

Efficient 3D elastic FWI using a spectral-element method on Cartesian-based mesh

P.T. Trinh^{1,2}, R. Brossier², L. Métivier^{2,3}, L. Tavar^{2,4}, J. Virieux², P. Wellington²

¹ Total EP, ² Univ. Grenoble Alpes, ISTerre, ³ Univ. Grenoble Alpes, CNRS, LJK ⁴ Univ. Grenoble Alpes, GRICAD

SUMMARY

Full Waveform Inversion offers the possibility to extract high-resolution quantitative multi-parameters models of the subsurface from seismic data. Heretofore, most of FWI applications at the crustal scale have been performed under the acoustic approximation, generally for marine environments. When considering challenging land problems, efficient strategies are required for moving toward elastic inversion. We present such approach for 3D elastic time-domain inversion, based on spectral element methods designed on cartesian-based meshes. The proposed workflow integrates an easy and accurate cartesian-based mesh building with high-order shape functions to capture rapid topography variations and an efficient workflow for the incident and adjoint fields computation. A nonstationary and anisotropic structure-oriented smoothing filter is implemented directly on the spectral element mesh, for preconditioning FWI by incorporating prior geological information such as coherent lengths, dip and azimuth angles. Numerical illustrations on Marmousi and SEAM II benchmarks illustrate the importance of each ingredient we have developed for making efficient and flexible elastic FWI for land applications.

INTRODUCTION

High-resolution quantitative multi-parameters models of the subsurface are essential for crustal exploration. By considering the entire information contained in seismic data, full waveform inversion (FWI) (Virieux and Operto, 2009) offers the possibility to extract such models. While most of FWI applications at the crustal scale have been performed in the acoustic approximation for a decade, mainly for marine-acquired data, it now becomes mandatory to tackle challenging land targets. The complex geology and possibly complex topography, found in many land environments, require to consider elastic effects, and therefore to develop an accurate and affordable 3D elastic FWI engine. Due to the complexity of the acquired seismic data, multiple frequency components are required to better constrain the inverse problem. Therefore, time-domain approaches are preferred for their ability to apply the time-windowing and to process the data (Brossier et al., 2009).

Moving toward elastic modeling could be done by a natural extension of the finite-difference (FD) methods widely-used for acoustic modeling. However, free-surface and near-surface representation, as well as free-surface effects, can be challenging to model with FD. Conversely, finite element (FE) methods can handle such kind of boundary conditions very accurately. For large scale problems, spectral element methods (SEM) appear to be accurate, efficient and flexible for 3D elastic modeling and FWI (Komatitsch and Tromp, 1999; Fichtner et al., 2008; Tape et al., 2010; Peter et al., 2011). One drawback of conventional SEM is the requirement of using and building hexahedral meshes, which can be a challenging and time-consuming task.

In this work, we present an efficient SEM-based 3D elastic FWI approach (SEM3D code) designed for crustal-scale exploration. This implementation relies on (1) Cartesian-based deformed mesh with high-order shape functions to capture complex topographies; (2) two Message Passing Interface (MPI)-based parallelism levels for tackling large scale and multiple shots experiments, associated with an efficient computation of incident and adjoint fields through optimized computing kernels (Deville et al., 2002); (3) structurally-based nonstationary and anisotropic smoothing filter implemented as a partial differential equation (PDE) solved with SEM on the modeling mesh (Trinh et al., 2017).

WAVE PROPAGATION AND FWI IN SEM3D

A classical hexahedra-based SEM frame is considered for elastic modeling (Komatitsch and Tromp, 1999): the physical domain Ω is decomposed into a set of non-overlapping hexahedral elements. Each element can be mapped to the unitary reference space of Gauss-Lobatto-Legendre (GLL) points, where the cube $[-1, 1] \otimes [-1, 1] \otimes [-1, 1]$ is discretized into a set of $(N+1)^3$ GLL points $(\xi_{k_1}, \eta_{k_2}, \zeta_{k_3})$; $k_1, k_2, k_3 = 0, \dots, N$, where N refers to the interpolation order. These collocation points define $(N+1)^3$ basis functions, which are triple products of Lagrange polynomials of degree N . Considering this choice of basis functions and the GLL quadrature for numerical integration, the weak form of the second-order PDE governing the elastic waves propagation can be written as:

$$\mathbf{M} \partial_{tt} u = -\mathbf{K} u + \mathbf{F}, \quad (1)$$

where the displacement field is denoted by u , the global mass and stiffness matrices by \mathbf{M} and \mathbf{K} , respectively, and the source term by \mathbf{F} . The global mass matrix \mathbf{M} is diagonal by construction. The free-surface condition is naturally taken into account by the weak formulation. A second order explicit Newmark scheme is implemented for the time integration (Komatitsch, 1997).

The inversion problem relies on a classical least-squares norm given by

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \|\mathbf{d}_{\text{obs}} - \mathbf{d}_{\text{cal}}\|^2, \quad (2)$$

which computes the L_2 distance between the recorded seismic data \mathbf{d}_{obs} and the modeled seismic data \mathbf{d}_{cal} . In the time domain, the gradient of $\mathcal{C}(\mathbf{m})$ with respect to the elastic tensor coefficients C_{ij} can be computed through the adjoint-state approach (Plessix, 2006; Vigh et al., 2014)

$$\mathbf{g}(\mathbf{x}) = \frac{\partial \mathcal{C}(\mathbf{m})}{\partial C_{ij}} = \left(\bar{\boldsymbol{\varepsilon}}, \frac{\partial C}{\partial C_{ij}} \boldsymbol{\varepsilon} \right)_{\Omega, t}, \quad (3)$$

where $\bar{\boldsymbol{\varepsilon}}$ and $\boldsymbol{\varepsilon}$ are respectively adjoint and incident strain fields. The matrix $C = (C_{ij})_{6 \times 6}$ contains the elastic tensor coefficients, with 21 independent components in the case of full anisotropy. The gradient for any parameter α (seismic velocity, anisotropic parameter, impedance, ...) can then be computed

by chain rule using the density ρ and C_{ij} elementary gradient

$$\frac{\partial \mathcal{C}}{\partial \alpha} = \sum_{i=1}^6 \sum_{j=i}^6 \frac{\partial \mathcal{C}}{\partial C_{ij}} \frac{\partial C_{ij}}{\partial \alpha} + \frac{\partial \mathcal{C}}{\partial \rho} \frac{\partial \rho}{\partial \alpha}. \quad (4)$$

CARTESIAN-BASED DEFORMED MESH

To combine the accurate representation of topography, allowed by FE meshes, and the easiness of implementation of FD grid, our SEM3D package considers a Cartesian-based mesh, with vertically deformed elements. The numbers of elements in x , y and z directions are constant. For the interpolation at order $N = 4$ or 5 , SEM allows to accurately model elastic waves propagation with around 5 GLL nodes per shortest wavelength (Komatitsch, 1997). This condition is referred as the *volume condition*.

When considering the presence of significant topography variation, hexahedral elements can be vertically deformed. For each element, a set of $(n + 1)$ control points in each direction is considered, leading to $(n + 1)^3$ control points and associated shape functions in 3D. These shape functions are triple products of Lagrange polynomials of degree n . The number of control points and shape functions $(n + 1)$ is not related to the interpolation order N of the test functions needed for solving the PDE.

Representing the surface with P_1 shape functions (linear functions with $n = 1$) leads to the use of the eight corners of the element as control points. Such simple representation cannot honor sharp spatial variation of the free surface, as shown in one example in Figure 1A. The rough P_1 approximation of the topography affect the accuracy of the simulation due to the interaction between elastic waves and the complex surface. Decreasing the element size is one way for following the rapid variation of the topography, namely the *surface condition*. This criterion might be stricter than the *volume condition*, and would significantly increase the computational cost.

This *surface condition* limitation can be overcome by P_n shape functions at the arbitrary order n , where the control points are $(n + 1)^3$ GLL points inside the element:

$$\mathbf{x}(\xi, \eta, \zeta) = \sum_{k_1=1}^{n+1} \sum_{k_2=1}^{n+1} \sum_{k_3=1}^{n+1} \ell_{\hat{k}}(\xi, \eta, \zeta) \mathbf{x}_{\hat{k}}, \quad (5)$$

where \hat{k} stands for the triple indexes k_1, k_2, k_3 . The associated shape function is a triple product of Lagrange polynomials of degree n : $\ell_{\hat{k}}(\xi, \eta, \zeta)$. Figure 1B highlights that with the same size of the element (at 100 m), the P_4 shape function provides a better representation of the complex topography (i.e. $(4 + 1)^2$ GLL points are used in each element to capture the topography map, instead of $(1 + 1)^2$ points for the P_1 case).

It should be noticed that only the volumetric Jacobian matrix associated with the mapping from the reference space to the Cartesian space is required for the wave propagation (together with the surface Jacobian for the radiative absorbing boundary condition). The mesh creation with P_n shape functions only affects the mesh construction and the computation of the Jacobian, which are computed only once in the FWI workflow. The computational cost of the wavefield modeling is unaltered,

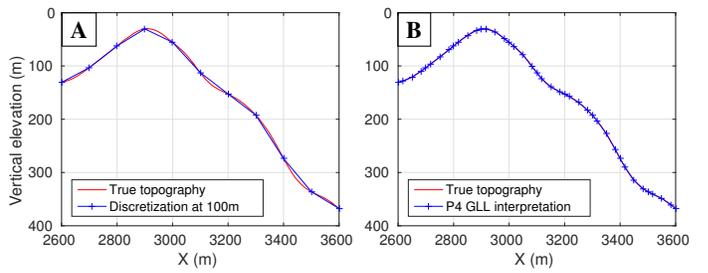


Figure 1: Topography description of a 2D cross-section (extracted from SEAM II model) using (A) the eight corners of each element and P_1 shape function and (B) the $(4 + 1)^3$ GLL control points associated with P_4 shape functions. The element size is 100m in both case.

while the simulation accuracy related to the complex wave-phenomena at the free-surface is significantly improved.

OPTIMIZED ARCHITECTURE

Modeling kernel

The SEM implementation used in our workflow is based on limited interpolation orders for test functions with $N = 4$ or 5 . It has been shown that these orders provide a good compromise between the numerical accuracy and the constraint on the CFL stability condition (Komatitsch, 1997). The key part of the modeling kernel is the computation of the stiffness-displacement matrix-vector product $\mathbf{K}u$. Our implementation benefits from the factorization of the stiffness matrix as

$$\mathbf{K} = \mathcal{D}^W \mathbf{C} \mathcal{D}, \quad (6)$$

where the operator \mathcal{D} estimates the spatial derivatives of a vector in the Cartesian space. The operator \mathcal{D}^W is equivalent to a weighted spatial derivatives operator. The application of these operators on a vector can be decomposed into two steps: the estimation of the spatial derivatives in the reference space and the projection back to the real space. The former step in the reference space can be estimated by using highly efficient algorithms developed by Deville et al. (2002), which take benefit from the tensorial properties of hexahedral elements, the optimization of cache usage, and the combination of efficient loop vectorization and manual unrolling. Similar strategies can be applied to accelerate the computation of the volumetric Jacobian associated with the P_n shape functions, when necessary.

Parallel implementation

Our implementation relies on a two-level MPI-based parallelization: one level is designed on Cartesian-based domain decomposition, allowing an efficient load-balancing thanks to the Cartesian-based mesh. This avoids the use of a third-party mesh-partitioner, even when the number of possible subdomains is constrained by the mesh split in each direction. The second MPI-level is over seismic shots (or “super-shots”) managed in parallel.

Inversion kernel

The inversion kernel relies on the reverse-communication interface provided by the SEISCOPE optimization toolbox (Métivier and Brossier, 2016), which includes various non-linear optimization methods. The FWI gradient, required as the input of the optimization process, is computed by the zero-lag cross-correlation of the incident and adjoint wavefields in the time-domain. The incident field is recomputed by the backward propagation in time from the stored wavefield in the boundaries,

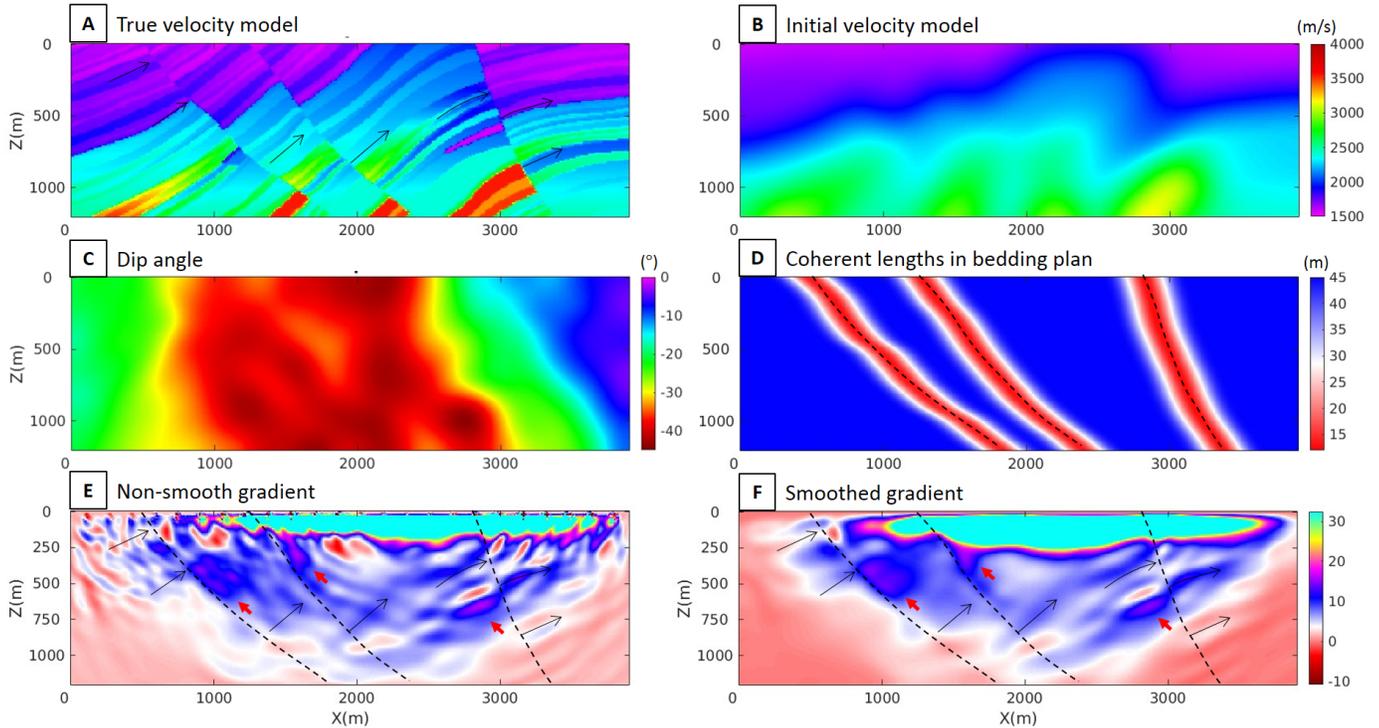


Figure 2: Example of the nonstationary filtering operator on FWI gradient, from pseudo-3D Marmousi model. (A) True velocity model. (B) Initial velocity model. (C) Dip field, estimated from the true velocity model. (D) Coherent lengths in bedding plan (L_u and L_w), which vary from 12 m at fault location to 45 m at other places. (E) Original scaled gradient without any smoothing. (F) Smoothed gradient with anisotropic nonstationary Laplace filter (approximated by the application of two Bessel filters): dip field as presented in Figure C, $L_u = L_w$ as presented in Figure D, and $L_v = 12$ m (≈ 0.15 of the shortest wavelength). Some interesting features are highlighted by black and red arrows, and faults are indicated by black dash-lines.

synchronously with the forward propagation of the adjoint field. As the gradient for the elastic tensor coefficient C_{ij} (Equation 3) involves the strain field, the gradient is directly accumulated during the simultaneous back and forth computation of incident and adjoint fields, resulting in a cheap operation (Dussaud et al., 2008).

BESSEL SMOOTHING FILTER

In many application, the gradient vector $\mathbf{g}(\mathbf{x})$ can exhibit artificial high wavenumber components, incompatible with the intrinsic resolution of FWI. Designing a non-stationary, anisotropic smoothing operator which can incorporate some prior knowledge of the geological structure, such as the local 3D rotation, becomes mandatory for practical applications. Such a filter has to be efficiently applied to the vector of interest.

To fulfill those requirements, we introduce the Bessel filter $B_{3D}(\mathbf{x})$, which can be directly and efficiently implemented with any FD or FE method (Trinh et al., 2017). Instead of convolving the original vector $\mathbf{g}(\mathbf{x})$ with the forward filter $B_{3D}(\mathbf{x})$ to get the smoothed vector $\mathbf{s}(\mathbf{x})$, we solve the following equation relying on the inverse operator

$$B_{3D}^{-1}(\mathbf{x}) * \mathbf{s}(\mathbf{x}) = \mathbf{g}(\mathbf{x}). \quad (7)$$

If the coherent lengths L_x , L_y and L_z in x , y and z directions are uniform over space, Equation (7) can be translated into the related-Bessel PDE

$$\mathbf{s}(\mathbf{x}) - \left(L_z^2 \frac{\partial^2}{\partial z^2} + L_x^2 \frac{\partial^2}{\partial x^2} + L_y^2 \frac{\partial^2}{\partial y^2} \right) \mathbf{s}(\mathbf{x}) = \mathbf{g}(\mathbf{x}), \quad (8)$$

in which the original gradient $\mathbf{g}(\mathbf{x})$ appears in the right hand side. Following the weak formulation of SEM, Equation (8)

naturally yields a symmetric, positive-definite and well-conditioned linear system

$$(\mathbf{M}_b + \mathbf{K}_b) \mathbf{s} = \mathbf{M}_b \mathbf{g}. \quad (9)$$

Similar to the wave propagation problem, the mass matrix \mathbf{M}_b associated with the application of Bessel filter is diagonal and the stiffness matrix \mathbf{K}_b is symmetric by definition.

The anisotropic nonstationary filter is defined by a 3D rotation defined by dip θ and azimuth φ angles and variable coherent lengths: L_v is associated with the direction perpendicular to the local bedding plan, L_u and L_w are related to the planar structure of geological structures. Under the assumption of slow variation of the filter parameters, their spatial derivatives can be neglected. The PDE governing the smoothing process can be approximated as

$$\mathbf{s}(\mathbf{x}) - \nabla_{z,x,y}^t \mathbf{P}(\mathbf{x}) \mathbf{P}^t(\mathbf{x}) \nabla_{z,x,y} \mathbf{s}(\mathbf{x}) = \mathbf{g}(\mathbf{x}), \quad (10)$$

where $\nabla_{z,x,y}$ is the spatial derivatives $(\partial/\partial z, \partial/\partial x, \partial/\partial y)^t$, and the upper symbol “ t ” stands for the transposed operator. The information related to the geological variation of the medium is preserved in the projection matrix

$$\mathbf{P}(\mathbf{x}) = \begin{bmatrix} L_v \cos \varphi & L_u \sin \varphi & 0 \\ -L_v \cos \theta \sin \varphi & L_u \cos \theta \cos \varphi & L_w \sin \theta \\ L_v \sin \theta \sin \varphi & -L_u \sin \theta \cos \varphi & L_w \cos \theta \end{bmatrix}, \quad (11)$$

between the real space and the locally rotated dimensionless coordinates system.

Similar to the system (9) for constant filter parameters, the weak form of Equation (10) yields a symmetric, positive-definite and well-conditioned linear system. We solve this linear system

(9) through a parallel conjugate gradient (CG) iterative solver, using the same high-performance-computing structure as the one for the wave equation. The most expensive operator is the product of the sparse stiffness-matrix \mathbf{K}_b with a given vector. Again, the factorization of this matrix-vector product is used to achieve an efficient implementation (Deville et al., 2002).

A double application of Bessel operators provides an accurate approximation of the Laplace filter. The overall scheme is highly efficient, as the algorithmic complexity is of order $O(L)$, for a given coherent length $L = L_x = L_y = L_z$, compared to the complexity $O(L^3)$ for the 3D explicit convolution approach (Trinh et al., 2017).

Figure 2 illustrates the application of an anisotropic nonstationary Laplace filter (approximated by double application of Bessel filters) on a gradient computed in the Marmousi benchmark. The 2D Marmousi has been extended to an elastic 3D volume for this test. A surface acquisition is used with a line of 24 sources, with distance 160 m between adjacent sources. The receivers are located on the whole surface, with 12.5 m between receivers. A Ricker wavelet centered at 8 Hz is used as the source signal. The 2D cross-section of the gradient without any smoothing filter underneath the source line is shown in Figure 2E. The gradient contains significant acquisition footprint at the near-surface and high wavenumber artifacts in the deeper part. An anisotropic nonstationary Laplace filter with parameters as described in Figures 2C and D is applied to produce the smoothed gradient as shown in Figure 2F. The near-surface acquisition footprint is effectively removed. The continuity of the features at greater depths is enhanced, because the horizontal oscillation artifacts are attenuated, as indicated by the black arrows. Due to the design of the coherent lengths, the energy is not smeared out across the faults, indicated by the red arrows in Figures 2E and F. In this example, the smoothing process costs about 2.2% running time comparing to the forward problem.

FULL WAVEFORM INVERSION EXAMPLE

Figure 3 shows a 3D elastic FWI example on a subset of SEAM Phase II foothills benchmark (Oristaglio, 2012). SEM is used at order $N = 4$ in this test. A surface acquisition is used with a line of 24 sources, each 300 m. Receivers are mapping the free surface each 12.5 m. A Ricker wavelet, centered at 3 Hz is used as the source signal. At this frequency, the *volume condition* implies that the element size should be 100 m. As illustrated in Figure 1A, shape functions of order 1 are not sufficient for such rapid variation of the surface. The P_4 shape functions are thus used, which provide an accurate representation of the topography (Figure 1B).

The 2D cross-section of the shear wave velocity (V_s) model underneath the source line is shown in Figure 3A. The initial model in Figure 3B is a smoothed version of the true model. Similar smoothed model is used for P wave and density as the input of the inversion process. Both surface waves and body waves are used for the inversion without any distinction. We invert for compressional and shear velocities, but only show the result for V_s in Figure 3C. We do not invert for density in this example. The inversion process consists of 60 iterations, with l -BFGS optimization method. Only a stationary Laplace

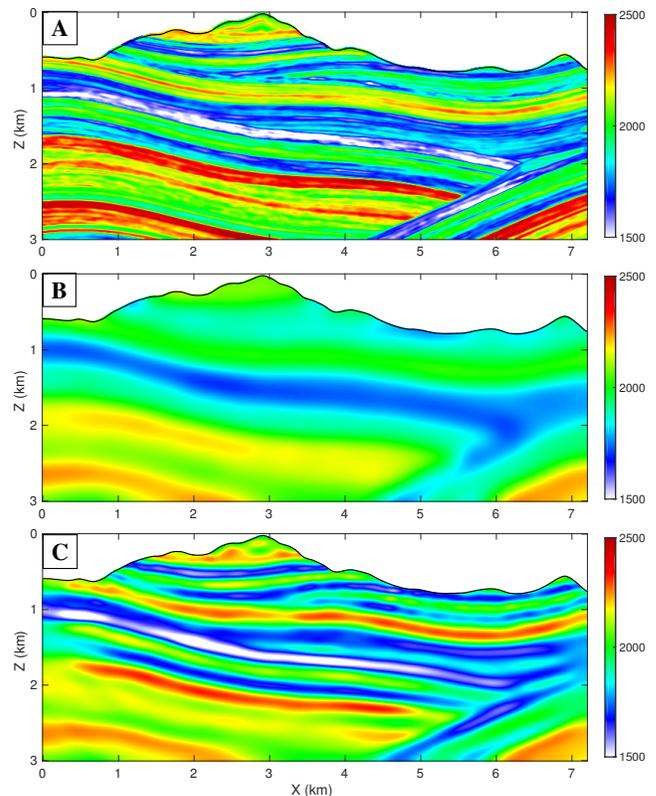


Figure 3: (A) True V_s model. (B) Initial V_s model. (C) Inverted V_s model.

smoothing filter with $L_z = 25$ m, $L_x = L_y = 75$ m is applied to the gradient to remove artifacts beyond the FWI intrinsic resolution. The inversion successfully recovers details in the V_s model, even with the presence of a complex topography. No acquisition footprint or high wavenumber artifact appears in the final model, thanks to the effectiveness of the smoothing filter.

CONCLUSIONS AND PERSPECTIVES

We present an integrated workflow capable of efficiently performing 3D elastic time-domain FWI for multi-parameters on a hexahedral mesh, based on a spectral-element method in 3D. The scheme relies on the use of deformed Cartesian-based mesh and high-order shape functions, simultaneous forward and adjoint fields for gradient computation coupled with a highly optimized kernel, MPI-based parallelism and a novel nonstationary and anisotropic structure-oriented smoothing filter solved efficiently and directly by SEM. Perspectives include the attenuation modeling, the coupling with an optimal wavefield decimation and reconstruction proposed by Yang et al. (2016a,b), and applications to large-scale complex land targets.

ACKNOWLEDGMENTS

The authors would like to thank Total E&P for financial support of PTT's PhD project, and for allowing to present and access the SEAM II model. This study was partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>). This study was granted access to the HPC resources of the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>) and CINES/IDRIS (allocation 046091 made by GENCI). Authors enjoy discussions with M. Appe, B. Duquet, J.L. Boelle and P. Williamson.

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2017 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

- Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: *Geophysics*, **74**, no. 6, WCC105–WCC118, <http://doi.org/10.1190/1.3215771>.
- Deville, M., P. Fischer, and E. Mund, 2002, High order methods for incompressional fluid flow: Cambridge University Press.
- Dussaud, E., W. W. Symes, P. Williamson, L. Lemaistre, P. Signer, B. Denel, and A. Cherrett, 2008, Computational strategies for reverse-time migration: 78th Annual International Meeting, SEG, Expanded Abstracts, 2267–2271, <http://doi.org/10.1190/1.3059336>.
- Fichtner, A., B. L. N. Kennett, H. Igel, and H. P. Bunge, 2008, Theoretical background for continental- and global-scale full-waveform inversion in the time-frequency domain: *Geophysical Journal International*, **175**, 665–685, <http://doi.org/10.1111/j.1365-246x.2008.03923.x>.
- Komatitsch, D., 1997, Méthodes spectrales et éléments spectraux pour l'équation de l'élastodynamique 2D et 3D en milieu hétérogène: Ph.D. thesis, Institut de Géophysique du Globe de Paris.
- Komatitsch, D., and J. Tromp, 1999, Introduction to the spectral element method for three-dimensional seismic wave propagation: *Geophysical Journal International*, **139**, 806–822, <http://doi.org/10.1046/j.1365-246x.1999.00967.x>.
- Métivier, L., and R. Brossier, 2016, The SEISCOPE optimization toolbox: A large-scale nonlinear optimization library based on reverse communication: *Geophysics*, **81**, no. 2, F11–F25, <http://doi.org/10.1190/geo2015-0031.1>.
- Oristaglio, M., 2012, SEAM phase II — Land seismic challenges: *The Leading Edge*, **31**, 264–266, <http://doi.org/10.1190/1.3694893>.
- Peter, D., D. Komatitsch, Y. Luo, R. Martin, N. Le Goff, E. Casarotti, P. Le Loher, F. Magnoni, Q. Liu, C. Blitz, T. Nissen-Meyer, P. Basini, and J. Tromp, 2011, Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes: *Geophysical Journal International*, **186**, 721–739, <http://doi.org/10.1111/j.1365-246x.2011.05044.x>.
- Plessix, R. E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: *Geophysical Journal International*, **167**, 495–503, <http://doi.org/10.1111/j.1365-246x.2006.02978.x>.
- Tape, C., Q. Liu, A. Maggi, and J. Tromp, 2010, Seismic tomography of the southern California crust based on spectral-element and adjoint methods: *Geophysical Journal International*, **180**, 433–462, <http://doi.org/10.1111/j.1365-246x.2009.04429.x>.
- Trinh, P. T., R. Brossier, L. Métivier, J. Virieux, and P. Wellington, 2017, Bessel smoothing filter for spectral element mesh: *Geophysical Journal International*, **in press**.
- Vigh, D., K. Jiao, D. Watts, and D. Sun, 2014, Elastic full-waveform inversion application using multicomponent measurements of seismic data collection: *Geophysics*, **79**, no. 2, R63–R77, <http://doi.org/10.1190/geo2013-0055.1>.
- Virieux, J., and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC1–WCC26, <http://doi.org/10.1190/1.3238367>.
- Yang, P., R. Brossier, L. Métivier, and J. Virieux, 2016a, Wavefield reconstruction in attenuating media: A checkpointing-assisted reverse-forward simulation method: *Geophysics*, **81**, no. 6, R349–R362, <http://doi.org/10.1190/geo2016-0082.1>.

Yang, P., R. Brossier, and J. Virieux, 2016b, Wavefield reconstruction from significantly decimated boundaries: *Geophysics*, **81**, no. 5, T197–T209, <http://doi.org/10.1190/geo2015-0711.1>