

Large-scale finite-difference and finite-element frequency-domain seismic wave modeling
with multi-level domain-decomposition preconditioner
V. Dolean, P. Jolivet, P.-H. Tournier, L. Combe, S. Operto, S. Rizzo

Summary

The emergence of long-offset sparse stationary-recording surveys carried out with ocean bottom nodes (OBN) makes frequency-domain full waveform inversion (FWI) attractive to manage compact volume of data and perform attenuation imaging. One challenge of frequency-domain FWI is the forward problem, which requires the solution of large and sparse linear systems with multiple right-hand sides. While direct methods are suitable for dense acquisitions and problems involving less than 100 million unknowns, iterative solvers are more suitable for large computational domains covered by sparse OBN surveys. Here, we solve these linear systems with a Krylov subspace method preconditioned with the two-level Optimized Restricted Additive Schwarz (ORAS) domain decomposition preconditioner, the prefix optimized referring to the use of absorbing conditions at the subdomain interfaces. We implement this method with finite differences on uniform grid and finite elements on unstructured tetrahedral meshes. A simulation in a model where the velocity linearly increases with depth allows us to validate the accuracy of the two schemes against an analytical solution while highlighting how their relative cost varies with the band of propagated wavelengths. A simulation in the overthrust model involving up to 2 billions of parameters allows us to tune the method and highlights its scalability.

Introduction

Most of the 3D Full Waveform Inversion (FWI) codes are developed today in the time domain because the time-marching forward modeling engines used to perform reverse time migration of towed-streamer data can be readily used for FWI. The emergence of ultra-long offset sparse stationary-recording acquisitions carried out with ocean bottom nodes (OBN) may put the frequency-domain formulation of FWI back in the spotlight because the inversion can be limited to a few frequencies for such acquisitions and attenuation can be straightforwardly taken into account in the forward and inverse problems. The main challenge of frequency-domain FWI is related to the forward problem (the solution of the time-harmonic wave equation), which requires the solution of large, sparse and ill-conditioned linear systems with multiple right-hand sides for each frequency. Today, direct solver for sparse matrices likely provide the most efficient tools to solve these systems for dense seabed acquisitions and computational domains involving less than 100 millions unknowns (Amestoy et al., 2016). For larger domains covered by sparse node arrays or for high frequency imaging, iterative solvers and domain decomposition preconditioner provide the replacement solution. While an up to date assessment of direct methods for 3D FWI is discussed in a companion study, we focus here on the development of an hybrid direct/iterative solver where the direct solver is used to solve the local problems in the subdomains of the preconditioner with incomplete Cholesky factorization. We develop the solver for a finite-element (FE) discretization on unstructured tetrahedral mesh with Lagrange elements of order 3 and an optimized 27-point finite difference (FD) scheme with adaptive coefficients (Operto et al., 2007; Turkel et al., 2013). The iterative method relies on the Krylov subspace GMRES solver (Saad, 2003) with the Optimized Restricted Additive Schwarz (ORAS) domain decomposition preconditioner (Graham et al., 2017; Bonazzoli et al., 2019). Compared to preconditioners based upon shifted Laplacian (Erlangga, 2008), ORAS is less sensitive to the shift (added attenuation) and can be used without it. In the following, we discuss the pros and cons of the two discretization methods with a numerical example. Then, we review the basic principles of the domain decomposition preconditioner before assessing the strong and weak scalability of the hybrid solver with the 3D SEG/EAGE Overthrust model and the FE discretization on unstructured mesh.

Tetrahedral Finite elements versus 27-point finite-differences

The simplest mathematical model of acoustic wave propagation is the Helmholtz equation

$$(\Delta + k^2(\mathbf{x})) u(\mathbf{x}, \omega) = b(\mathbf{x}, \omega), \text{ in a subsurface domain } \Omega, \quad (1)$$

where u is the monochromatic pressure wavefield, b the source, $k(\mathbf{x}, \omega) = \omega/c(\mathbf{x})$, with ω denoting frequency, $c(\mathbf{x})$ the wavespeed (which is complex valued in viscous media) and $\mathbf{x} = (x, y, z) \in \Omega$. After discretization, eq. (1) can be written in matrix form as

$$\mathbf{A}u = \mathbf{b}. \quad (2)$$

We implement the above equation with absorbing boundary conditions along the vertical and bottom faces of Ω and a homogeneous Dirichlet condition on the pressure along the top face.

We consider two discretizations of (1). The first relies on the 27-point FD stencil in which compact 2nd-order accurate stencils minimize the numerical bandwidth and maximizes the sparsity of \mathbf{A} while reaching a high-order accuracy by mixing consistent mass and stiffness matrices on different (rotated) coordinate systems (Operto et al., 2007). These sparsity and compactness properties are useful to minimize the matrix fill-in induced by a sparse direct solver, which is used to solve the local problems in each subdomains of the preconditioner (see next section). The stiffness and consistent mass matrices are weighted by coefficients that are computed by least-squares minimization of the numerical dispersion and anisotropy. Generally, the same coefficients are used in each row of the matrix. In this study, we implement adaptive coefficients in \mathbf{A} that are matched to the local wavespeed to optimize the accuracy of the stencil in heterogeneous media (Turkel et al., 2013). The second relies on Lagrange finite elements on a tetrahedral mesh Γ of the domain Ω . Compared to the 27-point FD method on uniform Cartesian grid, unstructured meshes in FE methods allow for the adaptation of the size of the elements to the local wavelength (the so-called h -adaptivity) and the conformal representation of complex known boundaries (topography, bathymetry). The drawback of the FE discretization is the higher number of degrees of freedom per element for a given accuracy. A numerical and dispersion analysis of the FE method lead to the conclusion that Polynomials of degree 3 (P3) are necessary to reach a sufficient accuracy for a discretization of four points per wavelength (*ppwl*) (Dolean et al., 2020), which is further supported by the dispersion analysis of Ainsworth and Wajid (2010) (Fig 1). This discretization is typically used for FWI application because it is the coarsest one allowing to sample a heterogeneity of size half a

wavelength. The phase velocity dispersion curves for the 27-point FD stencil compare well with those obtained for the P3 finite elements (Fig. 1). To gain a first-hand understanding of the relative cost and accuracy of the two methods, we perform a simulation in a 3D medium of size $2 \text{ km} \times 4 \text{ km} \times 12 \text{ km}$ where the velocity linearly increases with depth ($c(x, y, z) = c_0 + \alpha \times z$) with $c_0=1 \text{ km/s}$. We validate the numerical solutions against an analytical solution (Kuvshinov and Mulder, 2006) using $\alpha=0.8$ and $\alpha=2$ (Tab. 1). Source is at 1 km depth and frequency is 8 Hz. The grid interval in the FD grid ($h=31.25 \text{ m}$) corresponds to 4 *ppwl*, while the size of the tetrahedral elements is matched to the local wavelength for the FE simulation with a discretization rule of 3 *ppwl*. Fig. 1(a-e) highlights the high accuracy achieved by the two discretizations. The number of degrees of freedom (#*dof*) and the error involved in the two simulations highlight how the discretization method should be selected based on the specifications of the application in terms of size, structural complexity and dynamic of propagated wavelengths λ (Tab. 1).

$\alpha(s^{-1})$	$\lambda_{min}(m)$	$\lambda_{max}(m)$	#dof (M)		ℓ_2 error	
			FD	FE	FD	FE
0.8	125	1200	13	28	0.0079	0.034
2	125	3125	13	16	0.044	0.034

Table 1: FD .vs. FE. α : gradient of the wavespeed. $\lambda_{min,max}(m)$: Min./max. wavelength. #dof: number of degrees of freedom. ℓ_2 error: Least-squares error .wrt. analytical solution.

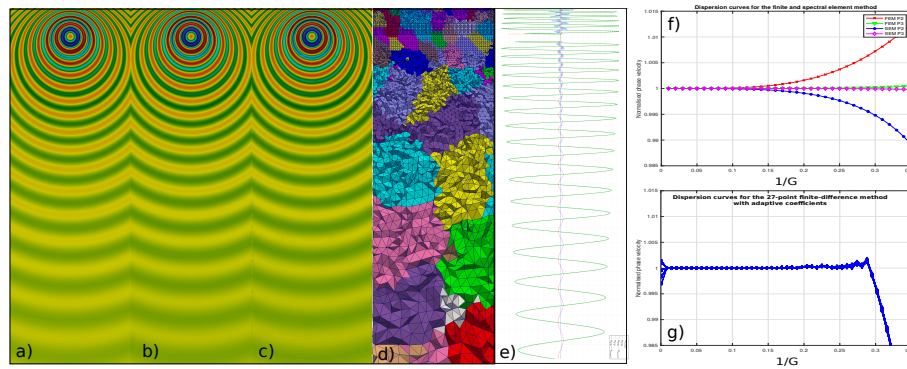


Figure 1: FE and FD solutions against analytical solution. (a) FD solution. (b) Analytical (A) solution. (c) FE solution. (d) FE mesh. (e) Comparison between (a-c) along a vertical profile cross-cutting source position. Gain with offset is applied. FD: dash blue; FE: dash red; A: green. FD-A: blue. FE-A: red. (f-g) Phase velocity dispersion curves for (f) FE and (g) FD. G : number of *ppwl*.

Domain decomposition preconditioner

We now review the preconditioner that we use to solve efficiently the linear system (2). A well-known iterative solver for this type of indefinite linear systems is the Krylov subspace Generalized Minimal RESidual Method (GMRES) (Saad, 2003). However, the Helmholtz operator requires efficient preconditioning which can be done by domain decomposition (Dolean et al., 2015, section 2.2.1).

In this study, we solve system (2) with a two-level domain decomposition preconditioner \mathbf{M}^{-1}

$$\mathbf{M}^{-1} = \mathbf{M}_1^{-1}(\mathbf{I} - \mathbf{A}\mathbf{Q}) + \mathbf{Q}, \quad \text{with } \mathbf{Q} = \mathbf{Z}\mathbf{E}^{-1}\mathbf{Z}^T, \quad \mathbf{E} = \mathbf{Z}^T\mathbf{A}\mathbf{Z}, \quad (3)$$

where \mathbf{M}_1^{-1} is the one-level domain decomposition preconditioner called Optimized Restricted Additive Schwarz (ORAS) and \mathbf{Z}^T is the interpolation matrix from the FE space defined on Γ onto a FE space defined on a coarse mesh Γ_H . The construction of the domain decomposition preconditioner is described in detail in Bonazzoli et al. (2019). Let $\{\Gamma_i\}_{1 \leq i \leq N_d}$ be an overlapping decomposition of the mesh Γ into N_d subdomains. Let $\{\mathbf{A}_i\}_{1 \leq i \leq N_d}$ denote local Helmholtz operators with absorbing (or transmission) boundary conditions at the subdomain interfaces. The one-level ORAS preconditioner is

$$\mathbf{M}_1^{-1} = \sum_{i=1}^{N_d} \mathbf{R}_i^T \mathbf{D}_i \mathbf{A}_i^{-1} \mathbf{R}_i, \quad (4)$$

where $\{\mathbf{R}_i\}_{1 \leq i \leq N_d}$ are the Boolean restriction matrices from the global to the local finite element spaces and $\{\mathbf{D}_i\}_{1 \leq i \leq N_d}$ are local diagonal matrices representing the partition of unity.

The key ingredient of the ORAS method is that the local matrices $\{\mathbf{A}_i\}_{1 \leq i \leq N_d}$ incorporate more efficient

boundary conditions (i.e. absorbing boundary conditions) than in the standard RAS preconditioner based on local Dirichlet boundary value problems. The coarse problem \mathbf{E} in (3) is also solved iteratively by performing 10 GMRES iterations with a one-level ORAS preconditioner. We use the same spatial subdomain partitioning for the coarse and fine meshes. Each computing core is assigned to one spatial subdomain and holds the corresponding coarse and fine local matrices. Each application of the global preconditioner \mathbf{M}^{-1} relies on local concurrent subdomain solves on the coarse and fine levels, which are performed by a direct solver. This hybrid direct/iterative solver requires careful strong scalability analysis to achieve the best compromise between parallel efficiency and memory storage.

Numerical results

The two-level solver is implemented using the high-performance domain decomposition library HPDMM (<http://github.com/hpddm/hpddm>) (High-Performance unified framework for Domain Decomposition Methods) (Jolivet et al., 2013). We assess the solver on the Irène supercomputer of TGCC (<http://www-hpc.cea.fr>) with the 3D $20 \times 20 \times 4.65$ km SEG/EAGE Overthrust model (Amizadeh et al., 1997). We perform wave simulation with P3 finite elements on regular and adaptive tetrahedral meshes (Fig. 2a) for the 5 Hz, 10 Hz and 20 Hz frequencies (Tab. 3) in double and single precision. The average length of the element edges is set to 5 nodes per minimum wavelength on the regular tetrahedral mesh, and 5 nodes per local wavelengths in the adaptive tetrahedral mesh (2.5 for the coarser mesh used in the two-level method). We use a homogeneous Dirichlet boundary condition at the surface

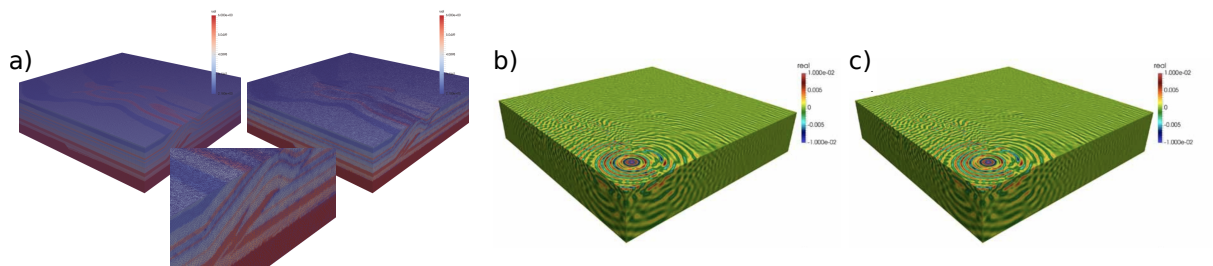


Figure 2: (a) Tetrahedral meshing of Overthrust. Solutions on (b) regular and (c) adaptive meshes.

and first-order absorbing boundary conditions along the other five faces of the model. The source is located at $(2.5, 2.5, 0.58)$ km. For weak scalability analysis, we keep #dofs per subdomain roughly constant from one frequency to the next (Tab. 3). The h -adaptivity in the unstructured tetrahedral mesh decreases #dofs relative to the regular mesh by a factor of 2.07. The stopping tolerance ε for GMRES is set to 10^{-3} . The consistency between the 10 Hz wavefields computed in the regular and adaptive tetrahedral meshes is shown in Fig. 2(b-c). First, we carry out a set of numerical simulations at 5 Hz on the regular mesh in order to illustrate the benefits of performing computations in single precision arithmetic (versus double precision), as well as using an approximate factorization for the fine local matrices to apply \mathbf{A}_i^{-1} in (4). More precisely, we compare incomplete Cholesky factorization (ICC) to complete Cholesky factorization performed by Intel MKL PARDISO. The experiments are performed on 1060 cores with P3 finite elements and 5 ppwl, resulting in 74 million dofs. Results are reported in Tab. 2. First, we can see that performing the whole computation in single precision instead of double precision yields a speedup of about 1.4 for the solution phase. The number of GMRES iterations is the same, there is no loss of accuracy or additional numerical instability. Additionally, the setup phase is drastically reduced (speedup 1.8) when performing Cholesky factorization in single precision. Second, we can see that using an incomplete Cholesky factorization for the fine local matrices yields a speedup of about 1.6 with respect to complete factorization, once again with no effect on the number of GMRES iterations. Moreover, the memory savings are pretty significant: with complete Cholesky factorization we run out of memory with 768 cores, while the simulation runs on 265 cores using ICC. In the rest of this paper, the experiments are performed in single precision and using incomplete Cholesky factorization for the fine local matrices. Timings for the adaptive tetrahedral mesh are around two times smaller than those obtained on the regular mesh (Tab. 3). The simulation at 20 Hz on the adaptive mesh involves 2,285 millions of dofs and requires 16,960 cores. The elapsed time achieved by the 2-level preconditioner is 15s and 37s for 10 Hz and 20 Hz respectively (Tab. 3).

Conclusions

We propose a highly-scalable hybrid direct/iterative solver based upon a domain decomposition preconditioner as a forward engine to perform large-scale 3D frequency domain FWI of sparse stationary-recording acquisitions. The method is implemented with a finite-difference and finite-element method

to select the most suitable scheme for the case study at hand. This forward engine should be used when the size of the problem outreaches the capability of leading-edge sparse direct solvers.

Acknowledgments:

This study was granted access to the HPC resources of SIGAMM (<http://crimson.oca.eu>) and CINES/IDRIS under the allocation 0596 made by GENCI. This study was partially funded by the WIND consortium (<https://www.geoazur.fr/WIND>) sponsored by Chevron, Shell and Total.

Cartesian grid, $f = 5\text{Hz}$				
precision	fine local solver	#it	setup(s)	$T_{GM}(s)$
double	Cholesky	10	92.5	15.5
double	ICC	10	30.2	8.9
single	Cholesky	10	50.3	10.3
single	ICC	10	25.8	6.3

Table 2: Comparison between Cholesky and incomplete Cholesky factorizations (ICC) of local matrices at the fine level, and single .vs. double precision arithmetic for the whole computation, at 5Hz with P3 FEs and 5 ppwl (74M dofs) on 1060 cores. #it: iteration count. setup: Elapsed time for the setup phase (assembly and factorization of local matrices). T_{GM} : Elapsed time in GMRES ($\epsilon = 10^{-3}$).

Regular tetrahedral mesh					
f(Hz)	#core	#elts (M)	#dofs (M)	#it	$T_{GM}(s)$
5	265	16	74	7	16s
10	2,120	131	575	15	33s
Adaptive tetrahedral mesh					
f(Hz)	#core	#elts (M)	#dofs (M)	#it	$T_{GM}(s)$
10	2,120	63	286	14	15s
20	16,960	506	2,285	30	37s

Table 3: Simulation cost in regular/adaptive meshes. f : frequency; #core: number of cores; #elts: number of elements; #dofs: number of dofs; #it: iteration count. T_{GM} : Elapsed time in GMRES.

References

- Ainsworth, M. and Wajid, H. [2010] Optimally Blended Spectral-Finite Element Scheme for Wave Propagation and NonStandard Reduced Integration. *SIAM Journal on Numerical Analysis*, **48**, 346–371.
- Amestoy, P., Brossier, R., Buttari, A., L’Excellent, J.Y., Mary, T., Métivier, L., Miniussi, A. and Operto, S. [2016] Fast 3D frequency-domain FWI with a parallel Block Low-Rank multifrontal direct solver: application to OBC data from the North Sea. *Geophysics*, **81**(6), R363 – R383.
- Aminzadeh, F., Brac, J. and Kunz, T. [1997] *3-D Salt/Overthrust models*. SEG/EAGE 3-D Modeling Series No.1.
- Bonazzoli, M., Dolean, V., Graham, I.G., Spence, E.A. and Tournier, P.H. [2019] A 2-level domain decomposition preconditioner for the time-harmonic Maxwell’s equations. *Math. Comp.*, **88**, 2559–2604.
- Dolean, V., Jolivet, P. and Nataf, F. [2015] *An introduction to domain decomposition methods*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. Algorithms, theory, and parallel implementation.
- Dolean, V., Tournier, P.H., Jolivet, P. and Operto, S. [2020] Large-scale frequency-domain seismic wave modeling on h-adaptive tetrahedral meshes with iterative solver and multi-level domain-decomposition preconditioners. In: *SEG Technical Program Expanded Abstracts 2020*. 2683–2688.
- Erlangga, Y. [2008] Advances in Iterative Methods and Preconditioners for the Helmholtz Equation. *Archives of Computational Methods in Engineering*, **15**, 37–66.
- Graham, I.G., Spence, E.A. and Vainikko, E. [2017] Recent Results on Domain Decomposition Preconditioning for the High-Frequency Helmholtz Equation Using Absorption. *Lahaye D., Tang J., Vuik K. (eds) Modern Solvers for Helmholtz Problems. Geosystems Mathematics. Birkhäuser, Cham*, 3–26.
- Jolivet, P., Hecht, F., Nataf, F. and Prud’homme, C. [2013] Scalable Domain Decomposition Preconditioners for Heterogeneous Elliptic Problems. In: *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, SC’13. ACM, NY, USA.
- Kuvshinov, B.N. and Mulder, W.A. [2006] The exact solution of the time-harmonic wave equation for a linear profile. *Geophysical Journal International*, **167**, 659–662.
- Operto, S., Virieux, J., Amestoy, P., L’Excellent, J.Y., Giraud, L. and Ben Hadj Ali, H. [2007] 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study. *Geophysics*, **72**(5), SM195–SM211.
- Saad, Y. [2003] *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia.
- Turkel, E., Gordon, D., Gordon, R. and Tsynkov, S. [2013] Compact 2D and 3D sixth order schemes for the Helmholtz equation with variable wave number. *Journal of Computational Physics*, **232**, 272–287.