MONITORING AND FORECASTING OF SEISMIC WAVEFIELDS IN THE SUBSURFACE

MONITORING AND FORECASTING OF SEISMIC WAVEFIELDS IN THE SUBSURFACE

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op maandag 4 januari 2021 om 12:30 uur

door

Joeri BRACKENHOFF

Master of Science in Toegepaste Geofysica,
Technische Universiteit Delft, Delft, Nederland,
geboren te Delft, Nederland.

Dit proefschrift is goedgekeurd door de

Promotor: Prof. dr. ir. C.P.A. Wapenaar

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. C.P.A. Wapenaar,	Technische Universiteit Delft
Onafhankelijke leden:	
Prof. dr. A. Fichtner,	ETH Zürich
Prof. dr. L.G. Evers,	Technische Universiteit Delft
Dr. H. Paulsen,	Universiteit van Utrecht
Prof. dr. ir. E.C. Slob,	Technische Universiteit Delft
Prof. dr. J.A. Trampert,	Universiteit van Utrecht
Dr. ir. D.J. Verschuur,	Technische Universiteit Delft

Dit werk heeft financiële steun gekregen van het Horizon 2020 onderzoeks en innovatie programma van de Europese Unie: European Research Council (toelage nummer 742703).

Keywords: Marchenko, Homogeneous Green's function, induced seismicity

Printed by: Gildeprint

Front & Back: Green's function contaminated by anti-symmetric artifacts

ISBN 978-94-6419-105-9

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

For my family

CONTENTS

SUMMARY

Monitoring seismic wavefields caused by induced seismicity in the subsurface is a difficult process. Ideally, it requires physical receivers in the subsurface, which is unpractical. Frequently, only measurements at the surface of the Earth are available, which give a limited amount of information about the subsurface. One way to improve the monitoring of the subsurface is through the use of virtual sources and receivers, which are not physically present but are created from the measured reflection data at the surface. This can be achieved through the use of the classical homogeneous Green's representation, however, this method requires two Green's functions measured on an enclosing boundary, which is an unrealistic requirement. Instead, a single-sided representation of the homogeneous Green's function can be used, where a focusing function, which is a wavefield that focuses from a single-sided boundary to a focal position in the subsurface without artifacts related to the internal multiples, is employed together with a Green's function. To obtain the Green's function and focusing function that are needed for this representation, the Marchenko method is used. This method employs reflection data, without free-surface multiples, at the surface of the Earth and an estimation of the first arrival, which can be modeled in a macro velocity model.

To test whether induced seismicity in the real subsurface can be monitored using the single-sided representation, synthetic data are first considered, which include a synthetic reflection response and macro velocity model. The Marchenko method is used in combination with these data to obtain the focusing functions and Green's functions that are required for the homogeneous Green's function representations. The classical representation and the single-sided representation of the homogeneous Green's function employ the Green's functions and focusing functions to obtain the homogeneous Green's function of the medium. The homogeneous Green's function is visualized by creating snapshots of the homogeneous Green's function and these snapshots are compared to a directly modeled reference wavefield. This demonstrates that the classical representation, when applied to data at an open acquisition boundary, yields significant artifacts in the results, while the single-sided representation obtains accurate results. It is also shown that the radiation pattern of a double-couple source can be included in the retrieval of the homogeneous Green's function. The synthetic reflection data are truncated by limiting the offsets and sampling distance and applying attenuation to simulate field conditions. These truncations show that the single-sided homogeneous Green's function contains artifacts and lacks physical events if the reflection data are not ideal. 2D field reflection data and a macro velocity model from the Vøring basin are considered and pre-processed to account for these truncations. The classical and the single-sided homogeneous Green's function representation are both applied to the field data and the results show that the retrieval of the homogeneous Green's function is possible for 2D field data using point sources while employing the single-sided representation. The results of the classical representation contain a large amount of errors. It is also shown

that a homogeneous Green's function can be retrieved that has a virtual source with a double-couple radiation pattern.

Next, the application of the single-sided representation is considered in greater detail. The representation is used to forecast a wavefield in the subsurface as well as to monitor a wavefield in the subsurface. For the monitoring of the wavefield, it is assumed that a physical source in the subsurface causes a wavefield which is measured at the surface of the Earth. The Marchenko method is used to create virtual receivers inside the subsurface, which are used in combination with the physical measurement in the single-sided representation. This is a one-step process, because the Marchenko method is only used to create the virtual receivers. The single-sided representation of the homogeneous Green's function requires the source wavelet to be symmetric in time, which is unlikely for physical sources. Hence, a different single-sided representation can be used, which retrieves the causal Green's function and does not require a symmetric source wavelet. The single-sided representation of the causal Green's function can retrieve a majority of the correct events, however, the results contain anti-symmetric artifacts when the physical source is located above the virtual receiver. To forecast a wavefield in the subsurface, given a specific source configuration, the single-sided representation of the homogeneous Green's function can be used. In this case, a two-step process is applied, where both the source and the receiver in the subsurface are created by the Marchenko method and are therefore both virtual. After the homogeneous Green's function is obtained, it can be convolved with a non-symmetric wavelet. To demonstrate the difference between the one-step monitoring process and the two-step forecasting process, 2D synthetic reflection data are utilized. For the source configuration, a rupture plane is considered, which is modeled by superposing and time-shifting point sources, which contain a double-couple radiation pattern and are all scaled differently to simulate the heterogeneity of the rupture plane. The total wavefield created by this rupture plane is monitored using the single-sided representation of the causal Green's function. There are anti-symmetric artifacts present in the result, related to each point source, however, the correct wavefield is retrieved above the shallowest source location and below this source location after the first arrivals of all sources. The single-sided representation of the homogeneous Green's function is applied to forecast a virtual rupture plane, by retrieving the homogeneous Green's function for each source separately. The retrieved homogeneous Green's functions are transformed to causal Green's functions, shifted in time and superposed to forecast the total wavefield, which is free of the anti-symmetric artifacts at any depth. Both the monitoring approach and the forecasting approach are tested on 2D field data and the retrieved wavefields show similar results as were seen when the synthetic data were used. When the total wavefield is forecasted, there are no anti-symmetric artifacts present and when the wavefield is monitored, there are artifacts, however, they are only present in part of the result, below the sources before and during the first arrival of each source.

To test the application of the single-sided representation in 3D, a 3D implementation of the Marchenko method is required. The implementation is straightforward from a theoretical standpoint, as the surface integrals are performed over two dimensions instead of just one. The practical implementation is more difficult, however. The Marchenko method requires that the reflection data are well sampled in both space and

time for sources and receivers, hence, the 3D reflection data are of a large size. As a result, not only a large amount of storage space is required, but the loading time of the reflection data is high, both of which are unpractical for efficient computation. We limit these problems by pre-transforming the reflection data to the frequency domain and compressing the data using floating point arrays, which reduces the storage space and loading time. Two datasets are considered, one modeled in a simple four layer model and the other in a subsection of the complex 3D Overthrust model. For both models, a Green's function inside the medium is retrieved, using a first arrival in the Marchenko method that was modeled in the exact medium, and compared to a reference Green's function that was directly modeled. The results for both models are accurate for the single Green's function. Next, imaging is performed for the models, however, instead of modeling the first arrivals, they are estimated using an Eikonal solver, because the modeling time of all the first arrivals is too high. The results of the imaging using the Marchenko method are compared to the results of conventional imaging, which demonstrates that artifacts, related to the internal multiples, are attenuated.

The 3D implementation of the Marchenko method is used to retrieve the Green's functions and focusing functions in 3D using 3D synthetic reflection data modeled in the Overhtrust model. The classical homogeneous Green's function representation and the single-sided representation of the causal Green's function and the homogeneous Green's function are all applied using these data, for three different combinations of a virtual source and a virtual receiver. The results are compared to a directly modeled wavefield, which shows that the result obtained by using the classical representation is contaminated by artifacts and lacks physical events. The result of the single-sided representation of the causal Green's function contains anti-symmetric artifacts related to the focusing function when the virtual receiver is located below the virtual source. The result of the single-sided representation of the homogeneous Green's function shows a good match to the reference result. The single-sided representation of the homogeneous Green's function is also applied using an Eikonal solver to obtain the first arrival that is required for the Marchenko method. The homogeneous Green's function that is obtained in this way shows a small decrease in quality for the result, however, this approach is more computationally feasible. The single-sided representation is used in combination with the Eikonal solver to retrieve a large amount of virtual receivers, so that the propagation of the wavefield in the subsurface can be visualized in time through the use of snapshots. This reveals that the part of the wavefield that is traveling at angles that are close to the normal of the surface is retrieved properly, while the part of the wavefield that is traveling at greater angles to the normal is reconstructed with less accuracy. This lack of proper retrieval is caused by the limited aperture of the reflection data. A rupture plane in 3D is considered and constructed in a similar way as is done for the 2D synthetic data. Point sources are used to model wavefields, which are time-shifted and superposed, however, to further represent the heterogeneity of the rupture plane, each wavefield is modeled using an unique causal wavelet. Both monitoring, using the single-sided causal Green's function representation, and forecasting, using the single-sided homogeneous Green's function representation, are performed on the rupture plane configuration. The two-step forecasting approach yields accurate results, for a given distribution of sources. The one-step monitoring approach retrieves accurate results above the shallowest source lo-

cation, however, the result contains artifacts at the locations below the shallowest source, before and during the first arrival of each source.

SAMENVATTING

Het monitoren van golfvelden die veroorzaakt worden door geïnduceerde seismiciteit is een moeilijk proces. Idealiter worden er fysieke ontvangers gebruikt om het golfveld in de ondergrond te meten, maar dit is onpraktisch. Meestal zijn er alleen metingen beschikbaar aan het aardoppervlak, en deze metingen geven beperkte informatie over de ondergrond. Het monitoren van de ondergrond kan worden versterkt met behulp van virtuele bronnen en ontvangers. Dit zijn bronnen en ontvangers die niet fysiek aanwezig zijn, maar gemaakt worden van gemeten reflectie data aan het aardoppervlak. Het maken van deze virtuele metingen kan gedaan worden met de klassieke representatie van de homogene Green's functie. Dit vereist echter wel dat er twee Green's functies gemeten zijn op een omsluitende grens, wat een onrealistische verwachting is. In plaats van de klassieke representatie van de homogene Green's functie kan een enkelzijdige representatie worden toegepast, die een focusserende functie samen met een Green's functie gebruikt. Een focusserende functie is een golfveld dat van een enkelzijdige grens focust naar een focus punt in de ondergrond zonder artefacten die gerelateerd zijn aan interne meervoudige reflecties. De Green's functie en focusserende functie die nodig zijn voor deze representatie kunnen verkregen worden met behulp van de Marchenko methode. Deze methode gebruikt reflectie data, die geen meervoudige reflecties van het vrije oppervlak bevatten en gemeten zijn aan het aardoppervlak, samen met een schatting van de eerste aankomst van een golfveld, die gemodelleerd kan worden in een macro snelheidsmodel.

Synthetische data worden gebruikt om te testen of de enkelzijdige representatie toegepast kan worden op veld data. De synthetische data bevatten een reflectie dataset en een macro snelheidsmodel. De Marchenko methode wordt samen met deze data gebruikt om Green's functies en focusserende functies te bepalen, die gebruikt kunnen worden voor de homogene Green's functie representaties. De klassieke en enkelzijdige representatie van de homogene Green's functie worden beiden gebruikt om de homogene Green's functie in het medium te bepalen. Deze homogene Green's functies worden gevisualiseerd door middel van momentopnames en deze momentopnames worden vergeleken met momentopnames van een direct gemodelleerd referentie golfveld. Dit toont aan dat de klassieke representatie veel fouten bevat als het gebruikt wordt op een enkelzijdige grens in plaats van een omsluitende grens en dat de enkelzijdige representatie nauwkeurig het golfveld weet te bepalen. Er wordt ook aangetoond dat het radiatie patroon van een dubbelkoppel bron kan worden meegenomen in het verkrijgen van de homogene Green's functie. De synthetische reflectie data worden verslechterd door aanpassingen te maken in de meetafstand en monstering van de data en door energieverlies in de data aan te brengen, waarmee veldcondities worden benaderd. Deze verslechteringen tonen aan dat de enkelzijdige representatie van de homogene Green's functie fouten bevat en fysieke delen van het golfveld mist als de data niet ideaal zijn. 2D veld data met een reflectie dataset en een macro snelheidsmodel worden bewerkt

om te compenseren voor deze verslechteringen. De klassieke en enkelzijdige representatie van de homogene Green's functie worden beiden toegepast in combinatie met deze data. De resultaten laten zien dat het mogelijk is om de homogene Green's functie te verkrijgen met gebruik van 2D veld data en de enkelzijdige representatie als puntbronnen worden gebruikt. De resulaten van de klassieke representatie bevatten grote fouten. Er wordt ook aangetoond dat een homogene Green's functie kan worden verkregen die een virtuele bron met een dubbelkoppel radiatie patroon bevat.

Vervolgens wordt de enkelzijdige representatie in groter detail bekeken. De enkelzijdige representatie wordt gebruikt om het golfveld in de ondergrond te voorspellen en te monitoren. Voor het monitoren van het golfveld wordt aangenomen dat een fysieke bron in de ondergrond een golfveld veroorzaakt dat gemeten wordt aan het aardoppervlak. De Marchenko methode wordt gebruikt om virtuele ontvangers in de ondergrond te creëren, die samen met de fysieke metingen gebruikt worden in de enkelzijdige representatie. Dit is een een-staps proces, omdat de Marchenko methode alleen wordt gebruikt om de virtuele ontvangers te maken. De enkelzijdige representatie van de homogene Green's functie heeft als eis dat het bronsignaal symmetrisch in tijd is, wat onwaarschijnlijk is voor fysieke bronnen. Daarom wordt een andere enkelzijdige representatie gebruikt, die een causale Green's functie verkrijgt en niet de eis heeft dat het bronsignaal symmetrisch moet zijn. De enkelzijdige representatie van de Green's functie kan een groot deel van het golfveld in de ondergrond bepalen, maar bevat antisymmetrische fouten als de fysieke bron zich boven de virtuele ontvanger bevindt. Voor het voorspellen van het golfveld in de ondergrond kan de enkelzijdige representatie van de homogene Green's functie worden gebruikt, als een specifieke bronopstelling wordt gebruikt. In dit geval is er sprake van een twee-staps proces omdat zowel de bron als de ontvanger virtueel zijn en door gebruik van de Marchenko methode worden verkregen. Als de homogene Green's functie is bepaald kan deze worden geconvolueerd met een niet symmetrisch bronsignaal. Het verschil tussen deze twee aanpakken worden gedemonstreerd met behulp van 2D synthetische reflectie data. Als bronopstelling wordt een breukvlak gebruikt, dat geconstrueerd wordt door middel van superpositie van puntbronnen die in de tijd verschoven worden. Al deze bronnen bevatten een dubbelkoppel radiatie patroon en zijn anders geschaald om de heterogeniteit van het breukvlak te simuleren. Het totale golfveld dat door dit breukvlak wordt veroorzaakt wordt gemonitord met behulp van de enkelzijdige representatie van de causale Green's functie. Er zijn antisymmetrische fouten aanwezig die gerelateerd zijn aan iedere puntbron, maar het correcte golfveld wordt verkregen op iedere diepte boven de ondiepste bron. Onder deze diepte zijn de fouten alleen aanwezig voor en tijdens de eerste aankomst van iedere bron. De enkelzijdige toepassing van de homogene Green's functie wordt gebruikt om het golfveld van een virtueel breukvlak te voorspellen. Dit wordt gedaan door eerst de homogene Green's functie gerelateerd aan iedere puntbron te maken en deze dan te veranderen in een causale Green's functie, dit te verschuiven in tijd en tenslotte alle golfvelden te superposeren om het total golfveld te maken. Dit totale golfveld is vrij van fouten op iedere diepte. Zowel de toepassing van het monitoren als van het voorspellen van het golfveld worden gebruikt op 2D veld data. De resulaten zijn vergelijkbaar met die van de synthetische data. Als het totale golfveld wordt voorspeld zijn er geen assymetrische fouten en als het golfveld wordt gemonitord dan zijn deze fouten wel aanwezig, maar alleen in een deel van het resultaat, namelijk

onder de bronnen voor en tijdens de eerste aankomst van iedere bron.

Voordat er getest kan worden of de enkelzijdige representatie in 3D kan worden gebruikt is eerst een 3D implementatie van de Marchenko method nodig. De uitbreiding van de implementatie van 2D naar 3D is in theorie niet lastig, aangezien de integralen nu over twee dimensies worden uitgevoerd in plaats van maar een. De praktische implementatie is echter een stuk ingewikkelder. Voor de Marchenko methode is het belangrijk dat de reflectie data fijn gemonsterd zijn in zowel ruimte als tijd, wat betekent dat de grootte van de dataset zeer hoog is. Hierdoor is niet alleen veel opslag ruimte voor de data nodig, maar kost het ook veel tijd om de data te laden, wat niet praktisch is voor een efficiënte berekening. Deze problemen worden beperkt door de data voor het laden naar het frequentie domein te veranderen en deze data te comprimeren met behulp van een zwevende punt compressie algorithme. Hierdoor neemt zowel de data grootte als de laadtijd af. Twee datasets worden met deze methode getest, een die in een simpel, vlak vierlaags model is gemodelleerd en een andere die in een deel van het complexe Overthrust model is gemodelleerd. In beide modelen wordt een Green's functie verkregen, met behulp van een eerste aankomst van het golfveld die in het exacte medium is gemodelleerd. De Green's functies worden vergeleken met een referentie resultaat en dit toont aan dat de verkregen Green's functies nauwkeurig genoeg zijn. Hierna wordt beeldvorming van de modelen toegepast met behulp van de Marchenko methode. De eerste aankomsten die nodig zijn voor de Marchenko method worden in dit geval bepaald met een Eikonale methode, omdat het modelleren van deze aankomsten te lang zou duren. De beeldvorming die wordt verkregen door de Marchenko method wordt vergeleken met conventionele methodes, wat aantoont dat de resulaten van de Marchenko method beter zijn, omdat fouten die veroorzaakt worden door interne meervoudige reflecties worden verwijderd.

De 3D implementatie van de Marchenko methode wordt gebruikt om 3D Green's functies en 3D focusserende functies te verkrijgen uit reflectie data die in het Overthrust gemodelleerd zijn. De klassieke representatie van de homogene Green's functie en de enkelzijdige representaties van de causale en homogene Green's functie worden allemaal toegepast met behulp van deze data. Iedere representatie wordt getest met drie verschillende configuraties van een virtuele bron en een virtuele ontvanger. Deze resulaten worden vergeleken met een direct gemodelleerd golfveld, wat aantoont dat de resultaten van de klassieke representatie vol zitten met fouten en fysieke delen van het golfveld missen. De resultaten van de enkelzijdige representatie van de causale Green's functie bevatten antisymmetrische fouten die gerelateerd zijn aan de focusserende functie als de virtuele ontvanger zich onder de virtuele bron bevindt. De resulaten van de enkelzijdige representatie van de homogene Green's functie zijn vergelijkbaar met de referentie resultaten. De enkelzijdige representatie van de homogene Green's functie wordt ook gebruikt met golfvelden die verkregen zijn met een Marchenko methode waarin de eerste aankomst is geschat met een Eikonale methode. Het resultaat van deze manier is minder nauwkeurig dan wanneer de eerste aankomst gemodelleerd wordt, maar het berekenen van dit resultaat is een stuk efficiënter. De enkelzijdige representatie wordt gebruikt in combinatie met de Eikonale methode om een groot aantal virtuele ontvangers te creëren. De beweging van het golfveld in de ondergrond wordt gevisualiseerd met behulp van momentopnames. Dit toont aan dat het deel van het golfveld dat beweegt in een

grote hoek ten opzichte van de normaal van het aardoppervlak niet geheel wordt gereconstrueerd. Het deel van het golfveld dat beweegt in een kleine hoek ten opzichte van deze normaal wordt wel goed gereconstrueerd. Deze berperking wordt bepaald door de gelimiteerde apertuur van de reflectie data. Het experiment met een breukvlak als bronopstelling wordt herhaald, maar ditmaal met 3D synthetische reflectie data in plaats van 2D synthetische reflectie data. Puntbronnen worden gebruikt om golvelden te modelleren, die worden verschoven en gesuperposeerd. Ieder golfveld wordt gemodelleerd met een uniek causaal bronsignaal. Met behulp van deze data wordt het golfveld van het breukvlak in de ondergrond zowel gemonitord met behulp van de enkelzijdige representatie van de causale Green's functie als voorspeld met behulp van de enkelzijdige representatie van de homogene Green's functie. De twee-staps voorspelling methode verkrijgt nauwkeurige resulaten, als de bronopstelling bekend is. De een-staps monitoring methode verkrijgt nauwkeurige resulaten in de dieptes boven de ondiepste bron, maar bevat fouten onder die diepte, voor en tijdens de eerste aankomst van iedere bron.

1

INTRODUCTION

In the past few centuries the human population on the planet Earth has increased considerably. Due to this large number of people there is an ever growing demand for resources, varying from basic necessities such as food and water, to more advanced needs, like proper shelter and transportation. While the Earth can provide many natural resources to satisfy these demands, obtaining these resources can increase the likelihood of so-called geohazards. Floods, climate change and earthquakes are among the many geohazards that humanity faces during its existence on the planet. Earthquakes form a significant kind of geohazard that can cause major damage to structures by creating seismic waves that propagate to the surface of the Earth [1]. The term seismicity describes the occurrences of Earthquakes, as well as its mechanisms, magnitude and geographical location, and is useful for determining the earthquake activity in a certain area. While earthquakes can occur naturally, studies have shown that a large amount of seismicity is induced, or in other words, caused by human activity, for example, during the recovery of hydrocarbons [2]. Because the demand for natural resources was and still is high, it is vital to understand what causes induced seismicity and how it can be possibly prevented.

Measuring and processing induced seismicity has seen significant developments in the past few decades [3]. To measure induced seismicity, arrays of geophones are used to record the seismic events in an area. These measurements can then be processed and studied to monitor active seismicity [4] or to forecast the occurence of seismicity [5]. In areas where human activity may cause induced seismicity, these measurements are often employed in a traffic light system (TLS) [6]. The TLS determines whether the induced seismicity passes certain levels, and when this happens, activities that potentially induce earthquakes are halted. While this is useful to prevent damaging activities, if the monitoring of the induced seismicity is inaccurate it can also prevent the application of useful techniques, such as geothermal systems [7]. To properly monitor induced seismicity, it is vital that the location and the source mechanism of the induced seismicity is resolved as good as possible. The source mechanism is often described by the moment tensor [8], which can describe the radiation pattern of a variety of source mechnisms,

such as faulting [9]. To improve on the determination of the moment tensor and location of the source, accurate recordings of the induced seismicity are required, which can be achieved by either employing downhole arrays [10], dense recording arrays [11] or, ideally, a combination of the two. However, while downhole arrays can provide additional insights, they are very expensive to create and maintain.

An alternative to using only the physical measurements is the use of virtual measurements, that is to say, data that are not recorded by a physical receiver, but rather obtained through advanced seismic processing techniques. Depending on the technique, virtual sources and/or receivers can be created anywhere inside the medium of interest, so virtual receivers could be created close to the source of the induced seismicity. One approach to create virtual sources is time-reversal mirroring [12]. This method assumes that if the Green's function; i.e. the impulse response of a medium; is measured on a boundary that encloses the medium where the source of the Green's function is located, the Green's function can be time-reversed and injected from the boundary. This wavefield will propagate towards the original source location and focus, thereby creating a virtual source at this focal location. In order for this method to function, it needs to be actually possible to inject the time-reversed Green's function into the physical medium, instead of simulating this numerically by injecting the time-reversed Green's function into a model of the medium. This physical injection approach is popular in the field of ultrasound [13, 14], however, for geophysical applications it is not as often used.

An alternative, similar approach is homogeneous Green's function retrieval. A homogeneous Green's function is the superposition of a Green's function with its time-reversal. The classical representation was derived by Porter [15] and has been applied for inverse source problems [16], inverse scattering methods [17], seismic imaging [18] and seismic holography [19], among others. Effectively, the representation states that if two Green's functions are measured on an enclosing boundary, they can be cross-correlated and integrated along the boundary to obtain the response of the medium between the source locations of the Green's functions. The advantage of this approach is that no medium information is required, instead all the information is created from the measured data. The principles used in this approach formed a basis for Seismic Interferometry, which can be used to obtain virtual sources [20–22] or virtual receivers [23]. However, the classical homogeneous Green's function representation relies on the assumption that the boundary around the sources is a closed boundary. In the case of seismic monitoring, this is hard to achieve as it is most common to have recording arrays on the surface of the Earth, hence, at a single side of the medium. The classical representation can still be applied using only a single-sided boundary, which will produce relatively accurate results if the overburden above the sources is very smooth. However, if the medium contains strong impedance contrasts, the homogeneous Green's function that is obtained will lack desired events, especially the downward propagating part of the wavefield, and contain significant artifacts related to the internal multiples of the medium. Because of these errors, using the classical representation for the monitoring and forecasting of induced seismicity would create a large amount of uncertainty in the results. To avoid some of the errors that are created by using the time-reversed Green's function, only the time-reversed direct arrival of the Green's function could be used in the representation, in combination with the full causal Green's function [24]. While this can decrease the

amount of artifacts caused by the internal multiples, it will not reconstruct the missing desired part of the homogeneous Green's function.

To overcome the limitations of the classical homogeneous Green's function representation, an alternative single-sided representation can be employed. This representation is designed to work with the single-sided open boundary [25] and makes use of one Green's function and a different wavefield, the so-called focusing function, rather than a time-reversed Green's function. A focusing function is a special wavefield that focuses from the single-sided boundary to a point in the medium without artifacts caused by the internal multiples. Estimating a focusing function is not a straightforward process, however. The single-sided focusing function is defined as the inverse of the transmission response that is truncated below the focal location of the wavefield [26]. If one wants to obtain the transmission response, physical receivers inside the Earth are required, which is unpractical. A better alternative is to obtain the focusing function through the use of the Marchenko method, which is based on the Marchenko equation of quantum mechanics by Marchenko [27]. The method is an advanced way to estimate the Green's function and focusing function from reflection data, that contain no free-surface multiples, at the surface of the Earth. Some of the original principles of single-sided focusing were shown by Rose [28], Broggini *et al.* [29] and Slob *et al.* [30], for 1D media. Wapenaar *et al.* [31], Broggini *et al.* [32] and Behura *et al.* [33] further developed the method for 2D and 3D media. They showed that when the focusing function is convolved with single-sided reflection data at the surface of the Earth, the result is a superposition of the Green's function and the focusing function that are related to a focal position inside the medium. The equation that describes this is a Green's function representation. If the medium of interest is acoustic, the two functions are separated in time except for the direct arrival of both wavefields, which overlay each other in time. By applying a temporal muting operator, the Green's function can be removed from the result and only the focusing function remains. In this way, an equation with only one unknown remains, namely the focusing function. This Marchenko-type equation can be solved using either an iterative scheme [34] or an inversion [35] and only requires a first estimation of the focusing function. For this purpose, often the direct arrival of the wavefield is employed, which can be estimated from a macro velocity model. After the focusing function has been obtained, it can be used in the Green's function representation to obtain the Green's function. Furthermore, the Green's function representation can also be decomposed to relate the upgoing and downgoing Green's function to the upgoing and downgoing focusing function. Similarly, the Marchenko equation can be decomposed into the coupled Marchenko equations so that the decomposed focusing functions can be obtained, which can be used in the decomposed Green's function representation to obtain the decomposed Green's functions [36]. All wavefields, the Green's functions and focusing functions, that are required for the single-sided representation of the homogeneous Green's function can be obtained through a 2D or 3D implementation of the Marchenko equations and the Green's function representations.

The Marchenko method has the advantage that a virtual source or receiver can be created at any point in the medium of interest, as long as this is covered by the aperture and recording length of the reflection data. The method has the disadvantage that evanescent waves are ignored and that the medium of interest is assumed to be lossless.

While this makes it challenging to apply the method to field data, the method has been successfully applied in the field for the purpose of imaging in both 2D [37–40] and 3D [41]. Furthermore, it has been shown that the method can be employed in a variety of ways and schemes. The method has been applied to elastic reflection data [42–44], to remove the internal multiples from reflection data [38, 45], to work with a lossy medium [46], to work with plane-waves [47], to obtain reflection data that only contain primaries [48], to work with reflection data that contain free-surface multiples [49, 50] or to handle media with very thin bedding [51, 52], among others. Wapenaar *et al.* [25] showed that the Marchenko method can be used in combination with the single-sided representation of the homogeneous Green's function to create a virtual source-receiver pair at any point in the medium and demonstrated this on numerical data.

The main interest of this thesis is the possibility of using the Marchenko method and the single-sided representation of the homogeneous Green's function for the purpose of monitoring and forecasting of wavefields related to induced seismicity. While the single-sided representation has been shown to work with 2D numerical data, previous work has not demonstrated that the method can be applied to 2D field data. Furthermore, while a 2D acquisition of the reflection data in a 3D medium can be achieved, this setup will introduce errors due to the fact that the 3D effects are not properly handled [53]. As such, the 2D application of the method will not suffice for complex 3D media. The goal of this thesis is to determine if the single-sided representation can be applied for realistic situations involving induced seismicity. The thesis contains the following chapters:

- Chapter 2: *Virtual sources and receivers in the real Earth*
  This chapter considers the application of the single-sided representation to a field dataset containing 2D reflection data and an interpreted velocity model using the Marchenko method. The single-sided representation is compared to the classical representation for point sources. Aside from an isotropic point source, the response from a point source that has a double-couple radiation pattern is considered.

- Chapter 3: *Monitoring of induced distributed double-couple sources using Marchenko-based virtual receivers*
  This chapter considers the difference between monitoring and forecasting the induced seismicity using the single-sided representation for the homogeneous Green's function and the Marchenko method. In the monitoring application, only the focusing functions are obtained through the use of the Marchenko method and the seismicity is measured. In the forecasting application, the seismicity is not measured, but obtained through the Marchenko method as well. Both approaches are demonstrated on numerical data and field data.

- Chapter 4: *Implementation of the 3D Marchenko method*
  In this chapter, the implementation of the Marchenko method in 3D is considered. This includes not only the theoretical implementation, but also the practical aspects, relating to data size and compute times. The method is used to retrieve the Green's functions inside a simple and complex 3D model and to create images of these two models.

- Chapter 5: *3D virtual seismology*
  In this chapter, the 3D implementation of the single-sided homogeneous Green's function representation is considered, which is applied using the 3D Marchenko implementation of the previous chapter. Both the monitoring and the forecasting of induced seismicity are demonstrated on 3D numerical data for a strongly heterogeneous rupture plane.

- Chapter 6: *Conclusions and recommendations*
  In this chapter, conclusions are drawn based on the results of the other chapters. Recommendations are given based on the results to further develop the ideas and techniques that are described in this thesis.

- Appendices
  In Appendix A, a method is shown to redatum the focusing functions from the surface of the Earth to locations in the subsurface, which can help to remove artifacts from the causal Green's function retrieval.
  In Appendix B, the single-sided homogeneous Green's function representation is used to retrieve the homogeneous Green's function using reflection data that were measured in a lab experiment.

2

VIRTUAL SOURCES AND RECEIVERS IN THE REAL EARTH

To enhance monitoring of the subsurface, virtual sources and receivers inside the subsurface can be created from seismic reflection data at the surface of the Earth using the Marchenko method. The response between these virtual sources and receivers can be obtained through the use of homogeneous Green's function retrieval. A homogeneous Green's function is a superposition of a Green's function and its time-reversal. The main aim of this chapter is to obtain accurate homogeneous Green's functions from field data. Classical homogeneous Green's function retrieval requires an unrealistic enclosing recording surface, however, by using a recently proposed single-sided retrieval scheme, this requirement can be avoided. We first demonstrate the principles of using the single-sided representation on synthetic data and show that different source signatures can be taken into account. Because the Marchenko method is sensitive to recording limitations of the reflection data, we study five cases of recording limitations with synthetic data and demonstrate their effects on the final result. Finally, the method is demonstrated on a pre-processed field dataset which fulfills the requirements for applying the single-sided Green's function retrieval scheme. The scheme has the potential to be used in future applications, such as source localization.

This chapter was published as J. Brackenhoff, J. Thorbecke, & K. Wapenaar, *Virtual sources and receivers in the real Earth: Considerations for practical applications.* Journal of Geophysical Research: Solid Earth, **124**, 11802-11821 (2019).
Minor modifications have been applied to the text and figures for the sake of consistency in the thesis.

2.1. INTRODUCTION

Seismic data can be used in a variety of ways to monitor and explore the subsurface of the Earth. Such data are obtained by measuring the wavefield that is propagating through the subsurface at physical receivers. Seismic data can be acquired using an active source at the surface of the Earth, in which case receivers are usually located on the same surface as the source, or in a borehole. The receivers measure the full wavefield, i.e., both primary and multiply scattered events. These measurements are often used to obtain information about the structure of the subsurface and its properties [54]. Alternatively, data can be acquired using a passive source, which is a source of the wavefield that occurs naturally in the subsurface of the Earth. In this setup, the wavefield is recorded by a continuously recording receiver array, usually at the surface of the Earth. These measurements can contain additional information about processes in the subsurface, such as induced seismicity [55]. These types of measurements are receiving more attention because of the potentially damaging effects of induced seismicity in residential areas [56, 57].

Active measurements can be employed to supplement the passive measurments. Using advanced seismic processing techniques, the wavefield that is measured at the surface of the Earth can be redatumed to locations inside the subsurface. By redatuming receivers from their physical location on the surface to locations at depth, virtual receivers are created. The advantage of such virtual receivers is that, by considering many of them, the evolution of the wavefield through the subsurface over time can be studied, which can provide relevant information about source mechanisms and the locations of scatterers in the subsurface. Similar to receiver redatuming, physical sources at the surface can be redatumed to create virtual sources at any location in the subsurface. Furthermore, the response between any combination of a virtual source and virtual receiver can be retrieved, a process we call homogeneous Green's function retrieval. Whereas a Green's function describes the response of a medium to a Dirac function, a homogeneous Green's function is a Green's function superposed by its time-reversal to avoid a source singularity. The classical representation for the homogeneous Green's function retrieval was derived by Porter [15]. This method was further extended for inverse source problems by Porter and Devaney [16] and inverse scattering methods by Oristaglio [17]. This classical representation has been employed as the theoretical basis in the field of seismic interferometry to create virtual sources [20–22] or virtual receivers [23]. However, in all of these applications, it appeared that a complete enclosing boundary is vital for retrieving a full homogeneous Green's function without artifacts.

Recently, a new single-sided representation for homogeneous Green's function retrieval has been derived. Instead of an enclosing boundary, it uses a single, non-enclosing boundary, typically the Earth's surface [58]. An example of the application of this method on synthetic data can be found in Wapenaar *et al.* [25]. In this approach, the data-driven Marchenko method is used to create virtual sources and receivers in the subsurface from reflection data at the Earth's surface. Using the homogeneous Green's function retrieval, the response between one selected virtual source and all virtual receivers is obtained. The Marchenko method, for the purpose of geophysial applications, was first proposed for 1D by Broggini *et al.* [29], based on work by Rose [59], and was later extended for 2D and 3D applications [26, 60]. The method uses two types of input. The first is active-

source single-sided seismic reflection data measured at the surface of the Earth. The second is an estimation of the first wavefield event, which is called the first arrival, that would be caused by a source from a location in the subsurface to receiver locations at the surface of the Earth (hence, the first arrival of a Green's function between a subsurface location and the surface). The locations of the receivers of the Green's function match the locations of the receivers of the reflection response. The Marchenko method uses these data to create a full waveform Green's function, including all multiple scattering, for a virtual source in the subsurface and receivers at the Earth's surface. To model the first arrival, only a background velocity model is required, which can be estimated by processing the reflection data. A dense array of virtual sources for Green's functions in the subsurface can be created through repeated use of this methodology. Aside from the Green's function, the Marchenko method is also capable of retrieving a focusing function, which is designed to focus from the single-sided surface, where the reflection response is measured, to a focal location in the subsurface without any reverberation artifacts. The single-sided representation uses the focusing function, together with a Green's function, to create the response between a virtual source and receiver. Due to the single-sided focusing properties of the focusing function, the retrieval can be done for a single-sided recording setup without any artifacts.

Employing the Marchenko method on field data for practical applications is challenging due to the sensitivity of the Marchenko method to recording limitations of the reflection response. The sensitivity is partially caused by the fact that in the derivation of the Marchenko method, evanescent waves are ignored and it is assumed that the medium of interest is lossless. In real media, the wavefield suffers from absorption, which violates the latter assumption. Furthermore, the method requires the reflection response to be well sampled and the aperture to be sufficiently large. The Marchenko method has been succesfully applied on field data, by pre-processing the reflection response. Examples for the purpose of imaging can be found in Ravasi *et al.* [37] and Staring *et al.* [38], who used adaptive corrections in the Marchenko method. Homogeneous Green's function retrieval using the single-sided representation on field data was achieved by Wapenaar *et al.* [61] and Brackenhoff *et al.* [62].

The aim of this chapter is to apply the single-sided representation on field data and to consider the influence of recording limitations of the reflection response on the retrieved homogeneous Green's functions. To this end, we consider a 2D field seismic dataset from the Vøring basin off the coast of Norway. Along with the field data, we also consider a subsurface model, that is designed to simulate the subsurface of the area where the actual reflection response is recorded. Using this model, synthetic reflection data are created. First, we use the synthetic data to make a comparison between the results that are obtained when the single-sided representation is used and when the classical representation is used. The results show that the homogeneous Green's function is more accurately retrieved when the single-sided representation is used. The first arrivals that are used in these tests in the Marchenko method are all modeled using a monopole source mechanism. To study the influence of the source mechanism on the final result, the experiment is repeated using first arrivals that were modeled using a double-couple source mechanism, which is more representative for small-scale earthquakes [8]. The homogeneous Green function that is obtained in this way still contains the correct events and

has a double-couple signature. Next, we determine the sensitivity of the result to five recording limitations on the reflection data, namely coarse source-receiver sampling, missing near offsets, small aperture, offsets missing in one direction and absorption of the reflection data. The results of these numerical experiments are taken into account so that the field reflection data can be pre-processed and the single-sided representation can be applied properly. We employ both the classical and single-sided representation to the field data in order to compare the results. The applications show the potential of the single-sided representation for field data, as well as the possibility of applying the representation to passive field recordings.

2.2. THEORY

In this section, we present an overview of the definitions and equations that are required for homogeneous Green's function retrieval. The Green's function and focusing function are reviewed, followed by the definitions of the classical enclosed boundary and single-sided representations for homogeneous Green's function retrieval. The Marchenko method and its limitations are considered, as well as the double-couple source mechanism.

2.2.1. GREEN'S FUNCTION

The Green's function is defined as the solution of the wave equation to a Dirac point source which can be written as [63, 64]:

$$\partial_i(\rho^{-1}\partial_i G) - \kappa\partial_t^2 G = -\delta(\mathbf{x}-\mathbf{x}_A)\partial_t\delta(t), \tag{2.1}$$

where $G = G(\mathbf{x},\mathbf{x}_A,t)$ describes the response of the medium, at time t, at location $\mathbf{x}$ to a source at location $\mathbf{x}_A$. The locations are defined in 3D such that $\mathbf{x} = (x_1,x_2,x_3)$. The symbols $\rho = \rho(\mathbf{x})$ and $\kappa = \kappa(\mathbf{x})$ indicate the density and compressibility of the medium, respectively, δ indicates a Dirac delta function, ∂_t a temporal derivative and ∂_i the partial derivative in the three principal directions. The repeated subscript i follows the Einstein summation convention. Note that the source at the right hand side is defined with a temporal derivative acting on the Dirac delta function. This choice is made to simulate a volume injection-rate source. According to the reciprocity principle, the source and receiver location of the Green's functions can be interchanged, $G(\mathbf{x},\mathbf{x}_A,t) = G(\mathbf{x}_A,\mathbf{x},t)$.

We also consider the Fourier-transformed Green's function $G(\mathbf{x},\mathbf{x}_A,\omega)$:

$$G(\mathbf{x},\mathbf{x}_A,\omega) = \int_{-\infty}^{\infty} G(\mathbf{x},\mathbf{x}_A,t)e^{i\omega t}\mathrm{d}t, \tag{2.2}$$

where ω denotes the angular frequency and i the imaginary unit. Note, that the sign in the exponential can be reversed, as long as the same is done for the inverse Fourier transform. Using Equation (2.2), Equation (2.1) is transformed to the frequency domain:

$$\partial_i(\rho^{-1}\partial_i G) + \omega^2\kappa G = i\omega\delta(\mathbf{x}-\mathbf{x}_A). \tag{2.3}$$

A schematic illustration of the Green's function is shown in Figure 2.1(b), where the receivers are placed at the surface of a medium and its source inside the medium. We

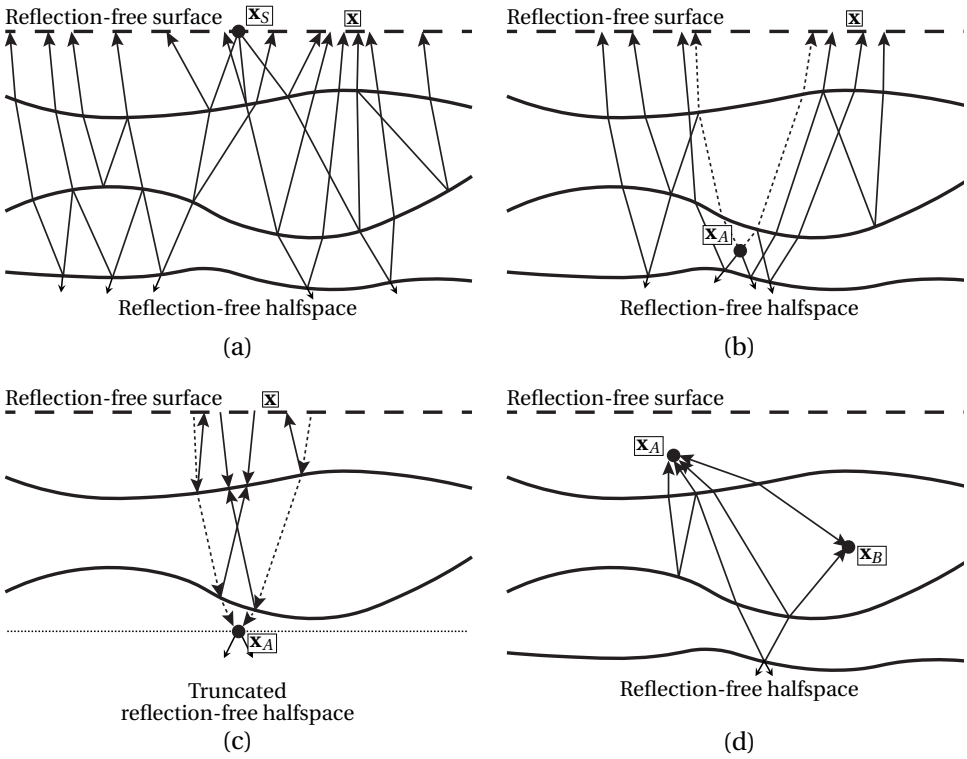

Figure 2.1: Possible raypaths drawn for, (a) a reflection response $R(\mathbf{x},\mathbf{x}_S,t)$, measured at varying receiver locations $\mathbf{x}$ at the surface, with a source at $\mathbf{x}_S$ also at the surface, (b) a Green's function $G(\mathbf{x},\mathbf{x}_A,t)$, measured at varying receiver location $\mathbf{x}$ at the surface with a source at $\mathbf{x}_A$ inside the medium, (c) a focusing function $f_1(\mathbf{x},\mathbf{x}_A,t)$, emitted from the surface at varying locations $\mathbf{x}$, focusing to a focal location $\mathbf{x}_A$ inside a medium that is truncated below $\mathbf{x}_A$ (indicated by the horizontal dotted line), and (d) a homogeneous Green's function $G_{\rm h}(\mathbf{x}_A,\mathbf{x}_B,t)$, between two locations, $\mathbf{x}_A$ and $\mathbf{x}_B$ inside the medium. The homogeneous Green's function is indicated with two-sided arrows to represent that it is the superposition of Green's function and its time-reversal. The dotted arrows in (b) and (c) indicate the first arrival for the Green's function and focusing function. The surfaces at the top of all figures are transparent, hence, there are no free-surface multiples.

assume that the medium has a transparent reflection-free halfspace at the top of the medium. In practice this situtation is obtained after the elimination of surface-related multiples. Some possible raypaths, including scattering, have been drawn in the figure. Figure 2.1(a) shows a special case of the Green's function, with both source and receivers placed at the surface of the medium. This is called the reflection response $R(\mathbf{x},\mathbf{x}_S,t)$ and contains all the reflections, both primaries and multiples, of the medium, however, we assume that the direct wave from the sources to the receivers are not present in the reflection response.

The homogeneous Green's function is defined as the superposition of the Green's function and its time-reversal. Because of the temporal derivative on the Dirac delta function in Equation (2.1), when time-reversal is applied, the source term will obtain an opposite sign. As a result, the superposition of the Green's function and its time-reversal

removes the source term, thereby avoiding a singularity at the source position:

$$G_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,t) = G(\mathbf{x},\mathbf{x}_A,t) + G(\mathbf{x},\mathbf{x}_A,-t), \tag{2.4}$$

$$\partial_i(\rho^{-1}\partial_i G_{\mathrm{h}}) - \kappa\partial_t^2 G_{\mathrm{h}} = 0, \tag{2.5}$$

and in the frequency domain:

$$G_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,\omega) = G(\mathbf{x},\mathbf{x}_A,\omega) + G^*(\mathbf{x},\mathbf{x}_A,\omega) = 2\Re\{G(\mathbf{x},\mathbf{x}_A,\omega)\}, \tag{2.6}$$

$$\partial_i(\rho^{-1}\partial_i G_{\mathrm{h}}) + \omega^2\kappa G_{\mathrm{h}} = 0, \tag{2.7}$$

where $G_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,t)$ and $G_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,\omega)$ denote the homogeneous Green's function in the time domain and frequency domain, respectively, $\Re$ is the real part of a complex function and the asterisk indicates complex conjugation. Figure 2.1(d) shows a schematic illustration of the homogeneous Green's function, $G_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,t)$, with some possible raypaths drawn, with both its source and receiver inside the medium. To reflect the superposition of the Green's function and its time-reversal, the raypaths are indicated with two-sided arrows.

2.2.2. FOCUSING FUNCTION

The focusing function $f_1(\mathbf{x},\mathbf{x}_A,t)$ describes a wavefield, at time t, at location $\mathbf{x}$, that focuses to a focal location $\mathbf{x}_A$ in the subsurface. The focusing function propagates in a medium that is truncated below $\mathbf{x}_A$, which means that there are no reflectors present below the focal location.

The focusing function can be decomposed into its upgoing and downgoing parts:

$$f_1(\mathbf{x},\mathbf{x}_A,t) = f_1^+(\mathbf{x},\mathbf{x}_A,t) + f_1^-(\mathbf{x},\mathbf{x}_A,t), \tag{2.8}$$

where $f_1^+(\mathbf{x},\mathbf{x}_A,t)$ denotes the downgoing focusing function and $f_1^-(\mathbf{x},\mathbf{x}_A,t)$ the upgoing focusing function. The downgoing part of the focusing function is defined as the inverse of the transmission response of the truncated medium [26].

The focusing function is schematically illustrated in Figure 2.1(c), where some possible raypaths have been drawn. The first arrival, which is indicated by the dotted raypath, propagates from the surface to the focal location and scatters at the reflectors, creating an upgoing wavefield. In order to ensure that these upgoing waves do not cause additional events arriving after the wavefield has focused, additional downgoing waves are injected from the surface, which cancel out these events. This occurs at the locations where arrowheads meet in Figure 2.1(c). Because of the reflection-free surface at $\mathbf{x}$, there are no events present in the focusing function to account for free-surface multiples. A more detailed description of the focusing function can be found in Slob *et al.* [30].

2.2.3. HOMOGENEOUS GREEN'S FUNCTION REPRESENTATION

The classical representation of the homogeneous Green's function states that the response between a source and receiver inside a medium can be retrieved from observations at a boundary. In order to achieve this, an enclosing boundary around the medium

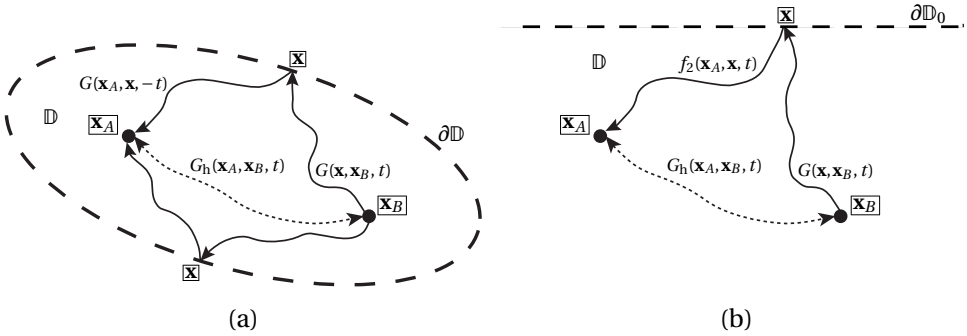

Figure 2.2: Recording setup for homogeneous Green's function retrieval using (a) the classical representation and (b) the single-sided representation. For both setups, a Green's function, $G(\mathbf{x},\mathbf{x}_B,t)$ is utilized. For the classical representation, an additional Green's function, $G(\mathbf{x}_A,\mathbf{x},t)$ is used to create a virtual receiver location. The medium of interest $\mathbb{D}$ is surrounded by an enclosing boundary $\partial\mathbb{D}$. Equation (2.9) is evaluated over this enclosing boundary. For the single-sided representation, the virtual receiver is not created using a Green's function, but rather a focusing function, $f_2(\mathbf{x}_A,\mathbf{x},t)$. The medium $\mathbb{D}$ is not enclosed, instead only a single-sided non-enclosing boundary $\partial\mathbb{D}_0$ is present at a single side. Equation (2.10) is evaluated over this single-sided boundary. The homogeneous Green's function is indicated by the dotted line.

of interest, over which the data can be recorded and/or injected, needs to be present [15–17]. The classical representation in the frequency domain can be written as follows:

$$G_h(\mathbf{x}_A,\mathbf{x}_B,\omega) = \oint_{\partial\mathbb{D}} \frac{-1}{i\omega\rho(\mathbf{x})}\{\partial_i G^*(\mathbf{x}_A,\mathbf{x},\omega)G(\mathbf{x},\mathbf{x}_B,\omega) - G^*(\mathbf{x}_A,\mathbf{x},\omega)\partial_i G(\mathbf{x},\mathbf{x}_B,\omega)\}n_i \mathrm{d}\mathbf{x}, \tag{2.9}$$

where n_i indicates the components of the normal vector in the three principal directions. The integral is evaluated over a boundary $\partial\mathbb{D}$ enclosing the medium $\mathbb{D}$. In Equation (2.9), the function $G(\mathbf{x},\mathbf{x}_B,\omega)$ describes the response of the medium at varying location $\mathbf{x}$ at the boundary to a source at location $\mathbf{x}_B$ inside the medium. The time-reversed function $G^*(\mathbf{x}_A,\mathbf{x},\omega)$ back-propagates the responses from the boundary to the receiver location $\mathbf{x}_A$, thereby creating a virtual receiver at $\mathbf{x}_A$. A schematic overview of the application of this representation is shown in Figure 2.2(a).

In practice, the classical representation is often not evaluated correctly, because acquisition on an enclosing boundary is not feasible and only measurements on a single-sided non-enclosing boundary, usually the Earth's surface, are available. As an approximation, Equation (2.9) can be evaluated over the single-sided boundary. Applying the representation in this way causes significant artifacts in the retrieved homogeneous Green's function, however. Due to the fact that few alternatives are available, the method is still widely applied to cases where no closed boundaries are present.

An alternative representation that can be employed uses a focusing function instead of a Green's function. This representation is capable of retrieving the full homogeneous Green's function with significantly less artifacts from a single-sided boundary, hence it is referred to as the single-sided representation. It can be written as [58, equation 30]:

$$G_h(\mathbf{x}_A,\mathbf{x}_B,\omega) = 4\Re\int_{\partial\mathbb{D}_0} \frac{1}{i\omega\rho_0} G(\mathbf{x},\mathbf{x}_B,\omega)\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega) - \{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}^2\mathbf{x}, \tag{2.10}$$

where $\partial\mathbb{D}_0$ denotes the single-sided boundary and ρ_0 is the density at the single-sided boundary. The right hand side of Equation (2.10) can be combined in a single focusing function, $f_2(\mathbf{x}_A,\mathbf{x},\omega)$, as:

$$f_2(\mathbf{x}_A,\mathbf{x},\omega) = f_1^+(\mathbf{x},\mathbf{x}_A,\omega) - \{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*. \tag{2.11}$$

In Equation (2.10), $G(\mathbf{x},\mathbf{x}_B,\omega)$ serves again as the response to a source location inside the medium, measured at the single-sided boundary $\partial\mathbb{D}_0$. The focusing function $f_2(\mathbf{x}_A,\mathbf{x},\omega)$ serves as the back-propagator of the responses from the boundary to the focal location inside the medium. A schematic representation of this procedure is shown in Figure 2.2(b).

The two representations in Equations (2.9) and (2.10) for homogeneous Green's function retrieval are similar in form, as both use a backward propagator on the response measured on the boundary. The main difference is that, for the single-sided representation, the backward propagator is a focusing function instead of a time-reversed Green's function. As one can interpret from Figure 2.1(c), the convergence of the focusing function to the focal location is ensured by the first arrival, whereas the coda of the focusing function removes unwanted reflections caused by the first arrival when it encounters reflectors while propagating to the focal location. The arrival times of the direct wave of the focusing function are the same as the arrival times of the direct wave of the time-reversed Green's function. The difference is that the coda of the focusing function is designed to cancel out the events that are introduced by the direct arrival, whereas the coda of the time-reversed Green's function introduces additional artifacts, in the form of reverberations.

2.2.4. MARCHENKO METHOD

We use the Marchenko method to retrieve the focusing functions and Green's functions required for the representations for homogeneous Green's function retrieval. A more detailed consideration of the method can be found in Wapenaar *et al.* [26]. Here we only consider the equations and properties of the method relevant for this chapter. The Green's function and focusing function of a medium are related via the reflection response:

$$G(\mathbf{x},\mathbf{x}_B,t) - f_2(\mathbf{x}_B,\mathbf{x},-t) = \int_{\partial\mathbb{D}_0}\int_{-\infty}^{\infty} R(\mathbf{x},\mathbf{x}_S,t')\, f_2(\mathbf{x}_B,\mathbf{x}_S,t-t')\,\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S, \tag{2.12}$$

where $\mathbf{x}_B$ is a location inside the medium $\mathbb{D}$ and $\mathbf{x}_S$ indicates the array of sources that are present on the non-enclosing surface $\partial\mathbb{D}_0$. Equation (2.12) states that if the reflection response R at a boundary $\partial\mathbb{D}_0$ and a focusing function with a focal location inside medium $\mathbb{D}$ are available, the Green's function with a source at the focal location can be retrieved. The retrieval of the focusing function inside the medium can be achieved using the iterative Marchenko equation:

$$\begin{aligned} &f_{2,k+1}(\mathbf{x}_B,\mathbf{x},-t) = \\ &\qquad D(\mathbf{x},\mathbf{x}_B,t) - w_t(\mathbf{x},\mathbf{x}_B,t)\int_{\partial\mathbb{D}_0}\int_{-\infty}^{\infty} R(\mathbf{x},\mathbf{x}_S,t')\, f_{2,k}(\mathbf{x}_B,\mathbf{x}_S,t-t')\,\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S, \end{aligned} \tag{2.13}$$

where $f_{2,k}(\mathbf{x}_B, \mathbf{x}, t)$ is the estimated focusing function after k iterations, $D(\mathbf{x}, \mathbf{x}_B, t)$ is the first arrival of the Green's function and $w_t(\mathbf{x}, \mathbf{x}_B, t)$ is a windowing function. The windowing function is used to mute the Green's function completely. When the windowing function is applied to Equation (2.12), the Green's function is removed. The arrival times of the first arrival of the Green's function are the same as the arrival times of the last arrival of the time-reversed focusing function, hence the windowing function also removes the last arrival of the time-reversed focusing function, however, it will not remove the coda of the time-reversed focusing function. Therefore, in order to obtain the full focusing function in Equation (2.13), $D(\mathbf{x}, \mathbf{x}_B, t)$ needs to be added after the windowing function has been applied. The windowing function $w_t(\mathbf{x}, \mathbf{x}_B, t)$ can be estimated from $D(\mathbf{x}, \mathbf{x}_B, t)$, as the edge of the muting area is located around the first arrival. In order to use Equation (2.13) and start the iterative scheme, a first estimation of the focusing function is required. The time-reversed first arrival $D(\mathbf{x}_S, \mathbf{x}_B, -t)$ of the Green's function is used as this first estimation of the focusing function $f_{2,0}(\mathbf{x}_B, \mathbf{x}_S, t)$. If this arrival is emitted into the medium, it will cause additional reflections that are not cancelled. By using Equation (2.13) iteratively, until convergence is achieved, the coda of the focusing function is retrieved, which will suppress the undesired reflections. The only required components for the iterative scheme are a reflection response measured at the single-sided boundary (i.e. the Earth's surface) and the direct arrival from the focal point to the same boundary. This direct arrival can be modeled using a smooth velocity model. Because only the direct arrival is of interest, the model requires no detailed features. Generally, a monopole point source is used to model these first arrivals. After the focusing function has been retrieved, it can be used in Equation (2.12) to compute the Green's function, $G(\mathbf{x}, \mathbf{x}_B, t)$. Subsequently, this Green's function and a similarly derived focusing function for focal point $\mathbf{x}_A$ are used in Equation (2.10) to retrieve the homogeneous Green's function, $G_{\mathrm{h}}(\mathbf{x}_A, \mathbf{x}_B, \omega)$. All the Green's functions and focusing functions in this chapter are retrieved using the Marchenko method to ensure the representations are applied to the field data and the synthetic data in the same way.

The Marchenko method has restrictions, particularly when it is applied on field data. An important underlying assumption of the Marchenko method that is considered in this chapter, is that no free-surface multiples are present in the reflection response. Hence, the free-surface multiples should be removed prior to applying the Marchenko method, for example by applying a surface-related multiple elimination scheme [65]. There are ways to incorporate these multiples in the Marchenko method as well, for an example, see Singh *et al.* [49]. Additionally, the reflection response that is used needs to be accurate, as issues with the quality of the recording have strong influences on the final result. An important requirement is that the medium of interest needs to be lossless, which is an unrealistic approximation in real media. Also, the reflection response needs, preferably, to be densely sampled, contain both positive and negative source-receiver offsets and have sufficient recording length and aperture. The effects of some of the limitations of the Marchenko method are considered in Ravasi *et al.* [37], Brackenhoff [66] and Staring *et al.* [38]. When synthetic data are used, the reflection response can be modeled without these limitations. However, when field data are recorded, not all of these requirements are fulfilled and appropriate pre-processing is required.

2.2.5. DOUBLE-COUPLE SOURCE

As mentioned before, the Green's functions and focusing functions that are retrieved using the Marchenko method usually have a monopole source signature. In the field, especially for passive recordings, there are many different types of source mechanisms. These different source mechanisms can be taken into account for the homogeneous Green's function retrieval as well. We show this by incorporating a double-couple source mechanism in the representation. This mechanism is chosen because it is considered to be representative for an earthquake reponse when the wavelength of the wavefield is larger than the dimensions of the source [8]. The mechanism differs from a monopole source in the sense that the radiation pattern is not homogeneous, rather, the polarity and amplitude vary depending on the radiation angle of the source. There will be four distinct polarity changes along the radiaton pattern, at 90, 180, 270 and 360 degrees. The source can be inclined at various angles, for example aligned along a fault, which changes the orientation of the polarity changes as well.

Operator $\mathfrak{D}_B^\theta\{\cdot\}$ is introduced, which transforms a monopole source signature to a double-couple source signature. This operator is defined as

$$\mathfrak{D}_B^\theta\{\cdot\} = (\theta_i^\parallel + \theta_i^\perp)\partial_{i,B}\{\cdot\}, \tag{2.14}$$

where $\partial_{i,B}$ is a component of the vector containing the partial derivatives acting on the monopole signal originating from source location $\mathbf{x}_B$, to transform the source signature to a double-couple mechanism, $\theta_i^\parallel$ is a component of the unit vector that orients one couple of the signal parallel to the fault plane and $\theta_i^\perp$ is a component of a similar vector that orients the other couple perpendicular to the fault plane. The operator is applied to the representation for homogeneous Green's function retrieval in Equation (2.10), which yields:

$$\begin{aligned}&\mathfrak{D}_B^\theta\{G_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega)\} = \\ &4\Re\int_{\partial\mathbb{D}_0}\frac{1}{\omega\rho(\mathbf{x})}\mathfrak{D}_B^\theta\{G(\mathbf{x},\mathbf{x}_B,\omega)\}\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega) - \{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}^2\mathbf{x},\end{aligned} \tag{2.15}$$

where the subscript B in the double-couple operator indicates that the operator acts on the source in location $\mathbf{x}_B$. In this representation, the double-couple operator is only applied to the Green's function for the virtual source location, $G(\mathbf{x},\mathbf{x}_B,\omega)$, and not to the focusing functions, $f_1^+(\mathbf{x},\mathbf{x}_A,\omega) - \{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*$, for the virtual receiver location. This is because the focusing function is retrieved using Equation (2.13), and thus has a monopole source signature. For the retrieval of the Green's function, $\mathfrak{D}_B^\theta\{G(\mathbf{x},\mathbf{x}_B,t)\}$, we apply the double-couple operator to Equations (2.12) and (2.13):

$$\begin{aligned}&\mathfrak{D}_B^\theta\{G(\mathbf{x},\mathbf{x}_B,t)\} - \mathfrak{D}_B^\theta\{f_2(\mathbf{x}_B,\mathbf{x},-t)\} = \\ &\qquad\int_{\partial\mathbb{D}_0}\int_{-\infty}^{\infty} R(\mathbf{x},\mathbf{x}_S,t')\mathfrak{D}_B^\theta\{f_2(\mathbf{x}_B,\mathbf{x}_S,t-t')\}\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S,\end{aligned} \tag{2.16}$$

$$\begin{aligned}&\mathfrak{D}_B^\theta\{f_{2,k+1}(\mathbf{x}_B,\mathbf{x},-t)\} = \mathfrak{D}_B^\theta\{D(\mathbf{x},\mathbf{x}_B,-t)\} - \\ &\qquad w_t(\mathbf{x},\mathbf{x}_B,t)\int_{\partial\mathbb{D}_0}\int_{-\infty}^{\infty} R(\mathbf{x},\mathbf{x}_S,t')\mathfrak{D}_B^\theta\{f_{2,k}(\mathbf{x}_B,\mathbf{x}_S,t-t')\}\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S.\end{aligned} \tag{2.17}$$

The first arrival is modeled using the double-couple source mechanism instead of a monopole source mechanism, and used in Equation (2.17), for k=0, as $\mathfrak{D}_B^\theta\{f_{2,0}(\mathbf{x}_B,\mathbf{x}_S,t)\}$. This results in an estimation of the focusing function $\mathfrak{D}_B^\theta\{f_2(\mathbf{x}_B,\mathbf{x},t)\}$, with a double-couple source mechanism. This focusing function is then used in Equation (2.16) to obtain the Green's function $\mathfrak{D}_B^\theta\{G(\mathbf{x},\mathbf{x}_B,t)\}$ with a double-couple source signature, which in turn is used in Equation (2.15).

2.3. DATASETS

We consider both a synthetic dataset and a field dataset. The synthetic dataset is based on the field dataset, hence we first consider the parameters of the field recording before the synthetic dataset is considered.

2.3.1. VØRING DATA

The considered field data were recorded in a marine setting over the Vøring basin in offshore Norway by SAGA Petroleum A.S., which is currently part of Equinor. The data consist of a reflection response acquired along a 2D line. For each source location, the receivers were moved along with the source position. The parameters of the acquisition can be found in Table 2.1. An example of a single common-source record is shown in Figure 2.3(a), where the wavelet on the data has been reshaped to a 30Hz Ricker wavelet for display purposes (a Ricker wavelet is defined as minus the second time-derivative of a Gaussian function). There are several events present in the common-source record, however it should be noted that the near offsets are missing. This is a common situation for marine acquisition, because receivers cannot be placed too close to active sources. The sources and receivers are located inside the water, and because S-waves cannot propagate in water, only P-waves are measured by the receivers. There are conversions from P-waves to S-waves and back in the subsurface below the water, so there are P-waves present in the reflection data that were converted from S-waves. The data also contain free-surface multiples, which are indicated by the black arrows.

Table 2.1: Acquisition parameters for the Vøring dataset.

Parameter	Value
Number of source positions	399
Source spacing	25m
First source position	5000m
Final source position	14950m
Number of receiver positions per source	180
Receiver spacing	25m
Minimum source-receiver offset	150m
Maximum source-receiver offset	4625m
Number of recording samples	2001
Sampling interval	0.004s
High-cut frequency	90 Hz

Aside from the reflection data, a smooth P-wave velocity model is provided by the

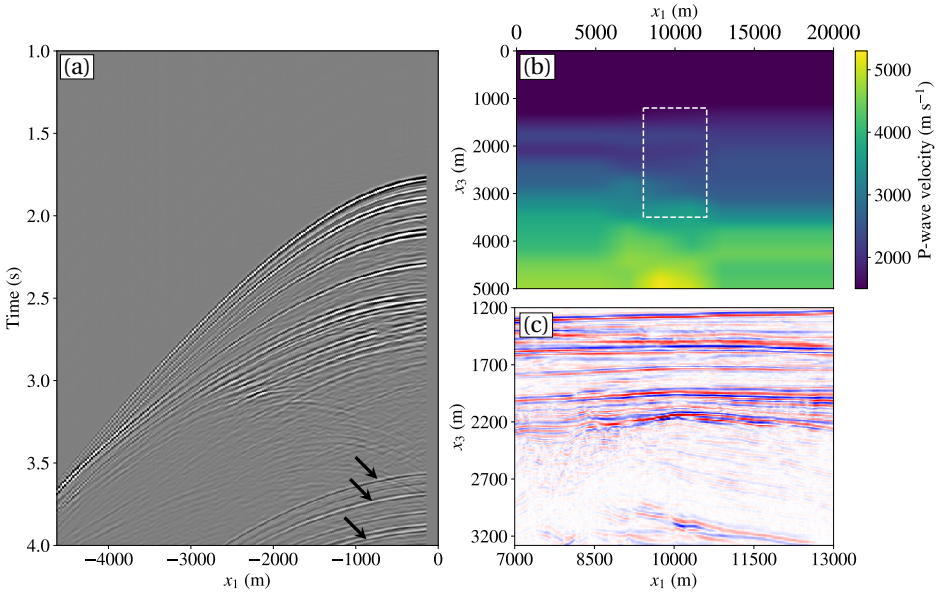

Figure 2.3: (a) Unprocessed common-source record of the reflection response recorded in the Vøring basin. The source is located at zero offset, where along with other near offsets no data could be recorded. Free-surface multiples are present in the data at late times, indicated by the black arrows. (b) Estimated P-wave velocity model in m s^{-1} of the subsurface region where the data in the Vøring basin were recorded. The white dashed box represents the region of interest. (c) Image of the region of interest, indicated by the white dashed box in (b). The wavelet in the data in (a) and (c) has been reshaped into a 30Hz Ricker wavelet for display purposes.

Delphi Consortium and displayed in Figure 2.3(b). This model is used to create the first arrivals required for the Marchenko method. The dashed white box in Figure 2.3(b) indicates the region of interest where the homogeneous Green's functions are retrieved in this chapter. An image of the region of interest was constructed, which is shown in Figure 2.3(c), using the reflection data, the velocity model and a one-way recursive depth migration based on Thorbecke *et al.* [67]. Imaging is not the main subject of this chapter, so the details of the construction of this image are not discussed, however it should be noted that free-surface multiples were removed prior to imaging. Also note that the retrieval of the homogeneous Green's function (discussed in the sections on synthetic data and field data) and the construction of the image were done independently of each other. The image is used to construct a subsurface model and to validate the homogeneous Green's function retrieval on the field data.

2.3.2. SYNTHETIC DATA

The synthetic models that we use in this chapter are based on the field dataset. Because there are no direct measurements of the subsurface available, an acoustic model is interpreted based on the smooth P-wave velocity model in Figure 2.3(b) and the image in Figure 2.3(c). This is done, because there is no S-wave velocity information of the sub-

surface available, hence we cannot construct an elastodynamic model. It is possible to use an elastodynamic representation for homogeneous Green's function retrieval, see Reinicke and Wapenaar [44], however for the Vøring basin we do not have the required multi-component data to do so.

To construct the model, we use the image to determine the locations of geological layer boundaries and determine the P-wave velocities by calculating the interval velocities in the smooth P-wave velocity model between the contrasts. The interpreted velocity model is displayed in Figure 2.4(b), which shows hard boundaries. Notice that below the area of interest there are no reflectors present in the medium. It is not possible to achieve reliable imaging in this area, therefore no structures are interpreted. Features outside the region of interest were extrapolated to create a full model. A density model is also constructed in order to ensure strong amplitudes in the reflection data. Because no direct measurements of density in the subsurface are available, it is not possible to create a density model that represents the subsurface accurately. Instead, the densities are chosen based on realistic ranges, that ensure strong contrasts between the layers. Figure 2.4(c) displays the density model.

The finite-difference wavefield modeling code by Thorbecke and Draganov [68] is used to create the single-sided reflection data as input for the Marchenko method. The reflection response of the interpreted model is modeled using the same measurement parameters as for the real dataset, shown in Table 2.1, with pressure receivers recording the modeled wavefield at the surface of the model, in response to force sources that are located at the same surface. However, all offsets are included, no absorption is modeled and free-surface multiples are ignored. When comparing Figure 2.3(a) and Figure 2.4(a), there are similar events, however fewer events are present in the synthetic data. Not all the reflectors in the subsurface can be properly imaged and interpreted, therefore only the major features are present. The field data is more complex than the synthetic data. Because the exact P-wave velocity and density information of the actual medium are not available, there is an amplitude mismatch. The converted waves due to elastic interactions from the actual recording are also not taken into account. Furthermore, no free-surface multiples are present in the modeled reflection data.

2.4. SYNTHETIC DATA

The synthetic data are used to retrieve the homogeneous Green's function. The single-sided reflection data and the velocity model are the only information employed. Furthermore, the velocity model is smoothed to create a background velocity model, without any detailed features. This smooth velocity model is used in an Eikonal solver that is based on the method by Vidale [69] to determine the arrival times from every location in the region of interest, that is indicated in Figures 2.4(b) and (c), to the single-sided surface. These arrival times are combined with amplitude estimations using the method by Spetzler and Angelov [70] and convolved with a wavelet to create a band-limited pressure wavefield with a monopole source signature. Note that only the smooth velocity model is used for this approach and no density information is used. The first arrivals and the reflection data are used as input for the iterative Marchenko equation, as described by Equation (2.13), using the code by Thorbecke *et al.* [34]. The windowing function is created by using the arrival times from the Eikonal solver. Through use of the iterative code,

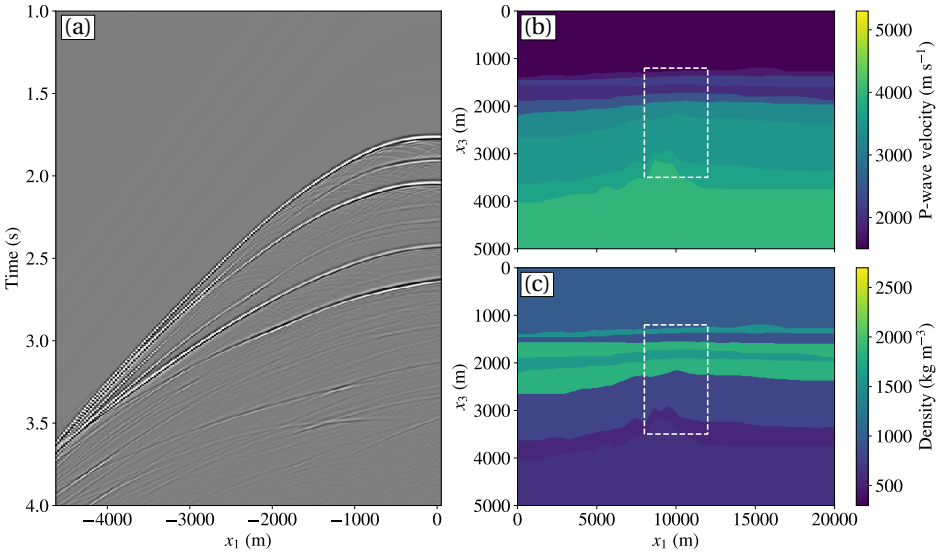

Figure 2.4: (a) Common-source record of the acoustic reflection response, modeled using finite-difference modeling in the P-wave velocity model from (b) and density model from (c), at the same location as the common-source record in Figure 2.3(a). The data have been convolved with a 30Hz Ricker wavelet for display purposes. (b) Synthetic P-wave velocity model in m s^{-1} based on the smooth velocity model from Figure 2.3(b) and the image from Figure 2.3(c). (c) Synthetic density model in kg m^{-3} based on the image from Figure 2.3(c). The white boxes in (b) and (c) indicate the region of interest.

Green's functions and focusing functions are obtained.

2.4.1. HOMOGENEOUS GREEN'S FUNCTION RETRIEVAL

To demonstrate the advantage of the single-sided representation, we compare different ways of retrieving the homogeneous Green's function. We aim to replicate a realistic situation as much as possible, hence the only data that are used are the smooth velocity model and the modeled reflection response. First, to create a benchmark for the homogeneous Green's function, we model the wavefield directly inside the medium by placing a volume injection-rate source and pressure receivers inside the region of interest. The wavefield is time-reversed and added to the original wavefield conforming to Equation (2.4) to create the homogeneous Green's function. Three snapshots of this result are shown in Figures 2.6(a)-(c) at 0ms, 200ms and 400ms, respectively. This is the result to which the homogeneous Green's functions that are obtained using the representations in Equations (2.9) and (2.10) are compared to. In Figure 2.6, the wavefield is convolved with a 15Hz Ricker wavelet and dotted black lines are shown for reference where we expect scattering to take place. This type of visualization is used for all the snapshots we produce for the synthetic data.

Now, we assume that we do not know the exact model and use the Marchenko method to retrieve Green's functions in the region of interest in the subsurface, from the reflec-

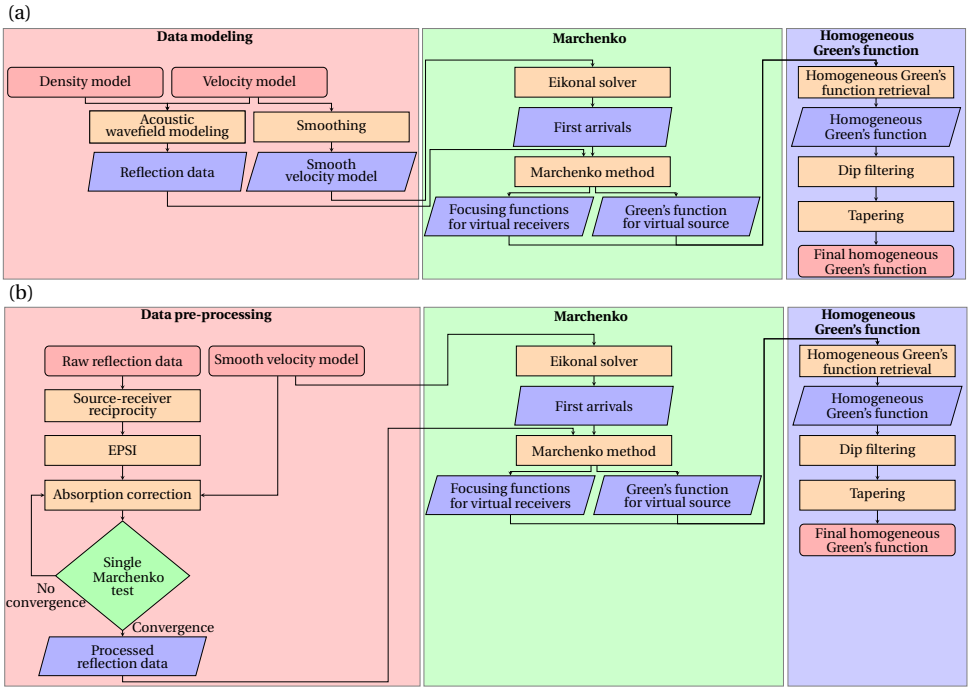

Figure 2.5: Flowchart for the retrieval of the homogeneous Green's function for (a) the synthetic data and (b) the field data.

tion data at the single-sided surface at the top of the model. These Green's functions are used to evaluate Equation (2.9) at the single-sided surface, not at an enclosing surface. The location $\mathbf{x}_B$ of one response, $G(\mathbf{x},\mathbf{x}_B,t)$, is kept constant as the virtual source position, while the location $\mathbf{x}_A$ of the other response, $G(\mathbf{x}_A,\mathbf{x},t)$, varies to serve as the virtual receiver positions. The positions of the virtual source and receivers are exactly the same as the respective source and receiver positions of the modeled wavefield. The retrieved homogeneous Green's function is shown in Figures 2.6(d)-(f) at 0ms, 200ms and 400ms, respectively. In Figure 2.6(d), for zero time, there is noise present in the entire snapshot, and while there is a focus of the wavefield to the source position, there are strong artifacts present directly above the source position. Furthermore, coherent artefacts with weaker amplitudes are present throughout the snapshot. The snapshot at 200ms in Figure 2.6(e) shows that the downgoing primary wavefield is retrieved, however, the upgoing primary wavefield is missing, instead forming into a weakening event with incorrect arrival times. Furthermore, while the coda of the wavefield has been retrieved with the correct upgoing reflections, the downgoing reflections are missing, and downgoing artifacts from above the source location are present in the snapshot. Aside from these problems, the snapshot is contaminated by artifacts. The final snapshot at 400ms in Figure 2.6(f) shows similar problems.

For comparsion, we repeat the experiment, however, instead of using a full Green's function $G(\mathbf{x},\mathbf{x}_B,t)$ for the virtual source, we only use its first arrival to reduce the num-

ber of artifacts. The results are shown in Figures 2.6(g)-(i) at 0ms, 200ms and 400ms, respectively. Compared to the previous experiment, the number of artifacts decreases, although not all are removed. The strong source artifacts at time zero remain present and the upgoing primary wavefield and downgoing coda are not restored.

Next we apply the single-sided representation using Equation (2.10). For the virtual source location $\mathbf{x}_S$ the same Green's function, $G(\mathbf{x},\mathbf{x}_B,t)$, is used as in the previous two experiments. However, the Green's function that was used to create the virtual receiver, $G(\mathbf{x}_A,\mathbf{x},t)$, is replaced by a focusing function, $f_2(\mathbf{x}_A,\mathbf{x},t)$. Figures 2.6(j)-(l) shows the result at 0ms, 200ms and 400ms, respectively. The improvement is noticeable, as artifacts are removed from the homogeneous Green's function. Aside from this, the primary wavefield is fully reconstructed, as is the coda of the downgoing wavefield. When comparing this result to the benchmark, it shows that the events are retrieved at the correct locations and times, although an amplitude mismatch is present. This is due to the fact that the amplitude of the first arrival is not exact, because we assume that we cannot model the first arrival in the real medium. When the amplitude of the first arrival is incorrect it scales the retrieved focusing functions and Green's function, although the correct relative amplitude, arrival times and events are obtained [62, 66]. Some of the events are not reconstructed, especially when the angle of the reflection is high. This is because the single-sided boundary is assumed to be infinite, while in reality the aperture is limited. The reflection response lacks certain angles of reflection, so these angles cannot be reconstructed for the homogeneous Green's function. At zero time the snapshot contains less artifacts, however, some remain, which contaminate the homogeneous Green's function at later times.

An analysis of the wavenumber-frequency spectrum of the homogeneous Green's function retrieved using Equation (2.10) revealed noise at locations corresponding to high angles of the wavefield. As mentioned before, the aperture of the reflection response is limited, so at these angles no events can be reconstructed. However, this also introduces noise into the final result as the Marchenko method tries to reconstruct these angles. To remove the noise at these high angles, dip filtering is applied to the homogeneous Green's function, which can be applied because no physical events are reconstructed for these angles. This creates some small artifacts around the first arrival of the homogeneous Green's function, which in turn are removed by applying a time-dependent taper. The improved result is shown in Figure 2.6(m)-(o) at 0ms, 200ms and 400ms, respectively. None of the desired events have been removed, however, the artifacts around the source position are gone. When this result is compared to the modeled response, we find the match of the arrival times quite satisfactory, while the amplitude mismatch remains, and the improvement over the homogeneous Green's function obtained using the classical representation is significant. When comparing the result of this retrieval to the modeled result from Figure 2.6(a)-(c), the match between desired events and their arrival times are strong. A notable difference occurs in the shape of the source at zero time. When the wavefield is modeled directly, the source radiates uniformly in all directions, however, when the wavefield is retrieved with the Marchenko method, this is not possible. This is caused by the limited aperture, which cannot capture the parts of the wavefield that propagate horizontally at this depth. Additionally, the dip filtering that is applied will also affect the source radiation. Even with these limitations, this type

of workflow yields the optimal result, and is used for the retrieval of the homogeneous Green's function from here onward. The workflow is summarized in Figure 2.5(a).

Finally, we consider the workflow in Figure 2.5(a) with a double-couple source mechanism. Instead of using Equation (2.10) for the homogeneous Green's function retrieval, we use Equation (2.15). This means that we use a Green's function, with a double-couple source signature, obtained with the Marchenko method, to investigate whether the source signature has a strong effect on the retrieved homogeneous Green's function. A first arrival caused by a double-couple source, needed to initiate the Marchenko method, cannot be retrieved using the Eikonal solver, therefore we model the required first arrival using the finite-difference code. The double-couple source mechanism is incorporated into the code using the moment tensor approach for finite-difference modeling, as described by Li *et al.* [71]. The double-couple source mechanism is an elastic mechanism, however, it is assumed that all the data are acoustic. To model an acoustic response to the double-couple source, we use the smoothed P-wave velocity model and a homogeneous density model of 1000kg m^{-3}, respectively, as well as a homogenous S-wave velocity model of 1000m s^{-1}, except for the top layer where the S-wave velocity is set to zero, simulating a marine setting. This means that no S-waves will arrive at the receiver locations. The coda of this modeling will be incorrect, however, as we only use the first arrival, this has no consequences for our results, aside from the scaling of the first arrival. The first arrival will be a pure P-wave, as the velocities of the P-wave model are larger than those of the S-wave model, which is a realistic situation.
We use a double-couple source that was inclined at -20 degrees to model the first arrival that it is used in the Marchenko method to create a Green's function. This Green's function replaces the one created from a monopole source that was used in the previous examples and the resulting homogeneous Green's function is shown in Figures 2.6(p)-(r). The arrival times of the events are the same as the ones in the previous experiment where only monopole signatures were employed. The main differences are found in the polarity of the events. Due to the double-couple source, the amplitude and polarity of the wavefield changes depending on the angle. Because the source we used to model the first arrival for the Green's function is inclined at -20 degrees, these polarity changes are not occuring at 90, 180, 270 and 360 degrees, but, shifted by -20 degrees, at approximately 70, 160, 250 and 340 degrees. All retrieved events, not just the first arrival, are affected by this inclination, without introducing any additional artifacts. This shows that different source signatures can be incorporated into the single-sided representation for homogeneous Green's function retrieval.

2.4.2. LIMITATIONS OF REFLECTION DATA

The reflection response that we use to retrieve the results in Figure 2.6 is nearly ideal, due to the recording setup and the absence of absorption. In the following experiment, we perform the homogeneous Green's function retrieval using the workflow in Figure 2.5(a), with five different types of acquisition limitations applied to the reflection response. In all five cases we perform the entire process that is outlined in the workflow, starting with the Marchenko method to retrieve the focusing function and Green's function from the reflection response, to which one of the limitations is applied, followed by applying Equation (2.10) to obtain the homogeneous Green's function. This demonstrates the ef-

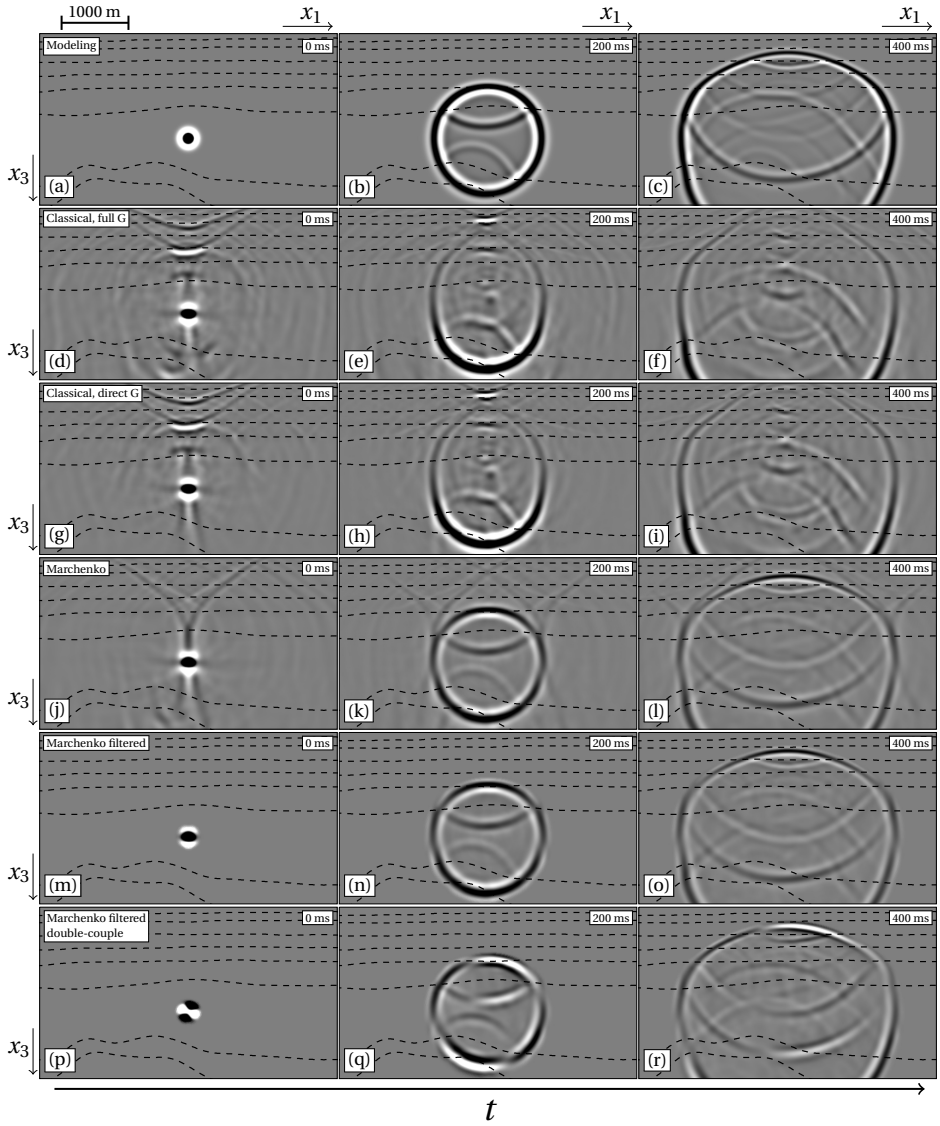

Figure 2.6: Snapshots of the wavefield at different times. Column 1 indicates t=0ms, column 2 t=200ms and column 3 t=400ms. All wavefields have been convolved with a 15Hz Ricker wavelet for display purposes. The black dotted lines indicate the locations of geological layer interfaces. (a)-(c) Directly modeled homogeneous Green's function in the subsurface used as a reference. (d)-(f) Homogeneous Green's function retrieval using Equation (2.9) with full Green's functions for the virtual source and receiver positions. (g)-(i) Idem, however with a full Green's function for the virtual receiver position and a direct arrival as the Green's function for the virtual source position. (j)-(l) Homogeneous Green's function retrieval using Equation (2.10), a Green's function for the virtual source position and a focusing function for the virtual receiver position, without filtering. (m)-(o) Idem, with dip filtering and tapering applied. (p)-(r) Idem, using a Green's function with a double-couple source signature inclined at -20 degrees.

fects of the limitations of the reflection response on the homogeneous Green's function that is obtained. The results of these tests are shown in Figure 2.7, where (a), (b) and (c), show the result from Figure 2.6(f), (i) and (o), respectively, which are used as a reference. All snapshots in Figure 2.7, are shown at 400ms. The results shown in the rows below the first one are achieved in the same way as the result from Figure 2.7(c), with different types of limitations applied to the reflection response. Each column shows a varying value of the limitation to indicate the sensitivity of the method to these limitations.

To determine the Nyquist spatial sampling interval for the data, we use the sampling criterion:

$$\Delta x < \frac{c_1}{2 f_{max} \sin(\alpha_{max})}, \tag{2.18}$$

where Δx is the spatial sampling, c_1 is the velocity of the top layer of the subsurface, f_{max} is the maximum frequency of the wavelet and α_{max} is the maximum angle of the waves in the top layer. To determine the sampling, we use the top layer velocity of the data, which corresponds to 1500 ms^{-1}, a "maximum" frequency of 30Hz for the 15Hz Ricker wavelet and a maximum angle of 60 degrees. This results in a recommended spatial sampling of about 29m. The current sampling of 25m should therefore be sufficient for our applications, which is supported by the quality of the retrieved homogeneous Green's function that have been obtained in the previous section.

In Figures 2.7(d)-(f), we display the result retrieved using a reflection response that has a coarse source-receiver sampling. The sampling values for the receiver and source spacing are 50m, 75m and 100m, which are double, triple and quadruple the original spacing, and all exceed the Nyquist sampling. When the spacing is doubled, noise is introduced into the homogeneous Green's function due to spatial aliasing. The physical events are distorted by this noise and background artifacts are present. This issue is worsened when the spacing distance is tripled. Some events are obscured and strong noise is present. Quadrupling the spacing produces a result that is unusable. It consists almost entirely of noise and the physical events cannot be distinguished. Hence, for successful use of the method the source-receiver sampling of the reflection response must not be larger than the Nyquist sampling.

Next, we consider the influence of missing small source-receiver offsets. The result is shown in Figures 2.7(g)-(i), where, respectively, the first 125m, 250m and 500m of the offsets are removed from the reflection response, for both positive and negative offsets, and replaced with empty traces. When 125m of offsets are missing, the homogeneous Green's function is still comparable to the ideal situation. There is a degradation in quality and a few artifacts are present. Removing 250m of near offsets produces a less accurate homogeneous Green's function, with a stronger degradation in quality. When 500m of near offsets are removed, the low angle reflections are not reconstructed properly, similarly to the high angle noise that is caused by the limited aperture of the data. The events below and above the virtual source position are missing and strong artifacts are present. These artifacts can be partially removed by adjusting the dip filtering to remove the low angle events, however, this procedure will also remove part of the physical events and is therefore not reccomendable to apply. The near offsets do have an impact on the final result and ideally should be reconstructed, if possible, before applying the Marchenko method.

Aside from missing near offsets, a reflection response can also be recorded exclusively in one direction. Figures 2.7(j) shows the result using only positive source-receiver offsets in the reflection response and Figure 2.7(k) does the same for negative source-receiver offsets. In both cases unwanted artifacts are present and, depending on the direction of the source-receiver offsets, large parts of the physical events are missing. This issue can be avoided by applying source-receiver reciprocity. We perform source-receiver reciprocity on the reflection response containing only the negative offsets and retrieve the result shown in Figure 2.7(l). This homogeneous Green's function is similar to the one produced in the ideal situation. We retrieve a similar result when we apply source-receiver reciprocity on the reflection response containing only positive offsets.

The final acquisition limitation that we review is the absence of large source-receiver offsets, or aperture limitation of the data. In Figures 2.7(m)-(o), we show the homogeneous Green's function when the largest source-receiver offset is, respectively, 2000m, 1000m and 500m. When the aperture is 2000m, the homogeneous Green's function is comparable to the one that is retrieved in the ideal situation, with some artifacts introduced. When the aperture is limited to 1000m, a homogeneous Green's function is produced that contains more artifacts and is missing a larger portion of the angles of the desired events. This is once again due to the fact that the angles of this part of the wavefield are not present in the reflection response. If only 500m of aperture are available, a large portion of the angles of events are missing. This is clear when Figure 2.7(o) is compared to Figure 2.7(i). The parts of the events that are missing due to the lack of near offsets are present in the case of limited aperture and vice versa. By applying a stronger dip filter, the artifacts can be suppressed, however, as mentioned before, this also removes part of the physical events.

Finally, we consider the case of absorption, which is a factor that cannot directly be influenced during the acquistion of the reflection response. Even if the recording setup is ideal, absorption of the wavefield is present and can degrade the result. This is demonstrated in Figures 2.7(p)-(r), where the loss is simulated by applying time-dependent absorption functions to the reflection data of $0.9e^{-0.2t}$, $0.8e^{-0.3t}$ and $0.7e^{-0.4t}$, respectively. In case of low absorption, the homogeneous Green's function still contains the physical events, although they have a lower amplitude. The artifacts are present with a low amplitude. If the absorption is increased, the physical events start to vanish and the artifacts are more pronounced. In case of high absorption, the physical events have very low amplitude and there are strong artifacts present.

2.5. FIELD DATA

2.5.1. PRE-PROCESSING

The raw seismic field reflection data cannot directly be used with the Marchenko method, because when the method uses these data it does not converge to a solution. The reflection data needs to be preprocessed to compensate for the limitations [38]. As was shown by the results in Figure 2.7, there are multiple effects that need to be taken into account. The reflection data are missing small source-receiver offsets, only have offsets in a negative direction, may be sub-sampled, possibly have insufficient aperture and are affected by absorption. Aside from this, the 2D line was recorded in a 3D setting, which can cause

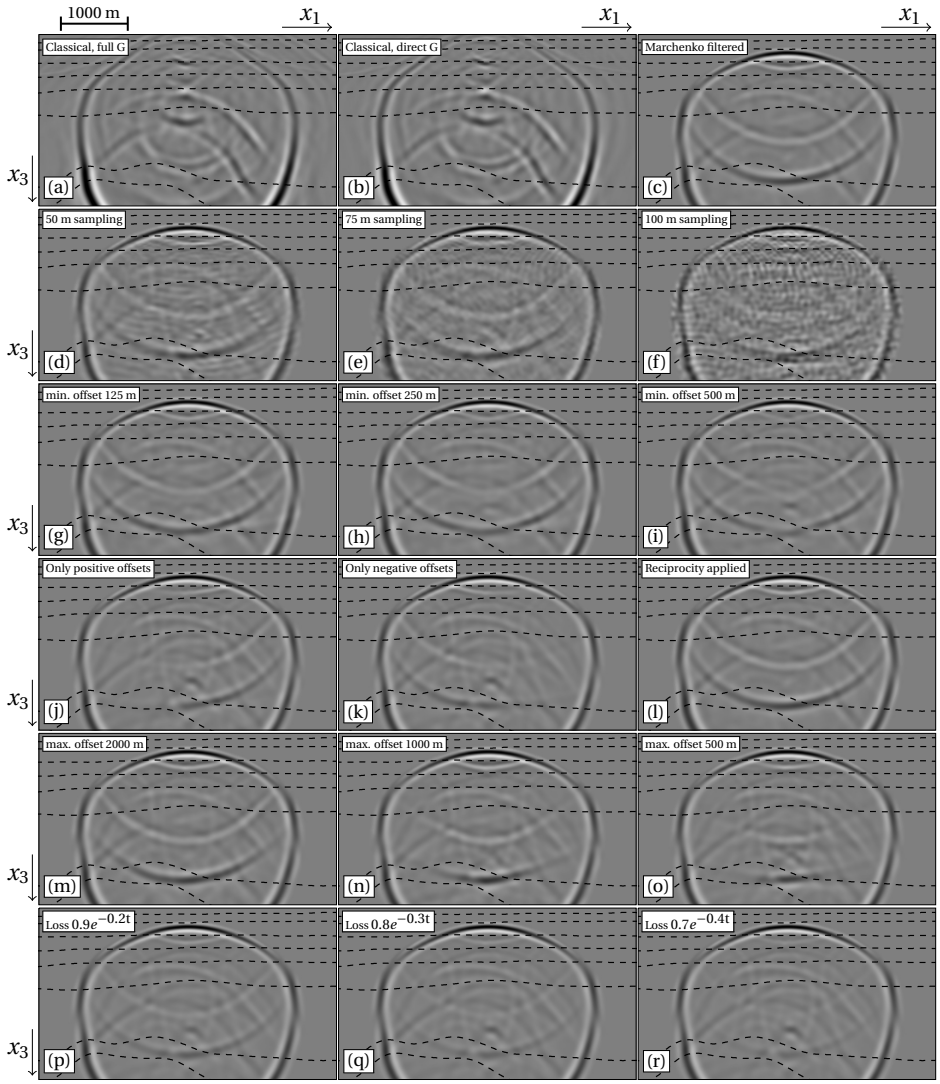

Figure 2.7: Snapshots of the homogeneous Green's function at t=400ms, retrieved using varying limitations of the reflection response, following the workflow in Figure 2.5(a). All wavefields have been convolved with a 15Hz Ricker wavelet. (a) Result from Figure 2.6(f), (b) result from Figure 2.6(i) and (c) result from Figure 2.6(o), the latter used as a benchmark for the other results. Second row: Homogeneous Green's function obtained, when the reflection response has a source and receiver spacing of (d) 50m, (e) 75m and (f) 100m. Third row: Homogeneous Green's function obtained when the reflection response is missing the small source-receiver offsets up to a distance of (g) 125m, (h) 250m and (i) 500m. Fourth row: Homogeneous Green's function obtained when the reflection response contains (j) only positive source-receiver offsets, (k) only negative source-receiver offsets and (l) has all source-receiver offsets restored using source-receiver reciprocity. Fifth row: Homogeneous Green's function obtained when the reflection response has an aperture limited to (m) 2000m, (n) 1000m and (o) 500m. Sixth row: Homogeneous Green's function obtained when the reflection response has a loss applied to it of (p) $0.9e^{-0.2t}$, (q) $0.8e^{-0.3t}$ and (r) $0.7e^{-0.4t}$.

complications due to out of plane effects.

The effects of having geometric spreading in a 3D setting while the Marchenko scheme that is applied is 2D, are corrected by applying a time-dependent gain on the data. This gain is equal to $\sqrt{t}$, as the geometrical spreading in 3D can be approximated as being t^{-1} and in 2D as $t^{-\frac{1}{2}}$ [72]. Next, source-receiver reciprocity is applied to the reflection data to create offsets in the positive direction. Having offsets in both positive and negative directions is vital for the next step, where the Estimation of Primaries through Sparse Inversion (EPSI) method is applied. The EPSI method estimates the primaries in the data, through the use of an inversion process, and estimates the free-surface multiples separately. This allows for the retrieval of a datset without the free-surface multiples. Furthermore, in this process, the information about the subsurface that is contained in the free-surface multiples is used to reconstruct the missing near offsets of the primaries. Simultaneously, an estimation of the source wavelet is made, which can be used to deconvolve the reflection data. Note, that the EPSI method only removes the free-surface multiples. The internal multiples remain in the data after the application of the EPSI method. See van Groenestijn and Verschuur [73] for a detailed overview of the EPSI method.

The effects of absorption on the data is adaptively corrected for by applying exponential time-gain and a scaling factor. Note that the exponential time-gain is applied in addition to the $\sqrt{t}$ gain that was applied to account for the geometrical spreading factor. This second gain is intended to counteract the effects of the absorption on the data and therefore is estimated separately from the first gain. The first estimations of the time-gain are based on the velocity model conforming to the method discussed by Draganov *et al.* [74]. The scaling factor is estimated by minimizing the cost functions proposed by Brackenhoff [66]. After these processing steps are applied, the reflection data are used to retrieve a Green's function for one location in the subsurface. If the method does not converge to a solution, where the artifacts are minimal, the exponential gain and scaling factor are adjusted and the test is run again. After a few iterations, we found that applying a gain of $1.73e^{1.3t}$ to the data resulted in a solution that converged with significant removal of artifacts.

Interpolation was also tested by interpolating the source-receiver spacing on the reflection data to smaller values, however we found that this did not significantly improve our results. Another aspect we cannot improve is the limited aperture of the data. The final workflow for the field data can be found in Figure 2.5(b). An example of a common-source record before and after the pre-processing is shown in Figure 2.8. Note the improvement of the common-source record, as the coverage of the data has been increased and the free-surface multiples have been weakened, as indicated by the black arrows.

2.5.2. HOMOGENEOUS GREEN'S FUNCTION RETRIEVAL

After applying all the corrections, the Marchenko method is utilized to retrieve the required Green's functions and focusing functions in the region of interest that is indicated in Figure 2.3(b). These functions are retrieved using only the pre-processed single-sided reflection response and a smooth velocity model. Next, we use these functions to retrieve the homogeneous Green's function using the single-sided boundary representation of Equation (2.10). Snapshots of the obtained result are shown in Figures 2.9(e)-(h), for

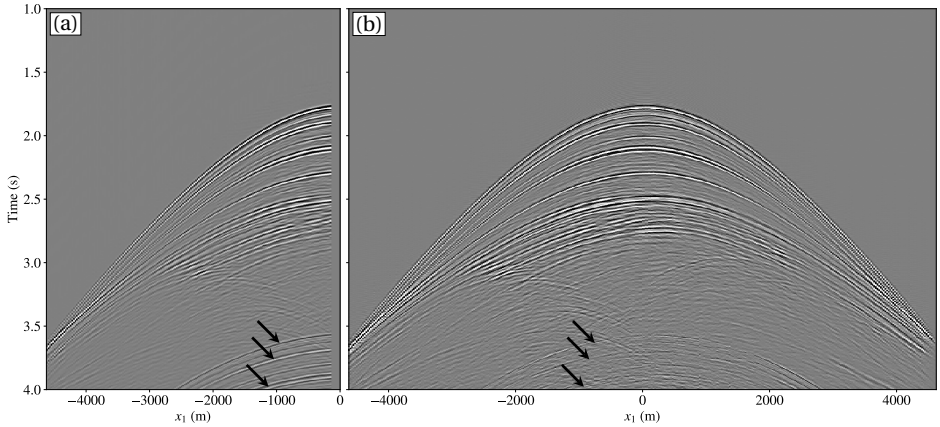

Figure 2.8: (a) Common-source record from Figure 2.3(a) before any processing is applied and (b) common-source record from (a) with source-receiver reciprocity, EPSI and an exponential gain of $1.73e^{1.3t}$ applied. Both common-source records have their wavelets reshaped to a 30Hz Ricker wavelet. Note the removal of the free-surface multiples at late times, as indicated by the black arrows.

0ms, 300ms, 600ms and 900ms, respectively. For comparison, the homogeneous Green's function obtained using the classical representation and using only the first arrival for the Green's function for the virtual source is shown in Figures 2.9(a)-(d), for 0ms, 300ms, 600ms and 900ms, respectively. In both cases a monopole source was used to model all the first arrivals.

We also test the creation of a virtual source that has a double-couple source signature on the field data. We use a double-couple source inclined at -20 degrees to create the first arrival that is used to obtain the Green's function for the virtual source using the Marchenko method. Similar to our approach on the synthetic data, we use the smooth P-wave velocity model, a homogeneous density model of 1000kg m^{-3} and a constant S-wave velocity model of 1000m s^{-1}, with the top layer of water, where the S-wave velocity is zero. We select the P-wave first arrival and use the Marchenko method to create the Green's function for the virtual source. The homogeneous Green's function is obtained through Equation (2.15) and the workflow in Figure 2.5(b). Snapshots of the result are shown in Figure 2.9(i)-(l) for 0ms, 300ms, 600ms and 900ms, respectively.

For all images, a transparent overlay of the image from Figure 2.3(c) is used to indicate locations where scattering is expected. This image is only used for verification and was not used for the retrieval of the homogeneous Green's function. The results of the single-sided representation for the monopole source from Figure 2.9(e)-(h) were previously shown in Wapenaar *et al.* [61], and the results of the single-sided representation for the double-couple source from Figure 2.9(i)-(l) were previously shown in Brackenhoff *et al.* [62]. The details about the pre-processing of the reflection data were not discussed before.

The snapshots of the homogeneous Green's functions obtained using the single-sided representation for both types of sources show multiple events, both upgoing and down-

going. The locations of the scattering and the layer interfaces in the image overlay have a strong match and, aside from the primary reflections, the multiple reflections can also be seen. All of these events are completely absent when the classical representation is used. Strong artifacts are present in this case and the coda of the downgoing wavefield is missing entirely. The primary downgoing wavefield is present, however the upgoing primary wavefield is absent, which is similar to the result that was obtained on the synthetic data. This shows that, as expected, the single-sided representation is an improvement over the classical representation on field data as well. The results for the double-couple source signature are similar to the ones that were obtained on the synthetic data. The polarity changes are present on both the primary wavefield as well as on the coda. Because of these changes there are a few events in the coda where the amplitude is low, which makes it harder to distinguish these events. The results in Figure 2.9 are different from those that were obtained in Figure 2.6. As stated in the Datasets section, this is because the synthetic data do not contain all the events that are present in the field data. The results on the field data contain more events, which could not be predicted from the image. This is an advantage of the data-driven approach of the Marchenko method. We find the overall results in Figure 2.9 encouraging.

However, even the homogeneous Green's functions obtained using the single-sided representation are not perfect as there are still artifacts present. This is partially due to the presence of background noise in the dataset, which distorts the final result. More coherent events are also present, which do not correlate with the primary wavefield and scattering locations from the image. Because we cannot be sure whether the reflection response has been pre-processed perfectly, there may be some low amplitude artifacts present that are nor fully compensated for or which are created by the Marchenko method. The Marchenko method that we applied was intended for 2D acoustic media, however, the true medium is 3D and elastic. As the geological layering for this region appears to be close to horizontal, the out-of-plane effects are assumed to be small. Due to conversion from P-waves to S-waves and back, some events are present in the reflection response that would not be present if the medium was purely acoustic and these are not handled correctly by our acoustic Marchenko implementation.

2.5.3. DISCUSSION

The results on the field data show promise for further applications of the single-sided representation, for example, by combining them with passive measurements. In this case, the Green's function for the virtual source, that was constructed through the use of the Marchenko method, would be replaced with a passive recording. The accurately retrieved propagation and scattering of the wavefield in the inhomogeneous medium also holds much promise, despite the presence of some artifacts. If we wanted to track the paths of the wavefronts emitted for example by an induced seismic source, a result like this could provide substantial insight. However, there are several limitations. In this chapter, we exclusively considered a 2D dataset that has few out-of-plane effects. For complex 3D media, where the out-of-plane effects are much more severe, 3D reflection data would be required, with sufficient coverage, as well as a 3D smooth P-wave velocity model. In this case, the issues with offset, sampling and aperture that were present for the 2D data would be more complex to deal with, and will be present in two directions.

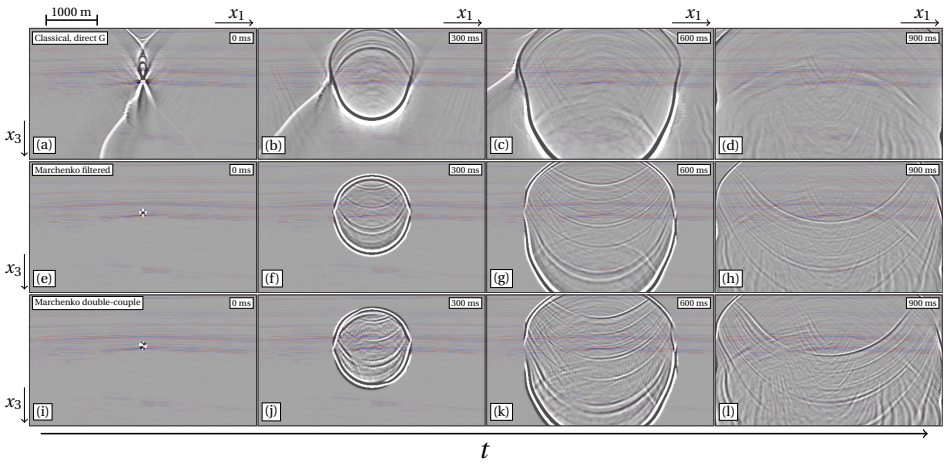

Figure 2.9: Snapshots of the homogeneous Green's function in the subsurface of the Vøring basin, obtained using the classical representation of Equation (2.9) for a monopole source at (a) 0ms, (b) 300ms, (c) 600ms and (d) 900ms. Idem, using the single-sided representation of Equation (2.10) for a monopole source at (e) 0ms, (f) 300ms, (g) 600ms and (h) 900ms. Idem, using the single-sided representation of Equation (2.15) for a double-couple source inclined at -20 degrees at (i) 0ms, (j) 300ms, (k) 600ms and (l) 900ms. All data contain a 30Hz Ricker wavelet.

Furthermore, passive measurement setups are usually sparser than for active measurements, which could yield additional complications if one wants to use a passive recording for the source of the homogeneous Green's function. These aspects are subject of ongoing research.

2.6. CONCLUSION

We demonstrated the data-driven generation of virtual receivers and a virtual source, which have the potential to improve, for example, the monitoring of the subsurface and the prediction of the complex response of different source mechanisms. We did this by utilizing a single-sided representation to retrieve the homogeneous Green's function in the subsurface. To this end, we applied the Marchenko method, which only requires a single-sided reflection response and a smooth velocity model. We showed that the single-sided representation produces significantly more accurate results than the classical representation when the reflection data are only available on a single-sided non-enclosing boundary, typically the Earth's surface. The sensitivity of the Marchenko method to limitations of the reflection data was investigated by manipulating synthetic reflection data. This showed that pre-processing of the reflection response to compensate for coarse source-receiver sampling, missing offsets and absorption is vital for the succesful application of the Marchenko method.

We considered a dataset from the Vøring basin (offshore Norway), which was affected by some of such limitations and processed the data using geometric spreading correction, source-receiver reciprocity, the EPSI method and applying a time-gain and scaling factor. The processed reflection response was used to obtain the focussing functions

and Green's functions, needed to apply the representations for homogeneous Green's function retrieval. The homogeneous Green's function obtained using the single-sided representation shows potential in our opinion for wavefield monitoring in the subsurface, as the complete coda of the wavefield is recovered, which is not the case when the classical representation is used. The scattering occurs at locations that correlate with possible reflector locations. Monopole source and double-couple source signatures can both be used in the Marchenko method and in the single-sided representation to obtain homogeneous Green's functions with the same source signature. To further explore this potential, more complex source mechanisms should be considered, such as dynamic fault planes, that are active over an extended area and time period. This includes taking into account the effects caused by elastic media instead of acoustic media.

3

MONITORING OF INDUCED DISTRIBUTED DOUBLE-COUPLE SOURCES USING MARCHENKO-BASED VIRTUAL RECEIVERS

We aim to monitor and characterize signals in the subsurface by combining these passive signals with recorded reflection data at the surface of the Earth. To achieve this, we propose a method to create virtual receivers from reflection data using the Marchenko method. By applying homogeneous Green's function retrieval, these virtual receivers are then used to monitor the responses from subsurface sources. We consider monopole point sources with a symmetric source signal, where the full wavefield without artifacts in the subsurface can be obtained. Responses from more complex source mechanisms, such as double-couple sources, can also be used and provide results with comparable quality as the monopole responses. If the source signal is not symmetric in time, our technique that is based on homogeneous Green's function retrieval provides an incomplete signal, with additional artifacts. The duration of these artifacts is limited and they are only present when the source of the signal is located above the virtual receiver. For sources along a fault rupture, this limitation is also present and more severe due to the source activating over a longer period of time. Part of the correct signal is still retrieved, as well as the source location of the signal. These artifacts do not occur in another method which creates virtual sources as well as receivers from reflection data at the surface. This second method can be used to forecast responses to possible future induced seismicity sources (monopoles, double-couple sources and fault ruptures). This method is applied to field data, where similar results to synthetic data are achieved, which shows the potential for the application on real data signals.

This chapter was published as J. Brackenhoff, J. Thorbecke, & K. Wapenaar, *Monitoring of induced distributed double-couple sources using Marchenko-based virtual receivers.* Solid Earth, **10**, 1301-1319 (2019).
Minor modifications have been applied to the text and figures for the sake of consistency in the thesis.

3.1. INTRODUCTION

Seismic monitoring of processes in the subsurface has been an active field of research for many years. Traditionally, most recording setups are limited to the surface of the Earth, although boreholes can also be utilized. The latter approach is more expensive and complicated, however. In case of monitoring with active sources, the receivers in these recording setups measure valuable reflection data, which provide quantifiable information about processes in the subsurface. Some examples of using this information are monitoring time-shifts in seismic data to predict the velocity-strain relation for a depleting reservoir [75] and the monitoring of geomechanics in the subsurface by using time-lapse data [76]. The responses from passive sources, such as when the signal is caused by an induced earthquake, can be measured as well. These passive measurements are more difficult to process due to the fact that the signal is complex and unknown [77], however, the information content in these induced seismic signals is of great interest. Induced seismicity has had a large impact in countries such as the Netherlands [56] and the USA [57] and there is much discussion about the cause and the effects. To determine the cause of induced seismicity, the source of the signal is of particular interest and consequently, inversions for the source mechanism [78] as well as the location of the source [79] are often performed. These methods can be carried out from surveys that are located at the surface of the Earth or inside boreholes, however, they are limited in accuracy. Ideally, one would use a dense network of receivers around the source location to directly monitor the wavefield.

Due to practical difficulties and expenses associated with placing a dense network of receivers in the subsurface, the wavefield can generally not be directly measured around the source location of the signal. An alternative to using physical receivers for these measurements is the use of virtual receivers. A virtual receiver is not physically present in the subsurface, rather, it is created through processing of measured signals at the surface. Virtual receivers can be created in a variety of ways. A mathematical basis for the retrieval of these virtual receivers is the so-called homogeneous Green's function representation. The classical form of this representation was proposed by Porter [15] and extended for inverse source problems by Porter and Devaney [16] and for inverse scattering methods by Oristaglio [17]. This representation states that if the responses from two signals are measured on an enclosing recording surface, the response between the two sources of the signals can be retrieved. It forms the basis for seismic interferometry to create virtual sources [80] or virtual receivers [23]. All of these approaches require access to the medium from an enclosing surface and introduce artifacts if this requirement is not met. Even though this limitation is well known, for many cases these approaches are still utilized.

A novel approach that can be used when the acquisition surface is not closed is the data-driven 3D Marchenko method. This method can create virtual sources and receivers in the subsurface [26, 30]. In order to achieve this, the method requires a reflection response recorded at the surface of the Earth, and an estimation of the first arrival of the signal from a location in the subsurface to the receiver locations in the measurement array. This first arrival can be estimated from a background velocity model, which requires no detailed information about the subsurface. Through the Marchenko method, the Green's function with a virtual receiver in the subsurface can be retrieved. Using this

method, many virtual receivers can be created in the subsurface, which can be used to monitor the wavefield from the virtual receiver locations to the receiver array. To obtain the signal between an induced signal from the subsurface and the virtual receiver locations, homogeneous Green's function retrieval can be employed, however, as pointed out before, the classical approach would include artifacts due to the open surface of the recording. These artifacts can disturb the interpretation of the signal. An alternative retrieval scheme was developed by Wapenaar *et al.* [25], who showed that if a focusing function is used in combination with a Green's function, an open surface can be used for the retrieval instead of an enclosing one, without the artifacts of the classical method when applied to an open surface. A focusing function is a wavefield that is designed to focus at a location in the subsurface and can be retrieved from reflection data using the Marchenko method [36]. This single-sided representation has been proven to work succesfully on both synthetic data and on field data [81].

Using the single-sided method, two approaches for monitoring induced seismicity can be taken. First, virtual receivers can be used in combination with a virtual source. In this case, all the signals are created from the reflection data using the Marchenko method. This has the benefit that the virtual source can be created at any location in the subsurface, where one expects induced seismicity to happen, and that the source signal can be controlled. This is the way that the method has been mostly applied in previous works. Another approach that can be taken is to create virtual receivers using the Marchenko method and to use a real induced seismic source signal instead of a virtual Green's function. This effectively allows for the monitoring of the actual signal in the subsurface, including the source location and mechanism. This could be a boon to induced seismicity monitoring, however, this approach does require some modifications. Induced seismicity often causes more complex source signals that evolve over a period of time and cover an extended area in the subsurface. These rupture planes or fault sources are the main topic of interest.

In this work, we aim to apply the single-sided homogeneous Green's function retrieval on both synthetic and field data for a distribution of virtual double-couple sources. We first apply the method on synthetic data for point sources and show the principles of the representation. We then use the same synthetic data to apply the representation with modifications to the sources originating from a fault plane and show the results that can be achieved. Finally, we also apply the representation on field data for both types of sources.

3.2. THEORY

3.2.1. GREEN'S FUNCTION AND FOCUSING FUNCTION

In this chapter, we present several representations for the retrieval of wavefields in the subsurface. First, we review the properties and quantities that are relevant for these representations. To this end, we consider a medium that is acoustic, lossless and inhomogenous with mass density $\rho = \rho(\mathbf{x})$ and compressibility $\kappa = \kappa(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, x_3)$ indicates the Cartesian coordinate vector. We make use of a Green's function in this medium that obeys the following wave equation:

$$\partial_i(\rho^{-1}\partial_i G) - \kappa\partial_t^2 G = -\delta(\mathbf{x}-\mathbf{x}_A)\partial_t\delta(t), \tag{3.1}$$

where $G(\mathbf{x},\mathbf{x}_A,t)$ indicates a Green's function that at time t describes the response of the medium at location $\mathbf{x}$ due to an unit impulsive point source of volume-injection rate density $\delta(\mathbf{x}-\mathbf{x}_A)\delta(t)$ at source location $\mathbf{x}_A$. $\delta(\cdot)$ is the Dirac delta function, ∂_t the temporal partial differential operator $\frac{\partial}{\partial t}$ and ∂_i a component of a vector containing the spatial partial differential operators in the three principal directions $\left(\frac{\partial}{\partial x_1},\frac{\partial}{\partial x_2},\frac{\partial}{\partial x_3}\right)$. Einstein's summation convention applies to repeated subscripts. The Green's function obeys source-receiver reciprocity, which allows the interchange of the source and receiver position, hence $G(\mathbf{x}_B,\mathbf{x}_A,t)=G(\mathbf{x}_A,\mathbf{x}_B,t)$. We impose causality on the Green's function, $G(\mathbf{x},\mathbf{x}_A,t)=0$ for $t<0$, such that it is forward propagating, originating from the source, and a causal solution to Equation (3.1). A schematic illustration of the Green's function is shown in Figure 3.1(a), where several possible raypaths are drawn for a heterogeneous model. This includes the direct arrival, primary reflections and multiple reflections.

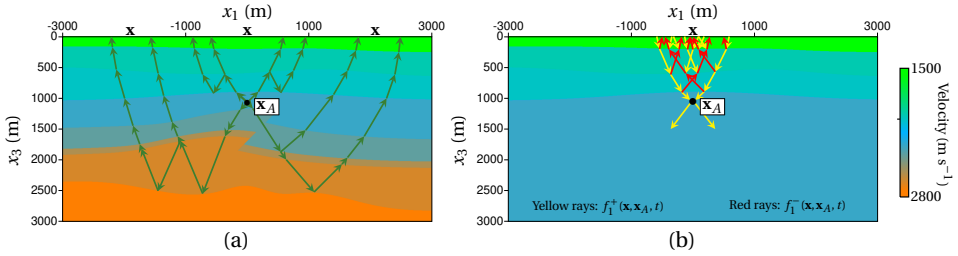

Figure 3.1: (a) Schematic representation of the Green's function $G(\mathbf{x},\mathbf{x}_A,t)$, defined in the physical medium, with a source located at $\mathbf{x}_A$, which is measured at varying location $\mathbf{x}$ at the surface. (b) Schematic representation of the focusing function $f_1(\mathbf{x},\mathbf{x}_A,t)$, defined in the truncated medium, where the wavefield propagates from $\mathbf{x}$ at the surface to the focal location $\mathbf{x}_A$. For both functions, several possible raypaths are drawn. For the focusing function the downgoing waves are marked with yellow arrows and the upgoing waves with red arrows.

We also consider the time-reversed Green's function $G(\mathbf{x},\mathbf{x}_A,-t)$, which is the acausal solution to Equation (3.1), where the causality condition implies $G(\mathbf{x},\mathbf{x}_A,-t)=0$ for $t>0$. Superposition of the causal and acausal Green's function yields the homogeneous Green's function:

$$G_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,t)=G(\mathbf{x},\mathbf{x}_A,t)+G(\mathbf{x},\mathbf{x}_A,-t), \tag{3.2}$$

where $G_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,t)$ obeys the homogeneous wave equation:

$$\partial_i(\rho^{-1}\partial_i G_{\mathrm{h}})-\kappa\partial_t^2 G_{\mathrm{h}}=0. \tag{3.3}$$

Equation (3.3) is similar to Equation (3.1), with the exception of the lack of a source singularity on the right hand side of the equation.

Aside from the Green's function, we consider the focusing function $f_1(\mathbf{x},\mathbf{x}_A,t)$, which describes a wavefield, at time t and location $\mathbf{x}$, that converges to a focal location $\mathbf{x}_A$ in the subsurface of a medium that is truncated below the focal location. The focusing function can be decomposed as,

$$f_1(\mathbf{x},\mathbf{x}_A,t)=f_1^+(\mathbf{x},\mathbf{x}_A,t)+f_1^-(\mathbf{x},\mathbf{x}_A,t), \tag{3.4}$$

where $f_1^+(\mathbf{x},\mathbf{x}_A,t)$ denotes the downgoing and $f_1^-(\mathbf{x},\mathbf{x}_A,t)$ the upgoing component of the focusing function. A schematic representation of the focusing function can be found in Figure 3.1(b). Similar to the Green's function, several possible raypaths are drawn, however, to distinguish the decomposed wavefields, the downgoing focusing function is marked with yellow rays and the upgoing focusing function with red rays. The medium of the focusing function and the Green's function are identical until the focal depth, after which the medium of the focusing function becomes truncated. The physical and truncated medium can be used in reciprocity theorems in order to relate the focusing function to the Green's function, which is shown in section 2 of the supplementary information. For moderately inhomogeneous media, the focusing function and Green's function can be separated from each other in time. The coda of the focusing function resides in the interval between the direct arrival of a related Green's function and its time reversal. The direct arrival of the focusing function coincides with the direct arrival of the time reversed Green's function. This difference in time intervals explains some of the effects that are present in the representations that are used in this chapter. Both the focusing function and Green's function can be retrieved for a heterogeneous medium from the reflection data and an estimate of the direct arrival, through use of the Marchenko method. We will not explain this method in detail in this chapter, instead we refer the reader to Wapenaar *et al.* [26] for a more detailed overview.

Due to the nature of some equations, we also make use of the frequency domain version of the time domain quantities. To obtain these transformation we make use of the Fourier transform. We define the Fourier transform of a space- and time-dependent function $u(\mathbf{x},t)$ as

$$u(\mathbf{x},\omega) = \int_{-\infty}^{\infty} u(\mathbf{x},t)\exp(i\omega t)\mathrm{d}t, \tag{3.5}$$

where $u(\mathbf{x},\omega)$ is the Fourier transformed version of $u(\mathbf{x},t)$ in the space-frequency domain, with ω as the angular frequency and i the imaginary unit. By using Equation (3.5) we obtain the space-frequency domain versions of Equation (3.1), (3.2), (3.3) and (3.4), respectively:

$$\partial_i(\rho^{-1}\partial_i G) + \kappa\omega^2 G = i\omega\delta(\mathbf{x}-\mathbf{x}_A), \tag{3.6}$$

$$G_\mathrm{h}(\mathbf{x},\mathbf{x}_A,\omega) = G(\mathbf{x},\mathbf{x}_A,\omega) + G^*(\mathbf{x},\mathbf{x}_A,\omega) = 2\Re\{G(\mathbf{x},\mathbf{x}_A,\omega)\}, \tag{3.7}$$

$$\partial_i(\rho^{-1}\partial_i G_\mathrm{h}) + \kappa\omega^2 G_\mathrm{h} = 0, \tag{3.8}$$

$$f_1(\mathbf{x},\mathbf{x}_A,\omega) = f_1^+(\mathbf{x},\mathbf{x}_A,\omega) + f_1^-(\mathbf{x},\mathbf{x}_A,\omega), \tag{3.9}$$

where $\Re$ indicates the real part of a complex function.

3.2.2. HOMOGENEOUS GREEN'S FUNCTION REPRESENTATION

The classical homogeneous Green's function representation was originally developed for a configuration where the Green's function was measured on an arbitrarily shaped surface enclosing the medium of interest [15–17]. The representation states that, if the responses from two sources inside the medium are recorded on the surface, the response

between the two source locations can be obtained. For seismic recording setups, the measurements are usually only available at the surface of the Earth, meaning that the surface is single-sided instead of closed, which will introduce significant errors into the final result.

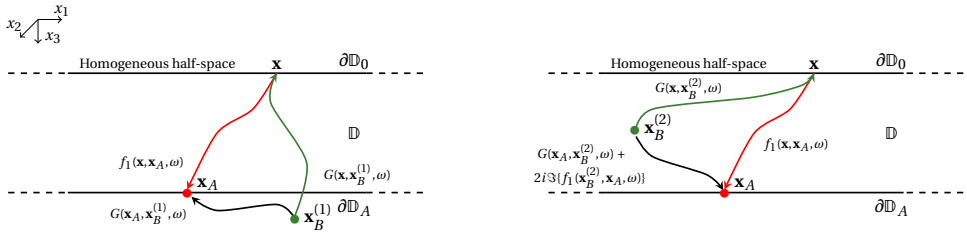

Figure 3.2: Setup for the single-sided Green's function representation for (a) a case where the source of the Green's function is located below the focal location and (b) a case where the source of the Green's function is located above the focal location. The rays in this figure indicate full Green's functions and focusing functions, including multiple scattering.

In recent years a new representation for homogeneous Green's function retrieval was developed that is designed to work with the single-sided surface, where a focusing function is used together with a Green's function [25]. Consider the setup in Figure 3.2, where a heterogeneous medium $\mathbb{D}$ is bounded by two horizontal surfaces $\partial\mathbb{D}_0$ and $\partial\mathbb{D}_A$ on two different levels in vertical direction x_3. The surfaces extend infinitely in the horizontal directions x_1 and x_2. The medium above $\partial\mathbb{D}_0$ is homogeneous, with mass density ρ_0 and compressibility κ_0, and the surface itself is non-reflecting, while the medium below $\partial\mathbb{D}_A$ can be heterogeneous. The upper surface $\partial\mathbb{D}_0$ corresponds to the surface where the receiver locations $\mathbf{x}$ of focusing functions and Green's functions are available. In this scenario, we assume that we have three functions available at the upper surface, a Green's function $G(\mathbf{x},\mathbf{x}_B^{(1)},\omega)$, that has a source location $\mathbf{x}_B^{(1)}$ below $\partial\mathbb{D}_A$, a Green's function $G(\mathbf{x},\mathbf{x}_B^{(2)},\omega)$, that has a source location $\mathbf{x}_B^{(2)}$ inside medium $\mathbb{D}$ and a focusing function $f_1(\mathbf{x},\mathbf{x}_A,\omega)$, that has a focal location $\mathbf{x}_A$, located at the depth of $\partial\mathbb{D}_A$.

The available functions can be used to obtain the response between two locations. To this end, we use the representation given by equation (35) of the supplementary information (for the derivation see section 2.3 of the supplementary material),

$$G(\mathbf{x}_A,\mathbf{x}_B,\omega)+\chi(\mathbf{x}_B)2i\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\} = \int_{\partial\mathbb{D}_0}\frac{2}{i\omega\rho_0}G(\mathbf{x},\mathbf{x}_B,\omega)\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}\mathbf{x}, \tag{3.10}$$

where $\Im$ is the imaginary part of a complex function and $\chi(\mathbf{x}_B)$ is the characteristic function,

$$\chi(\mathbf{x}_B)=\begin{cases}1, & \text{for } \mathbf{x}_B \text{ in } \mathbb{D},\\ \frac{1}{2}, & \text{for } \mathbf{x}_B \text{ on } \partial\mathbb{D}=\partial\mathbb{D}_0\cup\partial\mathbb{D}_A,\\ 0, & \text{for } \mathbf{x}_B \text{ outside } \mathbb{D}\cup\partial\mathbb{D}.\end{cases} \tag{3.11}$$

This representation states that, by applying the focusing function components to a Green's function at the upper surface, the Green's function between the focal location $\mathbf{x}_A$ of the focusing function and the source location $\mathbf{x}_B$ of the Green's function can be obtained. The focal location will become the receiver of this new Green's function, and the source location of the original Green's function on the right hand side of Equation (3.10) will become the source location of the new Green's function. However, contributions from the imaginary part of the focusing function between the source and receiver locations are present if the source location is located inside the medium $\mathbb{D}$, as is the case if the Green's function from Figure 3.2(b) with source location $\mathbf{x}_B^{(2)}$ is used. Because they are related to a focusing function, these artifacts will be present between the direct arrival of the Green's function and its time reversal. In this case, the source location is present above the focal location. These contributions vanish if the source location is present outside $\mathbb{D}$, in other words if it is located below the focal location, such as when the Green's function from Figure 3.2(a) with source location $\mathbf{x}_B^{(1)}$ is used. This would mean that, without knowledge of $\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\}$, we are limited in the correct application of the representation. To overcome this limitation, we substitute Equation (3.10) into the right hand side of Equation (3.7) to create the single-sided homogeneous Green's function representation:

$$G_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega) = 4\Re\int_{\partial\mathbb{D}_0}\frac{1}{i\omega\rho_0}G(\mathbf{x},\mathbf{x}_B,\omega)\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega) - \{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}^2\mathbf{x}, \tag{3.12}$$

which corresponds to equation (33) from our companion paper [24]. The additional contributions have vanished from this representation and the homogeneous Green's function will be obtained when it is evaluated, instead of the causal Green's function.

3.2.3. VIRTUAL SOURCES AND RECEIVERS

Generally, the focusing function and Green's function are not directly available. These functions can be obtained through the use of the Marchenko method [26, 29, 36], which is a data-driven method that requires only reflection data at the surface of the Earth and an estimation of the first arrival of the wavefield at the location of interest inside the medium. The method handles the primaries of the reflection data in the same way as conventional methods, however, unlike those methods, the Marchenko method can also correctly handle the multiples in the data. The first arrival can be estimated through the use of a macro-velocity model. The method cannot handle attenuation on the reflection data and ignores evanescent waves. On field data, the data requires additional processing to account for these and other requirements. The Marchenko method has been applied succesfully on both synthetic and field data, for examples see Ravasi *et al.* [37], Staring *et al.* [38] and Brackenhoff *et al.* [81].

The method can be used in the homogeneous Green's function retrieval scheme in two ways, which are schematically shown in Figure 3.3. The first approach is a two-step process, as shown in Figure 3.3(a), where both the source and receiver of the homogeneous Green's function are obtained by redatuming them from the reflection response. This type of source-receiver redatuming is discussed in section 3.4 of our companion paper by Wapenaar *et al.* [24]. First, we consider the fact that the data that we use in the field is bandlimited and therefore a source signal $s(t)$ is convolved with the Green's

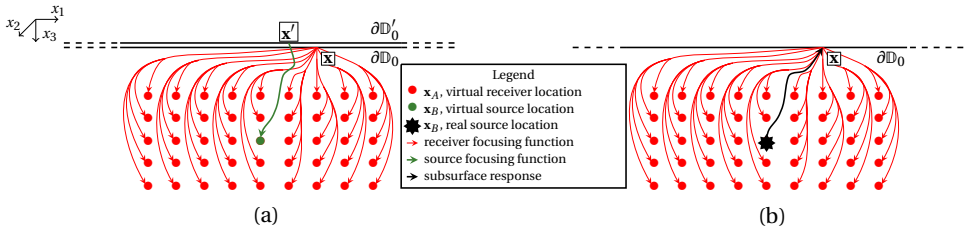

Figure 3.3: Schematic setup for (a) the two step process and (b) the one step process for retrieving the homogeneous Green's function in the subsurface. The red and green arrows show the focusing functions that are used to respectively create the virtual receiver and virtual source location. The red and green dots show the locations for the virtual receiver and virtual source, respectively. The black star indicates the source location of a real subsurface response, indicated with a black arrow, that is measured at the surface $\partial\mathbb{D}_0$ on the same receiver location $\mathbf{x}$ as the focusing and Green's functions. $\partial\mathbb{D}_0$' is a surface located just above $\partial\mathbb{D}_0$ on which the source locations $\mathbf{x}$' of the reflection response $p(\mathbf{x},\mathbf{x}',\omega)$ are located. The rays in this figure indicate full Green's functions and focusing functions, including multiple scattering.

function, which changes its phase and amplitude:

$$p(\mathbf{x},\mathbf{x}_B,t) = \int_{-\infty}^{\infty} G(\mathbf{x},\mathbf{x}_B,t-t')s(t')\mathrm{d}t', \tag{3.13a}$$

$$p(\mathbf{x},\mathbf{x}_B,\omega) = G(\mathbf{x},\mathbf{x}_B,\omega)s(\omega), \tag{3.13b}$$

where $p(\mathbf{x},\mathbf{x}_B,t)$ is a pressure wavefield in the medium and $s(\omega)$ is the Fourier transform of the source signal. For the first step, we introduce a second surface $\partial\mathbb{D}_0'$ that is located just above $\partial\mathbb{D}_0$ and assume that a reflection response $p(\mathbf{x},\mathbf{x}',\omega)$ of the medium has been measured, where $\mathbf{x}'$ is the source location on the surface $\partial\mathbb{D}_0'$. The reflection response is used to create a virtual source location in the subsurface. To this end, we utilize a modification of Equation (3.12), and use Equation (3.13b) to create an equivalent version for pressure wavefields, which is the same as equation (41) of our companion paper:

$$p(\mathbf{x},\mathbf{x}_B,\omega)+p^*(\mathbf{x},\mathbf{x}_B,\omega) = 4\Re\int_{\partial\mathbb{D}_0'}\frac{1}{i\omega\rho_0}p(\mathbf{x},\mathbf{x}',\omega)\partial_3\Big(f_1^+(\mathbf{x}',\mathbf{x}_B,\omega)-\{f_1^-(\mathbf{x}',\mathbf{x}_B,\omega)\}^*\Big)\mathrm{d}\mathbf{x}'. \tag{3.14}$$

In Equation (3.14), we assume that the source spectrum is strictly real-valued. The focusing function $f_1(\mathbf{x}',\mathbf{x}_A,\omega)$ is obtained through use of the Marchenko method and employed in Equation (3.14) to create a wavefield with a virtual source location, which is indicated by the green line in Figure 3.3(a). This function will be used to create a source location for the wavefield retrieved through the homogeneous Green's function representation. This source is called a virtual source because it is not physically present in the subsurface.

In the second step of the process, using the Marchenko method, many focusing functions are created for focal points at varying locations in the medium, that serve as the virtual receiver locations for the retrieved wavefield. This is indicated by the red dots and arrows in Figure 3.3(a). Similarly to the virtual source, these are called virtual receivers, again, because they are not physically present in the medium. We use these focusing

functions in Equation (3.10), which we modify using equation Equation (3.13b) as follows

$$p(\mathbf{x}_A,\mathbf{x}_B,\omega)+\chi(\mathbf{x}_B)2is(\omega)\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\} = \int_{\partial\mathbb{D}_0}\frac{2}{i\omega\rho_0}p(\mathbf{x},\mathbf{x}_B,\omega)\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}\mathbf{x}. \tag{3.15}$$

In this representation, we make use of the wavefield $p(\mathbf{x},\mathbf{x}_B,\omega)$ with the virtual source location that we obtained in the first step. The acausal part of the left hand side of the time-domain version of Equation (3.14) can be removed easily by applying causality through the use of a Heaviside function. Since we assumed $s(\omega)$ to be real-valued, substitution of Equation (3.15) into Equation (3.7) yields,

$$p_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega) = 4\Re\int_{\partial\mathbb{D}_0}\frac{1}{i\omega\rho_0}p(\mathbf{x},\mathbf{x}_B,\omega)\partial_3\big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\big)\mathrm{d}\mathbf{x}, \tag{3.16}$$

where $p_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega)=p(\mathbf{x}_A,\mathbf{x}_B,\omega)+p^*(\mathbf{x}_A,\mathbf{x}_B,\omega)$. This is a similar representation to equation (39) for modified back propagation from our companion paper by Wapenaar *et al.* [24].

The second way we can use the Marchenko method in the application of homogeneous Green's function retrieval is a one-step process, where the Marchenko method is only used to retrieve focusing functions to create virtual receivers. This is shown in Figure 3.3(b). Here, no virtual source is created from the reflection data using Equation (3.14), rather the actual response from a real source inside the medium is used, which is illustrated by the black star and arrow in Figure 3.3(b). The response that is monitored is used as $p(\mathbf{x},\mathbf{x}_B,\omega)$ in Equation (3.15). It can not generally be used in Equation (3.16), however. If the source spectrum of the response is not strictly real-valued, the signal is not symmetric in time, because $s(\omega)\neq s^*(\omega)$, and therefore there will be a phase difference between the causal and acausal wavefield, making the superposition of the signal with its time-reverse incorrect. Assuming that through processing of the signal, the type of wavelet that is applied to the data can be controlled, symmetry of the source signal can be ensured by using zero-phase wavelets. When this condition is fulfilled, Equation (3.16) can be used for the subsurface response. Monitoring real source signals is the eventual goal of this approach, such as for the case of induced seismicity. The boon of this method is that aside from the measured signal, no information about the source of the data is required. There are limitations to this approach as well, most pressing that to evaluate the integral, the signal needs to be recorded on the same receiver array that was used to record the reflection data.

3.2.4. MODIFICATIONS FOR REALISTIC INDUCED SEISMICITY SOURCES

DOUBLE-COUPLE POINT SOURCES

For the case of induced seismicity, the source signal can be more complex than just a single monopole point source. To include the mechanics for induced earthquakes more accurately, the double-couple source mechanism can be included in the representation. The double-couple source mechanism is accepted as representative for an

earthquake response if the wavelength of the signal is at least of the same dimension as the size of the fault that originated the earthquake [8]. It can be implemented through the use of a moment tensor, which is useful for the case of finite-difference modeling [71]. The response of a monopole source and double-couple source for a homogeneous medium is shown in Figure 3.4, along with their radiation patterns in the center. While the monopole source response has an uniform amplitude along the wavefront, the double-couple source response has a varying amplitude and polarity along the wavefront, due to the variation in the radiation pattern. Consequently, the orientation of the double-couple source affects the source signal, which is visible in the Figure 3.4(b), while the orientation of the monopole source does not matter. Hence, the orientation of the fault is crucial to the characteristics of the double-couple source signal. To include this orientation in the representation, we introduce the operator $\mathfrak{D}_B^\theta$, which acts on the wavefield and creates the double-couple source orientation from the monopole source signature. This operator is defined as

$$\mathfrak{D}_B^\theta = (\theta_i^\parallel + \theta_i^\perp)\partial_{i,B}, \tag{3.17}$$

where $\partial_{i,B}$ is a component of the vector containing the partial derivatives acting on the monopole signal originating from source location $\mathbf{x}_B$, that turns it into a double-couple source mechanism, $\theta_i^\parallel$ is a component of the unit vector that orients one couple of the signal parallel to the fault plane and $\theta_i^\perp$ is a component of the vector that orients the other couple perpendicular to the fault plane. The operator can be applied to Equation (3.15):

$$\begin{aligned}\mathfrak{D}_B^\theta\{p(\mathbf{x}_A,\mathbf{x}_B,\omega)\}+\mathfrak{D}_B^\theta\{\chi(\mathbf{x}_B)2is(\omega)\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\}\} = \\ \int_{\partial\mathbb{D}_0}\frac{2}{i\omega\rho_0}\mathfrak{D}_B^\theta\{p(\mathbf{x},\mathbf{x}_B,\omega)\}\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}\mathbf{x},\end{aligned} \tag{3.18}$$

and assuming that the source signal is symmetric in time, the operator is also applied to Equation (3.16)

$$\begin{aligned}\mathfrak{D}_B^\theta\{p_\mathrm{h}(\mathbf{x}_A,\mathbf{x}_B,\omega)\} = \\ 4\Re\int_{\partial\mathbb{D}_0}\frac{1}{i\omega\rho_0}\mathfrak{D}_B^\theta\{p(\mathbf{x},\mathbf{x}_B,\omega)\}\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}\mathbf{x}.\end{aligned} \tag{3.19}$$

In these two equations, the operator can be freely applied to both sides, because the integral is not evaluated over the source locations. Consequently, if the wavefield response used as a source for the homogeneous wavefield has a double-couple signature, the homogeneous wavefield will also have a double-couple signature. Note that the operator does not operate on the focusing functions, hence we can use the monopole responses for these signals.

DOUBLE-COUPLE SOURCES ALONG EXTENDED FAULTS

In case of induced seismicity, the fault or rupture plane that triggers the signal can be larger than the wavelength of the signal. In this case, the double-couple point source is no longer a valid approximation for the source of the signal. Studies of induced faults suggest that the signal develops over the fault during an extended period of time [82]. To

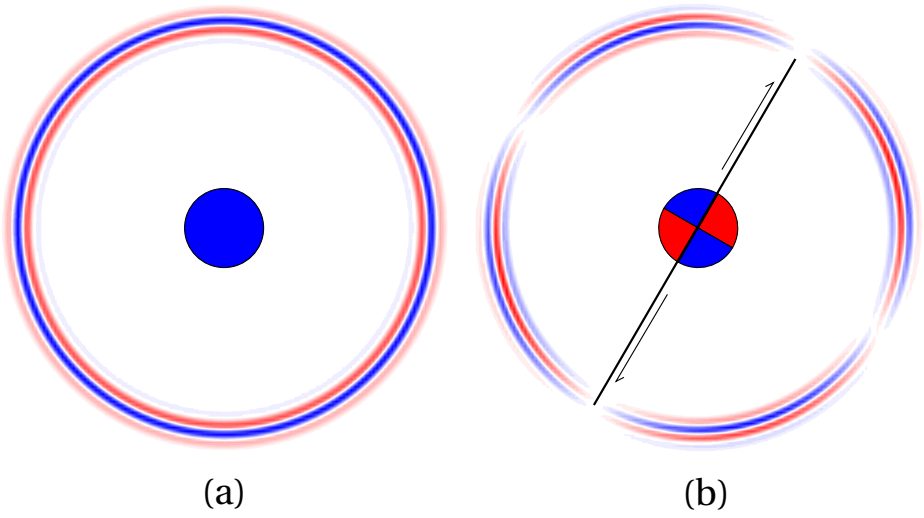

Figure 3.4: Comparison of the wavefields caused by (a) a monopole point source and (b) a double-couple point source tilted at an angle of 30 degrees. For both types of sources the radiation pattern of the source is shown in the center. The wavefields have been convolved with a 30Hz Ricker wavelet.

approximate this type of source, a superposition of many point sources can be utilized. The total signal of the resulting superposition can be written as the superposition of the individual signals,

$$P(\mathbf{x}_A,\omega) = \sum_{k=1}^{n_S} \mathfrak{D}_B^{\theta,(k)}\{p(\mathbf{x}_A,\mathbf{x}_B^{(k)},\omega)\} = \sum_{k=1}^{n_S} \mathfrak{D}_B^{\theta,(k)}\{G(\mathbf{x}_A,\mathbf{x}_B^{(k)},\omega)\, s^{(k)}(\omega)\}, \quad (3.20)$$

where the superscript k indicates the number of the source location $\mathbf{x}_B^{(k)}$ that has the source spectrum $s^{(k)}(\omega)$ and n_S is the total number of sources. The different source spectra include a linear phase term that determines the time at which the signal is triggered along the fault plane. $P(\mathbf{x}_A,\omega)$ can be created in two different ways, similar as before.

First, we consider the two-step process, where both the source and receiver are virtual. In this case, every source location can be treated separately to retrieve the homogeneous wavefield, and the superposition can be done after each signal has been retrieved through Equation (3.19) and then shifted over $t^{(k)}$,

$$P(\mathbf{x}_A,t) = \sum_{k=1}^{n_S} H(t-t^{(k)})\mathfrak{D}_B^{\theta,(k)}\{p_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B^{(k)},t-t^{(k)})\}, \quad (3.21)$$

where H is the Heaviside step function and $t^{(k)}$ is the time at which point the k-th signal originates on the fault. The Heaviside in Equation (3.21) selects the shifted causal signal from the shifted homogeneous (two-sided) signal before the superposition takes place, which is required to construct the correct signal. If the shifted homogeneous signals would be used instead, the shifted acausal part of later signals would overlap with the

causal part of signals that originated earlier. Through use of Equation (3.21) the correct signal can be retrieved.

In case the source signal is measured rather than virtually created, the same approach cannot be taken. This signal is by definition measured after superposition, therefore each point source cannot be evaluated seperately. To represent this, Equation (3.18) is adjusted to take the implicit superposition into account, according to

$$\begin{aligned} P(\mathbf{x}_A,\omega)+\sum_{k=1}^{n_S}\mathfrak{D}_B^{\theta,(k)}\{\chi(\mathbf{x}_B^{(k)})2is^{(k)}(\omega)\Im\{f_1(\mathbf{x}_B^{(k)},\mathbf{x}_A,\omega)\}\} = \\ \int_{\partial\mathbb{D}_0}\frac{2}{i\omega\rho_0}P(\mathbf{x},\omega)\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}\mathbf{x} = \\ \int_{\partial\mathbb{D}_0}\frac{2}{i\omega\rho_0}\sum_{k=1}^{n_S}\mathfrak{D}_B^{\theta,(k)}\{p(\mathbf{x},\mathbf{x}_B^{(k)},\omega)\}\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}\mathbf{x}. \end{aligned} \tag{3.22}$$

In this scenario, the sum is inside the integral and the entire signal is superposed before the focusing function is applied to it. This also results in a superposition of contributions of the focusing function between the virtual receiver location and the fault plane (i.e., the second term on the left-hand side). Substituting Equation (3.22) into Equation (3.7) will not lead to a cancellation of the focusing function on the left-hand side, as the wavefield does not have a symmetric source signal, due to the time differences between all the sources. As such, Equation (3.22) is the endpoint and we will not obtain a homogeneous wavefield, but rather a signal between the source and virtual receiver plus additional artifacts caused by the focusing funtion between the virtual receiver and the fault plane. Similar to the single source, each set of artifacts maps in between the shifted direct arrival of the wavefield and its time-reversal. Due to the different shift of each signal, the artifacts overlap with the shifted causal and acausal parts of other signals and cannot be easily separated. However, because of the limited duration of the artifacts, the signal at later times will be free from these artifacts. Additionally, due to the nature of the characteristic function, the artifacts also vanish when the source location $\mathbf{x}_B^{(k)}$ is outside the volume $\mathbb{D}$. In other words, if the virtual receiver location $\mathbf{x}_A$ is above the shallowest source location, the correct signal can be retrieved for this virtual receiver.

3.3. RESULTS

3.3.1. NUMERICAL RESULTS

MONOPOLE AND DOUBLE-COUPLE POINT SOURCES

To demonstrate the different approaches to homogeneous Green's function retrieval, we apply the methods first on synthetic data. Figure 3.5(a) shows a density model and Figure 3.5(b) shows the accompying P-wave velocity model. The model contains an area of faulting in the center of the model, which is highlighted with a black dashed line. To create the required reflection data, the model is used in a finite-difference modeling code for wavefield modeling [68]. An example of an acoustic common-source record from the center of the model is shown in Figure 3.5(c). This type of common-source records and a smoothed version of the velocity model in Figure 3.5(b), are the only input that we will use for our applications. To retrieve the required Green's functions and focusing functions with the Marchenko method, we model the first arrival from a point in the

subsurface to the surface of the medium using the smooth velocity model and a homogeneous density model. This first arrival is then used to initiate the Marchenko method to retrieve focusing functions and a Green's function from the reflection response at the surface (i.e., from the common source records). The scheme that we use is based on the Marchenko code created by Thorbecke *et al.* [34]. This is a code for an acoustic wavefield Marchenko method, excluding free-surface multiples in the reflection data. Free-surface multiples could be included in the scheme as was shown by Singh *et al.* [49], but this beyond the scope of the current chapter.

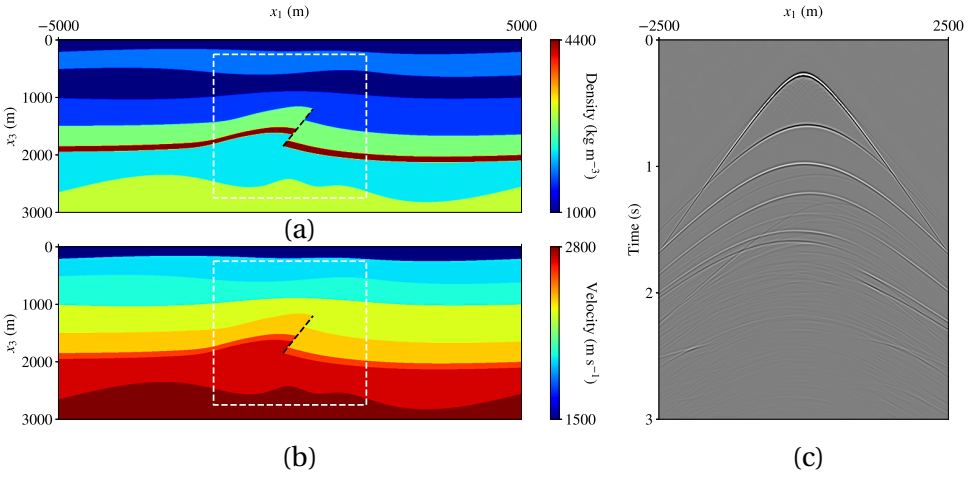

Figure 3.5: (a) Density in kg m^{-3} and (b) P-wave velocity in m s^{-1} of the numerical model used to create reflection data. The white box denotes the area of interest for the purpose of homogeneous Green's function retrieval. The black dashed line indicates a fault plane. (c) Common-source record, created using the model data in (a) and (b), with the source at the top center of the model, using a finite-difference modeling code and convolved with a 30Hz Ricker wavelet.

Figure 3.6 shows the results of the homogeneous Green's function retrieval. All snapshots show the same area in the subsurface, which is denoted by the white box in Figures 3.5(a) and (b). Note that the box does not show the true aspect ratio of the area, however, the snapshots in Figure 3.6 do. Each pixel in the image is a receiver location and the source location for all images is exactly the same. The columns show snapshots of the wavefield in the subsurface at four different points in time, 0ms, 150ms, 300ms and 450ms. Each row corresponds to a specific way the wavefield in the subsurface was constructed. In the first row, the source and the receivers of the wavefield are placed inside the model and the wavefield is directly modeled. This is the benchmark that the other results will be compared to. All snapshots contain an overlay of black dashed lines, which indicate the locations of geological layer interfaces. As can be seen in the figure, the wavefield of the modeling scatters at these lines.

The Marchenko based approach is an improvement over classical methods as was shown by Brackenhoff *et al.* [81], because of the focusing functions that are utilized. To demonstrate this, we first consider a more conventional approach, namely the classical back propagation method from section 2.4 of our companion paper by Wapenaar *et al.*

[24], from which we use equation (23):

$$p^-(\mathbf{x}_A,\mathbf{x}_B,\omega) \approx \int_{\partial\mathbb{D}_0} \frac{2}{i\omega\rho_0} p^-(\mathbf{x},\mathbf{x}_B,\omega)\partial_3 G_d^*(\mathbf{x},\mathbf{x}_A,\omega)\mathrm{d}\mathbf{x}, \tag{3.23}$$

where p^- is the upgoing component of the pressure wavefield at $\partial\mathbb{D}_0$ and $G_d^*(\mathbf{x},\mathbf{x}_A,\omega)$ is the time-reversed first arrival of the Green's function and is the same first arrival that is used as the initial estimation of the focusing function that is used in the Marchenko method. For more information about the method, we refer the reader to our companion paper. Here, we demonstrate the issues with this approach, which can be seen in Figures 3.6(e)-(h). The primary upgoing wavefield can be recovered using this method, however, the downgoing wavefield is missing and strong artifacts are present. This is due to the fact that the multiples and the downgoing wavefield are not taken into account properly using the back propagation method. To make a more detailed comparison between the result of this method and the modeling, we extract the measurements from two receiver locations. These locations are indicated in Figure 3.6(a), where the red dot is a receiver location above the source location and the blue dot a receiver location below the source location. Parts of these measurements are displayed in Figure 3.7, where the left column corresponds to the red dot and the right column to the blue dot. The results in the rows of Figure 3.7 correspond to the results of the rows in Figure 3.6. However, the normalized amplitudes of the traces are used instead of the exact amplitudes. This is done because the first arrivals that were used for the Marchenko method and back propagation were retrieved in a smooth velocity medium without any density information, which is realistic, considering the availibility of data in the field. Because of these limitations the absolute amplitude of the first arrival will be incorrect and while this has no effect on the relative amplitude, it does cause an incorrect overall scaling on the final retrieved wavefield. However, we can still use the normalized traces to analyze the events that are retrieved with the correct relative amplitude. The trace in Figure 3.7(c), located above the virtual source, shows that while some of the correct events are retrieved, a large amount of desired events are missing. These problems are more severe for the receiver below the source location. In Figure 3.7(d), physical events are missing and there are artifacts present all over the trace. The classical back propagation method lacks a great deal of accuracy.

The third row of Figure 3.6 shows the result of Green's function retrieval using the method described by Equation (3.15). The Green's function and focusing functions that are required for this method are retrieved using the Marchenko method. This means that all the receivers and the source are virtual. When the result is compared to the benchmark, it is clear that there are some issues. The wavefield below the source location, as indicated by the yellow dashed line, contains numerous artifacts and the downgoing direct arrival of the wavefield is missing, however, the coda of the wavefield is present both above and below the source location, which is a significant improvement over the back propagation. The remaining errors below the source location are caused by the fact that the focusing function between the virtual source and receiver is present and the lack of compensation for these contributions cause artifacts in the final result. When the virtual receivers are located above the virtual source location, the wavefield is comparable to the benchmark and the direct arrival is present. When the trace in Figure 3.7(a) is

compared to (e), the arrival times of the events match and there are no artifacts present, however there is a mismatch in amplitude. This is due to transmission losses in the reflection response, that the Marchenko method in its current form does not compensate. These effects have been partially compensated for through use of the method discussed by Brackenhoff [66], although not all the effects have been removed. Also, we expect some numerical issues due to the fact that the modeling and the retrieval of the data are two fundamentally different approaches and the data are discretized. The modeling of the first arrival in the smooth model does not only affect their amplitudes, also the arrival times will shift slightly. Due to this slight shift the sampling points of the modeling and the retrieved wavefield may not match exactly. We ensure that the wavelet is zero-phase for the modeling and the Marchenko method to fulfill the symmetric source signal requirement for the homogeneous Green's function representation. When the receiver location below the source is considered in Figure 3.7(f), the results are less accurate. The trace of the modeling contains no signal before the first arrival, whereas the trace for the Green's function retrieval contains numerous events and is lacking the first arrival. The coda of the traces shows a match that is comparable to the receiver location above the source. The arrival times of the events show a good match, while the amplitudes show errors. Because this receiver is located deeper inside the model, the transmission effects are stronger and therefore the error is larger.

Next, the homogeneous Green's function retrieval using Equation (3.16) is considered. The input for this approach is exactly the same as the one used for the previous approach using Equation (3.15), however, this time, we expect to retrieve the correct result. Looking at Figures 3.6(m)-(p), the result more closely matches the result of the benchmark. The improvement over the previous result for the deeper virtual receivers is clear. For some of the deeper receivers, part of the wavefield is still not completely present, however. This is the part of the wavefield that has a steep angle. The reason for this missing part is that the reflection response at the surface does not contain the reflections corresponding to the angles at larger depths, as they travel outside the aperture of the recording survey. Therefore, these steep angles cannot be reconstructed. As can be seen when the trace from Figure 3.7(e) is compared to (g), the result of the two approaches is exactly the same if the virtual receiver is located above the source. The improvement is noticeable when the receiver is located below the source. Figure 3.7(h) does contain the first arrival and lacks any signal before this arrival, and therefore shows a better match to Figure 3.7(b). While the amplitude mismatch is still present, the arrival times of the events match and no artifacts are present. This also shows that the coda of Figure 3.7(f) is correctly retrieved. We have indicated the moment that the correct coda is retrieved with a yellow line in this figure.

To make a more careful comparison between the modeled wavefield and the wavefields retrieved from the reflection data, we plotted the traces from Figures 3.7(a)-(h) together in Figure 3.8, where the left column shows the result for the traces above and the right column shows the result for the traces below the virtual source location. Each subplot contains the modeled response with an overlay of one of the retrieval methods. The back propagation method shows very large errors for both receiver locations as can be seen in Figures 3.8(a) and (b). Strong physical events are missing and artifacts are present on both traces. When comparing the results in Figure 3.8(c), the match of the events be-

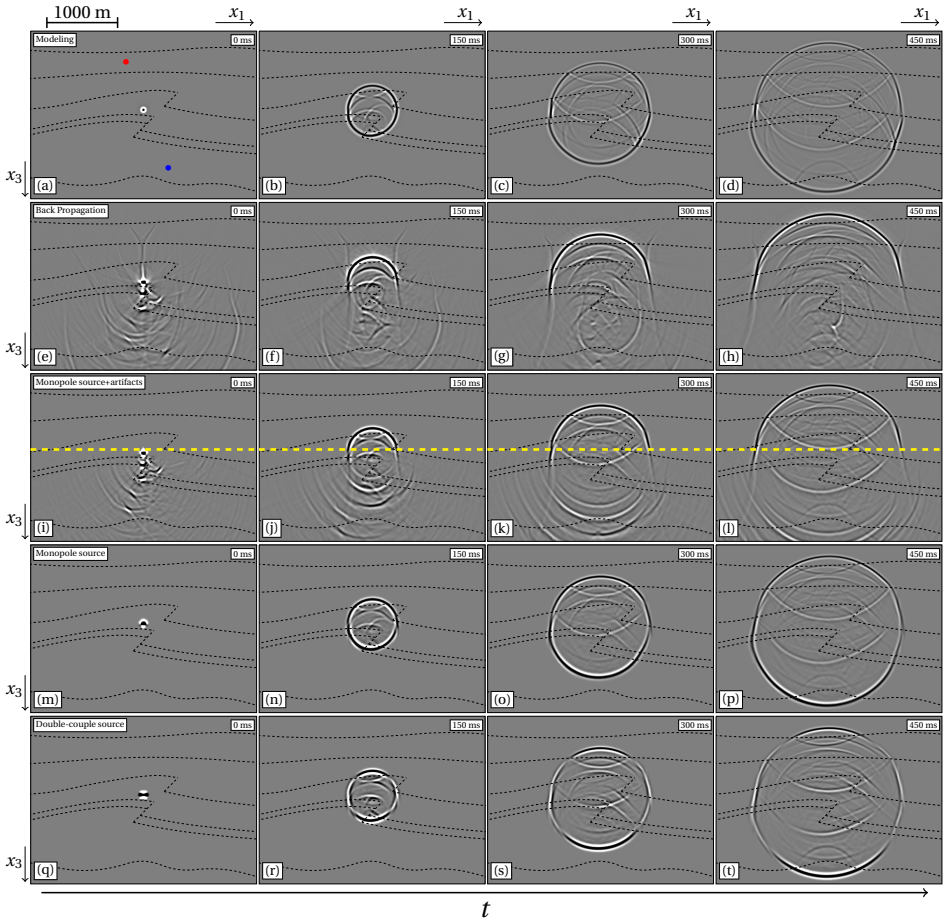

Figure 3.6: Snapshots of the wavefield inside the white box in Figure 3.5 for point sources. (a)-(d) Directly modeled wavefield using the exact model from Figures 3.5(a) and (b). (e)-(h) Back-propagated wavefield obtained using Equation (3.23). (i)-(l) Wavefield in the subsurface, retrieved for virtual receivers and a virtual source using Equation (3.15). The yellow line indicates the border between the area below and above the virtual source. (m)-(p) Similar as (i)-(l), for the homogeneous wavefield using Equation (3.16). (q)-(t) Similar as (m)-(p), using Equation (3.19) and a double couple signature inclined at an angle of 45 degrees. All wavefields have been convolved with a 30Hz Ricker wavelet. The red and blue dot indicate the locations of the traces in Figure 3.7. The black dashed lines indicate the locations of geological layer interfaces.

tween the modeled wavefield and the retrieved wavefield is not perfect. As mentioned before, this is due to the influence of the smooth model and numerical effects that occur. A similar match can be seen in Figure 3.8(e). The retrieval of the Green's function with the artifacts below the source location, which is displayed in Figure 3.8(d), shows the errors at early time, however, also demonstrates that the events in the coda are well captured. This error is not present in case of the homogeneous Green's function retrieval as shown in Figure 3.8(f). These results show that the approach using the Marchenko

method is capable of retrieving the relative amplitudes of the events and can retrieve arrival times that are very close to the actual arrival times, even if a smooth velocity model is used.

Finally, we consider the situation where the source mechanism is more complex, through the use of a double-couple signature. The retrieval in this case corresponds to the approach in Equation (3.19), using a virtual source. The double-couple is an elastic mechanism, however, as we only require the first arrival to initiate the Marchenko method, the coda of the wavefield is not of interest. The S-wave velocity used for the modeling of the first arrival is set to 500m s^{-1}, to ensure that all the S-wave events arrive after the first P-wave arrival. We incline the double-couple source at an angle of 45 degrees and use it to model the first arrival, which is used to initiate the Marchenko method to retrieve the wavefield response for the virtual source location. The focusing functions remain the same as the ones we used for the previous approaches in Figures 3.6(e)-(h). The result of this retrieval is shown in Figures 3.6(q)-(t). As Equation (3.19) states, because the Green's function contains a double-couple signature, the homogeneous Green's function contains the same signature, both in the direct arrival and in the coda of the wavefield. The double-couple signature affects the amplitude of the wavefield depending on the angle of the wavefront, however, the arrival times are similar to those when a monopole virtual source is used. This becomes clear when the traces from Figures 3.7(i)-(j) are considered. The arrival times for the events are similar to the previous result, however, there are apparent amplitude and phase differences, caused by the different types of source signature. Due to these differences, we have not included these traces in Figure 3.8, as a direct comparison between the events cannot be made. The result shows that the double-couple signature can be succesfully integrated in the Marchenko method.

DOUBLE-COUPLE SOURCES ALONG EXTENDED FAULTS

Until now, we only considered single point sources that have a symmetric signal. To study the situation of induced seismicity, we simulate a source that evolves over time over a larger area than a single point. We achieve this by placing a collection of sources along a line in the model. For this purpose, we place 131 sources along the fault plane that was indicated in Figure 3.5, starting at the bottom left corner, with a spacing of 7.07m. The time between the activation of the shots is 12ms, simulating a propagation speed of the source along the fault of 589m s^{-1}. The fault is inclined at 45 degrees, therefore we make use of double-couple sources that are inclined at the same angle. We consider two scenarios, one where we have virtual sources and one where we have a measurement of a real source.

For the first scenario, we approach the problem by considering each source position separately. We do this by retrieving the homogeneous wavefield for each virtual source location separately and by shifting and superposing the results, similar to Equation (3.21). Causality is applied to each individual wavefield before the superposition to avoid overlap between the causal and acausal part of the wavefields. Snapshots of the results are shown in Figures 3.9(a)-(d), for 0ms, 500ms, 1000ms and 1500ms. The reason for the large timesteps is to ensure that all the sources along the fault have been activated during the final snapshot. The propagation of the source location along the fault is clear in these snapshots, however, a propagating wavefield appears to be largely absent, with

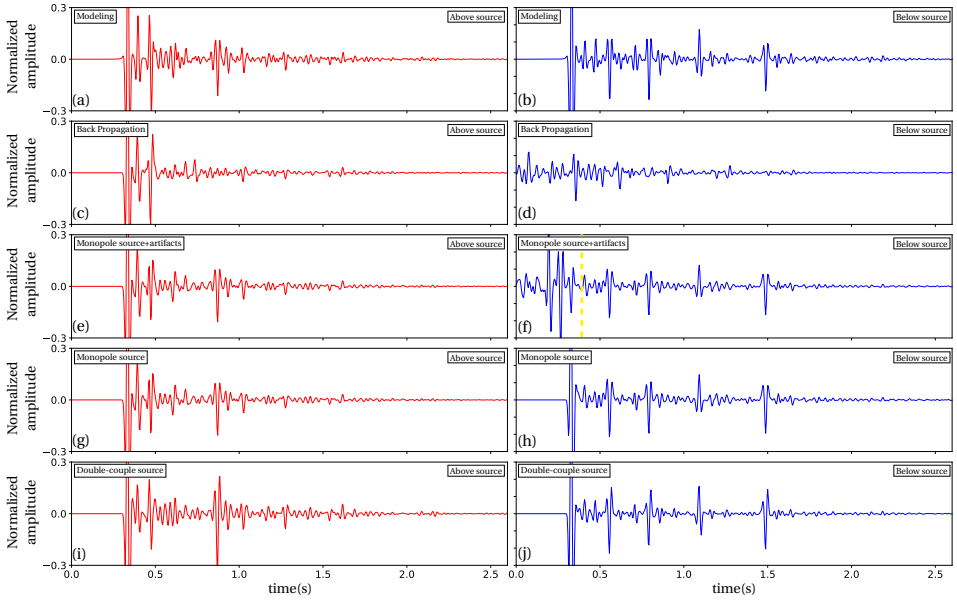

Figure 3.7: Traces from receivers in the subsurface at two locations, extracted from Figure 3.6. In the left column, the receiver is located above the source and corresponds to the red dot in Figure 3.6(a) and in the right column it is located below the source and corresponds to the blue dot in Figure 3.6(a). (a)-(b) Directly modeled wavefield using the exact model from Figures 3.5(a) and (b). (c)-(d) Back-propagated wavefield obtained using Equation (3.23). (e)-(f) Wavefield in the subsurface, retrieved for virtual receivers and a virtual source using Equation (3.15). The yellow line in (f) indicates the time after which the correct signal is retrieved. (g)-(h) Similar as (e)-(f), for the homogeneous wavefield using Equation (3.16). (i)-(j) Similar as (g)-(h), using Equation (3.19) and a double couple signature inclined at an angle of 45 degrees. All wavefields have been convolved with a 30Hz Ricker wavelet.

only a few events and ringing effects present. The reason for this phenomenon is that the velocity at which the sources are activated along the faults is lower than the propagation velocity of the medium. This effectively means that the phase velocity of the combined wavefield along the fault is lower than the propagation velocity of the medium and the emitted wavefield therefore becomes evanescent. This effect can be seen more clearly by considering the traces from two receiver positions. Similar to Figure 3.7, we extract the same receiver locations to consider the individual traces, as shown in Figure 3.10. In Figures 3.10(a)-(b), the trace for the receiver location above the shallowest source location shows a trace with few events, except for some high amplitude events. The receiver location below the deepest source shows a trace that contains more ringing effects with a uniform amplitude. Because the amplitudes are similar and the events located close together, little information can be gained from this trace.

In reality, faults are not uniform, rather they are strongly heterogeneous, which causes variations for the source amplitude along the fault plane. To account for this effect, we apply random scaling to each source location along the fault plane before the superposition takes place. Applying a random scaling factor to the wavefield only affects the amplitudes of the wavefields and does not affect the arrival times or presence of the events

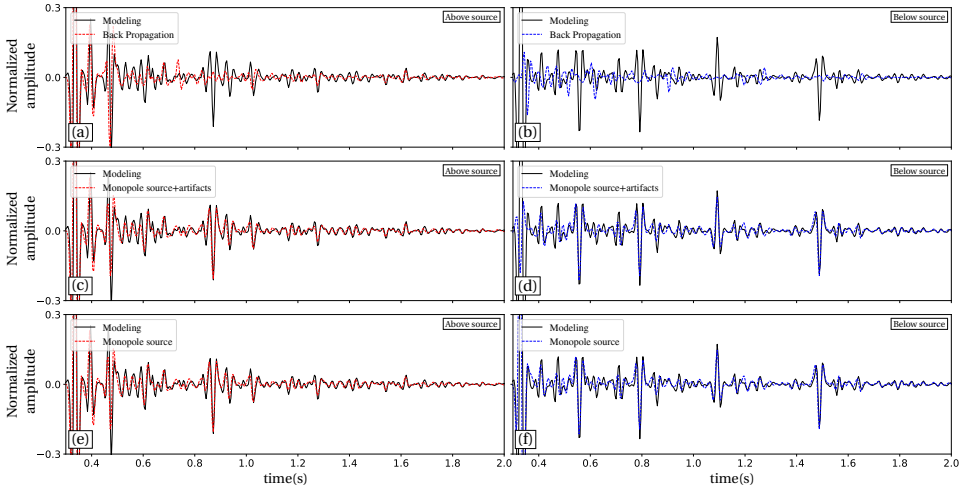

Figure 3.8: (a) Overlay of the traces from Figures 3.7(a) and (c). (b) Similar as (a) for the traces from Figures 3.7(b) and (d). (c) Similar as (a) for the traces from Figures 3.7(a) and (e). (d) Similar as (a) for the traces from Figures 3.7(b) and (f). (e) Similar as (a) for the traces from Figures 3.7(a) and (g). (f) Similar as (a) for the traces from Figures 3.7(b) and (h). All wavefields have been convolved with a 30Hz Ricker wavelet.

in the wavefield. The result of this approach is shown in Figures 3.9(e)-(h). The propagation of the source location along the fault is similar to the uniform amplitude approach, however, the individual wavefields are visible due to the random amplitude approach. Both the first arrivals and the codas can be seen, although there is much overlap between all the wavefields which makes distinguishing individual events at later times challenging. When the two receiver traces in Figures 3.10(c)-(d) are studied, this challenge is still present. The trace contains events, however, it is difficult to say whether these events correspond to the response of one source or another.

To make an estimation for the arrival times of the retrieved response, we numerically model the line source in the subsurface, using the same random amplitude distribution as in the previous case. As we lack the capability to model snapshots of the response to the double-couple source acoustically, we make use of monopole point sources, instead of double-couple sources. As a result, the amplitudes of the events should not be compared to the retrieved response, however, the arrival times can be compared. The wavefield in Figures 3.9(i)-(l) shows that the arrival times are well comparable between the modeling result and the retrieved response. This is further proven when the traces in Figures 3.10(e)-(f) are considered. The arrival times have a strong match, while the amplitudes are not comparable. This confirms that the correct events are retrieved through this approach.

Next, we consider a different scenario, with a real source instead of a virtual one. Here, we once again retrieve the wavefield response of each source separately. However, instead of retrieving a separate wavefield for each of these responses and then superposing these results together, we superpose the responses before the wavefield is retrieved, following Equation (3.22). By using this approach we obtain a response record

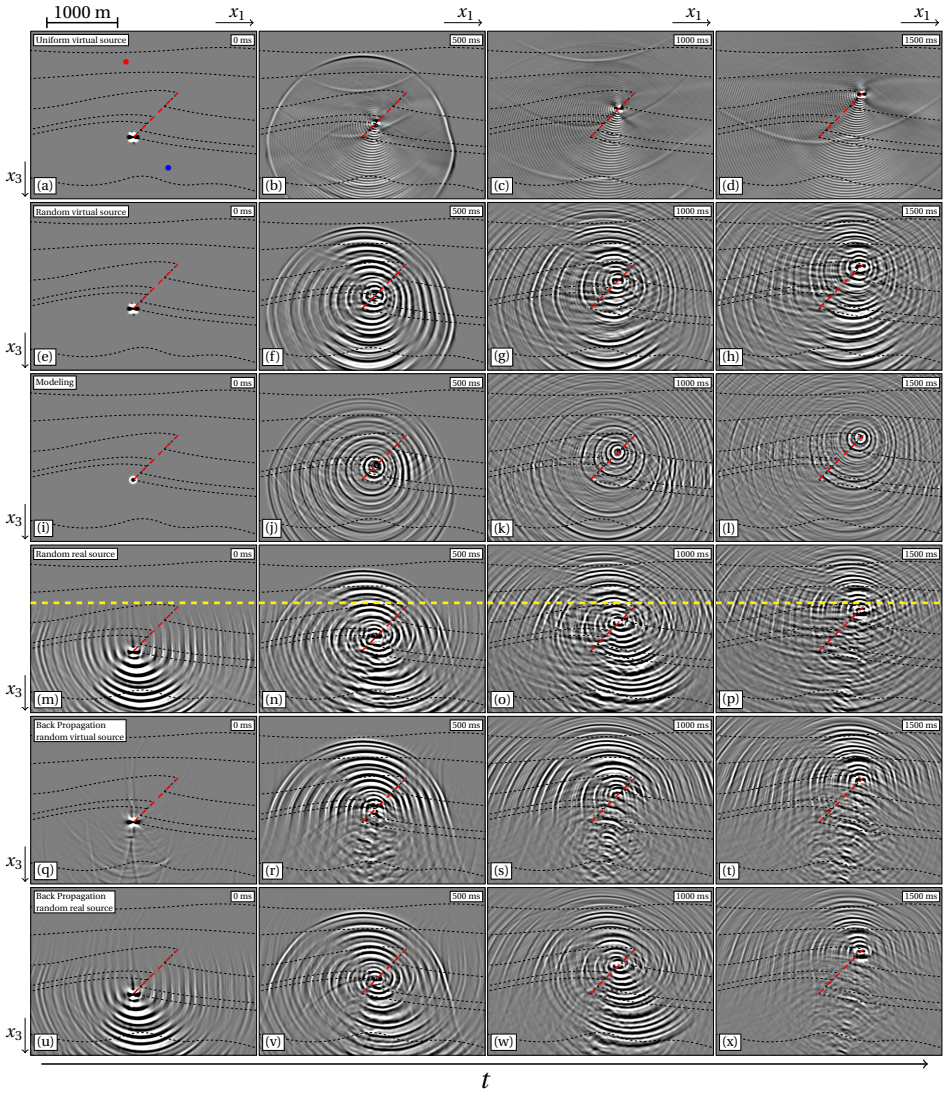

Figure 3.9: Snapshots of the wavefield inside the white box in Figure 3.5 for line sources. (a)-(d) Response in the subsurface, retrieved using Equation (3.21) for virtual receivers and virtual double-couple sources inclined at 45 degrees with an uniform amplitude. (e)-(h) Similar as (a)-(d), using random amplitudes for the source. (i)-(l) Directly modeled wavefield using the exact model from Figures 3.5(a) and (b) and monopole point sources with a random amplitude. (m)-(p) Similar as (e)-(h) using a superposition of double-couple sources with random amplitudes using Equation (3.22). The yellow line indicates the border between the area below and above the shallowest source. (q)-(t) Similar as (e)-(h), however instead of using the homogeneous Green's function retrieval, the back propagation using Equation (3.23) is used for each source position. (u)-(x) Similar as (m)-(p), however instead of using the Green's function retrieval, the back propagation using Equation (3.23) is used. All wavefields have been convolved with a 30Hz Ricker wavelet. The red and blue dot indicate the locations of the traces in Figure 3.10. The black dashed lines indicate the locations of geological layer contrasts.

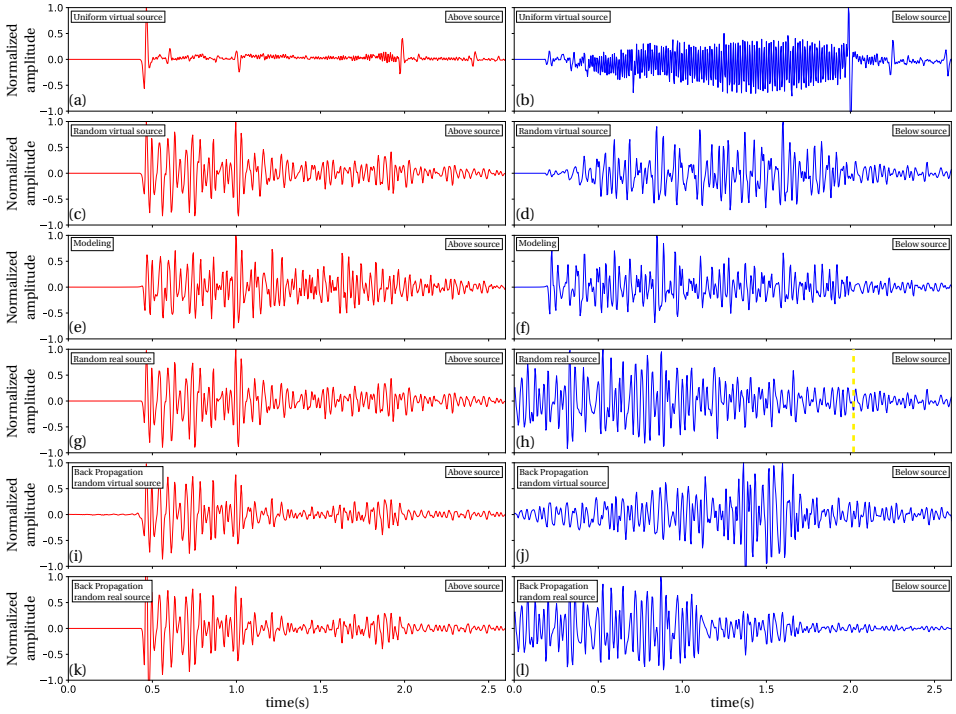

Figure 3.10: Traces of receivers in the subsurface at two locations, extracted from Figure 3.9. In the left column, the receiver is located above the source and corresponds to the red dot in Figure 3.9(a) and in the right column it is located below the source and corresponds to the blue dot in Figure 3.9(a). (a)-(b) Response in the subsurface, retrieved using Equation (3.21) for virtual receivers and virtual double-couple sources inclined at 45 degrees with an uniform amplitude. (c)-(d) Similar as (a)-(b), using random amplitudes for the source. (e)-(f) Directly modeled wavefield using the exact model from Figures 3.5(a) and (b) and monopole point sources with a random amplitude. (g)-(h) Similar as (c)-(d) using a superposition of double-couple sources with random amplitudes using Equation (3.22). The yellow line in (h) indicates the time after which the correct signal is retrieved. (i)-(j) Similar as (c)-(d), however instead of using the homogeneous Green's function retrieval, the back propagation using Equation (3.23) is used for each source position. (k)-(l) Similar as (g)-(h), however instead of using the Green's function retrieval, the back propagation using Equation (3.23) is used. All wavefields have been convolved with a 30Hz Ricker wavelet.

that matches the response of a real source recording in the subsurface. The same random amplitude distribution that we used for the previous two results is applied for this approach as well, to make the comparison fair. The wavefield that is obtained is shown in Figures 3.9(m)-(p), where we can see that the propagation of the source location along the fault is captured properly. There are issues with the approach due to the limitation of the representation that is used. The response to each source has artifacts that arrive before the first arrival when the virtual receiver is located below any of the source locations. These effects overlap with the causal wavefields of sources at other locations, and obscure the events that should be present. Additionally, the downgoing first arrival is missing for all source locations. These problems are inherent to the representation and cannot be easily avoided, however, the coda of the response for later times will be cor-

rect, as we saw already for the point source in Figures 3.7(e)-(h). When the traces for this approach from Figures 3.10(g)-(h) are studied, we can see that if the receiver is located below the source locations, individual events belonging to the sources are impossible to distinguish. If the receiver is located above all the sources, however, the response is retrieved correctly. The lower receiver contains the correct coda at later time. We indicated this moment with a yellow line in Figure 3.10(h), similar to Figure 3.7(d). This, combined with the fact that the source location of the signal can be clearly distinguished, shows that this approach has potential for field recordings.

Finally, as an example for the improvement of this approach over conventional methods, we repeat the retrieval of a fault plane source using the back propagation method. We consider both the approach for retrieving a virtual source and retrieving a real source. For the first approach, we retrieve the response for each source location, mute the acausal part of the response and shift it in time, to create one source signal. However, instead of using homogeneous Green's function retrieval to obtain the responses, we employ the classic back propagation and show the resulting wavefield in Figures 3.9(q)-(t). While the primay upgoing wavefield is still captured, the coda and the downgoing wavefield are absent. Aside from the missing events, artifacts are present at all times in the result. When the extracted traces are considered in Figures 3.10(i)-(j), we can see that the trace is completely different to the traces in Figures 3.10(c)-(d). Due to the fact that the missing events and the artifacts shift along with the source position, it masks the entire trace. The effects of the classical back propagation approach have a similar result when we repeat the experiment for our real source example. We use classical back propagation instead of Green's function retrieval on the simulated real source response and show the result in Figures 3.9(u)-(x). Similar problems with the coda and the downgoing wavefield are present and the artifacts in the wavefield are still ocurring. The extracted trace above the source locations in Figure 3.10(k) shows the same result as in Figure 3.10(i), which is consistent with the previous results. The extracted trace below the source locations in Figure 3.10(l) shows the strong degradation in quality and has no match with the desired result in Figure 3.10(d). This shows that for both types of sources, real or virtual, the single-sided approach with a focusing function is an improvement over the classical approach using back propagation. Therefore, the latter approach will not be used for the field data.

3.3.2. FIELD DATA RESULTS

To demonstrate that our approach is not limited to synthetic data, we also apply the method on field reflection data. The field data were recorded in the Vøring basin, in a marine setting by SAGA Petroleum A.S., which is currently part of Equinor. Due to the setting, the receivers only recorded P-waves. The data consist of 399 common-source records, an example of which is shown in Figure 3.11(c). The data were preprocessed before the application of the homogeneous Green's function retrieval, through the use of the Estimation of Primaries through Sparse Inversion (EPSI) method to remove the source wavelet, retrieve the near-offsets and remove the free-surface multiples [83]. Moreover, we applied source-receiver reciprocity to allow the retrieval of two directions of offset and adaptive corrections to compensate for attenuation and incorrect source strength. Along with the reflection data, a smooth P-wave velocity model was also provided, which

is shown in Figure 3.11(a). We indicate the region of interest, where we will perform homogeneous Green's function retrieval, with a white dashed box. The model is not displayed in a true to life aspect ratio. The reflection data and the velocity model are the only inputs that are available for the homogeneous Green's function retrieval. No direct information about the subsurface is available for this area, however, using the reflection data and the velocity model, an image of the subsurface was created using the Marchenko method, shown in Figure 3.11(b), which we will use as a reference for where scattering is expected to take place. This imaging was done independently of the homogeneous Green's function retrieval and is only used as a reference. More information about imaging using the Marchenko method, as well as an application on field data, can be found in Staring *et al.* [38]. The homogeneous Green's function retrieval for this dataset has been succesfully performed, as was shown in [81], however, in this work we will expand the results to include the line source configuration.

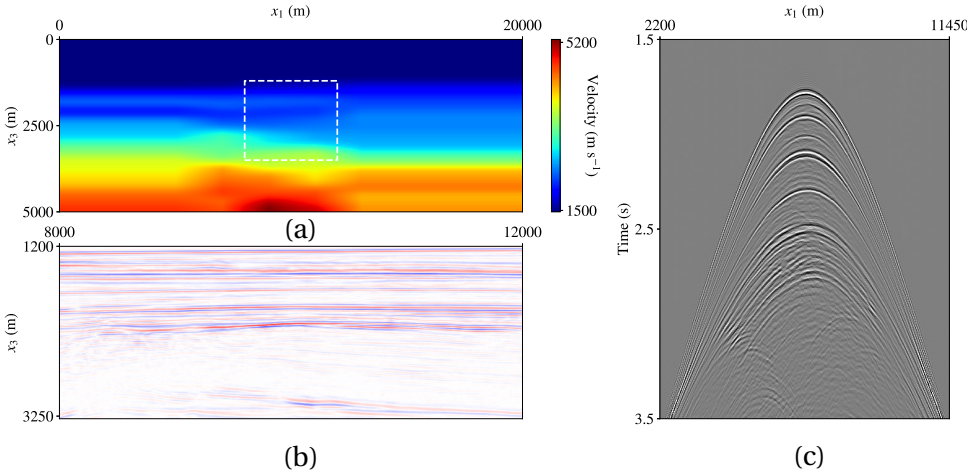

Figure 3.11: Real data example, (a) P-wave velocity in m s^{-1} of the field data. The white box denotes the area of interest for the purpose of homogeneous Green's function retrieval. (b) Image of the subsurface located in the region indicated by the white dashed box. (c) Common-source record of the field reflection data, processed for the purpose of applying the Marchenko method. The reflection data source wavelet was reshaped to a 30Hz Ricker wavelet. The data itself was recorded in the Vøring basin in Norway and was provided by Equinor.

Because there is no information about the subsurface available, we cannot directly model in the subsurface and therefore have no benchmark, however, we have shown with the previous examples that the method is capable of retrieving the correct result. We perform homogeneous Green's function retrieval in the subsurface for both a virtual source and virtual receivers. The virtual source is a double-couple source, inclined at 20 degrees. The result is shown in Figures 3.12(a)-(d) for 0ms, 300ms, 600ms and 900ms. The image of the subsurface from Figure 3.11(b) is used as an overlay to help indicate the region where scattering of the wavefield is expected. The scattering takes place along regions where high amplitudes are present for the subsurface image, which indicates a match between the image and the homogeneous wavefield. Aside from the direct arrival, there is also a coda present, which contains several events. The result is not as

clean as the synthetic data, however. This is due to the limitations of the field data. The data is attenuated, a problem that the Marchenko method cannot properly account for. The attenuation has been corrected for during the processing, however, this process is imperfect and will leave imperfections in the final result. There is also incoherent noise present in the field data, which has not been removed during the processing and will be present in the final result.

Figure 3.12(a) shows a red and blue dot, which indicate the location of traces that are extracted and are shown in the left and right column of Figure 3.13, respectively. No benchmark for these traces is available, and thus it cannot be directly validated. The results in Figures 3.13(a)-(b) show that the traces contain multiple well defined events, and that the noise on the trace is has lower amplitudes than these events. The amplitude of the first arrival is strong compared to the coda and the phase of all the events is similar. This shows that if the faults in the model are small compared to the wavelength, this approach can be useful for interpretation and characterisation of the source mechanism.

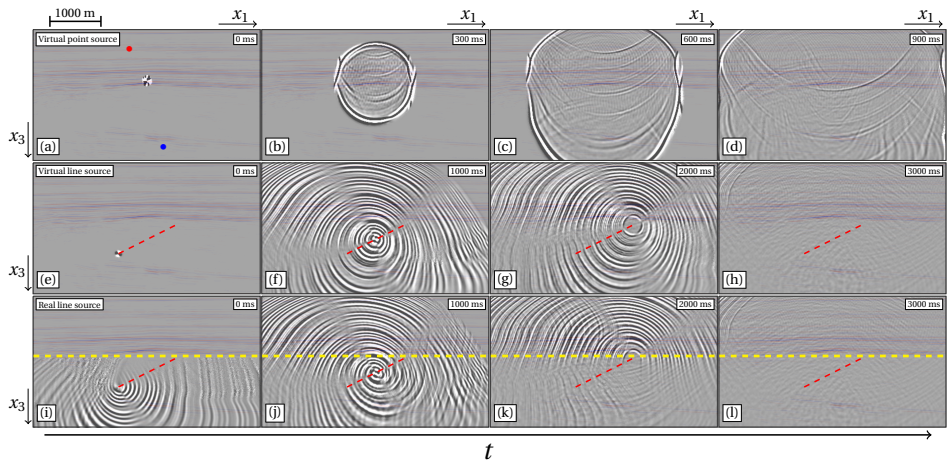

Figure 3.12: Snapshots of the wavefield inside the white box in Figure 3.11 for the field data. (a)-(d) Homogeneous wavefield in the subsurface, retrieved for virtual receivers and a virtual double-couple source inclined at -20 degrees using Equation (3.19). (e)-(h) Similar as (a)-(d), for a line source of double couple sources with random amplitudes inclined at 22.4 degrees using Equation (3.21). (i)-(l) Similar as (e)-(h), using a superposition of double-couple sources with random amplitudes using Equation (3.22). The yellow line indicates the border between the area below and above the shallowest source. The images are overlain with the image of the subsurface from Figure 3.11(b). All wavefields had their source wavelets reshaped to a 30Hz Ricker wavelet.

Next, we consider the two line source configurations for the virtual and the real source configuration. As there is no clear fault present in the model, the fault line is arbitrarily placed in the center of the model, inclined at an angle of 22.4 degrees. 161 sources are used with a spacing of 6.99m, where the time between the activation of the shots is 12ms, simulating a propagation speed of the source along the fault of 583m s^{-1}. A random amplitude is assigned to each of the source locations to generate propagating waves. The first situation we consider is using Equation (3.21), where homogeneous Green's function retrieval is performed for each location separately and the results are superposed and causality is imposed. The results of this approach are shown in Figures 3.12(e)-(h),

for 0ms, 1000ms, 2000ms and 3000ms. Similar to the synthetic data, the movement of the source is well captured and the first arrival and the coda are present in the signal. Part of the wavefield is not present, which corresponds to high angles at deeper depths, which, as we explained before, are not present in the reflection response and can therefore not be reconstructed. The result has a similar quality as the single double-couple source in Figures 3.12(a)-(d) and the results on the synthetic data Figure 3.9.

There is no induced seismicity signal present for this area, so a real source signal cannot be used, but we simulate this as follows. Similar to the approach for the synthetic data, we use the Marchenko method to retrieve a wavefield response with a double-couple signature for each source location. These signals are then superposed to create a single source record, as a substitute for a real source signal. This approach follows Equation (3.22), the results of which are shown in Figures 3.12(i)-(l). Similar to the results for the synthetic data, the match between the two approaches above the shallowest source location is strong. This is proven further when the traces above the source from Figures 3.13(c) and (e) are compared to each other. The traces are nearly identical. If we consider a location below the the deepest source location, the results are less comparable, again similar to the results that were achieved on the synthetic data. The traces for this location, shown in Figures 3.13(d) and (f), support this conclusion. The match in this situation is non-existent for earlier times, and the information is hard to appraise. At later times, as indicated by the yellow line, the coda of the two approaches match each other, similar as seen before. For both types of retrieval, the source locations are well-defined in both time and space and not obscured by artifacts that could cast doubt on the source locations. Using both types of approach shows potential for the determination of the source location and the coda and can help in the characterisation of the fault mechanism.

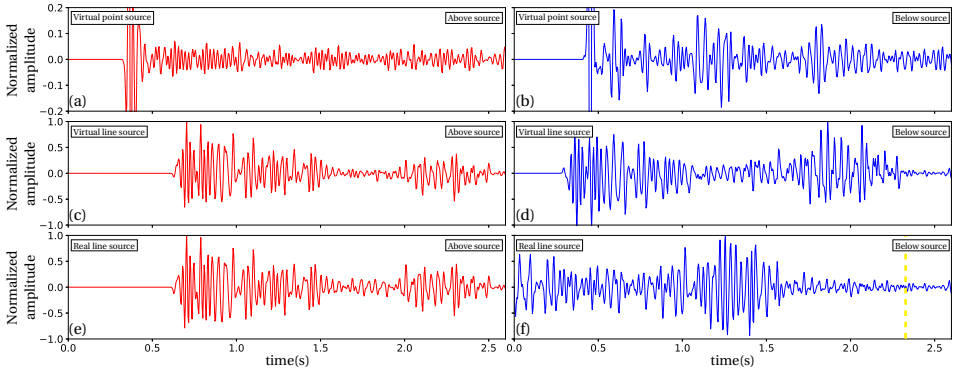

Figure 3.13: Traces of receivers in the subsurface at two locations, extracted from Figure 3.12. In the left column, the receiver is located above the source and corresponds to the red dot in Figure 3.12(a) and in the right column it is located below the source and corresponds to the blue dot in Figure 3.12(a). (a)-(b) Homogeneous wavefield in the subsurface, retrieved for virtual receivers and a virtual double-couple source inclined at -20 degrees using Equation (3.19). (c)-(d) Similar as (a)-(b), for a line source of double couple sources with random amplitudes inclined at 67.6 degrees using Equation (3.21). (e)-(f) Similar as (c)-(d), using a superposition of double-couple sources with random amplitudes using Equation (3.22). The yellow line in (f) indicates the time after which the correct signal is retrieved. All wavefields had their source wavelets reshaped to a 30Hz Ricker wavelet.

3.4. CONCLUSIONS

In this chapter, we considered two methods to monitor full wavefields in the subsurface using the Marchenko method and found that in both cases, the Marchenko based approach is an improvement over classical methods such as back propagation. The first method is based on the creation of both virtual receivers and virtual sources in the subsurface. In this case, all the signals are created from the reflection data at the surface, and no response from a real subsurface source is used. For virtual point sources, we showed that we can assure that the source signal is symmetric and that therefore the full homogeneous wavefield can be retrieved without artifacts. The main limitation is that the steepest part of the wavefield at large depths cannot be retrieved. This approach works for virtual sources, both with a monopole signature and a more complex double-couple signature, the latter of which was used as a model for a small scale induced seismicity signal. Larger scale induced seismicity signals, emitted from a fault plane, were considered as well, simulated by a series of individual point sources with a double-couple signature. For this case, the homogeneous wavefield was retrieved for all the sources separately, after which the causal parts were isolated, shifted in time and superposed together. This produces a response from an extended fault rupture that is operating over a larger window of time, which produces a far more complex signal. All the source locations can be distinguished using this method. This method can be used to forecast in a data-driven way the response to possible future induced seismic events.

The second method we considered creates virtual receivers in the subsurface that observe a real response from a subsurface source. To this end, we considered point sources where the source signal was not assumed to be symmetric in time. The causal wavefield that is retrieved in this case is missing a part of the direct arrival and contains artifacts. These problems are only present when the virtual receiver is located below the source location, and the artifacts map exclusively in the time interval between the direct arrival of the wavefield and its time reversal. The coda of the causal wavefield is retrieved in full, as well as the source location of the subsurface response. When considering the responses propagating from a fault, the artifacts are more severe. Unlike in the method with the virtual sources, to simulate the response to a real rupturing fault, we shifted and superposed the source responses before the Green's function retrieval. Because of this, the artifacts are present for each point source, however, due to the time shift, the artifacts of one response coincided with the causal coda of other responses. As a result the coda of the retrieved wavefield is only partially obtained. The source locations of the fault response are retrieved correctly. This method can be used to monitor in a data-driven way the response to actual induced seismic events everywhere between the surface and the source.

We applied the two methods to synthetic and field data. For the synthetic data we showed that the retrieved responses match very well with directly modelled responses. The results obtained from the field data are very similar to those obtained from the synthetic data. The results on the datasets show the potential for the application of the method on real source signals in the future.

4

IMPLEMENTATION OF THE 3D MARCHENKO METHOD

The Marchenko method can be applied on cross-sections of 3D data using a 2D algorithm, but only a full 3D implementation can properly retrieve all 3D effects present in the data. The 3D implementation of the iterative Marchenko method is in principle a straightforward extension of the 2D method, it only requires an additional surface integration dimension. We present a 3D implementation of the Marchenko method based on an earlier implementation of the 2D Marchenko method. In the discussed implementation, special attention is given to an efficient kernel implementation and limiting the amount of data to read in, by using data compression. The algorithm properly handles 3D events and this is illustrated in the intermediate results of individual iterations of the method. Due to model-time and data-size constraints, the aperture of the data is limited, which leads to finite aperture artefacts that cause a lower convergence rate. We demonstrate the 3D method by retrieving the Green's functions for two models and comparing these functions to reference solutions. The two models are a horizontally layered medium and the complex SEG/EAGE Overthrust model. Using the Marchenko method, imaging is applied to both models to show that false images caused by internal multiples are attenuated. Our 3D implementation is fully opensource and is intended to be used in future studies.

This chapter is in preparation for publication. A preprint is available as J. Brackenhoff, J. Thorbecke, V. Koehne, D. Barrera & K. Wapenaar, *Implementation of the 3D Marchenko method.* ArXiv, preprint number:2004.00896 (2020).
Minor modifications have been applied to the text and figures for the sake of consistency in the thesis.

4.1. INTRODUCTION

The Marchenko method is a novel method in the field of geophysics and was first considered by Broggini *et al.* [29] for 1D media and Wapenaar *et al.* [31], Broggini *et al.* [32] and Behura *et al.* [33] further derived the method for 2D and 3D applications. The Marchenko method uses reflection data without free-surface multiples at the surface of a medium together with a first arrival from a focal point in the medium to obtain the upgoing and downgoing Green's functions as if there was a virtual source at the focal location, without artefacts caused by internal multiples [26]. Furthermore, the method is also capable of retrieving focusing functions, which are wavefields that focus from the surface of the Earth to the focal location. The method can be used to predict and subtract internal multiples from the reflection data [38, 45], to obtain images of the subsurface without artefacts caused by internal multiples [30, 32, 84, 85] or to obtain the homogeneous Green's functions between any two focal points in the subsurface [62, 81]. The method is also known for its versatility, for example, Singh *et al.* [49] and Slob and Wapenaar [50] developed a Marchenko scheme that functions with reflection data that include free-surface multiples, Wapenaar and Slob [42], da Costa Filho *et al.* [43] and Reinicke and Wapenaar [44] applied the method to elastic media and Meles *et al.* [47] adjusted the scheme to work very efficiently with plane-waves. Although the theory is fully developed for 3D media, most published applications are based on the 2D Marchenko method, due to the need for well sampled data, which is difficult to achieve for 3D acquisitions. Lomas and Curtis [53] made a comparison between results obtained with a full 3D acquisition and results obtained with linear seismic acquisition arrays, both recorded over a 3D medium. The authors showed that while the 2D approximation can yield good results, if one wants to take into account the full 3D effects, especially the out-of-plane reflections, a 3D version of the Marchenko method is required.

The Marchenko method can be implemented in a variety of ways. Most applications make use of an iterative scheme, however, the Marchenko method has also been implemented as a least-squares inversion [36, 50, 86] or the iterative scheme can be combined with adaptive subtraction [38, 41]. The least-squares inversion is computationally feasible for 2D reflection data, however, for 3D data this method becomes computationally expensive. The adaptive subtraction is more robust to imperfections in the reflection data, however, due to its adaptive nature, the subtraction can attenuate physical events that are coinciding with multiples. In this chapter, we describe the iterative 3D Marchenko implementation, that is still feasible to employ while using limited computing resources. Thorbecke *et al.* [34] published an opensource 2D iterative Marchenko scheme that can be used in combination with reflection data that contain no free-surface multiples. In this chapter, we extend this iterative implementation to the 3D situation and discuss how the implementation and intermediate results are different from the 2D situation.

After a brief theoretical introduction of the Marchenko method, the 3D algorithm is explained. In the numerical examples section the basic processing steps are demonstrated with a four layer reflection model, followed by a more complicated example based on the SEG/EAGE Overthrust model by Aminzadeh *et al.* [87]. This includes the retrieval of a Green's function for a single focal point, which is then compared to a reference solution. Furthermore the method is also used to demonstrate the imaging of the model,

with removal of artefacts caused by the internal multiples. To obtain these results, the 3D reflection data are compressed using the ZFP algorithm by Lindstrom [88], in order to limit the storage size and the time required to read them into the Marchenko program. Several other modules are also implemented, including a finite difference modeling code, a 3D Eikonal solver and an imaging module. Appendix 4A explains how these applications and auxiliary programs, as well as the basic 3D Marchenko program, can be used.

4.2. MARCHENKO METHOD

In theory, the extension of the Marchenko method from 2D to 3D is straightforward, as it only requires the integration over a line in 2D to change to an integration over a surface in 3D. In practice, it is more complicated due to a variety of reasons, which will be discussed in the next section. Thorbecke *et al.* [34] discussed the implementation and theory of the Marchenko method in 2D in great detail. The extension to 3D in this chapter is based on the code that was presented in this previous work, along with the theory, which works in both 2D and 3D. Because of this, the Marchenko method will not be discussed in detail and only the most important equations will be reviewed.

The implementation is based on the use of the coupled Marchenko equations, which are given as [26]:

$$G^+(\mathbf{x}_B,\mathbf{x},t) = -\int_{\partial\mathbb{D}_0}\int_{t'=-\infty}^{t} R(\mathbf{x},\mathbf{x}_S,t-t')f_1^-(\mathbf{x}_S,\mathbf{x}_B,-t')\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S + f_1^+(\mathbf{x},\mathbf{x}_B,-t), \tag{4.1}$$

$$G^-(\mathbf{x}_B,\mathbf{x},t) = \int_{\partial\mathbb{D}_0}\int_{t'=-\infty}^{t} R(\mathbf{x},\mathbf{x}_S,t-t')f_1^+(\mathbf{x}_S,\mathbf{x}_B,t')\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S - f_1^-(\mathbf{x},\mathbf{x}_B,t), \tag{4.2}$$

where $R(\mathbf{x},\mathbf{x}_S,t)$ is a reflection response measured at receiver location $\mathbf{x}=(x_1,x_2,x_3)$ due to a dipole source at $\mathbf{x}_S$ at the surface $\partial\mathbb{D}_0$, often the surface of the Earth, $G^\pm(\mathbf{x}_B,\mathbf{x},t)$ are the upgoing and downgoing Green's functions at $\mathbf{x}_B$ due to a source at $\mathbf{x}$, $f_1^\pm(\mathbf{x},\mathbf{x}_B,-t)$ are the upgoing and downgoing focusing functions that focus from $\mathbf{x}$ to focal location $\mathbf{x}_B$, where the superscripts + and - indicate a downgoing and upgoing wavefield, respectively, and t is time. The downgoing focusing function is defined as the inverse of the transmission response, $T^{\text{inv}}(\mathbf{x}_B,\mathbf{x},t)$, in a medium that is truncated below the focal location [26]. For the purpose of the implementation, the downgoing focusing function is written as a combination of the direct arrival following by a scattering coda:

$$f_1^+(\mathbf{x},\mathbf{x}_B,t) = T_d^{\text{inv}}(\mathbf{x}_B,\mathbf{x},t) + M^+(\mathbf{x},\mathbf{x}_B,t), \tag{4.3}$$

where $T_d^{\text{inv}}(\mathbf{x}_B,\mathbf{x},t)$ is the inverse of the direct arrival of the transmission response and M^+ is the scattering coda of the focusing function. Instead of $T_d^{\text{inv}}(\mathbf{x}_B,\mathbf{x},t)$, often the time reversal of the direct arrival of the Green's function, $G_d(\mathbf{x},\mathbf{x}_B,-t)$, is used as the direct wave of the focusing function. This is done for practical reasons and if this approximation is used, equation (4.3) can be estimated as

$$f_1^+(\mathbf{x},\mathbf{x}_B,t) \approx G_d(\mathbf{x},\mathbf{x}_B,-t) + M^+(\mathbf{x},\mathbf{x}_B,t). \tag{4.4}$$

The focusing function in equation (4.4) will contain the correct travel times of the wavefield, however, the amplitudes will contain errors, which are proportional to the transmission losses of the medium.

In equations (4.1) and (4.2), there are four unknowns, the decomposed Green's functions and decomposed focusing functions, and only two equations, making it an underdetermined system. To separate the focusing functions and Green's functions, we introduce an offset-dependent time-windowing function:

$$w_t(\mathbf{x},\mathbf{x}_B,t) = \begin{cases} 1 & |t| < t_d^\epsilon(\mathbf{x},\mathbf{x}_B), \\ \frac{1}{2} & |t| = t_d^\epsilon(\mathbf{x},\mathbf{x}_B), \\ 0 & |t| > t_d^\epsilon(\mathbf{x},\mathbf{x}_B), \end{cases} \tag{4.5}$$

where $t_d^\epsilon = t_d - \epsilon$, t_d is the arrival time of the direct arrival $G_d(\mathbf{x},\mathbf{x}_B,t)$ and ϵ indicates a small constant. The time window makes use of the causality relations, the direct arrival is zero before t_d and the scattering coda of the focusing function is zero for $t \leq -t_d$ and for $t \geq t_d$ [30]. Applying the window, which we will refer to as w_t for simplicity, to equation (4.4) will cause only the scattering coda M^+ to remain. The window will not remove any events from the upgoing focusing function, however, if it applied to the Green's function it will completely remove this wavefield, as the Green's function exist for $t \geq t_d$ and is zero for $t < t_d$. The small constant ϵ is required because the seismic data are band-limited and a shift is necessary to avoid cutting into the wavelet. Because of this, ϵ should be at least equal to half the size of the wavelet. Applying w_t to Equation (4.1) and (4.2) yields

$$M^+(\mathbf{x},\mathbf{x}_B,-t) = w_t \int_{\partial\mathbb{D}_0}\int_{t'=-\infty}^{t} R(\mathbf{x},\mathbf{x}_S,t-t')f_1^-(\mathbf{x}_S,\mathbf{x}_B,-t')\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S, \tag{4.6}$$

$$f_1^-(\mathbf{x},\mathbf{x}_B,t) = w_t \int_{\partial\mathbb{D}_0}\int_{t'=-\infty}^{t} R(\mathbf{x},\mathbf{x}_S,t-t')f_1^+(\mathbf{x}_S,\mathbf{x}_B,t')\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S. \tag{4.7}$$

In this system, we have two unknowns and two equations, allowing us to solve it in an iterative manner. The reflection response is the known quantity and will not change, while the upgoing and downgoing focusing function will be updated according to

$$M_k^+(\mathbf{x},\mathbf{x}_B,-t) = w_t \int_{\partial\mathbb{D}_0}\int_{t'=-\infty}^{t} R(\mathbf{x},\mathbf{x}_S,t-t')f_{1,k-1}^-(\mathbf{x}_S,\mathbf{x}_B,-t')\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S, \tag{4.8}$$

$$f_{1,k}^-(\mathbf{x},\mathbf{x}_B,t) = w_t \int_{\partial\mathbb{D}_0}\int_{t'=-\infty}^{t} R(\mathbf{x},\mathbf{x}_S,t-t')f_{1,k}^+(\mathbf{x}_S,\mathbf{x}_B,t')\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S, \tag{4.9}$$

where k indicates the iteration number. The update to the downgoing focusing function can be estimated by combining Equation (4.8) with Equation (4.4):

$$f_{1,k}^+(\mathbf{x},\mathbf{x}_B,t) = G_d(\mathbf{x},\mathbf{x}_B,-t) + M_k^+(\mathbf{x},\mathbf{x}_B,t). \tag{4.10}$$

To start the scheme, a first estimation is required. For the first estimation, we assume that the scattering coda, $M_0^+(\mathbf{x},\mathbf{x}_B,t)$ is equal to zero, so that

$$f_{1,0}^+(\mathbf{x},\mathbf{x}_B,t) = G_d(\mathbf{x},\mathbf{x}_B,-t). \tag{4.11}$$

The first estimation of Equation (4.11) is used in Equation (4.9) to create a first estimation of the upgoing focusing function. This estimation is used in Equation (4.8) to update the scattering coda of the downgoing focusing function, which is used in Equation (4.10) to

create an update to the downgoing focusing function. The updated focusing function can then be used in equation (4.9) to repeat the process until convergence has been achieved. After the focusing functions converge to a solution, they can be employed in Equations (4.1) and (4.2) to obtain the Green's functions. Note that all of these results will have the incorrect scaling due to the use of $G_d(\mathbf{x},\mathbf{x}_B,-t)$ as the first estimation in Equation (4.11). If the time-reversed direct arrival is replaced by $T_d^{\text{inv}}(\mathbf{x}_B,\mathbf{x},t)$, the results will contain the correct scaling.

4.3. MARCHENKO ALGORITHM

Because of the similarity to the previous implementation in 2D, the flowchart in Figure 4.1 is similar to the one in Thorbecke *et al.* [34]. The differences of the implementation in 2D and 3D will be highlighted when they occur. For the implementation of the Marchenko method in both 2D and 3D, additional operations are performed next to the iterative Marchenko equations. To obtain not only the decomposed Green's functions, but also the full Green's function, we also estimate the focusing function of the second type [26, 34], which is related to the decomposed focusing functions by

$$f_2(\mathbf{x}_B,\mathbf{x},t) = f_1^+(\mathbf{x},\mathbf{x}_B,t) - f_1^-(\mathbf{x},\mathbf{x}_B,-t), \tag{4.12}$$

and to the full Green's function G as

$$G(\mathbf{x},\mathbf{x}_B,t) - f_2(\mathbf{x}_B,\mathbf{x},-t) = \int_{\partial\mathbb{D}_0}\int_{-\infty}^{\infty} R(\mathbf{x},\mathbf{x}_S,t')f_2(\mathbf{x}_B,\mathbf{x}_S,t-t')\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S. \tag{4.13}$$

In both equation (4.8) and (4.9), a convolution of a wavefield with the reflection data matrix R is performed, after which an integration is applied over the source coordinate of R, which is a linear array in 2D and a surface array in 3D, and finally w_t is applied to the result. Because these operations are so similar, they can be used to design a compute kernel and to this end, we rewrite the Equations (4.8) and (4.9) as a series expansion [26]:

$$M_k^+(\mathbf{x},\mathbf{x}_B,t) = \sum_{i=0}^{k} N_{2i+1}(\mathbf{x}_B,\mathbf{x},t), \tag{4.14}$$

$$-f_{1,k}^-(\mathbf{x},\mathbf{x}_B,-t) = \sum_{i=0}^{k} N_{2i}(\mathbf{x}_B,\mathbf{x},t), \tag{4.15}$$

where

$$N_i(\mathbf{x}_B,\mathbf{x},t) = -w_t RN_i(\mathbf{x}_B,\mathbf{x},-t), \tag{4.16}$$

$$RN_i(\mathbf{x}_B,\mathbf{x},t) = \int_{\partial\mathbb{D}_0}\int_{t'=-\infty}^{t} R(\mathbf{x},\mathbf{x}_S,t-t')N_{i-1}(\mathbf{x}_B,\mathbf{x}_S,t')\mathrm{d}t'\mathrm{d}^2\mathbf{x}_S, \tag{4.17}$$

$$N_{-1}(\mathbf{x}_B,\mathbf{x},t) = f_{1,0}^+(\mathbf{x},\mathbf{x}_B,t). \tag{4.18}$$

In our compute kernel, we update the new vectors N_i and RN_i through the convolution, instead of the focusing functions directly. Furthermore, we consider the computation

of N_i and RN_i as two separate operations, because one operation is performed in the frequency domain; the convolution with the reflection data; and the other operation is applied in the time domain; the application of the time window. The odd an even iterations of N_i are used in combination with equations (4.14) and (4.15) to update the downgoing and upgoing focusing function, respectively.

Before the iterative scheme is started, $f^+_{1,0}$, $f_{2,0}$, and N_{-1} are set equal to the time-reversed direct arrival of the Green's function $G_d(\mathbf{x},\mathbf{x}_B,-t)$, while all other vectors are set to zero. To compute the direct arrival, a smooth velocity model is required, and to start the iterative scheme, the reflection response R is also required. For convergence of the

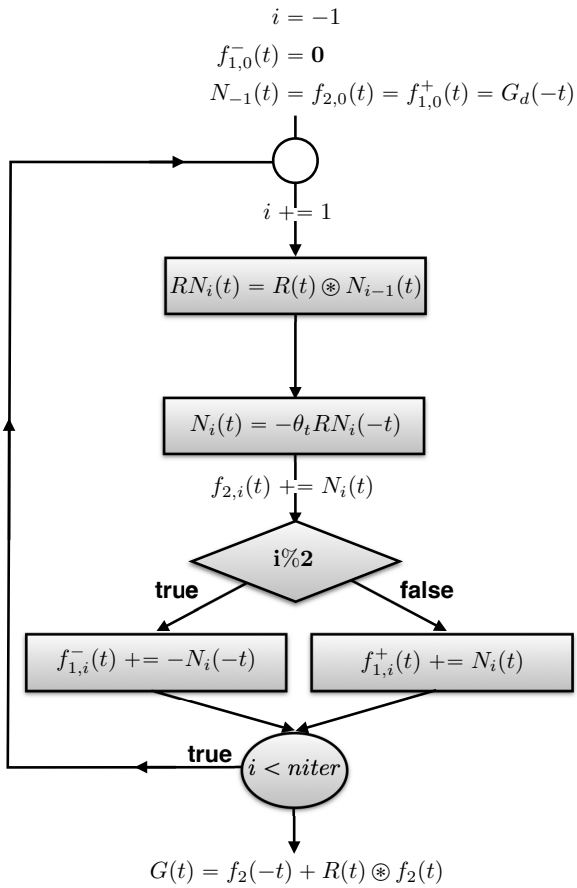

Figure 4.1: Flow chart of the Marchenko algorithm. In the notation the lateral coordinates are omitted for a more compact notation. Based on the flowchart from Thorbecke *et al.* [34].

iterative Marchenko method, the 3D reflection data must be pre-processed to comply with the assumptions made in the derivation of the Marchenko equations [89]. This processing has to at least include [81];

- removal of free-surface multiples,
- deconvolution with source wavelet,
- crossline interpolation to avoid aliasing.

Following Algorithm 1, the first step is to read the pre-processed reflection data from disk, which is generally an inexpensive task in 2D. The full 3D reflection data matrix, however, can be large and to reduce the data-size a ZFP based compression algorithm is used [88]. Before the reflection data are compressed, they are transformed to the frequency domain, as the reflection data is only used in the frequency domain for the convolution, and only the data in the frequency-band of interest are compressed and stored to disk. Typically this lossless compression reduces the 3D data size by a factor of 4, which decreases the storage space of the data and the read-in time to memory. The program `TWtransform` (explained in Appendix 4A) transforms uncompressed reflection time-data to the frequency (ω) domain, applies ZFP compression (based on tolerance) on a selected frequency range and writes the compressed data to disk. The data on disk contain a special compressed header that includes all location information present in the uncompressed Segy/SU headers that are needed in the 3D Marchenko program. The 3D Marchenko program has multiple options to read the reflection data, which can be done in the time-domain, frequency-domain, or compressed frequency-domain. In Appendix 4A, the most important options and parameter settings of the 3D Marchenko program are explained in more detail.

The first arrival time of the focal point of interest is read from disk as well. This first arrival time can be a full shot record e.g. modelled by finite difference, or the output of an Eikonal, or ray-based traveltime solver. In Appendix 4A, we give a description of our 3D finite difference modeling code `fdelmodc3D` and our 3D Eikonal solver `raytime3D`, which are used for these purposes. The first arrival times are used to construct a time window in order to separate the focusing function from the Green's function. A shift ϵ is applied to the time window to take into account the signal width and to avoid including a possible reflection event at time $|t_d|$ in the focusing function. This can happen if the focal point is located close to a reflector in the medium.

After the initializations of the focusing functions, the Marchenko iterations can start and there are no specific algorithmic changes needed for the 3D implementation. The main difference is the addition of a dimension in x_2-direction, to account for the change from 2D to 3D. After the program has computed the focusing function, equations (4.1) and (4.2) can be used to calculate the decomposed Green's function and equations (4.12) and (4.13) can be used to calculate the full focusing function and Green's function. In the iteration kernel, the *if* statement ensures that the correct focusing function is updated depending on whether the iteration is even or odd. The focusing function f_2 is updated in each iteration by N_i, as it is composed of both decomposed focusing functions.

The program can be run efficiently by computing the results for multiple focal locations simultaneously (an additional outer loop not shown in Algorithm 1), which means that the 3D reflection data only have to be read from the disk once. This is especially useful in 3D as reading in the 3D reflection data takes up a significant portion of the compute time.

Main **begin**
  Reading SU-style input parameters and Allocate arrays
  READ($R[N_{shots}, \omega, N_{recv}]$)
  makeWindow3D($G_d[N_{shots}, t], muteW[N_{shots}]$)
  Initialisation
  $N_{-1}[N_{shots}, t] = f^+_{2,0}[N_{shots}, t] = f^+_{1,0}[N_{shots}, t] = G_d[N_{shots}, -t]$
  $f^-_{1,0}[N_{shots}, t] = \mathbf{0}$
  for $i \leftarrow 0$ **to** n_i **do**
    synthesis3D(R, N_{i-1}, RN_i)
    $N_i[N_{shots}, t] = -RN_i[N_{shots}, -t]$
    applyMute3D($N_i, muteW[N_{shots}]$)
    $f_{2,i}[N_{shots}, t] = f^+_{2,i}[N_{shots}, t] + N_i[N_{shots}, t]$
    if *(iter % 2 == 0)* **then**
      $f^-_{1,i}[N_{shots}, t] = f^-_{1,i}[N_{shots}, t] - N_i[N_{shots}, -t]$
    else
      $f^+_{1,i}[N_{shots}, t] = f^+_{1,i}[N_{shots}, t] + N_i[N_{shots}, t]$
    end
  end
  synthesis3D(R, f_{2,n_i}, G_{n_i})
  $G_{n_i}[N_{shots}, t] = G_{n_i}[N_{shots}, t] + f^+_{2,n_i}[N_{shots}, -t]$
end

Algorithm 1: 3D Marchenko algorithm as implemented in the provided source code. The arrays in this algorithm are stored in C-order; the last (most right) addressed dimension is contiguous in memory. The dimensions of these arrays are within square brackets [...], the arguments of function calls are within regular brackets (...).

The overview in Algorithm 2 shows how the main compute kernel is implemented. The function `synthesis3D` transforms the term N_{i-1} to the frequency domain, where it is convolved with the reflection response and integrated over the receiver coordinates. The Fourier transform is indicated by $\mathscr{F}\{\ldots\}$. After the integration, the result is transformed back to the time domain to create RN_i so that the time window can be applied and N_i can be created. Thus, for each iteration i, an updated N_i is transformed to the frequency domain, convolved with R and the result is transformed back to time. The integration is carried out over all horizontal (x_1, x_2) source positions at the surface for each shot-record in the reflection data matrix R. The loop over the number of shots is parallelised using OpenMP so that the integration can be performed as quickly as possible for any number of focal positions.

Additionally, the `Marchenko3D` program contains a module `imaging3D` that computes a common application of the method, namely the imaging of the medium. The module can obtain an image of the medium at the location of the focal point using the double-focusing method, as described in Staring *et al.* [38]. We use Equation (11) from

synthesis3D($R[N_{shots}, \omega, N_{recv}), N_{i-1}[N_{shots}, t], RN_i[N_{shots}, t]$)
begin
  $Fop[\omega, N_{shots}] = \mathscr{F}\{N_{i-1}[N_{shots}, t]\}$
  $RN_i[N_{shots}, t] = 0$
  #pragma omp parallel for
  for $k \leftarrow 0$ **to** N_{shots} **do**
    for $\omega \leftarrow \omega_{min}$ **to** ω_{max} **do**
      for $j \leftarrow 0$ **to** N_{recv} **do**
        $sum[\omega] = sum[\omega] + R[k, \omega, j] * Fop[\omega, j]$
      end
    end
    $RN_i[k, t] = \mathscr{F}^{-1}\{sum[\omega]\}$
  end
end

Algorithm 2: Marchenko synthesis kernel with a parallel OpenMP loop over the number of shots.

this paper:

$$G^{-,+}(\mathbf{x}_B, \mathbf{x}_B, t) = \int_{\partial\mathbb{D}_0} \int_{t'=-\infty}^{t} G^{-}(\mathbf{x}_B, \mathbf{x}, t - t') f_1^{+}(\mathbf{x}, \mathbf{x}_B, t') \mathrm{d}t' \mathrm{d}^2\mathbf{x}, \tag{4.19}$$

where $G^{-,+}(\mathbf{x}_B, \mathbf{x}_B, t)$ is the upgoing Green's function measured by a virtual receiver at focal location $\mathbf{x}_B$ due to a downward radiating virtual source at the same focal location. By taking the zero-time sample, one can obtain the local reflectivity at the focal point. As the downgoing focusing function and the upgoing Green's function are obtained for the same focal position, the module `imaging3D` requires no additional inputs. In Appendix 4A, this module is explained further.

4.4. NUMERICAL EXAMPLES

We demonstrate the application of the 3D Marchenko scheme on two models. The first is a simple model with four layers that only varies in the vertical direction. This is the same model as was used in Thorbecke *et al.* [34], only extended to 3D, to demonstrate the basics of the 3D implementation. To take into account the complex scattering in 3 dimensions, we also apply the method on a subsection of the SEG/EAGE Overthrust model from Aminzadeh *et al.* [87]. For both models, we show iterations of the focusing functions and Green's function and use these to obtain an image of the model.

4.4.1. HORIZONTALLY LAYERED MODEL

The velocity and density of the horizontally layered model are shown in Figures 4.2(a) and 4.2(b), respectively. A 3D shot record with its source located at the surface in the center of the model is shown in Figure 4.2(c). For the full reflection response, we use

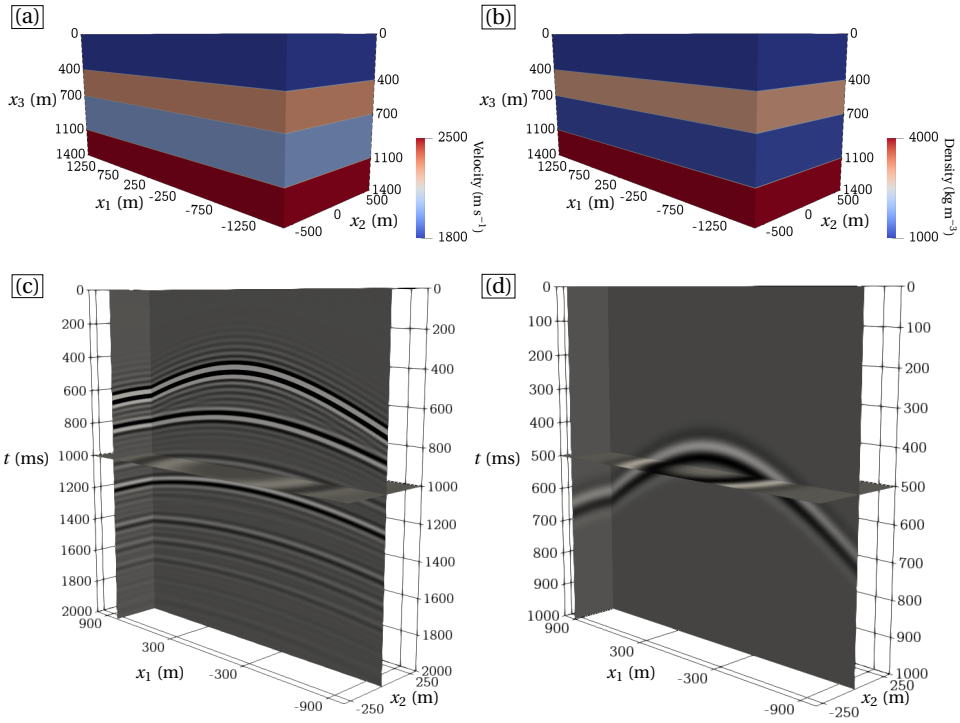

Figure 4.2: Four layer model with velocity (a) and density (b) contrasts. (c) A shot record, with source position $\mathbf{x}_s = (0,0,0)$ and receivers at $\mathbf{x} = (x_{1,r}, x_{2,r}, 0)$, convolved with a wavelet with a flat spectrum between 5 and 30Hz. (d) A shot record containing the first arrival response from a source at $\mathbf{x}_B = (0,0,900)$, convolved with a 15Hz Ricker wavelet. The inverse of this shot record is the initial estimate of the focusing function.

a fixed spread acquisition, where the source is modeled at every receiver position. For the modeling of the reflection data, we apply a wavelet with a flat frequency spectrum that introduces ringing in the time-domain. This approximates the pre-requisite of deconvolving for the source wavelet [34]. The frequency spectrum of this wavelet is flat between 5 and 30Hz, and tapered to zero outside this range. The range of the frequency spectrum is chosen for modeling runtime purposes. Figure 4.2(d) shows the shot record of the first arrival associated with a focal point at position $\mathbf{x}_B = (0,0,900)$, convolved with a 15Hz Ricker wavelet. The receivers are located at the surface of the model ($x_3 = 0$), with an offset in the inline x-direction of -1000 to 1000m and in the crossline x_2-direction -300 to 300m with receiver spacing of $\mathrm{d}x_1 = \mathrm{d}x_2 = 10$m. The receiver locations in the first arrival shot-record coincide with the receiver locations of the reflection data. All shot records in this example are computed using the finite difference program `fdelmodc3D`; see Appendix 4A for more information.

The data from Figure 4.2 are used in the Marchenko scheme. To demonstrate how the application works, we look at individual iterations of the program. Figure 4.3 shows the inline $x_1\,t$-slices and Figure 4.4 the crossline $x_2\,t$-slices of the initial data and the first,

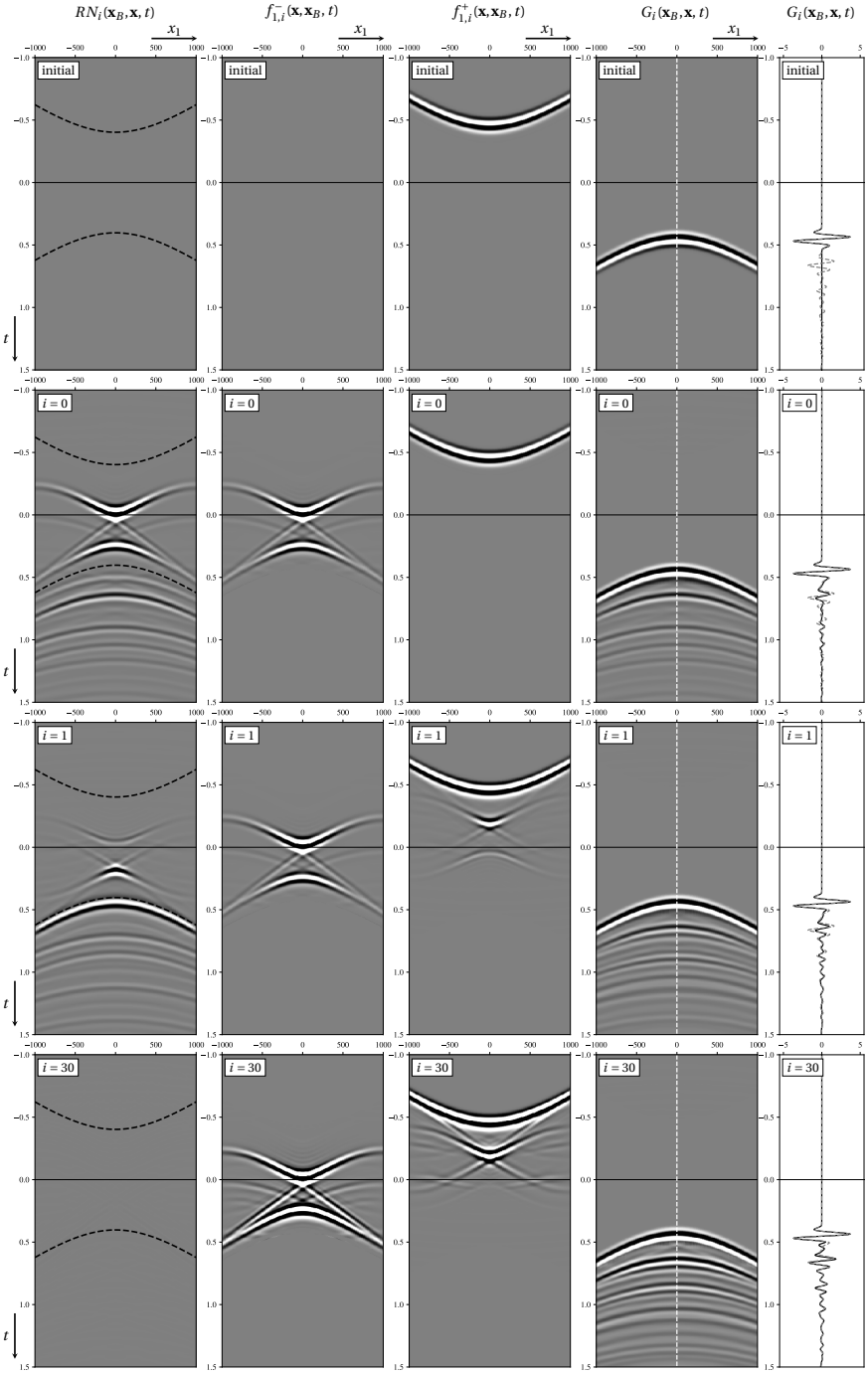

Figure 4.3: Four successive iterations of the Marchenko method for the horizontally layered model sliced along the x_1-direction, for a constant x_2-offset of 0m. The 4^{th} column shows the full Green's function for a source at the focal point and receivers at the surface. The white line in this column is a trace that is plotted in solid black in the 5^{th} column, where the dashed grey line is a directly modeled Green's function. The clip level is the same for all panels.

second and 30^{th} iteration of the 3D Marchenko method. For the inline slices, the panels are taken along a constant x_2-position of 0m and for the crossline direction the slices are taken along a constant x_1-position of 300m. The $f_{1,i}^{+}(\mathbf{x},\mathbf{x}_B,t)$ column for the initial data shows the first estimation and consists of the direct arrival associated with the focal point at $(0,0,900)$m. The first arrival is modeled in the exact medium and is inverted to create an accurate estimation of the first arrival of the focusing function. This is done to make a fair comparison between the retrieved Green's function and a directly modeled reference solution. The response of this first arrival with R is computed by the synthesis process (Algorithm 2) and results in the record in the first column, labeled RN_i. From the panels in Figures 4.3 and 4.4 (all with the same clipping factor), it can be observed that the amplitude of RN_i decreases as the iteration count increases. In the 30^{th} iteration, the update term RN_i does not give a visible contribution anymore. The dashed black lines in the RN_i figures represent the time window w_t that separates the Green's function, below the lower truncation line, and the focusing function, in between the truncation lines.

The upgoing focusing function $f_{1,i}^{-}(\mathbf{x},\mathbf{x}_B,t)$ is shown in the second column, while the downgoing focusing function $f_{1,i}^{+}(\mathbf{x},\mathbf{x}_B,t)$ is shown in the third column. The Green's function is shown in the fourth column. In the fifth column, the solid black line is the central trace from the Green's function in the fourth column, which is indicated by the vertical white dashed line in the fourth column. The dashed grey line is a reference solution of the Green's function that was obtained by modeling the wavefield in the exact medium. These traces show that while initially only the first arrival is obtained, subsequent iterations introduce additional events and artifacts are suppressed, while other events are amplified. After 30 iterations, the iterative scheme has converged and there is a strong match between the reference Green's function and the computed Green's function. Note that, compared to the 2D approach that was presented in Thorbecke *et al.* [34], there seem to be more events present in the upgoing focusing function. These are artefacts associated with the limited aperture of the data. Due to storage limitations and computational costs, the offsets of the 3D data that are considered in this chapter are limited, and therefore the stationary phase area is small. Because the layers in this model are all flat, the structure can still be resolved, however, for more complex structures, which include dipping layers, this will become a more significant problem, especially in the crossline direction. When the integration over all source positions is performed, additional artefacts will be introduced due to missing stationary point contributions.

The curve in Figure 4.5 shows the energy of the update N_i in Algorithm 1 relative to the energy of the first estimation N_0. If the relative energy decreases with increasing iterations, it indicates that the Marchenko method is converging towards a solution. The lower the relative energy in this update the smaller the contribution of N_i to the focusing functions and the smaller the update in the Green's function. Because the aperture is limited, there will be artefacts present in N_i that cannot be removed by the iterative scheme. As a result the energy in the update cannot converge to zero, but is bound by the energy of these aperture artefacts. Figure 4.5 shows the relative energy for each iteration, which is a continually decreasing logarithmically smooth convergence curve, indicating that successive iterations give smaller updates. After 30 iterations, the updates are 3 orders of magnitude smaller than the first update, which is low enough to conclude that the iterative scheme no longer gives significant energy updates anymore and has

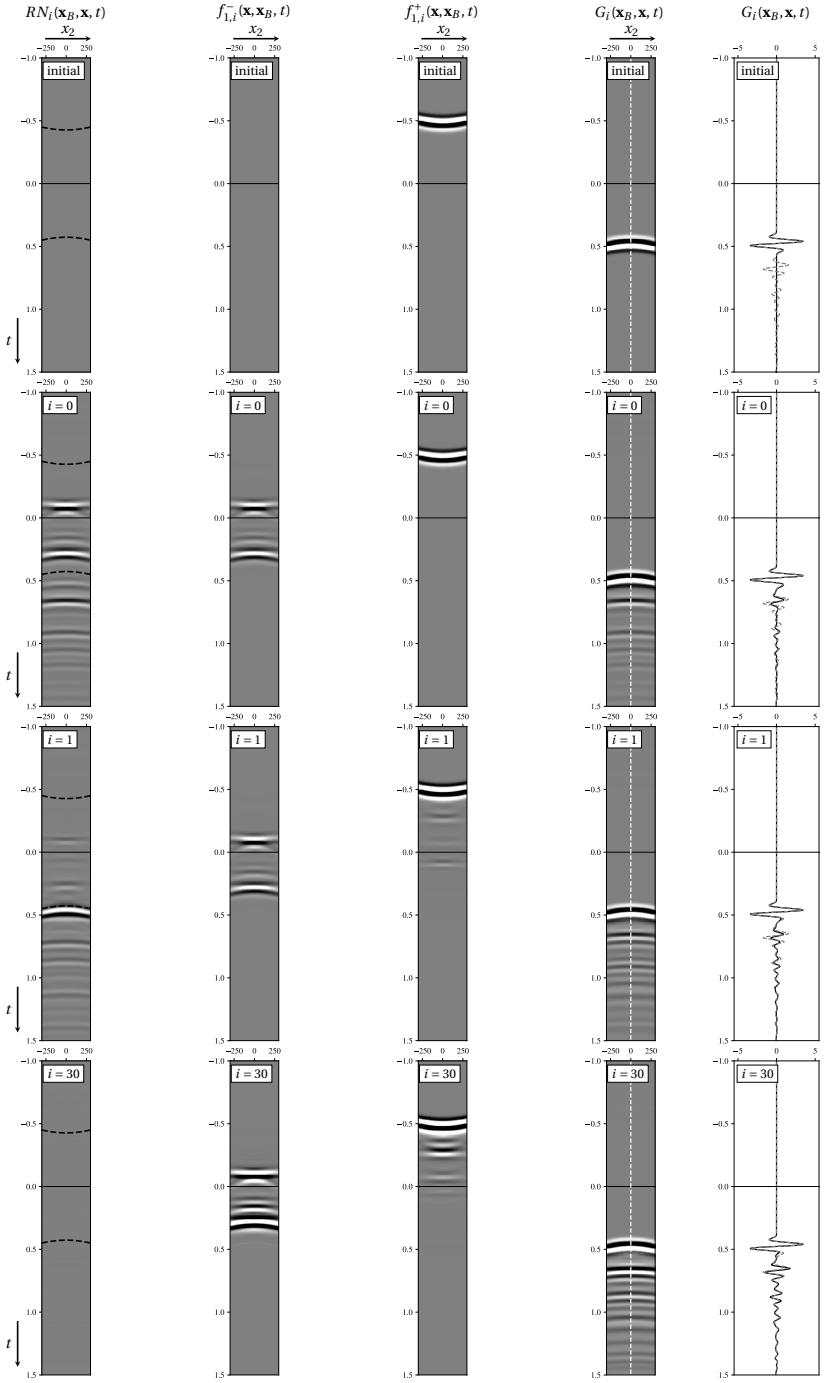

Figure 4.4: Four successive iterations of the Marchenko method for the horizontally layered model sliced along the x_2-direction, for a constant x_1-offset of 300m. The 4^{th} column shows the full Green's function for a source at the focal point and receivers at the surface. The white line in this column is a trace that is plotted in solid black in the 5^{th} column, where the dashed grey line is a directly modeled Green's function. The clip level is the same for all panels.

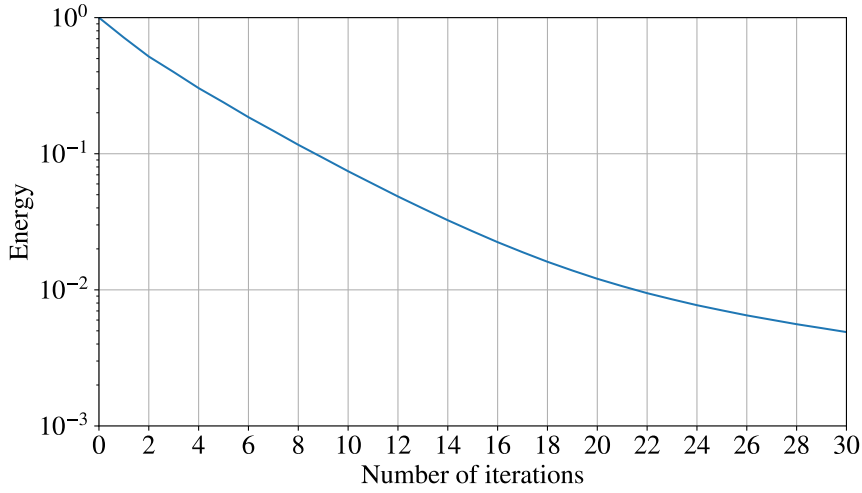

Figure 4.5: Logarithmic convergence rate of the `marchenko3D/demo/marchenko3D/oneD` example for 30 iterations using the first arrival from Figure 4.2(d). The energy level is an indication of the accuracy reached by the method.

therefore converged. This is supported by the results in the bottom rows of Figures 4.3 and 4.4.

To further demonstrate the application of our scheme, we image the model using Equation (4.19), which is implemented in the `imaging3D` module. Due to the large amount of focal points that are required for this approach, it is not feasible to model all the first arrivals with the finite-difference code. Instead, we make use of a 3D Eikonal solver, `raytime3D`, to obtain the first arrival times. Furthermore, to approximate field conditions, we smooth the velocity model in 3D and use no density information. The Eikonal solver is capable of retrieving a geometric spreading factor to account for amplitude variation along the wavefront. Note that this is just an approximation of the amplitude. We use the time-reversal of these arrivals as the first estimation instead of the inverted first arrivals, because it is not possible to obtain the exact scaling either way. Using the Eikonal solver, we compute travel-times for focal points in a depth range from 200 to 1200m with a depth spacing of 10m. The images are computed along two planes, one in the inline direction from -1000 to 1000m and one in the crossline direction from −300 to 300m, both with a horizontal spacing of 10m. The inline image has a constant offset in the crossline direction of 0m, and the crossline image has a constant offset in the inline direction of 0m, which means the images intersect each other in their respective centers. Due to the large amount of focal points, the number of iterations for each focal point is limited to 20. The results of the imaging are shown in Figures 4.6(b) and 4.6(d) for the inline direction and the crossline direction, respectively. For comparison, we also performed conventional imaging for the same planes, which are shown in Figures 4.6(a) and 4.6(c), respectively. The conventional imaging contains a clear artefact, caused by

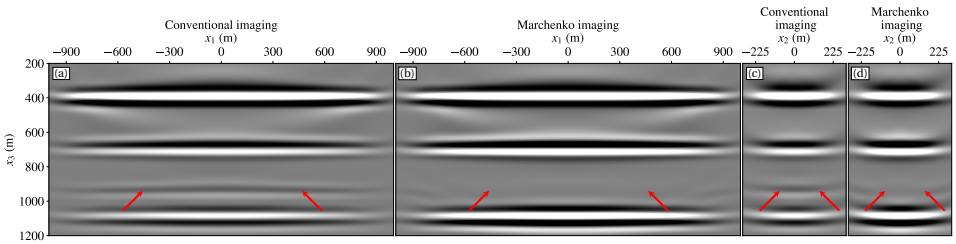

Figure 4.6: Image of the horizontally layered model along a fixed x_2 value of 0m using (a) conventional imaging and (b) Marchenko imaging after 20 iterations, and image of the horizontally layered model along a fixed x_1 value of 0m using (c) conventional imaging and (d) Marchenko imaging. The locations of artefacts that are attenuated by the Marchenko imaging are indicated by the red arrows.

4

the internal multiples, which is indicated by the red arrows. This artefact is attenuated by the Marchenko imaging. As the medium only has flat layers, the layer contrasts and the multiple artefact are present at the same depth in both the inline and the crossline direction. Due to the more limited aperture in the crossline direction, the image in this direction is less flat, especially at the edges of the aperture, which also limits the effect of the imaging in this direction. However, in all cases the Marchenko imaging shows a clear improvement over the conventional imaging.

4.4.2. SEG/EAGE OVERTHRUST MODEL

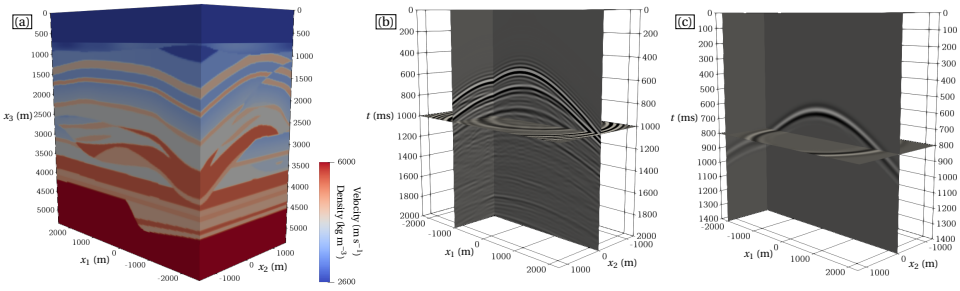

Figure 4.7: (a) Subsection of the Overthrust model used for both the velocity and density contrasts. (b) A shot record, with source position $\mathbf{x}_S = (0,0,0)$ and receivers at $\mathbf{x} = (x_{1,r}, x_{2,r}, 0)$, convolved with a wavelet with a flat spectrum between 5 and 30Hz. (c) A shot record containing the first arrival response from a source at $\mathbf{x}_B = (0,0,2050)$, convolved with a 15Hz Ricker wavelet. The inverse of this shot record is the initial estimate of the focusing function.

To validate our implementation for true 3D models, we use the SEG/EAGE Overthrust model by Aminzadeh *et al.* [87], that is publicly available from the SEG Wiki. We select a subsection of the model, as a recording setup over the full extent of the model is too large to fit in the memory of our compute nodes. Furthermore, this reduces the modeling time for the reflection response. We insert a layer with constant velocity and density above the model to simulate a water layer. The velocity and density of the sub-

section are shown in Figure 4.7(a). The density model is chosen the same as the velocity model to ensure strong reflections. Similarly to the horizontally layered model, we use the `fdelmodc3D` code to model the reflection response and the first arrival from the focal point. An example of a shot record from a source at the surface in the center of the model is shown in Figure 4.7(b), and the shot record containing the first arrival for a focal point at $\mathbf{x}_B = (0, 0, 2050)$ is shown in Figure 4.7(c). The reflection data are modeled using the same wavelet with a flat spectrum that we use for the horizontally layered medium and the first arrival is likewise modeled using the same Ricker wavelet as is used for the first arrival of the horizontally layered medium. Again, this first arrival is modeled in the exact model and inverted to create an accurate first arrival estimate for the focusing function. While the setup of the recording array is once again fixed spread, with the sources located at every receiver location, the array has coarser sampling than before. In the inline x_1-direction the source and receiver are distributed from -2250 to 2250m with a spacing of 25m and in the crossline x_2-direction the extent varies from -1250 to 1250m with a spacing of 50m. While the aperture in this case is larger than before, the sampling is much coarser, especially in the crossline direction, and the model is more complex. Note that the sampling in the inline and crossline directions is not equal, as is often the case for acquisition setups in the field. The results for the Overthrust model are shown for the inline x_1-direction in Figure 4.8 and for the crossline x_2-direction in Figure 4.9. The inline panels are located along a constant offset in the x_2-direction of 0m and the crossline panels are located along a constant offset in the x_1-direction of 500m. The Green's function G_i in Figure 4.8 shows many events arriving after the first arrival with varying arrival times and amplitude. Similarly, both the upgoing and downgoing focusing function contain more events with more variations than the focusing functions obtained in the horizontally layered medium. This demonstrates that the complexity of the model is taken into account by the Marchenko method, however, it also shows that there are more events to resolve. The iteration count for this focal position is higher than before, because of this complexity. After 40 iterations, the method converges to a result where the updates of RN_i contain little energy compared to the first iteration. While the traces of the Green's function in the fifth column show a strong match, there are some small errors, likely caused by the complexity of the model and the larger spatial sampling. Comparing the first iteration and the final iteration it becomes clear that the method improves the amplitude of the desired events and removes the undesired internal multiples. The results in the crossline direction, shown in Figure 4.9 display a similar quality of results, although they are not as smooth as the results in the inline direction. This is caused by the coarser spatial sampling in the crossline direction. There are artefacts present at the edges of the panels, caused by the limited aperture. The traces of the Green's function show a similar match as the ones in the inline direction, despite the coarser sampling. These results do show that the Marchenko method does not require equal sampling in the inline and crossline direction. However, if the reflection data become subsampled in either the inline or crossline direction, the result deteriorates in quality.

Compared to the convergence rate of the previous four layer model (Figure 4.5), the energy in Figure 4.10 converges at a slower rate and also to a higher energy level. The convergence to a higher energy level indicates that more limited-aperture artefacts are

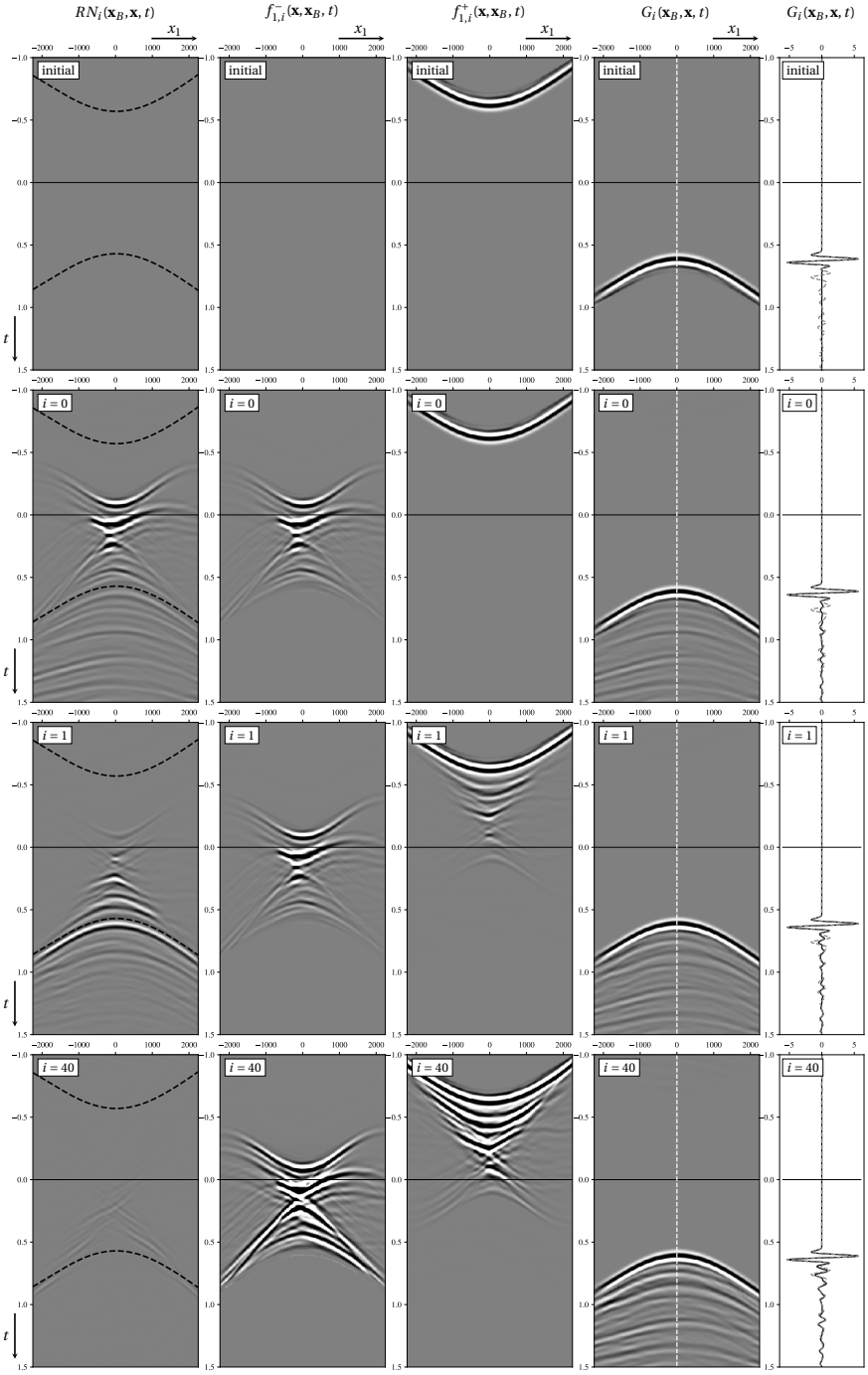

Figure 4.8: Four successive iterations of the Marchenko method for the Overthrust model sliced along the x_1-direction, for a constant x_2-offset of 0m. The 4^{th} column shows the full Green's function for a source at the focal point and receivers at the surface. The white line in this column is a trace that is plotted in solid black in the 5^{th} column, where the dashed grey line is a directly modeled Green's function. The clip level is the same for all panels.

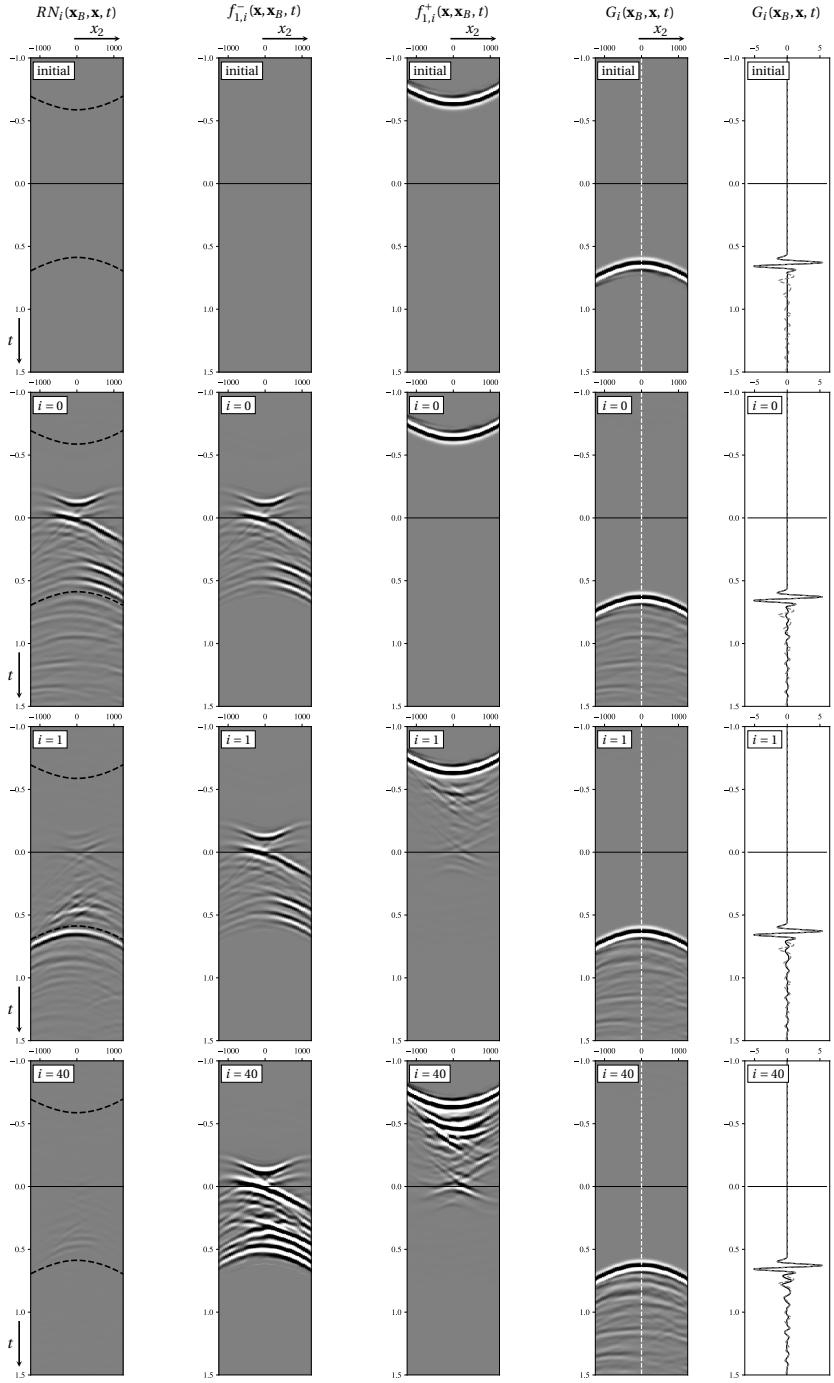

Figure 4.9: Four successive iterations of the Marchenko method for the Overthrust model sliced along the x_2-direction, for a constant x_1-offset of 500m. The 4^{th} column shows the full Green's function for a source at the focal point and receivers at the surface. The white line in this column is a trace that is plotted in solid black in the 5^{th} column, where the dashed grey line is a directly modeled Green's function. The clip level is the same for all panels.

present in the update fields N_i. These limited-aperture effects are caused by the acquisition footprint, especially in the crossline direction. While the aperture of the recording array is larger for the Overthrust model than for the horizontally layered medium, the complexity of the model causes scattering at larger angles, which means that a larger aperture is required to properly capture the reflection data. The reason for the increased amount of iterations that are required is due to the fact that more events need to be resolved for the Overthrust model. We also obtain an image of the Overthrust model using

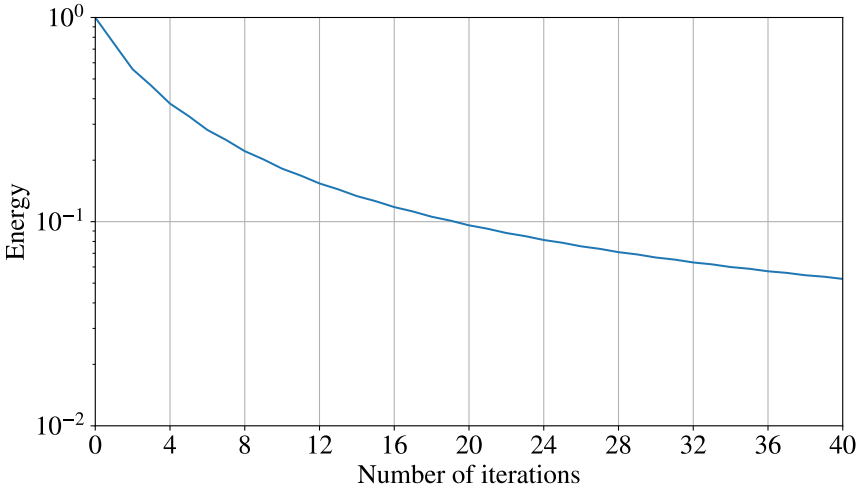

Figure 4.10: Logarithmic convergence rate of the Overthrust model for 40 iterations.

the 3D implementation. Again, we take two cross sections, one inline for a fixed x_2-offset of 0m and one crossline for a fixed x_1-offset of 0m, which intersect in their respective centers. The travel-times from the focal points to the surface are obtained using our 3D Eikonal solver, modeled in a smoothed version of the velocity model, along a depth range of 400 to 4400m with a sampling of 12.5m. The sampling in the inline direction is 25m and in the crossline direction it is 50m. The first arrivals are time-reversed instead of inverted, for the sake of simplicity. The amount of iterations for these focal points are limited to 30 due to the large computational costs. Figure 4.11(b) shows the result for the inline direction and 4.11(d) shows the result for the crossline direction. Conventional images for the inline and crossline directions are shown in Figures 4.11(a) and 4.11(c), respectively. As can be seen from the figures, the subsurface is much more complex and harder to resolve. Due to the small frequency bandwidth, the resolution of the images is limited, however, there are still artefacts present caused by the internal multiples, as indicated by the red arrows. The Marchenko imaging attenuates these artefacts, which shows that even on complex 3D models, the method can produce good results.

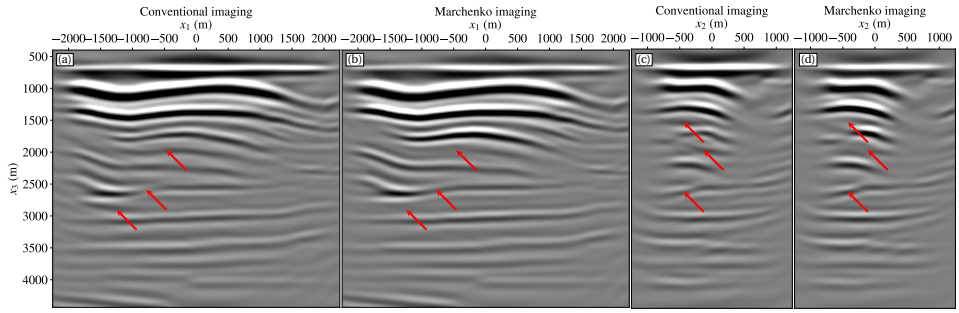

Figure 4.11: Image of the Overthrust model along a fixed x_2 value of 0m using (a) conventional imaging and (b) Marchenko imaging after 30 iterations, and image of the Overthrust model along a fixed x_1 value of 0m using (c) conventional imaging and (d) Marchenko imaging. The locations of artefacts that are attenuated by the Marchenko imaging are indicated by the red arrows.

4.5. CONCLUSIONS

In this chapter, we have shown that our Marchenko implementation can be applied on 3D reflection data. We achieved this by making use of an efficient computing kernel that updates the upgoing and downgoing focusing function during even and odd iterations, respectively. Furthermore, our application made use of the ability to compress the reflection data to limit the amount of data that needs to be read in. Due to the requirements of the Marchenko method for the reflection response, concessions had to be made for generating the reflection data, including limited frequency content, coarse spatial sampling and limited aperture. Despite these limitations, the method still converged to an accurate result. We demonstrated this using two models, a simple flat layered model and a complex model that caused full 3D scattering. On the simple model, the method retrieved a Green's function that showed a strong match to a reference solution. Furthermore, the imaging of the medium removed a clear internal multiple artefact, that was present when conventional imaging was used, for both the inline and crossline direction. On the second model that was considered, a subsection of the SEG/EAGE Overthrust model, less accurate results were achieved, and more iterations were required. The first effect was caused by the coarser sampling of the reflection response, while the second effect was caused by the increased complexity of the model. The results showed that the poorer sampling of the reflection data for this model in the crossline direction than in the inline direction did not cause the result to converge to an incorrect solution. The retrieval of the Green's function was still possible and the imaging of the model showed clear attenuation of artefacts.

4A INPUT FOR MARCHENKO3D AND AUXILIARY PROGRAMS

MARCHENKO3D

The marchenko3D program has the following parameters and options:

```
MARCHENKO3D - Iterative Green's function and focusing functions retrieval in 3D

marchenko3D file_tinv= file_shot= [optional parameters]

Required parameters:

  First arrival input options:
  file_tinv= ............... direct arrival from focal point: G_d
  file_ray= ................ direct arrival from raytimes
  Shot data input options:
  file_shot= ............... Reflection response (time data): R(t)
  file_shotw= .............. Reflection response (frequency data): R(w)
  file_shotzfp= ............ Reflection response (frequency compressed data): zfp[R(w)]

Optional parameters:

INTEGRATION
  ampest=0 ................. Estimate a scalar amplitude correction with depth (=1)
  tap=0 .................... lateral taper focusing(1), shot(2) or both(3)
  ntap=0 ................... number of taper points at boundaries
  fmin=0 ................... minimum frequency in the Fourier transform
  fmax=70 .................. maximum frequency in the Fourier transform
MARCHENKO ITERATIONS
  niter=10 ................. number of iterations
MUTE-WINDOW
  file_amp= ................ amplitudes for the raytime estimation
  file_wav= ................ Wavelet applied to the raytime data
  above=0 .................. mute above(1), around(0) or below(-1) the travel times of
  ........................... the first arrival
  shift=12 ................. number of points above(positive) / below(negative) travel
  ........................... time for mute
  hw=8 ..................... window in time samples to look for maximum in next trace
  smooth=5 ................. number of points to smooth mute with cosine window
  plane_wave=0 ............. enable plane-wave illumination function
REFLECTION RESPONSE CORRECTION
  scale=2 .................. scale factor of R for summation of Ni with G_d (only for
  ........................... time shot data)
  pad=0 .................... amount of samples to pad the reflection series
HOMOGENEOUS GREEN'S FUNCTION RETRIEVAL OPTIONS
  file_inp= ................ Input source function for the retrieval
  scheme=0 ................. Scheme for the retrieval
  ........................... scheme=0 Marchenko homogeneous Green's function retrieval
  ........................... with G source
  ........................... scheme=1 Marchenko homogeneous Green's function retrieval
  ........................... with f2 source
  ........................... scheme=2 Marchenko Green's function retrieval with source
  ........................... depending on virtual receiver location
  ........................... scheme=3 Marchenko Green's function retrieval with G
                               source
  ........................... scheme=4 Marchenko Green's function retrieval with f2
                               source
  ........................... scheme=5 Classical homogeneous Green's function retrieval
```

```
   ......................... scheme=6 Marchenko homogeneous Green's function retrieval
                              with multiple G sources
   ......................... scheme=7 Marchenko Green's function retrieval
                              with multiple G sources
   ......................... scheme=8 f1+ redatuming
   ......................... scheme=9 f1- redatuming
   ......................... scheme=10 2i IM(f1) redatuming
   cp=1000.0 ................ Velocity of upper layer for certain operations
   rho=1000.0 ............... Density of upper layer for certain operations
 OUTPUT DEFINITION
   file_green= .............. output file with full Green function(s)
   file_gplus= .............. output file with G+
   file_gmin= ............... output file with G-
   file_f1plus= ............. output file with f1+
   file_f1min= .............. output file with f1-
   file_f2= ................. output file with f2
   file_imag= ............... output file with image
   file_homg= ............... output file with homogeneous Green's function
   file_ampscl= ............. output file with estimated amplitudes
   file_iter= ............... output file with -Ni(-t) for each iteration
   compact=0 ................ Write out homg and imag in compact format
   ......................... WARNING! This write-out cannot be displayed with SU
   verbose=0 ................ silent option; >0 displays info
```

The input of the 3D Marchenko method is similar to the 2D implementation, requiring reflection data and the first arrival time from the focal point. However, due to the large size of 3D data, reading in the reflection data using the `file_shot` option can become time-consuming. To mitigate this problem, the 3D implementation gives two alternate options of reading in data. The first is reading in the shot data in the frequency domain using `file_shotw`. These data have been pre-transformed to the frequency domain, which avoids the Fourier transform that was required on the shot data in the time domain. Alternatively, the frequency data can be compressed using the ZFP algorithm [88] before reading in, to reduce the file size. The code requires one of these three data types as input, and the latter two options can be obtained using the `TWtransform` module.

The second required input, the first arrival time, can be passed to the code in two ways. The first is by reading in a shot record using the `file_tinv` option. An alternative is using the arrival times that are calculated by a 3D Eikonal solver, an example of which is the `raytime3D` program that is part of the software distribution, based on the work by Vidale [90]. The `raytime3D` program also computes a geometric spreading factor that can be used to estimate the amplitude of the first arrivals. This file can be read in using the `file_amp` option. To approximate seismic broadband data, a wavelet can be read into the code as well using `file_wav`. The file needs to contain a wavelet with no time shift and have the same temporal sampling as the reflection data. For both types of input for the first arrival, multiple focal points can be read in at the same time.

The number of iterations required for convergence depends on the reflection strengths and on the number of events in the model; a complex model will need more iterations. Typically the number of iterations is chosen between 8 and 20. Setting the `verbose=2` option will compute the convergence of the algorithm by printing the energy of the iteration update $N_i(t)$ relative to the initial value $N_0(t)$. The energy in the

update term $N_i(t)$ should become smaller in each iteration.

The `marchenko3D` program is capable of computing an image for the focal points and outputting them directly through use of the `imaging3D` module of the code. To do this, one can simply set the `file_imag` option and the code computes an image point for each given focal-point.

The program also contains an additional module called `homogeneous3D`, which is used for the purpose of retrieving the wavefield between two focal points in the subsurface. By setting the option `file_homg` to a correct path, the wavefield is computed according to the scheme set by the option `scheme`. Most of these schemes are explained in Brackenhoff *et al.* [62] and Brackenhoff *et al.* [81]. All of the schemes require that the module is given input data by the `file_inp` option. This input file needs to be sampled at the same positions as the data that are computed in the main `marchenko3D` program with the same sample length and distance. It is recommended to run the module for one focal position first before using it on a large amount of focal positions.

To compensate for the transmission losses, an approximate amplitude correction can be used [85]. This estimation can be added to the results by setting the `ampest` option equal to 1. By using the `file_ampscl` option, the estimated amplitude correction are written out for each focal point.

The code to reproduce the figures of the flat layered model in this chapter can be found in the directory `marchenko3D/demo/marchenko3D/oneD`. The `README` file in that directory explains in detail how to run the scripts. The SEG/EAGE Overthrust model by Aminzadeh *et al.* [87] can be found on the SEG wiki at https://wiki.seg.org/wiki/SEG/EAGE_Salt_and_Overthrust_Models.

FDELMODC3D

Based on the finite difference (FD) code in Thorbecke and Draganov [68] a 3D version of the acoustic implementation is made. The finite-difference kernels are extended to 3D and the code has the same functionality as the 2D code. In 3D it can be cumbersome to define 3D gridded models. In the 3D FD implementation the gridded velocity and density model can also be a 1D or 2D model and will be extended to a full 3D model by setting `ny=` for 2D models and `ny=nx=` for 1D models. In the implementation pointers are used to avoid copying full dimensions to 3D. The directory `utils` contains programs to calculate a gridded model (`makemod`), source wavelets (`makewave`) as well as programs for basic processing steps.

To validate the 3D modeling we compared the results with a 3D layer-based wavenumber-frequency code (not included). To check the accuracy of the amplitude of the modeling code (which is crucial in the Marchenko algorithm) an extra verification is made by integration of the 3D result over the y-axis and comparing that result with the 2D solution.

TWTRANSFORM

```
TWtransform - Transform data from uncompressed time domain to compressed frequency
              domain

TWtransform file_in= file_out= [optional parameters]
```

```
Required parameters:

  file_in= ................. File containing the uncompressed time domain data
  file_out= ................ Output for the (compressed) frequency domain data

Optional parameters:

  verbose=1 ................ silent option; >0 displays info
  fmin=0 ................... minimum frequency in the output
  fmax=70 .................. maximum frequency in the output
  mode=1 ................... sign of the frequency transform
  zfp=0 .................... (=1) compress the transformed data using zfp
  tolerance=1e-3 ........... accuracy of the zfp compression, smaller values
                             give more accuracy to the compressed data but
                             will decrease the compression rate
  weight=2.0 ............... scaling of the reflection data
```


The `TWtransform` program is intended to reduce the file size of the reflection data and to limit the amount of time that is required for reading the reflection data. The program transforms the time data to the frequency domain, which already reduces the file size depending on the frequency range set by `fmin` and `fmax`. The file size can be further reduced by using the `zfp` option, which will compress the frequency data using the ZFP compression by Lindstrom [88], with an accuracy set by the input parameter `tolerance`. The program also puts a custom header on the data to further reduce the file size.

5

3D VIRTUAL SEISMOLOGY

We create virtual sources and receivers in a 3D subsurface using the single-sided homogeneous Green's function representation. We employ Green's functions and focusing functions that are obtained with reflection data at the surface of the Earth, a macro velocity model and the Marchenko method. The homogeneous Green's function is a Green's function superposed with its time-reversal. Unlike the classical homogeneous Green's function representation, this approach requires no receivers on an enclosing boundary, however, it does require the source signal to be symmetric in time. We demonstrate that the single-sided representation is an improvement over the classical representation by applying the representations to numerical data that are modeled in a complex 3D model. We retrieve responses to virtual point sources with an isotropic and with a double-couple radiation pattern and compare the results to a directly modeled reference result. We also demonstrate the application of the single-sided representation for retrieving the response to a virtual rupture that consists of a superposition of point sources. This is achieved by obtaining the homogeneous Green's function for each source separately, before they are transformed to the causal Green's function, time-shifted and superposed. The single-sided representation is also used to monitor the complete wavefield that is caused by a numerically modeled rupture. However, the source signal of an actual rupture is not symmetric in time and the single-sided represenation can therefore only be used to obtain the causal Green's function. This approach leaves artifacts in the final result, however, these artifacts are limited in space and time.

This chapter has been submitted for publication to IEEE Transactions on Geoscience and Remote Sensing. Minor modifications have been applied to the text and figures for the sake of consistency in the thesis.

5.1. INTRODUCTION

Over the past few decades, the amount of induced seismicity has increased and is occurring at locations around the world [91]. While the effects of induced seismicity are often harmful, the measurements of these events can be used to gain more insight into the mechanics of earthquake rupture [92]. For example, the measurements can be used in an inversion process to obtain the seismic moment tensor, which describes the source mechanism of a seismic event [8]. The knowledge of the moment tensor as well as the location of the source can help to determine what caused the induced seismicity. These inversions often rely on an accurate velocity model of the subsurface to obtain the required wavefields [93], because errors in the velocity model can cause mistakes in the inversion result [94].

A recent development for obtaining accurate wavefields in the subsurface is the homogeneous Green's function retrieval method. A homogeneous Green's function is a Green's function superposed with its time-reversal. Porter [15] derived the original representation for the homogeneous Green's function, which was later used for inverse source problems by Porter and Devaney [16] and further researched for inverse scattering methods [17], seismic imaging [18] and seismic holography [19]. The classical representation of the homogeneous Green's function involves an integral over a closed boundary. In practical situations, data are usually available only on an open boundary. Methods like seismic imaging and holography still work well for this situation as long as only primary waves are considered. However, internal multiples are incorrectly handled and lead to artifacts when the classical representation is approximated by an integral along an open boundary.

Instead of the classical representation of the homogeneous Green's function, a single-sided representation can be used, which is designed to work with an open boundary, typically the surface of the Earth [25]. This single-sided representation is designed to correctly handle the internal multiples by employing so-called focusing functions. These focusing functions can be obtained through the use of the Marchenko method, which employs reflection data at the surface of the Earth [95]. Brackenhoff *et al.* [81] demonstrated the validity of this single-sided representation for field data.

While many applications of the Marchenko method have been performed on 2D data, recently more applications on 3D data have been achieved, for example by Pereira *et al.* [96] and Staring and Wapenaar [97]. Especially in areas where there are strong out-of-plane effects, the 2D approximation on 3D data can cause errors in the result [53]. To properly take into account the effects of wave propagation and scattering in 3D, the single-sided retrieval scheme for the homogeneous Green's function needs to be employed together with a 3D version of the Marchenko method.

In this chapter, we present the retrieval of the 3D homogeneous Green's function. We first review the classical and single-sided homogeneous Green's function retrieval schemes and apply the schemes to single source-receiver pairs. We use the 3D Open-source Marchenko method by Brackenhoff *et al.* [98] on a synthetic reflection response, that was modeled using a subset of the Overthrust model by Aminzadeh *et al.* [87], to create the required Green's functions and focusing functions for the retrieval schemes. We demonstrate the method for point sources that have an isotropic radiation pattern and compare the retrieved Green's functions to directly modeled data. Furthermore, we also

retrieve snapshots of wavefields at the virtual receivers in 3D to observe the propagation of the wavefield through the medium over time. Aside from considering an isotropic radiation pattern, we also consider the non-isotropic double-couple radiation pattern, which describes the seismic response to a pure shear fault [8]. Furthermore, we consider the retrieval of a response caused by a rupture in the subsurface by employing a series of superposed point sources with varying amplitudes and activation times and a double-couple radiation pattern, similar to Brackenhoff *et al.* [62], but extended to a 3D medium. For this latter situation we use two different approaches. One is a one-step process, where we assume that we measure the response from the rupture directly, so that we can monitor the wavefield as it propagates through the subsurface. Hence, in this one-step process we create virtual receivers to monitor the response to a real source. The other is a two-step process, where we use the Marchenko method to obtain the homogeneous Green's function for each virtual source point separately, and superpose them after each homogeneous Green's function has been obtained. Hence, in this two-step process we create virtual receivers and virtual sources. This is a way to forecast the wavefield that would be caused by the rupture, given the properties of the rupture and reflection data at the surface. We illustrate the methods with numerical examples. When we speak, for the sake of argument, of measurements of the response to a real source, in the examples these measurements are simulated by numerical modeling.

5.2. 3D VIRTUAL SEISMOLOGY

5.2.1. WAVEFIELDS

We consider a Green's function, $G = G(\mathbf{x}, \mathbf{x}_A, t)$, which describes the response of a medium at time t and position $\mathbf{x} = (x_1, x_2, x_3)$, due to an impulsive point source at $\mathbf{x}_A$, using a Cartesian coordinate system. In the coordinate system that we use, the third principal direction points downwards. The Green's function is the solution to the following acoustic wave equation:

$$\partial_i(\rho^{-1}\partial_i G) - \kappa\partial_t^2 G = -\delta(\mathbf{x}-\mathbf{x}_A)\partial_t\delta(t), \tag{5.1}$$

where $\rho = \rho(\mathbf{x})$ is the density of the medium in kg m^{-3}, $\kappa = \kappa(\mathbf{x})$ is the compressibility in kg^{-1} m s^2, $\partial_i = \frac{\partial}{\partial x_i}$ is the component of a vector consisting of the partial differential operators in the three principal directions of the coordinate system, $\partial_t = \frac{\partial}{\partial t}$ is the temporal partial differential operator and $\delta(\cdot)$ is a Dirac delta function. In case of repeating subscripts, Einstein's summation convention applies. The Green's function is causal; i.e. $G(\mathbf{x}, \mathbf{x}_A, t) = 0$ for $t < 0$, hence, it propagates away from the source location; and obeys source-receiver reciprocity so that $G(\mathbf{x}, \mathbf{x}_A, t) = G(\mathbf{x}_A, \mathbf{x}, t)$. Because the wave equation for the Green's function contains a temporal derivative in the source term, the source is defined as a volume injection rate source.

We also consider the homogeneous Green's function $G_\mathrm{h} = G_\mathrm{h}(\mathbf{x}, \mathbf{x}_A, t)$, which is defined as

$$G_\mathrm{h}(\mathbf{x}, \mathbf{x}_A, t) = G(\mathbf{x}, \mathbf{x}_A, t) + G(\mathbf{x}, \mathbf{x}_A, -t), \tag{5.2}$$

where $G(\mathbf{x}, \mathbf{x}_A, -t)$ is the time-reversed Green's function, which is acausal; i.e. $G(\mathbf{x}, \mathbf{x}_A, -t) = 0$ for $t > 0$, hence, it propagates towards the source. By combining

Equations (5.1) and (5.2), we obtain the acoustic homogeneous wave equation

$$\partial_i(\rho^{-1}\partial_i G_{\mathrm{h}}) - \kappa\partial_t^2 G_{\mathrm{h}} = 0, \tag{5.3}$$

where the right hand side vanishes, because the source term on the right hand side of Equation (5.1) contains a temporal derivative, hence, the wave equation for the time reversal of the Green's function causes the source term to obtain the opposite sign.

In this chapter, we will make use of the frequency domain versions of the Green's function and other quantities. A time-dependent function $u(\mathbf{x},t)$ is related to the frequency-dependent function $u(\mathbf{x},\omega)$ by

$$u(\mathbf{x},\omega) = \int_{-\infty}^{\infty} u(\mathbf{x},t)e^{i\omega t}\mathrm{d}t, \tag{5.4a}$$

$$u(\mathbf{x},t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(\mathbf{x},\omega)e^{-i\omega t}\mathrm{d}\omega, \tag{5.4b}$$

where ω is the angular frequency in rad s^{-1} and i is the imaginary unit. Using Equation (5.4a), the Green's function can be transformed to the frequency domain, for the sake of efficiently performing certain operations, such as convolution. The data that are considered in this chapter are band-limited and therefore we define a pressure wavefield $p(\mathbf{x},\mathbf{x}_A,t)$, which is related to the Green's function by

$$p(\mathbf{x},\mathbf{x}_A,t) = \int_{-\infty}^{\infty} G(\mathbf{x},\mathbf{x}_A,t-t')s(t')\mathrm{d}t', \tag{5.5a}$$

$$p(\mathbf{x},\mathbf{x}_A,\omega) = G(\mathbf{x},\mathbf{x}_A,\omega)s(\omega), \tag{5.5b}$$

where $s(t)$ and $s(\omega)$ are the time domain and frequency domain versions of a source signal. We also define a homogeneous pressure wavefield, similar to Equation (5.2),

$$p_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,t) = \int_{-\infty}^{\infty} G_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,t-t')s(t')\mathrm{d}t', \tag{5.6a}$$

$$\begin{aligned} p_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,\omega) &= G_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,\omega)s(\omega) = \\ \{G(\mathbf{x},\mathbf{x}_A,\omega) + G^*(\mathbf{x},\mathbf{x}_A,\omega)\}s(\omega) &= 2\Re\{G(\mathbf{x},\mathbf{x}_A,\omega)\}s(\omega), \end{aligned} \tag{5.6b}$$

where $\Re$ indicates the real part of a complex function and * indicates complex conjugation. Note that in Equation (5.6), we have defined that homogeneous wavefield as the convolution of the source wavelet with the homogeneous Green's function; i.e. the Green's function is superposed with its time-reversal before the convolution. If the Green's function is convolved with a wavelet before the superposition is applied, the time-reversal will affect the source wavelet as well. Only if $s(t) = s(-t)$ and hence $s(\omega) = s^*(\omega)$ can the convolution be applied before the superposition. In other words,

$$\{G(\mathbf{x},\mathbf{x}_A,\omega) + G^*(\mathbf{x},\mathbf{x}_A,\omega)\}s(\omega) = G(\mathbf{x},\mathbf{x}_A,\omega)s(\omega) + G^*(\mathbf{x},\mathbf{x}_A,\omega)s^*(\omega), \tag{5.7a}$$

$$p_{\mathrm{h}}(\mathbf{x},\mathbf{x}_A,\omega) = p(\mathbf{x},\mathbf{x}_A,\omega) + p^*(\mathbf{x},\mathbf{x}_A,\omega), \tag{5.7b}$$

only holds if the source spectrum $s(\omega)$ is purely real-valued.

5.2.2. HOMOGENEOUS GREEN'S FUNCTION RETRIEVAL

Homogeneous Green's function retrieval has been employed in the past to obtain the response between two locations inside a medium. The classical representation states that the response between a source and receiver inside a lossless medium can be obtained if observations are available on a closed boundary around the medium [15–17] and can be written as

$$G_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega) = \oint_{\partial\mathbb{D}} \frac{-1}{i\omega\rho(\mathbf{x})} \Big\{ \partial_i G^*(\mathbf{x}_A,\mathbf{x},\omega) G(\mathbf{x},\mathbf{x}_B,\omega) - G^*(\mathbf{x}_A,\mathbf{x},\omega)\partial_i G(\mathbf{x},\mathbf{x}_B,\omega) \Big\} n_i \mathrm{d}^2\mathbf{x}, \tag{5.8}$$

where n_i is the ith component of the outward pointing normal vector on $\partial\mathbb{D}$. In Equation (5.8), $G(\mathbf{x},\mathbf{x}_B,\omega)$ describes the response to a source at $\mathbf{x}_B$, inside the medium in $\mathbb{D}$, at location $\mathbf{x}$ on a boundary $\partial\mathbb{D}$, which encloses the medium. $G^*(\mathbf{x}_A,\mathbf{x},\omega)$ back propagates this response from location $\mathbf{x}$ at the boundary to receiver location $\mathbf{x}_A$ inside $\mathbb{D}$. This creates the response $G_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega)$, with a source at location $\mathbf{x}_B$ and a receiver at location $\mathbf{x}_A$. The main practical disadvantage of this approach is that a closed boundary around the medium is required, which is usually not feasible for seismological applications. More realistically, the boundary will be open and situated on a single side of the medium, which is often the surface of the Earth. In this case, the representation is approximated as

$$G_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega) \approx \int_{\partial\mathbb{D}_0} \frac{2}{i\omega\rho_0} \Big\{ \partial_3 G^*(\mathbf{x}_A,\mathbf{x},\omega) G(\mathbf{x},\mathbf{x}_B,\omega) \Big\} \mathrm{d}^2\mathbf{x}, \tag{5.9}$$

where ρ_0 is the density at a horizontal single open boundary $\partial\mathbb{D}_0$ and we used $\mathbf{n} = (0,0,-1)$. Note that we assume that the medium above $\partial\mathbb{D}_0$ is homogeneous. Applying the representation in this way introduces significant artifacts in the homogeneous Green's function [81].

In more recent years, the homogeneous Green's function representation has been adjusted to take into account the single-sided open boundary [25]. The scheme that is used in this chapter is taken from Equations (10) and (11) from Brackenhoff *et al.* [62],

$$\begin{aligned} &G(\mathbf{x}_A,\mathbf{x}_B,\omega) + \chi(\mathbf{x}_B) 2i\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\} = \\ &\int_{\partial\mathbb{D}_0} \frac{2}{i\omega\rho_0} G(\mathbf{x},\mathbf{x}_B,\omega)\partial_3 \Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega) - \{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^* \Big) \mathrm{d}^2\mathbf{x}, \end{aligned} \tag{5.10}$$

where $f_1(\mathbf{x}_B,\mathbf{x}_A,\omega) = f_1^+(\mathbf{x}_B,\mathbf{x}_A,\omega) + f_1^-(\mathbf{x}_B,\mathbf{x}_A,\omega)$, $\Im$ denotes the imaginary part of a complex function and $\chi(\mathbf{x}_B)$ is a characteristic function that is defined as

$$\chi(\mathbf{x}_B) = \begin{cases} 1, & \text{for } \mathbf{x}_B \text{ in } \mathbb{D}, \\ \frac{1}{2}, & \text{for } \mathbf{x}_B \text{ on } \partial\mathbb{D} = \partial\mathbb{D}_0 \cup \partial\mathbb{D}_A, \\ 0, & \text{for } \mathbf{x}_B \text{ outside } \mathbb{D} \cup \partial\mathbb{D}, \end{cases} \tag{5.11}$$

where $\partial\mathbb{D}_A$ is a horizontal open boundary inside the subsurface of the Earth at the same depth as $\mathbf{x}_A$. The medium in $\mathbb{D}$ is assumed to be lossless and evanescent waves are ignored. Note, that in Equation (5.10), we retrieve the causal Green's function instead

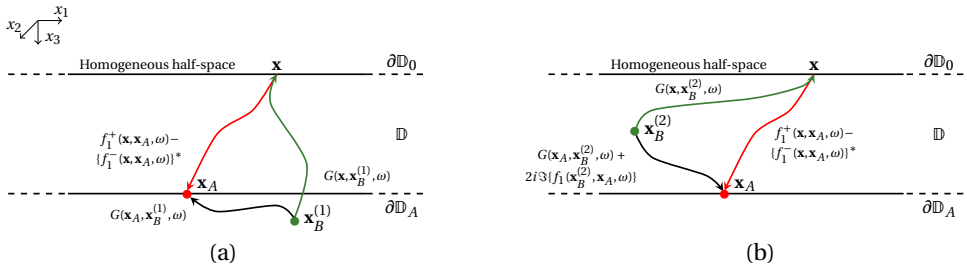

Figure 5.1: Schematic representation of the single-sided Green's function retrieval scheme from Equations (5.10) and (5.11). (a) Retrieval of $G(\mathbf{x}_A,\mathbf{x}_B,\omega)$ when the receiver location is located above the source and no artifacts are present and (b) retrieval of $G(\mathbf{x}_A,\mathbf{x}_B,\omega)$ when the virtual receiver location is located below the virtual source and artifacts are present in the form of $2i\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\}$. The figure is adapted from Brackenhoff *et al.* [62].

of the homogeneous Green's function. In this representation, no time-reversed Green's function is employed, but rather the decomposed focusing functions $f_1^+(\mathbf{x},\mathbf{x}_A,\omega)$ and $f_1^-(\mathbf{x},\mathbf{x}_A,\omega)$ are used, where the superscripts + and − indicate a downgoing and upgoing wavefield, respectively. These focusing functions are designed to focus from a single-sided open boundary $\partial\mathbb{D}_0$ to a location $\mathbf{x}_A$ inside the subsurface of the Earth, generally referred to as the focal location, without artifacts caused by multiple scattering in the overburden. The downgoing focusing function is defined as the inverse of the transmission response of a medium that is truncated below the focal location [26, 30]. In Equation (5.10), the focusing functions $f_1^+(\mathbf{x},\mathbf{x}_A,\omega)$ and $f_1^-(\mathbf{x},\mathbf{x}_A,\omega)$ operate in a similar way as the time-reversed Green's function $G^*(\mathbf{x}_A,\mathbf{x},\omega)$ does in Equation (5.9), backpropagating the response from the boundary $\partial\mathbb{D}_0$ to location $\mathbf{x}_A$. The main difference is that unlike Equation (5.9), Equation (5.10) is specifically designed for application to the open boundary.

The representation in Equation (5.10) does have an issue on the left hand side of the equation in the form of the term $\chi(\mathbf{x}_B)2i\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\}$. Depending on the relative locations of the receiver $\mathbf{x}_A$ and the source $\mathbf{x}_B$, as formulated by Equation (5.11), artifacts related to the focusing function between the two locations are introduced in the obtained Green's function. This is schematically shown in Figure 5.1, where in (a) the receiver is located above the source and the Green's function is retrieved without artifacts. When the virtual source is located at the same depth level or above the virtual receiver, artifacts are present in the retrieved Green's function. By combining Equations (5.10) and (5.2), we obtain the single-sided retrieval scheme for the homogeneous Green's function [24, Equation (33)]:

$$G_\mathrm{h}(\mathbf{x}_A,\mathbf{x}_B,\omega) = 4\Re\int_{\partial\mathbb{D}_0}\frac{1}{i\omega\rho_0}G(\mathbf{x},\mathbf{x}_B,\omega)\partial_3\Big(f_1^+(\mathbf{x},\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*\Big)\mathrm{d}^2\mathbf{x}. \tag{5.12}$$

Equation (5.12) expresses the retrieval of the homogeneous Green's function between two locations in the subsurface using a single-sided boundary, without any artifacts from the focusing function.

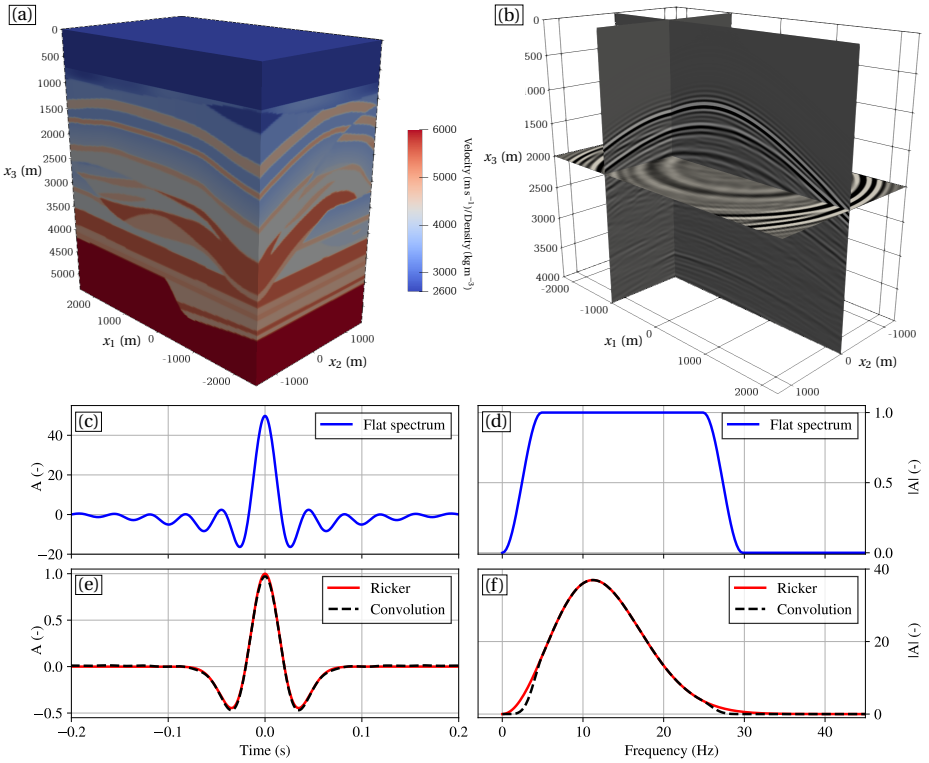

Figure 5.2: (a) Velocity and density model of the subsurface, based on the Overthrust model by Aminzadeh *et al.* [87]. (b) Common-source record of a source, located at (0,0,0)m, recorded at the surface of the model in (a). Wavelet with a flat spectrum in (c) the time domain and (d) the frequency domain. Ricker wavelets in (e) the time domain and (f) the frequency domain. The common-source record in (b) is modeled using the blue wavelet in (c) and (d). The black dashed wavelet in (e) and (f) is the result of the temporal convolution of the blue wavelet in (c) and (d) with the red wavelet in (e) and (f).

5.2.3. IMPLEMENTATION OF GREEN'S FUNCTION RETRIEVAL

We will demonstrate the results of the retrieval schemes in Equations (5.9), (5.10) and (5.12) with numerical examples. In order to obtain the required Green's functions and focusing function, we employ the 3D Marchenko method on acoustic reflection data. The Marchenko method that we use is based on the theory by Wapenaar *et al.* [26]. The scheme allows one to retrieve the Green's function and focusing function between receivers at the surface of the Earth and a focal location in the subsurface of the Earth. To obtain these functions, a reflection response without surface related multiples at the surface of the Earth is required, as well as an estimation of the direct arrival from the surface of the Earth to the focal location. Usually, the time-reversed direct arrival of the Green's function from the focal location to the surface is used for this, following Equation (4.4), even though this introduces errors proportional to the transmission losses into the final result [34].

In this chapter, we make use of the code that was developed by Brackenhoff *et al.*

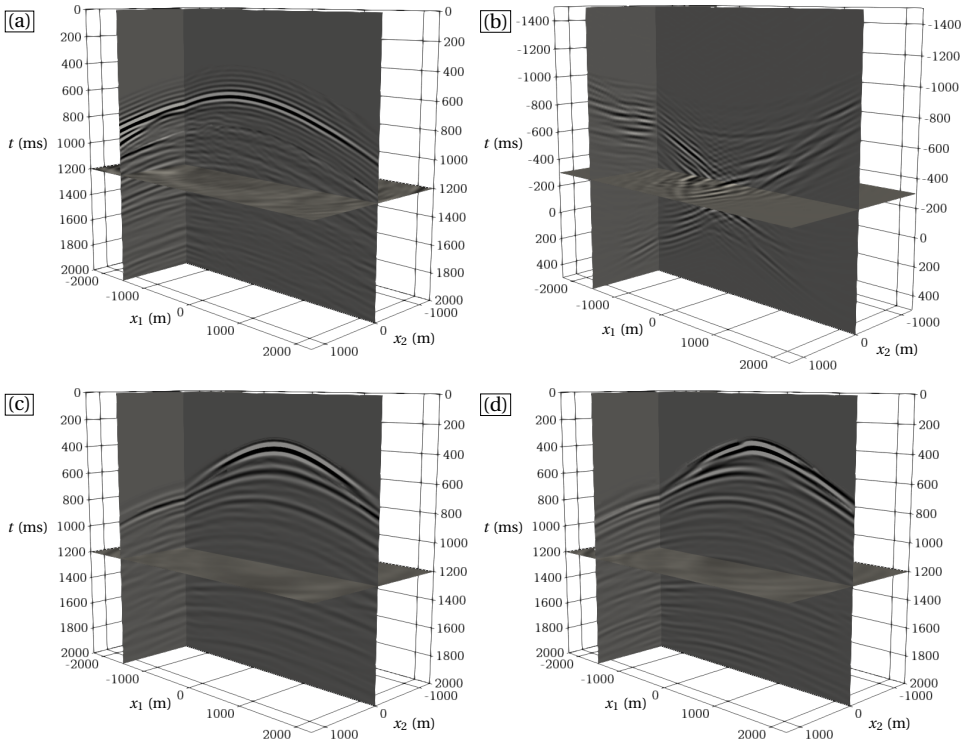

Figure 5.3: Examples of wavefields obtained using the Marchenko method. (a) Green's function $G(\mathbf{x}_A,\mathbf{x}_j,\omega)$ and (b) focusing function $f_1^+(\mathbf{x}_j,\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x}_j,\mathbf{x}_A,\omega)\}^*$ convolved with the flat spectrum wavelet from Figure 5.2(c), with $\mathbf{x}_A = (-350,100,2150)$. Pressure wavefield $p(\mathbf{x}_j,\mathbf{x}_B,\omega)$, i.e., a Green's function convolved with the Ricker wavelet from Figure 5.2(e), with (c) an isotropic source and (d) pressure wavefield $\mathfrak{D}_B^\theta\{p(\mathbf{x}_j,\mathbf{x}_B,\omega)\}$ with a double-couple source, both with $\mathbf{x}_B = (500,-150,1025)$.

[98], which is an opensource 3D implementation of the Marchenko method. To obtain the reflection data, we use the 3D finite-difference modeling code by Thorbecke and Brackenhoff [99] together with a subset of the 3D Overthrust model by Aminzadeh *et al.* [87], which is shown in Figure 5.2(a). To ensure strong reflections, the same model is used for the density and velocity values. To model the data, a fixed-spread acquisition is utilized, where a source is modeled at every receiver location. The source/receiver locations vary from -2250 to 2250m in the inline (x_1) direction, with a spacing of 25m, while the locations in the crossline (x_2) direction vary from -1250 to 1250m, with a spacing of 50m. We define a common-source record as the reflection response to a fixed source, observed by all receivers. The recording length of each common-source record is 4.0s with a temporal sampling of 4ms. An example of a common-source record is shown in Figure 5.2(b). The data are modeled using a wavelet with a flat spectrum, shown as the blue wavelet in Figure 5.2(c) and (d). Examples of Green's functions and a focusing function obtained from these data can be found in Figure 5.3.

Once we obtain the required Green's functions and focusing functions, we use them

in the various retrieval schemes. The schemes so far have been formulated assuming we have the analytical representations instead of the numerical data that we are actually dealing with. In reality, we will not have impulse responses, but rather wavefields with a band-limited signal as defined by Equations (5.5a) and (5.5b). To account for this, we use numerical approximations of the schemes and make use of pressure wavefields with a band-limited source signature. We rewrite Equations (5.9), (5.10) and (5.12) as

$$p_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega) \approx \sum_j^{n_R} \frac{2}{i\omega\rho_0}\Big\{\partial_3 G^*(\mathbf{x}_A,\mathbf{x}_j,\omega)\,p(\mathbf{x}_j,\mathbf{x}_B,\omega)\Big\}\Delta^2\mathbf{x}_j, \tag{5.13}$$

$$\begin{aligned} &p(\mathbf{x}_A,\mathbf{x}_B,\omega)+\chi(\mathbf{x}_B)2is(\omega)\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\} = \\ &\quad\sum_j^{n_R} \frac{2}{i\omega\rho_0}\,p(\mathbf{x}_j,\mathbf{x}_B,\omega)\partial_3\Big(f_1^+(\mathbf{x}_j,\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x}_j,\mathbf{x}_A,\omega)\}^*\Big)\Delta^2\mathbf{x}_j, \end{aligned} \tag{5.14}$$

$$\begin{aligned} &p_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega) = \\ &4\Re\sum_j^{n_R} \frac{1}{i\omega\rho_0}\,p(\mathbf{x}_j,\mathbf{x}_B,\omega)\partial_3\Big(f_1^+(\mathbf{x}_j,\mathbf{x}_A,\omega)-\{f_1^-(\mathbf{x}_j,\mathbf{x}_A,\omega)\}^*\Big)\Delta^2\mathbf{x}_j, \end{aligned} \tag{5.15}$$

where $\mathbf{x}_j$ is the location of the jth receiver at the surface of the Earth, n_R is the amount of receivers and $\Delta^2\mathbf{x}_j$ indicates the receiver sampling distance. While $\Delta^2\mathbf{x}_j$ can be unique for each receiver position, in our fixed spread acquisition the value is the same for all receivers, namely $\Delta^2\mathbf{x}_j = \Delta x_1\Delta x_2 = 25.0\cdot 50.0 = 1250$. Note that in all the numerical representations, we have replaced $G_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega)$, $G(\mathbf{x}_A,\mathbf{x}_B,\omega)$ and $G(\mathbf{x},\mathbf{x}_B,\omega)$ by $p_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega)$, $p(\mathbf{x}_A,\mathbf{x}_B,\omega)$ and $p(\mathbf{x},\mathbf{x}_B,\omega)$, respectively, while some of the other quantities are still denoted by their original symbol. In the application of Equations (5.13)-(5.15), we assume that $p(\mathbf{x},\mathbf{x}_B,\omega)$ is obtained either through the use of the Marchenko method (in the two-step method) or by a direct measurement (in the one-step method), while $G(\mathbf{x}_A,\mathbf{x},\omega)$, $f_1^+(\mathbf{x},\mathbf{x}_A,\omega)$ and $f_1^-(\mathbf{x},\mathbf{x}_A,\omega)$ are always obtained through the use of the Marchenko method. Therefore, we can control the source spectrum of the data that are used to generate the virtual receiver data. We ensure that $G(\mathbf{x}_A,\mathbf{x},\omega)$, $f_1^+(\mathbf{x},\mathbf{x}_A,\omega)$ and $f_1^-(\mathbf{x},\mathbf{x}_A,\omega)$ have a source signature with a flat spectrum of amplitude 1.0 for a certain frequency range, so that the convolution with a unique source signature in that frequency range will produce the response to the latter source signal. This is schematically shown in Figures 5.2(c)-(f). The reflection data in Figure 5.2(b) are convolved with the blue wavelet from Figures 5.2(c) and (d), which has a flat spectrum in a frequency range of 5Hz to 25Hz, and the estimation of the direct arrival is modeled with the same wavelet. The result of the convolution of these wavefields is a wavefield that contains a wavelet very similar to the original flat spectrum wavelet. If a wavefield with this source signature is convolved with a wavefield that contains a different source spectrum, for example, the Ricker wavelet that is shown as the red wavelet in Figures 5.2(e) and (f), the resulting wavefield will contain a wavelet that is almost identical to the original Ricker wavelet, as is shown by the black dashed line in Figures 5.2(e) and (f). There is some slight attenuation of the lowest and highest frequencies, as can be seen in Figure 5.2(f), however, this has little effect on the wavelet shape in Figure 5.2(e). The versions of $p_{\mathrm{h}}(\mathbf{x}_A,\mathbf{x}_B,\omega)$, $p(\mathbf{x}_A,\mathbf{x}_B,\omega)$ and $p(\mathbf{x},\mathbf{x}_B,\omega)$ that are used in Equations (5.13)-(5.15) all include the original

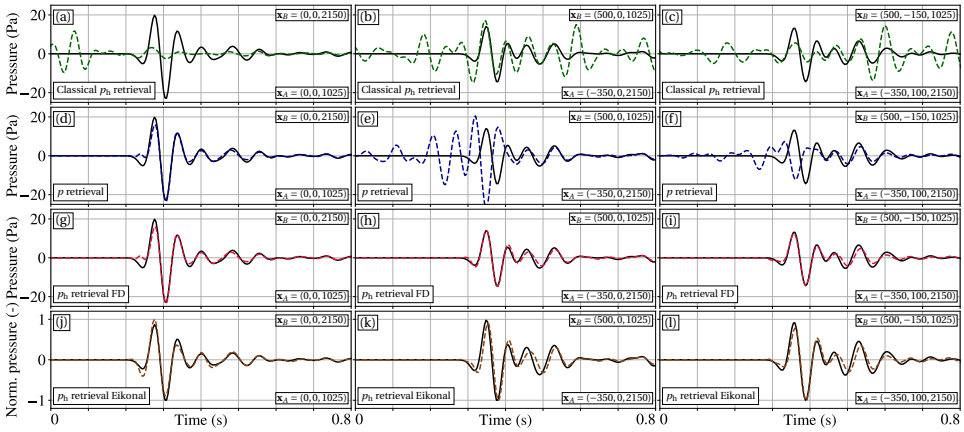

Figure 5.4: Green's funcions of pairs of virtual sources and virtual receivers for different locations and different types of retrieval scheme. The solid lines are the exact (directly modeled) Green's functions and the dotted lines are the retrieved functions. Each column corresponds to a different pair of locations. The first row corresponds to the classical retrieval scheme of Equation (5.13), the second row to the Marchenko retrieval scheme of Equation (5.14) and the third row to the homogeneous Marchenko retrieval scheme of Equation (5.15). For all of these rows, the first arrival required for the Marchenko method is obtained using finite-difference modeling. For the fourth row the same retrieval method is used as in the third row, except the first arrival is obtained using an Eikonal solver, instead of finite-difference modeling. All traces contain the Ricker wavelet from Figure 5.2(e).

Ricker wavelet from Figure 5.2(e); an example of such a pressure wavefield can be found in Figure 5.3(c), with its source at location $\mathbf{x}_B = (500, -150, 1025)$. All other quantities are convolved with the flat spectrum wavelet from Figure 5.2(c), examples of a Green's function and focusing function convolved with such a wavelet are shown in Figure 5.3(a) and (b), respectively, with their source present at location $\mathbf{x}_A = (-350, 100, 2150)$. The application of the band-limitation introduces one more complication, namely that Equation (5.15) is only valid if the source spectrum of $p(\mathbf{x}_A, \mathbf{x}_B, \omega)$ is purely real valued, which holds for the source spectrum of the zero-phase Ricker wavelet.

To demonstrate the validity of our implementation, we show the result of the retrieval schemes in Figure 5.4, using the two-step method. Each column corresponds to a different pair of virtual source and virtual receiver positions, while each row corresponds to a different type of retrieval method. The first column has a virtual receiver located above the virtual source and the positions only differ in depth. In the second column the virtual receiver is located below the virtual source and the locations differ in both the inline direction and depth. For the third column the virtual receiver is located below the virtual source and the locations differ in all three principal directions. The required Green's function and focusing function are obtained using the Marchenko method and a first arrival that was obtained by modeling in the exact medium. We invert the first arrival instead of only time-reversing it, to avoid the transmission losses. While this is not a realistic scenario, as for field data we would not be able to use the exact model, we wish to demonstrate that the method is, at least in theory, capable of obtaining the exact amplitudes. The source has an isotropic radiation pattern. For each panel, the result

that is obtained through the use of a retrieval scheme is plotted in a dashed color, while a directly modeled reference solution is shown in solid black.

The homogeneous wavefield that is obtained using Equation (5.13) is shown in dashed green in the first row of Figure 5.4. Both the Green's function for the virtual source and the virtual receiver were obtained using the Marchenko method. For all location pairs, the results are poor. While the order of magnitude of the retrieved wavefield is similar to that of the direct modeling, the exact amplitudes have a strong mismatch and there are artifacts present at all times. The exceptions are the first arrival in Figure 5.4(b) and the early coda in Figure 5.4(c). Aside from these events however, all other events are wrong and there are still significant artifacts for both examples.

The second row shows in dashed blue the pressure wavefield that is obtained using the open boundary retrieval scheme from Equation (5.14). When the source is located below the virtual receiver, as is the case in Figure 5.4(d), the result shows a good match to the reference solution in both amplitude and arrival time. Because the first arrival is isolated, we apply a muting window to the data before the first arrival to remove numerical artifacts. When the source is located above the receiver, however, the result degrades in quality. There are strong artifacts present in the result at times before the first arrival and the first arrival has the wrong polarity and amplitude. As these artifacts are of a similar magnitude as the first arrival, we cannot apply the muting window. The retrieved coda in these two latter cases is still accurate, with some slight mismatch in the amplitude of the events. This is caused by the different lateral positions of the source and receiver. The aperture and sampling of the data in both the inline and crossline direction are limited, so the exact events become harder to obtain. The overall coda shows a good match to the reference.

To improve the results of the retrieval, the representation from Equation (5.15) is used to retrieve the homogeneous Green's function, shown in dashed red in the third row of Figure 5.4. The results in Figure 5.4(d) and (g) are identical, which corresponds to the condition in Equation (5.11). The improvement is apparent when the source is located above the receiver as is the case in Figure 5.4(h) and (i). Compared to the results in Figure 5.4(e) and (f), the unwanted artifacts are removed and the first arrival is retrieved properly. Here, we once again apply the muting window before the first arrival. The amplitude mismatch in the coda is still present, indicating that this is a limitation caused by the aperture of the recording array and not of the type of retrieval method.

Finally, in the bottom row of Figure 5.4, we apply Equation (5.15) again, however, this time the first arrivals used in the Marchenko method are obtained using an Eikonal solver in a smoothed version of the velocity model. This is to simulate a more realistic situation, where accurate model information would not be available. Because the exact amplitude of the first arrival cannot be obtained in this case, the retrieved homogeneous Green's function is normalized and compared to a normalized version of the reference solution. This is intended to show that even when the exact amplitude cannot be obtained, the relative amplitude can be properly obtained. The matches for all three source-receiver pairs are good, but of a lesser quality than when the finite-difference modeling is employed. Due to the complexity of the model, as well as the smoothing, the Eikonal solver can encounter issues with obtaining the correct arrival times. Furthermore, we only use an estimation of the amplitude distribution along the wavefront,

which also will not properly represent the true effect that the subsurface would have on the amplitude. However, the results still support that use of an Eikonal solver for 3D media can yield useful results.

5.2.4. VISUALIZATION OF THE 3D RESULTS

While the traces in Figure 5.4 demonstrate the validity of our approach, they are limited in scope. To further test our approach in 3D, we obtain the results for not just a single source-receiver pair. Instead we retrieve a large amount of focusing functions and use these in Equation (5.15) to visualize the retrieved Green's functions evolving in time through the 3D medium. To obtain the results, we use the approach employing the Eikonal solver, similarly to how we obtained the results in the bottom row of Figure 5.4. The reason for this is that the computational load and storage space for the use of finite-difference modeling are not feasible for the amount of source-receiver pairs that we desire. The use of the Eikonal solver is a similar approach as was used by Brackenhoff *et al.* [62], however, in this work, we extend its use to 3D. We use the Eikonal solver to obtain focusing functions at locations along three slices through the 3D medium, one with a fixed depth at 2050m with an inline and crossline position from -2250 to 2250m and -1250 to 1250m respectively, one at a fixed inline position of 0m with a depth and crossline position from 400 to 4600m and -1250 to 1250m respectively and one slice at a fixed crossline position of 0m with a depth and inline position from 400 to 4600m and -2250 to 2250m, respectively. For all slices, the sampling in the depth, inline and crossline direction is 25, 25 and 50m, respectively. For the source wavefield $p(\mathbf{x}_j, \mathbf{x}_B, t)$, we obtain a single pressure wavefield due to a source with an isotropic radiation pattern at $\mathbf{x}_B = (0, 0, 2050)$m.

The results of the retrieval using these data and Equation (5.15) are shown in the left column of Figure 5.5. For comparison, we have created a reference homogeneous pressure wavefield by modeling the wavefield directly in the exact medium and superposing it with its time-reversal. We also apply the muting window before the first arrival at all the positions. The four rows each correspond to a different moment in time, namely 0, 300, 600 and 900ms. When comparing the results of the retrieval to the direct modeling, it can be seen that while the match in certain locations is strong, in other locations events appear to be missing. This is due to the finite aperture of the data. The theoretical representations in Equations (5.10) and (5.12) assume that the aperture of the data is infinite. In reality, the aperture is limited, especially in the crossline direction. The events in the homogeneous Green's function are reconstructed from the reflection data, so if an event is not present in the reflection data, it will not be reconstructed properly. The horizontally traveling wavefield, especially near the edges of the aperture, will not arrive at the surface within the range of the aperture. The deeper the target is in the medium, the more severe this problem can become. It should be noted that if the velocity of the medium is increasing with depth, the refraction of the waves will ensure that more angles of the wavefield arrive at the surface of the medium. In the Overthrust model, the propagation velocity of the medium is generally increasing with depth, however, there are some low velocity zones present at greater depths. Because of the general increasing trend, some of the horizontally traveling wavefield at greater depths is still recovered.

The part of the wavefield that is traveling at a smaller angle is reconstructed properly,

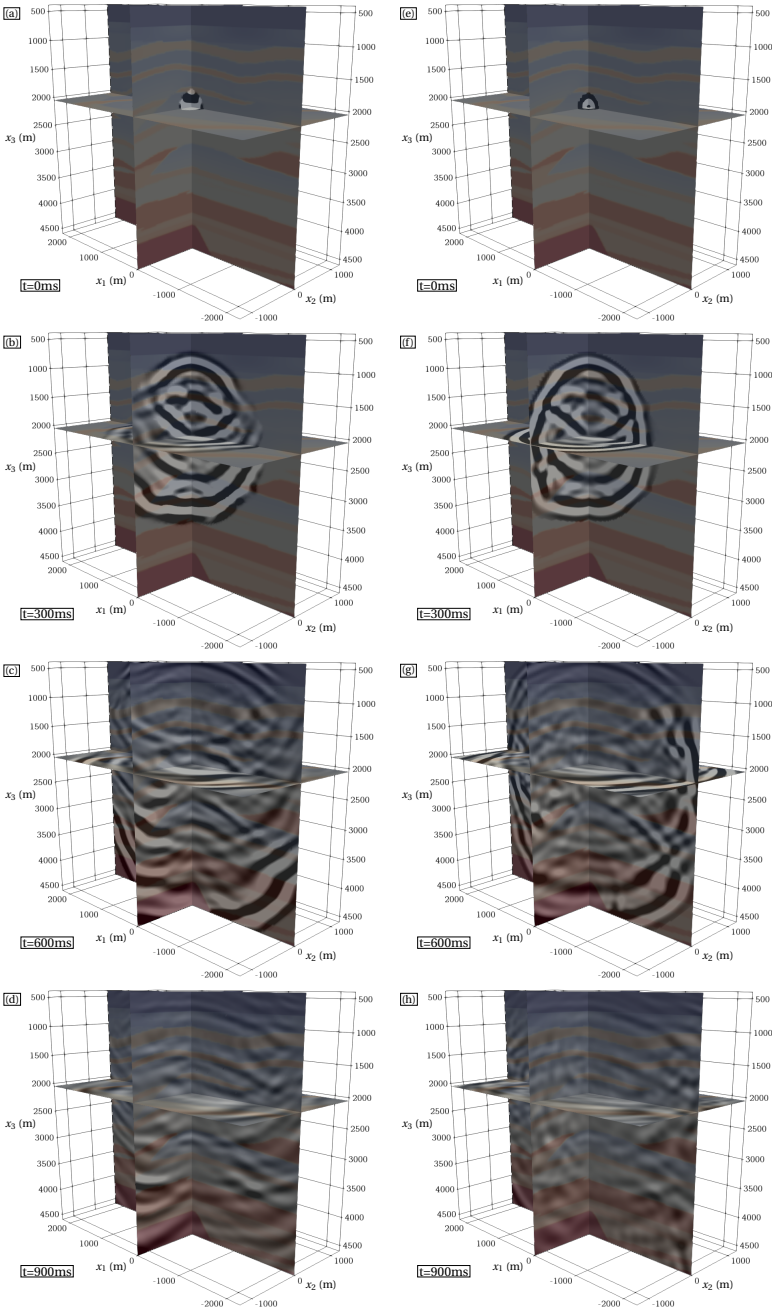

Figure 5.5: 3D snapshots of the homogeneous Green's function retrieval of an isotropic virtual source in the Overthrust model at (a) 0ms, (b) 300ms, (c) 600 ms and (d) 900ms. For comparison, (e), (f), (g) and (h) are snapshots of a directly modeled homogeneous pressure wavefield at 0ms, 300ms, 600ms and 900ms, respectively. The source is located at $\mathbf{x}_B = (0, 0, 2050)$m. The first arrivals for every virtual source-receiver pair were obtained using the Eikonal solver, similar to the results in Figures 5.4(j)-(l). All wavefields contain the Ricker wavelet from Figure 5.2(e) and (f) and contain an overlay of a cross section of the Overthrust model to indicate the locations where we expect scattering to take place.

even at large depths and at the edges of the aperture. The events in the center of the model are reconstructed properly. The amplitudes and arrival times of the events are not correct everywhere, which is caused by the use of a smooth velocity model and the Eikonal solver for the direct arrivals, instead of modeling these in the exact medium. However, the results still show the potential of the Marchenko method for 3D virtual seismology.

5.3. MOMENT TENSOR MONITORING

5.3.1. NON-ISOTROPIC POINT SOURCE

In reality, an event in the subsurface is seldom generated by an isotropic point source. Instead the source wavefield is often caused by faulting, the mechanism of which can be described by a moment tensor [8], which causes the amplitude along the wavefront to vary. The double-couple source mechanism is often used, which is a moment tensor that describes a pure shear fault, by its strike, rake and dip [71]. In previous work, by Brackenhoff *et al.* [81], the double-couple source mechanism was combined with the Marchenko method to obtain the virtual response of a double-couple point source in the subsurface, as well as that of a rupture plane. Here, we wish to demonstrate that similar results can be achieved in 3D. We repeat the examples of the isotropic point source, using Equations (5.13)-(5.15), however, we replace the isotropic source at $\mathbf{x}_B$ in $p(\mathbf{x}_j,\mathbf{x}_B,\omega)$ by a double-couple source generated by a moment tensor. We use an operator $\mathfrak{D}_B^\theta\{\cdot\}$, which transforms the radiation pattern of the source at $\mathbf{x}_B$ from an isotropic radiation pattern to a double-couple radiation pattern. It is defined as

$$\mathfrak{D}_B^\theta\{\cdot\} = (\theta_i^\parallel + \theta_i^\perp)\partial_{i,B}, \tag{5.16}$$

where $\partial_{i,B}$ is a component of the vector containing the partial derivatives acting on the monopole signal originating from source location $\mathbf{x}_B$, which alters the radiation pattern, $\theta_i^\parallel$ is a component of a vector that orients one couple of the signal parallel to the fault plane and $\theta_i^\perp$ is a component of a vector that orients the other couple perpendicular to the fault plane. Because we are dealing with acoustic reflection data, we only model the P-waves of the double-couple source, select the first arrival and use it in the Marchenko method to obtain the desired virtual double-couple response $\mathfrak{D}_B^\theta\{p(\mathbf{x}_j,\mathbf{x}_B,\omega)\}$. An example of such a wavefield can be found in Figure 5.3(d), which has a source at the same position as the pressure wavefield with an isotropic source in (c). This wavefield is then used in Equation (5.13) or (5.15) to obtain $\mathfrak{D}_B^\theta\{p_\mathrm{h}(\mathbf{x}_A,\mathbf{x}_B,\omega)\}$ or in Equation (5.14) to obtain $\mathfrak{D}_B^\theta\{p(\mathbf{x}_A,\mathbf{x}_B,\omega)\}$. Previous research has suggested that the double-couple source is not always a sufficient description of an earthquake source [9]. Our wavefield retrieval method is valid for any type of moment tensor, however, for the sake of simplicity, we stick with the double-couple representation.

For our example, we use a double-couple source with a strike, rake and dip of, 19, 68 and 25 degrees, respectively, and obtain the response between the virtual source-receiver pairs, similar to the examples in Figure 5.4, using the same color scheme as in that figure. The results are shown in Figure 5.6, at the same locations as the source-receiver pairs for the isotropic results. Each row shows a different retrieval method, while each column shows a different source-receiver pair. The first column shows very compa-

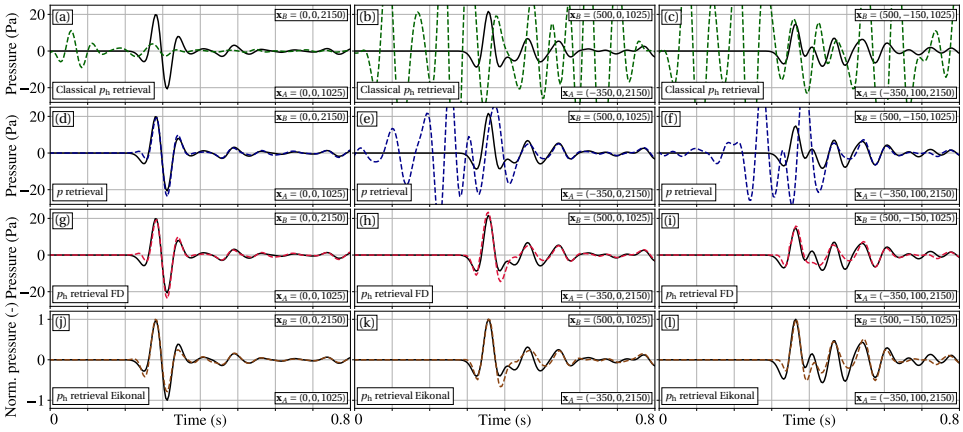

Figure 5.6: As Figure 5.4, but for a virtual double-couple source with a strike, rake and dip of 19, 68 and 25 degrees, respectively.

5

rable results to those of the first column of Figure 5.4, using the same color schemes. The amplitude and shape of the events are different, caused by the different source mechanism. Once again, the classical retrieval using Equation (5.13) shows poor results, while the results of the Green's function retrieval using Equation (5.14) and the homogeneous Green's function retrieval using Equation (5.15) are identical, because the virtual receiver is located above the virtual source, and the match to the direct modeling is good.

If the second and third column are considered, the retrieval using Equation (5.14) shows strong errors around the first arrival, while the coda is retrieved properly. When the homogeneous Green's function retrieval from Equation (5.15) is used, the result improves around the first arrival, as can be seen in the bottom two rows of Figure 5.6. There are still some errors in the early coda. The results in the final column show that parts of the coda are also not properly obtained. Due to the more complex source signature, it becomes harder for the method to resolve all events, especially with the limited aperture. The Eikonal solver once again can only obtain the relative amplitudes, and has similar issues with retrieving the proper events in the coda. Overall the results are encouraging.

To further investigate the effects of using a double-couple source mechanism, we retrieve the wavefield for the same three 3D slices as we did for Figure 5.5. The result is shown in Figure 5.7, where the left column shows the result for the retrieval using the Eikonal solver and Equation (5.15), while a direct modeling is shown in the right column for comparison. Note that the direct modeling contains a source artifact caused by the modeling of only the P-waves. When the results of the retrieval and the direct modeling are compared, most of the nearly vertically traveling events are properly retrieved, not only in arrival time, but also in polarity. For events traveling nearly horizontally, the retrieval is once again poorer. Overall, the result using the double-couple source have a similar quality as the results for the isotropic source, which demonstrates that in 3D, the double-couple source can be successfully integrated into the homogeneous Green's function retrieval.

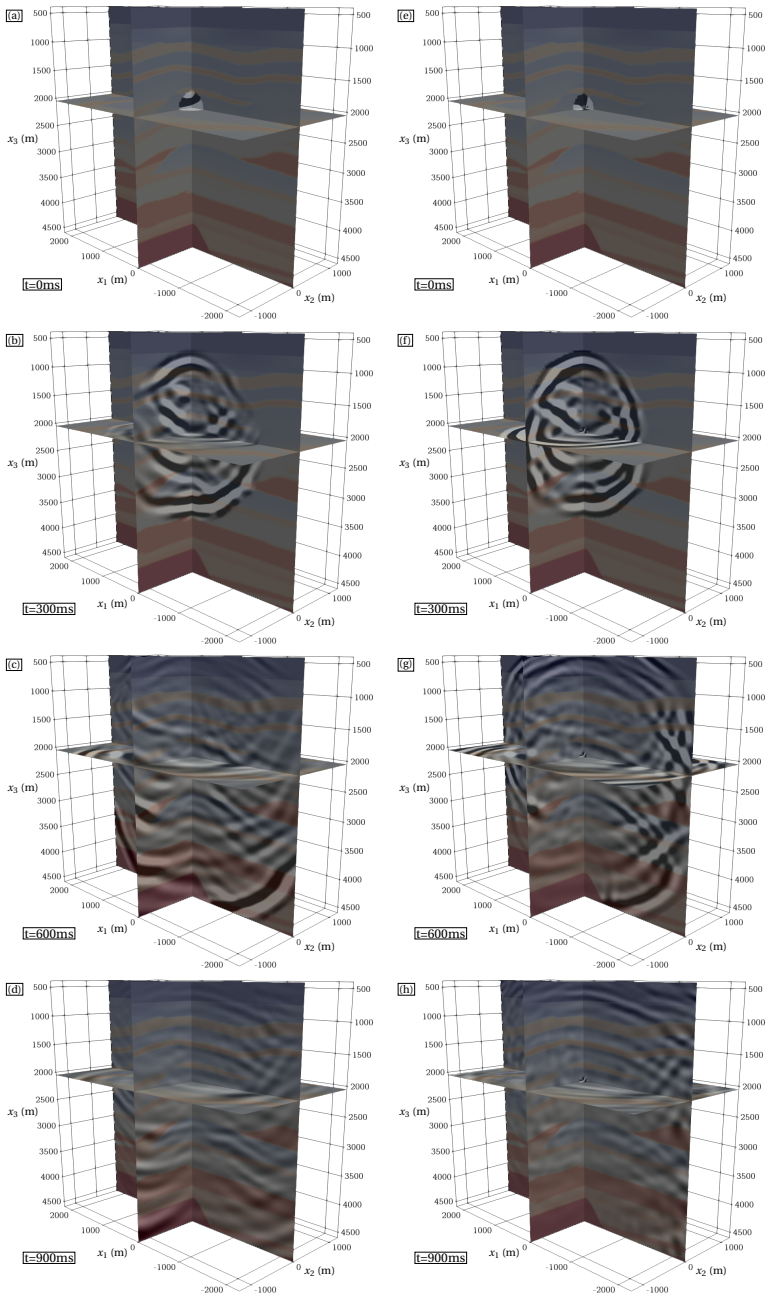

Figure 5.7: As Figure 5.5, but for a virtual double-couple source with a strike, rake and dip of 19, 68 and 25 degrees, respectively. Note that the direct modeling contains a source artifacts caused by only modeling the P-waves.

5.3.2. RUPTURE

In previous sections, we have only considered point sources, however, in the field, an earthquake is seldom a single event, rather, it consists of a cluster of several events that are activated over an area for a period of time [100]. Hence, the total wavefield of an earthquake is not the result of a single instantaneous source, instead, it consists of a superposition of wavefields caused by different sources that are activated at different times. To approximate this kind of wavefield, we define a total wavefield $P(\mathbf{x}_A, t)$ that consists of a superposition of wavefields that are caused by double-couple point sources. The superposition can be expressed as

$$P(\mathbf{x}_A, t) = \sum_k^{n_S} \mathfrak{D}_B^{\theta,(k)}\{p(\mathbf{x}_A, \mathbf{x}_B^{(k)}, t)\} = \sum_k^{n_S} \int_{-\infty}^{\infty} \mathfrak{D}_B^{\theta,(k)}\{G(\mathbf{x}_A, \mathbf{x}_B^{(k)}, t-t')\} s^{(k)}(t')\mathrm{d}t', \tag{5.17a}$$

$$P(\mathbf{x}_A, \omega) = \sum_k^{n_S} \mathfrak{D}_B^{\theta,(k)}\{p(\mathbf{x}_A, \mathbf{x}_B^{(k)}, \omega)\} = \sum_k^{n_S} \mathfrak{D}_B^{\theta,(k)}\{G(\mathbf{x}_A, \mathbf{x}_B^{(k)}, \omega)\} s^{(k)}(\omega), \tag{5.17b}$$

where $\mathbf{x}_B^{(k)}$ indicates the location of the kth source of a total of n_S sources, $\mathfrak{D}_B^{\theta,(k)}\{\cdot\}$ is the double-couple operator for each location and $s^{(k)}(t)$ is the corresponding source signal for each location that contains all the information for the source strength, activation time and duration. Because of the different activation times, the source spectrum of $P(\mathbf{x}_A, t)$ is no longer purely real-valued and can therefore not be used in Equation (5.15). However, using it in Equation (5.14) is still valid, as no time-reversal is applied. We rewrite Equation (5.14) for this purpose as

$$\begin{aligned} &P(\mathbf{x}_A, \omega) + \sum_k^{n_S} \mathfrak{D}_B^{\theta,(k)}\{\chi(\mathbf{x}_B^{(k)}) 2i s^{(k)}(\omega) \Im\{f_1(\mathbf{x}_B^{(k)}, \mathbf{x}_A, \omega)\}\} = \\ &\sum_j^{n_R} \frac{2}{i\omega\rho_0} P(\mathbf{x}_j, \omega) \partial_3 \Big(f_1^+(\mathbf{x}_j, \mathbf{x}_A, \omega) - \{f_1^-(\mathbf{x}_j, \mathbf{x}_A, \omega)\}^* \Big) \Delta^2 \mathbf{x}_j = \\ &\sum_j^{n_R} \frac{2}{i\omega\rho_0} \sum_k^{n_S} \mathfrak{D}_B^{\theta,(k)}\{p(\mathbf{x}_j, \mathbf{x}_B^{(k)}, \omega)\} \partial_3 \Big(f_1^+(\mathbf{x}_j, \mathbf{x}_A, \omega) - \{f_1^-(\mathbf{x}_j, \mathbf{x}_A, \omega)\}^* \Big) \Delta^2 \mathbf{x}_j. \end{aligned} \tag{5.18}$$

In Equation (5.18), we can retrieve $P(\mathbf{x}_A, \omega)$, however, we will also obtain the focusing function artifacts that are related to each source position, below the source depth. As there are multiple sources, that can have different depths, the artifacts related to one source can interfere with the part of the signal that originates from a deeper part of the medium. Consequently, only above the shallowest source depth can we expect to obtain the correct wavefield at all times. For deeper parts of the medium, we expect to retrieve artifacts before and around the first arrival time and the correct coda at later times, similar to the results that were shown in Figures 5.4 and 5.6. Obtaining the wavefield in this way is a one-step process, where we first measure the total wavefield of an actual rupture and use it in combination with the focusing functions, obtained through the Marchenko method, to monitor the subsurface with virtual receivers.

On the other hand, to obtain the response to a virtual rupture, we can use a two-step process to retrieve $P(\mathbf{x}_A, t)$. This has a significant advantage over the one-step process. Instead of measuring the resulting wavefield of the superposed sources, we use the

Marchenko method to retrieve the individual wavefields $\mathfrak{D}_B^{\theta,(k)}\{p_\mathrm{h}(\mathbf{x}_A,\mathbf{x}_B^{(k)},\omega)\}$ related to each source position. In this case we do not measure the total wavefield, but predict it by using the Marchenko method to obtain the source wavefield $\mathfrak{D}_B^{\theta,(k)}\{p(\mathbf{x}_j,\mathbf{x}_B^{(k)},\omega)\}$ before using it in Equation (5.15). Because of this, we can ensure that the source spectrum of $\mathfrak{D}_B^{\theta,(k)}\{p(\mathbf{x}_j,\mathbf{x}_B^{(k)},\omega)\}$ for each individual virtual source is purely real-valued, before we apply the time-reversal. The wavefields that we retrieve in this way are free of the artifacts related to the focusing function and can be combined to form $P(\mathbf{x}_A,t)$:

$$P(\mathbf{x}_A,t)=\sum_k^{n_S} H(t-t^{(k)})\mathfrak{D}_B^{\theta,(k)}\{p_\mathrm{h}(\mathbf{x}_A,\mathbf{x}_B^{(k)},t-t^{(j)})\}, \tag{5.19}$$

where H is the Heaviside function and $t^{(k)}$ is the activation time of the source. In Equation (5.19), we shift the signals in time by $t^{(k)}$ before superposition is applied. Because these wavefields are time-shifted and homogeneous, i.e. they contain time-shifted versions of $\mathfrak{D}_B^{\theta,(k)}\{p(\mathbf{x}_A,\mathbf{x}_B^{(k)},t)\}$ and $\mathfrak{D}_B^{\theta,(k)}\{p(\mathbf{x}_A,\mathbf{x}_B^{(k)},-t)\}$, the acausal part of one wavefield may interfere with the causal part of another wavefield. The Heaviside function is applied to remove all acausal parts of the wavefields to avoid such an issue. While this approach cannot be used for the monitoring of wavefields measured in the field that are caused by sources that are active over a period of time, the approach can be used to forecast the total wavefield of a virtual rupture, given a specific distribution of sources.

To demonstrate the monitoring and forecasting of the total wavefield, we consider a rupture plane in the Overthrust model that consists of a cluster of 61 point sources with a double-couple radiation pattern and that are activated at different points in time. Instead of retrieving wavefields that contain the zero-phase wavelet, like we have done in the previous examples, we retrieve wavefields that contain a unique causal wavelet for each source position, as wavefields in the real subsurface will be causal and not zero-phase. We choose the Berlage wavelet, which is defined as [101]:

$$W(t)=AH(t)t^n e^{-\alpha t}\cos(2\pi \mathrm{f}_0 \mathrm{t}+\phi_0), \tag{5.20}$$

where A is the amplitude of the wavelet. The time exponent n, exponential decay factor α, initial phase angle ϕ_0 and peak frequency f_0 control the shape of the wavelet. To ensure that the wavelet has an amplitude equal to zero at $t=0$, we use an initial phase angle of -90 degrees. For the peak frequency, we use the same peak frequency as we used for the Ricker wavelet in Figure 5.2(f). However, for the amplitude, time exponent and exponential decay factor, we take random values, to simulate a heterogeneous region along the rupture plane. The schematic overview for the rupture simulation can be found in Figure 5.8. The sources are located along a fault in the model, where each source has a strike and rake of 90 and 0 degrees, respectively, and is located at a fixed crossline position of 0m. The dip of the source is dictated by the fault orientation at each source location. Figure 5.8(a) contains the locations of the sources, while Figure 5.8(b) shows the activation time and random amplitude and Figure 5.8(c) shows the random time exponent and exponential decay factor that are used for the Berlage wavelets. The activation time for the sources is linear, with a time delay of 24ms between the activation of subsequent sources, except for the positions where the depth of the source changes. In these cases the time delay is increased to 32ms to account for the increase in step size.

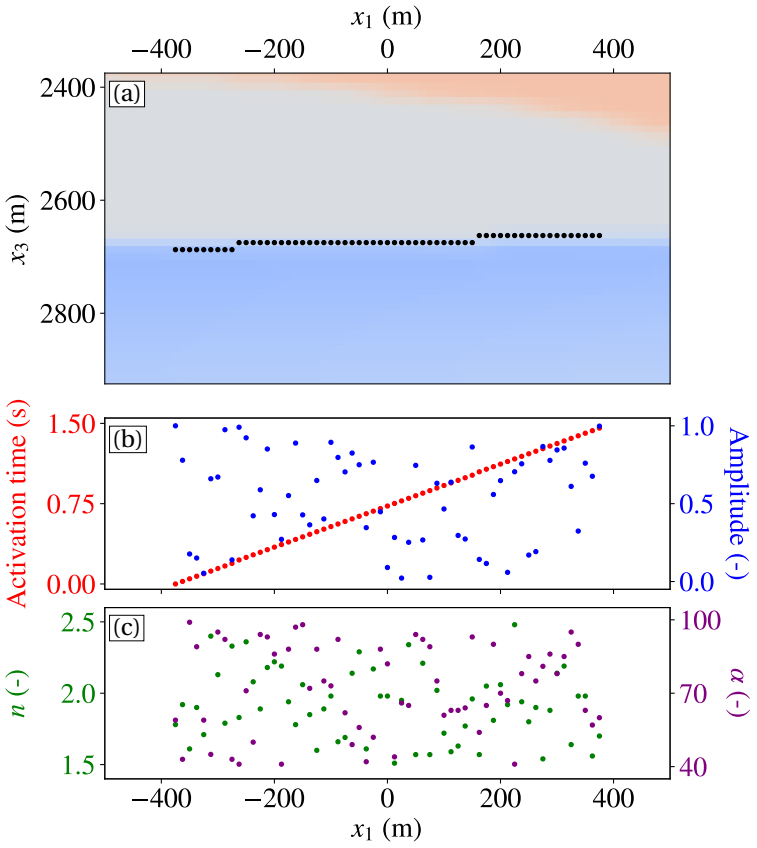

Figure 5.8: (a) Locations of individual double-couple sources with a strike and rake of 90 and 0 degrees, respectively. The dip is oriented along the fault direction for each location. The slice is located along a constant crossline position of 0m. (b) Activation time and amplitude in red and blue, respectively, and (c) time exponent n and exponential decay factor α in green and purple, respectively, for computing the Berlage wavelets using Equation (5.20). The horizontal positions of the sources in (a), (b) and (c) match.

In this way, we simulate a rupture activating and propagating along the rupture plane with a velocity of 520m s^{-1}.

The results of both the one-step and the two-step process can be found in Figure 5.9. The left column of this figure shows the result for the one-step process for monitoring a signal, using Equation (5.18), and the right column shows the result for the two-step process for forecasting a signal, using Equation (5.19). For the monitoring process, we convolve the Berlage wavelets with the source wavefield before we employ the causal Green's function retrieval. For the forecasting process, we use a flat spectrum wavelet to obtain the individual homogeneous Green's functions. Because we are creating everything from the data, this is a valid approach. After we obtain these homogeneous Green's functions, we convolve the functions with the Berlage wavelets, similar to Equation (5.6), to obtain the homogeneous wavefields. These wavefields are then utilized in

Equation (5.19). In both cases we apply the mute window again, for the monitoring this is only for the depths above the shallowest source and for the forecasting we apply it to the individual wavefields before they are superposed.

When comparing the results, it can be seen that, at 640ms, there is a strong difference between the monitoring and forecasting of the signal. Below the depth of the shallowest source location, the wavefield contains strong artifacts, however, above this depth, the wavefields of the two approaches are exactly the same. For later times, around 1280ms, the area below the shallowest source matches more between the two approaches, however, the deeper parts of the medium still shows significant differences. At 1920ms, the match between the two results is even closer, only the deepest parts of the model still contains artifacts for the monitoring approach. We showed that applying homogeneous Green's function retrieval for a single source is accurate, so the superposition of the homogeneous pressure wavefields yields a good result. While the monitoring approach does contain artifacts, we can use the method to monitor the wavefield in the subsurface between the surface and the shallowest source depth accurately. Moreover, we can also use this approach to obtain the coda of the signal for late times at all depths. Overall, the results support the potential of using the single-sided Green's function retrieval in 3D in the field.

5.4. CONCLUSIONS

We have shown that the Marchenko method can be applied to 3D reflection data at the surface of the Earth to obtain the responses for virtual source-receiver pairs in the subsurface in a data-driven way. We did this by considering the 3D single-sided representation for obtaining the homogeneous Green's function in the subsurface. For the input reflection data, we modeled a reflection dataset in a subsection of the 3D Overthrust model. We compared the single-sided approach to the classical representation and showed for a number of selected source-receiver pairs that the single-sided representation can obtain accurate results, whereas the result of the classical representation contains artifacts. The single-sided approach can be applied in two ways, one where the causal Green's function and one where the homogeneous Green's function is retrieved. The retrieval of the causal Green's function is accurate between the surface and the source location. When the virtual receiver is located below the source, artifacts are created that are related to the focusing function from the source to the virtual receiver. These artifacts are present before the coda, and therefore affect the first arrival, however, the coda is obtained accurately. When the homogeneous Green's function retrieval is applied, the result is accurate for all times everywhere in the subsurface, however, this requires the source spectrum to be purely real-valued.

We showed the retrieval of the homogeneous Green's function in a 3D view for a single point source using snapshots and compared these slices to a directly modeled result. The retrieved result showed a good match to the reference for the wavefield that was traveling at moderate angles with the vertical and in the shallow part of the medium. The waves that were traveling close to horizontally were not always retrieved properly, because these angles are not present in the events that are captured by the reflection data. We showed these results not only for an isotropic source, but also for a source that had a double-couple radiation pattern. In both cases, the match to the reference was

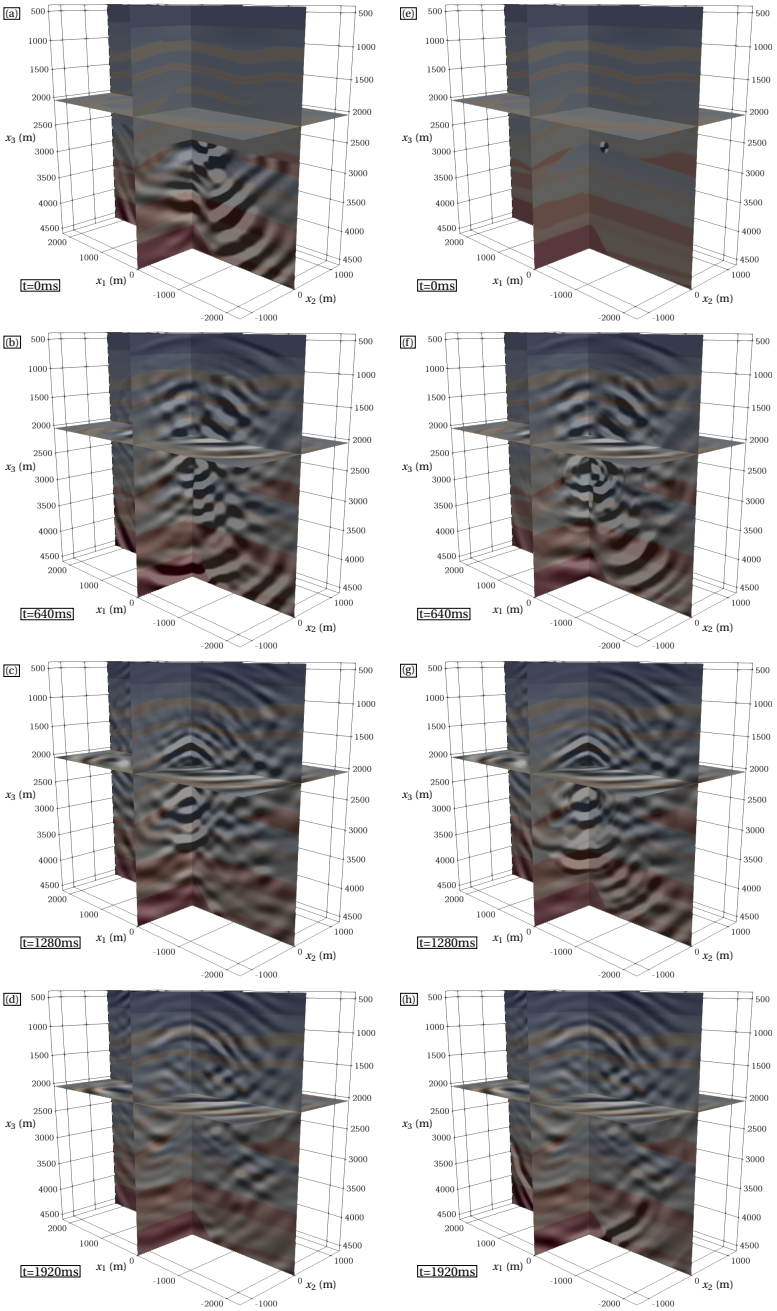

Figure 5.9: 3D snapshots of the Green's function retrieval for a wavefield caused by a rupture in the Overthrust model at (a) 0ms, (b) 640ms, (c) 1280 ms and (d) 1920 ms, using Equation (5.18), and 3D snapshots of superposed and time-shifted wavefields in the Overthrust model, obtained using homogeneous Green's function retrieval using Equation (5.19), at (e) 0ms, (f) 640ms, (g) 1280 ms and (h) 1920 ms. All wavefields have an overlay of a cross section of the Overthrust model to indicate the locations where we expect scattering to take place. Details about the locations, activation times and the wavelets of each source can be found in Figure 5.8.

good.

We also considered a source mechanism that was not a single point source. To this end, we simulated a rupture by combining point sources with a double-couple radiation pattern and source signals that were based on causal Berlage wavelets, with unique parameters for each source location. We considered two different ways of applying the representation. The first was for the purpose of monitoring, where we applied the method to the total wavefield caused by all sources. Because the composite source spectrum was not purely real-valued, we could only apply the causal Green's function retrieval approach, with its inherent limitations. The second method we used was for the purpose of forecasting the total wavefield, given a specific distribution of sources. In this case, we assumed that we did not measure the total wavefield, but simulated it by using the homogeneous Green's function retrieval for each virtual source separately. After the retrieval was applied, the acausal part of the wavefield was muted and the remaining causal part was shifted in time. Finally, the individual responses were summed together to obtain the forecasted total wavefield.

The monitoring approach contained artifacts from all source locations at varying times, although the late coda was properly retrieved. The forecasting approach yielded no such artifacts, however, it is based on a model of sources.

The results on the synthetic data demonstrated that the method has the potential to be applied in the field for monitoring and forecasting the wavefields associated to induced seismicity in a data-driven way. The results are accurate for point sources, for wavefields with an isotropic and double-couple radiation pattern, and for larger ruptures that are active over a period of time.

6

Conclusions and Recommendations

6.1. Conclusions

The primary research goal of this thesis is to determine whether the Marchenko method can be combined with the single-sided representation of the homogeneous Green's function to monitor and forecast seismic wavefields caused by induced seismicity. Both the theoretical and practical aspects have been considered in this thesis.

In chapter 2, the single-sided homogeneous Green's function representation was applied, using the Marchenko method and both 2D numerical reflection data and 2D field reflection data. Using the numerical data, it was shown that the single-sided representation produced a wavefield that closely resembled a directly modeled reference result, something which the classical representation could not achieve, particularly for the downward traveling part of the wavefield. Furthermore, it was shown that aside from a virtual source with an isotropic radiation pattern, a virtual source with a double-couple radiation pattern could also be retrieved using the single-sided representation. When the double-couple radiation pattern was used, it changed the amplitude behavior of the wavefield, but not the arrival times. To determine the sensitivity of the Marchenko method and of the single-sided homogeneous Green's function representation to imperfect reflection data, the numerical reflection data were truncated before the Marchenko method and single-sided representation were applied. The limitations that were applied were related to the offset, sampling and attenuation of the data. The results that were obtained when these truncated reflection data were used demonstrated the need for the pre-processing in case the reflection data are not ideal, as is often the case for field data. The homogeneous Green's function retrieval was performed on the field data after the dataset was pre-processed and the results showed that the single-sided representation was a significant improvement over the classical representation when field data are utilized. The virtual source with a double-couple radiation pattern was also retrieved using the homogeneous Green's function representation without complications.

In chapter 3, two approaches for applying the single-sided representation were considered. The first approach was a one-step process to monitor a wavefield in the subsurface, which employed the single-sided representation of the causal Green's function. The measurement of the response to a physical source was used in combination with virtual receivers created through the use of the Marchenko method. Because the Marchenko method was only used to create the virtual receivers, this approach was a one-step process. The second approach that was considered was a two-step process to forecast the wavefield in the subsurface, which employed the single-sided representation of the homogeneous Green's function. Both the source and receiver in this case were virtual and both were created using the Marchenko method, hence this was a two-step process. The results that were obtained using the single-sided representation of the causal Green's function contain artifacts in the wavefield below the shallowest source location. When the single-sided representation of the homogeneous Green' function was used, these artifacts were removed. However, the single-sided representation of the causal Green's function could be used for any type of source wavelet, while the single-sided representation of the homogeneous Green's function could only be used for symmetric source wavelets. Both of these approaches were applied using 2D synthetic reflection data and the Marchenko method for point sources with a monopole and double-couple radiation pattern and a rupture plane source. The simulation of a wavefield caused by a rupture plane was achieved by superposition of point sources with a double-couple radiation pattern. The wavefields caused by these point sources were time-shifted so that the rupture plane was active over a period of time rather than a single instant. The monitoring of the wavefield caused by the rupture plane contained artifacts that were related to each point source. These errors existed only before and during the first arrivals of the sources and the later coda could be properly retrieved. The forecasting of the wavefield caused by a virtual rupture plane was achieved by using the representation to retrieve the homogeneous Green's function for each source separately, removing the acausal part of the wavefield and applying the time shift. These causal Green's functions were then superposed to obtain the total wavefield caused by the virtual rupture plane.The monitoring and forecasting approaches were applied using 2D numerical reflection data and 2D field reflection data to obtain the total wavefield of a (virtual) rupture plane. The results of these approaches matched each other above the shallowest source location at all times and below this source location after the first arrivals of the sources.

In chapter 4, the opensource implementation of the 3D Marchenko method was discussed. The theoretical extension of the Marchenko method from 2D to 3D was a straightforward process, however, there were difficulties for the practical implementation. For the standard iterative implementation, the reflection data need to be well sampled in both the inline and the crossline direction, meaning that more sources and receivers are required. Consequently, there was a large increase in the size of the data in 3D compared to 2D, hence a large amount of storage space was required. Aside from the storage requirements, the time it took to load the data from the disk increased as did the overall computation time. To reduce the required storage space and loading time, the reflection data were compressed using a floating point algorithm. The 3D implementation was successfully tested on compressed 3D numerical reflection data, that were modeled in a simple layered model and the more complex Overthrust model. For both models,

the Green's function and focusing function were retrieved using a direct arrival that was modeled in the exact medium. The Green's function showed a good match to a directly modeled reference Green's function. To limit the modeling time, an Eikonal solver was used to obtain a large amount of first arrivals so that the Green's function and focusing function could be retrieved in many places in the models, while maintaining a feasible computation time. These data were used to create images of the subsurface and showed that internal multiples in the image were attenuated. These results demonstrated that the implementation worked properly and could be used for other applications beside imaging.

In chapter 5, the application of the single-sided homogeneous Green's function representation in 3D was considered. The representation was tested using the 3D Marchenko implementation that was discussed in chapter 4. The single-sided representation of the causal and the homogeneous Green's function, as well as the classical representation of the homogeneous Green's function, were tested using three pairs of a single virtual source and a single virtual receiver at different locations. The first arrivals that were required for the Marchenko method were obtained by inverting the first arrival of the Green's function that was created by finite-difference modeling in the exact model. The results obtained using the various representations were similar to those obtained using the 2D versions of the representations in chapter 2, with the single-sided homogeneous Green's function representation demonstrating the best result. The single-sided representation for the homogeneous Green's function was shown to be able to not only obtain the correct arrival times of the events, but also the correct amplitude of events, when detailed knowledge of the subsurface is available. Due to the high computation cost of the finite-difference modeling, an Eikonal solver was used to apply the single-sided homogeneous Green's function representation as well, which decreased the quality of the result by a small amount. The Eikonal solver did make it more feasible to apply the representation for a large amount of virtual receivers. The retrieval using the Eikonal solver was further demonstrated by visualizing the propagating wavefield in 3D in time, by retrieving snapshots in large cross-sections of the model. The results of the retrieval showed a good match to a directly modeled reference wavefield, for the vertically traveling events. The part of the wavefield traveling at a larger angle with the normal of the surface were limited in their quality due to the relatively low aperture of the reflection data, however. All these results were shown for sources with an isotropic and double-couple radiation pattern. A rupture plane in 3D was also considered, constructed in a similar way as was done in chapter 3, by superposing the time-shifted wavefields of individual point sources. Unlike chapter 3, however, all sources were modeled using unique causal Berlage wavelets, instead of identical zero-phase wavelets to simulate a more realistic situation. Both the forecasting and monitoring of the total wavefield were tested using the single-sided representation for the homogeneous Green's function and similar results were achieved as were seen when the 2D reflection data were used. The wavefields matched above the shallowest source location for all times and below this source location after all the first arrivals related to the sources had arrived.

Overall, it has been shown that the single-sided homogeneous Green's function representation can be applied for the purpose of monitoring and forecasting seismic wavefields that are caused by induced seismicity on 2D and 3D numerical data and on 2D

field data. The results for the 3D numerical data show potential for applying the developed method on 3D field data as well. The advantages of the single-sided representation are that the virtual receivers can be created at any location in the subsurface and that different source mechanisms can be incorporated in the scheme. The disadvantages are that the demand on the quality of the reflection data is high and that for the monitoring approach, artifacts appear in the final result below the source locations.

6.2. RECOMMENDATIONS

Considering the promise that was shown when the single-sided homogeneous Green's function representation was applied using 3D numerical reflection data, a logical next step would be to apply the method using 3D field reflection data. The method has been successfully applied to 3D field reflection data for the purpose of imaging [96, 97], which demonstrates that the Marchenko method can work on this type of data. The main challenge would be quality of the reflection data, as the recording array needs to be well-sampled in both space and time. As the results of the 2D field reflection data have shown, pre-processing of the 3D field reflection dataset will be of vital importance if the single-sided representation of the homogeneous Green's function representation is to be applied to this field dataset.

All field data that were considered in this thesis were recorded using an active source. This means that while the monitoring of induced seismicity could be simulated, no monitoring of a signal recorded in the field was performed. To prove the viability of the single-sided representation for the purpose of monitoring, a seismic survey should be combined with the measurement of a passive source inside the subsurface. To ensure a greater amount of quality control, a seismic experiment could be performed in either a lab or in the field. In Appendix B, the application of the single-sided homogeneous Green's function representation on a physical lab reflection dataset is shown. The results in this appendix can be validated with greater accuracy because the model is known. Additionally, in a setup like this, an active source can be activated inside the model, to simulate a passive recording, however, the source parameters are available for further quality control. If the lab experiment is successful, the approach should be tested on a field dataset.

When the single-sided representation has been successfully applied to field data to retrieve the wavefield in the subsurface, the retrieved wavefields can be used for other applications aside from just monitoring or forecasting the wavefield. An example would be using the single-sided representation to obtain the homogeneous Green's function in the subsurface at different points in time. These data can then be used for the purpose of time-lapse monitoring to see the change in properties in the subsurface. The evolution of the location and signature of the source can also be used to make predictive models of subsurface seismic activity. As of the writing of this thesis, there are plans to install a network of sensors in the North of the Netherlands, so that the seismic activity in the area can be measured using Distributed Acoustic Sensing (DAS). The network would be densely sampled, hence the measured data will be ideal for use in the single-sided representation.

In case of rupture planes, that are active over a period of time, there are artifacts present below the shallowest source location, related to the focusing function. This fo-

cusing function can be estimated from the data itself by redatuming the focusing function, a process which is shown in appendix A. However, this approach is currently not feasible to apply to field data, as knowledge of the location and activation time of the sources is required. Further development of the single-sided representation of the causal Green's function could yield a way to more effectively remove these artifacts from the data. Doing so would make it possible to monitor the area below the sources in more detail.

Another practical application of the single-sided representation of the homogeneous Green's function would be in the area of moment tensor inversion. This is a technique where the source mechanism of a seismic event is inverted for using the measured data. In this field, more measurements of the subsurface are of great benefit and physical measurements could be substituted by virtual measurements created from the reflection data. One of the most important aspects for moment tensor inversion and monitoring induced seismicity is that the source mechanisms of the seismic events are almost always elastic in nature instead of acoustic. It has been shown that the double-couple source mechanism can be acoustically approximated. However, in order to obtain the true homogeneous Green's function caused by a moment tensor source, an elastic version of the Marchenko method needs to be employed. Previous research by [102] has demonstrated that in layered elastic media, the single-sided homogeneous Green's function representation can be applied. There are still a lot of ongoing challenges in the field of the elastic implementation of the Marchenko method for its fundamental application on field data [103]. If breakthroughs in this field are achieved, the elastic version of the single-sided homogeneous Green's function representation should be tested on synthetic data. If these results are promising, the representation should then also be applied to elastic field reflection data. And if the field data results are of good quality, the representation should be applied for the purpose of enhancing the inversion of the moment tensor.

ACKNOWLEDGEMENTS

While this thesis bears my name, it should be noted that the work described in these pages could not have been achieved without the help of the many wonderful people who have supported me during my Ph.D. project. I owe all of you a debt of gratitude and I thank you from the bottom of my heart.

The first three people that had the most impact on my work are of course **Kees**, **Jan** and **Joost**, the supervisors of my project. This project simply would not have been possible without you.

Kees, I remember the day that you asked me if I was interested in a Ph.D. candidate position. At the time, I was working on my M.Sc. thesis and I was afraid that I had done something wrong, so I was very happy that I could not have been further of the mark. I tried to play it cool at the time, but I was so happy that I could have jumped for joy (which I did, after I had placed a respectable distance between me and your office). I can honestly say that it was a great decision to join the department and I am very proud of what we achieved in all those years. I am grateful that you were always there to answer any questions and that you were always so supportive of my research. Allowing me to switch my focus after my first year to induced seismicity was fantastic, and I don't think either of us regrets it. Thank you for also supporting my other ventures, such as the DOGS, the IBA, teaching, and whatever other mad ideas popped into my head.

Jan, if I tried to calculate all the times you helped me with my work, I probably would get an overflow error. I simply do not understand where you find the time to not only help me, but all other people in the department. All the software that was written for the research in this thesis is based on the things that you taught me. I hope that me teaching you some Python is just starting to make up for this. I will remember all the times we drank tea in your office, while discussing some better ways to implement something, nibbling on some nice pastries that you brought from home. We should probably have run that in parallel too.

Joost, you were a great supervisor during my M.Sc. thesis, seeing as that work was crowned by an award. It is a shame for me that you were so briefly involved with my Ph.D. project, considering how valuable all of your input has been. I respect your choice to look into a new career path however, and I am glad that we still got to see each other on occasion. Without your guidance, I would probably not have made it so far.

The research in this thesis would not have been possible without the funding of the European Union's Horizon 2020 research and innovation programme: **European Research Council** (grant agreement 742703). I also would like to extend my gratitude to **Equinor** for allowing me to use the 2D field data in my research and the **Delphi Consortium** for assisting me with the application to the field data and providing me with a 2D velocity model. I would like to thank **Senai Cimatec** for providing me with an opportunity to do an internship abroad.

I was very lucky to work in one of the nicest departments on the planet Earth, and I encountered so many wonderful people there. Working on Marchenko is quite difficult, however, I had a great group of people surrounding me. **Myrna**, thank you for being my main Marchenko buddy and for all your kind support throughout the years. **Lele**, you are one of the most intelligent and positive thinking human beings I have ever met. **Evert**, thanks for always bringing my attention to a new aspect I had not thought about yet, whether it is Marchenko or life in general. **Giovanni**, sorry that it took so long to get the plane waves running in 3D, but thanks for all the support for Marchenko and your competition in Mario Kart. **Johno**, thanks for taking me down a peg when I was getting too ahead of myself, which happened more frequently than you would think. **Aydin**, I hope our discussions have helped you during your own project too. **Chris**, thank you for all the interesting discussions and for showing me that I actually had it pretty good working with acoustic data instead of elastodynamic data.

I would like to take this time here to apologize to all the people who have been flooded with Marchenko research while they were working on such amazing other work too. **Jan-Willem**, the ultimate ghostbuster, who I never properly apologized to for having so many loud conversations at the desk next to him. Thank you for all the help with the Delphi software. **Max**, thanks for all the discussions about decomposition, the wave equation, how to run the DOGS board and how robot monkeys will eventually conquer the Earth. **Deyan**, thank you for all the good times at the movies, the ridiculously awesome birthday gifts, all the books and all the support you gave me. I will never forget it. **Kees Weemstra**, I have to write your full name as we have more than one Kees W. Sorry for often taking a bit too long to talk in your office with **Deyan** and **Jan**. I am glad we got to work together helping some students too, so I was more than just a nuisance. **Wim**, thanks for modeling that data in 3D and giving me advice on FD modeling. **Quinten, Rémi, Stephan** and **Tim**, thanks for all the good times with board games and for reminding me that Geology is important too. **Menno**, keep training on those eating skills, maybe one day you will be a challenge for me. **Milad**, stay in the real world. **Anne**, I had a great time working together to represent both the Ph.D. and Post-Doc groups in our department. **Guus** and **Joost**, sorry for blowing up the cluster so much and thanks for keeping it running so smoothly. And thanks to all the other people who I met along the way, **Helena, Martha, Santosh, Rahul, Iris, Ranjani, Karlien, Lisanne, Florencia, Faezeh, Shohei, Shotaro, Pieter, Gil, Shahar, Musab, Amin, Jingming, Jianhuan, Aukje, Reuben, Dieter, Hamed, Navid, Samantha, Karim, Baptiste, Runhai, Matteo, Siddarth, Shan, Dong, Junhai, Koen, Niels, Boris, Femke, Sixue, Timo, Carlos, Kevin, Nicolas, Guy, Gerrit, Aayush, Nelson, Apostolos, Dominique, Jens, Alber, Olivier, Diego, Sara, Atsushi** and **Siamak**.

I was lucky to be a part of the DOGS board for many years, and I would like all the people who I worked alongside there, **Myrna, Karlien, Reuben, Florencia, Johno, Billy, Aukje, Matteo, Sixue** and **Ranjani**. We did some great work while we were there and I am very proud of it. I also want to thank the members of the EAGE Local chapter Netherlands, **Diego, Hannes, Evert, Xander, Dorit, Rita** and **Panos**, for working together on creating the chapter.

I am very happy that I got to represent the Ph.D. group of our faculty on several occasions. I want to thank **Karlien** for getting me on board and **Kevin** for taking over from

me. I want to thank everyone who I was on the Ph.D. council with, namely **Ali, Haopeng, Mariska, Jonathan, Silke, Marco, Pengxu, Yan, Emanuele, Maria, Lara, Martijn** and **Bahareh**.

Thanks to the IBA team for the great experience we had in going to Prague and all the fun times we had in that little sweaty office filled with computers. **Myrna, Ella, Bas, Mohammed** and **Gerhard**, it was a great time and I would not have missed it for the world.

My gratitude to **Marlijn**, **Margot**, **Lydia**, **Ralf** and **Myrthe** for all the additional support during my thesis. I hope I showed enough appreciation while I was at the TU.

Outside of the department, I would like to express my gratitude to **Ivan, Marcin** and **Matteo** for their insight in Marchenko and their discussions on the topic, **Eric** for all the help with the data provided by Equinor, **Huub** for teaching me to never split the difference, **Nick** and **Tristan** for the trip in Houston, **Andreas** for the support in writing a research proposal for after my Ph.D and allowing me to continue my journey in academia.

I also would be remiss if I did not mention all the people who welcomed me while I was in Brazil for my internship at Senai Cimatec. Of course, **Diego** should be first as we worked together so much, he showed me around the beautiful city of Salvador and played me on the switch. We also saw some great football matches together and speaking of which, thanks to **Adhvan** for taking me to see Bahia play. It was a great atmosphere in the stadium and it made me feel very welcome. **Joao**, thanks for showing me the Flatfish and watching over me when we went out in the town together. Thanks to **Otávio** and **Victor** for teaching me new programming skills and showing me how Carnaval works in Brazil.

Thanks to all the people who kept me sane outside work. My mother, **Joke**, for all the support that she has shown over the years. I hope you are proud when you read this, this thesis would not have been written if you had not helped me to be the best that I can. My sister, **Stefanie**, I hope you still look up a little to your cute brother and I am so proud of all that you have done over the years. **Ed**, thanks for all the tips on gadgets and programming language. I am glad we still got to go on at least one fieldwork together. I also want to thank all my other family, I would mention you all here, but I don't think I would have enough space. However, I want to thank you for all your love and support over the years. I, of course, also had many good friends who supported me throughout the project, so thanks to **Bart, Rianne, Sander, Sander, Jet, Jp, Suzet, Michael, Robin, Annette, Thomas** and **Richard** and all the other guys and girls who I had good times with over the years. And thanks to everyone at **Scouting Willem de Zwijger** for all the fun and teaching me how to build a toilet.

And finally, thanks to **you**, whoever you are that is reading this thesis. Maybe you already saw your name, in which case, thanks again. Maybe your name was not there and it should have been, in which case, I am very sorry that I was so rude. And maybe you are just reading this thesis and appreciating how many people contributed to this work. Remember, collaboration is key! And having good people around you is what makes you happy in the end. Whoever you are, stay classy.

REFERENCES

[1] A. Rosti, M. Rota, and A. Penna, *Damage classification and derivation of damage probability matrices from L'Aquila (2009) post-earthquake survey data,* Bulletin of Earthquake Engineering **16**, 3687 (2018).

[2] D. Elsworth, C. J. Spiers, and A. R. Niemeijer, *Understanding induced seismicity,* Science **354**, 1380 (2016).

[3] L. Li, J. Tan, D. A. Wood, Z. Zhao, D. Becker, Q. Lyu, B. Shu, and H. Chen, *A review of the current status of induced seismicity monitoring for hydraulic fracturing in unconventional tight oil and gas reservoirs,* Fuel **242**, 195 (2019).

[4] T. S. Eyre, D. W. Eaton, M. Zecevic, D. D'Amico, and D. Kolos, *Microseismicity reveals fault activation before Mw 4.1 hydraulic-fracturing induced earthquake,* Geophysical Journal International **218**, 534 (2019).

[5] J. P. Verdon and J. Budge, *Examining the capability of statistical models to mitigate induced seismicity during hydraulic fracturing of shale gas reservoirs,* Bulletin of the Seismological Society of America **108**, 690 (2018).

[6] H. Clarke, J. P. Verdon, T. Kettlety, A. F. Baird, and J. Kendall, *Real-time imaging, forecasting, and management of human-induced seismicity at Preston New Road, Lancashire, England,* Seismological Research Letters **90**, 1902 (2019).

[7] E. L. Majer, R. Baria, M. Stark, S. Oates, J. Bommer, B. Smith, and H. Asanuma, *Induced seismicity associated with Enhanced Geothermal Systems,* Geothermics **36**, 185 (2007).

[8] K. Aki and P. G. Richards, *Quantitative Seismology,* (2002), W.H. Freeman and Company, San Fransisco, CA.

[9] B. R. Julian, A. D. Miller, and G. Foulger, *Non-double-couple earthquakes 1. Theory,* Reviews of Geophysics **36**, 525 (1998).

[10] S. C. Maxwell, J. Rutledge, R. Jones, and M. Fehler, *Petroleum reservoir characterization using downhole microseismic monitoring,* Geophysics **75**, 75A129 (2010).

[11] K. Chambers, J. Kendall, S. Brandsberg-Dahl, and J. Rueda, *Testing the ability of surface arrays to monitor microseismic activity,* Geophysical Prospecting **58**, 821 (2010).

[12] M. Fink, *Time reversal of ultrasonic fields. I. Basic principles,* IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **39**, 555 (1992).

[13] C. Draeger and M. Fink, *One-channel time reversal of elastic waves in a chaotic 2D-silicon cavity,* Physical Review Letters **79**, 407 (1997).

[14] E. Bossy, K. Daoudi, A.-C. Boccara, M. Tanter, J.-F. Aubry, G. Montaldo, and M. Fink, *Time reversal of photoacoustic waves,* Applied Physics Letters **89**, 184108 (2006).

[15] R. P. Porter, *Diffraction-limited, scalar image formation with holograms of arbitrary shape,* Journal of the Optical Society of America **60**, 1051 (1970).

[16] R. P. Porter and A. J. Devaney, *Holography and the inverse source problem,* Journal of the Optical Society of America **72**, 327 (1982).

[17] M. L. Oristaglio, *An inverse scattering formula that uses all the data,* Inverse Problems **5**, 1097 (1989).

[18] C. Esmersoy and M. Oristaglio, *Reverse-time wave-field extrapolation, imaging, and inversion,* Geophysics **53**, 920 (1988).

[19] C. Lindsey and D. Braun, *Principles of seismic holography for diagnostics of the shallow subphotosphere,* The Astrophysical Journal Supplement Series **155**, 209 (2004).

[20] K. Wapenaar, *Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation,* Physical Review Letters **93**, 254301 (2004).

[21] D.-J. van Manen, J. O. Robertsson, and A. Curtis, *Modeling of wave propagation in inhomogeneous media,* Physical Review Letters **94**, 164301 (2005).

[22] A. Bakulin and R. Calvert, *Virtual shear source: A new method for shear-wave seismic surveys,* in *SEG Technical Program Expanded Abstracts 2005* (Society of Exploration Geophysicists, Houston, Texas, 2005).

[23] A. Curtis, H. Nicolson, D. Halliday, J. Trampert, and B. Baptie, *Virtual seismometers in the subsurface of the Earth from seismic interferometry,* Nature Geoscience **2**, 700 (2009).

[24] K. Wapenaar, J. Brackenhoff, and J. Thorbecke, *Green's theorem in seismic imaging across the scales,* Solid Earth **10**, 517 (2019).

[25] K. Wapenaar, J. Thorbecke, and J. van der Neut, *A single-sided homogeneous Green's function representation for holographic imaging, inverse scattering, time-reversal acoustics and interferometric Green's function retrieval,* Geophysical Journal International **205**, 531 (2016).

[26] K. Wapenaar, J. Thorbecke, J. van Der Neut, F. Broggini, E. Slob, and R. Snieder, *Marchenko imaging,* Geophysics **79**, WA39 (2014).

[27] V. A. Marchenko, *On reconstruction of the potential energy from phases of the scattered waves,* Dokl. Akad. Nauk SSSR **104**, 695 (1955).

[28] J. H. Rose, *'Single-sided' autofocusing of sound in layered materials,* Inverse problems **18**, 1923 (2002).

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

[30] E. Slob, K. Wapenaar, F. Broggini, and R. Snieder, *Seismic reflector imaging using internal multiples with Marchenko-type equations,* Geophysics **79**, S63 (2014).

[31] K. Wapenaar, F. Broggini, E. Slob, and R. Snieder, *Three-dimensional single-sided Marchenko inverse scattering, data-driven focusing, Green's function retrieval, and their mutual relations,* Physical Review Letters **110**, 084301 (2013).

[32] F. Broggini, K. Wapenaar, J. van der Neut, and R. Snieder, *Data-driven Green's function retrieval and application to imaging with multidimensional deconvolution,* Journal of Geophysical Research: Solid Earth **119**, 425 (2014).

[33] J. Behura, K. Wapenaar, and R. Snieder, *Autofocus imaging: Image reconstruction based on inverse scattering theory,* Geophysics **79**, A19 (2014).

[34] J. Thorbecke, E. Slob, J. Brackenhoff, J. van der Neut, and K. Wapenaar, *Implementation of the Marchenko method,* Geophysics **82**, WB29 (2017).

[35] J. van der Neut, J. Thorbecke, K. Wapenaar, and E. Slob, *Inversion of the multidimensional Marchenko equation,* in *77th EAGE Conference and Exhibition 2015* (European Association of Geoscientists & Engineers, 2015).

[36] J. van der Neut, I. Vasconcelos, and K. Wapenaar, *On Green's function retrieval by iterative substitution of the coupled Marchenko equations,* Geophysical Journal International **203**, 792 (2015).

[37] M. Ravasi, I. Vasconcelos, A. Kritski, A. Curtis, C. A. da Costa Filho, and G. A. Meles, *Target-oriented Marchenko imaging of a North Sea field,* Geophysical Journal International **205**, 99 (2016).

[38] M. Staring, R. Pereira, H. Douma, J. van der Neut, and K. Wapenaar, *Source-receiver Marchenko redatuming on field data using an adaptive double-focusing method,* Geophysics **83**, 1 (2018).

[39] C. A. da Costa Filho, G. A. Meles, A. Curtis, M. Ravasi, and A. Kritski, *Imaging strategies using focusing functions with applications to a North Sea field,* Geophysical Journal International **213**, 561 (2017).

[40] X. Jia, A. Guitton, and R. Snieder, *A practical implementation of subsalt marchenko imaging with a gulf of mexico data set,* Geophysics **83**, S409 (2018).

[41] M. Staring and K. Wapenaar, *3D Marchenko internal multiple attenuation on narrow azimuth streamer data of the Santos Basin, Brazil,* Geophysical Prospecting **68**, 1864 (2020).

[42] K. Wapenaar and E. Slob, *On the Marchenko equation for multicomponent single-sided reflection data,* Geophysical Journal International **199**, 1367 (2014).

[43] C. A. da Costa Filho, M. Ravasi, and A. Curtis, *Elastic P- and S-wave autofocus imaging with primaries and internal multiples,* Geophysics **80**, S187 (2015).

[44] C. Reinicke and K. Wapenaar, *Elastodynamic single-sided homogeneous Green's function representation: Theory and numerical examples,* Wave Motion **89**, 245 (2019).

[45] L. Zhang and M. Staring, *Marchenko scheme based internal multiple reflection elimination in acoustic wavefield,* Journal of Applied Geophysics **159**, 429 (2018).

[46] T. Cui, T. S. Becker, D.-J. van Manen, J. E. Rickett, and I. Vasconcelos, *Marchenko redatuming in a dissipative medium: Numerical and experimental implementation,* Phys. Rev. Applied **10**, 044022 (2018).

[47] G. A. Meles, K. Wapenaar, and J. Thorbecke, *Virtual plane-wave imaging via Marchenko redatuming,* Geophysical Journal International **214**, 508 (2018).

[48] L. Zhang, J. Thorbecke, K. Wapenaar, and E. Slob, *Data-driven internal multiple elimination and its consequences for imaging: A comparison of strategies,* Geophysics **84**, S365 (2019).

[49] S. Singh, R. Snieder, J. Behura, J. van der Neut, K. Wapenaar, and E. Slob, *Marchenko imaging: Imaging with primaries, internal multiples, and free-surface multiples,* Geophysics **80**, S165 (2015).

[50] E. Slob and K. Wapenaar, *Theory for Marchenko imaging of marine seismic data with free surface multiple elimination,* in *79th EAGE Conference and Exhibition, 2017* (European Association of Geoscientists & Engineers, 2017).

[51] M. Dukalski, E. Mariani, and K. de Vos, *Handling short-period scattering using augmented Marchenko autofocusing,* Geophysical Journal International **216**, 2129 (2019).

[52] P. Elison, M. S. Dukalski, K. de Vos, D. J. van Manen, and J. O. A. Robertsson, *Data-driven control over short-period internal multiples in media with a horizontally layered overburden,* Geophysical Journal International **221**, 769 (2020).

[53] A. Lomas and A. Curtis, *Marchenko methods in a 3-D world,* Geophysical Journal International **220**, 296 (2020).

[54] Ö. Yilmaz, *Seismic Data Analysis,* (2001), Society of Exploration Geophysicists, Tulsa, OK.

[55] F. Grigoli, S. Cesca, E. Priolo, A. P. Rinaldi, J. F. Clinton, T. A. Stabile, B. Dost, M. G. Fernandez, S. Wiemer, and T. Dahm, *Current challenges in monitoring, discrimination, and management of induced seismicity related to underground industrial activities: A European perspective,* Reviews of Geophysics **55**, 310 (2017).

[56] K. van Thienen-Visser and J. Breunese, *Induced seismicity of the Groningen gas field: History and recent developments,* The Leading Edge **34**, 664 (2015).

[57] M. B. Magnani, M. L. Blanpied, H. R. DeShon, and M. J. Hornbach, *Discriminating between natural versus induced seismicity from long-term deformation history of intraplate faults,* Science Advances **3**, e1701593 (2017).

[58] K. Wapenaar, J. Thorbecke, J. van der Neut, E. Slob, and R. Snieder, *Virtual sources and their responses, part II: data-driven single-sided focusing,* Geophysical Prospecting **65**, 1430 (2017).

[59] J. H. Rose, *"Single-sided" focusing of the time-dependent Schrödinger equation,* Physical Review A **65**, 012707 (2001).

[60] K. Wapenaar, F. Broggini, E. Slob, and R. Snieder, *Three-dimensional single-sided Marchenko inverse scattering, data-driven focusing, Green's function retrieval, and their mutual relations,* Physical Review Letters **110**, 084301 (2013).

[61] K. Wapenaar, J. Brackenhoff, J. Thorbecke, J. van der Neut, E. Slob, and E. Verschuur, *Virtual acoustics in inhomogeneous media with single-sided access,* Scientific Reports **8**, 2497 (2018).

[62] J. Brackenhoff, J. Thorbecke, and K. Wapenaar, *Monitoring induced distributed double-couple sources using Marchenko-based virtual receivers,* Solid Earth **10**, 1301 (2019).

[63] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures On Physics, Vol. I: The New Millennium Edition: Mainly Mechanics, Radiation, And Heat,* (2011), Basic Books, New York, NY.

[64] P. M. Morse and H. Feshbach, *Methods Of Theoretical Physics, Vol. I,* (1953), McGraw-Hill Book Company Inc., New York, NY.

[65] D. J. Verschuur, A. Berkhout, and C. Wapenaar, *Adaptive surface-related multiple elimination,* Geophysics **57**, 1166 (1992).

[66] J. Brackenhoff, *Rescaling of incorrect source strength using Marchenko redatuming,* (2016), M.Sc. thesis, Delft University of Technology, Delft, Zuid-Holland, the Netherlands.

[67] J. W. Thorbecke, K. Wapenaar, and G. Swinnen, *Design of one-way wavefield extrapolation operators, using smooth functions in WLSQ optimization,* Geophysics **69**, 1037 (2004).

[68] J. Thorbecke and D. Draganov, *Finite-difference modeling experiments for seismic interferometry,* Geophysics **76**, H1 (2011).

[69] J. Vidale, *Finite-difference calculation of travel times,* Bulletin of the Seismological Society of America **78**, 2062 (1988).

[70] J. Spetzler and P. Angelov, *Ray perturbation theory for traveltime and amplitude attributes in 3D heterogeneous media with weak anisotropy,* in *75th SEG Annual Meeting* (Society of Exploration Geophysicists, Houston, Texas, 2005).

[71] D. Li, D. Helmberger, R. W. Clayton, and D. Sun, *Global synthetic seismograms using a 2-D finite-difference method,* Geophysical Journal International **197**, 1166 (2014).

[72] A. Berkhout, *Applied Seismic Wave Theory,* (1987), Elsevier, Amsterdam.

[73] G. van Groenestijn and D. Verschuur, *Estimating primaries by sparse inversion and application to near-offset data reconstruction,* Geophysics **74**, A23 (2009).

[74] D. Draganov, R. Ghose, E. Ruigrok, J. Thorbecke, and K. Wapenaar, *Seismic interferometry, intrinsic losses and Q-estimation,* Geophysical Prospecting **58**, 361 (2010).

[75] P. Hatchell and S. Bourne, *Rocks under strain: Strain-induced time-lapse time shifts are observed for depleting reservoirs,* The Leading Edge **24**, 1222 (2005).

[76] J. V. Herwanger and S. A. Horne, *Linking reservoir geomechanics and time-lapse seismics: Predicting anisotropic velocity changes and seismic attributes,* Geophysics **74**, W13 (2009).

[77] J. H. McClellan, L. Eisner, E. Liu, N. Iqbal, A. A. Al-Shuhail, and S. I. Kaka, *Array processing in microseismic monitoring: Detection, enhancement, and localization of induced seismicity,* IEEE Signal Processing Magazine **35**, 99 (2018).

[78] H. Zhang and D. W. Eaton, *Induced seismicity near Fox Creek, Alberta: Interpretation of source mechanisms,* in *Unconventional Resources Technology Conference, Houston, Texas, 23-25 July 2018* (Society of Exploration Geophysicists, American Association of Petroleum Geologists, Society of Petroleum Engineers, 2018) pp. 2577–2584.

[79] L. Eisner, B. Hulsey, P. Duncan, D. Jurick, H. Werner, and W. Keller, *Comparison of surface and borehole locations of induced seismicity,* Geophysical Prospecting **58**, 809 (2010).

[80] K. Wapenaar, J. Fokkema, and R. Snieder, *Retrieving the Green's function in an open system by cross correlation: A comparison of approaches (L),* The Journal of the Acoustical Society of America **118**, 2783 (2005).

[81] J. Brackenhoff, J. Thorbecke, and K. Wapenaar, *Virtual sources and receivers in the real Earth: Considerations for practical applications,* Journal of Geophysical Research: Solid Earth **124**, 11802 (2019).

[82] L. Buijze, P. A. van den Bogert, B. B. Wassing, B. Orlic, and J. ten Veen, *Fault reactivation mechanisms and dynamic rupture modelling of depletion-induced seismic events in a Rotliegend gas reservoir,* Netherlands Journal of Geosciences **96**, S131 (2017).

[83] G. van Groenestijn and D. Verschuur, *Estimation of primaries and near-offset reconstruction by sparse inversion: Marine data applications,* Geophysics **74**, R119 (2009).

[84] M. M. A. Matias, R. da C. Pestana, and J. van der Neut, *Marchenko imaging by unidimensional deconvolution,* Geophysical Prospecting **66**, 1653 (2018).

[85] J. van der Neut, J. Brackenhoff, M. Staring, L. Zhang, S. de Ridder, E. Slob, and K. Wapenaar, *Single- and double-Sided Marchenko imaging conditions in acoustic media,* IEEE Transactions on Computational Imaging **4**, 160 (2018).

[86] M. Ravasi, *Rayleigh-Marchenko redatuming for target-oriented, true-amplitude imaging,* Geophysics **82**, S439 (2017).

[87] F. Aminzadeh, J. Brac, and T. Kunz, *SEG/EAGE 3-D Salt and Overthrust models. SEG/EAGE 3-D modeling series, no. 1: Distribution CD of Salt and Overthrust models,* SEG book series (1997).

[88] P. Lindstrom, *Fixed-rate compressed floating-point arrays,* IEEE Transactions on Visualization and Computer Graphics **20**, 2674 (2014).

[89] K. Wapenaar, J. Thorbecke, J. van der Neut, F. Broggini, E. Slob, and R. Snieder, *Green's function retrieval from reflection data, in absence of a receiver at the virtual source position,* The Journal of the Acoustical Society of America **135**, 2847 (2014).

[90] J. E. Vidale, *Finite-difference calculation of traveltimes in three dimensions,* Geophysics **55**, 521 (1990).

[91] K. M. Keranen and M. Weingarten, *Induced seismicity,* Annual Review of Earth and Planetary Sciences **46**, 149 (2018).

[92] M. Galis, J. P. Ampuero, P. M. Mai, and F. Cappa, *Induced seismicity provides insight into why earthquake ruptures stop,* Science advances **3**, eaap7528 (2017).

[93] C. Willacy, E. van Dedem, S. Minisini, J. Li, J. W. Blokland, I. Das, and A. Droujinine, *Application of full-waveform event location and moment-tensor inversion for Groningen induced seismicity,* The Leading Edge **37**, 92 (2018).

[94] J. Šílenỳ, *Resolution of non-double-couple mechanisms: Simulation of hypocenter mislocation and velocity structure mismodeling,* Bulletin of the Seismological Society of America **99**, 2265 (2009).

[95] A. Lomas and A. Curtis, *An introduction to Marchenko methods for imaging,* Geophysics **84**, F35 (2019).

[96] R. Pereira, M. Ramzy, P. Griscenco, B. Huard, H. Huang, L. Cypriano, and A. Khalil, *Internal multiple attenuation for OBN data with overburden/target separation,* in *SEG annual meeting, 2019* (Society of Exploration Geophysicists, 2019) pp. 4520–4524.

[97] M. Staring and K. Wapenaar, *Interbed demultiple using Marchenko redatuming on 3D field data of the Santos Basin,* in *16th International Congress of the Brazilian Geophysical Society, 2019* (Brazilian Geophysical Society, SBGf, 2019).

[98] J. Brackenhoff, J. Thorbecke, V. Koehne, D. Barrera, and K. Wapenaar, *Implementation of the 3d marchenko method,* arXiv preprint arXiv:2004.00896 (2020).

[99] J. Thorbecke and J. Brackenhoff, *OpenSource: Code for geophysical 3D/2D Finite Difference modelling, Marchenko algorithms, 2D/3D x-w migration and utilities.* (2019), https://doi.org/10.5281/zenodo.3374728.

[100] D. Schorlemmer, S. Wiemer, and M. Wyss, *Variations in earthquake-size distribution across different stress regimes,* Nature **437**, 539 (2005).

[101] D. F. Aldridge, *The Berlage wavelet,* Geophysics **55**, 1508 (1990).

[102] C. Reinicke Urruticoechea and K. Wapenaar, *Elastodynamic single-sided homogeneous Green's function representation: Theory and examples,* in *79th EAGE Conference and Exhibition 2017* (European Association of Geoscientists and Engineers, Paris, France, 2017).

[103] C. Reinicke, *Elastodynamic Marchenko inverse scattering,* (2020), Ph.D. thesis, Delft University of Technology, Delft, Zuid-Holland, the Netherlands.

[104] K. Wapenaar, *Reciprocity and representation theorems for flux-and field-normalised decomposed wave fields,* Advances in Mathematical Physics **2020**, 9540135 (2020).

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

[106] K. Langenberg, M. Berger, T. Kreutter, K. Mayer, and V. Schmitz, *Synthetic aperture focusing technique signal processing,* NDT International **19**, 177 (1986).

[107] E. H. Saenger, G. K. Kocur, R. Jud, and M. Torrilhon, *Application of time reverse modeling on ultrasonic non-destructive testing of concrete,* Applied Mathematical Modelling **35**, 807 (2011).

[108] J. Cai, L. Shi, S. Yuan, and Z. Shao, *High spatial resolution imaging for structural health monitoring based on virtual time reversal,* Smart Materials and Structures **20**, 055018 (2011).

[109] S. Müller, E. Niederleithinger, and T. Bohlen, *Reverse time migration: A seismic imaging technique applied to synthetic ultrasonic data,* International Journal of Geophysics **2012**, 128465 (2012).

[110] R. Zhu, G. Huang, and F. Yuan, *Fast damage imaging using the time-reversal technique in the frequency-wavenumber domain,* Smart Materials and Structures **22**, 075028 (2013).

[111] M. Tanter and M. Fink, *Ultrafast imaging in biomedical ultrasound,* IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **61**, 102 (2014).

[112] O. A. Sapozhnikov, S. A. Tsysar, V. A. Khokhlova, and W. Kreider, *Acoustic holography as a metrological tool for characterizing medical ultrasound sources and fields,* The Journal of the Acoustical Society of America **138**, 1515 (2015).

[113] J. D. Maynard, E. G. Williams, and Y. Lee, *Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH,* The Journal of the Acoustical Society of America **78**, 1395 (1985).

[114] S. F. Wu, *Hybrid near-field acoustic holography,* The Journal of the Acoustical Society of America **115**, 207 (2004).

[115] H. Wu, W. Jiang, and H. Zhang, *A mapping relationship based near-field acoustic holography with spherical fundamental solutions for Helmholtz equation,* Journal of Sound and Vibration **373**, 66 (2016).

[116] D. Gajewski and E. Tessmer, *Reverse modelling for seismic event characterization,* Geophysical Journal International **163**, 276 (2005).

[117] G. A. Meles, K. Löer, M. Ravasi, A. Curtis, and C. A. da Costa Filho, *Internal multiple prediction and removal using Marchenko autofocusing and seismic interferometry,* Geophysics **80**, A7 (2015).

[118] S. Singh and R. Snieder, *Source-receiver Marchenko redatuming: Obtaining virtual receivers and virtual sources in the subsurface,* Geophysics **82**, Q13 (2017).

[119] G. L. Lamb Jr, *Elements Of Soliton Theory,* (1980).

[120] K. Chadan and P. C. Sabatier, *Inverse Problems In Quantum Scattering Theory,* (2012).

[121] A. Derode, E. Larose, M. Tanter, J. de Rosny, A. Tourin, M. Campillo, and M. Fink, *Recovering the Green's function from field-field correlations in an open scattering medium (L),* The Journal of the Acoustical Society of America **113**, 2973 (2003).

[122] J. van der Neut, J. L. Johnson, K. van Wijk, S. Singh, E. Slob, and K. Wapenaar, *A Marchenko equation for acoustic inverse source problems,* The Journal of the Acoustical Society of America **141**, 4332 (2017).

[123] W. S. Harlan, R. T. Langan, and T. Nemeth, *Introduction to the supplement on velocity estimation for depth imaging,* Geophysics **73**, VE1 (2008).

[124] K. Wapenaar, *Single-sided Marchenko focusing of compressional and shear waves,* Physical Review E **90**, 063202 (2014).

[125] C. A. da Costa Filho, M. Ravasi, A. Curtis, and G. A. Meles, *Elastodynamic Green's function retrieval through single-sided Marchenko inverse scattering,* Phys. Rev. E **90**, 063201 (2014).

[126] Y. Alkhimenkov, *Redatuming and quantifying attenuation from reflection data using the Marchenko equation: A novel approach to quantify Q-factor and seismic upscaling,* (2017), M.Sc. thesis, Delft University of Technology, Delft, Zuid-Holland, the Netherlands.

[127] G. Blacquière and A. E. Koek, *3-D seismic experiments with the Delft modeling facility,* in *SEG Technical Program Expanded Abstracts 1997* (Society of Exploration Geophysicists, 1997).

[128] L. V. Wang and J. Yao, *A practical guide to photoacoustic tomography in the life sciences,* Nature methods **13**, 627 (2016).

[129] M. K. A. Singh, M. Jaeger, M. Frenz, and W. Steenbergen, *In vivo demonstration of reflection artifact reduction in photoacoustic imaging using synthetic aperture photoacoustic-guided focused ultrasound (PAFUSion),* Biomedical Optical Express **7**, 2955 (2016).

[130] K. Daoudi, P. van den Berg, O. Rabot, A. Kohl, S. Tisserand, P. Brands, and W. Steenbergen, *Handheld probe integrating laser diode and ultrasound transducer array for ultrasound/photoacoustic dual modality imaging,* Optics Express **22**, 26365 (2014).

A

Removing focusing function related artifacts from Marchenko-based Green's function retrieval in a data-driven way

During the Marchenko-based retrieval of the Green's function with its source and receiver in the subsurface, artifacts are introduced into the final result, related to the redatumed focusing function. These artifacts can be removed by superposing the retrieved Green's function with its time-reversal, however, this approach is not always viable. An alternative way of removing the artifacts is to estimate the redatumed focusing function from the data and use it to remove the artifacts. By using the decomposed reciprocity theorems, a representation is obtained that can accurately estimate the redatumed focusing function from the data. To validate the representation, it is applied to 1D data, which demonstrates that the artifacts can be correctly estimated and removed from the Green's function.

A

A.1. INTRODUCTION

In the field of seismics, the Green's function in the subsurface is often used for target-oriented applications, such as imaging and monitoring, as it describes the response of a medium to an impulsive point source. The Green's function inside the subsurface is often not available, in which case it can be obtained from processing measurements at the surface of the Earth. For example, by employing the Marchenko method, the redatumed Green's function can be obtained from surface reflection data [58]. However, redatuming both the source and the receiver of the Green's function into the medium, thereby creating a virtual source and a virtual receiver, introduces artifacts into the result [62]. These artifacts can be removed by superposing the redatumed Green's function with its time-reversal to obtain the homogeneous Green's function. This approach has been applied on both synthetic and field data [58, 81]. However, this approach cannot always be used, for example, when the source signal is active over a period of time. In this paper, we present an alternative way of removing the artifacts, by redatuming the focusing function from the surface of the Earth in a data-driven way. We derive the methodology and demonstrate the approach on 1D examples.

A.2. THEORY

In this paper, we consider pressure-normalized wavefields that obey the acoustic wave equation. To redatum the Green's function from the surface of the earth into the medium, we use Equation A-26 from Wapenaar *et al.* [58]:

$$G(\mathbf{x}_A,\mathbf{x}_B,\omega)+\chi(\mathbf{x}_B)2i\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\} = \int_{\partial\mathbb{D}_0}\frac{2}{i\omega\rho_0}G(\mathbf{x},\mathbf{x}_B,\omega)\partial_3 f_2(\mathbf{x}_A,\mathbf{x},\omega)\mathrm{d}\mathbf{x}, \tag{A.1}$$

where $\mathbf{x}=(x_1,x_2,x_3)^T$ is the spatial position vector, $G(\mathbf{x}_A,\mathbf{x}_B,\omega)$ is the Green's function in the frequency domain, with its source at $\mathbf{x}_B$ and its receiver at $\mathbf{x}_A$, $f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)$ is the focusing function of the first type that is present at $\mathbf{x}_B$ and focuses downwards to focal location $\mathbf{x}_A$ and $f_2(\mathbf{x}_A,\mathbf{x},\omega)$ is the focusing function of the second type that is present at $\mathbf{x}_A$ and that focuses upwards to focal location $\mathbf{x}$. i is the imaginary unit, $\Im$ denotes the imaginary part of a complex function, ∂_3 is the spatial derivative in the depth direction, ω is the angular frequency and ρ_0 is the density at the top of the medium. $\partial\mathbb{D}_0$ is a transparent boundary at the top of the medium, above which is a reflection-free halfspace. $\chi(\mathbf{x}_B)$ is a characteristic function that is defined as:

$$\chi(\mathbf{x}_B)=\begin{cases}1, & \text{for } x_{3,B}<x_{3,A},\\ \frac{1}{2}, & \text{for } x_{3,B}=x_{3,A},\\ 0, & \text{for } x_{3,B}>x_{3,A}.\end{cases} \tag{A.2}$$

While $G(\mathbf{x}_A,\mathbf{x}_B,\omega)$ is the desired wavefield in Equation (A.1), it is contaminated by $2i\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\}$, depending on the positions of $\mathbf{x}_A$ and $\mathbf{x}_B$.

To obtain $f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)$, we consider the decomposed reciprocity theorems from Wapenaar [104]. We assume that we have two states, A and B. In both of these states there is a medium $\mathbb{D}$ that is bounded by two horizontal flat surfaces, $\partial\mathbb{D}_0$ and $\partial\mathbb{D}_B$. $\partial\mathbb{D}_B$ is located

below $\partial\mathbb{D}_0$ at the depth of $\mathbf{x}_B$. The reciprocity theorems allow one to relate the wavefields of the states if the medium $\mathbb{D}$ is exactly the same for each state. The decomposed reciprocity theorem of the convolution type is given as:

$$\begin{aligned}&\int_{\mathbb{D}}\{p_A^+q_B^- + p_A^-q_B^+ - q_A^+p_B^- - q_A^-p_B^+\}\mathrm{d}\mathbf{x} = \\ &\int_{\partial\mathbb{D}_0}\frac{2}{i\omega\rho}\{(\partial_3 p_A^+)p_B^- + (\partial_3 p_A^-)p_B^+\}\mathrm{d}^2\mathbf{x} - \int_{\partial\mathbb{D}_A}\frac{2}{i\omega\rho}\{(\partial_3 p_A^+)p_B^- + (\partial_3 p_A^-)p_B^+\}\mathrm{d}^2\mathbf{x},\end{aligned} \tag{A.3}$$

and the decomposed reciprocity theorem of the correlation type is given as:

$$\begin{aligned}&\int_{\mathbb{D}}\{p_A^{+,*}q_B^+ + p_A^{-,*}q_B^- + q_A^{+,*}p_B^+ + q_A^{-,*}p_B^-\}\mathrm{d}\mathbf{x} = \\ &\int_{\partial\mathbb{D}_0}\frac{2}{i\omega\rho}\{(\partial_3 p_A^+)^*p_B^+ + (\partial_3 p_A^-)^*p_B^-\}\mathrm{d}^2\mathbf{x} - \int_{\partial\mathbb{D}_A}\frac{2}{i\omega\rho}\{(\partial_3 p_A^+)^*p_B^+ + (\partial_3 p_A^-)^*p_B^-\}\mathrm{d}^2\mathbf{x},\end{aligned} \tag{A.4}$$

where $p^\pm$ is the decomposed acoustic pressure wave field and $q^\pm$ the decomposed source function. The subscripts A and B indicate whether a wavefield or source belongs to state A or B, the superscripts + and - indicate downward or upward propagation, respectively, $\rho = \rho(\mathbf{x})$ is the density of the medium and * indicates complex conjugation. For the reciprocity theorem of the correlation type in Equation (A.4), evanescent waves are ignored.

For both state A and B, we consider the decomposed downgoing focusing function of the first type, f_1^+, and the decomposed upgoing focusing function of the first type, f_1^-, which are related to the full focusing functions as

$$f_1(\mathbf{x},\mathbf{x}_A,\omega) = f_1^+(\mathbf{x},\mathbf{x}_A,\omega) + f_1^-(\mathbf{x},\mathbf{x}_A,\omega), \tag{A.5}$$

$$f_2(\mathbf{x}_A,\mathbf{x},\omega) = f_1^+(\mathbf{x},\mathbf{x}_A,\omega) - \{f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\}^*. \tag{A.6}$$

The focusing function of the first type is defined in a medium that is truncated below the focal location, which can cause problems in the reciprocity theorems. To ensure that in both states the medium $\mathbb{D}$ is the same, the focal locations of the focusing functions generally cannot be above $\partial\mathbb{D}_B$, as the truncation can cause a difference in the medium. The focal location in state B is located at the same depth as $\partial\mathbb{D}_B$, hence the depth of the focal location in state B cannot be greater than the depth of the focal location in state A. Assuming that the $\mathbb{D}$ is the same in both states, the wavefields and sources for the states are equal to the ones shown in Table A.1. There are no source terms present, which results in the volume integrals of the left hand sides of Equation (A.3) and Equation (A.4) becoming zero. Furthermore, Wapenaar *et al.* [26] showed that at the boundary $\partial\mathbb{D}_B$, the following relations exist:

$$\partial_3 f_1^+(\mathbf{x},\mathbf{x}_B,\omega)|_{x_3=x_{3,B}} = \frac{1}{2}i\omega\rho\delta(\mathbf{x}_\mathrm{H} - \mathbf{x}_{\mathrm{H},B}), \tag{A.7}$$

$$\partial_3 f_1^-(\mathbf{x},\mathbf{x}_B,\omega)|_{x_3=x_{3,B}} = 0, \tag{A.8}$$

where δ indicates the Dirac function and the subscript H indicates that only the hori-

	State A		State B	
	$\partial\mathbb{D}_0$	$\partial\mathbb{D}_B$	$\partial\mathbb{D}_0$	$\partial\mathbb{D}_B$
p^+	$f_1^+(\mathbf{x},\mathbf{x}_A,\omega)$	$f_1^+(\mathbf{x},\mathbf{x}_A,\omega)$	$f_1^+(\mathbf{x},\mathbf{x}_B,\omega)$	$f_1^+(\mathbf{x},\mathbf{x}_B,\omega)$
p^-	$f_1^-(\mathbf{x},\mathbf{x}_A,\omega)$	$f_1^-(\mathbf{x},\mathbf{x}_A,\omega)$	$f_1^-(\mathbf{x},\mathbf{x}_B,\omega)$	$f_1^-(\mathbf{x},\mathbf{x}_B,\omega)$
q^+	0	0	0	0
q^-	0	0	0	0

Table A.1: Wavefields and sources of state A and B.

zontal directions are considered. By substituting the results of Table A.1, Equation (A.7) and Equation (A.8) into Equation (A.3), we obtain:

$$f_1^-(\mathbf{x}_B,\mathbf{x}_A,\omega) = \int_{\partial\mathbb{D}_0} \frac{2}{i\omega\rho_0}\Big((\partial_3 f_1^+(\mathbf{x},\mathbf{x}_B,\omega))\, f_1^-(\mathbf{x},\mathbf{x}_A,\omega) + (\partial_3 f_1^-(\mathbf{x},\mathbf{x}_B,\omega))\, f_1^+(\mathbf{x},\mathbf{x}_A,\omega)\Big)\mathrm{d}\mathbf{x}, \tag{A.9}$$

and by substituting the same wavefields and sources into Equation (A.4), we obtain:

$$f_1^+(\mathbf{x}_B,\mathbf{x}_A,\omega) = -\int_{\partial\mathbb{D}_0} \frac{2}{i\omega\rho_0}\Big((\partial_3 f_1^+(\mathbf{x},\mathbf{x}_B,\omega))^*\, f_1^+(\mathbf{x},\mathbf{x}_A,\omega) + (\partial_3 f_1^-(\mathbf{x},\mathbf{x}_B,\omega))^*\, f_1^-(\mathbf{x},\mathbf{x}_A,\omega)\Big)\mathrm{d}\mathbf{x}. \tag{A.10}$$

Equation (A.9) and Equation (A.10) are substituted into Equation (A.5) to obtain the full focusing function, and by making use of Equation (A.6) we obtain:

$$f_1(\mathbf{x}_B,\mathbf{x}_A,\omega) = \int_{\partial\mathbb{D}_0} \frac{2}{i\omega\rho_0}\Big((\partial_3 f_2(\mathbf{x}_B,\mathbf{x},\omega))\, f_1^-(\mathbf{x},\mathbf{x}_A,\omega) - (\partial_3 f_2(\mathbf{x}_B,\mathbf{x},\omega))^*\, f_1^+(\mathbf{x},\mathbf{x}_A,\omega)\Big)\mathrm{d}\mathbf{x}. \tag{A.11}$$

Note that Equation (A.11) is only generally valid in case $x_{3,B} \leq x_{3,A}$. When comparing this requirement to Equation (A.2), we can see that Equation (A.11) is valid for the situation when artifacts are present in the data. Note, that the result in Equation (A.11) can be obtained from the reflection data without the need for a model, using the Marchenko method.

A.3. 1D RESULTS

To demonstrate the validity of our approach, we apply the methodology to a 1D example. Figure A.1 shows the velocity in (a) and the density in (b), which are used to generate 1D reflection data, which are shown in (c). The Marchenko method is used to obtain $G(\mathbf{x},\mathbf{x}_B,\omega)$, $f_2(\mathbf{x}_A,\mathbf{x},\omega)$, $f_2(\mathbf{x}_B,\mathbf{x},\omega)$, $f_1^+(\mathbf{x},\mathbf{x}_A,\omega)$ and $f_1^-(\mathbf{x},\mathbf{x}_A,\omega)$ using this reflection data [26]. Using the data obtained through the use of the Marchenko method, the response between a virtual source and a virtual receiver using Equation (A.1) is created. The virtual source is located at a depth of 1300 meters, we will refer to this location as $\mathbf{x}_B$, and the virtual receiver is located at a depth of 2500 meters, this location will be referred to as $\mathbf{x}_A$. Because $\mathbf{x}_A$ is located below $\mathbf{x}_B$, artifacts in the retrieval of $G(\mathbf{x}_A,\mathbf{x}_B,t)$ are expected. Figure A.2-(b) shows the retrieved response between the virtual source and

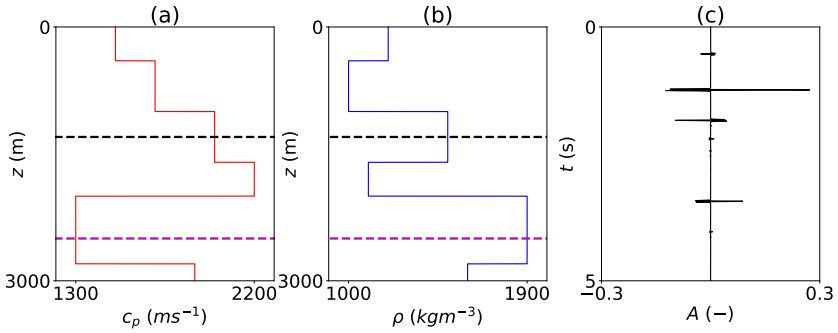

Figure A.1: (a) Velocity in ms^{-1} and (b) density in kgm^{-3} of the 1D medium used to create the 1D reflection data in (c). The dashed black and magenta line indicate the depth of the virtual source at $\mathbf{x}_B$ and the virtual receiver at $\mathbf{x}_A$ used in Figure A.2, respectively.

receiver. Comparing the result to a reference solution, shown in (a), proves that there are artifacts present in the response, as indicated by the red arrows. Furthemore, the first arrival, indicated by the black arrow, has incorrect amplitude and polarity. By using Equation (A.11), the focusing function artifacts are estimated, which are shown in (c). In (d) the result is shown after these artifacts have been subtracted from the response in (b). It can be clearly seen that the artifacts are removed and the first arrival has been retrieved correctly. A comparison in (e) between the response after the artifact removal, and the reference response shows a strong match.

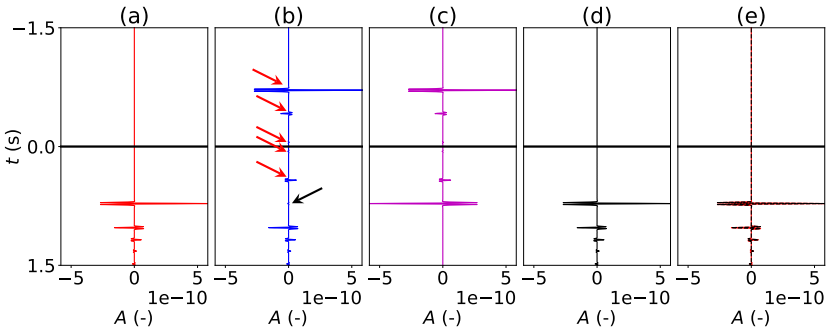

Figure A.2: Responses between a virtual source located at $\mathbf{x}_B = (0,0,1300)^T$ meters and a virtual receiver at $\mathbf{x}_A = (0,0,2500)^T$ meters, indicated by the dashed black and magenta line in Figure A.1, respectively. The solid horizontal black line indicates zero time. (a) Reference solution. (b) $G(\mathbf{x}_A,\mathbf{x}_B,t)$, obtained through use of Equation (A.1). The Green's function is contaminated by artifacts, indicated with the red arrows. Additionally, the first arrival, indicated by the black arrow, has incorrect amplitude and polarity. (c) $2i\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\}$, transformed to the time domain, obtained through use of Equation (A.11). (d) $G(\mathbf{x}_A,\mathbf{x}_B,t)$ obtained by subtracting the result in (c) from the result in (b). (e) Comparison between the result in (a), indicated by the solid red line, and the result in (d), indicated by the dashed black line.

To further demonstrate the method, we consider a scenario with multiple virtual receivers and a single virtual source using the same model and reflection data. The virtual

source is located at the same depth as in the previous example, while the virtual receivers vary over the entire depth of the model, with a spacing of 1 meter. The results are shown in Figure A.3, where the green dot indicates the location of $\mathbf{x}_B$ and the horizontal dashed lines indicate the depths of geological layer contrasts. (a) shows a reference solution, (b) shows the redatumed $G(\mathbf{x}_A,\mathbf{x}_B,t)$ containing the focusing function artifacts. We estimate these artifacts for $x_{3,B} \leq x_{3,A}$, as shown in (c). Subtracting these artifacts from (b) results in the correct redatumed $G(\mathbf{x}_A,\mathbf{x}_B,t)$ in (d).

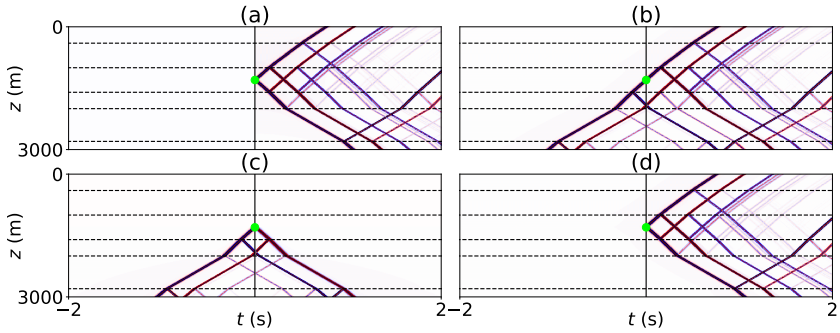

Figure A.3: Responses between a virtual source located at $\mathbf{x}_B = (0,0,1300)^T$ meters, indicated by the green dot, and virtual receiver locations varying in depth over the vertical axis. The horizontal axis denotes the time, the vertical solid black line indicates zero time and the dashed horizontal black lines indicate the depths of geological layer contrasts. (a) Reference solution. (b) $G(\mathbf{x}_A,\mathbf{x}_B,t)$ obtained through use of Equation (A.1). Note the acausal artifacts below the virtual source location and the incorrect downgoing first arrival. (c) $2i\Im\{f_1(\mathbf{x}_B,\mathbf{x}_A,\omega)\}$, transformed to the time domain, obtained through use of Equation (A.11). (d) $G(\mathbf{x}_A,\mathbf{x}_B,t)$ obtained by subtracting the result in (c) from the result in (b).

A.4. CONCLUSION

We have shown how to redatum the focusing function from the surface of the Earth to inside the medium of interest in a data-driven way. This redatumed focusing function can be used to remove the artifacts that occur during the retrieval of the Green's function using the Marchenko method. We have shown the validity of our approach by applying the methodology to two 1D examples, where the artifacts are completely removed. For a single instantaneous source the homogeneous Green's function retrieval is the easier approach. When multiple sources are used, the method in this paper can be used to remove the artifacts for all sources, provided that the time-delays and positions of the original sources are known.

B

VIRTUAL ACOUSTICS IN INHOMOGENEOUS MEDIA WITH SINGLE-SIDED ACCESS

A virtual acoustic source inside a medium can be created by emitting a time-reversed point-source response from the enclosing boundary into the medium. However, in many practical situations the medium can be accessed from one side only. In those cases the time-reversal approach is not exact. Here, we demonstrate the experimental design and use of complex focusing functions to create virtual acoustic sources and virtual receivers inside an inhomogeneous medium with single-sided access. The retrieved virtual acoustic responses between those sources and receivers mimic the complex propagation and multiple scattering paths of waves that would be ignited by physical sources and recorded by physical receivers inside the medium. The possibility to predict complex virtual acoustic responses between any two points inside an inhomogeneous medium, without needing a detailed model of the medium, has large potential for holographic imaging and monitoring of objects with single-sided access, ranging from photoacoustic medical imaging to the monitoring of induced-earthquake waves all the way from the source to the earth's surface.

This appendix was published as K. Wapenaar, J. Brackenhoff, J. Thorbecke, J. van der Neut, E. Slob & E. Verschuur. *Virtual acoustics in inhomogeneous media with single-sided access.* Scientific Reports **8**, 2497 (2018).

B.1. INTRODUCTION

In many acoustic applications, ranging from ultrasonics to seismology, virtual sources can be created by emitting a focusing wave field from the boundary into the medium[12, 16, 17, 105]. Time-reversal mirroring, developed by Fink and co-workers[12, 105], is a well-known approach to create a virtual source. It exploits the fact that the wave equation in a lossless medium is symmetric in time. In many practical situations, like in non-destructive testing[106–110], medical imaging[111, 112], near-field acoustic holography [113–115] or geophysical holography[18, 19, 116], the medium can be accessed from one side only. In those cases the time-reversal approach is not exact, and it breaks down in inhomogeneous media with strong impedance contrasts. Recent work by the authors[25, 30, 31] and others[32, 37, 117, 118] concerns the design of single-sided focusing functions. When emitted from the upper boundary into the medium, these focusing functions yield well-defined foci at predefined positions, which act as omnidirectional virtual sources. This work is inspired by the Marchenko equation of quantum mechanics[27, 119, 120] and its applications in 1D autofocusing[28, 29, 59].

We start this paper with a comparison of the time-reversal method and the single-sided focusing approach, at the hand of a number of numerical examples. Next, we discuss our approach for retrieving virtual sources and receivers from single-sided reflection data. We apply this methodology to ultrasonic physical model data and seismic reflection data. Finally, we discuss potential applications for photoacoustic medical imaging and for monitoring of induced-earthquake waves.

B.2. TIME-REVERSAL VERSUS SINGLE-SIDED FOCUSING

The time-reversal method is illustrated in the first column of Figure B.1, for a lossless layered medium with curved interfaces (denoted by the dashed lines in the grey panels) and different propagation velocities and mass densities in the layers between these interfaces. The top panel shows the time-reversal of the response $V(\mathbf{x},\mathbf{s},t)$ to a point source at $\mathbf{s}$ in the third layer of the medium, as a function of receiver position $\mathbf{x}=(x,z)$ along the boundary and time t. V stands for the normal component of the particle velocity. Only the response at the upper boundary is shown, but the response is available along the entire enclosing boundary $\mathbb{S}$.

The time-reversed response $V(\mathbf{x},\mathbf{s},-t)$ is fed to sources (the red dots) at the original positions of the receivers, which emit the wave field back into the medium. The other panels in column (a) show "snapshots" (i.e., wave fields frozen at constant time) of the wave field propagating through the medium. For negative time ($\cdots -t_2, -t_1 \cdots$), the field follows the same paths as the original field, but in opposite direction. Then, at $t=0$, the field focuses at the position $\mathbf{s}$ of the original source. Because there is no sink to absorb the focused field, the wave field continues its propagation, away from the focal point. Hence, the focal point acts as a virtual source. The snapshots for positive time ($\cdots +t_1$, $+t_2 \cdots$) show the response to this virtual source. The virtual source is omni-directional and radiates a perfect replica of the original field into the inhomogeneous medium.

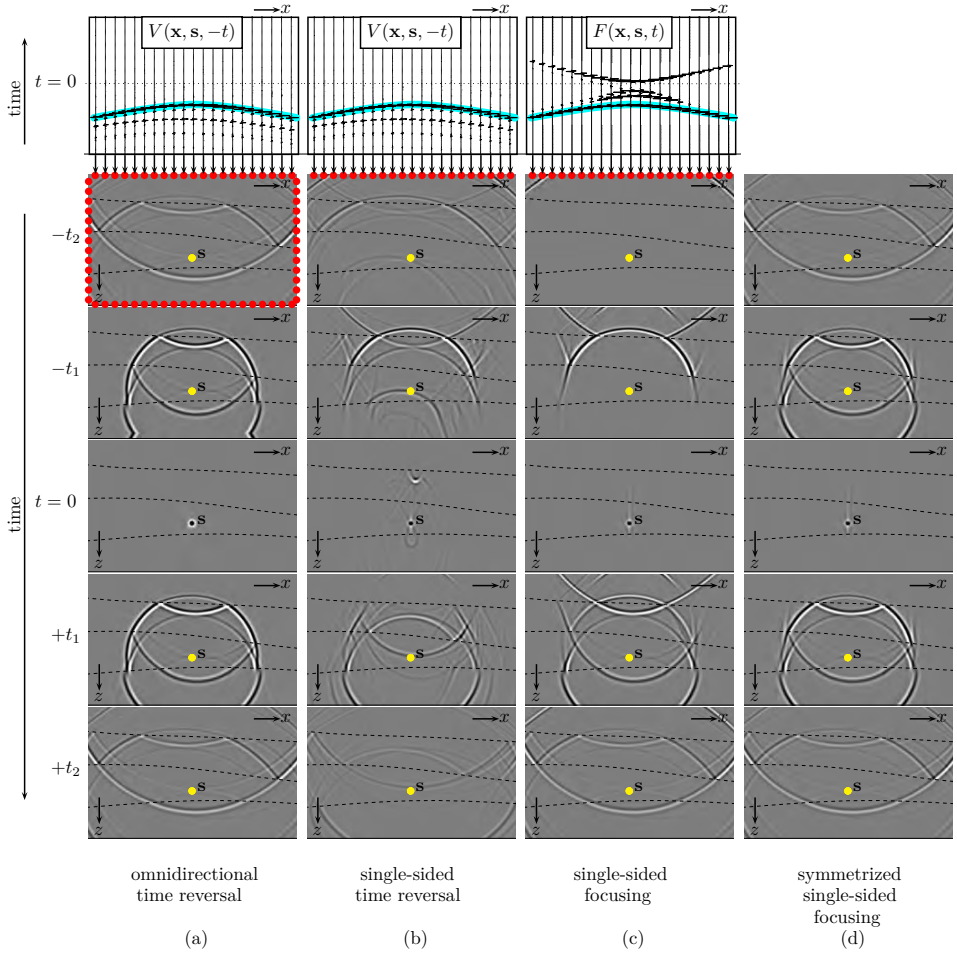

Figure B.1: Illustration of virtual-source methods. (a) A time-reversed point source response is emitted from the enclosing boundary into the inhomogeneous medium. For negative time, it converges towards the focal point, where it focuses at $t = 0$. Subsequently, the focal point acts as an omnidirectional radiating virtual source. (b) Emission of the time-reversed response from the upper boundary only. Ghost foci occur at $t = 0$. The virtual source radiates mainly downward. (c) Emission of a single-sided focusing function from the upper boundary only. No ghost foci occur at $t = 0$. The virtual source radiates mainly downward. (d) Symmetrizing the previous result. No ghost foci occur at $t = 0$. The virtual source is omnidirectional.

Mathematically, time-reversal acoustics is formulated as follows[121]:

$$G(\mathbf{r},\mathbf{s},t)+G(\mathbf{r},\mathbf{s},-t)=2\oint_{\mathbb{S}}\underbrace{G(\mathbf{r},\mathbf{x},t)}_{\text{“propagator”}}*\underbrace{V(\mathbf{x},\mathbf{s},-t)}_{\text{“secondary sources”}}\,\mathrm{d}\mathbf{x} \tag{B.1}$$

(see Supplementary Information). On the right-hand side, the time-reversed field $V(\mathbf{x},\mathbf{s},-t)$ is propagated through the medium by the Green's function $G(\mathbf{r},\mathbf{x},t)$ from the sources at $\mathbf{x}$ on the boundary $\mathbb{S}$ to any receiver position $\mathbf{r}$ inside the medium (the asterisk denotes convolution). The integral is taken along all sources $\mathbf{x}$ on the closed boundary. Note that the right-hand side resembles Huygens' principle, which states that each point of an incident wave field acts as a secondary source, except that here the secondary sources on $\mathbb{S}$ consist of time-reversed measurements rather than an actual incident field. On the left-hand side, the time-reversed Green's function $G(\mathbf{r},\mathbf{s},-t)$ represents the wave field at negative time that converges to the focal point $\mathbf{s}$; the Green's function $G(\mathbf{r},\mathbf{s},t)$ is the response at positive time to the virtual source at $\mathbf{s}$.

Figure B.1(b) shows what happens when the time-reversed response is emitted into the medium by sources (red dots) at the upper boundary only. The field still focuses at $t=0$, but in addition several ghost foci occur at $t=0$. The field at positive time is a virtual-source response, contaminated by artefacts, caused by the ghost foci. Moreover, because the focal point is illuminated mainly from above, the virtual source is far from isotropic: it radiates mainly downward.

We now introduce the single-sided focusing approach, which is designed to overcome the limitations of the time-reversal approach in inhomogeneous media with strong impedance contrasts. The upper panel in Figure B.1(c) shows a 2D focusing function $F(\mathbf{x},\mathbf{s},t)$, for the same focal point $\mathbf{s}$ as in the time-reversal example. Note that the main event (indicated in blue) is the same as that in $V(\mathbf{x},\mathbf{s},-t)$ in the upper panel in Figure B.1(b), but the other events in $F(\mathbf{x},\mathbf{s},t)$ come after the main event (instead of preceding it, like in $V(\mathbf{x},\mathbf{s},-t)$). The snapshots in Figure B.1(c) show the propagation of this focusing function through the medium. Mathematically, the emission of the focusing function $F(\mathbf{x},\mathbf{s},t)$ into the medium by sources at $\mathbf{x}$ at the upper boundary $\mathbb{S}_0$ is described by

$$G(\mathbf{r},\mathbf{s},t)+\text{anti-symmetric artefacts}=\int_{\mathbb{S}_0}G(\mathbf{r},\mathbf{x},t)*F(\mathbf{x},\mathbf{s},t)\mathrm{d}\mathbf{x} \tag{B.2}$$

(see Supplementary Information). The right-hand side resembles again Huygens' principle, this time with the focusing function defining secondary sources on $\mathbb{S}_0$ only. The left-hand side represents the virtual-source response $G(\mathbf{r},\mathbf{s},t)$, contaminated by artefacts that are anti-symmetric in time. Because the anti-symmetric term vanishes at $t=0$, the panel at $t=0$ in Figure B.1(c) shows a "clean" focus. Like in the time-reversal method, the focused field acts as a virtual source. The snapshots at positive time show that this virtual source radiates mainly downward.

Next, we symmetrize both sides of equation (B.2), by adding the time-reversal. This suppresses the anti-symmetric artefacts:

$$G(\mathbf{r},\mathbf{s},t)+G(\mathbf{r},\mathbf{s},-t)=\text{Symmetrize}\Big(\int_{\mathbb{S}_0}G(\mathbf{r},\mathbf{x},t)*F(\mathbf{x},\mathbf{s},t)\mathrm{d}\mathbf{x}\Big) \tag{B.3}$$

(see Supplementary Information). Note that the left-hand side is identical to that in equation (B.1). However, unlike equation (B.1), the right-hand side of equation (B.3) contains an integral along the accessible boundary $\mathbb{S}_0$ only. Symmetrization implies addition of the snapshots at negative times in Figure B.1(c) to those at the corresponding positive times and vice versa, see Figure B.1(d). Note that these superposed snapshots

are nearly identical to those obtained by emitting the time-reversed response into the medium from the entire enclosing boundary (Figure B.1(a)). The remaining artefacts are caused by the finite source aperture and the fact that evanescent waves are neglected in equations (B.2) and (B.3) (see Supplementary Information).

B.3. RETRIEVING VIRTUAL SOURCES AND RECEIVERS FROM SINGLE-SIDED REFLECTION DATA

B.3.1. VIRTUAL ACOUSTICS METHODOLOGY

The snapshots in Figure B.1 (for both methods) were obtained by numerically modelling the medium's response to fields emitted from (parts of) its boundary. These snapshots nicely visualise the propagation, scattering, focusing and defocusing of the fields inside the medium. In practical situations these fields are not visible, unless receivers would be placed throughout the medium, which is of course not feasible. However, our focusing methodology can be extended to create not only virtual sources, but also virtual receivers anywhere inside the medium. As input we need the reflection response of the medium, measured with sources and receivers at the accessible boundary $\mathbb{S}_0$ only (hence, no physical sources nor receivers are needed inside the medium). The reflection response is represented by the Green's function $G(\mathbf{x}',\mathbf{x},t)$, where $\mathbf{x}$ denotes the variable position of the source and $\mathbf{x}'$ that of the receiver, both at $\mathbb{S}_0$. Consider the following variant of equation (B.3)

$$G(\mathbf{r},\mathbf{x},t)+G(\mathbf{r},\mathbf{x},-t)=\text{Symmetrize}\Big(\int_{\mathbb{S}_0} G(\mathbf{x}',\mathbf{x},t)*F(\mathbf{x}',\mathbf{r},t)\,\mathrm{d}\mathbf{x}'\Big) \tag{B.4}$$

(see Supplementary Information). This expression shows how the recorded data $G(\mathbf{x}',\mathbf{x},t)$, measured at the upper boundary of the medium, are transformed into $G(\mathbf{r},\mathbf{x},t)$ and its time-reversal, being the response to a real source at $\mathbf{x}$, observed by a virtual receiver at $\mathbf{r}$ anywhere inside the medium. The focusing function $F(\mathbf{x}',\mathbf{r},t)$, required for this transformation, can be derived from the recorded data $G(\mathbf{x}',\mathbf{x},t)$, using the multidimensional Marchenko method[25, 30, 31, 66, 122]. We have implemented a 2D version of the Marchenko method as an iterative proces[34]. The time-reversal of the direct arriving wave between $\mathbf{x}$ and $\mathbf{x}'$ is used as an initial estimate of the focusing function $F(\mathbf{x}',\mathbf{r},t)$. This direct arrival, in turn, is based on an estimate of the propagation velocity of the medium. This does not require information about the layer interfaces, nor about the internal structure of the layers: a smooth background model suffices to compute the direct arrival[32]. Note that estimating a background model is state-of-the-art methodology in geophysical imaging [123]. Then, by evaluating equation (B.4) we obtain $G(\mathbf{r},\mathbf{x},t)$ for any virtual receiver position $\mathbf{r}$ inside the medium. Next, using the retrieved virtual-receiver data $G(\mathbf{r},\mathbf{x},t)$ in the right-hand side of equation (B.3), we obtain $G(\mathbf{r},\mathbf{s},t)$ and its time-reversal, being the response to a virtual source at $\mathbf{s}$, observed by virtual receivers at $\mathbf{r}$.

Theoretical research shows that this methodology can be generalised for vectorial wave fields in lossless media, such as electromagnetic waves, elastodynamic waves (after decomposition at the surface into P- and S-waves), etc.[124, 125]. Small to moderate propagation losses can be accommodated by applying loss corrections to the data before applying the Marchenko method[126]. In the following we apply the virtual acoustics

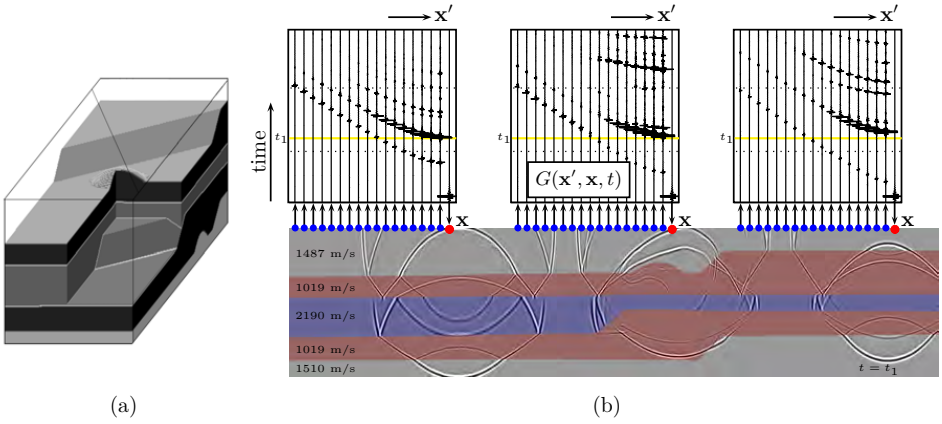

Figure B.2: (a) 3D physical model. The grey-levels indicate different propagation velocities and mass densities. Ultrasonic reflection experiments are carried out along the diagonal line above the model. (b) 2D cross-section of the physical model (with modelled snapshots, for visualisation only) and the actually recorded response at the surface, $G(\mathbf{x}', \mathbf{x}, t)$ (here shown for 3 source positions $\mathbf{x}$ and 3 × 17 receiver positions $\mathbf{x}'$).

methodology for scalar wave fields in lossless media, as outlined above, to ultrasonic physical model data and seismic reflection data.

B.3.2. APPLICATION TO ULTRASONIC PHYSICAL MODEL DATA

Figure B.2(a) shows a 3D physical model, composed of silicone gel and beeswax layers with different acoustic propagation velocities (their numerical values are tabulated in Figure B.2(b)). The size of the model is 70 × 600 × 600 mm. The model is placed in a watertank and probed with ultrasound, emitted and received by piezo-electric transducers in the water. The acquisition is carried out along a horizontal diagonal line (indicated in Figure B.2(a)), 12 mm above the upper boundary of the model and perpendicular to its main structures. A 2D cross-section of the model below the acquisition line is shown in Figure B.2(b). The emitting transducer sends a sweep signal in the frequency range 0.4 MHz to 1.8 MHz. The resulting wave field propagates through the water into the model, propagates and scatters inside the model, and propagates back through the water to the acquisition line, where it is recorded by a receiving transducer. The recorded response is deconvolved for the sweep signal, effectively compressing the source signal to a short zero-phase pulse with a central frequency of 1.1 MHz[127]. This experiment is repeated 106 times, with the source at the same position and the receiver moving along the acquisition line in steps of 1.25 mm. Next, the source is moved 1.25 mm along the line and again 106 traces are recorded. This whole process is carried out 301 times, leading to a recorded reflection response consisting of 301 × 106 = 31 906 traces. Figure B.2(b) shows 51 of those traces, for 3 source positions and 17 receivers per source position. Before further processing, source-receiver reciprocity is applied, effectively doubling the number of traces, and the data are interpolated to a twice as dense spatial grid (source and

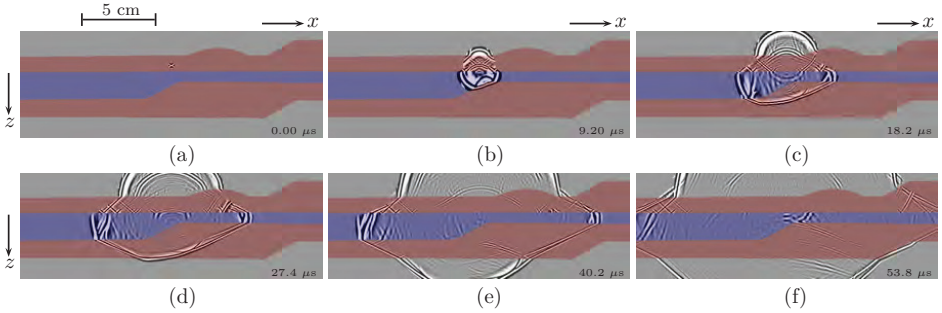

Figure B.3: Virtual response $G(\mathbf{r},\mathbf{s},t)+G(\mathbf{r},\mathbf{s},-t)$, retrieved from the single-sided ultrasonic reflection response $G(\mathbf{,}\mathbf{x},t)$ of the physical model in Figure B.2(a). (a) $t=0$ μs. (b) $t=9.2$ μs. (c) $t=18.2$ μs. (d) $t=27.4$ μs. (e) $t=40.2$ μs. (f) $t=53.8$ μs.

receiver spacing 0.625 mm) to suppress spatial aliasing.

We denote the recorded reflection response by Green's function $G(\mathbf{x}',\mathbf{x},t)$, where $\mathbf{x}$ denotes the variable position of the source and $\mathbf{x}'$ that of the receiver (actually the recorded response is the Green's function convolved with the compressed source pulse, but for the sake of simplicity we treat the recorded data as a Green's function). We apply the methodology discussed above to this response. Figure B.3 shows snapshots of the virtual acoustic response $G(\mathbf{r},\mathbf{s},t)+G(\mathbf{r},\mathbf{s},-t)$, for a fixed virtual source inside the second layer of the 3D physical model and variable virtual receiver positions $\mathbf{r}$ throughout the 2D cross-section of the model. The different colours in the background of this figure indicate the different layers. We used the velocities of these layers to model the direct arrivals, as initial estimates for the focusing functions. Note, however, that we did not use information about the layer interfaces for the retrieval of the virtual response: all scattering information comes directly from the recorded reflection response. The figure clearly shows the evolution of the wave field through the medium, including scattering at the layer interfaces. Imperfections are explained by the finite aperture, the limited radiation angles of the piezo-electric transducers, the negligence of evanescent waves and the fact that we used a 2D method to retrieve this virtual wave field in a 3D medium.

B.3.3. APPLICATION TO SEISMIC REFLECTION DATA

The proposed methodology can be applied to reflection data at a wide range of scales. Next we apply our methodology to vintage seismic reflection data, acquired in 1994 over the Vøring Basin by SAGA Petroleum A.S. (currently part of Statoil ASA). We use a smooth background model to define the initial estimates of the focusing functions. Figure B.4 shows snapshots of $G(\mathbf{r},\mathbf{s},t)+G(\mathbf{r},\mathbf{s},-t)$ obtained from these seismic data. Again, the evolution of the retrieved wave field clearly includes the primary and multiply scattered events, which have been obtained directly from the recorded reflection data. In the background these snapshots show an independently obtained seismic image of the interfaces between the geological layers, for visualisation only. Note the consistency between the position of these interfaces and the apparent origin of scattering in the snapshots.

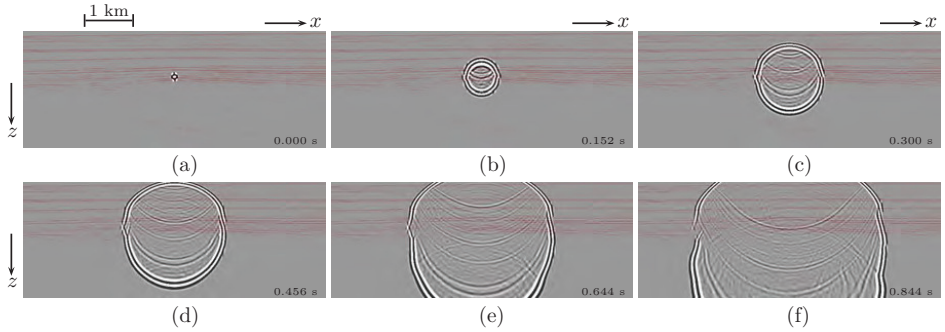

Figure B.4: Virtual response $G(\mathbf{r},\mathbf{s},t)+G(\mathbf{r},\mathbf{s},-t)$, retrieved from the single-sided seismic reflection response $G(\mathbf{x}',\mathbf{x},t)$ of the Vøring Basin. (a) $t=0$ ms. (b) $t=152$ ms. (c) $t=300$ ms. (d) $t=456$ ms. (e) $t=644$ ms. (f) $t=844$ ms.

B.4. DISCUSSION

The ability to create virtual sources and receivers inside a medium from single-sided reflection data opens new ways for imaging and monitoring. An exciting new field in medical imaging is photoacoustic (PA) imaging [128], a method which employs the conversion of optical energy into acoustic energy at those locations inside the medium where light is absorbed. The resulting acoustic wave field may be very complex because usually many PA sources go off simultaneously and inhomogeneities in the medium may cause reflection artefacts[129]. Our proposed virtual acoustics methodology could be applied to ultrasonic reflection measurements to predict the direct and scattered wave fields of (clusters of) virtual PA sources, thus improving the interpretation and imaging of the complex wave field of actual PA sources. With the emergence of dual-modality ultrasound and photoacoustic imaging tools [130] this becomes feasible and the first steps in this direction have already been made [122]. Note that in medical applications it is often sufficient to use a homogeneous background model, which means that analytical expressions can be used for the initial estimate of the focusing functions. Real-time application of our virtual acoustics methodology for medical imaging therefore seems feasible, particularly when the imaging is restricted to a finite region of interest.

Another exciting potential application is the investigation of induced seismicity. By acquiring high-resolution seismic reflection data in areas prone to induced seismicity, our virtual acoustics approach could forecast the wave field and the associated ground motion caused by possible future earthquakes. Moreover, when the same acquisition system is also used to passively record the response to actual induced earthquakes, our method could be used to create virtual seismometers in the subsurface around the actual earthquake and use these to retrieve accurate knowledge of the source mechanism of the earthquake, insight in the evolution of the geomechanical state of the subsurface (horizontal and vertical stress distribution, fault and fracture properties etc.), and deep understanding of the link between the earthquake and the observed ground motion.

CURRICULUM VITÆ

Joeri BRACKENHOFF

23-11-1992	Born in Delft, the Netherlands.

EDUCATION

2011–2014	B.Sc. in Applied Earth Sciences (*Cum Laude*) Delft University of Technology, Delft, the Netherlands
2014–2016	M.Sc. IDEA League in Applied Geophysics (*Cum Laude*) Delft University of Technology, Delft, the Netherlands Eidgenössische Technische Hochschule, Zürich, Switzerland Rheinisch-Westfälische Technische Hochschule, Aachen, Germany Thesis at Delft University of Technology, Delft, the Netherlands
2016–2020	Ph.D. project in Applied Geophysics *Thesis:* Monitoring and forecasting of seismic wavefields in the subsurface *Promotor:* Prof. dr. ir. C.P.A. Wapenaar

AWARDS

2015-2016	Dean's list of top 5% of students of RWTH Aachen
2016	Best Graduate of the Faculty of Civil Engineering and Geosciences, TU Delft
2016	Best Thesis award of the Department of Geoscience and Engineering of the Faculty of Civil Engineering and Geosciences, TU Delft
2019	Travel grant for the SEG/Chevron SLS symposium at the SEG annual conference 2019

INTERNSHIPS

2015	Internship at the Dutch Geological Survey, TNO, in Utrecht, the Netherlands
2019	Internship at SENAI CIMATEC in Salvador de Bahia, Brazil

STUDENTS SUPERVISED

2016-2017	M.Sc. student *Student:* Yury Alkhimenkov *Thesis:* Redatuming and quantifying attenuation from reflection data using the Marchenko equation: A novel approach to quantify Q-factor and seismic upscaling
2019-2020	M.Sc. student *Student:* Hassan Alfaraj *Thesis:* Imaging angle-dependent reflectivity using the Marchenko method

EXTRACURRICULAR ACTIVITIES

2017-2020	President of Delft Organization of Geophysics Students (D.O.G.S.), SEG and EAGE student chapters.
2018-2020	Ph.D. representative of the Department of Geoscience and Engineering
2018-2020	Member of the Ph.D. council of the Faculty of Civil Engineering and Geoscience, TU Delft.
2018-2020	Member of the University Ph.D. council of TU Delft.
2018	Participant in the AAPG Imperial Barrel Award.
2018	Participant in the SPE Petrobowl Challenge.
2019	Member of the SEG Europe Regional Advisory Committee.
2019-present	Founding member and student representative of the EAGE local chapter Netherlands

LIST OF PUBLICATIONS

JOURNAL PUBLICATIONS

1. J. Hunziker, J. Thorbecke, **J. Brackenhoff** & E Slob, *Inversion of controlled-source electromagnetic reflection responses*, Geophysics **81**, F49 (2016).

2. J. Thorbecke, E. Slob, **J. Brackenhoff**, J. van der Neut & K. Wapenaar, *Implementation of the Marchenko method*, Geophysics **82**, WB29 (2017).

3. J. van der Neut, **J. Brackenhoff**, M. Staring, L. Zhang, S. de Ridder, E. Slob & K. Wapenaar, *Single-and double-sided Marchenko imaging conditions in acoustic media*, IEEE Transactions on Computational Imaging **4**, 160 (2017).

4. K. Wapenaar, **J. Brackenhoff**, J. Thorbecke, J. van der Neut, E. Slob & E. Verschuur, *Virtual acoustics in inhomogeneous media with single-sided access*, Scientific Reports **8**, 2497 (2018).

5. **J. Brackenhoff**, K. Wapenaar & J. Thorbecke, *Monitoring of induced distributed double-couple sources using Marchenko-based virtual receivers*, Solid Earth **10**, 1301 (2019).

6. K. Wapenaar, **J. Brackenhoff** & J. Thorbecke, *Green's theorem in seismic imaging across the scales*, Solid Earth **10**, 517 (2019).

7. **J. Brackenhoff**, J. Thorbecke & K. Wapenaar, *Virtual sources and receivers in the real Earth: Considerations for practical applications*, Journal of Geophysical Research: Solid Earth **124**, 11802 (2019).

8. **J. Brackenhoff**, J. Thorbecke, V. Koehne, D. Barrera & K. Wapenaar, *Implementation of the 3D Marchenko method*. ArXiv, preprint number:2004.00896 (2020)

9. **J. Brackenhoff**, J. Thorbecke & K. Wapenaar, *3D virtual seismology.* Submitted for publication to IEEE Transactions on Geoscience and Remote Sensing (2020).

10. K. Wapenaar, **J. Brackenhoff**, M. Dukalski, G. Meles, C. Reinicke, E. Slob, M. Staring, J. Thorbecke, J. van der Neut & L. Zhang, *Marchenko redatuming, imaging and multiple elimination, and their mutual relations.* Submitted for publication to Geophysics (2020).

CONFERENCE ABSTRACTS

1. **J. Brackenhoff**, J. Thorbecke & K. Wapenaar, *Homogeneous Green's function retrieval using the Marchenko method*, SIAM Houston Imaging Sciences Symposium 1 (2017).

2. **J. Brackenhoff**, J. van der Neut & K. Wapenaar, *Decomposition of the Green's function using the Marchenko equation*, SEG Technical Program Expanded Abstracts 5593 (2017).

3. K. Wapenaar, **J. Brackenhoff**, J. Thorbecke, J. van der Neut, E. Slob & E. Verschuur, *Virtual Seismology: monitoring the Earth's crust with virtual sources and virtual receivers in the subsurface*, Conference Proceedings, 20th EGU General Assembly 8338 (2018).

4. K. Wapenaar, **J. Brackenhoff**, J. Thorbecke, J. van der Neut & E. Slob, *A single-sided representation for the homogeneous Green's function, accounting for all multiples*, Conference Proceedings, 80th EAGE Conference and Exhibition 1 (2018).

5. G.A. Meles, C. Reinicke, K. Wapenaar, **J. Brackenhoff** & J.W. Thorbecke, *Focusing conditions: a comparison between different Marchenko imaging strategies*, Conference Proceedings, 80th EAGE Conference and Exhibition Workshop Programme 1 (2018).

6. Y. Alkhimenkov, **J. Brackenhoff**, E. Slob & K. Wapenaar, *Q-factor Estimation and Redatuming in a Lossy Medium Using the Marchenko Equation*, Conference Proceedings, 80th EAGE Conference and Exhibition 1 (2018).

7. **J. Brackenhoff**, J. Thorbecke & K. Wapenaar, *Homogeneous Green's Function Retrieval on Field Data Using the Marchenko Method*, Conference Proceedings, 80th EAGE Conference and Exhibition 1 (2018).

8. I. Vasconcelos, C. Reinicke & **J. Brackenhoff**, *Estimating Overburden-Only Transmission Waveforms from Surface Reflection Data*, Conference Proceedings, 80th EAGE Conference and Exhibition 1 (2018).

9. K. Wapenaar, **J. Brackenhoff**, & J. Thorbecke, *Marchenko method for monitoring induced seismicity with virtual receivers*, SEG Technical Program Expanded Abstracts 5520 (2018).

10. K. Wapenaar, **J. Brackenhoff**, M. Staring, J. Thorbecke, & E. Slob, *Virtual seismology: from hydrocarbon reservoir imaging to induced earthquake monitoring*, American Geophysical Union, Fall Meeting S53A-03 (2018).

11. **J. Brackenhoff**, J. Thorbecke & K. Wapenaar, *Monitoring Induced Distributed Double-Couple Sources Using Marchenko-Based Virtual Receivers*, Conference Proceedings, 81st EAGE Conference and Exhibition 1 (2019).

12. K. Wapenaar, **J. Brackenhoff** & J. Thorbecke, *Green's Theorem in Time-Reversal Acoustics, Back Propagation and Source-Receiver Redatuming: a Tutorial*, Conference Proceedings, 81st EAGE Conference and Exhibition 1 (2019).

13. K. Wapenaar, **J. Brackenhoff** & J. Thorbecke, *Practical considerations for the creation of virtual sources and receivers in the 3D subsurface*, American Geophysical Union, Fall Meeting S24C-04 (2019).

14. H. Alfaraj, **J. Brackenhoff** & K Wapenaar, *Obtaining Angle-Dependent Reflectivity Using the Marchenko Redatuming Method*, Conference Proceedings, 82nd EAGE Conference and Exhibition Workshop Programme 1 (2020).

15. **J. Brackenhoff**, J. Van IJsseldijk & K. Wapenaar, *Data-driven internal multiple elimination applications using imperfectly sampled reflection data*, Conference Proceedings, 82nd EAGE Conference and Exhibition Workshop Programme 1 (2020).

BOOK CHAPTERS

1. D. Ngan-Tillard, D. Draganov, M. Warnaar, J. Liu, **J. Brackenhoff**, J. van de Berg, R. Ghose, A. Veltmeijer & H. Stöger, 'Exploring with GPR the frigidarium of the Byzantine baths in Ostia Antica after excavation, backfilling and floor re-tiling' in *Designating place: Archaeological perspectives on built environments in Ostia and Pompeii* (in press) 2021.

2. R. Ghose, J. Liu, D. Draganov, D. Ngan-Tillard, M. Warnaar, **J. Brackenhoff**, J. van de Berg & H. Stöger, 'Ultra-shallow shear-wave reflections locating near-surface buried structures in the unexcavated southern fringe of the ancient Ostia, Rome' in *Designating place: Archaeological perspectives on built environments in Ostia and Pompeii* (in press) 2021.