SPATIO-TEMPORAL DISTRIBUTION OF SEISMIC WAVE FIELDS AND INTENSITIES: ACTIVE SOURCE FOCUSING AND ENERGY TRANSFER IN STRONGLY SCATTERING MEDIA

by

Manuel Alejandro Jaimes Caballero

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Geophysics).

Golden, Colorado

Date _____

Signed: _____

Manuel Alejandro Jaimes Caballero

Signed: _____

Dr. Roel Snieder
Thesis Advisor

Golden, Colorado

Date _____

Signed: _____

Dr. Paul Sava
Department Head
Department of Geophysics

ABSTRACT

As waves propagate from a source to a receiver, traversing a scattering medium, they interact with the inhomogeneities present in the medium. Depending on the scattering strength of the medium, or the propagation time, different waves may be recorded at the receiver location. These waves range from direct to single scattered to weakly multiple scattered to diffuse waves, and they contain information about the medium through which they propagate. These different waves have been used to locate and image seismic sources in strongly scattering media using the time-reversal method. This method consists of recording, at discrete receivers, the wave forms generated by a source, time-reversing these signals, and sending these signals back into the propagating medium (either physical or virtual) as new sources. Once the time-reversed signals are propagated into the medium, all the way back to the excitation time of the (original) source, the wave field is concentrated at the source location, provided the wave field is known on a closed surface surrounding the medium. The resolution with which one images the source is limited by the bandwidth of the wave field and the recording aperture. To deal with these resolution limitations, I develop a weighted time-reversal method, that is aimed at optimally localizing the source in space and time, within a given area. The weights compensate for the frequency content of the wave field and the acquisition geometry of the experiment, at the expense of very large wave field amplitudes outside the area on which I localize the source.

In addition to source focusing, one may want to describe the propagation of a wave field in strongly scattering media. Such description, via wave field amplitudes and phases, is not always possible because one may not know the location or characteristics of individual scatterers. If instead of wave field amplitudes and phases, one considers wave energy, it is possible to describe the wave propagation on average using the radiative transfer equations, which are based on energy conservation. These equations, which are hard to solve numerically due to coupling between angular directions as well as wave modes, describe the propagation of energy in a scattering medium as a function of space, time, direction of propagation, and wave mode (if considering elastic waves). To solve these equations, I construct a time-stepping algorithm with which I evolve the energy over time. To validate my numerical method, I compare the solutions that I obtain with known solutions, and find good agreement.

Because my algorithm accounts for the direction of propagation and wave mode, I use my algorithm to study the distribution of energy among propagation directions (valid for acoustic and elastic waves), and among wave modes (valid only for elastic waves). I find that the distribution of energy among propagation directions and wave modes, also known as the extent of equipartitioning, strongly depends on space and

time. Thus, the equipartitioning of a wave field is not a global, but a local property. This local behavior has implications for Green's function reconstructions where one assumes that the noise field that one uses in the reconstruction is equipartitioned. Because equipartitioning is a local rather than a global concept, the accuracy of the Green's function retrieval may be a function of space and time, contingent upon the length of the time windows that are used in the reconstruction of the Green's function.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost I would like to thank God for giving me the physical and mental health to be where I am today. I would like to thank my family for their teachings and sacrifices, so that I could accomplish a new milestone; my dear wife Emily and our chow chow puppy Blue for being by my side on sleepless nights of work, for always supporting me, and for starting a new life with me on the other side of the world; my cousins Milena and Daniela for endless conversations; and countless friends at the geophysics and math departments at Colorado School of Mines and the earth sciences and physics departments at Memorial University of Newfoundland without whom I wouldn't be where I am. Also thanks to my non-geo friends Tomi, Enrique, Shady, Areeb, Jim, and others. I really value our friendship, and I hope I can see you again. Thank you Luis Gabriel for teaching me so much about STEM, team work, and the excitement of science. I cherish those high school years. Thank you Roel and Alison for teaching me so much about research, geophysics, and life in general, I was very lucky to have amazing mentors. Thank you Paul Martin and Samy Wu Fung for teaching me so much about math, I enjoyed learning from the both of you. Thank you Lisa, Lynn, and Michelle for all the help with paperwork. Lastly, I would like to thank and remember those family members that left this world way to soon. My Auntie Luz, thank you for all your love and care, I miss you deeply. My Grandma Herminda and Teresa, and My Grandpa Jose, while I did not get to know you as much as I would have wished, I thank you for all the sacrifice you endured over decades. I'm only where I am thanks to your hard work. To my cousin Leo, thank you for our conversations, meals, and family gatherings, I always enjoyed talking to you and I wish you had not left this world so young.

To my beloved wife Emily, thank you for being part of my life.

To my dear parents Manuel and Martha, thank you for all your teachings and sacrifices. I wouldn't be where I'm today without you.

To my dear siblings Valentina and Andres, I miss you both deeply.

CHAPTER 1

INTRODUCTION

As waves propagate through a medium with inhomogeneities, they scatter, and travel through a multitude of travel paths. Generally speaking, one can identify four types of wave propagation regimes, either in terms of increasing scattering strength, or increasing propagation time, which I illustrate in Figure 1.1. Firstly, there are direct (or ballistic) waves, which travel on the direct path connecting the source and receiver locations. Secondly, singly scattered waves, which scatter once on the path from source to receiver. Thirdly, weak to strong multiply scattered waves, which scatter multiple times on the path from source to receiver. Lastly, diffuse waves, which are a type of strong multiply scattered waves. These waves have scattered so much that they lose track of their initial propagation direction and follow a random-walk-like trajectory (Franceschetti et al., 2004).

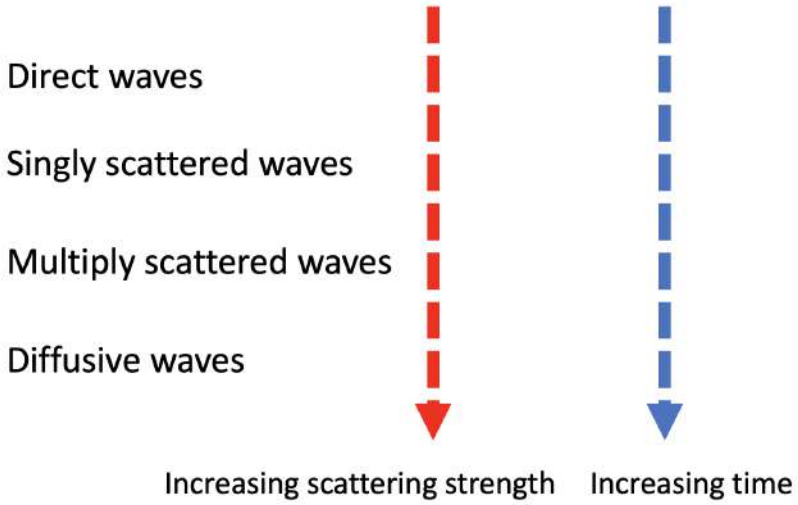

Figure 1.1 Illustration of scattering regimes in terms of either scattering strength (red arrow) or propagation time (blue arrow).

These different wave types, and the information that they contain, have been used in passive monitoring to locate and image seismic sources in strongly scattering media (Oren and Shragge, 2021), through wave field *focusing*. In this thesis, I refer to focusing as the process by which an incident wave field collapses, as well as possible, into a spatio-temporal delta function $\delta(\mathbf{x}-\mathbf{x}_0)\delta(t-t_0)$ (Rose, 2002). In simple terms, this definition says that focusing is the process by which a wave field becomes localized at a given spatial location $\mathbf{x_0}$ at a given time t_0. Other definitions of focusing also consider extended spatial distributions (Bazargani and Snieder, 2016). Methods such as inverse scattering (Broggini et al., 2012),

phase conjugation (Jackson and Dowling, 1991), inverse filtering (Tanter et al., 2000; Tanter et al., 2001), and time-reversal (Parvulescu and Clay, 1965; Fink, 1997) have been used to achieve source focusing. The last of these methods has been thoroughly tested and found to be robust in strongly scattering media (Fink et al., 2002). In **Chapter 2** of this thesis, I place emphasis on a modified version of the time-reversal method, which I call weighted time-reversal.

Time-reversal (TR) techniques have been widely used in seismology to study source focusing. Some applications include analysis of earthquake source mechanisms and location estimation (Lokmer et al., 2009), monitoring of nuclear explosions and environmental applications (Larmat et al., 2010), microseismic event location (Lu et al., 2008), reservoir monitoring (Shapiro, 2008), and reversed time migration (Schuster, 2002). TR techniques have also found use in other fields which include underwater acoustics communication (Edelmann et al., 2005), three-dimensional elastic media communication (Anderson et al., 2016), subsurface detection (Micolau et al., 2003), study of electromagnetic waves (Lerosey et al., 2004), ultrasonic nondestructive testing (Chakroun et al., 1995), and real-time tracking of gallstones during lithotripsy treatments (Fink et al., 2003). TR techniques were first experimentally developed by Parvulescu and Clay (1965) in the field of underwater acoustics, and later formalized by Fink (1997). TR techniques rely on the even order of the time differential operator in the wave equation and on spatial reciprocity (Fink et al., 2002; Snieder, 2004), and perform well in heterogeneous media (Cassereau and Fink, 1992; Blomgren et al., 2002). These techniques are based on the fact that one can transform the time variable $t \rightarrow -t$ to reverse the direction of time without altering the governing wave equation in non-dissipative environments (Fink et al., 2002). This reversibility property allows one to time-reverse the wave field recorded by a set of receivers all the way back to the point in time at which the wave field was generated. The TR process consists of recording the wave form (generated by a source) at discrete receivers, time-reversing the signal, and sending this signal back into the propagating medium (either physical or virtual) as a new source. Ideally, the wave field that is sent back into the medium should focus at the source location $\mathbf{x}_0$ at excitation time t_0, provided that one either has accurate knowledge of the medium properties (numerical time-reversal) or one propagates the time-reversed signal back into the medium in which the forward propagation took place (Fink et al., 2003). In practice, the wave field never truly collapses into a point in space and time. The focusing is spatially (and temporally) limited due to the finite bandwidth of the wave field and the imperfect sampling geometries (e.g., one-sided recording) that are used in seismology (Wapenaar et al., 2014). This focusing limitation is the *wave diffraction limit*, also known as Abbe's diffraction limit, which states that focusing can not be achieved with a resolution better than half a wavelength ($\lambda/2$) (Maznev and Wright, 2017).

It has been a longstanding research question whether it is possible to achieve resolution better than the diffraction limit to improve the spatial (and temporal) focal resolution. Some of the earliest ideas in the matter were postulated by Francia (1952). They proposed the idea of imaging beyond the diffraction limit using optimal circular array apertures. With these arrays, they optimally focused waves in a limited region of space. However, they found that outside this region, the wave field becomes very strong. They showed that if one increases the power of the circular arrays (with the goal of enhancing the focal resolution), the wave field outside the region of interest becomes increasingly large. More recently, novel developments in the optics community have shown that it is possible in practice to achieve resolution better than the diffraction limit, a phenomenon which is referred to as *super-resolution* (Rogers and Zheludev, 2013). Some of these developments include Near-Field Scanning Optical Microscopy (SNOM) where a scanning array collects information of propagating and evanescent waves near the source (Pohl et al., 1984), negative index super-lenses which allow transferring evanescent waves from the source to the image plane (Pendry, 2000), and super-oscillatory imaging lenses which allow for imaging beyond the diffraction limit but often produce adjacent side-bands (Rogers and Zheludev, 2013).

Achieving super-resolution would aid in enhancing focal resolution, which has implications for focusing through human tissue (Thomas and Fink, 1996), acoustic source reconstruction (Antoni, 2012), localization of moving sound sources in shallow water (Soares et al., 1999), and focusing in turbid media (Mosk et al., 2012). Outside of the field of optics, several researchers have developed novel super-resolution techniques. Lerosey et al. (2007) achieve super-resolution with far field time-reversal by placing random scatterers near the source location, Lemoult et al. (2011) use acoustic resonators in the far field to achieve sub-wavelength resolution of sound, Mimani et al. (2015) propose a sponge-layer time-reversal technique to improve aero-acoustic time-reversal imaging, Schuster et al. (2012) introduce the idea of a seismic scanning tunneling macroscope, analogous to SNOM, which requires scatterers in the vicinity of the seismic source, and Guo et al. (2016) show that in the far field it suffices to use resonant multiples in data migration to achieve sub-wavelength resolution. All of these developments rely, one way or another, on either the use of evanescent or strongly scattered waves. The former help retrieve the full wave field spectrum (Katrich, 2005), and the latter help with the source illumination (Fink, 2006). In practice, one does not always have access to evanescent or strongly scattered waves. The evanescent waves decay very rapidly beyond the near-field, and strongly scattered waves are only generated in strongly heterogeneous media.

To avoid using evanescent or strongly scattered waves in enhancing the focal resolution, I pose the following question in **Chapter 2**: is it possible to enhance the spatio-temporal focal resolution by modifying the TR process? I formulate this question as a modified Backus-Gilbert (BG) problem (Backus and Gilbert, 1968), in the sense that I look for the optimal weights with which I can reconstruct a

band-limited delta function in space and time (after time-reversal), robust to errors in the velocity model and estimated source location. In **Chapter 2** I provide the answer to this question through numerical experiments in homogenous as well as heterogeneous media. As I show in **Chapter 2**, it is indeed possible to enhance the focal resolution, to some extent, through a weighted TR process, but this enhancement can only be done locally in an area denoted the field-of-view (FOV). This FOV corresponds to a limited spatial area on which I optimize the source focusing. Outside this area, large wave field amplitudes arise. This wave field amplitude behavior suggests that the optimization of the focal resolution must be treated as a local rather than global process. The work that I present in **Chapter 2** is linked to earlier work in seismology such as that of Anderson et al. (2015) who achieve temporal focusing via deconvolution which is implemented as an inverse filter in the frequency domain, and that of Bazargani and Snieder (2016) who minimize the difference between the backpropagated wave field and the time-reversed displacement field in the near source region by searching for optimal signals to send into the medium.

While source focusing has been a useful tool in the passive monitoring of seismic sources (Anderson et al., 2015; Douma et al., 2015), it is also of interest to understand (qualitatively and quantitatively) wave field propagation in strongly scattering media. In such media, it is not always possible to compute wave field amplitudes and phases, since one may not know the location or characteristics of individual scatterers (for a discussion on the matter see Chapter 1 of Ishimaru, 1978). However, if instead of wave field displacements, one pays attention to wave field energies, it is possible to describe the wave field propagation on average, using the radiative transfer equations (RTE).

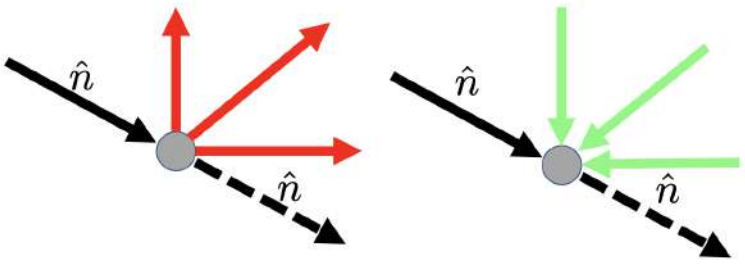

Figure 1.2 Schematic of the radiative transfer equations using energy conservation. The black arrow indicates energy incoming from a propagation direction $\hat{n}$. The dashed black arrow indicates energy that is scattered along the same propagation direction $\hat{n}$. The red arrows indicate energy scattered to other propagation directions. The green arrows indicate energy incident from directions different to $\hat{n}$. The grey dot represents a scatterer.

The RTE are based on the principle of energy conservation (Turner and Weaver, 1994). In simple terms one can represent the RTE through the following relation

$$[\partial_t + v\hat{n} \cdot \nabla]\text{intensity} = \text{source} - \text{loss} \ + \text{gain}. \tag{1.1}$$

In expression 1.1, ∂_t is the time derivative, v is the speed of propagation of the waves, $\hat{n}$ is the direction along which energy advects, and ∇ is the gradient operator. This expression states that as energy advects along the propagation direction $\hat{n}$ (left-hand side), it loses energy due to scattering to other directions and gains energy due to scattering from other directions and from any sources that may be present. I illustrate this energy principle in Figure 1.2. Formally, the RTE consist of a coupled system of integro-differential equations where one solves for the wave intensity as a function of space, time, and angular direction (and wave mode if dealing with elastic waves), assuming one knows the scattering mean free path(s), the angular dependence of the scattering pattern(s), and the speed(s) of energy propagation.

Both acoustic and elastic formulations to the RTE have been studied (Gaebler et al., 2015; Margerin et al., 2016; Zhang and Sens-Schönfelder, 2022), with most of the attention devoted to the acoustic formulation. The elastic formulation is harder to study than the acoustic formulation, because there is coupling not only in the angular direction but also in the mode of propagation due to $P \rightarrow S$ and $S \rightarrow P$ mode conversions. The acoustic description to RTE has been used in astrophysics to analyze radiation transport across cosmic dust in a wide range of astrophysical objects (Steinacker et al., 2002), in atmospheric sciences to model solar radiation across clouds to better understand the evolution of sea surface temperatures (Evans and Stephens, 1995), in optics to develop novel optical tomographic imaging algorithms which allow diagnosis and treatment of biological tissues (Klose et al., 2002), in acoustics to model the interaction of acoustic waves with the ocean bottom (Quijano and Zurk, 2009), in geophysics to model infrared radiation across volcanic ash clouds (Prata, 1989), heat transfer in the mantle (Hofmeister, 2005), and scattering kernels in coda wave interferometry (Pacheco and Snieder, 2005; Rossetto et al., 2011; Margerin et al., 2016; Snieder et al., 2019; Duran et al., 2020).

The last of these applications has gained much attention in the fields of passive imaging and non-destructive testing, because with coda wave interferometry it is possible to localize (to some extent) the time-lapse changes in medium properties (Pacheco and Snieder, 2005; Planès et al., 2014; and Margerin et al., 2016). To this end, one solves an inverse problem (i.e., $\mathbf{d} = \mathbf{G}\mathbf{m}$), where one seeks to retrieve a model $\mathbf{m}$, from data $\mathbf{d}$, using a forward operator $\mathbf{G}$. In the context of coda wave interferometry, the data is the average time-shift (within a given time window) between wave forms recorded before and after the time-lapse change, the model is the localized perturbation, and the forward operator is the scattering kernel (Snieder et al., 2019). This kernel is a function of the spatial coordinate $\mathbf{x}$, the location of the source

$\mathbf{x}_s$, the receiver $\mathbf{x}_r$, the time window over which the travel time shift is measured, as well as on the wave intensities, which are the solution to the RTE. Because with RTE it is possible to describe the spatio-temporal-angular distribution of the wave intensities, solving these equations has become important in the construction of accurate scattering kernels (Margerin et al., 2016; Zhang et al., 2021).

Several numerical techniques have been developed to model the RTE. These include the discontinuous Galerkin finite element method (Clarke et al., 2019; Han et al., 2010), Markov chain Monte Carlo (Iwabuchi, 2006; Xu et al., 2011; Camps and Baes, 2018; Noebauer and Sim, 2019; Przybilla and Korn, 2008; Yoshimoto, 2000), finite difference (Klose and Hielscher, 1999), and wave equation modeling, whereby one exploits the connection between the acoustic wave equation and the acoustic RTE (Przybilla et al., 2006; Kanu and Snieder, 2015; Snieder et al., 2019; Duran et al., 2020). In addition to these numerical techniques, several authors have derived analytical approximations to the acoustic RTE. These approximations include assuming a point-like, isotropic, impulsive source of intensity in a statistically homogeneous medium (Margerin et al., 2016), expanding the intensity and scattering function into a finite sum of Legendre polynomial and then solving a finite system of equations for the unknown coefficients appearing in a truncated expansion (Roberge, 1983), decomposing the specific intensities into a sum of partial intensities and then solving the RTE for each partial intensity assuming that scattering is angle-independent and that the source is isotropic (Paasschens, 1997), assuming a steady-state intensity field (Fan et al., 2019; Le Hardy et al., 2016), assuming azimuthal symmetry (Baes and Dejonghe, 2001; de Abreu, 2004), or assuming that the wave propagation is diffusive, which is only valid at times much larger than the transport mean free time (Rossetto et al., 2011; Planès et al., 2014). These theoretical developments, though insightful and meaningful in some particular applications, can not always be generalized due to the assumptions that they require.

To address some of the analytical limitations mentioned above, I propose an efficient time-stepping algorithm in **Chapter 3** to solve the acoustic RTE. This algorithm is based on transforming the acoustic RTE into integral equations for the specific intensities, following the work of Paasschens (1997), and then evolving the specific intensities over time. This algorithm consists of computing the specific intensities at time t, from the specific intensities at $t - \Delta t$. The relation between the intensities at time t and $t - \Delta t$ is given by a time-stepping matrix which arises due to the discretization over angular directions. The numerical algorithm that I propose is valid for different types of initial condition, as well as angle-dependent scattering, and non-uniform (in space) scattering. While with my time-stepping algorithm I do not currently handle boundary conditions, they can be incorporated into the algorithm using energy conservation.

The elastic RTE has been mostly restricted to seismological applications. Some of the earliest work was done by Wu and Aki (1985). They introduce the RTE in seismology to describe the generation of seismic coda as seismic waves scatter through the heterogeneous Earth's structure. Sato et al. (1997) used the elastic RTE to derive semi-analytical expressions of the energy density due to a point shear-dislocation source. Margerin et al. (2000) used Monte Carlo simulations to study the multiple scattering of elastic waves via the elastic RTE. Przybilla et al. (2006) solve the elastic RTE using a Monte Carlo method and compare their results to finite difference modeling of elastic waves in a two-dimensional (2-D) random medium. Yamamoto and Sato (2010) use elastic RTE to study the scattering properties of the Asama volcano in Japan. In addition to the applications mentioned above, the RTE have been used to study the late time behavior of acoustic and elastic energy in strongly scattering media. Ishimaru (1978) showed that for late times, when the wave propagation is almost independent of direction, and is nearly stationary in time, the acoustic RTE lead to diffusive energy transport. Similarly, Ryzhik et al. (1996) and Turner and Weaver (1994) derived the diffusion approximation to the elastic RTE. As elastic waves are scattered more and more, the energy of the waves equilibrates between wave modes and propagation directions. This equilibration is referred to as *equipartitioning*. Ryzhik et al. (1996) showed through the elastic RTE that the ratio of the total S to the total P energy of elastic waves converges to $2v_p^3/v_s^3$ in three dimensions (3-D) regardless of the scattering properties of the medium or the type of source that releases the energy. Weaver (1982) and Snieder (2002) also derived this ratio without invoking the RTE. Zhang et al. (2021) showed that in 2-D this ratio converges to v_p^2/v_s^2. Margerin et al. (2000) investigated the time evolution of the ratio $E_s(t)/E_p(t)$, via Monte Carlo simulations of the elastic RTE, as a marker of different scattering mechanisms.

The concept of energy equipartitioning has also been a topic of discussion in seismic interferometry. The main idea behind seismic interferometry is that one can retrieve the causal and anti-causal Green's function propagating between receiver locations A and B by cross-correlating noise recorded at such locations. In seismic interferometry, one assumes that the noise has no preferred propagation direction (Wapenaar et al., 2010). If these assumptions hold, one should retrieve, via cross-correlation, symmetric causal and anti-causal Green's functions (Weaver and Lobkis, 2004). In this thesis, I refer to equipartitioning as either the equal distribution of energy along angular directions (valid for both acoustic and elastic waves), or the stabilization of energy between the available wave modes (only valid for elastic waves). Snieder et al. (2010) showed that while equipartitioning is necessary for Green's function retrieval, it is not sufficient. This is because the process of Green's function extraction also requires that the noise sources are uncorrelated. Equipartitioning, of both acoustic and elastic waves, has been experimentally, numerically, and theoretically studied in fields such as acoustics and seismology due to its importance in

Green's function reconstruction. Weaver and Lobkis (2001) proposed to perform such reconstruction in an ultrasonic experiment using thermal field fluctuations, assuming that such field is equipartitioned. Snieder et al. (2007) illustrated, heuristically, the role that equipartitioning plays in the Green's function retrieval. Stehly et al. (2006) showed that the noise, which one may use to reconstruct Green's function near coastlines, radiates preferentially away from storms at the oceans. This preferential directionality of the noise radiation affects the symmetry of the Green's function that one retrieves, especially if the noise field that one uses is uni-directional and not evenly distributed. Malcolm et al. (2004) experimentally reconstructed, in a laboratory setting, Green's functions in a highly scattering medium using time windows ranging from ballistic to diffusive to equipartitioned waves. They showed that the reconstruction of the causal and anti-causal components of the Green's functions is only possible when the waves are equipartitioned. Sens-Schönfelder et al. (2015) studied the equipartitioning of the late coda in teleseismic events. They found that even at late times the seismic waves propagate with a preferred direction. This directionality causes seismic interferometry to fail in reconstructing the Green's function unless both the earthquake and receiver are located along the preferred direction of propagation. Generally speaking, in the field of seismic interferometry, researchers assume that equipartitioning occurs globally, which is almost never the case for open or very large systems, such as the earth. Weaver (2010) introduced conceptually the idea of local equipartitioning for open/large systems. Because the degree of equipartitioning is a function of space and time, depending on the amount of scattering, the mismatch between the numerical and true Green's function is also a function of space and time, contingent upon the length of the time intervals that one uses in the Green's function reconstruction.

Due to the significance of the elastic RTE in seismology, I extend the algorithm that I develop in **Chapter 3**, to include elastic waves in **Chapter 4**. I focus on the 2-D case, though one could extend the algorithm to 3-D. Similar to the acoustic algorithm, I transform the elastic RTE into integral equations for the P and S specific intensities, and then derive a time-stepping algorithm with which I evolve the P and S specific intensities. The time-stepping matrix that I obtain with this algorithm is a generalized version of the time-stepping matrix for the acoustic algorithm, and it accounts for $P \rightarrow P$, $S \rightarrow P$, $P \rightarrow S$, and $S \rightarrow S$ scattering. I test this algorithm against the diffusive approximation to the elastic RTE, the equipartitioning ratio, and expressions for the total P and S energy (Sato et al., 2012). Since with the acoustic and elastic time-stepping algorithms that I develop it is possible to describe the spatio-temporal-angular evolution of the energy, I use these algorithms to investigate the idea of local equipartitioning. In **Chapter 3** I investigate the angular equipartitioning of energy, via the acoustic RTE, as a function of space and time. I show, as suggested by Weaver (2010), that one needs to consider equipartition as a local rather than global process. I introduce the equipartitioning index δ in **Chapter 3**

to measure the distribution of energy among propagation directions as a function of space and time, and show numerically that for a localized source, even at times $t \gg \tau_s$, where τ_s is the time scale over which scattering occurs (i.e., the scattering mean free time), the energy field does not reach global equipartitioning for the simulation set-up that I consider. In **Chapter 4** I follow up the work conducted in **Chapter 3** to study the angular equipartitioning of P and S waves, via the elastic RTE, using the equipartitioning index δ. In addition to angular equipartitioning, I also numerically investigate the equipartition of energy between P and S wave modes as a function of space and time, and find that it takes much longer for the local equipartitioning ratio, as compared to the global equipartitioning ratio, to reach the theoretical value of v_p^2/v_s^2. As in **Chapter 3**, I show in **Chapter 4**, that even at times much larger than the P and S scattering mean free times, the energy field does not reach global equipartitioning for the numerical simulations that I consider. The local behavior of both angular and modal equipartitioning implies that the accuracy with which one retireves Green's function is also a function of space and time (Weaver, 2010).

In **Chapter 3** and **Chapter 4**, I only consider propagation times less than the time it takes the energy of the source to reach the computational boundary, and assume spatially uniform scattering. While such assumptions are valid for particular problems, there are scenarios where one might want to incorporate reflection at boundaries or spatially non-uniform scattering. It is possible in practice, though not simple, to enhance the algorithms in **Chapter 3** and **Chapter 4** to handle the scenarios mentioned above using energy conservation. For instance, to enforce a reflective boundary, one can require the energy reflecting off the boundary to be the same as the energy that is incident on the boundary. Similarly, one could extend the algorithms presented in **Chapter 3** and **Chapter 4** to handle layered velocity models. To do this, one would solve the RTE independently in each of the layers, and then at the interfaces enforce that the sum of the fluxes of the incoming waves be the same as the flux of the outgoing waves (Margerin, 2005). To enforce the conservation of the energy flux, one would need to specify reflection and transmission coefficients such that the incident energy at the interface equals the sum of reflected plus transmitted energy.

In **Chapter 5**, I explore the use of machine learning to model the RTE. In 2-D and 3-D the acoustic RTE consists of a system of many coupled integro-partial differential equations. In 2-D and 3-D, one truncates the acoustic RTE into a finite set of N coupled integro-partial differential equations, and then one numerically solves this system. Due to computational limitations, this large coupled system is currently very expensive to solve through machine learning. To proptype the application of machine learning to the RTE, I consider the 1-D RTE, which consists of a coupled system of two partial differential equations (right and left going waves). With machine learning one aims to learn the relationship between input and output data pairs. To learn this mapping one uses *Artificial Neural Networks* (ANNs), which were inspired by biological neural networks (Mishra and Srivastava, 2014). ANNs consist of a collection of

nodes, denoted *neurons*, connected through a series of layers. To inform ANNs about the desired mapping one constructs a *penalty function* which quantifies the difference between the current output and the desired output. One then updates (i.e., the training algorithm) the ANNs until the penalty function is below some threshold. Physics Informed Neural Networks (PINNs), a machine learning technique, have become a popular tool to compute the solutions to complicated ordinary and partial differential equations (Raissi et al., 2019), subject to initial and/or boundary conditions, and have found use in the modeling of the Navier-Stokes equations (Jin et al., 2021), Helmholtz equation (Song et al., 2021), Euler equations (Jagtap et al., 2020), Eikonal equation (Waheed et al., 2021), and Schrödinger's equation (Wu et al., 2022). The main idea behind PINNs is to train a neural network so that its output satisfies some given physical law, as well as initial and/or boundary conditions (Raissi et al., 2019). Historically speaking, two different approaches have been used in the field of PINNs. The simplest approach is that of Raissi et al. (2019), denoted as the *soft-constraints* approach. In this approach, one assesses the mismatch of the initial and boundary condition through a penalty function. As an alternative, Lu et al. (2021) proposed a *hard-constraints* approach, in which one incorporates the initial and boundary conditions within the neural network architecture, thereby satisfying the initial and boundary condition exactly. I propose a mixed-constraints approach to solve the 1-D RTE, where the initial condition is handled exactly, and the boundary condition mismatch is incorporated in the penalty function. In **Chapter 5** I show that this approach outperforms the soft-constraints approach. Although I focus on the 1-D RTE in **Chapter 5**, I perform several numerical tests of increasing difficulty to assess the future capability of PINNs to model RTE in realistic scenarios in higher dimensions.

1.1 Thesis Outcomes

I present the outcomes of this thesis in Chapters 2 through 5, listed as follows

1. Chapter 2: **Jaimes, M. A.** and R. Snieder, 2021. Spatio-temporal resolution improvement via weighted time-reversal. Wave Motion. **106**:102803.

2. Chapter 3: **Jaimes, M. A.** and R. Snieder, 2023. Illustration of Diffusion and Equipartitioning as Local Processes: A Numerical Study Using the Radiative Transfer Equations. The Journal of the Acoustical Society of America (*Accepted pending minor revisions*).

3. Chapter 4: **Jaimes, M. A.**, and R. Snieder, 2023. Angular and Modal Equipartitioning of Elastic Waves in Scattering Media: A Numerical Study Based on Energy Transport. The Journal of the Acoustical Society of America (*to be submitted*).

4. Chapter 5: **Jaimes, M. A.**, R. Snieder, and Wu Fung S., 2023. Application of Physics Informed Neural Networks to the 1-D Radiative Transfer Equations. The European Journal of Physics (*to be submitted*).

In **Chapter 6** I provide a general overview of this thesis, final conclusions, and possible future research directions.

CHAPTER 2

SPATIO-TEMPORAL RESOLUTION IMPROVEMENT VIA WEIGHTED TIME-REVERSAL

Reproduced with the permission of the Journal of Wave Motion[1].

Manuel Jaimes-Caballero[2], and Roel Snieder[3].

2.1 Abstract

We formulate resolution enhancement as a modified Backus-Gilbert inverse problem and determine the optimal complex weights that improve focusing of waves in space and time. The optimization corrects for receiver geometry. If we accurately know the location of a control point in the subsurface we can use the corresponding optimal weights to achieve enhanced focusing in a prescribed target zone surrounding the control point. Errors in the back propagation velocity and noisy data degrade the quality of focusing. The optimal wave field shows a blow-up behaviour outside the optimization area. We show different measures of resolution to estimate the compression of the focal spot. The optimized weights amplify the high frequencies, but the algorithm also improves the focusing for monochromatic waves. At all frequencies our algorithm improves the resolution of the focal spot. We also show that for a uniformly sampled line array and a homogenous medium, the weights used to enhance resolution have a negligible imaginary part and that they are oscillatory across the array used. To fully test the robustness of our algorithm, we also consider focusing in a heterogeneous medium with embedded scatterers and an irregular receiver line, and show that in this scenario we are also able to attain focusing improvement.

2.2 Introduction

Locating and imaging seismic sources has long been of interest in quantitative seismology and acoustics. Typically one uses the information contained in the wave motions caused by an earthquake to image that earthquake. Methods widely used to achieve source focusing include inverse scattering (Broggini et al., 2012), phase conjugation (Jackson and Dowling, 1991), the inverse filter (Tanter et al.; Tanter et al., 2001; 2000), time-reversal (Fink; Parvulescu and Clay, 1997; 1965), and beam forming (Billingsley and Kinns; Capon, 1976; 1969). The latter two methods can be thought of as related processes (Jin et al., 2007; Viteri-Mera and Teixeira, 2017; Xu et al., 2018). In this chapter we will mostly refer to time-reversal techniques but relevant analogies can be made to beam forming techniques. Further, we note that many other techniques in fields such as optics and wireless communications have been developed to enhance

[1]Reprinted with permission of Journal of Wave Motion, 106:102803. See appendix for details.
[2]Graduate student, primary researcher, and author at Department of Geophysics, Colorado School of Mines.
[3]Supervisor at Department of Geophysics, Colorado School of Mines.

focusing. One technique that is useful as an analogy to the method we propose is spatial light modulation where one modifies the amplitude and phase of incident mono-chromatic waves to achieve spatial focusing upon propagation through an imaging system (Chung and Kim, 1999).

Seismologists often use time-reversal (TR) techniques for source focusing. Applications in seismology include analysis of earthquake source mechanisms and location estimation (Lokmer et al., 2009), monitoring of nuclear explosions and environmental applications (Larmat et al., 2010), microseismic event location (Lu et al., 2008), reservoir monitoring (Shapiro, 2008), and reversed time migration (Schuster, 2002). Other applications of TR techniques include underwater acoustics communication (Edelmann et al., 2005), three-dimensional elastic media communication (Anderson et al., 2016), subsurface detection (Micolau et al., 2003), study of electromagnetic waves (Lerosey et al., 2004), ultrasonic nondestructive testing (Chakroun et al., 1995), and real time tracking of gallstones during lithotripsy treatments (Fink et al., 2003).

TR techniques, as first developed by Parvulescu and Clay (1965) and then expanded by Fink (1997), rely on the even order of the time differential operator in the wave equation and on spatial reciprocity (Fink et al., 2002; Snieder, 2004), and are valid both in homogeneous and inhomogeneous media (Cassereau and Fink, 1992; Blomgren et al., 2002). As a result, it is possible to reverse the direction of time without altering the governing wave equation in non-dissipative environments (Fink et al., 2002). This reversibility allows one to time-reverse data recorded at an array to estimate the location of sources. The time-reversal process consists of recording the physical signal at discrete receivers, time-reversing it, and sending back the time-reversed signal into the medium as a new source.

Optimally, the wave that one sends back to the medium should focus at the location where the event originated at time $t = 0$ provided one has an accurate knowledge of the velocity of the medium in which propagation takes place. However, even under ideal conditions, the focusing is spatially limited by the dominant wavelength of the wave field. This limit is the *wave diffraction limit*, which states that focusing can not be achieved with a resolution better than half a wavelength ($\lambda/2$) (Maznev and Wright, 2017).

Francia (1952) proposed the idea of imaging beyond the diffraction limit using optimal circular array apertures. However, this concept was not applicable at the time due to practical limitations regarding the large power required for the focusing. Recent tests in the optics community have shown that it is possible in practice to go beyond the diffraction limit to achieve *super-resolution*, which is of importance in the inverse source problem as well as conventional seismic imaging. Several seismological and non-seismological applications can benefit from the development of super-resolution techniques. These include the applications mentioned above but also other applications in which improving the resolution of the focusing at the source is desirable, such as focusing through human tissue (Thomas and Fink, 1996), acoustic source

reconstruction (Antoni, 2012), localization of moving sound sources in shallow water (Soares et al., 1999), and focusing in turbid media (Mosk et al., 2012). New optical super-resolution technologies include Near-Field Scanning Optical Microscopy (SNOM) where a scanning array collects information of propagating and evanescent waves near the source, negative index super-lenses which allow transferring evanescent waves from the source to the image plane, and super-oscillatory imaging lenses which allow for imaging beyond the diffraction limit but often produce adjacent side-bands (Rogers and Zheludev, 2013). Several researchers in the field of acoustics have also developed novel super-resolution techniques. de Rosny and Fink (2002) overcome the diffraction limit by combining a time reversal mirror with an acoustic sink, Lerosey et al. (2007) achieve super-resolution with far field time-reversal by placing random scatterers near the source location, Conti et al. (2007) propose a near-field time-reversal approach which combines an acoustic sink with near field acoustic holography (NAH), Lemoult et al. (2011) use acoustic resonators in the far field to achieve sub-wavelength resolution of sound, Mimani et al. (2015) propose a sponge-layer time-reversal technique to improve aeroacoustic time-reversal imaging, and Mimani (2021) propose an iterative version of the work by Mimani et al. (2015). In the seismological community Schuster et al. (2012) introduce the idea of a seismic scanning tunneling macroscope, analogous to SNOM, which requires scatterers in the vicinity of the seismic source, a near-field technique. To overcome the near-field limitation, Guo et al. (2016) show that in the far field it suffices to use resonant multiples in data migration to achieve sub-wavelength resolution. These developments, though insightful, depend on the presence of evanescent waves for near-field (Schuster et al., 2012) or strongly scattered multiples for far-field (Guo et al., 2016).

To understand how we can attain enhanced resolution without evanescent waves or a strongly heterogeneous medium we pose the following question: Can we find frequency-dependent complex weights for each of the discrete receivers such that after time reversal, the focal spot at the source location resembles a delta function as closely as possible?. We formulate this question as a modified Backus-Gilbert (BG) problem (Backus and Gilbert, 1968), in the sense that we search optimal weights that allow for reconstruction of a band-limited delta function in space and time, robust to errors in the velocity model and estimated source location. This is similar to weight optimization in beam forming where several authors have made use of BG techniques with the goal of focusing at a desired location (Wilson, 1995; Sekihara et al., 2005; Huang et al., 2003). Van Veen and Buckley (1988) give an extensive discussion on widely used beam forming techniques. Note that in conventional beam forming one typically addresses spatial focusing (monochromatic pulses or narrow band data) whereas this paper addresses spatio-temporal focusing (polychromatic pulses or wide band data). Some research has been developed in the area of wide band beam forming (Bucris et al., 2012; Hawes and Liu, 2014; Rasekh and Seydnejad, 2014; Liu et al., 2019) but more work remains to be done to attain an understanding of spatio-temporal focusing via beam

forming (Liu and Weiss, 2010). Our approach is also linked to earlier work in the seismological community. Anderson et al. (2015) achieve temporal focusing via deconvolution which is implemented as an inverse filter in the frequency domain. Bazargani and Snieder (2016) minimize the difference between the backpropagated wave field and the time-reversed displacement field in the near source region by searching for optimal signals to send into the medium.

This chapter is structured as follows: In section 2.3 we formulate the optimal weights needed to achieve improved focusing. Next, in section 2.4 we show the improvement of focusing in a standard seismic geometry when we use the optimal weights and estimate whether we obtain local or global focusing. We present a sensitivity analysis in section 2.5 where we consider errors in source location and medium velocity, as well as noisy data. We illustrate different metrics of resolution in section 2.6. Our results show not only the boosting of high frequencies but also the modulation of the source signal at all frequencies, which results in resolution improvement even for monochromatic waves. In section 2.7 we show the frequency variation of the spectrum of the weights for four receiver locations along with the variation of the weights across the receiver array for different frequencies. We relate the spatial dependence of the weights along the receiver line to Tchebyscheff polynomials used in beam forming, and discuss how our algorithm is advantageous in the case of irregular arrays. In section 2.8 we apply our method to focusing in a heterogeneous medium, and show how focusing is improved. We discuss the effectiveness of the optimal weights in time-reverse imaging in section 2.9.

2.3 Theory

Consider a source located at $\mathbf{x}_0$ which emits a pulse $S(t)$ at $t_0 = 0$. The time-reversal imaging operation in the time domain where the signal recorded by each receiver, $G(\mathbf{x}_i, t, \mathbf{x}_0) * S(t)$, is time reversed (i.e., $G(\mathbf{x}_i, \mathbf{x}_0, -t) * S(-t)$) and sent back into the medium from receiver i reads (Fink and Prada, 2001)

$$f_i(\mathbf{x}, \mathbf{x}_i, t) = G(\mathbf{x}, \mathbf{x}_i, t) * (G(\mathbf{x}_i, \mathbf{x}_0, -t) * S(-t)). \tag{2.1}$$

The index i specifies a particular receiver, the symbol "$*$" implies temporal convolution, and $\mathbf{x}$ specifies the location of the grid point in the field of view (FOV), see Figure 2.1. $G(\mathbf{x}, \mathbf{x}_i, t)$ corresponds to the Green's function between receiver i and the location $\mathbf{x}$ in the FOV. Notice that although we use a focusing function which corresponds to unweighted time-reversal, one may use other focusing functions. Following Francia (1952) and Rogers and Zheludev (2013), we only optimize the focusing in the FOV.

We use the following convention for the Inverse Fourier Transform

$$f(t) = \int_{-\infty}^{+\infty} F(\omega) e^{-i\omega t} d\omega. \tag{2.2}$$

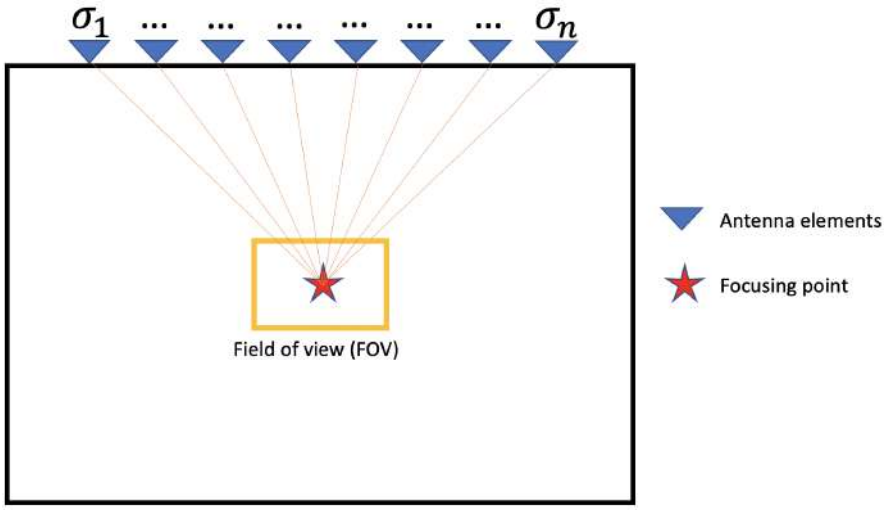

Figure 2.1 Illustration of the TR experiment geometry for n receivers. We optimize focusing in a small area (FOV) around the source within the propagation medium. σ_i is the frequency dependent complex weight assigned to the i^{th} receiver to optimize focusing.

For a point impulsive source, assuming perfect sampling in space and time (i.e., a continuous cavity and infinite bandwidth), a full aperture (i.e, all possible illumination angles), and a wave field whose focusing is not diffraction limited, the time-reversal operation leads to perfect focusing, namely

$$\sum_i f_i(\mathbf{x}, \mathbf{x}_i, t) = \delta(\mathbf{x} - \mathbf{x}_0)\delta(t), \tag{2.3}$$

which corresponds in the frequency domain to

$$\sum_i F_i(\mathbf{x}, \mathbf{x}_i, \omega) = \delta(\mathbf{x} - \mathbf{x}_0). \tag{2.4}$$

Recall that equation 2.4 only holds for a wave field which is not diffraction-limited and that is fully sampled in space and time. In the real world we deal with imperfect sampling geometries and diffraction-limited focusing. In order to enhance focusing we use frequency-dependent complex weights σ_i which improve the localization of the time-reversed wave field, replacing equation 2.4 by

$$\sum_i \sigma_i(\mathbf{x}_i, \omega) F_i(\mathbf{x}, \mathbf{x}_i, \omega) = \delta(\mathbf{x} - \mathbf{x}_0). \tag{2.5}$$

Notice that we want to achieve spatial focusing for each frequency under consideration. To construct a band-limited delta function given a finite set of measurements we use a modified version of the Backus-Gilbert (BG) method (Backus and Gilbert, 1968), in which we minimize the following objective function for every angular frequency ω

$$J(\mathbf{x}, \mathbf{x}_0, \omega, \sigma, \lambda) = \sum_i |\mathbf{x} - \mathbf{x}_0|^2 |\sigma_i(\mathbf{x}_i, \omega) F_i(\mathbf{x}, \mathbf{x}_i, \omega)|^2 + \lambda \left(\sum_i \sigma_i(\mathbf{x}_i, \omega) F_i(\mathbf{x}, \mathbf{x}_i, \omega) - \delta(\mathbf{x} - \mathbf{x}_0) \right). \tag{2.6}$$

We simultaneously minimize the objective function J at discrete points $\mathbf{x}$ in the FOV rather than minimizing the integral of J over FOV (as it is typically done in BG theory). This choice of minimization guarantees that we approach the desired band-limited delta function in a point-wise fashion. We minimize this objective function in terms of the complex weights σ_i and a Lagrange multiplier λ. By minimizing the first term of equation 2.6 we force the energy of the weighted TR wave field to be zero at all points except at the source location. The second term forces the amplitude of the TR wave field to be non-zero at the source.

Next, we consider the partial derivatives of J with respect to the optimization parameters

$$\frac{\partial J}{\partial \sigma_i^*} = \sum_j |\mathbf{x}-\mathbf{x}_0|^2 \sigma_j(\mathbf{x}_j,\omega) F_j(\mathbf{x},\mathbf{x}_j,\omega) F_i^*(\mathbf{x},\mathbf{x}_i,\omega) + \lambda F_i^*(\mathbf{x},\mathbf{x}_i,\omega), \tag{2.7}$$

$$\frac{\partial J}{\partial \lambda} = \sum_j \sigma_j(\mathbf{x}_j,\omega) F_j^*(\mathbf{x},\mathbf{x}_j,\omega) - \delta(\mathbf{x}-\mathbf{x}_0), \tag{2.8}$$

here the superscript * implies complex conjugation. The subscripts i, j specify receivers from which the incident wave field is time-reversed.

To minimize the objective function we set the partial derivatives in equations 2.7 and 2.8 equal to zero and solve this system of equations simultaneously for all grid points in FOV. These equations are written as

$$\underbrace{\begin{bmatrix}
|\mathbf{x}_1-\mathbf{x}_0|^2 F_1^1 F_1^{*1} & \cdots & |\mathbf{x}_1-\mathbf{x}_0|^2 F_N^1 F_1^{*1} & F_1^{*1} \\
\vdots & \vdots & \vdots & \vdots \\
|\mathbf{x}_1-\mathbf{x}_0|^2 F_1^1 F_N^{*1} & \cdots & |\mathbf{x}_1-\mathbf{x}_0|^2 F_N^1 F_N^{*1} & F_N^{*1} \\
F_1^{*1} & \cdots & F_N^{*1} & 0 \\
\vdots & \vdots & \vdots & \vdots \\
|\mathbf{x}_k-\mathbf{x}_0|^2 F_1^k F_1^{*k} & \cdots & |\mathbf{x}_k-\mathbf{x}_0|^2 F_N^k F_1^{*k} & F_1^{*k} \\
\vdots & \vdots & \vdots & \vdots \\
|\mathbf{x}_k-\mathbf{x}_0|^2 F_1^k F_N^{*k} & \cdots & |\mathbf{x}_k-\mathbf{x}_0|^2 F_N^k F_N^{*k} & F_N^{*k} \\
F_1^{*k} & \cdots & F_N^{*k} & 0 \\
\vdots & \vdots & \vdots & \vdots \\
|\mathbf{x}_M-\mathbf{x}_0|^2 F_1^M F_1^{*M} & \cdots & |\mathbf{x}_M-\mathbf{x}_0|^2 F_N^M F_1^{*M} & F_1^{*M} \\
\vdots & \vdots & \vdots & \vdots \\
|\mathbf{x}_M-\mathbf{x}_0|^2 F_1^M F_N^{*M} & \cdots & |\mathbf{x}_M-\mathbf{x}_0|^2 F_N^M F_N^{*M} & F_N^{*M} \\
F_1^{*M} & \cdots & F_N^{*M} & 0
\end{bmatrix}}_{\mathbf{A}}
\underbrace{\begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_N \\ \lambda \end{bmatrix}}_{\boldsymbol{\chi}}
=
\underbrace{\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \end{bmatrix}}_{\mathbf{b}}, \tag{2.9}$$

here $1 \le k \le M$, k is the grid point $\mathbf{x}$ in the field of view where we evaluate equations 2.7 and 2.8, M is the number of grid points in the field of view, and N is the number of receivers. The superscript in F specifies the corresponding grid point whereas the subscript specifies the receiver number. The only non-zero element in the vector on the right-hand side of equation 2.9 corresponds to setting $\mathbf{x} = \mathbf{x}_0$ in equation 2.8. Ultimately, we look for solutions to equation 2.9 that minimize $||\mathbf{A}\boldsymbol{\chi}-\mathbf{b}||^2$ in the least-squares sense. The

solution vector $\boldsymbol{\chi}$ is found for each frequency independently, and contains the frequency-dependent complex weights σ_i and the lagrange multiplier λ.

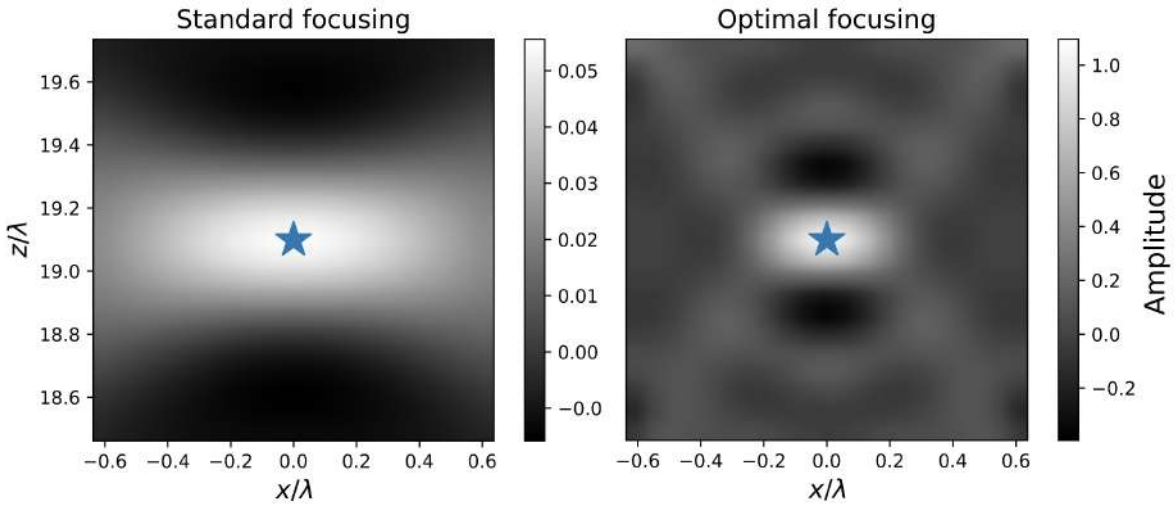

Figure 2.2 Comparison of standard TR (left) and optimal TR images (right) for $\lambda \sim 80$m at FOV. Notice the improvement of the focusing with the optimal weights.

2.4 Focusing in homogeneous media

To simulate a typical seismic scenario, consider Figure 2.1. This array illuminates the source at limited angles. We consider an array far away from the source (about 19 dominant wavelengths). Here we first consider a homogeneous acoustic medium. We perform the simulation and optimization using the Green's function $G^{2D}(\mathbf{r}, \mathbf{r}_0)$ which gives the outgoing wave solutions of the 2-D Helmholtz equation for our choice of Fourier transform convention,

$$G^{2D}(\mathbf{r}, \mathbf{r}_0) = \frac{-i}{4} H_0^1(k|\mathbf{r} - \mathbf{r}_0|), \tag{2.10}$$

where H_0^1 denotes the first Hankel function of degree zero (Snieder and van Wijk, 2015). We use 101 receivers, placed symmetrically above the source between -1000 and 1000 m, and located at a vertical distance of 1500 m from the source. For the FOV we use a box located from -50 to 50 m in the horizontal direction and 1450 to 1550 m in the vertical direction, with a grid spacing of 5 m in both directions. We use an acoustic medium with constant velocity v=1000 m/s, frequency ranging from 0.32 Hz to 40 Hz with a spacing of 0.32 Hz, and a source wavelet with peak frequency of 12.73 Hz. Throughout this section and the rest of the chapter we refer to pressure amplitude as amplitude.

We expect the weights to amplify the high frequency content of the signal, and correct for incomplete wave field sampling and illumination angles. Figure 2.2 shows the comparison of optimal and standard

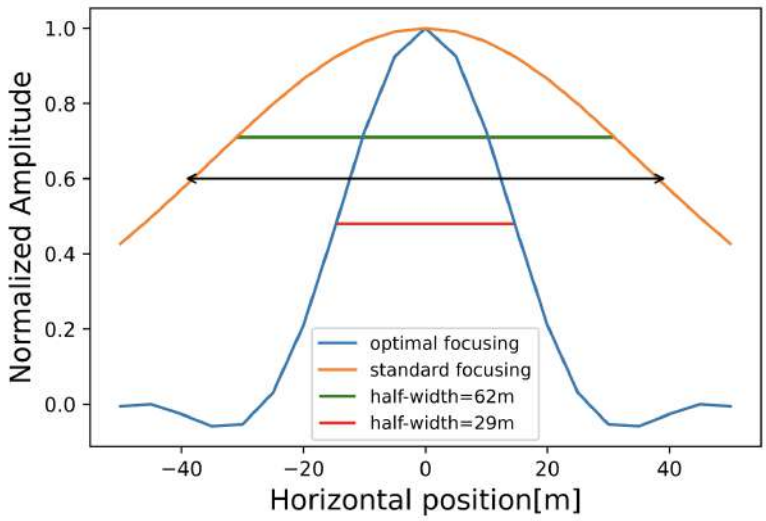

Figure 2.3 Cross section at source level for standard and optimal focusing. Optimal half-width (red line) is below half the dominant wavelength $\sim$ 40 m. Standard half-width (green line) is about 3/4 of the dominant wavelength. The two headed black arrow indicates one full dominant wave length.

focusing. The focal spot that we obtain with the algorithm that we propose, has a better resolution than the one we obtain with standard time-reversal. This figure demonstrates that the weights help achieve improved localization of the source. To show in detail the improvement of the focal resolution we show in Figure 2.3 a cross-section at the source level for the panels in Figure 2.2. The focal spot that we obtain with our algorithm shows a compression of about a factor of 2 relative to standard time-reversal. This focal spot is resolved with a resolution below the dominant wavelength. Besides compression of the focal spot, our algorithm reduces the side-lobes within the FOV. To estimate whether we achieve local or global focusing we calculate the weights for the FOV in Figure 2.1 and analyze the standard and optimal time-reversed wave fields in a larger area. Figure 2.4 shows the standard and optimal time-reversed wave fields for an area larger than FOV. The amplitude of the optimal time-reversed wave field inside FOV is masked by the amplitudes outside FOV. To better illustrate this behavior consider Figure 2.5 where we show the cross-section at the source level of the optimal time-reversed wave field from Figure 2.4. Notice that outside the FOV side lobes appear with amplitudes greater than the wave field inside the FOV, including the source location ($x = 0$), as pointed out above. This implies that we are only able to achieve local focusing in the vicinity of the source rather than everywhere in the medium. One could potentially increase the size of FOV to achieve a more global focusing at the expense of computational power (i.e., more data would be involved in the computation of the optimal weights assuming the grid spacing remains the same).

This finding regarding local focusing is not new. Francia (1952) showed that it is possible to increase the resolution at the focal spot at the expense of large amplitudes at other locations in the imaging area.

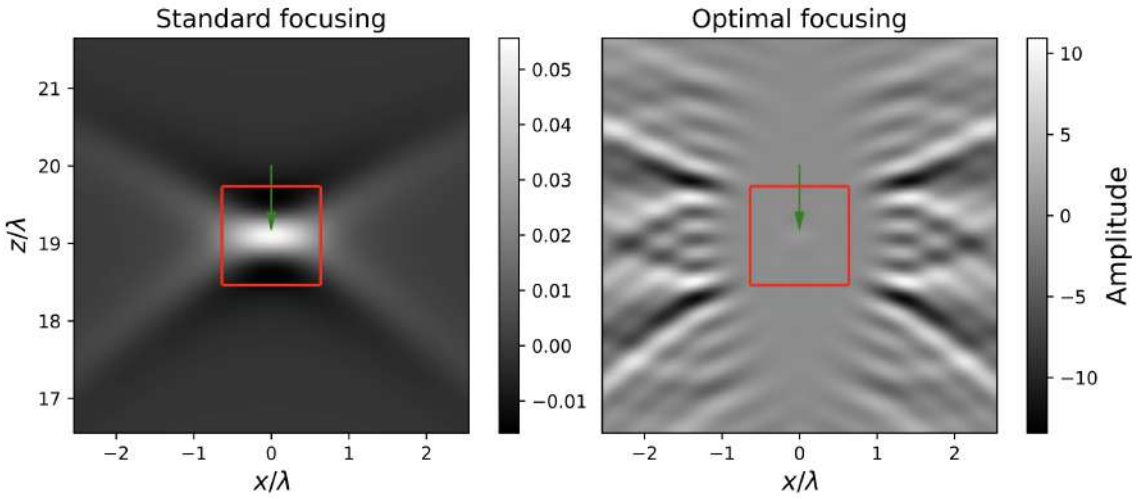

Figure 2.4 Comparison of standard TR (left) and optimal TR images (right) for $\lambda \sim 80$m for a region larger than FOV. The red box indicates the FOV which corresponds to the same dimensions as Figure 2.2. The green arrow indicates the location of the focal spot

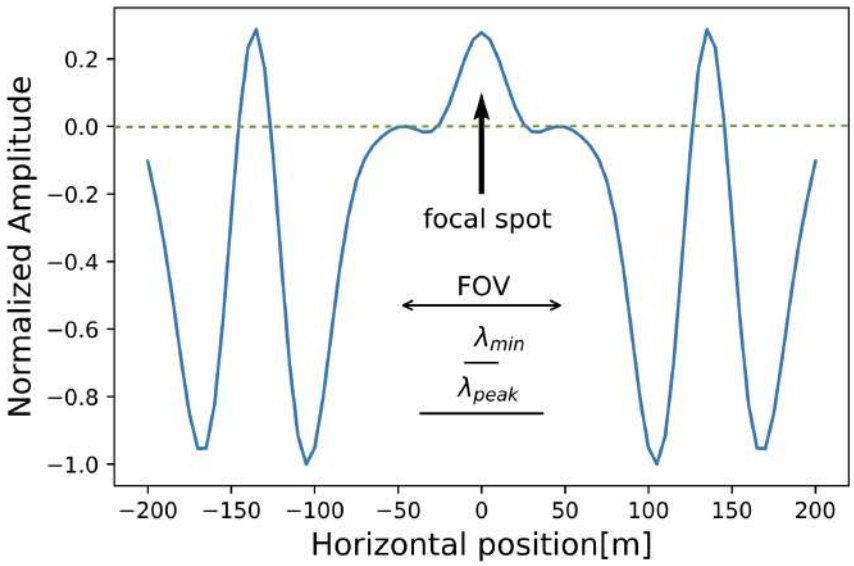

Figure 2.5 Cross section at source level of optimal focusing using an imaging area with horizontal size of 400 m around the source location (-200 to 200 m).

More recently, Rogers and Zheludev (2013) showed that super-oscillatory lenses, which are developed to achieve optimal focusing, allow for imaging beyond the diffraction limit but produce adjacent side-bands as the field of view is increased, a consequence of super-oscillations. These oscillations arise when the wave field oscillates faster than its fastest fourier component, and typically occur in regions where the wave field takes on small values (Berry and Popescu, 2006). To visualize super-oscillations we plot the natural

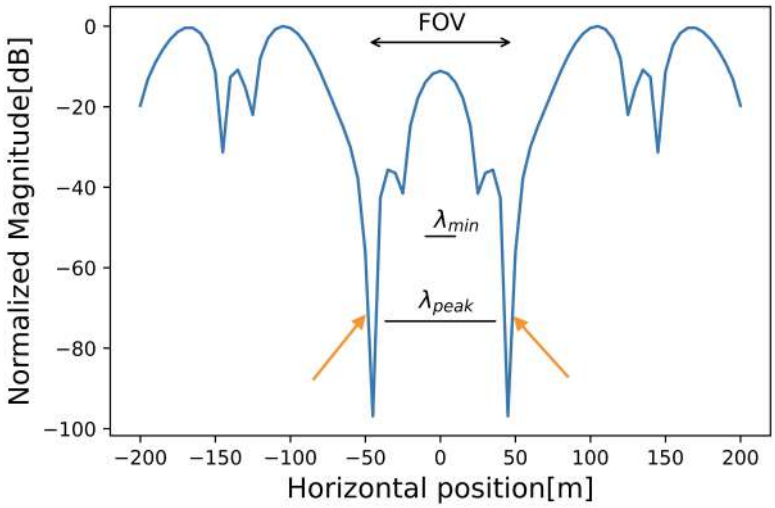

Figure 2.6 Magnitude plot of the cross-section in Figure 2.5 . The orange arrows indicate areas of super-oscillations.

logarithm of the absolute value of the time-reversed wave field as a function of space.

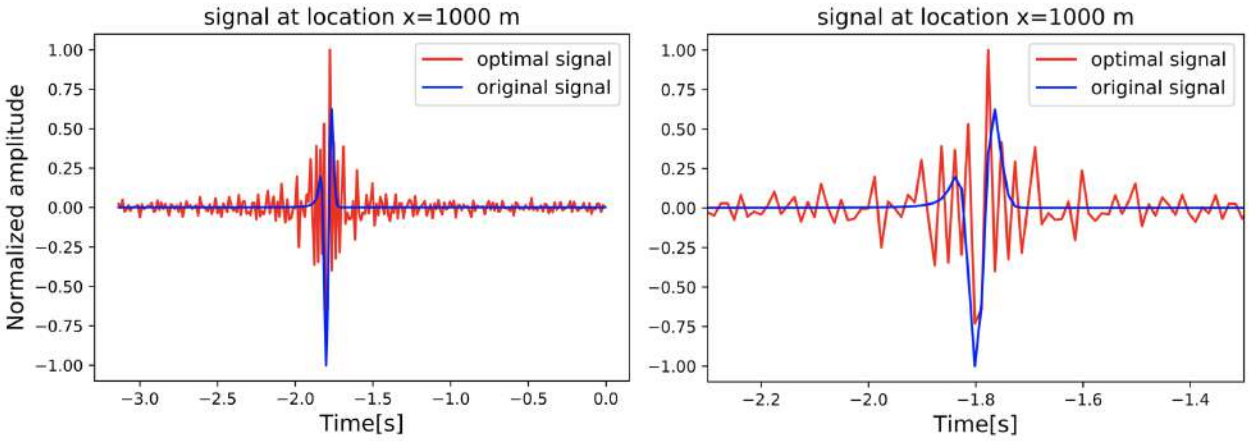

Figure 2.7 Comparison of original $\big($i.e., $G(\mathbf{x}_i, -t_0, \mathbf{x}_0, -t) * S(-t)\big)$ and optimal $\big($i.e., $\sigma_i(\mathbf{x}_i, t) * G(\mathbf{x}_i, -t_0, \mathbf{x}_0, -t) * S(-t)\big)$ signals for receiver at $x = 1000$ m with normal view (left) and zoom-in view (right). The focusing time at the source location corresponds to $t = 0$.

Figure 2.6 shows the magnitude plot in decibels (dB) of the cross-section in Figure 2.5. Notice that regardless of how large we make our imaging area, the optimal wave field oscillates most rapidly in the vicinity of the source. Near the source location, where the optimization takes place, the optimal wave field oscillates faster than the smallest wavelength. These fast oscillations have a high frequency content that ultimately allows us to improve the resolution of the focal spot, indicating a connection between temporal and spatial resolution.

Now we consider the optimal and standard time-domain wave forms to visualize the optimal and standard signals. Figure 2.7 shows the optimized time-reversed signal that is sent back from a receiver located at $x = 1000$ m. Notice the enrichment of high frequencies, which results in sharper focusing. However, this is not the only mechanism by which our algorithm improves the resolution of the focal spot. As we show in section 2.6, the focusing is also improved when the algorithm is applied to monochromatic data, indicating that the resolution is improved at all frequencies. Periodic oscillations appear before and after the direct pulse at $t = -1.8$ s, with different temporal spacing at different receivers. Only oscillations within one dominant period around the direct arrival are necessary for focusing at the source location and focusing time $t = 0$. For earlier and later times the reverberations do not contribute to focusing at $t = 0$ and are a result of imposing spatio-temporal focusing instead of only spatial focusing.

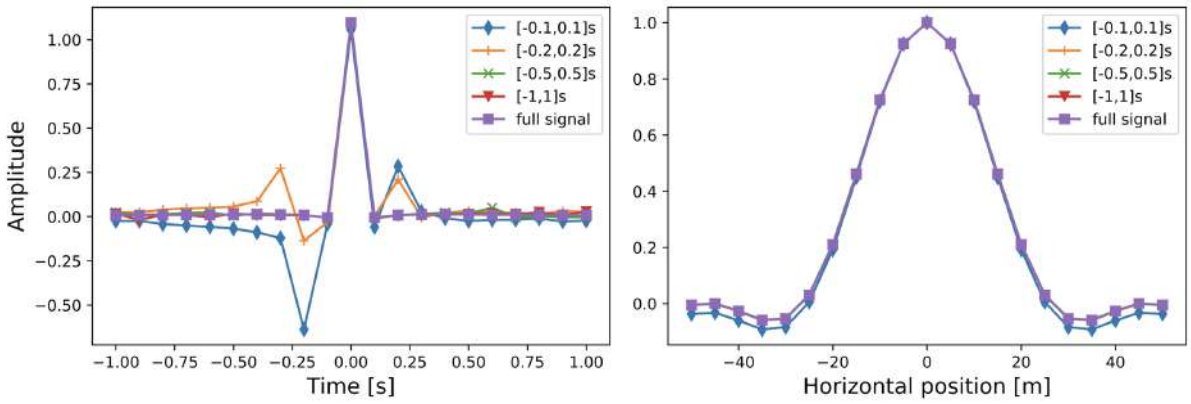

Figure 2.8 Amplitude at source location as a function of time for different window sizes around the direct arrival in Figure 2.7 (left). Cross-section at source level for different window sizes around the direct arrival in Figure 2.7 (right).

To validate this reasoning we consider the temporal compression of the optimal time-reversed wave field at the source location for different sizes of the time-window used of the signals that are sent into the medium. The left panel of Figure 2.8 shows the amplitude of the time-reversed wave field at the source location as a function of time for different window sizes around the direct arrival at t = -1.8 s, where the dominant period is 0.075 s. Notice that as we decrease the size of the window around the direct arrival the

temporal delta function is distorted (there are now nonzero amplitudes at $t \neq 0$). This observation implies that we need these reverberations to achieve temporal focusing. The right panel of Figure 2.8 shows cross sections at the source level for the wave field focused with the signals radiated from the different window sizes. Notice that as we decrease the window size the spatial focusing at time $t = 0$ remains mostly unaffected (i.e., no significant change in the focal spot). The reverberations outside the windows do not affect the amplitude of time-reversed wave field at the source location at $t = 0$ but distort the amplitude of the wave field at other times for the same location. This means that these reverberations are necessary to maintain the temporal focusing but do not affect spatial focusing.

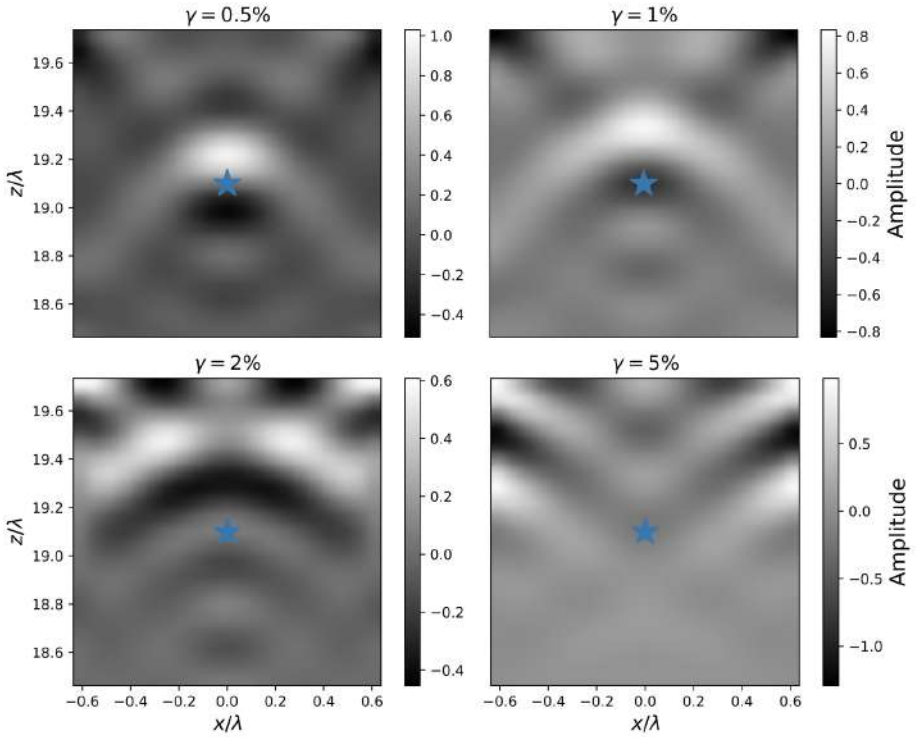

Figure 2.9 Optimal time-reversed images for different levels of velocity perturbation. The blue star indicates the true source location.

2.5 Sensitivity Analyses

To investigate the robustness of the optimal focusing we consider variations in the background velocity and source location, as well as the effect of adding band-limited noise to the recorded signal before time-reversal.

2.5.1 Velocity Perturbation

First, we calculate the weights for the true velocity of the medium. We then perform the back-propagation using a perturbed velocity. Denoting γ as the percentage variation of the true velocity, we consider $\gamma = 0.5$, 1.0, 2.0, and 5.0%. Figure 2.9 shows that the quality of focusing degrades with increasing velocity perturbation. An erroneous velocity gives the wrong source location and increases the width of the focal spot. This means that if one wants to achieve accurate focusing (both in terms of location and size of the focal spot), one needs to know the velocity of the medium with great accuracy. For our experimental set-up we need to know the velocity of the medium within a 1% error such that the focal spot is shifted but not degraded. This sensitivity to velocity errors is similar in the unweighted time-reversed wave field.

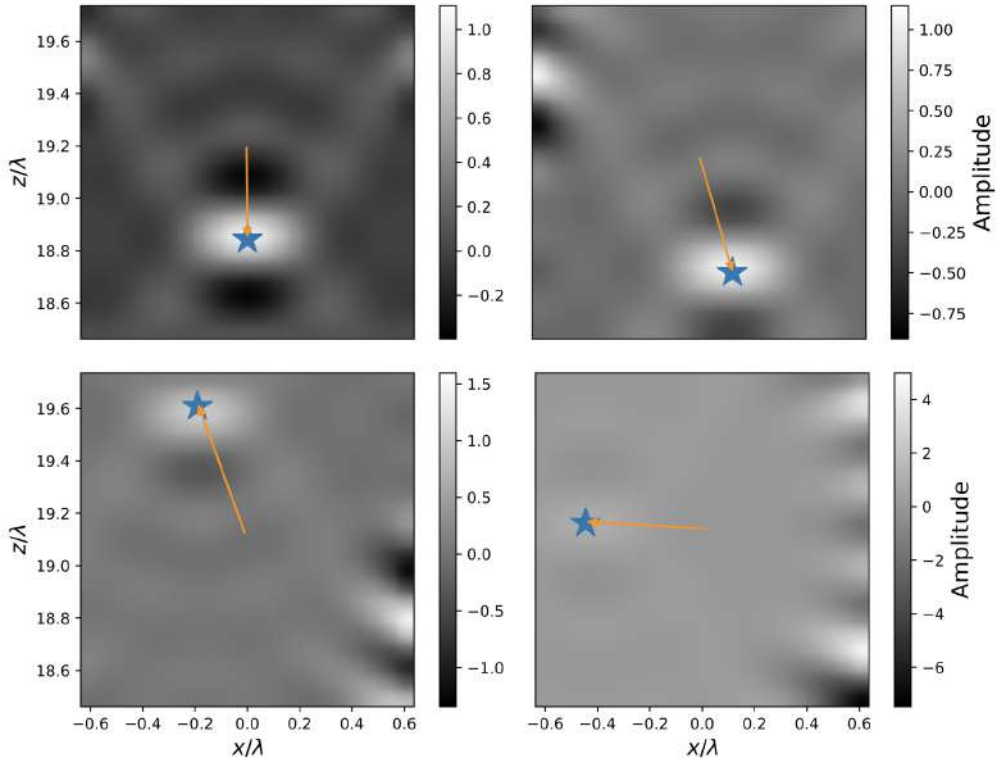

Figure 2.10 Optimal time-reversed images for different source locations using the weights corresponding to the reference source location (x_0, z_0)=(0,1500). The orange arrow provides a measure of the distance between the reference location used to compute the weights and the tested source location.

2.5.2 Source location

Now suppose that the tested and reference source locations are different. To study this scenario we first calculate the weights corresponding to a fixed source location (center of the FOV). Using those weights we perform the time-reversal process for waves which originate from other locations within the field of view.

Figure 2.10 shows that the optimal weights allow for focusing of events at locations different than the reference location for which the weights were calculated. Artifacts on the boundary of FOV arise when the distance between the new and original source location increases. Notice that the imaging is more sensitive to horizontal than vertical changes of the source location. This is due to the fact that a horizontal shift induces a relative phase shift among receivers, whereas a vertical shift in source location induces a constant phase shift across all receivers because of the symmetry of the experiment.

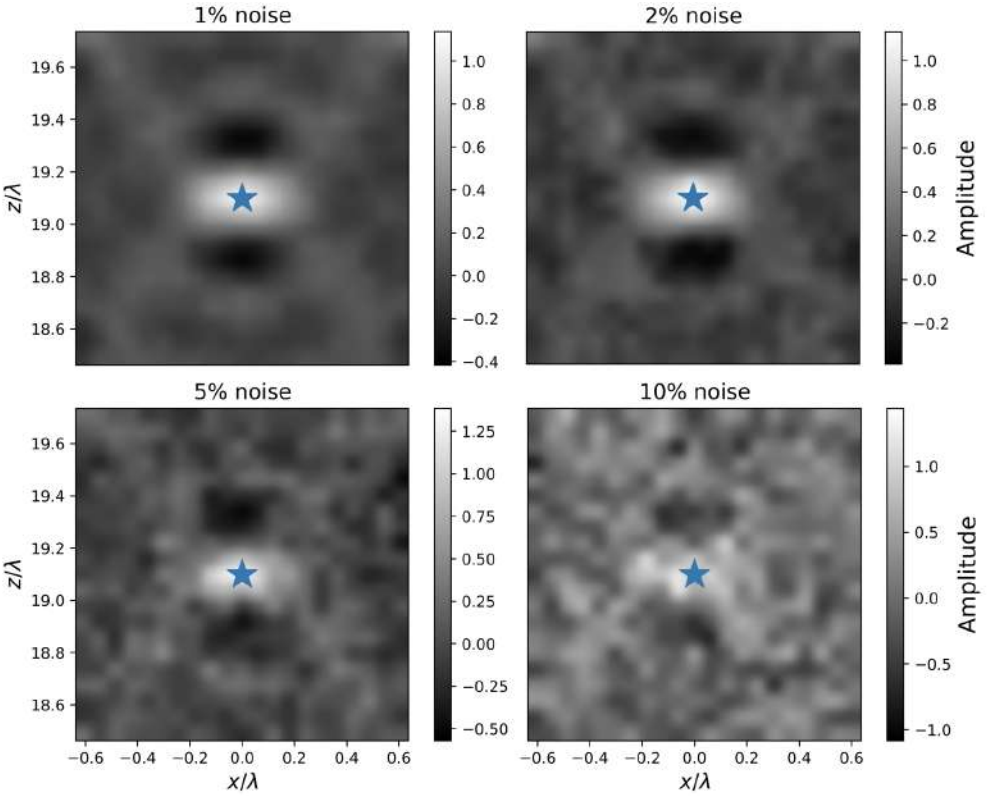

Figure 2.11 Optimal time-reversed wave field for different noise levels.

2.5.3 Noisy Inversion

Let us consider the effect of noisy recorded data. To this end we add band-limited (0.32 to 40 Hz) uncorrelated random noise to the recorded signal, defined as $\frac{\text{mean}(|\text{noise}|)}{\text{mean}(|\text{signal}|)} \times 100 = \text{noise } \%$. We then perform optimal imaging using the weights calculated for data without noise. One would expect the weights to amplify the noise, and consequently the optimal TR image to be degraded. Figure 2.11 and Figure 2.12 show that noise degrades the quality of focusing. We can introduce up to 5% noise without distorting the focal spot significantly. Our algorithm is sensitive to noise, especially if the noise has high frequency content. This occurs because we are amplifying higher frequencies. To make optimal imaging

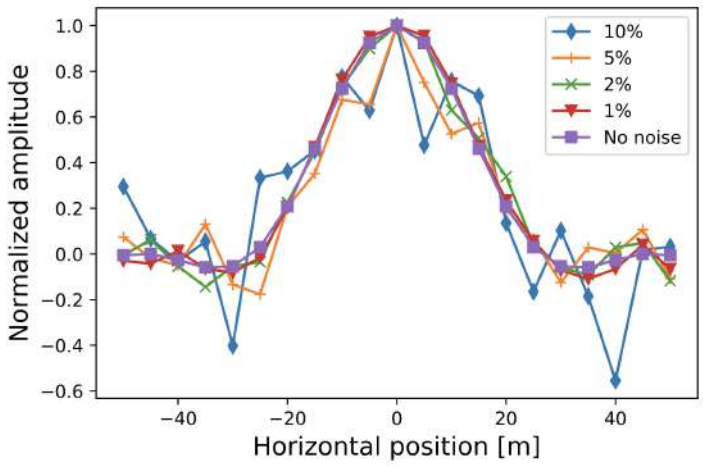

Figure 2.12 Cross sections of optimal time-reversed wave fields from Figure 2.11.

more robust to noise one could filter out the highest frequencies at the expense of resolution improvement.

2.6 Resolution Analyses

We make use of the full-width half-maximum (FWHM) to measure the resolution improvement. For simplicity we make the resolution comparisons in terms of amplitudes rather than intensities. As we discussed in section 2.4 (Figure 2.3), the standard focusing gives a FWHM of 62m whereas the optimal focusing gives a FWHM of 29m, a resolution improvement of more than a factor of 2. To avoid ambiguity in dealing with the resolution of a polychromatic wave field we show the cross-section at the source level of the TR wave field for several frequencies in Figure 2.13. This way of analyzing resolution allows us to understand how well our algorithm improves the resolution at the focal spot (i.e., considering a monochromatic rather than a polychromatic wave field), and confirms that rather than just boosting the high frequency content of the signal, we improve the resolution at all frequencies by the used weighting.

Figure 2.13 shows the resolution improvement at different individual frequencies. Even for each single frequency the focusing is improved. Notice that the improvement is most pronounced at low frequencies, where the resolution is poor. At higher frequencies there is still an improvement in the compression of the focal spot but not as significant as for lower frequencies. The way our algorithm is constructed (i.e., frequency by frequency) allows us to improve the resolution at every single frequency, with the amount of improvement varying across the frequency spectrum.

Alternatively, there are two other ways in which we can empirically study resolution improvement. The first approach consists of using a horizontal dipole source in generating the source-receiver data rather than an impulse monopole source, and perform the time-reversal operation using the optimal weights calculated

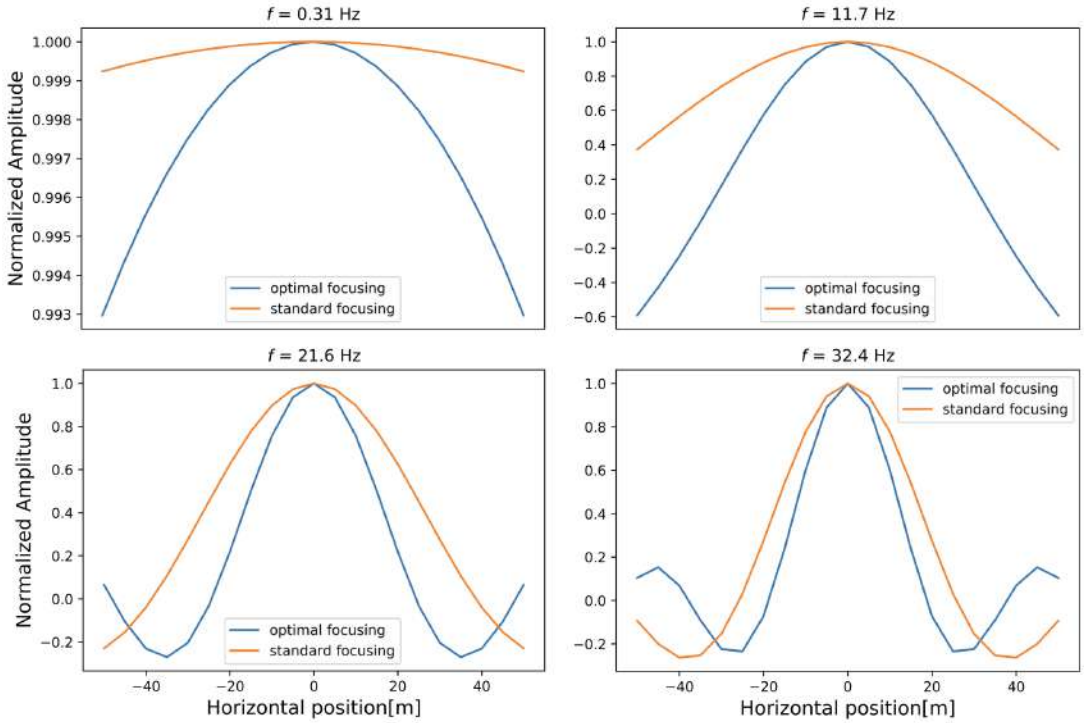

Figure 2.13 Cross-sections of time-reversed wave fields for standard and optimal wave fields at the source level at focusing time for different frequencies.

for the spatio-temporal delta function. Assuming infinite resolution, we should obtain a mathematical dipole (i.e., infinitesimally small) after performing the TR operation. However, since resolution is limited, we can use the horizontal separation between the positive and negative ends of the dipole to measure resolution. Figure 2.14 shows the standard and optimal TR wave fields for a horizontal dipole source, and their corresponding horizontal cross-sections at the source level. Standard imaging shows a separation of 70 m between minimum and maximum, whereas optimal imaging shows a separation of just 30 m. This corresponds to improving the resolution by a factor of 2.3. Notice also that the optimal time-reversed wave field shows improvement of the vertical resolution. The second approach consists of increasing the upper end of the frequency bandwidth as well as the peak frequency of the source wavelet until the cross-section of the standard TR wave field at the source level matches most closely that of the optimal wave field computed with the initial frequency bandwidth. Figure 2.15 shows that for the optimal and standard waveforms to match we need to increment the maximum and the peak frequency in the standard TR experiment. Notice that for the focal spots to match we need to go beyond the original maximum frequency of 40 Hz.

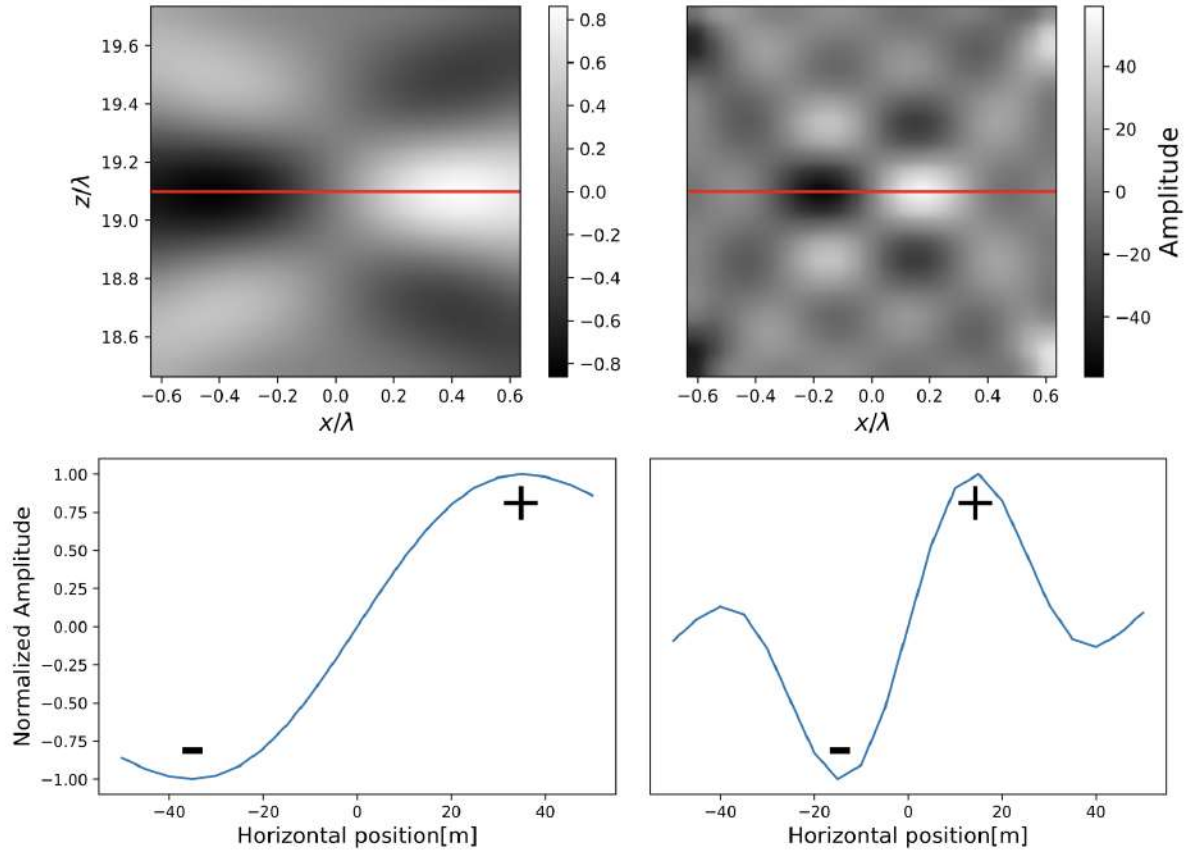

Figure 2.14 Time-reversed standard (upper left) and optimal (upper right) wave fields for a horizontal dipole source, and their corresponding standard (lower left) and optimal (lower right) horizontal cross-sections at the source level. The red horizontal line marks the source level and the location of the cross sections in the lower panels. The negative and positive signs indicate minimum and maximum, respectively.

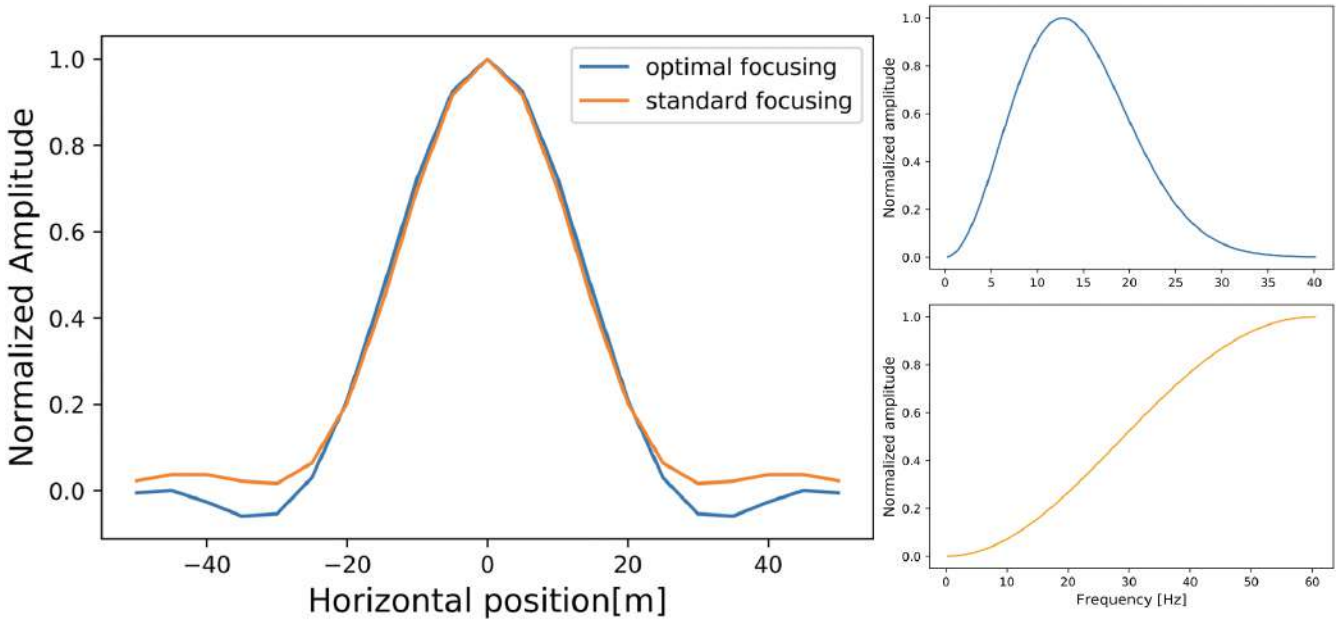

Figure 2.15 Cross section at source level for standard and optimal focusing. The standard focusing waveform is computed using the source wavelet in the lower right panel ($f_{peak} = f_{max} = 60.47$ Hz). The optimal waveform is computed using the source wavelet in the upper right panel ($f_{peak} = 12.73$ Hz, $f_{max} = 40$ Hz).

2.7 Variation of optimal weights

2.7.1 Frequency variation of the optimal weights

We begin by considering how the magnitude of the weights change with frequency for several spatial locations. This is useful to understand the effect of the weights in the improved resolution of the time-reversed wave field as a function of frequency. For illustrative purposes we only consider receivers located at $x = 200, 500, 800, 1000$ m. Figure 2.16 shows the natural logarithm of the frequency spectrum of the weights at four different locations. Despite their seemingly complicated nature, the high frequencies are amplified but the signal at lower frequencies is modulated, as previously explained in section 2.6. At all frequencies the resolution is improved which ultimately allows us to significantly reduce the size of the focal spot of the time-reversed wave field.

2.7.2 Spatial variation of the optimal weights

To appreciate the contribution of the different receivers to the improved resolution of the time-reversed wave field we show in Figure 2.17 the spatial variation of the real part of the weights across the receiver array for different frequencies. We omit the imaginary part since it has amplitudes much smaller than its real counterpart ($\simeq 5\%$). First, notice that the weights are symmetric around the source position ($x = 0$).

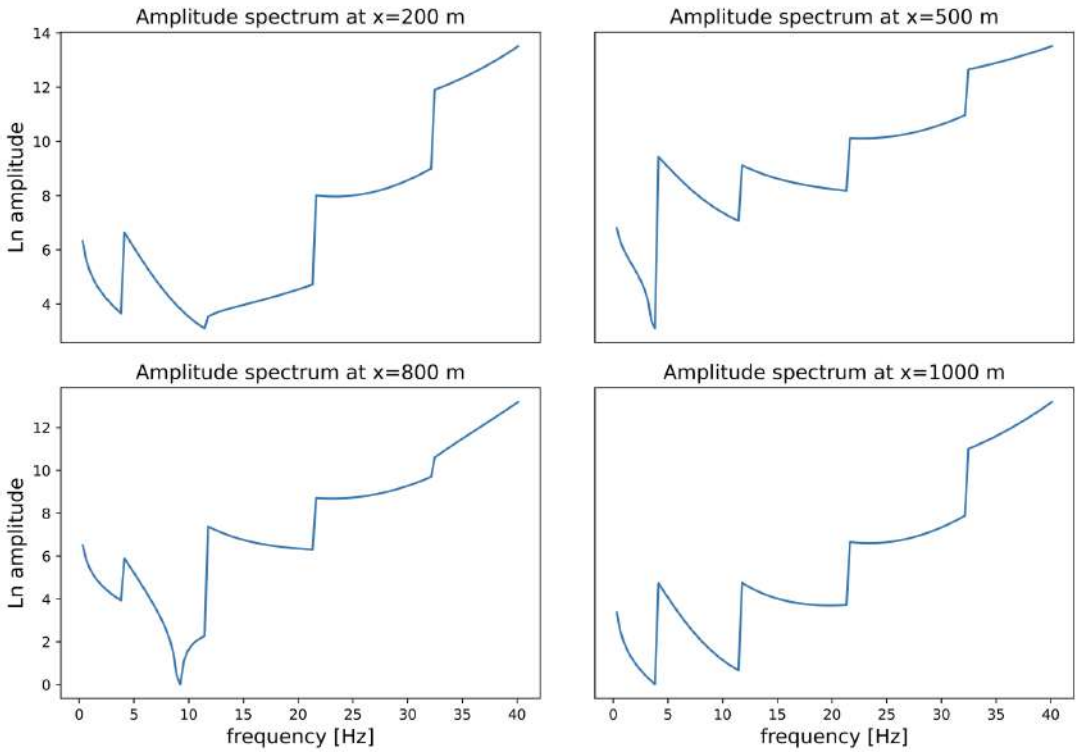

Figure 2.16 Natural logarithm of the frequency spectrum of the weights at four different locations.

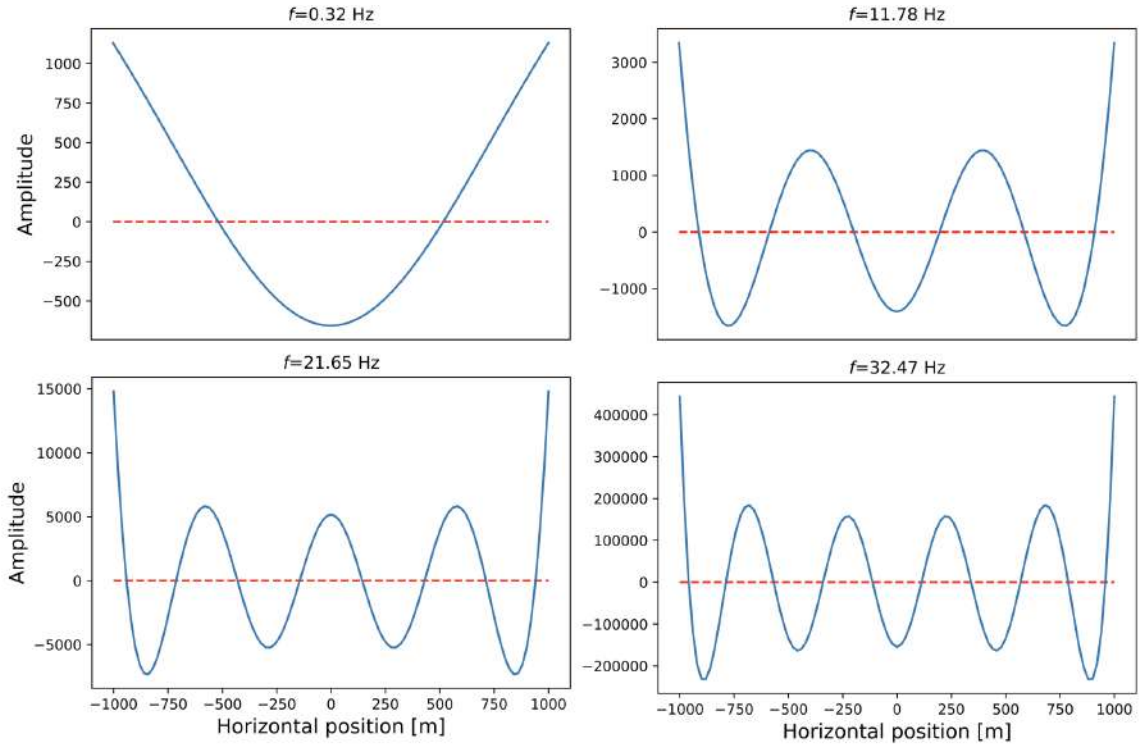

Figure 2.17 Spatial variation of the real part of weights across the receiver array for different frequencies.

This is expected from the symmetric array geometry. The weights are oscillatory in space with a fairly constant wavelength across the array with a weak amplitude modulation, except near the end points of the array. Furthermore, the weights oscillate faster in space as the temporal frequency f increases. Figure 2.17 also shows that the highest amplitude across the array increases with frequency. This means that the algorithm boosts the higher frequencies, which together with the modulation of all other frequencies results in resolution improvement of the focal spot for broadband waveforms.

As mentioned previously, the imaginary part of the weights is much smaller than the real part at all frequencies. This implies that rather than applying sophisticated phase changes, it is only necessary to multiply each receiver data by real weights with a variable sign. Francia (1952) showed that it is possible to improve the resolution of a focal spot by superimposing different ring apertures. To superimpose the different ring apertures he assigns real coefficients, which signify aperture transparency, to each of the rings. The similarity between his and our results suggests that, as mentioned before, multiplying the receiver data by real weights with a variable sign suffices to achieve improved resolution. In addition, the weights shown in Figure 2.17 resemble the Tchebyscheff polynomials widely used in beam forming Dolph (1946). As mentioned above, the weights retrieved by our algorithm are almost real, which is in agreement with the work of Dolph (1946). The fact that our algorithm produces these weights in the case of a homogeneous medium and a uniform line array establishes a connection with the optimal weights derived by Dolph (1946) for a regular array in a homogeneous medium. Our algorithm, which converges to the Tchebyscheff polynomials in our set-up, can also be applied to irregular arrays, as opposed to using the Tchebyscheff polynomials for beam forming. Gallaudet and de Moustier (2000) show an example of the modifications that one must perform when doing beam forming via Tchebyscheff polynomials with irregular arrays.

2.8 Focusing in scattering media

To test the robustness and flexibility of our algorithm we now consider an inhomogeneous medium which contains scatterers and an irregular receiver line. We remind the reader that rather than comparing the size of the focal spot to the results obtained in section 2.4, the aim of this section is to demonstrate that our algorithm improves the resolution of the focal spot even when the medium is complicated and the receiver line is irregular. We illustrate the geometrical set-up in Figure 2.18. We place 50 isotropic scatterers randomly between 400 and 1400 m in the vertical direction and -200 to 200 m in the horizontal direction. The location of the source and the coordinates of FOV as well as the grid spacing are the same as in the experiments in section 2.4. We place 30 receivers at random locations along a horizontal line at $z = 0$ between -300 and 300 m, which allows us to investigate focusing with irregular line arrays. We compute the wave field that propagates from the source to the receivers using the multiple isotropic point

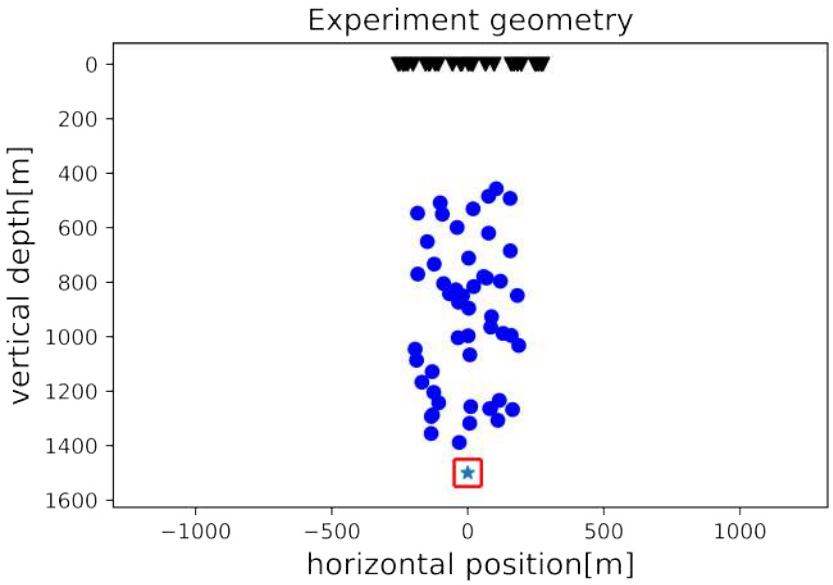

Figure 2.18 Illustration of the geometrical set-up for focusing with embedded scatterers and an irregular receiver array.The blue star indicates the location of the source, the red box the field of view, the blue dots the locations of the point scatterers, and the black triangles the receiver array elements.

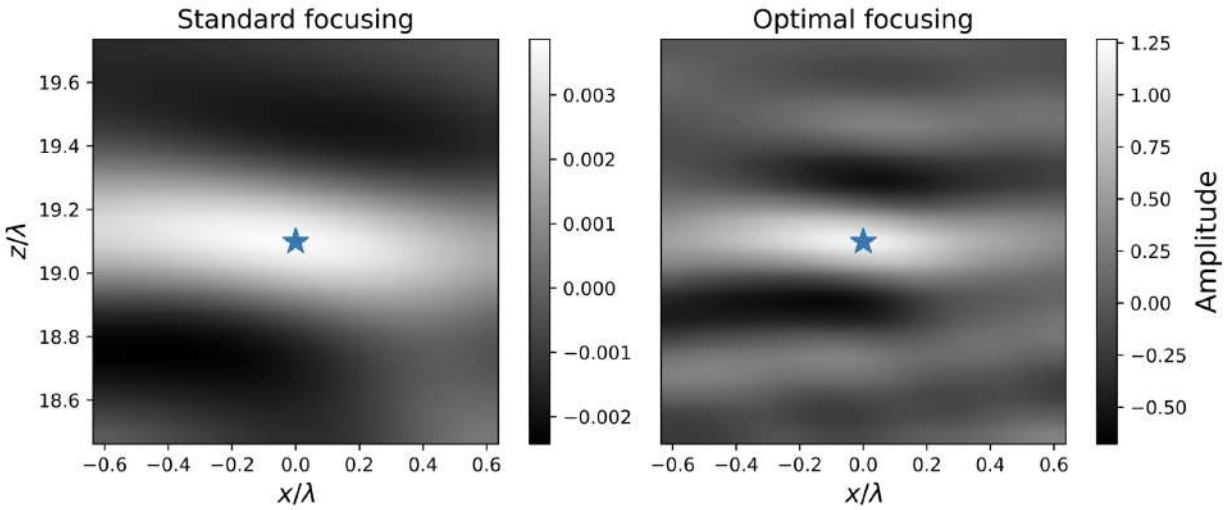

Figure 2.19 Comparison of standard TR (left) and optimal TR images (right) for $\lambda \sim 80$m at FOV in the presence of scatterers. The blue star indicates the source location.

scattering (MIPS) solution (Groenenboom and Snieder, 1995). This approximation is based on Foldy's method (Foldy, 1945), and consists of describing the total wave field in the frequency domain as a sum of the direct wave and the scattered wave field from a set of scatterers

$$\Psi(\mathbf{r}) = \Psi_0(\mathbf{r}) + \sum_{i=1}^{n} G^{(0)}(\mathbf{r}, \mathbf{r}_i) A_i \Psi(\mathbf{r}_i), \tag{2.11}$$

where n is the number of scatterers, A_i is the ith scattering amplitude which is a complex number in the frequency domain, and $G^{(0)}(\mathbf{r}, \mathbf{r}_i)$ is the 2-D Green's function used in section 2.4. For details regarding the numerical computations of the total wave field under the MIPS approximation see the appendix in (Groenenboom and Snieder, 1995). We choose $|A_i| = 3.0$ for all scatterers, and compute the imaginary and real components of the scattering amplitude as $\Im(A_i) = -|A_i|^2/4$ and $\Re(A_i) = \sqrt{-\Im(A_i)(4 + \Im(A_i))}$, respectively. The velocity of the background medium, the frequency parameters and the source signal are the same as in section 2.4. The procedure for computing the standard and optimal time-reversed wave fields is the same as before, which facilitates the optimization.

Figure 2.19 shows both standard (left-panel) and optimal (right-panel) focusing in the presence of embedded scatterers. As for the examples in section 2.4 for a homogenous medium and a regularly spaced receiver array, our algorithm improves the resolution of the focal spot. We attain resolution enhancement both along the vertical and horizontal directions. Our algorithm thus performs well in both homogeneous and inhomogeneous media, and we are able to reduce the size of the focal spot in both cases. In addition, our algorithm is able to handle irregular arrays without any modifications. Contrary to focusing in a homogeneous medium with a uniform line array, the optimal weights do not resemble Tchebyscheff polynomials, instead they have a complicated oscillatory structure which we refrain from showing here.

2.9 Discussion

The focusing method we propose allows us to improve the resolution of focusing in a time-reversal numerical simulation, and may be extended to other focusing functions and source distributions. The extent to which the resolution is improved will ultimately depend on the Green function which describes the propagation of the wave field to be time reversed and the type of model that we use (strongly vs weakly scattering media). The optimization of the focal spot is confined to the field of view, as shown in section 2.4, Figure 2.4. Outside of the field of view the amplitude of the optimal time-reversed wave field is significantly larger than inside the FOV and masks the focal spot. This is related to the concept of super-gain antennas with which it is possible to increase the resolution of the focal spot in a given FOV at the expense of larger amplitudes outside FOV (Francia, 1952). Section 2.5 addresses the applicability of our method. Firstly, we point out that we are able to handle weak velocity errors, as illustrated in

Figure 2.9. We are also able to obtain good focusing in test locations within FOV different to the reference source location (Figure 2.10). For the used geometry our method is able to handle band-limited noise up to 5% (Figure 2.11). One could filter out the highest frequencies at the expense of resolution improvement to make the optimal imaging more robust to noise. Our resolution analyses show that we do not only boost high frequencies but also improve the resolution at all frequencies (Figure 2.13). We are also able to obtain a significant resolution improvement for a horizontal dipole source which is very promising for applications that involve complex sources (Figure 2.14). We find that for standard TR to closely match optimal TR we have to increase the upper end of the frequency bandwidth as well as the peak frequency of the source wavelet by a significant amount (Figure 2.15). The frequency spectra of the optimal weights change in a complicated manner but illustrate that we do not just boost high frequencies; we improve the focusing at each individual frequency (Figure 2.16). Lastly, we show that the weights oscillate faster in space as the temporal frequency increases, providing a link between spatial and temporal frequency improvement. For a homogeneous medium and a uniform line array we find that for a given frequency the weights have a fairly constant wavelength across the array with weak amplitude modulation (Figure 2.17). In this set-up the optimal weights resemble the Tchebyscheff polynomials used in earlier beam forming. Our algorithm does not require any extra modifications when dealing with irregular arrays or inhomogeneous media, as shown in section 2.8.

2.10 Conclusions

We have shown that it is possible to locally improve the resolution of the TR wave field in both homogeneous and heterogeneous media, even at very large distances from the source (19 dominant wavelengths) by applying a complex filter to the signal before time-reversal. In practice this is similar to techniques such as beam forming, spatial light modulation or antenna focusing (Xiao et al., 2010), where some type of spatial filtering is applied. Our technique is directly applicable to polychromatic wave fields as opposed to the techniques mentioned above and the focusing is achieved in space and time. In this paper we optimize the signal to be time-reversed rather than asking for an arbitrarily optimal signal that allows focusing after propagation of the wave field through an imaging system. To calculate the optimal weights we use a modified Backus-Gilbert approach which guarantees the focal spot to be localized at the source. Neither evanescent waves nor resonant multiples are needed to achieve enhanced focusing if we know how the original signal must be modulated. We find artifacts remain in the optimal TR images. The weights partially compensate for the limited geometry and the incomplete wave field sampling by modifying the frequency spectrum of the signal that is time-reversed. Optimal imaging is robust to errors in the source location. This is important because we do not always know the exact source location. The

fact that we obtain proper focusing in locations different from the reference source implies that if we have full information of a control point in the field of view we can achieve focusing in other points within the FOV. To achieve proper focusing we also need to know the velocity of the medium accurately, this drawback arises from the nature of time-reversal. Additive noise degrades the quality of focusing, in particular when the noise has high frequency content. In addition, we find that even with optimization techniques, only local focusing is possible. Regardless of the metric used to study resolution enhancement the optimal weights improve the localization of the focal spot by several factors. We showed that the algorithm does not only boost the highest frequencies but that it also increases the resolution of the focal spot for monochromatic waves. The optimal weights oscillate faster across the receiver array as the temporal frequency increases. In addition, the amplitude of the weights increase with frequency, resulting in higher resolution. In the case of a homogeneous medium and a uniform line array our algorithm retrieves weights that resemble the Tchebyscheff polynomials widely used in beam forming. Our proposed method also shows focusing improvement in inhomogeneous media, and no extra modifications are needed when dealing with irregular arrays.

CHAPTER 3

ILLUSTRATION OF DIFFUSION AND EQUIPARTITIONING AS LOCAL PROCESSES: A NUMERICAL STUDY USING THE RADIATIVE TRANSFER EQUATIONS

To be published in the Journal of the Acoustical Society of America (JASA)[4].

Manuel Jaimes-Caballero[5], and Roel Snieder[6].

3.1 Abstract

We study the transition from ballistic to diffusive to equipartitioned waves in scattering media using the acoustic radiative transfer equations. To solve these equations we first transform them into integral equations for the specific intensities, and then construct a time stepping algorithm with which we evolve the specific intensities numerically in time. We handle the advection of energy analytically at the computational grid points and use numerical interpolation to deal with advection terms which do not lie on the grid points. This approach allows us to reduce the numerical dispersion, compared to standard numerical techniques. With this algorithm we are able to model various initial conditions for the intensity field, non-isotropic scattering, and uniform scatterer density. We test this algorithm for an isotropic initial condition, isotropic scattering and uniform scattering density, and find good agreement with analytical solutions. We compare our numerical solutions to known 2-D diffusion approximations, and find good agreement. We use this algorithm to numerically investigate the transition from ballistic to diffusive to equipartitioned wave propagation over space and time, for two different initial conditions. The first one corresponds to an isotropic Gaussian distribution in space and the second one to a plane wave segment. We find that diffusion and equipartitioning must be treated as local rather than global concepts. This local behavior of equipartitioning has implications for Green's functions reconstruction, which is of interest in acoustics and seismology.

3.2 Introduction

As waves propagate through a scattering medium they lose (or gain) energy due to scattering to (or from) other directions. In the absence of anelastic attenuation, this phenomenon obeys energy conservation, which one may mathematically describe using the acoustic radiative transfer equations (RTE). The RTE consist of a coupled system of integro-differential equations where one solves for the wave intensity as a function of space, time and angular direction, assuming one knows the scattering mean free

[4]Accepted pending minor revisions.
[5]Graduate student, primary researcher, and author at Department of Geophysics, Colorado School of Mines.
[6]Supervisor at Department of Geophysics, Colorado School of Mines.

path l, the angular dependence of scattering, and the speed of energy propagation v. Due to its usefulness RTE has been employed in several scientific fields. In astrophysics RTE has been used to analyze radiation transport across cosmic dust in a wide range of astrophysical objects (Steinacker et al., 2002; Narayanan et al., 2021; Wolf, 2003); in atmospheric sciences researchers use RTE to model solar radiation across clouds to better understand the evolution of sea surface temperatures (Evans and Stephens, 1995; Aumann et al., 2018; Manners et al., 2009); in optics RTE has been used to develop novel optical tomographic imaging algorithms which allow diagnosis and treatment of biological tissues (Klose et al., 2002; Abdoulaev, 2003; Ren et al., 2004); in acoustics RTE has been used for modeling the interaction of acoustic waves with the ocean bottom (Quijano and Zurk, 2009), modeling of forest acoustics (Ostashev et al., 2017), and modeling of acoustical diffractions by obstacles (Reboul et al., 2005); in geophysics RTE has been used to study infrared radiation across volcanic ash clouds (Prata, 1989; Francis et al., 2012; Lee et al., 2014), modeling of heat transfer in the mantle (Hofmeister, 2005), computation of scattering kernels in coda wave interferometry (Margerin et al., 2016; Duran et al., 2020; Dinther et al., 2021; Obermann et al., 2016), and estimation of the scattering properties of the crust from coda envelopes (Zeng, 2017; Lacombe et al., 2003; Gaebler et al., 2015).

Despite its usefulness in describing energy transport, RTE is complicated and numerically demanding due to its dependence on not only space and time, but also on the directions of wave propagation. In 2 dimensions the radiative transfer solution depends on 4 variables (time, two space variables, and one angle). In 3 dimensions it depends on 6 variables (time, three space variables, and two angles). Numerical techniques that have been used to solve the RTE include

- the Discontinuous Galerkin finite element method, which has a high order accuracy in space and time but requires one to add numerical viscosity in order to avoid nonphysical oscillations in the numerical solution and is computationally expensive (Clarke et al., 2019; Han et al., 2010).
- Markov Chain Monte Carlo techniques, which are quite accurate and applicable to complicated media but are computationally expensive and rely on using enough phonons to sample the intensity field as a function of space, time, and propagation directions (Iwabuchi, 2006; Xu et al., 2011; Camps and Baes, 2018; Noebauer and Sim, 2019; Przybilla and Korn, 2008; Yoshimoto, 2000).
- Finite Difference techniques, which suffer from numerical dispersion in space and time due to the discretization of spatial and temporal derivatives and may produce negative intensities which are nonphysical (Klose and Hielscher, 1999).

- Wave Equation modeling, whereby one exploits the connection between the acoustic (or elastic) wave equation and the scalar (or elastic) RTE (Przybilla et al., 2006; Kanu and Snieder, 2015; Snieder et al., 2019; Duran et al., 2020). One solves the acoustic wave equation for several realization of a medium and computes a statistical average of the intensity field to suppress statistical fluctuations of the intensity field. However, the wave field must be averaged over enough realizations of the medium and one must locally decompose the calculated wave field into different propagation directions to properly account for the directionality of the wave energy transport.

In addition to these numerical techniques there have been analytical approximations to RTE such as assuming diffusive wave propagation which is only valid at times much larger than the transport mean free time (Rossetto et al., 2011; Planès et al., 2014); assuming a point-like, isotropic, impulsive source of intensity in a statistically homogeneous medium (Margerin et al., 2016); expanding the intensity and scattering function into a finite sum of Legendre polynomial and then solving a finite system of equations for the unknown coefficients appearing in a truncated expansion (Roberge, 1983); decomposing the specific intensities into a sum of partial intensities and then solving the RTE for each partial intensity assuming that scattering is angle-independent and that the source is isotropic (Paasschens, 1997); assuming a steady-state intensity field (Fan et al., 2019; Le Hardy et al., 2016); or assuming azimuthal symmetry (Baes and Dejonghe, 2001; de Abreu, 2004). In search of a numerical method that overcomes some of the limitations mentioned above we propose a numerical scheme to compute the numerical intensities which solve the scalar RTE in 2-D. Our approach is based on transforming the 2-D scalar RTE into an integral equation for the specific intensities and then solving this integral equation as a time-stepping algorithm. Contrary to standard finite difference (FD) or discontinuous Galerkin finite element (DGFE) methods, we handle the advection of energy analytically at the computational grid points and use numerical interpolation to deal with the advection terms that do not lie on the grid points. This means that our numerical procedure for computing the wave energy transport reduces the amount of numerical dispersion, compared to the numerical schemes mentioned above. Further, since the time discretization depends on the transport mean free time, instead of the frequency of the waves, the space and time discretization in the numerical modeling of the radiative transfer equation can be much coarser than it is for wave field modeling.

Because with this algorithm we are able to describe the spatio-temporal-angular distribution of energy propagation of acoustic waves in a scattering medium, we use our numerical solutions to study the transition from ballistic to diffuse to equipartitioned waves. Since we only deal with acoustic waves, we refer to *equipartitioning* in this chapter as the equal angular distribution of energy. Notice that this

definition for equipartitioning is different to the usual definition for elastic waves which also considers the equilibration of energy between wave modes.

Equipartitioning, of both acoustic and elastic waves, has been experimentally, numerically, and theoretically studied in fields such as acoustics and seismology due to its importance in Green's functions (GF) reconstruction. Weaver and Lobkis (2001) proposed to perform such reconstruction in an ultrasonic experiment using thermal field fluctuations, assuming that such field is equipartitioned. This work inspired the development of *seismic interferometry* (Curtis et al., 2006; Roux et al., 2005; Snieder and Larose, 2013). In this technique one retrieves the causal and anti-causal GF between receiver locations A and B, by cross-correlating the wave fields recorded at such locations due to an even distribution of sources, assuming that the wave field has no preferred propagation direction. Snieder et al.(2007) illustrated, heuristically, the role that equipartitioning plays in the GF retrieval, and Snieder et al. (2010) showed that equipartitioning is necessary for GF retrieval but not sufficient. Paul (2005) computed field-field correlations of Alaskan earthquakes to retrieve the causal and anti-causal GF. In theory, if the noise wave field were equipartitioned, these GF should be symmetric in time. However, they found that the correlations they computed were asymmetric in time, which they suggest is due to the preferred direction of energy flow of the coda waves that they use to compute the field-field correlations. When they use late coda they find that this asymmetry remains, but to a lesser extent than when using earlier coda. Stehly et al. (2006) showed that the noise, which one may use to reconstruct GF near coastlines, radiates preferentially away from storms at the oceans. This preferential directionality of the noise radiation affects the symmetry of the GF that one retrieves, especially if the noise field that one uses is uni-directional and not evenly distributed. Malcolm et al. (2004) experimentally reconstructed, in a laboratory setting, GF in a highly scattering medium using time windows ranging from ballistic to diffusive to equipartitioned waves. They find that the reconstruction of the causal and anti-causal components of GF is only possible when the waves are equipartitioned. Sens-Schönfelder et al.(2015) studied the equipartitioning of the late coda in teleseismic events. They find that even at late times the seismic waves propagate with a preferred direction. This directionality causes seismic interferometry to fail in reconstructing GF unless both the earthquake and receiver are located along the preferred direction of propagation. In his educational paper Weaver (2010) introduced, conceptually, the idea of local equipartitioning for open systems. Notice that his definition of equipartitioning is broader than the one in this chapter. He argues that open systems can achieve local equipartitioning and that this gives cross-correlations which only resemble the GF locally.

The purpose of the work we present here is to numerically study this concept of local equipartitioning (i.e., equilibration along the angular directions) using the acoustic RTE. Local angular equipartitioning has not been thoroughly studied before since the observable that one has typically access to is the total

intensity rather than the individual specific intensities. In this chapter we refer to the *diffusive state* as that in which there is a net flux of energy away from the source, and the *equipartitioned state* as that in which the net flux goes to zero, which results in the energy being distributed equally among all propagation directions. This chapter is organized as follows: In subsection 3.3 we discuss the scalar RTE and derive a localized time-stepping algorithm. In subsection 3.4 we show numerical simulations for isotropic scattering, and compare the computed intensities to exact solutions. We also show how our algorithm handles arbitrary source distributions and angle-dependent scattering, and compare angle-dependent scattering numerical solutions to diffusive approximations. We discuss the transition from ballistic to diffusive to equipartitioned waves. To quantify this transition we introduce the equipartitioning index, a quantity which ranges from 0 (full equipartitioning) to 1 (uni-directional wave propagation), and emphasize that the transition from ballistic to diffuse to equipartitioned waves occurs locally rather than globally. In subsection 3.5 we discuss the algorithm that we develop, equiparitioning as a local process, and the implications that this has for Green's function retrieval.

3.3 Theory

3.3.1 Description of the Scalar Radiative Transfer Equations

The acoustic radiative transfer equations (RTE) are a system of coupled integro-differential equations which describe the distribution of energy in a scattering medium as a function of space, time, and direction $\hat{\mathbf{n}}$ of wave propagation (Chandrasekhar, 1960; Turner and Weaver, 1994; Ryzhik et al., 1996). For late times, when the wave propagation is almost independent of direction, and is nearly stationary in time, the equation of radiative transfer leads to diffusive wave transport (van Rossum and Nieuwenhuizen, 1999). The radiative transfer equations follow from the ladder approximation to the Bethe-Salpeter equation (van Rossum and Nieuwenhuizen, 1999) and accurately describe wave transport at both early and late times, as well as the transition from ballistic wave propagation to weak scattering to strong multiple scattering (Paasschens, 1997). In the derivation of the RTE one assumes that scattered waves have uncorrelated random phases, resulting in the description of scattering processes in terms of average intensities rather than in terms of stress or material displacement (Turner and Weaver, 1994). To derive the RTE one may either make use of energy balance in an infinitesimal volume containing scatterers, or directly derive these equations from the wave equations by estimating the ensemble average of the covariance of the Green's function in a random medium (Turner and Weaver, 1994). For a thorough description of the radiative transfer (transport) equations in random media we refer the reader to Ryzhik et al. (1996).

Assuming a uniform speed of transport and ignoring intrinsic attenuation the scalar radiative transfer equation in 2 dimensions for an intensity field reads (Liu, 1994)

$$\frac{\partial I(\mathbf{r},\hat{\mathbf{n}},t)}{\partial t} + v\hat{\mathbf{n}}\cdot\nabla I(\mathbf{r},\hat{\mathbf{n}},t) = -\frac{1}{\tau_s}I(\mathbf{r},\hat{\mathbf{n}},t) + \frac{1}{\tau_s}\oint f(\hat{\mathbf{n}},\hat{\mathbf{n}}')I(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}', \tag{3.1}$$

where the scattering function $f(\hat{\mathbf{n}},\hat{\mathbf{n}}')$ does not explicitly depend on frequency (i.e., the incident waves change angular direction upon scattering but not frequency), which holds for band-limited wavefields. Notice that in equation 3.1 we suppressed the frequency dependence of the waves. We outline in subsection 3.3.3 the role that frequency plays in defining the scattering function, and therefore the frequency dependence of the waves. In equation 3.1 the variable $I(\mathbf{r},\hat{\mathbf{n}},t)$ is the intensity of waves at a location $\mathbf{r}$ and time t propagating in the direction $\hat{\mathbf{n}}$. In the literature of radiative transfer the variable $I(\mathbf{r},\hat{\mathbf{n}},t)$ is referred to as the *specific intensity* (Chandrasekhar, 1960). The advection of the energy, propagating with a wave velocity v, is described by the term $v\hat{\mathbf{n}}\cdot\nabla I(\mathbf{r},\hat{\mathbf{n}},t)$. The parameter τ_s is the scattering mean free time which describes the average time between scattering events. The term $-I(\mathbf{r},\hat{\mathbf{n}},t)/\tau_s$ accounts for the energy that is lost to other directions in the scattering process. The term $\oint f(\hat{\mathbf{n}},\hat{\mathbf{n}}')I(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}'/\tau_s$ describes the gain due to energy scattered from other directions. The scattering function $f(\hat{\mathbf{n}},\hat{\mathbf{n}}')$ relates the incident intensity to scattered intensity. This function is normalized, assuming that there is no absorption (van Rossum and Nieuwenhuizen, 1999), such that

$$\oint f(\hat{\mathbf{n}},\hat{\mathbf{n}}')d^2\hat{\mathbf{n}}' = 1. \tag{3.2}$$

This normalization follows by analyzing equation 3.1 for a homogeneous and time-independent intensity.

3.3.2 Formulation of the Numerical Algorithm

We solve the 2-D scalar RTE

$$\partial_t I(\mathbf{r},\hat{\mathbf{n}},t) + v\hat{\mathbf{n}}\cdot\nabla I(\mathbf{r},\hat{\mathbf{n}},t) = -I(\mathbf{r},\hat{\mathbf{n}},t)/\tau_s + \oint f(\hat{\mathbf{n}},\hat{\mathbf{n}}')I(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}'/\tau_s, \text{for } t > t_0 \tag{3.3}$$

subject to the initial condition $I_0(\mathbf{r},\hat{\mathbf{n}})$ at an initial time t_0

$$I(\mathbf{r},\hat{\mathbf{n}},t=t_0) = I_0(\mathbf{r},\hat{\mathbf{n}}). \tag{3.4}$$

In equation 3.3 the term $\hat{\mathbf{n}} = (\cos\theta,\sin\theta)$ indicates the direction of propagation, with θ the counter-clockwise angle relative to the horizontal direction ($x-$axis). For simplicity we assume uniform τ_s. If τ_s is a function of space, an algorithm similar to the one we derive here follows. We assume that the computational boundaries are far enough from the source such that the intensities can be assumed to vanish at the boundary. This causes a small leakage of energy around the computational boundary. To avoid such leakage one can either extend the computational domain, so that the vanishing energy assumption holds, or implement a different boundary condition (i.e., reflection or advection) depending on

the problem one aims to solve. One can enforce these boundary conditions using energy conservation. For instance, for the reflective boundary condition, one can require the energy reflecting off the boundary to be the same as the energy that is incident on the boundary. We also assume that the speed of propagation is constant throughout the computational domain. One could in practice extend the algorithm that we show here to handle layered velocity models. To do this one would solve the acoustic RTE independently in each of the layers and then at the interfaces enforce that the sum of the fluxes of the incoming waves be the same as the flux of the outgoing waves. However, this discussion and implementation is out of the scope of this thesis. We refer the reader to Margerin (2005) for a discussion on how to deal with a multitude of boundary conditions using energy conservation arguments.

We transform the initial value problem in equation 3.3 into an integral equation for the specific intensities (Paasschens, 1997)

$$\begin{aligned} I(x,y,\hat{\mathbf{n}},t) &= I(x-v(t-t_0)\cos(\theta), y-v(t-t_0)\sin(\theta), \hat{\mathbf{n}}, t_0)e^{-(t-t_0)/\tau_s} \\ &+ \frac{1}{\tau_s}\int_{t_0}^{t}\oint f(\hat{\mathbf{n}},\hat{\mathbf{n}}')I(x-v(t-t')\cos(\theta), y-v(t-t')\sin(\theta), \hat{\mathbf{n}}', t')e^{-(t-t')/\tau_s}d^2\hat{\mathbf{n}}'dt'. \end{aligned} \tag{3.5}$$

Paasschens (1997) analytically solves for the specific intensities in equation 3.5 for a medium with isotropic scattering and constant scattering properties assuming that the initial condition for the specific intensity is isotropic. To do this, he decomposes the specific intensities into a sum of partial intensities, where each partial intensity signifies a number of scattering events (i.e., direct propagation, single scattering, double scattering, and so on). We show in section 3.7 that equation 3.5 solves the radiative transfer equations 3.3 for the arbitrary initial condition 3.4 and angle-dependent scattering. In this chapter we use equation 3.5 to integrate the specific intensities numerically in time. This approach allows us to use non-isotropic initial conditions and non-isotropic scattering radiation patterns.

In expression 3.5 the first term in the right hand side corresponds to the incident intensity which propagates and decays over time due to scattering losses to other directions. The second term corresponds to the scattered intensity which accounts for the energy gain due to scattering from all possible propagation directions. To develop an iterative algorithm which depends only on the current time t and the previous time step $t-\Delta t$ we replace t_0 by $t-\Delta t$ to obtain

$$\begin{aligned} I(x,y,\hat{\mathbf{n}},t) &= I(x-v\Delta t\cos(\theta), y-v\Delta t\sin(\theta), \hat{\mathbf{n}}, t-\Delta t)e^{-\Delta t/\tau_s} \\ &+ \frac{1}{\tau_s}\int_{t-\Delta t}^{t}\oint f(\hat{\mathbf{n}},\hat{\mathbf{n}}')I(x-v(t-t')\cos(\theta), y-v(t-t')\sin(\theta), \hat{\mathbf{n}}', t')e^{-(t-t')/\tau_s}d^2\hat{\mathbf{n}}'dt'. \end{aligned} \tag{3.6}$$

The contribution of the incident intensity at time $t-\Delta t$ in equation 3.5 now becomes the first term in equation 3.6. The second term in equation 3.6 corresponds to the scattered contributions at current and

previous times. Through this iterative algorithm we compute the advection and scattering of energy locally. Notice that the numerical algorithm could be extended from 2-D to 3-D by writing the angular integration in expression 3.6 in terms of the solid angle $d\hat{\mathbf{\Omega}}$. We proceed by applying the discrete ordinate method (Chandrasekhar, 1960) whereby one divides the angular integral into N equal segments. In two dimensions this method corresponds to setting $d^2\hat{\mathbf{n}} \rightarrow d\theta$ and then setting $d\theta = 2\pi/N$ in the angular integration

$$I(x, y, \hat{\mathbf{n}}, t) = I(x - v\Delta t\cos(\theta), y - v\Delta t\sin(\theta), \hat{\mathbf{n}}, t - \Delta t)e^{-\Delta t/\tau_s}$$
$$+ \frac{1}{\tau_s}\int_{t-\Delta t}^{t}\sum_{\hat{\mathbf{n}}'}\frac{2\pi}{N}f(\hat{\mathbf{n}}, \hat{\mathbf{n}}')I(x - v(t - t')\cos(\theta), y - v(t - t')\sin(\theta), \hat{\mathbf{n}}', t')e^{-(t-t')/\tau_s}dt'. \tag{3.7}$$

We then discretize the remaining time integral using the two-point trapezoidal quadrature rule, which makes our numerical scheme energy conserving, to obtain

$$I(x, y, \hat{\mathbf{n}}, t) = I(x - v\Delta t\cos(\theta), y - v\Delta t\sin(\theta), \hat{\mathbf{n}}, t - \Delta t)e^{-\Delta t/\tau_s}$$
$$+ \frac{1}{\tau_s}\sum_{\hat{\mathbf{n}}'}\frac{\pi}{N}f(\hat{\mathbf{n}}, \hat{\mathbf{n}}')I(x - v\Delta t\cos(\theta), y - v\Delta t\sin(\theta), \hat{\mathbf{n}}', t - \Delta t)e^{-\Delta t/\tau_s}\Delta t$$
$$+ \frac{1}{\tau_s}\sum_{\hat{\mathbf{n}}'}\frac{\pi}{N}f(\hat{\mathbf{n}}, \hat{\mathbf{n}}')I(x, y, \hat{\mathbf{n}}', t)\Delta t. \tag{3.8}$$

For the discretization in equation 3.8 we require $\Delta t \ll \tau_s$ to capture the scattering interactions. Lastly, we arrange equation 3.8 to obtain

$$I(x, y, \hat{\mathbf{n}}, t) - \frac{1}{\tau_s}\sum_{\hat{\mathbf{n}}'}\frac{\pi}{N}f(\hat{\mathbf{n}}, \hat{\mathbf{n}}')I(x, y, \hat{\mathbf{n}}', t)\Delta t =$$
$$I(x - v\Delta t\cos(\theta), y - v\Delta t\sin(\theta), \hat{\mathbf{n}}, t - \Delta t)e^{-\Delta t/\tau_s} +$$
$$\frac{1}{\tau_s}\sum_{\hat{\mathbf{n}}'}\frac{\pi}{N}f(\hat{\mathbf{n}}, \hat{\mathbf{n}}')I(x - v\Delta t\cos(\theta), y - v\Delta t\sin(\theta), \hat{\mathbf{n}}', t - \Delta t)e^{-\Delta t/\tau_s}\Delta t, \tag{3.9}$$

where the right-hand side is a known quantity since it only depends on previous time $t - \Delta t$.

Equation 3.9 consists of a system of N equations, with N variables (the intensities at every discretization angle). For an arbitrary number of N directions we write

$$\mathbf{A}\vec{\mathbf{I}} = \vec{\mathbf{S}}. \tag{3.10}$$

The entries of the vector on the right-hand side of equation 3.10 correspond to the right-hand side of equation 3.9. For a given direction $\hat{\mathbf{n}}_i$ we write

$$S(x, y, \hat{\mathbf{n}}_i, t - \Delta t) = I(x - v\Delta t\cos\theta, y - v\Delta t\sin\theta, \hat{\mathbf{n}}_i, t - \Delta t)e^{-\Delta t/\tau_s}$$
$$+ \frac{1}{\tau_s}\sum_{\hat{\mathbf{n}}'}\frac{\pi}{N}f(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}')I(x - v\Delta t\cos\theta, y - v\Delta t\sin\theta, \hat{\mathbf{n}}', t - \Delta t)e^{-\Delta t/\tau_s}\Delta t. \tag{3.11}$$

The $N \times N$ matrix on the left-hand side of equation 3.10 has entries $A_{ij} = \delta_{ij} - \frac{\pi \Delta t}{N \tau_s} f(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_j)$. The indices are introduced to emphasize that the scattering function $f(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_j)$ depends on the incident and scattering angles. This matrix multiplies the intensity vector $\vec{\mathbf{I}}$ that we solve for. The right hand-side of equation 3.10 corresponds to a source term with which we evolve the numerical solution, and depends on the intensities at the previous time, given by equation 3.11.

Directly solving the linear system 3.10 is expensive. For a computational domain of size (N_x, N_y, N_t), one would need to perform $N_x * N_y * N_t$ linear solves. To speed up the computation we compute the inverse of the matrix $\boldsymbol{A}$ before-hand via LU decomposition with partial pivoting. We quantify the stability of inverting the matrix $\boldsymbol{A}$ in the linear system 3.10 with the condition number $\kappa(\boldsymbol{A}) = ||\boldsymbol{A}||\ ||\boldsymbol{A}^{-1}||$, where $||\cdot||$ is the 2-norm. When the condition number $\kappa(\boldsymbol{A}) = 1$, the matrix $\boldsymbol{A}$ is well-conditioned. For the numerical experiments in this chapter we have a condition number $\kappa(\boldsymbol{A}) = 1.02$. Notice also that we can write $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{K}$, with $K_{ij} = \frac{\pi \Delta t}{N \tau_s} f(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_j)$. With this decomposition we can perform the matrix expansion $(\boldsymbol{I} - \boldsymbol{K})^{-1} = \boldsymbol{I} + \sum_{n=1}^{\infty} \boldsymbol{K}^n$, which is valid for $||K|| < 1$. We can truncate the infinite sum to some upper bound M, depending on the value of $||K||$, which we can control with Δt. If the medium has non-uniform scattering properties, we can pre-compute the inverses of $\boldsymbol{A}$ for the different sections of the numerical model, and then use these inverses in the time-stepping algorithm. Once we have access to the inverse we can parallelize the algorithm at each time-step.

3.3.3 Choice of Scattering Function and the Transport Mean Free Time

We use the Henyey-Greenstein (HG) scattering function in 2D to describe a medium with angle-dependent scattering (Margerin et al., 2016)

$$f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') = \frac{1}{2\pi} \frac{1 - g^2}{(1 + g^2 - 2g(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}'))}. \tag{3.12}$$

The variable g describes the level of angle-dependent scattering and it ranges from predominantly backward scattering ($g = -1$) to isotropic scattering ($g = 0$) to predominantly forward scattering ($g = 1$). For practical purposes one considers the range $-1 < g < 1$ to avoid the singularities for $g = \pm 1$ when $(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}') = \pm 1$. When scattering depends on angle, the time scale at which the randomization of the propagation direction of the waves occurs is referred to as the transport mean free time τ^*, and it is related to the scattering mean free τ_s time by $\tau^* = \tau_s/(1 - g)$ (van Rossum and Nieuwenhuizen, 1999). This relationship implies that if scattering is predominantly in the forward direction $\tau^* > \tau_s$, whereas when scattering is predominantly in the backward direction $\tau^* < \tau_s$. Margerin et al.(2016) noted that if one sets $g = (1 + 2k^2a^2 - \sqrt{1 + 4k^2a^2})$, the HG scattering function can be considered as an end-member of the more

general Von-Karman scattering phase function

$$f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') = \frac{N(ka, \kappa)}{(1 + 2k^2a^2(1 - \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}'))^{\kappa+1}}, \tag{3.13}$$

in the limit $\kappa \to 0$. In equation 3.13, k is the wavenumber, a is the correlation distance of the medium, κ is a free parameter, and $N(ka, \kappa)$ denotes a normalization factor so that the condition 3.2 is satisfied. As the frequency of the waves increase, the magnitude of g increases as the waves become more sensitive to the inhomogeneities in the medium (Chapter 4 , Sato et al., 2012).

We normalize the scattering function using a discretized version of equation 3.2

$$\sum_{\hat{\mathbf{n}}'} \frac{1}{CN} \frac{1 - g^2}{(1 + g^2 - 2g(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}'))} = 1, \tag{3.14}$$

where the discrete sum is over N angles. C thus corresponds to a constant

$$C = \sum_{\hat{\mathbf{n}}'} \frac{1}{N} \frac{1 - g^2}{(1 + g^2 - 2g(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}'))}, \tag{3.15}$$

which we introduce to satisfy the unity of the normalization condition 3.2 for the discretized scattering function. When $g = 0$ the constant $C = 1$, and when $g \neq 0$ the value of the constant $C \to 1$ as $N \to \infty$.

3.4 Numerical Simulations

3.4.1 Comparison of numerical and exact solutions for isotropic scattering

The initial condition for the intensity at $t_0 = 0$ consists of a 2-D Gaussian placed at the middle of the computational domain with a full width at half maximum of ≈ 2350 m, where at every point in space energy is emitted equally in all directions. Figure 3.1 shows this initial condition as well as the spatial coordinates of the computational domain. Both coordinates range from -20000 m to 20000 m, with a grid spacing $\Delta x = \Delta y = 200$ m. We use a constant propagation velocity $v = 1000$ m/s and scattering mean free time $\tau_s = 5$ s. We integrate the intensities from $t = 0$ to 20 s with $\Delta t = 0.2$ s, with the choice of temporal discretization small enough ($\Delta t \ll \tau_s$) to capture the scattering interactions. We use $N = 128$ directions of propagation, equally spaced in the range $[0, 2\pi)$. To compute the advection terms which do not lie on grid points we perform a numerical interpolation. Different interpolation techniques may be used depending on the accuracy and speed of computation that one wants to achieve. Here, for simplicity, we use a bilinear interpolation with neighboring points, which is a sufficiently good approximation provided that the initial condition is smooth. In our numerical experiments we find that ten sample points for the full-width at half-maximum are enough to avoid interpolation artifacts. As with finite differences, the modeling of true point sources is not possible with our algorithm. However, one could in practice set-up the initial condition to be a 2-D Gaussian with a small width, relative to the length scale over which scattering occurs, and use a very dense computational grid. At the boundary we set the intensities equal to zero, which does not

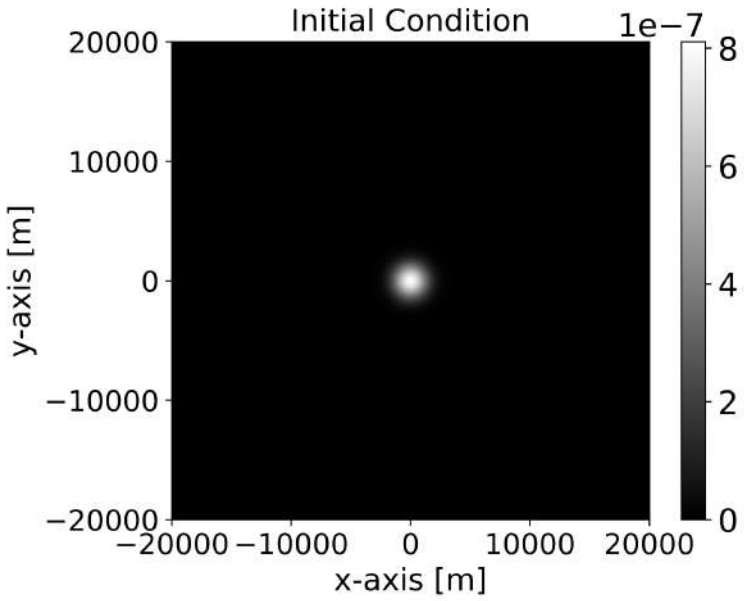

Figure 3.1 Total intensity at $t = 0$. This initial condition is normalized such that the total energy equals unity.

affect the accuracy of our results for the distances and propagation times that we investigate. After the energy reaches the boundary, some leakage occurs over time. We compare the analytical and numerical

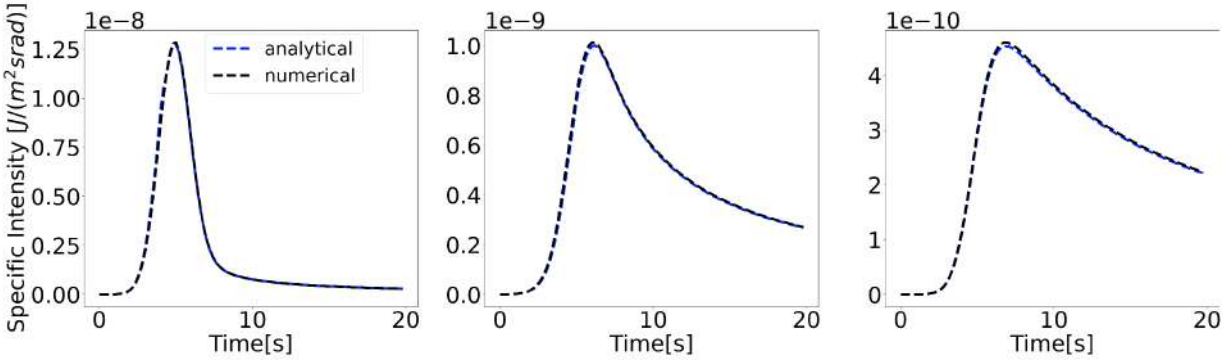

Figure 3.2 Comparison of analytical and numerical specific intensities at a distance of 1 scattering mean free length (5000 m) for isotropic scattering ($g = 0$). The three pannels from left to right show the specific intensities at an angle of $0, \pi/4, \pi/2$ relative to the radial direction, respectively.

specific intensities, rather than the total intensities, for a medium with isotropic scattering. This comparison allows us to assess the accuracy of our algorithm in retrieving the individual specific intensities since the total intensity may average out discrepancies between the numerical and analytical solutions of the specific intensities. The analytical intensity corresponds to the solution derived by Paasschens (1997) convolved with the initial condition shown in Figure 3.1

Figure 3.2 shows the comparison between analytical and numerical solutions for the specific intensities at different angular directions at a distance of 5000 m from the Gaussian peak in Figure 3.1. The direction aligned with the radial direction shows the largest amplitude, and the magnitude of the specific intensity decreases with increasing angle θ relative to the radial direction. The radial direction contains both incident and scattered energy, whereas off-radial directions contain only scattered energy. The scattered energy decreases in amplitude with increasing angle relative to the radial direction, as shown by Paasschens (1997). Our numerical solutions in Figure 3.2 agree within 3% with the exact solutions. The fact that we match the exact solution well indicates that our algorithm is able to model the specific intensities, rather than just the total intensity. Therefore, the algorithm that we are showing in this chapter is applicable for problems in which one wants to model the directionality of wave transport.

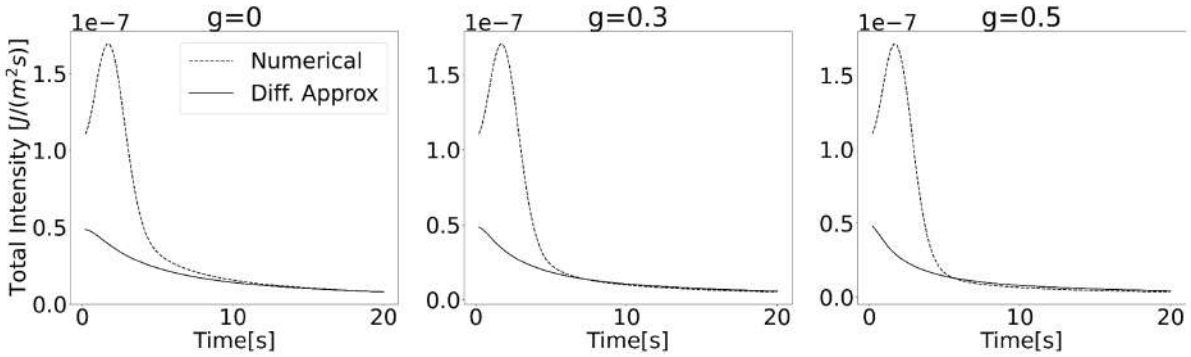

Figure 3.3 Comparison of numerical solutions for the total intensity with diffusive approximation at $(x, y) = (2000, 0)$ m, for different values of g, ranging from isotropic to predominantly forward scattering.

3.4.2 Comparison of numerical and diffusive solutions

To benchmark the long-time behavior of the algorithm we compare the numerical solutions that we obtain with diffusive approximations to RTE, which are valid after sufficient scattering has occurred. We use the 2-D version of the diffusion approximation to the scalar RTE (chapter 7, Sato et al., 2012).

Figure 3.3 shows the comparison of the numerical solutions for the Gaussian simulation for different values of g against the diffusive approximation at $(x, y) = (2000, 0)$ m. At $t = 0$ there is non-zero energy for the chosen location due to the source that we use. Notice that as the value of g increases, the coda decays more rapidly due to the preferred forward scattering of the outward propagating energy. As the value of g increases it takes a longer time for the scattered energy to reach a diffusive state. This is because as g increases so does the transport mean free time, the time scale over which the randomization of the direction of the waves occur. At late times the numerical solutions agree with the diffusive approximation.

3.4.3 Isotropic scattering

As waves propagate through scattering media their wave propagation regime changes from ballistic to single scattering to multiple scattering to diffusion to equipartitioning. To study the spatial dependence of the propagation regime we consider a numerical experiment with a Gaussian initial condition at the center of the computational domain, as in Figure 3.1, and isotropic scattering. The simulation parameters are the same as for Figure 3.2. Figure 3.4 shows cross-sections of the total intensity at $y = 0$ for different

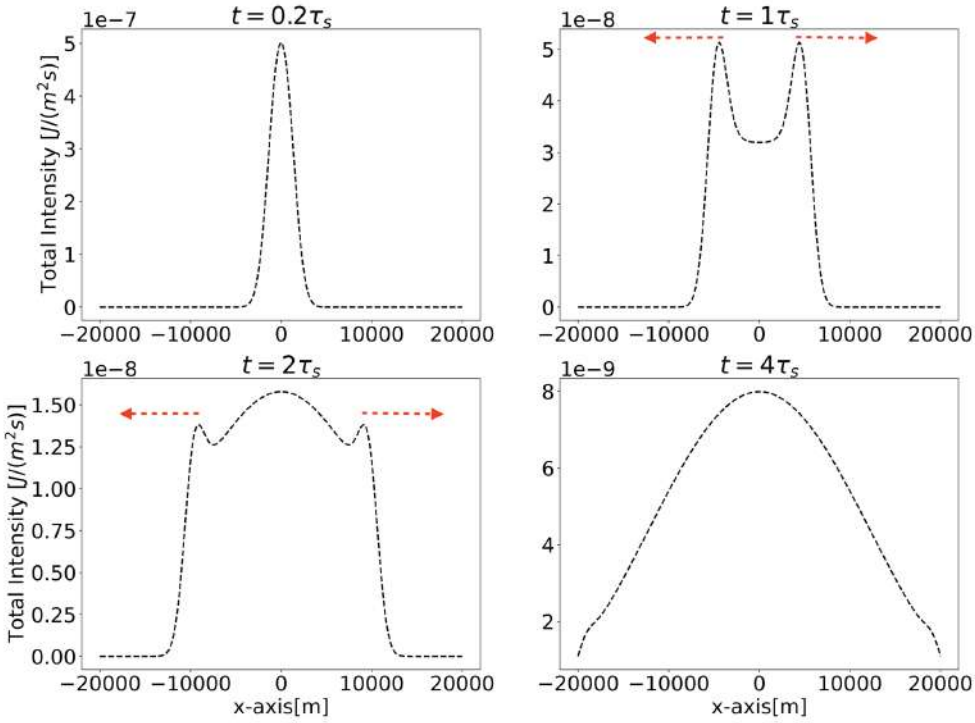

Figure 3.4 Cross-section of the total intensities for isotropic scattering. The red arrows indicate the outgoing ballistic energy.

simulation times. At early times ($t \ll \tau_s$) the energy is still concentrated near the source vicinity and little scattering has occurred. At later times ($t = \tau_s$) there is outgoing ballistic energy, with scattered energy in-between the outgoing wavefronts. This energy is generated by single and multiple scattering. At a later time of $t = 2\tau_s$ the scattered energy starts to dominate over the ballistic energy. This corresponds to a transfer of energy from incident to scattered waves. The scattering energy now contains single and multiple scattering energy, part of which has started to develop a smooth Gaussian-like shape in between the ballistic peaks. This smoothing indicates the beginning of the on-set of the diffusive behavior, as one expects from the solution to the diffusion equation (Sato et al., 2012). In this diffusive regime, the propagation direction of the waves starts to randomize over time. However, the rate at which the

randomization occurs changes with space and time. On the panel in the lower right the scattered energy dominates, there is nearly no ballistic energy left, and the diffusive propagation starts to dominate. The rapid drop at the edges of the computational domain occurs because of the vanishing boundary condition. Figure 3.5 shows a comparison of the cross-sections from fig. Figure 3.4 with the diffusive approximation.

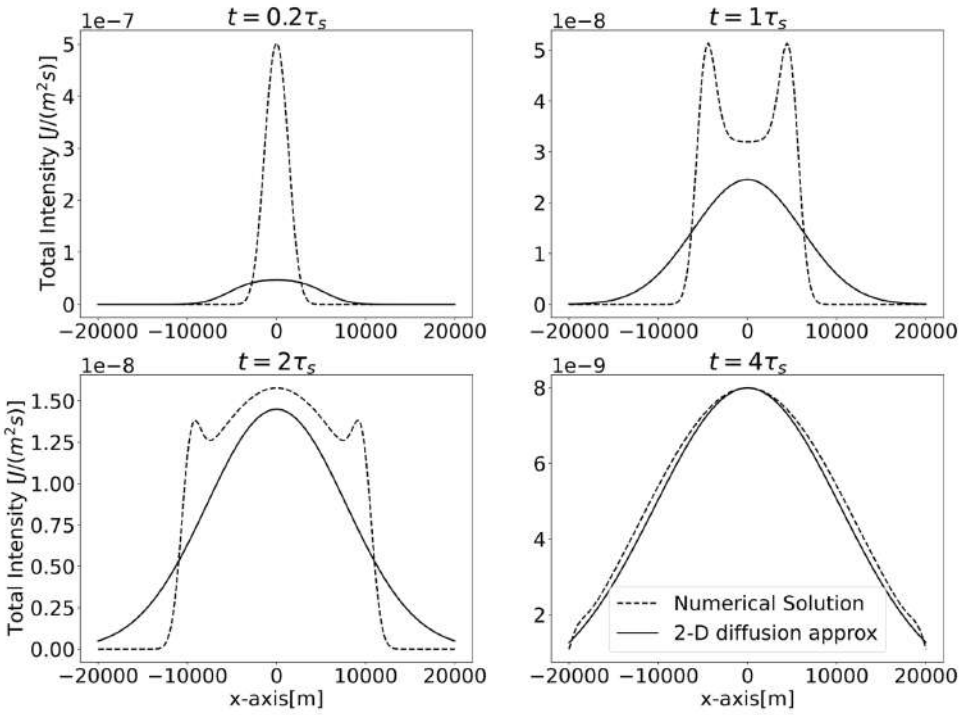

Figure 3.5 Comparison of cross-sections in Figure 3.4 with the diffusive approximation.

Only after two scattering mean free times, the diffusive behavior starts to develop. At later times ($t = 4\tau_s$), the diffusive behavior is evident. The energy first becomes diffusive near the source, where more scattering has occurred, relative to the ballistic wave. Far away from the source, and close to the ballistic peak, the energy does not become diffuse yet. Over time, the net flux of energy decreases and scattering homogenizes the angular distribution of the energy.

At late times, after sufficiently many scattering events, one assumes first that the wave field reaches a diffusive state (i.e., the net flux of energy is small away from the source) and follows a random-walk like behavior, and then that it becomes equipartitioned (i.e., the energy is the same in all propagation directions and the net flux of energy is zero). However, one usually assumes that diffusion, and then equipartitioning, occurs when the condition $t \gg \tau^*$ is satisfied, rather than the condition $t \gg r/v + \tau^*$. In this expression r is the source-receiver distance, v is the speed of transport, and τ^* is the transport mean free path, which for isotropic scattering equals the scattering mean free time τ_s. The second condition is

stricter because it accounts for the scattering events after the ballistic arrival. For practical applications,

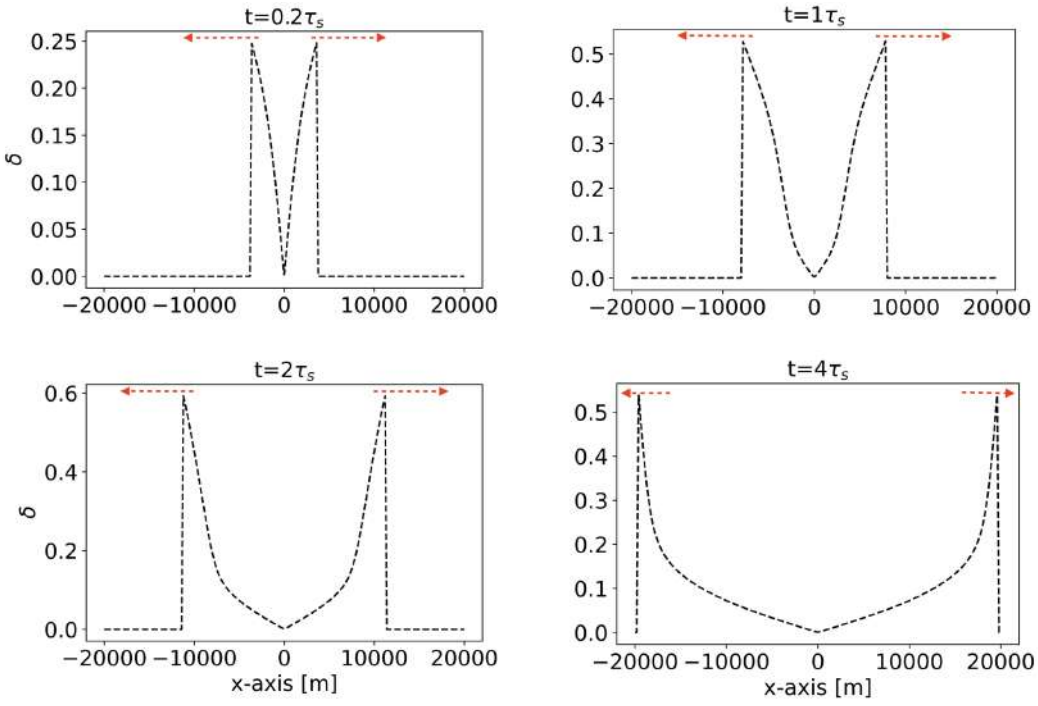

Figure 3.6 Cross-section of equipartitioning index δ for the isotropic source experiment. The red arrows point to outgoing ballistic energy.

one usually assumes first that the wave field is diffusive, and then equipartitioned, after enough scattering events (Campillo, 2006; Obermann et al., 2013). To quanitify the degree of equipartitioning we introduce the *equipartitioning index*

$$\delta(\mathbf{r},t) = \frac{\sigma(\mathbf{r},t)}{\sqrt{N}\mu(\mathbf{r},t)}, \tag{3.16}$$

where $\sigma(\mathbf{r},t)$ is the standard deviation of the specific intensities along the angular directions. This quantity is defined as $\sigma^2(\mathbf{r},t) = \frac{1}{N-1}\sum_{i=1}^{N}(I(\mathbf{r},t,\hat{\mathbf{n}}_i) - \mu(\mathbf{r},t))^2$, where $\mu(\mathbf{r},t) = \frac{1}{N}\sum_{i=1}^{N} I(\mathbf{r},t,\hat{\mathbf{n}}_i)$ is the mean of the specific intensities along the angular direction, and N is the number of angular directions (Snieder and van Wijk, 2015). The quantity $\delta(\mathbf{r},t)$ provides a measure of the variation of the specific intensities along the angular directions as a function of space and time. We compute the standard deviation and the mean of the specific intensities along the N propagation directions for a fixed point in space and time. The constant $1/\sqrt{N}$ is included so that $0 \leq \delta \leq 1$. Consider an uni-directional intensity field $I_i = I_0\delta_{i,1}$, where I_0 is the intensity along the only non-zero direction and $\delta_{i,j}$ is the kronecher delta. The mean for this intensity field is $\mu = I_0/N$. The standard deviation for the same intensity field is $\sigma = I_0/\sqrt{N}$, which gives $\sigma/\mu = \sqrt{N}$. The equipartitioning index is 0 when the intensity field is fully equipartitioned (i.e., same specific intensity

along all directions) and 1 when the intensity field is unidirectional (i.e., all of the specific intensities but one are equal to zero).

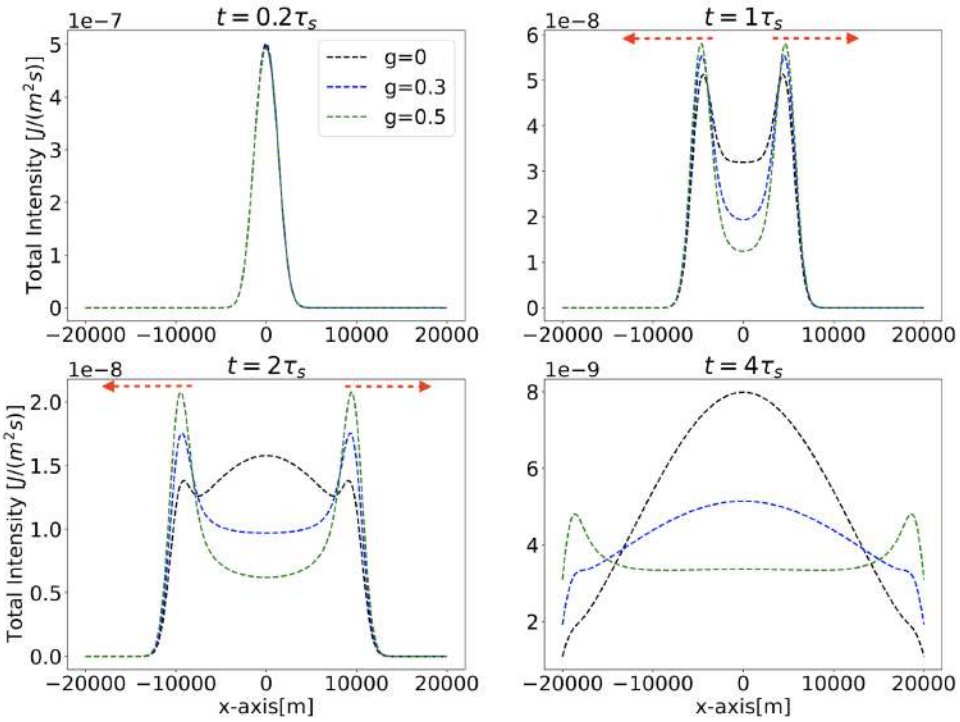

Figure 3.7 Cross-section of the total intensities for different levels of angle-dependent scattering and simulation times. The black, blue, and green curves correspond to $\tau^* = 1, 1.43, 2\tau_s$, respectively. The red arrows indicate the outgoing ballistic energy.

In the first numerical experiment we release the energy within the Gaussian equally along all directions, meaning that $\delta = 0$ everywhere at $t = 0$. However, due to the nonzero gradient of the intensity field part of the intensity propagates away from the initial Gaussian with a preferred propagation direction and $\delta \neq 0$ due to the directionality of the energy transport. Figure 3.6 shows δ for the same cross-section as Figure 3.4. The ballistic arrival shows the highest value of δ. For this simulation the ballistic peak shows $\delta < 1$ due to the spread of the source and the isotropic initial condition. At very early times ($t = 0.2\tau_s$), $\delta \neq 0$ despite the initial condition being isotropic, as explained above. The scattering energy in between the ballistic peaks starts to homogenize in the angular direction over time, with the lowest δ value at $x = 0$. Even at late times the value of δ is only small (< 0.1) between -10000 and 10000 m, where more scattering events have occurred relative to those events near the ballistic arrival. One should use the condition $t \gg r/v + \tau^*$ for assessing equipartitioning to account for multiple scattering after the ballistic arrival. To see why this condition is necessary, consider the two terms at the right hand side of the inequality. The first term r/v accounts for the travel time of the ballistic wave from the source to receiver locations. The

second term, τ^*, is the transport mean free path, which is a measure of the time scale over which the direction of propagation of the energy is randomized. Both of these terms together describe the time necessary for multiple scattering, and therefore diffusion and equipartitioning, to develop after the ballistic wave arrival. In particular, one can not generally say that the field is diffuse/equipartitioned because the level of diffusion/equipartitioning depends on location and time.

3.4.4 Angle-dependent scattering

For realistic applications of the scalar RTE, it is necessary to consider the effect of angle-dependent scattering. To this end, we study the spatial dependence of wave propagation when the scattering function depends on angular direction. Here, we investigate a medium in which forward scattering dominates. We initialize the intensities and set-up the computational domain as in the isotropic simulation to provide a just comparison between the numerical experiments.

Figure 3.7 shows a cross-section along the x-axis at $y = 0$ of the total intensity for different levels of angle-dependent scattering. The scattering function is given by equation 3.12 with $g = 0, 0.3, 0.5$ for isotropic, weak, and medium forward scattering, respectively. Overall, this numerical simulation shows the same behavior as for the isotropic case. However, as forward scattering becomes stronger, the ballistic peak retains more energy and there is less energy scattered to other directions, as compared to isotropic scattering. As forward scattering increases it takes a longer time for the propagation of the waves to be randomized. To see why this is the case consider a unidirectional intensity source. As forward scattering becomes stronger, the energy that scatters off the ballistic wave takes longer to randomize over the angular direction due to the preferred propagation direction and the energy in-between the outgoing ballistic peaks takes a longer time to become diffusive. Figure 3.8 illustrates the on-set of the diffusive behavior for different values of g. As forward scattering increases so does the transport mean free path, the length scale over which the randomization of the propagation direction occurs.

For these simulations we find agreement with two physical principles. The first one is that with increasing forward scattering the ballistic wave develops a pronounced peak along the forward direction and the scattered energy behind the ballistic peak decreases. The second one is that as forward scattering increases, the intensity field takes a longer time to become diffusive, and then equipartitioned, as it takes longer for the propagation direction to be randomized. As with the isotropic case there is a spatial dependence of the transition from ballistic to diffuse to equipartitioned waves. To explore the degree of equipartitioning when scattering depends on angle we compute the equipartitioning index δ for the same cross-section as in Figure 3.6. Figure 3.9 shows the equipartitioning index δ for this cross-section. Overall, the behavior is the same as for isotropic scattering. As before, the ballistic arrival shows the highest

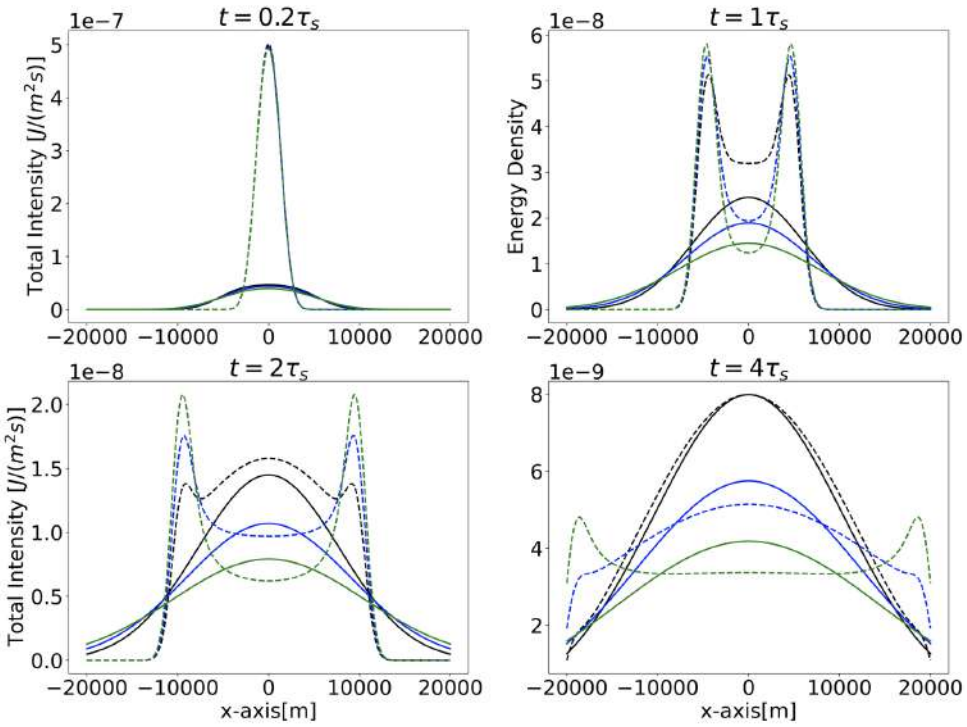

Figure 3.8 Comparison of cross-sections in fig. Figure 3.7 with the diffusive approximation. The solid line indicates numerical simulation, and the dashed line indicates diffusive approximation.

equipartitioning index δ, and the lowest δ is at $x = 0$, at the center of the computational domain. For all degrees of angle-dependent scattering the equipartition index is nearly the same. This occurs because the field starts with $\delta = 0$ at $t = 0$ so that different values of g do not significantly affect the value of the equipartitioning index δ in-between the outgoing ballistic energy.

Consider the bottom left panel in Figure 3.9. Between -5000 and 5000 m, $\delta < 0.1$. This implies that for these distances the intensity fields is close to having an isotropic angular distribution (i.e., δ is close to zero). The bottom right panel of Figure 3.8 shows that, for this same range, only the curve with $g = 0$ starts to show a diffusive behavior. This comparison shows that even if a field has $\delta << 1$ it is not necessarily diffuse because an extra condition related to the net flux of energy must be satisfied.

van Rossum and Nieuwenhuizen (1999) derive a diffusion equation from the 3-D version of equation 3.1. To this end they make two assumptions. The first assumption is that the intensity distribution is almost isotropic ($\delta << 1$), so that the energy current $\mathbf{J}$ is much smaller than the energy density I^{tot} (pg. 176, Chapter 9, Ishimaru, 1978). With this assumption they expand the specific intensities as

$$I(\mathbf{r}, \hat{\mathbf{n}}, t) \approx I^{\text{tot}}(\mathbf{r}, t) + 3v\hat{\mathbf{n}} \cdot \mathbf{J}(\mathbf{r}, t), \tag{3.17}$$

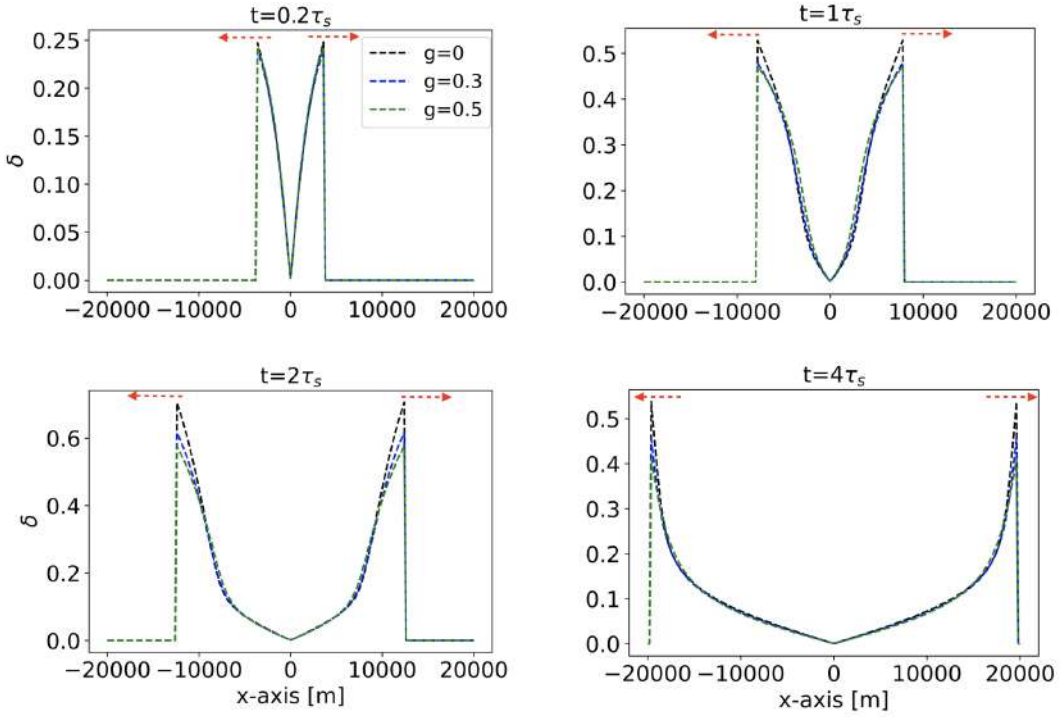

Figure 3.9 Cross-section of equipartitioning index δ for the isotropic source experiment. The red arrows indicate outgoing ballistic energy.

where they ignore the higher order terms since these terms become negligible. The variables I^{tot} and $\mathbf{J}$ are the local radiation and current densities, respectively. They then assume that the local current density of the intensity field satisfies

$$\left|\tau^* \partial_t \mathbf{J}\right| \ll \left|\mathbf{J}\right|, \tag{3.18}$$

with which they obtain the diffusive approximation to the scalar RTE. Expression 3.18 indicates that the current density of the field varies slowly in time, relative to the time scale over which the randomization of the direction of propagation of the wave occurs. To numerically study this diffusive condition we define

$$\epsilon = |\tau^* \partial_t \mathbf{J}/\mathbf{J}|. \tag{3.19}$$

Recall that in Figure 3.8 the diffuse character of the scattered waves only starts to develop after $t = 2\tau_s$. Figure 3.10 shows expression 3.19 for the bottom left and right panels, respectively, of Figure 3.8, with the x-axis now ranging from -6000 to 6000 m to focus on the diffuse character of the scattered waves. The left panel of Figure 3.10 shows the lowest value of ϵ for isotropic scattering, and this value increases as forward scattering dominates. This reflects that the energy that scatters isotropically reaches a diffuse state more rapidly than when forward scattering dominates, as we showed in Figure 3.8.

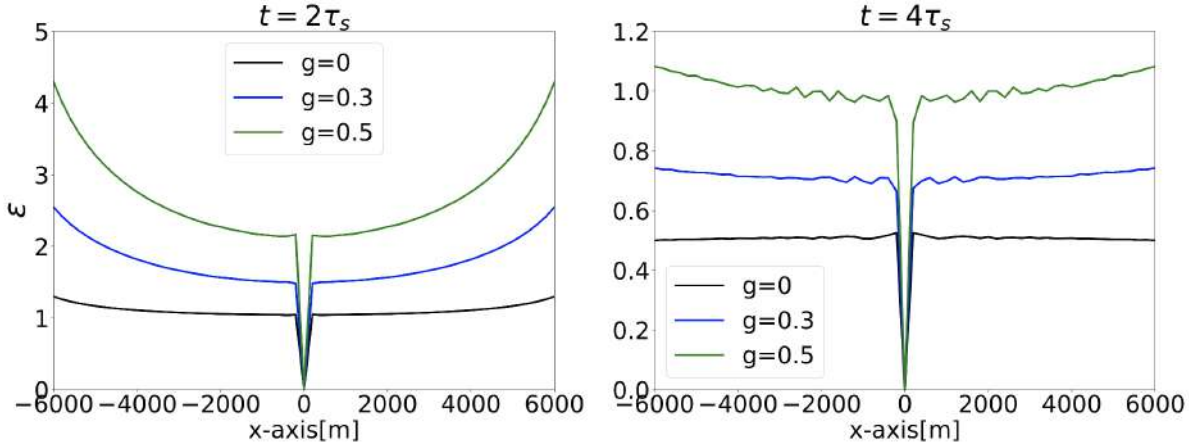

Figure 3.10 Diffusion condition, as measured by ϵ defined in expression 3.19, for the bottom left and bottom right panels in fig. Figure 3.7, with the x-axis now ranging from -6000 to 6000 m.

$\delta = 1$	$\longleftrightarrow$	Ballistic
$0 < \delta < 1, \epsilon \gg 0$	$\longleftrightarrow$	single/multiple scattering
$\delta \ll 1, \epsilon \ll 1$	$\longleftrightarrow$	multiple scattering/diffusion
$\delta = 0$	$\longleftrightarrow$	equipartitioning

Table 3.1 Wave propagation regimes. The transition over time among the different regimes starts with the ballistic and ends with the equipartitioned regime.

The left panel of Figure 3.10 also shows that, as with δ, ϵ changes locally. The smallest value of ϵ is near $x = 0$, where many scattering events have occurred, and increases away from this location. This local behavior is more evident with stronger forward scattering ($g = 0.5$). The right panel of Figure 3.10 also shows higher values of ϵ for stronger forward scattering. It also shows that ϵ decreases with time because the scattered waves start to become diffuse. As before, this diffuse behavior changes locally, but with less variations than for earlier times.

With the introduction of δ and the definition for ϵ we can identify 4 propagation regimes, depending on the extent of scattering, which we illustrate in Table 3.1. A ballistic wave, due to its directionality, has $\delta = 1$. As scattering occurs, the energy behind the ballistic peak (i.e., singly and multiply scattered waves) starts to distribute over the different propagation directions ($\delta < 1$). At this stage the net flux of energy is not necessarily small. Over time, the angular distribution of the intensity becomes almost isotropic ($\delta \ll 1$). If the field satisfies condition 3.18, the energy reaches a diffusive state. At even later times, the angular distribution of the intensity becomes isotropic as the net flux of energy goes to 0. To see this,

consider δ for expression 3.17

$$\delta(\mathbf{r},t) = \frac{\frac{1}{N-1}\sum_{i=1}^{N}\left(3v\hat{\mathbf{n}}_i \cdot \mathbf{J}(\mathbf{r},t)\right)^2}{I^{\text{tot}}(\mathbf{r},t)}. \tag{3.20}$$

Since each term on the sum in expression 3.20 is positive, it follows that for equipartitioning to happen ($\delta = 0$), each term in the sum must be equal to zero (i.e., net flux of energy equals zero).

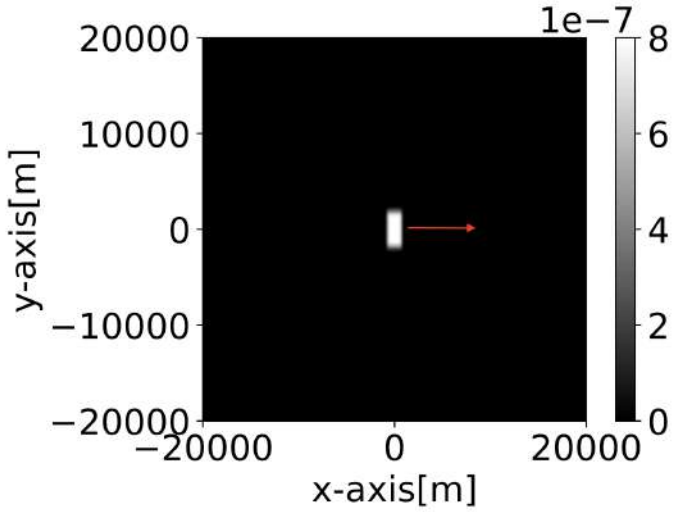

Figure 3.11 Normalized initial condition for the specific intensity of a plane-wave segment aligned with the positive x-axis, all other specific intensities are set equal to zero. The red arrow indicates the initial direction of wave propagation.

3.4.5 Directional source and angle-dependent scattering

So far we have explored the spatial dependence on wave propagation due to the scattering properties of the medium. We extend our numerical study by considering a segment of a plane-wave which propagates from left to right. As before, we test for isotropic, weak forward, and medium forward scattering ($g = 0.0, 0.3, 0.5$), which correspond to transport mean free times $\tau^* = 1, 1.43, 2\tau_s$, respectively. Figure 3.11 shows the initial condition and computational domain for this numerical test. We place the plane wave segment around the origin with the x,y-coordinate of the computational domain ranging from -20000 to 20000 m. The only non-zero specific intensity at time $t = 0$ corresponds to $\theta = 0$. We taper the plane-wave segment using the product of a horizontal and a vertical cosine taper. We use the following cosine taper

$$T(n,\alpha) = \left\{ \begin{array}{ll} 1 & , \quad \text{if } \alpha/2 \leq n \leq 1-\alpha/2, \\ 0.5(1-\cos(2\pi n/\alpha)), & \text{else,} \end{array} \right\} \tag{3.21}$$

where $0 \leq \alpha, n \leq 1$. α is a smoothing parameter which be set to 0.5, and n is a normalized spatial coordinate defined as $n = (x + 1000)/2000$ for the horizontal window and $n = (y + 2500)/5000$ for the

vertical window. As with the previous numerical simulations we scale the initial condition so that the total energy is normalized. Figure 3.12 shows a cross-section at the middle of the computational domain along

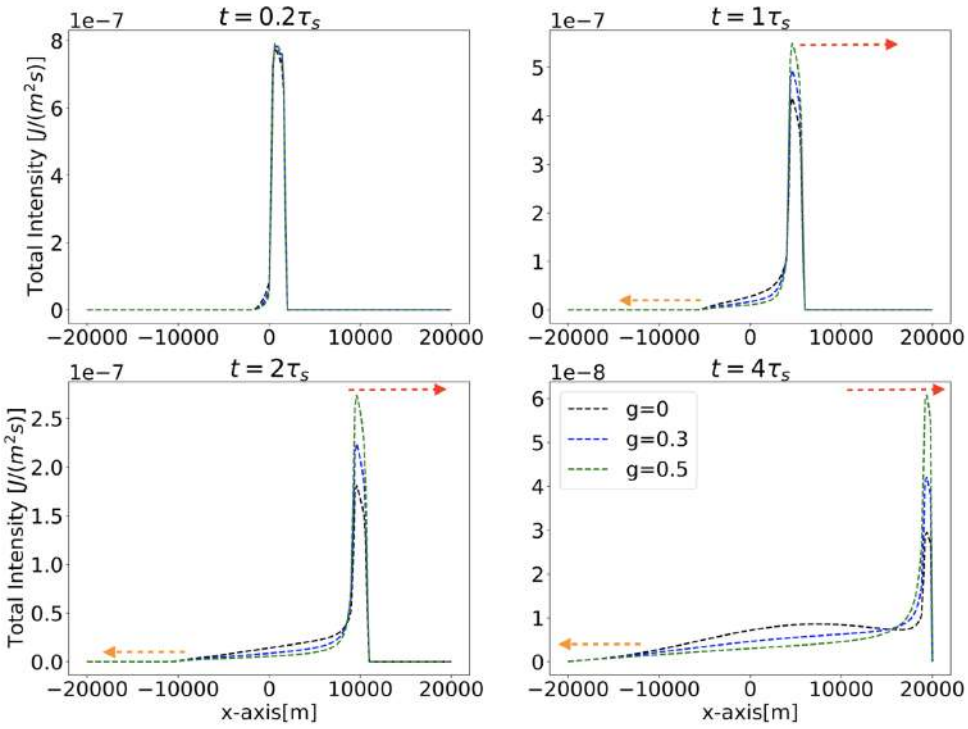

Figure 3.12 Cross-section of the total intensities for different levels of angle-dependent scattering and simulation times. The black, blue, and green curves correspond to $\tau^* = 1, 1.43, 2\tau_s$, respectively. The red arrow indicates the right-going ballistic wave, the orange arrow indicates a secondary "ballistic" arrival which arises due to backscattering.

the x-axis for the plane wave segment simulation. As with the isotropic initial condition, when time progresses the intensity with stronger forward scattering has more energy concentrated along the ballistic arrival and less trailing scattered energy. A secondary "ballistic" wave, caused by waves that are backscattered immediately after the plane wave segment is launched, is indicated by the orange arrow and propagates towards the left of the computational domain. This secondary arrival has a negligible amplitude. Figure 3.13 shows the equipartitioning index δ for the intensity cross-section in Figure 3.12. The ballistic arrival shows the highest equipartitioning index δ, as in the previous simulations. Ideally, the ballistic wave should show $\delta = 1$. However, because of the numerical discretization the right-going peak contains small amounts of scattered energy, which results in a lower δ. The secondary "ballistic" arirval that arises early in the simulation shows up as a secondary peak on the left side of Figure 3.13, but the intensity corresponding to this peak is negligible. The scattering energy directly behind the peak propagating to the right shows the lowest δ. The directionality in wave propagation occurs because the

initial condition has a preferred propagation direction. For this type of initial condition one should also use the condition $t \gg r/v + \tau^*$ to account for both the primary ballistic arrival (propagating from left to right) and the secondary wave (propagating to the left). Figure 3.13 also shows scattered energy with higher δ as forward scattering increases, which is most prominent at later times. This occurs because for this numerical test the initial condition has a preferred propagation direction so that g plays a key role in determining the angular randomization of the wave field.

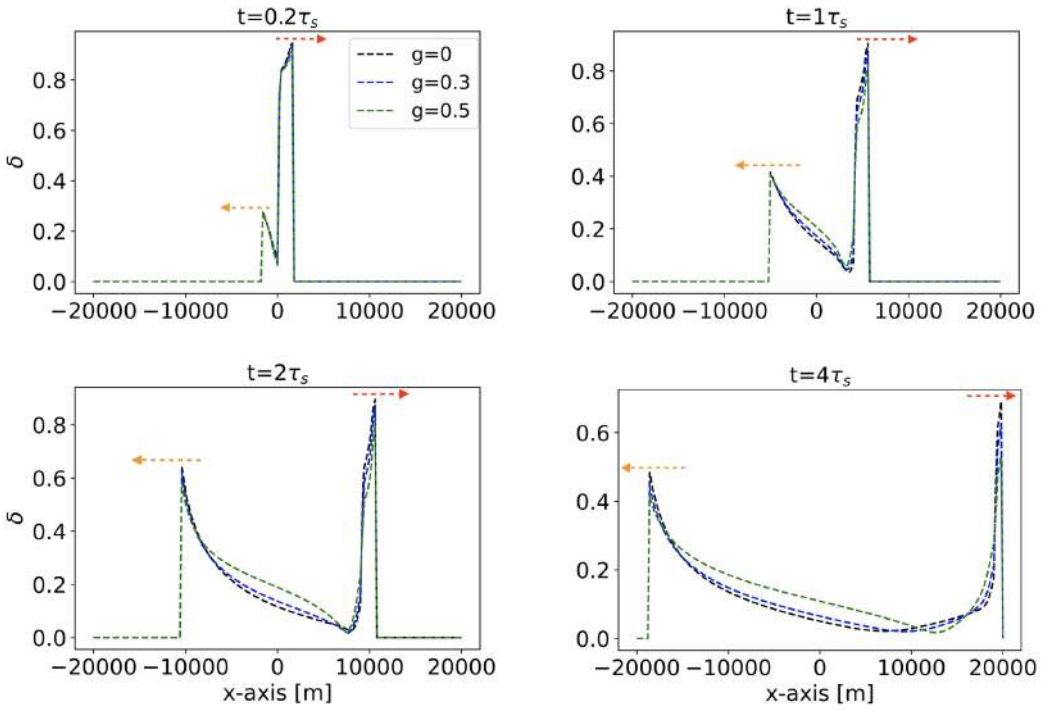

Figure 3.13 Cross-section of equipartitioning index δ for the plane source experiment. The red arrow indicates the right-going ballistic wave, the orange arrow indicates a secondary "ballistic" arrival which arises due to backscattering.

3.5 Discussion

The numerical algorithm that we propose to solve the scalar RTE in a medium with uniform speed of transport, uniform scattering mean free time, and arbitrary angle-dependent scattering allows us to compute the specific intensities, as opposed to just the total intensities. This algorithm can be used in applications where one needs to resolve the angular distribution of the wave energy transport. We can use our algorithm for an intensity with a sufficiently smooth initial condition (i.e., enough sampling points in space so that the interpolation does not introduce numerical artifacts), as we showed through numerical experiments with isotropic and plane wave segment initial conditions in a medium with angle-dependent scattering (Figure 3.7 and Figure 3.12). In our algorithm we discretize the angular integral and then

handle the advection of energy analytically at grid points, and through numerical interpolation between grid points. This approach allows us to reduce the numerical dispersion errors, compared to numerical techniques such as finite differences or discontinuous Galerkin finite element methods. This numerical dispersion may cause negative intensities which are non-physical. Since we compute the specific intensities, our algorithm is useful for testing and understanding the transition from ballistic to single/multiply scattered to diffuse to equipartitioned waves. We quantify the extent of randomization of an intensity field through the equipartitioning index δ. This equipartitioning index, together with the quantity ϵ, allows us to identify 4 propagation regimes, ranging from ballistic to single/multiply scattered waves to multiply scattered/diffuse waves to equipartitioned waves. We find that, in general, the level of randomization of the angular distribution of an intensity field and the net flux of energy is a function of space and time. Figure 3.9 and Figure 3.13 show that the ballistic arrival has the highest equipartitioning index and that the trailing scattered energy randomizes its propagation direction over time but not at the same rate throughout the computational domain. The difference in the rate at which the randomization occurs requires one to use the condition $t \gg r/v + \tau^*$ as a measure of randomization to account for the scattering events following the ballistic arrival. This condition indicates that the wave propagation regime varies spatially. Comparison of Figure 3.9 and Figure 3.13 shows that for assessing the randomization of the intensity field one also needs to consider, to some extent, the initial condition of the intensity field. In the plane wave segment simulation the energy directly behind the right-going peak quickly randomizes, but as one moves closer to the secondary "ballistic" wave (left-going) the wave field has a preferred propagation direction due to backscattering early in the simulation. This again indicates that the wave propagation regime depends on space. The spatial dependence of these different propagation regimes, and in particular the local behavior of equipartitioning, has implications for Green's functions retrieval. One particular example is *seismic interferometry* where the extent of the angular randomization of the wave field affects the accuracy of the Green's functions that one recovers. Since equipartitioning changes locally rather than globally, the accuracy of the Green's function reconstruction may depend on space and time.

3.6 Acknowledgements

This work is supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at the Colorado School of Mines.

3.7 RTE as a system of coupled integral equations

We show that the integral equation 3.5 satisfies the radiative transfer equations 3.3 with the initial condition 3.4. Inserting $t = t_0$ in equation 3.5 gives the initial condition 3.4. To simplify the proof we align

the x-axis with $\hat{\mathbf{n}}$ so that the integral equation 3.5 becomes

$$\begin{aligned} I(x, y=0, \hat{\mathbf{n}}, t) &= I(x - v(t-t_0), y=0, \hat{\mathbf{n}}, t_0) e^{-(t-t_0)/\tau_s} \\ &+ \frac{1}{\tau_s} \int_{t_0}^{t} \oint f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I(x - v(t-t'), y=0, \hat{\mathbf{n}}', t') e^{-(t-t')/\tau_s} d^2\hat{\mathbf{n}}' dt', \end{aligned} \tag{3.22}$$

and the radiative transfer equations 3.3 become

$$\partial_t I(x, y=0, \hat{\mathbf{n}}, t) + v\partial_x I(x, y=0, \hat{\mathbf{n}}, t) = -I(x, y=0, \hat{\mathbf{n}}, t)/\tau_s + \oint f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I(x, y=0, \hat{\mathbf{n}}', t) d^2\hat{\mathbf{n}}'/\tau_s, \tag{3.23}$$

for $t > t_0$. Next, we take the derivative of equation 3.22 with respect to t

$$\begin{aligned} \partial_t I(x, y=0, \hat{\mathbf{n}}, t) = &- v\partial_x I\big(x - v(t-t_0), y=0, \hat{\mathbf{n}}, t_0\big) e^{-(t-t_0)/\tau_s} \\ &- \frac{1}{\tau_s} I\big(x - v(t-t_0), y=0, \hat{\mathbf{n}}, t_0\big) e^{-(t-t_0)/\tau_s} \\ &+ \frac{1}{\tau_s} \oint f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I\big(x, y=0, \hat{\mathbf{n}}', t\big) d^2\hat{\mathbf{n}}' \\ &\frac{1}{\tau_s} \int_{t_0}^{t} \oint f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') \Big(- v\partial_x I\big(x - v(t-t'), y=0, \hat{\mathbf{n}}', t'\big) \Big) e^{-(t-t')/\tau_s} d^2\hat{\mathbf{n}}' dt' \\ &- \frac{1}{\tau_s^2} \int_{t_0}^{t} \oint f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I\big(x - v(t-t'), y=0, \hat{\mathbf{n}}', t'\big) e^{-(t-t')/\tau_s} d^2\hat{\mathbf{n}}' dt', \end{aligned} \tag{3.24}$$

and the derivative with respect to x

$$\begin{aligned} \partial_x I(x, y, \hat{\mathbf{n}}, t) &= \partial_x I\big(x - v(t-t_0), y=0, \hat{\mathbf{n}}, t_0\big) e^{-(t-t_0)/\tau_s} \\ &+ \frac{1}{\tau_s} \int_{t_0}^{t} \oint f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') \partial_x I\big(x - v(t-t'), y=0, \hat{\mathbf{n}}', t'\big) e^{-(t-t')/\tau_s} d^2\hat{\mathbf{n}}' dt'. \end{aligned} \tag{3.25}$$

Inserting equations 3.24 and 3.25 into the radiative transfer equations 3.23 gives

$$\begin{aligned} \frac{\partial I(x, y=0, \hat{\mathbf{n}}, t)}{\partial t} + v\partial_x I(x, y=0, \hat{\mathbf{n}}, t) = &- \frac{1}{\tau_s} \Bigg(I(x - v(t-t_0), y=0, \hat{\mathbf{n}}, t_0) e^{-(t-t_0)/\tau_s} \\ &+ \frac{1}{\tau_s} \int_{t_0}^{t} \oint f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I(x - v(t-t'), y=0, \hat{\mathbf{n}}', t') e^{-(t-t')/\tau_s} d^2\hat{\mathbf{n}}' dt' \Bigg) \\ &+ \frac{1}{\tau_s} \oint f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I\big(x, y=0, \hat{\mathbf{n}}', t\big) d^2\hat{\mathbf{n}}'. \end{aligned} \tag{3.26}$$

We note that the first and second terms on the right-hand side of equation 3.26 correspond to the right-hand side of the integral equation 3.22 times τ_s^{-1}. Equation 3.26 then becomes

$$\frac{\partial I(x, y=0, \hat{\mathbf{n}}, t)}{\partial t} + v\partial_x I(x, y=0, \hat{\mathbf{n}}, t) = -\frac{1}{\tau_s} I\big(x, y=0, \hat{\mathbf{n}}, t\big) + \frac{1}{\tau_s} \oint f(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I\big(x, y=0, \hat{\mathbf{n}}', t\big) d^2\hat{\mathbf{n}}', \tag{3.27}$$

which generalizes to the radiative transfer equations 3.3 since the orientation of the x-axis is arbitrary.

CHAPTER 4

ANGULAR AND MODAL EQUIPARTITIONING OF ELASTIC WAVES IN SCATTERING MEDIA: A NUMERICAL STUDY BASED ON ENERGY TRANSPORT

To be submitted to the Journal of the Acoustical Society of America.

Manuel Jaimes-Caballero[7], and Roel Snieder[8].

4.1 Abstract

We illustrate the angular and modal equipartitioning of elastic waves in scattering media using the elastic radiative transfer equations in 2-D. To solve these equations we decompose the P and S specific intensities into direct and scattered components. We handle the direct component analytically, and derive integral equations for the scattered components of the P and S specific intensities. We construct a time-stepping algorithm with which we evolve the scattered components of the specific intensities numerically in time. We handle the advection of P and S energy analytically at the computational grid points and use numerical interpolation to deal with advection terms which do not lie on the grid points. This approach allows us to reduce numerical dispersion, compared to standard numerical techniques. We test this algorithm for a pure P source, and a double couple which radiates both P and S energy. We compare our numerical solutions against known approximations, and find good agreement. We use this algorithm to numerically study the local behavior of equipartitioning over wave modes and angular directions. We find that both types of equipartitioning are a function of space and time, depending on the extent of scattering. This local behavior must be taken into account when studying diffusion and equipartitioning of elastic waves.

4.2 Introduction

As waves propagate through a scattering medium they lose (or gain) energy due to scattering to (or from) other directions. When dealing with elastic waves the scattering also accounts for mode conversion. In the absence of anelastic attenuation, scattering and mode conversion obey energy conservation, which one may mathematically describe using the radiative transfer equations (RTE). The RTE consist of a coupled system of integro-differential equations where one solves for the wave intensity as a function of space, time, and angular direction (and wave mode if dealing with elastic waves), assuming one knows the scattering mean free path(s), the angular dependence of the scattering pattern(s), and the speed(s) of

[7]Graduate student, primary researcher, and author at Department of Geophysics, Colorado School of Mines.
[8]Supervisor at Department of Geophysics, Colorado School of Mines.

energy propagation. Both acoustic and elastic formulations to the RTE have been studied (Gaebler et al., 2015; Margerin et al., 2016; Zhang and Sens-Schönfelder, 2022), with most of the attention devoted to the acoustic formulation. The elastic formulation is harder to study than the acoustic formulation, because there is coupling not only in the angular direction but also in the mode of propagation due to $P \rightarrow S$ and $S \rightarrow P$ mode conversions. In 2-D the acoustic RTE consists of a coupled system of N equations, where N refers to number of angular directions. In contrast, in 2-D the elastic RTE consists of a coupled system of $2N$ equations. In 3-D the elastic RTE consists of a coupled system of $5NM$ equations, where the factor of five corresponds to the size of the stokes vector, N is the number of angular directions in the azimuthal plane and M is the number of angular directions in the polar plane (Turner and Weaver, 1994).

The acoustic description to RTE has been used in astrophysics to analyze radiation transport across cosmic dust in a wide range of astrophysical objects (Steinacker et al., 2002; Narayanan et al., 2021; Wolf, 2003); in atmospheric sciences to model solar radiation across clouds to better understand the evolution of sea surface temperatures (Evans and Stephens, 1995; Aumann et al., 2018; Manners et al., 2009); in optics to develop novel optical tomographic imaging algorithms which allow diagnosis and treatment of biological tissues (Klose et al., 2002; Abdoulaev, 2003; Ren et al., 2004; Yodh and Chance, 1995); in acoustics to model the interaction of acoustic waves with the ocean bottom (Quijano and Zurk, 2009), forest acoustics (Ostashev et al., 2017), and acoustical diffractions by obstacles (Reboul et al., 2005); in geophysics to model infrared radiation across volcanic ash clouds (Prata, 1989; Francis et al., 2012; Lee et al., 2014), heat transfer in the mantle (Hofmeister, 2005), and scattering kernels in coda wave interferometry (Pacheco and Snieder, 2005; Rossetto et al., 2011; Margerin et al., 2016; Snieder et al., 2019; Duran et al., 2020; Dinther et al., 2021; Obermann et al., 2016).

Several numerical techniques have been developed to solve the acoustic RTE. These include the Discontinuous Galerkin finite element method (Clarke et al., 2019; Han et al., 2010), Markov Chain Monte Carlo (Iwabuchi, 2006; Xu et al., 2011; Camps and Baes, 2018; Noebauer and Sim, 2019; Przybilla and Korn, 2008; Yoshimoto, 2000), Finite Difference (Klose and Hielscher, 1999), and Wave Equation modeling, whereby one exploits the connection between the acoustic wave equation and the acoustic RTE (Przybilla et al., 2006; Kanu and Snieder, 2015; Snieder et al., 2019; Duran et al., 2020). In addition to these numerical techniques, several authors have derived analytical approximations to the acoustic RTE. These approximations include assuming a point-like, isotropic, impulsive source of intensity in a statistically homogeneous medium (Margerin et al., 2016), expanding the intensity and scattering function into a finite sum of Legendre polynomial and then solving a finite system of equations for the unknown coefficients appearing in a truncated expansion (Roberge, 1983), decomposing the specific intensities into a sum of partial intensities and then solving the RTE for each partial intensity assuming that scattering is

angle-independent and that the source is isotropic (Paasschens, 1997), assuming a steady-state intensity field (Fan et al., 2019; Le Hardy et al., 2016), assuming azimuthal symmetry (Baes and Dejonghe, 2001; de Abreu, 2004), or assuming that the wave propagation is diffusive, which is only valid at times much larger than the transport mean free time (Rossetto et al., 2011; Planès et al., 2014).

The elastic formulation to the RTE has been mostly restricted to seismology. Wu and Aki (1985) introduce the RTE in seismology to describe the generation of seismic coda as seismic waves scatter through the heterogeneous Earth's structure. Sato et al. (1997) use RTE to derive semi-analytical expressions of the energy density due to a point shear-dislocation source. Margerin et al. (2000) use Monte Carlo simulations to study the multiple scattering of elastic waves via the elastic RTE. Przybilla et al. (2006) solve the elastic RTE using a Monte Carlo method and compare their results to finite difference modeling of elastic waves in a 2-D random medium. Yamamoto and Sato (2010) use radiative transfer theory to study multiple scattering and mode conversions of seismic waves at Asama volcano in Japan. For a thorough overview of the analytical, numerical, and applied aspects of both the acoustic and elastic RTE we refer the reader to Sato et al. (2012).

The elastic RTE has also been used to study the late time behavior of elastic waves. At late times, after many scattering events, the ratio of S to P total energy equals a constant which depends on the P and S wave mean velocities and the dimension under consideration (Weaver, 1982; Snieder, 2002). Similarly, at late times, the angular distribution of the P and S energy homogenizes (Margerin et al., 2000). This late time behavior is usually referred to as the equipartitioned state of elastic waves. Khazaie et al. (2017) numerically studied the transition to the equipartitioned state in a randomly heterogeneous elastic medium for a wide range of modeling parameters; Sánchez-Sesma et al. (2008) theoretically studied equipartitioning as a necessary condition for the retrieval of Green's functions in seismic inteferometry; and Hennino et al. (2001) investigated the principle of equipartitioning using observational data of seismic coda. In this paper we refer to the equilibration of energy between P and S modes as modal equipartitioning, and the homogenization along the angular direction as angular equipartitioning.

Most studies have only considered the global behavior (i.e., time dependence without regard of spatial location) of the modal equipartitioning. In this paper we investigate both the local and global behavior of the modal and angular equipartitioning. The paper is organized as follows: In subsection 4.3 we provide a description of the elastic RTE, and derive a time-stepping algorithm based on an integral equation formalism that allows us to compute the spatio-temporal-angular distribution of the P and S energies. In subsection 4.4, we show numerical simulations for an explosive source which radiates P waves, and a double couple source which radiates P and S waves. We study the local and global behavior of equipartitioning for both of these sources. In subsection 4.6 we discuss our findings. In subsection 4.8 show

the diffusive approximation that we use to benchmark our algorithm. In subsection 4.9 we show the analytical expressions for the total P and S energies which we compare against our numerical results in subsection 4.4. In subsection 4.10 we show how we make our numerical scheme energy conserving.

4.3 Theory

4.3.1 Description of the Elastic Radiative Transfer Equations

The two dimensional elastic radiative transfer equations correspond to the following coupled system of equations (Zhang et al., 2021)

$$\frac{\partial I^P(\mathbf{r},\hat{\mathbf{n}},t)}{\partial t} + \alpha_0\hat{\mathbf{n}}\cdot\nabla I^P(\mathbf{r},\hat{\mathbf{n}},t) = -\alpha_0(g^0_{pp}+g^0_{ps})I^P(\mathbf{r},\hat{\mathbf{n}},t) + \alpha_0\oint g_{pp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}' + \beta_0\oint g_{sp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}', \quad (4.1)$$

$$\frac{\partial I^S(\mathbf{r},\hat{\mathbf{n}},t)}{\partial t} + \beta_0\hat{\mathbf{n}}\cdot\nabla I^S(\mathbf{r},\hat{\mathbf{n}},t) = -\beta_0(g^0_{ss}+g^0_{sp})I^S(\mathbf{r},\hat{\mathbf{n}},t) + \beta_0\oint g_{ss}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}' + \alpha_0\oint g_{ps}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}', \quad (4.2)$$

where the superscripts P and S indicate the specific intensity of the P and S-waves, respectively. In equations 4.1 and 4.2 we ignore intrinsic attenuation, and α_0 and β_0 describe the mean propagation velocity of P and S waves, respectively. The terms $\hat{\mathbf{n}}' = (\cos\gamma', \sin\gamma')$ and $\hat{\mathbf{n}} = (\cos\gamma, \sin\gamma)$, with γ'(or γ) the angle relative to the positive $x-$direction, denote the incident and scattered propagation directions, respectively, while $\nabla = (\partial/\partial x, \partial/\partial y)$ denotes the gradient operator. The terms $g_{ij}(\hat{\mathbf{n}},\hat{\mathbf{n}}')$, with $\{i,j = p,s\}$, denote the scattering functions, which describe scattering between different angles and mode conversion. They quantify the energy transfer from a wave of mode i incident in direction $\hat{\mathbf{n}}'$ to a wave that scatters as a mode j towards direction $\hat{\mathbf{n}}$. These scattering functions assume that scattering only depends on $\hat{\mathbf{n}}'\cdot\hat{\mathbf{n}} = \cos\theta$, where θ is the angle between the incident and scattered directions. Therefore, $\theta = 0$ corresponds to forward scattering, and $\theta = \pi$ corresponds to backward scattering. The terms g^0_{ij} denote the total scattering coefficients and are defined as $g^0_{ij} = \oint g_{ij}(\theta)d\theta$. They describe the energy transfer from a wave of mode i to a wave of mode j. The left-hand side of equations 4.1 and 4.2 describe the advection of the P and S specific intensities, respectively. The first term on the right-hand side of equations 4.1 and 4.2 describes scattering losses due to scattering to other angles and mode conversion. The integral terms describe energy gain due to scattering from other angles and mode conversion. Equations 4.1 and 4.2 follow from the ladder approximation to the Bethe-Salpeter equation assuming that $ka\epsilon \ll 1$, where k is the wavenumber, and a, ϵ^2 are the correlation length and variance of the random medium, respectively (Przybilla et al. ,2006). In this chapter we refer to $I^m(\mathbf{r},\hat{\mathbf{n}},t)$ as the *specific intensity*,

$E^m(\mathbf{r},t) = \oint I^m(\mathbf{r},\hat{\mathbf{n}},t)d^2\hat{\mathbf{n}}$ as the *energy density*, and $Y^m(t) = \int E^m(\mathbf{r},t)dxdy$ as the *total energy*. The superscript m indicates P or S wave mode.

For late times, when the wave propagation is almost independent of direction, and is nearly stationary in time, the elastic RTE equations 4.1 and 4.2 lead to diffusive wave transport (Sato et al., 2012). The conventional diffusion approximation to the elastic RTE describes the spatio-temporal evolution of the combined energy density $E(\mathbf{r},t) = E^P(\mathbf{r},t) + E^S(\mathbf{r},t)$. Describing the energies E^P and E^S independently of each other is not possible due to the coupling between wave modes (Trégourès and van Tiggelen, 2002). In the diffusive regime, after enough scattering events, the ratio of the total energy carried by S waves, to that carried by P waves, is $2(\alpha_0/\beta_0)^3$ in 3-D (Weaver, 1982; Snieder, 2002) and $(\alpha_0/\beta_0)^2$ in 2-D (Zhang et al., 2021). Despite the coupling between wave modes, one can derive expressions for the total energy, rather than the energy density, carried by the P and S waves. Sato et al. (2012) give expressions for the evolution of the total energies $Y^P(t)$ and $Y^S(t)$. To do this, they assume the source is isotropic, that the P and S specific intensities are almost isotropic, and that the spatial gradient of the P and S energy densities vanishes at large distances. With some modifications their expressions are also valid for non-isotropic sources, under the requirement that the spatial gradient of the P and S specific intensities vanishes at large distances from the source, as we show in subsection 4.9. These expressions show that the rate at which $Y^P(t)$ and $Y^S(t)$ vary over time is only a function of the scattering properties of the medium.

4.3.2 Formulation of the Time-Stepping Algorithm

We solve the system of equations 4.1 and 4.2, with the P and S specific intensities subject to the following initial conditions

$$\begin{aligned} I^P(\mathbf{r},\hat{\mathbf{n}},t=t_0) &= I_0^P(\mathbf{r},\hat{\mathbf{n}}), \\ I^S(\mathbf{r},\hat{\mathbf{n}},t=t_0) &= I_0^S(\mathbf{r},\hat{\mathbf{n}}). \end{aligned} \tag{4.3}$$

We carry out the simulation for times less than the propagation time for a P-wave to reach the boundary. In such case, the boundary values do not affect the simulation, as long as no energy is injected into the computational domain from the boundary. We decompose the specific intensities into the direct and scattered components

$$\begin{aligned} I^P(\mathbf{r},\hat{\mathbf{n}},t) &= I^P_{dir}(\mathbf{r},\hat{\mathbf{n}},t) + I^P_{scat}(\mathbf{r},\hat{\mathbf{n}},t), \\ I^S(\mathbf{r},\hat{\mathbf{n}},t) &= I^S_{dir}(\mathbf{r},\hat{\mathbf{n}},t) + I^S_{scat}(\mathbf{r},\hat{\mathbf{n}},t). \end{aligned} \tag{4.4}$$

Similarly, we decompose the initial conditions in equation 4.3

$$I^P_{dir}(\mathbf{r},\hat{\mathbf{n}},t=t_0)=I^P_0(\mathbf{r},\hat{\mathbf{n}});I^P_{scat}(\mathbf{r},\hat{\mathbf{n}},t=t_0)=0,$$
$$I^S_{dir}(\mathbf{r},\hat{\mathbf{n}},t=t_0)=I^S_0(\mathbf{r},\hat{\mathbf{n}});I^S_{scat}(\mathbf{r},\hat{\mathbf{n}},t=t_0)=0, \tag{4.5}$$

where the subscripts *dir* and *scat* denote the direct and scattered parts of the specific intensity, respectively. This decomposition allows us to reduce numerical errors since the direct component of the specific intensity can be handled analytically. The direct components of the P and S specific intensities satisfy

$$\frac{\partial I^P_{dir}(\mathbf{r},\hat{\mathbf{n}},t)}{\partial t}+\alpha_0\hat{\mathbf{n}}\cdot\nabla I^P_{dir}(\mathbf{r},\hat{\mathbf{n}},t)=-\alpha_0(g^0_{pp}+g^0_{ps})I^P_{dir}(\mathbf{r},\hat{\mathbf{n}},t), \tag{4.6}$$

and

$$\frac{\partial I^S_{dir}(\mathbf{r},\hat{\mathbf{n}},t)}{\partial t}+\beta_0\hat{\mathbf{n}}\cdot\nabla I^S_{dir}(\mathbf{r},\hat{\mathbf{n}},t)=-\beta_0(g^0_{ss}+g^0_{sp})I^S_{dir}(\mathbf{r},\hat{\mathbf{n}},t), \tag{4.7}$$

respectively. The exact solutions to equations 4.6 and 4.7, for the initial conditions in 4.5, are

$$I^P_{dir}(\mathbf{r},\hat{\mathbf{n}},t)=I^P_{dir}(\mathbf{r}-\alpha_0(t-t_0)\hat{\mathbf{n}},\hat{\mathbf{n}},t=t_0)e^{-\alpha_0(g^0_{pp}+g^0_{ps})(t-t_0)}, \tag{4.8}$$

and

$$I^S_{dir}(\mathbf{r},\hat{\mathbf{n}},t)=I^S_{dir}(\mathbf{r}-\beta_0(t-t_0)\hat{\mathbf{n}},\hat{\mathbf{n}},t=t_0)e^{-\beta_0(g^0_{ss}+g^0_{sp})(t-t_0)}. \tag{4.9}$$

Inserting the decomposition in equation 4.4 into equations 4.1 and 4.2, with the help of equations 4.6 and 4.7, gives

$$\begin{aligned}\frac{\partial I^P_{scat}(\mathbf{r},\hat{\mathbf{n}},t)}{\partial t}+\alpha_0\hat{\mathbf{n}}\cdot\nabla I^P_{scat}(\mathbf{r},\hat{\mathbf{n}},t)&=-\alpha_0(g^0_{pp}+g^0_{ps})I^P_{scat}(\mathbf{r},\hat{\mathbf{n}},t)\\&+\alpha_0\oint g_{pp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}'+\beta_0\oint g_{sp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}',\end{aligned} \tag{4.10}$$

$$\begin{aligned}\frac{\partial I^S_{scat}(\mathbf{r},\hat{\mathbf{n}},t)}{\partial t}+\beta_0\hat{\mathbf{n}}\cdot\nabla I^S_{scat}(\mathbf{r},\hat{\mathbf{n}},t)&=-\beta_0(g^0_{ss}+g^0_{sp})I^S_{scat}(\mathbf{r},\hat{\mathbf{n}},t)\\&+\beta_0\oint g_{ss}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}'+\alpha_0\oint g_{ps}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r},\hat{\mathbf{n}}',t)d^2\hat{\mathbf{n}}'.\end{aligned} \tag{4.11}$$

Note that in equations 4.10 and 4.11 the integral contributions still contain the total specific intensity for each of the modes. Using the work of Paasschens (1997) we transform equations 4.10 and 4.11 into the following integral equations

$$I^P_{scat}(\mathbf{r},\hat{\mathbf{n}},t) = I^P_{scat}(\mathbf{r}-\alpha_0(t-t_0)\hat{\mathbf{n}},\hat{\mathbf{n}},t_0)e^{-(t-t_0)\alpha_0(g^0_{pp}+g^0_{ps})} + \int_{t_0}^{t}\oint\Big[\alpha_0 g_{pp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\alpha_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t') + \beta_0 g_{sp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\alpha_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t')\Big]e^{-(t-t')\alpha_0(g^0_{pp}+g^0_{ps})}d^2\hat{\mathbf{n}}'dt', \tag{4.12}$$

$$I^S_{scat}(\mathbf{r},\hat{\mathbf{n}},t) = I^S_{scat}(\mathbf{r}-\beta_0(t-t_0)\hat{\mathbf{n}},\hat{\mathbf{n}},t_0)e^{-(t-t_0)\beta_0(g^0_{ss}+g^0_{sp})} + \int_{t_0}^{t}\oint\Big[\beta_0 g_{ss}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\beta_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t') + \alpha_0 g_{ps}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\beta_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t')\Big]e^{-(t-t')\beta_0(g^0_{ss}+g^0_{sp})}d^2\hat{\mathbf{n}}'dt'. \tag{4.13}$$

Paasschens (1997) analytically solve for the specific intensities in the acoustic version of equation 4.12 (i.e., they disregard the terms involving conversion from P to S waves) for a medium with isotropic scattering and constant scattering properties assuming that the initial condition for the specific intensity is isotropic. To do this, they decompose the specific intensities into a sum of partial intensities, where each partial intensity signifies a number of scattering events (i.e., direct propagation, single scattering, double scattering, and so on). In equations 4.12 and 4.13 we deal with the advection of the specific intensities exactly at the grid points, and through interpolation at points that lie within the grid points. This approach allows us to reduce the numerical dispersion, compared to standard numerical techniques. In this chapter we use equation 4.12 and 4.13 to integrate the specific intensities numerically in time. This procedure allows us to use non-isotropic initial conditions and non-isotropic scattering radiation patterns.

To develop a time-stepping algorithm which depends only on the current time t, and the previous time $t-\Delta t$, we replace t_0 by $t-\Delta t$ to obtain

$$I^P_{scat}(\mathbf{r},\hat{\mathbf{n}},t) = I^P_{scat}(\mathbf{r}-\alpha_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}},t-\Delta t)e^{-\Delta t\alpha_0(g^0_{pp}+g^0_{ps})} + \int_{t-\Delta t}^{t}\oint\Big[\alpha_0 g_{pp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\alpha_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t') + \beta_0 g_{sp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\alpha_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t')\Big]e^{-(t-t')\alpha_0(g^0_{pp}+g^0_{ps})}d^2\hat{\mathbf{n}}'dt', \tag{4.14}$$

$$I^S_{scat}(\mathbf{r},\hat{\mathbf{n}},t) = I^S_{scat}(\mathbf{r}-\beta_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}},t-\Delta t)e^{-\Delta t\beta_0(g^0_{ss}+g^0_{sp})} + \oint\int_{t-\Delta t}^{t}\Big[\beta_0 g_{ss}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\beta_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t') + \alpha_0 g_{ps}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\beta_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t')\Big]e^{-(t-t')\beta_0(g^0_{ss}+g^0_{sp})}d^2\hat{\mathbf{n}}'dt'. \tag{4.15}$$

We proceed by applying the discrete ordinate method (Chandrasekhar, 1960), whereby one divides the angular integral into N equal segments. In two dimensions this method corresponds to setting $d^2\hat{\mathbf{n}} \to d\theta$ and then setting $d\theta = 2\pi/N$ in the angular integration

$$\begin{aligned} I^P_{scat}(\mathbf{r},\hat{\mathbf{n}},t) &= I^P_{scat}(\mathbf{r}-\alpha_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}},t-\Delta t)e^{-\Delta t\alpha_0(g^0_{pp}+g^0_{ps})} \\ &+ \frac{2\pi}{N}\int_{t-\Delta t}^{t}\sum_{\hat{\mathbf{n}}'}\Big[\alpha_0 g_{pp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\alpha_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t') \\ &+ \beta_0 g_{sp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\alpha_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t')\Big]e^{-(t-t')\alpha_0(g^0_{pp}+g^0_{ps})}dt', \end{aligned} \tag{4.16}$$

$$\begin{aligned} I^S_{scat}(\mathbf{r},\hat{\mathbf{n}},t) &= I^S_{scat}(\mathbf{r}-\beta_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}},t-\Delta t)e^{-\Delta t\beta_0(g^0_{ss}+g^0_{sp})} \\ &+ \frac{2\pi}{N}\int_{t-\Delta t}^{t}\sum_{\hat{\mathbf{n}}'}\Big[\beta_0 g_{ss}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\beta_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t') \\ &+ \alpha_0 g_{ps}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\beta_0(t-t')\hat{\mathbf{n}},\hat{\mathbf{n}}',t')\Big]e^{-(t-t')\beta_0(g^0_{ss}+g^0_{sp})}dt'. \end{aligned} \tag{4.17}$$

We then discretize the remaining time integral using a two-point quadrature rule, with weights not necessarily equal to 1/2, by which we obtain

$$\begin{aligned} I^P_{scat}(\mathbf{r},\hat{\mathbf{n}},t) &= I^P_{scat}(\mathbf{r}-\alpha_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}},t-\Delta t)e^{-\Delta t\alpha_0(g^0_{pp}+g^0_{ps})} \\ &+ A_p\Delta t\frac{2\pi}{N}\sum_{\hat{\mathbf{n}}'}\Big[\alpha_0 g_{pp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\alpha_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}}',t-\Delta t) \\ &+ \beta_0 g_{sp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\alpha_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}}',t-\Delta t)\Big]e^{-\Delta t\alpha_0(g^0_{pp}+g^0_{ps})} \\ &+ B_p\Delta t\frac{2\pi}{N}\sum_{\hat{\mathbf{n}}'}\Big[\alpha_0 g_{pp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r},\hat{\mathbf{n}}',t) \\ &+ \beta_0 g_{sp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r},\hat{\mathbf{n}}',t)\Big], \end{aligned} \tag{4.18}$$

$$\begin{aligned} I^S_{scat}(\mathbf{r},\hat{\mathbf{n}},t) &= I^S_{scat}(\mathbf{r}-\beta_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}},t-\Delta t)e^{-\Delta t\beta_0(g^0_{ss}+g^0_{sp})} \\ &+ A_s\Delta t\frac{2\pi}{N}\sum_{\hat{\mathbf{n}}'}\Big[\beta_0 g_{ss}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\beta_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}}',t-\Delta t) \\ &+ \alpha_0 g_{ps}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\beta_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}}',t-\Delta t)\Big]e^{-\Delta t\beta_0(g^0_{ss}+g^0_{sp})} \\ &+ B_s\Delta t\frac{2\pi}{N}\sum_{\hat{\mathbf{n}}'}\Big[\beta_0 g_{ss}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r},\hat{\mathbf{n}}',t) \\ &+ \alpha_0 g_{ps}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r},\hat{\mathbf{n}}',t)\Big]. \end{aligned} \tag{4.19}$$

A_p and A_s denote the lower bound coefficients (corresponding to the intensity at time $t-\Delta t$) in the quadrature rule for the time integral in equations 4.16 and 4.17, respectively. Similarly, B_p and B_s denote

the upper bound coefficients (corresponding to the intensity at time t) in the quadrature rule. For the trapezoidal quadrature rule that we use in this paper we set $A_p = A_s = B_p = B_s = 1/2$. In subsection 4.10 we show that this quadrature gives a time-steppping scheme that is energy conserving, as long as the time step Δt is much smaller than the scattering mean free time of P and S waves. Equations 4.18 and 4.19 give the following system of equations

$$\begin{bmatrix} \boldsymbol{\sigma}^{PP} & \boldsymbol{\sigma}^{SP} \\ \boldsymbol{\sigma}^{PS} & \boldsymbol{\sigma}^{SS} \end{bmatrix} \begin{bmatrix} \vec{I}^P \\ \vec{I}^S \end{bmatrix} = \begin{bmatrix} \vec{W}^P \\ \vec{W}^S \end{bmatrix}. \tag{4.20}$$

where $\vec{I}^P, \vec{I}^S \in \mathcal{R}^{N\times 1}$ are the solution vectors for the P and S wave intensities, respectively. The vectors $\vec{W}^P, \vec{W}^S \in \mathcal{R}^{N\times 1}$ are

$$\begin{aligned} \vec{W}^P(\mathbf{r},\hat{\mathbf{n}},t) &= I^P_{scat}(\mathbf{r}-\alpha_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}},t-\Delta t)e^{-\Delta t\alpha_0(g^0_{pp}+g^0_{ps})} \\ &+ A_p\Delta t\frac{2\pi}{N}\sum_{\hat{\mathbf{n}}'}\Big[\alpha_0 g_{pp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\alpha_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}}',t-\Delta t) \\ &+ \beta_0 g_{sp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\alpha_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}}',t-\Delta t)\Big]e^{-\Delta t\alpha_0(g^0_{pp}+g^0_{ps})} \\ &+ B_p\Delta t\frac{2\pi}{N}\sum_{\hat{\mathbf{n}}'}\Big[\alpha_0 g_{pp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P_{dir}(\mathbf{r},\hat{\mathbf{n}}',t) \\ &+ \beta_0 g_{sp}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S_{dir}(\mathbf{r},\hat{\mathbf{n}}',t)\Big], \end{aligned} \tag{4.21}$$

$$\begin{aligned} \vec{W}^S(\mathbf{r},\hat{\mathbf{n}},t) &= I^S_{scat}(\mathbf{r}-\beta_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}},t-\Delta t)e^{-\Delta t\beta_0(g^0_{ss}+g^0_{sp})} \\ &+ A_s\Delta t\frac{2\pi}{N}\sum_{\hat{\mathbf{n}}'}\Big[\beta_0 g_{ss}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S(\mathbf{r}-\beta_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}}',t-\Delta t) \\ &+ \alpha_0 g_{ps}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P(\mathbf{r}-\beta_0\Delta t\hat{\mathbf{n}},\hat{\mathbf{n}}',t-\Delta t)\Big]e^{-\Delta t\beta_0(g^0_{ss}+g^0_{sp})} \\ &+ B_s\Delta t\frac{2\pi}{N}\sum_{\hat{\mathbf{n}}'}\Big[\beta_0 g_{ss}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^S_{dir}(\mathbf{r},\hat{\mathbf{n}}',t) \\ &+ \alpha_0 g_{ps}(\hat{\mathbf{n}},\hat{\mathbf{n}}')I^P_{dir}(\mathbf{r},\hat{\mathbf{n}}',t)\Big], \end{aligned} \tag{4.22}$$

and we define $\boldsymbol{\sigma}^{PP}$, $\boldsymbol{\sigma}^{SP}$, $\boldsymbol{\sigma}^{PS}$, $\boldsymbol{\sigma}^{SS} \in \mathcal{R}^{N\times N}$ as

$$\begin{aligned} &\sigma^{PP}_{ij} = \delta_{ij} - \alpha_0 B_p\frac{2\pi}{N}\Delta t g_{pp}(\hat{\mathbf{n}}_i,\hat{\mathbf{n}}_j);\ \sigma^{SP}_{ij} = -\beta_0 B_p\frac{2\pi}{N}\Delta t g_{sp}(\hat{\mathbf{n}}_i,\hat{\mathbf{n}}_j) \\ &\sigma^{PS}_{ij} = -\alpha_0 B_s\frac{2\pi}{N}\Delta t g_{ps}(\hat{\mathbf{n}}_i,\hat{\mathbf{n}}_j);\ \sigma^{SS}_{ij} = \delta_{ij} - \beta_0 B_s\frac{2\pi}{N}\Delta t g_{ss}(\hat{\mathbf{n}}_i,\hat{\mathbf{n}}_j). \end{aligned} \tag{4.23}$$

Directly solving the linear system 5.14 is expensive. For a computational domain of size (N_x, N_y, N_t), one would need to perform $N_x * N_y * N_t$ linear solves. To speed up the computation we compute the inverse of

the matrix

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{\sigma}^{PP} & \boldsymbol{\sigma}^{SP} \\ \boldsymbol{\sigma}^{PS} & \boldsymbol{\sigma}^{SS} \end{bmatrix} \tag{4.24}$$

before-hand via LU decomposition with partial pivoting (Chapter 3, Golub and Van Loan, 2013). We quantify the stability of inverting the matrix $\boldsymbol{A}$ in the linear system 5.14 with the condition number $\kappa(\boldsymbol{A}) = ||\boldsymbol{A}||\ ||\boldsymbol{A}^{-1}||$, where $||\cdot||$ is the 2-norm. When the condition number $\kappa(\boldsymbol{A}) = 1$, the matrix $\boldsymbol{A}$ is well-conditioned. We can control this condition number with the value of the time step Δt. For the numerical experiments in this chapter we have a condition number $\kappa(\boldsymbol{A}) = 1.03$. We re-write the linear system 5.14 as

$$\begin{bmatrix} \vec{I}^{P} \\ \vec{I}^{S} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\sigma}^{PP} & \boldsymbol{\sigma}^{SP} \\ \boldsymbol{\sigma}^{PS} & \boldsymbol{\sigma}^{SS} \end{bmatrix}^{-1} \begin{bmatrix} \vec{W}^{P} \\ \vec{W}^{S} \end{bmatrix}. \tag{4.25}$$

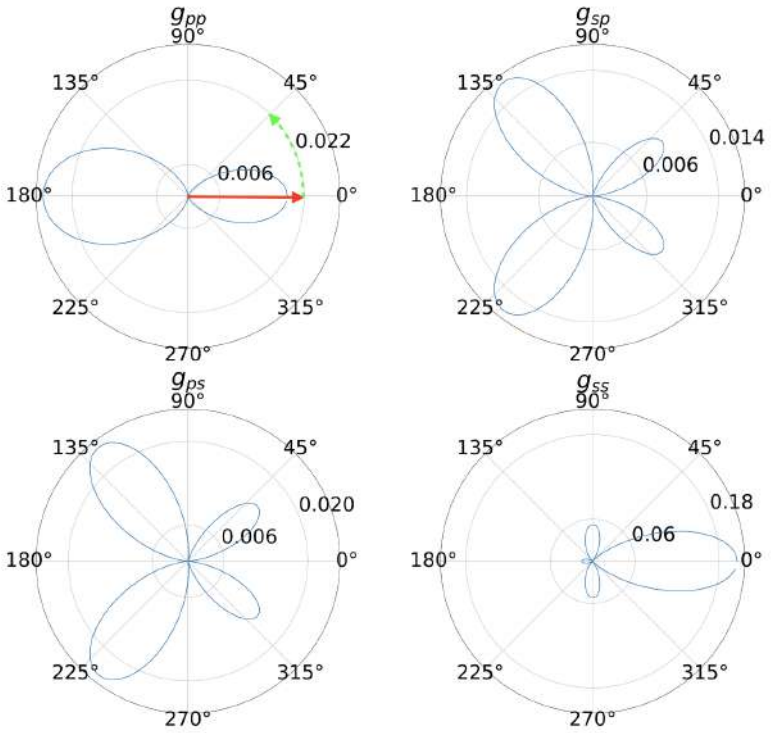

Figure 4.1 Polar plots of radiation patterns for an incident wave of mode i propagating along $\theta = 0$, which scatters to mode j. The red arrow indicates the incident direction, and the green dashed arrow indicates the direction of increasing scattering angle θ.

For the numerical simulations in the next section, we use the scattering functions g_{pp}, g_{sp}, g_{ps} and g_{ss} derived by Zhang et al. (2021) for a 2-D exponential random medium. We use $\alpha_0 = 6$ km/s, $\beta_0 = 4$ km/s,

and set the correlation distance $a = 0.1$ km, and the variance $\epsilon^2 = 0.0256$. We set the P and S wavenumbers to $k_p = 2\pi v_s/(\lambda_s v_p)$ and $k_s = 2\pi/\lambda_s$, respectively, with $\lambda_s = 1$ km. This choice of parameters gives a scattering mean free time of $\tau_p = 1.36$ s and $\tau_s = 0.82$ s for P and S waves, respectively. We illustrate the angular dependence of these scattering coefficients in Figure 4.1.

4.4 Numerical Simulations

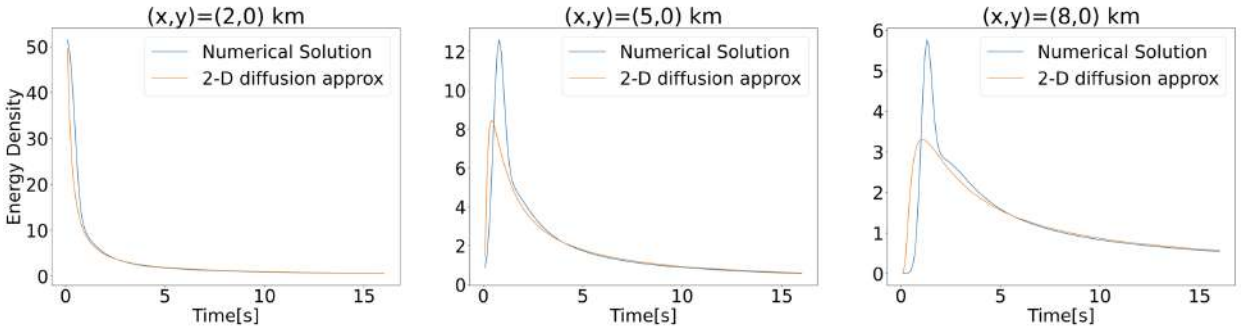

Figure 4.2 Comparison of $E(\mathbf{r}, t) = E^P(\mathbf{r}, t) + E^S(\mathbf{r}, t)$ for the pure P source for different distances (shown above in each panel). We compare our numerical solution against the diffusive approximation to the 2-D RTE.

In this section we show numerical simulations for a source which releases only P energy, and a double-couple source which releases both P and S energy. For both simulations we study the local (i.e., function of space and time) and global (i.e., function of time) behavior of modal and angular equipartitioning.

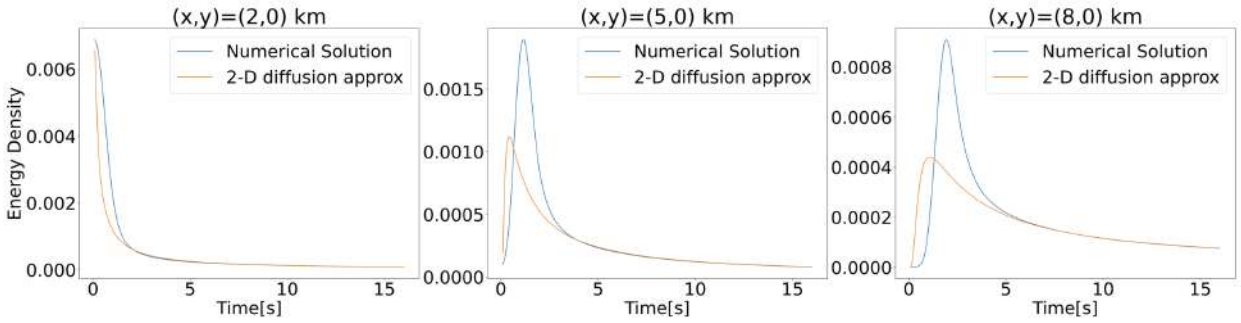

Figure 4.3 Comparison of $E(\mathbf{r}, t) = E^P(\mathbf{r}, t) + E^S(\mathbf{r}, t)$ for the double couple source.

For both simulations we use a spatial grid $\Delta x = \Delta y = 1000$ m, with $(x, y) \in [-100, 100]$ km. We use $N = 120$ angular directions. We evolve the P and S specific intensities from $t = 0$ to $T = 16$ s, with $\Delta t = 0.1$ s.

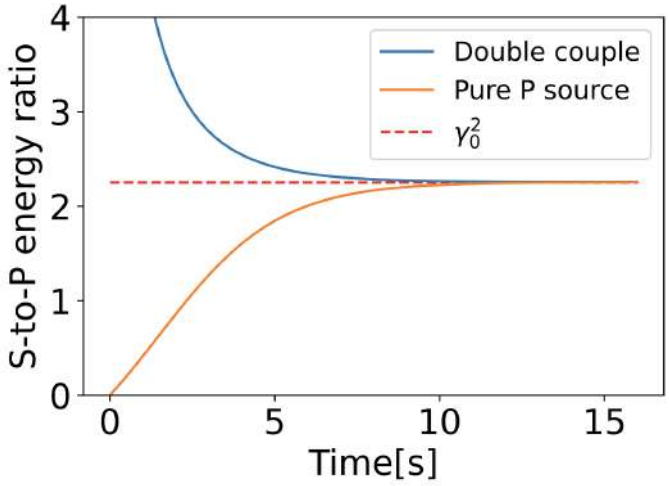

Figure 4.4 The ratio $Y^S(t)/Y^P(t)$ for the P source (orange line) and double couple (blue line) simulation. The dashed red line shows the asymptotic ratio $\gamma_0^2 = (\alpha_0/\beta_0)^2$.

4.4.1 Benchmarking of Algorithm

To benchmark the algorithm we use three different tests. First, we compare our numerical solution of the sum of the P and S energy densities, $E(\mathbf{r},t) = E^P(\mathbf{r},t) + E^S(\mathbf{r},t)$, to the diffusive approximation to the 2-D elastic RTE. As mentioned before, the diffusive approximation describes the spatio-temporal evolution of the energy density $E(\mathbf{r},t)$, rather than the individual P and S energy densities. Since most diffusive approximations to the elastic RTE are derived for a 3-D medium, we show in subsection 4.8 a heuristic 2-D analogous derivation. Second, we compare the global equipartitioning ratio of S to P total energy of our simulations to the theoretical value $(\alpha_0/\beta_0)^2$. This comparison allows us to test the long time behavior of the numerical simulations. Third, we compare the evolution of the P and S energies, $Y^P(t)$ and $Y^S(t)$, of our numerical simulations to analytical expressions. As mentioned before, these expressions are also valid for non-isotropic sources.We center the explosive source around $(x,y) = (0,0)$ and set the initial condition $I^P_{dir}(\mathbf{r},\hat{\mathbf{n}},t=0) = e^{-r^2/2\sigma^2}$ and $I^S_{dir}(\mathbf{r},\hat{\mathbf{n}},t=0) = 0$, with $\sigma = 1.55\times10^3$ m and $r^2 = x^2+y^2$. For the double couple source we set $I^P_{dir}(\mathbf{r},\hat{\mathbf{n}},t=0) = \sin^2(2\theta)e^{-r^2/2\sigma^2}/\alpha_0^6$, and $I^S_{dir}(\mathbf{r},\hat{\mathbf{n}},t=0) = \cos^2(2\theta)e^{-r^2/2\sigma^2}/\beta_0^6$, with $\theta\in[0,2\pi]$ and $\sigma = 1.55\times10^3$ m. The constants $1/\alpha_0^6$ and $1/\beta_0^6$ correspond to the square of the constants $1/\alpha_0^3$ and $1/\beta_0^3$ appearing in the far-field component of the displacement field in an infinite homogeneous medium due to a double couple radiation pattern (Chapter 4, Aki and Richards, 2009). We evolve the P and S scattered specific intensities using equation 4.25. We compute the direct specific intensities exactly using expressions 4.8 and 4.9.

Figure 4.2 shows a comparison of the explosive source numerical solution against the 2-D diffusive approximation to RTE for three different distances, $(x,y) = (2,0), (x,y) = (5,0)$, and $(x,y) = (8,0)$ km.

Figure 4.3 shows the comparison for the double couple source simulation. For all of the panels, in both Figure 4.2 and Figure 4.3, our numerical solution matches the diffusive approximation for times much larger than the P and S scattering mean free times ($t \gg \tau_s, \tau_p$). As one goes further away from the source it takes a longer time, relative to points close to the source, for the energy to become diffuse. This explains why, as we go from the left to the middle to the right panel, the solutions only match at late times. To

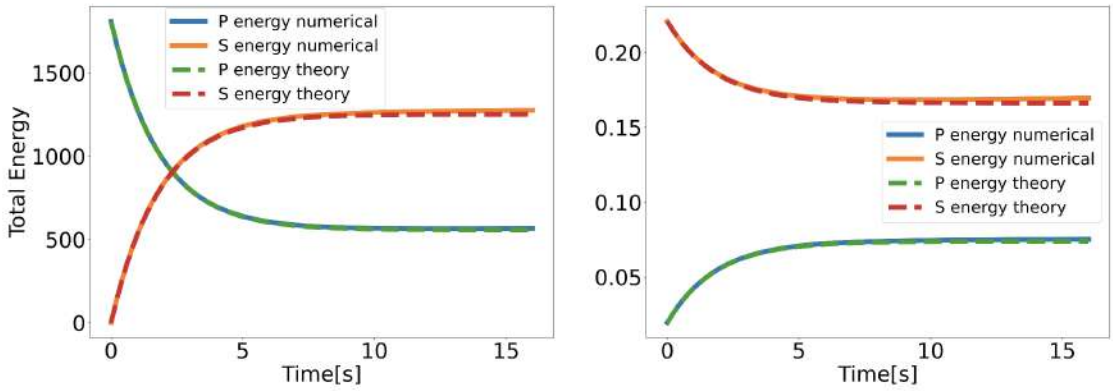

Figure 4.5 Comparison of the P and S energies, $Y^P(t)$ and $Y^S(t)$. The blue and orange solid lines show the numerical P and S energy, respectively. The green and red dashed lines show the analytical P and S energy, respectively. The left and right panel show the comparison of the explosive source with a Gaussian extent and double couple source simulations, respectively.

further test the long time behavior of our algorithm we study the asymptotic behavior of the modal equipartitioning ratio, which converges to $(\alpha_0/\beta_0)^2$ as $t \to \infty$. In practice, this ratio is reached for times much larger than the P and S scattering mean free times ($t \gg \tau_s, \tau_p$). Figure 4.4 shows the equipartitioning ratio $Y^S(t)/Y^P(t)$ for both the explosive and double couple source simulations. For times much larger than the P and S scattering times, the numerical ratio converges to the theoretical ratio. This convergence implies a stable long-time behavior of the simulations. For the double couple simulation, the numerical ratio converges to the theoretical ratio slightly faster than for the explosive source simulation. This is because at time $t = 0$ both wave modes carry energy, which speeds up the equilibration of energy. In addition to studying the long time behavior of the energy propagation, we analyze the global distribution of the P and S energies over the whole simulation range $t \in [0, T]$. Figure 4.5 shows a comparison of the numerical and analytical evolution of the P and S energies for both the explosive (left panel) and double couple source (right panel) simulation. Overall, our simulation is able to predict the analytical values in the whole simulation range, indicating that the algorithm is accurate. At late times, the numerical solution slightly deviates from analytical values. The accuracy of the solution may be improved by decreasing the value of Δt.

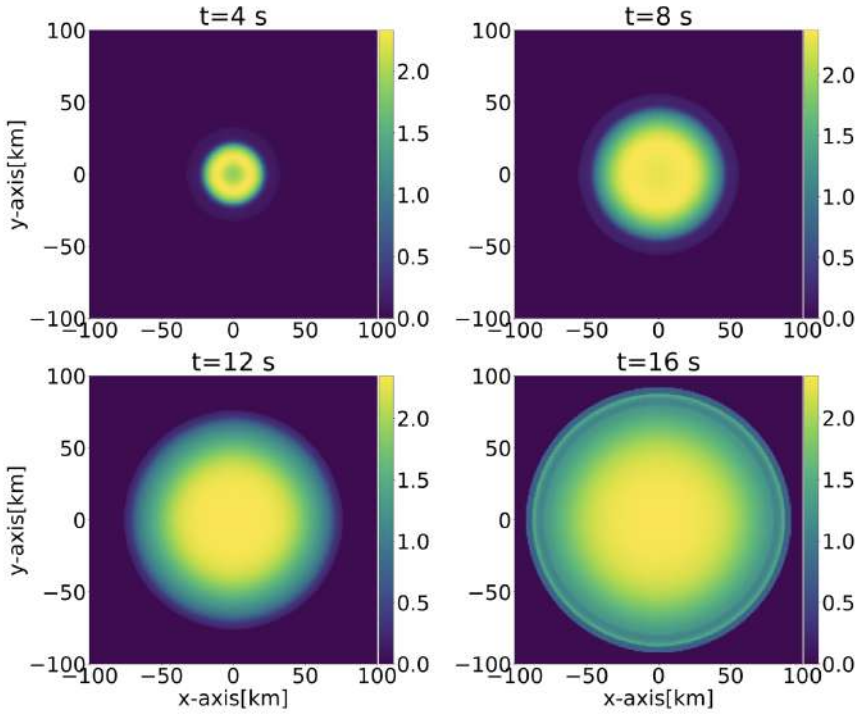

Figure 4.6 Snapshots of the local energy ratio $E^S(\mathbf{r},t)/E^P(\mathbf{r},t)$ for the pure P source. This ratio reaches an equilibrium value of 2.34 within the expanding circle in yellow.

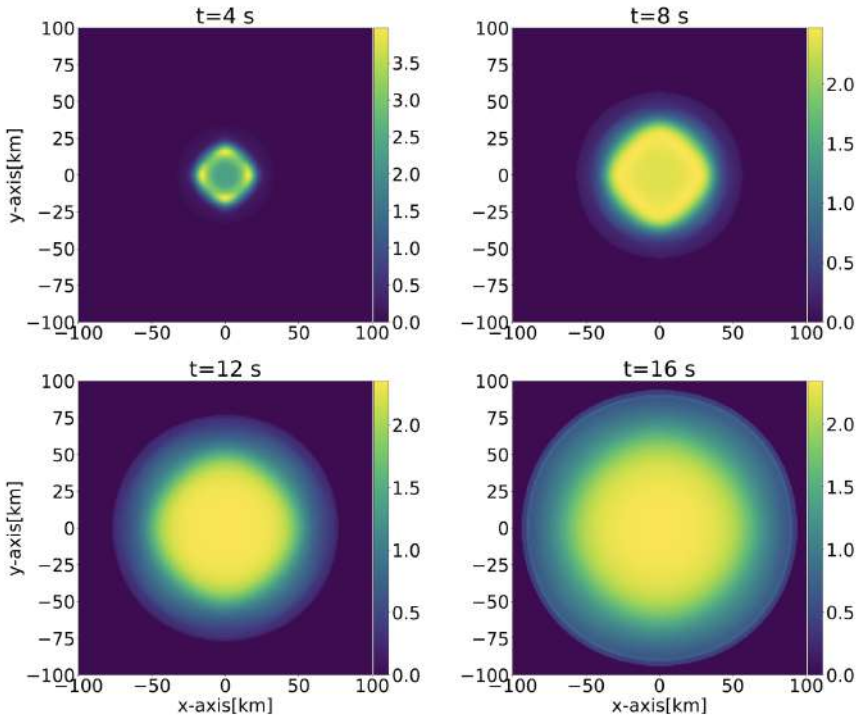

Figure 4.7 Snapshots of the local energy ratio $E^S(\mathbf{r},t)/E^P(\mathbf{r},t)$ for the double couple simulation. As in Figure 4.6, the local ratio reaches an equilibrium value of 2.34 within the expanding yellow circle.

4.5 Local Behavior of Modal Equipartitioning

In 2-D, $\lim_{t\to\infty} Y^S(t)/Y^P(t) = (\alpha_0/\beta_0)^2$, regardless of the source properties (Sato et al., 2012). This ratio is controlled by the repeated conversion between P and S waves, as the energy that is carried by each of the wave modes equilibrates over time. This is shown by the late time behavior of $Y^P(t)$ and $Y^S(t)$ in Figure 4.5. Locally, it also follows that $\lim_{t\to\infty} E^S(\mathbf{r},t)/E^P(\mathbf{r},t) = (\alpha_0/\beta_0)^2$ (Sato et al., 2012). Generally speaking, it takes a longer time to reach this local ratio than to reach the global energy ratio. This is because for the local ratio to converge to a constant, regardless of spatial position, the distribution of the energy density must be homogeneous, over space and angular direction, so that the net density flux is zero (i.e., equipartitioned state).

In this section we investigate numerically the spatio-temporal dependence of the local ratio. Figure 4.6 and Figure 4.7 show snapshots of the local ratio of the explosive and double couple source simulations, respectively, for simulation times $t = (4, 8, 12, 16)$ s. These figures show an inner ring corresponding to S energy propagation and an outer ring corresponding to P energy propagation. In both simulations the local ratio approaches a numerical value of 2.34 at $t = 16$ s within the expanding yellow circle centered around the peak of the Gaussian. As energy propagates outwards, it scatters and starts to equilibrate over the 2 available wave modes. Figure 4.6 shows that the area where the local ratio starts to converge towards $(\alpha_0/\beta_0)^2 = 2.34$ increases over time as the number of scattering events increase. However, the rate at which the local ratio converges to the theoretical value is a function of location, depending on the extent of scattering and on mode conversions. Figure 4.7 shows a local behavior for the modal ratio for the double couple source that is similar to that of the one in Figure 4.6 for the pure P-source. Even though globally the mode equilibration occurs more quickly for the double couple source than for the explosive source (see Figure 4.4), locally that is not the case. After $t > 8$ s there is little difference in the local behavior of the modal equipartitioning between the two numerical simulations. The last panel in Figure 4.6 and Figure 4.7 show that even though at very late times ($t = 16$ s) the global equipartitioning ratio is reached (see Figure 4.4), the local ratio is not yet reached for all locations behind the ballistic wave (i.e., only the locations within the exapnding yellow circle show a local ratio which approaches the global ratio). Only in locations where the energy density is homogenized with enough scattering events, does the local ratio approach the theoretical value. The last panel of Figure 4.6 and Figure 4.7 show that at locations close to the source (inner yellow ring), the local ratio approaches the theoretical value faster than for locations far from the source. For times much larger than for which we evolve the intensities, the local ratio eventually reaches $(\alpha_0/\beta_0)^2$, but it takes significantly longer than for the global ratio to reach the asymptotic value. For such propagation times, when the local ratio stabilizes, the energy distribution enters a locally

equipartitioned state for the energy distribution between P and S modes.

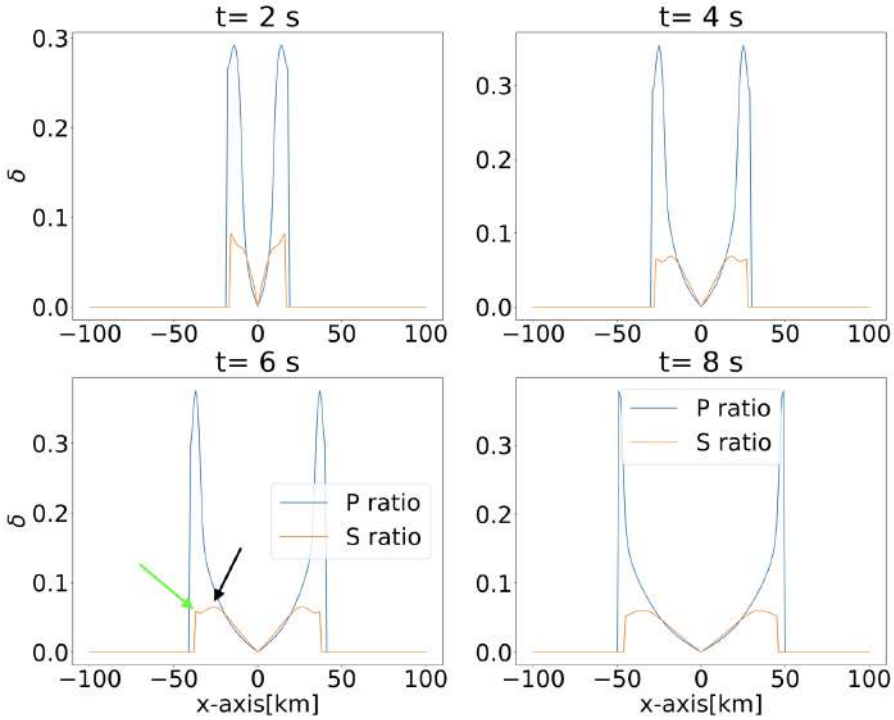

Figure 4.8 Cross sections of the angular equipartitioning ratio δ, for the pure P source simulation, at $y = 0$ and $x \in [-100, 100]$ km. The green and black arrow point to two peaks that arise in δ_S due to mode conversion.

4.5.1 Angular Equipartitioning

For elastic waves, one usually considers the equipartitioning between wave modes, as energy equilibrates due to mode conversion (see Figure 4.4, Figure 4.6 and Figure 4.7). In addition to modal equilibration, angular equilibration occurs as the energy density homogenizes.

Here, we numerically study the angular equipartitioning of the elastic energy as the number of scattering events increase. To quantify the degree of angular equipartitioning we use the *equipartitioning index*, introduced in Chapter 3,

$$\delta_{P,S}(\mathbf{r}, t) = \frac{\sigma_{P,S}(\mathbf{r}, t)}{\sqrt{N}\mu_{P,S}(\mathbf{r}, t)}, \tag{4.26}$$

where the subscript indicates P or S waves. In this expression $\sigma_{P,S}(\mathbf{r}, t)$ is the standard deviation of the P or S specific intensities along the angular directions defined with a division by $N - 1$, as $\sigma^2_{P,S}(\mathbf{r}, t) = \frac{1}{N-1}\sum_{i=1}^{N}(I_{P,S}(\mathbf{r}, t, \hat{\mathbf{n}}_i) - \mu_{P,S}(\mathbf{r}, t))^2$, while $\mu_{P,S}(\mathbf{r}, t) = \frac{1}{N}\sum_{i=1}^{N} I_{P,S}(\mathbf{r}, t, \hat{\mathbf{n}}_i)$ is the mean of the specific intensities along the angular direction, and N is the number of angular directions. The quantity $\delta_{P,S}$ provides a measure of the variation of the P or S specific intensities along the angular

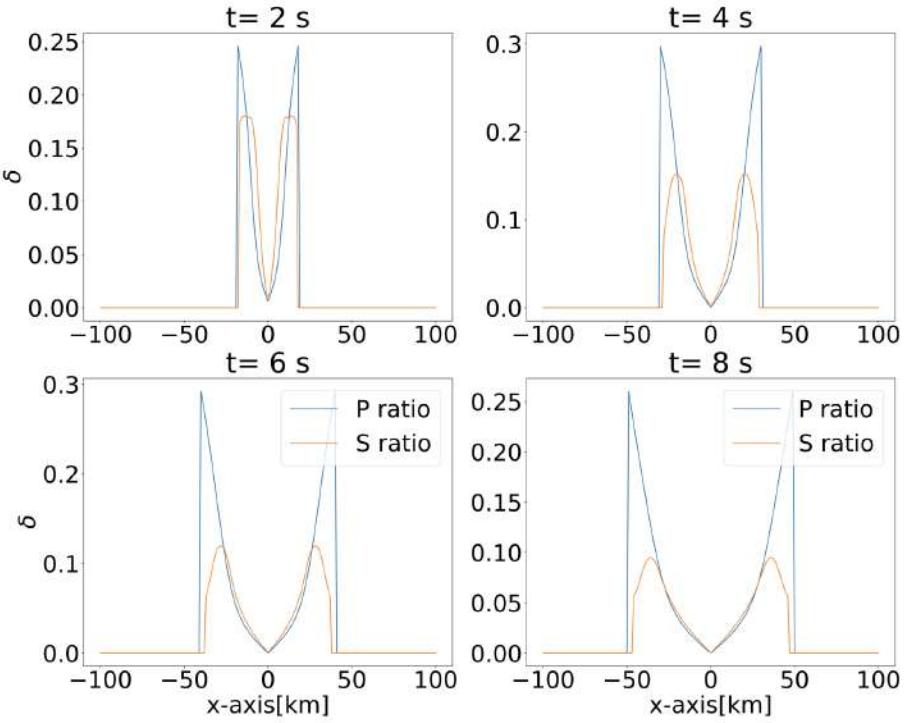

Figure 4.9 Same as Figure 4.8, but for a double couple.

directions as a function of space and time. We compute the standard deviation and the mean of the specific intensities along the N propagation directions for a fixed point in space and time. The constant $1/\sqrt{N}$ is included in equation 4.26 so that $0 \leq \delta_{P,S} \leq 1$. Consider an uni-directional intensity field $I_i = I_0 \delta_{i,1}$, where I_0 is the intensity along the only non-zero direction and $\delta_{i,j}$ is the Kronecher delta. The mean for this intensity field is $\mu = I_0/N$. The standard deviation for the same intensity field is $\sigma = I_0/\sqrt{N}$, which gives $\sigma/\mu = \sqrt{N}$. Hence, $\delta = 1$ for the most extreme case where all energy propagates in one direction. Thus, $\delta_{P,S} = 0$ when the intensity field (P, S) is fully equipartitioned in the angular sense (i.e., same specific intensity along all directions) and $\delta_{P,S} = 1$ when the intensity field is unidirectional (i.e., all of the specific intensities but one are equal to zero).

Figure 4.8 and Figure 4.9 show horizontal cross sections of the angular equipartitioning ratio $\delta_{P,S}$ at $y = 0$, of the explosive and double couple source simulations, respectively, for times $t = (2, 4, 6, 8)$ s. In the first simulation we release an explosive source with Gaussian spatial extent. Within this Gaussian we release the energy equally along all directions. Because of the nonzero gradient of the intensity field, part of the intensity propagates away from the initial Gaussian with a preferred propagation direction. As time progresses, this preferential propagation of the P wave becomes more evident, as the increase in the value of the peaks of δ_P in Figure 4.8 shows. This figure shows that δ_P takes the highest value at the location of

the ballistic wave. This value is significantly lower than one because of the spatial extent of the source (i.e., there is a preferrential direction but there is energy propagating in other directions). The value of δ_P decreases as one moves closer to the original source location centered at $x = 0$, where more scattering events have occurred relative to locations closer to the ballistic arrival. Some of this behavior also occurs with δ_S. However, at times $t = 4$ s and $t = 6$ s there are two characteristic peaks in the equipartitioning index for δ_S. The peak that is closer to the source (black arrow) corresponds to P energy within the initial Gaussian that over time converts to S energy. The peak that is further from the source (green arrow) corresponds to S energy at time t that propagates for most of the simulation (up to the current time t) as P energy. This peak has a travel time that is close to that of the ballistic P energy, and corresponds to single $P \rightarrow S$ scattering directly behind the ballistic P wave. Because of this scattering so close to the ballistic wave, the S energy still has some preferred directionality and produces a maximum in δ_S. Overall, the behavior of $\delta_{P,S}$ for the double couple simulation is similar to the one for the explosive simulation, with the exception of two differences. The first difference is that the highest value δ_P is slightly smaller than for the explosive source simulation. This is because this peak corresponds, not to P ballistic energy, but to $P \rightarrow P$ energy conversion early in the simulation. The second difference is that the values of δ_S are larger than for the first simulation. This is because for the explosive source, all S-energy is caused by scattering (i.e., the source only releases P-energy). This means that for the explosive source, the S-waves are excited over a range of propagation directions as a consequence of mode conversion (see the radiation pattern for $P \rightarrow S$ scattering in Figure 4.1). For the double couple source there is a ballistic S-wave (unlike for the explosive source), which propagates at every point away from the source, and it is thus directional (higher δ_S).

4.6 Discussion

We developed a time-stepping algorithm for solving the elastic RTE in expressions 4.1 and 4.2. Our approach is based on first splitting the total intensity into the direct and scattered contributions, and then re-writing the RTE for the scattered intensities as integral equations. With these integral equations we evolve the specific P and S specific intensities over time, under the assumption that the time-step $\Delta t \ll \tau_p, \tau_s$ so that we can capture the scattering of P and S waves. Here, we disregard the boundary conditions, under the assumption that for times less than the propagation time for a P-wave to reach the boundary, the boundary conditions do not influence the solution.

Contrary to many numerical developments, we resolve the angular distribution of both P and S energies. In our algorithm we discretize the angular integral and then handle the advection of P and S energy analytically at grid points and through interpolation within the numerical grid. This approach allows us to reduce dispersion errors as compared to standard numerical techniques such as finite

differences or discontinuous Galerkin finite element methods. This numerical dispersion may cause negative intensities which are non-physical. To test the algorithm, we solve the RTE for both an explosive and a double couple source. We first benchmark our results against the diffusive approximation. After many scattering events this approximation is valid for any initial condition. Figure 4.2 and Figure 4.3 show that, at late times, our solutions match the diffusive approximation. The double couple source simulation takes longer to reach a diffusive state, compared to the explosive source simulation, because of the radiation pattern of the double couple. Since the double couple does not release energy in all possible directions, it takes a longer time, compared to the pure P source, to redistribute the energy in all propagation directions. We test that the ratio between total S-wave energy to total P-wave energy converges to the global theoretical ratio $(\alpha_0/\beta_0)^2 = 2.25$ for the given choice of wave speeds. Figure 4.4 shows that our simulations do indeed converge to the theoretical ratio, with the double couple simulation reaching it faster since at $t = 0$ the energy is already distributed among the two available wave modes. We also compare our numerical simulations to analytical expressions for the total energies $Y^P(t)$ and $Y^S(t)$ to validate the energy evolution throughout the simulation range. Figure 4.5 shows this comparison. Overall, our numerical results for the global equipartitioning ratio, the diffusive approximation, and the total energy as a function of time, match the known analytical expressions given by γ_0^2, equation 4.30 and equation 4.31, respectively. There is a small mismatch in the total energy at late times (Figure 4.5), which we can reduce by decreasing the value of Δt. We studied the local behavior of the energy equilibration between wave modes. Figure 4.6 and Figure 4.7 show that it takes significantly longer for the local equipartitioning to converge to the theoretical ratio, compared to the global ratio. As explained above, this is because for the local ratio to be established, the intensity field must be locally equipartitioned.

In addition to modal equipartitioning, we study angular equipartitioning. We use the equipartitioning index to study the angular randomization of the intensity fields. We find that, in general, the level of equipartitioning of an intensity field is a function of space and time. Figure 4.8 and Figure 4.9 show that the ballistic arrival has the highest equipartitioning index and that the trailing scattered energy becomes equipartitioned over time but not at the same rate throughout the computational domain. If we define t_b as the arrival time of the ballistic wave, one can approximate the number of scattering events as $n \approx (t - t_b)/\tau_{av}$, where τ_{av} corresponds to a weighted average between the P (τ_p) and S (τ_s) scattering mean free times. For a given propagation time t, the larger t_b is (greater distance), the smaller n is. The algorithm that we introduce, together with our numerical simulations, allow us to study the propagation of P and S waves, and to investigate the local evolution of both modal and angular equipartitioning.

Our observations about the spatial and temporal dependence of equipartitioning have implications for the retrieval of Green's functions in seismic interferometry. The principle of equipartition is regarded as a

necessary, but not sufficient (Snieder et al., 2010), condition to retrieve Green's functions. Because the degree of equipartitioning varies locally, depending on the extent of scattering, this implies that the accuracy with which one retrieves the Green's functions is also a function of space and time (Weaver, 2010).

4.7 Acknowledgements

This work was supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at the Colorado School of Mines.

4.8 2-D Diffusive Approximation

For a 3-D infinite medium, assuming there is no intrinsic attenuation, the diffusion approximation to the elastic RTE reads (Sato et al., 2012)

$$E_{3d}(\mathbf{r},t) = \frac{W^P + W^S}{(4\pi D^*_{3d} t)^{3/2}} H(t) e^{-\frac{r^2}{4D^*_{3d}t}}, \tag{4.27}$$

where D^*_{3d} is the effective diffusion coefficient in 3-D, and W^P and W^S denote the P and S energy at $t = 0$, respectively. The diffusion coefficient is

$$D^*_{3d} = \frac{1}{3(1+2\gamma_0^3)}\left(\frac{\alpha_0}{g^{Pe}} + 2\gamma_0^3 \frac{\beta_0}{g^{Se}}\right), \tag{4.28}$$

where $\gamma_0 = \alpha_0/\beta_0$ is the ratio of mean P wave speed to mean S wave speed. g^{Pe} and g^{Se} are the P and S effective scattering coefficients. To obtain the diffusive approximation in 2-D we write the diffusion coefficient as

$$D^*_{2d} = \frac{1}{2(1+\gamma_0^2)}\left(\frac{\alpha_0}{g^{Pe}} + \gamma_0^2 \frac{\beta_0}{g^{Se}}\right), \tag{4.29}$$

where we replaced the 3-D equipartitioning ratio $2\gamma_0^3$ by its 2-D counterpart γ_0^2. We also replaced the factor 3 in the denominator of expression 4.28 by a factor 2, and replaced the exponent $3/2$ in the denominator of equation 4.27 by the exponent $2/2$, to account for the change from 3-D to 2-D (Snieder and van Wijk, 2015). The diffusive approximation in 2-D reads

$$E_{2d}(\mathbf{r},t) = \frac{W^P + W^S}{(4\pi D^*_{2d} t)} H(t) e^{-\frac{r^2}{4D^*_{2d}t}}. \tag{4.30}$$

4.9 Expressions for $Y^P(t)$ and $Y^S(t)$

Sato et al. (2012) derive the following expressions for the total P and S energies in 3-D

$$Y^P(t) = \frac{(g_0^{PS}\alpha_0 W^P - g_0^{SP}\beta_0 W^S)e^{-g_0^{PS}\alpha_0 t - g_0^{SP}\beta_0 t} + g_0^{SP}\beta_0(W^P + W^S)}{g_0^{PS}\alpha_0 + g_0^{SP}\beta_0},$$
$$Y^S(t) = \frac{(-g_0^{PS}\alpha_0 W^P + g_0^{SP}\beta_0 W^S)e^{-g_0^{PS}\alpha_0 t - g_0^{SP}\beta_0 t} + g_0^{PS}\alpha_0(W^P + W^S)}{g_0^{PS}\alpha_0 + g_0^{SP}\beta_0}. \tag{4.31}$$

To derive these expressions, Sato et al. (2012) make three assumptions. First, they assume that the source releases P energy as $W^P\delta(\mathbf{r})\delta(t)$ and S energy as $W^S\delta(\mathbf{r})\delta(t)$, where W^P and W^S are constants. Second, they assume that the angular dependence of the P and S specific intensities is small and almost isotropic. Third, they assume that the spatial gradient of both the P and S energy density vanishes at large distances from the source. If we define $W^P = \int \oint I_0^P(\mathbf{r},\theta)dxdy$ and $W^S = \int \oint I_0^S(\mathbf{r},\theta)d\theta dxdy$, and then assume that the gradient of both the P and S specific intensities vanishes at large distances from the source (without assuming that the P and S specific intensities are almost isotropic), the expressions in 4.31 remain valid. Expression 4.31 is valid in 2- and 3-D, with the scatteriing coefficients having a different definition depending on the dimension.

4.10 Conservation of energy

The choice of the parameters $A_{p,s}$ and $B_{p,s}$ in the linear system 5.14 determine the accuracy of the time stepping algorithm. To choose the values of $A_{p,s}$ and $B_{p,s}$ we impose energy conservation, so that the numerical algorithm conserves the energy as much as possible. We begin by re-defining expression 4.25 as

$$\begin{bmatrix}\vec{I}^P \\ \vec{I}^S\end{bmatrix} = \begin{bmatrix}\boldsymbol{\eta}^{PP} & \boldsymbol{\eta}^{SP} \\ \boldsymbol{\eta}^{PS} & \boldsymbol{\eta}^{SS}\end{bmatrix}\begin{bmatrix}\vec{W}^P \\ \vec{W}^S\end{bmatrix}. \tag{4.32}$$

We then write equation 4.32 as the following two equations

$$I^P(\mathbf{r},\hat{\mathbf{n}}_j,t) = \sum_{i=1}^{N}\eta_{ji}^{PP}W^P(\mathbf{r},\hat{\mathbf{n}}_i,t) + \sum_{i=1}^{N}\eta_{ji}^{SP}W^S(\mathbf{r},\hat{\mathbf{n}}_i,t) \tag{4.33}$$
$$I^S(\mathbf{r},\hat{\mathbf{n}}_j,t) = \sum_{i=1}^{N}\eta_{ji}^{PS}W^P(\mathbf{r},\hat{\mathbf{n}}_i,t) + \sum_{i=1}^{N}\eta_{ji}^{SS}W^S(\mathbf{r},\hat{\mathbf{n}}_i,t). \tag{4.34}$$

Adding equations 4.33 and 4.34 gives

$$I(\mathbf{r},\hat{\mathbf{n}}_j,t) = \sum_{i=1}^{N}(\eta_{ji}^{PP} + \eta_{ji}^{PS})W^P(\mathbf{r},\hat{\mathbf{n}}_i,t) + \sum_{i=1}^{N}(\eta_{ji}^{SP} + \eta_{ji}^{SS})W^S(\mathbf{r},\hat{\mathbf{n}}_i,t). \tag{4.35}$$

Summing both sides of equation 4.35 from $j = 1$ to N, and multiplying by $2\pi/N$ gives

$$I(\mathbf{r},t) = \frac{2\pi}{N}\sum_{j=1}^{N}(\eta_{ji}^{PP} + \eta_{ji}^{PS})\sum_{i=1}^{N}W^P(\mathbf{r},\hat{\mathbf{n}}_i,t) + \frac{2\pi}{N}\sum_{j=1}^{N}(\eta_{ji}^{SP} + \eta_{ji}^{SS})\sum_{i=1}^{N}W^S(\mathbf{r},\hat{\mathbf{n}}_i,t), \tag{4.36}$$

where we assumed that the terms $\sum_{j=1}^{N}\eta_{ji}^{PP}, \sum_{j=1}^{N}\eta_{ji}^{PS}, \sum_{j=1}^{N}\eta_{ji}^{SP}, \sum_{j=1}^{N}\eta_{ji}^{SS}$ are independent of i, for sufficiently small Δt. To show that this assumption is valid consider the block matrix $\boldsymbol{A}$ in the linear system 5.14 and its corresponding inverse $\boldsymbol{A}^{-1}$. As before, we write $\boldsymbol{A} = \boldsymbol{I} + \boldsymbol{K}$. Then, we apply the matrix expansion $(\boldsymbol{I} - \boldsymbol{K})^{-1} \approx \boldsymbol{I} + \boldsymbol{K}$, which is valid for $||\boldsymbol{K}|| \ll 1$. With this expansion the elements of $\boldsymbol{A}^{-1}$ are

$$\begin{aligned}
\eta_{ij}^{PP} &= \delta_{ij} + \alpha_0 B_p \frac{2\pi}{N}\Delta t g_{pp}(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_j);\ \eta_{ij}^{SP} = \beta_0 B_p \frac{2\pi}{N}\Delta t g_{sp}(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_j) \\
\eta_{ij}^{PS} &= \alpha_0 B_s \frac{2\pi}{N}\Delta t g_{ps}(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_j);\ \eta_{ij}^{SS} = \delta_{ij} + \beta_0 B_s \frac{2\pi}{N}\Delta t g_{ss}(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_j),
\end{aligned} \tag{4.37}$$

which gives $\sum_{j=1}^{N}(\eta_{ji}^{PP} + \eta_{ji}^{PS}) = 1 + \alpha_0 B_p \Delta t g_{pp}^0 + \alpha_0 B_s \Delta t g_{ps}^0$ and
$\sum_{j=1}^{N}(\eta_{ji}^{SP} + \eta_{ji}^{SS}) = 1 + \beta_0 B_p \Delta t g_{sp}^0 + \beta_0 B_s \Delta t g_{ss}^0$, justifying our assumption that the terms
$\sum_{j=1}^{N}\eta_{ji}^{PP}, \sum_{j=1}^{N}\eta_{ji}^{PS}, \sum_{j=1}^{N}\eta_{ji}^{SP}, \sum_{j=1}^{N}\eta_{ji}^{SS}$ are independent of i, for sufficiently small Δt. We re-write the definitions for W^P and W^S from equations 4.38 and 4.39 to include both the direct and scattered energy as

$$\begin{aligned}
\vec{W}_{tot}^{P}(\mathbf{r}, \hat{\mathbf{n}}, t) &= I^P(\mathbf{r} - \alpha_0 \Delta t \hat{\mathbf{n}}, \hat{\mathbf{n}}, t - \Delta t) e^{-\Delta t \alpha_0 (g_{pp}^0 + g_{ps}^0)} \\
&+ A_p \Delta t \frac{2\pi}{N} \sum_{\hat{\mathbf{n}}'} \Big[\alpha_0 g_{pp}(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I^P(\mathbf{r} - \alpha_0 \Delta t \hat{\mathbf{n}}, \hat{\mathbf{n}}', t - \Delta t) \\
&+ \beta_0 g_{sp}(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I^S(\mathbf{r} - \alpha_0 \Delta t \hat{\mathbf{n}}, \hat{\mathbf{n}}', t - \Delta t) \Big] e^{-\Delta t \alpha_0 (g_{pp}^0 + g_{ps}^0)},
\end{aligned} \tag{4.38}$$

$$\begin{aligned}
\vec{W}_{tot}^{S}(\mathbf{r}, \hat{\mathbf{n}}, t) &= I^S(\mathbf{r} - \beta_0 \Delta t \hat{\mathbf{n}}, \hat{\mathbf{n}}, t - \Delta t) e^{-\Delta t \beta_0 (g_{ss}^0 + g_{sp}^0)} \\
&+ A_s \Delta t \frac{2\pi}{N} \sum_{\hat{\mathbf{n}}'} \Big[\beta_0 g_{ss}(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I^S(\mathbf{r} - \beta_0 \Delta t \hat{\mathbf{n}}, \hat{\mathbf{n}}', t - \Delta t) \\
&+ \alpha_0 g_{ps}(\hat{\mathbf{n}}, \hat{\mathbf{n}}') I^P(\mathbf{r} - \beta_0 \Delta t \hat{\mathbf{n}}, \hat{\mathbf{n}}', t - \Delta t) \Big] e^{-\Delta t \beta_0 (g_{ss}^0 + g_{sp}^0)},
\end{aligned} \tag{4.39}$$

We Taylor-expand the specific intensities in equation 4.36 to first order around $(x, y, t - \Delta t)$, and integrate over space to obtain, with the help of the vanishing boundary conditions, the following system of equations

$$\begin{bmatrix} \eta_0^P(B_p, B_s)\alpha_0 \Delta t g_{pp}^0 e^{-\Delta t/\tau_p} & \eta_0^S(B_p, B_s)\alpha_0 \Delta t g_{ps}^0 e^{-\Delta t/\tau_s} \\ \eta_0^P(B_p, B_s)\beta_0 \Delta t g_{sp}^0 e^{-\Delta t/\tau_p} & \eta_0^S(B_p, B_s)\beta_0 \Delta t g_{ss}^0 e^{-\Delta t/\tau_s} \end{bmatrix} \begin{bmatrix} A_p \\ A_s \end{bmatrix} = \begin{bmatrix} 1 - \eta_0^P e^{-\Delta t/\tau_p} \\ 1 - \eta_0^S e^{-\Delta t/\tau_s} \end{bmatrix}, \tag{4.40}$$

where $\eta_0^P(B_p, B_s) = \sum_{j=1}^{N}(\eta_{ji}^{PP}(B_p, B_s) + \eta_{ji}^{PS}(B_p, B_s)),\ \eta_0^S(B_p, B_s) = \sum_{j=1}^{N}(\eta_{ji}^{SP} + \eta_{ji}^{SS}), \tau_p =$
$1/(\alpha_0(g_{pp}^0 + g_{ps}^0)),\ \tau_s = 1/(\beta_0(g_{sp}^0 + g_{ss}^0))$. Notice that the matrix in expression 4.40 is a function of B_p and B_s. Using $A_p + B_p = 1$ and $A_s + B_s = 1$, the solution vector in equation 4.40 is $[A_p, A_s] = [1/2, 1/2]$; hence we obtain $[B_p, B_s] = [A_p, A_s] = [1/2, 1/2]$, which corresponds to the trapezoidal rule.

CHAPTER 5

APPLICATION OF PHYSICS INFORMED NEURAL NETWORKS TO THE 1-D RADIATIVE TRANSFER EQUATIONS

To be submitted to the European Journal of Physics.

Manuel Jaimes-Caballero[9], Roel Snieder[10], and Samy Wu Fung[11].

5.1 Abstract

We solve the 1-D radiative transfer equations, subject to initial and boundary conditions, using physics informed neural networks, a deep learning technique. This technique is based on training a neural network to approximate the mapping between input (i.e., position, time, and direction of propagation in this chapter) and output (i.e., solution to the 1-D radiative transfer equations in this chapter). We consider three numerical examples of increasing complexity, and test two different physics informed neural network approaches: a soft-constraints, and a mixed-constraints approach. In the first approach, the penalty function that we use to train the neural network is the weighted sum of the radiative transfer equation, initial condition, and boundary condition losses. In the second approach, we incorporate the initial condition exactly within the neural network architecture, and construct a penalty function which is the weighted sum of the radiative transfer equation and boundary condition losses. We find the second approach to work better than the first, because we satisfy the initial condition exactly, and we only need to include a regularization term for the boundary condition. The solutions that we obtain using the mixed-constraints approach resemble the reference analytical/numerical solutions, with an error of less than 2% for the most complicated numerical example. The mixed-constraints approach appears to be a promising tool to model the radiative transfer equations in complicated numerical settings.

5.2 Introduction

Machine Learning (ML), a subset of Artifical Intelligence (AI), has revolutionized many research fields in the last two decades. Some common applications of ML include speech and pattern recognition (Nassif et al., 2019; Paolanti and Frontoni, 2020), natural language processing (Otter et al., 2021), autonomous vehicles (Aradi, 2022), facial recognition (Zong and Huang, 2011), robotics (Alsamhi et al., 2020), finance trading (Chen et al., 2020), smart assistants (Polyakov et al., 2018), and protein folding (Jo et al., 2015).

[9]Graduate student, primary researcher, and author at Department of Geophysics, Colorado School of Mines.
[10]Supervisor at Department of Geophysics, Colorado School of Mines.
[11]Assistant Professor at Department of Applied Math and Statistics, Colorado School of Mines.

With ML one aims to learn the relationship between input and output data pairs. *Artificial Neural Networks* (ANNs), which were inspired by biological neural networks (Mishra and Srivastava, 2014), are particularly useful to discover such mappings. ANNs consist of a collection of nodes, denoted *neurons*, connected via a series of layers. ANNs can approximate any continuous function due to the Universal Approximation theorem (Baker and Patil, 1998), and are therefore fundamental in ML techniques. To inform ANNs about the desired mapping one constructs a *penalty function* which quantifies the difference between the current and desired output. One then updates (i.e., the training algorithm) the ANNs until the penalty function is below some threshold, and the mapping is adequately approximated.

In addition to the areas mentioned above, ML techniques have found use in the physical sciences. One application that has become important is the use of ML techniques to describe complicated physical systems. While there are many ways to describe such systems (e.g., pattern recognition, forward modeling and imaging algorithms), we limit our discussion to the use of ML techniques in forward modeling (e.g., solving an ordinary or partial differential equation). Generally speaking, there are three types of ML techniques that are used for forward modeling; (1) data-driven, (2) physics-driven (Raissi et al., 2019), and (3) mixed techniques. In the data-driven approach, one penalizes the output of the neural network with data that are, either simulated through standard numerical methods (e.g., finite difference, finite volume, finite element, Monte Carlo based techniques), or collected in an experimental setting (e.g., fluid flow in a pipe). There are two major disadvantages to this approach. First, the numerical errors present in the training data are propagated to the neural network predictions. Second, the computation and/or collection of the training data may be expensive and only justified if one repeatedly uses the neural network output. Of the physics-driven approaches, we limit our discussion to Physics Informed Neural Networks (PINNs). The main idea behind PINNs is to train the neural network such that its output satisfies a given physical law (Raissi et al., 2019), rather than matching simulated/collected data as it is the case in the data-driven approach. The PINNs approach is advantageous because no training data is needed, and the neural network is trained to satisfy the given physical law as closely as possible.

In this chapter we use PINNs to solve the 1-D Radiative Transfer Equations (RTE). The RTE are a coupled system of integro-differential equations that describe the spatio-temporal-angular evolution of the wave intensity in a scattering medium. To solve for the wave intensity one must know the scattering mean free path l, the angular dependence of scattering, and the speed of energy propagation v. We only address the 1-D version of the RTE because the application of PINNs to 2-D and 3-D is not yet computationally feasible due to angular coupling (Mishra and Molinaro, 2021). In 1-D the RTE are a coupled system of two partial differential equations describing the propagation of right and left-going energy in a scattering medium.

We focus on the RTE since it has been employed in several scientific fields, including astrophysics (Steinacker et al., 2002; Narayanan et al., 2021; Wolf, 2003); atmospheric sciences (Evans and Stephens, 1995; Aumann et al., 2018; Manners et al., 2009); optics (Klose et al., 2002; Abdoulaev, 2003; Ren et al., 2004; Yodh and Chance, 1995); acoustics (Quijano and Zurk, 2009; Ostashev et al., 2017; Reboul et al., 2005); and geophysics (Prata, 1989; Francis et al., 2012; Lee et al., 2014; Hofmeister, 2005; Margerin et al., 2016; Duran et al., 2020; Dinther et al., 2021; Obermann et al., 2016). Despite their usefulness in describing energy transport, the RTE are complicated and numerically demanding due to their dependence on not only space and time, but also on the directions of wave propagation. In 1-D the RTE solution depends on 3 variables (time, space, and direction of propagation), in 2-D the solution depends on four variables (time, two space variables, and direction of propagation), and in 3-D it depends on six variables (time, three space variables, and two angular propagation directions).

Numerical techniques used to solve the RTE include the discontinuous Galerkin method (Clarke et al., 2019; Han et al., 2010), Markov chain Monte Carlo techniques (Iwabuchi, 2006; Xu et al., 2011; Camps and Baes, 2018; Noebauer and Sim, 2019; Przybilla and Korn, 2008; Yoshimoto, 2000), finite difference techniques (Klose and Hielscher, 1999), and wave equation modeling whereby one exploits the connection between the acoustic wave equation and the scalar RTE (Przybilla et al., 2006; Kanu and Snieder, 2015; Snieder et al., 2019; Duran et al., 2020). In addition to these numerical techniques, there have been analytical approximations to RTE such as assuming diffusive wave propagation which is only valid at times much larger than the transport mean free time (Rossetto et al., 2011; Planès et al., 2014), assuming a point-like, isotropic, impulsive source of intensity in a statistically homogeneous medium (Margerin et al., 2016), expanding the intensity and scattering function into a finite sum of Legendre polynomial and then solving a finite system of equations for the unknown coefficients appearing in the truncated expansion (Roberge, 1983), decomposing the specific intensities into a sum of partial intensities and then solving the RTE for each partial intensity by assuming that scattering is angle-independent and that the source is isotropic (Paasschens, 1997), assuming a steady-state intensity field (Fan et al., 2019; Le Hardy et al., 2016), or assuming azimuthal symmetry (Baes and Dejonghe, 2001; de Abreu, 2004).

Because of the theoretical and numerical attention that the RTE have received, we investigate the capability of PINNs for the forward modeling of the RTE in 1-D. The structure of this chapter is as follows: In section 5.3, we describe neural networks. In section 5.4, we describe the 1-D RTE, how we can use neural networks to solve these equations, the training algorithm, as well as soft and hard-constraints approaches of PINNs. The soft-constraints approach assesses the mismatch of the initial and boundary condition through a penalty function; whereas the hard-constraints approach incorporates the initial and boundary condition within the neural network architecture, thereby satisfying the initial and boundary

condition exactly. Since satisfying both the initial and boundary condition exactly is complicated, we show a mixed-constraints approach where the initial condition is handled exactly, and the boundary condition mismatch is incorporated in the penalty function.

In section 5.5 we show three numerical experiments where we solve the 1-D RTE. For the numerical experiment in section 5.5.1 there is a reference analytical solution. For the numerical experiments in sections 5.5.2 and 5.5.3, we define the reference solution to be the numerical solution computed with the time-stepping algorithm developed by Jaimes and Snieder (2023). For all of the experiments we compare the solution obtained with the soft- and mixed-constraints PINNs approach to the reference solution. In section 5.6 we discuss the capability of the PINNs algorithm, and future research directions. In section 5.8 we provide an overview of the reference solution.

5.3 Description of Neural Networks

ANNs, as described before, consist of a collection of neurons or nodes, connected through a series of layers. While there are many types of neural networks, in this chapter we focus on sequential or *forward-feed* neural networks. The term *forward-feed* neural network refers to the architecture in which the connections in the network are only made in the forward direction, starting with the input layer, traversing through interconnected layers of neurons, and finalizing with an output layer, which typically corresponds to the desired output. Figure 5.1 shows a representation of this type of network.

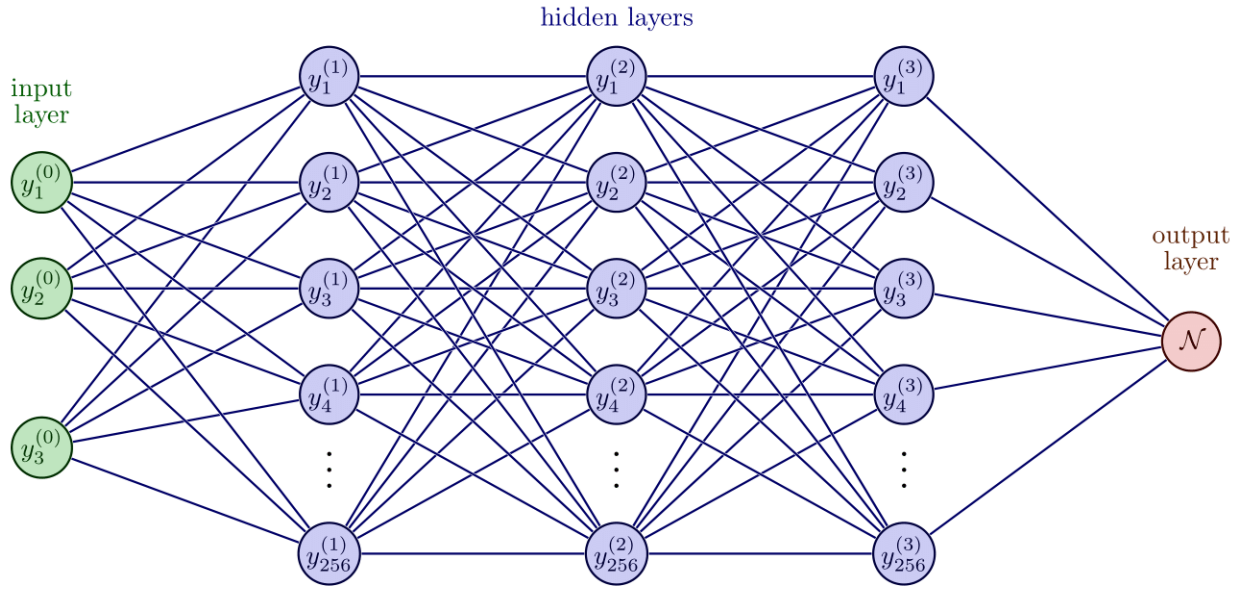

Figure 5.1 Representation of fully connected layer. The input layer contains the nodes $(y_1^{(0)}, y_2^{(0)}, y_3^{(0)})$. The term $y_n^{(j)}$, with $j \in \{1, 2, 3\}$ and $n \in \{1, \ldots, 256\}$, represents the n^{th} neuron at the j^{th} hidden layer.

The input nodes $(y_1^{(0)}, y_2^{(0)}, y_3^{(0)})$, or more generally $(\mathbf{y_1}^{(0)}, \mathbf{y_2}^{(0)}, \mathbf{y_3}^{(0)})$, are denoted as the *features*. One can think of these features as the number of variables in a given partial differential equation. In this section we define the input vector as $\mathbf{y} = (\mathbf{y_1}^{(0)}, \mathbf{y_2}^{(0)}, \mathbf{y_3}^{(0)})$. The forward-feed neural network transforms the vector $\mathbf{y}$, through layers of neurons denoted as *hidden layers*, into an output denoted by $\mathcal{N}$. In the context of the 1-D RTE, the input vector is $\mathbf{y} = [\mathbf{x}, \theta, \mathbf{t}]$. The entries of the vectors $\mathbf{x}$ and $\mathbf{t}$ are the spatial and temporal coordinates, respectively, on which we evaluate the neural network, and consist of discrete points in a space-time computational domain. In section 5.4.2 we provide specific details about how we choose these points. The entries of the vector θ are assigned a single value of either 0 (for a right-going wave) or π (for a left-going wave). In this chapter, we assume that the velocity and scattering models are fixed. The output $\mathcal{N}$ corresponds to the approximate solution to the 1-D RTE, the intensity of the waves at $(\mathbf{x}, \theta, \mathbf{t})$.

The term $y_n^{(j)}$ represent the n^{th} node at the j^{th} layer. Since $j = 1$ corresponds to the first hidden layer, the term $y_n^{(0)}$ corresponds to the n node of the input layer. For $y_n^{(j)}$, the n^{th} node in layer j, one can recursively define the output at the node y_n^{j+1} as

$$y_n^{(j+1)} = \sigma_j\Big(\sum_{m=1}^{N_j} W_m^{(j)} y_m^{(j)} + b_n^j\Big), \tag{5.1}$$

where N_j is the number of nodes in layer j, $W_m^{(j)}$ is the weight of the connection between nodes $y_m^{(j)}$ and $y_n^{(j+1)}$, b_n^j is the bias at node y_n^j, and σ_j is the activation function at layer j. With expression 5.1 we propagate the input data through the neural network. The weights and biases are referred to as the *tunable weights*, $\boldsymbol{\omega}$. Here, $\boldsymbol{\omega}$ contains all the weights and biases of the neural network as a whole. To achieve the desired output one strategically modifies $\boldsymbol{\omega}$ until the value of the penalty function achieves some threshold (i.e., an optimization problem). In the context of PINNs, this tuning corresponds to finding optimal $\boldsymbol{\omega}$ so that the output of the neural network satisfies the physical law of interest, given some constraints (e.g., initial and boundary conditions). These constraints can be either analytical (e.g., some given function that describes initial and boundary conditions) or numerical (e.g., data recorded at the boundary of a scattering medium). The term σ_j denotes an element-wise non-linear activation function at layer j (Goodfellow et al., 2016). Typical activation functions include the hyperbolic tangent function, the sigmoid function and the rectified linear unit (ReLU) function (Apicella et al., 2021). In this chapter we use the hyperbolic tangent function in all numerical examples. Different activation functions may be needed depending on the problem that one wants to solve. The non-linear activation function helps the neural network solve complex tasks.

5.4 The Deep Learning Approach to RTE

The RTE are a system of coupled integro-differential equations which describe the distribution of energy in a scattering medium as a function of space, time, and direction of wave propagation $\hat{\mathbf{n}}$

(Chandrasekhar, 1960; Turner and Weaver, 1994; Ryzhik et al., 1996). For late times, the waves scatter many times, and as a result, the wave propagation is almost independent of direction. Thus, the equation of radiative transfer leads to diffusive energy transport (van Rossum and Nieuwenhuizen, 1999). The RTE accurately describe energy transport at both early and late times, as well as the transition from ballistic wave propagation to weak scattering to strong multiple scattering (Paasschens, 1997). For a thorough description of the radiative transfer (transport) equations we refer the reader to Ryzhik et al. (1996).

Assuming a uniform speed of transport, and ignoring intrinsic attenuation, the 1-D radiative transfer equation for an intensity field reads (Chandrasekhar, 1960)

$$\underbrace{\frac{\partial I(x,\theta,t)}{\partial t} + v\cos\theta\frac{\partial I(x,\theta,t)}{\partial x}}_{\text{ADVECTION}} = \underbrace{-\frac{1}{\tau_s}I(x,\theta,t)}_{\text{LOSS}} + \underbrace{\frac{1}{\tau_s}\Big[FI(x,\theta,t) + BI(x,\theta+\pi,t)\Big]}_{\text{GAIN}}, \tag{5.2}$$

where $\theta = 0$ for right propagating waves or $\theta = \pi$ for left propagating waves. F and B refer to the forward and backward scattering coefficients, respectively. For a right-propagating wave, the term F quantifies the amount of energy that continues propagating towards the right, and B the amount of energy that scatters back and begins to propagate towards the left direction. Energy conservation implies that $F + B = 1$.

In equation 5.2 the variable $I(x,\theta,t)$ is the intensity of waves at location x and time t propagating towards the left ($\theta = \pi$) and towards the right ($\theta = 0$). In the literature of radiative transfer the variable $I(x,\theta,t)$ is referred to as the specific intensity (Chandrasekhar, 1960). The advection of the energy, propagating with a wave velocity v, is described by the term $\partial I(x,\theta,t)/\partial t + v\cos\theta\partial I(x,\theta,t)/\partial x$. The parameter τ_s is the scattering mean free time, which describes the time scale at which scattering occurs. The term $-I(x,\theta,t)/\tau_s$ accounts for the energy that is lost due to back reflection. The term $(FI(x,\theta,t) + BI(x,\theta+\pi,t))/\tau_s$ describes the gain due to energy scattered from the forward and backward directions. Equation 5.2 can be understood as a system of two coupled advection partial differential equations with a damping term $-I(x,\theta,t)/\tau_s$, and a source term $(FI(x,\theta,t) + BI(x,\theta+\pi,t))/\tau_s$. The similarity between the RTE and the damped advection equation is of importance to understand the meaning of the analytical solution in the first numerical example, which we demonstrate in section 5.5.1, and also to understand the spatio-temporal evolution of the right- and left-going wave intensities. Note that for the initial condition $u(x,t=0) = u_0(x)$, the solution to the 1-D damped advection equation $\partial_t u(x,t) + v\partial_x u(x,t) + qu(x,t) = 0$, is $u(x,t) = u_0(x-vt)e^{-qt}$.

5.4.1 Formulation of the Deep Learning Algorithm

Our goal is to solve the 1-D RTE, subject to the initial condition $I_0(x,\theta)$, at an initial time $t = 0$, with boundary condition $f(x,\theta,t)$ for $t > 0$

$$
\begin{aligned}
&\partial_t I(x,\theta,t) + v\cos\theta\partial_x I(x,\theta,t) = -I(x,\theta,t)/\tau_s + \Big[FI(x,\theta,t) + BI(x,\theta+\pi,t)\Big]/\tau_s,\\
&I(x,\theta,t=0) = I_0(x,\theta),\\
&I(x=x_b^l,\theta,t) = f(x_b^l,\theta,t),\ I(x=x_b^r,\theta,t) = f(x_b^r,\theta,t),\ (x,t)\in[x_b^l,x_b^r]\times(0,T].
\end{aligned}
\tag{5.3}
$$

where x_b^r and x_b^l denote the right and left boundary locations, respectively. The two equalities, in the last line of system 5.3, are the left and right boundary conditions, respectively. When the boundary points $(x_b^r, x_b^l) \to (-\infty, \infty)$, and the velocity and scattering properties are constant, equation 5.3 has an analytical solution. If one defines the initial condition by a general directional point source $[I_0(x,\theta=0), I_0(x,\theta=\pi)] = \delta(x)\delta(t)[1+c, 1-c]$, with c a parameter that controls the directionality of the source (i.e., $c=0$ for isotropic, $c=1$ for a right-propagating wave, and $c=-1$ for a left-propagating wave), then the Green's function for the total intensity is given by Haney et al. (2005) as

$$
\begin{aligned}
I_t(x,t) = \underbrace{\exp\left(-\frac{Bvt}{l_s}\right)}_{\text{DAMPING}}\Bigg[&\underbrace{(1-c)\delta(vt+x) + (1+c)\delta(vt-x)}_{\text{ADVECTION}}\\
&+\underbrace{\frac{B}{l_s}H(vt-|x|)\Bigg[I_0\left(\frac{B\sqrt{v^2t^2-x^2}}{l_s}\right)}_{\text{SCATTERING}}\\
&+\underbrace{\frac{vt+cx}{\sqrt{v^2t^2-x^2}}I_1\left(\frac{B\sqrt{v^2t^2-x^2}}{l_s}\right)\Bigg]}_{\text{SCATTERING}}\Bigg],
\end{aligned}
\tag{5.4}
$$

where I_0 and I_1 are the modified Bessel functions of the zeroth and first orders, respectively. H is the Heaviside step function, and l_s is the scattering mean free length, defined as $l_s = v\tau_s$. Solution 5.4 can be understood as an extension of the solution to the 1-D damped advection equation. The δ terms describe the advection of the initial condition, the exponential in the front accounts for the decay of the energy as scattering occurs, and the modified Bessel functions account for the scattering gain. Even for this very simple 1-D example with uniform v and τ_s and no boundary conditions, the total intensity (see expression 5.4) has a complicated form. The scattering behavior for this simplest problem is given by modified Bessel functions. Deep learning may be able to learn and model this complicated scattering behavior, when there are no analytical solutions.

The main idea behind PINNs is to approximate the solution of equation 5.3, $I(x,\theta,t)$, with a neural network $\mathcal{N}_{\boldsymbol{\omega}}(x,\theta,t)$, with trainable weights $\boldsymbol{\omega}$. The subscript $\boldsymbol{\omega}$ indicates that the network depends on these weights. Replacing the solution $I(x,\theta,t)$ by the neural network approximation $\mathcal{N}_{\boldsymbol{\omega}}(x,\theta,t)$ gives

$$\partial_t \mathcal{N}_{\boldsymbol{\omega}}(x,\theta,t) + v\cos\theta\partial_x \mathcal{N}_{\boldsymbol{\omega}}(x,\theta,t) = -\mathcal{N}_{\boldsymbol{\omega}}(x,\theta,t)/\tau_s + \Big[F\mathcal{N}_{\boldsymbol{\omega}}(x,\theta,t) + B\mathcal{N}_{\boldsymbol{\omega}}(x,\theta+\pi,t)\Big]/\tau_s,$$
$$\mathcal{N}_{\boldsymbol{\omega}}(x,\theta,t=0) = I_0(x,\theta),$$
$$\mathcal{N}_{\boldsymbol{\omega}}(x=x_b^l,\theta,t) = f(x_b^l,\theta,t),\ \mathcal{N}_{\boldsymbol{\omega}}(x=x_b^r,\theta,t) = f(x_b^r,\theta,t),\ (x,t)\in[x_b^l,x_b^r]\times(0,T]. \tag{5.5}$$

Approximating the solution $I(\mathbf{x},\theta,\mathbf{t})$ to the system 5.3 using a neural network $\mathcal{N}_{\boldsymbol{\omega}}(\mathbf{x},\theta,\mathbf{t})$ is convenient for two reasons. First, contrary to standard numerical approaches (including the one that we use as a reference in this chapter), we can evaluate the solution $\mathcal{N}_{\boldsymbol{\omega}}(\mathbf{x},\theta,\mathbf{t})$ at any point within the computational domain in which we train the neural network, rather than at fixed points. Since the goal of the neural network is to learn the mapping between the input points $(\mathbf{x},\theta,\mathbf{t})$ and the output $\mathcal{N}$ representing the intensity of the waves, rather than evolving the numerical solution at fixed points (as it is done in finite difference methods), we are free to choose any values for $(\mathbf{x},\theta,\mathbf{t})$ which are within the range of values that we use to train the network. However, the network would produce poor results outside these ranges. Second, we can evaluate the derivatives in system 5.5 exactly using automatic differentiation, a computational technique based on the chain-rule for evaluating the derivatives of a function with respect to its input variables. Automatic differentiation relies on the fact that all computer programs use a finite set of elementary operations. Thus, any function that is computed by such programs, must be a composition of elementary functions/operations. Since the partial derivatives of the simple elementary functions are known, one can compute the derivative of the function of interest (i.e., the intensity of the waves at $(\mathbf{x},\theta,\mathbf{t})$) via the chain rule (Verma, 2000).

We want to determine the neural network $\mathcal{N}_{\boldsymbol{\omega}}$, which takes $(\mathbf{x},\theta,\mathbf{t})$ as its input, and approximates the solution to equation 5.3 as its output. To determine this mapping we need to tell the neural network (i.e., the optimization problem and training algorithm) which physical law, initial condition, and boundary conditions to satisfy (i.e., imposing equation 5.3). In the next section we discuss the training algorithm and the neural network architecture.

5.4.2 Training Algorithm and Neural Network Architecture

To learn the approximate mapping, we need to inform the neural network about the desired output. Hence, we need to define a penalty function $J(\boldsymbol{\omega})$ to guide the optimization of the tunable weights $\boldsymbol{\omega}$. There are different approaches to define the penalty function $J(\boldsymbol{\omega})$ in the context of PINNs. The simplest approach, and the earliest mentioned in the literature, is that of Raissi et al. (2019), denoted as the *soft-constraints* approach. With this approach, one defines the penalty function as the weighted sum of the individual penalty functions for the partial differential equation (PDE), the initial condition and/or the

boundary condition.

To define the penalty function via the soft-constraints approach we first define the PDE loss using the residual of equation 5.5 as

$$L_{\theta_i}(x_j,t_j;\boldsymbol{\omega}) = \partial_t \mathcal{N}_{\boldsymbol{\omega}}(x_j,\theta_i,t_j) + v\cos\theta_i \partial_x \mathcal{N}_{\boldsymbol{\omega}}(x_j,\theta_i,t_j) + \mathcal{N}_{\boldsymbol{\omega}}(x_j,\theta_i,t_j)/\tau_s \\ - [F\mathcal{N}_{\boldsymbol{\omega}}(x_j,\theta_i,t_j) + B\mathcal{N}_{\boldsymbol{\omega}}(x_j,\theta_i+\pi,t_j)]/\tau_s, \tag{5.6}$$

where the subscript $\theta_1 = 0$ (right-going waves) or $\theta_2 = \pi$ (left-going waves). The terms x_j and t_j denote the j^{th} spatial and temporal coordinates, respectively, on which we evaluate the PDE residual. We write the penalty function for the soft-constraints approach as

$$\begin{aligned} \mathrm{J}_{\mathrm{soft}}(\boldsymbol{\omega}) = \frac{1}{2N_{pde}} \sum_{i=1}^{2}\sum_{j=1}^{N_{pde}} |L_{\theta_i}(x_j,t_j;\boldsymbol{\omega})|^2 &+ \frac{\lambda_{ic}}{2N_{ic}} \sum_{i=1}^{2}\sum_{j=1}^{N_{ic}} |\mathcal{N}_{\boldsymbol{\omega}}(x_j,\theta_i,t=0) - I_0(x_j,\theta_i)|^2 \\ &+ \frac{\lambda_{bc}}{2N_{bc}^l} \sum_{i=1}^{2}\sum_{j=1}^{N_{bc}^l} |\mathcal{N}_{\boldsymbol{\omega}}(x_j = x_b^l,\theta_i,t) - f(x = x_b^l,\theta_i,t)|^2 \\ &+ \frac{\lambda_{bc}}{2N_{bc}^r} \sum_{i=1}^{2}\sum_{j=1}^{N_{bc}^r} |\mathcal{N}_{\boldsymbol{\omega}}(x_j = x_b^r,\theta_i,t) - f(x = x_b^r,\theta_i,t)|^2, \end{aligned} \tag{5.7}$$

where the first double sum denotes the PDE loss, the second double sum denotes the initial condition loss, and the third and fourth double sums denote the left and boundary condition losses, respectively. The terms λ_{ic} and λ_{bc} denote the regularization parameters for the initial and boundary condition losses, respectively. Notice that to avoid biasing the values of the penalty function 5.7 we use the same regularization term λ_{bc} for the left and right boundary conditions. The constants $N_{pde}, N_{ic}, N_{bc}^l$, and N_{bc}^r denote the number of sampling points for the PDE, initial condition, left boundary condition, and right boundary condition losses, respectively. During the training process, as the value of the penalty function is updated, these sampling points are referred to as the *training points*, the data set on which we train the neural network. In this chapter we refer to the penalty function 5.7 as the soft-constraints penalty function. The main disadvantage of the soft-constraints approach is that finding the proper weighting of the different losses is challenging. For instance, one could choose a high value of λ_{ic}, with which the neural network output satisfies the initial condition, but that does not satisfy the PDE. The same idea applies to choosing the correct value of λ_{bc}. Methods such as the L-curve or generalized cross-validation (Hansen, 1992) can be used to determine the optimal parameters λ_{ic} and λ_{bc}. However, to retrieve these parameters, one would have to solve the system 5.5 multiple times, which is computationally demanding and time-consuming. For the numerical experiments in this chapter we choose λ_{ic} values which range over two orders of magnitude $(0.1, 1, 10)$, and set $\lambda_{bc} = 10$ for the numerical simulations in sections 5.5.2 and 5.5.3.

We find this value of λ_{bc} to work reasonable well for the numerical experiments in this chapter.

Due to the cost of tuning the regularization parameters in the penalty function 5.7, a new type of PINNs has been devised (Lu et al., 2021), denoted in this chapter as the *hard-constraints* approach. In this approach one incorporates the initial and boundary conditions as part of the neural network architecture. The idea with this approach is to apply one or more extra operations to the output of the neural network such that after these operations the initial and/or boundary conditions are satisfied exactly. This approach gives a penalty function which only contains the PDE loss

$$\mathrm{J}_{\mathrm{hard}}(\boldsymbol{\omega}) = \frac{1}{2N_{pde}} \sum_{i=1}^{2} \sum_{j=1}^{N_{pde}} |L_{\theta_i}(x_j, t_j; \boldsymbol{\omega})|^2, \tag{5.8}$$

where the output $\mathcal{N}_{\boldsymbol{\omega}}$, that is used in the PDE loss function 5.6, is modified (through some operations) to satisfy the initial and boundary conditions exactly. In practice, satisfying both initial and boundary conditions exactly makes the training complicated. In this chapter we use a combination of soft and hard-constraints approach. We encode the initial condition in the architecture as follows. First, define an auxiliary neural network output $\mathcal{N}^{\mathrm{aux}}_{\boldsymbol{\omega}}(x, t, \theta)$. Then, define the output $\mathcal{N}_{\boldsymbol{\omega}}(x, t, \theta)$ as

$$\mathcal{N}_{\boldsymbol{\omega}}(x, t, \theta) = \left[(1 - t/T) I_0(x, \theta) + \frac{t}{T} \mathcal{N}^{\mathrm{aux}}_{\boldsymbol{\omega}}(x, t, \theta) \right] e^{-\gamma t/T}, \tag{5.9}$$

where T indicates the maximum simulation time for which we evolve the intensity, and γ is a trainable parameter of size 1×1. At time $t = 0$, equation 5.9 reduces to $\mathcal{N}_{\boldsymbol{\omega}}(x, t, \theta) = I_0(x, \theta)$, thereby automatically satisfying the initial condition of choice.

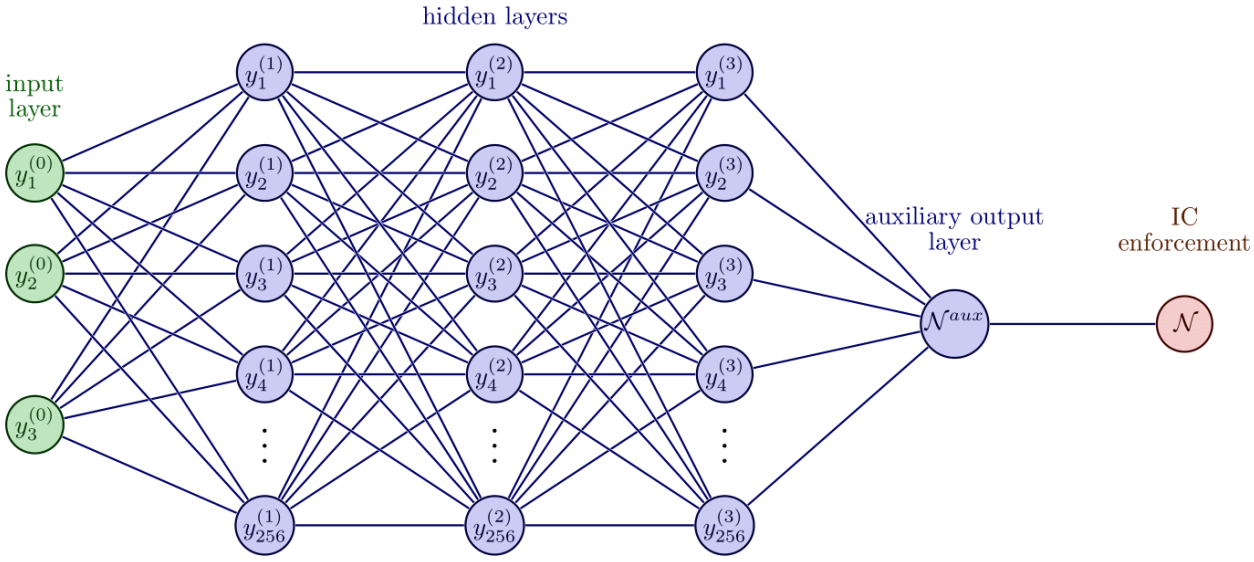

Figure 5.2 Representation of initial condition (IC) enforcement. This neural network is an extension of the neural network in Figure 5.1 with the extra operation given by equation 5.9.

Figure 5.2 shows the modification of the neural network output, using expression 5.9, to enforce the initial condition exactly. This extra operation is simple to implement with a single line of code, and can be changed accordingly depending on the type of initial or boundary condition that one aims to satisfy. We train the neural network so that the output $\mathcal{N}_{\boldsymbol{\omega}}$ solves the RTE; the output $\mathcal{N}^{\text{aux}}_{\boldsymbol{\omega}}(x,t,\theta)$ serves as an auxiliary output which we are not interested in individually retrieving, since it satisfies a PDE which is different to the RTE. We define a new penalty function

$$\mathrm{J}_{\text{mixed}}(\boldsymbol{\omega}) = \frac{1}{2N_{pde}}\sum_{i=1}^{2}\sum_{j=1}^{N_{pde}} |L_{\theta_i}(x_j,t_j;\boldsymbol{\omega})|^2 + \frac{\lambda_{bc}}{2N^l_{bc}}\sum_{i=1}^{2}\sum_{j=1}^{N^l_{bc}} |\mathcal{N}_{\boldsymbol{\omega}}(x_j = x^l_b,\theta_i,t) - f(x = x^l_b,\theta_i,t)|^2 + \frac{\lambda_{bc}}{2N^r_{bc}}\sum_{i=1}^{2}\sum_{j=1}^{N^r_{bc}} |\mathcal{N}_{\boldsymbol{\omega}}(x_j = x^r_b,\theta_i,t) - f(x = x^r_b,\theta_i,t)|^2, \quad (5.10)$$

where $\mathcal{N}_{\boldsymbol{\omega}}(x,t,\theta)$ is computed using expression 5.9. In equation 5.10 the output $\mathcal{N}_{\boldsymbol{\omega}}(x,t,\theta)$ (given by the linear combination in equation 5.9), must satisfy the PDE and boundary condition, and is thus the solution to the system 5.5 that we seek to solve. We denote the penalty function 5.10 as the *mixed-constraints* penalty function.

We optimize the tunable parameters $\boldsymbol{\omega}$ by minimizing either the soft-constraints penalty function 5.7 or the mixed-constraints penalty function 5.10. We solve this minimization problem with the ADAM optimizer, a first-order gradient-based optimizer of stochastic objective functions (Kingma and Ba, 2014). This optimizer is based on stochastic gradient descent, an optimization algorithm with which one updates the tunable parameters $\boldsymbol{\omega}$ using a subset of the training points, contrary to gradient descent which requires evaluating all the training points before a paramater update is determined. Hence, if one uses a large number of training points, gradient descent can be very costly at every parameter update, as compared to stochastic gradient descent (Bottou, 2012).

For all of the numerical experiments we use three nodes in the input layer (i.e., each node represents a column vector in the input vector $\mathbf{y} = [\mathbf{x}, \theta, \mathbf{t}]$), three hidden layers with 256 nodes in each layer, and an output layer containing one node (see Figure 5.1). The number of layers and neurons that are used in the hidden section of the network may be modified to reduce the training time at the expense of accuracy. Following definition 5.1 we apply the non-linearity in all but the last layer. In the last layer we set $\sigma = 1$. For the numerical simulations we set $T = 1$ s, $x^l_b = -1$ km, and $x^r_b = 1$ km. We use $N_{pde} = 400$ training points per iteration (i.e., penalty function update) for the PDE loss, randomly chosen at each iteration from a uniform distribution in the space-time domain $(x,t) \in [-1,1] \times [0,1]$. For the initial condition (penalty function 5.7) we use $N_{ic} = 100$ training points randomly chosen from a uniform distribution in the space domain $x \in [-0.25, 0.25]$. Notice that the range on which we evaluate the points N_{ic} is different to

the one on which the evaluate the points N_{pde}, because we want to heavily sample the initial condition in the region where the initial intensity is localized. We use $N_{bc}^{l} = 50$ and $N_{bc}^{l} = 50$ training points at each of the space boundaries located at x_b^l and x_b^r, respectively. We change the values of the training points with iteration (while being in a fixed range), so that the neural network can extract information about the different parts of the computational domain. Anytime we evaluate the neural network we define the input vector as $(\mathbf{x}, \theta, \mathbf{t})$, which is of size $3 \times N$, where the N corresponds to the number of ordered pairs in the spatio-temporal plane (i.e., $[(x_1, t_1), (x_2, t_2), (x_3, t_3), \cdots, (x_{N-1}, t_{N-1}), (x_N, t_N)]$), and takes a value of $N_{pde}, N_{ic}, N_{bc}^{l}$, or N_{bc}^{r}, depending on whether we evaluate the network $\mathcal{N}_{\boldsymbol{\omega}}$ for the PDE, initial condition, left boundary condition, or right boundary condition loss, respectively. The number 3 corresponds to the number of variables that govern the 1-D RTE (i.e., x, θ, t).

In the next section we provide a comparison between the soft (equation 5.7) and mixed-constraints (equation 5.10) approach for three different examples. We illustrate that, generally speaking, the mixed-constraints approach is more accurate than the soft-constraints approach. This is because we do not need sample points for the initial condition, and we do not need to consider regularization parameter λ_{ic}.

penalty function	experiment 1	experiment 2	experiment 3
equation 5.7, λ_{ic}=0.1	2.49×10^{-3}	2.96×10^{-3}	9.68×10^{-4}
equation 5.7, λ_{ic}=1.0	3.06×10^{-3}	9.82×10^{-4}	5.39×10^{-4}
equation 5.7, λ_{ic}=10	3.82×10^{-4}	2×10^{-3}	5.52×10^{-4}
equation 5.10	8.12×10^{-5}	1.32×10^{-5}	2.26×10^{-5}

Table 5.1 L2 error, scaled by the number of samples (101×101), between deep learning and reference solutions.

5.5 Numerical Examples

To illustrate the capability of both the soft- (equation 5.7) and mixed-constraints (equation 5.10) approach, we show the intensity (obtained via deep learning) for three different numerical simulations with increasing complexity. In the first simulation we use an isotropic initial condition ($c = 0$ in equation 5.4) and assume the boundary is located at $(x_b^l, x_b^r) = (-\infty, \infty)$ (i.e., the boundary condition is not imposed). For this numerical example there is a closed-form analytical solution. The second simulation is the same as the first simulation with reflective boundary conditions at $(x_b^l, x_b^r) = (-1, 1)$. In the third simulation we set the initial condition to be non-zero only for the right-going intensity ($c = 1$ in equation 5.4), and apply reflective boundary conditions at $(x_b^l, x_b^r) = (-1, 1)$ as in the second simulation, and set τ_s to be a function of space. These last two simulations can not be solved analytically. We compare these simulations to a time-stepping algorithm developed by Jaimes and Snieder (2023). We outline in the section 5.8 the details

of this algorithm.

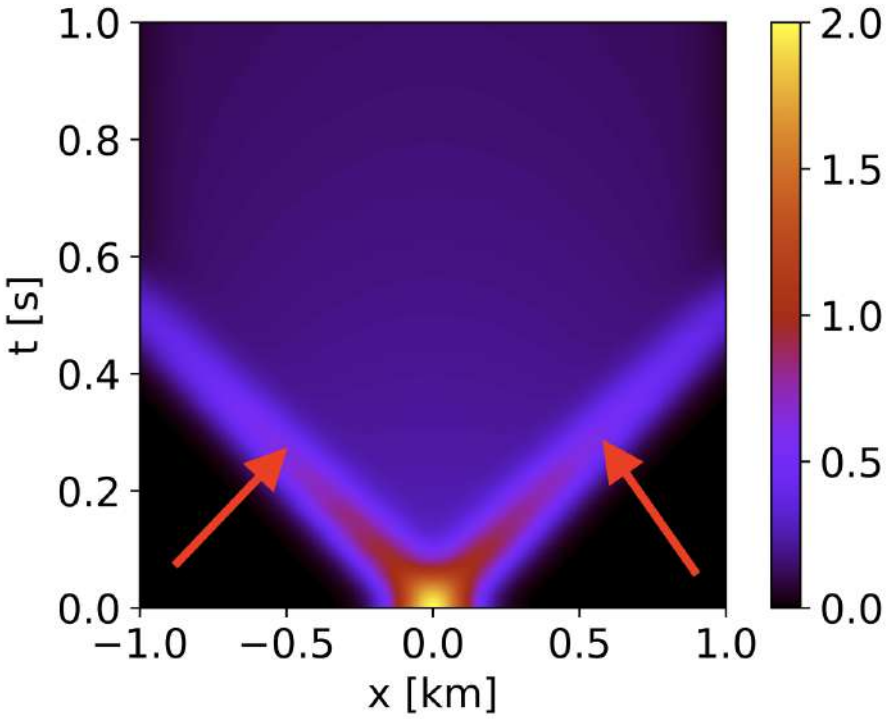

Figure 5.3 Total intensity for the initial condition defined by $[I_0(x, \theta = 0), I_0(x, \theta = \pi)] = e^{-x^2/2\sigma}\delta(t)[1, 1]$. The red arrows point to the direct energy which advects and decays over time.

For all the numerical examples we set $v = 2.0$ km/s, $F = B = 0.5$, and train the network for 5×10^4 iterations. We initialize the network with a learning rate of 10^{-3}, and decrease the rate by a factor of 0.8 every 1000 iterations up to 10^4 iterations, after which we fix the rate to 1×10^{-4}. We find these values for the learning rate to work reasonable well for the numerical experiments in this chapter. The learning rate is the step-size with which we update the weights in the stochastic gradient descent algorithm, and it is a user-defined hyper-parameter that is crucial in the optimization algorithm. Too big of a learning rate makes the optimization algorithm jump around the desired minimum; and too small of a learning rate leads to small or negligible weight updates (Gardner and Dorling, 1998). We retrieve the trainable parameters $\gamma = 1.47, 1.3, 1.34$ in equation 5.9 for the first, second, and third simulation, respectively. While we do not know the reason we retrieve these particular values, including the exponential term in expression 5.9 allows us to reduce (as compared to not including the exponential term) the error between the reference solution and the deep learning solution obtained from the penalty function 5.10 by 30% to 90%, depending on the numerical simulation. For all numerical experiments we show the total intensity, defined as $I_t(x, t) = I(x, , \theta = 0, t) + I(x, \theta = \pi, t)$. We display the total intensity in the space-time domain $[-1, 1]$ km

$\times$ $[0, 1]$ s, with time along the vertical axis, and space along the horizontal axis, and use 101 points uniformly distributed in each of the axes.

5.5.1 Example 1: Isotropic Initial Condition and No Boundary

We set $I(x, \theta, t = 0) = e^{-x^2/2L^2}$, with $L = 0.1$ km, and the scattering mean free time $\tau_s = 0.2$ s. Since the initial condition is isotropic, and the values of the forward (F) and backward (B) scattering coefficients are the same, the left and right going intensities are symmetric around $x = 0$. Figure 5.3 shows the exact total intensity for the initial condition $[I_0(x, \theta = 0), I_0(x, \theta = \pi)] = e^{-x^2/2L^2}\delta(t)[1, 1]$ (i.e., equation 5.4 after spatial convolution). This solution is symmetric due to the choice of initial condition and scattering pattern. The red arrows point to the direct energy which propagates towards the right (or left) direction and decays over time due to scattering. Due to causality there is no energy before the direct wave. Over time, the wave energy in the region between the direct waves increases, as scattering occurs. The magnitude and rate of decay of this scattered energy is a function of the scattering properties of the medium. Figure 5.5 shows total intensity obtained via deep learning for the numerical set-up in this

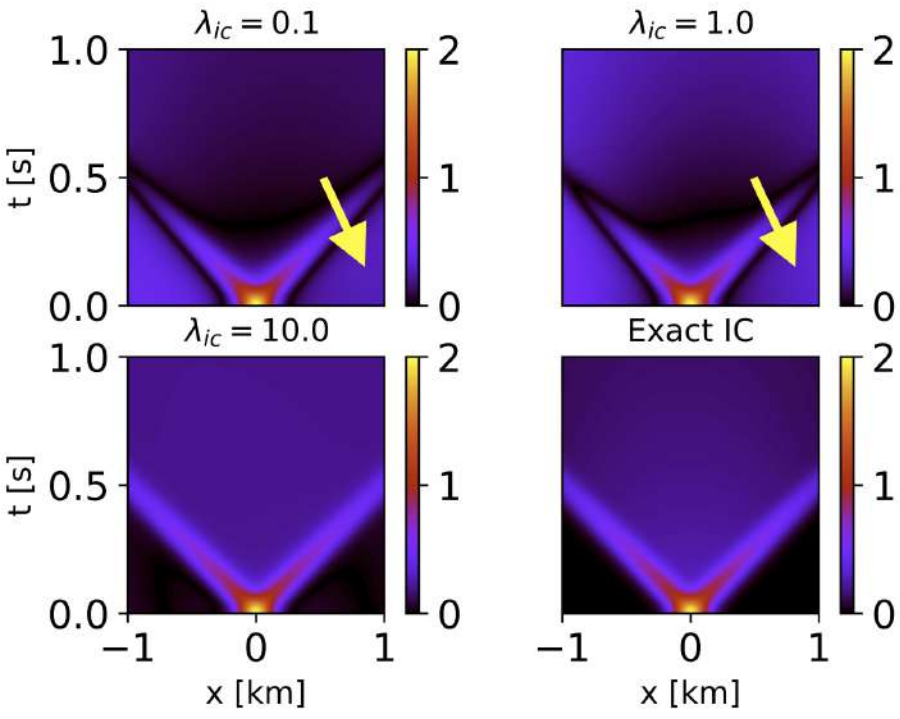

Figure 5.4 Total intensity obtained via deep learning for the numerical set-up in section 5.5.1. The yellow arrow points to energy arriving before the direct wave

section, based on either the penalty function 5.7 or 5.10. The panels with λ_{ic} show the PINNs solution using the penalty function 5.7 (soft-constraints approach). The panel labeled *Exact IC* shows the PINNs

solution using the penalty function 5.10 (mixed-constraints approach). We use this naming convention for the latter because the initial condition is handled exactly via the neural network architecture. The top panels show energy that arrives before the direct wave. This acausal energy is non-physical. The bottom panels resemble, to some extent, the reference solution in Figure 5.3. To quantify the error of the total intensity computed through deep learning, we show in Figure 5.5 the absolute difference between the deep learning solutions in Figure 5.4 and the analytical solution in Figure 5.3. The absolute difference of the

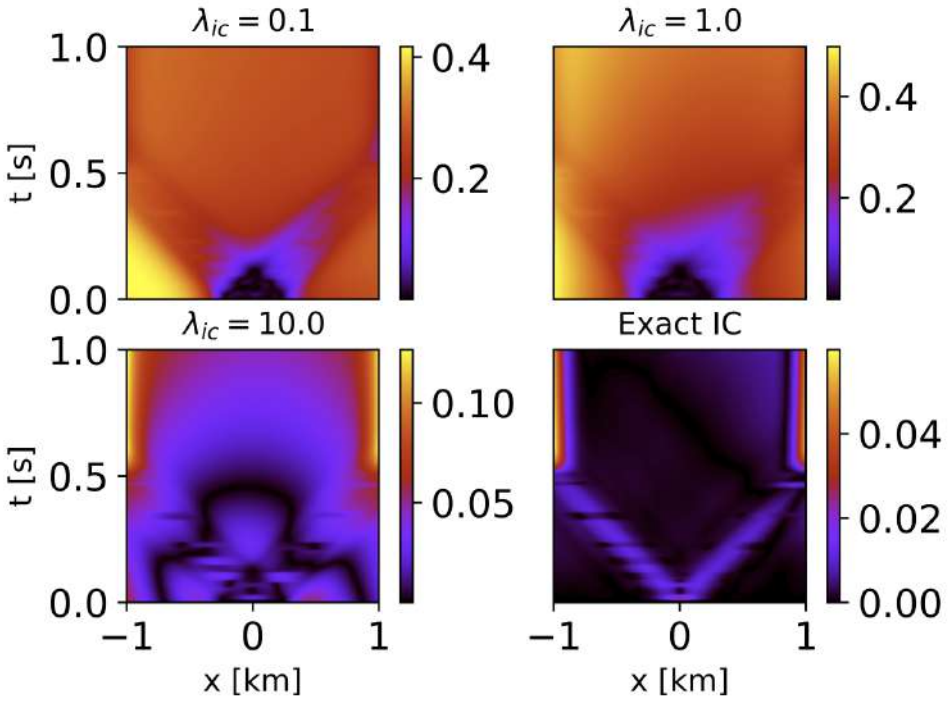

Figure 5.5 Difference of total intensity obtained through deep learning with the exact total intensity in Figure 5.3.

intensities in the top panels is of the same scale as the reference intensity in Figure 5.3. This means that for these two panels the relative error is around 50 to 100%. The error in the bottom left panel is lower (1% to 6%) than in the top panels, but not as good as the bottom right panel. One could iterate over many more values of λ_{ic} and λ_{bc}, and choose the optimal parameters based on an L-curve or generalized cross-validation, as mentioned before, but doing so is expensive.

A simple modification of the neural network and penalty function (i.e., mixed-constraints approach via equation 5.10) allows the deep learning solution to resemble the analytical solution as illustrated by the bottom right panel. Since we do not have training data along the boundary we fail to accurately retrieve the solution at $(x_b^l, x_b^r) = (-1, 1)$ (indicated by the vertical red bands in the bottom right panel). There is also a small error of about 1% along the direct arrival. Figure 5.6 shows a comparison of the total intensity

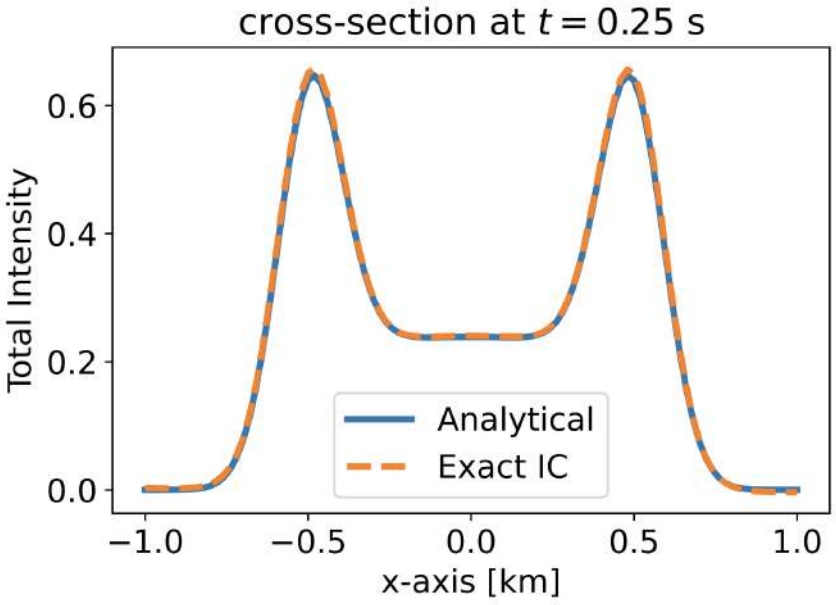

Figure 5.6 Total intensity at a time of $t = 0.25$ s. The solid blue and orange lines are the analytical (equation 5.4) convolved with the initial condition and neural network (using penalty function 5.10) solutions, respectively.

at $t = 0.25$ s between the analytical solution in Figure 5.3 and the solution obtained from the mixed-constraints approach (Exact IC) in the bottom right panel of Figure 5.4. This neural network solution resembles the analytical solution, despite the complicated nature of equation 5.4. Table 3.1 shows that the mixed-constraints approach decreases the error of the deep learning solution, as compared to the soft-constraints approach, by multiple orders of magnitude.

5.5.2 Example 2: Isotropic Initial Condition and Reflective Boundary

For this example, we slightly modified the numerical example in section 5.5.1 to make the numerical modeling more challenging. We impose fully reflective boundary conditions at $(x_b^l, x_b^r) = (-1, 1)$, by requiring that $I(x_b^l, \theta = 0, t) - I(x_b^l, \theta = \pi, t) = 0$ and $I(x_b^r, \theta = 0, t) - I(x_b^r, \theta = \pi, t) = 0$. This boundary condition follows from conservation of energy by enforcing that the local current density $\boldsymbol{J}$ at the boundary vanishes (Margerin et al., 1998). This boundary condition is challenging because at $(x_b^l, x_b^r) = (-1, 1)$ the left and right going intensities are independent of direction. For this example there is no closed-form analytical solution, and equation 5.3 must be solved numerically. As mentioned earlier, we use the 1-D version of the time-stepping algorithm developed by Jaimes and Snieder (2023) as our reference. We evolve the left and right-going intensities over time using expression 5.15. In this this numerical example, we expect the scattering behavior to be more complicated than for the previous numerical example, because of the energy that reflects off of the boundaries. At times before the direct wave reaches the boundary, the

solution is the same as the previous example. After the wave reflects off of the boundary, the solutions are different.

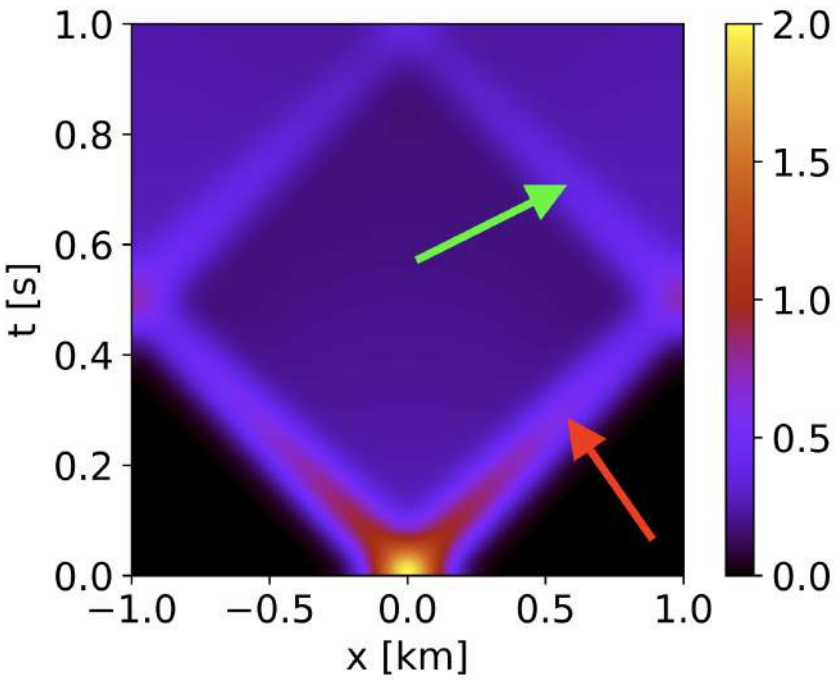

Figure 5.7 Total intensity computed with Time-stepping algorithm that we outline in section 5.8. The red arrow points to the right-propagating direct wave. The green arrow points to the right-propagating direct wave which reflects off the boundary at $x_b^r = 1$, and then propagates towards the left.

Figure 5.7 shows the time-stepping solution for this example. As in the numerical example in section 5.5.1, the solution is symmetric because of the isotropic initial condition, isotropic scattering pattern, and the same boundary condition on $(x_b^l, x_b^r) = (-1, 1)$. Over time the direct energy decays due to scattering and scattered energy in-between the direct wave starts to develop. After reaching the boundary the direct wave reflects off the boundary, reverses propagation direction, and keeps decaying as it propagates.

We can handle the boundary conditions via PINNs, by allocating training points at the left and right boundary locations, denoted by N_{bc}^l and N_{bc}^r (see expression 5.10). In this numerical example we are hoping for the neural network to learn the PDE in equation 5.2, and the initial and boundary conditions that we are imposing. The penalty function and the training algorithm becomes more complicated than that in the previous numerical experiment. For instance, with the soft-constraints (equation 5.7) approach we need to choose the correct combination of λ_{ic} and λ_{bc} values, and with the mixed-constraints approach (equation 5.10) we need to choose a suitable λ_{bc}. The latter approach becomes useful for this set-up since it is easier to choose a suitable value for λ_{bc}, than to choose the proper combination of λ_{ic} and λ_{bc} values.

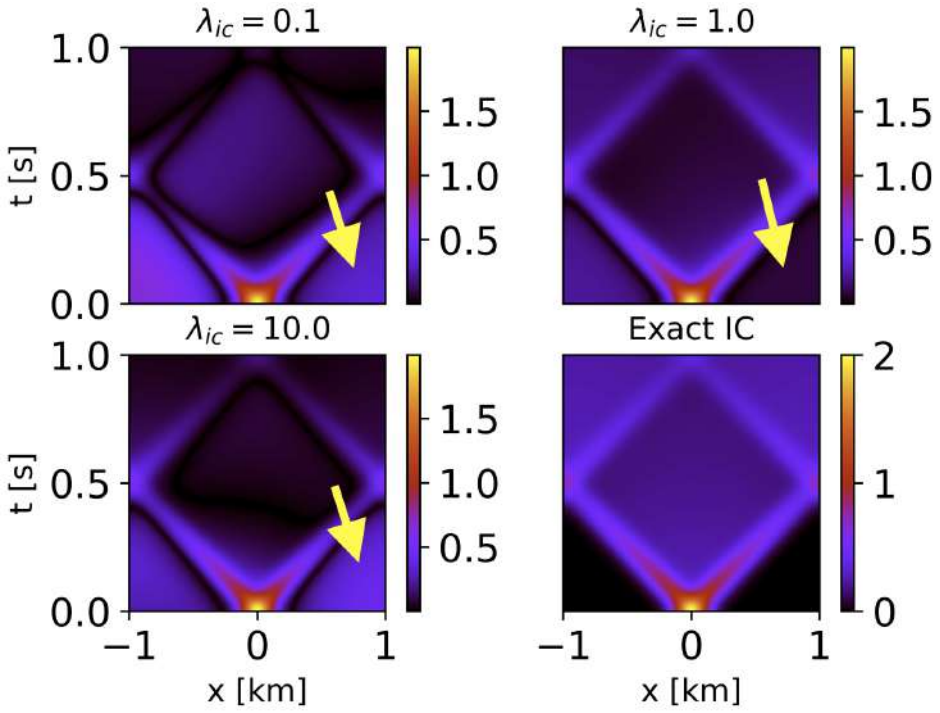

Figure 5.8 Total intensity obtained via deep learning for the numerical set-up in section 5.5.2. The yellow arrow points to energy arriving before the direct wave.

Figure 5.8 shows the total intensity computed via deep learning for the example in this section. All of the solutions obtained via the soft-constraints approach (top and bottom left panels) show energy that arrives before the direct wave, which is non-physical. The solution obtained via the mixed-constraints approach (bottom right panel) resembles the reference solution in Figure 5.7 the most. For this deep learning solution there is no acausal energy, and there is symmetry in the solution, which we expect for this numerical experiment. Figure 5.9 shows the difference between the deep learning solutions in Figure 5.8 and the time-stepping solution in Figure 5.7. For the values of λ_{ic} that we use, the error between the solutions obtained via the soft-constraints approach and the reference solution is large (relative to the intensity in Figure 5.7). As before, a simple modification of the neural network architecture and penalty function (equation 5.10) gives a solution which is close to the time-stepping solution (bottom right panel). The absolute error for the right- and left-propagating direct wave is about 10 times smaller than in Figure 5.5, and the relative error of the reflected wave is around 1%. We are most likely able to significantly reduce the error along the direct wave because the additional constraint (i.e., the boundary condition) further informs the neural network $\mathcal{N}$ about the desired mapping. Table 5.1 shows that the mixed-constraints approach significantly reduces the error between the deep learning and time-stepping solutions.

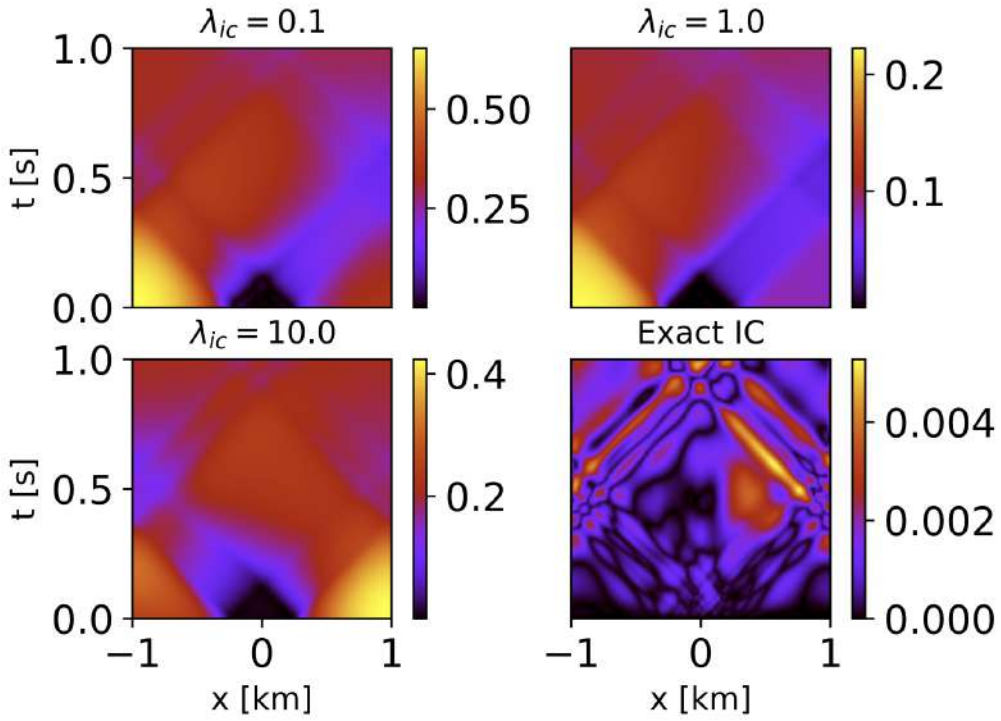

Figure 5.9 Difference between deep learning solutions and time-stepping numerical solution in Figure 5.7. The naming convention is the same as Figure 5.5.

5.5.3 Example 3: Directional Initial Condition, Reflective Boundary and Variable Scattering

In the previous numerical examples, the PINNs mixed-constraints approach gave solutions that resembled the analytical/numerical solutions (Figure 5.5, Figure 5.6, and Figure 5.9). To illustrate the capability of this deep learning approach, we consider a third numerical example with a directional initial condition, reflective boundary at $(x_b^l, x_b^r) = (-1, 1)$, and variable scattering. This example is useful to see the applicability of deep learning to model the RTE in complicated settings. We define the directional initial condition by $I_0(x, \theta = 0) = e^{-x^2/2L^2}$, with $L = 0.1$ km^2, and $I_0(x, \theta = \pi) = 0$. We set τ_s to be a piece-wise function, defined by $\tau_s = 0.2$ s for $|x| > 0.4$, and $\tau_s = 0.4$ s otherwise. There is less scattering near the source ($|x| \leq 0.4$), and more scattering far away from it ($|x| > 0.4$).

Figure 5.10 shows the time-stepping solution for the set-up in this numerical example. The red arrow points to the right-propagating direct wave. There is no left-propagating direct wave, and the energy that propagates to the left is only due to scattering. For $|x| > 0.4$ the direct wave decays more than for $|x| \leq 0.4$. The green arrow points to the left-propagating wave which arises due to the reflection off the boundary of the initial right-propagating direct wave. Over time trailing energy develops due to scattering.

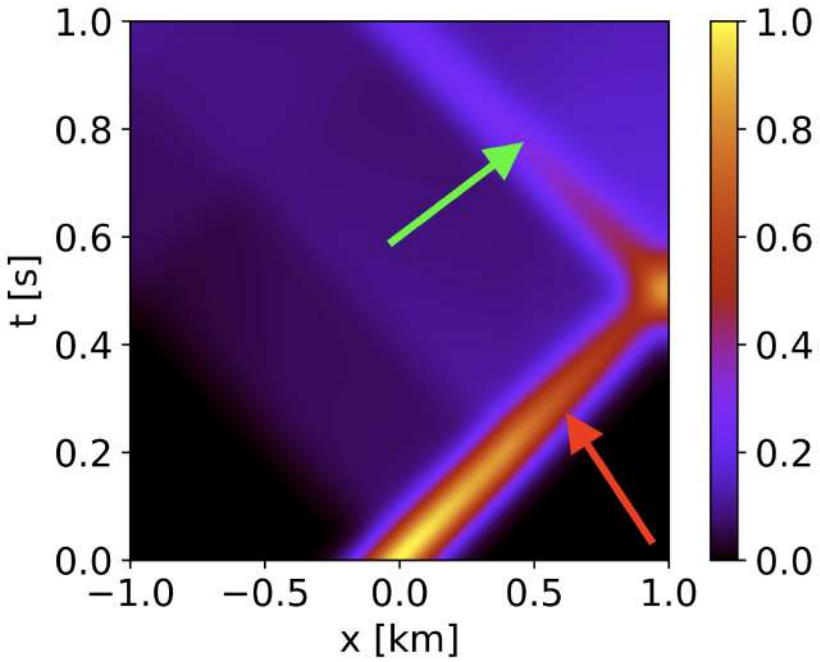

Figure 5.10 Total intensity computed using the time-stepping algorithm outlined in section 5.8. The red arrow points to the right-propagating direct wave. The green arrow points to the right-propagating direct wave which reflects off the boundary at $x_b^r = 1$, and begins to propagate towards the left.

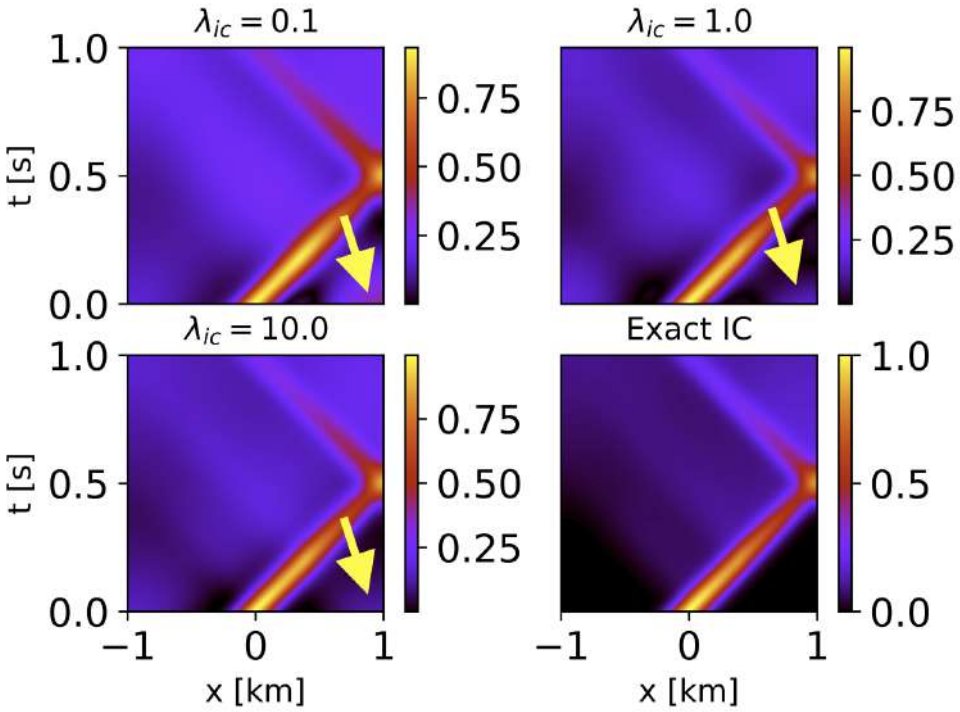

Figure 5.11 Deep learning solutions for the numerical set-up in section 5.5.3. The yellow arrow points to energy arriving before the direct wave.

Figure 5.11 shows the total intensity computed via deep learning for the numerical set-up in this section. Once again, the solutions obtained via the soft-constraints approach show non-physical energy. The mixed-constraints approach resembles the reference solution in Figure 5.10 the most. Figure 5.12 shows the absolute difference between the deep learning solutions in Figure 5.11 and the time-stepping solution in Figure 5.10. There is a large error between the solution obtained via the soft-constraints approach and the reference solution. This might indicate that one needs to iterate over many different values of λ_{ic} and λ_{bc} to find the best combination of regularization parameters that minimizes the objective function 5.7. As in the previous two numerical examples, modifying the neural network to match the initial

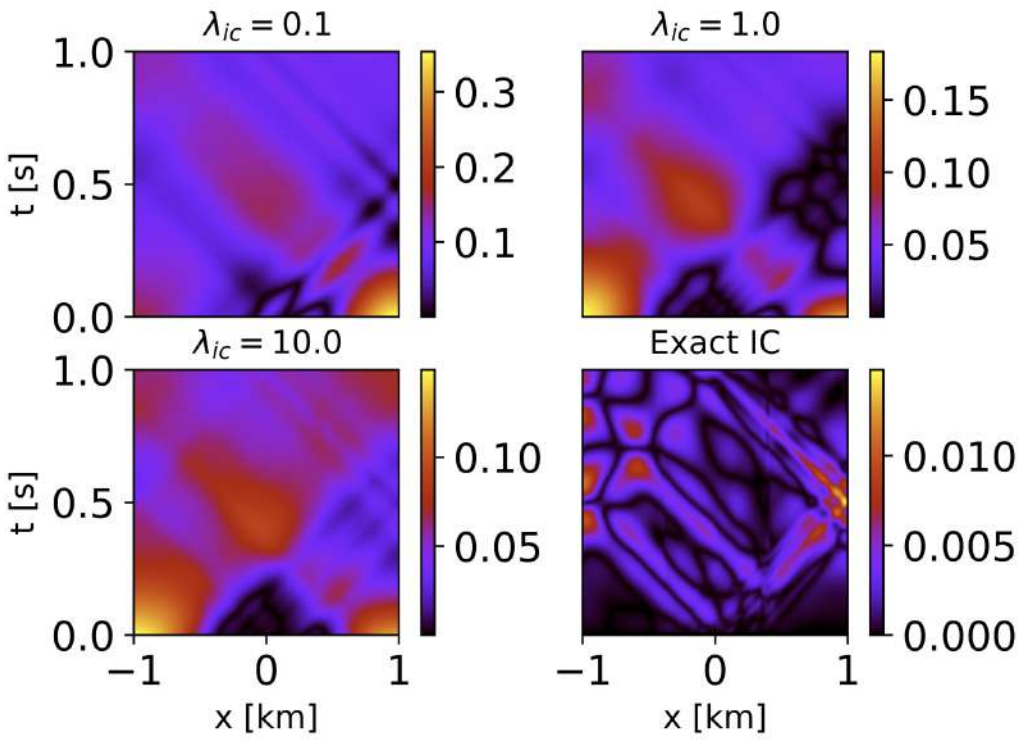

Figure 5.12 Difference of deep learning solutions with time-stepping numerical solution in Figure 5.10. The naming convention is the same as Figure 5.5.

condition exactly (mixed-constraints approach), allows the PINNs to closely match the numerical solution (bottom right panel of Figure 5.12). The goal of the mixed-constraints approach is to focus only on the boundary condition and PDE losses, and let the neural network satisfy the initial condition exactly. As for the other two numerical examples, Table 5.1 shows that the mixed-constraints approach significantly reduces the error between the deep learning and time-stepping solutions.

5.6 Discussion

This chapter provides insight into the capability of deep learning, based on PINNs, to model the 1-D RTE. There are two major advantages to using the PINNs approach as compared to the reference solution outlined in section 5.8. First, we can evaluate the neural network solution anywhere within the computational domain, rather than at fixed grid points. Second, the derivatives that we compute in equation 5.3 are handled exactly via automatic differentiation. This computational technique is based on the chain-rule, and allows the computation of the derivatives of an arbitrary function with respect to its input variables.

A soft-constraints implementation of PINNs (i.e., penalty function 5.7) does not generally allow for the accurate modeling of the physical law under consideration, because to minimize the penalty function 5.7 one needs to find the most optimal set of regularization parameters λ_{ic} and λ_{bc}. The process of finding these optimal parameters can be time consuming, the optimal parameters may change between simulations, and even if one finds the optimal regularization parameters, the fit to the PDE, initial condition, and boundary condition may be poor. To make the application of PINNs feasible, we can modify the neural network architecture to incorporate the initial condition exactly (i.e., penalty function 5.10). Table Table 5.1 shows that the mixed-constraints approach reduces the error between the reference and deep learning solutions by several orders of magnitude, depending on the numerical example. This improvement occurs for two reasons. First, because it is easier to choose a suitable weight λ_{bc} than to choose the proper combination of λ_{ic} and λ_{bc} values. Second, because when we modify the output of the neural network following expression 5.9, we exactly enforce the initial condition and also tell the network how the solution evolves over time.

To illustrate the capability of this mixed-constraints approach we considered three examples of increasing complexity. The first example (section 5.5.1) consisted of an isotropic initial condition, no boundary, and uniform scattering. The second example (section 5.5.2) was the same as the first example with a reflective boundary at $(x_b^l, x_b^r) = (-1, 1)$. The third example (section 5.5.3) consisted of a directional initial condition, reflective boundary, and variable scattering. All of these examples show that with the mixed-constraints approach we are able to match, to some extent, analytical/numerical solutions (Figure 5.5, Figure 5.6, Figure 5.9, and Figure 5.12). Even for the last example, which is the most challenging, the neural network constructed via the mixed-constraints approach is able to model the solution of the 1-D RTE within 2%. There is some error for this set-up due to the piece-wise scattering model, which one can reduce by increasing the number of training samples along the model discontinuity.

PINNs, under some modifications, appear to be a promising tool in the modeling of the RTE in complicated numerical settings (i.e., variable scattering model and involved boundary conditions). While we focus only on the forward modeling problem, a similar PINNs approach can be used for inverse problems (i.e., PDE-constrained inversion) where the observable is the solution to the PDE (or an approximation to it) and one seeks the parameters of the PDE (v and τ_s in the case of the RTE). Even though we focus on the 1-D RTE, the deep learning algorithm is easily transferable to other type of equations/physical systems. For the RTE, using PINNs for higher dimensions is not yet computationally feasible due to sampling complexity as a result of the the angular coupling (Mishra and Molinaro, 2021). However, for other equations such as the Navier-Stokes equations (e.g., Liu et al., 2022), Maxwell's equations (e.g., Lu et al., 2021), and the acoustic wave equation (e.g., Huang and Alkhalifah, 2022), the mixed-constraints PINNs approach can be readily implemented in 2- and 3-D.

5.7 Acknowledgements

This work was supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at the Colorado School of Mines.

5.8 Time-Stepping Algorithm

We want to numerically solve equation 5.3, where, for simplicity, we assume that the scattering mean free time τ_s is uniform. A similar treatment to the one below follows when τ_s is a function of space. To solve system 5.3, following the work of Paasschens (1997), we transform the system 5.3 into a system of integral equations

$$I(x,\theta,t) = \underbrace{I(x - v(t-t_0)\cos\theta, \theta, t_0)e^{-(t-t_0)/\tau_s}}_{\text{DAMPED ADVECTION}} + \underbrace{\frac{1}{\tau_s}\int_{t_0}^{t}\Big[FI(x - v(t-t')\cos\theta, \theta, t') + BI(x - v(t-t')\cos\theta, \theta+\pi, t')\Big]e^{-(t-t')/\tau_s}dt'}_{\text{SCATTERING GAIN}}, \quad (5.11)$$

where $\theta = 0, \pi$. This transformation is of use because we are able to handle the advection terms (i.e., the spatial and temporal derivatives) analytically. Equation 5.11 constitutes a set of two coupled integral equations. The first term on the right-hand side is the energy of the direct propagating wave, which decays over time due to scattering. The second term contains the gain in energy due to scattering from the forward and backward directions.

To develop an iterative scheme that depends only on the current time t and previous time $t - \Delta t$, we let $t_0 = t - \Delta t$ to obtain

$$I(x,\theta,t) = I(x - v\Delta t\cos\theta, \theta, t-\Delta t)e^{-\Delta t/\tau_s} + \frac{1}{\tau_s}\int_{t-\Delta t}^{t}\Big[FI(x - v(t-t')\cos\theta, \theta, t') + BI(x - v(t-t')\cos\theta, \theta+\pi, t')\Big]e^{-(t-t')/\tau_s}dt'. \quad (5.12)$$

We discretize the time-stepping integral using a two-point quadrature rule, with lower and upper bounds A_1 and A_2, which gives

$$\begin{aligned} I(x,\theta,t) =& I(x - v\Delta t\cos\theta, \theta, t-\Delta t)e^{-\Delta t/\tau_s} \\ &+\frac{A_1\Delta t}{\tau_s}\Big[FI(x - v\Delta t\cos\theta, \theta, t-\Delta t) + BI(x - v\Delta t\cos\theta, \theta+\pi, t-\Delta t)\Big]e^{-\Delta t/\tau_s} \\ &\frac{A_2\Delta t}{\tau_s}\Big[FI(x,\theta,t) + BI(x,\theta+\pi,t)\Big]. \end{aligned} \quad (5.13)$$

From equation 5.13 we construct the linear system

$$\begin{bmatrix} 1-\frac{A_2\Delta tF}{\tau_s} & -\frac{A_2\Delta tB}{\tau_s} \\ -\frac{A_2\Delta tB}{\tau_s} & 1-\frac{A_2\Delta tF}{\tau_s} \end{bmatrix}\begin{bmatrix} I(x,\theta=0,t) \\ I(x,\theta=\pi,t)\end{bmatrix} = \begin{bmatrix} W_r \\ W_l\end{bmatrix}, \quad (5.14)$$

where the components of the right-hand vector are

$$\begin{aligned} W_r &= I(x - v\Delta t, \theta=0, t-\Delta t)e^{-\Delta t/\tau_s} \\ &+\frac{A_1\Delta t}{\tau_s}\Big[FI(x - v\Delta t, \theta=0, t-\Delta t) + BI(x - v\Delta t, \theta=\pi, t-\Delta t)\Big]e^{-\Delta t/\tau_s}, \\ W_l &= I(x + v\Delta t, \theta=\pi, t-\Delta t)e^{-\Delta t/\tau_s} \\ &+\frac{A_1\Delta t}{\tau_s}\Big[FI(x + v\Delta t, \theta=\pi, t-\Delta t) + BI(x + v\Delta t, \theta=0, t-\Delta t)\Big]e^{-\Delta t/\tau_s}. \end{aligned}$$

The solution to the linear system 5.14 is

$$\begin{bmatrix} I(x,\theta=0,t) \\ I(x,\theta=\pi,t)\end{bmatrix} = \frac{1}{\left(1-\frac{A_2\Delta tF}{\tau_s}\right)^2 - \left(\frac{A_2\Delta tB}{\tau_s}\right)^2}\begin{bmatrix} 1-\frac{A_2\Delta tF}{\tau_s} & \frac{A_2\Delta tB}{\tau_s} \\ \frac{A_2\Delta tB}{\tau_s} & 1-\frac{A_2\Delta tF}{\tau_s} \end{bmatrix}\begin{bmatrix} W_r \\ W_l\end{bmatrix}. \quad (5.15)$$

We initialize the left- and right-going intensities with the corresponding initial condition, and evolve the intensities over time with expression 5.15. In this chapter we use the trapezoidal rule (i.e., $A_1 = A_2 = 1/2$). This quadrature makes the numerical scheme energy conserving. To handle the reflective boundary condition we introduce a buffer zone, which we illustrate in Figure 5.13. In the blue area we evolve the specific intensities using expression 5.15. We enforce the advection of energy from the start of the buffer zone to the boundary location. Once the energy reaches the boundary, we reverse the direction, and propagate the energy back into the computational domain. We assume that no energy comes from outside the computational domain.

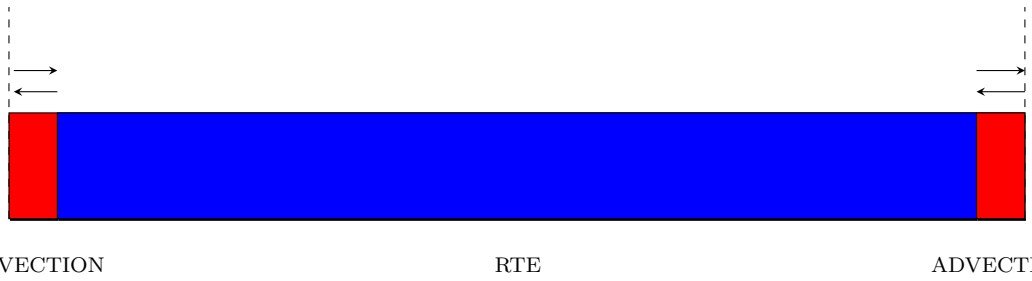

Figure 5.13 Illustration of the computational domain for the time-stepping algorithm. In the blue area we evolve the specific intensities with expression 5.15. In the red area we only impose advection of the specific intensities. The dashed lines indicate the spatial boundary. The black arrows indicate the reflection that occurs at the boundary.

CHAPTER 6

GENERAL OVERVIEW AND FUTURE RESEARCH DIRECTIONS

In this chapter I provide a general overview of the outcome of this thesis, and future research directions. To study the spatio-temporal distribution of seismic wave fields and intensities in scattering media, I divide my thesis into two parts. In the first part (**Chapter 2**), I address active source focusing in homogeneous and inhomogeneous media, and in the second part (**Chapters 3, 4, and 5**), I address energy propagation.

In **Chapter 2**, to address active source focusing, I introduce a weighted (in the frequency domain) time-reversal algorithm, which is valid in both homogeneous and heterogeneous media. With this algorithm I can locally improve the spatio-temporal resolution of the focal spot of back-propagated waves, compared to standard time-reversal. There are three major advantages to this algorithm: (1) it is directly applicable to band-limited wave fields, (2) the focusing is achieved in both space and time, and (3) it does not require regularly spaced receivers. The optimal weights, which I obtain through my algorithm, compensate for the limited acquisition geometry, the incomplete wave field sampling, and the bandwidth of the wave fields that are recorded. The modified algorithm is robust to errors in the source location, and requires accurate knowledge of the velocity of the medium. Outside the spatial area in which I localize the focusing, the wave field amplitudes become very large, which agrees with the early observations by Francia (1952), and the more recent observations in the optics community by Rogers and Zheludev (2013). Thus, if one only cares about optimizing the focal spot in a localized region, one can disregard the wave field amplitudes outside this region. However, one must keep in mind that, as Francia (1952) showed experimentally, the more one tries to localize the wave field within a limited spatial region, the more energy is wasted in the large wave field amplitudes outside the region of interest, which may degrade the focused wave field in the presence of noise.

In **Chapter 3**, I develop a time-stepping algorithm to solve the acoustic RTE in 2-D. The algorithm that I develop is valid for media with non-uniform scattering properties, arbitrary angle-dependent scattering, and a uniform speed of transport. This algorithm accounts for the angular distribution of the energy; and models direct, single scattered, weakly multiple scattered, and diffuse waves. Comparisons with exact solutions show that the algorithm is accurate. I introduce the equipartitioning index δ to quantify the extent of angular randomization of the energy. This index ranges from $\delta = 0$ for equally distributed (along angular directions) energy, to $\delta = 1$ for uni-directional energy. The direct wave, due to its preferred direction, has the equipartitioning index $\delta = 1$. However, due to discretization, the numerical values of δ that I show in **Chapter 3** and **Chapter 4** are below one. As scattering occurs, energy is

redistributed among the different propagation directions, and the value of δ decreases. At times much larger than the scattering mean free time, after many scattering events have occurred, the equipartitioning index shows that the wave field is nearly equipartitioned ($\delta \ll 1$). As waves are scattered more and more, the wave field enters a diffusive state as the current density of the field varies slowly in time, relative to the time scale over which the direction of propagation of the wave is randomized. Beyond diffusion, after the energy has equilibrated, the wave field becomes equipartitioned ($\delta \approx 0$) and stationary. I find, however, that the rate at which such transitions (diffusion and equipartitioning) occur depends strongly on space and time. Thus, equipartitioning and diffusion are not global, but local properties.

In **Chapter 4**, I extend the acoustic algorithm to include elastic waves. With this new algorithm it is possible to describe the energy propagation as a function of space, time, propagation direction, and wave mode. The numerical solutions that I obtain with this algorithm agree with the modal equipartitioning ratio, the diffusion approximation, and global expressions for the total P and S energies. As in **Chapter 3**, I study the local behavior of the angular equipartitioning using the equipartitioning index δ for each of the P and S wave modes. For elastic waves, as with acoustic waves, the equipartitioning index δ depends on space and time. The value of the equipartitioning index δ, as well as the rate at which it changes, also depends on the $P \to P$, $P \to S$, $S \to P$, and $S \to S$ scattering patterns. In addition to angular equipartitioning, I study the spatio-temporal dependence of the S to P energy ratio. After many scattering events, this local energy ratio converges in 2-D to the theoretical global ratio v_p^2/v_s^2, but the rate at which this convergence occurs is a function of space and time. The local behavior of both the angular and modal equipartitioning has implications for Green's function retrieval, which relies on the use of equipartitioned noise fields. Since equipartitioning changes locally rather than globally, the accuracy of the Green's function reconstruction may be a function of space and time, depending on the length of the time windows that are used in the reconstruction.

In **Chapter 5**, I explore the use of PINNs, a deep learning technique, to solve the 1-D RTE. To implement PINNs, I consider two different approaches: (1) soft-constraints, and (2) mixed-constraints. In the soft-constraints approach, one constructs and minimizes a penalty function which is the weighted sum of the PDE (i.e., 1-D RTE in this thesis), initial, and boundary condition losses; whereas with the mixed-constraints approach one satisfies the initial condition exactly, which results in a penalty function that only contains the PDE and boundary condition losses. I perform several tests of increasing complexity, and find that the mixed-constraints approach outperforms the soft-constraints approach. This difference in performance occurs because with the mixed-constraints approach one handles the initial condition exactly, and one only has to deal with regularization for the boundary conditions; contrary to the soft-constraints approach in which the initial condition is not satisfied exactly, and one has to deal with the regularization

of both the initial and boundary conditions. The numerical solutions that I compute through the mixed-constraints PINNs approach resemble the reference analytical/numerical solutions, even in the complicated numerical test in section 5.5.3. While I focus only on the 1-D RTE, deep learning appears as a promising tool to model the RTE in 2,3-D in complicated settings (e.g., layered velocity models and partial reflection/transmission at boundaries). Due to practical limitations it is not yet computationally feasible to use deep learning to solve the equations of radiative transfer in 2D, let alone in 3D.

6.1 Future Research

The work that I present in this thesis serves as a starting point for future research opportunities, which I outline below.

6.1.1 Weighted Time-Reversal with Field Data

In **Chapter 2**, I develop a weighted time-reversal algorithm that improves the spatio-temporal resolution of the focal spot. In that chapter, I also show that with the modified algorithm, I obtain proper focusing at multiple locations (within a given spatial area) using the optimal weights computed for a fixed location. This means that if one has access to a control point within the given spatial area, then one can achieve focusing at other points within the same area using the optimal weights computed for the control point. The improved resolution, as well as the focusing at different source locations within the localized area, would be of use in microseismic event localization (Li et al., 2019; Zhu et al., 2019), characterization of earthquake rupture (Yang and Zhu, 2019; An and Meng, 2017; Hossen et al., 2015), nondestructive imaging and localization of fractures for structural health monitoring (Leutenegger and Dual, 2004; Gliozzi et al., 2006; Douma et al., 2015), and imaging/localization of sources in random media (Kawakatsu and Montagner, 2008; Yuanwei Jin et al., 2008; Fouda and Teixeira, 2012). While using time-reversal techniques for the areas mentioned above is not new, the algorithm that I develop partially compensates for the limited acquisition geometry and the incomplete wave field sampling; thus, my algorithm produces a focal spot with enhanced resolution as compared to the standard time-reversal method. This enhancement is of importance not only to locate the source in space and time, but also to characterize the source mechanism. The improved resolution of my algorithm means that one could better describe, as compared to standard time-reversal, complicated rupture patterns, for which high resolution images are required.

The algorithm that I develop serves as a starting point for several open research questions. First, can one use the information contained in the enhanced time-reversed image to improve the characterization of complicated rupture patterns? Second, can one use the same optimal weights for similar experimental settings (e.g., small perturbations in source location, medium velocity, or acquisition geometry)? I partly

address this question in **Chapter 2**, where I explore focusing in different locations within the field-of-view using the weights computed from a control location, and imaging a horizontal dipole using the weights computed for a point source. Third, is it possible to extend the algorithm to elastic media, complicated velocity/scattering models, and noisy environments? Last, but not least, would it be possible to incorporate the optimal weights that I propose into the image-domain tomography adjoint-state method? In this method, one aims to retrieve the slowness perturbation, which optimally focuses the image, and represents the difference between assumed and true models (Shragge et al., 2013). To this end, one computes cross-correlations between the forward propagated source wave fields, and backward propagated receiver wave fields, which results in extended image gathers. To quantify the accuracy of the slowness update one measures imperfections in the image gathers through a given penalty operator. Different penalty operators have been proposed, with most attention given to those that deal with irregular illumination (Yang et al., 2013; Díaz and Sava, 2017). Hence, one can pose the following question: could one construct an imaging operator, based on the optimal weights that I develop in **Chapter 2**, which account for the irregular illumination and the band-limited nature of the source and receiver wave fields? and if so, how would one perform such construction?

6.1.2 Construction of Sensitivity Kernels for Coda Wave Interferometry

With the algorithms that I develop in **Chapter 3** and **Chapter 4**, I can compute the spatio-temporal-angular distribution of the wave intensity in strongly scattering media. Since the sensitivity kernels that are used to image localized changes are a function of the intensity of the waves, the numerical solutions that I obtain can be used to construct such kernels.

There are several advantages to using my algorithm in constructing the sensitivity kernels compared to using analytical approximations such as assuming diffuse propagation (Pacheco and Snieder, 2005) or assuming isotropic scattering (Planès et al., 2014). First, my algorithms are valid in both 2-D acoustic and elastic media; hence, it is possible to construct both the acoustic and elastic kernels without disregarding mode conversions. Second, my algorithms are valid for different source distributions, radiation patterns, and non-uniform scattering properties. Third, my algorithms account for all propagation regimes ranging from ballistic to single to multiple scattered to diffuse waves; thus, at early times, the intensities are not as smooth as those given by the diffusive approximation. This means that the sensitivity kernel that one can construct using my algorithm is strongly peaked (relative to the diffuse sensitivity kernel) at early times (i.e., direct and single scattered wave arrival). Thus, because the sensitivity kernels are strongly peaked, one could localize the time-lapse medium perturbation better than with diffuse kernels, which are only appropriate when one uses later parts of the wave forms. Fourth, because my algorithms account for the

directional wave transport, it is possible in practice to use my algorithms to retrieve velocity perturbations which depend on space as well as on angular direction, namely anisotropic perturbations (Snieder et al., 2019). Last, because my algorithms are based on the radiative transfer equations, they account for the direction of the incoming and outgoing waves at the locations of the velocity perturbations, whereas the diffuse kernels disregard the direction of the incoming and outgoing waves since by definition diffuse waves propagate with nearly equal intensity in all directions (Snieder et al., 2019). Note, however, that the radiative transfer kernels assume that the direction of the incoming and outgoing waves at the perturbation location is the same, meaning that the velocity perturbation only leaves an imprint on the arrival time of the scattered waves (Snieder et al., 2019), and does not scatter incoming waves. After constructing the sensitivity kernels with the numerical solutions obtained from my algorithms, one could first test the accuracy of the retrieval of the time-lapse perturbations with synthetic data, and then use the kernels to image time-lapse changes with field data. For a real world application, my algorithms are advantageous, compared to analytical approximations, since they do not rely on specific assumptions such as diffuse propagation, isotropic scattering, or isotropic source radiation.

6.1.3 Extension of Time-Stepping Algorithm

For practical applications, one may want to simulate the energy propagation in 3-D, account for boundary conditions, or handle layered velocity models. Theoretically, it is possible to re-formulate my algorithms in 3-D, by writing the RTE in terms of a solid angle $\hat{\Omega}$, or by parameterizing the angular dependence, for example, with an expansion in spherical harmonics (Mengüç and Viskanta, 1985). However, computationally, the numerical algorithm becomes very expensive. For acoustic waves, the time-stepping matrix would have a size $MN \times MN$, where M and N are the number of angles along the azimuthal and polar directions, respectively. For elastic waves, the time-stepping matrix would have a size of either $3MN \times 3MN$ if one ignores the polarization of the elastic waves, or $5MN \times 5MN$, if one accounts for the polarization (Turner and Weaver, 1994). To remedy this computational complication, one could use domain decomposition techniques (Champaney et al., 1997) to parallelize the numerical algorithms.

To handle boundary conditions, or piece-wise layered velocity models, one could incorporate into my algorithms energy conservation at interfaces, following the work of Margerin (2005). As I show in **Chapter 5**, it is possible with my acoustic algorithm (in 1-D) to enforce full reflection at the boundary, by requiring that the energy incident into the boundary is the same as the energy which reflects off of the boundary. A similar approach may be taken in 2- and 3-D. To this end one would take two steps: (1) keep track of the change of propagation direction and decay of the direct wave as it encounters interfaces, and (2) keep track of the incoming and outgoing directions of the scattered waves at the interfaces, and the energy that these

waves carry. For a continuous velocity model $v(z)$, where the $z-$axis is the vertical axis, the time-stepping algorithm must be modified. As Ryzhik et al.(1996), and later Tualle and Tinet (2003) show, even for a simple velocity model $v(z)$, the radiative transfer equations become complicated due to terms involving the spatial gradient of the velocity, and the derivative of the intensity with respect to the angular direction. This last term arises from the fact that, when the velocity is a function of space, the energy rays are not straight lines but curves (Ferwerda, 1999). Thus, one would need to reformulate the radiative transfer equations (including the scattering functions), and the corresponding time-stepping algorithm, in terms of the ray parameter, say p, instead of the angular direction $\hat{n}$.

6.1.4 Estimation of the Scattering Properties of the Crust

The acoustic and elastic RTE have been used to estimate the scattering properties of the crust from seismic coda (Lacombe et al., 2003; Gaebler et al., 2015; Margerin et al., 1998; Sato and Fukushima, 2013; Przybilla et al., 2009). To estimate these properties, one first computes theoretical energy envelopes for a range of models with distinct scattering properties. Then, one compares these envelopes to the energy envelopes computed from seismic coda, and chooses the scattering model whose coda has the lowest misfit with the experimental coda. To compute the theoretical envelopes, one usually assumes that the seismic source radiates energy isotropically, and that scattering is not a function of angle (e.g., Fielitz and Wegler, 2015; Ugalde et al., 2010). Depending on the source-receiver geometry, and the scattering properties of the medium, the assumptions mentioned above may not hold. Thus, the scattering properties that one retrieves with a model that assumes isotropic scattering and energy radiation, may not be representative of the subsurface. Since the properties of the subsurface dictate the scattering of energy, one could pose the following research question: is it possible to unravel spatial variations in scattering properties from the variations in scattering patterns? And if so, how would one best exploit the information contained in the variations in the scattering patterns?

The algorithms that I develop in **Chapter 3** and **Chapter 4** can be used to compute realistic energy envelopes that account for the source distribution, the angle-dependent scattering, and the different regimes of wave propagation. Some applications where my algorithms could be used include the estimation of the scattering properties of the crust of the moon (Garcia et al., 2019; Zhang et al., 2022), that of mars (Menina et al., 2021), and the earth (Przybilla et al., 2009). Similarly, the algorithm could be used to estimate the scattering properties of planetary atmospheres (Rozanov et al., 2001), as well as the estimation of scattering losses in the ocean (Li et al., 2015). For these last two applications, the angular quadrature that is used in the time-stepping algorithm may need to be modified depending on the extent of forward scattering.

6.1.5 PINNs for the 2,3-D RTE

In **Chapter 5**, I propose a PINNs algorithm to solve the 1-D RTE. In higher dimensions, the algorithm that I propose is not yet computationally feasible due to the cost of the neural network evaluations because of the angular coupling in the scattering gain term in the RTE (Mishra and Molinaro, 2021). The easiest alternative to solve the computational issues is to use parallel PINNs (Shukla et al., 2021). This new type of PINNs is based on domain decomposition, which allows for parallelization of the neural network training, and resolves the memory issues that arise from the two and three-dimensional RTE.

A second alternative is to smartly choose the training points in the PINNs algorithm, thereby reducing the sampling complexity. Raissi (2018) propose a deep learning approach to solving high-dimensional PDEs by exploting their connection with forward-backward stochastic differential equations (FBSDEs). In the forward stage, one starts with a set of training points, and evolves these points over time using the information contained in the PDE that one seeks to solve, rather than randomly choosing the training points. In the backward stage, one solves (through deep learning) an ordinary differential equation, whose solution is (theoretically) the same as that of the PDE that one wants to solve. With this approach, one can reduce the sampling redundancy, which decreases the computational cost of PINNs.

REFERENCES

Abdoulaev, G. S. (2003). Three-dimensional optical tomography with the equation of radiative transfer. *Journal of Electronic Imaging*, 12(4):594.

Aki, K. and Richards, P. G. (2009). *Quantitative seismology*. University Science Books, Mill Valley, California New York, 2. edition, corrected printing edition.

Alsamhi, S. H., Ma, O., and Ansari, M. S. (2020). Convergence of Machine Learning and Robotics Communication in Collaborative Assembly: Mobility, Connectivity and Future Perspectives. *Journal of Intelligent & Robotic Systems*, 98(3-4):541–566.

An, C. and Meng, L. (2017). Time reversal imaging of the 2015 Illapel tsunami source: Time Reversal Imaging of Illapel Tsunami. *Geophysical Research Letters*.

Anderson, B. E., Douma, J., Ulrich, T., and Snieder, R. (2015). Improving spatio-temporal focusing and source reconstruction through deconvolution. *Wave Motion*, 52:151–159.

Anderson, B. E., Ulrich, T. J., Le Bas, P.-Y., and Ten Cate, J. A. (2016). Three-dimensional time reversal communications in elastic media. *The Journal of the Acoustical Society of America*, 139(2):EL25–EL30.

Antoni, J. (2012). A Bayesian approach to sound source reconstruction: Optimal basis, regularization, and focusing. *The Journal of the Acoustical Society of America*, 131(4):2873–2890.

Apicella, A., Donnarumma, F., Isgrò, F., and Prevete, R. (2021). A survey on modern trainable activation functions. *Neural Networks*, 138:14–32.

Aradi, S. (2022). Survey of deep reinforcement learning for motion planning of autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 23(2):740–759.

Aumann, H. H., Chen, X., Fishbein, E., Geer, A., Havemann, S., Huang, X., Liu, X., Liuzzi, G., DeSouza-Machado, S., Manning, E. M., Masiello, G., Matricardi, M., Moradi, I., Natraj, V., Serio, C., Strow, L., Vidot, J., Chris Wilson, R., Wu, W., Yang, Q., and Yung, Y. L. (2018). Evaluation of Radiative Transfer Models With Clouds. *Journal of Geophysical Research: Atmospheres*, 123(11):6142–6157.

Backus, G. and Gilbert, F. (1968). The Resolving Power of Gross Earth Data. *Geophys J Int*, 16(2):169–205.

Baes, M. and Dejonghe, H. (2001). Radiative transfer in disc galaxies - I. A comparison of four methods to solve the transfer equation in plane-parallel geometry. *Monthly Notices of the Royal Astronomical Society*, 326(2):722–732.

Baker, M. R. and Patil, R. B. (1998). Universal Approximation Theorem for Interval Neural Networks. *Reliable Computing*, 4(3):235–239.

Bazargani, F. and Snieder, R. (2016). Optimal source imaging in elastic media. *Geophys. J. Int.*, 204(2):1134–1147.

Berry, M. V. and Popescu, S. (2006). Evolution of quantum superoscillations and optical superresolution without evanescent waves. *Journal of Physics A: Mathematical and General*, 39(22):6965–6977.

Billingsley, J. and Kinns, R. (1976). The acoustic telescope. *Journal of Sound and Vibration*, 48(4):485–510.

Blomgren, P., Papanicolaou, G., and Zhao, H. (2002). Super-resolution in time-reversal acoustics. *The Journal of the Acoustical Society of America*, 111(1):230–248.

Bottou, L. (2012). Stochastic Gradient Descent Tricks. In Montavon, G., Orr, G. B., and Müller, K.-R., editors, *Neural Networks: Tricks of the Trade*, volume 7700, pages 421–436. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Computer Science.

Broggini, F., Snieder, R., and Wapenaar, K. (2012). Focusing the wavefield inside an unknown 1D medium: Beyond seismic interferometry. *Geophysics*, 77(5):A25–A28.

Bucris, Y., Cohen, I., and Doron, M. A. (2012). Bayesian Focusing for Coherent Wideband Beamforming. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(4):1282–1296.

Campillo, M. (2006). Phase and Correlation in 'Random' Seismic Fields and the Reconstruction of the Green Function. *Pure and Applied Geophysics*, 163(2-3):475–502.

Camps, P. and Baes, M. (2018). The Failure of Monte Carlo Radiative Transfer at Medium to High Optical Depths. *The Astrophysical Journal*, 861(2):80.

Capon, J. (1969). High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418.

Cassereau, D. and Fink, M. (1992). Time-reversal of ultrasonic fields. III. Theory of the closed time-reversal cavity. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 39(5):579–592.

Chakroun, N., Fink, M., and Wu, F. (1995). Time reversal processing in ultrasonic nondestructive testing. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 42(6):1087–1098.

Champaney, L., Cognard, J. Y., Dureisseix, D., and Ladevèze, P. (1997). Large scale applications on parallel computers of a mixed domain decomposition method. *Computational Mechanics*, 19(4):253–263.

Chandrasekhar, S. (1960). *Radiative transfer*. Dover Publications, New York.

Chen, Y., Liu, K., Xie, Y., and Hu, M. (2020). Financial Trading Strategy System Based on Machine Learning. *Mathematical Problems in Engineering*, 2020:1–13.

Chung, S.-W. and Kim, Y.-K. (1999). Design and fabrication of 10x10 micro-spatial light modulator array for phase and amplitude modulation. *Sensors and Actuators A: Physical*, 78(1):63 – 70.

Clarke, P., Wang, H., Garrard, J., Abedi, R., and Mudaliar, S. (2019). Space-angle discontinuous Galerkin method for plane-parallel radiative transfer equation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 233:87–98.

Conti, S. G., Roux, P., and Kuperman, W. A. (2007). Near-field time-reversal amplification. *The Journal of the Acoustical Society of America*, 121(6):3602.

Curtis, A., Gerstoft, P., Sato, H., Snieder, R., and Wapenaar, K. (2006). Seismic interferometry—turning noise into signal. *The Leading Edge*, 25(9):1082–1092.

de Abreu, M. P. (2004). A two-component method for solving multislab problems in radiative transfer. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 85(3-4):311–336.

de Rosny, J. and Fink, M. (2002). Overcoming the Diffraction Limit in Wave Physics Using a Time-Reversal Mirror and a Novel Acoustic Sink. *Physical Review Letters*, 89(12):124301.

Dinther, C., Margerin, L., and Campillo, M. (2021). Implications of Laterally Varying Scattering Properties for Subsurface Monitoring With Coda Wave Sensitivity Kernels: Application to Volcanic and Fault Zone Setting. *Journal of Geophysical Research: Solid Earth*, 126(12).

Dolph, C. (1946). A Current Distribution for Broadside Arrays Which Optimizes the Relationship between Beam Width and Side-Lobe Level. *Proceedings of the IRE*, 34(6):335–348.

Douma, J., Niederleithinger, E., and Snieder, R. (2015). Locating Events Using Time Reversal and Deconvolution: Experimental Application and Analysis. *Journal of Nondestructive Evaluation*, 34(1):2.

Duran, A., Planès, T., and Obermann, A. (2020). Coda-wave decorrelation sensitivity kernels in 2-D elastic media: a numerical approach. *Geophysical Journal International*, 223(2):934–943.

Díaz, E. and Sava, P. (2017). Cascaded wavefield tomography and inversion using extended common-image-point gathers: A case study. *GEOPHYSICS*, 82(5):S391–S401.

Edelmann, G., Song, H., Kim, S., Hodgkiss, W., Kuperman, W., and Akal, T. (2005). Underwater Acoustic Communications Using Time Reversal. *IEEE Journal of Oceanic Engineering*, 30(4):852–864.

Evans, K. F. and Stephens, G. L. (1995). Microweve Radiative Transfer through Clouds Composed of Realistically Shaped Ice Crystals. Part II. Remote Sensing of Ice Clouds. *Journal of the Atmospheric Sciences*, 52(11):2058–2072.

Fan, Y., An, J., and Ying, L. (2019). Fast algorithms for integral formulations of steady-state radiative transfer equation. *Journal of Computational Physics*, 380:191–211.

Ferwerda, H. A. (1999). The radiative transfer equation for scattering media with a spatially varying refractive index. *Journal of Optics A: Pure and Applied Optics*, 1(3):L1–L2.

Fielitz, D. and Wegler, U. (2015). Intrinsic and scattering attenuation as derived from fluid induced microseismicity at the German Continental Deep Drilling site. *Geophysical Journal International*, 201(3):1346–1361.

Fink, M. (1997). Time Reversed Acoustics. *Physics Today*, 50(3):34–40.

Fink, M. (2006). Time-reversal acoustics in complex environments. *Geophysics*, 71(4):SI151–SI164.

Fink, M., Kuperman, W. A., Montagner, J.-P., and Tourin, A., editors (2002). *Imaging of Complex Media with Acoustic and Seismic Waves*, volume 84 of *Topics in Applied Physics*. Springer, Berlin.

Fink, M., Montaldo, G., and Tanter, M. (2003). Time-Reversal Acoustics in Biomedical Engineering. *Annual Review of Biomedical Engineering*, 5(1):465–497.

Fink, M. and Prada, C. (2001). Acoustic time-reversal mirrors. *Inverse Problems*, 17(1):R1–R38.

Foldy, L. L. (1945). The Multiple Scattering of Waves. I. General Theory of Isotropic Scattering by Randomly Distributed Scatterers. *Physical Review*, 67(3-4):107–119.

Fouda, A. E. and Teixeira, F. L. (2012). Imaging and tracking of targets in clutter using differential time-reversal techniques. *Waves in Random and Complex Media*, 22(1):66–108.

Franceschetti, M., Bruck, J., and Schulman, L. (2004). A random walk model of wave propagation. *IEEE Transactions on Antennas and Propagation*, 52(5):1304–1317.

Francia, G. (1952). Super-gain antennas and optical resolving power. *Il Nuovo Cimento*, 9(3):426–438.

Francis, P. N., Cooke, M. C., and Saunders, R. W. (2012). Retrieval of physical properties of volcanic ash using Meteosat: A case study from the 2010 Eyjafjallajökull eruption: Meteosat Volcanic Ash Retrievals. *Journal of Geophysical Research: Atmospheres*, 117(D20).

Gaebler, P. J., Eulenfeld, T., and Wegler, U. (2015). Seismic scattering and absorption parameters in the W-Bohemia/Vogtland region from elastic and acoustic radiative transfer theory. *Geophysical Journal International*, 203(3):1471–1481.

Gallaudet, T. and de Moustier, C. (2000). On optimal shading for arrays of irregularly-spaced or noncoplanar elements. *IEEE Journal of Oceanic Engineering*, 25(4):553–567.

Garcia, R. F., Khan, A., Drilleau, M., Margerin, L., Kawamura, T., Sun, D., Wieczorek, M. A., Rivoldini, A., Nunn, C., Weber, R. C., Marusiak, A. G., Lognonné, P., Nakamura, Y., and Zhu, P. (2019). Lunar Seismology: An Update on Interior Structure Models. *Space Science Reviews*, 215(8):50.

Gardner, M. and Dorling, S. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14-15):2627–2636.

Gliozzi, A. S., Griffa, M., and Scalerandi, M. (2006). Efficiency of time-reversed acoustics for nonlinear damage detection in solids. *The Journal of the Acoustical Society of America*, 120(5):2506–2517.

Golub, G. H. and Van Loan, C. F. (2013). *Matrix computations*. Johns Hopkins studies in the mathematical sciences. The Johns Hopkins University Press, Baltimore, fourth edition.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.

Groenenboom, J. and Snieder, R. (1995). Attenuation, dispersion, and anisotropy by multiple scattering of transmitted waves through distributions of scatterers. *The Journal of the Acoustical Society of America*, 98(6):3482–3492.

Guo, B., Huang, Y., Røstad, A., and Schuster, G. (2016). Far-field super-resolution imaging of resonant multiples. *Sci. Adv.*, 2(5):e1501439.

Han, W., Huang, J., and Eichholz, J. A. (2010). Discrete-Ordinate Discontinuous Galerkin Methods for Solving the Radiative Transfer Equation. *SIAM Journal on Scientific Computing*, 32(2):477–497.

Haney, M. M., van Wijk, K., and Snieder, R. (2005). Radiative transfer in layered media and its connection to the O'Doherty-Anstey formula. *Geophysics*, 70(1):T1–T11.

Hansen, P. C. (1992). Analysis of Discrete Ill-Posed Problems by Means of the L-Curve. *SIAM Review*, 34(4):561–580.

Hawes, M. B. and Liu, W. (2014). Sparse Array Design for Wideband Beamforming With Reduced Complexity in Tapped Delay-Lines. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(8):1236–1247.

Hennino, R., Trégourès, N., Shapiro, N. M., Margerin, L., Campillo, M., van Tiggelen, B. A., and Weaver, R. L. (2001). Observation of Equipartition of Seismic Waves. *Physical Review Letters*, 86(15):3447–3450.

Hofmeister, A. (2005). Dependence of diffusive radiative transfer on grain-size, temperature, and Fe-content: Implications for mantle processes. *Journal of Geodynamics*, 40(1):51–72.

Hossen, M. J., Cummins, P. R., Dettmer, J., and Baba, T. (2015). Time reverse imaging for far-field tsunami forecasting: 2011 Tohoku earthquake case study: Far-Field Tsuname Forecasting. *Geophysical Research Letters*, 42(22):9906–9915.

Huang, M.-X., Shih, J., Lee, R., Harrington, D., Thoma, R., Weisend, M., Hanlon, F., Paulson, K., Li, T., Martin, K., Miller, G., and Canive, J. (2003). Commonalities and Differences Among Vectorized Beamformers in Electromagnetic Source Imaging. *Brain Topography*, 16(3):139–158.

Huang, X. and Alkhalifah, T. (2022). Source location using physics-informed neural networks with hard constraints. In *Second International Meeting for Applied Geoscience & Energy*, pages 1770–1774, Houston, Texas. Society of Exploration Geophysicists and American Association of Petroleum Geologists.

Ishimaru, A. (1978). *Wave Propagation and Scattering in Random Media*. Elsevier.

Iwabuchi, H. (2006). Efficient Monte Carlo Methods for Radiative Transfer Modeling. *Journal of the Atmospheric Sciences*, 63(9):2324–2339.

Jackson, D. R. and Dowling, D. R. (1991). Phase conjugation in underwater acoustics. *The Journal of the Acoustical Society of America*, 89(1):171–181.

Jagtap, A. D., Kharazmi, E., and Karniadakis, G. E. (2020). Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, 365:113028.

Jaimes, M. A. and Snieder, R. (2021). Spatio-temporal resolution improvement via weighted time-reversal. *Wave Motion*, 106:102803.

Jaimes, M. A. and Snieder, R. (2023). Illustration of diffusion and equipartitioning as local processes: A numerical study using the radiative transfer equations. *Submitted to The Journal of the Acoustical Society of America*.

Jin, X., Cai, S., Li, H., and Karniadakis, G. E. (2021). Nsfnets (navier-stokes flow nets): Physics-informed neural networks for the incompressible navier-stokes equations. *Journal of Computational Physics*, 426:109951.

Jin, Y., Jiang, Y., and Moura, J. M. (2007). Time Reversal Beamforming for Microwave Breast Cancer Detection. In *2007 IEEE International Conference on Image Processing*, pages V – 13–V – 16, San Antonio, TX, USA. IEEE.

Jo, T., Hou, J., Eickholt, J., and Cheng, J. (2015). Improving Protein Fold Recognition by Deep Learning Networks. *Scientific Reports*, 5(1):17573.

Kanu, C. and Snieder, R. (2015). Numerical computation of the sensitivity kernel for monitoring weak changes with multiply scattered acoustic waves. *Geophysical Journal International*, 203(3):1923–1936.

Katrich, A. (2005). Do evanescent waves really exist in free space? *Optics Communications*, 255(4-6):169–174.

Kawakatsu, H. and Montagner, J.-P. (2008). Time-reversal seismic-source imaging and moment-tensor inversion. *Geophysical Journal International*, 175(2):686–688.

Khazaie, S., Cottereau, R., and Clouteau, D. (2017). Numerical observation of the equipartition regime in a 3D random elastic medium, and discussion of the limiting parameters. *Computers & Geosciences*, 102:56–67.

Kingma, D. P. and Ba, J. (2014). Adam: A Method for Stochastic Optimization. Publisher: arXiv Version Number: 9.

Klose, A. D. and Hielscher, A. H. (1999). Iterative reconstruction scheme for optical tomography based on the equation of radiative transfer. *Medical Physics*, 26(8):1698–1707.

Klose, A. D., Netz, U., Beuthan, J., and Hielscher, A. H. (2002). Optical tomography using the time-independent equation of radiative transfer — Part 1: forward model. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 72(5):691–713.

Lacombe, C., Campillo, M., Paul, A., and Margerin, L. (2003). Separation of intrinsic absorption and scattering attenuation from *Lg* coda decay in central France using acoustic radiative transfer theory. *Geophysical Journal International*, 154(2):417–425.

Larmat, C. S., Guyer, R. A., and Johnson, P. A. (2010). Time-reversal methods in geophysics. *Physics Today*, 63(8):31–35.

Le Hardy, D., Favennec, Y., and Rousseau, B. (2016). Solution of the 2-D steady-state radiative transfer equation in participating media with specular reflections using SUPG and DG finite elements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 179:149–164.

Lee, K. H., Wong, M. S., Chung, S.-R., and Sohn, E. (2014). Improved volcanic ash detection based on a hybrid reverse absorption technique. *Atmospheric Research*, 143:31–42.

Lemoult, F., Fink, M., and Lerosey, G. (2011). Acoustic Resonators for Far-Field Control of Sound on a Subwavelength Scale. *Physical Review Letters*, 107(6):064301.

Lerosey, G., de Rosny, J., Tourin, A., Derode, A., Montaldo, G., and Fink, M. (2004). Time Reversal of Electromagnetic Waves. *Physical Review Letters*, 92(19):193904.

Lerosey, G., de Rosny, J., Tourin, A., and Fink, M. (2007). Focusing Beyond the Diffraction Limit with Far-Field Time Reversal. *Science*, 315(5815):1120–1122.

Leutenegger, T. and Dual, J. (2004). Non-destructive testing of tubes using a time reverse numerical simulation (TRNS) method. *Ultrasonics*, 41(10):811–822.

Li, C., Park, K.-H., and Alouini, M.-S. (2015). On the Use of a Direct Radiative Transfer Equation Solver for Path Loss Calculation in Underwater Optical Wireless Channels. *IEEE Wireless Communications Letters*, 4(5):561–564.

Li, M., Li, H., Tao, G., Ali, M., and Guo, Y. (2019). Microseismic event location using multi-scale time reversed imaging. *Journal of Petroleum Science and Engineering*, 174:144–160.

Liu, P. (1994). A new phase function approximating to mie scattering for radiative transport equations. *Physics in Medicine and Biology*, 39(6):1025–1036.

Liu, S., Hao, Z., Ying, C., Su, H., Zhu, J., and Cheng, Z. (2022). A Unified Hard-Constraint Framework for Solving Geometrically Complex PDEs. Publisher: arXiv Version Number: 5.

Liu, W. and Weiss, S. (2010). *Wideband Beamforming*. John Wiley & Sons, Ltd, Chichester, UK.

Liu, Y., Liu, C., Hu, D., and Zhao, Y. (2019). Robust Adaptive Wideband Beamforming Based on Time Frequency Distribution. *IEEE Transactions on Signal Processing*, 67(16):4370–4382.

Lokmer, I., O'Brien, G. S., Stich, D., and Bean, C. J. (2009). Time reversal imaging of synthetic volcanic tremor sources. *Geophys. Res. Lett.*, 36(12):L12308.

Lu, L., Pestourie, R., Yao, W., Wang, Z., Verdugo, F., and Johnson, S. G. (2021). Physics-Informed Neural Networks with Hard Constraints for Inverse Design. *SIAM Journal on Scientific Computing*, 43(6):B1105–B1132.

Lu, R., Toksöz, M. N., and Willis, M. E. (2008). Locating microseismic events with time reversed acoustics: A synthetic case study. In *SEG Technical Program Expanded Abstracts 2008*, pages 1342–1346. Society of Exploration Geophysicists.

Malcolm, A. E., Scales, J. A., and van Tiggelen, B. A. (2004). Extracting the Green function from diffuse, equipartitioned waves. *Physical Review E*, 70(1):015601.

Manners, J., Thelen, J.-C., Petch, J., Hill, P., and Edwards, J. (2009). Two fast radiative transfer methods to improve the temporal sampling of clouds in numerical weather prediction and climate models: Fast RT Methods for Temporal Sampling of Cloud. *Quarterly Journal of the Royal Meteorological Society*, 135(639):457–468.

Margerin, L. (2005). Introduction to radiative transfer of seismic waves. In Levander, A. and Nolet, G., editors, *Geophysical Monograph Series*, volume 157, pages 229–252. American Geophysical Union, Washington, D. C.

Margerin, L., Campillo, M., and Tiggelen, B. (1998). Radiative transfer and diffusion of waves in a layered medium: new insight into coda Q. *Geophysical Journal International*, 134(2):596–612.

Margerin, L., Campillo, M., and Van Tiggelen, B. (2000). Monte Carlo simulation of multiple scattering of elastic waves. *Journal of Geophysical Research: Solid Earth*, 105(B4):7873–7892.

Margerin, L., Planès, T., Mayor, J., and Calvet, M. (2016). Sensitivity kernels for coda-wave interferometry and scattering tomography: theory and numerical evaluation in two-dimensional anisotropically scattering media. *Geophysical Journal International*, 204(1):650–666.

Maznev, A. and Wright, O. (2017). Upholding the diffraction limit in the focusing of light and sound. *Wave Motion*, 68:182–189.

Mengüç, M. and Viskanta, R. (1985). Radiative transfer in three-dimensional rectangular enclosures containing inhomogeneous, anisotropically scattering media. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 33(6):533–549.

Menina, S., Margerin, L., Kawamura, T., Lognonné, P., Marti, J., Drilleau, M., Calvet, M., Compaire, N., Garcia, R., Karakostas, F., Schmerr, N., van Driel, M., Stähler, S. C., Plasman, M., Giardini, D., Carrasco, S., Knapmeyer-Endrun, B., Sainton, G., and Banerdt, W. B. (2021). Energy Envelope and Attenuation Characteristics of High-Frequency (HF) and Very-High-Frequency (VF) Martian Events. *Bulletin of the Seismological Society of America*, 111(6):3016–3034.

Micolau, G., Saillard, M., and Borderies, P. (2003). DORT method as applied to ultrawideband signals for detection of buried objects. *IEEE Transactions on Geoscience and Remote Sensing*, 41(8):1813–1820.

Mimani, A. (2021). A point-like enhanced resolution of experimental Aeolian tone using an iterative point-time-reversal-sponge-layer damping technique. *Mechanical Systems and Signal Processing*, 151:107411.

Mimani, A., Prime, Z., Doolan, C., and Medwell, P. (2015). A sponge-layer damping technique for aeroacoustic Time-Reversal. *Journal of Sound and Vibration*, 342:124–151.

Mishra, M. and Srivastava, M. (2014). A view of artificial neural network. In *2014 International Conference on Advances in Engineering Technology Research (ICAETR - 2014)*, pages 1–3.

Mishra, S. and Molinaro, R. (2021). Physics informed neural networks for simulating radiative transfer. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 270:107705.

Mosk, A. P., Lagendijk, A., Lerosey, G., and Fink, M. (2012). Controlling waves in space and time for imaging and focusing in complex media. *Nature Photonics*, 6(5):283–292.

Narayanan, D., Turk, M. J., Robitaille, T., Kelly, A. J., McClellan, B. C., Sharma, R. S., Garg, P., Abruzzo, M., Choi, E., Conroy, C., Johnson, B. D., Kimock, B., Li, Q., Lovell, C. C., Lower, S., Privon, G. C., Roberts, J., Sethuram, S., Snyder, G. F., Thompson, R., and Wise, J. H. (2021). powderday: Dust Radiative Transfer for Galaxy Simulations. *The Astrophysical Journal Supplement Series*, 252(1):12.

Nassif, A. B., Shahin, I., Attili, I., Azzeh, M., and Shaalan, K. (2019). Speech Recognition Using Deep Neural Networks: A Systematic Review. *IEEE Access*, 7:19143–19165.

Noebauer, U. M. and Sim, S. A. (2019). Monte Carlo radiative transfer. *Living Reviews in Computational Astrophysics*, 5(1):1.

Obermann, A., Planès, T., Hadziioannou, C., and Campillo, M. (2016). Lapse-time-dependent coda-wave depth sensitivity to local velocity perturbations in 3-D heterogeneous elastic media. *Geophysical Journal International*, 207(1):59–66.

Obermann, A., Planès, T., Larose, E., Sens-Schönfelder, C., and Campillo, M. (2013). Depth sensitivity of seismic coda waves to velocity perturbations in an elastic heterogeneous medium. *Geophysical Journal International*, 194(1):372–382.

Oren, C. and Shragge, J. (2021). Passive-seismic image-domain elastic wavefield tomography. *Geophysical Journal International*, 228(3):1512–1529.

Ostashev, V. E., Muhlestein, M. B., and Wilson, D. K. (2017). Radiative transfer formulation for forest acoustics. *The Journal of the Acoustical Society of America*, 142(6):3767–3780.

Otter, D. W., Medina, J. R., and Kalita, J. K. (2021). A survey of the usages of deep learning for natural language processing. *IEEE Transactions on Neural Networks and Learning Systems*, 32(2):604–624.

Paasschens, J. C. J. (1997). Solution of the time-dependent Boltzmann equation. *Physical Review E*, 56(1):1135–1141.

Pacheco, C. and Snieder, R. (2005). Time-lapse travel time change of multiply scattered acoustic waves. *The Journal of the Acoustical Society of America*, 118(3):1300–1310.

Paolanti, M. and Frontoni, E. (2020). Multidisciplinary pattern recognition applications: A review. *Computer Science Review*, 37:100276.

Parvulescu, A. and Clay, C. (1965). Reproducibility of signal transmissions in the ocean. *Radio and Electronic Engineer*, 29(4):223.

Paul, A. (2005). Empirical synthesis of time-asymmetrical Green functions from the correlation of coda waves. *Journal of Geophysical Research*, 110(B8):B08302.

Pendry, J. B. (2000). Negative Refraction Makes a Perfect Lens. *Physical Review Letters*, 85(18):3966–3969.

Planès, T., Larose, E., Margerin, L., Rossetto, V., and Sens-Schönfelder, C. (2014). Decorrelation and phase-shift of coda waves induced by local changes: multiple scattering approach and numerical validation. *Waves in Random and Complex Media*, 24(2):99–125.

Pohl, D. W., Denk, W., and Lanz, M. (1984). Optical stethoscopy: Image recording with resolution /20. *Applied Physics Letters*, 44(7):651–653.

Polyakov, E. V., Mazhanov, M. S., Rolich, A. Y., Voskov, L. S., Kachalova, M. V., and Polyakov, S. V. (2018). Investigation and development of the intelligent voice assistant for the internet of things using machine learning. In *2018 Moscow Workshop on Electronic and Networking Technologies (MWENT)*, pages 1–5.

Prata, A. J. (1989). Infrared radiative transfer calculations for volcanic ash clouds. *Geophysical Research Letters*, 16(11):1293–1296.

Przybilla, J. and Korn, M. (2008). Monte Carlo simulation of radiative energy transfer in continuous elastic random media-three-component envelopes and numerical validation. *Geophysical Journal International*, 173(2):566–576.

Przybilla, J., Korn, M., and Wegler, U. (2006). Radiative transfer of elastic waves versus finite difference simulations in two-dimensional random media. *Journal of Geophysical Research*, 111(B4):B04305.

Przybilla, J., Wegler, U., and Korn, M. (2009). Estimation of crustal scattering parameters with elastic radiative transfer theory. *Geophysical Journal International*, 178(2):1105–1111.

Quijano, J. E. and Zurk, L. M. (2009). Radiative transfer theory applied to ocean bottom modeling. *The Journal of the Acoustical Society of America*, 126(4):1711.

Raissi, M. (2018). Forward-Backward Stochastic Neural Networks: Deep Learning of High-dimensional Partial Differential Equations. Publisher: arXiv Version Number: 1.

Raissi, M., Perdikaris, P., and Karniadakis, G. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707.

Rasekh, M. and Seydnejad, S. R. (2014). Design of an Adaptive Wideband Beamforming Algorithm for Conformal Arrays. *IEEE Communications Letters*, 18(11):1955–1958.

Reboul, E., Le Bot, A., and Perret-Liaudet, J. (2005). Radiative transfer equation for multiple diffraction. *The Journal of the Acoustical Society of America*, 118(3):1326–1334.

Ren, K., Abdoulaev, G. S., Bal, G., and Hielscher, A. H. (2004). Algorithm for solving the equation of radiative transfer in the frequency domain. *Optics Letters*, 29(6):578.

Roberge, W. G. (1983). The spherical harmonics solution for the radiation field in plane-parallel clouds with embedded sources. *The Astrophysical Journal*, 275:292.

Rogers, E. T. F. and Zheludev, N. I. (2013). Optical super-oscillations: sub-wavelength light focusing and super-resolution imaging. *J. Opt.*, 15(9):094008.

Rose, J. H. (2002). *Time Reversal, Focusing and Exact Inverse Scattering*, pages 97–106. Springer Berlin Heidelberg, Berlin, Heidelberg.

Rossetto, V., Margerin, L., Planès, T., and Larose, (2011). Locating a weak change using diffuse waves: Theoretical approach and inversion procedure. *Journal of Applied Physics*, 109(3):034903.

Roux, P., Sabra, K. G., Kuperman, W. A., and Roux, A. (2005). Ambient noise cross correlation in free space: Theoretical approach. *The Journal of the Acoustical Society of America*, 117(1):79–84.

Rozanov, A., Rozanov, V., and Burrows, J. (2001). A numerical radiative transfer model for a spherical planetary atmosphere: combined differential–integral approach involving the Picard iterative approximation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 69(4):491–512.

Ryzhik, L., Papanicolaou, G., and Keller, J. B. (1996). Transport equations for elastic and other waves in random media. *Wave Motion*, 24(4):327–370.

Sato, H., Fehler, M. C., and Maeda, T. (2012). *Seismic Wave Propagation and Scattering in the Heterogeneous Earth : Second Edition*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Sato, H. and Fukushima, R. (2013). Radiative transfer theory for the fractal structure and power-law decay characteristics of short-period seismograms. *Geophysical Journal International*, 195(3):1831–1842.

Sato, H., Nakahara, H., and Ohtake, M. (1997). Synthesis of scattered energy density for nonspherical radiation from a point shear-dislocation source based on the radiative transfer theory. *Physics of the Earth and Planetary Interiors*, 104(1):1–13. Stochastic Seismology Stochastic Seismic Wave Fields and Realistic Media.

Schuster, G. T. (2002). Reverse-time migration = generalized diffraction stack migration. In *SEG Technical Program Expanded Abstracts 2002*, pages 1280–1283. Society of Exploration Geophysicists.

Schuster, G. T., Hanafy, S., and Huang, Y. (2012). Theory and feasibility tests for a seismic scanning tunnelling macroscope: Seismic scanning tunnelling macroscope. *Geophysical Journal International*, 190(3):1593–1606.

Sekihara, K., Sahani, M., and Nagarajan, S. S. (2005). Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. *NeuroImage*, 25(4):1056–1067.

Sens-Schönfelder, C., Snieder, R., and Stähler, S. C. (2015). The lack of equipartitioning in global body wave coda. *Geophysical Research Letters*, 42(18):7483–7489.

Shapiro, S. A. (2008). *Microseismicity: a tool for reservoir characterization*. Number 2 in Education tour series CIS. Eage, Houten.

Shragge, J., Yang, T., and Sava, P. (2013). Time-lapse image-domain tomography using adjoint-state methods. *GEOPHYSICS*, 78(4):A29–A33.

Shukla, K., Jagtap, A. D., and Karniadakis, G. E. (2021). Parallel physics-informed neural networks via domain decomposition. *Journal of Computational Physics*, 447:110683.

Snieder, R. (2002). Coda wave interferometry and the equilibration of energy in elastic media. *Physical Review E*, 66(4):046615.

Snieder, R. (2004). *A Guided Tour of Mathematical Methods: For the Physical Sciences*. Cambridge University Press, 2 edition.

Snieder, R., Duran, A., and Obermann, A. (2019). Locating velocity changes in elastic media with coda wave interferometry. In Nakata, N., Gualtieri, L., and Fichtner, A., editors, *Seismic ambient noise*, pages 188–217. Cambridge University Press, Cambridge, UK. Section: 6.

Snieder, R., Fan, Y., Slob, E., and Wapenaar, K. (2010). Equipartitioning is not sufficient for Green's function extraction. *Earthquake Science*, 23(5):403–415.

Snieder, R. and Larose, E. (2013). Extracting Earth's Elastic Wave Response from Noise Measurements. *Annual Review of Earth and Planetary Sciences*, 41(1):183–206.

Snieder, R. and van Wijk, K. (2015). *A Guided Tour of Mathematical Methods for the Physical Sciences*. Cambridge University Press, Cambridge, 3 edition.

Snieder, R., Wapenaar, K., and Wegler, U. (2007). Unified Green's function retrieval by cross-correlation; connection with energy principles. *Physical Review E*, 75(3):036103.

Soares, C., Waldhorst, A., and Jesus, S. (1999). Matched field processing: environmental focusing and source tracking with application to the North Elba data set. In *Oceans '99. MTS/IEEE. Riding the Crest into the 21st Century. Conference and Exhibition. Conference Proceedings (IEEE Cat. No.99CH37008)*, volume 3, pages 1598–1602, Seattle, WA, USA. IEEE & Marine Technol. Soc.

Song, C., Alkhalifah, T., and Waheed, U. B. (2021). A versatile framework to solve the Helmholtz equation using physics-informed neural networks. *Geophysical Journal International*, 228(3):1750–1762.

Stehly, L., Campillo, M., and Shapiro, N. M. (2006). A study of the seismic noise from its long-range correlation properties. *Journal of Geophysical Research*, 111(B10):B10306.

Steinacker, J., Bacmann, A., and Henning, T. (2002). Application of adaptive multi-frequency grids to three-dimensional astrophysical radiative transfer. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 75(6):765–786.

Sánchez-Sesma, F. J., Pérez-Ruiz, J. A., Luzón, F., Campillo, M., and Rodríguez-Castellanos, A. (2008). Diffuse fields in dynamic elasticity. *Wave Motion*, 45(5):641–654.

Tanter, M., Aubry, J.-F., Gerber, J., Thomas, J.-L., and Fink, M. (2001). Optimal focusing by spatio-temporal inverse filter. I. Basic principles. *The Journal of the Acoustical Society of America*, 110(1):37–47.

Tanter, M., Thomas, J.-L., and Fink, M. (2000). Time reversal and the inverse filter. *The Journal of the Acoustical Society of America*, 108(1):223–234.

Thomas, J.-L. and Fink, M. (1996). Ultrasonic beam focusing through tissue inhomogeneities with a time reversal mirror: application to transskull therapy. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 43(6):1122–1129.

Trégourès, N. P. and van Tiggelen, B. A. (2002). Generalized diffusion equation for multiple scattered elastic waves. *Waves in Random Media*, 12(1):21–38.

Tualle, J.-M. and Tinet, E. (2003). Derivation of the radiative transfer equation for scattering media with a spatially varying refractive index. *Optics Communications*, 228(1-3):33–38.

Turner, J. A. and Weaver, R. L. (1994). Radiative transfer of ultrasound. *The Journal of the Acoustical Society of America*, 96(6):3654–3674.

Ugalde, A., Carcolé, E., and Vargas, C. A. (2010). S-wave attenuation characteristics in the Galeras volcanic complex (south western Colombia). *Physics of the Earth and Planetary Interiors*, 181(3-4):73–81.

van Rossum, M. C. W. and Nieuwenhuizen, T. M. (1999). Multiple scattering of classical waves: microscopy, mesoscopy, and diffusion. *Reviews of Modern Physics*, 71(1):313–371.

Van Veen, B. and Buckley, K. (1988). Beamforming: a versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5(2):4–24.

Verma, A. (2000). An introduction to automatic differentiation. *Current Science*, 78(7):804–807.

Viteri-Mera, C. A. and Teixeira, F. L. (2017). Equalized Time Reversal Beamforming for Frequency-Selective Indoor MISO Channels. *IEEE Access*, 5:3944–3957.

Waheed, U. b., Haghighat, E., Alkhalifah, T., Song, C., and Hao, Q. (2021). PINNeik: Eikonal solution using physics-informed neural networks. *Computers & Geosciences*, 155:104833.

Wapenaar, K., Draganov, D., Snieder, R., Campman, X., and Verdel, A. (2010). Tutorial on seismic interferometry: Part 1 — basic principles and applications. *Geophysics*, 75(5):75A195–75A209.

Wapenaar, K., Thorbecke, J., van der Neut, J., Vasconcelos, I., van Manen, D.-J., and Ravasi, M. (2014). On the focusing conditions in time-reversed acoustics, seismic interferometry, and Marchenko imaging. In *SEG Technical Program Expanded Abstracts 2014*, pages 4613–4619, Denver, Colorado. Society of Exploration Geophysicists.

Weaver, R. L. (1982). On diffuse waves in solid media. *The Journal of the Acoustical Society of America*, 71(6):1608–1609.

Weaver, R. L. (2010). Equipartition and retrieval of Green's function. *Earthquake Science*, 23(5):397–402.

Weaver, R. L. and Lobkis, O. I. (2001). Ultrasonics without a Source: Thermal Fluctuation Correlations at MHz Frequencies. *Physical Review Letters*, 87(13):134301.

Weaver, R. L. and Lobkis, O. I. (2004). Diffuse fields in open systems and the emergence of the Green's function (L). *The Journal of the Acoustical Society of America*, 116(5):2731–2734.

Wilson, J. H. (1995). Applications of inverse beamforming theory. *The Journal of the Acoustical Society of America*, 98(6):3250–3261.

Wolf, S. (2003). Efficient Radiative Transfer in Dust Grain Mixtures. *The Astrophysical Journal*, 582(2):859–868.

Wu, G.-Z., Fang, Y., Kudryashov, N. A., Wang, Y.-Y., and Dai, C.-Q. (2022). Prediction of optical solitons using an improved physics-informed neural network method with the conservation law constraint. *Chaos, Solitons & Fractals*, 159:112143.

Wu, R. S. and Aki, K. (1985). Elastic wave scattering by a random medium and the small-scale inhomogeneities in the lithosphere. *Journal of Geophysical Research*, 90(B12):10261.

Xiao, S., Altunc, S., Kumar, P., Baum, C. E., and Schoenbach, K. H. (2010). A Reflector Antenna for Focusing Subnanosecond Pulses in the Near Field. *IEEE Antennas and Wireless Propagation Letters*, 9:12–15.

Xu, F., Davis, A. B., West, R. A., Martonchik, J. V., and Diner, D. J. (2011). Markov chain formalism for vector radiative transfer in a plane-parallel atmosphere overlying a polarizing surface. *Optics Letters*, 36(11):2083.

Xu, Q., Jiang, C., Han, Y., Wang, B., and Liu, K. J. R. (2018). Waveforming: An Overview With Beamforming. *IEEE Communications Surveys & Tutorials*, 20(1):132–149.

Yamamoto, M. and Sato, H. (2010). Multiple scattering and mode conversion revealed by an active seismic experiment at asama volcano, japan. *Journal of Geophysical Research: Solid Earth*, 115(B7).

Yang, J. and Zhu, H. (2019). Locating and monitoring microseismicity, hydraulic fracture and earthquake rupture using elastic time-reversal imaging. *Geophysical Journal International*, 216(1):726–744.

Yang, T., Shragge, J., and Sava, P. (2013). Illumination compensation for image-domain wavefield tomography. *Geophysics*, 78(5):U65–U76.

Yodh, A. and Chance, B. (1995). Spectroscopy and Imaging with Diffusing Light. *Physics Today*, 48(3):34–40.

Yoshimoto, K. (2000). Monte Carlo simulation of seismogram envelopes in scattering media. *Journal of Geophysical Research: Solid Earth*, 105(B3):6153–6161.

Yuanwei Jin, Moura, J. M., Yi Jiang, Wahl, M., He Zhu, and Qiuhong He (2008). Breast cancer detection by time reversal imaging. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 816–819, Paris, France. IEEE.

Zeng, Y. (2017). Modeling of High-Frequency Seismic-Wave Scattering and Propagation Using Radiative Transfer Theory. *Bulletin of the Seismological Society of America*, 107(6):2948–2962.

Zhang, T. and Sens-Schönfelder, C. (2022). Adjoint envelope tomography for scattering and absorption using radiative transfer theory. *Geophysical Journal International*, 229(1):566–588.

Zhang, T., Sens-Schönfelder, C., and Margerin, L. (2021). Sensitivity kernels for static and dynamic tomography of scattering and absorbing media with elastic waves: a probabilistic approach. *Geophysical Journal International*, 225(3):1824–1853.

Zhang, X., Zhang, L., Zhang, J., and Mitchell, R. N. (2022). Strong Heterogeneity in Shallow Lunar Subsurface Detected by Apollo Seismic Data. *Journal of Geophysical Research: Planets*, 127(11).

Zhu, T., Sun, J., Gei, D., Carcione, J. M., Cance, P., and Huang, C. (2019). Hybrid multiplicative time-reversal imaging reveals the evolution of microseismic events: Theory and field-data tests. *Geophysics*, 84(3):KS71–KS83.

Zong, W. and Huang, G.-B. (2011). Face recognition based on extreme learning machine. *Neurocomputing*, 74(16):2541–2551. Advances in Extreme Learning Machine: Theory and Applications Biological Inspired Systems. Computational and Ambient Intelligence.

APPENDIX

COPYRIGHT PERMISSIONS

This appendix includes the permission from the Journal of Wave Motion to use my published paper (Jaimes and Snieder, 2021) as **chapter 2** of my thesis. Figure A.1 shows the authorization from the Journal of Wave Motion to use the published manuscript in this thesis. Figure A.2 and Figure A.3 show a description of the institution's and author's rights, respectively, for manuscripts published in the Journal of Wave Motion.

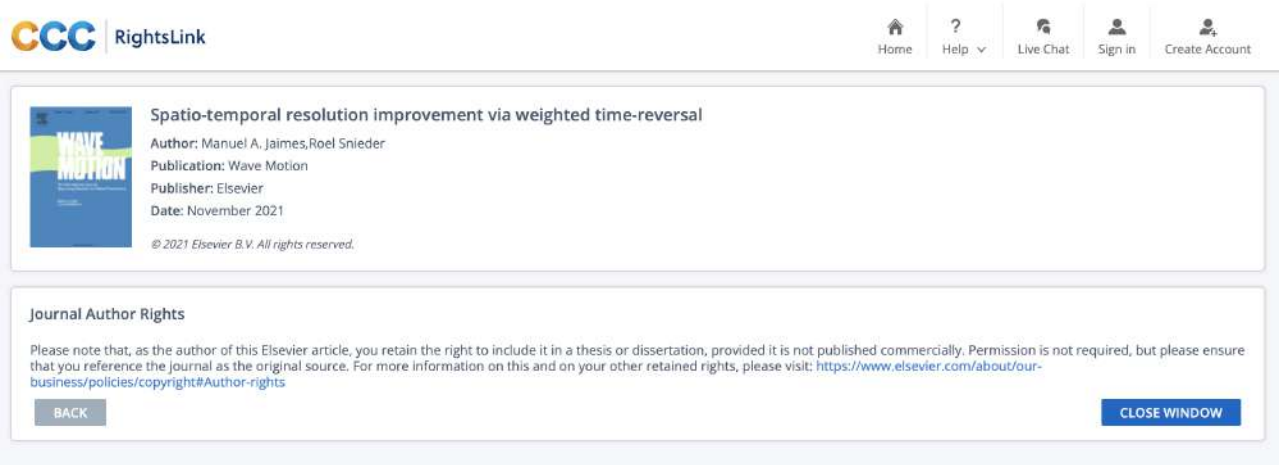

Figure A.1 Journal of Wave Motion authorization to use the published manuscript in this thesis.

Institution rights

Regardless of how the author chooses to publish with Elsevier, their institution has the right to use articles for classroom teaching and internal training. Articles can be used for these purposes throughout the author's institution, not just by the author:

Institution rights in Elsevier's proprietary journals (providing full acknowledgement of the original article is given)	**All articles**
Copies can be distributed electronically as well as in physical form for classroom teaching and internal training purposes	√
Material can be included in coursework and courseware programs for use within the institution (but not in Massive Open Online Courses)	√
Articles can be included in applications for grant funding	√
Theses and dissertations which contain embedded final published articles as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publication on ScienceDirect	√

Figure A.2 Description of institution's rights from the Journal of Wave Motion.

Author rights

The below table explains the rights that authors have when they publish with Elsevier, for authors who choose to publish either open access or subscription. These apply to the corresponding author and all co-authors.

Author rights in Elsevier's proprietary journals	Published open access	Published subscription
Retain patent and trademark rights	√	√
Retain the rights to use their research data freely without any restriction	√	√
Receive proper attribution and credit for their published work	√	√
Re-use their own material in new works without permission or payment (with full acknowledgement of the original article): 1. Extend an article to book length 2. Include an article in a subsequent compilation of their own work 3. Re-use portions, excerpts, and their own figures or tables in other works.	√	√
Use and share their works for scholarly purposes (with full acknowledgement of the original article): 1. In their own classroom teaching. Electronic and physical distribution of copies is permitted 2. If an author is speaking at a conference, they can present the article and distribute copies to the attendees 3. Distribute the article, including by email, to their students and to research colleagues who they know for their personal use 4. Share and publicize the article via Share Links, which offers 50 days' free access for anyone, without signup or registration 5. Include in a thesis or dissertation (provided this is not published commercially) 6. Share copies of their article privately as part of an invitation-only work group on commercial sites with which the publisher has a hosting agreement	√	√
Publicly share the preprint on any website or repository at any time.	√	√
Publicly share the accepted manuscript on non-commercial sites	√	√ using a CC BY-NC-ND license and usually only after an embargo period (see Sharing Policy for more information)
Publicly share the final published article	√ in line with the author's choice of end user license	×
Retain copyright	√	×

Figure A.3 Description of author's rights from the Journal of Wave Motion.