Multi-level, Multi-frequency, compressed Full Waveform Inversion

Florian Faucher, Maarten V. de Hoop, Henri Calandra, Lingyun Qiu, Christian Rivera

Proceedings of the Project Review, Geo-Mathematical Imaging Group (Purdue University, West Lafayette IN), Vol. 1 (2013) pp. 1-?.

MULTI-LEVEL, MULTI-FREQUENCY, COMPRESSED FULL WAVEFORM INVERSION*

FLORIAN FAUCHER†, MAARTEN V. DE HOOP‡, HENRI CALANDRA§, LINGYUN QIU¶, AND CHRISTIAN RIVERA‖

Abstract. We study the inverse boundary value problem for the wave equation and recovery of the wavespeed upon taking a time-Fourier transform of the data. We design a hierarchical compressed reconstruction in a multi-level scheme for the inverse boundary value problem associated with the Helmholtz equation using the Dirichlet-to-Neumann map, or the single-layer potential operator, at selected frequencies as the data. The compression is based on a domain partitioning of the subsurface, while the hierarchy is straightforwardedly established through refinement. The coefficients are assumed to be piecewise constant functions following the domain partitioning, thus allowing the presence of conormal singularities (causing reflections and diffractions); these guarantee a Lipschitz stability estimate for the inverse problem which gives a radius of convergence, related to the stability constant, of the scheme even in the case of partial data. The stability constant grows exponentially with the number of subdomains in the domain partitioning, whence the compression. The frequencies are selected to control the approximation errors in the data due to the compression at the different levels in the scheme. The stability constants (number of subdomains) and approximations errors (frequencies) are coupled through a condition between levels for convergence of the multi-level scheme. We use Haar wavelets for the compression in alternative strategies providing a gradual increase in the number of subdomains in the domain parititioning via adaptive, local multi-scale refinement. We include visco-acoustic behavior and carry out numerical experiments.

1. Introduction. We study the inverse boundary value problem for the wave equation and recovery of the wavespeed, including attenuation, upon taking a time-Fourier transform of the data. We follow a hierarchical compressed reconstruction in a multi-level scheme for the inverse boundary value problem associated with the Helmholtz equation using the Dirichlet-to-Neumann map as the data [16]. The effectiveness and convergence of this scheme was analysed in [7, 8]. The wavespeed can be complex and frequency dependent and, hence, visco-acoustic behaviour is included. Density can, in principle, be recovered via the reconstruction of a potential (using its frequency dependency) in the corresponding Schrödinger equation. We do not address the inverse source (signature) problem here.

Iterative methods for the above mentioned inverse boundary value problem have been collectively referred to, in reflection seismology, as full waveform inversion (FWI). (The term 'full waveform inversion' was supposedly introduced in [31] with reference to the use of full seismograms information.) Lailly [24] and Tarantola [43, 44] introduced the formulation of the seismic inverse problem as a local optimization problem encompassing a least-squares (L^2) minimization of a misfit functional. We also mention the original work of Bamberger, Chavent & Lailly [3, 4] in the one-dimensional case. Tarantola noticed that the evaluation of the gradient of the mentioned functional corresponds with an imaging procedure with a cross correlation imaging condition. Initial numerical experiments in the two-dimensional case were carried out by Gauthier [19]. Since then, a range of alternative misfit functionals have been considered; we mention, here, the criterion derived from the instantaneous phase [11].

The time-harmonic formulation was initially promoted by Pratt and other collaborators in [37, 34] and specifically for wide-angle reflection data in [36]. The Newton method using the discretized Helmholtz equation as a point of departure was developed by Pratt, Shin & Hicks [35]. Later the use of complex frequencies was studied in [38, 20]; misfit functions using a logarithmic norm were

*This research was supported by Total.

†Department of Mathematics, Purdue University 150 N. University Street, West Lafayette IN 47907, USA (ffaucher@math.purdue.edu).

‡Department of Mathematics, Purdue University 150 N. University Street, West Lafayette IN 47907, USA (mdehoop@purdue.edu).

§Total Research Exploration & Production, Pau, France (henri.calandra@total.com).

¶Department of Mathematics, Purdue University 150 N. University Street, West Lafayette IN 47907, USA (qiu@math.purdue.edu).

‖Total Research Exploration & Production, Houston, TX 77002 (christian.rivera@total.com).

explored in [44, 40, 41].

The nonlinearity of FWI has motivated studies to develop hierarchical multiscale strategies to mitigate this nonlinearity. In the time domain, Bunks *et al.* [13] proposed successive inversion of data subsets of increasing frequency contents following the intuition that low frequencies are less sensitive to cycle skipping. This multiscale approach can, for example, be related to the subspace search method advocated by Kennett, Sambridge & Williamson [22]. The time-harmonic formulation, however, has been recognized as the more natural one for this purpose. The idea of frequency progression has been considered, amongst others, by Sirgue & Pratt [42] and extensively studied in the electromagnetic waves case, for example, by Bao and Li [5, 6]. The application of wavelet bases in compressing the successive models in the iteration has been considered by [27, 26] in wave-equation tomography (sparsity promoting optimization) and in FWI by Lin, Abubakar & Habashy [25] (for the purpose of reducing the size of the Jacobian). In our approach the compression, in particular, the compression rate, plays a fundamental role in guaranteeing convergence of our multi-level scheme.

Malinowski, Operto & Ribodetti [28] demonstrate high-resolution imaging of attenuation and phase velocity in the visco-acoustic case. In this context, we also mention the work of Askan, Akcelik, Bielak & Ghattas [2]. Our convergence analysis applies to complex wavespeeds.

In FWI one commonly applies a 'nonlinear' conjugate gradient method, a Gauss-Newton method, or a quasi-Newton method (L-BFGS; for a review, see Brossier [12]). For the application of multi-scale Newton methods, see Akcelik [1]. In this paper, we carry out numerical experiments with the nonlinear conjugate gradient method. In the gradient method, the step length is typically estimated by a simple line search for which a linearization of the direct problem is used (Gauthier, Virieux & Tarantola [19]). This estimation is challenging in practice and may lead to a failure of convergence. For this purpose, research based on trust region has been studied by Eisenstat & Walker in [17], for FWI see Métivier and others in [30], the method is detailed in [15]. In various approaches one accounts for the Hessian, at least its diagonal, in the Gauss-Newton scheme [39]. The second term in the Hessian (Newton method) can be interpreted in terms of multiple scattering (see, for example, Pratt, Shin & Hicks [35]). In certain earthquake seismology applications, one builds the Fréchet derivative or Jacobian (sensitivity, for example, [14]) explicitly and then applies LSQR.

1.1. Time-harmonic waves. We formulate FWI in the frequency domain as an inverse boundary value problem for the Helmholtz equation. We consider a bounded domain Ω of $\mathbb{R}^3$ representing the region of interest in the subsurface; the data are acquired on an open portion of the boundary, $\partial\Omega$. Time-harmonic waves are solutions, u, of

$$\begin{cases} -\left(\Delta + q(x)\right) u = 0 \text{ on } \Omega, \\ \qquad\qquad\quad u = g \text{ on } \partial\Omega. \end{cases} \tag{1.1}$$

Here, $q = q(x) = k(x)^2 = \omega^2 c(x)^{-2}$ with ω representing the (possibly complex) angular frequency and $c = c(x)$ the wavespeed; $g = g(x, \omega)$ is a boundary source. For the discretization of the Helmholtz operator we employ a finite difference method with a compact stencil, and a massively parallel structured multifrontal solver, *Hsolver* [46, 47, 48]. This solver is naturally designed for solving the Helmholtz equation for many right-hand sides.

Attenuation. To incorporate visco-acoustic behaviour of subsurface materials, we introduce complex wavespeeds leading to attenuation. For a comprehensive review of models of attenuation, we refer to [45]. Suppressing the x-dependencies in the notation, the wave number k becomes:

$$k(\omega) = \frac{\omega}{c(\omega)} = \frac{\omega}{c_p(\omega)} + \mathrm{i}\alpha(\omega),$$

where c_p represents the phase velocity. We consider the Kolsky-Futterman model ([23], [18]) which has been widely used in applications and is given by

$$\begin{cases} \dfrac{1}{c_p(\omega)} = \dfrac{1}{c_r} + \dfrac{\log\left|\dfrac{\omega_r}{\omega}\right|}{\pi c_r Q_r}, \\ \alpha(\omega) = \dfrac{|\omega|}{2c_r Q_r}, \end{cases}$$

where c_r and Q_r represent the values of the the phase velocity c_p and the quality factor Q at a reference angular frequency, ω_r. If $\alpha c_p \ll 2|\omega|$,

$$Q(\omega) = Q_r + \frac{1}{\pi}\log\left|\frac{\omega_r}{\omega}\right|.$$

If Q and c_p were both frequency independent, then $c_p = c_r$, $Q_r = Q$, and the equation for the complex wavespeed may be simplified to

$$c(\omega) = \frac{4c_pQ^2 - 2\mathrm{i}\,\mathrm{sgn}(\omega)\,c_pQ}{4Q^2+1}.$$

In the case of weak attenuation, $4Q^2 \gg 1$, the expression on the right-hand side can be approximated as

$$c(\omega) = c_p\left(1 - \mathrm{i}\,\mathrm{sgn}(\omega)\,\frac{1}{2Q}\right). \tag{1.2}$$

This simplified representation was used in [28] while studying full waveform inversion with attenuation.

1.2. Modelling the data and mismatch functional. The data generated by the boundary sources, g, can be represented by the so-called single layer potential operator, S_q. That is,

$$S_q:\ g \mapsto \int_{\partial\Omega} G_q(x,y)g(y)\,\mathrm{d}\sigma(y), \quad x \in \partial\Omega, \tag{1.3}$$

where G_q is the fundamental solution of $-(\Delta + q(x))$ in $\mathbb{R}^3$. This operator is discussed and analyzed in [32]. We note that the single-layer potential operator and the so-called Dirichlet-to-Neumann map represent equivalent data [32]. The action of S_q has been referred to as source blending [9]. The forward problem is then described by

$$F:\ \frac{\omega^2}{c(x)^2} = q(x) \mapsto S_q$$

In [32] it was shown that $S_{q_1} - S_{q_2}$ is a Hilbert-Schmidt (HS) operator if $c_2(x) = c_1(x)$ in a small neighborhood of $\partial\Omega$. Thus, if $q^\dagger$ represents the 'true' model and, hence, $F(q^\dagger) = S_{q^\dagger}$ the data, we can introduce the mismatch functional

$$\mathcal{J}(q) = \tfrac{1}{2}\|S_q - S_{q^\dagger}\|^2 \tag{1.4}$$

using the HS norm

$$\|S_q - S_{q^\dagger}\| = \left[\sum_{j=1}^{\infty}\|(S_q - S_{q^\dagger})\psi_j\|^2\right]^{1/2}, \tag{1.5}$$

where $\{\psi_j\}$ is a countable basis in $H^{1/2}(\partial\Omega)$, and could be chosen to be right singular functions of the residual operator. Note that the norm of the residual operator, $\|S_q - S_{q^\dagger}\|$, does not depend on the choice of basis $\{\psi_j\}$. As an alternative, the log criterion replacing

$$\|(S_q - S_{q^\dagger})\psi_j\|^2 \quad \text{by} \quad \|\log(S_q\psi_j) - \log(S_{q^\dagger}\psi_j)\|^2,$$

has been proposed [44, 40]. This is well-defined considering $S_q - S_{q^\dagger}$ as a densely defined operator and using a basis $\{\psi_j\}$ of $H^{3/2}(\Omega)$. Intuitively, the log criterion may capture better the components of the residual which belong to the small singular values; however, it does not give a proper norm for the residual operator. We will consider both functionals in our examples, where we will truncate the sum at N, a fixed number of terms.

2. Iterative scheme.

2.1. Gradient descent. We consider time-harmonic waves, described by solutions, u say, of the Helmholtz equation on $\mathbb{R}^n$, $n \geq 3$. We write

$$q(x) = -\omega^2 c^{-2}(x, \omega);$$

$\omega \in \mathbb{R}$. We introduce a bounded open (computational) domain $\Theta \subset \mathbb{R}^n$ with boundary in $C^{(1,1)}$. We assume that $\Theta' = \mathbb{R}^n \backslash \overline{\Theta}$ is connected. We consider the problem,

$$\left\{ \begin{array}{rcll} (-\Delta - \quad k^2)u & = & 0, \quad x \in \Theta' & , \\ (-\Delta + q(x))u & = & f, \quad x \in \Theta & , \\ \lim\limits_{r\to\infty} r^{(n-1)/2}\left(\dfrac{\partial u}{\partial r} - \mathrm{i}ku\right) & = & 0 & . \end{array} \right. \tag{2.1}$$

Here, $q \in L^\infty(\Theta)$ while $-k^2 = -\omega^2 c_\infty^{-2}$, where c_∞ is constant.

We replace (2.1) by the equivalent problem on Θ,

$$\left\{ \begin{array}{rcll} (-\Delta + q(x))u & = & f, \quad x \in \Theta & , \\ \dfrac{\partial u}{\partial \nu} & = & -\Lambda_e u, \quad x \in \partial\Theta & , \end{array} \right. \tag{2.2}$$

where Λ_e is the exterior Dirichlet-to-Neumann map for the Helmholtz equation on Θ'. We introduce an open bounded domain, D, with $D \subset \Theta$ and a boundary in $C^{(1,1)}$, and $D' = \mathbb{R}^n \backslash \overline{D}$ being connected. Let Σ be an open portion of ∂D. In the above, furthermore, f represents volume sources supported in ∂D; this property will be implicit through the introduction of an appropriate Dirac measure and factoring out ψ (defined on ∂D) in f .

The solution, u, of (2.2), in $H^1(\Theta)$, also solves the variational equation

$$b(q; u, v) = s(v) \quad \text{for all } v \in H^1(\Theta), \tag{2.3}$$

where

$$b(q; u, v) = \int_\Theta (\nabla u \cdot \nabla \bar{v} - q\, u\bar{v})\, \mathrm{d}x, \tag{2.4}$$

and

$$s(v) = \int_\Theta f\, \bar{v}\, \mathrm{d}x - \int_{\partial\Theta} (\Lambda_e u)\, \bar{v}\, \mathrm{d}\sigma. \tag{2.5}$$

We consider the one-parameter family of models, q_t, with $q_0 = q$, and a set of source functions, $(f_k(x))_{k=1}^N$; we collect $(u_{k,t}(x))_{k=1}^N$ satisfying

$$b(q_t; u_{k,t}, v) = s_k(v) \quad \text{for all } v \in H^1(\Theta), \tag{2.6}$$

with

$$s_k(v) = \int_\Theta f_k\,\bar{v}\,\mathrm{d}x - \int_{\partial\Theta} (\Lambda_e u)\,\bar{v}\,\mathrm{d}\sigma. \tag{2.7}$$

(The first integral on the right-hand side is effectively an integral over ∂D.) We write the solution, $u_{k,t}$, restricted to ∂D as an operator, $\mathcal{S}(q_t)$, acting on f_k, and introduce the functional,

$$\mathcal{J}(q_t) = \frac{1}{2}\sum_{k=1}^{N} \|\mathcal{P}\left(\mathcal{S}(q_t)f_k - d_{f,k}\right)\|^2_{L^2(\partial D)}; \tag{2.8}$$

here, $(d_{f,k}(x))_{k=1}^N$ represents the data, and $\mathcal{P}$ is an elliptic operator such that $\|\mathcal{P}\,d\|_{L^2(\partial D)} = \|d\|_{H^{3/2}(\partial D)}$.

Taking the derivative with respect to t yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}(q_t)\,\Big|_{t=0} = \sum_{k=1}^{N} \mathrm{Re}\,\int_\Sigma [\mathcal{P}^*\mathcal{P}(\mathcal{S}(q_t)f_k - d_{f,k})]\,\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(q_t)f_k\,\Big|_{t=0}\,\mathrm{d}\sigma. \tag{2.9}$$

We introduce $v = (v_k(x))_{k=1}^N$, Lagrange multipliers $(w_k(x))_{k=1}^N$ and a family of functionals,

$$\mathcal{L}(q_t; v, w) = \mathcal{J}(q_t) + \sum_{k=1}^{N} \mathrm{Re}\,\left(b(q_t; v_k, w_k) - s_k(w_k)\right), \tag{2.10}$$

where $\mathcal{S}(q_t)f_k$ in $\mathcal{J}(q_t)$ is replaced by v_k. If $v_k = u_{k,t}$ is the solution of the direct problem (2.6),

$$\mathcal{L}(q_t; u_t, w) = \mathcal{J}(q_t) \quad \text{for all } w_k \in H^1(\Theta),\ k = 1, \ldots, N,$$

then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}(q_t; u_t, w)\,\Big|_{t=0} = \sum_{k=1}^{N} \mathrm{Re}\,\Big\{ \Big[\int_\Sigma [\mathcal{P}^*\mathcal{P}(\mathcal{S}(q_t)f_k - d_{f,k})]\,\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(q_t)f_k\,\Big|_{t=0}\,\mathrm{d}\sigma \\ + b\left(q; \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(q_t)f_k\,\Big|_{t=0}, w_k\right)\Big] + \frac{\mathrm{d}}{\mathrm{d}t} b(q_t; u_k, w_k)\,\Big|_{t=0}\Big\}\end{aligned}$$

for all $w_k \in H^1(\Theta)$. If w_k solves

$$\begin{aligned} b(q; v, w_k) &= -\int_\Sigma [\mathcal{P}^*\mathcal{P}(\mathcal{S}(q_t)f_k - d_{f,k})]\,v\,\mathrm{d}\sigma \\ &\qquad \text{for all } v \in H^1(\Theta),\ k = 1, \ldots, N, \end{aligned} \tag{2.11}$$

then

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}(q_t, u_t, w)\,\Big|_{t=0} = \sum_{k=1}^{N} \mathrm{Re}\,\frac{\mathrm{d}}{\mathrm{d}t} b(q_t; u_k, w_k)\,\Big|_{t=0}. \tag{2.12}$$

We can write (2.11) in the form of the direct problems:

$$\begin{aligned} b(q; \bar{w}_k, v) &= -\int_\Sigma [\mathcal{P}^*\mathcal{P}(\mathcal{S}(q_t)f_k - d_{f,k})]\,\bar{v}\,\mathrm{d}\sigma \\ &\qquad \text{for all } v \in H^1(\Theta),\ k = 1, \ldots, N; \end{aligned} \tag{2.13}$$

here, $(\bar{w}_k)_{k=1}^N$ is the so-called adjoint state, with $\bar{w}_k \in H^1(\Theta)$ being the solution to

$$\left\{\begin{array}{rcll} (-\Delta - \quad k^2)\bar{w}_k & = & 0, \quad x \in \Theta' & , \\ (-\Delta + q(x))\bar{w}_k & = & [\mathcal{P}^*\mathcal{P}(\mathcal{S}(q_t)f_k - d_{f,k})]\,\delta_\Sigma, \quad x \in \Theta & , \\ \lim\limits_{r\to\infty} r^{(n-1)/2}\left(\dfrac{\partial \bar{w}_k}{\partial r} - \mathrm{i}k\bar{w}_k\right) & = & 0 & . \end{array}\right. \tag{2.14}$$

Again, we have the equivalent problems

$$\left\{\begin{array}{rcll} (-\Delta + q(x))\bar{w}_k & = & [\mathcal{P}^*\mathcal{P}(\mathcal{S}(q_t)f_k - d_{f,k})]\,\delta_\Sigma, \quad x \in \Theta & , \\ \dfrac{\partial \bar{w}_k}{\partial \nu} & = & -\Lambda_e \bar{w}_k, \quad x \in \partial\Theta & . \end{array}\right. \tag{2.15}$$

We have made use of the fact that $\Lambda_e^*(\bar{w}_k) = \overline{\Lambda_e w_k}$. We then evaluate

$$\frac{\mathrm{d}}{\mathrm{d}t} b(q_t; u_k, w_k)\,\Big|_{t=0} = -\int_\Theta u_k \bar{w}_k\, \frac{\mathrm{d}}{\mathrm{d}t} q_t\,\Big|_{t=0}\,\mathrm{d}x. \tag{2.16}$$

We identify

$$\frac{\mathrm{d}}{\mathrm{d}t} q_t\,\Big|_{t=0} \quad \text{with} \quad \gamma.$$

Thus

$$d\mathcal{J}(q,\gamma) = -\sum_{k=1}^N \mathrm{Re}\int_\Theta u_k \bar{w}_k\,\gamma\,\mathrm{d}x =: (\rho,\gamma)_{L^2(\Theta)}. \tag{2.17}$$

In the case of attenuation in $\frac{\mathrm{d}}{\mathrm{d}t}q_t$ above, we can distinguish its real and imaginary parts. Instead, we distinguish between $\frac{\partial q}{\partial c_p}\frac{\mathrm{d}}{\mathrm{d}t}c_{p;t}$ and $\frac{\partial q}{\partial Q}\frac{\mathrm{d}}{\mathrm{d}t}Q_t$, where

$$\left\{\begin{array}{l} \dfrac{\partial q}{\partial c_p} = -\dfrac{2\omega^2}{c_p^3\left(1 - \dfrac{\mathrm{i}\,\mathrm{sgn}(\omega)}{2Q}\right)^2}, \\ \dfrac{\partial q}{\partial Q} = -\dfrac{\mathrm{i}\,\mathrm{sgn}(\omega)\omega^2}{c_p^2 Q^2\left(1 - \dfrac{\mathrm{i}\,\mathrm{sgn}(\omega)}{2Q}\right)^3}. \end{array}\right.$$

A comprehensive review for gradient computations may be found in [35, 33].

2.2. Convergence. The convergence rate and convergence radius of our iterative scheme are derived based on a conditional Lipschitz-type stability estimate for the inverse problem. In [8] it was proven that

$$\|q - q^\dagger\| \le \mathcal{C}\,\|F(q) - F(q^\dagger)\|, \tag{2.18}$$

in particular, for piecewise constant functions,

$$q(x) = \sum_{j=1}^J q_j \chi_{D_j}(x),$$

incorporating interfaces, pinchouts etc., where the q_j are constant coefficients and the χ characteristic functions; here, we pre-assume a domain partitioning $\{D_j\}_{j=1}^J$ of Ω. Moreover, in [8] it was

shown that $\mathcal{C}$ increases exponentially with increasing J. Increasing J implies an increasing spatial resolution.

The convergence radius is also determined by the lower bound of $\|DF(q^\dagger)\|$, where DF signifies the Fréchet derivative of F. Indeed, such a bound implies the above mentioned Lipschitz stability estimate. In the case of weak attenuation, it is clear that the lower bound of the Fréchet derivative with respect to Q decreases significantly, which implies a rapidly increasing stability constant, with decreasing frequency.

We mitigate the above mentioned growth of the stability constant by introducing a multi-level approach based on a hierarchical compression of the model during the reconstruction. The compression at any level leads to an approximation error: Let $q_J^\dagger$ denote the compressed approximation of $q^\dagger$ using J domains, then the error, $\|F(q_J^\dagger) - F(q^\dagger)\|$, decreases as ω^2 as $\omega \downarrow 0$. A condition on the coupling between approximation errors and stability constants from level to level guarantees convergence of our scheme, see [16]. This condition is satisfied if $\|q_J^\dagger - q^\dagger\|$ decreases at least exponentially with J. Given this compression rate, we can then choose an appropriate set of frequencies which control the approximation errors at the different levels. A straightforward choice is facilitated by a gradual increase in J. For the analysis, precise conditions and expressions for the convergence radii, see [8].

In our implementation we use the nonlinear conjugate gradient method including projections. The step size is computed using a backtracking line search. (By successively increasing the step size one finds the smallest one that reduces the functional, which is then used to update the model.)

3. Adaptive domain partitioning and model compression. In our applications, we will employ domain partitionings based on boxes and piecewise constant functions derived from the Haar basis. The boxes naturally fit a grid. Haar wavelets have been used for compressed model representations, for example, in tomography [27, 26]. In subsection 3.1 we briefly review the basis of Haar wavelets and the notion of scale refinement. The increase of the number of domains by refining the scale one level is large and hence an obstruction to a proper use of our multi-level approach. To arrive at a more gradual increase of the number of domains, still using Haar wavelets, we introduce in subsection 3.2 two adaptive strategies based on local refinement. We give numerical examples for a test model. We will apply the adaptive strategies to the gradient, periodically, in the iterative scheme; that is, we use the current gradient to determine, periodically, the next domain partitioning in our multi-level approach; see subsection 3.3.

3.1. Haar wavelets. The Haar basis [21] is defined using a scaling function ϕ and a mother wavelet ψ. In the one-dimensional case,

$$\psi(x) = \begin{cases} 1 & \text{if } 0 \le x < \frac{1}{2} \\ -1 & \text{if } \frac{1}{2} \le x < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

and the basis is formed by functions $\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)$, $j, k \in \mathbb{Z}$. This basis is orthogonal in L^2. The scaling function is simply the characteristic function on the interval $[0; 1]$, while

$$\phi_{j,k}(x) = \begin{cases} |I_{j,k}|^{-1/2} & \text{if } x \in I_{j,k} \\ 0 & \text{elsewhere,} \end{cases}$$

where $I_{j,k} = [2^j(k-1), 2^j k]$. The Haar coefficients of a function f are then given by

$$d_k^j = \langle f, \psi_{j,k} \rangle,$$

where $\langle ., . \rangle$ denotes the inner product in L^2, and the 'averages' of f on $I_{j,k}$ by

$$s_k^j = \langle f, \phi_{j,k} \rangle.$$

These satisfy the relationships,

$$d_k^{j+1} = \frac{1}{\sqrt{2}}(s_{2k-1}^j - s_{2k}^j)$$

and

$$s_k^{j+1} = \frac{1}{\sqrt{2}}(s_{2k-1}^j + s_{2k}^j),$$

which are used in the design of a fast transform [10]; [29]. We follow an extension to higher dimensions based on taking tensor products.

In this section, we use a model example on a regular grid which we refer to as louro3d. The model contains irregularly shaped salt bodies, is of size $124 \times 79 \times 31$, and represents a wavespeed. We will illustrate multi-scale approximations of the model by injecting it into the original grid, which corresponds with the 0 level compression. In Figure 1 we illustrate the decomposition of louro3d into Haar wavelets over 4 levels of compression. The number of coefficients for each level is given in Table 3.1.

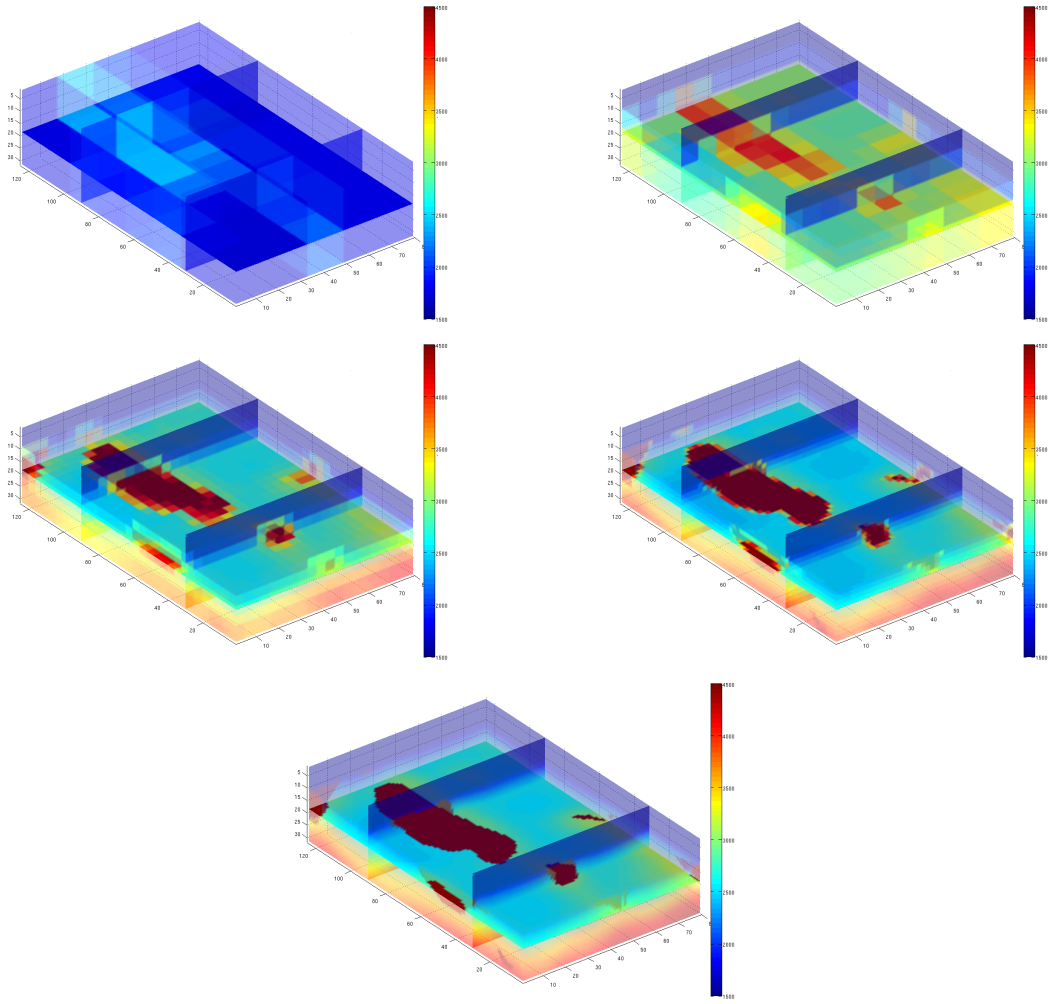

FIG. 1. *Decomposition into Haar wavelets, and implied domain partitionings, from level* 4 *(upper left figure) to level* 0 *(bottom figure). The level* 0 *corresponds with the original louro3d model.*

level of compression	number of coefficients or domains	percentage
0	303 676	100%
1	36 270	12%
2	4 123	1.4%
3	405	0.1%
4	28	0.009%

TABLE 3.1

Number of coefficients or domains in the domain partitioning depending on the level of compression for the louro3d model shown in Figure 1.

It is immediate from the structure of the Haar basis that refining the scale increases J in the domain partitioning by a factor 8. For the multi-level scheme, which requires a gradual increase of J, we now introduce an adaptive local refinement.

3.2. Local refinement. We discuss two methods, which assume a given coefficient model on a (computational) grid.

Method 1. In this method, we simply refine only where the deviation from the local average of the wavespeed evaluated through the scaling functions appears to be above a certain threshold. In case a refinement is applied, the box (domain) is split into 8 new boxes. Furthermore, we can distinguish the directions and subject them to refinement independently.

We apply this method to the louro3d model. We illustrate it for compression level 2 highlighting the implied domain partitionings. In Figure 2 we show a horizontal slice of the results; the left figure corresponds with Figure 1 (level 2) while the middle figure represents the method discussed here. The number of domains or coefficients has been moderately increased, from 4123 (1.4% of the original) to 7455 (2.5% of the original), still far below the number for Haar level 1 (12% of the original). In Figure 4 we show the logarithmic approximation error in L^2 as a function of the number of domains, J, in the domain partitioning, and illustrate the effect of a more gradual increase (the red dots).

Method 2. The strategy of our second method is based on gluing boxes together forming larger domains. The criterion for gluing follows from comparing adjacent averages at a current level. If the averages are sufficiently close we glue to underlying domains together. The threshold in the criterion essentially controls the increase in the number of domains in the domain partitioning from level to level.

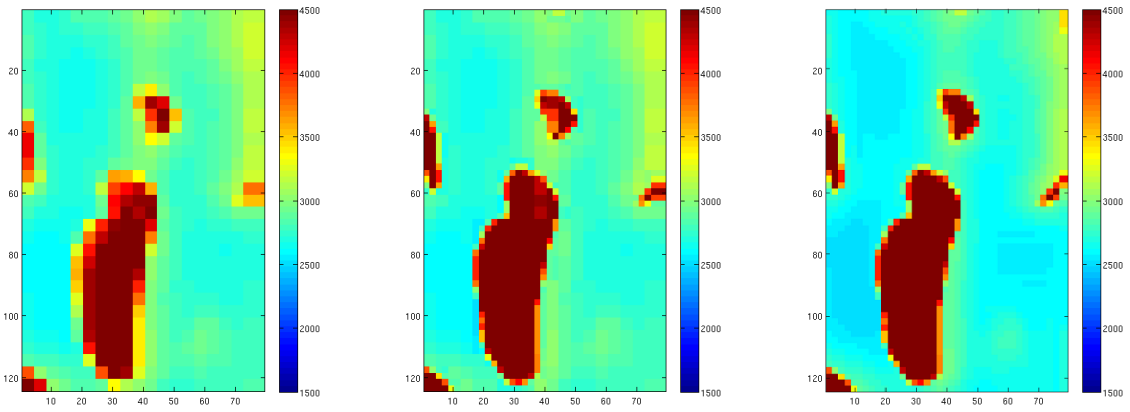

FIG. 2. *A horizontal slice of the compressed louro3d model at level* 2*; the color bar indicates wavespeed in m/s. Left: Haar basis (*4123 *domains); middle: method 1 (7 455 domains); right: method 2 (*9704 *domains).*

We show an example using the the louro3d model in Figure 2 (right). Here, the level 1 Haar compression yields the initial paritioning. We apply the gluing in the horizontal directions only. In Figure 4 we show the logarithmic approximation error in L^2 as a function of the number of domains, J, in the domain partitioning, and illustrate the effect of a more gradual increase (the green dots).

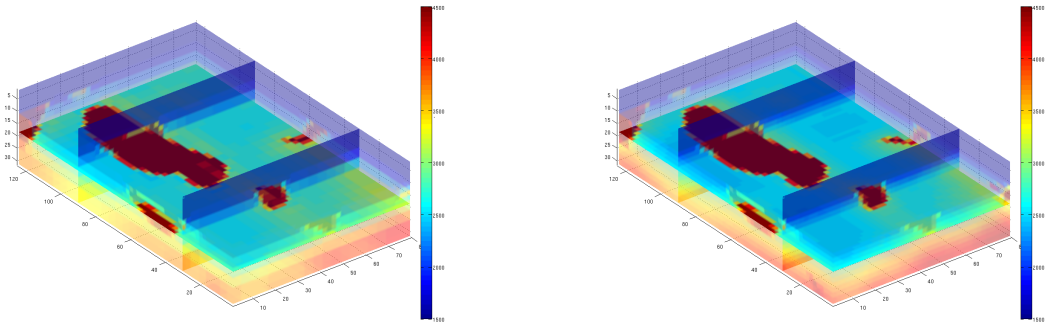

FIG. 3. *Representation of the local refinement methods for compression and domain partitioning. Left: method 1 (*7455 *domains); right: method 2 (*9704 *domains).*

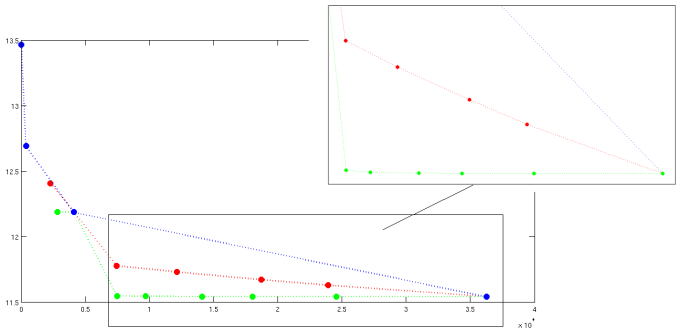

FIG. 4. *The error, in L^2 norm, as a function of the number of domains in the domain partitioning. The blue dots correspond with the different levels following the compression with Haar wavelets. The red dots correspond with method 1, and the green dots correspond with method 2; these methods generate intermediate domain partitionings as compared with the one using the Haar wavelets.*

While the first method keeps the structure of the Haar basis, the second method, of course, provides more flexibility, while both methods can be applied directionally. We illustrate this in Figure 5 for a 2-dimensional grid. Method 1 starts from a coarser scale representation (left subfigure); a local refinement to a finer scale is illustrated in the two top subfigures. Method 2 starts from a finer scale scale representation (right subfigure); the gluing to domains corresponding with a coarser scale scale is illustrated in the two bottom subfigures.

3.3. Gradient adaptive compression. In our multi-level iterative scheme, we refine the domain partitioning at each new frequency in the (discrete) frequency progression. In the refinement, we use the current gradient as a point of departure and project it onto two neighboring scales; the finest of these two scales corresponds with the next level. Method 1 uses the coarsest scale to construct the new domain partitioning while Method 2 uses the finest scale for this purpose.

4. Numerical experiments. We carry out numerical experiments illustrating the performance of our multi-level, multi-frequency iterative scheme using modelled data. We consider the data to be noise-free, though, in view of the mentioned stability estimates, additive noise through the approximation errors will not influence the results in a significant way. As mentioned before, the underlying theory including a condition for convergence can be found in [8]. We invoke the gradient

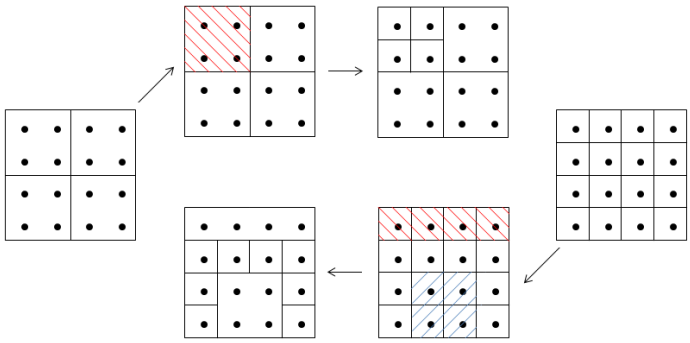

FIG. 5. *Example of the local refinement methods in 2D. Method 1 is illustrated in the upper part, while method 2 is illustrated in the bottom part. Left: the Haar level* 1 *represention; right: the Haar level* 0 *representation. Method 1 takes the coarser level as a starting point, selects which part should be refined and then refines as if it were Haar with one level down on the selected part (upper right subfigure). Method 2 starts from the finer level, selects which domains should be glued together (two in this example) and merges domains accordingly (lower left subfigure).*

adaptive refinement of the domain partitioning. We start with the lowest available frequency and a model representation on the coarsest scale using the computational grid as a reference for the finest scale. After a prescribed number of iterations, the frequency is increased; we carry out a domain re-partitioning as outlined in the previous section, and then restart the iterations. This process progresses through a set of frequencies determined by the condition on the approximation errors which involves the upper bound of the stability constants, and the stability constants themselves. These implicitly depend on the 'true' solution which we do not know in practice. However, through some trial iterations, a reasonable choice of frequencies can be obtained. In subsections 4.1 and 4.2 the wavespeed is real-valued and the stability constants and approximation errors control the convergence of our multi-level, multi-frequency iterative scheme; in subsection 4.3 we consider complex wavespeeds and attenuation where the (frequency dependent) lower bound of the Fréchet derivative also strongly influences the convergence in particular for the recovery of the quality factor.

4.1. Salt bodies: louro3d. The louro3d model was illustrated in Figure 1 (bottom); the step size is 20m and the wavespeed varies between 1517 m/s and 4527 m/s. We generate data between 1 Hz and 15 Hz with our *Hsolver*.

We employ, here, the non-adaptive approach using the Haar basis for the domain partitioning and refinement. The initial model is shown Figures 6 (right) and 7 (right), and does not contain any indication of the salt bodies.

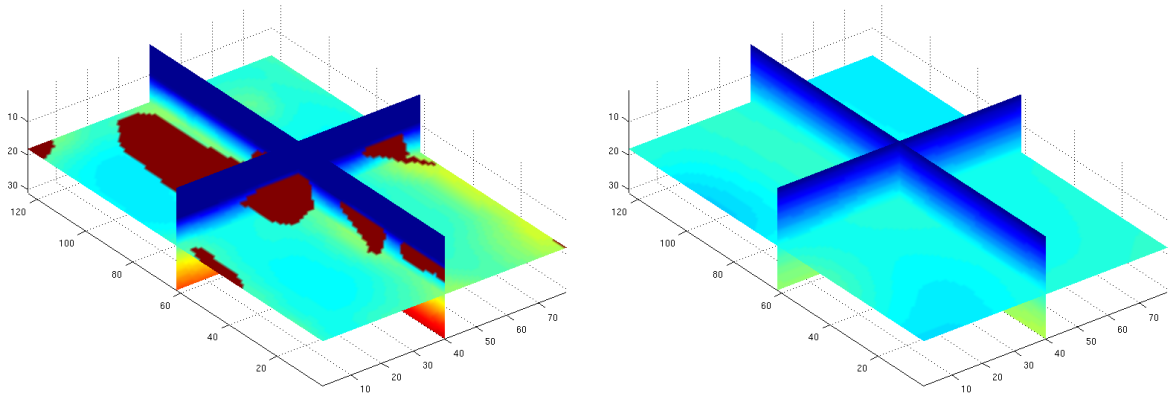

FIG. 6. *Left: True wavespeed model; right: initial wavespeed model.*

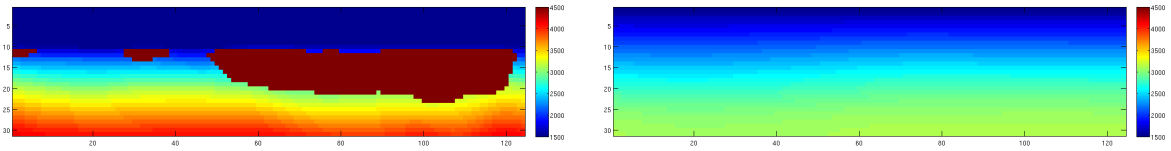

FIG. 7. *Vertical slices of the true wavespeed model (left) and initial wavespeed model (right).*

We begin with illustrating the role of compression in the convergence of our scheme. This role becomes pronounced at relatively higher frequencies, as expected from the explicit expressions for the convergence rates [16]. In Figure 8, the mismatch functionals are plotted as a function of the iteration, at 2 Hz (left) and 3 Hz (right), with compression (level 2) and without compression (level 0).

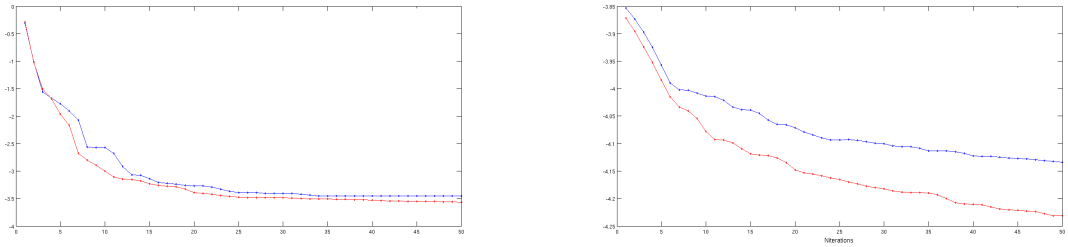

FIG. 8. *The mismatch functionals (logarithmic scale) as a function of iteration at* 2 *Hz (left) and* 3 *Hz (right), with compression (level* 2*, the red curve) and without (level* 0*, the blue curve).*

We proceed with applying our multi-level iterative scheme progressively for frequencies in the set $\{2, 3, 4, 5, 6, 7, 8, 9, 10\}$ (in Hz). Here, we simply use the Haar basis for compression and domain partitioning. At 2 Hz we start with level 2; at 4 Hz we refine the domain partitioning to level 1 and at 8 Hz to level 0. Figure 9 shows a horizontal slice of the (partial) recovery at level 2 and Figure 10 shows the same horizontal slice of the (partial) recovery at level 1. In Figures 11 and 12 we show the (partial) recovery obtained at level 0.

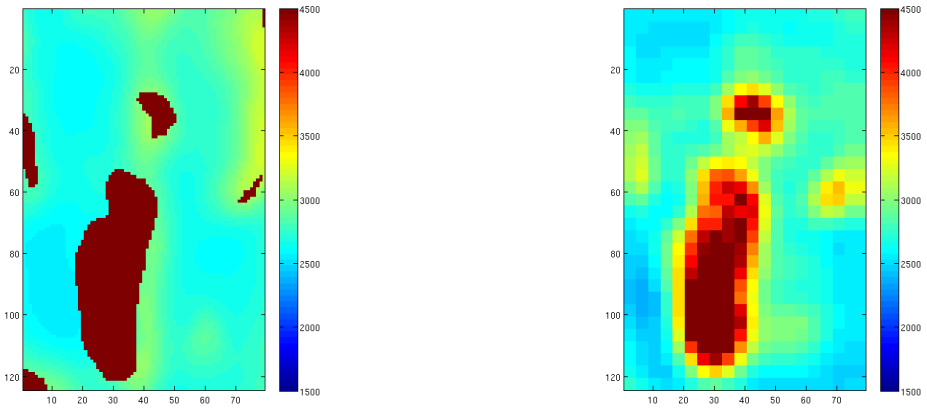

FIG. 9. *A horizontal slice of the true wavespeed model (left) and the (partial) recovery at level* 2 *(right).*

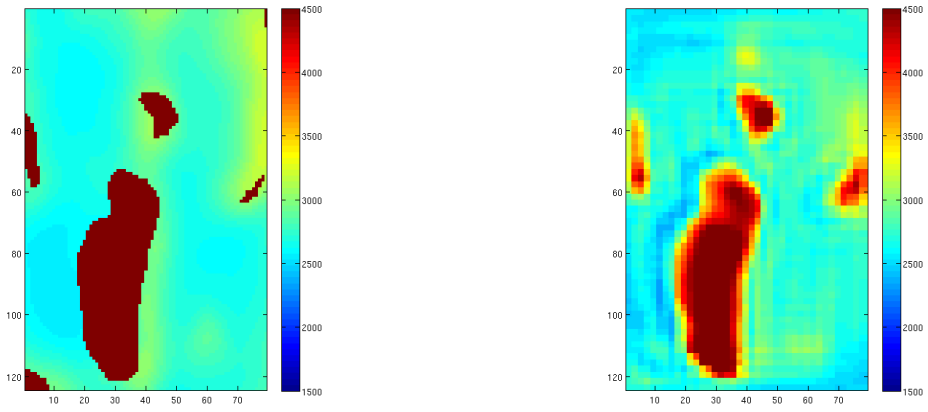

FIG. 10. *A horizontal slice of the true wavespeed model (left) and the (partial) recovery at level* 1 *(right).*

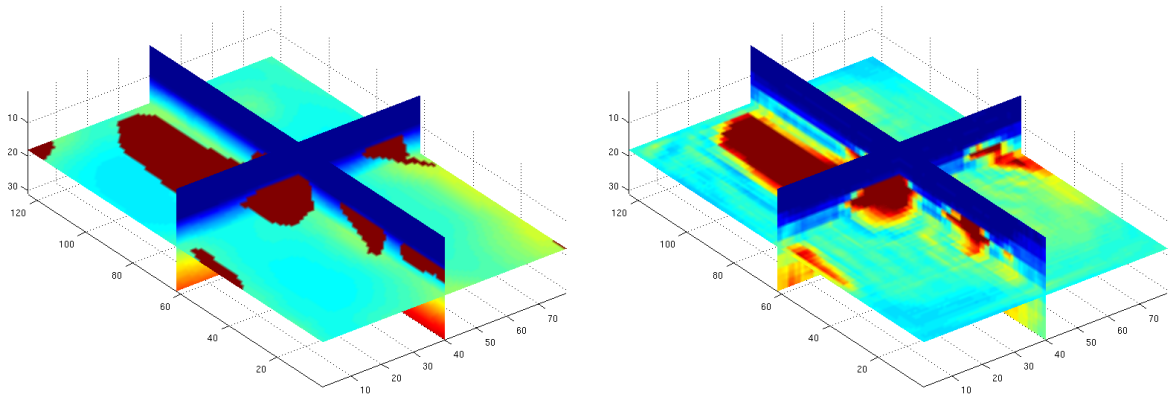

FIG. 11. *The true wavespeed model (left) and the (partial) recovery at level* 0 *(right).*

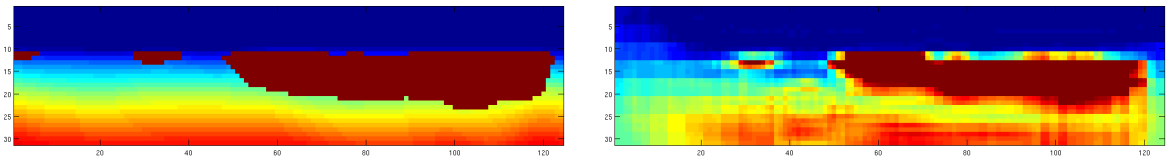

FIG. 12. *A vertical slice of the true wavespeed model (left) and the (partial) recovery at level* 0 *(right).*

The model consists of multiple disconnected salt bodies of different sizes. Even after the first iterations at the coarsest level, in Figure 9, the largest bodies appear with a wavespeed close to the true wavespeed at their centers. Following the progression in our multi-level, multi-frequency scheme we begin to resolve the smaller bodies, as well as the boundaries of the bodies and the wavespeeds in their interiors; see Figures 10, 11 and 12.

Figure 13 presents the residuals corresponding with a centrally located shot (real part) during the 5 Hz reconstructive iterations, Figure 14 during the 10 Hz iterations. The left figure shows the starting residual, in the middle after 5 iterations at the considered frequency, on right after 20 iterations. The residuals are decreasing while the salt bodies are partly being resolved. The residual does not decrease near the boundary of the computational domain; indeed, we cannot resolve the wavespeed near this boundary.

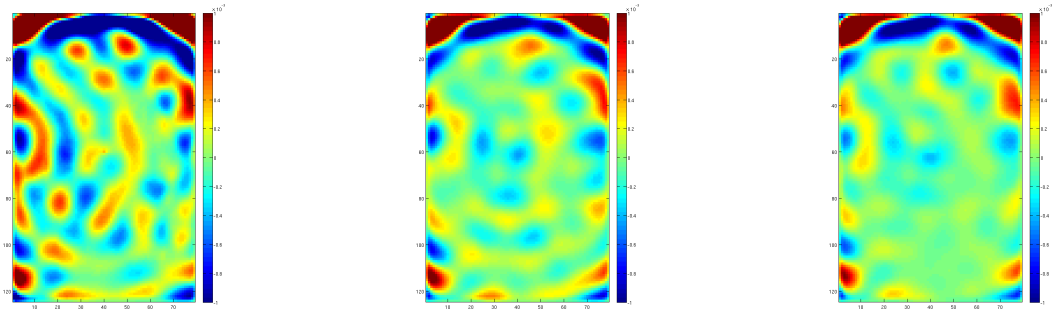

FIG. 13. *The residual for a central shot at different stages of the 5Hz frequency iterative reconstruction of the louro3d model. The color bar indicates the difference, all Figures have the same scale. Left: at the first iteration; middle: after the fifth iteration; right: after the twentieth iteration.*

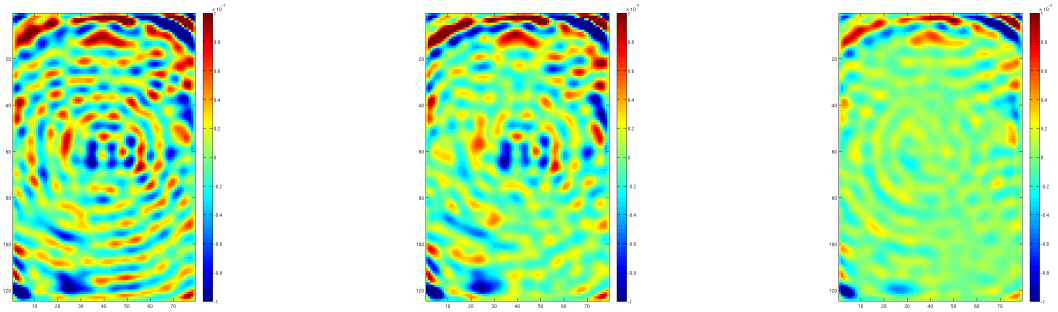

FIG. 14. *The residual for a central shot at different stages of the 10Hz frequency iterative reconstruction of the louro3d model. The color bar indicates the difference, all Figures have the same scale. Left: at the first iteration; middle: after the fifth iteration; right: after the twentieth iteration.*

4.2. Statoil model. We use a second test model (courtesy Statoil) representing a different geological environment with high wavespeed contrasts to illustrate our adaptive compression methods. The model is of size $256 \times 146 \times 123$ and the stepsize of the grid is 10 m; see Figures 15 and 16. The set of frequencies used in our scheme is $\{3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 18, 20, 25\}$ (in Hz). We carry out a maximum of 20 iterations per frequency. The initial model is built from a one-dimensional wavespeed profile in depth.

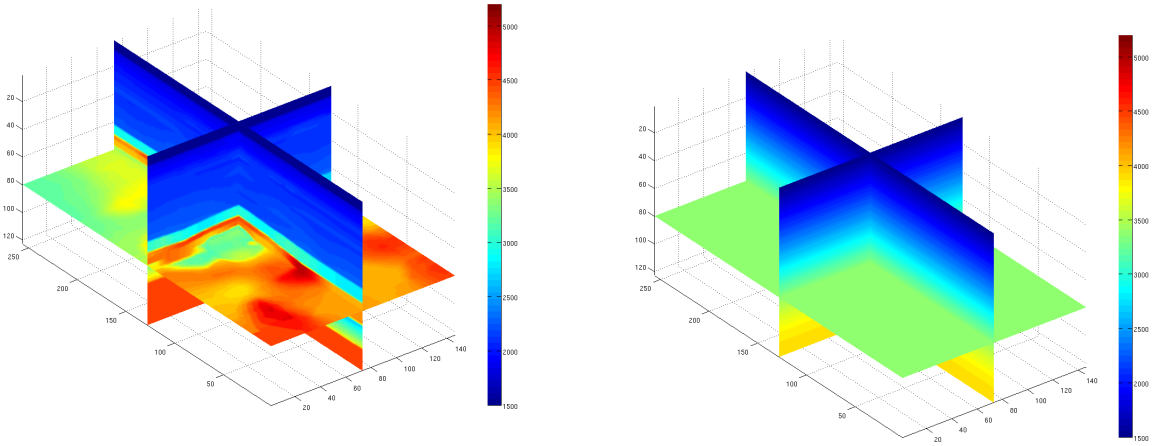

FIG. 15. *True wavespeed model (left) and initial model (right).*

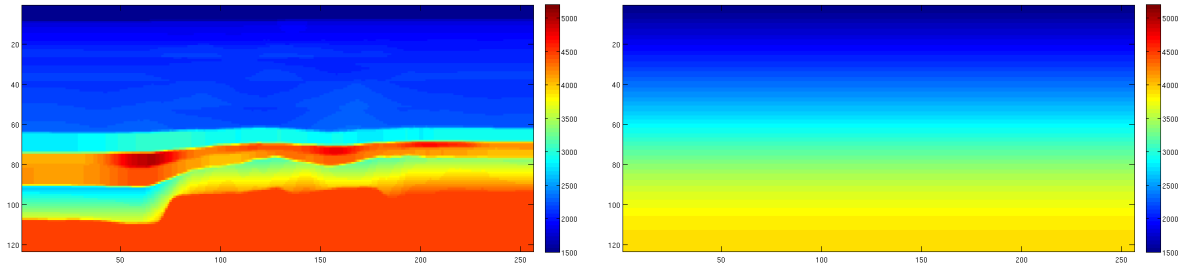

FIG. 16. *A vertical slice of the true wavespeed model (left) and the initial model (right).*

Nonlinear compressed reconstruction using the Haar basis. The Haar level 3 is the initial compression. It is switched to level 2 at 4 Hz, to level 1 at 6 Hz and to level 0 at 10 Hz. Figures 17 and 18 show the finite reconstruction at 25 Hz. Figure 19 gives a three-dimensional view and also shows wavespeed profiles in depth. We note the wavespeed in the high wavespeed layer has not completely been resolved. We expect that an iterative scheme based on high-frequency data will further resolve such a detail.

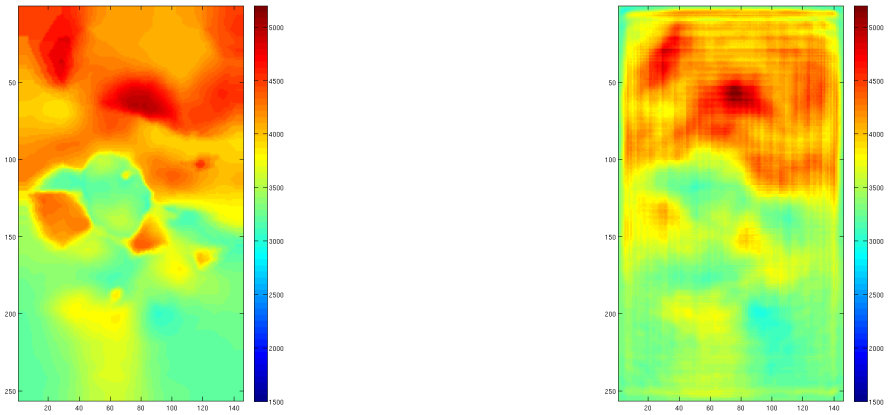

FIG. 17. *Horizontal slice of the true wavespeed model (left) and the reconstruction at* 25 *Hz (right).*

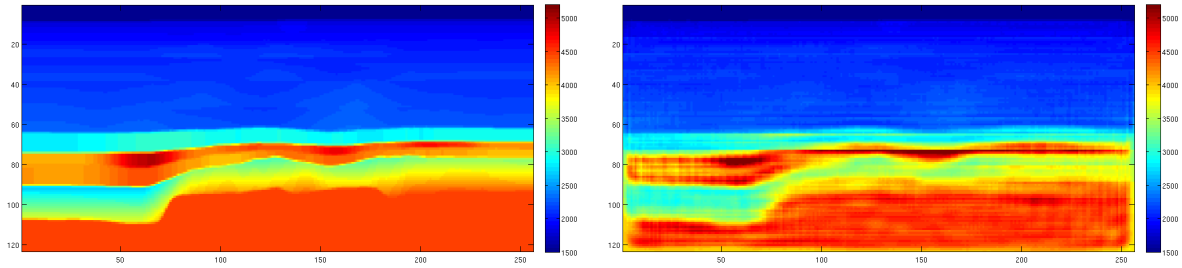

FIG. 18. *Vertical slice of the true wavespeed model (left) and the reconstruction at* 25 *Hz (right).*

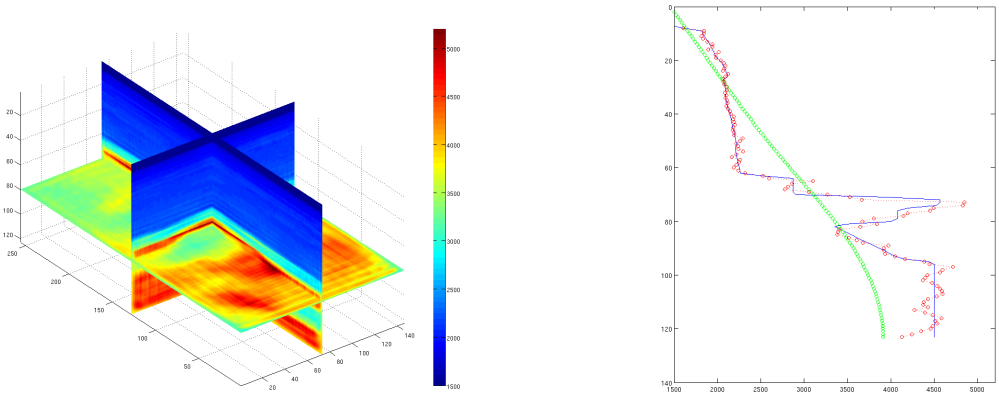

FIG. 19. *The partial reconstruction at* 25 *Hz (left) and wavespeed profiles in depth at* $x = 1.3km$; $y = 400m$*: the true wavespeed in blue, the initial model in green, and the reconstruction using Haar basis in red circles.*

Nonlinear compressed reconstruction using method 1. Again, the Haar level 3 is the initial compression. We apply method 1 for the domain re-partitioning. Table 4.1 shows the implied gradual increase in the number of domains in the multi-level approach while increasing the frequency. Figures 20, 22 (left) and 23 (left) show the finite reconstruction obtained with method 1.

frequency	level of compression	number of domains	percentage
3Hz	Haar level 3	8 640	0.19%
4Hz	Haar level 2	69 120	1.5 %
5Hz	method 1	348 847	7.6 %
6Hz	Haar level 1	569 984	12.4 %
7Hz	method 1	1 242 950	27 %
8Hz	method 1	2 090 685	45.5 %
9Hz	method 1	3 052 765	66.4 %
10Hz-25Hz	level 0	4 597 248	100 %

TABLE 4.1

Number of domains versus frequency in the multi-level scheme for the model depicted in Figure 15 of original size $256 \times 146 \times 123$.

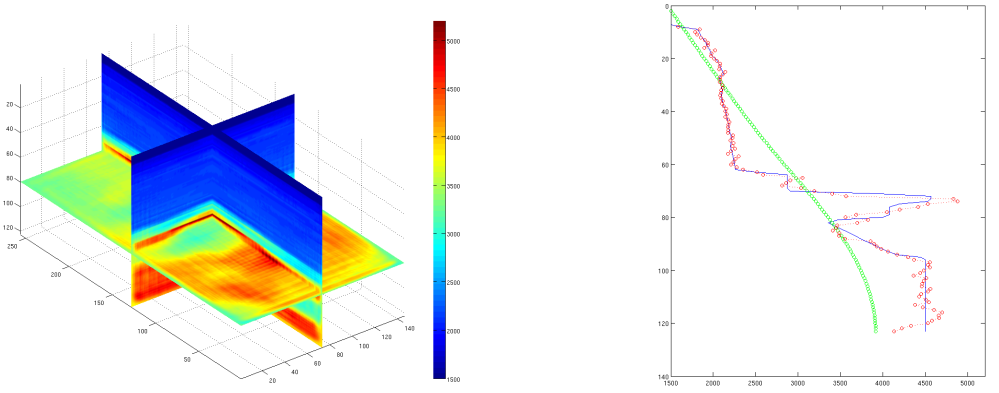

FIG. 20. *The partial reconstruction at* 25 *Hz using method 1 (left) and wavespeed profiles in depth at* $x = 1.3km$; $y = 400m$*: the true wavespeed in blue, the initial model in green, and the reconstruction in red circles using method 1.*

Nonlinear compressed reconstruction using method 2. Similarly we now apply method 2 for the domain re-partitioning. The number of domains in the multi-level approach is represented by the Table 4.2. The finite reconstruction is shown in Figures 21, 22 (right) and 23 (right).

frequency	level of compression	number of domains	percentage
3Hz	Haar level 3	8 640	0.19%
4Hz	Haar level 2	69 120	1.5 %
5Hz	method 2	469 224	10.2 %
6Hz	Haar level 1	569 984	12.4 %
7Hz	method 2	1 232 733	26.8 %
8Hz	method 2	2 176 304	47 %
9Hz	method 2	3 080 955	67 %
10Hz-25Hz	level 0	4 597 248	100 %

TABLE 4.2

Number of domains versus frequency in the multi-level scheme for the model depicted in Figure 15 of original size $256 \times 146 \times 123$.

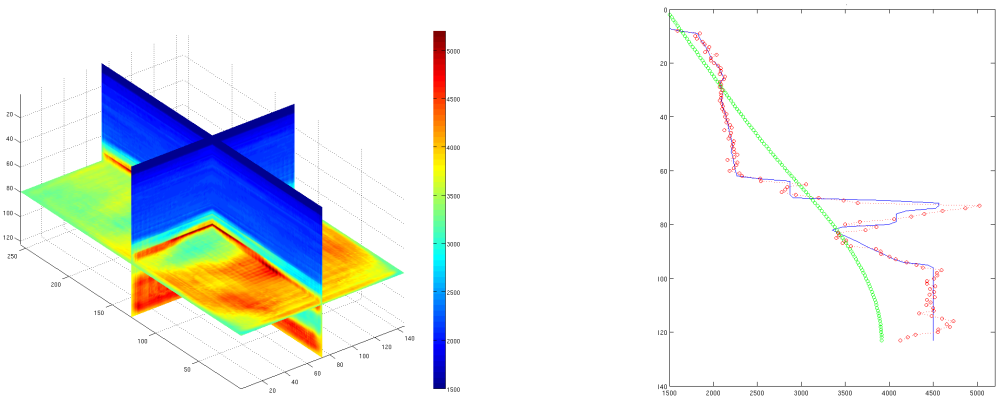

FIG. 21. *The partial reconstruction at* 25 *Hz using method 1 (left) and wavespeed profiles in depth at* $x = 1.3km$*;* $y = 400m$*: the true wavespeed in blue, the initial model in green, and the reconstruction in red circles using method 2.*

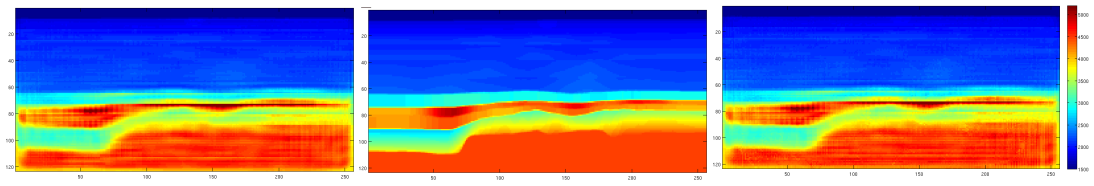

FIG. 22. *Vertical slice of the true wavespeed model (middle) and the reconstruction at* 25*Hz using method 1 (left) and method 2 (right).*

The differences between the results obtained with the Haar basis and method 1 or 2 for domain re-partitioning seem to be small. However, differences show up in the mismatch functional; see Figure 24. The differences become less in the progression to higher frequencies for a given domain partitioning; we illustrate this for the final iterations at $20 - 25$ Hz. Method 1 gives the best result.

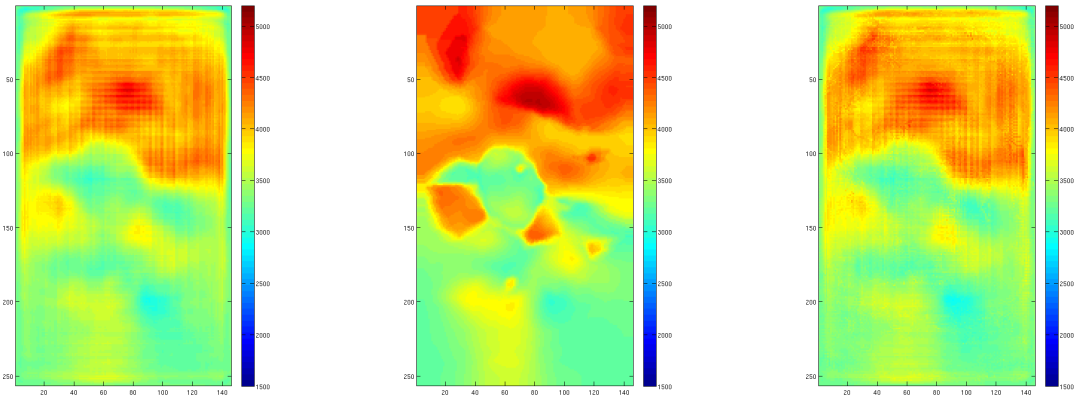

FIG. 23. *Horizontal slice of the true wavespeed model (middle) and the reconstruction at 25Hz using method 1 (left) and method 2 (right).*

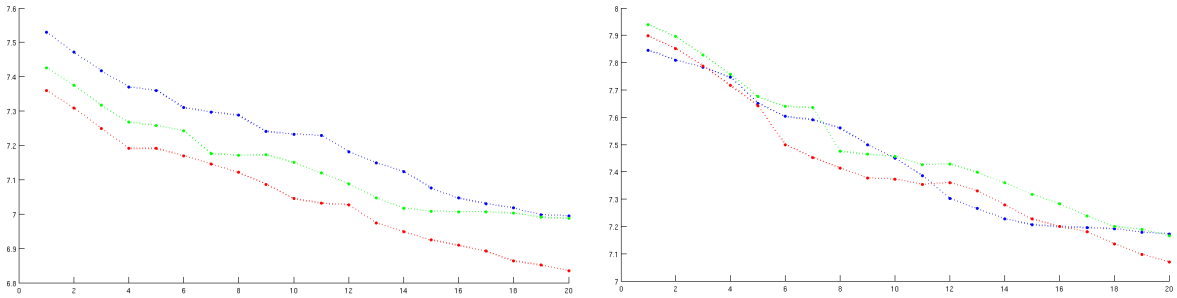

FIG. 24. *Convergence rates: The mismatch functional (logarithmic scale) as a function of iterate. Left: at 20 Hz; right: at 25 Hz. The blue curve corresponds with the refinement using the Haar basis while the red curve corresponds with the refinement using method 1. The green curve corresponds with the refinement using method 2. The domain partitioning is kept fixed in both cases.*

4.3. Attenuation. Here, we consider complex wavespeeds and attenuation. We study the partial recovery of the phase velocity and quality factor (cf. (1.2)), or the real and imaginary parts of the wavespeed. For our numerical experiments, we use a model containing an artificial geobody with the shape of a cross; see Figures 25 and 26. Outside the body, the phase velocity increases linearly in depth; also, the phase velocity is significantly higher inside the body than outside it. The quality factor is chosen to increase with increasing phase velocity. The model is of size $124 \times 79 \times 31$ with a step size of 20 m.

There is no fundamental obstruction to recover a complex wavespeed, and, indeed, our stability estimates hold for this case. However, the key challenge, here, is that the Fréchet derivative with respect to the quality factor becomes significant in magnitude only for relatively high frequencies. We encounter the effect of the (lower) bound of the Fréchet derivative on the radius of convergence and convergence rates. In our model, a sufficiently high frequency is reached at 15 Hz. At this frequency, partial recovery of the quality factor becomes possible; see Figure 28. In Figure 27 we show the partial recovery of the phase velocity. An appropriate set of frequencies follows to be $\{3, 4, 5, 6, 7, 8, 10, 12, 15\}$ Hz. We begin with the Haar basis at level 2; at 4 Hz we refine the domain partitioning to level 1 and then to level 0 at 8 Hz.

5. Discussion. We have implemented and applied the analysis and construction of a multi-level, multi-frequency iterative scheme including conditions for convergence developed in [16]. The

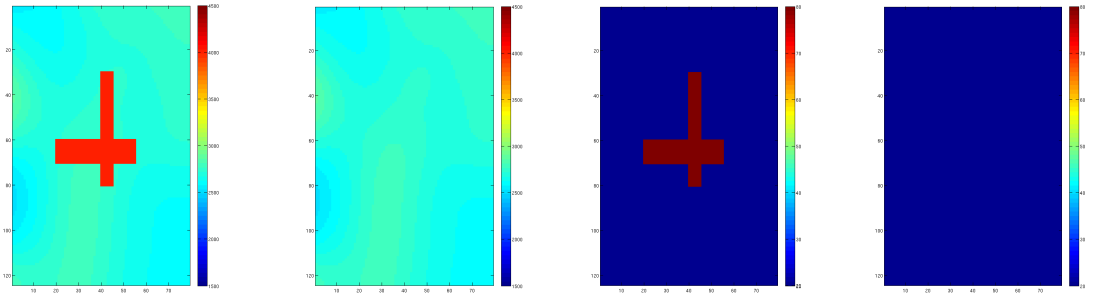

FIG. 25. *Horizontal slice of the true phase velocity (left) and quality factor (middle right) and the initial model (middle left and right).*

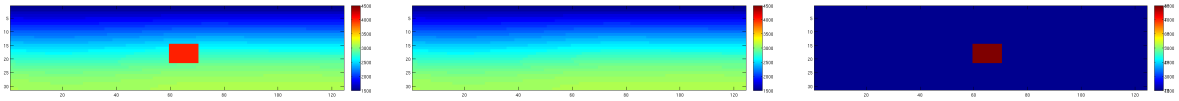

FIG. 26. *Vertical slice of the true phase velocity (left) and quality factor (right) and the initial model (middle).*

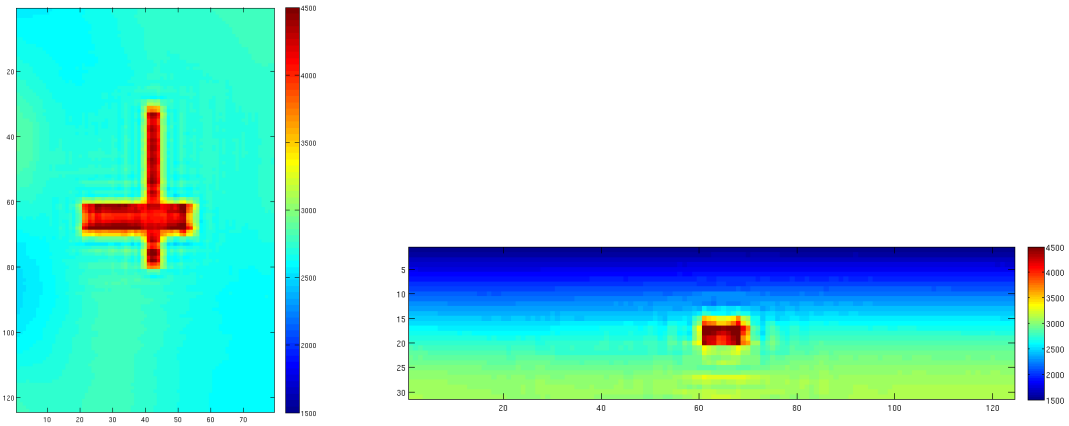

FIG. 27. *Partial reconstruction of the phase velocity at* 15 *Hz.*

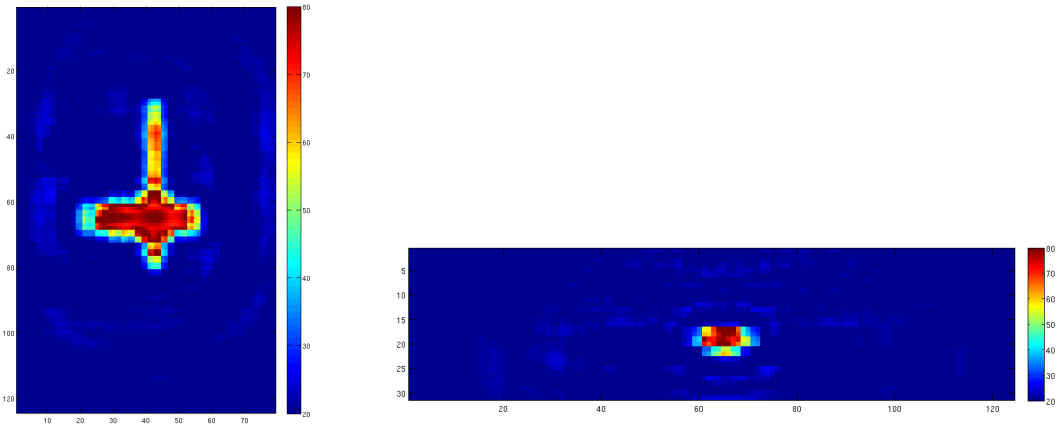

FIG. 28. *Partial reconstruction of the quality factor at* 15 *Hz.*

analysis provided an understanding and a precise characterization of, for example, frequency progression in full waveform inversion; it also showed that this progression should be coupled to hierarchical model compression in a particular way to guarantee convergence. The analysis originated from a Lipschitz stability estimate for the inverse problem [7] using the single layer potential operator as the data. This estimate was obtained for piecewise constant (blocky) models incorporating interfaces, pinchouts, etc., corresponding with a given domain partitioning. We introduced two adaptive methods of (directional) compression and compared their performance in computational experiments using two geologically different models. The compression overcomes in part the need for very-low frequency data while allowing initial models not containing prior information about the structure of the 'true' model. Our analysis and application are based on the behaviors of the stability constants, and approximation errors at low frequencies. The stability analysis, which will be local, and, hence, convergence, at high frequencies requires the use of different techniques, which is a current subject of research.

REFERENCES

[1] V. Akcelik, G. Biros, and O. Ghattas, *Parallel multiscale gauss-newton-krylov methods for inverse wave propagation*, in Supercomputing, ACM/IEEE 2002 Conference, 2002, pp. 41–41.

[2] A. Askan, V. Akcelik, J. Bielak, and O. Ghattas, *Full waveform inversion for seismic velocity and anelastic losses in heterogeneous structures*, Bulletin of the seismological of America, 97 (2007), pp. 1990–2008.

[3] A. Bamberger, G. Chavent, and P. Lailly, *Une application de la théorie du contrôle à un problème inverse de sismique*, Annales de Géophysique, 33 (1977), pp. 183–200.

[4] ———, *About the stability of the inverse problem in the 1-d wave equation*, Journal of Applied Mathematics and Optimisation, 5 (1979), pp. 1–47.

[5] G. Bao and P. Li, *Inverse medium scattering for the helmholtz equation at fixed frequency*, Inverse Problems, 21 (2005), p. 1621.

[6] ———, *Numerical solution of an inverse medium scattering problem for maxwell's equations at fixed frequency*, J. Comput. Phys., 228 (2009), pp. 4638–4648.

[7] E. Beretta, M. V. de Hoop, and L. Qiu, *Lipschitz stability of an inverse boundary value problem for a Schr\"{o}dinger type equation*, ArXiv e-prints, (2012).

[8] E. Beretta, M. V. de Hoop, L. Qiu, and O. Scherzer, *Inverse boundary value problem for the helmholtz equation with multi-frequency data*. preprinted, 2013.

[9] A. J. Berkhout, *Changing the mindset in seismic data acquisition*, The Leading Edge, 27 (2008), pp. 924–938.

[10] G. Beylkin, R. Coifman, and V. Rokhlin, *Fast wavelet transforms and numerical algorithms i*, Communications on Pure and Applied Mathematics, 44 (1991), pp. 141–183.

[11] E. Bozdag, J. Trampert, and J. Tromp, *Misfit functions for full waveform inversions based on instantaneous phase and envelope measurements*, Geophysical Journal International, 185 (2011), pp. 845–870.

[12] R. Brossier, *Imagerie sismique à deux dimensions des milieux viso-élastiques par inversion des formes d'ondes : développements méthodologiques et applications*, PhD thesis, Universit'e de Nice-Sophia Antipolis, 2009.

[13] C. Bunks, F. M. Saleck, S. Zaleski, and G. Chavent, *Multiscale seismic waveform inversion*, Geophysics, 60 (1995), pp. 1457–1473.

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

[15] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust Region Methods*, Society for Industrial and Applied Mathematics, Jan. 2000.

[16] M. V. de Hoop, L. Qiu, and O. Scherzer, *A convergence analysis of a multi-level projected steepest descent iteration for nonlinear inverse problems in Banach spaces subject to stability constraints*, ArXiv e-prints, (2012).

[17] S. C. Eisenstat and H. F. Walker, *Globally convergent inexact newton methods*, SIAM Journal on Optimization, 4 (1994), pp. 393–422.

[18] W. Futterman, *Dispersive body waves*, Journal of Geophysical Research, 67 (1962), pp. 5279–5291.

[19] O. Gauthier, J.Vieurx, and A. Tarantola, *Two-dimensional nonlinear inversion of seismic waveforms: Numerical results*, Geophysics, 51 (1986), pp. 1387–1403.

[20] W. Ha, S. Pyun, J. Yoo, and C. Shin, *Acoustic full waveform inversion of synthetic land and marine data in the laplace domain*, Geophysical Prospecting, 58 (2010), pp. 1033–1047.

[21] A. Haar, *Zur theorie der orthogonalen funktionensysteme*, Mathematische Annalen, 71 (1911), pp. 38–53.

[22] B. L. N. Kennett, M. S. Sambridge, and P. R. Williamson, *Subspace methods for large inverse problems with multiple parameter classes*, Geophysical Journal International, 94 (1988), pp. 237–247.

[23] H. Kolsky, *The propagation of stress pulses in viscoelastic solids*, Philosophical magazine, 1 (1956), pp. 693–710.

[24] P. Lailly, *The seismic inverse problem as a sequence of before stack migrations*, in Conference on Inverse

Scattering: Theory and Application, J. B. Bednar, ed., Society for Industrial and Applied Mathematics, 1983, pp. 206–220.
[25] Y. Lin, A. Abubakar, and T. M. Habashy, *Seismic full-waveform inversion using truncated wavelet representations*, (2012), pp. 1–6.
[26] I. Loris, H. Douma, G. Nolet, I. Daubechies, and C. Regone, *Nonlinear regularization techniques for seismic tomography*, Journal of Computational Physics, 229 (2010), pp. 890–905.
[27] I. Loris, G. Nolet, I. Daubechies, and F.A. Dahlen, *Tomographic inversion using l1-norm regularization of wavelet coefficients*, Geophysical Journal International, 170 (2007), pp. 359–370.
[28] M. Malinowski, S. Operto, and A. Ribodetti, *High-resolution seismic attenuation imaging from wid-aperture onshore data by visco-acoustc frequency-domain full-waveform inversion*, Geophysical Journal International, 186 (2011), pp. 1179–1204.
[29] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 2008.
[30] L. Métivier, R. Brossier, J. Virieux, and S. Operto, *Full waveform inversion and the truncated newton method*, SIAM Journal on Scientific Computing, to appear.
[31] G. S. Pan, R. A. Phinney, and R. I. Odom, *Full-waveform inversion of plane-wave seismograms in stratified acoustic media: Theory and feasibility*, Geophysics, 53 (1988), pp. 21–31.
[32] S. Park, L. Qiu, M. V. de Hoop, and C. Shin, *On time-harmonic seismic data and blending in full waveform inversion*, in Geo-Mathematical Imaging Group - Project Review, 2012.
[33] R.-E. Plessix, *A review of the adjoint-state method for computing the gradient of a functional with geophysical applications*, Geophysical Journal International, 167 (2006), pp. 495–503.
[34] G. Pratt and N. Goulty, *Combining wave-equation imaging with traveltime tomography to form high-resolution images from crosshole data*, Geophysics, 56 (1991), pp. 208–224.
[35] R. G. Pratt, C. Shin, and G. J. Hicks, *Gauss-newton and full newton methods in frequency-space seismic waveform inversion*, Geophysical Journal International, 133 (1998), pp. 341–362.
[36] R. G. Pratt, Z.-M. Song, P. Williamson, and M. Warner, *Two-dimensional velocity models from wide-angle seismic data by wavefield inversion*, Geophysical Journal International, 124 (1996), pp. 323–340.
[37] R. G. Pratt and M. H. Worthington, *Inverse theory applied to multi-source cross-hole tomography. part 1: Acoustic wave-equation method*, Geophysical Prospecting, 38 (1990), pp. 287–310.
[38] C. Shin and Y. H. Cha, *Waveform inversion in the laplace-fourier domain*, Geophysical Journal International, 177 (2009), pp. 1067–1079.
[39] C. Shin, S. Jang, and D.-J. Min, *Improved amplitude preservation for prestack depth migration by inverse scattering theory*, Geophysical Prospecting, 49 (2001), pp. 592–606.
[40] C. Shin and D.-J. Min, *Waveform inversion using a logarithmic wavefield*, Geophysics, 71 (2006), pp. 31–42.
[41] C. Shin, S. Pyun, and J. B. Bednar, *Comparison of waveform inversion, part 1: conventional wavefield vs logarithmic wavefiled*, Geophysical Prospecting, 55 (2007), pp. 449–464.
[42] L. Sirgue and R. G. Pratt, *Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies*, Geophysics, 69 (2004), pp. 231–248.
[43] A. Tarantola, *Inversion of seismic reflection data in the acoustic approximation*, Geophysics, 49 (1984), pp. 1259–1266.
[44] ———, *Inverse Problem Theory: methods for data fitting and and model parameter estimation*, Elsevier, 1987.
[45] B. Ursin and T. Toverud, *Comparison of seismic dispersion and attenuation models*, Studia Geophysica et Geodaetica, 46 (2002), pp. 293–320.
[46] S. Wang, M. V. de Hoop, and J. Xia, *Acoustic inverse scattering via helmholtz operator factorization and optimization*, Journal of Computational Physics, 229 (2010), pp. 8445–8462.
[47] ———, *On 3d modeling of seismic wave propagation via a structured parallel multifrontal direct helmholtz solver*, Geophysical Prospecting, 59 (2011), pp. 857–873.
[48] S. Wang, M. V. de Hoop, J. Xia, and X. S. Li, *Massively parallel structured multifrontal solver for time-harmonic elastic waves in 3-d anisotropic media*, Geophysical Journal International, 191 (2012), pp. 346–366.