ ,  $G_0(\mathbf{x}_j, \mathbf{x}_i, \omega)$  and  $G_S(\mathbf{x}_j, \mathbf{x}_i, \omega)$ , respectively, where  $G(\mathbf{x}_j, \mathbf{x}_i, \omega)$  propagates from  $\mathbf{x}_i$  to  $\mathbf{x}_j$ . Wave propagation between a source at  $\mathbf{x}_s$  and a receiver at  $\mathbf{x}_r$  in the actual and the reference medium is described by the differential equations

$$\begin{aligned} \mathbf{L}G &= -\delta(\mathbf{x}_r - \mathbf{x}_s) \\ \mathbf{L}_0G_0 &= -\delta(\mathbf{x}_r - \mathbf{x}_s) \end{aligned} \quad (5.2)$$

with the differential operators  $\mathbf{L} = \frac{\omega^2}{\kappa(\mathbf{x})} + \nabla \cdot \left( \frac{1}{\rho(\mathbf{x})} \nabla \right)$  and  $\mathbf{L}_0 = \frac{\omega^2}{\kappa_0(\mathbf{x})} + \nabla \cdot \left( \frac{1}{\rho_0(\mathbf{x})} \nabla \right)$ , where  $\kappa$  and  $\kappa_0$  as well as  $\rho$  and  $\rho_0$  are the actual and the reference bulk modulus and density of the medium, respectively. The perturbed component  $\mathbf{G}_S$  is called

the scattered field and is the part that carries information about the perturbations to the medium which can be diffractors or reflectors in general. The three Green's functions in Eq. 5.1 are related by the Lippmann-Schwinger equation, which in the frequency domain is

$$\mathbf{G}_S = \mathbf{G}_0 \mathbf{V} \mathbf{G} \quad (5.3)$$

Here  $\mathbf{V}$  is the perturbation operator defined as the difference between the two differential operators, i.e.,  $\mathbf{V} = \mathbf{L} - \mathbf{L}_0$ . When the problem is discretized,  $\mathbf{V}$  is a diagonal matrix in which the diagonal entries are non-zero if  $\kappa(\mathbf{x}) \neq \kappa_0(\mathbf{x})$  and/or  $\rho(\mathbf{x}) \neq \rho_0(\mathbf{x})$ . Substituting Eq. 5.1 into Eq. 5.3, and Eq. 5.3 into itself repeatedly, the Lippmann-Schwinger equation can be expanded into an infinite series of terms of increasing scattering order according to

$$\begin{aligned} \mathbf{D} &= \mathbf{G}_0 \mathbf{V} \mathbf{G}_0 + \mathbf{G}_0 \mathbf{V} \mathbf{G}_0 \mathbf{V} \mathbf{G}_0 + \mathbf{G}_0 \mathbf{V} \mathbf{G}_0 \mathbf{V} \mathbf{G}_0 \mathbf{V} \mathbf{G}_0 + \dots \\ &= \mathbf{D}_1 + \mathbf{D}_2 + \mathbf{D}_3 + \dots \end{aligned} \quad (5.4)$$

where  $\mathbf{D}_i$  is the  $i^{th}$  term in the scattering series. In Eq. 5.4, the scattered field  $\mathbf{G}_S$  on the left-hand side has been replaced by the data  $\mathbf{D}$ , which (after source signature deconvolution and subtraction of the reference field  $G_0$ ) we assume is equivalent to the scattered field at the measurement surface. The first term on the right-hand side of Eq. 5.4 accounts for first-order scattering and is also known as the single-scattering or Born approximation (Born and Wolf, 1999) and is also referred to as the primary wavefield. Our goal is to predict internal multiples, so we focus on the third term that describes third-order scattering as depicted in Fig. 5.1a for a medium with individual point scatterers. Fig. 5.1a shows a special case of third-order scattering, which was chosen deliberately as it resembles the geometry of a typical internal multiple in a horizontally layered reflecting medium (Fig. 5.1b). It satisfies the so-called "lower-higher-lower (LHL) condition" that ensures that the first scatterer  $(x_1, z_1)$  and the last scatterer  $(x_3, z_3)$  are located below the second scatterer  $(x_2, z_2)$ , expressed as

$$\begin{aligned} z_1 &> z_2 \\ z_3 &> z_2 \end{aligned} \quad (5.5)$$

Under the LHL condition the third-order scattering term  $\mathbf{G}_0 \mathbf{V} \mathbf{G}_0 \mathbf{V} \mathbf{G}_0 \mathbf{V} \mathbf{G}_0$  is the first-order internal-multiple generator in the forward scattering series. For a fixed source at  $\mathbf{x}_s$  and a fixed receiver at  $\mathbf{x}_r$ , Ten Kroode (2002) shows that

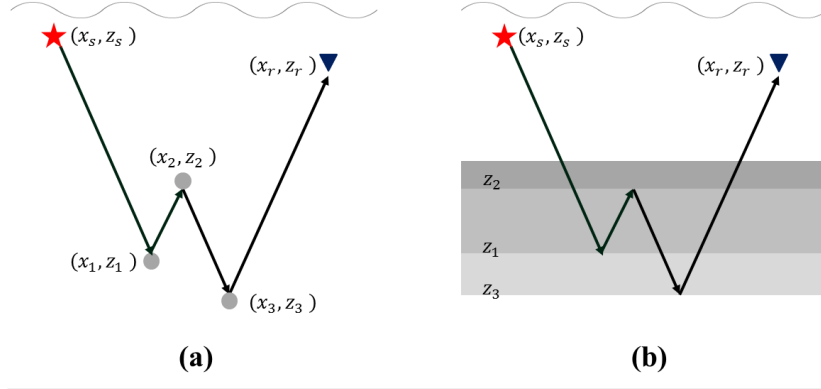


Figure 5.1: Example geometry for a third-order scattering event satisfying the lower-higher-lower condition in a medium with (a) individual point scatterers and (b) horizontal reflectors. The star indicates the source location and the inverted triangle indicates the receiver location; the background medium is water. The event in (b) represents an internal multiple of the kind we aim to predict.

$\mathbf{G}_0 \mathbf{V} \mathbf{G}_0 \mathbf{V} \mathbf{G}_0 \mathbf{V} \mathbf{G}_0$  can be written in the integral form

$$d_3^{IM} \approx \int_{\substack{z_1 > z_2 \\ z_3 > z_2}} G_0(\mathbf{x}_1, \mathbf{x}_s) V(\mathbf{x}_1) G_0(\mathbf{x}_2, \mathbf{x}_1) V(\mathbf{x}_2) G_0(\mathbf{x}_3, \mathbf{x}_2) \times V(\mathbf{x}_3) G_0(\mathbf{x}_r, \mathbf{x}_3) d\mathbf{x}_1 d\mathbf{x}_2 d\mathbf{x}_3 \quad (5.6)$$

where  $d_3^{IM}$  denotes the set of all first-order internal multiples recorded at receiver  $\mathbf{x}_r$  originating from a source at  $\mathbf{x}_s$  (dependency on frequency  $\omega$  has been omitted on the right-hand side for conciseness). Unfortunately in geophysical applications Eq. 5.6 cannot usually be computed directly as it requires a-priori information about the perturbation operator  $\mathbf{V}$ , also known as the reflectivity, at points throughout the volume spanned by the integral, which is not available (the reference Green's function  $G_0$  may be computed as the reference model is generally known). Hence, we need to solve the inverse problem of Eq. 5.4 first to obtain  $\mathbf{V}$  as a function of the data  $\mathbf{D}$ .

### 5.2.2 Inverse scattering and the internal multiples generator

Eq. 5.4 states that the data  $\mathbf{D}$  can be expanded into an infinite series  $\mathbf{D} = \sum_i \mathbf{D}_i$ , where  $\mathbf{D}_i$  is the portion of the data that is  $i^{th}$  order in  $\mathbf{V}$ . Invoking the properties of the geometric series, Weglein et al. (1997) argue that in the inverse series the reflectivity  $\mathbf{V}$  can be expanded equivalently into  $\mathbf{V} = \sum_i \mathbf{V}_i$ , where  $\mathbf{V}_i$  is the portion of  $\mathbf{V}$  that is  $i^{th}$  order in the data. Substituting  $\mathbf{V} = \sum_i \mathbf{V}_i$  into Eq. 5.4

they find that

$$\mathbf{D} = \mathbf{G}_0(\mathbf{V}_1 + \mathbf{V}_2 + \dots)\mathbf{G}_0 \quad (5.7)$$

$$\begin{aligned} &+ \mathbf{G}_0(\mathbf{V}_1 + \mathbf{V}_2 + \dots)\mathbf{G}_0(\mathbf{V}_1 + \mathbf{V}_2 + \dots)\mathbf{G}_0 + \dots \\ &= \mathbf{G}_0\mathbf{V}_1\mathbf{G}_0 \end{aligned} \quad (5.8)$$

since all higher-order terms cancel each other, for example

$$0 = \mathbf{G}_0\mathbf{V}_2\mathbf{G}_0 + \mathbf{G}_0\mathbf{V}_1\mathbf{G}_0\mathbf{V}_1\mathbf{G}_0 \quad (5.9)$$

$$\begin{aligned} 0 &= \mathbf{G}_0\mathbf{V}_3\mathbf{G}_0 + \mathbf{G}_0\mathbf{V}_1\mathbf{G}_0\mathbf{V}_2\mathbf{G}_0 + \mathbf{G}_0\mathbf{V}_2\mathbf{G}_0\mathbf{V}_1\mathbf{G}_0 \\ &+ \mathbf{G}_0\mathbf{V}_1\mathbf{G}_0\mathbf{V}_1\mathbf{G}_0\mathbf{V}_1\mathbf{G}_0 \end{aligned} \quad (5.10)$$

Eq. 5.8 allows us to infer the subseries  $\mathbf{V}_1$  directly from the data  $\mathbf{D}$ . In theory, all the other subseries could now be obtained sequentially:  $\mathbf{V}_2$  from  $\mathbf{V}_1$  using Eq. 5.9,  $\mathbf{V}_3$  from  $\mathbf{V}_1$  and  $\mathbf{V}_2$  using Eq. 5.10 and so on. The sum of all subseries then provides the true reflectivity  $\mathbf{V}$ , which is the missing component in the internal-multiple generator (Eq. 5.6). Weglein et al. (1997), however, approximate  $\mathbf{V}$  with its first subseries  $\mathbf{V}_1$ , which has the advantage that it can be directly computed from Eq. 5.8. The new multiples generator is thus  $\mathbf{G}_0\mathbf{V}_1\mathbf{G}_0\mathbf{V}_1\mathbf{G}_0\mathbf{V}_1\mathbf{G}_0$ . Note that this term is in fact cancelled by other terms (see Eq. 5.10) and therefore does not contribute physically to the internal multiples in the data  $\mathbf{D}$ . We will elaborate on this topic in the discussion and explain why it is still reasonable to infer information about internal multiples from this term.

### 5.2.3 A migration-demigration process

Following a more heuristic argument, Ten Kroode (2002) states that  $\mathbf{V}_1$  simply represents the common-shot migrated data, which can be used as an approximation for the unknown reflectivity function  $\mathbf{V}$  in Eq. 5.6. He points out that for fixed shots  $\mathbf{x}_s$  and under certain assumptions (for details see Ten Kroode, 2002) a function  $\mathcal{F}(\mathbf{x}_s)$  exists that maps the reflectivity onto the primaries  $\mathbf{D}_1 = \mathcal{F}(\mathbf{x}_s)\mathbf{V}$ . The inverse of this function is known as the common-shot migration and provides the reflectivity  $\mathbf{V}$  given the primaries  $\mathbf{D}_1$ :

$$\mathbf{V} = \mathcal{F}(\mathbf{x}_s)^{-1} \cdot \mathbf{D}_1 \quad (5.11)$$

The problem with Eq. 5.11 is that  $\mathbf{D}_1$  is typically not available (otherwise there was no need to eliminate the multiples). If we replace  $\mathbf{D}_1$  by the full recorded data  $\mathbf{D}$  including multiples we obtain

$$\mathbf{V}_1 = \mathcal{F}(\mathbf{x}_s)^{-1} \cdot \mathbf{D} \quad (5.12)$$

(which is the inverse of Eq. 5.8) where  $\mathbf{V}_1$  is the common-shot migrated data containing erroneously imaged multiples. Hence, using  $\mathbf{V}_1$  instead of  $\mathbf{V}$  in Eq. 5.6 will introduce errors; Ten Kroode (2002) argues that these will be smaller in amplitude than the internal multiples we aim to construct.

In summary, the migration process (Eq. 5.12) provides an estimate of the reflectivity, given the data  $\mathbf{D}$ . The migrated data (or the reflectivity estimate)  $\mathbf{V}_1$  can then be used in the internal-multiple generator (Eq. 5.6) to replace the true reflectivity  $\mathbf{V}$ . The internal-multiple generator itself can be regarded as a demigration process that generates (part of) the data, given the reflectivity; Verschuur (2013) thus refers to Weglein’s multiple estimation technique as a “migration-demigration process”. This also makes clear that the parameters used in the migration (e.g., the velocity of the reference medium) do not play an important role as errors committed in migration will be balanced by the inverse process (demigration). The internal multiples can therefore be estimated from the recorded data alone.

#### 5.2.4 An equation for internal multiples derived from the ISS

The final internal-multiple equation emerges from Eq. 5.6 after just one further modification: the second and third Green’s function on the right-hand side of Eq. 5.6 are rewritten such that each of them could have been emitted or recorded at the measurement surface:

$$\begin{aligned} G_0(\mathbf{x}_2, \mathbf{x}_1) &\approx \int G_0(\mathbf{x}_{r'}, \mathbf{x}_1) G_0^*(\mathbf{x}_{r'}, \mathbf{x}_2) d\mathbf{x}_{r'} \\ G_0(\mathbf{x}_3, \mathbf{x}_2) &\approx \int G_0(\mathbf{x}_3, \mathbf{x}_{s'}) G_0^*(\mathbf{x}_2, \mathbf{x}_{s'}) d\mathbf{x}_{s'} \end{aligned} \quad (5.13)$$

where  $\mathbf{x}_{r'}$  is an additional receiver location at the surface,  $\mathbf{x}_{s'}$  is an additional source location, and  $*$  indicates complex conjugation. The term in the integrand of Eq. 5.13 is a cross-correlation, which subtracts travel times in the phase term. Since  $\mathbf{x}_{r'}$  and  $\mathbf{x}_{s'}$  are not known in advance, integration over a source and a

receiver boundary at the surface ( $z = 0$ ) is required that provides contributions at the correct travel times following stationary phase arguments (Ten Kroode, 2002). Using Eq. 5.13 and substituting  $\mathbf{V}$  for  $\mathbf{V}_1$ , Eq. 5.6 can be rearranged to

$$d_3^{IM}(\mathbf{x}_r, \mathbf{x}_s) \approx \int_{\substack{z_1 > z_2 \\ z_3 > z_2}} [G_0(\mathbf{x}_1, \mathbf{x}_s)V_1(\mathbf{x}_1)G_0(\mathbf{x}_{r'}, \mathbf{x}_1)] \\ \times [G_0(\mathbf{x}_2, \mathbf{x}_s)V_1(\mathbf{x}_2)G_0(\mathbf{x}_{r'}, \mathbf{x}_2)] \\ \times [G_0(\mathbf{x}_3, \mathbf{x}_s)V_1(\mathbf{x}_3)G_0(\mathbf{x}_{r'}, \mathbf{x}_3)]d\mathbf{x}_1d\mathbf{x}_2d\mathbf{x}_3dx_{r'}x_{s'} \quad (5.14)$$

Each term in brackets now stands for a primary wave reflected at  $\mathbf{x}_1, \mathbf{x}_2$ , and  $\mathbf{x}_3$ , respectively. Finally, Ten Kroode (2002) expresses Eq. 5.14 in terms of the data  $\mathbf{D}$  recorded at times  $t$  for the 1.5D case (2D wave propagation in a medium that varies only in one dimension) as

$$d_3^{IM}(\mathbf{x}_r, \mathbf{x}_s, \omega) = (-i\omega)^2 \int_{\substack{t_1 > t_2 \\ t_3 > t_2}} \mathcal{A}(\mathbf{x}_s, \mathbf{x}_{s'}, \mathbf{x}_r, \mathbf{x}_{r'}, t_1, t_2, t_3)e^{i\omega(t_1 - t_2 + t_3)} \\ \times D(\mathbf{x}_s, \mathbf{x}_{r'}, t_1) \\ \times D(\mathbf{x}_{s'}, \mathbf{x}_{r'}, t_2) \\ \times D(\mathbf{x}_{s'}, \mathbf{x}_r, t_3)dt_1dt_2dt_3dx_{r'}dx_{s'} \quad (5.15)$$

For details about the amplitude factor  $\mathcal{A}(\mathbf{x}_s, \mathbf{x}_{s'}, \mathbf{x}_r, \mathbf{x}_{r'}, t_1, t_2, t_3)$  the reader is referred to Ten Kroode (2002). However, even without any knowledge of  $\mathcal{A}$ , correct kinematic information about first-order internal multiples can be inferred from Eq. 5.15. Note that for a constant background velocity  $c_0$  and a collocated source and receiver ( $\mathbf{x}_r = \mathbf{x}_s$ ),  $\mathcal{A}$  reduces to  $(\frac{2}{c_0})^2$ . For this case, the stationary points  $\mathbf{x}_{s'}$  and  $\mathbf{x}_{r'}$  coincide with the source-receiver pair ( $\mathbf{x}_s = \mathbf{x}_r = \mathbf{x}_{s'} = \mathbf{x}_{r'}$ ) and integration over source and receiver boundaries can be omitted. Thus, Eq. 5.15 becomes

$$d_3^{IM}(\omega) = \left(\frac{-2i\omega}{c_0(0)}\right)^2 \int_{\substack{t_1 > t_2 \\ t_3 > t_2}} D(t_1)D(t_2)D(t_3)e^{i\omega(t_1 - t_2 + t_3)}dt_1dt_2dt_3 \quad (5.16)$$

In Eqs. 5.15 and 5.16 the LHL condition (Eq. 5.5) has been transferred from depth to time, which is valid under the assumption of travel-time monotonicity

$$z_1 > z_2 \Leftrightarrow \tau(\mathbf{x}_s; \alpha, z_1) > \tau(\mathbf{x}_s; \alpha, z_2) \quad (5.17)$$

where  $\tau(\mathbf{x}_s; \alpha, z_i)$  is the travel time of a ray starting at  $\mathbf{x}_s$  under angle  $\alpha$ , reflecting at depths  $z_i$  and travelling back to the surface. This condition tends to hold in media without strong lateral velocity variations, down to the maximum depth at which a ray with take-off angle  $\alpha$  reflects.

Although Eq. 5.15 is derived from the first-order internal-multiple scattering term, the transformation of the LHL condition from depth to time makes it clear why also higher-order multiples are automatically generated by this equation: since the data  $D$  contain both primaries and multiples, an event arriving at travel time  $t_3 > t_2$  could either be a primary that has a longer travel time due to a deeper reflection point, or an internal multiple that has a longer travel time due to a longer (multiply reflected) propagation path. In the latter case, the combination of two primaries ( $D(t_1)$  and  $D(t_2)$ ) and a first order multiple ( $D(t_3)$ ) according to Eq. 5.15 results in an estimate of a second-order internal multiple. It should be noticed that there can be unfavourable combinations of primaries and internal multiples that, although consistent with the LHL criterion, may result in artefacts in the internal-multiple estimate, as discussed in Liang et al. (2013) and Ma and Weglein (2014).

Eq. 5.15 is also valid in 2D cases under the assumptions of travel-time monotonicity and conormal reflectivity (Ten Kroode, 2002), which means that the reflectivity  $\mathbf{V}(\mathbf{x})$  is singular in the direction  $\mathbf{n}(\mathbf{x})$  only, where  $\mathbf{x} \rightarrow \mathbf{n}(\mathbf{x})$  is a smooth map from  $\mathbb{R}^2$  to the unit circle. This is generally the case if the subsurface has a predominantly layered structure without point scatterers or angular boundaries.

Ten Kroode (2002) shows that the internal-multiple generator provided in Eq. 5.16 is equivalent to the 1D formula provided by Weglein et al. (1997, 2003)

$$b_3(k) = \int_{-\infty}^{\infty} dz_1 e^{ikz_1} b(z_1) \int_{-\infty}^{z_1} dz_2 e^{-ikz_2} b(z_2) \int_{z_2}^{\infty} dz_3 e^{ikz_3} b(z_3) \quad (5.18)$$

where  $k$  denotes the vertical wavenumber and  $b(z_i)$  is data with primaries and internal multiples in the so-called pseudo-depth domain, i.e., after migration with the reference velocity of water.

In the following, the time domain representation (Eq. 5.16) by Ten Kroode (2002) will be the basis for the comparison with the internal-multiple generator derived from SRI.



### 5.3 An equation for internal multiples derived from source-receiver interferometry

We will now derive the internal-multiple generator from source-receiver interferometry (SRI) starting with a brief revision of the key equation and the main assumptions in SRI.

#### 5.3.1 Introduction to SRI

The standard SRI equation in the monopole approximation gives an estimate of the homogeneous Green's function between a source at  $\mathbf{x}_s$  and a receiver at  $\mathbf{x}_r$  according to

$$G(\mathbf{x}_r, \mathbf{x}_s, \omega) + G^*(\mathbf{x}_r, \mathbf{x}_s, \omega) \approx \frac{4}{(c\rho)^2} \int_S \int_{S'} G(\mathbf{x}_{r'}, \mathbf{x}_s, \omega) G^*(\mathbf{x}_{r'}, \mathbf{x}_{s'}, \omega) G(\mathbf{x}_r, \mathbf{x}_{s'}, \omega) d\mathbf{x}_{r'} d\mathbf{x}_{s'} \quad (5.19)$$

(Curtis and Halliday, 2010). The double integral is over two closed surface boundaries ( $S$  over sources  $\mathbf{x}_{s'}$  and  $S'$  over receivers  $\mathbf{x}_{r'}$ ). The integrand comprises the product of three Green's functions representing the wavefields propagating between the central source ( $\mathbf{x}_s$ ) and the boundary receivers ( $\mathbf{x}_{r'}$ ), between the boundary sources ( $\mathbf{x}_{s'}$ ) and the boundary receivers ( $\mathbf{x}_{r'}$ ), and between the boundary sources ( $\mathbf{x}_{s'}$ ) and the central receiver ( $\mathbf{x}_r$ ), respectively (Fig. 5.2a). The symbol  $*$  denotes complex conjugation. Eq. 5.19 is an approximation to an exact equation given in (Curtis and Halliday, 2010). This holds if the boundaries  $S$  and  $S'$  are located in the far-field of each other and of the central source-receiver pair, such that all raypaths can be assumed to be perpendicular to the boundaries. Though in theory the boundaries are required to completely surround the central source-receiver pair ( $\mathbf{x}_s, \mathbf{x}_r$ ), Snieder (2004a) showed for inter-receiver interferometry that the main contributions to the integral come from the parts of the boundary where the integrand has approximately stationary phase, and contributions from elsewhere on the boundary cancel each other destructively. The same argument holds for SRI, which can be regarded as a concatenation of inter-receiver interferometry and inter-source interferometry (Curtis and Halliday, 2010; Curtis et al., 2012). Moreover, it has been shown (Löer et al., 2014) that linear surface boundaries (Fig. 5.2b), as typically used in seismic exploration, span the stationary points needed to construct pseudo-physical scattered

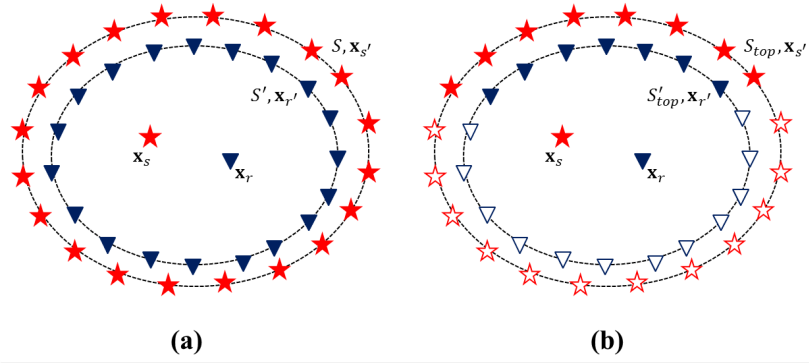


Figure 5.2: Conceptual geometry for source-receiver interferometry with **(a)** complete boundaries and **(b)** partial surface boundaries. Stars are sources, triangles are receivers.  $\mathbf{x}_s$  and  $\mathbf{x}_r$  denote the central source-receiver pair. Open symbols denote missing sources and receivers.

wave energy from SRI, assuming that the scattering part of the medium is located below the source-receiver pair. The term pseudo-physical denotes the fact that although travel times of the physical scattered wave energy are obtained correctly, amplitude and phase information of the wavefield constructed by Eq. 5.19 can be incorrect. Also if full boundaries were used, the pseudo-physical parts would cancel with contributions from other terms and not contribute to the final physical result. To obtain the physical scattered wavefield, stationary points that lie in the subsurface would have to be spanned by boundary sources and receivers, which is not practical in most experiments. It is therefore convenient to use *pseudo*-physical energy as an estimate of the physical scattered wavefield.

By separating the direct arrival  $G_0$  from the scattered field  $G_S$  using  $G = G_0 + G_S$  (Eq. 5.1 assuming that the reference field contains no scattering heterogeneities) the SRI equation can be written as the sum of eight terms, in each of which different combinations of perturbed (scattered) and unperturbed fields are cross-correlated and convolved. L  er et al. (2014) show for a single scatterer that the causal pseudo-physical scattered wavefield can be constructed using only one of the eight terms, namely the term that involves only scattered fields  $G_S$ . This also applies for media containing multiple scatterers or reflecting interfaces, the only drawback being that additional non-physical energy is introduced in these cases. The latter notwithstanding, SRI therefore constructs the pseudo-physical scattered wavefield between  $\mathbf{x}_s$  and  $\mathbf{x}_r$  from only the scattered components of the recorded wavefield and using only partial surface boundaries, and Eq. 5.19

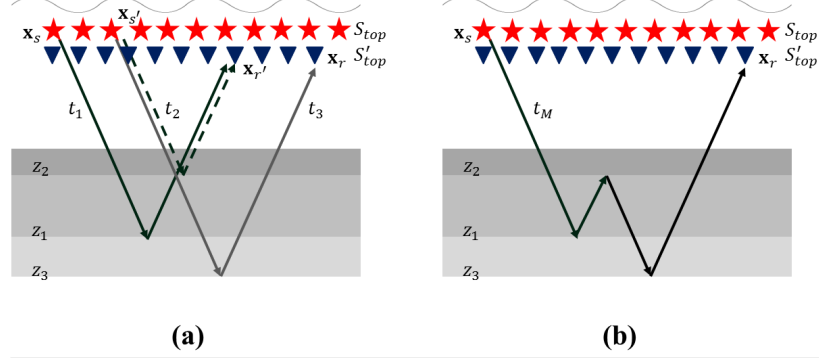


Figure 5.3: **(a)** Raypaths of primaries used in the multiple condition (Eq. 5.21 and 5.22). Portions of dashed and solid raypaths that run parallel cancel each other in SRI. **(b)** Raypath of the internal multiple with travel time  $t_M$  that can be calculated from the travel times of the primaries in (a). Symbol key as in Fig. 5.2.

becomes

$$\Phi(\mathbf{x}_r, \mathbf{x}_s, \omega) \approx \frac{4}{(c\rho)^2} \int_{S_{top}} \int_{S'_{top}} G_S(\mathbf{x}_{r'}, \mathbf{x}_s, \omega) G_S^*(\mathbf{x}_{r'}, \mathbf{x}_{s'}, \omega) G_S(\mathbf{x}_r, \mathbf{x}_{s'}, \omega) d\mathbf{x}_{r'} d\mathbf{x}_{s'} \quad (5.20)$$

where  $S_{top}$  and  $S'_{top}$  denote the partial surface boundaries of sources and receivers as shown in Fig. 5.2b. The scattered wavefields  $G_S$  on the right-hand side correspond to the data  $D$  recorded at the measurement surface comprising primaries and internal multiples (the direct wave arrival and all surface related multiples are assumed to have been removed). On the left-hand side,  $\Phi(\mathbf{x}_r, \mathbf{x}_s, \omega)$  denotes the constructed wavefield consisting of pseudo-physical primaries and internal multiples, and some non-physical energy.

### 5.3.2 The multiple condition

To estimate internal multiples only (thus not primaries or non-physical events), parts of the scattered Green's functions  $G_S$  are used that satisfy a “multiple condition”. If three primary events have been recorded as depicted in Fig. 5.3 with travel times

$$\begin{aligned} t_1 &> t_2 \\ t_3 &> t_2 \end{aligned} \quad (5.21)$$

Keydar et al. (1997) show that their travel times can be added and subtracted according to

$$t_M = t_1 - t_2 + t_3 \quad (5.22)$$

to yield the travel time of an internal multiple,  $t_M$ . The primaries are reflected at different depths and each primary is associated with a different source-receiver pair. Note that the first primary and the second primary (with travel times  $t_1$  and  $t_2$ , respectively) arrive at the same receiver location ( $\mathbf{x}_{r'}$ ) and that the second primary and the third primary (with travel times  $t_2$  and  $t_3$ , respectively) start at the same source location ( $\mathbf{x}_{s'}$ ). These locations must be chosen such that parts of the raypaths of the first and the second primary, as well as of the second and the third primary, run parallel so that the corresponding travel times exactly cancel each other (see dashed and solid lines in Fig. 5.3). In SRI these locations correspond to the stationary points of the surface integrals, some of which are included automatically if we integrate over  $S_{top}$  and  $S'_{top}$ . This is also true if the central source-receiver pair is not separated from the boundary, as in Fig. 5.3, and notice that if travel-time monotonicity (Eq. 5.17) holds, Eq. 5.21 is equivalent to the LHL condition in Eq. 5.5.

### 5.3.3 The equation for internal multiples from SRI

The following example shows how the LHL condition in Eq. 5.21 can be implemented inside the SRI equation to construct internal multiples only. Further, it helps to establish a clear link between SRI and the internal-multiple equation derived from the ISS.

Let us consider the 1.5D case (2D wave propagation in a medium that varies only in one dimension) and a collocated source and receiver ( $\mathbf{x}_s = \mathbf{x}_r$ ). For this case, the stationary points  $\mathbf{x}_{s'}$  and  $\mathbf{x}_{r'}$  coincide with the source-receiver pair ( $\mathbf{x}_s = \mathbf{x}_r = \mathbf{x}_{s'} = \mathbf{x}_{r'}$ ) and integration over source and receiver boundaries can be omitted (in the infinite frequency approximation when the Fresnel zone is infinitesimal). Thus, Eq. 5.20 reduces to

$$\Phi(\omega) = \frac{4}{(c\rho)^2} D(\omega) D^*(\omega) D(\omega) \quad (5.23)$$

Here, the scattered wavefield  $G_S$  has been replaced by the data  $D$ , which equals  $G_S$  at the measurement surface. Eq. 5.23 reproduces all primaries and multiples at frequency  $\omega$  between a collocated source and receiver, plus some non-physical

artefacts. Next, we express Eq. 5.23 as the Fourier transform of the time domain Green's functions:

$$\begin{aligned} \Phi(\omega) = & \frac{4}{(c\rho)^2} \int_{-\infty}^{\infty} dt_1 e^{i\omega t_1} D(t_1) \int_{-\infty}^{\infty} dt_2 e^{-i\omega t_2} D(t_2) \\ & \times \int_{-\infty}^{\infty} dt_3 e^{i\omega t_3} D(t_3) \end{aligned} \quad (5.24)$$

This time domain representation allows us to implement the multiple condition by modifying the integration boundaries of the second and the third integral such that  $t_1 > t_2$  and  $t_3 > t_2$ :

$$\begin{aligned} \Phi_{IM}(\omega) = & \frac{4}{(c\rho)^2} \int_{-\infty}^{\infty} dt_1 e^{i\omega t_1} D(t_1) \int_{-\infty}^{t_1 - \epsilon_1} dt_2 e^{-i\omega t_2} D(t_2) \\ & \times \int_{t_2 + \epsilon_2}^{\infty} dt_3 e^{i\omega t_3} D(t_3) \end{aligned} \quad (5.25)$$

where  $\epsilon_1$  and  $\epsilon_2$  are small positive numbers and  $\Phi_{IM}(\omega)$  denotes the internal multiples. Eq. 5.25 is the internal-multiple generator for the 1.5D case with a collocated source and receiver as derived from the SRI equation. As for the internal-multiple generator derived from the ISS, Eq. 5.25 can also predict higher-order multiples due to the presence of multiples in the data  $D$  and the ambiguity inherent in the LHL condition in the time domain.

It is easy to show that Eq. 5.25 is equivalent to Eq. 5.16 given by Ten Kroode (2002), which can be rearranged to

$$\begin{aligned} \Phi_{IM}(\omega) = & \left( \frac{-2i\omega}{c_0(0)} \right)^2 \int_{-\infty}^{\infty} dt_1 e^{i\omega t_1} D(t_1) \int_{-\infty}^{t_1 - \epsilon_1} dt_2 e^{-i\omega t_2} D(t_2) \\ & \times \int_{t_2 + \epsilon_2}^{\infty} dt_3 e^{i\omega t_3} D(t_3) \end{aligned} \quad (5.26)$$

The different pre-factors in ours and Ten Kroode's formula ( $\frac{4}{(c\rho)^2}$  and  $(\frac{-2i\omega}{c_0(0)})^2$ , respectively) are due to different source types (volume injection rate in ours versus volume injection in Ten Kroode's derivation—for details see Wapenaar and Fokkema, 2006) and the fact that Ten Kroode assumes a medium with constant density from the beginning and therefore does not include  $\rho$  in his wave equation (Eq. 1 in Ten Kroode, 2002).

Due to the pseudo-physical nature of the single SRI term that was used in Eq. 5.20, Eq. 5.25 provides the correct travel times of internal-multiple energy but amplitude and phase information of each multiple can be erroneous. Similarly,

Weglein et al. (2003) argue that amplitudes of internal multiples estimated from the ISS are wrong because only a subseries of the inverse-scattering series is considered that estimates internal multiples to first order only. Indeed, either by considering stationary points within SRI integral representations, or by using only subseries of the ISS, respectively L er et al. (2014) and Weglein et al. (2003) showed that the terms included in Eqs. 5.25 and 5.26 would in fact be cancelled if the full theory was used in each case. Thus in neither method is the final equation used the one that represents how multiples are actually constructed by SRI or the ISS, explaining from two quite different points of view why Eqs. 5.25 and 5.26 do not produce exactly correct physical results. Nevertheless we show next that the insights gained from linking the ISS to SRI allow a significantly different and computationally far more efficient representation for multiple prediction in both the ISS and the SRI frameworks.

## 5.4 Alternative representation

The SRI equation (Eq. 5.20 or 5.23) is a combination of correlation-type inter-receiver interferometry and convolution-type inter-source interferometry. Cross-correlation and convolution in the frequency domain are simply vector multiplications that are fast and cheap to compute. It would therefore be favourable if we could rewrite the internal-multiple equation (Eq. 5.25), which is based on the SRI equation, in a way that allows us to carry out these two operations separately, one after the other, in the frequency domain to reduce computational cost. In this chapter a new representation is derived that achieves this by using an alternative implementation of the multiple condition.

### 5.4.1 Correlation- and convolution-type representation of internal multiples

We start by identifying the correlational and convolutional operations in Eq. 5.23 and the derived internal-multiple equation (Eq. 5.25). In the frequency domain, a cross-correlation  $C(\omega)$  corresponds to a simple multiplication of two functions  $f(\omega)$  and  $g(\omega)$ , where one of the two functions has been complex conjugated:

$$C(\omega) = f(\omega)g^*(\omega) \tag{5.27}$$

In the time domain, cross-correlation of two functions  $f(t)$  and  $g(t)$  is given by

$$C(\tau) = \int_{-\infty}^{\infty} f(t)g(t + \tau)dt \quad (5.28)$$

where  $\tau$  is a time shift parameter. In Eq. 5.23 the correlational part is given by the multiplication of the term  $D^*(\omega)$  with either of the other two terms  $D(\omega)$ ; without loss of generality we will take the rightmost term. The corresponding time domain integrals in Eq. 5.25 can be rewritten by analogy with Eq. 5.28 by introducing a time shift parameter  $\tau$  and substituting the integration variable  $t_3$  for  $t_2 + \tau$ . Interchanging the order of integration, the rightmost two integrals can be written

$$\begin{aligned} & \int_{-\infty}^{t_1 - \epsilon_1} dt_2 e^{-i\omega t_2} D(t_2) \int_{t_2 + \epsilon_2}^{\infty} dt_3 e^{i\omega t_3} D(t_3) \\ &= \int_{\epsilon_2}^{\infty} e^{i\omega \tau} \left[ \int_{-\infty}^{t_1 - \epsilon_1} D(t_2) D(t_2 + \tau) dt_2 \right] d\tau \end{aligned} \quad (5.29)$$

The term in square brackets on the right-hand side of Eq. 5.29 resembles a cross-correlation at time shift  $\tau$ . However, the upper integration boundary is set to  $t_1 - \epsilon_1$  rather than to  $\infty$ . This is an important point that will be addressed later. Similarly, the outer integral over  $\tau$  resembles a Fourier transform but again one integration boundary, here the lower one, is incorrect.

The purpose of the restricted lower boundary is to ensure that when the correlation results for different time shifts  $\tau$  are summed, only positive time shifts  $\tau \geq \epsilon_2$  (or  $\tau > 0$ ) are considered. With  $t_3 = t_2 + \tau$ , this condition is equivalent to the part of the multiple condition stating that  $t_3 > t_2$ . An alternative way to implement this condition is to invoke the Heaviside step-function  $H(\tau)$  that takes value one for positive (causal) values of  $\tau$ , and zero for negative (acausal) values of  $\tau$ . Hence, the restriction of the lower integration boundary becomes obsolete and the cross-correlation part can be written as

$$\int_{-\infty}^{\infty} e^{i\omega \tau} \Gamma_{causal}(\tau) d\tau = \int_{-\infty}^{\infty} e^{i\omega \tau} \left[ H(\tau) \int_{-\infty}^{t_1 - \epsilon_1} D(t_2) D(t_2 + \tau) dt_2 \right] d\tau \quad (5.30)$$

where the function  $\Gamma_{causal}(\tau)$  represents the causal part of the correlation-type integral over  $t_2$  that is zero for negative values of  $\tau$ . The left-hand side of Eq. 5.30 is used to replace the latter two terms in the internal-multiple equation, Eq.

5.25, which yields

$$\Phi_{IM}(\omega) = \frac{4}{(c\rho)^2} \int_{-\infty}^{\infty} e^{i\omega t_1} D(t_1) \int_{-\infty}^{\infty} e^{i\omega\tau} \Gamma_{causal}(\tau) d\tau dt_1 \quad (5.31)$$

Thus we have expressed part of the internal-multiple equation by the correlation-type function  $\Gamma_{causal}(\tau)$ . Next, we show that Eq. 5.31 is equivalent to a convolution-type function. Convolution in the time domain is defined as

$$C'(\tau') = \int_{-\infty}^{\infty} f(t)g(\tau' - t)dt \quad (5.32)$$

As before, we change the integration variable of the inner integral in Eq. 5.31 by introducing a new time shift parameter  $\tau = \tau' - t_1$ . Rearranging the terms yields

$$\Phi_{IM}(\omega) = \frac{4}{(c\rho)^2} \int_{-\infty}^{\infty} e^{i\omega\tau'} \int_{-\infty}^{\infty} D(t_1)\Gamma_{causal}(\tau' - t_1)dt_1 d\tau' \quad (5.33)$$

The inner integral is equivalent to a convolution in the time domain, whereas the outer integral is a standard Fourier transformation from time to frequency domain.

#### 5.4.1.1 Cross-correlation and convolution uncoupled

Thus far we have shown explicitly how the formula that estimates internal multiples (Eq. 5.25) can be written as a concatenation of a correlation-type function (Eq. 5.30) and a convolution-type function (Eq. 5.33). However, in the current form of Eq. 5.33 the two are coupled through the argument of the correlation-type function  $\Gamma$ , which depends on the integration variable of the convolution-type function,  $t_1$ . We now show how the two operations can be uncoupled and performed independently in the frequency domain in order to save computational cost.

Firstly, we substitute the correlation-type function  $\Gamma_{causal}$  in Eq. 5.33 for the corresponding integral expression (right-hand side of Eq. 5.30):

$$\begin{aligned} \Phi_{IM}(\omega) = & \frac{4}{(c\rho)^2} \int_{-\infty}^{\infty} e^{i\omega\tau'} \int_{-\infty}^{\infty} D(t_1) \\ & \times H(\tau' - t_1) \int_{-\infty}^{t_1-\epsilon} D(t_2)D(t_2 + (\tau' - t_1))dt_2 dt_1 d\tau' \end{aligned} \quad (5.34)$$



The coupling of convolution and cross-correlation is implicit in the upper boundary of the innermost integral, which depends on the integration variable of the surrounding integral ( $t_1$ ). With a few modifications, however, this can be changed: due to the Heaviside function  $H$  the contribution to the outermost integral is zero if  $\tau' - t_1 < 0$ , hence for all  $t_1 > \tau'$ . We can thus cut the first data vector  $D(t_1)$  at  $t_1 = \tau'$  and keep only the part of  $D$  with  $t_1 < \tau'$  (reducing the length of the data also saves computation time). As  $t_2 < t_1$  (restricted by the upper boundary of the innermost integral) and consequently  $t_2 < \tau'$  we can remove the part of the second data vector  $D$  for which  $t_2 \geq \tau'$ . The argument of the third  $D$  is  $t_3 = t_2 + \tau' - t_1$ . From  $t_2 < t_1$  it follows that  $t_3 < \tau'$  and again we can remove the part of  $D$  for which  $t_3 \geq \tau'$ . The restricted integration boundaries are now obsolete and we can write

$$\begin{aligned} \Phi_{IM}(\omega) = & \frac{4}{(c\rho)^2} \int_{-\infty}^{\infty} e^{i\omega\tau'} \int_{-\infty}^{\infty} D_{m(\tau')}(t_1) \\ & \times H(\tau' - t_1) \int_{-\infty}^{\infty} D_{m(\tau')}(t_2) D_{m(\tau')}(t_2 + (\tau' - t_1)) dt_2 dt_1 d\tau' \end{aligned} \quad (5.35)$$

where notation  $D_{m(\tau')}(t_i)$  means that the vector  $D$  is muted for times  $t_i \geq \tau'$ . For each  $\tau'$  the correlational integral (over  $t_2$ ) and the convolutional integral (over  $t_1$ ) can now be evaluated separately, one after the other. This allows us to perform these operations in the frequency domain where they correspond to simple vector-multiplications that are fast and cheap to compute.

The Fourier integral over  $\tau'$  can be interpreted as an integral over possible multiple travel times  $t_M$ : only if  $\tau' = t_M$  is the multiple condition  $t_M = t_1 - t_2 + t_3$  fulfilled and the contribution of the integral is non-zero. Note that the multiple condition (Eq. 5.22) can be found explicitly in the argument of the third  $D$ :  $t_2 + (t_M - t_1) = t_3 \Leftrightarrow t_M = t_1 - t_2 + t_3$ .

The additional constraints described by the LHL condition,  $t_1 > t_2$  and  $t_3 > t_2$ , are also implicit by allowing only  $t < \tau'$  because

$$t_3 < \tau' \wedge t_3 = t_2 + (\tau' - t_1) \Rightarrow t_1 > t_2 \quad (5.36)$$

and by considering only the causal part of the cross-correlation

$$\tau' - t_1 > 0 \wedge t_3 = t_2 + (\tau' - t_1) \Rightarrow t_3 > t_2 \quad (5.37)$$

In summary, for a collocated source and receiver the internal multiple arriving at

time  $\tau'$  can be computed by performing the following steps:

- Cut the data  $D(t)$  at  $t = \tau'$  to obtain  $D(t < \tau')$ .
- Transform  $D(t < \tau')$  to the frequency domain to obtain  $D(\omega)$ , then multiply  $D(\omega)$  with its complex conjugate  $D^*(\omega)$  (corresponds to an auto-correlation in the time domain), the result of which is  $\Gamma(\omega)$ .
- Transform  $\Gamma(\omega)$  to the time domain to obtain  $\Gamma(\tau)$ . Keep only the causal part  $\Gamma_{causal}(\tau)$  by setting  $\Gamma(\tau < 0)$  to zero.
- Transform  $\Gamma_{causal}(\tau)$  to the frequency domain and multiply  $\Gamma_{causal}(\omega)$  with  $D(\omega)$  (corresponds to a convolution in the time domain).
- Integrate the result of step 4 over all  $\tau'$  to obtain an estimate of all internal multiples.

When considering a common-shot gather we need to reassign the geometrical parameters (source and receiver positions) to the data and integrate over source and receiver boundaries  $S$  and  $S'$ , respectively, as shown by SRI in Eqs. 5.19 and 5.20. This yields

$$\begin{aligned}
\Phi_{IM}(\mathbf{x}_r, \mathbf{x}_s, \omega) &= \frac{4}{(c\rho)^2} \int_S d\mathbf{x}_{s'} \int_{S'} d\mathbf{x}_{r'} \int_{-\infty}^{\infty} d\tau' e^{i\omega\tau'} \\
&\quad \times \int_{-\infty}^{\infty} dt_1 D_{m(\tau')}(\mathbf{x}_{r'}, \mathbf{x}_s, t_1) \\
&\quad \times H(\tau' - t_1) \int_{-\infty}^{\infty} dt_2 D_{m(\tau')}(\mathbf{x}_{r'}, \mathbf{x}_{s'}, t_2) \\
&\quad \times D_{m(\tau')}(\mathbf{x}_r, \mathbf{x}_{s'}, t_2 + (\tau' - t_1))
\end{aligned} \tag{5.38}$$

Note that in this case instead of performing an auto-correlation in the 2<sup>nd</sup> step we need to cross-correlate the correct portions of the data, namely  $D_{m(\tau')}(\mathbf{x}_{r'}, \mathbf{x}_{s'}, t)$  and  $D_{m(\tau')}(\mathbf{x}_r, \mathbf{x}_{s'}, t)$ . In the 3D case the integration boundaries over  $\mathbf{x}_{s'}$  and  $\mathbf{x}_{r'}$  span 2D surfaces rather than 1D lines.

## 5.5 Examples from a synthetic data set

We estimate internal multiples for the 1.5D case following a) Weglein's formulation from the ISS-version, (Eqs. 5.16 and 5.15) and b) the new representation from SRI (Eqs. 5.35 and 5.38) using a set of Matlab codes. We use a synthetic data set based on an acoustic 10-layers model (Fig. 5.4) provided by Total S.A.

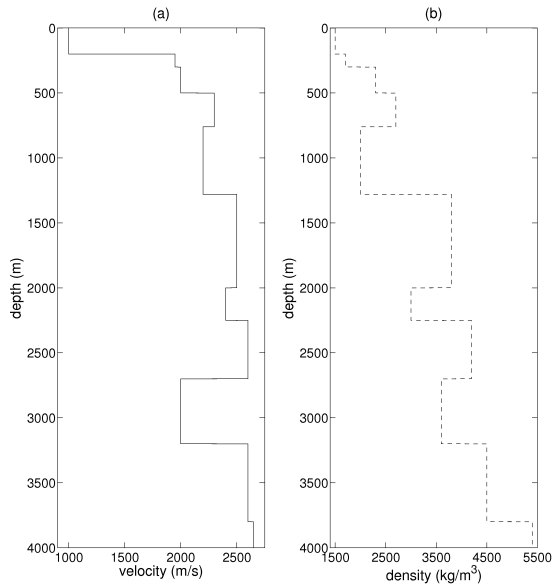


Figure 5.4: (a) Velocity profile and (b) density profile of the acoustic model used to generate synthetic data sets.

to compare the ISS and the SRI results both qualitatively in terms of the results and quantitatively in terms of computational cost.

Fig. 5.5 and Fig. 5.6 show the results for a collocated source and receiver using the ISS (Eq. 5.16) and SRI (Eq. 5.35), respectively. Displayed are in (a) the modeled scattered wavefield  $G_S$  (grey) in the time window between 0 s and 2 s. Surface-related multiples have not been modeled. Superimposed is the internal-multiple estimate (black). The amplitude factors in both equations have been ignored; instead the first arriving multiple was normalized with respect to the amplitude of the original trace at the same arrival time. This underlines the similarity between true and estimated multiples regardless of errors in the absolute amplitude. In (b) the internal-multiple estimate has been subtracted from the scattered wavefield and the result (black) is compared to the directly modeled primary wavefield (grey).

Fig. 5.7 shows internal-multiple estimates for a common-shot gather in the time window between 0 s and 1 s from a source at  $z = 5$  m and  $x = 12.5$  m recorded at 9 receivers located at the same depth between  $x = 12.5$  m and  $x = 212.5$  m with inter-receiver spacing of 25 m, where  $x$  is the horizontal coordinate axis. Boundary sources and receivers were distributed as linear arrays at depth  $z = 5$  m between  $x = 0$  m and  $x = 237.5$  m with a spacing of 12.5 m. The same velocity-density model has been used (Fig. 5.4). Figs. 5.7a and b show the full scattered wavefield (grey) superimposed by the internal multiples estimated using the ISS (Eq. 5.15)

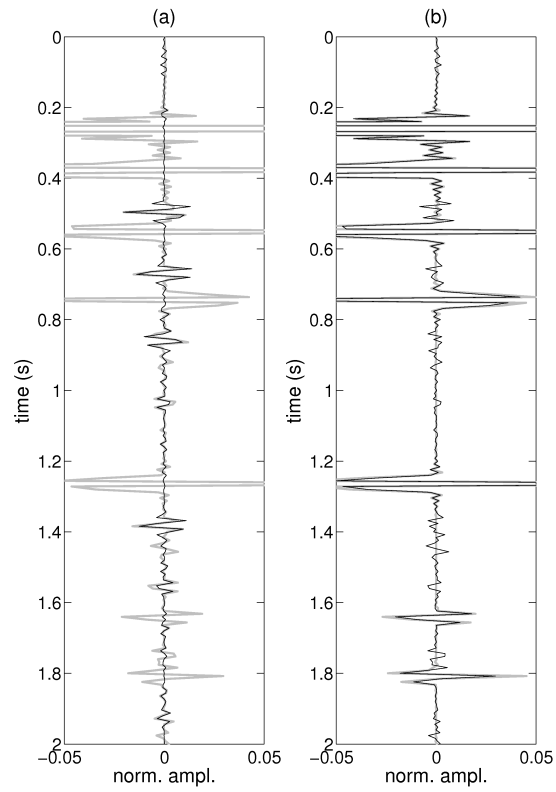


Figure 5.5: Internal multiples estimated from the ISS (Eq. 5.16) for a collocated source and receiver. (a) Estimate of internal multiples (black) compared to the full scattered wavefield (grey) (b) Demultiplied scattered field (black) compared to the primary wavefield (grey).

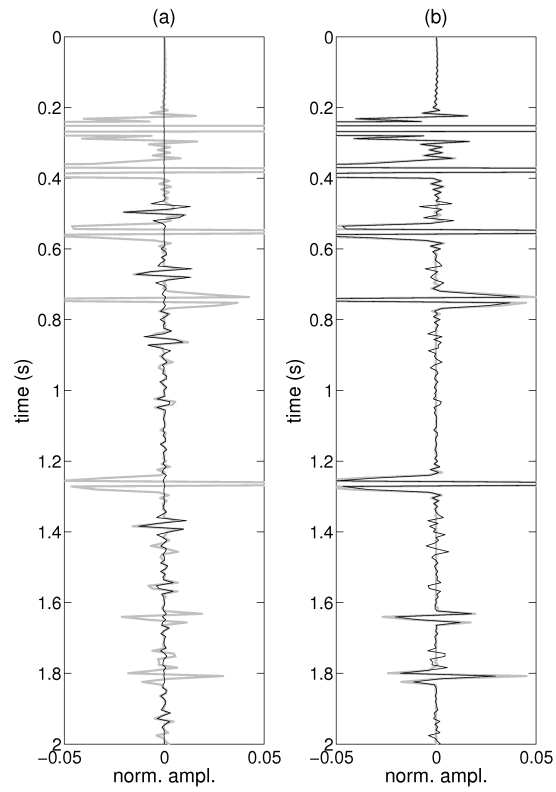


Figure 5.6: Internal multiples estimated from SRI (Eq. 5.35) for a collocated source and receiver. **(a)** Estimate of internal multiples (black) compared to the full scattered wavefield (grey). **(b)** Demultiplied scattered field (black) compared to the primary wavefield (grey).

and SRI (Eq. 5.38), respectively. Fig. 5.7c compares the internal multiples from the two different equations directly. The few small differences are likely to be due to the finite number of frequencies computed using the ISS, which affects the Fourier transformation to the time domain. Fig. 5.8 shows the demultiplied common-shot gathers compared to the directly modeled primary wavefield using the ISS-equation (a) and the SRI-equation (b).

## 5.6 Computational cost

We calculated the computational cost (number of operations) and time for both methods. For a single trace (the case of a collocated source and receiver) with 256 time samples (or 256 frequencies), the number of operations performed by the ISS equation is of the order of  $10^9$ , whereas the SRI equation requires of the order of  $10^5$  operations. For the ISS-version, this results in a computing time of about 245 s ( $\approx 4$  min) on a standard desktop computer without parallelization (about 75 s when parallelized over 4 workers), whereas the SRI-version provides the same result in 0.25 s unparallelized. These differences in computation time or operation count become more important when the source and the receiver are not collocated. To compute the internal multiples between a source at  $\mathbf{x}_s$  and a receiver at  $\mathbf{x}_r \neq \mathbf{x}_s$  both methods require integration over a source boundary and a receiver boundary to cover the additional source receiver pair  $(\mathbf{x}_{s'}, \mathbf{x}_{r'})$  that is involved in the corresponding equations. The computation time for a single source-receiver pair is multiplied by the product of the number of sources  $N_{s'}$  and the number of receivers  $N_{r'}$  on the boundaries. Hence, for  $N_{s'} = 30$  boundary sources and  $N_{r'} = 30$  boundary receivers the computation time is  $N_{s'} \cdot N_{r'} = 900$  times as long in both SRI and ISS as in the single-trace case. Moreover, to compute multiples in a common-shot gather rather than for a single trace, the computation time is multiplied by the number of traces  $N_r$  in the gather. Since both methods are multiplied by the same factor  $N_{s'} \cdot N_{r'} \cdot N_r$ , the relative computational cost remains  $\frac{cost_{ISS}}{cost_{SRI}} \approx \frac{10^9}{10^5} = 10^4$ , while the cost in absolute number of operations taken for a gather is  $\approx 10^4 \cdot N_{s'} \cdot N_{r'} \cdot N_r$ , which for the case above with  $N_{s'} = N_{r'} = 30$  and for a gather of only  $N_r = 20$  traces results in a saving of  $10^4 \cdot 30 \cdot 30 \cdot 20 \approx 10^8$  operations per gather using SRI compared to the ISS. In the 3D case the factors  $N_{s'}$ ,  $N_{r'}$  and  $N_r$  would usually increase by an order of magnitude each, leading to another saving of  $10^3$  operations per gather.

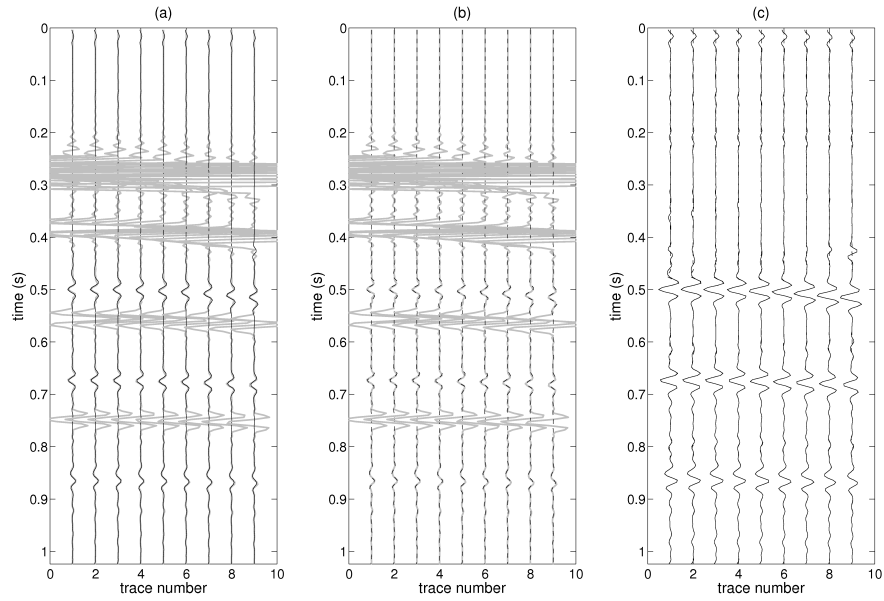


Figure 5.7: Estimate of internal multiples for a common-shot gather. Details on the geometry are described in the main text. **(a)** Full scattered wavefield (grey) versus internal multiples estimated from the ISS equation (black). **(b)** Full scattered wavefield (grey) versus internal multiples estimated from the SRI equation (dashed black). **(c)** Internal multiples from the ISS (solid) versus internal multiples from SRI (dashed).

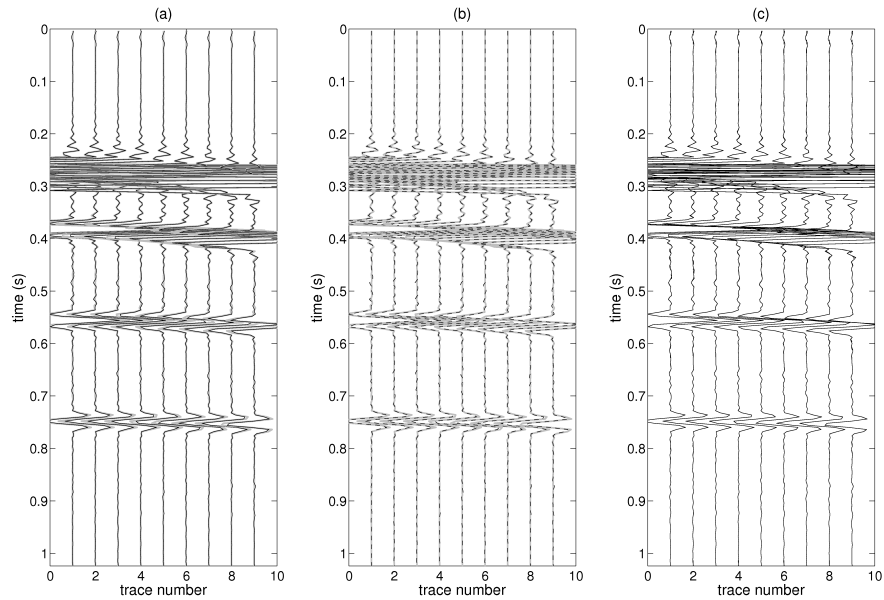


Figure 5.8: Demultiplied scattered wavefield (black) for a common-shot gather as described in the main text using **(a)** the ISS-equation, and **(b)** the SRI-equation compared to the directly modeled primary wavefield. **(c)** Demultiplied data from the ISS (solid) versus demultiplied data from SRI (dashed).

## 5.7 Discussion

Forward scattering theory as described by the Lippmann-Schwinger equation, provides an estimate of the data given the reflectivity of the medium. The equation can be expanded into an infinite series, and Weglein et al. (1997) take one term out of this series, namely the third-order scattering term, as the basis for internal-multiple prediction. However, the third-order scattering term provides only a first-order estimate to first-order internal scattering; higher-order contributions come from the fifth, the seventh, etc. order terms. The fifth-order term has been analysed by Ramirez and Weglein (2005) who showed that indeed amplitude estimates can be improved by taking such higher-order terms into account.

Further, we highlighted that the term used to estimate internal multiples by Weglein et al. (1997) does not contribute physically to the internal multiples in the data. We now discuss why the term still contains information about multiples and how this provides another link to interferometry.

In the forward series (Eq. 5.4), each component ( $\mathbf{D}_1, \mathbf{D}_2, \dots$ ) of the data  $\mathbf{D}$  can be related to a different order of scattering. For example, the first term in Eq. 5.4,  $\mathbf{D}_1 = \mathbf{G}_0 \mathbf{V} \mathbf{G}_0$ , constructs primaries only (first-order scattering). When we replace  $\mathbf{V}$  by the sum of its components ( $\mathbf{V}_1, \mathbf{V}_2, \dots$ ) in the primaries-only term we obtain

$$\mathbf{D}_1 = \mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0 + \mathbf{G}_0 \mathbf{V}_2 \mathbf{G}_0 + \mathbf{G}_0 \mathbf{V}_3 \mathbf{G}_0 + \dots \quad (5.39)$$

The first term on the right-hand side,  $\mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0$ , gives the full data including the multiples (see Eq. 5.8); thus, to obtain primaries only, all of the other terms ( $\mathbf{G}_0 \mathbf{V}_2 \mathbf{G}_0 + \mathbf{G}_0 \mathbf{V}_3 \mathbf{G}_0 + \dots$ ) must cancel the multiples in  $\mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0$ . Considering that

$$\begin{aligned} \mathbf{G}_0 \mathbf{V}_3 \mathbf{G}_0 = & -(\mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0 \mathbf{V}_2 \mathbf{G}_0 + \mathbf{G}_0 \mathbf{V}_2 \mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0 \\ & + \mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0) \end{aligned} \quad (5.40)$$

(see Eq. 5.10) and by appealing to the similarity between the third-order scattering term in Eq. 5.40 ( $\mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0$ ) and the multiple generator in the forward series ( $\mathbf{G}_0 \mathbf{V} \mathbf{G}_0 \mathbf{V} \mathbf{G}_0 \mathbf{V} \mathbf{G}_0$ ; Eq. 5.6), it seems reasonable to infer that the term  $\mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0 \mathbf{V}_1 \mathbf{G}_0$  provides a contribution that at least partly cancels first-order internal multiples in Eq. 5.39. If this term is evaluated on its own as in Eqs. 5.15 and 5.16 it provides an estimate of the internal-multiple energy.

Similarly, in interferometry so-called *pseudo-physical* energy (Löer et al., 2014)



can be used as an estimate for physically scattered waves. Pseudo-physical energy, which has the same kinematics as physical energy but the wrong amplitude and phase, does not contribute to the construction of a physical wavefield but is required in order to obtain correct amplitudes. We mentioned above that the SRI equation can be written as the sum of eight terms, in which different combinations of perturbed and unperturbed (scattered) fields are cross-correlated and convolved. Meles and Curtis (2013) and L er et al. (2014) investigate the different contributions from each term to the final result for a single scatterer case. They show, for example, that the term that provides the physical scattered wave also gives rise to a non-physical contribution from a different stationary point pair on the source and receiver boundaries. The non-physical part is cancelled by contributions from other terms and in fact there is a second term that provides the physical scattered wave. Summing the two physical contributions would, however, result in incorrect (double) amplitudes. The physical contribution from the second term must be cancelled by *pseudo*-physical contributions from other terms. Thus, while pseudo-physical energy arrives at the same time as physical energy, an individual pseudo-physical term does not necessarily provide the correct (physical) amplitude and phase—it is the *sum* of all pseudo-physical contributions that cancels exactly with the extra, redundant physical energy.

If not all of the terms are used, or not all stationary point pairs are spanned by the portions of boundaries included in SRI, pseudo-physical energy may not cancel out and will then remain in the interferometric wavefield estimate, while physical energy may not be constructed at all. This is exactly what happens when SRI is used to estimate internal multiples herein: only one out of eight terms is considered, and boundaries are available only at the surface. Under these conditions only pseudo-physical energy can be obtained, which emulates the kinematics of a physical scattered wavefield but does not provide the correct amplitude and phase information.

Thus, in both the SRI and the ISS approach internal multiples are estimated from pseudo-physical energy and hence cannot be used to estimate and eliminate the internal multiples in the data directly. An adaptive subtraction algorithm is required to find the best fit between estimated and true multiples and eventually attenuate the multiples in the data.

Finally, notice the similarities between the pseudo-physical parts of SRI and the ISS in the numerical examples presented above. This implies that there might be further links between the two representations of scattered fields, for example

an implicit relationship between the other terms in SRI and higher-order contributions in the ISS. Part of our future research will explore this relationship, in the hope that we may gain new insight into each of the two methods by taking information from the other.

## 5.8 Conclusions

We summarize the derivation of the ISS equation for estimating internal multiples and present a second derivation of the same equation starting from SRI. For the first time this provides an explicit relationship between the two domains of inverse scattering and interferometry. The use of pseudo-physical energy for internal-multiple estimation in both methods is highlighted and compared. Using the interferometric perspective, we propose an alternative representation for internal-multiple prediction that takes advantage of convolution and correlation operations in the frequency domain. It allows us to compute internal multiples more efficiently, which is confirmed by a comparison of both representations applied to a synthetic single-trace example and a synthetic common-shot gather.

# Chapter 6

## Discussion

This chapter reflects on open issues in the fields of pseudo-physical energy, multiple scattering and internal multiple prediction in the context of SRI. I expand on topics such as anisotropy, non-physical energy and alternative forms of SRI, establish connections to seismic imaging methods, and suggest potential directions of future research with the aim to shed new light on different aspects in wavefield theory and interpretation using the results from this thesis.

### 6.1 Cancellation of pseudo-physical energy in SRI and the optical theorem

One of the major outcomes of chapter 2 was that SRI generates so-called pseudo-physical energy in the Green's function estimate when boundaries are incomplete. The derivation was shown explicitly for one term of the SRI equation (Eq. 2.7), namely for  $G_S G_S G_S^*$ , which is also used in the prediction of multiply diffracted waves and internal multiple reflections in chapters 3, 4 and 5. We argued that if complete boundaries were used, this energy would be cancelled out by (pseudo-physical) contributions from other scattered terms, but did not provide evidence for this. We will now show that, similar to the cancellation of non-physical energy in standard interferometry, the sum of all these contributions provides exactly the causal scattered wave arrival with scattering amplitude  $A$  on account of the optical theorem.

Table 6.1 shows the travel times and amplitudes of the events generated by the seven scattered terms in Eq. 2.7. They were derived using the following 3D Green's function definition, representing particle velocity resulting from a volume injection rate source and assuming that the diffractor is isotropic and located at

Terms in the integrand of Eq. 2.7	Travel time	Amplitude
$G_0 G_S G_0^*$	$t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$A$
	$-t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$A$
$G_S G_0 G_0^*$	$t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$A$
	$t(\mathbf{x}_s, \mathbf{x}_1) - t(\mathbf{x}_2, \mathbf{x}_s)$	$A$
$G_S G_S G_0^*$	$t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$(2ik)AA$
$G_0 G_0 G_S^*$	$t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$-A^*$
	$t(\mathbf{x}_s, \mathbf{x}_1) - t(\mathbf{x}_2, \mathbf{x}_s)$	$-A^*$
	$-t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$-A^*$
	$-t(\mathbf{x}_s, \mathbf{x}_1) - t(\mathbf{x}_2, \mathbf{x}_s)$	$-A^*$
$G_0 G_S G_S^*$	$t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$(2ik)AA^*$
	$-t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$(2ik)AA^*$
$G_S G_0 G_S^*$	$t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$(2ik)AA^*$
	$t(\mathbf{x}_s, \mathbf{x}_1) - t(\mathbf{x}_2, \mathbf{x}_s)$	$(2ik)AA^*$
$G_S G_S G_S^*$	$t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$	$(2ik)^2 AAA^*$

Table 6.1: Travel times and amplitudes of physical, non-physical and pseudo-physical events provided by the seven scattered terms in Eq. 2.7 using complete boundaries. Non-physical events have non-physical travel times  $-t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$  and  $t(\mathbf{x}_s, \mathbf{x}_1) - t(\mathbf{x}_2, \mathbf{x}_s)$ ; pseudo-physical events have physical travel time  $t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$  but non-physical amplitudes. Amplitudes of non-physical and pseudo-physical waves are supposed to cancel each other when all terms are summed.

the origin  $\mathbf{x}_d = 0$ :

$$\begin{aligned}
G(\mathbf{x}_2, \mathbf{x}) &= G_0(\mathbf{x}_2, \mathbf{x}) + G_S(\mathbf{x}_2, \mathbf{x}) \\
&= \frac{i\omega\rho e^{i\omega k|\mathbf{x}_2 - \mathbf{x}|}}{4\pi |\mathbf{x}_2 - \mathbf{x}|} + \frac{i\omega\rho e^{i\omega k|\mathbf{x}|}}{4\pi |\mathbf{x}|} A \frac{e^{i\omega k|\mathbf{x}_2|}}{|\mathbf{x}_2|}
\end{aligned} \tag{6.1}$$

The amplitudes in Table 6.1 apply when complete boundaries are used. If partial boundaries are used, different events are affected differently depending on the coverage of stationary points; some may not be constructed at all, as was shown in chapter 2. Summing the amplitudes of all terms with travel time  $t(\mathbf{x}_s, \mathbf{x}_1) + t(\mathbf{x}_2, \mathbf{x}_s)$ , i.e., the travel time of the causal scattered wave, gives

$$\begin{aligned}
&A + A - A^* + (2ik)AA^* - (2ik)AA + (2ik)AA^* - (2ik)^2 AAA^* \\
&= A + [A - A^* + (2ik)AA^*] - (2ik)A[A - A^* + (2ik)AA^*]
\end{aligned} \tag{6.2}$$

where  $k = \frac{\omega}{c}$  the wavenumber. For the pseudo-physical energy to cancel, the

expressions in square brackets must equate to zero, hence

$$\begin{aligned}
A - A^* &= -(2jk)AA^* \\
-\frac{1}{2i}[A - A^*] &= k|A|^2 \\
-\Im\{A\} &= k|A|^2
\end{aligned} \tag{6.3}$$

Eq. 6.3 is the optical theorem for isotropic scattering (e.g. Wapenaar et al., 2010b). The optical theorem was derived in quantum physics (Heisenberg, 1943; Glauber and Schomaker, 1953) and in acoustics (Marston, 2001) and describes the conservation of energy in scattering processes. It links the forward scattering amplitude  $A$  of an object to the total scattered energy (the so-called scattering cross section).

Note that Wapenaar et al. (2010b) use a different Green's function definition (representing particle displacement  $u$  rather than particle velocity  $v$ ) when deriving the optical theorem from the cancellation of non-physical energy in standard inter-receiver interferometry. This leads to a representation of the imaginary part of the Green's function ( $\Im(G_u) = \frac{1}{2i}(G_u - G_u^*)$ ) in the interferometric equation rather than the real part ( $\Re(G_v) = \frac{1}{2}(G_v + G_v^*)$ ). The two Green's function definitions are mutually related via  $G_v = i\omega G_u$  (Wapenaar and Fokkema, 2006).

When scattering is not isotropic the scattering amplitude  $A$  depends on the wave vectors  $\mathbf{k}_0$  and  $\mathbf{k}$  of incoming and outgoing energy, respectively, and in this case the optical theorem reads

$$-\Im\{A(\mathbf{k}_0, \mathbf{k}_0)\} = \frac{k}{4\pi} \int |A(\mathbf{k}, \mathbf{k}_0)|^2 d\Omega \tag{6.4}$$

In theory, interferometry also works for anisotropic media; that is, it should construct scattered waves with correct amplitudes regardless of the scattering angle and the degree of anisotropy. However, most applications in interferometry focus on travel time analysis disregarding the information contained in amplitudes and phases of retrieved Green's functions.

Our study of pseudo-physical waves implies that these waves contribute more information on medium parameters and scattering properties than is contained in just the travel times. Moreover, while our analysis focuses on one particular term, other terms may be more useful in other geometries. This potential has not been fully explored yet but could eventually provide novel, valuable information, for example on scattering amplitudes and anisotropic behaviour, not only for

point scatterers but also reflectors in stratified media.

The laboratory data sets used in section 4 might provide a good basis for testing SRI, and in particular the term  $G_S G_S G_S^*$ , on an anisotropic medium. We showed for the single-scatterer case that using larger rod sizes (or different materials) results in anisotropic diffraction move-outs. Using these data it could be investigated whether the anisotropic features of a single diffraction move-out are reconstructed in the interferometric estimate provided by the pseudo-physical term  $G_S G_S G_S^*$ , and if the amplitude estimate can be used to infer information about the forward scattering amplitude of the scatterer. However, for a first test, it would be ideal if the medium was surrounded completely by sources and receivers to eliminate the amplitude reduction caused by partial boundaries.

## 6.2 Alternative forms of SRI

When Halliday and Curtis (2010) derived the link between Oristaglio’s imaging condition and SRI, they used a complex conjugated form of Eq. 2.7. In particular, they use the term  $G_0^* G_0^* G_S$ , which is the complex conjugate of the term  $G_0 G_0 G_S^*$ . As shown in Table 6.1, the term  $G_0 G_0 G_S^*$  provides a physical estimate of the *acausal* scattered wave, a pseudo-physical estimate of the causal scattered wave, and two non-physical contributions. Halliday and Curtis (2010) show that its conjugated version provides a physical estimate of the *causal* scattered wave, a pseudo-physical estimate of the *acausal* scattered wave, and two non-physical contributions.

Similarly, the contributions from all other terms differ in the conjugated form, however, the final SRI result, i.e., the sum of all terms, is identical in both forms. This can be understood by considering the left-hand side of Eq. 2.7, which is a superposition of the causal Green’s function and its conjugate, the *acausal* Green’s function. Since complex conjugation is an associative operation, which acts on both the normal and the conjugated function, it simply reverses the order of the two functions. The superposition of both functions still provides the sum of the causal and the *acausal* part, hence, the homogeneous Green’s function.

Using conjugated forms of the SRI equation we can thus vary the contributions of individual terms while preserving the total result. Of course, this also applies to inter-receiver and inter-source interferometry (e.g. Eqs. 2.3 and 2.4, respectively). Since SRI is a two-step procedure combining inter-source and inter-receiver in-

Complex conjugated forms of SRI
$G(\mathbf{x}', \mathbf{x}_1)G(\mathbf{x}_2, \mathbf{x})G^*(\mathbf{x}', \mathbf{x})$
$G(\mathbf{x}', \mathbf{x}_1)G^*(\mathbf{x}_2, \mathbf{x})G(\mathbf{x}', \mathbf{x})$
$G^*(\mathbf{x}', \mathbf{x}_1)G(\mathbf{x}_2, \mathbf{x})G^*(\mathbf{x}', \mathbf{x})$
$G^*(\mathbf{x}', \mathbf{x}_1)G^*(\mathbf{x}_2, \mathbf{x})G(\mathbf{x}', \mathbf{x})$

Table 6.2: Overview of complex conjugated forms of the SRI equation in short-hand notation.  $G^*(\mathbf{x}', \mathbf{x}_1)G^*(\mathbf{x}_2, \mathbf{x})G(\mathbf{x}', \mathbf{x})$ , for example, stands for the integral  $\frac{4}{(\rho c)^2} \int_{S'} \int_S G^*(\mathbf{x}', \mathbf{x}_1)G^*(\mathbf{x}_2, \mathbf{x})G(\mathbf{x}', \mathbf{x})dSdS'$ . When theoretical requirements of SRI are met, all of them provide approximations to the homogeneous Green's function  $G(\mathbf{x}_2, \mathbf{x}_1) + G^*(\mathbf{x}_2, \mathbf{x}_1)$ .

terferometry, one could apply complex conjugation in just one of the two steps, which provides two more versions of the SRI equation (Table 6.2). Different versions of SRI might be useful in different scenarios of boundary conditions and geometries and it could be the subject of future research to explore the contributions of individual terms and the location of stationary points for physical, non-physical and pseudo-physical energy in conjugated forms of SRI.

### 6.3 SRI and Marchenko autofocusing

Besides its explicit relationship to standard imaging techniques based on the work by Oristaglio (1989), SRI also has a connection to a rather new imaging method called Marchenko imaging, which provides a new perspective on the multiple scattering problem in migration. It is based on single-sided autofocusing (Rose, 2002; Broggini et al., 2011; Wapenaar et al., 2011), a method that allows one to focus a wavefield emitted and recorded on a surface, at any focusing point in the subsurface.

Assuming that after focusing the wavefield propagates outwards again, the focal point becomes a virtual source located inside the medium. The wavefield obtained from the virtual source contains all internal and surface related multiples and provides separate estimates of up- and down-going wavefield components, which are used to define a new imaging condition. The method, described in detail for example in Wapenaar et al. (2014), uses operations similar to interferometry but claims to go *beyond interferometry* as it provides a virtual source at a location in the subsurface without the necessity of a receiver at that location.

Marchenko imaging involves multi-dimensional deconvolution (MDD) of up- and down-going so-called focusing functions ( $f^-$  and  $f^+$ , respectively), obtained from

background Green's functions, and the reflection response of the medium measured at the surface. Wapenaar et al. (2014) show that approximating MDD by cross-correlating the initial estimates of up- and down-going focusing functions yields

$$\begin{aligned}
R(\mathbf{x}_j, \mathbf{x}'_j, t) &\approx C(\mathbf{x}_j, \mathbf{x}'_j, t) \\
&= \int_{\partial D_0} f_{2,0}^+(\mathbf{x}_j, \mathbf{x}''_0, t) f_{2,0}^-(\mathbf{x}_j, \mathbf{x}''_0, -t) d\mathbf{x}_0 \\
&= - \int_{\partial D_0} \int_{\partial D_0} G_d(\mathbf{x}''_0, \mathbf{x}'_j, t) R(\mathbf{x}''_0, \mathbf{x}_0, -t) G_d(\mathbf{x}_0, \mathbf{x}_j, t) d\mathbf{x}_0 d\mathbf{x}''_0
\end{aligned} \tag{6.5}$$

The authors point out that the resulting double integral is a form of SRI. In fact, from the analysis in chapter 2 we know that here the term  $G_0 G_0 G_S^*$  of the SRI equation is used, since  $G_d = G_0$  represents the direct or reference Green's function and  $R = G_S$  is the scattered wavefield or the reflection response measured at the surface. For the geometry considered throughout chapter 2 (boundaries at the surface, scatterer below the source-receiver pair), the term  $G_0 G_0 G_S^*$  provides the *acausal* scattered wavefield. However, in the geometry of Wapenaar et al. (2014) the scatterer is located above the (virtual) source-receiver pair and thus Eq. 6.5 provides an estimate of the *causal* scattered field. Moreover, the amplitude analysis of the SRI term (Table 6.1) suggests that the scattered field constructed from Eq. 6.5 has a pseudo-physical amplitude, similarly to the event constructed from  $G_0 G_0 G_S^*$ .

Insights from SRI could provide a new perspective on amplitude estimation in imaging techniques and help to interpret erroneous amplitudes. Further, the link between SRI and Marchenko imaging by MDD suggests that a deconvolutional version of SRI exists, as also discussed in Entwistle et al. (2015). Convolutional and deconvolutional techniques have successfully been used in inter-source and inter-receiver interferometry (e.g. Slob and Wapenaar, 2007; Vasconcelos and Snieder, 2008), however, they have not been analysed and tested in detail for SRI. Deconvolutional interferometry works better in dissipative media, where cross-correlational methods often show deficiencies; hence, using deconvolution could be advantageous for SRI when used in highly dissipative environments or over large distances.



## 6.4 Primary move-out estimation and interferometry

The method used for primary move-out estimation from multiply diffracted wavefields presented in chapter 3, section 3.3.1, is in fact a demonstration of how physical information about a scatterer can be obtained from non-physical energy generated in interferometry. Non-physical energy from the cross terms of diffracted, refracted and reflected waves has been used in previous studies to estimate scatterer locations (Harmankaya et al., 2013; Kaslilar et al., 2014) and analyse velocities (Mikesell et al., 2009; King and Curtis, 2011). Here, we use travel times of non-physical events to discriminate individual primary diffraction curves from a multiply scattered wavefield. The method described is based on cross-correlation of two common-source gathers or two common-receiver gathers and resembles standard interferometry methods: taking the case of two CSGs, the data in each gather represents the scattered wavefield emitted by one source recorded at a surface boundary of receivers. Cross-correlating the two wavefields and summing over receiver positions corresponds exactly to the two steps in inter-source interferometry. For the case of two CRGs the method is equivalent to inter-receiver interferometry.

Notice, however, that when only the scattered components of the wavefields ( $G_S$ ) are considered, non-physical energy is generated in inter-source or inter-receiver interferometry (Snieder et al., 2008). The non-physical event arrives at travel time  $t(\mathbf{x}_d, \mathbf{x}_1) - t(\mathbf{x}_d, \mathbf{x}_2)$ , where  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are the locations of the two sources and  $\mathbf{x}_d$  is the location of a diffractor. Given that  $\mathbf{x}_1$  and  $\mathbf{x}_2$  do not change, the travel time depends solely on the location of the diffractor and hence each diffractor in the medium provides one non-physical event. The arrival times of non-physical events are equivalent to the time shifts under which the two gathers are most alike, i.e., the time shifts for which one move-out arrives at the same time in both gathers. This information is used by the automated algorithm to isolate individual move-outs associated with different time-shifts from the CSG or CRG. Note that stationary, non-physical events also occur when a primary is cross-correlated with a multiple that has the same last scatterer, i.e., the same fingerprint in the CSG. These events, however, have much lower amplitudes and are automatically filtered out by the algorithm presented in chapter 3.

The geometry considered in chapter 3 differs from the geometry typically used in interferometry since the two sources are not spatially separated from the receiver boundary. This violates the assumption made in the monopole approximation

that all raypaths are perpendicular to the boundary, which is typically the case if the receiver boundary is in the far-field of the central pair of sources. However, as we have demonstrated in chapter 2, for scattered energy this requirement is less strict, because scattered waves travel from the source to the scatterer first before reaching the receiver boundary; thus the angle of incoming energy at the boundary depends on the scatterer location rather than on the source location.

## 6.5 Open issues in internal multiple prediction from SRI

The proposed equation to predict internal multiples more efficiently has not been implemented yet in a code that can operate on more than four worker processors simultaneously in order to deal with large quantities of data in a reasonable time. Thus the method remains to be tested on real data sets with larger offsets and longer recording times. Also, we have not tested the performance of the algorithm on more complex models comprising dipping layers or synclines. However, since we could demonstrate its equivalence to the ISS scheme, we assume that it will perform equally well in different scenarios at lower computational cost.

Both methods still suffer from the fact that amplitudes of internal multiples are not predicted perfectly (although quite accurately, for example in the scheme by Ramirez and Weglein, 2005), because the generated multiples are in fact pseudo-physical events. In general, the relationship between the amplitude of a pseudo-physical wave and the scattering amplitude of an object discussed in section 2.3 and above could be of interest for “true amplitude” internal multiple prediction. Note, however, that the analysis in section 2.3 does not account for a multiple scattering scenario: multiple scatterers generate multiple (multiply scattered) non-physical and pseudo-physical events and the interpretation of the corresponding amplitudes becomes more complicated. Also, limited source and receiver boundaries affect the amplitudes and need to be taken into account. Nevertheless, further investigation of this topic would enhance our understanding of how amplitudes of predicted multiples can be improved and of how they are related to parameters of the medium.

# Chapter 7

## Conclusions

This thesis investigated the ability of source-receiver interferometry (SRI) to construct and analyse multiply scattered wavefields in increasingly complex scenarios when source and receiver boundaries are only partially available, as in most practical applications.

For the case of a single point diffractor, I examined each of the eight terms of the SRI equation individually, both kinematically and dynamically, with respect to their contribution of physical and non-physical energy to the final interferometric estimate. I discovered that one term of the SRI equation generates a new type of non-physical energy called pseudo-physical energy, which emulates the kinematics but not the dynamics of a physically scattered wave. The integrand of this term is stationary for any source-receiver pair on the boundary and provides a robust pseudo-physical estimate of the causal scattered wave even when very limited source and receiver boundaries are used. When complete boundaries are used, pseudo-physical energy is produced by several terms of the SRI equation and I demonstrated that these pseudo-physical contributions cancel each other on account of the optical theorem. Vice versa, the cancellation of pseudo-physical energy in SRI proves the optical theorem. I identified one term of the SRI equation that is also used in seismic imaging and that constructs the physical acausal scattered field. I showed that for this term partial boundaries are in fact a positive advantage, since they omit the stationary points of non-physical events and thus suppress non-physical events in the constructed wavefield. Using only these two terms, a robust estimate of the causal and acausal scattered field was obtained, even if boundaries were limited to linear surface arrays and only sparsely populated by sources and receivers.

The results from the single-scatterer case were applied in a multiple scattering scenario, where the pseudo-physical term was used to predict travel times and scattering paths of multiply diffracted waves. I developed a fully automated,

purely data-driven algorithm that estimates primary diffraction move-outs, identifies secondary waves and uses these events in the pseudo-physical term to predict higher-order multiply diffracted waves. The method to estimate individual primary move-outs, which emerged as a by-product of this algorithm, relies on cross-correlation of data gathers. Unlike other schemes that use, for example, semblance analysis of travel time curves, the new method does not require a model of the background medium and works also for heterogeneous background media. The algorithm was successfully tested on synthetic data sets obtained from numerical models containing multiple isotropic point scatterers, contaminated with different noise-to-signal ratios.

Finite-difference modeling experiments showed the deficiencies of the method when finite-sized or anisotropic scatterers were used. The modeling results helped to design a laboratory scattering experiment and to obtain a real data set to test the automatic algorithm on. In the laboratory experiment steel rods with sub-wavelength diameter were used as isotropic scatterers in a homogeneous, quasi-acoustic gel. Despite the presence of spurious scattered energy in the data, introduced by the experimental setup itself, the algorithm estimated primary move-outs quite reliably for up to four scatterers, and I also presented examples of correctly identified secondary and tertiary arrivals.

The construction of high-order multiply diffracted waves from low-order diffracted waves naturally lead to an expansion to reflecting media. Using the same pseudo-physical term of the SRI equation as in the multiply diffracted wavefield I derived an equation that predicts internal multiples in reflection seismic data. In linear migration schemes internal multiples typically cause problems since they introduce false reflectors in the seismic image. Hence, they need to be identified and removed from the data prior to migration. I showed that the internal multiple equation based on SRI is equivalent to an existing internal multiple equation based on the inverse-scattering series (ISS). This provided an explicit link between the two quite different concepts of SRI and the ISS. From the insight gained from SRI I inferred that amplitudes of estimated multiples derived from both SRI and the ISS must be incorrect because they are in fact pseudo-physical. The equation also has the downside that it is computationally expensive, which can be a limiting factor especially in three-dimensional applications. Again, using the SRI perspective I derived an alternative representation of the internal multiple equation that computes internal multiples more efficiently using computationally cheap cross-correlation and convolution, just as in interferometry.

Following up the results of this thesis, future work could explore whether anisotropic features are preserved in non-physical or pseudo-physical interferometric estimates and how this would complement the predominantly kinematic wavefield analysis of most interferometric methods. This could also provide insight into how the automatic algorithm for diffracted wave identification could be made applicable to anisotropic wavefields and finite-sized scatterers. Finally, both standard linear migration schemes, which depend on multiple removal, and novel non-linear imaging techniques, which integrate internal multiples into the migration scheme, could benefit from a deeper analysis of the link between SRI and the ISS in the context of internal multiple prediction.

# Bibliography

- F.V. Araújo, A.B. Weglein, P.M. Carvalho, and R.H. Stolt. Inverse scattering series for multiple attenuation: An example with surface and internal multiples. Technical report, Society of Exploration Geophysicists, Tulsa, OK (United States), 1994.
- A. Aubry and A. Derode. Multiple scattering of ultrasound in weakly inhomogeneous media: Application to human soft tissue. *Journal of the Acoustical Society of America*, 129:225–233, 2010.
- A. Bakulin and R. Calvert. The virtual source method: Theory and case study. *Geophysics*, 71(4):SI139–SI150, 2006.
- J. Behura and F. Forghani. A practical approach to prediction of internal multiples and ghosts. In *SEG Technical Program Expanded Abstracts 2012*, pages 1–5. Society of Exploration Geophysicists, 2012. doi: 10.1190/segam2012-1269.1.
- G.D. Bensen, M.H. Ritzwoller, and N.M. Shapiro. Broadband ambient noise surface wave tomography across the united states. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 113(B5), 2008. doi: 10.1029/2007JB005248.
- A.J. Berkhout and D.J. Verschuur. Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations. *Geophysics*, 62(5):1586–1595, 1997.
- A. Berkovitch, I. Belfer, Y. Hassin, and E. Landa. Diffraction imaging by multifocusing. *Geophysics*, 74:WCA75–WCA81, 2009.
- J.M. Bordier, M. Fink, A. Le Brun, and F. Cohen-Tenoudji. The influence of multiple scattering in incoherent ultrasonic inspection of course grain stainless steel. *Ultrasonics Symposium*, 2:803–808, 1991.
- M. Born and E. Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 1999.
- L. M. Brekhovskikh and Y. P. Lysanov. *Fundamentals of Ocean Acoustics*. Springer-Verlag, New York, 2003.

- F. Broggini, R. Snieder, and K. Wapenaar. Connection of scattering principles: Focusing the wavefield without source or receiver. In *SEG Technical Program Expanded Abstracts 2011*. Society of Exploration Geophysicists, 2011.
- M. Campillo and A. Paul. Long-range correlations in the diffuse seismic coda. *Science*, 299(5606):547–549, 2003.
- J.F. Claerbout. Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, 33:264, 1968.
- J.F. Claerbout. *Imaging the earth's interior*. Blackwell Scientific Publications, New York, 1985.
- A. Curtis and D. Halliday. Source-receiver wave field interferometry. *Physical Review E*, 81(4):046601, 2010.
- A. Curtis, P. Gerstoft, H. Sato, R. Snieder, and K. Wapenaar. Seismic interferometry—turning noise into signal. *The Leading Edge*, 25:1082–1092, 2006.
- A. Curtis, H. Nicolson, D. Halliday, J. Trampert, and B. Baptie. Virtual seismometers in the subsurface of the earth from seismic interferometry. *Nature Geoscience*, 2(10):700–704, 2009.
- A. Curtis, Y. Behr, E. Entwistle, E. Galetti, J. Townend, and S. Bannister. The benefit of hindsight in observational science: Retrospective seismological observations. *Earth and Planetary Science Letters*, 345:212–220, 2012.
- J. De Rosny and P. Roux. Multiple scattering in a reflecting cavity: Application to fish counting in a tank. *Journal of the Acoustical Society of America*, 109:2587–2597, 2001.
- A. Derode, E. Larose, M. Campillo, and M. Fink. How to estimate the green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves. *Applied Physics Letters*, 83(15):3054–3056, 2003a.
- A. Derode, E. Larose, M. Tanter, J. De Rosny, A. Tourin, M. Campillo, and M. Fink. Recovering the Green's function from field-field correlations in an open scattering medium (1). *The Journal of the Acoustical Society of America*, 113:2973, 2003b.
- D. Draganov, K. Heller, and R. Ghose. Monitoring CO<sub>2</sub> storage using ghost reflections retrieved from seismic interferometry. *International Journal of Greenhouse Gas Control*, 11:S35–S46, 2012.

- C. Duguid, D. Halliday, and A. Curtis. Source-receiver interferometry for seismic wavefield construction and ground-roll removal. *The Leading Edge*, 30(8):838–843, 2011.
- T.L. Duvall, S.M. Jefferies, J.W. Harvey, and M.A. Pomerantz. Time–distance helioseismology. *Nature*, 362:430–432, 1993.
- E. Entwistle, A. Curtis, E. Galetti, B. Baptie, and G.A. Meles. Constructing new seismograms from old earthquakes: Retrospective seismology at multiple length scales. *Journal of Geophysical Research: Solid Earth*, 120(4):2466–2490, 2015.
- J.H. Faccipieri, D. Rueda Serrano, L.-J. Gelius, and M. Tygel. Recovering diffractions in CRS stacked sections. *First Break*, 31:65–69, 2013.
- A. Ferretti, C. Prati, and F. Rocca. Permanent scatterers in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 39:8–20, 2001.
- M. Fink, G. Montaldo, and M. Tanter. Time-reversal acoustics in biomedical engineering. *Annual Review of Biomedical Engineering*, 5:465–497, 2003.
- C. Fleury. Increasing illumination and sensitivity of reverse-time migration with internal multiples. *Geophysical Prospecting*, 61(5):891–906, 2013.
- C. Fleury and I. Vasconcelos. Imaging condition for nonlinear scattering-based imaging: Estimate of power loss in scattering. *Geophysics*, 77:S1–S18, 2012.
- L. Foldy. The multiple scattering of waves. I. General theory of isotropic scattering by randomly distributed scatterers. *Physical Review*, 67:107–119, 1945.
- K.G. Foote. Acoustic scattering by marine organisms. In *Encyclopaedia of Ocean Sciences*, pages 62–70. Academic Press, London, 2<sup>nd</sup> edition, 2008.
- H. Friedrich. *Theoretical Atomic Physics*. Springer-Verlag, New York, 2006.
- E. Galetti and A. Curtis. Generalised receiver functions and seismic interferometry. *Tectonophysics*, 532-535:1–26, 2012.
- E. Galetti, D. Halliday, and A. Curtis. A simple and exact acoustic wavefield modelling code for data processing, imaging and interferometry applications. *Geophysics*, 78(6):F17–F27, 2013.



- L.S. Gao, L.C. Lee, N. Biswas, and K. Aki. Comparison of the effects between single and multiple scattering on coda waves for local earthquakes. *Bulletin of the Seismological Society of America*, 73:377–388, 1983.
- R. Glauber and V. Schomaker. The theory of electron diffraction. *Physical Review*, 89(4):667, 1953.
- J. Groenenboom and R. Snieder. Attenuation, dispersion, and anisotropy by multiple scattering of transmitted waves through distributions of scatterers. *Journal of the Acoustical Society of America*, 98:3482–3492, 1995.
- D. Halliday and A. Curtis. Seismic interferometry of scattered surface waves in attenuative media. *Geophysical Journal International*, 178:419–446, 2009.
- D. Halliday and A. Curtis. An interferometric theory of source-receiver scattering and imaging. *Geophysics*, 75(6):SA95–SA103, 2010.
- D. Halliday, A. Curtis, J.O.A. Robertsson, and D.-J. van Manen. Interferometric surface-wave isolation and removal. *Geophysics*, 72(5):A69–A73, 2007.
- D. Halliday, A. Curtis, P. Vermeer, C. Strobbia, A. Glushchenko, D. J. van Manen, and J. O. Robertsson. Interferometric ground-roll removal: Attenuation of scattered surface waves in single-sensor data. *Geophysics*, 75(2):SA15–SA25, 2010.
- D. Halliday, P. Bilsby, L. West, E. Kragh, and J. Quigley. Scattered ground-roll attenuation using model-driven interferometry. *Geophysical Prospecting*, 63(1):116–132, 2015.
- D. Hampson. Inverse velocity stacking for multiple elimination. In *SEG Technical Program Expanded Abstracts 1986*. Society of Exploration Geophysicists, 1986.
- U. Harmankaya, A. Kaslilar, J. Thorbecke, K. Wapenaar, and D. Draganov. Locating near-surface scatterers using non-physical scattered waves resulting from seismic interferometry. *Journal of Applied Geophysics*, 91:66–81, 2013.
- W. Heisenberg. Die beobachtbaren Größen in der Theorie der Elementarteilchen. *Zeitschrift für Physik*, 120:513–538, 1943.
- T. Hong and W. Menke. Tomographic investigation of the wear along the San Jacinto fault, southern California. *Physics of the Earth and Planetary Interiors*, 155:236–248, 2006.

- B. Hung and M. Wang. Internal demultiple methodology without identifying the multiple generators. In *SEG Technical Program Expanded Abstracts 2012*, pages 1–5. Society of Exploration Geophysicists, 2012.
- J. Hunziker, E. Slob, Y. Fan, R. Snieder, and K. Wapenaar. Two-dimensional controlled-source electromagnetic interferometry by multidimensional deconvolution: Spatial sampling aspects. *Geophysical Prospecting*, 60(5):974–994, 2012.
- M.F. Insana, R.F. Wagner, D.G. Brown, and T.J. Hall. Describing small-scale structure in random media using pulse-echo ultrasound. *Journal of the Acoustical Society of America*, 87:179–192, 1990.
- H. Jakubowicz. Wave equation prediction and removal of interbed multiples. In *SEG Technical Program Expanded Abstracts 1998*. Society of Exploration Geophysicists, 1998.
- Z. Jiang, J. Sheng, J. Yu, G.T. Schuster, and B.E. Hornby. Migration methods for imaging different-order multiples. *Geophysical Prospecting*, 55(1):1–19, 2007.
- X. Jixiang, B.F. McLean, and S. Xuejuan. A near-surface seismic scattered wave separation method. *Petroleum Exploration and Development*, 41(6):771–777, 2014.
- A. Kaslilar, U. Harmankaya, K. van Wijk, K. Wapenaar, and D. Draganov. Estimating location of scatterers using seismic interferometry of scattered rayleigh waves. *Near Surface Geophysics*, 12(6):721–730, 2014.
- S. Keydar, E. Landa, B. Gurevich, and B. Gelchinsky. Multiple prediction using wavefront characteristics of primary reflections. In *59<sup>th</sup> EAGE Conference & Exhibition*, 1997.
- V. Khaidukov, E. Landa, and T. Moser. Diffraction imaging by focusing-defocusing: An outlook on seismic superresolution. *Geophysics*, 69:1478–1490, 2004.
- S. King and A. Curtis. Velocity analysis using both reflections and refractions from seismic interferometry. *Geophysics*, 76:SA83–SA96, 2011.
- S. King and A. Curtis. Suppressing nonphysical reflections in Green’s function estimates using source-receiver interferometry. *Geophysics*, 77:Q15, 2012.

- S. King, A. Curtis, and T. Poole. Interferometric velocity analysis using physical and non-physical energy. *Geophysics*, 76:SA34–SA49, 2011.
- V. Korneev and A. Bakulin. On the fundamentals of the virtual source method. *Geophysics*, 71(3):A13–A17, 2006.
- E. Landa, V. Shtivelman, and B. Gelchinsky. A method for detection of diffracted waves on common-offset sections. *Geophysical Prospecting*, 35:359–374, 1987.
- E. Larose, A. Derode, D. Clorennec, L. Margerin, and M. Campillo. Passive retrieval of rayleigh waves in disordered elastic media. *Physical Review E*, 72(4):046607, 2005.
- E. Larose, J. de Rosny, L. Margerin, D. Anache, P. Gouedard, M. Campillo, and B. van Tiggelen. Observation of multiple scattering of kHz vibrations in a concrete structure and application to monitoring weak changes. *Physical Review E*, 73:016609, 2006.
- H. Liang, C. Ma, and A.B. Weglein. General theory for accommodating primaries and multiples in internal multiple algorithm: Analysis and numerical tests. In *SEG Technical Program Expanded Abstracts 2013*. Society of Exploration Geophysicists, 2013.
- B.A. Lippmann and J. Schwinger. Variational principles for scattering processes. I. *Physical Review*, 79(3):469, 1950.
- O.I. Lobkis and R.L. Weaver. On the emergence of the Green’s function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America*, 110(6):3011–3017, 2001.
- K. Lör, G.A. Meles, A. Curtis, and I. Vasconcelos. Diffracted and pseudo-physical waves from spatially limited arrays using source–receiver interferometry (SRI). *Geophysical Journal International*, 196(2):1043–1059, 2014.
- K. Lör, G.A. Meles, and A. Curtis. Automatic identification of multiply diffracted waves and their ordered scattering paths. *The Journal of the Acoustical Society of America*, 137(4):1834–1845, 2015.
- C. Ma and A.B. Weglein. Including higher-order inverse scattering series terms to address a serious shortcoming/problem of the internal-multiple attenuator: Exemplifying the problem and its resolution. In *SEG Technical Program Expanded Abstracts 2014*. Society of Exploration Geophysicists, 2014.

- A.E. Malcolm, B. Ursin, and V. Maarten. Seismic imaging and illumination with internal multiples. *Geophysical Journal International*, 176(3):847–864, 2009.
- P.L. Marston. Generalized optical theorem for scatterers having inversion symmetry: Applications to acoustic backscattering. *The Journal of the Acoustical Society of America*, 109(4):1291–1295, 2001.
- K. Mehta, R. Snieder, R. Calvert, and J. Sheiman. Acquisition geometry requirements for generating virtual-source data. *The Leading Edge*, 27:620–629, 2008.
- G.A. Meles and A. Curtis. Physical and non-physical energy in scattered wave interferometry. *Journal of the Acoustical Society of America*, 133:3790–3801, 2013.
- G.A. Meles and A. Curtis. Fingerprinting ordered diffractions in multiply-scattered waves. *Geophysical Journal International*, 198:1701–1713, 2014a.
- G.A. Meles and A. Curtis. Discriminating physical and non-physical diffracted energy in source-receiver interferometry. *Geophysical Journal International*, 197:1642–1659, 2014b.
- G.A. Meles, K. L  er, M. Ravasi, A. Curtis, and C.A. da Costa Filho. Internal multiple prediction and removal using Marchenko autofocusing and seismic interferometry. *Geophysics*, 80(1):A7–A11, 2014.
- D. Mikesell, K. van Wijk, A. Calvert, and M. Haney. The virtual refraction: Useful spurious energy in seismic interferometry. *Geophysics*, 74:A13–A17, 2009.
- D. Miller, M. Oristaglio, and G. Beylkin. A new slant on seismic imaging: Classical migration and integral geometry. *Geophysics*, 52:943–964, 1987.
- G. Montaldo, M. Tanter, and M. Fink. Revisiting iterative time reversal processing: Application to detection of multiple targets. *Journal of the Acoustical Society of America*, 115:776–784, 2004.
- H. Nicolson, A. Curtis, B. Baptie, and E. Galetti. Seismic interferometry and ambient noise tomography in the British Isles. *Proceedings of the Geologists’ Association*, 123(1):74–86, 2012.
- M.L. Oristaglio. An inverse scattering formula that uses all the data. *Inverse Problems*, 5:1097, 1989.

- C. Pacheco and R. Snieder. Time-lapse travelttime change of singly scattered acoustic waves. *Geophysical Journal International*, 165(2):485–500, 2006.
- U. Papziner and K.P. Nick. Automatic detection of hyperbolas in georadargrams by slant-stack processing and migration. *First Break*, 16:219–223, 1998.
- O.V. Poliannikov. Retrieving reflections by source-receiver wavefield interferometry. *Geophysics*, 76:SA1–SA8, 2011.
- O.V. Poliannikov, S. Rondenay, and L. Chen. Interferometric imaging of the underside of a subducting crust. *Geophysical Journal International*, 189:681–690, 2012.
- C. Prada and M. Fink. Eigenmodes of the time-reversal operator: A solution to selective focusing in multiple-target media. *Wave Motion*, 20:151–163, 1994.
- C. Prada, E. Kerbrat, D. Cassereau, and M. Fink. Time reversal techniques in ultrasonic nondestructive testing of scattering media. *Inverse Problems*, 18:1761–1773, 2002.
- R.G. Pratt. Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. *Geophysics*, 64(3):888–901, 1999.
- A.C. Ramirez and A. Weglein. An inverse scattering internal multiple elimination method: Beyond attenuation, a new algorithm and initial tests. In *SEG Technical Program Expanded Abstracts 2011*, pages 2115–2118. Society of Exploration Geophysicists, 2005.
- M. Ravasi and A. Curtis. Non-linear scattering based imaging in elastic media: Theory, theorems and imaging conditions. *Geophysics*, 78(3):S137–S155, 2013.
- M. Ravasi, I. Vasconcelos, and A. Curtis. Beyond conventional migration: Non-linear elastic subsalt imaging with transmissions and two-sided illumination. *Geophysical Journal International*, 198(2):1173–1185, 2014.
- J.H. Rose. Single-sided autofocusing of sound in layered materials. *Inverse Problems*, 18(6):1923–1934, 2002.
- K.G. Sabra, P. Gerstoft, P. Roux, W.A. Kuperman, and M.C. Fehler. Extracting time-domain green’s function estimates from ambient seismic noise. *Geophysical Research Letters*, 32(3), 2005a.

- K.G. Sabra, P. Gerstoft, P. Roux, W.A. Kuperman, and M.C. Fehler. Surface wave tomography from microseisms in southern california. *Geophysical Research Letters*, 32(14), 2005b.
- K.G. Sabra, S. Conti, P. Roux, and W.A. Kuperman. Passive in vivo elastography from skeletal muscle noise. *Applied physics letters*, 90(19):194101, 2007.
- P. Sava and I. Vasconcelos. Extended imaging conditions for wave-equation migration. *Geophysical Prospecting*, 59(1):35–55, 2011.
- G.T. Schuster. *Seismic Interferometry*. Cambridge University Press, 2009.
- C. Sens-Schönfelder and U. Wegler. Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia. *Geophysical research letters*, 33(21), 2006.
- N.M. Shapiro, M. Campillo, L. Stehly, and M.H. Ritzwoller. High-resolution surface-wave tomography from ambient seismic noise. *Science*, 307(5715):1615–1618, 2005.
- R. E. Sheriff and L. P. Geldart. *Exploration Seismology*. Cambridge University Press, 2<sup>nd</sup> edition, 1995.
- E. Slob and K. Wapenaar. Electromagnetic Green’s functions retrieval by cross-correlation and cross-convolution in media with losses. *Geophysical Research Letters*, 34(5), 2007. doi: 10.1029/2006GL029097.
- R. Snieder. Imaging and averaging in complex media. *Diffuse Waves in Complex Media*, 531(NATO Science Series - Series C: Mathematical and Physical Sciences):405–454, 1999.
- R. Snieder. Coda wave interferometry and the equilibration of energy in elastic media. *Physical review E*, 66(4):046615, 2002.
- R. Snieder. Extracting the Green’s function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, 69:046610, 2004a.
- R. Snieder. *A Guided Tour of Mathematical Methods for the Physical Sciences*. Cambridge University Press, 2004b.
- R. Snieder and C. Fleury. Cancellation of spurious arrivals in Green’s function retrieval of multiple scattered waves. *Journal of the Acoustical Society of America*, 128:1598–1605, 2010.

- R. Snieder, K. Wapenaar, and K. Larner. Spurious multiples in seismic interferometry of primaries. *Geophysics*, 71(4):SI111–SI124, 2006.
- R. Snieder, K. van Wijk, M. Haney, and R. Calvert. Cancellation of spurious arrivals in Green’s function extraction and the generalized optical theorem. *Physical Review E*, 78(3):036606, 2008.
- T.K. Stanton. Multiple scattering with applications to fish-echo processing. *Journal of the Acoustical Society of America*, 73:1164–1169, 1982.
- L. Stehly, B. Fry, M. Campillo, N.M. Shapiro, J. Guilbert, L. Boschi, and D. Giardini. Tomography of the Alpine region from observations of seismic ambient noise. *Geophysical Journal International*, 178(1):338–350, 2009.
- H. Tadayyon, A. Sadeghi-Naini, L. Wirtzfeld, F.C. Wright, and G. Czarnota. Quantitative ultrasound characterization of locally advanced breast cancer by estimation of its scatterer properties. *Medical Physics*, 41:0129031, 2014.
- F. Ten Kroode. Prediction of internal multiples. *Wave Motion*, 35(4):315–338, 2002.
- J. Thorbecke and K. Wapenaar. On the relation between seismic interferometry and the migration resolution function. *Geophysics*, 72(6):T61–T66, 2007.
- R. Thorson and J.F. Claerbout. Velocity-stack and slant-stack stochastic inversion. *Geophysics*, 50:2727–2741, 1985.
- D.J. van Manen, J.O.A. Robertsson, and A. Curtis. Modeling of wave propagation in inhomogeneous media. *Physical Review Letters*, 94(16):164301, 2005.
- D.J. van Manen, A. Curtis, and J.O.A. Robertsson. Interferometric modeling of wave propagation in inhomogeneous elastic media using time reversal and reciprocity. *Geophysics*, 71(4):SI47–SI60, 2006.
- I. Vasconcelos. Generalized representations of perturbed fields—applications in seismic interferometry and migration. In *SEG Technical Program Expanded Abstracts 2008*, pages 2927–2931, 2008.
- I. Vasconcelos. Source-receiver, reverse-time imaging of dual-source, vector-acoustic seismic data. *Geophysics*, 78:WA123–WA145, 2013.
- I. Vasconcelos and R. Snieder. Interferometry by deconvolution: Part 1—Theory for acoustic waves and numerical examples. *Geophysics*, 73(3):S115–S128, 2008.

- I. Vasconcelos, R. Snieder, and H. Douma. Representation theorems and Green's function retrieval for scattering in acoustic media. *Physical Review E*, 80(3):036605, 2009.
- I. Vasconcelos, P. Sava, and H. Douma. Nonlinear extended images via image-domain interferometry. *Geophysics*, 75(6):SA105–SA115, 2010.
- D.J. Verschuur. *Seismic multiple removal techniques: past, present and future*. EAGE publications bv, Houten, The Netherlands, Revised edition, 2013.
- D.J. Verschuur, A.J. Berkhout, and K. Wapenaar. Adaptive surface-related multiple elimination. *Geophysics*, 57(9):1166–1177, 1992.
- K. Wapenaar. Synthesis of an inhomogeneous medium from its acoustic transmission response. *Geophysics*, 68(5):1756–1759, 2003.
- K. Wapenaar. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation. *Physical Review Letters*, 93:254301, 2004.
- K. Wapenaar and J. Fokkema. Green's function representations for seismic interferometry. *Geophysics*, 71(4):SI33–SI46, 2006.
- K. Wapenaar, J. Fokkema, and R. Snieder. Retrieving the Green's function in an open system by cross correlation: A comparison of approaches. *Journal of the Acoustical Society of America*, 118:2783–2786, 2005.
- K. Wapenaar, D. Draganov, R. Snieder, X. Campman, and A. Verdel. Tutorial on seismic interferometry. Part 1—basic principles and applications. *Geophysics*, 75:A195–A209, 2010a.
- K. Wapenaar, E. Slob, and R. Snieder. On seismic interferometry, the generalized optical theorem, and the scattering matrix of a point scatterer. *Geophysics*, 75(3):SA27–SA35, 2010b.
- K. Wapenaar, E. Slob, R. Snieder, and A. Curtis. Tutorial on seismic interferometry. Part 2—underlying theory and new advances. *Geophysics*, 75:A211–A227, 2010c.
- K. Wapenaar, F. Broggini, and R. Snieder. A proposal for model-independent 3D wave field reconstruction from reflection data. In *SEG Technical Program Expanded Abstracts 2011*. Society of Exploration Geophysicists, 2011.



- K. Wapenaar, J. Thorbecke, J. van der Neut, F. Broggini, E. Slob, and R. Snieder. Marchenko imaging. *Geophysics*, 79(3):WA39–WA57, 2014.
- A.B. Weglein, F. Gasparotto, P.M. Carvalho, and R.H. Stolt. An inverse-scattering series method for attenuating multiples in seismic reflection data. *Geophysics*, 62(6):1975–1989, 1997.
- A.B. Weglein, F.V. Araújo, P.M. Carvalho, R.H. Stolt, K.H. Matson, R.T. Coates, D. Corrigan, D.J. Foster, S.A. Shaw, and H. Zhang. Inverse scattering series and seismic exploration. *Inverse problems*, 19(6):R27, 2003.
- R. Wu and K. Aki. Introduction: Seismic wave scattering in three-dimensionally heterogeneous earth. In *Scattering and Attenuations of Seismic Waves, Part I*. Birkhäuser, Basel, 1988.
- Y. Yang, M.H. Ritzwoller, A.L. Levshin, and N.M. Shapiro. Ambient noise Rayleigh wave tomography across Europe. *Geophysical Journal International*, 168(1):259–274, 2007.
- Ö. Yilmaz. *Seismic data analysis*. Society of Exploration Geophysicists, 2001.

# Appendix A

## The method of stationary phase

The method of stationary phase is a procedure that provides an approximate evaluation of integrals of the form

$$I = \int_{-\infty}^{\infty} F(x)e^{-i\varphi(x)}dx \quad (\text{A.1})$$

where the function  $F(x)$  varies slowly with  $x$  compared to the phase term  $\varphi(x)$  (Snieder, 2004a; Schuster, 2009). As the exponential term is rapidly oscillating over most of the range of integration it can be shown that its contribution to the integral will be zero apart from the so-called points of stationary phase  $x_s$ , where  $\varphi'(x_s) = 0$ . The Taylor series expansion for  $\varphi(x)$  around the stationary point  $x_s$  up to second order is given by

$$\varphi(x) \approx \varphi(x_s) + \varphi''(x_s)\frac{(x - x_s)^2}{2} \quad (\text{A.2})$$

Note that because  $\varphi'(x_s) = 0$  it is omitted in Eq. A.2. Because the function  $F(x)$  is slowly varying with  $x$ , close to each stationary phase point it can be replaced by its value at the stationary point,  $F(x_s)$ , and taken outside of the integral. Substituting the Taylor expansion into Eq. A.1 this yields

$$\begin{aligned} I &\approx F(x_s)e^{-i\varphi(x_s)} \int_{-\infty}^{\infty} e^{-i\varphi''(x_s)\frac{(x-x_s)^2}{2}} dx \\ &\approx F(x_s)e^{-i\varphi(x_s)} \sqrt{\frac{2\pi}{i\varphi''(x_s)}} \end{aligned} \quad (\text{A.3})$$

There will be one such approximation for each stationary point  $x_s$ , and the set of such approximations may be summed. Eq. A.3 shows that the main contribution to the integral in Eq. A.1 comes from the points  $x_s$  where the phase is stationary, in short, the stationary points.

# Appendix B

## Pseudo-physical energy

In this study we show that the term  $G_S G_S G_S^*$  constructs an event that arrives at the traveltime of the causal scattered wave  $G_S(\mathbf{x}_2, \mathbf{x}_1)$  and—after normalization of amplitudes—perfectly matches the waveform of the causal part of the modelled scattered Green’s function (e.g. Figs. 2.15c and d). This appendix provides an alternative mathematical development that explains why, despite its physical appearance, the event constructed from  $G_S G_S G_S^*$  is in fact nonphysical, and moreover why it can still be used as an estimate of a physical event.

The derivation is based on representations for scattered fields introduced by Vasconcelos et al. (2009), for performing intersource and thereafter inter-receiver interferometry. Halliday and Curtis (2009) and Vasconcelos et al. (2009) show that in inter-source interferometry the causal scattered wave is provided by stationary point contributions from the top boundary of receivers ( $S'_t$ ) according to

$$G_S(\mathbf{x}_2, \mathbf{x}_1) = \frac{2}{\rho c} \int_{S'_t} G_S(\mathbf{x}', \mathbf{x}_1) G_0(\mathbf{x}', \mathbf{x}_2) dS'_t \quad (\text{B.1})$$

A sketch of the corresponding ray paths is given in Fig. B.1. We now assume that, as before,  $\mathbf{x}_2$  is a receiver location; hence,  $G_0^*(\mathbf{x}', \mathbf{x}_2)$  is not available and needs to be constructed from interreceiver interferometry using

$$G_0^*(\mathbf{x}', \mathbf{x}_2) = \frac{2}{\rho c} \int_{S_b} G_0(\mathbf{x}_2, \mathbf{x}) G_0^*(\mathbf{x}', \mathbf{x}) dS_b \quad (\text{B.2})$$

where  $S_b$  denotes a bottom boundary of sources, as shown in Fig. B.2. However, if  $S_b$  is not available because the source locations are restricted to the surface,  $G_0^*(\mathbf{x}', \mathbf{x}_2)$  cannot be constructed and  $G_S(\mathbf{x}_2, \mathbf{x}_1)$  in Eq. B.1 is not retrieved.

Note that when substituting Eq. B.2 into Eq. B.1 we obtain a double surface integral corresponding to the term  $G_S G_S G_0^*$ . According to Fig. 2.3 and Table 2.1 this term accounts for the construction of the causal scattered wave; however,

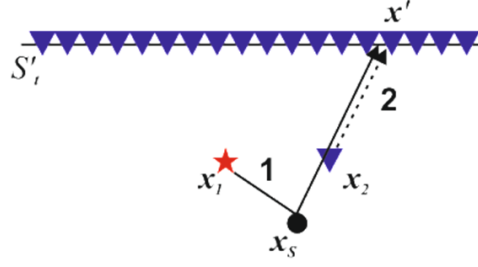


Figure B.1: Sketch of ray paths constructing  $G_S(\mathbf{x}_2, \mathbf{x}_1)$  from inter-source interferometry according to Eq. B.1. Ray paths shown are (1)  $G_S(\mathbf{x}', \mathbf{x}_1)$  and (2)  $G_0^*(\mathbf{x}', \mathbf{x}_2)$ . Symbol key as in Fig. 2.2. The dashed line indicates complex conjugation, that is, solid and dashed lines that run parallel cancel each other in phase. Since  $\mathbf{x}_2$  is in fact a receiver location,  $G_0^*(\mathbf{x}', \mathbf{x}_2)$  needs to be constructed from inter-receiver interferometry (see Fig. B.2 or B.3), which turns the receiver at  $\mathbf{x}_2$  into a virtual source.

as the associated stationary point pair ( $\mathbf{x} = \mathbf{d}, \mathbf{x}' = \mathbf{c}$  in Fig. Fig. 2.3) is not included in the surface boundaries this wave is not retrieved in our examples (Fig. 2.9h). In this case we can still construct an event that *looks like* the required direct wave by cross-correlating scattered fields according to

$$\hat{G}_0^*(\mathbf{x}', \mathbf{x}_2) = \frac{2}{\rho c} \int_{S_t} G_S(\mathbf{x}_2, \mathbf{x}) G_S^*(\mathbf{x}', \mathbf{x}) dS_t \quad (\text{B.3})$$

Fig. B.3 shows a sketch of the corresponding ray paths. However, the Green's functions on the right-hand side of Eq. B.3 interact with the scatterer and hence carry information about the scattering matrix. Therefore the left-hand side of Eq. B.3 is not equal to a direct wave since it must also contain information about the scatterer. Snieder et al. (2006) apply a similar argument for waves reflected at an interface: the apparent direct wave constructed from crosscorrelation of reflected wavefields contains a factor proportional to the reflection coefficient. Inserting Eq. B.3 into Eq. B.1 gives the double surface integral representing the term  $G_S G_S G_S^*$ .

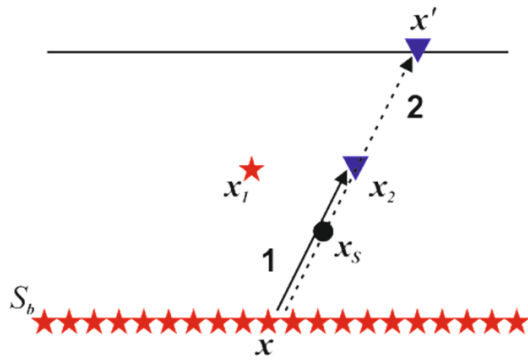


Figure B.2: Sketch of ray paths constructing  $G_0^*(\mathbf{x}', \mathbf{x}_2)$  from interreceiver interferometry according to Eq. B.2. Ray paths shown are (1)  $G_0(\mathbf{x}_2, \mathbf{x})$  and (2)  $G_0^*(\mathbf{x}', \mathbf{x})$ . Symbol key as in Fig. B.1.

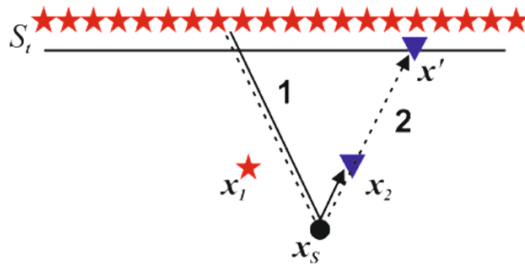


Figure B.3: Sketch of ray paths constructing  $\hat{G}_0^*(\mathbf{x}', \mathbf{x}_2)$  from interreceiver interferometry according to Eq. B.3. Ray paths shown are 1:  $G_S(\mathbf{x}_2, \mathbf{x})$  and 2:  $G_S^*(\mathbf{x}', \mathbf{x})$ . Symbol key as in Fig. B.1.

# Appendix C

## Supporting Information

The following sections analyse interferometric estimates of source-receiver Green's functions between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  with boundary sources and receivers confined to linear surface arrays, as shown in Fig. 2.7. Each estimate is constructed using only the single term in the integrand of Eq. 2.7 that is shown in the section title (i.e., the estimate is constructed by ignoring all other terms). Thus we illustrate the contribution and properties of each term individually. The only exception is the first section below which analyses the case where all terms are considered (i.e., where the full Eq. 2.7 is used).

### *GGG*<sup>\*</sup>: No wavefield separation

The solid line in Fig. C.1 shows the interferometric estimate using all of Eq. 2.7, with the boundaries and constrained to the surface as in Fig. 2.7. Both the causal and acausal wave between and appear to be constructed; however, we show in Appendix B that what looks like the causal scattered wave is in fact a pseudo-physical arrival. The direct wave is missing since stationary points are not sampled in this geometry. Instead a number of non-physical events occur that are not predicted by the directly modelled Green's function (dashed line). These events can be stationary or non-stationary and occur due to the lack of boundary sources and receivers in the subsurface. The origin of the different types of non-physical arrivals will be investigated in detail by the following numerical examples. Some non-physical events can be suppressed by tapering the contributions from the end points of the boundaries (Fig. C.1b). Fig. C.1c displays the interferometric estimate using both monopole and dipole sources (i.e., using Eqs. 2.1 and 2.2 instead of the approximate Eqs. 2.3 and 2.4). This reduces the amplitudes of non-physical arrivals (e.g. at 0.25 s) relative to the amplitudes of physical arrivals. Using a sparser source and receiver coverage along the bound-

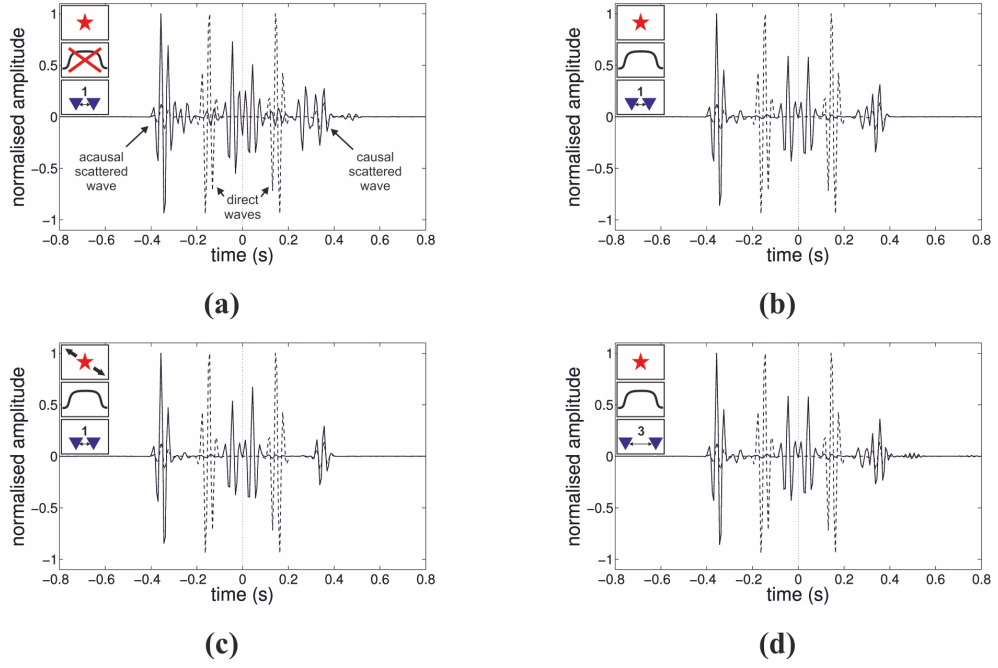


Figure C.1: Interferometric result of the full integral  $GGG^*$  of Eq. 2.7 (solid line) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  (dashed line) for different parameter constellations (see Table 2.2). Maximum amplitudes have been normalized to one.

aries evokes additional non-physical events (e.g., at 0.5s, Fig. C.1d). As the following examples will show, each of these events can be attributed to particular terms of Eq. 2.7.

$$G_0G_0G_S^*$$

Table 2.1 shows that the term  $G_0G_0G_S^*$  accounts for the construction of four events related to different stationary pairs. In the geometry used (Fig. C.2), however, only the stationary points  $\underline{\mathbf{c}}$  and  $\underline{\mathbf{e}}$  are spanned by sources and receivers, hence only the acausal scattered wave is constructed (Fig. C.3). Besides this physical arrival, non-physical events are observed at approx. (1)  $-0.05$  s, (2)  $-0.15$  s, and (3)  $-0.25$  s. Events (2) and (3) originate from the end points of the receiver boundary where the contributions of non-stationary boundary points are not cancelled due to the abrupt truncation of the summation—see Fig. 2.10b. Similarly, event (1) originates from the end points of the source boundary: Fig. 2.10a shows that it is generated in the first step and then contributes to the second step. All such non-physical arrivals can be significantly suppressed by a spatial taper (Fig. C.3b). Moreover, the non-physical events are non-stationary

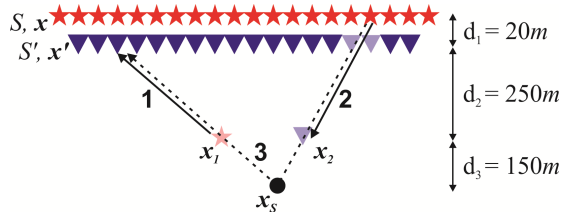


Figure C.2: Example set of raypaths used in  $G_0G_0G_S^*$ , symbol key as in Fig. 2.2. The dashed line represents the ray of the complex conjugated term. Faded symbols are used to make raypaths visible. Rays shown are 1:  $G_0(\mathbf{x}', \mathbf{x}_1)$ , 2:  $G_0(\mathbf{x}_2, \mathbf{x})$ , 3:  $G_S^*(\mathbf{x}', \mathbf{x})$ .

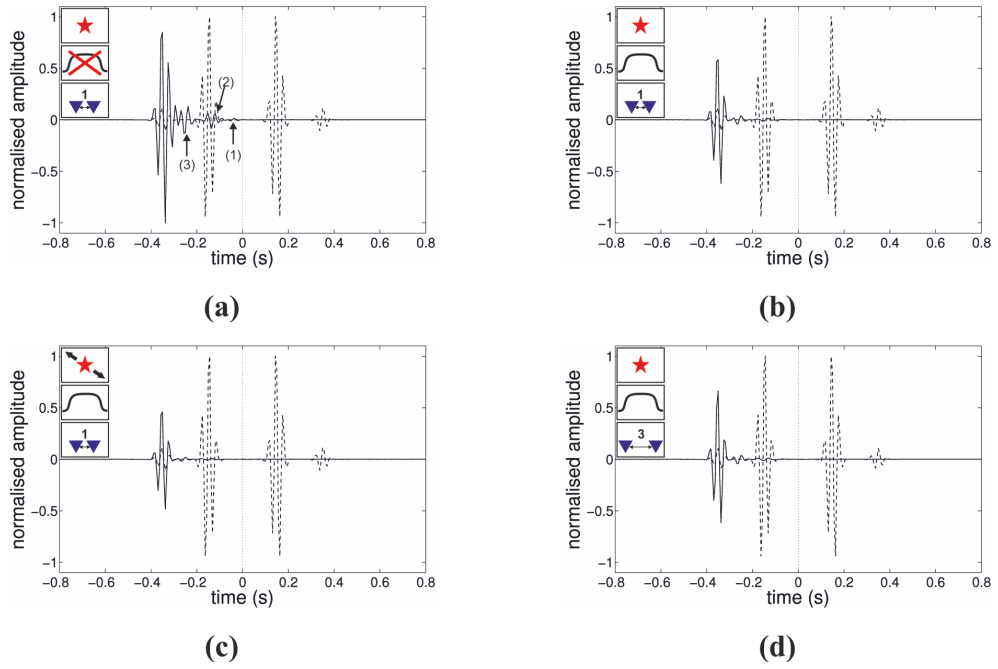


Figure C.3: Interferometric result of the cross term  $G_0G_0G_S^*$  in Eq. 2.7 (solid line) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  (dashed line) for different parameter constellations (see Table 2.2). Traces have amplitudes that are normalized relative to case (a). Number labels are referred to in the text.



and will vary in travel time when, for example, the lengths of the boundaries are changed. Therefore they can easily be discriminated from the stationary scattered wave arrival. Using dipole sources enhances the ratio of physical to non-physical energy (Fig. C.3c). Increasing the inter-source and inter-receiver distance along the boundaries still retrieves a good estimate of the acausal scattered wave (Fig. C.3d). In the presented example the decimation of the boundary points by up to a factor of six did not create additional non-physical events, but this depends on the slope of the travel time curve in the correlation gather (Mehta et al., 2008; also compare Figs. 2.13c and 2.14c).

$$G_S G_S G_S^*$$

Stationary phase analysis predicts that the term  $G_S G_S G_S^*$  constructs an event with the same travel time as the causal scattered wave, and that any source-receiver pair is stationary and contributes energy to the constructed arrival (Table 2.1). This is confirmed in the numerical examples (Fig. C.5): a single event appears on the trace where the causal wave is expected, and unlike in other terms, no spurious events associated with the boundaries' end points are observed regardless of the parameter constellation. This can be understood in terms of the correlation gathers (Fig. 2.10c and d): in inter-receiver interferometry as well as inter-source interferometry the travel time curve in the correlation gather is a constant, i.e., the travel time of each individual event constructed from a single source or receiver, respectively, is independent of the source or receiver position along the boundaries. The stationary points are actually stationary planes, and all events sum up constructively to give a single arrival. However, the constructed event is not the causal scattered wave but in fact a non-physical event that arrives at the same travel time, as we discuss in detail in Appendix B. Moreover, we show that in 2D it matches also the waveform of the physical arrival, and that its amplitude differs by a factor proportional to the imaginary part of the scattering matrix. Hence, the *pseudo*-physical event constructed from  $G_S G_S G_S^*$  gives a good estimate of the causal scattered wave even when using only a partial integration boundary. In fact, a single source-receiver pair on the surface is sufficient to construct the pseudo-physical wave, and varying the parameters does not affect the quality of the constructed wave, nor does it introduce spurious energy (compare Figs. C.5a to d).

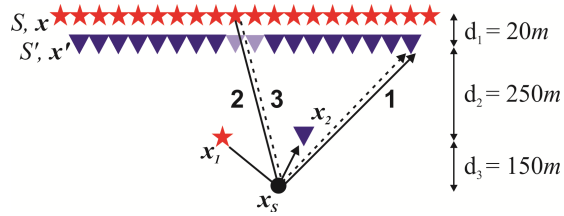


Figure C.4: Example set of raypaths used in  $G_S G_S G_S^*$ , symbol key as in Fig. C.2. Rays shown are 1:  $G_S(\mathbf{x}', \mathbf{x}_1)$ , 2:  $G_S(\mathbf{x}_2, \mathbf{x})$ , 3:  $G_S^*(\mathbf{x}', \mathbf{x})$ .

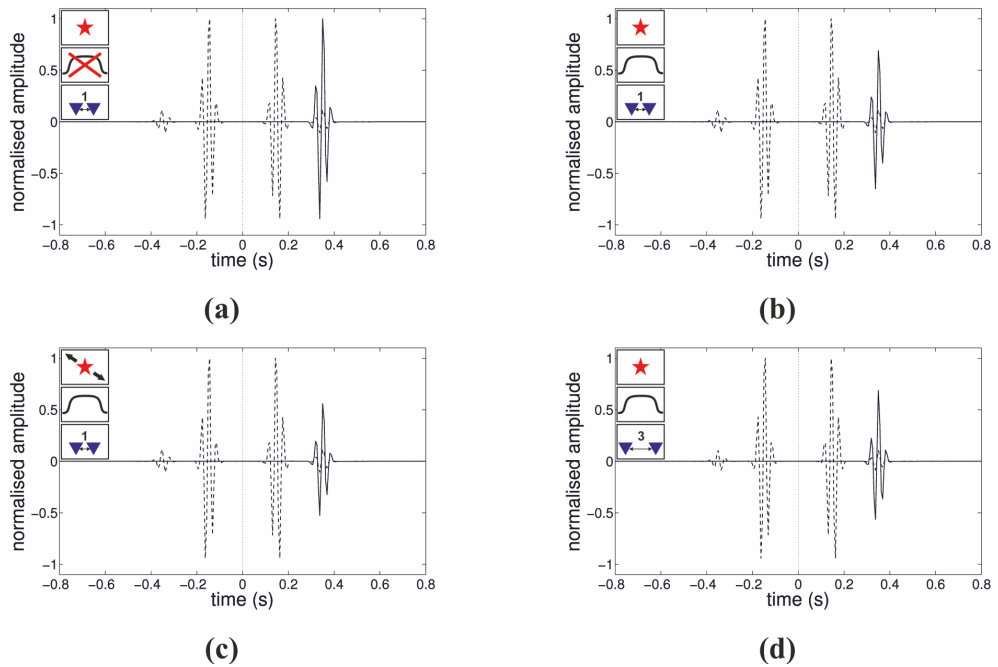


Figure C.5: Interferometric result of the cross term  $G_S G_S G_S^*$  in Eq. 2.7 (solid line) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_1$  (dashed line) for different parameter constellations (see Table 2.2). Traces have amplitudes that are normalized relative to case (a).

$$G_0 G_S G_S^*$$

The receiver array in the geometry used (Fig. 2.7) samples the stationary point  $\mathbf{x}' = \underline{\mathbf{e}}$  (Fig. 2.3), which for the term  $G_0 G_S G_S^*$  combined with any source on  $S$  provides a stationary non-physical arrival at  $\approx -0.05$  s (Table 2.1, Fig. C.8). The second event around 0.2 s is non-stationary and non-physical as it originates from the abrupt truncation of the summation over receivers. Fig. C.6 shows that only the acausal event is stationary under changing boundary conditions. Although this stationary event is termed non-physical its travel time has a useful physical interpretation: while the travel time of a (physical) scattered wave between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  (e.g. from term  $G_0 G_0 G_S^*$ ) is given by the *sum* of the travel times from  $\mathbf{x}_1$  to  $\mathbf{x}_s$  and from  $\mathbf{x}_s$  to  $\mathbf{x}_2$ , the travel time of the non-physical event corresponds to the difference of the travel times from  $\mathbf{x}_s$  to  $\mathbf{x}_2$  and from  $\mathbf{x}_1$  to  $\mathbf{x}_s$ . The combination of both, the sum and the difference, can be solved for the travel times of each part (from  $\mathbf{x}_1$  to  $\mathbf{x}_s$  and from  $\mathbf{x}_s$  to  $\mathbf{x}_2$ , respectively) and uniquely defines the location of a scatterer located below  $\mathbf{x}_1$  and  $\mathbf{x}_2$ .

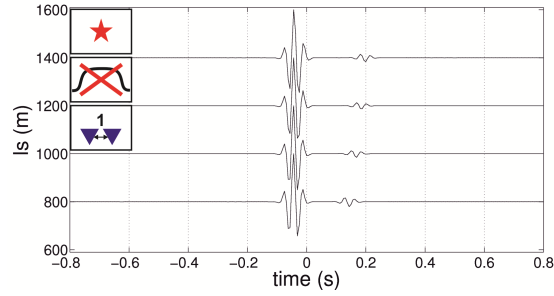


Figure C.6:  $G_0 G_S G_S^*$  for different apertures (length  $l_s$ ) of both source ( $l_s$ ) and receiver ( $l_s - 100$ ) boundary showing a stationary ( $-0.05$  s) and a non-stationary (between 0.1 s and 0.2 s) non-physical arrival.

$$G_S G_0 G_S^*$$

The term  $G_S G_0 G_S^*$  can be interpreted analogously to term  $G_0 G_S G_S^*$ : only the stationary points that account for the construction of the non-physical event ( $\forall \mathbf{x}', \mathbf{x} = \underline{\mathbf{c}}$ , Fig. 2.3) are spanned by the source and receiver arrays and provide a stationary event at approximately 0.05 s (Table 2.1, Fig. C.10). The travel time of this event corresponds to the difference between the travel times from  $\mathbf{x}_1$  to  $\mathbf{x}_s$  and from  $\mathbf{x}_s$  to  $\mathbf{x}_2$ . The second event at approximately 0.15 s is also non-physical but is non-stationary as it is an artefact from the abrupt truncation of the

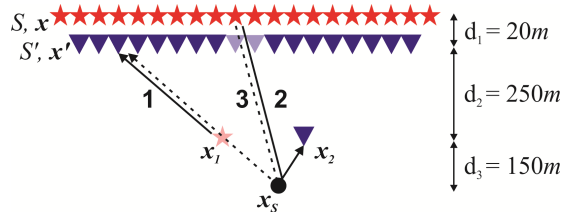


Figure C.7: Example set of raypaths used in  $G_0 G_S G_S^*$ , symbol key as in Fig. C.2. Rays shown are 1:  $G_0(\mathbf{x}', \mathbf{x}_1)$ , 2:  $G_S(\mathbf{x}_2, \mathbf{x})$ , 3:  $G_S^*(\mathbf{x}', \mathbf{x})$ .

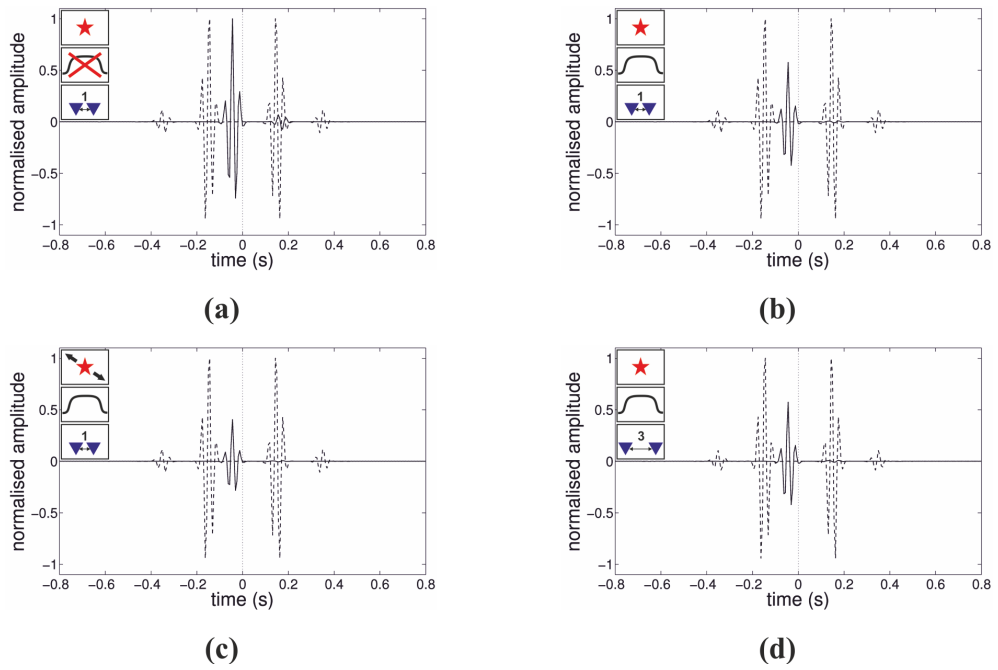


Figure C.8: Interferometric result of the cross term  $G_0 G_S G_S^*$  in Eq. 2.7 (solid line) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_1$  (dashed line) for different parameter constellations (see Table 2.2). Traces have amplitudes that are normalized relative to case (a).

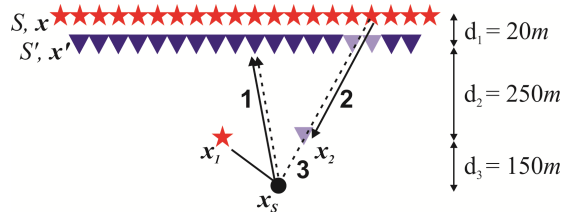


Figure C.9: Example set of raypaths used in  $G_S G_0 G_S^*$ , symbol key as in Fig. C.2. Rays shown are 1:  $G_S(\mathbf{x}', \mathbf{x}_1)$ , 2:  $G_0(\mathbf{x}_2, \mathbf{x})$ , 3:  $G_S^*(\mathbf{x}', \mathbf{x})$ .

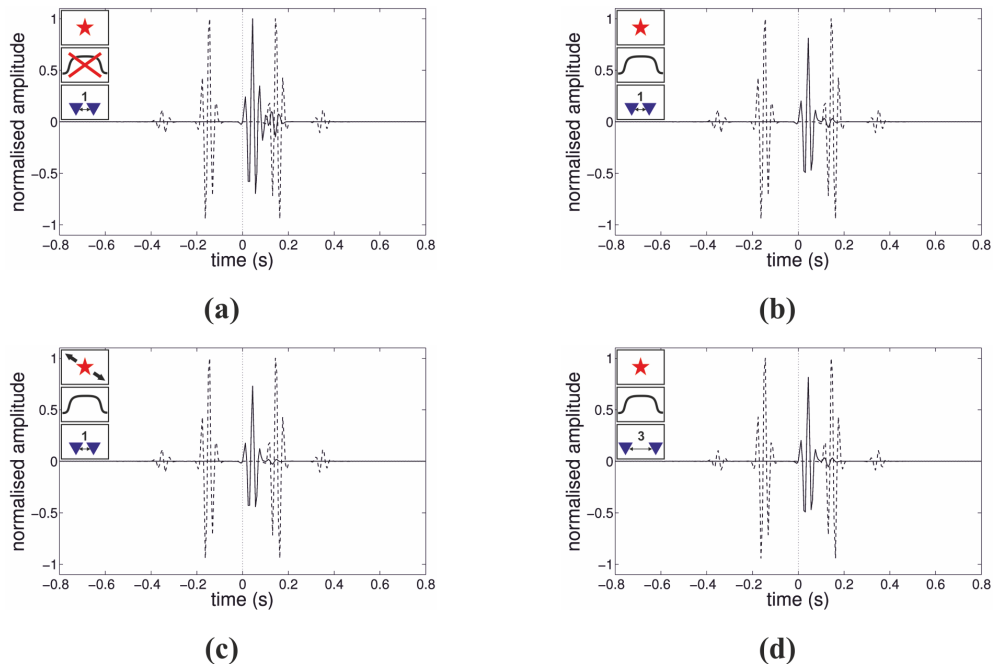


Figure C.10: Interferometric result of the cross term  $G_S G_0 G_S^*$  in Eq. 2.7 (solid line) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_1$  (dashed line) for different parameter constellations (see Table 2.2). Traces have amplitudes that are normalized relative to case (a).

source boundary (see correlation gather in Fig. 2.10g). Its travel time therefore changes with varying aperture of the source array, which makes it distinguishable from the stationary event (see also Fig. C.6), and its amplitude can be reduced by tapering on the boundary (Fig. C.10b) and by using dipole sources (Fig. C.10c). Decimating the source and receiver boundary by a factor of three does not introduce additional non-physical energy (Fig. C.10d).

$$G_0 G_0 G_0^*$$

Because the stationary points associated with  $G_0 G_0 G_0^*$  (**a** and **b** in Fig. 2.3) are not spanned by the surface boundaries the signal that is constructed (Fig. C.13) does not match the direct wave but consists of (at least two) non-physical arrivals. As can be seen in Fig. C.11 these events are not stationary but move out with increasing depth in Fig. C.12. For smaller depths the event converges towards the true direct arrivals at 0.15 s and  $-0.15$  s, respectively, whereas for larger depths an asymmetric move out towards positive travel times is observed. The cross-correlation gather (Fig. 2.10j) reveals that the non-physical energy originates from the contribution of non-stationary points at the edges of the boundaries. The events' amplitudes are significantly reduced by tapering the end point sources and receivers (Fig. C.13b), and are further reduced by including dipole sources (Fig. C.13c). Increasing the inter-source and inter-receiver spacing along the boundaries  $S$  and  $S'$ , respectively, introduces further spurious events on the causal side (Fig. C.13d), which occur due to the incomplete cancellation of non-stationary events in the inter-receiver step when the spacing on boundary  $S$  is too large (compare Figs. 2.14a and c).

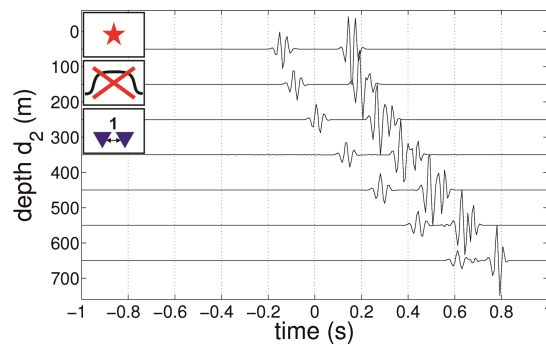


Figure C.11:  $G_0 G_0 G_0^*$  for different depths  $d_2$  showing the move-out of non-stationary events.

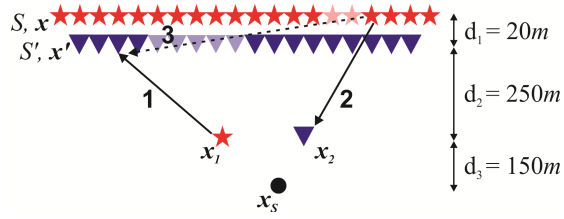


Figure C.12: Example set of raypaths used in  $G_0 G_0 G_0^*$ , symbol key as in Fig. C.2. Rays shown are 1:  $G_0(\mathbf{x}', \mathbf{x}_1)$ , 2:  $G_0(\mathbf{x}_2, \mathbf{x})$ , 3:  $G_0^*(\mathbf{x}', \mathbf{x})$ .

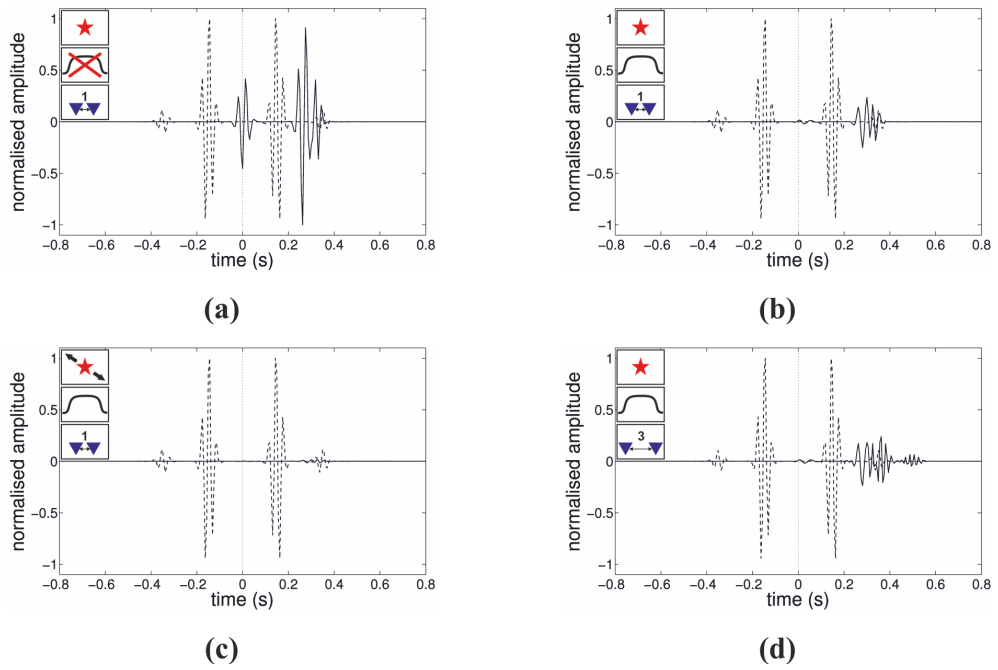


Figure C.13: Interferometric result of the cross term  $G_0 G_0 G_0^*$  in Eq. 2.7 (solid line) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_1$  (dashed line) for different parameter constellations (see Table 2.2). Traces have amplitudes that are normalized relative to case (a).

$G_S G_S G_0^*$ 

Meles and Curtis (2013) showed that for the term  $G_S G_S G_0^*$  any pair of a boundary source  $\mathbf{x}$  and a boundary receiver  $\mathbf{x}'$  connected by a straight line passing through the scatterer is stationary and provides a stationary arrival at the travel time of the causal scattered wave (Table 2.1). This configuration cannot, however, be realised when using surface boundaries only: both  $\mathbf{x}$  and  $\mathbf{x}'$  are always located on the same side of, namely above, the scatter. Therefore no stationary event is constructed, and as before the incomplete boundaries generate non-physical arrivals due to incomplete destructive interference of non-stationary events (Fig. C.15a). Applying a taper (Fig. C.15b) and using dipole sources (Fig. C.15c) suppresses the amplitude of the signal almost completely. A sparse source and receiver coverage along the boundaries (Fig. C.15d) generates further spurious events; however, they are not observed in the given time window.

 $G_0 G_S G_0^*$ 

As before, the linear source and receiver arrays do not span the stationary points required to create the events associated with the term  $G_0 G_S G_0^*$  (Table 2.1); instead non-physical arrivals occur due to the incomplete cancellation of non-stationary signals. This can be proved following previous arguments by looking at the cross-correlation gathers and the behaviour of the constructed signals under varying source/receiver boundary conditions. The traces in Fig. C.17 show the effect of different parameter constellations. Most noticeable is the strong amplitude reduction by down weighting the contribution of end point sources and receivers with a taper (Fig. C.17b), and the additional reduction by using dipole sources (Fig. C.17c). Increasing the inter-source and inter-receiver distance on boundaries  $S$  and  $S'$ , respectively, produces new spurious events, which, however, lie outside the considered time window (Fig. C.17d).



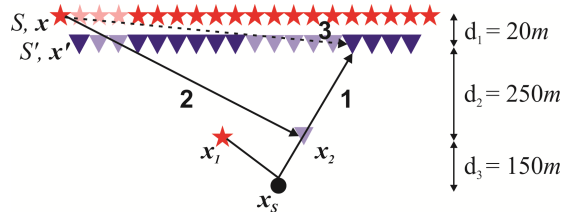


Figure C.14: Example set of raypaths used in  $G_S G_S G_0^*$ , symbol key as in Fig. C.2. Rays shown are 1:  $G_S(\mathbf{x}', \mathbf{x}_1)$ , 2:  $G_S(\mathbf{x}_2, \mathbf{x})$ , 3:  $G_0^*(\mathbf{x}', \mathbf{x})$ .

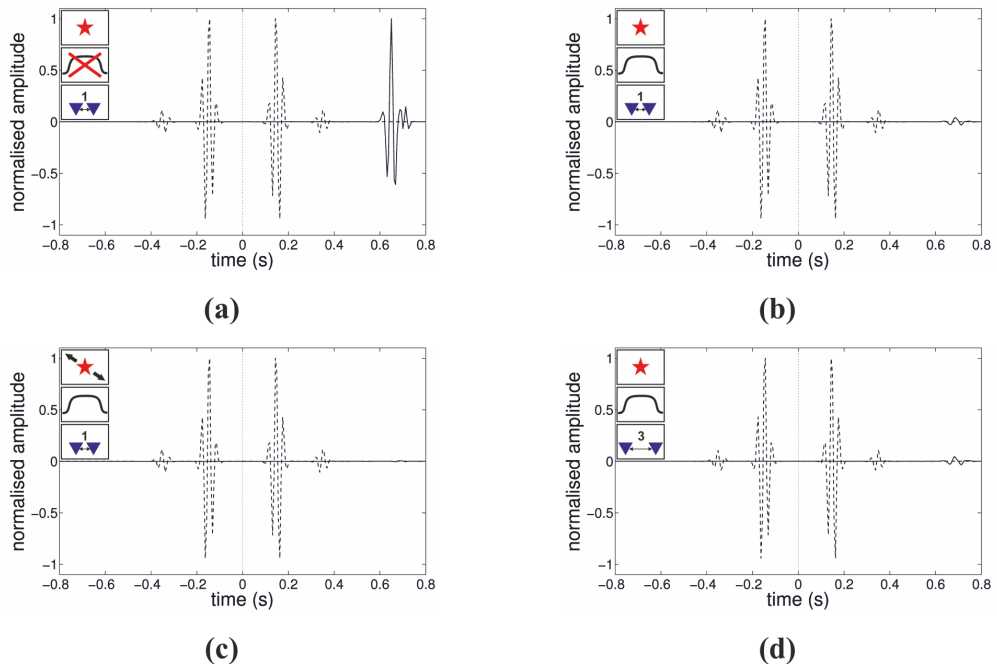


Figure C.15: Interferometric result of the cross term  $G_S G_S G_0^*$  in Eq. 2.7 (solid line) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_1$  (dashed line) for different parameter constellations (see Table 2.2). Traces have amplitudes that are normalized relative to case (a).

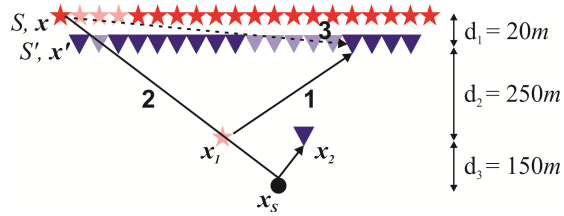


Figure C.16: Example set of raypaths used in  $G_0 G_S G_0^*$ , symbol key as in Fig. C.2. Rays shown are 1:  $G_0(\mathbf{x}', \mathbf{x}_1)$ , 2:  $G_S(\mathbf{x}_2, \mathbf{x})$ , 3:  $G_0^*(\mathbf{x}', \mathbf{x})$ .

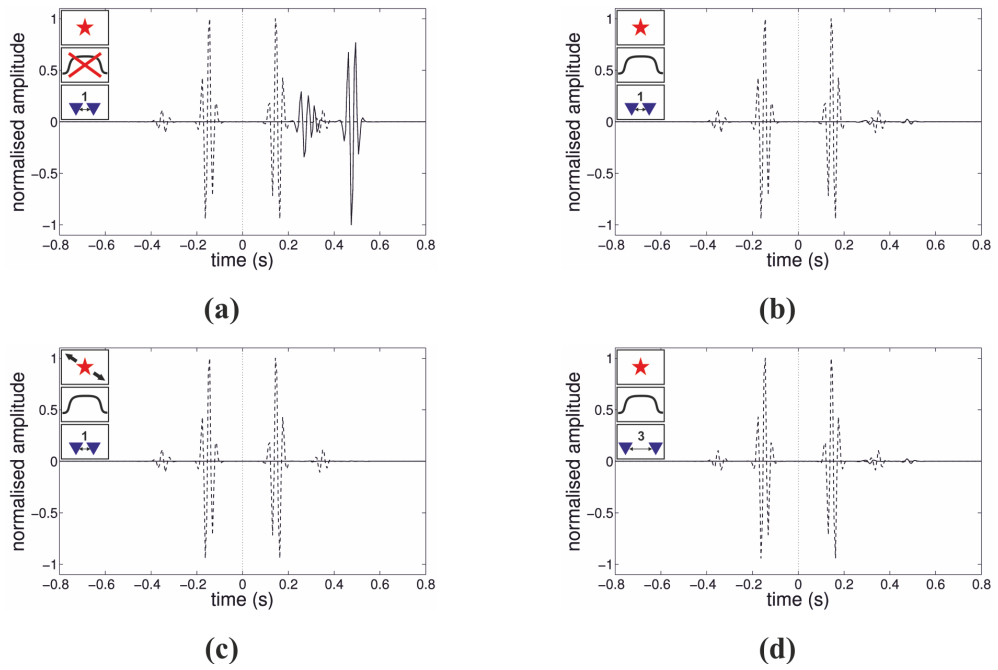


Figure C.17: Interferometric result of the cross term  $G_0 G_S G_0^*$  in Eq. 2.7 (solid line) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_1$  (dashed line) for different parameter constellations (see Table 2.2). Traces have amplitudes that are normalized relative to case (a).

$$G_S G_0 G_0^*$$

The stationary points for  $G_S G_0 G_0^*$  are **c** and **d** in Fig. 2.3, which account for the construction of the causal scattered wave and a non-physical, but stationary event (Table 2.1). Using surface boundaries only, point **d** is not spanned and the expected signals are not constructed. The arrivals shown in Fig. C.19 are non-physical and non-stationary and originate from incomplete cancellations in the summation of non-stationary events. As before, this can be shown from the cross-correlation gathers. The amplitude can noticeably be reduced by applying a spatial taper (Fig. C.19b) and by using dipole sources (Fig. C.19c). A coarser source and receiver coverage along the boundaries introduces new spurious arrivals at large travel times (Fig. C.19d). See previous sections for a detailed analysis of the generation of non-stationary events.

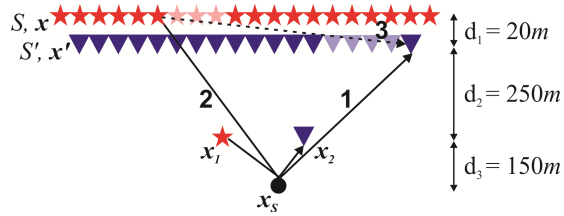


Figure C.18: Example set of raypaths used in  $G_S G_0 G_0^*$ , symbol key as in Fig. C.2. Rays shown are 1:  $G_S(\mathbf{x}', \mathbf{x}_1)$ , 2:  $G_0(\mathbf{x}_2, \mathbf{x})$ , 3:  $G_0^*(\mathbf{x}', \mathbf{x})$ .

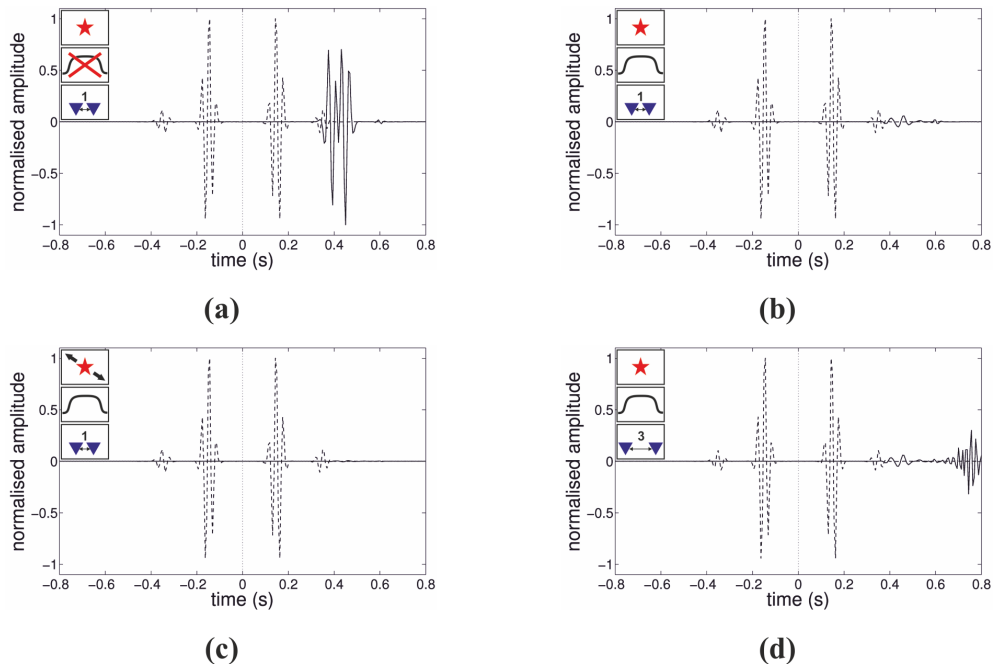


Figure C.19: Interferometric result of the cross term  $G_S G_0 G_0^*$  in Eq. 2.7 (solid line) compared to the true Green's function between  $\mathbf{x}_1$  and  $\mathbf{x}_1$  (dashed line) for different parameter constellations (see Table 2.2). Traces have amplitudes that are normalized relative to case (a).