Geophysical Journal International

Geophys. J. Int. (2013) **193,** 788–797 doi: 10.1093/gji/ggt021

Towards high resolution quantitative subsurface models by full waveform inversion

P. Haffinger, A. Gisolf and P. M. van den Berg

Department of Imaging Science & Technology, Section Acoustical Wavefield Imaging, Delft University of Technology, Delft, Netherlands.
E-mail: P.R.Haffinger@tudelft.nl

Accepted 2013 January 21. Received 2013 January 21; in original form 2012 March 27

SUMMARY

Full waveform inversion (FWI) has the potential to recover detailed quantitative property models of the subsurface, but the process is computationally expensive. Currently available computer systems do not allow to use the full bandwidth of the acquired seismic data, which effectively reduces the resolution that can be obtained. In this paper, we propose a novel approach to obtain high resolution subsurface models from broad-band FWI. The method is based on localization of the inversion, while subsequently the interaction between local domains is estimated. A global field update is calculated which honours the non-linear relationship between the subsurface properties and the measured seismic data. By using this non-linearity, the spectral gap between the *a priori* background model and the seismic bandwidth will be closed and spatially broad-band properties can be estimated from a band-limited seismic signal.

Key words: Image processing; Numerical solutions; Inverse theory.

1 INTRODUCTION

In principle, full waveform inversion (FWI) is a suitable tool to obtain information about the properties of the subsurface on a very detailed scale and the process was described for seismic application already more than three decades ago (Bamberger *et al.* 1977; Symes, 1981; Bamberger *et al.* 1982; Lailly, 1983; Tarantola, 1984, 1987). The initial work on waveform inversion has been developed further over the years and today Gauss–Newton (Pratt *et al.* 1998; Epanomeritakis *et al.* 2008; Fichtner, 2011; Métivier *et al.* 2012) or gradient based methods (Mora, 1987; Crase *et al.* 1990) are most commonly used to recover subsurface models from seismic data. A bottleneck of these approaches is the large memory requirement and the necessity to solve several full forward problems during the process. Even currently available huge computer clusters cannot provide the capacity that would be required for FWI while using the full bandwidth of the measured data. By full bandwidth the authors refer to a spectrum with maximum frequencies between 50 and 100 Hz. To overcome this limitation, recent approaches have been made by encoding sources and receivers (Krebs *et al.* 2009; Abubakar *et al.* 2011; Routh *et al.* 2011), or by using compressive sensing, if the system to be solved is sparse in some sense (Moghaddam & Herrmann, 2010). Also multi-scale approaches could be deployed to decrease the problem size while at the same time incorporating higher frequencies (Bunks, 1995; Brossier *et al.* 2009).

Interestingly, the seismic community seems to focus on reduction of the data space while reduction of the model space is not really investigated yet. Despite all efforts in reducing the amount of data that is simultaneously needed, maximum frequencies used for inversion (∼20 Hz in 2-D and several Hz in 3-D) are still far away from the maximum frequencies measured in the seismic surface data. Consequently, the role of FWI is still limited to providing background velocity models that can be used as input to a migration scheme (Plessix *et al.* 2010; Prieux *et al.* 2010; Vigh *et al.* 2011). This means that the final output of the seismic processing chain remains a structural image that will be used for interpretation.

In this paper, we describe a gradient based FWI scheme with the potential to notably increase the upper frequency limit that can be taken into account for FWI. We discuss mainly two new aspects: reduction of the memory requirements by localizing the inversion and iterative calculation of the wave fields in the inversion domain to avoid solving the full forward problem several times. The purpose of our work is not to speed-up waveform inversion but to reduce it's computational complexity, especially in terms of simultaneous memory requirements. As a result, broad-band FWI of seismic surface data should become feasible with a very fine subsurface discretization (≈5 m) while inverting frequencies of 50 Hz and beyond. Consequently, FWI could become feasible to provide a resolution that is competitive to structural imaging by migration.

In the described approach we split up the large-scale inversion problem into many small, and highly parallelizable, local inversions. In a first step seismic data is back-propagated into the subsurface (Wapenaar 1989) which allows us to invert for the medium properties of a local domain only, while the rest of the subsurface can be disregarded for the moment. The method combines redatuming (Berryhill 1984; Shtivelman & Canning 1988; Mulder 2005) and target-oriented inversion approaches (Valenciano *et al.* 2006; Staal *et al.* 2010). For back-propagation a background medium is needed that sufficiently describes the kinematics and we assume that such a model, derived from existing methods, exist. We fully realize that

obtaining such a background model is a non-trivial task and we refer the reader to an excellent overview on existing FWI work given by Virieux & Operto (2009) because a discussion of all challenges is out of scope of this paper. Since the local back-propagated data set is much smaller than the entire surface data set and a local domain is only a subdomain of the global subsurface, the problem size gets reduced significantly and broad-band FWI on a finely discretized grid becomes possible. It should be realized that domain decomposition methods are generally used to speed-up the forward problem (Bamberger *et al.* 1997). This is different to our approach where we use a similar concept to split up the inverse problem into many local domains. Subsequently, these local inversion results are combined into a global high-resolution property model of the subsurface. This procedure has been successfully demonstrated under the assumption that seismic data is fully linear in the model properties (Haffinger *et al.* 2011). If this assumption holds and the medium is supposed to generate primary reflections only, the described localized inversion might lead to a satisfactory result, because field interaction between the local domains, for example, multiple scattering and transmission, can be neglected. However, in real life the relationship between the subsurface properties and the measured data is mostly significantly non-linear and this fact has to be taken into account. This means that a total field update has to be applied that takes intra domain and inter domain interaction into account to make the wave propagation in the subsurface consistent with the latest obtained subsurface model. This procedure effectively estimates the field interaction between the local domains and hence inter domain multiples as well as propagation through the overburden above the local domains is accounted for. Our total field update incorporates all propagation effects, including internal multiples, transmission and the true traveltimes in the inverted medium.

The total inversion scheme involves an iterative procedure of localized inversions and global field updates, which converges when the obtained subsurface model and the fields at each gridpoint in the subsurface do not change any more. The method is supposed not only to recover the structure, but also quantitative property information of the subsurface with a resolution well beyond the structural images that are obtained from current migration schemes. This implies that results from our FWI approach could be directly used for structural interpretation but also for reservoir characterization. Furthermore it should be understood that we do not aim at replacing current FWI schemes but the described approach has to be seen as another step, building on current results by direct inversion of surface data. The subsurface models coming from direct FWI of surface data then become new background models for the scheme as outlined in this paper. A sketch of the full scheme is shown in Fig. 1.

In this paper, we limit ourselves to the acoustic case in which the density is assumed to be constant and known and the method will be illustrated using a synthetic data set. Future research will be devoted to extending the scheme to a multiparameter FWI inverting for the full elastic set of subsurface properties. Gisolf & van den Berg (2012) discussed the elastic implementation of our approach in a reservoir oriented 1.5-D setting but implementation of the elastic 2-D and finally 3-D schemes still needs to be investigated.

2 METHODOLOGY

Although the method will be exemplified on a 2-D data set, derivation of the scheme will be for the general 3-D case. Before we describe the methodology we present a summary of the variables that will be used throughout this paper in Table 1.

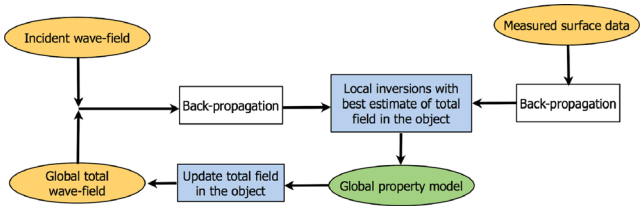

Figure 1. Schematic view of our novel FWI approach. By back-propagation the inversion can be localized and global high resolution property images can be obtained by combining the local results. While the inversion is initialized assuming fields propagating in a smooth background model only (background fields), subsequently the true field at each gridpoint in the subsurface is updated based on the currently best known subsurface model. In this way the non-linear relationship between the subsurface properties and the measured data can be fully honoured.

Table 1. Most important variables used throughout the paper.

ω	=	temporal angular frequency,
$\mathcal{D}$	=	scattering domain (subsurface),
$\mathcal{S}$	=	surface domain,
$\mathbf{x}_s$	=	source location along $\mathcal{S}$,
$\mathbf{x}_r$	=	receiver location along $\mathcal{S}$,
$\mathbf{x}$, $\mathbf{x}'$	=	scattering locations in $\mathcal{D}$,
$c_0(\mathbf{x})$	=	acoustic wave velocity of the background medium,
$c(\mathbf{x})$	=	acoustic wave velocity of the true medium,
ρ_0	=	constant mass density of the background medium,
$\chi(\mathbf{x})$	=	contrast function,
$\mathcal{G}$	=	Green's function in background medium,
p_{sct}	=	measured seismic data,
p_0	=	background field at each point in the subsurface,
p_{tot}	=	total field at each point in the subsurface.

In the acoustic approximation with constant density we have to solve a system of two coupled equations, the data equation and the domain equation. The data equation describes the seismic data set, measured along the Earth's surface, in terms of a total field at each gridpoint in the subsurface, the contrast function and the Green's functions in a non-scattering background medium:

$$p_{\text{sct}}(\mathbf{x}_r, \mathbf{x}_s, \omega) = \int_{\mathbf{x}' \in \mathcal{D}} \mathcal{G}(\mathbf{x}_r, \mathbf{x}', \omega)\, \chi(\mathbf{x}')\, p_{\text{tot}}(\mathbf{x}', \mathbf{x}_s, \omega)\, \mathrm{d}\mathbf{x}'. \qquad (1)$$

The contrast function that is supposed to be determined by FWI is defined as

$$\chi(\mathbf{x}) = 1 - \left[\frac{c_0(\mathbf{x})}{c(\mathbf{x})}\right]^2, \qquad (2)$$

and depends on the difference between a known background medium and the true but unknown subsurface model. From eq. (1) it can be seen that the total field is non-linear in the contrast χ. If the dependency of the total field on the contrast function is neglected, the data equation can be linearized and approximate subsurface properties can be obtained by a simple linear inversion. On the other hand, if the exact total field would be known, the exact contrast function could be obtained by again a simple linear inversion.

In practice the total field is usually approximated by a background field that propagates in a smooth medium only, while multiple scattering and transmission in the subsurface is neglected. Note that this assumption is made by all current migration algorithms, which implies that the obtained structural images assume a linear relationship between the data and the subsurface properties. It should

be clear that in many situations the true field propagation in the subsurface is far more complex than could be described by a linear data model. To overcome this limitation and to obtain a non-linear inversion scheme we make use of the domain equation which describes the total field at each gridpoint in the subsurface in terms of a field propagating in the background, the contrast and the Green's functions in a background medium:

$$p_{\text{tot}}(\mathbf{x}, \mathbf{x}_s, \omega) = p_0(\mathbf{x}, \mathbf{x}_s, \omega) + \int_{\mathbf{x}' \in \mathcal{D}} \mathcal{G}(\mathbf{x}, \mathbf{x}', \omega)\, \chi(\mathbf{x}')\, p_{\text{tot}}(\mathbf{x}', \mathbf{x}_s, \omega)\, \mathrm{d}\mathbf{x}'. \quad (3)$$

The inverse scattering problem consists of determining $\chi(\mathbf{x}')$ from the best estimate of the total fields p_{tot} and the measured seismic data p_{sct}. We recast the inverse problem as an optimization problem of finding χ to minimize the error in eq. (1), subject to the constraint that eq. (3) is satisfied in some sense.

Eqs (1) and (3) are the Lippmann–Schwinger equation (Lippmann & Schwinger 1950) for acoustic scattering problems with constant density. If the contrast function is known, the true total field, including all complex propagation in the true subsurface, can be obtained by solving eq. (3) numerically. The problem in seismic inversion is that the contrast function is supposed to be found from the seismic data and is generally unknown. The challenge in FWI is to find the total field at each point in the subsurface as well as the unknown contrast function. In the following we explain an iterative approach in which we solve eqs (1) and (3) in an alternating manner. We estimate the contrast function from the data equation assuming that the total field does not change, followed by updating the total field based on the domain equation, but assuming that the contrast function is known. In this way we obtain a non-linear inversion result and we stop the iterative procedure when the contrast function and the total field do not change any more.

The scheme could in principle be directly applied to the measured surface data, but it was mentioned earlier that current computer systems do not allow for high-resolution full bandwidth FWI. To overcome this limitation we localize the inversion process and combine the local results to a global property model.

3 LOCALIZED LINEAR INVERSIONS

In the previous section the data equation and its role in FWI was discussed. If the total field during the inversion is fixed, a linear inversion scheme can be used to retrieve a subsurface model from the measured seismic data. Note that this statement even holds if the total field in eq. (1) is not just the background field in a smooth medium, but becomes a total field that contains non-linear wave phenomena, estimated in a more complex subsurface model. The only restriction is that during inversion of eq. (1) the total field remains fixed.

The main problem that remains is that the numerical costs of this linear inversion grow exponentially with the amount of data, as well as with the extension and the discretization of the unknown subsurface model. With large seismic data sets, and the huge amount of unknowns in a finely discretized subsurface model that is required for full bandwidth FWI, the process cannot be performed on today's computer systems.

If a high resolution property model needs to be recovered by full bandwidth inversion, the problem size has to be decreased significantly. Current efforts are made by the seismic community to reduce the amount of data that is simultaneously involved in the inversion. Although reducing the data size is a powerful tool, all current schemes invert for the properties of the entire subsurface at once. In our scheme we back-propagate seismic sources and receivers into the subsurface to localize the inversion process. This means that the full surface acquisition is smartly combined into a back-propagated local data set and by subsequent time-windowing we can invert for the properties of a limited local domain while neglecting the rest of the subsurface for the moment. The scheme involves also back-propagation of the total fields which makes the localization effectively a pre-conditioning operation since it is applied to both sides of eq. (1).

Defining the back-propagation operators in the frequency domain as

$$[\mathcal{F}_r\, p_{\text{sct}}](\mathbf{x}_{\text{br}}, \omega) = 2 \int_{\mathbf{x}_r \in \mathcal{S}} \frac{\partial}{\partial z_r} \mathcal{G}^*(\mathbf{x}_{\text{br}}, \mathbf{x}_r, \omega) p_{\text{sct}}(\mathbf{x}_r, \omega)\, \mathrm{d}\mathbf{x}_r\,,$$

$$[\mathcal{F}_s\, p_{\text{sct}}](\mathbf{x}_{\text{bs}}, \omega) = 2 \int_{\mathbf{x}_s \in \mathcal{S}} \frac{\partial}{\partial z_s} \mathcal{G}^*(\mathbf{x}_{\text{bs}}, \mathbf{x}_s, \omega) p_{\text{sct}}(\mathbf{x}_s, \omega)\, \mathrm{d}\mathbf{x}_s\,, \quad (4)$$

where $\mathbf{x}_r$ and $\mathbf{x}_s$ stand for the actual source and receiver positions, while $\mathbf{x}_{\text{br}}$ and $\mathbf{x}_{\text{bs}}$ denote the back-propagated source–receiver positions, we apply a dual back-propagation to eq. (1)

$$[\mathcal{F}_r(\mathcal{F}_s\, p_{\text{sct}})]\,(\mathbf{x}_{\text{br}}, \mathbf{x}_{\text{bs}}, \omega) = \int_{\mathbf{x}' \in \mathcal{D}} [\mathcal{F}_r \mathcal{G}](\mathbf{x}_{\text{br}}, \mathbf{x}', \omega) \chi(\mathbf{x}') [\mathcal{F}_s\, p_{\text{tot}}](\mathbf{x}', \mathbf{x}_{\text{bs}}, \omega)\, \mathrm{d}\mathbf{x}'\,. \quad (5)$$

Back-propagation as described in eq. (4) is dependent on the anticausal Green's function, $\mathcal{G}^*$, between the locations of the back-propagated sources/receivers in the subsurface and the locations of the actual sources/receivers located along the surface. Because back-propagation is applied to the left- and right-hand side (RHS) of the data equation, the process becomes a pre-conditioning operation and no model error is made. A similar reasoning tells us that the integral in eq. (4) may be replaced by a finite summation over the actual sources/receivers, without introducing a model error. This allows us to replace the exact Green's functions in eq. (4) by ones that are parametrized in terms of traveltimes and approximate amplitudes only. It is crucial to understand that back-propagation is solely supposed to increase the sensitivity of the inversion with respect to a certain local domain of the subsurface. Handling of the complex overburden will be part of the global non-linear field update as will be described later in this paper. Nevertheless, since after back-propagation a certain time-window of the data will be chosen, the Green's functions have to describe the kinematic propagation reasonably well and we assume that for this purpose a background model, derived from standard velocity analysis or current FWI techniques, exists. Interestingly, pre-conditioning by back-propagation could be incorporated in any other FWI scheme to increase the sensitivity of the error functional with respect to an arbitrary target.

Although in this paper we describe the method for surface seismic data, the same concept can in principle also be applied to other acquisition geometries, for example, seismic well data. Still, a local back-propagated data set has to be generated above the domain of interest and such a technique was recently proposed for VSP data by Soni *et al.* (2012).

After carrying out the back-propagation we obtain the data equation with back-propagated sources located at $\mathbf{x}_{\text{bs}}$ and back-propagated receivers located at $\mathbf{x}_{\text{br}}$

$$\hat{p}_{\text{sct}}(\mathbf{x}_{\text{br}}, \mathbf{x}_{\text{bs}}, \omega) = \int_{\mathbf{x}' \in \mathcal{D}} \hat{\mathcal{G}}(\mathbf{x}_{\text{br}}, \mathbf{x}', \omega) \chi(\mathbf{x}') \hat{p}_{\text{tot}}(\mathbf{x}', \mathbf{x}_{\text{bs}}, \omega)\, \mathrm{d}\mathbf{x}', \quad (6)$$

where $\hat{p}_{\text{sct}} = [\mathcal{F}_r[\mathcal{F}_s\, p_{\text{sct}}]]$, $\hat{\mathcal{G}} = [\mathcal{F}_r \mathcal{G}]$ and $\hat{p}_{\text{tot}} = [\mathcal{F}_s\, p_{\text{tot}}]$. Eq. (6) is equivalent to eq. (1) with the difference that the back-propagated

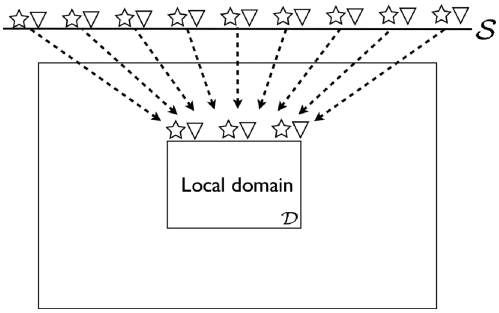

Figure 2. By back-propagating the surface acquisition, a local data set is generated at an arbitrary depth level. If only a limited time-window of this back-propagated data set is used the medium properties of a local domain can be obtained from FWI while neglecting the rest of the subsurface.

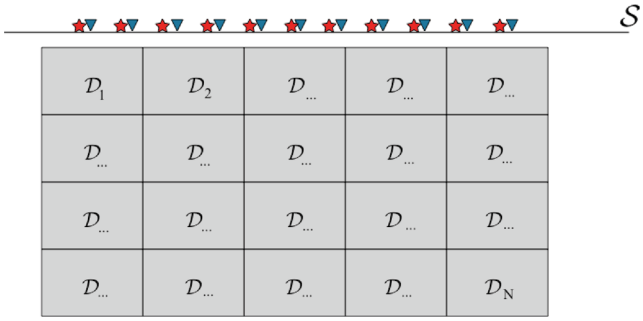

Figure 3. A number of local domains can be combined to obtain a global property model of the subsurface.

sources and receivers are located within the subsurface domain $\mathcal{D}$ instead along the surface domain $\mathcal{S}$. The situation after back-propagation is shown in Fig. 2.

The back-propagated data consists of a causal and an anticausal part. The causal part contains reflections from points in the subsurface below the new acquisition level, while the reflections from above show up in the anticausal part. If we use the causal part only and also apply a time-window, we can invert for the subsurface properties of a limited local domain located just below the back-propagated sources and receivers. This can be justified by the fact that wave fields can propagate only a limited distance from the source into the medium and back to the receivers within a limited time (i.e. causality). The time-window is calculated based on the source–receiver distance and the maximum depth of the local domain. The traveltime in the background model from each source to the midpoint of a source–receiver pair at the maximum depth of the local domain and back to the receiver, becomes the maximum window length for a particular trace. We choose the midpoint to calculate the maximum time, to ensure that the reflection point of the inverted data lies within the local domain. One could also decide to use the maximum traveltime to any point in the local domain and back to the receiver. Choosing the time-window is a trade off between maximum illumination of the local domain and possibly generating artefacts because the time-window contains data that cannot be explained by the properties of the local domain. Nevertheless, time-windowing is in only necessary for the very first iterations and we show later how this procedure becomes in principle oblique by combining the local results and taking their interaction into account.

The whole scheme reduces the amount of data and the size of the inversion domain significantly. This reduction of the problem size allows us to perform the inversion on a very fine grid and we can use the full seismic bandwidth to obtain high resolution property images of the subsurface. For inversion of eq. (6) we use a multiplicative regularized conjugate gradient (MR-CG) scheme. In principle any suitable inversion scheme can be used and we refer the interested reader to Abubakar *et al.* (2004) for more detailed information on multiplicative regularization. Although we believe that multiplicative regularization is superior to additive regularization, for example, hands-off estimation of the regularization parameter and a regularization factor that goes to one with increasing iterations, a comparison of both methods is not the objective of this paper.

To obtain a global high resolution property model of the subsurface, we repeat the described FWI approach for many overlapping local domains and finally the results are combined to form a global model. In complex geologies, neglecting inter domain effects during the first iteration, can potentially prevent the CG-scheme from minimizing the error functional, effectively leading to zero or very small contrast values in these local domains. By choosing for overlapping local domains, every subsurface gridpoint is covered multiple times and the probability is increased that every gridpoint gets assigned a reasonable contrast value. This is of interest especially early in the iterative scheme, while later inter domain effects are estimated and properly taken into account. Since all local domains are independent we can run the inversions in parallel, allowing very fast computation of a high resolution property model. Combining the local domains to obtain a global model is sketched in Fig. 3.

4 GLOBAL NON-LINEAR FIELD UPDATE

In the previous section it was described how, by localizing FWI, a high resolution property model of the subsurface can be obtained. Although in our approach always a linear inversion is applied to obtain the contrasts from the data, the non-linear relationship between the subsurface properties and the measured wave fields will be incorporated in updating the total fields. Note that if the exact total field can be found, a linear inversion of the data equation is supposed to yield the exact property model of the subsurface. Updating the total fields in the subsurface represents a forward modelling problem that could be performed by several numerical methods. The best known and also most widely used modelling techniques are based on finite-differences. This method can provide a good approximation of the total field for a given contrast function with the drawbacks of being computationally rather expensive and being subject to numerical dispersion. Since, especially during the very first iterations, the contrast is known not to be correct yet, we argue that it is unnecessary and undesirable to calculate an exact solution based on inexact properties. Still, finite-difference methods could potentially be used for updating the fields in the medium under consideration but during development of the scheme it turned out that an iterative approach shows better convergence behaviour. To fully prove this observation further research is needed and a comparison between both forward modelling methods would certainly be of interest. In the following we describe a more effective and intrinsically stable method for updating the total fields in the subsurface. To this purpose we write the total field as a sum of the background field and a number of basis functions:

$$p^{(N)}_{\mathrm{tot}}(\mathbf{x}, \mathbf{x}_s, \omega) = p_0(\mathbf{x}, \mathbf{x}_s, \omega) + \sum_{n=1}^{(N)} \alpha^{(N)}_n(\omega)\phi_n(\mathbf{x}, \mathbf{x}_s, \omega), \qquad (7)$$

where $\alpha^{(N)}_n$ are frequency dependent weighting factors and depend on the number N of basis functions taken into account. If the

weighting factors are equal to one, eq. (7) becomes the Neumann series which is known to be unstable for large contrast, high frequencies or large scattering domains. Note that N is equivalent to the iteration number and for each iteration one basis function is added. The basis functions are defined as

$$\phi_n(\mathbf{x}, \mathbf{x}_s, \omega) = \int_{\mathbf{x}' \in \mathcal{D}} \mathcal{G}(\mathbf{x}_r, \mathbf{x}', \omega) \partial \mathcal{W}_n \, \mathrm{d}\mathbf{x}', \tag{8}$$

and the incremental contrast sources $\partial \mathcal{W}_n$ are given by

$$\partial \mathcal{W}_1 = \chi_1 p_{\mathrm{tot}}^{(0)}, \partial \mathcal{W}_n = \chi_n \, p_{\mathrm{tot}}^{(n-1)} - \chi_{n-1} \, p_{\mathrm{tot}}^{(n-2)}, \quad n > 1. \tag{9}$$

Substituting eq. (7) into eq. (3) and including the dependency of χ on the iteration number n leads to

$$\sum_{n=1}^{(N)} \alpha_n^{(N)}(\omega) \phi_n(\mathbf{x}, \mathbf{x}_s, \omega) = \int_{\mathbf{x}' \in \mathcal{D}} \mathcal{G}(\mathbf{x}, \mathbf{x}', \omega) \chi_n(\mathbf{x}') p_0(\mathbf{x}', \mathbf{x}_s, \omega) \, \mathrm{d}\mathbf{x}'$$

$$+ \int_{\mathbf{x}' \in \mathcal{D}} \mathcal{G}(\mathbf{x}, \mathbf{x}', \omega) \chi_n(\mathbf{x}') \sum_{n=1}^{(N)} \alpha_n^{(N)}(\omega) \phi_n(\mathbf{x}', \mathbf{x}_s, \omega) \, \mathrm{d}\mathbf{x}'. \tag{10}$$

Eq. (10) is a linear system of equations including the unknown weighting factors $\alpha_n^{(N)}$. Let we assume that the number of points $\mathbf{x}$ in $\mathcal{D}$ is a finite number J. For each frequency ω and each source position $\mathbf{x}_s$ we define a J by N matrix ($J \geq N$) as

$$\mathbf{M} = \phi_n(\mathbf{x}, \mathbf{x}_s, \omega) - \int_{\mathbf{x}' \in \mathcal{D}} \mathcal{G}(\mathbf{x}, \mathbf{x}', \omega) \chi_n(\mathbf{x}') \phi_n(\mathbf{x}', \mathbf{x}_s, \omega) \, \mathrm{d}\mathbf{x}' \tag{11}$$

and a known vector $\mathbf{b}$ as

$$\mathbf{b} = \int_{\mathbf{x}' \in \mathcal{D}} \mathcal{G}(\mathbf{x}, \mathbf{x}', \omega) \chi_n(\mathbf{x}') p_0(\mathbf{x}', \mathbf{x}_s, \omega) \, \mathrm{d}\mathbf{x}'. \tag{12}$$

We now can re-write eq. (10) in matrix vector form to

$$\mathbf{M} \, \boldsymbol{\alpha} = \mathbf{b}. \tag{13}$$

Note that the weighting factors $\alpha_n^{(N)}$ have been combined to the vector $\boldsymbol{\alpha}$ and that the summation over the basis functions is now represented as a vector matrix multiplication. Inverting eq. (13) in the least-squares sense leads to optimal weighting factors $\alpha_n^{(N)}$:

$$\boldsymbol{\alpha} = \left(\mathbf{M}^H \mathbf{M}\right)^{-1} \mathbf{M}^H \mathbf{b}, \tag{14}$$

where H denotes the Hermitian of a matrix. Least-squares estimation of α ensures that the total field update as described in eq. (7) cannot diverge. This is beneficial compared to the straightforward Neumann series or finite-difference methods, which can suffer from instability issues, especially at higher frequencies. It is important to realize that each time a new basis function is added to the total field, N is increased by one and a complete new set of optimized $\alpha_n^{(N)}$ is calculated. Note that if the contrast is not changing, that is, $\chi(\mathbf{x}) = \chi_n(\mathbf{x})$, the scheme becomes equivalent to the Krylov subspace method. A detailed description of the Krylov method can be found in Kleinman & van den Berg (1991).

5 OVERBURDEN AND INTER DOMAIN MULTIPLES

We described earlier how by back-propagating sources and receivers into the subsurface a localized inversion can be performed. The concept is based on the fact that back-propagation, followed by time-windowing, can separate the reflection responses associated with the local domain, from the data that is generated from the surroundings of the local domain. However, back-propagation is a linear concept that cannot separate multiples that were generated in a shallower local domain, but appear in the time-window that is used for the current local domain. Although in the first iteration we have no other option than ignoring this fact, we subsequently will build up knowledge of the contrasts and the total field at every gridpoint in the subsurface. This allows us to predict overburden multiples and inter domain wave phenomena which means we can subtract them from the surface data before localization of the inversion. Mathematically the process can be described by splitting up the data equation:

$$p_{\mathrm{sct}}(\mathbf{x}_r, \mathbf{x}_s, \omega) = \int_{\mathbf{x}' \in \mathcal{D}_\cup} \mathcal{G}(\mathbf{x}_r, \mathbf{x}', \omega) \chi(\mathbf{x}') p_{\mathrm{tot}}(\mathbf{x}', \mathbf{x}_s, \omega) \, \mathrm{d}\mathbf{x}'$$

$$+ \int_{\mathbf{x}' \in \mathcal{D}_\cap} \mathcal{G}(\mathbf{x}_r, \mathbf{x}', \omega) \chi(\mathbf{x}') p_{\mathrm{tot}}(\mathbf{x}', \mathbf{x}_s, \omega) \, \mathrm{d}\mathbf{x}'. \tag{15}$$

where the first term on the RHS is the contribution from the current local domain, indicated by $\mathcal{D}_\cup$ and the second term is the contribution from the complement of the local domain, indicated by $\mathcal{D}_\cap$. It should be noted that in the equation above sources and receivers are still located along the surface. Although initially not possible, after every following iteration the complementary part of the data can be calculated and subtracted from the surface data in order to back-propagate only the part of the seismic data that is related to the local domain under consideration. If we define the surface data related to a specific local domain L, after subtracting the complementary part and at iteration n as

$$p_n^L(\mathbf{x}_r, \mathbf{x}_s, \omega) = p_{\mathrm{sct}}(\mathbf{x}_r, \mathbf{x}_s, \omega)$$

$$- \int_{\mathbf{x}' \in \mathcal{D}_\cap} \mathcal{G}(\mathbf{x}_r, \mathbf{x}', \omega) \chi_{n-1}(\mathbf{x}') p_{\mathrm{tot}}^{n-1}(\mathbf{x}', \mathbf{x}_s, \omega) \, \mathrm{d}\mathbf{x}' \tag{16}$$

we can write the data equation for a specific local domain as

$$p_n^L(\mathbf{x}_r, \mathbf{x}_s, \omega) = \int_{\mathbf{x}' \in \mathcal{D}_\cup} \mathcal{G}(\mathbf{x}_r, \mathbf{x}', \omega) \chi_{n-1}(\mathbf{x}') p_{\mathrm{tot}}^{n-1}(\mathbf{x}', \mathbf{x}_s, \omega) \, \mathrm{d}\mathbf{x}'. \tag{17}$$

This means that after the first iteration, eq. (17) will be used and back-propagated to localize the inversion process. In this way overburden multiples and inter domain effects are effectively removed in an iterative procedure. Furthermore, time-windowing can finally be abandoned and multiples that are generated within the local domain but arrive outside the primary time-window can be used and can therefore contribute to the inversion as well.

6 NUMERICAL RESULTS

We illustrate the feasibility of the approach discussed on a small part of the Marmousi2 model. We utilize Marmousi2 instead of its original version because of the much finer spatial discretization. Since current implementation of our scheme can only use a homogeneous background model we reduce the low wavenumber content of the true Marmousi2 model. In a homogeneous background the Green's functions $\mathcal{G}$ can be calculated analytically which significantly simplifies the computation. Nonetheless, for real data applications the use of smooth inhomogeneous background media is inevitable. This means that exact Green's functions in a smooth background medium have to be calculated. Since this can be a rather expensive task, an effective modelling of Green's functions in inhomogeneous media is necessary. An interesting effort was made by Fokkema & van den Berg (2012) for medium stretching to effectively make it homogeneous. Application of this method to a smooth background medium would then allow direct use of our current implementation. Another

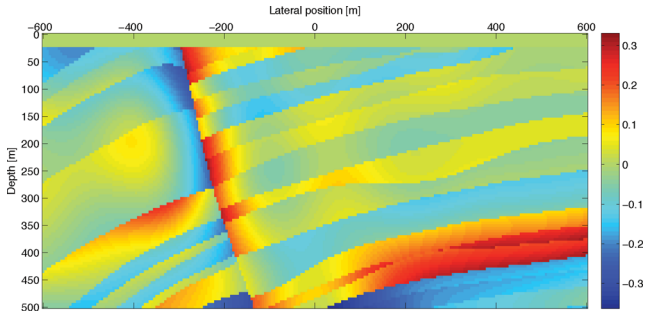

Figure 4. Exact contrast function that was used to generate synthetic data with sources and receivers located along the surface.

promising method to incorporate a smooth background medium is given by Abubakar *et al.* (2009). Storing these Green's functions is not an option, so they have to be either recalculated for each iteration, or written to disk and loaded whenever needed. Since storage is usually extensively available, we currently use the latter option in 2-D. However, for a realistically sized 3-D inversion object, storing the Green's functions will not be feasible and another smart solution needs to be thought of. Again, we are not so much concerned about efficiency yet, but about making FWI possible on a much finer discretization as currently possible. It has to be realized that although the subsurface model is updated for each iteration, the background model is kept the same. Updating the background is optional but requires calculation of the Green's functions in the new background medium. The contrast function with respect to the homogeneous background model with $c_0 = 2000$ m s^{-1} is displayed in Fig. 4. The total width of the modelling and the inversion domain was 150 m wider on each side but in the following we only display the part where coverage was well enough to get a reliable inversion result. For modelling we use a constant density of $\rho = 2000$ kg m^{-3} and for inversion this density is assumed to be known. As mentioned earlier, extension to elastic FWI in 2-D and finally 3-D will be the subject of future research.

By solving the full acoustic wave equation for the shown part of the Marmousi2 model, we generate an acoustic data set with 51 sources and 51 receivers located along the surface, while source and receiver separation is 30 m each. A selection of shot records is displayed in Fig. 5. Since we remove the water layer from the Marmousi2 model, surface multiples will have a minimal influence. Demonstration of the method on different subsurface models, including ones generating strong free-surface multiples, will be part of a follow-up publication. Still, all types of multiples, surface as well as internal, are properly modelled and present in the surface data. The wavelet that was used to compute synthetic data is displayed in Fig. 6. The maximum frequency in the data is as high as 55 Hz and the full bandwidth will be used for inversion.

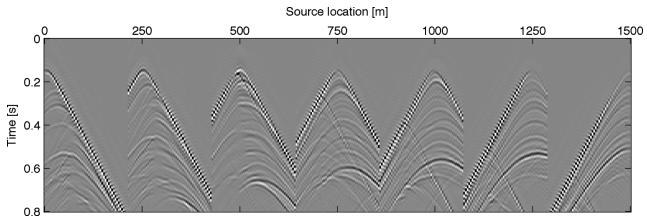

Figure 5. Some shot records from the synthetic data that was generated with sources and receivers located along the surface of the model shown in Fig. 4. The complete data set consists of 51 sources and 51 receivers with 30 m sampling each.

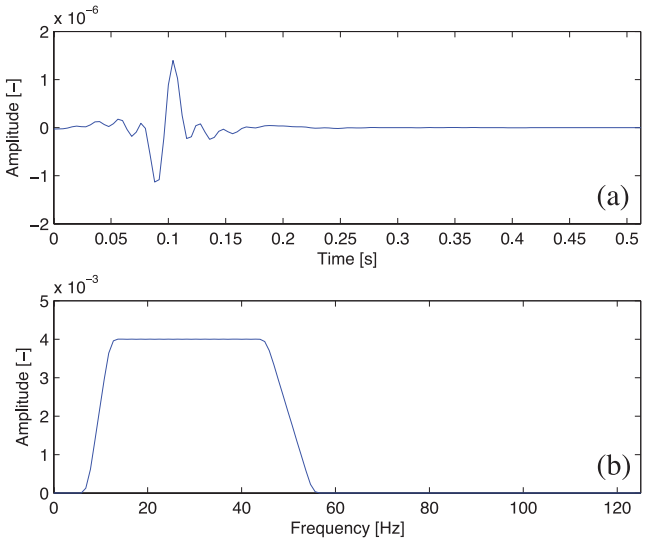

Figure 6. The wavelet used to generate synthetic data is displayed in the time- (a) and frequency-domain (b). Our inversion scheme utilizes the full bandwidth of the seismic data to obtain high resolution property images of the subsurface.

6.1 Localized linear inversions

The very first step in our inversion approach is to perform local broad-band inversions by assuming the total field to be equivalent to the background field in the background medium. Hence, we initialize the scheme with localized linear inversions, where the local domains overlap by a factor of two in depth and in the lateral direction. Finally, the results can be combined by weighted averaging to a global model. In this example we split up the entire subsurface into 21 local domains. The fact that the local inversions are independent of each other allows us to obtain many local results in a rather short time, by parallelization. The total number of local domains consists of seven lateral domains and three different depth levels starting at 0, 150 and 300 m, respectively. The overlap ensures that, especially during the first iterations, where the linear approximation is used and inter domain effects are neglected, still a reliable property model can be obtained. For inversion the full seismic bandwidth is used and the inversion domains are sampled on a 5 m grid in both dimensions.

Localization of the inversion involves back-propagation of the entire data set to the respective depth level, choosing a subset of the data to achieve a lateral localization and time-windowing to get a limitation of the inversion domain with regard to depth. It should be realized that the first layer starting at $z = 0$ m can be inverted locally without back-propagation, but by simply choosing a laterally limited subset of the entire surface data and subsequent time-windowing. In Fig. 7, we show two single shots belonging to a local domain starting from the surface and the third depth level. Although only a single shot for two independent local inversion domains is shown, every local inversion uses 11 shots measured by 21 receivers. In Fig. 8, local inversion results for two independent domains are shown as well as the global model that was obtained by combining all 21 local domains by weighted averaging.

It becomes obvious that the obtained contrast function is quite different from the true model that was shown in Fig. 4. Although the

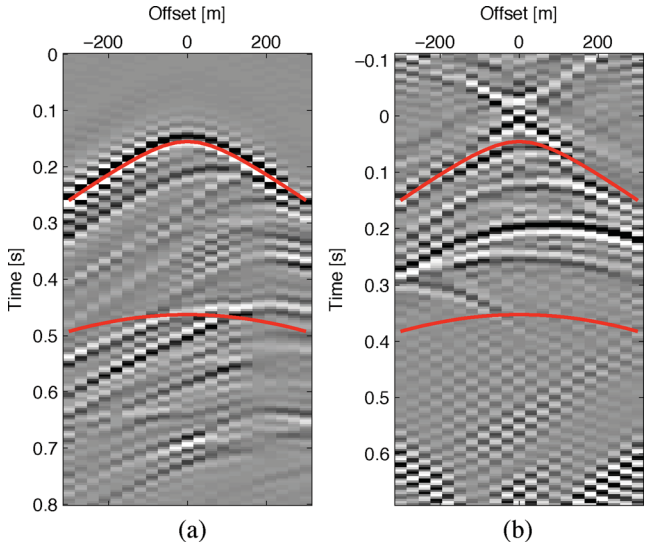

Figure 7. Shots corresponding to local inversion domains starting at $z = 0$ m (a) and $z = 300$ m (b) depth. Note that for local domains which don't start at the surface the data has to be back-propagated. For inversion, only the data within the red time window is used. This primary time window is dependent on the extension of the local domain.

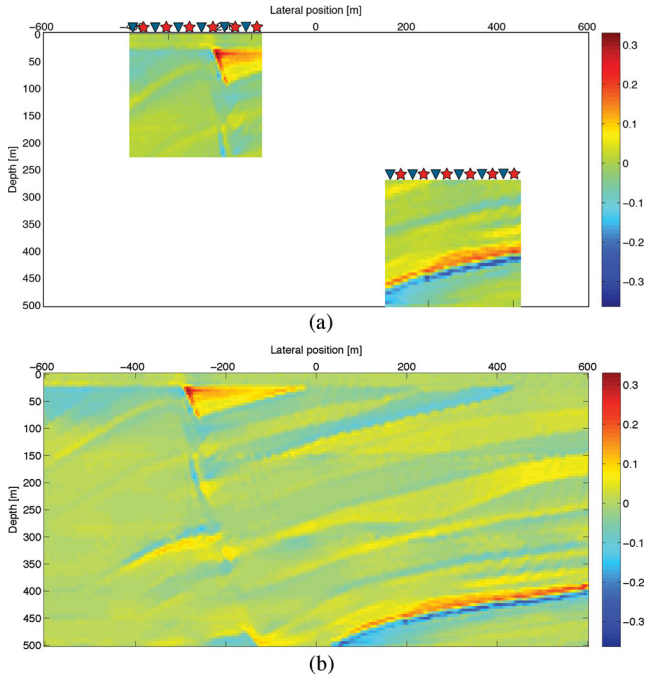

Figure 8. Local inversion results for two separate local domains (a) as well as the global result obtained by combining 27 local domains (b). Since the first inversion assumes the total field in the subsurface to be the background field in the background medium and because interaction between local domains was neglected an unsatisfactory result is obtained.

structure of the subsurface was recovered reasonably well, the actual properties of the subsurface could not be estimated accurately. There are two main reasons for this observation. First, the initial inversion of the iterative scheme, as it was sketched in Fig. 1, assumes that $p_{\text{tot}} = p_0$ in eq. (6). This is clearly an approximation of reality and this obviously shows up in the inversion result. Since the data was modelled by solving the full wave equation it contains the non-linear relationship between the subsurface properties and the measured fields and an inversion that assumes a linear relationship does not lead to a satisfactory result. Second, the first iteration does not take any interaction between the local domains into account. By inter domain interaction we refer to the propagation of the fields through the overburden and to multiple scattering between different local domains. In real life field propagation always takes place in the global medium and we have to account for that fact if an accurate inversion result is to be obtained. We can do this by performing a global field update based on eq. (3).

6.2 Global non-linear field update

Previously we have seen that a linear inversion, which is based on the Born approximation, was able to recover structural information of the subsurface while quantitative information, the actual property values, were inaccurate. In this section we show that this objective can be achieved by FWI. Non-linear inversion honours the non-linear relationship between the subsurface properties and the measured seismic data. Non-linearity in seismic data can show up as different wave phenomena, where the most well known one is multiple scattering. Additionally, transmission effects and true traveltimes as opposed to background traveltimes can be interpreted as non-linear effects. Because in the first iteration the subsurface was unknown there was no other choice than using the background field as the best approximation of the exact total field. However, since with this approximation we obtained a first estimate of the subsurface properties we can subsequently update the total field, based on the latest obtained subsurface model. This can in principle be performed by substituting the obtained contrast function χ into eq. (3) and solving with any suitable numerical method for p_{tot}. Since this would be a computationally rather expensive task, we build up the total field as a sum of basis function and after every iteration only a single basis function, containing only one order of scattering each, is added. It is most important to realize that the field update is always a global process, which uses the global model coming from the combination of the latest local inversion results. Only in this way the field propagation as it occurs in the real medium, can accurately be estimated. The initial and updated total field after 35 iterations and for a single depth level are displayed in Fig. 9. The field update is always performed for the original source geometry, which is in most cases along the surface and for subsequent localized inversion the updated fields have to be back-propagated into the subsurface again. The total field as it is shown in Fig. 9 was obtained after 35 iterations. This means that it is a sum of 35 basis functions of which each contains information of the property model that was obtained during the respective iteration. In Fig. 10 we plot the final property model that was obtained by combining 27 local inversions but now based on the total field after the last iteration.

It is interesting to realize that the local results still come from a linear inversion of eq. (6). The improvement with respect to the linear inversion result that was shown in Fig. 8 comes solely from the fact that the total field honours the non-linear relationship with the subsurface properties. The same reasoning also explains that although the homogeneous background is not changed during the procedure, still a very good result can be obtained. Using the full non-linearity in the seismic data allows us to reconstruct spatial wave numbers that are outside the equivalent temporal frequencies in the seismic data, consequently closing the spectral gap (Haffinger *et al.* 2011).

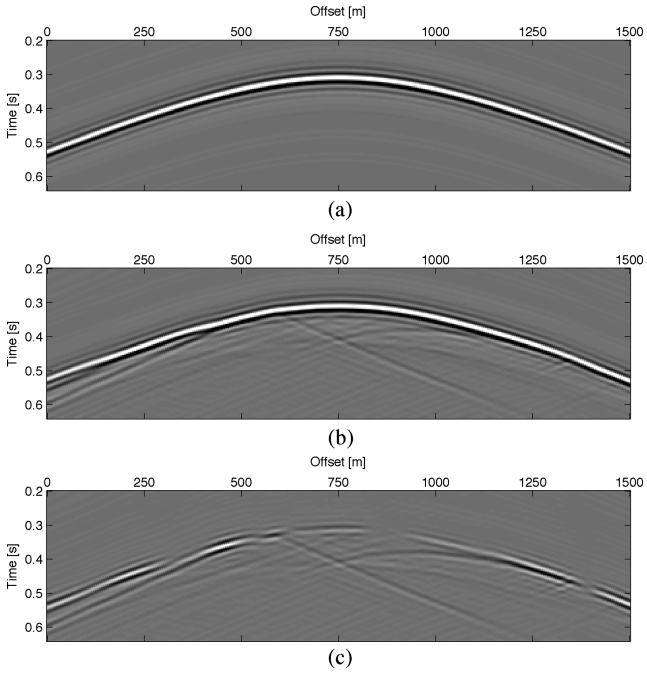

Figure 9. The background field at $z = 400$ m (a) is very simple since it propagated in the homogeneous background medium only. After 35 iterations of alternating local inversions followed by a global field update (b), the total field becomes much more complex since it describes the true field propagation in the best-known medium. Updating the total field estimates multiple scattering as well as the correct traveltimes and amplitudes in the inverted subsurface model. This becomes clear by plotting the difference in (c).

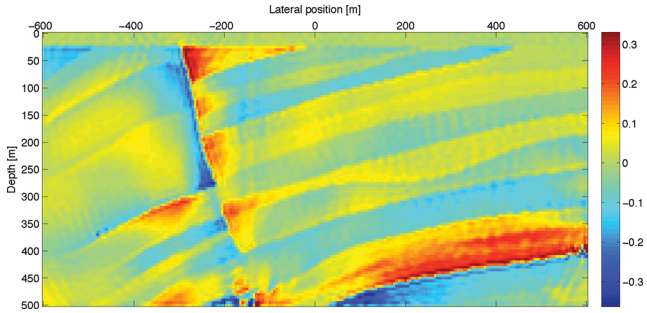

Figure 10. Final contrast function coming from our non-linear inversion scheme. Note how well the structure is reconstructed but also how the quantitative information of the subsurface is nicely estimated. The fault in the model is clearly visible, which demonstrates that the resolution of non-linear inversion is far beyond what can be obtained by linear imaging/inversion techniques.

6.3 Overburden and inter domain multiples

Finally we illustrate how overburden multiples and inter-domain scattering were estimated and subtracted from the surface data before localization. The data we show in Fig. 11 belongs to a local domain that starts at $z = 300$ m depth and with the centre being located at $x = 150$ m. For the first iteration the entire measured surface data was back-propagated to localize the inversion, because initially the contrast function is unknown and we are not able to subtract contributions from the complement of the local domain. After the first iteration we can use the obtained contrast function and the initial total field $p_{\text{tot}} = p_0$ to estimate the part of the data that does not belong to the local domain, but comes from the overburden or the surrounding of the local domain. This can be done

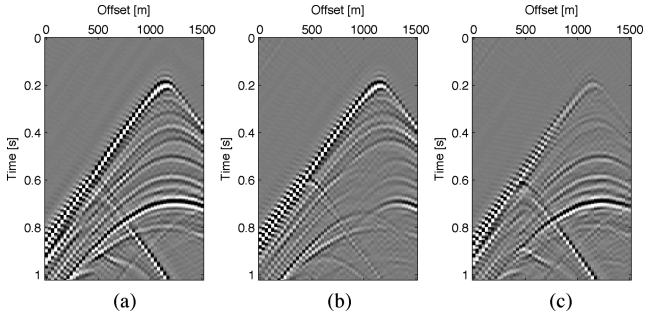

Figure 11. Surface data for a single shot at the surface, located just above a local domain starting at $z = 300$ m and with the centre being located at $x = 150$ m (a). The complementary data (b) obtained by integrating over gridpoints after the first iteration, but over the domain $\mathcal{D}_\cap$ only. Difference between (a) and (b) which is the current best estimate of the data associated with the local domain $\mathcal{D}_\cup$, is shown in (c). Only the difference will be used for the next local inversion of this specific domain. Note that for other domains the complementary data will differ but the concept will be equally valid.

by modelling surface data while integrating over gridpoints in the domain $\mathcal{D}_\cap$ (the complement of the local domain) only. Note that this is equivalent to the second term on the RHS of eq. (17). The corresponding data is shown in Fig. 11. It is clear that the first estimate of the complementary data explains the surface data partly, but that quite some residual energy remains. Therefore, we run the inversion for 35 iterations, in which we update the contrast function and the total fields in the subsurface alternatingly. In Fig. 12, we show the same panels as in Fig. 11 but after 35 iterations. It is obvious that after 35 iterations, the complementary data matches the surface data much better. Wave fields reflected in the overburden are nearly perfectly estimated and subtracted. Please note the event going obliquely through the data, being generated by the fault. After 35 iterations it is nicely predicted and subtracted because it is not generated within the local domain we want to invert.

It is worthwhile mentioning that by performing a global field update and subsequently subtracting the complementary data, an exact inversion scheme is obtained, in which the localization takes only the role of pre-conditioning. It also becomes clear that if the complementary data can be estimated very accurately, time windowing the back-propagated data is not necessary any more and even the

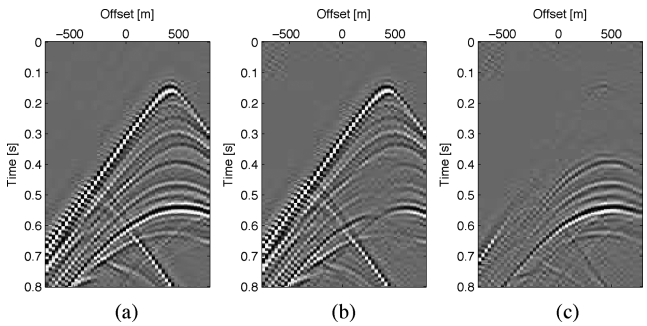

Figure 12. Same data as shown in Fig. 11 but now after 35 iterations. Note how well the complementary data (b) matches the measured data (a) in the parts that do not belong to the local domain. The difference (c) that is used for subsequent inversions of the particular local domain only contains events that are related to the local domain. Note how the linear event, associated with the fault in the subsurface, was nicely predicted and removed after 35 iterations. This is a desired result since the fault lies outside the local domain under consideration and reflections from this part of the subsurface should not be taken into account by local inversion.

coda that arrives outside the primary time window can be used for inversion. This is possible because the interference of the coda from a local domain with reflections from the subsurface below the local domain will be subtracted as well.

7 COMPUTATIONAL COMPLEXITY

Before getting to the conclusions, we compare the computational complexity of direct inversion of surface data to the proposed localized scheme. Using the variables given in Table 2 the number of computations for direct inversion of surface data becomes $Computations_{\text{surface}} \approx O(N_s N_r N_{fg} N_{\chi g} N_{CG}) + O(N_s N_{fm} N \log N N_{FU}(N_{FU}+1)/2)$. The first term on the RHS accounts for the inversion of the data equation with the help of a CG-scheme and the second term is related to the field updates. The complexity of the localized scheme described in this paper then becomes $Computations_{\text{local}} \approx O(N_{bs} N_{br} N_{fl} N_{\chi l} N_{CG} N_l) + O(N_s N_{fm} N \log N N_{FU}(N_{FU}+1)/2)$. In both cases, the second term is the same because the field update is always applied to the global domain and no localization is involved. Computationally efficiency of the second term could be improved by conventional domain decomposition methods, which usually aim at speeding-up the forward problem. Nonetheless, this is not within the scope of this paper.

Instead, the proposed localization acts on the inverse problem, described by the first term on the RHS. While the first four factors are significantly decreased by our scheme, a new factor based on the number of local domains gets introduced. If the number of calculations is decreased by the described scheme depends on the number of local domains and hence the chosen overlap. Again, localized inversion does not aim at making FWI more efficient but at making it feasible on a much finer grid while taking the full seismic bandwidth into account. That requires a closer investigation of the memory requirements. Memory is mainly needed for the inversion because computation of the field updates can be heavily parallelized. For direct inversion of surface data, simultaneous memory required is in the order of $Memory_{\text{surface}} \approx O(N_s N_r N_{fg} N_{\chi g})$ while for localized inversion it can be significantly reduced to $Memory_{\text{local}} \approx O(N_{bs} N_{br} N_{fl} N_{\chi l})$. The strength of localized inversion comes from the fact that all numbers in $Memory_{\text{local}}$ are smaller than their equivalents in $Memory_{\text{surface}}$ and that the factor N_l drops out because all local domains are independent and can therefore be inverted sequentially or in parallel.

8 CONCLUSIONS

In this paper a novel FWI approach was proposed and successfully illustrated on synthetic data modelled above a small subset of the Marmousi2 model. It was shown that by localizing the inversion process, global property models of the subsurface can be obtained with a resolution far beyond current standards. The non-linear relationship between the subsurface properties and the wave fields was successfully accounted for by a total field update. The scheme was able to estimate internal multiples, transmission and the true traveltimes in the obtained model in an iterative manner. By using the updated fields a non-linear inversion result was obtained with structural and quantitative property information much more accurate than it could be recovered by a linear scheme, for example, linear inversion or migration.

Table 2. Variables used for complexity analysis.

N_s	=	number of sources along surface,
N_r	=	number of receivers along surface,
N_{bs}	=	number of back-propagated sources above local domain,
N_{br}	=	number of back-propagated receivers above local domain,
N_{fm}	=	number of frequencies of measured surface data,
N_{fg}	=	number of frequencies used for global inversion,
N_{fl}	=	number of frequencies used for local inversion,
$N_{\chi g}$	=	number of unknowns in global domain,
$N_{\chi l}$	=	number of unknowns in local domain,
N_l	=	number of local domains,
N_{CG}	=	number of CG-iterations,
N_{FU}	=	number of total field updates.

ACKNOWLEDGEMENTS

The authors thank the sponsors of the Delphi consortium for their continuing support. Furthermore we appreciate the comments of two anonymous reviewers that significantly increased the quality of this paper.

REFERENCES

Abubakar, A., van den Berg, P.M., Habashy, T.M. & Braunisch, H., 2004. A multiplicative regularization approach for deblurring problems, *IEEE Trans. Image Process.,* **13,** 1524–1532.

Abubakar, A., Hu, W., Habashy, T.M. & van den Berg, P.M., 2009. Application of the finite-difference contrast-source inversion algorithm to seismic full-waveform data, *Geophysics,* **74**(6), WCC163–WCC174.

Abubakar, A., Habashy, T.M. & Pan, G., 2011. Source-receiver compression approach for 3D full-waveform inversion with an iterative forward solver, in *Proceedings of the 81st Conference and Exhibition of Soc. Expl. Geophys., Expand. Abstracts,* San Antonio.

Bamberger, A., Chavent, G. & Lailly, P., 1977. Une application de la theórie du contrôle à un problème inverse sismique, *Ann. Geophys.,* **33,** 183–200.

Bamberger, A., Chavent, G., Hemons, C. & Lailly, P., 1982. Inversion of normal incidence seismograms, *Geophysics,* **47,** 757–770.

Bamberger, A., Glowinski, R. & Tran, U.H., 1997. A domain decomposition method for the acoustic wave equation with discontinuous coefficients and grid change, *SIAM J. Numer. Anal.,* **34**(2), 603–639.

Berryhill, J.R., 1984. Wave equation datuming before stack (short note) *Geophysics,* **49**(11), 2064–2066.

Brossier, R., Operto, S. & Virieux, J., 2009. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion, *Geophysics,* **74,** WCC63–WCC76.

Bunks, C., Saleck, F.M., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion *Geophysics,* **60**(5), 1457–1473.

Crase, E., Pica, A., Noble, M., McDonald, J. & Tarantola, A., 1990. Robust elastic non-linear waveform inversion: application to real data, *Geophysics,* **55,** 527–538.

Epanomeritakis, I., Akelik, V., Ghattas, O. & Bielak, J., 2008. A Newton-CG method for large-scale three-dimensional elastic full waveform seismic inversion, *Inverse Problems,* **24,** 1–26.

Fichtner, A. & Trampert, J., 2011. Hessian kernels of seismic data functionals based upon adjoint techniques *Geophys. J. Int.,* **185,** 775–798.

Fokkema, J.T. & van den Berg, P.M., 2012. Stretched backgrounds for acoustic scattering models, *J. Comput. Phys.,* **231,** 1728–1742.

Gisolf, D. & van den Berg, P.M., 2012. Target-oriented elastic full wave form inversion, in *Proceedings of the 74th Ann. Internat. Mtg., Eur. Ass. of Geosc. and Eng., Expanded Abstracts,* Copenhagen.

Haffinger, P., Gisolf, D. & van den Berg, P.M., 2011. Towards broadband non-linear full-waveform inversion with the help of shot/receiver refocussing, in *Proceedings of the 81st Conference and Exhibition, Soc. Expl. Geophys., Expanded Abstracts,* San Antonio.

Kleinman, R.E. & van den Berg, P.M., 1991. Iterative methods for solving integral equations, *IEEE Trans. Image Process.,* **26**(1), 175–181.

Krebs, J.R., Anderson, J.E., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A. & Lacasse, M.D., 2009. Fast full-wavefield seismic inversion using encoded sources, *Geophysics,* **74**(6), WB41–WB51.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations, in *Proceedings of the Conference on Inverse Scatterring: Theory and Application,* Society of Industrial and Applied Mathematics.

Lippmann, B.A. & Schwinger, J., 1950. Variational principles for scattering processes. I, *Phys. Rev.,* **79,** 469.

Métivier, L., Brossier, R., Virieux, J. & Operto, S., 2012. Toward Gauss-Newton and exact newton optimization for full waveform inversion, in *Proceedings of the 74th Ann. Internat. Mtg., Eur. Ass. of Geosc. and Eng., Expanded Abstracts,* Copenhagen.

Moghaddam, P.P. & Herrmann, F.J., 2010. Randomized full-waveform inversion: a dimenstionality-reduction approach, in *Proceedings of the 80th Conference and Exhibition, Soc. Expl. Geophys., Expanded Abstracts,* Denver.

Mora, P.R., 1987. Nonlinear two-dimensional elastic inversion of multi-offset seismic data, *Geophysics,* **52,** 1211–1228.

Mulder, W.A., 2005. Rigorous redatuming *Geophys. J. Int.,* **161,** 401–415.

Plessix, R.E., Baeten, G., de Maag, J.W., Klaassen, M., Rujie, Z. & Zhifei, T., 2010. Application of acoustic full waveform inversion to a low frequency large offset land data set, in *Proceedings of the at the 80th Conference and Exhibition, Soc. Expl. Geophys.,* Denver.

Pratt, R.G., Shin, C. & Hicks, G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion *Geophys. J. Int.,* **133,** 341–362.

Prieux, V., Brossier, R., Kommendal, J.R., Barkved, O.I., Operto, S. & Virieux, J., 2010. Application of 2D acoustic frequency-domain full-waveform inversion to OBC wide-aperture data from the Valhall field, in *Proceedings of the 80th Conference and Exhibition, Soc. Expl. Geophys., Expanded Abstracts,* Denver.

Routh, P. *et al.*, 2011. Encoded simultaneous source full- wavefield inversion for spectrally shaped marine streamer data, in *Proceedings of the the 81st Conference and Exhibition, Soc. Expl. Geophys., Expanded Abstracts,* San Antonio.

Shtivelman, V. & Canning, A., 1988. Datum correction by wave equation extrapolation, *Geophysics,* **53**(10), 1311–1322.

Soni, A.K., Wouters, W. & Verschuur, D.J., 2012. Target oriented VSP imaging—a sparse-inversion approach, in *Proceedings of the 74th Ann. Internat. Mtg., Eur. Ass. of Geosc. and Eng., Expanded Abstracts,* Copenhagen.

Staal, X.R., Verschuur, D.J., Gisolf, A., Tonellot, T. & Burnstad, R., 2010. Improved target-oriented linear full waveform inversion, in *Proceedings of the 72th Ann. Internat. Mtg., Eur. Ass. of Geosc. and Eng., Expanded Abstracts,* Barcelona.

Symes, W., 1981. Stable solution of the inverse reflection problem for a stratified elastic medium, *SIAM, J. Math. Anal.,* **12**(3).

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

Tarantola, A., 1987. *Inverse Problem Theory, Methods for Data Fitting and Model Parameter Estimation,* Elsevier Science Publ. Co., Inc.

Valenciano, A., Biondi, B. & Guitton, A., 2006. Target-oriented wave-equation inversion, *Geophysics,* **71,** A35–A38.

Vigh, D., Kapoor, J., Moldoveanu, N. & Hongyan, L., 2011. Breakthrough acquisition and technologies for subsalt imaging, *Geophysics,* **76**(5), WCC177–WCC188.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics,* **74**(6), WCC1–WCC26.

Wapenaar, C.P.A. & Berkhout, A.J., 1989. *Elastic Wave Field Extrapolation: Redatuming of Single- and Multi-Component Seismic Data,* Elsevier Science Publ. Co., Inc.