AES/TGP/13-24 Surface-wave retrieval using Seismic Interferometry by multi-dimensional deconvolution

29-08-13 Vincent de Waard

Title	:	Surface-wave retrieval using Seismic Interferometry by multi-dimensional deconvolution

Author(s)	:	Vincent M. de Waard

Date	:	29 August 2013
Professor(s)	:	
Supervisor(s)	:	Dr.Ir. Karel N. van Dalen, Dr. Ir. Elmer N Ruigrok
TA Report number	:	AES/TGP/13-23

Postal Address	:	Section for Applied Geophysics and Petrophysics Department of Geoscience & Engineering Delft University of Technology P.O. Box 5028 The Netherlands
Telephone	:	(31) 15 2781328 (secretary)
Telefax	:	(31) 15 2781189

Abstract

In this thesis, seismic interferometry by multi dimensional deconvolution is used to retrieve the surface-wave response from a virtual source to a receiver by using the observations due to, often passive, noise sources located elsewhere. The virtual source response is used to retrieve the phase velocity, a property used for shallow subsurface velocity models. Seismic interferometry by multi dimensional deconvolution is an existing method, but has not been applied on actual surface waves recordings. It is not known if: the existing seismic receiver-array networks are fit for surface wave retrieval; the receiver spacing is small enough; it is possible to select the correct passive sources; the results improved with respect to seismic interferometry by crosscorrelation. Three existing arrays, with data in the IRIS database, are modelled to see if it is possible to retrieve an accurate phase velocity that matches the input velocity. Under the right circumstances, MDD is indeed suitable to retrieve an accurate virtual source response. However, not all the arrays have the correct proportions, and it proves to be difficult to separate the desired from undesired sources. A small adjustment to the one-sided MDD theory allows for full wavefield MDD with two-sided illumination. The alternative MDD theory seems to work well on modelled data, but more extensive testing is required.

Acknowledgements

I want to thank all the people that supported me during this thesis. Special thanks goes to Elmer and Karel for their support and patience.

Delft, University of Technology
August 29, 2013

Vincent de Waard

Table of Contents

Chapter 1

Introduction

In this thesis, we look at surface wave retrieval by means of seismic interferometry. Surface waves are used to obtain a velocity model of the shallow subsurface, which is often used for improving imaging at greater depth. Surface wave retrieval knows many successful applications in tomographic imaging of the earth crust, see for instance Sabra et al. (2005), Gerstoft et al. (2005). The basic idea is simple: assume uniformly distributed sources on a closed surface, and a lossless medium. The crosscorrelation of the observations at two seismometers inside this closed surface, due to all sources, converges to the Green's function between these seismometers. Basically, one of the seismometers becomes a virtual source. In theory, the full surface wave could be retrieved (Wapenaar, 2004), but in practical situations often only the fundamental mode of the direct wave is obtained. Furthermore, the uniform source distribution assumption is often not met, this reduces the accuracy of the retrieved surface waves. A way to improve the Green's function retrieval is via multi-dimensional deconvolution (MDD). The concept of MDD is based on inverting an exact convolution Green's function representation (Wapenaar et al., 2011b) and avoids the restrictions of a lossless medium and uniformly distributed sources in a natural way. MDD uses the crosscorrelation on an array of seismometers to a single seismometers. The pointspread function, function of the array and source distribution, tries to compensate for possible deviations from the uniformly distributed sources, therefore the retrieved Green's function could improve with respect to the crosscorrelation method. Wapenaar et al. (2011a) show that MDD is effective for surface wave retrieval.
In this thesis, we use MDD for surface wave retrieval. A model study is performed to compare the results of the regular crosscorrelation method with the MDD method for three existing seismological arrays in North-America. What are the conditions that the arrays should fulfil in order to get an improved result compared to the crosscorrelation method. The surface waves used here are so called microseisms, and the related frequency bands of the microseisms used in this thesis are the so called single and double frequency bands. The single frequency band ranges from 0.03 to 0.09 hertz, and the double frequency band ranges from 0.09 to 0.5 hertz. These microseisms are ambient seismic vibrations related to swell waves in oceans (Ruigrok (2012)). The generation of microseisms is strongly affected by strong oceanic storms and therefore, has a seasonal pattern. There are two distinctive peaks in the microseism band. The first falls in the single frequency band and is related to strong ocean waves. The second peak falls in the double frequency band and is also related to strong ocean waves, but is a result of interaction between waves and their relative velocity. In general it contains stronger

amplitdes than the first peak. Microseisms are used in this thesis, because they are easy to measure and present almost anywhere on earth at any time. Preferably the single frequency band is used, the wavelenghts are longer; which allows larger receiver spacings in the MDD array.
Chapter 2 explains the basic interferometry approach by crosscorrelation; shows the general derivation of the MDD approach, and explains how these equations are used for modelling. A feasibility study is performed on three arrays on the North-American continent in chapter 3. In chapter 4 follows a slight modification to the general MDD theory in order to use ingoing and outgoing wavefields, thereby avoiding one of the restrictions of regular MDD. The modification follows from changing the boundary condition in the reference medium. Due to this change, also the multiples of the virtual source response are constructed, but these are separated in time and should cause no problem to remove if desired. A small model study is performed in chapter 5 with this modified MDD theory. Finally followed by the conclusions in the last chapter.

Chapter 2

Theory

The basic theory of seismic interferometry by crosscorrelation and multidimensional deconvolution, and in this case surface wave retrieval will be described in short. For a full review of interferometry in general, the reader is referred to Wapenaar et al. (2011b), and for a explanation with respect to surface wave retrieval the reader is referred to Wapenaar et al. (2011a).

2-1 Seismic interferometry by crosscorrelation

Interferometry is suitable to apply to controlled source experiments, or noise observations. Assuming the ambient noise field is equipartitioned, or the noise sources are uniformly distributed, and the medium is lossless, then the crosscorrelation of noise observations at two seismometers yields the Green's function between these seismometers. In many situations, only the fundamental mode of the direct surface wave is retrieved and the accuracy of the retrieved waves is limited. The uniform distribution of noise and a lossless medium are serious restrictions, and Multi-Dimensional Deconvolution avoids these restrictions.

In other words, seismic interferometry aims at retrieving a response from receiver $\mathbf{x}_A$ to receiver $\mathbf{x}_B$ while using sources located elsewhere. Note that $\mathbf{x}$ denotes a position vector. Basically, the observations are used to create a virtual source at receiver $\mathbf{x}_A$, and the response is measured in receiver $\mathbf{x}_B$. In order to do so, the signals generated by real sources need to be measured both in $\mathbf{x}_A$ and $\mathbf{x}_B$.

In this thesis, only transient sources are considered. For ambient noise applications, similar equations are obtained, see Wapenaar and Fokkema (2006) and Wapenaar et al. (2011b). The crosscorrelation function of two responses $\hat{u}\left(\mathbf{x}_A, \mathbf{x}_S^{(i)}, \omega\right)$ and $\hat{u}\left(\mathbf{x}_B, \mathbf{x}_S^{(i)}, \omega\right)$ for two receivers $\mathbf{x}_A$ and $\mathbf{x}_B$, illuminated by the same sources, is given by

$$\hat{C}\left(\mathbf{x}_B, \mathbf{x}_A\right) = \sum_i \hat{u}\left(\mathbf{x}_B, \mathbf{x}_S^{(i)}\right) \left\{\hat{u}\left(\mathbf{x}_A, \mathbf{x}_S^{(i)}\right)\right\}^* . \tag{2-1}$$

The hats indicate that the functions are in the frequency space domain, therefore ω is omitted. A correlation in the time domain turns into a multiplication with the complex conjugate in the frequency domain. The crosscorrelation function from $\mathbf{x}_A$ to $\mathbf{x}_B$ is given by the sum over

State A	State B
$\hat{q}_A(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_S)$	$\hat{q}_B(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_B)$
$\hat{p}_A(\mathbf{x}) = \hat{G}^{p,q}(\mathbf{x}, \mathbf{x}_S)$	$\hat{p}_B(\mathbf{x}) = \hat{G}^{p,q}(\mathbf{x}, \mathbf{x}_B)$
$\hat{v}_{j,A}(\mathbf{x}) = \hat{G}_j^{v,q}(\mathbf{x}, \mathbf{x}_S)$	$\hat{v}_{j,B}(\mathbf{x}) = \hat{G}_j^{v,q}(\mathbf{x}, \mathbf{x}_B)$
$\hat{G}_j^{v,q}(\mathbf{x}, \mathbf{x}_S) = \frac{-1}{i\omega\rho(\mathbf{x})}\partial_j \hat{G}^{p,q}(\mathbf{x}, \mathbf{x}_S)$	$\hat{G}_j^{v,q}(\mathbf{x}, \mathbf{x}_B)\frac{-1}{i\omega\rho(\mathbf{x})}\partial_j \hat{G}^{p,q}(\mathbf{x}, \mathbf{x}_B)$

Table 2-1: This table shows the definition of the different terms introduced

all individual crosscorrelated responses for all sources to $\mathbf{x}_A$ and $\mathbf{x}_B$. In case of an irregular source distribution, and the source locations are unknown; the crosscorrelation function in equation 2-1 is related to the sought Green's functions between receiver $\mathbf{x}_A$ and $\mathbf{x}_B$ with the following equation according to Wapenaar et al. (2011b)

$$\left(\hat{G}(\mathbf{x}_B, \mathbf{x}_A) + \hat{G}^*(\mathbf{x}_B, \mathbf{x}_A)\right)\hat{S}(\omega) \approx \frac{2}{\rho c}\hat{C}(\mathbf{x}_B, \mathbf{x}_A) \tag{2-2}$$

In this equation $\hat{S}(\omega)$ denotes the autocorrelation of the source wavelet. Normally the crosscorrelation theory is applied on a closed surface around the receivers. In case of one sided illumination, only one of the two Green's functions in equation 2-2 is retrieved.

2-2 Derivation of MDD relation for seismic interferometry

The MDD relation is based on the convolution reciprocity equation. See for instance Wapenaar and Fokkema (2006) for an explanation. The convolution reciprocity equation reads

$$\int_V \left[\hat{p}_A\hat{q}_B - \hat{v}_{j,A}\hat{f}_{j,B} - \hat{q}_A\hat{p}_B + \hat{f}_{j,A}\hat{v}_{j,B}\right] dV = \oint_{\partial V} \left[\hat{p}_A\hat{v}_{j,B} - \hat{v}_{j,A}\hat{p}_B\right] n_j dS \tag{2-3}$$

where $\hat{p}$ is the pressure response, $\hat{v}_j$ is the particle velocity response in j direction, $\hat{q}$ is an impulsive point source of volume injection rate and $\hat{f}$ is an impulsive point source of force. The relations between the terms in equation 2-3 and the Green's functions are defined in table 2-1. The sub- and superscripts in $\hat{G}_j^{v,q}(\mathbf{x}_B, \mathbf{x}_S)$ denote the observation of v measured in direction j at location $\mathbf{x}_B$ due to a source q at $\mathbf{x}_S$. The situation is slightly modified with respect to the crosscorrelation and $\mathbf{x}_A$ is taken outside volume V and renamed as $\mathbf{x}_S$. Assume that only volume injection sources exist. This means that $\hat{f}$ is zero, and $\hat{q}_{A,B}$ is replaced by $\delta(\mathbf{x} - \mathbf{x}_{S,B})$. The integration over V shall only yield $\hat{q}_B\hat{p}_A$; since $\delta(\mathbf{x} - \mathbf{x}_S)$ is located outside the volume V, $\hat{q}_A\hat{p}_B$ is zero. Substituting the terms in table 2-1 into equation 2-3 and integrating over volume V gives:

$$\hat{G}^{p,q}(\mathbf{x}_B, \mathbf{x}_S) = \oint_{\partial V} \left[\hat{G}^{p,q}(\mathbf{x}, \mathbf{x}_S)\, n_j \hat{\hat{G}}_j^{v,q}(\mathbf{x}, \mathbf{x}_B) - \hat{\hat{G}}^{p,q}(\mathbf{x}, \mathbf{x}_B)\, n_j \hat{G}_j^{v,q}(\mathbf{x}, \mathbf{x}_S)\right] dS \tag{2-4}$$

$\hat{G}$ denotes the Green's function in the true medium, and $\hat{\hat{G}}$ denotes the Green's function in the reference medium; shown as state B in table 2-1. $\mathbf{x}_S$ are source locations in the true medium, outside V. $n_j\hat{\hat{G}}_j^{v,q}(\mathbf{x}, \mathbf{x}_B)$ is the projected velocity on the unitvector perpendicular

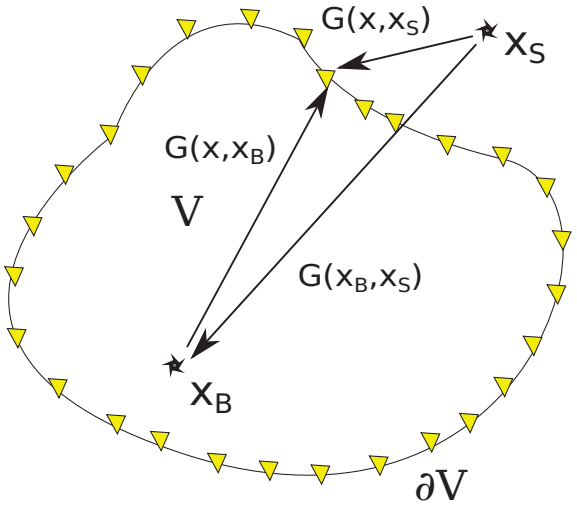


Figure 2-1: Illustration of Greens functions and the sources and receivers. ∂V is an enclosing surface around $\mathbf{x}_B$ and is well sampled with receivers (yellow triangles). There is a source outside ∂V and a source at location $\mathbf{x}_B$ inside ∂V.

to the boundary ∂V at location $\mathbf{x}$ due to sources in $\mathbf{x}_B$. Figure 2-1 illustrates equation 2-4. ∂V is well sampled with receivers, and sources are located outside, but also inside ∂V. The arrows illustrate the different Green's functions. Equation 2-4 is the starting point for interferometry by multi dimensional deconvolution. It is valid in media with losses, which is an advantage compared to the crosscorrelation approach. Equation 2-4 contains two convolution terms under the integral, however rewriting the equation to one term simplifies solving it for $\hat{\tilde{G}}^{p,q}(\mathbf{x},\mathbf{x}_B)$. For the real medium, reality defines the boundary conditions, but for the reference medium other boundary conditions can be chosen. Assume that the reference medium is bounded by an absorbing boundary; and split up $\hat{G}^{p,q}(\mathbf{x},\mathbf{x}_S)$ in an in and out going field: $\hat{G}^{p,q}(\mathbf{x},\mathbf{x}_S) = \hat{G}^{p,q;\text{in}}(\mathbf{x},\mathbf{x}_S) + \hat{G}^{p,q;\text{out}}(\mathbf{x},\mathbf{x}_S)$. Equation 2-4 becomes (Wapenaar et al., 2011b):

$$\hat{G}^{p,q}(\mathbf{x}_B,\mathbf{x}_S) = 2\oint_{\partial V}\left[n_j\hat{\tilde{G}}_j^{v,q}(\mathbf{x},\mathbf{x}_B)\,\hat{G}^{p,q;\text{in}}(\mathbf{x},\mathbf{x}_S)\right]dS = \oint_{\partial V}\frac{-2}{i\omega\rho}\left[\left(n_j\partial_j\hat{\tilde{G}}^{p,q}(\mathbf{x},\mathbf{x}_B)\right)\hat{G}^{p,q;\text{in}}(\mathbf{x},\mathbf{x}_S)\right]dS \quad (2\text{-}5)$$

Defining a dipole Green's function $\hat{\tilde{G}}_d(\mathbf{x},\mathbf{x}_B) = \frac{-2}{i\omega\rho}n_j\partial_j\hat{\tilde{G}}^{p,q}(\mathbf{x},\mathbf{x}_B)$ and applying source receiver reciprocity $\hat{\tilde{G}}^{p,q}(\mathbf{x},\mathbf{x}_B) = \hat{\tilde{G}}^{p,q}(\mathbf{x}_B,\mathbf{x})$ simplifies equation 2-5 to:

$$\hat{G}^{p,q}(\mathbf{x}_B,\mathbf{x}_S) = \oint_{\partial V}\left[\hat{\tilde{G}}_d(\mathbf{x}_B,\mathbf{x})\,\hat{G}^{p,q;\text{in}}(\mathbf{x},\mathbf{x}_S)\right]dS \quad (2\text{-}6)$$

In practical situations there are often no well sampled receivers on a closed boundary; integration over the sampled part of ∂V, or S_{rec}, is sufficient, because the radiation conditions apply on the remaining half-sphere closing the boundary. Contribution of the integral disappears

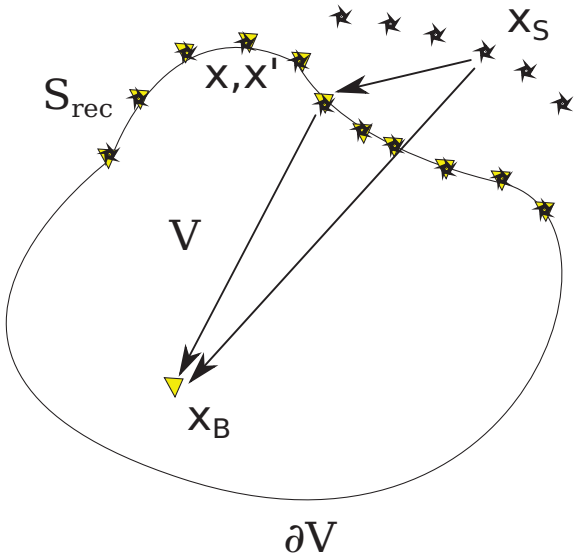


Figure 2-2: Illustration of the situation with S_{rec} as a part of ∂V.

when the half-sphere is large enough (Wapenaar and Fokkema, 2006). Hence, equation 2-6 can be rewritten to

$$\hat{G}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S\right)=\int_{S_{\text{rec}}}\left[\hat{\hat{G}}_d\left(\mathbf{x}_B,\mathbf{x}\right)\hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S\right)\right]dS, \tag{2-7}$$

where S_{rec} is the part of ∂V covered by the receiver array. Figure 2-2 shows the new situation after source-receiver reciprocity and S_{rec} is the part of ∂V covered with receivers. Furthermore the image shows what $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{x}_B$ and $\mathbf{x}_S$ indicate. Equation 2-7 is the equation that needs to be solved. Ideally, inverting $\hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S\right)$ should yield the sought Green's function, namely $\hat{\hat{G}}_d^{p,q}\left(\mathbf{x}_B,\mathbf{x}\right)$. Practically this inversion is often not possible; for instance the inverse is not well defined, or the matrix is not square. Therefore equation 2-7 needs to be manipulated. Numerical inversions are often solved with the linear least squares inversion, or the minimum norm solution. Both options will be explained in this chapter.

2-2-1 Linear Least Squares Inversion

A possible way to avoid a direct inversion of $\hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S\right)$ is to apply the linear least squares inversion. The linear least squares inversion can be applied to overdetermined systems of equations, so there are more equations than unknowns. The linear least squares inversion multiplies both sides with the complex conjugate transpose of $\hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S\right)$, thereby inverting $\hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S\right)\left\{\hat{G}^{p,q;\text{in}}\left(\mathbf{x}',\mathbf{x}_S\right)\right\}^*$ instead; where $\mathbf{x}'$ is an additional coordinate vector over S_{rec}. The final MDD equation follows from equation 2-7 by multiplying both sides with

$\left[\hat{G}^{p,q;\text{in}}\left(\mathbf{x}',\mathbf{x}_S\right)\right]^*$, and summing over the sources gives:

$$\sum_i \left[\hat{G}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S^{(i)}\right)\left\{\hat{G}^{p,q;\text{in}}\left(\mathbf{x}',\mathbf{x}_S^{(i)}\right)\right\}^*\right] = \int\limits_{S_{\text{rec}}}\left[\hat{\tilde{G}}_d^{p,q}\left(\mathbf{x}_B,\mathbf{x}\right)\sum_i\left[\hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S^{(i)}\right)\left\{\hat{G}^{p,q;\text{in}}\left(\mathbf{x}',\mathbf{x}_S^{(i)}\right)\right\}^*\right]\right]dS \quad (2\text{-}8)$$

Equation 2-8 consists of Green's functions, but real observations are Green's functions convolved with a source wavelet.

$$\sum_i \left[\hat{G}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S^{(i)}\right)\left\{\hat{G}^{p,q;\text{in}}\left(\mathbf{x}',\mathbf{x}_S^{(i)}\right)\right\}^*\hat{S}\left(\omega\right)^{(i)}\right] = \int\limits_{S_{\text{rec}}}\left[\hat{\tilde{G}}_d^{p,q}\left(\mathbf{x}_B,\mathbf{x}\right)\sum_i\left[\hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S^{(i)}\right)\left\{\hat{G}^{p,q;\text{in}}\left(\mathbf{x}',\mathbf{x}_S^{(i)}\right)\right\}^*\hat{S}\left(\omega\right)^{(i)}\right]\right]dS \quad (2\text{-}9)$$

where $\hat{S}\left(\omega\right)^{(i)}$ is the autocorrelation of the source wavelet for the source at location $\mathbf{x}_S^{(i)}$. Equation 2-9 yields the linear least squares solution, and can be found in Wapenaar et al. (2011b,a) with a few small modifications. First, the crosscorrelation- and pointspread-function are inserted, where the crosscorrelation function is defined as:

$$\hat{C}\left(\mathbf{x}_B,\mathbf{x}'\right) = \sum_i\left(\hat{u}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S^{(i)}\right)\left\{\hat{u}^{p,q;\text{in}}\left(\mathbf{x}',\mathbf{x}_S^{(i)}\right)\right\}^*\right) = \sum_i\left(\hat{G}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S^{(i)}\right)\left\{\hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S^{(i)}\right)\right\}^*\hat{S}\left(\omega\right)^{(i)}\right) \quad (2\text{-}10)$$

and the pointspread function is defined as:

$$\hat{\Gamma}\left(\mathbf{x},\mathbf{x}'\right) = \sum_i\left(\hat{u}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S^{(i)}\right)\left\{\hat{u}^{p,q;\text{in}}\left(\mathbf{x}',\mathbf{x}_S^{(i)}\right)\right\}^*\right) = \sum_i\left(\hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}_S^{(i)}\right)\left\{\hat{G}^{p,q;\text{in}}\left(\mathbf{x}',\mathbf{x}_S^{(i)}\right)\right\}^*S\left(\omega\right)^{(i)}\right) \quad (2\text{-}11)$$

Second, the remaining integral is not continuous since receivers have a discrete size, and a discrete spacing; therefore the integral can be replaced by a sum given that the sampling is small enough. Applying the changes described above give:

$$\hat{C}\left(\mathbf{x}_B,\mathbf{x}'\right) = \sum_{S_{\text{rec}}}\left(\hat{\tilde{G}}_d^{p,q}\left(\mathbf{x}_B,\mathbf{x}\right)\hat{\Gamma}\left(\mathbf{x},\mathbf{x}'\right)\right) \quad (2\text{-}12)$$

$\hat{\tilde{G}}_d$ is the sought virtual-source response. The pointspread function tries to focus the initial estimate of the virtual source response, which is the left hand side of equation 2-12; thereby deghosting and deblurring the initially retrieved response (Wapenaar et al. (2011b)). The MDD approach is similar to the crosscorrelation approach, except that receiver $\mathbf{x}_A$ is

replaced by a whole array $\mathbf{x}$ and an additional inversion is applied. For computational purposes, $\hat{C}(\mathbf{x}_B, \mathbf{x}')$, $\hat{\tilde{G}}_d(\mathbf{x}_B, \mathbf{x})$ and $\hat{\Gamma}(\mathbf{x}', \mathbf{x})$ can be written as matrices. The first and second dimension contain the number of receivers, and the third dimension contains the frequency components. In that case, equation 2-12 becomes a matrix multiplication per frequency component; thereby summing over the desired receiver dimension $\mathbf{x}$.

$$\underline{\underline{C}} = \underline{\underline{G}}_d \underline{\underline{\Gamma}} \tag{2-13}$$

Equation 2-13 can be solved by a stabilized inversion per frequency component, see equation 2-14.

$$\underline{\underline{G}}_d = \underline{\underline{C}}(\underline{\underline{\Gamma}} + \epsilon^2 \underline{\underline{I}})^{-1} \tag{2-14}$$

where ϵ is a small stabilization factor, and $\underline{\underline{I}}$ is an identity matrix. MDD is fit for one-way retrieval only, therefore only sources that have their ray-paths via the MDD array $\mathbf{x}$ to the receiver array $\mathbf{x}_B$ can be used.

2-2-2 Minimum Norm Solution

An alternate solution used in chapter 3, to avoid the direct inversion of $\hat{G}^{p,q;\text{in}}(\mathbf{x}, \mathbf{x}_S)$, is the minimum norm solution. The minimum norm solution is fit to solve under-determined systems of equations, so more unknowns than equations. Instead of multiplying with $\underline{\underline{G}}^\dagger$ on both sides of the equation, note that the dagger indicates the complex conjugate transpose of matrix G, the minimum norm solution tries to solve:

$$2\underline{\underline{m}} = \underline{\underline{G}}^\dagger \underline{\lambda}; \quad \underline{\lambda} = 2\left[\underline{\underline{G}}\underline{\underline{G}}^\dagger\right]^{-1} \underline{\underline{d}} \tag{2-15}$$

Substituting for $\underline{\lambda}$ yields:

$$\underline{\underline{m}} = \underline{\underline{G}}^\dagger \left[\underline{\underline{G}}\underline{\underline{G}}^\dagger\right]^{-1} \underline{\underline{d}} \tag{2-16}$$

Note that the original inversion problem is defined as:

$$\underline{\underline{d}} = \underline{\underline{G}}\underline{\underline{m}} \tag{2-17}$$

$\underline{\underline{d}}$ is the observed data $\hat{G}^{p,q}(\mathbf{x}_B, \mathbf{x}_S)$, $\underline{\underline{G}}$ is the model matrix $\hat{G}^{p,q;\text{in}}(\mathbf{x}, \mathbf{x}_S)$ and $\underline{\underline{m}}$ are the model parameters $\hat{\tilde{G}}_d^{p,q}(\mathbf{x}_B, \mathbf{x})$. $\underline{\lambda}$ is the Lagrange multiplier. The derivation of the minimum norm solution is described in for instance Menke (1989). Notice that in equation 2-17 $\underline{\underline{G}}$ is located before $\underline{\underline{m}}$, while in equation 2-7 $\underline{\underline{G}}$ follows after $\underline{\underline{m}}$. The order of multiplication is switched. The minimum norm solution in equation 2-16 only works when equation 2-17 and 2-7 have the same order of multiplication. However $\underline{\underline{G}}$ and $\underline{\underline{m}}$ are matrices, therefore switching the order of multiplication is not allowed. Transposing equation 2-17 flips the order of $\underline{\underline{G}}$ and $\underline{\underline{m}}$ and turns equation 2-16 into:

$$\underline{\underline{m}}^\dagger = \underline{\underline{d}}^\dagger \left[\underline{\underline{G}}\underline{\underline{G}}^\dagger\right]^{-1} \underline{\underline{G}} \tag{2-18}$$

Substituting the corresponding terms into equation 2-18 yields the MDD equation for the minimum norm solution:

$$\left\{\hat{\tilde{G}}_d^{p,q}\left(\mathbf{x}_B,\mathbf{x}\right)\right\}^* = \int\limits_{S_{\text{source}}} \left[\left\{\hat{G}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S\right)\right\}^* \sum_i \left[\hat{G}^{p,q;\text{in}}\left(\mathbf{x}^{(i)},\mathbf{x}_S\right)\left\{\hat{G}^{p,q;\text{in}}\left(\mathbf{x}^{(i)},\mathbf{x}'_S\right)\right\}^*\right]^{-1} \hat{G}^{p,q;\text{in}}\left(\mathbf{x},\mathbf{x}'_S\right)\right] dS \tag{2-19}$$

The terms in equation 2-19 can be replaced by matrices, and solving is possible by replacing the integral over S_{source} by a sum over the individual sources. Of course the matrix inversion needs to be stabilized in a similar way as equation 2-14. Note that the solution of equation 2-7 via the minimum norm solution is the very similar to the least squares solution, the only difference is that instead of summing the raypaths over array $\mathbf{x}$, they are summarized over the source array $\mathbf{x}_S$.

2-3 Surface-wave responses

Since the single-mode surface waves behave as scalar waves in the surface planes; the scalar MDD scheme as introduced in this chapter can be applied to retrieve single-mode surface wave responses. In the surface plane, a single-mode surface wave behaves very similar to a two dimensional acoustic wave in an infinite space. Two dimensional acoustic waves in an infinite space can be described by Hankel functions. Therefore, in this thesis, we model the single-mode surface waves using the two dimensional infinite space Green's functions; cf. de Hoop (1995). The two dimensional infinite space Green's functions used in the modelling are listed in table 2-2

Two dimensional infinite space Green's functions	Definition of radial distance
$G^{p,q}\left(\mathbf{x},\mathbf{x}_S\right) = i\omega\rho\frac{-i}{4}H_0^{(2)}\left(kr\right)$	$\sqrt{\left(x-x_S\right)^2+\left(y-y_S\right)^2}$
$G_j^{v,q}\left(\mathbf{x},\mathbf{x}_S\right) = i\omega\rho\frac{-i}{4}k\frac{\partial x_j}{r}H_1^{(2)}\left(kr\right)$	$\sqrt{\left(x-x_S\right)^2+\left(y-y_S\right)^2}$
$G_d^{p,q}\left(\mathbf{x}_B,\mathbf{x}\right) = i\omega\rho\frac{-i}{2}kH_1^{(2)}\left(kr\right)$	$\sqrt{\left(x_B-x\right)^2+\left(y_B-y\right)^2}$

Table 2-2: List of all two dimensional infinite space Green's functions used in the modelling

$H_n^{(2)}$ denotes the nth Hankel function of the second kind, k is the wavenumber, defined as: $\frac{\omega}{c}$; r is the radial distance from the source, c is the velocity and ω is the radial frequency. The solutions used here are the standard solutions multiplied with $i\omega\rho$, which follows from the source definition used in (Wapenaar and Fokkema, 2006).

Chapter 3

Feasibility of MDD

for different existing seismological arrays, plus comparison with crosscorrelation results

To check the feasibility of MDD for different existing arrays, a model needs to be constructed that has the same outline as the real seismological array. This model is used for testing the MDD procedure. There is a number of properties that the set-up must fulfil:

- The spatial sampling of the MDD array must be small enough to avoid aliasing, hence $2dx \leq \frac{c_a}{f}$, whereby dx is the receiver spacing, c_a is the apparent wave velocity and f is the frequency,
- The spatial extent of the MDD array must be large enough to gather enough angles for a proper illumination, otherwise there is no improvement with respect to the crosscorrelation method,
- There must be a number of sources present that emit waves in the single frequency band and together create a wavefield that contains enough angles for proper illumination. See chapter 1 for the definition of the single frequency band.

3-1 CANOE set-up

The first set-up to test is the CANOE array, which stands for Canadian Northwest Experiment. See reference CANOE (2005) for some more information of the experiment and the actual location. The model consists of two receiver lines. One line could be seen as the MDD array, S_{rec} sampled by x in equation 2-12, and an auxiliary array, sampled by $\mathbf{x}_B$ in equation 2-12. The goal is to retrieve responses between the MDD and the auxiliary array. The receiver lines have an average spacing of about 50 kilometers. The sources are on a parallel line at a distance of about 400 km from the MDD array, they represent ocean-related single-frequency band noise-sources close to the coast, see chapter 1. However, the sources have a transient

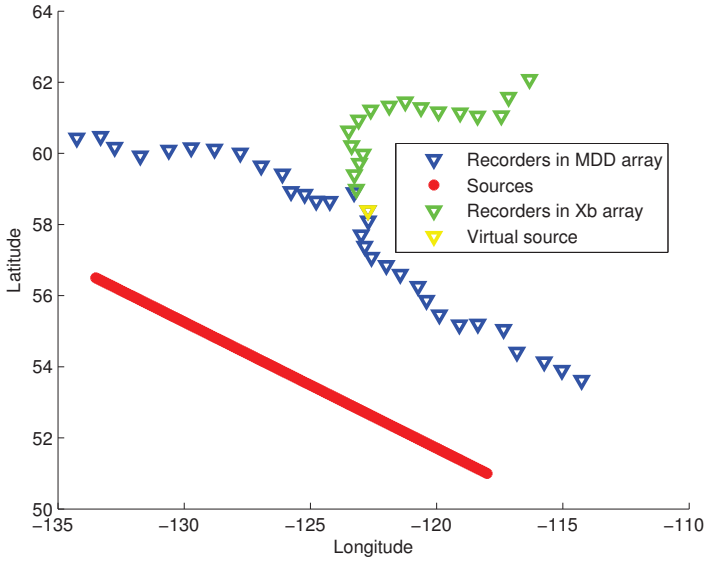
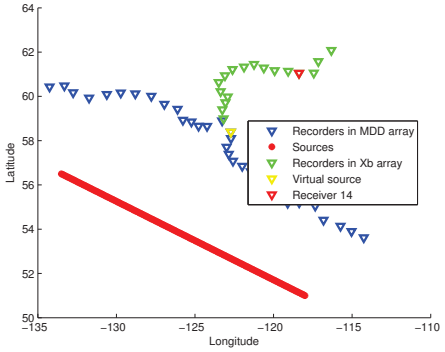
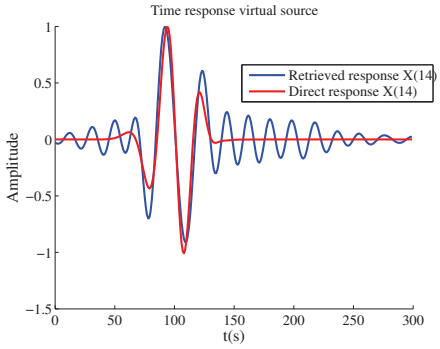


Figure 3-1: Setup of the CANOE receiver lines and the source locations.

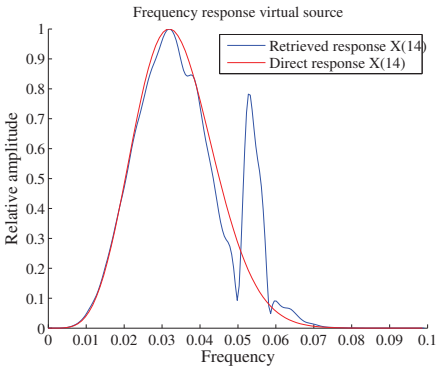
character in the model, since there is no limitation in the source type (chapter 2). See figure 3-1 for an overview of the source and receiver locations. The MDD array, depicted by the blue triangles in figure 3-1, consists of 32 receivers, the auxiliary array, depicted by the green triangles, consists of 17 receivers and here we use 127 sources depicted by the red dots. First a test of the robustness of the model is performed. The retrieved response of the model for the virtual source, in the middle of the MDD array $\mathbf{x}$, to a receiver in the receiver array $\mathbf{x}_B$ is compared to the directly modelled response. This comparison is performed for changes of the source to receiver angles and the spatial extent of the source line. In order to retrieve a good MDD result, the spacing of the MDD array should be small enough to record all frequencies without aliasing. Recall that a signal with a certain wavelength and angle of incidence needs to be measured twice per apparent wavelength in order to be reconstructed properly. The resulting receiver spacing is defined as: $2dx = \frac{c}{f \sin\phi}$. dx is the receiver spacing of the MDD array, ϕ is the angle of incidence compared to the orthogonal of the MDD array, c is the velocity, and f is the frequency. Using this relation, it turns out that the receiver spacing of the MDD array is too coarse. For the upper range of the single frequency band, the maximum allowed angle of incidence is about 30 degrees, which is smaller than the occurring angles of incidence. This could explain why the results in figure 3-2 do not match the response obtained with a real source at the virtual source location. The virtual source location is indicated in figure 3-2a together with receiver 14, which is used to compare the responses. Note that figures 3-2 and 3-3 only show relative amplitudes. The results in figure 3-2 look not very promising. For instance, figure 3-2b shows the time response for a virtual source in the middle of the MDD array to receiver 14 in the auxiliary array; see figure 3-2a for source and receiver location. The retrieved response does somewhat look like the real source response, but is accompanied by high frequency waves before and after the real event. Looking at the same figure in the frequency domain, see figure 3-2c, shows a reasonably retrieved amplitude spectrum, with a sharp frequency peak around 0.055 hertz. Apart from that peak, the frequency content of the response is quite accurate. The virtual source responses for the other receivers in the auxiliary



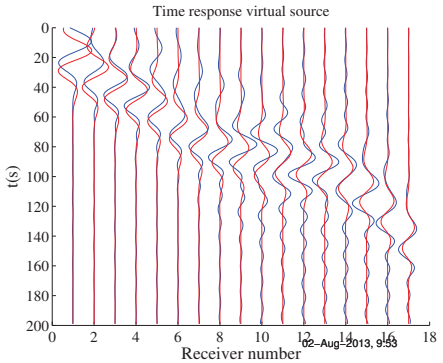
(a) Setup used for figures b, c, d and e. Shows location of virtual source and receiver 14.



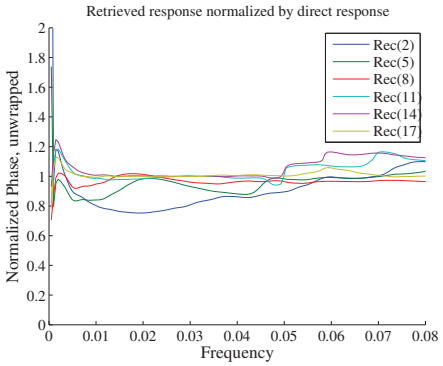
(b) Time response at receiver 14, for the retrieved virtual source response, and the actual source response.



(c) Frequency response at receiver 14, for the retrieved virtual source response, and the actual source response.



(d) Time response at all receivers in the auxiliary array, due to the virtual source in blue, and the actual source in red.



(e) Virtual-source phase-response normalized by the actual-source phase-response. For a few receivers in the auxiliary array.

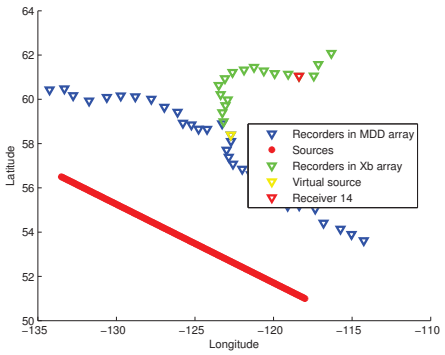
Figure 3-2: MDD results for the CANOE model, with a virtual source in the middle of the MDD array. Note that the actual source, used to compare responses, is located at the same location as the virtual source. Ideally, the results for the virtual source should match the results for the actual source.

array does not show any improvement compared to receiver 14; see figure 3-2d. The virtual source response, in blue, changes per location. For the lower receiver numbers, closer to the source, the high frequency waves are less pronounced; but the dominant response does not match with the real source response. For the higher receiver numbers, further away from the virtual source, the high frequency waves become more dominant. On the other hand, when ignoring the high frequency waves, the dominant response matches reasonably well with the reference response. Figure 3-2e shows the virtual-source phase-response normalized by the real-source phase-response. A value closer to one means a more accurate retrieved phase. The phase accuracy clearly differs per receiver. In general it looks like the overall phase becomes more accurate for the higher receiver numbers, so further away from the source. Note that all receives have inaccurate phase for the very low frequencies, and it tends to get better for increasing frequencies. The highest receiver numbers are fairly accurate for the middle range, and the accuracy decreases again for the higher frequencies. Even for the best response, probably receiver 17, the virtual phase response shows deviations up to 10% compared to the real phase.

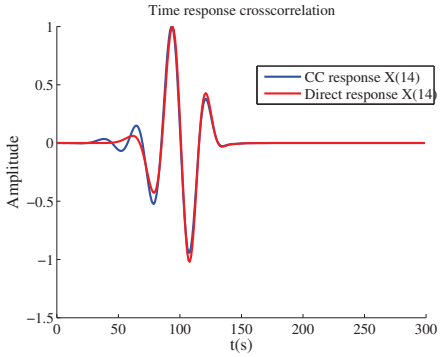
When comparing the MDD result to the seismic interferometry by crosscorrelation result, or CC in short, shown in figure 3-3; it turns out that the CC results are better. Recall from equation 2-6 that the virtual source response for MDD is a dipole function, while all the other responses are monopole. The CC result is therefore compared to a monopole real source located at the same virtual source location.

Compared to the MDD results in figure 3-2, the crosscorrelation results in figure 3-3 look much better. The wavelets in figure 3-3b look very similar. The crosscorrelation response has an extra wiggle, and there are some small amplitude deviations, so the peak amplitude occurs at a different time, but overall it looks very good. A similar story applies to figure 3-3c, there are some small deviations in for instance the peak frequency, and the flanks have a slightly different shape; but the crosscorrelation response has a similar shape, and no additional amplitude peaks, compared to the real source response. In figure 3-3d, the time response for the crosscorrelation response is compared to the real source response. For all traces, the response seems to match without any high frequency waves disturbing the result. Again the reference response is shown in red, and the crosscorrelation result is shown in blue. Note that the whole image is scaled according to one maximum value; apparently the real source response decays faster over the same distance than the crosscorrelation response, which explains the amplitude difference building up for the higher receiver numbers. Figure 3-3e shows the normalized phase response. There are a few differences compared to the MDD phase response in figure 3-2e. First, the overall phase response is more accurate for all the receivers; here also applies the rule that the phases are more accurate for receivers further away. Second, the phase response for the very low frequencies, up to 0.01 hertz, is quite bad; even worse than the MDD phase response. Receiver 2 is remarkable, the phase response is much worse for this receiver compared to the other receivers in figure 3-3e. A likely reason for this could be that, the separation distance between the virtual source and receiver 2 is small; which causes edge effects due to integration over a spatially limited source array. The wavelength of the dominant frequency, 132 kilometer, is longer than the separation distance, 116 kilometer, therefore not part of the energy for that frequency is reconstructed at negative times, which in turn has an influence on the phase response in the frequency domain.

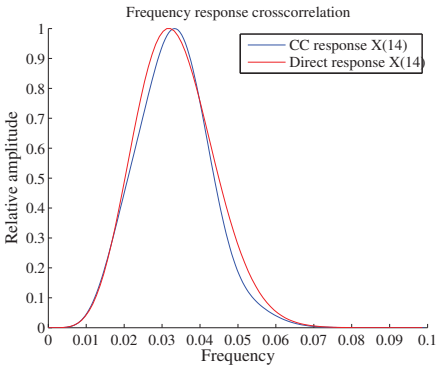
Figure 3-4 shows the retrieved velocity that is recovered from the phase response of the virtual source. The black line is the theoretical velocity which was used as an input into the



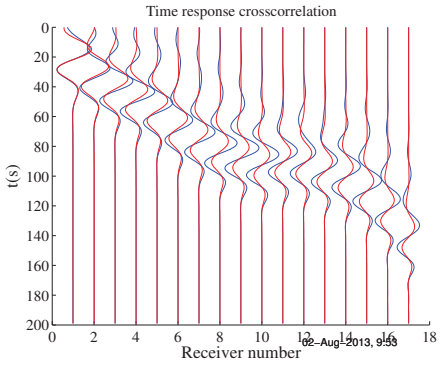
(a) Setup used for figures b, c, d and e. Shows location of virtual source and receiver 14.



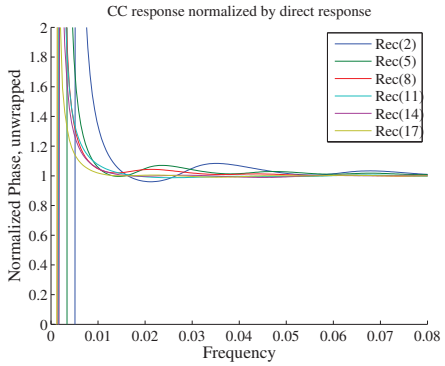
(b) Time response at receiver 14, for the crosscorrelation response, and the actual source response.



(c) Frequency response at receiver 14, for the crosscorrelation response, and the actual source response.



(d) Time response for all receivers in the auxiliary array, due to the crosscorrelation in blue, and the actual source in red.



(e) Crosscorrelated phase-response normalized by the regular-source phase-response. For a few receivers in the auxiliary array.

Figure 3-3: CC results for the CANOE model. The crosscorrelation source is located in the middle of the MDD array, at the virtual source location.

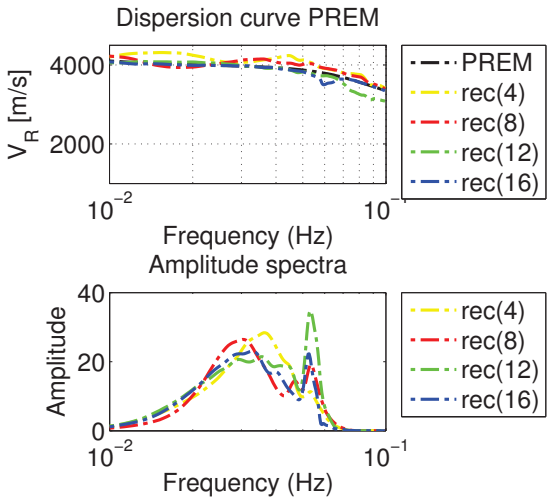


Figure 3-4: Retrieved velocity with the MDD response, second picture shows the frequency spectra.

model. The other colours are the obtained velocities for different receivers, and the same virtual source. The retrieved phase velocity is fairly accurate. There are some deviations, most of them occur where the retrieved frequency varies from the reference response, but the general trend is good. The crosscorrelation result gives a very similar figure; the result is even smoother than the MDD result with less distortions; the only deviations occur for receiver 4 and frequencies up to 0.015 hertz. This means that the far field assumption for phase velocity reconstruction works fine for most frequencies and distances. The dipole far-field phase response is defined in equation 3-1.

$$\phi(\omega) = -\frac{\omega r}{c(\omega)} + \frac{\pi}{4} \tag{3-1}$$

Using the velocities as defined in the PREM model, the phase goes up to about ten cycles for the higher frequencies and larger distances, but goes to zero for the shorter distances. However, the violation of the far-field approximation seems to have a rather small influence on the retrieved velocities.

In short, it seems impossible to retrieve an accurate virtual source response in the CANOE experiment via MDD. Experimenting with the extend of the source line, and the number of MDD receivers taken into account did not improve the results. Changing the extend of the source line could reduce the angle of incidence waves, which in turn should prevent aliasing. The same applies to a smaller MDD array, the angle of incidence from the MDD array to the auxiliary array reduces. However this did not improve the result, especially a smaller MDD array degrades the results quickly. Increasing the size of the source array does not work, because this will only increase the aliasing at the MDD array. For all results, relatively high frequency waves appear in the reconstruction that should not be there. The crosscorrelation results seem to be more accurate.

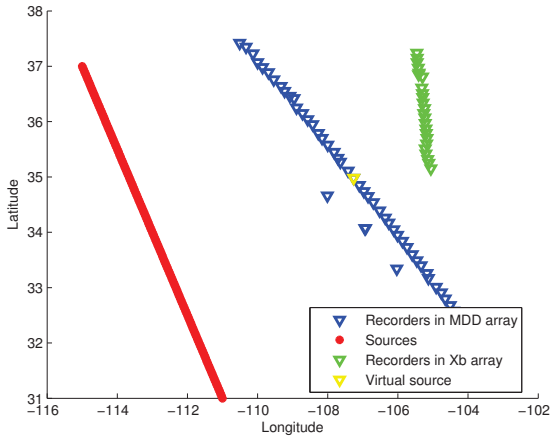
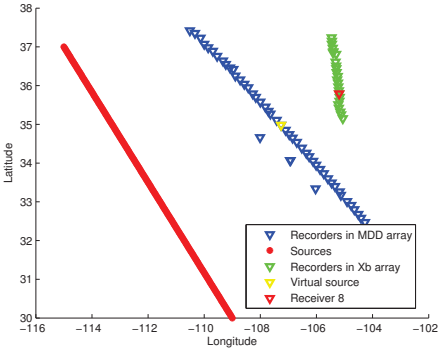


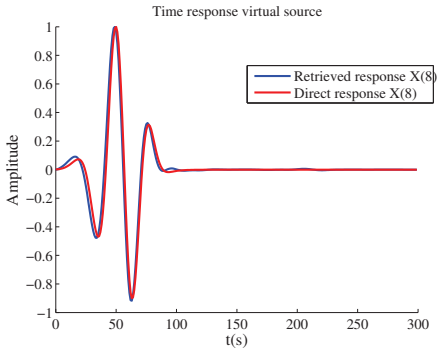
Figure 3-5: Setting of the arrays in New Mexico.

3-2 New Mexico set-up

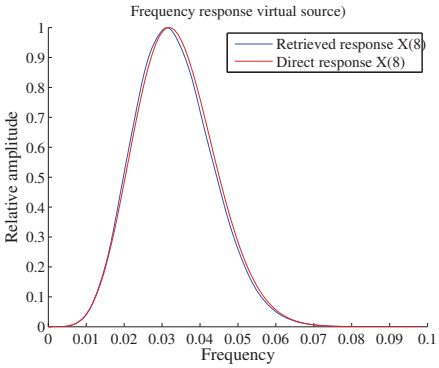
The second set-up to test is in New Mexico and surroundings. A combination of two arrays, the RGST- and CDROM-network, is used in order to create a similar situation with a MDD- and an auxiliary array. See CDROM (2000) and RGST (2000) for more information about the arrays and their location. Again the sources are microseism noise sources off the coast. See figure 3-5 for a picture of the setting. The spacing of the arrays in New Mexico is more favourable compared to the CANOE setting, the receiver spacing is smaller. Therefore, larger angles of incoming waves are registered correctly, and this should improve the MDD response. The New Mexico model contains 25 receivers in the auxiliary array, 58 receivers in the MDD array and 200 sources. The model study of the New Mexico array yields the results shown in figure 3-6. The figures 3-6c,3-6b show the response in receiver 8, indicated in figure 3-6a. The MDD response in time looks much better compared to the MDD result in figure 3-2b. The high frequency disturbances, visible in the CANOE experiment, are not present in the New Mexico setting. The wavelet matches almost perfectly with the reference response. When comparing the frequency spectrum in figure 3-6c with the directly modelled spectrum, there is still a very small difference visible. Compared to the spectrum in figure 3-2c, this time there is no high amplitude peak reconstructed and the spectrum is much smoother. Figure 3-6d shows the virtual source response, in blue, compared to the real source response, in red, for all receivers in the auxiliary array. The quality of the response seems similar for all receivers, all the retrieved wavelets match very well with the reference; for the receivers further away from virtual source, a similar kind of high frequency wavelets appear as observed in the CANOE experiment, but much weaker. Overall the result is very good. Figure 3-7a shows the virtual source phase response normalized by the real source phase response. The result is remarkably better than the retrieved phases in the CANOE experiment, figure 3-2e, and there is no huge difference in the quality of the phase response for the different receivers. Deviations are less than 10 percent, ignoring the first 0.01 hertz. Again the response is more accurate for the higher frequencies. Comparing the phase response in figure 3-7a with the crosscorrelated phase response, in figure 3-7b, shows that the phase response of the MDD procedure is actually more accurate than the crosscorrelation procedure for the low frequencies. Up to



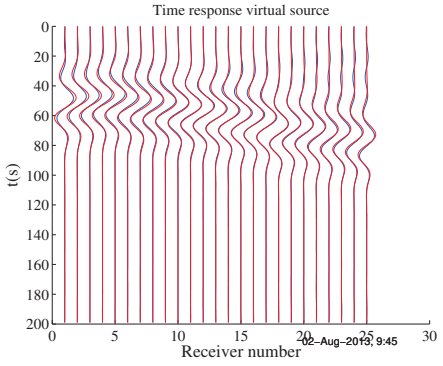
(a) Setup used for figures b, c, d and e. Shows location of virtual source and receiver 8.



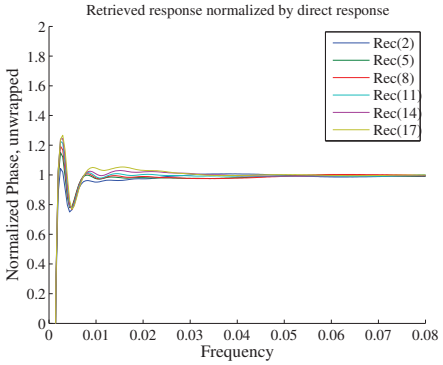
(b) Time response at receiver 8, for the retrieved virtual source response, and the actual source response.



(c) Frequency response at receiver 8, for the retrieved virtual source response, and the actual source response.

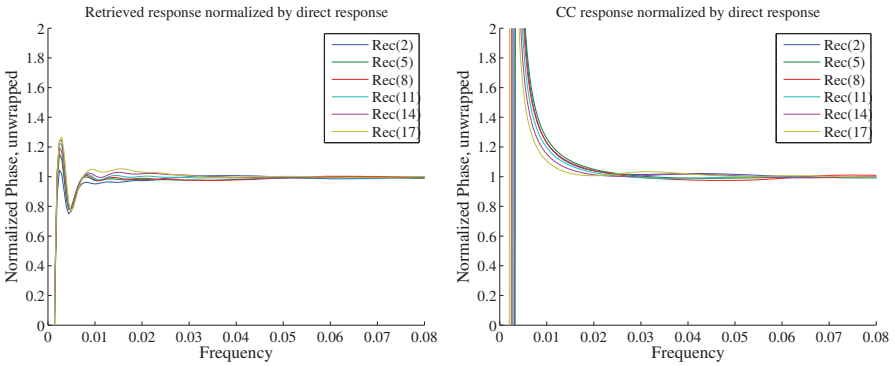


(d) Time response at all receivers in the auxiliary array, due to the virtual source in blue, and the actual source in red.



(e) Virtual-source phase-response normalized by the actual-source phase-response. For a few receivers in the auxiliary array.

Figure 3-6: Results for the New Mexico model, with a virtual source in the middle of the MDD array. Note that the actual source, used to compare responses, is located at the same location as the virtual source.



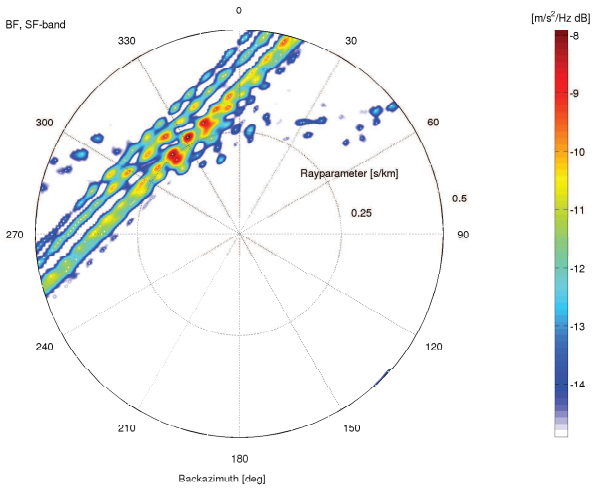
(a) Virtual-source phase-response normalized by the actual-source phase-response. For a few receivers in the auxiliary array.

(b) Crosscorrelated phase-response normalized by the actual-source phase-response. For a few receivers in the auxiliary array.

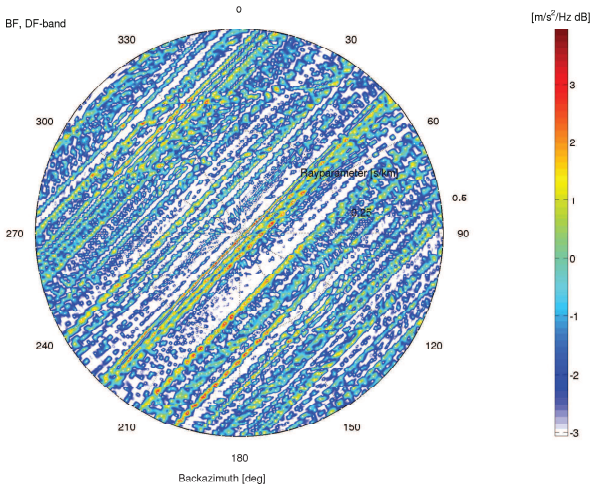
Figure 3-7: Virtual-source phase-response for MDD and Crosscorrelation method.

about 0.02 hertz, the crosscorrelated response is not very accurate, but the MDD response seems accurate from about 0.01 hertz and higher.

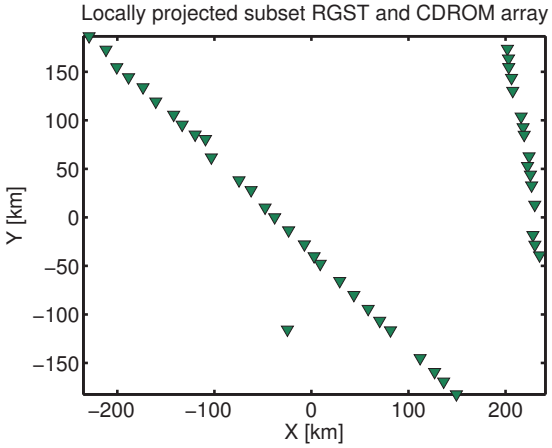
In case that the New Mexico set-up is used for real data, some data processing needs to be done before MDD. At the latitude of the New Mexico array, the north-American continent is not as wide as at latitude of the CANOE experiment. This means that the noise very likely originates from both sides of the continent, however the MDD relation assumes one-sided illumination only. In order to use the New Mexico arrays, it is necessary to single out the preferred sources. Equation 2-12 shows that only $G^{p,q;\text{in}}$ is needed. Furthermore, by selecting source locations, it is possible to prevent angles of incidence that are too high; however the model does not seem to be very sensitive to aliasing. A way to find the angle of incidence of a wave, and the direction of propagation, is with beamforming. Beamforming is described in two basic steps in Ruigrok (2012), more extensive information can be found in for instance Rost and Thomas (2002). The first step is to mutually crosscorrelate all traces. This gives a crosscorrelation matrix which contains all the time differences between all wave-arrivals at different stations. The second step is to fit these time differences with a forward plane-wave model to figure out where the waves originate from. Figure Figure 3-8a shows the result of beamforming for 6.5 minutes of noise in the single-frequency band, recorded on 3 may 2000 at 0:15 AM. Only the receivers above 34 degrees latitude are used, and of course only the working receivers. In the ideal case, the beamforming result should indicate one or two spots where the noise originates from. However, the result in this figure shows a complete band. One of the assumptions in beamforming is that the subsurface does not show any lateral velocity changes. The spatial extent of the array is about 950 kilometer, and over complicated terrain, and therefore the assumption of no lateral velocity variations is probably not valid; this could explain why a complete source band appears instead of individual locations. Figure 3-8b shows the result of beamforming for 6.5 minutes of noise in the double-frequency band, recorded at the same moment as the data in the SF band. The result for the double-frequency is even poorer than the single-frequency result. Noise seems to originate from almost all di-



(a) Beamforming result for about 6.5 minutes of noise in the single-frequency band.



(b) Beamforming result for about 6.5 minutes of noise in the double-frequency band.



(c) Used receivers for beamforming.

Figure 3-8: Beamforming results. Only receivers above 34 degrees latitude, and containing data, are included in the beamforming process.

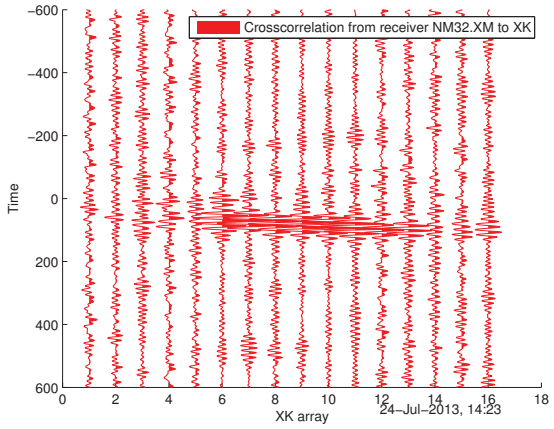


Figure 3-9: Crosscorrelation interferometry result for receiver 32 in the MDD array to auxiliary array, using a full day of noise in the single-frequency band.

rections and arrive with various apparent velocities. When looking at the data, there seems to be no coherent arrivals at the receiver arrays. This is probably due to scattering in the area, therefore it appears that there is no dominant source direction. This could explain the results in figure 3-8b.
Interferometry with crosscorrelation seems to work well for the real data. Since the crosscorrelation method is not hindered by aliasing over the MDD array or noise sources on the wrong side, no data selection needs to take place. Figure 3-9 shows the result of a full day of single-frequency noise crosscorrelation seismic-interferometry. Receiver NM.32 is the same receiver used as a virtual source in the modelling. The result clearly shows a dominant arrival in the causal time domain, and the arrival time matches approximately with the modelled response in figure 3-6b. Exact comparison is not possible, due to unknown local velocity variations in the subsurface and the irregular retrieved waveform on real data.

In short, the modelled results for the New Mexico model are better than the CANOE model. The phase response is much more accurate, even more accurate than the crosscorrelation response; although no parameters were introduced that would make the crosscorrelation inaccurate, for instance losses or an irregular source distribution. The dominant retrieved response matches with the reference response. In order to apply the MDD procedure on real data, beamforming is used to separate desired from undesired signals. In practice, the assumption of no lateral velocity variations is probably violated. Beamforming did not help in selecting the wanted source locations. The crosscorrelation method does not need source selection in order to retrieve a virtual source response, and the result is pretty good. Since the source selection did not work out and the highly scattering environment would cause illumination from both sides, even if only primary illumination from the West could be selected, no MDD was performed on real data.

3-3 USarray set-up

The third and final set-up is located throughout the complete USA. The arrays used are part of the US-array network. See figure 3-10 for an overview of the considered source and

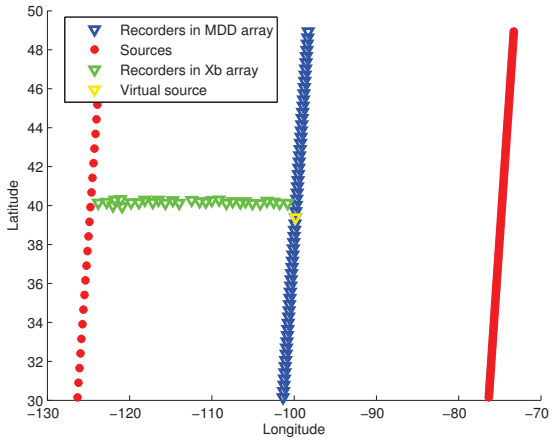


Figure 3-10: Overview of the US array model. Sources located on both sides to simulate noise from both sides of the continent.

receiver locations; see USArray (2009) for more information. Here illumination from both sides is implemented, so the model can be used to test if it is possible to select the desired sources only. The MDD array originally consists of 30 receivers over a distance of about 1000 km, but for modelling purposes the number is increased to 60. This way the results can't be degraded by aliasing. The number of receivers in receiver array $\mathbf{x}_B$ is 31 receivers. The number of sources used is 151 on the correct side, and another 26 on the other side to simulate unwanted input. The MDD array is located halfway the USA, which makes it unlikely time windows can be chosen, with e.g. beamforming, containing only illumination from one side. Instead, with this array we test wavefield-decompositon to obtain one-sided illumination from a recording with two-sided illumination

3-3-1 Wavefield decomposition

Wavefield decomposition is a well known method, see for instance Wapenaar and Berkhout (1989) for more information. We assumed that surface waves can be modelled the same way as acoustic wavefields, see chapter 2; therefore the wavefield decomposition uses a scalar scheme. Furthermore, we assume a laterally invariant medium. Wavefield decomposition splits up the total wavefield in, in this case, a left and a right travelling wavefield. The wavefield decomposition procedure makes use of pressure and velocity data to identify the propagation direction. Equation 3-2 gives the equation for wavefield decomposition in the frequency-wavenumber domain. The tildes denote frequency-wavenumber domain. The Fourier transformation takes place over the y direction, along the arrays, and the arrays are separated from each other in the x direction.

$$\begin{aligned}\tilde{P}^{+}(x) &= \frac{1}{2}\left[\tilde{P}(x) + \frac{\omega\rho}{k_x(x)}\tilde{V}_x(x)\right] \\ \tilde{P}^{-}(x) &= \frac{1}{2}\left[\tilde{P}(x) - \frac{\omega\rho}{k_x(x)}\tilde{V}_x(x)\right]\end{aligned} \tag{3-2}$$

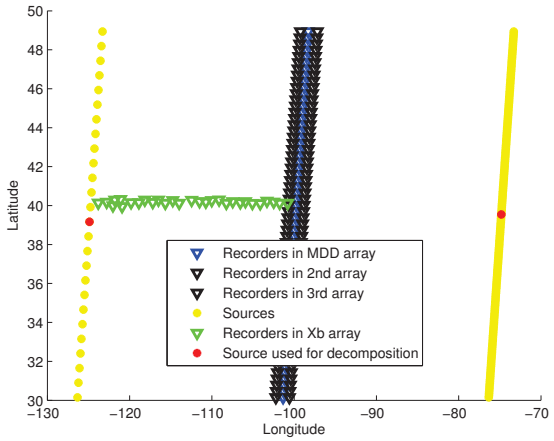


Figure 3-11: Setting with three parallel arrays for estimating the derivatives. The red sources are used for the wavefield decomposition example.

$\tilde{V}_x(x)$ stands for particle velocity in the x direction at location x, and $\tilde{P}(x)$ is the pressure at location x.

Since the real data only consists of particle velocity measurements, only geophone data is collected and pressure is zero per definition at a free surface, the pressure component needs to be replaced by an apparent pressure. The apparent pressure is related to the particle velocity, in the space-frequency domain, via the following equation(Wapenaar and Berkhout (1989)),

$$\frac{\partial \hat{V}_x(x)}{\partial x} + \frac{\partial \hat{V}_y(x)}{\partial y} = \frac{1}{j\omega\rho} k^2 \hat{P}(x). \tag{3-3}$$

Using equations 3-2 and 3-3, the arriving wavefield can be subdivided into two separated wavefields.

Wavefield decomposition is necessary because MDD only uses the ingoing field, see equation 2-12. Equation 3-2 is applied on the model visible in figure 3-10. Both situations are tested, with both pressure and velocity data and with only velocity data; using equation 3-3 next to equation 3-2. In order to retrieve the derivative of the velocity, two extra receiver arrays are placed parallel to the MDD array. See figure 3-11 for an overview of the new situation. The derivatives of $\hat{V}_x$ and $\hat{V}_y$ are obtained via a central derivative scheme:

$$\frac{df\,(x)}{dx}\,(x_i) = \frac{f\,(x_{i+1}) - f\,(x_{i-1})}{2\Delta x} \tag{3-4}$$

A smaller separation distance should yield a more accurate derivative, but matching reality forces a separation distance of approximately 100 kilometer. Figures 3-12,3-13 and 3-14 show the results for the wavefield decomposition with respectively both velocity and pressure data available, and with an apparent pressure approximation. All the images are generated with a source on each side of the arrays, the sources indicated in red in figure 3-11. These source responses are combined to one wavefield. The first figure has a separation distance of only 800 meter between the parallel arrays, the second and third figures a separation distance of 100 kilometers. The first and second figure have a center frequency of 0.03 hertz, and the

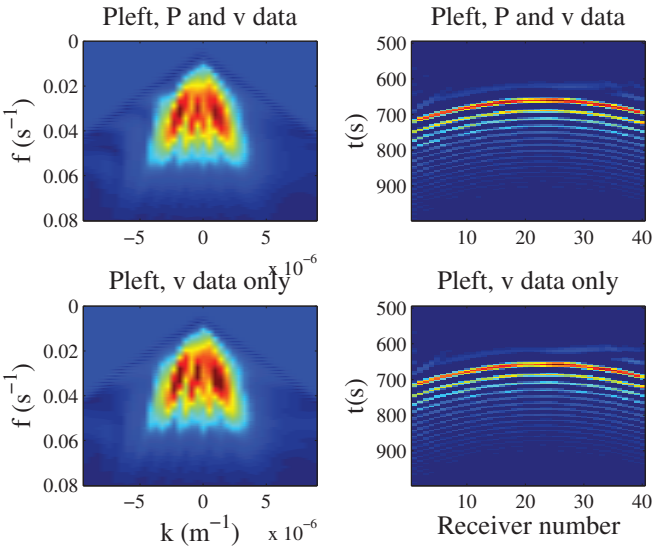


Figure 3-12: This illustration shows the result of the wavefield decomposition in the wavenumber frequency domain on the left, and in the space time domain on the right. The upper two images show the decomposed wavefield for the optimal case where both pressure and particle velocity data are available for the wavefield decomposition. The lower two images show the result where the pressure data is substituted by the derivative of the velocity data. The spacing between the arrays is 800 meter

third image has a center frequency of only 0.015 hertz. The images shown here are with a taper in the wavenumber-frequency domain to remove spurious events that exist due to the strong amplitude differences at the boundary from evanescent to propagating waves, and appear after transforming back to the space-time domain . The taper is a simple cosine squared function in the vicinity of $k_x^2 = (\frac{w}{c})^2$. The images show the left-going decomposed wavefield at the MDD array. Figure 3-12 shows the decomposed result in both space-time and wavenumber-frequency domain for a wavefield travelling from right to left. The two upper images show the result for the ideal case where both particle velocity and pressure data is available on the main array. The two lower images show the result when only particle velocity observations are available on three parallel arrays with a spacing of 0.8 kilometer and the apparent pressure is estimated via the derivative of the velocity. The retrieved wavefield is good in both cases. It clearly extracts the desired direction, and only a few artefacts are visible. Realistic array spacings in the US-array force the separation distance to about 100 kilometer. This time there is a clear difference between the results for both pressure and velocity data, or just velocity data available. The results in figure 3-13 show a big quality change for the wavefield decomposition with velocity data only. This time it is not possible to decompose the wavefields successfully. The result shows two wavefields instead of just one. This is due to the fact that the spacing between the three arrays is now 100 kilometer, which is about equal to the dominant wavelength. This means that the discrete derivative is not that accurate compared to the real derivative at the central array. Therefore it is not possible to accurately retrieve the dominant wave direction for a center frequency of 0.03 hertz with

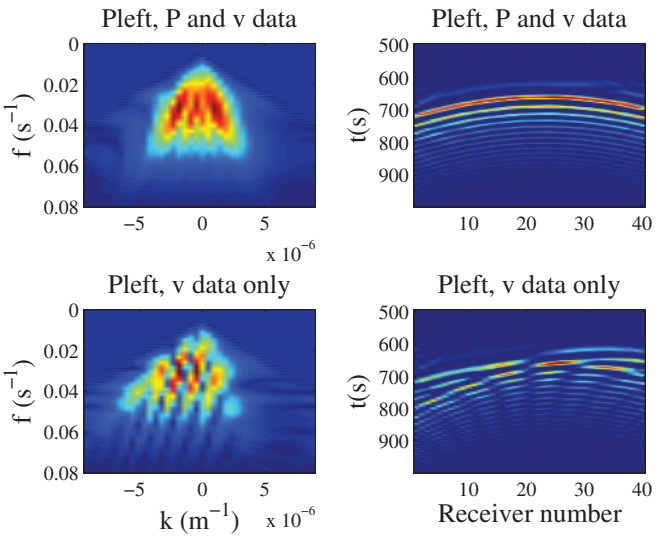


Figure 3-13: This illustration shows the result of the wavefield decomposition in the wavenumber frequency domain on the left, and in the space time domain on the right. The upper two images show the decomposed wavefield for the optimal case where both pressure and particle velocity data are available for the wavefield decomposition. The lower two images show the result where the pressure data is substituted by the derivative of the velocity data. The spacing between the arrays is 100 kilometer

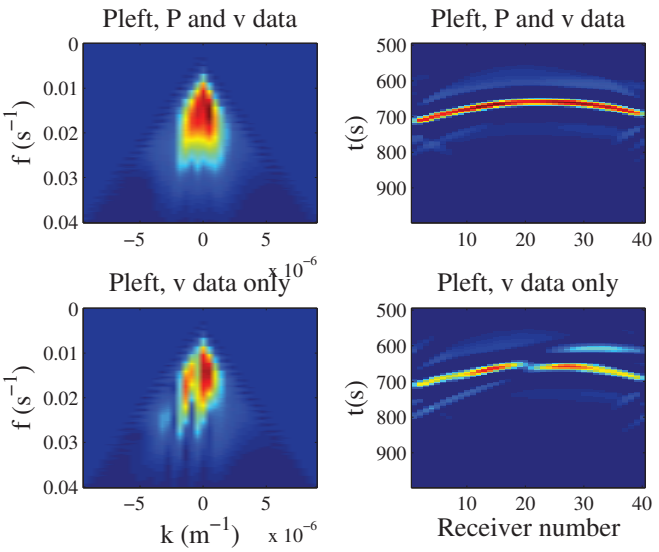


Figure 3-14: Same as the caption in figure 3-13, but with a center frequency of 0.015 hertz instead of 0.030 hertz.

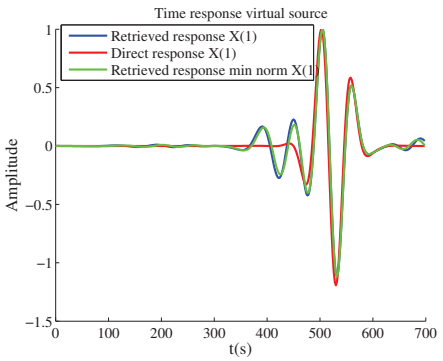
an array spacing of 100 kilometers or more. A possible way out is to increase the dominant wavelength, or reduce the center frequency. Reducing the center frequency from 0.03 to 0.015 hertz improve the results significantly. Figure 3-14 shows the decomposed wavefields. The retrieved wavefield for velocity data only in figure 3-14 looks a lot better compared to the retrieved wavefield in figure 3-13. The longer wavelength allows a more accurate estimate of the derivative of the velocity for an array spacing of 100 kilometer, which results in a better wavefield decomposition. The resulting wavefield in the time-domain can be separated in a dominant direction, which is exactly needed for the MDD procedure.

3-3-2 MDD

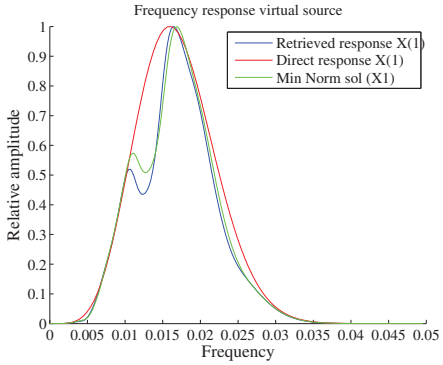
The next step is to use the result of the wavefield decomposition, and use this for the MDD procedure. The decomposed wavefield is disturbed in comparison to the original field. Remnants of the counter travelling field are present, and the taper function used above introduces zeros in the wavenumber-frequency spectrum that were not there before. Using the decomposed wavefield without taper is not possible, the truncation effects distort the result badly. The zeros introduced due to the taper blow up in the MDD procedure and especially the low frequencies are boosted during the MDD procedure. One way to stabilize is to add a small value to the tapered values in the wavenumber-frequency domain, thereby removing the zeros introduced by the taper.

The MDD results for the US array are shown in figure 3-15, here only the ingoing wavefields are used. No decomposition is applied, and only sources on the desired side are active.

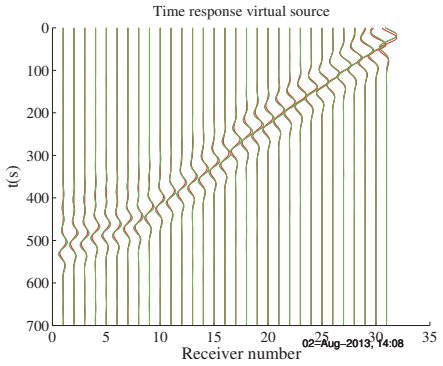
The result in figures 3-15a and 3-15b give a similar result as the New Mexico setting. Note that there is also a minimum norm solution given. The green lines are retrieved by implementing the minimum norm solution described in chapter 2. See for instance equation 2-19. The result



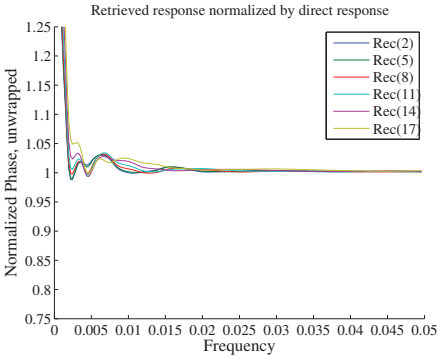
(a) Time response at receiver 1, for the retrieved virtual source response, and the actual source response.



(b) Frequency response at receiver 1, for the retrieved virtual source response, and the actual source response.



(c) Time response for all receivers at the auxiliary array, due to the virtual source in blue, the minimum norm solution in green and the actual source in red.

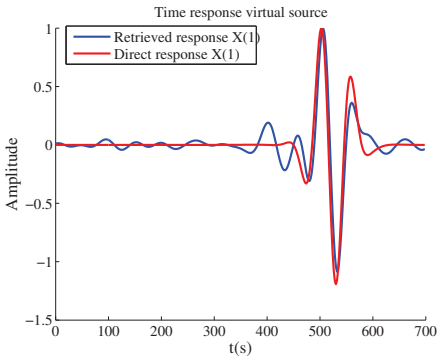


(d) Virtual-source phase-response normalized by the regular-source phase-response. For a few receivers in the auxiliary array.

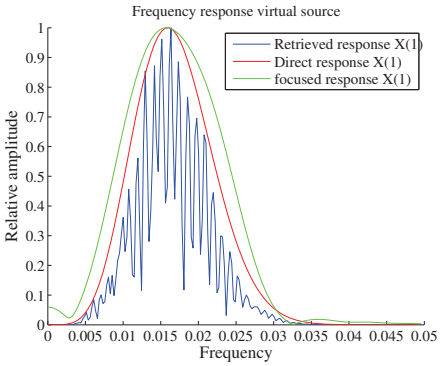
Figure 3-15: Results for the US array model, with a virtual source in the middle of the MDD array. Note that the actual source, used to compare responses, is located at the same location as the virtual source.

is similar to the regular MDD solution by least squares inversion, perhaps the minimum norm solution is even a bit less stable. The regular MDD solution is given in blue, and red is again the reference response. The time response in figure 3-15a shows a very nice match for receiver 1, furthest away from the virtual source, and a virtual source in the middle of the MDD array. Both the minimum norm and least squares MDD scheme retrieve the virtual source response well. The real event is reconstructed almost perfectly, with just a small additional event. In frequency domain, see figure 3-15b, a similar result appears. The retrieved frequency spectrum matches really well, except for the dip around 0.0125 hertz and a small dip around 0.025. The distortions of the frequency spectrum, and the extra wiggles in the time domain, are probably edge effects of the finite source array that overlap with the virtual source response in time. The time response has a similar quality for all receivers in the auxiliary array, see figure 3-15c, the retrieved virtual source response matches in all traces with the real source response. This accurate retrieval also follows from the results in figure 3-15d, all the retrieved phases show deviations of 5 percent or less. Again the response is more accurate for the receivers further away, the lower receiver numbers.
Now the decomposed wavefield is used, as shown in figure 3-14, for the MDD scheme. This means that the three parallel arrays, used for estimating the apparent pressure, are separated by 100 kilometer. This time, there are 26 sources on the wrong side and 151 on the correct side. The results are shown in figure 3-16 for a virtual source in the middle of the array. See for instance the setting in figure 3-10. The results in figure 3-16a look promising, the main event is retrieved fairly well. Unfortunately, there is quite some noise around the retrieved event, and the second peak has a smaller amplitude than the reference wavelet. Also an anti causal event is retrieved, due to not completely removal of the counter travelling wavefield. Therefore, the frequency response in figure 3-16b looks pretty bad initially, but selecting a short time window around the retrieved event, referred to as the focussed response in the images, improves the result a lot; the final frequency spectrum resembles the reference spectrum, although it is not as good as figure 3-15b. Of course the frequency spectrum still contains edge effect, the used virtual source is in the middle, so these overlap in time. More about the edge effect follows in A-1. Looking at the time response for all receivers in the auxiliary array, see figure 3-16c, the response seems similar for all traces, only for the very close receivers the result degrades, but the rest looks fairly similar. When using the time-windowed part for phase analysis, the phase seems to match the reference phase really well, see figure 3-16d. The retrieved phase is not as accurate as for the regular MDD, see figure 3-15d, especially for the higher frequencies it starts to drift off from the reference phase.

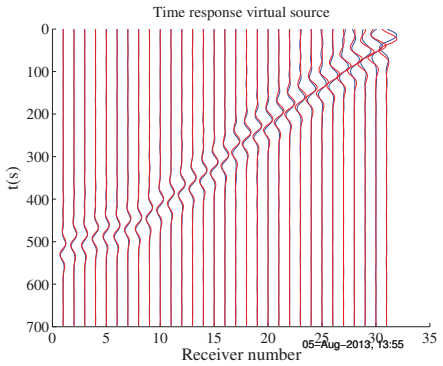
For a fair comparison, the MDD results are shown for the same sources without wavefield decomposition in figure 3-17. The results are remarkable. In figure 3-17a, the retrieved wavelet looks very similar to decomposed wavefield result. The only big difference is that the second peak in the wavelet is now much larger than the reference response, while it was smaller in the decomposed wavefield result. Figure 3-17b shows the result in the frequency domain. Again the frequency of the full trace is shown, and the frequency content of a small window around the retrieved event. This time, it appears that the frequency spectrum is more accurate compared to the frequency spectrum in figure 3-16b. On the other hand, it could be that the edge effect is less pronounced here, and removing it might correct the difference. The response in time is, similar to the decomposed wavefield MDD, constant for all receivers; see figure 3-17c. The result only degrades for the receivers close to the virtual source. The phase response, see figure 3-17d, is very similar to the phase response in the decomposed wavefield.



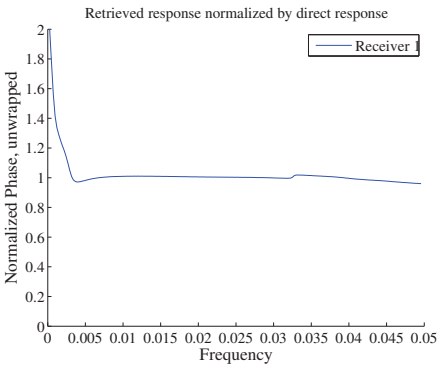
(a) Time response at receiver 1, for the retrieved virtual source response, and the actual source response.



(b) Frequency response at receiver 1, for the retrieved virtual source response, and the actual source response.

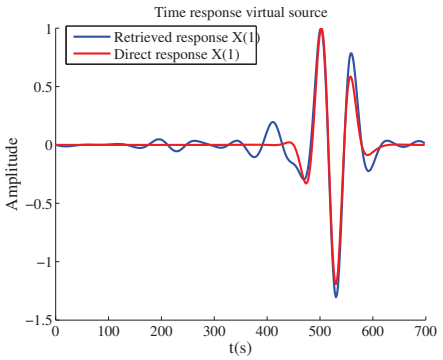


(c) Time response for all receivers at the auxiliary array, due to the virtual source in blue, and the actual source in red.

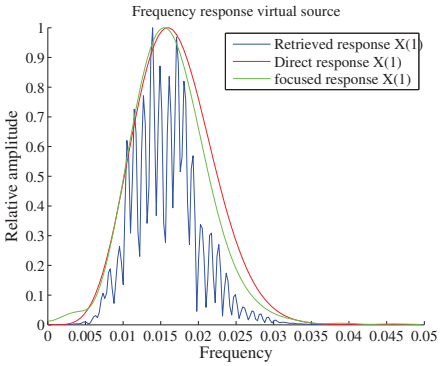


(d) Virtual-source phase-response normalized by the actual-source phase-response. For the last receiver in the auxiliary array.

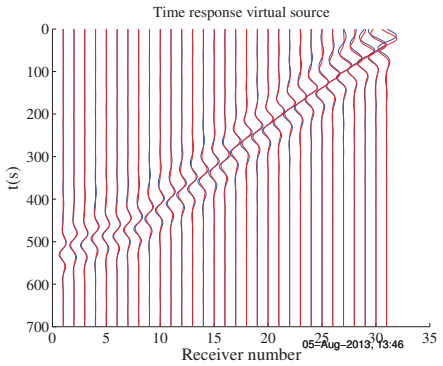
Figure 3-16: Results for the US array model, with a virtual source in the middle of the MDD array. Decomposed wavefields used for MDD procedure.



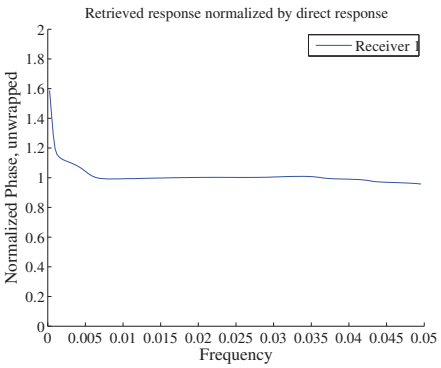
(a) Time response at receiver 1, for the retrieved virtual source response, and the actual source response.



(b) Frequency response at receiver 1, for the retrieved virtual source response, and the actual source response.



(c) Time response for all receivers at the auxiliary array, due to the virtual source in blue, and the actual source in red.



(d) Virtual-source phase-response normalized by the regular-source phase-response. For the last receiver in the auxiliary array.

Figure 3-17: Results for the US array model, with a virtual source in the middle of the MDD array. Full wavefields used for MDD procedure.

Again a time window is used for the response in receiver 1. As before, the result degrades for the higher frequencies. It appears that the result is slightly worse compared to the result for the decomposed wavefield. The MDD results with a better decomposed wavefield improve a bit, the second peak in the time-domain response matches almost perfectly with the reference response; but that requires the three parallel arrays to be separated by only 800 meters, so the optimal situation in figure 3-12 for 0.015 hertz. Even for the optimal wavefield decomposition, time windowing is necessary to improve the result and remove the leaking counter travelling wavefield.

In short, the USarray setup retrieves a decent virtual source response, when only one sided illumination is used. In time, the dominant response is retrieved well; some extra wiggles are present, but relatively small. The phase of the virtual source response matches the real source response. Overall a good result. When starting off with two-sided illumination and using wavefield decomposition to approximate one-sided illumination, the results are not that good. The leaking counter travelling wavefields cause spurious events. Selecting a small timewindow around the retrieved result shows that the retrieved frequency spectrum is quite okay. However, applying no wavefield decomposition gives very similar results in this situation; probably due to the fact that the amount of undesired signal is small compared to the desired signal; here it was 26 versus 151. This might change when the ratio is less favourable. Also, a better wavefield decomposition, so ignoring a realistic setup, shows that the result improves a bit. Still, the regular MDD without wavefield decomposition and just a 'natural' availability of only one-sided illumination performs much better. For the wavefield decompositioin, the derivative of the particle velocity is, in this thesis, estimated with a central derivative scheme. Which is a rather crude method. For instance the observations in the middle array are ignored here. It should be possible to obtain a more elegant derivative scheme, e.q. a two dimensional derivative scheme. When the decomposed wavefield is more accurate, it might improve the one-side MDD scheme with two-sided illumination.

Chapter 4

Total field MDD

The MDD method described in chapter 2 uses only the incoming wavefields. This is due to the assumption that the boundary in the reference state is absorbing. A different way to solve the equations for a virtual source response is to choose different boundary conditions in the reference state. Instead of an absorbing boundary, we choose a reflecting boundary around the reference medium. With this approach, the full wavefields can be used at each receiver and avoid the need of applying wavefield decomposition or beamforming to select the desired observations. The reciprocity theorem of the convolution type, see equation 2-3, is the basis of the MDD theory.

4-1 Injection type sources

As a starting point for the derivation of the full wavefield MDD scheme, we reprint equation 2-4

$$\hat{G}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S\right)=\oint_{\partial V}\left[\hat{G}^{p,q}\left(\mathbf{x},\mathbf{x}_S\right)n_j\hat{\tilde{G}}_j^{v,q}\left(\mathbf{x},\mathbf{x}_B\right)-\hat{\tilde{G}}^{p,q}\left(\mathbf{x},\mathbf{x}_B\right)n_j\hat{G}_j^{v,q}\left(\mathbf{x},\mathbf{x}_S\right)\right]dS \tag{4-1}$$

Where $\hat{G}$ denotes the Greens function in the true medium, and $\hat{\tilde{G}}$ denotes the Green's function in the reference medium. $\mathbf{x}_S$ are source locations in the true medium. $n_j\hat{\tilde{G}}_j^{v,q}\left(\mathbf{x},\mathbf{x}_B\right)$ is the projected velocity on the unitvector normal to the boundary ∂V at location $\mathbf{x}$ due to sources in $\mathbf{x}_B$; $\hat{G}_j^{v,q}$ can be rewritten to $\frac{-1}{i\omega\rho(x_B)}\partial_j^B G^{p,q}$. Often, a well sampled, closed, boundary ∂V is not available. Next, we assume either a Neumann or Dirichlet boundary condition, pressure or velocity zero at boundary, for the reference state in order to simplify equation 4-1 so that one term remains under the integral; then, the remaining equation can be solved by using MDD. For two dimensional in-plane modelled surface waves, the assumption that the apparent pressure is zero implies that $\partial_x v_x=-\partial_y v_y$, see equation 3-3, at the boundary ∂V. This seems a bit odd, therefore $v_j n_j=0$ at the boundary is preferred; in other words, the normal particle velocity is zero at the boundary. This boundary condition is only applied at the sampled part of ∂V, or S_{rec}; for the infinite part, the non sampled part, the contribution of the integral disappears due to the radiation condition. Using the Dirichlet boundary condition in equation

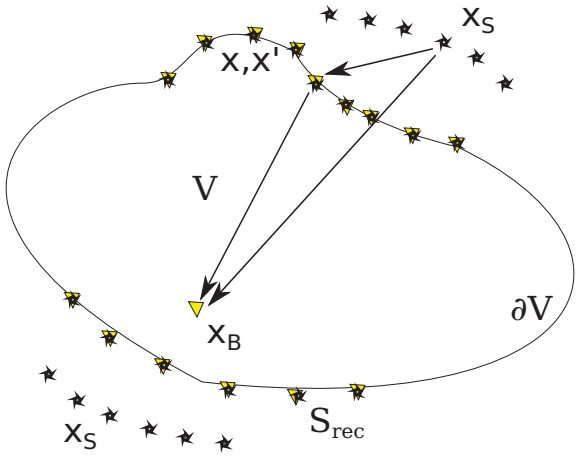


Figure 4-1: Illustration of Green's functions and the sources and receivers. ∂V is a enclosing surface around x_B, and S_{rec} is the well sampled part of ∂V with receivers. There is a source outside V and now a receiver at location x_B inside V. Furthermore there is now also a well distributed source distribution at S_{rec}.

4-1 gives the following equation:

$$\hat{G}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S\right) = -\oint_{S_{\text{rec}}}\left[\hat{\bar{G}}^{p,q}\left(\mathbf{x},\mathbf{x}_B\right)n_j\hat{G}_j^{v,q}\left(\mathbf{x},\mathbf{x}_S\right)\right]dS \tag{4-2}$$

For this application, it is necessary to apply the source receiver reciprocity theorem which allows to interchange the source and receiver locations. We use $\hat{\bar{G}}^{p,q}\left(\mathbf{x}_B,\mathbf{x}\right) = \hat{\bar{G}}^{p,q}\left(\mathbf{x},\mathbf{x}_B\right)$, see for instance de Hoop (1995) for a reference. Applying the reciprocity relation on the Green's functions in the reference medium changes equation 4-2 to:

$$\hat{G}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S\right) = -\oint_{S_{\text{rec}}}\left[\hat{\bar{G}}^{p,q}\left(\mathbf{x}_B,\mathbf{x}\right)n_j\hat{G}_j^{v,q}\left(\mathbf{x},\mathbf{x}_S\right)\right]dS \tag{4-3}$$

This equation is similar to 4-1; now for the pressure observations in the reference medium, the sources are at the boundary of the medium and not inside the medium. Figure 4-1 illustrates the Green's functions in 4-3, the subspace S_{rec} and the additional coordinate vector used later in this chapter.

Equation 4-3 consists of just one term under the integral with the unknown virtual source response given by: $\hat{\bar{G}}^{p,q}\left(\mathbf{x}_B,\mathbf{x}\right)$. Next, equation 4-3 is not ideal, because it contains pressure and velocity observations: $\hat{G}^{p,q}\left(\mathbf{x}_B,\mathbf{x}_S\right)$ and $\hat{G}_j^{v,q}\left(\mathbf{x},\mathbf{x}_s\right)$. However, the apparent pressure cannot be recorded; multiplying both sides by $\frac{-1}{i\omega\rho(\mathbf{x}_B)}\partial_j^B$ turns the apparent pressure measurements into velocity measurements. Equation 4-3 turns into:

$$\hat{G}_j^{v,q}\left(\mathbf{x}_B,\mathbf{x}_S^{(i)}\right) = -\oint_{S_{\text{rec}}}\hat{\bar{G}}_j^{v,q}\left(\mathbf{x}_B,\mathbf{x}\right)\hat{G}_{\perp}^{v,q}\left(\mathbf{x},\mathbf{x}_S^{(i)}\right)dS \tag{4-4}$$

State A	State B
$\hat{f}_{j,A}(\mathbf{x}) = \delta(\mathbf{x}-\mathbf{x}_S)\,\delta_{jp}$	$\hat{f}_{j,B}(\mathbf{x}) = \delta(\mathbf{x}-\mathbf{x}_B)\,\delta_{jr}$
$\hat{p}_A(\mathbf{x}) = \hat{G}_p^{p,f}(\mathbf{x},\mathbf{x}_S)$	$\hat{p}_B(\mathbf{x}) = \hat{\tilde{G}}_r^{p,f}(\mathbf{x},\mathbf{x}_B)$
$\hat{v}_{j,A}(\mathbf{x}) = \hat{G}_{j,p}^{v,f}(\mathbf{x},\mathbf{x}_S)$	$\hat{v}_{j,B}(\mathbf{x}) = \hat{\tilde{G}}_{j,r}^{v,f}(\mathbf{x},\mathbf{x}_B)$

Table 4-1: This table shows the definition of the different terms

Where $\hat{G}_{\perp}^{v,q}\left(\mathbf{x},\mathbf{x}_S^{(i)}\right) = n_j\hat{G}_j^{v,q}\left(\mathbf{x},\mathbf{x}_S^{(i)}\right)$; Equation 4-4 holds for each of the sources, therefore $\mathbf{x}_S$ is replaced by $\mathbf{x}_S^{(i)}$. Next, multiplying both sides by the observations at ∂V by $\hat{G}_{\perp}^{v,q}(\mathbf{x}',\mathbf{x}_S)^*$, whereby $\mathbf{x}'$ is an additional coordinate vector over S_{rec}.

$$\sum_i\left[\hat{G}_j^{v,q}\left(\mathbf{x}_B,\mathbf{x}_S^{(i)}\right)\left(\hat{G}_{\perp}^{v,q}\left(\mathbf{x}',\mathbf{x}_S^{(i)}\right)\right)^*\right] = \oint_{S_{\text{rec}}}\hat{\tilde{G}}_j^{v,q}(\mathbf{x}_B,\mathbf{x})\sum_i\left[-\hat{G}_{\perp}^{v,q}\left(\mathbf{x},\mathbf{x}_S^{(i)}\right)\left(\hat{G}_{\perp}^{v,q}\left(\mathbf{x}',\mathbf{x}_S^{(i)}\right)\right)^*\right]dS \tag{4-5}$$

Equation 4-5 can be regrouped to a similar equation as 2-12 with a crosscorrelation and pointspread function:

$$\hat{C}_j(\mathbf{x}_B,\mathbf{x}') = \oint_{S_{\text{rec}}}\hat{\tilde{G}}_j^{v,q}(\mathbf{x}_B,\mathbf{x})\,\hat{\Gamma}(\mathbf{x},\mathbf{x}')\,dS \tag{4-6}$$

The main difference, with respect to 2-12, is that 4-6 contains only full wavefields, and all the pressure observations are replaced by velocity observations. In a similar way as in equation 2-13, the integral can be replaced by a summation given that the sampling is small enough.

4-2 Force sources

A similar derivation exists starting form equation 2-3 and using the impulsive sources of the force type. Table 4-1 shows the relevant substitutions for the pressure, velocity and impulsive sources of the force type. A is again the actual medium, and B is the reference medium. The subscripts p and r in the source functions and Green's functions indicate direction of forces. For surface waves, p and r are both in the horizontal plane. Again p is the pressure observation and v is the velocity observation. $\hat{G}_{j,p}^{v,f}(\mathbf{x}_A,\mathbf{x}_B)$ denotes observation v in j-direction at $\mathbf{x}_A$ due to a source f in p direction at location $\mathbf{x}_B$. Assuming that only impulsive sources of the force type exits. Substituting the terms in 4-1 into equation 2-3 gives:

$$-\hat{G}_{r,p}^{v,f}(\mathbf{x}_B,\mathbf{x}_S) = \oint_{\partial V}\left[\hat{G}_p^{p,f}(\mathbf{x},\mathbf{x}_S)\,n_j\hat{\tilde{G}}_{j,r}^{v,f}(\mathbf{x},\mathbf{x}_B) - n_j\hat{G}_{j,p}^{v,f}(\mathbf{x},\mathbf{x}_S)\,\hat{\tilde{G}}_r^{p,f}(\mathbf{x},\mathbf{x}_B)\right]dS \tag{4-7}$$

Equation 4-7 contains two terms under the integral and sources both inside and outside volume V. For the reference medium, it is possible to choose a boundary condition. We choose to use the Dirichlet boundary condition, $v_j n_j = 0$ at ∂V, on the reference medium to simplify equation 4-7. The normal velocity at ∂V is zero; hence, the observation $\hat{\tilde{G}}_{j,r}^{v,f}(\mathbf{x},\mathbf{x}_B)$ is zero. This boundary condition is only applied at the finite part of ∂V; for the infinite part,

contribution vanishes due to the radiation condition.

$$\hat{G}^{v,f}_{r,p}\left(\mathbf{x}_B,\mathbf{x}_S\right) = -\oint_{S_{\text{rec}}}\left[\hat{\bar{G}}^{v,q}_{r}\left(\mathbf{x},\mathbf{x}_B\right)n_j\hat{G}^{v,f}_{j,p}\left(\mathbf{x},\mathbf{x}_S\right)\right]dS \tag{4-8}$$

Applying the source-receiver reciprocity relations defined by de Hoop (1995), $-\hat{\bar{G}}^{v,q}_{j}\left(\mathbf{x}_B,\mathbf{x}\right) = \hat{\bar{G}}^{p,f}_{j}\left(\mathbf{x},\mathbf{x}_B\right)$, yields the following equation:

$$\hat{G}^{v,f}_{r,p}\left(\mathbf{x}_B,\mathbf{x}_S\right) = -\oint_{S_{\text{rec}}}\left[\hat{\bar{G}}^{v,q}_{r}\left(\mathbf{x}_B,\mathbf{x}\right)n_j\hat{G}^{v,f}_{j,p}\left(\mathbf{x},\mathbf{x}_S\right)\right]dS \tag{4-9}$$

Equation 4-9 contains different observations compared to 4-4, but the same propagator $\hat{\bar{G}}^{v,q}_{r}\left(\mathbf{x}_B,\mathbf{x}\right)$. Again only one term remains under the integral, and solving is possible. Correlating both sides with $\hat{G}^{v,f}_{\perp,p}\left(\mathbf{x},\mathbf{x}_S\right)^*$, which is defined as: $n_j\hat{G}^{v,f}_{j,p}\left(\mathbf{x},\mathbf{x}_S\right)^*$, yields:

$$\sum_i \hat{G}^{v,f}_{r,p}\left(\mathbf{x}_B,\mathbf{x}^{(i)}_S\right)\left[\hat{G}^{v,f}_{\perp,p}\left(\mathbf{x}',\mathbf{x}^{(i)}_S\right)\right]^* = \oint_{S_{\text{rec}}}\left[\hat{\bar{G}}^{v,q}_{r}\left(\mathbf{x}_B,\mathbf{x}\right)\sum_i\left(-1\right)\hat{G}^{v,f}_{\perp,p}\left(\mathbf{x},\mathbf{x}^{(i)}_S\right)\left[\hat{G}^{v,f}_{\perp,p}\left(\mathbf{x}',\mathbf{x}^{(i)}_S\right)\right]^*\right]dS \tag{4-10}$$

Finally equation 4-10 is the equivalent of equation 4-5 for force type functions. Introducing a pointspread and crosscorrelation function gives:

$$\hat{C}_r\left(\mathbf{x}_B,\mathbf{x}\right) = \oint_{S_{\text{rec}}}\left[\hat{\bar{G}}^{v,q}_{r}\left(\mathbf{x}_B,\mathbf{x}\right)\hat{\Gamma}_p\left(\mathbf{x},\mathbf{x}'\right)\right]dS \tag{4-11}$$

Equation 4-10 is a bit more complicated compared to 4-5, because it requires normal particle velocity observations due to two component force sources, while equation 4-5 only uses normal particle velocity observations due to volume injection sources. It is possible to use either type of sources for full waveform MDD, but equation 4-5 requires less measurements to be able to solve it. Therefore it seems best to use the volume injection source for the model to test the theory.

Chapter 5

Feasibility of MDD using total fields

In chapter 4 a variation on the MDD theory is explained. Since it has not been tested before, it is interesting to test for instance what happens in a one dimensional setting with a scatterer present. Does it work, and what kind of multiples will be reconstructed? Another option is what will happen in a two dimensional situation similar to the ones tested for the regular MDD. For the one dimensional setting, the homogeneous case is treated with and without a point-scatterer. Due to the determinant of the pointspread matrix going to zero for one-sided illumination, only two sided illumination is tested. Since the equation for the volume injection source is easier compared to the force rate source, equation 4-6 is implemented. The Green's functions required are similar to the Green's functions used in chapter 3 and are listed in chapter 2.

5-1 One dimensional setting

A one dimensional setting is fairly simple. Only one source is needed for proper illumination on one side, and only one receiver is needed to create a deconvolution array. The setup in this experiment, see figure 5-1, consists of two sources, in red, on opposite sides of the medium, two deconvolution receivers, in blue and black, on the inside and finally a scatterer, in yellow, somewhere in the middle. The results shown in figure 5-2, are retrieved at the receiver array, in green, in the middle. Note that the obtained results are for one-dimensional Green's functions. The full wavefield MDD scheme retrieves the direct wave plus the multiples since the reference medium is bounded by a rigid boundary. In figure 5-1 there is a source on both sides of the MDD receivers, which means energy travelling from and to the MDD receiver inside V, therefore it should be possible to retrieve the causal and anti-causal wavefield. First, the wavefield is shown for a homogeneous medium without a scatterer in figure 5-2a and figure 5-2b shows the same wavefield but this time with a scatterer inside V. Both images are for a virtual source at the MDD receiver on the right. Figure 5-2a shows a symmetrical picture. The direct wave travels away from the source on the right side in both the causal and anti-causal direction, then hits the boundary and reflects back; after that, it hits the next boundary and reflects again. Figure 5-2b shows the same principle as in figure 5-2a, except that now the same happens as well for the reflected wave due to the scatterer. This creates a lot more multiples, and the resulting figure is crowded. This clearly shows that both boundaries in the

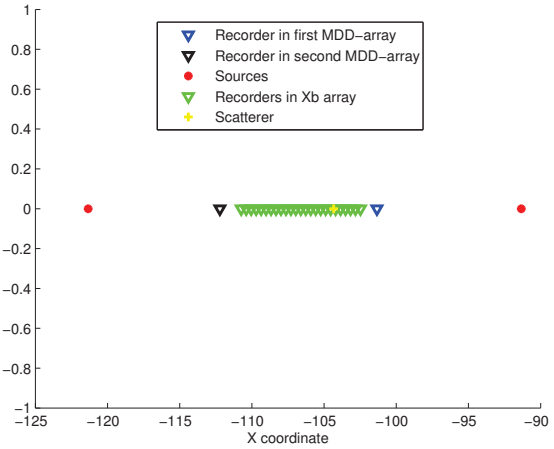
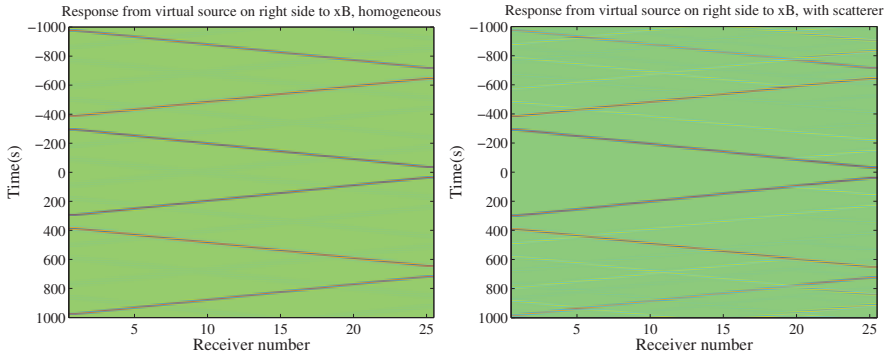


Figure 5-1: One dimensional setting for the total wavefield MDD procedure.



(a) Full wavefield MDD result for two sided illumination in a one dimensional setting. Causal and anti-causal result for a virtual source on the right side.

(b) Full wavefield MDD result for two sided illumination in a one dimensional setting with a scatterer. Causal and anti-causal result for a virtual source on the right side.

Figure 5-2: Retrieved result at center array due to virtual source on the right, with and without a point scatterer

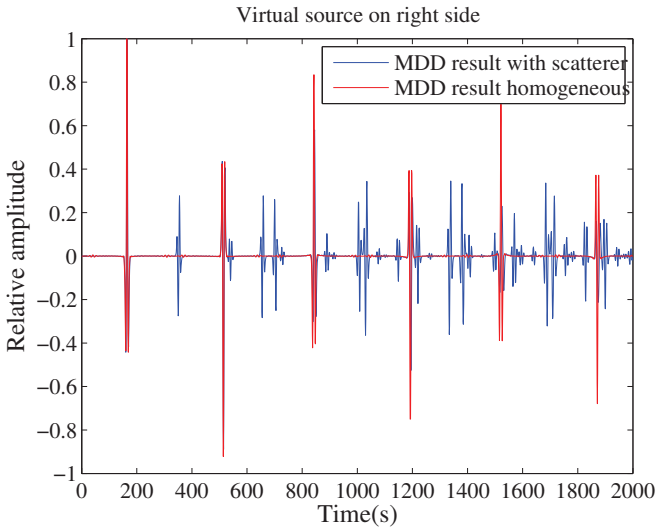


Figure 5-3: Arrivals for the receiver in the middle, in figure 5-2a and 5-2b

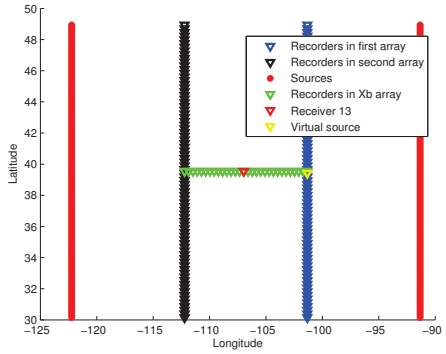
reference state act as reflectors, for both the primary and the scattered arrivals. Note that the time axis are longer than displayed in figures 5-2a and 5-2b in order to reduce the wraparound. If the time axis is long enough, about 16 times longer than displayed, the reflections get weak enough to not visibly wrap around. Figure 5-3 shows the arrivals for the middle trace in figure 5-2a and 5-2b. The image shows the arrivals for the homogeneous situation and with a scatterer present. The response for the medium with scatterer matches the response for the homogeneous medium, plus the additional reflections caused by the scatterer.

Using one-sided illumination would be interesting in order to compare the results with results for the regular one-dimensional MDD scheme; and possible gives a results with less reflections. However, one-sided illumination gives a pointspread function with a determinant that is zero; therefore it is impossible to retrieve the desired Green's function.

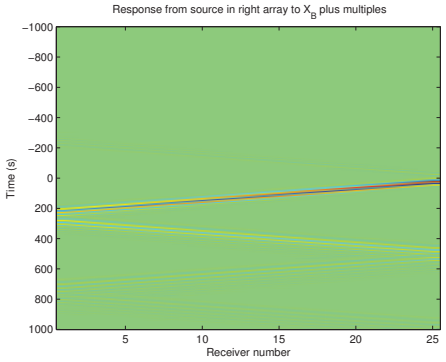
5-2 Two dimensional setting

The results in the one-dimensional setting look fine. So now the theory can be tested in a two-dimensional situation. The situation for a two-dimensional test is very similar to the MDD models in chapter 3, the biggest change is that a second MDD array is needed on the other side of the auxiliary array. Furthermore the frequency spectrum of the input signal is increased to 0.04 hertz to match the original single frequency spectrum, and consequently the spacing of the MDD array and the source line is smaller in order to reduce aliasing. Figure 5-4a shows the setting that is used. Both source arrays consist of 151 sources; both MDD arrays consist of 80 receivers and the auxiliary array consists of 10 receivers.

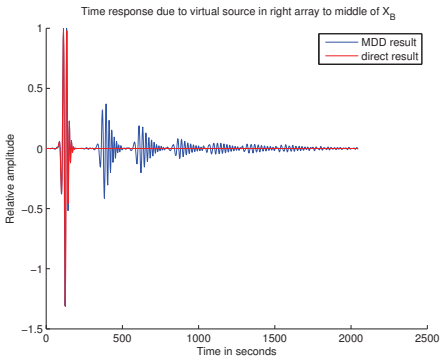
The result in figure 5-4b shows that primarily the causal result is retrieved. The anti-causal part is much weaker compared to the results in the one dimensional situation, see figure 5-2a. Again, the retrieved result includes reflections from the boundaries, which are the MDD arrays, and a long time axis is needed to prevent strong wraparound effects. Taking



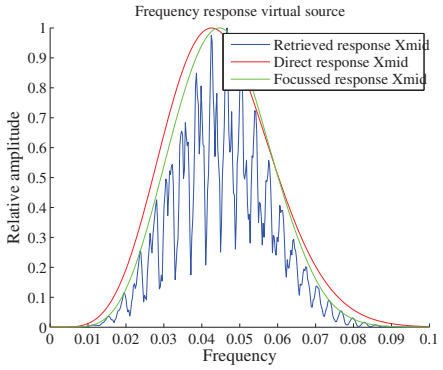
(a) Setup used in 2D situation with virtual source 40 highlighted



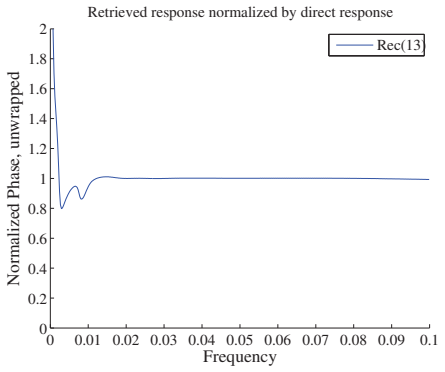
(b) Full wavefield MDD result for two sided illumination in a two dimensional setting.



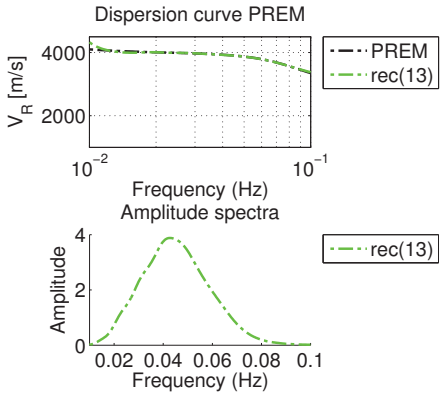
(c) Time response for receiver 13 in the auxiliary array due to virtual source 40



(d) Frequency response for receiver 13 in the auxiliary array due to virtual source 40



(e) Retrieved phase-response normalized by reference phase-response

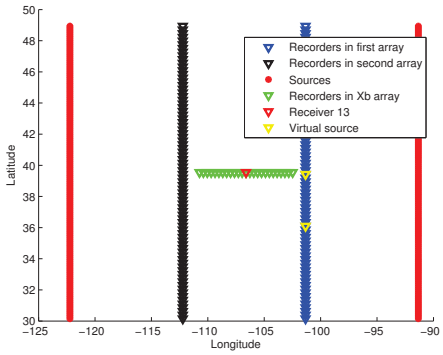


(f) Retrieved phase velocity

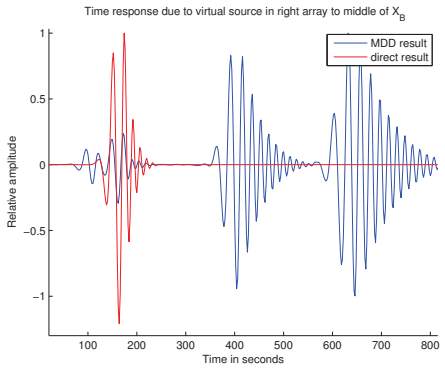
Figure 5-4: Retrieved result at center array due to virtual source in the middle of the right array.

the results for the middle trace in the auxiliary array, receiver 13, gives figure 5-4c. The first arrival overlaps perfectly with the reference arrival, but is followed by a large number of reflections. Selecting a small time window over the first arrival, and transforming that to the frequency domain, gives figure 5-4d. The blue line is the frequency content of the whole trace, and the green line is the frequency content of the small window. The result is not yet perfect, but it looks good. The phase response of that window normalized by the reference response is shown in figure 5-4e. The phase response for receiver 13 shows almost no differences with the reference phase. Finally, the retrieved phase is used to compare the retrieved phase velocity with the input velocity model, see figure 5-4f. Again, for receiver 13 the results are good; the retrieved phase velocity is very close to the input velocity.
In the current setup, the MDD result for full wavefields is very sensitive to the location of the virtual source. Moving the virtual source away from the center of the MDD array yields the opportunity to separate the edge effect in regular MDD, see figure A-3, for the fullwave form MDD this does not apply. The first arrival is very sensitive to the location of the virtual source. The multiples appear to be more consistent. See figure 5-5 for the effect of changing the virtual source from 40 to 26. Figure 5-5a shows the location of virtual source 26 and 40. Figures 5-5b and 5-5c compare the time response in receiver 13 in the auxiliary array for virtual source 26 and 40. The main difference between the responses is that for virtual source 40 the strongest retrieval the direct arrival is, but for virtual source 26 the strongest retrieval is the second reflection. The frequency content of the arrivals in figure 5-5b is different. Figures 5-5d and 5-5e show the frequency content of respectively the direct arrival and second reflection. The blue line is again the frequency content of the whole trace, the green line contains the frequency content of either the direct arrival, or the second reflection; the selection takes place with a time window around the arrival. The frequency content of the first arrival does not match at all with the reference response. However, when looking at the first arrival in the time-domain, there is a wiggle visible before the expected event. This looks very similar to the edge effect, briefly treated in A-1, but it apparently overlaps with the retrieved response and can't be separated. The frequency content of the third arrival looks much better. The frequency spectrum is smooth and almost similar to the reference frequency content. The quality is comparable to the results in figure 5-4d for virtual source 40.
Increasing the distance between the two MDD arrays, and consequently moving the receiver line as well, drastically improves the results for the virtual source at receiver 26. See figure 5-6a for the used setup. Note that the distance between the virtual source and the auxiliary array also increased. With this setup, the retrieved wavelet fits the reference response much better, and the amplitude of the first arrival is larger than before. See figure 5-6b for the comparison of the first retrieved arrival to the reference response. In the frequency domain, see figure 5-6c, the first arrival looks very similar to the reference response. Again, when selecting the correct time window, the retrieved phase and phase velocity match the reference response and velocity model.

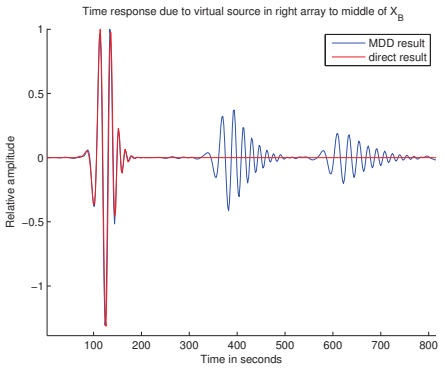
In short, the full wavefield MDD scheme seems to work. The direct arrival and the corresponding multiples are retrieved. The time response matches very well with the reference response, and also in the frequency domain a decent result is obtained when the multiples are removed. This also applies to the retrieved phase, and corresponding phase velocities. Sensitivity to the edge effects depends on virtual source position in the MDD array, and the distance between the receiver considered and the MDD array. For a receiver further away, the sensitivity decreases and a decent first arrival can be retrieved. In the examples considered,



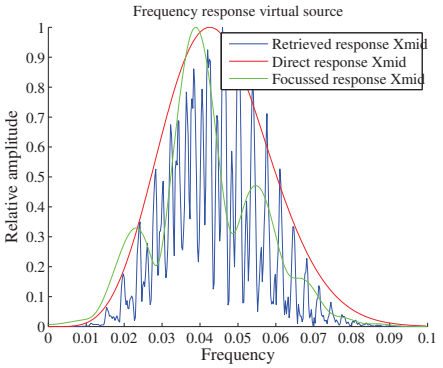
(a) Setup with virtual source 26 highlighted



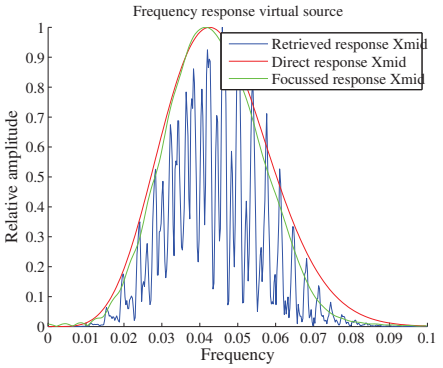
(b) Time response for receiver 13 in the auxiliary array due to virtual source 26



(c) Time response for receiver 13 in the auxiliary array due to virtual source 40

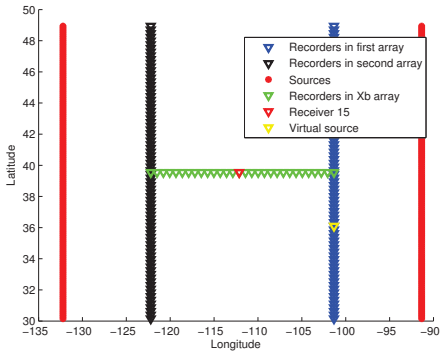


(d) Frequency response of first arrival for receiver 13 in the auxiliary array due to virtual source 26

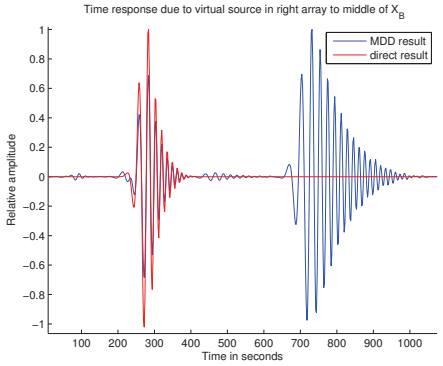


(e) Frequency response of third arrival for receiver 13 in the auxiliary array due to virtual source 26

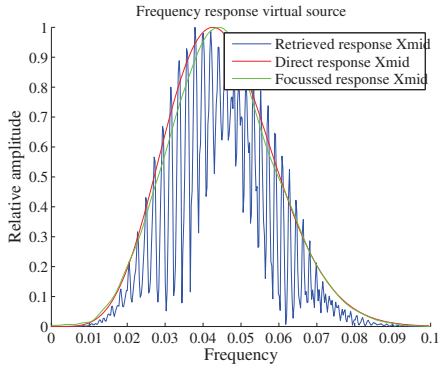
Figure 5-5: Retrieved result at center array due to virtual source at receiver 26 of the right array.



(a) Setup with virtual source 26 highlighted



(b) Time response for receiver 15 in the auxiliary array due to virtual source 26



(c) Frequency response of first arrival for receiver 15 in the auxiliary array due to virtual source 26

Figure 5-6: Retrieved result at center array due to virtual source at receiver 26 of the right array.

the multiples were less sensitive to the location of the virtual source and the distance between the auxiliary array and the MDD array.

Chapter 6

Conclusion

In conclusion, it appears to be possible to apply one-sided MDD to existing, publicly available, seismic array configurations. However, not all the array configurations are fit for proper MDD seismic interferometry. Recall the CANOE experiment, here the virtual source response did not match the expected result. Probably, the receiver spacing of the MDD array, in the CANOE experiment, is to coarse. Another drawback is that the collected data is not always of good quality. For instance in the New Mexico experiment, the noise sources originated from both sides of the continent. Standard seismic interferometry by MDD has been derived for one-sided illumination only, which means that the undesired measurements need to be removed. Beamforming did not work in the New Mexico experiment, due to violation of the assumption that no lateral velocity variations exist at the location of the array. The other option, wavefield decomposition did work. With decomposition it was possible to filter out one-sides illumination from recordings with two-sides illumination, but the quality of the decomposition depended on how accurate the apparent pressure was estimated. Reality forced a fairly large separation distance between the arrays used to estimate the apparent pressure. The decomposed result was therefore not perfect, and that reflects in the MDD result. Leakage of the wavefield that travels in the wrong direction distorted the frequency spectrum of the MDD result. When the right data was available, and the setting was good; the MDD result was sometimes distorted by an edge effect, caused by the limited extent of the MDD array. When this effect was removed, the remaining frequency content matched the reference frequency content. Unfortunately, it was difficult to completely separate the edge effect from the retrieved event in time.

Due to the difficulty of obtaining one-sided wavefields in some settings, we derived an alternative full-wavefield MDD scheme. The full wavefield MDD scheme seemed to work. The boundaries, both MDD arrays, acted as reflectors. Since the multiples were separated in time from the direct retrieval, this posed no problem to single out the direct arrival. Selecting the primary worked well with a small time window, and used that to retrieve the amplitude and phase spectrum. The full wavefield MDD scheme worked in both one- and two-dimensional settings. Especially the one dimensional setting was sensitive to wrap-around effects: the reflected wavefield was still strong when it reached the end of the time axis and pops back in at the end of the negative time axis. In situations with low damping, it became a fairly expensive operation to prevent wrap-around, the time axis became very long. This in turn made the model slower. In the used configurations, retrieving the frequency spectrum worked

well for the direct arrival when the virtual source was located in the center of the MDD array, but degraded quickly for a virtual source elsewhere. The frequency spectrum of the retrieved multiples seemed to be more stable.

The full wavefield MDD scheme seems promising, and is a good alternative for using wavefield decomposition together with the one-sided MDD scheme. The one-sided MDD result with wavefield decomposition is distorted, but the full wavefield MDD gives a result without distortions for a virtual source in the middle of the MDD array. For virtual sources away from the center of the MDD array; the direct arrival is a bit unstable, but the multiples contained the correct frequency spectrum.

For future research, it is still interesting to apply MDD on real data. The one-sided MDD theory should work, and is easiest to implement; however, it is sensitive to noise. Only the ingoing wavefields can be used, and wavefield decomposition does not look promising; it is a complex operation that requires atleast two, preferably three, parallel arrays to estimate the apparent pressure. However, in this thesis a crude derivative of the particle velocity is obtained. A more elegant solution might improve the overall results. Beamforming could work, if subsurface velocity is homogeneous in the vicinity of the MDD array, but it requires quality control of the data in small time windows. The other option is the full wavefield MDD procedure, which does not need to distinguish between in and outgoing observations, but it needs the normal velocity on two MDD arrays. Which means that the x and y component of the velocity needs to be measured, but this should pose no problem. It is interesting to see if MDD also retrieves the accurate phase velocity on real data.

Bibliography

CANOE (2005). http://www.iris.edu/gmap/xn?timewindow=2003-2005.

CDROM (2000). http://www.iris.edu/gmap/xk.

de Hoop, A. T. (1995). *Handbook of Radiation and Scattering of Waves*. Academic Press.

Gerstoft, P., Sabra, K. G., Roux, P., Kuperman, W. A., and Fehler, M. C. (2005). Greens functions extraction and surfacewave tomography from microseisms in southern california. *Geophysics*, 71:SI23SI31.

Menke, W. (1989). *Geophysical Data Analysis: Discrete Inverse Theory*, volume 45 of *international geophysics series*. Academic Press.

RGST (2000). http://www.iris.edu/gmap/xm.

Rost, S. and Thomas, C. (2002). Array seismology: methods and applications. *Reviews of Geophysics*, 40:1008.

Ruigrok, E. (2012). *Body-wave seismic interferometry applied to earthquake- and storm-induced wavefields*. PhD thesis, Technische Universiteit Delft.

Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W. A., and Fehler, M. C. (2005). Surface wave tomography from microseisms in southern california. *Geophysical research letters*, 32:L14311.

USArray (2009). http://www.iris.edu/gmap/ta.

Wapenaar, C. and Berkhout, A. (1989). *Elastic Wavefield Extrapolation: Redatuming of Single- and Multi-component seismic data*. Elsevier.

Wapenaar, K. (2004). Retrieving the elastodynamic greens function of an arbitrary inhomogeneous medium by cross correlation. *Physical Review Letters*, 93:254301.

Wapenaar, K. and Fokkema, J. (2006). Greens function representations for seismic interferometry. *Geophysics*, 71:SI33–SI46.

Wapenaar, K., Ruigrok, E., van der Neut, J., and Draganov, D. (2011a). Improved surface-wave retrieval from ambient seismic noise by multi-dimensional deconvolution. *Geophysical research letters*, 38.

Wapenaar, K., van der Neut, J., Ruigrok, E., Draganov, D., Hunziker, J., Slob, E., Thorbecke, J., and Snieder, R. (2011b). Seismic interferometry by crosscorrelation and multidimensional deconvolution: a systematic comparison. *Geophysical Journal International*, 185:1335–1364.

Appendix A

Appendix

A-1 MDD quality

During the implementation of the MDD scheme some problems arose. The problem was that none of the arrays above showed a well retrieved virtual source response. For instance the US array set-up should preform well, since the receiver spacing is small enough to prevent aliasing. All sorts of testing was done, for instance checking if the forward problem is well defined. Recall equation 2-12; instead of inverting it and solving for G_d, solving for the crosscorrelation function is possible the pointspread function and the Green's function response from the virtual source to the receiver array are given. Comparing this crosscorrelation function to the regular crosscorrelation gives an indication if the problem is well defined or not. The retrieved crosscorrelation function for the US array matches the real function well. See figure A-1 for the comparison. Finally, the problem is narrowed down to the following: the retrieved results in the MDD scheme degrade for lower frequencies and larger distances. The next set of images in figure A-2 show the US array model, where for each image the distance, time and dominant frequency are scaled by a different factor. When the distance (receiver spacing, source spacing and distance between the arrays and the source locations) decreases by a factor ten, the center frequency increases by a factor ten; Therefore the relative distance, the amount of wavelengths between the virtual source and receiver: $\frac{\lambda}{r}$, does not change and the results should be the same.

The results in figure A-2 are remarkable. From (a) to (d) the retrieved virtual source response due to multi dimensional deconvolution changes drastically. For the smallest scale, the retrieved frequency spectrum result matches almost perfect with the expected frequency spectrum. In these models, when the distance decreases, the frequency increases with a same factor; while the velocity remains constant. The relative distance, for instance the number of wavelengths from source to receiver, remains constant; but the results are different. It turns out that when adjusting the code to work with low frequency waves, larger distances and realistic velocities, the definition of the source wavelet broke down. The source wavelet used to generate the data did not fit in the defined source wavelet window, thereby introducing errors that were clearly visible in the MDD result only.

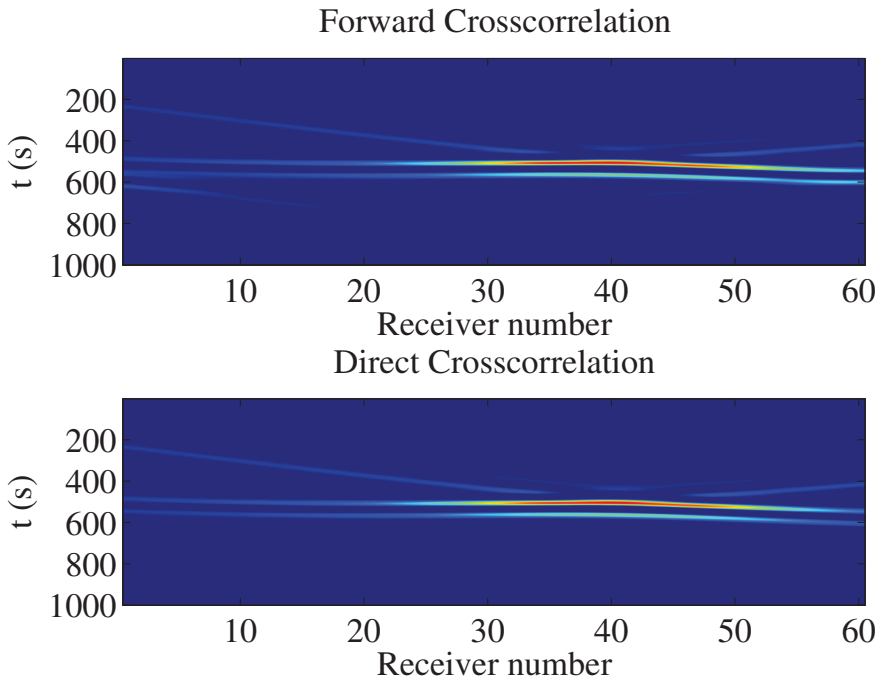
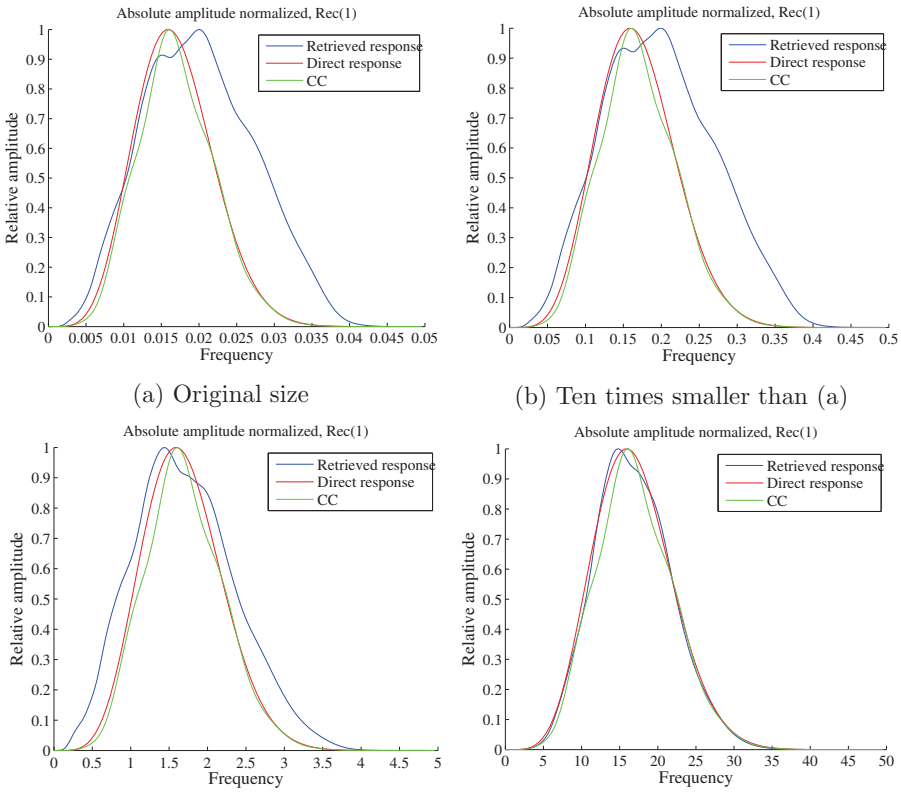


Figure A-1: Comparison of the retrieved crosscorrelation and directly calculated crosscorrelation

Furthermore, the edge effect, due to the finite extent of the MDD array, causes an event that overlaps with the retrieved response for a virtual source in the center of the MDD array. Picking a virtual source away from the center, separates the edge effect in time from the retrieved response. Now it is possible to, at least partially, remove the edge effect to see what happens to the retrieved frequency spectrum. See figure A-3 for the used setup, the edge effect in time, and the effect of removing the edge effect to the retrieved frequency spectrum. The result A-3b shows that the MDD response has an edge-effect well before the reference response starts. Removing this effect, via time-windowing, shows a much smoother frequency spectrum in figure A-3c compared to the original retrieved result. Withouth the edge effect, the retrieved frequency spectrum is very similar. So MDD is capable of retrieving a virtual source response that perfectly matches the reference source. The edge effect showed up in all settings tested above, but in some arrays stronger than in others.



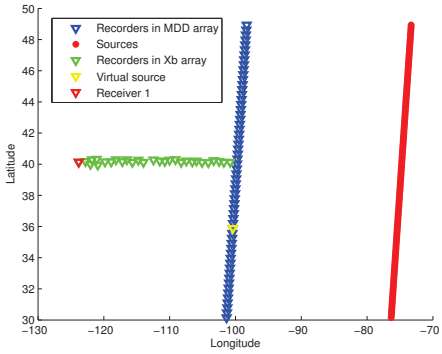
(a) Original size

(b) Ten times smaller than (a)

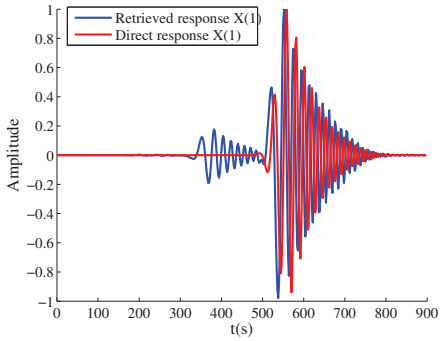
(c) Hundred times smaller than (a)

(d) Thousand times smaller than (a)

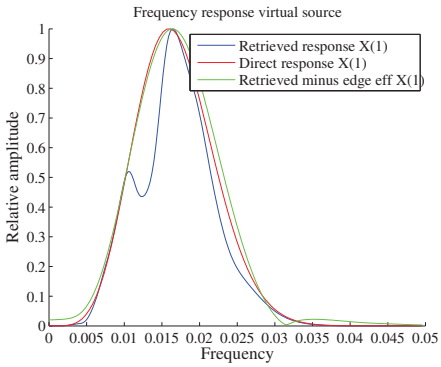
Figure A-2: Retrieved virtual source response for the same setting at different scales. Result increases drastically when the distances decrease, and the dominant frequency goes up.



(a) Used setup and virtual source location



(b) Edge effect in time, visible before retrieved event



(c) Edge effect in frequency domain

Figure A-3: Effect of edge effect shown in time- and frequency-domain.