Towards anisotropic and anelastic global adjoint models : Improving measurements and parametrization for global full-waveform inversion

Ridvan Örsvuran

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES FONDAMENTALES ET APPLIQUÉES

THÈSE DE DOCTORAT

Vers des modèles anisotropes et anélastiques de la Terre globale: Observables et la paramétrisation de l'inversion des formes d'ondes complètes

Rıdvan ÖRSVURAN

Géoazur

Présentée en vue de l'obtention du grade de docteur en Sciences de la Planète et de l'Univers
d'Université Côte d'Azur
Dirigée par : Ebru BOZDAĞ
Co-encadrée par : Stéphane OPERTO
Soutenue le : 31/03/2021

Devant le jury, composé de :
Jean Paul Ampuero, Directeur de Recherche IRD, Université Côte d'Azur
Andreas Fichtner, Assistant Professor, ETH Zürich
Frederik J Simons, Pr., Princeton University
Alessia Maggi, Pr., Université de Strasbourg
Jeroen Ritsema, Pr., University of Michigan
Guust Nolet, Pr. Émérite, Université Côte d'Azur

Vers des modèles anisotropes et anélastiques de la Terre globale: Observables et la paramétrisation de l'inversion des formes d'ondes complètes

Towards anisotropic and anelastic global adjoint models: Improving measurements and parametrization for global full-waveform inversion

Jury:

Président du jury
Jean-Paul Ampuero, Directeur de Recherche IRD, Université Côte d'Azur

Rapporteurs
Andreas Fichtner, Associate Professor, ETH Zürich
Frederik J. Simons, Professeur, Princeton University

Examinateurs
Alessia Maggi, Professeur, Université de Strasbourg
Jeroen Ritsema, Professeur, University of Michigan

Invités
Guust Nolet, Professeur Émérite, Université Côte d'Azur

Directeur de Thèse
Ebru Bozdağ, MCF, Université Côte d'Azur, Colorado School of Mines

Co-encadrée de Thèse
Stéphane Operto, Directeur de Recherche CNRS, Université Côte d'Azur

Acknowledgements

I would first like to thank my advisor Ebru Bozdağ for her guidance and support during my PhD. Her input was invaluable for all of the work that went into this thesis. I would also like to thank my co-advisor Stéphane Operto for sharing his scientific expertise and providing a valuable perspective.

I would like to thank Guust Nolet for his time and the helpful discussions. I would be remiss If I forgot to thank Dimitri Komatitsch whose work was essential for the numerical simulation that are performed in this thesis. I was also grateful have him in my thesis committee before his departure. I would also like to thank Daniel B. Peter, who is the current maintainer of *SPECFEM3D_GLOBE*, for his help.

I would like to thank Jeroen Tromp and his group in Princeton for including me in their meetings and discussions. It was quite beneficial especially at the start of my PhD. I would also like to thank Wenjie Lei, Youyi Ruan and Ryan Modrak for their aid navigating adjoint tomography workflow. I would like to thank Yanhua Yuan, for the discussions and help for the double-different misfit.

I would like to thank my friends Deniz Ertuncay and Ekrem Bekin. Although, we have parted ways after doing our undergraduate and master degrees together and ended up at different universities in different countries for our doctorate, their work and support was inspiring.

Lastly, I would like to express my gratitude to my family whom I am beholden to for all of my achievements. I thank my brothers Alaattin and Mustafa, my mother Tülin and my father Erdem.

Summary

Seismic waves are our primary tools to see the Earth's interior and draw inferences on its structural, thermal and chemical properties. Seismic tomography, similar to medical tomography, is a powerful technique to obtain 3D computed tomography scan (CT scan) images of the Earth's interior using seismic waves generated by seismic sources such as earthquakes, ambient noise or controlled explosions. It is crucial to improve the resolution of tomographic images to better understand the internal dynamics of our planet driven by the mantle convection, that directly control surface processes, such as plate tectonics. To this end, at the current resolution of seismic tomography, full physics of (an)elastic wave propagation must be taken into account.

The adjoint method is an efficient full-waveform inversion (FWI) technique to take 3D seismic wave propagation into account in tomography to construct high-resolution seismic images. In this thesis, I develop and demonstrate new measurements for global-scale adjoint inversions such as the implementation of double-difference traveltime and waveform misfits. Furthermore, I investigate different parameterizations to better capture Earth's physics in the inverse problem, such as addressing the azimuthal anisotropy and anelasticity in the Earth's mantle.

My results suggest that double-difference misfits applied to dense seismic networks speed up the convergence of FWI and help increase the resolution underneath station clusters. I further observe that double-difference measurements can also help reduce the bias in data coverage towards the cluster of stations.

Earth's lithosphere and upper mantle show significant evidence of anisotropy as a result of its composition and deformation. Starting from the recent global adjoint tomography model GLAD-M25, which is the successor of GLAD-M15 and transversely isotropic in the upper mantle, my goal is to construct an azimuthally anisotropic global model of the upper mantle. I performed 10 iterations using the multitaper traveltimes combined with double difference measurements made on paired stations of minor- and major-arc surface waves. The results after 10 iterations, in general, show the global anisotropic pattern consistent with plate motions and achieve higher resolution in areas with dense seismic coverage such as in North America and Europe.

Attenuation is also another key parameter for determining the partial melt, water content and thermal variations in the mantle. In the last chapter, I investigate anelastic adjoint inversions to ultimately construct a global attenuation mantle model by the simultaneous inversion of elastic and anelastic parameters assimilating both the phase and amplitude information, which will lead to exact FWI at the global scale. I investigate the trade-off between elastic and anelastic parameters based on 2D synthetic tests to define a strategy for 3D global FWIs. I also explore the effect of different measurements for simultaneously and sequentially inverted elastic and anelastic parameters. The 2D test results suggest that the envelope misfit performs best at earlier iterations by reducing the nonlinearity of the FWI. After analyzing the effect of different radially-symmetric attenuation models on seismic waveforms by performing forward simulations in various 1D and 3D elastic/anelastic models, the results suggest the necessity of simultaneous elastic/anelastic inversions to also improve

the elastic structure as attenuation cause not only amplitude anomalies but also significant physical dispersion, particularly on surface waves. I performed one global simultaneous iteration of elastic and anelastic parameters using GLAD-M25 and its 1D anelastic model QRF12 as the starting models with a dataset of 253 earthquakes. The preliminary results are promising depicting, for instance, the high and low attenuation in the West and East coasts of North America.

Keywords: Inverse theory, Computational seismology, Seismic tomography, Wave propagation

Résumé

Les ondes sismiques constituent le principal outil pour scanner l'intérieur de la Terre et en déduire des informations sur sa structure, son état thermique et ses propriétés chimiques. La tomographie sismique, de manière analogue à la tomographie médicale, construit des images en trois dimensions de l'intérieur de la Terre à partir des ondes sismiques déclenchées par des sources naturelles (tremblements de terre, bruit ambiant) ou contrôlées (explosions, ...).

La méthode de l'état adjoint permet une implémentation efficace de l'inversion des formes d'ondes complètes (FWI : Full Waveform Inversion), une méthode d'imagerie qui exploite potentiellement toute la complexité du champ d'onde en trois dimensions pour construire des images haute résolution de l'intérieur de la Terre.

Dans cette thèse, je propose de nouveaux observables fondés sur des doubles différences des temps de trajet et des formes d'ondes pour appliquer la FWI à l'échelle globale. Par ailleurs, je teste différentes paramétrisations du problème inverse pour extraire les propriétés physiques de la Terre comme l'anisotropie azimuthale et l'atténuation dans le manteau.

Mes résultats suggèrent que les doubles différences utilisés avec des dispositifs denses de stations accélèrent la convergence de la FWI, améliore la résolution de l'imagerie sous le dispositif et réduisent les artefacts générés par la couverture hétérogène de la Terre par les données sismologiques.

Il est connu que la composition et la déformation de la lithosphère et du manteau supérieur génèrent de l'anisotropie lors de la propagation des ondes. En partant du modèle de Terre globale GLAD-M25 développé par tomographie adjointe, le successeur du modèle GLAD-M15 et en) et en paramétrant l'inversion avec une anisotropie transverse, j'ai construit un premier modèle global du manteau supérieur anisotrope. J'ai effectué 10 itérations de la FWI adjointe en sélectionnant par fenêtrage les temps de trajet des ondes de surface combinés avec des doubles différences formés par des paires de stations. Les résultats révèlent l'empreinte au premier ordre de l'anisotropie et une résolution accrue dans les régions bénéficiant d'une forte couverture comme en Amérique du Nord et en Europe.

L'atténuation est un autre paramètre physique clef pour identifier de la fusion partielle, la présence d'eau et cartographier des variations thermiques dans le manteau. Dans le dernier chapitre, j'effectue une première évaluation de la tomographie adjointe anélastique dans la perspective de construire un modèle d'atténuation du manteau par inversion conjointe des paramètres élastiques et anélastiques à partir de la phase et de l'amplitude des signaux. J'étudie les couplages entre les différentes classes de paramètres avec des tests synthétiques 2D afin de définir la meilleure stratégie pour la FWI anélastique à l'échelle globale. J'évalue également différents observables pour la reconstruction simultanée ou alternée des paramètres élastiques / anélastiques. Les tests 2D suggèrent qu'une fonction coût fondée sur l'enveloppe des signaux fournit les meilleurs résultats lors des premières itérations en réduisant la non linéarité de la FWI. Après avoir évalué l'empreinte de différents modèles d'atténuation sur les formes d'onde avec des simulations numériques dans différents

modèles élastiques/anélastiques 1D/3D, j'ai conclu que la reconstruction conjointe des paramètres élastiques et anélastiques était nécessaire car l'atténuation affecte non seulement l'amplitude mais génère aussi une dispersion significative, notamment des ondes de surface. J'ai effectué une itération de la tomographie adjointe élastique/anélastique à l'échelle globale en partant du modèle élastique GLAD-M25 et du modèle anélastique 1D QRF12 et en utilisant 253 tremblements de terre. Les résultats préliminaires sont prometteurs et révèlent par exemple des zones de forte et faible atténuation sur les côtes ouest et est de l'Amérique du Nord.

Mots clés: Théorie inverse, Sismologie Computationnelle, Sismique Tomographie, Propagation des Ondes

Contents

List of Figures

Chapter 1

Introduction

Most of our knowledge about Earth's deep interior is based on seismic waves generated by earthquakes that propagate through the Earth. Seismologists systematically analyze these waves to interpret Earth's dynamic processes and the thermochemical structure which are needed to explain surface processes such as plate tectonics, earthquakes and volcanic activities. Based on measurements made on observed and synthetic seismograms, seismologists also construct 3D computed tomography scan (CT scan) images of the Earth using a technique called "seismic tomography", similar to medical tomography to investigate human body. These images are constructed in terms of seismic parameters, such as compressional (P) and shear (S) wavespeeds, density, anisotropy and anelasticity. Classically, this process has been based on ray theory, where seismic waves are assumed to propagate along a ray path as in optics. However, finite-frequency effects can be important because of the band-limited data I use during the construction of tomographic images. Recent advances in numerical methods and high-performance computing (HPC) allow seismologists to increase the resolution of seismic images by taking the full complexity of wave propagation into account in seismic tomography by using 3D numerical wave simulations, which requires careful measurements and strategies for successful convergence of tomographic problems. In this thesis, I demonstrate new measurements and parameterizations for global-scale full-waveform inversion based on the adjoint methods and 3D seismic wave simulations, following the first-generation global adjoint tomography models GLAD-M15 (Bozdağ et al., 2016) and GLAD-M25 (Lei et al., 2020) models. Although the focus is on global-scale tomography to investigate the Earth's deep mantle, results can may also be extended to other scales (i.e., exploration to regional) and domains (i.e., medical imaging). To this end, I first demonstrate and implement double-difference measurements which were first introduced in adjoint tomography by Yuan et al. (2016), into global-scale full-waveform inversions after appropriate adjustments. I then use double-difference measurements for surface waves to construct an azimuthally anisotropic global upper-mantle model based on adjoint tomography. Following that, I show 2D numerical experiments to see how well one can retrieve the anelastic structure simultaneously with the elastic one by demonstrating different misfit functions and strategies. Lastly, I demonstrate elastic & anelastic simultaneous adjoint inversions at the global scale including amplitude information from seismic waveforms to perform full-waveform inversion in more exact sense which includes both phase and amplitude information. Finally I summarize the results and observations discussing the potential challenges and future directions.

1.1 Seismic tomography

1.1.1 Earth's interior from seismic observations

Studies for investigation of Earth's inner structure based on seismic observations began in early 20th century with the availability of data from global seismic instruments worldwide. The first-order discontinuities of Earth's interior, such as the core (Oldham, 1906), the Moho discontinuity (Mohorovicic, 1909), and the inner core (Lehmann, 1936), were identified using the traveltime information of body waves. Following the first discoveries of the Earth's interior, Jeffreys and Bullen (1940) compiled travel-time catalogs which allowed for the construction of wavespeed profiles as a function of radius. Such profiles allowed for constructing "radially symmetric" Earth models which represent the mean average of Earth's P- and S-wavespeed, and density structures.

First seismic tomographic models to image lateral variations in Earth's mantle came out in late 70s (e.g., Aki et al., 1977; Sengupta and Toksöz, 1977; Dziewonski et al., 1977). Since then seismic tomography has been a powerful tool to image the multi-scale structure of Earth's interior from reservoir-scale for hydrocarbon exploration with active sources (i.e., explosions) (e.g., Jarchow et al., 1994; Operto et al., 2006) to crustal to global scales to investigate the deep Earth's interior commonly with passive sources such as earthquakes and seismic ambient noise (e.g., Shapiro et al., 2005; Sager et al., 2020). It was clear from the early observations that the radially symmetric models were not fully able to explain the complexity of waveforms or traveltime & amplitude anomalies, where lateral variations in structure must be taken into account. However, radially-symmetric models of Earth's interior have been used as reference earth models up to now. For instance, PREM (Preliminary Reference Earth Model, Dziewonski and Anderson (1981)) was constructed using body-wave traveltimes and normal-mode data, IASP91 of Kennett and Engdahl (1991) and ak135 (Kennett et al., 1995) which were constructed based on body-wave traveltimes only, which explain specifically the teleseismic body-wave arrivals reasonably well. Either global or regional radially symmetric models are also still commonly being used in, for instance, earthquake location studies (e.g., Engdahl et al., 1998). However, recently developed algorithms use path corrections to incorporate the complexity of wave propagation due to lateral variations (e.g., Schweitzer, 2001; Ekström et al., 2012). Radially-symmetric models are also used to define 3D models as perturbations around them since the classical tomographic models are constructed based on the perturbation theory. However, recent advances in global seismic tomographic models (e.g., French and Romanowicz, 2014; Bozdağ et al., 2016), which will be discussed below, indicate the necessity to go beyond the classical methods where the models can be constructed iteratively taking the full complexity of wave propagation into account (e.g., Kustowski et al., 2008; Ritsema et al., 2011; Simmons et al., 2012).

1.1.2 Seismic data

Seismic data can be seismic waveforms or some secondary observables such as amplitudes, traveltimes of body waves, group and phase velocities of surface waves or normal-mode splitting functions. Depending on the objective of tomographic studies measurements of body waves, surface waves or normal mode splitting functions can be used separately or combined to improve the data coverage. For instance,

Woodhouse and Dziewonski (1984) attempted to map lateral variations of the upper mantle using Rayleigh and Love waves with periods longer than 135 s, whereas Montagner (1986) also used such long-period surface waves to investigate the 3D structure beneath the Indian Ocean, and Hilst and Spakman (1989) used P-wave delay times to image subduction below the Caribbean Plate.

Several recent global tomography studies have capitalized on the wealth of data over the past few decades, using a broad range of body and surface waves, and normal-mode observations to image either shear-wave (Vs) (e.g., Houser et al., 2008; Ritsema et al., 2011; Schaeffer and Lebedev, 2013; Chang et al., 2014), compressional-wave (Vp) (e.g., Boschi and Dziewonski, 2000; Li et al., 2008) or both bulk-sound (Vc) and Vs variations (e.g., Masters et al., 2000; Antolik et al., 2003). Most of our knowledge of density variations in the mantle relies on long-wavelength normal-mode studies (e.g., Ishii and Tromp, 1999, 2001; Trampert et al., 2004; Mosca et al., 2012; Koelemeijer et al., 2017).

1.1.3 Ray theory vs finite-frequency theory

Classically, tomographic studies are based on the ray theoretical approximation which assumes that seismic waves propagate along ray paths as in optics. However, the ray theory is a high-frequency approximation that whenever the scale-length of heterogeneities become smaller than the width of the first Fresnel zone it breaks down (Wang and Dahlen, 1995). Most of our understanding of Earth's interior is based on ray theory and it is still being preferred because of its ease of implementation to tomography. However, the finite-frequency effects have become important at the current resolution of seismic tomographic models that they have to be addressed properly in tomographic studies. After Woodhouse and Girnius (1982) showed the finite-frequency effects on surface waves, Marquering et al. (1999) showed that finite frequency body-wave sensitivities are counter-intuitively zero along the ray paths, which are also known as "banana-doughnut" kernels because of the shape of well-isolated P- and S-wave sensitivity kernels. Finite-frequency kernels are frequency dependent, thus in the high-frequency limit they converge to ray paths. Dahlen et al. (2000) and Hung et al. (2000) built upon and extended this theory.

To take finite-frequency effects into account in seismic tomography, Li and Romanowicz (1995) introduced the non-linear asymptotic coupling theory (NACT) to invert surface and shallow-turning body wave which led to more robust 3D structure while having similar computational time to path average approximation method. Using P-wave traveltimes, Montelli et al. (2004) showed the effect of finite-frequency theory compared to the ray theory in global seismic tomography which led to better identification of mantle plumes in tomographic models. Following that Hung et al. (2004) used a similar finite-frequency approach to image the Iceland hotspot. In all these studies finite-frequency effects were calculated using a radially-symmetric Earth models.

The potential bias might be introduced into modeling body and surface waves using ray theory may already be understood by looking at finite-frequency kernels which becomes more prominent gradually at longer periods. However, there are some contradictory observations in the literature if taking finite-frequency effects into account also improves the resolution of tomographic images (e.g., Ritzwoller et al., 2002; Zhou et al., 2004; Sieminski et al., 2004). Godfrey et al. (2019) reported that exact ray tracing, where a ray-tracing algorithm is used to define exact ray paths in 3D structures avoiding great-circle approximation, give better estimation of the phase than 2D finite-frequency kernels. Similar observations were also reported

by Bozdağ and Trampert (2008). These observations likely to indicate the necessity to compute finite-frequency kernels in fully 3D background models with lateral variations which requires the computation of finite-frequency kernels numerically (Tromp et al., 2005; Fichtner et al., 2006). A major consequence of full 3D kernels is that tomographic models then can be updated iteratively. On the other hand, another important factor hindering the advantage of finite-frequency kernels could be the uneven data coverage due to the distribution of seismic sources and stations on the globe which was also reported by Trampert and Spetzler (2006).

1.2 Full-Waveform Inversion

Synthetic seismograms can be computed either using semi-analytic methods such as normal-mode summation (e.g., Gilbert, 1970; Dahlen and Tromp, 1998), reflectivity methods (e.g., Fuchs, 1968), asymptotic techniques such as ray theory (e.g., Cerveny, 2001) or asymptotic coupling theory (e.g., Li and Tanimoto, 1993) or numerical methods such as finite difference (e.g., Igel and Weber, 1996) or finite elements (e.g., Toshinawa and Ohmachi, 1992; Bao et al., 1998). Semi-analytic methods are used commonly to simulate wave propagation in radially-symmetric models. To take lateral variations into account, normal-mode coupling (e.g., Park, 1986; Lognonné and Romanowicz, 1990; Capdeville et al., 2000) is used which becomes inefficient in case of large variations in lateral heterogeneities. Ray theory is commonly used for its ease of implementation and reasonable computational requirements, however, has limitations when the scale-length of heterogeneities becomes comparable to the wavelength of seismic signals. Numerical methods allow for modelling full wave propagation.

There are several methods for solving the wave equation numerically. Finite difference (e.g., Olsen, 1995; Igel and Weber, 1995, 1996; Thomas et al., 2000), pseudo-spectral (e.g., Tessmer et al., 1992; Carcione and Wang, 1993; Furumura et al., 1998a; Igel, 1999), finite-element (e.g., Toshinawa and Ohmachi, 1992; Bao et al., 1998) and spectral-element (Patera, 1984; Priolo et al., 1994) methods are the most commonly used in seismology. Finite-difference is preferred for its fast and easy implementation to seismic problems which is commonly used in exploration seismology or seismic-hazard studies (e.g. Saenger and Bohlen, 2004; Ren et al., 2014; Maeda et al., 2016). Pseudo-spectral elements have high accuracy, however, may have problems with periodic boundary conditions. Finite-elements have flexibility to mesh complex structures. To solve the wave equation at the global scale numerically, the biggest challenge is the cost of simulations. In addition, one has to take the rapid surface and internal topographic variations into account which is easier with finite-element type of methods. To this end, spectral-element method stands out for global seismic wave simulations in terms of combining the accuracy with the choice of basis functions and flexibility of finite-element methods (Komatitsch and Vilotte, 1998). The chosen Gauss-Lobatto-Legendre polynomials lead to a diagonal mass matrix which speed up spectral-element solution of the wave equation and makes global wave simulations feasible. Spectral-element simulations to seismic wave propagation started in 2D (e.g., Priolo et al., 1994) and extended to 3D regional (e.g., Komatitsch and Vilotte, 1998; Seriani, 1998; Komatitsch and Tromp, 1999; Komatitsch et al., 2004) and global (e.g., Komatitsch and Tromp, 2002a,b; Komatitsch et al., 2002; Capdeville et al., 2003; Chaljub et al., 2003) scales.

The ultimate goal in seismology has been to use 3D seismic waveforms in seismic imaging since the beginning of seismic tomographic studies. Tarantola (1984)

introduced the adjoint methods based on numerical simulations to seismic problems. Adjoint tomography, a full-waveform inversion technique, has the following advantages: 1) synthetic seismograms and data sensitivity kernels (Fréchet kernels) are computed in 3D background models, 2) any wiggle in seismograms can be used which increases the amount of information that goes into inversions, 3) any approximation is avoided other than the chosen numerical method to simulate wave propagation, 4) commonly used “crustal corrections" are avoided, which can potentially bias specifically the anisotropic mantle models, 5) seismic models can be iteratively updated until convergence. In the adjoint method, Fréchet derivatives of the model parameters are computed by the correlation between forward wavefield due to a regular seismic source and a reversed-time wavefield that is generated by the data residual at the receiver location (Chen et al., 2007). Secondary wavefield (due to data residuals) is often referred as the “adjoint wavefield" and the fictious source based on data residuals is called the “adjoint source" (Tromp et al., 2005). The adjoint method is also used outside of seismology (Gauthier et al., 1986; Plessix, 2006; Virieux and Operto, 2009), for instance, in medical imaging (e.g., Pratt et al., 2007; Bernard et al., 2017; Sandhu et al., 2017; Boehm et al., 2018; Bachmann and Tromp, 2020) and in fluid dynamics (e.g., Talagrand and Courtier, 1987; Kim et al., 2001; McNamara et al., 2004).

First examples of using spectral-element seismograms in global tomography are based on coupling with 2D asymptotic kernels (Lekić and Romanowicz, 2011; French and Romanowicz, 2014). After successful applications at regional (Tape et al., 2009) and continental scales (Fichtner et al., 2009; Zhu et al., 2015), first global adjoint tomography models (Bozdağ et al., 2016; Lei et al., 2020) finally came out more than 30 years after the theory was introduced to seismology by Tarantola (1984).

1.2.1 Measurements in Full Waveform Inversion

Misfit, or objective, functions are defined to measure distance between observed data and synthetics, which are seismic waveforms or their secondary observables in this case, computed for a chosen source and structural model. Seismic tomography is an optimization technique that aims to minimize the chosen misfit function. Ideally, one would like to have synthetics that explain observed data perfectly yielding zero misfit. However, in practice, this is not possible due to noisy data and insufficient data coverage as well as other challenges coming from the nature of chosen inverse method to tackle a nonlinear problem, such as linearized or nonlinear inversion techniques, etc. As mentioned above, the adjoint method provides an efficient way to incorporate 3D wave simulations in full-waveform inversions where misfit between observed and synthetic data is minimized iteratively.

Misfit functions play an important role in success of the inversions (Modrak et al., 2016) because of the nonlinear nature of seismic problems. In exploration geophysics, the term FWI addresses making use of full waveforms which is performed by using an L_2-norm waveform misfit between observed and synthetic data (Tarantola, 1984; Virieux and Operto, 2009). In general L_2 norm misfit functionals are used in seismological problems which are based on standard least squares approach. L_1-norm misfit functions are discussed in the context of exploration studies by Brossier et al. (2010) which shown to good alternative especially for decimated data sets and there are also studies that uses other distance such as optimal transport which is more resilient to time shift and noise (Engquist et al., 2016; Métivier et al., 2018). The waveform misfit has also been used in global seismology (Nolet et al., 1986; Li and Romanowicz, 1996; French and Romanowicz, 2014). The waveform misfit can

be applied to any wave train or complete data. However, one of the issues is that in case of using seismograms involving seismic phases with significantly different amplitudes, such as body and surface waves which can be the case specifically in earthquake seismology, the measurements can be dominated by the high-amplitude signals. To overcome this problem, it is common to window data where each wave package or train is normalized by its own energy (e.g., Li and Romanowicz, 1996). Although, it is straightforward to compute the waveform misfit to incorporate full waveforms in inversions it is prone to behave nonlinearly especially when the starting model is not close enough to the actual model (Prieux et al., 2013a). To mitigate the nonlinearity of the inverse problem, one can start the inversion using long period data and decrease the period content gradually (Ekström et al., 1997; Zhu et al., 2015; Pageot et al., 2013). This approach requires to have access to low-frequency or long-period data which is easier for earthquake-seismology studies with the availability of broadband data. There are several studies that uses this multiscale inversion approach (Bunks et al., 1995; Sirgue and Pratt, 2004; Fichtner et al., 2013b; Yuan et al., 2015).

To reduce nonlinearities, it is also common to separate the phase and amplitudes of waveforms and tackle the phase information first which is more linearly related to Earth's parameters (Woodhouse and Wong, 1986). Cross-correlation traveltime measurements are commonly used to obtain (Luo and Schuster, 1991; Marquering et al., 1999) Fréchet derivatives of the model parameters which provide robust measurements, however, require to use multiple measurement windows to highlight low-amplitude phases by normalizing each window by its own energy.

The counterpart measurement of amplitudes can also be computed based on cross-correlations (Ritsema et al., 2002). However, due to higher non-linearity of the amplitude measurements, they are harder incorporate in the inversions.

To capture the dispersive behaviour of seismic waves, frequency dependent traveltime and amplitude measurements, also known as multitaper measurements, are proposed by Laske and Masters (1996). Multitaper traveltime measurements were used by Tape et al. (2010) in a regional adjoint tomography study in Southern California, and by Bozdağ et al. (2016) and Lei et al. (2020) in global adjoint tomography. Some other earlier examples of adjoint tomography in earthquake seismology by, for instance, Fichtner et al. (2009) used time-frequency analysis which allowed for taking wavetrains into account without the loss of small amplitudes with an intrinsic weighting term introduced into the adjoint source. Bozdağ et al. (2011) proposed a similar misfit function based on instantaneous phase (IP) measurements which minimizes the number of windows to maximize the information from a single seismic trace. Rickers et al. (2012) used the IP misfit in a synthetic experiment to image mantle plumes, then applied for imaging the Icelandic plume (Rickers et al., 2013), since IP measurements highlight the scattered waves better which are generally ignored in cross-correlation measurements. Although instantaneous phase measurements are useful, they are also prone to cycle skip problems like any other phase measurement. Yuan et al. (2020) introduced a variation of the method called exponentiated phase to alleviate this problem, where the misfit is defined between the exponential terms of the phase rather than computing the difference directly between observed and synthetic instantaneous phases. Similarly, the misfit function can also be defined based on instantaneous amplitude difference (Bozdağ et al., 2011) which is also known as "envelope misfit" (Kristeková et al., 2006; Fichtner et al., 2008). Envelope measurements used by (Romanowicz, 1994b; Karaoğlu and Romanowicz, 2018a) to construct anelastic global mantle models in waveform inversions whereas Zhu et al. (2013) used multitaper traveltime and amplitude measurements to invert elastic and

anelastic models of Europe simultaneously.

Generally, measurements are made on observed and synthetic data recorded at a single station. Another way to perform traveltime measurements to minimize potential source uncertainities, such as origin times, differential measurements between different phases are also commonly used (Hung et al., 2000; Tromp et al., 2005). On the other hand, the Double-Difference (DD) technique, utilizes measurements on station or source pairs. It was first introduced by Waldhauser and Ellsworth (2000) to improve earthquake locations. Since then, a similar approach has been used in body wave tomographic studies (Zhang and Thurber, 2003; Zhang et al., 2004; Monteiller et al., 2005). Yuan et al. (2016) introduced the DD traveltime measurements into adjoint tomography which are shown to be less sensitive to uncertainties in source origin time and source time function. There are also variations of this concept which use a reference trace instead of receiver/source pairs (Choi and Alkhalifah, 2011; Zhang et al., 2016). Örsvuran et al. (2020) extended the DD measurements to global adjoint tomography by introducing a weighting term to the misfit function that balances dense station arrays. It was also shown that DD measurements speed up the convergence at densely sampled areas.

1.2.2 Parametrization in Full Waveform Inversion

3D numerical wave propagation solvers allow for taking the full complexity of wave propagation into account in the forward and adjoint simulations. However, for the success of the FWI the physics of the medium and the source must also be incorporated in the inverse problem through appropriate parametrization of model parameters (Prieux et al., 2013b,c). For instance, specifically in global seismology, isotropic representation of the upper-mantle may not be enough since Earth's upper mantle shows significant anisotropy. In case of performing global FWI including amplitude information, one has to address anelasticity in the inverse problem as well to reduce the well-known source-structure trade off.

Earth's mantle, specifically the upper mantle, shows significant anisotropy. In anisotropic media, wavespeeds change based on the direction of the propagation. Mineral structure and orientation can cause the discrepancy between body and surface wavespeeds for different orientations (Forsyth, 1975; Nicolas and Christensen, 1987). Anisotropy has been showed to have significant effects on the phase velocities and converted waves (Crampin, 1981). The elastic tensor has 21 independent parameters due to symmetries. It is desirable to invert for all 21 parameters, however, some simplifications have to be made as it is not possible to invert them uniquely simultaneously (Beller and Chevrot, 2020). In an isotropic medium, elastic tensor can be represented by two Lamé parameters (λ and μ). For a layered medium, there is a one distinct direction and other two are equivalent to each other (Thomsen, 1986). The symmetry axis for this type of medium can be in the vertical direction which is defined as vertical transversely isotropic medium (VTI) which is called transverse isotropy or radial anisotropy in global seismology. A transversely isotropic model is represented by five Love parameters (A, C, L, N, F) (Love, 1911). If the symmetry axis is different than the vertical direction, two additional parameters are required to define the symmetry axis. This type of parameterization is called tilted transverse isotropy (TTI) which is the general case of the VTI medium (Tsvankin et al., 2010). There can also be the azimuthal dependency of wavespeeds which occurs in the plane of two identical directions of transversely isotropic media (Forsyth, 1975).

In the crust, it has been shown that cracks can induce anisotropy (Crampin and Booth, 1985). In the upper mantle, anisotropy is attributed to the strain which causes

lattice preferred orientation (Nicolas and Christensen, 2013). The discrepancy between Rayleigh and Love waves can be modeled using transversely isotropic models (Dziewonski and Anderson, 1981). However, it has been shown that azimuthal variations also play an important role in Earth's mantle (Montagner, 1985). Both body waves and surface waves exhibit transverse isotropy and azimuthal anisotropy. Body-wave anisotropy has generally been studied using shear-wave splitting measurements for seismic phases like S, SKS, SKKS, etc. (Silver and Chan, 1991; Vinnik et al., 1992) and in laboratory environments (Babuška, 1984). Surface waves are preferred to map the anisotropy in the upper mantle because they provide better data coverage (Nataf et al., 1984; Montagner and Tanimoto, 1991). There have been several regional and global anisotropic models that take either the transverse isotropy (Kustowski et al., 2008; French and Romanowicz, 2014; Bozdağ et al., 2016; Lei et al., 2020) or azimuthal anisotropy (Montagner, 1985; Montagner and Tanimoto, 1991; Trampert and Heijst, 2002; Schaeffer et al., 2016; Zhu et al., 2020) into account. First-generation global adjoint tomography models (Bozdağ et al., 2016; Lei et al., 2020) have transverse isotropy confined to the upper mantle.

Attenuation is a measure of energy loss as the wave propagates. It is particularly sensitive to thermal variations, partial melt and water content which can provide additional constraints of the physical and chemical content of Earth's interior together with wavespeeds. Since attenuation affects both the phase and amplitude of seismic waveforms (Zhou, 2009; Ruan and Zhou, 2010, 2012), elastic focusing effects, source uncertainties have to be properly addressed in inversions using full waveforms. Amplitudes can be more sensitive to the uncertainties in source radiation pattern and instrument response. However, if amplitudes are measured properly to be used in inversions, they also provide excellent constrains not only on the anelastic structure but also on elastic heterogeneities (e.g., Laske and Masters, 1996). Therefore, the attenuation tomography is a harder problem which is the main reason of fewer attenuation studies compared to elastic models. There are several radially symmetric attenuation models which do not show consensus especially in the crust and upper mantle (Durek and Ekström, 1996; Resovsky et al., 2005). The lateral variations in mantle attenuation were mapped using surface wave amplitudes (Romanowicz, 1995; Selby and Woodhouse, 2002; Gung and Romanowicz, 2004; Dalton and Ekström, 2006; Karaoğlu and Romanowicz, 2018b) and to a lesser extent by body waveforms (e.g., Bhattacharyya et al., 1996; Reid et al., 2001; Warren & Shearer, 2002). It is an important parameter for investigating the properties of the Earth's mantle which can provide additional constraints (Dalton and Faul, 2010; Cobden et al., 2018; Debayle et al., 2020). Karaoğlu and Romanowicz (2018a) constructed a global attenuation mantle model by inverting elastic shear wavespeed and anelastic models sequentially using 3D numerical wave simulations coupled with 2D asymptotic kernels. Anelastic adjoint inversions in earthquake seismology so far have been performed only at the European scale by Zhu et al. (2015). Attenuation is also used in exploration geophysics due to being sensitive to fluid content which can help processes like monitoring gas storage (Carcione et al., 2006; Métivier et al., 2015).

On the other hand, source effects can be comparable to path effects in global seismology (Ferreira and Woodhouse, 2007). Ideally source and structure should be updated simultaneously (Valentine and Woodhouse, 2010a). Kim et al. (2011) presented the adjoint source inversions and demonstrated the cost of such inversions compared to inversions based on computation of Green's functions of source parameters (i.e., moment tensor, origin time, depth, latitude, longitude) in full 3D Earth models. Nevertheless, it is not straightforward to simultaneously invert source and structure simultaneously because of the parameter trade-offs.

1.2.3 Potential Challenges

Seismic data has grown exponentially in recent years (e.g., IRIS DMC Data Statistics, 2020). However, the global data coverage is sparse because of the uneven source and station distribution. Earthquakes occur mostly on the plate boundaries and stations are mostly located on lands. We do not have enough sampling of the structure beneath the oceans which cover more than 70% of the surface area. Additionally, although dense networks like USArray in North America and Hi-Net in Japan provide great opportunities to have high-resolution lithospheric and upper mantle structure underneath continents, they disturb the balance of the measurement distribution further. The ultimate solution would be to cover the oceans with ocean-bottom seismometers (OBS). Meanwhile emerging datasets such as those from acoustic floating robots MERMAIDs (Nolet et al., 2019) could help better sample oceans. Furthermore, underwater fiber-optic cables might help improve global data coverage in future (Marra et al., 2018).

One possible solution to balance the global data coverage is to reduce the number of stations used from dense seismic networks avoiding redundant data. However, in this way potentially useful information might be missed. In order to mitigate uneven data coverage, some applications used summary rays which are defined as one average ray from a cluster to a station (Morelli and Dziewonski, 1987). Another technique has been to use a composite ray approach which uses the weighted average of single rays (Spakman, 1991; Bijwaard et al., 1998). Li and Romanowicz (1996), on the other hand, used error characterization and redundancy information to introduce a weighting term to the measurements. Ruan et al. (2019) used geographical distributions of source and stations to determine weighting terms introduced into a chosen misfit function to balance the data coverage.

Using ambient noise as seismic data in inversions is another promising technique to improve data coverage which potentially overcomes the global earthquake distribution problem (Shapiro et al., 2005; Lu et al., 2018). To this end, noise FWI is promising to specifically improve the data coverage in the lithosphere and upper mantle (Tromp et al., 2010a; Sager et al., 2020). There are also exciting results to extract body waves (including core phases) from ambient noise data (Gerstoft et al., 2008; Landès et al., 2010; Poli et al., 2012) which may in future can be used to sample the lower mantle.

Uncertainty quantification and resolution analysis have been the major problems for seismic tomography in general, which becomes more challenging for FWI because of the computational requirements. Resolution and uncertainty quantification are critical for the interpretation of the robustness of the tomographic models and potential bias and error introduced. Generally, checkerboard tests are used in classical tomographic studies to assess the resolution power of data coverage. However there are also concerns on the robustness of checkerboard tests which are shown to be susceptible to misleading results (Lévêque et al., 1993). One should ideally perform Bayesian type probabilistic inversions. However, they are not possible in FWI due their computational cost. Therefore, resolution analyses mostly done using approximations of the Hessian kernel (Fichtner and Trampert, 2011b; Fichtner and van Leeuwen, 2015) based on Point-Spread Function (PSF) tests. There are also optimization techniques which allow for approximating the Hessian kernel (Liu and Peter, 2019). Fichtner et al. (2019) also introduced a probabilistic Hamiltonian Monte Carlo solutions for the adjoint tomography.

FWI is a computationally intensive technique. Numerical simulations are generally performed for each source independent from number of stations. To reduce

the computational cost, van Herwaarden et al. (2020) uses a quasi-random subset of the sources in each iteration. Another approach is to use source stacking (Capdeville et al., 2005; Romanowicz et al., 2020). Source encoding method, which uses combination of sources to so-called supersources, is another method to reduce the computational cost (Krebs et al., 2009; Ben-Hadj-Ali et al., 2011; Castellanos et al., 2015; Tromp and Bachmann, 2019). Recent improvements reduce the downsides of this method which can lead to more adoption in the field. For example, crosstalk caused by the combination of the sources, and requirement for the all of sources to be recorded at all stations have been overcome by the recent studies (Dai et al., 2013; Krebs et al., 2018; Huang and Schuster, 2018; Zhang et al., 2018).

Numerical solvers like SPECFEM3D_GLOBE (Komatitsch and Tromp, 2002b), are able to capture full complexity of wave propagation accurately and efficiently for a given 3D global Earth model including anisotropy, anelasticity, tomography, bathymetry, gravity, rotation and the ellipticity of the planet. Most expensive part of wave simulations is coming from the crust. To speed up specifically the global simulations, homogenization (Capdeville et al., 2010a,b) is preferred by some researchers, where the crust is replaced by its smoothly varying anisotropic long-wavelength equivalent (Lekić and Romanowicz, 2011; French and Romanowicz, 2014). Homogenization is also used to aid multi-scale approach where a region of the model is finely meshed and upscaled for the coarsely meshed whole model decreasing the overall computational cost (Fichtner et al., 2013a). The AxiSEM3D package deals with this problem considering wave propagation on 2D axi-symmetric planes and analytically computing the azimuthal third dimension which significantly decreases the computational cost (Nissen-Meyer et al., 2014). Successful benchmarks of AxiSEM3D against SPECFEM3D_GLOBE simulations can be found in Leng et al. (2019). Another approach is to adapt meshes to the expected complexity of the wavefield (van Driel et al., 2020; Thrastarson et al., 2020).

1.3 Thesis Outline

In this thesis, I define strategies for global FWI by revising and demonstrating new misfit misfits, such as the implementation of double-difference measurements, and work towards constructing next-generation global adjoint tomography models including azimuthal anisotropy and anelasticity.

In **Chapter 2**, I give a brief overview of 2D and 3D spectral-element wave simulations and the adjoint method and the adjoint tomography workflow I used in this thesis. In **Chapter 3**, I present adaptation of of double-difference (DD) measurements to global adjoint tomography by defining misfit functions with appropriate weigthing terms to balance uneven data coverage. In **Chapter 4** I use double-difference multitaper traveltime measurements for minor- and major-arc surface waves to construct a global azimuthally anisotropic upper-mantle model. In **Chapter 5** I demonstrate anelastic adjoint inversions based on 2D examples to define a strategy for global-scale adjoint inversions. I test the effect of chosen misfit functions and simultaneous and sequential inversion of elastic and anelastic models. I also present a test iteration of simultaneous elastic and anelastic global adjoint tomography in 3D. Finally, I discuss the results in **Chapter 6**.

Chapter 2

3D wave simulations and full-waveform inversion

The inverse problem relies on the accuracy of the forward modeling. In this thesis, the ultimate goal is to improve the images of the Earth's mantle to gain more insight into the dynamics of our planet using 3D numerical waveform modeling of seismic waves by capturing the full complexity of wave propagation in realistic 3D media. Using 3D wave simulations in seismic imaging has the following major consequences: 1) we can avoid approximations commonly used for seismic wave propagation, such as ray theory, to take the full complexity of wave propagation into account which has become prominent at the current resolution of seismic images, 2) we can avoid commonly used crustal corrections that may potentially bias our inference of the upper mantle structure and anisotropy, 3) we can make use of any wiggles in seismograms with appropriately designed measurement techniques.

In this chapter I give brief background information on 3D numerical wave simulations, the adjoint method and measurements for full-waveform inversions. All the numerical simulations in this thesis were performed with the SPECFEM2D and SPECFEM3D_GLOBE (Komatitsch and Tromp, 2002a; Komatitsch et al., 2002) packages which are freely available from the CIG (Computational Infrastructure for Geodynamics) website.

2.1 Seismological background

The wave equation governs the behavior of seismic waves emanating due to an active (i.e., explosions) or passive (i.e., earthquakes, ambient noise) seismic source. The equation of motion can be expressed in terms of the source and medium properties as

$$\rho \delta_t^2 \mathbf{s} = \nabla \cdot \mathbf{T} + \mathbf{f} \,, \tag{2.1}$$

where ρ denotes the density of the medium, $\mathbf{s}$ is the displacement vector, $\mathbf{T}$ is the stress tensor and $\mathbf{f}$ is the source term. The stress tensor $\mathbf{T}$ is related to the gradient of the displacement, which is the strain tensor, by Hooke's law

$$\mathbf{T} = \mathbf{c} : \nabla \mathbf{s} \,, \tag{2.2}$$

where $\mathbf{c}$ is the fourth-order elasticity tensor. Due to the internal symmetries any anisotropic medium can be described by 21 unique parameters. For instance, an isotropic medium can be represented by only two Lamé parameters: Lamé constant λ and shear modulus μ. A transversely isotropic medium can be represented by 5

Love parameters, A, C, F, L, N (Love, 1911). Depending on the medium's characteristics, different parameterizations may also be considered such as the azimuthal dependence of surface wave velocities (Tanimoto and Anderson, 1985), or tilted transversely isotropy (TTI), where the symmetry axis is changed from the vertical, commonly used in exploration geophysics (Tsvankin et al., 2010). Recently, Beller and Chevrot (2020) used a projection to estimate the elastic coefficients instead of explicitly defining the symmetry axes. All these simplifications are done mainly because 21 anisotropic parameters cannot be uniquely inverted.

The source term is an earthquake for the purpose of this thesis, which can be given by

$$\mathbf{f} = \mathbf{M} \cdot \nabla S(t) \delta(\mathbf{x} - \mathbf{x_s}) \,, \tag{2.3}$$

where $\mathbf{M}$ and $S(t)$ are the moment tensor and the source time function as a function of time t for an earthquake located at $\mathbf{x}_s$. To solve the partial differential equation the surface boundary condition is defined as

$$\hat{\mathbf{n}} \cdot \mathbf{T} = 0 \tag{2.4}$$

where the traction goes to zero at the free surface and the initial conditions are set to

$$\mathbf{s}(\mathbf{x}, 0) = 0 \,, \tag{2.5}$$

and

$$\partial_t \mathbf{s}(\mathbf{x}, 0) = 0 \,. \tag{2.6}$$

The wave equation can be solved analytically for the simplest media like homogeneous whole space and half space models or the Lamb's problem. However, for full-waveform modeling only numerical solutions are available for any arbitrary 3D medium with lateral variations.

2.2 Numerical Solution to the Wave Equation

The finite difference is one of the most commonly used methods in seismology which incorporates the finite-difference approximation of derivatives, which is also known as the strong formulation (e.g. Olsen, 1995; Igel and Weber, 1996). Because of its ease of implementation and computational efficiency it is specifically preferred in seismic exploration problems (Saenger and Bohlen, 2004; Virieux and Operto, 2009; Chaljub et al., 2010; Ren et al., 2014). As the complexity of the medium increases, the run time of computations may also significantly increase because of the challenges related to the mesh. In addition, specifically surface waves may suffer from numerical dispersion. First applications of the finite-difference method used layered 2D mediums (Alterman and Karal, 1968; Boore et al., 1971). Virieux and Madariaga (1982) studied the dynamic shear cracks using more efficient staggered schemes both in space and time. The velocity-stress formulation is also studied in 2D both for SH and P-SV wave propagation (Virieux, 1984, 1986). Frankel and Vidale (1992) and Graves (1996) used 3D finite-difference modeling to generate synthetic seismograms which showed great promise for the future work. Igel et al. (1995) introduced anisotropy to finite difference schemes. Attenuation can also be incorporated to the finite difference via viscoacoustic and viscoelastic parametrization (Day and Minster, 1984; Robertsson et al., 1994; Bohlen, 2002; Yang et al., 2016).

Pseudo-spectral methods (e.g., Tessmer et al., 1992; Furumura et al., 1998b; Igel, 1999) have high accuracy to solve the wave equation which is desirable to accurately simulate seismic wave propagation. However, it is specifically challenging to deal with boundary conditions with reasonable run times. Another commonly used method in geophysical problems is the finite-element method (FEM) (e.g., Toshinawa and Ohmachi, 1992; Bao et al., 1998) which has the flexibility to mesh complex domains with finite elements. FEM was developed for engineering problems such as structural analysis. A large problem is subdivided into smaller finite elements. The solution for each element is assembled into a complete solution via common points between elements. Finite-element methods rely on the weak formulation of the partial differential equations. For the wave equation, displacement field is defined in terms of selected basis functions which leads to an approximation but a continuous solution where the displacement values are exact at the grid points. The weak formulation is implemented by multiplying the partial differential equation by a test function (with the same basis) and integrating over the whole spatial domain. The solution requires solving for the mass and stiffness matrices and the mass matrix should be inverted which can be computationally costly. One of the distinct advantages of FEM is that there is no additional step needed for the free surface boundary condition. It also allows for the use of hexahedral and tetrahedral meshes which are efficient to capture lateral heterogeneities. There are also examples of hybrid modeling with the finite-difference method to capture boundary conditions and complex topography (Moczo et al., 1997; Ma et al., 2004; Galis et al., 2008; Tarrass et al., 2011; de la Puente et al., 2014). However, increasing the polynomial degree to increase the accuracy may also significantly increase the computation time which still makes FEM unfeasible for simulating the global wave propagation. The spectral-element method, on the other hand, combines the accuracy of the pseudo-spectral method (Patera, 1984; Priolo et al., 1994; Komatitsch and Vilotte, 1998) and the meshing flexibility of the finite-element method (Komatitsch and Tromp, 1999). The Gauss-Lobatto-Legendre quadrature is used as the interpolation scheme which leads to a diagonal mass matrix. This property of the spectral-element method makes large-scale simulations, such as the global wave propagation, feasible. It also shares the free surface boundary advantage of FEM. There are several publicly available tools that uses the spectral-element method (Komatitsch and Tromp, 2002b,a; Fichtner and Igel, 2008). Nissen-Meyer et al. (2014) introduced AxiSEM3D which compute 3D wavefields in 2D axisymmetric global Earth's models while the azimuthal dimension is accounted for analytically, which reduces computational cost. Due to the efficiency of the spectral-element method, it has been extensively used in seismological applications (Komatitsch et al., 2004; Capdeville et al., 2005; Chaljub et al., 2007; Fichtner and Igel, 2008; Tape et al., 2010; Fichtner et al., 2013b; French and Romanowicz, 2014; Bozdağ et al., 2016; Lei et al., 2020).

There are also other numerical methods, like the Discontinuous Galerkin method, which is proven to be more efficient than spectral elements for certain applications (Igel, 2017). It has been used for modeling complex structures such as fluid-solid interfaces, faults and surface and internal topographies (Käser and Dumbser, 2008; Gallovič et al., 2010). It has been also used to study earthquake dynamics (e.g. dynamic rupture) (Käser et al., 2007; Puente et al., 2009; Pelties et al., 2012; Tago et al., 2012). In this thesis, 2D and 3D spectral-element methods are used for the numerical simulations.

2.2.1 The Spectral Element Method for global wave simulations

The spectral element method is based on Legendre polynomials and Gauss-Lobatto-Legendre quadrature (Komatitsch and Vilotte, 1998). This factorization results in the diagonal mass-matrix which increases the computational efficiency and parallelization capabilities. With the advances in computational resources, the method made large-scale simulations feasible and was eventually extended to global-scale seismic wave simulations (Komatitsch and Tromp, 2002b,a). In this thesis, I use the 2D and 3D global spectral element solvers SPECFEM2D (Komatitsch and Vilotte, 1998) and SPECFEM3D_GLOBE (Komatitsch and Tromp, 2002a; Komatitsch et al., 2002) for numerical wave simulations.

The spectral element method uses the integral or the *weak* formulation of the wave equation,

$$\int_V \rho \mathbf{w} \cdot \delta_t^2 \mathbf{s} d^3\mathbf{r} = -\int_V \nabla \mathbf{w} : \mathbf{T} d^3\mathbf{r} + \mathbf{M} : \nabla \mathbf{w}(\mathbf{r}_s) S(t)\,, \tag{2.7}$$

where $\mathbf{w}$ is a test function, $\mathbf{M}$ is the moment tensor and $S(t)$ is the source time function defined over the volume V. The Hooke's law in an attenuating medium requires a modification to include the entire strain history:

$$\mathbf{T}(t) = \int_{-\infty}^{t} \delta_t \mathbf{c}(t-t') : \nabla \mathbf{s}(t') dt'\,. \tag{2.8}$$

Assuming that attenuation is constant over the seismic frequency band (Liu et al., 1976) it is represented by three standard linear solids (Carcione et al., 1988) in SPECFEM3D_GLOBE.

For global Earth simulations with the SPECFEM3D_GLOBE package, crust & mantle, outer core and inner core regions are solved separately and then combined using the boundary terms. This allows for decoupling the solid and fluid regions. The formulation of the crustal and mantle region is

$$\int_M \rho \mathbf{w} \cdot \delta_t^2 \mathbf{s} d^3\mathbf{r} = -\int_M \nabla \mathbf{w} : \mathbf{T} d^3\mathbf{r} + \mathbf{M} : \nabla \mathbf{w}(\mathbf{r}_s) S(t) + \int_{CMB} p \hat{\mathbf{n}} \cdot \mathbf{w} d^2\mathbf{r}\,, \tag{2.9}$$

where p is the fluid pressure and last term is the integral traction of $-p\hat{\mathbf{n}}$ in the fluid ($\hat{\mathbf{n}}$ is the unit outward normal on the core-mantle boundary). For the fluid outer core, the wave equation is solved as follows:

$$\int_{OC} \kappa^{-1} w \delta_t^2 \chi d^3\mathbf{r} = -\int_{OC} \rho^{-1} \nabla w \cdot \nabla \chi d^3\mathbf{r} + \int_{CMB} w \hat{\mathbf{n}} \cdot \delta_t \mathbf{s} d^2\mathbf{r} - \int_{ICB} w \hat{\mathbf{n}} \cdot \delta_t \mathbf{s} d^2\mathbf{r}\,, \tag{2.10}$$

where w is the scalar test function, κ is the adiabatic bulk modulus, and χ is the scalar potential. For the solid inner core equation becomes

$$\int_{IC} \rho \mathbf{w} \cdot \delta_t^2 \mathbf{s} d^3\mathbf{r} = -\int_{IC} \nabla \mathbf{w} : \mathbf{T} d^3\mathbf{r} - \int_{ICB} p \hat{\mathbf{n}} \cdot \mathbf{w} d^2\mathbf{r}\,. \tag{2.11}$$

2.3 The Adjoint Method

Using the reciprocity and the time reversible nature of the wave propagation, Tarantola (1984) introduced the adjoint method, a full-waveform inversion technique, into seismic inversions to invert for the structural and source parameters of the Earth

based on full waveform simulations. Full-waveform inversions has been widely used in exploration geophysics (Virieux and Operto, 2009), generally based on an acoustic wave approximation to reduce the computational cost. With the advances in computational facilities and the 3D wave propagation solvers adjoint inversions both in elastic and anelastic media have become a routine tool, also in earthquake seismology.

Following Tromp et al. (2005), finite frequency kernels and the adjoint method may be formulated using the Born approximation. Let's assume $\mathbf{\Theta}(\mathbf{x_r}, t, \mathbf{m})$ is a chosen metric defined to measure the difference between observed and synthetic data recorded at a receiver location $\mathbf{x_r}$, where synthetic data is calculated for model parameters $\mathbf{m}$ and t denotes time:

$$\chi(\mathbf{m}) = \sum_s \sum_r \int_0^T \mathbf{\Theta}(\mathbf{x_r}, t, \mathbf{m}) \, \mathrm{d}t \, . \tag{2.12}$$

The misfit function is simply summed over a group of sources, s, and receivers, r. The gradient of the misfit function $\chi(\mathbf{m})$ may be written as:

$$\delta\chi(\mathbf{m}) = \sum_s \sum_r \int_0^T \partial_{\mathbf{s}} \mathbf{\Theta}(\mathbf{x_r}, t, \mathbf{m}) \delta\mathbf{s}(t, \mathbf{m}) \, \mathrm{d}t \, , \tag{2.13}$$

where $\mathbf{s}$ denotes synthetic displacement seismograms, $\delta\mathbf{s}$ is the perturbations in synthetic displacement due to a change in model parameters $\mathbf{m}$. Following Tromp et al. (2005), $\delta\mathbf{s}$ may be written using the Born approximation (Hudson, 1977; Wu and Aki, 1985):

$$\begin{aligned} \delta\chi(\mathbf{m}) = \sum_s \sum_r \int_0^T \partial_{s_i} \mathbf{\Theta}(\mathbf{x_r}, t, \mathbf{m}) \int_0^T \int_V [\delta\rho(\mathbf{x}') G_{ij}(\mathbf{x_r}, \mathbf{x}'; t - t') \partial_{t'}^2 s_j(\mathbf{x}', t')] \\ + [\delta c_{jklm}(\mathbf{x}') \partial_k' G_{ij}(\mathbf{x_r}, \mathbf{x}'; t - t') \partial_l' s_m(\mathbf{x}', t')] \mathrm{d}^3\mathbf{x}' \mathrm{d}t' \mathrm{d}t \, . \end{aligned} \tag{2.14}$$

Using the reciprocity of the Green's function, a time-reversed displacement wavefield, called the "adjoint wavefield", $s^\dagger$, may be defined,

$$s_k^\dagger(\mathbf{x}', t') = \int_0^{t'} \int_V G_{ki}(\mathbf{x}', \mathbf{x_r}; t' - t) f_i^\dagger(\mathbf{x}, t) \mathrm{d}^3\mathbf{x} \mathrm{d}t \, , \tag{2.15}$$

which is initiated by an "adjoint source" based on the misfit between observed and synthetic data defined for every station component i and injected at the location of the receiver

$$f_i^\dagger(\mathbf{x}, t) = \sum_{s=1}^{M} \sum_{r=1}^{N} \partial_{s_i} \mathbf{\Theta}(\mathbf{x_r}, T - t, \mathbf{m}) \delta(\mathbf{x} - \mathbf{x_r}) \, . \tag{2.16}$$

The gradient of a misfit function can also be expressed in terms of perturbations in model parameters depending on the chosen parameterization in the inverse problem. For an isotropic model, the gradient may be written in terms of perturbations of P and S wavespeeds (α and β, respectively), and density (ρ)

$$\delta\chi(\mathbf{m}) = \int_V [K_\rho'(\mathbf{x}) \delta\ln\rho + K_\alpha'(\mathbf{x}) \delta\ln\alpha + K_\beta'(\mathbf{x}) \delta\ln\beta] \, \mathrm{d}^3\mathbf{x} \tag{2.17}$$

where the volume integral reduces to the surface integral in 2D experiments and the kernels for density, P and S wavespeeds (K_ρ', K_α and K_β, respectively) are given as

$$K_{\rho}^{'}(\mathbf{x}) = K_{\rho} + K_{\kappa} + K_{\mu}\,, \tag{2.18}$$

$$K_{\alpha}(\mathbf{x}) = 2\left(\frac{\kappa + 4/3\mu}{\kappa}\right) K_{\kappa}\,, \tag{2.19}$$

$$K_{\beta}(\mathbf{x}) = 2\left(K_{\mu} - \frac{4\mu}{3\kappa}K_{\kappa}\right)\,, \tag{2.20}$$

where κ and μ are the bulk and shear moduli, respectively, and K_{μ}, K_{κ} are the associated Fréchet kernels. Depending on the parameterization source, attenuation and anisotropic parameters can be added to the gradient as well. The kernels for isotropic parameters denoted by $K(\mathbf{x})$ can be computed as (Tromp et al., 2005)

$$K_{\rho}(\mathbf{x}) = -\int_0^T \rho(\mathbf{x})\mathbf{s}^{\dagger}(\mathbf{x}, T-t)\cdot\mathbf{s}(\mathbf{x},t)\,\mathrm{d}t\,, \tag{2.21}$$

$$K_{\kappa}(\mathbf{x}) = -\int_0^T \kappa(\mathbf{x})[\nabla\cdot\mathbf{s}^{\dagger}(\mathbf{x}, T-t)][\nabla\cdot\mathbf{s}(\mathbf{x},t)]\,\mathrm{d}t\,, \tag{2.22}$$

$$K_{\beta}(\mathbf{x}) = -\int_0^T 2\mu(\mathbf{x})\mathbf{D}^{\dagger}(\mathbf{x}, T-t) : \mathbf{D}(\mathbf{x}, T-t)\,\mathrm{d}t\,, \tag{2.23}$$

where $\mathbf{D}$ is the deviatoric strain and $\dagger$ denotes the adjoint counterpart of the deviatoric strain and displacement. As clearly seen from the equations, the data sensitivity (Fréchet) kernels can be computed by performing two numerical simulations for the forward and adjoint displacement wavefields and the deviatoric strain. Since the Green's functions in both forward and adjoint wavefields are the same (only the source terms differ), adjoint simulations can readily be performed by the numerical solver used for forward simulations. The kernels will strictly be dependent on the chosen misfit function and any normalization or weighting applied to the misfit function will be reflected in the adjoint source.

Similarly the transverse isotropy can be taken into account in the inverse problem by the following five parameters: density ρ, bulk sound speed c, vertically polarized S-wave speed β_v, horizontally polarized S-wave speed β_h and the dimensionless parameter η:

$$\delta\chi(\mathbf{m}) = \int_V [K_c(\mathbf{x})\delta\ln c + K_{\beta_v}(\mathbf{x})\delta\ln\beta_v + K_{\beta_h}(\mathbf{x})\delta\ln\beta_h + K_{\eta}(\mathbf{x})\delta\ln\eta]\,\mathrm{d}^3\mathbf{x}\,, \tag{2.24}$$

where the K_c, K_{β_v}, K_{β_h} and K_{η} are the Fréchet derivaties of the transversely isotropic model parameters. Instead of the bulk-sound speed one may prefer to use vertically and horizontally polarized P wavespeeds. In this thesis, following Zhu et al. (2012) and Bozdağ et al. (2016) the bulk-sound speed is preferred to reduce the trade-off between P and S wavespeeds (Kennett et al., 1998) where it is more challenging to determine the P-wave anisotropy.

For the azimuthally anisotropic global inversions, vertically polarized S-wavespeed β_v, horizontally polarized S-wavespeed β_h and the azimuthally anisotropic parameters G_c' and G_s', which are the normalized G_c and G_s by the shear modulus of the reference background model are used (Sieminski et al., 2007a; Zhu et al., 2015):

$$\delta\chi(\mathbf{m}) = \int_V [K_{\beta_v}(\mathbf{x})\delta\ln\beta_v + K_{\beta_h}(\mathbf{x})\delta\ln\beta_h + K_{G_c'}(\mathbf{x})\delta G_c' + K_{G_s'}(\mathbf{x})\delta G_s']\,\mathrm{d}^3\mathbf{x}\,. \tag{2.25}$$

In theory, one may also use the full elastic tensor (21 independent parameters)in the parameterization. However, practically, it is not possible to solve 21 parameters independently. Thus, it is common to focus on inverting the dominant parameters in a medium (Beller and Chevrot, 2020).

Attenuation is also another key parameter to constrain thermochemical variations and the water content inside the Earth (Romanowicz, 1995; Bhattacharyya et al., 1996; Reid et al., 2001; Dalton and Ekström, 2006; Dalton et al., 2008). In Chapter 5, attenuation is included in the parameterization to simultaneously invert for elastic and anelastic parameters. Following (Tromp et al., 2005), the gradient of a misfit function in terms of perturbations in attenuation Q^{-1}, where Q is the quality factor (inverse of attenuation), is given by

$$\delta\chi(\mathbf{m}) = \int_V K_\mu(\mathbf{x})\delta Q_\mu^{-1}(\mathbf{x})d\mathbf{x}\,. \tag{2.26}$$

μ denotes the shear attenuation where Q_κ^{-1} is ignored since the intrinsic attenuation is dominated by Q_μ^{-1}. Following Liu et al. (1976), it is assumed that the shear attenuation is constant over the seismic frequency band and the relation between the frequency dependent shear modulus $\mu(\omega)$ and Q_μ^{-1} may be given as

$$\mu(\omega) = \mu(\omega_0)[1 + (2/\pi)Q_\mu^{-1}\ln(|\omega|/\omega_0) - isgn(\omega)Q_\mu^{-1}]\,, \tag{2.27}$$

where ω_0 is the reference frequency. The change in shear modulus due to a change in shear attenuation would then be

$$\delta\mu(\omega) = \mu(\omega_0)[(2/\pi)\ln(|\omega|/\omega_0) - isgn(\omega)]\delta Q_\mu^{-1}\,, \tag{2.28}$$

where the first term $(2/\pi)\ln(|\omega|/\omega_0)$ captures the physical dispersion around a reference angular frequency ω_0. Then the anelastic adjoint wavefield may be computed by the anelastic adjoint source

$$\tilde{f}^\dagger(\mathbf{x}, t, \mathbf{m}) = \frac{1}{2\pi}\int_{-\infty}^{\infty}[(2/\pi)\ln(|\omega|/\omega_0) - isgn(\omega)]^* f^\dagger(\mathbf{x}, \omega, \mathbf{m})\ln(\mathrm{i}\omega t)\mathrm{d}\omega\,, \tag{2.29}$$

where $f^\dagger(\mathbf{x}, \omega, \mathbf{m})$ is the Fourier transform of the elastic adjoint source. As can be seen anelastic kernels can be computed by an anelastic adjoint source which is obtained by the Hilbert transform of the elastic adjoint source without modifying the 3D solver (Tromp et al., 2005; Bozdağ et al., 2011). However, with this approach one needs to perform two sets of numerical simulations to compute elastic and anelastic gradients.

Waveform discrepancies can also be caused by source parameters. Ideally, source and structural inversions need to be performed simultaneously (Valentine and Woodhouse, 2010b). However, in practice, it is challenging because of the potential trade-offs between seismic parameters. Source parameters (i.e., location, depth, moment tensor, origin times) may be updated by computing source Green's functions in 3D background models numerically (Liu et al., 2004). Alternatively, source parameters can also be incorporated into the adjoint tomography. For instance, Kim et al. (2011) explored centroid-moment tensor inversions using the adjoint method. One way to reduce the trade-off between source and structural parameters may be to apply source corrections, such as based on a grid search to adjust the origin time and the scalar moment at every iteration (e.g., Zhu et al., 2015; Lei et al., 2020). In this thesis,

I preferred the latter in global inversions where the inversion of source parameters needs further investigation.

2.3.1 Measurements in adjoint tomography

The choice of the misfit function closely affects the success of the inversion (Gee and Jordan, 1992; Brossier et al., 2010; Bozdağ et al., 2011). In this section, I summarize commonly used misfit functions in adjoint inversions which are also used in the rest of the thesis.

Waveform misfit

The waveform misfit based on the L_2 norm between synthetic and observed data may be given as (Tarantola, 1984)

$$\chi_{wf} = \frac{1}{2}\sum_{r=1}^{N}\int_0^T ||d_r(\mathbf{x},t) - s_r(\mathbf{x},t,\mathbf{m})||^2 \mathrm{d}t\,, \tag{2.30}$$

where the associated adjoint source may be obtained as (Tromp et al., 2005),

$$f^{\dagger}_{wf_i}(\mathbf{x},t) = -\sum_{r=1}^{N}\left[d_r(\mathbf{x}_r,T-t) - s_r(\mathbf{x},T-t,\mathbf{m})\right]\delta(\mathbf{x}-\mathbf{x}_r)\,. \tag{2.31}$$

The waveform misfit is commonly used in exploration geophysics (Virieux and Operto, 2009) and in some global tomographic studies as well (e.g., Li and Romanowicz, 1996). However, one needs to have a starting model close enough to the actual one to avoid nonlinearities.

Cross-correlation traveltime misfit

The cross-correaltion traveltime misfit is defined as

$$\chi_{tt} = \frac{1}{2}\sum_{r=1}^{N}||T_r^{obsd} - T_r(\mathbf{m})||^2 \mathrm{d}t\,. \tag{2.32}$$

The associated adjoint source then given as

$$f^{\dagger}_{cc_i}(\mathbf{x},t) = -\sum_{r=1}^{N}\Delta T_r\frac{1}{N_r}w_r(T-t)\delta_t s_i(\mathbf{x}_r,T-t,\mathbf{m})\delta(\mathbf{x}-\mathbf{x}_r)\,, \tag{2.33}$$

where w_r is the measurement window, ΔT_r the traveltime difference between observed and synthetic data, N_r is the normalization factor ($N_r = \int_0^T w_r(t)s_i(\mathbf{x}_r,T-t,\mathbf{m})\delta_t^2 s_i(\mathbf{x}_r,T-t,\mathbf{m})dt$), and $\delta_t s_i$ is the time derivative of the synthetic data (e.g., Luo and Schuster, 1991; Marquering et al., 1999; Dahlen et al., 2000). The shortcoming of this misfit is that one needs to use multiple measurement windows to maximize the information extracted from each time series. Because of the cross-correlation measurements naturally favor the main arrival where scattered waves may be suppressed. To capture the wave dispersion the frequency-dependent traveltime measurements are also preferred which are known as multitaper misfits (Laske and Masters, 1996; Tape et al., 2010).

Amplitude misfit

The cross-correlation amplitude misfit may be defined as the ratio between the observed and synthetic amplitudes (Ritsema et al., 2002)

$$\chi_{amp} = \frac{1}{2}\sum_{r=1}^{N}\left[\ln\frac{A_r^{obsd}}{A_r(\mathbf{m})}\right]^2 \mathrm{d}t\,. \tag{2.34}$$

The adjoint source then becomes

$$f^{\dagger}_{amp_i}(\mathbf{x},t) = -\sum_{r=1}^{N}\ln(\Delta A_r)\frac{1}{N_r}w_r(T-t)s_i(\mathbf{x}_r,T-t,\mathbf{m})\delta(\mathbf{x}-\mathbf{x}_r)\,, \tag{2.35}$$

where w is the time window and the N_r is the normalization factor ($N_r = \int_0^T w_r(t)s_i^2(\mathbf{x}_r, T-t, \mathbf{m})dt$). Similar to the cross-correlation traveltimes the amplitude misfit also requires multiple time windows to maximize the information from each seismic trace. Its multitaper version may help capture dispersive waves better (Tape, 2009).

Envelope and instantaneous phase misfits

Working in the Hilbert domain is efficient to separate the phase and amplitude measurements from each other which is desirable to better linearize the seismic inverse problem (e.g., Bozdağ et al., 2011). The Hilbert transform of an analytic signal f is defined as

$$\mathcal{H}\{f(t)\} = -\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{f(\tau)}{t-\tau}d\tau\,, \tag{2.36}$$

where P is the Cauchy principal value. The analytical signal $\tilde{f}(t)$ can then be written as

$$\tilde{f}(t) = f(t) - \mathcal{H}\{f(t)\}\,, \tag{2.37}$$

where $\mathcal{H}$ denotes the Hilbert transform which can be computed using the Fourier transform by setting the negative frequencies to zero:

$$\mathcal{H}\{f(t)\} = \mathcal{F}^{-1}\{\mathcal{F}\{f(t)\}(-isgn(\omega)\}\}\,, \tag{2.38}$$

where $\mathcal{F}$ is the forward, $\mathcal{F}^{-1}$ is the backward Fourier transforms and sgn denotes the signum function. The analytic signal can be written in terms of an amplitude and a phase term,

$$\tilde{f}(t) = E(t)\exp(i\phi(t))\,, \tag{2.39}$$

where $E(t)$ is the instantaneous amplitude (envelope) which can be defined based on the real and imaginary parts of the analytic signal

$$E(t) = \sqrt{\mathcal{R}\{\tilde{f}(t)\}^2 + \mathcal{I}\{\tilde{f}(t)\}^2}\,, \tag{2.40}$$

and $\phi(t)$ is the instantaneous phase given as

$$\phi(t) = \arctan\frac{\mathcal{I}\{\tilde{f}(t)\}}{\mathcal{R}\{\tilde{f}(t)\}}\,. \tag{2.41}$$

One can then define an **envelope misfit** based on the difference between the envelopes of observed and synthetic waveforms (Bozdağ et al., 2011)

$$\chi_{env} = \frac{1}{2}\sum_{r=1}^{N}\int_0^T \ln\left(\frac{E_i^{obs}(t)}{E_i(\mathbf{m},t)}\right)\mathrm{d}t\,. \tag{2.42}$$

The associated adjoint source for the envelope misfit can be obtained as

$$\begin{aligned} f^{\dagger}_{env_i}(\mathbf{x},t) = -\sum_{r=1}^{N}\Bigg[&\ln\left[\frac{E_r^{obs}(\mathbf{x}_r,t)}{E_r(\mathbf{x}_r,\mathbf{m},t)}\right]\frac{w_r(t)s_i(\mathbf{x}_r,T-t,\mathbf{m})}{E_i(\mathbf{x}_r,\mathbf{m},T-t)^2} \\ &-\mathcal{H}\left\{\ln\left[\frac{E_r^{obs}(\mathbf{x}_r,t)}{E_r(\mathbf{x}_r,\mathbf{m},t)}\right]\frac{w_r(t)\mathcal{H}\{s_i(\mathbf{x}_r,T-t,\mathbf{m})\}}{E_i(\mathbf{x}_r,\mathbf{m},T-t)^2}\right\}\Bigg]\delta(\mathbf{x}-\mathbf{x}_r)\,. \end{aligned} \tag{2.43}$$

The advantage of the envelope misfit is that it can be applied to any wave train and, in the ideal case to entire seismic trace without cutting seismograms into small windows because of the intrinsic normalization term by the square of envelopes (Bozdağ et al., 2011). A similar misfit was also proposed in the time-frequency domain by Kristeková et al. (2006) and Fichtner et al. (2009). Envelope misfits can be useful to increase the linear regime of the full waveform inversion especially when there is a lack of good starting models (Yuan et al., 2015). It has also been reported that envelope misfits may be a better choice in attenuation inversions (Karaoğlu and Romanowicz, 2018b) which is also discussed in the last chapter.

Similarly the **instantaneous phase misfit** may be defined as the difference between the instantaneous phase of observed and synthetic waveforms as

$$\chi_{ip} = \frac{1}{2}\sum_{r=1}^{N}\int_0^T \left[\phi_r^{obs}(t)-\phi_r(\mathbf{m},t)\right]^2\mathrm{d}t\,. \tag{2.44}$$

Then the associated instantaneous phase adjoint can be written as (Bozdağ et al., 2011)

$$\begin{aligned} f^{\dagger}_{ip_i}(\mathbf{x},t) = -\sum_{r=1}^{N}\Bigg[&\left[\phi_i^{obs}(\mathbf{x}_r,T-t)-\phi_i(\mathbf{x}_r,\mathbf{m},T-t)\right]\frac{w_r(T-t)\mathcal{H}\{s_i(\mathbf{x}_r,T-t,\mathbf{m})\}}{E_i(\mathbf{x}_r,\mathbf{m},T-t)^2} \\ &-\mathcal{H}\left\{\left[\phi_i^{obs}(\mathbf{x}_r,T-t)-\phi_i(\mathbf{x}_r,\mathbf{m},T-t)\right]\frac{w_r(T-t)s_i(\mathbf{x}_r,T-t,\mathbf{m})}{E_i(\mathbf{x}_r,\mathbf{m},T-t)^2}\right\}\Bigg]\delta(\mathbf{x}-\mathbf{x}_r)\,. \end{aligned} \tag{2.45}$$

Similar to the envelope misfit the instantaneous misfit can also be applied to any wave train successfully highlighting low amplitude signals. It has been shown that instantaneous phase measurements improve the resolution of mantle plumes compared to cross-correlation traveltime measurements by capturing more information from diffracted waves (Rickers et al., 2012). The instantaneous phase misfit may suffer from cycle skips similar to any other phase measurement which may make its implementation to adjoint inversions challenging.

Exponentiated phase misfit

The exponentiated phase misfit (Yuan et al., 2020) was introduced as an extension to the instantaneous phase misfit to alleviate cycle skip problems. To this end, the instantaneous misfit is revised by defining the misfit based on normalized analytic

signals which leads to difference between the exponential terms of observed and syntethic waveforms:

$$\chi_{ep} = \frac{1}{2}\sum_i \int_0^T ||\tilde{d}_i(\mathbf{x},t,\mathbf{m}) - \tilde{s}_i(\mathbf{x},t)||^2 \mathrm{d}t\,, \tag{2.46}$$

where $\tilde{d}_i = e^{i\phi(t)}$ and $\tilde{s}_i = e^{i\phi_s(t,\mathbf{m})}$ are the normalized analytical signals for observed and synthetic data, respectively, where ϕ and ϕ_s are the instantaneous phase of observed and syntehtic data, respectively. The associated adjoint source for the exponentiated misfit then becomes

$$\begin{aligned} f^{\dagger}_{ep_i}(\mathbf{x},t) = -\sum_{r=1}^{N} \Bigg[& \Delta I(t) \frac{w_r(t) s_i(\mathbf{x}_r,t,\mathbf{m}) \mathcal{H}\{s_i(\mathbf{x}_r,t,\mathbf{m})\}}{E_i(\mathbf{x}_r,\mathbf{m},t)^3} - \Delta R(t) \frac{w_r(t) \mathcal{H}\{s_i(\mathbf{x}_r,t,\mathbf{m})\}^2}{E_i(\mathbf{x}_r,\mathbf{m},t)^3} \\ & + \mathcal{H}\left\{ \Delta I(t) \frac{w_r(t) s_i(\mathbf{x}_r,t,\mathbf{m})^2}{E_i(\mathbf{x}_r,\mathbf{m},t)^3} - \Delta R(t) \frac{w_r(t) s_i(\mathbf{x}_r,t,\mathbf{m}) \mathcal{H}\{s_i(\mathbf{x}_r,t,\mathbf{m})\}}{E_i(\mathbf{x}_r,\mathbf{m},t)^3} \right\} \Bigg] \\ & \delta(\mathbf{x}-\mathbf{x}_r)\,, \end{aligned} \tag{2.47}$$

where $\Delta R(t)$ stands for differences in real part and $\Delta I(t)$ stands for the difference in the imaginary part of the analytic signal (Yuan et al., 2020).

Choice of the misfit function

Figure 2.1 shows adjoint kernels for a first arrival S wave on the transverse component computed for different misfit functions described above. The traveltime kernel shows a typical 'banana-doughnut' sensitivity (Marquering et al., 1999). The exponentiated phase misfit gives a similar sensitivity as it is also based on the phase information and there are not much observed scattered waves within the measurement window. The amplitude and envelope misfits show similar sensitivities whereas the waveform misfit looks more similar to traveltime kernels in this case. Each misfit function has advantages and disadvantages which are briefly discussed above. Therefore, they must be chosen carefully depending on the target problem. In full-waveform inversions, the goal is to use full phase and amplitude information to better constrain the structural and source parameters. However, to reduce the trade-off, specifically in earthquake seismology, it is common to use the phase information only when the elastic structure is targeted. Therefore, it is desirable to separate phase and amplitude information rather than directly start using the waveform misfit. This is the main motivation of the first-generation adjoint models in earthquake seismology which are based on traveltime or phase measurements (Tape et al., 2009; Fichtner et al., 2009; Zhu et al., 2012; Bozdağ et al., 2016; Lei et al., 2020). Amplitudes must be taken into account while constraining anelastic and source parameters (Karaoğlu and Romanowicz, 2018a,b; Kim et al., 2011). If amplitudes can be properly used, they can also provide invaluable constraints to locate elastic heterogeneities (Woodhouse and Wong, 1986; Laske and Masters, 1996).

2.3.2 Adjoint tomography workflow

The adjoint tomography is an iterative full-waveform inversion scheme. Figure 2.2 shows a typical adjoint tomography workflow used in this thesis. The adjoint tomography workflow has three basic stages: 1) Numerical simulations of forward

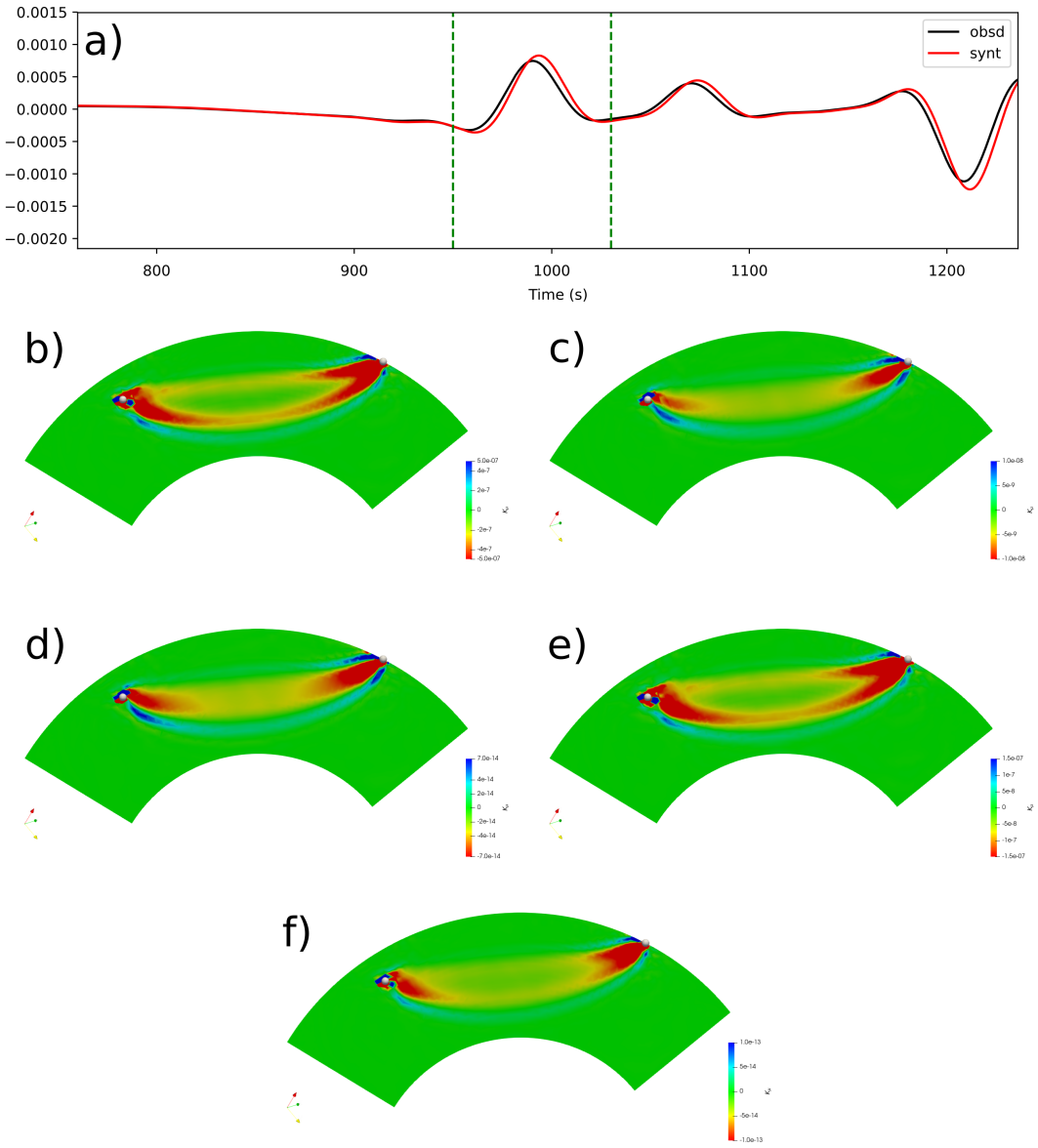

FIGURE 2.1: First arrival S-wave sensitivity kernels computed for the 1994 Bolivia Earthquake ($M_w = 8.2$) at a station witha n epicentral distance of 60°. (a) The selected time window showing S waveforms, (b) cross-correlation traveltime misfit sensitivity, (c) cross-correlation amplitude misfit sensitivity, (d) envelope misfit sensitivity, (e) exponentiated phase misfit sensitivity, (f) waveform misfit sensitivity.

and adjoint simulations, 2) the pre-processing stage (i.e., data processing, selecting measurement windows, computing adjoint sources), 3) the post-processing stage (i.e., summation of event kernels, smoothing and pre-conditioning the gradient, determining the search direction and the step length for model updates).

Pre-processing stage

Observed seismic data are obtained using public repositories such as *IRIS*, *ORFEUS* or local seismic networks. After the computation of synthetic data by a numerical solver, measurements are made between observed data and synthetic data based on a chosen misfit function.

I use the automated window selection algorithm developed by Maggi et al. (2009) to select the usable parts of seismic waveforms to compute adjoint sources. The selection criteria are defined according to the target problem. Figure 2.3 shows sample waveforms and selected windows for the minor- and major-arc Rayleigh waves used in the azimuthally anisotropic inversions in Chapter 4. In addition the effect of oceanic and continental crust on waveforms can be clearly seen on surface waves.

Adjoint sources can be computed using the measurement windows for the desired misfit function. If measurements have different categories, such as different period bands, different type of wave selection (body and surface wave) and different components, category weighting should be considered to obtain a balanced representation across categories. Also other weighting terms can be introduced for specific measurements (see Chapter 3) or to balance the geographical distribution (e.g., Ruan et al., 2019). Final adjoint sources can be computed by summing each individual category with their respective weights.

Numerical simulations

The iterations start by performing forward simulations to compute synthetic seismograms, which then used to make measurements compared to observed data to construct adjoint sources. Then, sensitivity kernels for model parameters are computed using adjoint sources performing adjoint simulations. These simulations are performed for each event (e.g., earthquake) in the database.

Post-processing stage

Figure 2.4 shows a sample event kernel for vertically polarized shear wavespeed. It is observed that raw kernels may have significant small scale fluctuations which can partly be due to the numerical noise which is generally eliminated by smoothing the gradient. The gradient is the summation of all event kernels. Kernels are then smoothed and preconditioned to speed up the convergence. Figure 2.4b shows a total gradient after the post processing steps.

Model updates are performed by updating old model based on a search direction (d_i).

$$m_{i+1} = m_i + \alpha d_i \ ,$$

where m_{i+1} is the new model, m_i is the current model α is the step length and d_i is the search direction. There are several optimization techniques for computing the search direction. Simplest case is called steepest descent which can be written as $d_i = -g_i$ where g_i is the gradient. The iterations typically start with steepest descent but the subsequent iterations generally involve more advanced techniques like Nonlinear

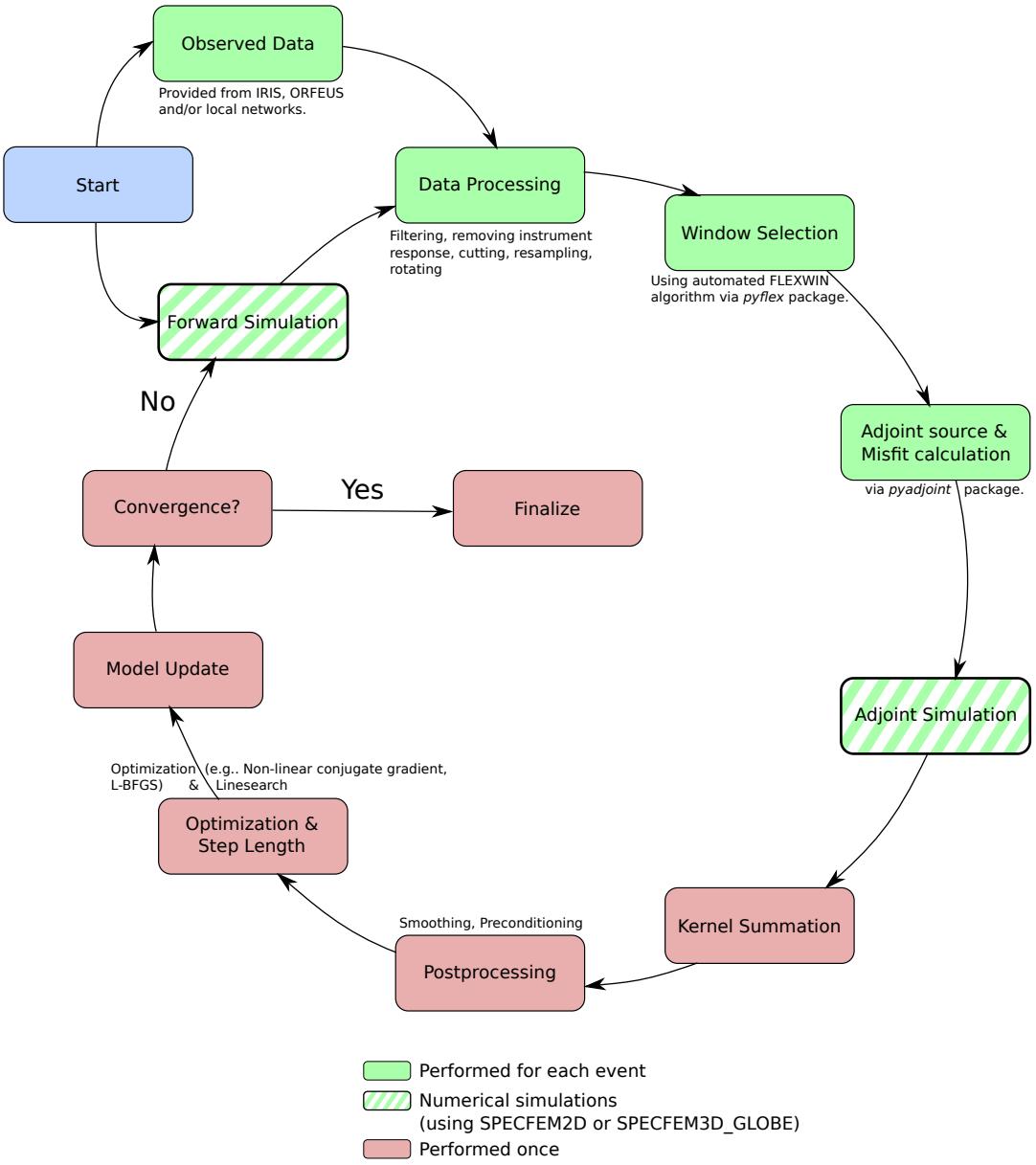

FIGURE 2.2: The adjoint tomography workflow used in this thesis. Iterative iteration steps are repeated until convergence.

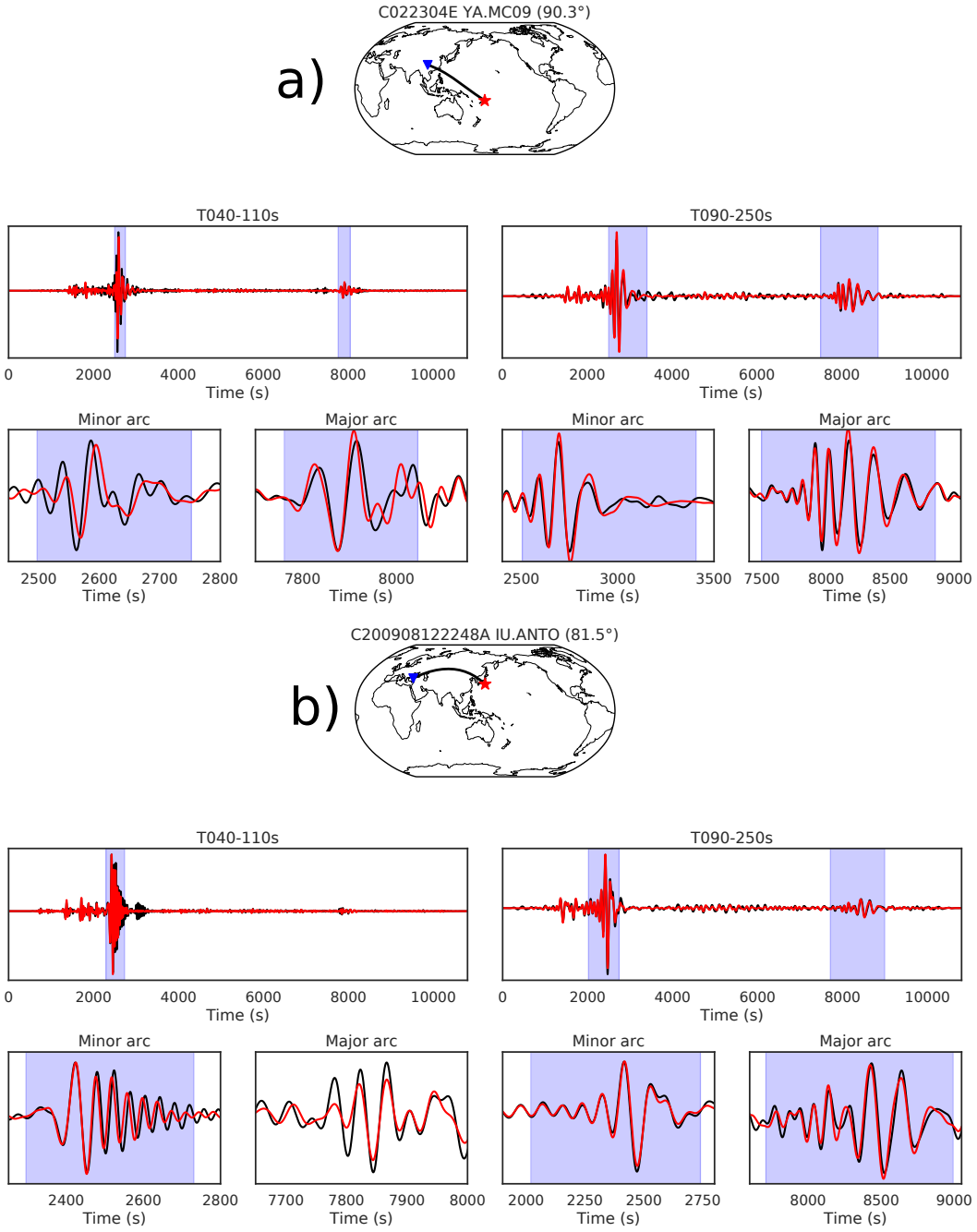

FIGURE 2.3: Sample seismic waveforms recorded on vertical components of two stations for two period bands. The blue windows denote the measurement windows automatically selected by FLEXWIN (Maggi et al., 2009). The seismic paths are as shown in the maps where the sources and receivers are denoted by red stars and blue triangles, respectively. The first rows of waveforms show the complete seismograms. The second rows of waveforms highlight the minor- and major-arc surface wave windows. Differences between oceanic and continental paths can be seen in minor arc windows especially for 40-110 s period range. It can also be seen that waveform fits are much better in the longer periods. (a) February 23, 2004 Samoa Islands ($M_w = 6.1$) earthquake recorded at 90° distance with mostly oceanic minor arc path. (b) August 12, 2009 Japan ($M_w = 6.6$) earthquake recorded at 81° distance with continental minor arc path. Major arc window couldn't be selected in 40-110 seconds period range due to larger amplitude difference.

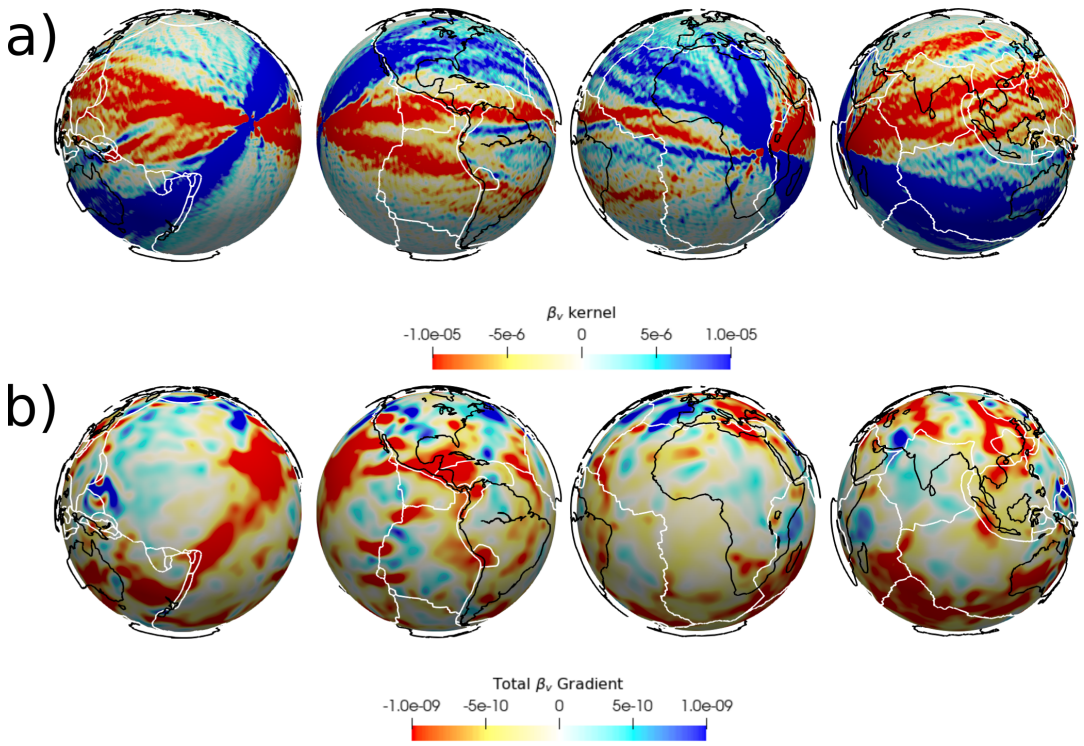

FIGURE 2.4: Event and total gradients. (a) An example vertically polarized shear wave (β_v) event kernel for 2005 Lake Tanganyika Earthquake (Mw: 6.8) at 200 km depth (b) An example vertically post-processed polarized shear wave (β_v) total gradient for the azimuthally anisotropic model GLAD-AZI M29 (see Chapter 4).

Conjugate Gradient (NLCG) or L-BFGS method (Fletcher and Reeves, 1964; Nocedal and Wright, 2006). The step length α may be determined using a line search. For 3D applications, a subset of the database might be used for this process to reduce the computational cost.

In this thesis, in 2D experiments, forward and adjoint simulations are done using the spectral-element based software package *SPECFEM2D* (Komatitsch and Vilotte, 1998). For synthetic examples, models are designed to perform in a specific frequency range, and target and initial model seismic data is obtained using forward simulations. Pre-processing steps are performed using *python* based tools like *obspy* and *pyadjoint* (Beyreuther et al., 2010). Since these steps can be performed in an embarrassingly parallel way, they are parallelized using Message Passing Interface (MPI). In order to reduce the number of files, which is advised for High Performance Computing (HPC) systems, Seismic Unix (SU) binary data format is used for the seismic data. Unformatted Fortran binary data format is used for the model and kernels for numerical simulations.

For 3D global simulations, I used Oak Ridge National Laboratory's IBM AC922 Summit system. In 3D experiments, since data are much more sophisticated and bigger, other formats that enable parallel input/output operation are used. HDF5 based Adaptaple Seismic Data Format (ASDF) is used for seismograms which holds seismic data, instrument and event information as a single container (Krischer et al., 2016). Models and kernels are stored in the Adaptable IO System (ADIOS) file format (Liu et al., 2014). For the numerical simulations GPU version of the *SPECFEM3D_GLOBE* is used which is faster than the CPU version because of the more parallelizable hardware infrastructure. Since I/O still takes considerable amount

of time during the simulations, there are also studies which explore the use of multi-tier storage systems and using faster Non-volatile memory as a means to speed up the I/O processes further [1].

[1] The content of this work was submitted: Mehta, K.; Wolf, M.; Podhorzki, N.; **Örsvuran, R.**; Logan, J.; Wan, L.; Choi, J. Y.; Huck, K.; Yakushin, I.; Munson, T.; Foster, I.; Klasky, S. Co-Design Evaluation of HPC I/O Middleware for Multi-Tier Storage Architectures. International Parallel & Distributed Processing Symposium

Chapter 3

Double-difference measurements in global full-waveform inversions[1][2]

Summary

We demonstrate double-difference (DD) tomography, a method involving differential measurements between stations, for 2-D and 3-D adjoint inversions based on realistic source–receiver distributions, with a focus on the global scale. We first present 2-D synthetic inversion results using cross-correlation traveltime and L2 waveform difference objective functions. Introducing a weighting term to DD objective functions based on the number of measurement pairs per station speeds up convergence and reduces bias in the final inverted model due to uneven data coverage. We next demonstrate frequency-dependent multitaper DD measurements in a 3-D experiment with real earthquake data by computing global-scale gradients. At the global scale, careful selection of station pairs is required for differential measurements in terms of geographical distance or geological context. In our technique, if no suitable pairs are found for a particular station, the DD measurement reduces to a classical misfit measurement. Furthermore, we compare 2-D and 3-D DD results with those from corresponding conventional misfits. By exploiting previously unused information in the recorded wavefield, DD tomography shows promise for balancing the gradient and speeding up convergence, especially around dense regional seismic networks.

3.1 Introduction

Dense seismic networks, such as USArray in North America or Hi-net in Japan, provide opportunities for detailed imaging of the lithosphere and upper mantle (e.g., Shapiro et al., 2005; Burdick et al., 2008; Zhao et al., 2012; Schaeffer and Lebedev, 2014; Lu et al., 2018). However, such networks also introduce significant data imbalance challenges to global-scale adjoint inversions. Besides long-term efforts to improve global data coverage through ocean bottom seismometers (e.g., Operto et al., 2006; Thomas and Laske, 2015), floating acoustic robots (e.g., Nolet et al., 2019) and

[1] The content of this chapter was published in: **Örsvuran, R.**, Bozdağ, E., Modrak, R., Lei, W., Ruan, Y., 2020. Double-difference measurements in full-waveform inversion, *Geophy. J. Int.*, 20(1), 661–680, https://doi.org/10.1093/gji/ggz444.

[2] Geographical weighting based on source-receiver distribution was published in an accompanying paper: Ruan, Y., Lei, W., Modrak, R., **Örsvuran, R.**, Bozdağ, E., Tromp, J., 2019. Balancing Unevenly Distributed Data in Seismic Tomography: A Global Adjoint Tomography Example *Geophy. J. Int.*, 219(2), 1225-1236, https://doi.org/10.1093/gji/ggz356.

underwater fiber-optic cables (e.g., Sladen et al., 2019), data imbalance challenges may be addressed through the measurement strategy and choice of misfit function.

The adjoint method (Tarantola, 1984; Tromp et al., 2005; Virieux and Operto, 2009) is an iterative full-waveform inversion technique that has become a routine tool in exploration and earthquake tomography. After successful applications to regional- to continental-scale studies (e.g., Fichtner et al., 2009; Tape et al., 2009; Zhu et al., 2012), adjoint tomography has also become feasible at the global scale as demonstrated by Bozdağ et al. (2016) and most recently by Lei et al. (2020).

The success of adjoint inversions is closely linked to the chosen misfit function (e.g., Modrak et al., 2016). It is an active research field in seismology to design suitable misfit functions for the problem at hand. L_2 waveform misfit is commonly used for exploration (e.g., Tarantola, 1984) and global tomographic studies (e.g., Nolet, 1987; Li and Romanowicz, 1996). Although amplitudes constrain the elastic structure more effectively, their complex sensitivity to multiple parameters makes their inversion more challenging. The nonlinear behavior of waveform misfit is a well-known issue in full-waveform inversions which may cause global convergence difficulties. Potential remedies that have been proposed include starting iterations from longer periods (e.g., Ekström et al., 1997; Zhu et al., 2015; Pageot et al., 2013) if there is enough low-frequency signal and linearizing measurements by carefully chosen misfit functions, such as based on a phase measurement or envelopes, during the early iterations (e.g., Yuan et al., 2015; Bozdağ et al., 2016). The cross-correlation traveltime misfit is widely used in seismic tomography (e.g., Luo and Schuster, 1991) which highlights the traveltimes of the highest-amplitude phase in a selected time window. It provides very robust measurements. However, one needs to isolate waveforms to maximize the information extracted from each time series, which is also the case for its counterpart the cross-correlation amplitude misfit (e.g. Ritsema et al., 2002). Multitaper misfits (e.g., Laske and Masters, 1996), which are frequency-dependent measurements, are preferred to better take advantage of dispersive waves in inversions (e.g., Tape et al., 2009). Similar to any other phase measurement, multitaper and classical cross-correlation travel time measurements require a certain degree of correlation between observed and synthetic data to avoid cycle-skipping problems. Fichtner et al. (2008) proposed using time-frequency analysis (Kristeková et al., 2006) in adjoint inversions to allow for measuring the phase and amplitude of a wavetrain without losing information from lower-amplitude phases. Similarly, Bozdağ et al. (2011) proposed instantaneous phase and amplitude (envelope) misfits which can also be applied to wavetrains in the time domain. Rickers et al. (2012) and Yuan et al. (2020) used variants of the instantaneous phase misfit which better deal with the cycle-skipping problem.

Classically, seismologists make measurements on observed and synthetic data at a given station independently of data from other stations. Double-Difference (DD) methods, in contrast, involve measurements on station or source pairs, as introduced by Waldhauser and Ellsworth (2000) to improve earthquake locations. This approach was then used in body-wave tomography by Zhang and Thurber (2003) and Monteiller et al. (2005). Choi and Alkhalifah (2011) and Zhang et al. (2016) introduced a similar concept based on relative measurements with respect to a reference trace rather than using pairs of receivers or sources while Hu and Menke (1992) used differential measurements to obtain the polarization direction of seismic waves. In the adjoint tomography context, Yuan et al. (2016) proposed a DD traveltime misfit involving differences between cross-correlated observed and synthetic data over groups of stations and showed that DD traveltime measurements speed up convergence in iterative inversions while reducing bias due to source uncertainties (i.e.,

origin time and source-time function).

In this study, our goal is to adapt DD measurements to realistic global-scale adjoint inversions. We take the work of Yuan et al. (2016) further and perform 2D and 3D tests with global data coverage and measurement windows based on real data. We revise the DD misfit of Yuan et al. (2016) by introducing a normalization term based on the number of receiver (or measurement) pairs per station. In tomographic studies, it is common to apply a weighting scheme based on the distribution of sources and receivers to balance data coverage (e.g., Li and Romanowicz, 1996; Bijwaard et al., 1998; Lebedev and Van Der Hilst, 2008; Visser et al., 2008; Schaeffer and Lebedev, 2013; Ruan et al., 2019). To develop insight on the performance of DD measurements, we compare DD results to classical results with and without the receiver weighting (e.g., Bozdağ et al., 2016; Lei et al., 2020).

The paper is organized as follows. In Section 2 we define the DD traveltime and waveform misfits and their associated adjoint sources. We then discuss strategies for pairing stations, or measurement windows, for DD measurements to reduce the computational cost and potential nonlinearity in measurements. We next introduce the pair-wise weighting into DD misfits based on the number of measurement pairs. For completeness, we briefly discuss the general weighting strategy, and the receiver and source weightings based on their geographical distribution to balance gradients and speed up the convergence given by Ruan et al. (2019). In Sections 3 & 4, we perform 2D and 3D experiments with a realistic data coverage, using minor- and major-arc surface waves only. We then compare the DD results to those from classical misfit functions with and without the receiver weighting. We further demonstrate combining the misfit functions of DD and classical measurements to maximize the information retrieved from waveforms. We finally discuss and summarize our findings in Sections 5 & 6.

3.2 Methodology

In this section, we first give brief background information on classical cross-correlation traveltime measurements and their DD variants described by Yuan et al. (2016), and define the DD version of the classical L_2-norm waveform misfit (Tarantola, 1984; Nolet, 1987) with its associated adjoint source. We then discuss how DD measurements should be used in practice with an emphasis on global data coverage. Finally, we introduce “pair-wise weighting" to DD misfits and compare them to classical ones with and without the receiver weighting based on the geographical distribution of stations presented in a counterpart paper by Ruan et al. (2019).

3.2.1 Double-difference traveltime misfit

In theory, DD measurements can be applied to any misfit function. In seismic tomography, a major advantage of applying DD to traveltime measurements is to mitigate source uncertainty (Yuan et al., 2016). In this study, another reason we focus on DD traveltime measurements is to make direct comparisons with the types of phase measurements used in global adjoint tomography models (e.g., Bozdağ et al., 2016; Lei et al., 2020).

The classical traveltime misfit function (CC) is defined as

$$\chi_{cc} = \frac{1}{2}\sum_i [\Delta T_i]^2 \,, \tag{3.1}$$

where ΔT_i is the difference between the traveltimes of observed (T_i^{obs}) and synthetic ($T_i^{syn}(\mathbf{m})$) waveforms recorded at a component of station i, i.e., $\Delta T_i = T_i^{obs} - T_i^{syn}(\mathbf{m})$. Model parameters are denoted by $\mathbf{m}$ and to avoid clutter, we drop the summation over sources. The associated adjoint source becomes (see Tromp et al. (2005) for complete derivation)

$$f^{\dagger}_{cc_i}(\mathbf{x}, t) = -\sum_i \Delta T_i \frac{1}{N_i} w_i(T-t)\delta_t s_i(\mathbf{x}, T-t, \mathbf{m})\delta(\mathbf{x}-\mathbf{x}_i)\,, \tag{3.2}$$

where δs_i denotes perturbations in displacement synthetics due to change in model parameters $\mathbf{m}$ within a selected time window w. The normalization term N_i is given as

$$N_i = \int_0^T w_i s_i(\mathbf{x}, T-t, \mathbf{m})\partial_t^2 s_i(\mathbf{x}, T-t, \mathbf{m}) dt\,. \tag{3.3}$$

Following Yuan et al. (2016) the double-difference cross-correlation traveltime misfit (CC^{DD}) is defined as the differential traveltimes of observed and synthetic waveforms paired over a group of receivers, such that

$$\chi_{cc}^{dd} = \frac{1}{2}\sum_i \sum_{j>i} [\Delta\Delta T_{ij}]^2\,, \tag{3.4}$$

where $\Delta\Delta T_{ij} = \Delta T_{ij}^{obs} - \Delta T_{ij}^{syn}$, and ΔT_{ij}^{obs} and ΔT_{ij}^{syn} are cross-correlations of observed and synthetic traveltimes, respectively, computed over a group of station pairs denoted by indices i and j. The adjoint source for the CC^{DD} misfit made at a set of paired stations becomes (Yuan et al., 2016)

$$f^{dd^{\dagger}}_{cc_i}(\mathbf{x}, t) = -\sum_{j>i} \Delta\Delta T_{ij} \frac{1}{N_{ij}} \delta_t s_j(\mathbf{x}, T-[t-\Delta T_{ij}^{syn}], \mathbf{m})\ \delta(\mathbf{x}-\mathbf{x}_i)\,, \tag{3.5}$$

$$f^{dd^{\dagger}}_{cc_j}(\mathbf{x}, t) = +\sum_{i<j} \Delta\Delta T_{ij} \frac{1}{N_{ij}} \delta_t s_i(\mathbf{x}, T-[t+\Delta T_{ij}^{syn}], \mathbf{m})\ \delta(\mathbf{x}-\mathbf{x}_j)\,, \tag{3.6}$$

where s_i and s_j are the synthetic displacements at station i and j, respectively, and N_{ij} is the normalization factor given as

$$N_{ij} = \int_0^T \delta_t^2 s_i(\mathbf{x}, t+\Delta t_{ij}^{syn}, \mathbf{m}) s_j(\mathbf{x}, t, \mathbf{m})\ \mathrm{d}t\,. \tag{3.7}$$

3.2.2 Double-difference waveform misfit

The classical L_2-norm waveform misfit (WF) (e.g., Tarantola, 1984) is defined as

$$\chi_{wf} = \frac{1}{2}\sum_i \int_0^T ||d_i(\mathbf{x}, t, \mathbf{m}) - s_i(\mathbf{x}, t)||^2 \mathrm{d}t\,, \tag{3.8}$$

where d_i and s_i are observed and synthetic displacement data, respectively, recorded at station i. Following Tromp et al. (2005) the corresponding adjoint source may be written as

$$f^{\dagger}_{wf_i}(\mathbf{x}, t) = -\sum_i \left[d_i(\mathbf{x}, T-t) - s_i(\mathbf{x}, T-t, \mathbf{m})\right] \delta(\mathbf{x}-\mathbf{x}_i)\,. \tag{3.9}$$

We define the double-difference waveform misfit ($\mathrm{WF^{DD}}$) as the difference between differential observed and synthetic displacements between the ith and the jth stations

$$\chi_{wf}^{dd} = \frac{1}{2}\sum_i\sum_{j>i}\int_0^T || \left[d_i(\mathbf{x},t) - d_j(\mathbf{x},t)\right] - \left[s_i(\mathbf{x},t,\mathbf{m}) - s_j(\mathbf{x},t,\mathbf{m})\right] ||^2 \mathrm{d}t\,. \tag{3.10}$$

The gradient then becomes

$$\delta\chi_{wf}^{dd} = -\sum_i\sum_{j>i}\int_0^T \{\left[d_i(\mathbf{x},t) - d_j(\mathbf{x},t)\right] - \left[s_i(\mathbf{x},t,\mathbf{m}) - s_j(\mathbf{x},t,\mathbf{m})\right]\}[\delta s_i(\mathbf{x},t,\mathbf{m}) - \delta s_j(\mathbf{x},t,\mathbf{m})]\mathrm{d}t\,, \tag{3.11}$$

where δs_i and δs_j are the perturbations of synthetic displacements at the ith and the jth stations, respectively. We can rewrite the gradient by dividing it into two parts,

$$\delta\chi_{wf}^{dd} = -\sum_i\sum_{j>i}\int_0^T \{\left[d_i(\mathbf{x},t) - d_j(\mathbf{x},t)\right] - \left[s_i(\mathbf{x},t,\mathbf{m}) - s_j(\mathbf{x},t,\mathbf{m})\right]\}\delta s_i(\mathbf{x},t,\mathbf{m})\mathrm{d}t \tag{3.12}$$

$$+\sum_i\sum_{i<j}\int_0^T \{\left[d_i(\mathbf{x},t) - d_j(\mathbf{x},t)\right] - \left[s_i(\mathbf{x},t,\mathbf{m}) - s_j(\mathbf{x},t,\mathbf{m})\right]\}\delta s_j(\mathbf{x},t,\mathbf{m})\mathrm{d}t\,. \tag{3.13}$$

Following Tromp et al. (2005) and Yuan et al. (2016), we may formulate the $\mathrm{WF^{DD}}$ adjoint source in two parts at the location of receivers i and j,

$$f_i^\dagger(\mathbf{x},t) = -\sum_{j>i}\int_0^T \{\left[d_i(\mathbf{x},T-t) - d_j(\mathbf{x},T-t)\right] - \left[s_i(\mathbf{x},T-t,\mathbf{m}) - s_j(\mathbf{x},T-t,\mathbf{m})\right\}]\,\delta(\mathbf{x}-\mathbf{x}_i) \tag{3.14}$$

$$f_j^\dagger(\mathbf{x},t) = +\sum_{i<j}\int_0^T \{\left[d_i(\mathbf{x},T-t) - d_j(\mathbf{x},T-t)\right] - \left[s_i(\mathbf{x},T-t,\mathbf{m}) - s_j(\mathbf{x},T-t,\mathbf{m})\right]\}\delta(\mathbf{x}-\mathbf{x}_j)\,. \tag{3.15}$$

3.2.3 Pairing measurements

It is well known that the seismic inversion is prone to cycle skipping and problems with local minima. To mitigate these problems, it is essential to have good starting models (Mulder and Plessix, 2008; Brossier et al., 2009; Prieux et al., 2013c, e.g.,) and choose appropriate misfit functions (e.g., Gauthier et al., 1986; Luo and Schuster, 1991; Bozdağ et al., 2011; Modrak and Tromp, 2016) or start with longer-period data and gradually involve shorter periods as the inversion proceeds (e.g., Sirgue and Pratt, 2004; Bozdağ et al., 2016). Although in one sense it is desirable to have as many station pairs as possible, in DD measurements one also has to make sure that waveforms from paired stations are reasonably well correlated to minimize potential nonlinearities. We suggest the following three selection criteria to pair stations based on practical experience.

1. **Measurement window:** In theory, it is possible to pair any signal in DD measurements. However, to make robust, well-behaved measurements, it is sensible to consider similar seismic phases or time windows. For instance, minor-arc surface waves should not be directly compared with major-arc surface waves; both should be considered separately within their own categories.

2. **Distance:** Typically, it is useful to consider stations only within a certain distance from one another. Following Yuan et al. (2016), a selection radius may be determined based on the scale length of heterogeneities and the width of the first Fresnel zone, which depends on the wavelength and path length. If the size of the problem is small enough (e.g., regional or reservoir scale), this criterion may be skipped and poorly-behaved pairs may directly be eliminated based on the waveform similarity criterion in the next item. However, the search radius may also affect the convergence rate (see Appendix). On the other hand in large-size problems, such as the global-scale tomography, it is not practical to correlate data from all stations distributed worldwide with this criterion, the search area can be narrowed down to check waveform similarities.

3. **Waveform similarity:** It is desirable to have well-correlated waveforms to mitigate nonlinearity in full-waveform inversions. As determined by cross-correlation of paired waveforms, waveform similarity can be used to exclude poorly-correlated station pairs.

In this study, we demonstrate DD measurements for global adjoint inversions based on two sets of data: minor- and major-arc surface waves. For measurements, minor- and major-arc surface waves are paired in their own categories, i.e., no pairing between minor- and major-arc waves. In Appendix A, we show how the selection criteria to pair stations may affect the performance of the inversion based on 2D checkerboard examples.

3.2.4 Balancing measurements

In earthquake seismology, data is typically recorded on three components. To maximize the information extracted from data and to reduce the nonlinearity of adjoint inversions, it is common to make measurements on pre-selected time windows on multiple period bands (e.g., Zhu et al., 2015; Bozdağ et al., 2016). Measurements made on every component (typically vertical, radial and transverse) for every period band form a measurement category. Different measurement categories may then be balanced based on their number of measurements with a weighting term introduced into the misfit function to give similar importance to all wave types in the inversion (e.g., Bozdağ et al., 2016). At the global scale, the cluster of sources and stations may further be weighted to reduce the bias in tomographic images due to uneven data coverage or speed up convergence of iterative inversions (e.g., Li and Romanowicz, 1996; Lebedev and Van Der Hilst, 2008; Clouzet et al., 2018). To balance different measurement categories by letting every wave type (i.e., body and surface waves) have similar contributions in inversions, a general weighting may be defined (Ruan et al., 2019; Lei et al., 2020) as

$$\Phi = \sum_{s}^{S} \omega_s \sum_{c}^{C} \omega_c \sum_{r}^{R_{sc}} \omega_{scr} \sum_{w}^{N_{scr}} \omega_{scrw} \chi_{scrw} \,, \tag{3.16}$$

where S, C, R, and N denote the number of sources, the number of categories (period bands and components), the number of receivers and the number of measurement windows, respectively. ω_s, ω_c, ω_{scr} and ω_{scrw} are the source, category, receiver and measurement weights, respectively. Let the data misfit from source s, measurement category c, receiver r, and measurement window w be

$$\chi_{srcw} = \left(\frac{\Delta d_{scrw}}{\sigma_{scrw}} \right)^2 , \tag{3.17}$$

where Δd_{scrw} is a measurement with an associated standard deviation σ_{scrw}. When the model fits the data to within one standard deviation, we expect that

$$\chi_{scrw} \sim 1 . \tag{3.18}$$

Pair-wise weighting for double-difference measurements

In DD measurements, each station, or measurement window, may have different number of measurement pairs. As a result of the summation over paired receivers in the adjoint source, the paths towards the cluster of receivers become more pronounced in the gradient.
We introduce a pair-wise weighting term to the DD misfit function that weights every measurement by the number of its pairs including itself such that when a station does not have any other pairs, the measurement reduces to classical misfit. The pair-wise weighting is applied to each measurement window.
Let us first suppose the measurement weight ω_{scrw} and the receiver weight ω_{scr} are equal to one. Then for a given source s and the category c, the total weight becomes,

$$\sum_{r}^{R_{sc}} \omega_{scr} \sum_{w}^{N_{scr}} \omega_{scrw} = N_{sc} . \tag{3.19}$$

The pair-wise weighting should satisfy this equation. The weight of non-paired stations - where classical misfits are used - is set to one and the paired measurements are down-weighted by their number of pairs, such that

$$\omega^{dd}_{srcw} = \frac{1}{p_{scrw}} , \tag{3.20}$$

where ω^{dd}_{scrw} is the weight of the measurement w of source s, category c and receiver r, and p_{scrw} is the measurement's number of pairs. To satisfy eq. (3.19), we need to define the normalization parameter α

$$\alpha \sum_{r}^{R_{sc}} \omega_{scr} \sum_{i}^{N_{scr}} \omega^{dd}_{scrw} = N_{sc} , \tag{3.21}$$

where α becomes

$$\alpha = \frac{N_{sc}}{\sum_{r}^{R_{sc}} \omega_{scr} \sum_{i}^{N_{scr}} \omega^{dd}_{scrw}} . \tag{3.22}$$

In Fig. 3.1, we show the effect of pair-wise weighting in a 2D experiment with a small and a bigger array of receivers. The event kernels are computed for CC^{DD} with and without the pair-wise weighting and compared to those from the classical cross-correlation traveltime misfit. The DD misfit gives higher resolution underneath receiver arrays compared to the CC misfit, consistent with the results of Yuan

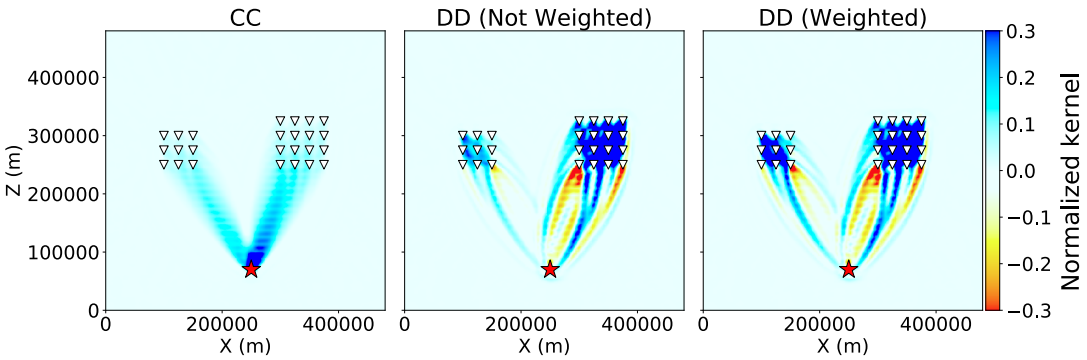

FIGURE 3.1: Sensitivity kernels computed for cross-correlation traveltime misfit (CC) (*left*) and double-difference cross-correlation traveltime CC^{DD} misfit without (*middle*) and with (*right*) the pair-wise weighting. Stations are paired within their own arrays for DD measurements.

et al. (2016). The pair-wise weighting, on the other hand, reduces the dominance of the bigger array and highlights the structure underneath both station clusters.

Geographical weighting

We compare DD misfits with the pair-wise weighting to classical misfits with a "geographical weighting" proposed by Ruan et al. (2019) and used by Lei et al. (2020) in a recent global adjoint tomography study. The geographical weighting can be applied to weight both the source and receiver distribution and is defined by

$$\omega_i^{-1} = \sum_{j=1}^{N} \exp\left[-\left(\frac{\Delta_{ij}}{\Delta_0}\right)^2\right], \tag{3.23}$$

where Δ_{ij} is the distance between stations i and j and Δ_0 is the reference distance. We use 21 earthquakes distributed worldwide in our 2D inversion tests and 3D gradient simulations avoiding the cluster of sources. Therefore, the source weights are set to one in all our experiments balancing measurements with the receiver weighting only.

Combining Classical and Double-Difference Measurements

In a real experiment, there may be non-paired stations such as oceanic island stations in global inversions. DD misfits with the pair-wise weighting may deal with such a scenario by applying classical misfits to non-paired stations and setting their weights to one. On the other hand, DD measurements may also be combined with classical measurements applied to all stations. For instance, two different misfits may be applied to the same set of measurement windows as multiple misfits may complement each other to extract more information from data. Alternatively, one may prefer to use classical misfits for body-wave phases combined with DD measurements of surface waves.

We formulate the combination of paired and non-paired (single) measurements using the general weighting (eq. 3.16). We may also introduce an additional weighting to balance these two types of measurements if needed. We define the total weight of measurements from single stations for source s and category c as

$$\omega_{total}^{single} = \sum_{r}^{R_{sc}} \omega_{scr} \sum_{w}^{N_{scr}} \omega_{scrw} \,, \tag{3.24}$$

where ω_{scrw} is the weight of classical measurements which is set to one due to the pair-wise weighting. For DD measurements the total weight becomes

$$\omega_{total}^{dd} = \sum_{r}^{R_{sc}} \omega_{scr} \sum_{w}^{N_{scr}^{dd}} \omega_{scrw}^{dd} \,, \tag{3.25}$$

where ω_{scrw}^{dd} is the pair-wise weighting. Again, we need to satisfy eq. (3.19) and the total weight must be equal to N_{sc} such that

$$\alpha[\omega^{single}\omega_{total}^{single} + \omega^{dd}\omega_{total}^{dd}] = N_{sc} \,, \tag{3.26}$$

where ω^{single} and ω^{dd} are the relative weights for classical and DD measurements, respectively. If these weights are the same, both measurement types will have equal contribution to the total misfit. To satisfy the equation, we need to define α as

$$\alpha = \frac{N_{sc}}{\omega^{single}\omega_{total}^{single} + \omega^{dd}\omega_{total}^{dd}} \,. \tag{3.27}$$

Hereafter we will call this type of combination of classical and DD misfits for cross-correlation traveltime and waveform measurements $\mathrm{CC}^{\mathrm{DD}} + \mathrm{CC}^{\mathrm{single}}$ and $\mathrm{WF}^{\mathrm{DD}} + \mathrm{WF}^{\mathrm{single}}$, respectively.
When classical misfits are applied to all stations and combined with DD misfits, DD measurements may be introduced as a new category, i.e., every station has a classical measurement and DD measurements are made on paired stations only. Thus for each period band and component two categories are defined, one for the conventional misfit and one for the DD misfit where each category may have different number of measurements. From here on, we call this type of combination $\mathrm{CC}^{\mathrm{DD}} + \mathrm{CC}^{\mathrm{all}}$ and $\mathrm{WF}^{\mathrm{DD}} + \mathrm{WF}^{\mathrm{all}}$ for traveltime and waveform measurements, respectively.

Relation between pair-wise weighting and geographical receiver weighting

If we do not take the waveform similarity into account and assume that each station has only one measurement, the receiver and the pair-wise weighting behave similarly. We illustrate such a case in Fig 3.2 how the receiver and pair-wise weightings vary globally for a reference or pairing distance, respectively, of 10 and 30 degrees. Both weighting strategies show similar results where the receiver weighting behave like the smoother version of the pair-wise weighting.
Fig. 3.3 shows how the receiver and the pair-wise weightings change as a function of the reference or pairing distance, respectively, depending on the location of a chosen reference station on the globe. When the reference station is in a seismic network and well-paired with neighboring stations, as can be the case for USArray stations, both weightings give similar results (Fig. 3.3, top plots). The similarity decreases for fewer pairs of stations. For a reference station in Asia, although both weighting schemes have similar trends, the pair-wise weighting shows some steps as the number of pairs does not gradually increase by the pairing distance (Fig. 3.3, middle plots). When the reference station is located on an oceanic island, both weightings become one as there is no other station nearby unless a significantly high reference distance is set to search for neighboring stations (Fig. 3.3, bottom plots).

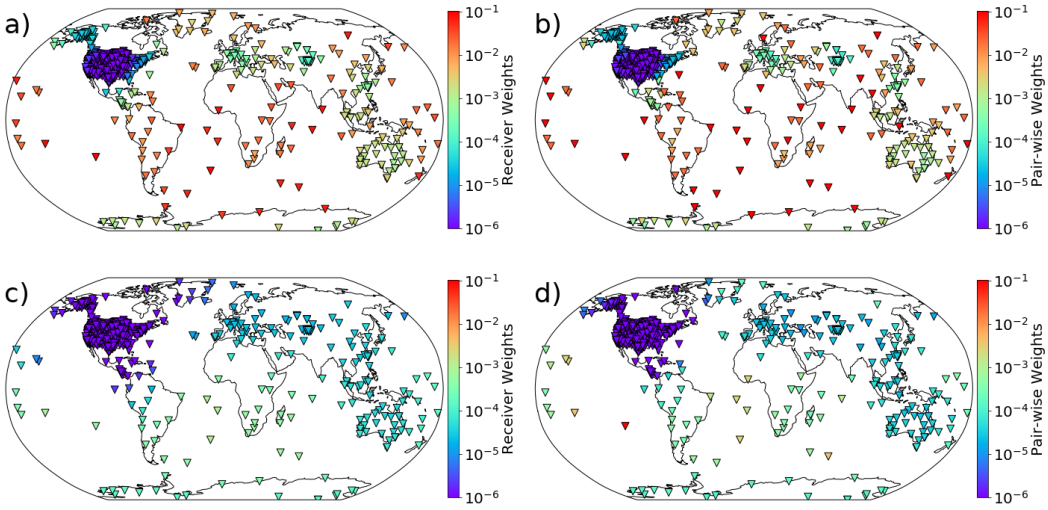

FIGURE 3.2: Receiver and pair-wise weights as a function of reference and pairing distances, respectively. a) Receiver weighting for reference distance of 10 degrees. b) Pair-wise weighting for pairing distance of 10 degrees. c) Receiver weighting for reference distance of 30 degrees. d) Pair-wise weighting for pairing distance of 30 degrees.

In theory, the receiver weighting can be used for DD measurements instead of or with the pair-wise weighting. However, the pair-wise weighting is a natural term introduced into the DD misfit function and applied to each measurement window rather than each station. Therefore it is more suitable to balance DD measurements with the pair-wise weighting, especially in case of having multiple measurement windows per station.

3.3 2D Experiments

In this section, we perform 2D experiments to test DD measurements with a realistic global data coverage. We first invert for a global phase-speed model (Trampert and Woodhouse, 1995). Then we use Picasso's Guernica painting as a target model which is independent of the global data coverage and has more smaller-scale heterogeneities. In addition to DD traveltime measurements we also use the classical waveform misfit and its DD version for the Guernica example. To gain a better understanding of DD measurements with the newly introduced pair-wise weighting scheme, we compare the results to those from classical misfits with and without the geographical receiver weighting. Furthermore, we test combining classical and DD measurements to maximize the information we can extract from seismograms. The `SPECFEM2D` package (Komatitsch and Vilotte, 1998) is used for wave simulations in SH mode such that membrane waves propagate in x-z plane while the particle motion is in y direction. Periodic boundary conditions are used in order to mimic global ray paths in the 2D plane.

3.3.1 CASE I: Inverting for a global phase-speed model

In this example, our target model in inversions is the same phase-speed map used in the experiments of Yuan et al. (2016). It is the 40 s Rayleigh wave phase-speed map of Trampert and Woodhouse (1995) adapted to 2D numerical simulations and used to generate synthetic seismograms to serve as *observed* data in our tests. The

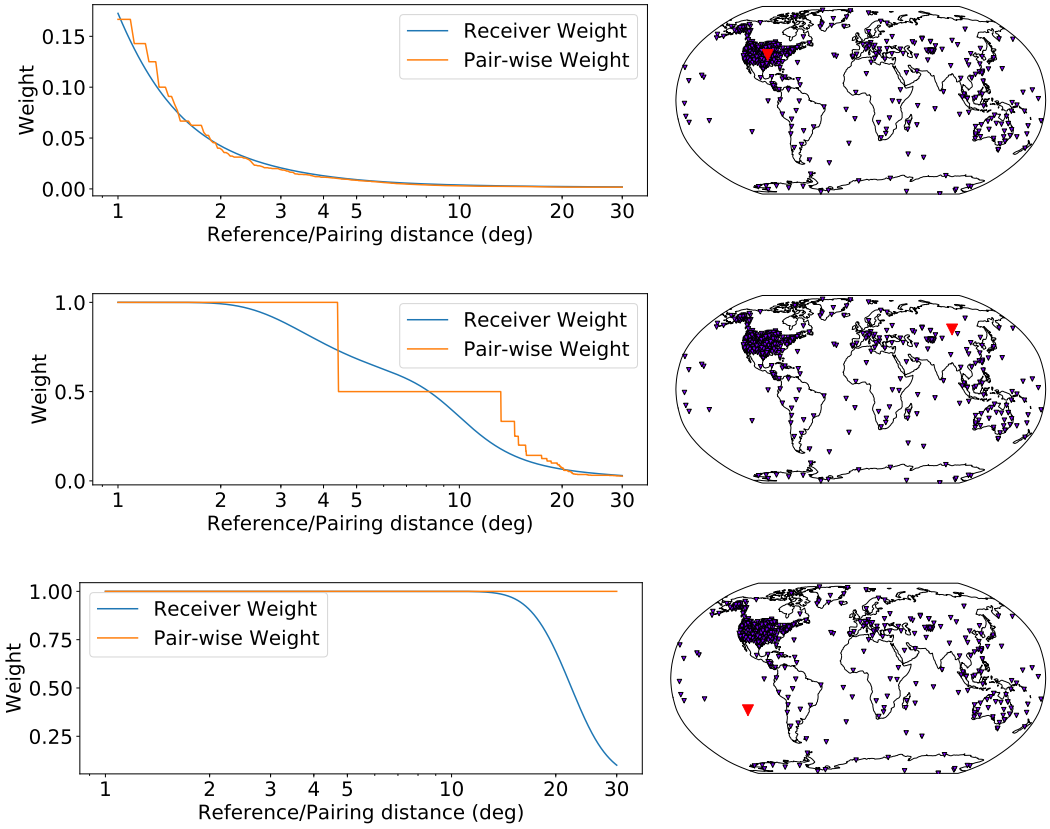

FIGURE 3.3: Receiver and pair-wise weights for three stations (red triangles) located in a dense seismic network in North America (*top*), in a moderately sampled region in Talaya, Russia (*middle*), and on the Easter Island in Pacific (*bottom*). Agreement between two weighting schemes depends on the number of stations in the vicinity of the reference station.

iterations start from a homogeneous model whose numerical seismograms are used to mimic *synthetic* data. We selected 21 global CMT earthquakes and their associated stations used in Bozdağ et al. (2016). We use the paths only where surface waves would be selected in a real 3D global-scale inversion. To this end, we use the same time windows selected by *pyflex* (the python version of the `FLEXWIN` algorithm by Maggi et al. (2009)) for the GLAD-M15 model (Bozdağ et al., 2016). DD pairs are selected based on the criteria described in Section 3.2.3. The measurement windows are paired if the stations are within a 5-degree radius area and there is 90% similarity between the waveforms. The selection parameters are chosen conservatively to mimic a potential real 3D scenario.
As the number of paired stations is limited, we combine DD measurements made on paired stations with classical traveltime measurements made on non-paired stations with the appropriate weighting schemes. Five misfit functions used in our tests with different weighting schemes are as given below:

1. $\mathrm{CC}^{\mathrm{all}}$ is the classical cross-correlation traveltime misfit applied to all stations without the receiver weighting.

2. $\mathrm{CC}^{\mathrm{all}}_{\mathrm{rec}}$ is the classical cross-correlation traveltime misfit applied to all stations with the receiver weighting (rec) introduced by Ruan et al. (2019).

3. $\mathrm{CC}^{\mathrm{DD}} + \mathrm{CC}^{\mathrm{single}}$ is the double-difference cross-correlation traveltime misfit on paired stations with the classical cross-correlation traveltime misfit without the receiver weighting on remaining non-paired (single) stations.

4. $\mathrm{CC}^{\mathrm{DD}}$+$\mathrm{CC}^{\mathrm{all}}$ is the double-difference cross-correlation traveltime misfit on paired stations combined with the classical traveltime misfit on all stations without the receiver weighting.

5. $\mathrm{CC}^{\mathrm{DD}}$+$\mathrm{CC}^{\mathrm{all}}_{\mathrm{rec}}$ is the double-difference cross-correlation traveltime misfit on paired stations combined with the classical traveltime misfit with the receiver weighting on all stations.

Note that hereafter in all 2D and 3D experiments $\mathrm{CC}^{\mathrm{DD}}$ measurements are always made on paired stations with the pair-wise weighting. Fig. 3.4 shows the final models obtained after 20 iterations using an L-BFGS optimization technique (e.g., Nocedal, 1980; Liu and Nocedal, 1989). The model misfit reduction is computed using the root-mean-square metric between the target and the inverted models for all measurement types (Fig. 3.5). Balancing the receiver distribution speeds up the convergence for the $\mathrm{CC}^{\mathrm{all}}_{\mathrm{rec}}$ measurements which is consistent with the results of Ruan et al. (2019). $\mathrm{CC}^{\mathrm{DD}} + \mathrm{CC}^{\mathrm{all}}$ has better results than $\mathrm{CC}^{\mathrm{all}}$ but cannot improve the performance as good as other misfits since the CC measurements without the receiver weighting dominate the gradient. The primary goal of DD measurements is to extract more information underneath densely covered regions which is clearly observed, for instance, in North America due to nicely paired USArray stations. However, it is promising to observe that the pair-wise weighting of DD measurements, which weights every measurement window by its number of pairs, naturally balances the data coverage by reducing the artifacts towards the cluster of stations. As a result, the convergence rate becomes comparable to that of $\mathrm{CC}^{\mathrm{all}}_{\mathrm{rec}}$ as seen by the misfit reduction of $\mathrm{CC}^{\mathrm{DD}} + \mathrm{CC}^{\mathrm{single}}$. All measurements with the receiver weighting (for CC) and/or pair-wise weighting (for DD) give similar misfit reductions after 20 iterations.

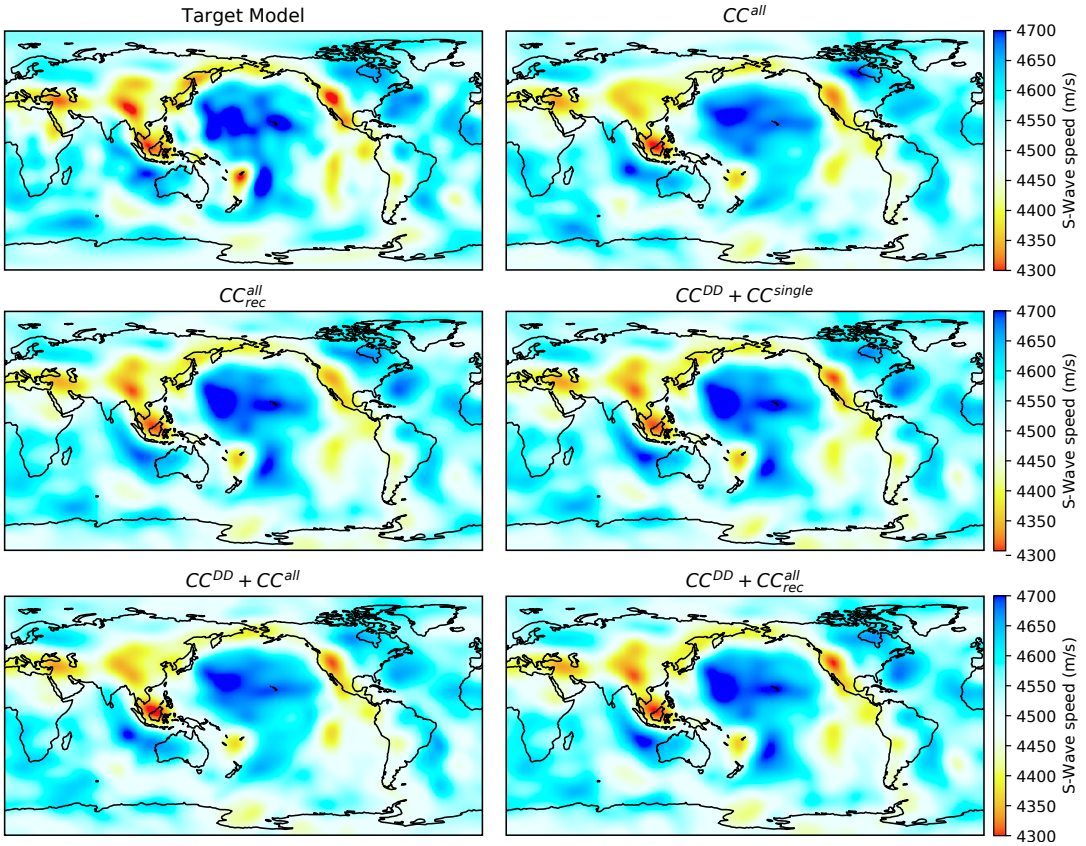

FIGURE 3.4: Inversion results after 20 iterations for cross-correlation traveltimes (CC^{all}), cross-correlation traveltimes with the receiver weighting (CC^{all}_{rec}), double-difference cross-correlation traveltimes with classical measurements on non-paired stations ($CC^{DD} + CC^{single}$), double-difference cross-correlation traveltimes with classical measurements on all stations without ($CC^{DD} + CC^{all}$) and with ($CC^{DD} + CC^{all}_{rec}$) the receiver weighting. DD measurements are always made on paired stations with the pair-wise weighting. The starting model is homogeneous with the average wavespeed of the target model. The target model is re-constructed from the 40 s Rayleigh phase-speed global model by Trampert and Woodhouse (1995)). The distribution of 21 earthquakes (red stars) and the stations (triangles) are as shown on the *top-left* plot.

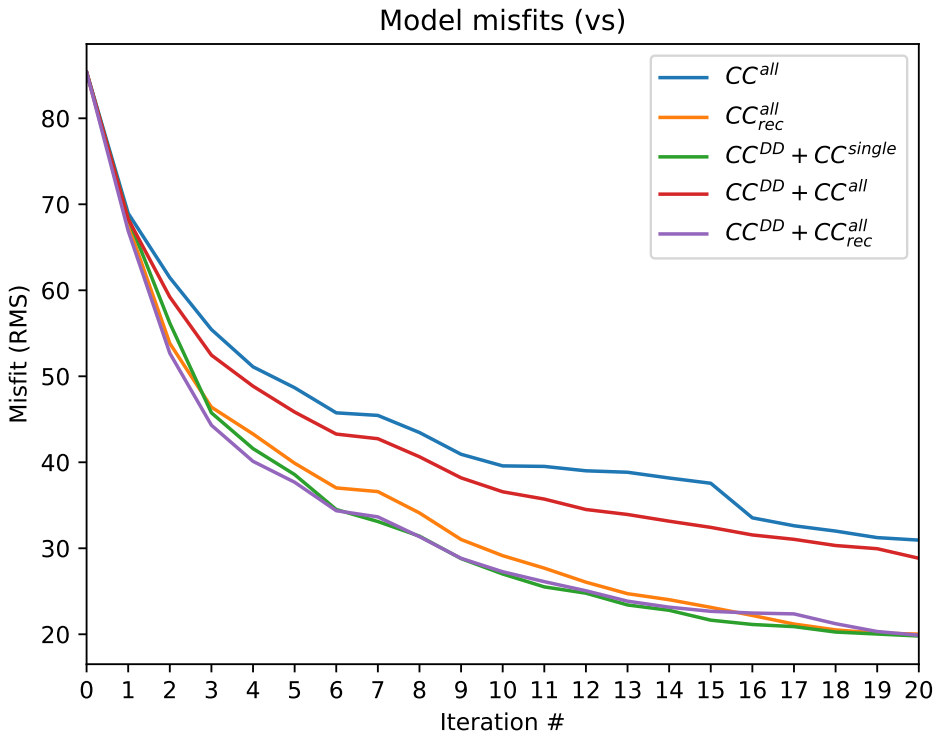

FIGURE 3.5: RMS model misfits for 2D inversions after 20 iterations presented in Fig. 3.4.

FIGURE 3.6: Guernica by Pablo Picasso (© 2019 Estate of Pablo Picasso / Artists Rights Society (ARS), New York).

3.3.2 CASE II: Guernica

We extend our tests to invert for Picasso's Guernica painting (Figure 3.6) which is independent of the global source-station distribution and has more smaller-scale heterogeneities. Fig. 3.7 shows the power spectral densities of the global phase-speed map used in the previous section and the Guernica model to compare their scale-length of heterogeneities. Welch's method (Welch, 1967) is used to compute the power spectral densities. Spatial frequency contents, which are inversely proportional to the scale-lengths of heterogeneities, show the richer spectrum of Guernica.

The same global source-receiver distribution used as in Section 3.3.1 and the initial model is homogeneous with the average wavespeed of the target model. We perform 15 iterations with traveltime measurements where the smoothing is decreased gradually every five iterations from a half-width of 5 degrees to 3 degrees to 1 degree.

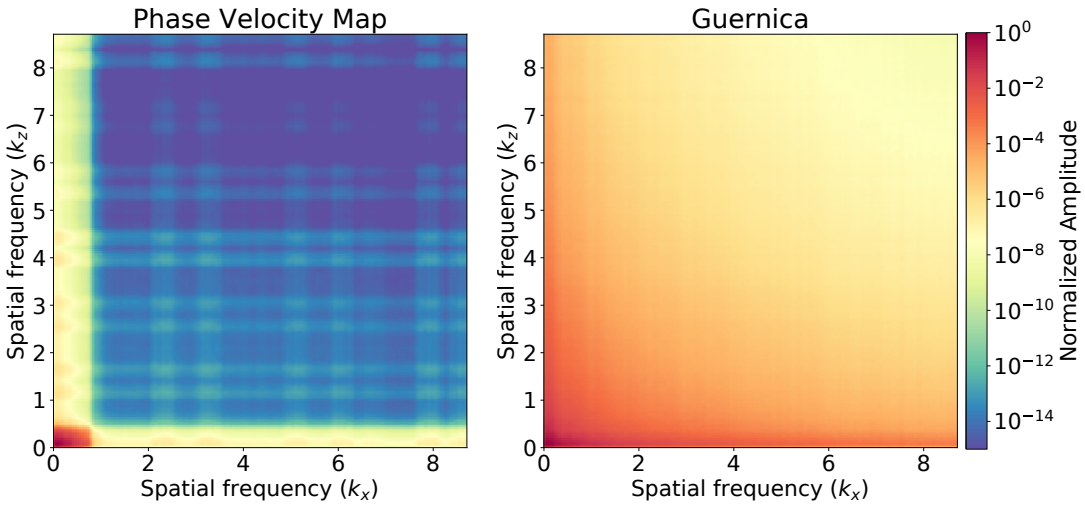

FIGURE 3.7: Power spectral density plots of the 40 s Rayleigh phase-speed map and the Guernica model used in 2D experiments. Axes of the plots denote spatial frequencies in x and y directions where higher values correspond to smaller-scale heterogeneities.

Fig. 3.8 shows the results for five misfit functions described in Section 3.3.1. The receiver weighting is again essential for speeding up the convergence. CC^{DD} measurements are showing significant improvement visually on densely covered regions, such as North America, compared to CC measurements (Fig. 3.9). CC^{DD} combined with a receiver-weighted CC misfit is promising to speed up the convergence and extract more information from data, specifically underneath densely covered areas. To further understand what each measurement exactly does we look at the regional misfit variations. Fig. 3.10 shows the regional logarithmic model misfits of four measurements computed with respect to those from CC^{all} (i.e., $\ln(\text{Model}/\text{Model}_{CC^{all}})$) at every 10×10-degree cell. CC^{all}_{rec} has worse model misfits in North America than those of any measurement with CC^{DD} because the receiver weighting down-weights the cluster of stations to balance the total gradient. Indeed, CC^{DD} measurements may significantly speed up the convergence underneath dense seismic stations as can be seen in regional experiments (see Appendix) which may not be fully observed in total model misfits in global inversions due to the limited station clusters worldwide.
After 15 iterations with traveltime-based misfits, we switch to L_2 waveform measurements using the result of $CC^{DD} + CC^{all}_{rec}$ as the initial model (Fig. 3.11). Including amplitude information in inversions significantly increases the resolution as expected. Overall, we observe similar results to those from traveltime measurements. WF^{all}_{rec} gives slightly better results than the others (Fig. 3.12). Fig. 3.13 shows that the measurements with WF^{DD} again better improve the regions with dense coverage.

3.4 3D global-scale experiments

3D wave propagation in laterally heterogeneous models can be more complex than what we have presented in 2D so far. Because of the computational cost, it is not feasible to perform 3D adjoint inversions for every misfit used in 2D experiments. However, 3D gradients may give an idea of how each measurement would behave in a real 3D set up. We specifically aim to observe if DD measurements combined with classical misfits are robust enough to be used in a real global adjoint inversion context. To this end, we compute sample 3D gradients with the same source-receiver

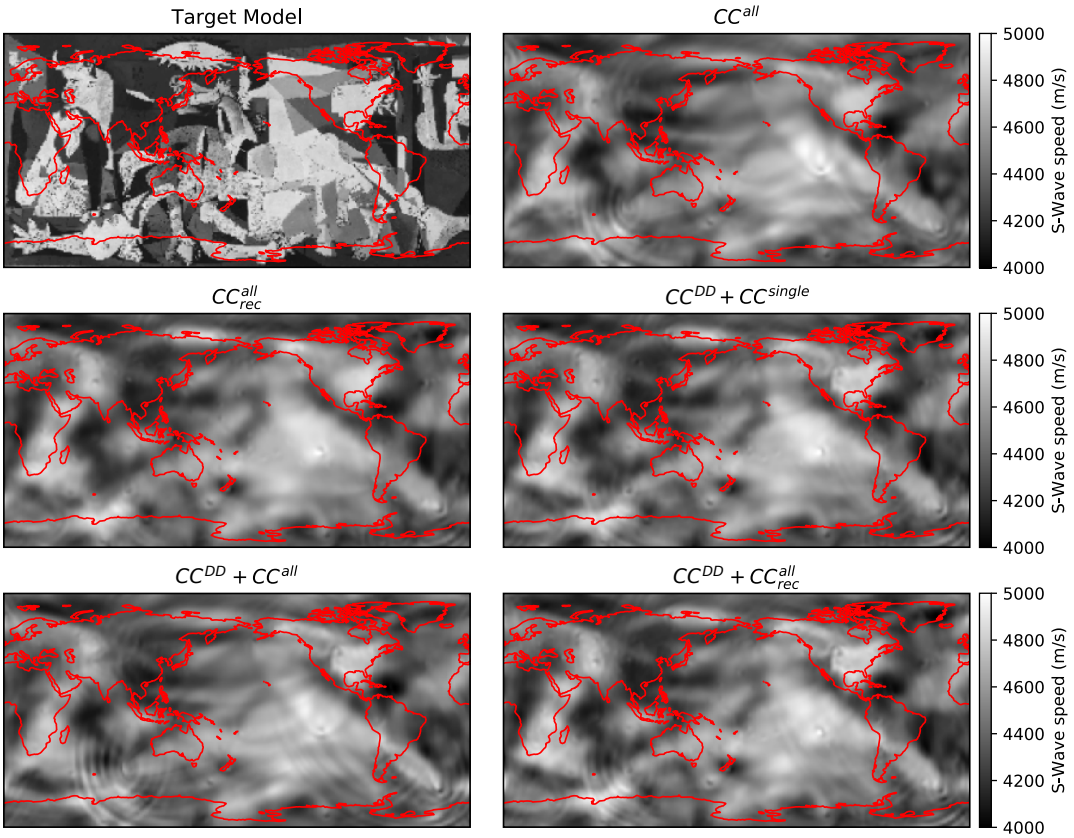

FIGURE 3.8: Same as Fig. 3.4 but for the Guernica model.

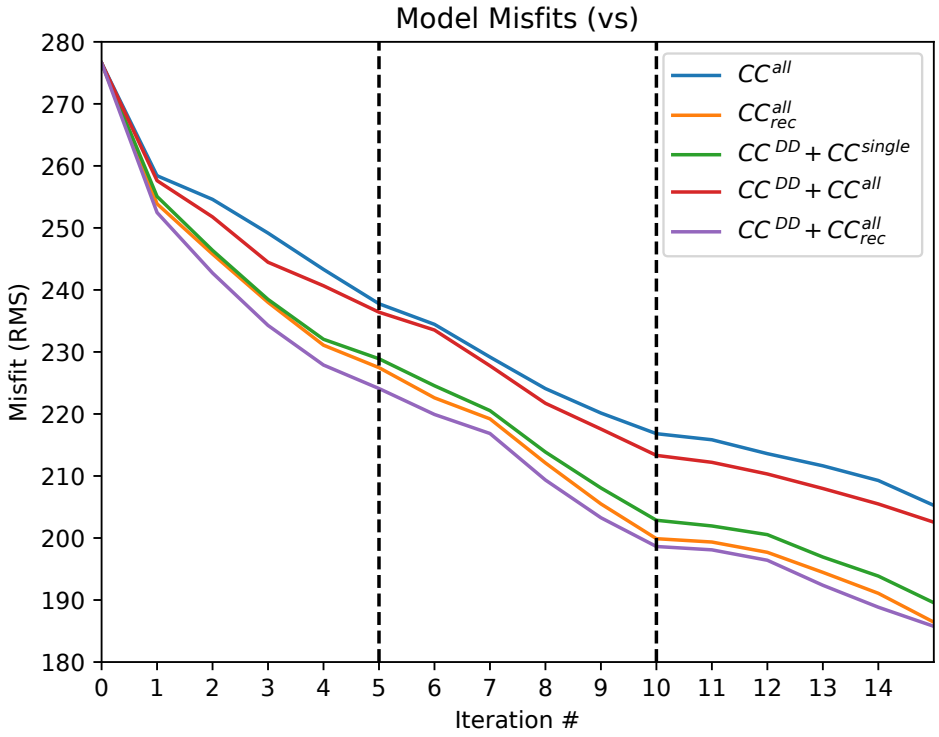

FIGURE 3.9: Same as Fig. 3.5 but for the Guernica model.

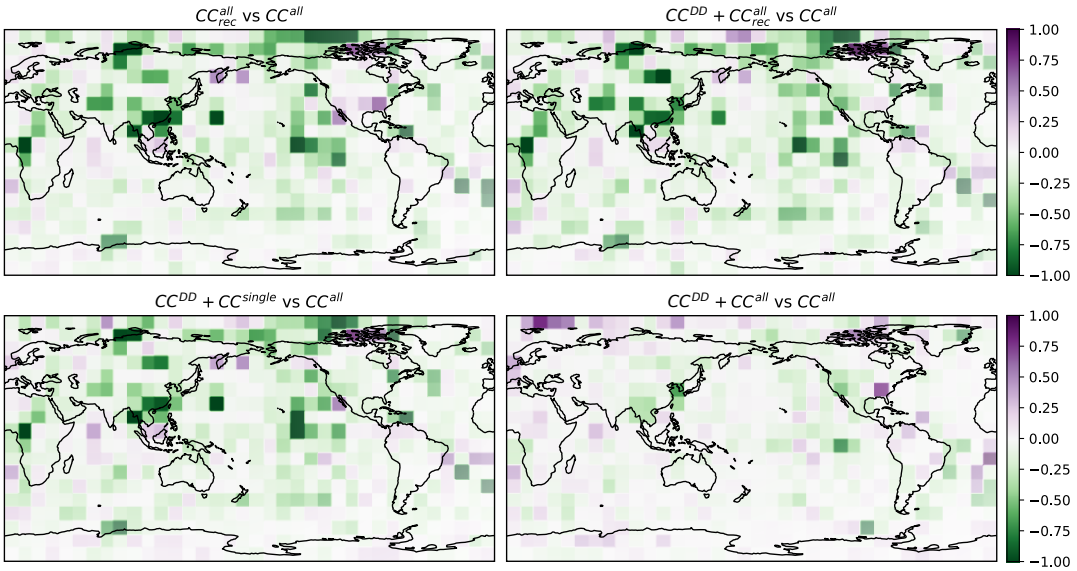

FIGURE 3.10: Logarithmic model misfits for four misfit functions computed with respect to CC^{all} (i.e., $\ln(\text{Model}/\text{Model}_{CC^{all}})$) showing the regional variations. The model misfits are computed at every 10×10-degree cell. Green and purple colors denote better (lower) and worse (higher) misfits, respectively, than those from CC^{all} measurements.

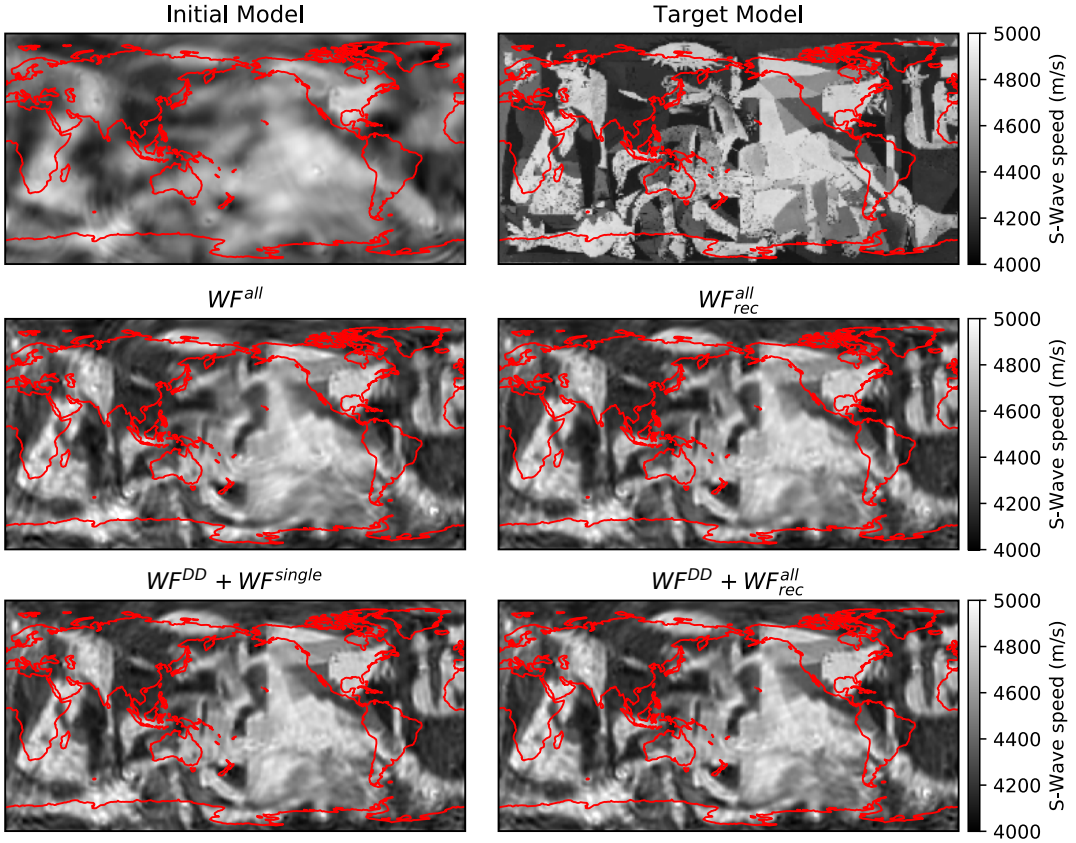

FIGURE 3.11: Same as Fig. 3.4 but for the Guernica model based on classical and double-difference waveform measurements. Initial model is the final model presented in Fig. 3.8 obtained from $CC^{DD} + CC^{all}_{rec}$ measurements.

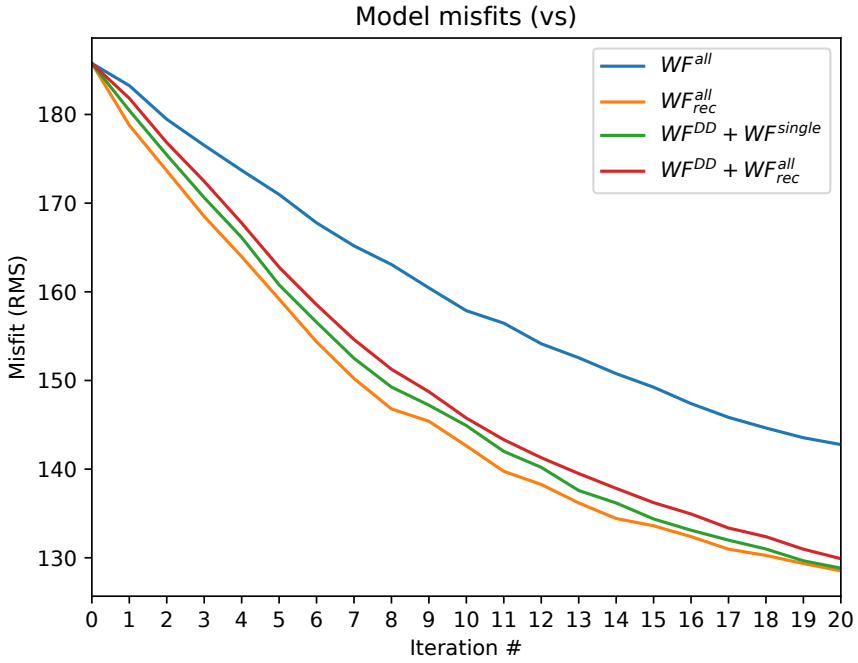

FIGURE 3.12: Same as Fig. 3.5 but for the Guernica model based on classical and double-difference waveform measurements.

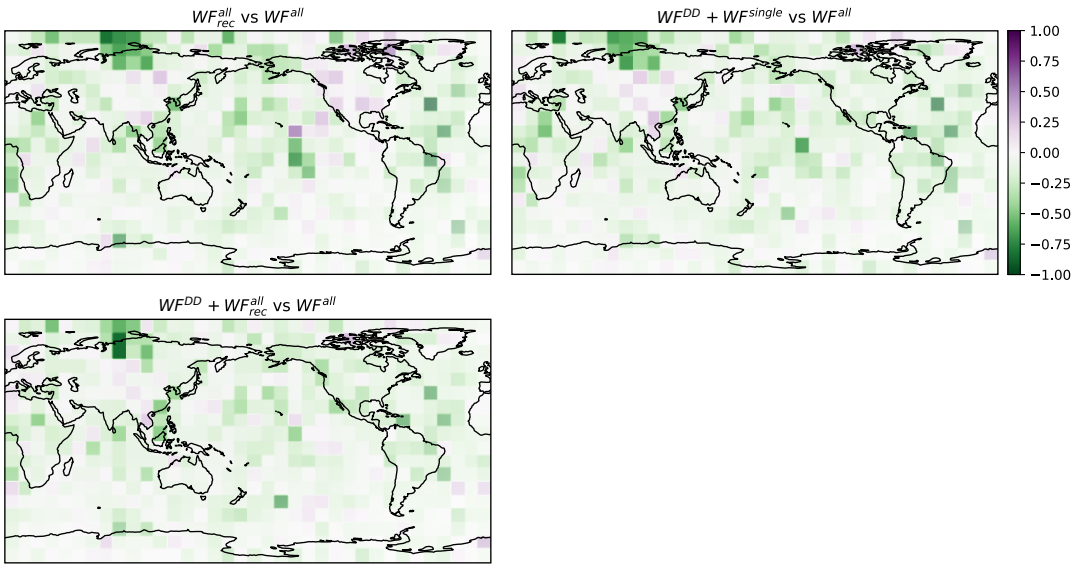

FIGURE 3.13: Same as Fig. 3.10 but for waveform measurements without $WF^{DD} + WF^{all}$.

TABLE 3.1: Number of pairs for different regions and components. R, T and Z denote radial, transverse and vertical components, respectively.

	Asia	Europe	North America	South America	Oceania
R	107	201	24714	42	42
T	460	301	50871	531	63
Z	2683	747	384677	3284	190

distribution. We measure multitaper traveltime misfits for minor- and major-arc Rayleigh and Love waves with the same misfit configurations for traveltime measurements presented in Section 3.1. GLAD-M15 (Bozdağ et al., 2016) is used as the background model. The forward and adjoint simulations are performed by the `SPECFEM3D_GLOBE` software package (Komatitsch and Tromp, 2002b,a). Topography, bathymetry, Earth's rotation, self-gravity, the ocean load, and attenuation are all taken into account during the simulations.

3.4.1 Computation of 3D adjoint sources and kernels

`FLEXWIN` (Maggi et al., 2009) is configured to select minor- and major-arc surface waves only. Following Bozdağ et al. (2016), surface-wave measurements are done in two period bands: 45–110 seconds and 90–250 seconds. The measurements are based on frequency-dependent (multitaper) traveltimes. To linearize the problem and avoid cycle skips in phase measurements the time shifts greater than half of the minimum period and the measurements larger than 4-standard deviations are rejected. The advantage of DD measurements increases parallel to the number of selected pairs. We observe that 90% waveform similarity used in 2D synthetic experiments degrades the number of paired stations significantly in 3D. As we can model every wiggle in our selected windows with the help of 3D numerical wave simulations, we set the waveform similarity to 80%. We pair stations within a diameter of 500 km distance where minor- and major-arc surface waves are paired in their own categories. USArray stations results in the highest number of pairs (see Table 3.1). The stations in Europe and Asia are also paired to a lesser extent (Fig. 3.14).

3.4.2 3D gradients

3D global gradients are computed using four traveltime misfit functions similar to those used in 2D experiments but for multitaper misfits (Fig. 3.15). Although it is not enough to perform global inversions, to reduce the computational cost we used data only from the same evenly distributed 21 earthquakes in 2D experiments. We observe that multitaper measurements with no receiver weighting ($\mathrm{MT}^{\mathrm{all}}$) have the most visible wave-path signature, especially across Pacific due to the USArray stations in North America. The receiver weighting ($\mathrm{MT}^{\mathrm{all}}_{\mathrm{rec}}$) significantly reduces these path signatures while down-weighting the sensitivity underneath USArray. On the contrary, the DD measurements emphasize the structure underneath North America, and the path-signature outside of the station clusters is also significantly reduced, even for the $\mathrm{MT}^{\mathrm{DD}} + \mathrm{MT}^{\mathrm{single}}$ misfit. In case of $\mathrm{MT}^{\mathrm{DD}} + \mathrm{MT}^{\mathrm{all}}_{\mathrm{rec}}$ the gradient looks more balanced. However, the advantage of the DD measurements seems slightly suppressed underneath USArray. Using MT measurements for only the non-paired stations ($\mathrm{MT}^{\mathrm{DD}} + \mathrm{MT}^{\mathrm{single}}$) may display the advantage of DD better. This is likely because of the interference of two misfit functions where classical measurements may degrade the resolution power of DD.

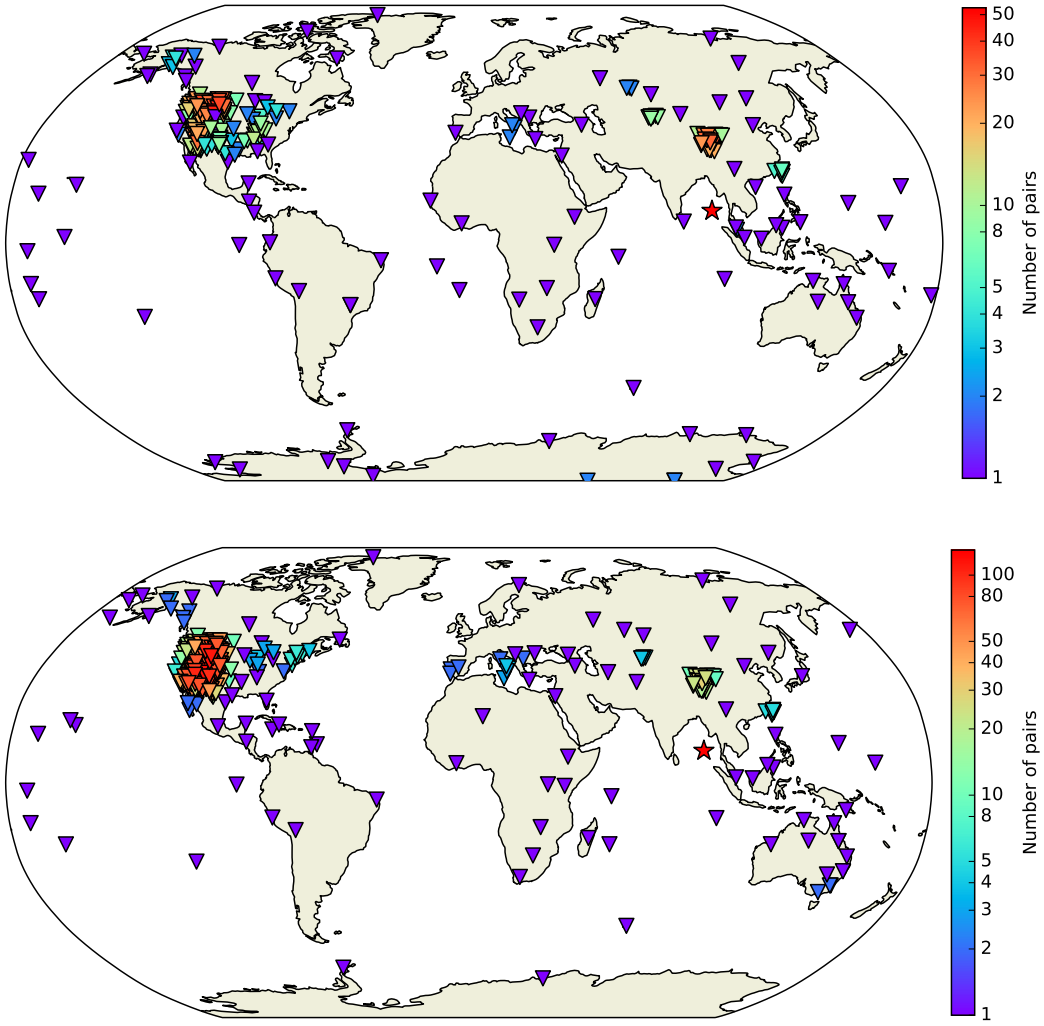

FIGURE 3.14: Number of pairs for double-difference measurements in period bands of 45-110 s (*top*) and 90-250 s (*bottom*) for 2008 Andaman Islands, India earthquake ($M_w = 6.1$, depth = 15 km).

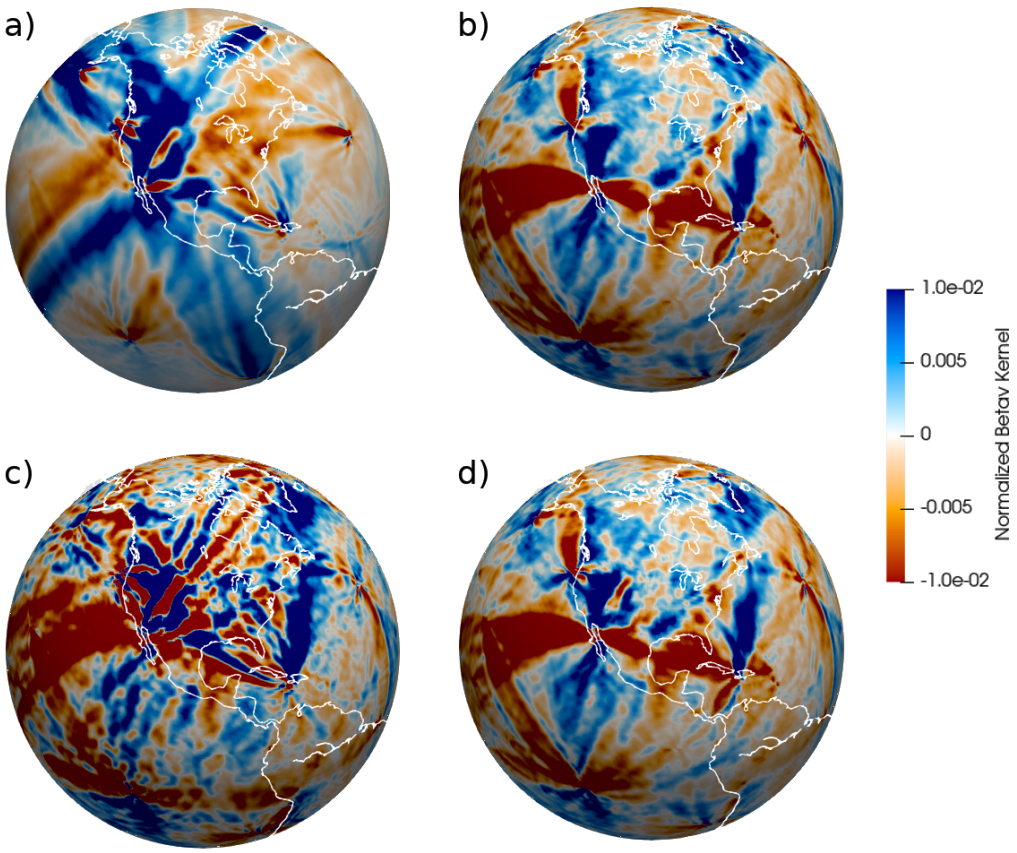

FIGURE 3.15: Horizontal cross-sections of 3D global gradients of vertically-polarized shear wavespeeds at 74 km computed for a) multitaper traveltime measurements $\mathrm{MT}^{\mathrm{all}}$ on all stations, b) multitaper measurements on all stations with the receiver weighting ($\mathrm{MT}^{\mathrm{all}}_{\mathrm{rec}}$), c) double-difference multitaper measurements combined with classical multitaper measurements on non-paired stations ($\mathrm{MT}^{\mathrm{DD}} + \mathrm{MT}^{\mathrm{single}}$), d) double-difference multitaper measurements combined with classical multitaper measurements on all stations with the receiver weighting ($\mathrm{MT}^{\mathrm{DD}} + \mathrm{MT}^{\mathrm{all}}_{\mathrm{rec}}$). All double-difference measurements are made on paired stations with the pair-wise weighting.

3.5 Discussion

The above experiments show that it is easier to pair stations in 2D examples due to their relatively simple waveforms. In 3D, however, mainly the crustal complexity can significantly affect waveforms, especially those of surface waves. Therefore the station pairs must be selected carefully depending on geological context. In particular, it is useful to separate stations based on whether they lie on oceanic or continental crust, and to employ geographical distance and correlation selection criteria, which help mitigate nonlinearity. In addition, although the selected windows do not necessarily have to isolate single phases, minor- and major-arc surface waves or any body-wave window of interest should be paired separately in their own categories to make measurements tractable and avoid the nonlinearity problem of adjoint inversions.

DD measurements applied to clusters of stations can very easily exacerbate bias in tomographic images due to uneven station distribution. Introducing pair-wise weighting to DD misfits not only alleviates this problem, but also helps to geographically balance gradients. In 2D experiments, the model misfit reduction is comparable to those of classical misfits with the receiver weighting. We recommend using DD misfits always with the pair-wise weighting to balance measurements by their own number of pairs.

In 2D global experiments, we observe that DD measurements speed up the convergence specifically underneath station clusters in iterative inversions. This result is mainly due to 1) the increased convergence rate underneath densely covered regions, 2) balanced measurements and gradients with the pair-wised weighting, which is comparable to any weighting scheme based on station distribution. Even though our 2D global results show that classical measurements with the receiver weighting may give slightly better model misfit reductions, DD model misfits are comparable to those and give better resolution in areas with dense seismic networks. 2D regional experiments confirm these results where DD measurements can significantly increase the convergence rate compared to classical measurements with and without the receiver weighting (see Appendix). DD gives slightly worse model misfit reduction in global inversions likely because of the limited area on the globe covered by dense networks. On the other hand, combining two or more misfit functions, for instance DD for paired stations and classical misfits for all stations with the receiver weighting, may allow for extracting more information from data, as suggested by 2D experiments. However, the combination of multiple misfit functions may require additional weighting not to conceal the advantage of DD measurements. Alternatively, one may also consider using classical and DD measurements separately in subsequent iterations. Furthermore, it is common to smooth gradients at the global scale to further regularize the imperfect coverage which may also degrade the resolution power of DD measurements. In such a case a multi-scale smoothing strategy (e.g., Bozdağ et al., 2016) may be a good choice, where the smoothing operator may be defined as a function of data coverage such that densely covered regions are smoothed less.

In this study, the experiments were performed using 21 well-distributed earthquakes. Therefore, it was not crucial to balance the source distribution, and we set the weight terms to one, giving the same importance to all earthquakes. In full-scale global tomographic applications, balancing the distribution of sources is also recommended (e.g., Ruan et al., 2019) because earthquakes are generally distributed in a lopsided way along plate boundaries.

Given similar selection criteria, the number of selected pairs is generally less in 3D compared to 2D due to the complexity of waveforms. As a result, the advantage of DD measurements is less pronounced in 3D gradients, which may be improved by relaxing the pairing criteria, but with a caution to avoid the nonlinearity problem in full-waveform inversions. The DD version of the exponentiated-phase misfit, a variant of the instantaneous phase misfit (Bozdağ et al., 2011), may be a good alternative which better deals with the cycle-skip problem (Yuan et al., 2020). The advantage of DD measurements will naturally be more noticeable by increased number of stations, for instance, with the inclusion of PASSCAL arrays and regional networks after careful data quality check.
We considered surface waves only in this study, using both minor- and major-arc measurement windows. The result of applying DD measurements to surface waves is very intuitive, with the paired measurements directly highlighting structure beneath station clusters. Both theoretically and in practice, DD measurements can be applied to any measurement window, including body waves. However, use of body-wave DD measurements requires further investigation to understand its practical effects in real inversions. Meanwhile, a careful combination of DD surface-wave measurements and classical body-wave measurements with receiver weighting may be a good alternative to assimilate all wave types in inversions. While addition of classical measurements to DD traveltime measurements removes some of the source uncertainty advantages shown by Yuan et al. (2016), DD measurements remain useful for reducing potential systematic time errors within individual station clusters.

3.6 Conclusions

DD measurements can easily be adapted to the adjoint tomography workflow. They do not cost anything extra in wavefield simulations (forward and adjoint simulations), since measurements from all selected windows and misfit functions are combined during the construction of adjoint sources. Only the expense of computing adjoint sources increases, but this data processing expense remains insignificant compared with the cost of numerical wavefield simulations.

We demonstrate DD measurements with a realistic global source-receiver distribution with 2D synthetic adjoint inversions and 3D global gradients based on real data. We introduce a pair-wise weighting to DD misfits based on the number of pairs of each station, or measurement window. To develop insight on how DD measurements should be used in practice, we compare the results to those from classical inversions with and without receiver weighting. Having used DD measurements with surface waves only, application to body wave windows remains a topic for future investigation.

Our experiments show that pair-wise weighted DD measurements are useful for assimilating data from station clusters and for faster convergence under densely-covered regions. The pair-wise weighting term also helps balance global data coverage by reducing the bias around dense station clusters. From the standpoint of the model misfit reduction, the effect of the pair-wise weighting in DD inversions is comparable to the effect of the receiver weighting in classical inversions at the global scale. To draw on advantages of both techniques and maximize the information extracted from data, DD measurements may be combined with classical measurements by including both types of terms in the misfit function.

Acknowledgements

We gratefully acknowledge Editor Jean Virieux and two anonymous reviewers for their constructive feedback. We thank Yanhua Yuan for her initial help and making the 2D synthetic tests from Yuan et al. (2016) available. We gratefully acknowledge the fruitful discussions with Dimitri Komatitsch, Guust Nolet, Stephané Operto and Jeroen Tromp. For 3D simulations, this research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under contract DE-AC05-00OR22725. Additional computational resources were provided by Laboratoire Géoazur, Université Côte d'Azur, Colorado School of Mines and the Princeton Institute for Computational Science & Engineering (PICSciE). We acknowledge IRIS (iris.edu) and ORFEUS (orfeus-eu.org) for providing the data used in this study. The open source spectral-element software package SPECFEM3D_GLOBE and the seismic measurement software package FLEXWIN used for this article are freely available via the Computational Infrastructure for Geodynamics (CIG; geodynamics.org). RO's doctoral studies are supported by the French Ministry of Education. EB is supported by her UNS–CNRS Chaire d'Excellence grant at Université Côte d'Azur and the start-up grant at Colorado School of Mines. The 2D examples presented in this manuscript and the computer codes are available at https://github.com/rdno/dd_global. Picasso's Guernica used in the manuscript as a target model is slightly modified and converted to greyscale while implementing it into the 2D wave propagation solver for numerical reasons. Picasso's Guernica is used in this manuscript with the permission of © 2019 Estate of Pablo Picasso / Artists Rights Society (ARS), New York.

3.7 Appendix A: Regional checkerboard tests

In this appendix, we compare classical and double-difference cross-correlation traveltime measurements with various weighting strategies on a regional example to have better insight on the performance of each measurement and the effect of distance and waveform similarity criteria on inversions. Starting from a homogeneous background model we invert for a target checkerboard model using a realistic data coverage from Southern California (Tape et al., 2007) as shown in Figure 3.16 (top first two figures).

In Figure 3.16, final models after 10 iterations with classical cross-correlation traveltimes without (CC^{all}) and with (CC^{all}_{rec}) the receiver weighting for two different reference distances (see Section 3.2.4), and double-difference traveltime measurements (CC^{DD}) of all stations and paired stations with various pairing criteria (see Sections 3.2.3 & 3.2.4) are presented. The receiver weighting based on the geographical distribution of stations (Ruan et al., 2019) speeds up the convergence compared to CC^{all} with no receiver weighting. Increasing the reference distance for computing the receiver weights gives slightly better results as the gradients are better balanced. Unlike the 2D examples with global coverage, where the advantage of double-difference measurements is not well pronounced due to limited cluster of stations globally, CC^{DD} measurements show better performance where the winner is CC^{DD} with pairing criteria 50% correlation between the waveforms within a 100-km search radius. This is mainly because the receiver weighting is suppressing station clusters to balance the gradient whereas double-difference measurements are taking their advantage into account. Figure 3.18 shows the receiver weight of each station for each reference distance considered in the tests. The reference distances

are selected based on the condition number (the ratio between the maximum and minimum weights). The reference distance of 43 km is related to the maximum condition number whereas 17.5 km is the value related to one third of the maximum condition number which is recommended in practice (Ruan et al., 2019). Increasing the reference distance more than 43 km decreases the condition number such that the receiver weighting loses its importance (i.e., results become similar to measurements with no receiver weighting).

Although in 3D experiments waveform similarity may play an important role in mitigating nonlinearity, in 2D 50% and 95% waveform correlations give similar results. On the other hand, increasing the search radius to 200 km clearly worsens performance as it is more challenging to decrease misfit by correlating waveforms which are significantly different from each other. Similarly, when all the stations are paired we observe the worst performance within all CC^{DD} measurements. Figure 3.19 shows the number of pairs of each station for each pairing criterion considered here. In a real 3D experiment we have no chance to demonstrate such experiments to find optimum pairing criteria for double-difference measurements. A 100 km search radius is close to the width of the first Fresnel zone $\sqrt{\lambda L}$ suggested by Yuan et al. (2016) (the mean path length and the dominant wavelength related to the source frequency are 340 km and 42 km, respectively, give an average width of the first Fresnel zone of about 120 km), which may be used as a starting point for search radius in addition to geological information in 3D studies followed by the selection criteria given in Section 3.2.3.

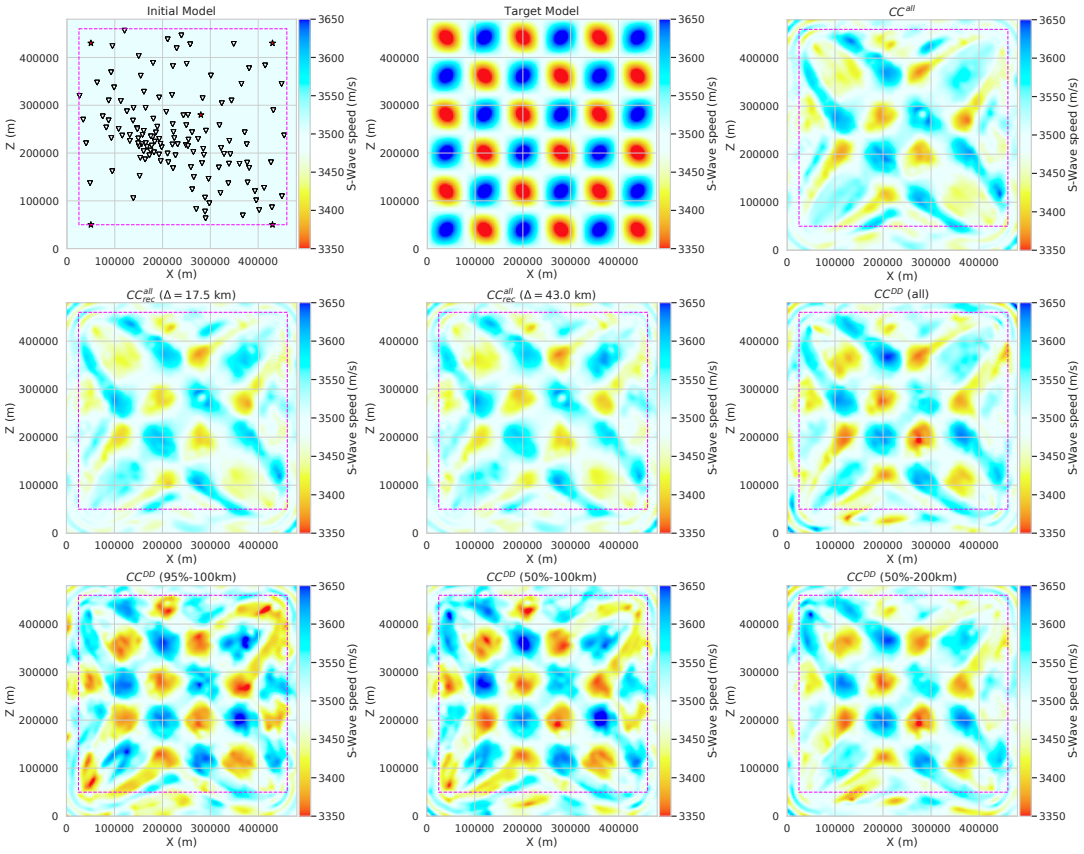

FIGURE 3.16: Inversion results after 10 iterations. CC^{all}: Cross-correlation traveltimes without the receiver weighting. CC^{all}_{rec}: Cross-correlation traveltimes with the receiver weighting with two different reference distances ($\Delta = 17.5$ km, $\Delta = 43.0$ km). CC^{DD} (all): Double-difference cross-correlation traveltimes where all stations are paired. CC^{DD} (95%-100km): DD measurements where stations within 100 km and with 95% waveform similarity are paired. CC^{DD} (50%-100km): DD measurements where stations within 100 km and with 50% waveform similarity are paired. CC^{DD} (50%-200km): DD measurements where stations within 200 km and with 50% waveform similarity are paired. Initial and target models are as shown in the top first two figures.

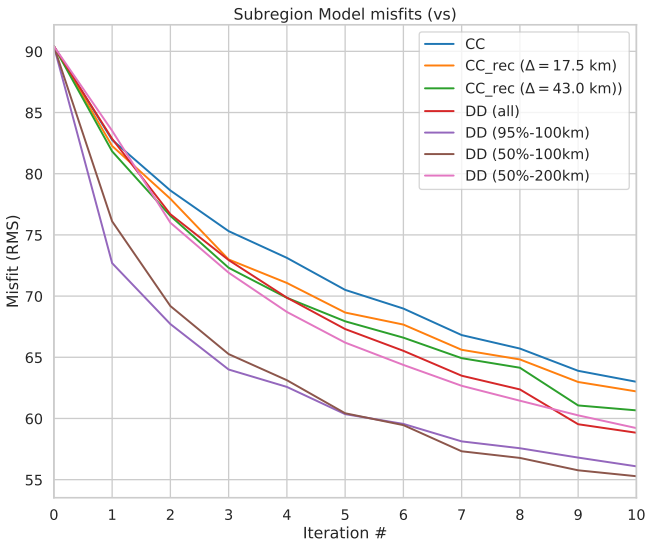

FIGURE 3.17: RMS model misfits for the models shown in Fig. 3.16. Misfits are computed for the region within the magenta rectangle shown in the first plot of Fig. 3.16.

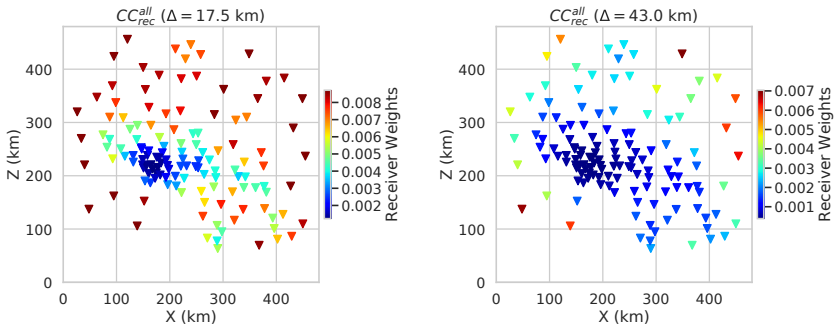

FIGURE 3.18: Receiver weights for the two reference distances used in cross-correlation traveltime measurements in Fig. 3.16.

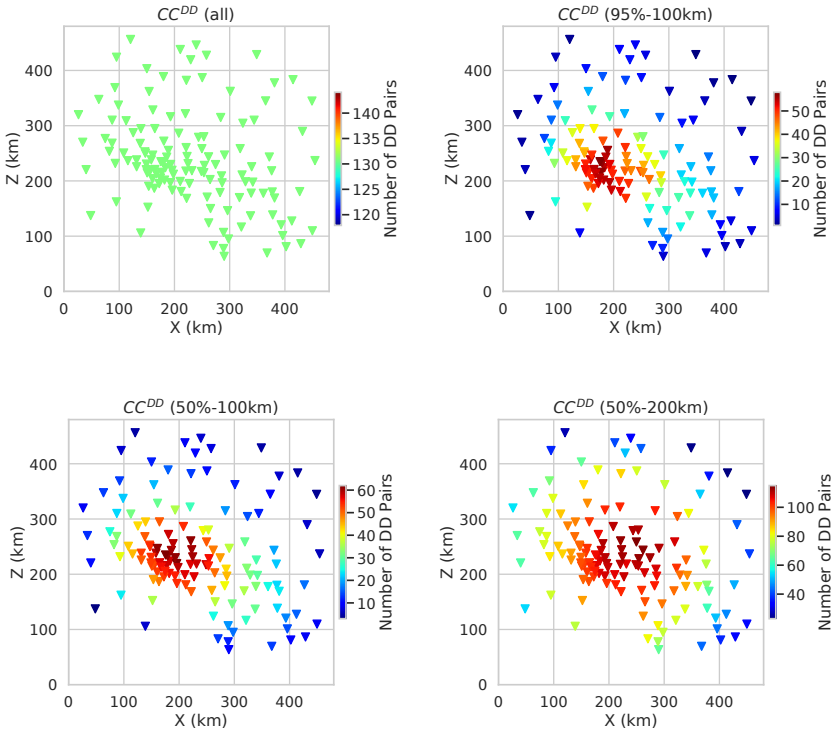

FIGURE 3.19: Number of station pairs used as pair-wise weighting for various pairing criteria used in double-difference traveltime measurements in Fig. 3.16.

Chapter 4

Surface-wave azimuthally anisotropic adjoint inversions based on double-difference measurements[1]

Summary

Earth's lithosphere and upper mantle show significant evidence of anisotropy as a result of its composition and deformation. Starting from the first-generation global adjoint tomography model GLAD-M25 (Lei et al., 2020), which is the successor of GLAD-M15 (Bozdağ et al., 2016) and transversely isotropic in the upper mantle, our goal is to construct a global adjoint model of the upper mantle by taking azimuthal anisotropy into account in the parameterization of the inverse problem. We focus on four elastic parameters that surface waves are known to be most sensitive to, namely, vertically and horizontally polarized shear waves and the shear modulus-normalized anisotropic parameters Gc' & Gs'. We performed 10 iterations using the frequency-dependent traveltime (multitaper) measurements of minor- and major-arc Rayleigh and Love waves on three components, where each waveform is normalized by its number of pairs in the period ranges 40-110 s & 90-250 s. In addition, source and receiver weightings based on their geographical distribution are applied to the measurements to balance the gradient of the misfit function. After the third iteration, we started combining multitaper traveltime measurements with double-difference multitaper measurements made on paired stations. New measurements result in better balanced gradients with the expectation of extracting more information underneath clusters of stations, such as USArray. Our initial results reveal multi-scale anisotorpic signals depending on data (kernel) coverage and getting closer to the continental-scale resolution in areas with dense coverage, consistent with previous studies. This is also the first time that we demonstrated double-difference measurements in a global-scale adjoint inversions as outlined in Örsvuran et al. (2020). The anisotropy signal is still not as strong as expected likely suggesting that more iterations are needed to be performed. Furthermore, the tradeoff between elastic and anelastic parameters shouuld be investigated by, for instance, point-spread function tests. The global azimuthally anisotropic simulations have been performed on the Oak Ridge National Laboratory's IBM AC922 Summit system using a dataset of 263 earthquakes where 253 of them were used in the construction of GLAD-M15.

[1]The content of this chapter will be submitted for publication: **Örsvuran, R.**, Bozdağ, E., & Peter, D., 2021. Global azimuthally anisotropic adjoint tomography of surface waves, *in prep.*

4.1 Introduction

Seismic anisotropy is a phenomenon where seismic wave velocities can be dependent on the propagation direction. Anisotropy is caused by the structural properties of the medium. In the upper mantle, anisotropy is largely attributed to the strain which causes lattice preferred orientation (LPO) of minerals (Nicolas and Christensen, 1987). Seismic anisotropy can also caused by the spatial organization (e.g. fine layering) of the materials in the medium which are defined by shape-preferred orientation (SPO). In the crust, it has been shown that cracks can also induce anisotropy (Crampin and Booth, 1985). Anisotropy has also been used for monitoring stress build-up by observing shear-wave splitting changes with time in the crust (Crampin et al., 1999), determining the deformation in the mantle due to plate tectonics (Silver and Chan, 1991; Debayle et al., 2005) and the mantle convection (Montagner, 1994).

The anisotropic behavior of seismic waves is observed for both body and surface waves. SKS and SKKS phases are used in shear wave splitting measurements to determine the anisotropy in the mantle (Silver and Chan, 1991; Vinnik et al., 1992). P-wave anisotropy have been observed both in laboratory (Babuška, 1984) and used to determine the crack parameters in the crust (Crampin et al., 1980). Surface waves are used for mapping the transverse isotropy and the azimuthal anisotropy in the upper mantle from regional to global scales (Nataf et al., 1984; Montagner and Tanimoto, 1991; Montagner, 1985; Trampert and Heijst, 2002; Kustowski et al., 2008; Schaeffer et al., 2016; Zhu et al., 2020).

Although they are constructed from different data sets, it has been shown that shear wave splitting and surface wave anisotropy studies can be related to each other which provides additional avenues to confirm and intrepret the models (Montagner et al., 2000). Becker et al. (2012a) suggests that long-wavelength features are consistent between surface wave tomography and SKS splitting measurements which are useful for geodynamical interpretations. They also suggest the use of dense seismic arrays for more detailed regional anisotropic tomography.

Ray based methods are commonly used in anisotropy studies since they are relatively inexpensive (Debayle et al., 2005; Kustowski et al., 2008). However, it has been shown that accurately accounting for crustal complexity is crucial for the success of the inversions (Bozdağ and Trampert, 2008). Full waveform inversion techniques are able to take the physics of wave propagation correctly thus avoiding the crustal corrections.

GLAD-M15 (Bozdağ et al., 2016) and GLAD-M25 (Lei et al., 2020) used transversely isotropic upper mantle. The goal of this study is to introduce the azimuthal anisotropy into upper mantle models using the full waveform inversion method (Tromp et al., 2005). Using global adjoint tomography model GLAD-M25 (Lei et al., 2020), we performed 10 conjugate-gradient iterations including azimuthal anisotropy in our parameterization. Minor and major arc surface waves are used for three component seismograms. Double-difference measurements are used to increase resolution underneath the dense station networks. Results show large scale agreement with other models while being limited variance with the depth.

4.2 3D numerical simulations and the adjoint inversion

4.2.1 Starting model

We started global azimuthally anisotropic adjoint inversions using the recent global mantle model GLAD-M25 (Lei et al., 2020), which is the successor of the first-generation global adjoint tomography model GLAD-M15 (Bozdağ et al., 2016). GLAD-M25 was constructed by performing 10 more L-BFGS iterations starting from GLAD-M15 and increasing the data used in inversions by about 6 times compared to those used during the construction of its starting model. GLAD-M15 was obtained after 15 conjugate gradient iterations starting from the ray-based global mantle model S362ANI (Kustowski et al., 2008) together with 3D crustal model Crust2.0 (Bassin et al., 2000). GLAD-M25 has the same parameterization of GLAD-M15 and its starting model S362ANI with transverse isotropy confined to the upper mantle (between Moho down to 670 km discontinuity). GLAD-M25 is the result of the simultaneous inversion of the mantle and crust avoiding crustal corrections.

4.2.2 Numerical simulations

We use the GPU version of the 3D global wave propagation solver SPECFEM3D_GLOBE package (Komatitsch and Tromp, 2002a; Komatitsch et al., 2002) to simulate synthetic seismograms and to compute the gradient of the chosen misfit function (i.e., data sensitivity, Fréchet, kernels) using 3D global mantle and crustal models during iterative inversions. The resolution of numerical simulations is NEX=256 (NEX denotes the number of spectral elements at the surface on one side of each of 6 chunks that form the globe in spectral-element simulations) which resolves seismic waveforms down to about 17 s. Topography, bathymetry, attenuation, gravity (Cowling approximation), ellipticity and the rotation of the Earth are all taken into account in all forward and adjoint simulations.

We used Oak Ridge National Laboratory's IBM AC922 Summit system to perform 10 iterations. For a single event, we used 384 GPUs. Forward simulations took $\sim$ 3 minutes and adjoint simulations took $\sim$ 9 minutes in real time. The preprocessing stage, that involves data processing, measurement window selection and computation of adjoint sources, was performed on the Oak Ridge National Laboratory's Rhea system. Using 16 CPU cores, processing of synthetic data took $\sim$ 6 minutes whereas that of observed data took $\sim$ 12 minutes which involved the removal of instrument responses). The measurement window selection stage took $\sim$ 10 minutes whereas the adjoint source calculations took $\sim$ 5 minutes per event with $\sim$ 1000 stations in real time. Batch jobs were submitted to multiple nodes speed up the effective real time waiting period for the data processing and numerical simulations.

4.2.3 Data & measurements

In this study, our target is to image the azimuthal anisotropy in the upper mantle. Considering our parameterization (see next section) in the inverse problem, we use minor- and major-arc surface waves only excluding body waves. Using surface waves only also give us the opportunity to directly demonstrate double-difference measurements in global adjoint tomography as described in the previous chapter.

We started iterations with the dataset of 253 earthquakes used in the construction of GLAD-M15 (Bozdağ et al., 2016). After the fourth iteration, 10 more earthquakes were added to the database. We also included additional data from the European

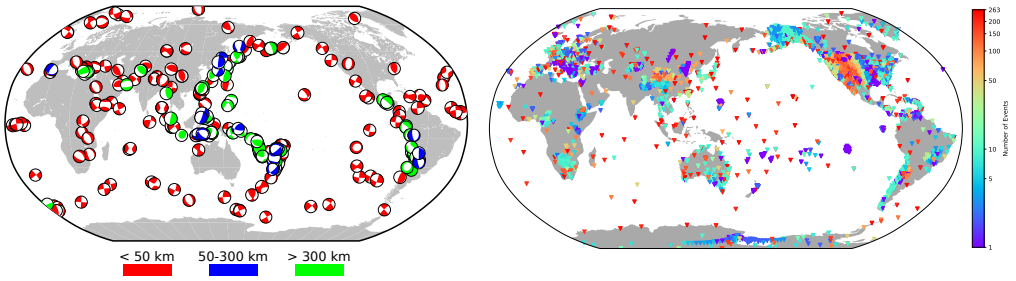

FIGURE 4.1: The distribution of 263 earthquakes (left) and seismic stations (right) that are used to construct the azimuthally anisotropic global adjoint model. 253 of the events were also used in the construction of the "first-generation" global adjoint tomography model GLAD-M15 (Bozdağ et al., 2016). The events are selected within the moment magnitude range 5.5–7.0. The colorbar of seismic stations denote how many times each station is used for 263 earthquakes.

networks through the ORFEUS data center. Figure 4.1 shows the distribution of 263 earthquakes and seismic stations used in this study.

We combine minor- and major-arc surface waves in the period ranges 40–110 s & 90–250 s. In Figure 4.2, we show sample seismic traces where surface waves are selected using the Python version of the automated measurement window selection algorithm FLEXWIN (Maggi et al., 2009) by restricting the selection criteria to select minor- and major-arc surface waves only. We used frequency-dependent cross-correlation traveltime measurements (multitaper traveltime misfit (Laske and Masters, 1996)) targeting the elastic structure only. We balanced the global data coverage as much as possible using a weighting scheme based on the distribution of seismic sources and stations as described in Ruan et al. (2019) and also in the previous chapter. After the third iteration, we started incorporating double-difference measurements of multitaper traveltimes (Yuan et al., 2016) with the normalization term of the number of measurement pairs used to balance the sensitivity of paired stations (or measurements) (Örsvuran et al., 2020). In this case we defined our misfit function as the summation of multitaper traveltimes weighted by the source and receiver distribution, and double-difference multitaper measurements made on paired stations with the weighting term of number of measurement pairs. All measurement categories (i.e., three-component data at different period bands) are also normalized based on their number of measurements to make sure that each measurement category is contributing to the model updates equally. The number of measurements for each category are summarized in Table 4.1. Figure 4.3 shows the number of measurement pairs for each measurement category used in double-difference measurements. Minor- and major-arc Rayleigh and Love waves were paired separately in their own measurement categories where we have the most measurement pairs in North America due to the USArray stations. In addition, Love-wave pairs are, in general, lower than those of Rayleigh waves, likely because of the stronger effect of the crust on Love waveforms. Similarly, the number of major-arc pairs are much less than those of minor-arc. The stations within a diameter of 250 km were paired and the waveforms with less than 80% correlation were rejected. Each double-difference measurement was then weighted by its number of pairs.

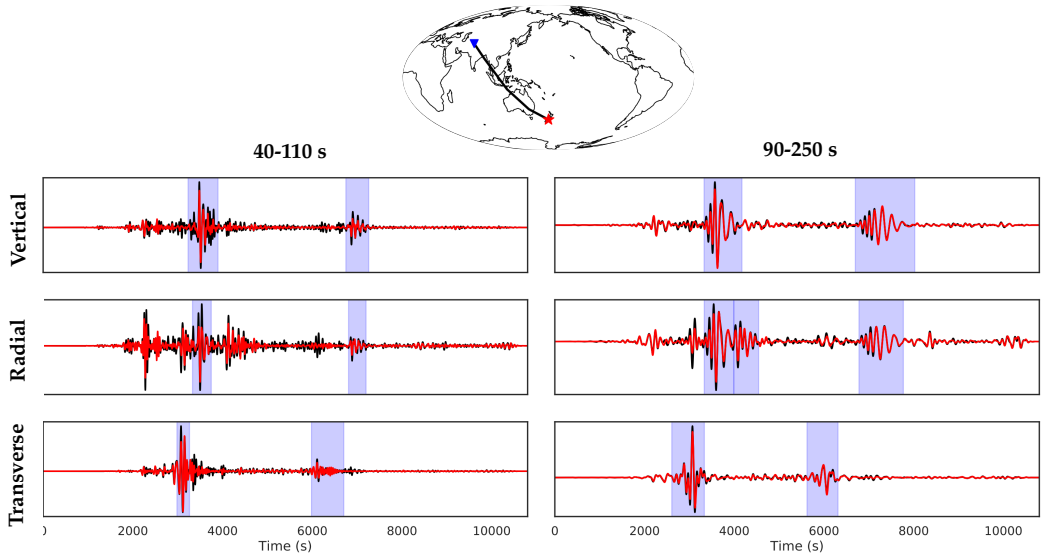

FIGURE 4.2: Automated window selection (using FLEXWIN (Maggi et al., 2009) algorithm and pyflex package) using M35 model. Traces are shown is for 3rd September 2010 New Zealand Earthquake at station KBL at Kabul, Afghanistan. Both minor and major arc surface selections are made possible with the iterative inversions.

TABLE 4.1: Number of measurements selected for two period bands and three components (vertical (Z), radial (R) and transverse (T)) at three inversion stages. First selections are done using the database with 253 events, second selections are done with the introduction of new event and stations. Third selections are done with the same dataset but with improved synthetics.

	40-110 s Z	40-110 s R	40-110 s T	90-250 s Z	90-250 s R	90-250 s T
M25	126,127	61,106	101,596	139,946	20,963	34142
M28	128,360	62,555	104,171	142,029	21,742	35251
M35	131,876	64,713	107,798	142,411	21,740	35,556

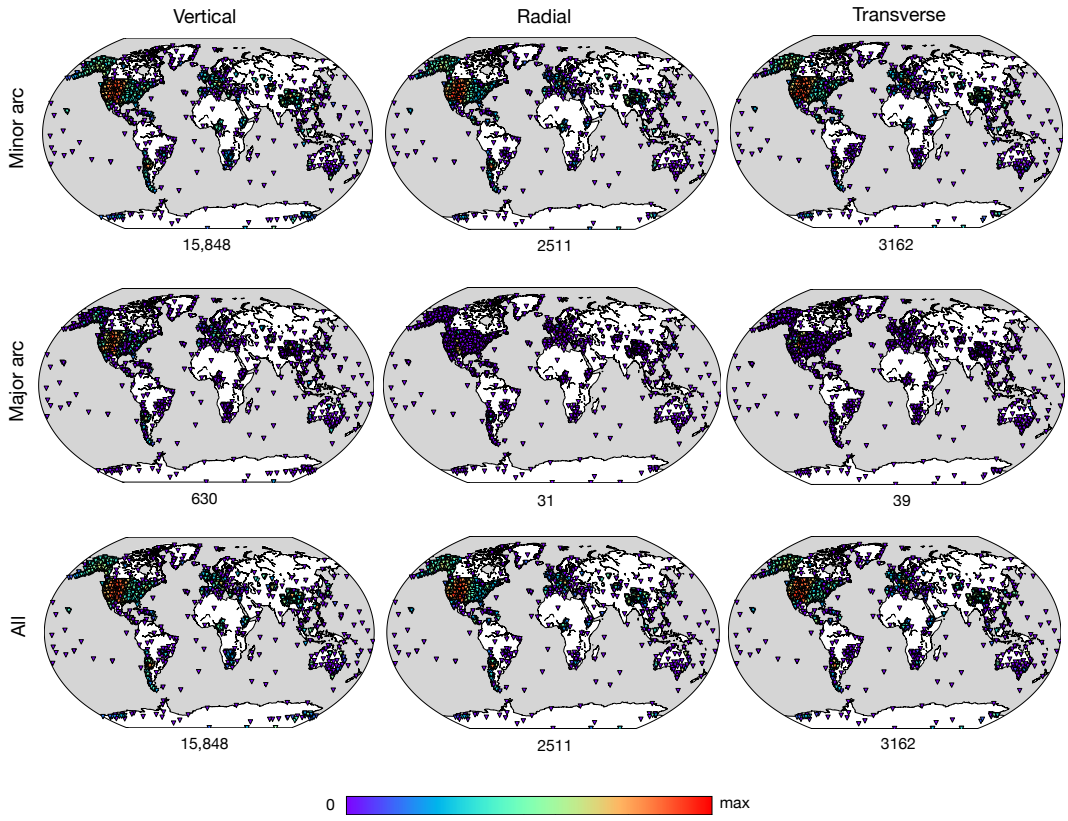

FIGURE 4.3: Minor arc, major arc and total number of pairs for vertical, radial and transverse components. Pairs are shown in color logarithmically.

4.2.4 Parameterization

Following Smith and Dahlen (1973) and Montagner and Nataf (1986) surface-wave phase speed in a weakly anisotropic medium may be expressed by

$$c(\omega,\theta) = c_0(\omega) + c_1(\omega)\cos(2\theta) + c_2(\omega)\sin(2\theta) + c_3(\omega)\cos(4\theta) + c_4(\omega)\sin(4\theta)\,. \tag{4.1}$$

Surface-wave anisotropy may be represented by 13 parameters: c_0 term denotes the radial anisotropy (Love parameters A, C, L, N, F), 2θ terms c_1 and c_2 ($G_{c,s}$, $H_{c,s}$, $B_{c,s}$) and 4θ terms c_3 and c_4 ($E_{c,s}$). Since it is not possible to solve all 13 parameters independently, following previous azimuthally anisotropic surface wave tomographic studies (Zhu et al., 2015; Schaeffer et al., 2016) we parameterized our inverse problem combining the radial anisotropy with the normalized azimuthally anisotropic parameters G_c' and G_s' which are given as

$$G_c' = G_c/(\rho\beta_0^2)\,, \tag{4.2}$$

$$G_s' = G_s/(\rho\beta_0^2)\,, \tag{4.3}$$

where ρ and β_0 are the density and isotropic shear wavespeed of the reference model. Then we define the variation in the misfit as

$$\delta\chi = \int \delta\ln\beta_h\, K_{\beta_h} + \delta\ln\beta_v\, K_{\beta_v} + \delta G_c'\, K_{G_c'} + \delta G_s'\, K_{G_s'}\,\mathrm{d}^3\mathbf{x} \quad , \tag{4.4}$$

where β_h and β_v denote the horizontally and vertically polarized shear wavespeeds, respectively, and G_c' and G_s' denote the normalized azimuthally anisotropic parameters G_c and G_s. Kernels for azimuthal anisotropy terms can be defined in terms of the kernels of the elastic tensor coefficients (Sieminski et al., 2007a,b)

$$K_{G_c} = K_{C_{55}} - K_{C_{44}}\,, \tag{4.5}$$

$$K_{G_s} = -K_{C_{45}}\,, \tag{4.6}$$

where the $K_{C_{ij}}$ kernel may be computed as described in Tromp et al. (2005). Note that we keep the bulk-sound speed and the dimensionless transverse isotropy parameter η fixed in these iterations while restricting transverse isotropy to the upper mantle.

4.2.5 Post-processing of gradients, model updates

We obtain the gradient of the misfit function by summing up event kernels (Fréchet kernels of each earthquake) at each iteration. Before we update the model we smooth and precondition the gradient to speed up the convergence.

We smooth the gradient using a Gaussian smoothing operator in horizontal and vertical directions. We used a Gaussian function with a half width of 120 km (started from 180 km in the first iteration) in the horizontal direction. We gradually increase the smoothing in the radial direction from 5 km half width in the lithosphere to 75 km half width in the lower mantle. We then use the diagonal pseudo-Hessian kernel as a preconditioner computed by the interaction of the forward and adjoint acceleration wavefields (Luo et al., 2014):

$$P(\mathbf{x}) = \int \partial_t^2\mathbf{s}(\mathbf{x},t)\cdot\partial_t^2\mathbf{s}^\dagger(\mathbf{x},T-t)\mathrm{d}t\,, \tag{4.7}$$

where $\mathbf{s}$ and $\mathbf{s}^\dagger$ are the forward and adjoint displacement wavefields, respectively, integrated over time.
We then update models using the conjugate gradient method (Fletcher and Reeves, 1964):

$$\mathbf{d}_i = \mathbf{g}_i + \beta \mathbf{d}_{i-1}\,, \tag{4.8}$$

where $\mathbf{d}$ and $\mathbf{g}$ are the search direction and the gradient of the misfit function, respectively, and the indices i and $i-1$ denote the current and the previous iteration numbers. β is given as

$$\beta = \frac{\mathbf{g}_i^T.(\mathbf{g}_i - g_{i-1})}{\mathbf{g}_{i-1}^T \cdot g_{i-1}}\,. \tag{4.9}$$

Following Tape et al. (2009) and Bozdağ et al. (2016) we determined the step length for model updates by performing a line search using a subset of data of 24 earthquakes as representative of our entire database.

4.3 Results after 10 iterations

In Figure 4.4, horizontal cross-sections of model perturbations for β_h, β_v, G_c' and G_s' for the first iteration at a depth of 150 km are presented. Since there is no azimuthal anisotropy in our starting model GLAD-M25 and the 1D reference model 1DREF (Kustowski et al., 2008) used to compute the perturbations, G_c' and G_s' perturbations are small and are the same as the conjugate gradient direction.

Figure 4.5 shows the total misfit together with the misfit reduction in each measurement category after 10 conjugate gradient iterations. We observe that the misfit curve has started flattening at the 9th iteration. We plan to continue iterations by decreasing the smoothing further and adjusting the strategy for iterations such as decreasing the period content of surface waves, implementing multi-scale smoothing (Bozdağ et al., 2016) and the L-BFGS optimization technique.

In Figure 4.9 we present our results after 10 iterations by comparing them to a global and continental models. We observe that after 10 iterations the magnitude of azimuthal anisotropy in the upper mantle has not been captured well yet. However, the overall global pattern is promising. Our adjoint models are sampled on the mesh used in numerical simulations. To make them comparable to current global and continental models we resampled them at every 5×5 and 2×2 degrees, respectively. The comparison to a recent global anisotropic upper mantle model by Schaeffer et al. (2016) show good agreement at 200 km. Our adjoint model does not show significant change in the radial direction which may suggest that the depth resolution may be limited. The comparison to the European model by Zhu et al. (2015) is promising to achieve higher resolution underneath continents with good sampling which is also partially the result of the double-difference measurements.

In Figure 4.10, we compare our results at 50 km underneath North America to those from a collection of SKS splitting measurements by (Becker et al., 2012b). Anisotropy determined from surface waves and shear wave splitting measurements do not necessarily agree with each other as both wave types sample different parts of the Earth. However, the similarity between the anisotropy signal especially on the west coast likely suggest that the source of anisotropy in SKS splitting measurements is related to the lithosphere and might be explained by the deformation and flow motions related to subduction (Zhu et al., 2020).

These results are preliminary and we plan to perform 5–10 more iterations to finalize the first azimuthally anisotropic model and perform resolution tests before publishing the model.

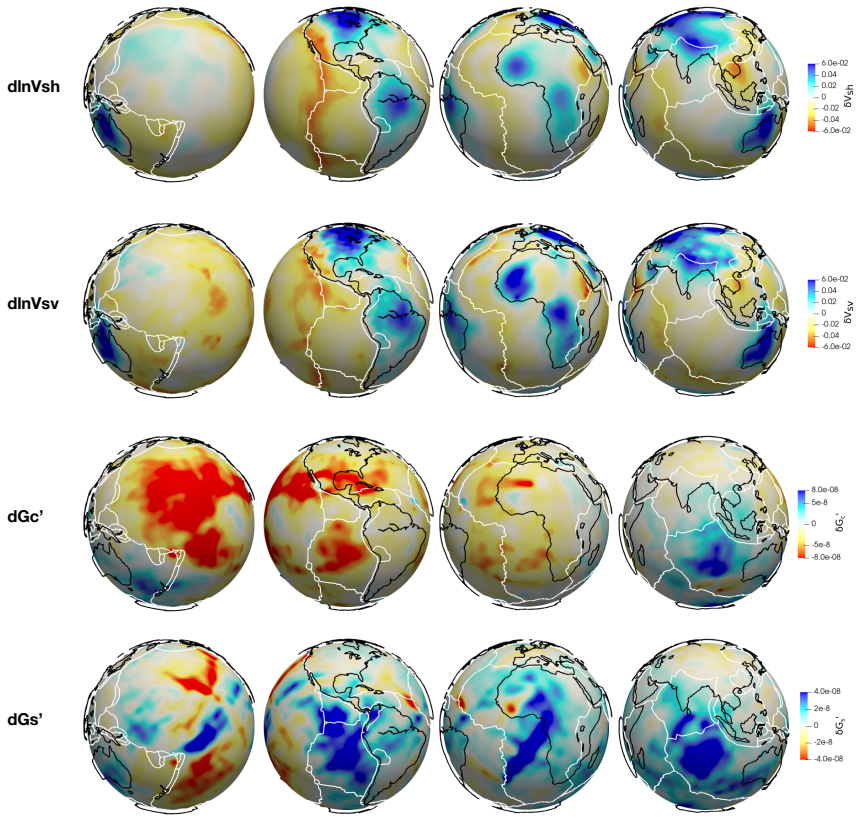

FIGURE 4.4: Perturbations in horizontally (δlnVsh) and vertically (δlnVsv) polarized shear wavespeeds and azimuthally anisotropic parameters $\delta G_c'$ and $\delta G_s'$. Note that the perturbations of the parameters are different. Since there is no azimuthal anisotropy in the starting model GLAD-M25, $\delta G_c'$ and $\delta G_s'$ are the same as the computed conjugate directions. $\delta G_c'$ and $\delta G_s'$ are normalized versions of δG_c and δG_c to be consistent with the other parameters. δlnVsh) and δlnVsv) lead to transverse isotropy in the upper mantle.

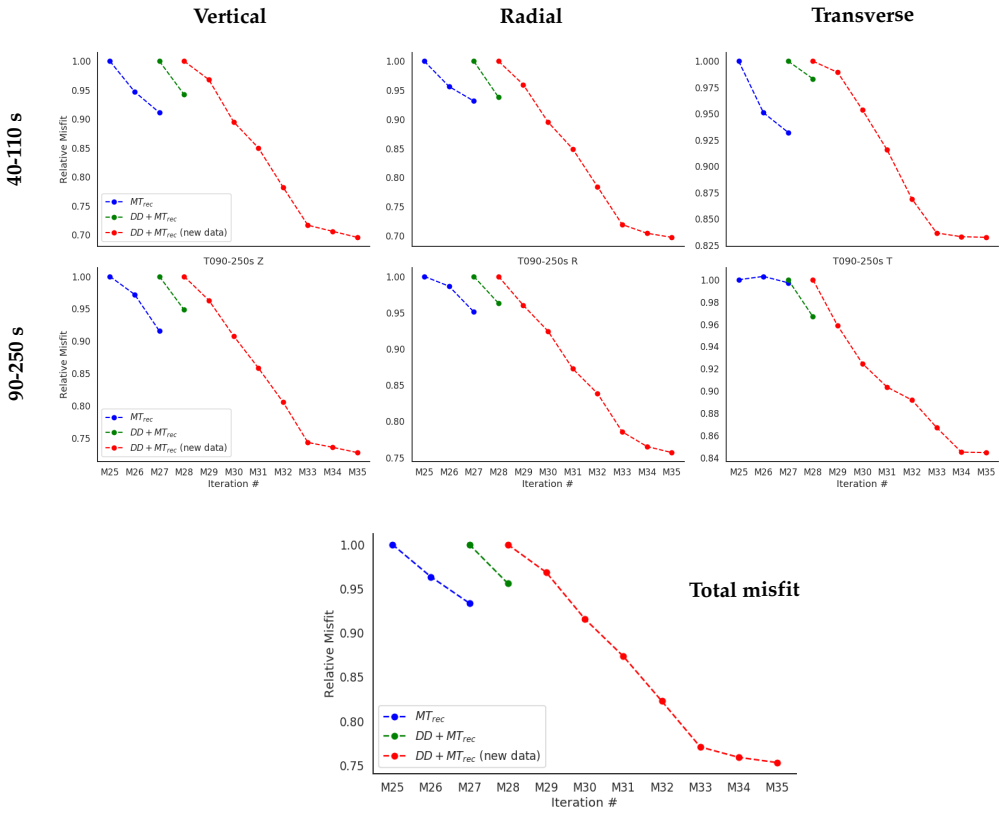

FIGURE 4.5: Total misfit reduction for three components and two period bands. After the second iteration regular multitaper measurements were combined with double-difference measurements made on paired stations, and after the third iteration 10 more earthquakes and ORFEUS stations were added to the database.

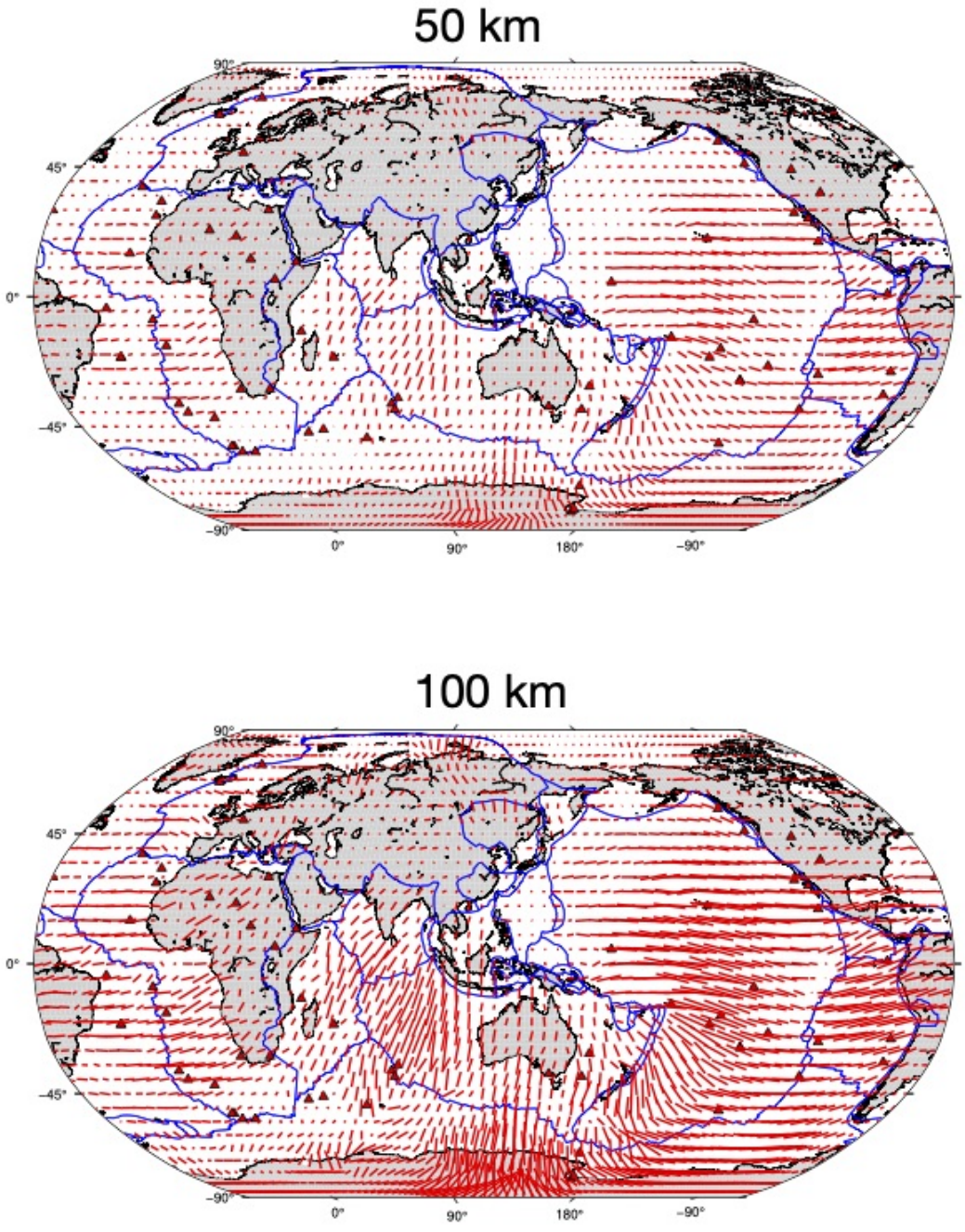

FIGURE 4.6: Cross-sections of azimuthal anisotropy at 50 km and 100 km. Anisotropy is sampled at every $5^\circ \times 5^\circ$. Red triangles denote hotspots.

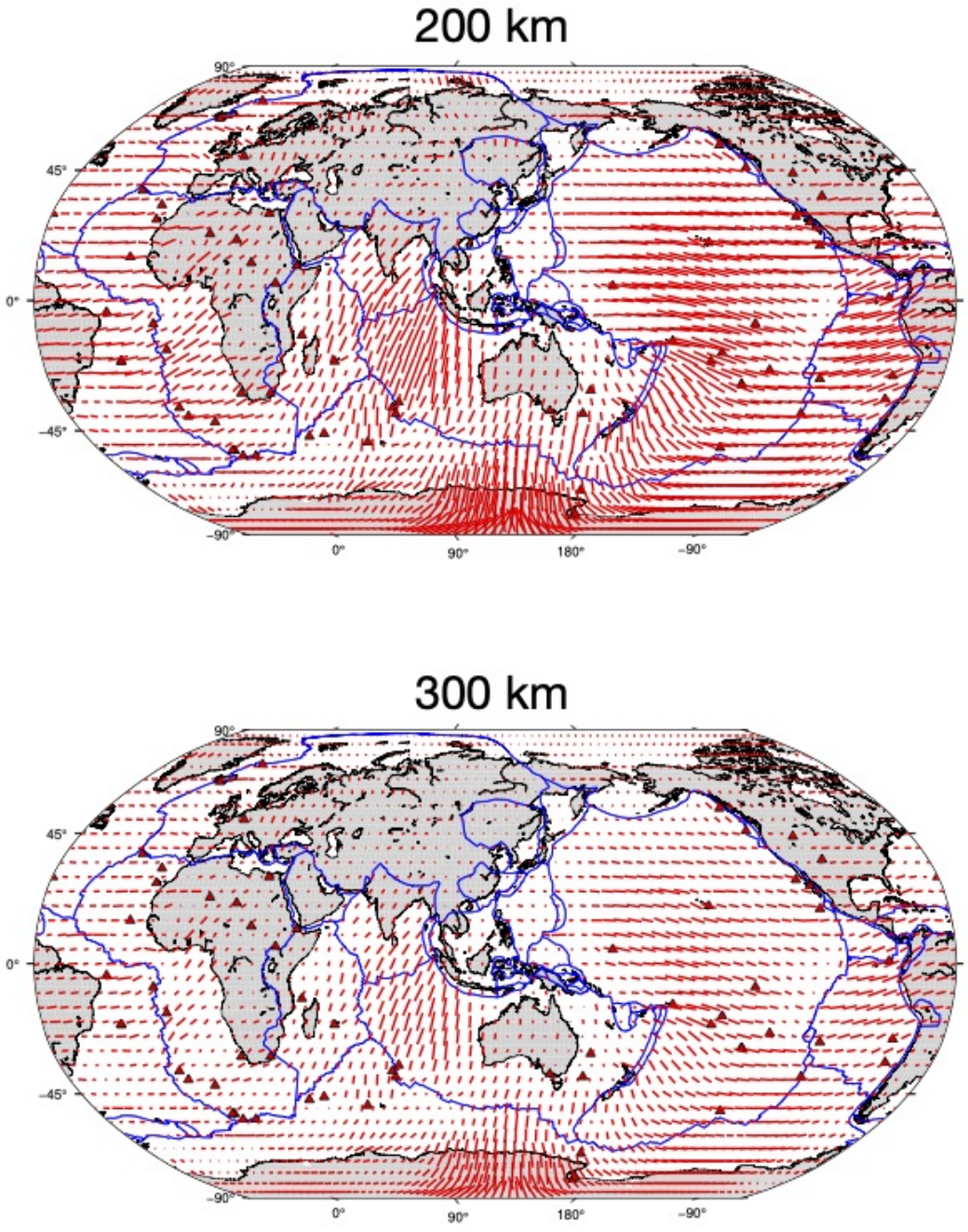

FIGURE 4.7: Same as Figure 4.6 but for 200 km and 300 km.

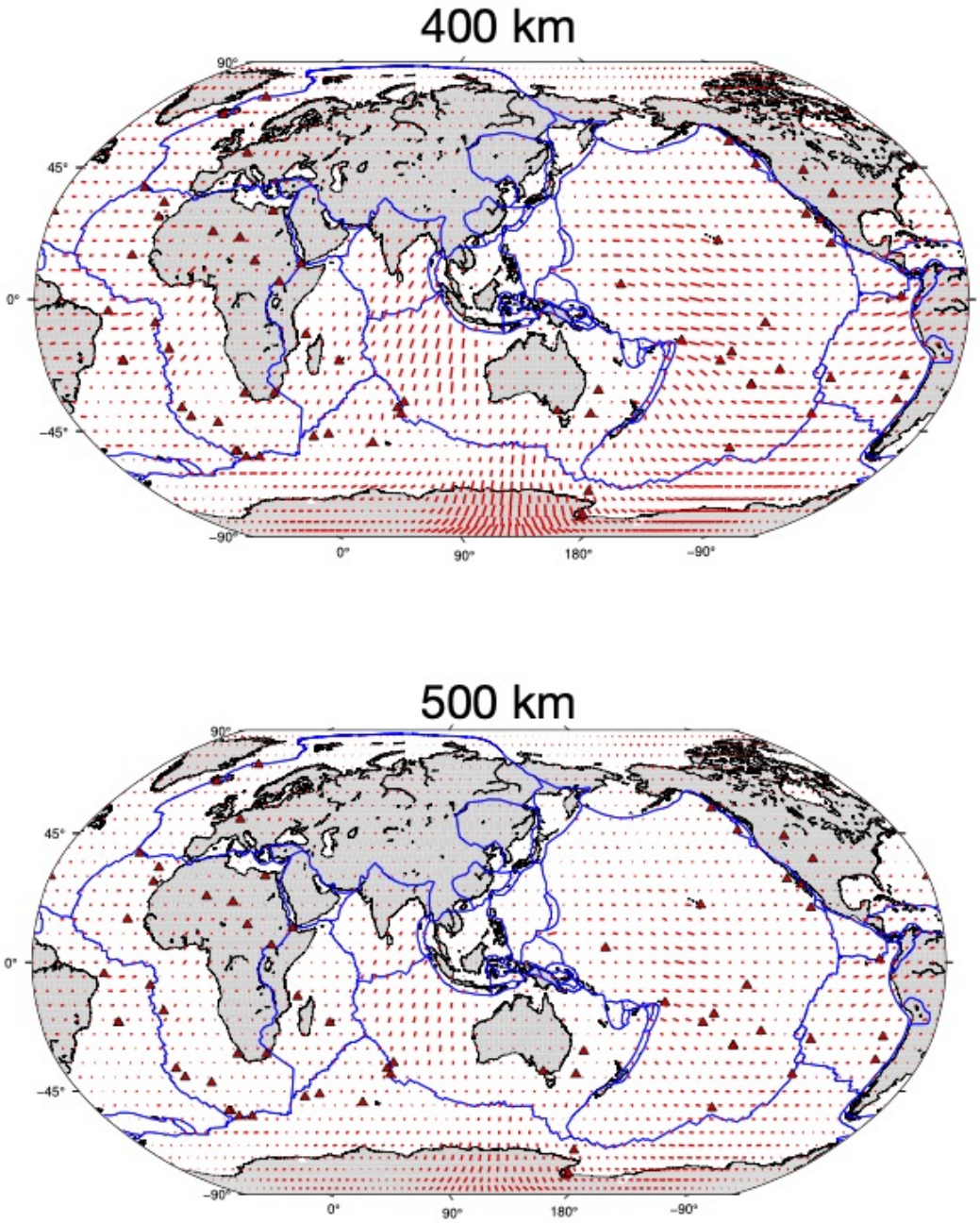

FIGURE 4.8: Same as Figure 4.6 but for 400 km and 500 km.

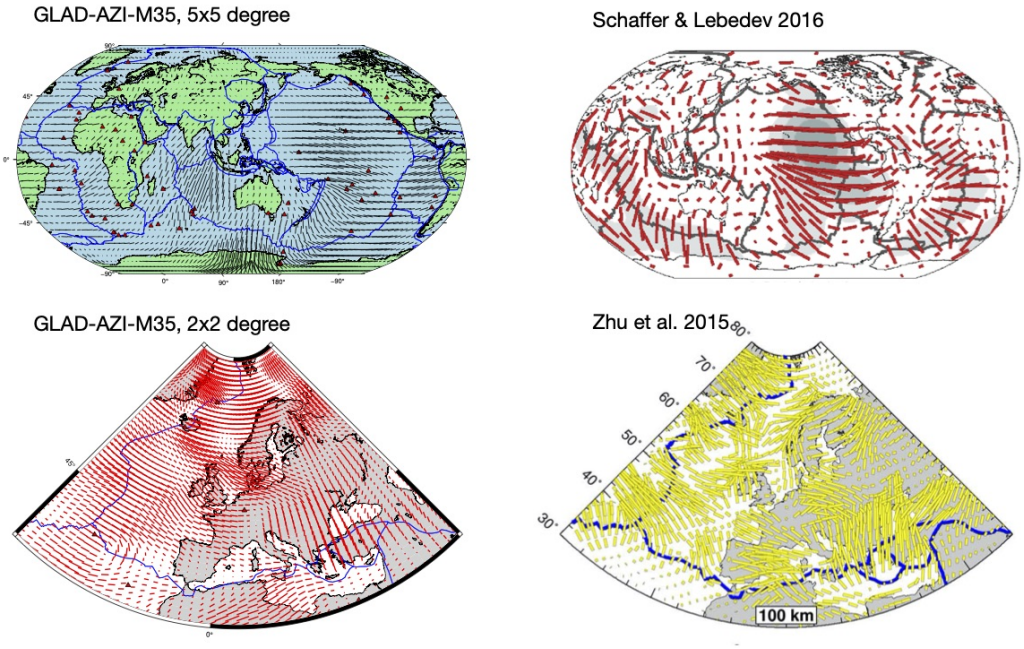

FIGURE 4.9: Comparisons of azimuthally anisotropic global adjoint results after 10 iterations (GLAD-M35-AZI) with recent azimuthally anisotropic models from classical surface-wave inversions by Schaeffer et al. (2016) at 50 km, 100 km and 200 km. SL2016svA and SL2016svAr are the same models by Schaeffer et al. (2016) but for the latter less damping was applied during its inversion. GLAD-M35-AZI sections are sampled at every 5×5 degrees to make them comparable to the resolution of global models.

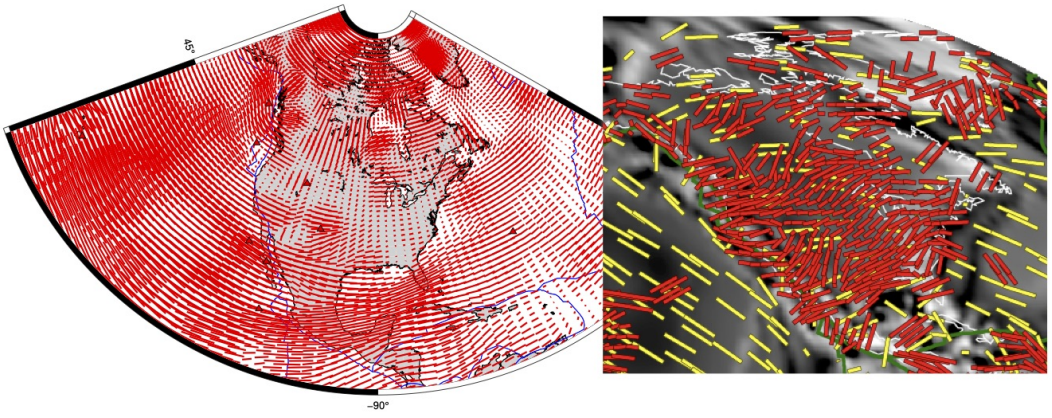

FIGURE 4.10: Comparison of GLAD-M35-AZI underneath North America at 50 km (left) with a compilation of SKS splitting measurement results (right, red dashes) from Thorsten Becker. Yellow dashes on the left figure are from geodynamical modeling.

4.4 Discussions

There is strong evidence from seismic data that Earth's upper mantle is anisotropic mainly due to the lithospheric deformation and mantle flow controlled by plate tectonics or mantle convection (e.g., Montagner and Tanimoto, 1991; Trampert and Woodhouse, 2003; Schaeffer et al., 2016). We demonstrated taking azimuthal anisotropy into account in global adjoint tomography based on 3D waveform simulations. To reduce the trade-off between seismic parameters in this research we only focus on surface wave anisotropy, and following previous azimuthally anisotropic tomography studies we considered only the 2θ terms and attempted to invert for normalized anisotropic parameters G_c' and G_s' (Montagner and Tanimoto, 1991; Schaeffer et al., 2016) along with the vertically and horizontally polarized shear wavespeeds (Zhu et al., 2015). Despite the great opportunity of taking full complexity of the wave propagation into account to further improve the resolution of tomographic results the effect of ignored 4θ azimuthal anisotropy terms needs to be investigated as particularly Love wave azimuthal anisotropy may be dominated by them (Montagner and Tanimoto, 1991; Trampert and Woodhouse, 2003; Ekström, 2011; Adam and Lebedev, 2012). Adam and Lebedev (2012) also reported that 4θ terms may also be important for shorter period Rayleigh waves because of their sensitivity to the continental crust.

After 10 iterations our results show good agreement with plate tectonics and overall global pattern reported in previous studies (Schaeffer et al., 2016). For instance, the azimuthal anisotropy in the Pacific plate correlates well with the plate motions where we also observe largest values of anisotropy within the plate. The northeast motion of the Nazca plate and subduction underneath South America, the mid-ocean ridges etc. are also well captured. The anisotropy in Africa shows significant spatial variations which can be related to the plumes and complex tectonics of the region as well as sparse data coverage that may hinder the resolution. Where we have dense seismic networks we are able to observe more details such as the circular pattern in anisotropy in the east coast of North America which may be related to the subduction related mantle flow (e.g., Zhu et al., 2020), or the counter-clockwise rotation of the Anatolian plate (e.g., Endrun et al., 2008).

Although our results after 10 iterations show consistent results with the previous global- and continental-scale azimuthally anisotropic models there are also significant differences. In fact, there are significant differences in details between all azimuthally anisotropic models (see comparisons in Schaeffer et al., 2016) similar to isotropic or transversely isotropic models which agree with each other at long scales. This is indeed one of the main motivations of this study aiming to improve the resolution by capturing full complexity of wave propagation. Furthermore, in our inversions crust and mantle are inverted simultaneously avoiding commonly used crustal corrections for which concerns have been reported that the accuracy of determining the crustal phase with first-order approximations and the chosen crustal model may have an impact specifically on the upper mantle structure and anisotropy (e.g., Bozdağ and Trampert, 2008; Ferreira et al., 2010; Lekić and Romanowicz, 2011). However, there are other prominent factors which may affect the resolution. Our 10-iteration results can capture the direction of the fast direction at larger scales which are, in general, consistent with the plate motions. However, the magnitude of azimuthal anisotropy is still weaker compared to what has been reported in previous studies. This is likely because we need to continue our iterations as the inversion has not converged yet. Although the misfit reduction has started to flatten after the 8th iteration (Figure 4.5) the total misfit reduction compared to

the misfit of the starting model is slightly less than 25%. This is likely pointing out that the bulk of the global misfit has decreased. However, the imperfect global data coverage is making it difficult to decrease the misfit further. On the other hand the smoothing we have been using to balance the gradient must be decreased (half width of the Gaussian operator is 120 km in the horizontal direction) to explain 40 s surface waves. Furthermore, consider applying a multi-scale smoothing by defining the width of the Gaussian smoothing operator as a function of the Hessian kernel which resembles the kernel coverage (Bozdağ et al., 2016).

Our current results show the maximum strength of azimuthal anisotropy around 100-250 km which start decreasing below 300 km. On the other hand the results do not show significant difference in the fast direction of anisotropy in the radial direction. The radial change in azimuthal anisotropy is mostly observed underneath densely covered regions such as North America and Europe. This may be due to: 1) the model has not converged yet, 2) we need to readjust the smoothing in the radial direction and perform more iterations, 3) we need better measurements to increase the depth resolution. Although in our measurements we adjust the measurement window lengths to capture both higher and fundamental mode minor- and major-arc Love and Rayleigh waves the multitaper traveltime misfit we have been using generally highlights the fundamental mode surface waves. To improve the depth resolution, in addition to decrease and readjust the smoothing, we should highlight the contribution of higher modes in our measurements. To this end, the exponentiated phase misfit (Yuan et al., 2020), which is a variant of the instantaneous phase measurements (Bozdağ et al., 2011) that better deals with the phase cycle skips, might be a better alternative to cross-correlation traveltimes. The exponentiated misfit has the potential to highlight higher modes better in measurement because of the intrinsic weighting introduced into associated adjoint sources that balances amplitude differences between different seismic phases. Alternatively, the contribution of higher modes in the total misfit may be increased by carefully selecting them and adjusting their weighting in the misfit function appropriately (i.e, van Heijst and Woodhouse, 1997). We first plan to implement the exponentiated phase into our adjoint tomography workflow which is straightforward. Meanwhile it is worthwhile to further investigate better measurement techniques for higher modes which is essential to extract optimum information from waveforms and improve the depth resolution of surface waves, specifically for the transition zone.

4.5 Conclusions

We performed 10 azimuthally anisotropic global adjoint inversions using a multitaper traveltime misfit of minor- and major-arc surface waves combined with double-difference multitaper traveltime measurements made on paired measurement windows. We observe that the double-difference measurements help increase the resolution underneath densely covered regions such as in the continental lithosphere underneath North America and Europe. Our results show that after 10 iterations we capture multi-scale anisotropic signal which correlate well with plate motions where continental-scale resolution is achieved underneath densely covered regions. The observed anisotropic strength is maximum around 1% which is an indication that the model has not converged yet and more iterations are needed. To further increase the depth resolution we consider reducing the smoothing and use a multi-scale smoothing strategy to smooth densely covered regions less and vice versa. Furthermore we will continue iterations using the exponentiated phase and its double-difference

version to increase the contribution of higher modes to better image specifically the transition zone.

Acknowledgements

SPECFEM3D_GLOBE simulations and 3D adjoint tomography workflow so far have been performed on Oak Ridge National Laboratory's (ORNL) Summit system. We gratefully acknowledge the additional computational resources provided by The Princeton Institute for Computational Science and Engineering (PICSciE) and Colorado School of Mines High Performance Computing Center. We thank Jeroen Tromp for discussions to set up anisotropic adjoint inversions and his help to access ORNL and Princeton computational resources. SPECFEM3D_GLOBE packages are freely available from Computational Infrastructure for Geodynamics (CIG) where we routinely commit our changes related to anisotropic simulations to the developers version.

Chapter 5

Anelastic full-waveform inversion in earthquake seismology[1]

Summary

Attenuation, a measure of the energy loss of seismic waves, is a crucial parameter to detect partial melt, water content, and thermal variations in the mantle, which is difficult to be constrained by seismic wavespeeds alone. Although there is almost agreement on the long wavelength elastic structure of the mantle at the global scale, there is no such consensus on the anelastic structure. First-generation global adjoint tomography models GLAD-M15 (Bozdağ et al., 2016), and its successor GLAD-M25 (Lei et al., 2020), are elastic models with transverse isotropy confined to the upper mantle. They were constructed using the phase information of waveforms only. Our ultimate goal is to construct an anelastic upper-mantle model by updating elastic and anelastic parameters simultaneously in adjoint inversions based on 3D wave simulations by assimilating both the phase and amplitude information in inversions which will lead to exact[2] full-waveform inversion (FWI). The trade-off between seismic parameters are well known in seismic tomography. To this end, we first perform 2D tests to investigate the trade-off between wavespeeds and Q (quality factor, inverse of attenuation) parameters to define a strategy for 3D global inversions. We also explore the effect of different measurements on anelastic inversions using cross-correlation traveltimes and amplitudes, instantaneous phase and envelopes, and classical waveform misfit functions, inverting elastic and anelastic parameters simultaneously and sequentially. Our 2D test results suggest that the envelope misfit performs best at earlier iterations by helping reduce the nonlinearity in FWI. In addition, we compared simultaneous inversions to the sequential ones, i.e., inverting elastic or anelastic parameters consecutively, where simultaneous inversions behave slightly better in our 2D experiments.

In the second part, we first analyzed the effect of different 1D Q models on seismic waveforms by performing forward simulations of various 1D and 3D elastic/anelastic models, which suggest the necessity of simultaneous elastic/anelastic inversions also to construct the elastic structure, as attenuation cause not only amplitude anomalies but also significant physical dispersion, particularly on surface waves. We then performed one simultaneous iteration of elastic and anelastic models using the starting models GLAD-M25 and 1D Q model QRF12, and the dataset of 253 earthquakes used in the construction of GLAD-M15. The preliminary results are promising, such as depicting the high and low attenuation in the West and East coasts of North America. We will initially focus on the upper-mantle attenuation

[1] The content of this chapter will be submitted for publication: **Örsvuran, R.** & Bozdağ, E., 2021. Strategies for anelastic adjoint waveform inversions, *in prep*.

[2] as opposed to only using phase information

where we have the better coverage compared to the lower mantle and consider proceeding global iterations based on the experience we gain from 2D tests. We also consider applying a correction to source parameters (scalar moment and origin time) rather than inverting them simultaneously with the elastic and anelastic models to reduce the trade-off between parameters. The global simulations have been performed on Oak Ridge National Laboratory's IBM AC922 Summit system.

5.1 Introduction

In an elastic medium, it is assumed that the seismic energy is conserved and attenuation is observed only due to spatial propagation. In an anelastic medium, on the other hand, seismic waves can lose their energy also due to the internal friction which is the source of the anelastic attenuation. This effect is generally represented by the dimensionless quantity "quality factor" Q (Aki and Richards, 2002), which is the inverse of attenuation (Q^{-1}). It is observed that anelastic attenuation can cause both physical dispersion and amplitude anomalies in seismic waveforms (Anderson, 1967; Liu et al., 1976). Therefore, it is crucial to have accurate anelasticity models while simulating the wave propagation.

Anelastic attenuation provides important information about the physical and chemical properties of the Earth's inner structure which enables the detection of partial melt, water content and thermal variations with the elastic wavespeed (Karato, 1993; Faul and Jackson, 2005).

Anelastic models generally have lower resolution than elastic models. This is partly due to challenges related to amplitude measurements. For the seismological problems, it is common to construct elastic velocities using the phase of seismic waveforms only since the seismic velocities are more linearly related to the phase information. On the other hand, the anelastic information is generally derived using amplitudes which are more prone to nonlinearity due to their more complex relation to elastic, anelastic and source parameters (Romanowicz, 1994a). Several studies constructed radially symmetric anelastic models which showed an order of magnitude variations of Q values, specifically in the upper mantle and crust, which has led the discussions that highlights the importance of lateral variations in the Earth's anelastic structure (Durek and Ekström, 1996; Resovsky et al., 2005).

There have been several studies which attempted to construct laterally heterogeneous attenuation models. Romanowicz (1994a) used minor- and major-arc Rayleigh waves to construct an upper mantle attenuation model which revealed upwellings under Hawaii and northeastern Africa. Bhattacharyya et al. (1996) used S-SS differential waveforms to constrain a shear-wave attenuation model and showed some correlation with models using surface wave analysis. Reid et al. (2001) also used differential S waveforms and combined them with surface-wave measurements to enhance the depth resolution, while Warren and Shearer (2002) used P-PP differential amplitudes. Selby and Woodhouse (2002) and Dalton et al. (2008) also used Rayleigh wave amplitudes to image the lateral variations in the anelastic mantle. Gung and Romanowicz (2004) used non-linear asymptotic coupling theory for anelastic inversions using surface-wave amplitudes by aligning them in phase, while Karaoğlu and Romanowicz (2018b) used envelope measurements without the need for phase alignment. The model of Karaoğlu and Romanowicz (2018b) is the most advanced model in terms of imaging techniques by simultaneously inverting shear wavespeeds and Q using the waveforms from spectral-element simulations coupled with 2D

asymptotic kernels. There are also studies that use ambient noise to construct attenuation models (Prieto G. A. et al., 2009) which can be advantages to improve the path coverage. These studies show that the attenuation variation can be correlated with structures due to plate mechanics such as hotspots and ridges. However, there is still no consensus on neither 1D nor 3D attenuation models of the Earth's mantle (Resovsky et al., 2005; Karaoğlu and Romanowicz, 2018a).

Since the fluid content may cause attenuation, anelasticity is also active research field in exploration geophysics (Chapman et al., 2006; Kamei and Pratt, 2013; Kurzmann et al., 2013). It has been shown that attenuation is an important parameter for monitoring gas storage and it can also provide constraints on determining soft sediment structures (Carcione et al., 2006; Operto and Miniussi, 2018). For numerical simulations, frequency-domain methods are generally used (Carcione et al., 1988; Robertsson et al., 1994; Keating and Innanen, 2019) while there are also examples with time-domain solvers (Bai et al., 2014; Yang et al., 2016; Fabien-Ouellet et al., 2017).

With the advancement of computational resources, full-waveform inversion was first demonstrated at the European scale in earthquake seismology by Zhu et al. (2013). Full waveform inversion method has an advantage of using full physics of the wave propagation which eliminates the need for doing corrections for elastic focusing effects, crustal structure etc. There are also discussions in the literature on the use of full-waveform inversion in anelastic tomography (Cobden et al., 2018). One of the important milestones towards anelastic full-waveform inversion was the implementation of attenuation in adjoint simulations based on check-pointing that uses parsimonious storage to compute exact anelastic sensitivity kernels by Komatitsch et al. (2016) which is specifically crucial in anelastic and global-scale elastic/anelastic adjoint inversions.

As mentioned above, one of the main challenges to invert for the anelastic structure is coming from the measurement complexities. Both amplitudes and phases are sensitive to multiple parameters (i.e., elastic parameters, anelasticity, source parameters, internal topographies, etc.) where amplitude measurements are more nonlinearly related to them. In the ideal case, all these parameters should be inverted simultaneously, as, for instance, demonstrated by Valentine and Woodhouse (2010b) for the simultaneous source and structural inversions. However, the trade-off between them makes it impossible to invert them uniquely (i.e., Beller and Chevrot, 2020). To mitigate the trade-off between source and structural parameters source corrections are also considered to be applied such as to reduce the bias in origin times and scalar moments (Zhu et al., 2015; Lei et al., 2020). On the other hand, most of the 3D anelastic studies are performed by fixing a chosen elastic model to take focusing and defocusing effects into account (e.g, Dalton et al., 2008). However, it is clear that elastic and anelastic models should be updated iteratively to avoid or minimize the bias in both models. This is one of the main motivations of this research to investigate the trade-off between seismic wavespeeds and attenuation. In addition, amplitudes may be more prone to the potential instrument response problems as well as noise that the data selection for inversions requires more attention. Automated window selection algorithms are helpful to some extent (e.g., Maggi et al., 2009; Lee and Chen, 2013). Recently, machine learning algorithms have started to gain interest for this purpose, as well (e.g., Chen et al., 2017).

In this study, we explore inversion strategies for anelastic full-waveform inversion. Our ultimate goal is to construct an anelastic upper-mantle model by updating elastic and anelastic parameters simultaneously in adjoint inversions based on 3D wave simulations by assimilating both the phase and amplitude information

in inversions which will lead to exact full-waveform inversion (FWI). To this end, first, using 2D spectral-element simulations, we test various misfit functions such as, cross-correlation traveltimes and amplitudes, instantaneous phase and envelope misfits, and the L2-norm waveform misfit. In addition, we investigate simultaneous and sequential elastic and anelastic inversions to see which strategy is favored by the full-waveform inversion. In the second part, we present our preliminary results of the first iteration simultaneous elastic and anelastic full-waveform inversion at the global scale based on observed data following our observations based on global-scale forward modeling with 1D/3D elastic and anelastic models.

5.2 Methodology

Seismic waveforms are affected by various factors, such as path and source effects as well as the instrument response of receivers. If $A(\omega)$ and $\phi(\omega)$ represent the amplitude and phase of a seismic waveform as a function of the angular frequency ω in the frequency domain, they may be written in terms of these factors (Dalton et al., 2008),

$$\phi(\omega) = \phi_S(\omega) + \phi_I(\omega) + \phi_F(\omega) + \phi_Q(\omega)\,, \tag{5.1}$$

$$A(\omega) = A_s(\omega)A_I(\omega)A_F(\omega)A_Q(\omega) \tag{5.2}$$

where the subscripts S, I, F, Q denote the source, instrument response, elastic focusing effects and anelastic attenuation, respectively. It is common to use phase information only to construct seismic tomographic models to linearize the inverse problem focusing only on the elastic structure (Luo and Schuster, 1991), where a radially symmetric Q model is used in the forward part of the problem. In many exploration or smaller-scale regional earthquake seismology studies, attenuation may also be neglected in simulations. However, at the continental to global scale tomographic studies, the effect of low Q layer at 200 km cannot be excluded. Therefore, generally all elastic models are constructed based on a chosen radially 1D Q model. Nevertheless, there is no consensus on radially symmetric Q models that highlights the importance of constructing 3D Q models which not only affects the amplitude but also the phase of waveforms as can also be seen equations 5.1 & 5.2. This is also the reason why in classical attenuation only tomographic inversions, it is assumed that the elastic focusing effects are correctly predicted.

Quality factor Q, which is the inverse of attenuation, depends on the physical and chemical conditions of the medium and it can modeled using (Anderson and Archambeau, 1964):

$$Q(\omega, T) = A \exp\left(\frac{\alpha(E^* + PV^*)}{RT}\right)\,, \tag{5.3}$$

where T is the temperature, A is a constant related to the properties of the medium, E^* is the activation energy, V^* is the activation volume of anelastic relaxation, P is the pressure, R is the gas constant. Thus, attenuation is sensitive to temperature and pressure of the medium. Velocity is also dependent on these parameters and the relation between the velocity and attenuation can be written as:

$$V(\omega) = V(\omega_0)[1 + (1/\pi)Q^{-1}\ln(|\omega|/\omega_0) - isgn(\omega)Q^{-1}]\,, \tag{5.4}$$

where ω_0 is a chosen reference frequency. Hence, the physical nature of the attenuation and seismic wave velocity also points out that there is a trade-off between these parameters.

In this study, our goal is to ultimately construct a global upper-mantle Q model without any approximation or correction to the wave propagation using full-waveform inversion. We aim to invert elastic and anelastic model parameters simultaneously. Assuming that the instrument response information we get from data repositories is known, we need to address the source uncertainties which directly affect both amplitudes (i.e., radiation pattern) and phase (i.e., depth and origin time uncertainties). Ideally, source and structural parameters should also be inverted simultaneously (Valentine and Woodhouse, 2010b). However, due to potential trade-offs between seismic parameters, to simplify and make the problem more tractable we prefer to apply source corrections to waveforms by determining the origin time and scalar moment of each earthquakes through grid search (e.g., Zhu et al., 2015; Lei et al., 2020). In addition, we use more strict selection criteria to determine measurement windows with the automated window selection algorithm, FLEXWIN (Maggi et al., 2009), to use high-quality data from database providers.

We describe how we compute elastic and anelastic kernels using the adjoint method in the following section. All 2D and 3D forward and adjoint simulations are performed with the 2D and 3D numerical wave propagation solvers SPECFEM2D (Komatitsch and Vilotte, 1998) and SPECFEM3D_GLOBE (Komatitsch and Tromp, 2002b) packages, respectively.

5.2.1 Anelastic adjoint kernels

To further simplify the inverse problem we ignore the bulk attenuation as the anelastic attenuation is dominated by the shear processes (Romanowicz, 1990). Assuming that the shear Q (Q_μ) is constant over the seismic-frequency band, the frequency-dependent shear modulus $\mu(\omega)$ may be written as (Liu et al., 1976),

$$\mu(\omega) = \mu(\omega_0)[1 + (2/\pi)Q_\mu^{-1}\ln(|\omega|/\omega_0) - isgn(\omega)Q_\mu^{-1}]\,, \tag{5.5}$$

where ω is the angular frequency, ω_0 is a chosen reference angular frequency, Q_μ is the quality factor for the shear modulus. The gradient of the shear modulus may be written as (Tromp et al., 2005),

$$\delta\mu(\omega) = \mu(\omega_0)[(2/\pi)\ln(|\omega|/\omega_0) - isgn(\omega)]\delta Q_\mu^{-1}\,, \tag{5.6}$$

where the $(2/\pi)\ln(|\omega|/\omega_0)$ captures the physical dispersion and $isgn(\omega)$ controls the amplitude term.

In the adjoint method, the Fréchet kernels are computed numerically through the interaction of forward and adjoint wavefields (Tarantola, 1984; Tromp et al., 2005). Forward wavefieds are computed based on a regular seismic source recorded by a station at a specific point. Adjoint wavefields are initiated by a fictitious source, called *adjoint source*, based on the misfit between observed and synthetic data inserted at the location of the receiver where the wavefield simulations are performed backward in time from receiver to the regular source.

Let $f_i^\dagger(\mathbf{x}, t)$ be the elastic adjoint source in the time domain for the ith component at position $\mathbf{x}$. Using the Born approximation in the frequency domain, which is accurate for the Rayleigh wave scattering (Wu and Aki, 1985), the adjoint wavefield may be rewritten by inserting the frequency dependent shear modulus $\delta\mu(\omega)$ perturbations,

$$s_k^\dagger(\mathbf{x}',t') = \int_0^{t'} \int_V G_{ki}(\mathbf{x}',\mathbf{x_r};t'-t) \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} [(2/\pi)\ln(|\omega|/\omega_0) - isgn(\omega)]^* f_i^\dagger(\mathbf{x},\omega) \exp(i\omega t) d\omega \right) \mathrm{d}^3\mathbf{x}\mathrm{d}t\,, \quad (5.7)$$

where G_{ki} is the Green's function. The anelastic adjoint source may then be defined as (Tromp et al., 2005; Bozdağ et al., 2011)

$$\tilde{f}_i^\dagger(\mathbf{x},t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [(2/\pi)\ln(|\omega|/\omega_0) - isgn(\omega)]^* f_i^\dagger(\mathbf{x},\omega) \exp(i\omega t) d\omega \quad (5.8)$$

which shows that one can compute anelastic kernels just by defining the anelastic adjoint source based on the Hilbert transform of the elastic adjoint source without modifying the numerical solver. The disadvantage is that this approach doubles the simulations and the computation time as we need to perform two adjoint simulations to compute elastic and anelastic gradients for their simultaneous inversions. The gradient in terms of Q perturbations can be written as

$$\delta\chi = \int_V K_\mu(\mathbf{x}) \delta Q_\mu^{-1}(\mathbf{x}) d\mathbf{x}\,. \quad (5.9)$$

In Figure 5.1, sample waveforms are presented to illustrate the elastic and anelastic effects on the waveforms. 2D spectral-element simulations were computed using the SPECFEM2D package (Komatitsch and Vilotte, 1998) for homogeneous background models (V_s = 3.5 km/s, Q_μ = 10) recorded by a receiver at 20 km distance. First row shows, two seismograms computed for the background model (black) and 3% perturbed V_s model (V_s = 3.4 km/s) (red) with the same attenuation model. The source is defined as an elastic force and its source time function is a Ricker wavelet with 0.5 Hz dominant frequency. The wavespeed affects traveltimes, as expected, producing a time delay between the fast and slow model waveforms while the amplitude change is small. The second row shows the difference between waveforms computed for different Q_μ models with a low (Q_μ = 20) and high (Q_μ = 15) attenuation models. The amplitudes are larger for the low attenuation model. Due to physical dispersion, phase shifts can be observed where high attenuation slows down the seismic phase. It should be noted, attenuation perturbations need to be much higher compared to wavespeed perturbations to produce measurable waveform difference. Since distances are generally much longer in global applications, smaller attenuation differences can produce significant effects.

5.3 2D elastic-anelastic inversions

In this section, we explore inversion strategies for the joint inversion of elastic and anelastic parameters using 2D test models. We test the effect of various misfit functions (i.e., waveform misfit, cross-correlation traveltime and amplitude measurements, exponentiated phase and envelopes) and compare the inversion results obtained from each misfit function. In addition, we compare the results from simultaneous and sequential inversions of elastic and anelastic parameters.

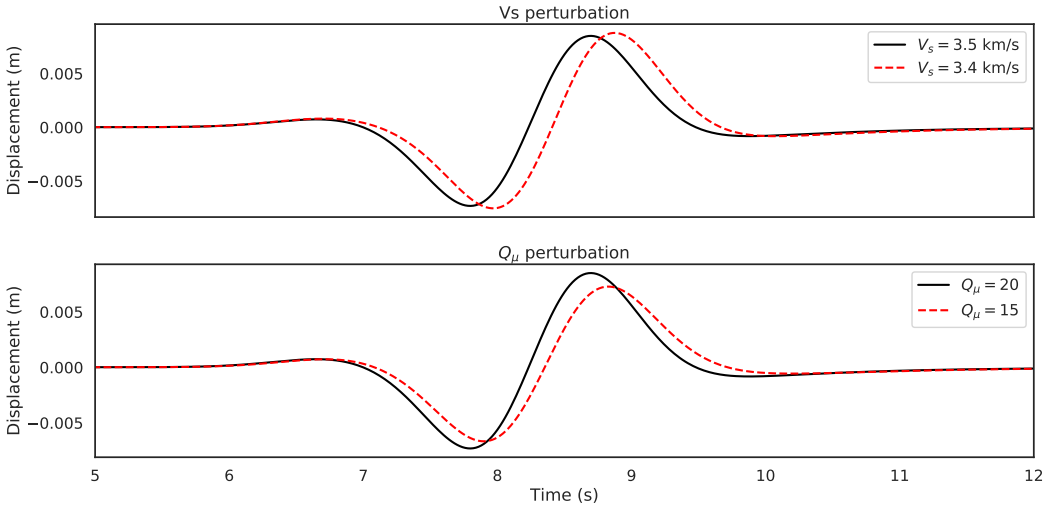

FIGURE 5.1: Synthetic waveforms computed for different homogeneous background models. First row shows two waveforms for fixed anelastic model (Q_μ = 20) and different Vs models. Second rows shows two waveforms for fixed wave velocity model (V_s = 3.5 km/s) and varying attenuation models. Velocity changes effect the traveltime as expected and produce slight amplitude change. Changing attenuation model mainly effects amplitude where high Q (low attenuation) waveform has bigger amplitude but also effects the traveltime due to physical dispersion.

5.3.1 Numerical simulations

We compute synthetic seismograms and adjoint kernels with SPECFEM2D (Komatitsch and Vilotte, 1998). SPECFEM2D is a 2D spectral-element solver for acoustic, elastic and anelastic seismic wave propagation (Komatitsch and Vilotte, 1998). In this section, viscoelastic models are used in the P-SV mode with the free surface boundary at the top of the mesh. The Perfectly Matched Layer (PML) absorbing boundaries are set to the other three sides of the model (Komatitsch and Tromp, 2003). Elastic forces are used as seismic sources. The first derivative of the Gaussian is used as the source time function with the dominant source frequency set to 0.1 Hz. The physical dispersion due to attenuation is defined around the same frequency (i.e., the reference angular frequency ω_0 is set to 0.628 radian Hz). The meshes for simulations are designed to resolve a highest frequency of 0.2 Hz. Measurement windows are selected to include the most of the direct, reflected and refracted waves. Misfits are computed after low-pass filtering of seismic data considering the highest resolvable frequency in simulations. For simultaneous inversions, the elastic and anelastic gradients are computed based on two sets of simulations as described in the previous section. For the sequential inversions, we start iterations by updating the elastic model while fixing the anelastic one (odd iteration numbers) followed by an anelastic model update while the elastic model is fixed (even iteration numbers). The elastic and anelastic gradients, which are the summation of event kernels, are smoothed, and then preconditioned using the pseudo-Hessian kernels. The L-BFGS (Nocedal, 1980) optimization technique is used to update models while keeping track of the last 5 iterations (memory value). Taking the advantage of synthetic tests we compute model misfits after each iteration in addition to data misfits. The iterations stopped when the misfit could not be reduced by more than 0.1%.

5.3.2 Misfit functions

One of the goals of this study is to experiment the effect of the chosen misfit function on anelastic inversions. Our ultimate goal is to choose the best measurements for elastic/anelastic model updates at the global scale where any speed up in convergence is valuable because of the cost of global-scale full-waveform inversion as well as the robustness of iterations.

The L_2 norm waveform misfit (WF) is the difference between observed and synthetic waveforms, such that,

$$\chi_{wf} = \frac{1}{2}\sum_{r=1}^{N}\int_0^T ||d_r(\mathbf{x},t) - s_r(\mathbf{x},t,\mathbf{m})||^2 \mathrm{d}t \,, \tag{5.10}$$

where d and s are the observed and synthetic waveforms, respectively, recorded at station r as a function of position $\mathbf{x}$ and time t. $\mathbf{m}$ denotes model parameters. The waveform misfit is commonly used in exploration geophysics (Virieux and Operto, 2009) and some global tomography studies (Li and Romanowicz, 1996; French and Romanowicz, 2014) It may behave nonlinearly for complicated waveforms (i.e. surface waves).

Traveltime measurements are widely used alternative to the waveform misfit in seismology. Cross-correlation traveltime (TT) measurements can be used to determine the phase shift between waveforms where they can be defined as,

$$\chi^{TT} = \frac{1}{2}\sum_{r=1}^{N} ||T_r^{obs} - T_r(\mathbf{m})||^2 \mathrm{d}t \,, \tag{5.11}$$

where the $T_r(\mathbf{m})$ and T_r^{obs} are the arrival times for synthetic and observed waveforms, respectively. Since cross-correlation cannot distinguish between different arrivals, it is mainly used for simpler waveforms and separate measurement are made for distinct arrivals. Frequency-dependent traveltime measurements like the multitaper misfit are better fit for complex waveforms such as dispersive surface waves. (Laske and Masters, 1996).

The cross-correlation amplitude misfit (AMP) is a common measurement for amplitude anomalies. It is common to use the logarithmic ratio between the amplitudes of observed (A_r^{obsd}) and synthetic ($A_r(\mathbf{m})$)waveforms (Dahlen and Baig, 2002), which may be given by

$$\chi^{AMP} = \frac{1}{2}\sum_{r=1}^{N}\left[\ln\left(\frac{A_r^{obsd}}{A_r(\mathbf{m})}\right)\right]^2 \mathrm{d}t \,. \tag{5.12}$$

Traveltime and amplitude (TT+AMP) measurements can be combined to recover both phase and amplitude information which was previously used at the continental scale anelastic full-waveform inversion (Zhu et al., 2015). This is also one of our main strategies to check in this study.

The envelope misfit (ENV) (Bozdağ et al., 2011) uses the difference between the envelopes of observed and synthetic waveforms, which may be defined as their logarithmic ratios,

$$\chi^{ENV} = \frac{1}{2}\sum_{r=1}^{N}\int_0^T \ln\left(\frac{E_r^{obs}(t)}{E_r(\mathbf{m},t)}\right)\mathrm{d}t \,, \tag{5.13}$$

where envelope $E_r^{obs}(t)$ and $E_r(\mathbf{m},t)$ denote the envelopes of observed and synthetic waveforms computed by the Hilbert transform which turns a real signal into a

complex-valued analytical signal. Envelopes are the amplitude of the analytical signal which can be computed by the square root of the square of real and imaginary parts.

The Instantaneous phase (IP) (Bozdağ et al., 2011) can also be computed using the Hilbert transform similar to the envelope misfit, and can be applied to wavetrains. It is defined as the difference between the instantaneous phase of observed ($\phi_r^{obs}(t)$) and synthetic ($\phi_r(\mathbf{m}, t)$) waveforms,

$$\chi^{IP} = \frac{1}{2}\sum_{r=1}^{N}\int_0^T \left[\phi_r^{obs}(t) - \phi_r(\mathbf{m}, t)\right]^2 \mathrm{d}t\,. \tag{5.14}$$

The instantaneous phase misfit (IP) has been shown to recover diffracted waves better than cross-correlation traveltime measurements (Rickers et al., 2012). Like other phase measurements, the instantaneous misfit may suffer from the cycle-skips where the exponentiated phase misfit (Yuan et al., 2016) may help alleviate this problem. The EP misfit is defined as the difference between the exponential terms of observed and synthetic phases rather than directly computing the instantaneous phase shift,

$$\chi^{EP} = \frac{1}{2}\sum_{i}\int_0^T ||e^{i\phi_r^{obs}(\mathbf{x},t)} - e^{i\phi_r(\mathbf{x},t,\mathbf{m})}||^2 \mathrm{d}t\,. \tag{5.15}$$

Similar to TT+AMP measurements, one can also use the combination of EP+ENV measurements to take both the phase and amplitude into account in inversions. A similar separation of phase and amplitude measurements was proposed by Fichtner et al. (2008) in the time-frequency domain.

The adjoint method can be used to compute Fréchet derivatives of a chosen misfit function with respect to the model parameters via the interaction between the forward wavefield, due to a regular seismic source, and a secondary time-reversed adjoint wavefield, initiated by adjoint sources based on data residuals at the location of receivers. Using these misfit functions and their combinations, we present our 2D synthetic tests below. We test their effectiveness in anelastic inversion context.

5.3.3 Case study I: A layered model

In this 2D example, we adapted the upper-mantle of the radially symmetric elastic and Q PREM model (Dziewonski and Anderson, 1981) to our experiment, which is used as the initial model. We fixed the elastic part of the model during the simulations and decreased the Q_μ values by 50% (Figure 5.2) in the target model. In both the initial and final models there is a low velocity and low Q layer in the middle. Therefore, we only invert for the anelastic structure as we fix the elastic model avoiding the trade-off between Vs and Q_μ parameters. In our setup, we use 9 sources located at the edges and in the center, and the stations are evenly distributed throughout the medium as shown in Figure 5.2. It is an idealistic setup which is not possible in the real world. It is constructed to investigate the performance differences between various measurement types. The measurement windows are selected to include complete waveforms due to the simplistic nature of this example. We then performed iterations with six different measurements (i.e., cross-correlation traveltimes (TT), cross-correlation traveltimes and amplitudes (TT+AMP), exponentiated phase (EP), envelope (ENV), exponentiated phase and envelope (EP+ENV), waveform (WF) misfits) to observe the effect of the choice of misfit functions on anelastic inversions.

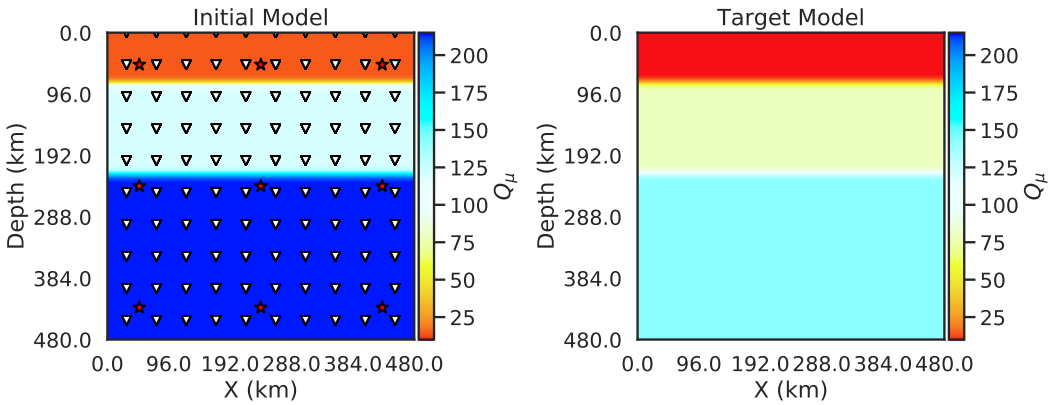

FIGURE 5.2: Initial and target Q_μ models. Starting model is modeled after PREM and target model is the 50% perturbation of the initial model. Elastic model is fixed. Nine sources (red stars) and 100 stations (white downward triangles) are superimposed on the initial model.

We present our inversion results in Figure 5.3 and their data misfits in Figure 5.4. From the visual inspection of final models iterated by the misfit functions mentioned above it is observed that all misfits resolve the first layer reasonably well taking the advantage of phases reflected from the free surface. ENV and EP+ENV seem to have the best convergence in all layers followed by WF which is slightly worse in the middle and the third layers. Indeed, the major discrepancy in the resolving power of all misfits become more visible in the third layer where the illumination is more sparse.

Figure 5.5 shows the model misfit reductions for each inverted model where we observe that EP+ENV gives the best convergence rate by decreasing both phase and amplitude misfits simultaneously. The worst result is obtained for TT measurements which is understandable as cross-correlation TT measurements highlight the time-shift information of the maximum amplitude signals. As both the phase and amplitudes are sensitive to the variations in anelastic structure the results suggest inverting them simultaneously. ENV gives better results compared to WF likely because of the intrinsic weighting term in its adjoint source which better balances seismic phases with different amplitudes in a wavetrain. On the other hand, WF tends to be dominated by maximum-amplitude signals if windowing is not applied. However, ENV may have some limitations as it tries to fit the envelope of signals which are not sensitive to polarity of waveforms (Yuan et al., 2020). WF can extract more information from both phase and amplitudes compared to ENV. In Figure 5.5, the model-misfit reduction for ENV after 40 iterations is shown where additionally the iterations proceeded after 30th iteration with the WF misfit. Using WF after 30th iterations speeds up the misfit reduction compared to ENV showing the higher sensitivity of WF to waveform variations than ENV.

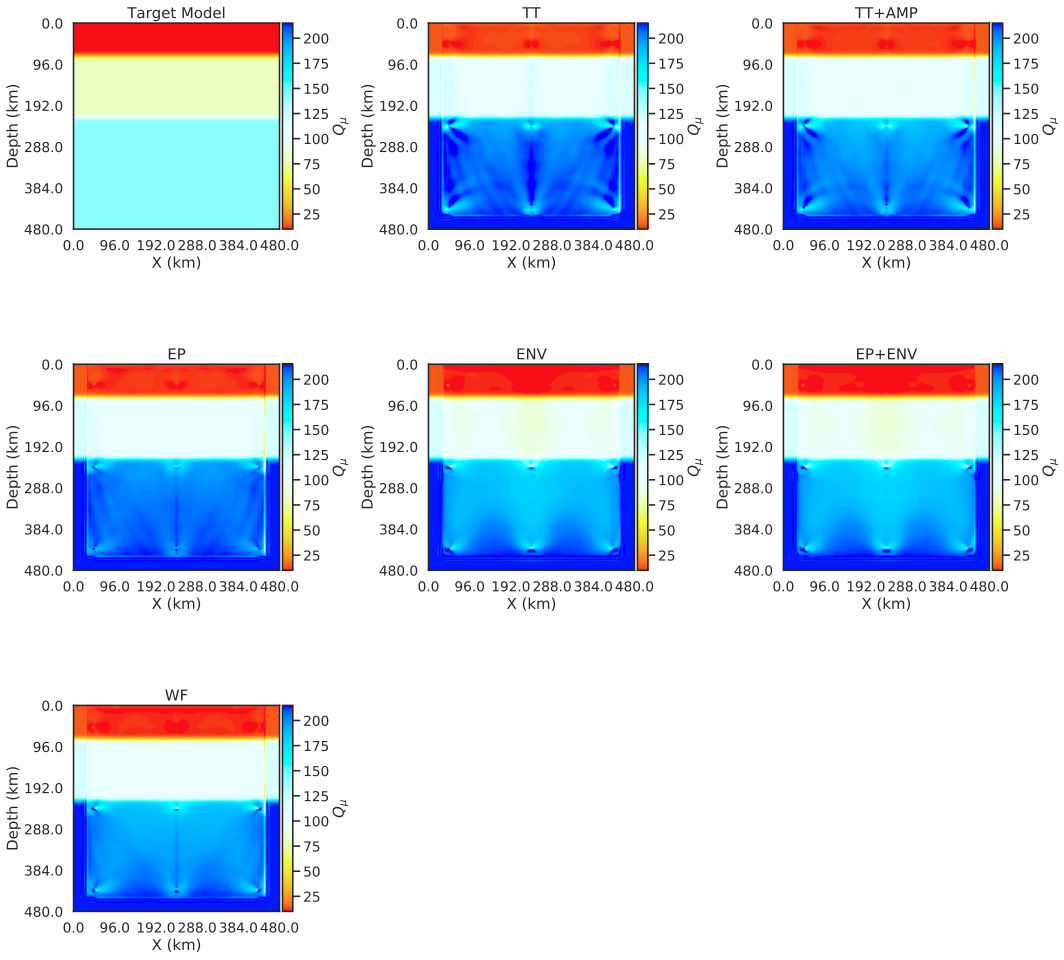

FIGURE 5.3: Inversion results for CC traveltime (TT), CC traveltime and amplitude (TT+AMP), exponentiated phase (EP), envelope (ENV), combined exponentiated phase and envelope (EP+ENV) and waveform (WF) misfits. Including both phase and amplitude gives better results.

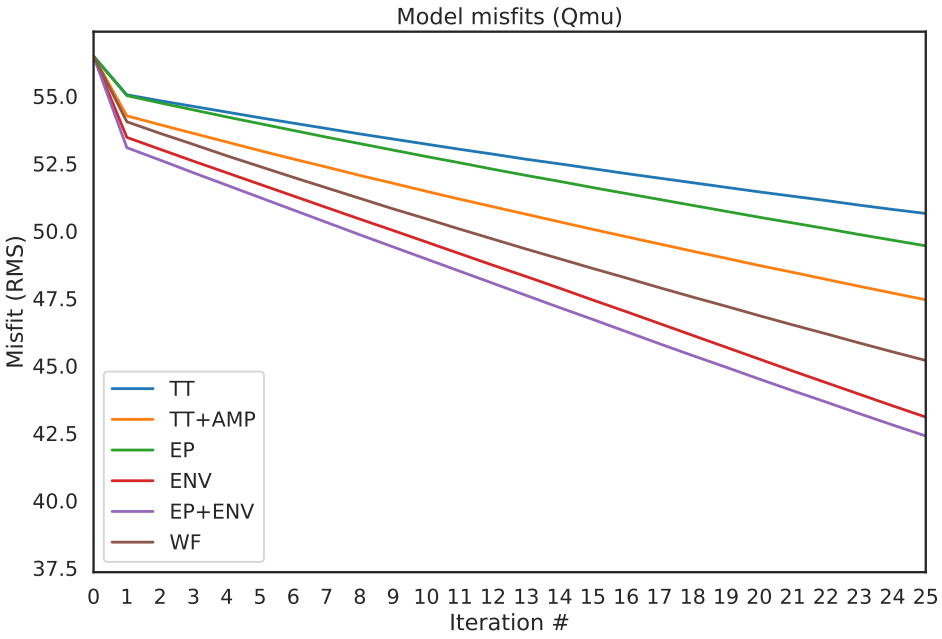

FIGURE 5.4: Q_μ model misfits for the simple layered model for the misfits: CC traveltime (TT), CC traveltime and amplitude (TT+AMP), exponentiated phase (EP), envelope (ENV), combined exponentiated phase and envelope (EP+ENV) and waveform (WF)

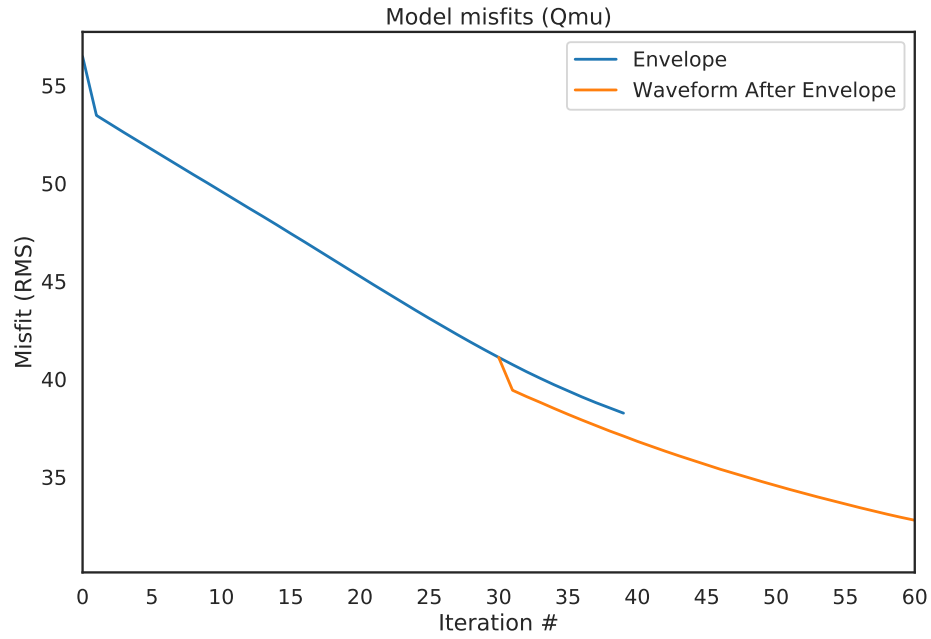

FIGURE 5.5: Q_μ model misfits for the simple layered model for the envelope misfit and waveform misfit which is started from the 30 iteration on. Waveform misfit decreases further than envelope misfit if we use a better starting model.

5.3.4 Case study II: Joint Inversion of elastic and anelastic models

In this section, we jointly invert elastic and anelastic parameters either simultaneously or sequentially (i.e., fixing either elastic or anelastic model at each iteration). We again use a three-layer model simplified from the elastic and anelastic PREM models with a low-velocity and low-Q layer in the middle, respectively (Figure 5.6). In the target models, S wavespeeds and Q_μ values are perturbed by 1% and 10% , respectively, around the corresponding three-layer PREM models. In both elastic and anelastic models the layer thicknesses are assumed to be known and fixed throughout the iterations where we invert S wavespeeds and Q_μ only.

Similar to the previous section, we again investigate the effect of different misfit functions on the final models by performing elastic and anelastic inversions simultaneously or sequentially. In the sequential inversion strategy, one can alternate the models at a chosen frequency of iterations. Here, we alternate the elastic and anelastic models at each iteration.

Figure 5.8 shows the data misfit reduction for both simultaneous and sequential inversions using the ENV misfit. Both strategies are able to decrease the misfit consistently although simultaneous inversions perform slightly better. More jagged misfit reduction for the sequential inversions is the result of the alternated updates of S-wavespeed and Q_μ models However, when we compare the model misfit reductions for Q_μ the sequential inversion strategy clearly outperforms the simultaneous counterpart (Figure 5.9) where the iterations start with the elastic model update. The first model update of the S wavespeeds seems to help decrease the Q_μ model misfit in the second iteration sharply. However, the model misfit reduction of S-wavespeed model does not show the same result where simultaneous inversions converge faster (Figure 5.10) It looks like it takes about 5 iterations for simultaneous iterations to settle down with the elastic model updates whereas the sequential strategy flattens from 5th to 12th iterations followed by a decrease in the model misfit. This non-monotonic decrease in model misfits is likely due to the trade-off between model parameters.

Figure 5.11 shows the inversion results after 19 iterations. S-wave models look very similar to each other but the model misfit suggest that simultaneous inversion results are slightly better. On the other hand, sequential inversion for the Q_μ model shows much better convergence than the simultaneous inversion. The close similarity between seismic waveforms computed for initial and target models (Figure 5.12) show that data misfits are saturated which may also be responsible for the oscillated model-misfit reductions.

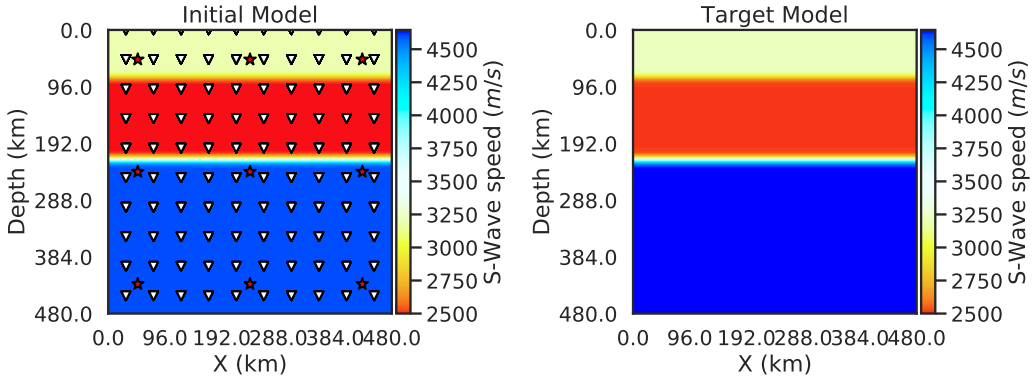

FIGURE 5.6: Initial and target S-wave velocity models. Three layered models with low-velocity layer in the middle. Target model is 1% perturbed version of the initial model. Nine sources (red stars) and 100 stations (white downward triangles) are superimposed on the initial model.

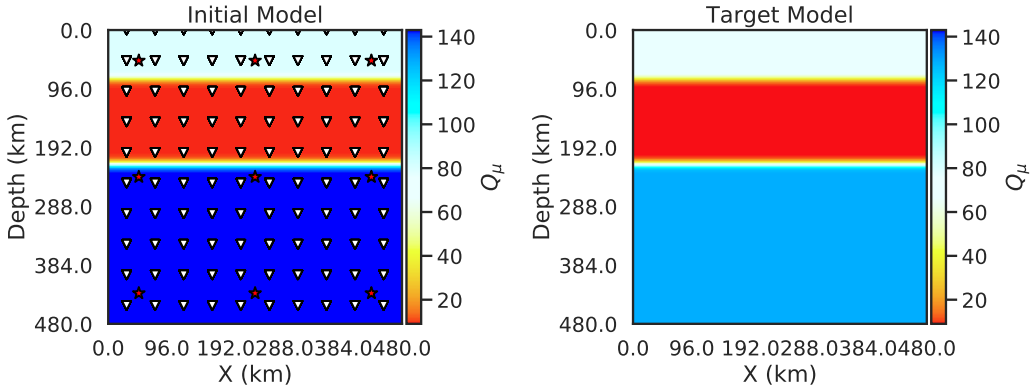

FIGURE 5.7: Initial and target Q_μ models. Target model is 10% perturbed version of the initial model. Nine sources (red stars) and 100 stations (white downward triangles) are superimposed on the initial model.

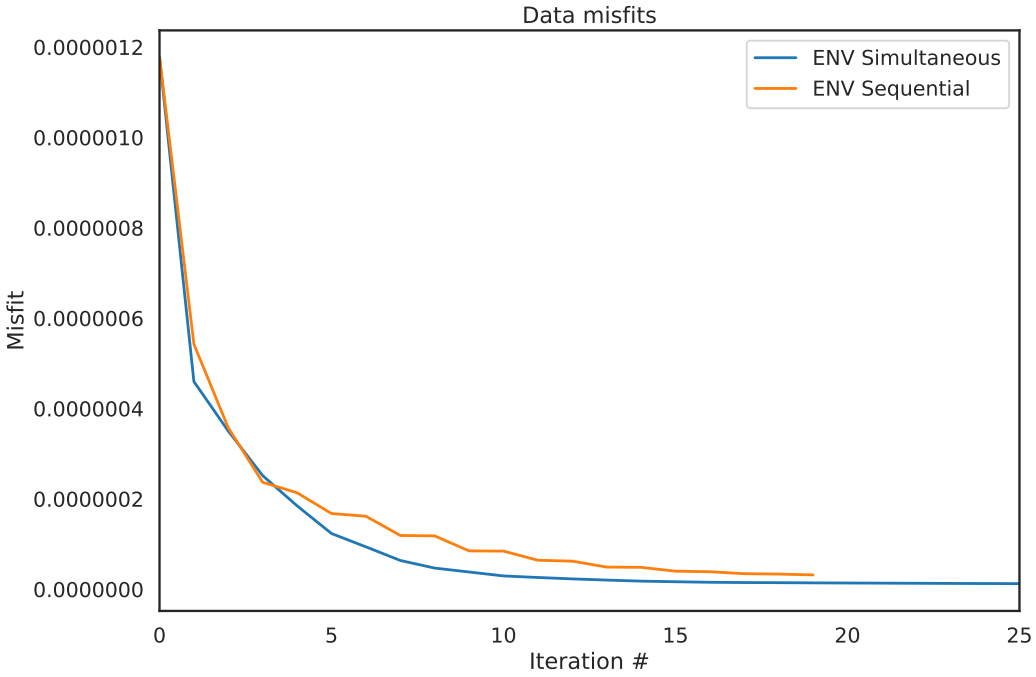

FIGURE 5.8: Data Misfits for the different inversion strategies. Sequential inversion method was not able to decrease the misfit after 19 iterations.

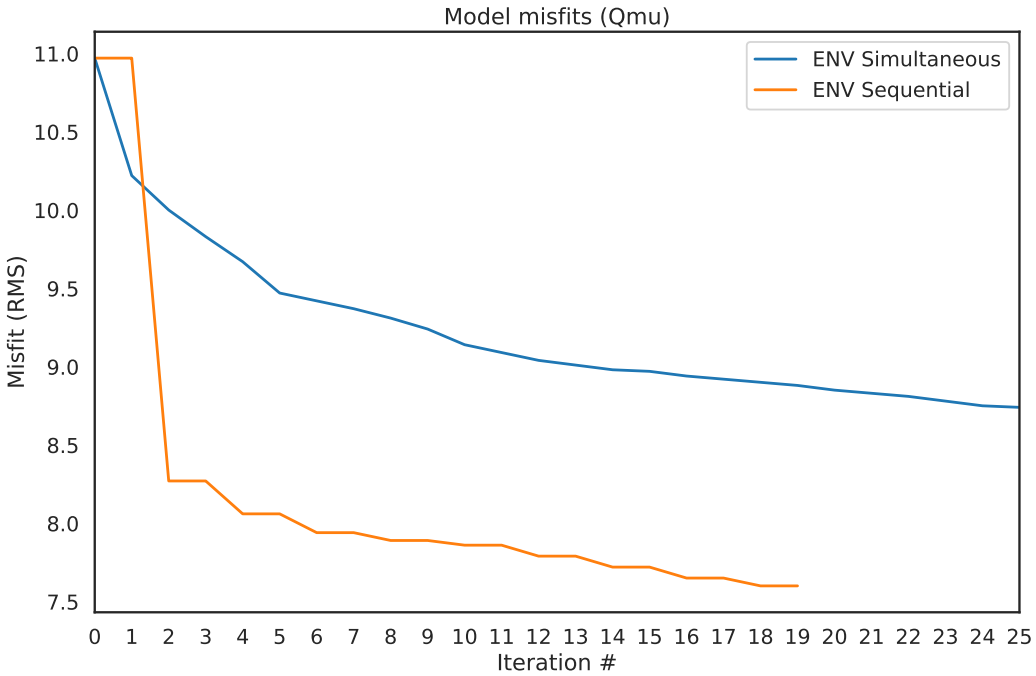

FIGURE 5.9: Q_μ model misfits for sequential and simultaneous inversion. simultaneous inversion reduces the anelastic model misfit more slowly than sequential version. Sequential version does not update the anelastic model on odd iterations which can be seen as the constant model misfits.

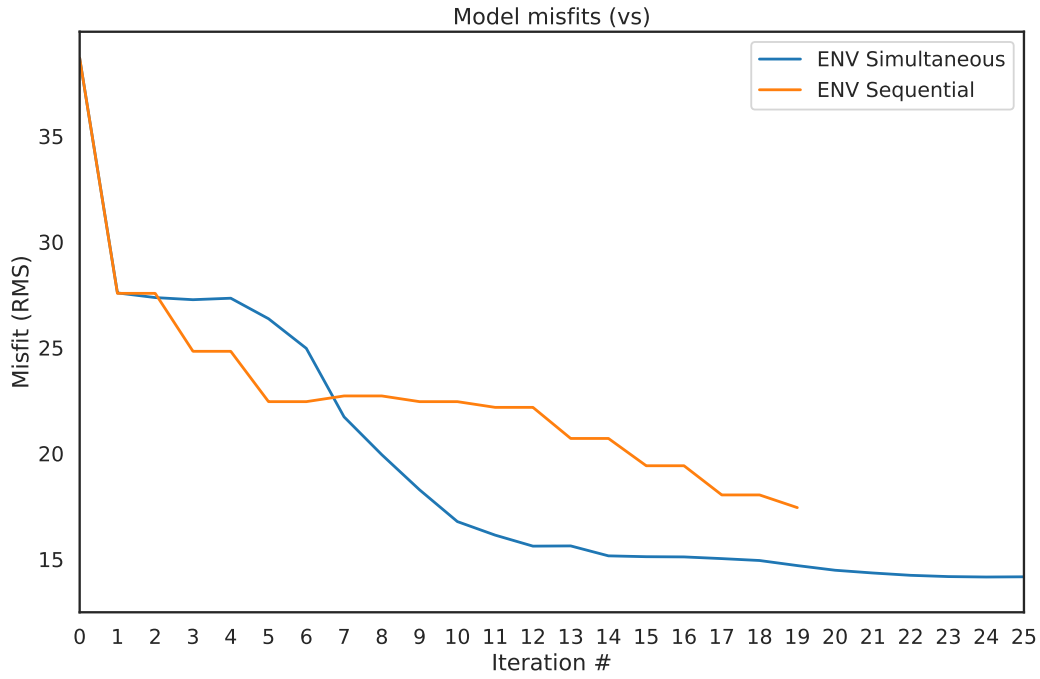

FIGURE 5.10: S-wave model misfits for sequential and simultaneous inversion. Simultaneous inversion was able to decrease the elastic model misfit slightly better for this example. Sequential version does not update the elastic model on even iterations which can be seen as the constant model misfits.

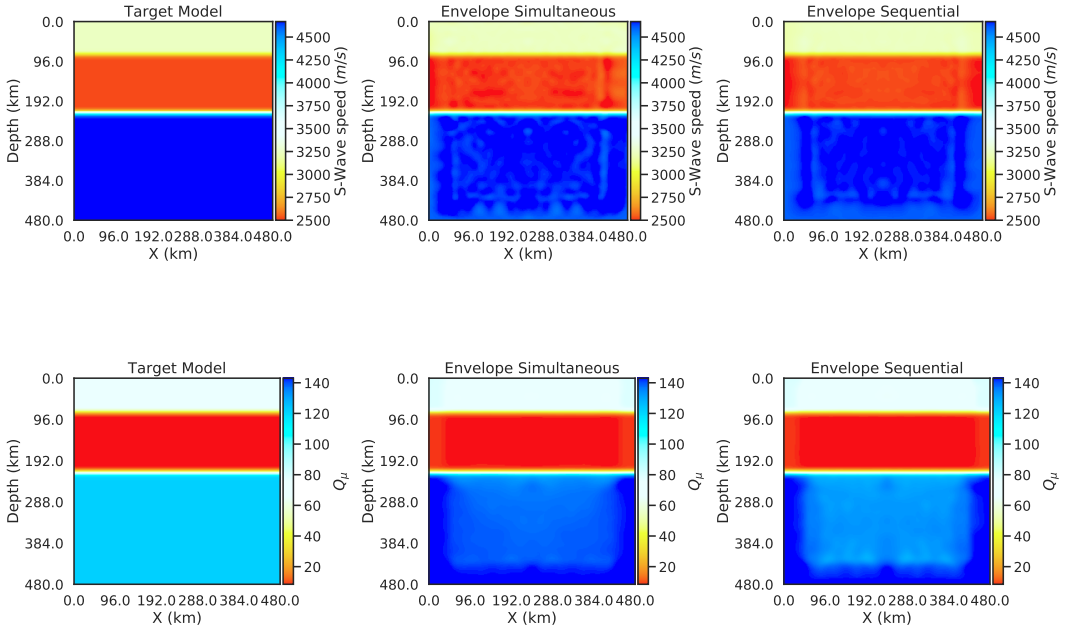

FIGURE 5.11: Inversion results after 19 iterations for S-wave velocity (top) and Q_μ (bottom) using simultaneous and sequential inversions with envelope misfit.

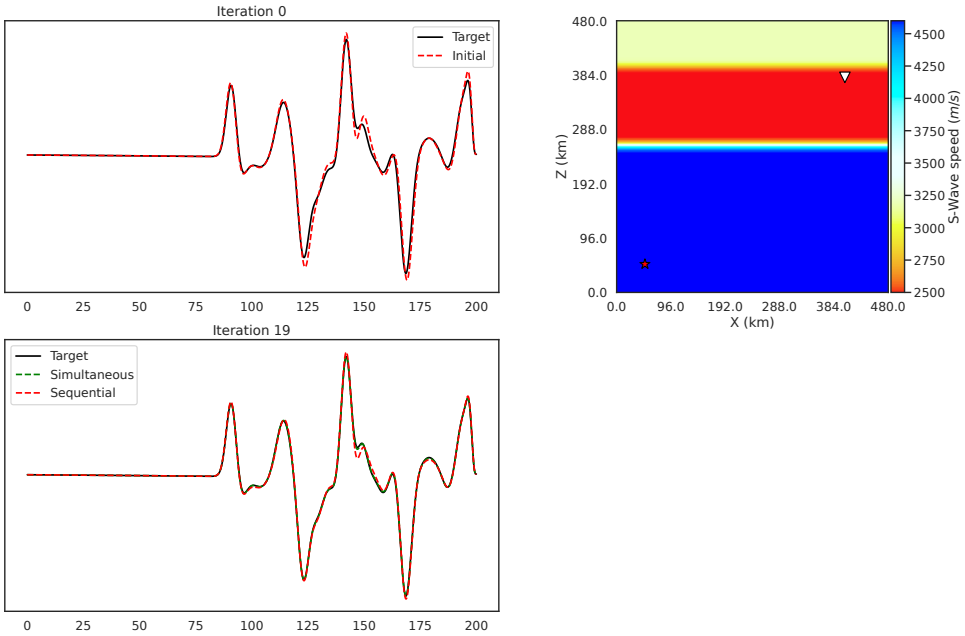

FIGURE 5.12: Waveforms for different models. Top-left shows the target and initial model waveforms, top-right shows the source and station associated with the waveform and bottom row shows the waveforms for envelope misfit inversion for sequential (red) and simultaneous (green) methods. Simultaneous inversion's waveform shows better fit to target waveform (black).

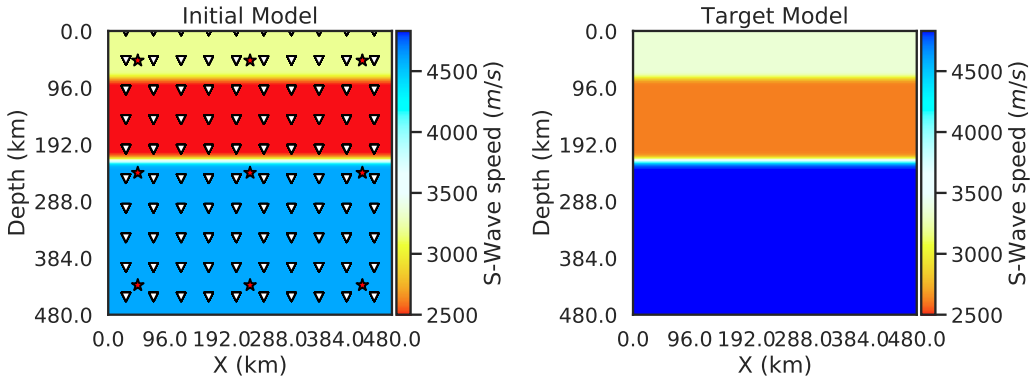

FIGURE 5.13: Initial and target S-wave velocity models. Three layered models with low-velocity layer in the middle. Target model is 5% perturbed version of the initial model. Nine sources (red stars) and 100 stations (white downward triangles) are superimposed on the initial model.

5.3.5 Case study III: Joint inversion of elastic and anelastic parameters with higher perturbations

In this example, we again have a three-layer model adapted from PREM as used in the previous section, with a low-velocity and a low-Q layer in the middle in elastic and anelastic models, respectively. In the target model, S wavespeeds (Figure 5.13) and Q_μ values (Figure 5.14) are perturbed by 5% and 20% around the corresponding PREM values. We aim to examine how iterations behave for larger model perturbations. Figure 5.15 shows the data misfits for sequential and simultaneous inversions performed with ENV. Both strategies are able to decrease the misfit consistently while the simultaneous inversion of elastic and anelastic parameters gives slightly better data misfit reduction. On the other hand, the sequential inversion give slightly better Q_μ model misfit reduction (Figure 5.16). On the contrary, the S-wavespeed model misfit is better for the simultaneous inversion (Figure 5.17). The inverted models after 21 iterations (Figure 5.18) show similar results to those in the previous section. There are some fluctuations on model misfits which might likely be due to the trade-off between parameters throughout the iterations. In Figure 5.19, we show sample seismograms computed for target and initial models as well as final models from simultaneous and sequential inversions where the waveform misfit nicely reduced by both strategies.

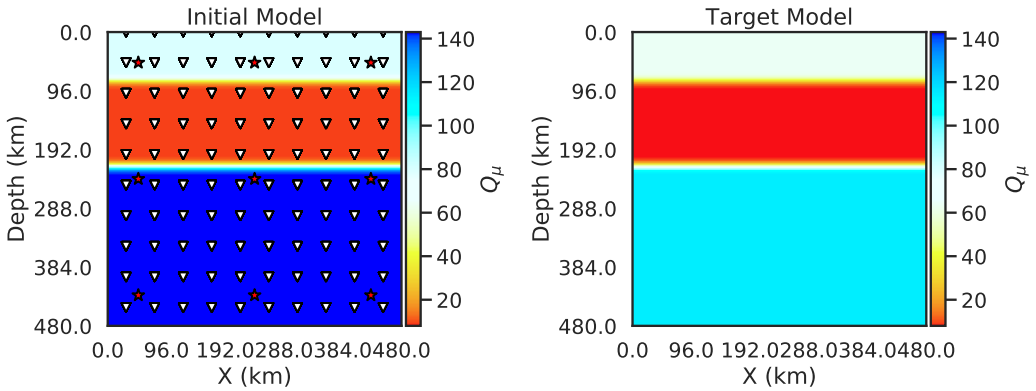

FIGURE 5.14: Initial and target Q_μ models. Target model is 20% perturbed version of the initial model. Nine sources (red stars) and 100 stations (white downward triangles) are superimposed on the initial model.

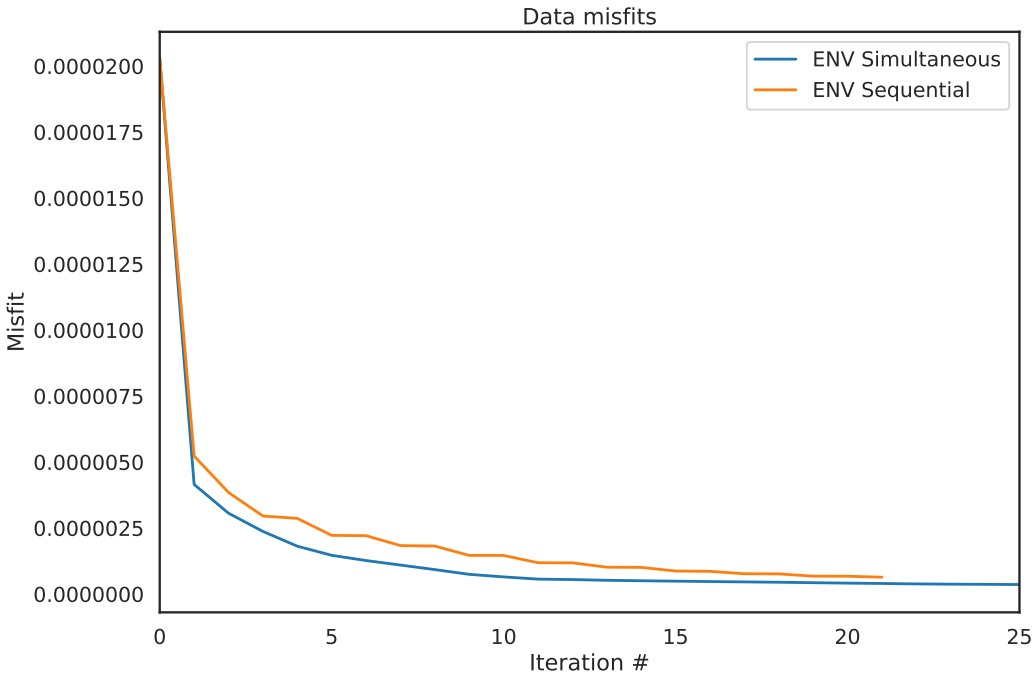

FIGURE 5.15: Data Misfits for the different inversion strategies. Sequential inversion method was not able to decrease the misfit after 21 iterations.

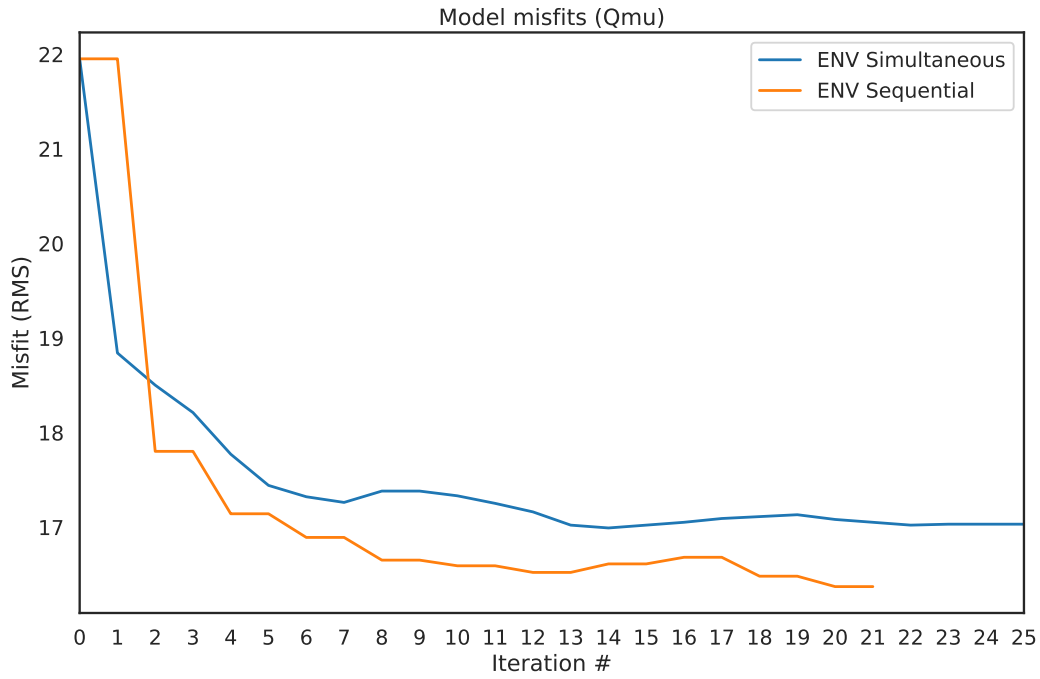

FIGURE 5.16: Q_μ model misfits for sequential and simultaneous inversion. Simultaneous inversion reduces the anelastic model misfit more slowly than sequential version. Sequential version does not update the anelastic model on odd iterations which can be seen as the constant model misfits.

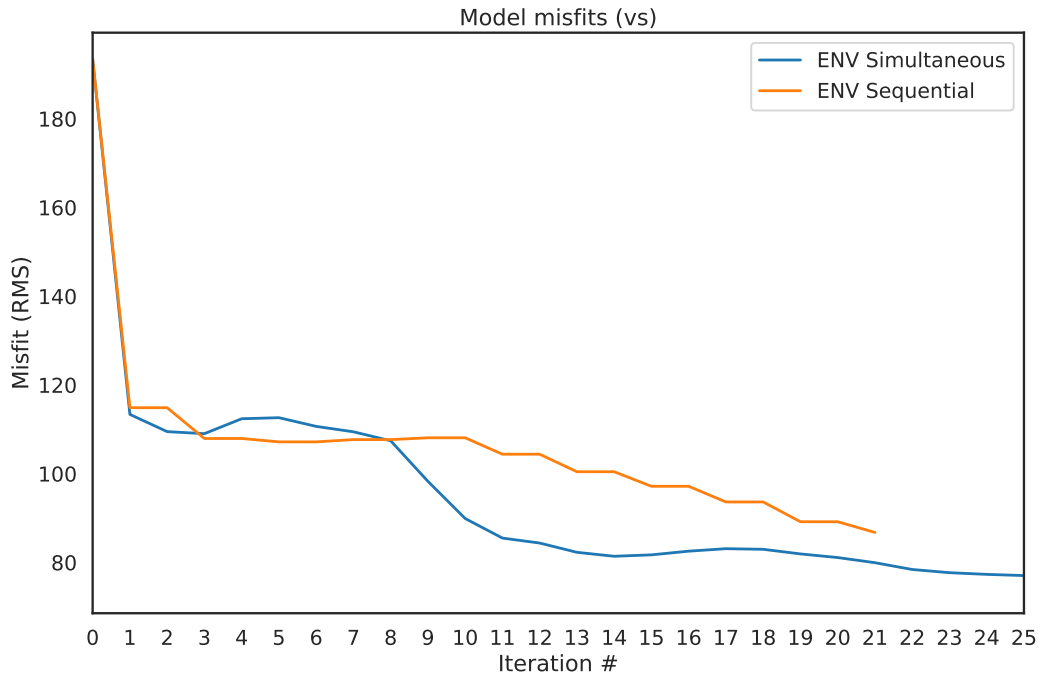

FIGURE 5.17: S-wave model misfits for sequential and simultaneous inversion. simultaneous inversion was able to decrease the elastic model misfit slightly better for this example. Sequential version does not update the elastic model on even iterations which can be seen as the constant model misfits.

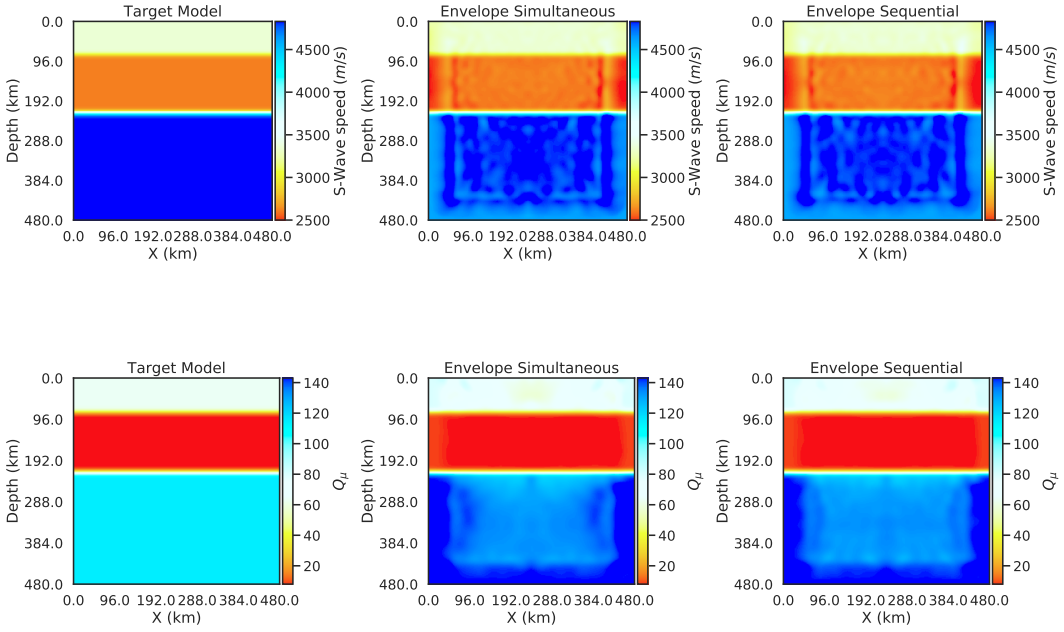

FIGURE 5.18: Inversion results after 21 iterations for S-wave velocity (top) and Q_μ (bottom) using simultaneous and sequential inversions with envelope misfit.

FIGURE 5.19: Waveforms for different stations. Left column shows the target and initial waveforms and right column shows the simultaneous inversion (green) and sequential inversion (red) waveforms for Iteration 21. Difference between the inversion methods are not easily visible but data misfits shows slightly better data misfits for simultaneous inversion results.

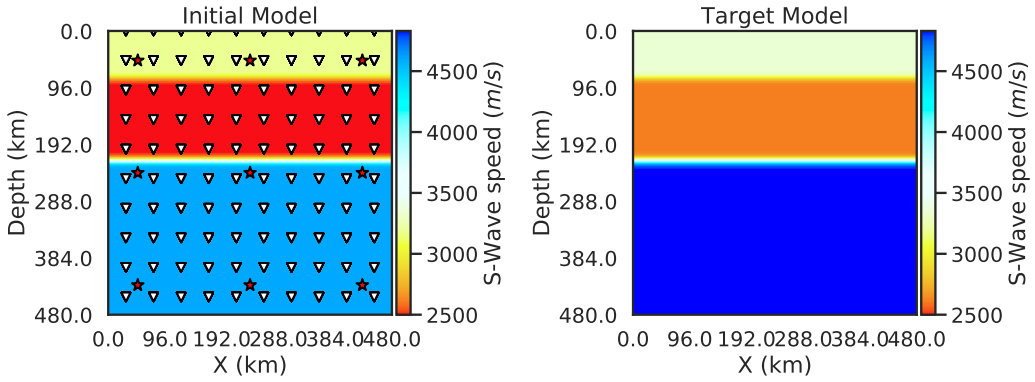

FIGURE 5.20: Initial and target S-wave velocity models. Three layered models with low-velocity layer in the middle. Target model is 5% perturbed version of the initial model. Nine sources (red stars) and 100 stations (white downward triangles) are superimposed on the initial model.

5.3.6 Case study IV: Limitations of joint inversions

Using the same elastic and anelastic layered models we increase the perturbations of S wavespeed and Q_μ to 5% (Figure 5.20), and 50% (Figure 5.21), respectively. Figure 5.22 shows the data misfits where both simultaneous and sequential inversions nicely improve the waveform fits. In Figure 5.24, the Q_μ model misfit from the simultaneous inversion shows a smooth reduction whereas the sequential inversions start being problematic after the 9th iteration although it starts decreasing again after the 19th iteration. The model misfit reduction for S wavespeeds (Figure 5.23) is more challenging since both strategies show some oscillations at the beginning with slightly better behaviour of the simultaneous inversion. Figure 5.26 shows the inverted models. In this case Q_μ model is inverted better using the simultaneous inversion but S-wave model model misfit is slightly lower for sequential inversions. These results suggest that 50% perturbation in Q_μ we are likely running into nonlinearity problems where it becomes more challenging to solve the trade-off between parameters. Sample waveforms before and after inversions are presented in Figure 5.26. This example confirms the need for a good starting model for the success of the anelastic inversions where measurement strategies, such as starting iterations with the ENV misfit, may help to decrease the nonlinear behaviour with caution.

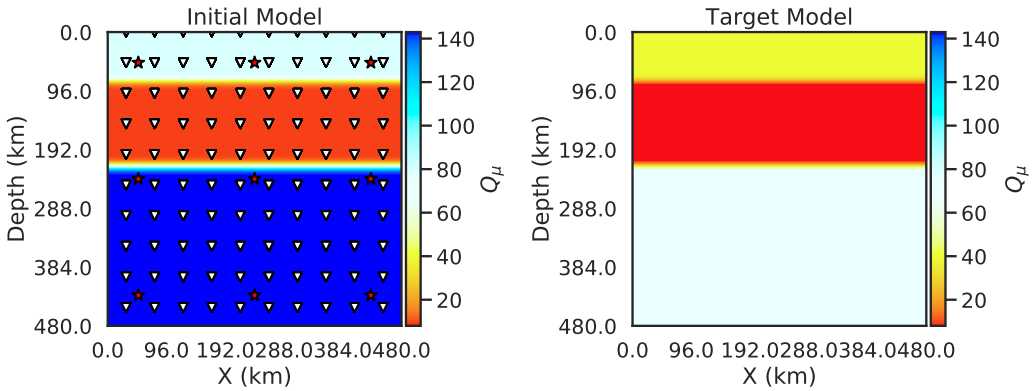

FIGURE 5.21: Initial and target Q_μ models. Target model is 50% perturbed version of the initial model. Nine sources (red stars) and 100 stations (white downward triangles) are superimposed on the initial model.

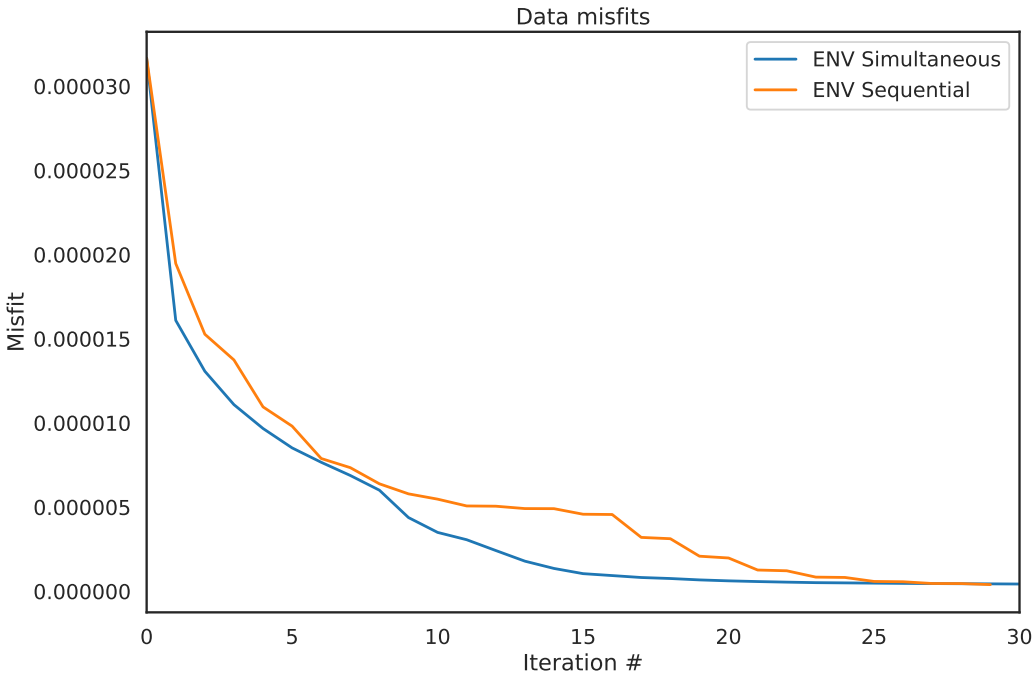

FIGURE 5.22: Data Misfits for the different inversion strategies. Sequential inversion method gives slighlty worse results and was not able to decrease the misfit after 29 iterations.

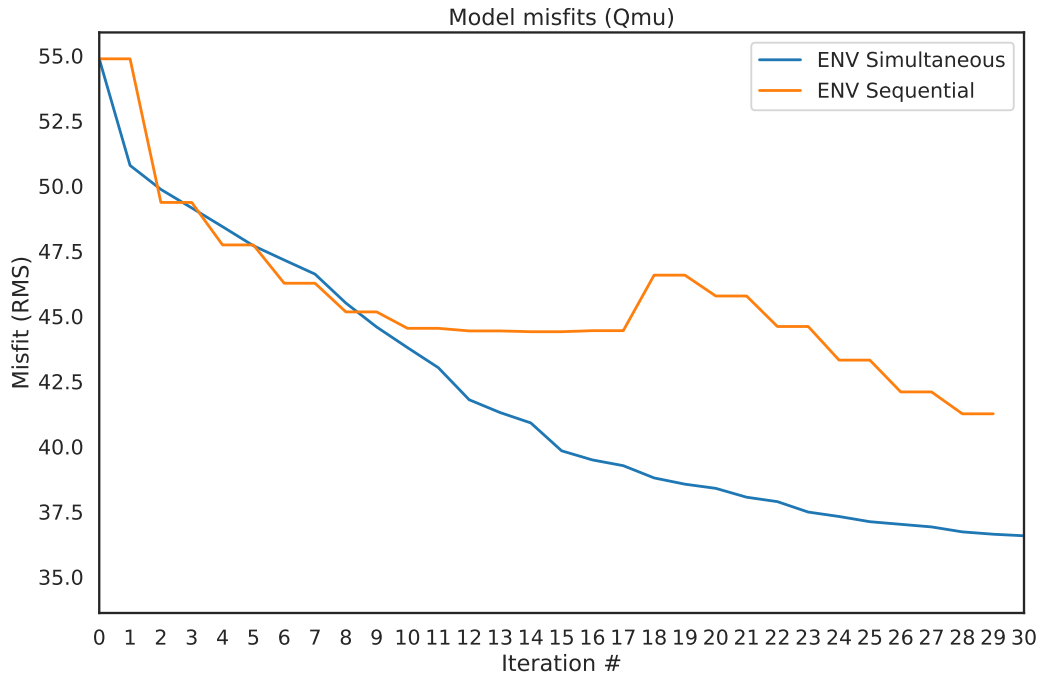

FIGURE 5.23: Q_μ model misfits for sequential and simultaneous inversion. Simultaneous inversion reduces was able to decrease the model misfit correctly however sequential inversion was not able decrease the model misfit after 9th iteration and there is an increase of model misfit after 17 iteration.

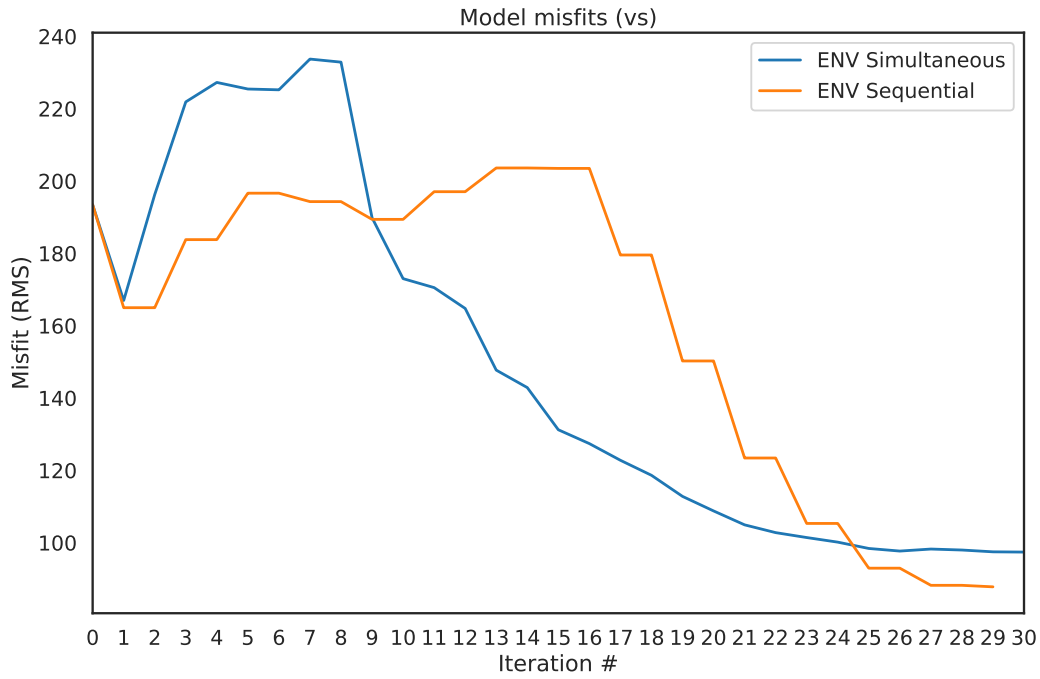

FIGURE 5.24: S-wave model misfits for sequential and simultaneous inversion. Both inversion methods show erratic behaviour for the velocity model misfit. This can be explained by the effect of the high Q_μ perturbation. High perturbation leads to increased non-linearity and tradeoff between the elastic and anelastic parameters leads to unstable behavior.

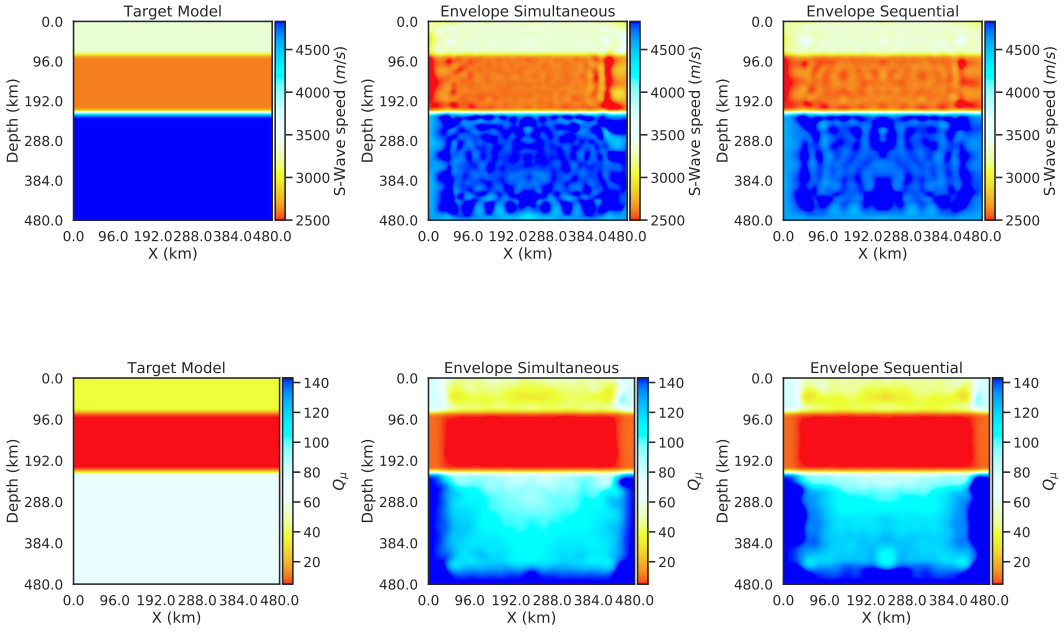

FIGURE 5.25: Inversion results after 29 iterations for S-wave velocity (top) and Q_μ (bottom) using simultaneous and sequential inversions with envelope misfit.

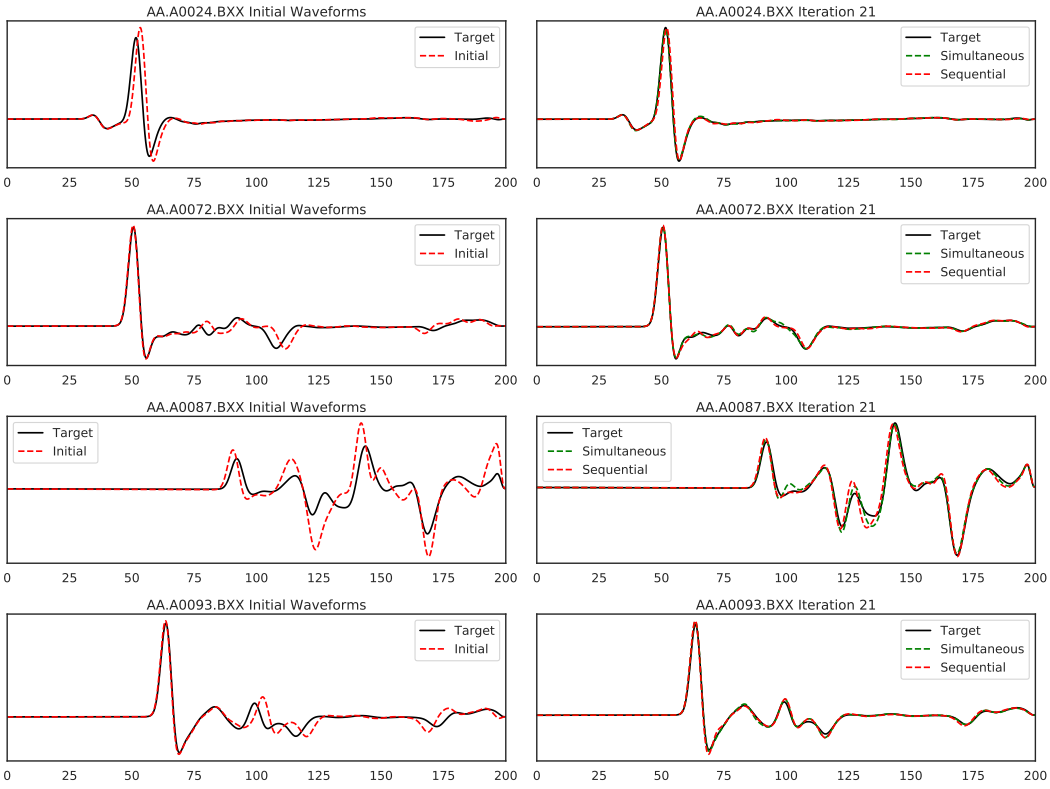

FIGURE 5.26: Waveforms for different stations. Left column shows the target and initial waveforms and right column shows the simultaneous inversion (green) and sequential inversion (red) waveforms for Iteration 29. Although total data misfit is better for simultaneous inversion results, AA.A0087 station which is located in the low-velocity layer show better data misfit for sequential inversion.

5.4 Towards global anelastic adjoint inversions

Based on our experience and observations from the 2D experiements performed in the previous section, our ultimate goal is to construct a global anelastic model of the mantle using full-waveform inversion. Considering the global data coverage, our initial focus will be on the upper mantle. To this end, in this section, we performed one test iteration making cross-correlation multitaper traveltime and amplitude measurements on observed and 3D synthetic data at the global scale. Since the anelasticity affects both the phase and amplitudes of waveforms they both must be considered in the inversion, which leads to the "full-waveform inversion" in the exact sense. The challenge is that the phase and amplitudes are also sensitive to the lateral variations in the elastic structure as well as source effects. Generally, global elastic inversions are performed with a fixed anelastic model and anelastic inversions are done with a fixed elastic model. Due to the “cross-talk” between seismic parameters simultaneous inversions would be the better option.

Commonly, 3D elastic models are constructed as a perturbation from 1D Q models. The first-order discontinuities of the Earth are well captured in 1D reference models (e.g., PREM (Dziewonski and Anderson, 1981)) and there is almost a consensus on the global long-wavelength elastic structure (Ritzwoller et al., 2002; Lekic et al., 2012). However, error bars on the radially symmetric Q structure is large, specifically in the crust and upper mantle (Resovsky et al., 2005). Furthermore, despite the progress in anelastic inversions, there is a limited agreement between the current 3D Q models (Karaoğlu and Romanowicz, 2018a), which needs to be investigated further.

Amplitudes provide sharper elastic images compared to those from travel-time measurements only. However, in 3D global experiments, upper-mantle attenuation has a profound affect on seismic waveforms, specifically those of surface waves (Bozdağ and Trampert, 2010), that we have to invert for elastic and anelastic structures simultaneously when amplitudes are used while more carefully treating the source parameters, which make the inversions more challenging due to potential trade offs between model parameters. In both GLAD-M15 (Bozdağ et al., 2016) and GLAD-M25 (Lei et al., 2020), the same radially-symmetric Q model, QL6 (Durek and Ekström, 1996), is used, which was fixed throughout the iterations. The reason for using QL6 was that it was also the Q model of S362ANI (Kustowski et al., 2008), which was the starting model of GLAD-M15. It is crucial to update the anelastic structure also for the elastic model updates as attenuation can cause significant physical dispersion, specifically for teleseismic surface waves.

Anelastic adjoint tomography previously demonstrated at the European scale (Zhu et al., 2013) where potential water content and partial melting at subduction zones were exhibited by high attenuation. The trade-off between parameters and the success of the constructed models needs to be investigated further. At the global scale, the most advanced technique so far was used by Karaoğlu and Romanowicz (2018a) for surface-wave modeling to construct an upper-mantle anelastic model using a hybrid approach based on spectral-element simulations and asymptotic data sensitivity kernels where shear-wavespeeds and attenuation were inverted simultaneously. In this study, our ultimate goal is to construct a mantle attenuation model based on adjoint tomography. It is computationally expensive, however, in the current resolution of tomographic images, such expenses are justified, even necessary to further improve the resolution (Bozdağ et al., 2016; Komatitsch et al., 2016). Anelastic adjoint inversions have not been performed at the global scale yet through the simultaneous inversion of elastic (both Vp & Vs) and anelastic parameters, which would

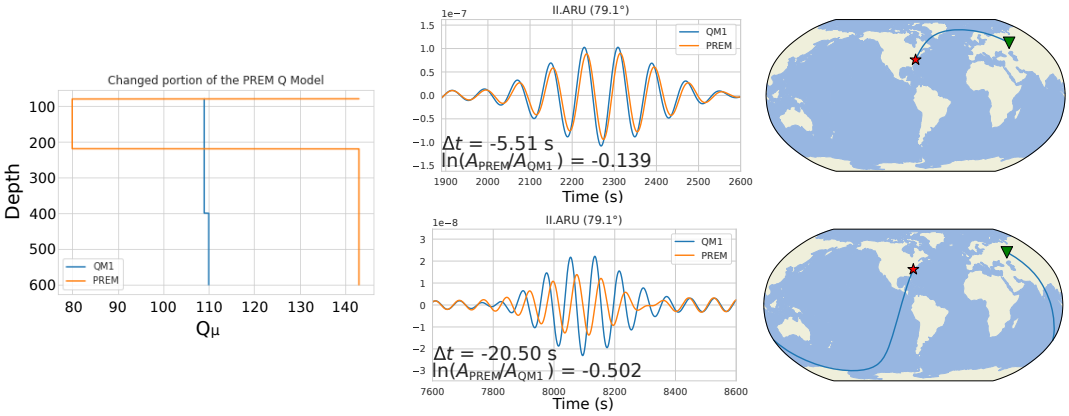

FIGURE 5.27: Minor and major arc Rayleigh waves for the fixed PREM elastic model (Dziewonski and Anderson, 1981) and two different Q_μ upper mantle models PREM & QM1 (Widmer et al., 1991) (left). Significant phase and amplitude changes can be seen, especially in the major arc. Seismogram is bandpassed filtered between 45-110 s and located at 79 degrees.

lead to the first global full-waveform inversion model in the exact sense by assimilating both the phase and amplitude information in inversions. We initially target to retrieve a long-wavelength anelastic upper-mantle model with lateral variations using the dataset of 253 events that were used in the construction of GLAD-M15 (Bozdağ et al., 2016).

5.4.1 Effect of attenuation on waveforms

We computed two sets of synthetic seismograms for two different Q models. In the first set, we used the PREM Q model and in the second set, we replaced the upper-mantle Q values of PREM with those of QM1 (Widmer et al., 1991) (Figure 5.27). The PREM Q model has a significantly high-attenuation layer above 200 km which is not pronounced in QM1. When we look at the surface waveforms of Rayleigh waves, the difference in the upper-mantle Q models creates about 6 s time shift for minor-arc waves at a distance of 79° where PREM Q attenuates the amplitude slightly more. As the propagation distance increases the impact of different Q models becomes more visible. For major-arc surface waves, the phase shift becomes more than 20 s, which is almost half of the minimum period of the waveforms (45 s), and amplitudes of the PREM Q synthetics decrease almost by half of that of QM1. Thus, upper mantle Q models can cause considerable difference on amplitudes but also large phase shifts which can even lead to cycle-skipping issues in measurements.

We performed another test to further analyze the 3D Q effects on waveforms. To this end, we selected 50 earthquakes which were not used in the construction of model GLAD-M15 (Bozdağ et al., 2016). Following its starting model S362ANI (Kustowski et al., 2008), GLAD-M15 was also constructed based on the 1D Q model of QL6 (Durek and Ekström, 1996). We computed two sets of seismograms with a resolution down to 17 s. In the first set, we used GLAD-M15 with its Q model QL6 and in the second set, we used GLAD-M15 with the 3D Q model of Dalton et al. (2008). We made cross-correlation traveltime and amplitude measurements between the computed synthetics and the associated observed data for the 45-110 s period band using both minor- and major-arc body and surface waves. In Figure 5.28, the histograms of cross-correlation amplitude measurements on three components (vertical, radial

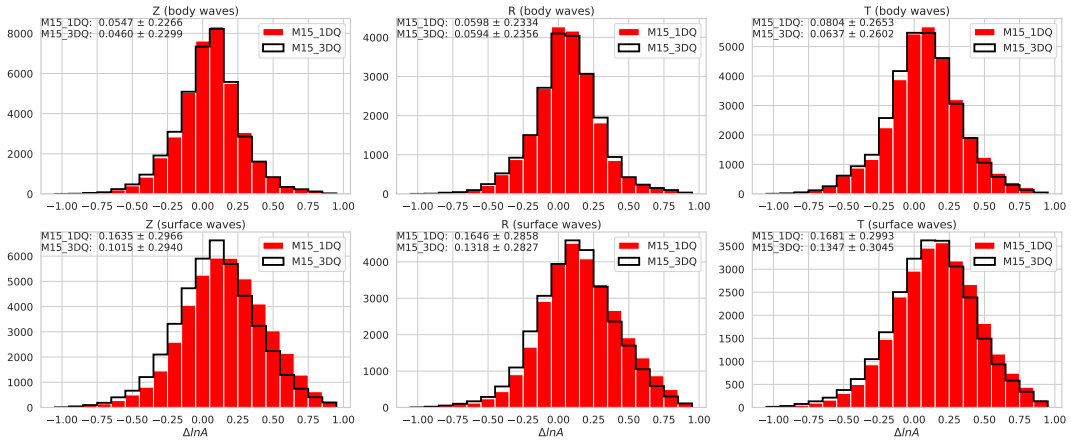

FIGURE 5.28: Body and surface wave amplitude measurements for GLAD-M15 elastic model with 1D Q model QL6 (Durek and Ekström, 1996) which was used in the construction of the elastic model and 3D Q model QRFSI12 (Dalton et al., 2008) in 45 to 110 seconds period range. Mean and standard deviation of the measurements are written at the top left of the plots. Using 3D Q model slightly improves the results.

and transverse) for body and surface waves are presented. The amplitude measurements of the 3D Q synthetics slightly work better by centering the histograms, which specifically becomes more prominent for surface waves. However, the cross-correlation traveltime measurements do not exhibit a similar pattern (Figure 5.29). The effect of Q models on surface waves are more pronounced and the 3D Q model increases the traveltime shifts compared to those from the 1D Q model. These results illustrate that elastic models can not be combined with arbitrary Q models since they are constructed based on a specific Q model where the wavespeeds are shifted according to the physical dispersion at a specific reference frequency. This demonstration is also crucial for choosing a starting Q model for the adjoint inversions. The results suggest that the starting model should be fixed to the reference Q model of the starting elastic model, which is in our case QL6 (Durek and Ekström, 1996), to continue iterations from GLAD-M15 (Bozdağ et al., 2016) or GLAD-M25 (Lei et al., 2020), which is the successor of GLAD-M15 as a result of 10 more iterations with a larger data set. The results also highlight that the full-waveform inversion is essential by simultaneously inverting elastic and anelastic parameters to further improve also the elastic structure where shear velocities involve the physical dispersion due to the reference Q model.

5.4.2 3D numerical forward & adjoint simulations

In this study, we use the global wave propagation solver the SPECFEM3D_GLOBE package (Komatitsch and Tromp, 2002a; Komatitsch et al., 2002) to perform forward and adjoint simulations to compute 3D seismic wave propagation and 3D data sensitivities (Fréchet kernels), respectively. GLAD-M15 and its successor GLAD-M25 were constructed based on wave simulations with accuracy down to 17 s (NEX=256, where NEX is the number of spectral elements along one side of each chunk at the surface. Note that the globe consists of six chunks in the global spectral-element simulations). The Moho variations, which directly significantly affect surface-wave propagation, are honored in numerical simulations as described in Tromp et al. (2010b) specifically for better sampling the thin oceanic crust (on average 7 km). Topography, bathymetry, ellipticity, rotation, gravity (Cowling approx.), attenuation

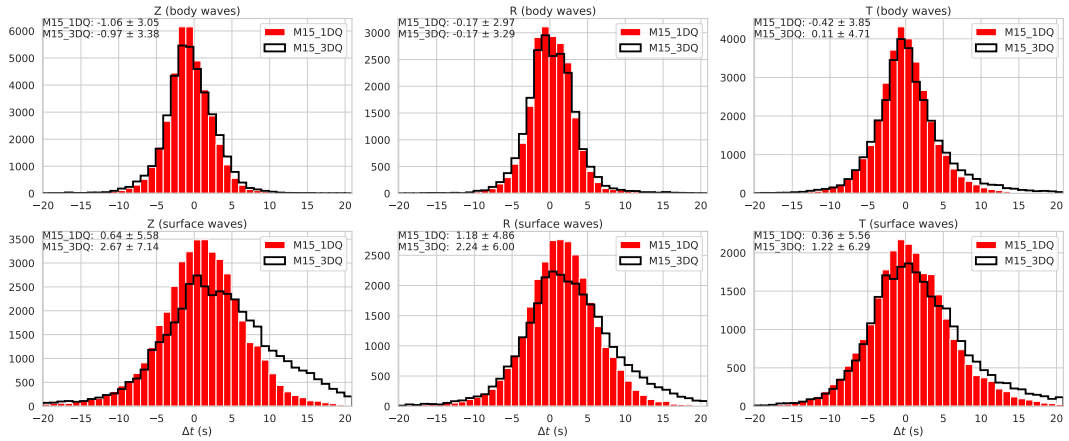

FIGURE 5.29: Body and Surface wave traveltime measurements for GLAD-M15 elastic model with 1D Q model QL6 (Durek and Ekström, 1996) which was used in the construction of the elastic model and 3D Q model QRFSI12 (Dalton et al., 2008) in 45 to 110 seconds period range. Mean and standard deviation of the measurements are written at the top left of the plots. Measurement distributions do not drastically change using 3D Q model. On the other hand, for surface wave traveltimes mean values are bigger and standard deviations are higher using the 3D Q model.

and the ocean load are all taken into account during forward and adjoint simulations. Simulations are performed for each earthquake, the computational cost is independent of the number of seismic stations.

In adjoint inversions, Fréchet derivatives can be computed based on two numerical simulations: (1) a forward simulation initiated by a regular source, such as an earthquake, and recorded at a receiver, and (2) an adjoint simulation initiated by placing an adjoint source based on the data misfit at the location of the station and recorded at the location of the regular source (Tarantola, 1984; Tromp et al., 2005). However, two adjoint simulations are needed to compute elastic and anelastic gradients for the current setup of the solver as we describe below.

In this set up, we choose GLAD-M25 (Lei et al., 2020) as the elastic starting model and its radially-symmetric Q model QL6 (Durek and Ekström, 1996) as the anelastic starting model for FWI. The total misfit function may be defined as the sum of the phase and amplitude misfits (Zhu et al., 2015) or a direct waveform misfit may also be used (Virieux and Operto, 2009; Li and Romanowicz, 1996). Our results based on 2D inversions suggest that it might be a good strategy to start iterations with an envelope misfit to linearize the problem, especially while starting form a 1D Q model. Here, to be consistent with the previous global adjoint tomography studies we performed a test iteration using multitaper traveltime and amplitude measurements which have been demonstrated to be robust, and with appropriate windowing strategies (i.e., using an automated phase picking algorithm FLEXWIN (Maggi et al., 2009)) the measurements can be linearized while information going into inversions can be maximized (e.g., Bozdağ et al., 2011).

To perform anelastic-elastic simultaneous inversions, we need to calculate two sets of gradients without the need of changing the numerical solver: one for elastic and one for anelastic parameters. Assuming that attenuation is constant over the seismic frequency band (Liu et al., 1976), the anelastic gradient can be computed by the anelastic adjoint source modified from the elastic adjoint sources by taking its Hilbert transform (e.g., Tromp et al., 2005; Bozdağ et al., 2011). In this study, we ignore the bulk attenuation since intrinsic attenuation is dominated by shear

attenuation. The models can then be iteratively updated using a gradient-based optimization method, such as conjugate gradient or L-BFGS (Nocedal, 1980).

5.4.3 Measurements

For this experiment, we use three period bands 17-40 s, 45-110 s and 90-250 s, similar to the measurements of previous GLAD models. Measurement windows are selected for each period band using the automated window selection algorithm (Maggi et al., 2009). In the shortest period range, we only select body waves. For the 45-110 seconds, we select both body and surface waves and for the longest period band we select surface waves only. 253 events that were used in the construction of the GLAD-M15 (Bozdağ et al., 2016) are used in this experiment. For each earthquake, 180-minute seismograms are computed to capture major-arc body and surface waves. Frequency-dependent multitaper measurements (Laske and Masters, 1996) are used for measuring traveltimes and amplitudes. The 2D experiments favor the envelope misfit (Bozdağ et al., 2011) which should be considered in the further iterations, together with the exponentiated phase misfit (Yuan et al., 2020).

5.4.4 Parameterization, model updates

In the ideal case, we should also be inverting source parameters together with anelastic and elastic parameters (Valentine and Woodhouse, 2010b). However, because of the potential trade-offs between seismic parameters we prefer to apply source corrections to the origin time and the scalar moment of each earthquake based on a grid search (Zhu et al., 2015; Lei et al., 2020). We use both the phase and amplitude measurements which moves our adjoint inversions to complete 'full-waveform' inversions. To this end, we define our misfit function as the sum of phase and amplitude misfits:

$$\chi = \chi_{\text{phase}} + \chi_{\text{amplitude}} \,. \tag{5.16}$$

One may also use a simple L2-norm waveform misfit or other combinations of phase and amplitude measurements such as instantaneous phase and envelopes. Our 2D tests with exponentiated phase and envelope misfits are promising that we consider using them in future iterations. To perform elastic-anelastic simultaneous inversions we need to calculate two sets of gradients: one for elastic parameters and another for attenuation. To compute the anelastic gradient, we just modify the elastic adjoint sources by taking their Hilbert transform (e.g., Tromp et al., 2005; Bozdağ et al., 2011) such that,

$$\tilde{f}_i^\dagger(\mathbf{x},t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}[(2/\pi)\,\ln(|\omega|/\omega_0) - i\,\text{sgn}(\omega)]^* f_i^\dagger(\mathbf{x},\omega)\exp(i\omega t)\,\mathrm{d}\omega\,, \tag{5.17}$$

where $\tilde{f}^\dagger$ and $f^\dagger$ are anelastic and elastic adjoint sources, respectively, and ω_0 is the reference angular frequency of the model. The $(2/\pi)\,\ln(|\omega|/\omega_0)$ term captures the physical dispersion around the reference frequency and $i\,\text{sgn}(\omega)$ is responsible for amplitude variations. We can then write our total gradient as the sum of elastic and anelastic gradients computed both for phase and amplitude measurements:

$$\chi^{\text{total}} = \chi^{\text{elastic}}_{\text{phase}} + \chi^{\text{elastic}}_{\text{amp}} + \chi^{\text{anelastic}}_{\text{phase}} + \chi^{\text{anelastic}}_{\text{amp}}\,. \tag{5.18}$$

For simultaneous ealstic and anelastic inversions we use the following parameterization in inversions following Zhu et al. (2013):

$$\delta\chi(\mathbf{m}) = \int_V K_c(\mathbf{x})\delta\ln c + K_{\beta_v}(\mathbf{x})\delta\ln\beta_v + K_{\beta_h}(\mathbf{x})\delta\ln\beta_h + K_\eta(\mathbf{x})\delta\ln\eta + K_\mu(\mathbf{x})\delta Q_\mu^{-1}(\mathbf{x})\ \mathrm{d}^3\mathbf{x}\,, \tag{5.19}$$

where we kept the original transversely isotropic parameterization of GLAD-M25. c is the bulk sound speed, β_v and β_h are the vertically and horizontally polarized shear wave speeds, respectively, η is the dimensionless parameter related to the incidence angle and anisotropy (Anderson, 1961), and Q_μ^{-1} is the shear attenuation. Density of the model, (ρ), is not directly included in the parametrization to reduce the trade-off and it is updated using the empirical scaling relation with isotropic (Voight averaged) shear-wave perturbations ($\delta \ln \rho = 0.33\ \delta \ln \beta$) (Montagner and Anderson, 1989).

To remove the numerical noise and balance the imperfect data coverage we smooth our gradients. In this case, since we start from a 1D Q model with potentially large uncertainties, we used larger smoothing for the anelastic gradient. We smoothed the anelastic gradient with a Gaussian operator with half widths of 450 km and 50 km in the horizontal and vertical directions, respectively. For the elastic gradients we used 120 km and 50 km half widths of the Gaussian smoothing operator in the horizontal and vertical directions, respectively, considering the resolution of GLAD-M25 (Lei et al., 2020) as the elastic starting model. The gradients were then preconditioned with the diagonal pseudo-Hessian kernel which is obtained by the interaction between forward and adjoint acceleration fields and mimic the kernel coverage (Luo, 2012). The elastic model is updated at the first iteration using the steepest descent:

$$\mathrm{m}_{i+1} = \mathrm{m}_i \exp\left(\alpha \mathrm{d}_i\right)\,, \tag{5.20}$$

where the next model m_{i+1} is constructed perturbing the current model m_i by α determined by the line search in the direction d_i using the exponential function. We performed the line search using 25 earthquakes from our database distributed worldwide to reduce the computational cost. The attenuation model, on the other hand, is updated using the absolute values of the gradient direction rather than using the logarithmic ratio of the models

$$\mathrm{m}_{i+1} = \mathrm{m}_i + \alpha \mathrm{d}_i\,. \tag{5.21}$$

5.4.5 Source corrections

To minimize the effect of source uncertainties, one can invert moment-tensor source parameters using 3D wave simulations either by computing source Green's function in 3D background models (Liu et al., 2004) or solve them in the adjoint sense (Kim et al., 2011). Here, we prefer to apply source corrections to the origin time and scalar moment of earthquakes similar to Zhu et al. (2015) and Lei et al. (2020) to focus on the inversion of elastic and anelastic parameters. The source correction terms for origin times and scalar moments are determined for each earthquake by performing a grid search that minimizes the total misfit. To reduce the potential bias due to the uneven distribution of seismic stations a weighting scheme for measurements is introduced as described in Liu et al. (2004), such as

$$w = N_a^{-0.5}(\text{dist})^{-1}\text{N}_\text{c}^{-\text{c}},$$

where N_a is the number of stations in each azimuthal bin (azimuthal bins are assigned for each degree), N_c is the number of traces in each component and the c exponent is selected as 1, 2, 1 for vertical, radial and transverse components, respectively, which introduce a relative weighting between components. Since vertical and transverse component fits are typically better than radial components it is weighted less than other components. The distance is also considered in the weighting which is anti-correlated with the weight of the measurement. In each measurement window, cross-correlation traveltime and amplitude misfits are computed and they are combined with the weights $w_{tt} = 0.143$ and $w_{amp} = 0.857$ to balance out the misfit values which are selected based on the experimental setup.

Figure 5.30 shows an example of grid search for a selected earthquake in New Zealand where nearly 20% reduction in total amplitude misfit is observed after source corrections were applied. The positive and negative values of origin time shifts mean the that origin time is delayed (synthetics were lagging behind the observed data) and advanced (synthetics were faster than observed data), respectively. The positive and negative scalar moment changes denote smaller and larger synthetic amplitudes than observed ones, respectively. Figure 5.31 shows the distribution of correction terms for 253 earthquakes where most of the time shift corrections are less than 1 second and the total scalar moment is predominantly increasing. After the source correction is applied, traveltime and amplitude measurements recomputed to test the efficacy of the corrections. Figure 5.32 and 5.33 show the distribution of the measurements before and after the source corrections. The traveltime histograms center around zero, especially for the shorter period bands. Since it has been shown that attenuation is more apparent in longer periods due to large variation in the low Q zone at depth 80-200 km, better anelastic model might improve the long period measurements further (Ruan and Zhou, 2010). On the other hand, effect of the corrections on the amplitude measurements is much more apparent with the reduced skewness of the amplitude histograms toward the positive side. Such source corrections are expected to reduce the bias in source parameters as well as increase the number of measurement windows at each iteration by reducing amplitude differences between observed and synthetic waveforms. However, it should be noted that to reduce the trade-off, source corrections are suggested to be performed before each iteration (Zhu et al., 2015).

5.4.6 First iteration results

We demonstrated a simultaneous elastic and anelastic full-waveform inversion in 3D by performing a test iteration. We performed two adjoint simulations to obtain the elastic and anelastic gradients as described above. We used the frequency-dependent cross-correlation traveltime and amplitude measurements between observed and synthetic waveforms of minor- and major-arc body and surface waves using the dataset of 253 earthquakes used in the construction of GLAD-M15 (Bozdağ et al., 2016). The measurement categories were weighted to make sure that each category has an equal contribution to the total misfit in addition to the applied geographical weighting based on the source-receiver distribution to balance the data coverage as described in Ruan et al. (2019) and also in Chapter 3.

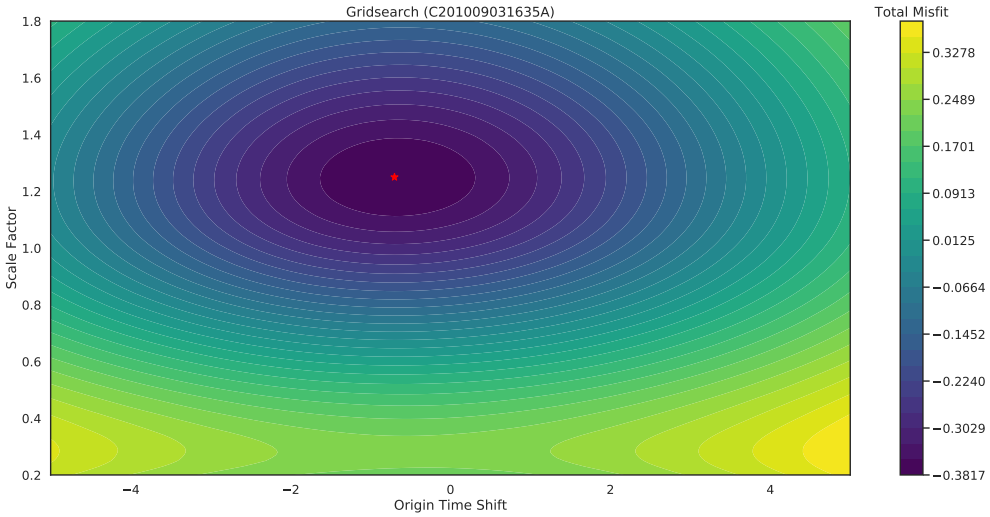

FIGURE 5.30: Grid search results for September 3rd, 2010 New Zealand Earthquake (M_w = 7.1). Optimal origin time shift and scalar moment scale factor are found to be -0.7 seconds (origin time is advanced) and 1.25 (scalar moment is increased by 25%) respectively. Source correction resulted in nearly 20% reduction in the total misfit.

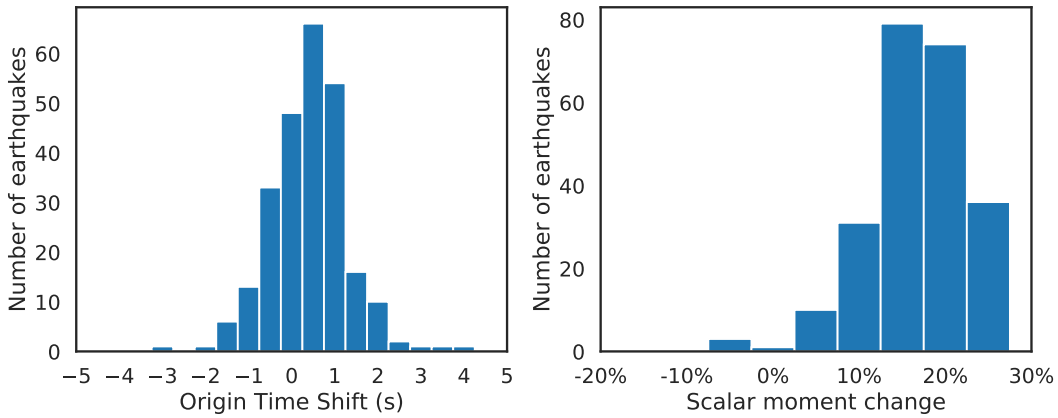

FIGURE 5.31: Histograms of source correction terms for origin time and scalar moment changes. Positive origin times corrections means delay in the origin times (slower synthetics), positive scalar moment changes point to increase in scalar moments (smaller amplitude synthetics).

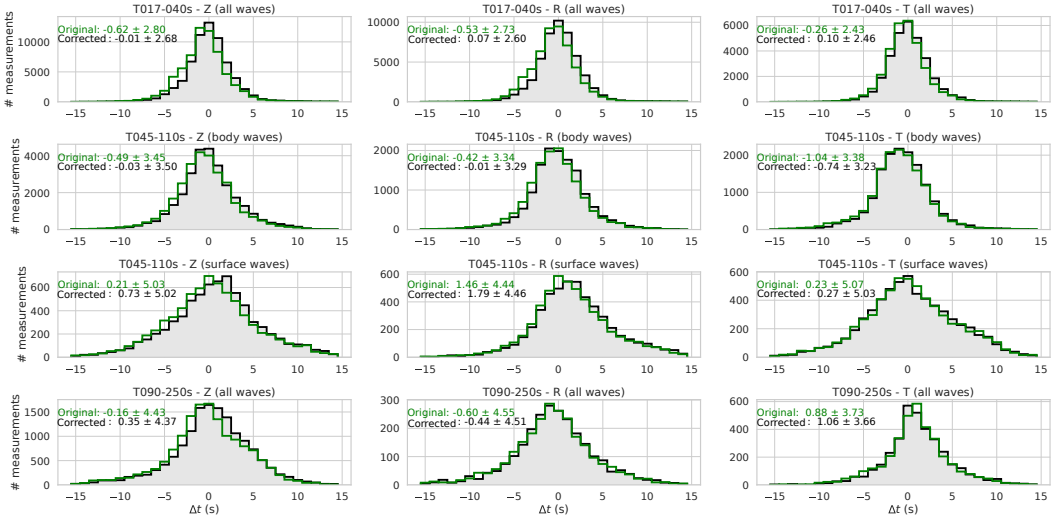

FIGURE 5.32: Traveltime measurement histograms before and after source correction. Mean values and standard deviations are written in the top-left. Measurement mean values generally gets closer to zero, especially for the shorter period body waves.

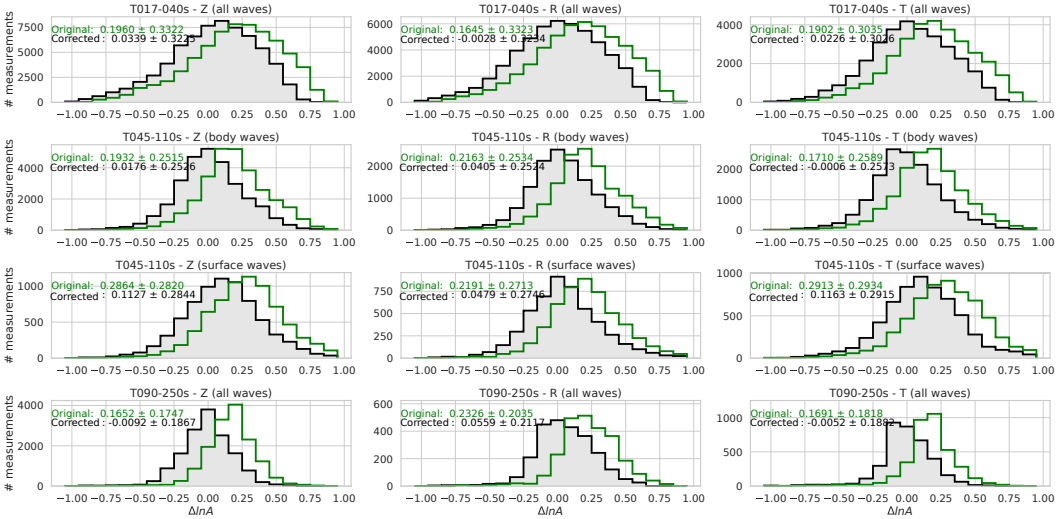

FIGURE 5.33: Amplitude measurement histograms before and after source correction. Mean values and standard deviations are written in the top-left. Correction effect is much more visible in the amplitude measurements.

After the first iteration model updates, we observe a modest decrease in data misfits in all measurement categories at all period bands, components (vertical, radial and transverse) for traveltimes and amplitudes (Figure 5.34). Part of the modest decrease in data misfits should be due to that the elastic and anelastic gradients were smoothed differently, specifically the large smoothing applied to the anelastic gradient. In addition, it is challenging but required to make sure that all measurement categories improve at each iteration. However, a few more iterations are likely needed to stabilize the simultaneous inversions by making the anelastic model comparable to the elastic one in terms of resolution.

In Figures 5.35 and 5.36, we present the cross-sections of the vertically-polarized shear wavespeed (Vsv) and shear Q models taken at 100 km and 300 km depths, respectively. The models are perturbed by 0.5% around their starting models. The anelastic model update is smoother due to the large smoothing applied to the associated gradient and the Vsv model update naturally resembles its starting model GLAD-M25 (Lei et al., 2020). Overall there is a good correlation between the Vsv and shear Q models. For instance, the slow and fast wavespeeds in the east and west costs of North America are denoted by low and high shear Q values, or high attenuation at some plume regions at 300 km such as Hawaii, etc. Similar correlations were also reported, for instance, by Dalton et al. (2009) between shear-velocity model S362ANI (Kustowski et al., 2008) and the attenuation model QRFSI12 (Dalton et al., 2008). Although these observations are promising, it is early to draw a conclusion as they are the results after only one iteration. To make reliable physical interpretations as well as to comment on the success of the simultaneous inversions more iterations must be performed showing successful decrease in misfit. In addition, the trade-off between elastic and anelastic parameters must be carefully investigated by performing, for instance, point-spread function (PSF) tests (Fichtner and Trampert, 2011a).

5.5 Discussion

In this study, we investigate strategies for anelastic adjoint inversions in 2D and 3D. We first demonstrated 2D adjoint waveform inversions to retrieve the anelastic structure for synthetic models. Then we performed one elastic and anelastic global adjoint inversion using observed data. The 2D experiments confirm that both the elastic and anelastic starting models are crucial for successful inversions. Even the simple 2D synthetic models of layered structures show strong trade-off between parameters. Potential remedies are carefully choosing appropriate measurements for full-waveform inversions and introducing constraints to the full-waveform inversion problem which might help reduce the trade-off (Aghamiry et al., 2020; Aragao and Sava, 2020).

The global elastic models are constructed based on a chosen, generally, radially-symmetric Q model. It is also common to use the phase information only to linearize measurements assuming that the phase anomaly is due to only the elastic structure. Since traveltime anomalies are also affected by physical dispersion (Kanamori and Anderson, 1977; Romanowicz, 1990) it is now a requirement to simultaneously update elastic and anelastic structures not only to construct anelastic Earth models but also to further improve elastic models.

It is well-known that misfit functions play an important role in the success of full-waveform inversions (e.g., Modrak et al., 2016). Our 2D experiments favor the envelope misfit in anelastic adjoint tomography supporting the results of Karaoğlu

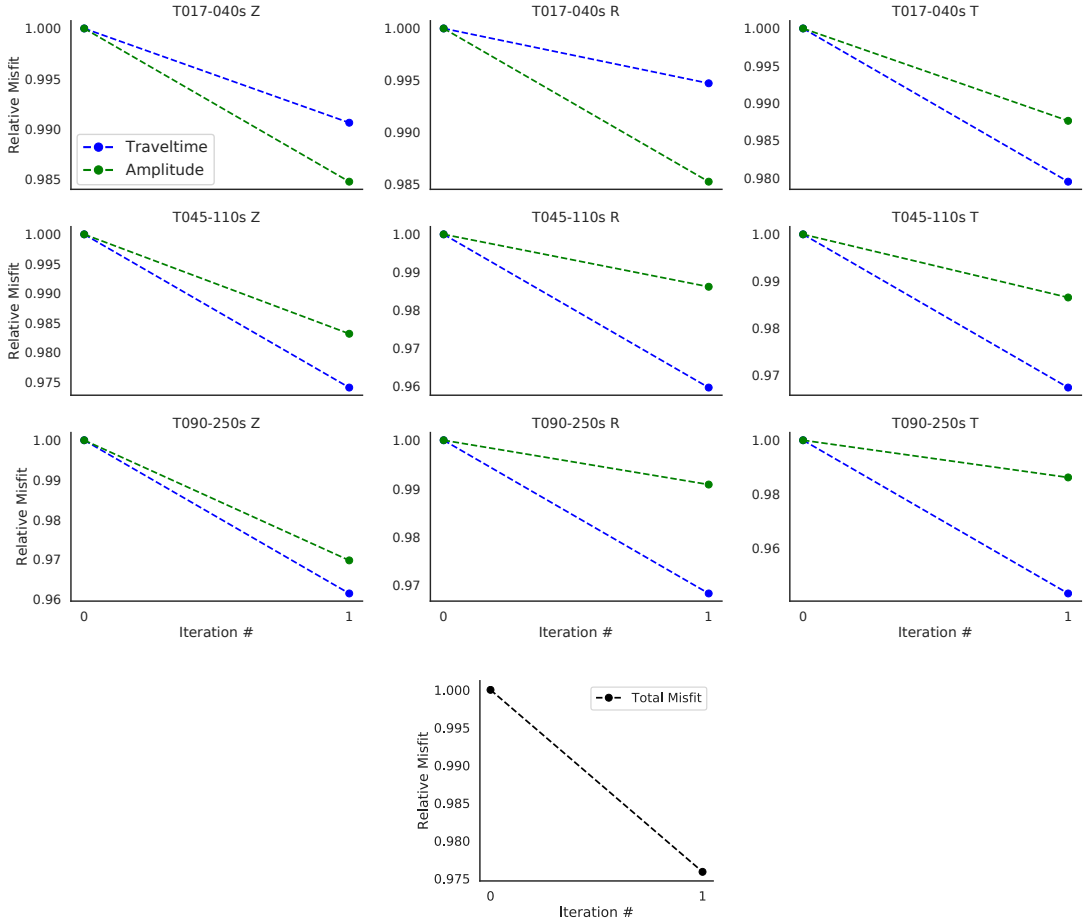

FIGURE 5.34: Misfit change for the first iteration for 3 period bands and for 3 components individually and in total.

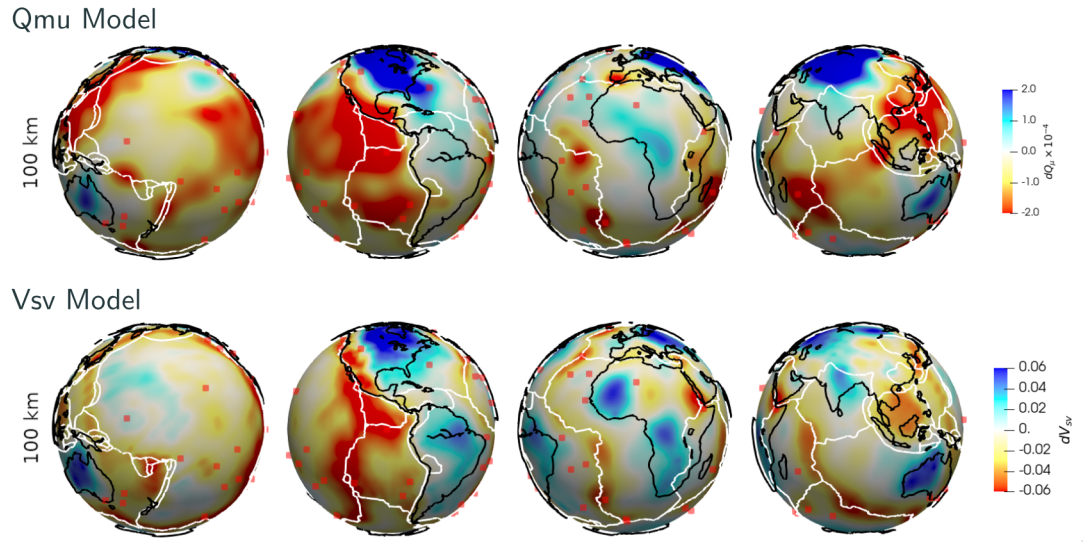

FIGURE 5.35: First iteration Vsv and attenuation models at 100 km depth.

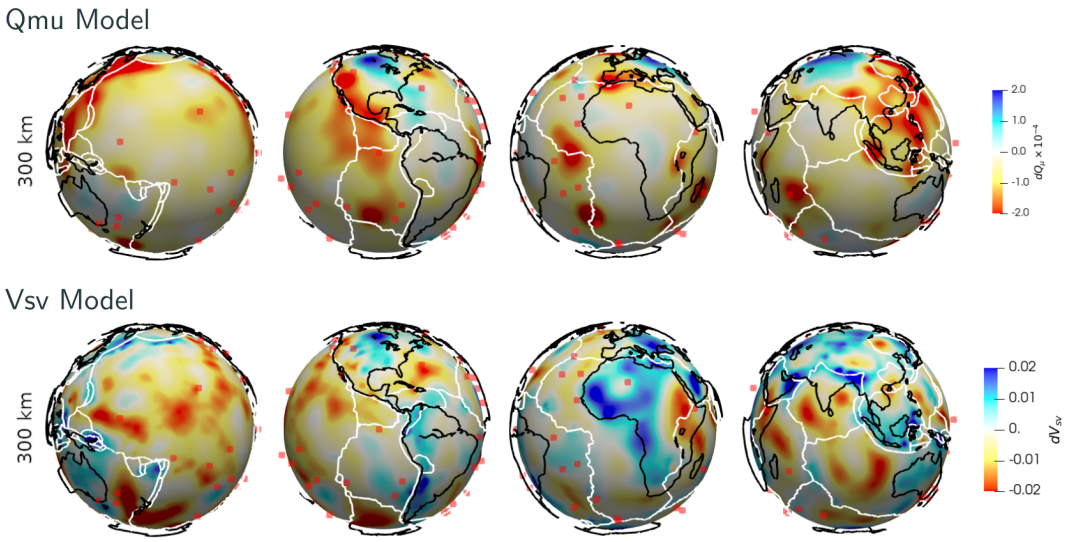

FIGURE 5.36: First iteration Vsv and attenuation models at 300 km depth.

and Romanowicz (2018a) who reported that the envelope based misfit behaves better than the classical waveform misfit for anelastic inversions. The advantage of the envelope misfit has also been reported in other studies (e.g., Yuan et al., 2015) where a multi-step approach can be a good strategy, such as starting with the envelope misfit followed by the classical L2-norm waveform misfit to better model entire waveforms. On the other hand, a combination of the envelope misfit with its counterpart exponentiated phase misfit can be a good alternative to increase resolution, specifically over the frequency-dependent (multitaper) cross-correlation traveltimes and amplitudes where the measurements generally favor main phases (Bozdağ et al., 2011).

We demonstrated to simultaneously invert for the elastic and anelastic structure. However, it is well known that source uncertainties can also affect the accuracy of measurements and source effects can be as large as path effects (e.g., Ferreira and Woodhouse, 2007). Ideally the source and structural parameters should be inverted simultaneously as well (Valentine and Woodhouse, 2010b). However, the strong trade-off between seismic parameters makes it challenging to invert everything simultaneously. An alternative approach might be to invert source and structural parameters sequentially. For instance, source parameters can be updated by computing source Green's functions in 3D background models (Liu et al., 2004) or estimating source parameters using the adjoint method similar to the structural inversions we performed here (e.g., Kim et al., 2011). However, the challenge of using 3D simulations as well as the adjoint inversions for source parameters may increase the computational cost significantly. Part of the source uncertainties, such as those in origin times, may be addressed by using differential measurements (Romanowicz, 1994b; Bhattacharyya et al., 1996; Warren and Shearer, 2002; Yuan et al., 2016). In our 3D application, to keep inversions tractable, we preferred to apply source corrections to correct the origin times and scalar moments at every iteration instead of performing source inversions to reduce the trade-off between source and structural parameters and not to increase the computational cost of numerical simulations further following (Lei et al., 2020).

Adjont inversions based on 3D wave simulations offer great opportunities to further improve the resolution of both elastic and anelastic structural parameters.

However, at the global scale, one of the main shortcomings to further improve the resolution specifically in the lower mantle and D" region as well as the structure underneath the oceans (i.e., oceanic mantle plumes) is still the data coverage due to the uneven distribution of sources and stations on the globe. The coverage in the upper-mantle is much better due to surface waves. Therefore the imperfect data coverage will be a bigger concern to image the lower-mantle attenuation where the body-wave coverage is sparse. A solution would be to combine body waves with normal-modes, however, it is not a straightforward process since this requires setting up a hybrid approach to involve normal modes because of the potential cost and stability issues of simulating normal-modes numerically. The ultimate solution would be to cover the oceans with three-component Ocean Bottom Seismometers (OBS). Meanwhile, long-lived seismic floats called MERMAIDs (e.g., Simons et al., 2009; Sukhovich et al., 2011) detect acoustic conversions from P-waves from regional and teleseismic events are promising to sample the oceans. MERMAIDs have been used in arrival-time inversions, most recently to image the Galapagos plume (Nolet et al., 2019). An interesting question is if we could use them to constrain the P-wavespeed elastic structure in adjoint inversions, which could eventually help constrain both the elastic and the anelastic structure. For now, a good strategy would be to use as many data as possible from the catalogue (Lei et al., 2020). This approach can also introduce redundant data as we cannot change the location of earthquakes and broad-band stations significantly in the near future. Using appropriate measurements to extract as much information as possible from each seismic trace could help improve the data coverage (e.g, Fichtner et al., 2009; Bozdağ et al., 2011; Yuan et al., 2016; Örsvuran et al., 2020). Furthermore, in global inversions, a multi-scale smoothing strategy can be used to smooth gradients in regions with good and sparse data coverage less and more, respectively, following Bozdağ et al. (2016). In addition, regional inversions with half-Earth simulations can be considered to target regions with good coverage in the lower mantle and the core-mantle boundary in adjoint inversions.

The uncertainty quantification and resolution analysis are long-standing challenge in seismic tomography since Backus and Gilbert (1968) and Backus et al. (1970). Checkerboard tests are commonly used in classical inversions for resolution analysis. However, they are not affordable in full-waveform inversions due to their computational cost. They are not relevant to uncertainty quantification, and their validity is also questionable for assessing the resolution (e.g., Lévêque et al., 1993). Bayesian inference would be a preferred method of inversion, but this is not feasible for 3D global adjoint inversions, even for classical 3D inversions. For resolution tests, most commonly Point Spread Function (PSF) tests (Fichtner and Trampert, 2011b) and random probing are used (Fichtner and van Leeuwen, 2015) to asses the resolution. Uncertainty quantification (UQ) for a smaller number of model parameters has been performed in non-linear inverse problems (e.g, Mosegaard and Tarantola, 1995; Sambridge, 1999; Meier et al., 2007), but the practicality of UQ is still an issue for full-waveform inversion problems due to the size of model space. In addition to assessing the quality of the constructed models with analyzing the misfit with data from an independent set of earthquakes and PSF tests, performing low-dimensional Monte-Carlo type of sampling might provide a better understanding of the resolution of certain features, for instance, oceanic mantle plumes whose 3D geometry and size may potentially be biased due to uneven data coverage.

5.6 Conclusions

Attenuation is a key parameter to constrain the thermochemical variations and the water content in Earth's mantle. It is more challenging to construct the anelastic models compared to the elastic ones due to the more complex nature of measurements involving the amplitudes of waveforms. However, attenuation not only affects amplitudes but also the phase of waveforms where elastic models are constructed as perturbations around a chosen reference attenuation model. Therefore, at the current resolution of global seismology it is now required to simultaneously invert for the elastic and anelastic models not only to construct anelastic models, that would provide invaluable constraints to interpret the composition and dynamics of the mantle, but also to further improve the resolution of elastic models.

Performing simultaneous elastic and anelastic adjoint inversions is straightforward in the current adjoint tomography workflows, which requires only an additional adjoint simulation to construct the anelastic gradient by modifying the elastic adjoint source taking its Hilbert transform. The potential trade-off between seismic parameters may be reduced by choosing appropriate measurements where our 2D experiements results suggest that it would be a good strategy to start such inversions with an envelope-based misfit function.

Our first 3D global iteration attenuation model shows promising correlation with the elastic shear wavespeed structure for the long-wavelength structure, which needs further investigation by performing more iterations and resolution tests to track the trade-off between parameters. In addition, a 3D synthetic global tomography based on simultaneous inversion of elastic and anelastic parameters at longer periods can be informative which should be possible with the current computational facilities.

5.7 Acknowledgements

The SPECFEM2D simulations and 2D inversions were performed on the Colorado School of Mines' cluster Wendian and the multi-processor Dell workstations of the Computational & Global Seismology Group. The SPECFEM3D_GLOBE simulations and the 3D adjoint tomography workflow were performed on the Oak Ridge National Laboratory's "Summit" system. The SPECFEM2D and SPECFEM3D_GLOBE packages are freely available from the Computational Infrastructure for Geodynamics (CIG).

Chapter 6

Discussions & Conclusions

In this thesis, I investigated the effect of different measurements and parameterizations in global-scale adjoint waveform inversions. I first adapted double-difference measurements to global adjoint tomography. Then, I used these measurements to construct an azimuthally anisotropic global upper mantle model based on adjoint tomography. In the last part, I demonstrated the effect of various misfit functions on anleastic inversions and performed simultaneous and sequential elastic and anelastic iterations in 2D to set up a strategy for global full-waveform inversion. Based on the observations from 2D tests and using 3D global wave simulations and full-waveform inversion, I performed one elastic and anelastic simultaneous global iteration with real data.

In Chapter 3, I demonstrated double-difference measurements, introduced into the adjoint tomography by Yuan et al. (2016), for real experiements. Double-difference measurements have great advantages such as minimizing source uncertainties for traveltime misfits, speeding up the convergence and increasing the resolution in densely covered areas. Since double-difference measurements are defined on measurement pairs from neighboring stations, they naturally increase the data sensitivity around and towards station clusters.

For global applications, some data, specifically from oceanic island stations, may not be easily paired because of the lack of neighbouring stations or rapid variations in Earth structure. Therefore, combining double-difference measurements with those from classical misfits can be useful to extract more information from seismograms. In this thesis, I mainly focused on double-difference measurements of surface waves which are intuitively highlight the structure underneath paired stations. The extension of double-difference measurements to body waves is theoretically possible and straightforward, however, needs further investigation for practical applications.

In the remaining of the thesis, I demonstrated double-difference measurements in global full-waveform inversion and explored different parametrizations for imaging Earth's mantle. In Chapter 4, I demonstrated double-difference measurements to construct a global azimuthally anisotropic upper-mantle model by performing 10 conjugate-gradient iterations using minor- and major-arc surface waves. Our preliminary results are promising to capture global pattern of azimuthal anizotropy consistent with plate motions and previous studies at the large scale (e.g., Schaeffer et al., 2016) and achieve continental-scale resolution in global inversions underneath densely covered regions (e.g., Becker et al., 2012b; Zhu et al., 2015). This is likely because of our parameterization of the models on the mesh of numerical simulations and taking the advantage of 3D wave simulations in imaging. Nevertheless, the initial results suggest that the model has not converged yet as the magnitude of azimuthal anisotropy is still weak compared to observations. We plan to perform 5–10 more iterations to finalize the inversions. However, the trade-off between shear

wavespeeds and anisotropy needs to be explored, for instance, by point-spread function (PSF) tests (Fichtner and Trampert, 2011a). Furthermore, a multi-stage smoothing (e.g., Bozdağ et al., 2016) may be a good strategy to highlight the resolution of the structure underneath the station clusters. The depth resolution in our inversions may also be increased by highlighting higher modes by appropriate measurements. We plan to continue iterations with the exponentiated phase misfit (Yuan et al., 2020), an instantaneous phase measurement (Bozdağ et al., 2011), which amplifies low-amplitude signals naturally by the normalization term in its adjoint source coming from the derivation. If need be, higher-modes may also be measured separately increasing their contribution in inversions with appropriate weightings in the misfit function (van Heijst and Woodhouse, 1997). In this thesis, we focused on surface-wave anisotropy in the upper mantle (e.g., Montagner and Tanimoto, 1991). Future studies should also address using body-wave anisotropy and how it can be incorporated in global-scale full-waveform inversions. Specifically to address the potential anisotropy in the lower mantle and core-mantle boundary body waves (i.e., shear-wave splitting) or normal modes must be used. The theory of body-wave anisotropy has extensively been discussed in literature and used in full-waveform inversions from exploration to earthquake seismology (e.g., Sieminski et al., 2007a). However, the parameter trade-off and the data coverage are the main challenges to be examined carefully to constrain the lower-mantle anisotropy.

In Chapter 5, we investigated the trade-off between wavespeeds and attenuation and different misfit functions for anelastic full-waveform inversion based on 2D numerical experiments, with an ultimate goal to construct a global attenuation mantle model through the simultaneous inversion of elastic and anelastic parameters. Since anelasticity can cause significant physical dispersion specifically for teleseismic surface waves, at the current resolution of tomographic images elastic and anelastic models have to be updated simultaneously to also further improve elastic parameters. Our 2D tests suggest that the envelope misfit help reduce nonlinearities of full-waveform inversion and together with the exponentiated phase can be a good misfit for elastic and anelastic simultaneous inversions. We do not observe a very significant difference between simultaneous and sequential inversions of elastic and anelastic parameters where simultaneous inversions perform slightly better. Our one simultaneous iteration of elastic and anelastic parameters at the global scale with real data shows promising preliminary results consistent with large-scale attenuation variations in 3D. However, the 2D experiements show strong trade-off between shear wavespeeds and attenuation which is a good reminder for the global-scale where inversions would be expected to be more challenging because of the uneven data coverage and the effect of other seismic parameters, where the source is likely to be the most prominent one.

Looking forward, to further improve the global images of the Earth's mantle we need to take better physics in the parameterization of the inverse problem which has to be complemented by better inversion and measurement strategies (i.e, addressing the trade-off between elastic, anealstic and source parameters, extracting more information from existing data) as well as data from sparsely covered regions (i.e., oceans). We cannot change the data coverage in the near future at the global scale but there is scope for extracting more information from every time series with appropriate measurements targeted to the scientific problem of interest. A valid approach might also be to focus on densely covered regions in the lower mantle and the core-mantle boundary, or highlighting such regions in multi-scale inversions as have been demonstrated in the literature (e.g., Afanasiev et al., 2016) mainly by incorporating high-resolution lithospheric structures up to now. Global inversions have two major

challenges that we need to tackle to further improve the resolution and reliability of tomographic images: data coverage and resolution & uncertainty quantification. While there are promising approaches for the resolution & uncertainty quantification (e.g., Liu and Peter, 2019; Fichtner and Zunino, 2019) the cost is still prohibitive to have a full assessment. One of the next biggest opportunities will likely be assimilating new datasets in inversions together with classical broadband earthquake seismograms, such as using ambient noise (e.g., Poli et al., 2012), data from emerging instruments MERMAIDs (e.g., Nolet et al., 2019) and distributed acoustic sensors (e.g., Marra et al., 2018), which have the potential to close the gap in global data coverage and help increase the resolution specifically underneath oceans.

Bibliography

Adam, J.M.-C. and S. Lebedev (2012). "Azimuthal anisotropy beneath southern Africa from very broad-band surface-wave dispersion measurements". In: *Geophysical Journal International* 191(1), pp. 155–174. DOI: 10.1111/j.1365-246X.2012.05583.x.

Afanasiev, Michael, Daniel Peter, Korbinian Sager, Saulė Simutė, Laura Ermert, Lion Krischer, and Andreas Fichtner (Jan. 2016). "Foundations for a Multiscale Collaborative Earth Model". en. In: *Geophysical Journal International* 204.1, pp. 39–58. ISSN: 0956-540X. DOI: 10.1093/gji/ggv439.

Aghamiry, Hossein S, Ali Gholami, and Stéphane Operto (2020). "Multiparameter wavefield reconstruction inversion for wavespeed and attenuation with bound constraints and total variation regularization". In: *Geophysics* 85.4, R381–R396.

Aki, Keiiti, Anders Christoffersson, and Eystein S. Husebye (1977). "Determination of the Three-Dimensional Seismic Structure of the Lithosphere". en. In: *Journal of Geophysical Research (1896-1977)* 82.2, pp. 277–296. ISSN: 2156-2202. DOI: 10.1029/JB082i002p00277.

Aki, Keiiti and Paul G. Richards (2002). *Quantitative Seismology*. 2nd ed. Sausalito, California: University Science Books. ISBN: 0-935702-96-2.

Alterman, Z. and F. C. Karal (Feb. 1968). "Propagation of Elastic Waves in Layered Media by Finite Difference Methods". en. In: *Bulletin of the Seismological Society of America* 58.1, pp. 367–398. ISSN: 0037-1106.

Anderson, Don L. (1961). "Elastic Wave Propagation in Layered Anisotropic Media". en. In: *Journal of Geophysical Research (1896-1977)* 66.9, pp. 2953–2963. ISSN: 2156-2202. DOI: 10.1029/JZ066i009p02953.

Anderson, Don L. (Nov. 1967). "The Anelasticity of the Mantle". en. In: *Geophysical Journal International* 14.1-4, pp. 135–163. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1967.tb06232.x.

Anderson, Don L. and C. B. Archambeau (1964). "The Anelasticity of the Earth". en. In: *Journal of Geophysical Research (1896-1977)* 69.10, pp. 2071–2084. ISSN: 2156-2202. DOI: 10.1029/JZ069i010p02071.

Antolik, Michael, Yu J. Gu, Göran Ekström, and Adam M. Dziewonski (May 2003). "J362D28: A New Joint Model of Compressional and Shear Velocity in the Earth's Mantle". en. In: *Geophysical Journal International* 153.2, pp. 443–466. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.2003.01910.x.

Aragao, Odette and Paul Sava (2020). "Elastic full-waveform inversion with probabilistic petrophysical model constraints". In: *Geophysics* 85.2, R101–R111.

Babuška, Vladislav (Jan. 1984). "P-Wave Velocity Anisotropy in Crystalline Rocks". en. In: *Geophysical Journal International* 76.1, pp. 113–119. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1984.tb05026.x.

Bachmann, Etienne and Jeroen Tromp (May 2020). "Source Encoding for Viscoacoustic Ultrasound Computed Tomography". In: *The Journal of the Acoustical Society of America* 147.5, pp. 3221–3235. ISSN: 0001-4966. DOI: 10.1121/10.0001191.

Backus, G., F. Gilbert, and Edward Crisp Bullard (Mar. 1970). "Uniqueness in the Inversion of Inaccurate Gross Earth Data". In: *Philosophical Transactions of the Royal*

Society of London. Series A, Mathematical and Physical Sciences 266.1173, pp. 123–192. DOI: 10.1098/rsta.1970.0005.

Backus, George and Freeman Gilbert (Oct. 1968). "The Resolving Power of Gross Earth Data". en. In: *Geophysical Journal International* 16.2, pp. 169–205. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1968.tb00216.x.

Bai, Jianyong, David Yingst, Robert Bloor, and Jacques Leveille (May 2014). "Viscoacoustic Waveform Inversion of Velocity Structures in the Time Domain". en. In: *Geophysics* 79.3, R103–R119. ISSN: 0016-8033. DOI: 10.1190/geo2013-0030.1.

Bao, Hesheng, Jacobo Bielak, Omar Ghattas, Loukas F. Kallivokas, David R. O'Hallaron, Jonathan R. Shewchuk, and Jifeng Xu (Jan. 1998). "Large-Scale Simulation of Elastic Wave Propagation in Heterogeneous Media on Parallel Computers". en. In: *Computer Methods in Applied Mechanics and Engineering*. Containing Papers Presented at the Symposium on Advances in Computational Mechanics 152.1, pp. 85–102. ISSN: 0045-7825. DOI: 10.1016/S0045-7825(97)00183-7.

Bassin, C., G. Laske, and G. Masters (2000). "The current limits of resolution for surface wave tomography in North America". In: *EOS*. Vol. F897, p. 81.

Becker, T. W., S. Lebedev, and M. D. Long (2012a). "On the Relationship between Azimuthal Anisotropy from Shear Wave Splitting and Surface Wave Tomography". en. In: *Journal of Geophysical Research: Solid Earth* 117.B1. ISSN: 2156-2202. DOI: 10.1029/2011JB008705.

Becker, Thorsten, Sergei Lebedev, and Maureen Long (2012b). "On the relationship between azimuthal anisotropy from shear wave splitting and surface wave tomography". In: *Journal of Geophysical Research* 117, B01306. DOI: doi:10.1029/2011JB008705.

Beller, S. and S. Chevrot (2020). "Probing depth and lateral variations of upper-mantle anisotropy from FWI of teleseismic body waves". In: *Geophysical Journal International* 222.1, 352–387. DOI: 10.1093/gji/ggaa069.

Ben-Hadj-Ali, Hafedh, Stéphane Operto, and Jean Virieux (June 2011). "An Efficient Frequency-Domain Full Waveform Inversion Method Using Simultaneous Encoded Sources". In: *GEOPHYSICS* 76.4, R109–R124. ISSN: 0016-8033. DOI: 10.1190/1.3581357.

Bernard, Simon, Vadim Monteiller, Dimitri Komatitsch, and Philippe Lasaygues (Aug. 2017). "Ultrasonic Computed Tomography Based on Full-Waveform Inversion for Bone Quantitative Imaging". en. In: *Physics in Medicine & Biology* 62.17, pp. 7011–7035. ISSN: 0031-9155. DOI: 10.1088/1361-6560/aa7e5a.

Beyreuther, Moritz, Robert Barsch, Lion Krischer, Tobias Megies, Yannik Behr, and Joachim Wassermann (May 2010). "ObsPy: A Python Toolbox for Seismology". en. In: *Seismological Research Letters* 81.3, pp. 530–533. ISSN: 0895-0695. DOI: 10.1785/gssrl.81.3.530.

Bhattacharyya, Joydeep, Guy Masters, and Peter Shearer (1996). "Global Lateral Variations of Shear Wave Attenuation in the Upper Mantle". en. In: *Journal of Geophysical Research: Solid Earth* 101.B10, pp. 22273–22289. ISSN: 2156-2202. DOI: 10.1029/96JB01782.

Bijwaard, Harmen, Wim Spakman, and E. Robert Engdahl (1998). "Closing the Gap between Regional and Global Travel Time Tomography". In: *Journal of Geophysical Research: Solid Earth* 103.B12, pp. 30055–30078.

Boehm, Christian, Naiara Korta Martiartu, Nicolas Vinard, Ivana Jovanović Balic, and Andreas Fichtner (Mar. 2018). "Time-Domain Spectral-Element Ultrasound Waveform Tomography Using a Stochastic Quasi-Newton Method". In: *Proc.SPIE*. Vol. 10580. DOI: 10.1117/12.2293299.

Bohlen, Thomas (Oct. 2002). "Parallel 3-D Viscoelastic Finite Difference Seismic Modelling". en. In: *Computers & Geosciences* 28.8, pp. 887–899. ISSN: 0098-3004. DOI: 10.1016/S0098-3004(02)00006-7.

Boore, David M., Kenneth L. Larner, and Keiiti Aki (1971). "Comparison of Two Independent Methods for the Solution of Wave-Scattering Problems: Response of a Sedimentary Basin to Vertically Incident SH Waves". en. In: *Journal of Geophysical Research (1896-1977)* 76.2, pp. 558–569. ISSN: 2156-2202. DOI: 10.1029/JB076i002p00558.

Boschi, Lapo and Adam M. Dziewonski (2000). "Whole Earth Tomography from Delay Times of P, PcP, and PKP Phases: Lateral Heterogeneities in the Outer Core or Radial Anisotropy in the Mantle?" en. In: *Journal of Geophysical Research: Solid Earth* 105.B6, pp. 13675–13696. ISSN: 2156-2202. DOI: 10.1029/2000JB900059.

Bozdağ, Ebru, Daniel Peter, Matthieu Lefebvre, Dimitri Komatitsch, Jeroen Tromp, Judith Hill, Norbert Podhorszki, and David Pugmire (Dec. 2016). "Global Adjoint Tomography: First-Generation Model". en. In: *Geophysical Journal International* 207.3, pp. 1739–1766. ISSN: 0956-540X. DOI: 10.1093/gji/ggw356.

Bozdağ, Ebru and Jeannot Trampert (Mar. 2008). "On Crustal Corrections in Surface Wave Tomography". en. In: *Geophysical Journal International* 172.3, pp. 1066–1082. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2007.03690.x.

Bozdağ, Ebru and Jeannot Trampert (Mar. 2010). "Assessment of Tomographic Mantle Models Using Spectral Element Seismograms". en. In: *Geophysical Journal International* 180.3, pp. 1187–1199. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2009.04468.x.

Bozdağ, Ebru, Jeannot Trampert, and Jeroen Tromp (2011). "Misfit Functions for Full Waveform Inversion Based on Instantaneous Phase and Envelope Measurements". In: *Geophysical Journal International* 185.2, pp. 845–870.

Brossier, Romain, Stéphane Operto, and Jean Virieux (Nov. 2009). "Seismic Imaging of Complex Onshore Structures by 2D Elastic Frequency-Domain Full-Waveform Inversion". In: *Geophysics* 74.6, WCC105–WCC118. ISSN: 0016-8033. DOI: 10.1190/1.3215771.

Brossier, Romain, Stéphane Operto, and Jean Virieux (Apr. 2010). "Which Data Residual Norm for Robust Elastic Frequency-Domain Full Waveform Inversion?" In: *Geophysics* 75.3, R37–R46. ISSN: 0016-8033. DOI: 10.1190/1.3379323.

Bunks, Carey, Fatimetou M. Saleck, S. Zaleski, and G. Chavent (Oct. 1995). "Multiscale Seismic Waveform Inversion". en. In: *Geophysics* 60.5, pp. 1457–1473. ISSN: 0016-8033. DOI: 10.1190/1.1443880.

Burdick, Scott et al. (May 2008). "Upper Mantle Heterogeneity beneath North America from Travel Time Tomography with Global and USArray Transportable Array Data". en. In: *Seismological Research Letters* 79.3, pp. 384–392. ISSN: 0895-0695. DOI: 10.1785/gssrl.79.3.384.

Capdeville, Y., E. Chaljub, and J. P. Montagner (Jan. 2003). "Coupling the Spectral Element Method with a Modal Solution for Elastic Wave Propagation in Global Earth Models". en. In: *Geophysical Journal International* 152.1, pp. 34–67. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.2003.01808.x.

Capdeville, Y., Y. Gung, and B. Romanowicz (Aug. 2005). "Towards Global Earth Tomography Using the Spectral Element Method: A Technique Based on Source Stacking". en. In: *Geophysical Journal International* 162.2, pp. 541–554. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2005.02689.x.

Capdeville, Yann, Laurent Guillot, and Jean-Jacques Marigo (May 2010a). "1-D Non-Periodic Homogenization for the Seismic Wave Equation". en. In: *Geophysical*

Journal International 181.2, pp. 897–910. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2010.04529.x.

Capdeville, Yann, Laurent Guillot, and Jean-Jacques Marigo (Aug. 2010b). "2-D Non-Periodic Homogenization to Upscale Elastic Media for P–SV Waves". en. In: *Geophysical Journal International* 182.2, pp. 903–922. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2010.04636.x.

Capdeville, Yann, Eléonore Stutzmann, and Jean Paul Montagner (Apr. 2000). "Effect of a Plume on Long Period Surface Waves Computed with Normal Modes Coupling". en. In: *Physics of the Earth and Planetary Interiors* 119.1, pp. 57–74. ISSN: 0031-9201. DOI: 10.1016/S0031-9201(99)00153-3.

Carcione, J. M. and P. J. Wang (1993). "A Chebyshev Collocation Method for the Wave Equation in Generalized Coordinates". In: *Comp. Fluid Dyn. J* 2, pp. 269–290.

Carcione, José M., Dan Kosloff, and Ronnie Kosloff (Dec. 1988). "Wave Propagation Simulation in a Linear Viscoelastic Medium". en. In: *Geophysical Journal International* 95.3, pp. 597–611. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1988.tb06706.x.

Carcione, José M., Stefano Picotti, Davide Gei, and Giuliana Rossi (Jan. 2006). "Physics and Seismic Modeling for Monitoring CO2 Storage". en. In: *pure and applied geophysics* 163.1, pp. 175–207. ISSN: 1420-9136. DOI: 10.1007/s00024-005-0002-1.

Castellanos, Clara, Ludovic Métivier, Stéphane Operto, Romain Brossier, and Jean Virieux (Feb. 2015). "Fast Full Waveform Inversion with Source Encoding and Second-Order Optimization Methods". en. In: *Geophysical Journal International* 200.2, pp. 720–744. ISSN: 0956-540X. DOI: 10.1093/gji/ggu427.

Cerveny, V. (2001). *Seismic Ray Theory*. en. Cambridge University Press. ISBN: 978-0-521-01822-7.

Chaljub, Emmanuel, Yann Capdeville, and Jean-Pierre Vilotte (May 2003). "Solving Elastodynamics in a Fluid–Solid Heterogeneous Sphere: A Parallel Spectral Element Approximation on Non-Conforming Grids". en. In: *Journal of Computational Physics* 187.2, pp. 457–491. ISSN: 0021-9991. DOI: 10.1016/S0021-9991(03)00119-0.

Chaljub, Emmanuel, Dimitri Komatitsch, Jean-Pierre Vilotte, Yann Capdeville, Bernard Valette, and Gaetano Festa (Jan. 2007). "Spectral-Element Analysis in Seismology". en. In: *Advances in Geophysics*. Ed. by Ru-Shan Wu, Valerie Maupin, and Renata Dmowska. Vol. 48. Advances in Wave Propagation in Heterogenous Earth. Elsevier, pp. 365–419. DOI: 10.1016/S0065-2687(06)48007-9.

Chaljub, Emmanuel, Peter Moczo, Seiji Tsuno, Pierre-Yves Bard, Jozef Kristek, Martin Käser, Marco Stupazzini, and Miriam Kristekova (Aug. 2010). "Quantitative Comparison of Four Numerical Predictions of 3D Ground Motion in the Grenoble Valley, France". en. In: *Bulletin of the Seismological Society of America* 100.4, pp. 1427–1455. ISSN: 0037-1106. DOI: 10.1785/0120090052.

Chang, Sung-Joon, Ana M. G. Ferreira, Jeroen Ritsema, Hendrik J. van Heijst, and John H. Woodhouse (Mar. 2014). "Global Radially Anisotropic Mantle Structure from Multiple Datasets: A Review, Current Challenges, and Outlook". en. In: *Tectonophysics* 617, pp. 1–19. ISSN: 0040-1951. DOI: 10.1016/j.tecto.2014.01.033.

Chapman, Mark, Enru Liu, and Xiang-Yang Li (Oct. 2006). "The Influence of Fluid Sensitive Dispersion and Attenuation on AVO Analysis". en. In: *Geophysical Journal International* 167.1, pp. 89–105. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2006.02919.x.

Chen, P., T.H. Jordan, and L. Zhao (2007). "Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods". In: *Geophysical Journal International* 170, pp. 175–181.

Chen, Yangkang, Judith Hill, Wenjie Lei, Matthieu Lefebvre, Jeroen Tromp, Ebru Bozdag, and Dimitri Komatitsch (Aug. 2017). "Automated Time-Window Selection Based on Machine Learning for Full-Waveform Inversion". In: *SEG Technical Program Expanded Abstracts 2017*. SEG Technical Program Expanded Abstracts. Society of Exploration Geophysicists, pp. 1604–1609. DOI: 10.1190/segam2017-17734162.1.

Choi, Yunseok and Tarik Alkhalifah (Sept. 2011). "Source-Independent Time-Domain Waveform Inversion Using Convolved Wavefields: Application to the Encoded Multisource Waveform Inversion". In: *Geophysics* 76.5, R125–R134. ISSN: 0016-8033. DOI: 10.1190/geo2010-0210.1.

Clouzet, P., Y. Masson, and B. Romanowicz (June 2018). "Box Tomography: First Application to the Imaging of Upper-Mantle Shear Velocity and Radial Anisotropy Structure beneath the North American Continent". en. In: *Geophysical Journal International* 213.3, pp. 1849–1875. ISSN: 0956-540X. DOI: 10.1093/gji/ggy078.

Cobden, Laura, Jeannot Trampert, and Andreas Fichtner (2018). "Insights on Upper Mantle Melting, Rheology, and Anelastic Behavior From Seismic Shear Wave Tomography". en. In: *Geochemistry, Geophysics, Geosystems* 19.10, pp. 3892–3916. ISSN: 1525-2027. DOI: 10.1029/2017GC007370.

Crampin, Stuart (Oct. 1981). "A Review of Wave Motion in Anisotropic and Cracked Elastic-Media". en. In: *Wave Motion* 3.4, pp. 343–391. ISSN: 0165-2125. DOI: 10.1016/0165-2125(81)90026-3.

Crampin, Stuart and David C. Booth (Oct. 1985). "Shear-Wave Polarizations near the North Anatolian Fault – II. Interpretation in Terms of Crack-Induced Anisotropy". en. In: *Geophysical Journal International* 83.1, pp. 75–92. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1985.tb05157.x.

Crampin, Stuart, Robert McGonigle, and David Bamford (Mar. 1980). "Estimating Crack Parameters from Observations of P-wave Velocity Anisotropy". In: *GEOPHYSICS* 45.3, pp. 345–360. ISSN: 0016-8033. DOI: 10.1190/1.1441086.

Crampin, Stuart, Theodora Volti, and Ragnar Stefánsson (July 1999). "A Successfully Stress-Forecast Earthquake". en. In: *Geophysical Journal International* 138.1, F1–F5. ISSN: 0956-540X. DOI: 10.1046/j.1365-246x.1999.00891.x.

Dahlen, F. A. and Adam M. Baig (Aug. 2002). "Fréchet Kernels for Body-Wave Amplitudes". In: *Geophysical Journal International* 150.2, pp. 440–466. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.2002.01718.x.

Dahlen, F. A., S.-H. Hung, and Guust Nolet (Apr. 2000). "Fréchet Kernels for Finite-Frequency Traveltimes—I. Theory". In: *Geophysical Journal International* 141.1, pp. 157–174. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.2000.00070.x.

Dahlen, F. A. and Jeroen Tromp (Oct. 1998). *Theoretical Global Seismology*. en. Princeton University Press. ISBN: 978-0-691-00124-1.

Dai, Wei, Yunsong Huang, and Gerard T Schuster (2013). "Least-squares reverse time migration of marine data with frequency-selection encoding". In: *Geophysics* 78.4, S233–S242. DOI: 10.1190/geo2013-0003.1.

Dalton, Colleen A. and Göran Ekström (May 2006). "Global Models of Surface Wave Attenuation". en. In: *Journal of Geophysical Research: Solid Earth* 111.B5. ISSN: 0148-0227. DOI: 10.1029/2005JB003997.

Dalton, Colleen A., Göran Ekström, and Adam M. Dziewonski (June 2009). "Global Seismological Shear Velocity and Attenuation: A Comparison with Experimental

Observations". In: *Earth and Planetary Science Letters* 284.1, pp. 65–75. ISSN: 0012-821X. DOI: 10.1016/j.epsl.2009.04.009.

Dalton, Colleen A. and Ulrich H. Faul (Nov. 2010). "The Oceanic and Cratonic Upper Mantle: Clues from Joint Interpretation of Global Velocity and Attenuation Models". In: *Lithos*. The Lithosphere/Asthenosphere Boundary: Nature, Formation and Evolution 120.1, pp. 160–172. ISSN: 0024-4937. DOI: 10.1016/j.lithos.2010.08.020.

Dalton, Colleen A., Ekström Göran, and Adam M. Dziewonski (Sept. 2008). "The Global Attenuation Structure of the Upper Mantle". In: *Journal of Geophysical Research: Solid Earth* 113.B9. ISSN: 0148-0227. DOI: 10.1029/2007JB005429.

Day, Steven M. and J. Bernard Minster (July 1984). "Numerical Simulation of Attenuated Wavefields Using a Padé Approximant Method". en. In: *Geophysical Journal International* 78.1, pp. 105–118. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1984.tb06474.x.

de la Puente, J., M. Ferrer, M. Hanzich, J. E. Castillo, and J. M. Cela (2014). "Mimetic seismic wave modeling including topography on deformed staggered grids". In: *Geophysics* 79(3), T125–T141. DOI: 10.1190/geo2013-0371.1.

Debayle, Eric, Thomas Bodin, Stéphanie Durand, and Yanick Ricard (Oct. 2020). "Seismic Evidence for Partial Melt below Tectonic Plates". en. In: *Nature* 586.7830, pp. 555–559. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2809-4.

Debayle, Eric, Brian Kennett, and Keith Priestley (Feb. 2005). "Global Azimuthal Seismic Anisotropy and the Unique Plate-Motion Deformation of Australia". en. In: *Nature* 433.7025, pp. 509–512. ISSN: 1476-4687. DOI: 10.1038/nature03247.

Durek, Joseph J. and Göran Ekström (Feb. 1996). "A Radial Model of Anelasticity Consistent with Long-Period Surface-Wave Attenuation". en. In: *Bulletin of the Seismological Society of America* 86.1A, pp. 144–158. ISSN: 0037-1106.

Dziewonski, Adam M. and Don L. Anderson (June 1981). "Preliminary Reference Earth Model". In: *Physics of the Earth and Planetary Interiors* 25.4, pp. 297–356. ISSN: 0031-9201. DOI: 10.1016/0031-9201(81)90046-7.

Dziewonski, Adam M., Bradford H. Hager, and Richard J. O'Connell (1977). "Large-Scale Heterogeneities in the Lower Mantle". en. In: *Journal of Geophysical Research (1896-1977)* 82.2, pp. 239–255. ISSN: 2156-2202. DOI: 10.1029/JB082i002p00239.

Ekström, G., M. Nettles, and A. M. Dziewoński (June 2012). "The Global CMT Project 2004–2010: Centroid-Moment Tensors for 13,017 Earthquakes". en. In: *Physics of the Earth and Planetary Interiors* 200-201, pp. 1–9. ISSN: 0031-9201. DOI: 10.1016/j.pepi.2012.04.002.

Ekström, Göran, Jeroen Tromp, and Erik W. F. Larson (1997). "Measurements and Global Models of Surface Wave Propagation". en. In: *Journal of Geophysical Research: Solid Earth* 102.B4, pp. 8137–8157. ISSN: 2156-2202. DOI: 10.1029/96JB03729.

Ekström, Göran (2011). "A global model of Love and Rayleigh surface wave dispersion and anisotropy, 25-250 s". In: *Geophysical Journal International* 187, 1668–1686.

Endrun, B., T. Meier, S. Lebedev, M. Bohnhoff, G. Stavrakakis, and H.-P. Harjes (2008). "S velocity structure and radial anisotropy in the Aegean region from surface wave dispersion". In: *Geophys. J. Int.* 174, pp. 593–616.

Engdahl, E. Robert, Rob van der Hilst, and Raymond Buland (June 1998). "Global Teleseismic Earthquake Relocation with Improved Travel Times and Procedures for Depth Determination". en. In: *Bulletin of the Seismological Society of America* 88.3, pp. 722–743. ISSN: 0037-1106.

Engquist, B., B. D. Froese, and Y. Yang. (2016). "Optimal transport for seismic full waveform inversion". In: *Communications in Mathematical Sciences* 14.8, pp. 2309–2330.

Fabien-Ouellet, Gabriel, Erwan Gloaguen, and Bernard Giroux (June 2017). "Time Domain Viscoelastic Full Waveform Inversion". en. In: *Geophysical Journal International* 209.3, pp. 1718–1734. ISSN: 0956-540X. DOI: 10.1093/gji/ggx110.

Faul, Ulrich H. and Ian Jackson (May 2005). "The Seismological Signature of Temperature and Grain Size Variations in the Upper Mantle". In: *Earth and Planetary Science Letters* 234.1, pp. 119–134. ISSN: 0012-821X. DOI: 10.1016/j.epsl.2005.02.008.

Ferreira, AMG, JH Woodhouse, K Visser, and J Trampert (2010). "On the robustness of global radially anisotropic surface wave tomography". In: *Journal of Geophysical Research: Solid Earth* 115.B4.

Ferreira, Ana M. G. and John H. Woodhouse (Jan. 2007). "Source, Path and Receiver Effects on Seismic Surface Waves". en. In: *Geophysical Journal International* 168.1, pp. 109–132. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2006.03092.x.

Fichtner, A., H. P. Bunge, and H. Igel (Aug. 2006). "The Adjoint Method in Seismology: I. Theory". In: *Physics of the Earth and Planetary Interiors* 157.1, pp. 86–104. ISSN: 0031-9201. DOI: 10.1016/j.pepi.2006.03.016.

Fichtner, Andreas and Heiner Igel (May 2008). "Efficient Numerical Surface Wave Propagation through the Optimization of Discrete Crustal Models—a Technique Based on Non-Linear Dispersion Curve Matching (DCM)". en. In: *Geophysical Journal International* 173.2, pp. 519–533. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2008.03746.x.

Fichtner, Andreas, Brian L. N. Kennett, Heiner Igel, and Hans-Peter Bunge (Nov. 2008). "Theoretical Background for Continental- and Global-Scale Full-Waveform Inversion in the Time–Frequency Domain". en. In: *Geophysical Journal International* 175.2, pp. 665–685. ISSN: 0956-540X, 1365-246X. DOI: 10.1111/j.1365-246X.2008.03923.x.

Fichtner, Andreas, Brian L. N. Kennett, Heiner Igel, and Hans-Peter Bunge (Dec. 2009). "Full Seismic Waveform Tomography for Upper-Mantle Structure in the Australasian Region Using Adjoint Methods". en. In: *Geophysical Journal International* 179.3, pp. 1703–1725. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2009.04368.x.

Fichtner, Andreas, Erdinc Saygin, Tuncay Taymaz, Paul Cupillard, Yann Capdeville, and Jeannot Trampert (July 2013a). "The Deep Structure of the North Anatolian Fault Zone". en. In: *Earth and Planetary Science Letters* 373, pp. 109–117. ISSN: 0012-821X. DOI: 10.1016/j.epsl.2013.04.027.

Fichtner, Andreas and Jeannot Trampert (May 2011a). "Hessian Kernels of Seismic Data Functionals Based upon Adjoint Techniques". In: *Geophysical Journal International* 185.2, pp. 775–798. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2011.04966.x.

Fichtner, Andreas and Jeannot Trampert (Dec. 2011b). "Resolution Analysis in Full Waveform Inversion: Resolution in Full Waveform Inversion". en. In: *Geophysical Journal International* 187.3, pp. 1604–1624. ISSN: 0956540X. DOI: 10.1111/j.1365-246X.2011.05218.x.

Fichtner, Andreas, Jeannot Trampert, Paul Cupillard, Erdinc Saygin, Tuncay Taymaz, Yann Capdeville, and Antonio Villaseñor (July 2013b). "Multiscale Full Waveform Inversion". en. In: *Geophysical Journal International* 194.1, pp. 534–556. ISSN: 0956-540X. DOI: 10.1093/gji/ggt118.

Fichtner, Andreas and Tristan van Leeuwen (2015). "Resolution Analysis by Random Probing". In: *Journal of Geophysical Research: Solid Earth* 120.8, pp. 5549–5573.

Fichtner, Andreas and Andrea Zunino (2019). "Hamiltonian Nullspace Shuttles". en. In: *Geophysical Research Letters* 46.2, pp. 644–651. ISSN: 1944-8007. DOI: 10.1029/2018GL080931.

Fichtner, Andreas, Andrea Zunino, and Lars Gebraad (Feb. 2019). "Hamiltonian Monte Carlo Solution of Tomographic Inverse Problems". en. In: *Geophysical Journal International* 216.2, pp. 1344–1363. ISSN: 0956-540X. DOI: 10.1093/gji/ggy496.

Fletcher, R. and C. M. Reeves (Jan. 1964). "Function Minimization by Conjugate Gradients". In: *The Computer Journal* 7.2, pp. 149–154. ISSN: 0010-4620. DOI: 10.1093/comjnl/7.2.149.

Forsyth, Donald W. (Oct. 1975). "The Early Structural Evolution and Anisotropy of the Oceanic Upper Mantle". en. In: *Geophysical Journal International* 43.1, pp. 103–162. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1975.tb00630.x.

Frankel, Arthur and John Vidale (Oct. 1992). "A Three-Dimensional Simulation of Seismic Waves in the Santa Clara Valley, California, from a Loma Prieta Aftershock". en. In: *Bulletin of the Seismological Society of America* 82.5, pp. 2045–2074. ISSN: 0037-1106.

French, S. W. and B. A. Romanowicz (Dec. 2014). "Whole-Mantle Radially Anisotropic Shear Velocity Structure from Spectral-Element Waveform Tomography". en. In: *Geophysical Journal International* 199.3, pp. 1303–1327. ISSN: 0956-540X. DOI: 10.1093/gji/ggu334.

Fuchs, Karl (1968). "The Reflection of Spherical Waves from Transition Zones with Arbitrary Depth-Dependent Elastic Moduli and Density". In: *Journal of Physics of the Earth* 16.Special, pp. 27–41. DOI: 10.4294/jpe1952.16.Special_27.

Furumura, T., B. L. N. Kennett, and M. Furumura (Dec. 1998a). "Seismic Wavefield Calculation for Laterally Heterogeneous Whole Earth Models Using the Pseudospectral Method". en. In: *Geophysical Journal International* 135.3, pp. 845–860. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.1998.00682.x.

Furumura, Takashi, B. L. N. Kennett, and Hiroshi Takenaka (Jan. 1998b). "Parallel 3-D Pseudospectral Simulation of Seismic Wave Propagation". In: *GEOPHYSICS* 63.1, pp. 279–288. ISSN: 0016-8033. DOI: 10.1190/1.1444322.

Galis, Martin, Peter Moczo, and J. Kristek (Oct. 2008). "A 3-D Hybrid Finite-Difference—Finite-Element Viscoelastic Modelling of Seismic Wave Motion". en. In: *Geophysical Journal International* 175.1, pp. 153–184. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2008.03866.x.

Gallovič, F., M. Käser, J. Burjánek, and C. Papaioannou (2010). "Three-Dimensional Modeling of near-Fault Ground Motions with Nonplanar Rupture Models and Topography: Case of the 2004 Parkfield Earthquake". en. In: *Journal of Geophysical Research: Solid Earth* 115.B3. ISSN: 2156-2202. DOI: 10.1029/2008JB006171.

Gauthier, Odile, Jean Virieux, and Albert Tarantola (1986). "Two-Dimensional Nonlinear Inversion of Seismic Waveforms: Numerical Results". In: *Geophysics* 51.7, pp. 1387–1403.

Gee, Lind S. and Thomas H. Jordan (1992). "Generalized Seismological Data Functionals". en. In: *Geophysical Journal International* 111.2, pp. 363–390. ISSN: 1365-246X. DOI: 10.1111/j.1365-246X.1992.tb00584.x.

Gerstoft, Peter, Peter M. Shearer, Nick Harmon, and Jian Zhang (2008). "Global P, PP, and PKP Wave Microseisms Observed from Distant Storms". en. In: *Geophysical Research Letters* 35.23. ISSN: 1944-8007. DOI: 10.1029/2008GL036111.

Gilbert, Freeman (1970). "Excitation of the Normal Modes of the Earth by Earthquake Sources". en. In: *Geophysical Journal International* 22.2, pp. 223–226. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1971.tb03593.x.

Godfrey, Karen E., Colleen A. Dalton, Zhitu Ma, Vala Hjörleifsdóttir, and Göran Ekström (June 2019). "A Comparison of Approaches for the Prediction and Inversion of Surface Wave Phase Delays". en. In: *Geophysical Journal International* 217.3, pp. 1496–1514. ISSN: 0956-540X. DOI: 10.1093/gji/ggz096.

Graves, R. (1996). "Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences". In: *Bulletin of the Seismological Society of America* 86, pp. 1091–1106.

Gung, Y. and B. Romanowicz (2004). "Q Tomography of the Upper Mantle Using Three-Component Long-Period Waveforms". In: *Geophysical Journal International* 157.2, pp. 813–830.

Hilst, Rob D. van der and Wim Spakman (1989). "Importance of the Reference Model in Linearized Tomography and Images of Subduction below the Caribbean Plate". en. In: *Geophysical Research Letters* 16.10, pp. 1093–1096. ISSN: 1944-8007. DOI: 10.1029/GL016i010p01093.

Houser, C., G. Masters, P. Shearer, and G. Laske (July 2008). "Shear and Compressional Velocity Models of the Mantle from Cluster Analysis of Long-Period Waveforms". en. In: *Geophysical Journal International* 174.1, pp. 195–212. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2008.03763.x.

Hu, Ge and William Menke (1992). "Formal Inversion of Laterally Heterogeneous Velocity Structure from P-Wave Polarization Data". In: *Geophysical Journal International* 110.1, pp. 63–69.

Huang, Yunsong and Gerard T Schuster (2018). "Full-waveform inversion with multisource frequency selection of marine streamer data". In: *Geophysical Prospecting* 66.7, pp. 1243–1257. DOI: 10.1111/1365-2478.12588.

Hudson, J. A. (1977). "Scattered Waves in Coda of P". In: *Journal of Geophysics-Zeitschrift Fur Geophysik* 43.1-2, pp. 359–374.

Hung, S.-H., F. A. Dahlen, and Guust Nolet (Apr. 2000). "Fréchet Kernels for Finite-Frequency Traveltimes—II. Examples". In: *Geophysical Journal International* 141.1, pp. 175–203. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.2000.00072.x.

Hung, Shu-Huei, Yang Shen, and Ling-Yun Chiao (Aug. 2004). "Imaging Seismic Velocity Structure beneath the Iceland Hot Spot: A Finite Frequency Approach". en. In: *Journal of Geophysical Research: Solid Earth* 109.B8, B08305. ISSN: 2156-2202. DOI: 10.1029/2003JB002889.

Igel, Heiner (Mar. 1999). "Wave Propagation in Three-Dimensional Spherical Sections by the Chebyshev Spectral Method". en. In: *Geophysical Journal International* 136.3, pp. 559–566. ISSN: 0956-540X. DOI: 10.1046/j.1365-246x.1999.00758.x.

Igel, Heiner (2017). *Computational Seismology: A Practical Introduction*. en. Oxford University Press. ISBN: 978-0-19-871740-9.

Igel, Heiner, Peter Mora, and Bruno Riollet (July 1995). "Anisotropic Wave Propagation through Finite-difference Grids". In: *GEOPHYSICS* 60.4, pp. 1203–1216. ISSN: 0016-8033. DOI: 10.1190/1.1443849.

Igel, Heiner and Michael Weber (1995). "SH-Wave Propagation in the Whole Mantle Using High-Order Finite Differences". en. In: *Geophysical Research Letters* 22.6, pp. 731–734. ISSN: 1944-8007. DOI: 10.1029/95GL00312.

Igel, Heiner and Michael Weber (1996). "P-SV Wave Propagation in the Earth's Mantle Using Finite Differences: Application to Heterogeneous Lowermost Mantle Structure". en. In: *Geophysical Research Letters* 23.5, pp. 415–418. ISSN: 1944-8007. DOI: 10.1029/96GL00422.

Ishii, Miaki and Jeroen Tromp (Aug. 1999). "Normal-Mode and Free-Air Gravity Constraints on Lateral Variations in Velocity and Density of Earth's Mantle". en.

In: *Science* 285.5431, pp. 1231–1236. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.285.5431.1231.

Ishii, Miaki and Jeroen Tromp (Apr. 2001). "Even-degree Lateral Variations in the Earth's Mantle Constrained by Free Oscillations and the Free-air Gravity Anomaly". en. In: *Geophysical Journal International* 145.1, pp. 77–96. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2001.00385.x.

Jarchow, Craig M., Rufus D. Catchings, and William J. Lutter (Feb. 1994). "Large-explosive Source, Wide-recording Aperture, Seismic Profiling on the Columbia Plateau, Washington". In: *GEOPHYSICS* 59.2, pp. 259–271. ISSN: 0016-8033. DOI: 10.1190/1.1443588.

Jeffreys, H. and K. E. Bullen (1940). *Seismological Tables*. British Association for the Advancement of.

Kamei, R. and R. G. Pratt (Aug. 2013). "Inversion Strategies for Visco-Acoustic Waveform Inversion". en. In: *Geophysical Journal International* 194.2, pp. 859–884. ISSN: 0956-540X. DOI: 10.1093/gji/ggt109.

Kanamori, Hiroo and Don L. Anderson (1977). "Importance of Physical Dispersion in Surface Wave and Free Oscillation Problems: Review". en. In: *Reviews of Geophysics* 15.1, pp. 105–112. ISSN: 1944-9208. DOI: 10.1029/RG015i001p00105.

Karaoğlu, Haydar and Barbara Romanowicz (Feb. 2018a). "Global Seismic Attenuation Imaging Using Full-Waveform Inversion: A Comparative Assessment of Different Choices of Misfit Functionals". en. In: *Geophysical Journal International* 212.2, pp. 807–826. ISSN: 0956-540X. DOI: 10.1093/gji/ggx442.

Karaoğlu, Haydar and Barbara Romanowicz (June 2018b). "Inferring Global Upper-Mantle Shear Attenuation Structure by Waveform Tomography Using the Spectral Element Method". en. In: *Geophysical Journal International* 213.3, pp. 1536–1558. ISSN: 0956-540X. DOI: 10.1093/gji/ggy030.

Karato, Shun-ichiro (1993). "Importance of Anelasticity in the Interpretation of Seismic Tomography". en. In: *Geophysical Research Letters* 20.15, pp. 1623–1626. ISSN: 1944-8007. DOI: 10.1029/93GL01767.

Käser, Martin and Michael Dumbser (Mar. 2008). "A Highly Accurate Discontinuous Galerkin Method for Complex Interfaces between Solids and Moving Fluids". In: *GEOPHYSICS* 73.3, T23–T35. ISSN: 0016-8033. DOI: 10.1190/1.2870081.

Käser, Martin, P. Martin Mai, and Michael Dumbser (Oct. 2007). "Accurate Calculation of Fault-Rupture Models Using the High-Order Discontinuous Galerkin Method on Tetrahedral Meshes". en. In: *Bulletin of the Seismological Society of America* 97.5, pp. 1570–1586. ISSN: 0037-1106. DOI: 10.1785/0120060253.

Keating, Scott and Kristopher A Innanen (2019). "Parameter crosstalk and modeling errors in viscoacoustic seismic full-waveform inversion". In: *Geophysics* 84.4, R641–R653. DOI: 10.1190/geo2018-0410.1.

Kennett, B. L. N. and E. R. Engdahl (May 1991). "Traveltimes for Global Earthquake Location and Phase Identification". en. In: *Geophysical Journal International* 105.2, pp. 429–465. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1991.tb06724.x.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (July 1995). "Constraints on Seismic Velocities in the Earth from Traveltimes". en. In: *Geophysical Journal International* 122.1, pp. 108–124. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1995.tb03540.x.

Kennett, B. L. N., S. Widiyantoro, and R. D. van der Hilst (1998). "Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle". In: *Journal of Geophysical Research: Solid Earth* 103.B6, pp. 12469–12493. DOI: https://doi.org/10.1029/98JB00150. eprint: https://agupubs.onlinelibrary.wiley.

com/doi/pdf/10.1029/98JB00150. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JB00150.

Kim, Hyoung-Jin, Daisuke Sasaki, Shigeru Obayashi, and Kazuhiro Nakahashi (2001). "Aerodynamic Optimization of Supersonic Transport Wing Using Unstructured Adjoint Method". In: *AIAA Journal* 39.6, pp. 1011–1020. ISSN: 0001-1452. DOI: 10.2514/2.1441.

Kim, YoungHee, Qinya Liu, and Jeroen Tromp (July 2011). "Adjoint Centroid-Moment Tensor Inversions". en. In: *Geophysical Journal International* 186.1, pp. 264–278. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2011.05027.x.

Koelemeijer, Paula, Arwen Deuss, and Jeroen Ritsema (May 2017). "Density Structure of Earth's Lowermost Mantle from Stoneley Mode Splitting Observations". en. In: *Nature Communications* 8.1, p. 15241. ISSN: 2041-1723. DOI: 10.1038/ncomms15241.

Komatitsch, Dimitri, Qinya Liu, Jeroen Tromp, Peter Süss, Christiane Stidham, and John H. Shaw (Feb. 2004). "Simulations of Ground Motion in the Los Angeles Basin Based upon the Spectral-Element Method". en. In: *Bulletin of the Seismological Society of America* 94.1, pp. 187–206. ISSN: 0037-1106. DOI: 10.1785/0120030077.

Komatitsch, Dimitri, Jeroen Ritsema, and Jeroen Tromp (Nov. 2002). "The Spectral-Element Method, Beowulf Computing, and Global Seismology". en. In: *Science* 298.5599, pp. 1737–1742. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1076024.

Komatitsch, Dimitri and Jeroen Tromp (Dec. 1999). "Introduction to the Spectral Element Method for Three-Dimensional Seismic Wave Propagation". en. In: *Geophysical Journal International* 139.3, pp. 806–822. ISSN: 0956-540X. DOI: 10.1046/j.1365-246x.1999.00967.x.

Komatitsch, Dimitri and Jeroen Tromp (May 2002a). "Spectral-Element Simulations of Global Seismic Wave Propagation—I. Validation". en. In: *Geophysical Journal International* 149.2, pp. 390–412. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.2002.01653.x.

Komatitsch, Dimitri and Jeroen Tromp (July 2002b). "Spectral-Element Simulations of Global Seismic Wave Propagation—II. Three-Dimensional Models, Oceans, Rotation and Self-Gravitation". en. In: *Geophysical Journal International* 150.1, pp. 303–318. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.2002.01716.x.

Komatitsch, Dimitri and Jeroen Tromp (July 2003). "A Perfectly Matched Layer Absorbing Boundary Condition for the Second-Order Seismic Wave Equation". en. In: *Geophysical Journal International* 154.1, pp. 146–153. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.2003.01950.x.

Komatitsch, Dimitri and Jean-Pierre Vilotte (Apr. 1998). "The Spectral Element Method: An Efficient Tool to Simulate the Seismic Response of 2D and 3D Geological Structures". en. In: *Bulletin of the Seismological Society of America* 88.2, pp. 368–392. ISSN: 0037-1106.

Komatitsch, Dimitri, Zhinan Xie, Ebru Bozdağ, Elliott Sales de Andrade, Daniel Peter, Qinya Liu, and Jeroen Tromp (Sept. 2016). "Anelastic Sensitivity Kernels with Parsimonious Storage for Adjoint Tomography and Full Waveform Inversion". en. In: *Geophysical Journal International* 206.3, pp. 1467–1478. ISSN: 0956-540X. DOI: 10.1093/gji/ggw224.

Krebs, Jerome R., John E. Anderson, David Hinkley, Ramesh Neelamani, Sunwoong Lee, Anatoly Baumstein, and Martin-Daniel Lacasse (Nov. 2009). "Fast Full-Wavefield Seismic Inversion Using Encoded Sources". In: *Geophysics* 74.6, WCC177–WCC188. ISSN: 0016-8033. DOI: 10.1190/1.3230502.

Krebs, Jerome R, Young Ho Cha, Sunwoong Lee, Pavel Dimitrov, Anoop A Mullur, Nathan J Downey, and Partha S Routh (2018). *Orthogonal source and receiver encoding*. US Patent 10,012,745.

Krischer, Lion, James Smith, Wenjie Lei, Matthieu Lefebvre, Youyi Ruan, Elliott Sales de Andrade, Norbert Podhorszki, Ebru Bozdağ, and Jeroen Tromp (2016). "An Adaptable Seismic Data Format". In: *Geophysical Journal International*, ggw319.

Kristeková, Miriam, Jozef Kristek, Peter Moczo, and Steven M. Day (Oct. 2006). "Misfit Criteria for Quantitative Comparison of Seismograms". en. In: *Bulletin of the Seismological Society of America* 96.5, pp. 1836–1850. ISSN: 0037-1106. DOI: 10.1785/0120060012.

Kurzmann, A., A. Przebindowska, D. Kohn, and T. Bohlen (2013). "Acoustic full waveform tomography in the presence of attenuation: a sensitivity analysis". In: *Geophysical Journal International* 195(2), pp. 985–1000.

Kustowski, B., G. Ekström, and A. M. Dziewoński (June 2008). "Anisotropic Shear-Wave Velocity Structure of the Earth's Mantle: A Global Model". In: *Journal of Geophysical Research: Solid Earth* 113.B6. ISSN: 0148-0227. DOI: 10.1029/2007JB005169.

Landès, Matthieu, Fabien Hubans, Nikolai M. Shapiro, Anne Paul, and Michel Campillo (2010). "Origin of Deep Ocean Microseisms by Using Teleseismic Body Waves". en. In: *Journal of Geophysical Research: Solid Earth* 115.B5. ISSN: 2156-2202. DOI: 10.1029/2009JB006918.

Laske, G. and G. Masters (1996). "Constraints on Global Phase Velocity Maps from Long-Period Polarization Data". en. In: *Journal of Geophysical Research: Solid Earth* 101.B7, pp. 16059–16075. ISSN: 2156-2202. DOI: 10.1029/96JB00526.

Lebedev, Sergei and Rob D. Van Der Hilst (May 2008). "Global Upper-Mantle Tomography with the Automated Multimode Inversion of Surface and S-Wave Forms". en. In: *Geophysical Journal International* 173.2, pp. 505–518. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2008.03721.x.

Lee, E.-J. and P. Chen (2013). "Automating seismic waveform analysis for full 3-D waveform inversions". In: *Geophys. J. Int.* 194, pp. 572–589.

Lehmann, Inge (1936). "P', Publ". In: *Bur. Centr. Seism. Internat. Serie A* 14, pp. 87–115.

Lei, Wenjie et al. (Oct. 2020). "Global Adjoint Tomography—Model GLAD-M25". In: *Geophysical Journal International* 223.1, pp. 1–21. ISSN: 0956-540X. DOI: 10.1093/gji/ggaa253.

Lekić, V. and B. Romanowicz (May 2011). "Inferring Upper-Mantle Structure by Full Waveform Tomography with the Spectral Element Method". en. In: *Geophysical Journal International* 185.2, pp. 799–831. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2011.04969.x.

Lekic, Vedran, Sanne Cottaar, Adam Dziewonski, and Barbara Romanowicz (Dec. 2012). "Cluster Analysis of Global Lower Mantle Tomography: A New Class of Structure and Implications for Chemical Heterogeneity". en. In: *Earth and Planetary Science Letters* 357-358, pp. 68–77. ISSN: 0012-821X. DOI: 10.1016/j.epsl.2012.09.014.

Leng, Kuangdai, Tarje Nissen-Meyer, Martin van Driel, Kasra Hosseini, and David Al-Attar (June 2019). "AxiSEM3D: Broad-Band Seismic Wavefields in 3-D Global Earth Models with Undulating Discontinuities". en. In: *Geophysical Journal International* 217.3, pp. 2125–2146. ISSN: 0956-540X. DOI: 10.1093/gji/ggz092.

Lévêque, Jean-Jacques, Luis Rivera, and Gérard Wittlinger (Oct. 1993). "On the Use of the Checker-Board Test to Assess the Resolution of Tomographic Inversions". en. In: *Geophysical Journal International* 115.1, pp. 313–318. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1993.tb05605.x.

Li, Chang, Robert D. van der Hilst, E. Robert Engdahl, and Scott Burdick (2008). "A New Global Model for P Wave Speed Variations in Earth's Mantle". en. In: *Geochemistry, Geophysics, Geosystems* 9.5. ISSN: 1525-2027. DOI: 10.1029/2007GC001806.

Li, Xiang-Dong and Barbara Romanowicz (June 1995). "Comparison of Global Waveform Inversions with and without Considering Cross-Branch Modal Coupling". en. In: *Geophysical Journal International* 121.3, pp. 695–709. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1995.tb06432.x.

Li, Xiang-Dong and Barbara Romanowicz (Oct. 1996). "Global Mantle Shear Velocity Model Developed Using Nonlinear Asymptotic Coupling Theory". en. In: *Journal of Geophysical Research: Solid Earth* 101.B10, pp. 22245–22272. ISSN: 2156-2202. DOI: 10.1029/96JB01306.

Li, Xiang-Dong and Toshiro Tanimoto (Jan. 1993). "Waveforms of Long-Period Body Waves In A Slightly Aspherical Earth Model". en. In: *Geophysical Journal International* 112.1, pp. 92–102. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1993.tb01439.x.

Liu, Dong C. and Jorge Nocedal (1989). "On the Limited Memory BFGS Method for Large Scale Optimization". In: *Mathematical programming* 45.1-3, pp. 503–528.

Liu, Hsi-Ping, Don L. Anderson, and Hiroo Kanamori (Oct. 1976). "Velocity Dispersion Due to Anelasticity; Implications for Seismology and Mantle Composition". en. In: *Geophysical Journal of the Royal Astronomical Society* 47.1, pp. 41–58. ISSN: 1365-246X. DOI: 10.1111/j.1365-246X.1976.tb01261.x.

Liu, Qiancheng and Daniel Peter (Aug. 2019). "Square-Root Variable Metric Based Elastic Full-Waveform Inversion—Part 2: Uncertainty Estimation". en. In: *Geophysical Journal International* 218.2, pp. 1100–1120. ISSN: 0956-540X. DOI: 10.1093/gji/ggz137.

Liu, Qing et al. (2014). "Hello ADIOS: The Challenges and Lessons of Developing Leadership Class I/O Frameworks". en. In: *Concurrency and Computation: Practice and Experience* 26.7, pp. 1453–1473. ISSN: 1532-0634. DOI: 10.1002/cpe.3125.

Liu, Qinya, Jascha Polet, Dimitri Komatitsch, and Jeroen Tromp (Oct. 2004). "Spectral-Element Moment Tensor Inversions for Earthquakes in Southern California". en. In: *Bulletin of the Seismological Society of America* 94.5, pp. 1748–1761. ISSN: 0037-1106. DOI: 10.1785/012004038.

Lognonné, P. and B. Romanowicz (Aug. 1990). "Modelling of Coupled Normal Modes of the Earth: The Spectral Method". en. In: *Geophysical Journal International* 102.2, pp. 365–395. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1990.tb04472.x.

Love, Augustus Edward Hough (1911). *Some Problems of Geodynamics: Being an Essay to Which the Adams Prize in the University of Cambridge Was Adjudged in 1911*. CUP Archive.

Lu, Yang, Laurent Stehly, and Anne Paul (Aug. 2018). "High-Resolution Surface Wave Tomography of the European Crust and Uppermost Mantle from Ambient Seismic Noise". en. In: *Geophysical Journal International* 214.2, pp. 1136–1150. ISSN: 0956-540X. DOI: 10.1093/gji/ggy188.

Luo, Yang (2012). "Seismic Imaging and Inversion Based on Spectral-Element and Adjoint Methods". en. PhD thesis.

Luo, Yang, Ryan Modrak, and Jeroen Tromp (2014). "Strategies in Adjoint Tomography". en. In: *Handbook of Geomathematics*. Ed. by Willi Freeden, M. Zuhair Nashed, and Thomas Sonar. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–52. ISBN: 978-3-642-27793-1. DOI: 10.1007/978-3-642-27793-1_96-2.

Luo, Yi and Gerard T. Schuster (1991). "Wave-Equation Traveltime Inversion". In: *Geophysics* 56.5, pp. 645–653.

Ma, Shuo, Ralph J. Archuleta, and Pengcheng Liu (Aug. 2004). "Hybrid Modeling of Elastic P-SV Wave Motion: A Combined Finite-Element and Staggered-Grid Finite-Difference Approach". en. In: *Bulletin of the Seismological Society of America* 94.4, pp. 1557–1563. ISSN: 0037-1106. DOI: 10.1785/012003087.

Maeda, Takahiro, Asako Iwaki, Nobuyuki Morikawa, Shin Aoi, and Hiroyuki Fujiwara (Nov. 2016). "Seismic-Hazard Analysis of Long-Period Ground Motion of Megathrust Earthquakes in the Nankai Trough Based on 3D Finite-Difference Simulation". en. In: *Seismological Research Letters* 87.6, pp. 1265–1273. ISSN: 0895-0695. DOI: 10.1785/0220160093.

Maggi, Alessia, Carl Tape, Min Chen, Daniel Chao, and Jeroen Tromp (2009). "An Automated Time-Window Selection Algorithm for Seismic Tomography". In: *Geophysical Journal International* 178.1, pp. 257–281.

Marquering, Henk, F. A. Dahlen, and Guust Nolet (June 1999). "Three-Dimensional Sensitivity Kernels for Finite-Frequency Traveltimes: The Banana-Doughnut Paradox". en. In: *Geophysical Journal International* 137.3, pp. 805–815. ISSN: 0956-540X. DOI: 10.1046/j.1365-246x.1999.00837.x.

Marra, Giuseppe et al. (Aug. 2018). "Ultrastable Laser Interferometry for Earthquake Detection with Terrestrial and Submarine Cables". en. In: *Science* 361.6401, pp. 486–490. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aat4458.

Masters, Guy, Gabi Laske, Harold Bolton, and Adam Dziewonski (2000). "The Relative Behavior of Shear Velocity, Bulk Sound Speed, and Compressional Velocity in the Mantle: Implications for Chemical and Thermal Structure". In: *Washington DC American Geophysical Union Geophysical Monograph Series* 117, pp. 63–87. DOI: 10.1029/GM117p0063.

McNamara, Antoine, Adrien Treuille, Zoran Popović, and Jos Stam (Aug. 2004). "Fluid Control Using the Adjoint Method". In: *ACM Transactions on Graphics* 23.3, pp. 449–456. ISSN: 0730-0301. DOI: 10.1145/1015706.1015744.

Meier, Ueli, Andrew Curtis, and Jeannot Trampert (2007). "Global Crustal Thickness from Neural Network Inversion of Surface Wave Data". In: *Geophysical Journal International* 169.2, pp. 706–722.

Métivier, L., R. Brossier, S. Operto, and J. Virieux (Aug. 2015). "Acoustic Multi-Parameter FWI for the Reconstruction of P-Wave Velocity, Density and Attenuation: Preconditioned Truncated Newton Approach". In: *SEG Technical Program Expanded Abstracts 2015*. SEG Technical Program Expanded Abstracts. Society of Exploration Geophysicists, pp. 1198–1203. DOI: 10.1190/segam2015-5875643.1.

Métivier, Ludovic, Aude Allain, Romain Brossier, Quentin Mérigot, Edouard Oudet, and Jean Virieux (2018). "Optimal transport for mitigating cycle skipping in full waveform inversion: a graph space transform approach". In: *Geophysics* 83.5, R515–R540. DOI: 10.1190/geo2017-0807.1.

Moczo, Peter, Erik Bystrický, Jozef Kristek, José M. Carcione, and Michel Bouchon (Oct. 1997). "Hybrid Modeling of P-SV Seismic Motion at Inhomogeneous Viscoelastic Topographic Structures". en. In: *Bulletin of the Seismological Society of America* 87.5, pp. 1305–1323. ISSN: 0037-1106.

Modrak, Ryan, Jeoren Tromp, and Yanhua Yuan (Sept. 2016). "On the Choice of Material Parameters for Elastic Waveform Inversion". In: *SEG Technical Program Expanded Abstracts 2016*. SEG Technical Program Expanded Abstracts. Society of Exploration Geophysicists, pp. 1115–1119. DOI: 10.1190/segam2016-13970232.1.

Modrak, Ryan and Jeroen Tromp (Sept. 2016). "Seismic Waveform Inversion Best Practices: Regional, Global and Exploration Test Cases". In: *Geophysical Journal International* 206.3, pp. 1864–1889. ISSN: 0956-540X. DOI: 10.1093/gji/ggw202.

Mohorovicic, A. (1909). "Das Beben Vom 8". In: *X, Jahrbuch met. obs. Zagreb* 9, pp. 1–63.

Montagner, J. P. (1986). "Regional Three-Dimensional Structures Using Long-Period Surface Waves". In: *Ann. Geophys* 4.B3, pp. 283–294.

Montagner, Jean-Paul (Mar. 1985). "Seismic Anisotropy of the Pacific Ocean Inferred from Long-Period Surface Waves Dispersion". en. In: *Physics of the Earth and Planetary Interiors* 38.1, pp. 28–50. ISSN: 0031-9201. DOI: 10.1016/0031-9201(85)90120-7.

Montagner, Jean-Paul (1994). "Can Seismology Tell Us Anything about Convection in the Mantle?" en. In: *Reviews of Geophysics* 32.2, pp. 115–137. ISSN: 1944-9208. DOI: 10.1029/94RG00099.

Montagner, Jean-Paul and Don L. Anderson (Apr. 1989). "Petrological Constraints on Seismic Anisotropy". en. In: *Physics of the Earth and Planetary Interiors* 54.1, pp. 82–105. ISSN: 0031-9201. DOI: 10.1016/0031-9201(89)90189-1.

Montagner, Jean-Paul, Daphné-Anne Griot-Pommera, and Jérôme Lavé (2000). "How to Relate Body Wave and Surface Wave Anisotropy?" en. In: *Journal of Geophysical Research: Solid Earth* 105.B8, pp. 19015–19027. ISSN: 2156-2202. DOI: 10.1029/2000JB900015.

Montagner, Jean-Paul and Henri-Claude Nataf (1986). "A Simple Method for Inverting the Azimuthal Anisotropy of Surface Waves". en. In: *Journal of Geophysical Research: Solid Earth* 91.B1, pp. 511–520. ISSN: 2156-2202. DOI: 10.1029/JB091iB01p00511.

Montagner, Jean-Paul and Toshiro Tanimoto (1991). "Global Upper Mantle Tomography of Seismic Velocities and Anisotropies". en. In: *Journal of Geophysical Research: Solid Earth* 96.B12, pp. 20337–20351. ISSN: 2156-2202. DOI: 10.1029/91JB01890.

Monteiller, Vadim, Jean-Luc Got, Jean Virieux, and Paul Okubo (Dec. 2005). "An Efficient Algorithm for Double-Difference Tomography and Location in Heterogeneous Media, with an Application to the Kilauea Volcano". en. In: *Journal of Geophysical Research: Solid Earth* 110.B12, B12306. ISSN: 2156-2202. DOI: 10.1029/2004JB003466.

Montelli, R., G. Nolet, G. Masters, F. A. Dahlen, and S.-H. Hung (Aug. 2004). "Global P and PP Traveltime Tomography: Rays versus Waves". en. In: *Geophysical Journal International* 158.2, pp. 637–654. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2004.02346.x.

Morelli, A. and A. M. Dziewonski (1987). "The Harmonic Expansion Approach to the Retrieval of Deep Earth Structure". en. In: *Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics*. Ed. by Guust Nolet. Seismology and Exploration Geophysics. Dordrecht: Springer Netherlands, pp. 251–274. ISBN: 978-94-009-3899-1. DOI: 10.1007/978-94-009-3899-1_11.

Mosca, I., L. Cobden, A. Deuss, J. Ritsema, and J. Trampert (2012). "Seismic and Mineralogical Structures of the Lower Mantle from Probabilistic Tomography". en. In: *Journal of Geophysical Research: Solid Earth* 117.B6. ISSN: 2156-2202. DOI: 10.1029/2011JB008851.

Mosegaard, Klaus and Albert Tarantola (1995). "Monte Carlo Sampling of Solutions to Inverse Problems". en. In: *Journal of Geophysical Research: Solid Earth* 100.B7, pp. 12431–12447. ISSN: 2156-2202. DOI: 10.1029/94JB03097.

Mulder, W. A. and R.-E. Plessix (2008). "Exploring Some Issues in Acoustic Full Waveform Inversion". en. In: *Geophysical Prospecting* 56.6, pp. 827–841. ISSN: 1365-2478. DOI: 10.1111/j.1365-2478.2008.00708.x.

Nataf, H.-C., I. Nakanishi, and Don L. Anderson (1984). "Anisotropy and Shear-Velocity Heterogeneities in the Upper Mantle". en. In: *Geophysical Research Letters* 11.2, pp. 109–112. ISSN: 1944-8007. DOI: 10.1029/GL011i002p00109.

Nicolas, Adolphe and Nikolas I. Christensen (1987). "Formation of Anisotropy in Upper Mantle Peridotites-A Review". In: *Composition, structure and dynamics of the lithosphere-asthenosphere system* 16, pp. 111–123.

Nicolas, Adolphe and Nikolas I. Christensen (2013). "Formation of Anisotropy in Upper Mantle Peridotites - A Review". en. In: *Composition, Structure and Dynamics of the Lithosphere-Asthenosphere System*. American Geophysical Union (AGU), pp. 111–123. ISBN: 978-1-118-67041-5. DOI: 10.1029/GD016p0111.

Nissen-Meyer, T., M. van Driel, S. C. Stähler, K. Hosseini, S. Hempel, L. Auer, A. Colombi, and A. Fournier (June 2014). "AxiSEM: Broadband 3-D Seismic Wavefields in Axisymmetric Media". English. In: *Solid Earth* 5.1, pp. 425–445. ISSN: 1869-9510. DOI: 10.5194/se-5-425-2014.

Nocedal, J. and S. J. Wright (2006). *Numerical Optimization*. 2nd. Springer.

Nocedal, Jorge (1980). "Updating Quasi-Newton Matrices with Limited Storage". In: *Mathematics of computation* 35.151, pp. 773–782.

Nolet, G. (1987). "Waveform Tomography". en. In: *Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics*. Ed. by Guust Nolet. Seismology and Exploration Geophysics. Dordrecht: Springer Netherlands, pp. 301–322. ISBN: 978-94-009-3899-1. DOI: 10.1007/978-94-009-3899-1_13.

Nolet, Guust, Jos van Trier, and Reina Huisman (1986). "A Formalism for Nonlinear Inversion of Seismic Surface Waves". In: *Geophysical Research Letters* 13.1, pp. 26–29.

Nolet, Guust et al. (Feb. 2019). "Imaging the Galápagos Mantle Plume with an Unconventional Application of Floating Seismometers". En. In: *Scientific Reports* 9.1, p. 1326. ISSN: 2045-2322. DOI: 10.1038/s41598-018-36835-w.

Oldham, Richard Dixon (Feb. 1906). "The Constitution of the Interior of the Earth, as Revealed by Earthquakes". en. In: *Quarterly Journal of the Geological Society* 62.1-4, pp. 456–475. ISSN: 0370-291X. DOI: 10.1144/GSL.JGS.1906.062.01-04.21.

Olsen, Kim B. (1995). "Magnitude 7.75 Earthquake on the San Andreas Fault: Three-Dimensional Ground Motion in Los Angeles". In: *Science* 270, pp. 1628–1632.

Operto, S. and A. Miniussi (June 2018). "On the Role of Density and Attenuation in Three-Dimensional Multiparameter Viscoacoustic VTI Frequency-Domain FWI: An OBC Case Study from the North Sea". en. In: *Geophysical Journal International* 213.3, pp. 2037–2059. ISSN: 0956-540X. DOI: 10.1093/gji/ggy103.

Operto, Stéphane, J. Virieux, J.-X. Dessa, and G. Pascal (2006). "Crustal Seismic Imaging from Multifold Ocean Bottom Seismometer Data by Frequency Domain Full Waveform Tomography: Application to the Eastern Nankai Trough". In: *Journal of Geophysical Research: Solid Earth* 111.B9.

Örsvuran, Rıdvan, Ebru Bozdağ, Ryan Modrak, Wenjie Lei, and Youyi Ruan (Jan. 2020). "Double-Difference Measurements in Global Full-Waveform Inversions". In: *Geophysical Journal International* 220.1, pp. 661–680. ISSN: 0956-540X. DOI: 10.1093/gji/ggz444.

Pageot, Damien, Stéphane Operto, Martin Vallée, Romain Brossier, and Jean Virieux (June 2013). "A Parametric Analysis of Two-Dimensional Elastic Full Waveform Inversion of Teleseismic Data for Lithospheric Imaging". en. In: *Geophysical Journal International* 193.3, pp. 1479–1505. ISSN: 0956-540X. DOI: 10.1093/gji/ggs132.

Park, Jeffrey (1986). "Synthetic Seismograms from Coupled Free Oscillations: Effects of Lateral Structure and Rotation". en. In: *Journal of Geophysical Research: Solid Earth* 91.B6, pp. 6441–6464. ISSN: 2156-2202. DOI: 10.1029/JB091iB06p06441.

Patera, Anthony T (June 1984). "A Spectral Element Method for Fluid Dynamics: Laminar Flow in a Channel Expansion". en. In: *Journal of Computational Physics* 54.3, pp. 468–488. ISSN: 0021-9991. DOI: 10.1016/0021-9991(84)90128-1.

Pelties, Christian, Josep de la Puente, Jean-Paul Ampuero, Gilbert B. Brietzke, and Martin Käser (2012). "Three-Dimensional Dynamic Rupture Simulation with a High-Order Discontinuous Galerkin Method on Unstructured Tetrahedral Meshes". en. In: *Journal of Geophysical Research: Solid Earth* 117.B2. ISSN: 2156-2202. DOI: 10.1029/2011JB008857.

Plessix, R.-E. (Nov. 2006). "A Review of the Adjoint-State Method for Computing the Gradient of a Functional with Geophysical Applications". en. In: *Geophysical Journal International* 167.2, pp. 495–503. ISSN: 0956540X, 1365246X. DOI: 10.1111/j.1365-246X.2006.02978.x.

Poli, P., H. A. Pedersen, and M. Campillo (Feb. 2012). "Emergence of Body Waves from Cross-Correlation of Short Period Seismic Noise". en. In: *Geophysical Journal International* 188.2, pp. 549–558. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2011.05271.x.

Pratt, R. Gerhard, Lianjie Huang, Neb Duric, and Peter Littrup (Mar. 2007). "Sound-Speed and Attenuation Imaging of Breast Tissue Using Waveform Tomography of Transmission Ultrasound Data". In: *Medical Imaging 2007: Physics of Medical Imaging*. Vol. 6510. International Society for Optics and Photonics, 65104S. DOI: 10.1117/12.708789.

Prieto G. A., Lawrence J. F., and Beroza G. C. (July 2009). "Anelastic Earth Structure from the Coherency of the Ambient Seismic Field". In: *Journal of Geophysical Research: Solid Earth* 114.B7. ISSN: 0148-0227. DOI: 10.1029/2008JB006067.

Prieux, V., G. Lambaré, S. Operto, and J. Virieux (June 2013a). "Building Starting Models for Full Waveform Inversion from Wide-Aperture Data by Stereotomography". English. In: *Geophysical Prospecting* 61.sup1, pp. 109–137. ISSN: 1365-2478. DOI: 10.1111/j.1365-2478.2012.01099.x.

Prieux, Vincent, Romain Brossier, Stéphane Operto, and Jean Virieux (Sept. 2013b). "Multiparameter Full Waveform Inversion of Multicomponent Ocean-Bottom-Cable Data from the Valhall Field. Part 1: Imaging Compressional Wave Speed, Density and Attenuation". en. In: *Geophysical Journal International* 194.3, pp. 1640–1664. ISSN: 0956-540X. DOI: 10.1093/gji/ggt177.

Prieux, Vincent, Romain Brossier, Stéphane Operto, and Jean Virieux (Sept. 2013c). "Multiparameter Full Waveform Inversion of Multicomponent Ocean-Bottom-Cable Data from the Valhall Field. Part 2: Imaging Compressive-Wave and Shear-Wave Velocities". en. In: *Geophysical Journal International* 194.3, pp. 1665–1681. ISSN: 0956-540X. DOI: 10.1093/gji/ggt178.

Priolo, Enrico, José M. Carcione, and Géza Seriani (Feb. 1994). "Numerical Simulation of Interface Waves by High-order Spectral Modeling Techniques". In: *The Journal of the Acoustical Society of America* 95.2, pp. 681–693. ISSN: 0001-4966. DOI: 10.1121/1.408428.

Puente, J. de la, J.-P. Ampuero, and M. Käser (2009). "Dynamic Rupture Modeling on Unstructured Meshes Using a Discontinuous Galerkin Method". en. In: *Journal of Geophysical Research: Solid Earth* 114.B10. ISSN: 2156-2202. DOI: 10.1029/2008JB006271.

Reid, F. J. L., J. H. Woodhouse, and H. J. Van Heijst (June 2001). "Upper Mantle Attenuation and Velocity Structure from Measurements of Differential S Phases". en. In: *Geophysical Journal International* 145.3, pp. 615–630. ISSN: 0956-540X. DOI: 10.1046/j.1365-246x.2001.01395.x.

Ren, Zhiming, Yang Liu, and Qunshan Zhang (May 2014). "Multiscale Viscoacoustic Waveform Inversion with the Second Generation Wavelet Transform and Adaptive Time–Space Domain Finite-Difference Method". en. In: *Geophysical Journal International* 197.2, pp. 948–974. ISSN: 0956-540X. DOI: 10.1093/gji/ggu024.

Resovsky, Joseph, Jeannot Trampert, and R. D. Van der Hilst (Feb. 2005). "Error Bars for the Global Seismic Q Profile". In: *Earth and Planetary Science Letters* 230.3, pp. 413–423. ISSN: 0012-821X. DOI: 10.1016/j.epsl.2004.12.008.

Rickers, Florian, Andreas Fichtner, and Jeannot Trampert (July 2012). "Imaging Mantle Plumes with Instantaneous Phase Measurements of Diffracted Waves". en. In: *Geophysical Journal International* 190.1, pp. 650–664. ISSN: 0956-540X, 1365-246X. DOI: 10.1111/j.1365-246X.2012.05515.x.

Rickers, Florian, Andreas Fichtner, and Jeannot Trampert (Apr. 2013). "The Iceland–Jan Mayen Plume System and Its Impact on Mantle Dynamics in the North Atlantic Region: Evidence from Full-Waveform Inversion". en. In: *Earth and Planetary Science Letters* 367, pp. 39–51. ISSN: 0012-821X. DOI: 10.1016/j.epsl.2013.02.022.

Ritsema, J., A. Deuss, H. J. van Heijst, and J. H. Woodhouse (Mar. 2011). "S40RTS: A Degree-40 Shear-Velocity Model for the Mantle from New Rayleigh Wave Dispersion, Teleseismic Traveltime and Normal-Mode Splitting Function Measurements". en. In: *Geophysical Journal International* 184.3, pp. 1223–1236. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2010.04884.x.

Ritsema, Jeroen, Luis A. Rivera, Dimitri Komatitsch, Jeroen Tromp, and Hendrik-Jan van Heijst (2002). "Effects of Crust and Mantle Heterogeneity on PP/P and SS/S Amplitude Ratios". en. In: *Geophysical Research Letters* 29.10, pp. 72–1–72–4. ISSN: 1944-8007. DOI: 10.1029/2001GL013831.

Ritzwoller, Michael H., Nikolai M. Shapiro, Mikhail P. Barmin, and Anatoli L. Levshin (2002). "Global Surface Wave Diffraction Tomography". en. In: *Journal of Geophysical Research: Solid Earth* 107.B12, ESE 4–1–ESE 4–13. ISSN: 2156-2202. DOI: 10.1029/2002JB001777.

Robertsson, Johan O. A., Joakim O. Blanch, and William W. Symes (Sept. 1994). "Viscoelastic Finite-difference Modeling". In: *GEOPHYSICS* 59.9, pp. 1444–1456. ISSN: 0016-8033. DOI: 10.1190/1.1443701.

Romanowicz, B. (1995). "A Global Tomographic Model of Shear Attenuation in the Upper Mantle". In: *Journal of Geophysical Research: Solid Earth* 100.B7, pp. 12375–12394.

Romanowicz, Barbara (1990). "The Upper Mantle Degree 2: Constraints and Inferences from Global Mantle Wave Attenuation Measurements". en. In: *Journal of Geophysical Research: Solid Earth* 95.B7, pp. 11051–11071. ISSN: 2156-2202. DOI: 10.1029/JB095iB07p11051.

Romanowicz, Barbara (1994a). "Anelastic Tomography: A New Perspective on Upper Mantle Thermal Structure". In: *Earth and Planetary Science Letters* 128.3-4, pp. 113–121.

Romanowicz, Barbara (July 1994b). "On the Measurement of Anelastic Attenuation Using Amplitudes of Low-Frequency Surface Waves". In: *Physics of the Earth and Planetary Interiors* 84.1, pp. 179–191. ISSN: 0031-9201. DOI: 10.1016/0031-9201(94)90040-X.

Romanowicz, Barbara, Li-Wei Chen, and Scott W. French (Jan. 2020). "Accelerating Full Waveform Inversion via Source Stacking and Cross-Correlations". en. In: *Geophysical Journal International* 220.1, pp. 308–322. ISSN: 0956-540X. DOI: 10.1093/gji/ggz437.

Ruan, Youyi, Wenjie Lei, Ryan Modrak, Rıdvan Örsvuran, Ebru Bozdağ, and Jeroen Tromp (Nov. 2019). "Balancing Unevenly Distributed Data in Seismic Tomography: A Global Adjoint Tomography Example". en. In: *Geophysical Journal International* 219.2, pp. 1225–1236. ISSN: 0956-540X. DOI: 10.1093/gji/ggz356.

Ruan, Youyi and Ying Zhou (Apr. 2010). "The Effects of 3-D Anelasticity (Q) Structure on Surface Wave Phase Delays". en. In: *Geophysical Journal International* 181.1, pp. 479–492. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2010.04514.x.

Ruan, Youyi and Ying Zhou (May 2012). "The Effects of 3-D Anelasticity (Q) Structure on Surface Wave Amplitudes". en. In: *Geophysical Journal International* 189.2, pp. 967–983. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2011.05356.x.

Saenger, Erik H. and Thomas Bohlen (Mar. 2004). "Finite-Difference Modeling of Viscoelastic and Anisotropic Wave Propagation Using the Rotated Staggered Grid". en. In: *Geophysics* 69.2, pp. 583–591. ISSN: 0016-8033. DOI: 10.1190/1.1707078.

Sager, Korbinian, Christian Boehm, Laura Ermert, Lion Krischer, and Andreas Fichtner (2020). "Global-Scale Full-Waveform Ambient Noise Inversion". en. In: *Journal of Geophysical Research: Solid Earth* 125.4, e2019JB018644. ISSN: 2169-9356. DOI: 10.1029/2019JB018644.

Sambridge, Malcolm (Aug. 1999). "Geophysical Inversion with a Neighbourhood Algorithm—I. Searching a Parameter Space". en. In: *Geophysical Journal International* 138.2, pp. 479–494. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.1999.00876.x.

Sandhu, Gursharan Yash, Erik West, Cuiping Li, Olivier Roy, and Neb Duric (Mar. 2017). "3D Frequency-Domain Ultrasound Waveform Tomography Breast Imaging". In: *Medical Imaging 2017: Ultrasonic Imaging and Tomography*. Vol. 10139. International Society for Optics and Photonics, p. 1013909. DOI: 10.1117/12.2254399.

Schaeffer, A. J. and S. Lebedev (July 2013). "Global Shear Speed Structure of the Upper Mantle and Transition Zone". en. In: *Geophysical Journal International* 194.1, pp. 417–449. ISSN: 0956-540X. DOI: 10.1093/gji/ggt095.

Schaeffer, A. J. and S. Lebedev (Sept. 2014). "Imaging the North American Continent Using Waveform Inversion of Global and USArray Data". In: *Earth and Planetary Science Letters*. Special Issue on USArray Science 402, pp. 26–41. ISSN: 0012-821X. DOI: 10.1016/j.epsl.2014.05.014.

Schaeffer, A. J., S. Lebedev, and T. W. Becker (Nov. 2016). "Azimuthal Seismic Anisotropy in the Earth's Upper Mantle and the Thickness of Tectonic Plates". en. In: *Geophysical Journal International* 207.2, pp. 901–933. ISSN: 0956-540X. DOI: 10.1093/gji/ggw309.

Schweitzer, J. (Feb. 2001). "HYPOSAT – An Enhanced Routine to Locate Seismic Events". en. In: *pure and applied geophysics* 158.1, pp. 277–289. ISSN: 1420-9136. DOI: 10.1007/PL00001160.

Selby, Neil D. and John H. Woodhouse (2002). "The Q Structure of the Upper Mantle: Constraints from Rayleigh Wave Amplitudes". en. In: *Journal of Geophysical Research: Solid Earth* 107.B5, ESE 5–1–ESE 5–11. ISSN: 2156-2202. DOI: 10.1029/2001JB000257.

Sengupta, Mrinal K. and M. Nafi Toksöz (1977). "The Amplitudes of P Waves and Magnitude Corrections for Deep Focus Earthquakes". en. In: *Journal of Geophysical Research (1896-1977)* 82.20, pp. 2971–2980. ISSN: 2156-2202. DOI: 10.1029/JB082i020p02971.

Seriani, Géza (Oct. 1998). "3-D Large-Scale Wave Propagation Modeling by Spectral Element Method on Cray T3E Multiprocessor". en. In: *Computer Methods in Applied Mechanics and Engineering*. Exterior Problems of Wave Propagation 164.1, pp. 235–247. ISSN: 0045-7825. DOI: 10.1016/S0045-7825(98)00057-7.

Shapiro, Nikolai M., Michel Campillo, Laurent Stehly, and Michael H. Ritzwoller (Mar. 2005). "High-Resolution Surface-Wave Tomography from Ambient Seismic

Noise". en. In: *Science* 307.5715, pp. 1615–1618. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1108339.

Sieminski, Anne, Jean-Jacques Lévêque, and Eric Debayle (2004). "Can Finite-Frequency Effects Be Accounted for in Ray Theory Surface Wave Tomography?" en. In: *Geophysical Research Letters* 31.24. ISSN: 1944-8007. DOI: 10.1029/2004GL021402.

Sieminski, Anne, Qinya Liu, Jeannot Trampert, and Jeroen Tromp (Oct. 2007a). "Finite-Frequency Sensitivity of Body Waves to Anisotropy Based upon Adjoint Methods". en. In: *Geophysical Journal International* 171.1, pp. 368–389. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2007.03528.x.

Sieminski, Anne, Qinya Liu, Jeannot Trampert, and Jeroen Tromp (2007b). "Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods". en. In: *Geophysical Journal International* 168(3).1, 1153–1174. DOI: 10.1111/j.1365-246X.2006.03261.x.

Silver, Paul G. and W. Winston Chan (1991). "Shear Wave Splitting and Subcontinental Mantle Deformation". en. In: *Journal of Geophysical Research: Solid Earth* 96.B10, pp. 16429–16454. ISSN: 2156-2202. DOI: 10.1029/91JB00899.

Simmons, N. A., S. C. Myers, G. Johannesson, and E. Matzel (2012). "LLNL-G3Dv3: Global P Wave Tomography Model for Improved Regional and Teleseismic Travel Time Prediction". en. In: *Journal of Geophysical Research: Solid Earth* 117.B10. ISSN: 2156-2202. DOI: 10.1029/2012JB009525.

Simons, Frederik J., Guust Nolet, Paul Georgief, Jeff M. Babcock, Lloyd A. Regier, and Russ E. Davis (2009). "On the Potential of Recording Earthquakes for Global Seismic Tomography by Low-Cost Autonomous Instruments in the Oceans". en. In: *Journal of Geophysical Research: Solid Earth* 114.B5. ISSN: 2156-2202. DOI: 10.1029/2008JB006088.

Sirgue, Laurent and R. Gerhard Pratt (Jan. 2004). "Efficient Waveform Inversion and Imaging: A Strategy for Selecting Temporal Frequencies". In: *Geophysics* 69.1, pp. 231–248. ISSN: 0016-8033. DOI: 10.1190/1.1649391.

Sladen, A., D. Rivet, J. P. Ampuero, L. De Barros, Y. Hello, G. Calbris, and P. Lamare (Dec. 18, 2019). "Distributed Sensing of Earthquakes and Ocean-Solid Earth Interactions on Seafloor Telecom Cables". In: *Nature Communications* 10.1 (1), p. 5777. ISSN: 2041-1723. DOI: 10.1038/s41467-019-13793-z. URL: https://www.nature.com/articles/s41467-019-13793-z (visited on 06/11/2021).

Smith, Martin L and FA Dahlen (1973). "The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium". In: *Journal of Geophysical Research* 78.17, pp. 3321–3333.

Spakman, Wim (Nov. 1991). "Delay-Time Tomography of the Upper Mantle below Europe, the Mediterranean, and Asia Minor". en. In: *Geophysical Journal International* 107.2, pp. 309–332. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1991.tb00828.x.

Sukhovich, Alexey, Jean-Olivier Irisson, Frederik J. Simons, Anthony Ogé, Yann Hello, Anne Deschamps, and Guust Nolet (2011). "Automatic Discrimination of Underwater Acoustic Signals Generated by Teleseismic P-Waves: A Probabilistic Approach". en. In: *Geophysical Research Letters* 38.18. ISSN: 1944-8007. DOI: 10.1029/2011GL048474.

Tago, J., V. M. Cruz-Atienza, J. Virieux, V. Etienne, and F. J. Sánchez-Sesma (2012). "A 3D Hp-Adaptive Discontinuous Galerkin Method for Modeling Earthquake Dynamics". en. In: *Journal of Geophysical Research: Solid Earth* 117.B9. ISSN: 2156-2202. DOI: 10.1029/2012JB009313.

Talagrand, Olivier and Philippe Courtier (1987). "Variational Assimilation of Meteorological Observations With the Adjoint Vorticity Equation. I: Theory". en. In:

Quarterly Journal of the Royal Meteorological Society 113.478, pp. 1311–1328. ISSN: 1477-870X. DOI: 10.1002/qj.49711347812.

Tanimoto, Toshiro and Don L. Anderson (1985). "Lateral Heterogeneity and Azimuthal Anisotropy of the Upper Mantle: Love and Rayleigh Waves 100–250 s". en. In: *Journal of Geophysical Research: Solid Earth* 90.B2, pp. 1842–1858. ISSN: 2156-2202. DOI: 10.1029/JB090iB02p01842.

Tape, Carl (2009). "Seismic Tomography of Southern California Using Adjoint Methods". PhD thesis. California Institute of Technology. DOI: https://thesis.library.caltech.edu/1656/10/09_CHT2009_Ch2.pdf.

Tape, Carl, Qinya Liu, Alessia Maggi, and Jeroen Tromp (2009). "Adjoint Tomography of the Southern California Crust". In: *Science* 325.5943, pp. 988–992.

Tape, Carl, Qinya Liu, Alessia Maggi, and Jeroen Tromp (Jan. 2010). "Seismic Tomography of the Southern California Crust Based on Spectral-Element and Adjoint Methods". In: *Geophysical Journal International* 180.1, pp. 433–462. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2009.04429.x.

Tape, Carl, Qinya Liu, and Jeroen Tromp (Mar. 2007). "Finite-Frequency Tomography Using Adjoint Methods-Methodology and Examples Using Membrane Surface Waves". en. In: *Geophysical Journal International* 168.3, pp. 1105–1129. ISSN: 0956540X, 1365246X. DOI: 10.1111/j.1365-246X.2006.03191.x.

Tarantola, Albert (1984). "Inversion of Seismic Reflection Data in the Acoustic Approximation". In: *Geophysics* 49.8, pp. 1259–1266.

Tarrass, I., L. Giraud, and P. Thore (2011). "New curvilinear scheme for elastic wave propagation in presence of curved topography". In: *Geophysical Prospecting* 59.5, pp. 889–906. DOI: 10.1111/j.1365-2478.2011.00972.x.

Tessmer, Ekkehart, David Kessler, Dan Kosloff, and Alfred Behle (June 1992). "Multi-Domain Chebyshev-Fourier Method for the Solution of the Equations of Motion of Dynamic Elasticity". en. In: *Journal of Computational Physics* 100.2, pp. 355–363. ISSN: 0021-9991. DOI: 10.1016/0021-9991(92)90241-P.

Thomas, Ch, H. Igel, M. Weber, and F. Scherbaum (May 2000). "Acoustic Simulation of P-Wave Propagation in a Heterogeneous Spherical Earth: Numerical Method and Application to Precursor Waves to PKPdf". en. In: *Geophysical Journal International* 141.2, pp. 307–320. ISSN: 0956-540X. DOI: 10.1046/j.1365-246x.2000.00079.x.

Thomas, Christine and Gabi Laske (Feb. 2015). "D″ Observations in the Pacific from PLUME Ocean Bottom Seismometer Recordings". en. In: *Geophysical Journal International* 200.2, pp. 851–862. ISSN: 0956-540X. DOI: 10.1093/gji/ggu441.

Thomsen, Leon (Oct. 1986). "Weak Elastic Anisotropy". In: *GEOPHYSICS* 51.10, pp. 1954–1966. ISSN: 0016-8033. DOI: 10.1190/1.1442051.

Thrastarson, Solvi, Martin van Driel, Lion Krischer, Christian Boehm, Michael Afanasiev, Dirk-Philip van Herwaarden, and Andreas Fichtner (June 2020). "Accelerating Numerical Wave Propagation by Wavefield Adapted Meshes. Part II: Full-Waveform Inversion". en. In: *Geophysical Journal International* 221.3, pp. 1591–1604. ISSN: 0956-540X. DOI: 10.1093/gji/ggaa065.

Toshinawa, Takumi and Tatsuo Ohmachi (Aug. 1992). "Love-Wave Propagation in a Three-Dimensional Sedimentary Basin". en. In: *Bulletin of the Seismological Society of America* 82.4, pp. 1661–1677. ISSN: 0037-1106.

Trampert, Jeannot, Frédéric Deschamps, Joseph Resovsky, and Dave Yuen (Oct. 2004). "Probabilistic Tomography Maps Chemical Heterogeneities Throughout the Lower Mantle". en. In: *Science* 306.5697, pp. 853–856. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1101996.

Trampert, Jeannot and Hendrik Jan van Heijst (May 2002). "Global Azimuthal Anisotropy in the Transition Zone". en. In: *Science* 296.5571, pp. 1297–1299. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1070264.

Trampert, Jeannot and Jesper Spetzler (Feb. 2006). "Surface Wave Tomography: Finite-Frequency Effects Lost in the Null Space". en. In: *Geophysical Journal International* 164.2, pp. 394–400. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2006.02864.x.

Trampert, Jeannot and John Woodhouse (2003). "Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s". In: *Geophysical Journal International* 154, 154–165.

Trampert, Jeannot and John H. Woodhouse (1995). "Global Phase Velocity Maps of Love and Rayleigh Waves between 40 and 150 Seconds". In: *Geophysical Journal International* 122.2, pp. 675–690.

Tromp, Jeroen and Etienne Bachmann (Sept. 2019). "Source Encoding for Adjoint Tomography". en. In: *Geophysical Journal International* 218.3, pp. 2019–2044. ISSN: 0956-540X. DOI: 10.1093/gji/ggz271.

Tromp, Jeroen, Yang Luo, Shravan Hanasoge, and Daniel Peter (Nov. 2010a). "Noise Cross-Correlation Sensitivity Kernels". en. In: *Geophysical Journal International* 183.2, pp. 791–819. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2010.04721.x.

Tromp, Jeroen, Carl Tape, and Qinya Liu (Dec. 2005). "Seismic Tomography, Adjoint Methods, Time Reversal and Banana-Doughnut Kernels: Seismic Tomography, Adjoint Methods, Time Reversal and Banana-Doughnut Kernels". en. In: *Geophysical Journal International* 160.1, pp. 195–216. ISSN: 0956540X, 1365246X. DOI: 10.1111/j.1365-246X.2004.02453.x.

Tromp, Jeroen et al. (Oct. 2010b). "Near Real-Time Simulations of Global CMT Earthquakes". en. In: *Geophysical Journal International* 183.1, pp. 381–389. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2010.04734.x.

Tsvankin, Ilya, James Gaiser, Vladimir Grechka, Mirko van der Baan, and Leon Thomsen (Sept. 2010). "Seismic Anisotropy in Exploration and Reservoir Characterization: An Overview". In: *GEOPHYSICS* 75.5, 75A15–75A29. ISSN: 0016-8033. DOI: 10.1190/1.3481775.

Valentine, Andrew P. and John H. Woodhouse (2010a). "Approaches to Automated Data Selection for Global Seismic Tomography". In: *Geophysical Journal International* 182.2, pp. 1001–1012.

Valentine, Andrew P. and John H. Woodhouse (Feb. 2010b). "Reducing Errors in Seismic Tomography: Combined Inversion for Sources and Structure". en. In: *Geophysical Journal International* 180.2, pp. 847–857. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2009.04452.x.

van Driel, M., C. Boehm, L. Krischer, and M. Afanasiev (June 2020). "Accelerating Numerical Wave Propagation Using Wavefield Adapted Meshes. Part I: Forward and Adjoint Modelling". en. In: *Geophysical Journal International* 221.3, pp. 1580–1590. ISSN: 0956-540X. DOI: 10.1093/gji/ggaa058.

van Heijst, Hendrik Jan and John Woodhouse (Nov. 1997). "Measuring Surface-Wave Overtone Phase Velocities Using a Mode-Branch Stripping Technique". en. In: *Geophysical Journal International* 131.2, pp. 209–230. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1997.tb01217.x.

van Herwaarden, Dirk Philip, Christian Boehm, Michael Afanasiev, Solvi Thrastarson, Lion Krischer, Jeannot Trampert, and Andreas Fichtner (May 2020). "Accelerated Full-Waveform Inversion Using Dynamic Mini-Batches". en. In: *Geophysical Journal International* 221.2, pp. 1427–1438. ISSN: 0956-540X. DOI: 10.1093/gji/ggaa079.

Vinnik, L. P., L. I. Makeyeva, A. Milev, and A. Yu Usenko (Dec. 1992). "Global Patterns of Azimuthal Anisotropy and Deformations in the Continental Mantle". en. In: *Geophysical Journal International* 111.3, pp. 433–447. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1992.tb02102.x.

Virieux, J. and S. Operto (Nov. 2009). "An Overview of Full-Waveform Inversion in Exploration Geophysics". en. In: *Geophysics* 74.6, WCC1–WCC26. ISSN: 0016-8033, 1942-2156. DOI: 10.1190/1.3238367.

Virieux, Jean (Nov. 1984). "SH-Wave Propagation in Heterogeneous Media: Velocity-stress Finite-difference Method". In: *GEOPHYSICS* 49.11, pp. 1933–1942. ISSN: 0016-8033. DOI: 10.1190/1.1441605.

Virieux, Jean (Apr. 1986). "P-SV Wave Propagation in Heterogeneous Media: Velocity-stress Finite-difference Method". In: *GEOPHYSICS* 51.4, pp. 889–901. ISSN: 0016-8033. DOI: 10.1190/1.1442147.

Virieux, Jean and Raul Madariaga (Apr. 1982). "Dynamic Faulting Studied by a Finite Difference Method". en. In: *Bulletin of the Seismological Society of America* 72.2, pp. 345–369. ISSN: 0037-1106.

Visser, K., Jeannot Trampert, and B. L. N. Kennett (2008). "Global Anisotropic Phase Velocity Maps for Higher Mode Love and Rayleigh Waves". In: *Geophysical Journal International* 172.3, pp. 1016–1032.

Waldhauser, Felix and William L. Ellsworth (2000). "A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California". In: *Bulletin of the Seismological Society of America* 90.6, pp. 1353–1368.

Wang, Zheng and F. A. Dahlen (Dec. 1995). "Validity of Surface-Wave Ray Theory on a Laterally Heterogeneous Earth". en. In: *Geophysical Journal International* 123.3, pp. 757–773. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1995.tb06888.x.

Warren, Linda M. and Peter M. Shearer (2002). "Mapping Lateral Variations in Upper Mantle Attenuation by Stacking P and PP Spectra". en. In: *Journal of Geophysical Research: Solid Earth* 107.B12, ESE 6–1–ESE 6–11. ISSN: 2156-2202. DOI: 10.1029/2001JB001195.

Welch, P. (June 1967). "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging over Short, Modified Periodograms". In: *IEEE Transactions on Audio and Electroacoustics* 15.2, pp. 70–73. ISSN: 0018-9278. DOI: 10.1109/TAU.1967.1161901.

Widmer, Rudolf, Guy Masters, and Freeman Gilbert (Mar. 1991). "Spherically Symmetric Attenuation within the Earth from Normal Mode Data". en. In: *Geophysical Journal International* 104.3, pp. 541–553. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1991.tb05700.x.

Woodhouse, J. H. and T. P. Girnius (Mar. 1982). "Surface Waves and Free Oscillations in a Regionalized Earth Model". en. In: *Geophysical Journal International* 68.3, pp. 653–673. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1982.tb04921.x.

Woodhouse, J. H. and Y. K. Wong (Dec. 1986). "Amplitude, Phase and Path Anomalies of Mantle Waves". en. In: *Geophysical Journal International* 87.3, pp. 753–773. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1986.tb01970.x.

Woodhouse, John H. and Adam M. Dziewonski (1984). "Mapping the Upper Mantle: Three-Dimensional Modeling of Earth Structure by Inversion of Seismic Waveforms". In: *Journal of Geophysical Research: Solid Earth* 89.B7, pp. 5953–5986.

Wu, R. and K. Aki (Apr. 1985). "Scattering Characteristics of Elastic Waves by an Elastic Heterogeneity". In: *Geophysics* 50.4, pp. 582–595. ISSN: 0016-8033. DOI: 10.1190/1.1441934.

Yang, Pengliang, Romain Brossier, Ludovic Métivier, and Jean Virieux (Oct. 2016). "A review on the systematic formulation of 3D multiparameter full waveform inversion in viscoelastic medium". In: *Geophysical Journal International* 207.1, pp. 129–149. DOI: 10.1093/gji/ggw262.

Yuan, Yanhua O., Ebru Bozdağ, Caio Ciardelli, Fuchun Gao, and F. J. Simons (May 2020). "The Exponentiated Phase Measurement, and Objective-Function Hybridization for Adjoint Waveform Tomography". en. In: *Geophysical Journal International* 221.2, pp. 1145–1164. ISSN: 0956-540X. DOI: 10.1093/gji/ggaa063.

Yuan, Yanhua O., Frederik J. Simons, and Ebru Bozdağ (2015). "Multiscale Adjoint Waveform Tomography for Surface and Body Waves". In: *Geophysics* 80.5, R281–R302.

Yuan, Yanhua O., Frederik J. Simons, and Jeroen Tromp (2016). "Double-Difference Adjoint Seismic Tomography". In: *Geophysical Journal International* 206.3, pp. 1599–1618.

Zhang, Haijiang and Clifford H. Thurber (Oct. 2003). "Double-Difference Tomography: The Method and Its Application to the Hayward Fault, California". en. In: *Bulletin of the Seismological Society of America* 93.5, pp. 1875–1889. ISSN: 0037-1106. DOI: 10.1785/0120020190.

Zhang, Haijiang, Clifford H. Thurber, David Shelly, Satoshi Ide, Gregory C. Beroza, and Akira Hasegawa (Apr. 2004). "High-Resolution Subducting-Slab Structure beneath Northern Honshu, Japan, Revealed by Double-Difference Tomography". en. In: *Geology* 32.4, pp. 361–364. ISSN: 0091-7613. DOI: 10.1130/G20261.2.

Zhang, Qingchen, Weijian Mao, Hui Zhou, Hongjing Zhang, and Yangkang Chen (Sept. 2018). "Hybrid-domain simultaneous-source full waveform inversion without crosstalk noise". In: *Geophysical Journal International* 215.3, pp. 1659–1681. ISSN: 0956-540X. DOI: 10.1093/gji/ggy366. eprint: https://academic.oup.com/gji/article-pdf/215/3/1659/27719104/ggy366.pdf. URL: https://doi.org/10.1093/gji/ggy366.

Zhang, Qingchen, Hui Zhou, Qingqing Li, Hanming Chen, and Jie Wang (2016). "Robust Source-Independent Elastic Full-Waveform Inversion in the Time Domain". In: *Geophysics* 81.2, R29–R44.

Zhao, Dapeng, T. Yanada, A. Hasegawa, N. Umino, and Wei Wei (Aug. 2012). "Imaging the Subducting Slabs and Mantle Upwelling under the Japan Islands". en. In: *Geophysical Journal International* 190.2, pp. 816–828. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2012.05550.x.

Zhou, Ying (Sept. 2009). "Surface-Wave Sensitivity to 3-D Anelasticity". en. In: *Geophysical Journal International* 178.3, pp. 1403–1410. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2009.04230.x.

Zhou, Ying, F. A. Dahlen, and Guust Nolet (July 2004). "Three-dimensional Sensitivity Kernels for Surface Wave Observables". en. In: *Geophysical Journal International* 158.1, pp. 142–168. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2004.02324.x.

Zhu, Hejun, Ebru Bozdağ, Thomas S. Duffy, and Jeroen Tromp (Nov. 2013). "Seismic Attenuation beneath Europe and the North Atlantic: Implications for Water in the Mantle". In: *Earth and Planetary Science Letters* 381, pp. 1–11. ISSN: 0012-821X. DOI: 10.1016/j.epsl.2013.08.030.

Zhu, Hejun, Ebru Bozdağ, Daniel Peter, and Jeroen Tromp (2012). "Structure of the European Upper Mantle Revealed by Adjoint Tomography". In: *Nature Geoscience* 5.7, pp. 493–498.

Zhu, Hejun, Ebru Bozdağ, and Jeroen Tromp (Apr. 2015). "Seismic Structure of the European Upper Mantle Based on Adjoint Tomography". en. In: *Geophysical Journal International* 201.1, pp. 18–52. ISSN: 0956-540X. DOI: 10.1093/gji/ggu492.

Zhu, Hejun, Jidong Yang, and Xueyan Li (2020). "Azimuthal Anisotropy of the North American Upper Mantle Based on Full Waveform Inversion". en. In: *Journal of Geophysical Research: Solid Earth* 125.2, e2019JB018432. ISSN: 2169-9356. DOI: 10.1029/2019JB018432.