A full waveform inversion hybrid method based on short period teleseismic body waves for lithospheric imaging

Dimitri Komatitsch, Vadim Monteiller and Sébastien Chevrot

Laboratory of Mechanics and Acoustics
CNRS, Marseille, France
and Observatoire Midi-Pyrénées, CNRS, Toulouse, France

4th QUEST Meeting, Bénodet, France
May 21, 2013

Adjoint methods for tomography and imaging

Problem is self-adjoint, thus no need for automatic differentiation (AD, autodiff), can be solved with the same code.

$$\chi_1(\mathbf{m}) = \frac{1}{2}\sum_{r=1}^{N_r}\int_0^T w_r(t)||\mathbf{s}(\mathbf{x}_r,t;\mathbf{m}) - \mathbf{d}(\mathbf{x}_r,t)||^2\,\mathrm{d}t,$$

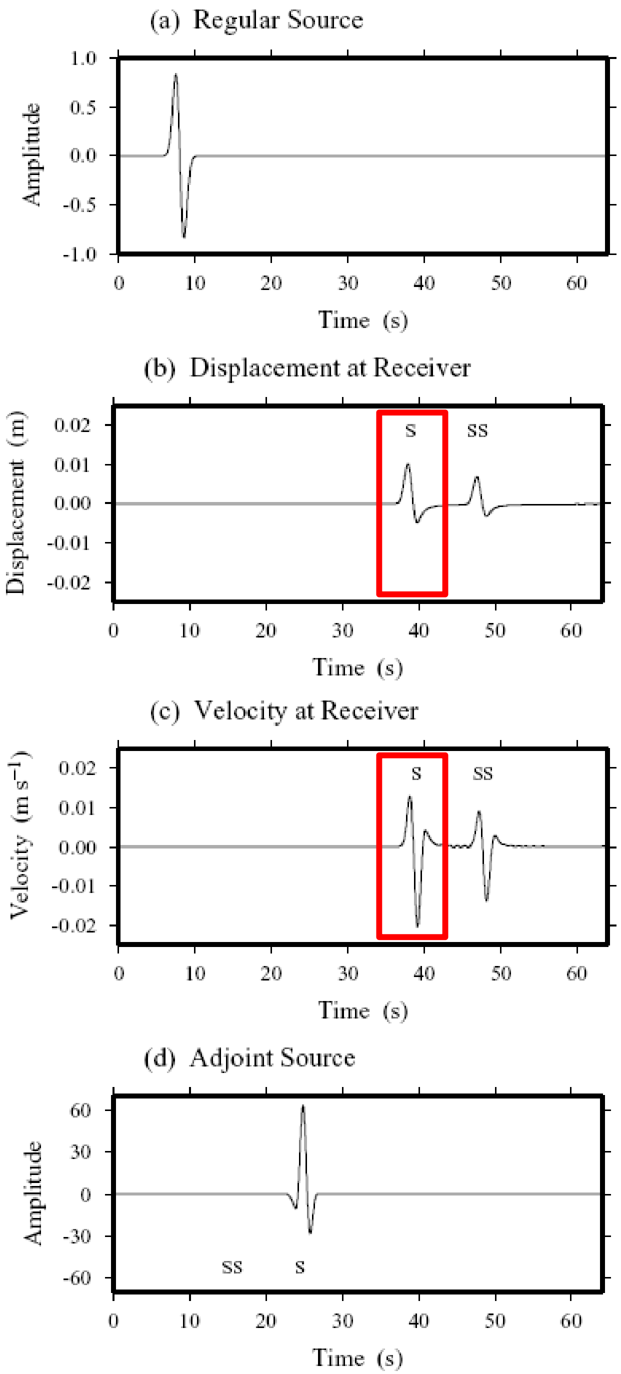
$$\delta\chi_1 = \int_V [K_\rho(\mathbf{x})\,\delta\ln\rho(\mathbf{x}) + K_\mu(\mathbf{x})\,\delta\ln\mu(\mathbf{x}) + K_\kappa(\mathbf{x})\,\delta\ln\kappa(\mathbf{x})]\,\mathrm{d}^3\mathbf{x},$$

$$K_\kappa(\mathbf{x}) = -\int_0^T \kappa(\mathbf{x})\,[\nabla\cdot\mathbf{s}^\dagger(\mathbf{x},T-t)]\,[\nabla\cdot\mathbf{s}(\mathbf{x},t)]\,\mathrm{d}t,$$

Theory: P. Lailly (1983), A. Tarantola (1984), Talagrand and Courtier (1987), ‘Banana-Donut’ kernels (Tony Dahlen et al., Princeton, 2000). Some similarities with time reversal (Mathias Fink et al.).

Apply this to tomography of the full Earth (current ANR / NSF contract with Princeton University, USA), and in acoustic tomography: ocean acoustics, non destructive testing.

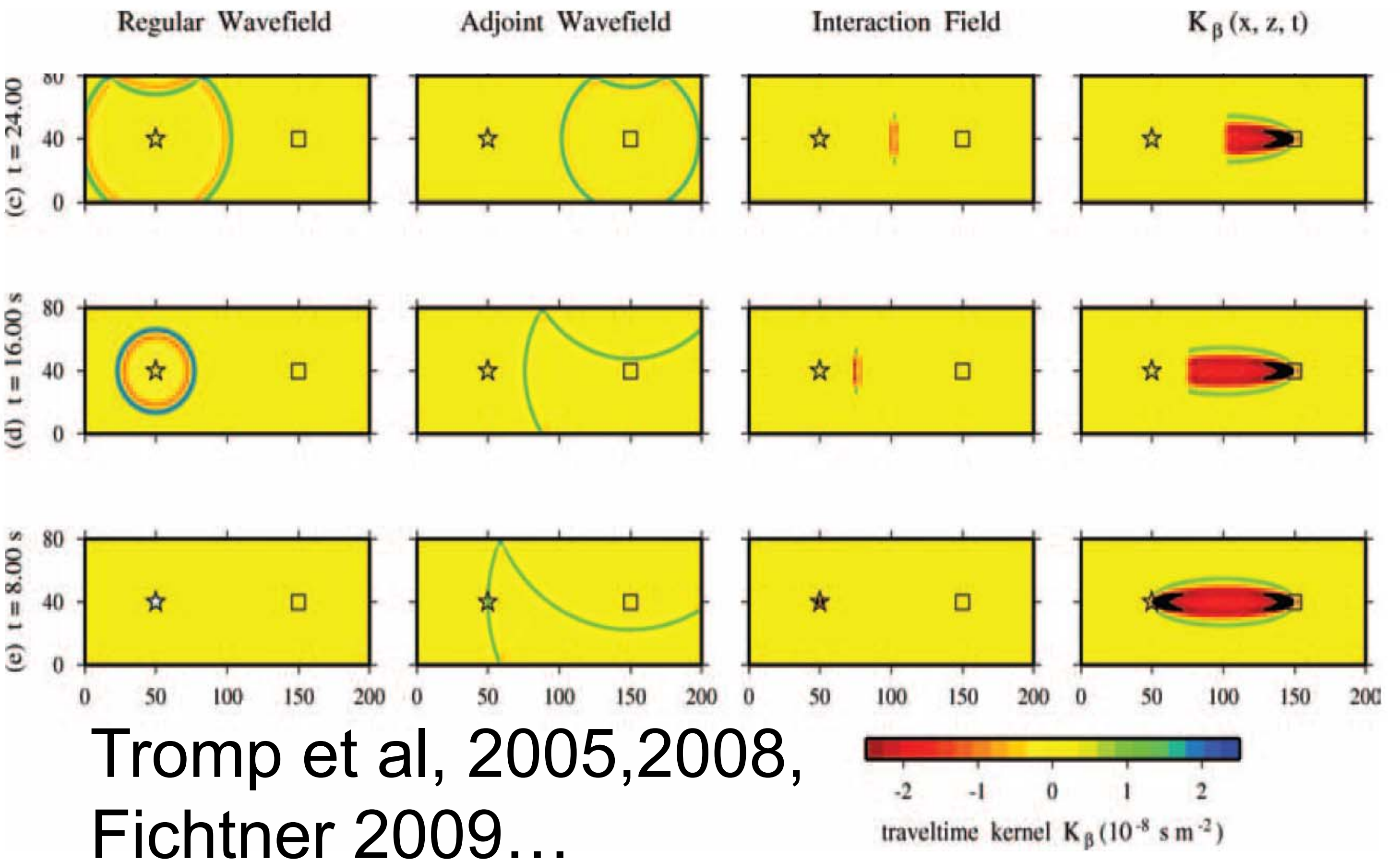
Adjoint method in seismic tomography, `Banana-Donut' kernels



$$\chi(m) = \frac{1}{2}\sum_{r=1}^{N}\left[T_r(m) - T_r^{\mathrm{obs}}\right]^2, \qquad \delta\chi = \sum_{r=1}^{N}\left[T_r(m) - T_r^{\mathrm{obs}}\right]\delta T_r.$$

$$\delta T_r = \int_V \left[\bar{K}_\rho(\mathbf{x},\mathbf{x}_r)\delta\ln\rho(\mathbf{x}) + \bar{K}_\mu(\mathbf{x},\mathbf{x}_r)\delta\ln\mu(\mathbf{x}) + \bar{K}_\kappa(\mathbf{x},\mathbf{x}_r)\delta\ln\kappa(\mathbf{x})\right] d^3\mathbf{x}.$$

$$\bar{K}_\kappa(\mathbf{x},\mathbf{x}_r) = -\int_0^T \kappa(\mathbf{x})\left[\nabla\cdot\bar{\mathbf{s}}^\dagger(\mathbf{x},\mathbf{x}_r,T-t)\right]\left[\nabla\cdot\mathbf{s}(\mathbf{x},t)\right]dt,$$



Tromp et al, 2005,2008, Fichtner 2009…

Imaging at the global scale

- Improving images of the lower mantle (D", super-plumes)
  → Pdiff sensitivity kernels

- Improve the resolution of images of the mantle
  → full waveform imaging

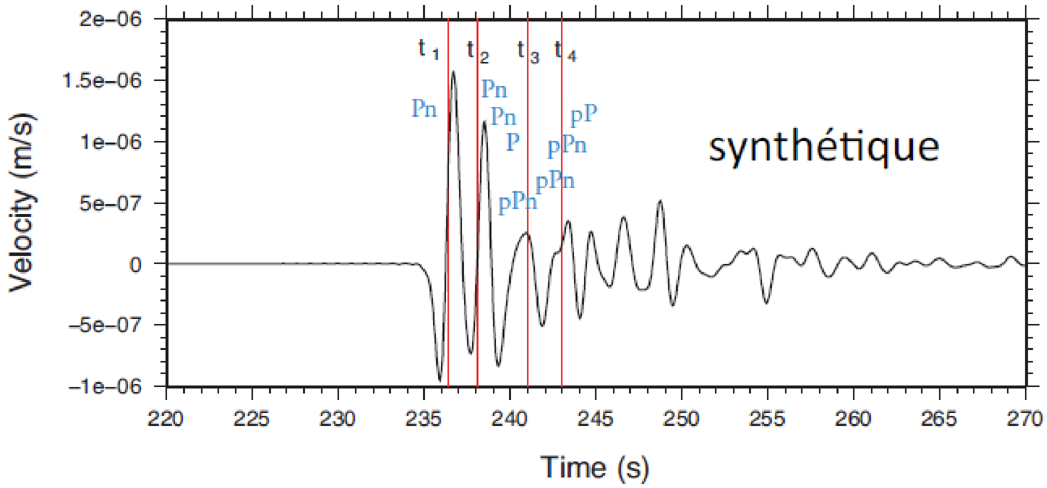
Pdiff and PKP (caustics) do not work with classical ray tracing;
and normal modes are very difficult below 5 to 8 s. Thus resort to DSM.

- Tools : DSM (Geller & Ohminato 94, Takeuchi and Kawai)
  → synthetic seismograms at 1s in a 1D global Earth model
  → sensitivity kernels
  → partial derivatives

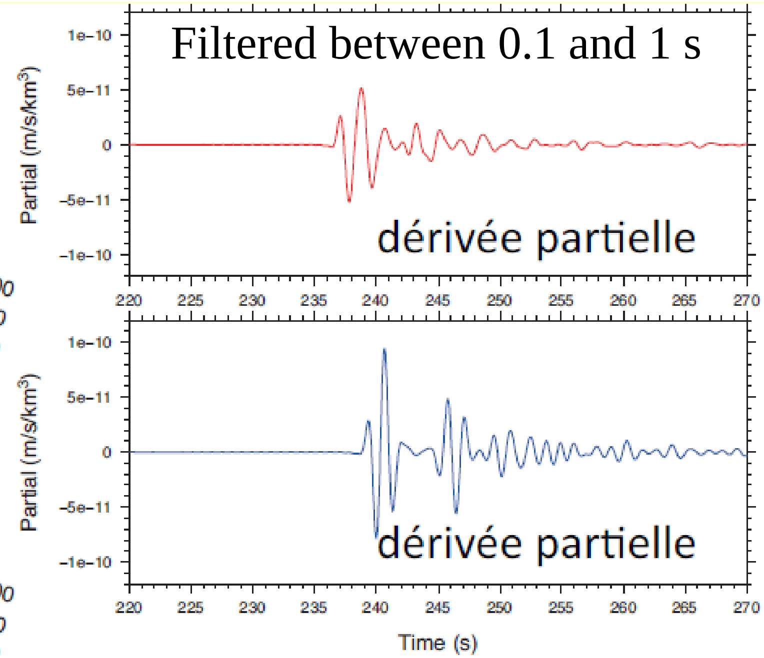
Database of such full-waveform short-period kernels for 1D models
built at CNRS Toulouse (France) by Fuji and Chevrot using DSM.

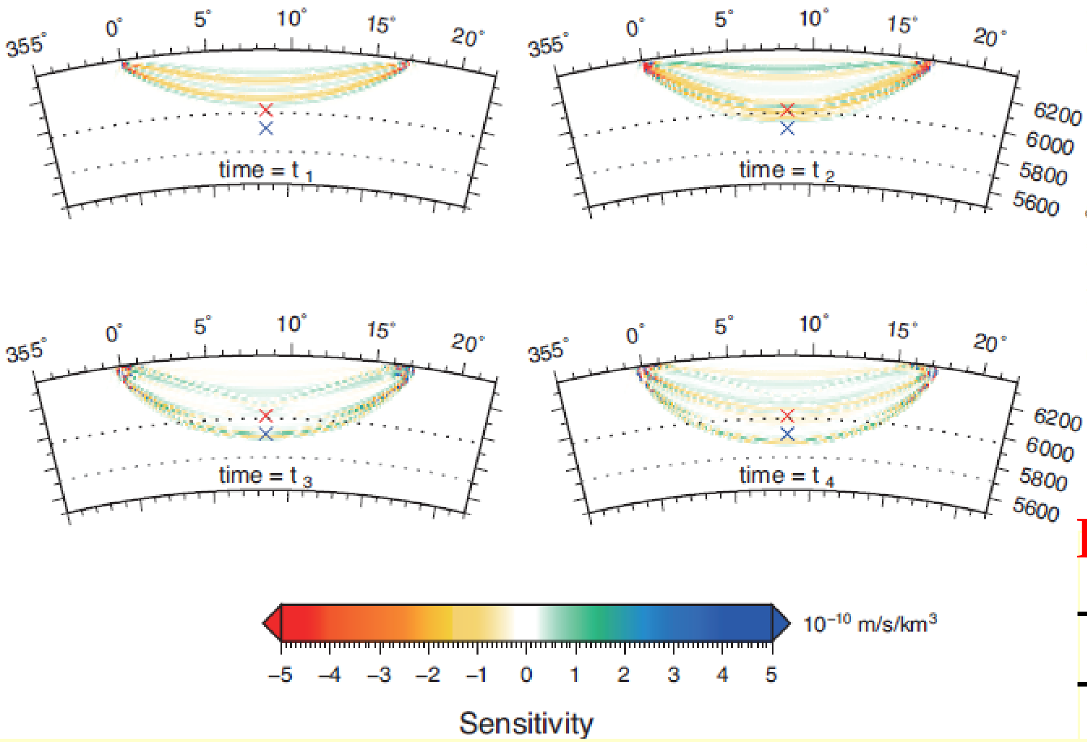
Sensitivity kernels for full waveform at high frequency

Sensitivity kernels for full waveform: 1 kernel per time step (Fuji et al, 2012)



Partial derivatives in 3D at short period (1s) but for a 1D model using DSM





Big limitations:

- Tens of terabytes of storage
- 1D reference background model

Motivation to go toward full waveform inversion for imaging

1/ Classical « low cost » migration (e.g. Reverse Time Migration, widely used in the oil industry): we get the interfaces (wave speed jumps) if the geometry is not too distorted / not too tilted; thus difficult to use in mountain ranges or subduction zones for instance; and we need Vp and Vs → mineralogical content, nature of the materials

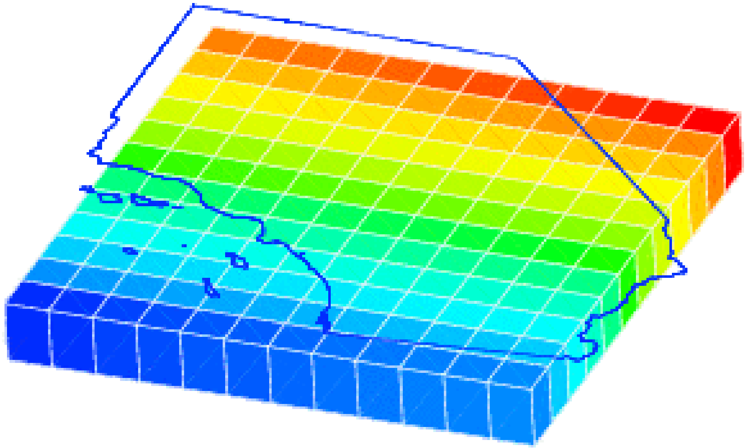
2/ If we want to image deep structures, we would like to be able to use short-period body waves; surface waves at short periods will only sample part of the crust but not deeper mantle structures; thus we would like to target e.g. P waves at ~ 1 s.

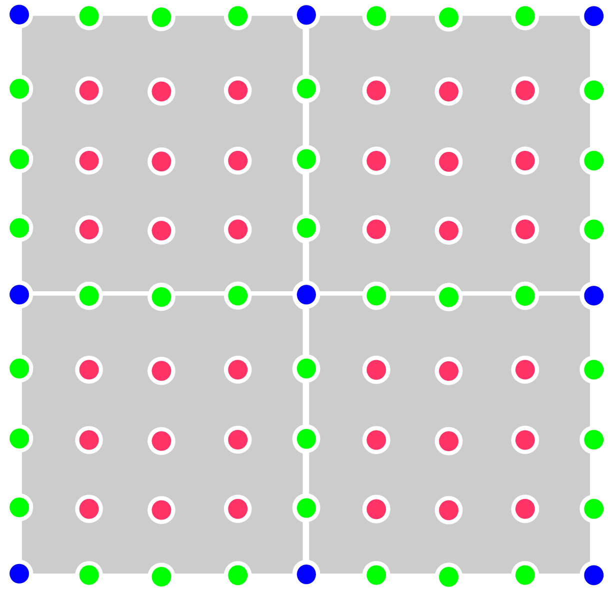
Thus we will need dense regional networks.

3/ Widely used traveltime tomography limited by width of the Fresnel zone, while full waveform inversion (FWI) can theoretically go to half the wavelength; thus for instance in the Pyrénées traveltime tomography would give a resolution of a few tens of km, while FWI with P and S waves at 1 s could maybe reach a few km.

Price to pay is huge technical difficulties: for instance using SPECFEM3D at 1 s in the full Earth is barely possible on the largest supercomputers in the world for a single (forward) run, thus solving inverse problems is not an option for the next 10+ years.

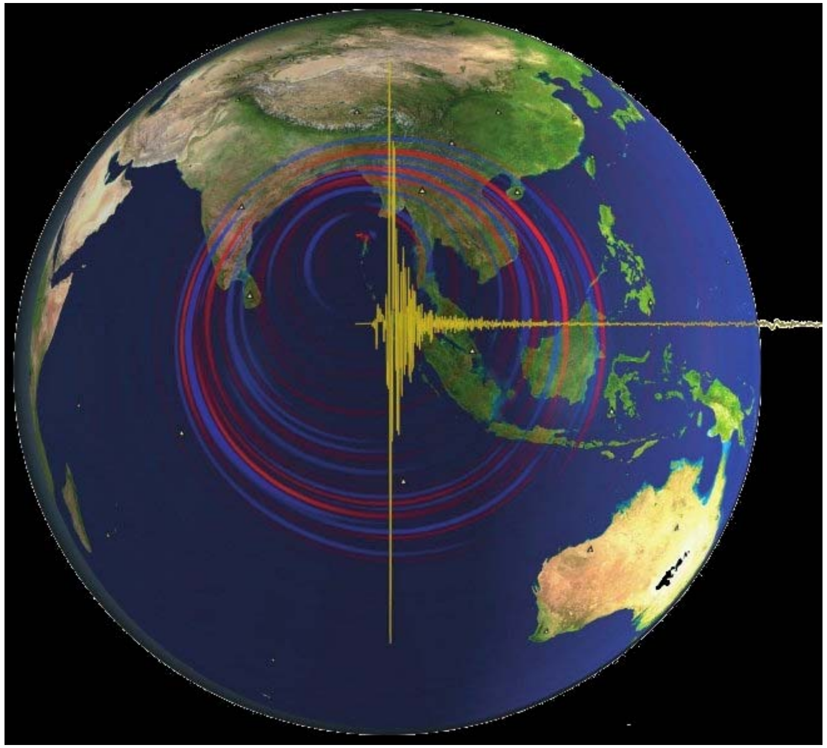
High-performance computing





Our goal is to use moderate-size clusters that are easily available in research labs rather than very big Tier-0 systems (not so usual for us ☺).

Our SPECFEM3D software package



Goal: model acoustic / elastic / viscoelastic / poroelastic / seismic wave propagation in the Earth (earthquakes, oil industry), in ocean acoustics, in non destructive testing…

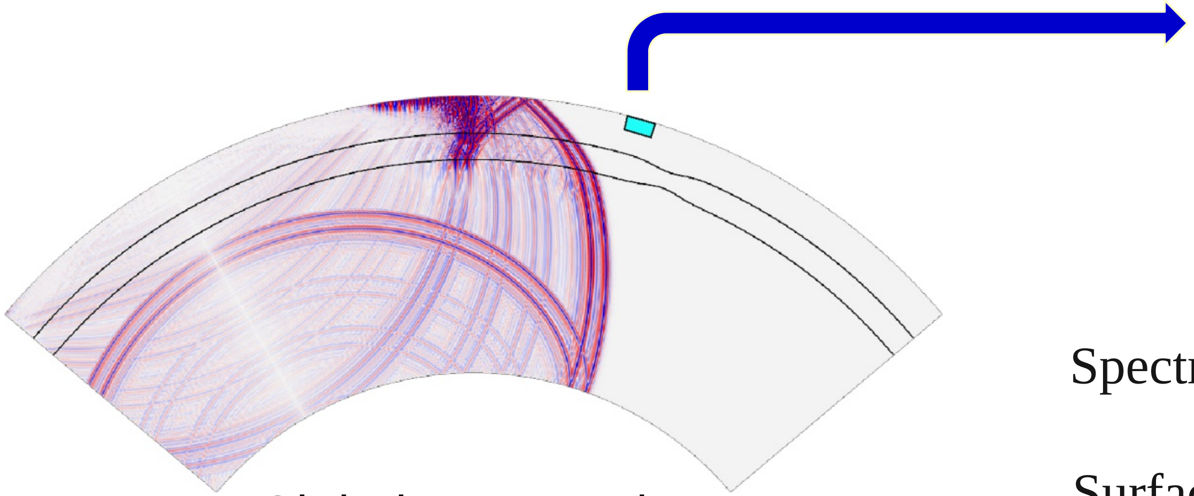
The SPECFEM3D source code is open (GNU GPL v2)

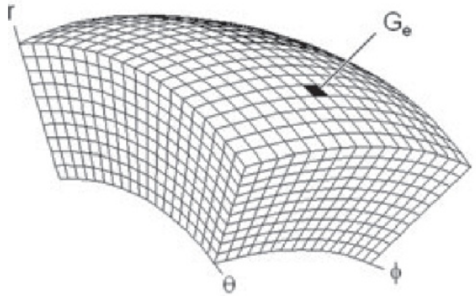
Initially IPG Paris (France), mostly developed by Dimitri Komatitsch and Jeroen Tromp at Harvard University, then Caltech, Princeton (USA) and University of Pau / CNRS (France) since 1996.

Improved with INRIA (Pau, France), CNRS (Marseille, France), the Barcelona Supercomputing Center (Spain), University of Basel (Switzerland), NVIDIA…

A hybrid approach: Coupling global and regional propagation

A hybrid technique for 3-D waveform modeling and inversion of high-frequency teleseismic body waves





Regional propagation
in a 3-D spherical shell

Spectral elements (Komatitsch et al, 2002)

- Surface topography
- Interface topography
- Isotropic or anisotropic heterogeneities
- Attenuation

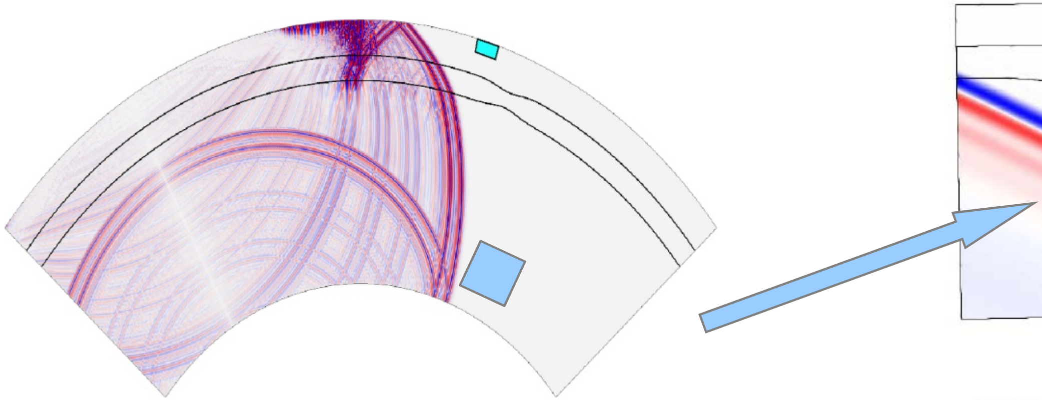
Global propagation
in a spherically-symmetric Earth model
DSM technique: simulations up to1.25 Hz
(Geller et al, 1996)

→ Drastically reduce the size of the region to model

→ Fast 3D modeling of short-period teleseismic body waves (period ~ 1s)

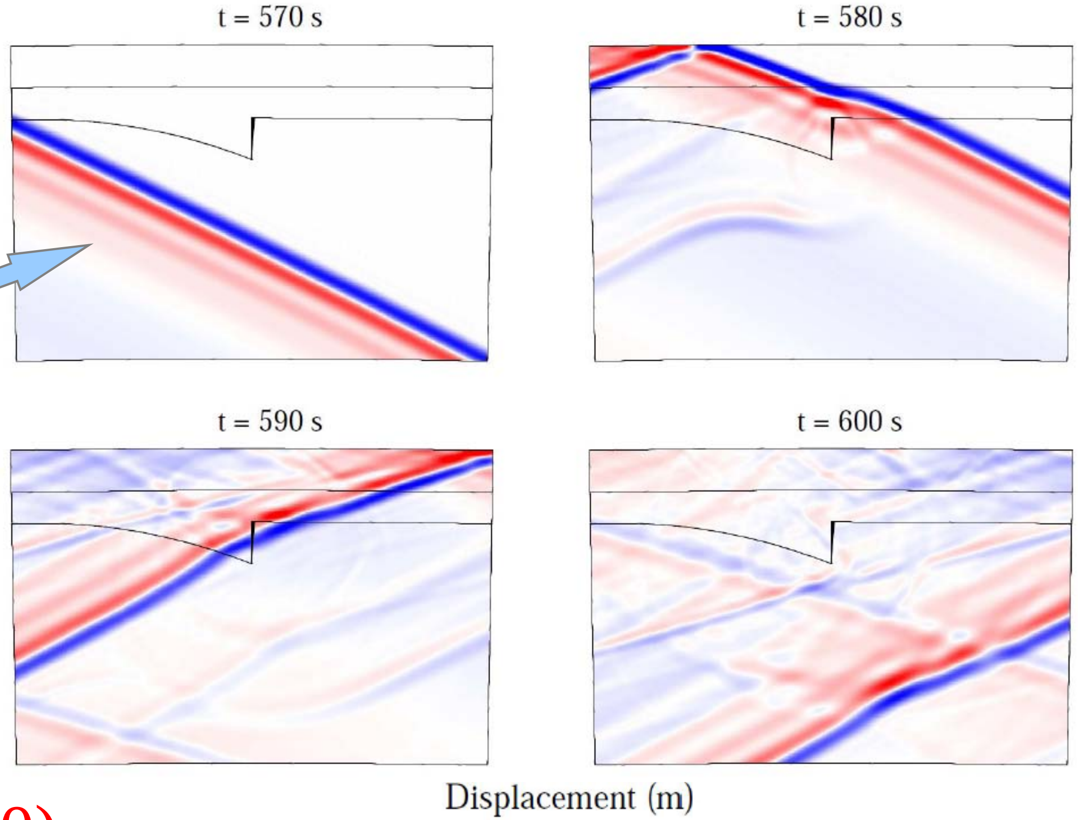
→ We had to speedup DSM by 40x (otherwise very expensive)

Local 3D modeling of teleseismic waveforms based on spectral elements



1D propagation at the global scale (DSM)

Some similarities with the coupling of Capdeville et al. (2000) with normal modes, but DSM can go beyond.

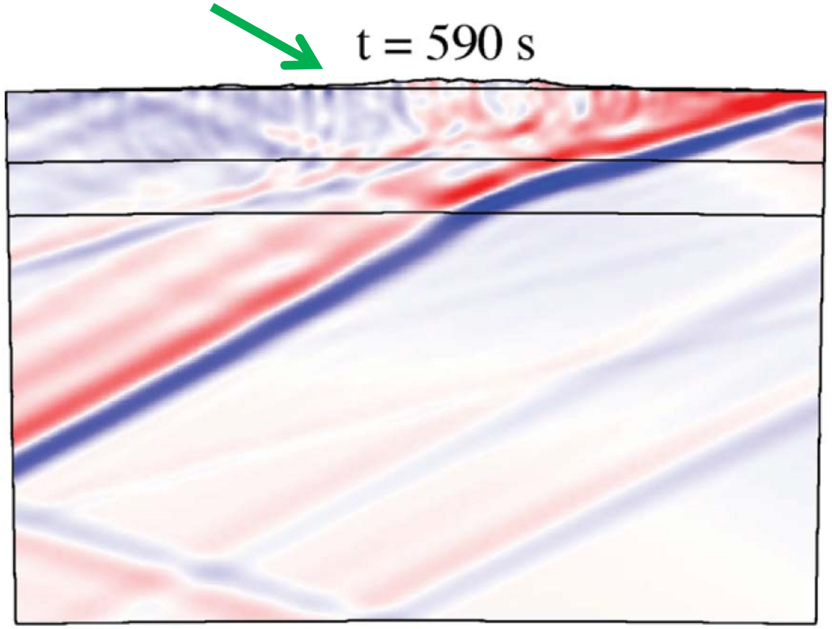


3D propagation at the regional scale (spectral elements)

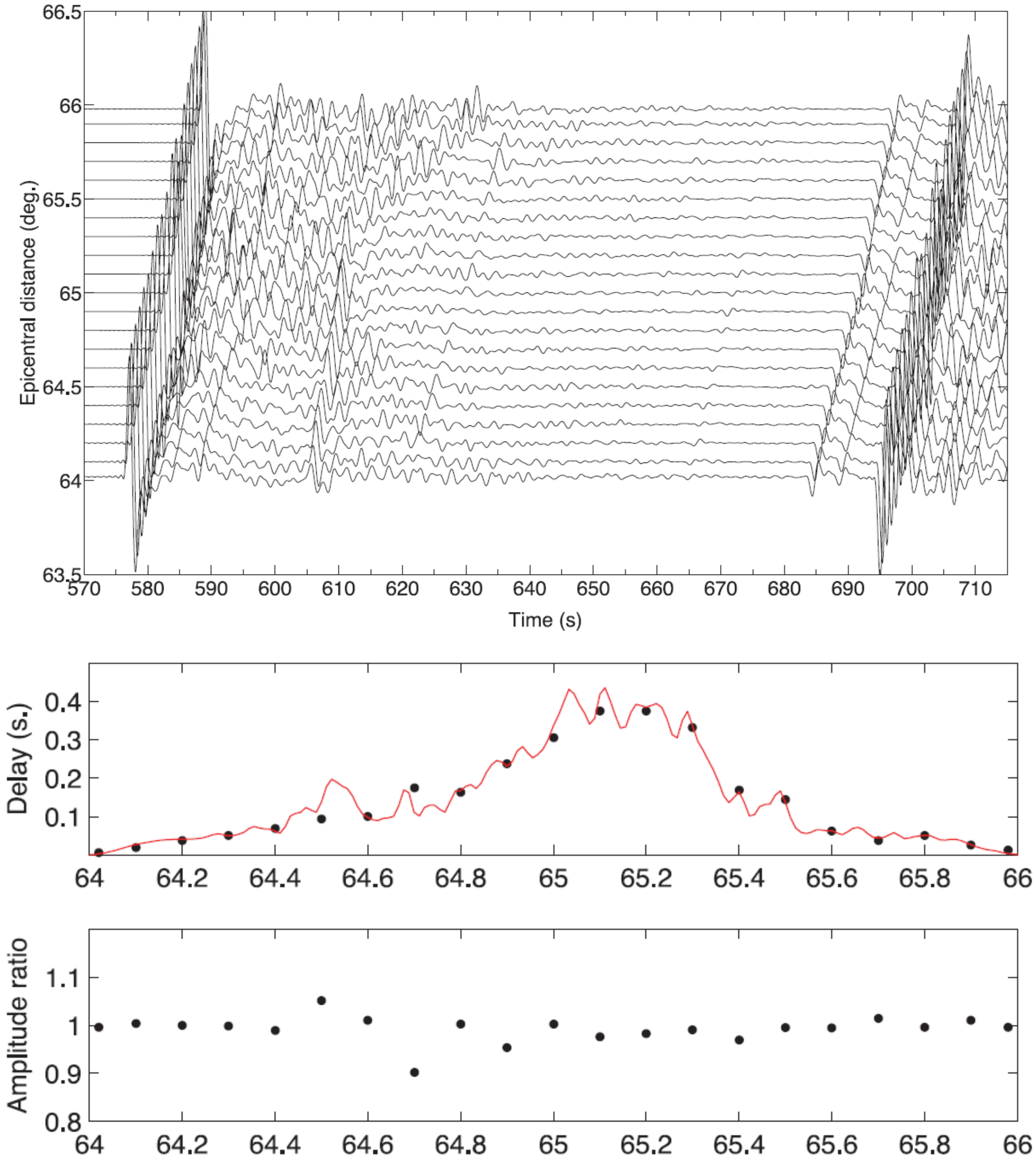
A 3D run takes less than 1 minute on 512 CPU cores (mostly I/Os, will be improved), thus inexpensive. ***Monteiller** et al, GJI 2013*

We can thus use many sources and perform imaging.

We can take the effect of topography into account



Effect of topography

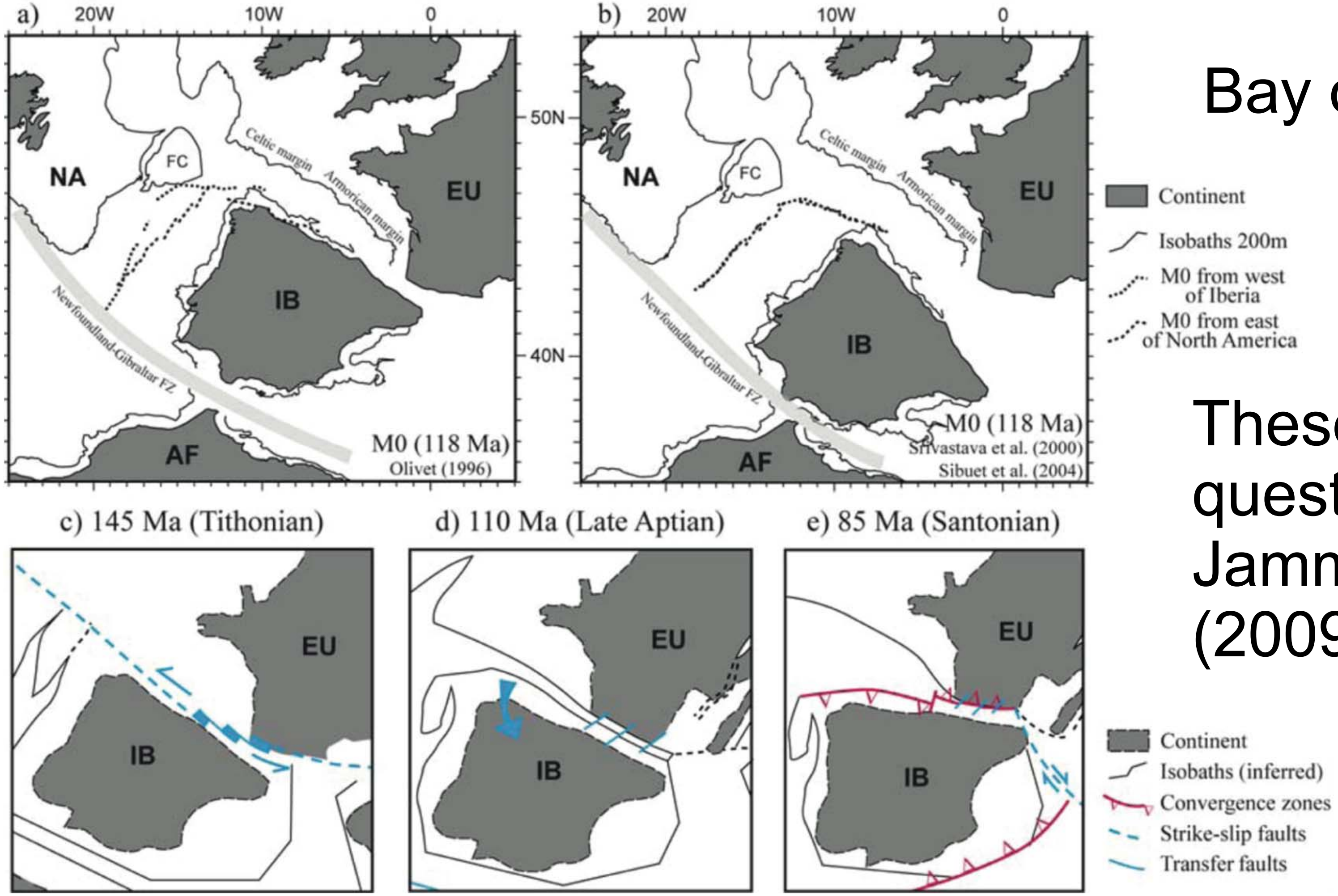


Effect on traveltime and on amplitude

Focusing / defocusing effect on multiples (change of traveltime and amplitude)

Interesting for imaging based on multiples + conversions P → surface waves

Models of Golfe de Gascogne opening and Pyrénées mountains creation

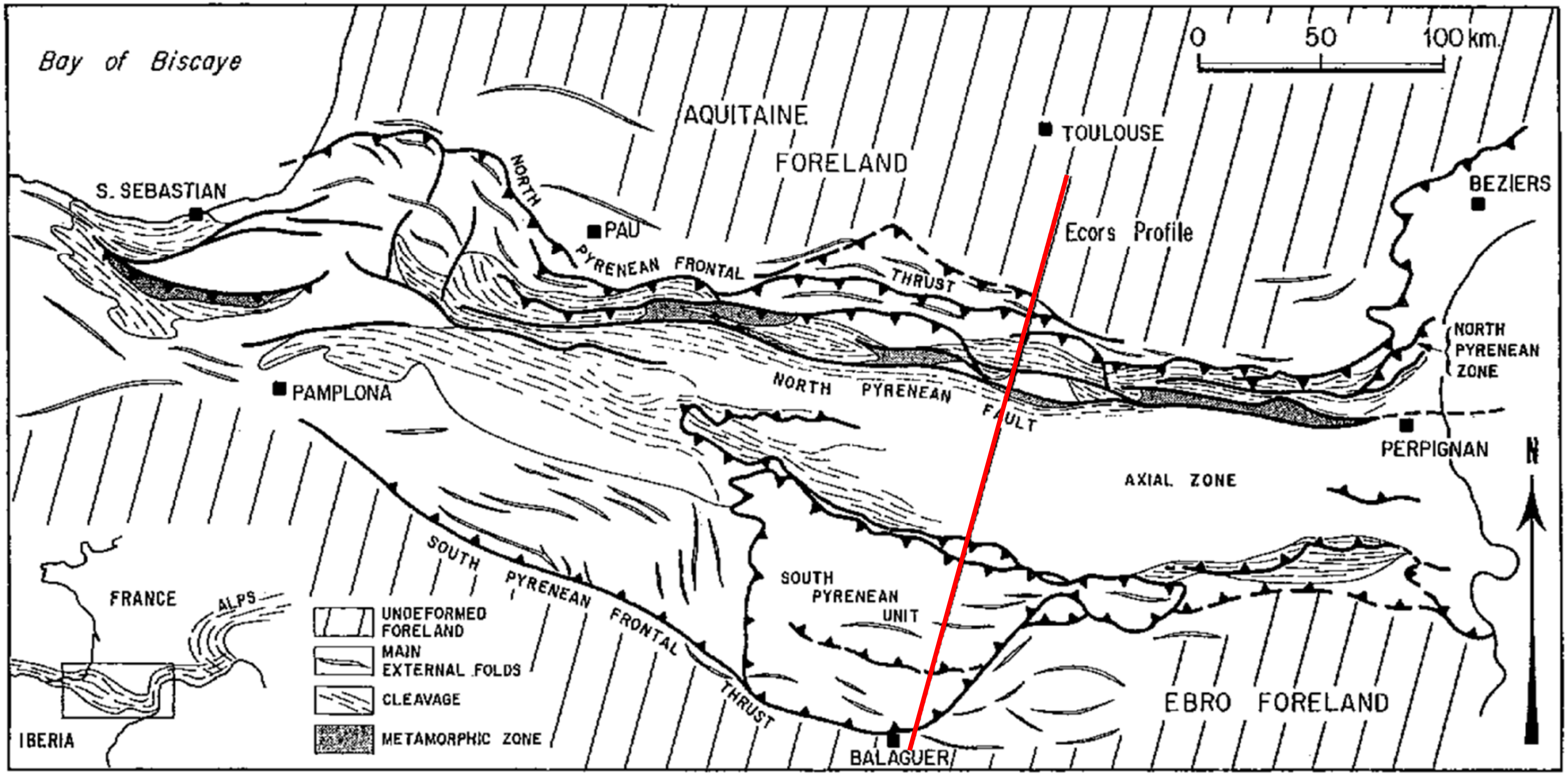


Bay of Biscay

These are current questions: e.g. Jammes et al. (2009)

- ❑ The nature, timing and location of relative movements between Iberia and Europe during the Cretaceous period are still debated
- ❑ Deep structure imaging could provide crucial constraints to discriminate between these models

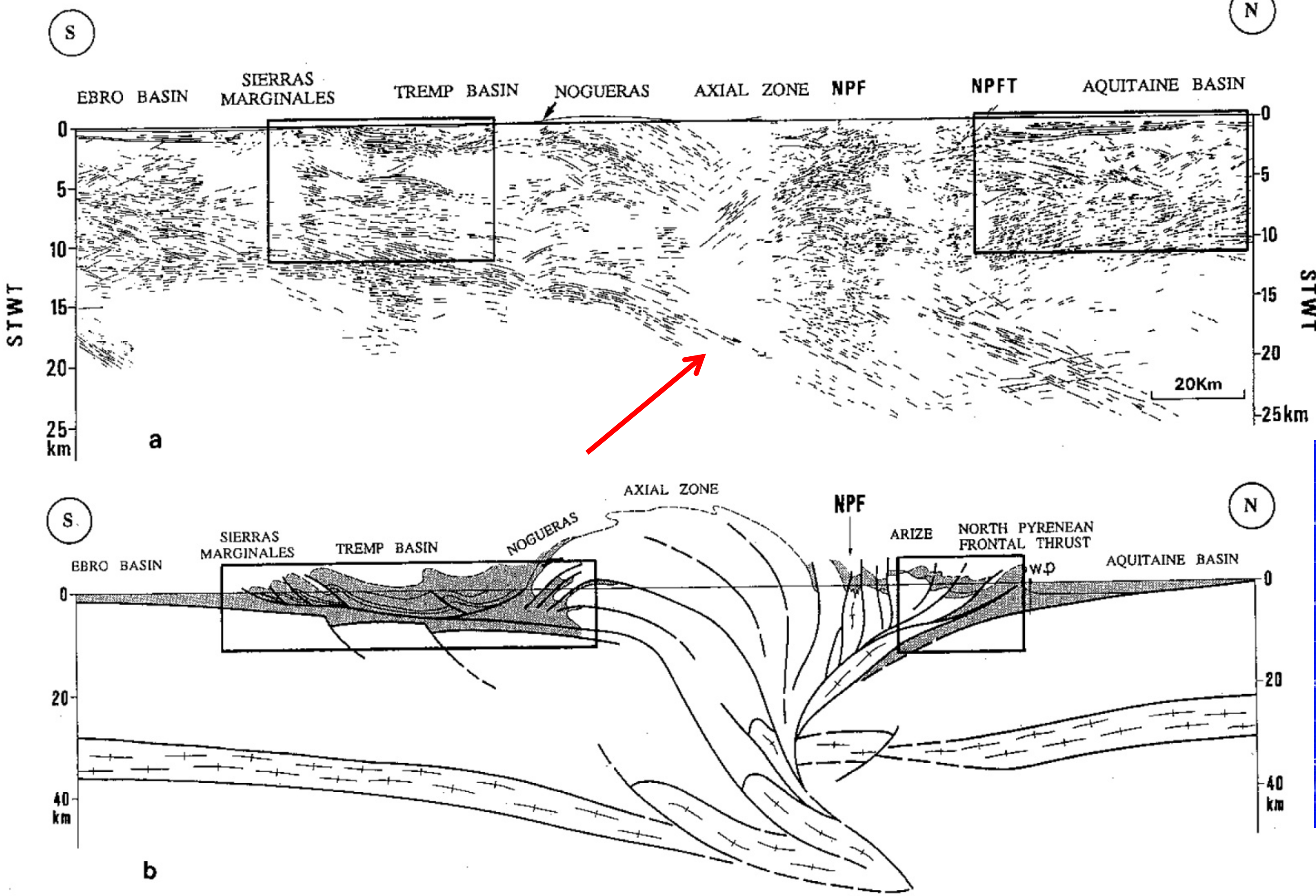
The ECORS seismic profile (1985-1986)

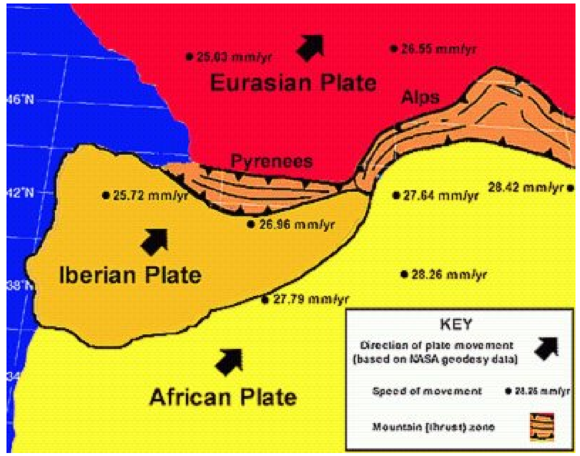


ECORS team, Nature, vol. 331 (1988)

Roure & Choukroune (1989)

Interpretation by Choukroune et al. (1989)

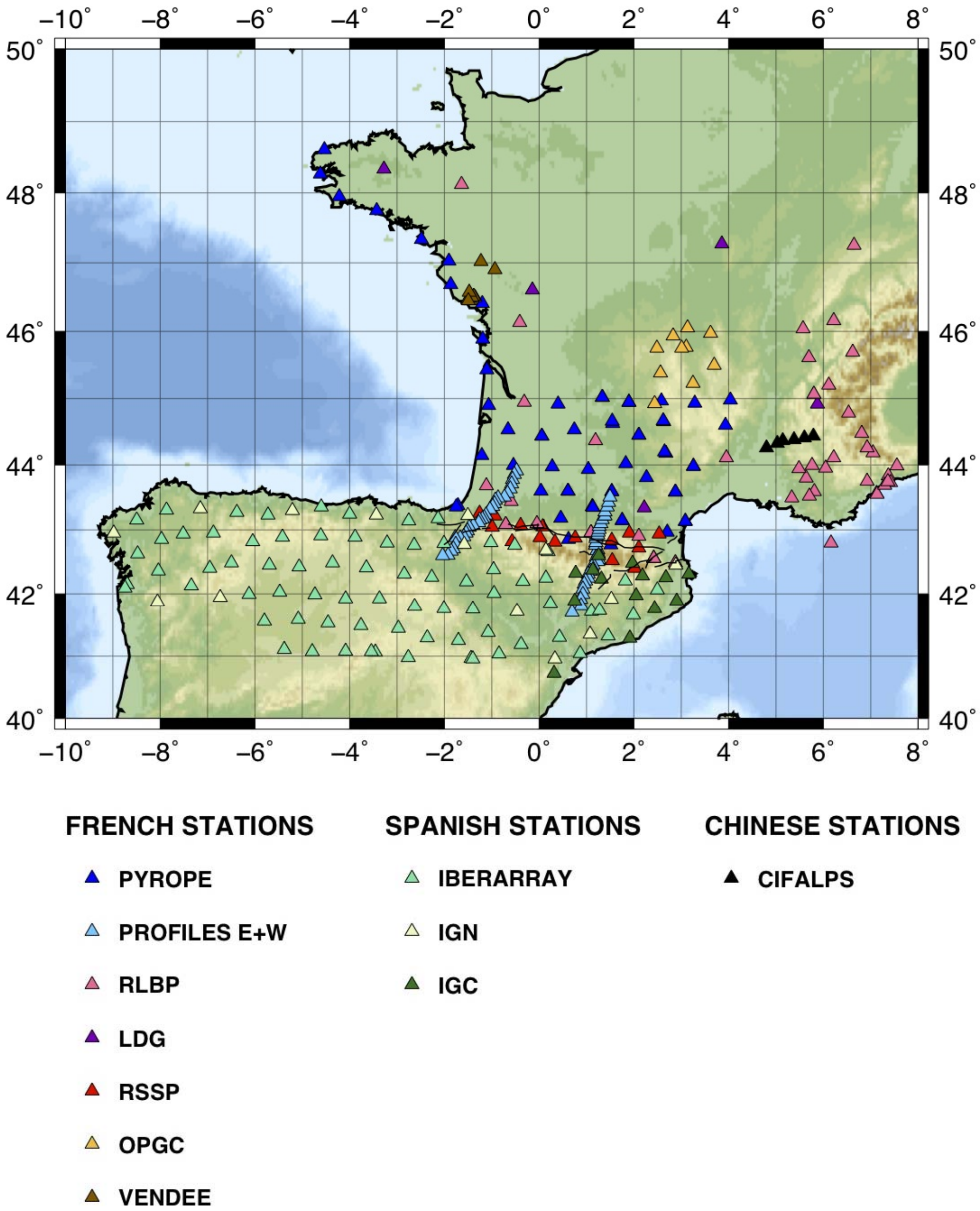




Back then they saw the Iberian plate going under the Eurasian plate (Nature, 1988)

Current question: what is the detailed structure in the lower part?

The PYROPE + IBERARRAY experiments



- French/Spanish initiative, supported by the French ANR
- ~150 temporary + 50 permanent broadband stations (~ 250 stations in total)
- Interstation spacing ~ 60 km
- Dense transects across the Pyrénées

Main goal: image the lithosphere (deep structures) under the Pyrénées based on teleseismic full waveform inversion

Gauss-Newton method

- Minimize the misfit between data and synthetics

$$J(m)=\frac{1}{2}\sum_r \int_0^T (s(x_r,m,t)-d(x_r,t))^2 dt+(m-m_0)^T C_m^{-1}(m-m_0)$$

→ We look for a model that cancels the gradient of the cost function

$$\nabla J(m_e)=0$$

- Gauss-Newton iterative algorithm

$$m_{k+1}=m_k+(G_k^t G_k+C_m^{-1})^{-1}\nabla J(m_k)$$

G: matrix of the Fréchet derivatives (kernels) of $m \rightarrow s(x_r,m,t)$

Problem : for full waveform inversion, in practice the G matrix is far too big

BFGS method

- Iterative Gauss-Newton algorithm $m_{k+1} = m_k + \left(G_k^t G_k + C_m^{-1}\right)^{-1} \nabla J(m_k)$

  δm_k

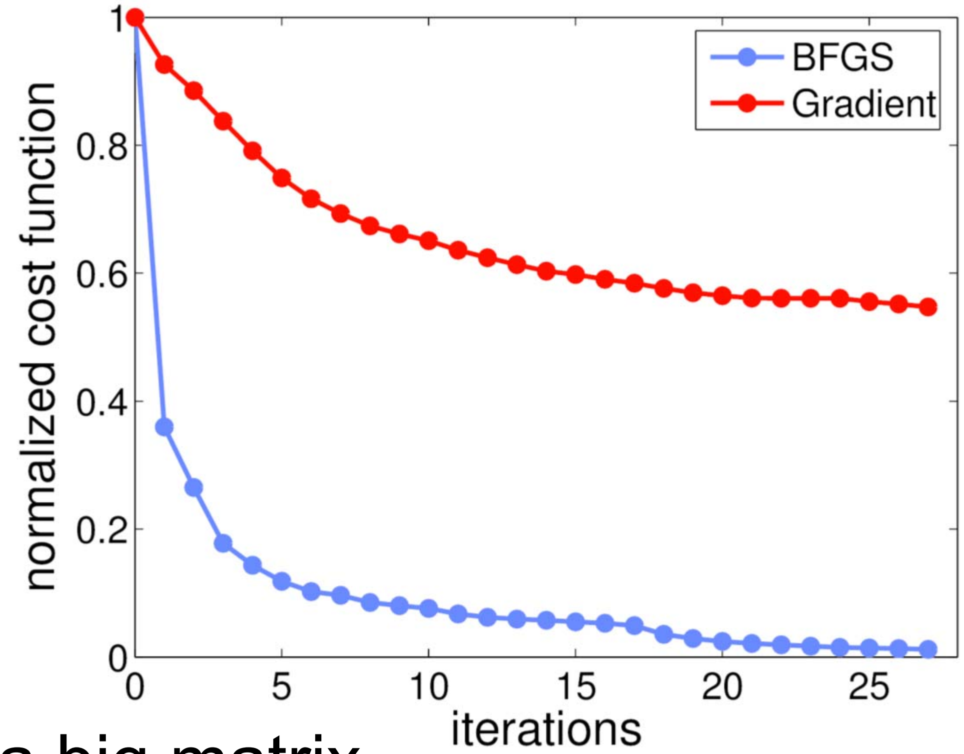
- BFGS (Broyden–Fletcher–Goldfarb–Shanno): approximate

  δm_k

  from:

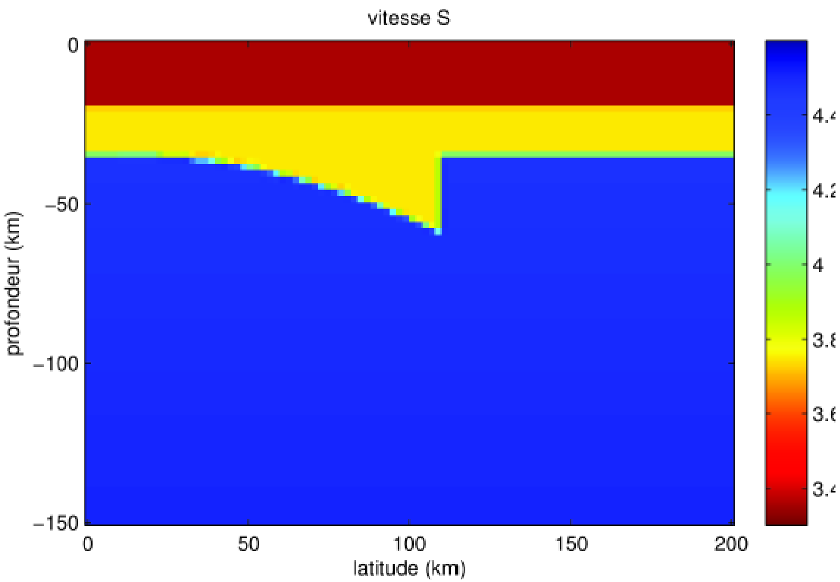
  $m_{k-1}, m_{k-2}, m_{k-3}, \ldots, m_0$

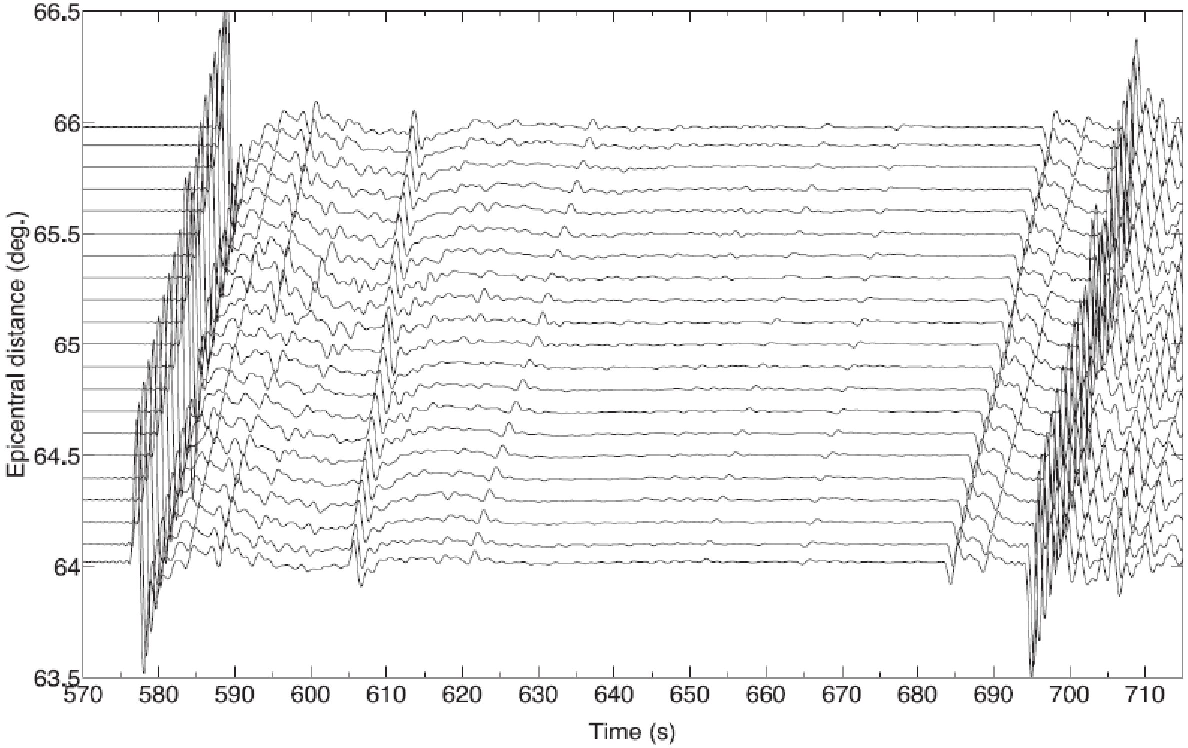
  $\nabla J(m_{k-1}), \nabla J(m_{k-2}), \ldots, \nabla J(m_0)$



→ no need to invert or even build a big matrix

Synthetic example: Moho jump

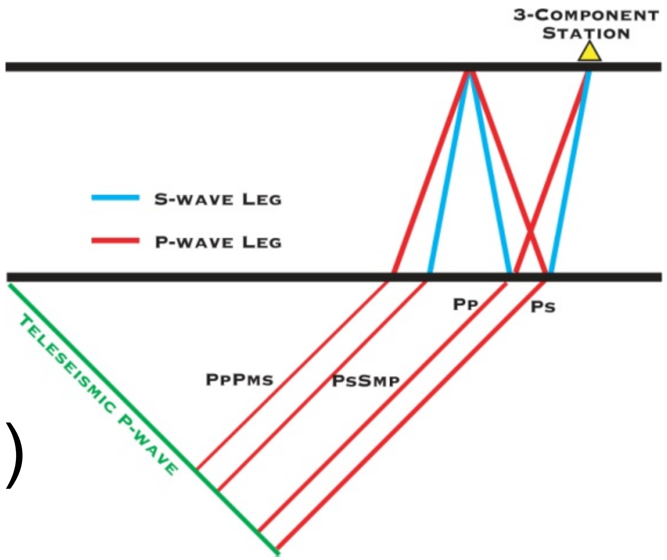




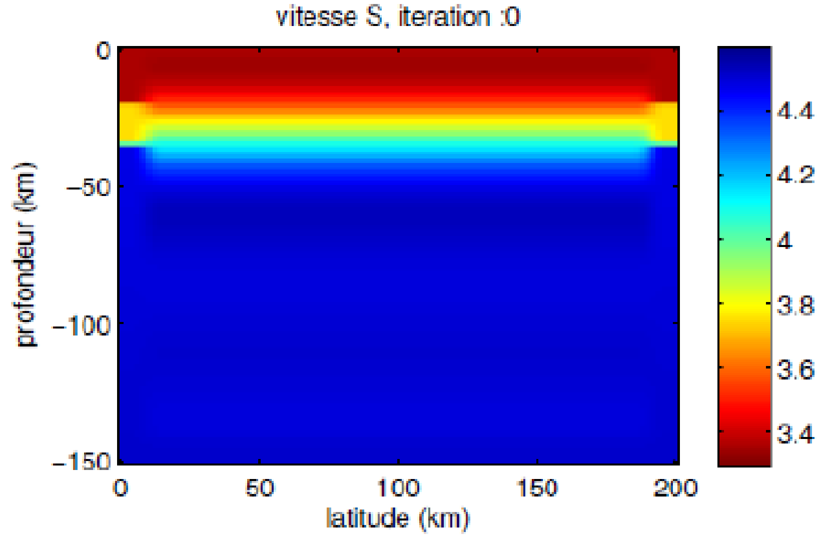
Full wavefield modeling includes:

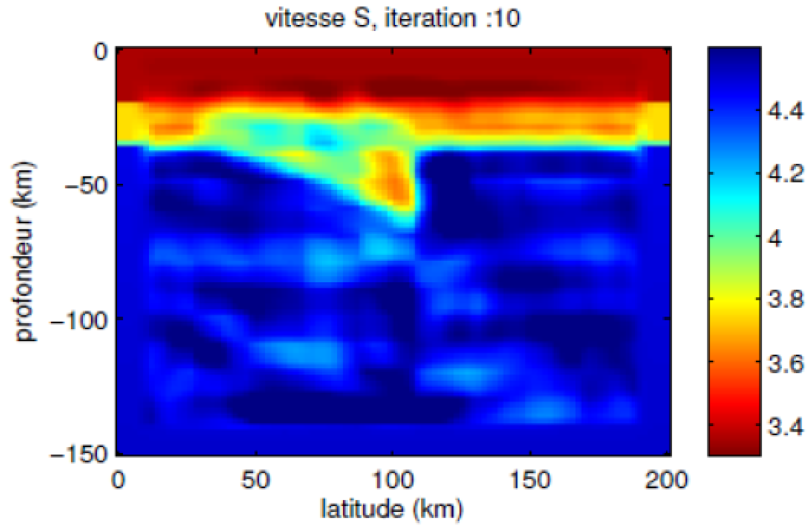
- direct P waves
- converted waves
- reflected waves

→ better constraints (e.g. on thickness)

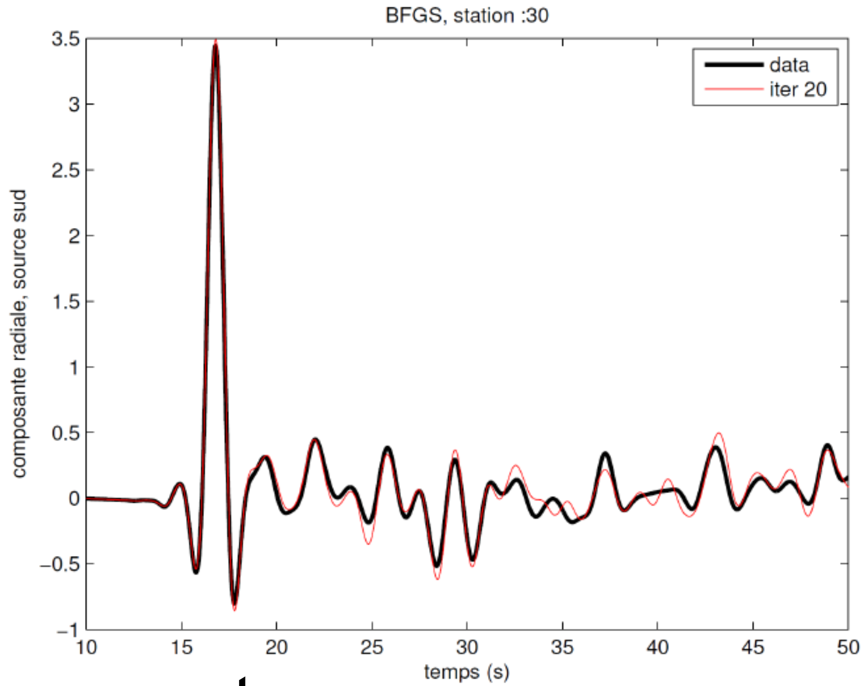


Full waveform inversion: preliminary results





Smooth initial model; tapering to match ak135 for DSM (smoothing only inside, keeping vertical slowness i.e. location of discontinuities)



We get the discontinuity at 20 km and the vertical Moho jump back

Position of the interfaces → geometry

We get the P and S velocities → type of material

Current work: Undoing attenuation

- Constitutive relationship:

$$\mathbf{T}(t) = \int_{-\infty}^{t} \partial_t \mathbf{c}(t - t') : \nabla \mathbf{s}(t')\, \mathrm{d}t'$$

Difficult in time domain methods because of convolution

- Use L Zener body standard linear solids to make an absorption-band model:

$$\mu(t) = \mu_R \left[1 - \sum_{\ell=1}^{L} \left(1 - \tau_\ell^\varepsilon / \tau_\ell^\sigma\right) e^{-t/\tau_\ell^\sigma} \right] H(t)$$

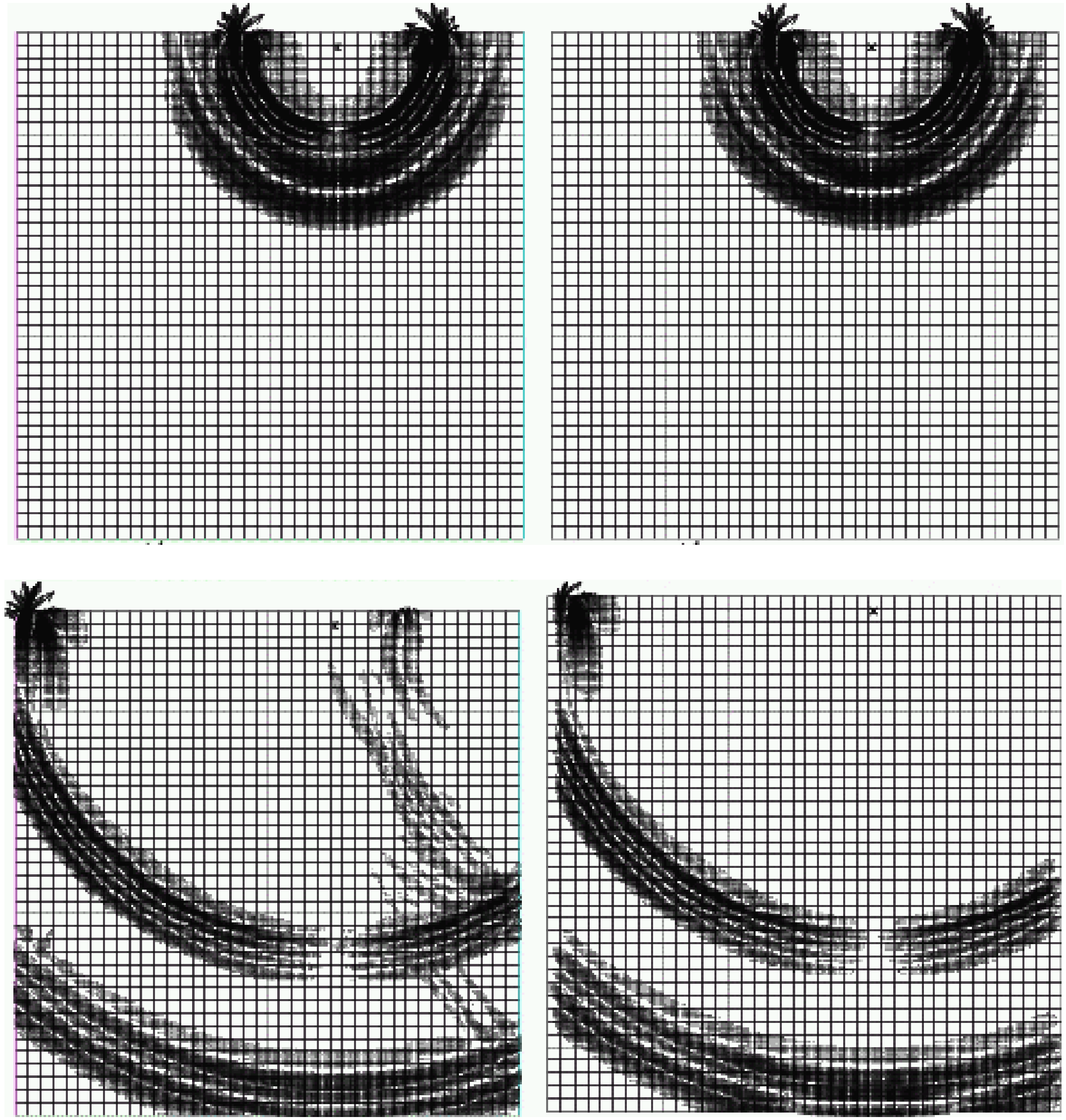
Current work: Undoing attenuation

- Several options:

  - Store all the time steps of the forward run: we do this in 2D, but currently impossible (unrealistic / too expensive) in 3D (hundreds of terabytes of storage to disk for a single run)
  - Gear et al. (2004): go back and forth (20 steps backwards, 5 steps forward, 20 steps backwards, 5 steps forward…) to stabilize: works fine only for slowly-evolving systems (*NOT* our case)
  - Filtering / regularization to stabilize the system: uneasy for Zener attenuation bodies, and requires too heavy filtering $\Rightarrow$ accuracy is lost and thus the sensitivity kernels are inaccurate
  - $\Rightarrow$ current idea: use only partial storage; I have started to test that, more about this in a few months

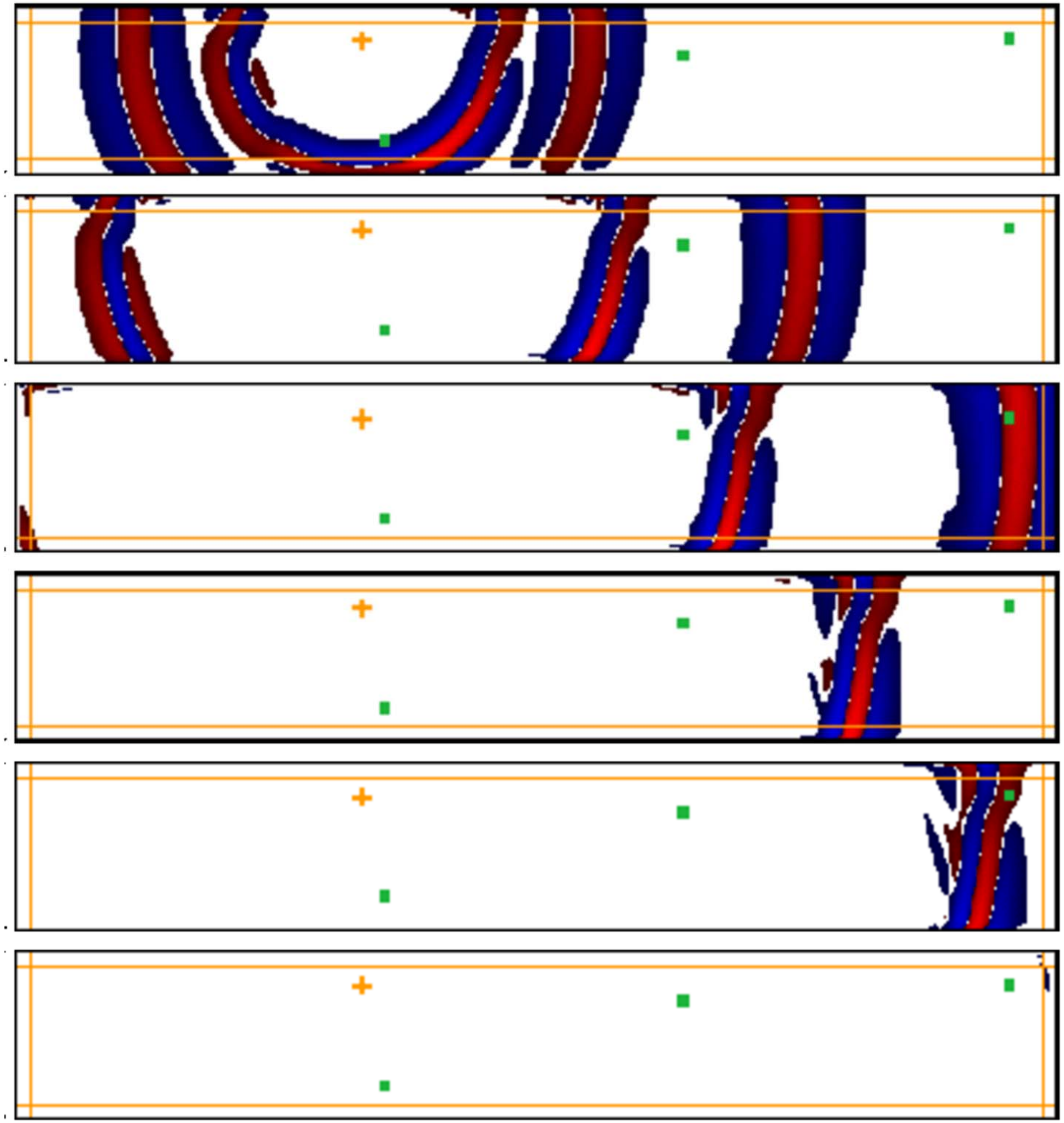
Reducing the cost: absorbing conditions

- Used to be a big problem
- Bérenger (1994)
- INRIA (Collino, Cohen)
- Extended to second-order systems by Komatitsch and Tromp (2003)
- C-PML (Convolution Perfectly Matched Layer)



With Stacey (1988) With C-PML

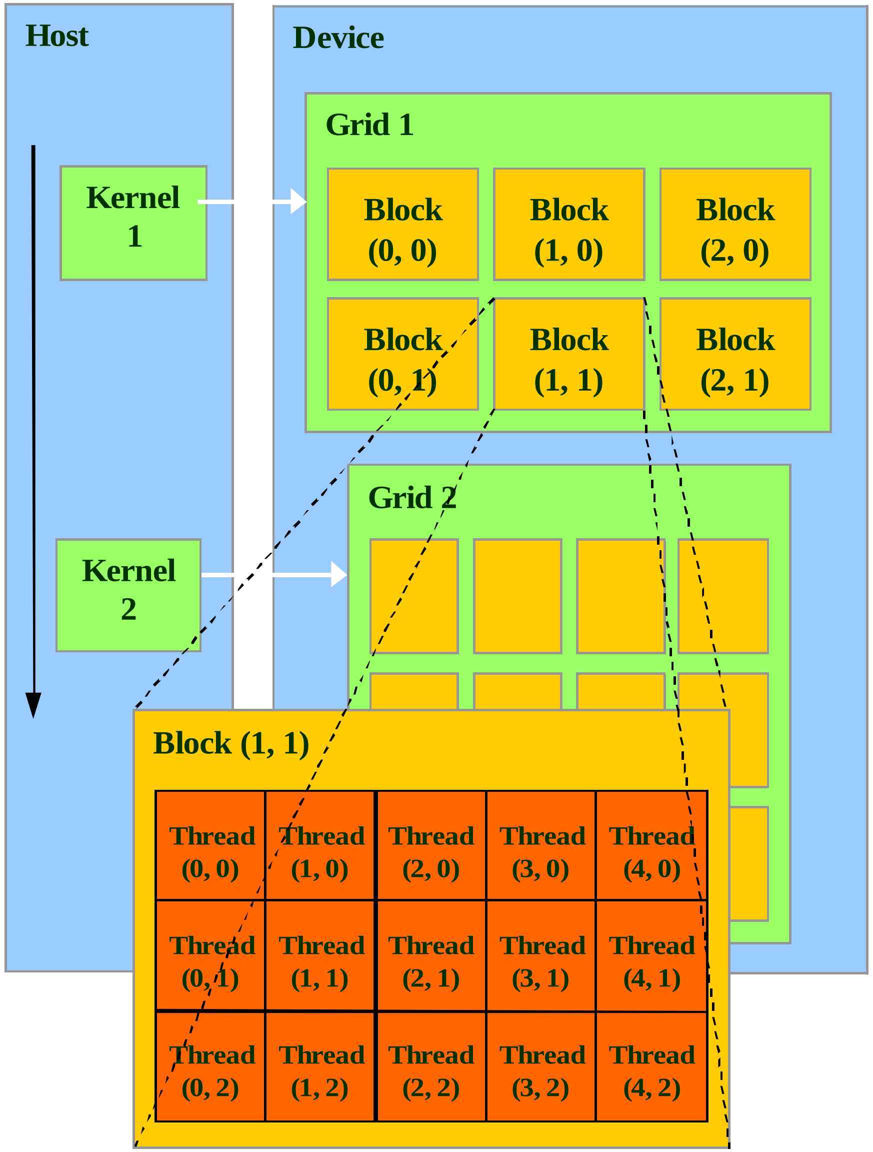
Convolution-PML for seismic waves



- **Optimized for grazing incidence**
- **Not split**
- **Use recursive convolution based on memory variables (Luebbers and Hunsberger 1992)**
- **Thin slices: « 3D at the cost of 2D », Komatitsch and Martin (2007).**

GPU graphics cards

NVIDIA



Why are they so powerful for scientific computing?

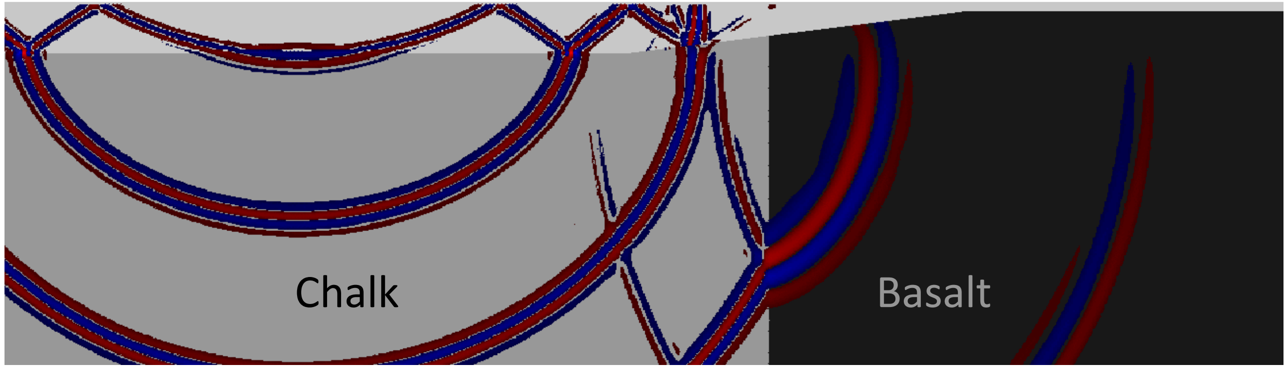
Compute all pixels simultaneously, massive multithreading.

Ocean acoustics

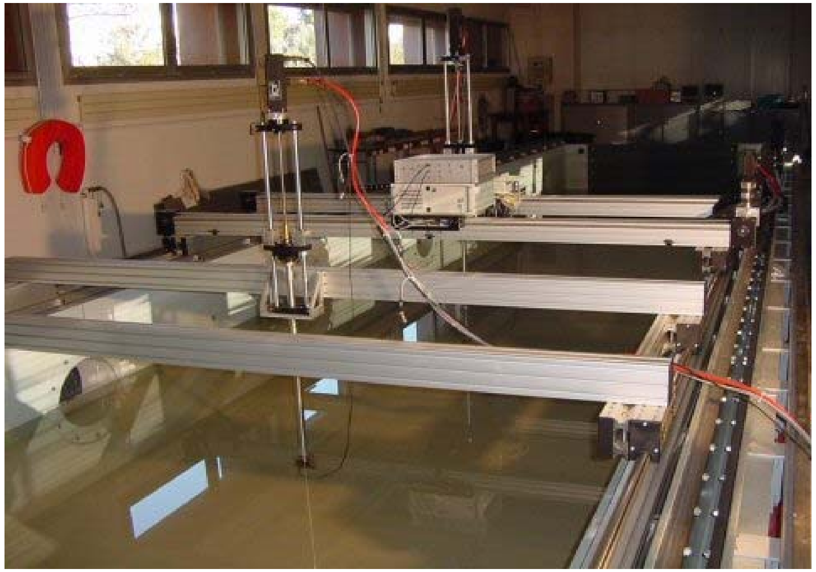
Numerical simulation Collaboration with Paul Cristini and Mark Asch (CNRS).

Wave propagation across an impedance discontinuity.

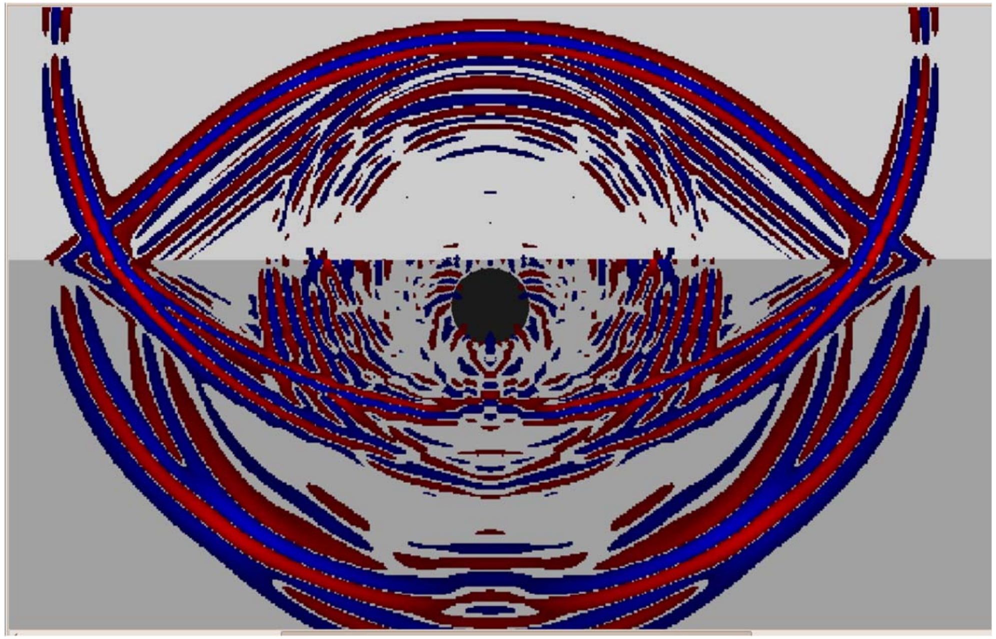
Influence on interface waves.



Experiments performed in tanks



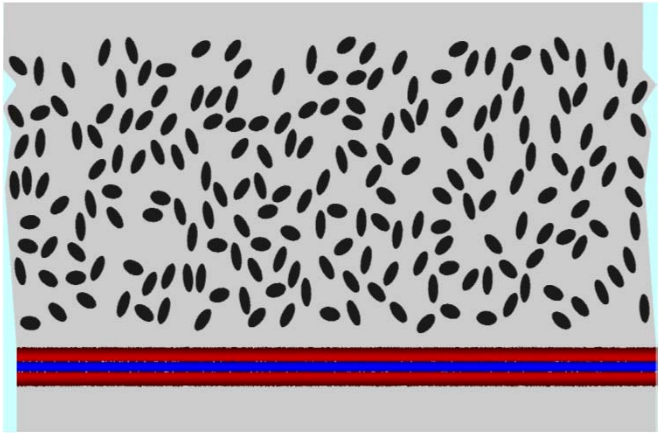
Experimental tanks in Marseille

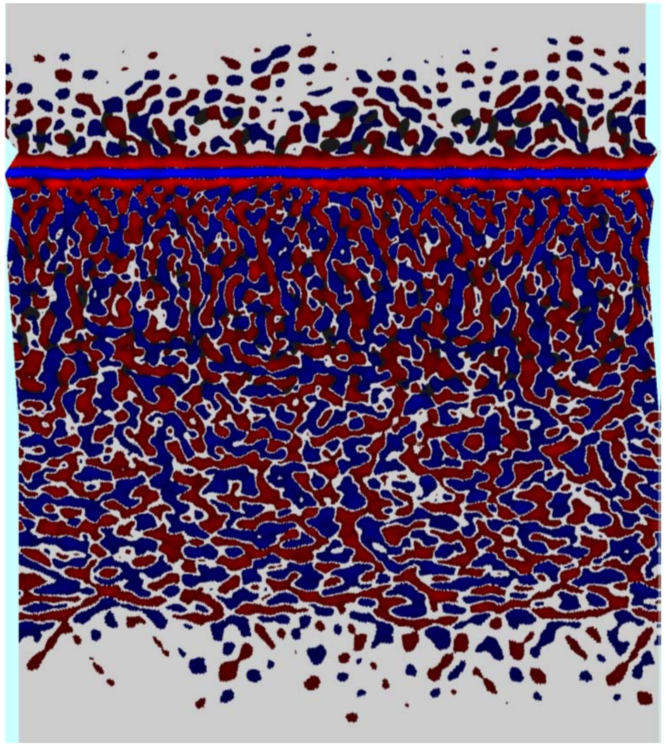


Experiments in known environment / setup

Perform experimental benchmarks

Non destructive testing of materials





Collaboration with LCND lab (CNRS) in Aix, France.

Currently : Physical modeling based on diffusion functions for objects of complex shape, cracks or multiple cavities in concrete, metals, or composite materials. Experiments on samples.

⇒ Very accurate calculations without homogenization can validate (or not) these diffusion functions and extend them beyond their domain of validity.

⇒ Reliable modeling of the “coda” part of the signal, which contains useful information on the medium.

Conclusions and future work

- On modern computers, large 3D full-waveform forward modeling problems can be solved at high resolution in the time domain for acoustic / elastic / viscoelastic / poroelastic / seismic waves

- Inverse (adjoint) tomography and full waveform imaging problems can also be studied, although the cost is still high

- Regional scale is now OK and relatively inexpensive thanks to coupling with DSM on moderate-size clusters ;
  Wavelet compression could be helpful in future work (Simons et al. 2010, Chevrot, Martin and Komatitsch 2012)

- Useful in different industries in addition to academia: oil and gas, ocean acoustics / sonars, non destructive testing (concrete, composite media, fractures, cracks).

Note: the SPECFEM3D source code is freely available open source at http://www.geodynamics.org