

A 2-D spectral-element method for computing spherical-earth seismograms—II. Waves in solid–fluid media

Tarje Nissen-Meyer,¹ Alexandre Fournier² and F. A. Dahlen¹

¹Department of Geosciences, Princeton University, Princeton, NJ 08544, USA. E-mail: tarje@alumni.princeton.edu

²Laboratoire de Géophysique Interne et Tectonophysique, Université Joseph Fourier, CNRS, 38041 Grenoble Cedex 9, France

Accepted 2008 April 4. Received 2008 February 17; in original form 2007 October 23

SUMMARY

We portray a dedicated spectral-element method to solve the elastodynamic wave equation upon spherically symmetric earth models at the expense of a 2-D domain. Using this method, 3-D wavefields of arbitrary resolution may be computed to obtain Fréchet sensitivity kernels, especially for diffracted arrivals. The meshing process is presented for varying frequencies in terms of its efficiency as measured by the total number of elements, their spacing variations and stability criteria. We assess the mesh quantitatively by defining these numerical parameters in a general non-dimensionalized form such that comparisons to other grid-based methods are straightforward. Efficient-mesh generation for the PREM example and a minimum-messaging domain decomposition and parallelization strategy lay foundations for waveforms up to frequencies of 1 Hz on moderate PC clusters. The discretization of fluid, solid and respective boundary regions is similar to previous spectral-element implementations, save for a fluid potential formulation that incorporates the density, thereby yielding identical boundary terms on fluid and solid sides. We compare the second-order Newmark time extrapolation scheme with a newly implemented fourth-order symplectic scheme and argue in favour of the latter in cases of propagation over many wavelengths due to drastic accuracy improvements. Various validation examples such as full moment-tensor seismograms, wavefield snapshots, and energy conservation illustrate the favourable behaviour and potential of the method.

Key words: Numerical solutions; Computational seismology; Wave propagation.

1 INTRODUCTION

The solution to the 3-D elastodynamic wave equation in spherically symmetric background models has been addressed from various angles in the past such as normal-mode summation (Lognonné & Romanowicz 1990; Dahlen & Tromp 1998), direct solution method (Cummins *et al.* 1994; Geller & Ohminato 1994; Kawai *et al.* 2006), GEMINI (Friederich & Dalkolmo 1995), or axisymmetric finite differences (Igel & Weber 1995, 1996; Chaljub & Tarantola 1997; Toyokuni *et al.* 2005; Toyokuni & Takenaka 2006). While these methods present reasonable approximations or precise solutions to the problem of singular seismograms upon such media, the task of saving full time–space wavefields to construct exact, arbitrary-resolution seismic sensitivity kernels at the global scale (Nissen-Meyer *et al.* 2007a) has remained unfulfilled, mostly due to the computational load of such high-frequency fields (e.g. 1 Hz) in terms of CPU time (up to years), run-time memory occupation (Terabytes) and disk storage (10^{18} bytes). As proposed in Nissen-Meyer *et al.* (2007b), we argue for usage of a 2-D spectral-element method (SEM) on these grounds due to its favourable cost–accuracy trade-off for whole wavefields at the desired resolutions with its relative ease to accurately compute all necessary displacement, velocity, and strain tensor fields throughout the domain. In that introductory, methodological paper, we showed how to solve and implement a 2-D

weak (variational) form of elastodynamics for a full moment-tensor source and validated the solution against normal-mode summation for a homogeneous solid domain. Here, we shall address all aspects related to realistic earth models, that is, the solid–fluid discretization given a number of arbitrary discontinuity radii and velocities in a spherical semi-disk as shown in Fig. 1 (compare similar finite difference domains, e.g. Igel & Weber 1996; Thomas *et al.* 2000; Toyokuni *et al.* 2005). One generally constructs spheroidal skeletons of quadrilateral grid cells for respective finite depth ranges and connects these either conformingly or non-conformingly. While this is certainly not an ideal approach to accommodate grid spacing variations and seismic velocities, it has gained widest popularity due to its simplicity. Alternatively, one may optimize the mesh construction based on triangles (Käser & Dumbser 2006; Mercierat *et al.* 2006) or irregular grids (Käser & Igel 2001), both of which are superior for meshing but carry an inevitable computational burden since either interpolations, topological book-keeping, and index arrays need to be invoked, or the system may lose advantages such as the explicit time scheme, or tensorized matrix terms. Exploiting the advantageous simplicity of this 2-D computational domain, the meshing technique presented here is tailored towards flexible input such as the maximal frequencies and a mere list of discontinuities. The problem that velocities increase, while spherical-grid arc lengths decrease with depth has been tackled in 3-D SEM incarnations

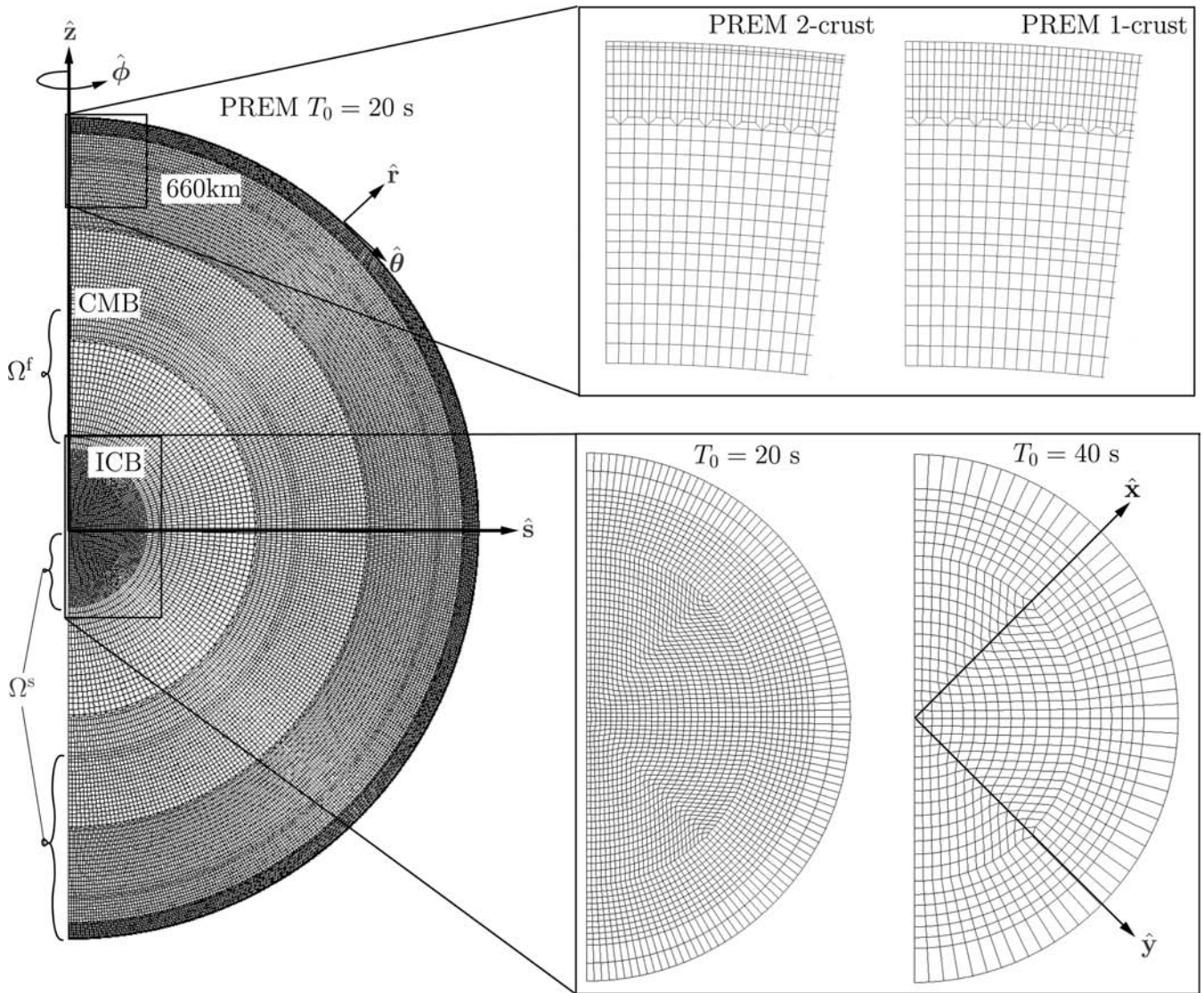


Figure 1. Left-hand panel: The semicircular, solid–fluid domain $\Omega = \Omega^s + \Omega^f$ discretized for the PREM background model using quadrilateral elements Ω_e for dominant source period $T_0 = 20$ s. Note that all discontinuities are honoured and several conforming coarsening levels are included to maintain a relatively constant resolution throughout the domain. Top right-hand panel: Enlargement of the crust for one (right-hand side) and two (left-hand side) crustal layers, and the upper mantle, including one mesh coarsening region. Note the variable vertical spacing due to discontinuities. Bottom right-hand panel: To circumvent the singularity at the centre, we apply eq. (5) to reshape rectangular elements, warranting an easy handle on grid spacing which varies maximally at the outermost elements of this central region.

by including coarsening regions such that the lateral number of elements above these is twice the number of elements below (Komatitsch & Tromp 2002) which we adopted for our simpler 2-D case (Nissen-Meyer *et al.* 2007b). Even in 2-D, we need a *divide et impera* parallelization concept for source periods below about 5 s. The convenient disk topology suggests a cake-piece decomposition to account for maximal load balancing and minimal communication, following the one-message approach of Tufo & Fischer (2001) and Deville *et al.* (2002). With the generation of high-frequency global wavefields in mind, we aim at maintaining high accuracy after propagation over many wavelengths. We therefore adapt a symplectic fourth-order time scheme which enjoys wide popularity in numerical astrophysics and molecular dynamics (e.g. Channell & Scovel 1990; McLachlan & Quispel 2006) and has recently been suggested for the elastodynamic case (Ampuero & Nissen-Meyer 2008). Its CPU cost is about 2.5 times that of the conventional second-order

Newmark scheme, but the accuracy significantly higher. Finally, we show seismograms for various settings and validate their accuracy against normal-mode summation. In addition, we examine the method via wavefield snapshots at critical depths such as the solid–fluid interface and the crust in light of relying on spatio-temporal global wavefields as the backbone for sensitivity kernels. The appendix provides some background on symplectic schemes.

2 MESHING OF 1-D BACKGROUND MODELS

2.1 Preliminaries and general constraints

Let us first define some geometric notation. As shown in Fig. 1 on the left-hand panel, we work in a semi-disk $\Omega = \Omega^s + \Omega^f$ of outer radius r_0 spanned by $\hat{\mathbf{x}}^{2D} = \hat{\mathbf{s}} + \hat{\mathbf{z}} = \hat{\mathbf{r}} + \hat{\boldsymbol{\theta}}$, where $s \in [0, r_0]$,

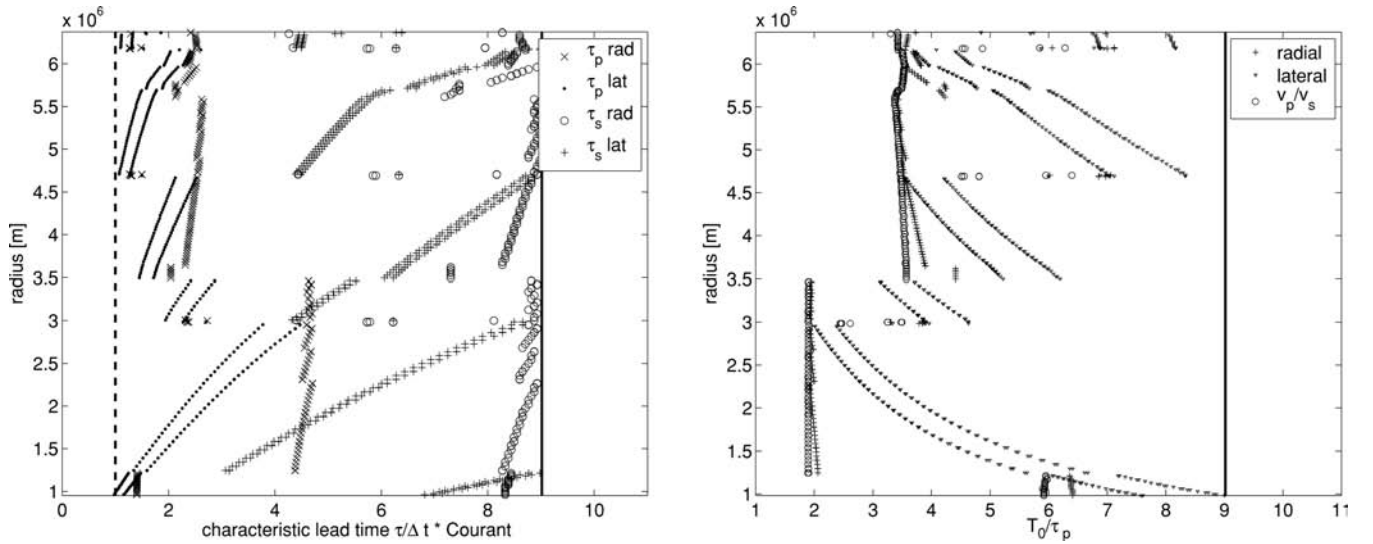


Figure 2. Left-hand panel: Elementally minimal and maximal characteristic lead times scaled by the time step and Courant number, $\tau_{p,s}C^0/\Delta t$ in the spherical part of the model space as a function of radius for PREM with source period $T_0 = 10$ s. The vertical line to the left denotes unity, that is, minimal possible $\min[\tau_p]$ due to the definition of Δt in eq. (2), and the vertical line to the right the corresponding maximal value given by the relationship for the source period, eq. (1), that is, $T_0C^0/(n_\Lambda^0\Delta t)$. Right-hand panel: Local temporal oversampling in terms of period and time step. Unity would be equivalent to the ideal case of no variation for the characteristic lead time τ whatsoever, impossible for any elastic medium. A more realistic aim, the empty circles to the left denote the ratio τ_p/τ_s which represents the limiting case of ideal meshing for the given background model and the inevitable grid spacing variation due to the polynomial basis of Gauss–Lobatto–Legendre points (Nissen-Meyer *et al.* 2007b), that is, with no oversampling due to suboptimal meshing. The vertical line to the right is the constant ratio $(T_0/\Delta t)(C^0/n_\Lambda^0)$, that is, the global worst case scenario after which these parameters have been chosen.

$z \in [-r_0, r_0]$ refer to cylindrical coordinates, and radius $r \in [0, r_0]$ and colatitude $\theta \in [0, \pi]$ denote the spherical system. The longitude $\phi \in [0, 2\pi]$ is identical for both coordinate systems, $\hat{\mathbf{x}}^{3D} = \hat{\mathbf{x}}^{2D} + \hat{\phi}$. Material properties are invariant in ϕ , and the seismic wavefields either invariant (monopole sources) or analytically continued from the (s, z) plane (Nissen-Meyer *et al.* 2007a); as a result we do not need to discretize this third dimension. The grid structure represents the tiling into non-overlapping solid (Ω_s^e) and fluid (Ω_f^e) elements within which functions are analytically mapped to a reference square $[-1, +1]^2$ to be expanded upon a polynomial basis (see Section 3.2).

The generation of a tangible mesh is no trivial matter. The following factors need to be considered as specific constraints from the spectral-element approach (see also Nissen-Meyer *et al.* 2007b, Sections 3.1 and 4.1):

- (i) Quadrilateral element shapes.
- (ii) Non-overlapping element boundaries.
- (iii) Discontinuities to coincide with element boundaries.

Point (i) is a minor issue inasmuch as the nature of sharp global boundaries is spherical, that is, mildly deformed and we do not need to accommodate sharp wedges such as in regional velocity models which may considerably influence the spacing variability and accuracy. We however face the complication that the seismic velocity $v_{p,s}(r)$ generally increases, while horizontal grid spacing (for spheroidal topologies) decreases with depth which works as a doubly detrimental effect for meshing purposes. Thus, we need to employ mesh coarsening to remediate these issues. Unless the expensive mortar element method (Chaljub *et al.* 2003) or a discontinuous Galerkin approach based on numerical fluxes (Käser & Dumbser 2006) are used, constraint (ii) only permits conforming coarsening architectures (Komatitsch & Tromp 2002), which is what we chose (Nissen-Meyer *et al.* 2007b). Due to the latter point (iii), however, we are faced with problems such as discretizing the thin crustal layers even for long periods, or the existence of the fluid outer core

(FOC), and drastic velocity changes across the inner-core boundary (ICB, see Section 2.3). For elastic-wave phenomena, one seeks a discretization in which the local P - or S -velocity-dependent wavelength $\Lambda[v_{p,s}(r)]$ is sampled by a minimal variation in the number of gridpoints per wavelength n_Λ anywhere in the spatial domain, and, simultaneously, minimal variation in the number of time samples $T_0/\Delta t$ for the source-induced, constant dominant (i.e. peak spectrum) period $T_0 = \Lambda[v(r)]/v(r)$. The limiting values for the number of gridpoints per wavelength, $n_\Lambda^0 = \min[n_\Lambda(r, \theta)]$, controlling the resolution, and the Courant number, $C^0 = \max[C(r, \theta)]$, controlling the stability, are inherent properties of the numerical scheme. They are related to T_0 and Δt via the characteristic lead time $\tau_{p,s}(r, \theta) = \Delta x(r, \theta)/v_{p,s}(r)^1$

$$T_0 = n_\Lambda^0 \max[\tau_p(r, \theta)], \quad (1)$$

$$\Delta t = C^0 \min[\tau_p(r, \theta)], \quad (2)$$

where for each element, grid spacing $\Delta x(r, \theta)$ also depends on θ due to the irregular Gauss–Lobatto–Legendre (GLL) point distribution across elements (Nissen-Meyer *et al.* 2007b), the occurrence of coarsening levels and the discretization of the central region. Meshing in light of these constraints and ideas principally aims at accommodating two factors: (i) obtaining a smallest possible total number of elements to minimize the run-time memory cost with (ii) the least variation in $\tau_{p,s}(r, \theta)$. This non-unique meshing procedure can be addressed from many angles.

¹ Subscripts p,s are mnemonic indicators for adhering to the respectively largest and smallest velocities for a given location; τ_s in the fluid for example is defined by the overarching P -wave velocity, and near the surface by surface-wave velocities.

Table 1. Characteristic lead times and time steps for various mesh resolutions and Courant number $C^0 = 0.6$ and $n_\Lambda^0 = 6$.

T_0 [s]	$\min(\tau_p^{\text{lat}})$ [s]	$\min(\tau_p^{\text{rad}})$ [s]	$\max(\tau_s^{\text{lat}})$ [s]	Δt [s]	$\max(\tau_s)C^0/\Delta t$	$T_0/\min(\Delta t^{\text{eff}})$
5	0.0917	0.117	0.832	0.0555	9.08	8.99
10	0.182	0.234	1.66	0.111	8.97	9.16
20	0.356	0.239	3.32	0.121	16.46	13.9
40	0.687	0.239	6.45	0.121	32.0	27.9

2.2 The background-model based spherical mesh

We chose a straightforward approach to define the mesh for our 2-D geometry: Given an anticipated period T_0 to be resolved, n_Λ^0 , and the location and velocity jump across discontinuities, we calculate the smallest global grid spacing and relate that to an appropriate number of lateral elements at the surface via the number of coarsening levels and the number of processors N_{proc} (see Section 2.5). The global time step is found given the limiting Courant number C^0 and $\min[\tau_p]$ from the newly created mesh. For the SEM, the minimal possible number of points per wavelength is $n_\Lambda^0 \approx 4$ (e.g. Ampuero & Nissen-Meyer 2008), whereas C^0 is usually determined empirically and depends on the type and order of the time extrapolation scheme (see Section 3.4) and model-space dimension. We obtain stable, accurate simulations up to $C^0 = 0.6$, but one should treat this choice carefully for each different mesh and time scheme (see Section 3.4). Clearly, for any mesh containing model property or grid spacing variations, Δt and T_0 are the global worst-case combinations, and the actual, local Courant number and number of gridpoints per wavelength are variables across the mesh. All applications in this paper are undertaken with a polynomial order $N_{\text{pol}} = 4$. Note that small polynomial orders ease matters in two ways: the smallest grid spacing within elements upon a GLL basis varies as N_{pol}^{-2} , and, secondly, smaller element sizes allow for easier adaptation to material interfaces, especially regarding closely spaced discontinuities like in the crust. Fig. 1 shows an example of a PREM model discretization for $T_0 = 20$ s (left-hand panel), and two magnified regions of interest, that is, the crust and upper mantle (top right-hand panel) and the central region (bottom right-hand panel, see next section). Lower resolutions are less effective and expose larger spacing variations, as seen in the crust. Note the variable vertical spacing due to the discretization of adjacent discontinuities such as 600 and 660 km. Generally, meshing is done only upon the lowest velocities (surface waves near the surface, S waves in deeper solid, P waves in fluid), but one subsequently determines the time step based on largest (P wave) velocities. It is thus desirable to assess this non-unique, potentially iterative process in terms of efficiency and overall cost. We present the mesh realization in terms of non-dimensionalized parameters, namely the characteristic lead time

$$\frac{\tau_{p,s}}{\min[\tau_p]} = \tau_{p,s} \frac{C^0}{\Delta t}, \quad (3)$$

and the local oversampling ratio

$$\frac{T_0}{\Delta t^{\text{eff}}(\mathbf{x})} = \frac{T_0}{n_\Lambda^0 \tau_p(\mathbf{x})}. \quad (4)$$

In both cases, the definition contains all relevant constant parameters such that this analysis can be viewed as independent both of the actual mesh resolution and the choice of the spatio-temporal discretization scheme (spectral elements, finite differences, etc.) and may, therefore, be useful in estimating the mesh quality for varying resolutions and methods. In Fig. 2 on the left-hand panel, we show

the variation of the characteristic lead time for the spherical part of the domain as a function of radius for a PREM model realization with $T_0 = 10$ s. We plot radial/lateral and maximal/minimal values of $\tau_{p,s}$ for each element, respectively. Clearly, the mesh coarsening is reflected by the jumps in lateral spacing, whereas the radial structure is smoothly kept within the bounds set by the lateral values. Another apparent feature is the fact that the lateral characteristic lead times τ_p seem to follow two profiles, that is, for a given radius, there are two minimal spacings. This is merely the result of the fact that we employ a different polynomial basis within axial elements [Gauss–Lobatto–Jacobi (GLJ) points] than for non-axial elements (GLL points). The profile with smaller values is due to axial elements, and all others follow the profile for larger values (see Nissen-Meyer *et al.* 2007b, Section 3). The distinction only exists for lateral τ as we utilize the same GLL points in the radial direction for all axial and non-axial elements. The right-hand panel in Fig. 2 shows the local oversampling ratio as a function of radius, and denotes the local numerical gap between defining largest grid spacing upon the period, but simultaneously noting the smallest spacing to abide stability for the temporal extrapolation. The lower bound is the mere τ_p/τ_s ratio, that is, the ideal case with local oversampling due to meshing, and the higher bound is the scaled ratio of period and time step $T_0/\Delta t$, that is, the global worst case. As before, the radial spacing follows a relatively smooth profile along the idealized case of the characteristic lead time ratio, whereas the lateral spacing is subject to coarsening layers and the main contributor to spacing variability or deviation from the smooth profile. Both these plots give an idea about the mesh quality, inasmuch as they delineate excursions from the respective minimal values (compare Komatitsch & Tromp 2002). Table 1 summarizes all relevant parameters for resolutions between $T_0 = 5$ s and $T_0 = 40$ s. Evidently, the smallest radial spacing is constant for meshes above 10 s such that the time step does not increase with the period, and the ratio of minimal versus maximal τ increases with the period. Apart from these cases of $T_0 > 10$ s, the meshing process is consistent and independent of the resolution in that source period T_0 and τ are linearly related and their ratio remains constant. The non-dimensionalized oversampling ratio in the last column is equally independent of the actual resolution as long as discontinuity locations do not determine the grid spacing as for $T_0 \geq 10$ s.

2.3 The central region

The singularity at $r = 0$ is circumvented using linear elements (Chaljub 2000) which introduces an additional problem: the vast spacing variations between linear $r < r_{\text{sl}}$ and circular regions $r > r_{\text{sl}}$. We accommodate this by defining

$$|x|^p + |y|^p = |r|^p, \quad x = s + z, \quad y = s - z, \quad 1 \leq p \leq 2, \quad (5)$$

where the exponent p can vary either linearly, quadratically, or cubically with the radius r (see Fig. 1). We then maintain an acceptable

$\min[\tau_p] \leq \tau_{p,s}(r < r_{sl}, \theta) \leq \max[\tau_s]$ where we take the extremal values of $\tau_{p,s}$ from the spherical domain $r \geq r_{sl}$, thereby avoiding the case that these linear elements determine the overall cost. Applying the mapping eq. (5), elements along the diagonal directions $x = 0$ and $y = 0$ are extremely deformed, and therefore, subsequently stretched outwards by half an element size, respectively (see Fig. 1c). The high mesh density in the inner-core stems from the fact that we effectively follow a drastic velocity drop across the inner-core boundary (ICB) since we only need to resolve outer-core P waves but inner-core S waves.

2.4 Mesh scaling and computational cost

Based on the two previous sections, we now revert to estimating the overall cost of the scheme that is launched upon a mesh as defined here. In recent work, Ampuero & Nissen-Meyer (2008) suggest for the cost on a topologically and structurally homogeneous mesh the factor $[L/(N_{\text{pol}}\Delta x)]^D [(L/v)/\Delta t]$ for fixed polynomial order N_{pol} , where L is the total propagation length, and D the dimension. Assessing the highly non-regular mesh, we modify this expression such that we can determine the quality and scaling properties as a function of the resolution. Let the cost be factorized into instantaneous cost (or an indicator for run-time memory occupation) \mathcal{I} formed by the total number of elements N_{el} , and the extrapolation cost \mathcal{E} , that is, the number of time steps:

$$\mathcal{I} = \frac{N_{\text{el}}}{N_{\text{el}}^{\text{ref}}} = \frac{2}{\pi r_0^2} (N_{\text{pol}} \Delta x_{\text{max}})^2 N_{\text{el}}, \quad (6)$$

$$\mathcal{E} = \frac{\Delta t^{\text{ref}}}{\Delta t} = \frac{T_0}{2\Delta t}. \quad (7)$$

In eq. (6), we scale N_{el} with an approximate reference number of fewest possible elements, assuming global coverage by the largest grid spacing such that $\min[\mathcal{I}] = 1$ represents the idealized scenario where there no velocity and spacing variation for a given period T_0 . In eq. (7), we take the sampling rate from the Nyquist sampling theorem as $\Delta t^{\text{ref}} = T_0/2$, such that again $\min[\mathcal{E}] = 1$ represents the cheapest possible setting. Fig. 3 depicts the instantaneous and extrapolation cost, respectively, as a function of period for PREM and PREM-1-crust realizations of the meshing technique. Evidently, low resolution is relatively ineffective inasmuch as the time step is controlled by the smallest grid spacing which in turn is the mere radial separation of the two closest discontinuities, mostly in the crust. For both cost functions, we enter another regime below 10 s, where the meshing process itself controls the computational cost and the two models are equally expensive. The slight increase of instantaneous cost with higher resolution below 7 s represents the cutoff resolution below which the instantaneous cost does not decrease any further due to the background model heterogeneities, overall velocity contrasts and the necessity to honor smaller velocities with smaller element sizes, independent from the actual resolution. The asymptotic convergence to 50 for the extrapolation cost reflects the ratio $T_0/\Delta t$ (as shown on the right-hand side of Fig. 2) which is then independent of the resolution but merely the difference in maximal and minimal characteristic lead times. In both instantaneous and extrapolating cost functions, the mesh is more cost-effective with increasing resolution, which is expected as we asymptotically mimic a continuous medium when increasing resolution at fixed background structure: we are progressively less limited by grid constraints such as discontinuity locations, permissible lateral number of elements, coarsening levels, and hemispheric mirror symmetry.

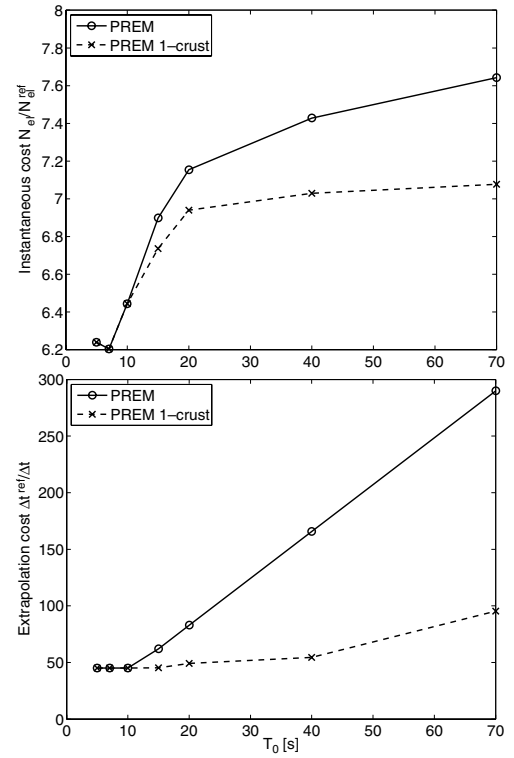


Figure 3. Computational cost of the mesh. Top panel: Instantaneous cost, that is, the size of the mesh. Bottom panel: Extrapolation cost, that is, the number of time steps. Both functions are constructed in a non-dimensionalized fashion such that they cancel out resolution dependencies when plotted as a function of the source period T_0 .

2.5 Parallelization and domain decomposition

The 2-D meshes described in the previous section are dense enough to reach typical memory limits when going below $T_0 \sim 5$ s. Such a mesh contains about 10^7 gridpoints, each carrying up to 40 (single-precision) global, scalar floating-point numbers at run-time memory, summing up to 1.5 GByte shared memory. This shall not remain a centennial problem in the wake of hardware improvements, but as we do strive to compute sensitivity kernels down to 1 s, it is indispensable to parallelize the method, especially in light of the extensive wavefield output. With the limited life-time of the parallelized version in mind, the code is written such that all message passing, communication and domain decomposition issues are confined to a single module and small amount of call statements to allow for a trivial denunciation of the parallel world. When spreading the computational workload across several CPUs and memory units, one seeks to minimize *inefficiency* (work imbalance), *latency* (number of neighbouring processors) and *bandwidth* (vector length to be communicated). We follow an approach by Tufo & Fischer (2001) (see also, Deville *et al.* 2002) with predefined index vectors of arbitrarily located gridpoints for each processor-processor pair to be exchanged such that processors may harbour separate subdomains or share multiple edges but always exchange one message with any neighbour. Fig. 4 shows examples of the domain decomposition strategy for the mesh down to $T_0 = 10$ s for four, eight and 16 processors from left- to right-hand side. We chose to split the domain laterally such that each processor has an equal amount of fluid and solid elements at the inevitable expense of focusing the axial workload (additional terms and source) onto two processors that occupy most of the axis (e.g. black and green domains in the

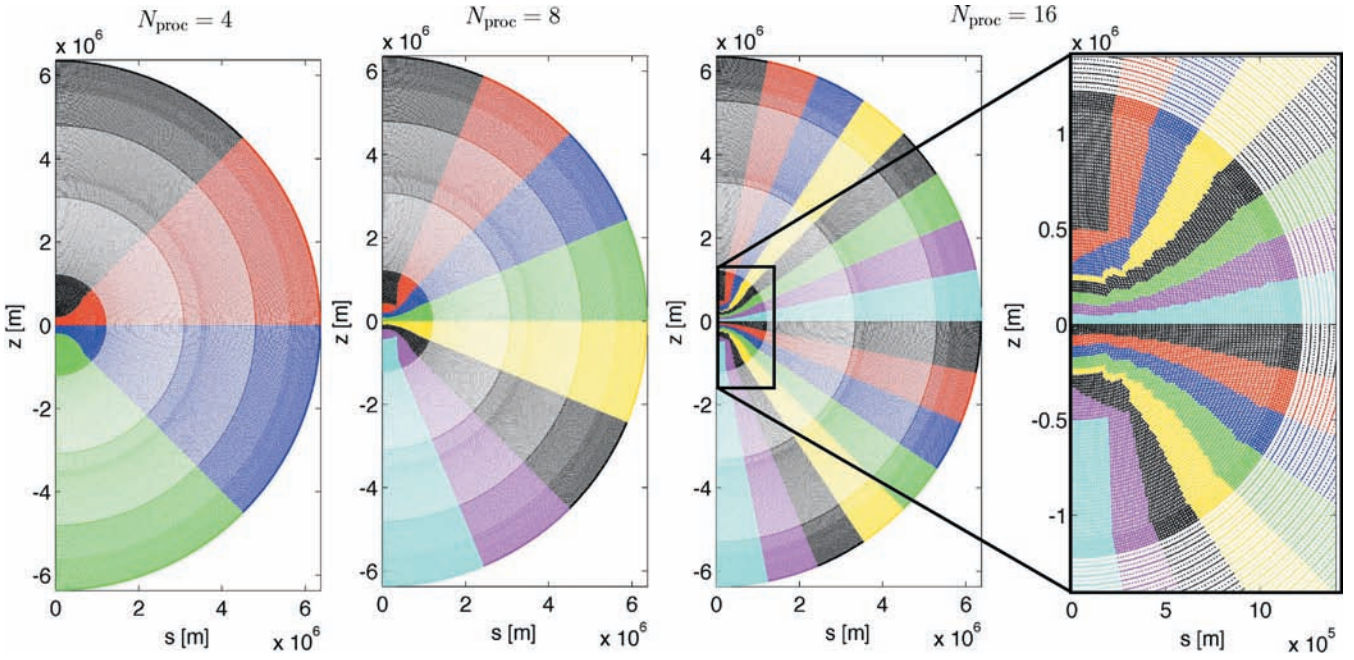


Figure 4. Domain decomposition for $T_0 = 10$ s and four (left-hand panel), eight (middle panel) and sixteen (right-hand side two panels) processors. Each processor has the same number of elements (load-balance), and maximally two neighbours (minimal latency) with small shared boundaries (small bandwidth). Each processor touches the axis, albeit to varying amounts. Axial terms are however 1-D, and do not add any countable CPU time to the overall scheme. The central region is subdivided such that each processor has exactly the same amount of elements.

left-hand panel). With axis treatment being 1-D, this does not hamper the favourable scaling of the parallelization. At this point, it is important to recall the application of saving spatio-temporal wavefields for the computation of sensitivity kernels, that is, such extensive I/O shall overshadow any slight processor imbalance due to additional axial terms or the source by and large. The actual number of processors needs to be an even multiplicative of 2 due to the equatorial symmetry and the connection of the spheroidal with the central region. While the former is trivial to decompose simply based upon colatitude, the latter central part poses a problem in keeping with the load-balance philosophy. We tackled this issue as seen in the magnified right-hand panel of Fig. 4 by defining discrete polynomials $z(s)$ under the premises of covering the same discrete area, maintaining a respective vertical and diagonal ‘thickness’ of at least one element, choosing outer-end locations given by the spheroidal processor latitudes, and touching the axis with at least one element. Most importantly, this approach guarantees exact load balancing in terms of number of elements, and minimal communication such that each processor only touches maximally two neighbours, taking an inherent topological advantage of the 2-D semi-disk bounded by the axis. Detailed analyses of fractional run-times and quantitative scaling can be found in Nissen-Meyer (2007).

3 THE SOLID-FLUID SYSTEM

3.1 3-D weak form

For the remainder of this paper, we will treat a non-rotating, non-gravitating, isotropic (Lamé parameters λ, μ) earth model. The 3-D elastodynamic weak form comprises the following three integral equalities for the displacement-potential vector (\mathbf{u}, χ) in an appropriate space with square-integrable derivatives for all admissible test

or trial functions (\mathbf{w}, w) (Nissen-Meyer *et al.* 2007a):

$$0 = \int_{\text{SIC}} \rho \mathbf{w} \cdot \ddot{\mathbf{u}} d^3 \mathbf{x} - \int_{\text{ICB}} (\hat{\mathbf{r}} \cdot \mathbf{w}) \ddot{\chi} d^2 \mathbf{x} + \int_{\text{SIC}} \{ \lambda (\nabla \cdot \mathbf{w}) (\nabla \cdot \mathbf{u}) + \mu \nabla \mathbf{w} : [\nabla \mathbf{u} + (\nabla \mathbf{u})^T] \} d^3 \mathbf{x}, \quad (8)$$

$$0 = \int_{\text{FOC}} \kappa^{-1} w \ddot{\chi} d^3 \mathbf{x} + \int_{\text{ICB}} w (\hat{\mathbf{r}} \cdot \mathbf{u}) d^2 \mathbf{x} - \int_{\text{CMB}} w (\hat{\mathbf{r}} \cdot \mathbf{u}) d^2 \mathbf{x} + \int_{\text{FOC}} \rho^{-1} \nabla w \cdot \nabla \chi d^3 \mathbf{x}, \quad (9)$$

$$\begin{aligned} & \left\{ \begin{array}{l} \hat{\mathbf{p}} \cdot \mathbf{w}(r_r \hat{\mathbf{z}}) \delta(t) \\ \mathbf{M} : \nabla \mathbf{w}(r_s \hat{\mathbf{z}}) H(t) \end{array} \right\} \\ &= \int_{\text{SMC}} \rho \mathbf{w} \cdot \ddot{\mathbf{u}} d^3 \mathbf{x} + \int_{\text{CMB}} (\hat{\mathbf{r}} \cdot \mathbf{w}) \ddot{\chi} d^2 \mathbf{x} \\ &+ \int_{\text{SMC}} \{ \lambda (\nabla \cdot \mathbf{w}) (\nabla \cdot \mathbf{u}) + \mu \nabla \mathbf{w} : [\nabla \mathbf{u} + (\nabla \mathbf{u})^T] \} d^3 \mathbf{x}, \quad (10) \end{aligned}$$

where ρ is the mass density, κ the incompressibility, $\hat{\mathbf{p}} \delta(t)$ the impulsive unity single-force vector acting at the receiver at axial radius r_r and $\mathbf{M} H(t)$ the moment tensor at the source at axial radius r_s as a step function in time, and second time derivatives are abbreviated as $\ddot{\chi} = \partial_t^2 \chi$. Regions solid-inner core (SIC), fluid-outer core (FOC), and solid mantle and crust (SMC) in eqs (8)–(10) are coupled by virtue of the surface integrals over inner-core boundary (ICB) and core–mantle boundary (CMB). The fluid region eq. (9) is treated via a displacement potential formulation (Nissen-Meyer *et al.* 2007a)

$$\mathbf{u} = \rho^{-1} \nabla \chi, \quad (11)$$

such that the fluid-domain wave equation becomes

$$\kappa^{-1} \ddot{\chi} = \nabla \cdot (\rho^{-1} \nabla \chi). \quad (12)$$

Note that the inclusion of the inverse density in this definition of the potential χ results in boundary integrals in eqs (8)–(10) that are independent of any intrinsic material properties. This is different from other spectral-element applications (Komatitsch & Tromp 2002; Chaljub & Valette 2004), in which the density-free potential leads to solid–fluid boundary terms that depend upon the discontinuity jumps. This is necessitated by their inclusion of gravitation, but in our gravity-free case it is indeed feasible to pursue this convenient definition in eq. (11). For spherically symmetric models, the 3-D weak system eqs (8)–(10) may be reduced to its 2-D equivalent by factorizing (\mathbf{u}, χ) into their (s, z) and ϕ dependencies and analytically evaluating the longitudinal integrals such as $\int_0^{2\pi} \sin^2 \phi \, d\phi = \pi$ as shown in Nissen-Meyer *et al.* (2007a, Section 4.4).

3.2 Spatial discretization: the fluid core and its boundaries

The dimension reduction yields a collection of integrals to be evaluated over the fluid and solid parts of the D-shaped domain, as well as along their boundaries. These integrals are broken into a sum of elemental integrals, the evaluation of which requires knowledge of the geometrical mapping, and usage of a quadrature rule in the parent domain $[-1, +1]^2$. Following Nissen-Meyer *et al.* (2007b, Sections 3.2–3.3), we approximate integrals over a fluid element Ω_e^f by

$$\int_{\Omega_e^f} \chi(s, z, t) s \, ds \, dz \approx \sum_{p,q=0}^N \sigma_p^N \sigma_q^N s(\xi_p, \eta_q) \mathcal{J}(\xi_p, \eta_q) \sum_{i,j=0}^N l_i^N(\xi_p) l_j^N(\eta_q) \chi^{ij}(t), \quad (13)$$

where $\sigma_{p,q}$ are integration weights, \mathcal{J} the Jacobian, and $l_i^N(\xi_p)$ the i th Lagrange interpolating function of polynomial order N evaluated at GLJ points ξ_p in an element touching the axis, or GLL points elsewhere in the domain (Fournier *et al.* 2004). We will refrain from a notational distinction between axial and non-axial elements for the sake of brevity, and note special axial treatment when necessary.

3.2.1 Fluid mass and stiffness terms

Using the notation of Nissen-Meyer *et al.* (2007b, Section 4), the mass term for a non-axial element inside the fluid reads

$$\int_{\Omega_e^f} \kappa^{-1} w \ddot{\chi} s \, ds \, dz \approx \sum_{IJ} w^{IJ} \sum_{ij} M_f^{IJij} \ddot{\chi}^{ij} = \sum_{IJ} w^{IJ} (\kappa^{IJ})^{-1} \sigma_I \sigma_J s^{IJ} \mathcal{J}^{IJ} \ddot{\chi}^{IJ}. \quad (14)$$

Axial elements are treated equivalently after the substitution $\sigma_I \rightarrow \bar{\sigma}_I (1 + \xi_I)^{-1}$ and application of L'Hospital's rule on the axis $s^{0J} (1 + \xi_0)^{-1} = (\partial_\xi s)^{0J}$ (Nissen-Meyer *et al.* 2007b, Section 3.3). Note that the mass terms are the same for monopole, dipole and quadrupole source types, respectively, since the average over $\sin^2 \phi$ and $\cos^2 \phi$ is the same as over $\sin^2 2\phi$ and $\cos^2 2\phi$. Compared to the lengthy solid stiffness terms in Nissen-Meyer *et al.* (2007b, appendix B), the fluid stiffness terms take on a straightforward, computationally light form. All source types contain the basic scalar, monopole

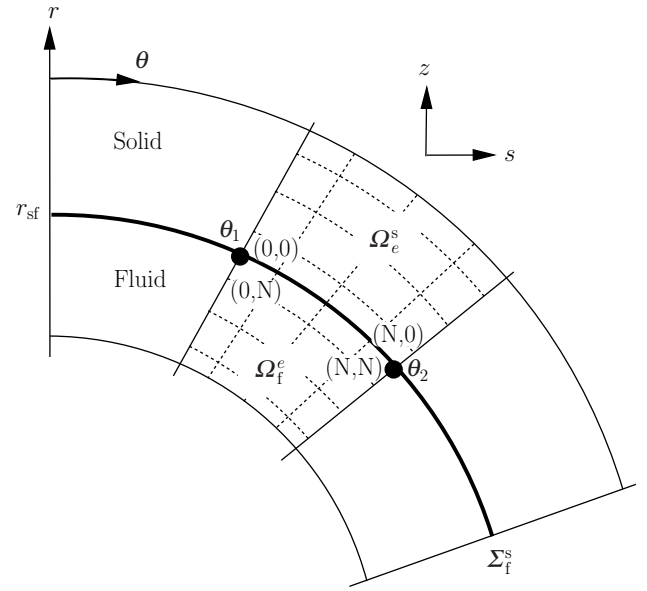


Figure 5. Sketch to illustrate our nomenclature around a solid–fluid boundary (in this case CMB). r_{sf} is the radius, and Σ_f^s the surface of the boundary. (...) denote the respective elemental indices (I, J) of Gauss–Lobatto–Legendre points of the boundary-hugging corner points of elements above and below.

Laplacian integral

$$\int_{\Omega_e^f} \rho^{-1} (\partial_s w \partial_s \chi + \partial_z w \partial_z \chi) s \, ds \, dz \approx \sum_{IJ} w^{IJ} \left[\sum_i D_\xi^{Li} C^{iJ} G_1^{iJ} \sum_j D_\eta^{jJ} \chi^{ij} + \sum_p D_\xi^{Lp} C^{pJ} G_2^{pJ} \sum_i D_\xi^{ip} \chi^{iJ} + \sum_j D_\eta^{Jj} C^{jJ} G_3^{jJ} \sum_i D_\xi^{iJ} \chi^{ij} + \sum_q D_\eta^{Jq} C^{Jq} G_4^{Jq} \sum_j D_\eta^{jq} \chi^{iJ} \right]. \quad (15)$$

The dipole fluid stiffness term is completed by adding to eq. (15) once, and the quadrupole fluid stiffness term four times, the expression

$$\int_{\Omega_e^f} \rho^{-1} s^{-2} w \chi s \, ds \, dz \approx \sum_{IJ} w^{IJ} (\rho^{IJ})^{-1} \sigma_I \sigma_J (s^{IJ})^{-1} \mathcal{J}^{IJ} \chi^{IJ} \quad (16)$$

for non-axial elements and

$$\int_{\Omega_e^f} \rho^{-1} s^{-2} w \chi s \, ds \, dz \approx \sum_{I>0J} w^{IJ} \sigma_J \left[(\rho^{IJ})^{-1} \sigma_I (s^{IJ})^{-1} \mathcal{J}^{IJ} \chi^{IJ} + D_\xi^{I0} (\rho^{0J})^{-1} \bar{\sigma}_0 \mathcal{J}^{0J} (\partial_\xi s^{0J})^{-1} \sum_{i>0} D_\xi^{i0} \chi^{iJ} \right] \quad (17)$$

for axial elements. We abbreviated several quantities as defined in Nissen-Meyer *et al.* (2007b, Table 1):

$$C^{ij} = \sigma_i \sigma_j s^{ij} (\mathcal{J}^{ij})^{-1} \text{ (non-axial),}$$

$$C^{ij} = \bar{\sigma}_i \sigma_j s^{ij} (1 + \xi_i)^{-1} (\mathcal{J}^{ij})^{-1} \text{ (axial, } i > 0),$$

$$C^{0j} = \bar{\sigma}_0 \sigma_j s_\xi^{0j} (\mathcal{J}^{0j})^{-1} \text{ (axial, } i = 0),$$

$$D_\xi^{Li} = \partial_\xi l_i(\xi_i), \quad D_\eta^{Lj} = \partial_\eta l_j(\eta_j) = \partial_\xi l_j(\xi_j),$$

$$G_1^{iJ} = z_\xi^{iJ} z_\eta^{iJ} + s_\xi^{iJ} s_\eta^{iJ}, \quad G_2^{pJ} = z_\eta^{pJ} z_\eta^{pJ} + s_\eta^{pJ} s_\eta^{pJ},$$

$$G_3^{Ij} = z_\xi^{Ij} z_\eta^{Ij} + s_\xi^{Ij} s_\eta^{Ij}, \quad G_4^{Jq} = z_\xi^{Jq} z_\xi^{Jq} + s_\xi^{Jq} s_\xi^{Jq},$$

where for example, $s_\xi = \partial_\xi s$.

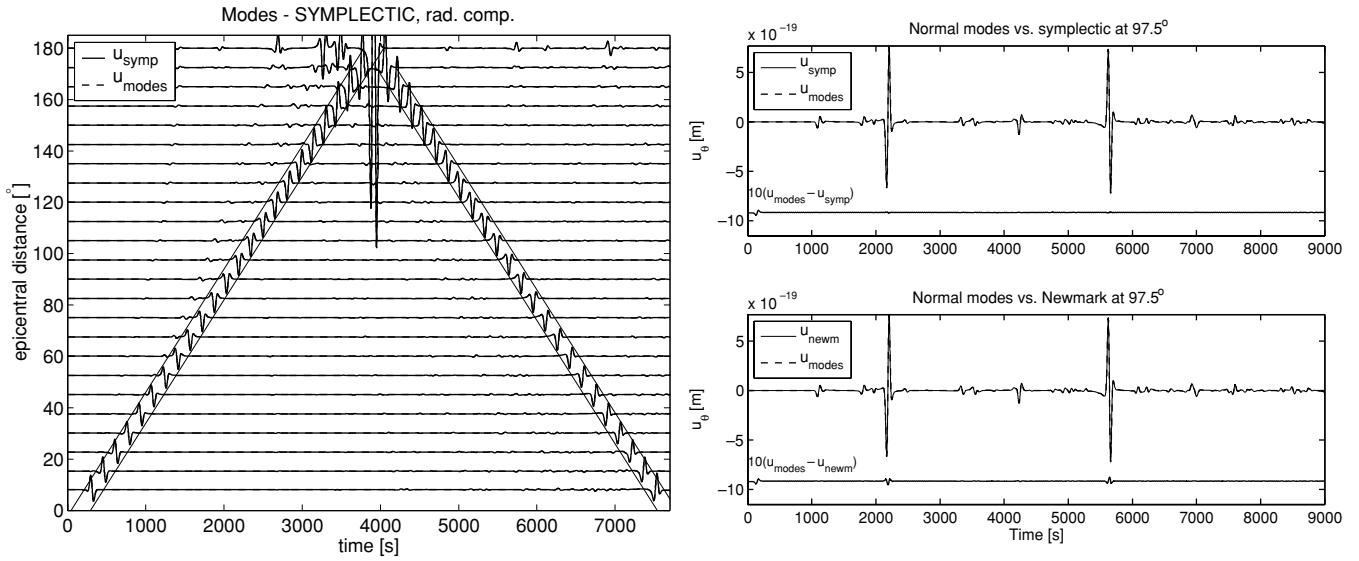


Figure 6. Left-hand panel: Global seismograms from the symplectic SEM and normal-mode summation for a spherical version of Lamb's problem, with the surface wave window used to compute the errors. Differences between the methods are indistinguishable. Right-hand panels: Longitudinal-component seismograms at $\Delta = 97.5^\circ$ for the symplectic scheme (top panel) and the Newmark scheme (bottom panel) compared to mode synthetics. Both panels contain the ten-fold exaggerated residual between the two traces below the seismogram.

3.2.2 Solid–fluid coupling terms

Let $\Sigma = \Sigma_f^s \cup \Sigma_s^f$ denote the union of solid–fluid boundaries such that we can specify, for example, those non-axial element edges in the solid that touch the boundary at radius r_{sf} from above by Σ_e^s and those axial element edges in the fluid that touch the boundary from below by Σ_f^e , as shown in Fig. 5. Note that only purely circular element types (see Nissen-Meyer *et al.* 2007b, Appendix A) may be located at discontinuities. We also define the colatitudinal element boundaries such that within an element $\theta_1 \leq \theta \leq \theta_2$ for the northern hemisphere and vice versa for the southern hemisphere. We can then represent the two types of boundary terms as

$$\int_{\Sigma_e^s} (w_s \sin \theta + w_z \cos \theta) \ddot{\chi} r^2 \sin \theta d\theta \approx \frac{r_{sf}^2}{2} |\theta_2 - \theta_1| \sum_I \sigma_I (\sin \theta)_{I0} [w_s^{I0} (\sin \theta)_{I0} + w_z^{I0} (\cos \theta)_{I0}] \ddot{\chi}_{I0} \quad (18)$$

for non-axial elements inside the solid region, and

$$\int_{\Sigma_f^e} w (u_s \sin \theta + u_z \cos \theta) r^2 \sin \theta d\theta \approx \frac{r_{sf}^2}{2} |\theta_2 - \theta_1| \sum_I w^{I0} \sigma_I (\sin \theta)_{I0} [u_s^{I0} (\sin \theta)_{I0} + u_z^{I0} (\cos \theta)_{I0}] \quad (19)$$

for non-axial elements inside the fluid. Note that the statically pre-computable factors such as $(r_{sf}^2/2)|\theta_2 - \theta_1|\sigma_I(\sin \theta)_{I0}^2$ are identical in both cases and on both sides of the boundary since our potential formulation eq. (11) results in such material-independent boundary terms. Furthermore, the coupling terms are identical for all excitation types for the same reason as the mass terms. For axial elements, we again replace σ_I by $\bar{\sigma}_I(1 + \xi_0)^{-1}$ in eqs (18) and (19), respectively. For the axis $I = 0$ itself, we apply L'Hospital's rule to remove the singularity: $(\sin \theta)_{00}(1 + \xi_0)^{-1} = (\partial_{\xi s})_{00} = \frac{1}{2}\theta_2 r_{sf}$. In practice, the solid–fluid model leads to two distinct domains with respective global numbers such that the corresponding indices from both sides in the boundary terms need to be found based on a coordinate search. Being a 1-D operator, the boundary term is compu-

tationally invisible to the overall cost of the scheme, similar to axial stiffness terms.

3.3 Global system in time

Upon inserting the polynomial expansion outlined above, the 2-D set of weak equations translates to its discretized counterpart, a coupled, global system of ordinary differential equations in time, which, following the same procedure as in Nissen-Meyer *et al.* (2007b, Section 4), reads

$$\mathbf{M}_s \ddot{\mathbf{u}}(t) + \mathbf{K}_s \mathbf{u}(t) + \mathbf{B}_f^s \dot{\chi}(t) = \mathbf{0}, \quad (20)$$

$$\mathbf{M}_f \ddot{\chi}(t) + \mathbf{K}_f \chi(t) - \mathbf{B}_f^s \mathbf{u}(t) + \mathbf{B}_f^f \mathbf{u}(t) = \mathbf{0}, \quad (21)$$

$$\mathbf{M}_s \ddot{\mathbf{u}}(t) + \mathbf{K}_s \mathbf{u}(t) + \mathbf{B}_f^s \dot{\chi}(t) = \mathbf{f}(t). \quad (22)$$

We discussed the action of the fluid mass matrix \mathbf{M}_f , fluid stiffness matrix \mathbf{K}_f , and the solid–fluid coupling matrices $\mathbf{B}_{s,f}$ on \mathbf{u} and $\dot{\chi}$ in Section 3.2. The solid-domain terms involving the mass matrix \mathbf{M}_s , stiffness matrix \mathbf{K}_s , and source terms \mathbf{f} , as well as the relation to their elemental counterparts are detailed in Nissen-Meyer *et al.* (2007b, Section 4 and appendix B). We now turn our attention on the temporal discretization, using two different approaches.

3.4 Temporal discretization

Temporal evolution problems such as eqs (20)–(22) are ubiquitous in a variety of physical phenomena, hence a zoo of possible numerical approximations exists. The common principle is to find a numerical expression for the ODE type $\ddot{u} = f(u)$ at time t discretized by steps Δt after applying the inverse mass matrix (\mathbf{M}^{-1}), and calculating internal (e.g. $\mathbf{K}\mathbf{u}$) and external (\mathbf{f}) forces. In seismology, a second-order Newmark scheme has primarily been applied to solve such systems (Komatitsch & Vilotte 1998). Here, we employ this Newmark scheme, but will also argue for symplectic integration based on

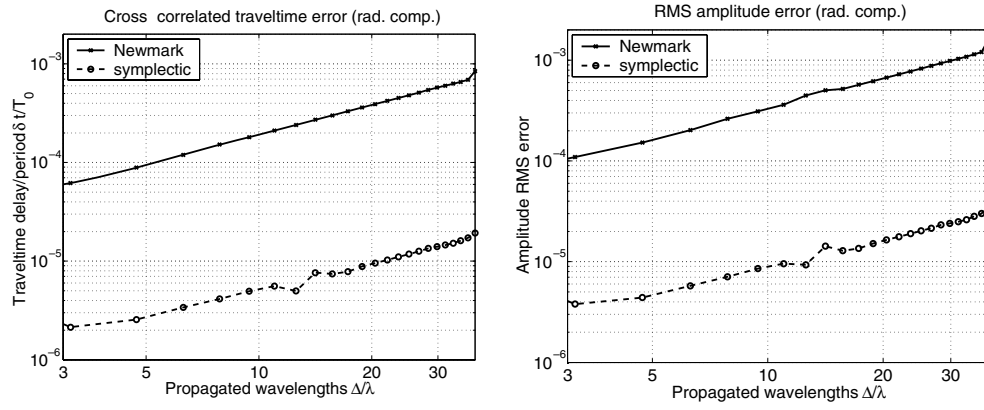


Figure 7. Left-hand panel: Relative cross-correlated traveltine error (with respect to normal modes) for the surface wave time window in Fig. 6 and both time schemes. Right-hand panel: Surface wave rms amplitude error *after* shifting the traces to eliminate the traveltine error.

superior accuracy at small additional cost and straightforward implementation. We explicitly present the respective schemes in this section and provide a tutorial-style introduction and summary of symplectic integration in Appendix A. For brevity, let us concatenate internal and external forces for the fluid as $\Xi_f(\mathbf{u}, \chi) = -\mathbf{K}_f \chi + \mathbf{B}_f \mathbf{u}$ and solid domains as $\Xi_s(\mathbf{u}, \ddot{\chi}) = -\mathbf{K}_s \mathbf{u} - \mathbf{B}_s \ddot{\chi} + \mathbf{f}$.

3.4.1 Second-order Newmark scheme

The second-order, acceleration-driven explicit Newmark scheme we employed in Nissen-Meyer *et al.* (2007b) is extended to the solid–fluid case as (Chaljub & Valette 2004)

$$\chi(t + \Delta t) = \chi(t) + \Delta t \dot{\chi}(t) + \frac{1}{2} \Delta t^2 \ddot{\chi}(t), \quad (23)$$

$$\mathbf{u}(t + \Delta t) = \mathbf{u}(t) + \Delta t \dot{\mathbf{u}}(t) + \frac{1}{2} \Delta t^2 \ddot{\mathbf{u}}(t), \quad (24)$$

$$\ddot{\chi}(t + \Delta t) = \mathbf{M}_f^{-1} \Xi_f[\mathbf{u}(t + \Delta t), \chi(t + \Delta t)], \quad (25)$$

$$\ddot{\mathbf{u}}(t + \Delta t) = \mathbf{M}_s^{-1} \Xi_s[\mathbf{u}(t + \Delta t), \ddot{\chi}(t + \Delta t)], \quad (26)$$

$$\dot{\chi}(t + \Delta t) = \dot{\chi}(t) + \frac{1}{2} \Delta t [\ddot{\chi}(t) + \ddot{\chi}(t + \Delta t)], \quad (27)$$

$$\dot{\mathbf{u}}(t + \Delta t) = \dot{\mathbf{u}}(t) + \frac{1}{2} \Delta t [\ddot{\mathbf{u}}(t) + \ddot{\mathbf{u}}(t + \Delta t)]. \quad (28)$$

This scheme is fully explicit and requires one evaluation of the global forces per time step which is the dominant factor for CPU time. Note that the order of evaluation is important as we need $\ddot{\chi}$ on solid–fluid boundaries from eq. (25) for the internal forces on the solid side eq. (26).

3.4.2 Fourth-order symplectic scheme

Alternatively, we implemented a fourth-order symplectic scheme which has been suggested for the elastodynamic spectral-element method by Ampuero & Nissen-Meyer (2008), following Omelyan *et al.* (2002, eq. 22) and Omelyan *et al.* (2003). This means that we evaluate all terms four times in between t and $t + \Delta t$. Setting $\mathbf{u}_0 = \mathbf{u}(t)$, $\dot{\mathbf{u}}_0 = \dot{\mathbf{u}}(t)$ and $\mathbf{u}_5 = \mathbf{u}(t + \Delta t)$, $\dot{\mathbf{u}}_4 = \dot{\mathbf{u}}(t + \Delta t)$ (ditto for χ , $\dot{\chi}$) and rewriting $\Xi_s[\mathbf{u}, \Xi_f(\mathbf{u}, \chi)] = -\mathbf{K}_s \mathbf{u} - \mathbf{B}_s \Xi_f(\mathbf{u}, \chi) + \mathbf{f}$ to eliminate $\ddot{\chi}$, the scheme reads

$$\chi_i = \chi_{i-1} + \kappa_i \Delta t \dot{\chi}_{i-1}, \quad (29)$$

$$\mathbf{u}_i = \mathbf{u}_{i-1} + \kappa_i \Delta t \dot{\mathbf{u}}_{i-1}, \quad (30)$$

$$\dot{\chi}_i = \dot{\chi}_{i-1} + \pi_i \Delta t \mathbf{M}_f^{-1} \Xi_f(\mathbf{u}_i, \chi_i), \quad (31)$$

$$\dot{\mathbf{u}}_i = \dot{\mathbf{u}}_{i-1} + \pi_i \Delta t \mathbf{M}_s^{-1} \Xi_s[\mathbf{u}_i, \Xi_f(\mathbf{u}_i, \chi_i)], \quad (32)$$

$$\chi_5 = \chi_4 + \kappa_5 \Delta t \dot{\chi}_4, \quad (33)$$

$$\mathbf{u}_5 = \mathbf{u}_4 + \kappa_5 \Delta t \dot{\mathbf{u}}_4, \quad (34)$$

where $i = 1, \dots, 4$ denotes an inner loop of updates that requires a four-fold evaluation of all forces per time step. This scheme is explicit, does not require storage of past time steps and avoids the usage of the acceleration resulting in less run-time memory occupation than the Newmark scheme eqs (23)–(28). Factorization coefficients κ_i, π_i are usually determined by symbolic algebra; the position extended Forest-Ruth like (PEFRL) scheme we adopted takes the values

$$\kappa_1 = \kappa_5 = \alpha, \quad \kappa_2 = \kappa_4 = \gamma, \quad \kappa_3 = 1 - 2(\gamma + \alpha), \quad (35)$$

$$\pi_1 = \pi_4 = 1/2 - \beta, \quad \pi_2 = \pi_3 = \beta, \quad (36)$$

where

$$\alpha = +0.1786178958448091, \quad (37)$$

$$\beta = -0.2123418310626054, \quad \text{and} \quad (38)$$

$$\gamma = -0.06626458266981849. \quad (39)$$

See Appendix A for details on symplectic time integration and an outline to obtain factorization coefficients. In general, this scheme satisfies more stringent conditions on preserving geometrical properties such as total energy, (angular) momentum or time-reversal symmetry which are not respected by other schemes such as Newmark. While short-time phenomena are expected to be sufficiently discretized using standard, non-symplectic techniques, any long-time issues including propagation over many wavelengths shall be much better represented by minimizing numerical dissipation of these invariants. More importantly, as shown in Ampuero & Nissen-Meyer (2008), it is conceivable to consider that seismic-wave propagation is at, or at least will soon enter, the stage of deducing information from high-frequency arrivals at large offsets for which minimal dispersion and accurate techniques are crucial.

Fig. 6 compares Newmark to symplectic time schemes through a spherical version of Lamb's problem (vertical single-force acting on the surface above a homogeneous model space). The left-hand panel shows both normal-mode summation seismograms and synthetics from the SEM using the symplectic scheme for $T_0 = 100$ s. The right section depicts single seismograms for both Newmark and symplectic time schemes versus normal modes, respectively, and the residual times 10 below. Evidently, both methods are excellent fits to the reference solution, but the symplectic scheme on top is virtually perfect in that even the magnified residual trace shows no sign of inaccuracies whereas the residual between modes and Newmark scheme (bottom panel) is characterized by a small difference for the two surface wave passings. By windowing on one such specific phase (surface wave, left-hand panel in Fig. 6), one can furthermore deduce the behaviour and accuracy as a function of propagated wavelengths, and apply quantitative measures such as cross correlation, the same way as used in obtaining traveltimes differences between observed and synthetic data. As seen in Fig. 7, both schemes exhibit a linear increase in traveltimes error with dis-

tance on a log-log plot, but with different values. The symplectic scheme is about 40 times more accurate, and requires 2.5 times more CPU time since (i) it involves a four-fold force-evaluation and (ii) the Courant number C^0 may be chosen 1.5 times larger compared to the Newmark case. For propagation distances over 100 wavelengths (linearly extrapolating the cross-correlated traveltimes error function in Fig. 7), the symplectic scheme becomes feasible in cost if one decreases the time step of the Newmark scheme to obtain similar accuracies. A quantitative cost-accuracy trade-off to deduce optimal simulation parameters and facilitate the decision on which spatial and temporal approximation orders are most appropriate is detailed in Ampuero & Nissen-Meyer (2008). The right-hand panel in Fig. 7 shows a similar error behaviour, but for the amplitude RMS fit after aligning the two time-series by the traveltimes error obtained in the cross-correlation plot. The respective 'bumps' at about 14 propagated wavelengths are from receivers near the antipode, where amplitudes are much larger due to the coincidental arrival of the surface wave from all directions (see Fig. 6, left-hand panel), and hence these error ratios slightly more sensitive to inaccuracies. For a

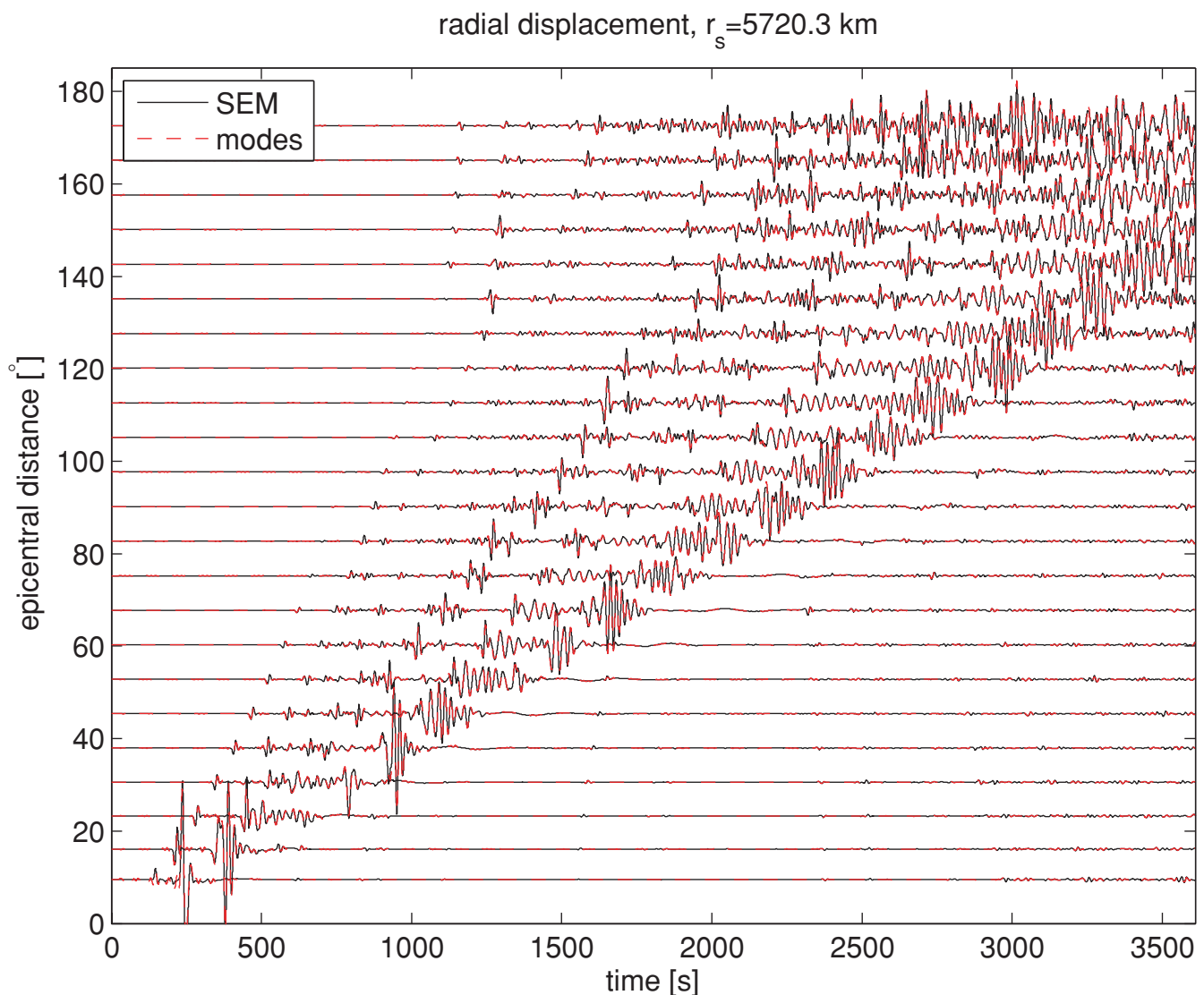


Figure 8. Radial displacement time-series distributed along the Earth's surface, calculated using both the SEM and normal-mode summation for the isotropic PREM background model at $T_0 = 20$ s and a full moment-tensor event at radius $r_s = 5720.3$ km. The traces are virtually indistinguishable, that is, the SEM agrees with the reference solution for all distances and phases. Small (acausal) 'noise' comes from truncated normal-mode summation.

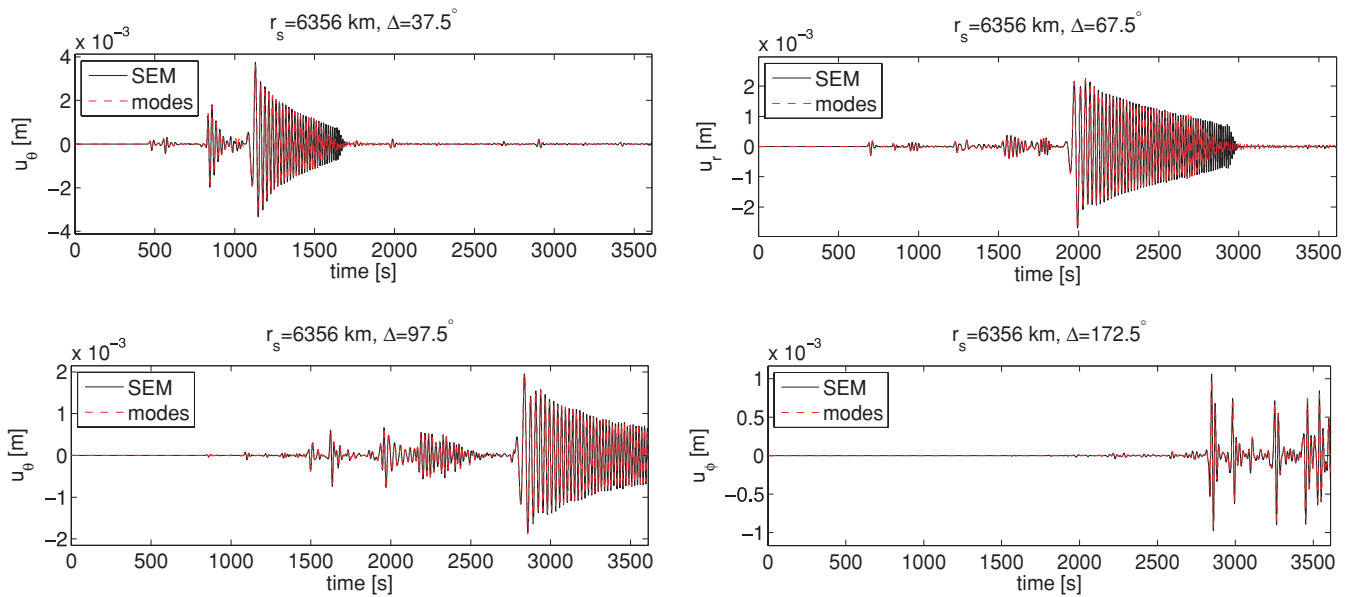


Figure 9. Selected SEM and normal-mode seismograms for the Vanuatu event at 15 km depth and the PREM model and $T_0 = 30$ s. The vertical axes denote the respective components of the actual ground displacement, and the titles the epicentral distances. All complex details in the seismograms including direct body waves, dispersive surface waves, and reflected phases, are correctly modelled except for small differences at the highest frequencies of the surface waves for which the mode catalogue is insufficient and hence the reference solution not entirely reliable.

detailed analysis of dispersion errors and the symplectic time scheme including derivations of appropriate schemes, expressions of error and cost functions, and an approach to choose cost-minimizing, ‘optimal’ simulation parameters, consult Ampuero & Nissen-Meyer (2008).

4 SNREI VALIDATION

After detailing the meshing process, the cost efficiency of the scheme, discretization of the solid–fluid configuration, and the implementation and validation of a symplectic time scheme, we now turn to comparing the SEM to reference solutions for spherically symmetric background models. For the specific case of non-gravitating SNREI (spherically symmetric, non-rotating, elastic and isotropic) earth models, we use a normal-mode summation method that allows for neglecting gravity completely. We will consider PREM (two-layered crust, isotropic mantle) for varying resolutions down to 10 s. Simulations are (unless otherwise noted) done using temporal Dirac-delta functions for various full moment-tensor sources and seismograms convolved a posteriori with a Gaussian of half-width T_0 . In an attempt to cover the most critical regions and configurations with regards to whole-wavefield applications for sensitivity kernels, we consider shallow and deep sources, show seismograms upon a full moment tensor collected over the 3-D globe, show wavefield snapshots around the CMB and upper mantle to verify the lack of significant amounts of spurious energy in the wavefield, and time-dependent energy plots to validate both the time scheme and the absence of any numerical dispersion throughout the domain.

4.1 Global synthetics

In this section, we show simulations of events at a shallow (15 km) and larger (650 km) depth to validate both strong surface waves and weaker body waves (e.g. P_{diff}) and cover the approximate depth ranges of earthquakes and surface distribution of receivers to be

included in a database for arbitrary source–receiver pairs. We will not differentiate between different source types as this has been extensively treated in the preceding paper but show results for full moment tensors taken from actual CMT solutions. Additionally, we do not deem it appropriate to quantify the misalignment by any fixed-station metric (such as the ‘energy misfit’) due to the abundance of complex-geometry paths and different accuracies for each path and phase. The simulations in this section were done on four processors for a mesh resolution down to 10 s (see Fig. 4, left-hand panel) and took 6 hours. Fig. 8 shows a section of radial displacement time-series distributed along a great circle path of epicentral distances between 7.5° and 172.5° for a CMT solution from an event in Brazil but moved to 650 km depth. The traces compare the SEM (straight lines) to mode summation (dashed lines) for $T_0 = 20$ s and are virtually identical for all distances and phases. This means that the meshing process is approved and the solid–fluid configuration is correctly formulated and implemented. Note that any slight differences are due to acausal noise from the mode summation. This simulation has been undertaken using the Newmark time scheme eqs (23)–(28), but the symplectic scheme gives indistinguishable results for these settings.

Fig. 9 compares single seismograms from simulations of a Vanuatu event (depth: 15 km) with strong and dispersive surface waves for $T_0 = 30$ s for both SEM and normal-mode summation. All parts of the seismograms expose an excellent fit between the SEM and normal-mode summation even for the dispersive surface wave trains at all distances and on all three channels with the exception for the last few parts of the surface wave train in the top panels (r and θ components) for which the mode summation is inaccurate due to truncation. In any case, this event within the crust is arguably one of the most difficult settings for global simulation to obtain correct results due to the thin crustal layer sensitivity and strong dispersion. In conclusion, this sampling of two extreme event depths and globe-spanning receiver locations strongly suggests that the method runs very accurately for the entire wavefield. The restriction to dominant

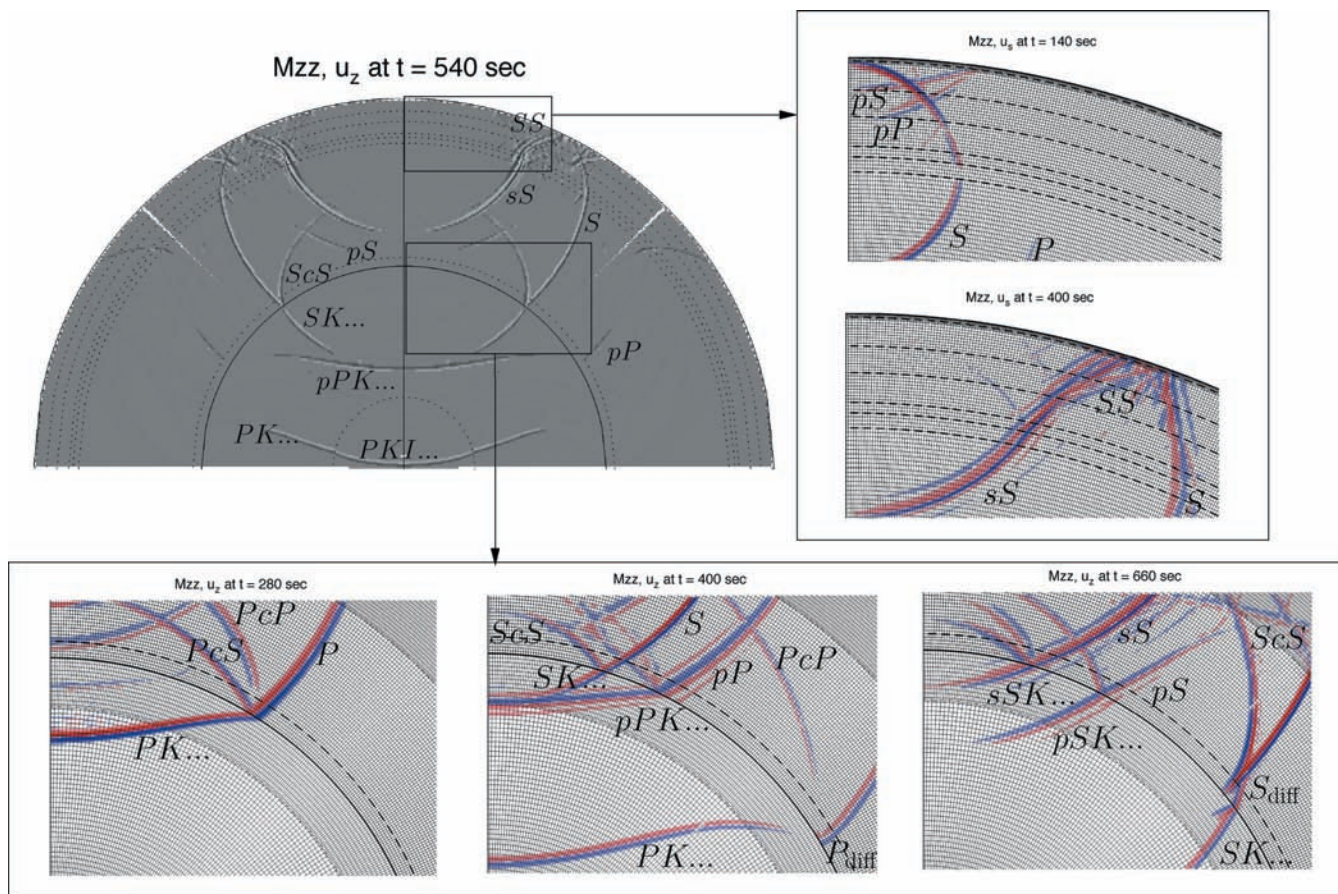


Figure 10. Wavefield snapshots for a monopole M_{zz} source at 650 km depth for $T_0 = 10$ s, including magnified panels of the crust–upper mantle and core–mantle boundary regions; labels refer to major seismic phases.

periods greater than 20 s is solely based upon our inability to obtain reliable normal-mode reference solutions for shorter periods.

4.2 Wavefield snapshots

In this sense, we shall find it helpful to additionally illuminate waves at depth via wavefield snapshots. In light of saving entire time–space wavefields to construct sensitivity kernels, it is also important to zoom into critical regions (e.g. surface, upper mantle discontinuities, CMB, core, coarsening levels, processor boundaries).

Fig. 10 (top left-hand panel) shows the northern hemisphere of the z component wavefield upon a M_{zz} source at 650 km for $T_0 = 10$ s. All phases (see labels) are present, and no unrefined noise is apparent. The straight, white lines appear due to processor-specific plotting interpolation and denote inter-processor boundaries for this four-processor simulation. A close-up view of the upper mantle in the upper right-hand panel for the s component at two different times reveals some of the reflections off the surface and discontinuities such as depth phases (pS , pP) and surface multiples (SS). Dotted lines are discontinuities, and the actual mesh is plotted in the background, including a coarsening level just below the crust. No apparent unphysical waveform distortion occurs, and any numerical noise is suppressed by the input Gaussian source time function of appropriate resolution. Finally, the bottom panel is a zoom into the CMB region (straight line, D' discontinuity: dotted line), including a processor boundary. The z -component wavefield ex-

hibits various reflections (PcP , PcS , ScS), converted transmissions (SK), refractions and diffractions (P_{diff}) as expected from this region. The velocity contrast across the CMB for P waves is evident by the wave front kinks for PKP , and the smooth passing for converted waves such as SK verifies that the velocity contrast for S -to- P conversions is small. The passing of the fluid-core coarsening and the crossing of the processor boundary leave, as expected, no artificial noise or amplitude deprecations. Clearly, these snapshots contain no quantitative control over the numerical scheme; they mainly serve as a conceptually accessible view on propagation details, and may furthermore, due to their ease in plotting in 2-D, be a useful tool to assess high-frequency 3-D wavefield behaviour anywhere in the (spherically symmetric) earth.

4.3 Energy conservation

We now examine the total energy conservation, which for purely elastic-wave propagation within enclosed domains and especially with various time schemes is another tangent of validation: Dissipation over time may then stem from either non-conserving time schemes (e.g. Newmark) or leaking (and hence incorrect) boundary and discontinuity conditions (such as the free surface, the axial, masked boundary, or solid–fluid interface terms). With our case of both essential and natural boundary conditions, and a new potential formulation resulting in material-independent boundary terms in addition to two different types of time schemes, it is indeed conceivable

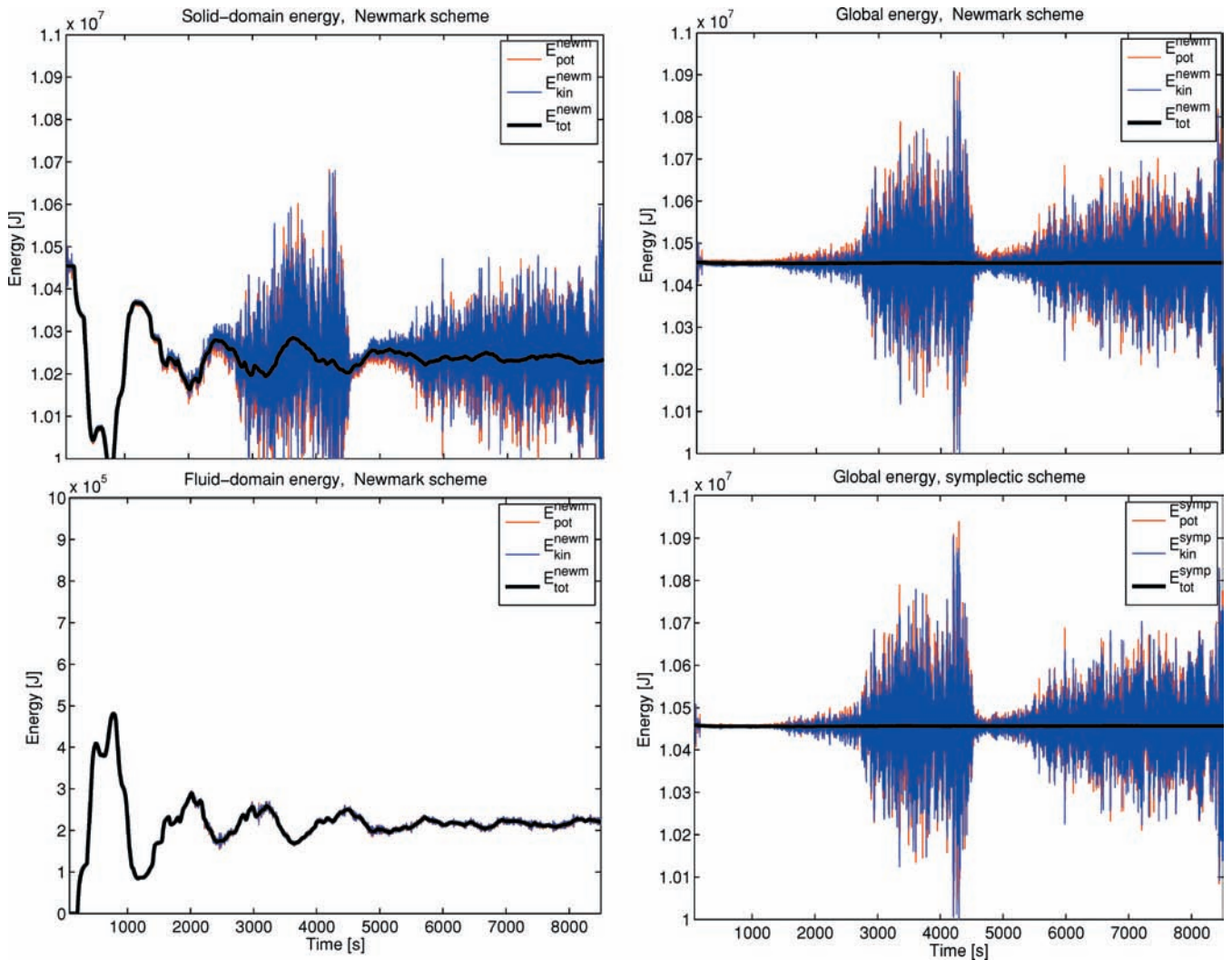


Figure 11. Kinetic, potential and total energy of seismic waves in a solid–fluid sphere for a M_{zz} source with $T_0 = 20$ s at 650 km depth. Left-hand panels: Energy in the solid (top panel) and fluid (bottom panel) domains, respectively. Right-hand panel: Global energy for the Newmark (top panel) and the symplectic (bottom panel) time schemes.

to address total energy conservation. The kinetic energy at time t is given by

$$E_{\text{kin}} = \frac{1}{2} \int_{\oplus} \rho \dot{\mathbf{u}}^2 d^3 \mathbf{x} \approx \frac{1}{2} \dot{\mathbf{u}}^T \mathbf{M}_s \dot{\mathbf{u}} + \frac{1}{2} \dot{\chi}^T \mathbf{K}_f \dot{\chi}, \quad (40)$$

where we have used the fluid potential definition given by eq. (11). The stored, potential or elastic energy in an isotropic, non-gravitating earth at time t is

$$E_{\text{pot}} = \frac{1}{2} \int_{\oplus} \left\{ \lambda [\nabla \cdot \mathbf{u}]^2 + \mu \nabla \mathbf{u} : [\nabla \mathbf{u} + (\nabla \mathbf{u})^T] \right\} d^3 \mathbf{x} \\ \approx \frac{1}{2} \mathbf{u}^T \mathbf{K}_s \mathbf{u} + \frac{1}{2} \chi^T \mathbf{M}_f \chi, \quad (41)$$

where we have invoked both the potential formulation eq. (11) and the fluid-domain wave equation eq. (12). Note that the expressions for the fluid potential (i.e. the respectively second terms) are given by the opposite actions of stiffness and mass matrices compared to the solid displacement terms. Fig. 11 shows some examples for the kinetic, potential, and total energy of the solid (top left-hand panel), fluid (bottom left-hand panel), and global domains (right-

hand panel), for the Newmark (top right-hand panel) and the symplectic scheme (bottom right-hand panel) using a M_{zz} source at 650 km depth with $T_0 = 20$ s. While the breakdown of energy content into fluid and solid domains bears no inference on the accuracy of the method, it is nevertheless an added ‘gift’ of the calculation and included mainly as a curious divergence. Note the sudden onset upon the entry of the P wave into the fluid core at 200 s. The prominent peak in the fluid domain between 500 and 1000 s is related to the propagation of the direct P and converted SK waves, whereas later peaks refer to reflected phases. The total energy inside the FOC (and conversely, the solid regions) oscillates with a recurrence period of about 1000 s, and converges to a constant value (about 2.2×10^5 J) for later times. Both time schemes are consistent in that the total energy is conserved, and only exhibits negligible fluctuations at later times which are related to the globe-circling surface waves and multiples that converge at the antipode at those times. The differences between the time schemes are too small for these settings to differentiate them in any illustrative manner (see top and bottom right-hand panels in Fig. 11).

5 CONCLUSIONS

This paper presents the constituents that are necessary to discretize a spherically symmetric earth model and its accompanying elastodynamic weak form using a spectral-element method upon a 2-D semi-disk after collapsing the third dimension by exploiting symmetries in the radiation pattern. The meshing process is described in detail to highlight its flexibility with regard to arbitrary background models and resolutions. The central region is discretized via a rotated-coordinate polynomial that minimizes grid spacing variations. We quantify various numerically important parameters such as the *characteristic lead time* and the *local oversampling ratio* in a non-dimensionalized fashion, allowing for easy adaption and comparison to virtually any other meshing procedures in similar domains. Meshing determines the numerical cost of the simulation inasmuch as it defines the grid spacing and total number of elements (*instantaneous cost*) along with the time step (*extrapolation cost*). Again, we quantify the cost non-dimensionally to remain independent of actual resolutions and transparent to other schemes. The meshing approach is relatively cheaper with increasing resolution, that is, primarily depends on the locations of discontinuities which in turn hamper the efficiency at low resolution but converge to a steady cost for resolutions that exceed the geometric constraints imposed by discontinuities. Arbitrary resolutions force us to parallelize the method even in its 2-D world. We take advantage of the D-shaped domain and decompose it into exactly load-balanced cake-piece processor domains which contain maximally two neighbours, that is, the ideal minimization of the number of network messages. The method scales extremely well, and at moderate resolutions is largely blind to the network traffic introduced by the parallelization (Nissen-Meyer 2007). We anticipate this benevolent ignorance to be even enhanced when extensive I/O of time–space wavefields shall take up significant amounts of CPU time.

The solid–fluid system is discretized closely following the previous solid-only incarnation of the method, using a displacement-potential formulation for the fluid that differs from other spectral-element methods by the inclusion of the density, resulting in solid–fluid boundary terms that are independent of material properties and identical on fluid and solid sides in their discretized form. The synthetics expose excellent overlaps with normal-mode summation for all distances and phases. We also implemented a symplectic fourth-order time extrapolation scheme which proves useful for certain applications needing high accuracy over large propagation distances. The total energy of the method is conserved, and wavefield snapshots near the CMB illustrate the lack of any spurious energy.

This method is to be considered complete for non-gravitating SNREI earth models. Future developments on these grounds shall entail transverse isotropy, gravity, viscoelasticity, oceans, and rotation, but with respect to seismic sensitivity at the global scale, it is now worthwhile to compute Green functions for the purpose of full-wave based finite-frequency inversions.

ACKNOWLEDGMENTS

We are grateful to Jean-Paul Ampuero for his remarks on an earlier draft of the paper and initiating the work on symplectic time integration. Reviews by Heiner Igel, Dimitri Komatitsch and GJI editor Yanick Ricard improved the manuscript, and discussions with Jeroen Tromp, Guust Nolet, and the Princeton seismology group have been beneficial for this work. TNM thanks the Laboratoire de Géophysique Interne et Tectonophysique, Grenoble, and the

Seismolab at California Institute of Technology, Pasadena, for their hospitality during his stays there. Financial support has been provided by the US National Science Foundation under Grants EAR-0105387 and EAR-0113050 and the Harold W. Dodds Honorific Fellowship by the Graduate School at Princeton. Simulations were undertaken at Princeton's departmental cluster; open-source software through CIG (geodynamics.org) and SPICE (spice-rtn.org) is greatly appreciated. TNM and AF completed the last bits of this paper upon Tony Dahlen's sad passing in June 2007; his active participation, thoughts and comments on this work continued until the week before his untimely departure.

REFERENCES

- Ampuero, J.-P. & Nissen-Meyer, T., 2008. High-order conservative time schemes in spectral-element methods for seismic wave propagation, *Geophys. J. Int.*, submitted.
- Chaljub, E., 2000. Modélisation numérique de la propagation d'ondes sismiques en géométrie sphérique: application à la sismologie globale, *PhD thesis*, Université de Paris 7.
- Chaljub, E. & Tarantola, A., 1997. Sensitivity of SS precursors to topography on the upper-mantle 660 km discontinuity, *Geophys. Res. Lett.*, **24**, 2613–2616.
- Chaljub, E. & Valette, B., 2004. Spectral element modeling of three dimensional wave propagation in a self-gravitating Earth with an arbitrarily stratified outer core, *Geophys. J. Int.*, **158**, 131–141.
- Chaljub, E., Capdeville, Y. & Vilotte, J.-P., 2003. Solving elastodynamics in a fluid-solid heterogeneous sphere: a parallel spectral element approximation on non-conforming grids, *J. Comp. Phys.*, **187**, 457–491.
- Channell, P.J. & Scovel, C., 1990. Symplectic integration of Hamiltonian systems, *Nonlinearity*, **3**, 231–259.
- Chin, S.A., 2006. Complete characterization of fourth-order symplectic integrators with extended-linear coefficients, *Phys. Rev. E*, **73**, 026705.
- Cummins, P., Geller, R. & Takeuchi, N., 1994. DSM complete synthetic seismograms: P-SV, spherically symmetric, case, *Geophys. Res. Lett.*, **21**, 1663–1666.
- Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton University Press, Princeton.
- Deville, M.O., Fischer, P.F. & Mund, E.H., 2002. *High-Order Methods for Incompressible Fluid Flow*, Vol. 2 of Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge.
- Dixon, M. & Reich, S., 2004. Symplectic time-stepping for particle methods, *GAMM Mitteilungen*, **27**, 9–24.
- Fournier, A., Bunge, H.-P., Hollerbach, R. & Vilotte, J.-P., 2004. Application of the spectral element method to the axisymmetric Navier-Stokes equation, *Geophys. J. Int.*, **156**, 682–700.
- Friederich, W. & Dalkolmo, J., 1995. Complete synthetic seismograms for a spherically symmetric Earth by a numerical computation of the Green's function in the frequency domain, *Geophys. J. Int.*, **122**, 537–550.
- Geller, R.J. & Ohminato, T., 1994. Computation of synthetic seismograms and their partial derivatives for heterogeneous media with arbitrary natural boundary conditions using the Direct Solution Method, *Geophys. J. Int.*, **116**, 421–446.
- Igel, H. & Weber, M., 1995. SH-wave propagation in the whole mantle using high-order finite differences, *Geophys. Res. Lett.*, **22**, 731–734.
- Igel, H. & Weber, M., 1996. P-SV wave propagation in the Earth's mantle using finite differences: application to heterogeneous lowermost mantle structure, *Geophys. Res. Lett.*, **23**, 415–418.
- Käser, M. & Dumbser, M., 2006. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes – I. The two-dimensional isotropic case, with external source terms, *Geophys. J. Int.*, **166**, 855–877.
- Käser, M. & Igel, H., 2001. Numerical simulation of 2D wave propagation on unstructured grids using explicit differential operators, *Geophys. Prospect.*, **49**, 607–619.

- Kawai, K., Takeuchi, N. & Geller, R.J., 2006. Complete synthetic seismograms up to 2 Hz for transversely isotropic spherically symmetric media, *Geophys. J. Int.*, **164**, 411–424.
- Komatitsch, D. & Tromp, J., 2002. Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.*, **149**, 390–412.
- Komatitsch, D. & Vilotte, J.P., 1998. The spectral-element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. seism. Soc. Am.*, **88**, 368–392.
- Lognonné, P. & Romanowicz, B., 1990. Modeling of coupled normal modes of the Earth: the spectral method, *Geophys. J. Int.*, **102**, 365–395.
- McLachlan, R.I., 2002. Families of high-order composition methods, *Numer. Algorithms*, **31**, 233–246.
- McLachlan, R.I. & Quispel, G.R.W., 2006. Geometric integration of ODEs, *J. Phys. A*, **A39**, 5251–5286.
- Mercerat, E.D., Vilotte, J.-P. & Sanchez-Sesma, F.J., 2006. Triangular spectral element simulation of two-dimensional elastic wave propagation using unstructured triangular grids, *Geophys. J. Int.*, **166**, 679–698.
- Nissen-Meyer, T., 2007. Full-wave seismic sensitivity in a spherical Earth, *PhD thesis*, Princeton University.
- Nissen-Meyer, T., Dahlen, F.A. & Fournier, A., 2007a. Spherical-earth Fréchet sensitivity kernels, *Geophys. J. Int.*, **168**, 1051–1066.
- Nissen-Meyer, T., Fournier, A. & Dahlen, F.A., 2007b. A 2-D spectral-element method for computing spherical-earth seismograms—I. Moment-tensor source, *Geophys. J. Int.*, **168**, 1067–1093.
- Omelyan, I.P., Mryglod, I.M. & Folk, R., 2002. Optimized Forest-Ruth- and Suzuki-like algorithms for integration of motion in many-body systems, *Comput. Phys. Commun.*, **146**, 188–202.
- Omelyan, I.P., Mryglod, I.M. & Folk, R., 2003. Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations, *Comput. Phys. Commun.*, **151**, 272–314.
- Thomas, C., Igel, H., Weber, M. & Scherbaum, F., 2000. Acoustic simulation of P-wave propagation in a heterogeneous spherical Earth: numerical method and application to precursor energy to PKPdf, *Geophys. J. Int.*, **141**, 307–320.
- Toyokuni, G. & Takenaka, H., 2006. FDM computation of seismic wavefield for an axisymmetric earth with a moment tensor point source, *Earth Planets Space*, **58**, e29–e32.
- Toyokuni, G., Takenaka, H., Wang, Y. & Kennett, B.L.N., 2005. Quasi-spherical approach for seismic wave modeling in a 2D slice of a global Earth model with lateral heterogeneity, *Geophys. Res. Lett.*, **32**, doi:10.1029/2004GL022180.
- Tufo, H.M. & Fischer, P.F., 2001. Fast parallel direct solvers for coarse grid problems, *J. Par. Dist. Comput.*, **61**, 151–177.
- Yoshida, H., 1990. Construction of higher order symplectic integrators, *Phys. Lett.*, **150**, 262–268.

APPENDIX A: SYMPLECTIC INTEGRATION

Dealing with a linear version of solid–fluid elastodynamics, we can readily extend concepts from celestial mechanics, molecular dynamics, electrodynamics, accelerator physics, statistical or quantum wave mechanics to our specific case, drawing similarities within the ubiquitous kinetic, potential and source terms. In this section, we will summarize and outline the symplectic integrator scheme as used in this paper. *Symplectic* refers to systems in which the phase-space volume is preserved such that temporal motions of phase-space points $\mathcal{Y} = (q, p)$ (q : displacement, p : momentum) from 0 to t preserve the Poincaré invariants $\int \int \sum_i dq_i dp_i$ and $\int \int \int \sum_{i \neq j} dq_i dp_i dq_j dp_j$ (i.e. encapsulating Liouville’s Theorem $\frac{d\mathcal{Y}}{dt} = 0$). All topics revolving around cost-effective temporal integration of this kind are vast, active research areas; here we denote some simple concepts to justify the usage of the scheme for our settings. See for example, (Omelyan *et al.* 2002; Channell &

Scovel 1990; McLachlan & Quispel 2006; Dixon & Reich 2004; Chin 2006; McLachlan 2002) for more in-depth information.

A1 Hamiltonian dynamics & conservation of geometric structure

Let us start by rewriting the temporal system of ODEs eqs (20)–(22) as

$$\partial_t \dot{\chi} = \mathbf{M}_f^{-1} \Xi_f(\mathbf{u}, \chi), \quad \partial_t \dot{\mathbf{u}} = \mathbf{M}_s^{-1} \Xi_s(\mathbf{u}, \chi), \quad (\text{A1})$$

where $\Xi_f(\mathbf{u}, \chi) = -\mathbf{K}_f \chi + \mathbf{B}_f \mathbf{u}$ and

$$\Xi_s(\mathbf{u}, \chi) = -\mathbf{K}_s \mathbf{u} - \mathbf{B}_s \mathbf{M}_f^{-1} (-\mathbf{K}_f \chi + \mathbf{B}_f \mathbf{u}) + \mathbf{f}.$$

To utilize symplectic integrators, we need to examine eqs (A1) in terms of Hamilton’s canonical equations of motion

$$\dot{\mathbf{q}} = \partial_{\mathbf{p}} H, \quad \dot{\mathbf{p}} = -\partial_{\mathbf{q}} H, \quad (\text{A2})$$

where the Hamiltonian (energy function) takes the separable form $H(\mathbf{q}, \mathbf{p}) = H_1(\mathbf{p}) + H_2(\mathbf{q})$. Let now $\mathbf{q} = \mathbf{u}$, $\mathbf{v} = \dot{\mathbf{u}}$, $\Xi = \Xi_s$ in the solid or $\mathbf{q} = \chi$, $\mathbf{v} = \dot{\chi}$, $\Xi = \Xi_f$ in the fluid, then, considering our kinetic and potential energy in the discretized form (e.g. Nissen-Meyer *et al.* 2007b), respectively,

$$E_k(\mathbf{v}) = \frac{1}{2} \mathbf{v}^T \mathbf{M} \mathbf{v}, \quad E_p(\mathbf{q}) = \frac{1}{2} \mathbf{q}^T \mathbf{K} \mathbf{q} + \frac{1}{2} \mathbf{q}^T \mathbf{B} \mathbf{q}, \quad (\text{A3})$$

we obtain

$$\mathbf{v} = \mathbf{M}^{-1} \mathbf{p}, \quad \dot{\mathbf{p}} = -\Xi(\mathbf{q}), \quad (\text{A4})$$

where $\mathbf{p} = \partial_{\mathbf{v}} H = \mathbf{M} \mathbf{v}$.

We can now express the equations of motion as

$$\partial_t \mathcal{Y} = \{\mathcal{Y}, H(\mathcal{Y})\} = \partial_{\mathbf{q}} \mathcal{Y} \partial_{\mathbf{p}} H - \partial_{\mathbf{p}} \mathcal{Y} \partial_{\mathbf{q}} H = \hat{L}_H \mathcal{Y}, \quad (\text{A5})$$

where the concatenated full set of phase variables $\mathcal{Y} = (\mathbf{q}, \mathbf{v})$ comprises the canonical, global coordinates, $\{\cdot, \cdot\}$ is the Poisson bracket and $\hat{L}_H = \{\cdot, H\} = \partial_{\mathbf{p}} E_k \partial_{\mathbf{q}} + \partial_{\mathbf{p}} E_p \partial_{\mathbf{p}} = \mathbf{v} \cdot \partial_{\mathbf{q}} - \Xi(\mathbf{q}) \cdot \partial_{\mathbf{p}}$ is the Liouville operator. Due to the conservation of such global geometric properties as energy, time-reversal symmetry or (angular) momentum, any discretization based upon representing the Hamiltonian such as symplectic schemes belongs to the class of *geometric integration*. Note that discretization inevitably introduces energy dissipation; the symplectic approach however lacks any long-term drifts and is only subject to small periodic errors.

A2 Splitting/decomposition schemes

Systems of the form $\partial_t \mathcal{Y}(t) = \hat{L}_H \mathcal{Y}(t) = [\hat{L}_k(t) + \hat{L}_p] \mathcal{Y}(t)$ have a unique solution at time t after $n_t = t/\Delta t$ integrations

$$\mathcal{Y}(t) = \left[\exp \left((\hat{L}_k + \hat{L}_p) \Delta t \right) \right]^{n_t} \mathcal{Y}(0), \quad (\text{A6})$$

if initial condition $\mathcal{Y}(0)$ is known. We now decompose and approximate the ‘total propagator’ to order K ,

$$e^{(\hat{L}_k + \hat{L}_p) \Delta t + \mathcal{O}(\Delta t^{K+1})} = \prod_{j=1}^J e^{\hat{L}_k \kappa_j \Delta t} e^{\hat{L}_p \pi_j \Delta t}. \quad (\text{A7})$$

The splitting into the ‘subpropagator’ terms $e^{\hat{L}_k \kappa_j \Delta t}$ and $e^{\hat{L}_p \pi_j \Delta t}$ in eq. (A7) leads to analytical solutions:

$$e^{\hat{L}_k \kappa_j \Delta t} (\mathbf{q}, \mathbf{v}) = e^{\mathbf{v} \cdot \partial_{\mathbf{q}} \kappa_j \Delta t} = (\mathbf{q} + \kappa_j \Delta t \mathbf{v}, \mathbf{v}), \quad (\text{A8})$$

$$e^{\hat{L}_p \pi_j \Delta t} (\mathbf{q}, \mathbf{v}) = e^{\Xi \cdot \partial_{\mathbf{p}} \pi_j \Delta t} = (\mathbf{q}, \mathbf{v} + \pi_j \Delta t \Xi). \quad (\text{A9})$$

These resultant coordinate translations outline the preservation of the phase-space volume, that is, symplectic structure. Note that $J \geq K$ controls the ‘cost’ of the scheme by virtue of the number of force Ξ evaluations per time step. For any given order of accuracy K , we now need to determine the factorization coefficients κ_j and π_j , $j = 1, \dots, J$. At first, we shall appropriate constraints to reproduce self-adjoint time reversibility of solutions $S^{-1}(t) = S(-t)$ by imposing symmetry conditions on the respective subpropagators:

$$\kappa_1 = 0, \kappa_{j+1} = \kappa_{J-j+1}, \pi_j = \pi_{J-j+1}, \text{ or} \quad (\text{A10})$$

$$\kappa_j = \kappa_{J-j+1}, \pi_j = \pi_{J-j}, \pi_J = 0, \quad (\text{A11})$$

where the former version leads to so-called velocity schemes, and the latter to position schemes. Additionally, this choice cancels all odd-order terms within $\mathcal{O}(\Delta^{K+1})$ automatically such that only even-order schemes may be developed in this context. We henceforth aim at cancelling all even-order terms up to a finite number $k > K$, starting for first-order terms with $\sum_{j=1}^J \kappa_j = \sum_{j=1}^J \pi_j = 1$. The $K = 4$ algorithm we adopted is based upon $J = 5$, resulting in $J - 1 = 4$ force evaluations per time step. Using time-reversibility constraints eq. (A10) and the cancellation of first-order terms, the

propagator then takes the form

$$\begin{aligned} e^{(\hat{L}_k + \hat{L}_p)\Delta t + C_3 \Delta t^3 + C_5 \Delta t^5 + \mathcal{O}(\Delta t^7)} = \\ e^{\hat{L}_p \xi \Delta t} e^{\hat{L}_k (1-2\lambda)\Delta t/2} e^{\hat{L}_p \chi \Delta t} e^{\hat{L}_k \lambda \Delta t} e^{\hat{L}_p [1-2(\chi+\xi)]\Delta t} \\ \times e^{\hat{L}_k \lambda \Delta t} e^{\hat{L}_p \chi \Delta t} e^{\hat{L}_k (1-2\lambda)\Delta t/2} e^{\hat{L}_p \xi \Delta t}, \end{aligned} \quad (\text{A12})$$

where the number of single-exponential operators reflects the number of stages in the resultant scheme, $2J - 1 = 9$. What remains to be done to arrive at an explicit scheme is to determine the optimal choice of coefficients ξ, χ, λ , which is done by expanding both sides of eq. (A7) into Taylor series and collecting terms with the same powers of Δt to obtain a set of non-linear algebraic equations for κ_j, π_j . This can readily be solved for low orders, but becomes elaborate even for $K = 4$ such as eq. (A12). Methods to determine optimal factorization coefficients can be found in Yoshida (1990), Chin (2006), Omelyan *et al.* (2003). Our choice of coefficients in Section 3.4.2 follows the latter article, table 2 algorithm No. 20. Optimal coefficients depend on the definition of the cost function to minimize, and may be specified for each Hamiltonian separately. In this regard, our choice of coefficients and overall scheme is not necessarily optimal.