# Probabilistic Seismic Tomography with Scattered Waves

**Tuo Zhang**

GFZ Helmholtz Centre Potsdam

**Dissertation**

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

Freie Universität Berlin

vorgelegt im Fachbereich Geowissenschaften

Freie Universität Berlin

Berlin, 2022

Dissertation eingereicht am: 05. Oktober 2022
Disputation am: 08. Dezember 2022

**Prüfungskommission:**
Prof. Dr. Serge Shapiro (Vorsitz)
Prof. Dr. Frederik Tilmann
Prof. Dr. Ludovic Margerin
Dr. Christoph Sens-Schönfelder
Prof. Dr. Lena Noack

**Erstgutachter:** Prof. Dr. Frederik Tilmann (FU Berlin und GFZ Potsdam, Germany)

**Zweitgutachter:** Prof. Dr. Ludovic Margerin (Univ. Toulouse III - Paul Sabatier, Toulouse, France)

# Zusammenfassung

In dieser Arbeit wird hauptsächlich eine neue probabilistische seismische Tomographiemethode mit gestreuten Wellen vorgestellt, um die Streuung und Absorption abzubilden, die die kleinskaligen Strukturen der Erde beschreiben, genannt Adjoint Envelope Tomography (AET).

Die kleinräumige elastische Struktur des Untergrunds auf Längenskalen unterhalb der Auflösungsgrenzen der Wellenform kann in der Theorie der Zufallsmedien durch statistische Eigenschaften wie den intrinsischen Qualitätsfaktor $Q^{-1}$ oder die Fluktuationsstärke $\varepsilon$ beschrieben werden, die die Stärke der Heterogenität charakterisiert. Hochfrequent gestreute seismische Wellen werden häufig zur Charakterisierung kleinräumiger Heterogenität, intrinsischer Dämpfung oder zeitlicher Veränderungen der Wellengeschwindigkeit mit einer Vielzahl von Ansätzen verwendet. Es fehlt jedoch ein strenger Rahmen für die iterative Inversion gestreuter Wellen zur Abbildung von Heterogenität und hochfrequenter Dämpfung, vergleichbar mit der Full-waveform inversion (FWI). In dieser Arbeit wird ein mathematischer Rahmen für eine iterative Inversion unter Verwendung von Vorwärts- und adjungierten Simulationen der Strahlungstransportgleichungen vorgestellt, in voller Analogie zur FWI, die auf der Wellengleichung basiert.

Das Vorwärts- und das adjungierte Problem werden durch die Modellierung von 2-D-Mehrfachstreuung in einem zufälligen elastischen Medium mit räumlich variabler Heterogenität und Dämpfung unter Verwendung der Strahlungstransportgleichungen gelöst, die mit der Monte-Carlo-Methode gelöst werden. Die Aufzeichnung der spezifischen Energiedichte des Wellenfelds, die die vollständige Information über die Energiedichte an einem bestimmten Ort, zu einer bestimmten Zeit und in einer bestimmten Ausbreitungsrichtung enthält, ermöglicht die Berechnung von Sensitivitätskernen nach strengen theoretischen Ableitungen. Die praktische Berechnung der Kerne erfordert die Lösung der adjungierten Strahlungstransportgleichungen. In dieser Arbeit werden Sensitivitätskerne untersucht, die die Beziehungen zwischen den Änderungen des Modells in der P- und S-Wellengeschwindigkeit, der P- und S-Wellendämpfung und der Stärke der Fluktuation einerseits und der Seismogrammhüllkurve, den Laufzeitänderungen und der Wellenformdekorrelation als Beobachtungsgrößen andererseits beschreiben. Diese Sensitivitätskerne spiegeln die Auswirkungen der räumlichen Variationen der Eigenschaften des Mediums auf das Wellenfeld wider und stellen den ersten Schritt in der Entwicklung dieses tomographischen Inversionsansatzes für die Verteilung kleinräumiger Heterogenität auf der Grundlage gestreuter Wellen dar. Dann werden die Misfit-Kerne für die quadratische Differenz zwischen den vollständigen beobachteten und modellierten Hüllkurven abgeleitet, die iterativ mit der L-BFGS-Methode minimiert werden.

Zunächst wird AET numerisch in der akustischen Näherung getestet und zeigt, dass es möglich ist, die räumliche Verteilung der kleinräumigen Heterogenität und Dämpfung in iterativen Inversionen abzubilden. Die Analyse zeigt, dass die relative Bedeutung von Streu- und Dämpfungsanomalien bei der Bewertung der Modellauflösung berücksichtigt werden muss. Die Inversionen bestätigen, dass die frühe Coda für die Abbildung der Verteilung der Heterogenität wichtig ist, während spätere Coda-Wellen empfindlicher auf die intrinsische Dämpfung reagieren, und diese Abhängigkeit kann genutzt werden, um den Kompromiss zwischen beiden Materialeigenschaften zu bewältigen. Zweitens wird die Anwendung von AET in einem Laborexperiment umgesetzt. Unter Verwendung von

Ultraschalldaten von eingebetteten Sensoren in einer meterlangen Betonprobe wird die Verteilung von Absorption und Heterogenität in einem späteren bzw. früheren Coda-Zeitfenster abgebildet, was den Schlussfolgerungen aus den numerischen Tests entspricht. Die Ergebnisse lokalisieren erfolgreich einen Bereich des Salzbetons mit erhöhter Streuung und konzentrischen Anomalien der intrinsischen Dämpfung. Die Auflösungsprüfung zeigt, dass die gefundenen Anomalien die innere Struktur des Probekörpers angemessen wiedergeben. Sowohl die numerischen als auch die experimentellen Ergebnisse zeigen die Wirksamkeit von AET. Schließlich wird ein erster Versuch unternommen, AET für die Untersuchung des sphärisch symmetrischen 1D-Streumodells der Erde einzusetzen. Die Vorarbeiten zur globalen Stapelung der $P_{\mathrm{diff}}$-Koda sind abgeschlossen. Die globale Erdbebenstreuung wird durch die Monte-Carlo-Methode mit einem vorgeschlagenen Streumodell simuliert, das bis zu einem gewissen Grad mit der Beobachtung von $P_{\mathrm{diff}}$ coda übereinstimmt, die in Zukunft mit der AET-Methode verbessert werden soll. Eine Analyse über den Ursprung von $P_{\mathrm{diff}}$ coda wird durch die Simulation mit verschiedenen Streuschichten von der Lithosphäre bis zum CMB diskutiert und die verfügbare Streufläche für $P_{\mathrm{diff}}$ coda und ihre Dauer werden unter der Annahme von Einzelstreuung demonstriert.

# Summary

This thesis presents a new probabilistic seismic tomography method with scattered waves to image the scattering and absorption that describe the small-scale structures of the Earth. I call it *Adjoint Envelope Tomography* (AET).

The small-scale elastic structure of the subsurface at length scales below the resolution limits of waveform-based imaging methods can be described by random medium theory using statistical properties like correlation length of the fluctuations $a$ or their amplitude $\varepsilon$ that characterizes the strength of the heterogeneity. Additionally, the quality factor $Q$ or its inverse $Q^{-1}$ is used to describe the intrinsic attenuation. High-frequency scattered seismic waves are frequently used to characterize small-scale heterogeneity, intrinsic attenuation or temporal changes of wave velocity with a variety of approaches. However, a rigorous framework for the iterative inversion of scatted waves to image the spatial variability of heterogeneity and high-frequency attenuation comparable to full-waveform inversion (FWI) is missing. A mathematical framework for an iterative inversion using forward and adjoint simulations of the radiative transfer equations is presented in this thesis, in full analogy to FWI which is based on the wave equation.

The forward and adjoint problems are solved by modelling non-isotropic scattering in a random elastic medium with spatially variable heterogeneity and attenuation. The radiative transfer equations are solved with the Monte Carlo method. Recording of the specific energy density of the wavefield that contains the complete information about the energy density at a given position, time and propagation direction allows for calculating sensitivity kernels according to rigorous theoretical derivations. The practical calculation of the kernels involves the solution of the adjoint radiative transport equations. This thesis investigates sensitivity kernels that describe the relationships between changes of the model in P- and S-wave velocity, P- and S-wave attenuation and the strength of fluctuation on the one hand and seismogram envelope, traveltime changes and waveform decorrelation as observables on the other hand. These sensitivity kernels reflect the effect of the spatial variations of medium properties on the wavefield and constitute the first step in the development of the tomographic inversion approach for the distribution of small-scale heterogeneity based on scattered waves. To connect data misfits observed at multiple stations and extended lapse time windows to required model changes, misfit kernels are derived from the sensitivity kernels. The squared difference between observed and modelled envelopes is used as misfit function which is iteratively minimized with the Limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS).

In a first step AET is tested numerically in the acoustic approximation and shows that it is possible to image the spatial distribution of small-scale heterogeneity and attenuation in iterative inversions. The analysis shows that the relative importance of scattering and attenuation anomalies needs to be considered when the model resolution is assessed. The inversions confirm that the early coda is important for imaging the distribution of heterogeneity while later coda waves are more sensitive to intrinsic attenuation and this dependency can be used to cope with the trade-off that exists between both material properties. Building on the conclusions from the numerical test, AET is applied to a laboratory experiment in a second step. Using ultrasound data from embedded sensors in a meter-sized concrete specimen, the distribution of absorption and heterogeneity is respectively imaged using a later and earlier coda time window. The

results successfully locate an area of salt concrete with increased scattering and concentric anomalies of intrinsic attenuation. The resolution test shows that the recovered anomalies constitute reasonable representations of the internal structure of the specimen. Both numerical and experimental results show the effectiveness of AET. Lastly, an initial attempt to investigate the 1D spherically symmetric scattering model of the Earth is conducted. The preliminary work about global stacking of $P_{\mathrm{diff}}$ coda is finished. The global scattering of seismic waves from large deep earthquakes is simulated with the Monte Carlo method with a published heterogeneity model that matches the observations of $P_{\mathrm{diff}}$ coda to some degree. This model is waiting for improvement with AET in the future. An analysis about the origin of $P_{\mathrm{diff}}$ coda is discussed using the simulation with different scattering layers from lithosphere to Core-Mantle Boundary (CMB) and the available scattering area for $P_{\mathrm{diff}}$ coda and its duration are demonstrated under the assumption of single-scattering showing that the $P_{\mathrm{diff}}$ coda at high frequency originates from whole-mantle scattering.

# Declaration

Hiermit versichere ich, dass ich die vorliegende Dissertation ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Literatur angefertigt wurde. Die Stellen der Arbeit, die anderen Werken wörtlich oder inhaltlich entnommen sind, wurden durch entsprechende Angaben der Quellen kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as declared specified in the text.

Tuo Zhang
Potsdam, 23 September 2022

# Contents

Chapter 1

# Introduction

## 1.1 Motivation

Seismic waves are an important tool to probe the interior of the Earth, geotechnical targets and man-made structures. The interaction of seismic waves with the propagation medium is a complex process that occurs on a range of length scales characterized by elastic and inelastic contributions. Interpretations of seismic observations therefore use simplifying assumptions about the dominant effects. A simplification commonly used in seismic imaging is based on the assumption of weak interaction, that is to say, one assumes that a particular wave is perturbed only at one specific location on its way from source to receiver. Single-scattering imaging techniques are based on this assumption (Nishigami, 1991, 1997, 2000). Apart from the interaction with this perturbation, the propagation path is assumed to be predictable based on knowledge of the background medium (Pacheco & Snieder, 2006). Furthermore, heterogeneities in the medium have to be rare or weak enough such that reflections from different structures can be disentangled. Imaging methods have evolved in exactly that parameter range of wave propagation where the requirements for this approximation are fulfilled. This allows to deterministically image the large-scale structure with resolution limits determined by wavelength and ray coverage.

Outside this niche of the weak interaction, the wavefield can be dominated by multiple scattering such that the signal recorded at any time consists of a complex superposition of waves from different directions, therefore carrying information about different location in the medium at the same time. In this regime, the waveform is made up of continuous oscillations which cannot be associated with any particular contrast in the medium. Yet the ensemble of scattered waves that forms the coda of high-frequency seismograms carries information about the statistical properties of the heterogeneity in the propagation medium (Sato et al., 2012).

The properties of the scattered wavefield allow obtaining information about the medium that is complementary to conventional seismic investigations. In particular, one can (A) characterize statistically the structure of the propagation medium at small spatial scales (in seismology typically in the range $10^2 - 10^4$m) where traditional methods fail, (B) monitor very weak changes of the elastic properties in the medium due to the long propagation time of the waves leading to a high sensitivity and (C) monitor localized changes anywhere in the medium due to the extended sampling of the medium by scattered waves. This thesis focuses on the target (A).

Conventional wavefield-based imaging methods are intrinsically limited by wavelength. Imaging small spatial scales requires a high frequency that leads to very high cost of computation. However, the application of the adjoint method with the wave equations, e.g. the full-waveform inversion (FWI), has successfully imaged many large-scale structure of the Earth (Tarantola, 1984; Tromp et al., 2005; Fichtner et al., 2006; Fichtner, 2010). It has become a mature tool in seismology (Tape et al., 2009; Fichtner et al., 2010; Zhu et al., 2012; Tao et al., 2018). The natural idea originating from this success is the introduction of the adjoint method to the radiative transfer equations for the inversion of statistical properties that

describe the small-scale structure of the Earth.

Different from FWI with the wave equations, radiative transfer equations are used to model the propagation of scattered seismic energy in a medium that contains random fluctuations of elastic properties. To implement the adjoint method in this context, the sensitivity kernels that describe the expected changes of coda wavefield observables in response to localized changes of medium properties are required, to allow for iterative tomographic imaging of the spatial distribution of heterogeneity and intrinsic attenuation. To calculate these sensitivity kernels, directional information about the energy propagation and extension of the radiative transfer modeling are necessary to model the energy propagation in a medium with spatially variable heterogeneity and attenuation. Those tasks constitute the first step of this work. Based on these sensitivity kernels, the misfit kernels that describe the gradients of difference between the full envelopes of synthetic scattered seismic energy and real observation are available to be derived with the adjoint method. The second step is the implementation of this methodology to image the small-scale structures of material, and the application to numerical and experimental data. The effectiveness of the method can be verified from the numerical and experimental tests, which gives the confidence to investigate the final target - the small-scale structure of the Earth.

The small-scale structure of the Earth can extend from the crust and mantle to the inner core. The heterogeneity in the shallower Earth has more influence on the seismic scattering. The tomography for the crust shows strong lateral variations of scattering and intrinsic attenuation (Sens-Schönfelder et al., 2009; Calvet et al., 2013; Mayor et al., 2016), especially in volcanoes (De Siena et al., 2016). The deep Earth scattering is more challenging and is primarily observed in the $PKP$ precursors. The heterogeneity existing near the Core-Mantle Boundary (CMB) that causes the single-scattering energy to arrive earlier than $PKP$ at the surface in the distance range $110° < \Delta < 145°$ explains the observation of the precursor of $PKP$ (Cleary & Haddon, 1972; Haddon & Cleary, 1974). The $PKP$ precursor was further used to investigate scattering through the whole mantle. The proposed scattering structures of the whole Earth are still some simple 1D spherically symmetric models that only assume two or four layers from the lithosphere to CMB (Earle & Shearer, 2001; Shearer & Earle, 2004; Bentham et al., 2017). Such models are used for interpreting the observation of scattering energy in the high-frequency seismogram of $P$ coda, $P_{\text{diff}}$ coda or $PP$ precursor. At high frequency, the coda after $P_{\text{diff}}$ at long distances originates from seismic scattering energy since the time is too early for other phases and the diffracted $P_{\text{diff}}$ phase itself does not propagate effectively. The $P_{\text{diff}}$ coda thus offers a suitable observation to investigate the scattering of the whole mantle and the crust.

## 1.2 Thesis Contributions

In this thesis I approach the development of Adjoint Envelope Tomography (AET) in three tasks step by step: (1) proposing a methodology for adjoint tomography of the small-scale structure, (2) verifying this method numerically and experimentally, and (3) promoting this method to the Earth. This work led to four manuscripts during my doctoral study stage (at the time of submission two published, one under review, and one in preparation).

- **Paper I**. Tuo Zhang, Christoph Sens-Schönfelder, Ludovic Margerin, Sensitivity kernels for static and dynamic tomography of scattering and absorbing media with elastic waves: a probabilistic approach, Geophysical Journal International, Volume 225, Issue 3, June 2021, Pages 1824–1853, https://doi.org/10.1093/gji/ggab048

- **Paper II**. Tuo Zhang, Christoph Sens-Schönfelder, Adjoint envelope tomography for scattering and absorption using radiative transfer theory, Geophysical Journal International, Volume 229, Issue 1,

April 2022, Pages 566–588, https://doi.org/10.1093/gji/ggab457

- **Paper III**. Tuo Zhang, Christoph Sens-Schönfelder, Niklas Epple, Ernst Niederleithinger, Imaging of small-scale heterogeneity and absorption using Adjoint Envelope Tomography: results from laboratory experiments, under review, Journal of Geophysical Research - Solid Earth, https://doi.org/10.1002/essoar.10511632.1

- **Paper IV**. Tuo Zhang, Christoph Sens-Schönfelder, Marcelo Bianchi, Klaus Bataille, $P_{\text{diff}}$ coda in the core shadow at high frequencies.

This thesis is structured in the following way.

**Chapter** 2 gives a review of scattering theory based on Sato et al. (2012); Sens-Schönfelder et al. (2009) and the extension of the methodology to a medium with spatially variable heterogeneity and attenuation with radiative transfer equations in 2D.

**Chapter** 3 numerically solves the radiative transfer equations using the Monte Carlo method. This chapter shows the simulation results in the spatially variable heterogeneity and attenuation media and the recording of the specific energy density of the wavefield. The reciprocity theory of Green's function in 2D elastic radiative transfer theory is derived and numerically verified.

**Chapter** 4 investigates the sensitivity kernels that describe the relationships between changes of the model in P- and S-wave velocity, P- and S-wave attenuation and the strength of fluctuation on the one hand and seismogram envelope, traveltime changes and waveform decorrelation as observables on the other hand. Recording of the specific energy density of the wavefield that contains the complete information about the energy density at a given position, time and propagation direction allows us to calculate sensitivity kernels according to rigorous theoretical derivations. These sensitivity kernels reflect the effect of the spatial variations of medium properties on the wavefield and constitute the first step in the development of a tomographic inversion approach for the distribution of small-scale heterogeneity based on scattered waves.

Most parts of **Chapter** 2, the whole **Chapter** 3, 4 and **Appendix** A correspond to **Paper I**. I completed the main work under the supervision of the coauthor Christoph Sens-Schönfelder. **Appendix** A and parts of **Chapter** 3.3 that correspond to Appendix A and C in **Paper I** are contributed by the other coauthor Ludovic Margerin who also contributed to discussions of the results and improvements of the manuscript.

**Chapter** 5 presents the mathematical framework for an iterative full envelope inversion using forward and adjoint simulations of the radiative transfer equations, called Adjoint Envelope Tomography (AET), in full analogy to Full-Waveform Inversion (FWI) . Sensitivity kernels are derived for the squared difference between the full observed and modelled envelopes which is iteratively minimized with the L-BFGS method. This chapter describes an application of this algorithm in some numerical tests in the acoustic approximation and discusses that the relative importance of scattering and attenuation anomalies needs to be considered when the model resolution is assessed. The inversions confirm that the early coda is important for imaging the distribution of heterogeneity while later coda waves are more sensitive to intrinsic attenuation and this dependency can be used to cope with the trade-off that exists between both material properties. This chapter corresponds to **Paper II** that is finished under the supervision of the coauthor Christoph Sens-Schönfelder.

**Chapter** 6 presents the application of AET to experimental data. Using ultrasound data from embedded sensors in a meter-sized concrete specimen, the distributions of absorption expressed by the intrinsic quality factor $Q^{-1}$ and the fluctuation strength $\varepsilon$ that characterizes the strength of the heterogeneity are both imaged. The late and earlier coda time windows are used to invert for absorption and the distribution of heterogeneity, respectively. The results successfully locate an area of

salt concrete with increased scattering and concentric anomalies of intrinsic attenuation. The resolution test shows that the recovered anomalies can constitute reasonable representations of internal structure of the specimen. This chapter corresponds to **Paper III**. I completed the data processing, inversion and results analysis under the supervision of the coauthor Christoph Sens-Schönfelder, while the experiment is conducted by the two coauthors Niklas Epple and Ernst Niederleithinger who also contributed to discussions of the results and improvements of the manuscript.

**Chapter** 7 presents the global stacking of several earthquakes with huge magnitude in the frequency range $1 \sim 2$ Hz in which $P_{\text{diff}}$ coda can be observed at very long distances in these events. The observation of main early phases reaching the core shadow beyond 102° is known as coming from the core of the Earth like the $PKP$ wave that surpasses the $PP$ arrival. Besides this, the $P_{\text{diff}}$ wave that diffracts along the Core-Mantle Boundary (CMB) can also propagate into the core shadow, but its amplitude decreases with increasing frequency due to the vanishing effect of diffraction at high frequencies. More detailed studies show that $PKP$ is not really the first wave that arrives in the shadow zone. Scattering of $PKP$ at the CMB or the lower mantle creates the precursor to $PKP$ arriving a few seconds before actual $PKP$. However, an even faster way for seismic energy to propagate to more than 150° distance without interacting with the core or its boundary to the mantle is shown in this chapter. Scattering of $P$ wave in the lower- and mid-mantle allows waves to arrive more than 100 s prior to the $PKP$ precursor. Since this energy tails the arrival time of the $P_{\text{diff}}$ phase, we still refer to it as the $P_{\text{diff}}$ coda although its origin is not the scattering of $P_{\text{diff}}$. The Monte Carlo method is used to simulate the global earthquake scattering in a 1D spherically symmetric heterogeneity model which matches the observation of $P_{\text{diff}}$ coda. An analysis about the origin of $P_{\text{diff}}$ coda is discussed using the simulation with different scattering layers from lithosphere to CMB. Under the assumption of single-scattering, the available scattering volume for $P_{\text{diff}}$ coda and its duration are demonstrated. This chapter corresponds to the draft of **Paper IV**. I completed the main work under the supervision of the coauthor Christoph Sens-Schönfelder, the code of downloading the data is implemented in a python package from the coauthor Marcelo Bianchi who also contributed together with Klaus Bataille to discussions of the results and improvements of the manuscript.

Chapter 2

# Seismic Scattering

## 2.1 Introduction

Seismic scattering has early been recognized as the origin of coda waves that are observed in the seismograms following ballistic wave arrivals. Aki (1969) first identified the coda as the signature of backscattered waves from distributed scatterers. Aki & Chouet (1975) then proposed two classical models to describe the amplitude of the scattered wave field, the single backscattering model and the diffusion model, which represent approximations for the weak and strong scattering processes, respectively. On the one hand, the single scattering model considers scattering to happen only once between the source and the receiver. The diffusion model on the other hand requires scattering to occur often enough for the wave to lose information about its initial direction of propagation and polarization, such that it is usually applicable to seismogram modeling at long lapse-time only. Therefore, the multiple-scattering model that considers the continuum from single scattering to many scattering events is more suitable to describe the realistic scattering processes. Wu (1985) first employed the multiple scattering model for isotropic scattering and strong forward scattering. Isotropic scattering indicates that the distribution of scattering angles is uniform and does not show any preferred direction. This is a special case, and in fact non-isotropic scattering widely exists. Gusev & Abubakirov (1987), Abubakirov & Gusev (1990) and Hoshiba (1995) found evidence of non-isotropic scattering in the data. However, all the above mentioned investigations worked under the assumption of acoustic scattering which does not consider the energy conversion between P- and S-waves and the polarization of the S-waves.

Since Weaver (1990) and Ryzhik et al. (1996) derived the radiative transfer equations for elastic waves, Margerin et al. (2000) gave a detailed description of the multiple scattering of elastic waves. Gaebler et al. (2015) applied it in elastic media to model the translational and rotational motion seismogram envelopes.

A further challenge for the simulation of wave scattering arises when the scattering properties of the medium are spatially variable. Hoshiba (1994) used the Monte-Carlo simulation method to synthesize the energy density of the coda in depth-dependent scattering structure under the assumption of isotropic scattering. These efforts were continued by Margerin et al. (1998) who introduced a discontinuity in both scattering and velocity properties in multiple scattering simulations. Wegler (2004) used a model containing a scattering layer over a half space to take the depth dependency of scattering into account. Using an analytic solution of the diffusion equation in a simple model with spatially variable diffusivity, Friedrich & Wegler (2005) could improve the modeling of the spatio-temporal coda decay at Merapi volcano showing that scattering is concentrated within the edifice. A block of increased scattering strength embedded in a less heterogeneous crust above a transparent mantle was modeled with the Monte-Carlo method by Sens-Schönfelder et al. (2009) to explain Lg-wave blockage. De Siena et al. (2013) combined the modeling of multiple scattering with boundary conditions obtained from diffusion in a circular structure of highly heterogeneous material to explain spatially variable coda decay at Campi Flegrei.

Among various methods for the description of seismic scattering, radiative transfer theory is a flexible tool that can be adapted to different approximations of the scattering process. Many pioneering works employed this method to describe the wave scattering (Wu, 1985; Hoshiba, 1991; Zeng, 1991). A detailed introduction to radiative transfer theory was given by Margerin (2005). Further work used it to simulate the energy transfer for more complicated problems (Wegler et al., 2006; Przybilla et al., 2006; Sens-Schönfelder & Wegler, 2006). To numerically solve the radiative transfer equations, the Monte-Carlo method has been proposed by Hoshiba (1991); Gusev & Abubakirov (1987) and has since then been widely used to simulate the envelopes of waves (Gusev & Abubakirov, 1996; Margerin et al., 2000; Yoshimoto, 2000; Mancinelli & Shearer, 2016; Sanborn et al., 2017).

## 2.2 Random Media

Heterogeneity is widely observed in Earth materials. To describe these small-scale fluctuations of elastic properties in media, the concept of random media can be used (Sato et al., 2012). The wave velocity of heterogeneous media can be expressed as the mean velocity $V_0$ plus the perturbation $\delta V(\mathbf{x})$ which depends on the location $\mathbf{x}$:

$$V(\mathbf{x}) = V_0 + \delta V(\mathbf{x}) = V_0[1 + \xi(\mathbf{x})] \tag{2.1}$$

where $\xi(\mathbf{x}) = \delta V(\mathbf{x})/V_0$ is the fractional fluctuation of the velocity. When $\xi(\mathbf{x})$ is a random variable the medium is described as a random medium with $\langle V(\mathbf{x})\rangle = V_0$ and $\langle \xi(\mathbf{x})\rangle = 0$, where the sign $\langle\rangle$ indicates the spatial average. If this randomness is statistically uniform, i.e. does not depend on locations, the autocorrelation function (ACF) of $\xi(\mathbf{x})$ only depends on the distance $\mathbf{x}$ between two locations $\mathbf{y}$ and $\mathbf{y}+\mathbf{x}$ as:

$$R(r) = R(\mathbf{x}) = \langle \xi(\mathbf{y})\xi(\mathbf{y}+\mathbf{x})\rangle \tag{2.2}$$

where $r = |\mathbf{x}|$ when the randomness is isotropic.

The strength of fractional fluctuation $\varepsilon$ and correlation length $a$ can be defined as $\varepsilon^2 = R(0)$ and $R(a) = \varepsilon^2/e$. The ACF is related to the power spectral density function (PSDF) with the Fourier transform:

$$P(\mathbf{m}) = \tilde{R}(\mathbf{x}) = \iiint_{-\infty}^{\infty} R(\mathbf{x}) e^{-i\mathbf{m}\mathbf{x}} d\mathbf{x} \tag{2.3}$$

There are several ACF and PSDF tpyes to mathematically describe the random media in 3D:

- **Gaussian**

$$\begin{aligned} R(r) &= \varepsilon^2 e^{-r^2/a^2} \\ P(m) &= \varepsilon^2 \sqrt{\pi^3} a^3 e^{-m^2 a^2/4} \end{aligned} \tag{2.4}$$

- **Exponential**

$$\begin{aligned} R(r) &= \varepsilon^2 e^{-r/a} \\ P(m) &= \frac{8\pi\varepsilon^2 a^3}{(1 + a^2 m^2)^2} \end{aligned} \tag{2.5}$$

- **von Kármán**

$$R(r) = \frac{\varepsilon^2 2^{1-\kappa}}{\Gamma(\kappa)} \left(\frac{r}{a}\right)^{\kappa} K_{\kappa}\left(\frac{r}{a}\right)$$
$$P(r) = \frac{8\pi^{3/2}\Gamma(\kappa+3/2)\varepsilon^2 a^3}{\Gamma(\kappa)(1+a^2m^2)^{\kappa+3/2}} \tag{2.6}$$

where $\Gamma(\kappa)$ is gamma function and $K_{\kappa}$ is the modified Bessel function of the second kind of order $\kappa = 0 \sim 1$.

Sato et al. (2012) also give the realization of random media with different PSDF:

$$\xi(\mathbf{x}) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \sqrt{P(\mathbf{m})} e^{i\phi(\mathbf{m})} e^{i\mathbf{m}\mathbf{x}} d\mathbf{m} \tag{2.7}$$

where $\phi(\mathbf{m})$ is a random phase between 0 and $2\pi$.

Figure 2.1 shows the cross sections of the fractional fluctuation $\xi(\mathbf{x})$ of the velocity in the 3D random media with different PSDF types, where $a = 5$ km and $\varepsilon = 0.05$. All types have a similar perturbation of the velocity since $\varepsilon$ is the same, while the Gaussian random medium looks smoother and the von Kármán medium is also determined by a further parameter $\kappa$ that determines the short wavelength fluctuations. Note that the Exponential type is actually a special von Kármán medium with $\kappa$= 0.5.

In this thesis, I use random media with PSDF of exponential type (von Kármán $\kappa = 0.5$). Figure 2.2 shows the the cross sections of the fractional fluctuation $\xi(\mathbf{x})$ of the velocity in the 3D random media of exponential type. Figure 2.2(a) and (c) have the same correlation length $a = 1$ km which look very similar but the color scale in Figure 2.2(c) is about 10 times of Figure 2.2(a) since the $\varepsilon$ is exactly 10 times larger. The same can be observed in Figure 2.2(b) and (d) that have the same correlation length $a$ = 5 km. With the same fluctuation strength $\varepsilon$ (Figure 2.2(a,b) or (c,d)), the perturbation of the velocity has almost the same amplitude, but the scale of the heterogeneity is determined by correlation length $a$.

## 2.3 Elastic Radiative Transfer Equations

The transport of energy through a random medium is described by the theory of radiative transfer (Apresyan & Kravtsov, 1996; Wu, 1985; Hoshiba, 1991). The fundamental quantity of radiative transfer is the specific energy density $E(\mathbf{r}, \mathbf{n}, t)$ which describes the energy flux density at location $\mathbf{r}$, time $t$ in directions $\mathbf{n}$. Following the work of Sens-Schönfelder et al. (2009), the coupled radiative transfer equations for P- and S-waves in 2-D is we rewritten as:

$$\begin{aligned}\left(\frac{\partial}{\partial t} + \mathbf{n}\alpha_0 \cdot \nabla\right) E_P(\mathbf{r}, \mathbf{n}, t) = &-\left(\alpha_0 g_0^{P\to P} + \alpha_0 g_0^{P\to S} + \frac{\omega}{Q_P}\right) E_P(\mathbf{r}, \mathbf{n}, t) \\ &+ \int_{2\pi} \alpha_0 g^{P\to P}(\mathbf{n}, \mathbf{n}') E_P(\mathbf{r}, \mathbf{n}', t)\, d\mathbf{n}' \\ &+ \int_{2\pi} \beta_0 g^{S\to P}(\mathbf{n}, \mathbf{n}') E_S(\mathbf{r}, \mathbf{n}', t)\, d\mathbf{n}'\end{aligned} \tag{2.8}$$

$$\begin{aligned}\left(\frac{\partial}{\partial t} + \mathbf{n}\beta_0 \cdot \nabla\right) E_S(\mathbf{r}, \mathbf{n}, t) = &-\left(\beta_0 g_0^{S\to S} + \beta_0 g_0^{S\to P} + \frac{\omega}{Q_S}\right) E_S(\mathbf{r}, \mathbf{n}, t) \\ &+ \int_{2\pi} \beta_0 g^{S\to S}(\mathbf{n}, \mathbf{n}') E_S(\mathbf{r}, \mathbf{n}', t)\, d\mathbf{n}' \\ &+ \int_{2\pi} \alpha_0 g^{P\to S}(\mathbf{n}, \mathbf{n}') E_P(\mathbf{r}, \mathbf{n}', t)\, d\mathbf{n}'\end{aligned} \tag{2.9}$$

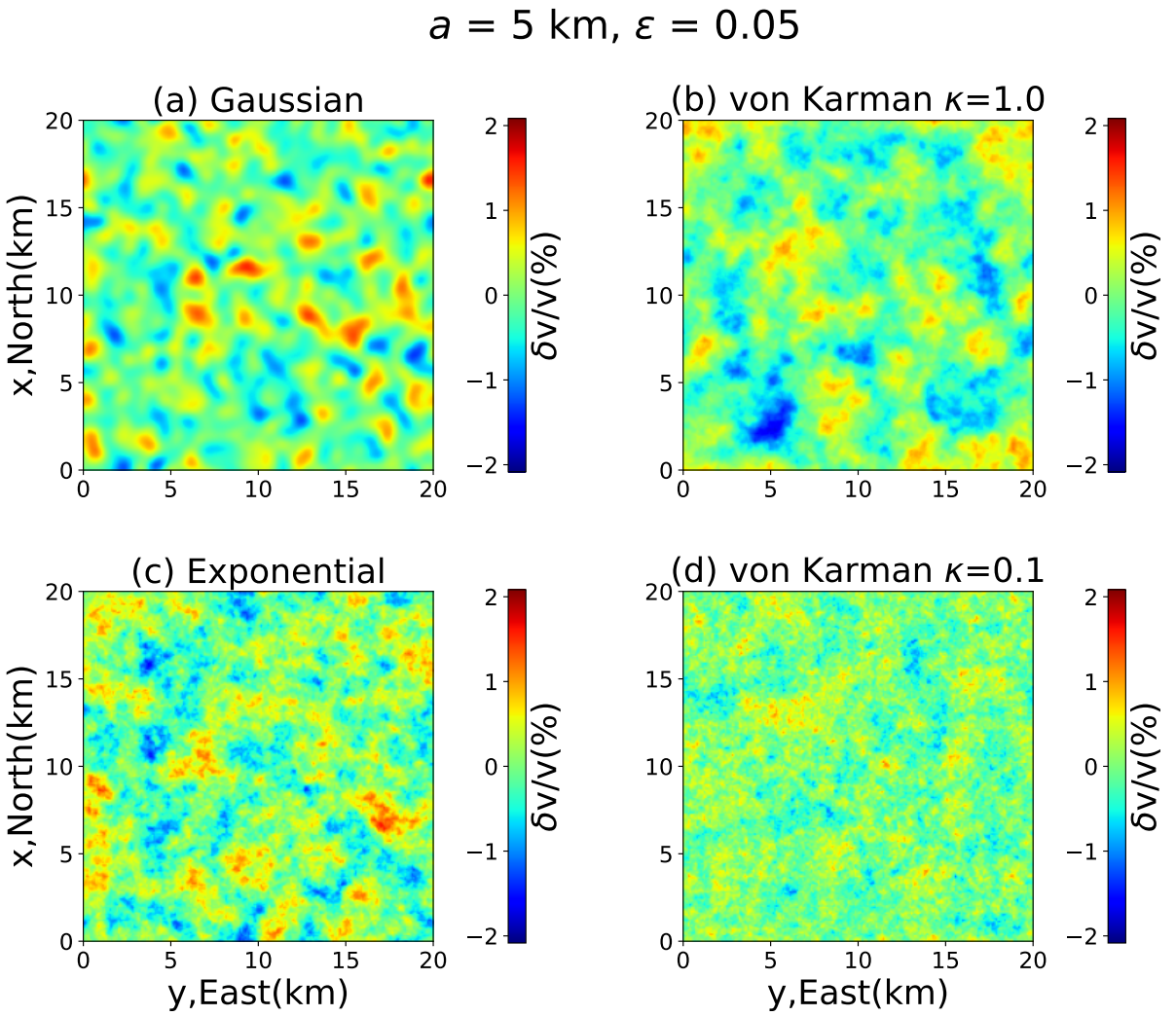



**Fig. 2.1:** The cross sections of the fractional fluctuation $\xi(\mathbf{x})$ of the velocity in the 3D random media ($a = 5$ km and $\varepsilon = 0.05$) with different PSDF types: (a)Gaussian, (b)von Kármán $\kappa = 1.0$, (c)Exponential, and (d)von Kármán $\kappa = 0.1$.

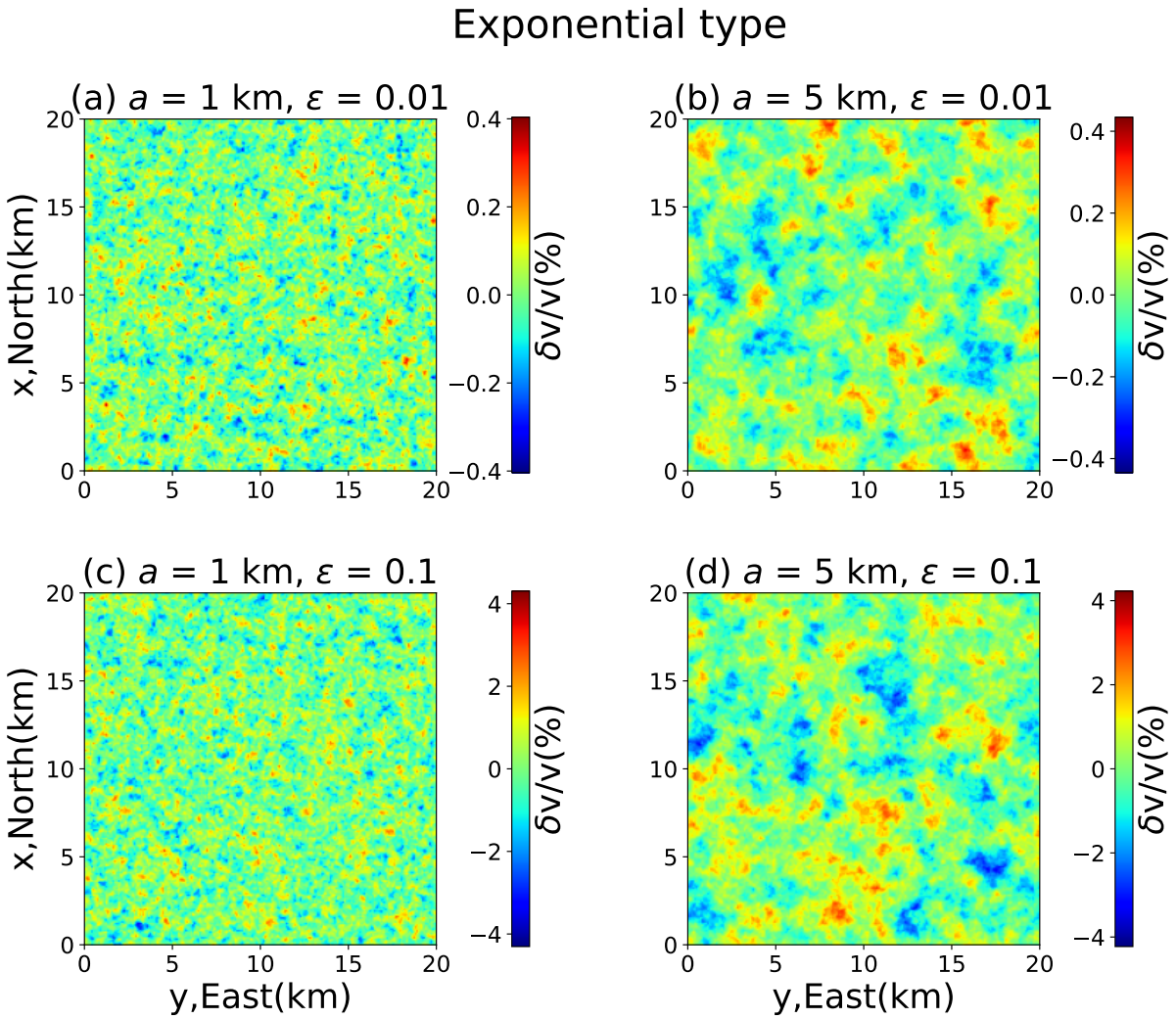



**Fig. 2.2:** The cross sections of the fractional fluctuation $\xi(\mathbf{x})$ of the velocity in the 3D random media with Exponential type : (a)$a = 1$ km and $\varepsilon = 0.01$, (b)$a = 5$ km and $\varepsilon = 0.01$, (c)$a = 1$ km and $\varepsilon = 0.1$, and (d)$a = 5$ km and $\varepsilon = 0.1$. Note the color scale of them is different.

where $\alpha_0$ and $\beta_0$ are the mean velocities of P-waves and S-waves, respectively. The subscript $P$ or $S$ indicate the energy type of the energy density. $g_0^{V\to W}$ denotes the total scattering coefficient which is the total probability of scattering a wave of type $V$ to type $W$ per unit length of propagation path. $Q_P$ and $Q_S$ are the intrinsic quality factors of P- and S-waves, respectively and $\omega$ is the angular frequency. The scattering coefficient $g^{V\to W}(\mathbf{n},\mathbf{n}')$ defines the probability of a wave of type $V$ incident in direction $\mathbf{n}'$ to be scattered into the outgoing direction $\mathbf{n}$ as type $W$. For isotropic random media $g^{V\to W}(\mathbf{n},\mathbf{n}') = g^{V\to W}(\theta)$ where $\theta$ is the scattering angle between $\mathbf{n}$ and $\mathbf{n}'$. The expressions of the 2D scattering coefficients are derived in Appendix A. They differ from the coefficients given by Przybilla et al. (2006) in some powers of the $c_p/c_s$ velocity ratio $\gamma$ and a factor of $2\pi$ in the definition of the total scattering coefficients. The total scattering coefficient is defined as:

$$g_0^{V\to W} = \int_{2\pi} g^{V\to W}(\theta)\, d\theta \,. \tag{2.10}$$

which is the integral overall scattering angles and the reciprocal of the mean free path which determines the total probability of scattering. More discussion is found in Appendix D5 of Sens-Schönfelder et al. (2009).

In this thesis, I focus on 2D infinite in-plane elastodynamics in which the S-wave is polarized in the plane of propagation (SV-wave). In this setting, SH-waves are polarized normal to the plane of propagation and are not coupled to P or SV-waves. They can thus be treated in the acoustic approximation.

## 2.4 Spatially Variable Heterogeneity and Attenuation

Sens-Schönfelder et al. (2009) describe an algorithm to simulate multiple non-isotropic scattering of elastic waves at randomly distributed heterogeneities. However, they do not consider spatially variable scattering and intrinsic attenuation properties. With elastic radiative transfer equations and Monte-Carlo method, we simulate a 2-D multiple non-isotropic scattering process in spatially heterogeneous media to model the complete waveform envelope with the specific energy density. The specific energy density contains the information about the angular dependence which has not been considered before.

The radiative transfer equations are normally used in spatially homogeneous random media in which the power spectrum of the fluctuation is independent of location. In accord with Takeuchi (2016), a more general situation is considered here , in which the standard deviation of the fluctuations $\varepsilon$ is a function of location ($\varepsilon(r)$), while all other parameters of the fluctuations remain constant. For a given wavenumber $|\mathbf{m}|$, the local power spectral density function (PSDF) of random media is rewritten as

$$\Phi\left(|\mathbf{m}|, \varepsilon^2(\mathbf{r})\right) = \frac{\varepsilon^2(\mathbf{r})}{\varepsilon_0^2}\Phi\left(|\mathbf{m}|, \varepsilon_0^2\right) \tag{2.11}$$

where $\varepsilon_0$ is a background value that is homogeneous in space. $\Phi\left(|\mathbf{m}|, \varepsilon_0^2\right)$ is the PSDF of the background fluctuation $\varepsilon_0$. Note that perturbing variance of the fluctuations only affects proportionally the magnitude of the scattering coefficient $g^{V\to W}(\theta)$ but not its angular distribution. In consequence the total scattering coefficients are now functions of the position $g_0^{V\to W}(\varepsilon^2(\mathbf{r}))$.

Spatial variations of intrinsic attenuation may be incorporated similarly by replacing $Q_V$ with $Q_V(\mathbf{r})$.

The right-hand side in eq. (2.8) and eq. (2.9) may accordingly be rewritten as:

$$\begin{aligned} &- \left( \alpha_0 g_0^{P\to P}(\varepsilon^2(\mathbf{r})) + \alpha_0 g_0^{P\to S}(\varepsilon^2(\mathbf{r})) + \frac{\omega}{Q_P(\mathbf{r})} \right) E_P\left(\mathbf{r},\mathbf{n},t\right) \\ &+ \int_{2\pi} \alpha_0 g^{P\to P}\left(\theta,\varepsilon^2(\mathbf{r})\right) E_P\left(\mathbf{r},\mathbf{n}',t\right) d\mathbf{n}' \\ &+ \int_{2\pi} \beta_0 g^{S\to P}\left(\theta,\varepsilon^2(\mathbf{r})\right) E_S\left(\mathbf{r},\mathbf{n}',t\right) d\mathbf{n}' \end{aligned} \tag{2.12}$$

and

$$\begin{aligned} &- \left( \beta_0 g_0^{S\to S}(\varepsilon^2(\mathbf{r})) + \beta_0 g_0^{S\to P}(\varepsilon^2(\mathbf{r})) + \frac{\omega}{Q_S(\mathbf{r})} \right) E_S\left(\mathbf{r},\mathbf{n},t\right) \\ &+ \int_{2\pi} \beta_0 g^{S\to S}\left(\theta,\varepsilon^2(\mathbf{r})\right) E_S\left(\mathbf{r},\mathbf{n}',t\right) d\mathbf{n}' \\ &+ \int_{2\pi} \alpha_0 g^{P\to S}\left(\theta,\varepsilon^2(\mathbf{r})\right) E_P\left(\mathbf{r},\mathbf{n}',t\right) d\mathbf{n}' \ . \end{aligned} \tag{2.13}$$

To solve the radiative transfer equations in the presence of spatially variable medium properties we implemented a Monte-Carlo algorithm that is detailed in Chapter 3.

## 2.5 Acoustic Case

We use the acoustic radiative transfer equation in 2D which is written as:

$$\begin{aligned} \left( \frac{\partial}{\partial t} + \mathbf{n}\alpha_0 \cdot \nabla \right) E\left(\mathbf{r},\mathbf{n},t\right) = &- \left( \alpha_0 g_0(\varepsilon^2(\mathbf{r})) + \frac{\omega}{Q(\mathbf{r})} \right) E\left(\mathbf{r},\mathbf{n},t\right) \\ &+ \int_{2\pi} \alpha_0 g\left(\mathbf{n},\mathbf{n}',\boldsymbol{\varepsilon}^{\mathbf{2}}(\mathbf{r})\right) E\left(\mathbf{r},\mathbf{n}',t\right) d\mathbf{n}' \ . \end{aligned} \tag{2.14}$$

The acoustic scattering coefficient is given by Wegler et al. (2006):

$$g\left(\theta,\varepsilon^2(\mathbf{r})\right) = k_0^3 \Phi\left(2k_0 \sin\frac{\theta}{2}, \varepsilon^2(\mathbf{r})\right) \tag{2.15}$$

where $k_0$ is the wavenumber and $\theta$ is the scattering angle between the directions $\mathbf{n}$ and $\mathbf{n}'$. $\alpha_0$ indicates the mean background velocity. $\Phi$ is the local power spectral density function (PSDF) of the parameter fluctuations in the random medium which is of exponential type in this paper:

$$\Phi(\mathbf{k},\varepsilon^2(\mathbf{r})) = \frac{2\pi a^2 \varepsilon^2(\mathbf{r})}{(1+a^2\mathbf{k}^2)^{3/2}} \ , \tag{2.16}$$

with the wave vector $\mathbf{k}$. This algorithm allows us to simulate the acoustic wave propagating in the spatially variable model $\mathbf{m} = (\varepsilon^2(\mathbf{r}), Q^{-1}(\mathbf{r}))$ that can consists of the two parts $\mathbf{m}^{\varepsilon} = \varepsilon^2(\mathbf{r})$ and $\mathbf{m}^{Q} = Q^{-1}(\mathbf{r})$ describing the scattering and attenuation properties, respectively.

# Chapter 3

# Monte-Carlo Simulation

For numerically solving the radiative transfer equations we use the Monte-Carlo method. The idea of the Monte-Carlo method is based on the concept of wave packets or seismic phonons that carry information about the wave energy but neglect phase information. The propagation of these phonons is governed by ray theory during unperturbed propagation and by the scattering coefficients upon the occurrence of scattering events. Large numbers of particles are propagated through the model to obtain a smooth representation of the energy distribution. The wave energy is represented by the number density of particles $N(\mathbf{r}, \mathbf{n}, t)$, their propagation mode and their weight, which decays during propagation due to intrinsic attenuation.

The left-hand side of eq. (2.8) is the material derivative of the specific energy density around the propagation direction $\mathbf{n}$. For clarity, we recall that the material derivative is the rate of change of a physical quantity followed on its path. With the Monte-Carlo method, it is modeled by the change of the number of particles tracked during a time interval $dt$ along their propagation path of length $\Delta l$ such that:

$$\Delta l = dl = v dt \; . \tag{3.1}$$

Defining $\Delta N_P(\mathbf{r}, \mathbf{n}, t) = N_P(\mathbf{r} + \mathbf{n}\Delta l_P, \mathbf{n}, t + dt) - N_P(\mathbf{r}, \mathbf{n}, t)$, we rewrite the radiative transfer equation for a P-phonon:

$$\begin{aligned} \Delta N_P(\mathbf{r}, \mathbf{n}, t) = &- \Delta l_P \left( g_0^{P\to P}\left(\varepsilon^2(\mathbf{r})\right) + g_0^{P\to S}\left(\varepsilon^2(\mathbf{r})\right) + \frac{\omega}{\alpha_0 Q_P(\mathbf{r})} \right) N_P(\mathbf{r}, \mathbf{n}, t) \\ &+ \Delta l_P \int_{2\pi} g^{P\to P}\left(\theta, \varepsilon^2(\mathbf{r})\right) N_P(\mathbf{r}, \mathbf{n}', t) \\ &+ \Delta l_S \int_{2\pi} g^{S\to P}\left(\theta, \varepsilon^2(\mathbf{r})\right) N_S(\mathbf{r}, \mathbf{n}', t) \; . \end{aligned} \tag{3.2}$$

The analogous expression for S-phonons is obtained by interchanging the indices P and S. Eq. (3.2) expresses the change of particle numbers as they propagate during a time $dt$ through the medium. The first term on the right hand describes the number of particles lost by the scattering and intrinsic attenuation per unit distance. The second and third terms describe the increase of particle numbers by scattering from all other directions into direction $\mathbf{n}$ of P-wave and S-wave, respectively.

The strategy of the MC simulation is to calculate $N_V(\mathbf{r}, \mathbf{n}, t)$ not sequentially in time for the whole spatial domain as done in time domain wavefield simulations, but rather propagate the seismic phonons one by one through the domain and accumulate their contribution to $N_V(\mathbf{r}, \mathbf{n}, t)$. Propagating many phonons with stochastically distributed scattering events leads to an increasingly better estimation of $N_V(\mathbf{r}, \mathbf{n}, t)$.

Figure 3.1 shows a flowchart of the Monte-Carlo algorithm. In the beginning, a particle is initialized and launched at the source with either P or S mode and a take-off direction. We use an isotropic source

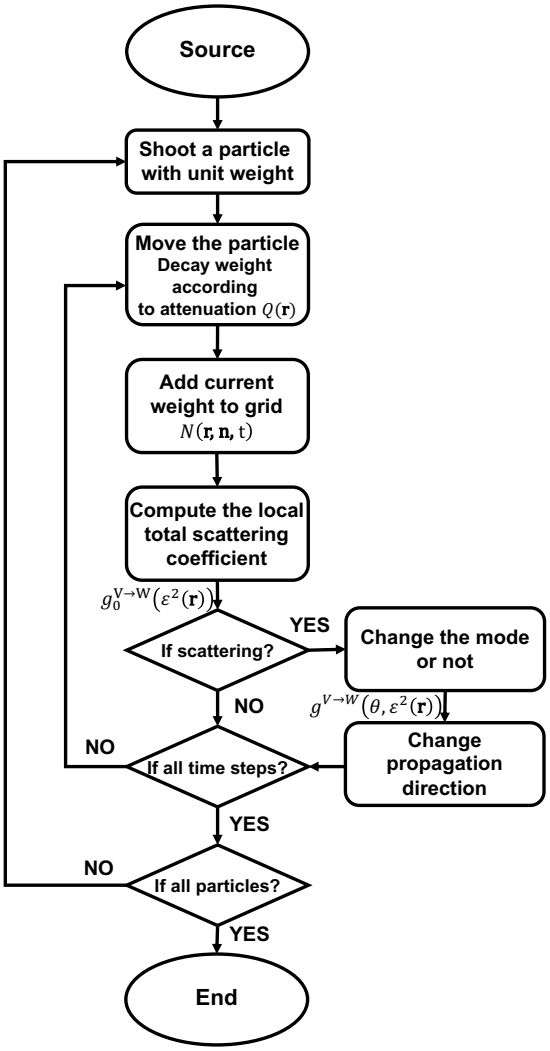



**Fig. 3.1:** Flowchart of Monte-Carlo simulation.

meaning that the initial direction is drawn from a uniform distribution. The propagation mode follows a Bernouilli probability law with the parameter: fraction of $P$ (or equivalently $S$) energy released at the source. The particle is then moved in this direction at the speed corresponding to its mode for one time increment. At the new position, $N_V(\mathbf{r}, \mathbf{n}, t)$ is updated with the weight of the present phonon and the local total scattering coefficient is evaluated. This coefficient determines whether scattering occurs or not. This is achieved by simulating a Bernouilli random variable with parameter $g^V dl$, where $V$ is the propagation mode, $dl$ the path length travelled by the particle during $dt$ and $g^V = \sum_W g^{V\to W}$.

If no scattering occurs and the total simulation time is not reached yet, the particle continues to move with the previous propagation direction. If scattering occurs, the scattering coefficients are used to determine which type of scattering occurs, i.e., if the mode of the particle is converted or remains unchanged. This is done again by simulating a Bernouilli random variable with the parameter: conversion rate from mode $V$ to mode $W$.

The scattering coefficient of the selected scattering type determines the probability distribution for selecting the scattering angle and consequently the new propagation direction of the particle. The selection of the angle uses a trial and reject method to generate angles with the required probability distributions (Sens-Schönfelder et al., 2009). This cycle will be repeated for all particles. Since we recorded the position in each time step for all particles, we obtain the number of particles $N(\mathbf{r}, \mathbf{n}, t)$

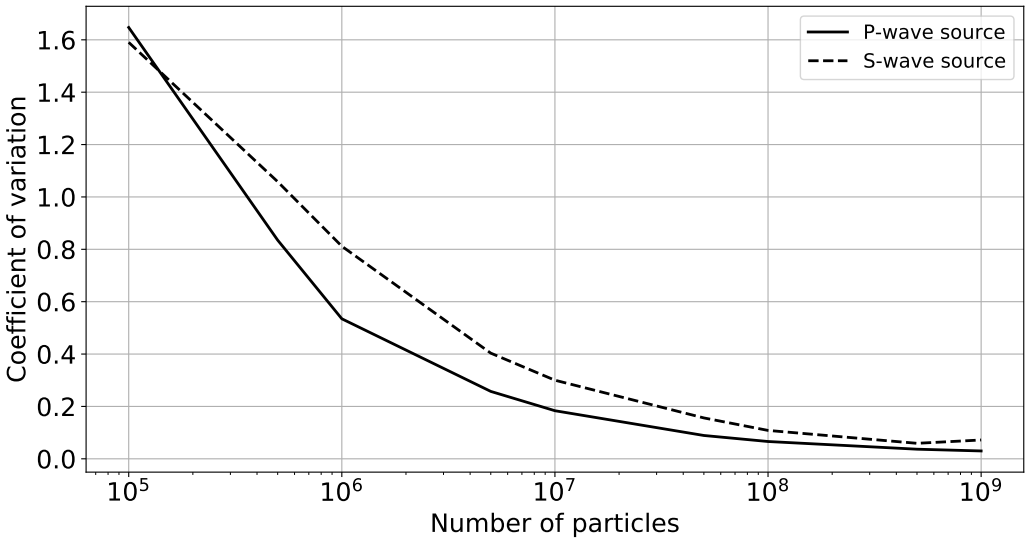


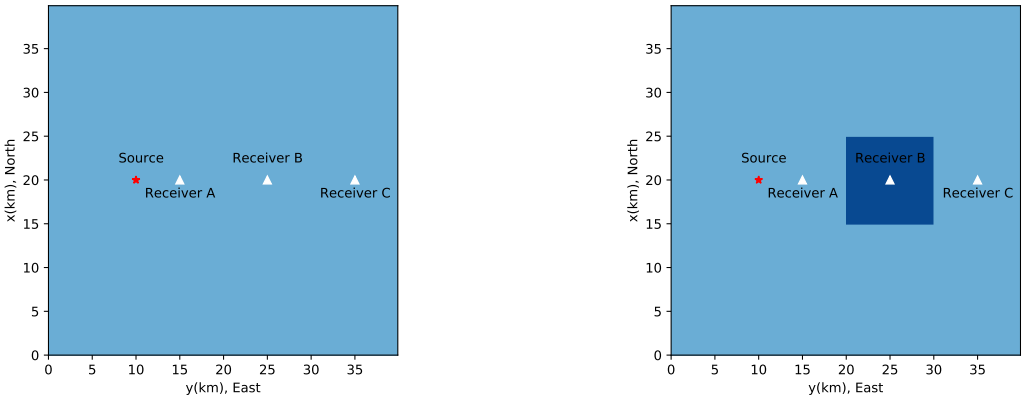

**Fig. 3.2:** Coefficient of variation changed with different number of particles.

moving in direction $\mathbf{n}$ at the position $\mathbf{r}$ at any time $t$. Actually we also record the propagation direction through the simulation to obtain $N(\mathbf{r}, \mathbf{n}, t)$.

As a stochastic method the Monte-Carlo simulations converge to a stable solution but exhibit fluctuations. These fluctuations determine the accuracy of the energy density estimates and we cannot hope to resolve structure in the medium when its effect on the envelopes is smaller than the fluctuations from the Monte-Carlo solution. However, the fluctuations of the solution are controlled by the number of particles used in the simulation. We provide a test for the amplitude of the fluctuations as a function of the number of particles in Figure 3.2. It shows the coefficient of variation, i.e. the ratio of standard deviation and mean value obtained for 24 identical simulations in the uniformly heterogeneous model. Figure 3.2 shows that the coefficient of variation decreases with increasing number of particles. $10^8$ particles are used in the simulations of this paper for which the coefficient of variation is about 5%.

## 3.1 Modelling with Spatially Variable Heterogeneity and Attenuation

We build a model with 2-D space and time on a $400 \times 400 \times 100$ grid. The size of each grid cell is 0.1 km $\times$ 0.1 km $\times$ 0.1 s. The two illustrations in Figure 3.3 show three sets of simulations: (1) baseline simulation with uniform $\varepsilon = 0.05$ and no intrinsic attenuation, (2) scattering anomaly for fluctuation strength $\varepsilon = 0.09$ and (3) attenuation anomaly for intrinsic quality factor $Q_P^{-1} = 0.17, Q_S^{-1} = 0.1$. The values of $Q_P^{-1}$ and $Q_S^{-1}$ assumed here imply much stronger attenuation than typically found in the lithosphere. The background velocity of all models is $\alpha_0 = 6$ km/s, $\beta_0 = 3.46$ km/s. The density is $\rho = 2.7$ g cm$^{-3}$and the correlation length is $a = 0.3$ km. Random fluctuations of velocity and density in this study are always of exponential type. With a length scale of 10 km the anomalous region is significantly larger than the length scale of the random fluctuations $a$, ensuring a meaningful statistical description of the fluctuations. As discussed by Cormier & Sanborn (2019) the determination of $a$ is subject to a trade-off with $\varepsilon$. The angular frequency in this study is 30 rad/s so the frequency is 4.78 Hz adapted to observations of high frequency wave scattering in the crust. Resulting wavelengths of P-wave and S-wave are 1.257 km and 0.726 km. In the Monte-Carlo Method, 100 million particles are




**(a)** homogeneous **(b)** with anomaly

**Fig. 3.3:** Illustrations of the model setup. (a) Homogeneous model with background $\varepsilon = 0.05$ and $Q_p^{-1} = Q_s^{-1} = 0$ (simulation 1) and (b) Anomalous model with $\varepsilon = 0.09$ (simulation 2) or intrinsic quality factors $Q_p^{-1} = 0.17, Q_s^{-1} = 0.1$ inside the anomaly (simulation 3). The background velocity of all models is $V_p = 6$ km/s, $V_s = 3.46$ km/s. The background density is $\rho = 2.7$ g cm$^{-3}$. The correlation length is $a = 0.3$ km in all simulations. The red star indicates the source and three white triangles indicate receivers that are located before, within and behind the anomaly as seen from the source.

used to simulate the energy transport in these 2-D models.

### 3.1.1 Scattering Anomaly Simulation

Results of the simulation in the homogeneously scattering medium and scattering anomaly medium are shown for pure P-wave and pure S-wave source with isotropic radiation pattern in Figures 3.4 and 3.5, respectively. Firstly, we analyze the results with P source. Figures 3.4a and 3.4b show the snapshots of the P and S energy density at lapse time of 2 s and 5 s in the homogeneous medium and the scattering anomaly medium with P source, respectively. There is obviously a reduction of the ballistic energy (P energy) and an increase of the scattered energy (S energy) where there is a stronger scattering area. An apparent ballistic S-wave can be seen emanating from the source in Figure 3.4a(S energy). It also shows the S-coda of the ballistic P-wave front. Strong S-wave generation inside the scattering anomaly is also apparent in Figure 3.4b(S energy). In particular there is a secondary S-wave generated at the time when the P-wave hits the anomaly. Furthermore, there is an obvious Mach cone emanating from the edges of the anomaly. Because the velocity of P-waves is larger than that of the S-wave the speed of the source of scattered S-waves is faster than their propagation. To examine in greater details the energy conversion in the scattering process, we subtract the results of the simulation in the homogeneous model from the results of the simulation in the scattering anomaly model. The difference is shown in Figure 3.4c. Before the P-wave hits the anomaly area, there should be no difference. Since the initial directions of all particles are random, small fluctuations are observed but can be ignored. With increasing lapse time, the energy conversion is more and more obvious. In addition, the envelopes at three receivers for the homogeneous model(dotted) and the anomalous model(solid) are compared in Figure 3.6a. Since receiver A (corresponding to the red curves) is located closer to the source than the anomaly area, the ballistic waves show no difference between the two models but with the lapse time increasing, the energy is higher in the anomaly model due to back-scattering. Receiver B (shown by blue curves) is in the centre of the anomaly. There is obviously more S energy when the wave arrives. We can see strong differences in S energy between the two models also at receiver C (shown by green curves). The $S$ energy

in the scattering anomaly peaks in between the arrival times of the ballistic $P$ and apparently ballistic $S$ arrivals, confirming that the origin of this extra pulse of $S$ energy is located in the anomaly.

In Figure 3.5 we show the results of the simulation in the case of an S-wave source. There are many similar phenomena that can be seen in the snapshots of simulated energy densities as compared to the case of a P source. However, no Mach cone can be seen in Figure 3.5b(P energy) because the velocity of S-wave is less than that of P-wave. Besides, there exist some cells where the direct S-wave has not arrived but that already contain P energy. This can be explained by the fact that a fraction of P-wave energy has been converted from the S-wave source and forms a precursor to the ballistic S wave. We can also see that there is stronger scattering when the direct S-wave arrives at the anomaly area in Figure 3.5c. The energy envelopes for a pure S source are shown in Figure 3.6b. Notice that for the S energy at the receiver B (shown by blue curves) there is some energy arriving before the direct S wave (at a lapse-time of about three seconds). This illustrates the S-to-P energy conversions that form precursors to the direct S wave. It can also be observed in Figure 3.6b(P energy).

### 3.1.2 Intrinsic Attenuation Anomaly Simulation

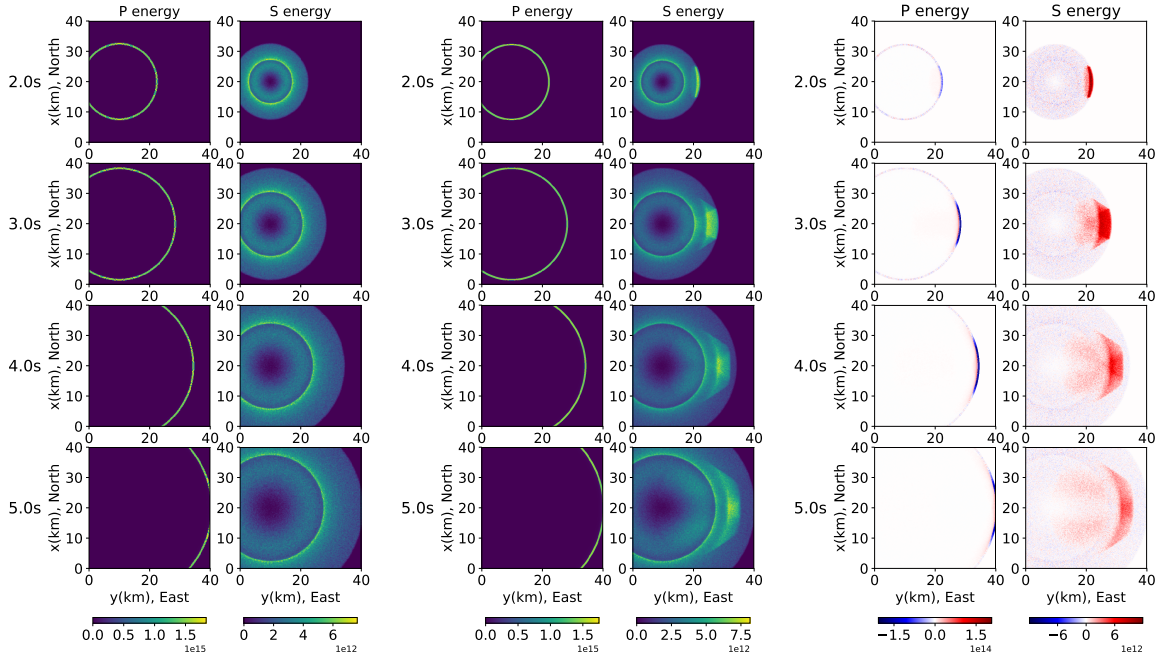
Energy lost through the propagation is normally due to scattering and intrinsic attenuation. In Section 3.1.1 we discussed the scattering simulation results in a scattering anomaly model without intrinsic attenuation. We are also able to simulate the wave scattering in a specific area with intrinsic attenuation. In order to show the attenuation more obviously, we select an anomaly with a high value of intrinsic inverse quality factors for both P-wave $Q_P^{-1}$ and S-wave $Q_S^{-1}$ and background RMS fluctuations $\varepsilon = 0.05$ as in the background medium. The location of the anomaly is the same as Figure 3.3b. Since $Q_P^{-1} = 0.17, Q_S^{-1} = 0.1$, the intrinsic attenuation is so strong in our model that a gap in the energy field develops as the ballistic waves propagate through the strongly absorbing anomaly (see Figure 3.7). Furthermore, both P-wave energy and S-wave energy are rapidly absorbed as time increases. The exceptionally low $Q$-values in the anomalous region are for illustration purposes.

### 3.1.3 Irregular Anomaly Simulation

In Section 3.1.1 and Section 3.1.2, the shape of the anomaly is a square as shown in Figure 3.3b. The grids we employ offer the possibility to models with various spatial distributions of heterogeneity. Here we build a scattering anomaly model with irregular shapes as shown in Figure 3.8. Note that there is no intrinsic absorption in this model and the background RMS fluctuations $\varepsilon$ is 0.05. The light blue and dark blue areas indicate, respectively, regions with $\varepsilon$ of 0.02 and 0.09. The source is located in the center with $x = 20$ km, $y = 20$ km. The simulation results with P source and S source are shown in Figure 3.9 and Figure 3.10. Compared with the uniform model the regions with perturbed fluctuation strength leave a clear imprint on the energy field, mostly by increased/decreased conversion of ballistic energy into scattered energy by larger/smaller fluctuation strength.

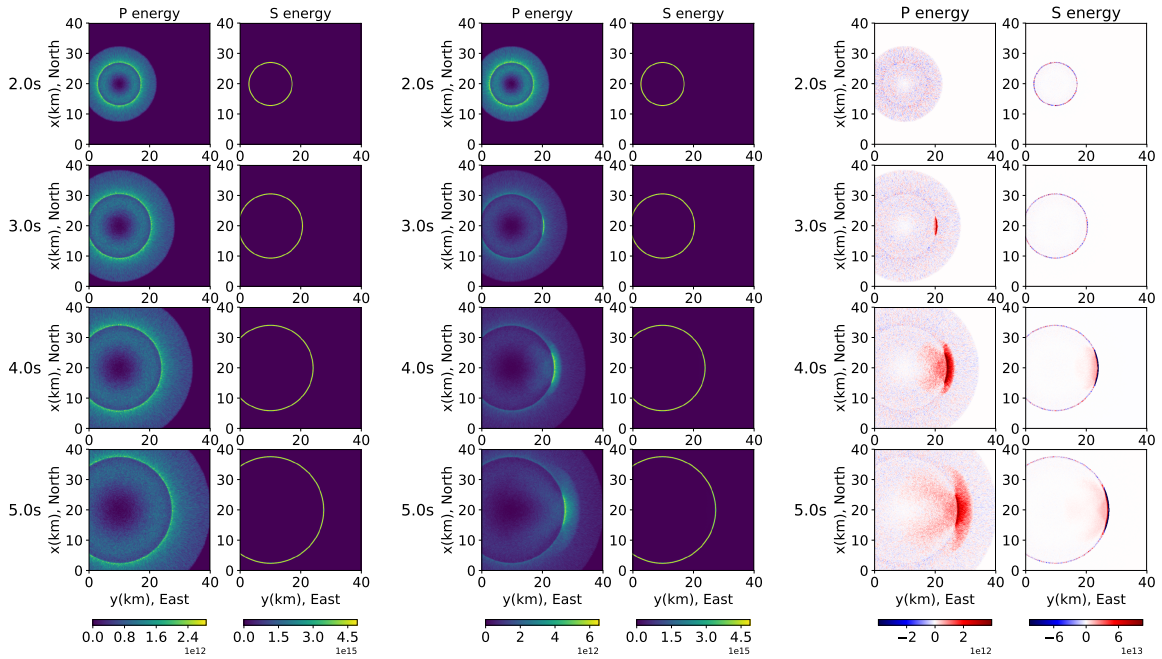
## 3.2 Modelling the Specific Energy Density

With radiative transfer theory, the wave energy propagation is encapsulated in the specific energy density $E_{YX}(\mathbf{r}, \mathbf{n}, t)$ where $YX$ indicate the mode of excitation ($X$) and recording ($Y$). The benefit of this description, is that we have access to the complete information about the energy distribution at any position $\mathbf{r}$ including its time dependence and propagation directions $\mathbf{n}$. Figure 3.11 shows the specific energy density of a simulation in a uniform medium at lapse times $t = 2$ s and $t = 4$ s in the propagation direction $az(\mathbf{n}) = 45°$ and $az(\mathbf{n}) = 90°$. $az$ defines the azimuth of the propagation direction measured clockwise from north. In the following we omit the $az(\cdot)$ notation for brevity as it is clear from the



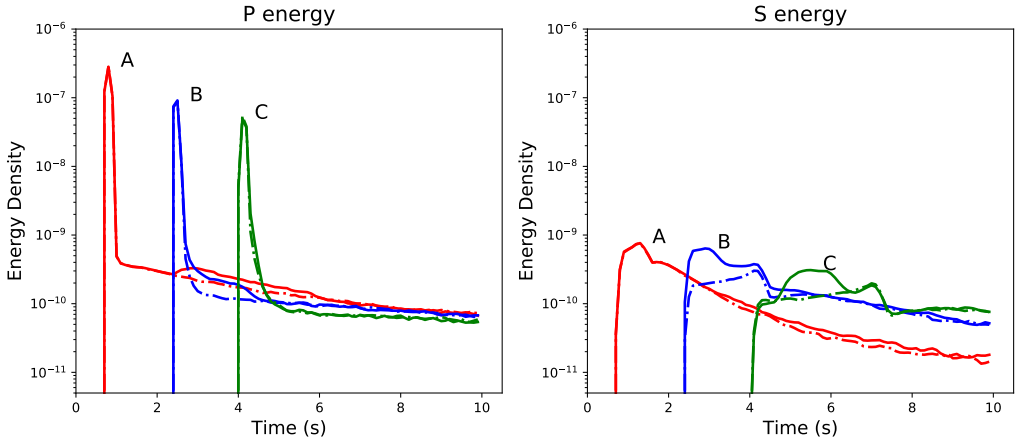

**(a)** uniform medium **(b)** medium with scattering anomaly **(c)** energy differences

**Fig. 3.4:** Snapshots ($2s - 5s$) of the simulated energy field in (a) the uniform medium and (b) the scattering anomaly medium. (c) differences between (a) and (b). The source emits pure P-wave energy. Both the P energy and the S energy are recorded.

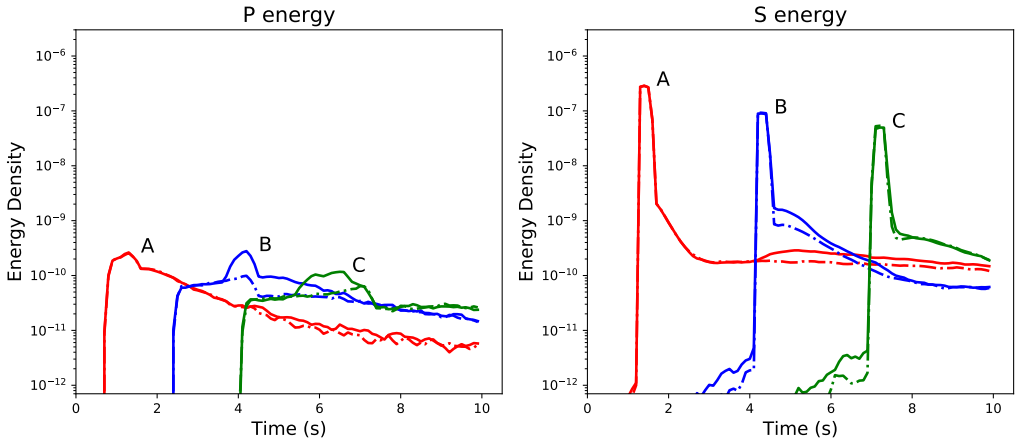



**(a)** uniform medium **(b)** medium with scattering anomaly **(c)** energy differences

**Fig. 3.5:** Snapshots ($2s - 5s$) of the simulated energy field in (a) the uniform medium and (b) the scattering anomaly medium. (c) differences between (a) and (b). The source emits pure S-wave energy. Both the P energy and the S energy are recorded. Note that the maximum of plot scale for the S-wave energy in the differences is 20% of the real maximum.

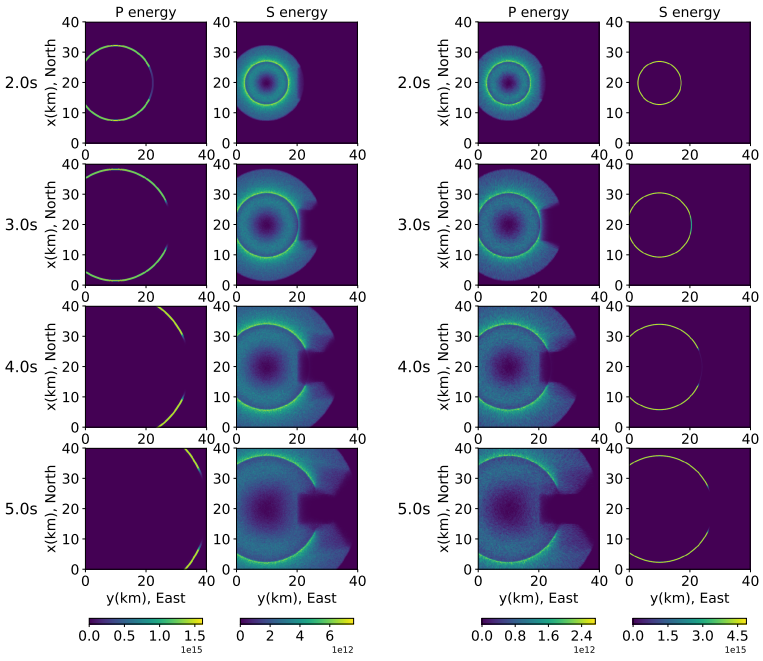



**(a)** The source emits pure P-wave energy.




**(b)** The source emits pure S-wave energy.

**Fig. 3.6:** Envelopes at three receivers for the uniform medium (dotted) and the scattering anomaly medium (solid). The red, blue and green curves indicate the energy that arrives at the receiver A, B and C respectively. Both the P energy and the S energy are recorded.




**(a)** P source **(b)** S source

**Fig. 3.7:** Snapshots ($2s-5s$) of the simulated energy field in the model with the anomaly in intrinsic attenuation. The intrinsic quality factors $Q_p^{-1}$ and $Q_s^{-1}$ are 0.17 and 0.1, respectively. The source is (a) P-wave and (b) S-wave. Both the P energy and the S energy are recorded.

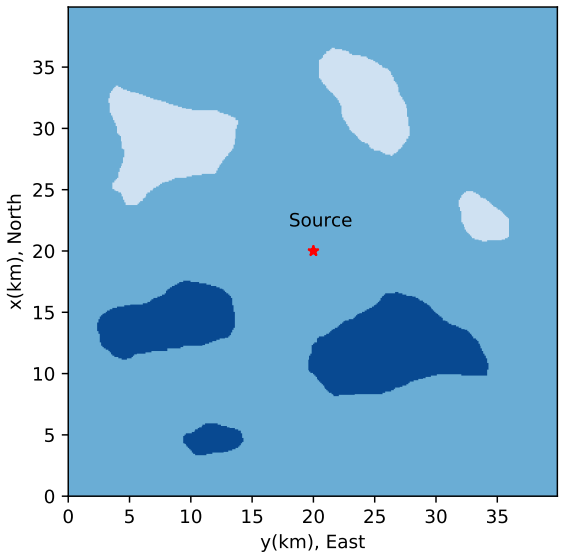
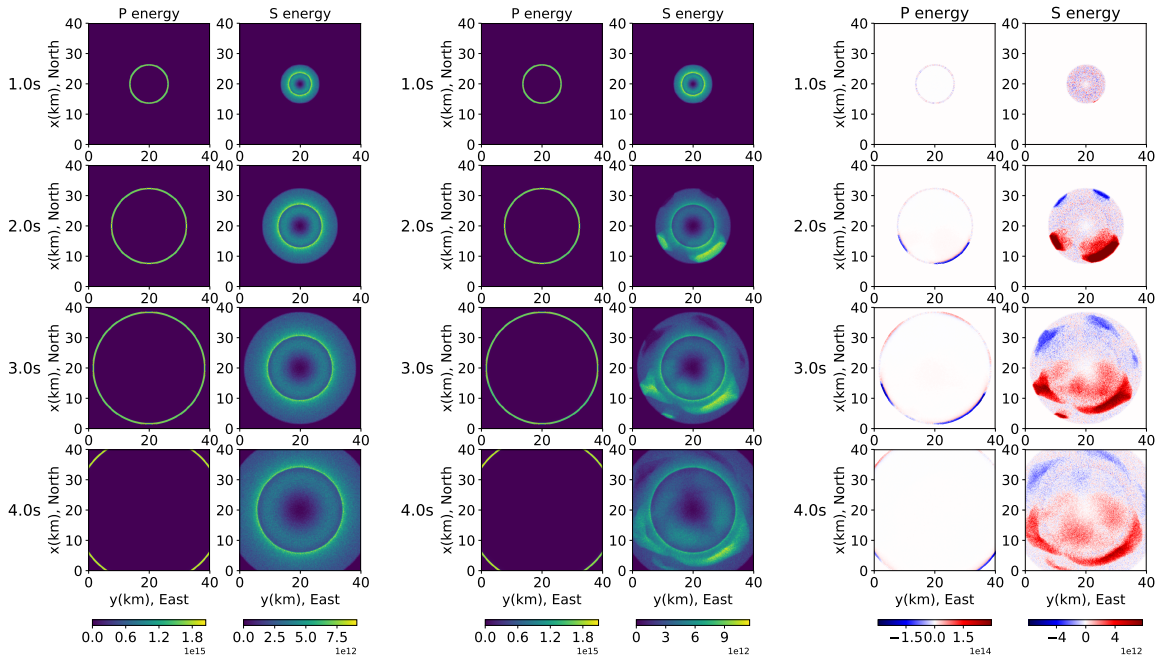
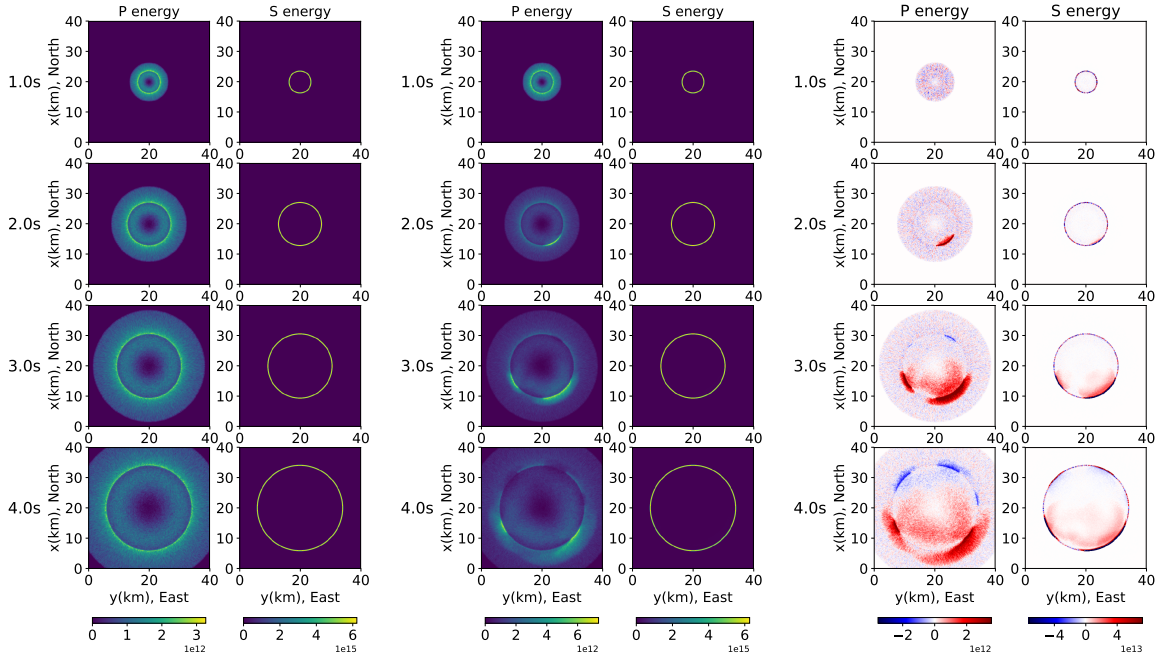



**Fig. 3.8:** Illustrations of the source (red star) in the irregular scattering anomaly model. The background value of $\varepsilon$ is 0.05. The light blue and the dark blue colors indicate areas with $\varepsilon = 0.02$ and 0.09, respectively.




**(a)** uniform medium **(b)** medium with scattering anomaly **(c)** energy differences

**Fig. 3.9:** Snapshots $(1s - 4s)$ of the simulated energy field in (a) the uniform medium and (b) the medium with the irregular scattering anomaly and (c) the differences between (a) and (b). The source emits P-wave energy.




**(a)** uniform medium **(b)** medium with scattering anomaly **(c)** energy differences

**Fig. 3.10:** Snapshots $(1s - 4s)$ of the simulated energy field in (a) the uniform medium and (b) the medium with the irregular scattering anomaly and (c) the differences between (a) and (b). The source emits S-wave energy. Note that the maximum of plot scale for the S-wave energy in the differences is 20% of the real maximum.

context when the azimuth of the direction $\mathbf{n}$ is referred to. The parameters of the medium are the same as discussed before. The RMS fluctuation is $\varepsilon = 0.05$ and there is no intrinsic attenuation. Here we employ a slightly coarser grid to reduce the numerical fluctuations inherent to Monte-Carlo simulations. We model in 2-D space and time on a $100 \times 100 \times 100$ grid and the size of each grid cell is 0.5 km $\times$ 0.5 km $\times$ 0.1 s. The propagation direction is recorded in 72 non-overlapping angular bins of width 5°.

Focusing on the plot of $E_{YX}(\mathbf{r}, 90°, 4s)$, it illustrates how the energy propagating in direction $\mathbf{n} = 90°$ at lapse-time $t = 4$ s is distributed in space. Most energy is located to the east of the source which exactly reflects the scattering patterns, as detailed in Section 3.4. Because the direct wave is strong, it is difficult to observe the scattered wave energy. To enhance its visibility, we deliberately saturated the color scale in Figure 3.11.

Instead of investigating the spatial distribution of energy propagating in a specific direction it is interesting to illustrate the directional distribution of the energy propagation at a specific location. Figure 3.12 shows polar plots of the energy density as a function of propagation directions $\mathbf{n}$ at a receiver located to the east of the source at $\mathbf{r} : x = 25$ km, $y = 35$ km. The top row of Figure 3.12 illustrates an early lapse time ($t = 3.4$ s) during the arrival of the ballistic P-energy which obviously propagates in the forward direction ($E_{PP}$). Interestingly $E_{PS}$ is also peaked in the forward direction as the conversion must have happened close to the source for the energy to arrive at this early lapse time. For the same reason the scattering that generated $E_{SP}$ must have happened close to the observation point, creating a specific energy density that resembles the scattering coefficient $g^{P \to S}$. $E_{SS}$ is poorly sampled as it requires both, scattering close to the sources and close to the receiver which occurs seldom. At lapse time $t = 5.5$ s shown in the second row of Figure 3.12 the ballistic S-wave arrives, resulting in a forward peaked $E_{SS}$. Caused by the dominance of the focused conversion at the source $E_{SP}$ is also peaked in the forward direction at this lapse time. $E_{PS}$ resembles a biased version of the conversion scattering coefficient as scattering may have occurred not only close to the receiver but along an ellipse-like single scattering line around the source. At later lapse time when both ballistic waves have passed the receiver ($t = 7$ s, bottom row of Figure 3.12), the specific energy densities show complicated patterns which are far from being isotropic. This anisotropy of the energy flux density that persists for late lapse times documents the importance of considering the specific energy density for the calculation of probabilistic wavefield sensitivities.

Since the specific energy density contains the information about the amplitude of the energy flux for any combination of position, time and propagation direction, recording the specific energy density requires a large amount of memory and storage. In the simulations shown here, P- and S-wave are both recorded with double precision resulting in a memory usage of 576 MB. This is easy to handle for now but when the approach is extended to the 3-D case, the space has three dimensions and the propagation direction has two degrees of freedom which means that the memory requirements inflate. On the other hand, the Monte-Carlo method needs more particles for a sufficient description of the energy density field in 3 dimensions which translates into more CPU hours. Here we launched 100 millions particles and simulate for 15 s propagation time with 0.01 s time-step. This setup required about half an hour for each simulation on 64 cores.

## 3.3 Reciprocity Theorem

We have shown how the radiative transfer equations in the 2D elastic random media can be solved by directly simulating the scattering process. This solves the forward problem of the inversion process which requires to estimate the energy field everywhere in space as it is created by a source at $\mathbf{r_0}$. The inverse problem requires to calculate the specific energy density at the receiver location $\mathbf{r}$ excited by a source at $\mathbf{r}'$, where $\mathbf{r}'$ can be anywhere in space. Direct simulations are infeasible as they would require a separate

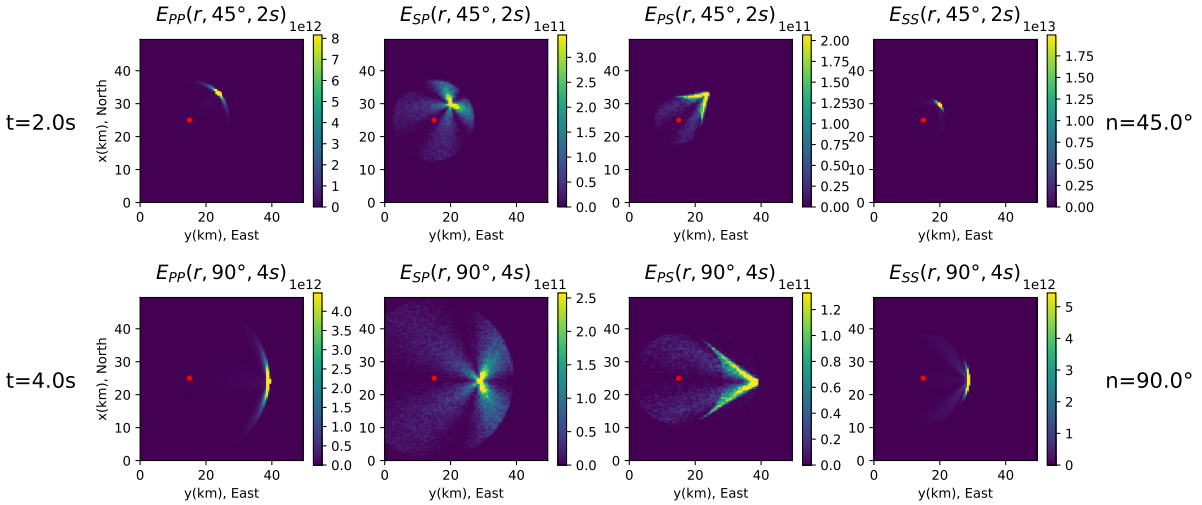



**Fig. 3.11:** Snapshots $(2s, 4s)$ of the specific energy density $E_{YX}(\mathbf{r}, \mathbf{n}, t)$ for propagation directions $\mathbf{n} = 45°$ and $\mathbf{n} = 90°$ in a uniform medium with background $\varepsilon = 0.05$. The red point indicates the source. Note that the maximum of the color scale is clipped to avoid the high values of the ballistic energy.

simulation for a source at every point in space. In waveform tomography this problem is solved using the adjoint equation which effectively allows to interchange source and receiver (Tromp et al., 2005; Fichtner et al., 2006) thereby solving the problem for all points in space with a single simulation in which the adjoint source is placed at the location of the receiver. Here we introduce reciprocity relations of the radiative transfer equation that serve the same purpose for the tomographic inversion of envelope observations. Following Margerin (2017), adjoint transport equations are introduced by:

$$\begin{aligned}\left(\frac{\partial}{\partial t} + \mathbf{n}\alpha_0 \cdot \nabla\right) E_P^\dagger(\mathbf{r}, \mathbf{n}, t) = &- \left(\alpha_0 g_0^{P\to P}(\varepsilon^2(\mathbf{r})) + \alpha_0 g_0^{\dagger P\to S}(\varepsilon^2(\mathbf{r})) + \frac{\omega}{Q_P}\right) E_P^\dagger(\mathbf{r}, \mathbf{n}, t) \\ &+ \int_{2\pi} \alpha_0 g^{P\to P}\left(\theta, \varepsilon^2(\mathbf{r})\right) E_P^\dagger(\mathbf{r}, \mathbf{n}', t)\, d\mathbf{n}' \\ &+ \int_{2\pi} \beta_0 g^{\dagger S\to P}\left(\theta, \varepsilon^2(\mathbf{r})\right) E_S^\dagger(\mathbf{r}, \mathbf{n}', t)\, d\mathbf{n}'\end{aligned} \tag{3.3}$$

$$\begin{aligned}\left(\frac{\partial}{\partial t} + \mathbf{n}\beta_0 \cdot \nabla\right) E_S^\dagger(\mathbf{r}, \mathbf{n}, t) = &- \left(\beta_0 g_0^{S\to S}(\varepsilon^2(\mathbf{r})) + \beta_0 g_0^{\dagger S\to P}(\varepsilon^2(\mathbf{r})) + \frac{\omega}{Q_S}\right) E_S^\dagger(\mathbf{r}, \mathbf{n}, t) \\ &+ \int_{2\pi} \beta_0 g^{S\to S}\left(\theta, \varepsilon^2(\mathbf{r})\right) E_S^\dagger(\mathbf{r}, \mathbf{n}', t)\, d\mathbf{n}' \\ &+ \int_{2\pi} \alpha_0 g^{\dagger P\to S}\left(\theta, \varepsilon^2(\mathbf{r})\right) E_P^\dagger(\mathbf{r}, \mathbf{n}', t)\, d\mathbf{n}'\end{aligned} \tag{3.4}$$

Different from the wave equations in FWI, radiative transfer equations describe the scattering process depending on scattering coefficients. The probability of scattering per time is given by the mean free time $1/(c_V g_0^{V\to W}(\varepsilon^2(\mathbf{r})))$. Each scattering event happening in the forward and adjoint wavefield must have the same probability leading to $c_V g^{V\to W} = c_W g^{\dagger W\to V}$. Consequently we have:

$$g^{\dagger P\to S}\left(\theta, \varepsilon^2(\mathbf{r})\right) = \frac{\beta_0}{\alpha_0} g^{S\to P}\left(\theta, \varepsilon^2(\mathbf{r})\right) \tag{3.5}$$

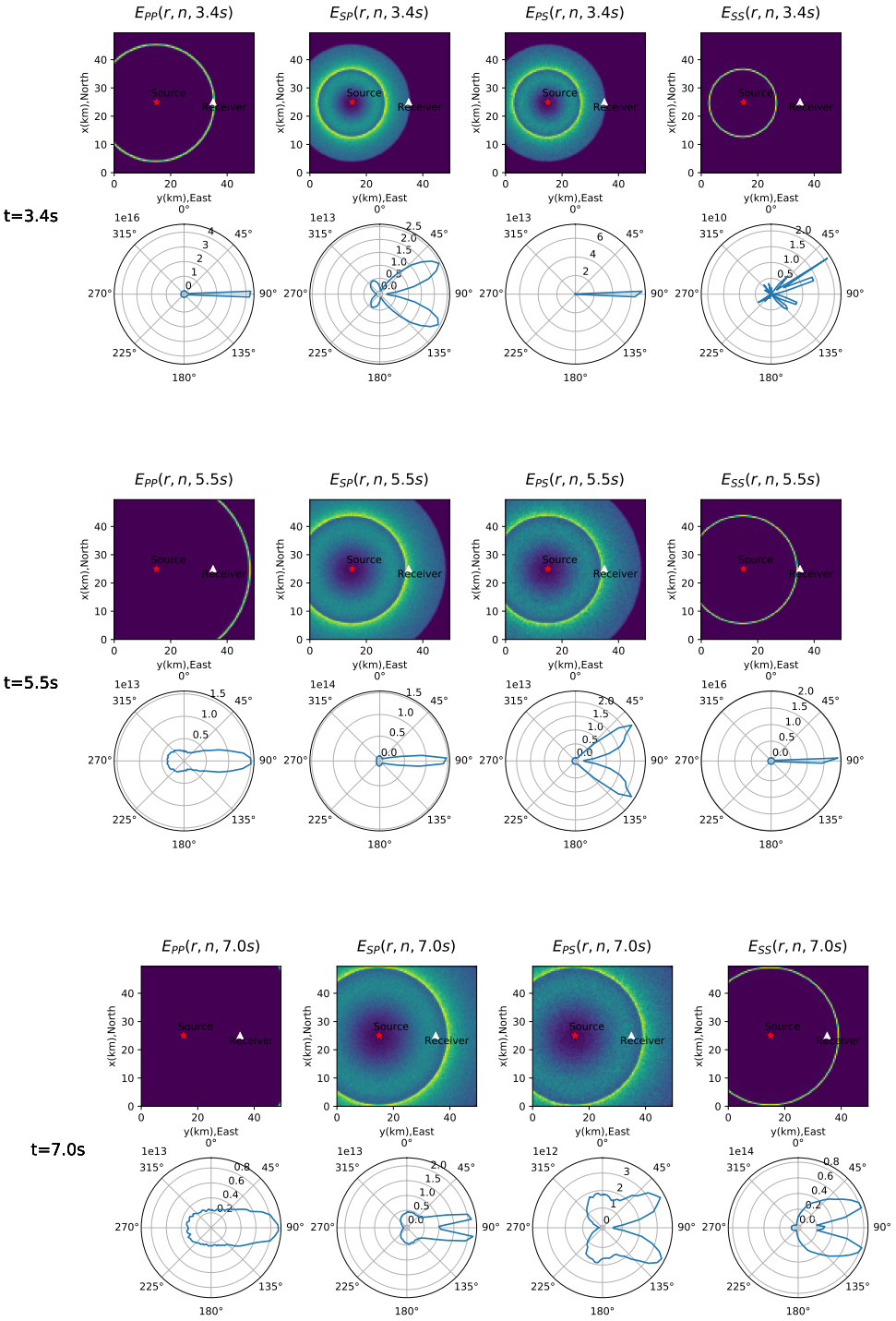



**Fig. 3.12:** Polar plots of the energy density at receiver $\mathbf{r}$: $x = 25$ km, $y = 35$ km (white triangle) as a function of propagation directions $\mathbf{n}$ in a uniform medium (background $\varepsilon = 0.05$). The red point denotes the source. The time increases from top to bottom. The corresponding snapshots of the spatial energy distribution are shown to illustrate the wave propagation. Combination of emitted ($X$) and recorded ($Y$) wave type is indicated above as $E_{X,Y}$ above each panel.

$$g^{\dagger S\to P}\left(\theta,\varepsilon^2(\mathbf{r})\right)=\frac{\alpha_0}{\beta_0}g^{P\to S}\left(\theta,\varepsilon^2(\mathbf{r})\right) \tag{3.6}$$

We note that our definition of the adjoint transport equation does not follow the one adopted in mathematical treatments. It is indeed well known that the formal adjoints of the operators $\partial/\partial_t$ and $\mathbf{n}\cdot\nabla$ are $-\partial/\partial_t$ and $-\mathbf{n}\cdot\nabla$ . However, whereas the relation (3.5) is physically essential, the sign differences for the partial derivative operators is not. It only entails notational changes in the symmetry relations between forward and adjoint intensities to be derived below. Using the representation theorem given in Margerin (2017), we gain the following reciprocity theorem:

$$E^{\dagger}_{XY}\left(\mathbf{r_1},t,-\mathbf{n_1};\mathbf{r_2},-\mathbf{n_2}\right)=E_{YX}\left(\mathbf{r_2},t,\mathbf{n_2};\mathbf{r_1},\mathbf{n_1}\right) \tag{3.7}$$

For an omnidirectional receiver at $\mathbf{r}$ which integrates the specific energy density over directions we rewrite the reciprocity relation as:

$$E^{\dagger}_{XY}\left(\mathbf{r}',t,-\mathbf{n}';\mathbf{r}\right)=\frac{1}{S^d}E_{YX}\left(\mathbf{r},t;\mathbf{r}',\mathbf{n}'\right) \tag{3.8}$$

The normalization $1/S^d$ results from the different sources and receivers on the left and right hand side. $S^d$ is the area of the unit sphere in space dimension $d$. In 2-D case, $S^d=2\pi$.

Upon noticing that $E^{\dagger}_P,\beta_0^2E^{\dagger}_S$ solve the forward transport equations with source terms $s_P,\beta_0^2s_S/\alpha^2$ provided that they solve the adjoint equation with source terms $s_P,s_S$, we deduce:

$$E_{XY}\left(\mathbf{r_1},t,-\mathbf{n_1};\mathbf{r_2},-\mathbf{n_2}\right)=const_{XY}\cdot E_{YX}\left(\mathbf{r_2},t,\mathbf{n_2};\mathbf{r_1},\mathbf{n_1}\right) \tag{3.9}$$

where we have introduced:

$$const_{XY}=\begin{cases}1 & X=Y\\ \left(\alpha_0/\beta_0\right)^2 & X=S,\ Y=P\end{cases} \tag{3.10}$$

To verify the reciprocity relations numerically, we used the anomaly model shown in Figure 3.13. The background value of the fluctuation is $\varepsilon=0.05$ and the other simulation parameters are the same as discussed above for the uniform model. However, there are two anomalous areas, one of which has stronger fluctuations (blue square, $\varepsilon=0.09$) and the other area has weaker fluctuations (light blue square, $\varepsilon=0.02$). The simulation of the forward energy field $E_{YX}\left(\mathbf{r_0},t;\mathbf{r}',\mathbf{n}'\right)$ is generated from a unit source with mode $X$ at $\mathbf{r}'$ with initial direction $\mathbf{n}'$. Since the bin width we used for recording the directional dependence of the energy density is 5°, we use the same angular range for the initial direction. The simulation of the adjoint wavefield $E^{\dagger}_{XY}\left(\mathbf{r}',t,\mathbf{n}';\mathbf{r_0}\right)$ is generated from an isotropic unit source at the receiver location $\mathbf{r_0}$ with mode $Y$ employing eqs. 3.3 and 3.4.

For the comparison of the different simulations in Figure 3.14, we select $r'$ within the strongly scattering anomaly region (c.f. Figure 3.13). Energy densities are shown for the North and South directions in Figure 3.14(a) and (b), respectively. For reference the panels in Figure 3.14 include the curves of $const_{XY}\cdot E_{YX}\left(\mathbf{r_2},t,\mathbf{n_2};\mathbf{r_1},\mathbf{n_1}\right)$ and the modes $X$ and $Y$ are given in the legend of each panel. The agreement between the three curves in each panel confirms the reciprocity relations 3.7 and 3.9. Differences between the curves can be attributed to the stochastic nature of the simulations and vanish for more accurate simulations with larger numbers of particles.

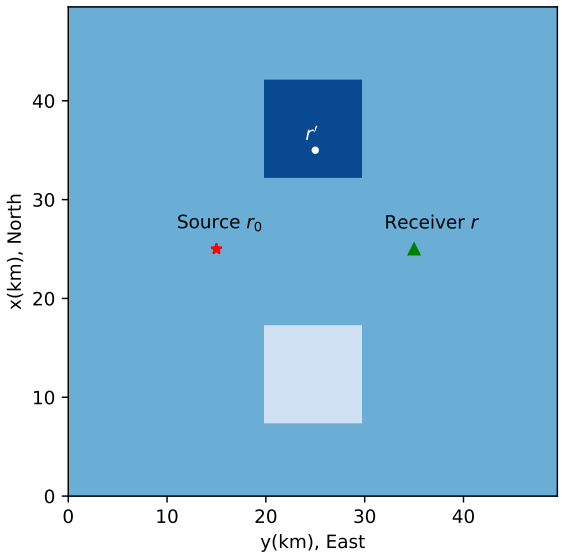



**Fig. 3.13:** Illustrations of the scattering anomaly model. The background of $\varepsilon$ is 0.05. The light blue and the dark blue colors indicate areas with $\varepsilon$ of 0.02 and 0.09, respectively. Source $\mathbf{r_0}$: $x = 25$ km, $y = 15$ km, receiver $\mathbf{r}$: $x = 25$ km, $y = 35$ km and one position $\mathbf{r}'$: $x = 35$ km, $y = 25$ km

## 3.4 Scattering Patterns

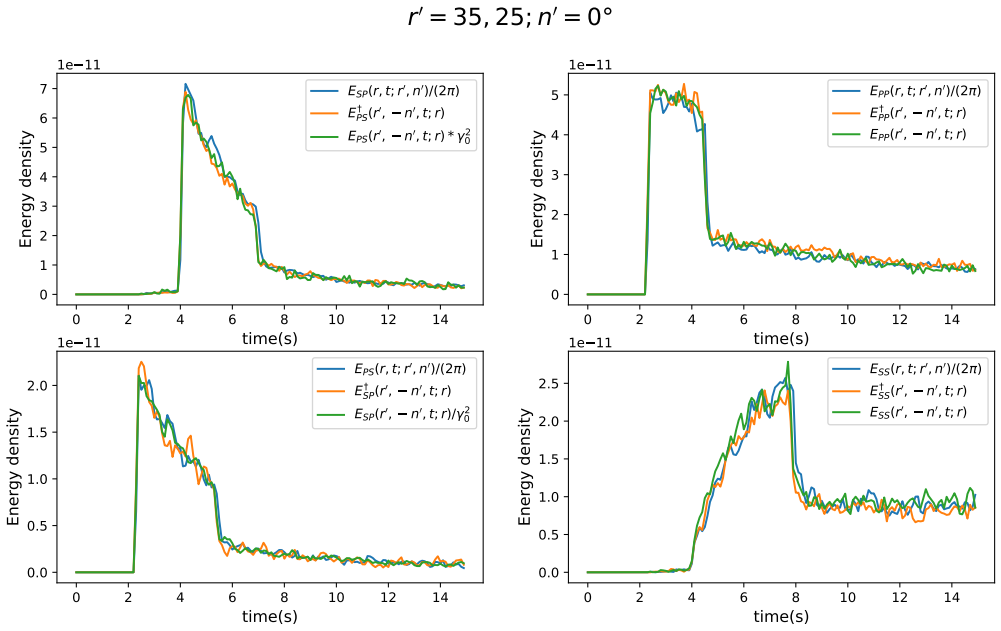
The angular probability distributions of scattering directions i.e. the scattering patterns are defined by scattering coefficients eq. (A.20). To verify the implementation in the scattering in simulation and to illustrate the functioning of the Monte-Carlo simulations, we design a special model for numerically comparing the implemented scattering coefficients with the theoretical expressions. In this model the background is homogeneous ($\varepsilon = 0$) but there is a point-like heterogeneous region with $\varepsilon = 0.05$ in the center of the model. Scattering can only occur in this region. Other parameters are as discussed in section 3.2. All particles from the source have the same initial direction towards the scatterer. Counting the particles scattered into different directions from the scatterer, we can compute the numerical scattering coefficients by normalization for the total number of particles. Note that the direct wave has the same direction as the forward scattered wave. We therefore remove the particles representing the direct waves without scattering. The result is shown for the different types of scattering in Figure 3.15 verifying that our implementation is in good agreement with the theoretical values.

## 3.5 Equipartition Ratio in Scattering Simulations

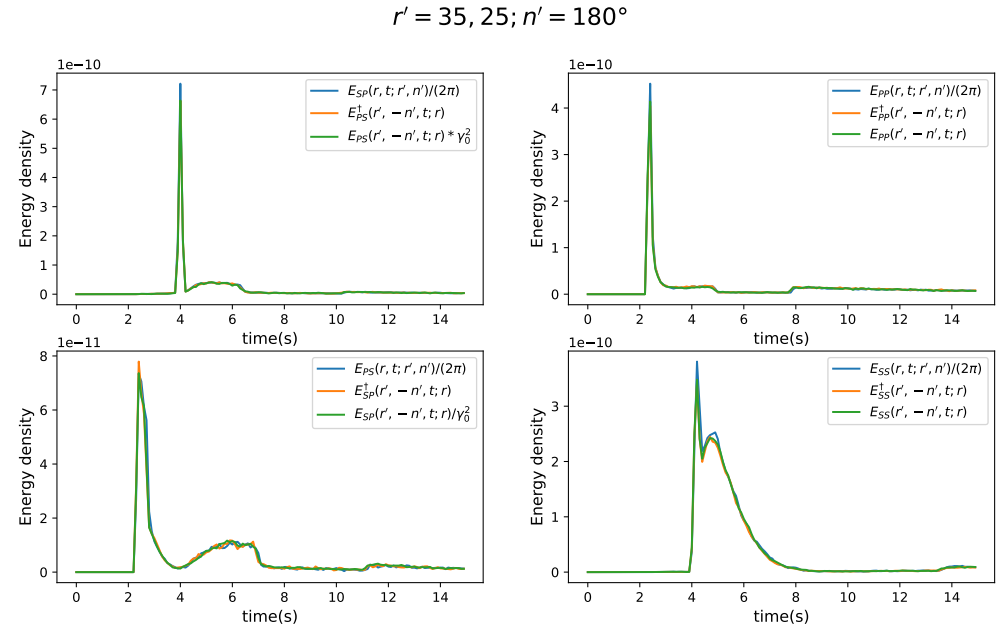
Energy equipartitioning is the intrinsic property of scattered wavefields that the ratio of S-wave to P-wave energy approaches constant value at large lapse time. Equipartition is thus a useful check for simulations of wave scattering. In 2-D elastic case the P/S energy ratio obeys:

$$R_{PS} = \frac{E_S}{E_P} = \frac{g_0^{P\to S}\alpha_0}{g_0^{S\to P}\beta_0} = \gamma_0^2 \,. \tag{3.11}$$

The values of the total scattering coefficients in our simulations are $g_0^{P\to P} = 4.953 \times 10^{-3}$ km$^{-1}$, $g_0^{P\to S} = 1.256 \times 10^{-3}$ km$^{-1}$, $g_0^{S\to P} = 7.24 \times 10^{-4}$ km$^{-1}$ and $g_0^{S\to S} = 1.3418 \times 10^{-2}$ km$^{-1}$. Figure 3.16 shows the temporal evolution of the P to S energy ratio in the region for long lapse time simulations with a P and




**(a)**




**(b)**

**Fig. 3.14:** The comparisons between forward wavefield $E_{YX}(\mathbf{r},t;\mathbf{r}',\mathbf{n}')/S^d$, (blue curves) and adjoint wavefield $E^{\dagger}_{XY}(\mathbf{r}',t,-\mathbf{n}';\mathbf{r})$ (orange curves) in the anomaly model. $E_{XY}(\mathbf{r}',t,-\mathbf{n}';\mathbf{r})$ (green curves) is also compared for reference. The direction is respectively (a)$\mathbf{n}' = 0°$ and (b)$\mathbf{n}' = 180°$.

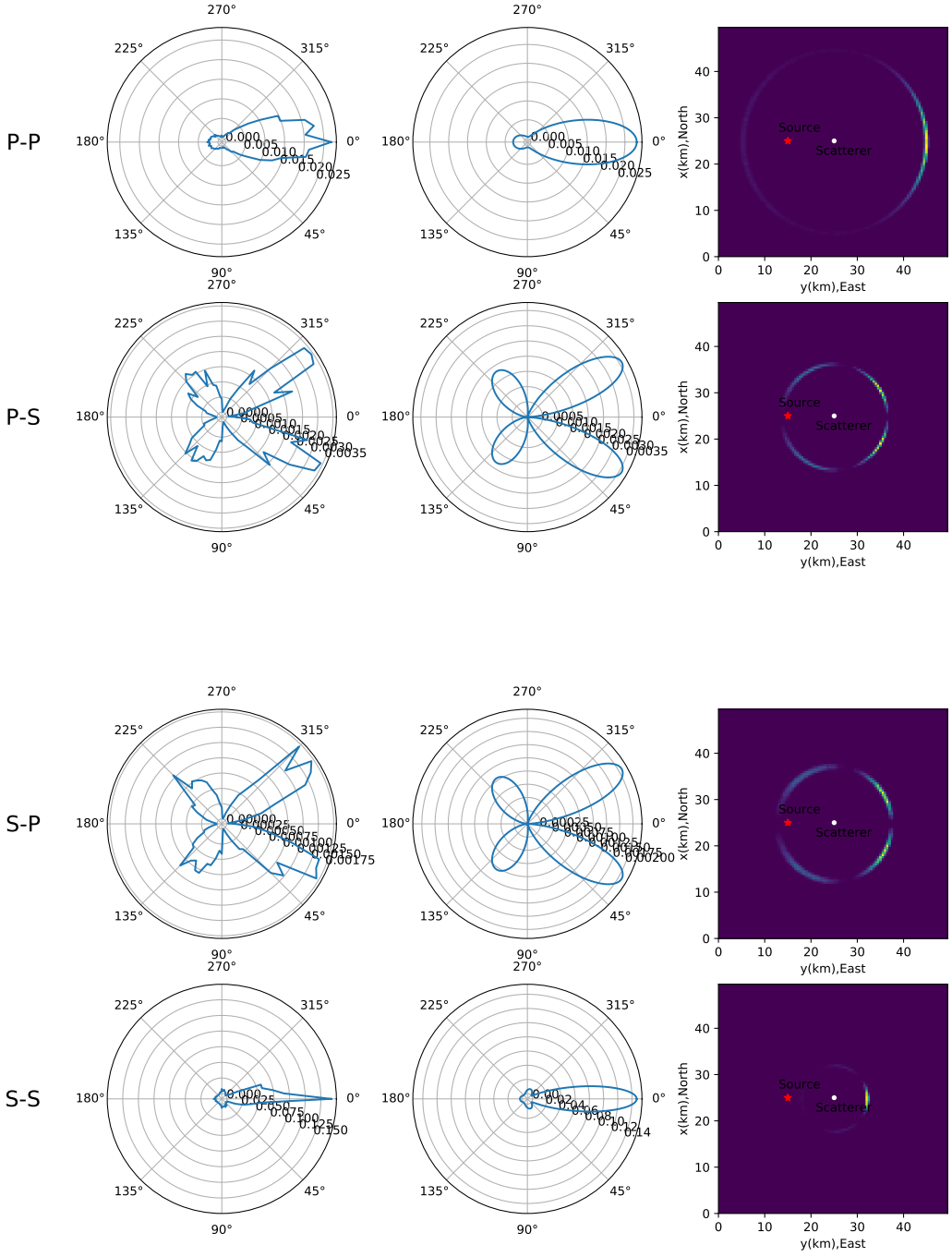
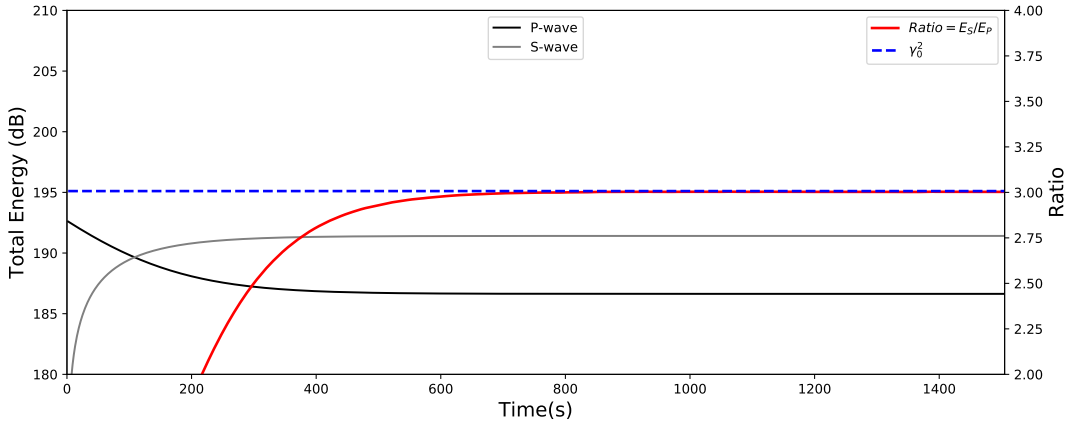
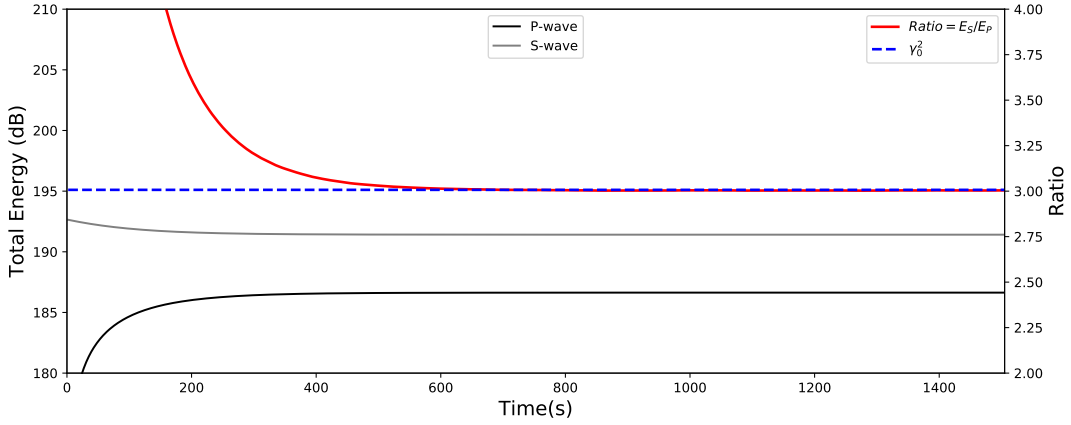



**Fig. 3.15:** The polar plot of scattering patterns for the mode conversions $P \to P$, $P \to S$, $S \to P$, $S \to S$ (from top to bottom). The left column shows the numerical results generated from the Monte-Carlo scattering simulation. The middle column shows the theoretical values, and the right column shows the distribution of scattered energy at time $t = 5s$ which was used to evaluate the angular distribution. The red star and the white point in the right column indicate the source and the scatterer, respectively.




**(a)** P source




**(b)** S source

**Fig. 3.16:** Development of the ratio of S to P energy as a function of time. The black and gray lines are respectively the total energy of P-wave and S-wave. The scale of energy is on the left. The blue dashed line indicates the theoretical value of $\gamma_0^2$ and the red line shows the ratio of gray to black lines. The scale for the energy ratios is on the right.

a S-wave source. In both cases the ratio approaches the blue dashed line which indicates the theoretical prediction of $\gamma_0^2$. This verifies the correct energy equipartition of our scattering simulation.

Chapter 4

# Sensitivity Kernels

## 4.1 Introduction

Sensitivity kernels of the coda provide the connection between a localized spatial perturbation of some propagation properties in the medium (e.g., wave-speed, attenuation, scattering strength) and the changes of a certain waveform property that we observe in the coda wave. This means the sensitivity kernels solve the forward problem of predicting the effect of a medium change on the observable and are thus a tool to localize the perturbations in the Earth based on seismogram changes.

Pacheco & Snieder (2005) first discussed the traveltime changes in a scattered wave field due to localized changes of the velocity under the assumption of scattering in the diffusion regime. The kernel they proposed directly expresses the relationship between the mean traveltime change of the wavefield and the perturbation in the slowness. For weakly scattering media they use the single-scattering approximation (Pacheco & Snieder, 2006). Larose et al. (2010) locate the region of new scattering based on the decorrelation between diffuse scattered waveforms recorded before and after a change was experimentally introduced. Rossetto et al. (2011) derive the spatial sensitivity of the decorrelation observable in the diffusion regime. Planès et al. (2014) argue that the diffusion regime is not applicable to a perturbation close to the source or near the stations and carefully compare the similarities and differences in kernels obtained using the multiple scattering and the diffusion approximations. Most discussions above are based on the assumption of point-like perturbations.

A separate problem is the depth localization of a change in the presence of body and surface waves. Obermann et al. (2013b) introduced an empirical depth sensitivity kernel of coda waves to perturbations within a thin layer using numerical wavefield simulations in a heterogeneous medium. According to Obermann et al. (2013b) the depth sensitivity of coda waves at the surface is a combination of the bulk wave sensitivity and the surface wave sensitivity with temporally varying contributions. These empirical kernels can be used to discriminate between shallow and deep velocity variations. This approach was later extended to the depth sensitivity calculations in the 3-D case (Obermann et al., 2016).

Mayor et al. (2014) derive theoretical expressions for scattering and absorption kernels which describe the direct relationship between the distribution of scattering and absorption properties in space on the one hand and the observed intensity on the other hand. Without assumptions about the propagation regime Margerin et al. (2016) proposed two concepts of passive and active medium perturbations to derive the traveltime and decorrelation sensitivity kernels, which provides a complete mathematical formulation for the spatial sensitivities. These expressions equal the derivation of Planès et al. (2014) in the diffusion regime but extend beyond it. Zhang et al. (2016) use laboratory experiments in a concrete specimen to verify that the decorrelation kernels have high sensitivity to changes in the material. Snieder et al. (2019) gave the expressions for traveltime sensitivity kernels of elastic waves which consider perturbations of P-waves and S-waves velocities. Also considering the elastic case we use the formulation of Margerin et al. (2016) to derive new expressions for traveltime, decorrelation, scattering and attenuation sensitivity

kernels in this paper. We develop a Monte-Carlo approach to calculate the specific energy density distribution, which is central in the theoretical formulation. Our approach is very general and is able to incorporate important ingredients such as non-isotropic scattering of elastic waves and spatially variable scattering and attenuation properties. This is a prerequisite for an iterative tomographic inversion for material properties and property changes of geological or geotechnical structures with scattered seismic waves.

Previous work on estimating medium properties from coda waves often either assumed spatially homogeneous properties (Fehler et al., 1992; Lacombe et al., 2003; Sens-Schönfelder & Wegler, 2006) or made very simple assumptions about the spatial sensitivity. Imaging of the total attenuation is often based on the coda-normalization method with straight-path ray theory (Del Pezzo et al., 2006). This method is widely applied in attenuation tomography of volcanoes (De Siena et al., 2009, 2014a,b; Prudencio et al., 2015b). More recently the method was improved by integrating sensitivity kernels based on assumptions of diffusion or multiple scattering into the inversion (De Siena et al., 2017; Del Pezzo et al., 2018). Using radiative transfer with the approximation of isotropic acoustic wave scattering, Obermann et al. (2013a) calculated the decorrelation sensitivity kernel to image changes in the structure of a volcano. This algorithm was also applied to locate the velocity variations after an earthquake or the eruption of a volcano in other studies (Obermann et al., 2014; Budi-Santoso & Lesage, 2016). Takeuchi (2016) developed a differential Monte-Carlo method that allows to directly calculate the sensitivity kernel in a Monte-Carlo simulation. This approach requires only a single simulation instead of two to obtain the derivative of the envelope with respect to a certain perturbation of the medium. Ogiso (2019) used this approach to map scattering and attenuation in Japan and Sens-Schönfelder et al. (2021) used it to study the effect of scattering at the core mantle boundary.

The assumption of an initially uniform distribution of heterogeneity is common to most studies that investigated the spatial variability of heterogeneity and attenuation properties with scattered coda waves. Therefore, previous studies are restricted to a first order mapping of deviations from uniform heterogeneity (De Siena et al., 2014a,b; Prudencio et al., 2015a; Zieger et al., 2016; Gabrielli et al., 2020; Sketsiou et al., 2020). An iterative tomography of this nonlinear problem is therefore impossible so far and requires developments such as those presented here.

A tomographic inversion tries to obtain the model of the target medium which best describes observed data. The model is encapsulated in a model vector $\mathbf{m}$ that may for example contain the elastic parameters of the medium at the nodes of a spatial grid. Observations are contained in a data vector $\mathbf{d_{obs}}$ consisting, for example, of traveltime measurements or waveform data. Data and model vectors are connected by the forward operator $\mathbf{F}(\mathbf{m}) = \mathbf{d_{syn}}$ where $\mathbf{d_{syn}}$ is a synthetic data vector. How well the model $\mathbf{m}$ describes the target medium is evaluated by an objective function $\chi(\mathbf{F}(\mathbf{m}), \mathbf{d_{obs}})$ which compares the synthetic data predicted from the model to the observed data. Usually $\chi$ is designed such that it assumes a minimum for the best possible model.

Finding a model that improves the description of the data requires knowledge of the gradient of the objective function with respect to the model parameters - the Fréchet derivative $\nabla_m \chi$. We can write the Fréchet derivative as (Fichtner, 2010):

$$\nabla_m \chi(\mathbf{F}(\mathbf{m}), \mathbf{d}) = \nabla_{\mathbf{F}} \chi(\mathbf{F}(\mathbf{m}), \mathbf{d}) \nabla_m \mathbf{F}(\mathbf{m}) \; . \tag{4.1}$$

$\nabla_{\mathbf{F}} \chi$ is the gradient of the objective function with respect to the forward operator, $\nabla_m \mathbf{F}$ is the gradient of the forward operator with respect to the model parameters. The particular difficulty in this formulation is the presence of the forward operator on the right-hand side which renders an explicit calculation of $\nabla_m \chi$ almost impossible for all possible variations of the model ($\delta \mathbf{m}$) because in a finite difference approach each model parameter would require at least one forward simulation. In this section we discuss an alternative

way to derive $\nabla_m \chi$ for the transport equations with different objective functions $\chi$ and different models $\mathbf{m}$ using Bayes' theorem.

Irrespective of the strategy to derive $\nabla_m \chi$ it has the same structure as the model space which is a regular spatial grid of the investigated domain, here. For small perturbations of the model parameters the variations of the forward operator can be assumed linear which leads to the following integral formulation for the change of the objective function in response to a model perturbation $\delta\mathbf{m}$

$$\nabla_m \chi \cdot \delta\mathbf{m} = \int_V K(\mathbf{r}) \delta\mathbf{m}(\mathbf{r}) dV(\mathbf{r}) \tag{4.2}$$

where $K(\mathbf{r})$ is the sensitivity kernel.

The sensitivity kernels encapsulate the information about the spatial sensitivity of the objective function to local changes of model parameters. This information is used to decide where the model needs to be changed to achieve a certain change in the modeled data and improve the model fit. The model $\mathbf{m}$ describes certain material properties which usually are the velocities of P- and S-waves and attenuation for classical seismic tomography. Lamé parameters and quality factors might similarly be used. In our case the medium has additional properties that describe the small scale heterogeneity which is responsible for the scattering. Here we assume that changes of the scattering properties occur in the form of changes in $\varepsilon^2$ only, which we use together with $\alpha$ and $\beta$, $Q_P$ and $Q_S$ as model parameters.

Observations or observables on the other hand are functions of the observed seismic data. In fact sensitivity kernels relate *changes* of the model to *changes* of the data. Here we have to differentiate between two possible applications. The first application is the tomographic question of determining the spatial distribution of material properties. In this case the kernels describe the change in the simulated data resulting from a change in the model. The second application is the monitoring of changes in the medium. This application gained significance with the development of continuous subsurface monitoring using ambient seismic noise. In this case the observable is derived from the seismic observations before and after a change in the medium occurred and the result is the change in medium properties rather than their absolute values.

So observations either describe the changes between two seismic traces observed before and after a change occurred in the medium (monitoring) or the misfit between simulated and measured data (tomography). For scattered waves, commonly used types of observables are the seismic energy difference $\delta E$ (difference between the envelopes of the seismic traces), the travel time change $\delta tt$ and the decorrelation of two traces $dc$. The travel time change can be measured from the seismograms with the time-windowed cross-correlation method (Poupinet et al., 1984; Snieder, 2006). The decorrelation is defined as one minus the normalized correlation coefficient of two seismograms recorded before and after a change occurred in the medium (Larose et al., 2010; Planès et al., 2014). Only the energy difference $\delta E$ can be used for tomography since measurements of travel time changes and decorrelation involve phase information that cannot be obtained with the forward simulation used here. These observables would require wavefield simulations if one aims at absolute parameters of the medium.

For elastic waves the observations can be made with the two different wave modes. Hence, it is natural to first decompose the sensitivity into elementary contributions depending on the emission mode $X$ at the source and the detection mode $Y$ at the receiver. Such a decomposition offers insight into the physical interpretation of the sensitivity. In a second step the elementary kernels will be recombined linearly to take into account the partitioning of the energy emitted at the source and detected at the receiver. This will yield expressions that are directly applicable to observations in seismology or acoustics. Some typical examples will be given later in the paper.

The elementary sensitivity kernels are of the form ${}^{\psi}K^{\phi}_{YX}(\mathbf{r}, t)$. Here $\psi$ stands for the model parameter ($\alpha$, $\beta$, $Q_P$, $Q_S$, $\varepsilon$). $\phi$ indicates the observable ($E$, $tt$, $dc$). $Y$ and $X$ indicate the mode of excitation

$(X)$ and recording $(Y)$. The notation with the receiving mode as first and the excitation mode as second subscript is chosen in accordance with the order of the arguments which lists the conditional arguments of the source in the last position. Not all combinations of $\phi$ and $\psi$ are physically meaningful as there should be a significant influence of the medium parameter on the observable. In the following we discuss the most useful combinations in which the observable has a first order dependence on the model parameter.

The notation of sensitivity kernels we used here is determined by the type of observation and the type of perturbation. The decorrelation and energy change are actually functions of changes in $\varepsilon^2$, but for the sake of simplicity we use $\varepsilon$ in the notation. Similarly, the change of intrinsic attenuation involves the factor $1/Q$ but for simplicity of notation we use $Q$. The notations $\alpha$ and $\beta$ in the traveltime sensitivity kernels also indicate $\delta\alpha/\alpha_0$ and $\delta\beta/\beta_0$, respectively.

## 4.2 Traveltime Sensitivity Kernels

We first look at the travel time sensitivity kernels ${}^{\alpha}K^{tt}_{YX}$, ${}^{\beta}K^{tt}_{YX}$ that describe the travel time change due to a perturbation of the P- or S-wave velocities, respectively. In the case of vectorial waves, the observable may be defined more precisely as the travel time perturbation averaged over each cartesian component of the field.

We utilize Bayes' theorem to derive the traveltime sensitivity kernels for the elastic case. It is the same probability-based method as used by Margerin et al. (2016) in which they define two events $A$ and $B$. $A$ denotes the event that: a seismic phonon is detected at time $t'$ in the volume $dV(\mathbf{r}')$ with a propagation direction $\mathbf{n}'$. In the elastic case, we use a subscript $A_V$ to indicate that the seismic phonon has mode $V$ in the event $A$. Similarly, $B_Y$ denotes the event: a seismic phonon of mode $Y$ reaches $\mathbf{r}$ at time $t$ in any direction.

We denote the energy density detected at the receiver position $\mathbf{r}$ from the source position $\mathbf{r_0}$ at the time $t$ by $E_{YX}(\mathbf{r}, t; \mathbf{r_0})$. It corresponds to event $B_Y$. $Y$ and $X$ indicate the mode of excitation $(X)$ and recording $(Y)$. For any position $\mathbf{r}'$ at any time $t'$ the probability density function relevant to event $A_V$ is (up to a normalization factor) given by the specific energy density $E_{VX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r_0})$. Hence, the probability that a phonon which we observe at location $\mathbf{r}$ in mode $Y$ has visited $dV(\mathbf{r}')$ in P-mode after being launched in mode $X$ from location $\mathbf{r_0}$ is given by:

$$\begin{aligned} P(A_P \mid B_Y) &= \frac{P(B_Y \mid A_P)\, P(A_P)}{P(B_Y)} \\ &= \frac{E_{YP}(\mathbf{r}, t-t'; \mathbf{r}', \mathbf{n}')\, E_{PX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r_0})\, dV(\mathbf{r}')}{E_{YX}(\mathbf{r}, t; \mathbf{r_0})} \end{aligned} \tag{4.3}$$

$E_{YP}(\mathbf{r}, t-t'; \mathbf{r}', \mathbf{n}')$ is the energy density recorded at the receiver at $\mathbf{r}$ from a source radiating in direction $\mathbf{n}'$ at location $\mathbf{r}'$. It correspond to the probability that the phonon at $\mathbf{r}'$ continues to propagate in direction $\mathbf{n}'$ and reaches the receiver in the remaining time $t - t'$ in mode $Y$. The probability that the phonon visited $dV(\mathbf{r}')$ in S-mode is given by:

$$\begin{aligned} P(A_S \mid B_Y) &= \frac{P(B_Y \mid A_S)\, P(A_S)}{P(B_Y)} \\ &= \frac{E_{YS}(\mathbf{r}, t-t'; \mathbf{r}', \mathbf{n}')\, E_{SX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r_0})\, dV(\mathbf{r}')}{E_{YX}(\mathbf{r}, t; \mathbf{r_0})} \end{aligned} \tag{4.4}$$

Thanks to these probability distributions we may compute, at a given lapse-time $t$ in the signal, the typical time spent in the volume $dV(\mathbf{r}')$ by either P-mode or S-mode seismic phonons propagating in

direction $\mathbf{n}'$:

$$ {}^{\alpha}T_{YX}(dV(\mathbf{r}'),\mathbf{n}',t;\mathbf{r},\mathbf{r_0}) = dV(\mathbf{r}')\int_0^t \frac{E_{YP}(\mathbf{r},t-t';\mathbf{r}',\mathbf{n}')E_{PX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_0})}{E_{YX}(\mathbf{r},t;\mathbf{r_0})}dt' \tag{4.5} $$

$$ {}^{\beta}T_{YX}(dV(\mathbf{r}'),\mathbf{n}',\mathbf{t};\mathbf{r},\mathbf{r_0}) = dV(\mathbf{r}')\int_0^t \frac{E_{YS}(\mathbf{r},t-t';\mathbf{r}',\mathbf{n}')E_{SX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_0})}{E_{YX}(\mathbf{r},t;\mathbf{r_0})}dt' \tag{4.6} $$

To obtain the time spent in $dV(\mathbf{r}')$ irrespective of propagation direction, eq. (4.5) and eq. (4.6) have to be integrated over directions $\mathbf{n}'$. The travel time change that a wave accumulates during its propagation in $dV(\mathbf{r}')$ is $-\delta c/c(\mathbf{r}')\cdot T(dV(\mathbf{r}'))$ with $\delta c/c$ the fractional change of the wave velocity. We assume that perturbations are small and travel time changes from different locations can be superimposed. Now we can obtain the travel time shift for spatially distributed changes of P- and S-wave velocities:

$$ \begin{aligned} \delta t_{YX}(t) &= -\int_{V^d}[\frac{\delta\alpha}{\alpha}(\mathbf{r}')\int_{S^d} {}^{\alpha}T_{YX}(dV(\mathbf{r}'),\mathbf{n}',t;\mathbf{r},\mathbf{r_0})d\mathbf{n}'+ \\ &\qquad \frac{\delta\beta}{\beta}(\mathbf{r}')\int_{S^d} {}^{\beta}T_{YX}(dV(\mathbf{r}'),\mathbf{n}',t;\mathbf{r},\mathbf{r_0})d\mathbf{n}']dV(\mathbf{r}') \\ &= -\int_{V^d}[\frac{\delta\alpha}{\alpha}(\mathbf{r}'){}^{\alpha}K^{tt}_{YX}(\mathbf{r}',\mathbf{n}',t;\mathbf{r},\mathbf{r_0})+ \\ &\qquad \frac{\delta\beta}{\beta}(\mathbf{r}'){}^{\beta}K^{tt}_{YX}(\mathbf{r}',\mathbf{n}',t;\mathbf{r},\mathbf{r_0})]dV(\mathbf{r}') \end{aligned} \tag{4.7} $$

which defines the sensitivity kernels ${}^{\alpha}K^{tt}_{YX}$ and ${}^{\beta}K^{tt}_{YX}$. In Eq. (4.7) $S^d$ denotes the unit sphere in space dimension $d$ and $V^d$ is the full space. Measurements of the travel time shifts can be performed on any seismogram component and may be averaged.

The calculation of the sensitivity kernels on the basis of eq. (4.5) and eq. (4.6) is impractical as it involves the energy density $E(\mathbf{r};\mathbf{r}')$ that originates from a source at $\mathbf{r}'$ which can be anywhere in the domain. To avoid the necessity of simulating sources throughout the medium we use the reciprocity relation of transport theory discussed in Section 3.3. It allows to replace the large number of simulation to obtain the signals recorded from sources everywhere in the medium at one particular receiver by a single simulation with a source at the original receiver and a large number of receivers everywhere in the medium. Using the adjoint energy density $E^{\dagger}(\mathbf{r}';\mathbf{r})$ introduced in eq. (3.8) to replace $E(\mathbf{r};\mathbf{r}')$ we obtain the following expressions for the elastic travel time sensitivity kernels:

$$ {}^{\alpha}K^{tt}_{YX}(\mathbf{r}',t;\mathbf{r},\mathbf{r_0}) = S^d\int_{S^d}\int_0^t \frac{E^{\dagger}_{PY}(\mathbf{r}',t-t',-\mathbf{n}';\mathbf{r})\,E_{PX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_0})\,dt'd\mathbf{n}'}{E_{YX}(\mathbf{r},t;\mathbf{r_0})} \tag{4.8} $$

$$ {}^{\beta}K^{tt}_{YX}(\mathbf{r}',t;\mathbf{r},\mathbf{r_0}) = S^d\int_{S^d}\int_0^t \frac{E^{\dagger}_{SY}(\mathbf{r}',t-t',-\mathbf{n}';\mathbf{r})\,E_{SX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_0})\,dt'd\mathbf{n}'}{E_{YX}(\mathbf{r},t;\mathbf{r_0})} \tag{4.9} $$

Both kernels describe waves that are launched in mode $X$ and recorded in mode $Y$.

To illustrate the kernel calculation, we employ the specific energy density results of the modeling in the statistically homogeneous model. The traveltime sensitivity kernels are shown in Figure 4.1 and Figure 4.2 for P- and S-wave source, respectively. Although the kernels are obviously affected by the positions of the source and the receiver and the lapse time, they are also sensitive to the scattering process. ${}^{\alpha}K^{tt}_{PP}$ describes the effect of a local perturbation of P-wave velocity on the P-wave recording from a P-wave source. The result is symmetric since the source and the receiver have the same type which means that we can interchange the locations. ${}^{\beta}K^{tt}_{PP}$ has the same reason for its symmetry but the effect of S-wave velocity leads to its value is much weaker than ${}^{\alpha}K^{tt}_{PP}$. Because it must at least convert twice before arriving at the receiver. ${}^{\alpha}K^{tt}_{SP}$ and ${}^{\beta}K^{tt}_{SP}$ highlight the role of P-wave and S-wave velocity perturbations, respectively, observed with an S-wave recording from a P-excitation. Since the source

excites the P-waves, the simplest case is that there is only one energy conversion. But for ${}^{\alpha}K^{tt}_{SP}$ the conversion happens after passing through $dV(\mathbf{r}')$ while for ${}^{\beta}K^{tt}_{SP}$ it occurs before. This is why the higher sensitivity is closer to the source for ${}^{\alpha}K^{tt}_{SP}$ but for ${}^{\beta}K^{tt}_{SP}$ it is closer to the receiver, resulting in a strong asymmetry. Figure 4.2 shows the traveltime sensitivity kernels for an S-wave source. It reveals the same type of asymmetry as discussed above for a P-wave source. The apparent asymmetry of ${}^{\beta}K^{tt}_{SS}$ at 5 s lapse time is due to poor sampling prior to the arrival of ballistic S-energy. The excited S-energy has to be converted to P in order to reach the receiver prior to the ballistic S-phase, but it has to be converted back to S-energy before being recording which makes these events very unlikely. The comparably high amplitude of ${}^{\alpha}K^{tt}_{SS}$ is due to the normalization by a very small total energy density.

## 4.3 Decorrelation Sensitivity Kernels

We now turn to the decorrelation sensitivity kernels ${}^{\varepsilon}K^{dc}_{YX}$ that describe the decorrelation between two wavefield measurements as defined by Planès et al. (2014) due to a change of the mechanical properties of the medium. We assume that the change in the mechanical properties can be described as a change in the fractional variation $\varepsilon$ of the fluctuations of the medium. This means that the change only influences the amplitude of the power spectrum. The angular pattern of the scattering coefficient is not affected by the change.

The definition and expression of the decorrelation coefficient given by Margerin et al. (2016) in the case of scalar waves may be extended to the vectorial case as follows:

$$\begin{aligned} dc(t) &= 1 - \frac{\langle u_i(t)\tilde{u}_i(t)\rangle}{\sqrt{\langle u_i(t)u_i(t)\rangle\langle \tilde{u}_i(t)\tilde{u}_i(t)\rangle}} \\ &\approx \frac{\langle (u_i(t)-\tilde{u}_i(t))(u_i(t)-\tilde{u}_i(t))\rangle}{2\langle u_i(t)u_i(t)\rangle}\,, \end{aligned} \tag{4.10}$$

where $u_i(t)$ and $\tilde{u}_i(t)$ refer to the $i$-th component of the wavefield before and after the perturbation, respectively. In Eq.(4.10) we have assumed that the difference between $u_i(t)$ and $\tilde{u}_i(t)$ is small and uncorrelated with $u_i(t)$. Note that the Einstein summation convention is employed in Eq.(4.10). In the second equality, we recognize (up to a constant pre-factor equal to the product $\rho\omega^2$ of local mass density and squared circular frequency) the extra energy density emitted by the scattering perturbation. Similarly, the denominator $\langle u^2(t)\rangle$ is up to the same pre-factor the total energy density of the wavefield.

The extra energy scattered by the local change in $\varepsilon$ observed in the lapse time interval $(t', t'+dt')$ is given by

$$\delta E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_0}) = \sum_{V=P,S} \int_{S^d} dt'\; c_V \;\left| g^{V\to W}(\mathbf{n}, \mathbf{n}'; \tilde{\varepsilon}^2(\mathbf{r}')) - g^{V\to W}(\mathbf{n}, \mathbf{n}'; \varepsilon^2(\mathbf{r}'))\right| E_{VX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r_0}) d\mathbf{n}' \tag{4.11}$$

The summation over wave mode $V$ takes care of the fact that both the $P$ and $S$ phonons that are incident on the volume element at $\mathbf{r}'$ contribute to the decorrelation. $c_V$ is the velocity of the wave mode $V$ and $c_P \equiv \alpha_0, c_S \equiv \beta_0$. $\mathbf{n}'$ and $\mathbf{n}$ denote the incoming and outgoing direction, respectively. $\tilde{\varepsilon}$ is the fractional

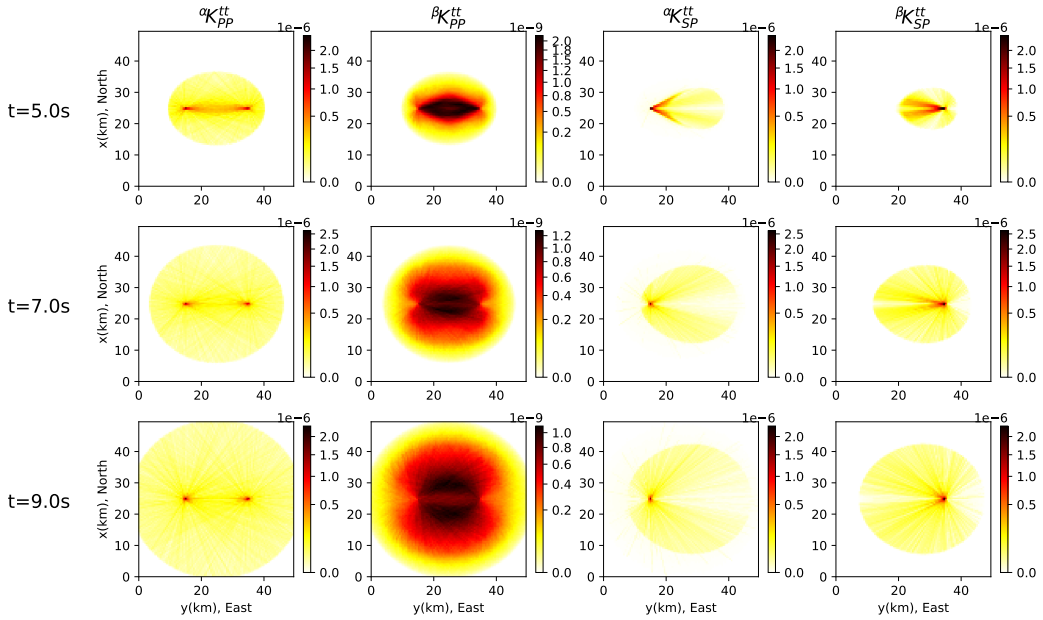



**Fig. 4.1:** Traveltime sensitivity kernels in uniform model: ${}^{\alpha}K^{tt}_{PP}$ (the 1st column), ${}^{\beta}K^{tt}_{PP}$ (the 2nd column), ${}^{\alpha}K^{tt}_{SP}$ (the 3rd column) and ${}^{\beta}K^{tt}_{SP}$ (the 4th column) at different lapse times with P-wave source. Note the 2nd column has a different color scale and all scales are nonlinear. Source $\mathbf{r_0}$: $x = 25$ km, $y = 15$ km and receiver $\mathbf{r}$: $x = 25$ km, $y = 35$ km.

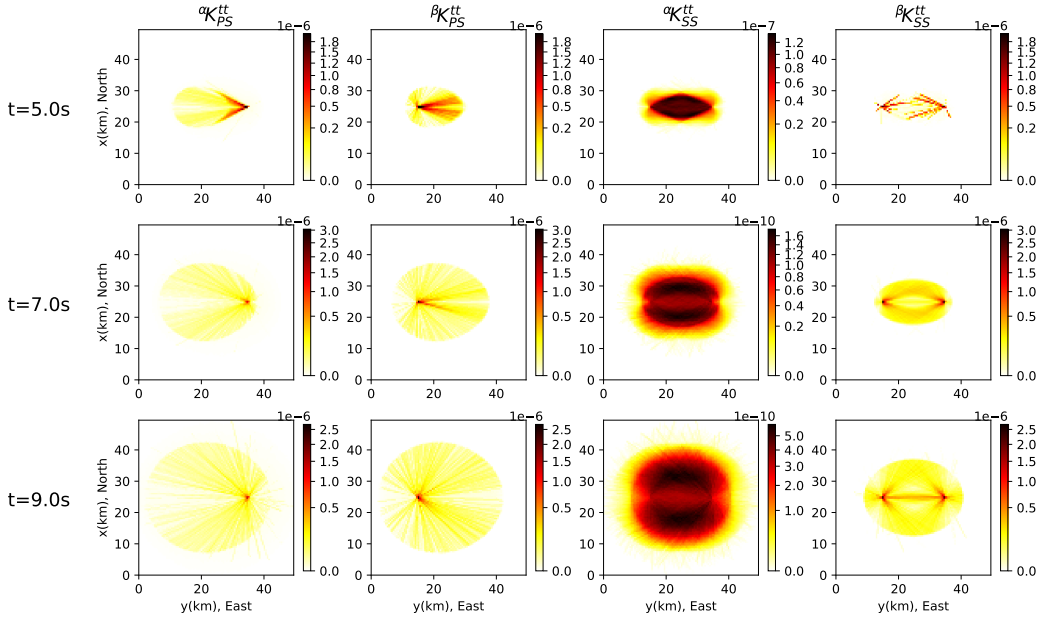



**Fig. 4.2:** Traveltime sensitivity kernels in uniform model: ${}^{\alpha}K^{tt}_{PS}$ (the 1st column), ${}^{\beta}K^{tt}_{PS}$ (the 2nd column), ${}^{\alpha}K^{tt}_{SS}$ (the 3rd column) and ${}^{\beta}K^{tt}_{SS}$ (the 4th column) at different lapse times with S-wave source. Note the 3rd column in each row has the different color scale and all scales are nonlinear.

fluctuation after the perturbation. Since the scattering coefficient is proportional to $\varepsilon^2$, we obtain

$$\begin{aligned}\delta E_{WX}(\mathbf{r}',t',\mathbf{n};\mathbf{r_0}) &= \sum_{V=P,S}\int_{S^d} dt'\ c_V\ \frac{\left|\delta\varepsilon^2(\mathbf{r}')\right|}{\varepsilon^2(\mathbf{r}')} g^{V\to W}(\mathbf{n},\mathbf{n}';\varepsilon^2(\mathbf{r}'))E_{VX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_0})d\mathbf{n}' \\ &= \sum_{V=P,S}\int_{S^d} dt'\ c_V\ \frac{\left|\delta\varepsilon^2(\mathbf{r}')\right|}{\varepsilon^2(\mathbf{r}')}\frac{\varepsilon^2(\mathbf{r}')}{\varepsilon_0^2} g_0^{V\to W}(\varepsilon_0^2) f^{V\to W}(\mathbf{n},\mathbf{n}')E_{VX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_0})d\mathbf{n}'\end{aligned} \tag{4.12}$$

Here the definition of $\delta\varepsilon(\mathbf{r}')^2$ is the perturbation of local scattering strength which is defined as $\delta\varepsilon^2(\mathbf{r}') = \tilde{\varepsilon}^2(\mathbf{r}') - \varepsilon^2(\mathbf{r}')$. $\varepsilon^2(\mathbf{r}')/\varepsilon_0^2$ is the ratio between the local value and the homogeneous background which describes the structure of the model. $g_0^{V\to W}(\varepsilon_0^2)$ is the total scattering coefficient of the background which determines the probability of scattering. The superscript $V \to W$ in $g_0^{V\to W}(\varepsilon_0^2)$ indicate the wave mode conversion from mode $V$ to $W$. Here we define

$$f^{V\to W}(\mathbf{n},\mathbf{n}') = \frac{g^{V\to W}(\mathbf{n},\mathbf{n}';\varepsilon_0^2)}{g_0^{V\to W}(\varepsilon_0^2)} \tag{4.13}$$

where $f^{V\to W}(\mathbf{n},\mathbf{n}')$ is the normalized differential scattering cross-section (Sato et al., 2012). In another words $f^{V\to W}(\mathbf{n},\mathbf{n}')$ is the probability density for an incoming $V$ phonon propagating in direction $\mathbf{n}'$ to be mode converted to a $W$ phonon propagating in direction $\mathbf{n}$.

Propagating this extra energy that acts as a secondary source at $\mathbf{r}'$ further to the receiver at $\mathbf{r}$ and integrating over time $t'$ at which the phonons visit the perturbed volume yields

$$\delta E_Y(\mathbf{r},t;\mathbf{r}';\mathbf{r_0}) = \sum_{W=P,S}\int_{S^d}\int_0^t E_{YW}(\mathbf{r},t-t';\mathbf{r}',\mathbf{n})\delta E_{WX}(\mathbf{r}',t',\mathbf{n};\mathbf{r_0})dt'd\mathbf{n} \tag{4.14}$$

The quantity $\delta E_Y(\mathbf{r},t;\mathbf{r}';\mathbf{r_0})$ depends on the location of the perturbation and has to be integrated over space $V^d$ to obtain the total change of energy.

Inserting eq. (4.14) in the numerator of eq. (4.10) yields for the decorrelation of the two wavefields recorded before and after the perturbation of the mechanical properties

$$\begin{aligned}dc_{YX}(t) =& \frac{\delta E_Y(\mathbf{r},t;\mathbf{r}';\mathbf{r_0})}{2E_{YX}(\mathbf{r},t;\mathbf{r_0})} \\ =& \frac{1}{2\varepsilon_0^2}\int_{V^d}\left|\delta\varepsilon^2(\mathbf{r}')\right|\ {}^{\varepsilon}K_{YX}^{dc}(\mathbf{r}',t;\mathbf{r},\mathbf{r_0})dV(\mathbf{r}')\end{aligned} \tag{4.15}$$

where:

$$\begin{aligned}{}^{\varepsilon}K_{YX}^{dc}(\mathbf{r}',t;\mathbf{r},\mathbf{r_0}) =& S^d\sum_W\sum_V\int_{S^d}\int_{S^d}\int_0^t c_V g_0^{V\to W}(\varepsilon_0^2) \\ &\times\frac{E_{WY}^{\dagger}(\mathbf{r}',t-t',-\mathbf{n};\mathbf{r})f^{V\to W}(\mathbf{n},\mathbf{n}')E_{VX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_0})}{E_{YX}(\mathbf{r},t;\mathbf{r_0})}dt'd\mathbf{n}'d\mathbf{n}\end{aligned} \tag{4.16}$$

Here we have used again the reciprocity relation eq. (3.8) and the summation over modes $W$ and $V$ corresponds to the four different modes of scattering that connect incident P- and S-waves each to outgoing P- and S-waves.

The illustration of decorrelation sensitivity kernels in a statistically homogeneous model are shown in Figure 4.3. The single scattering ellipse with the highest sensitivity in ${}^{\varepsilon}K_{PP}^{dc}$ is clearly marked since the strongest influence on P-wave recordings from a P-wave source comes from single-scattering. This interpretation also holds for the ellipse in ${}^{\varepsilon}K_{SS}^{dc}$ of S- into S-wave single-scattering. The regions of higher

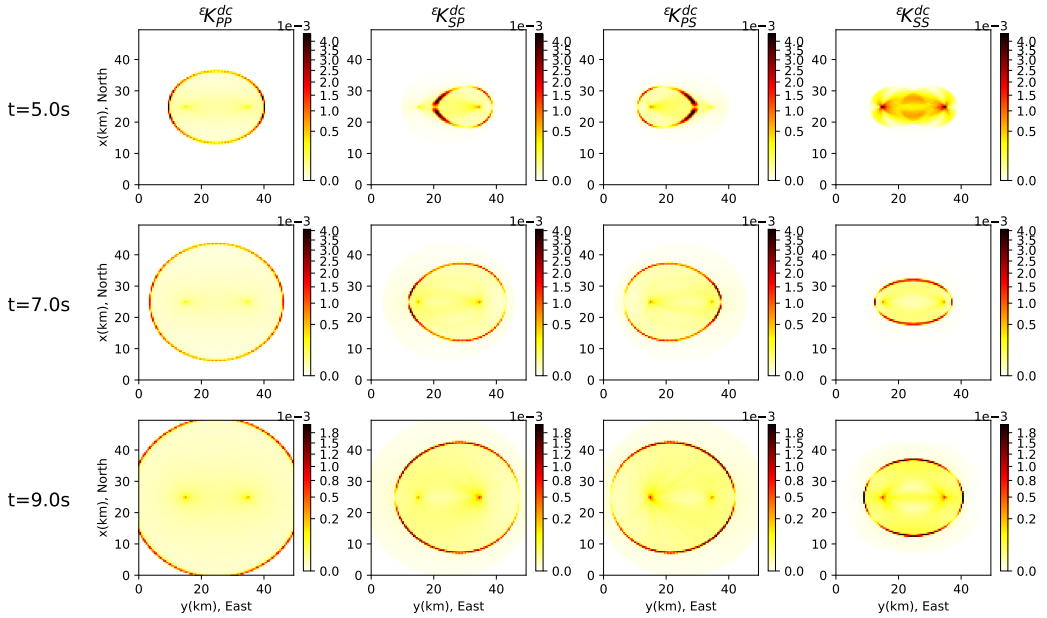



**Fig. 4.3:** Decorrelation sensitivity kernels in uniform model: $^{\varepsilon}K^{dc}_{PP}$ (the 1st column), $^{\varepsilon}K^{dc}_{SP}$ (the 2nd column), $^{\varepsilon}K^{dc}_{PS}$ (the 3rd column) and $^{\varepsilon}K^{dc}_{SS}$ (the 4th column) at different lapse times. Note all scales are nonlinear.

sensitivity in $^{\varepsilon}K^{dc}_{SP}$ and $^{\varepsilon}K^{dc}_{PS}$ have the same explanation but due to the change of modes which propagate with different velocities they are not elliptical. It is also easy to understand that $^{\varepsilon}K^{dc}_{SP}$ is symmetrical to $^{\varepsilon}K^{dc}_{PS}$ due to the interchange of the source and the receiver.

## 4.4 Energy Sensitivity Kernels for Changes in Scattering and Intrinsic Attenuation

In this section we derive sensitivity kernels for energy observations, i.e. the spatio temporal distribution of coda wave energy. Since this observable is influenced by variations of attenuation as well as in scattering strength we will calculate the kernels $^{Q}K^{E}_{YX}$ and $^{\varepsilon}K^{E}_{YX}$.

A perturbation of scattering properties in a local volume $dV(\mathbf{r}')$ has two effects on the propagation of energy (Margerin et al., 2016). Temporarily ignoring the intrinsic attenuation and assuming a local increase in scattering strength eq. (2.12) and eq. (2.13) reveal that one effect is the loss of energy due to the stronger scattering of $E_{YX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_0})$ into directions other than $\mathbf{n}'$, and the other effect is the increase of energy by scattering from other direction $\mathbf{n}'$ into direction $\mathbf{n}$. The second effect has been discussed in eq. (4.12) when we derived the decorrelation sensitivity kernels. In the following we denote the energy increase due to scattering described in eq. (4.12) by $^{2}\delta E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_0})$. The first mentioned effect of a local increase in scattering strength -the decrease of ballistic energy across the volume element- $dV(\mathbf{r}')$ is given by:

$$^{1}\delta E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_0}) = -\sum_{V=P,S} dt'\; c_W\; \frac{\delta\varepsilon^2(\mathbf{r}')}{\varepsilon^2(\mathbf{r}')} \frac{\varepsilon^2(\mathbf{r}')}{\varepsilon_0^2} g_0^{W\to V}(\varepsilon_0^2) E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_0}) \tag{4.17}$$

Considering both effects, we add eq. (4.12) and eq. (4.17) to obtain the secondary source at $\mathbf{r}'$. This energy further propagates to the receiver at $\mathbf{r}$. We obtain the seismic energy change at $\mathbf{r}$ due to a local

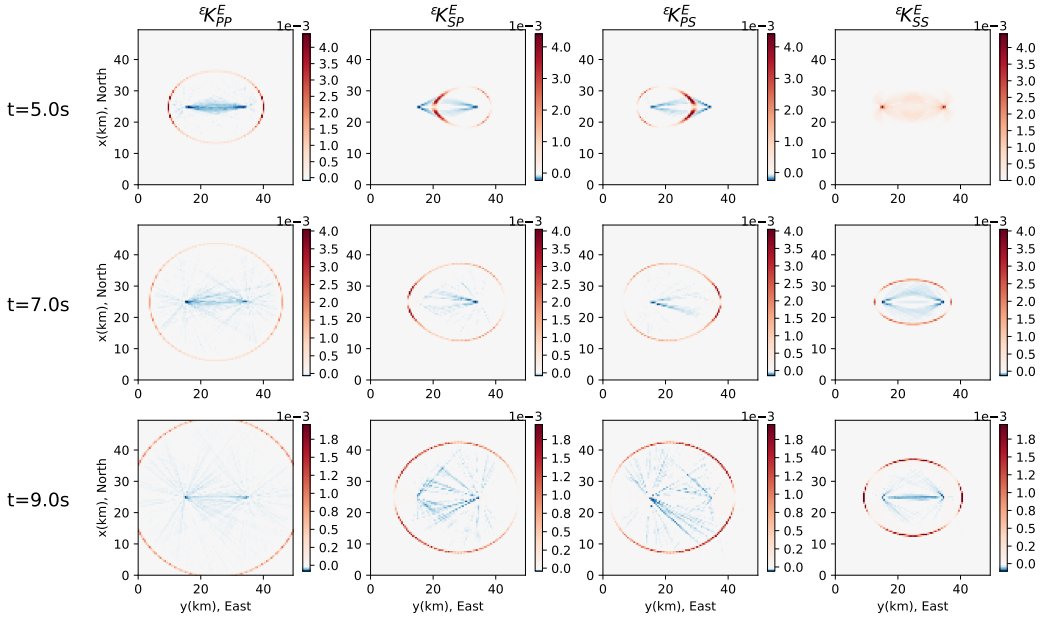



**Fig. 4.4:** Scattering sensitivity kernels in uniform model: ${}^{\varepsilon}K^{E}_{PP}$ (the 1st column), ${}^{\varepsilon}K^{E}_{SP}$ (the 2nd column), ${}^{\varepsilon}K^{E}_{PS}$ (the 3rd column) and ${}^{\varepsilon}K^{E}_{SS}$ (the 4th column) at different lapse times. The color of white indicates the value of 0, the red is positive and the blue is negative.

weak perturbation of scattering properties:

$$\delta E_{YX}(\mathbf{r},t;\mathbf{r}';\mathbf{r_0}) = \sum_{W=P,S}\int_{S^d}\int_0^t E_{YW}(\mathbf{r},t-t';\mathbf{r}',\mathbf{n})[{}^1\delta E_{WX}(\mathbf{r}',t',\mathbf{n};\mathbf{r_0}) + {}^2\delta E_{WX}(\mathbf{r}',t',\mathbf{n};\mathbf{r_0})]dt'd\mathbf{n}\ . \tag{4.18}$$

This expression is a combination of the traveltime and decorrelation sensitivity kernels given by eq. (4.8), eq. (4.9) and eq. (4.16). By substituting them into eq. (4.18), the perturbation of energy in the coda may be expressed as:

$$\frac{\delta E_{YX}}{E_{YX}}(\mathbf{r},t;\mathbf{r_0}) = \frac{1}{\varepsilon_0^2}\int_{V^d}\delta\varepsilon^2(\mathbf{r}')\ {}^{\varepsilon}K^{E}_{YX}(\mathbf{r}',t;\mathbf{r},\mathbf{r_0})dV(\mathbf{r}') \tag{4.19}$$

where the scattering sensitivity kernels may in turn be expressed in terms of the travel time and decorrelation kernels as follows:

$$\begin{aligned}{}^{\varepsilon}K^{E}_{YX}(\mathbf{r}',t;\mathbf{r},\mathbf{r_0}) = {}^{\varepsilon}K^{dc}_{YX}(\mathbf{r}',t;\mathbf{r},\mathbf{r_0}) - [&\alpha_0(g_0^{P\to P}(\varepsilon_0^2) + g_0^{P\to S}(\varepsilon_0^2))\ {}^{\alpha}K^{tt}_{YX}(\mathbf{r}',t;\mathbf{r},\mathbf{r_0}) \\ &+ \beta_0(g_0^{S\to P}(\varepsilon_0^2) + g_0^{S\to S}(\varepsilon_0^2))\ {}^{\beta}K^{tt}_{YX}(\mathbf{r}',t;\mathbf{r},\mathbf{r_0})]\end{aligned} \tag{4.20}$$

Figure 4.4 shows the sensitivity kernels calculated from eq. (4.20). Positive (resp. negative) values of the kernel are shown in red (resp. blue). Similar to the decorrelation sensitivity kernels, the positive sensitivity is dominated by single-scattering where an increase of scattering strength causes more scattered energy to be recorded at the receiver. Negative sensitivity indicates that increased scattering strength decreases the recorded energy due to multiple scattering that spreads out energy in space.

We now consider the effect of changes in intrinsic attenuation on the recorded energy. According to eq. (2.12) and eq. (2.13), a local change of intrinsic attenuation affects the weight of the phonons propagating through the perturbation, independent of the direction. From this consideration, we deduce

the change of energy due to a perturbation of attenuation in the volume $dV(\mathbf{r}')$:

$$\delta E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_0}) = -\sum_{W=P,S} dt' \omega \delta Q_W^{-1}(\mathbf{r}') E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_0}) \tag{4.21}$$

where $\delta Q_W^{-1}(\mathbf{r}') = \tilde{Q}_W^{-1}(\mathbf{r}') - Q_W^{-1}(\mathbf{r}')$ and $\tilde{Q}_W^{-1}(\mathbf{r}')$ denotes the local intrinsic quality factors after the perturbation. So the perturbation of energy density at $\mathbf{r}$ caused by a local perturbation of intrinsic attenuation $\delta Q_P^{-1}(\mathbf{r}')$ and $\delta Q_S^{-1}(\mathbf{r}')$ is:

$$\frac{\delta E_{YX}}{E_{YX}}(\mathbf{r}, t; \mathbf{r_0}) = -\omega \int_{V^d} [\delta Q_P^{-1}(\mathbf{r}') \ {}^{Q_P}K^E_{YX}(\mathbf{r}', t; \mathbf{r}, \mathbf{r_0}) + \delta Q_S^{-1}(\mathbf{r}') \ {}^{Q_S}K^E_{YX}(\mathbf{r}', t; \mathbf{r}, \mathbf{r_0})] dV(\mathbf{r}') \tag{4.22}$$

As could be anticipated, the attenuation sensitivity kernels are identical to the traveltime sensitivity kernels:

$${}^{Q_P}K^E_{YX}(\mathbf{r}', t; \mathbf{r}, \mathbf{r_0}) = {}^{\alpha}K^{tt}_{YX}(\mathbf{r}', t; \mathbf{r}, \mathbf{r_0}) \tag{4.23}$$

$${}^{Q_S}K^E_{YX}(\mathbf{r}', t; \mathbf{r}, \mathbf{r_0}) = {}^{\beta}K^{tt}_{YX}(\mathbf{r}', t; \mathbf{r}, \mathbf{r_0}) \tag{4.24}$$

## 4.5 Combination of Sensitivity Kernels

We derived different sensitivity kernels for the same perturbation depending on the different modes of excitation ($X$) and recording ($Y$). These expressions are derived theoretically but might not be practical in applications since P- and S -energy density are usually jointly excited and recorded. We will first consider the combination of kernels in a simple scenario where an idealized source emits a single type of energy but the $P$ and $S$ waves cannot be separated at the receiver. In a second step the results will be generalized to arbitrary source and detection.

The energy density detected at receiver is a mixture of P and S modes and is given by $E_{RX}(\mathbf{r}, t; \mathbf{r_0}) = E_{PX}(\mathbf{r}, t; \mathbf{r_0}) + E_{SX}(\mathbf{r}, t; \mathbf{r_0})$, where $X$ denotes a single emission mode ($P$ or $S$) and $R$ is the *recorded* energy density. After proper normalization, the energy densities $E_{PX}$ and $E_{SX}$ can be interpreted as the probability of detection of a $P$ or $S$ mode at the receiver. Hence the key quantity for the combination of kernels is the ratio of the two types of energies $R_{SP} = E_{SX}/E_{PX}$ which depends in general on the lapse-time in the coda. The sensitivity of the measurement that combines the different detection modes can now be obtained as a weighted sum of the P and S wave sensitivities with relative weights that depend on the ratio $R_{SP}$. For this ratio one can either assume a reasonable value or it can be obtained directly from the simulations used to calculate the sensitivity kernels. For example the traveltime sensitivity kernel for a measurement that mixes P and S modes can be written as:

$${}^{\alpha}K^{tt}_{RX} = W_P \, {}^{\alpha}K^{tt}_{PX} + W_S \, {}^{\alpha}K^{tt}_{SX} \tag{4.25}$$

and

$${}^{\beta}K^{tt}_{RX} = W_P \, {}^{\beta}K^{tt}_{PX} + W_S \, {}^{\beta}K^{tt}_{SX} \ , \tag{4.26}$$

where $W_P = 1/(1 + R_{SP})$ and $W_S = R_{SP}/(1 + R_{SP})$. For large lapse times we can assume wave propagation in the diffusion regime. In this case $g_0^{P\to S} = \gamma_0 g_0^{S\to P}$ where $\gamma_0 = \alpha_0/\beta_0$. Since the energy ratio of S- to P-waves obeys $R_{SP} = (g_0^{P\to S}\alpha_0)/(g_0^{S\to P}\beta_0)$ (Sato et al., 2012, pp. 241), $R_{SP} = \gamma_0^2$. So eq. (4.25) and eq. (4.26) have the same meaning as the expression of the velocity change weighted average of changes in the P- and S-wave velocities by Snieder (2006). The other sensitivity kernels can

be combined in the same way:

$$ {}^{\varepsilon}K^{dc}_{RX} = W_P {}^{\varepsilon}K^{dc}_{PX} + W_S {}^{\varepsilon}K^{dc}_{SX} \tag{4.27} $$

$$ {}^{\varepsilon}K^{E}_{RX} = W_P {}^{\varepsilon}K^{E}_{PX} + W_S {}^{\varepsilon}K^{E}_{SX} \, . \tag{4.28} $$

The sensitivity kernels we derived before are very general which means that we can combine some of them for different specific situations, for instance, if we can assume that the perturbations of P- and S-wave velocities are the same, i.e.:

$$ \frac{\delta\alpha\left(\mathbf{r}\right)}{\alpha_0} = \frac{\delta\beta\left(\mathbf{r}\right)}{\beta_0} \, . \tag{4.29} $$

This allows us to simplify the inversion problem.

In the case of a general source which emits simultaneously P- and S- waves in proportions $S_P$, $S_S$ (with $S_P + S_S = 1$), we further generalize the decomposition of the kernels into elementary components. The procedure is best explained with the aid of an example. For the general travel time sensitivity kernel we write:

$$ {}^{v}K^{tt} = \sum_{Y=P,S} \sum_{X=P,S} \sum_{v=\alpha,\beta} W_{YX} {}^{v}K^{tt}_{YX} \, , \tag{4.30} $$

where the time-dependent weight of each mode $YX$ is given by

$$ W_{YX} = \frac{S_X E_{YX}}{\sum\limits_{X=P,S} \sum\limits_{Y=P,S} S_X E_{YX}} \, . \tag{4.31} $$

Using these weights a single sensitivity kernel can be obtained that appears similar to the kernel in the acoustic case for one wave mode and one velocity. However, the weights are lapse time dependent which changes the relative contribution of the different elastic kernels over time.

For further work like inversion, our new sensitivity kernels provide options to include information about the wave mode which can be obtained from array observations or other measurements of wavefield gradients for example with rotation sensors (Gaebler et al., 2015).

## 4.6 Computation of Sensitivity Kernels in Scattering Anomaly Model

Previous works have obtained sensitivity kernels with different assumptions about the scattering process (Pacheco & Snieder, 2006; Larose et al., 2010; Obermann et al., 2013b; Planès et al., 2014). However, most previous studies assumed that the scattering and attenuation properties are spatially homogeneous. Locations within stronger heterogeneity tend to concentrate seismic energy and thereby alter the sensitivity of the wavefield to perturbations in that region. Our approach is based on the energy density obtained from the radiative transfer simulations and allows us to take into account this change of the sensitivity by calculating the sensitivity kernels in media with spatial variations of the attenuation and scattering properties. To demonstrate this influence we calculate sensitivity kernels in a model with an anomaly in the scattering properties shown in Figure 3.13.

Figure 4.5 and Figure 4.6 show the resulting traveltime sensitivity kernels. Compared with Figure 4.1 and Figure 4.2, ${}^{\beta}K^{tt}_{PP}$ and ${}^{\alpha}K^{tt}_{SS}$ show significant differences in the areas of anomalous scattering. Please note that figures 4.5 and 4.6 show the kernels also at 13 s when the ballistic S-wave has passed through

the anomalous regions. To intuitively understand these differences, we take ${}^{\beta}K^{tt}_{PP}$ as an example in which the modes of excitation ($X$) and recording ($Y$) are both P-wave used to observe changes in S-wave velocity $\beta$. As discussed in section 4.2, the simplest case to create this sensitivity requires two conversion scattering events to happen, one before and the other after passing through $dV(\mathbf{r}')$. The probability of such an event must be higher in the anomaly area with stronger scattering which leads to more sensitivity to changes in S-wave velocity. It is different from ${}^{\alpha}K^{tt}_{PP}$ since there is no conversion needed. ${}^{\alpha}K^{tt}_{SP}$ and ${}^{\beta}K^{tt}_{SP}$ require a single scattering event only and are thus only weakly affected by local changes of the scattering properties. The same explanation applies to the kernels representing the traveltime sensitivity of measurements employing an S-wave source as shown in figure 4.6.

Decorrelation and energy sensitivity kernels for changes in fluctuation strength in this model with the spatially varying scattering are shown in Figure 4.7 and Figure 4.8, respectively. Compared with Figure 4.3 and Figure 4.4 showing the kernels in the statistically homogeneous model, the differences of all decorrelation and energy sensitivity kernels are minor. But since ${}^{\varepsilon}K^{E}_{YX}$ considers not only extra energy gained from more scattering with $f^{W\to V}(\mathbf{n},\mathbf{n}')E_{VX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_0})$ which ${}^{\varepsilon}K^{dc}_{YX}$ only considered but also the loss of energy due to the stronger scattering on $E_{YX}(\mathbf{r}',\mathbf{n},t';\mathbf{r_0})$, the scattering sensitivity kernels are more affected by scattering perturbation. Notice that at $t = 5$ s ${}^{\varepsilon}K^{dc}_{SS}$ and ${}^{\varepsilon}K^{E}_{SS}$ in the anomaly model looks stronger than in the homogeneous model since the direct S-wave has not arrived and the denominator of expressions has more influence on kernels.

## 4.7 Discussion

The scattering sensitivity kernels express the relation between a local change in the medium and an observation made on the wavefield. They describe how strongly an observation responds to a change at a particular location. We derive a number of kernels that relate different types of changes in the medium to different types of observations leading to a multitude of possible combinations. However, the equation of radiative transfer does only allow for two distinct mechanism to perturb the wavefield. The wave can be perturbed (A) during ballistic propagation and (B) while being scattered. Both mechanisms have their own spatial sensitivities, but all sensitivity kernels can be related to these two fundamental forms.

Mechanism A, i.e. the perturbation of wavefield attributes during unperturbed propagation is described by the passive kernel (Margerin et al., 2016). The travel time kernel ${}^{v}K^{tt}_{X,Y}$ and the attenuation kernel ${}^{Q}K^{E}_{X,Y}$ share the spatial shape of the passive kernel which describes the time that the waves have spent in a certain volume. The active kernel (Margerin et al., 2016) describes changes introduced by the mechanism B in which the attributes are unchanged but the wave propagation is perturbed. The decorrelation ${}^{\varepsilon}K^{dc}_{X,Y}$ has the shape of the active kernel. The energy kernels for changes in the scattering properties (${}^{\varepsilon}K^{E}_{X,Y}$) involve both, the active and the passive kernels. While the passive kernel describes the loss/gain of ballistic energy due to an increase/decrease of energy by scattering from the current ($\mathbf{n}$) into new propagation directions ($\mathbf{n}'$), the active kernel describes the increase/decrease of energy due to increased/decreased scattering from all other propagation directions ($\mathbf{n}'$) into the current direction ($\mathbf{n}$). Consequently ${}^{\varepsilon}K^{E}_{X,Y} = K_{active} - K_{passive}$.

The sensitivity kernels that we compute here are very detailed in terms of the wave mode that is excited and recorded. Since this degree of detail can hardly be used in any practical application, we give the recipe for combining the kernels to describe realistic situations. However, the kernels are derived independently of the incident direction as $K(r',t)$. Since all computations are based on the specific energy density which contains all information about propagation direction of the wavefield, we could easily derive expressions for kernels of the form $K(r',n,t)$ that describe the spatial sensitivity of a measurement performed on the specific energy density in a particular direction $\mathbf{n}$. Such information can be used to investigate the scattered field with seismic arrays and beamforming as it is typically used to study scattering in the deep

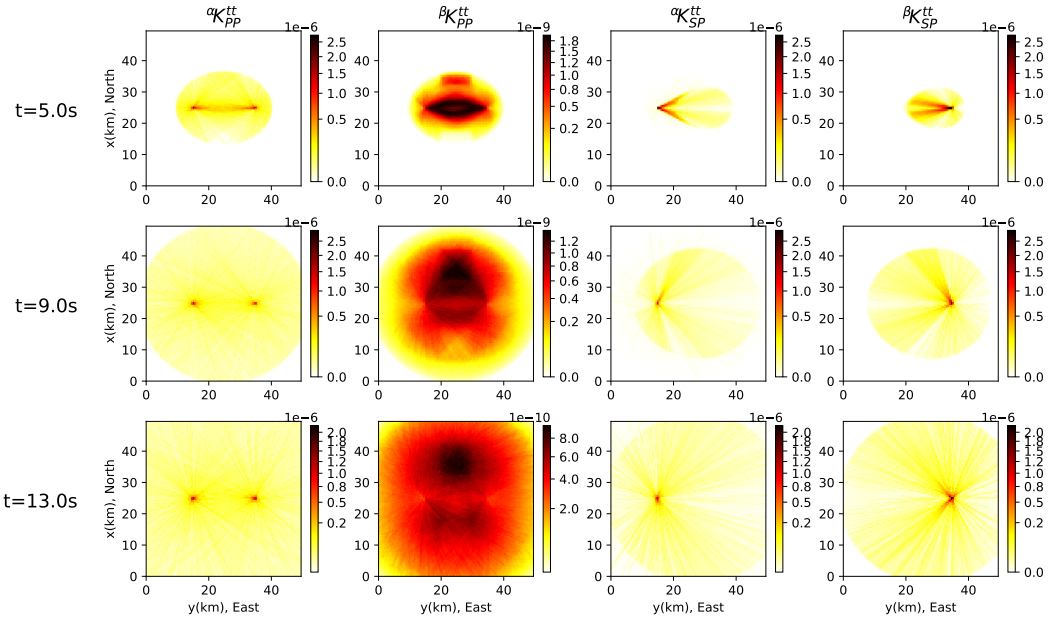



**Fig. 4.5:** Traveltime sensitivity kernels in medium with scattering anomaly: $^{\alpha}K^{tt}_{PP}$ (the 1st column), $^{\beta}K^{tt}_{PP}$ (the 2nd column), $^{\alpha}K^{tt}_{SP}$ (the 3rd column) and $^{\beta}K^{tt}_{SP}$ (the 4th column) at different lapse times with P-wave source. Note the 3rd column has a different color scale and all scales are nonlinear.

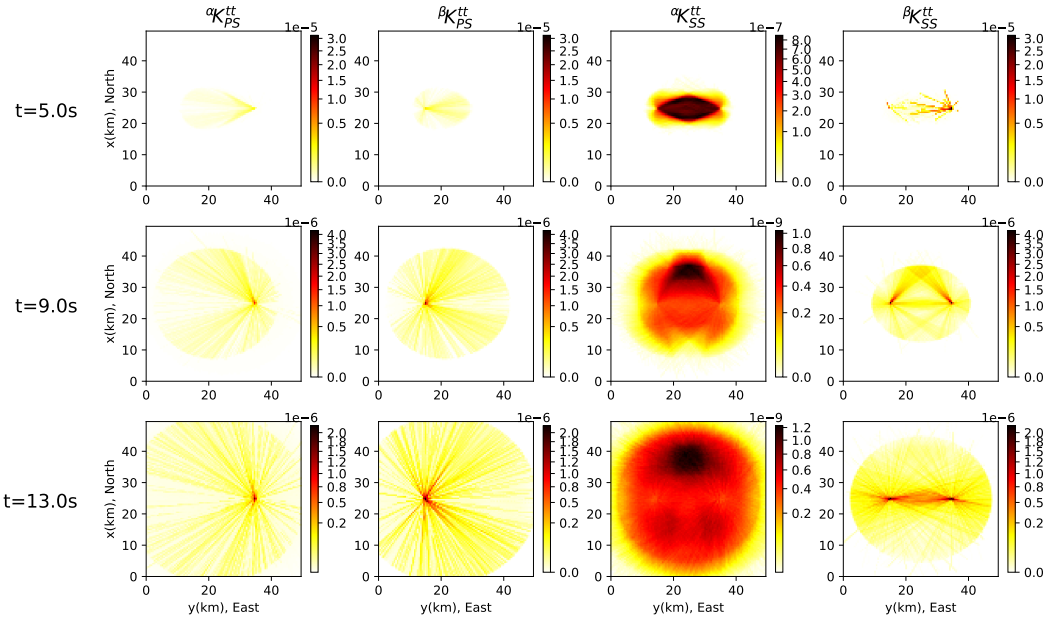



**Fig. 4.6:** Traveltime sensitivity kernels in medium with scattering anomaly: $^{\alpha}K^{tt}_{PS}$ (the 1st column), $^{\beta}K^{tt}_{PS}$ (the 2nd column), $^{\alpha}K^{tt}_{SS}$ (the 3rd column) and $^{\beta}K^{tt}_{SS}$ (the 4th column) at different lapse times with S-wave source. Note the 3rd column in each row has the different color scale and all scales are nonlinear.

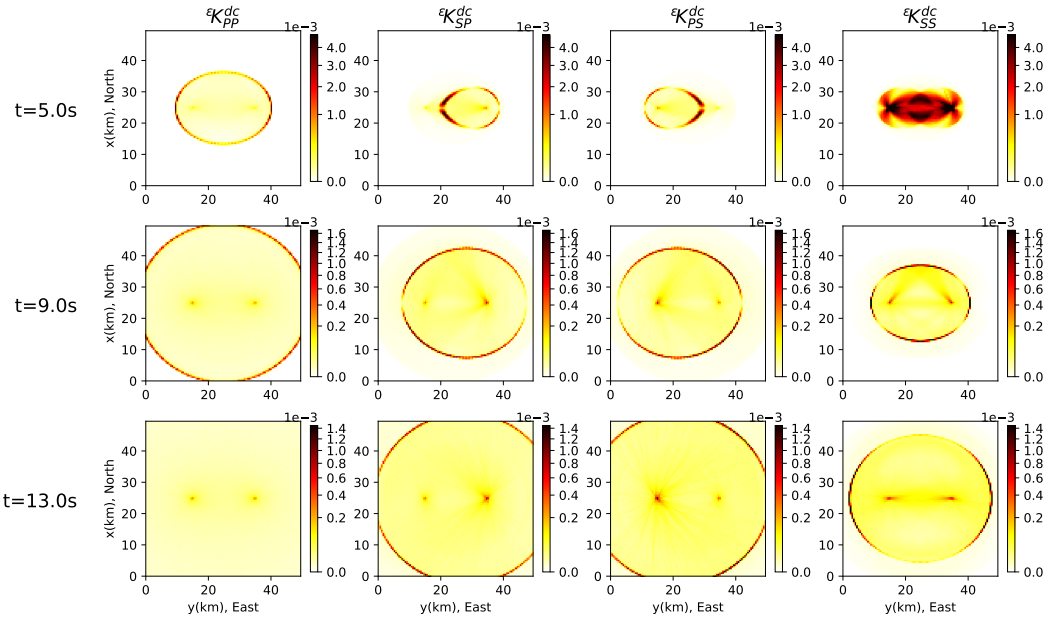



**Fig. 4.7:** Decorrelation sensitivity kernels in medium with scattering anomaly: ${}^{\varepsilon}K^{dc}_{PP}$ (the 1st column), ${}^{\varepsilon}K^{dc}_{SP}$ (the 2nd column), ${}^{\varepsilon}K^{dc}_{PS}$ (the 3rd column) and ${}^{\varepsilon}K^{dc}_{SS}$ (the 4th column) at different lapse times. Note all scales are nonlinear.

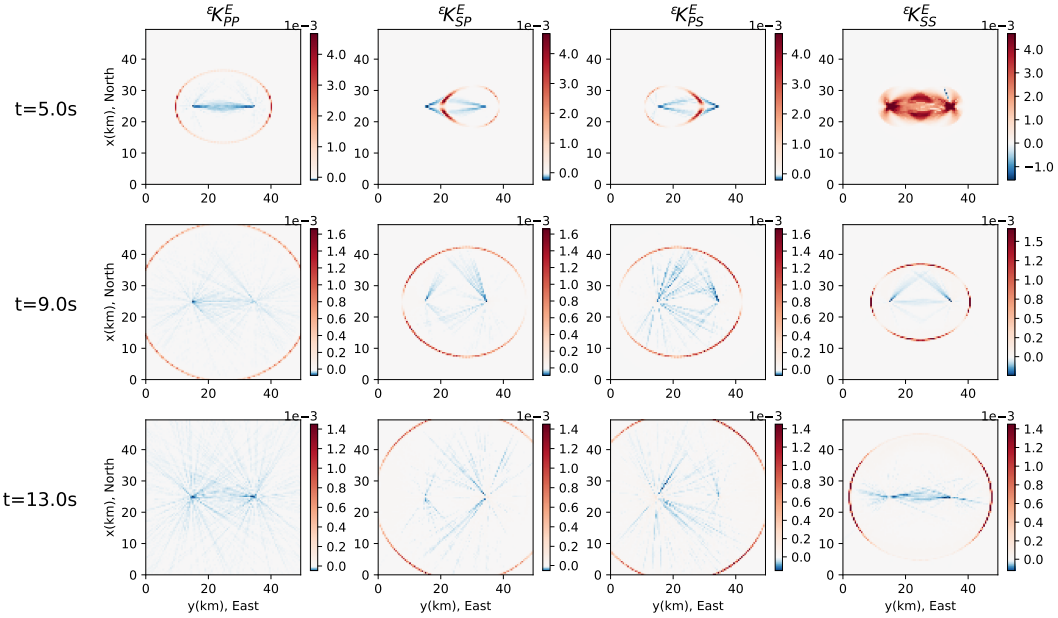



**Fig. 4.8:** Scattering sensitivity kernels in medium with scattering anomaly: ${}^{\varepsilon}K^{E}_{PP}$ (the 1st column), ${}^{\varepsilon}K^{E}_{SP}$ (the 2nd column), ${}^{\varepsilon}K^{E}_{PS}$ (the 3rd column) and ${}^{\varepsilon}K^{E}_{SS}$ (the 4th column) at different lapse times. The color of white indicates the value of 0, the red is positive and the blue is negative.

Earth (Lay & Garnero, 2011).

Compared with the previous studies (Mayor et al., 2014; Margerin et al., 2016) in the acoustic case, the elastic sensitivity kernels shown in Figures 4.1 through 4.4 constitute a major extension. The energy conversion and anisotropic scattering considered in this study result in an energy distribution that is very different from the previous work in the acoustic approximation. However, for example ${}^{\alpha}K^{tt}_{PP}$, ${}^{\varepsilon}K^{dc}_{PP}$ and ${}^{\varepsilon}K^{E}_{PP}$ show similar features as the acoustic case (Mayor et al., 2014), demonstrating that the acoustic approximation is reasonable. On the other hand the importance of treating anisotropic scattering was demonstrated by Margerin et al. (2016). Moreover, the acoustic approximation ignores conversion scattering which is especially important at short lapse times and results in asymmetric kernels. Together with the presence of two different wave speeds this causes four separate single-scattering ellipses which have strongly focused sensitivities in the active kernels. Considering only one of these ellipses in the acoustic approximation necessarily affects the tomographic inversion. To what extent a tomography is affected needs to be tested in a separate investigation, for which our work provides the means.

Implications of our results for previous studies that used sensitivity kernels derived from diffusion, multiple scattering or even based on empirical considerations are numerous (Sketsiou et al., 2020; Del Pezzo & Ibáñez, 2020). However in simplistic inversion approaches like imaging with space weighting functions or least square inversions of model misfit, the dependence of the results on the precise nature of the kernels is usually rather weak and we expects that the use of the elastic kernels would not have a major impact. But the method developed in the present paper allows for an iterative tomography in which the kernels are successively adapted to the improving model. In such an inverse problem the sensitivities need to be calculated with an accuracy that is comparable to the solution of the forward problem.

The elastic scattering process is complex and controlled by many parameters. In the present paper, we only consider changes of scattering strength in the form of the strength of the fluctuations. The correlation distance $a$ and the wavenumber $m$ directly affect the angular distribution of scattering angles. Since the PSDF is $\phi(m) = 2\pi\varepsilon^2 a^2 \left(1 + a^2 m^2\right)^{-3/2}$, it tends to a constant when $am \ll 1$ and the scattering is isotropic. When the scattering becomes isotropic, the traveltime sensitivity kernels we proposed approach the expression of Snieder et al. (2019). The PSDF here is the Fourier transform of an exponential ACF. Other possibilities are the Gaussian or von Kármán ACFs. The choice of the ACF influences the scattering process but our approach for the calculation of the sensitivity kernels is unaffected and can be applied with different ACFs.

The simulation is done in an infinite 2-D model and we do not consider the structure of the background velocity and density model either. Reflection and transmission will happen due to the impedance differences in a more complicated background structure. This is not to be confused with the stochastic scattering process used in section 3.1.1 and 4.6. Large-scale structure in the background velocity model can be included in the Monte-Carlo model as in Sanborn (2017); Sens-Schönfelder et al. (2009); Takeuchi (2016); Sens-Schönfelder et al. (2021).

Sensitivity kernels derived in this chapter have physical units. Table 4.1 therefore gives the units of the quantities involved in the calculation of the sensitivity kernels. The energy density is considered as the probability density of energy carrying particles. According to this table, we can confirm the units of

**Table 4.1:** The unit of different parameters.

| $E_{YX}(\mathbf{r},t;\mathbf{r_0})$ | $E_{YX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_0})$ | $c_W$ | $g_0^{V\to W}$ | $f^{V\to W}(\mathbf{n},\mathbf{n}')$ | $S^d$ | $dt$ | $d\mathbf{n}$ | $dV(\mathbf{r}')$ |
|---|---|---|---|---|---|---|---|---|
| $m^{-2}$ | $m^{-2}\cdot rad^{-1}$ | $m\cdot s^{-1}$ | $m^{-1}$ | $rad^{-1}$ | $rad$ | $s$ | $rad$ | $m^2$ |

the traveltime and decorrelation sensitivity kernels:

$$\begin{aligned}\left[{}^{v}K_{YX}^{tt}\right] =& rad\times\frac{m^{-2}\cdot rad^{-1}\times m^{-2}\cdot rad^{-1}\times s\times rad}{m^{-2}}\\ =& s\cdot m^{-2}\\ \left[{}^{\varepsilon}K_{YX}^{dc}\right] =& rad\times\frac{m\cdot s^{-1}\times m^{-1}\times m^{-2}\cdot rad^{-1}\times rad^{-1}\times m^{-2}\cdot rad^{-1}}{m^{-2}}\times s\times rad\times rad\\ =& m^{-2}\end{aligned} \tag{4.32}$$

where $[\cdot]$ denotes *Unit of* $\cdot$. The unit of attenuation sensitivity kernel is the same as that of the traveltime which together with the additional factor $\omega$ in equation 4.22 leads to the proper unit of relative energy density change. The units of the scattering sensitivity kernels are same as those of the decorrelation kernel.

## 4.8 Conclusions

With elastic radiative transfer theory, we simulate the propagation of seismic energy in the presence of wave scattering. The Monte-Carlo method is used to numerically solve the radiative transfer equations. Here we assume that the random velocity and density fluctuations of the medium have an exponential ACF and the scattering is anisotropic. In order to simulate energy transport in the presence of spatially variable fluctuation strength and intrinsic attenuation we separate the effects of fluctuation strength and ACF on the scattering coefficients and allow for location dependent fluctuation strength and quality factors. Two models are shown with spatially variable scattering and intrinsic attenuation to be compared with the statistically homogeneous model. The effects of stronger scattering and attenuation can be clearly observed in the two anomaly models.

As a further development we present the simulation of the specific energy density of the wavefield. The specific energy density $E_{YX}(\mathbf{r},\mathbf{n},t)$ describes the angularly resolved energy density at position $\mathbf{r}$ at time $t$ with the propagation direction $\mathbf{n}$. In the elastic case the mode of excitation $X$ and recording $Y$ can either be P- or S-wave. This quantity provides complete information about the energy transfer in an elastic medium with spatially variable randomness and intrinsic attenuation.

The complete information about the energy propagation allows for the computation of sensitivity kernels of scattered elastic waves including ballistic and scattered waves. For the efficient computation of the kernels we employ the reciprocity relation of an adjoint transport equation. We investigate sensitivity kernels in the form ${}^{\psi}K_{YX}^{\phi}$, where $\psi$ denotes the medium perturbation, $\phi$ denotes observable and $Y,X$ denote the excited wave mode $X$ and recorded wave mode $Y$. Both, the observable and medium property need to be specified to identify the kernels. ${}^{\alpha}K_{YX}^{tt}$ and ${}^{\beta}K_{YX}^{tt}$ are traveltime-velocity sensitivity kernels to describe the effect of P- and S-wave velocity perturbation in space on the traveltime perturbations of the seismogram. The derivation of these kernels is based on Bayes' theorem with a probabilistic interpretation of specific energy density and the reciprocity relation in transport theory. It leads to an expression that involves the convolution of the forward propagating field that is excited at the source and the adjoint propagating specific energy density that is excited at the receiver. By considering all combinations of the modes at source, receiver and perturbed location, the elastic sensitivity kernels turn out to be more complicated than the ones in the acoustic case proposed by Margerin et al. (2016). We show the eight

possible types of traveltime sensitivity kernels that result from the propagation of the two elastic wave modes.

Changes in the strength of random velocity and density fluctuations $\varepsilon$ can be observed as changes of the trace envelopes and decorrelation of the waveforms leading to the energy-scattering kernel ${}^{\varepsilon}K^{E}_{YX}$ and the decorrelation-scattering kernel ${}^{\varepsilon}K^{dc}_{YX}$. It is interesting to note that the energy-scattering kernel ${}^{\varepsilon}K^{E}_{YX}$ has positive and negative polarity whereas the decorrelation-scattering kernel ${}^{\varepsilon}K^{dc}_{YX}$ is strictly positive since any change in the scattering coefficient $\varepsilon$ (independent of its sign) will lead to an *increase* in decorrelation. Based on the observation of the decorrelation alone, it is thus not possible to discern an increase and a decrease of heterogeneity.

The functional form of the energy-attenuation kernels ${}^{Q}K^{E}_{YX}$ is the same as that of the velocity kernels. We also obtain eight different kernels for the combinations of the quality factors for P- and S-waves and the modes of excitation and recording. If the actual sources in an experiment emit both, P- and S-waves simultaneously or/and the receiver does not separate between P- and S-waves the different kernels can be superimposed with the suitable weighting which can be obtained from the simulation of the specific energy density.

We demonstrate the effect of spatial variations in scattering strength by comparing sensitivity kernels in a statistically homogeneous model with kernels calculated in a model that contains anomalies of the scattering properties. Obvious differences between kernels in the homogeneous and anomaly models exist in ${}^{\alpha}K^{tt}_{SS}$ and ${}^{\beta}K^{tt}_{PP}$ which are strongly affected by the scattering process because of the required mode conversion.

Localization of property changes still remains the challenge in the crust and in volcanic structures. However, inversion of the spatial perturbation of properties can be considered as an intuitive solution. Our work provides the technical basis for a probabilistic approach to tomography using the scattered elastic wavefield.

Chapter 5

# Adjoint Envelope Tomography

## 5.1 Introduction

Tomography as the method of inferring the spatial distribution of medium parameters from measurements of some physical observable has been successfully applied in different fields, like medicine, civil engineering and geophysics. In particular, seismic tomography uses seismograms to explore the structure of the Earth's interior (Dziewonski & Anderson, 1981). The different types of observations that can be made in seismic records offer various approaches to invert for physical properties of the Earth, leading to different classes of tomography. Since the traveltime of seismic waves is the most intuitive observation in seismic data, traveltime tomography was early developed to image the velocity or slowness structure (Aki & Lee, 1976; Thurber, 1983; Zhao et al., 1992; Li & Van Der Hilst, 2010). However, the obvious limitation of the traveltime tomography is that only the information of first arriving waves is used (Rawlinson et al., 2003) and information from later reflected or refracted wave arrivals is neglected. Exploiting all information contained in a seismogram including shape and amplitude of all arriving waves is the goal of full-waveform inversion (FWI) also referred to as adjoint tomography (Tarantola, 1984; Tromp et al., 2005; Fichtner et al., 2006; Fichtner, 2010). The inversion problem in FWI is solved iteratively by calculating the Fréchet derivative of the misfit function with respect to the model parameters that shows how the current model can be improved. The term *adjoint* tomography indicates that the interaction of the forward field with the adjoint field that satisfies the adjoint equation with the adjoint source derived from the misfit between the observed and the synthetic data is used to calculate the model gradient (Fichtner, 2010). With several decades of development, FWI has become a mature tool to reveal the velocity and attenuation structure of the medium (Tape et al., 2009; Fichtner et al., 2010; Zhu et al., 2012; Tao et al., 2018). Although FWI was successfully applied in global and regional imaging on multiple scales, its resolution is inherently limited by the wavelength of the observed wavefield. Structure on a length scale below this resolution limit cannot be resolved and causes misfit between modeled and observed wavefields that cannot be explained (Mancinelli et al., 2016). Therefore minor phases in the seismograms besides the main body wave or surface wave arrivals are often excluded from FWI (Liu & Gu, 2012).

The lithosphere and Earth's mantle, however, contain multi-scale heterogeneity ranging from the large-scale that can be interrogated deterministically to smallest scales far below the wavelength of seismic waves (Kennett et al., 2017). Well-log data provides the most direct evidence of this small-scale velocity heterogeneity (Holliger, 1996). Improving the resolution limits of FWI to image smaller scales faces the following limitations (A) computation costs for wavefield modelling increase drastically with frequency and (B) higher resolution requires an increase of the station and event density (Chauris, 2021).

For many applications, knowledge of the precise deterministic structure of the small-scale heterogeneity might not be required. Disregarding the geometric arrangement, the small-scale structure can be characterized by its statistical properties. In random medium theory it is assumed that heterogeneity can be regarded as a realization of a random process described by the amplitude of

fluctuations and a spatial correlation structure (Batchelor, 1953; Sato, 1982). Different from the velocities and density used to describe the elastic properties in deterministic structure, the random medium heterogeneity is characterized by the strength $\varepsilon$ and the spatial autocorrelation function of the fluctuations. Different correlation functions have been used in the literature but all contain a characteristic length scale called correlation length $a$. The statistical parameters $\varepsilon$ and $a$ of the small scale heterogeneity provide a description of the material that is complementary to the classical characterization of Earth materials based on macroscopic velocity (Sato et al., 2012).

The secondary phases resulting from the interaction of the wavefield with the small-scale heterogeneity form a continuous wave train that usually follows the first-arriving direct waves. These waves were first analyzed by Aki (1969) who coined the term coda waves and suggested that these are backscattered waves from the lateral heterogeneity. Aki & Chouet (1975) then proposed the single backscattering model and the diffusion model to interpret the transfer of seismic energy by the random heterogeneities. The scattering during seismic wave propagating from the source to the station causes a redistribution of seismic energy which involves loss of direct wave energy and the excitation of coda waves as a consequence of energy conservation (Sato et al., 2012). Another process that causes a decay of the seismic wave amplitude is intrinsic attenuation or absorption that converts seismic energy into other forms of energy. As a result, scattering attenuation and intrinsic attenuation are collectively referred to as seismic attenuation. Both intrinsic and scattering attenuation lead to an exponential decay of wave amplitude with traveled distance as described by the respective quality factors $Q_i$ and $Q_{sc}$. The temporal decay of coda wave amplitude can be described with the coda-$Q$ (Aki, 1980; Yoshimoto et al., 1993). However, the relationship between coda-$Q$ and scattering or intrinsic attenuation is variable, which naturally leads to the question of how to distinguish between the simultaneous influences of scattering and intrinsic attenuation on seismic waveforms.

In order to separate scattering from the effect of absorption, the multiple scattering model was proposed by Wu (1985). For the first time Wu (1985) introduced the radiative transfer theory (RTT) to seismology. RTT describes the multiple scattering process including back and forward scattering, which has the advantage of solving much more complicated problems (Wegler et al., 2006). To numerically solve the radiative transfer equation, the Monte-Carlo method was used by Gusev & Abubakirov (1987) and Hoshiba (1991). Following Wu (1985) and Hoshiba (1991), Fehler et al. (1992) proposed the multiple lapse-time window analysis method (MLTWA) to separate absorption and scattering, in which the seismic energy is integrated over different time windows since each attenuation parameter has a different influence on the individual lapse times. RTT is proved as an efficient tool to investigate absorption and seismic wave scattering of small-scale heterogeneity. Especially with the elastic radiative transfer equations derived by Weaver (1990) and Ryzhik et al. (1996) and the Monte-Carlo method applied in the elastic case (Margerin et al., 2000; Yoshimoto, 2000), RTT was widely used to synthesize seismogram envelopes (Wegler et al., 2006; Przybilla et al., 2006) and investigate scattering and intrinsic attenuation (Sens-Schönfelder & Wegler, 2006; Padhy et al., 2007; Eulenfeld & Wegler, 2016).

The approaches mentioned above are all based on the assumption of a uniform distribution of heterogeneity and constant intrinsic attenuation. Imaging the spatial distribution of non-uniform scattering and absorption properties has thus been a challenge. Using the MLTWA and following the work of Hoshiba (1993), Carcolé & Sato (2010) mapped the scattering loss parameter $Q_{sc}^{-1}$, intrinsic absorption $Q_i^{-1}$ and the seismic albedo $B_0$ (i.e. the ratio $Q_{sc}^{-1}/(Q_{sc}^{-1}+Q_i^{-1})$ of scattering attenuation to total attenuation) in Japan. A block with stronger scattering in a more homogeneous medium was used to model the lateral differences in the propagation of Lg-waves in the western Pyreneean crust with a Monte-Carlo simulation algorithm for the 3D elastic radiative transfer equations (Sens-Schönfelder et al., 2009). Another study of the spatially variable attenuation and scattering structure in the western Pyrenees was conducted by Calvet et al. (2013) based on the large lapse time coda decay as a proxy for

intrinsic attenuation factor and peak delay time as the indicator of scattering.

Scattering in volcanoes has been studied extensively as their internal structure is very heterogeneous and wave scattering ubiquitous. Spatial variations of heterogeneity have been investigated by De Siena et al. (2013) who employed a rim model with reflective diffusive boundary conditions to describe the distribution of heterogeneity at Campi Flegrei caldera, Italy. Prudencio et al. (2013) studied scattering and attenuation in Tenerife island and mount St Helens was investigated by De Siena et al. (2014b) and De Siena et al. (2016).

Compared to the FWI, these approaches for imaging heterogeneity and absorption have so far been rather simple. The above-mentioned approaches either prescribed a certain geometry of the anomalous regions (Sens-Schönfelder et al., 2009; De Siena et al., 2013) or estimated the parameters for a certain combination of source and station that is attributed to the straight line connecting source and receiver (Calvet et al., 2013; De Siena et al., 2014b). Although more complicated space-weighting functions like Gaussian weighting function (Prudencio et al., 2013) or functions based on scattering and absorption sensitivity kernels (Del Pezzo et al., 2016; Del Pezzo & Ibáñez, 2020) have been used, this approach is mostly a heuristic regionalization, that is in some cases applied to the non-physical medium parameter coda-Q that combines the effects of absorption and scattering (Jin & Aki, 2005; Bianco et al., 2002; Soergel et al., 2020). It is not an inversion of the nonlinear tomographic problem for the spatial distribution of heterogeneity and absorption that best describes the observations.

Takeuchi (2016) developed a differential Monte-Carlo method for simulating the perturbations of seismogram envelopes due to the spatial changes in scattering and intrinsic attenuation parameters, which offers a direct inversion method for the structure. This method approaches some ideas of the conventional waveform inversion and was applied by Ogiso (2019) for the 3D estimation of scattering and intrinsic attenuation in southwestern Japan.

This long-lasting quest for a reliable imaging method for attenuation and scattering on the one hand and the mature imaging technology of the full waveform inversion for velocity on the other hand inspired us to develop a new tomography method for the small-scale heterogeneity and high frequency absorption. The application of the adjoint method in FWI firstly benefited from the precise numerical solution of the wave equation (Komatitsch & Tromp, 1999; Fichtner & Igel, 2008), which allows simulating the seismic waves propagation in arbitrary models of velocities, density or other elastic properties.

The statistical properties that describe the small-scale heterogeneity have no direct expression in the wave equation which limits the applicability of FWI for spatial imaging of absorption and scattering. In our study, the forward problem is solved by modelling the multiple nonisotropic scattering in a random elastic medium based on the Radiative Transfer Equation using the Monte-Carlo method (Zhang et al., 2021). The spatial variability of scattering and absorption is described by the spatial distribution of fluctuation strength $\varepsilon$ and intrinsic quality factors $Q_P^{-1}$ and $Q_S^{-1}$ in the random medium. Key to the success of FWI is the efficient calculation of the gradient of the misfit function with respect to changes in the model parameters (Fréchet derivative) which is used to iteratively update the model for minimization of the misfit function. Although the choice of the misfit function in FWI can differ (Tromp et al., 2005; Yuan et al., 2016; Tao et al., 2017), it is in general used as the adjoint source to generate the adjoint wavefield which in turn is used to obtain the gradient for model updates. In the present paper we follow the same approach and use the least-squares misfit between the observed and modelled envelopes as adjoint source for the adjoint energy field to obtain the Fréchet derivatives.

## 5.2 Adjoint Tomography with the Radiative Transfer Equation

Full-waveform inversion (FWI) uses the discrepancy between observed seismograms and synthetic waveforms calculated for a particular model to iteratively update the model until it fits the observations

(Fichtner, 2010). The model usually describes the spatial distribution of elastic parameters like seismic velocities and the synthetics are obtained by solving the wave equation in this model. Various possibilities exist to quantify the discrepancy between the synthetics and the observation. Mathematically the discrepancy is described by the misfit function that maps the observed and synthetic waveforms to a scalar misfit value. Finding the gradient of the misfit with respect to all the different model parameters is the core of the adjoint method and allows to obtain an image of the elastic parameters in the subsurface.

Here we aim to formulate a similar approach for the inversion of seismogram envelopes using the adjoint method with the Radiative Transfer Equation (RTE). The model in our case does not describe the macroscopic seismic velocities as in FWI but the scattering and attenuation properties caused by the small-scale heterogeneity. Instead of comparing the waveforms, envelope inversion compares the seismogram envelopes with synthetic envelopes calculated on the basis of the RTE.

In scattering and absorbing media, spatial heterogeneity and intrinsic attenuation influence the energy propagation and absorption, which finally results in the energy density observed as seismogram envelope. To investigate these properties, we utilize the envelope of the full waveform including the ballistic wave and the scattered coda wave. We use the least-squares misfit function:

$$\chi_{YX}(\mathbf{m}) = \sum_i \sum_j \frac{1}{2} \int_0^T \left\| E_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \mathbf{m}) - D_{YX}(\mathbf{r_j}, t; \mathbf{r_i}) \right\|^2 dt \ . \tag{5.1}$$

$E_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \mathbf{m})$ is the synthetic energy density simulated in the current model $\mathbf{m}$, where $\mathbf{r_i}$ and $\mathbf{r_j}$ respectively represent the positions of the $i$-th source and the $j$-th receiver. In what follows, the subscript $YX$ of the energy density always indicates the wave mode $X$ of emission at the source and the mode of detection $Y$ at the receiver. For the S-wave energy we use the total S-energy as the sum of the energies of the two S-wave polarizations. This means that we disregard the specific influence of the S-wave polarization on the sensitivity kernels. $D_{YX}(\mathbf{r_j}, t; \mathbf{r_i})$ designates the observation data. In most practical applications it is impossible to disentangle the recorded energy with respect to its mode and the mode of excitation. But in order to facilitate the derivation, $\chi_{YX}(\mathbf{m})$ is used to represent each component of the total misfit function $\chi(\mathbf{m})$. After the derivation we will discuss how to extend this formulation to applications. The time window of the integral is somewhat arbitrary and indicates the range of lapse times that is to be chosen in the inversion but is assumed to start at time 0. Then the differential of misfit function reads:

$$\delta\chi_{YX}(\mathbf{m}) = \sum_i \sum_j \int_0^T [E_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \mathbf{m}) - D_{YX}(\mathbf{r_j}, t; \mathbf{r_i})] \delta E_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \mathbf{m}) dt \ , \tag{5.2}$$

where $\delta E_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \mathbf{m})$ is the perturbation of energy density due to the perturbation of the model $\mathbf{m}$. We assume here that there are no changes in the macroscopic velocity structure and the shape of the power spectral density function that describes the small scale fluctuations (see section 2.5). Under these assumptions, perturbations of the recorded energy density are either caused by changes in the amplitude of the small-scale fluctuations around the macroscopic velocity (e.g. due to the opening of cracks) that lead to changes of the scattering strength or due to changes of intrinsic attenuation described in transfer theory by the parameters fluctuation strength and intrinsic quality factors, respectively. A similar parameterization has been used for example by Takeuchi (2016). These dependencies are discussed in detail in Zhang et al. (2021) and can be expressed as

$$\delta E_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \mathbf{m}) = \delta E^{\varepsilon}_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \varepsilon(\mathbf{r}')) + \delta E^{Q}_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, Q(\mathbf{r}')) \tag{5.3}$$

where

$$\delta E^{\varepsilon}_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \varepsilon(\mathbf{r}')) = \frac{1}{\varepsilon_0^2} \int_{V^d} \delta\varepsilon^2(\mathbf{r}')\ {}^{\varepsilon}K^E_{YX}(\mathbf{r}', t; \mathbf{r_j}, \mathbf{r_i}) dV(\mathbf{r}') \tag{5.4}$$

and

$$\begin{aligned}\delta E^{Q}_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, Q(\mathbf{r}')) = -\omega \int_{V^d} [&\delta Q_P^{-1}(\mathbf{r}')\ {}^{Q_P}K^E_{YX}(\mathbf{r}', t; \mathbf{r_j}, \mathbf{r_i}) + \\ &\delta Q_S^{-1}(\mathbf{r}')\ {}^{Q_S}K^E_{YX}(\mathbf{r}', t; \mathbf{r_j}, \mathbf{r_i})] dV(\mathbf{r}')\ .\end{aligned} \tag{5.5}$$

Here $\varepsilon_0$ is the background value of the fluctuation strength. $Q_P$ and $Q_S$ are the intrinsic quality factors of P- and S-waves, respectively. $\omega$ is the angular frequency. $\delta\varepsilon^2(\mathbf{r}')$, $\delta Q_P^{-1}(\mathbf{r}')$ and $\delta Q_S^{-1}(\mathbf{r}')$ indicate the perturbation of $\varepsilon^2$, $1/Q_P$ and $1/Q_S$ at the position $\mathbf{r}'$ respectively. $V^d$ is the full space in dimension $d$. According to Zhang et al. (2021), the scattering sensitivity kernel and the attenuation sensitivity kernels read:

$$\begin{aligned}{}^{\varepsilon}K^E_{YX}(\mathbf{r}', t; \mathbf{r_j}, \mathbf{r_i}) = \\ \sum_W \int_{S^d} \int_0^t \Bigg\{ &G_{YW}(\mathbf{r_j}, t - t'; \mathbf{r}', \mathbf{n}) \times \\ \sum_V \Bigg[ \int_{S^d} &c_V\ g_0^{V\to W}(\varepsilon_0^2) f^{V\to W}(\mathbf{n}, \mathbf{n}') E_{VX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r_i}) d\mathbf{n}' \\ &- c_W g_0^{W\to V}(\varepsilon_0^2) E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i}) \Bigg] \Bigg\} dt' d\mathbf{n}\end{aligned} \tag{5.6}$$

and

$${}^{Q_P}K^E_{YX}\left(\mathbf{r}', t; \mathbf{r_j}, \mathbf{r_i}\right) = \int_{S^d} \int_0^t G_{YP}\left(\mathbf{r_j}, t - t'; \mathbf{r}', \mathbf{n}\right) E_{PX}\left(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i}\right) dt' d\mathbf{n} \tag{5.7}$$

$${}^{Q_S}K^E_{YX}\left(\mathbf{r}', t; \mathbf{r_j}, \mathbf{r_i}\right) = \int_{S^d} \int_0^t G_{YS}\left(\mathbf{r_j}, t - t'; \mathbf{r}', \mathbf{n}\right) E_{SX}\left(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i}\right) dt' d\mathbf{n}\ . \tag{5.8}$$

The notation of the specific energy density $E_{YX}(\mathbf{r}', t', \mathbf{n}')$ is used to indicate the dependence on the propagation direction $\mathbf{n}'$. $G_{YX}(\mathbf{r_j}, t - t'; \mathbf{r}', \mathbf{n})$ is Green's function of the radiative transfer equation with the unit source at the position $\mathbf{r}'$ and an initial direction $\mathbf{n}$. $g_0^{W\to V}(\varepsilon_0^2)$ is the total scattering coefficient that indicates the scattering probability of wave scattered from mode $W$ to $V$ with the background fluctuation strength $\varepsilon_0^2$. $f^{V\to W}(\mathbf{n}, \mathbf{n}')$ is the normalized differential scattering cross section. $c_V$ is the macroscopic velocity of wave mode $V$. $S^d$ denotes the unit sphere in space dimension $d$.

The iterative inversion requires Fréchet derivatives of the misfit function with respect to the model changes that can be obtained with the adjoint method (Tromp et al., 2005). We first discuss the misfit caused by variations of the scattering properties, that is the first term on the right-hand side of eq. (5.3):

$$\delta\chi^{\varepsilon}_{YX}(\mathbf{m}) = \sum_i \sum_j \int_0^T [E_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \mathbf{m}) - D_{YX}(\mathbf{r_j}, t; \mathbf{r_i})] \int_{V^d} \frac{1}{\varepsilon_0^2} \delta\varepsilon^2(\mathbf{r}')\ {}^{\varepsilon}K^E_{YX}(\mathbf{r}', t; \mathbf{r_j}, \mathbf{r_i}) dV(\mathbf{r}') dt \tag{5.9}$$

Inserting eq. (5.6) into eq. (5.9), exchanging the order of the integrals and the integration bounds we

rewrite eq. (5.9) as:

$$\begin{aligned}\delta\chi^{\varepsilon}_{YX}(\mathbf{m}) = &\sum_i\sum_j\int_{V^d}\int_0^T\int_{S^d}\sum_W\int_{t'}^T [E_{YX}(\mathbf{r_j},t;\mathbf{r_i},\mathbf{m}) - D_{YX}(\mathbf{r_j},t;\mathbf{r_i})]G_{YW}\left(\mathbf{r_j},t-t';\mathbf{r}',\mathbf{n}\right)\\ &\times\sum_V\Bigg[\int_{S^d} c_V\,\frac{g_0^{V\to W}(\varepsilon_0^2)}{\varepsilon_0^2}f^{V\to W}(\mathbf{n},\mathbf{n}')E_{VX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_i})d\mathbf{n}'\\ &- c_W\frac{g_0^{W\to V}(\varepsilon_0^2)}{\varepsilon_0^2}E_{WX}(\mathbf{r}',t',\mathbf{n};\mathbf{r_i})\Bigg]d\mathbf{n}dtdt'\delta\varepsilon^2(\mathbf{r}')dV(\mathbf{r}')\end{aligned} \tag{5.10}$$

This equation allows us to estimate the change of the misfit function due to a local change $\delta\varepsilon^2(\mathbf{r}')$. However, it involves the Green's function $G_{YW}\left(\mathbf{r_j},t-t';\mathbf{r}',\mathbf{n}\right)$ which in practical terms requires a simulation with a source at $\mathbf{r}'$ which is every point in space where a perturbation $\delta\varepsilon^2(\mathbf{r}')$ can occur. Using the adjoint method we can circumvent this problem and generate the Fréchet derivative with one simulation of the adjoint equation that has its source at the location of the original receiver. We omit the derivation that is almost the same as shown in Tromp et al. (2005) for the wave equation. Note that the reciprocity relationship of the elastic radiative transfer equations is (Zhang et al., 2021):

$$\frac{1}{S^d}G_{YW}\left(\mathbf{r_j},t-t';\mathbf{r}',\mathbf{n}\right) = G^{\dagger}_{WY}\left(\mathbf{r}',t-t',-\mathbf{n};\mathbf{r_j}\right)\ , \tag{5.11}$$

where $G^{\dagger}_{WY}\left(\mathbf{r}',t-t',-\mathbf{n};\mathbf{r_j}\right)$ is the Green function of an elastic radiative transfer equations with the new scattering coefficient $g^{\dagger W\to V}$ that obeys $c_V g^{V\to W} = c_W g^{\dagger W\to V}$. With this Green's function, the adjoint wavefield is written as:

$$E^{\dagger}_{WY}(\mathbf{r}',t',\mathbf{n};\mathbf{r_j}) = \int_{V^d}\int_0^{t'} G^{\dagger}_{WY}\left(\mathbf{r}',t'-t,\mathbf{n};\mathbf{r_j}\right)F^{\dagger}_{YX}(t,\mathbf{r}'')dtdV(\mathbf{r}'') \tag{5.12}$$

with the adjoint source:

$$F^{\dagger}_{YX}(t,\mathbf{r}'') = [E_{YX}(\mathbf{r_j},T-t;\mathbf{r_i},\mathbf{m}) - D_{YX}(\mathbf{r_j},T-t;\mathbf{r_i})]\delta(\mathbf{r}''-\mathbf{r_j})\ , \tag{5.13}$$

where $\delta(\mathbf{r}''-\mathbf{r_j})$ is the Dirac function. Note that the $\delta(\mathbf{r}''-\mathbf{r_j})$ in the expression of $F^{\dagger}_{YX}(t,\mathbf{r}'')$ indicates that the source position of the adjoint wavefield is the receiver position $\mathbf{r_j}$ of the forward wavefield. The expression of eq. (5.13) shows that the adjoint source is generated from the forward wavefield excited at $\mathbf{r_i}$ via the data misfit. Eq. (5.10) now can be written in a simplified form:

$$\delta\chi^{\varepsilon}_{YX}(\mathbf{m}) = \int_{V^d} {}^{\varepsilon}K^{\chi}_{YX}(\mathbf{r}')\delta\varepsilon^2(\mathbf{r}')dV(\mathbf{r}') \tag{5.14}$$

where ${}^{\varepsilon}K^{\chi}_{YX}(\mathbf{r}')$ is the misfit kernel with respect to fluctuation strength $\varepsilon$ which is:

$$\begin{aligned}{}^{\varepsilon}K^{\chi}_{YX}(\mathbf{r}') = &\sum_i\sum_j S^d\sum_W\int_0^T\int_{S^d}E^{\dagger}_{WY}(\mathbf{r}',T-t',-\mathbf{n};\mathbf{r_j})\times\\ &\sum_V\Bigg[\int_{S^d}c_V\,\frac{g_0^{V\to W}(\varepsilon_0^2)}{\varepsilon_0^2}f^{V\to W}(\mathbf{n},\mathbf{n}')E_{VX}(\mathbf{r}',t',\mathbf{n}';\mathbf{r_i})d\mathbf{n}'\\ &- c_W\frac{g_0^{W\to V}(\varepsilon_0^2)}{\varepsilon_0^2}E_{WX}(\mathbf{r}',t',\mathbf{n};\mathbf{r_i})\Bigg]d\mathbf{n}dt'\ .\end{aligned} \tag{5.15}$$

Note the difference to the energy sensitivity kernel ${}^{\varepsilon}K^{E}_{YX}(\mathbf{r}')$ of eq. 5.6 which does not depend on the data.

Similarly, the misfit caused by the intrinsic attenuation is obtained. Eq. (5.2) is finally written as:

$$\begin{aligned}\delta\chi_{YX}(\mathbf{m}) = &\int_{V^d} {}^{\varepsilon}K^{\chi}_{YX}(\mathbf{r}')\delta\varepsilon^2(\mathbf{r}')dV(\mathbf{r}') \\ &+ \int_{V^d}\left[{}^{Q_P}K^{\chi}_{YX}(\mathbf{r}')\delta Q_P^{-1}(\mathbf{r}') + {}^{Q_S}K^{\chi}_{YX}(\mathbf{r}')\delta Q_S^{-1}(\mathbf{r}')\right]dV(\mathbf{r}')\end{aligned} \tag{5.16}$$

where the intrinsic attenuation misfit kernels are:

$${}^{Q_P}K^{\chi}_{YX}(\mathbf{r}') = -\sum_i\sum_j S^d\omega\int_0^T\int_{S^d} E^{\dagger}_{PY}(\mathbf{r}', T-t', -\mathbf{n}; \mathbf{r_j})E_{PX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i})\,dt'd\mathbf{n} \tag{5.17}$$

$${}^{Q_S}K^{\chi}_{YX}(\mathbf{r}') = -\sum_i\sum_j S^d\omega\int_0^T\int_{S^d} E^{\dagger}_{SY}(\mathbf{r}', T-t', -\mathbf{n}; \mathbf{r_j})E_{SX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i})\,dt'd\mathbf{n}\,. \tag{5.18}$$

Eqs. (5.15), (5.17) and (5.18) describe the kernels with respect to the misfit function $\chi_{YX}(\mathbf{m})$ which compares the observed and simulated energy of recorded mode $Y$ due to excitation of mode $X$. These expressions are hard to apply since in most cases only the total energy $D(\mathbf{r_j}, t; \mathbf{r_i})$ is accessible. However, with some common assumptions about the composition of the wavefield we can solve this problem.

Without considering the S-wave polarization, we split the total energy into two parts as:

$$D(\mathbf{r_j}, t; \mathbf{r_i}) = D_P(\mathbf{r_j}, t; \mathbf{r_i}) + D_S(\mathbf{r_j}, t; \mathbf{r_i}) \tag{5.19}$$

The subscript of the observable indicates the recording mode of the energy. Zhang et al. (2021) discussed the combination of the sensitivity kernels in Eqs. (5.6), (5.7) and (5.8) based on the ratio of the two types of energies $R_{SP} = D_S/D_P$. This ratio can be reasonably approximated under the assumption of equipartition or it can be estimated using energy transfer simulations in the best available model. When the lapse time is large enough to reach the equipartition state $R_{SP} = 2\gamma^3$ in 3D, where $\gamma = c_P/c_S$ (Sato et al., 2012). Moreover, a general source simultaneously emits P and S energy in the proportions $S_P$ and $S_S$ with $S_P + S_S = 1$ for a unit source. The ratio between $S_P$ and $S_S$ that is required here, can either be estimated from the data or based on the assumption of a theoretical source model. So we can rewrite the misfit function:

$$\chi(\mathbf{m}) = \sum_i\sum_j\sum_Y\frac{1}{2}\int_0^T\left\|\sum_X S_X E_{YX}(\mathbf{r_j}, t; \mathbf{r_i}, \mathbf{m}) - D_Y(\mathbf{r_j}, t; \mathbf{r_i})\right\|^2 dt\,. \tag{5.20}$$

The differential of the misfit function becomes:

$$\begin{aligned}\delta\chi(\mathbf{m}) = &\int_{V^d}\sum_Y {}^{\varepsilon}K^{\chi}_{Y}(\mathbf{r}')\delta\varepsilon^2(\mathbf{r}')dV(\mathbf{r}') \\ &+ \int_{V^d}\left[\sum_Y {}^{Q_P}K^{\chi}_{Y}(\mathbf{r}')\delta Q_P^{-1}(\mathbf{r}') + \sum_Y {}^{Q_S}K^{\chi}_{Y}(\mathbf{r}')\delta Q_S^{-1}(\mathbf{r}')\right]dV(\mathbf{r}')\end{aligned} \tag{5.21}$$

where the scattering total misfit kernel is:

$$
\begin{aligned}
{}^{\varepsilon}K_Y^{\chi}(\mathbf{r}') = & \sum_i \sum_j S^d \sum_W \int_0^T \int_{S^d} E^{\dagger}_{WY}(\mathbf{r}', T-t', -\mathbf{n}; \mathbf{r_j}) \\
& \times \sum_X S_X \sum_V \left[ \int_{S^d} c_V \, \frac{g_0^{V\to W}(\varepsilon_0^2)}{\varepsilon_0^2} f^{V\to W}(\mathbf{n}, \mathbf{n}') E_{VX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r_i}) d\mathbf{n}' \right. \\
& \left. - c_W \frac{g_0^{W\to V}(\varepsilon_0^2)}{\varepsilon_0^2} E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i}) \right] d\mathbf{n} dt' \; .
\end{aligned}
\tag{5.22}
$$

Note the adjoint wavefield in this kernel is still based on Eq. (5.12) but the adjoint source now becomes:

$$
F_Y^{\dagger}(t, \mathbf{r}'') = [\sum_X S_X E_{YX}(\mathbf{r_j}, T-t; \mathbf{r_i}, \mathbf{m}) - D_Y(\mathbf{r_j}, T-t; \mathbf{r_i})] \delta(\mathbf{r}'' - \mathbf{r_j}) \; . \tag{5.23}
$$

With the same adjoint source applied, the total misfit kernels with respect to intrinsic attenuation are easy to obtain:

$$
{}^{Q_P}K_Y^{\chi}(\mathbf{r}' = -\sum_i \sum_j S^d \omega \sum_X S_X \int_0^T \int_{S^d} E^{\dagger}_{PY}(\mathbf{r}', T-t', -\mathbf{n}; \mathbf{r_j}) E_{PX}\left(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i}\right) dt' d\mathbf{n} \tag{5.24}
$$

$$
{}^{Q_S}K_Y^{\chi}(\mathbf{r}') = -\sum_i \sum_j S^d \omega \sum_X S_X \int_0^T \int_{S^d} E^{\dagger}_{SY}(\mathbf{r}', T-t', -\mathbf{n}; \mathbf{r_j}) E_{SX}\left(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i}\right) dt' d\mathbf{n} \; . \tag{5.25}
$$

## 5.3 Acoustic Case

For the propagation of a single wave mode in the acoustic case, the misfit function simplifies to:

$$
\chi(\mathbf{m}) = \sum_i \sum_j \frac{1}{2} \int_0^T \left\| E(\mathbf{r_j}, t; \mathbf{r_i}, \mathbf{m}) - D(\mathbf{r_j}, t; \mathbf{r_i}) \right\|^2 dt \; . \tag{5.26}
$$

We rewrite the Fréchet derivative as in section 5.2:

$$
\delta\chi(\mathbf{m}) = \int_{V^d} {}^{\varepsilon}K^{\chi}(\mathbf{r}') \delta\varepsilon^2(\mathbf{r}') dV(\mathbf{r}') + \int_{V^d} {}^{Q}K^{\chi}(\mathbf{r}') \delta Q^{-1}(\mathbf{r}') dV(\mathbf{r}') \; , \tag{5.27}
$$

where the misfit kernels are:

$$
\begin{aligned}
{}^{\varepsilon}K^{\chi}(\mathbf{r}') = & \sum_i \sum_j S^d \alpha_0 \frac{g_0(\varepsilon_0^2)}{\varepsilon_0^2} \int_0^T \int_{S^d} E^{\dagger}(\mathbf{r}', T-t', -\mathbf{n}; \mathbf{r_j}) \\
& \times \left[ \int_{S^d} f(\mathbf{n}, \mathbf{n}') E(\mathbf{r}', t', \mathbf{n}'; \mathbf{r_i}) d\mathbf{n}' - E(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i}) \right] d\mathbf{n} dt'
\end{aligned}
\tag{5.28}
$$

and

$$
{}^{Q}K^{\chi}(\mathbf{r}') = -\sum_i \sum_j S^d \omega \int_0^T \int_{S^d} E^{\dagger}(\mathbf{r}', T-t', -\mathbf{n}; \mathbf{r_j}) E(\mathbf{r}', t', \mathbf{n}; \mathbf{r_i}) d\mathbf{n} dt' \; . \tag{5.29}
$$

The adjoint wavefield is written as:

$$
E^{\dagger}(\mathbf{r}', t', \mathbf{n}; \mathbf{r_j}) = \int_{V^d} \int_0^{t'} G\left(\mathbf{r}', t'-t, \mathbf{n}; \mathbf{r_j}\right) F^{\dagger}(t, \mathbf{r}'') dt dV(\mathbf{r}'') \tag{5.30}
$$

with the adjoint source:

$$F^{\dagger}(t,\mathbf{r}'') = [E(\mathbf{r_j}, T-t; \mathbf{r_i}, \mathbf{m}) - D(\mathbf{r_j}, T-t; \mathbf{r_i})]\delta(\mathbf{r}'' - \mathbf{r_j}) \tag{5.31}$$

Note that the reciprocity relationship in the acoustic case is simpler than the elastic case without changing the scattering coefficient (Margerin et al., 2016):

$$\frac{1}{S^d} G\left(\mathbf{r_j}, t; \mathbf{r}', \mathbf{n}\right) = G\left(\mathbf{r}', t, -\mathbf{n}; \mathbf{r_j}\right) \tag{5.32}$$

## 5.4 Iterative Inversion

The target of the inversion is the minimization of the misfit function. This is achieved by iterative updates of the model converging toward the true model. The iterative inversion starts with the initial model ($\mathbf{m}_0^{\varepsilon}$, $\mathbf{m}_0^{Q}$) and uses the following initial search directions to update the model:

$$\mathbf{m}_{k+1}^{\varepsilon} = \mathbf{m}_k^{\varepsilon} + \eta_k^{\varepsilon}\mathbf{h}_k^{\varepsilon} \tag{5.33}$$

$$\mathbf{m}_{k+1}^{Q} = \mathbf{m}_k^{Q} + \eta_k^{Q}\mathbf{h}_k^{Q} \ . \tag{5.34}$$

Here, $\mathbf{m}_k^{\varepsilon}$ and $\mathbf{m}_k^{Q}$ are the models used in the simulation of the (k+1)-th iteration so that $k = 0, 1, 2, 3 \cdots$. $\mathbf{h}_k^{\varepsilon}$ and $\mathbf{h}_k^{Q}$ are the search directions derived from the the gradients of misfit function i.e. the misfit kernels ${}^{\varepsilon}K_k^{\chi}$ and ${}^{Q}K_k^{\chi}$. To derive the search directions from the gradients we employ the quasi-Newton L-BFGS method (Liu & Nocedal, 1989). The benefit of this method is that we obtain an approximation of the Hessian matrix of the misfit function to improve the convergence in comparison to a steepest decent method. More details are discussed in Section 5.5. $\eta_k^{\varepsilon}$ and $\eta_k^{Q}$ are the step lengths of the model updates in the resulting search directions. The step lengths are determined under the assumption that the misfit is a quadratic function of the step length (Fichtner, 2010). Using two more simulations with different trial step lengths in the $\mathbf{h}^{\varepsilon}$ and $\mathbf{h}^{Q}$ directions we approximate these quadratic functions and estimate the step lengths to reach their minima.

## 5.5 L-BFGS Method

The L-BFGS method is applied to approximate the Hessian matrix that allows us to generate the search direction with the gradient (kernel). For the search direction $\mathbf{h}_k$ of one parameter, we define the start model as $\mathbf{m}_k$ with the the gradient $\mathbf{K}_k$ in the (k+1)-th iteration. The L-BFGS algorithm is shown in Algorithm 1 (Nocedal & Wright, 2006). The number of most recent iterations $n$ is chosen for saving memory in this quasi-Newton method. In our tests, it is set between three and the maximum since the number of iterations is not huge. Note that $k$ in this algorithm must be more than 0 so it does not work for the first iteration. We use the steepest descent method to calculate the first search direction $\mathbf{h}_0$.

## 5.6 Numerical Examples

In the previous section we have outlined the mathematical formalism for an iterative adjoint tomography of the radiative transfer equation to infer the distribution of heterogeneity in the subsurface with scattered seismic waves. In this section we use numerical examples to demonstrate the performance of this approach. We employ three synthetic examples in a 2D acoustic setting to illustrate how to realize the adjoint tomography with the envelope for scattering and intrinsic attenuation. Two individual inversions are

**Algorithm 1** L-BFGS

$\mathbf{q}_k \leftarrow \mathbf{K}_k$,
chose the number of most recent iterations $n$, $n \leq k$
**for** $i = k-1$ to $k-n$ **do**
  $\mathbf{s}_i = \mathbf{m}_{i+1} - \mathbf{m}_i$,
  $\mathbf{y}_i = \mathbf{K}_{i+1} - \mathbf{K}_i$,
  $\rho_i = 1/(\mathbf{y}_i^T \mathbf{s}_i)$,
  $\alpha_i = \rho_i \mathbf{s}_i^T \mathbf{q}_{i+1}$,
  $\mathbf{q}_i = \mathbf{q}_{i+1} - \alpha_i \mathbf{y}_i$,
**end for**
$\mathbf{r}_{k-n} = \left[(\mathbf{s}_{k-1}^T \mathbf{y}_{k-1})/(\mathbf{y}_{k-1}^T \mathbf{y}_{k-1})\right] \mathbf{q}_{k-n}$,
**for** $i = k-n$ to $k-1$ **do**
  $\beta_i = \rho_i \mathbf{y}_i^T \mathbf{r}_i$,
  $\mathbf{r}_{i+1} = \mathbf{r}_i + (\alpha_i - \beta_i) s_i$,
**end for**
$\mathbf{h}_k \leftarrow \mathbf{r}_k$.

performed separately for the scattering and intrinsic attenuation structures separately with a uniform distribution of the other property. The trade-off between heterogeneity and attenuation structure is investigated in a third inversion in which the seismogram envelopes are inverted for the spatial distribution of both parameters simultaneously. The macroscopic model parameters mean seismic velocity and mean density are uniform in space. Forward and adjoint simulations are performed based on the radiative transfer equation that was introduced by section 2.5 using the algorithm presented by Zhang et al. (2021).

### 5.6.1 Scattering Inversion

The aim of the first inversion is the reconstruction of the medium heterogeneity, i.e. the spatial distribution of $\varepsilon$ which characterizes the amplitude of the small-scale fluctuation of the elastic parameters around the large scale macroscopic mean. We call this *scattering inversion* as the heterogeneity directly causes scattering. Intrinsic attenuation as described by $Q^{-1}$ is constant throughout the domain.

The observation data in this synthetic test is obtained with a Monte-Carlo simulation using Eq. (2.14) in the true model that is shown in Figure 5.1a. The intrinsic attenuation $Q^{-1}(\mathbf{r})$ is uniform in space (Figure 5.1b). For the application of FWI, a starting model is required that is a fair representation of the target (Fichtner, 2010) and has to be inferred using a complementary method. In our case we assume that the diffusion approximation (Wegler, 2004) or the multiple isotropic scattering approximation (Sens-Schönfelder & Wegler, 2006) can be used to infer the background values of heterogeneity and attenuation.

Inspired by the scale of local to regional seismological experiments (Gaebler et al., 2019; Izgi et al., 2020; van Dinther et al., 2021), the model extends over a 50 km by 50 km area and is discretized in a 100 by 100 cells grid (Figure 5.1). There are two square anomaly areas in the $\varepsilon$ model with a side length of 10 km. Twenty-four numbered locations around the anomaly areas in Figure 5.1 represent the co-located positions of the sources and receivers. When the source is at one location, all other locations are the receivers, a layout inspired by seismic interferometry (Curtis et al., 2006) that can turn receivers into virtual sources and the use of transducers in acoustic experiments that can act either as source or receiver (Niederleithinger et al., 2015).

The macroscopic velocity is spatially uniform at 6 km/s and the density is 2.7 g·cm$^{-3}$. The correlation length $a$ is constant 0.3 km. With an angular frequency $\omega = 30$ rad·s$^{-1}$ (i.e. $f = 4.78$ Hz) representing high frequency seismic waves, the normalized wavenumber $ak_0$ is 1.5. Since the forward modelling is conducted with the Monte-Carlo method, 100 million particles are used to simulate the energy transport

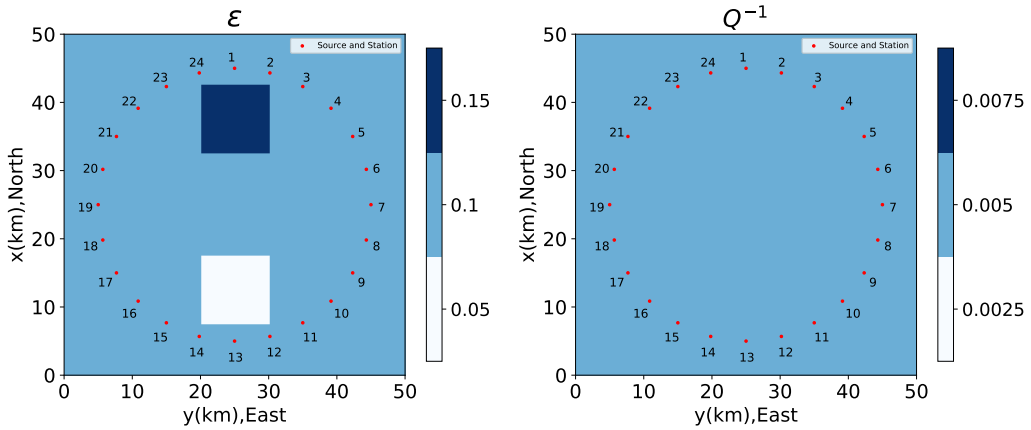



**Fig. 5.1:** The illustrations of the scattering anomaly model. The background of $\varepsilon$ (left) is 0.1. The white and the dark blue are two anomaly areas, respectively with the value 0.05 and 0.15. The attenuation model (right) is homogeneous with the value 0.005. The red points represent both the sources and receivers. Note that the model we demonstrated here is the partial space and there is no boundary for the simulation.

in the forward simulation for the true model to create the synthetic data. Although the propagation of particles is free and there is no boundary of the model, a $100 \times 100 \times 150$ grid of cell size 0.5 km $\times$ 0.5 km $\times$ 0.1 s is used for recording the energy and the propagation direction is recorded in 72 non-overlapping angular bins of 5° width.

The iterative inversion starts with a uniform initial model of $\varepsilon = 0.1$ that has to be inferred independently. The corresponding total scattering coefficient $g_0$ is 0.157 km$^{-1}$ which is equivalent to an inverse scattering quality factor $Q_{sc}^{-1} = 0.0314$. The misfit obtained from the differences between the simulations in the initial model and the observation (here it is the synthetic data simulated in the true model) is used in the adjoint method that we introduced in section 5.3 to obtain the gradient of the misfit function (sensitivity kernel) for each source-receiver couple. All kernels for every combination are then summed to compute the misfit kernel for updating the model. The energy at the sources and the stations (sources of the adjoint field) is much higher than elsewhere in the model resulting in singularities of the gradients at these positions. In FWI this influence can be removed by a pre-conditioner that was designed to approximate the Hessian matrix or its diagonal terms (Zhu et al., 2015). In this study, we use a simple approach: a normalized Gaussian distribution subtracted from 1 $(1 - \exp[-(\mathbf{r} - \mathbf{r_i})^2/2\sigma^2])$ is multiplied with the kernel for each source/sensor location $r_j$ to suppress the singularities of the gradient. Additionally, we also apply a Gaussian smoothing filter to the kernel, resulting in final misfit gradient. This step is usually regarded as a regularization procedure that guarantees the stability of inversion (Zhu et al., 2015; Tao et al., 2018). For details about the regularization which we applied here, please refer to Appendix 5.7.5.

Using the L-BFGS method the search direction of the $\varepsilon^2$ model is obtained. In each iteration, two tentative step lengths are tested by separate forward simulations and calculation of the misfit. The best step length is determined as the minimum of a quadratic function fitted to the three misfit values obtained for the initial model of the current iteration (step length zero) and the two tested step lengths. With this step length, the initial model is updated each iteration. This workflow is repeated several times and the model is iteratively improved until the misfit converges. Figure 5.2 shows the evolution of the normalized misfit over 8 iterations which shows that after about 5 iterations the misfit starts to converge to 36%. The slight increase in the misfit curve at the 6th and 8th iteration is due to changes in the details of the model during the model update and will be discussed later. The rather rapid convergence results from

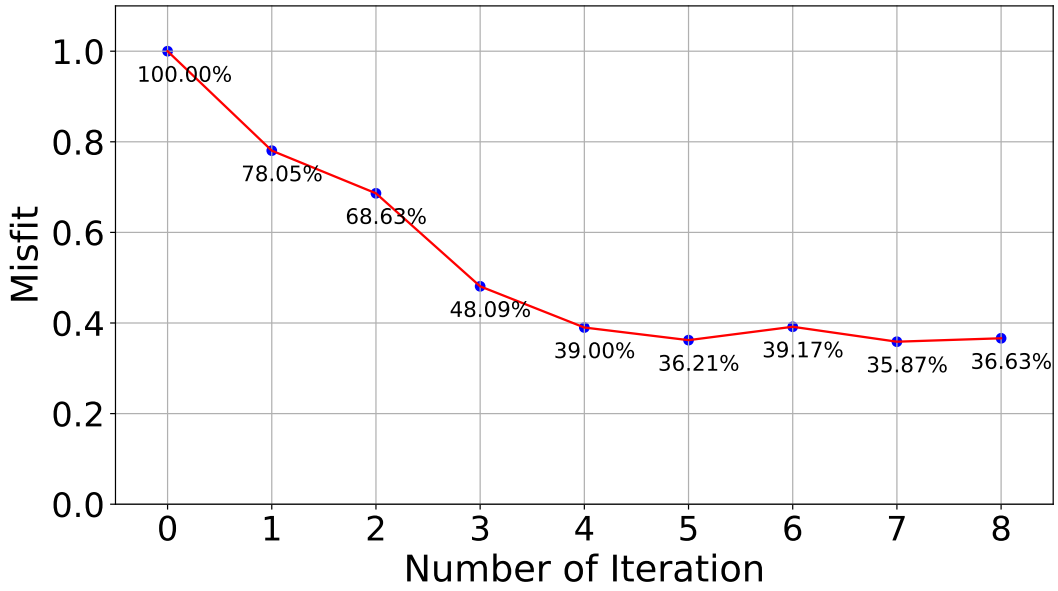


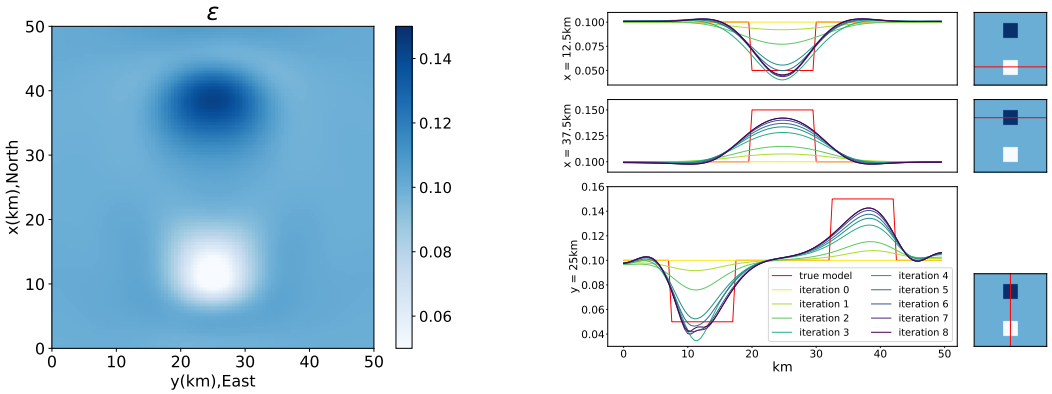

**Fig. 5.2:** The misfit in the scattering inversion experiment alters with the number of iterations. The values are normalized by the misfit between the uniform initial model and the true model shown in Figure 5.1.

the use of the L-BFGS method, the testing of the step length, and the good ray coverage (24 sources $\times$ 23 receivers). But it is also clear that a significant improvement of the model is achieved after the first iteration, documenting the nonlinearity of the problem and the benefit of the iterative inversion over the single-step kernel based approaches (e.g. Obermann et al., 2013a; Sánchez-Pastor et al., 2019).

Figure 5.3 illustrates the inversion result of $\varepsilon$ after 8 iterations. Although the initial model is uniform which means that there is no a prior information given about the perturbations, the two anomaly areas of the true model (Figure 5.3a) are successfully retrieved in location, shape and amplitude. To quantitatively evaluate the results, three profiles across the anomaly at $x = 12.5\ km, x = 37.5\ km, y = 25\ km$ are shown in Figure 5.3b. The evolution of the model in each iteration is shown with differently colored curves ranging from the initial model (light yellow) to the final result (dark purple) compared with the true model (red).

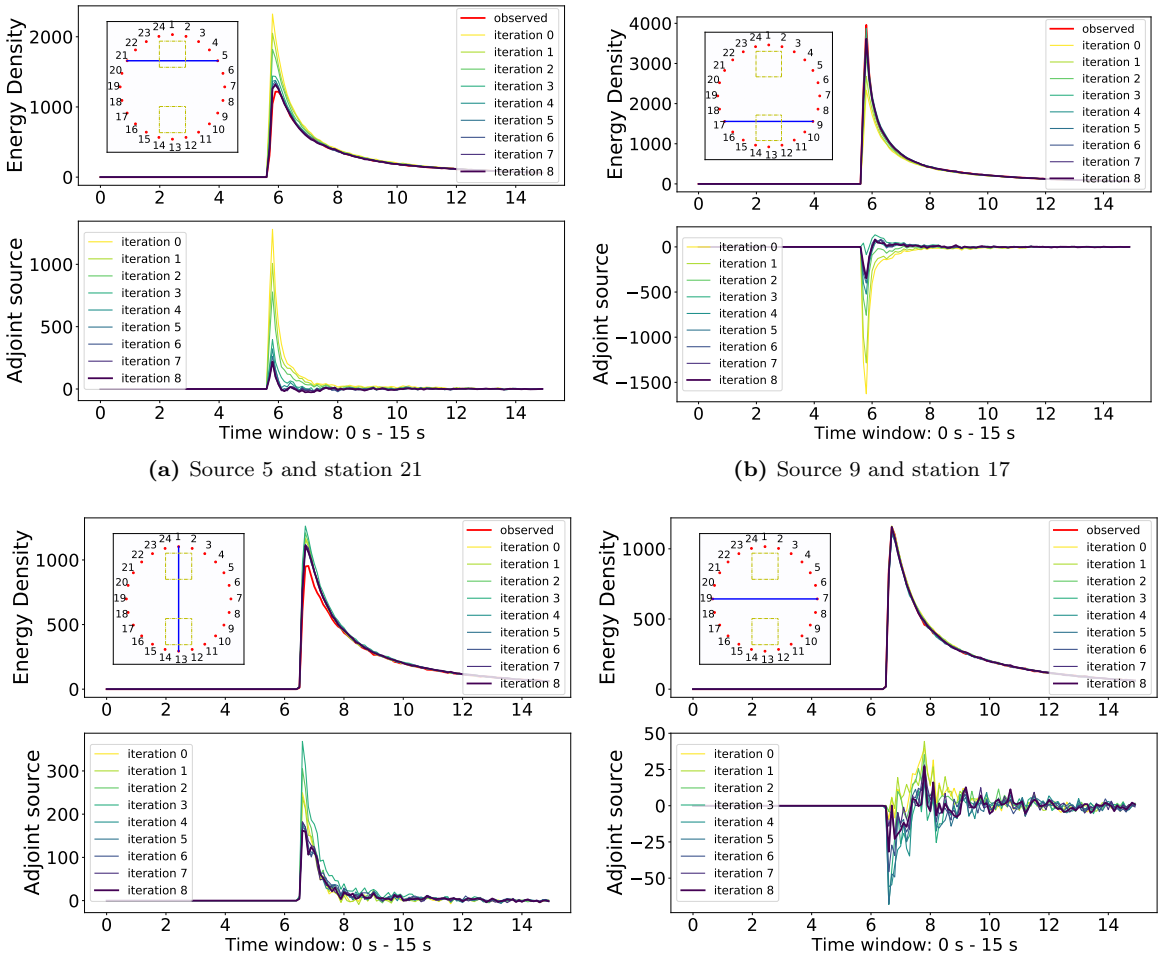
The location of the anomalies is picked up already in the 1st iteration while the amplitudes are gradually fitted with successive iterations. This process is not monotonic but can include some oscillations as seen for the negative anomaly where some overshoot in the central part after the fifth iteration is corrected in subsequent iterations. For this reason the misfit curve slightly fluctuates around 36% after the fifth iteration. Smaller step lengths can improve this behaviour but will increase the computational costs and are more likely to result in local minima. Both in Figure 5.3a and Figure 5.3b, the final result shows that the anomaly areas are smooth and do not exactly match the true model. This limitation is partially due to the Gaussian smoothing we use to in the preconditioning of the gradient, but also due to the setup of sources and receivers around the target area. Gaussian smoothing is chosen here as a simple method for demonstration and should be replaced in further applications by more advanced regularization techniques that are, for example, more suitable to resolve sharp-sided anomalies (Akcelik et al., 2002).

In a realistic inversion scenario, one cannot compare the inverted model with the true model. The only possibility to assess the final result, is the comparison of the modeled and measured envelopes. Figure 5.4 shows envelopes of four source-receiver combinations and the adjoint sources (misfit) in the different iteration from the initial (light yellow) to the final (dark purple) compared with the observation




**(a)** **(b)**

**Fig. 5.3:** The scattering inversion results: (a) the final model of $\varepsilon$ after 8 iterations; (b) three profiles across the anomalous areas (illustrated in the right column) of the updated model in each iteration from the initial model (light yellow) to the final model (dark purple) compared with the true model (red).




**(a)** Source 5 and station 21

**(b)** Source 9 and station 17

**(c)** Source 1 and station 13

**(d)** Source 7 and station 19

**Fig. 5.4:** The comparison of observed (red) and synthetic envelopes in each iteration from the initial (light yellow) to the final (dark purple) as well as the corresponding adjoint sources in the scattering inversion.

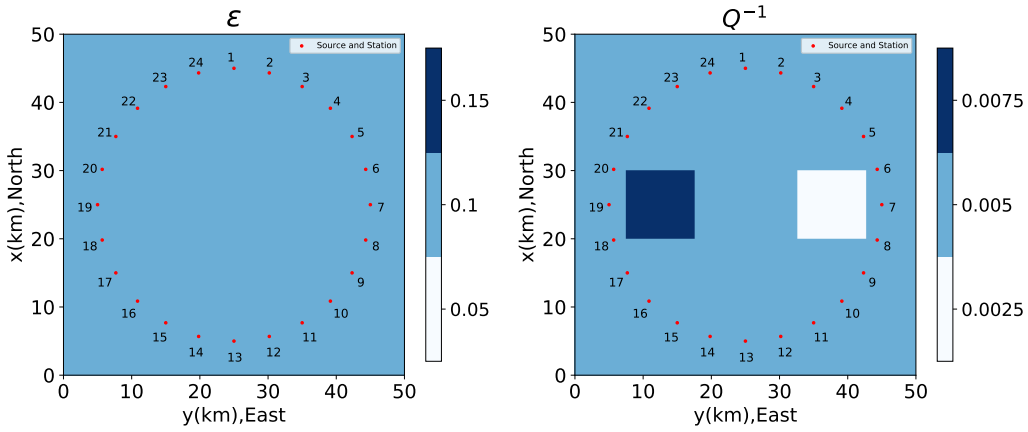



**Fig. 5.5:** The illustrations of the intrinsic attenuation anomaly model. The background of $Q^{-1}$ (right) is 0.005. The white and the dark blue are two anomaly areas, respectively with the value 0.0075 and 0.0025. The $\varepsilon$ model (left) is homogeneous with the value 0.1.

(red). Since the initial model is homogeneous, the synthetic energy density of direct wave is larger than the observations in Figure 5.4a that shows a source-receiver combination in which the ballistic wave passes through the stronger scattering anomaly. This results in increased energy loss of the ballistic wave. The amplitude of the adjoint source gradually decreases indicating convergence of the modelled envelope towards the data obtained in the true model. The same but with opposite sign can be observed in Figure 5.4b that shows a source-receiver configuration for which the ballistic wave passes through the anomaly with reduced heterogeneity. The energy of the ballistic peak is underestimated in the initial model and gradually increases during the inversion.

Figure 5.4c shows a source-receiver combination that is similarly affected by both anomalies as the ballistic wave passes through both anomaly areas. Since the effects of both anomalies compensate to some extent the absolute amplitude of the adjoint source is smaller than in figures 5.4a and 5.4b. However, as the effect of the increased heterogeneity on $\varepsilon^2$ is larger than the effect of the reduced heterogeneity for equal $|\Delta\varepsilon|$ the adjoint source is positive and decreases in amplitude during the inversion.

The absolute amplitude of the adjoint source for the source-receiver combination for which the direct wave does not pass through any anomalous region (Figure 5.4d) is even smaller. It has positive and negative segments that are improved during the inversion but it is devoid of a signal associated to the ballistic peak that is correctly represented by the initial model. The fluctuation of the adjoint source in Figure 5.4d mainly represents the statistical fluctuations of the simulations with the Monte-Carlo method and also contains some scattered waves from distant scatterers.

### 5.6.2 Intrinsic Attenuation Inversion

In this section, we conduct an inversion for the intrinsic attenuation only and fix the strength of the heterogeneity $\varepsilon$. We call it *intrinsic attenuation inversion*. Figure 5.5 shows the new model in which there are two anomaly areas in the $Q^{-1}$ model with values of 0.0075 and 0.0025, respectively whereas $\varepsilon$ is uniform in space. The parameters of the background and the setup of the grid are the same as in section 5.6.1.

Again the iterative inversion starts form a uniform initial model and the workflow follows section 5.6.1 but now the model of $Q^{-1}$ is updated in each iteration. Figure 5.6 shows the evolution of the model misfit during the inversion. It converges to a value of about 40% after the 7th iteration. The resulting

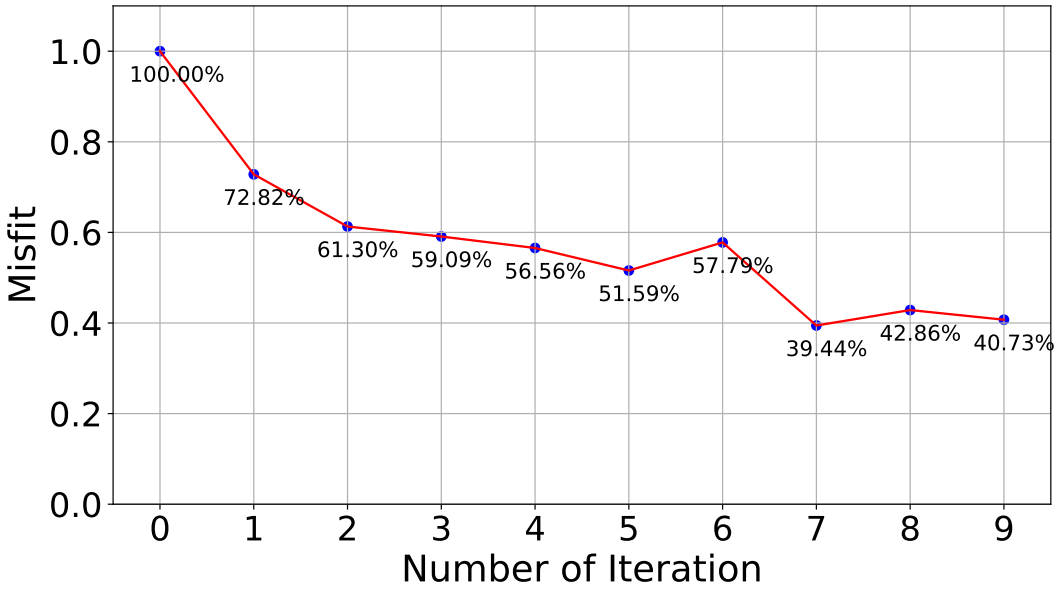



**Fig. 5.6:** The misfit in the intrinsic attenuation inversion experiment alters with the number of iterations. The values are normalized by the misfit between the uniform initial model and the true model shown in Figure 5.5.

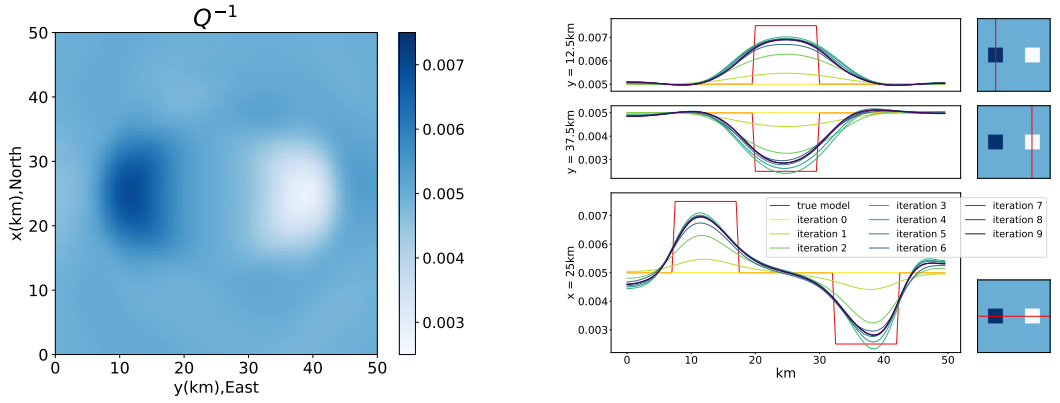
model of the 9th iteration is shown in Figure 5.7. Both anomaly areas are recovered (Figure 5.7a) but their shape is smoother compared to the inversion result for heterogeneity in section 5.6.1.

The envelope fits are shown in Figure 5.8 together with the adjoint sources. Figures 5.8a and 5.8b show the source-receiver configurations with ballistic wave passing through the areas of smaller and larger intrinsic attenuation, respectively. Both envelopes are fitted in a similar way over the course of the iterations leading to a vanishing adjoint source of the ballistic waves and its early coda. The two source-receiver combinations shown in Figures 5.8c and 5.8d that are affected by both anomaly areas again have significantly smaller adjoint sources fluctuating around zero. No significant changes occur here during the course of the inversion. This is to be expected since both anomalous areas have similar influence on the envelope as a shift of $Q^{-1}$ in both anomaly areas is 0.0025.

### 5.6.3 Simultaneous Inversion

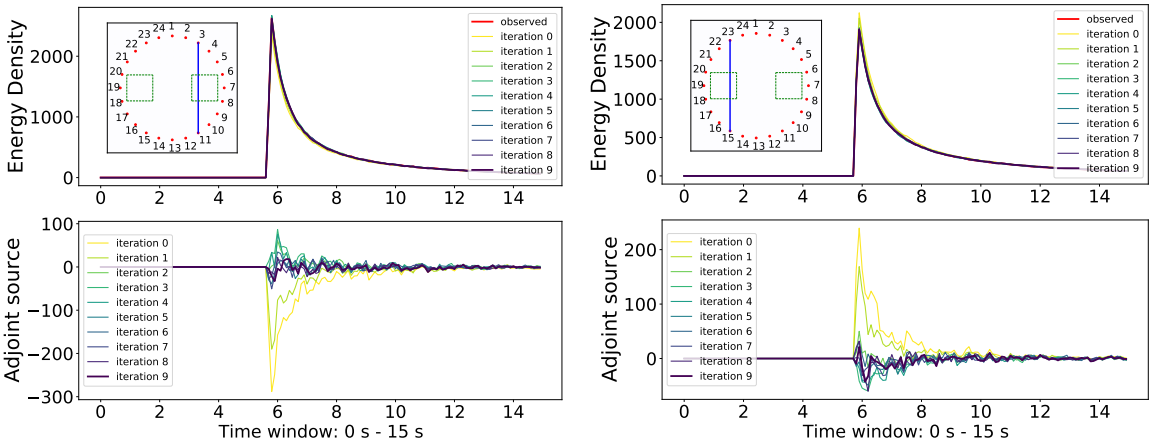
In the previous sections we demonstrated that the iterative adjoint envelope tomography allows to recover anomalies in the distribution of small-scale heterogeneity as well as in intrinsic attenuation. However, in both cases we used a uniform model for the other property and did not update the respective part of the model. This is not possible in a real application in which it is a prior unknown whether an observed misfit between model and data is caused by attenuation or scattering. Consequently both parameters of the model have to be updated simultaneously.

This simultaneous inversion of scattering and intrinsic attenuation is much more complicated than the separate inversions. In this section, we combine the $\varepsilon$ model that we used in section 5.6.1 for scattering inversion and the $Q^{-1}$ model from section 5.6.2 for the inversion of intrinsic attenuation. Both scattering and intrinsic attenuation of the true model have two anomaly areas as shown in Figure 5.9. The workflow of the simultaneous inversion is almost the same as the separate inversion. The only difference is that both $\varepsilon$ and $Q^{-1}$ are updated in each iteration. As a result, the simultaneous update of the models requires two different step lengths respectively for $\varepsilon$ and $Q^{-1}$. We have introduced in section 5.4 the assumption that the relationship between the misfit and each step fits the quadratic function. Here, we estimate



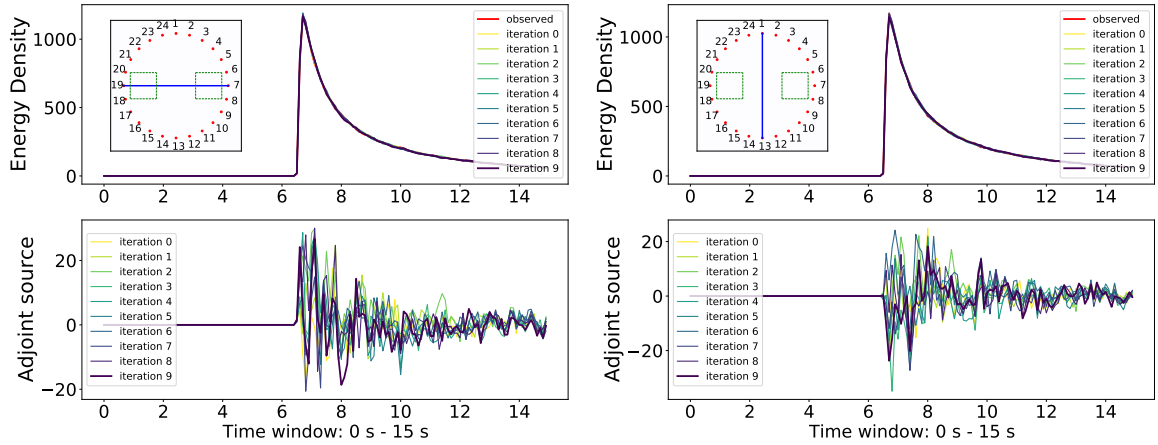

**(a)** **(b)**

**Fig. 5.7:** The intrinsic attenuation inversion results: (a) the final model of $\varepsilon$ after 9 iterations; (b) three profiles across the anomalous areas (illustrated in the right column) of the updated model in each iteration from the initial model (light yellow) to the final model (dark purple) compared with the true model (red).




**(a)** Source 3 and station 11

**(b)** Source 15 and station 23




**(c)** Source 7 and station 19

**(d)** Source 1 and station 13

**Fig. 5.8:** The comparison of observed (red) and synthetic envelopes in each iteration from the initial (light yellow) to the final (dark purple) as well as the corresponding adjoint sources in the intrinsic attenuation inversion.

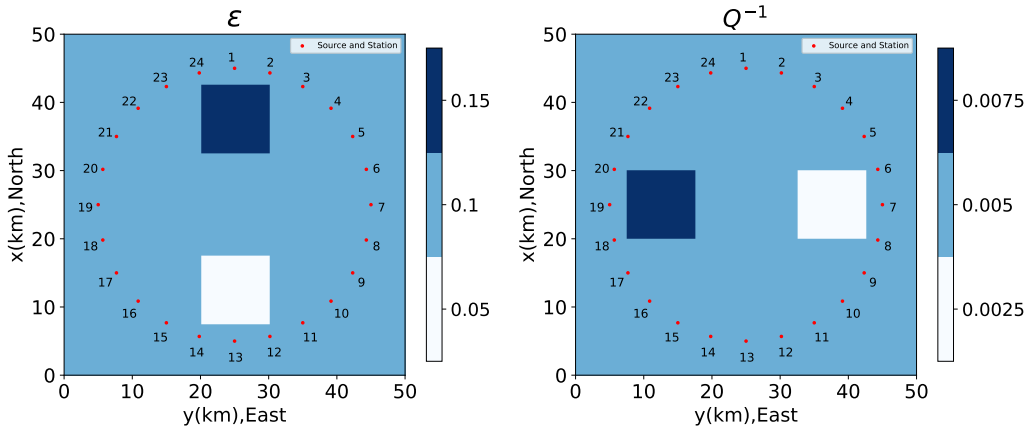



**Fig. 5.9:** The illustrations of the scattering and intrinsic attenuation anomaly model in the simultaneous inversion. The background of $\varepsilon$ and $Q^{-1}$ is 0.1 and 0.005, respectively. The white and the dark blue are two anomaly areas, with the value of 0.05 and 0.15 in the scattering model and the value of 0.0075 and 0.0025 in the intrinsic attenuation model, respectively.

the step length independently using two trial step lengths for each parameter to determine the best step length. Finally, four test simulations are required for each iteration.

As shown in Figure 5.10 the normalized misfit decreases to 24.81% after 8 iterations which is lower than in the separated inversions. Although the absolute value of the misfit in each numerical test is different, the faster decrease illustrates more significant changes in the envelopes due to the recovery of the anomalies. The inversion results are shown in Figure 5.11. Figure 5.11a shows the final inverted models after 8 iterations. The anomaly areas in the true model of $\varepsilon$ and $Q^{-1}$ are located in the north-south and east-west directions, respectively. The model of $\varepsilon$ is well recovered – similar to the result in section 5.6.1. Profiles along the four lines shown in Figure 5.11b and Figure 5.11d illustrate that the inverted model agrees well with the true $\varepsilon$ model.

However, in the inverted model of $Q^{-1}$ one can observe four anomaly areas. Two of them are located in the east and west which agrees with the locations in the true model. The profiles across these two areas in Figure 5.11c show that the amplitudes of these anomalies are underestimated. Normally the underestimation of the amplitudes is reflected by the misfit and improved by more iterations. Yet, in the joint inversion this process seems to be impaired by the presence of the other two anomaly areas in the north and south. Since these areas with anomalies in $\varepsilon$ introduce changes in the envelopes the misfit can be reduced also by adapting the $Q^{-1}$. Profiles through the $Q^{-1}$ model across the locations of the $\varepsilon$ anomalies in Figure 5.11e show these misplaced anomalies. The trade-off between improvements of the $Q^{-1}$ and $\varepsilon$ models will be discussed in the next section in detail.

The envelope fit and the adjoint sources in the simultaneous inversion are shown in Figure 5.12. Both the redistribution of energy due to the changes in heterogeneity and the energy loss caused by intrinsic attenuation anomaly contribute the changes in the envelopes. Disentangling these influences is the challenge of the joint inversion which complicates the model updates.

Let us first consider the source-receiver combinations that are affected strongest by the anomalies in heterogeneity because of the ballistic path traversing through the $\varepsilon$ anomalies in Figures 5.12a, 5.12b and 5.12c. The convergence of these envelopes is similar to the separate inversion for $\varepsilon$ in section 5.6.1.

As expected from the inverted model the evolution of envelopes affected by the attenuation anomalies is different. The source-receiver combination with ballistic path across the stronger intrinsic attenuation anomaly (Figure 5.12d) shows that the adjoint source improves in the first few iterations but then grows

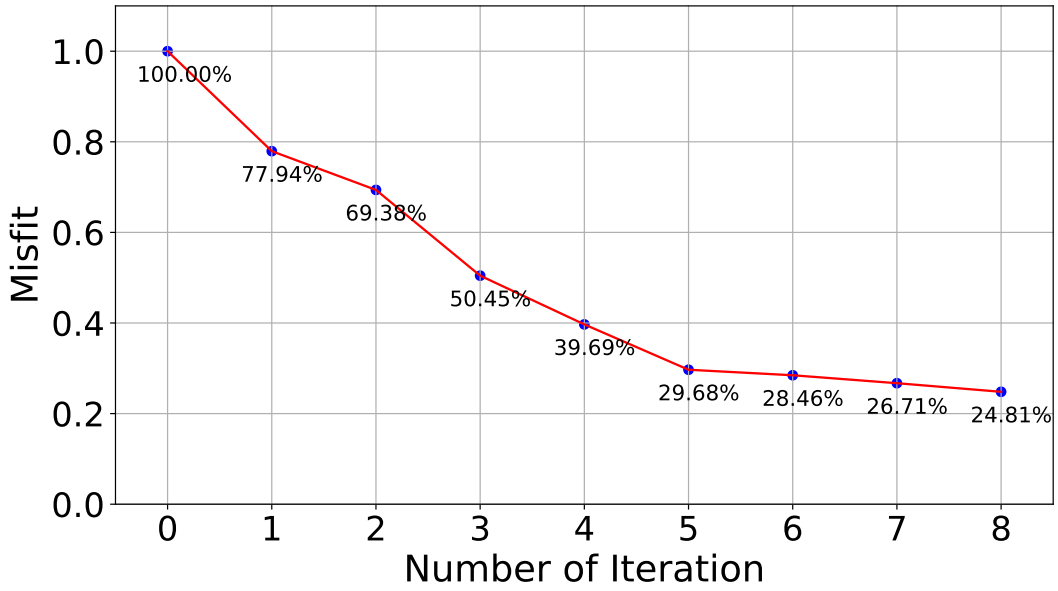


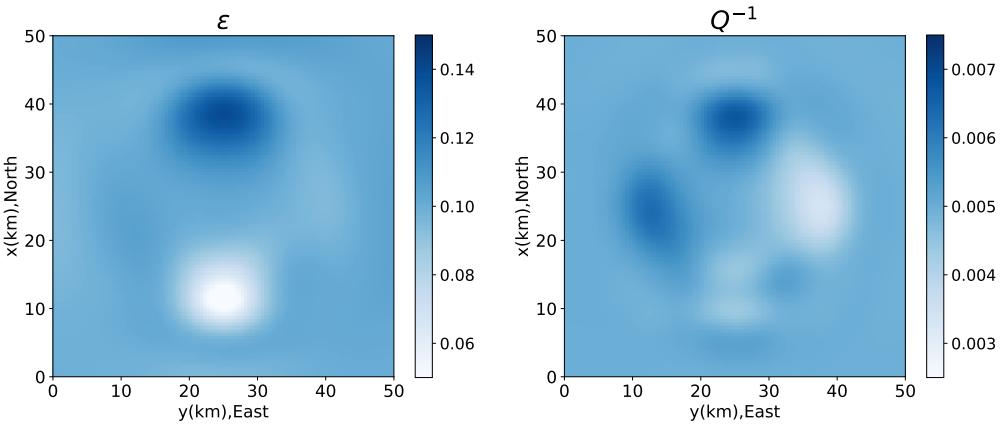

**Fig. 5.10:** The misfit in the simultaneous inversion experiment alters with the number of iterations. The values are normalized by the misfit between the uniform initial model and the true model shown in Figure 5.9.

again (with opposite sign) in the later iterations. In contrast the envelope across the weak attenuation area in Figure 5.12f nicely converges and the adjoint source remains minimal in the last iterations. In general the amplitudes of the adjoint sources of source-receiver combinations with ballistic path through the $Q^{-1}$ anomalies in the east and west are smaller than those affected by the $\varepsilon$ anomalies.

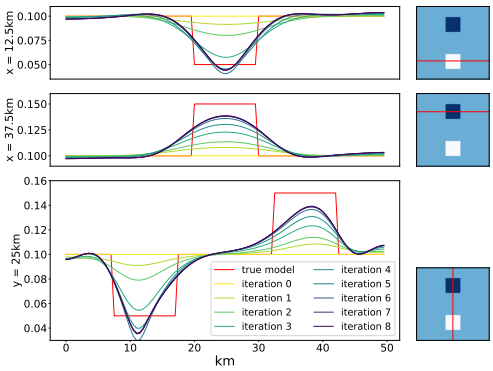
## 5.7 Discussion

To demonstrate that the mathematical framework of adjoint envelope tomography with the radiative transfer equation which we have presented in sections 5.2∼5.5 is able to recover material properties, we performed different synthetic tests. We want to remark first, that we used the radiative transfer MC-algorithm for the inversion and for the simulation of the synthetic data. We chose this strategy over the more realistic scenario of generating the synthetic data with wave field simulations to avoid the additional fluctuations that would result from the specific realization of the medium heterogeneity that has to be generated for the wave field simulation. This effect will be investigated in the future, but in this first application we want to focus on the methodological aspects of the adjoint tomography. For the same reason we restrict ourselves to the acoustic case in two dimensions.

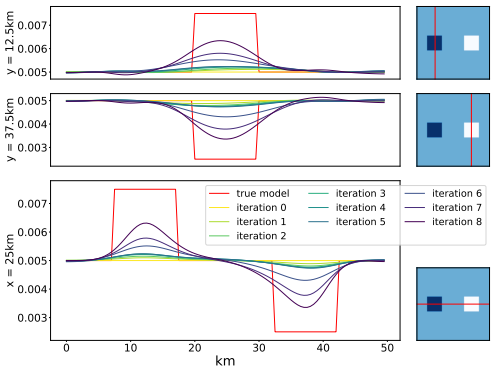
In the three numerical experiments we assume that first order estimates of the large scale average material properties are available form separate investigations, a prerequisite shared with FWI. In the first two experiments we show that the adjoint tomography successfully recovers the heterogeneity and absorption structures if the other material property is known and fixed. Significant improvements of the model are achieved following the first iteration. These improvements cannot be achieved with existing single step inversion approaches where sensitivity kernels are calculated in a uniform model. In both experiments with the separate inversion for either heterogeneity or absorption the misfit converges after a few iterations and the remaining misfit can be attributed to the enforced smoothness of the retrieved model that prohibits further improvements. Successively lowering the smoothness constraint imposed by the filtering of the gradient could be a strategy to reduce the misfit further. Such a strategy would be similar to successively raising the high frequency limit in FWI.



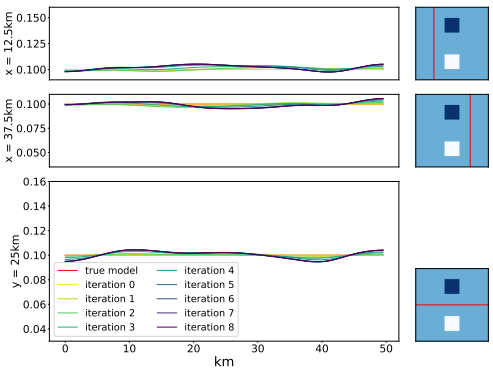

**(a)** the final inverted models



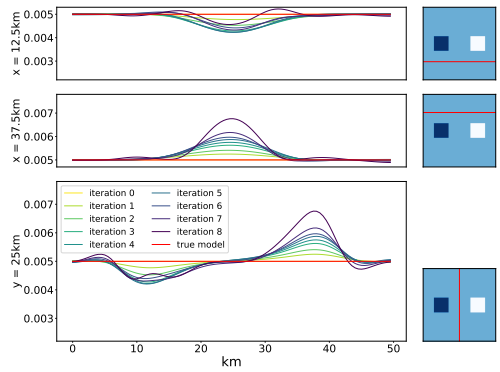

**(b)** the profiles of $\varepsilon$ across where the $\varepsilon$ anomalies located




**(c)** the profiles of $Q^{-1}$ across where the $Q^{-1}$ anomalies located

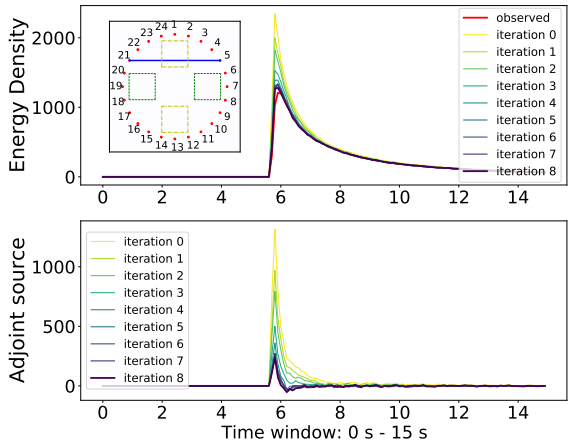



**(d)** the profiles of $\varepsilon$ across where the $Q^{-1}$ anomalies located

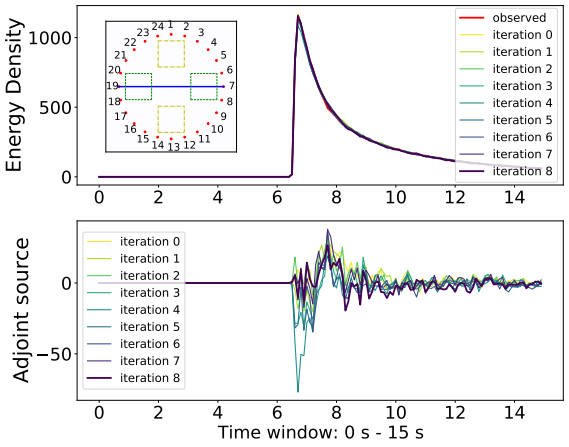



**(e)** the profiles of $Q^{-1}$ across where the $\varepsilon$ anomalies located

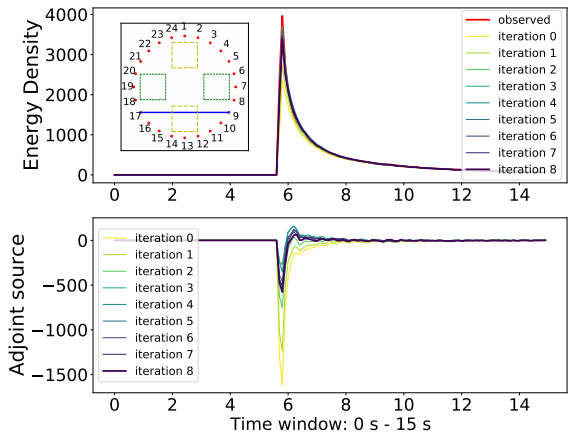
**Fig. 5.11:** The simultaneous inversion results: (a) the final model of $\varepsilon$ and $Q^{-1}$ after 8 iterations; (b)-(d) the profiles across the anomalous areas (illustrated in the right column) of the updated model in each iteration from the initial model (light yellow) to the final model (dark purple) compared with the true model (red).



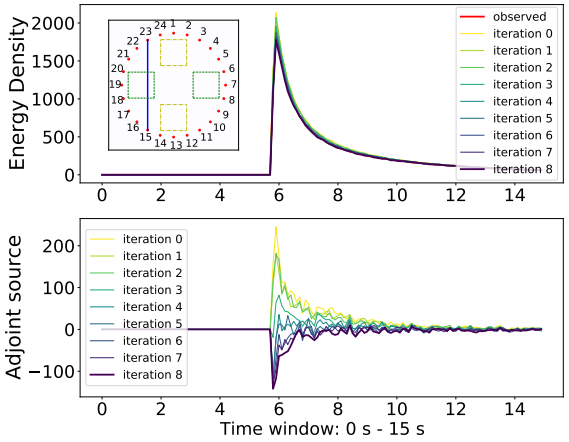

**(a)** Source 5 and station 21



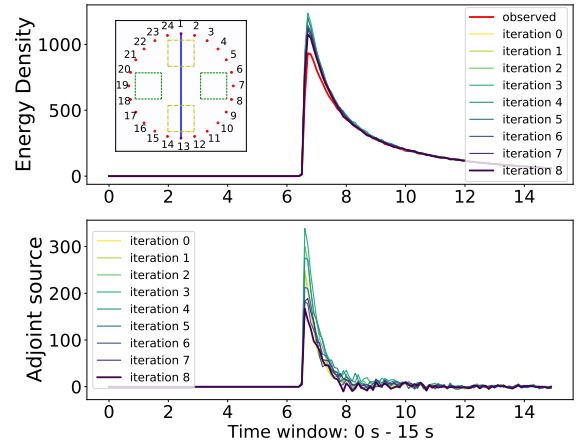

**(b)** Source 7 and station 19



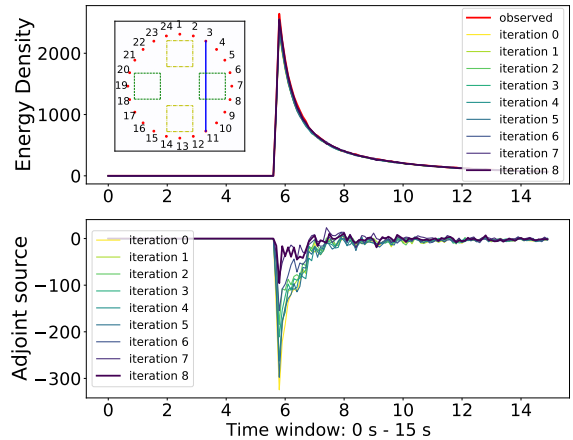

**(c)** Source 9 and station 17




**(d)** Source 15 and station 23




**(e)** Source 1 and station 13




**(f)** Source 3 and station 11

**Fig. 5.12:** The comparison of observed (red) and synthetic envelopes in each iteration from the initial (light yellow) to the final (dark purple) as well as the corresponding adjoint sources in the simultaneous inversion.

The situation is different in the third experiment in which we aim to invert the heterogeneity and attenuation structures simultaneously. Figure 5.11a gives the impression that the intrinsic attenuation is more difficult to constrain than heterogeneity. We will show in the following that this difficulty is due to the intrinsic trade-off between $\varepsilon$ and $Q^{-1}$ in the inversion process and depends on the specific parameter distribution that is to be imaged.

### 5.7.1 Trade-off

The fundamental problem we meet in the simultaneous inversion for heterogeneity and attenuation comes from the trade-off between changes in $\varepsilon$ and $Q^{-1}$. Eq. (5.28) shows that the scattering kernel is composed of two terms: (A) the *active kernel* that describes the energy gained by scattering from another directions; and (B) the *passive kernel* that describes the loss of energy due to scattering in other directions (Margerin et al., 2016; Zhang et al., 2021). The absorption kernel (Eq. (5.29)) has the shape of the passive kernel. Separating the effects of $\varepsilon$ and $Q^{-1}$ in the inversion is complicated by the passive kernel that is part of both kernels.

In order to quantify the influence of $\varepsilon$ and $Q^{-1}$ on the misfit, we design a test to calculate the misfit caused by the changes of $\varepsilon^2$ or $Q^{-1}$. Note that the expression (5.28) is a function of $\varepsilon^2$ while we use $\varepsilon$ for illustration. The model is shown in Figure 5.13. It contains only one anomaly area that is used for $\varepsilon$ and $Q^{-1}$ separately, while the other parameter is kept uniform. Background values are as before ($\varepsilon = 0.1$ and $Q^{-1} = 0.005$). We compare the adjoint sources caused by the two types of changes in the anomaly area ($\Delta\varepsilon^2 = 0.02$ or $\Delta Q^{-1} = 0.02$). Two source-receiver combinations are shown for which the ballistic wave crosses the anomaly or travels far away from the anomaly area as shown in Figure 5.14a and 5.14b, respectively. Although the difference between the two perturbations (the blue and red curves) is clearly observed in the later coda waves, it is impossible to estimate the character of the anomaly from the ballistic wave. In the case of distant sensors in Figure 5.14b the wavefield obviously does not encounter the anomaly prior to approximately 9 s making it impossible to learn anything about the anomaly. But even in the case of ballistic propagation through the anomaly (Figure 5.14a) the effect of increased scattering and increased attenuation on the ballistic wave and early coda is approximately the same as both perturbations decrease the coherent energy.

The time window chosen in the simultaneous inversion test includes the whole period from 0 to 15 s so the misfit is dominated by the large amplitude ballistic waves. Since the scattering kernel for the ballistic wave is dominated by the passive kernel and attenuation is only influenced by the passive kernel variation of both scattering and attenuation can each explain misfit of the ballistic wave resulting in a trade-off. This can only be resolved if sufficient weight is given to the late coda in which the active kernel contributes to the scattering kernel.

Figure 5.14c shows the misfit for varying amplitudes of the anomalies. It allows to estimate the values of $\varepsilon$ and $Q^{-1}$ at which both parameters have comparable influence on the misfit. The relative influence on the misfit can be illustrated in the $\varepsilon - Q^{-1}$ domain as shown in Figure 5.15 where the red line represents the relationship $Misfit(|\Delta\varepsilon^2|) = Misfit(|\Delta Q^{-1}|)$ which splits the domain into two parts. The green and pink areas represent combinations of $\varepsilon - Q^{-1}$ in which the misfit is dominated by the scattering and intrinsic attenuation, respectively. A similar discussion of this trade-off on the fit to the envelopes was conducted by Cormier & Sanborn (2019). Different from the adjoint tomography focused on the misfit function, their work compared the influence of the two parameters on the measurements with MLTWA which is affected by the choice of time windows. We will discuss the effects of the time window in our approach in section 5.7.3.

The values of $|\Delta\varepsilon^2|$ and $|\Delta Q^{-1}|$ that we used in the numerical experiment for the simultaneous inversion in section 5.6.3 are illustrated in Figure 5.15 by the blue dot and the red triangle for the

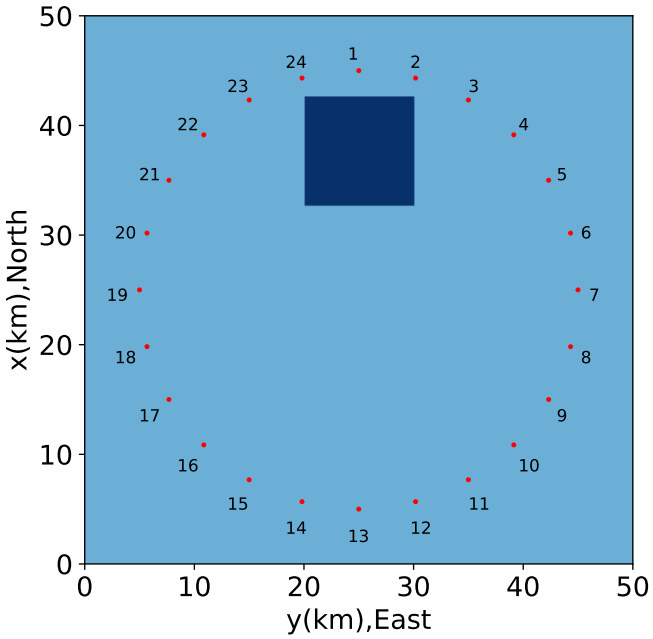



**Fig. 5.13:** Model used for trade-off test. The anomaly area is separately perturbed by changes in $\varepsilon$ and $Q^{-1}$. The other simulation parameters are the same as before.

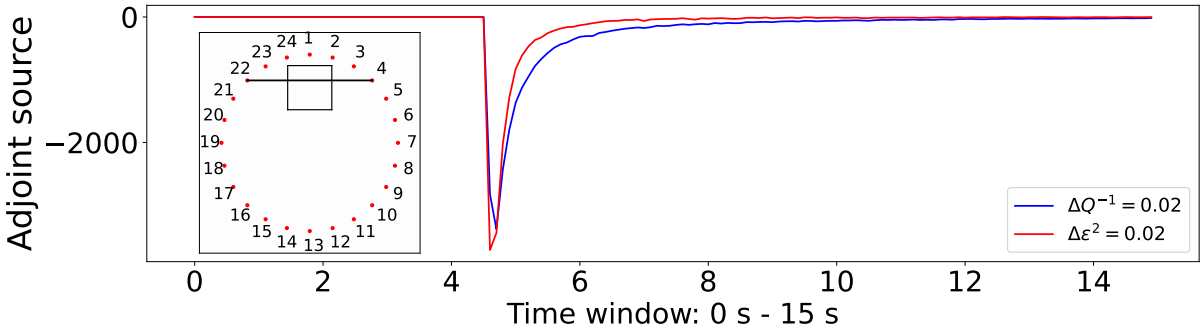
positive and negative anomalies, respectively. It is clear that the misfit in this situation is dominated by the anomalies in $\varepsilon$. To illustrate how this dominance affects the inversion Figure 5.16a shows the kernels in the first iteration in this simultaneous inversion experiment. While the $\varepsilon$ kernel correctly highlights the areas of the $\varepsilon$ anomalies in the north and south, the $Q^{-1}$ kernels also highlights the $\varepsilon$ anomalies supplemented by moderate values at the locations of the $Q^{-1}$ anomalies in the east and west. As a consequence the $Q^{-1}$ update starts off in the wrong direction. A similar effect but to a much smaller extend can be observed for the $\varepsilon$ kernel that shows a slight increase at the locations of the $Q^{-1}$ anomalies in the east and west. It is obvious that the misfit is dominated by the heterogeneity anomalies which results in a bias of the absorption kernels and leads to the inverted model in Figure 5.11a.

On the contrary we can try to recover a model in which the anomalies of intrinsic attenuation dominate the inversion as shown in left column of Figure 5.16b. The values of the $|\Delta\varepsilon^2|$ anomaly are 0.0019 and 0.0021 while that for $|\Delta Q^{-1}|$ is 0.005 as indicated by the green square and orange diamond in Figure 5.15. As expected the resulting kernels of the first iteration shown in Figure 5.16b reflect the dominance of the $Q^{-1}$ anomalies. The anomaly areas of $Q^{-1}$ are correctly recognized while the scattering kernel is biased by the anomalies in absorption.

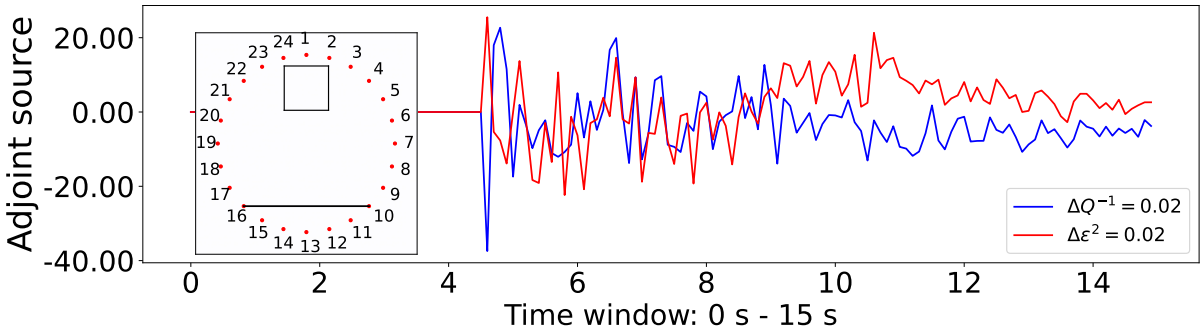
### 5.7.2 Influence of the Initial model

In the two individual inversions and the simultaneous inversion above, we assumed that the initial models are all uniform meaning that there is no prior information about the location, shape or amplitude of the anomaly. In the separate inversions for scattering and intrinsic attenuation the respectively other parameter distribution is uniform, equals the true model and is not updated, leading to a good recovery of each parameter as shown in Figure 5.3a and Figure 5.7a. If updates of both parameters are required, the joint calculation may project the data misfit into updates of both $\varepsilon$ and $Q^{-1}$ and it is instructive to investigate how the knowledge about one parameter influences the recovery of the other.

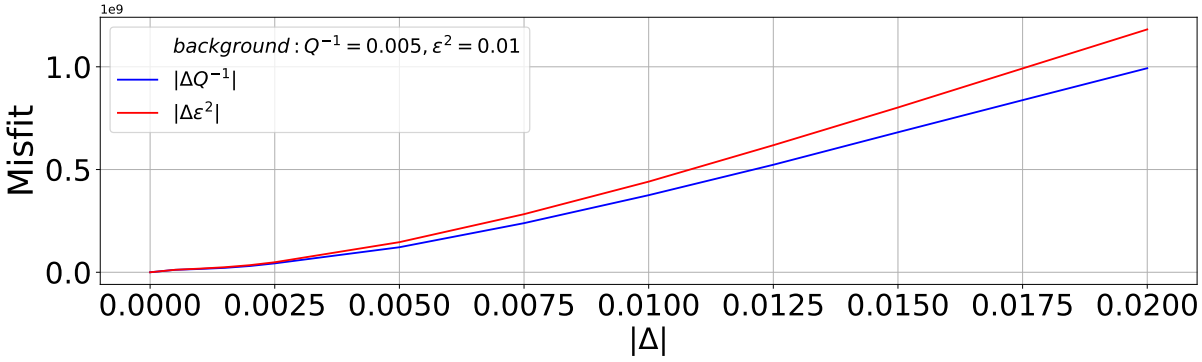
We assume that we have additional prior information about the $\varepsilon$ structure such that we can start




**(a)** source: 4, station: 22




**(b)** source: 10, station: 16




**(c)** misfits caused by varing anomalies

**Fig. 5.14:** Adjoint sources caused by two types of the changes in the anomaly area ($\Delta\varepsilon^2 = 0.02$ or $\Delta Q^{-1} = 0.02$): (a) and (b) show two source-receiver configurations for which the ballistic wave traverses the anomaly or travels far away from the anomaly area, respectively. (c) Misfits caused by varying amplitudes of the $\Delta\varepsilon^2$ and $\Delta Q^{-1}$ anomalies.

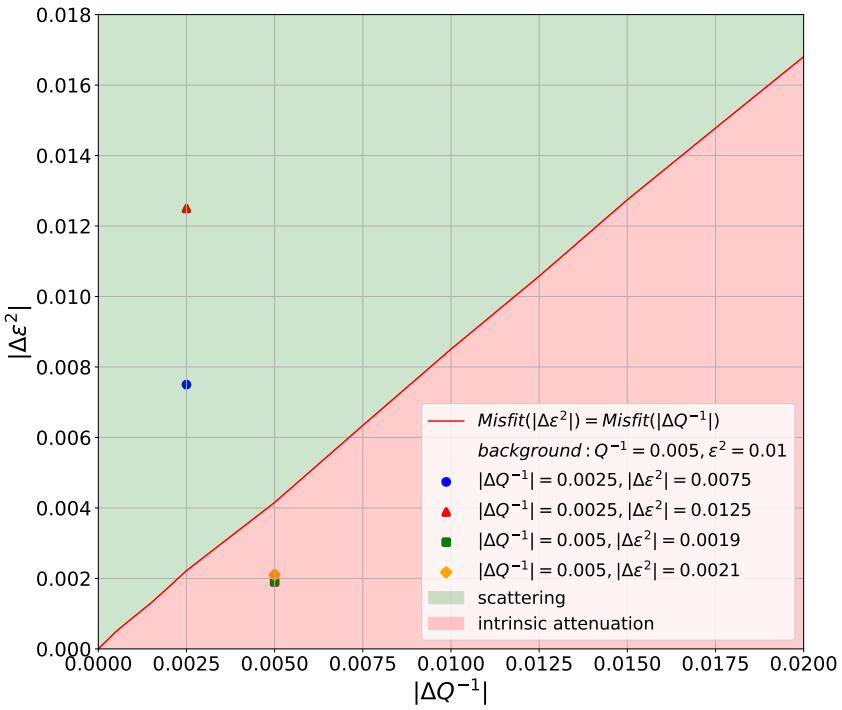



**Fig. 5.15:** The influence on the misfit caused by $|\Delta\varepsilon^2|$ and $|\Delta Q^{-1}|$. The red curve $Misfit(|\Delta\varepsilon^2|) = Misfit(|\Delta Q^{-1}|)$ is calculated by Figure 5.14 under the assumption that the relationship between the misfit and changes is piecewise linear.

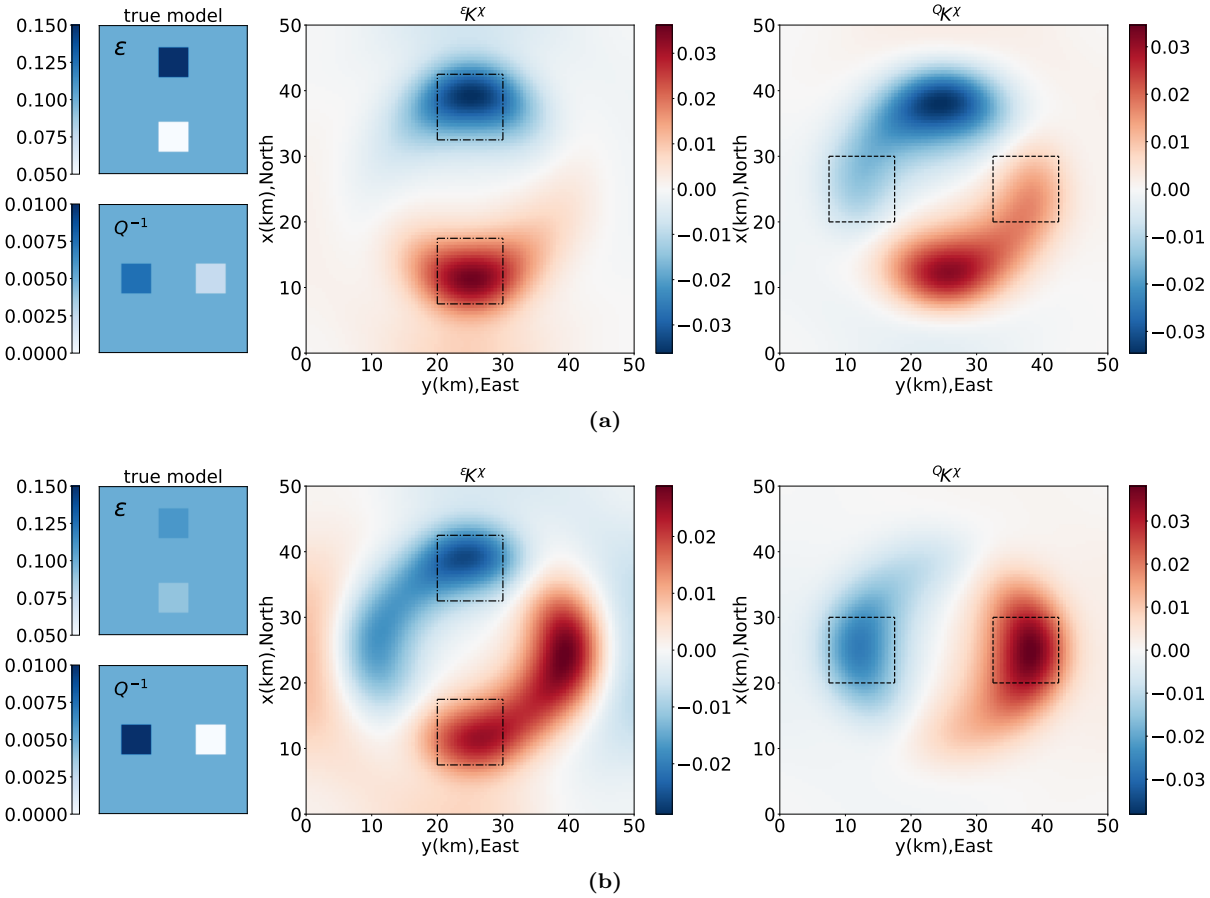



**Fig. 5.16:** The kernels of the first iteration in two simultaneous inversion cases: (a) the misfit is dominated by $\varepsilon$ which is the case conducted in section 5.6.3, (b) the misfit is dominated by $Q^{-1}$.

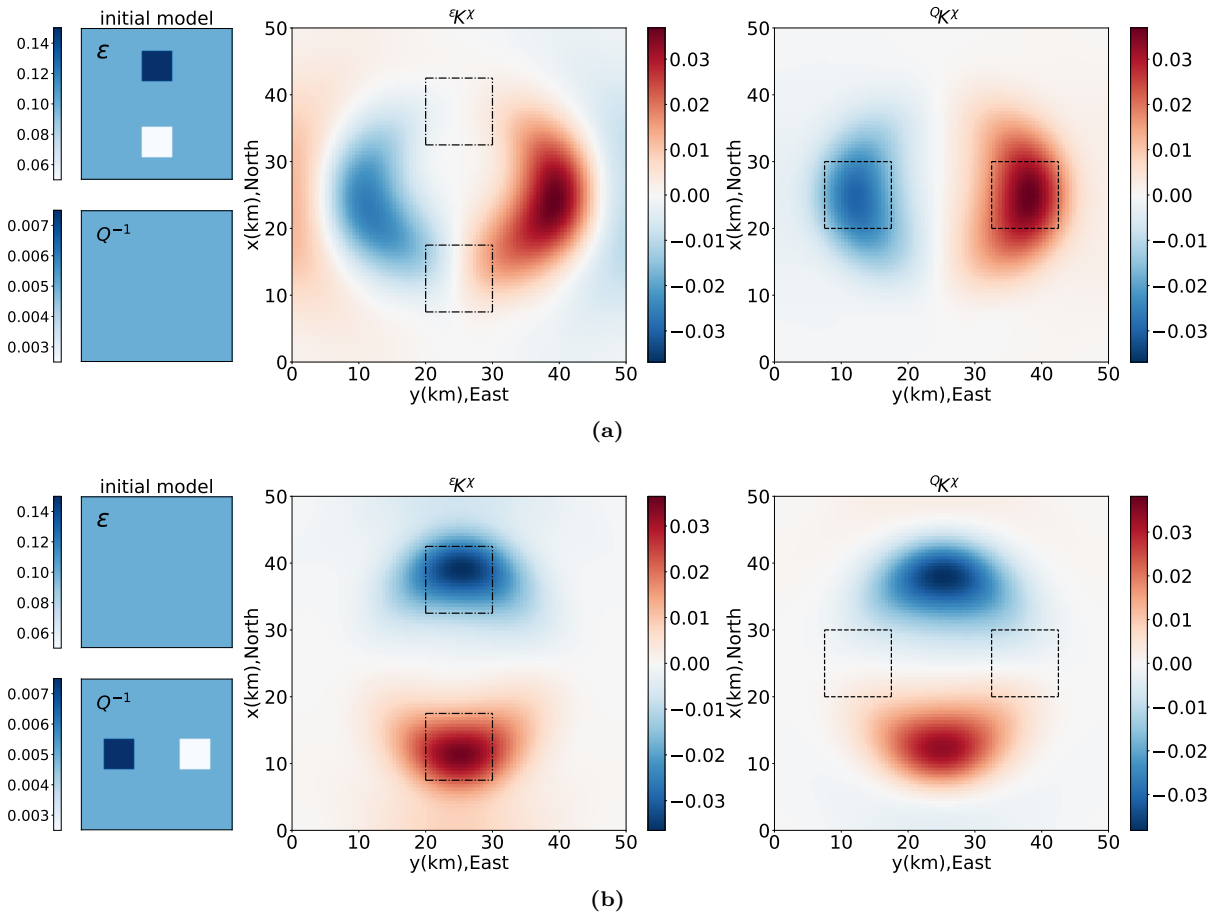



**Fig. 5.17:** The kernels of the first iteration in the simultaneous inversion conducted in section 5.6.3 with different initial models: (a) $\varepsilon$ is the same as the true model while $Q$ is homogeneous; (b) $Q^{-1}$ is the same as the true model while $\varepsilon$ is homogeneous.

the inversion with a non-uniform but correct $\varepsilon$ model. The kernels in the first iteration under this condition are shown in Figure 5.17a. Now the intrinsic attenuation kernel highlights the correct anomaly areas because there is no misfit from a wrong $\varepsilon$ model. However, the $Q^{-1}$ anomalies also show up in the $\varepsilon$ model, meaning that the correct $\varepsilon$ model would be altered in a simultaneous inversion for both parameters. The same applies when the correct $Q^{-1}$ model but a uniform $\varepsilon$ model is used to obtain the kernels as shown in Figure 5.17b. The $\varepsilon$ kernel highlights the correct anomalies in the north and south while the $Q^{-1}$ kernel would alter the correct model of absorption in a simultaneous inversion for both parameters.

These tests show that the kernels of absorption and heterogeneity are both correctly recovered also if the other parameter has a complicated spatial structure that is well described already. However, if the structure of one or both parameters is not known and both parameters are inverted for, the contributions to the data misfit of both parameters show up in both kernels to a variable degree.

### 5.7.3 Time Window for Full Envelope Inversion

In the experiments described above we use a time window that contains the whole envelope ranging from the source time until the late coda at more than twice the travel time of the ballistic wave. This choice corresponds to setting $T$ in eq. (5.28) and (5.29) to the end of the time window. Although the full information of wave propagation through the medium is utilized in this time window, we show in the following how absorption and heterogeneity have different imprints on the envelopes at different lapse

times.

Scattering reduces the energy of the ballistic wave and increases the coda wave while the intrinsic attenuation absorbs the energy during the whole lapse time including both the ballistic wave and coda waves. Since the effect of absorption and heterogeneity is the same for the ballistic wave one can anticipate that the coda is crucial to distinguish between both effects which has been illustrated in Figures 5.14a and 5.14b. In this section we use the models discussed in section 5.7.1 with anomalies in both heterogeneity and absorption and uniform initial models to study the effects of the ballistic wave and the coda wave by separating them in different time windows.

Figure 5.18 and Figure 5.19 show results under the condition that the misfit is dominated by heterogeneity and absorption, respectively. Figure 5.18a and Figure 5.18b show the misfit kernels in the first iteration of the inversion with time windows restricted to the ballistic waves and coda waves, respectively. The results for the ballistic time window shown in Figure 5.18a are similar to the ones in Figure 5.16a with the full lapse time range and show the strong bias of the absorption kernel from the anomalies in heterogeneity. This is reasonable since the energy peak of the ballistic wave dominates the inversion when the full envelope is inverted and the effect on the ballistic wave energy is the same for scattering and intrinsic attenuation. A significant improvement of the intrinsic attenuation kernel is achieved by restricting the time window to the coda wave as shown in Figure 5.18b. The effect of the heterogeneity anomaly on the absorption kernel is drastically reduced and reversed in polarity compared to Figure 5.18a. This reflects the different effects of anomalies in heterogeneity and absorption on coda waves and indicates a possibility to steer the inversion by a lapse time dependent weighting of the misfit in the calculation of the misfit kernels.

Figures 5.19a and 5.19b show the absorption dominated misfit kernels restricted to ballistic and coda wave time window, respectively. Here the absorption kernel is rather independent of the time window used due to its dominant effect. The scattering kernel, however, is very different for the two time windows. The kernels derived from both the coda and ballistic time windows are biased by the absorption anomalies, but the bias has opposite polarity for the ballistic and coda windows while the polarity at the location of the scattering anomalies is the same. An appropriate weighting of both time windows can thus improve the inversion also in this situation. In conclusion, separating the ballistic wave and coda wave into separate time windows indicates a potential strategy to improve the trade-off.

### 5.7.4 The Effect of Noise

The Monte-Carlo simulation which we use for the forward modelling is a stochastic method that introduces fluctuations in the synthetic data. These random fluctuations can be regarded as the noise of observation. However, the use of a sufficient amount of particles ensures convergence of the simulation suppressing this random noise to a desired degree. For more details refer to Appendix B of Zhang et al. (2021). To study a more realistic situation, we artificially add random noise to the synthetic data in our numerical tests. Gaussian noise is generated with a specific signal-to-noise ratio (SNR) defined as:

$$SNR = 10 \times \log_{10} \frac{E_{MS}}{N_{MS}} \tag{5.35}$$

where $MS$ indicates the mean square of the noise free signal envelope $E$ and the noise amplitude $N$. The addition of Gaussian noise can lead to negative values of the envelope which seems unrealistic. However, it is common practice in the analysis of scattered waves to estimate the noise energy from time windows prior to the arrival of direct waves and subtract this noise energy from the signal. This process does also lead to partially negative values in observed energy time series, consistent with our approach.

Figure 5.20 shows examples of the kernels in the first iteration for different SNR. The left column

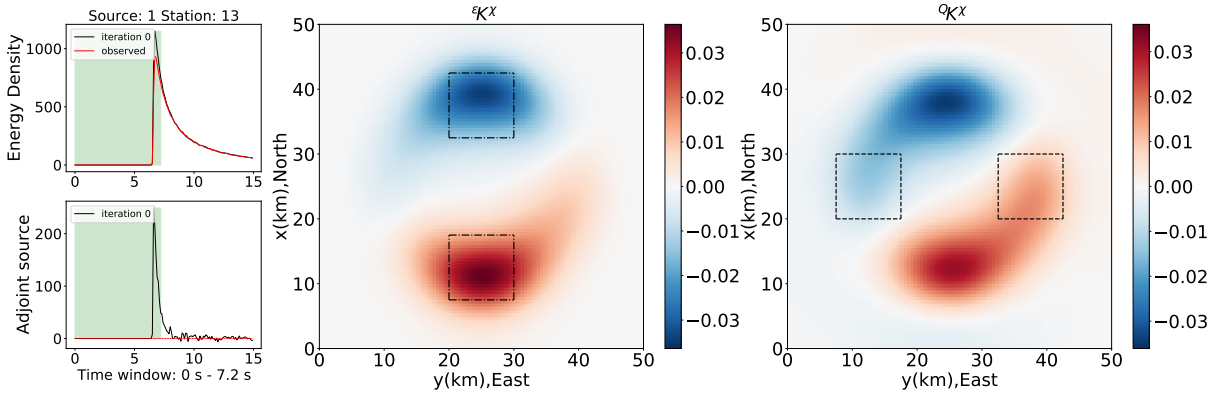
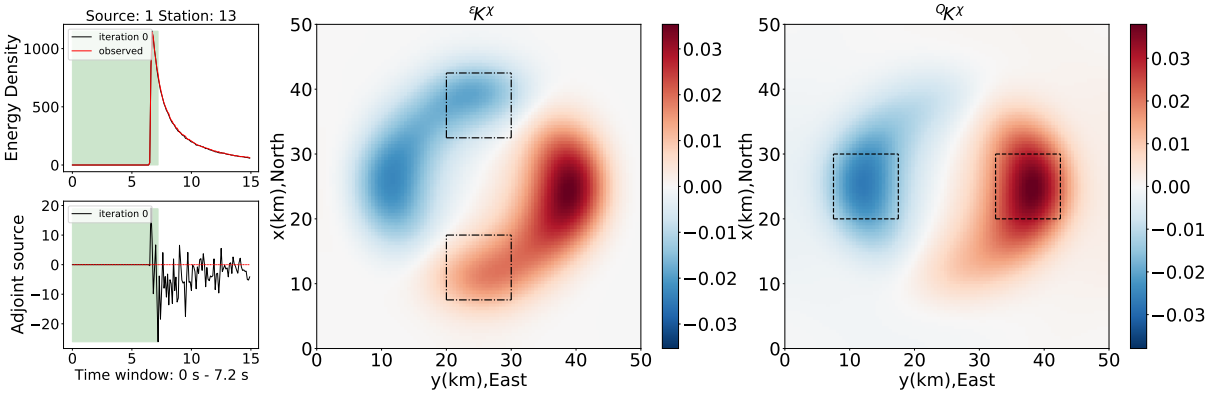
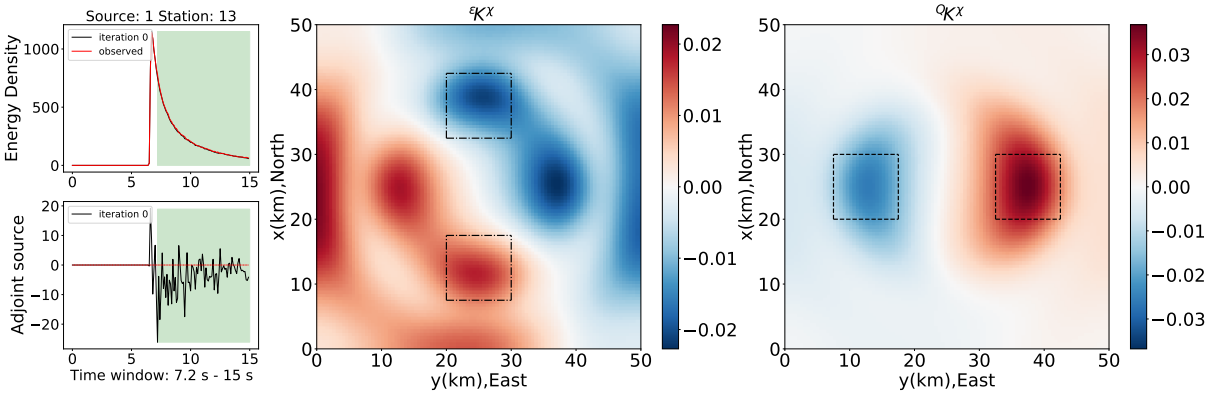



**Fig. 5.18:** The kernels of the first iteration in the simultaneous inversion whose misfit is dominated by $\varepsilon$ with different time windows: (a) ballistic wave; (b) coda wave. The results with the time window containing both ballistic wave and coda wave is shown in Figures 5.16a.




(a)




(b)

**Fig. 5.19:** The kernels of the first iteration in the simultaneous inversion whose misfit is dominated by $Q^{-1}$ with different time windows: (a) ballistic wave; (b) coda wave. The results with the time window containing both ballistic wave and coda wave is shown in Figures 5.16b.

illustrates the energy density and the adjoint source of one source-receiver couple. The black and red curves indicate the original data and that with added noise, respectively. From top to bottom the SNR is 20, 10 and 1, respectively. At the lowest level of noise (SNR = 20) shown in Figure 5.20a, the noise has rather little influence on the energy density. Also the adjoint source is still clearly visible and there is almost no difference between Figure 5.20a and Figure 5.16a that shows the result without any noise. With increasing noise (SNR = 10), the coda part of the adjoint source is drowned in the noise but the ballistic wave can be still recognized. Resulting changes in the scattering kernel shown in Figure 5.20b are still minor but the intrinsic attenuation kernel starts to degrade. At the high level of noise (SNR = 1), the additional fluctuations are comparable to the signal level. Although it is almost impossible to recognize the adjoint source in the data of a single source-station combination the scattering kernel shown in Figure 5.20c still reflects the anomaly areas in $\varepsilon$ to some degree. Apparently information of the numerous records that are combined in the kernel still facilitates recovery of a somewhat degraded image even under the influence of strong noise.

### 5.7.5 Conditioning of the Misfit Kernels

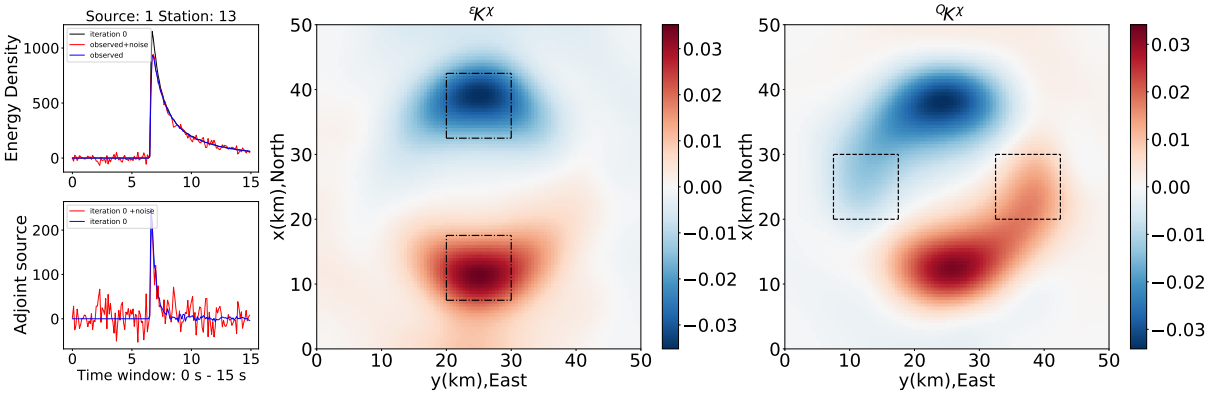
The regularization in our inversion has two parts, which are (A) removing the singularities at the sources/stations and (B) smoothing the results.

The source/receiver singularities are an intrinsic problem of the adjoint method. Their effect is significant and it is almost impossible to recover the anomaly without removing them. We multiply the kernels with a 2D Gaussian notch centered on each station and truncated at 2km distance. The kernels before and after removing the singularities are shown in Figure 5.21 bottom left and upper middle, respectively. Actually the influence of this step on kernels is mainly determined by the setup of sources and stations, which can be reduced with high enough station coverage.

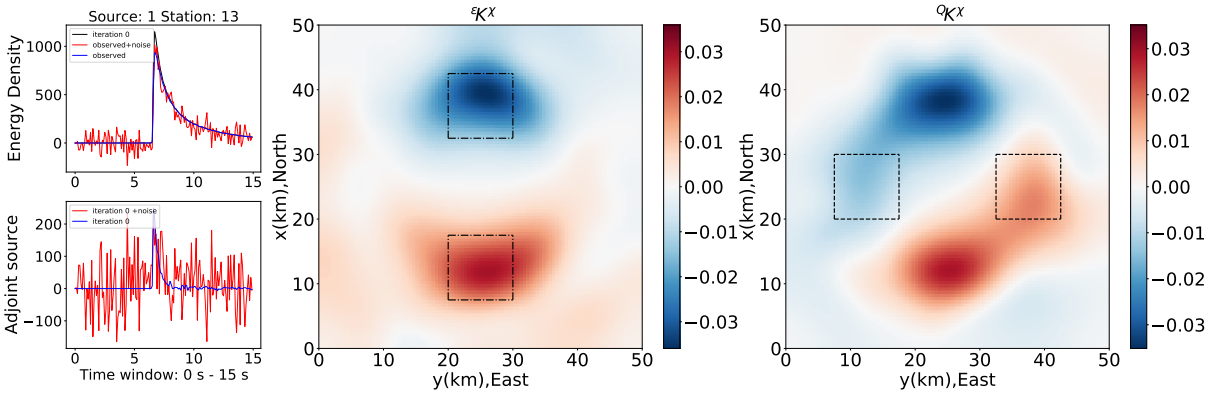
Moreover, the kernels are smoothed by a 2-D Gaussian filter with the same x- and y-axis standard deviations $\sigma_x = \sigma_y = \sigma$. The length of $\sigma$ is normally determined by the wavelength and the structure in FWI and decreases with the model improving (Zhu et al., 2015). Our approach is based on RTE that is beyond the limitation of the wavelength. But the model is discretized in a 100 by 100 cells grid with the size of 0.5 km $\times$ 0.5 km. The Figures 5.21 shows the Gaussian filters with different $\sigma$ (the upper left) and the corresponding kernels after smoothing (the right two columns). The one with $\sigma = 0$ km is the original kernel after removing the singularities at the sources and receivers. In the beginning, we set the $\sigma$ as 3 km and after more iterations reduce it to 2 km. The $\sigma$ less than 2 km has never been used which can result in the inversion becoming unstable. In order to demonstrate the effects from smoothing with different $\sigma$ on the adjoint sources, five simulations are conducted in these five models including the initial model and four updated models based on four different smoothed kernels in Figures 5.21 using the steepest descend. The resulting envelopes are shown in Figure 5.22. The comparison shows, that the differences originating from the different smoothing of the kernels are insignificant compared to the change of the envelopes resulting from the model update. This indicates that the precise choice of the smoothing filter has minor influence on the result and our choice of the filter size of the order of the station distance is acceptable.

## 5.8 Conclusion

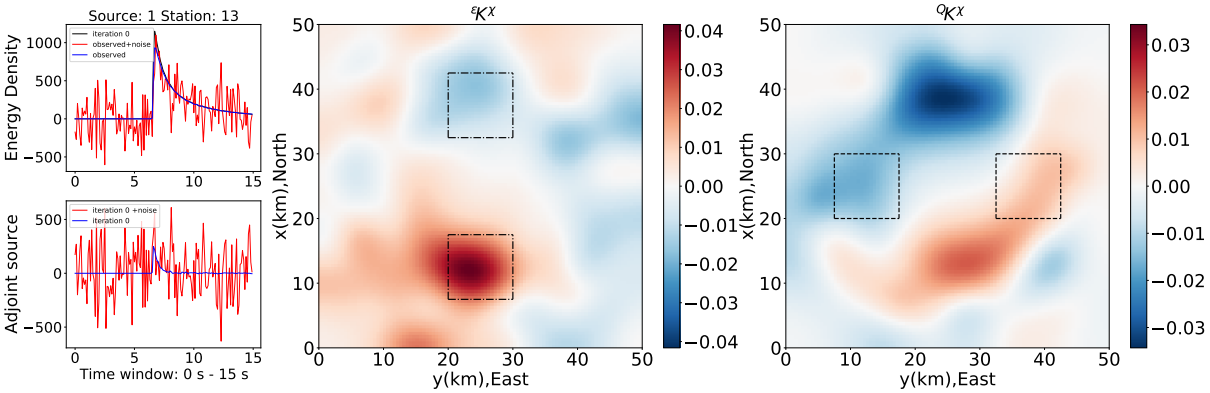
We have presented a new physically rigorous tomography method for heterogeneity and attenuation. The iterative tomography combines forward modelling of the radiative transfer equations with simulations of the adjoint equations to obtain the Fréchet derivatives that guide the iterative model improvements. This full envelope inversion is thus the radiative transfer counterpart to full waveform tomography but




**(a)** SNR=20




**(b)** SNR=10




**(c)** SNR=1

**Fig. 5.20:** The kernels of the first iteration in the simultaneous inversion conducted in section 5.6.3 with different added noise: (a) SNR =20; (b) SNR =10; (c) SNR =1.

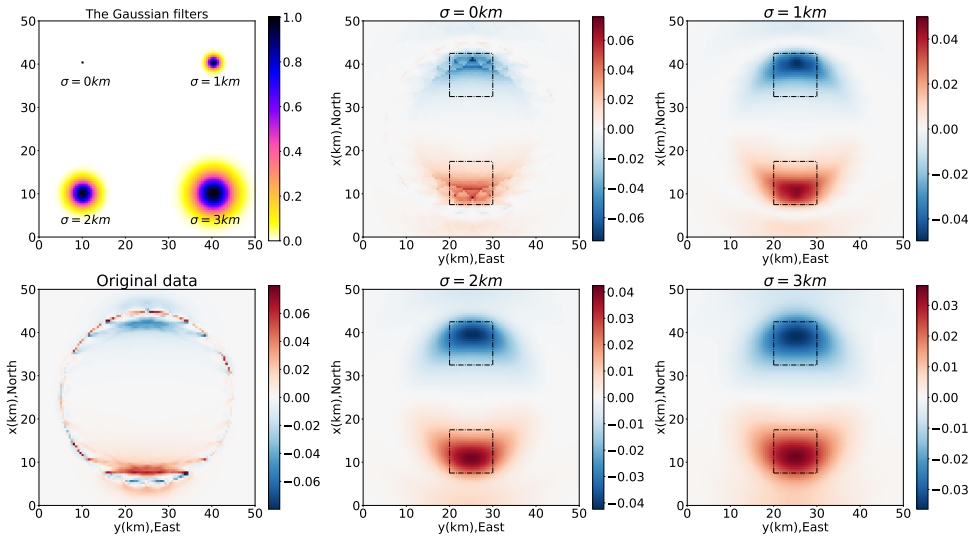
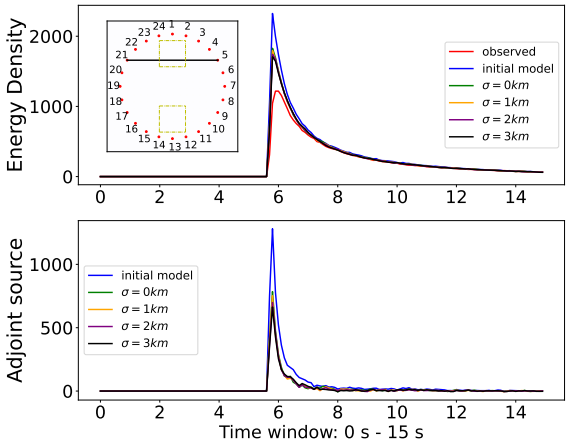
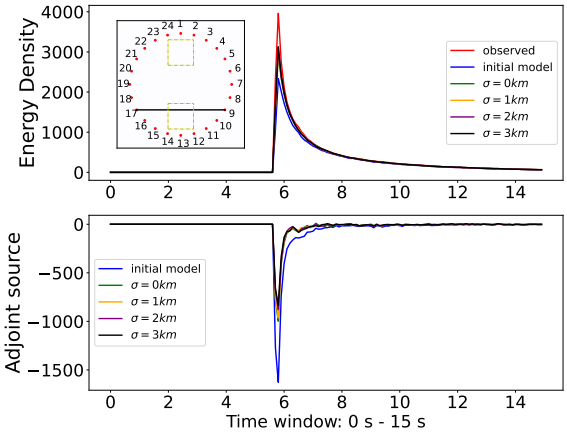


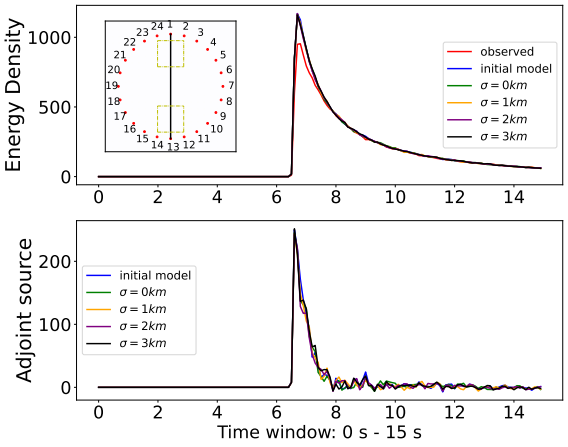

**Fig. 5.21:** Gaussian smoothing filters with different $\sigma$ (the upper left), the original date without removing the singularities (the bottom left) and the corresponding kernels after removing the singularities and smoothing (the right two columns).



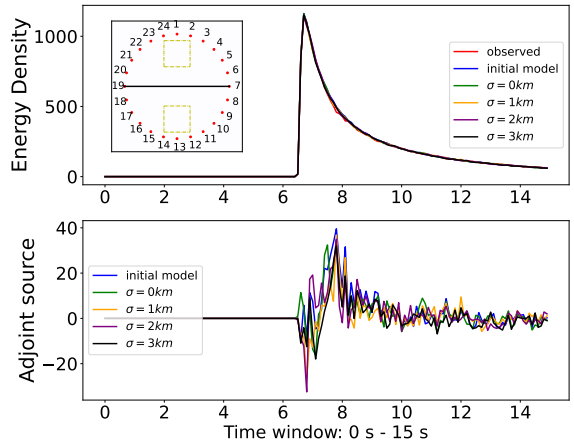

**(a)** Source 5 and station 21




**(b)** Source 9 and station 17




**(c)** Source 1 and station 13




**(d)** Source 7 and station 19

**Fig. 5.22:** Comparison of observed (red) and synthetic envelopes with the corresponding adjoint sources in the initial model (blue) and four updated models. The four updated models are obtained with the steepest descend method using the four differently smoothed kernels shown in the two rightmost columns of Figure 5.21. Source-receiver configurations are shown in the insets.

allows us to obtain information about the structure of the propagation medium on length scales below the resolution limits of waveform tomography. We have demonstrated the success of this approach in a synthetic test of 2D acoustic scattering.

The small-scale structure is statistically described by the spatially variable scattering and intrinsic attenuation properties of the medium described by the fluctuation strength $\varepsilon$ and intrinsic quality factor $Q^{-1}$. Seismic energy propagation in this medium is described by the radiative transfer equation with uniform macroscopic velocity and density. The presence of nonuniform scattering and intrinsic attenuation causes spatially variable seismic energy redistribution and loss with complex influence on the observed seismic waveform envelopes. We therefore use the full envelopes of the observed seismic waveforms for the inversion.

As objective function we use the squared difference between observed/synthetic and modeled data. Inspired by FWI we use the adjoint method to derive the Fréchet derivatives of the least-square misfit function. The radiative transfer equation is solved with the Monte-Carlo method in 2D with multiple non-isotropic scattering in an acoustic medium with spatially variable heterogeneity and attenuation. This algorithm is used to generate the synthetic data as well as for the forward and adjoint simulations. The synthetic data could have been generated also with wavefield simulations which would have been a more realistic experiment. However, we decided to use the same approach for generation of the synthetic data that we use in the inversion to focus on the performance of the adjoint formalism that we present here and leave the test against wavefield data for a later stage.

The numerical experiments demonstrate the success of the presented inversion concept. Fréchet derivatives of the two experiments with separate inversions for either $\varepsilon$ or $Q^{-1}$ reliably guide the inversion towards lower values of the misfit function. Models which better reproduce the synthetic envelopes as expressed by lower misfit also converge towards the true model in the model space. We have also shown that inversion is well behaved with respect to the addition of noise.

The iterative inversion correctly recovers the locations and amplitudes of the anomalies in each parameter, if the other parameter is fixed at the correct value. If both parameters are free in the inversion we observe a trade-off that is common to most inverse problems with multiple parameters. In this case the inversion process is dominated by the parameter that has the stronger anomalies, i.e. dominates the misfit function.

Although the trade-off between scattering and absorption cannot be avoided, some strategies exist to improve the simultaneous inversion. At first, correct knowledge of one of the parameters improves the recovery of the other. This means that independent knowledge of one parameter can be incorporated to stabilize the inversion. Secondly, the time window used to construct the adjoint source affects the recovery of the two parameters differently. While the late coda is essential to infer the distribution of absorption, the ballistic wave and early coda is important to locate anomalies in the small-scale heterogeneity. Applying a weighting between the influence of early and late coda in the inversion can help to improve the recovery of both parameters.

Future developments of adjoint envelope tomography for scattering and absorption include the application to real data (a laboratory study on a concrete test specimen is on the way), an extension to 3D and the treatment of elastic wave scattering. We hope that the approach presented here allows for more detailed analysis of heterogeneous media, providing more accurate knowledge of heterogeneity and attenuation structures of targets such as volcanoes, the Earth's crust on a regional scale, and even the deep Earth.

Chapter 6

# Laboratory Experiment of AET

## 6.1 Introduction

For the first test of AET on real data, we selected an environment that is more controlled than a seismological field experiment. We chose a metric-sized concrete specimen with embedded ultrasound transducers and the known internal structure. Imaging man-made material, like concrete, is normally done with ultrasound which is analog to seismic waves. Ultrasound imaging in concrete mainly utilizes primary reflections but is strongly affected by scattering and intrinsic attenuation (Anugonda et al., 2001; Turner & Anugonda, 2001). Consequently, imaging strong reflection anomalies, like tendon ducts in the concrete, is primarily implemented using the ultrasonic pulse-echo method with the synthetic aperture focusing technique (SAFT) (Schickert et al., 2003; Schickert, 2005). However, this approach requires an array of transducers with a specific geometry of the sensors and high impedance contrast of the anomaly, which restricts the ability to locate small-scale or weak contrasts in concrete. Recently more advanced imaging methods, such as Reverse Time Migration (RTM) have been adopted from oil exploration, improving the image quality, but still with significant limitations (Grohmann et al., 2017). The same applies to the monitoring of weak changes in the medium (Planès & Larose, 2013) where scattered coda waves have superior sensitivity as it has been mentioned for seismic waves in the Earth. A localized change in the velocity causing perturbations in the ultrasound wavefield can be detected using coda wave interferometry (Poupinet et al., 1984; Snieder et al., 2002) (CWI) and also located (Pacheco & Snieder, 2005). With several years development, CWI has been widely applied to localize small or weak changes in different media in response to stress (Larose et al., 2006; Niederleithinger & Wunderlich, 2013) or damage induced changes (Schurr et al., 2011; Wang et al., 2020). For a detailed review refer to Planès & Larose (2013). Recently, the results of several successful large scale evaluation experiments (Zhang et al., 2016; Niederleithinger et al., 2018; Zhong et al., 2021) as well as an extension towards nonlinear material parameters (e.g. Xue et al. (2021)) have been published.

Besides the direct way of simulating the ultrasound wave propagation in concrete, there are alternative ways to describe the elastic energy distribution in space and time. In the diffusion theory, the diffusion constant and dissipation are used to describe the ultrasonic scattering and intrinsic attenuation (Anugonda et al., 2001; Becker et al., 2003). These two parameters can be estimated by comparison between the experimental data and theoretical predictions using the diffusion model. This allows people to describe the effect of uniformly distributed material damage (Ramamoorthy et al., 2004; Deroo et al., 2010). On the other hand, the diffusion model has also been used for calculating the sensitivity kernel of CWI in velocity changes or decorrelation which allows the imaging of the spatial distribution of the changes (Rossetto et al., 2011; Zhang et al., 2016).

Although the diffusion model has been successfully implemented to simulate the wave scattering and absorption, it is a simplification of the multiple-scattering process and hard to extend to more realistic cases, like the early coda, short source-receiver distances, anisotropic scattering or spatially variable

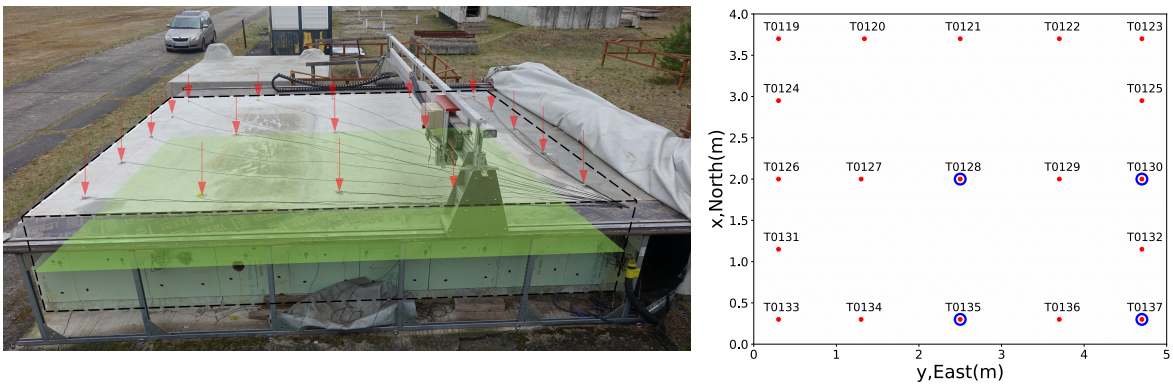



**Fig. 6.1:** Left panel: the photo of the concrete specimen. The red arrows indicate the location of ultrasonic transducers and the green shading indicates the plane where transducers were embedded. Right panel: the illustration of the transducer locations in the concrete specimen. The red points indicate nineteen ultrasonic transducers that serve as the energy sources as well as receivers. The blue circles are four temperature sensors that will be discussed in Section. 6.4.2. The orientation defined here is not the natural geographic coordinate.

heterogeneity. Wu (1985) first proposed the multiple scattering model and introduced the radiative transfer theory to seismology. To numerically solve the radiative transfer equations, the Monte Carlo method was introduced (Gusev & Abubakirov, 1987; Hoshiba, 1991), which allows for the ability to simulate wave scattering in the spatially variable heterogeneity and intrinsic attenuation media (Zhang et al., 2021). Instead of the diffusion constant and dissipation used in the diffusion model, the spatial distribution of fluctuation strength $\varepsilon$ in the random medium and the intrinsic quality factor $Q^{-1}$ describe the spatial variability of scattering and absorption.

The simulation of energy propagation with spatially variable properties using RTT allowed us to introduce the adjoint method initially developed in FWI (Tarantola, 1984; Tromp et al., 2005; Fichtner et al., 2006; Fichtner, 2010) for the imaging of scattering and absorption properties with scattered waves (Zhang & Sens-Schönfelder, 2022).

## 6.2 Experiments

To test the adjoint envelope tomography against real data we choose an acoustic experiment conducted in a reinforced concrete specimen at the German Federal Institute for Material Science and Testing (BAM). The sample has a size of 4 m × 5 m with a height of 0.8 m as shown in Fig. 6.1 (Epple et al., 2020). In this paper, the orientation defined as shown in Fig. 6.1 is for convenience to discuss and not the natural geographic coordinate. All directions in the following discussion refer to this definition of orientation. 19 ultrasonic transducers are embedded in the central layer of the specimen at 0.4 m height. The transducers serve as the energy sources of ultrasound with a center frequency of 60 kHz and as receivers. Both emission and recording of acoustic waves is laterally isotropic. This setup provides for 19 × 18 source-receiver combinations. The experiment has the following advantages for the present purpose: (A) due to the rather flat shape of the specimen and the placement of the transducers in its central plane we can restrict the energy propagation to the lateral directions and simplify the problem to 2D. (B) The boundary conditions of the lateral edges of the specimen can easily be modeled using mirror sources. (C) The embedded sensors that are located at least 30 cm away from the free surfaces reduces the excitation of surface waves which are not treated in our approach.

### 6.2.1 Data Processing

The ultrasound signals were recorded with sampling interval of 0.5 $\mu s$ for a lapse time of 5 ms. Seven identical experiments were performed on three consecutive days in October between 7:00 and 8:00 am. An illustration of original data excited at source T0120 and recorded by receiver T0135 is shown in Fig. 6.2(a). The first 200 samples, i.e. 0 - 0.1 ms precede the signal transmission and are recorded to control the noise level (Niederleithinger et al., 2018). The impulse at 1 ms lapse time is visible on all sensors and is caused by cross-talk between the high voltage source signal and the recording sensors. Data is detrended by subtracting its mean (Fig. 6.2(b)). The cross-talk is used to extract the envelope of the source signal and is then removed from the record (Fig. 6.2(c)) which band-pass filtered between 60 and 120 kHz to remove the high-frequency noise (Fig. 6.2(d)). The envelope of the filtered signal is extracted using the Hilbert transform (Fig. 6.2(e)). Envelopes of the repeated experiments are averaged to obtain the final envelope (Fig. 6.2(f)) for the inversion. The same processing is applied to the cross-talk to obtain the final envelope of the source signal (Fig. 6.2(f) inset).

From the processed envelopes we noticed that certain sensors systematically recorded smaller amplitudes than others, or excited less energetic waves. We attribute this to variable sensor coupling including the conversion between electrical and mechanical signals as well as the mechanical coupling between the transducer and the concrete. We estimate the coupling using the coda normalization method (Sato et al., 2012) which states that the signal envelopes in the late coda should be independent of location due to the equal distribution of elastic energy. We estimate one coupling coefficient for each transducer acting as source and receiver, separately by averaging the late coda envelope (3.5 - 4.7 ms) from the respective source or recorded at the respective station (Fig. 6.3(a)). Since the transducers act both as source and receiver, the coupling should have similar effects on both the emission and recording. This is consistent with the observations in Fig. 6.3(a). The influence on the envelope data from the i-th source to the j-th station is eliminated by dividing by the corresponding values in Fig. 6.3(a). An illustration of the coupling effect is shown in Fig. 6.3(b). The blue and red curves indicate two combinations exchanging the source and the station, which should be identical due to reciprocity. However, the sensor coupling introduces a difference between two curves shown in Fig. 6.3(b) but can be corrected using the coupling corrections (Fig. 6.3(c)).

### 6.2.2 Diffusion Model

The iterative inversion starts with an initial model. The density of the concrete is provided by Niederleithinger (2017) as 2.4 g$\cdot$ cm$^{-3}$ and the velocity of the wave is estimated from the arrival time of the ballistic waves as shown in Fig. 6.4. The transition from the noise level (blue) to wave signal (red) indicates the arrival of the ballistic wave (dashed line) with a velocity of 4.475 $m \cdot ms^{-1}$.

We have no prior information about the scattering and absorption properties, for the concrete in the present experiment. A simple description of multiple-scattering and intrinsic attenuation of ultrasound in concrete is provided by the diffusion model (Anugonda et al., 2001; Ramamoorthy et al., 2004). The 2D diffusion equation describes the energy radiating isotropically from a source (Wegler et al., 2006):

$$E_D(\mathbf{r}, t; \mathbf{r_i}) = E_0 \frac{1}{4\pi D t} e^{-\frac{r^2}{4Dt}} e^{-\frac{\omega}{Q} t} \,. \tag{6.1}$$

The diffusion energy density $E_D(\mathbf{r}, t; \mathbf{r_i})$ at position $\mathbf{r}$ with the lapse time $t$ is determined by the source energy $E_0$, diffusion constant $D$ and intrinsic factor $Q^{-1}$ at the specific angular frequency $\omega$. $\mathbf{r_i}$ is the position of the source while the distance between the source and receiver is $r = |\mathbf{r} - \mathbf{r_i}|$. To account for the existence of boundaries that reflect the acoustic energy, $E_D(\mathbf{r}, t; \mathbf{r_i})$ is summed for all mirror sources $\mathbf{r_i}^{mirr}$ corresponding to $\mathbf{r_i}$ (Zhang et al., 2018). Benefiting from the analytic solution of the diffusion

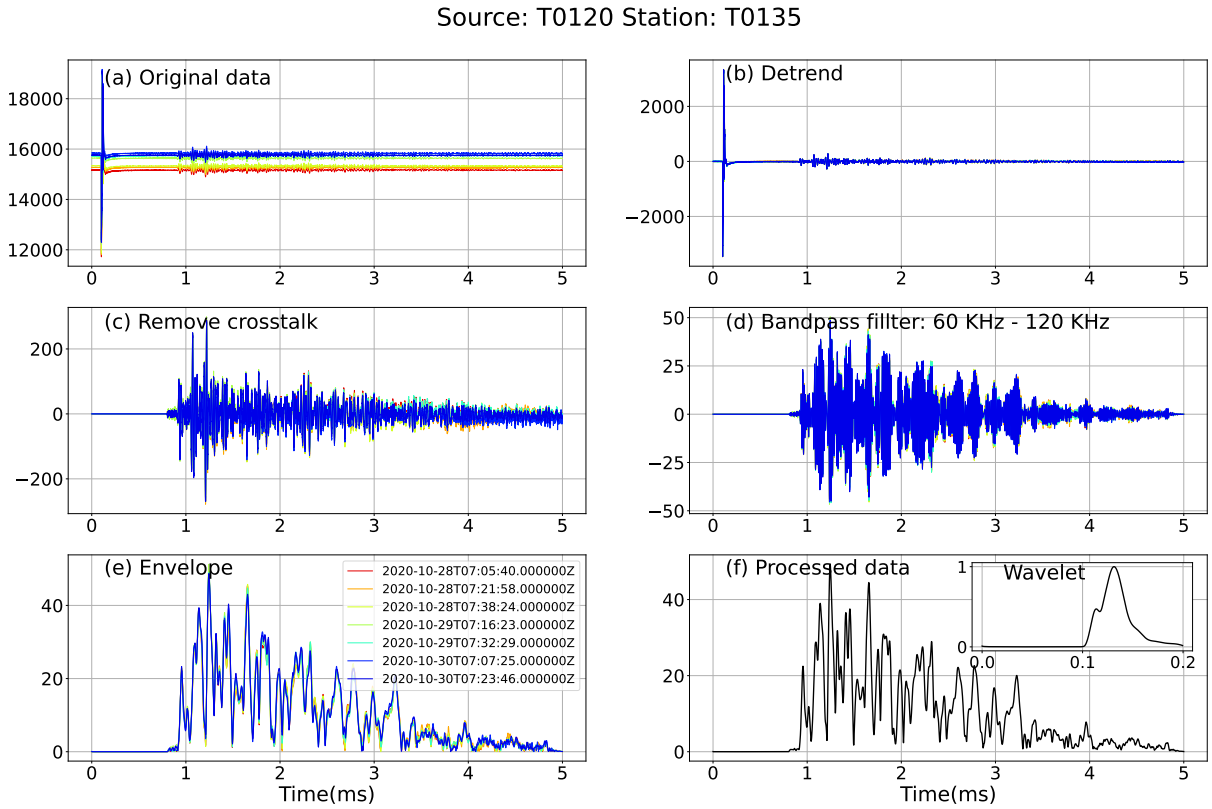



**Fig. 6.2:** The illustration of data processing for one source-receiver combination(Source T0120 and Receiver T0135): (a) the original data recorded in seven identical experiments; (b) the detrended data by subtracting its mean; (c) the cross-talk removed from the record; (d) the filtered data with band-pass filtered between 60 and 120 $kHz$; (e) the envelope of the filtered signal using the Hilbert transform; (f) the averaged envelope of the repeated experiments and the inset is the final envelope of the source signal.

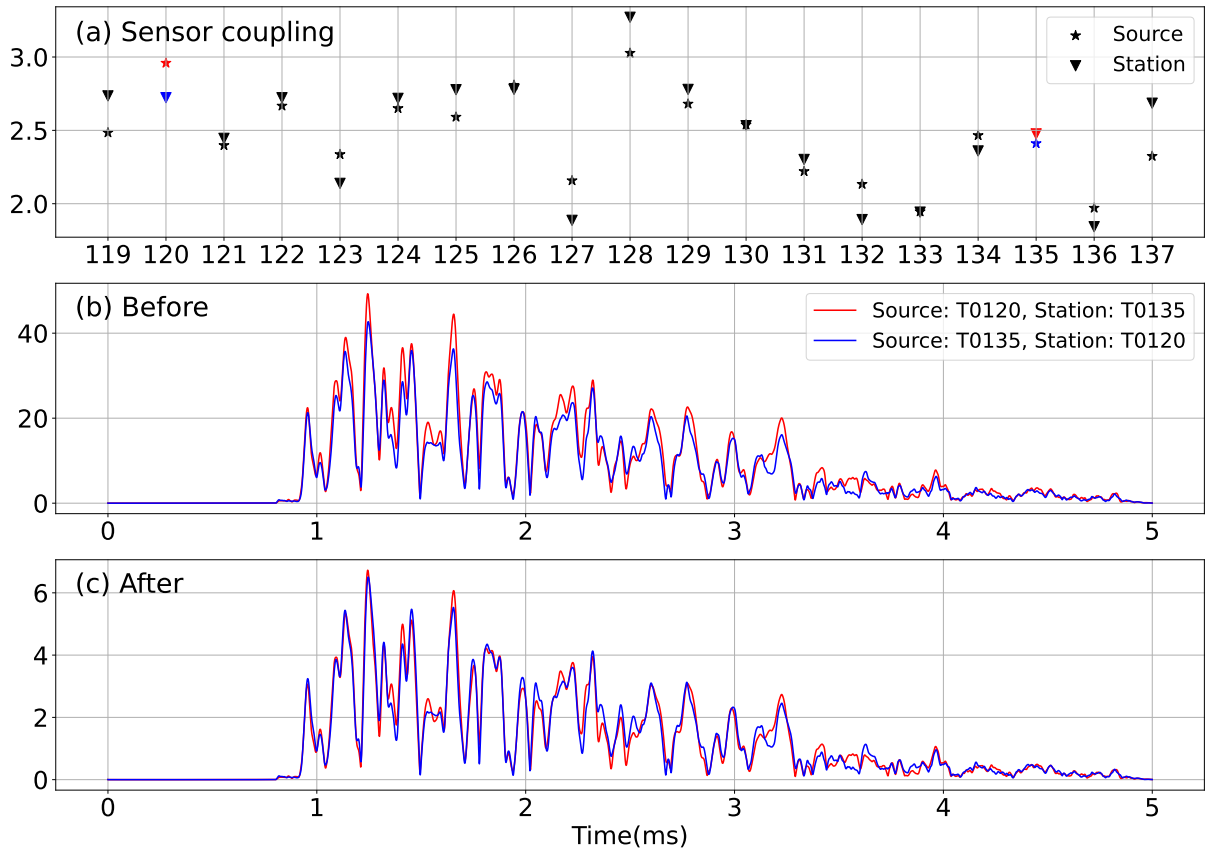



**Fig. 6.3:** The illustration of removing the effect from sensor coupling: (a) the coupling coefficients for each transducer acting as source (stars) and receiver (inverted triangles); (b) the blue and red curves indicate two combinations exchanging the source and the station, which should be identical due to reciprocity but not because of sensor coupling; (c) the corrected traces after using the coupling corrections (coupling coefficients used for the blue and red curves shown in (a) with the same color).

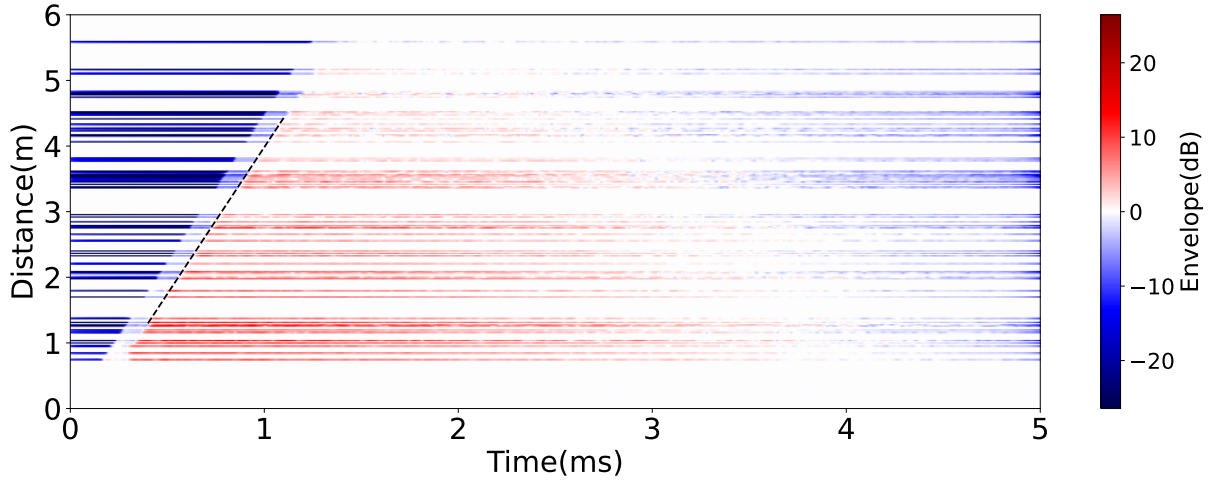



**Fig. 6.4:** The estimation for the velocity of the wave in the concrete specimen. All traces are displayed by the logarithmic scale with the exact distance between the source and the receiver. The dashed line indicates the boundary between the noise level (blue) and the wave signal (red), whose slope is 4.475 $m \cdot ms^{-1}$.

**Table 6.1:** All background parameters estimated for the radiative transfer equation.

| $\omega$ | $\rho$ | $\alpha_0$ | $Q^{-1}$ | $E_0$ | $\varepsilon$ | $a$ |
|---|---|---|---|---|---|---|
| $2\pi\cdot$ 60 $kHz$ | 2.4 $g\cdot cm^{-3}$ | 4.475 $m\cdot ms^{-1}$ | 0.003 | 12 | 0.13 | 0.011$m$ |

equation, we can estimate the parameters of the model by an interval search. Eq. 6.1 is rewritten as:

$$ln\left[E_D(\mathbf{r},t;\mathbf{r_i})\right] = lnE_0 - \left[ln(4\pi Dt) + \frac{r^2}{4Dt}\right] - \frac{\omega}{Q}t\ . \tag{6.2}$$

Expression 6.2 consists of three terms in which $lnE_0$ is constant. To speed up the process, we separately estimate $Q^{-1}$ from the later coda wave (3.5 - 4.7 ms) since the later coda wave is more sensitive to the intrinsic attenuation (Zhang & Sens-Schönfelder, 2022). $\left[ln(4\pi Dt) + r^2/4Dt\right]$ varies slowly in the late coda. Therefore, $-\omega/Q$ is easily estimated from the slope of the logarithmic envelope in the late coda. Fig. 6.5(a) shows the distribution of the estimated $Q^{-1}$ values from all source-sensor combinations. The mean and median value of this distribution are both 0.003 that will be used to estimate $D$ in the diffusion modeling and as initial model for inversion.

With the fixed value of $Q^{-1}$, the source energy $E_0$ can be extracted as the offset from the envelopes for each assumed diffusion constant $D$. An interval of $[50, 5000]\,mm^2/s$ with step-length 10 $mm^2/s$ is searched for the diffusion constant $D$. The diffusion model generated by Eq. 6.1 is convoluted with the wavelet shown in Fig. 6.2(f) to compare with the observable. For all source-receiver combinations, the distributions of $D$ and $E_0$ are shown in Fig. 6.5(b,c). According to this distribution, we fix the source energy $E_0$ in this study to 12. The diffusion constant $D$ does not directly correspond to the parameters used for the non-isotropic scattering in RTT. It corresponds to the transport scattering coefficient $g^*$ which is a version of $g$ that is weighted by the cosine of the scattering angle $\theta$. The relationship between $D$ and $g^*$ is given as (Wegler et al., 2006):

$$g_0^* = \frac{\alpha_0}{2D} \tag{6.3}$$

where $g_0^*$ is the average transport scattering coefficient that is defined as:

$$g_0^* = \frac{1}{2\pi}\int_{2\pi} g(\theta)\left[1-\cos(\theta)\right]d\theta\ . \tag{6.4}$$

$g(\theta)$ has been introduced as a function of scatter strength $\varepsilon$ and correlation length $a$ in Eq. 2.15. Assuming that the correlation length $a$ is uniform with $a = 0.011$ m (Anugonda et al., 2001) we calculate the values of $\varepsilon$ corresponding to the estimated values of $D$ using expressions 2.15, 6.3 and 6.4. The distribution of $\varepsilon$ is shown in Fig. 6.5(d). We fix $\varepsilon = 0.13$ as background parameter describing the small scale heterogeneity in the concrete specimen. Tab. 6.1 summarizes all background parameters estimated for the use with Eq. 2.14.

### 6.2.3 Monte Carlo Simulation

The radiative transfer equation is solved using the Monte-Carlo method to simulate the energy propagation (Zhang et al., 2021). To account for the free surface boundary conditions in the Monte-Carlo simulations the particles are reflected at the four sides of the model. In this study, 100 million particles are used for each simulation. The field generated by source T0120 in the initial model is illustrated in Fig. 6.6(a-f). Although the algorithm allows us to simulate in models with spatially variable $\varepsilon^2(\mathbf{r})$ and $Q^{-1}(\mathbf{r})$, here we only illustrate propagation in an uniform model with the background parameters given in Tab. 6.1. Note Fig. 6.6 only shows the energy density $E\,(\mathbf{r},t)$, while we actually

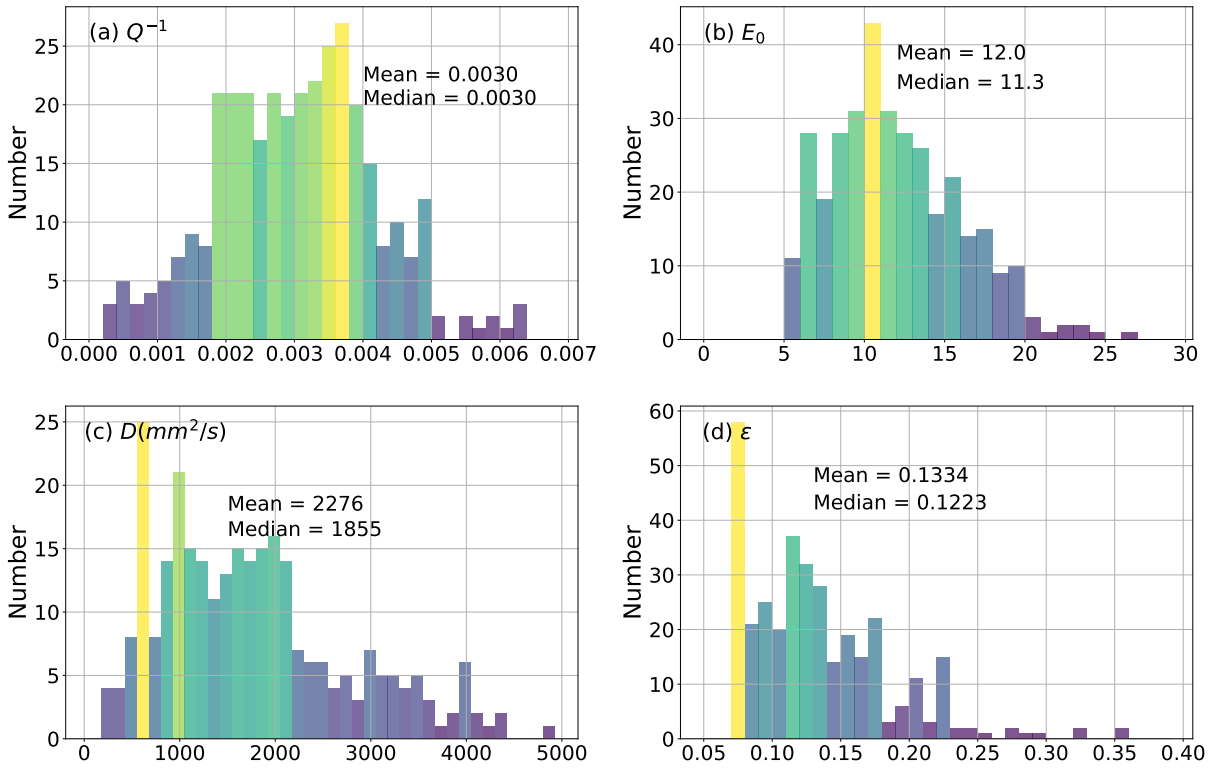



**Fig. 6.5:** The statistics histograms of the parameters: (a) $Q^{-1}$, (b) $E_0$ and (c) $D$ estimated from all traces in the diffusion model and (d) $\varepsilon$ calculated according to (c).

simulate the specific energy density $E(\mathbf{r}, \mathbf{n}, t)$ with information about the propagation direction.

The Monte Carlo method simulates a point-source in space and time. The simulation result is therefore convolved with the source wavelet and multiplied with the same source energy $E_0$ as diffusion model. Fig. 6.6(g) shows a comparison of one observed envelope with the diffusion model and the MC simulation in the background model. The blue and red curves represent the energy simulated with the diffusion model and radiative transfer equation, respectively.

## 6.3 Imaging

Starting from the initial model with uniform parameters estimated with the diffusion approximation, we use AET to infer the spatial distribution of the strength of heterogeneity and attenuation. Both material properties influence the energy propagation causing a the trade-off between changes in the scattering and absorption properties in a simultaneous inversion for both parameters as discussed in Zhang & Sens-Schönfelder (2022). For the ballistic wavefield, i.e. the energy that propagates without being scattered, the effect of scattering and attenuation is identical - leading to the impossibility of discerning both effects with direct waves. But the trade-off also exists for arbitrary sub-segments of the propagation path of coda waves. Only the combination between the energy that propagates directly between two points in the medium and the energy that is scattered between these points allows us to resolve the trade-off since heterogeneity increases the scattered part of the wavefield at the cost of the direct part. This trade-off means that strong spatial differences of one parameter unavoidably map into the other parameter to some extent (Cormier & Sanborn, 2019; Zhang & Sens-Schönfelder, 2022).

However, the fact that the early coda is important to image the heterogeneity while the later coda is more sensitive to intrinsic attenuation (Calvet et al., 2013; Zhang & Sens-Schönfelder, 2022) helps us to separately invert $\varepsilon^2(\mathbf{r})$ and $Q^{-1}(\mathbf{r})$ using the early and later coda, respectively. In this experiment, we

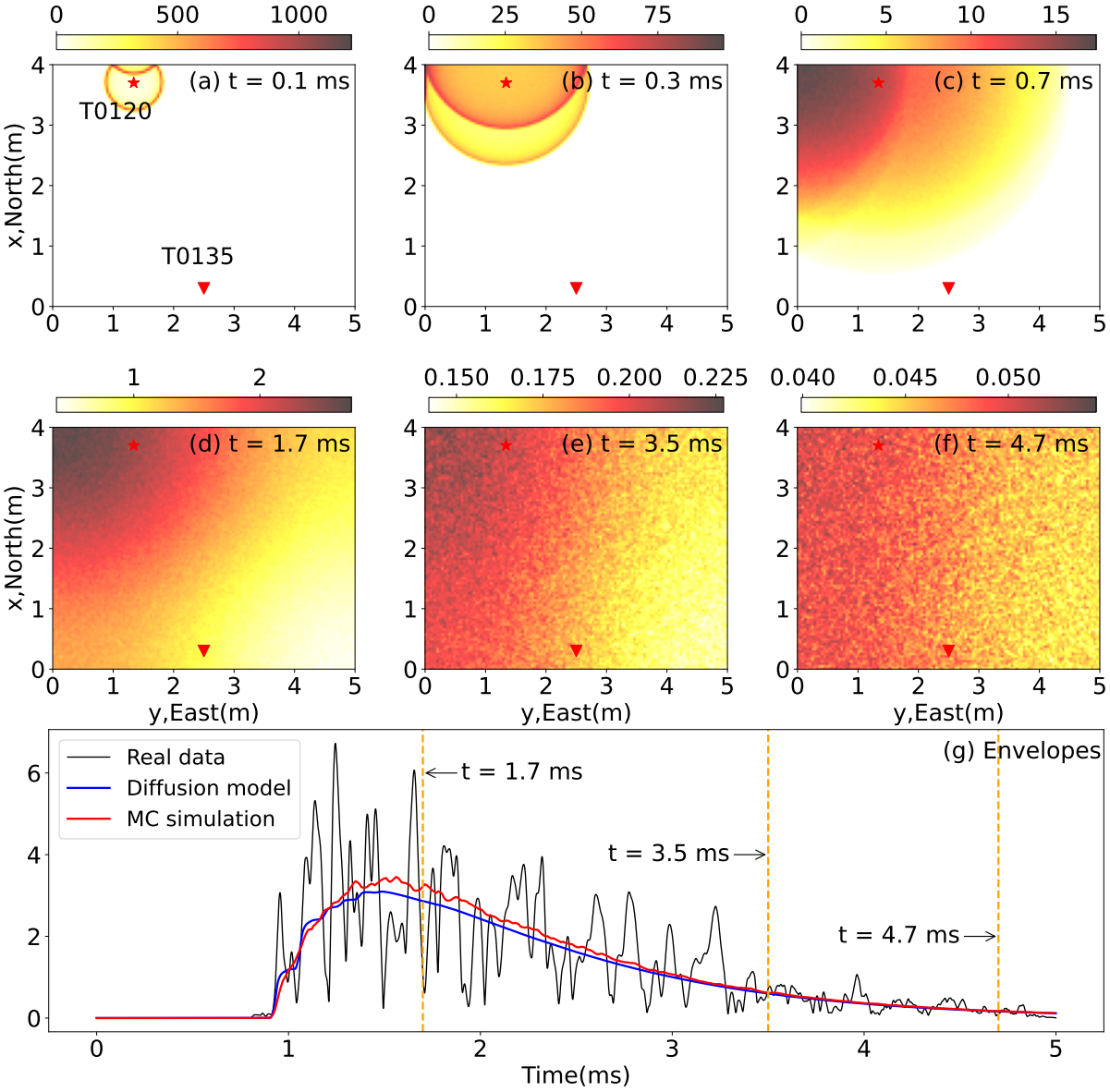



**Fig. 6.6:** (a-f) The snapshots of the simulated energy density field from T0120 (red star)at different lapse times. The scattering mean free path is 0.36 m and the mean free time is 0.08 ms. (g) the comparison among the envelopes recorded at T0135 (red inverted triangle) from the Monte Carlo simulation (red curve), the diffusion model (blue curve), and the real data from the concrete experiment (black curve). Note that the color scale range of each time in (a-f) is different and the energy density field has been multiplied with $E_0$ but not convolved with the source so that the values are not the same as the envelopes shown in (g).

simply define the early and later coda intervals by 1.7 - 3.5 ms and 3.5 - 4.7 ms respectively as shown in Fig. 6.6(g) and use these time windows to image the absorption and scattering structures successively.

### 6.3.1 Intrinsic Attenuation Inversion

We first focus on the intrinsic attenuation inversion with $Q^{-1}(\mathbf{r})$ since the absorption influences the whole envelope. The later coda wave (3.5 - 4.7 ms) is chosen as the time window to evaluate the misfit function and the initial model $Q_0^{-1}(\mathbf{r})$ is uniform with $Q_0^{-1} = 0.003$. The other parameters and the model of $\varepsilon^2(\mathbf{r})$ are all uniform based on Tab. 6.1 and remain constant during the inversion, meaning that only $Q_k^{-1}(\mathbf{r})$ is updated.

After 11 iterations of AET, the normalized misfit between the observed envelopes and synthetic data converges to 66% as shown in Fig. 6.7(a). The decrease of the misfit is very fast in the beginning since the initial model is uniform, slows down and stagnates from iteration 7. The benefit of iterative inversion as compared to a linear kernel-based inversion (Ogiso, 2019) is that the model is further improved after the first iteration based on the results of earlier iterations. The final inversion result is shown in Fig. 6.7(b). The distribution of $Q^{-1}(\mathbf{r})$ shows a dominant first order structure with a maximum in the center and a symmetry in the west-east and north-south directions. The decrease towards the sides is not isotropic with the east-west direction showing faster decrease than the north-south direction. We will discuss the interpretation of this result in the next section.

The misfit is the integral of the differences between the observed and modeled results in the specified time window. However, we can also directly check the data fit of the envelopes. Fig. 6.8 shows the data fit for some source-receiver combinations. The simulated envelopes in the final inverted model (red solid curves) are compared with the initial model (blue dashed curves) and the observed data (black solid curves). The locations of source and receiver in each combination are shown on the right side. For the north-south oriented combinations T0119 -126, T0123 -130, T0126 -133 and T0130 -137 which are located in the west and east, the envelopes of the inverted model become more similar to the observation compared with the initial model, as expected for a successful inversion that minimizes the misfit. This is caused by the decrease of $Q^{-1}$ along the western and eastern sides of the model. There are no significant improvements for station combinations T0120 -122 and T0134 -136 because already the initial model fits the observations reasonably well in these areas and the model update during the inversion is marginal. Envelope fits of the combinations T0124 -125 and T0131 -132 that transect through the whole specimen do not improve. In fact the fit of these long distance east-west combinations slightly degrades in favor of significant improvements of other pairs.

### 6.3.2 Scattering Inversion

Although the early coda wave is more sensitive to scattering, scattering inversion can benefit from using a more reasonable model of $Q^{-1}$ to suppress the influence of the absorption. In this step, we employ the inversion result shown in Fig. 6.7(b) as the model of $Q^{-1}(\mathbf{r})$ and keep it constant throughout the inversions for $\varepsilon(\mathbf{r})$. The initial model of $\varepsilon(\mathbf{r})$ is uniform and we use the earlier time window with lapse times 1.7 - 3.5 ms (c.f. Fig. 6.6).

9 iterations were conducted until the normalized misfit converged to 77% which is shown in Fig. 6.9(a). Note that although the simulation in the initial model in Fig. 6.9(a) is the same as the last one in Fig. 6.7(a), the absolute value of misfit is not since the time windows are different. Fig. 6.9(b) shows the inversion result of $\varepsilon(\mathbf{r})$. The dominant value of it is about 0.14 which is a little higher than the initial uniform model 0.13. The inferred distribution of heterogeneity has a more complex structure than the attenuation structure. Stronger scattering is inferred in two areas at the western and eastern boundaries and also in one anomaly of higher value in the south at about $y = 2\ m$. An elongated features

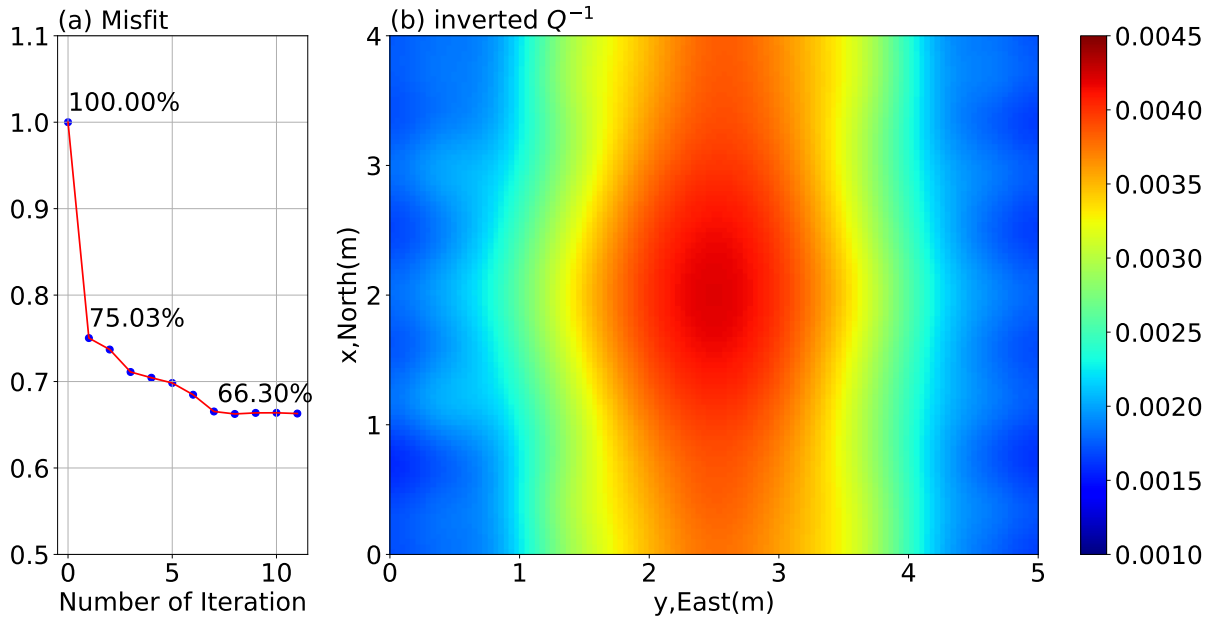



**Fig. 6.7:** (a)The misfits of the later coda time window varied with iterations for the absorption inversion. (b) The inversion result of $Q^{-1}(\mathbf{r})$ after 11 iterations.

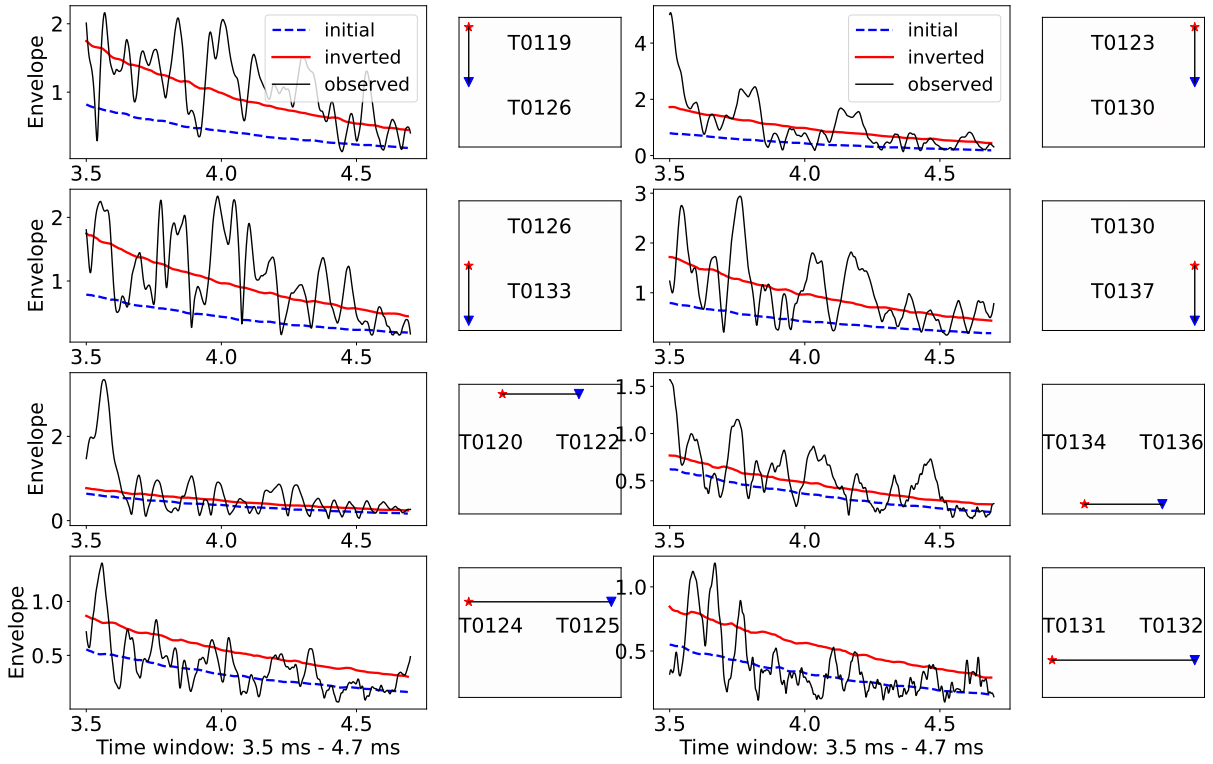



**Fig. 6.8:** The data fitting of different combinations (illustrated on the right side, red star and blue inverted triangle are source and receiver respectively) in the later coda wave (3.5 ms - 4.7 ms). The blue dashed and red solid curves indicate the envelopes simulated in the initial model and the inverted model shown in Fig. 6.7(b), respectively. The black curve is the real data from the concrete experiment.

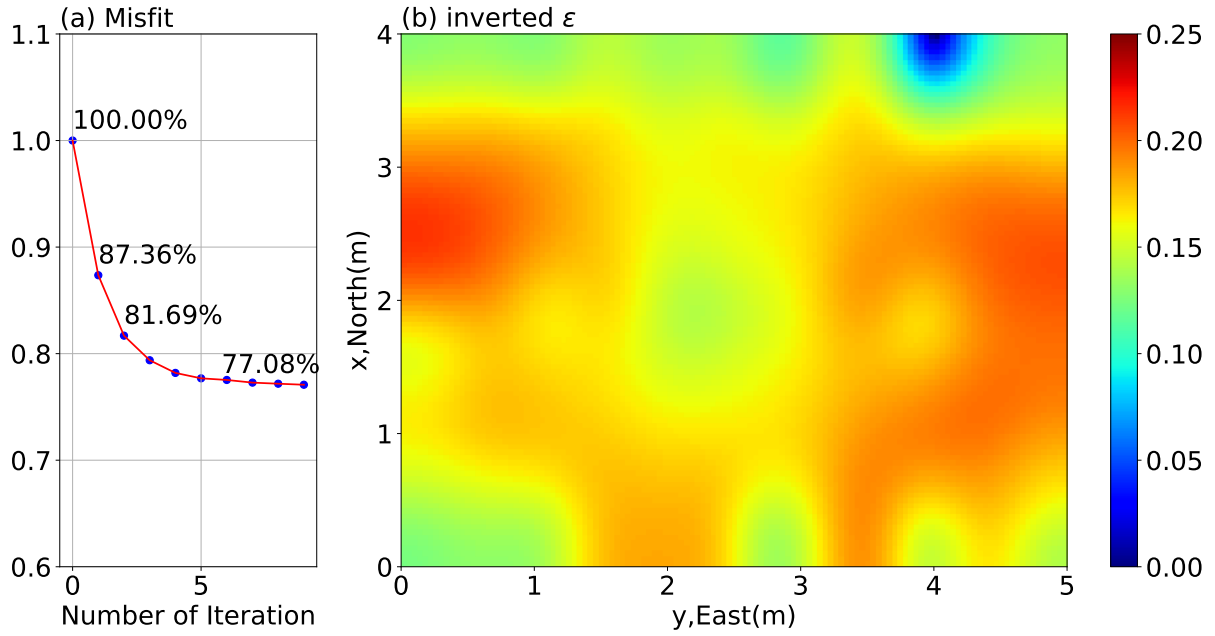



**Fig. 6.9:** (a)The misfits of the early coda time window varied with iterations for the scattering inversion. (b) The inversion result of $\varepsilon(\mathbf{r})$ after 9 iterations.

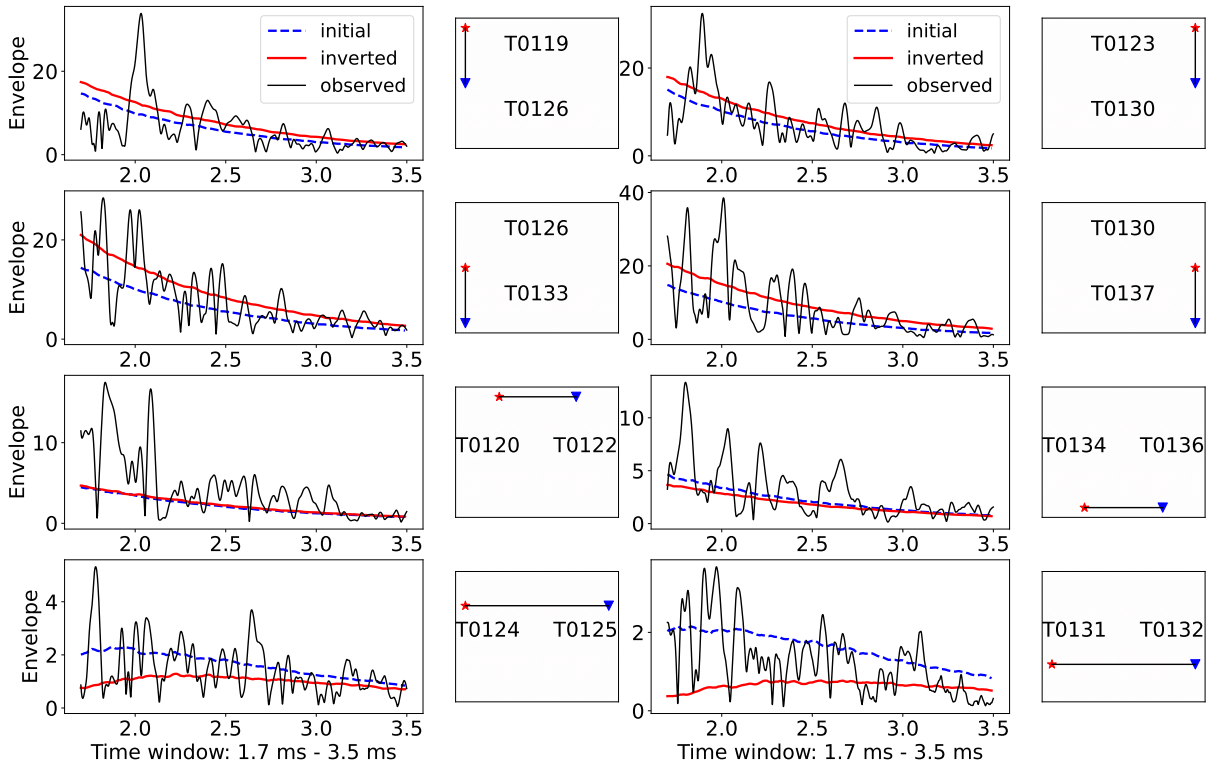



**Fig. 6.10:** The data fitting of different combinations (illustrated on the right side, red star and blue inverted triangle are source and receiver respectively) in the early coda wave (1..7 ms - 3.5 ms). The blue dashed and red solid curves indicate the envelopes simulated in the initial model and the inverted model shown in Fig. 6.9(b), respectively. The black curve is the real data from the concrete experiment.

extends from the northern to the southern edge at about $y = 3.4\ m$. A very low-value anomaly that indicates reduced heterogeneity is located in the north-east corner. Interpretations are discussed in the next section.

The data fits are shown in Fig. 6.10. Similar to what we discussed in Fig. 6.8, the different-distance combinations are compared in the early coda waves time window. The inversion result is dominated by the short-distance combinations which achieve a significantly improved data fit during the inversion. The medium- and long-distance combinations do not improve clearly. Note that the y-scale of the graphs in Fig. 6.10 is variable and combinations T0124-125 and T0131-132 have far smaller amplitudes.

## 6.4 Discussions

### 6.4.1 Misfit Evolution

In Section. 6.3, we have described two successive inversion runs for $\varepsilon^2(\mathbf{r})$ and $Q^{-1}(\mathbf{r})$ using the early and later coda, respectively. We start with uniform models of both parameters, firstly update the model of $Q^{-1}(\mathbf{r})$ only, and then fix the $Q^{-1}(\mathbf{r})$-model and continue to update $\varepsilon^2(\mathbf{r})$. The time windows of the misfit are chosen to use only the later coda for intrinsic attenuation inversion and only the early coda for scattering inversion because of their sensitivities. Of course, the misfits of both time windows varied in both inversions. Fig. 6.11 shows the evolution of the misfits of both time windows for the whole inversion process. The red and blue curves indicate the misfits of the later and early coda, respectively. The solid parts of the curves show the misfits that are optimized for during the inversion (they have been shown in Fig. 6.7(a) and Fig. 6.9(a)) while the dashed lines indicate the misfit during the optimization of the other time window.

The whole inversion is separated into two periods shown in Fig. 6.11. In the first period when we only update $Q^{-1}(\mathbf{r})$ (red domain), the misfit of the later coda (the red solid curve) decreases since the misfit kernel is based on this time window. Reasonably, with the improvement of the attenuation model the misfit of the early coda (the blue dashed curve) decreases as well although it is not used to guide the inversion. During the subsequent updating of $\varepsilon^2(\mathbf{r})$ (blue domain) the misfit of early coda time window continues to decrease since it is used to calculate the adjoint source. On the contrary the misfit of the late time window which is not used in this step re-increases slightly which is not surprising since this time window was already optimized for in the $Q^{-1}(\mathbf{r})$-inversion and does not inform the $\varepsilon^2(\mathbf{r})$-inversion. However, the mistfit change in the second run is dominated by the decrease of the misfit in the early time window. Using both time windows together to guide the second part of the inversion run would possibly have damped the misfit increase in the late time window, at the expense of smaller improvements in the early time window.

### 6.4.2 Interpretation

We begin the discussion with an interpretation of the inferred attenuation. The attenuation anomaly (Fig. 6.7b) is symmetric with respect to west-east and north-south axis in the center of the specimen and appears to be affected by some large scale influence on the specimen rather than internal small scale differences. Three processes could globally affect the specimen and result in a perturbation with the symmetry observed in the attenuation structure: (A) diffusion of humidity, (B) temperature changes and (C) stress distribution.

To investigate this hypothesis we make use of supplemental instrumentation. Additional to the 19 ultrasonic sensors, there were four temperature sensors embedded in the concrete specimen (shown in Fig. 6.1) which measured the internal temperatures on three consecutive days in the morning between 6

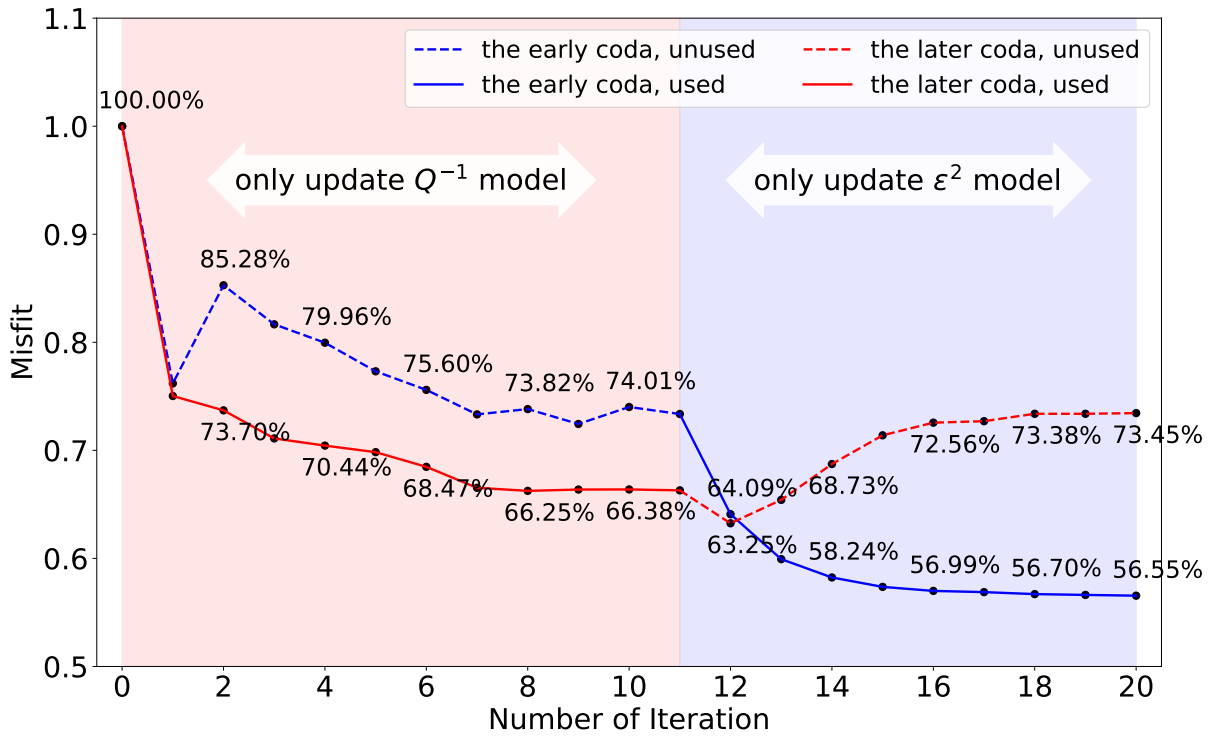



**Fig. 6.11:** The misfits of the later (the red curve) and early (the blue curve) coda waves as the time window varied with the whole inversion. The red and blue domains respectively indicate the absorption and scattering inversion conducted in Section. 6.3. The dashed parts of the curves are not used in the inversion.

a.m and 8 a.m as shown in Fig. 6.12(a). This experiment was conducted during a phase of decreasing temperatures in autumn. The temperature at each sensor decreased during the three successive days but the sensors maintained rather constant offsets from one another. The central sensor T0128 shows highest temperatures compared to the sensors closer to the rim. Smallest temperatures are observed in the corner of the specimen at sensor T0137 while intermediate temperatures are observed along the sides. We use the temperature measured on Oct. 28th at these four sensors to obtain an idea of the temperature distribution within the specimen. We therefore use the geometric symmetry of the sensor locations to interpolate the observations throughout the whole concrete in 2D using adjustable tension continuous curvature splines by Generic Mapping Tools (GMT) (Wessel et al., 2013; Smith & Wessel, 1990). The resulting temperature distribution within the concrete is shown in Fig. 6.12(b). This is clearly a rough estimate of the internal temperature distribution, but it shares clear similarity with the inversion result of $Q^{-1}(\mathbf{r})$ shown in Fig. 6.7(b).

It has been demonstrated that the temperature changes of the concrete can result in the velocity perturbation but the sensitivity is only about 0.05 $\%K^{-1}$– 0.15 $\%K^{-1}$ (Niederleithinger & Wunderlich, 2013; Epple et al., 2020; Larose et al., 2006). Since the maximum temperature change during the experiment is only 0.5 $K(^{\circ}C)$ the observed temperature changes will thus have a negligible influence on the propagation velocity and thus leave the envelopes unaffected which warrants the assumption of uniform and constant velocity in this experiment. We did not find conclusive evidence in the literature for the influence of temperature on attenuation in concrete or similar aggregates (at the present temperature (Zong et al., 2020)).

The influence of humidity on attenuation has been clearly documented by a number of authors (Clark et al., 1980; Green et al., 1993; Tisato & Quintal, 2014). Unfortunately in-situ observations of humidity are not available to us and the specimen is insulated from the sides and covered for protection against rain so that the humidity might be more or less uniform in the volume.

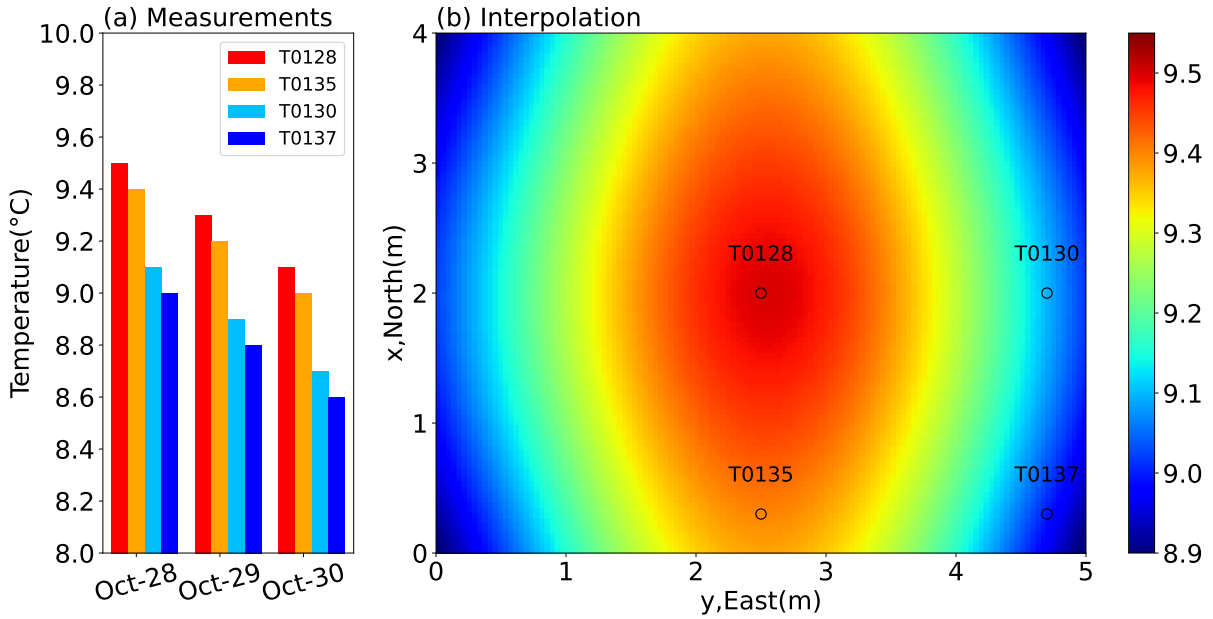



**Fig. 6.12:** (a) The temperature measured on three consecutive days morning between 6 a.m and 8 a.m from four embedded temperature sensors. (b) The temperature distribution of the whole concrete interpolated with four temperature sensors (the black circles) based on the geometric symmetry.

A distribution of absorption with a very similar symmetry pattern in a sample of comparable size was found by Liu & Guo (2005). These authors imaged the attenuation in a reinforced concrete block under the highway bridge pier cap which had a size of 6 m × 8 m with a height of 1.5 m. Using direct waves Liu & Guo (2005) inferred an inverse intrinsic quality factor of 0.0063 in the center of the block. This value is close to our result 0.0045. Towards the sides of their block, attenuation increases 7 times while it decreases 4 times in our results.

Different from absorption, the heterogeneity of the medium appears to be governed by internal structure rather than an external influence since the inferred distribution is much more structured. Fig. 6.13 shows the construction drawing of this concrete specimen. The strongest anomaly of increased heterogeneity is found at the western edge of the specimen. This area corresponds to a volume of the specimen that was cast with a different kind of concrete (salt concrete: 1600×1000×250 mm). Here, salt was added to the concrete mix to be able to provoke rapid corrosion of rebar at a later stage. As the concrete was poured separately by a different team and cured under different conditions, a different density and porosity can be expected. We interpret the increased scattering inferred in this region to be caused by the different properties of the salt-concrete.

The second prominent area of increased heterogeneity located in the west does not directly correspond to model features from the construction plans. During the installation of the embedded sensors an anomaly was detected in this area. While the calculated quantity of grout was sufficient to completely fill the boreholes in all other locations, almost three times the amount was required for refilling the borehole of sensor T0132. It can therefore be assumed that cavities were unintentionally created in this area during concreting, which now contribute to the increased scattering.

Before this experiment, there were three heating cartridges inserted in the east, south and northwest (Heating Cartridge A, Band C respectively in Fig. 6.13). Heating Cartridge A had been used to heat the concrete to 510 °C (Niederleithinger, 2017) while the other two had not been activated. The concrete after high-temperature heating generated thermal cracking and stress changes (Hager, 2013) that increase scattering.

Three autoclaved aerated cube concretes with size of 0.3 m, four horizontal plastic pipes and one

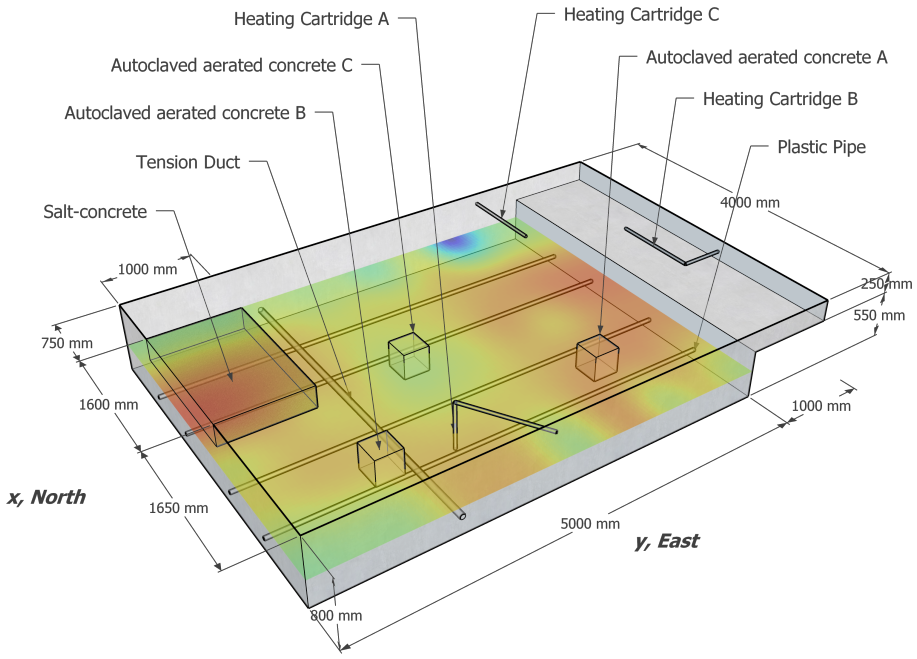



**Fig. 6.13:** The construction drawing of the concrete specimen. The western anomaly block is the salt concrete that is different from the background material with the size of 1600×1000×250 mm. Four horizontal plastic pipes, one vertical clamping channel, three autoclaved aerated cube concretes and three heating cartridges are embedded.

horizontal perpendicular to the plastic pipes tension duct are also embedded in the concrete. Structures of these sizes are out of the inversion resolution but can also affect the scattering to some degree.

A very prominent anomaly that is left to be discussed is the low-$\delta\varepsilon$ anomaly close to sensor T0122. This anomaly is located right at the boundary of the inversion domain and converges towards extremely low values of heterogeneity, i.e. locally homogeneous material. Its location directly on the boundary close to a corner of the model leads us to the interpretation as an artifact. Fitting envelopes of waveforms always requires significant averaging. In theory this averaging should be achieved by repeated observations in statistically identical realizations of the experiment. In reality there is only a single specimen and the averaging is realized on the one hand based on ergodicity by using long time windows for the comparison between observations and synthetics and on the other hand by using multiple source and receiver combinations. While the effect of long time windows is the same everywhere in the sample the averaging by different sensor combinations is not. The reflecting boundary conditions reduce the effective averaging by a factor of two along the edges and by a factor of four in the corners. A prominent wiggle in the waveform that can coincidentally originate from the constructive interference of scattered waves results in a strong pulse in the envelope (cf. Figs. 6.6, 6.8 and 6.10). Such a pulse can push the inversion into a certain direction and cannot effectively be compensated by other sensor combinations with sensitivity to the same location since the mirror sources have identical waveforms.

### 6.4.3 Resolution Test

Different tools exist to study the capabilities of the combination of a measurement setup and an inversion method. Checkerboard tests (Lévěque et al., 1993) use a periodic pattern of variable wavelengths to infer the minimum size of a feature to be resolved in different parts of the domain. Analytical approaches use the sensitivity of the misfit function to changing perturbations (the Hessian) at the different locations in

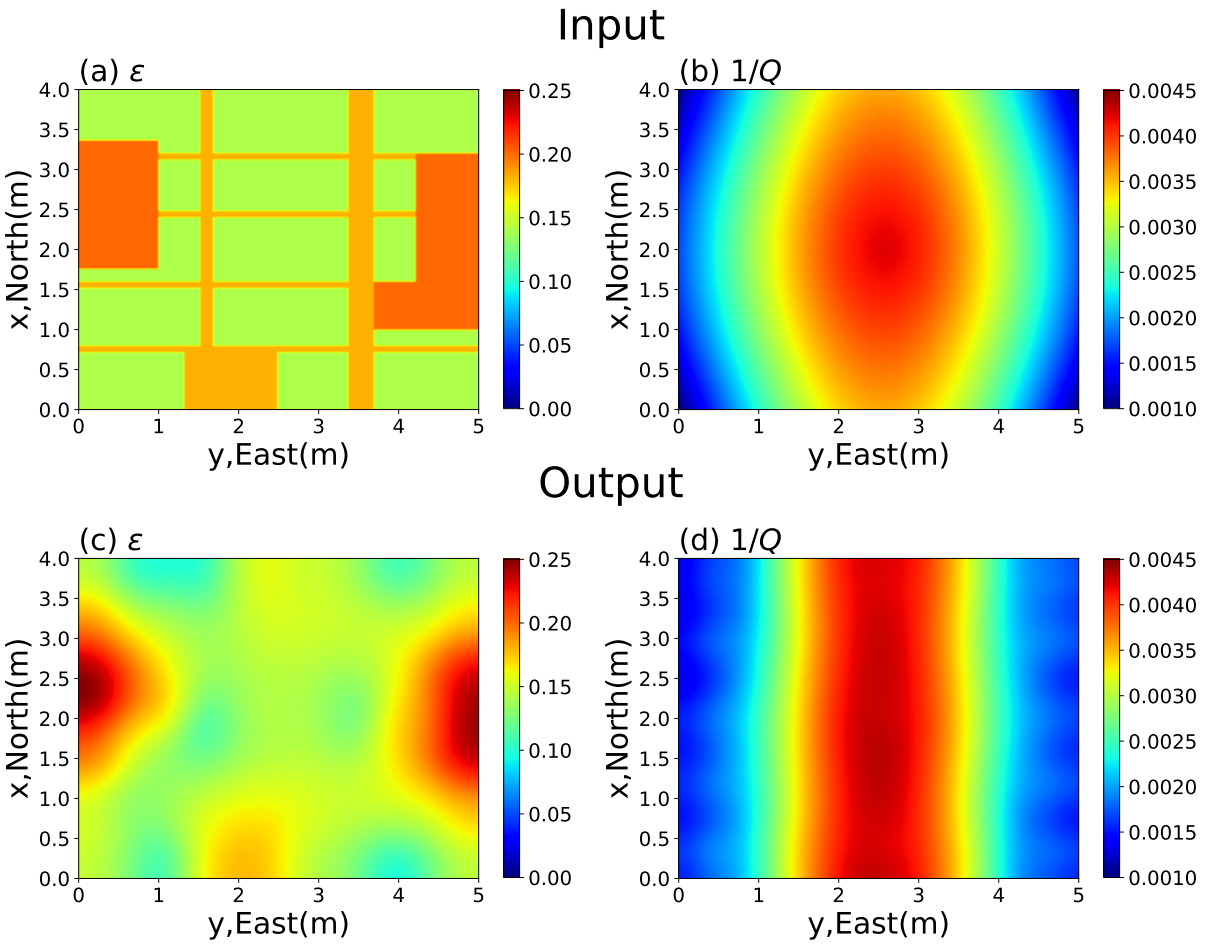



**Fig. 6.14:** The resolution test: (a) the input model of $\varepsilon(\mathbf{r})$ based on the construction plans of the concrete specimen; (b) the input model of $Q^{-1}(\mathbf{r})$ based on the temperature distribution shown in Fig. 6.12(a). (c) and (d) the inversion results of two parameters following the same workflow as the inversion for the laboratory experiment.

the domain to estimate the resolution capabilities (Fichtner & Trampert, 2011).

We take a different approach for the following reasons. Since we use reflecting boundary conditions in a domain with an regular distribution of sensors we can assume that also the resolution capabilities are rather uniform which would limit the value of a checkerboard test. The analytic approach using the Hessian is either computationally very expensive or requires further development, that is beyond the present scope. Here we ask the question: What would the inversion obtain if the structures were as we interpret it from the actual imaging. Technically this question is answered by inverting a simplified version of the obtained result that contains all structures which are regarded as relevant and interpreted. This approach is often used in tomography to confirm that the interpreted structure could indeed be resolved by the imaging (Koulakov et al., 2009; Jiang et al., 2014). For nonlinear problems such statements are more useful than theoretical values of resolution length in a homogeneous background model.

The resolution test is conducted with the exact same procedure as used in Sec. 6.3 for the inversion, including locations of sources and receivers, the parameters of the initial model and time window choices. The input and output models are both shown in Fig. 6.14.

The test model of $\varepsilon(\mathbf{r})$ is based on the construction plans of the concrete specimen and the inversion result. The background value of $\varepsilon(\mathbf{r})$ is designed not to be the same as the initial model but taken from the inversion result as 0.14. Fig. 6.14(a) shows the input model for the resolution test that contains the structures obtained in the inversion and some small elongated anomalies along the locations of channels and reinforcement bars in the specimen. The input model of $Q^{-1}(\mathbf{r})$ (Fig. 6.14(b)) is based on the

temperature distribution shown in Fig. 6.12(a).

Panels (c) and (d) of Fig. 6.14 show the resulting outputs of the synthetic inversion test. The output of $\varepsilon(\mathbf{r})$ shows that the background value is recovered well although it was different from the the initial model. The three larger anomalies are localized well, but their shapes are not recovered in detail due to limitations imposed by the by the number and the setup of sources and receivers and the intrinsic smoothing of imaging with the envelope information, only. Likewise the thin elongated anomalies are not resolved as could be expected from the locations of the 19 sensors of which only three are not arranged along the rim of the specimen. The incorrectly inferred shape of the anomalies is connected to their peak amplitudes which are partially overestimated during the inversion. Since the scattered energy depends to first order on an integral scattering strength of the anomaly higher values in the centers of the larger anomalies compensate for the lower strength along the edges of these anomalies. The inversion result of $Q^{-1}(\mathbf{r})$ recovered the input structure well. However, decay in the north-south direction is underestimated and the peak anomaly is overestimated.

From this test we conclude that the first order features interpreted from the imaged attenuation and scattering structures would indeed show up as observed in the results. Due to ambiguity and limited resolution we cannot exclude that smaller anomalies are present in the specimen.

## 6.5 Conclusions

This research presents the analysis of an acoustic experiment conducted in a 4 m by 5 m large concrete specimen equipped with embedded acoustic sensors. We applied adjoint envelope tomography to image the distribution of small-scale heterogeneity and intrinsic attenuation inside the specimen. To interrogate the structure below the resolution limit of conventional tomography, AET was proposed to invert for the statistical properties of the small scale heterogeneity as complementary information to the deterministic structures that can only be imaged at larger scales. Although AET had been successfully tested in numerical experiments, the application to experimental data in the present paper increases confidence in the methodology in view of further applications to seismic imaging of the Earth.

We performed this experiment with ultrasonic transmission form embedded transducers in reinforced concrete in analogy to seismic wave propagation in the Earth. The data recorded by 19 transducers are compared with simulations of energy propagation based on the Radiative Transfer Equation. This forward problem is solved by modelling the 2-D multiple nonisotropic scattering in an acoustic medium with spatially variable heterogeneity and attenuation using the Monte-Carlo method. The misfit between the observed and modeled envelopes is minimized by iteratively updating the model with the adjoint method. The whole workflow of AET for the real data is introduced including the processing of the data and the investigation of background values with the diffusion model. The fluctuation strength $\varepsilon$ and intrinsic quality factor $Q^{-1}$ respectively representing the spatial variability of scattering and absorption are separately inverted from different time windows. On the one side, the absorption inversion result shows a strong point-symmetric geometry which we interpret as some large-scale spatially variable in the specimen, but without a direct evidences for the causative process, e.g. temperature, humidity or stress.

The inverted distribution of scattering properties shows a more complex structure that can – to some extent – be interpreted in terms of the known internal structure of the test specimen. The largest anomaly of increased heterogeneity corresponds to a volume containing salt-concrete. Other anomalies are not as clearly linked to the known features of the concrete and a strong anomaly of decreased heterogeneity exists at the edge of the specimen that is interpreted as an artifact from envelope fluctuation that are insufficiently averaged at the reflecting boundaries of the model domain.

Despite obvious room for improvement in terms of spatial resolution and power to resolve the trade-off between scattering and attenuation the present results are encouraging. The spatial variability of

attenuation and scattering strength improved the data fit by about 35% when averaged over both time windows. This number appears small but cannot directly be compared to improvements known from waveform inversion. Two effects contribute to the limitation of the data fit. Firstly the observed envelopes are obtained in a real experiment and cannot be averaged over an ensemble of test specimens and thus show fluctuations introduced by the interference of scattered waves that cannot be fit. Secondly also the simulated envelopes contain additional fluctuations from the Monte-Carlo type simulation.

Future investigations to test the performance of the AET on real data will have to include dedicated test specimens with known scattering and attenuation properties. Even though the present concrete block with the embedded sensors was well suited for an application of AET is was already cast and the different types of concrete could not the analyzed separately to obtain ground truth. An important field of application for the presented approach is the monitoring of medium perturbations with coda waves (Sens-Schönfelder & Brenguier, 2019). The spatial sensitivity of coda wave based monitoring depends on the distribution of heterogeneity (Kanu & Snieder, 2015) and can thus be improved with the presented method. We hope that AET will contribute to non-destructive testing of civil engineering structures and investigations of wave propagation in the Earth.

# Chapter 7

# $P_{\text{diff}}$ Coda in the Core Shadow at High Frequencies

## 7.1 Introduction

The observation of complexity in the seismograms, on the one hand, always inspires seismologists to figure out the origin of different seismic phases. On the other hand, the investigation of these phases revealed many details about the interior structure of the Earth. The traveltime and amplitude of direct phases, like $P$-, $S$-, or surface-waves, are used for tomography of the large-scale structural features (Aki & Lee, 1976; Li & Van Der Hilst, 2010; Zhu et al., 2012; Bozdağ et al., 2016; Lei et al., 2020; Simmons et al., 2021). These studies show that Earth exhibits heterogeneity from the lithosphere down to the deep mantle Earth. Two large low-shear-velocity provinces (LLSVPs) above the Core-Mantle Boundary (CMB) indicate the presence of superplumes under the Pacific Ocean and Africa (Maruyama et al., 2007) which are important for Earth's dynamic processes and the thermal and chemical evolution of the mantle (Garnero, 2004; French & Romanowicz, 2015; Koelemeijer et al., 2017).

Additionally there is seismic energy that does not arrive as one of the main seismic phases. Such waves that usually tail the main seismic phases were first interpreted as scattered waves named *seismic coda* by Aki (1969). Seismologists observed coda waves following most direct wave arrival and also found some scattered energy arriving before the direct phases forming a *precursors* to the seismic phase. Scattered energy is most prominently observed in short-period seismograms, like $P$ coda, $P_{\text{diff}}$ coda, $PP$ precursors , $PKP$ precursors and $PKiKP$ coda (Shearer, 2015). However, in the seismology there is still some uncertainty about the precise origin of some of these seismic observation. The $PKP$ precursors were firstly interpreted as scattered waves caused by small-scale heterogeneity near the CMB (Cleary & Haddon, 1972; Haddon & Cleary, 1974; Hiemer & Thomas, 2022). Later, some studies interpreted them as result of scattering that happened throughout the whole mantle (Hedlin et al., 1997; Margerin & Nolet, 2003; Mancinelli & Shearer, 2013). The existence of small-scale heterogeneity in the whole mantle also matches results obtained from the stacking of $PP$ precursors signals (Bentham et al., 2017) which are thought to originate from scattering in the crust and uppermost mantle (Cleary et al., 1975; King et al., 1975) or even including the middle mantle (Rost et al., 2008).

The strong small-scale heterogeneity in the near-surface is regarded as the main cause of the $P$ coda (Aki, 1973). However, synthetic models of scattering in the whole mantle also fit global stacks of teleseismic $P$ coda (Shearer & Earle, 2004). With the epicentral distance increasing above $\approx 100°$, the core shadow inhibits the arrival of $P$ waves at the surface. Diffraction along CMB generating the $P_{\text{diff}}$ phase allows to observe waves propagating with mantle velocities at the longer distances. However, the diffraction process is most effective at long wavelength and diminishes at high frequencies (Rost et al., 2006). Besides the diffracted waves, the waves passing through the core e.g. $PKP$ constitute the first arriving energy in the long-period recording at long distances. Bataille & Lund (1996) observed

that the ratio between the amplitudes of the ballistic phase and the coda changes drastically during the transition from $P$ to $P_{\mathrm{diff}}$. Therefore, $P_{\mathrm{diff}}$ coda was suggested to originate from multiple scattering near the CMB (Bataille et al., 1990; Tono & Yomogida, 1996). A more detailed investigation of $P_{\mathrm{diff}}$ coda was conducted by Earle & Shearer (2001). Their global stacking of short-period recordings for $P_{\mathrm{diff}}$ coda extends to long epicentral distances up to 130°. They modeled the observation with single-scattering theory but considered small-scale heterogeneity throughout the whole mantle. $P_{\mathrm{diff}}$ coda are generated from the scattering of both $P$ and $P_{\mathrm{diff}}$.

In the present study, we focus on the $P_{\mathrm{diff}}$ coda that used to be defined as the waves following the $P_{\mathrm{diff}}$ arrival time. We present new observation of $P_{\mathrm{diff}}$ coda at high frequencies at a very long distance (beyond 150°) in stacked records of huge-magnitude global earthquakes in a time window between the arrival times of $P_{\mathrm{diff}}$ and $PKIKP$. We generate synthetic data with a proposed heterogeneity model (Bentham et al., 2017) to interpret the origin of the $P_{\mathrm{diff}}$ coda.

## 7.2 Observation

For the investigation of $P_{\mathrm{diff}}$ coda in the core shadow, the epicentral distance range is chosen from 70° to 160°. Long-distance $P_{\mathrm{diff}}$ coda waves are hard to detect using small earthquakes since the energy of these waves is small compared to the usual noise level at seismic stations. As a result, we select earthquakes with magnitude $\geq 7.9$ $Mw$ from years 1994 to 2021 resulting in 52 events. The time window is chosen from $280s$ before the arrival time of the $P$ or $P_{\mathrm{diff}}$ phase to $280s$ after $PKiKP$ or $PKIKP$ arrival which allows us to evaluate the noise level and the signal-noise ratio. For each event, the data is downloaded from all available stations in the IRIS and GFZ archives. After removing the instrument response and detrending, the vertical component of the data is filtered by a band-pass filter within the frequency band $1.0 \sim 2.0$ $Hz$. The envelopes of the seismograms are computed using the Hilbert-Transform and then are smoothed twice by computing a central moving average, respectively with time windows of 2.5 s and 8 s length.

Figure 7.1 shows the theoretical arrival times of seismic phases from a 600-km depth earthquake in the $ak135$ velocity model calculated using the Python package *obspy.taup* (Crotwell et al., 1999). Additionally it shows processed envelopes in different distance regimes from a 600-km-deep event with 8.2 $Mw$ on August 19, 2018, in the Fiji Islands region. The traces are all aligned to the arriving time of the earliest core phase (either $PKiKP$ or $PKIKP$). In accordance with common naming we define the $P$ coda or $P_{\mathrm{diff}}$ coda as the coda waves following the theoretical arrival time of the $P$ or $P_{\mathrm{diff}}$ phases, even though this is somewhat misleading in terms of the origin of these waves. At short distances, e.g. 70° - 102° (the gray area), $P$ coda window contains scattered $P$ waves but additionally also coda waves or precursors of $pP$, $sP$ and $PP$ phases. With distance increasing, the $PKiKP$ or $PKIKP$ becomes the first arrival in the core shadow since the $P_{\mathrm{diff}}$ do not propagate at so high frequency (Rost et al., 2006). The same holds for $pP_{\mathrm{diff}}$ and $sP_{\mathrm{diff}}$. However, scattered $P$-energy can arrive in the 200 $s$ long $P_{\mathrm{diff}}$-coda window before $PKiKP$ or $PKIKP$ arrivals (the light orange area). This energy has been studied up to 130° distance by Earle & Shearer (2001). The constituents of this intermediate $P_{\mathrm{diff}}$ coda window are also manifold. It may contain scattered $P$ or $P_{\mathrm{diff}}$ energy but also some later parts or precursors of $PP$ and $PKP$. The frequently studied $PKP$ precursor starts in this distance range, but it precedes the $PKP$ arrival by up to 20 $s$, only whereas the $P_{\mathrm{diff}}$ coda starts more than 100 $s$ prior to $PKP$. At even larger distances ($\Delta > 130°$) we show that energy in the $P_{\mathrm{diff}}$ coda window can be detected even more than 150° away from the epicentre more than 100 $s$ before the core phase (the orange area).

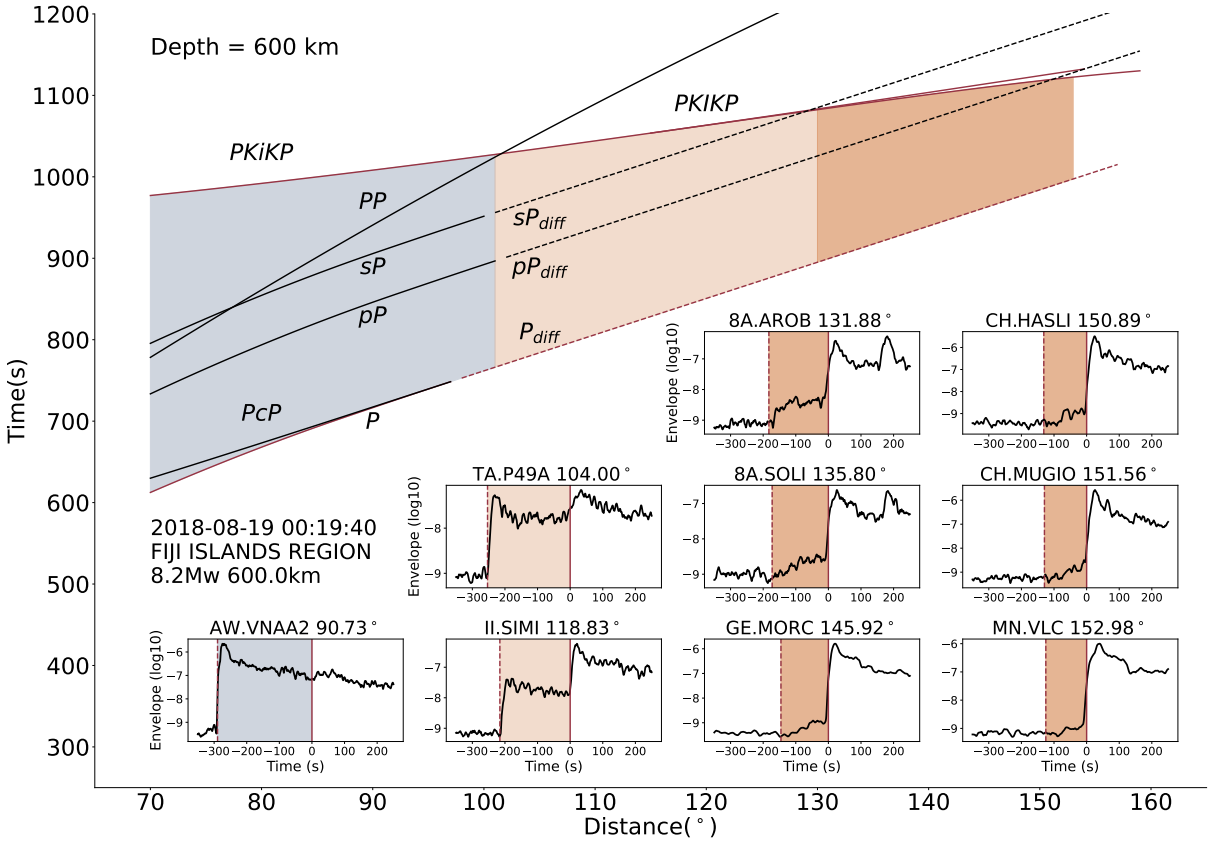



**Fig. 7.1:** The theoretical arrival times of seismic phases from a 600-km depth earthquake calculated with the $ak135$ model. The epicentral distance range is separated at 102° and 130° into three regions. For distances larger than 102° the core phase is the earliest arrival as it surpasses the $PP$ arrival. 130° is the maximum distance of previous observations of energy in the with $P_{\mathrm{diff}}$ coda window. The time window between arrival time of $P$ or $P_{\mathrm{diff}}$ and $PKiKP$ or $PKIKP$ phases which is investigated here is indicated by color shading. Insets show the logarithmic envelopes of single station vertical records from the 600 $km$ deep event that occurred with 8.2 $Mw$ on August 19, 2018, in the Fiji Islands region. The traces are all aligned to the earliest arriving core phase ($PKiKP$ or $PKIKP$). The color shadings indicates the $P$ coda or $P_{\mathrm{diff}}$ coda corresponding arrival time graph.

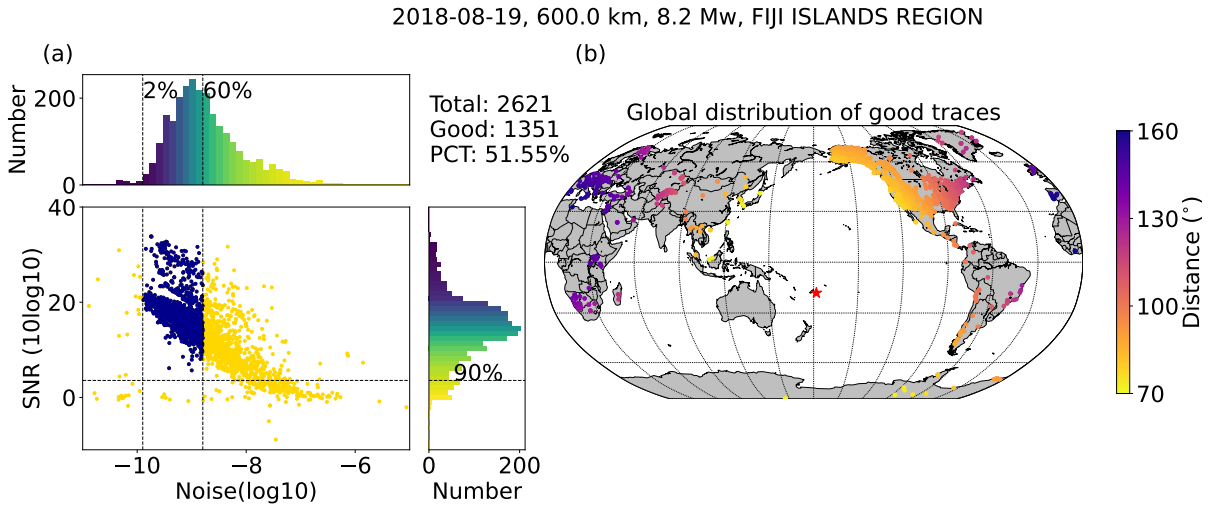



**Fig. 7.2:** (a) Scatterer plot between SNR and noise level of all available stations from the Fiji Island event (August 19, 2018). Each dot indicates one processed envelope in which the blue ones with the low noise (selected by the horizontal histogram) and the high SNR (selected by the vertical histogram) are selected for stacking. (b) The global distribution of good stations. Colorbar indicates the epicenter distance.

## 7.3 Stacking

Figure 7.1 shows that several orders of magnitude can separate the level of $P_{\mathrm{diff}}$ coda energy from the energy of the core phase. Global stacking can thus significantly improve the quality of the traces. However, the $P_{\mathrm{diff}}$ coda waves with their low energy are very sensitive to the noise level at each station. To improve detection of this tiny signal, both the absolute noise level and signal-to-noise ratio (SNR) are considered at each station individually. Noise level is measured in the time window of 200 to 50 $s$ before the theoretical arrival time of the $P$ or $P_{\mathrm{diff}}$ phase while the signal used in SNR is selected in the time window of 0 to 150 $s$ after $PKiKP$ or $PKIKP$ phase. Figure 7.2a shows a scatterer plot between SNR ($10\log_{10}(Signal/Noise)$) and noise level ($\log_{10} Noise$) of all available stations from the Fiji Island earthquake (August 19, 2018). Each dot indicates one record in which the blue ones with the low noise and the high SNR have been selected for stacking. Almost half of the traces are selected and the global distribution of stations with selected traces is shown in Figure 7.2b. The color bar indicates the the epicentral distance. The histograms in Figure 7.2a show that the selection is dominantly based on the noise level of the stations, while the SNR criterion only removes some spurious stations with extremely low noise level. For each event, we align the selected traces on the $PKiKP$ or $PKIKP$ arrival time and average the logarithmic envelopes in every distance bin (2°). We have stacked records from all 52 events individually using the same processing. Figure 7.3 shows the stacking results for two shallow, one middle-depth, and two deep events. Besides the deep event we mentioned before (August 19, 2018 Fiji Islands region earthquake, Figure 7.3a), another one is a 685.5 $km$ deep event with 7.9 $Mw$ on May 30, 2015, in the Bonin Islands, Japan region (Figure 7.3e). The shallow events respectively happened on January 23, 2018 in the Gulf of Alaska with magnitude 7.9 $Mw$ (Figure 7.3b) and on April 11, 2012 off the west coast of northern Sumatra with 8.6 $Mw$ (Figure 7.3c). The intermediate depth event occurred on May 26, 2019 in northern Peru with 8.0 $Mw$, 122.4 km depth (Figure 7.3d). The $P$ or $P_{\mathrm{diff}}$ coda is highlighted by the shading in the time window from the $P$ or $P_{\mathrm{diff}}$ phase to $PKiKP$ or $PKIKP$ phase using the noise level as the baseline with the colors corresponding to Figure 7.1.

From almost all of these events we detect the $P_{\mathrm{diff}}$ coda in the core shadow including very long distances (the orange area). However, signals of shallow events are strongly influenced on the source side and it is hard to identify $P_{\mathrm{diff}}$ coda (Figure 7.3b) unless events with huge magnitude are used (Figure 7.3c, 8.6

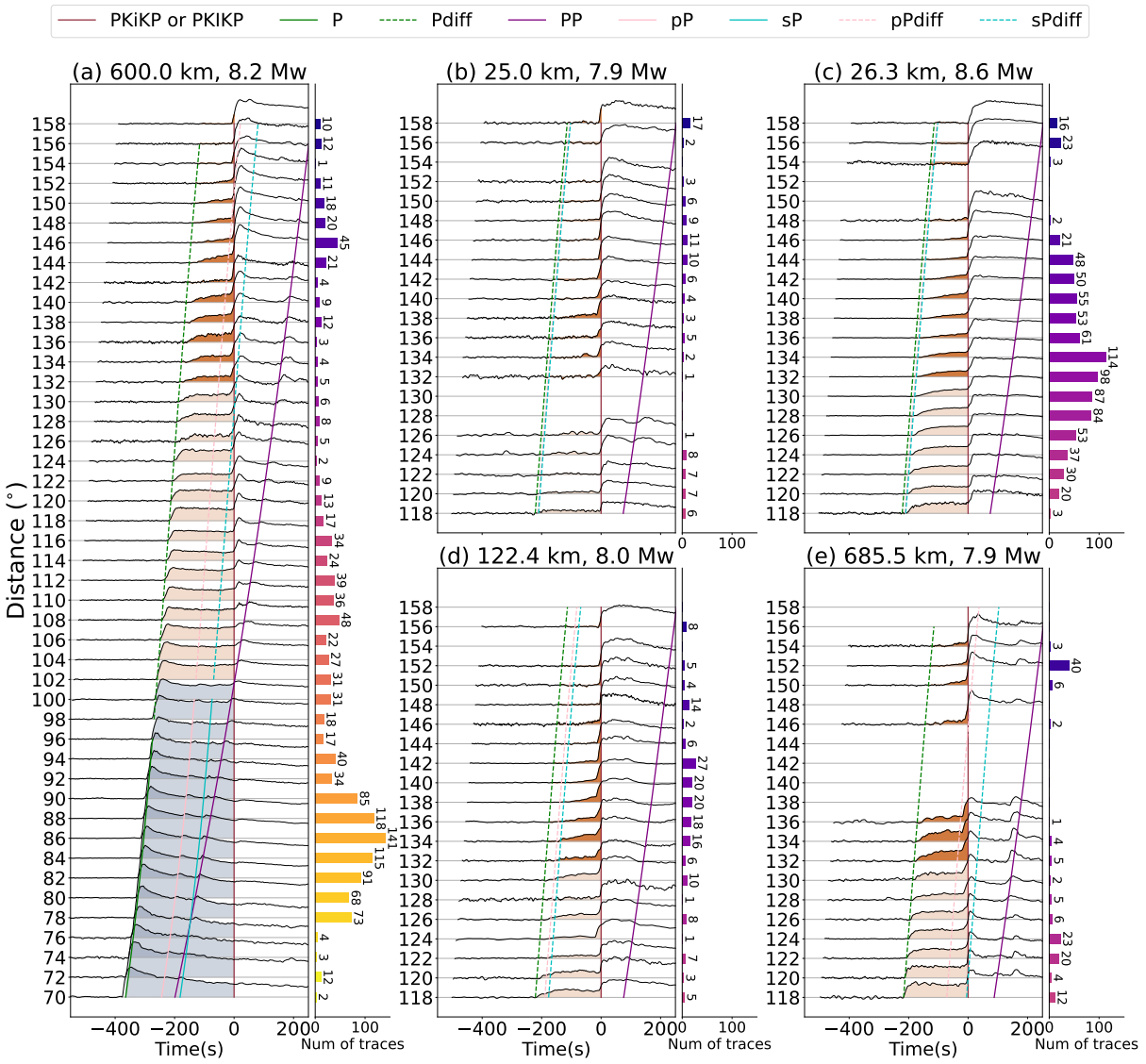



**Fig. 7.3:** The stacking results for two shallow, one intermediate depth, and two deep events. The $P$ or $P_{\text{diff}}$ coda is highlighted by shading in the time window from the $P$ or $P_{\text{diff}}$ phase to the $PKiKP$ or $PKIKP$ phase using the noise level as the baseline with colors corresponding to Figure 7.1. The histogram of the number of traces used in the stacking for each distance bin is shown on the right with color corresponding epicentral distances in Figure 7.2(b).

$Mw$). With the source depth increasing, the $P_{\text{diff}}$ coda becomes more clear even though the earthquake magnitude is less (Figure 7.3d), especially for very deep earthquake (Figure 7.3e). In conclusion, although we have investigated 52 events, the traces from the deep earthquakes with the huge magnitude constitute the best observations of $P_{\text{diff}}$ coda (Figure 7.3a).

This dependence on the absolute signal level together with the frequency-magnitude distribution means that a stack of different earthquakes is unlikely to be better than the record of the best event alone. It is thus not reasonable to stack the traces from different events that have different depths and magnitudes. Another reason why we use the events individually is that the shape of traces from the shallow and deep events are different due to the arriving times of the $pP$, $sP$ and $PP$ phases, especially at shorter distances. In this study, we focus on the 600 $km$ deep Fiji Island event (Figure 7.3a) to study the $P_{\text{diff}}$ coda.

## 7.4 Modeling

### 7.4.1 Current Models of Whole Earth Scattering

The deterministic structure of the Earth has been investigated globally and regionally in many studies. Seismic waveforms e.g. $P$- or $S$-waves or surface waves, inform about the elastics parameters of the Earth since the elastic structure determines the seismic wave propagation including travel times and amplitudes. Especially for the synthetic reconstruction of global seismograms, 1D spherically symmetric velocity models of the Earth have been proposed based on the main teleseismic phases e.g. PREM (Dziewonski & Anderson, 1981), iasp91 (Kennet, 1991) or ak135-F (Kennett et al., 1995; Montagner & Kennett, 1996). More precise 3D Earth models have also been derived more recently, like SPiRaL (Simmons et al., 2021). Compared to the deterministic large-scale structure of the Earth that can be investigated with the major seismic phases, the distribution of small-scale heterogeneity is more difficult to study since the resulting scattered energy is a secondary observation in global seismograms. The efficiency of small-scale heterogeneity to scatter seismic waves increase with frequency such that scattered waves are generally studied at frequencies above the range used for the investigation of ballistic waves. However, the increase of intrinsic attenuation with frequency limits the interesting range towards very high frequencies. Scattered coda waves or precursory arrivals are therefore often studied in the frequency range of a few hertz.

This small-scale structure in the Earth can extend in the crust and mantle or the inner core. In the liquid outer core small-scale heterogeneity is assumed to be erased by convection. The heterogeneity in the shallow Earth has a strong influence on the scattered seismic wavefield. The tomography for the crust shows strong lateral variations of scattering and intrinsic attenuation (Sens-Schönfelder et al., 2009; Calvet et al., 2013; Mayor et al., 2016), especially in volcanoes (De Siena et al., 2016). A model of scattering properties in the crust with $\varepsilon = 2.9\%$ and $a = 0.27\ km$ was derived by Gaebler et al. (2015) using teleseismic $P$-wave coda as well as coda from local events. The deep Earth scattering is mostly investigated using the $PKP$ precursors. The heterogeneity existing near the CMB that causes the single-scattering energy to arrive earlier than $PKP$ at the surface explains the observation of precursors to $PKP$ (Cleary & Haddon, 1972; Haddon & Cleary, 1974). Later, scattering in the whole mantle was invoked to explain the $PKP$ precursor. Hedlin et al. (1997) proposed heterogeneity in the mantle characterized by $\varepsilon = 1\%$ and $a = 8\ km$ while Margerin & Nolet (2003) and Mancinelli & Shearer (2013) proposed weaker heterogeneity with $\varepsilon = 0.1\%$. Whole mantle scattering was also used to explain the $P_{\mathrm{diff}}$ coda with $\varepsilon = 1\%$ and $a = 2\ km$ (Earle & Shearer, 2001) and a two-layers model with $\varepsilon = 3 \sim 4\%$ and $a = 4\ km$ in the upper mantle and $\varepsilon = 0.5\%$ and $a = 8\ km$ in the lower mantle to fit the observations of $P$ coda (Shearer & Earle, 2004).

In this paper, we consider the heterogeneity model developed by Bentham et al. (2017) to which we refer as *BRT2017*. It includes scattering in the lithosphere and three-layers in the mantle as shown in the left panel of Figure 7.4. Bentham et al. (2017) used a grid search including previous models to obtain the BRT2017 model which is adjusted to fit the observation of globally stacked $PP$ precursors in the distance range form 70° to 120°. Inner core scattering is excluded from this model.

### 7.4.2 Monte-Carlo Simulation

For modeling of the global seismic scattering, the Monte-Carlo simulation method which employs many particles to represent the energy propagation is used to simulate the Earthquake in a 1D spherically symmetric model with the main frequency 1 Hz. (Sens-Schönfelder et al., 2021). The reflection and refraction obey the Snell's law. Scattered waves are modeled using the radiative transfer equations in 3-

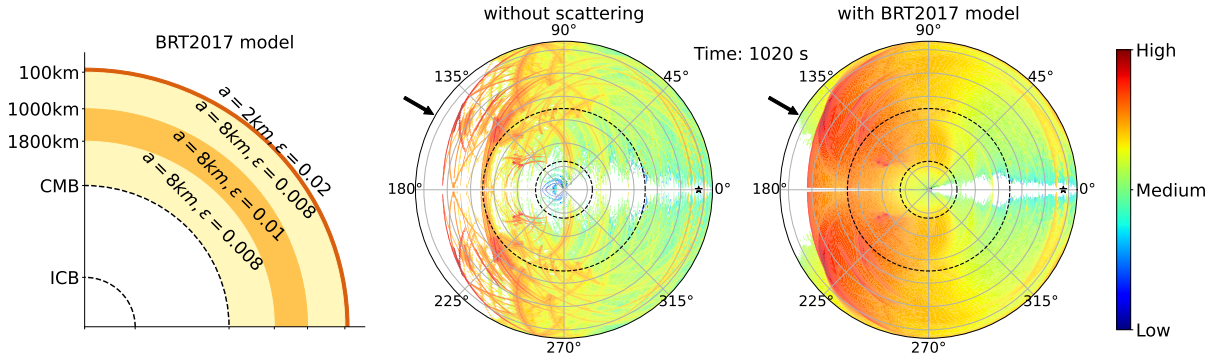



**Fig. 7.4:** Left panel: the heterogeneity model used in this study - BRT2017 model: 0-100km, a = 2 km, $\varepsilon = 0.02$; 100-1000km, a = 8 km, $\varepsilon = 0.008$; 1000-1800km, a = 8 km, $\varepsilon = 0.01$; 1800km-CMB, a = 8 km, $\varepsilon = 0.008$. Right panels: The cross sections of the Monte-Carlo simulation without scattering and with 3D multiple nonisotropic scattering using the BRT2017 model at lapse time 1020s. The color indicates the logarithm of the seismic energy and the arrow indicates a location at epicentral distance of 150°.

D that describe multiple nonisotropic scattering in an elastic medium (Sens-Schönfelder et al., 2009). The velocity, density and intrinsic attenuation models use the ak135-F spherical average model. The intrinsic quality factors described in the ak135-F model state $Q_\kappa$ and $Q_\mu$. The radiative transfer equations require the P-wave $Q_P$ and S-wave $Q_S$ quality factors. They have the relationships (Shearer, 2009):

$$Q_P^{-1} = LQ_\mu^{-1} + (1-L)Q_\kappa^{-1} \tag{7.1}$$

$$Q_S = Q_\mu \tag{7.2}$$

where $L = (4/3)(\beta/\alpha)^2$ and in the outer core $Q_P = Q_\kappa$. $\alpha$ and $\beta$ are the velocities of P-wave and S-wave, respectively.

The middle and right panels of Figure 7.4, respectively, show the snapshots of the energy field from Monte-Carlo simulation without scattering on the one hand and with 3D multiple nonisotropic scattering using the heterogeneity model - the BRT2017 model on the other hand. Although the simulation is conducted in 3D space, the model and recording of the wavefield are spherically symmetric. The source is 600 km deep located at the right center of the cross section through the Earth (the black star). The lapse time in Figure 7.4 is 1020 s. The simulation in the heterogeneous model shows a smoother energy distribution compared to the homogeneous model. The energy filling the space between the main seismic phases generated by scattering off the heterogeneity. The arrows in the figure indicates the location at epicenter distance of 150°, where some energy has already arrived at the illustrated 1020$s$ lapse time in the BRT2017 while no energy has reached the receiver in the model without scattering, yet.

The Monte-Carlo simulation uses the point-like source in space and time. In order to compare with the observation, we convolved the envelope with a Gaussian source time function whose center of time is 15 $s$ and the standard deviation is 1.95 $s$. Figure 7.5 compares the observation form the Fiji event (solid black curves) and the synthetics with scattering in the BRT2017 model (the solid red curves) and without any scattering (the thin gray curves). The simulation results are normalized by the maximum value of each trace before the time -50 $s$. Zero time is aligned to the arrival time of $PKiKP$ or $PKIKP$. To ease the comparison of the misfits between observations and simulations in the BRT2017 model are filled with blue shading in the $P_{\text{diff}}$ time window ranging from the arrival time of the $P$ or $P_{\text{diff}}$ phase to 20 $s$ after the $PKiKP$ or $PKIKP$ phase.

Generally the modeled envelopes fit the observations well. Considering the fact that the BRT2017 model was derived from $PP$ precursors especially, the fit of the $P_{\text{diff}}$ coda at large distances is remarkable. A systematic misfit can be observed at about -50$s$ at intermediate distances of about 96° $\sim$ 112°. We

discuss this issue in Section 7.7. For the $P_{\mathrm{diff}}$ coda, the simulation results from the BRT2017 model matched the observation very well and one can see that the first arrival of energy in the simulation without scattering becomes later and later with increasing distance until the core phase $PKIKP$ is the first arriving phase. There is no $P_{\mathrm{diff}}$ coda in the model without heterogeneity.

## 7.5 Origin of $P_{\mathrm{diff}}$ Coda

In accordance with previous work we refer to the energy arriving more than 100 s before $PKIKP$ at long distances as $P_{\mathrm{diff}}$ *coda* since it arrives after the theoretical arrival of the $P_{\mathrm{diff}}$ phase. This terminology avoids ambiguity with the term $PKP$ *precursor* used for the energy arriving a few seconds before $PKP$ in the distance range $110° < \Delta < 140°$. Early interpretations of $P_{\mathrm{diff}}$ coda believed it originates from multiple scattering near the CMB in the D" layer (Bataille et al., 1990). This process would be rather similar to diffraction that generates $P_{\mathrm{diff}}$, but would be more effective at high frequencies (Bataille & Lund, 1996). With the scattering of the whole mantle proposed to interpret the $PKP$ precursors (Hedlin et al., 1997), Earle & Shearer (2001) presented the single-scattering modes like $P$-to-$P$, $P_{\mathrm{diff}}$-to-$P$ or $P$-to-$P_{\mathrm{diff}}$ happening in the whole mantle to generate the $P_{\mathrm{diff}}$ coda. However, Rost et al. (2006) reported that $P_{\mathrm{diff}}$ is not detectable at large distance in the short period.

These differing models pose the question whether $P_{\mathrm{diff}}$ coda at long distances really originates from scattering of $P_{\mathrm{diff}}$.

The simulation of scattered waves with the BRT2017 model matches the observation of $P_{\mathrm{diff}}$ coda very well. To figure out which part of the Earth contributes to the scattered energy in the different time windows, we design many single-layer heterogeneity models based on the BRT2017 model. We divide the mantle into layers of 200 km thickness overlain by a 100 km thick lithosphere layer. The simulation is conducted for every single-layer heterogeneity model in which 3D multiple non-isotropic scattering can only happen in this layer and whose scattering parameters correspond to the one of the BRT2017 model at the respective depth while there is no scattering in other layers. The results are shown in Figure 7.6. Each color curve indicates the simulation result from one single-layer heterogeneity model. The observation is shown by the black curve and the simulation without any scattering and with scattering in the full BRT2017 model (already shown in Figure 7.5) are shown by the red and gray curves, respectively.

In Figure 7.6 at distances before the core shadow e.g. 90°, the $P$ coda is dominated by scattering at lithospheric heterogeneity (darkest blue curve). At the distance 100°, $P_{\mathrm{diff}}$ and its coda appears instead of $P$. The Monte Carlo simulation does not reproduce diffraction but still matches the observation of what is usually referred as $P_{\mathrm{diff}}$ in the present frequency band above 1Hz. The energy of $P_{\mathrm{diff}}$ mostly originates from scattering in the lower mantle with an onset of all curves (about -250 s) at the time of the theoretical $P_{\mathrm{diff}}$ arrival resulting in a strong peak reproducing the observation at the $P_{\mathrm{diff}}$ arrival time and its early coda. Note that there is no energy at this time in the simulation without scattering (the gray curve). With the distance increasing to 110° even 120°, the peaks of the different layers separate and arrive at different times and the later part of the coda energy stems from the shallower layers such that all layers of the mantle contribute to $P_{\mathrm{diff}}$ coda but with varying contributions at different travel times. As consequence the peak at the $P_{\mathrm{diff}}$ arrival time decreases in amplitude and the decay of its coda is reduces towards larger distances until the energy level remains almost stable. This confirms Bataille & Lund (1996)'s recognition that scattering is the main effect in the generation of high frequency energy following the theoretical arrival time of $P_{\mathrm{diff}}$. It explains the observation of peaks decreasing with increasing distance (Bataille & Lund, 1996; Earle & Shearer, 2001). What is usually called $P_{\mathrm{diff}}$ at high frequencies is actually scattered energy mostly from the lowermost mantle and should better be called $P_{\mathrm{scatt}}$. When the distance is larger than 130°, it is hard to observe the so-called $P_{\mathrm{diff}}$ arrival. Instead there is a gradual increase of energy towards the $P_{\mathrm{diff}}$ coda. Scattering near the CMB which contributes

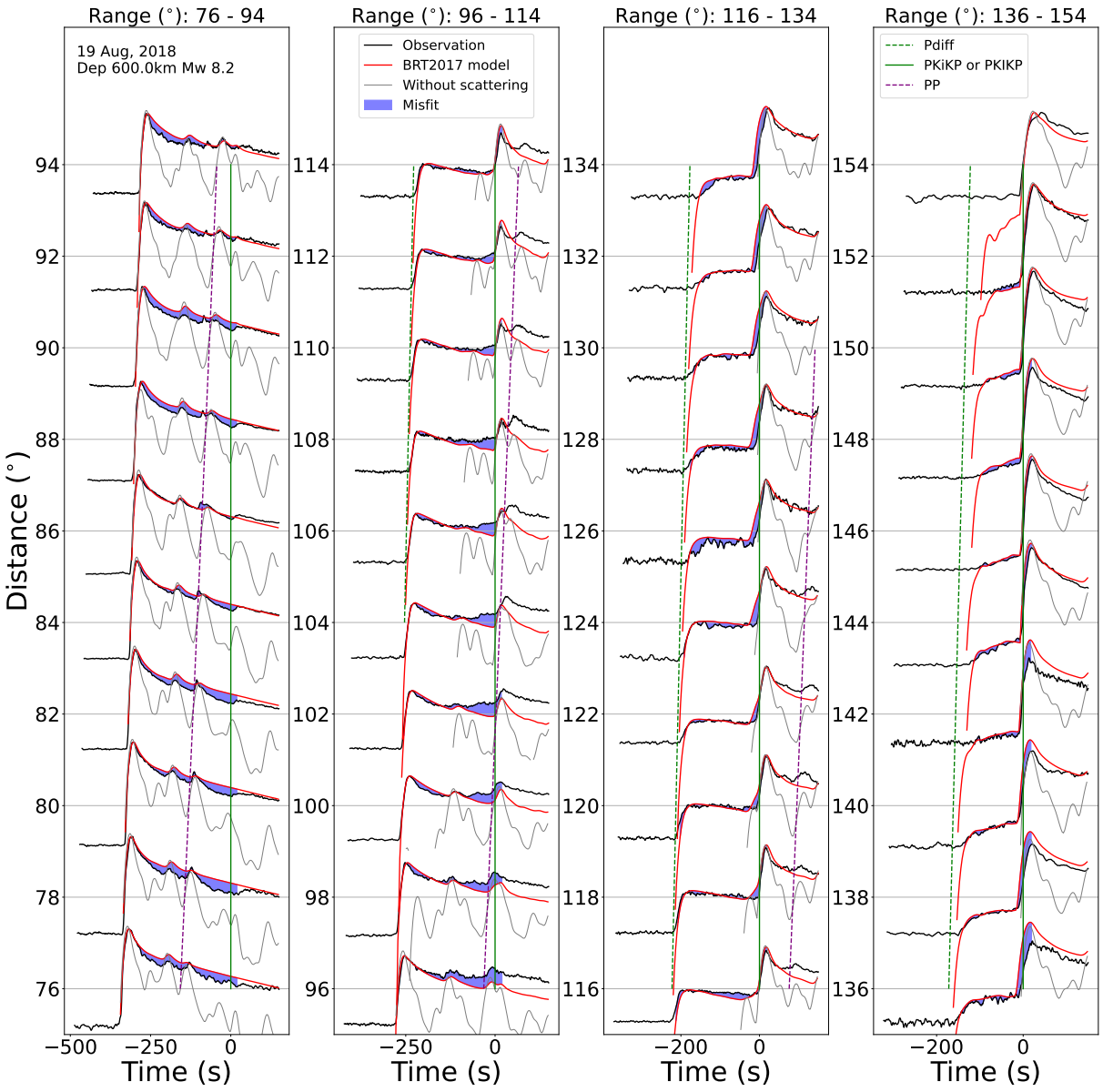



**Fig. 7.5:** Observed (solid black curves) and the synthetic seismogram envelopes with heterogeneity in the BRT2017 model (solid red curves) or without any heterogeneity (thin gray curves). As visual aid, the misfits between observation and simulation in the BRT2017 model are filled with the blue shading in the $P_{\mathrm{diff}}$ coda time window between the arrival time of the P or $P_{\mathrm{diff}}$ phase and 20 $s$ following the PKiKP or PKIKP arrival. Shading is only applied above the noise level.

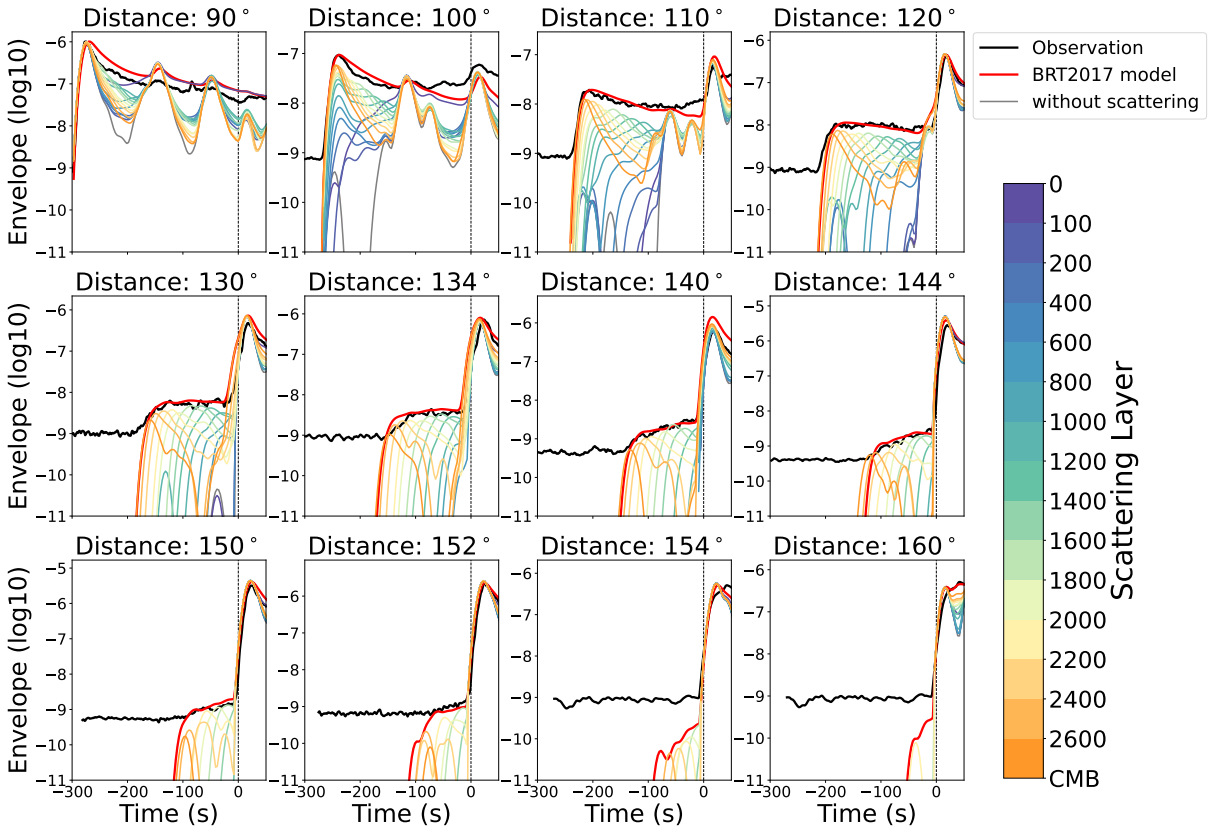



**Fig. 7.6:** The simulation results from each single-layer heterogeneity model (colorful curves corresponding to the colorbar that represents the depth of each layer) based on the BRT2017 model compared with the observation (black curve) and the simulation without any scattering (gray curve) and scattering in the BRT2017 model (red curve).

to the earliest coda energy decreases (orange curves). The increase shortly before the peak of the curve at the $PKIKP$ arrival at intermediate distances originates from the deeper layer, as expected for the $PKP$ precursor. At 160°, the $P_{\mathrm{diff}}$ coda can only originate from scattering in the middle mantle about 2000 km deep (yellow curve) as deeper scattering will not propagate energy around the core to these distances. Such energy is not observed in the seismograms since its energy is below the level of the noise even at the best stations. In conclusion, $P_{\mathrm{diff}}$ and $P_{\mathrm{diff}}$ coda at high frequencies are intrinsically the same energy scattered from the mantle at different depth. Different depths contribute to different arriving times generating the variable shapes of $P_{\mathrm{diff}}$ and $P_{\mathrm{diff}}$ coda at different distances. No diffraction is involved in the propagation of this energy.

To illustrate the regions in which scattering can occur to contribute to the $P_{\mathrm{diff}}$ coda, Figure 7.7 shows possible travel paths of $P_{\mathrm{diff}}$ coda for nine different distances. The colored areas indicate potential locations of single scattering ($P * P$) in the great circle plane with color encoding the resulting arrival time at the station. The time is aligned to the arrival time of the $PKIKP$ phase at each distance so the red area is the scattering region that generates the $P_{\mathrm{diff}}$ coda while scattering in the blue regions leads to arrivals later than $PKIKP$. As the distance increases, the boundary between red and blue regions shifts from shallow depth towards the deep Earth indicating that the large distance observations of $P_{\mathrm{diff}}$ rely on mid to lower mantle scattering.

Meanwhile, a shadow zone above the CMB starts to appear in the single-scattering region. Scattering close to the CMB (e.g. in D") cannot contribute to the $P_{\mathrm{diff}}$ coda at larger than $\approx 150°$ any more. Consequently, the region in which single scattering can generate $P_{\mathrm{diff}}$ coda narrows towards larger distances. The upper limit lowers to maintain short travel times and the lower limit rises for the

energy to pass around the core as illustrated in the bottom left panel of Figure 7.7. At a distance beyond 168° no $P_{\text{diff}}$ coda can be generated by single scattering ($P * P$) any more as confirmed by the observations in Figure 7.6 We calculate the earliest possible time of this single-scattering at each distance and compare it to the theoretical arrival time of $P_{\text{diff}}$ and the observed $P_{\text{diff}}$ coda in the right bottom of Figure 7.7. The $P_{\text{diff}}$ phase lags behind the single-scattering onset time curve. This is to be expected for large distances at which the region generating the earliest arrival becomes very narrow with signal levels below the station noise leading to a gradual increase of energy while the computation considers any possibility for the energy to reach the receiver . However, the observations follow this curve more closely than the theoretical $P_{\text{diff}}$ arrival time confirming that there is no relation of the observation to actual $P_{\text{diff}}$ propagation.

## 7.6 Sensitivity Kernels

The comparison of the data with the BRT2017 model in Figures 7.5 and 7.6 shows a good agreement between modeled and observed envelopes. Inferring more detailed information about the mantle heterogeneity is possible with the tools developed in Chapter 5 and Chapter 6.

The key to the application of AET is the calculation of the misfit kernels that are in turn derived from the scattering sensitivity kernels which describe the relationships between changes in the strength of fluctuation and seismogram envelopes as observable. The calculation of the scattering sensitivity kernels requires the specific energy density $E_{XY}(\mathbf{r}', t, \mathbf{n}; \mathbf{r_0})$ that contains the information about propagation direction. For now, the memory requirements of the 3D specific energy density for the whole Earth with so long lapse time is too huge so that we can only record the wavefield of the energy density $E_{XY}(\mathbf{r}', t; \mathbf{r_0})$. Luckily, the changes in scattering in the Earth are basically the creation of new propagation paths that causes the distortion of the coda waves, which can be quantified by the decorrelation sensitivity kernel (Margerin et al., 2016). The $P_{\text{diff}}$ coda arrives so early at teleseismic distances that propagation as shear energy can be excluded at that time since the S-wave energy has not enough time to arrive at the surface. Therefore, the scattering mode can be simplified to multiple $P$-to-$P$ scattering excluding $S$ propagation. The decorrelation sensitivity kernels proposed in Chapter 4 now are simplified as in the acoustic case. If scattering is further assumed to be isotropic, the kernel can be rewritten as (Margerin et al., 2016):

$$ {}^{\varepsilon}K^{dc}_{PP}(\mathbf{r}', t; \mathbf{r}, \mathbf{r_0}) = \int_0^t \frac{E_{PP}(\mathbf{r}', t - t'; \mathbf{r}) E_{PP}(\mathbf{r}', t'; \mathbf{r_0})}{E_{PP}(\mathbf{r}, t; \mathbf{r_0})} dt' \tag{7.3} $$

Figure 7.8 shows the decorrelation sensitivity kernels plotted with logarithmic values. The wavefield $E_{PP}(\mathbf{r}', t; \mathbf{r_0})$ is simulated with the BRT2017 model in 3D using the elastic multiple nonisotropic scattering shown in the right panel of Fig. 7.4. The time is aligned to $PKIKP$ arrival time at each distance. The decorrelation sensitivity kernels show the influences from changes in the strength of the fluctuation on the decorrelation of the seismograms. In fact, they also reflect how much of the energy arriving at $\mathbf{r}'$ from the source $\mathbf{r_0}$ can reach the station $\mathbf{r}$ with the lapse time $t$. Not like the assumption of single scattering used in Figure 7.7, multiple scattering implemented in the BRT2017 model allows for scattering across the entire mantle. As the distance increases and for earlier time, the high values of sensitivity kernels approach the CMB and the scattering area become smaller. If the simulation of the wavefield used for calculation of sensitivity kernels is without scattering, only single scattering at $\mathbf{r}'$ can generate energy at the receiver. This situation is illustrated in Figure 7.9 that is rather similar to Figure 7.7. Both of them show a cone shape in the mantle above the CMB that are the possible areas for single scattering.

Figure 7.8 and Figure 7.9 only show the cross-sections through the great-circle plane of the source

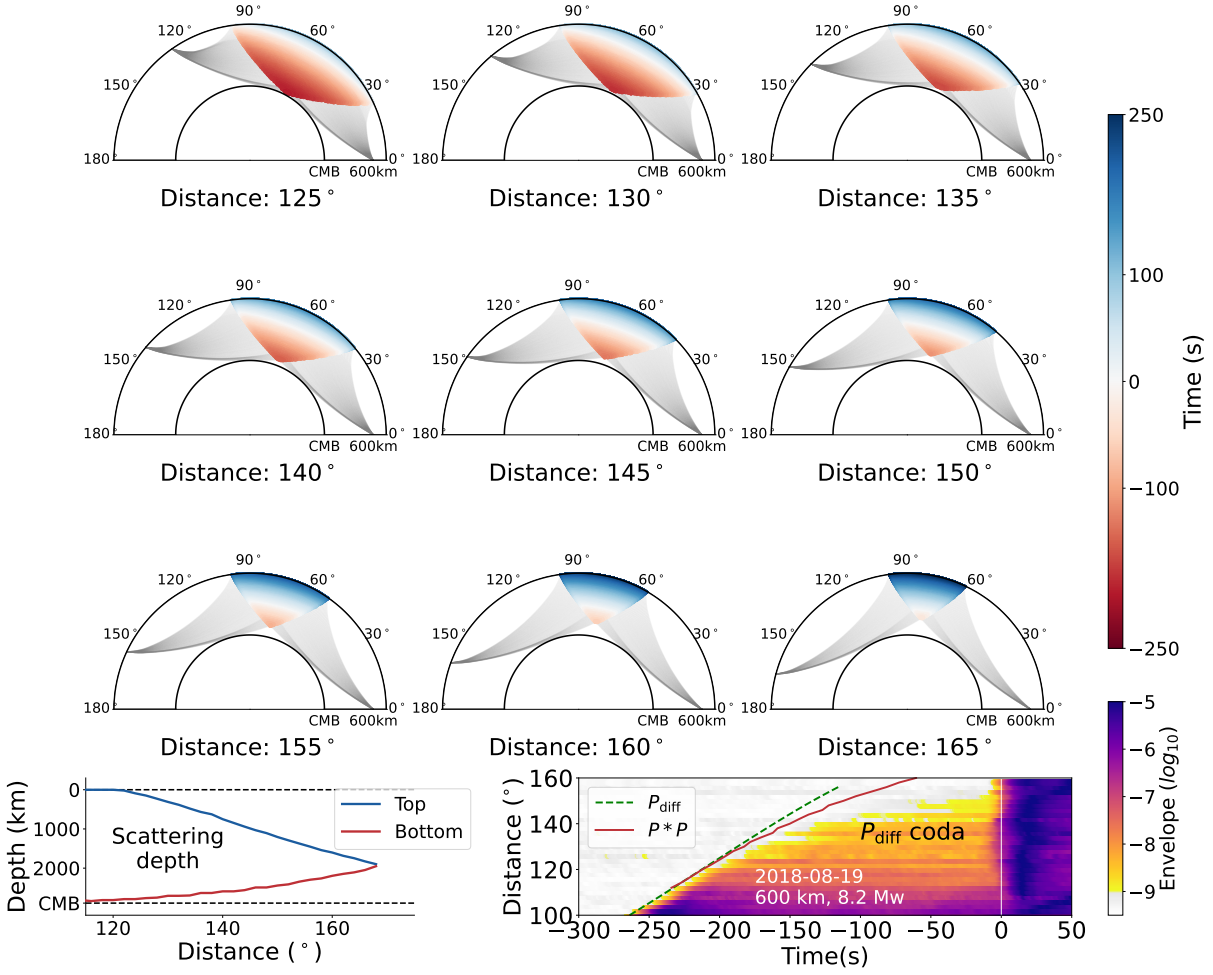



**Fig. 7.7:** Top nine panels: the cross sections (the great circle plane) of all possible single-scattering positions (the colored areas) for different epicenter distances. The color indicates the travel times resulting from single scattering at different locations. Times are aligned to the arrival time of the $PKIKP$ phase at each distance. Red represents the scatterers for $P_{\mathrm{diff}}$ coda arriving before $PKIKP$. Bottom left: The top of the red region (blue curve) and the bottom of possible single-scattering region (red curve) approach each other in the middle mantle at about 2000 km depth for an epicentral distances of 168°. Bottom right: the earliest possible time (onset time) of single scattering energy (red curve) at each distance compared with the theoretical arrival time of $P_{\mathrm{diff}}$ (dashed green curve) and the observed $P_{\mathrm{diff}}$ coda energy (the color filling in which the gray area indicates logarithmic energy below -9 ).

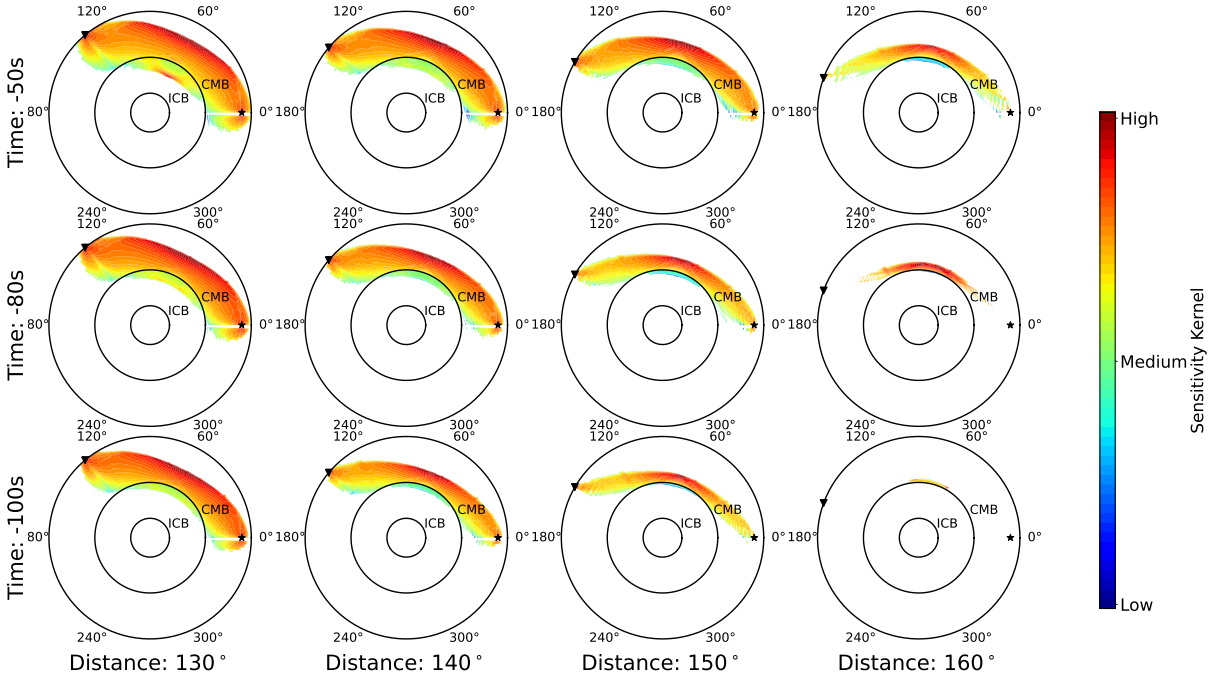



**Fig. 7.8:** Cross sections (great circle plane, $\Phi$=0°) of the decorrelation sensitivity kernels. The scattering is assumed to be isotropic. The wavefield is modeled in the BRT2017 model using the radiative transfer equations that are the 3-D elastic multiple nonisotropic scattering shown in the right panel of Fig. 7.4. Note the color bar indicates the logarithm of sensitivity kernels that are individually normalized by the maximum value of each panel. The time is aligned to $PKIKP$ arrival time of each distance.

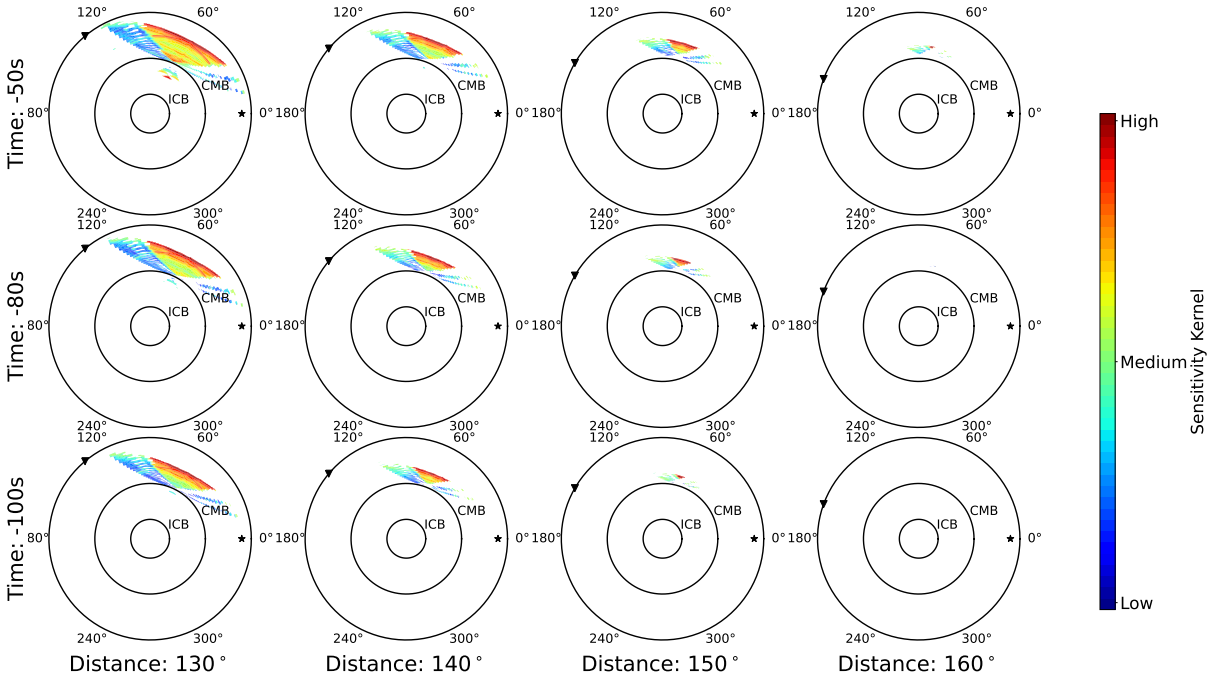



**Fig. 7.9:** Same as Fig. 7.8, but the wavefield is modeled in the homogeneous model.

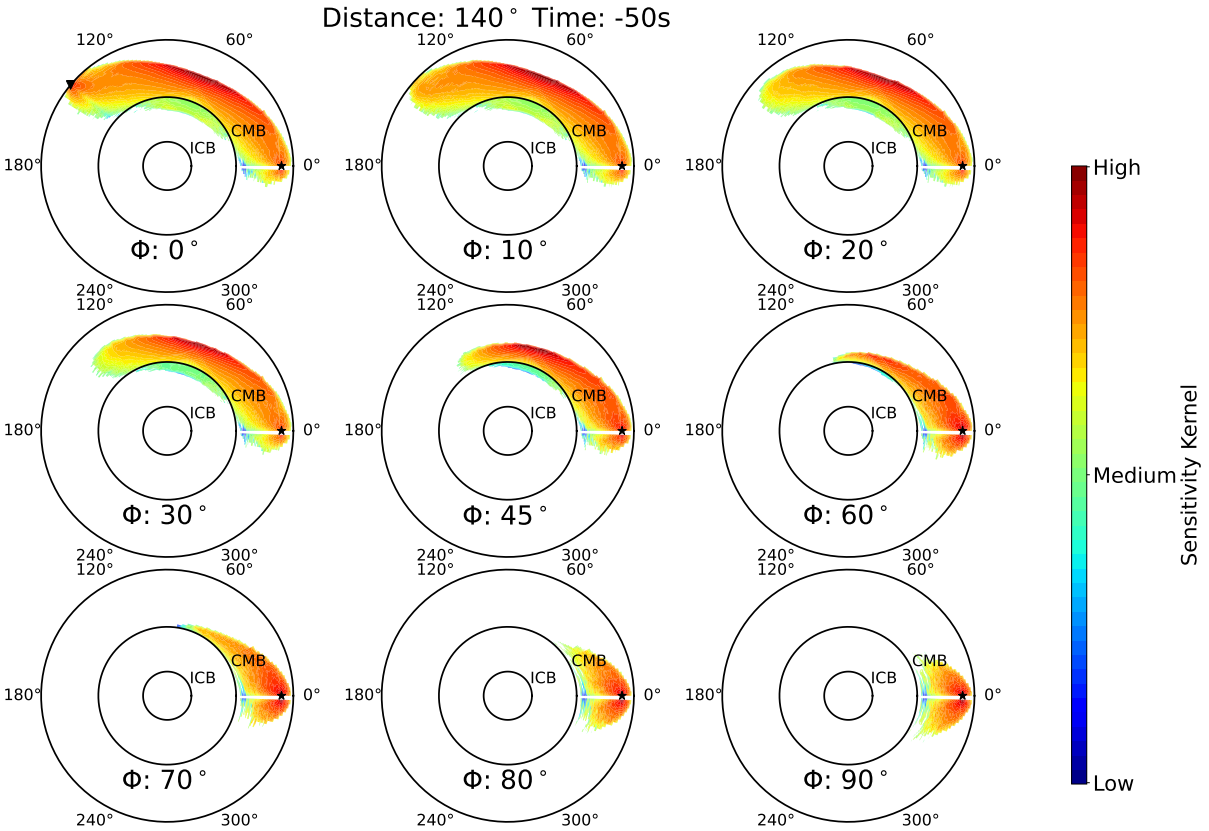



**Fig. 7.10:** Cross sections of decorrelation sensitivity kernels with different azimuth Φ tilted from the great circle plane. The wavefield is modeled in the BRT2017 model. The first panel is the same as the first one in the second column shown in the Fig. 7.8. Note the color bar indicates the logarithm of sensitivity kernels that are individually normalized by the maximum value of each panel. The time -50 s is aligned to $PKIKP$ arrival time at 140° .

and station. However the kernels are 3D features and Figure 7.10 and Figure 7.11 show cross sections planes that are tilted from the great circle plane by azimuth Φ . The fastest travel paths are definitely on the great-circle cross-section that includes source and station. With the azimuth Φ from 0° (the source-station plane) to 90° (perpendicular to the source-station plane) in Figure 7.10, the high values gradually concentrate at the source since the plane of this cross-section becomes farther away from the station. The effective volume of the sensitivity kernels can be imagined as a clam shell lying bent above the CMB. For single scattering (Figure 7.11), there is less possibility for scattering occurring farther off the source-station plane and the effective volume gradually disappears near the CMB.

## 7.7 Discussion

We show observations of scattered seismic energy arriving several tens of seconds prior to the core phase at distances as large as 150°. We present Monte Carlo simulation for the global seismic energy propagation in a heterogeneity model derived from $PP$ precursors which match the $P_{\mathrm{diff}}$ coda very well in Figure 7.5. However, the misfits during 96° to 112° are likely related to precursors of $PP$ from reflectiosn below the surface which are not perfectly modelled. The 1D spherically symmetric model used here is ak135-F in which the crust is considered continental. The intrinsic attenuation of the crust has strong lateral variations and the thicknesses between continental and oceanic crust are different which can influence the $PP$ travel times as well as the shape of precursory signals. The stations shown in Figure 7.2 are mostly distributed on the continents so that reflection points of the $PP$ wave are probably located in oceanic lithosphere. So we consider the misfit illustrated in Figure 7.5 to result from lateral differences that are

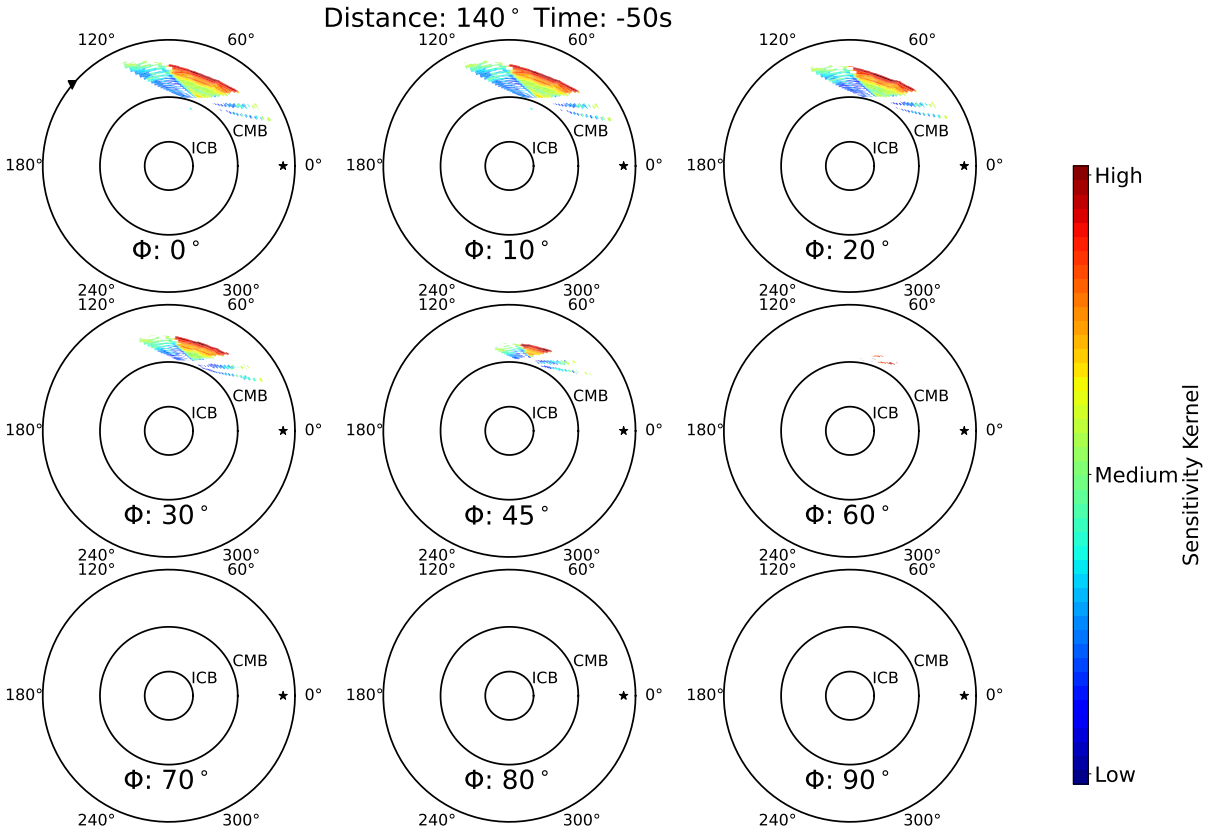



**Fig. 7.11:** See the caption in Fig. 7.10, but the wavefield is modeled in the homogeneous model.

not captured in the 1D heterogeneity model BRT2017. Our observation targets, e.g. $P_{\text{diff}}$ and $PKP$ that are mostly affected by the receiver site determine the choice of the model as the continental crust. Fortunately, the discussion of $P_{\text{diff}}$ coda mostly refers to arrival times well before $PP$ such that there is only marginal influence of the oceanic lithosphere on our conclusion.

Theoretically, $pP$ can only propagate to 100° and then develops into $pP_{\text{diff}}$ so the simulation without scattering and diffraction cannot generate $pP$ at larger distances. The earliest tiny phases shown in Figure 7.5 by the gray curves at distances larger than 100° are not $pP$ or $pP_{\text{diff}}$. These arrivals may come from the reflections on the shallow Earth which are mixed with $pP$ at short distances and can bypass the core shadow and propagating to longer distances.

Our simulation only uses the P-energy as source since for the $P_{\text{diff}}$ coda the time is too early for S-energy so that $S$-to-$P$ scattering makes a very minor contribution to the energy in the time distance window considered here.

## 7.8 Conlusions

We have for the first time observed the $P_{\text{diff}}$ coda at high frequencies (1∼2 Hz) at a very large epicenter distance (even more than 150°). The stacking of global earthquakes shows this signal is best observed for earthquakes with huge magnitude and deep source. The Monte Carlo simulation using 3D non-isotropic multiple scattering in a 1D spherically symmetric heterogeneity model - BRT2017 (Bentham et al., 2017) that has a 4-layers structure from the lithosphere to CMB is compared with the global stacked high-SNR $P_{\text{diff}}$ coda from a deep event (8.2 Mw, 600 km, August 19, 2018 Fiji Islands region). The synthetic data coincides with the observation very well in the time window between arrival times of the $P_{\text{diff}}$ and $PKIKP$, especially compared with the simulation without scattering which has a huge discrepancy to

the observations. The good fit supports the existent of small-scale heterogeneity in the whole mantle. To figure out the origin of $P_{\mathrm{diff}}$ coda, we separate this model into multiple single-layer models to simulate the scattered energy contributed from each layer. We believe the high-frequency $P_{\mathrm{diff}}$ at long distances ($\geq$100°) is intrinsically generated by the same process as the $P_{\mathrm{diff}}$ coda. The lowermost-mantle scattering contributes to the earliest part of $P_{\mathrm{diff}}$ coda that used to be recognized as $P_{\mathrm{diff}}$ when the distance is not too large. With the scattering layer rising, the energy arrives coherently one after another which causes the $P_{\mathrm{diff}}$ coda to be devoid of the typical coda decrease. The depth range contributing to the generation of the $P_{\mathrm{diff}}$ coda narrows towards larger epicentral epicenter distance ($\geq$130°). Both, the lowermost mantle and uppermost mantle may not contribute to $P_{\mathrm{diff}}$ coda at long distances. We use single-scattering to demonstrate all possible areas of scattering and their traveltimes compared with $PKIKP$. The top layer of the region generating the $P_{\mathrm{diff}}$ coda and the bottom of possible single-scattering region approach each other in the middle mantle at about 2000 km depth at the epicenter distance of 168°. The estimated earliest time of single-scattering arrivals does not perfectly agree with $P_{\mathrm{diff}}$ coda onset at large distances most likely due to the gradual increase of coda energy and the low signal level that does not increase above the noise level immediately after the onset time.

# Chapter 8

# Conclusion & Outlook

## 8.1 Conclusion

This thesis has proposed a new probabilistic seismic tomography method with scattered waves. Its target is to image the small-scale heterogeneity structure of the Earth that affects the scattering and intrinsic attenuation properties. I call this tomography method, Adjoint Envelope Tomography (AET).

The small-scale structures of the Earth below the wavelength are impossible to characterize by deterministic tomography. Chapter 2 briefly introduced the concept of random media that are statistically described by the fluctuation strength $\varepsilon$ and correlation length $a$ which respectively represent the perturbation range of the velocity and the length scale of heterogeneity. The type of random media is determined by the autocorrelation function (ACF). With random media, wave scattering is solved using the elastic radiative transfer theory. The random velocity and density fluctuations of the medium are assumed to follow an exponential ACF and the scattering is nonisotropic. Based on the above theories, I extend the elastic radiative transfer equations to spatially variable heterogeneity such that the fluctuation strength $\varepsilon(\mathbf{r})$ and intrinsic quality factor $Q^{-1}(\mathbf{r})$ become functions of location r. The simplification of the expressions in the acoustic approximation is also given.

The Monte Carlo method is used to numerically solve the radiative transfer equations in Chapter 3. In order to simulate energy transport in the presence of spatially variable fluctuation strength and intrinsic attenuation, I separate the effects of scattering and intrinsic attenuation in two models with individual fluctuation strength and quality factor anomalies. Two models with spatially variable scattering and intrinsic attenuation are compared with the statistically homogeneous model. The effects of stronger scattering and attenuation can be clearly observed in the two anomaly models.

The specific energy density of the simulated wavefield is recorded for further development. The specific energy density $E_{YX}(\mathbf{r}, \mathbf{n}, t)$ describes the angularly resolved energy density at position $\mathbf{r}$ at time $t$ with the propagation direction $\mathbf{n}$. In the elastic case, the mode of excitation $X$ and recording $Y$ can either be P- or S-wave. This quantity provides complete information about the energy transfer in an elastic medium with spatially variable randomness and intrinsic attenuation. Following Margerin (2017), the reciprocity theory of Green's function in 2D elastic radiative transfer theory is derived which is an essential step in the further application of the adjoint theory. The reciprocity theorem in the elastic radiative transfer theory is similar to FWI but the simulation of the adjoint wavefield requires an exchange of the scattering coefficients which is numerically implemented in the Monte Carlo method. The scattering pattern and equipartition ratio both show good agreement between the theoretical expectation and the presented Monte Carlo simulation, which verifies this method.

The complete information about the energy propagation allows for the computation of sensitivity kernels of scattered elastic waves including ballistic and scattered waves. Chapter 4 investigates sensitivity kernels in the form ${}^{\psi}K^{\phi}_{YX}$, where $\psi$ denotes the medium perturbation, $\phi$ denotes observable and $Y, X$ denote the excited wave mode $X$ and recorded wave mode $Y$. Both, the observable and medium property

need to be specified to identify the kernels. ${}^{\alpha}K^{tt}_{YX}$ and ${}^{\beta}K^{tt}_{YX}$ are traveltime-velocity sensitivity kernels to describe the effect of P- and S-wave velocity perturbation in space on the traveltime perturbations of the seismogram. There are eight possible types of traveltime sensitivity kernels that result from the propagation of the two elastic wave modes. Changes in the strength of random velocity and density fluctuations $\varepsilon$ can be observed as changes of the trace envelopes and decorrelation of the waveforms leading to the energy-scattering kernel ${}^{\varepsilon}K^{E}_{YX}$ and the decorrelation-scattering kernel ${}^{\varepsilon}K^{dc}_{YX}$. The energy-scattering kernel ${}^{\varepsilon}K^{E}_{YX}$ has positive and negative polarity whereas the decorrelation-scattering kernel ${}^{\varepsilon}K^{dc}_{YX}$ is strictly positive since any change in the scattering coefficient $\varepsilon$ (independent of its sign) will lead to an *increase* in decorrelation. Based on the observation of the decorrelation alone, one cannot know whether it is due to an increase or a decrease of heterogeneity. The functional form of the energy-attenuation kernels ${}^{Q}K^{E}_{YX}$ is the same as that of the traveltime-velocity kernels. Eight different kernels for the combinations of the quality factors for P- and S-waves and the modes of excitation and recording are obtained. The different kernels can be superimposed with the suitable weighting which can be obtained from the simulation of the specific energy density if the actual sources in an experiment emit both, P- and S-waves simultaneously or/and the receiver does not separate between P- and S-waves. I compared sensitivity kernels in a statistically homogeneous model with kernels calculated in a model that contains anomalies of the scattering properties to show the influence of spatial variations in scattering strength. Obvious differences between kernels in the homogeneous and anomaly models exist in ${}^{\alpha}K^{tt}_{SS}$ and ${}^{\beta}K^{tt}_{PP}$ which are strongly affected by the scattering process because of the required mode conversion.

Based on the achievements from the modeling and the derivation of the sensitivity kernels, in chapter 5 I used the adjoint method to derive the Fréchet derivatives of the least-square misfit function which is defined as the squared difference between observed/synthetic and modeled data as the objective function. The Monte Carlo method is used to generate the synthetic data as well as for the forward and adjoint simulations. The misfit kernels are derived with the format as ${}^{\varepsilon}K^{\chi}_{YX}(\mathbf{r})$ and ${}^{Q}K^{\chi}_{YX}(\mathbf{r})$ corresponding to the sensitivity kernels, where $\chi$ is the misfit function. The iterative update directions for the models $\varepsilon(\mathbf{r})$ or $Q^{-1}(\mathbf{r})$ from the initial models are calculated with the L-BFGS method using the misfit kernels. AET has been verified in numerical and laboratory experiments in chapter 5 and chapter 6, respectively.

The numerical experiments demonstrate the success of the presented inversion concept. Fréchet derivatives of the two experiments with separate inversions for either $\varepsilon$ or $Q^{-1}$ reliably guide the inversion towards lower values of the misfit function. Models which better reproduce the synthetic envelopes as expressed by lower misfit also converge towards the true model in the model space. The inversion is also well behaved with respect to the addition of noise. The iterative inversion correctly recovers the locations and amplitudes of the anomalies in each parameter, if the other parameter is fixed at the correct value. If both parameters are free in the inversion, a trade-off is observed which is common to most inverse problems with multiple parameters. In this case, the inversion process is dominated by the parameter that has the stronger anomalies, i.e. dominates the misfit function.

Although the trade-off between scattering and absorption cannot be avoided, I proposed some strategies to improve the simultaneous inversion. At first, correct knowledge of one of the parameters improves the recovery of the other. This means that independent knowledge of one parameter can be incorporated to stabilize the inversion. Secondly, the time window used to construct the adjoint source affects the recovery of the two parameters differently. While the late coda is essential to infer the distribution of absorption, the ballistic wave and early coda is important to locate anomalies in the small-scale heterogeneity. Applying a weighting between the influence of early and late coda in the inversion can help to improve the recovery of both parameters.

The laboratory experiment is conducted in a 4 m by 5 m large reinforced concrete specimen equipped with embedded ultrasonic transducers. The data recorded by 19 transducers are compared with simulations of energy propagation based on the Radiative Transfer Equation. The forward

problem is solved by modelling the 2-D multiple nonisotropic scattering in an acoustic medium with spatially variable heterogeneity and attenuation. The misfit between observed and modeled envelopes is minimized by iteratively updating the model with the adjoint method. Here I introduced a whole workflow of AET for the real data including the processing of the data and the investigation of background values with the diffusion model. The fluctuation strength $\varepsilon$ and intrinsic quality factor $Q^{-1}$ respectively representing the spatial variability of scattering and absorption are separately inverted from different time windows. On the one side, the absorption inversion result shows a strong point-symmetric geometry which we interpret as some large-scale parameter variation in the specimen, but without a direct evidences for the causative process, e.g. temperature, humidity or stress a definite interpretation is impossible. On the other side, the inverted distribution of scattering properties shows a more complex structure that can – to some extent – be interpreted in terms of the known internal structure of the test specimen. The largest anomaly of increased heterogeneity corresponds to a volume containing salt-concrete. Other anomalies are not as clearly linked to the known features of the concrete and a strong anomaly of decreased heterogeneity exists at the edge of the specimen that is interpreted as an artifact from envelope fluctuation that are insufficiently averaged at the reflecting boundaries of the model domain. Despite obvious room for improvement in terms of spatial resolution and power to resolve the trade-off between scattering and attenuation the present results are encouraging.

The code for Monte Carlo simulation, scripts used to calculate the misfit kernels and the processed data of the laboratory experiment is archived at `https://github.com/TuoZhang-seism/AET_concrete`.

An initial attempt to promote AET to investigate the 1D spherically symmetric heterogeneity model of the Earth is conducted. The preliminary work about global records of $P_{\text{diff}}$ coda is presented in chapter 7.

Existence of $P_{\text{diff}}$ coda at high frequencies (1∼2 Hz) is for the first time demonstrated at a very large epicenter distance (even more than 150°). The stacking results of global earthquakes show that observations are best for the earthquakes with huge magnitude and deep source. Similar to the modeling in the laboratory experiment, the Monte Carlo simulation is used but based on 3D non-isotropic multiple scattering in a 1D spherically symmetric heterogeneity model. The synthetic data produced with the heterogeneity model by Bentham et al. (2017) coincides well with the global stacked high-SNR $P_{\text{diff}}$ coda from a deep event (8.2 Mw, 600 km, August 19, 2018 Fiji Islands region earthquake) in the time window between arrival times of the $P_{\text{diff}}$ and $PKIKP$. Compared with the simulation without scattering that shows a huge discrepancy the fit is remarkable. Separating the model into multiple single-layer, I simulated the scattered energy that is contributed from each of theses depth layers. From this comparison, I conclude (A) the high-frequency $P_{\text{diff}}$ at long distances (≥100°) should be intrinsically the same as $P_{\text{diff}}$ coda, (B) the lowermost-mantle scattering contributes to the earliest part of $P_{\text{diff}}$ coda that used to be recognized as $P_{\text{diff}}$ when the distance is not too large, and (C) with the scattered layer rising above the CMB, the energy arrives sequentially one after another, jointly forming the $P_{\text{diff}}$ coda. The volume that can generate the $P_{\text{diff}}$ coda becomes narrower when the epicenter distance becomes larger (≥130°). Both the lowermost mantle and uppermost mantle lose the conditions to scatter the energy into the $P_{\text{diff}}$ coda time window at long distances. I use single-scattering to demonstrate all possible areas of scattering and their traveltime compared with $PKIKP$. Beyond 168° no $P_{\text{diff}}$ coda can be generated by single scattering any more. Onset times of single scattered $P * P$ are followed by the observed seismic energy with some delay at large distances resulting from the finite noise level that has to be exceeded to generate an observable energy signal.

## 8.2 Outlook

This thesis has succeeded in developing the methodology of Adjoint Envelope Tomography (AET) and has verified the efficiency by numerical and laboratory experiments. It has taken a large step towards the final goal of an application of AET to imaging the small-scale structures of the Earth.

The preliminary work about the investigation of the scattered waves in the Earth has been finished with the collection of observations that can be used later in an inversion based on the developed methodologies. This step is analogous to collecting the ultrasonic data in the concrete experiment. The tomography of the whole Earth with AET still faces some challenges. The calculation of the precise sensitivity kernels requires the knowledge of the specific energy density. Recording of the specific energy density that contains the information about propagation direction for the whole Earth requires a very huge memory. Although the model can be spherically symmetric, the simulation and recording have to deal with 3D in space. Including the time and two-dimensional directional vector, the data extend to 6 dimensions. Although this problem can be solved with growing high-performance computer, now a 1D spherically symmetric heterogeneity Earth model is more realistic to achieve based on isotropic scattering in which the directional information is not necessary. The inversion can start with an existing initial model (e.g. BRT2017) combined with a fixed intrinsic attenuation model, like $ak135 - F$. For the small-scale heterogeneity of the Earth, the observed coda waves or precursors are believed to be most sensitive. The $P_{\text{diff}}$ coda chosen in this thesis originates from the $P$-to-$P$ scattering. Since this very early arriving energy has the advantage that it is not disturbed by other seismic phases and that it almost exclusively depends on $P$ to $P$ scattering which simplifies the inversion process. Future work can consider more scattering pattern, including $P$-to-$S$, $S$-to-$P$ and $S$-to-$S$ that are contained in different time-distance windows.

In addition to the global tomography, it is also interesting to work on the regional imaging of the near-surface heterogeneity with AET. The crust exhibits stronger small-scale heterogeneity, like rock fractures, small regional faults, or volcanic areas. The application of AET faces two challenges. On the one side, the investigation of small-scale structures requires sufficient knowledge of deterministic structures. On the other side, the forward modeling for RTT needs more adaptability in the lateral variation of elastic properties. The development of Monte Carlo simulation needs to consider more complex velocity models. Moreover, combined with other techniques, like CWI, the sensitivity kernels calculated with the precise modeling can more accurately locate the changes in such media.

Lastly, in the laboratory experiment AET shows great potential in the field of engineering, like non-destructive testing. Although I successfully inverted the scattering and absorption structures, the lack of connection between the statistical properties mathematically described in RTT and the physical material of the constructions makes the interpretation indirect. On the one hand, more experiments should be conducted to strengthen this connection. On the other hand, ultrasonic experiments under different conditions can be used in AET to monitor the changes in the material.

# Bibliography


Abubakirov, I. & Gusev, A., 1990. Estimation of scattering properties of lithosphere of kamchatka based on monte-carlo simulation of record envelope of a near earthquake, *Physics of the Earth and Planetary Interiors*, **64**(1), 52–67.

Akcelik, V., Biros, G., & Ghattas, O., 2002. Parallel multiscale gauss-newton-krylov methods for inverse wave propagation, in *SC'02: Proceedings of the 2002 ACM/IEEE Conference on Supercomputing*, pp. 41–41, IEEE.

Aki, K., 1969. Analysis of the seismic coda of local earthquakes as scattered waves, *Journal of Geophysical Research*, **74**(2), 615–631.

Aki, K., 1973. Scattering of p waves under the montana lasa, *Journal of Geophysical Research (1896-1977)*, **78**(8), 1334–1346.

Aki, K., 1980. Attenuation of shear-waves in the lithosphere for frequencies from 0.05 to 25 hz, *Physics of the Earth and Planetary Interiors*, **21**(1), 50–60.

Aki, K. & Chouet, B., 1975. Origin of coda waves: source, attenuation, and scattering effects, *Journal of Geophysical Research*, **80**(23), 3322–3342.

Aki, K. & Lee, W., 1976. Determination of three-dimensional velocity anomalies under a seismic array using first p arrival times from local earthquakes: 1. a homogeneous initial model, *Journal of Geophysical Research*, **81**(23), 4381–4399.

Anugonda, P., Wiehn, J. S., & Turner, J. A., 2001. Diffusion of ultrasound in concrete, *Ultrasonics*, **39**(6), 429–435.

Apresyan, L. A. & Kravtsov, I. A., 1996. *Radiation transfer statistical and wave aspects*, CRC Press, London.

Bataille, K. & Lund, F., 1996. Strong scattering of short-period seismic waves by the core-mantle boundary and the p-diffracted wave, *Geophysical Research Letters*, **23**(18), 2413–2416.

Bataille, K., Wu, R., & Flatte, S., 1990. Inhomogeneities near the core-mantle boundary evidenced from scattered waves: A review, *Pure and Applied Geophysics*, **132**(1), 151–173.

Batchelor, G. K., 1953. *The theory of homogeneous turbulence*, Cambridge University Press.

Becker, J., Jacobs, L. J., & Qu, J., 2003. Characterization of cement-based materials using diffuse ultrasound, *Journal of Engineering Mechanics*, **129**(12), 1478–1484.

Bentham, H., Rost, S., & Thorne, M., 2017. Fine-scale structure of the mid-mantle characterised by global stacks of pp precursors, *Earth and Planetary Science Letters*, **472**, 164–173.

Bianco, F., Del Pezzo, E., Castellano, M., Ibanez, J., & Di Luccio, F., 2002. Separation of intrinsic and scattering seismic attenuation in the southern apennine zone, italy, *Geophysical Journal International*, **150**(1), 10–22.

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., & Pugmire, D., 2016. Global adjoint tomography: first-generation model, *Geophysical Journal International*, **207**(3), 1739–1766.

Budi-Santoso, A. & Lesage, P., 2016. Velocity variations associated with the large 2010 eruption of merapi volcano, java, retrieved from seismic multiplets and ambient noise cross-correlation, *Geophysical Journal International*, **206**(1), 221–240.

Calvet, M., Sylvander, M., Margerin, L., & Villaseñor, A., 2013. Spatial variations of seismic attenuation and heterogeneity in the pyrenees: Coda q and peak delay time analysis, *Tectonophysics*, **608**, 428–439.

Carcolé, E. & Sato, H., 2010. Spatial distribution of scattering loss and intrinsic absorption of short-period s waves in the lithosphere of japan on the basis of the multiple lapse time window analysis of hi-net data, *Geophysical Journal International*, **180**(1), 268–290.

Chauris, H., 2021. *Chapter 5 Full waveform inversion*, pp. 123–146, EDP Sciences.

Clark, V. A., Tittmann, B. R., & Spencer, T. W., 1980. Effect of volatiles on attenuation (q- 1) and velocity in sedimentary rocks, *Journal of Geophysical Research: Solid Earth*, **85**(B10), 5190–5198.

Cleary, J. & Haddon, R., 1972. Seismic wave scattering near the core-mantle boundary: a new interpretation of precursors to pkp, *Nature*, **240**(5383), 549–551.

Cleary, J. R., King, D. W., & Haddon, R. A. W., 1975. P-Wave Scattering in the Earth's Crust and Upper Mantle, *Geophysical Journal International*, **43**(3), 861–872.

Cormier, V. F. & Sanborn, C. J., 2019. Trade-offs in parameters describing crustal heterogeneity and intrinsic attenuation from radiative transport modeling of high-frequency regional seismogramstrade-offs in parameters describing crustal heterogeneity and intrinsic attenuation, *Bulletin of the Seismological Society of America*, **109**(1), 312–321.

Crotwell, H. P., Owens, T. J., & Ritsema, J., 1999. The TauP Toolkit: Flexible Seismic Travel-time and Ray-path Utilities, *Seismological Research Letters*, **70**(2), 154–160.

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

De Siena, L., Del Pezzo, E., Bianco, F., & Tramelli, A., 2009. Multiple resolution seismic attenuation imaging at mt. vesuvius, *Physics of the Earth and Planetary Interiors*, **173**(1-2), 17–32.

De Siena, L., Del Pezzo, E., Thomas, C., Curtis, A., & Margerin, L., 2013. Seismic energy envelopes in volcanic media: in need of boundary conditions, *Geophysical Journal International*, **195**(2), 1102–1119.

De Siena, L., Thomas, C., & Aster, R., 2014a. Multi-scale reasonable attenuation tomography analysis (murat): An imaging algorithm designed for volcanic regions, *Journal of Volcanology and Geothermal Research*, **277**, 22–35.

De Siena, L., Thomas, C., Waite, G. P., Moran, S. C., & Klemme, S., 2014b. Attenuation and scattering tomography of the deep plumbing system of mount st. helens, *Journal of Geophysical Research: Solid Earth*, **119**(11), 8223–8238.

De Siena, L., Calvet, M., Watson, K. J., Jonkers, A., & Thomas, C., 2016. Seismic scattering and absorption mapping of debris flows, feeding paths, and tectonic units at mount st. helens volcano, *Earth and Planetary Science Letters*, **442**, 21–31.

De Siena, L., Amoruso, A., Pezzo, E. D., Wakeford, Z., Castellano, M., & Crescentini, L., 2017. Space-weighted seismic attenuation mapping of the aseismic source of campi flegrei 1983–1984 unrest, *Geophysical Research Letters*, **44**(4), 1740–1748.

Del Pezzo, E. & Ibáñez, J. M., 2020. Seismic coda-waves imaging based on sensitivity kernels calculated using an heuristic approach, *Geosciences*, **10**(8), 304.

Del Pezzo, E., Bianco, F., De Siena, L., & Zollo, A., 2006. Small scale shallow attenuation structure at mt. vesuvius, italy, *Physics of the Earth and Planetary Interiors*, **157**(3-4), 257–268.

Del Pezzo, E., Ibanez, J., Prudencio, J., Bianco, F., & De Siena, L., 2016. Absorption and scattering 2-d volcano images from numerically calculated space-weighting functions, *Geophysical Journal International*, **206**(2), 742–756.

Del Pezzo, E., De La Torre, A., Bianco, F., Ibanez, J., Gabrielli, S., & De Siena, L., 2018. Numerically calculated 3d space-weighting functions to image crustal volcanic structures using diffuse coda waves, *Geosciences*, **8**(5), 175.

Deroo, F., Kim, J.-Y., Qu, J., Sabra, K., & Jacobs, L. J., 2010. Detection of damage in concrete using diffuse ultrasound, *The Journal of the Acoustical Society of America*, **127**(6), 3315–3318.

Domínguez, J. & Abascal, R., 1984. On fundamental solutions for the boundary integral equations method in static and dynamic elasticity, *Engineering Analysis*, **1**(3), 128–134.

Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary reference earth model, *Physics of the earth and planetary interiors*, **25**(4), 297–356.

Earle, P. S. & Shearer, P. M., 2001. Distribution of Fine-Scale Mantle Heterogeneity from Observations of Pdiff Coda, *Bulletin of the Seismological Society of America*, **91**(6), 1875–1881.

Epple, N., Barroso, D. F., & Niederleithinger, E., 2020. Towards monitoring of concrete structures with embedded ultrasound sensors and coda waves–first results of dfg for coda, in *European Workshop on Structural Health Monitoring*, pp. 266–275, Springer.

Eulenfeld, T. & Wegler, U., 2016. Measurement of intrinsic and scattering attenuation of shear waves in two sedimentary basins and comparison to crystalline sites in germany, *Geophysical Journal International*, **205**(2), 744–757.

Fehler, M., Hoshiba, M., Sato, H., & Obara, K., 1992. Separation of scattering and intrinsic attenuation for the kanto-tokai region, japan, using measurements of s-wave energy versus hypocentral distance, *Geophysical Journal International*, **108**(3), 787–800.

Fichtner, A., 2010. *Full seismic waveform modelling and inversion*, Springer Science & Business Media.

Fichtner, A. & Igel, H., 2008. Efficient numerical surface wave propagation through the optimization of discrete crustal models—a technique based on non-linear dispersion curve matching (dcm), *Geophysical Journal International*, **173**(2), 519–533.

Fichtner, A. & Trampert, J., 2011. Hessian kernels of seismic data functionals based upon adjoint techniques, *Geophysical Journal International*, **185**(2), 775–798.

Fichtner, A., Bunge, H.-P., & Igel, H., 2006. The adjoint method in seismology: I. theory, *Physics of the Earth and Planetary Interiors*, **157**(1-2), 86–104.

Fichtner, A., Kennett, B. L., Igel, H., & Bunge, H.-P., 2010. Full waveform tomography for radially anisotropic structure: new insights into present and past states of the australasian upper mantle, *Earth and Planetary Science Letters*, **290**(3-4), 270–280.

French, S. W. & Romanowicz, B., 2015. Broad plumes rooted at the base of the earth's mantle beneath major hotspots, *Nature*, **525**(7567), 95–99.

Friedrich, C. & Wegler, U., 2005. Localization of seismic coda at merapi volcano (indonesia), *Geophysical Research Letters*, **32**(14).

Gabrielli, S., De Siena, L., Napolitano, F., & Del Pezzo, E., 2020. Understanding seismic path biases and magmatic activity at mount st helens volcano before its 2004 eruption, *Geophysical Journal International*, **222**(1), 169–188.

Gaebler, P., Eken, T., Bektaş, H. Ö., Eulenfeld, T., Wegler, U., & Taymaz, T., 2019. Imaging of shear wave attenuation along the central part of the north anatolian fault zone, turkey, *Journal of Seismology*, **23**(4), 913–927.

Gaebler, P. J., Sens-Schönfelder, C., & Korn, M., 2015. The influence of crustal scattering on translational and rotational motions in regional and teleseismic coda waves, *Geophysical Journal International*, **201**(1), 355–371.

Garnero, E. J., 2004. A new paradigm for earth's core-mantle boundary, *Science*, **304**(5672), 834–836.

Green, D., Wang, H., & Bonner, B., 1993. Shear wave attenuation in dry and saturated sandstone at seismic to ultrasonic frequencies, in *International journal of rock mechanics and mining sciences & geomechanics abstracts*, vol. 30, pp. 755–761, Elsevier.

Grohmann, M., Müller, S., Niederleithinger, E., & Sieber, S., 2017. Reverse time migration: introducing a new imaging technique for ultrasonic measurements in civil engineering, *Near Surface Geophysics*, **15**(3), 242–258.

Gusev, A. & Abubakirov, I., 1987. Monte-carlo simulation of record envelope of a near earthquake, *Physics of the Earth and Planetary Interiors*, **49**(1-2), 30–36.

Gusev, A. A. & Abubakirov, I. R., 1996. Simulated envelopes of non-isotropically scattered body waves as compared to observed ones: another manifestation of fractal heterogeneity, *Geophysical Journal International*, **127**(1), 49–60.

Haddon, R. & Cleary, J., 1974. Evidence for scattering of seismic pkp waves near the mantle-core boundary, *Physics of the Earth and Planetary Interiors*, **8**(3), 211–234.

Hager, I., 2013. Behaviour of cement concrete at high temperature, *Bulletin of the Polish Academy of Sciences. Technical Sciences*, **61**(1), 145–154.

Hedlin, M. A., Shearer, P. M., & Earle, P. S., 1997. Seismic evidence for small-scale heterogeneity throughout the earth's mantle, *Nature*, **387**(6629), 145–150.

Hiemer, V. & Thomas, C., 2022. Generation of reflections and pkp precursors from a scattering layer in d", *Geophysical Research Letters*, **49**(4), e2021GL096900.

Holliger, K., 1996. Upper-crustal seismic velocity heterogeneity as derived from a variety of p-wave sonic logs, *Geophysical Journal International*, **125**(3), 813–829.

Hoshiba, M., 1991. Simulation of multiple-scattered coda wave excitation based on the energy conservation law, *Physics of the Earth and Planetary Interiors*, **67**(1-2), 123–136.

Hoshiba, M., 1993. Separation of scattering attenuation and intrinsic absorption in japan using the multiple lapse time window analysis of full seismogram envelope, *Journal of Geophysical Research: Solid Earth*, **98**(B9), 15809–15824.

Hoshiba, M., 1994. Simulation of coda wave envelope in depth dependent scattering and absorption structure, *Geophysical Research Letters*, **21**(25), 2853–2856.

Hoshiba, M., 1995. Estimation of nonisotropic scattering in western japan using coda wave envelopes: Application of a multiple nonisotropic scattering model, *Journal of Geophysical Research: Solid Earth*, **100**(B1), 645–657.

Izgi, G., Eken, T., Gaebler, P., Eulenfeld, T., & Taymaz, T., 2020. Crustal seismic attenuation parameters in the western region of the north anatolian fault zone, *Journal of Geodynamics*, **134**, 101694.

Jiang, C., Yang, Y., & Zheng, Y., 2014. Penetration of mid-crustal low velocity zone across the kunlun fault in the ne tibetan plateau revealed by ambient noise tomography, *Earth and Planetary Science Letters*, **406**, 81–92.

Jin, A. & Aki, K., 2005. High-resolution maps of coda q in japan and their interpretation by the brittle-ductile interaction hypothesis, *Earth, Planets and Space*, **57**(5), 403–409.

Kanu, C. & 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.

Kennet, B. L. N., 1991. Iaspei 1991 seismological tables, *Terra Nova*, **3**(2), 122–122.

Kennett, B., Yoshizawa, K., & Furumura, T., 2017. Interactions of multi-scale heterogeneity in the lithosphere: Australia, *Tectonophysics*, **717**, 193–213.

Kennett, B. L. N., Engdahl, E. R., & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes, *Geophysical Journal International*, **122**(1), 108–124.

King, D., Haddon, R., & Husebye, E., 1975. Precursors to pp, *Physics of the Earth and Planetary Interiors*, **10**(2), 103–127.

Koelemeijer, P., Deuss, A., & Ritsema, J., 2017. Density structure of earth's lowermost mantle from stoneley mode splitting observations, *Nature Communications*, **8**(1), 1–10.

Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophysical Journal International*, **139**(3), 806–822.

Koulakov, I., Kaban, M., Tesauro, M., & Cloetingh, S., 2009. P-and s-velocity anomalies in the upper mantle beneath europe from tomographic inversion of isc data, *Geophysical Journal International*, **179**(1), 345–366.

Lacombe, C., Campillo, M., Paul, A., & 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.

Larose, E., de Rosny, J., Margerin, L., Anache, D., Gouedard, P., Campillo, M., & van Tiggelen, B., 2006. Observation of multiple scattering of khz vibrations in a concrete structure and application to monitoring weak changes, *Physical Review E*, **73**(1), 016609.

Larose, E., Planes, T., Rossetto, V., & Margerin, L., 2010. Locating a small change in a multiple scattering environment, *Applied Physics Letters*, **96**(20), 204101.

Lay, T. & Garnero, E. J., 2011. Deep mantle seismic modeling and imaging, *Annual Review of Earth and Planetary Sciences*, **39**, 91–123.

Lei, W., Ruan, Y., Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., & Pugmire, D., 2020. Global adjoint tomography—model GLAD-M25, *Geophysical Journal International*, **223**(1), 1–21.

Lévěque, J.-J., Rivera, L., & Wittlinger, G., 1993. On the use of the checker-board test to assess the resolution of tomographic inversions, *Geophysical Journal International*, **115**(1), 313–318.

Li, C. & Van Der Hilst, R. D., 2010. Structure of the upper mantle and transition zone beneath southeast asia from traveltime tomography, *Journal of Geophysical Research: Solid Earth*, **115**(B7).

Liu, D. C. & Nocedal, J., 1989. On the limited memory bfgs method for large scale optimization, *Mathematical Programming*, **45**(1-3), 503–528.

Liu, L. & Guo, T., 2005. Seismic non-destructive testing on a reinforced concrete bridge column using tomographic imaging techniques, *Journal of Geophysics and Engineering*, **2**(1), 23–31.

Liu, Q. & Gu, Y., 2012. Seismic imaging: From classical to adjoint tomography, *Tectonophysics*, **566**, 31–66.

Mancinelli, N. & Shearer, P., 2016. Scattered energy from a rough core-mantle boundary modeled by a monte carlo seismic particle method: Application to pkkp precursors, *Geophysical Research Letters*, **43**(15), 7963–7972.

Mancinelli, N., Shearer, P., & Liu, Q., 2016. Constraints on the heterogeneity spectrum of earth's upper mantle, *Journal of Geophysical Research: Solid Earth*, **121**(5), 3703–3721.

Mancinelli, N. J. & Shearer, P. M., 2013. Reconciling discrepancies among estimates of small-scale mantle heterogeneity from PKP precursors, *Geophysical Journal International*, **195**(3), 1721–1729.

Margerin, L., 2005. Introduction to radiative transfer of seismic waves, *Geophysical Monograph-American Geophysical Union*, **157**, 229.

Margerin, L., 2017. Computation of green's function of 3-d radiative transport equations for non-isotropic scattering of p and unpolarized s waves, *Pure and Applied Geophysics*, **174**(11), 4057–4075.

Margerin, L. & Nolet, G., 2003. Multiple scattering of high-frequency seismic waves in the deep earth: Pkp precursor analysis and inversion for mantle granularity, *Journal of Geophysical Research: Solid Earth*, **108**(B11).

Margerin, L., Campillo, M., & 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., & 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., & 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.

Maruyama, S., Santosh, M., & Zhao, D., 2007. Superplume, supercontinent, and post-perovskite: Mantle dynamics and anti-plate tectonics on the core–mantle boundary, *Gondwana Research*, **11**(1), 7–37, Island Arcs: Past and Present.

Mayor, J., Margerin, L., & Calvet, M., 2014. Sensitivity of coda waves to spatial variations of absorption and scattering: radiative transfer theory and 2-d examples, *Geophysical Journal International*, **197**(2), 1117–1137.

Mayor, J., Calvet, M., Margerin, L., Vanderhaeghe, O., & Traversa, P., 2016. Crustal structure of the alps as seen by attenuation tomography, *Earth and Planetary Science Letters*, **439**, 71–80.

Montagner, J.-P. & Kennett, B. L. N., 1996. How to reconcile body-wave and normal-mode reference earth models, *Geophysical Journal International*, **125**(1), 229–248.

Niederleithinger, E., 2017. *Seismic Methods Applied to Ultrasonic Testing in Civil Engineering (Habilitation Thesis)*, habilitation, RWTH Aachen.

Niederleithinger, E. & Wunderlich, C., 2013. Influence of small temperature variations on the ultrasonic velocity in concrete, in *AIP Conference Proceedings*, vol. 1511, pp. 390–397, American Institute of Physics.

Niederleithinger, E., Wolf, J., Mielentz, F., Wiggenhauser, H., & Pirskawetz, S., 2015. Embedded ultrasonic transducers for active and passive concrete monitoring, *Sensors*, **15**(5), 9756–9772.

Niederleithinger, E., Wang, X., Herbrand, M., & Müller, M., 2018. Processing ultrasonic data by coda wave interferometry to monitor load tests of concrete beams, *Sensors*, **18**(6), 1971.

Nishigami, K., 1991. A new inversion method of coda waveforms to determine spatial distribution of coda scatterers in the crust and uppermost mantle, *Geophysical Research Letters*, **18**(12), 2225–2228.

Nishigami, K., 1997. Spatial distribution of coda scatterers in the crust around two active volcanoes and one active fault system in central japan: Inversion analysis of coda envelope, *Physics of the Earth and Planetary Interiors*, **104**(1-3), 75–89.

Nishigami, K., 2000. Deep crustal heterogeneity along and around the san andreas fault system in central california and its relation to the segmentation, *Journal of Geophysical Research: Solid Earth*, **105**(B4), 7983–7998.

Nocedal, J. & Wright, S., 2006. *Numerical Optimization*, Springer Science & Business Media.

Obermann, A., Planès, T., Larose, E., & Campillo, M., 2013a. Imaging preeruptive and coeruptive structural and mechanical changes of a volcano with ambient seismic noise, *Journal of Geophysical Research: Solid Earth*, **118**(12), 6285–6294.

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

Obermann, A., Froment, B., Campillo, M., Larose, E., Planes, T., Valette, B., Chen, J., & Liu, Q., 2014. Seismic noise correlations to image structural and mechanical changes associated with the mw 7.9 2008 wenchuan earthquake, *Journal of Geophysical Research: Solid Earth*, **119**(4), 3155–3168.

Obermann, A., Planès, T., Hadziioannou, C., & 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.

Ogiso, M., 2019. A method for mapping intrinsic attenuation factors and scattering coefficients of s waves in 3-d space and its application in southwestern japan, *Geophysical Journal International*, **216**(2), 948–957.

Pacheco, C. & 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.

Pacheco, C. & Snieder, R., 2006. Time-lapse traveltime change of singly scattered acoustic waves, *Geophysical Journal International*, **165**(2), 485–500.

Padhy, S., Wegler, U., & Korn, M., 2007. Seismogram envelope inversion using a multiple isotropic scattering model: Application to aftershocks of the 2001 bhuj earthquake, *Bulletin of the Seismological Society of America*, **97**(1B), 222–233.

Planès, T. & Larose, E., 2013. A review of ultrasonic coda wave interferometry in concrete, *Cement and Concrete Research*, **53**, 248–255.

Planès, T., Larose, E., Margerin, L., Rossetto, V., & 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.

Poupinet, G., Ellsworth, W., & Frechet, J., 1984. Monitoring velocity variations in the crust using earthquake doublets: An application to the calaveras fault, california, *Journal of Geophysical Research: Solid Earth*, **89**(B7), 5719–5731.

Prudencio, J., Del Pezzo, E., García-Yeguas, A., & Ibáñez, J. M., 2013. Spatial distribution of intrinsic and scattering seismic attenuation in active volcanic islands–i: model and the case of tenerife island, *Geophysical Journal International*, **195**(3), 1942–1956.

Prudencio, J., Del Pezzo, E., Ibáñez, J., Giampiccolo, E., & Patané, D., 2015a. Two-dimensional seismic attenuation images of stromboli island using active data, *Geophysical Research Letters*, **42**(6), 1717–1724.

Prudencio, J., Ibáñez, J., Del Pezzo, E., Martí, J., García-Yeguas, A., & De Siena, L., 2015b. 3d attenuation tomography of the volcanic island of tenerife (canary islands), *Surveys in Geophysics*, **36**(5), 693–716.

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

Ramamoorthy, S. K., Kane, Y., & Turner, J. A., 2004. Ultrasound diffusion for crack depth determination in concrete, *The Journal of the Acoustical Society of America*, **115**(2), 523–529.

Rawlinson, N., Sambridge, M., et al., 2003. Seismic traveltime tomography of the crust and lithosphere, *Advances in Geophysics*, **46**, 81–199.

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

Rost, S., Thorne, M. S., & Garnero, E. J., 2006. Imaging Global Seismic Phase Arrivals by Stacking Array Processed Short-Period Data, *Seismological Research Letters*, **77**(6), 697–707.

Rost, S., Garnero, E. J., & Williams, Q., 2008. Seismic array detection of subducted oceanic crust in the lower mantle, *Journal of Geophysical Research: Solid Earth*, **113**(B6).

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

Sanborn, C. J., 2017. *Simulations with Radiative3D: A software tool for radiative transport in 3-D Earth models*, Doctoral dissertations, UCONN.

Sanborn, C. J., Cormier, V. F., & Fitzpatrick, M., 2017. Combined effects of deterministic and statistical structure on high-frequency regional seismograms, *Geophysical Journal International*, **210**(2), 1143–1159.

Sánchez-Pastor, P., Obermann, A., Schimmel, M., Weemstra, C., Verdel, A., & Jousset, P., 2019. Short-and long-term variations in the reykjanes geothermal reservoir from seismic noise interferometry, *Geophysical Research Letters*, **46**(11), 5788–5798.

Sato, H., 1982. Attenuation of s waves in the lithosphere due to scattering by its random velocity structure, *Journal of Geophysical Research: Solid Earth*, **87**(B9), 7779–7785.

Sato, H., Fehler, M. C., & Maeda, T., 2012. *Seismic wave propagation and scattering in the heterogeneous earth*, vol. 496, Springer.

Schickert, M., 2005. Progress in ultrasonic imaging of concrete, *Materials and Structures*, **38**(9), 807–815.

Schickert, M., Krause, M., & Müller, W., 2003. Ultrasonic imaging of concrete elements using reconstruction by synthetic aperture focusing technique, *Journal of Materials in Civil Engineering*, **15**(3), 235–246.

Schurr, D. P., Kim, J.-Y., Sabra, K. G., & Jacobs, L. J., 2011. Damage detection in concrete using coda wave interferometry, *NDT & E International*, **44**(8), 728–735.

Sens-Schönfelder, C. & Brenguier, F., 2019. Noise-based monitoring, in *Seismic Ambient Noise*, pp. 267–301, eds Nakata, N., Gualtieri, L., & Fichtner, A., Cambridge University Press.

Sens-Schönfelder, C. & Wegler, U., 2006. Radiative transfer theory for estimation of the seismic moment, *Geophysical Journal International*, **167**(3), 1363–1372.

Sens-Schönfelder, C., Margerin, L., & Campillo, M., 2009. Laterally heterogeneous scattering explains lg blockage in the pyrenees, *Journal of Geophysical Research: Solid Earth*, **114**(B7).

Sens-Schönfelder, C., Bataille, K., & Bianchi, M., 2021. High-frequency (6 hz) pkpab precursors and their sensitivity to deep earth heterogeneity, *Geophysical Research Letters*, **48**(2), e2020GL89203.

Shearer, P. M., 2009. *Introduction to Seismology*, Cambridge University Press, Cambridge, 2nd edn.

Shearer, P. M., 2015. 1.24 - deep earth structure: Seismic scattering in the deep earth, in *Treatise on Geophysics (Second Edition)*, pp. 759–787, ed. Schubert, G., Elsevier, Oxford, second edition edn.

Shearer, P. M. & Earle, P. S., 2004. The global short-period wavefield modelled with a Monte Carlo seismic phonon method, *Geophysical Journal International*, **158**(3), 1103–1117.

Simmons, N. A., Myers, S. C., Morency, C., Chiang, A., & Knapp, D. R., 2021. SPiRaL: a multiresolution global tomography model of seismic wave speeds and radial anisotropy variations in the crust and mantle, *Geophysical Journal International*, **227**(2), 1366–1391.

Sketsiou, P., Napolitano, F., Zenonos, A., & De Siena, L., 2020. New insights into seismic absorption imaging, *Physics of the Earth and Planetary Interiors*, **298**, 106337.

Smith, W. & Wessel, P., 1990. Gridding with continuous curvature splines in tension, *Geophysics*, **55**(3), 293–305.

Snieder, R., 2006. The theory of coda wave interferometry, *Pure and Applied Geophysics*, **163**(2-3), 455–473.

Snieder, R., Grêt, A., Douma, H., & Scales, J., 2002. Coda wave interferometry for estimating nonlinear behavior in seismic velocity, *Science*, **295**(5563), 2253–2255.

Snieder, R., Duran, A., & Obermann, A., 2019. Locating velocity changes in elastic media with coda wave interferometry, *Seismic Ambient Noise*, p. 188.

Soergel, D., Pedersen, H., Stehly, L., Margerin, L., Paul, A., & Group, A. W., 2020. Coda-q in the 2.5–20 s period band from seismic noise: application to the greater alpine area, *Geophysical Journal International*, **220**(1), 202–217.

Takeuchi, N., 2016. Differential monte carlo method for computing seismogram envelopes and their partial derivatives, *Journal of Geophysical Research: Solid Earth*, **121**(5), 3428–3444.

Tao, K., Grand, S. P., & Niu, F., 2017. Full-waveform inversion of triplicated data using a normalized-correlation-coefficient-based misfit function, *Geophysical Journal International*, **210**(3), 1517–1524.

Tao, K., Grand, S. P., & Niu, F., 2018. Seismic structure of the upper mantle beneath eastern asia from full waveform seismic tomography, *Geochemistry, Geophysics, Geosystems*, **19**(8), 2732–2763.

Tape, C., Liu, Q., Maggi, A., & Tromp, J., 2009. Adjoint tomography of the southern california crust, *Science*, **325**(5943), 988–992.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**(8), 1259–1266.

Thurber, C. H., 1983. Earthquake locations and three-dimensional crustal structure in the coyote lake area, central california, *Journal of Geophysical Research: Solid Earth*, **88**(B10), 8226–8236.

Tisato, N. & Quintal, B., 2014. Laboratory measurements of seismic attenuation in sandstone: Strain versus fluid saturation effects, *Geophysics*, **79**(5), WB9–WB14.

Tono, Y. & Yomogida, K., 1996. Complex scattering at the core-mantle boundary observed in short-period diffracted p-waves, *Journal of Physics of the Earth*, **44**(6), 729–744.

Tromp, J., Tape, C., & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophysical Journal International*, **160**(1), 195–216.

Turner, J. A. & Anugonda, P., 2001. Scattering of elastic waves in heterogeneous media with local isotropy, *The Journal of the Acoustical Society of America*, **109**(5), 1787–1795.

van Dinther, C., Margerin, L., & Campillo, M., 2021. Laterally varying scattering properties in the north anatolian fault zone from ambient noise cross-correlations, *Geophysical Journal International*, **225**(1), 589–607.

Wang, X., Chakraborty, J., Bassil, A., & Niederleithinger, E., 2020. Detection of multiple cracks in four-point bending tests using the coda wave interferometry method, *Sensors*, **20**(7), 1986.

Weaver, R. L., 1990. Diffusivity of ultrasound in polycrystals, *Journal of the Mechanics and Physics of Solids*, **38**(1), 55–86.

Wegler, U., 2004. Diffusion of seismic waves in a thick layer: Theory and application to vesuvius volcano, *Journal of Geophysical Research: Solid Earth*, **109**(B7).

Wegler, U., Korn, M., & Przybilla, J., 2006. Modeling full seismogram envelopes using radiative transfer theory with born scattering coefficients, *Pure and Applied Geophysics*, **163**(2-3), 503–531.

Wessel, P., Smith, W. H., Scharroo, R., Luis, J., & Wobbe, F., 2013. Generic mapping tools: improved version released, *Eos, Transactions American Geophysical Union*, **94**(45), 409–410.

Wu, R.-S., 1985. Multiple scattering and energy transfer of seismic waves—separation of scattering effect from intrinsic attenuation—i. theoretical modelling, *Geophysical Journal International*, **82**(1), 57–80.

Wu, R.-S., Xu, Z., & Li, X.-P., 1994. Heterogeneity spectrum and scale-anisotropy in the upper crust revealed by the german continental deep-drilling (ktb) holes, *Geophysical Research Letters*, **21**(10), 911–914.

Xue, Q., Larose, É., Moreau, L., Thery, R., Abraham, O., & Henault, J.-M., 2021. Ultrasonic monitoring of stress and cracks of the 1/3 scale mock-up of nuclear reactor concrete containment structure, *Structural Health Monitoring*, p. 14759217211034729.

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

Yoshimoto, K., Sato, H., & Ohtake, M., 1993. Frequency-dependent attenuation of p and s waves in the kanto area, japan, based on the coda-normalization method, *Geophysical Journal International*, **114**(1), 165–174.

Yuan, Y. O., Simons, F. J., & Tromp, J., 2016. Double-difference adjoint seismic tomography, *Geophysical Journal International*, **206**(3), 1599–1618.

Zeng, Y., 1991. Compact solutions for multiple scattered wave energy in time domain, *Bulletin of the Seismological Society of America*, **81**(3), 1022–1029.

Zhang, T. & 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., & 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, Y., Planes, T., Larose, E., Obermann, A., Rospars, C., & Moreau, G., 2016. Diffuse ultrasound monitoring of stress and damage development on a 15-ton concrete beam, *The Journal of the Acoustical Society of America*, **139**(4), 1691–1701.

Zhang, Y., Larose, E., Moreau, L., & d'Ozouville, G., 2018. Three-dimensional in-situ imaging of cracks in concrete using diffuse ultrasound, *Structural Health Monitoring*, **17**(2), 279–284.

Zhao, D., Hasegawa, A., & Horiuchi, S., 1992. Tomographic imaging of p and s wave velocity structure beneath northeastern japan, *Journal of Geophysical Research: Solid Earth*, **97**(B13), 19909–19928.

Zhong, B., Zhu, J., & Morcous, G., 2021. Measuring acoustoelastic coefficients for stress evaluation in concrete, *Construction and Building Materials*, **309**, 125127.

Zhu, H., Bozdağ, E., Peter, D., & Tromp, J., 2012. Structure of the european upper mantle revealed by adjoint tomography, *Nature Geoscience*, **5**(7), 493–498.

Zhu, H., Bozdağ, E., & Tromp, J., 2015. Seismic structure of the european upper mantle based on adjoint tomography, *Geophysical Journal International*, **201**(1), 18–52.

Zieger, T., Sens-Schönfelder, C., Ritter, J. R., Lühr, B., & Dahm, T., 2016. P-wave scattering and the distribution of heterogeneity around etna volcano, *Annals of Geophysics*, **59**(4).

Zong, J., Stewart, R. R., & Dyaur, N., 2020. Attenuation of rock salt: ultrasonic lab analysis of gulf of mexico coastal samples, *Journal of Geophysical Research: Solid Earth*, **125**(7), e2019JB019025.

# Appendix A

# Born approximation of scattering coefficients in 2-D in-plane elastodynamics

In this Appendix, we outline the computation of the scattering coefficients in a 2-D random elastic medium based on the Born approximation. Our starting point is the following perturbed elastic wave equation for the 2-D displacement field $u_i$ at circular frequency $\omega$ and position $\mathbf{x}$:

$$-\rho^0\omega^2 u_i(\mathbf{x}) - C^0_{ijkl}\partial_j\partial_k u_l(\mathbf{x}) - \int_S V_{ij}(\mathbf{x},\mathbf{x}')u_j(\mathbf{x}')d^2x' = 0 \tag{A.1}$$

where $\rho^0$ and $C^0_{ijkl}$ denote, respectively, the density and elastic tensor of the homogeneous background medium. Indices $\{i,j,\cdots\}$ refer to components of a vector or tensor in a 2-D cartesian system and the Einstein summation convention is adopted. Below, we also employ the latine letters $p,q$ and the greek letter $\alpha$ to denote indices. $V$ represents the scattering potential of the random fluctuations superposed on the background. These fluctuations are assumed to be enclosed in the surface $S$. $V$ may be expressed in terms of the density and elastic tensor fluctuations, denoted by $\delta\rho(\mathbf{x})$ and $\delta C_{ijkl}(\mathbf{x})$, as follows:

$$V_{ij}(\mathbf{x},\mathbf{x}') = \delta\rho(\mathbf{x})\omega^2\delta(\mathbf{x}-\mathbf{x}')\delta_{ij} + \partial_p\left(\delta C_{ipqj}(\mathbf{x})\delta(\mathbf{x}-\mathbf{x}')\partial'_q\right) \tag{A.2}$$

The symbol $\partial'_q$ indicates that the partial derivative acts on the $q^{\text{th}}$ coordinate of the position vector $\mathbf{x}'$.

We seek solutions to Eq. (A.1) of the form $u = u^0 + u^{sc}$ where $u^0$ is a solution of the elastodynamic Eq. in the homogeneous background -i.e., a solution of (A.1) with $V = 0$- and $u^{sc}$ is the field scattered by the inhomogeneities contained in $S$. The formal solution to this problem is known as the Lippman-Schwinger Eq.:

$$u_\alpha(\mathbf{r}) = u^0_\alpha(\mathbf{r}) + \int_S G^0_{\alpha i}(\mathbf{r},\mathbf{x})V_{ij}(\mathbf{x},\mathbf{x}')u_j(\mathbf{x}')d^2x' \tag{A.3}$$

where $G^0$ denotes the elastodynamic Green's function of the background medium and $\mathbf{r}$ is the observation point which we take at a large distance from the surface $S$. For sufficiently weak perturbations (to be further discussed below), the field $u$ in the inhomogeneous region may be replaced by the unperturbed wavefield $u^0$. This is known as the Born approximation in the literature and it forms the basis of our calculation of the scattering coefficients. After substituting in Eq. (A.3) the scattering potential $V$ by its expression (A.2), we perform an integration by part to remove the partial derivative from the elastic perturbations and obtain the following formal expression for the scattered field $u^{sc}$ in the Born approximation:

$$\begin{aligned} u^{sc}_\alpha(\mathbf{r}) =& \omega^2\int_S G^0_{\alpha i}(\mathbf{r},\mathbf{x})\delta\rho(\mathbf{x})u_i(\mathbf{x})d^2x \\ &- \int_S \partial_j G^0_{\alpha i}(\mathbf{r},\mathbf{x})\delta C_{ijkl}(\mathbf{x})\partial_k u^0_l(\mathbf{x})d^2x, \end{aligned} \tag{A.4}$$

where the partial derivatives act on the $\mathbf{x}$ variable. In the case of an isotropic background medium with longitudinal and shear waves speeds given by $c_{p,s}$, the far-field Green's function ($r \to \infty$) is given by (Domínguez & Abascal, 1984):

$$\begin{aligned} G^0_{\alpha_i}(\mathbf{r},\mathbf{x}) \approx & \frac{i\hat{r}_\alpha \hat{r}_i}{4\rho^0 c_p^2}\sqrt{\frac{2}{\pi k_p r}} e^{ik_p(r-\hat{\mathbf{r}}\cdot\mathbf{x})-i\pi/4} \\ & + \frac{i\hat{\theta}_\alpha \hat{\theta}_i}{4\rho^0 c_s^2}\sqrt{\frac{2}{\pi k_s r}} e^{ik_s(r-\hat{\mathbf{r}}\cdot\mathbf{x})-i\pi/4}, \end{aligned} \tag{A.5}$$

where $k_{p,s} = \omega/c_{p,s}$ and $\hat{\mathbf{r}}$ indicates a unit vector in the direction of $\mathbf{r}$. The symbol $\hat{\boldsymbol{\theta}}$ denotes a unit vector perpendicular to $\mathbf{r}$. Eq. (A.5) splits the far-field Green's tensor into its longitudinal and transverse parts. In the case where the unperturbed field is a plane wave with polarization vector $\hat{\mathbf{p}}$, wavenumber $k^{\text{in}} \in k_p, k_s$ and propagation direction $\hat{\mathbf{k}}$, we write:

$$\mathbf{u}^0(\mathbf{x}) = \hat{\mathbf{p}} e^{ik^{\text{in}}\hat{\mathbf{k}}\cdot\mathbf{x}} \tag{A.6}$$

The scattered waves detected in the far-field with polarization vector $\hat{\mathbf{s}} \in \{\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}\}$, wave number $k^{\text{out}} \in \{k_p, k_s\}$ and corresponding propagation speed $c^{\text{out}} \in \{c_p, c_s\}$ will be noted $\mathbf{u}^{\text{in}\to\text{out}}(\mathbf{r})$. The superscripts $\{\text{in}, \text{out}\}$ are shortcuts for the 'incoming' and 'outgoing' wave modes. Using Eq. (A.4)-(A.6), we find:

$$\begin{aligned} u_\alpha^{\text{in}\to\text{out}}(\mathbf{r}) = & \frac{\hat{s}_\alpha}{4\rho^0 (c^{\text{out}})^2}\sqrt{\frac{2}{\pi k^{\text{out}} r}} e^{ik^{\text{out}} r - i\pi/4} \\ & \times \int_S \left( \delta\rho(\mathbf{x})\omega^2 \hat{\mathbf{s}}\cdot\hat{\mathbf{p}} - k^{\text{in}} k^{\text{out}} \delta C_{ijkl}(\mathbf{x}) \hat{s}_i \hat{r}_j \hat{k}_k \hat{p}_l \right) \\ & \times e^{i(k^{\text{in}}\hat{\mathbf{k}} - k^{\text{out}}\hat{\mathbf{r}})\cdot\mathbf{x}} d^2x \end{aligned} \tag{A.7}$$

It is worth noting that Eq. (A.7) is valid for an arbitrary elastic perturbations $\delta C$. We now specialize to the isotropic case. By substituting the following expression:

$$\delta C_{ijkl}(\mathbf{x}) = \delta\lambda(\mathbf{x})\delta_{ij}\delta_{kl} + \delta\mu(\mathbf{x})(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{A.8}$$

into Eq. (A.7), computing the inner products of the polarization and wave propagation vectors with the elastic perturbation tensor and making slight re-arrangements, we obtain:

$$\begin{aligned} u_\alpha^{\text{in}\to\text{out}}(\mathbf{r}) = & \hat{s}_\alpha \sqrt{\frac{(k^{\text{out}})^3}{8\pi r}} e^{ik^{\text{out}} r - i\pi/4} \\ & \times \int_S \left[ \frac{\delta\rho(\mathbf{x})}{\rho^0} \hat{\mathbf{s}}\cdot\hat{\mathbf{p}} - \frac{\delta\lambda(\mathbf{x})}{\rho^0 c^{\text{in}} c^{\text{out}}} (\hat{\mathbf{s}}\cdot\hat{\mathbf{r}})(\hat{\mathbf{p}}\cdot\hat{\mathbf{k}}) \right. \\ & \left. - \frac{\delta\mu(\mathbf{x})}{\rho^0 c^{\text{in}} c^{\text{out}}} ((\hat{\mathbf{s}}\cdot\hat{\mathbf{k}})(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}) + (\hat{\mathbf{s}}\cdot\hat{\mathbf{p}})(\hat{\mathbf{r}}\cdot\hat{\mathbf{k}})) \right] \\ & e^{i(k^{\text{in}}\hat{\mathbf{k}} - k^{\text{out}}\hat{\mathbf{r}})\cdot\mathbf{x}} d^2x \end{aligned} \tag{A.9}$$

Eq.(A.9) is the most general form of Born's approximation in 2-D isotropic in-plane elastodynamics. In seismologcal applications, it is common to assume that certain correlations exist between the elastic parameters. To facilitate the application of this assumption, broadly known as 'Birch law' in the literature, we adopt $\{\rho, c_p, c_s\}$ as new independent variables. This is simply achieved by making the following substitutions: $\delta\lambda \to (c_p^2 - 2c_s^2)\delta\rho + 2\rho^0 c_p \delta c_p - 4\rho^0 c_s \delta c_s$, $\delta\mu \to c_s^2\delta\rho + 2\rho^0 c_s \delta c_s$ in Eq. (A.9). Furthermore, we make the assumption that the fluctuations of velocities and density may be described by a single zero-mean random function $\phi(\mathbf{x})$ with variance $\langle\varepsilon^2\rangle$. We may nevertheless allow for different level of

fluctuations for the density and velocities by writing:

$$\frac{\delta\rho(\mathbf{x})}{\rho_0} = \lambda_\rho \phi(\mathbf{x}) \ , \ \frac{\delta c_{p,s}(\mathbf{x})}{c_{p,s}} = \lambda_{c_{p,s}} \phi(\mathbf{x}). \tag{A.10}$$

In other words, the fluctuations of velocities and density are supposed to be perfectly correlated but have possibly different variances. This assumption allows us to factorize the expression of the scattered field as follows:

$$\begin{aligned} u_\alpha^{\text{in}\to\text{out}}(\mathbf{r}) =& \hat{s}_\alpha \sqrt{\frac{(k^{\text{out}})^3}{8\pi r}} X^{\text{in}\to\text{out}}(\theta) e^{ik^{\text{out}} r - i\pi/4} \\ & \times \int_S \phi(\mathbf{x}) e^{i(k^{\text{in}}\hat{\mathbf{k}} - k^{\text{out}}\hat{\mathbf{r}})\cdot\mathbf{x}} d^2x, \end{aligned} \tag{A.11}$$

where $\theta$ denotes the angle between $\hat{\mathbf{k}}$ and $\hat{\mathbf{r}}$ and $X^{\text{in}\to\text{out}}(\theta)$ are scattering patterns that depend solely on the velocities and density perturbations. As suggested by logging data (Wu et al., 1994), we will further assume that $P$ and $S$ velocity fluctuations share the same variance $\langle\varepsilon^2\rangle$, i.e. $\lambda_{c_s} = \lambda_{c_p} = 1$, and introduce a parameter $\nu = \lambda_\rho / \lambda_{c_{p,s}}$ that determines the relative amplitude of density and velocity perturbations. Below, we detail out the scattering patterns for all possible incoming and outgoing modes:

$$\begin{aligned} X^{p\to p}(\theta) =& \lambda_\rho(\cos\theta + 2\gamma^{-2}\sin^2\theta - 1) + 4\gamma^{-2}\lambda_{c_s}\sin^2\theta - 2\lambda_{c_p} \\ =& \left[\nu\cos\theta + (2+\nu)(2\gamma^{-2}\sin^2\theta - 1)\right] \\ X^{p\to s}(\theta) =& \lambda_\rho \sin\theta(2\gamma^{-1}\cos\theta - 1) + 4\lambda_{c_s}\gamma^{-1}\sin\theta\cos\theta \\ =& \sin\theta\left[2\gamma^{-1}(2+\nu)\cos\theta - \nu\right] \\ X^{s\to p}(\theta) =& \lambda_\rho \sin\theta(1 - 2\gamma^{-1}\cos\theta) - 4\lambda_{c_s}\gamma^{-1}\sin\theta\cos\theta \\ =& -\sin\theta\left[2\gamma^{-1}(2+\nu)\cos\theta - \nu\right] \\ X^{s\to s}(\theta) =& \lambda_\rho(\cos\theta + 2\cos^2\theta - 1) + 2\lambda_{c_s}(1 - 2\cos^2\theta) \\ =& \left[\nu\cos\theta + (\nu+2)(2\cos^2\theta - 1)\right] \end{aligned} \ , \tag{A.12}$$

where $\gamma = c_p/c_s$. We note that the scattering patterns involving mode conversions differ only by a sign as a consequence of reciprocity. For each of the scattering patterns, the first Eq. allows one to keep track of the contributions of each type of perturbation separately. It could be used to introduce more general hypotheses than the simple Birch law employed in our work. The next step in the derivation is the calculation of the mean-squared fields. Since all the terms in front of the integral in Eq. (A.9) are deterministic, the key is to evaluate the following multiple integral:

$$I = \int_S \int_S \langle\phi(\mathbf{x})\phi(\mathbf{y})\rangle e^{i(k^{\text{in}}\hat{\mathbf{k}} - k^{\text{out}}\hat{\mathbf{r}})\cdot\mathbf{x} - i(k^{\text{in}}\hat{\mathbf{k}} - k^{\text{out}}\hat{\mathbf{r}})\cdot\mathbf{y}} d^2x d^2y \tag{A.13}$$

Assuming that the random process $\phi$ is spatially homogeneous, we write the spatial correlation function:

$$\langle\phi(\mathbf{x})\phi(\mathbf{y})\rangle = C(\mathbf{x} - \mathbf{y}) \tag{A.14}$$

It is then natural to introduce a new set of barycentric coordinates such that:

$$\begin{aligned} \mathbf{x} =& \boldsymbol{\xi} + \boldsymbol{\Delta}/2 \\ \mathbf{y} =& \boldsymbol{\xi} - \boldsymbol{\Delta}/2 \end{aligned} \tag{A.15}$$

It may be verified that the determinant of the Jacobian of the transformation (A.15) equals 1. Assuming

that the set $S$ is convex, we may rewrite the integral $I$ as follows:

$$I = \int_S d^2\xi \int_{S_\Delta} C(\mathbf{\Delta}) e^{-i(k^{\text{in}}\hat{\mathbf{k}} - k^{\text{out}}\hat{\mathbf{r}})\cdot\mathbf{\Delta}} d^2\Delta, \tag{A.16}$$

where $S_\Delta$ is the domain of integration for the variable $\Delta$. We now require (1) that the typical linear dimension $L$ of the domain $S$ (and therefore also $S_\Delta$) be much larger than the correlation length $a$ of the fluctuations; (2) that $L$ is sufficiently small for the perturbative approach to apply. Physically, this requires the following scaling relation: $a \ll L \ll g_{p,s}^{-1}$, where $g_{p,s}$ is the total scattering coefficient of $P, S$ waves. This relation is known to break down in the high-frequency limit ($\omega \to \infty$). When assumptions (1) applies, we may extend the domain of integration of the variable $\Delta$ over the entire plane to obtain:

$$I \approx S\Phi(k^{\text{in}}\hat{\mathbf{k}} - k^{\text{out}}\hat{\mathbf{r}}) \tag{A.17}$$

where $\Phi$ is the power spectrum of the fluctuations, i.e., the Fourier transform of the spatial correlation function $C$. In the case of a statistically isotropic random medium, $\Phi$ depends solely on the modulus of its argument. As an illustration, in the popular case of a 2-D exponential random medium with correlation length $a$, one has:

$$\Phi(|\mathbf{m}|) = \frac{2\pi a^2 \langle \varepsilon^2 \rangle}{(1 + a^2\mathbf{m}^2)^{3/2}} \tag{A.18}$$

We now define the scattering coefficients $g^{\text{in}\to\text{out}}(\hat{\mathbf{r}}, \hat{\mathbf{k}})$ as the ensemble averaged energy scattered per unit time and unit angle into direction $\hat{\mathbf{r}}$, normalized by the incident energy flux density and the area $S$ of the inhomogeneous zone. This quantity has the unit of inverse length in 2-D and may be interpreted as the attenuation factor of the incident plane wave due to the presence of the inhomogeneities. In the case of a statistically isotropic medium, $g$ is a function of the angle $\theta$ between $\hat{\mathbf{r}}$ and $\hat{\mathbf{k}}$ only. In the far-field of the inhomogeneous zone, we may locally use a plane wave approximation for the scattered wave and write its flux of energy across an elementary line element $dl = rd\theta$ as follows:

$$\frac{dE^{\text{in}\to\text{out}}}{dt} = \frac{\rho^0 \omega^2 c^{\text{out}} \langle |\mathbf{u}^{\text{in}\to\text{out}}(\mathbf{r})|^2 \rangle dl}{2} \tag{A.19}$$

After normalization by the incident energy flux $J = \rho^0\omega^2 c^{\text{in}}/2$, angular aperture $d\theta$, and surface $S$, we find with the aid of the intermediate results (A.9), (A.17) and (A.19) the following expressions of the scattering coefficients for all possible mode conversions:

$$\begin{aligned}
g^{p\to p}(\theta) =& \frac{k_p^3 X^{p\to p}(\theta)^2}{8\pi} \Phi(2k_p \sin(\theta/2)) \\
g^{p\to s}(\theta) =& \frac{k_s^3 X^{p\to s}(\theta)^2}{8\pi\gamma} \Phi\left(\sqrt{k_p^2 - 2k_p k_s \cos(\theta) + k_s^2}\right) \\
g^{s\to p}(\theta) =& \frac{\gamma k_p^3 X^{s\to p}(\theta)^2}{8\pi} \Phi\left(\sqrt{k_p^2 - 2k_p k_s \cos(\theta) + k_s^2}\right) \\
g^{s\to s}(\theta) =& \frac{k_s^3 X^{s\to s}(\theta)^2}{8\pi} \Phi\left(2k_s \sin(\theta/2)\right)
\end{aligned} \tag{A.20}$$

# Acknowledgements

This thesis arose from the funding by China Scholarship Council (CSC) under the Freie Universität Berlin - China Scholarship Council Program 2018 (FUB-CSC Program 2018). I wrote this thesis while I was working at GFZ Potsdam and enrolled at the Freie Universität Berlin.

First of all, I would like to express my sincere gratitude to my primary supervisor, Christoph Sens-Schönfelder. This thesis would not have been possible without his patient and careful guidance. I was fortunate to have four wonderful years working with him and I really enjoyed this research journey. As a "problem" student, I made many mistakes during my research and always had endless questions to ask every week. He encouraged me not to worry about those mistakes and helped me correct them. He always patiently answered my stupid questions that I should have known the answers to, or explained to me, with a deep insight, some esoteric questions that almost reached the edge of knowledge. I was inspired by his love and ambition for science. Not only a knowledgeable supervisor, most of the time he was like a teammate, working with me to derive formulas, program, figure out solutions, discuss results, and revise my bad original manuscripts. He also encouraged me to be active in the community and provided me with many opportunities to participate in various academic activities. Lastly, I also express my admiration for his concerning on the climate issue.

I would like to sincerely thank my supervisor, Frederik Tilmann. I really appreciate the great opportunity that he and Christoph gave me to study at GFZ and FU. Four years at GFZ, as long as I had anything to ask him he would always take the time to help me out although he was quite busy. I really admire his professional and academic knowledge. When I finished the draft of my first paper, he carefully reviewed my manuscript and insightfully found an imperceptible but very critical problem. I feel really lucky with his guidance so that I am always on the right way.

I also greatly thank all coauthors who collaborated on all papers in this thesis which would not have been completed without their contributions. My first published paper was a great challenge in my PhD career. But thanks to Ludovic Margerin, his deep understanding of nature and mastery of mathematics helped us overcome many difficulties and we made it. Niklas Epple and Ernst Niederleithinger helped me extend my study to the engineering field, which greatly encouraged me to pursue more different challenges. Marcelo Bianchi and Klaus Bataille taught me how to deal with seismic data and brought me into the mysterious deep Earth. Thank all of them for the discussions and improvements of the manuscripts.

I would like to thank all members of my working group Sektion 2.4: Seismologie at GFZ, with whom I spent an unforgettable time. Special thanks to Xiaohui Yuan and his wife Xueqing Li. Although more than 10,000 kilometers from my hometown, they often made me feel at home. I also thank all postdocs and PhD students who helped me a lot no matter on academic or in life, including Yajian Gao, Wei Li, Felipe Orlando Vera Sanhueza, Stefan Mroczek, Kuan-Yu Ke, Peter Makus, Luc Illien, Jannes Münchmeyer, Felix Schneider, Yanling Liang, Yiming Bai, Jing Chen, Manuel Asnar, Laura Hillmann, Alexandra Mauerberger. I also thank our section's secretary, Liane Lauterjung as well as other scientists, Benjamin Heit, Joachim Saul, Peter Evans, Angelo Strollo. I really like the regular activities we have, Section Seminar, Seismology Science Chat and Seismo Running. Lastly, I also thank the discussion with Manuel Alejandro Jaimes Caballero and really enjoyed the times in Berlin with my friends, Chongzhou

Zhang, Yu You, Yalei Hu, Lingyu Liu, Yulan Su and Zezhenyang Shi. No matter how many thanks are not enough to my neighbors Weronika Malinowska and her boyfriend Felix Gonschorreck who really saved my life.

Finally, I would like to express my deep love for my family. The pandemic of COVID-19 had prevented me from normal working and living for a long time which almost occupied half time of my PhD career. This also prevented us from seeing each other for over three years. Thanks to my parents, Jun Zhang and Shuiying Peng, my grandma Shuhua Peng, and my girlfriend Yuanyi Sun for their massive love supporting me to last through the toughest times and also in memory of my grandpa Guanghua Peng who is always blessing me.

# Curriculum Vitae

For reasons of data protection, the curriculum vitae is not published in the electronic version.