Geophys. J. Int. (2003) **155,** 953–980

A multi-level direct-iterative solver for seismic wave propagation modelling: space and wavelet approaches

Bernhard Hustedt,* Stéphane Operto and Jean Virieux

Géosciences Azur, OOV, CNRS UMR 6526*, La Darse BP* 48, 06235 *Villefranche-sur-Mer, France. E-mail: operto@obs-vlfr.fr*

Accepted 2003 August 4. Received 2003 July 23; in original form 2002 December 13

SUMMARY

We present a new numerical modelling approach for frequency-domain finite-difference (FDFD) wave simulations. The new approach is developed as an extension to standard FDFD modelling schemes, when wave propagation simulations are performed in large-scale 2-D or 3-D models with complex heterogeneous rheology. Partial differential equations are presented in matrix-type form. Wavefield solutions are computed on different coarse- and fine-discretized numerical grids by a combination of a direct solver with an iterative solver. Two different connection strategies are designed. Both compute a coarse-grid wavefield solution using a direct matrix solver. The obtained solution is projected on a fine-discretized grid, which is used as an initial solution for an iterative solver to compute the desired fine-grid solution. The wavefield projection that combines coarse and fine grids, is either based on a space interpolation scheme, called the direct iterative space solver (DISS) or on a multi-scale wavelet expansion, called the direct iterative wavelet solver (DIWS). The DISS scheme mimics a nested iteration scheme of a full multi-grid method, since numerical grids are prolonged by a simple bilinear interpolation scheme. The simple grid combination leads to wavefield solutions that are affected by spatial phase-shift artefacts (aliasing), which may be suppressed by a large number of iteration steps or a standard V- and W-cycles sequence between grids. The actual DIWS matrix construction implementation is computationally more expensive, though the wavelet iteration scheme guarantees fast and stable iterative convergence. Coarse-grid wavefield solutions are combined with fine-grid solutions through the multi-resolution scaling property of a standard orthogonal wavelet expansion. Since the wavelet transformation accounts for grid interactions, phase-shift artefacts are greatly reduced and significantly fewer iteration steps are required for convergence. We demonstrate the performance and accuracy of the DISS and DIWS strategies for two complex 2-D heterogeneous wave simulation examples.

Key words: direct and iterative matrix solver, finite-difference methods, frequency-domain wave modelling, multi-grid method, seismic wave propagation, wavelet-multi-grid iterative solver.

1 INTRODUCTION

Seismic wave propagation modelling is a useful tool for studying the response of complex geological structures under excitation of seismic waves. Moreover, synthetic seismograms that result from efficient forward modelling schemes are necessary to invert, for example, multi-source–receiver seismic exploration data. Unfortunately, synthetic wave propagation simulations either allow fast and accurate computations but are incomplete, or they provide the required precision for the complete wavefield solution but then traditional schemes based on Cartesian coordinate systems are slow and dispersive, therefore demanding unrealistically high computer memory and CPU time resources. Many different approaches have been tried to resolve this discrepancy in order to make highly accurate numerical modelling feasible.

Nowadays 3-D and 4-D time-lapse data acquisition in seismic exploration and/or the instrumentation of sedimentary basins and foothills by large seismic arrays have become standard tools in seismology. In order to benefit from these enhanced data acquisition techniques, geoscientists need to simulate large frequency bandwidth data with increased numerical precision for increasingly larger complex models. In recent years, the choice of a forward modelling method best adapted for simulating large-scale geological models was dominated by

*Now at: Shell International Exploration and Production BV Volmerlaan 8, Postbus 60, 2280 AB Rijswijk, The Netherlands.

asymptotic methods, such as the ray-tracing approach (Cervený *et al.* 1977). Their main interest is high computational efficiency to allow the geological interpretation of large data volumes on a standard basis. Despite many advantages, these methods are based on a high-frequency approximation of the wave equation, and therefore provide incomplete results especially for complex models.

Full-waveform methods such as finite-difference (FD) techniques on regular grids provide complete solutions to the wave equation. As a result of developments of massively parallel computers their usage became feasible. Consequently, FD modelling schemes enjoy great popularity for 2-D problems since they provide accurate waveform results for complex structures, while, at the same time, their implementation stays rather simple (Virieux 1984, 1986). Even though FD methods are also applied to realistic 3-D models (Graves 1996; Olsen & Archuleta 1996) with a rather low-frequency content ($<$0.5 Hz), their widespread application is mainly limited by the huge computer memory and CPU time requirements. Simulations for large multi-source experiments stay an exception and are far from being a standard procedure (Mulder & Plessix 2002).

Over the years, advances in numerical wave propagation modelling were mainly aimed at either enhancing the performance of standard FD modelling schemes or at investigating new numerical approaches in order to handle the problems related to the simulation of large data volumes with sufficiently high accuracy. Graves (1996) and Moczo *et al.* (1999) developed memory optimization techniques that permit 3-D FD modellings on a single-processor desktop workstation. Other FD optimization techniques vary the spatial and temporal discretization of the simulation parameters (e.g. Falk *et al.* 1998; Tessmer 2000). The common idea behind these techniques is the distribution of computational power to model regions corresponding to the local model complexity. For example, zones with complex structure need small grid spacing and time stepping, while zones with smooth parameter variations can be modelled with relatively coarse spatial and temporal discretizations. Varying the spatial grid size in the model (Pitarka 1999), combining FD grids with a different spatial grid interval (Moczo 1980; Jastram & Behle 1992) and a varying time step (Falk *et al.* 1998; Tessmer 2000) enhances the numerical performance. Unfortunately, the wavefield transformation from one grid to the next is critical and rather complicated. Moreover, the accuracy of proposed implementations for strongly heterogeneous complex models is questionable. In order to prevent problems caused by manual grid changes, Operto *et al.* (2002) used the wavelet transform (WT) to introduce a time-domain FD (FDTD) modelling scheme on grids of different resolution. In their approach, grids of different spatial resolution and their interactions were controlled by a multi-resolution wavelet transform. Although in their approach spatial discretization steps are adapted to each multi-resolution grid (i.e. space adaptivity), grid couplings in the wavelet domain and convolution-type projection procedures resulted in rather high computational cost compared with standard FD techniques.

Hestholm & Ruud (1998) and Xie *et al.* (2002) formulated FD modelling schemes on curvilinear grids. Since free-surface boundary conditions for arbitrary topographies are implemented exactly, no further grid refinement is needed, which increases the efficiency of the FD scheme.

Pseudospectral methods enhance the computational efficiency and accuracy by computing spatial derivatives in one domain while the principal equations are solved in another domain (see Fuchs & Müller 1971). Moreover, they address problems related to the correct free-surface representation (Tessmer & Kosloff 1994; Igel 1999), which is the main drawback of FD methods defined on rectangular Cartesian grids. Unfortunately, pseudospectral methods are restricted to smooth surface variations only.

In recent years, other techniques that rely on mesh generation have dominated developments of forward modelling schemes. Faccioli *et al.* (1996), Komatitsch & Vilotte (1998) and Komatitsch & Tromp (1999) developed a spectral element scheme for wave propagation modelling, which is based on finite-element (FE) discretizations (Zienkiewicz & Morgan 1982). The main advantages are exact propagation of surface waves in the presence of topography and fast implementation schemes (domain decomposition using parallel computations). Fluid–solid interfaces and anisotropy have been included. More recently, their approach has been extended to mixed spectral elements (Cohen & Fauqueux 2000), which provide even higher computational efficiency and flexibility than the spectral elements approach.

Finite-element techniques require the computational grid to be constructed before the modelling starts. Therefore, mesh generation is a key step of FE approaches that controls the accuracy of the scheme (Sambridge *et al.* 1995). Inaccuracies may occur when the mesh is not constructed correctly. This will cause errors in the final wavefield solution that are not detectable by the user. Therefore, LeVeque (1997) formulated an FE modelling scheme for hyperbolic systems of equations based on standard Cartesian grids that dynamically applies grid refinements in both space and time. Still, if repeated wave propagation simulations have to be performed because small model parts have changed due to newly acquired information, e.g. for 4-D time lapse simulations, the complete mesh generation and wave propagation simulation have to be recomputed.

Another group of methods, the multi-grid approaches (Hackbusch 1978), perform simulations using direct and iterative matrix equation solvers on a set of numerical grids with different spatial resolution. Low-frequency (smooth) wavefield components on a given (fine-discretized) grid are difficult to eliminate by an iterative solver. Therefore, multi-grid techniques introduce a sequence of grids having different resolution, in order to treat the components of the approximate solution with respect to their characteristic frequency content on the best adapted discrete numerical grid. Ideally on each resolution grid, the grid spacing is chosen such that the represented component of the solution (e.g. that may be a rather low-frequency component on the finest discretized grid) becomes a high-frequency one with respect to the local grid spacing. Then iterative solvers may be applied effectively. Numerical grids are combined by numerous passages through so-called V- and W-cycles, where the combination of V- and W-structures depend on the behaviour of the solution during iteration steps. A good introduction to the multi-grid philosophy is given by Briggs (1987).

In this article, we propose a new FD scheme that addresses the problem of large-scale wave propagation simulations in complex media for multi-source experiments carried out on a standard basis. A typical example would be the exploration for hydrocarbons but also the instrumentation of a zone with high seismic risk such as, for example, Mexico City (Sanchez-Sesma 1983). Even though today spectral elements are superior over FD methods in terms of CPU time requirements and accuracy of the solution (free surface), the ambiguity caused

by the mesh generation and necessary complete recalculations for small model changes that are not required for FD modelling (Robertsson & Chapman 2000), motivates a deeper investigation of FD methods.

To the best of our knowledge, 3-D wave propagation simulations are mainly carried out by the FDTD rather than by frequency-domain finite differences (FDFD), because of higher computational efficiency (Mulder & Plessix 2002) and easy extension to distributed memory computers (Olsen *et al.* 1995). For multi-source experiments within the framework of waveform inversions, the method of choice has been the FDFD technique, because solutions for multiple right-hand sides are computed at minimal additional cost (Pratt 1990), and non-linear rheology, such as attenuation, is easily incorporated into the scheme. The main step in FDFD is the inversion of a large system of equations, and therefore FDFD schemes are limited by the performance of either direct matrix factorization or iterative solver schemes. Despite their advantages, both solver approaches (direct and iterative) have only been competitive for rather limited 2-D models (Mulder & Plessix 2002).

In order to benefit from the classic advantages of FDFD over FDTD modelling for large 2-D and 3-D media and extended frequency ranges ($f = 0$–10 Hz), we propose the combination of a direct and an iterative solver scheme in a multi-grid framework. The direct solver is mainly limited by the computer memory available. Therefore, the new direct iterative solver (DIS) will be used to compute an approximate FDFD solution for a coarse-grid model, that results from smoothing the fine-discretized model and accordingly adapting the spatial sampling to allow the run of the direct solver scheme. Then the iterative scheme is used to fill the gap between the coarse-grid solution and the desired fine-grid solution. Since the performance of an iterative solver scheme strongly depends on a stable preconditioned matrix and the initial solution provided, we will use the exact reduced-model solution to approximate the desired fine solution. This strategy is similar to a *nested iteration* procedure in multi-grid methods (Briggs 1987), and it permits accelerated convergence of the iterative part with respect to a zero initial solution. The critical point in our approach will be the combination of the two solver schemes, such that the iterative solver takes maximum advantage of the multi-grid framework.

In this paper, we first present a new FDFD modelling scheme that is based on a first-order hyperbolic formulation of the 2-D *SH*-wave equation. By means of this formulation we discuss the general problem of FDFD modelling. Then we introduce the general idea of a combination of a direct and an iterative solver scheme to overcome the limits of separate usage of only one of the schemes. Moreover, we present two different DIS multi-level approaches; the first one, called the direct iterative space solver (DISS), is similar to a *nested iteration* scheme and is based on space linear interpolation when moving from one grid to the other. The second one, called the direct iterative wavelet solver (DIWS), uses an intrinsic feature of an orthogonal wavelet transform and can be directly compared with a *full multi-grid* method (FMG). Both schemes will be discussed in detail and their performances are tested on two heterogeneous 2-D models. Although we show wavefield solutions for 2-D media only, the underlying aim of our developments are FDFD modellings of realistic complex 3-D media. Since developments carried out for 2-D models set important directions leading to the actual DIS approach, we present them here.

2 FREQUENCY-DOMAIN FORWARD MODELLING

Frequency-domain forward modelling in 2-D media for *SH*-wave propagation may be formulated using a velocity–stress finite-difference scheme that automatically leads to the staggered grid geometry developed in FDTD modelling (Virieux 1984). We write the *SH*-wave equation as a first-order system and construct the corresponding complex impedance matrix in the frequency domain.

We introduce the elastodynamic equations combined with Hooke's law for *SH*-wave propagation in the *y*-direction (Virieux 1984):

$$\begin{aligned}
\rho(x,z)\frac{\partial v_y(x,z,t)}{\partial t} &= \frac{\partial \sigma_{xy}(x,z,t)}{\partial x} + \frac{\partial \sigma_{zy}(x,z,t)}{\partial z} + S(x,z,t)\\
\frac{1}{\mu(x,z)}\frac{\partial \sigma_{xy}(x,z,t)}{\partial t} &= \frac{\partial v_y(x,z,t)}{\partial x}\\
\frac{1}{\mu(x,z)}\frac{\partial \sigma_{zy}(x,z,t)}{\partial t} &= \frac{\partial v_y(x,z,t)}{\partial z},
\end{aligned} \tag{1}$$

where $v_y(x, z, t)$ is the velocity field in the time domain, $\rho(x, z)$ is the density, $\mu(x, z)$ is the shear modulus, $\sigma_{xy}(x, z, t)$ and $\sigma_{zy}(x, z, t)$ are shear stresses, and $S(x, z, t)$ is the external source. With a slight abuse of notation we assume the shear stress and velocity components to be normalized (e.g. $\sigma_{xy} = \sigma_0 \cdot \sigma'_{xy}$ and $v_y = v_0 \cdot v'_y$, where σ'_{xy} and v'_y are the non-normalized wavefield components). The normalization becomes necessary in order to stabilize subsequent matrix computations. Following FDTD developments for electromagnetic waves, we add perfectly matched layer (PML) absorbing boundary conditions (Berenger 1994) to simulate infinite media boundaries. We separate spatial derivatives appearing in the first equation for the *x*- and *z*-directions and add a damping function term (γ) for each spatial component that only takes effect in the absorbing layer (see Operto *et al.* 2002, for details):

$$\begin{aligned}
\rho(x,z)\left[\frac{\partial v_y^{(x)}(x,z,t)}{\partial t} + \gamma_x(x)\cdot v_y^{(x)}(x,z,t)\right] &= \frac{\partial \sigma_{xy}(x,z,t)}{\partial x} + S(x,z,t)\\
\rho(x,z)\left[\frac{\partial v_y^{(z)}(x,z,t)}{\partial t} + \gamma_z(z)\cdot v_y^{(z)}(x,z,t)\right] &= \frac{\partial \sigma_{zy}(x,z,t)}{\partial z}\\
\frac{1}{\mu(x,z)}\left[\frac{\partial \sigma_{xy}(x,z,t)}{\partial t} + \gamma_x(x)\cdot \sigma_{xy}(x,z,t)\right] &= \frac{\partial v_y^{(x)}(x,z,t)}{\partial x} + \frac{\partial v_y^{(z)}(x,z,t)}{\partial x}\\
\frac{1}{\mu(x,z)}\left[\frac{\partial \sigma_{zy}(x,z,t)}{\partial t} + \gamma_z(z)\cdot \sigma_{zy}(x,z,t)\right] &= \frac{\partial v_y^{(x)}(x,z,t)}{\partial z} + \frac{\partial v_y^{(z)}(x,z,t)}{\partial z}.
\end{aligned} \tag{2}$$

The two velocity wavefields $v_y^{(x)}$ and $v_y^{(z)}$ can be combined to the physical v_y wavefield through simple addition $v_y = v_y^{(x)} + v_y^{(z)}$. We transform the system of equations in the Fourier domain and introduce the new variables $\xi_x(x, \omega) = 1 + i\gamma_x(x)/\omega$ and $\xi_z(z, \omega) = 1 + i\gamma_z(z)/\omega$ to simplify the equations:

$$
\begin{aligned}
&i\omega\xi_x(x,\omega)\rho(x,z)\cdot v_y^{(x)}(x,z,\omega) + \frac{\partial\sigma_{xy}(x,z,\omega)}{\partial x} = -S(x,z,\omega)\\
&i\omega\xi_z(z,\omega)\rho(x,z)\cdot v_y^{(z)}(x,z,\omega) + \frac{\partial\sigma_{zy}(x,z,\omega)}{\partial z} = 0\\
&i\omega\xi_x(x,\omega)\frac{1}{\mu(x,z)}\cdot\sigma_{xy}(x,z,\omega) + \frac{\partial v_y^{(x)}(x,z,\omega)}{\partial x} + \frac{\partial v_y^{(z)}(x,z,\omega)}{\partial x} = 0\\
&i\omega\xi_z(z,\omega)\frac{1}{\mu(x,z)}\cdot\sigma_{zy}(x,z,\omega) + \frac{\partial v_y^{(x)}(x,z,\omega)}{\partial z} + \frac{\partial v_y^{(z)}(x,z,\omega)}{\partial z} = 0.
\end{aligned}
\tag{3}
$$

The resulting system of eqs (3) is solved for each angular frequency component ω_i separately.

2.1 Matrix construction

The system of eqs (3) is discretized using a staggered grid geometry (Virieux 1984; Levander 1988) and recast into matrix-type form ($\mathbf{A} \times \mathbf{x} = \mathbf{b}$). In the case of the velocity–stress formulation of the first-order hyperbolic wave equation, mainly two different matrix construction techniques are possible. One may follow standard FDTD modelling approaches (Graves 1996) and multiply density $\rho(x, z)$, shear modulus $\mu(x, z)$, angular frequency ω and the PML functions $\xi_x(x, \omega)$ and $\xi_z(z, \omega)$ by the partial derivatives of the wavefield components. We follow another strategy that combines all model and PML contributions on the diagonal matrix entries, because it allows the development of rather simple and computationally fast algorithms. Then the matrix-type equation that is equivalent to the system of eqs (3) reads as

$$
\underbrace{\begin{pmatrix} i\omega\xi_x\rho & 0 & \partial/\partial x & 0\\ 0 & i\omega\xi_z\rho & 0 & \partial/\partial z\\ \partial/\partial x & \partial/\partial x & i\omega\xi_x\frac{1}{\mu} & 0\\ \partial/\partial z & \partial/\partial z & 0 & i\omega\xi_z\frac{1}{\mu}\end{pmatrix}}_{\mathbf{A}}
\underbrace{\begin{pmatrix} v_y^{(x)}\\ v_y^{(z)}\\ \sigma_{xy}\\ \sigma_{zy}\end{pmatrix}}_{\mathbf{x}} =
\underbrace{\begin{pmatrix} -S\\ 0\\ 0\\ 0\end{pmatrix}}_{\mathbf{b}}.
\tag{4}
$$

The elements of the impedance matrix $\mathbf{A}$ are complex valued and depend on the medium properties $\rho(x, z)$ and $\mu(x, z)$ and on the angular frequency ω. Since each point on the FD grid is coupled to its nearest neighbours only, the matrix $\mathbf{A}$ is sparse. We denote the model dimensions by n_x and n_z for the x- and z-directions, respectively. The full matrix contains $(4n_xn_z)^2$ elements of which, in the case of second-order staggered grid FD derivative approximations, $6n_x^2 + 6n_z^2 + 4n_{nz}^x$ are different from zero. Only diagonal matrix blocks depend on the physical model parameters and on the frequency. The off-diagonal blocks i.e. spatial derivative operators, are constructed only once for all frequency components ω_i before the modelling starts. Moreover, the non-zero frequency-dependent coefficients are all distributed on the diagonal. Therefore, constructing the complex impedance matrix for several angular frequency components can easily be performed by simply replacing diagonal matrix coefficients. In the current implementation, the matrix is constructed in sparse compact sparse row (CSR) or Harwell–Boeing (HB) format, which allows fast matrix computations, such as matrix–vector products.

2.2 The direct iterative solver approach

The bottleneck in any frequency-domain modelling scheme is the factorization of the complex impedance matrix (Štekl & Pratt 1998). In general, the desirable matrix inversion is not performed, since the sparse matrix structure will result in a rather dense inverse matrix, and therefore the number of non-zero coefficients to compute or even to hold the inverse matrix in memory is too large to be acceptable. To avoid the computation of the inverse matrix, typically two different solution strategies are applied. An effective decomposition might be obtained by computing the $\mathbf{LU}$ factorization of the impedance matrix. In this case, which is called the direct solver approach, the system $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ is converted into two triangular systems using Gaussian elimination, which then permits fast computations of solutions (Golub & van Loan 1996). The algorithm separates the matrix $\mathbf{A}$ into a lower triangular ($\mathbf{L}$-factor) and an upper triangular ($\mathbf{U}$-factor) part of the form $\mathbf{L} \cdot \mathbf{U} = \mathbf{A}$. The fill-in that is created during the separation procedure (the number of non-zero coefficients exceeding the number of non-zero coefficients of the initial matrix), is significantly reduced compared with the computation of the true inverse, and highly optimized preconditioners and scaling techniques make this procedure fast and convenient (Štekl & Pratt 1998). Once the $\mathbf{LU}$ factors have been computed for each frequency component separately, they can be stored and wavefield solutions for multiple right-hand sides are obtained at low computational cost (Pratt 1990). The direct solver approach is only limited by the computational fill-in to compute and store the $\mathbf{LU}$ factors.

The second approach provides an approximate matrix equation solution by an iterative scheme. The direct solver approach might be impractical because the $\mathbf{LU}$ factors need to be constructed and additional matrix fill-in occurs during the factorization, which may significantly increase memory requirements. In contrast, iterative methods generate a sequence of approximate solutions x^i, which are supposed to converge to the true solutions for some fixed tolerance error. An update of solutions is mainly performed through matrix–vector multiplications and dot-product computations and therefore fast optimized computer structures can be exploited to the maximum. Moreover, significantly larger model

sizes may be treated, because numerical fill-in does not occur. The drawback is the finding of a good initial solution that closely resembles the desired true solution and an adapted iterative framework (e.g. multi-grids), such that fast iterative convergence can be guaranteed. In general, this is not at all a trivial task and modelling approaches exist that start iterations from a zero initial solution (Janod 1999). Moreover, for each right-hand side the iteration procedure has to be repeated and one of the main advantageous features of frequency-domain modelling is lost, if convergence of the iterative scheme is not fast enough.

3 THE DIRECT ITERATIVE SPACE SOLVER

Suppose we want to perform FDFD simulations for huge 2-D models, large number of right-hand sides and frequency ranges where, for example, the corresponding real data could result from a marine seismic experiment. Even though for one right-hand side (source), FDTD would be the most suitable method for the simulation, the number of repeated computations for the numerous right-hand sides renders the FDTD approach inefficient. If we further assume that FDFD modelling using a direct solver technique is excluded because of the size to hold or factorize the complex impedance matrix, the only possibility of performing an FDFD simulation would be via an iterative procedure. Such a scenario, which becomes even more true for 3-D modelling, underlines the need for an FDFD modelling scheme that is capable of treating large volume data. Nevertheless, FDFD modelling using an iterative approach suffers from similar restrictions as encountered for FDTD modellings for multi-source experiments, since iterations have to be computed for each right-hand side. The new scheme should profit to a maximum from the performance and uniqueness of wavefield solutions obtained by a direct solver, while at the same time it should have the capability to treat large earth models for high-frequency ranges at an acceptable computational cost for multiple right-hand sides.

In view of this demand and the limits and strengths of the two existing solver techniques (direct and iterative solver) we propose their combination in a multi-grid formulation to form the direct iterative solver (DIS) approach. The key point in this combination will be the connection of the two solver schemes, i.e. how a good initial iterative solution may be obtained and how the iterative convergence may be accelerated significantly.

3.1 Combination of two solver schemes

We combine the direct solver with the iterative solver by a *nested iteration* formulation. Typically, *nested iteration* is applied in a V-cycle of a *full multi-grid* approach. 2-D *SH*-wavefield solutions (u^{2h}) are computed for a restricted coarse discretized model (V^{2h}) by the direct solver ($\mathbf{A}^{2h}\ u^{2h} = b^{2h}$), where h is the fine-grid spatial discretization step. Coarse, in the sense used here, implies that a finer model (V^h) FD discretization would be desirable but is computationally too expensive to be computed by the direct solver approach. We use a highly optimized multi-frontal technique to compute the **LU** matrix factorization, called MUMPS (Amestoy *et al.* 2001).

The next step involves the combination of the coarse-grid solution with an iterative solver scheme in order to obtain the desired solution for a fine model discretization (u^h). Following the *nested iteration* approach, we prolongate the exact coarse-grid solution (u^{2h}) on a fine discretized grid using a simple bilinear interpolation scheme ($V^{2h} \Rightarrow V^h$). The prolongation operator is denoted by $\mathbf{I}_{2h}^h$. The prolongation result will then serve as an initial solution (v^h) for a GMRES iterative solver scheme developed by Frayssé *et al.* (1997):

$$\mathbf{I}_{2h}^h u^{2h} = v^h. \tag{5}$$

We denote this approach as the direct iterative space solver. The DISS approach is similar to the *nested iteration* part of the multi-grid V-cycle since the prolonged coarse-grid solution provides an acceptable initial solution for the GMRES iterative solver. A complete FMG V-cycle scheme applies sequentially the nested iteration with V-cycles (Briggs & Henson 1993). The nested iteration part provides an initial guess on the fine grid to initiate the subsequent V-cycles. The V-cycles define a path with several V-shapes joining grids of various resolution (see Briggs & Henson 1993; Press *et al.* 1994, for an illustration of these paths). The change from one grid to a finer one is achieved through a prolongation operator $\mathbf{I}_{2h}^h$ (eq. 5). The change to a coarser grid is obtained through the restriction operator $\mathbf{R}_h^{2h}$:

$$\mathbf{R}_h^{2h} u^h = v^{2h}. \tag{6}$$

We will show in the next section the close connection between FMG V-cycle methods and the DIWS algorithm and how the wavelet transform handles the restriction and prolongation operations.

A typical FMG V-cycle sequence requires performing six main elementary tasks:

(1) compute the initial solution v^h on a grid Ω^h through relaxation of $\mathbf{A}^h\ u^h = f^h$;
(2) compute the residual $r^h = f - \mathbf{A}v^h$;
(3) restrict the residual r^h on the coarser grid Ω^{2h}: $\mathbf{R}_h^{2h} r^h = r^{2h}$;
(4) obtain the coarse-grid correction e^{2h} by solving the residual equation $\mathbf{A}^{2h}\ e^{2h} = r$ on Ω^{2h}; the residual equation can be solved either by an iterative or a direct method if possible;
(5) interpolate the correction e^{2h} on Ω^h: $\mathbf{I}_{2h}^h e^{2h} = e^h$;
(6) correct the approximation on Ω^h : $v^h \leftarrow v^h + e^h$.

We compared the convergence of the iterative scheme using the prolonged coarse-grid wavefield solution and a zero solution as a first initial approximation for the iterative scheme. Even though in both cases the iterative scheme converged within the given tolerance limits, the simulation using the *nested iteration* formulation required 50 per cent less CPU time.

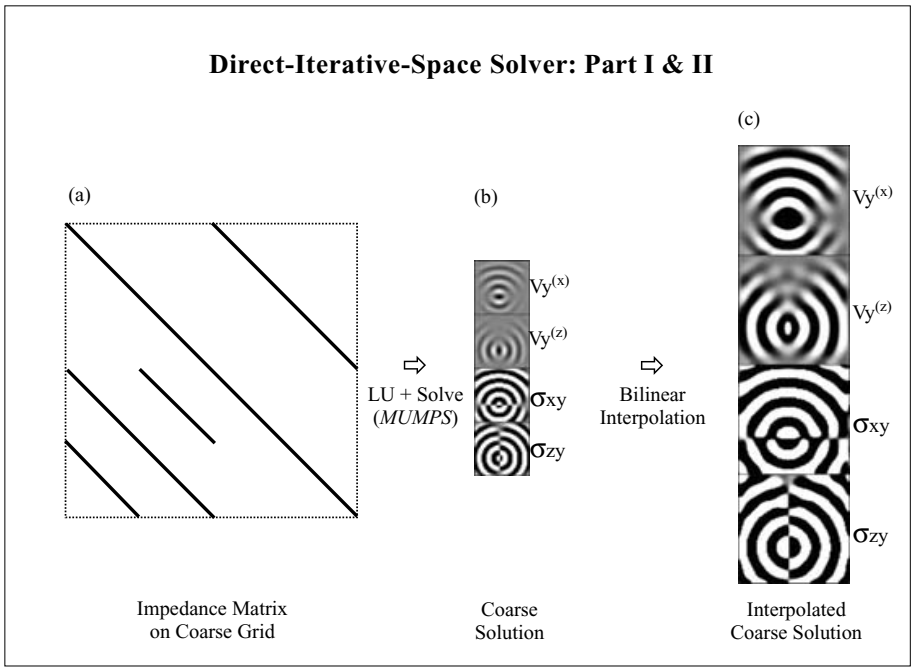

Figure 1. Flowchart showing the direct iterative solver approach: part I and II. (a) We construct the complex impedance matrix for a coarse model discretization and compute its **LU** factorization using the MUMPS direct solver. (b) We compute coarse-grid solutions for each source position and frequency component. (c) Coarse solutions are interpolated on a fine discretized grid using a simple bilinear interpolation procedure.

For both direct and iterative solver schemes, we construct the complex impedance matrix **A** directly for the corresponding grid discretization. In the case where the number of non-zero elements of the matrix for the fine-grid iterative part exceeds the actual free memory available, construction *on the fly*, though costly, is possible. This is to say, that the iterative scheme does not necessarily need the matrix available in core memory at a given iteration step. In contrast, the matrix only contributes to the iteration through matrix–vector multiplications, dot-products and matrix preconditionings. Therefore, matrix coefficients contributing to the actual computations may be constructed when needed (Pessel 2000). The complete DISS procedure is shown in Figs 1 and 2.

4 THE DIRECT ITERATIVE WAVELET SOLVER

We present a second DIS strategy that instead of space grid prolongations, relies on a rather natural multi-level formulation through orthogonal wavelet expansions. The approach is denoted by DIWS. The wavelet transform permits us to construct wavelet projected fine-grid wavefield solutions directly, where contributions from all spatial frequency components and resolution approximation grids are incorporated in the desired adapted way (Briggs & Henson 1993; Beylkin 1998). The wavelet-based DIWS approach is directly comparable to a full multi-grid method, where iterations are performed simultaneously on a sequence of adapted resolution grids without the necessity to perform artificial V-and W-cycles. In the following, we demonstrate the enhanced iterative convergence of the DIWS scheme. For completeness, we use Daubechies wavelets with two vanishing moments (referred to as the Daubechies-4 wavelets in the following) throughout. The problem of how to find the best adapted wavelet basis for the DIWS scheme is not addressed here.

4.1 The $V_J \times V_J$ space in a wavelet expansion

Without reviewing the theory of the discrete wavelet transform (DWT) we introduce some of its features that are especially important for our study.A short and comprehensive introduction of the wavelet transform can be found in Operto *et al.* (2002). For a detailed discussion of the DWT, we refer to Mallat (1989) or Daubechies (1992) for further readings.

The DWT decomposes a signal on a sequence of nested finite-difference grids ($V_0 \supset V_1 \supset V_2 \supset \ldots \supset V_j \supset V_{j+1} \supset \ldots \supset V_J$) of decreasing resolution ($V_0 \Rightarrow V_J$). The finest-resolution grid is denoted by V_0 and the coarsest by V_J. The basis for each resolution approximation space V_j is constructed through dilations and translations of the so-called *scaling function*. The signal decomposition on scaling spaces of varying resolution is the basis of the multi-resolution analysis (MRA). The scaling spaces are complemented by the wavelet spaces W_j. The wavelet spaces are constructed by the *wavelet function*. The wavelet space encodes the information that is necessary to interchange between different resolution scaling spaces ($V_j \oplus W_j = V_{j-1}$). Together, the *scaling* and *wavelet* functions build an orthogonal non-redundant wavelet transform.

With a slight abuse of notation, we set the $V_0 \times V_0$ space equal to the initial signal discretized with the finest spatial discretization step. This assumes equality of discretized function values and scaling function coefficients $s_{0,k} = f_k = f(x_0 + k \cdot \Delta)$, where the finest

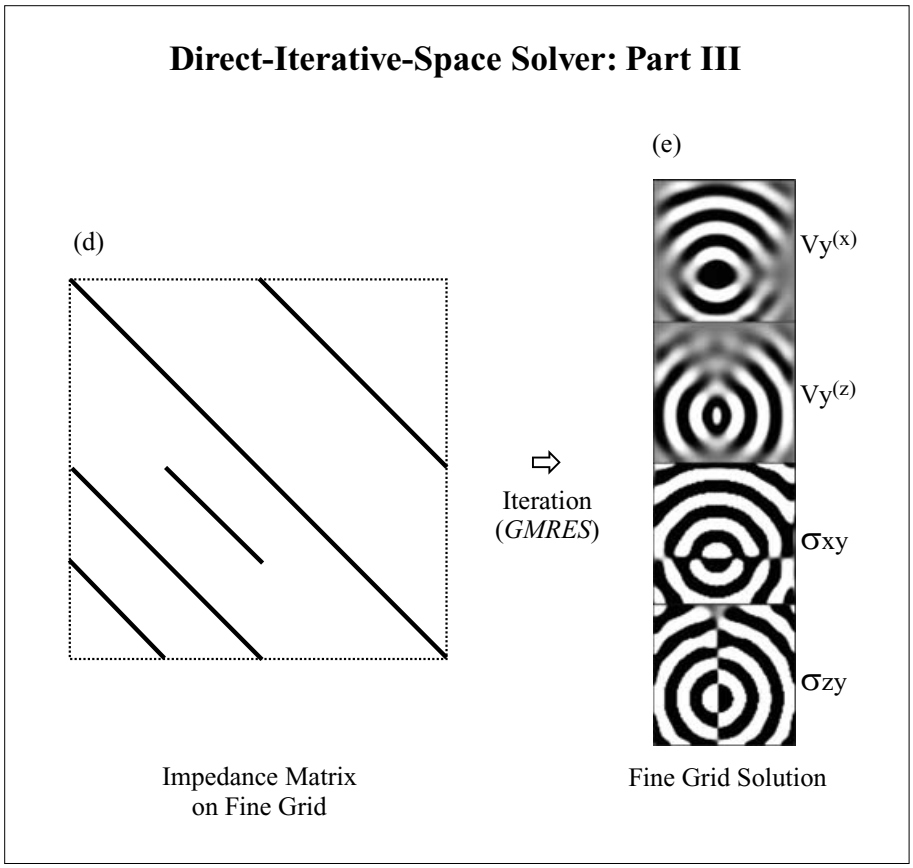

Figure 2. Flowchart showing the direct iterative solver approach: part III. (d) We construct the complex impedance matrix for the fine discretized grid. The interpolated coarse solution in Fig. 1(c) approximates the desired fine-grid solution. Therefore, we use it as initial solution in a GMRES iterative solver scheme (e), which gives the final frequency map solution.

approximation is set to $j_0 = 0$ and Δ is the spatial sampling, which obviously is not right. For Daubechies wavelets, the so-called *wavelet crime* is permitted, because the Daubechies scaling function approximately verifies an *interpolating* property of the form $s_{0,k} = \delta_{0,k} \cdot f$. For details we refer to the discussions by Strang & Nguyen (1996).

For example, we suppose a signal being projected on an orthogonal Daubechies wavelet basis consisting of $J = 3$ resolution scales:

$$P_0 f = \sum_{k\epsilon Z} s_{1,k}\phi_{1,k} + \sum_{j=1}^{J=3}\sum_{k\epsilon Z} d_{j,k}\psi_{j,k}, \tag{7}$$

where f_k are discrete function samples, k is the spatial position parameter, and s_k and d_k are the corresponding discrete scaling and wavelet functions samples.

In Fig. 3(a) we show a snapshot of a wavefield propagating in a 2-D homogeneous medium and in Fig. 3(c) the corresponding projection on an orthogonal wavelet basis. In the wavelet terminology, $J = 3$ resolution scales mean that the signal is distributed corresponding to its frequency content on three different grids of decreasing resolution, where each subspace $W_j \times W_j$ encodes the information that is lost when going from a fine to the next coarsest scale $V_j \Rightarrow V_{j+1}$. The spatial discretization of the signal is doubled while changing to the next coarsest grid. Additionally, the DWT is accompanied by the $V_J \times V_J$ space, which represents an average of the signal initially discretized on the $V_0 \times V_0$ space (see Fig. 3d). The projection of the signal on different resolution grids is shown schematically in Fig. 3(b).

The orthogonal wavelet expansion of a signal provides the following important features.

(1) We find the signal being decomposed on several grids of decreasing resolution up to a pre-defined coarsest representation of the initial signal (scaling coefficients).

(2) The wavelet and interaction spaces ($W_j \times W_j$, $W_j \times V_j$ and $V_j \times W_j$) control the information transfer between the coarsest V_J and the finest-resolution V_0 level grid. Therefore, spatial aliasing that typically occurs in FMG schemes is avoided.

(3) As was pointed out by Briggs & Henson (1993) and Beylkin (1998), the wavelet transform therefore provides a tool to suppress V- and W-cycles in FMG techniques, since grids of decreasing resolution and the corresponding grid interactions are treated simultaneously.

(4) The wavelet expansion provides an adapted signal decomposition where frequency components are automatically represented in the best suited resolution grid, with respect to the spatial discretization. Misinterpretation of spatial frequency components may be significantly reduced.

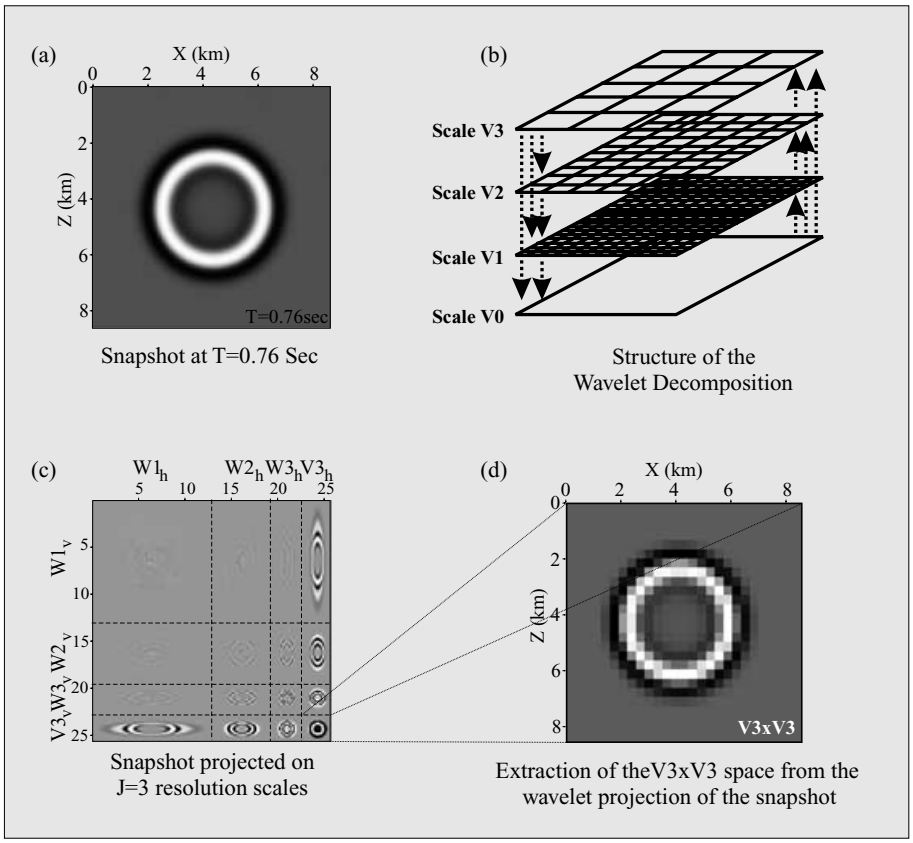

Figure 3. A 2-D signal is projected on an orthogonal wavelet basis with $J = 3$ resolution scales. (a) Time-snapshot of wave propagation in a homogeneous medium after $T = 0.76$ s. (b) Structure of the wavelet transform on $J = 3$ resolution scales. The initial signal in the $V_0 \times V_0$ space is decomposed on three subsequently coarser grids ($V_{i=1,...,3} \times V_{i=1,...,3}$ spaces). Arrows denote interaction between different resolution scales. (c) The initial signal decomposed on $J = 3$ resolution scales. The $W_{i=1,...,3} \times W_{i=1,...,3}$ spaces store the coefficients that are needed to change from one to the next resolution scale. The $V_3 \times V_3$ space is the coarsest discretized representation of the initial signal. All other contributions in spaces $V_{i=1,...,3} \times W_{i=1,...,3}$ and $W_{i=1,...,3} \times V_{i=1,...,3}$ control the interaction between different resolution grids. The wavelet transform stores a low-frequency representation in the $V_3 \times V_3$ space and subsequently higher frequency contributions in the $W_{i=1,...,3} \times W_{i=1,...,3}$ spaces. When the $V_{i=1,...,3} \times V_{i=1,...,3}$ spaces are summed up with the $W_{i=1,...,3} \times W_{i=1,...,3}$ spaces and including interactions ($W_{i=1,...,3} \times V_{i=1,...,3}$ and $V_{i=1,...,3} \times W_{i=1,...,3}$) spaces we obtain the initial signal projection in the $V_0 \times V_0$ space. (d) We extract the $V_3 \times V_3$ space from the wavelet transform in (c). This signal is a low-frequency (coarse discretized) version of the initial signal shown in (a).

4.2 The DIWS algorithm

The DIWS algorithm, in general, resembles that from the DISS, though now the iterative scheme will be applied to the wavelet expansion coefficients rather than to the prolonged coarse-grid solution in the physical space. We propose the application of the wavelet transform to construct fine-grid solutions instead of manual prolongation by bilinear interpolation. Then, the DIWS approach resembles a wavelet-based FMG method, where V- and W-cycles become obsolete. Since spatial frequency components are automatically decomposed on the corresponding wavelet resolution spaces optimal iteration may be performed and therefore the computational efficiency may be increased significantly. We compute coarse-grid wavefield solutions for each frequency component ω_i and the right-hand side using the MUMPS direct solver (see Figs 4a and b). The obtained coarse-grid solutions are recast in a wavelet expansion of the desired fine-grid solution, where all wavelet coefficients are initially zero: $W_1 \times W_1 = V_1 \times W_1 = W_1 \times V_1 = 0$ (see Fig. 4c). In contrast, the scaling coefficients $V_1 \times V_1$, are approximated by the coarse-grid solution (see Fig. 4c). The impedance matrix on the coarse grid is constructed through a wavelet-projection procedure. Initially, we define the physical model parameters, such as the density, on the desired fine discretized grid, e.g. the density model ρ^{h_0}, where h_0 is the fine-grid discretization step. In order to obtain a coarse-grid model representation ρ^{h_J} (J denotes the coarse grid), we transform the fine-grid models on the same wavelet basis as that used later for the iterative scheme. From the wavelet expansion we extract the coarse-grid model contribution in the scaling space and use it as a physical parameter input model for the coarse-grid matrix construction. For example, the fine-grid density model ρ^{h_0} is projected on J resolution scales:

$$P_0(\rho) = \sum_{k \epsilon Z} s_{J,k} \phi_{J,k} + \sum_{j=1}^{J} \sum_{k \epsilon Z} d_{j,k} \psi_{j,k} \tag{8}$$

where $\rho^{h_J} = s_J \epsilon V_J \times V_J$. The scaling space $V_J \times V_J$ provides the coarse-grid density model ρ^{h_J}. The model construction process is illustrated in Fig. 5. The right-hand side model is obtained in the same way. Note that the amplitude of the coarse-grid solution is scaled with respect to the wavelet projection on a scaling resolution approximation space ($V_J \times V_J$).

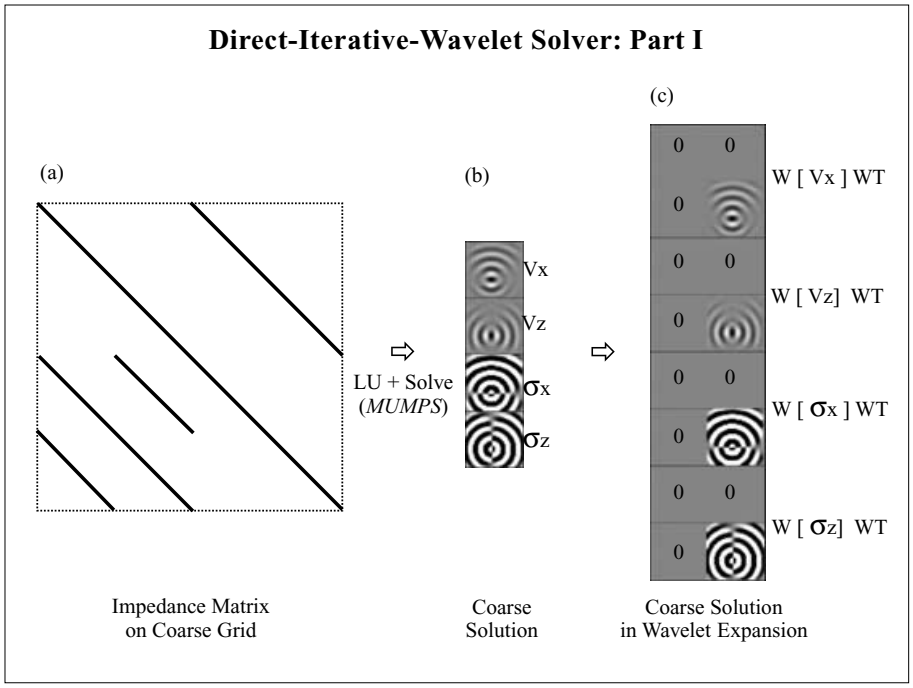

Figure 4. The DIWS: part I. (a) The complex impedance matrix is constructed for a coarse-grid discretization in the physical space. The matrix is factorized into its **LU** factors by the MUMPS direct solver and a wavefield solution for one frequency component is obtained by solving the resulting system of equations (b). In (c) the calculated coarse-grid wavefield solution is then recast into the scaling coefficients of the wavelet expansion of the desired fine-grid solution ($V_1 \times V_1$ space), where all wavelet and interaction coefficients are initially zero ($W_1 \times W_1 = V_1 \times W_1 = W_1 \times V_1 = 0$).

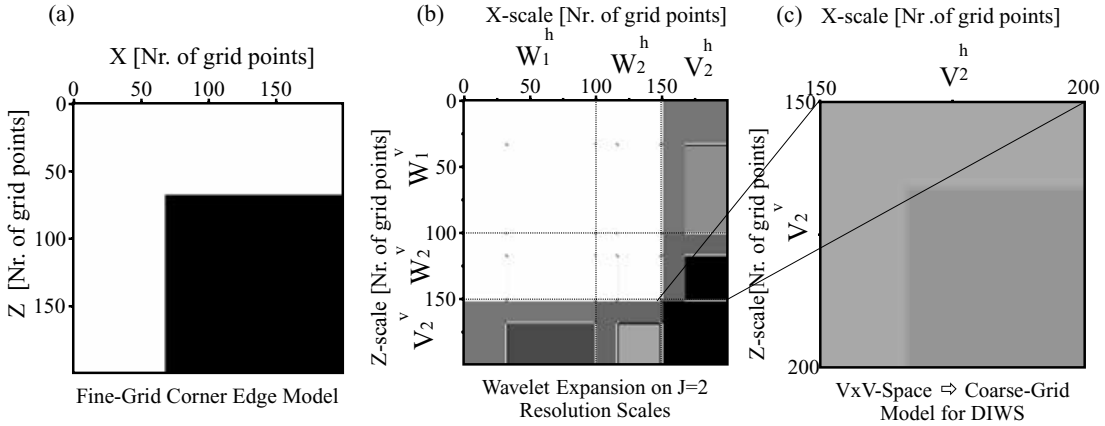

Figure 5. Illustration of the procedure to define a physical parameter model used for the coarse-grid impedance matrix construction of the DIWS modelling scheme. The corner edge model is initially defined for the fine-grid DIWS configuration. Note that in this illustration we plotted the density model, that is $\rho =$ 2000 kg m^{-3} in the upper and $\rho = 2500$ kg m^{-3} in the lower medium. In (a) the fine-grid model consists of 200 $\times$ 200 grid-nodes. The model is projected on the Daubechies-4 wavelet basis with $J = 2$ resolution scales (b). From the expansion we extract the $V_2 \times V_2$-scaling space. This coarse-scale representation of the initial fine-grid corner edge model is used for the coarse-grid impedance matrix construction (c). A similar coarse-grid model construction is applied for the source function.

In this wavelet construction, all wavelet and interaction coefficients are not known and need to be added somehow. We apply the iterative solver scheme to construct the missing fine-grid wavelet coefficients, while the initial coarse-scale solution may be updated if necessary (see Fig. 6e). The complex impedance matrix is directly constructed in the wavelet domain (see Fig. 6d).

The application of the iterative solver on the wavelet basis brings some significant advantages compared with the application in the physical domain. As was discussed before, the missing wavelet coefficients represent the projection of the true solution on different resolution scales incorporating all interactions between the scales. *A priori*, we construct wavelet coefficients simultaneously on all wavelet resolution scales. The wavelet expansion provides a natural decomposition with respect to the frequency content of the signal on a sequence of scales of decreasing resolution. We may apply the DIWS projection scheme, such that local wavelet coefficients on a given (coarse) resolution scale become relatively high frequency, even though they are low-frequency components when discretized on the initial fine grid. Then, the iterative solver effectively prolongates the numerical solution on the fine grid and, consequently, V- and W-cycles become obsolete. The simultaneous

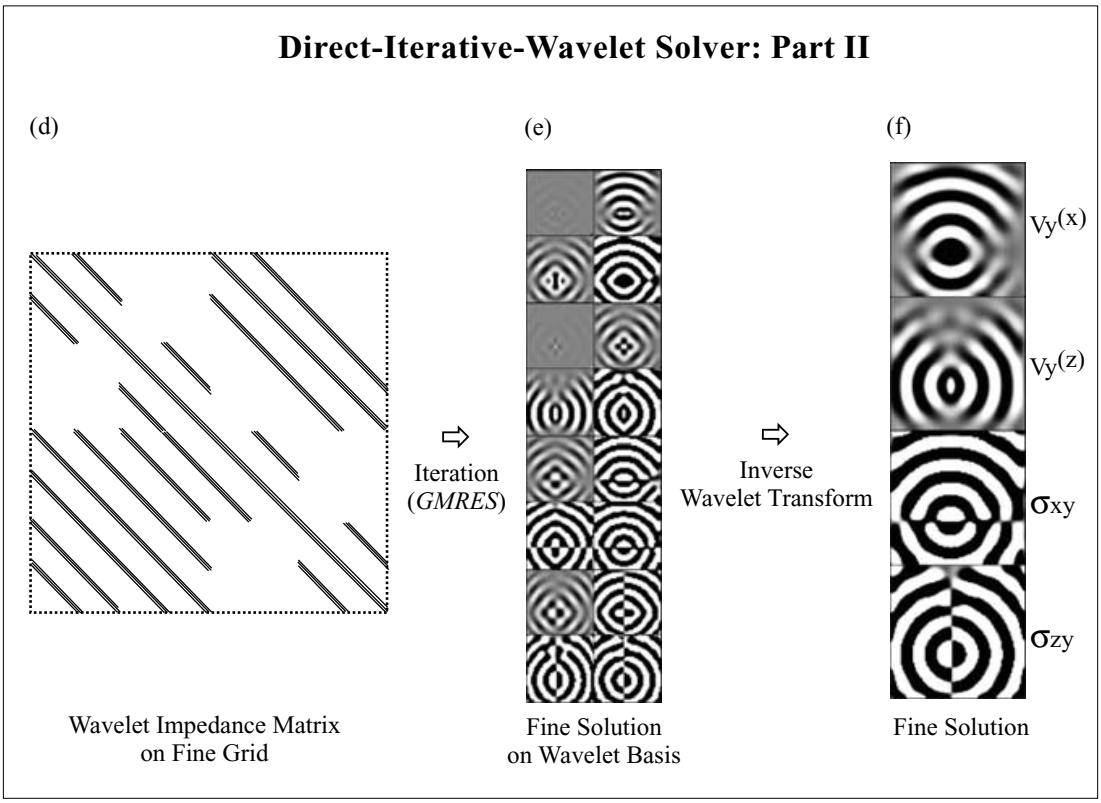

Figure 6. The DIWS: part II. (d) We construct the complex impedance matrix for the fine-grid discretization on the orthogonal Daubechies-4 wavelet basis. Wavefield solutions are computed by a GMRES iterative solver, where the initial solution is the coarse-grid solution expanded in the wavelet basis (see Fig. 4c). The iterative solver constructs all missing wavelet and interaction coefficients (e). No additional V- and W-cycles are necessary to correct for spatial aliasing and interpolation phase shifts. The wavefield in the physical space domain is obtained by an inverse wavelet transform (f).

iteration on all resolution scales may then significantly accelerate the iterative convergence. Finally, the desired fine-grid solution in the physical space can be obtained by applying the inverse wavelet transform (see Fig. 6f).

4.3 First-order formulation

Initially, we compute a coarse-grid wavefield solution using the direct solver scheme for the largest possible model configuration that can be handled by the presently available computer resources. We then construct the complex impedance matrix in the wavelet domain. This is explained in detail in the following.

We formulate the *SH*-wave equation as a first-order hyperbolic system of equations with PML absorbing boundary conditions in the frequency domain (see eq. 3). In order to construct the complex impedance matrix in the wavelet domain, we project the system of eqs (3) on a (2-D) orthogonal wavelet basis with respect to the spatial coordinates x and z:

$$
\begin{aligned}
&\mathbf{W}\left[i\omega\xi_x\rho\cdot v_y^{(x)}\right]\mathbf{W}^{\mathrm{T}}+\mathbf{W}\left[\frac{\partial\sigma_{xy}}{\partial x}\right]\mathbf{W}^{\mathrm{T}}=\mathbf{W}[-S]\mathbf{W}^{\mathrm{T}}\\
&\mathbf{W}\left[i\omega\xi_z\rho\cdot v_y^{(z)}\right]\mathbf{W}^{\mathrm{T}}+\mathbf{W}\left[\frac{\partial\sigma_{zy}}{\partial z}\right]\mathbf{W}^{\mathrm{T}}=0\\
&\mathbf{W}\left[i\omega\xi_x\frac{1}{\mu}\cdot\sigma_{xy}\right]\mathbf{W}^{\mathrm{T}}+\mathbf{W}\left[\frac{\partial v_y^{(x)}}{\partial x}\right]\mathbf{W}^{\mathrm{T}}+\mathbf{W}\left[\frac{\partial v_y^{(z)}}{\partial x}\right]\mathbf{W}^{\mathrm{T}}=0\\
&\mathbf{W}\left[i\omega\xi_z\frac{1}{\mu}\cdot\sigma_{zy}\right]\mathbf{W}^{\mathrm{T}}+\mathbf{W}\left[\frac{\partial v_y^{(x)}}{\partial z}\right]\mathbf{W}^{\mathrm{T}}+\mathbf{W}\left[\frac{\partial v_y^{(z)}}{\partial z}\right]\mathbf{W}^{\mathrm{T}}=0,
\end{aligned}
\tag{9}
$$

where $\mathbf{W}$ and $\mathbf{W}^{\mathrm{T}}$ denote direct and inverse wavelet transform operators, respectively. In the system of eqs (9) three different kinds of terms need to be projected on the wavelet basis. The first term is the product of medium parameters ρ and μ with the wavefield components e.g. $v_y^{(x)}$, etc. The second and third terms are spatial derivative operators $\partial/\partial x$ and $\partial/\partial z$ in x- and z-directions, respectively. We will discuss the projection on the orthogonal wavelet basis for each term separately. The complete wavelet transformation is provided in the Appendix. The

equations in matrix-type form are:

$$\begin{pmatrix} \mathbf{W}[i\omega\xi_x\rho]\mathbf{W}^{\mathrm{T}} & 0 & \mathbf{W}\left[\frac{\partial}{\partial x}\sigma_{xy}\right]\mathbf{W}^{\mathrm{T}} & 0 \\ 0 & \mathbf{W}[i\omega\xi_z\rho]\mathbf{W}^{\mathrm{T}} & 0 & \mathbf{W}\left[\frac{\partial}{\partial z}\sigma_{zy}\right]\mathbf{W}^{\mathrm{T}} \\ \mathbf{W}\left[\frac{\partial v_y^{(x)}}{\partial x}\right]\mathbf{W}^{\mathrm{T}} & \mathbf{W}\left[\frac{\partial v_y^{(z)}}{\partial x}\right]\mathbf{W}^{\mathrm{T}} & \mathbf{W}\left[i\omega\xi_x\frac{1}{\mu}\right]\mathbf{W}^{\mathrm{T}} & 0 \\ \mathbf{W}\left[\frac{\partial v_y^{(x)}}{\partial z}\right]\mathbf{W}^{\mathrm{T}} & \mathbf{W}\left[\frac{\partial v_y^{(z)}}{\partial z}\right]\mathbf{W}^{\mathrm{T}} & 0 & \mathbf{W}\left[i\omega\xi_z\frac{1}{\mu}\right]\mathbf{W}^{\mathrm{T}} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{W}[v_y^{(x)}]\mathbf{W}^{\mathrm{T}} \\ \mathbf{W}[v_y^{(z)}]\mathbf{W}^{\mathrm{T}} \\ \mathbf{W}[\sigma_{xy}]\mathbf{W}^{\mathrm{T}} \\ \mathbf{W}[\sigma_{zy}]\mathbf{W}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} \mathbf{W}[-S]\mathbf{W}^{\mathrm{T}} \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{10}$$

The products of physical medium parameters with the wavefield components are located around the diagonal in the matrix. The wavelet projection of these products requires the computation of a convolution-type system:

$$\begin{aligned} &\mathbf{W}[i\omega\xi_x(x,\omega)\rho(x,z)v_y^{(x)}(x,z,\omega)]\mathbf{W}^{\mathrm{T}} \\ &\quad = \sum_{\alpha,\beta}\sum_{i,j}\left\langle v_y^{(x)}(x,z,\omega),\psi_j\psi_\beta\right\rangle\langle\psi_j\psi_\beta, i\omega\xi_x(x,\omega)\rho(x,z)\psi_i\psi_\alpha\rangle\psi_i\psi_\alpha, \end{aligned} \tag{11}$$

where $\psi_{i,j}$ and $\psi_{\alpha,\beta}$ denote wavelet and scaling basis functions in the x- and z-directions, respectively. The combined inner products form a term-to-term product of the wavelet expansions of medium parameters $i\omega\xi_x\rho$ and wavefield components $v_y^{(x)}$. The wavelet projection of the diagonal terms (eq. 11) plays a key role for the CPU time requirements of the impedance matrix construction on an orthogonal wavelet basis. A similar computational procedure was encountered by Operto *et al.* (2002), who performed wavelet-based FD modelling in the time domain. The DIWS requires convolution projections (eq. 11) to be computed only once before the modelling starts.

However, computational inefficiency related to costly convolution projections are likely to become obsolete or, simplifying, the standard situation. The incorporation of increasingly complex wave phenomena in modelling schemes, in the end will lead to non-linear wave propagation simulations. In this context, the rather simple linear partial differential equation used here, is replaced by a non-linear one, which by definition will require the computation of products as encountered in eq. (11). While standard techniques at this point will have to be modified significantly, the wavelet-based DIWS will not require more computational power than before.

We have preferred the development of the DIWS scheme using a first-order hyperbolic formulation instead of the second-order standard FDFD formulation. The reason becomes clear when estimating the computational effort necessary to construct the complex impedance matrix in the wavelet domain. In the case of the first-order formulation, the construction of one coefficient $c_{i\alpha} = \mathbf{W}[i\omega\xi_i\rho_{i\alpha}]\mathbf{W}^{\mathrm{T}}$ in the matrix (eq. 10) requires the multiplication of medium parameters with the wavelet basis functions ψ_j and ψ_β. The term-to-term product is then computed via the matrix–vector product through corresponding storage of transformed terms in the matrix. We are interested in fast computations of the diagonal terms in order to keep the computational cost of the DIWS scheme at a minimum. The Daubechies-4 quadrature mirror filters that define the wavelet expansion consist of four coefficients only and therefore lead to fast convolution projection algorithms.

In order to evaluate the computational cost for one coefficient $\mathbf{W}[i\omega\xi_x\rho\cdot v_y^{(x)}]\mathbf{W}^{\mathrm{T}}$ we assume dense basis functions. Then the convolution computation for the first-order hyperbolic system is proportional to N^4 floating point operations, where N denotes the spatial model dimensions $n_x = n_z = N$.

We compare the first-order hyperbolic matrix to the standard second-order FDFD formulation combined with a similar projection on an orthogonal wavelet basis (see the Appendix for details). The coupled spatial derivatives applied to the product of medium parameters and wavefield components is projected on the wavelet basis. This gives:

$$\begin{aligned} &\mathbf{W}\left\{\frac{\partial}{\partial x}\left[\rho(x,z)\frac{\partial}{\partial x}\xi_x(x,\omega)\right]v_y^{(x)}(x,z,\omega)\right\}\mathbf{W}^{\mathrm{T}} = \sum_{\alpha,\beta}\sum_{i,j,k,l}\left\langle v_y^{(x)}(x,z,\omega),\psi_l\psi_\beta\right\rangle \\ &\quad\times\left\langle\frac{\partial}{\partial x}\psi_k,\psi_l\right\rangle\left\langle\frac{\partial}{\partial x}\psi_i,\psi_j\right\rangle\langle\psi_k\psi_\beta,\xi_x(x,\omega)\rho(x,z)\psi_j\psi_\alpha\rangle\psi_i\psi_\alpha. \end{aligned} \tag{12}$$

The construction of the second-order analogues of the complex impedance matrix requires the computation of two term-to-term products plus the product of medium parameters and wavelet basis functions. As before, we assume dense basis functions to evaluate the computational cost, and obtain N^6.

Since the second-order hyperbolic wave equation combined with the wavelet transformation is proportional to N^6 floating point operations compared with N^4 for the first-order hyperbolic formulation, we developed the DISS and DIWS solver schemes using the latter approach. Moreover, the first-order hyperbolic velocity–stress formulation of the wave equation allows the introduction of more complex rheological laws (non-linear elasticity) at almost no additional cost.

4.4 Derivative projection on wavelet basis

To complete the wavelet expansion of terms appearing in the complex impedance matrix we discuss the projection of spatially staggered derivative operators on an orthogonal Daubechies wavelet basis. The entire development is shown in the Appendix.

We apply the orthogonal wavelet transform to the product of spatial derivatives in x- and z-directions and wavefields $v_y^{(x)}$, $v_y^{(z)}$, σ_{xy} and σ_{zy}. As an example, we give the results for the velocity wavefield components $v_y^{(x)}$ and $v_y^{(z)}$:

$$\mathbf{W}\left[\frac{\partial}{\partial x}v_y^{(x)}(x,z,\omega)\right]\mathbf{W}^{\mathrm{T}} = -\sum_{\alpha}\sum_{i,j}\left\langle\frac{\partial}{\partial x}\psi_i,\psi_j\right\rangle\left\langle v_y^{(x)}(x,z,\omega),\psi_i\psi_\alpha\right\rangle\psi_i\psi_\alpha \tag{13}$$

$$\mathbf{W}\left[\frac{\partial}{\partial z}v_y^{(z)}(x,z,\omega)\right]\mathbf{W}^{\mathrm{T}} = -\sum_{\alpha,\beta}\sum_{i}\left\langle\frac{\partial}{\partial z}\psi_\alpha,\psi_\beta\right\rangle\left\langle v_y^{(z)}(x,z,\omega),\psi_i\psi_\beta\right\rangle\psi_i\psi_\alpha. \tag{14}$$

In practice, the derivative operator may be constructed in the wavelet domain following two different approaches. The first one was developed by Beylkin *et al.* (1991) and Jameson (1993). They directly derived the operator coefficients for FD derivatives from the definition of scaling and wavelet interpolating functions of several different orthogonal Daubechies wavelet bases.

In contrast, Wu & McMechan (1998) discretized the spatial FD operator in the physical domain (which approximately corresponds to the FD operator in the V_0 space) and then computed its projection in the wavelet domain. Their approach is independent of the properties of scaling and wavelet functions, as long as they form an orthogonal basis.

In the case of the Daubechies-4 wavelet basis, both approaches lead to the same operator coefficients in the wavelet basis, if the standard FD operator in the V_0 space is the non-staggered centred difference stencil $(-1/12, +2/3, 0, -2/3, +1/12)$. Moreover, Jameson (1993) showed that the projection of this FD operator on a Daubechies-4 orthogonal wavelet basis results in a local FD operator on each resolution scale together with interaction spaces that relate different FD approximations on different resolution scales.

For simplicity, we follow the latter approach by Wu & McMechan (1998) and construct discrete staggered derivative operators in the physical space and transform then in the wavelet space.

5 NUMERICAL EXAMPLES

We show two *SH*-wave propagation simulations in strongly heterogeneous media. Both examples have been computed for the DISS and DIWS using similar physical model parameters and source and receiver setups in order to allow a relative comparison of the two approaches. Even though we show 2-D examples only, we would like to remind the reader that the DISS and DIWS approaches were developed in view of large-scale wavefield simulations in complex 3-D media. However, at this stage we address a verification of the code together with first insights to the convergence performance of the space bilinear interpolation and wavelet-based preconditioning of the complex impedance matrix. The wave propagation simulations currently run on a single desktop computer in sequential mode. This significantly limits the actual model size treated, though first wave simulations and proofs of the method are not negatively influenced by the 2-D simulation. The transition from 2-D to 3-D necessarily includes the transformation of the sequential code to run on large state-of-the-art supercomputers (e.g. PC clusters) using MPI- (message passing interface, Forum 1994) based parallelization of the matrix construction and the direct and iterative solver. Provided the outcome of the 2-D numerical tests, we will develop strategies for the desired real-case 3-D simulations that may be compared with other highly optimized modelling schemes, such as a spectral-element code.

We used the same source configuration for all model simulations. The source is spatially set on one grid-node without smoothing over neighbouring points:

$$S(x,z,t) = S(t)\delta(x-x_s)\delta(z-z_s). \tag{15}$$

Thanks to the staggered-grid geometry of the spatial stencils, this simple implementation of the source provides quite accurate simulations. Although we observed a slight overestimation of the amplitude of the FD seismograms at receivers located in the vicinity of the source when compared with analytical seismograms (see Hustedt *et al.* 2003, for more details).

The time signature is the derivative of a Gaussian,

$$S(t) = -2\alpha(t-t_0)\mathrm{e}^{-\alpha(t-t_0)^2} \tag{16}$$

with $\alpha = 200$, $t_0 = 0.31$ s and a maximum frequency of 10 Hz. In order to simulate infinite model dimensions, we added PML absorbing boundary layers on all four edges of the model. Problems related to the free-surface implementation are not addressed at this stage. The PML attenuating function, denoted by γ in eq. (2) ($\xi = 1 + i\gamma/\omega$), smoothly rises from 0 at the interior model–PML boundary up to some empirically estimated value at the outer PML boundary (Operto *et al.* 2002).

5.1 The corner edge model

The first wave simulation is performed for the corner edge model for two different direct iterative grid combinations. The corner edge model consists of two homogeneous spaces, which are separated by a horizontal and a vertical interface that build a corner. The medium surrounding the corner edge has a wave velocity of $v = 2000$ m s^{-1} and a density of $\rho = 2000$ kg m^{-3}, while the medium inside the corner has a wave velocity of $v = 8000$ m s^{-1} and a density of $\rho = 2500$ kg m^{-3}. The significant physical parameter contrast between the two spaces creates a strong reflection from the interface and a diffraction from the corner point. The physical model size is 6400×6400 m^2 with an additional PML layer of 800 m thickness on all four edges of the model. All wave simulations were carried out using fourth-order derivative approximations.

Frequency maps are calculated separately for 103 frequency components that equidistantly sample a frequency band from 0 to 10 Hz. This corresponds to a time sampling of $dt = 0.01$ s and a time-length of the seismogram of 6 s.

5.1.1 Simulation 1: two-step case

The first numerical corner edge simulation for the DISS/DIWS schemes is set up for a so-called *two-step* geometry, that is equivalent to a two-grid multi-grid scenario. The coarse-grid wavefield solution is calculated on a numerical grid with a spatial discretization of at least 2.5 points per shortest wavelength for the slowest velocity present in the model. We then halved the discretization step and computed the fine-grid contributions using the iterative solver part. The name *two-step* denotes the fact that the spatial discretization for the direct and iterative computations of the DISS scheme changes by a factor of 2, which is similar to a DIWS projection on *two* resolution scales ($J = 1$).

Both solver approaches compute coarse-grid wavefield solutions using the direct solver *MUMPS* in sequential mode. The coarse-grid spatial sampling was set to $dx = dz = 80$ m, which gives a numerical grid of 100×100 grid-nodes including 10 nodes for the PML layers on each side of the model. The coarse-grid discretization leads to 2.5 grid-nodes per shortest wavelength. Fine-grid solutions are calculated on a grid having double size, i.e. 200×200 grid-nodes including 20 nodes for the PML. The fine-grid spatial discretization is reduced from 80 to 40 m. The source is located at the position ($x = 4000$ m, $z = 250$ m).

Wavefield solutions are computed for a fixed number of iteration steps of $n_{\mathrm{it}} = 1500$. In Fig. 7, we show an example of a fine-grid wavefield solution computed by the DISS for the frequency $f \approx 8$ Hz. From left to right, we plot the frequency maps for the wavefield solutions v_y, σ_{xy} and σ_{zy} (see eq. 4). Note that the velocity wavefield has been recombined from its PML components $v_y = v_y^{(x)} + v_y^{(z)}$. Clearly, visible are wavefield reflections (**R**), head wave (**H**), transmitted waves (**T**) and diffractions (**D**) from the horizontal interface and the corner point where physical model parameters change significantly from one grid-node to the next.

In Figs 8 and 9, we show DISS seismograms computed from the coarse-grid direct solutions (Fig. 8) and the results from the fine-grid iterative solver (Fig. 9). 80 receivers were placed along the upper surface (without PML boundaries) with a spacing of 80 m. In addition to the complete seismograms we also include three extracted traces, namely trace numbers 10, 50 and 75. The extracted traces are compared with an analytical solution that was computed for a homogeneous medium with the parameters of the upper full space in the corner edge model. The analytical solution is added to verify the correct simulation of the direct wave, and moreover to demonstrate possible numerical dispersion on far offset traces.

The seismograms in Fig. 8 are affected by a significant amount of noise, because numerical dispersion limits for FDFD modelling in strongly contrasted media are not respected properly. Still, the direct first arrival and the reflection of the interface between the two media are clearly visible. The interpolated coarse-grid frequency maps serve as an initial solution for the iterative scheme. From Fig. 9, we observe that halving the spatial discretization step and applying an iterative solver, produces a dispersion-free wavefield simulation. We detect direct and reflected (**R**) arrivals and note the presence of a weak diffraction event (**D**), which is excited by the corner edge.

In general, the PML absorbing boundary conditions work well for the corner edge model, though in Fig. 9 we clearly remark a spurious reflection (**N**) around trace 75 and $t = 3$ s, which propagates from left to right. Numerical tests performed to evaluate the performance and accuracy of wave absorption using PML worked well for structurally simple models. For more complex models, such as the corner edge, our rather empirical PML study did not reveal a simple and sufficiently accurate PML attenuating function and needs to be investigated in more

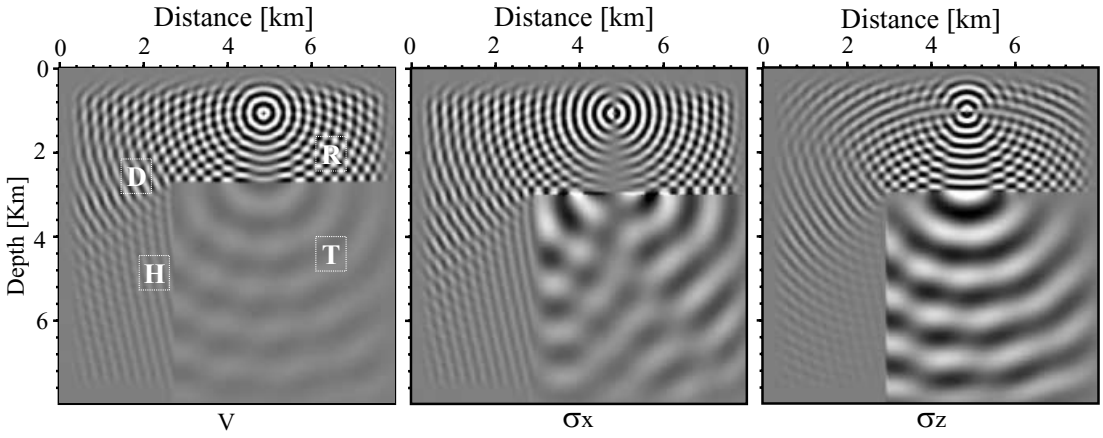

Figure 7. Wavefield solution for a fine-grid corner edge simulation. Frequency maps are computed by the DISS for the frequency $f \approx 8$ Hz. From left to right, we plot frequency maps for the three wavefield solutions v_y, σ_{xy} and σ_{zy} for the first-order hyperbolic system of equations in eq. (4). Note that the velocity wavefield is recombined from its PML components ($v_y = v_y^{(x)} + v_y^{(z)}$). Clearly, visible are strong reflections (**R**) as well ashead (**H**) and transmitted (**T**) waves from the horizontal interface. Diffraction phenomena from the corner point where the wave speed and model density change significantly from one gridpoint to the next are indicated by (**D**). On all four edges of the model we note wave attenuation in the PML absorbing layers.

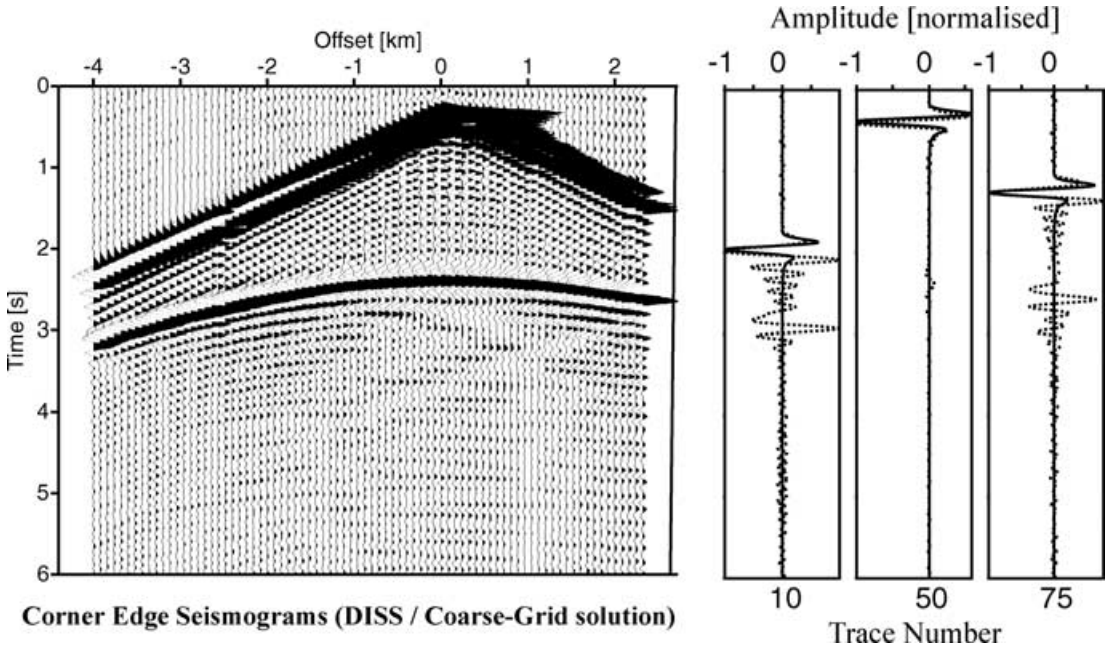

Figure 8. Seismograms for the corner edge model computed from the coarse-grid direct solution of the DISS. The seismogram includes arrivals from 80 receiver stations equally separated right across the model (receiver spacing, 80 m; receiver depth, 150 m). We also plotted three extracted traces (trace nos 10, 50, 75) together with an analytic solution (solid lines) for the upper model full space, which allows the verification of the direct wave only. We used 2.5 gridpoints per shortest wavelength to fix the coarse discretization step. Therefore, the time seismogram shows a significant amount of noise due to wave dispersion.

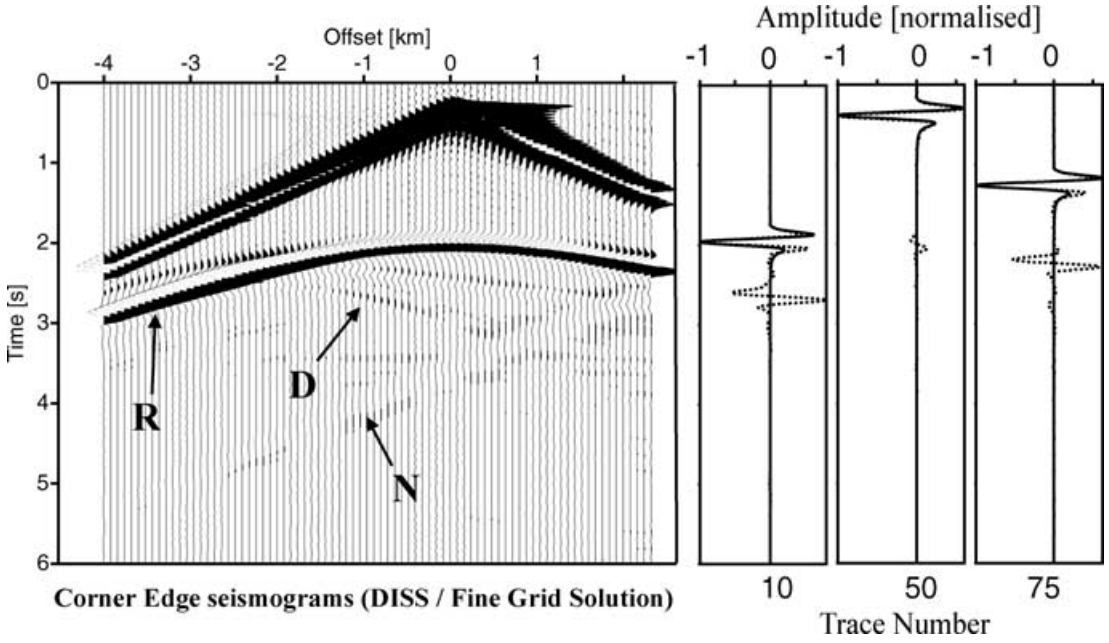

Figure 9. Seismograms for the corner edge model computed from the fine-grid iterative solution of the DISS. We clearly distinguish between wavefield reflections (R) and diffractions (D) from the horizontal interface and the corner point, where physical model parameters change significantly from one grid-node to the next. We observe good attenuation of the wave amplitude for all four PML absorbing layers surrounding the model. Note the spurious reflection that travels from right (trace 75, $t = 3$ s) to the left (N). This noise is caused by improper treatment of the PML absorbing boundary conditions. Its amplitude, though largely enhanced here, is approximately similar to the rather weak diffraction event from the corner edge located in the middle of the seismogram. On the right, we have plotted three extracted traces (trace nos 10, 50, 75) together with an analytical solution (solid lines) for the upper model full space.

detail. In this article, we are mainly interested in the comparison of the DISS and DIWS method with respect to performance and accuracy of the iterative scheme, and therefore leave the PML problem aside.

In Fig. 10, we show the DIWS corner edge seismogram that is computed using similar modelling parameters to those used for the DISS simulation for a *two-scale* geometry. In wavelet terminology, a *two-scale* geometry corresponds to a projection of the wavefield on two resolution scales ($J = 1$).

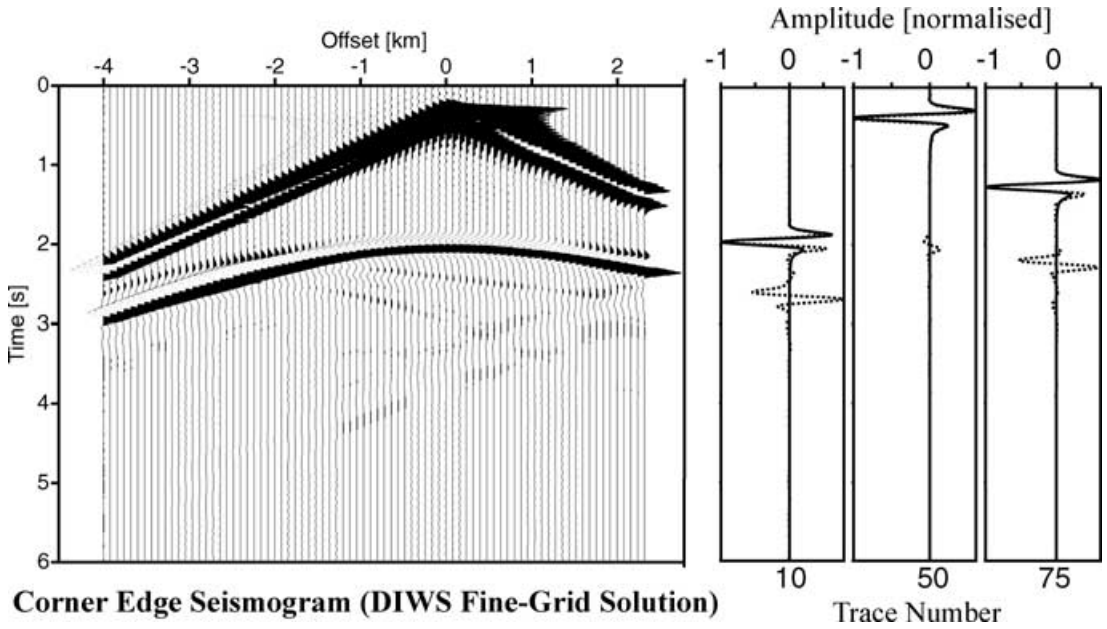

Figure 10. Seismogram for the corner edge model computed from the fine-grid iterative solution of the DIWS. Wave arrivals have been computed for 80 receiver stations having a horizontal distance of 80 m and depth of 150 m. We plot three extracted traces (trace nos 10, 50, 75) together with an analytic solution for the upper model full space, which allows the verification of the direct wave only.

The initial solution for the iterative solver in the wavelet domain consists of the coarse-grid solution recast in the $V_1 \times V_1$ space, where fine grid and interaction coefficients in spaces $W_1 \times W_1$, $W_1 \times V_1$ and $V_1 \times W_1$ are initially zero. Then, the *GMRES* iterative solver constructs the missing wavelet coefficients on all resolution scales $W_j \times W_j$ simultaneously and the grid-interaction coefficients in spaces $W_1 \times V_1$ and $V_1 \times W_1$. More important is that the iterative solver also modifies the coarse-grid scaling coefficients, such that the low-frequency components are adapted on the fine grid automatically with respect to the fine-grid high-frequency content. The modification of the scaling space is similar to the *coarse-grid correction* scheme of a *full multi-grid* method, while the construction of wavelet coefficients corresponds to a *relaxation* process on the fine grid. The scale interactions may be interpreted as the *interpolation* and *restriction* procedures in a V-and W-cycle. DIWS-wavelet iteration performs complete multi-grid cycles simultaneously at each iteration step, and therefore may provide superior iterative convergence rates.

In Fig. 11, we plot a final frequency map ($f \approx 8$ Hz) in the wavelet basis iteration (from left to right, we plot v_y, σ_{xy} and σ_{zy}). Each wavefield in the wavelet domain consists of the $V_1 \times V_1$ space, the $W_1 \times W_1$ space, and two interaction spaces $V_1 \times W_1$ and $W_1 \times V_1$. We note that coefficients with strong amplitude are constructed in the wavelet and interaction spaces for the upper medium, while the corner edge

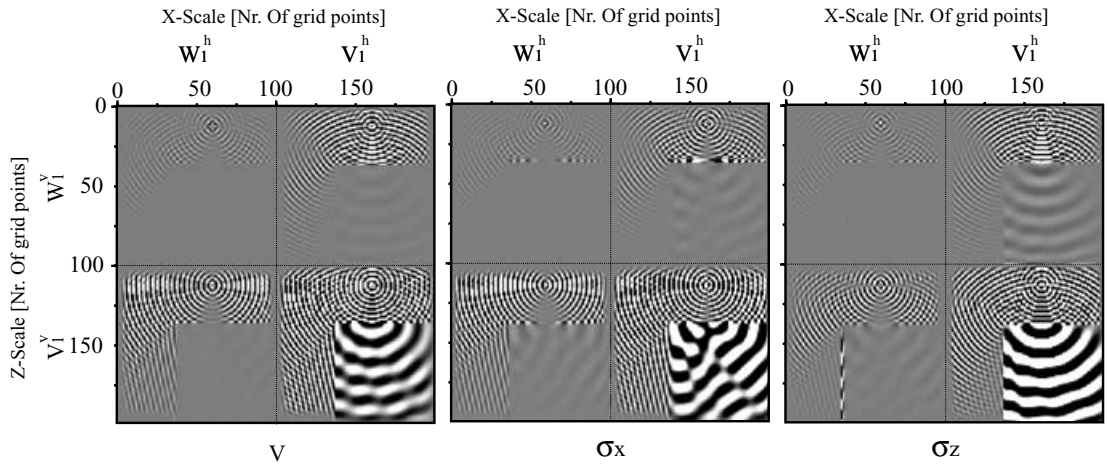

Figure 11. Wavefield solution for the DIWS fine-grid corner edge simulation. Frequency maps are computed by the DIWS for the frequency $f \approx 8$ Hz. From left to right, we plot frequency maps for all wavefield solutions v_y, σ_{xy} and σ_{zy} on two wavelet resolution scales ($J = 1$) using the Daubechies-4 wavelet basis. The upper left-hand block is the wavelet space ($W_1 \times W_1$), upper right and lower left blocks interaction spaces ($V_1 \times W_1$ and $W_1 \times V_1$) and the lower right block the scaling space ($V_1 \times V_1$). Clearly, visible are strong reflections from the horizontal interface and diffraction phenomena from the corner point where the wave speed and model density change significantly from one gridpoint to the next. On all four edges of the model we also note the wave attenuation in the PML absorbing layers.

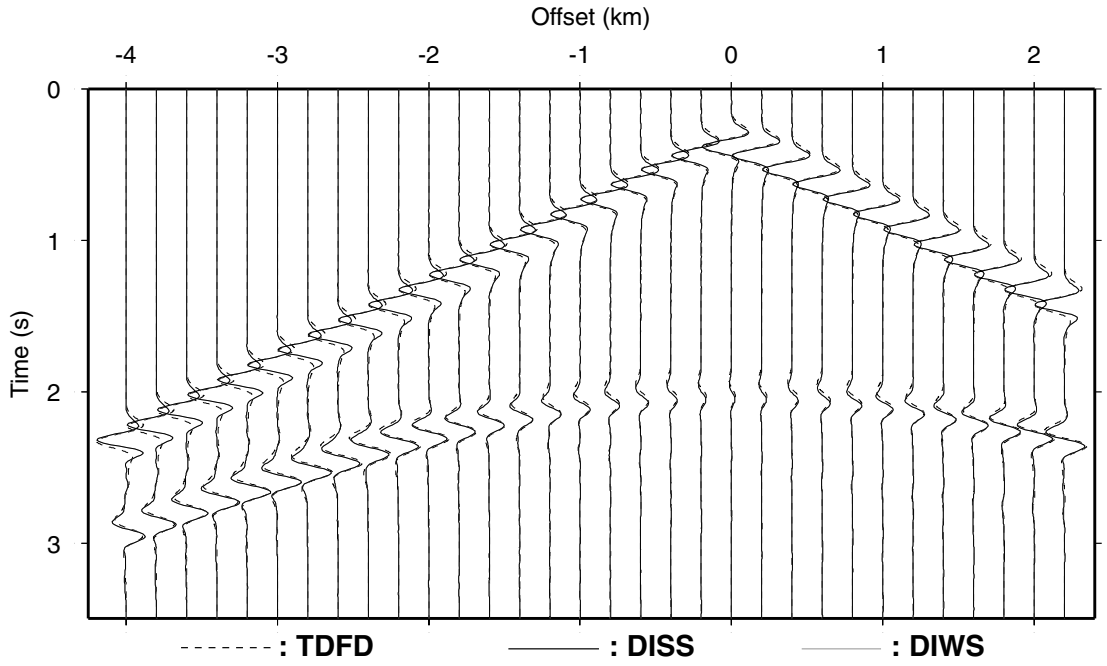

Figure 12. Comparison of time seismograms for the corner edge model computed by the DISS, DIWS and FDTD modelling schemes. The dotted line shows the FDTD solution. The black and grey solid lines show the DISS and DIWS solutions, respectively.

medium is represented throughout by coefficients with weak amplitude. Moreover, the wavelet decomposition clearly identifies the strong velocity contrast at the interface between the two corner edge spaces. The heterogeneous distribution of coefficients with varying amplitude illustrates the strong influence of the wavelength change ($v = \lambda \times f$) due to the change in physical parameters between the two media of the corner edge model.

The distribution of wavelet coefficients with respect to the complexity of the model structure and the inherent frequency content in a wavelet formulation allows the definition of space adaptivity through the introduction of numerical masks (Operto *et al.* 2002). We may define model regions where wavelet coefficients do not contribute to the wavefield simulation in certain approximation spaces. An adaptation to a local physical parameter that characterizes the behaviour of the wavefield solution may be obtained by inspecting the smoothness of the model structure. A typical region to apply a mask would be the corner edge space on the fine-grid approximation scales in Fig. 11, where wavelet coefficients are small and therefore do not provide significant contributions to the wavefield propagation at this frequency component. Although we expect great CPU time savings from the definition of masks, at this initial stage of the DISS–DIWS comparison we did not address its implementation in the DIWS algorithm.

We compare the DISS and DIWS wavefield solutions to a time seismogram that was computed using a standard fourth-order staggered FDTD modelling scheme. The DISS, DIWS and FDTD seismograms are shown in Fig. 12. The overall *fit* of the three solutions is good. We observe slight phase shifts between the FDTD and the FDFD solutions for the reflection event from the corner edge interface. This illustrates the approximate character of the iterative wavefield solutions computed by the DISS and DIWS. An increase of *GMRES* iteration steps eliminates these artefacts.

Differences in CPU time between the two DIS methods are related to the iterative part of the algorithm, since both use the MUMPS solver for the coarse-grid computations. Moreover, the DISS and DIWS prolongation of the coarse-grid solution is negligible from a computational point of view. Therefore, we compare the DISS and DIWS performance for the matrix construction and convergence of the iterative scheme. At this stage, we are mainly interested in the comparison of the DISS and DIWS iterative convergence behaviour under spatial and wavelet-based preconditioning by an initial coarse-grid solution (nested iteration).

The DIWS impedance matrix needs to be constructed analytically in the wavelet domain by projecting matrix subblocks in the wavelet space (see eq. 10). Spatial derivative expansions and convolution-type wavelet projections for diagonal matrix blocks (see eq. 11) are performed only once before the modelling starts. For each frequency component, the DIWS matrix is then easily constructed by combining pre-computed terms with the complex angular frequency ω_i. In the current implementation, we did not seek to optimize the sequential convolution computations which results in large DIWS matrix construction times. For example, the construction of all diagonal matrix blocks in eq. (10) for the DIWS corner edge simulation on $J = 1$ resolution scales, requires $t_{A_{\mathrm{DIWS}}} \approx 70$ min. In contrast, the DISS fine-grid impedance matrix construction in the space–frequency domain for the same corner edge model is done at practically no computational cost ($t_{A_{\mathrm{DISS}}} \approx 1$ s). Note that all computations are carried out on a stand-alone Linux PC with 1 GByte of RAM and 1 GHz speed.

The *GMRES* iterative solver is set to perform a fixed number of iterations $n_{\mathrm{iter}}^{\max} = 1500$ with a tolerance (backward error) of $\mathrm{tol} = 1 \times \mathrm{e}^{-7}$. Note that the case where the tolerance is fixed and the number of iterations required for convergence is investigated, is not addressed at

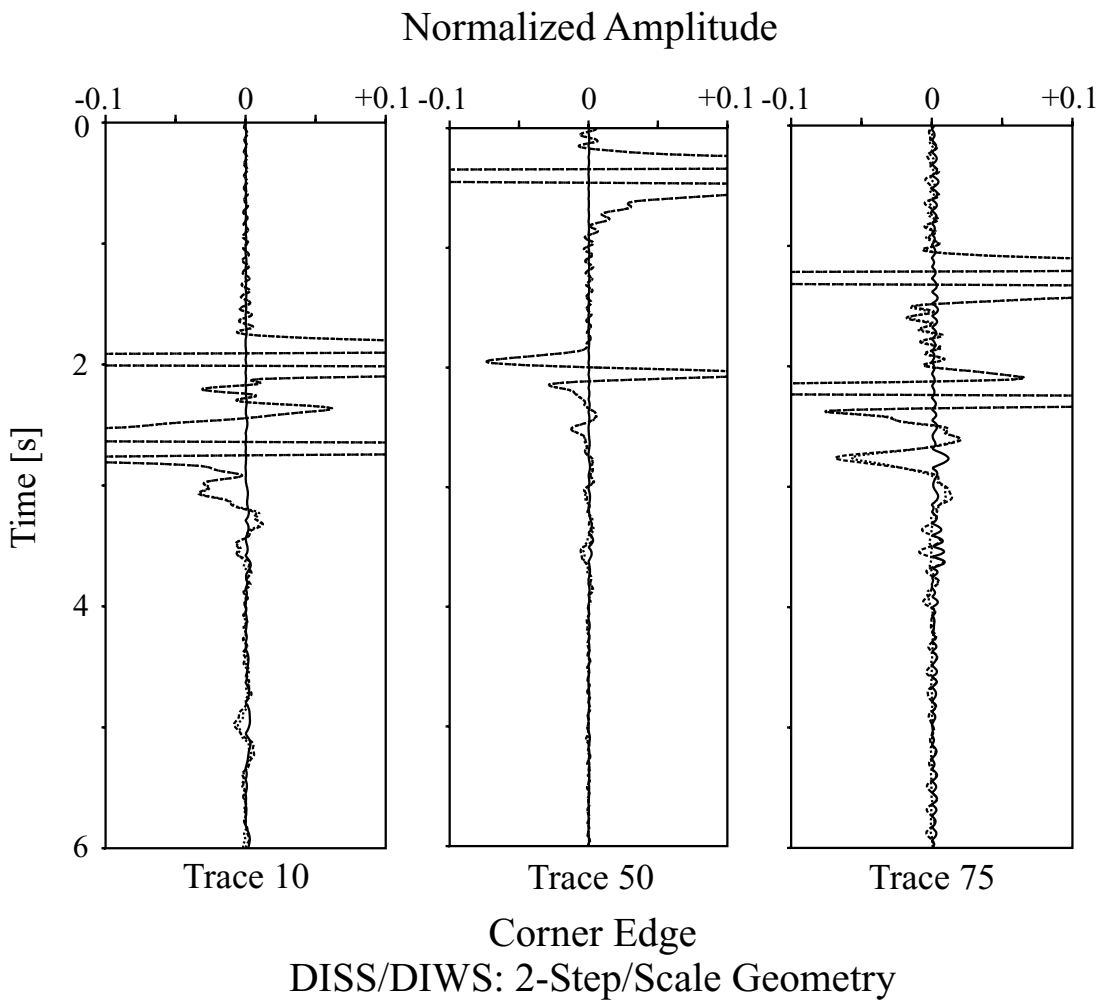

Figure 13. Comparison of three traces that are extracted from the DISS and DIWS seismograms of the corner edge model. Fine-grid solutions are obtained after $n_{\rm it} = 1500$ iteration loops. The dotted line is the DISS, the dashed line the DIWS solution. The continuous line is the difference between DISS and DIWS solutions. In order to visualize any differences we limited the amplitude to 10 per cent of the maximum amplitude. Only on trace 75 can slight differences be observed, while the overall fit is practically perfect.

this initial stage. The DISS and DIWS CPU time for $n_{\rm iter}^{\rm max} = 1500$ iterative steps for a single-frequency simulation on the 200×200 fine grid is ranging from $t_{f\approx 1\,{\rm Hz}} \approx 11/10$ min to $t_{f\approx 10\,{\rm Hz}} \approx 12/11$ min, respectively.

For the DISS *two-step* and DIWS *two-scale* simulations, no significant difference in CPU time is observed, even though the number of non-zero coefficients in the DIWS matrix compared with the DISS matrix is significantly larger. Normally we would expect an increased number of matrix coefficients to be accompanied by an increase of CPU time to perform the fixed number of iterations. This is not confirmed by our simulations and gives a first indication of an accelerated convergence rate for the wavelet preconditioning.

In Fig. 13, we show the DISS and DIWS comparison of the fine-grid corner edge solution for three extracted traces (trace nos 10, 50 and 75). Note that the trace amplitude is cut-off at 10 per cent of the maximum amplitude. The ringing on the traces, for example, on trace 10 for $0 < t < 1.8$ s, is due to the wavefield solution being computed by an iterative solver in the frequency domain with a limited number of iteration steps. Increasing the number of iterations may suppress ringing completely. The overall fit for the first arrival and reflection is good. More interesting is the difference of the DISS and DIWS solutions that we observe on trace 75 for times $0 < t < 1$ s and $3 < t < 4$ s. The DISS solution (dashed line) shows stronger ringing compared with the DIWS solution (dotted line).

5.1.2 Simulation 2: four-step case

We performed a second wave simulation for the corner edge model using the same model, source and receiver geometry as before. In contrast to the first simulation, we now compute wavefield solutions on a coarse grid that has half the size of the previous one. The prolongation of the DISS and DIWS coarse-grid frequency solutions is performed on the same fine grid used for the two-step scenario (200×200). This is similar to a two-grid situation, where the spatial discretization of the coarse grid is divided by a factor of 4 for the fine-grid simulation. For the DIWS scheme, this implies a wavelet decomposition on $J = 2$ resolution scales. The grid combination set-up is

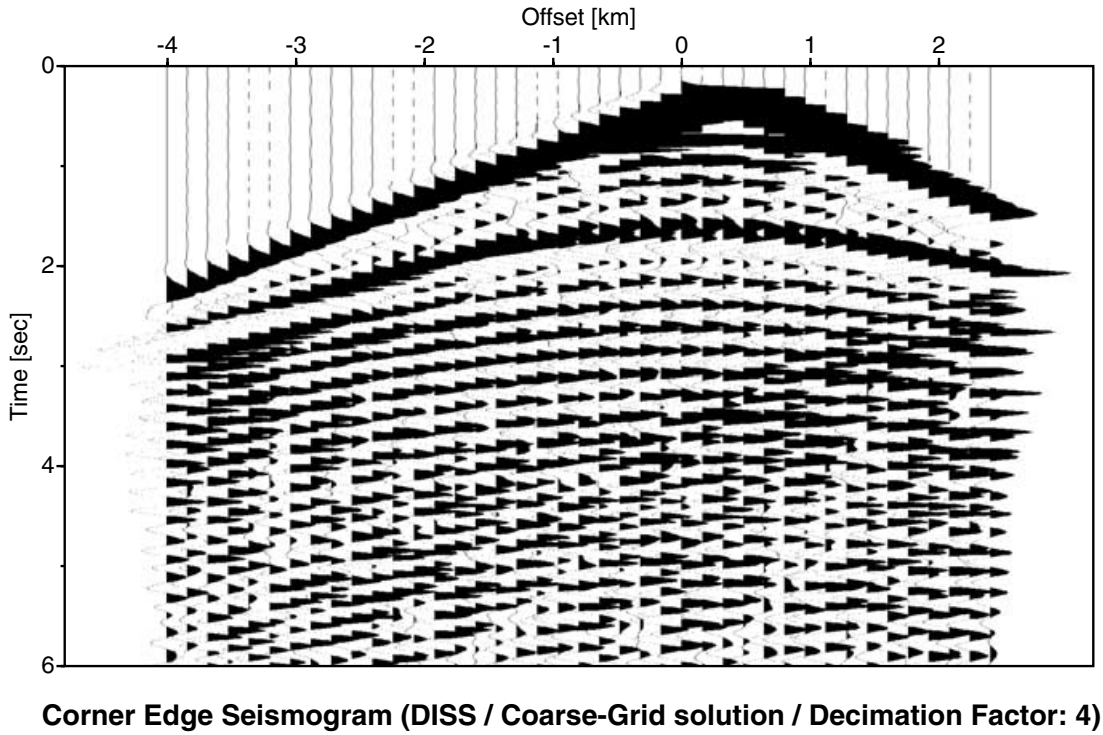

Figure 14. Seismogram for the corner edge model computed from the DISS coarse-grid solution. The spatial discretization step was set to $\Delta x = \Delta z =$ 160 m, which corresponds to 1.25 gridpoints per shortest wavelength. Because of the large undersampling of wavelength components strong wave dispersion dominates the seismogram.

therefore denoted by *four-step/three-scales* geometry. The coarse grid has a size of 50 × 50 grid nodes, of which five nodes on each side of the model are PML absorbing layers. The spatial discretization on the coarse grid is $dx = dz = 160$ m, which corresponds to 1.25 gridpoints per shortest wavelength. Applying the decimation factor of 4 then gives a spatial discretization for the fine grid of $dx = dz =$ 40 m.

In Fig. 14, we show the seismogram obtained by the coarse-grid wave simulation. Because of the large spatial discretization step with respect to the highest frequency modelled, we observe strong wave dispersion that significantly pollutes wave arrivals in the seismogram. As for the two-step case, the DISS scheme required $t_f \approx 11/12$ min CPU time to perform the fixed number of iteration steps ($n_{\text{iter}}^{\max} = 1500$), since despite the modified phase error of the initial iterative solution, the physical dimensions of the fine grid remained unchanged (the number of non-zero coefficients in the DISS fine-grid matrix is similar for *two-step* and *four-step* geometry). In contrast, the CPU time necessary to construct the DIWS fine-grid matrix increases to 80 min, due to an increased number of non-zero coefficients in the wavelet decomposition on three scales. Moreover, this increase automatically leads to less efficient matrix–vector and dot-products during the iteration steps, i.e. $t_f \approx 20/21$ min CPU time for $n_{\text{iter}}^{\max} = 1500$ DIWS iteration steps.

In Fig. 15(a), we plot the comparison of the DISS and DIWS fine-grid solutions for the *four-step* geometry. The DISS solution significantly improved when compared with the previous *two-step/two-scales* simulation example, and now fits much better with the DIWS solution. This is due to the DISS solution being less ringing than for the *two-step* case. The improvement is related to the DISS low-frequency components being better represented by the interpolated coarse-grid solution with respect to the convergence of the GMRES iterative solver. This behaviour is not surprising, since standard multi-grid V- and W-cycles descend from a fine to sets of coarser grids in order to obtain favourable iterative convergence conditions for the fine-grid low-frequency components.

In order to better quantify differences between the *two-* and *four-step* simulation examples, we plot separately the direct comparison of the two DISS and DIWS solutions in Figs 15(b) and (c), respectively. Moreover, in both Figs 15(b) and (c), we add the approximate *exact* solution that is obtained by the DISS scheme using $n_{\text{it}} = 5000$ iteration steps. As before, we observe an overall good agreement for both DISS and DIWS comparisons for traces 10 and 50. In contrast, trace 75, which is sensitive to the strong reflection from the horizontal corner edge interface, shows remarkable differences for the DISS solutions. We find the *four-step* DISS solution (dotted line) significantly less ringing than the *two-step* solution (dashed line), which is related to the better representation of fine-grid low-frequency components on the four-step rather than the two-step coarse grid. Nevertheless, increasing the number of DISS iteration steps (up to $n_{\text{it}} = 5000$) provides a *ringing-free* seismogram (the continuous line solution in Fig. 15b), but also increases significantly the CPU time from $t_{1500} \approx 10$ min to $t_{5000} \approx 20$ min.

The DIWS wave simulations for the *two-/three-scale* setups do not suffer from similar artefacts. Both simulation results fit well and practically no differences in amplitude are observed. Note that wiggles on trace 10 at $t \approx 5$ s and trace 75 at $t \approx 3.5$ s are artificial noise due to the model–PML reflection. The multi-scale formulation effectively combines numerical grids of varying resolution where grid interactions are

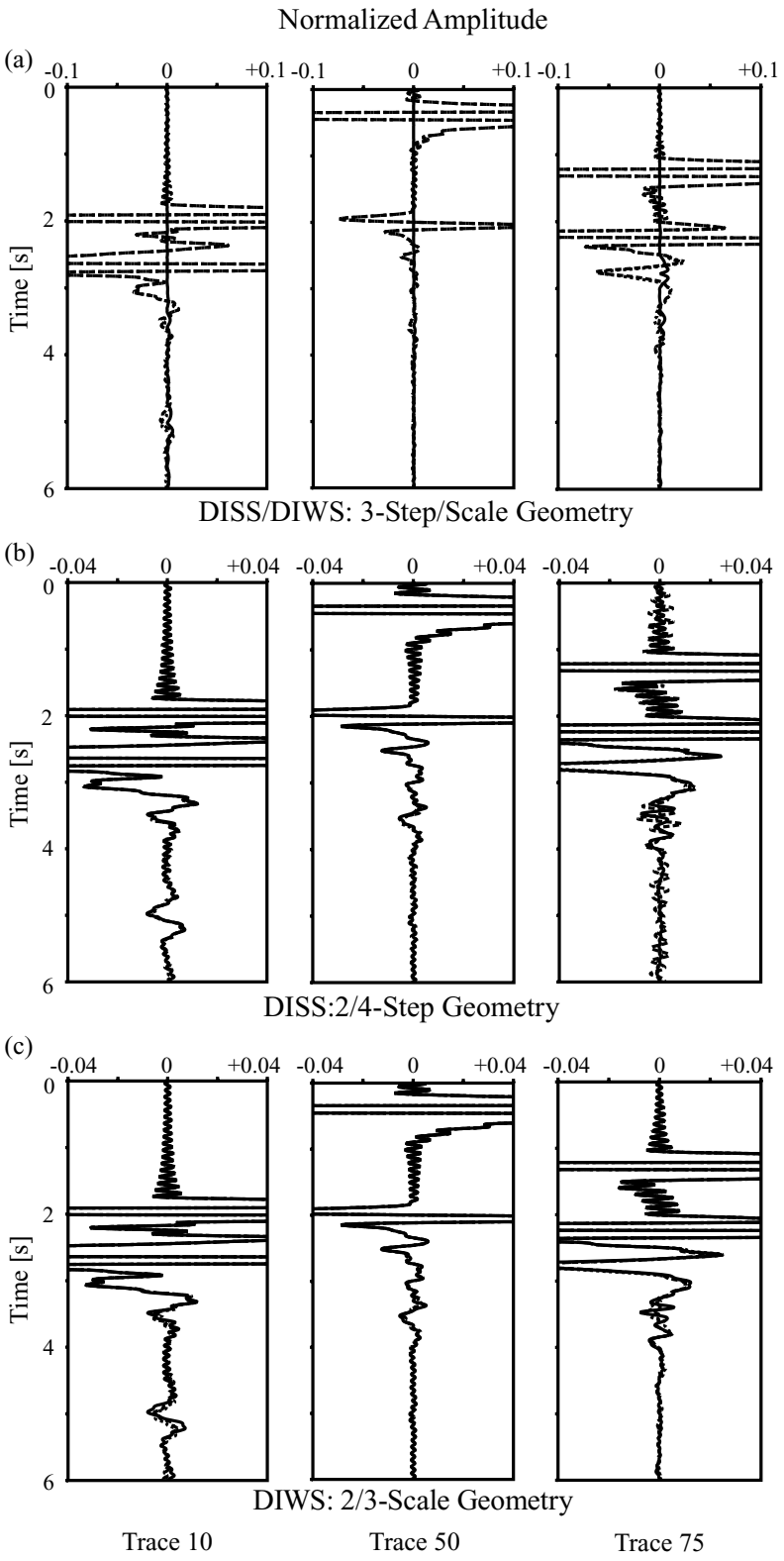

Figure 15. Comparison of traces 10, 50 and 75 for different wavefield simulations of the corner edge model. Part (a) shows the DISS and DIWS modelling results for the *four-step* (4ST) geometry after $n_{\mathrm{it}} = 1500$ iteration loops. The dotted line is the DISS solution and the dashed line the DIWS solution. The continuous line is the difference between the two solutions. The amplitude is limited to 10 per cent of the maximum amplitude in order to visualize the small differences between the two solutions. Part (b) shows the DISS wavefields for the *two-step* (2ST) and 4ST simulations. The dotted line is the fine-grid solution for the 2ST simulation and the dashed line for the 4ST simulation. The continuous line is the DISS solution after $n_{\mathrm{it}} = 5000$, which serves as the exact final solution. Only a slight differences between the two DISS seismograms can be observed on trace 75. Part (c) gives the simulation results for the DIWS *two-scales* (2SC) and *three-scales* (3SC) runs. The dotted and dashed lines denote fine-grid wavefield solutions for the 2SC and 3SC simulation, respectively. As in (b), the continuous line is the DISS solution after $n_{\mathrm{it}} = 5000$.

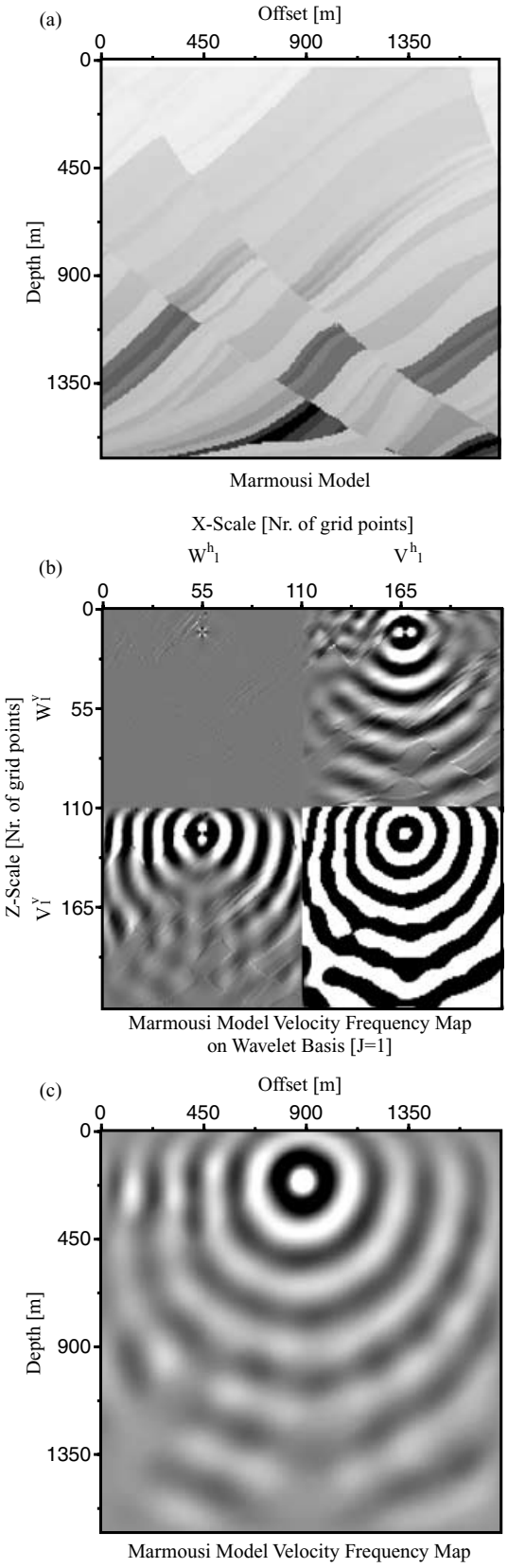

Figure 16. (a) Close-up of the Marmousi model. (b) Frequency map ($f \approx 9.5$ Hz) computed by the DIWS scheme for the windowed Marmousi model. We show the frequency distribution of the wavefield velocity component ($v_y = v_y^{(x)} + v_y^{(z)}$) on the wavelet basis with $J = 1$ resolution scales. Note that we only plotted the real part of complex valued wavefield. The gain is significantly increased in order to enhance the interaction coefficients that appear on the boundaries of different layers in the model with varying physical parameters. Application of the inverse wavelet transform gives the velocity wavefield component in the physical domain in (c).

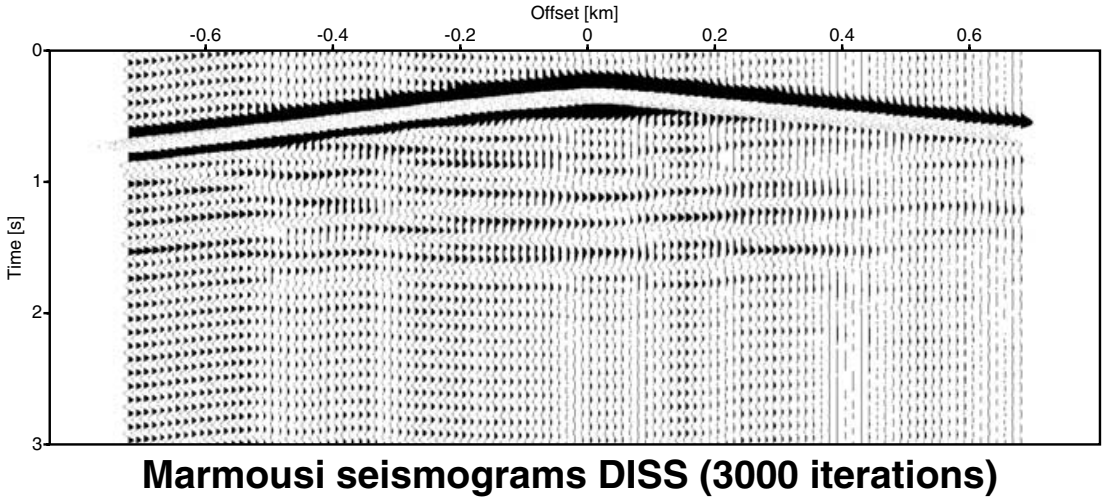

Figure 17. Time seismogram for the Marmousi model. The coarse-grid solution is extrapolated from 105×110 nodes ($dx = dz = 16$ m) to the fine-grid 210×220 ($dx = dz = 8$ m) by a bilinear interpolation. The seismogram is computed after $n_{\text{it}} = 3000$ iteration steps in the frequency domain.

accounted for automatically. Therefore, the performance of DIWS iterations on the wavelet coefficients is significantly less dependent on the number of iteration steps performed, since DIWS is comparable to a full multi-grid method where resolutions grids are treated simultaneously.

5.2 The Marmousi model

We perform a second wave simulation example for a windowed part of the Marmousi model (Bourgeois *et al.* 1991). This case study illustrates the iterative performance of the DIS approach for a wave simulation in a structurally complex heterogeneous media. DISS and DIWS wavefield simulations have been performed for a *two-step* situation. We applied PML absorbing boundary conditions on all four edges of the model (10, 20 grid nodes of the coarse grid and fine grid). The velocity in the model ranges from 1500 up to 4500 m s^{-1} (see Fig. 16a) and the density from 1000 to 2500 kg m^{-3}. We use a spatial grid interval of $dx = dz = 16$ and 8 m for the coarse- and the fine-grid simulations, respectively. This corresponds to approximately 10 and 20 grid-nodes per shortest wavelength, respectively. The source, located at ($x_s = 720$ m, $z_s = 50$ m), is the derivative of a Gaussian having similar properties to the one used for the corner edge simulations.

Wavefield solutions are computed for 52 frequency components distributed over a range of frequencies from $f_{\min} \approx 0$ Hz to $f_{\max} = 10$ Hz. Receivers are spread out over the whole model at a depth of 100 m and spacing of 12.5 m. An example of a frequency map ($f \approx 9.5$ Hz) for the v_y-velocity component ($v_y = v_y^{(x)} + v_y^{(z)}$) is shown in Fig. 16(b). In Fig. 16(b), we plot the wavefield solution after $n_{\text{it}} = 3000$ fixed *GMRES* iteration steps on $J = 1$ resolution scales. In the bottom right-hand block of Fig. 16(b), we see the wavefield projection on the $V_1 \times V_1$ scaling approximation space. The upper left-hand block is the fine-grid contribution to the v_y wavefield ($W_1 \times W_1$) with respect to the initial $V_0 \times V_0$ wavefield approximation. Note the high-frequency contribution on the fine wavelet scale around the source location.

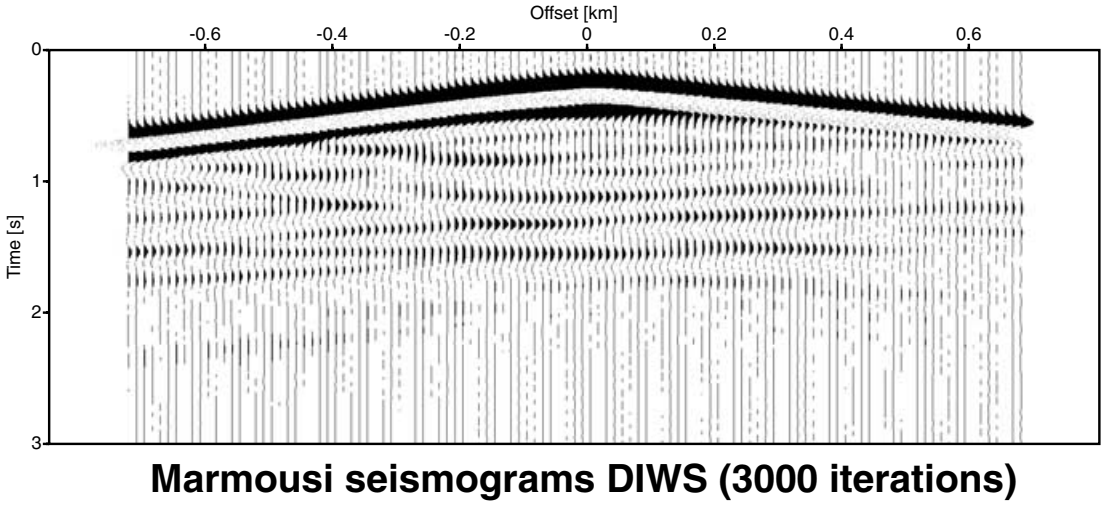

Figure 18. Time seismogram for the Marmousi model computed by the DIWS scheme. We use $J = 1$ wavelet resolution spaces (coarse grid, $105 \times 110/dx = dz = 16$ m; fine grid, $210 \times 220/dx = dz = 8$ m). The maximum number of iteration steps is set to 3000.

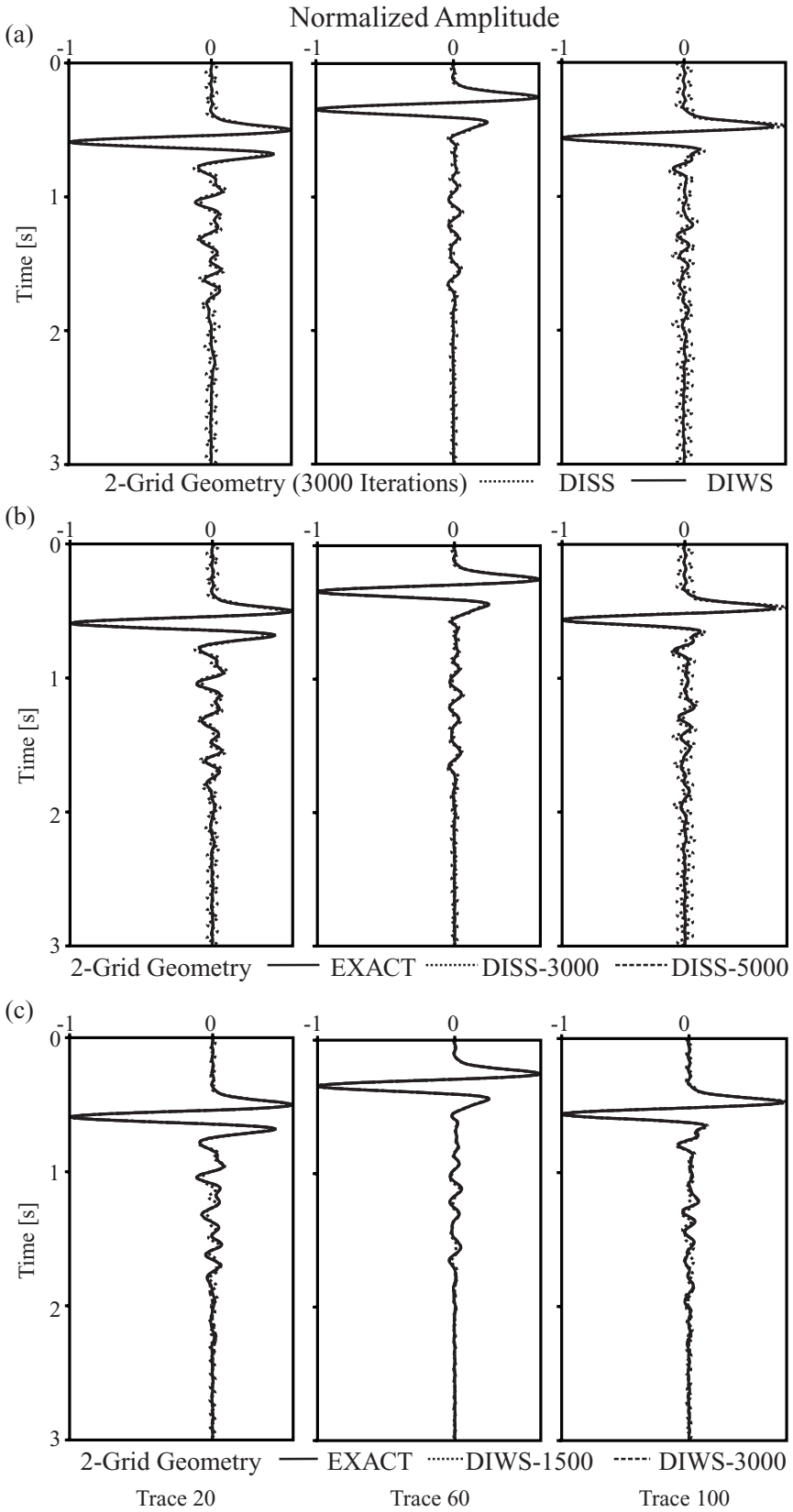

Figure 19. Comparison of traces 20, 60 and 100 for three different wavefield simulations in the Marmousi model. (a) The dotted and continuous lines are the DISS and DIWS fine-grid solution after $n_{it} = 3000$ iteration steps, respectively. (b) The DISS seismograms after $n_{it} = 3000$ and 5000 iteration steps. The continuous line is the exact wavefield solution, the dotted and dashed lines are the DISS solutions for $n_{it} = 3000$ and 5000 iterations, respectively. (c) The DIWS seismograms after $n_{it} = 1500$ and 3000 iteration steps. The continuous line is the exact wavefield solution, the dotted and dashed lines are the DIWS solutions for $n_{it} = 1500$ and 3000 iterations, respectively.

The other two blocks illustrate the coupling between the two wavefield approximation spaces. Fig. 16(c) is the same wavefield inverse wavelet transformed into the physical space.

Wavefield simulations are performed for different grid-model dimensions and number of fixed *GMRES* iterations (*GMRES* tolerance: $\mathrm{tol} = 1 \times \mathrm{e}^{-7}$). We find the number of DISS iteration steps required to provide dispersion-free wave simulations strongly increased. In contrast, the DIWS simulations show stable iteration rates that are independent of grid size and the number of resolution grids used. For example, we perform a DISS and DIWS wave simulation using a coarse grid and fine grid of size 105×110 and 210×220, respectively. The DISS and DIWS seismograms for $n_{\mathrm{it}} = 3000$ iterations are plotted in Figs 17 and 18, respectively. The DISS seismogram contains strong vertically vibrating noise that hides interesting reflection events. In contrast, the DIWS simulation provides a dispersion-free seismogram. We extract three traces from each of the seismograms and plot the direct DISS–DIWS comparison in Fig. 19(a). On all traces, we observe ringing of the DISS solution. Far offset traces, with respect to the source position, show stronger ringing artefacts.

We increase the maximum number of iterations from $n_{\mathrm{it}} = 3000$ to 5000 and recompute DISS iterative wavefield solutions. Moreover, we compute the exact seismogram using the direct solver for the fine-grid model setup. In Fig. 19(b), we show the comparison of three traces obtained by the exact direct solver and the DISS after $n_{\mathrm{it}} = 3000$ and 5000 iteration steps. The increase of iteration steps significantly reduces the previously detected ringing, and the DISS solution fits pretty well with the exact wavefield solution.

As for the DISS, we compare extracted traces for the DIWS simulation with the exact Marmousi seismogram. In Fig. 19(c), we plot the exact solution and two DIWS iterative solution approximations after $n_{\mathrm{it}} = 1500$ and 3000 iteration steps. Even though the DIWS iterative wavefield solution after $n_{\mathrm{it}} = 1500$ still has deficiencies, at this *early* stage (in terms of the number of iterations) it already reproduces the main structure of the final seismogram. Note that we call this effect *homogenization*. After $n_{\mathrm{it}} = 3000$ iteration steps the DIWS provides a seismogram that fits the exact solution very well.

We compare the CPU time needed by the DISS and DIWS scheme for a fixed numbers of iteration steps and a given frequency component. Computation times for different numbers of DISS and DIWS iteration steps are given in Table 1. For a fixed number of iterations $n_{\mathrm{it}} = 1500$, the DISS requires $\approx$40 per cent less computation time than the DIWS, though the obtained DISS wavefield solution is strongly distorted by noise. In order to obtain a qualitatively dispersion-free wavefield solution, the DISS scheme needs $\approx$30 per cent more computation time than the DIWS scheme. The increase in iterative computation time is significant if one has to perform iterations for many right-hand sides.

The CPU times indicated in Table 1 are mainly provided to differentiate the performance of the DISS and the DIWS approach with respect to a wavelet-based preconditioning. These CPU times are not revealing in absolute terms because no further numerical preconditioning strategies have been addressed at this stage. A significant improvement of the convergence rate is expected once a suitable preconditioning has been implemented.

For information, an FDTD simulation in the windowed Marmousi model takes only 4.5 min. At present, the DISS and DIWS algorithms require on average 15 min per frequency component (this latter time must be multiplied by the number of frequencies to have the CPU time required to generate time-domain seismograms). Then, the computational cost of FDTD simulation is much lower than that of the DISS and DIWS simulations because of the large number of iterations required by the non-preconditioned iterative solver (see Table 1).

We would also like to remind the reader that time-domain and frequency-domain methods rely on quite different numerical approaches. Our main motivation associated with frequency-domain approaches is to develop algorithms suited for multi-source experiments.

Time-domain methods require complete recalculations from the start when a new source position is considered. The computational cost of time-domain modelling in the case of multiple sources is equal to the CPU time required to compute a single-source simulation multiplied by the number of sources. Therefore, time-domain methods are poorly adapted to this configuration.

In the DISS and DIWS frequency-domain approaches, a direct method (**LU** factorization followed by forward and backward substitutions) is first applied on a coarse grid. When they can be used, direct methods are the method of choice for multiple sources since the most expensive operation (**LU** factorization) needs to be computed only once. One can obtain efficiently an approximate solution for a new source position by simple forward and backward substitutions. Therefore, the DISS and DIWS algorithms mainly take advantage of the multi-source configuration during the construction of approximate solutions that are subsequently used as initial guesses by the iterative solver. Accurate starting solutions in addition to efficient preconditioning allows the minimization of the number of iterations for the iterative solver, which does not take advantage of the source multiplicity directly.

Since up to now, we have used a non-preconditioned GMRES solver, it is presently not possible to rank FDTD methods with respect to the DISS and the DIWS methods in the case of multi-source experiments. Moreover, the relative efficiency of the two classes of methods

Table 1. Computation times for the complex Marmousi Model for the DISS and DIWS schemes. DISS wavefield solutions have been computed for four increasing number of iteration steps ($n_{\mathrm{it}}^{\max} = 1500, 2500, 3500, 5000$) while the DIWS simulation was performed using $n_{\mathrm{it}}^{\max} = 1500$ iteration steps only. Note that the CPU time is provided for one single-frequency component. This CPU time must be roughly multiplied by the number of frequencies to obtain the CPU time required to compute time-domain seismograms.

Solver	No of iterations	CPU time/frequency (min)
DISS	1500	6
DISS	2500	12
DISS	3500	17
DISS	5000	23
DIWS	1500	16

may be rather case dependent: among others, it will depend on the number of sources and the size of the problem (i.e. the dimension of the impedance matrix). The size of the problem will control the ratio between the coarse and the fine-grid sizes used by the direct and iterative parts, respectively. This ratio will probably have an influence on the accuracy of the starting solutions used by the iterative solver.

We also note that one of the most promising applications of frequency-domain modelling is frequency- domain waveform inversions in the frame of seismic tomography applications (Pratt & Shipp 1999; Ravaut 2003). In frequency-domain wavefield inversion, only a limited number of selected frequency components need to be inverted if the acquisition geometry contains sufficiently wide apertures (Pratt & Worthington 1990). In that case, there is a strong redundancy in the data, which can be partly decimated. Therefore, only these selected frequency components need to be modelled to solve the inverse problem. Time-domain solutions (which compute the solution at all frequencies) may not be the best adapted modelling tool for this purpose.

5.3 Discussion

The DISS and DIWS are tested for two complex modelling scenarios. Wavefield simulations in the corner edge model are performed for a *two-step* and a *four-step* setup, where spatial discretization steps change by a factor of 2 and 4 between the coarse and the fine grid. The second wave propagation simulation for the Marmousi model is performed using a *two-step* combination.

The spatial DISS approach permits fast and efficient matrix constructions. In general, the computation time related to the DIWS matrix construction is greatly increased, because of convolution-type projections in the wavelet space. Although expensive, the convolution calculations are carried out only once before the modelling starts.

The DISS matrix is sparse and therefore provides efficient computations of *GMRES* iteration steps. In contrast, the number of non-zero coefficients in the DIWS matrix increases with respect to the number of wavelet resolution scales used. Therefore, iterative matrix–vector and dot-product computations are slowed down, when compared with similar products in the physical space. We did not address any space adaptivity implementation at this stage of the development, although the wavelet formulation naturally supports the definition of spatial masks that significantly increase the computational performance.

Nevertheless, the multi-scale formulation through wavelet-based iterations provides a full multi-grid environment, where intermediate iterative solutions on all resolution scales are directed versus the correct final solution in a desirable automatic manner. Already a relatively small number of iteration steps provide the essential features of the final solution. With increasing iterations the wavelet and scaling coefficients are modified to finally construct the fine-grid wavefield solution.

The iteration on the wavelet basis is favourable, because interactions of the fine-grid and coarse-grid approximations are automatically accounted for. Even though the DIWS matrix construction is expensive, it is only performed once. More important is the reduction of necessary iteration steps required to compute a dispersion-free wavefield solution, if one simulates large-scale multi-source experiments. The required CPU time may be significantly decreased in the future through the invention of space adaptivity and optimized parallel wavelet transform algorithms.

In contrast, the DISS matrix construction is performed at practically no computational cost. The matrix sparsity is guaranteed by the staggered finite-difference stencil applied in the physical domain. Therefore, fast and efficient matrix–vector and dot-products may be implemented. The main drawback of the DISS approach is related to the simple-grid formulation that can only treat coarse- or fine-grid errors at a time. Dispersion-free wave simulations are strongly dependent on the number of iteration steps performed. Even though phase-corrected frequency maps may be obtained when large numbers of iteration steps are performed, the ambiguity for choosing the right number of iterations is rather cumbersome.

One possibility for introducing more variability in the DISS scheme to suppress interpolation-related phase shifts and the convergence of low-frequency components on the fine grid, is the extension of the current DISS *nested iteration* approach to a (space) FMG method through loops over sequences of V- and W-cycles.

The role of wavelet resolution scale interactions is illustrated in Figs 20(a) and (b) for a single-frequency map $f = 9.5$ Hz of the Marmousi simulation. In Fig. 20(a) from left to right, we show the coarse-grid solution recast as an initial iterative solution on $J = 1$ resolution scales, the result from the iteration process, and the difference between the initial and the final solution on the wavelet basis. Fig. 20(b) is the same sequence, now after inverse wavelet transformation in the physical space.

It is interesting to note that the DIWS modified the coarse-grid solution in the $V_1 \times V_1$ space. This indicates the necessity of coarse-grid corrections in the iterative scheme to suppress fine-scale low-frequency components that otherwise may cause severe difficulties in the iteration process. Even though fine-grid wavelet and interaction contributions are less strong, they still influence the iteration process at each iteration step. The wavelet multi-scale formulation naturally incorporates these effects, while the DISS nested iteration requires excessive CPU time or *manual* corrections through V- and W-cycles that are difficult to perform.

6 CONCLUSION AND PERSPECTIVES

We perform wave propagation simulations in structurally complex media with strong velocity contrasts. Wavefield solutions are computed using the DISS and DIWS schemes for a set of frequency components. The space- and wavelet-based solvers are compared with respect to CPU time performance and accuracy of the solution.

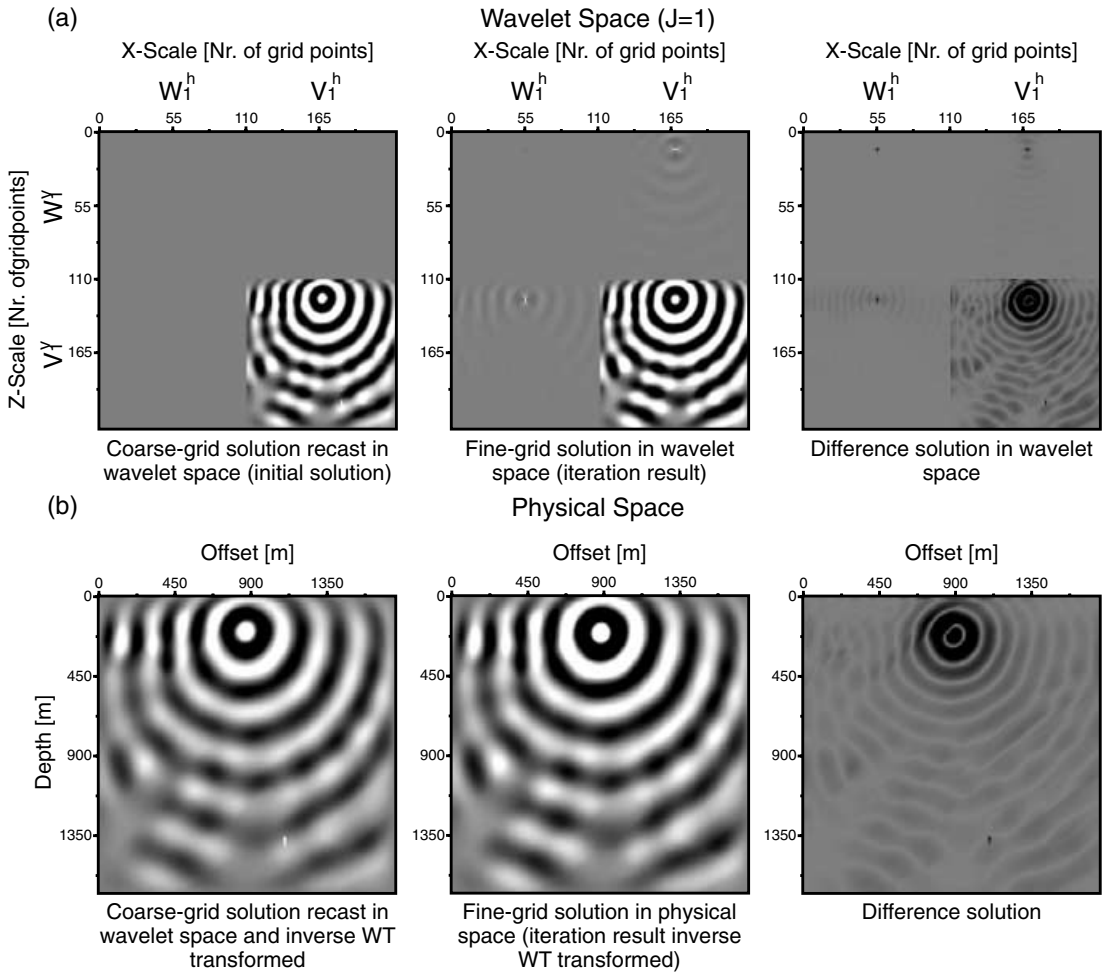

Figure 20. Illustration of grid interactions occurring during the iteration process on a wavelet expansion for the DIWS scheme. The example is shown for a frequency map ($f = 9.5$ Hz) computed by the DIWS on $J = 1$ resolution scales for the Marmousi model. In (a) we plot the coarse-grid wavefield solution recast in the wavelet expansion, where all wavelet coefficients are initially zero (left-hand panel). The middle panel shows the output of the *GMRES* iterative solver after $n_{\mathrm{it}} = 3000$ iteration steps. The right-hand panel is the difference of the initial and final wavefield solution. Part (b) is similar to (a), though now the velocity wavefield solutions have been inverse transformed in the physical domain. Interesting to note is that the iterative solver modified the $V_1 \times V_1$ space in the wavelet expansion (see the right-hand panel in (a)). This signifies that the fine-grid wavelet coefficients interact and alter the coarse-grid representation of the wavefield solution.

The combination of a direct with an iterative solver allows wavefield simulations in synthetic models of larger size, compared with applications of a single direct solver. The performance of the DIS approach depends mainly on the formulation of the iterative solution scheme.

We found that space linear grid combinations lead to fast and efficient matrix constructions, matrix–vector and dot-product algorithms. The DISS approach is similar to the nested iteration part of a standard multi-grid method. The phase shifts introduced by simple bilinear interpolation require large numbers of iteration steps in order to provide dispersion-free wave simulation results.

The second DIS approach uses the multi-scale formulation of a standard orthogonal Daubechies-4 wavelet transform. The analytic construction of the principal matrix terms in the spectral wavelet space, though performed only once before the modelling starts, requires the calculation of expensive convolution-type projection algorithms and leads to less efficient matrix computations than in the physical space.

The main advantage of the wavelet-based preconditioning is the simultaneous iteration on all wavelet resolution grids, where the interactions between scales are accounted for automatically. In a way, the DIWS scheme therefore performs FMG V-cycles at each iteration step. The wavelet iteration was found to be largely independent of the number of iterations steps performed. Already a small number of iteration steps provided quantitative dispersion-free seismograms where principal wave phenomena may be distinguished.

We illustrated the accuracy and performance of DISS and DIWS approaches for the corner edge model and the Marmousi model with respect to varying grid sizes and the number of iteration steps.

The numerical tests clearly indicate superior iteration behaviour for the DIWS multi-scale formulation rather than for the DISS nested iteration approach. The combination of the wavelet multi-grid construction with an iterative solver eliminates the ambiguity of manual grid combinations of standard nested iteration techniques.

The computational cost for matrix constructions in the spectral wavelet space may be reduced when space adaptivity strategies are implemented. Moreover, a significant speed-up is obtained by the extension of the DISS and DIWS sequential codes to run on parallel computer structures.

Nevertheless, large 2-D and 3-D wave propagation simulations require significantly increased memory and computational power, such that the transition of the existing sequential code to a parallel version seems unavoidable. Although it was not considered in this study, standard numerical preconditioning is expected to provide another important speed-up for the iterative scheme.

Altogether, the DIWS approach provides a powerful tool for investigating wave propagation phenomena for discrete frequency components or complete sets of frequencies for multiple right-hand sides in complex heterogeneous media. Typical applications are site effect studies, where the response to single frequencies is of great interest. Moreover, wavefield simulations of discrete frequency components provide powerful information for full waveform inversion schemes. Other applications include standard forward modelling of large-scale marine multi-source experiments for broad frequency bands.

The next development step of the DIS strategy includes the iterative convergence for a fixed tolerance level that was not addressed here. Furthermore, the numerical code will be parallelized together with an extension to 3-D modelling for elastic wave propagation.

ACKNOWLEDGMENTS

We are very grateful to Patrick Amestoy (ENSEEIHT-IRIT), Luc Giraud (CERFACS) and Jean-Yves Excellent (UMR CNRS-ENS Lyon-INRIA) for providing us with the multi-frontal matrix ordering software MUMPS, the iterative solver GMRES and some fruitful discussions regarding the direct-iterative strategy for wave simulations. We thank Stéphane Lanteri and INRIA Sophia Antipolis for the access to their PC-Cluster environment. We further acknowledge the usage of the IDRIS PC-Cluster. The research leading to this paper was partly founded by DASE/CEA through the PROSIS3D group. This paper is contribution no 612 from UMR Geosciences Azur, CNRS, France.

REFERENCES

Amestoy, P.R., Duff, I.S., Koster, J. & L'Excellent, J.-Y., 2001. A fully asynchronous multifrontal solver using distributed dynamic scheduling, *SIAM J. Matrix Anal. Applic.,* **23,** 15–41.

Berenger, J.P., 1994. A perfectly matched layer for the absorption of electromagnetic waves, *J. Comput. Phys.,* **114,** 185–200.

Beylkin, G., 1998. On multiresolution methods in numerical analysis, *Docum. Math.,* Extra Volume ICM III, 481–490.

Beylkin, G., Coifman, R. & Rokhlin, V., 1991. Fast wavelet transforms and numerical algorithms, *Commun. Pure Appl. Math.,* **44,** 141–183.

Bourgeois, A., Bourget, M., Lailly, P., Poulet, M., Ricarte, P. & Versteeg, R., 1991. Marmousi, model and data, in, *The Marmousi Experience,* pp. 5–16, Eur. Ass. Expl. Geophys.

Briggs, W.L., 1987. *A Multigrid Tutorial,* SIAM, Philadelphia.

Briggs, W.L. & Henson, V.E., 1993. Wavelets and multigrid, *SIAM J. Comput.,* **14,** 506–510.

Cervený, V., Molotkov, I.A. & Psencik, I., 1977. *Ray theory in Seismology,* Charles University Press, Praha.

Cohen, G. & Fauqueux, S., 2000. Mixed finite elements with mass-lumping for the transient wave equation, *J. Comput. Acoust.,* **8,** 171–188.

Daubechies, I., 1992. *Ten Lectures on Wavelets,* CBMS61, SIAM, Philadelphia.

Faccioli, E., Maggio, F., Quarteroni, A. & Tagliani, A., 1996. Spectral-domain deomposition methods for the solution of acoustic and elastic wave equations, *Geophysics,* **61,** 1160–1174.

Falk, J., Tessmer, E. & Gajewski, D., 1998. Efficient finite-difference modelling of seismic waves using locally adjustable time steps, *Geophys. Prospect.,* **46,** 603–616.

Frayssé, V., Giraud, L. & Gratton, S., 1997. A set of GMRES routines for real and complex arithmetics, *CERFACS Technical Report TR/PA/97/49*.

Fuchs, K. & Müller, K., 1971. Computations of synthetic seismograms with the reflectivity method and comparison with observations, *Geophys. J. R. astr. Soc.,* **23,** 417–433.

Golub, G.H. & van Loan, C.F., 1996. *Matrix Computations,* 3rd edn, John Hopkins University Press, Baltimore.

Graves, R.W., 1996. Simulating seismic wave propagation in 3-D elastic media using staggered-grid finite differences, *Bull. seism. Soc. Am.,* **86,** 1091–1106.

Hackbusch, W., 1978. On the multi-grid method applied to difference equations, *Computing,* **20,** 291–306.

Hestholm, S. & Ruud, B., 1998. 3-D finite-difference elastic wave modeling including surface topography, *Geophysics*, **63,** 613–622.

Hustedt, B., Operto, S. & Virieux, J., 2003. Mixed-grid and staggered-grid finite difference methods for frequency-domain acoustic wave modelling, *Geophys. J. Int.,* submitted

Igel, H., 1999. Wave propagation in three-dimensional spherical sections by the Chebyshev spectral method, *Geophys. J. Int.,* **136,** 559–566.

Jameson, L., 1993. On the wavelet based differentiation matrix, *J. Sci. Comput.,* **8,** 267–305.

Janod, F., 1999. Propagation des ondes large bande dans des milieux élastiques 3D: méthodes d'eléments de frontière et couplage avec les différences finies, PhD thesis, Université Joseph Fourier, Grenoble I.

Jastram, C. & Behle, A., 1992. Acoustic modelling on a grid of vertically varying spacing, *Geophys. Prospect.,* **40,** 157–169.

Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Int.,* **139,** 806–822.

Komatitsch, D. & Vilotte, J.P., 1998. The spectral element method: an efficient tool to simulate the response of 2-D and 3-D geological structures, *Bull. seism. Soc. Am.,* **88,** 368–392.

Levander, A.R., 1988. Fourth-order finite-difference, *P–SV* seismograms, *Geophysics,* **53,** 1425–1436.

LeVeque, R.J., 1997. Wave propagation algorithms for multidimensional hyperbolic systems, *J. Comput. Phys.,* **131,** 327–353.

Mallat, S., 1989. Multiresolution approximation and wavelets, *Trans. Am. Math. Soc.,* **315,** 69–88.

Message Passing Interface Forum, M., 1994. A message-passing interface standard, *Int. J. Supercomp. Appl. High Perform. Comput.,* **8,** 3/4.

Moczo, P., 1980. Finite difference technique for *SH* waves in 2-D media using irregular grids, *Geophys. J. R. astr. Soc.,* **99,** 321–329.

Moczo, P., Lucká, M., Kristek, J. & Kristeková, M., 1999. 3-D displacement finite differences and a combined memory optimization, *Bull. seism. Soc. Am.,* **89,** 69–79.

Mulder, W. & Plessix, R.-E., 2002. Time- versus frequency domain modelling of seismic wave propagation, in *Extended Abstracts of the 64th Annual EAEG Meeting,* p. E–15, Florence. Eur. Ass. of Geosc. and Eng.

Olsen, K.B. & Archuleta, R.J., 1996. Three-dimensional simulation of earthquakes on the Los Angeles fault system, *Bull. seism. Soc. Am.,* **86,** 575–596.

Olsen, K.B., Archuleta, R.J. & Matarese, J.R., 1995. Magnitude 7.75 earthquake on the San Andreas fault: three-dimensional ground motion in Los Angeles, *Science,* **270,** 1628–1632.

Operto, S., Virieux, J., Hustedt, B. & Malfanti, F., 2002. Adaptive wavelet-based finite-difference modelling of *SH*-wave propagation, *Geophys. J. Int.,* **148,** 1–28.

Pessel, M., 2000. Tomographie électrique: Dévelopements méthodologiques et application, *PhD thesis,* Université de Rennes 1.

Pitarka, A., 1999. 3-D elastic finite-difference modeling of seismic motion using staggered grids with nonuniform spacing, *Bull. seism. Soc. Am.,* **89,** 54–68.

Pratt, R.G., 1990. Frequency-domain elastic wave modeling by finite-differences: a tool for crosshole seismic imaging, *Geophysics,* **55,** 626–632.

Pratt, R.G. & Shipp, R.M., 1999. Seismic waveform inversion in the frequency domain, Part 2: Fault delineation in sediments crosshole data, *Geophysics,* **64,** 902–914.

Pratt, R.G. & Worthington, M.H., 1990. Inverse theory applied to multi-source cross-hole tomography. part 1: acoustic wave-equation method, *Geophys. Prospect.,* **38,** 287–310.

Press, W.H., Vettering, W.T., Teukolsky, S.A. & Flannery, B.P., 1994. *Numerical Recipies in Fortran,* 2nd edn, Cambridge University Press, Cambridge.

Ravaut, C., 2003. Tomographie sismique haute résolution de la croûte terrestre: inversion combinée des temps de trajet et des formes d'ondes de données de sismique réfraction-réflexion grand-angle multitrace, *PhD thesis,* Université Paris 6.

Robertsson, J.O.A. & Chapman, C.H., 2000. An efficient method for calculating finite-difference seismograms after model alterations, *Geophysics,* **65,** 907–918.

Sambridge, M., Braun, J. & McQueen, H., 1995. Geophysical parametrization and interpolation of irregular data using natural neighbours, *Geophys. J. Int.,* **122,** 837–857.

Sanchez-Sesma, F.J., 1983. Diffraction of elastic waves by three-dimensional surface irregularities, *Bull. seism. Soc. Am.,* **73,** 1621–1636.

Strang, G. & Nguyen, T., 1996. *Wavelet and Filter Banks,* Wellesley-Cambridge Press, Wellesley.

Tessmer, E., 2000. Seismic finite-difference modeling with spatially varying time steps, *Geophysics,* **65,** 1290–1293.

Tessmer, E. & Kosloff, D., 1994. 3-D elastic modeling with surface topography by a Chebychev spectral method, *Geophysics,* **59,** 464–473.

Virieux, J., 1984. *SH* wave propagation in heterogeneous media, velocity–stress finite difference method, *Geophysics,* **49,** 1259–1266.

Virieux, J., 1986. *P–SV* wave propagation in heterogeneous media, velocity–stress finite difference method, *Geophysics,* **51,** 889–901.

Štekl, I. & Pratt, R.G., 1998. Accurate viscoelastic modeling by frequency-domain finite differences using rotated operators, *Geophysics,* **63,** 1779–1794.

Wu, Y. & McMechan, G., 1998. Wave extrapolation in the spatial wavelet domain with application to postack reverse-time migration, *Geophysics,* **63,** 589–600.

Xie, Z., Chan, C.-H. & Zhang, B., 2002. An explicit fourth-order orthogonal curvilinear staggered-grid FDTD method for Maxwell's equations, *J. Comput. Phys.,* **175,** 739–763.

Zienkiewicz, O.C. & Morgan, K., 1982. *Finite Elements and Approximation,* 2nd edn, McGraw-Hill, London.

APPENDIX: MATRIX PROJECTION ON ORTHOGONAL WAVELET BASIS

A1 Projection of physical medium parameters (diagonal blocks)

We first derive eq. (10), that is the product of medium properties with wavefield components projected on an orthogonal wavelet basis. Note that wavefield projections on the wavelet basis in the x-direction are denoted by indices i, j, k, l and for the z-direction by α, β.

We start the wavelet transformation by expanding the expression $i\omega\xi(x, \omega)\rho(x, z)v_y^{(x)}(x, z)$ first on an orthogonal wavelet basis with respect to the x-direction:

$$\sum_i \left\langle i\omega\xi_x \rho(x, z)v_y^{(x)}(x, z), \psi_i \right\rangle \psi_i. \tag{A1}$$

We separate the wavefield component $v_y^{(x)}$ and the density term $i\omega\xi_x\rho(x, z)$ through permutation and subsequent application of a second wavelet expansion to the wavefield component $v_y^{(x)}$. The new basis, which is also a projection in the x-direction, is denoted by ψ_j. We obtain the double sum:

$$\sum_{i,j} \left\langle v_y^{(x)}(x, z), \psi_j \right\rangle \langle \psi_j, i\omega\xi_x\rho(x, z)\psi_i\rangle \psi_i. \tag{A2}$$

We are now ready to project eq. (A2) on a wavelet basis in the z-direction, denoted by ψ_α. This gives:

$$\sum_\alpha \sum_{i,j} \left\langle \left\langle v_y^{(x)}(x, z), \psi_j \right\rangle \langle \psi_j, i\omega\xi_x\rho(x, z)\psi_i\rangle, \psi_\alpha \right\rangle \psi_i\psi_\alpha. \tag{A3}$$

Again, we need to separate terms depending on z and therefore apply a second wavelet projection following the same procedure as before, but now in the z-direction (ψ_β):

$$\sum_{\alpha,\beta} \sum_{i,j} \left\langle v_y^{(x)}(x, z), \psi_j\psi_\beta \right\rangle \langle \psi_j\psi_\beta, i\omega\xi_x\rho(x, z)\psi_i\psi_\alpha\rangle \psi_i\psi_\alpha. \tag{A4}$$

The inner products in eq. (A4) form a term-to-term product that has to be recast to verify the matrix equation $A \cdot x = b$, where x is the desired solution constructed by the GMRES iterative solver. In order to compute the necessary terms, we first need to construct the wavelet basis functions ψ_i and ψ_α. For each pair (i, α) we compute the product of medium parameters $i\omega\rho(x, z)$ with the wavelet basis functions ψ_i and ψ_α and then expand the resulting term on the wavelet basis through subsequent projections in the x- and z-directions. The term-to-term product will then be verified through correct storage of the corresponding term $\langle \psi_j\psi_\beta, i\omega\xi_x\rho(x, z)\psi_i\psi_\alpha\rangle$ in the impedance matrix A and $\langle v_y^{(x)}(x, z), \psi_j\psi_\beta\rangle$ in the desired solution x to give a standard matrix–vector product.

A2 Projection of differential operators

We compute the projection of staggered grid derivative operators on an orthogonal wavelet basis. For simplicity, we show the development for the x-derivative operator only, since replacing x by z and following a similar procedure leads to the z-derivative operator.

We project the derivative operator on an initial basis ψ_i in the x-direction and obtain:

$$\sum_i \left\langle \frac{\partial}{\partial x} v_y^{(x)}, \psi_i \right\rangle \psi_i. \tag{A5}$$

As before, we need to separate contributions depending on the x-direction, namely the operator $\partial/\partial x$ and the wavefield $v_y^{(x)}(x, z)$. We go back to the definition of the inner product and apply an integration by parts, which gives

$$\begin{aligned}\sum_i \left\langle \frac{\partial}{\partial x} v_y^{(x)}, \psi_i \right\rangle \psi_i &= \int_{-\infty}^{+\infty} \frac{\partial v_y^{(x)}(x, z)}{\partial x} \cdot \psi(x)\, dx \\ &= \left[\psi(x) \cdot u_y^{(x)}(x, z)\right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \frac{\partial \psi(x)}{\partial x} \cdot v_y^{(x)}(x, z)\, dx. \end{aligned} \tag{A6}$$

Because of the local support of the Daubechies wavelets, the first term in eq. (A6) is zero. We therefore obtain:

$$\sum_i \left\langle \frac{\partial}{\partial x} v_y^{(x)}, \psi_i \right\rangle \psi_i = -\sum_i \left\langle \frac{\partial \psi_i}{\partial x}, v_y^{(x)}(x, z) \right\rangle \psi_i. \tag{A7}$$

We separate the remaining expression by projecting the derivative term $\frac{\partial}{\partial x}\psi_i$ on to a new wavelet basis ψ_j, and obtain:

$$-\sum_{i,j} \left\langle \frac{\partial}{\partial x} \psi_i, \psi_j \right\rangle \left\langle \psi_j, v_y^{(x)} \right\rangle \psi_i. \tag{A8}$$

Finally, we perform the wavelet expansion in the z-direction. Since the derivative operator depends on the x-component only, no further projections have to be performed and we obtain the final wavelet expansion of the x-derivative operator on a 2-D orthogonal wavelet basis:

$$-\sum_{\alpha} \sum_{i,j} \left\langle \frac{\partial}{\partial x} \psi_i, \psi_j \right\rangle \left\langle v_y^{(x)}, \psi_j \psi_\alpha \right\rangle \psi_i \psi_\alpha. \tag{A9}$$