www.geoazur.fr/WIND

MUMPS-aided frequency-domain FWI for high-resolution seismic imaging

Solving the time-harmonic wave equation with multiple sparse right-hand sides in large computational meshes

S. Operto[1], H. Aghamiry[1], S. Beller[1], L. Combe[1], G. Guo[1], F. Mamfoumbi[1], A. Ribodetti[1]
P. Amestoy[2], A. Buttari[3], M. Gerest[4], J.-Y. L'Excellent[2], T. Mary[5], C. Puglisi[2]
V. Dolean[1], P. Jolivet[5], P.-H. Tournier

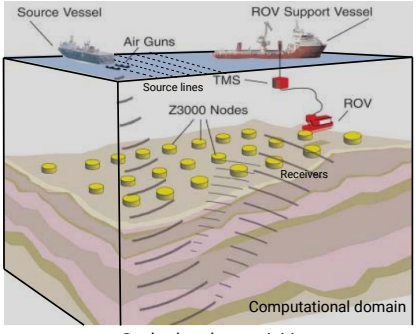
[1] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, IRD, France
[2] Mumps Technologies (https://mumps.tech.com), Lyon, France
[3] Université de Toulouse, CNRS, IRIT, France.
[4] EDF R & D, Paris, France.
[5] Sorbonne Université, LIP6, Paris, France
[6] Sorbonne Université, LJLL, Paris, France

Outline

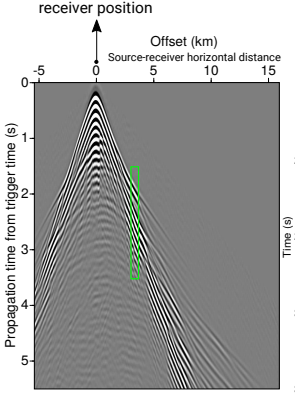
Outline

Illustration of seismic survey for oil exploration - seabed acquisition with dense node deployments

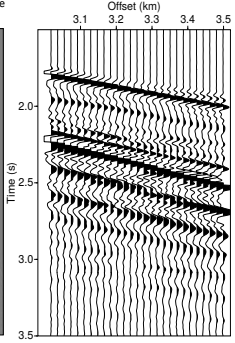
- Shooting vessel covers top of geological targeted area while triggering pressure sources at regular intervals. Acoustic/elastic waves propagate inside the medium, interact with heterogeneities before being recorded over time by receivers (here, a grid of autonomous nodes deployed on the seabed). Wave recordings provide indirect measurements of subsurface properties. Seismic imaging aims at drawing some inferences on the subsurface from these measurements.



Seabed node acquisition



Common-receiver gather



Samples of seismic traces

General definition of Full Waveform Inversion (FWI) (Tarantola, 1984; Pratt et al., 1998; Virieux and Operto, 2009)

- FWI: A nonlinear and ill-posed inverse problem aiming at converting waves into mechanical properties governing their propagation.
- Optimization problem: Minimize a distance between recorded data (wave measurements) and numerically-simulated data.
- Forward problem: Numerical simulation of wave propagation. Solve the wave equation (a linear PDE) in the framework of linear elasticity.
- Optimization variables: Spatially-varying parameters contained in the PDE coefficients.
- The forward problem for multiple RHS is the most computationally demanding task. MUMPS gets in the game at this stage.

- An initial-value evolution problem

$$\mathbf{M}\frac{\partial^2 \mathbf{w}(\mathbf{x},t)}{\partial t^2} + \mathbf{C}\frac{\partial \mathbf{w}(\mathbf{x},t)}{\partial t} + \mathbf{K}\mathbf{w}(\mathbf{x},t) = \mathbf{b}(\mathbf{x},t)$$

$$\text{State space} \quad \mathcal{U} := \left\{ \mathbf{w} \in \left[H^1(\Omega \times (0,T))\right]^3 : \mathbf{w}(\cdot,0) = 0; \dot{\mathbf{w}}(\cdot,0) = 0 \right\}, \tag{1}$$

  where Ω is the spatial computational domain containing the object domain (target to be imaged).

  1. Matrix-free explicit time-marching schemes.
  2. Cost scales to the number of right-hand sides ([reciprocal] sources).
  3. Attenuation in $\mathbf{D}$ generates computational ovearheads (factor 2 to 3).

- Forward engine of Reverse Time Migration (RTM) can be readily used for FWI. Convince the industry it is the right approach to take.

$\mathbf{M}$: mass matrix, $\mathbf{D}$: damping matrix, $\mathbf{K}$: stiffness matrix, $\mathbf{w}(\mathbf{x},t)$: wavefield, $\mathbf{b}(\mathbf{x},t)$: source term.

- A boundary-value problem (Generalized Helmholtz problem).

$$\left(-\omega^2\mathbf{M} + i\omega\mathbf{C} + \mathbf{K}\right)\mathbf{W}_\omega(\mathbf{x}) = \mathbf{B}_\omega(\mathbf{x})$$

$$\mathbf{AW} = \mathbf{B}.$$

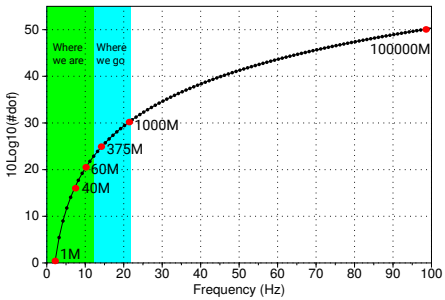
$$\text{State space} \quad \mathcal{U} := \left\{\mathbf{w} \in \left[H^1(\Omega)\right]^3 : \mathbf{w}(\partial\Omega_{z_1}) = 0; \mathbf{w}(\partial\Omega_{x,y,z_2}) = PML\right\} \tag{2}$$

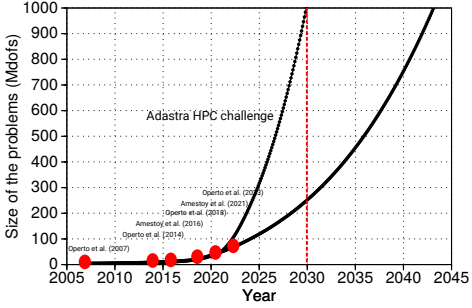
1. Large and sparse linear system / frequency with multiple right-hand sides.
2. Block processing of right-hand sides.
3. Implementation of attenuation is free.

- Two families of linear algebra approaches
  1. Direct methods for sparse matrices (Duff et al., 1986).
  2. Preconditioned iterative methods (Saad, 2003). When domain-decomposition preconditioner, hybrid direct-iterative methods (Tournier et al., 2022).

The curse of dimensionality in exploration geophysics

- Physics: acoustic approximation (no shear) - Scalar Helmholtz equation: $\left(\frac{\omega^2}{c^2(\mathbf{x})}\mathbf{I} + \Delta\right)\mathbf{w} = \mathbf{b}$.
- Frequency bandwidth: 2 Hz - 100 Hz for a minimum P wavespeed of 1.5 km/s
- A representative physical domain: $23.5\ km \times 30\ km \times 8\ km$
- Discretization rule for FWI: 4 grid points / min. wavelength.

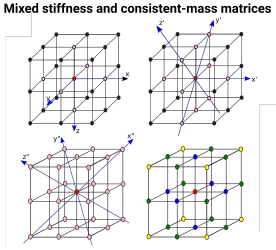




- Key functions of MUMPS: [1] BLR MUMPS with mixed-precision arithmetic, [2] efficient processing of multiple RHSs, [3] efficient multithreading (mitigate number of MPI process to minimize memory overheads).

Discretizing A with adaptive, compact and accurate finite-difference stencil:

27-point mixed-grid stencil (Operto et al., 2007; Brossier et al., 2010; Operto et al., 2014)



- Mass and stiffness operators

$$\left(-\omega^2\mathbf{M} + i\omega\mathbf{C} + \mathbf{K}\right)\mathbf{W}_\omega(\mathbf{x}) = \mathbf{B}_\omega(\mathbf{x}) \rightarrow \mathbf{AW} = \mathbf{B}.$$

- Compound stiffness matrix

$$\mathbf{K} = w_{s_1}\mathbf{K}_1 + \frac{w_{s_2}}{3}\sum_{i=1}^{3}\mathbf{K}_{2,i} + \frac{w_{s_3}}{4}\sum_{i=1}^{4}\mathbf{K}_{3,i},$$

where $\sum_{i=1}^{3} w_{s_i} = 1$.

- Consistent mass

$$(\mathbf{Mw})_0 = \omega^2\left(w_{m_1}\mathbf{w}_0 + w_{m_2}\sum_{i=1}^{6}\left[\frac{\mathbf{w}}{\kappa}\right]_{1,i} + w_{m_3}\sum_{i=1}^{12}\left[\frac{\mathbf{w}}{\kappa}\right]_{2,i} + w_{m_4}\sum_{i=1}^{8}\left[\frac{\mathbf{w}}{\kappa}\right]_{3,i}\right),$$

where $w_{m_1} + 6w_{m_2} + 12w_{m_3} + 8w_{m_4} = 1$.

- Ordering

Nested dissection based permutation (`ICNTL(7)=1`).

- A wave-equation constrained nonlinear optimization problem

$$\min_{\mathbf{w}_{s,f},\mathbf{m}} \sum_{f=1}^{N_f}\sum_{s=1}^{N_s} \| \overbrace{\mathbf{P}\mathbf{w}_{s,f} - \mathbf{d}^*_{s,f}}^{obs.\ equation} \|_2^2 \quad \text{subject to} \quad \overbrace{\mathbf{A}_f(\mathbf{m})\mathbf{w}_{s,f} - \mathbf{b}^*_{s,f} = \mathbf{0}}^{wave\ equation}, \overbrace{s = 1, ..., N_s}^{multi\ RHS}, \overbrace{f = 1, ..., N_f}^{multi\ freq.}. \quad (3)$$

- Full-space approach: Lagrange multiplier method (Akçelik, 2002)

$$\min_{\mathbf{w}_{s,f},\mathbf{m}} \max_{\mathbf{v}_{s,f}} \sum_{f=1}^{N_f}\sum_{s=1}^{N_s} \|\mathbf{P}\mathbf{w}_{s,f} - \mathbf{d}^*_{s,f}\|_2^2 + \sum_{f=1}^{N_f}\sum_{s=1}^{N_s} \left\langle \mathbf{v}_{s,f}, \mathbf{A}(\mathbf{m})_f \mathbf{w}_{s,f} - \mathbf{b}^*_{s,f} \right\rangle_{\mathcal{U}}, \quad (4)$$

- Reduced-space approach: Variable-projection approach (Golub and Pereyra, 2003) $\rightarrow$ data-fitting problem

$$\min_{\mathbf{m}} \phi^{FWI}(\mathbf{m}) = \sum_{f=1}^{N_f}\sum_{s=1}^{N_s} \|\mathbf{S}_f(\mathbf{m})\mathbf{b}^*_{s,f} - \mathbf{d}^*_{s,f}\|_2^2, \quad (5)$$

$\mathbf{S}_f(\mathbf{m}) = \mathbf{P}\mathbf{A}_f^{-1}(\mathbf{m})$: forward modeling operator; $\mathbf{d}_{s,f}(\mathbf{m}) = \mathbf{S}_f(\mathbf{m})\mathbf{b}^*_{s,f}$: simulated data.

- Linearization around $\mathbf{m}_k$ (Newton-type method)

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha \delta \mathbf{m}^{(k)}. \tag{6}$$

where

$$\frac{\partial^2 \phi^{FWI}(\mathbf{m}^{(k)})}{\partial \mathbf{m}^2} \delta \mathbf{m}^{(k)} = -\frac{\partial \phi^{FWI}(\mathbf{m}^{(k)})}{\partial \mathbf{m}}$$

$$\phi^{FWI}(\mathbf{m}) = \sum_{f=1}^{N_f} \sum_{s=1}^{N_s} \|\mathbf{S}_f(\mathbf{m})\mathbf{b}^*_{s,f} - \mathbf{d}^*_{s,f}\|_2^2 = \sum_{f=1}^{N_f} \sum_{s=1}^{N_s} \|\delta \mathbf{d}^*_{s,f}(\mathbf{m})\|_2^2.$$

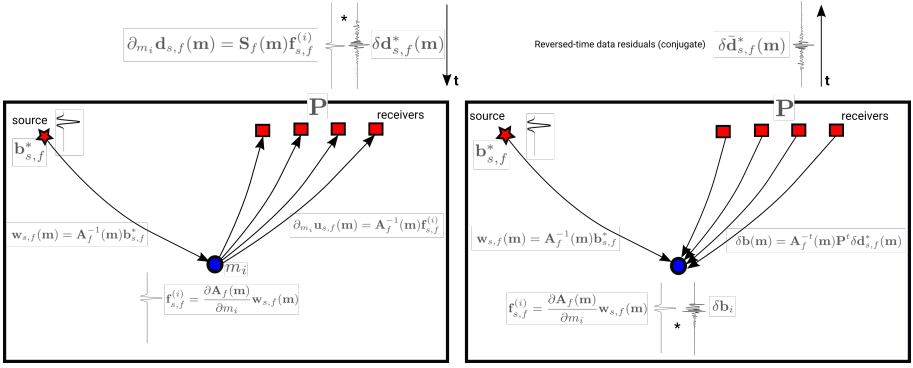
- Gradient of misfit function: diffraction-stack formulation

$$\nabla_{m_i} \phi^{FWI}(\mathbf{m}) = \sum_{f=1}^{N_f} \sum_{s=1}^{N_s} \left(\mathbf{S}_f(\mathbf{m})\mathbf{f}^{(i)}_{s,f}\right)^T \left(\delta \mathbf{d}^*_{s,f}(\mathbf{m})\right) \quad \text{with} \quad \mathbf{f}^{(i)}_{s,f} = \frac{\partial \mathbf{A}_f(\mathbf{m})}{\partial m_i} \mathbf{w}_{s,f}(\mathbf{m}). \tag{7}$$

- Gradient of misfit function: Computationally-efficient adjoint-state formulation

$$\nabla_{m_i} \phi^{FWI}(\mathbf{m}) = \sum_{f=1}^{N_f} \sum_{s=1}^{N_s} \left(\mathbf{f}^{(i)}_{s,f}\right)^T \left(\mathbf{S}^T_f(\mathbf{m})\delta \mathbf{d}^*_{s,f}(\mathbf{m})\right) \quad \text{with} \quad \mathbf{f}^{(i)}_{s,f} = \frac{\partial \mathbf{A}_f(\mathbf{m})}{\partial m_i} \mathbf{w}_{s,f}(\mathbf{m}). \tag{8}$$

Computing the gradient with the adjoint-state method by time reversal

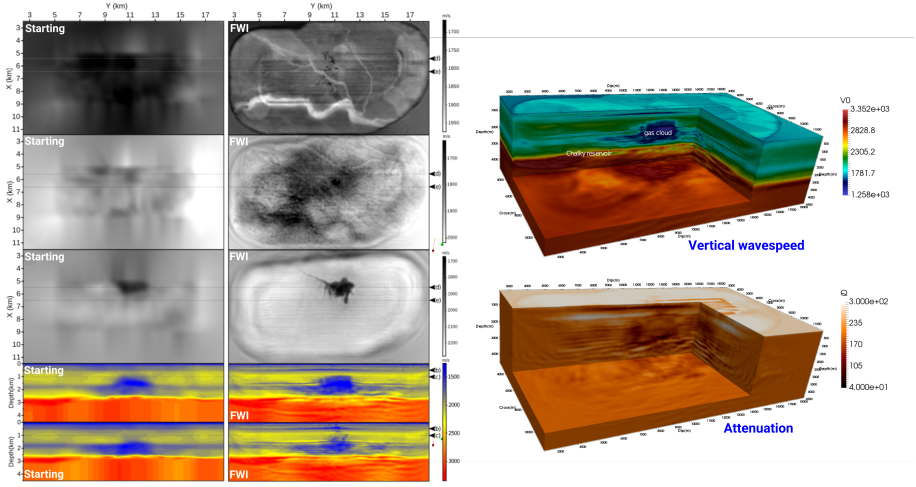


- Conclusion: Two wave simulations / (reciprocal) source & FWI iteration using $\mathbf{A}$ and $\mathbf{A}^t$ as matrices, respectively.
- First set of RHSs are (reciprocal) point sources $\mathbf{b}^*_{s,f}$ on the sea bed (nodes); second set of RHS are distributed-source below sea surface at (reciprocal) receiver positions $\delta\mathbf{d}^*_{s,f}(\mathbf{m})$.

Outline

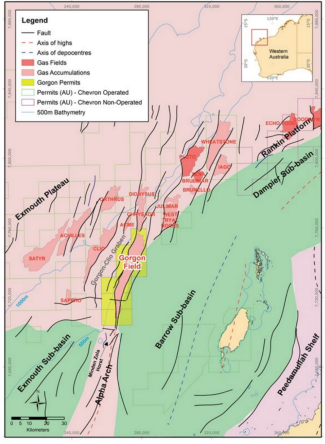
The first real-data case study of 3D frequency-domain FWI based on direct solver

(Operto et al., 2015; Amestoy et al., 2016; Operto and Miniussi, 2018; Amestoy et al., 2021)

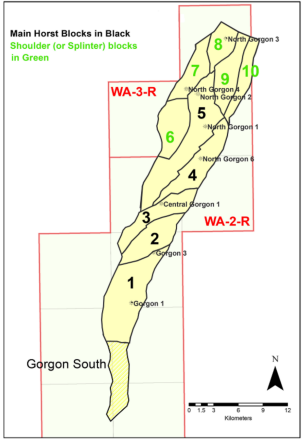


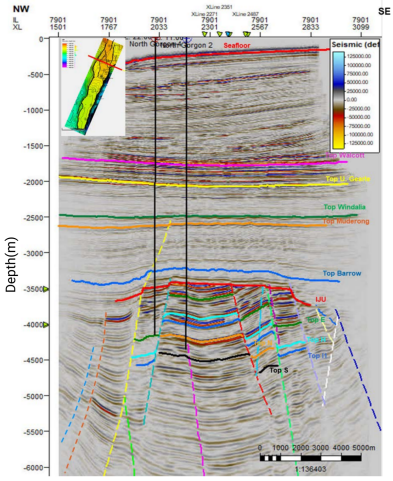
- **Target**: Elongated horst block trending N-NE in the North West Shelf of Australia with intra-horst en-echelon faults separating field into fault blocks.

Structural map

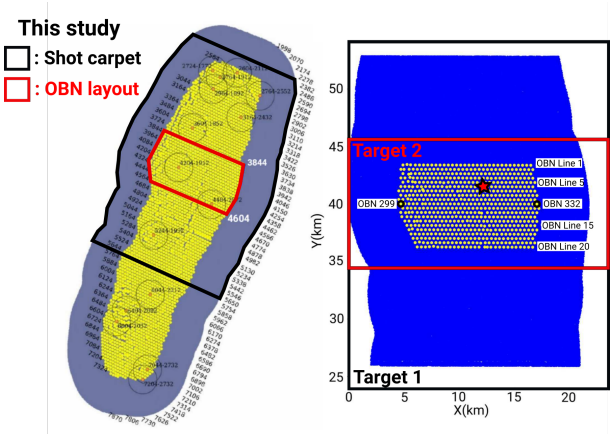


Gorgon field fault blocks



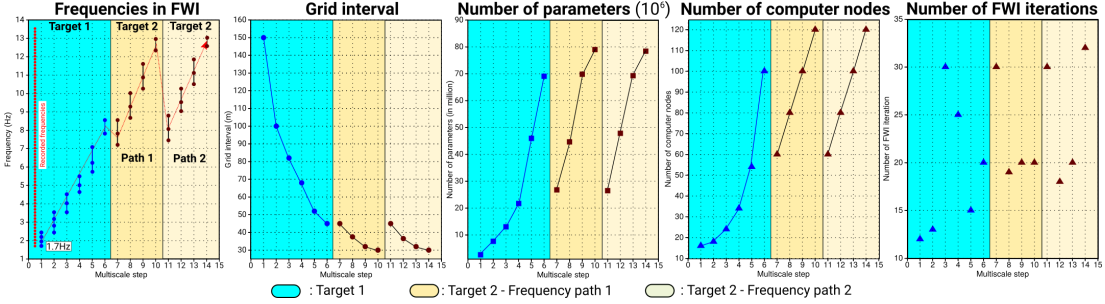


- **Full survey**:
  3100 4C OBns; 735 source lines; 697,345 shot points ($980\ km^2$); $240\ km^2$ full fold area.
- **Available dataset**:
  650 OBNs (OBN spacing: 375 m);
  Source carpet: 400,258 shots. **Area**: $\sim 705\ km^2$
- **Target 1 for FWI:**
  Full shot carpet involved in FWI.
  $23.5\ km(dip) \times 30\ km(cross) \times 8\ km(depth)$;
  Area: $\sim 705\ km^2$.
  Maximum frequency for FWI: 8.55 Hz.
- **Target 2 for FWI**:
  Shot carpet above OBN layout.
  $23.5\ km(dip) \times 11\ km(cross) \times 8\ km(depth)$;
  Area: $\sim 258.5\ km^2$.
  Maximum frequency for FWI: 13.2 Hz.
- **Recording length**: 6 s.

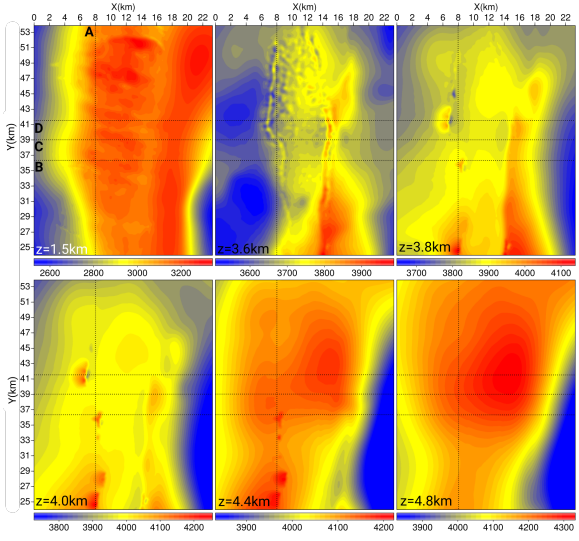


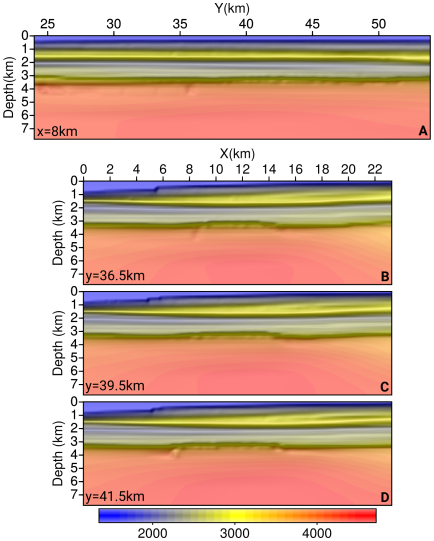
FWI experimental setup

- **Hardware**: Jean-Zay supercomputer of IDRIS, France, 1528 CPU nodes, 192 Gb / node.
- 2 processors Intel Cascade Lake 6248 (20 cores at 2,5 GHz), 40 cores /node.
- **Job design**: 1 MPI process / node with 40 threads → Mitigate memory overhead.
- All nodes & shots involved at each FWI iteration.
- **Frequency management**: Frequency patches with matched grid interval.

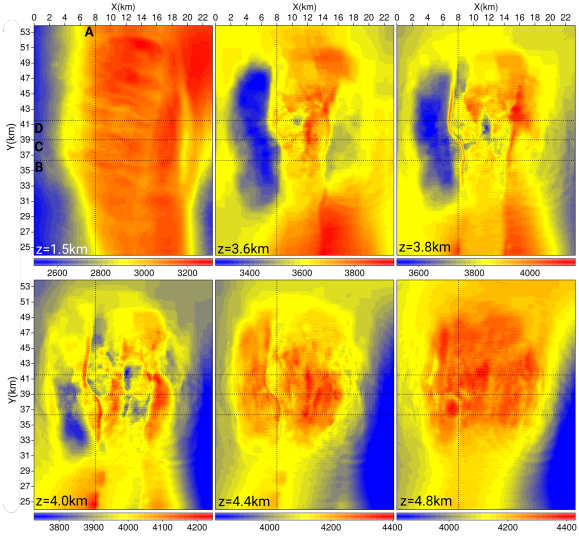


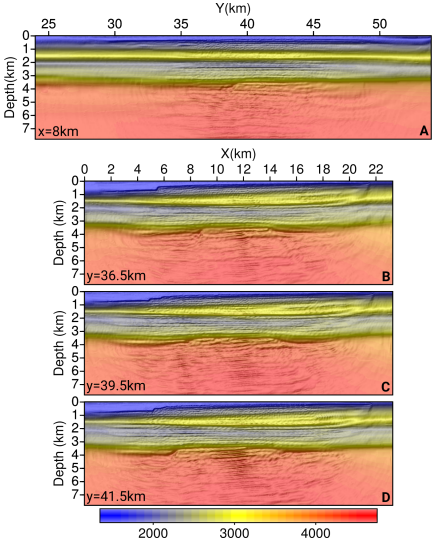
Imaging the Gorgon horst (Target 1): Starting model



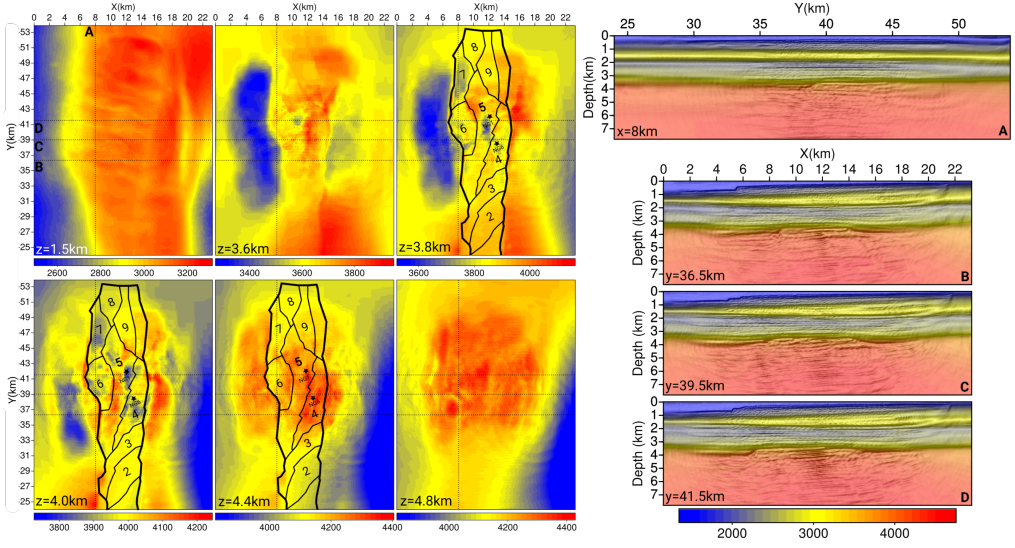


Imaging the Gorgon horst (Target 1): FWI model - f = 8.5 Hz (70 Mdof)

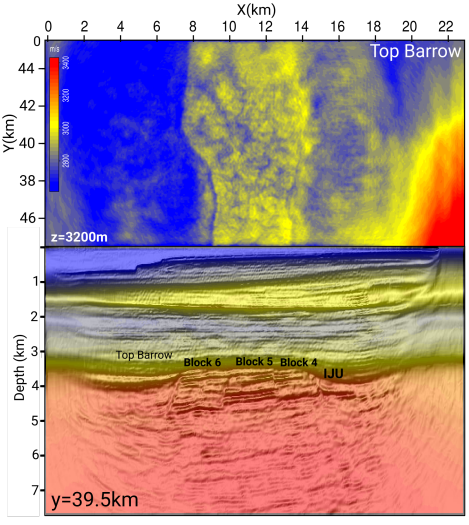


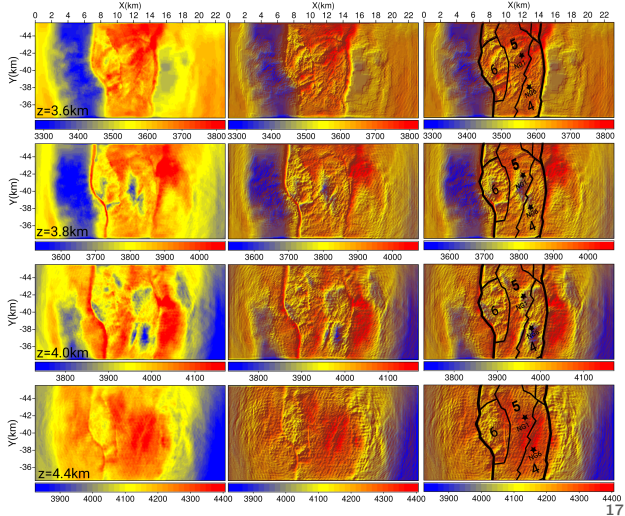


Imaging the Gorgon horst (Target 1): FWI model - f = 8.5 Hz (70 Mdof)

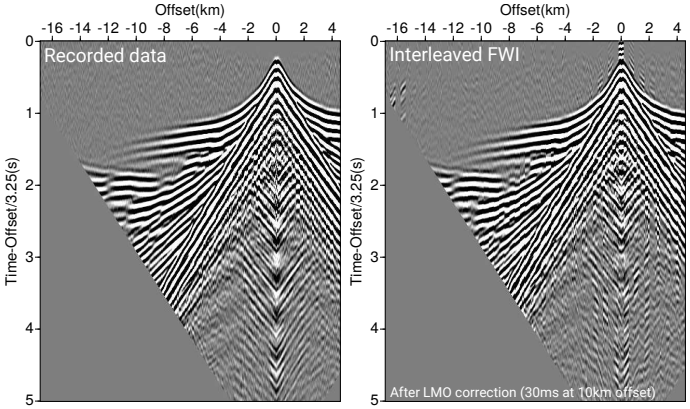


Imaging the Gorgon horst (Target 2): FWI model (Path 1) - f = 13.20 Hz (80 Mdof)





QC of FWI results: Data fit - Direct comparison between recorded and simulated data (FWI)
Bandwidth: 1.7-2-11-13 Hz



- QC of FWI results: time-domain simulated data with frequency-domain forward engine.
- Account for attenuation accurately.
- 80 LU factorizations & solution steps on the finest grid (around 30 hours elapsed time).
- Solutions needed only at receiver surface (shot carpet). Speedup of backward step?

MUMPS performances

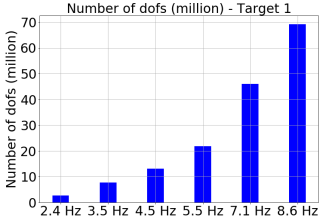
In the next slides,

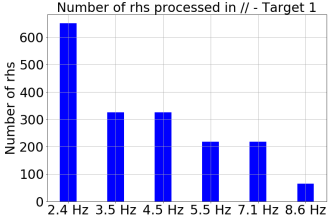
- Comparison of FR, BLR, MP-BLR for targets 1 and 2 (ICNTL(35)=2; ICNTL(36)=ICNTL(37)=1. CNTL(7)=10^{-5}.
- Comparison when source sparsity is used or not for targets 1 and 2. ICNTL(20)=2/3.
- Comparison of solution step when sources are processed sequentially or in parallel. ICNTL(27)=1/217.

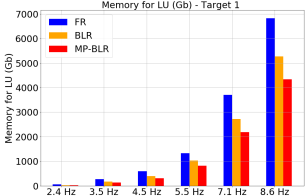
Fixed parameters in MUMPS:

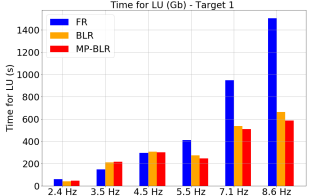
- Nested dissection
- Pivoting: KEEP(268)=-2
- Multithreading: KEEP(401)=1
- Scheduler: KEEP(370)=KEEP(371)=1

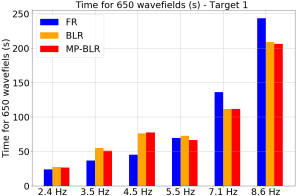
Target 1 ($23.5\ km \times 30\ km \times 8km$) - BLR with mixed-precision arithmetic

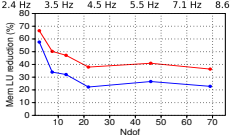


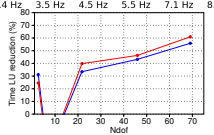




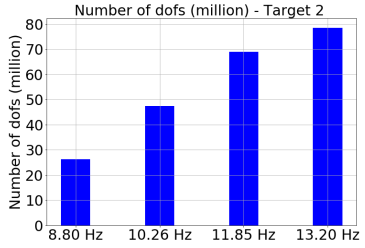


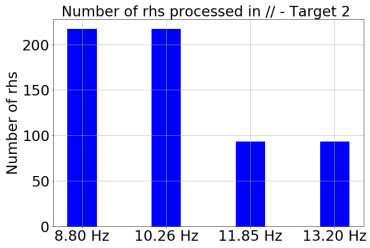


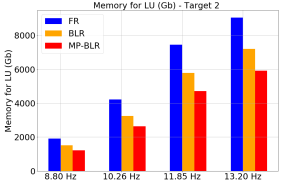


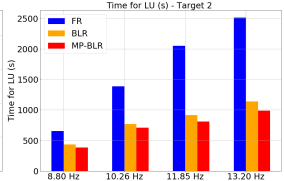


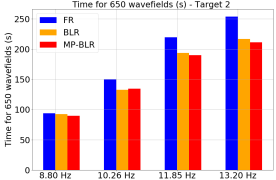
Target 2 ($23.5\ km \times 11\ km \times 8km$) - BLR with mixed-precision arithmetic

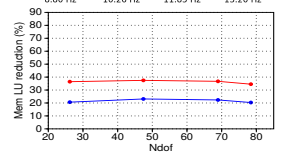


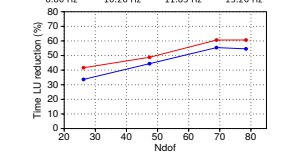




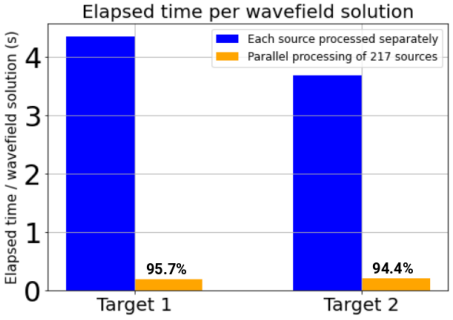




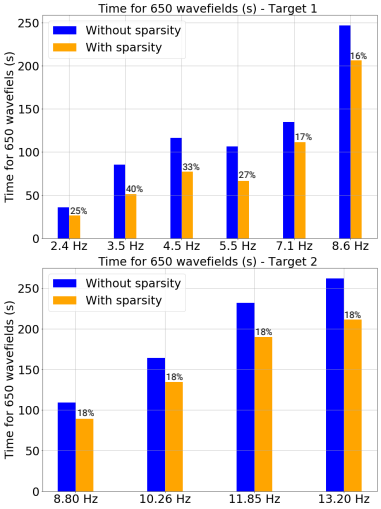




- **Target 1**: $23.5\ km \times 30\ km \times 8km$; Frequency: 7.1 Hz; #dofs: 46 millions. 54 MPI process.
- **Target 2**: $23.5\ km \times 11\ km \times 8km$; Frequency: 10.02 Hz; #dofs: 44.7 millions. 54 MPI process.



Computation done with MP$_B LR$.

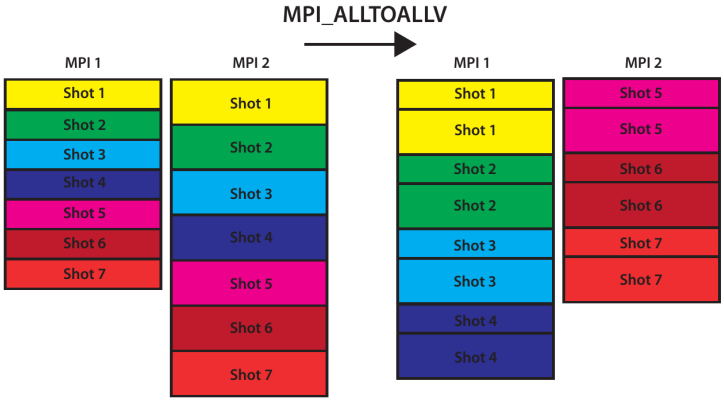


First-order conclusions

- Costs of LU and solution phases are of the same order of magnitude (consistent with theoretical complexities of LU). However, the cost of LU remains higher due to the small number of RHSs for this sparse node survey.
- The memory used during the factorization step and the solution step is of the same order of magnitude. INFOG(19) vs INFOG(31).
- BLR and MP-BLR lead to a memory reduction during LU of the order of 20% and 40%, respectively.
- BLR and MP-BLR lead to a elapsed-time reduction during LU of the order of 30-55% and 40-60%, respectively. These speed-ups increases with the size of the problem.
- Processing of mult RHS is speed up significantly with BLAS.
- Exploiting source sparsity further decreases the computational time by a factor of 18% when the sources cover a significant part of the surface. Better speed up is obtained for large computational domain?

Additional comment: Resorting distributed solutions

- It is beneficial for our FWI application to resort the distributed multi-rhs solution returned by MUMPS such the FWI gradient is computed with en embarrassing parallelism over sources.
- We recently face some problems with the collective communication MPI_ALLTOALLV due the size of some buffers associated with the largest domain.



Outline

Hybrid direct/iterative solver based domain-decomposition preconditioner for large-scale problems Tournier et al. (2022)

- Tackling large computational domains (several hundreds of millions of parameters to few billions of parameters) with sparse acquisitions (reasonable number of right-hand sides, few hundreds to few thousands) direct us toward more scalable and less memory demanding solvers such as domain-decomposition based hybrid direct/iterative solver.
- The Optimized Restricted Additive Schwarz (ORAS) preconditioner was assessed for large scale seismic modeling by Tournier et al. (2022) using both the wavelength-adaptive finite-difference method and P3 finite-element method on unstructured tetrahedral mesh.
- The global preconditioned system is solved with the iterative GMRES solver (Saad, 1986) while the preconditioner is built by solving local problems with MUMPS with absorbing boundary conditions (Robin conditions or PML) between subdomains.

Hybrid direct/iterative solver based domain-decomposition preconditioner: Principles

- Helmholtz system

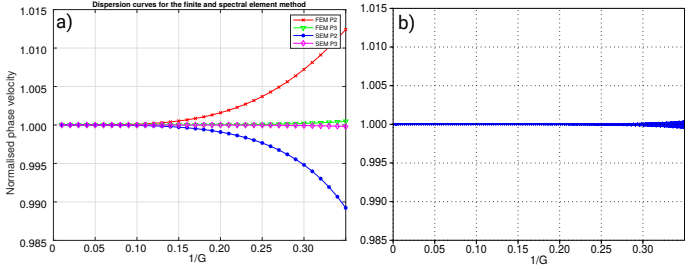
$$\mathbf{A}\mathbf{w} = \mathbf{b}. \tag{9}$$

- Right-preconditioned system

$$\mathbf{A}\mathbf{M}^{-1}\mathbf{y} = \mathbf{b} \quad \text{where } \mathbf{w} = \mathbf{M}^{-1}\mathbf{y} \quad \text{and} M_{\mathsf{ORAS}}^{-1} = \sum_{j=1}^{N} R_j^T D_j B_j^{-1} R_j. \tag{10}$$

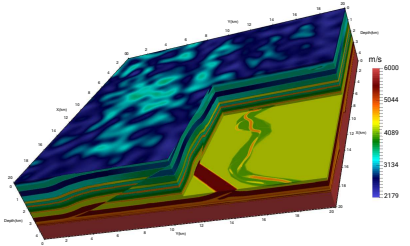
- Partitioning of Ω into N overlapping subdomains $\{\Omega_j\}_{j=1}^N$.
- Boolean matrix $R_j \in \mathbb{R}^{n_j \times n}$ where n is the total number of degrees of freedom and n_j is the number of degrees of freedom in Ω_j.
- Partition of unity having matrix form $D_j \in \mathbb{R}^{n_j \times n_j}$
- R_j and D_j are built such that $\sum_{j=1}^N R_j^T D_j R_j = I$.
- B_j is the local matrix on Ω_j and Robin or impedance boundary conditions are assumed on $\partial\Omega_j \setminus \partial\Omega$.

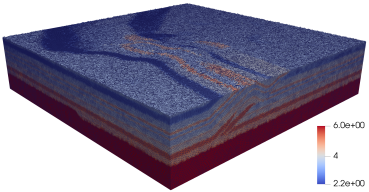
- Objectives from the practitioner perspective: Tune the solver with representative large-scale benchmarks to find the best compromise between accuracy and efficiency.
- Discretization: 4 grid points per minimum wavelength (finite differences) and $h = \lambda_{local}/4$ where h denotes the average element edges where λ_{local} is the local wavelength (h-adaptivity in finute elements. This is optimal for the helf wavelength resolution of FWI.
- Find the polynomial order such that accuracy of FE and FD roughly equivalent: P3 elements.

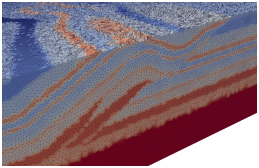


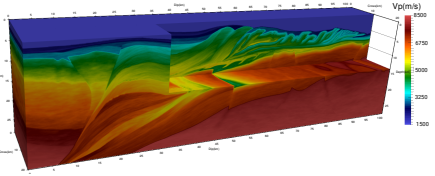
3D SEG/EAGE overthrust (Aminzadeh et al., 1997) & GO_3D_OBS benchmarks (Górszczyk and Operto, 2021)

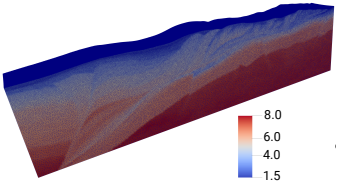
Unstructured tetrahedral meshing with Mmg (Dapogny et al., 2014) re-meshing software











- Stopping criterion of GMRES iteration

$$\varepsilon_{\text{GMRES}} = \frac{\|Aw - b\|_2^2}{\|b\|_2^2} = 10^{-4}.$$

- Arithmetic precision (single versus double complex precision)
- Local solver (FR/BLR MUMPS, INTEL MKM Pardiso). $\epsilon_{BLR} = 10^{-3}$.
- Orthogonalization of the Krylov basis: Classical Gram-Schmidt (CGS) and modified Gram-Schmidt (MGS) algorithms (Saad, 2003).
- Multi-RHS processing: Pseudo-block Krylov method, a block Krylov method and block Krylov with recycling by processing the 130 RHSs in group of 20 (Parks et al., 2006; Jolivet and Tournier, 2016).

Table 1: $Solver$: Intel MKL PARDISO versus MUMPS$_{\text{FR/BLR}}$. $arith$: $single$ versus $double$ precision. $ortho$: CGS versus MGS orthogonalization. #it: Number of iterations. $T_f(s)$: Elapsed time for local LU factorizations. $T_s(s)$: Elapsed time for all GMRES iterations for one RHS and 130 RHSs processed with a pseudo-block Krylov method. $T_{tot}(s) = T_f(s) + T_s(s)$: Total elapsed time for the simulation.

				1 RHS			130 RHS		
Solver	arith.	ortho.	$T_f(s)$	#it	$T_s(s)$	$T_{tot}(s)$	#it	$T_s(s)$	$T_{tot}(s)$
PARDISO	double	MGS	12.1	48	16.5	28.6	50	724.4	736.5
PARDISO	double	CGS	12.1	48	16.7	28.8	50	575.5	587.6
MUMPS$_{\text{FR}}$	double	MGS	11.3	48	17.2	28.5	50	654.9	666.2
MUMPS$_{\text{FR}}$	double	CGS	11.3	48	16.9	28.2	50	482.2	493.5
PARDISO	single	MGS	6.5	48	9.9	16.4	50	403.6	410.1
PARDISO	single	CGS	6.5	48	9.6	16.1	50	334.8	341.3
MUMPS$_{\text{FR}}$	single	MGS	6.2	48	8.9	15.1	50	345.8	352.0
MUMPS$_{\text{FR}}$	single	CGS	6.2	48	8.9	15.1	50	275.3	281.5
MUMPS$_{\text{BLR}}$	single	CGS	4.9	53	7.2	12.1	55	256.8	261.7

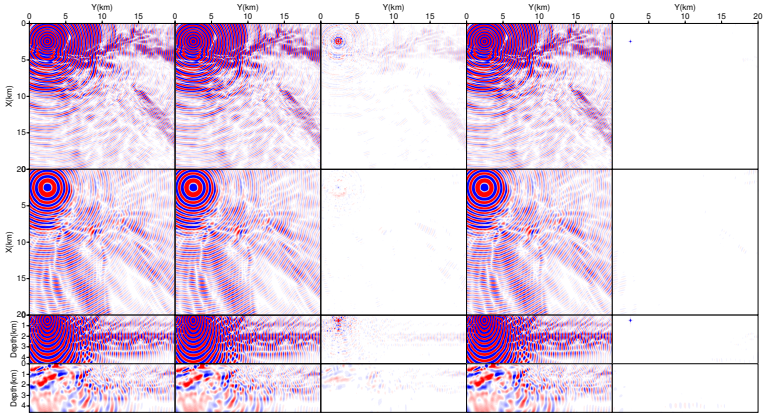
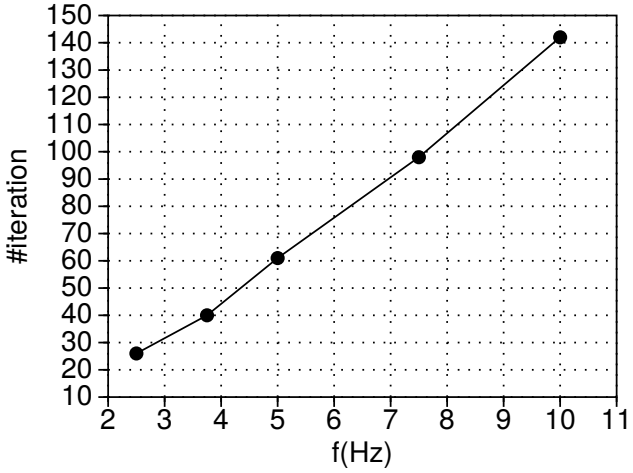


Figure 1: 3D EAGE/SEG overthrust model. Comparison between the wavefields computed with the CBS, FEFD and FDFD methods. The rows shown from top to bottom two depth slices at 500 m depth (across the source) and 2 km depth, and two vertical sections at x=2.5 km (across the source) and 15 km. From left to right, the columns show the CBS wavefield, the FEFD wavefield, the differences between the two, the FDFD wavefield and its differences with the CBS wavefield.

Table 2: Results of the four benchmarks obtained with the FEFD and FDFD methods. $\#d$: Number of dofs (including PMLs). $\#cores$: Number of cores. $\#it$: Number of GMRES iterations (for FEFD, average number of inner iterations to solve the coarse problem in parentheses). $T_f(s)$: Elapsed time for local LU factorizations. $T_s(s)$: Elapsed time for all GMRES iterations. $T_{tot}(s) = T_f(s) + T_s(s)$: Total elapsed time for the simulation. $T_{hc}(s) = \#cores \times T_{tot}$: Scalar time. Err: Solution error.

	FEFD method							
Benchmark	$\#d(M)$	#cores	#it	$T_f(s)$	$T_s(s)$	$T_{tot}(s)$	$T_{hc}(h)$	Err
Homogeneous	526.2	2400	31(13)	64.1	73.0	137.1	91.4	0.0891
Gradient	147.1	2400	5(14)	12.4	6.1	18.5	12.3	0.0485
Gradient$^{(c)}$	526.2	2400	8(16)	59.8	22.3	82.1	54.7	0.0119
Overthrust	157.1	2400	6(17)	13.5	6.7	20.2	13.5	0.2690
Overthrust$^{(c)}$	516.5	2400	6(20)	57.6	19.5	77.1	51.4	0.2351
GO_3D_OBS	597.7	2400	10(14)	62.3	26.1	88.4	58.9	0.2807
	FDFD method							
Benchmark	$\#d(M)$	#cores	#it	$T_f(s)$	$T_s(s)$	$T_{tot}(s)$	$T_{hc}(h)$	Err
Homogeneous	19.6	396	41	5.4	5.7	11.1	1.2	0.0317
Gradient	23.5	396	42	5.7	6.4	12.1	1.3	0.0114
Overthrust	19.1	363	33	5.0	4.5	9.5	0.96	0.1188
GO_3D_OBS	67.5	660	45	10.8	11.5	22.3	4.1	0.2096


#iteration
f(Hz)
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
2
3
4
5
6
7
8
9
10
11

Outline

- Representation industrial FWI case studies can be tackled with sparse direct solvers like MUMPS. A statement that was considered as utopian two decades ago. I did not know it was impossible so MUMPS did it.
- Limits of MUMPS were not reached in terms of problem size. Projet MUMPS4FWI supported by Adastra HPC challenge (CINES) will contribute to reach these limits in the next six months.
- Future works: update our FFWI code with
  1. Interface ORAS-based hybrid direct/iterative solver in FFWI.
  2. Finite-element schemes (spectral element method with consistent mass and Chebychev polynomials) (Seriani and Oliveira, 2008). Elemental assemblage of the matrix.
  3. Elastic physics (vertorial wave equation).
  4. New FWI formulation extending its linear regime (cycle skipping pathology).

Outline

Acknowledgements

www.geoazur.fr/WIND

- This study was partially funded by the sponsors of the WIND consortium.
- We thank Chevron and DUG for offering a copy of the preprocessed Gorgon OBN data and peripherical products to the WIND project.
- We are grateful to the OPAL (Observatoire Pluridisciplinaire des Alpes Maritimes) infrastructure from Observatoire de la Côte d'Azur for providing resources and support. This work was granted access to the high-performance computing resources of GENCI under allocation 0596.
- **Thanks to the MUMPS team for sharing with us their solver, their expertise and their good spirit)**
- **... and thank you for listening me**.

Aghamiry, H., Gholami, A., Combe, L., and Operto, S. (2022). Accurate 3D frequency-domain seismic wave modeling with the wavelength-adaptive 27-point finite-difference stencil: a tool for full waveform inversion. Geophysics, 87(3):1–66.

Akçelik, V. (2002). Multiscale Newton-Krylov methods for inverse acoustic wave propagation. PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania.

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.

Amestoy, P. R., Buttari, A., Combe, L., Gerest, M., L'Excellent, J.-Y., Mary, T., Operto, S., and Puglisi, C. (2021). Up to date assessment of 3D frequency-domain full waveform inversion based on the sparse multifrontal solver MUMPS. In Fifth EAGE Workshop on High Performance Computing for Upstream, volume 5.

Aminzadeh, F., Brac, J., and Kunz, T. (1997). 3-D Salt and Overthrust models. SEG/EAGE 3-D Modeling Series No.1.

Brossier, R., Etienne, V., Operto, S., and Virieux, J. (2010). Frequency-domain numerical modelling of visco-acoustic waves based on finite-difference and finite-element discontinuous galerkin methods. In Dissanayake, D. W., editor, Acoustic Waves, pages 125–158. SCIYO.

Dapogny, C., Dobrzynski, C., and Frey, P. (2014). Three-dimensional adaptive domain remeshing, implicit domain meshing, and applications to free and moving boundary problems. Journal of computational physics, 262:358–378.

Duff, I. S., Erisman, A. M., and Reid, J. K. (1986). Direct methods for sparse matrices, second edition. Oxford Science Publications, Oxford, U. K.

Golub, G. and Pereyra, V. (2003). Separable nonlinear least squares: the variable projection method and its applications. Inverse problems, 19(2):R1.

Górszczyk, A. and Operto, S. (2021). Go_3d_obs: the multi-parameter benchmark geomodel for seismic imaging method assessment and next-generation 3d survey design (version 1.0). Geoscientific Model Development, 14:1773–1799.

Jolivet, P. and Tournier, P. H. (2016). Block iterative methods and recycling for improved scalability of linear solvers. In SC16: International Conference for High Performance Computing, Networking, Storage and Analysis, pages 190–203.

Operto, S., Brossier, R., Combe, L., Métivier, L., Ribodetti, A., and Virieux, J. (2014). Computationally-efficient three-dimensional visco-acoustic finite-difference frequency-domain seismic modeling in vertical transversely isotropic media with sparse direct solver. Geophysics, 79(5):T257–T275.

Operto, S. and Miniussi, A. (2018). On the role of density and attenuation in 3D multi-parameter visco-acoustic VTI frequency-domain FWI: an OBC case study from the North Sea. Geophysical Journal International, 213:2037–2059.

Operto, S., Miniussi, A., Brossier, R., Combe, L., Métivier, L., Monteiller, V., Ribodetti, A., and Virieux, J. (2015). Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation. Geophysical Journal International, 202(2):1362–1391.

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.

Parks, M., de Sturler, E., Mackey, G., Johnson, D., and Maiti, S. (2006). Recycling krylov subspaces for sequences of linear systems. SIAM Journal of Scientific Computing, 28(5):1651–1674.

Plessix, R.-E. (2017). Modern solvers for Helmholtz problems, Geosystems Mathematics, chapter Some computational aspects of the time and frequency domain formulations of seismic waveform inversion, pages 159–187. Springer.

Pratt, R. G., Shin, C., and Hicks, G. J. (1998). Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. Geophysical Journal International, 133:341–362.

Saad, Y. (1986). GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM Journal on Scientific and Statistical Computing, 7(3):856–869.

Saad, Y. (2003). Iterative Methods for Sparse Linear Systems. SIAM, Philadelphia.

Seriani, G. and Oliveira, S. (2008). Dispersion analysis of spectral element methods for elastic wave propagation. Wave Motion, 45(6):729–744.

Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. Geophysics, 49(8):1259–1266.

Tournier, P.-H., Jolivet, P., Dolean, V., Aghamiry, H., Operto, S., and Riffo, S. (2022). Three-dimensional finite-difference & finite-element frequency-domain wave simulation with multi-level optimized additive schwarz domain-decomposition preconditioner: A tool for fwi of sparse node datasets. Geophysics, 87(5):T381–T402.

Virieux, J. and Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. Geophysics, 74(6):WCC1–WCC26.

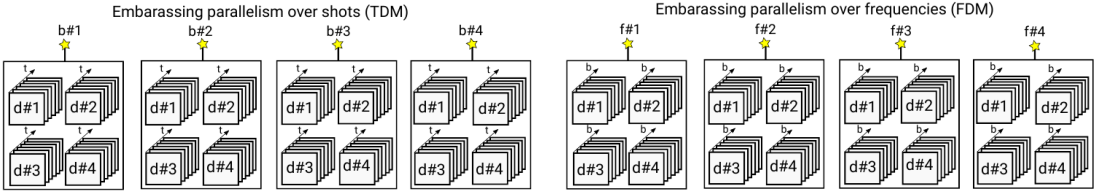
Virieux, J., Operto, S., Ben Hadj Ali, H., Brossier, R., Etienne, V., Sourbier, F., Giraud, L., and Haidar, A. (2009). Seismic wave modeling for seismic imaging. The Leading Edge, 28(5):538–544.

Outline

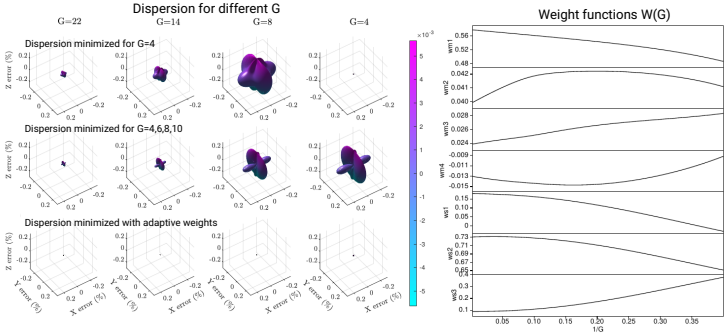
Forward problem: Complexity analysis of time-domain versus frequency-domain approaches

For a n^3 computational grid and n^2 sources,

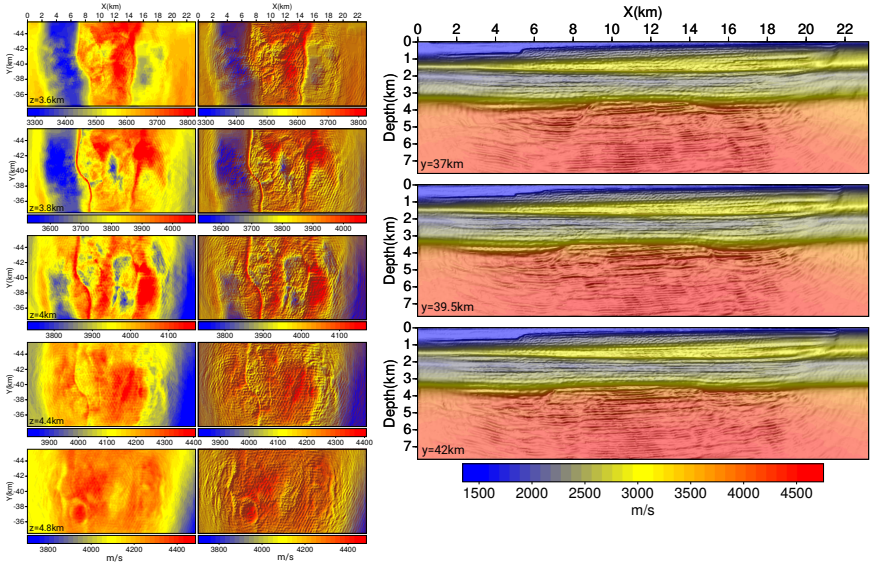
- **Time complexity** of FD and TD approaches scale to $\mathcal{O}(n^6)$ (for 1 frequency in the FD case).
- **Frequency-domain FWI should be performed with a limited number of frequencies**.
- **Memory complexity**: TD approach scales to $\mathcal{O}(n^5)$ for storage of n^2 wavefield snapshots (more memory demanding thn the storage of the LU factors $\mathcal{O}(n^4)$)
  $\rightarrow$ Random subsets of sources are processed at each iteration.
- For direct methods, few batches of the full shot carpet are processed sequentially thanks to **highly-efficient solution step**, while sources of a batch are processed in parallel with block methods.



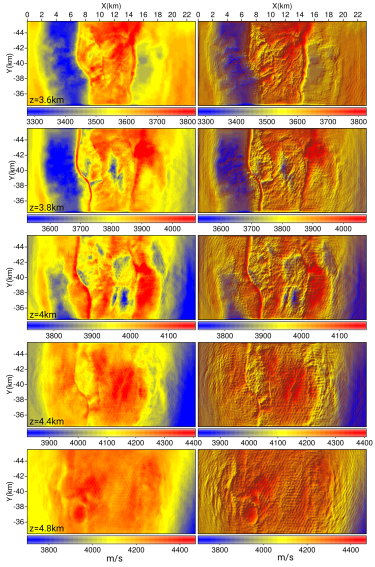
Wavelength-adaptive 27-point stencil: Making accuracy uniform (Aghamiry et al., 2022)

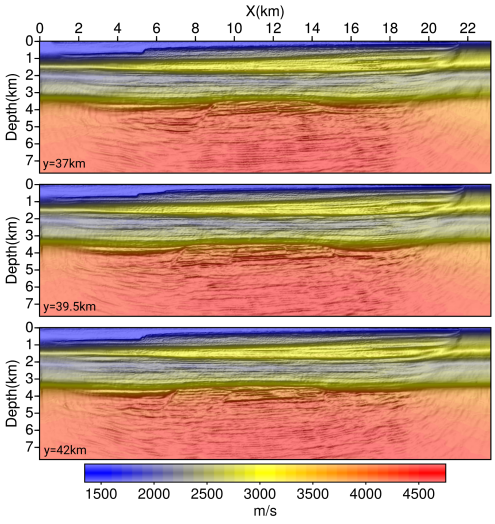


Imaging the Gorgon horst (Target 2): FWI model (Path 2) - f = 8.79 Hz (26 Mdof)

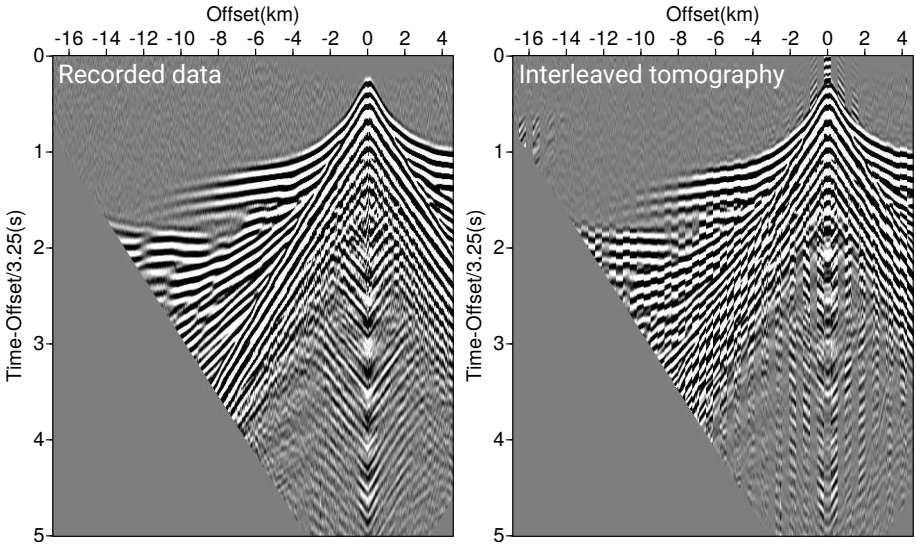


Imaging the Gorgon horst (Target 2): FWI model (Path 2) - f = 13.20 Hz (80 Mdof)





QC of FWI results: Direct comparison between recorded and simulated data (tomography)
Bandwidth: 1.7-2-11-13 Hz



Computational cost and comparison with FDTD modeling

- **FDTD**: $\mathcal{O}(\Delta t^2, \Delta x^8)$ accuracy.
- **FDTD**: Two level parallelism (source distribution + domain decomposition)
- **FDTD**: 2 sources processed on a node with 20 threads (domains) per source.
- **FDTD**: modeling without attenuation. Multiply cost by a factor of 2 to 3 for attenuation (Plessix, 2017).
- **FDFD**: modeling for 1 frequency. Multiply cost by number of jointly-processed frequencies.

Target	f(Hz)	$\#dof$	$\#sources$	$\#n$ (TD)	$\#n$ (FD)
1	8.55	80	650	325	100
1	8.55	80	3000	1500	100
2	13	78.4	650	325	120

Table 3: $\#sources$: number of reciprocal sources (OBN); $\#n$: number of computer nodes.