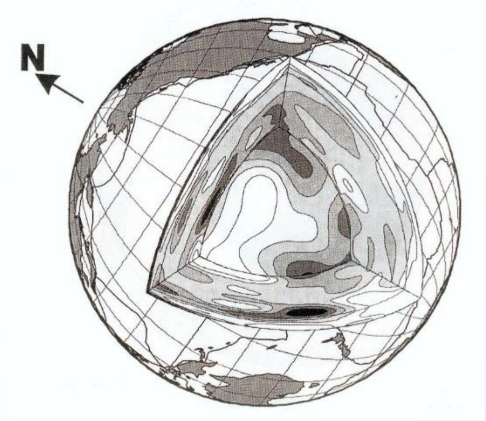
"Multi-parameter seismic inversion".

Joint Inversion –velocity/hypocenter



By J. Virieux, professor UGA

ISTerre
Institut des Sciences de la Terre

American Petroleum Institute, 1986

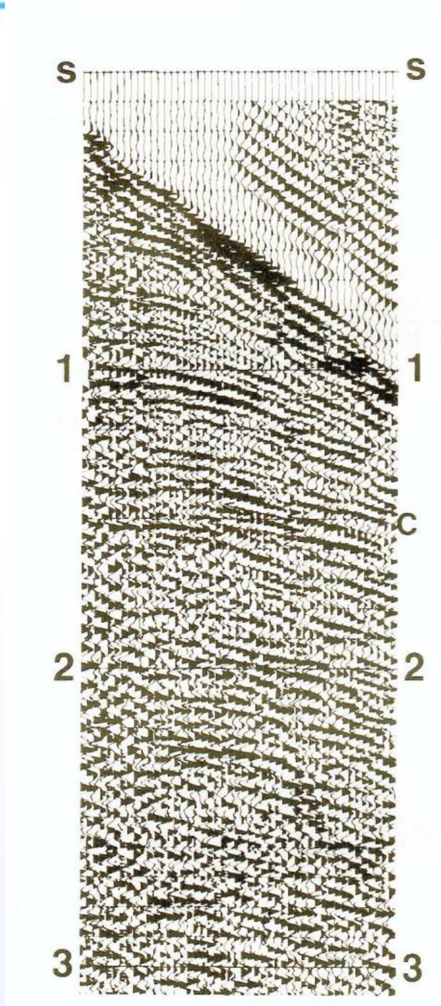
Understanding …

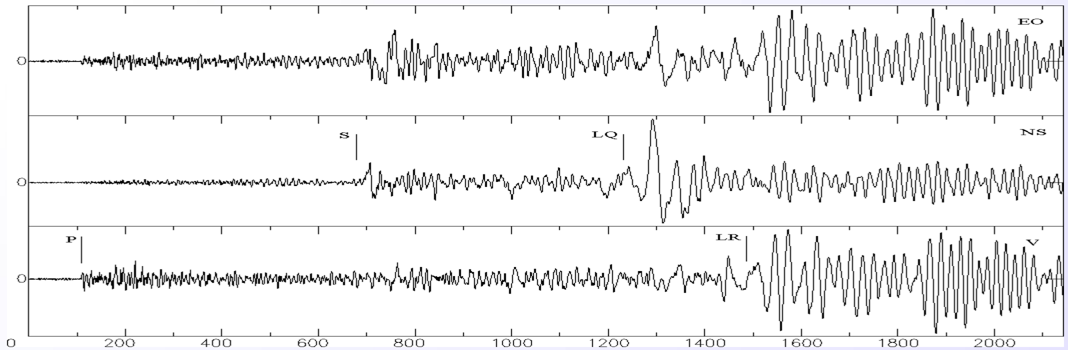
Performing travel time tomography requires

- ray theory for the forward problem
- optimization theory for the inverse problem

If not, you may have a chance to play with many parameters … with limited success.

Time scales





- Source time

  from 0.1 sec to 100 sec (rupture vel.)

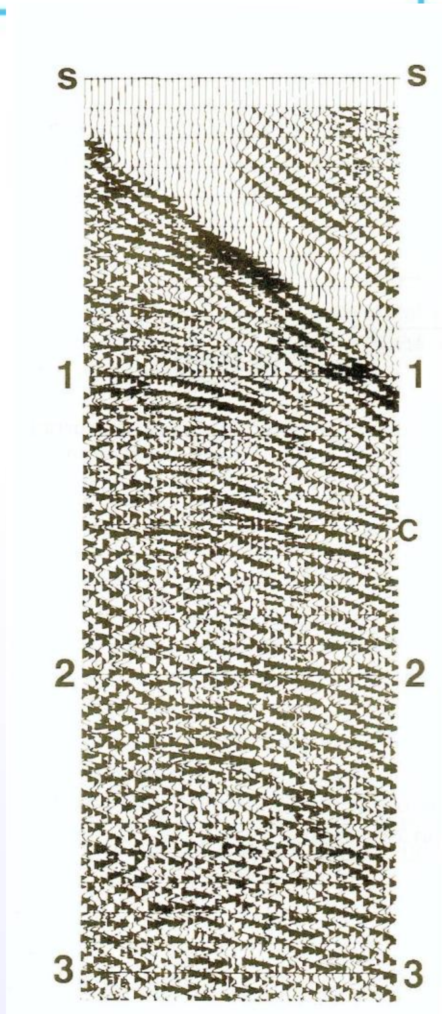
- Wave time

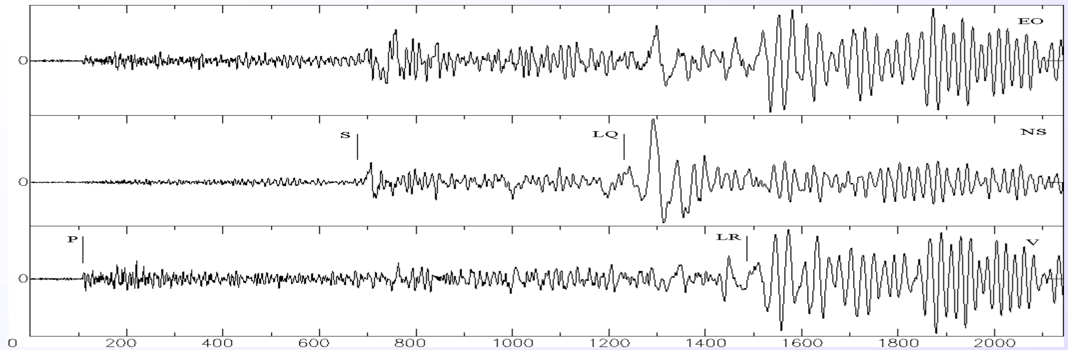
  from secondes to hours

- Window time

  from few secondes to days

Length scales





- Fault length

  200 km for vr=2 km/s

- Discontinuity distance

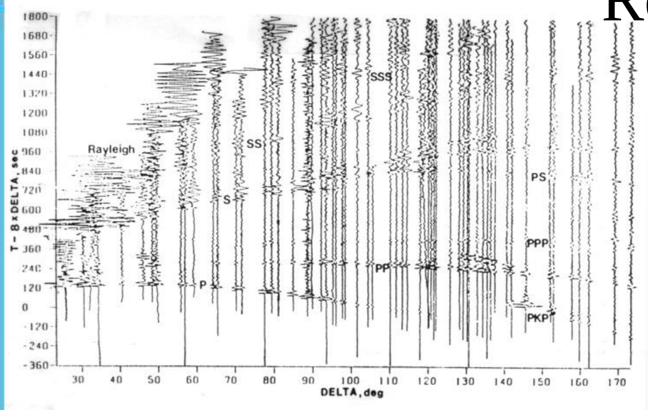
  from few meters to few 100 kms

- Volume sampling

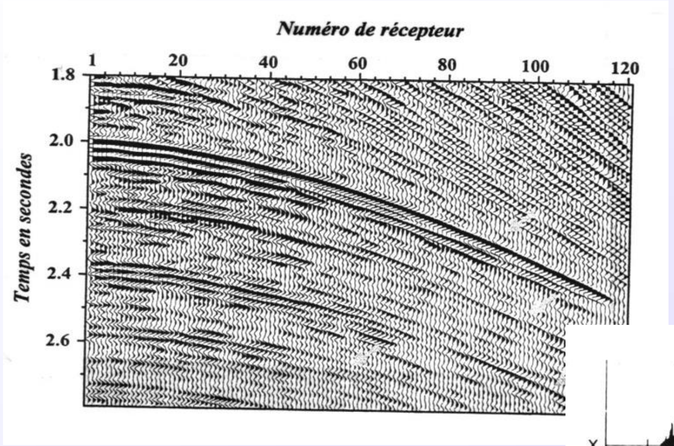
  from few kms to few 1000 kms

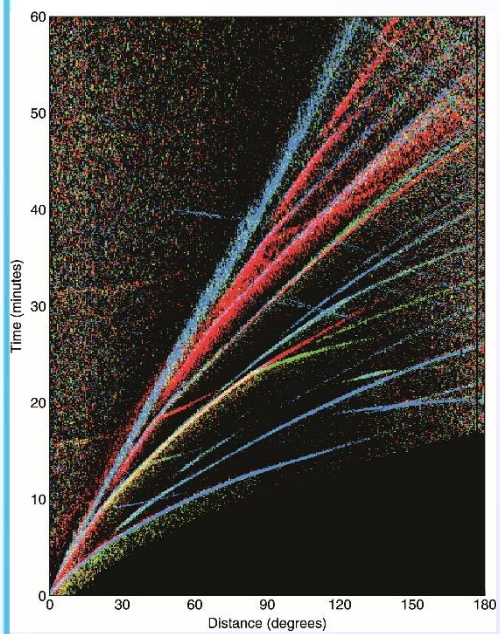
Examples

Record of a far earthquake (Müller and Kind, 1976)

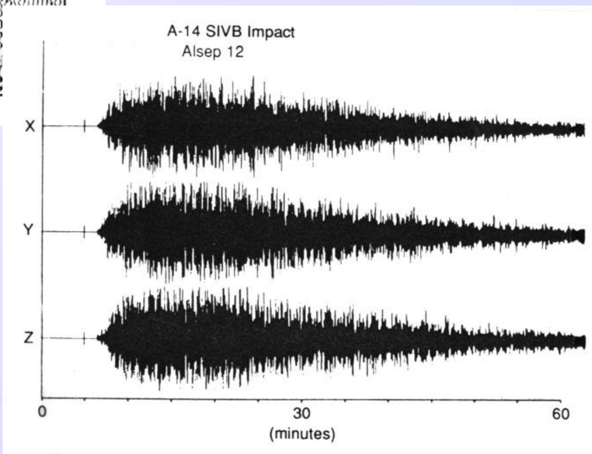


Traces from a oil reservoir (Thierry, 1997)





Seismic imaging is a tough problem on the Moon !



Records on the Moon (meteoritic impact)
(Latham et al., 1971)

Motivation

- Delayed travel-time tomography source:receiver
  - Fitting times => estimation of times

    => 2-pts rays: millions!
- Full Waveform Inversion/Migration src/rec:focal point
  - Same challenge with an order increase in magnitude
- Seismogram modeling
  - Handling multiple arrivals and filling gaps (shadows)

Rays ⇔ Geometrical Optics ⇔Infinite Frequency ⇔Singularities topology

A link with the Catastrophe Theory (René Thom)
Do we need this complexity?

Ray+Born approach

- Ray tracing efficient only in smooth media
- Updating the velocity structure through iterations induces a dramatic increasing complexity in ray tracing: people stays often (always) with the initial ray tracing once done (Background model~Born)
- Asymptotic theory (or high frequency approximation) introduces a complexity in our finite frequency seismic wave propagation : all scales are there while they are not present in the seismic data

In fluid mechanism:

physics below a given small scale is parametrized (upscaled)

References

- Červený, V., Seismic ray theory, Cambridge University Press, 2001
- Chapman, C., Fundamentals of seismic wave propagation, Cambridge University Press, 2004
- Slawinski, M.A., Seismic waves and rays in elastic media, Handbook of geophysical exploration, seismic, exploration, volume 34, Pergamon, 2003
- Abdullaev, S.S., Chaos and dynamics of rays in waveguide media, Gordon and Breach Science Publishers, 1993
- Goldstein, H, Classical mechanics, Addison-Wesley Publishing company, second edition,1980
- Glassner, A.S. (editor), An introduction to ray tracing, Academic Press, second edition, 1991
- Gilmore, R, Catastrophe theory for scientists and engineers, Wiley & Sons, Inc, 1981.
- Sethian, J.A., Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science, Cambridge University Press, 1999.
- Virieux, J. and Lambaré, G., Theory and Observations-Body waves: Ray methods and finite frequency effecs in Treatrise on Geophysics, Tome I by A. Diewonski and B. Romanowicz, Elsevier, 2010

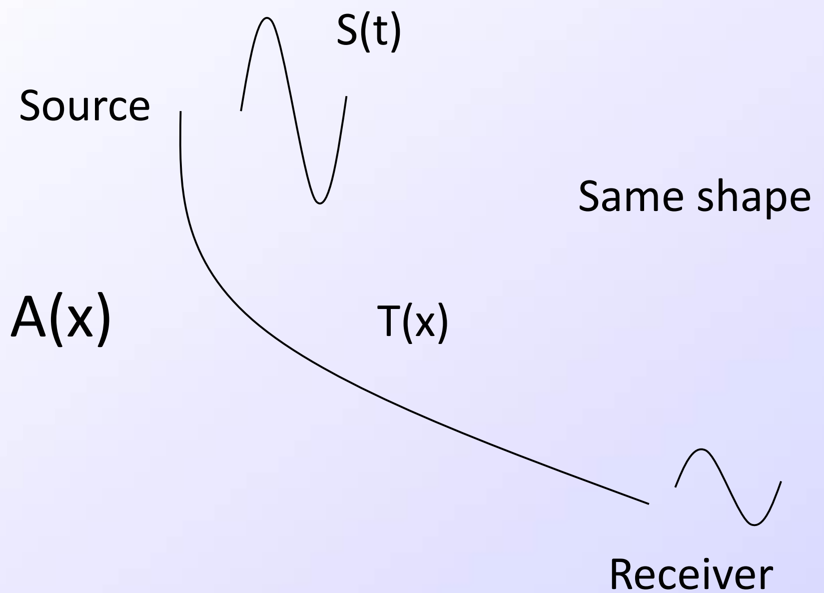
References (selection)

- Červený, V., Seismic ray theory, Cambridge University Press, 2001
- Chapman, C., Fundamentals of seismic wave propagation, Cambridge University Press, 2004

Translucid Earth

Vibration in a point $u(x,t)$

$$u(x,t) = A(x)\, S\big(t - T(x)\big)$$
$$u(x,\omega) = A(x)\, S(\omega)\, e^{i\omega T(x)}$$



Travel-time T(x) and Amplitude A(x)

Diffracting medium: Wavefront preserved

wavefront coherence lost !

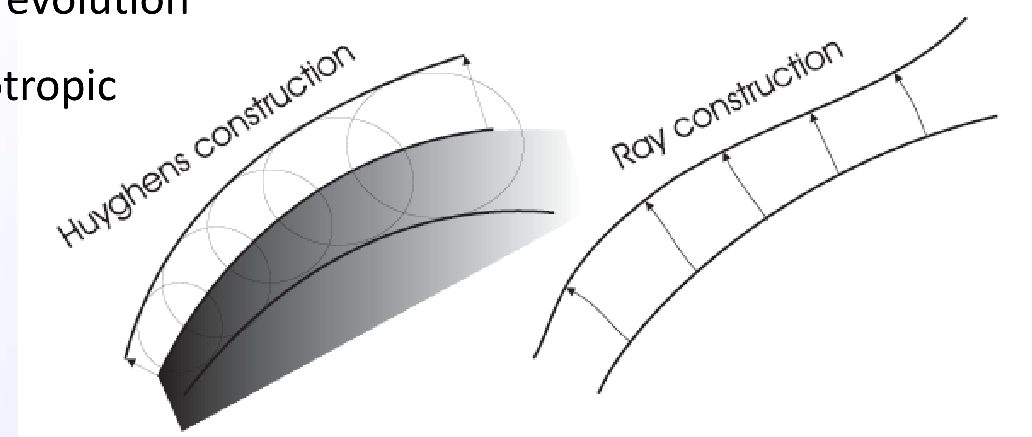
$\omega T(x)$ is sometimes called the phase

Wavefront : $T(x)=T_0$

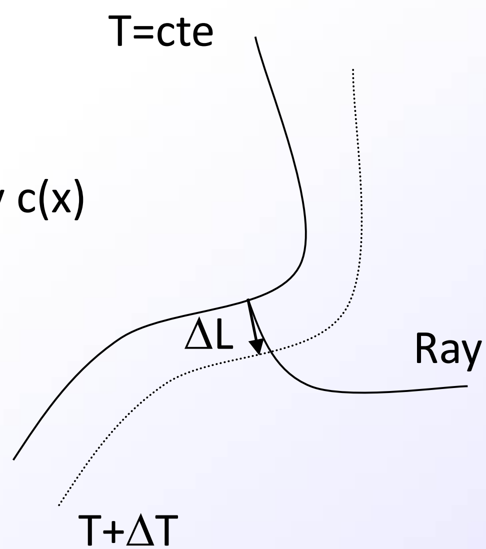
Eikonal equation

Two simple interpretations of wavefront evolution

Orthogonal trajectories are rays in an isotropic medium



Velocity c(x)



Grad(T)=$\nabla_x T$ orthogonal to wavefront

$$c(x) = \frac{\Delta L}{\Delta T} \rightarrow \frac{\Delta T}{\Delta L} = \frac{1}{c(x)} \rightarrow \nabla_x T(x) = \frac{1}{c(x)}$$

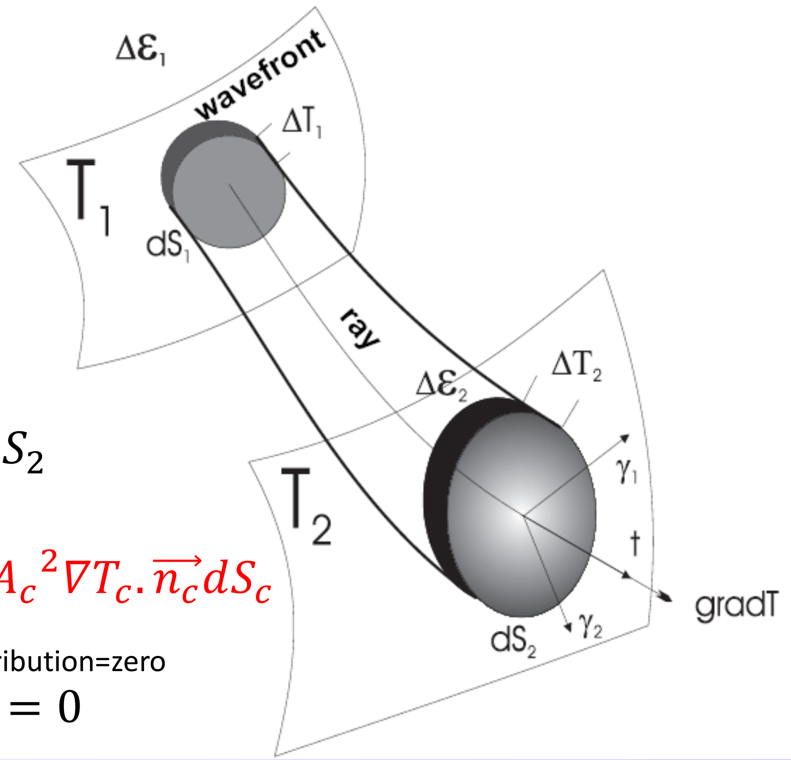
Direction ? : abs or square

The orientation of the wavefront could not be guessed from the local information on a specific wavefront

$$\boxed{(\nabla_x T(x))^2 = \frac{1}{c^2(x)}}$$

Transport Equation

Tracing neighboring rays defines a ray tube : variation of amplitude depends on energy flux conservation through sections.



Energy flux same at section one and at section two

$$\Delta\mathcal{E}_1 = {A_1}^2 dS_1 \Delta T_1 = {A_2}^2 dS_2 \Delta T_2 = \Delta\mathcal{E}_2$$

$$\rightarrow {A_1}^2 \nabla T_1 . \vec{n} dS_1 = {A_2}^2 \nabla T_2 . \vec{n} dS_2$$

$$0 = \iint_{Section}^{Ray} -{A_1}^2 \nabla T_1 . \vec{n} dS_1 + {A_2}^2 \nabla T_2 . \vec{n} dS_2$$

$$0 = \iint_{Section}^{Ray} {A_1}^2 \nabla T_1 . \vec{n'} dS_1 + {A_2}^2 \nabla T_2 . \vec{n} dS_2 + \iint_{Tube}^{Ray} {A_c}^2 \nabla T_c . \vec{n_c} dS_c$$

net contribution=zero

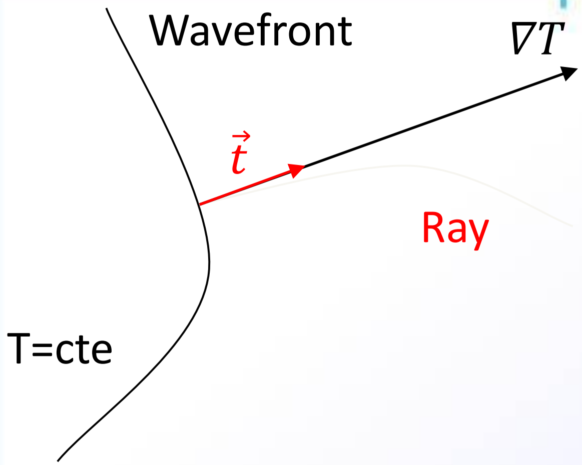
$$0 = \iiint_{Volume}^{Ray} div(A^2\, \nabla T) dV \Longrightarrow div(A^2 \nabla T) = 0$$

$$A(x)(2\nabla A(x) . \nabla T(x) + A(x) \nabla^2 T(x)) = 0$$

$$2\nabla A(x) . \nabla T(x) + A(x) \nabla^2 T(x) = 0$$

Ray equation

$\overrightarrow{x(s)}$ *with s as curvilinear abscisse* $\vec{t} = \frac{d\vec{x}}{ds} \rightarrow \|\vec{t}\| = 1$



- Evolution of $\vec{x}$ is given by $\frac{d\vec{x}}{ds}$

$$\frac{d\vec{x}}{ds} \parallel \nabla T \rightarrow \frac{d\vec{x}}{ds} = c(x)\nabla T(x)$$

- Evolution of ∇T is given by $\frac{d\nabla T}{ds}$

Ray equations

$$\frac{d\nabla T}{ds} = \nabla\left(\frac{1}{c(x)}\right)$$

but the operator $\frac{d}{ds} = \vec{t}.\nabla \; = c\nabla T.\nabla$

and, therefore, $\frac{d\nabla T}{ds} = c\nabla T.\nabla \; (\nabla T)$

leading to $\frac{d\nabla T}{ds} = \frac{c}{2}\nabla(\nabla T.\nabla T) = \frac{c}{2}\nabla(\nabla T)^2 = \frac{c}{2}\nabla\left(\frac{1}{c^2}\right)$

Curvature equation

$$\frac{d}{ds}\left(\frac{1}{c(x)}\frac{d\vec{x}}{ds}\right) = \nabla\left(\frac{1}{c(x)}\right)$$

We define the slownss vector $\vec{p} = \nabla T(x)$ and the position $\vec{q} = \vec{x}(s)$ along the ray

Various non-linear ray equations

Curvilinear stepping

$$\frac{d\vec{q}(s)}{ds} = c(\vec{q})\vec{p}$$
$$\frac{d\vec{p}(s)}{ds} = \nabla_{\vec{q}} \frac{1}{c(\vec{q})}$$
$$\frac{dT(s)}{ds} = \frac{1}{c(\vec{q})}$$

$$dT = \frac{1}{c(\vec{q})} ds = \frac{1}{c(\vec{q})^2} d\xi$$

Time stepping

$$\frac{d\vec{q}(t)}{dt} = c^2(\vec{q})\vec{p}$$
$$\frac{d\vec{p}(t)}{dt} = c(\vec{q})\nabla \frac{1}{c(\vec{q})}$$

The simplest set

Particule stepping

$$\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})} \nabla \frac{1}{c(\vec{q})}$$
$$\frac{dT(\xi)}{d\xi} = \frac{1}{c^2(\vec{q})}$$

under the condition of the eikonal
$p^2 = 1/c^2(\vec{q})$

Which ODE to select for numerical solving ? Either t or ξ sampling.
Many analytical solutions (gradient of velocity; gradient of slowness square)

How to solve these equations

ODE versus PDE !

ODE

Lagrangian formulation

Wavefront complexity

Methods of characteristics

Differential geometry (Courant & Hilbert, 1966)

- ❖ Non-linear ordinary differential equations
- ❖ Lagrangian formulation as we integrate along rays

In opposition to Eulerian formulation where we compute (ray) quantities at fixed positions

Properties of these ODEs

- Intrinsic solutions independent of the coordinate sytem used to solve it
- If dummy variable for velocity, use it as the variable stepping (often x coordinate)

$$\nabla_x \frac{1}{c(q_z)} = 0 \Rightarrow p_x = cte \Rightarrow q_x = {q_x}^0 + \xi p_x \text{ (1)}$$

- Eikonal equation: a good proxy for testing the accuracy of the ray tracing (not enough used)

In 3D: six or seven equations
In 2D: four or five equations

(1): rectilinear motion of a particle along this axis in mechanics

Hamilton's (ray) equations

$$\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})}\nabla_{\vec{q}}\frac{1}{c(\vec{q})}$$

$$\mathcal{H}(\vec{q},\vec{p}) = \frac{1}{2}(p^2 - \frac{1}{c^2(\vec{q})})$$

$$\frac{d\vec{q}(\xi)}{d\xi} = \nabla_{\vec{p}}\mathcal{H}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = -\nabla_{\vec{q}}\mathcal{H}$$
$$\frac{dT}{d\xi} = \vec{p}.\nabla_{\vec{p}}\mathcal{H}$$

sympletic structure (FUN!)

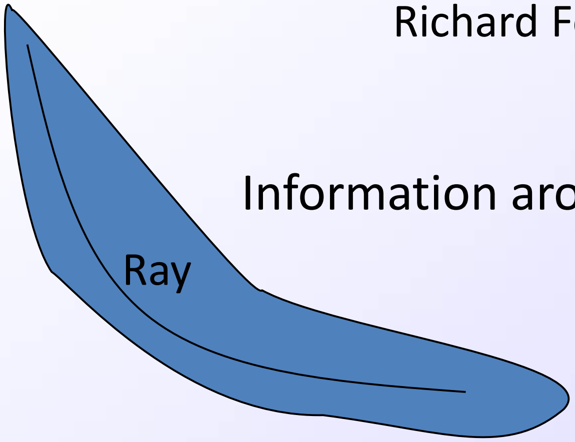
Mechanics : ray tracing as a particular balistic problem

Hamilton approach suitable for perturbation
(Henri Poincaré en 1907 « Mécanique céleste »,
Richard Feymann Prix Nobel 1965)

Information around the ray

$$\vec{q}_0 + \delta\vec{q}$$
$$\vec{p}_0 + \delta\vec{p}$$

$\delta q \; and \; \delta p$ "small"



Meaning of the neighborhood zone
Fresnel zone if finite frequency
Any zone depending on your problem GBS

Velocity variation v(z)

Ray equations are

$$\frac{dq_x}{d\tau} = p_x; \frac{dq_y}{d\tau} = p_y; \frac{dq_z}{d\tau} = p_z$$

$$\frac{dp_x}{d\tau} = 0; \frac{dp_y}{d\tau} = 0; \frac{dp_z}{d\tau} = u(z)\frac{du(z)}{dz}$$

The horizontal component of the slowness vector is constant: the trajectory is inside a plan which is called the plan of propagation. We may define the frame (xoz) as this plane.

$$\frac{dq_x}{d\tau} = p_x; \frac{dq_z}{d\tau} = p_z$$

$$\frac{dp_x}{d\tau} = 0; \frac{dp_z}{d\tau} = u(z)\frac{du(z)}{dz}$$

$$\frac{dq_x}{dq_z} = \frac{p_x}{p_z} = \frac{p_x}{\pm\sqrt{u^2(z) - p_x^2}}$$

Where p_x is a constante

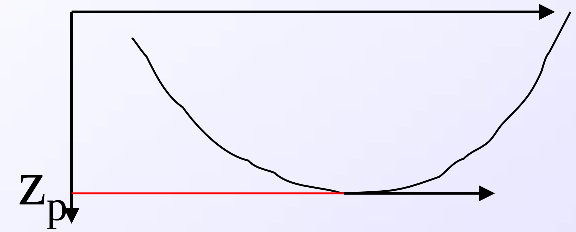
$$q_x(z_1, p_{x1}) = q_x(z_0, p_{x0}) + \int_{z_0}^{z_1} \frac{p_x}{\sqrt{u^2(z) - p_x^2}} dz$$

For a ray towards the depth

Velocity variation v(z)

$$q_x(z_1,p_{x1}) = q_{x0} + \int_{z_0}^{z_p} \frac{p_x}{\sqrt{u^2(z)-p_x^2}}dz + \int_{z_1}^{z_p} \frac{p_x}{\sqrt{u^2(z)-p_x^2}}dz$$

$$T(z_1,p_{x1}) = T_0 + \int_{z_0}^{z_p} \frac{u^2(z)}{\sqrt{u^2(z)-p_x^2}}dz + \int_{z_1}^{z_p} \frac{u^2(z)}{\sqrt{u^2(z)-p_x^2}}dz$$



At a given maximum depth z_p, the slowness vector is horizontal following the equation $p_x^2 = p^2 = u^2(z_p)$

If we consider a source at the free surface as well as the receiver, we get

In Cartesian frame

$$X(p) = 2\int_0^{z_p} \frac{p}{\sqrt{u^2(z)-p^2}}dz$$

$$T(p) = 2\int_0^{z_p} \frac{u^2(z)}{\sqrt{u^2(z)-p^2}}dz$$

with p = usini

In Spherical frame

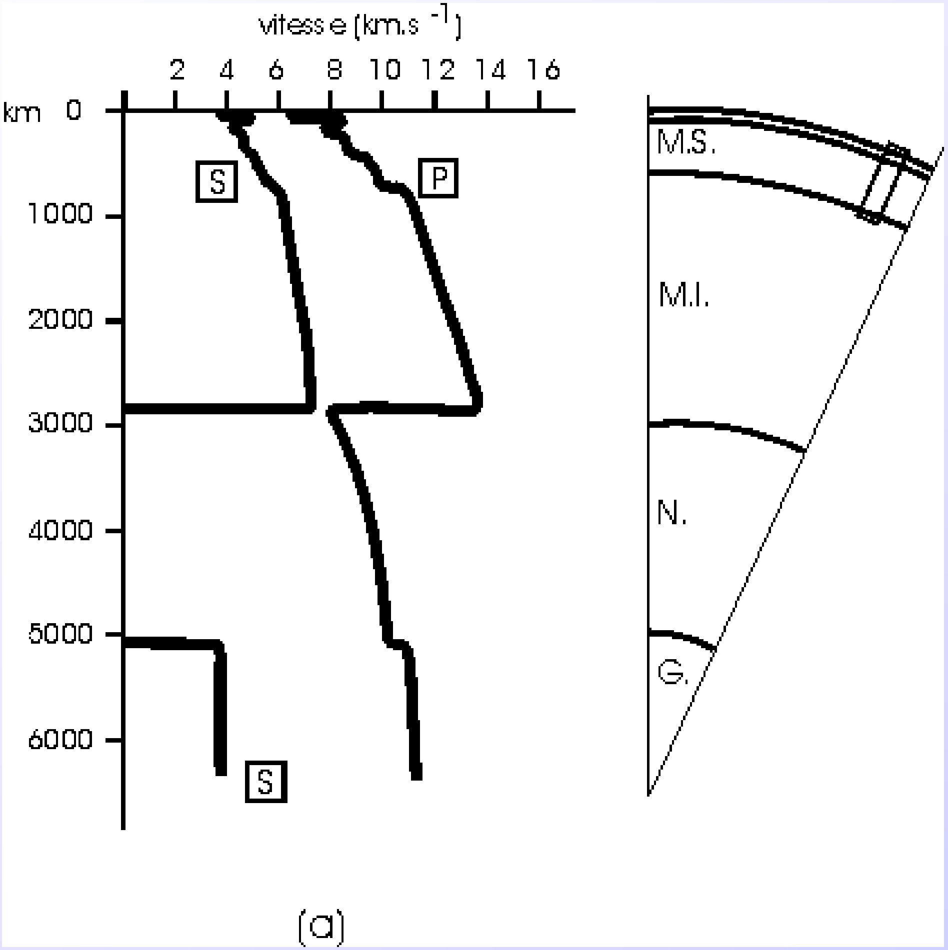
$$\Delta = 2\int_{r_p}^{a} \frac{p}{\sqrt{r^2u(r)^2-p^2}}\frac{dr}{r}$$

$$T = 2\int_{r_p}^{a} \frac{r^2u(r)^2}{\sqrt{r^2u(r)^2-p^2}}\frac{dr}{r}$$

$$with \quad p = ru\sin i$$

Velocity structure in the Earth

- Radial Structure (PREM, JB …)



System of ray tracing equations

Curvilinear stepping

$$\frac{d\vec{q}(s)}{ds} = c(\vec{q})\vec{p}$$
$$\frac{d\vec{p}(s)}{ds} = \nabla_{\vec{q}}\frac{1}{c(\vec{q})}$$
$$\frac{dT(s)}{ds} = \frac{1}{c(\vec{q})}$$

Time stepping

$$\frac{d\vec{q}(t)}{dt} = c^2(\vec{q})\vec{p}$$
$$\frac{d\vec{p}(t)}{dt} = c(\vec{q})\nabla\frac{1}{c(\vec{q})}$$

The simplest set

Particule stepping

$$\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})}\nabla\frac{1}{c(\vec{q})}$$
$$\frac{dT(\xi)}{d\xi} = \frac{1}{c^2(\vec{q})}$$

$$dT = \frac{1}{c(\vec{q})}ds = \frac{1}{c(\vec{q})^2}d\xi$$

under the condition of the eikonal
$p^2 = 1/c^2(\vec{q})$

Which ODE to select for numerical solving ? Either t or ξ sampling.

Many analytical solutions (gradient of velocity; gradient of slowness square)

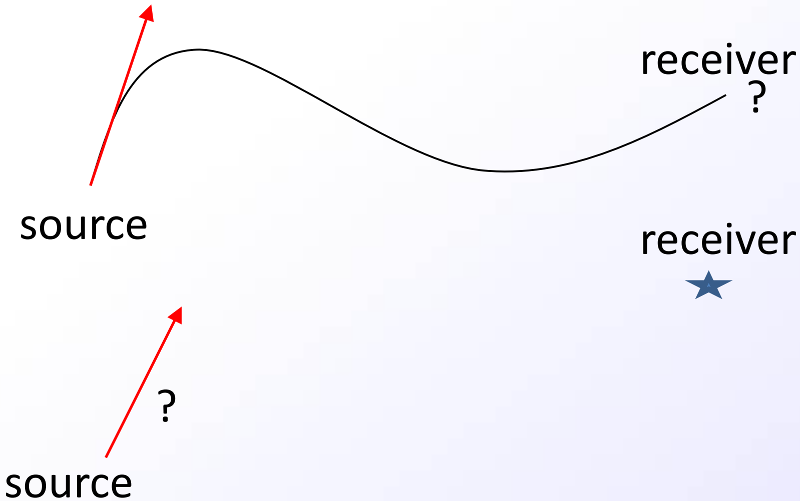
Time integration of ray equations

Runge-Kutta second-order integration

Predictor-Corrector integration

stiffness

1D sampling of 2D/3D medium : FAST



Initial conditions EASY

Boundary conditions VERY DIFFICULT

Shooting δp ?

Bending δx ?

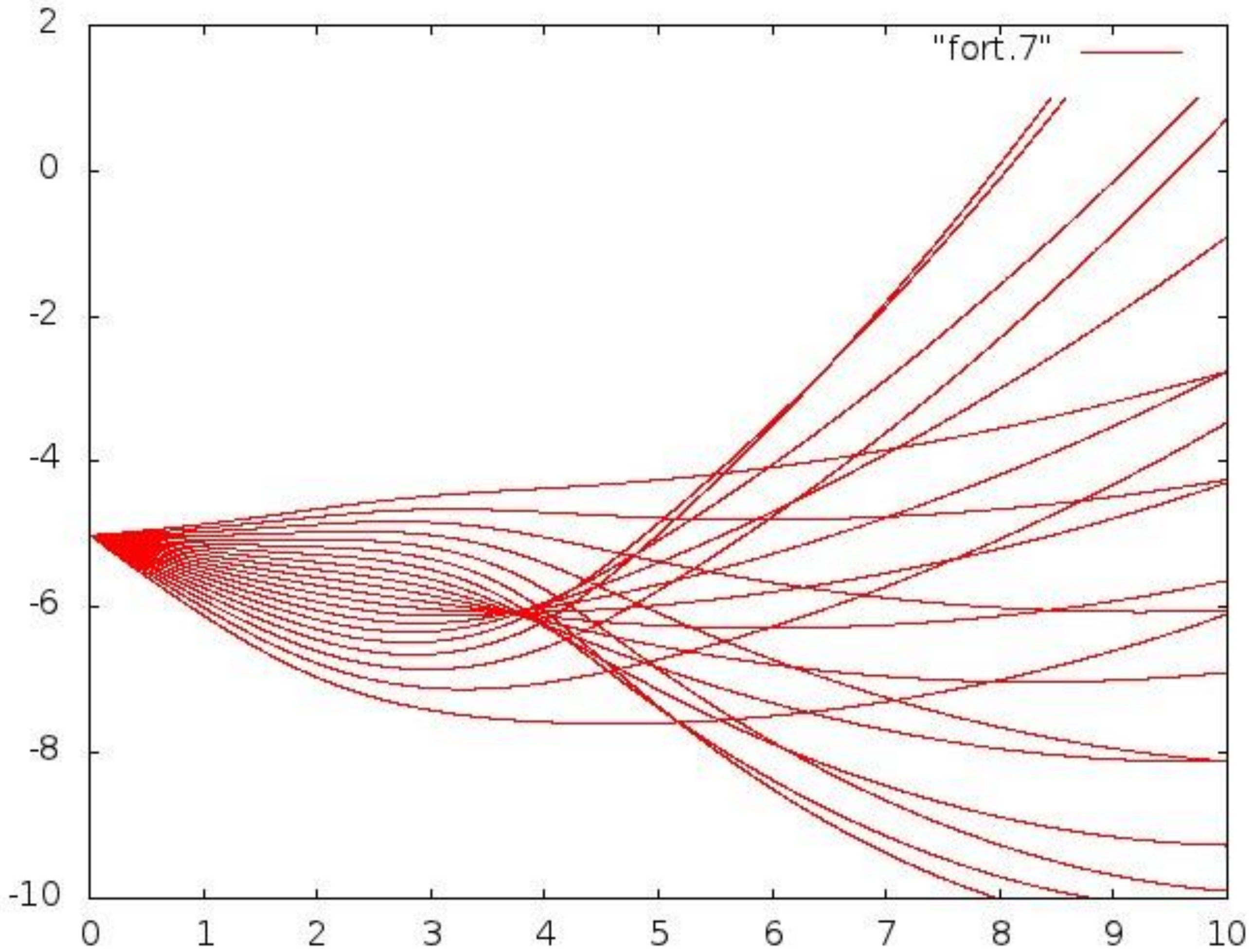
Continuing δc ?

Save p conditions if possible !

AND FROM TIME TO TIME IT FAILS !

But we need 2 points ray tracing because we have a source and a receiver to be connected in seismology!

Ray tracing by rays

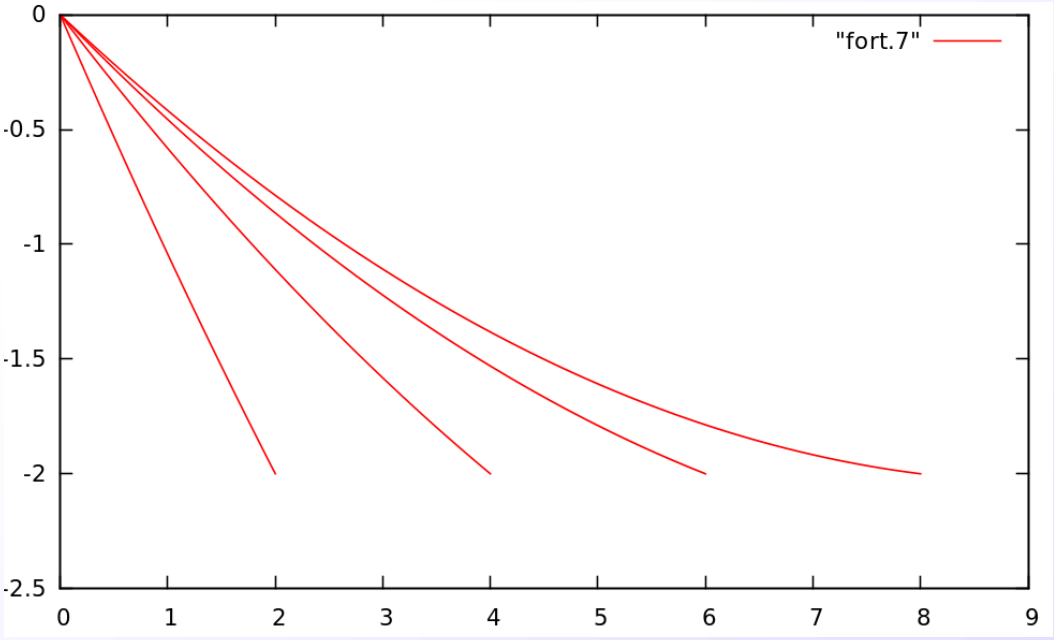


Ray tracing

Constant angle step

Ray tracing by rays

Two-points ray tracing



Source (0,0)

Receivers (2,-2),
(4,-2),(6,-2)

(depth is considered as negative)

Please note the irregular angle stepping.

Hamilton’s (ray) equations

$$\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})} \nabla_{\vec{q}} \frac{1}{c(\vec{q})}$$

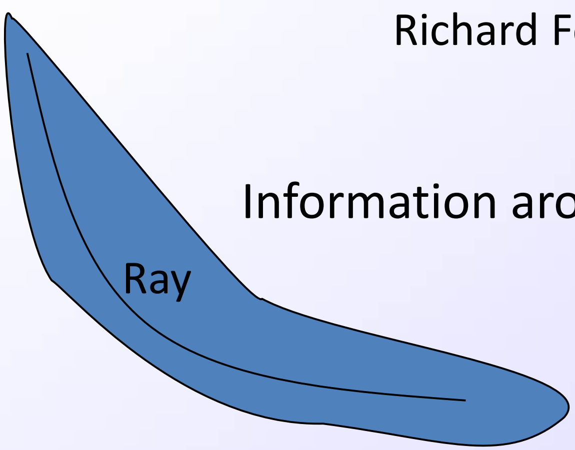
$$\mathcal{H}(\vec{q},\vec{p}) = \frac{1}{2}\left(p^2 - \frac{1}{c^2(\vec{q})}\right)$$

$$\frac{d\vec{q}(\xi)}{d\xi} = \nabla_{\vec{p}}\mathcal{H}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = -\nabla_{\vec{q}}\mathcal{H}$$
$$\frac{dT}{d\xi} = \vec{p}.\nabla_{\vec{p}}\mathcal{H}$$

Mechanics : ray tracing as a particular balistic problem

sympletic structure (FUN!)

Hamilton approach suitable for perturbation
(Henri Poincaré en 1907 « Mécanique céleste »,
Richard Feymann Prix Nobel 1965)



Information around the ray

$$\vec{q}_0 + \delta\vec{q}$$
$$\vec{p}_0 + \delta\vec{p}$$

δq *and* δp *"small"*

Meaning of the neighborhood zone
Fresnel zone if finite frequency
Any zone depending on your problem GBS

ODE resolution

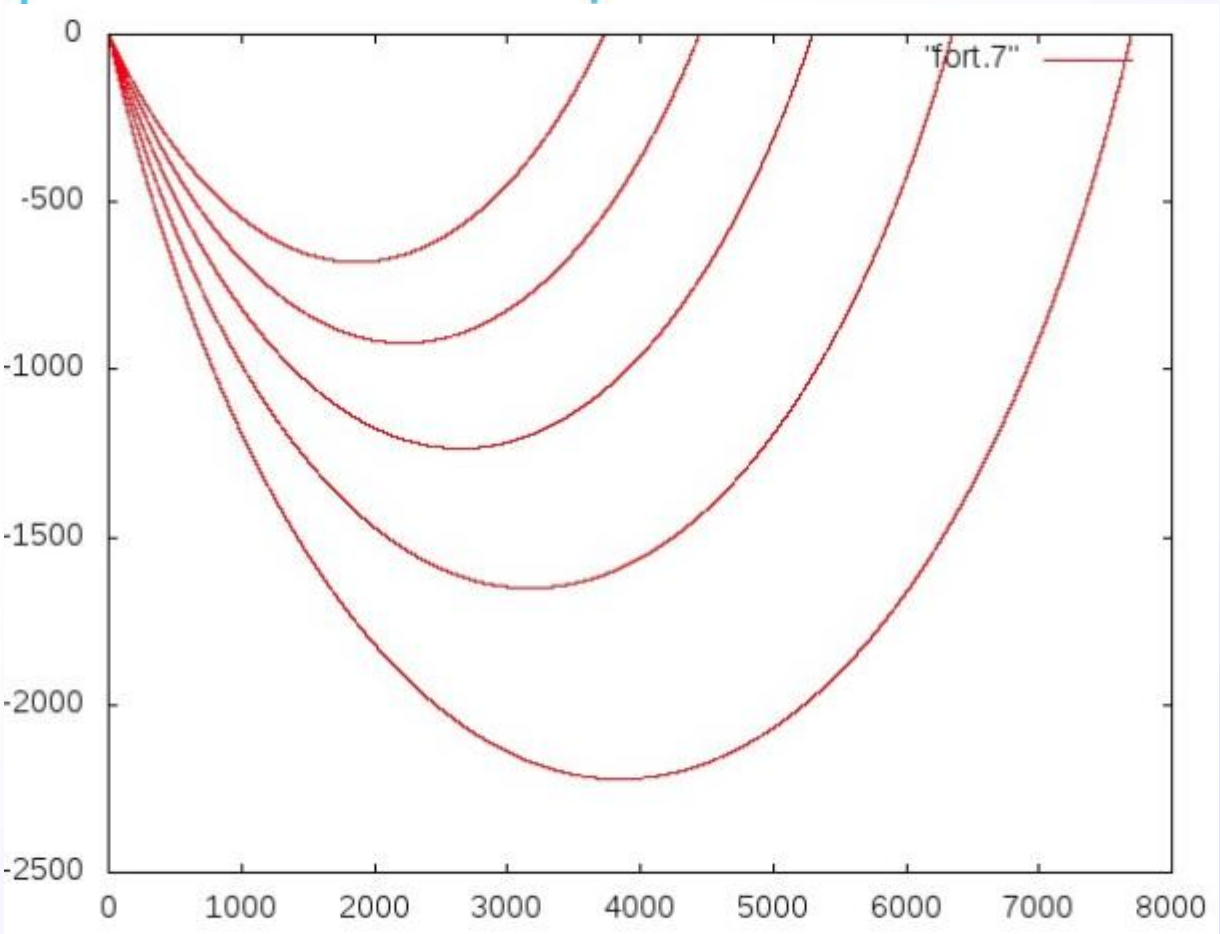
- Runge-Kutta of second order
- Write a computer program for an analytical law for the velocity: take a gradient with a component along x and a component along z

Home work : redo the same thing with a Runge-Kutta of fourth order

(look after its definition)

Consider a gradient of the square of slowness

Analytical solutions



Vertical gradient of the velocity

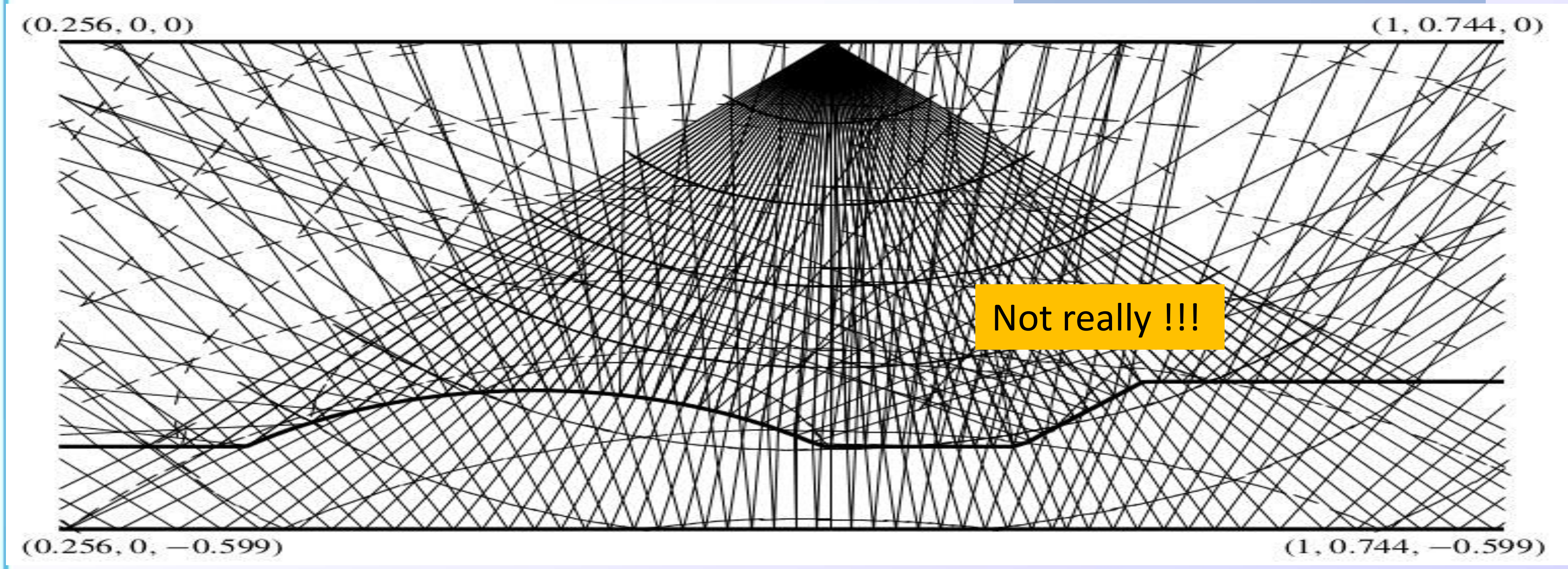
Gradient of the slowness square

Keeping complexity low

- Ray tracing is a fast 1D integration in 2D/3D
- Ray tracing equations as ODEs may sample the medium quite evenly
- They are lagrangian formulation: we follow a point while tracing rays without regarding the density of rays

Keeping complexity low

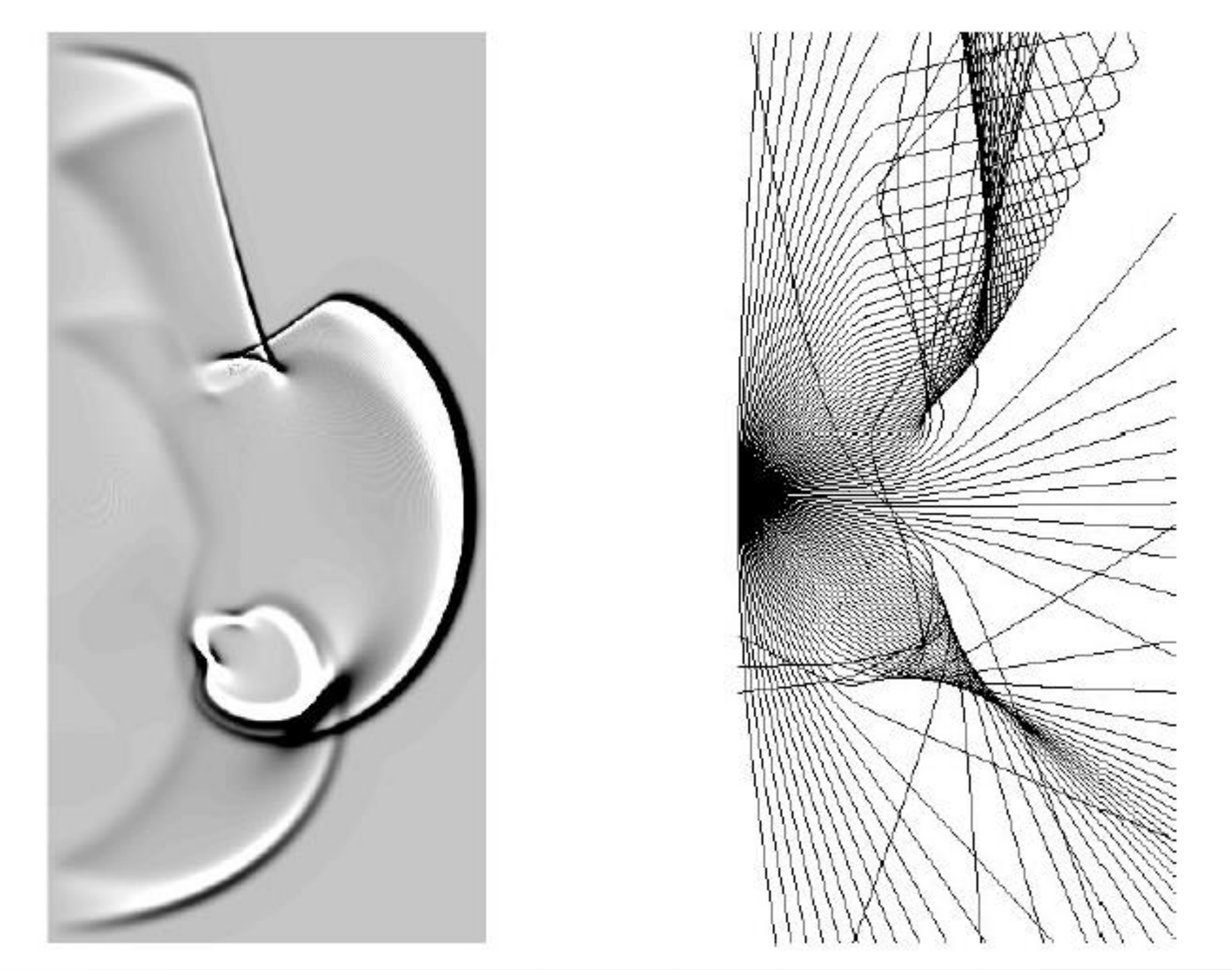
3D French's model



Solutions: moving from ODEs to PDEs

(Osher et al, 2002) for adequate spatial sampling of the wavefront. Grids control the complexity!

Interpolation challenge



How to find travel times where poor ray sampling!

Mitigating the interpolation issue

We cannot guarantee that two rays will stay nearby for small shooting angles changes

Differential approach will be an answer: the perturbed ray will stay nearby the reference ray …

Key step for seismic tomography (more than ray tracing !)

Hamilton's perturbation

$$\mathcal{H}(\vec{q},\vec{p}) = \frac{1}{2}\left(p^2 - \frac{1}{c^2(\vec{q})}\right)$$

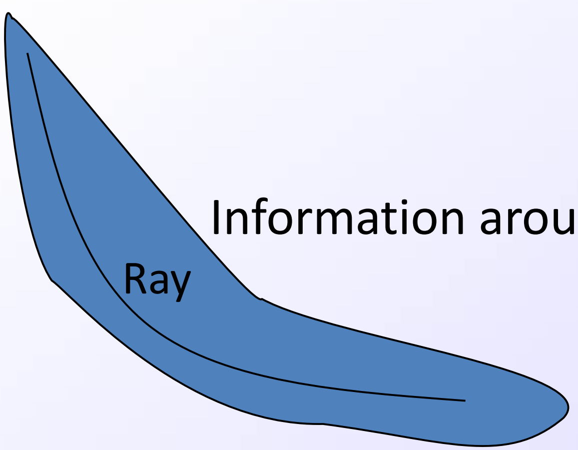
$$\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})}\nabla_{\vec{q}}\frac{1}{c(\vec{q})}$$

$$\frac{d\vec{q}(\xi)}{d\xi} = \nabla_{\vec{p}}\mathcal{H}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = -\nabla_{\vec{q}}\mathcal{H}$$
$$\frac{dT}{d\xi} = \vec{p}.\nabla_{\vec{p}}\mathcal{H}$$

$$y = \begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix}$$

$$\begin{matrix} \vec{q}_0 + \delta\vec{q} \\ \vec{p}_0 + \delta\vec{p} \end{matrix}$$

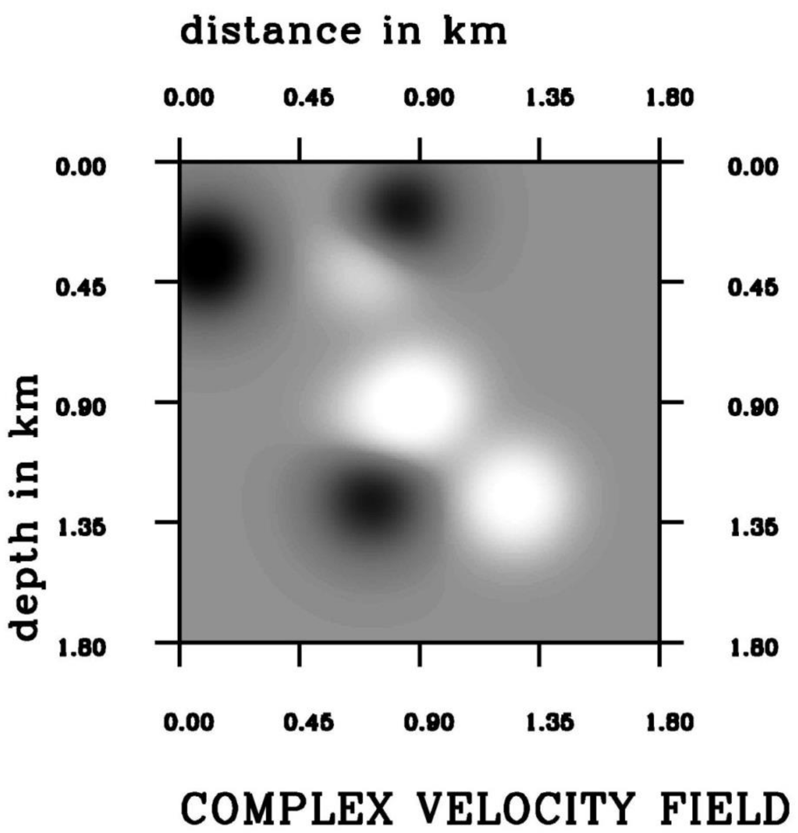
$\delta\vec{q}$ *and* $\delta\vec{p}$ *"small"*

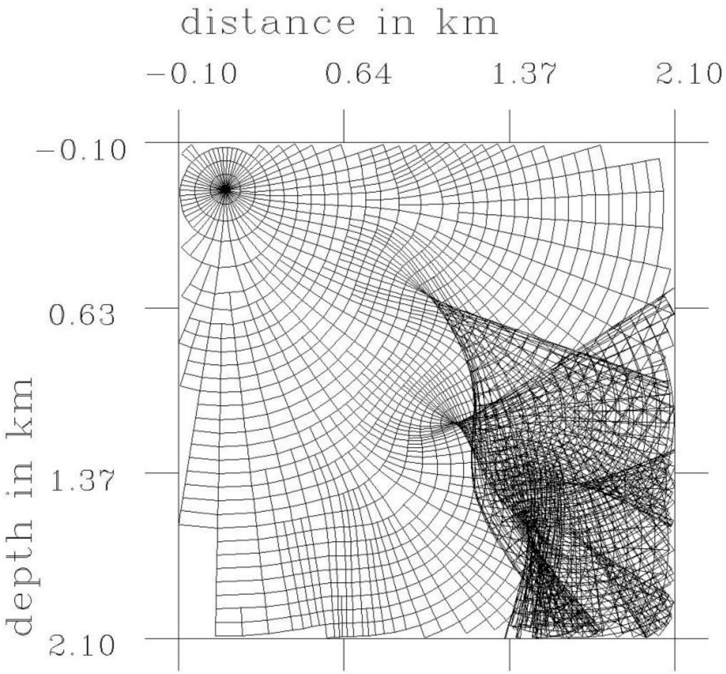


Information around the ray y_0

$$\delta y = \begin{pmatrix} \overrightarrow{\delta q} \\ \overrightarrow{\delta p} \end{pmatrix}$$

STEP ONE: ray tracing





(Lambaré et al., 1996)

Interpolation issue

Local ODE

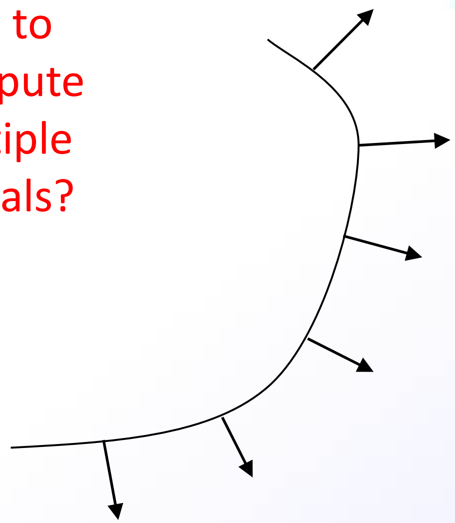
Semi-Lagrangian approach

Sampling control

Ray tracing by wavefronts

(Ray tracing by rays)

How to compute multiple arrivals?



Lambaré et al (1996)

Vinje et al (1993) widely used in NORSAR software

Evolution over time:

folding of the wavefront is allowed (still a significant curse of complexity!)

Dynamic sampling:

undersampling of ray fans

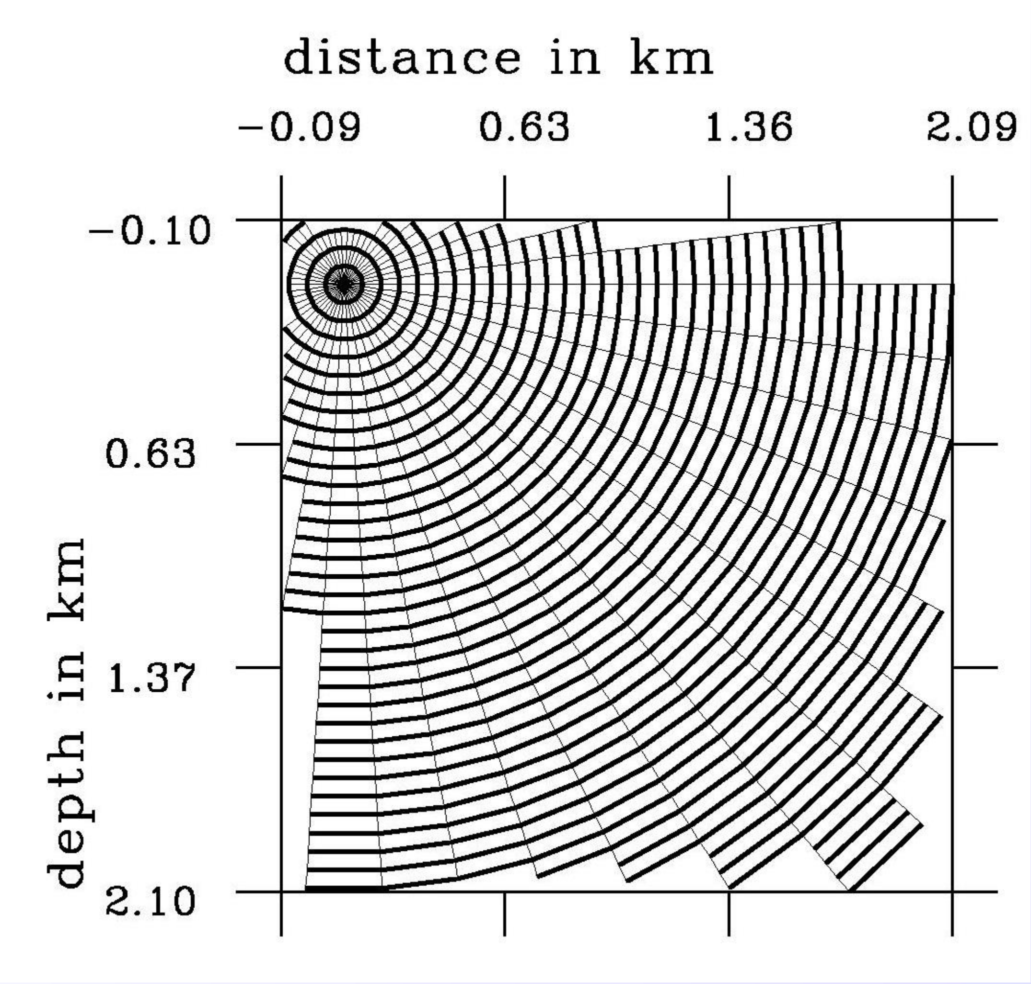
oversampling of ray fans

Keep an « uniform » sampling of the medium by rays

by tracking the surrounding density of rays

by estimating through paraxial approach the ray density

Ray tracing by wavefronts



An ODE is solved at each point of the wavefront while it is spanned

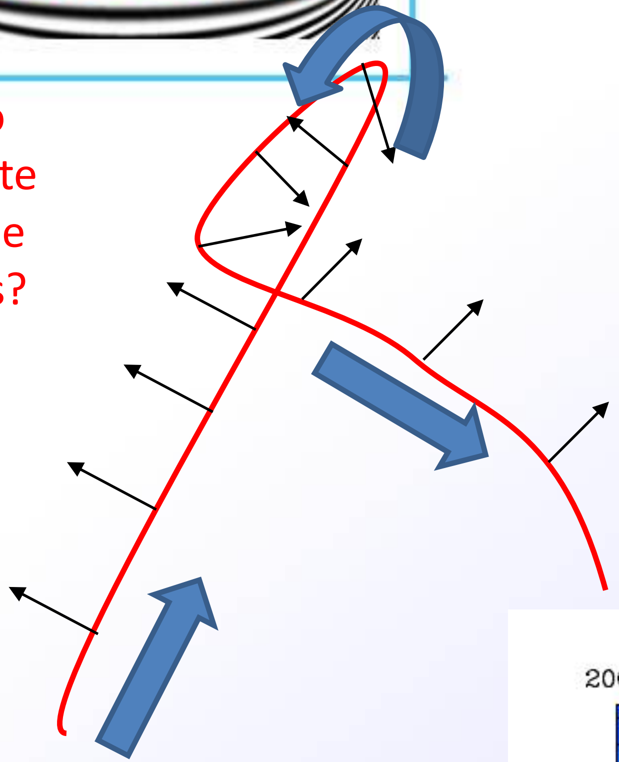
Keeping the sampling of the medium more or less uniform

Rays and wavefronts in an homogeneous medium.

(Lambaré et al., 1996)

Ray tracing by wavefro

How to compute multiple arrivals?

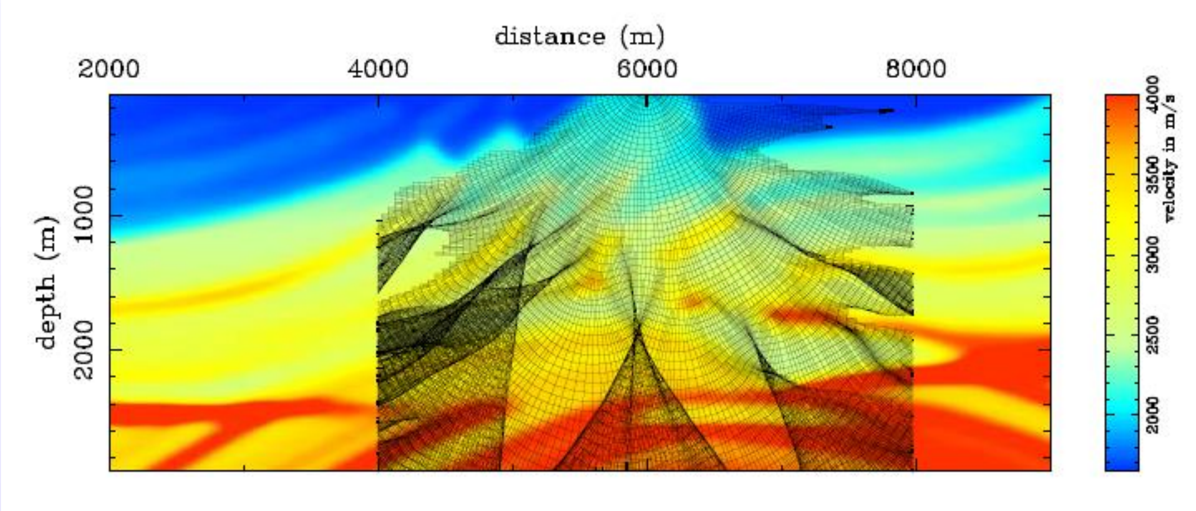


Sampling the wavefront is an heavy task in 2D & 3D !

Still better than oversampling through ray tracing by rays

Example of wavefront evolution in a smooth version of the Marmousi model

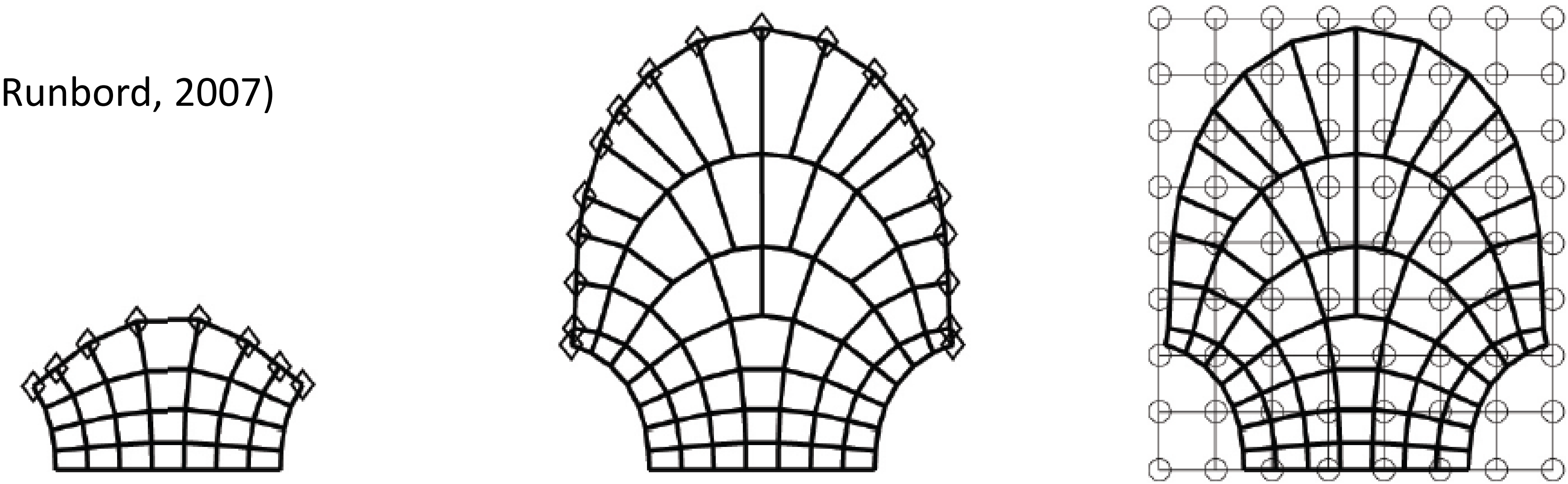
Smoothness is required !



Wavefront construction [Vinje, Iversen, Gjöystdal, Lambaré, . . .]

- Solve for $\boldsymbol{x}(t, \alpha)$ and $\boldsymbol{p}(t, \alpha)$. Discretize in α and trace rays for $\alpha_1, \alpha_2, \alpha_3, \ldots$ where $\alpha_j = j\Delta\alpha$.
- Insert new rays adaptively by interpolation when front resolution deteriorates. E.g.:
  If $|\boldsymbol{x}(t_n, \alpha_{j+1}) - \boldsymbol{x}(t_n, \alpha_j)| \geq tol$ then insert new ray at $\alpha_{j+1/2}$.
- Interpolate traveltime/phase/amplitude onto regular grid.

(Runbord, 2007)



$\boldsymbol{x}(t_n, \alpha_j)$

Geometrical optics resolution

Do we need this complexity of tracking seismic wavefronts!

No scale as we are at infinite frequency !

Is it a fair assumption while we have finite frequency wave content?

We must proceed down to a given resolution length under which we do not want to decipher the wavefront: wavefront healing related to so-called viscous solution.

Few attempts

The medium should be smooth enough for avoiding the wavefront tracking …

Usually done once in the initial model when performing imaging procedure …

A better strategy? Yes, for first arrivals; maybe for multiple arrivals.

PDE

Eulerian approach

Grid control

Level-set functions

Eikonal equation

$$p^2 = \frac{1}{c^2(\vec{x})}$$

$$\mathcal{H}(\vec{x},\vec{p}) = \frac{1}{2}\left(p^2 - \frac{1}{c^2(\vec{x})}\right) = 0$$

The eikonal equation is a first-order non-linear partial differential equation of static Hamilton-Jacobi equations: Crandall & Lions (1983, 1984) have shown that « viscous » solutions of such equation can be obtained.

« Computing first-arrival times is equivalent to tracking an interface advancing at a local speed normal to itself » from Sethian & Popovici (1999)

- Level-set methods (Osher & Sethian, 1988)
- Fast marching methods (Sethian, 1996a; Sethian and Popovici, 1999) $\mathcal{O}(NLogN)$
- Fast sweeping methods (Zhao, 2005) $\mathcal{O}(N)$
- Discontinuous Finite-Element methods (Li et al, 2008; Yan & Osher, 2011)

$$|\vec{p}| = \frac{1}{c(\vec{x})} = |\nabla_{\vec{x}} T(\vec{x})| \qquad \vec{x} \in \Omega$$

$$\mathrm{T}(\vec{x}) = T_0(\vec{x}) \qquad \vec{x} \in \Gamma \subset \partial\Omega$$

Using an upwind, monotone and consistent discretization of $|\nabla_{\vec{x}} T(\vec{x})|$

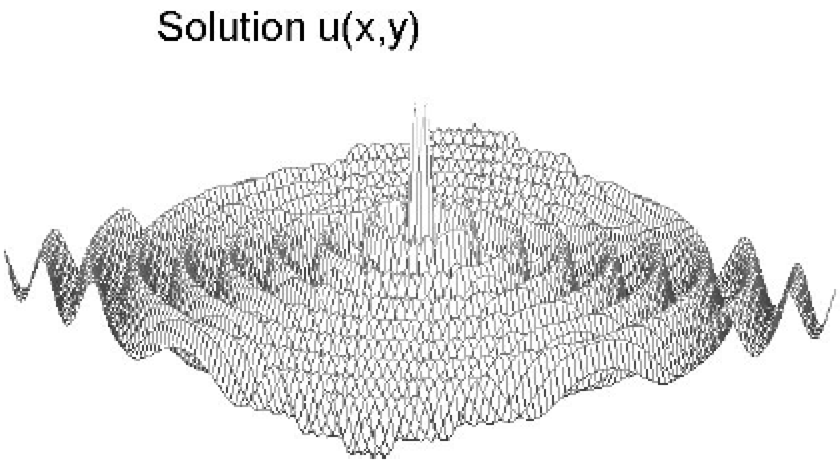
Geometrical optics

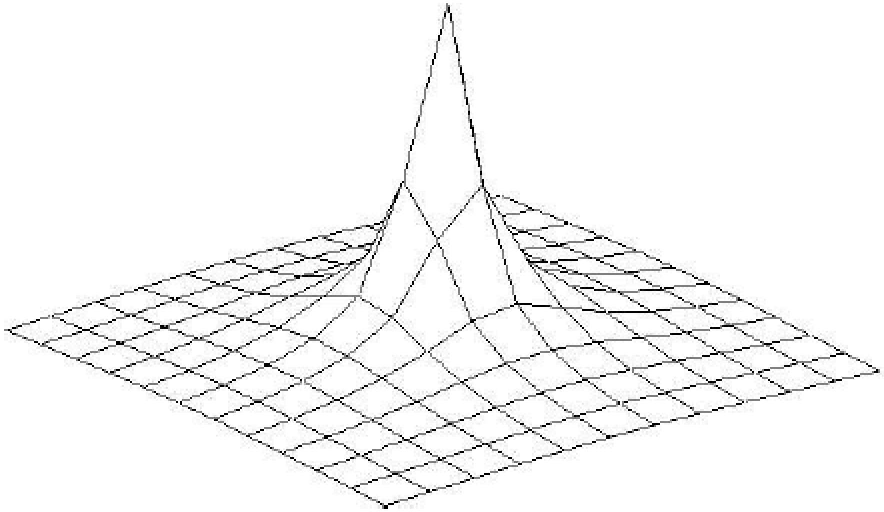
Helmholtz equation

$$\Delta u + \omega^2 n(\boldsymbol{x})^2 u = 0.$$

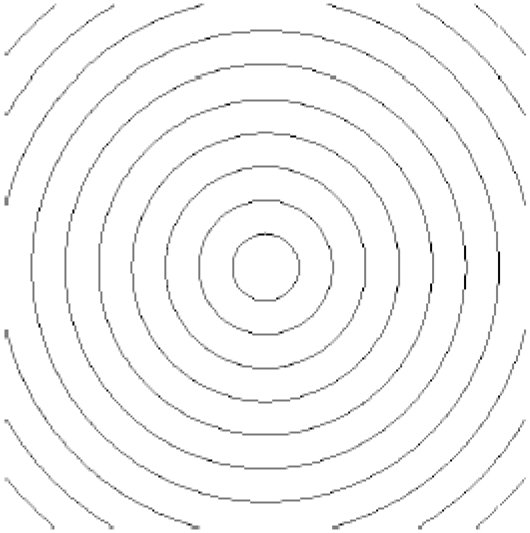
Write solution on the form

$$u(x) = A(x, \omega) e^{i\omega\phi(x)}.$$



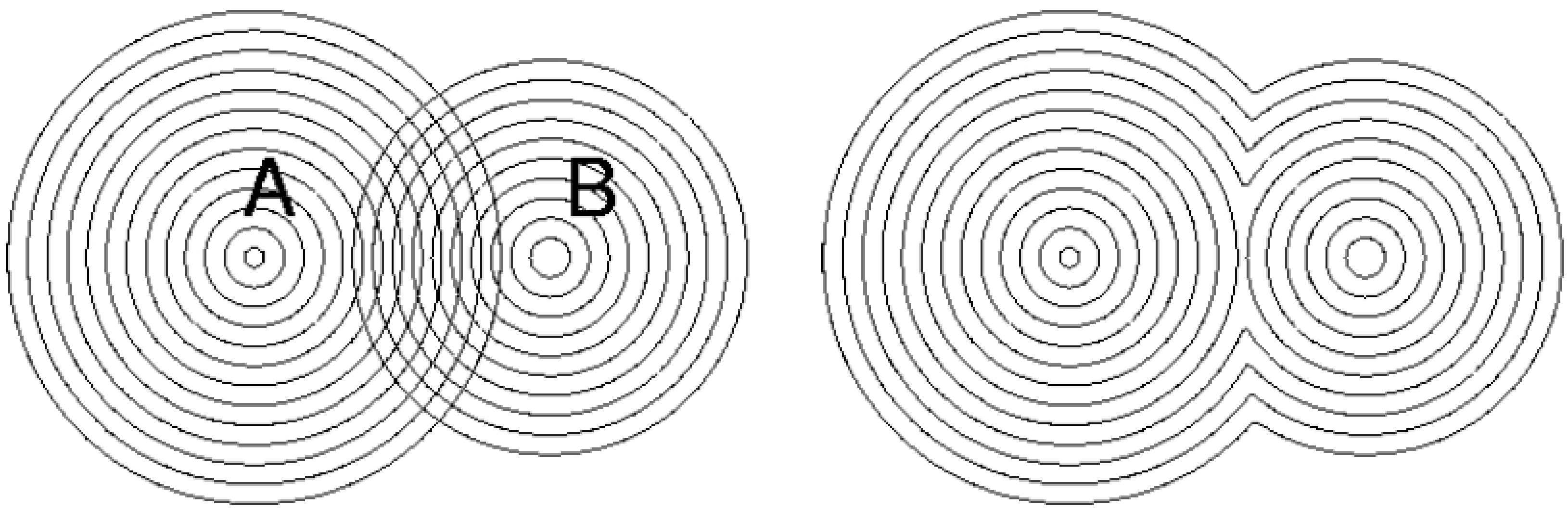


(a) Amplitude $A(x)$



(b) Phase $\phi(x)$

Eikonal equation
First arrival property



(a) Correct solution (b) Eikonal equation

Note:

- $\phi \sim$ traveltime of the wave
- $\phi(x)$ =constant (level sets) represent wave fronts

Eikonal Solvers

Fast marching method (FMM)

In a 2D layered medium

(tracking interface/wavefront motion: curve and surface evolution)

$$\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2 = \frac{1}{c(z)^2}$$

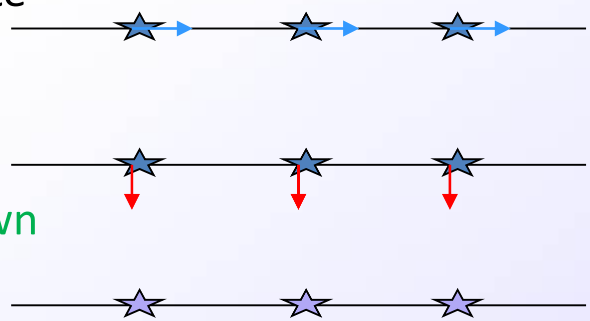
$$\frac{\partial T}{\partial z} = \pm\sqrt{\frac{1}{c(z)^2} - \left(\frac{\partial T}{\partial x}\right)^2}$$

Let us assume T is known at a level z=cte

a) $\frac{\partial T}{\partial x}$

Compute $\partial T/\partial x$ along z=cte by a finite difference approximation

z=cte



Deduce $\partial T/\partial z$ from eikonal

b) $\frac{\partial T}{\partial z}$

z

c) down

Extend T estimation at depth z+dz

z + dz

EIKONAL SOLVER

ONLY FIRST-ARRIVAL!

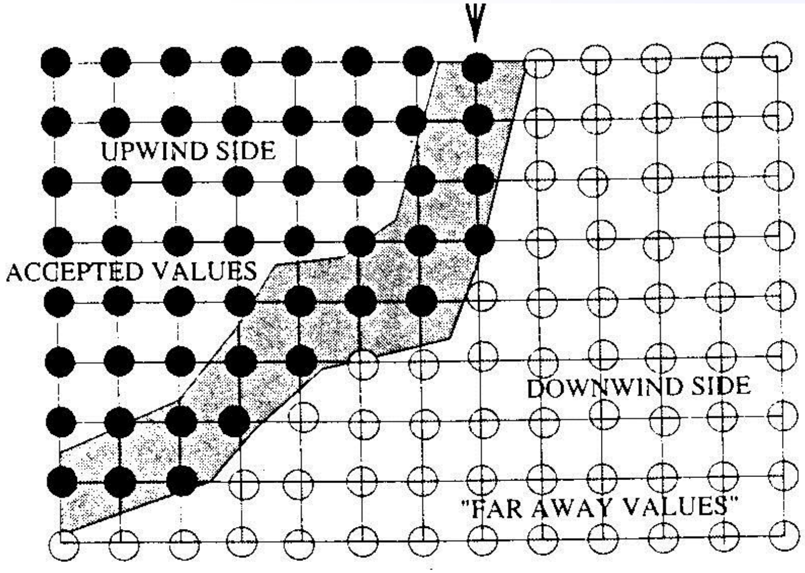
From T at a depth of z, we have been able to estimate T at a depth of z+dz

Fast marching method (FMM)

2D case

$$\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2 = \frac{1}{c(x,z)^2}$$

$$\frac{\partial T}{\partial z} = \pm\sqrt{\frac{1}{c(x,z)^2} - \left(\frac{\partial T}{\partial x}\right)^2}$$



From Sethian & Podovici (1999)

Computing T(x,z) is a stationay boundary problem: discretize it on a grid and find an efficient numerical method to solve it.

The solution is updated by following the causality in a sequential way: updated pointwise in the order the solution is strictly increasing (upwind difference scheme and a heap-sort algorithm)

Sharp interfaces are difficult to describe

$$\mathcal{O}(N\ LogN)$$

where N is the number of grid points in a direction

ENO or WENO stencil

How to estimate the discrete $|\nabla T|$?

A first-order Godunov upwind difference scheme

$$\left[\left(\frac{T_{i,j} - T_{i,j}{}^{xmin}}{h}\right)^{+}\right]^{2} + \left[\left(\frac{T_{i,j} - T_{i,j}{}^{ymin}}{h}\right)^{+}\right]^{2} = \frac{1}{c^{2}(x,z)}$$

with $T_{i,j}{}^{xmin} = \min(T_{i-1,j}, T_{i+1,j})$ and $T_{i,j}{}^{ymin} = \min(T_{i,j-1}, T_{i,j+1})$
and with $(x)^{+} = x \; if \; x > 0$ or $(x)^{+} = 0 \; if \; x \; \leq 0$

High-order improved ENO or WENO stencils (Liu et al, 1994; Jiang and Shu, 1996)

See Sethian's book (1999)

Fast sweeping method (FSM)

2D case

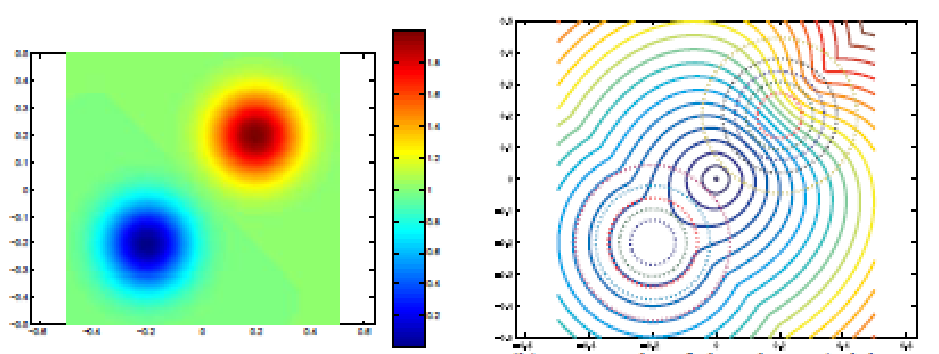
$$(\frac{\partial T}{\partial x})^2+(\frac{\partial T}{\partial z})^2=\frac{1}{c^2(x,z)}$$

Computing T(x,z) is a stationay boundary problem

Gauss-Seidel iterations with alternating direction sweepings are incorporated into same upwind finite difference stencil: complexity in $\mathcal{O}(N)$.

Velocity

Travel time



(a) graph of $f(x, y)$

(b) contour plot of the solution (solid line) and $f(x, y)$ (dashed line)

In 2D at least four sweeps are needed and six sweeps in 3D

Iterations are independent of the grid size

From Zhao (2005)

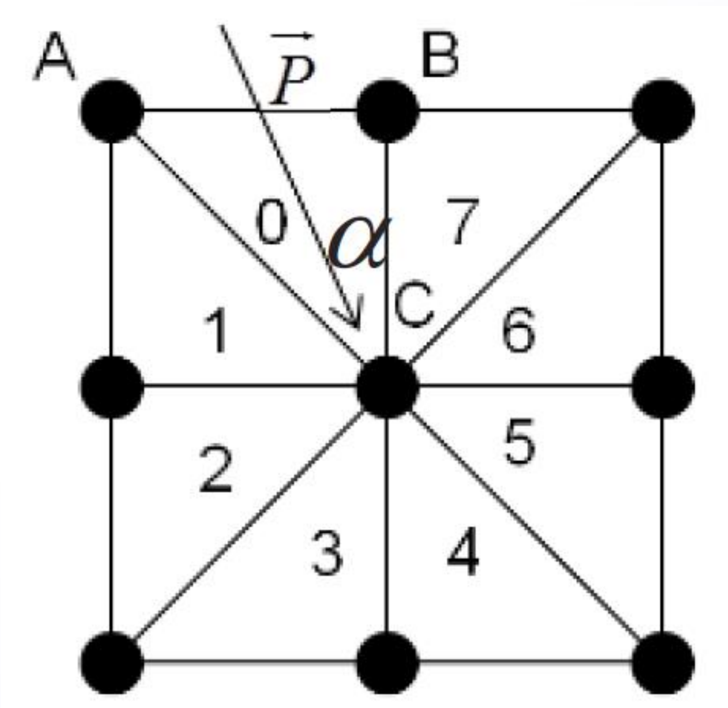
Singularities at the boundary may induce errors, especially for a point source (Luo & Qian, 2010)

Any mesh could be used

Fast sweeping method (FSM)

2D case $(p_x)^2+(p_z\ )^2=\dfrac{1}{c^2(x,z)}$

Any mesh could be used



$$T_C = T_A + \frac{\overrightarrow{AC}.\vec{p}}{AC}\mathrm{AC} \qquad T_C > T_A$$

$$T_C = T_B + \frac{\overrightarrow{BC}.\vec{p}}{BC}\mathrm{BC} \qquad T_C > T_B$$

Waves travel along AC or BC direction with an apparent slowness which is the projected slowness value from the true slowness vector.

$$\begin{pmatrix} \dfrac{x_C - x_A}{AC} & \dfrac{z_C - z_A}{AC} \\ \dfrac{x_C - x_B}{BC} & \dfrac{z_C - z_B}{BC} \end{pmatrix} \begin{pmatrix} p_x \\ p_z \end{pmatrix} = \begin{pmatrix} \dfrac{T_C - T_A}{AC} \\ \dfrac{T_C - T_B}{BC} \end{pmatrix}$$

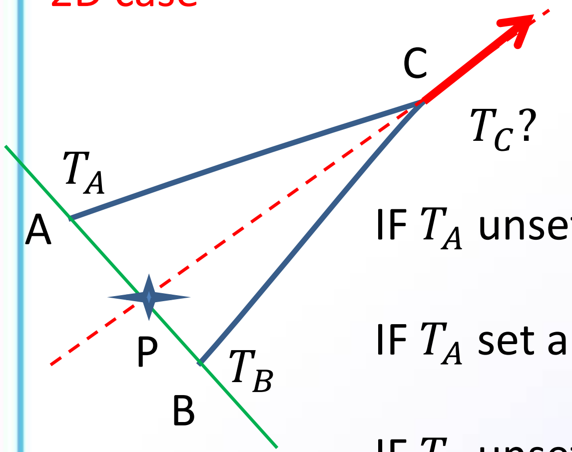
$$\begin{pmatrix} p_x \\ p_z \end{pmatrix} = \begin{pmatrix} MT_C + N \\ PT_C + Q \end{pmatrix} \Longrightarrow (MT_C + N)^2 + (PT_C + Q)^2 = \frac{1}{c^2(x,z)}$$

Han et al (2015)

Quadratic equation: real solution T_C needed!

Causality and viscosity

2D case



$$(MT_C + N)^2 + (PT_C + Q)^2 = \frac{1}{c^2(x,z)}$$

$$(M^2 + P^2){T_C}^2 + 2(MN + PQ)T_C + N^2 + Q^2 - \frac{1}{c^2(x,z)} = 0$$

equation (1)

IF T_A unset and IF T_B unset, do nothing

IF T_A set and IF T_B unset, compute T_C as if wave comes from A

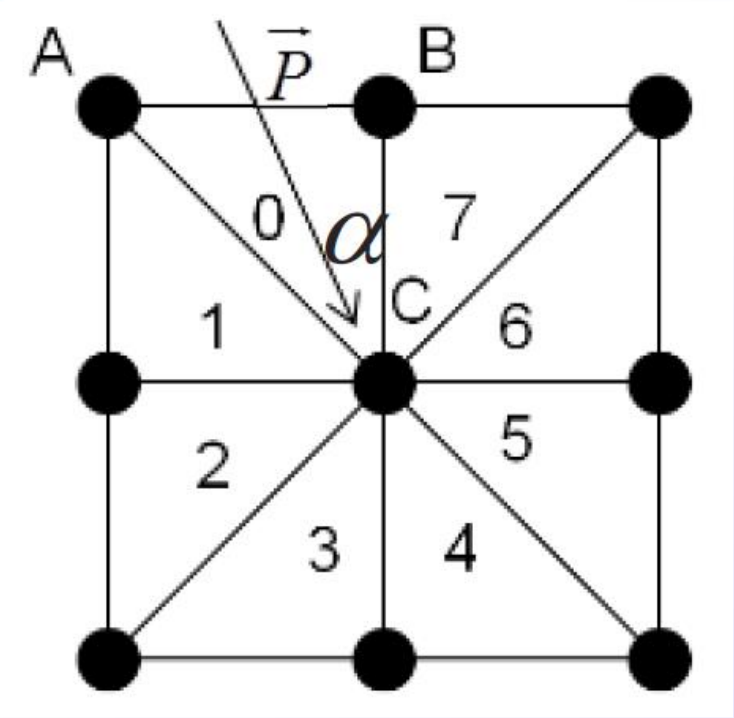
IF T_A unset and IF T_B set, compute T_C as if wave comes from B

IF T_A set and IF T_B set, do
compute roots of equation (1)
if real solution
check if the « backward ray » intersects the segment [AB]
if yes, update T_C by this value if it is smaller
if no real solution, compute the viscous solution
as wave is coming from A or from B: select the smallest value
enddo

We provide always a value if A and/or B have a value
\# ray tracing

Fast sweeping method (FSM)

2D case



We need to setup a set of values which will be fixed:
at the source, $T = 0$, for example

Eight points stencil:

From the possible eight values (if one is set), take the smallest one.

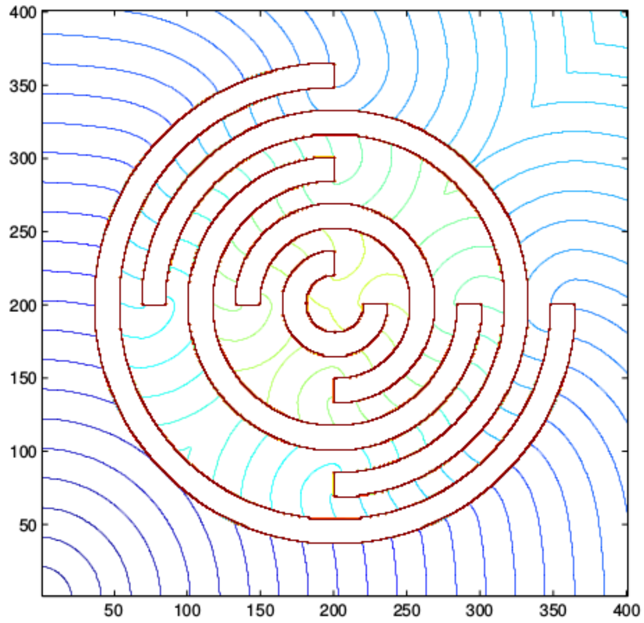
Sweeping technique:

Four sweeping when applying the stencil

Sweep 1: $i_i = 1, n_1, 1; i_2 = 1, n_2, 1$
Sweep 2: $i_i = n_1, 1, -1; i_2 = 1, n_2, 1$
Sweep 3: $i_i = n_1, 1, -1; i_2 = n_2, 1, -1$
Sweep 4: $i_i = 1, n_1, 1; i_2 = n_2, 1, -1$

Iteration over sweeps until convergence (no more updating of T_C)

Number of iterations



(c) distance to a point with five ring obstacle

The number of iterations depend on the medium structure: one must be aware that the characteristics of the hyperbolic system should be sampled at least once by the sweeping loop.

In seismics or seismology, we have less dramatic configuration than the one shown in this figure

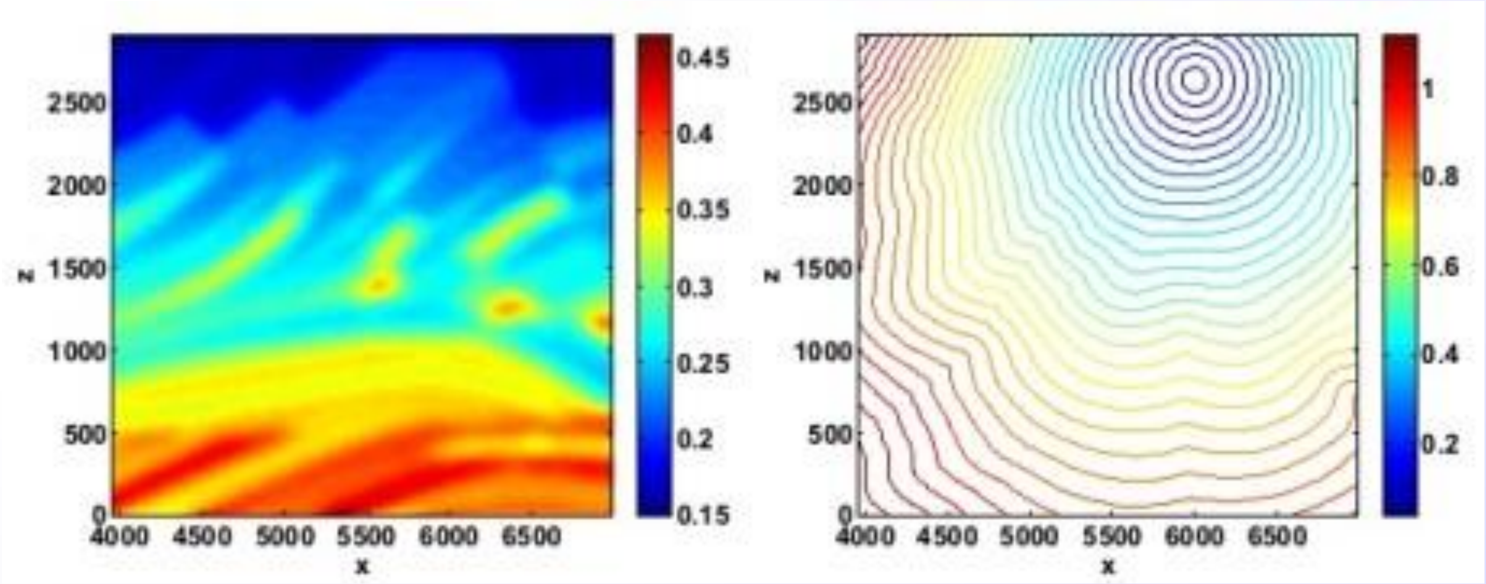
Few iterations are necessary to achieve convergence

Easy extension to anisotropy …

2D examples

A smooth model constructed from Marmousi

Wavefronts as deduced from travel-time computation



(Luo and Qian, 2008)

Only first-arrival times

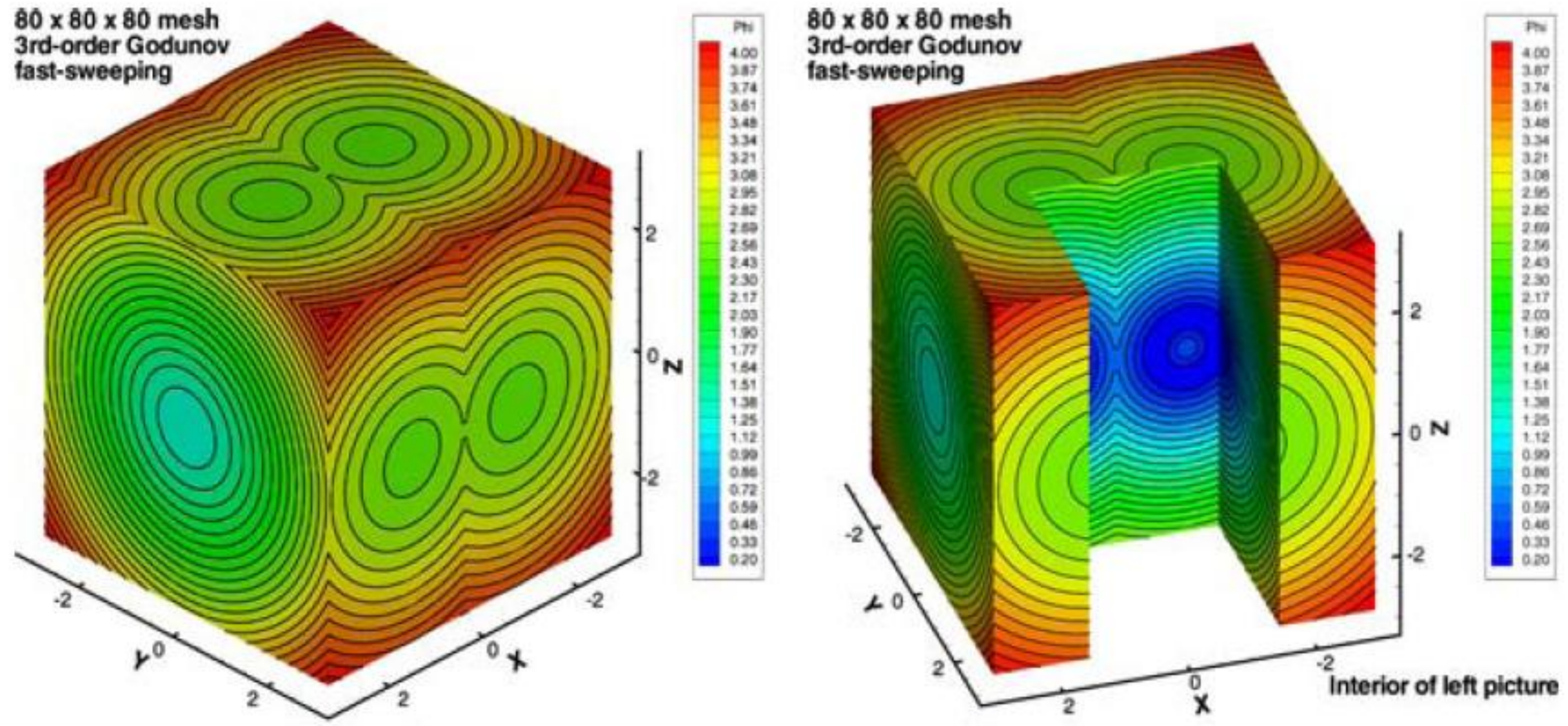
See Taillandier et al (1999) for an application to first-arrival time tomography using the adjoint formulation

3D example

$$\begin{cases} H(\phi_{x_1}, \ldots, \phi_{x_d}, x) = 0, & x \in \Omega \setminus \Gamma, \\ \phi(x) = g(x), & x \in \Gamma \subset \Omega, \end{cases}$$

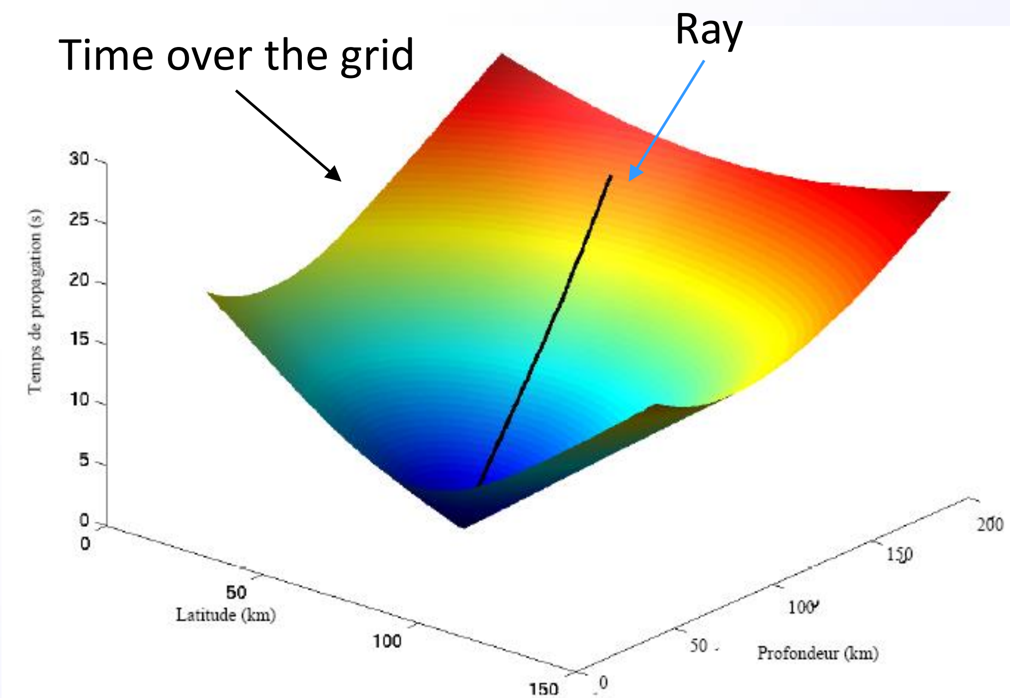
Static Hamilton–Jacobi Equations

Zhang, Zhao and Qian (2005)



Drawback: only first-arrival times!

Still very usefuf to back-raytracing once we know times: getting the Jacobian matrix



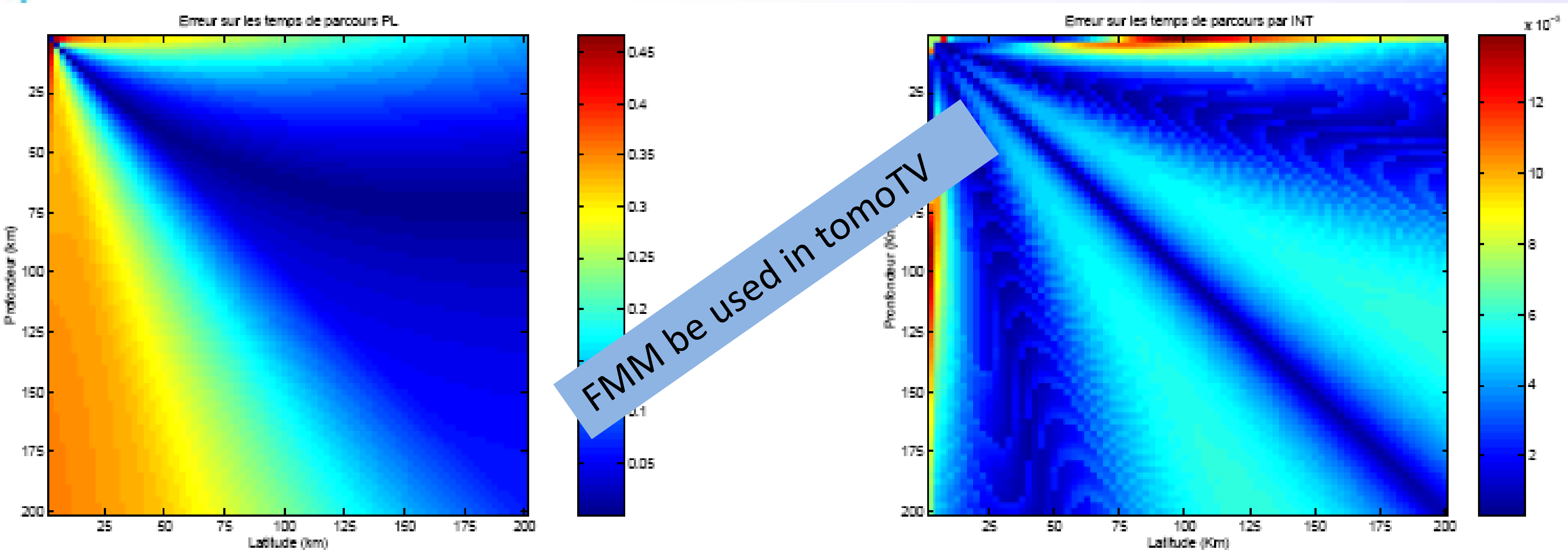
Once traveltime T is computed over the grid for one source, we may backtrace using the gradient of T from any point of the medium towards the source (should be applied from each receiver)

The surface {MIN TIME} is convex as time increases from the source : one solution !

Back to inversion through rays

Could we do better: multi-arrival times and amplitudes?

Time error over the grid (FMM)



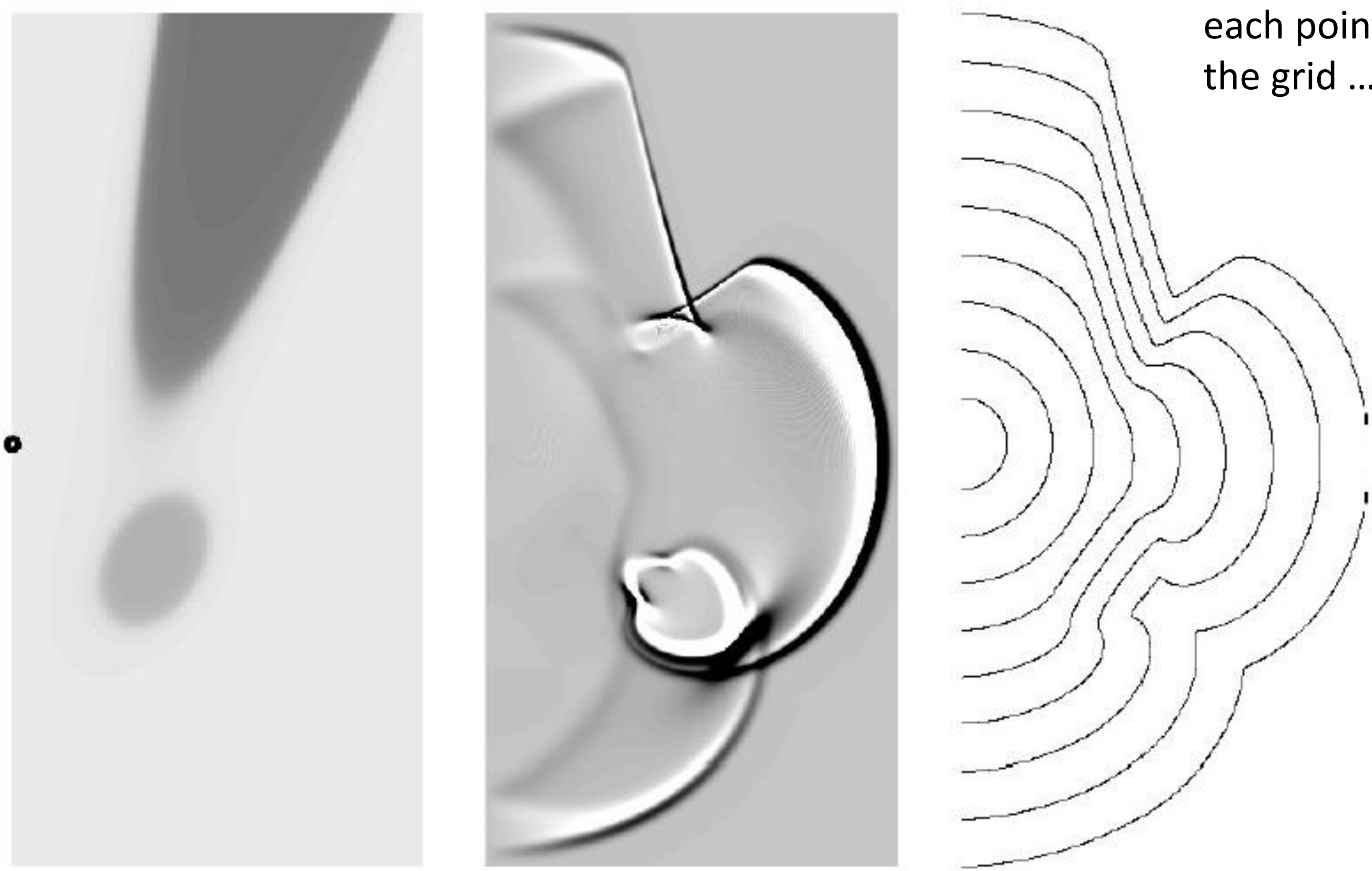
Errors through FMM times

Errors through rays deduced after FMM times

NOT THE SAME COLOR SCALE (factor 100)

Coarser grid for computation

Eikonal equation
Example

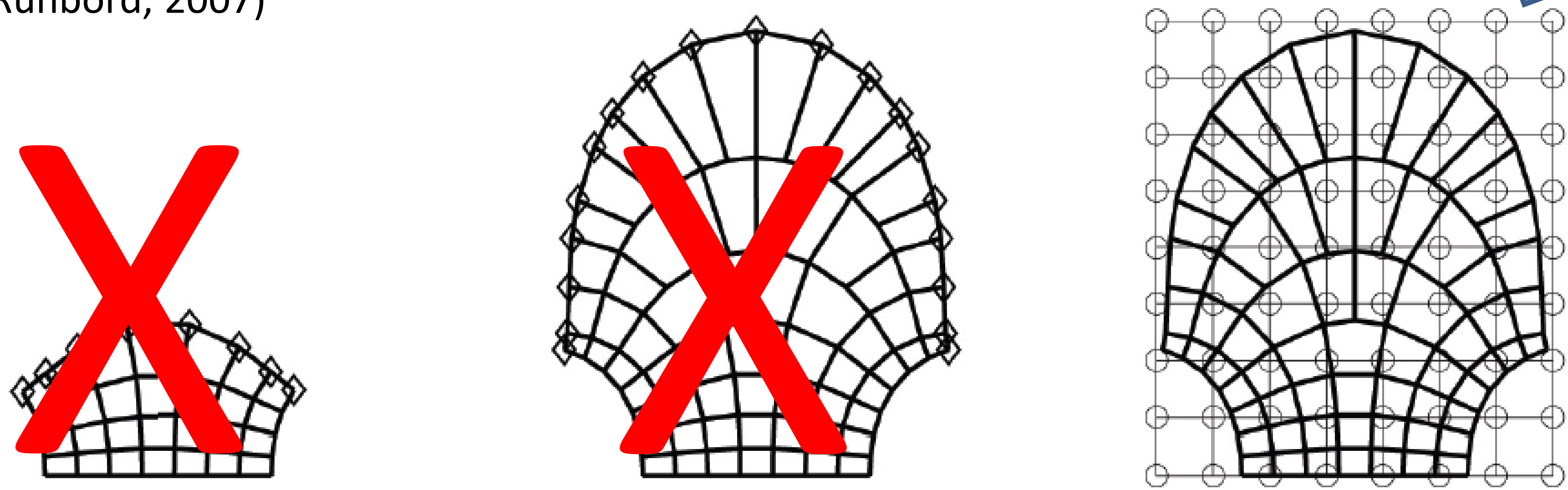


In fact, time defines at each point of the grid …

Wavefront construction [Vinje, Iversen, Gjöystdal, Lambaré, ...]

- Solve for $\boldsymbol{x}(t, \alpha)$ and $\boldsymbol{p}(t, \alpha)$. Discretize in α and trace rays for $\alpha_1, \alpha_2, \alpha_3, \ldots$ where $\alpha_j = j\Delta\alpha$.
- Insert new rays adaptively by interpolation when front resolution deteriorates. E.g.:
  If $|\boldsymbol{x}(t_n, \alpha_{j+1}) - \boldsymbol{x}(t_n, \alpha_j)| \geq tol$ then insert new ray at $\alpha_{j+1/2}$.
- Interpolate traveltime/phase/amplitude onto regular grid.

(Runbord, 2007)



$\boldsymbol{x}(t_n, \alpha_j)$

Conclusion Rays and Waves

- Geometrical optics: ODE versus PDE
  - Choose PDE when possible!

- ODE: tracing one (paraxial) ray is fast
  - Please trace paraxial rays as incremental cost

- Keep complexity low (seismic waves are finite frequency waves)
  - Do not drown yourself into the no-scale optical singularities

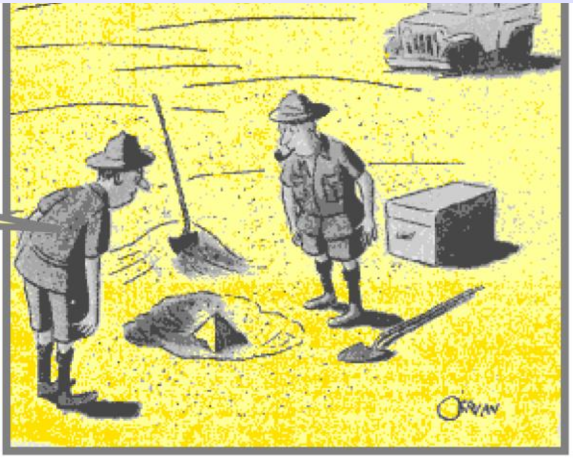
- Identification of rays is a key problem using Fréchet kernel

Perspectives Rays and Waves

- Fast sweeping method O(N) for travel-times and for amplitudes
  - Amplitude equations have been designed
- Finite element methods put into the scene
  - Stencils are moving to higher orders and h-adaptivity
- Discontinuous Galerkin methods
  - This is the road to take for interface investigation in the frame of PDE.

Inverse problem

This could be the discovery of the century. Depending, of course, on how far down it goes…



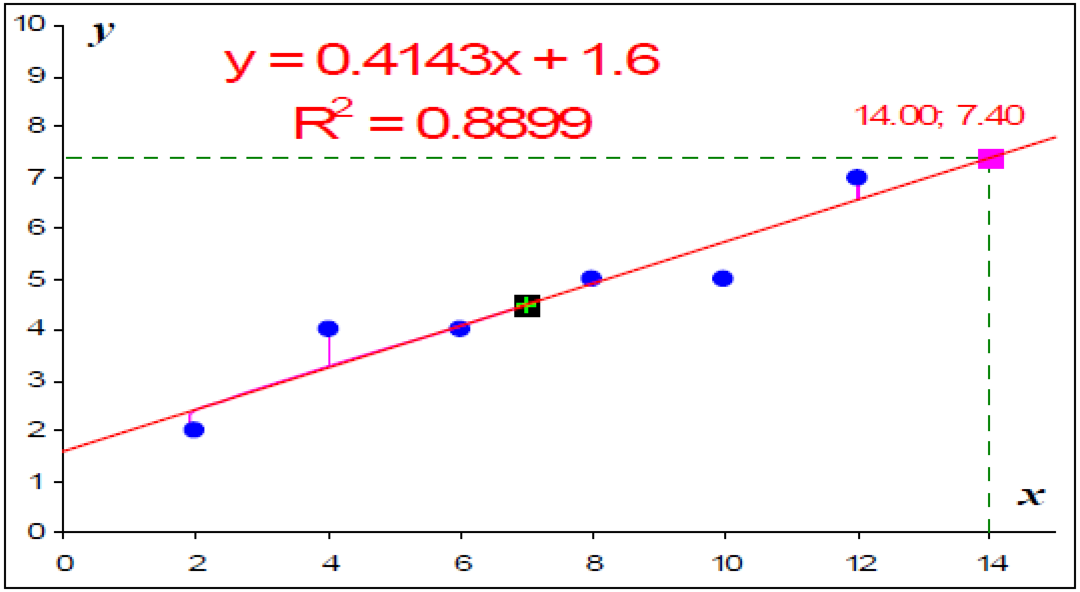
Other inverse problems?

Few hints

Least-squares method

From Excel

DATA		MODEL
x	y	$y' = ax+b$
2	2	2.42857
4	4	3.25714
6	4	4.08571
8	5	4.91429
10	5	5.74286
12	7	6.57143
14	??	**7.4**



Sum of vertical distances between data points and expected y values from the unknown line y=ax+b should be minimum: find a and b?

It is an inversion ….

Occam's razor

- Do not use more complicated maths than the data deserves

- Approximate the least constrained quantity

Given: data (observed and modeled)
Assumed: wave propagation
Unknown: Earth structure

- Occam's Razor: parcimonious principle

When you have many explanations for predicting exactly the same quantities and that there is no way to distinguish them, select the simplest one… until you end up with a contradiction.

Ockham (~1295-~1349)

Approximations can't be avoided …

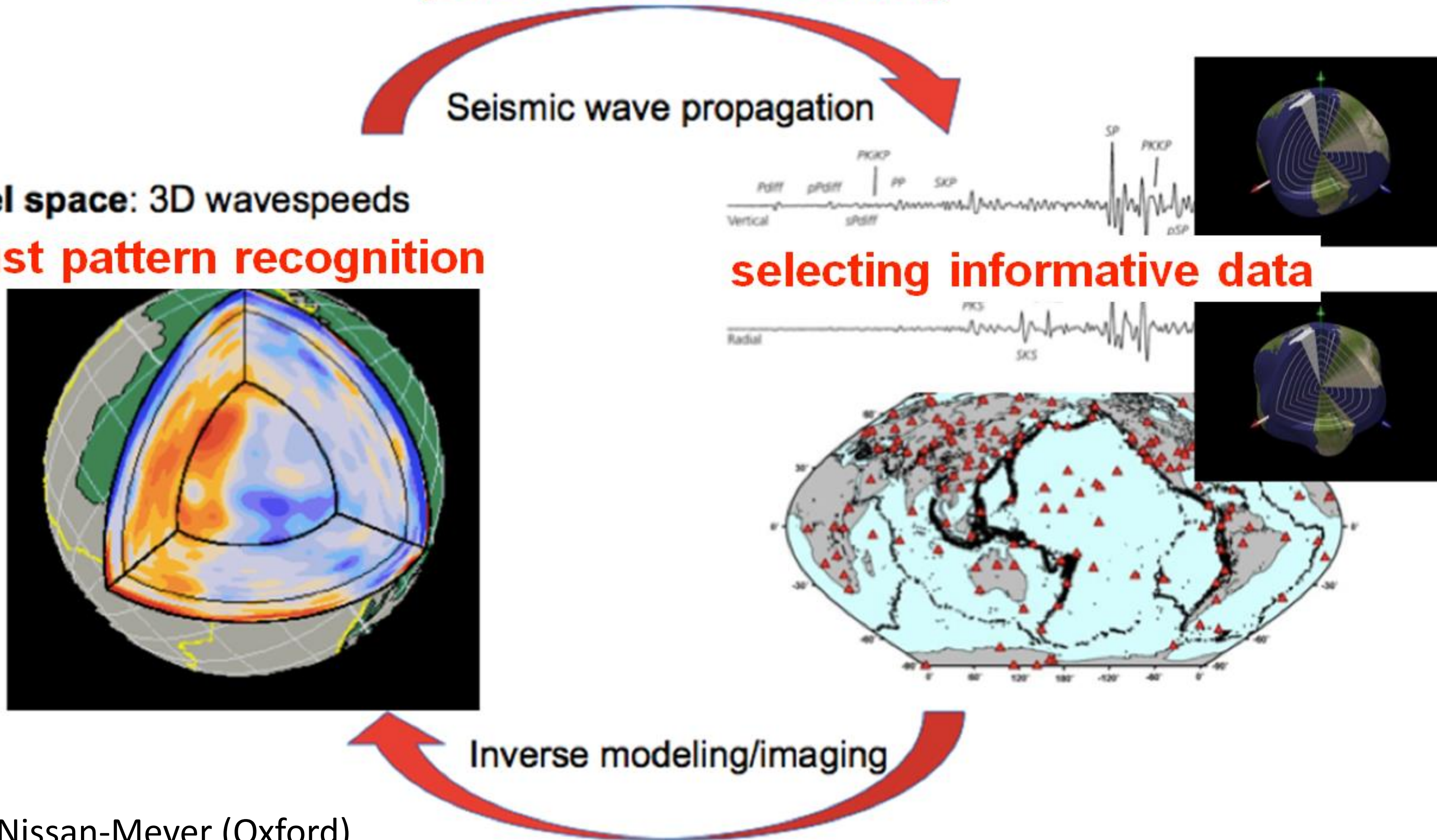
- ❑ Incomplete data
- ❑ Limited understanding
- ❑ Limited crunching capacity

« An approximate solution to a real problem is better than an exact solution to an ideal problem » (Tarantola, 2007)

- ➢ Limited frequency range: translucid Earth
- ➢ Limited dataset: direct access impossible (or difficult)
- ➢ Data windowing (observables): extraction of robust information
- ➢ Dimensionality (2D vs 3D …)
- ➢ Structural complexity (multi-scale heterogeneities)
- ➢ Physics approximation (elastic vs acoustic)

The sparse analysis cycle ...

parsimonious modeling



robust pattern recognition

selecting informative data

sparsity constraints for inference

From Nissan-Meyer (Oxford)

How to reconstruct the velocity structure ?

- Forward problem (easy)

  from a known velocity structure, it is possible to compute travel times, emergent distance and amplitudes.

- Inverse problem (difficult)

  from travel times (or similarly emergent distances), it is possible to deduce the velocity structure : this is the time tomography

  even more difficult is the diffraction tomography related to the waveform and/or the preserved/true amplitude.

Tomographic approach

- Very general problem

  medicine; oceanography, climatology; physics …

- Difficult problem when unknown prior medium (travel time tomography)

- Easier problem if a first medium could be constructed: perturbation techniques can be used for improving the reconstruction (delayed travel time tomography)

Travel Time tomography

v(z) ~ u(z) : u slowness
z(u) is the one we got

- We must « invert » the travel time or the emergent distance for getting z(u): we select the distance.

$$X(p) = 2\int_0^{z_p} \frac{p}{\sqrt{u^2(z) - p^2}}\,dz\ ; \quad \frac{X(p)}{2p} = \int_{u_0^2}^{p^2} \frac{dz/du^2}{\sqrt{u^2 - p^2}}\,du^2$$

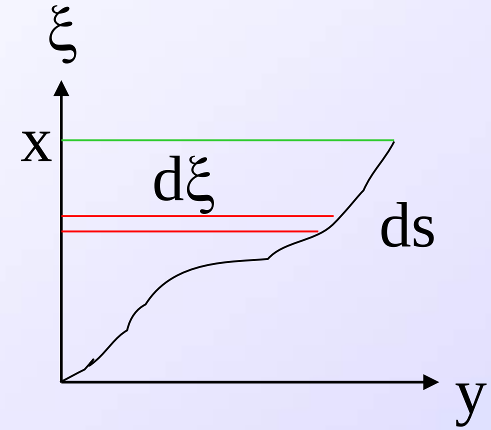
- Abel problem (1826)

Determination of the shape of a hill from travel times of a ball launched at the bottom of the hill with various initial velocity and coming back at the initial position.

ABEL PROBLEM

A point of mass ONE and initial velocity v_0 reaches a maximal height x given by

$$gx = \frac{1}{2} v_0^2$$



We shall take as the zero value for the potential energy: this gives us the following equations and its integration:

$$\left(\frac{ds}{dt} \right)^2 = 2g(x - \xi) \;\; ; \;\; t(x) = \int_0^x \frac{ds/d\xi}{\sqrt{2g(x-\xi)}} d\xi$$

We may transform it into the so-called Abel integral

$$t(x) = \int_0^x \frac{f(\xi)}{\sqrt{x - \xi}} d\xi$$

where t(x) is known and f(ξ) is the shape of the hill to be found: this is an integrale equation.

The exact inverse solution

We multiply and we integrate

$$\int_0^\eta \frac{t(x)}{\sqrt{\eta - x}} dx = \int_0^\eta \frac{dx}{\sqrt{\eta - x}} \int_0^x \frac{f(\xi)}{\sqrt{x - \xi}} d\xi$$

We inverse the order of integration

$$\int_0^\eta \frac{t(x)}{\sqrt{\eta - x}} dx = \int_0^\eta f(\xi)\, d\xi \int_\xi^\eta \frac{dx}{\sqrt{\eta - x}\sqrt{x - \xi}}$$

We change variable of integration

$$x = \xi \cos^2\theta + \eta \sin^2\theta$$

$$\int_0^\eta \frac{t(x)}{\sqrt{\eta - x}} dx = \pi \int_0^\eta f(\xi)\, d\xi$$

We differentiate and write it down the final expression

$$\frac{d}{d\eta} \int_0^\eta \frac{t(x)\, dx}{\sqrt{\eta - x}} = \pi \cdot f(\eta)$$

$$f(\xi) = \frac{1}{\pi} \frac{d}{d\xi} \int_0^\xi \frac{t(x)\, dx}{\sqrt{\xi - x}}$$

THE ABEL SOLUTION

By changing variable ξ in a-ξ and x in a-x, we get the standard formulae

$$t(x) = \int_a^x \frac{f(\xi)}{\sqrt{\xi - x}} d\xi$$

$$f(\xi) = -\frac{1}{\pi}\frac{d}{d\xi}\int_\xi^a \frac{t(x)\,dx}{\sqrt{x-\xi}}$$

We must have

t(x) should be continuous,

t(0)=0

t(x) should have a finit derivative with a finite number of discontinuities.

The most restrictive assumption is the continuity of the function t(x).

The solution HWB:
HERGLOTZ-WIECHERT-BATEMAN

$$\frac{X(p)}{2p} = \int_{u_0^2}^{p^2} \frac{dz/du^2}{\sqrt{u^2 - p^2}} du^2$$

$$z(v) = -\frac{1}{\pi} \int_{u_0^2}^{u^2} \frac{X(p)/2p}{\sqrt{p^2 - u^2}} dp^2$$

$$z(v) = -\frac{1}{\pi} \int_{u_0}^{u} \frac{X(p)}{\sqrt{p^2 - u^2}} dp$$

$$z(v) = \frac{1}{\pi} \int_0^{X(u)} \frac{dX}{\cosh(pv)}$$

In Cartesian frame

From the direct solution, we can deduce the inverson solution

After few manipulations, we can move from the Cartesian expression towards the Spherical expression

$$\ln\left(\frac{R}{r(v)}\right) = \frac{1}{\pi} \int_0^{\Delta(r/v)} \frac{d\Delta}{\cosh(pv/r)}$$

We find r(v) as a value of r/v

In spherical frame

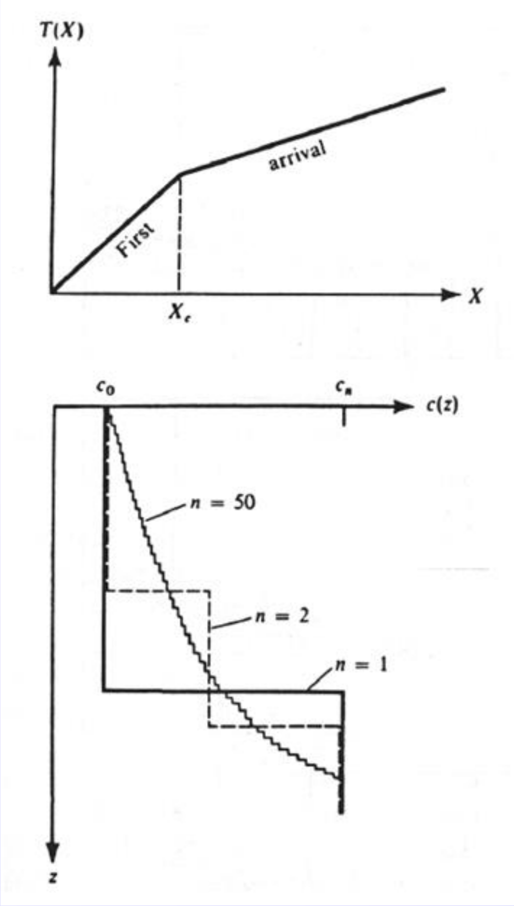
Stratified medium

When considering discontinuities, ABEL/HWB method based on first arrival times has not an unique solution.

More over there will be an ambiguity when the velocity decreases (can only defined the velocity jump in this zone)

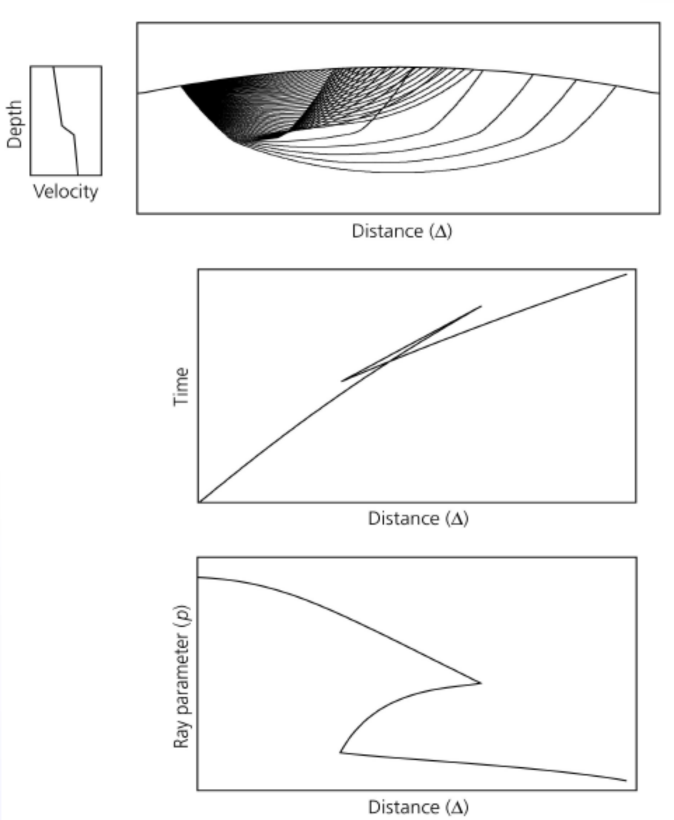
In fact, we avec an infinity of solutions when considering only direct and refracted waves.

We may find interfaces when considering all waves (including reflection waves)

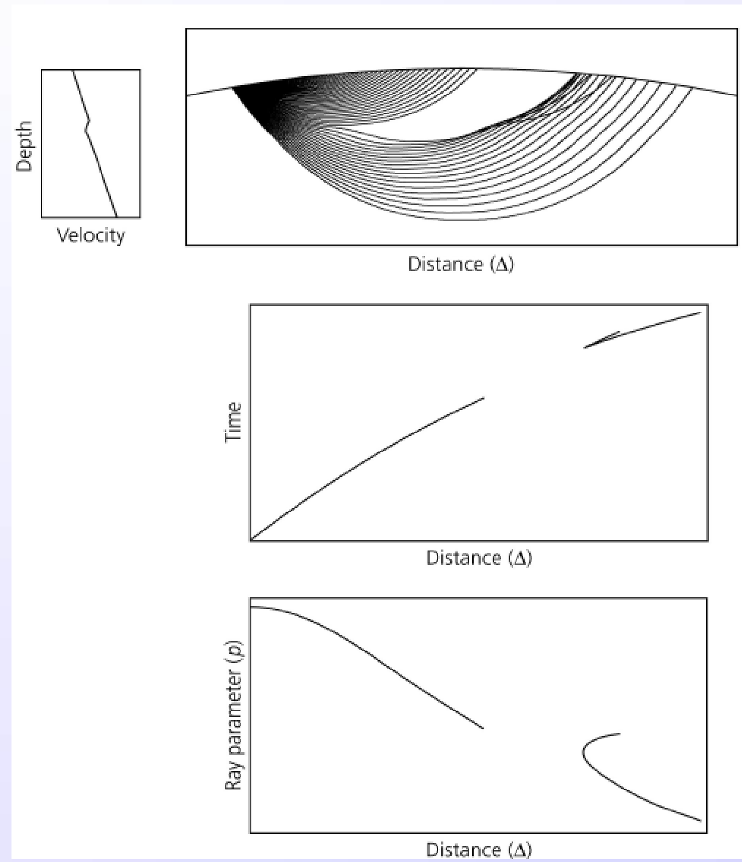


From Aki & Richards (1980)

Triplication & shadow zone

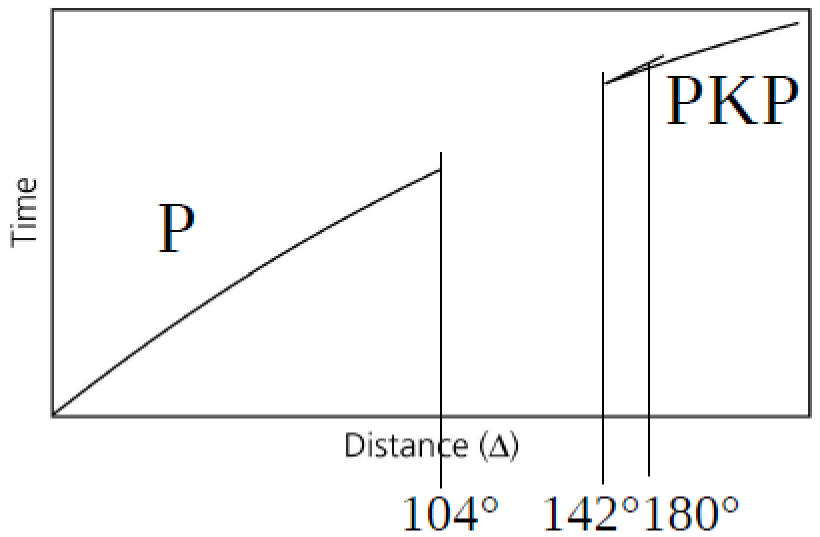


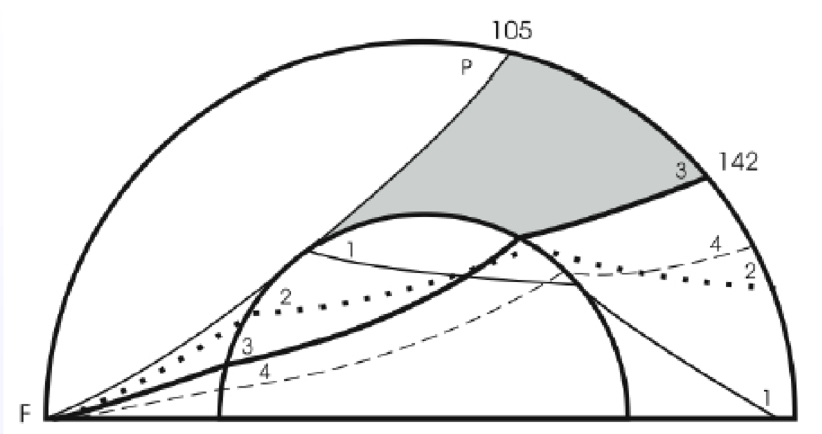
Increase of the velocity: triplication and retrograde branch



Decrease of the velocity: shadow zone (and a retrograde branch)

Shadow zone in the Earth



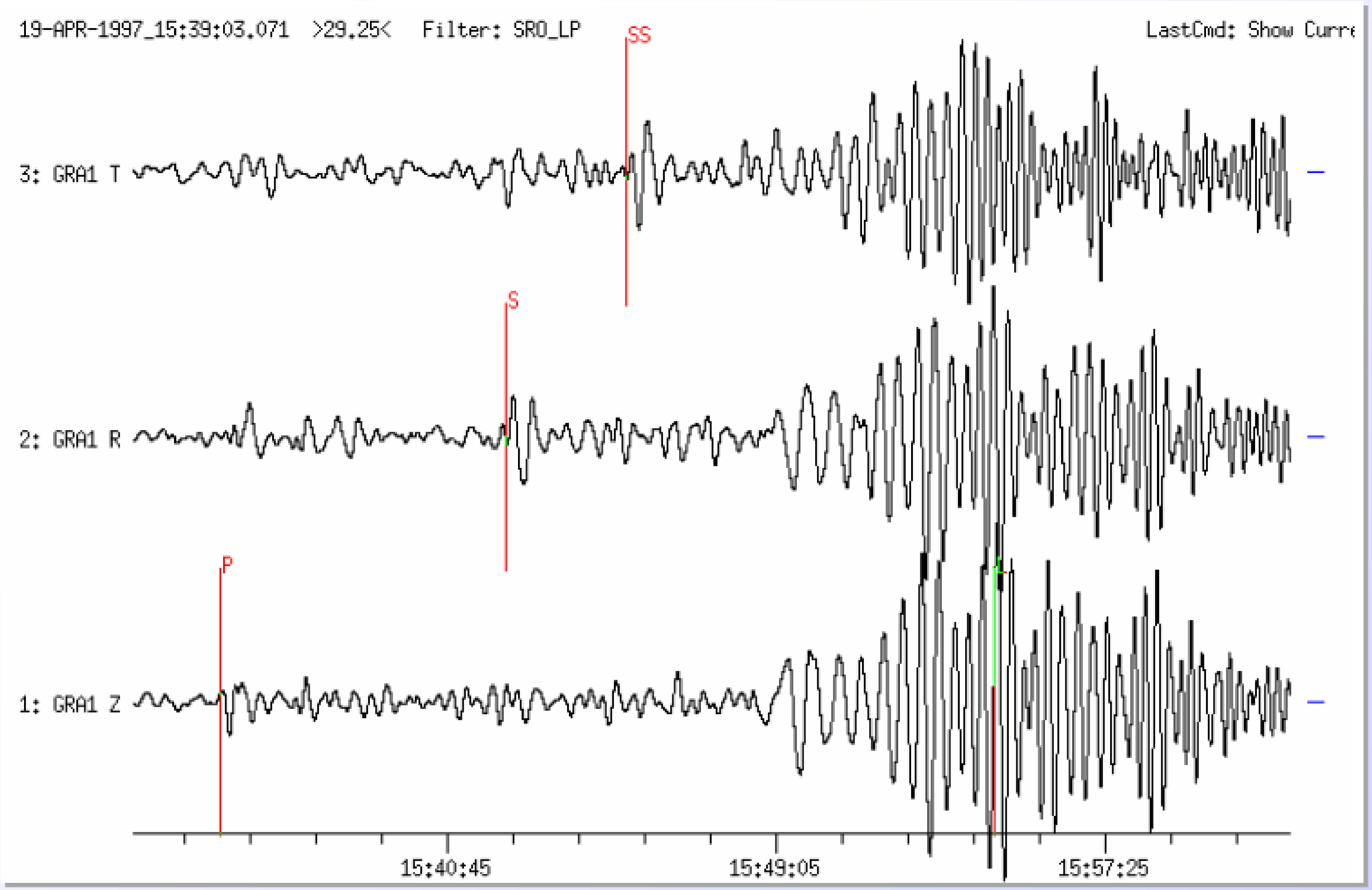


Abrupt decrease of velocity at the CMB (2900 km

(Oldham, 1906, Gutenberg, 1912)

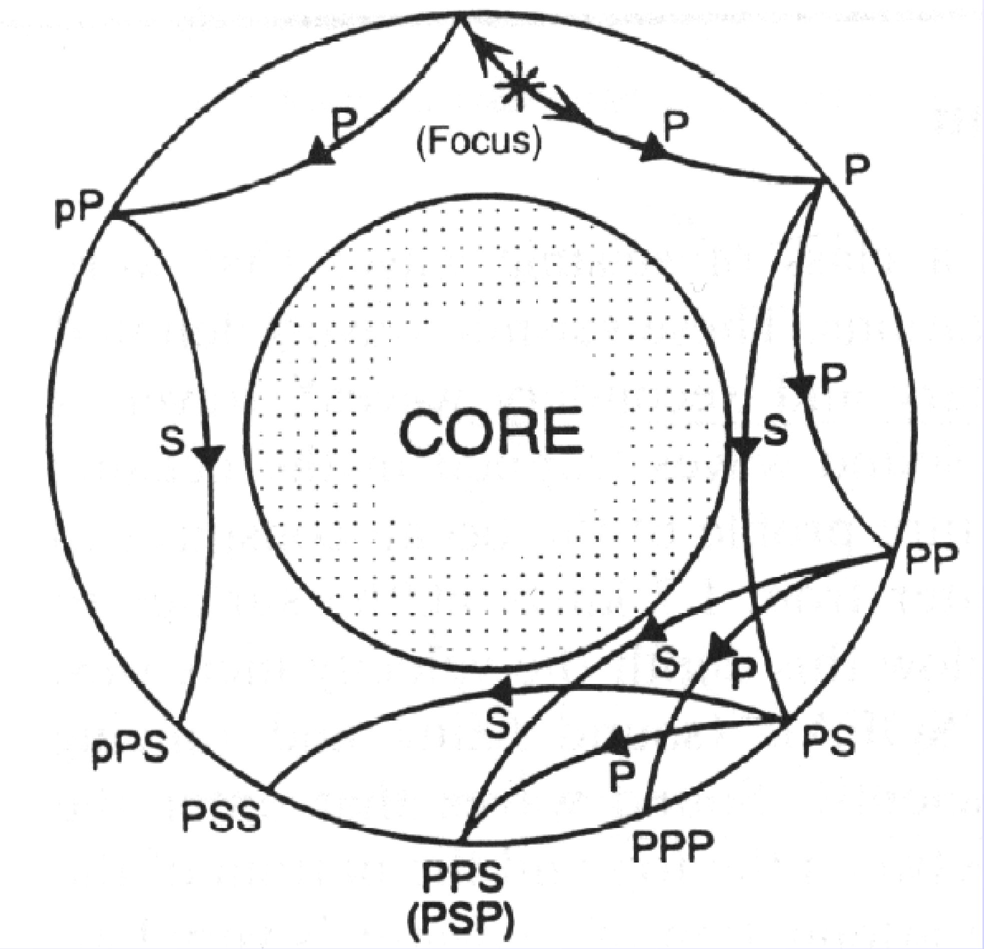
The most annoying
problem: no output!

Pick & Identify phases ?



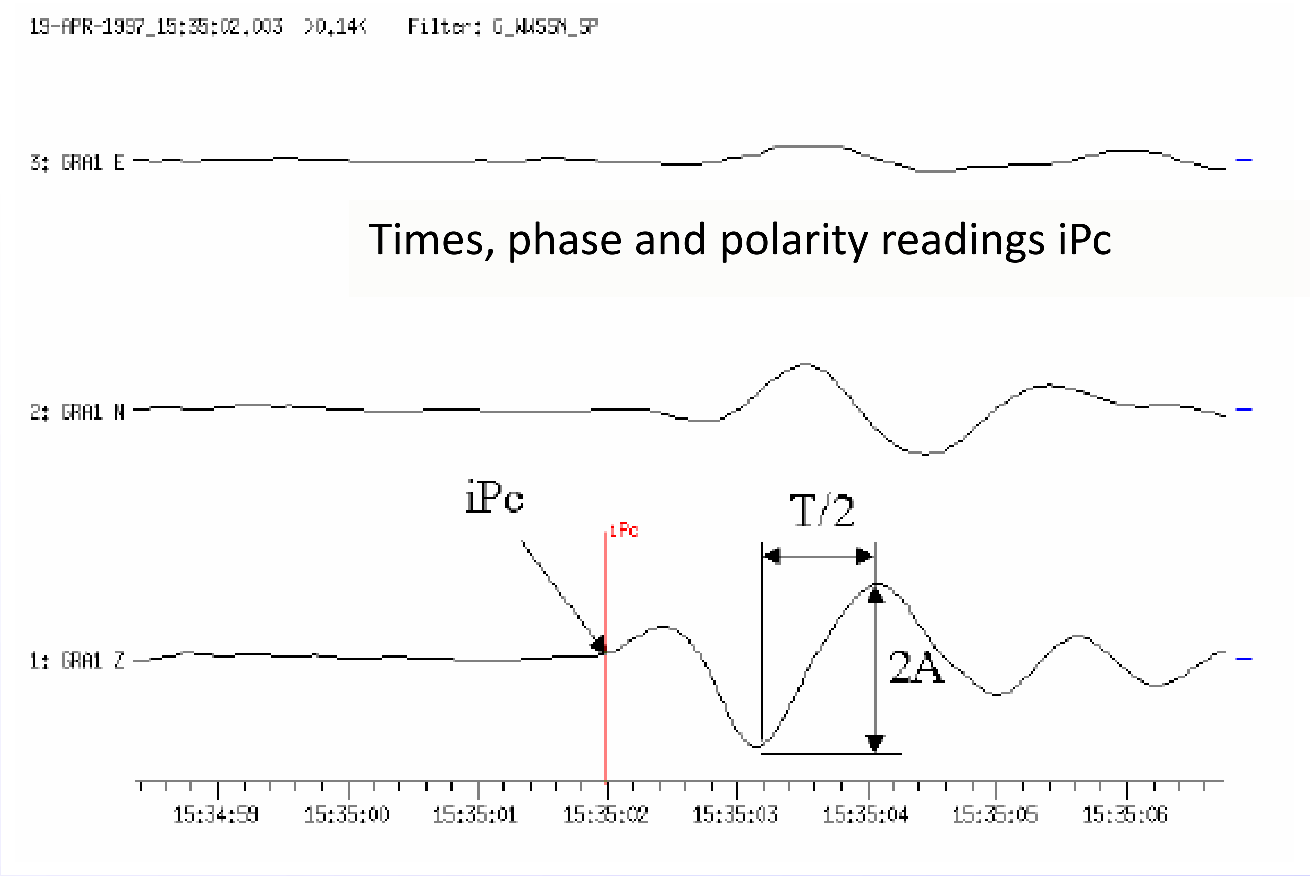
Identify phases ?

Draw the SS phase?



Picking time?

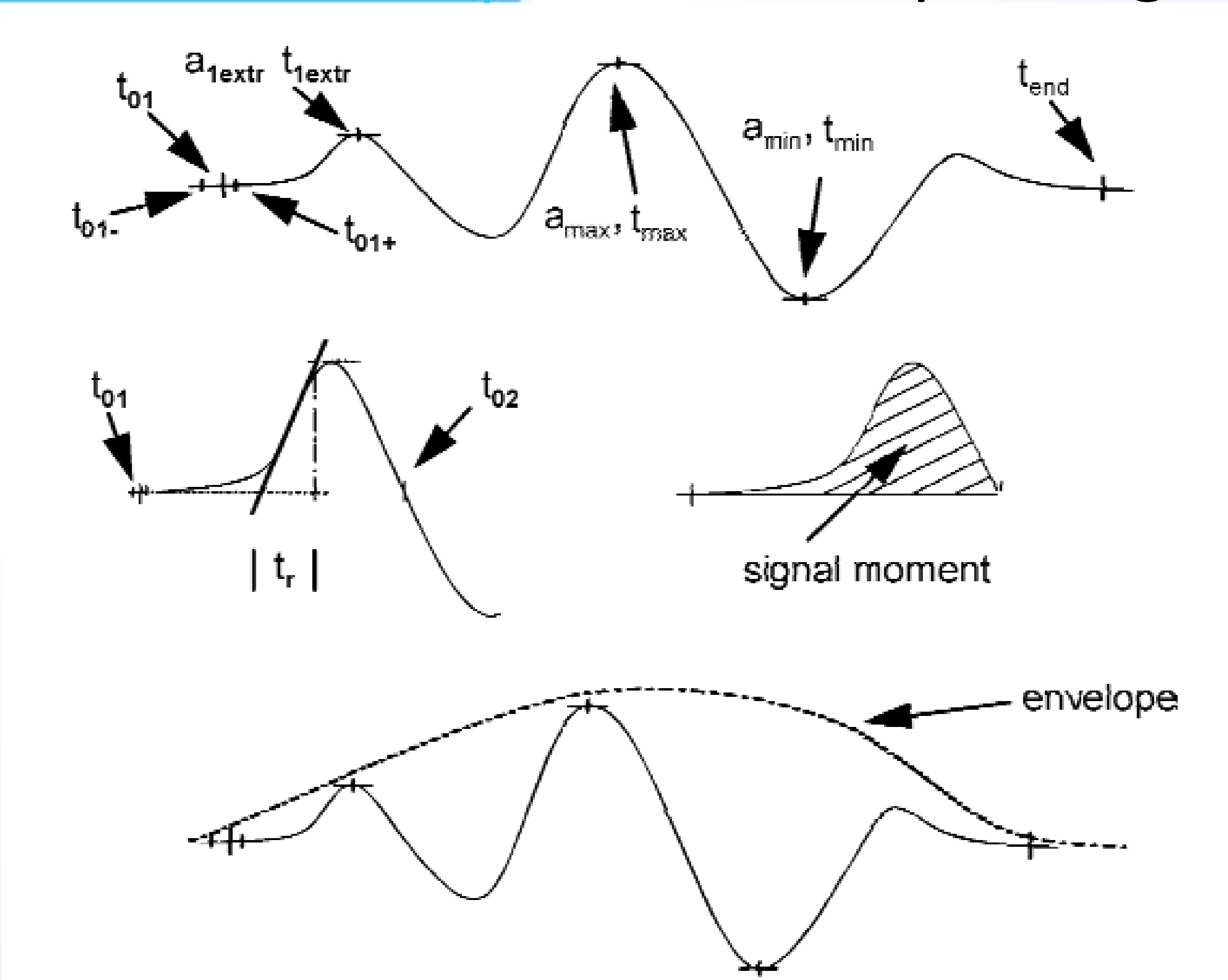
- Onset time for first-arrival phases?



Times, phase and polarity readings iPc

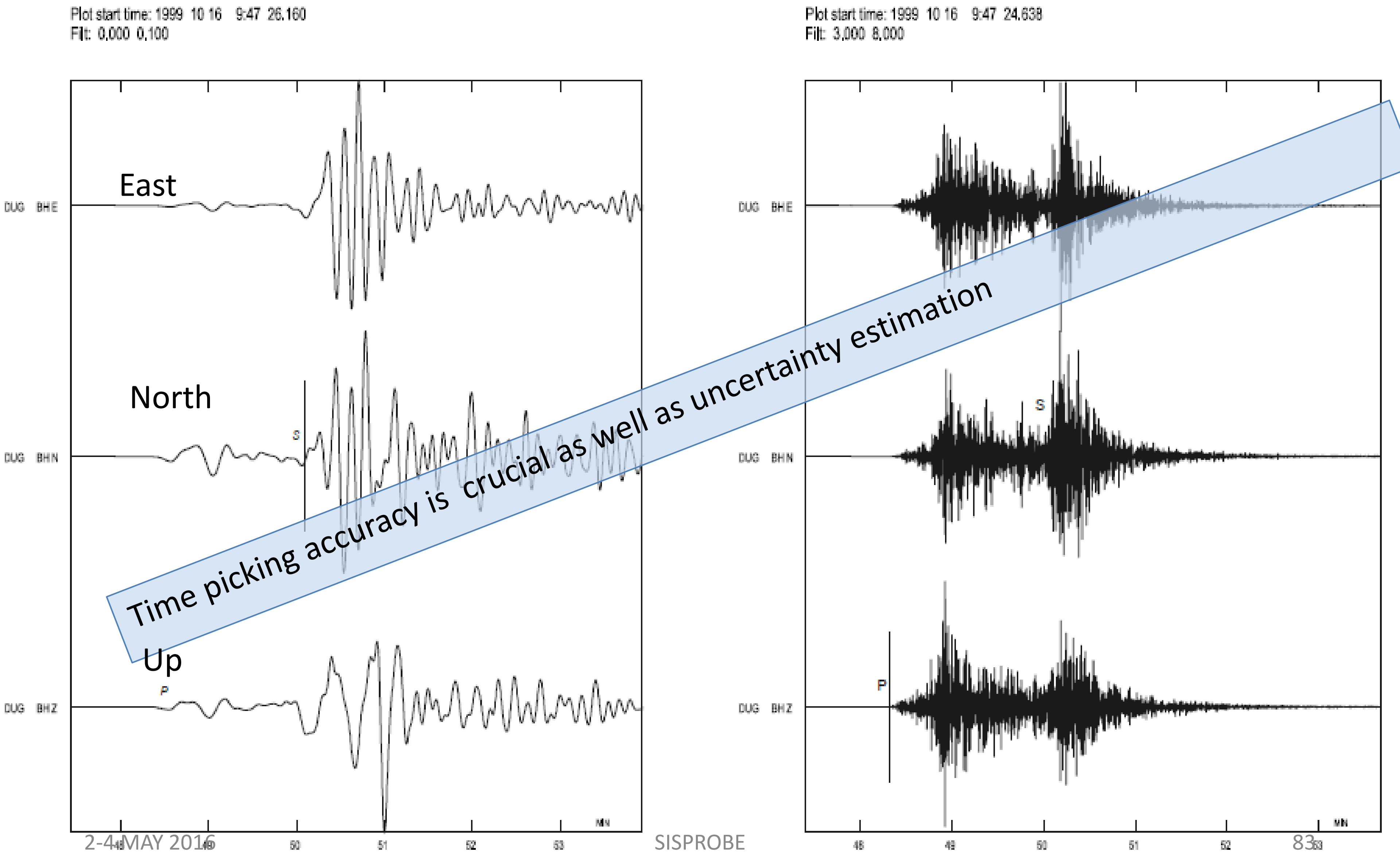
Picking time?

- Deciphering the waveform



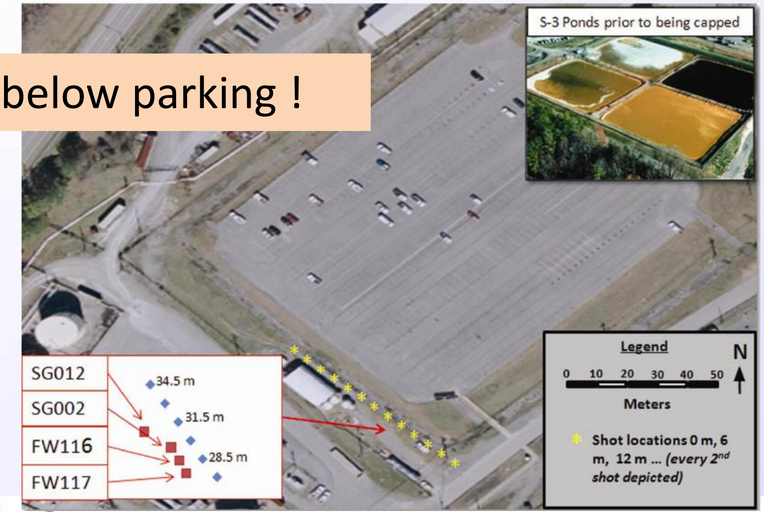
Finite frequency content of the seismic signal

Low-pass or band-pass



Local shots

Water ponds below parking !

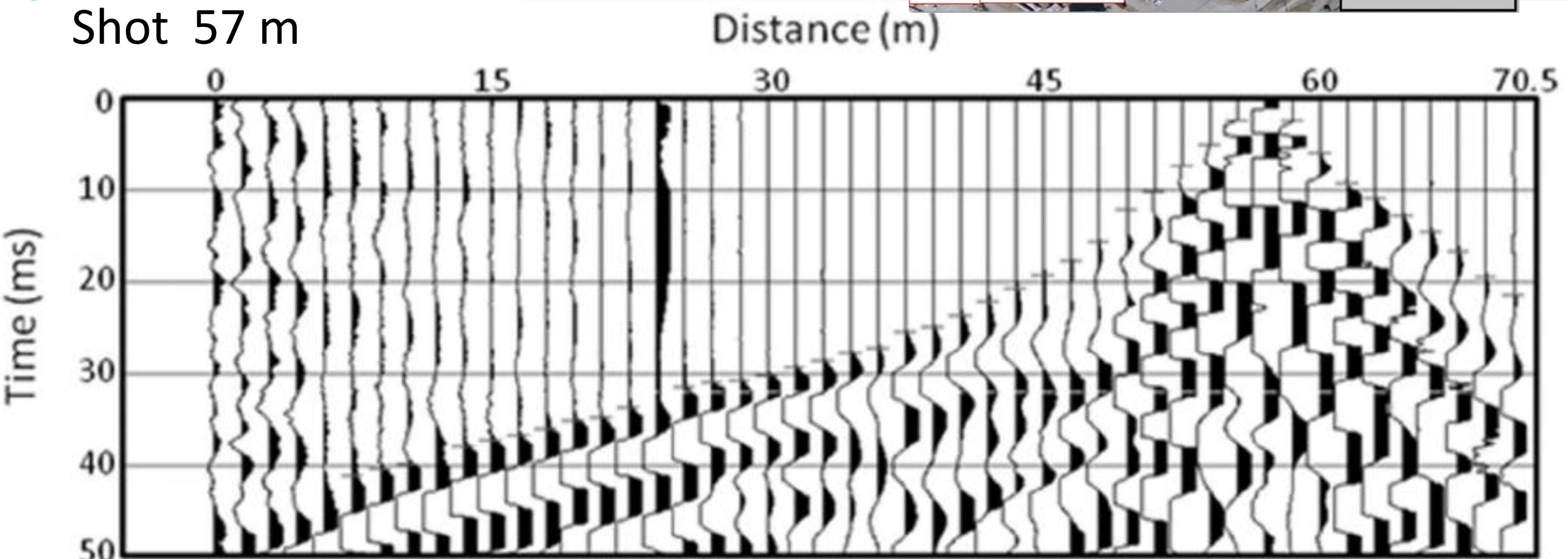


- Picking onset times

Profile 70.5 m

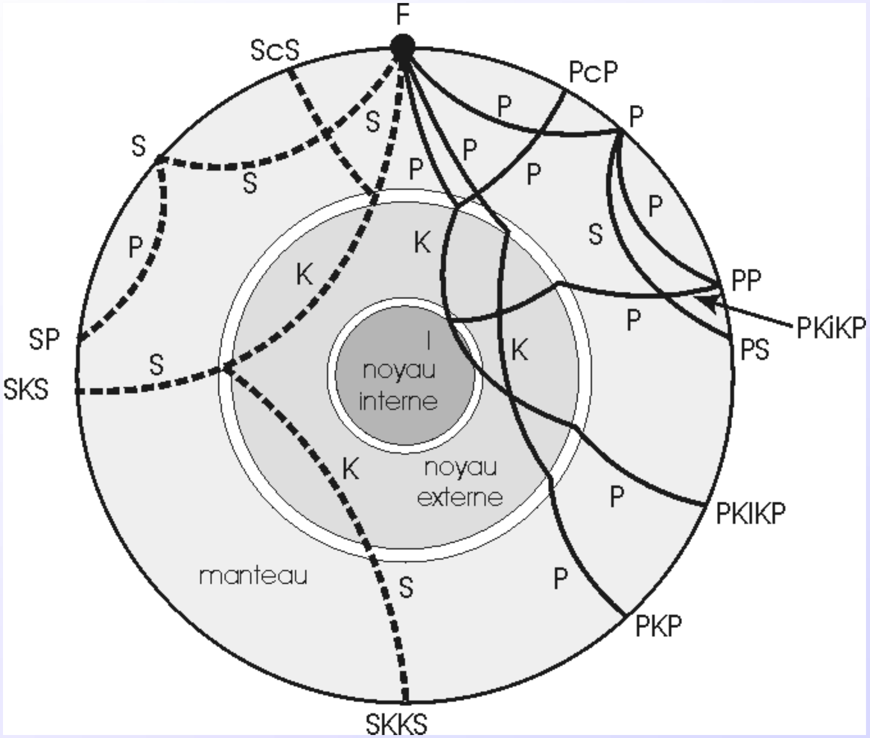
Geophone distance 1.5 m

Shot 57 m



Phase signature

- P : mantle P wave
- S : mantle S wave
- K : Outer core P wave
- I : Inner core P wave
- J : Inner core S wave
- c : Reflected P wave (outer core)
- i : Reflected P wave (inner core)
- m : number of reflections

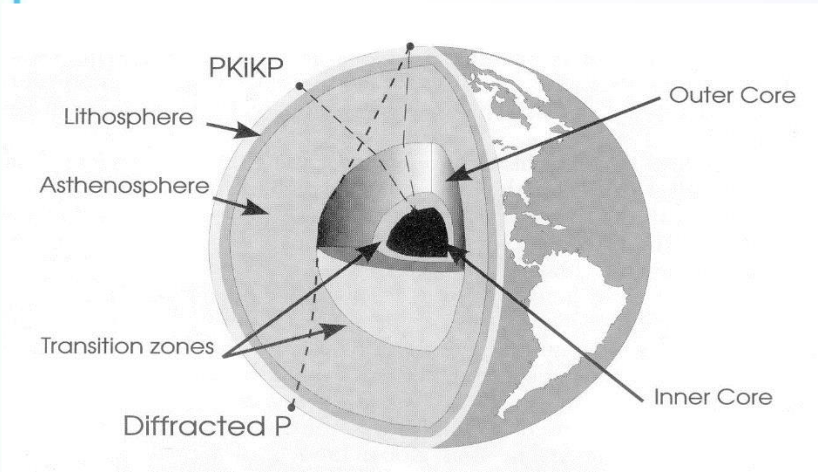


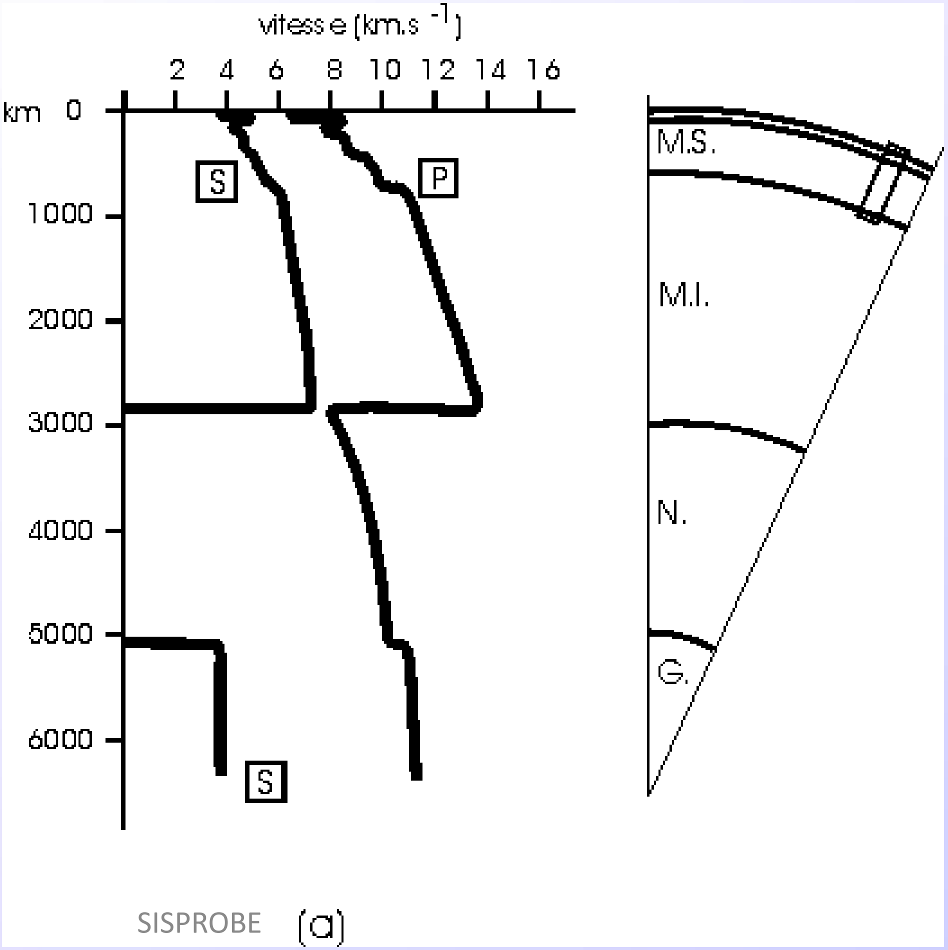
Please, describe the wave with signature

PKP, PKKP, SKKKS=S3KS

Depth velocity structure

- Velocity profile built without any prior





An difficulty arises when the velocity decreases.

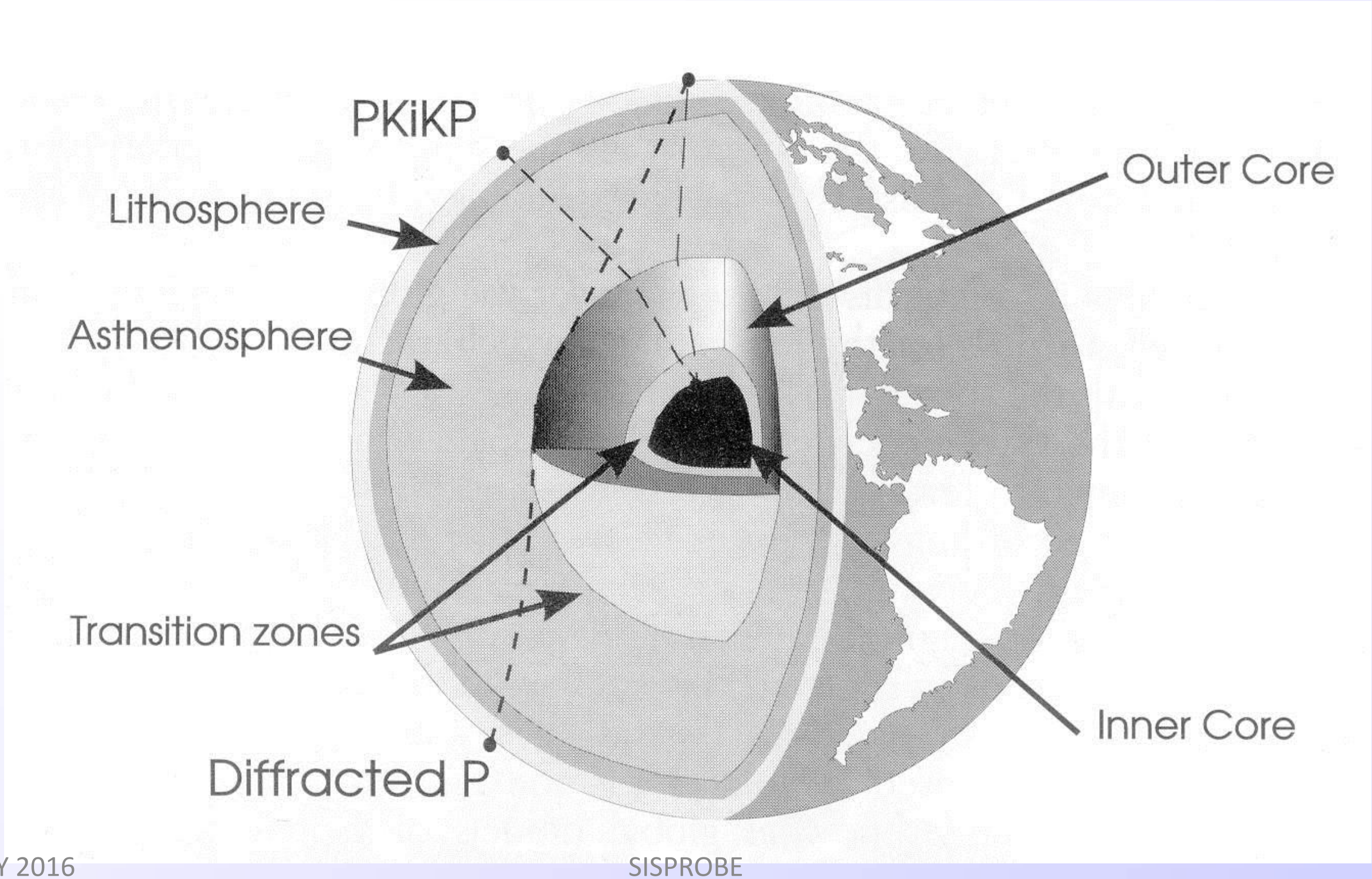
An initial model through the HWB method

- An initial model can be built (Orcut, 1980)
- The exact inverse formulae does not allow to introduce additional information,

  F. Press in 1968 has preferred the exhaustive exploration of possible profiles (5 millions !). The quality of the profile is appreciated using a misfit function as the sum of the square of delayed times as well as total volumic mass and inertial moments well constrained from celestial mechanics ...

- Exploration through grid search, Monte Carlo search, simulated annealing, genetic algorithm, tabou method, hant search and so on…

The symmetrical radial EARTH



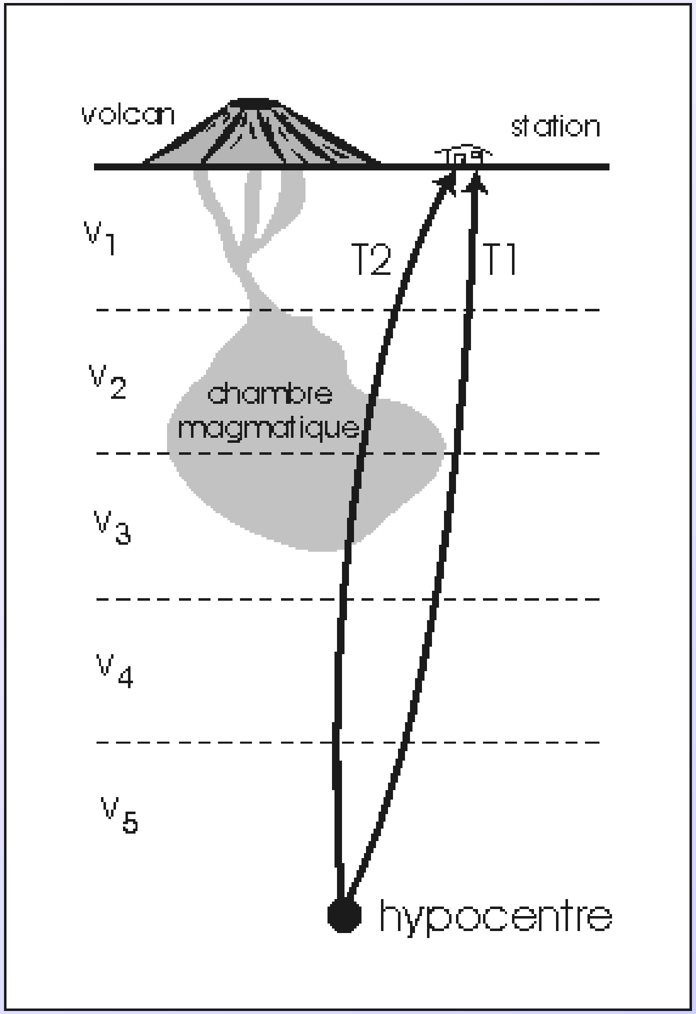
True time tomography

Only for one dimension …
uptonow

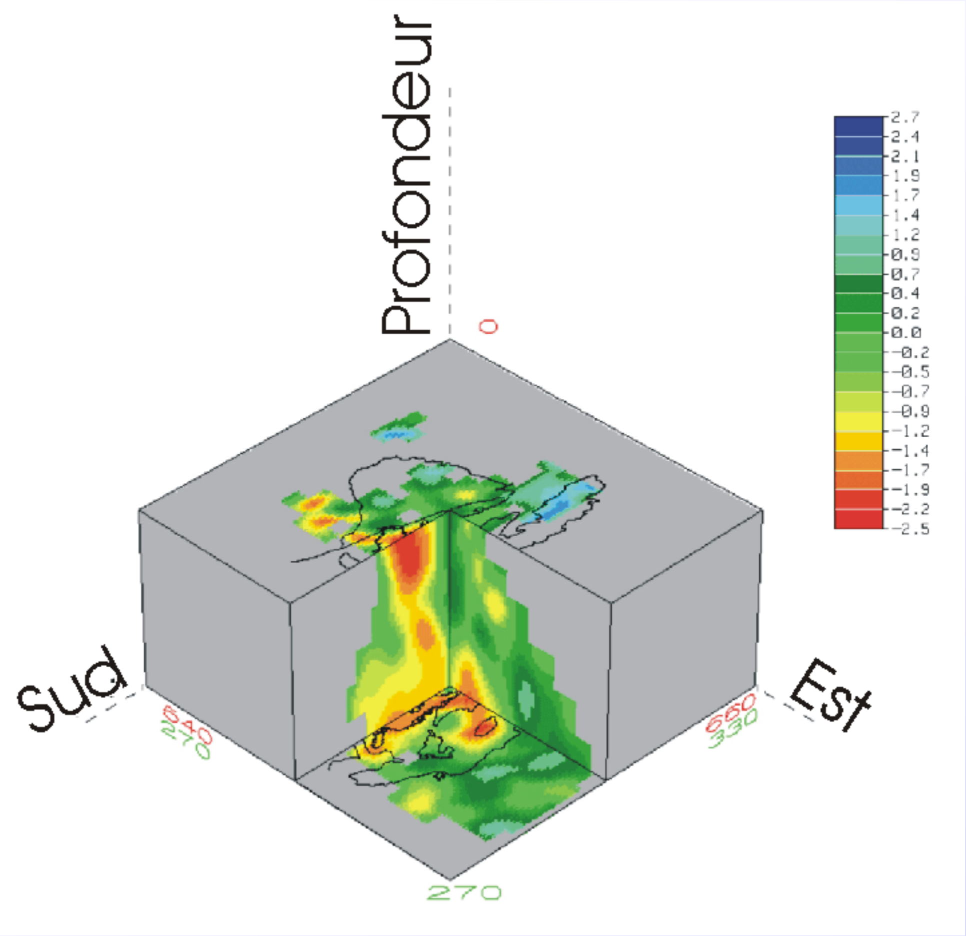
What can we do in 2D, 3D and 4D?

Simple strategy: small perturbation

- Initiale structure of velocity
- Search of small variation of velocity or slowness
- Linear approach



Example: Massif Central

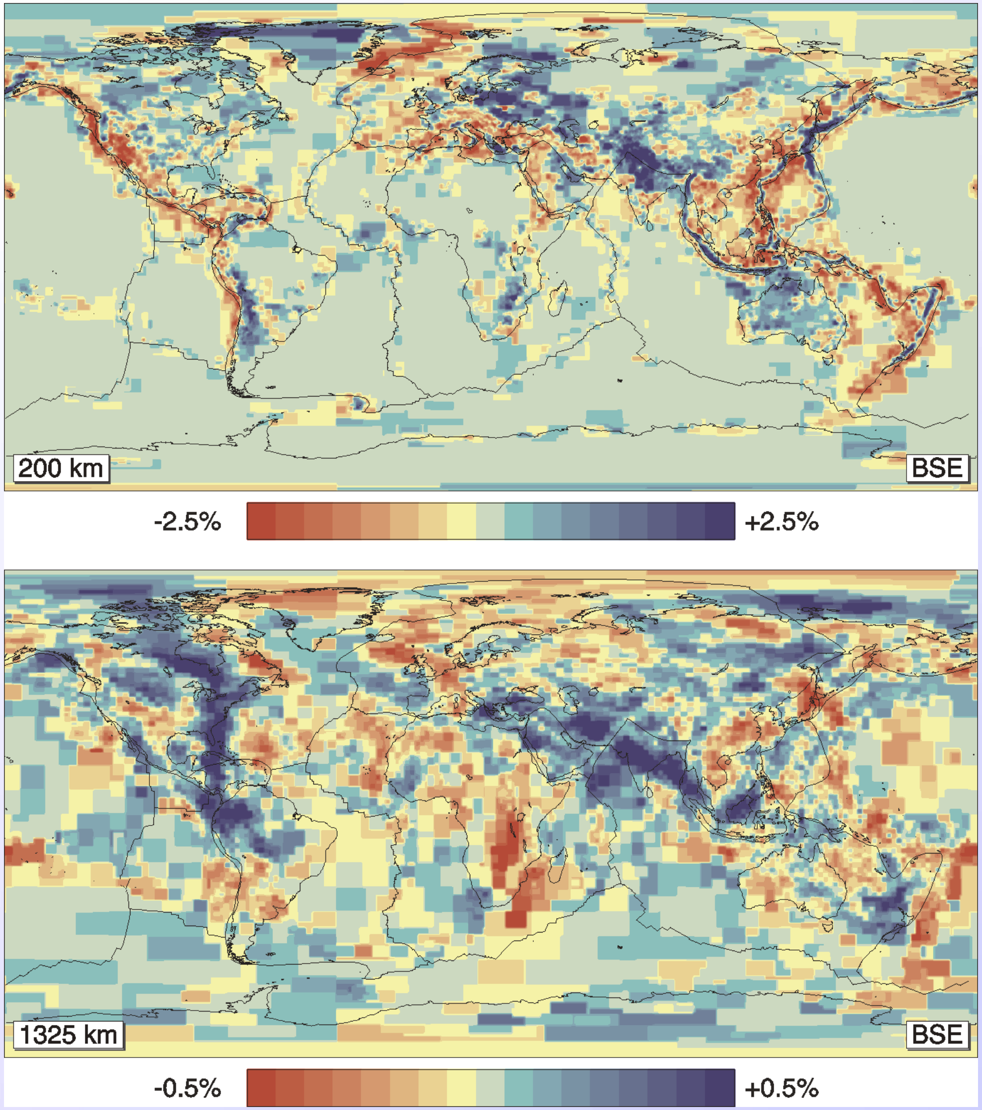


GLOBAL Tomography

- Velocity variation at a depth of 200 km : good correlation with superficial structures.

- Velocity variations at a depth of 1325 km : good correlation with the Geoid.

Courtesy of W. Spakman



Delayed Travel-time tomography

$$t(source, receiver) = \int_s^r u(x,y,z)dl$$

Finding the slowness u(x,y,z) from t(s,r) is a difficult problem: only techniques for one variable u(z) (Abel) !

Consider small perturbations $\delta u(x,y,z)$ of the slowness field $u_0(x,y,z)$

$$t(s,r) = \int_s^r u(x,y,z)dl = \int_s^r u_0(x,y,z)dl + \int_s^r \delta u(x,y,z)dl$$

$$t(s,r) = \int_{s_0}^{r_0} u_0(x,y,z)dl + \int_{s_0}^{r_0} \delta u(x,y,z)dl$$

« frozen » ray approximation (ray connecting source/receiver for the known slowness s_0)

$$t(s,r) - t_0(s,r) = \int_{s_0}^{r_0} \delta u(x,y,z)dl$$

$$\boxed{\delta t(s,r) = \int_{s_0}^{r_0} \delta u(x,y,z)dl}$$

LINEARIZED PROBLEM $\delta t(d) = J(u)\,\delta u$
from the model domain to the data domain

Discretization of the slowness perturbation

The velocity perturbation field (or the slowness field) $\delta u(x,y,z)$ can be described into a meshed cube regularly spaced in the three directions.

For each node, we specify a value $\delta u_{i,j,k}$. The interpolation will be performed with functions as step functions. For each grid point (i,j,k), shape functions $h_{i,j,k} = 1$ for i,j,k, and zero for other indices.

$$\delta u(x,y,z) = \sum_{cube} \delta u_{i,j,k}\, h_{i,j,k}$$

Nodal approach

Other shape functions are possible with two end members (nodal versus global):

fourier functions (cos,sin), chebychev, spline … and so on

Discrete Model Space

$$\delta t(\mathrm{s,r}) = \int_{ray_0} \sum_{cube} \delta u_{i,j,k}\, h_{i,j,k}\, dl = \sum_{cube} \delta u_{i,j,k} \boxed{\int_{ray_0} h_{i,j,k} dl}$$

Weighted ray segment

$$\delta t(s,r) = \sum_{i,j,k} \delta u_{i,j,k} \Delta l_{i,j,k} = \sum_{i,j,k} \frac{\partial t}{\partial s_{i,j,k}} \delta u_{i,j,k}$$

Discretization of the medium fats the ray

$$\delta t(s,r) = \sum_{i,j,k} J_{i,j,k}\ \delta u_{i,j,k}$$

$\delta t(d) = J(u)\ \delta u$

To be solved in least squares sense

Sensitivity matrix J is a sparse matrix

also named Fréchet derivative or Jacobian matrix or sensitivity matrix …

$$\begin{pmatrix} \delta t_1 \\ \delta t_2 \\ \vdots \\ \delta t_{n-1} \\ \delta t_n \end{pmatrix} = \begin{pmatrix} \frac{\partial t_1}{\partial u_1} & \cdots & \frac{\partial t_1}{\partial u_m} \\ \frac{\partial t_2}{\partial u_1} & \cdots & \frac{\partial t_2}{\partial u_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial t_{n-1}}{\partial u_1} & \cdots & \frac{\partial t_{n-1}}{\partial u_m} \\ \frac{\partial t_n}{\partial u_1} & \cdots & \frac{\partial t_n}{\partial u_m} \end{pmatrix} \begin{pmatrix} \delta u_1 \\ \delta u_2 \\ \vdots \\ \delta u_{m-1} \\ \delta u_m \end{pmatrix}$$

Joint hypo-velocity inversion

$$\delta t(src,rec) = \sum_{cube} \frac{\partial t}{\partial u} \delta u_{i,j,k} + \overrightarrow{p_s}.\overrightarrow{\delta x_s} + 1.\delta t_{0s}$$

$$\begin{bmatrix} \delta t_1 \\ \delta t_2 \\ \delta t_3 \\ \vdots \\ \delta t_{N-2} \\ \delta t_{N-1} \\ \delta t_N \end{bmatrix} = \begin{bmatrix} \partial t/\partial u & 0 & \partial t/\partial u & \cdots & p_{xs} & p_{ys} & p_{zs} & \cdots & 1 \\ 0 & \partial t/\partial u & 0 & \cdots & p_{ys} & p_{ys} & p_{zs} & \cdots & 1 \\ 0 & 0 & 0 & \cdots & p_{ys} & p_{ys} & p_{zs} & \cdots & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \partial t/\partial u & 0 & \partial t/\partial u & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \partial t/\partial u & \partial t/\partial u & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \partial t/\partial u & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \delta u_{1,1,1} \\ \delta u_{2,1,1} \\ \delta u_{3,1,1} \\ \vdots \\ \vdots \\ \delta x_s \\ \delta y_s \\ \delta z_s \\ \vdots \\ \delta t_{os} \\ \vdots \end{bmatrix} =$$

Draw on the Bboard

Linear algebra

$$\delta t = J\delta u$$
$$d = Gm$$
$$b = Ax$$

Least-squares problem

•The rectangular system can be recast into a square system (sometimes called normal equations).

• Solving this square linear system gives the so-called least-squares solution.

•Another interesting solution

Least-squares solution

$$J^t \delta t = J^t J \delta u$$
$$\delta u = (J^t J)^{-1} J^t \delta t$$

Remark

Least-norm solution

$\delta t = J J^t \delta y$ with $\delta u = J^t \delta y$

$$\delta u = J^t (J J^t)^{-1} \delta t$$

The system is both under-determined and over-determined depending on the considered zone (and the number of rays going through). This leads to a ray count around a node …

LINEAR INVERSE PROBLEM

At current medium m_0

Updating slowness perturbation values from time residuals

Formally one can write $\delta u = G_0^{-1} \delta t \rightarrow m = G_0^{-1} d$

with the forward problem $\delta t = G_0 \delta u \rightarrow d = G_0 m$

Existence, Uniqueness, Stability, Robustness

Discretisation

Identifiability

of the model

Small errors propagates

Outliers effects

NON-UNIQUENESS & NON-STABILITY : ILL-POSED PROBLEM

REGULARISATION : ILL-POSED -> WELL-POSED

G_0^{-1} *is the generalized inverse and it is difficult to obtain*

Weighted least-squares problem

$$E_k(m) = \frac{1}{2}\left((\delta t - G_k(m_k))^t(\delta t - G_k(m_k)\right) + \frac{1}{2}\varepsilon(m_k^t D m_k)$$

Tikhonov term with a Laplacian D

$$J^t \delta t = (J^t J + \varepsilon I)\,\delta u$$
$$\delta u = (J^t J + \varepsilon I)^{-1}\, J^t \delta t$$

The damping parameter ε provides stable inverse

Remark

Least-norm solution

Suppression of small eigenvalues by introducting a damping parameter

$\delta t = J J^t \delta y$ with $\delta u = J^t \delta y$

$$\delta u = J^t (J J^t + \epsilon I)^{-1} \delta t$$

More theory

For better understanding the influence of the damping parameter and the tikhonov term

Two guys: Resolution matrix & Importance matrix

LEAST SQUARES METHOD

L_2 norm

$$E(m) = (d - G_0 m)^t (d - G_0 m)$$

locates the minimum of E

$$\frac{\partial E(m)}{\partial m} = 0$$

G_0 is a N by M matrice

normal equations

$$G_0^t G_0 m = G_0^t d$$

$\left[G_0^t G_0\right]^{-1}$ is a M by M matrice

$$m_{est} = \left[G_0^t G_0\right]^{-1} G_0^t d \quad \text{if } \left[G_0^t G_0\right]^{-1} \text{exists}$$

Least-squares estimation

Operator $\left[G_0^t G_0\right]^{-1} G_0^t$ on data will derive a new model : this is called the generalized inverse G_0^g

Under-determination M > N

Over-determination N > M

Mixed-determination – seismic tomography

SVD analysis for stability and uniqueness

SVD decomposition : $$G_0 = U \Lambda V^t$$

U : (N x N) orthogonal $U^t=U^{-1}$

V : (M x M) orthogonal $V^t=V^{-1}$

Λ : (N x M) diagonal matrice Null space for $\Lambda_i=0$

$U^tU=I$ and $V^tV=I$ (not the inverse !)

$$V = [V_p | V_0]$$

$$U = [U_p | U_0]$$

$$G_0 = \left[U_p | U_0\right] \begin{bmatrix} \Lambda_p & 0 \\ 0 & 0 \end{bmatrix} \left[V_p | V_0\right]^t$$

V_p and V_0 determine the uniqueness while U_p and U_0 determine the existence of the solution

$$G_0 = U_p \Lambda_p V_p^t$$

$$G_0^{-1} = V_p \Lambda_p^{-1} U_p^t$$

Up and Vp have now inverses !

Solution, model & data resolution

The solution is

$$m_{est} = G_0^{-1} d = G_0^{-1}(G_0 m) = V_p \Lambda_p^{-1} U_p^t U_p \Lambda_p V_p^t m = \left[V V_p^t \right] m = Rm$$

where $R = \left[V_p V_p^t \right]$ Model resolution matrice : if $V_0=0$ then $R=VV^t=I$

$$d_{est} = G_0 m_{est} = \left[U_p U_p^t \right] d = Nd$$

where $N = \left[U_p U_p^t \right] d$ Data resolution matrice → if $U_0=0$ then $N=UU^t=I$

importance matrice

Goodness of resolution

SPREAD(R)= $\left\| R - I \right\|^2$

Good tools for quality estimation

SPREAD(N)= $\left\| N - I \right\|^2$

Spreading functions

Maximum Likelihood method

One assume a gaussian distribution of data

Data distribution could be written $$p(d) \propto exp\left[-\frac{1}{2}(d - G_0\, m)^t {C_d}^{-1}(d - G_0 m)\right]$$

where $G_0 m_{est}$ is the data mean and C_d is the data covariance matrice.

This method is very similar to the least squares method where we define the following misfit function E_1.

$$E(m) = (d - G_0\, m)^t (d - G_0 m) \rightarrow E_1(m) = (d - G_0\, m)^t {C_d}^{-1}(d - G_0 m)$$

Even without knowing the matrice C_d, we may consider data weight W_d through the misfit function $E_2(m) = (d - G_0 m)^t W_d (d - G_0 m)$

The inferred model distribution will be gaussian with the main difficulty of estimating the posterior model covariance C_m connected with the curvature of the misfit function E_2.

PRIOR INFORMATION

Hard bounds $A < m_i < B$ Seismic velocity should be positive

Prior model

$$E_3(m) = (d - G_0 m)^t (d - G_0 m) + \varepsilon (m - m_p)^t (m - m_p)$$

The parameter ε is the damping parameter controlling the importance of the model m_p

Gaussian distribution $E_4(m) = (d - G_0 m)^t C_d^{-1} (d - G_0 m) + (m - m_p)^t C_m^{-1} (m - m_p)$

$$G_0^g = \left[G_0^t C_d^{-1} G_0 + C_m^{-1} \right]^{-1} G_0^t$$

Model smoothness $E_5(m) = (d - G_0 m)^t W_d (d - G_0 m) + (m - m_p)^t W_m (m - m_p)$

with W_d data weighting and W_m model weighting

Penalty approach

add additional relations between model parameters (new lines)

Weighted joint hypo-velocity inversion

Penalty strategy: add lines to the linear system to be solved

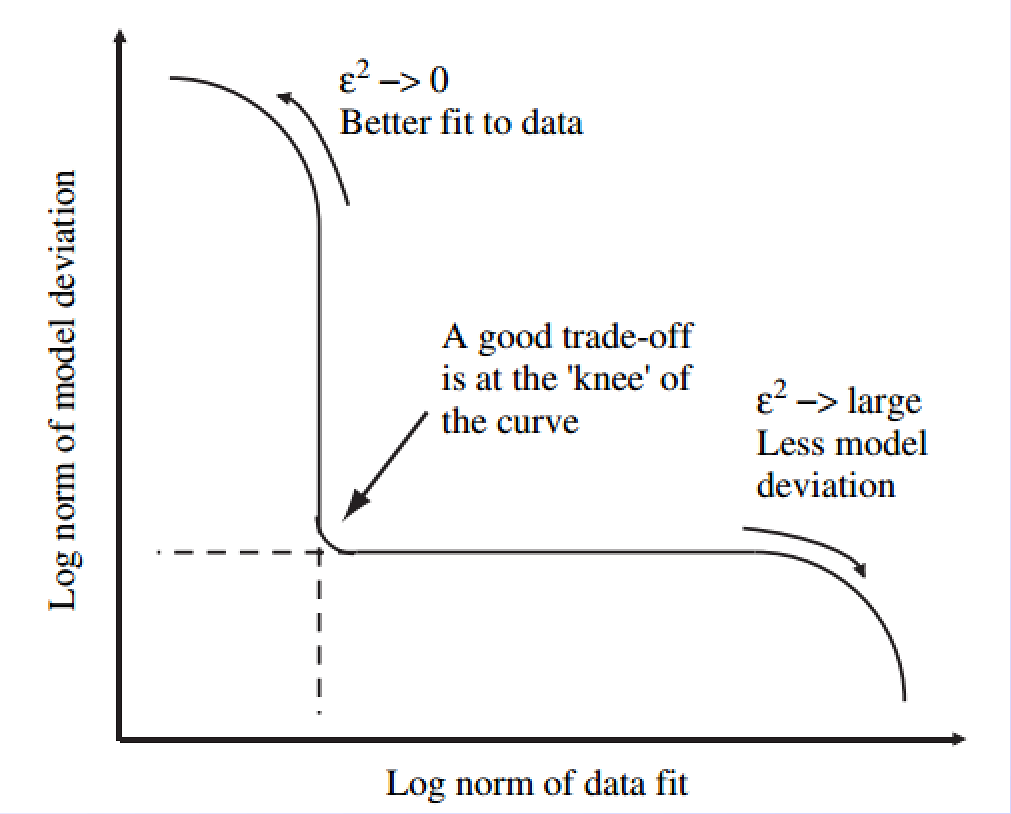
$$\begin{bmatrix} \delta t_1 \\ \delta t_2 \\ \delta t_3 \\ \vdots \\ \delta t_{N-2} \\ \delta t_{N-1} \\ \delta t_N \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \partial t/\partial u & 0 & \partial t/\partial u & \cdots & p_{xs} & p_{ys} & p_{zs} & \dots & 1 \\ 0 & \partial t/\partial u & 0 & \dots & p_{ys} & p_{ys} & p_{zs} & \cdots & 1 \\ 0 & 0 & 0 & \cdots & p_{ys} & p_{ys} & p_{zs} & \cdots & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \partial t/\partial u & 0 & \partial t/\partial u & \cdots & 0 & 0 & 0 & \dots & 0 \\ 0 & \partial t/\partial u & \partial t/\partial u & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \partial t/\partial u & \cdots & 0 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \delta u_{1,1,1} \\ \delta u_{2,1,1} \\ \delta u_{3,1,1} \\ \vdots \\ \vdots \\ \delta x_s \\ \delta y_s \\ \delta z_s \\ \vdots \\ \delta t_{os} \end{bmatrix} =$$

L curve

Estimating the parameter ε is always difficult

We can test different values which provide a L curve

If a knee appears, this could be an adequat value … although this knee is often difficult to identify in seismic tomography.



Other techniques are available as LASSO approach but they are intensive and sensitive to noise.

UNCERTAINTY ESTIMATION

Least squares solution

$$m_{est} = \left[G_0^t G_0\right]^{-1} G_0^t d = G_0^g d$$

$$\left[\mathrm{cov}\ m_{est}\right] = G_0^g \left[\mathrm{cov}\ d\right] G_0^{gt} = G_0^g C_d G_0^{gt}$$

$C_d = \sigma_d^2 I$ Uncorrelated data

$$\left[\mathrm{cov}\ m_{est}\right] = \sigma_d^2 \left[G_0^t G_0\right]^{-1}$$

$$\left[\mathrm{cov}\ m_{est}\right] = \sigma_d^2 \left[\frac{1}{2}\frac{\partial^2 E}{\partial m^2}\right]^{-1}_{m=m_{est}}$$

Model covariance : uncertainty in the data

curvature of the error function

Sampling the misfit function around the estimated model: often done numerically

A posteriori model covariance matrice

If one can decompose this matrice

$$G_0^t C_d^{-1} G_0 + C_m^{-1} = USU^t$$

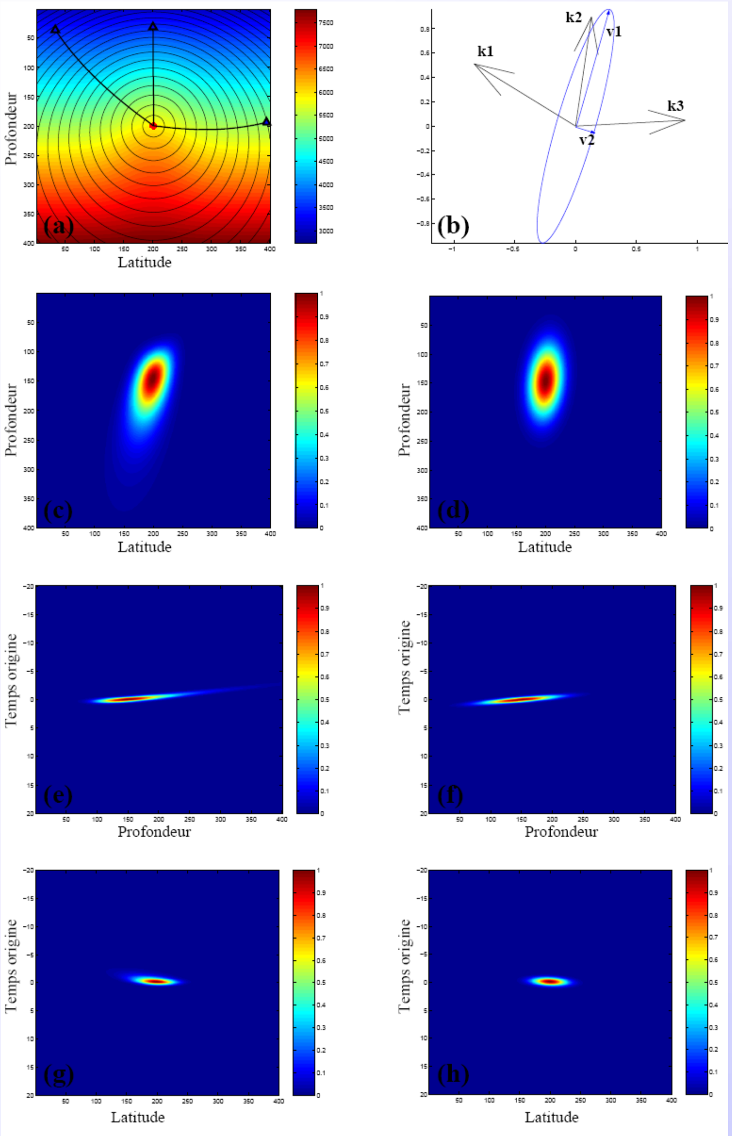
S diagonal matrice eigenvalues

U orthogonal matrice eigenvectors

Error ellipsoidal could be estimated

WARNING : formal estimation related to the gaussian distribution hypothesis

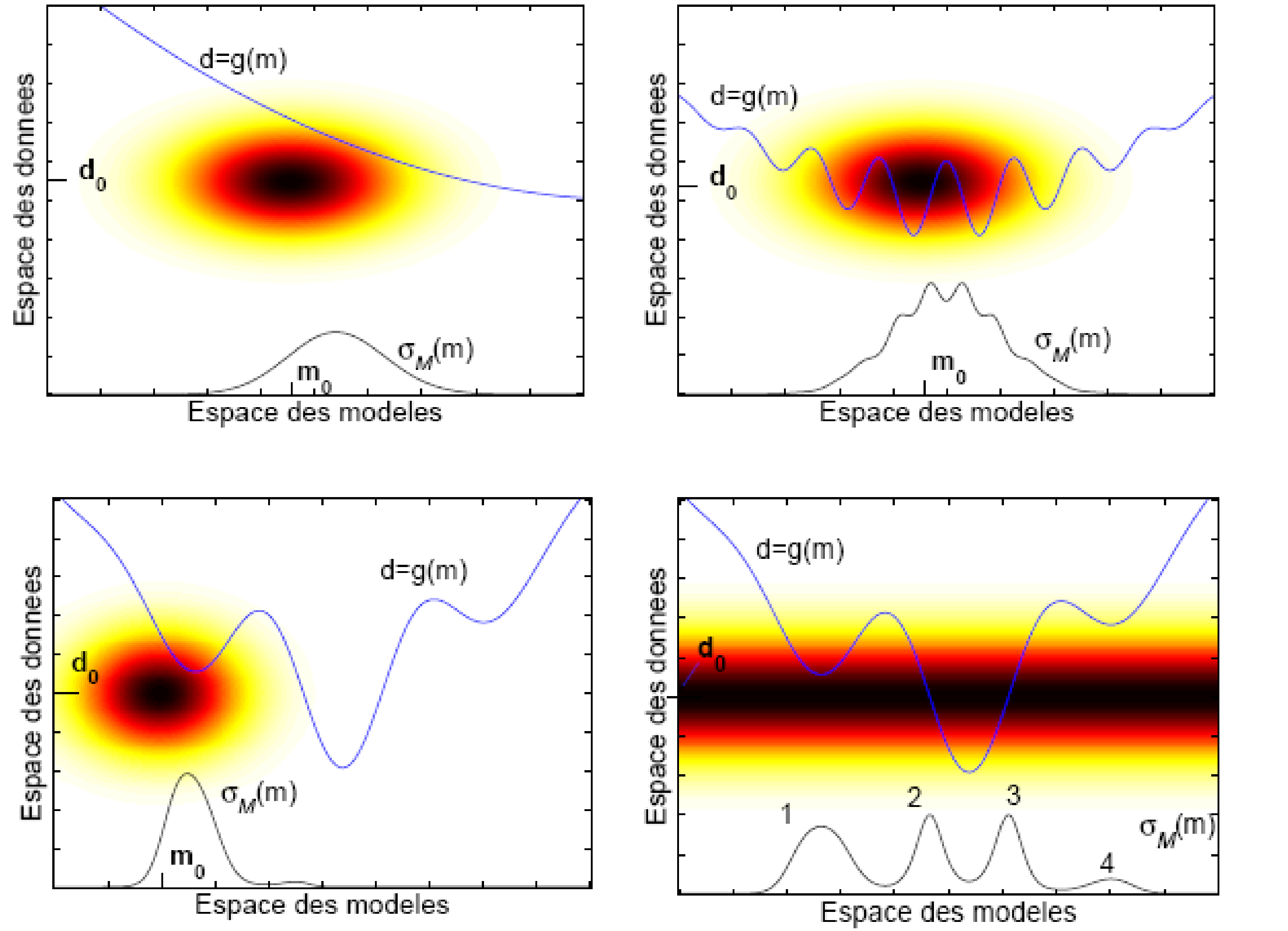
True a posteriori distribution



Tangent gaussian distribution

Prior & posterior information

What is the meaning of the « final » model we provide ?



From Monteiller (2005)

acceptable

Local Newton methods

$$E(m_{k+1}) < E(m_k)$$

Steepest Gradient method

$$E(m_k - t\nabla E(m_k)) < E(m_k)$$

$$d = -\nabla E(m_k)$$

Conjugate gradient

$$d_k = \begin{Bmatrix} -\nabla E_0 \\ -\nabla E_k + \beta_k d_{k-1} \end{Bmatrix}$$

Newton

$$d = -\nabla^2 E(m_k)^{-1} \nabla E(m_k)$$

quadratic approximation of E

Quasi-Newton

$$\nabla^2 E(m_k) \approx D_k$$

Gauss-Newton is Quasi-Newton for L^2 norm

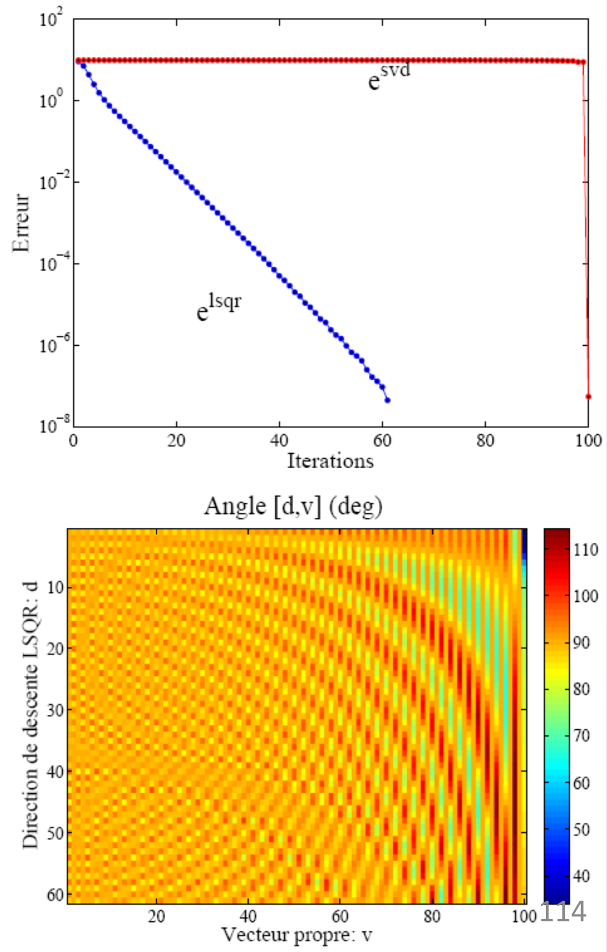
LSQR method

The LSQR method is a conjugate gradient method developped by Paige & Saunders (1982)

Good numerical behaviour for ill-conditioned matrices

When compared to an SVD exact solution, LSQR gives main components of the solution while SVD requires the entire set of eigenvectors

Fast convergence and minimal norm solution (zero components in the null space if any)



Linearized Inverse Problem

Misfit function $\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$

- First loop over models : for current model s_k (iteration k)

Two-loops procedure

  - Forward problem
    A. Solve eikonal equation
    B. Compute rays and synthetic travel-times
  - Build Fréchet derivative matrix J_k and delayed times δt_k

- Second loop over linear system: iteratively solve $J_k \delta u_k = \delta t_k$, ie $J_k{}^t J_k \delta u_k = \delta t_k$ using conjugate gradient (LSQR, for example)

Update the model $u_{k+1} = u_k + \delta u_k$

No explicit evaluation of the Hessian $J_k{}^t J_k$:
only products « $J\ \delta u$ » and « $J^t \delta t$ » are required in LSQR algorithm
Paige & Sanders (1982)

Flow chart

d_{obs} collected data

m starting model

loop

$$\boxed{d_{syn} = g(m)}$$

true ray tracing

$$\Delta d = d_{obs} - d_{syn}$$

data residual

$$G_0 = \partial g / \partial m$$

sensitivity matrice

$$\Delta m = G_0^{-1} \Delta d$$

model update

$$m = m + \Delta m$$

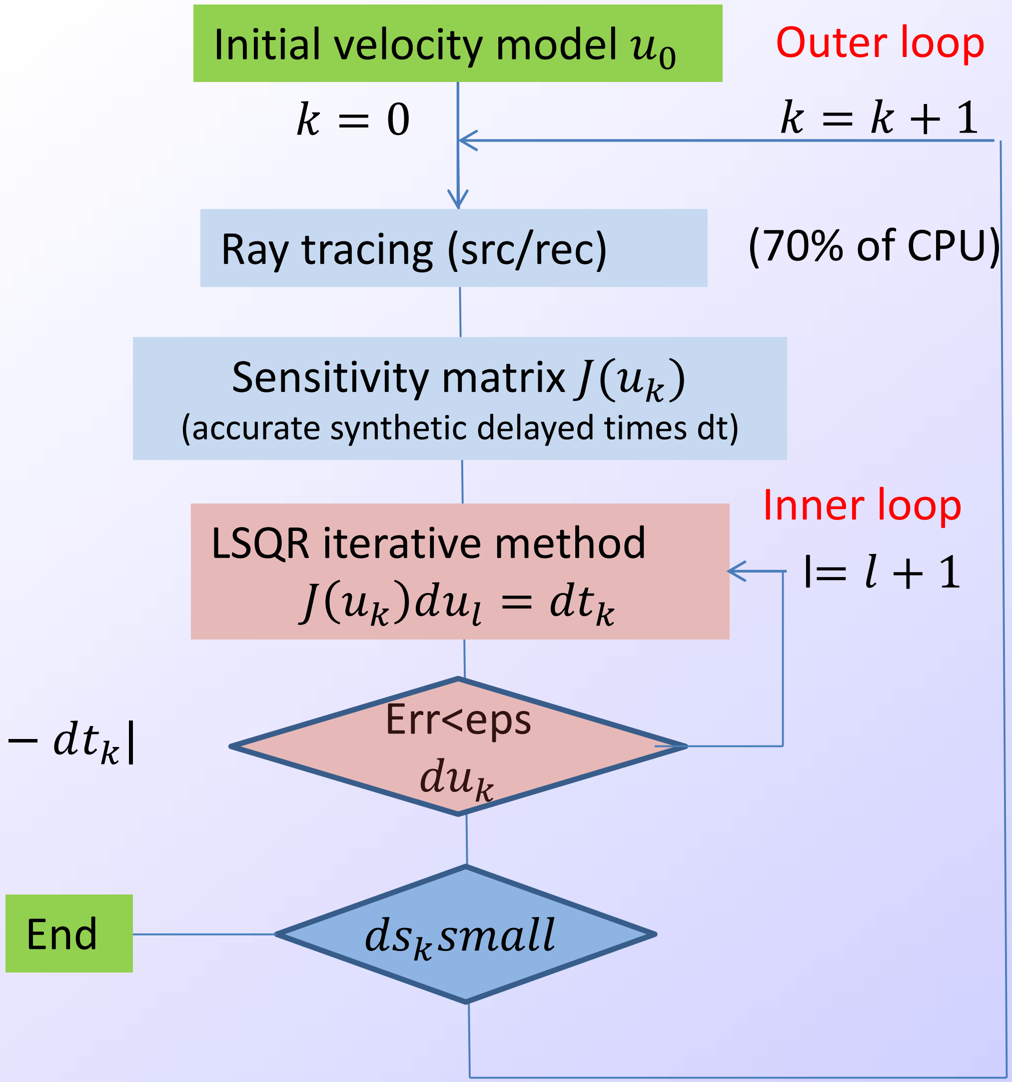
new model

small model variation or small errors exit

Calculate $\partial^2 E / \partial m^2$ for formal uncertainty estimation

Linearized Inverse Problem

Two-loops procedure



$\mathrm{Err} = |J(u_k)du_l - dt_k|$

The slowness field is denoted by s and it is often the one we reconstruct

Linearized Inverse Problem

$$\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$$

- First loop over models : for current model s_k (iteration k)

Two-loops procedure

  - Forward problem
    A. Solve eikonal equation
    B. Compute rays and synthetic travel-times

  - Build Fréchet derivative matrix $J_k$and delayed times δt_k

- Second loop over linear system: iteratively solve $J_k \delta u_k = \delta t_k$, ie $J_k{}^t J_k \delta u_k = \delta t_k$ using conjugate gradient (LSQR, for example)

Update the model $u_{k+1} = u_k + \delta u_k$

Complexity: $\mathcal{O}(N_{src} * N_{rec})$ for forward/adjoint formulation and

$\mathcal{O}(N_m * N_{src} * N_{rec} * k)$ for gradient storage (k: sparsity)

Linearized Inverse Problem

$$\mathbb{C}(s) = \frac{1}{2} dt^t dt$$

Two-loops procedure

Complexity: $\mathcal{O}(N_{src} * m * \ln(m))$ for eikonal formulation

$\mathcal{O}(k * m * N_{src} * N_{rec})$ for back-ray tracing

$\mathcal{O}(k * m * N_{src} * N_{rec})$ for sensitivity matrix storage J

The sparsity is expressed through k with only few points of the grid involved for each data

Linearized Inverse Problem

$$\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$$

One-loop procedure

- First loop over models : for current model s_k (iteration k)
  - Forward problem
    A. Solve eikonal equation
    B. Compute synthetic travel-times and delayed times δt_k
  - Adjoint problem
    A. Solve adjoint equation
    B. Build gradient γ_k
- Descent method (steepest gradient or conjugate gradient or l-BFGS) δs_k

Update the model $u_{k+1} = u_k + \alpha_k \delta u_k$
(Sei & Symes, 1994; Leung & Qian, 2006; Taillandier et al, 2009)

No explicit evaluation of the Hessian
l-BFGS provides an evaluation of the H^{-1} from previous stored gradients

Sampling model posterior distribution

Resolution estimation : spike test

Boot-Strapping

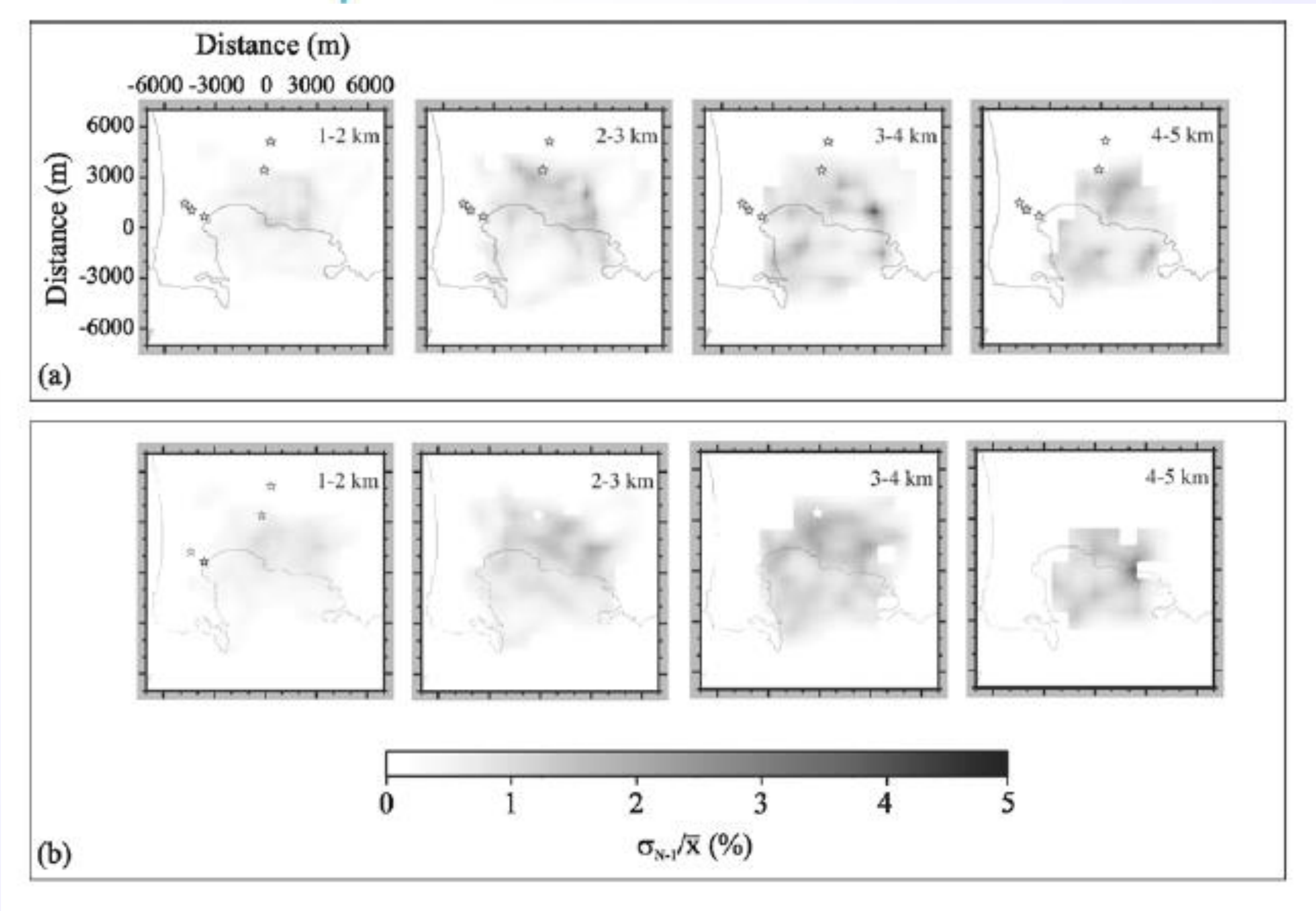
Jack-kniffing

Natural Neighboring

Monte-Carlo sampling

Checker-board test

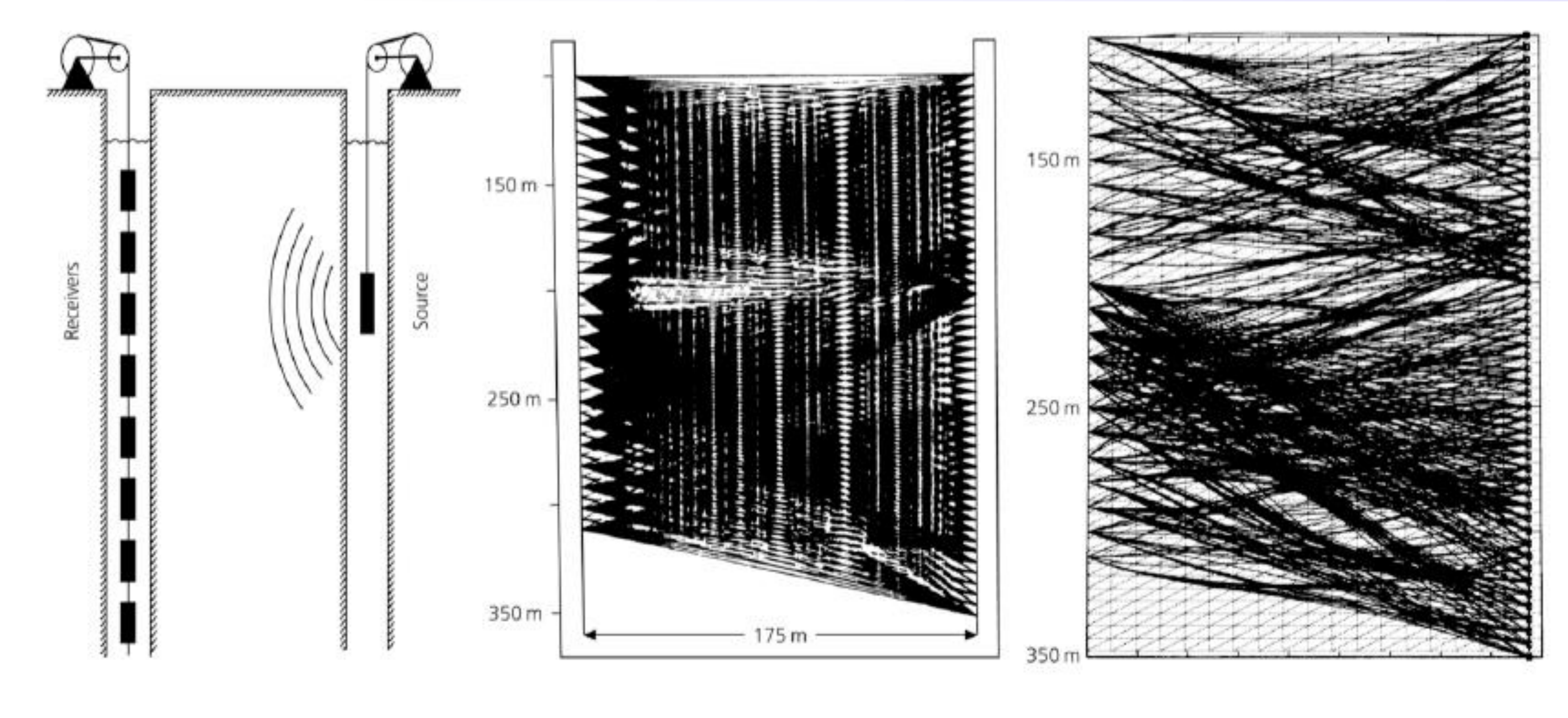
Sampling posterior distribution



Uncertainty estimation for P and S velocities using boot-strapping techniques

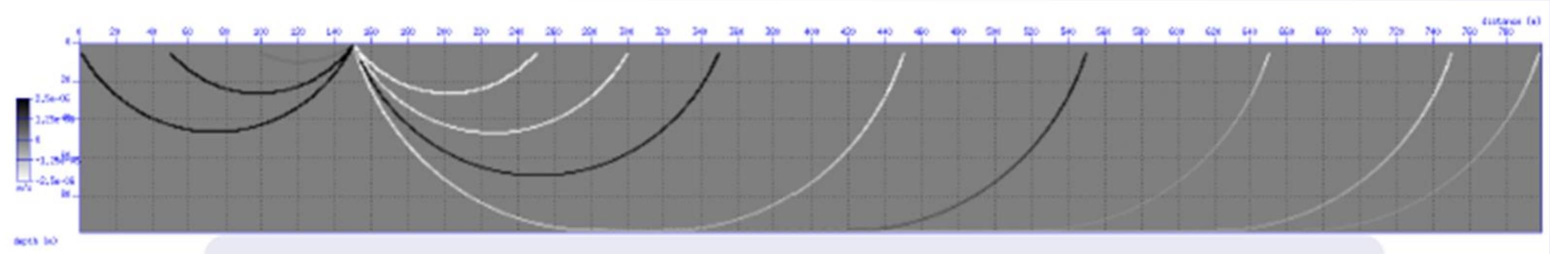
Coverage influence

Perturbation is spread along rays, coverage is important to constrain the inversion

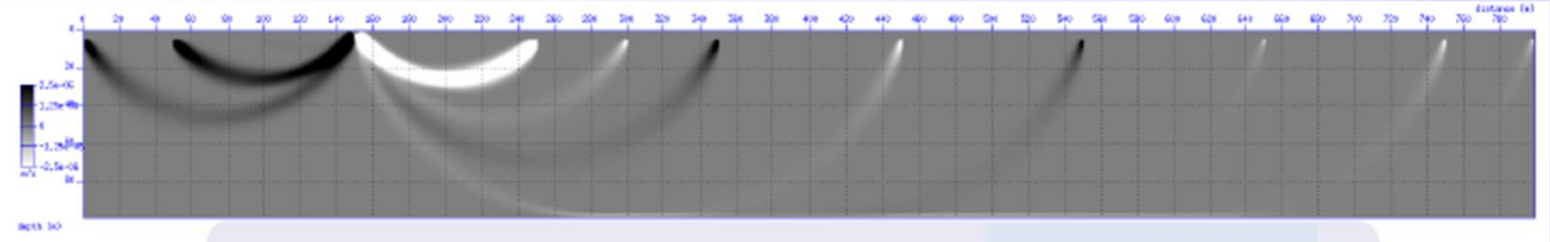


Ray approximation

$$\delta T(s,r) \approx \int u(x(l))dl \approx \iiint u(x)\delta\big(x - x(l)\big)K(s,r,x)dv$$



Line



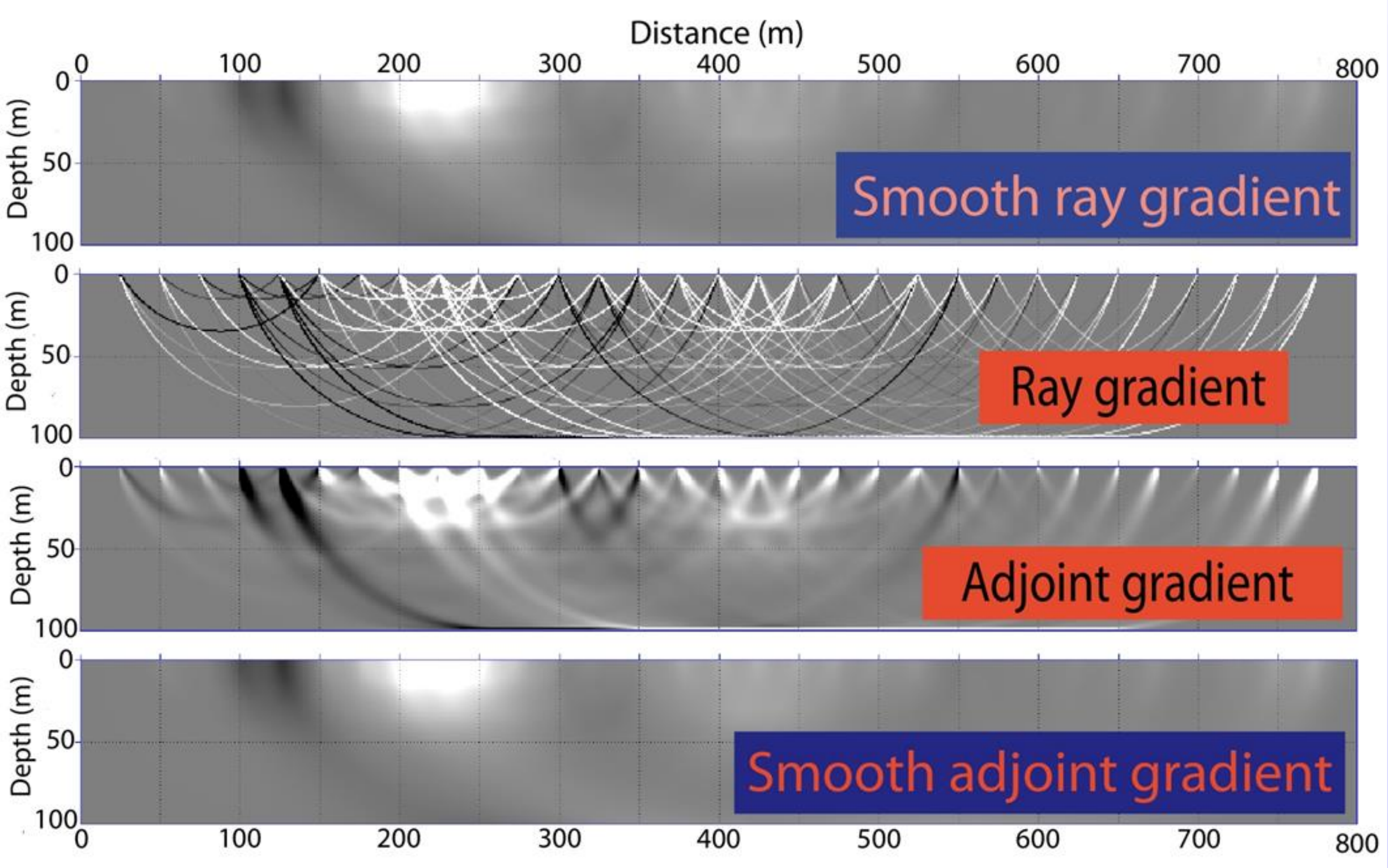
Volume

$K(x,y,z)$ is the sensitivity kernel

Ray approximation reasonnable approximation

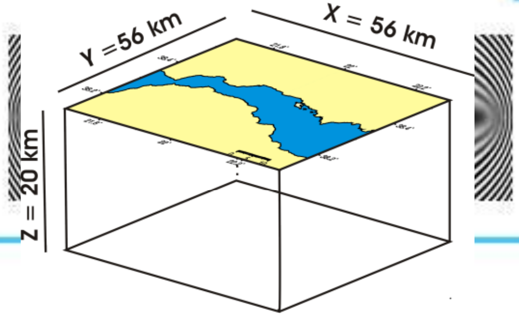
Gradient

$$J^t \delta t$$



Same gradient once estimated on the model discretization grid

Applications

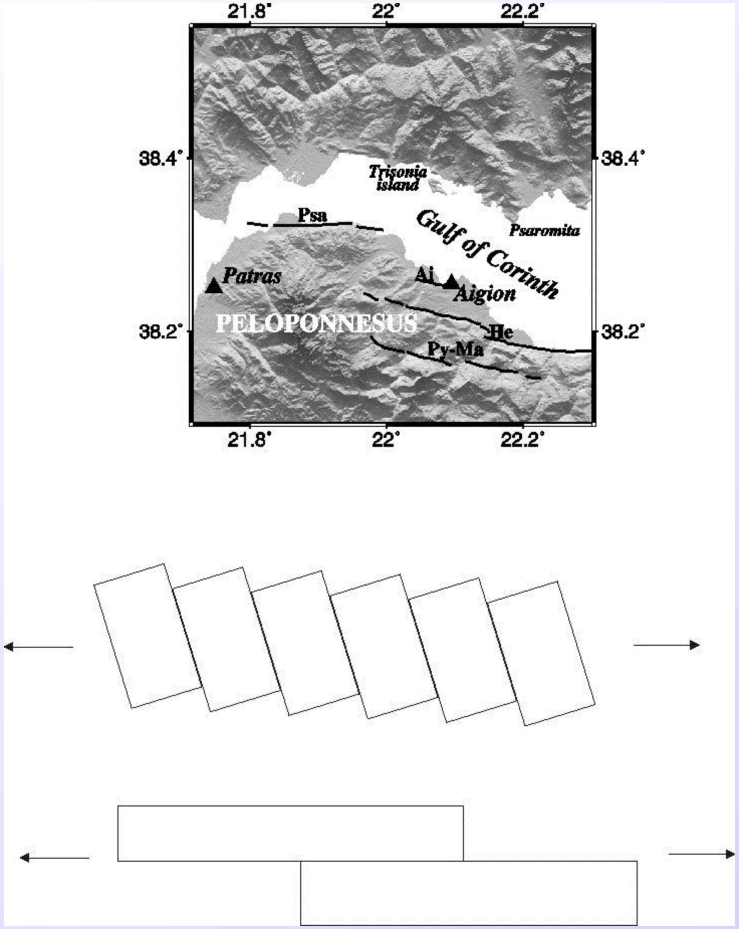


Corinth Gulf

An extension zone where there is a deep drilling project.

How this rift is opening?

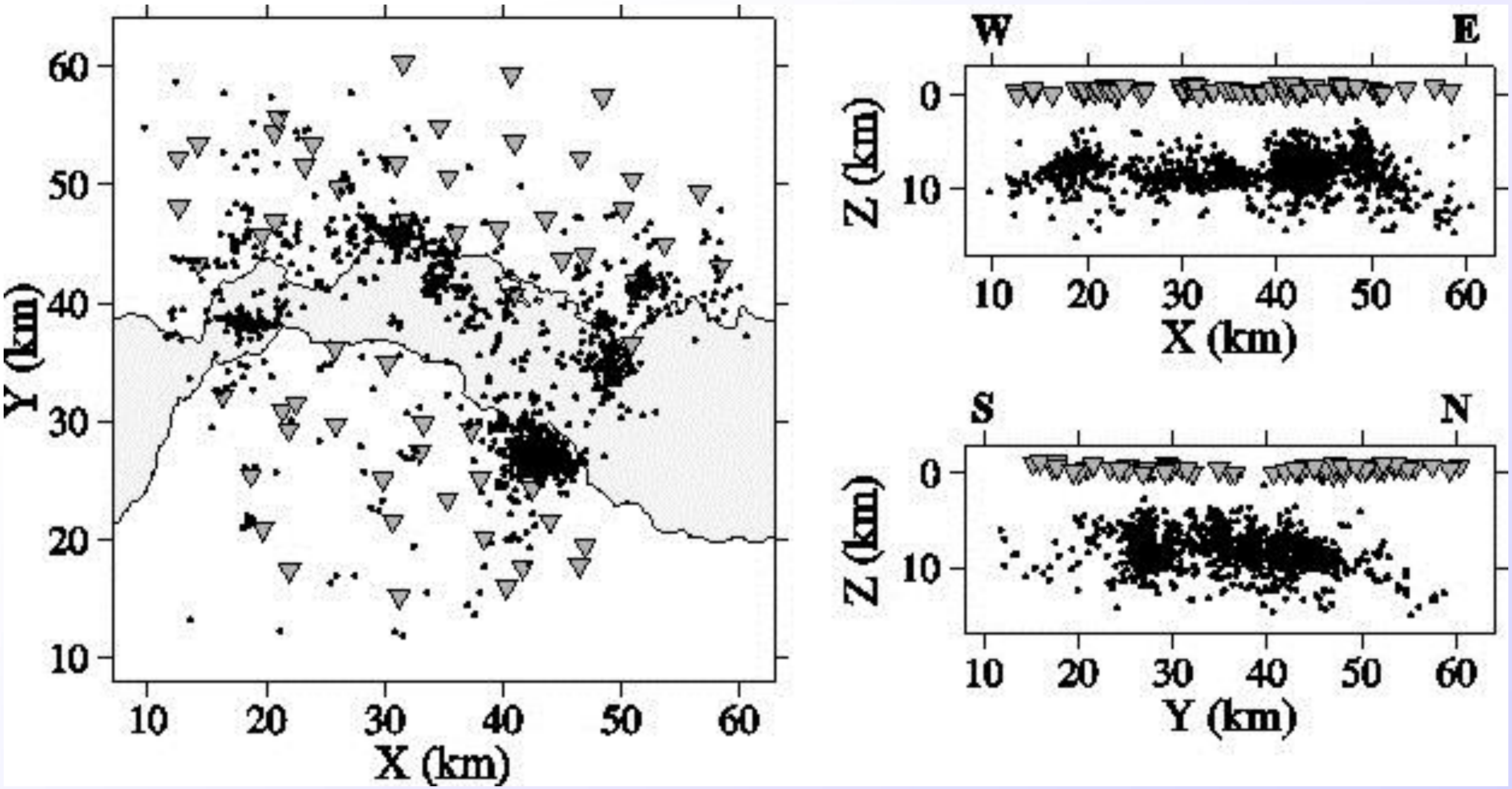
What are the physical mechanisms of extension (fractures, fluides, isostatic equilibrium)



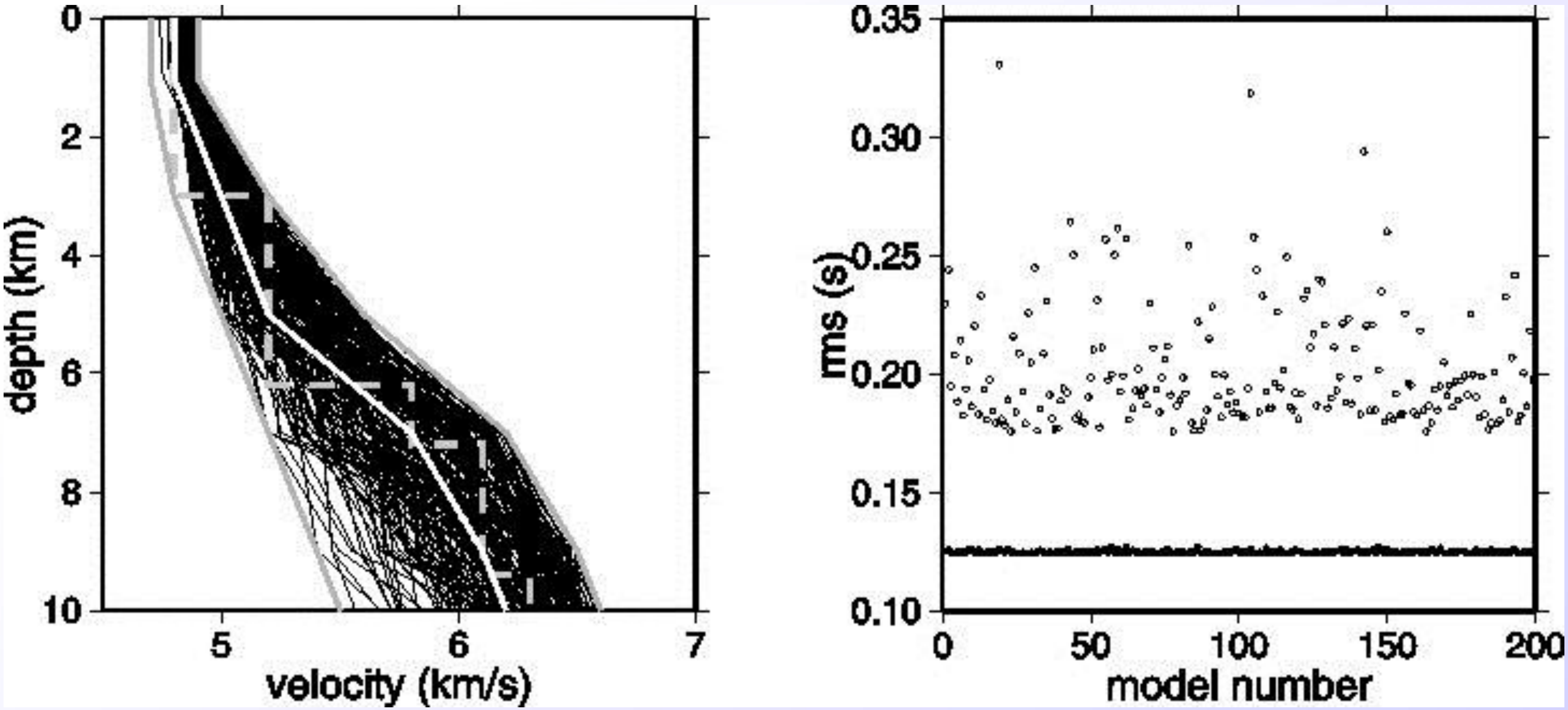
Work of Diana Latorre

and of Vadim Monteiller

Seismic experiment 1991 (and one in 2001)



MEDIUM 1D : HWB AND RANDOM SELECTION

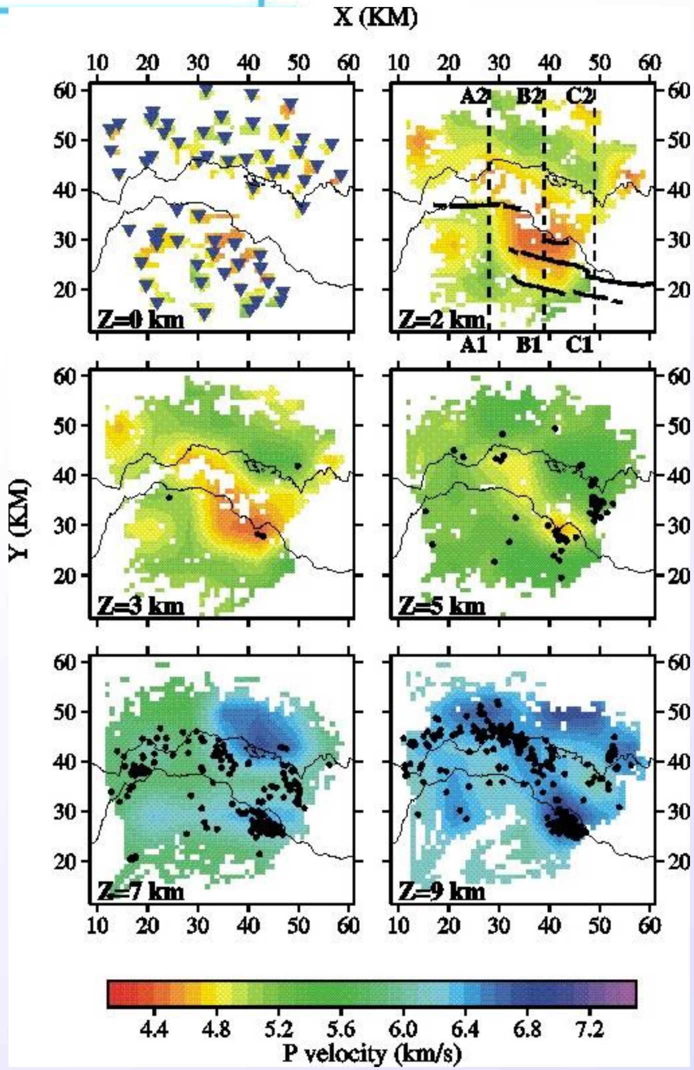


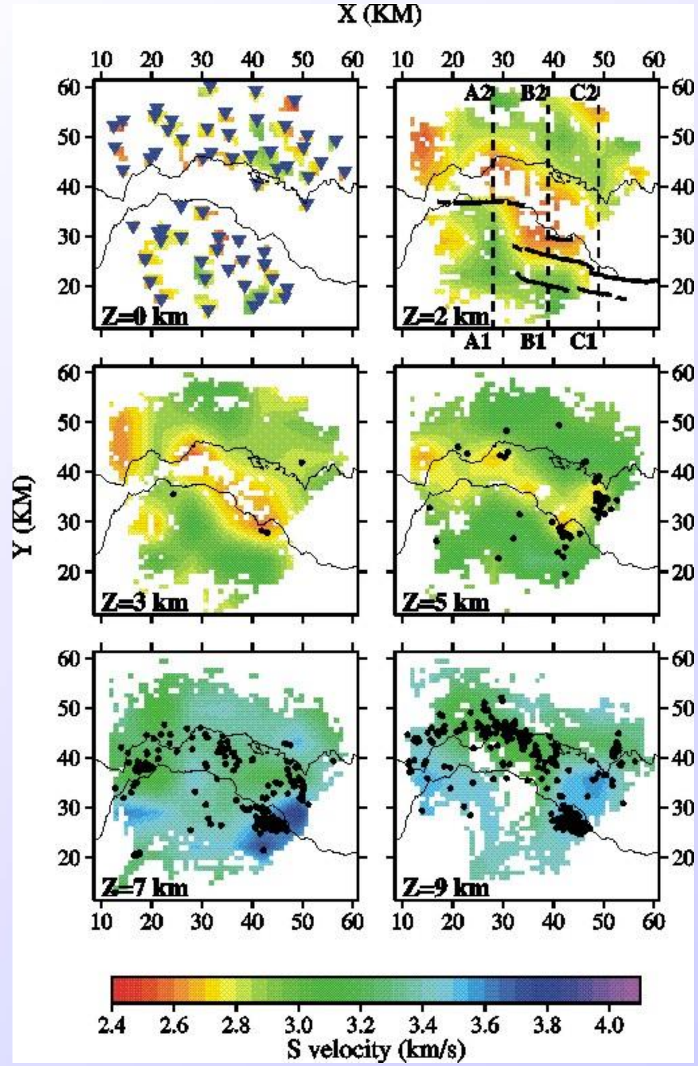
Velocity structure image

Horizontal sections

Basin structure

Fast variation

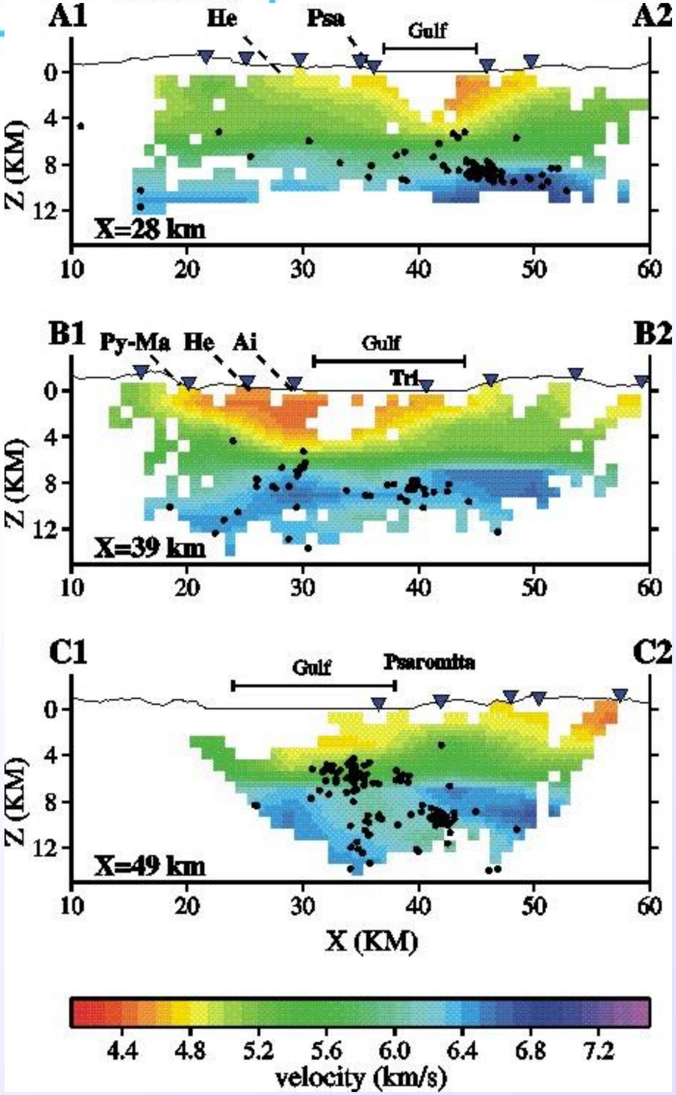




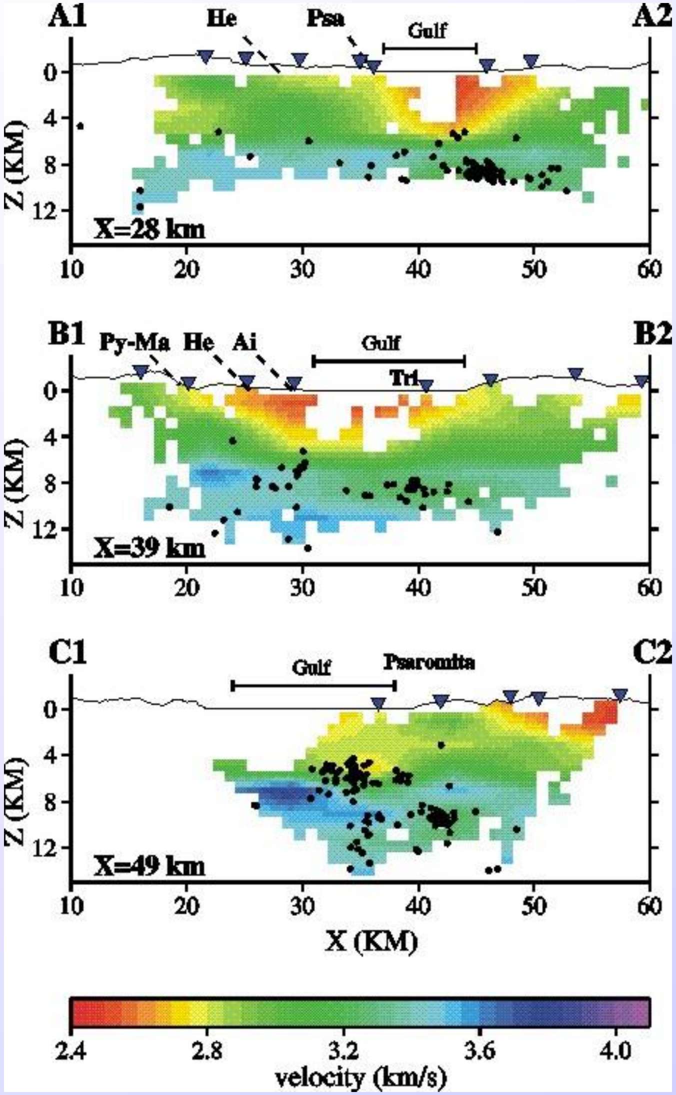
Velocity structure image

Vertical sections

P



S

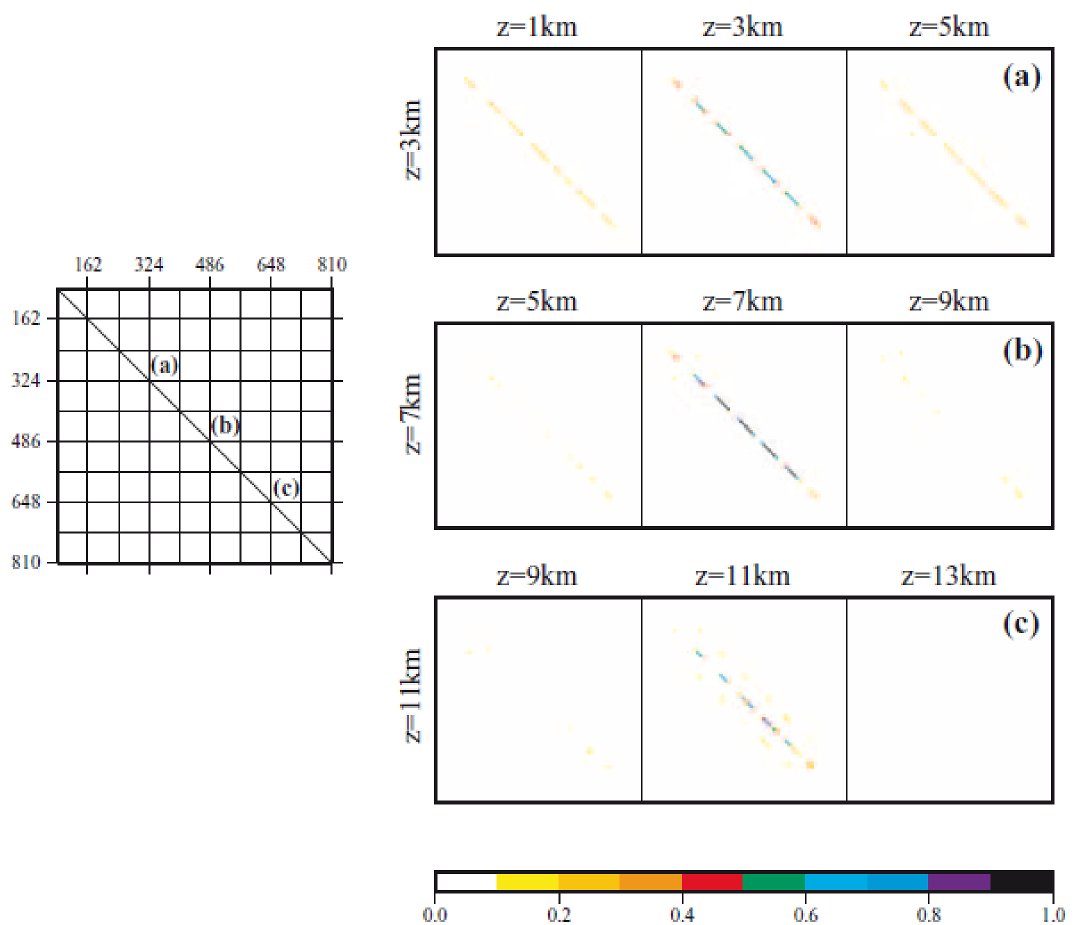


Resolution matrix

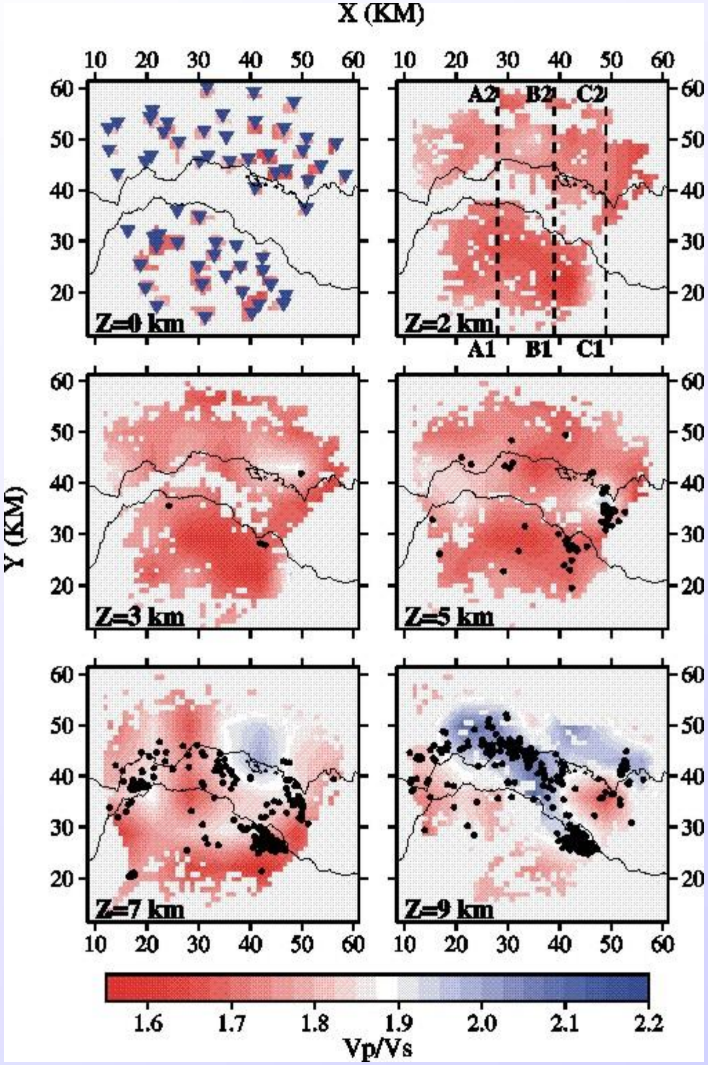
Workflow:

On the expected final solution, perform a small spike perturbation and invert for the synthetic new dataset

One line of the resolution matrix



Vp/Vs ratio:
fluid existence ?

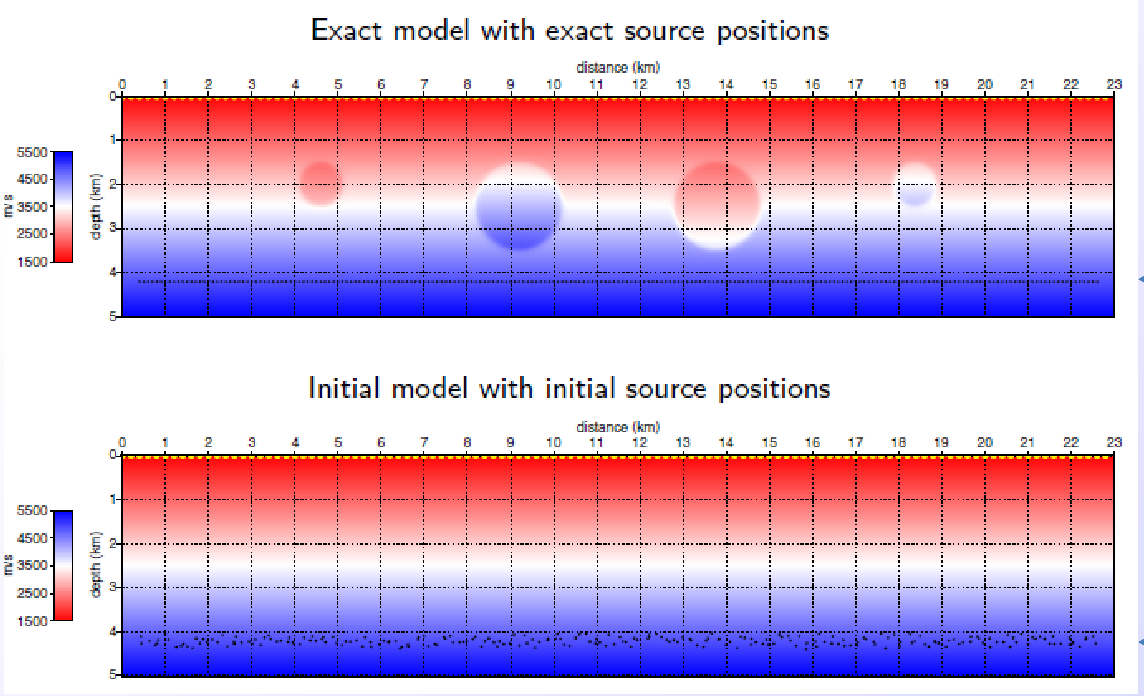


Recovered parameters might have diferent interpretation and the ratio Vp/Vs has a strong relation with the presence of fluids or the relation Vp*Vs may be related to porosity

Joint hypocenter/velocity

Strong trade-off between velocities and hypocenter parameters

- Calibration of weights between these parameters (Spakman & Nolet, 1988; Le Meur et al., 1997)
- Annealing operator (Pavlis and Booker, 1980; Julian and Gubbins, 1980)



Events at 4.2 km

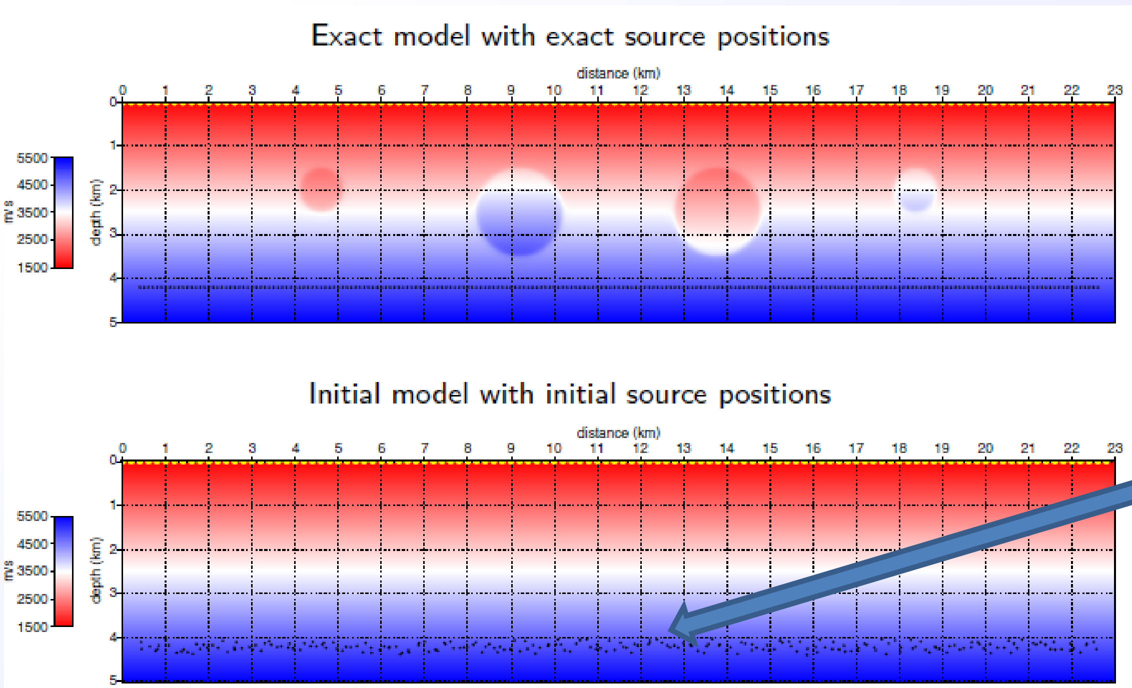
Toy example

Random initial hypocenters

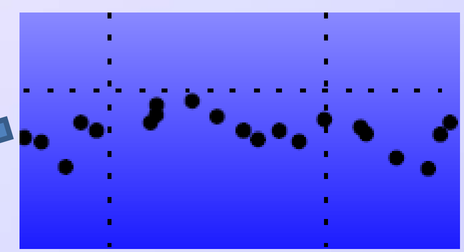
Joint hypocenter/velocity

Strong trade-off between velocities and hypocenter parameters

- Calibration of weights between these parameters (Spakman & Nolet, 1988; Le Meur et al., 1997)
- Annealing operator (Pavlis and Booker, 1980; Julian and Gubbins, 1980)



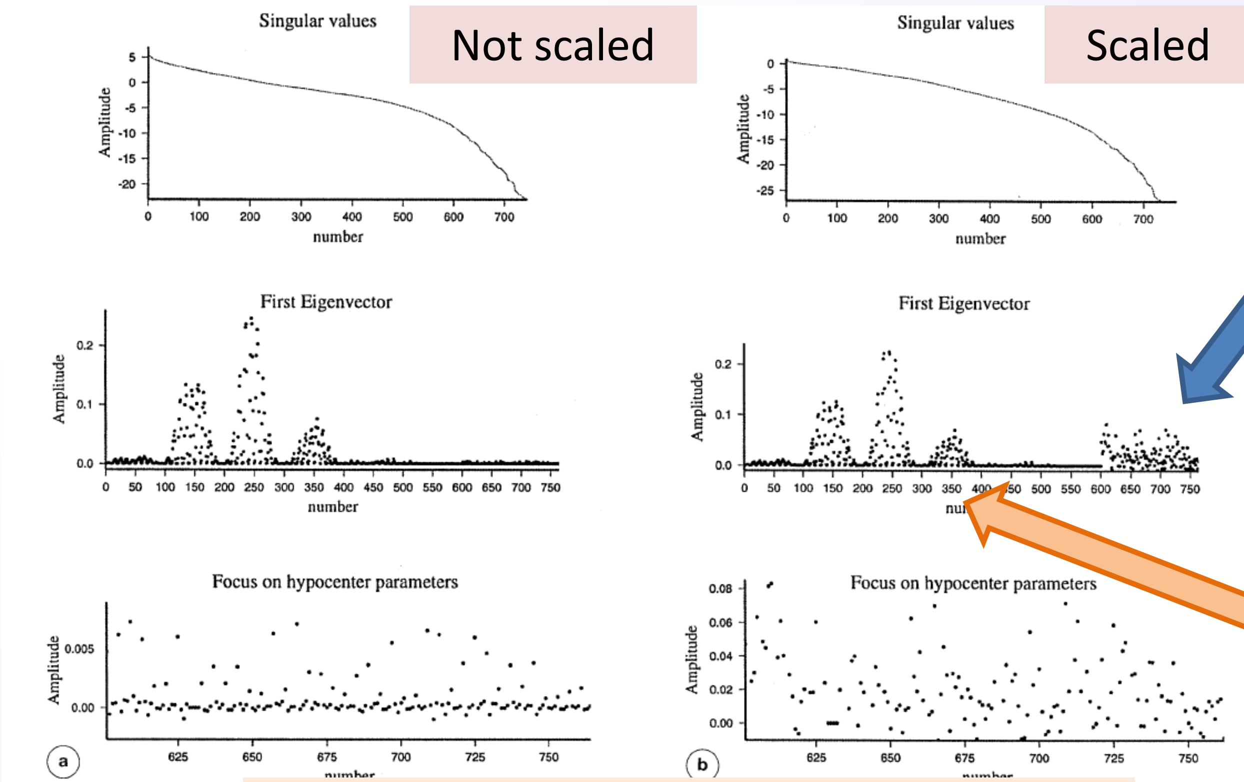
Events at 4.2 km



Random initial hypocenters

$Vp/XYZT_0$ inversion

SVD decomposition of sensitivity matrix



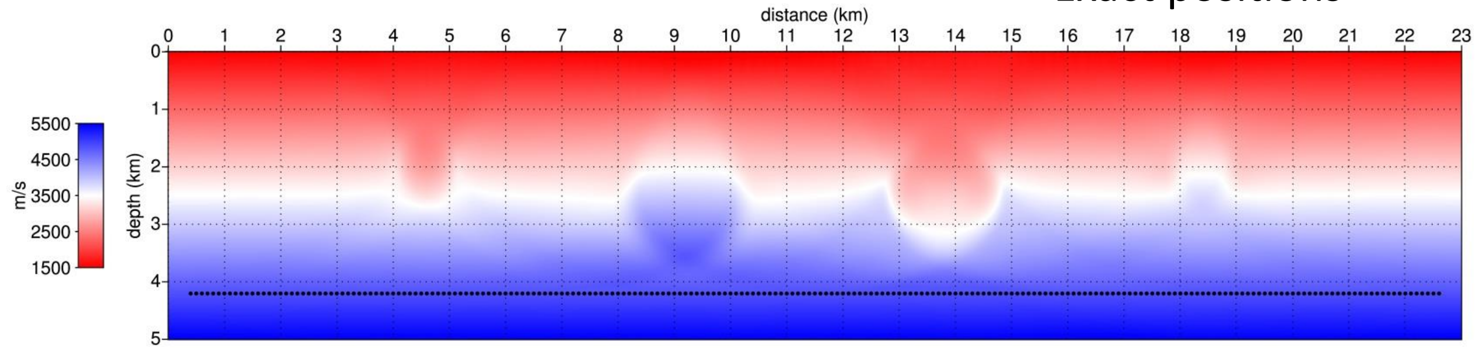
Zoom on hypocenter parameters (x,y,z,t0)

Hypocenter parameter (x,y,z,t0) for each event

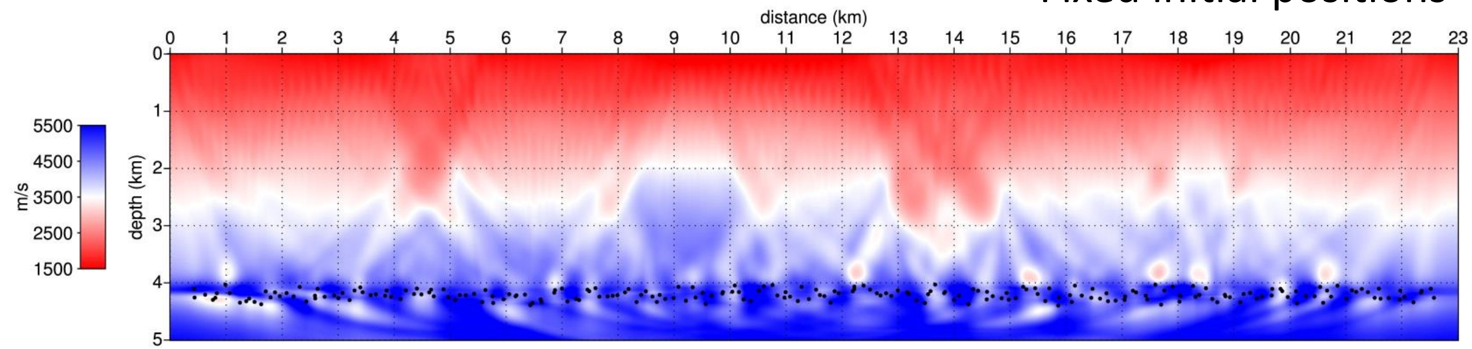
Vp parameter on a grid (640 dofs)

Vp inversion only

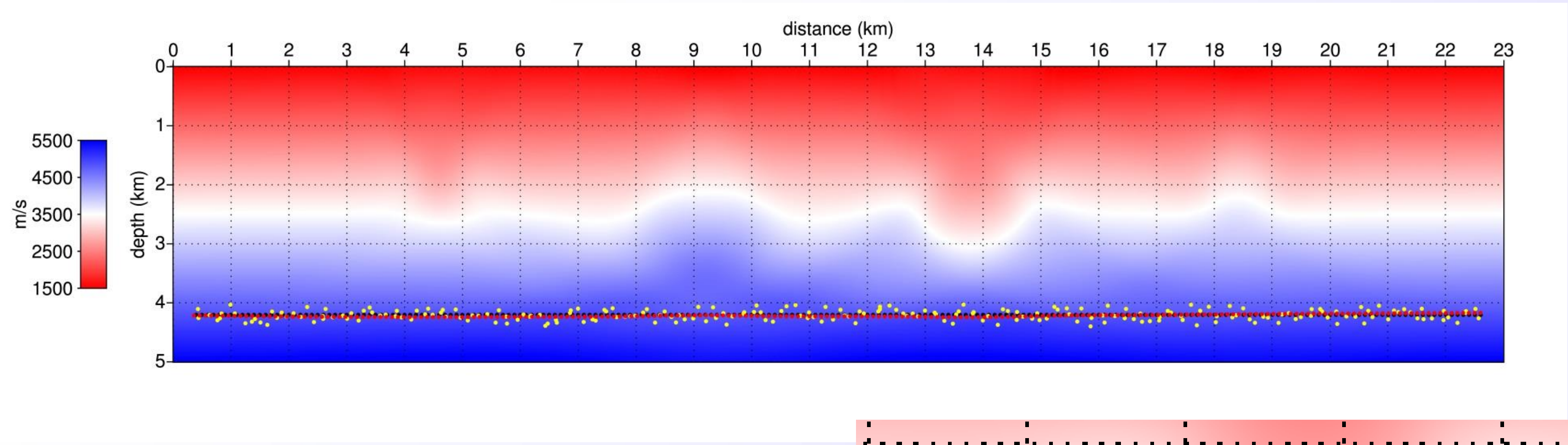
Exact positions

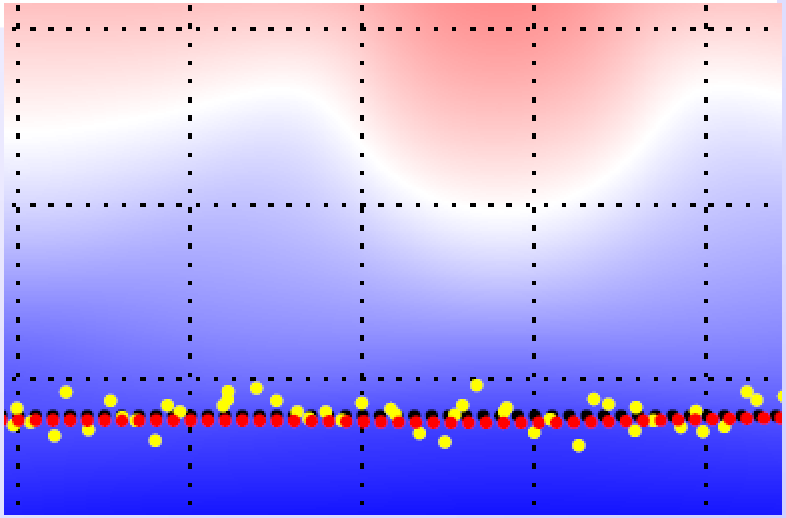


Fixed initial positions



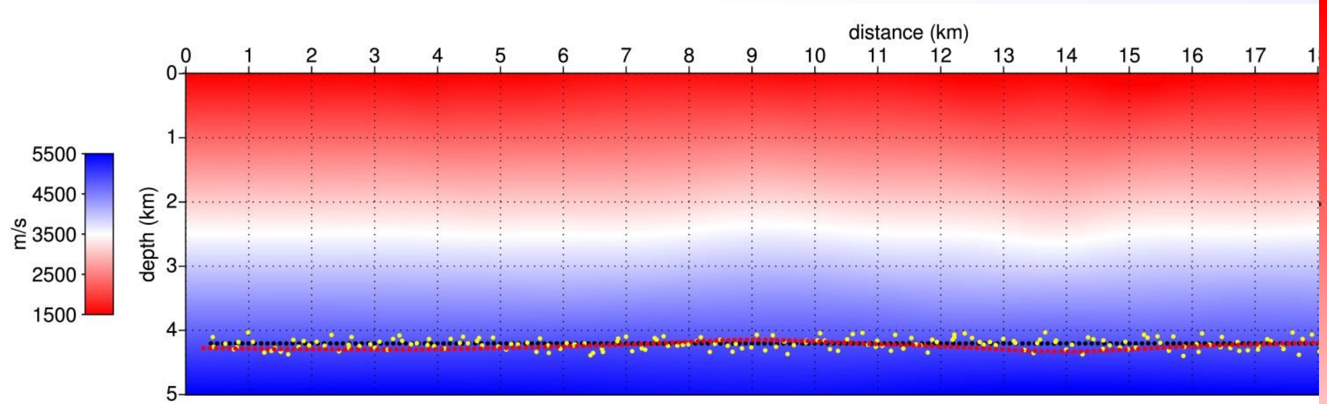
Vp and XYZ through l-BFGS

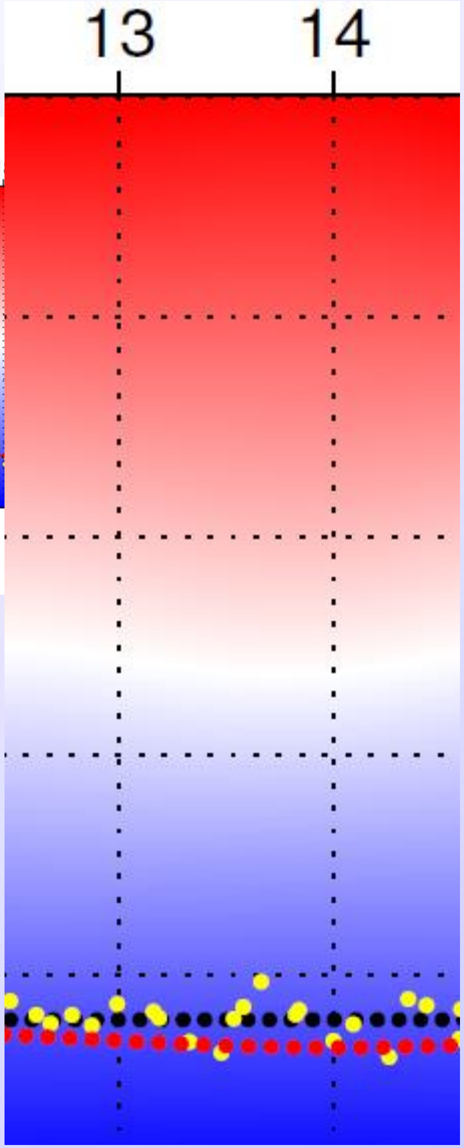




Same results with LSQR or TCN …

Vp and XYZT_0 through LSQR



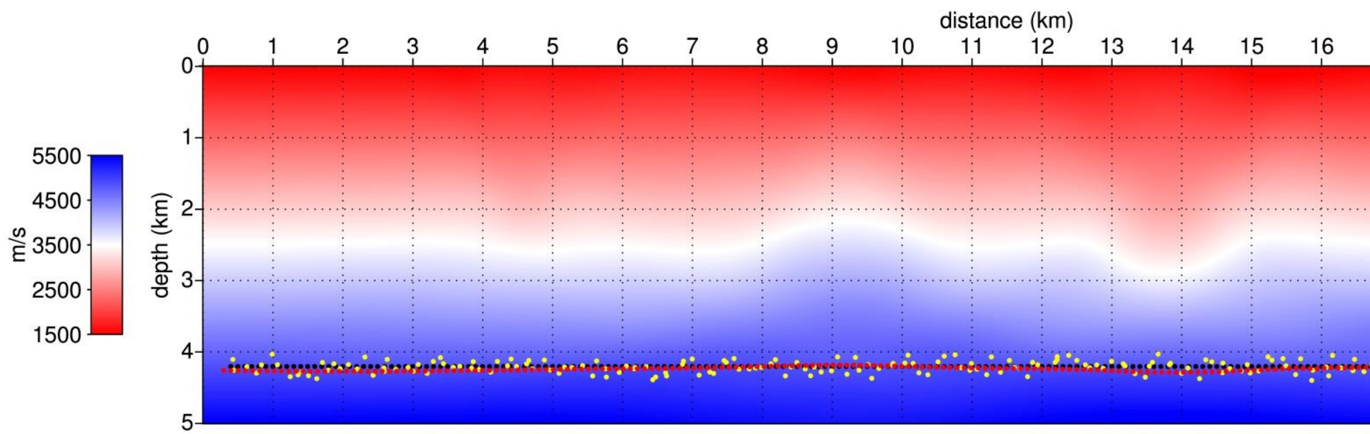


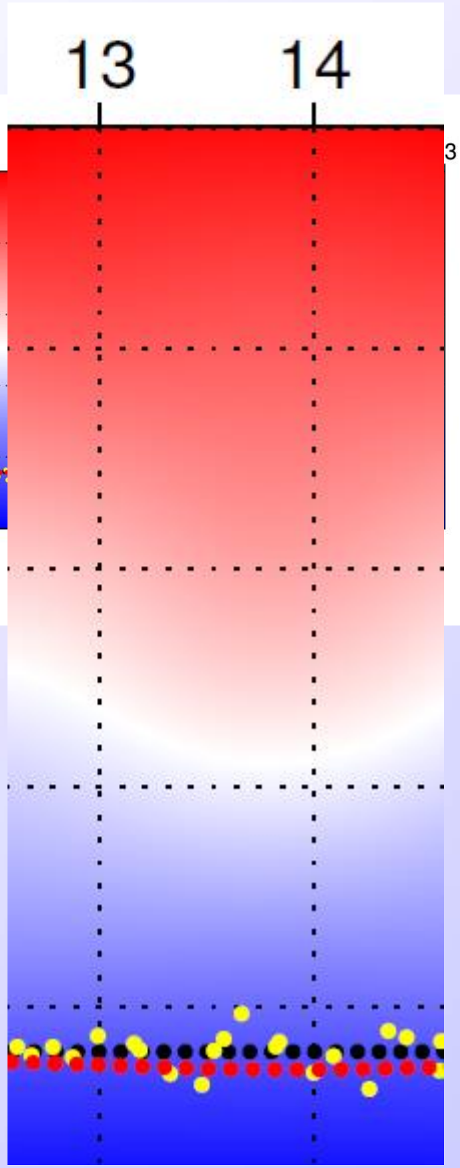
LSQR-like approach

solving the Gauss-Newton equation …

The introduction of the origin time seems to modify dramatically the multi-parameter inversion …

Vp and XYZT$_0$ through TCN



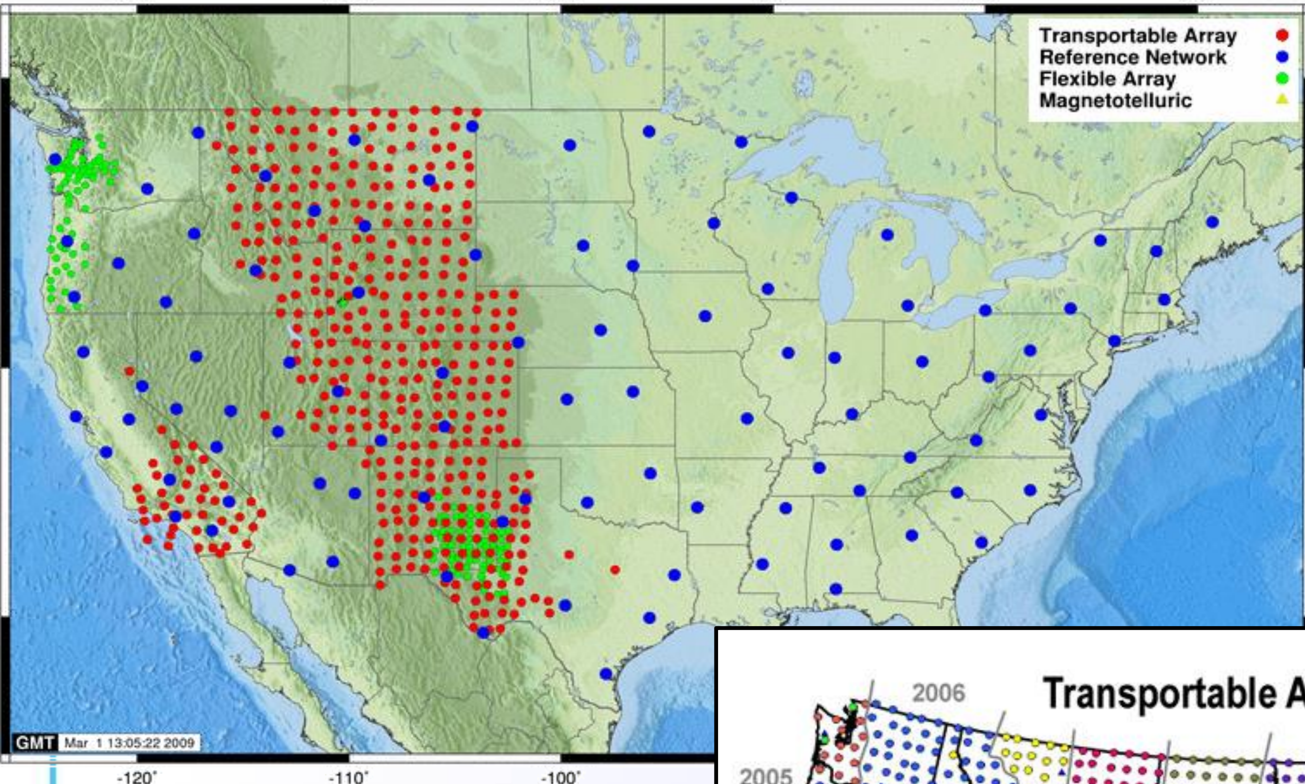


Only TCN seems to work more efficiently … and turns out to recover both hypocenter parameters (four) and velocity structure …

Expected improvements by preconditioning the normal equation …

Conclusion FATT

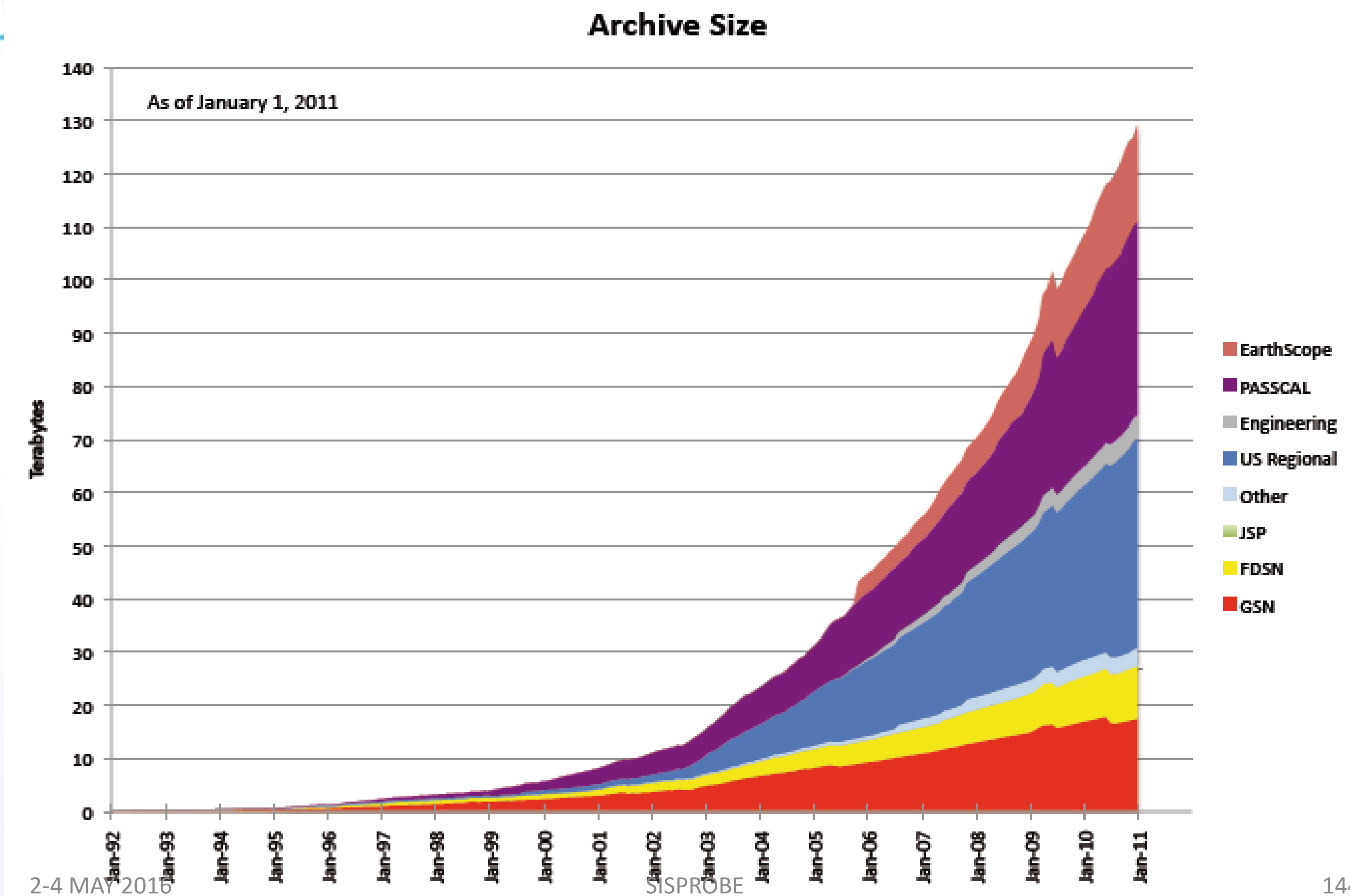
- Selection of an enough fine grid
- Selection of the a priori model information
- Selection of an initial model
- FMM and BRT for 2PT-RT
- Time and derivatives estimation
- LSQR inversion
- Update the model
- Uncertainty analysis (Lanzos or numerical)



General trend
for
seismic tomography

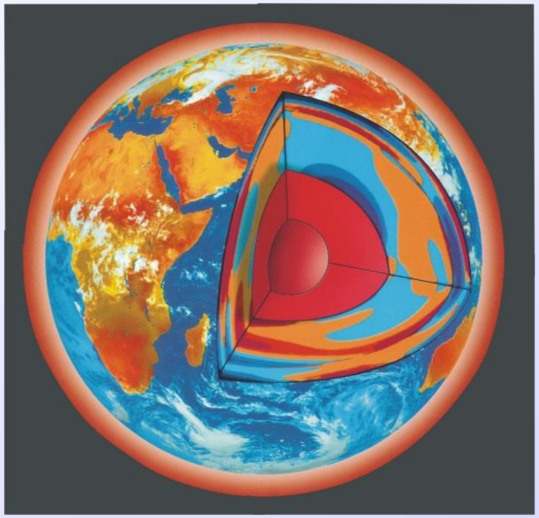
Dramatic increase of acquisition density both in the academic world (increasing density and in the industrial world (continuous recording)

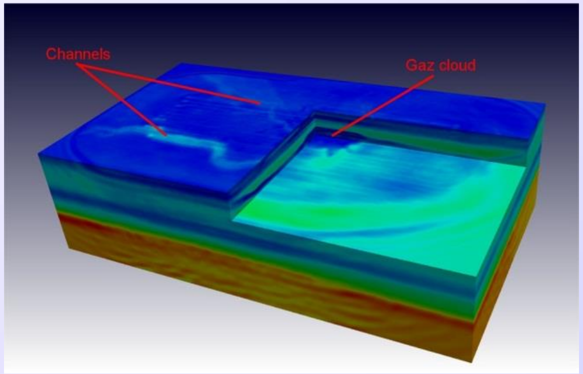
IRIS DATA CENTER



http://seiscope.oca.eu

THANK YOU !





THANK YOU !

Many figures have come from people I have worked with:

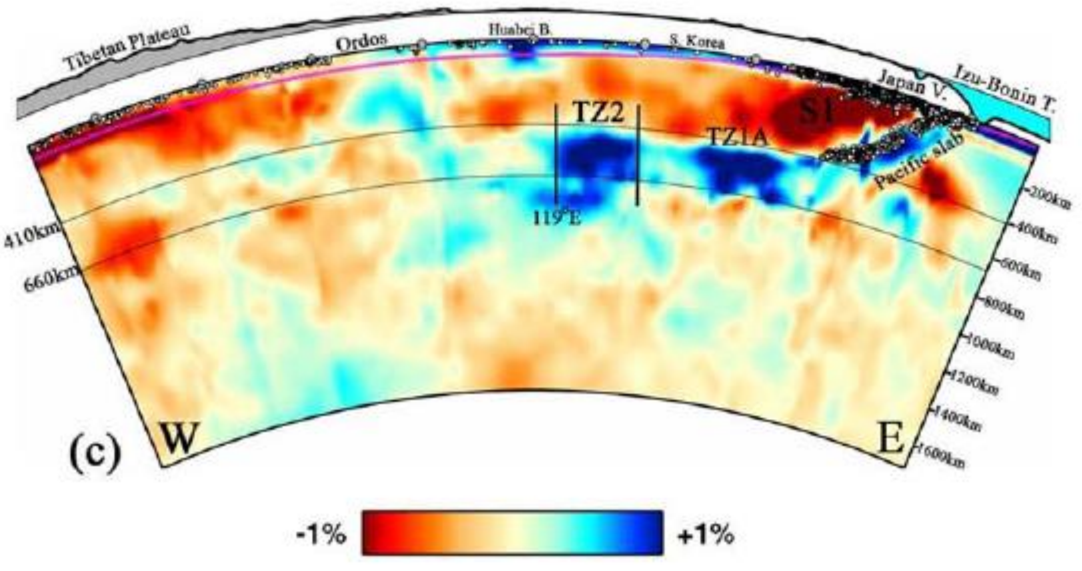
many thanks to them !

Other methods of exploration

- Grid search
- Monte-Carlo (ponctual or continuous)
- Genetic algorithm
- Simulated annealing and co
- Tabou method
- Natural Neighboring method

Delayed Travel-time Tomography

$$\delta \mathrm{T}(s,r) = \int u(x(l))dl = \iiint u(x)\delta\big(x - x(l)\big)K(s,r,x)dv$$



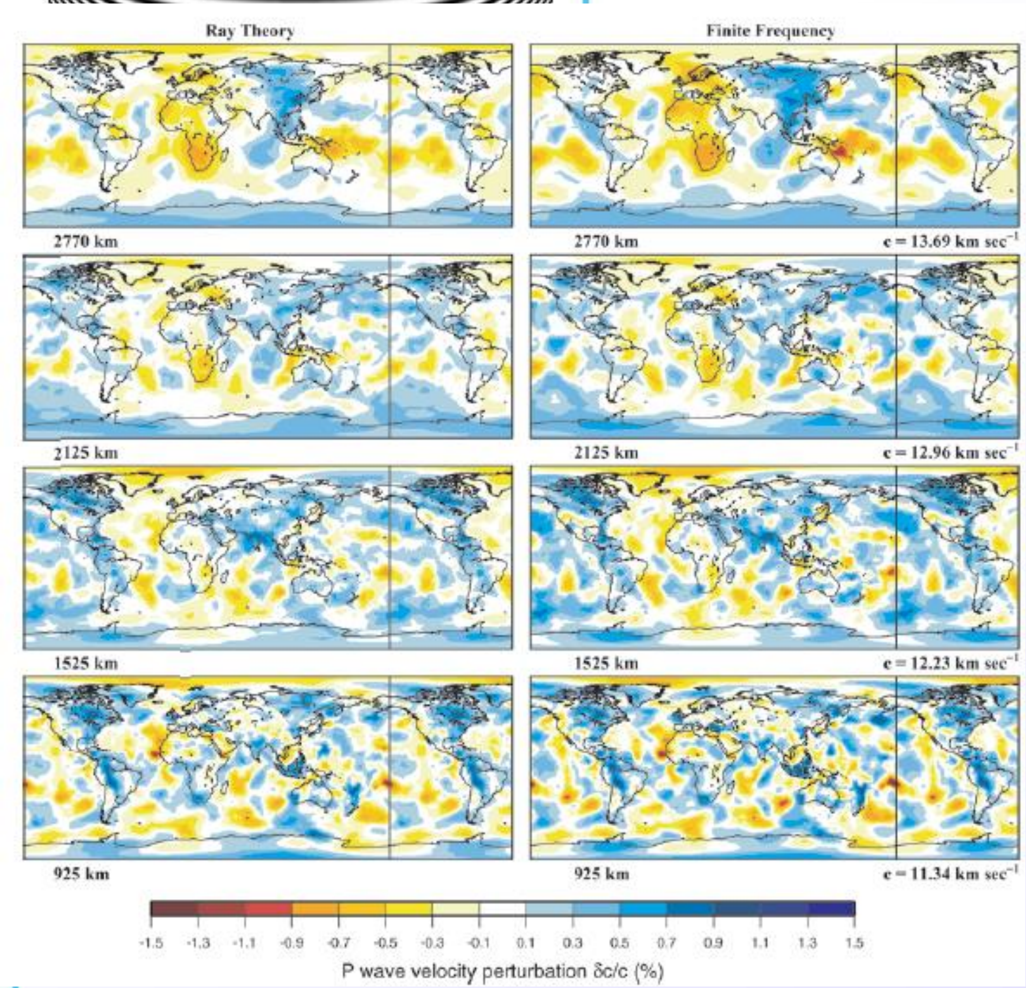
Still DTT provides impressive images while we do believe that DFT would provide better images in the future, thanks to finer discretization coming with the densification of the available data.

(Li & van der Hilst, 2010)

$$\delta \mathrm{T}(s,r) = \iiint u(x)\,K(s,r,x)dv$$

Delayed Fresnel Tomography

Delayed Fresnel Tomography



$$\delta \mathrm{T}(s,r) = \iiint u(x)\, K(s,r,x) dv$$

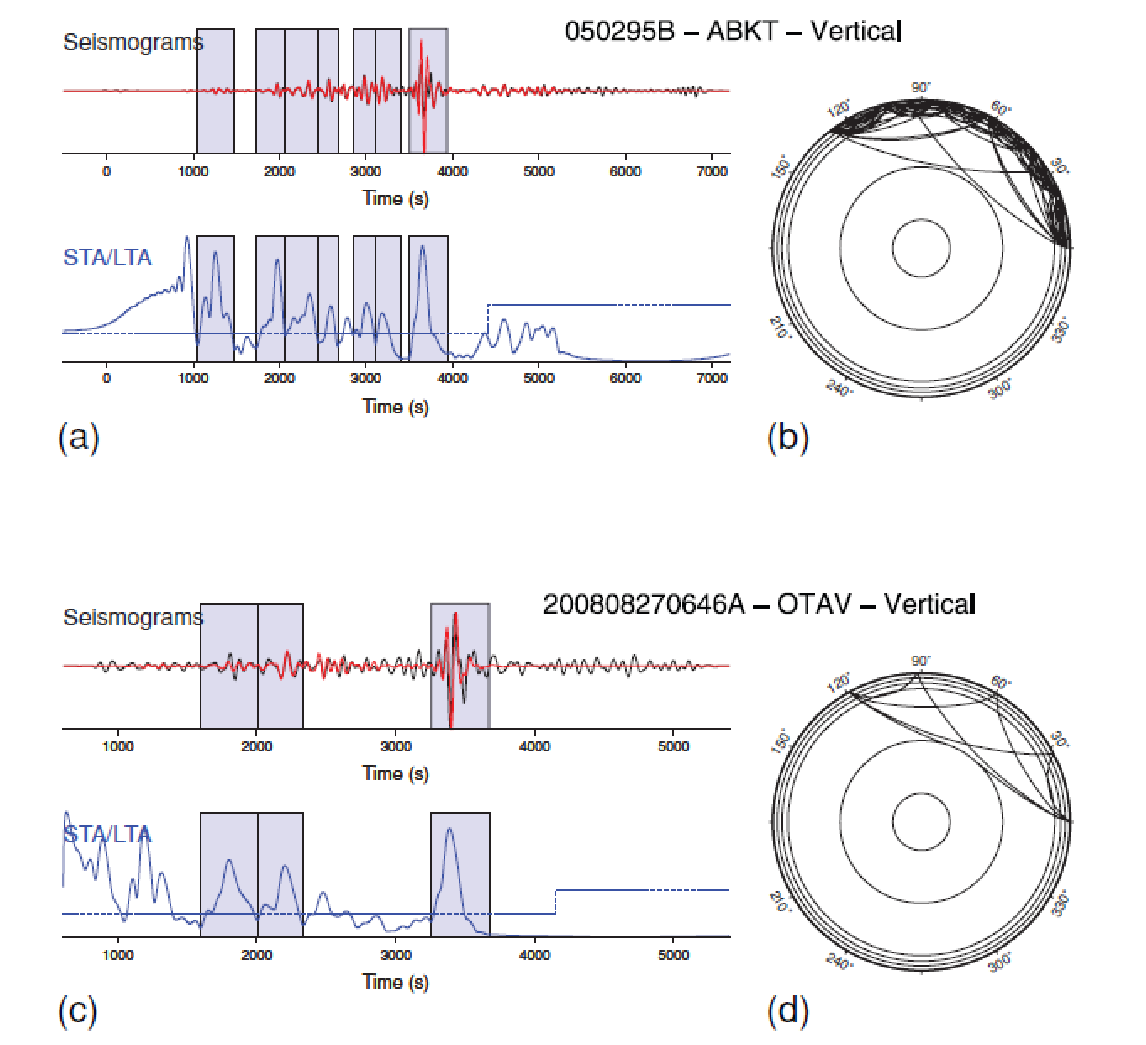
No amplitude measurements

Only phase information

Valid also for dispersive waves

Improved resolution?

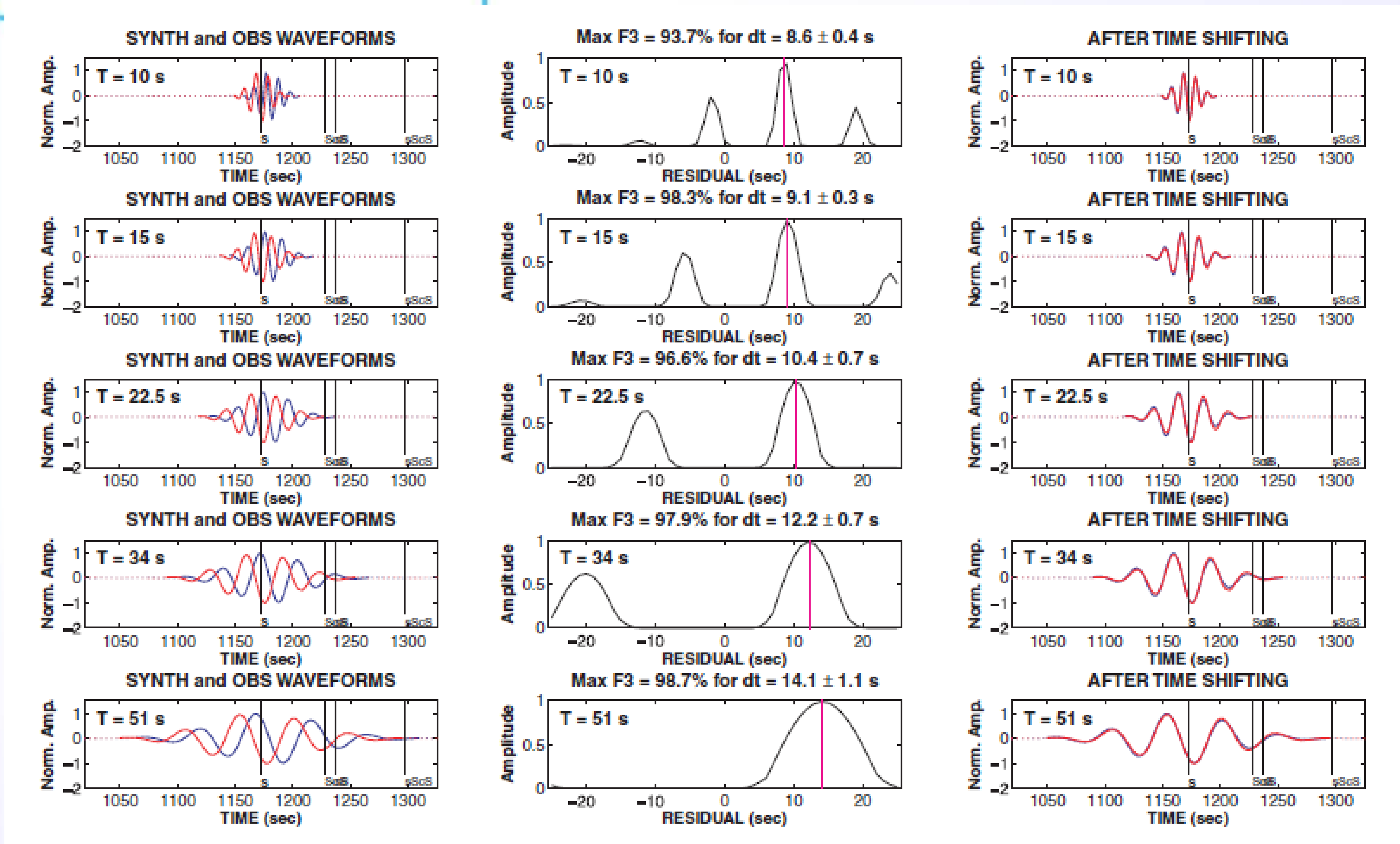
Windows definition



(LTA/STA;
Maggi et al, 2009)

(Wavelet Freq/time;
Lee and Chen, 2013)

Phases measurements

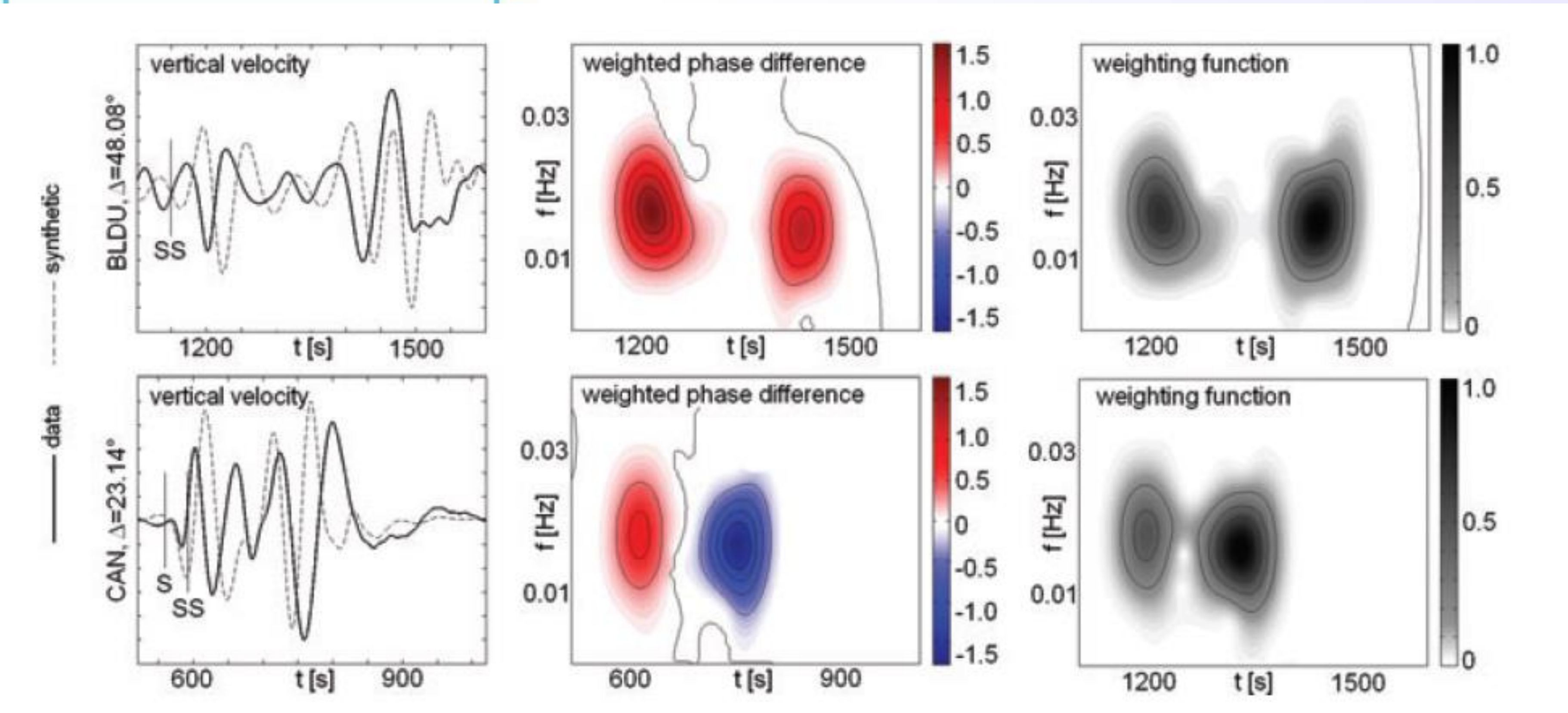


(Zaroli et al, 2010)

Workflow for DFT

- Taking seismograms
- Windowing phases
- Picking phases (cross-correlation)
- Definition of a misfit function (L2)
- Delayed Fresnel Tomography
- Model perturbation (velocity)

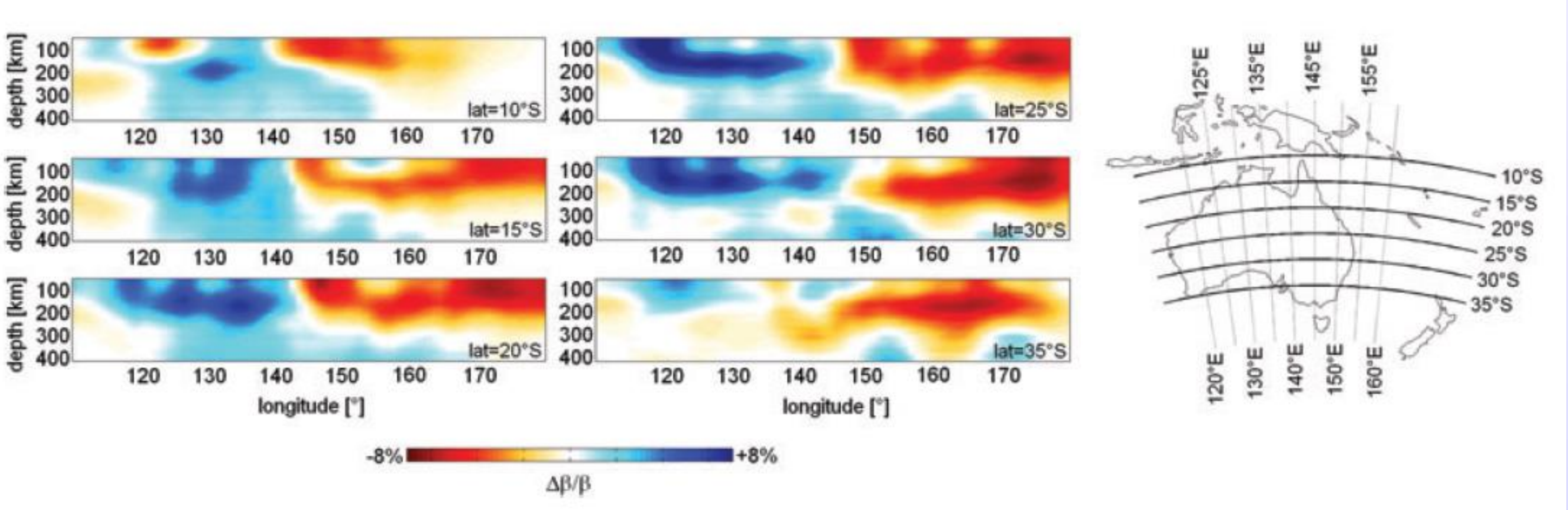
Phase differences



(Fichtner et al, 2009)

We do not pick phases as for DFT:

Full Phases Inversion



(Fichtner et al, 2009)

Do we improve the resolution compared to DFT?

If equivalent, very expensive; if not, smooth transition to FWI

Workflow for FPI

- Taking seismograms
- Windowing phases
- Definition of a misfit function (time differences)
- Full phases inversion
- Model perturbation (velocity)

END HERE

Tomographic descent

Minimisation of this vector

$$\frac{1}{2}\left\| \begin{bmatrix} C_d^{-1/2} d \\ -C_m^{-1/2} m_p \end{bmatrix} - \begin{bmatrix} C_d^{-1/2} g(m) \\ -C_m^{-1/2} m \end{bmatrix} \right\|^2$$

If one computes

$$A_k = \begin{bmatrix} -C_d^{-1/2} G_k \\ C_m^{-1/2} \end{bmatrix}$$

then

$$\left(A_k^t A_k\right)\delta m = A_k^t \begin{bmatrix} C_d^{-1/2}(g(m_k) - d) \\ C_m^{-1/2}(m_0 - m_k) \end{bmatrix}$$

Gaussian error distribution of data and of a posteriori model

Easy implementation once G_k has been computed

Extension using Sech transformation (reducing outliers effects while keeping L^2 norm simplicity)

THE $C_m^{-1/2}$ MATRICE

Shape independent of λ

Values depend on λ

The matrice Cm has a band diagonal shape

- σ is the standard error (same for all nodes)

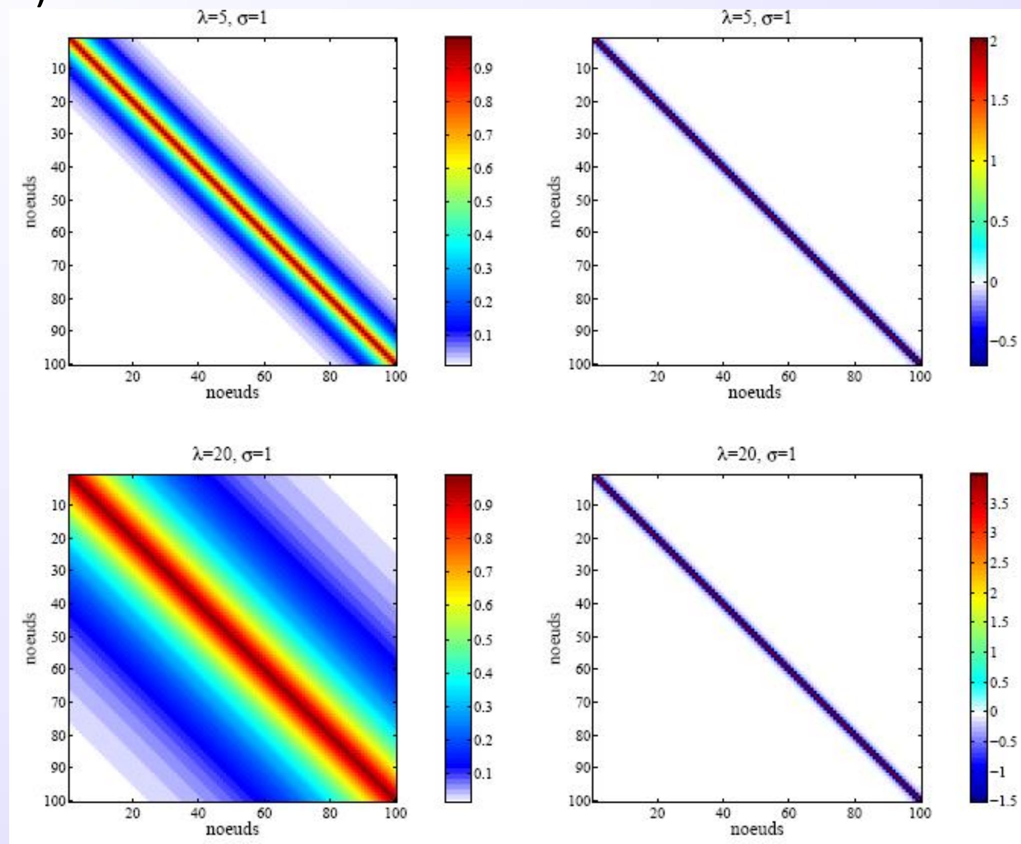
– λ is the correlation length

$$c_{ij} = \sigma^2 \exp\left(-\frac{|x_i - x_j|}{\lambda}\right)$$

n=nx.ny.nz=10^4 $C_m = USU^t$

(Lanzos decomposition)

$$C_m^{-1/2} = US^{-1/2}U^t$$



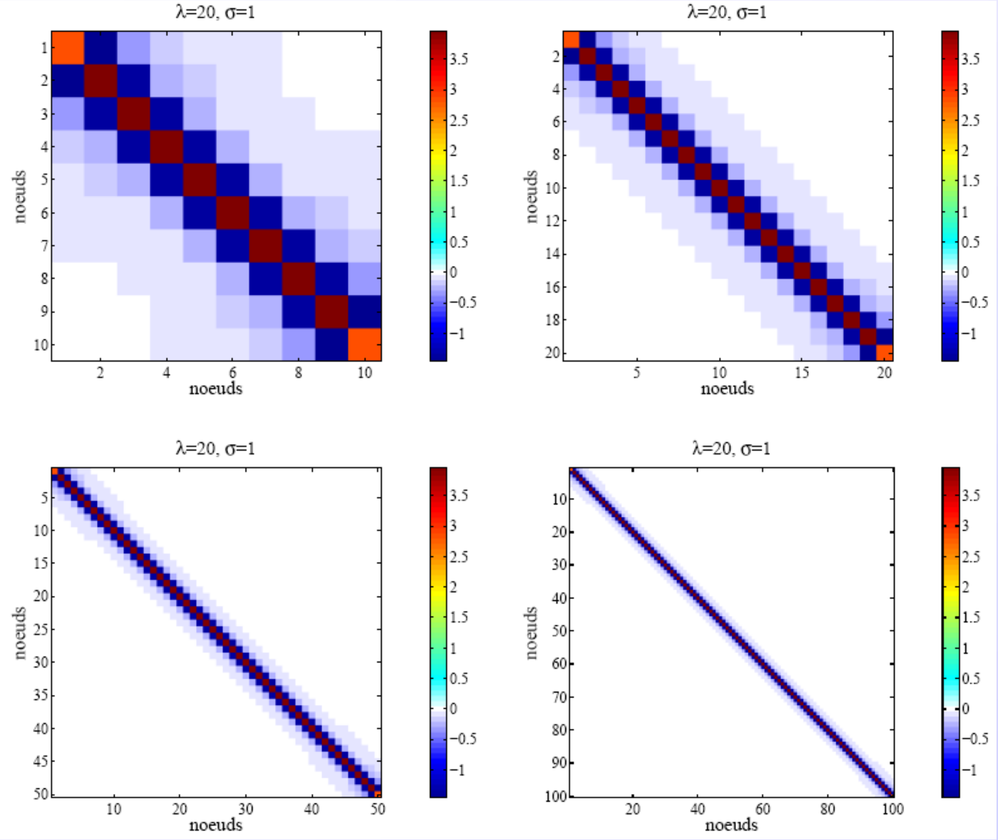
SATURATION

Analysis of coefficients

Values are only related to σ and λ

$$ {}^{n_0}C_m^{-1/2} \rightarrow {}^{n}\tilde{C}_m^{-1/2} $$

Values independent of n (n>5000)



Typical sizes 200x200x50

deduced from 20x20x5 (few minutes)

Strategy of libraries of $C_m^{-1/2}$ for various λ and σ=1

Other coefficients could be deduced

R: $C_m^{-1/2}$ sparse matrice

An example

σ_v=100 km/s

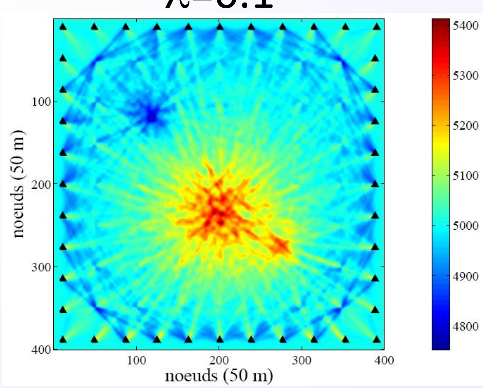
σ_x=100 km

σ_t=100 s

Same numerical grid for all simulations (either 100x100 or 400x400)

Same results at the limit of numerical precision related to the estimation of the sensitivity matrice

λ=0.1



Ray imprints

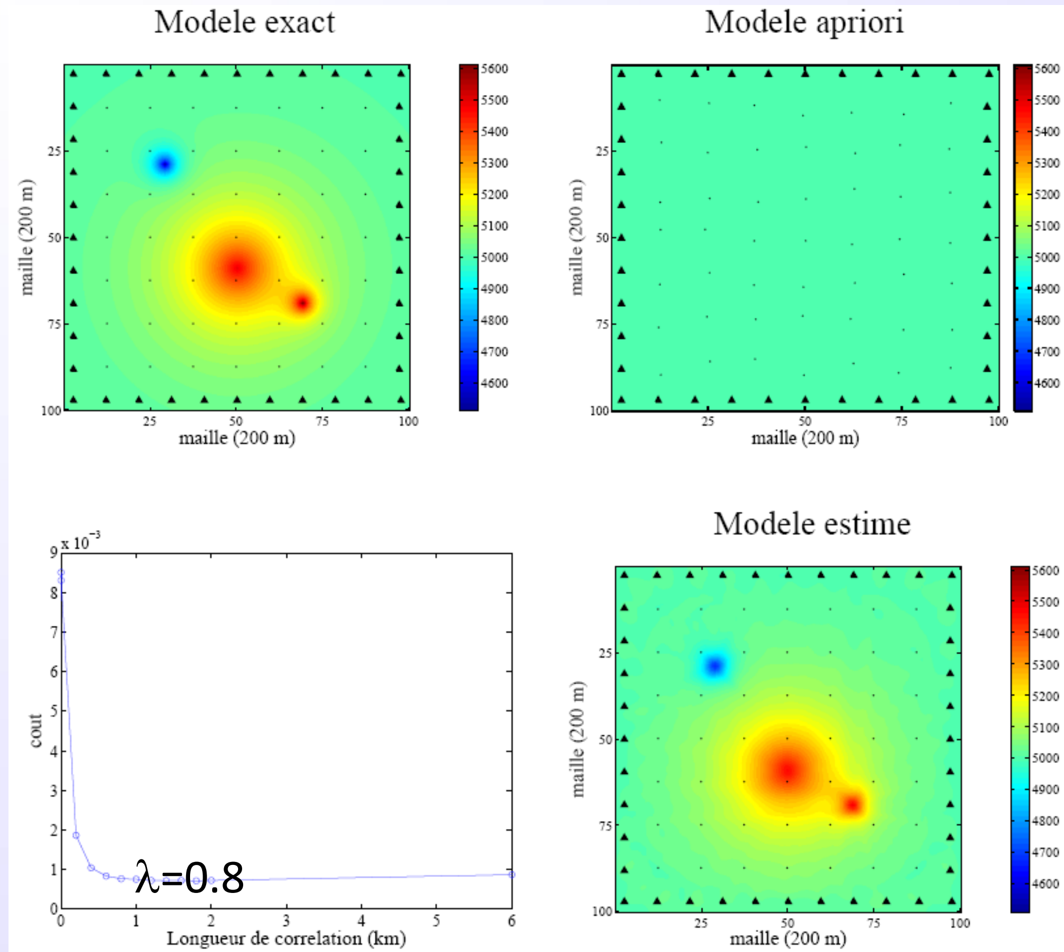
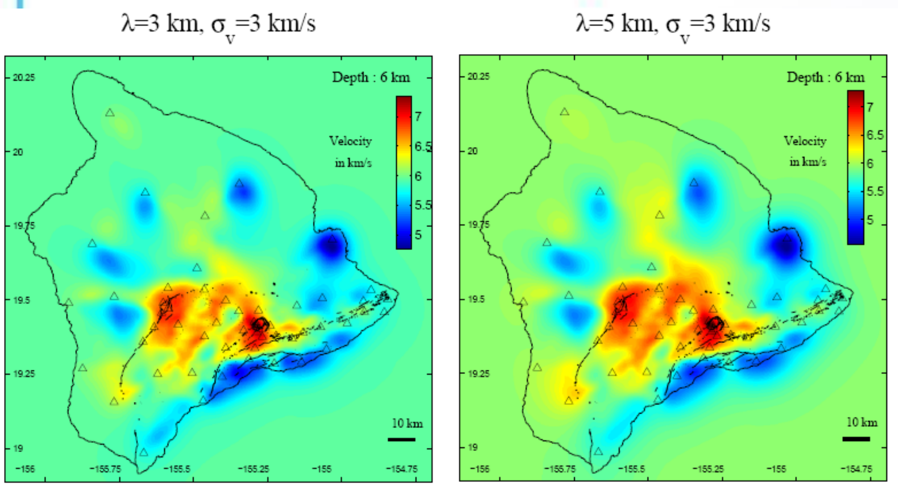
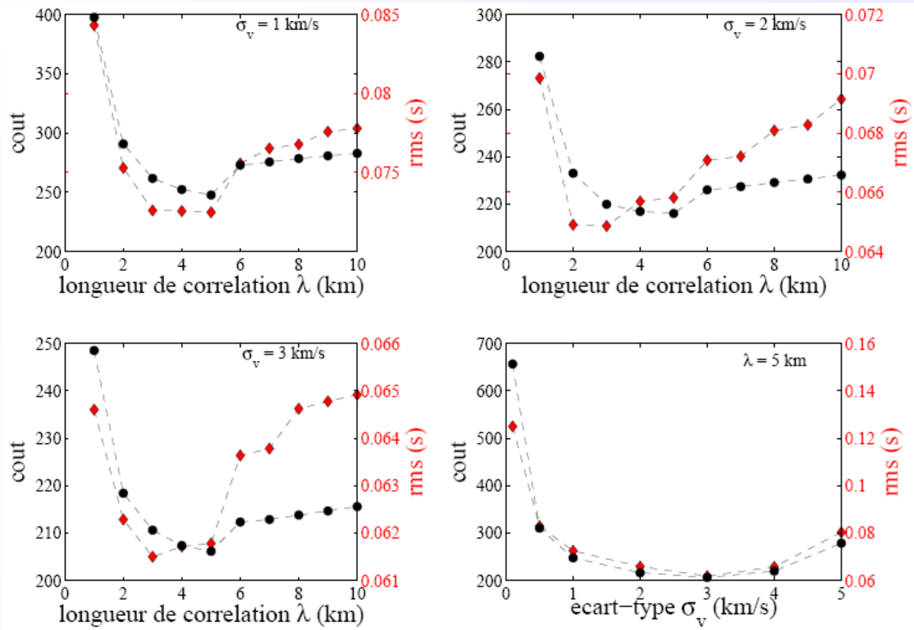


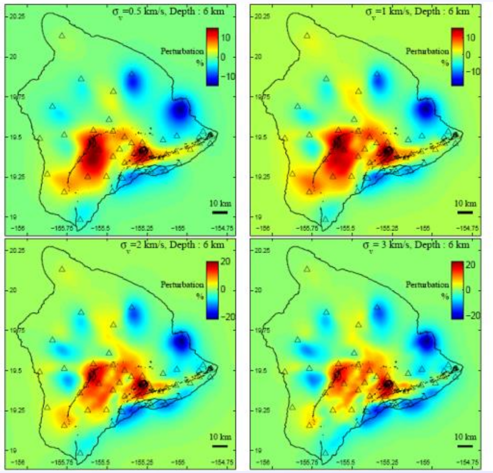
Illustration of selection $\{\lambda,\sigma_v\}$



λ influence



λ= 5 km and σ_v= 3 km/s

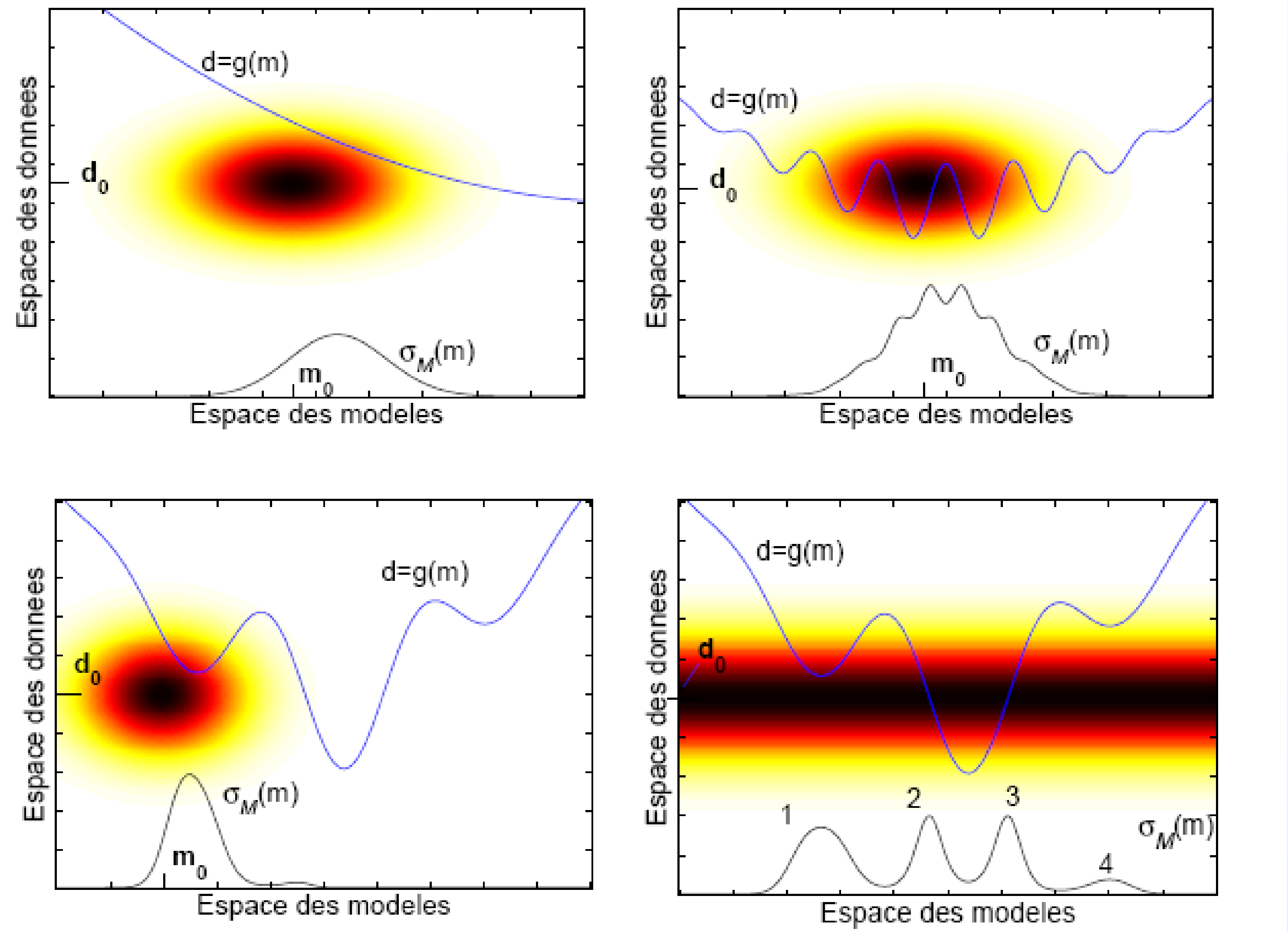


σ_v influence

Error function analysis will give us optimal values of a priori standard error and correlation length (2D analysis)

A posteriori information

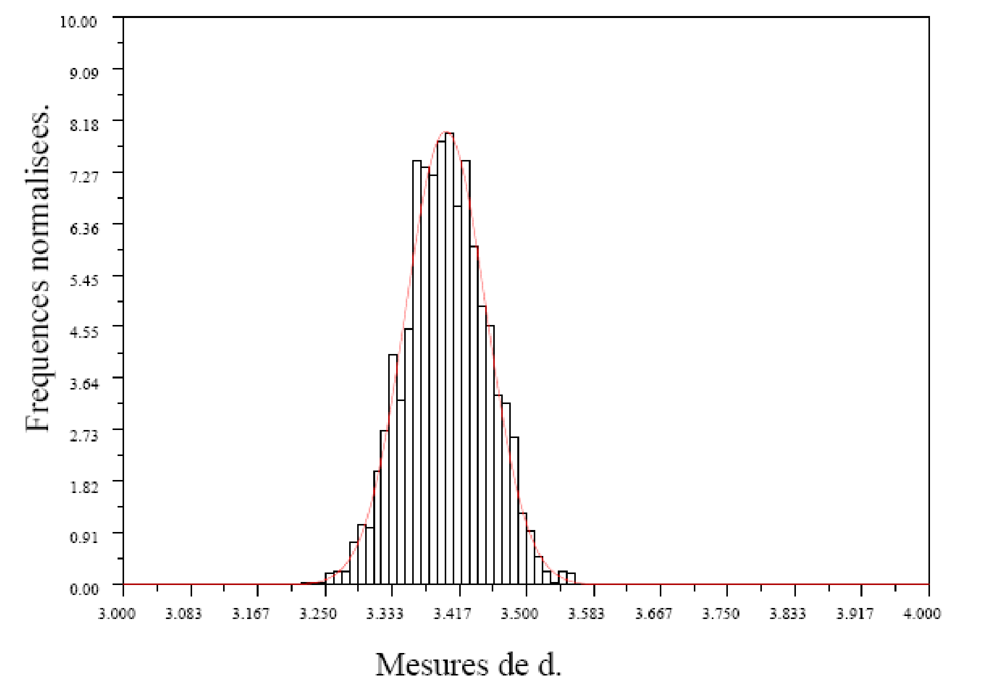
What is the meaning of the « final » model we provide ?



acceptable

Probability density

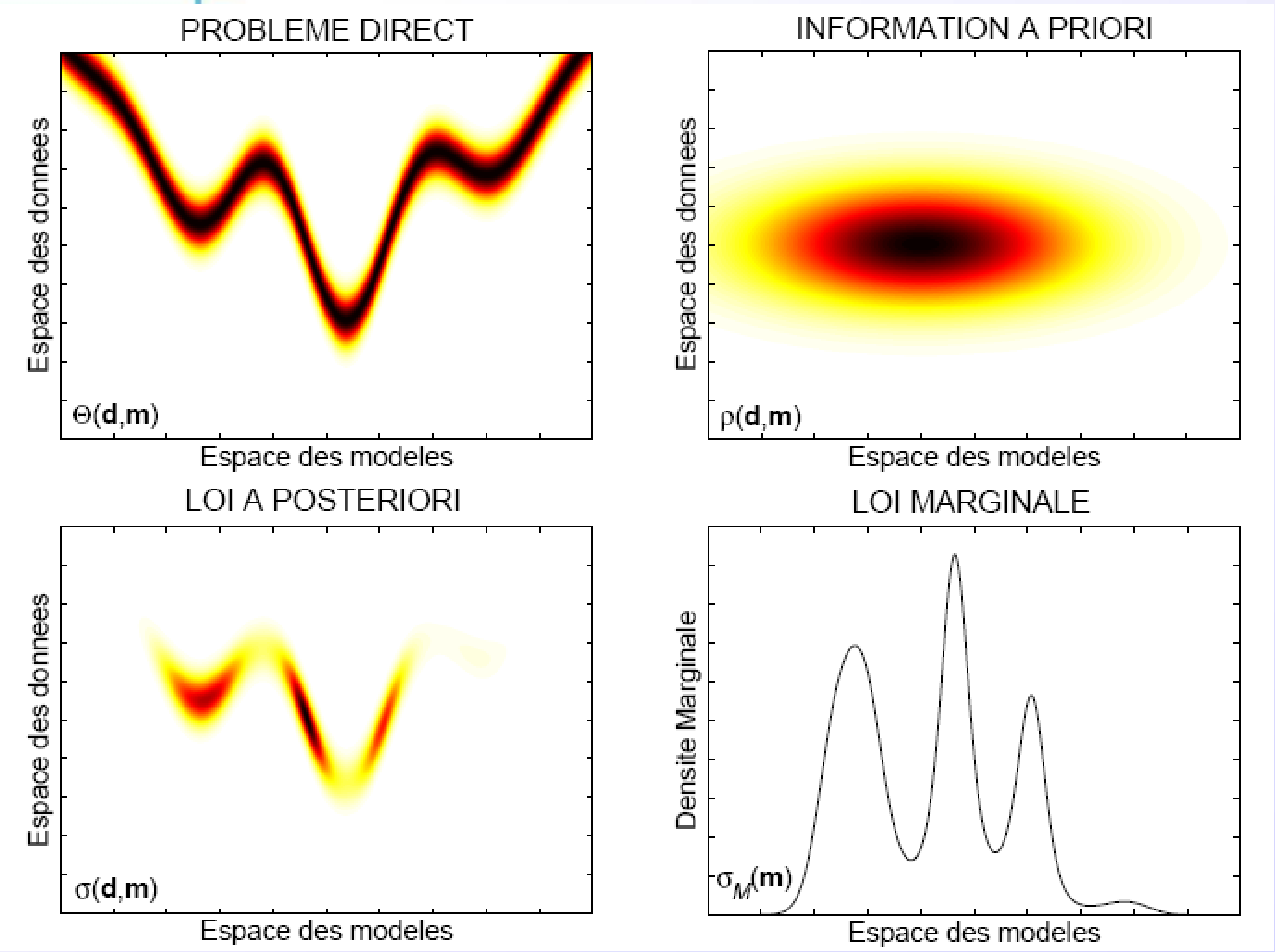
Data space



$$\rho(d, m) = \rho_D(d)\,\rho_M(m)$$

The probability density in the model space is related to the

Different informations



Perspectives

- A priori model covariance : location dependent
  wavelet decomposition allows local analysis

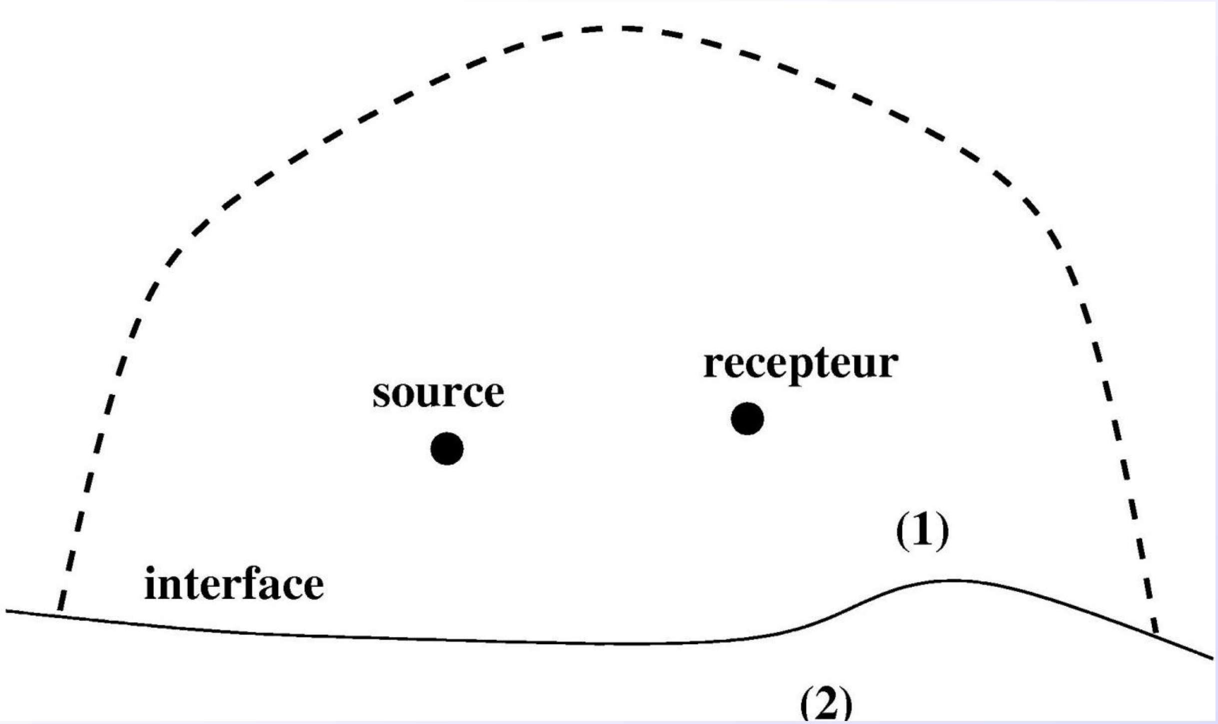
(work of Matthieu Delost-Geoazur)

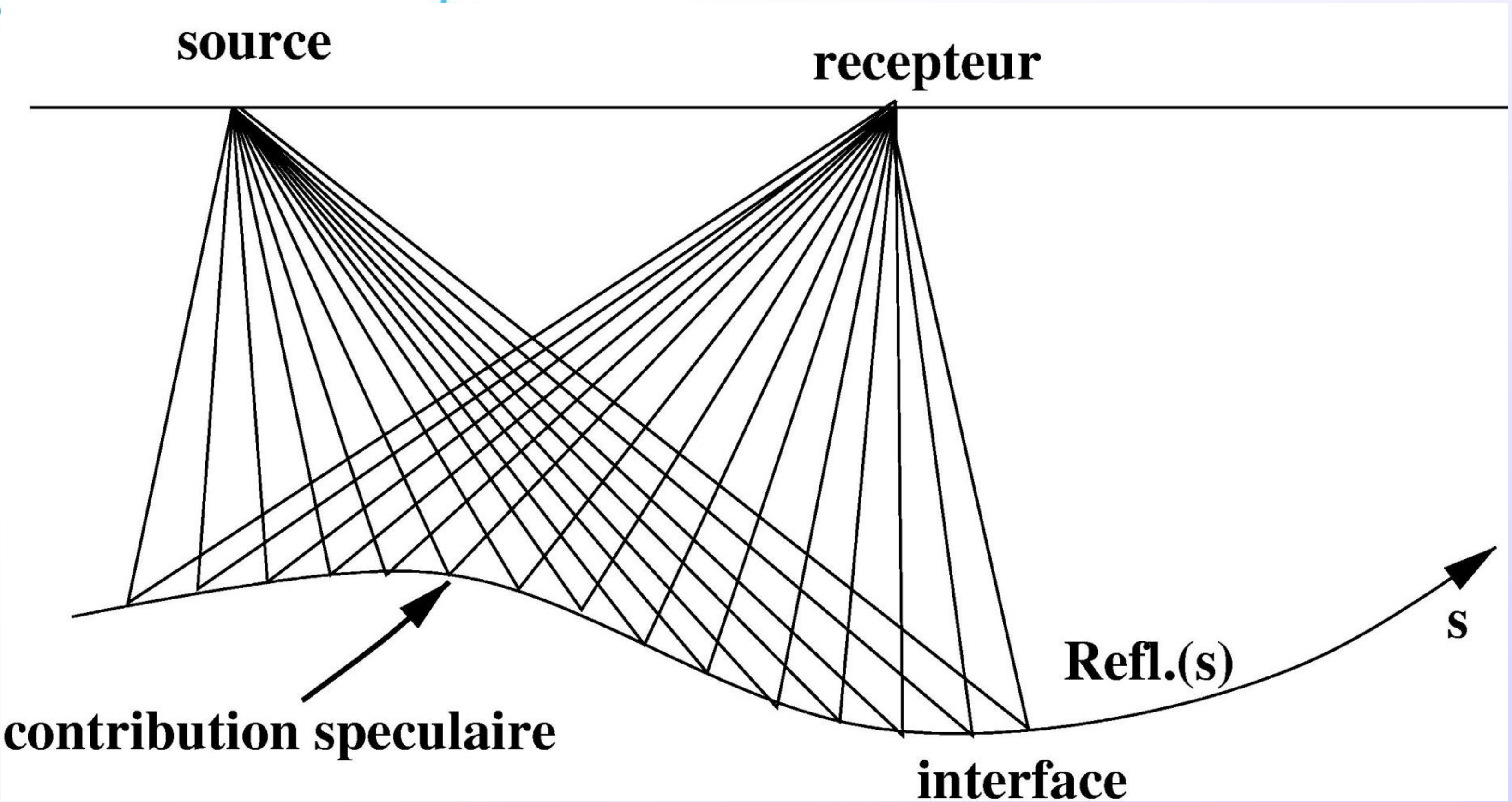
- Fresnel tomography : introduction of the first Fresnel zone influence in the forward problem

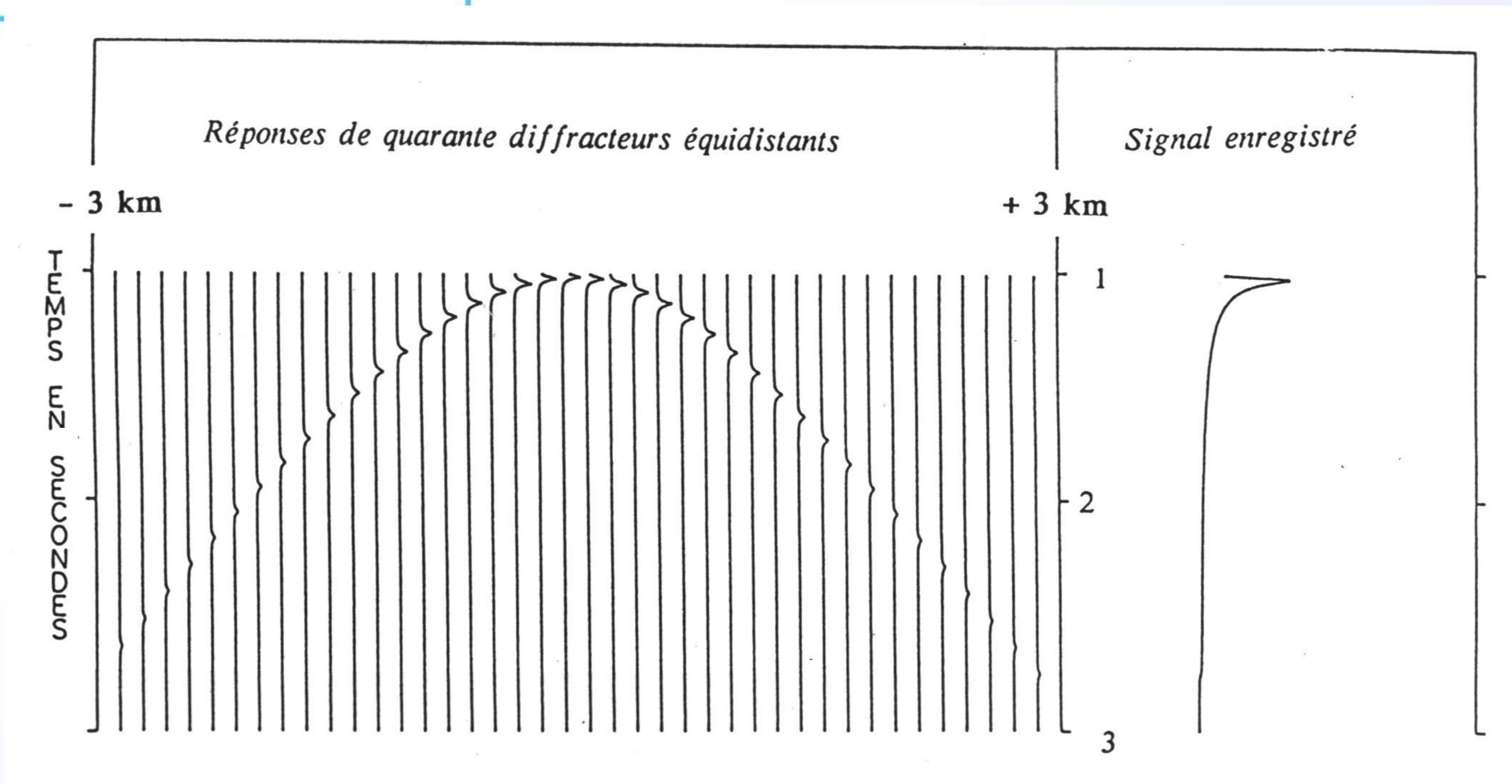
(work of Stéphanie Gautier-Montpellier)

Kirchhoff approximation

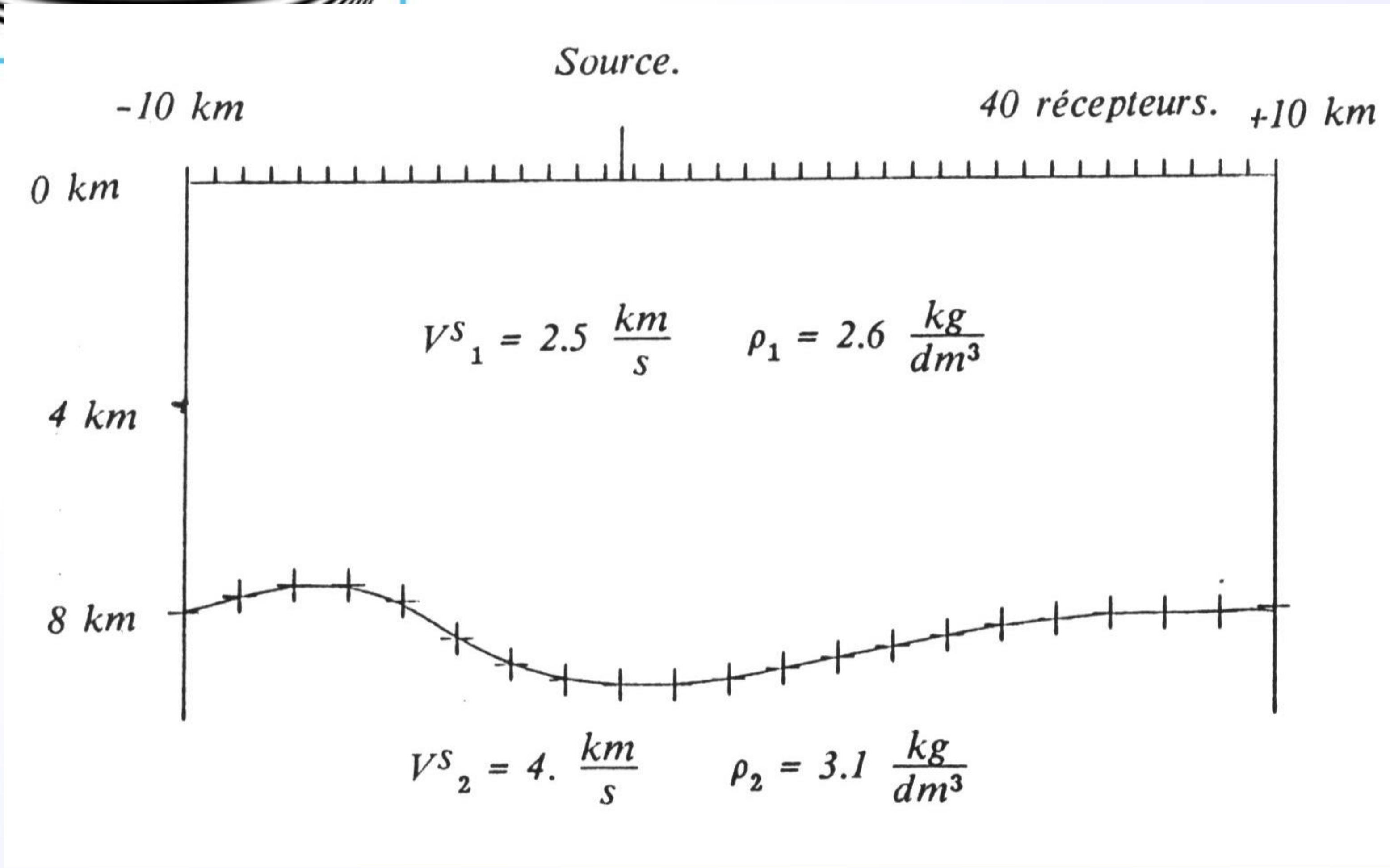
1) Representation theorem
2) Kirchhoff summation
3) Reciprocity


recepteur
source
(1)
interface
(2)

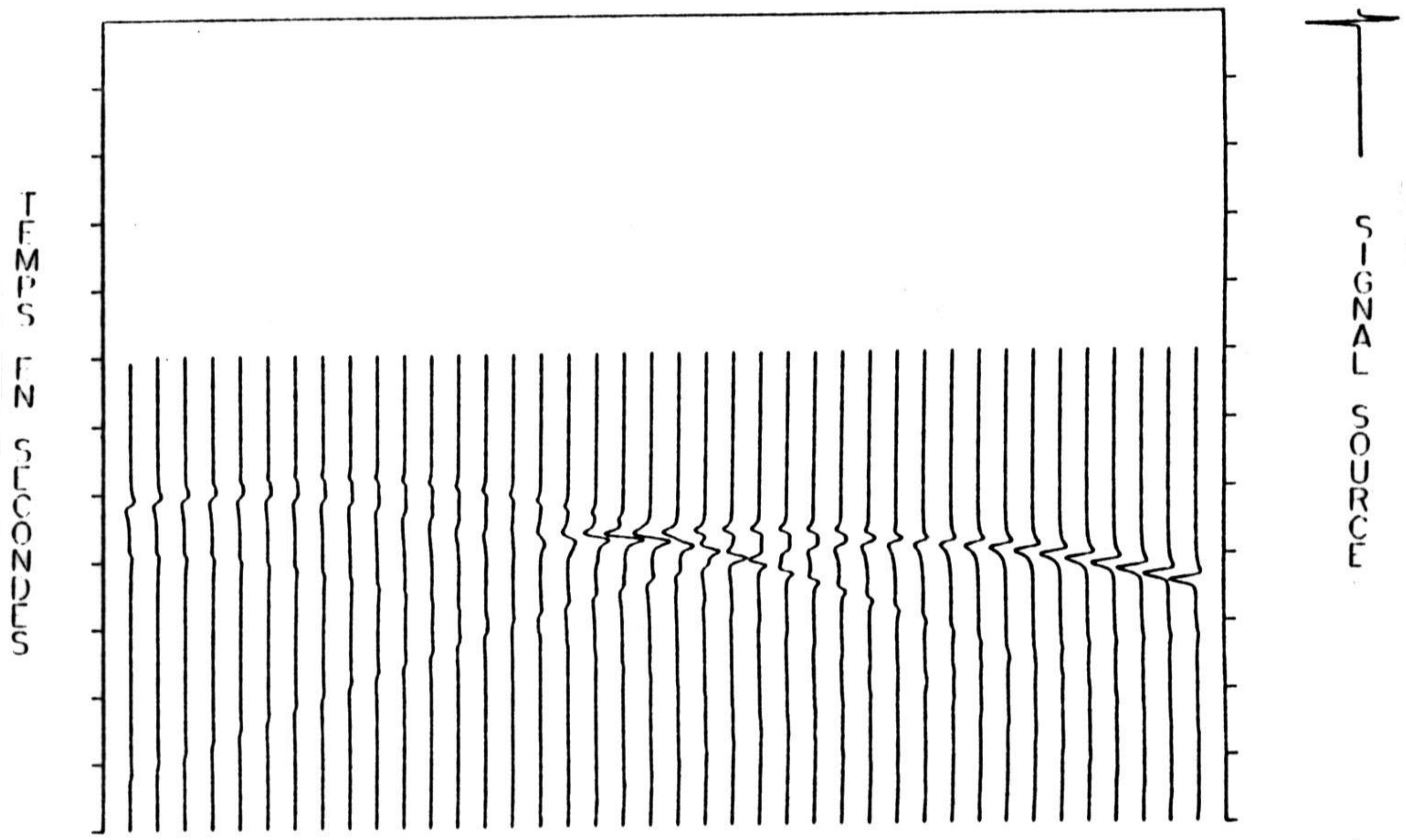

source
recepteur
s
Refl.(s)
contribution speculaire
interface



Réponses de quarante diffracteurs équidistants
Signal enregistré
- 3 km
+ 3 km
TEMPS EN SECONDES
1
2
3

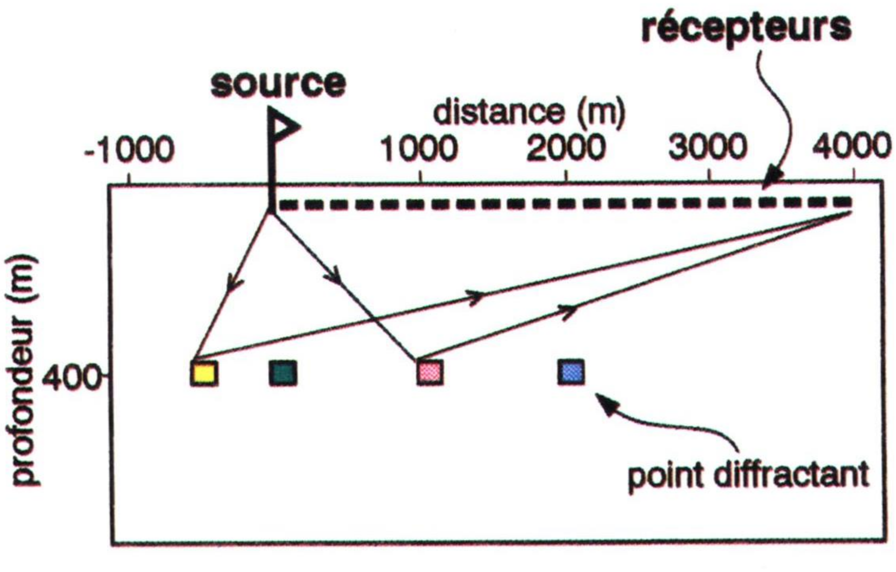

Source.
-10 km
40 récepteurs.
+10 km
0 km
4 km
8 km
$V^S{}_1 = 2.5 \frac{km}{s}$
$\rho_1 = 2.6 \frac{kg}{dm^3}$
$V^S{}_2 = 4. \frac{km}{s}$
$\rho_2 = 3.1 \frac{kg}{dm^3}$

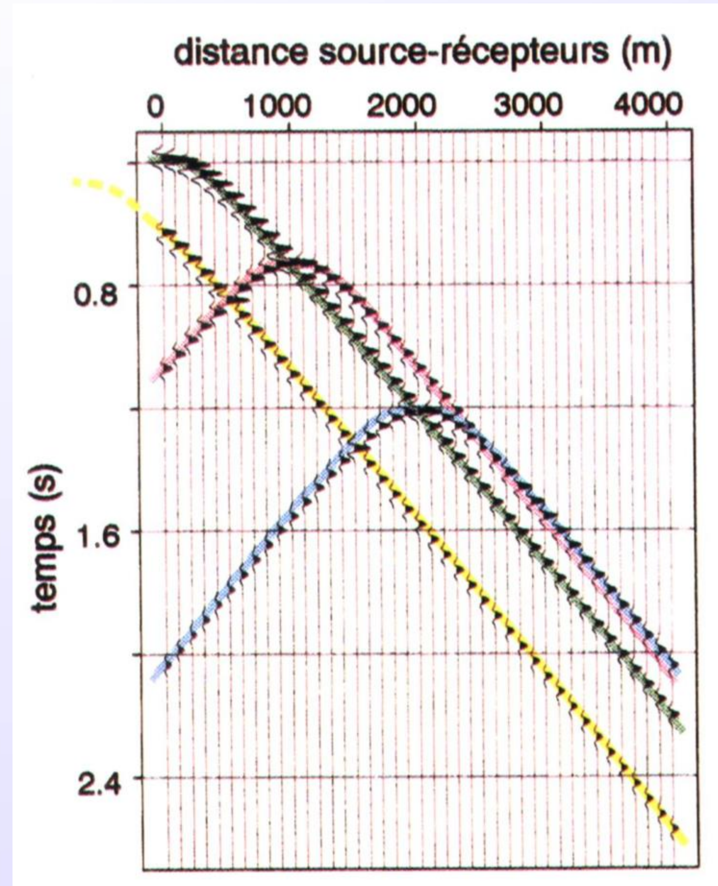
SISMOGRAMME CONVOLUE PAR UN SIGNAL SOURCE.

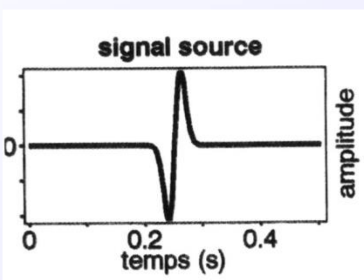


Born approximation

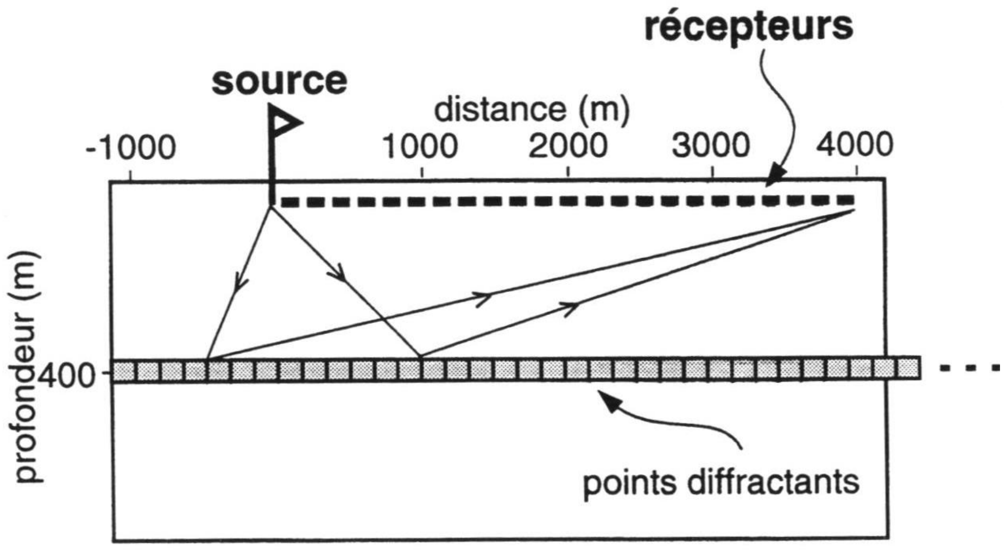
1) Single scattering approximation
2) Surface approximation

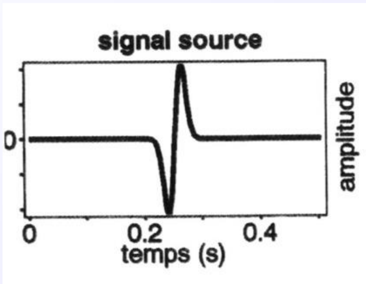


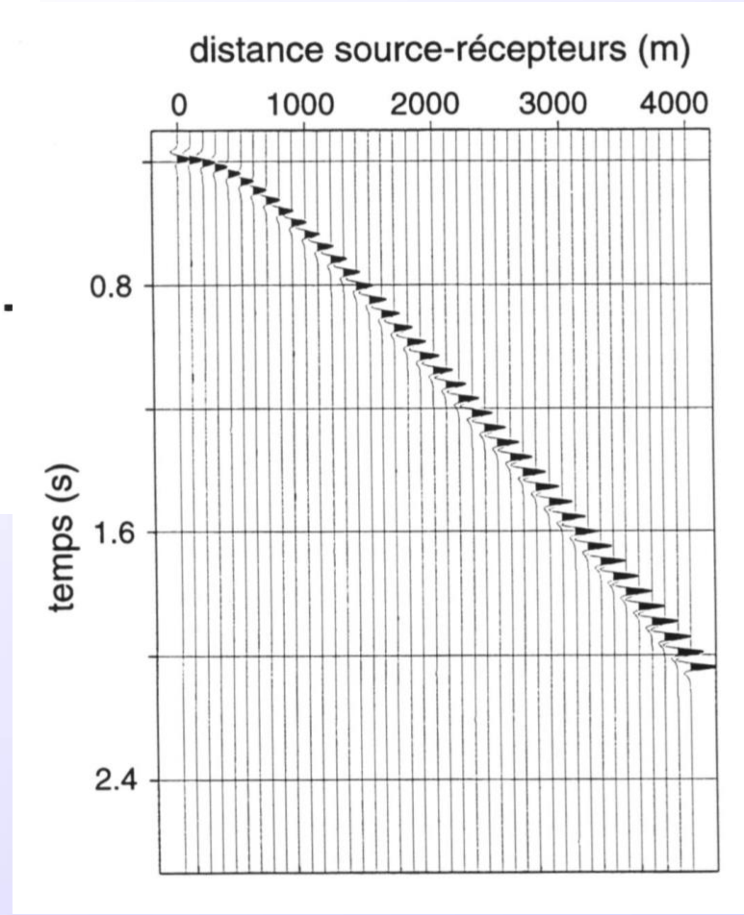




(Forgues et al., 1996)



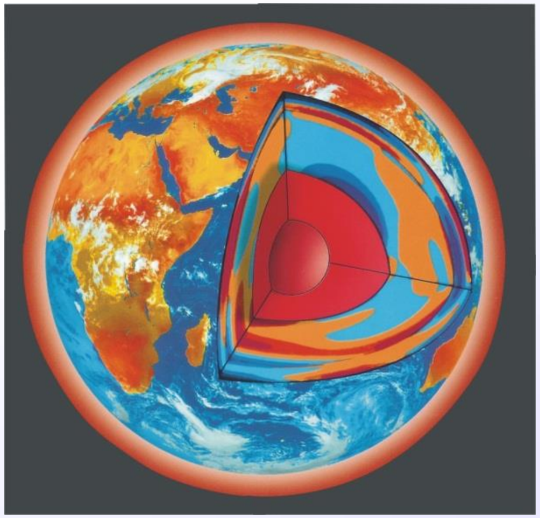


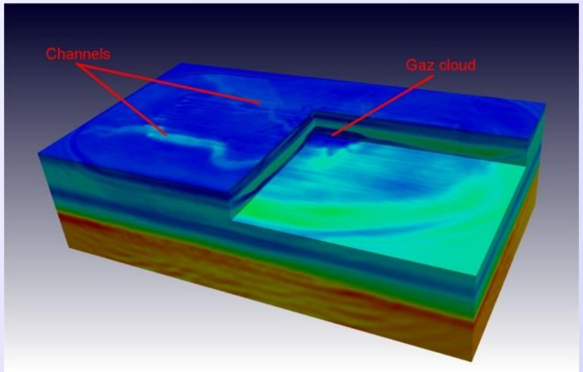


(Forgues et al., 1996)

http://seiscope2.osug.fr

THANK YOU !





Linearized Inverse Problem

Misfit function

$$\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$$

One-loop procedure

- First loop over models : for current model s_k (iteration k)
  - Forward problem
    A. Solve eikonal equation
    B. Compute synthetic travel-times and delayed times δt_k
  - Adjoint problem
    A. Solve adjoint equation
    B. Build gradient γ_k
- Descent method (one-step gradient or conjugate gradient or l-BFGS) δs_k

Update the model $s_{k+1} = s_k + \alpha_k \delta s_k$
(Sei & Symes, 1994; Leung & Qian, 2006; Taillandier et al, 2009)

Complexity: $\mathcal{O}(N_{src})$ for forward/adjoint formulation and $\mathcal{O}(N_m)$ for gradient

Linearized Inverse Problem

Misfit function

$$\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$$

Two-loop procedure

- First loop over models : for current model s_k (iteration k)
  - Forward problem
    A. Solve eikonal equation
    B. Compute synthetic travel-times and delayed times δt_k
  - Adjoint problem
    A. Solve adjoint equation
    B. Build gradient γ_k
- Second loop over linear system : $H_k\ \delta s_k = -\gamma_k$
  - Truncated Newton method based on iterative conjugate gradient
  (matrix free approach for product $H_k\ v$, thanks to second-order adjoint formulation)

Update the model $s_{k+1} = s_k + \alpha_k \delta s_k$ (Métivier et al, 2013)

Full Hessian impact $H = J^t J + \frac{\partial J}{\partial s}\delta t$ is evaluated

Linearized Inverse Problem

Two-loops procedure

Misfit function

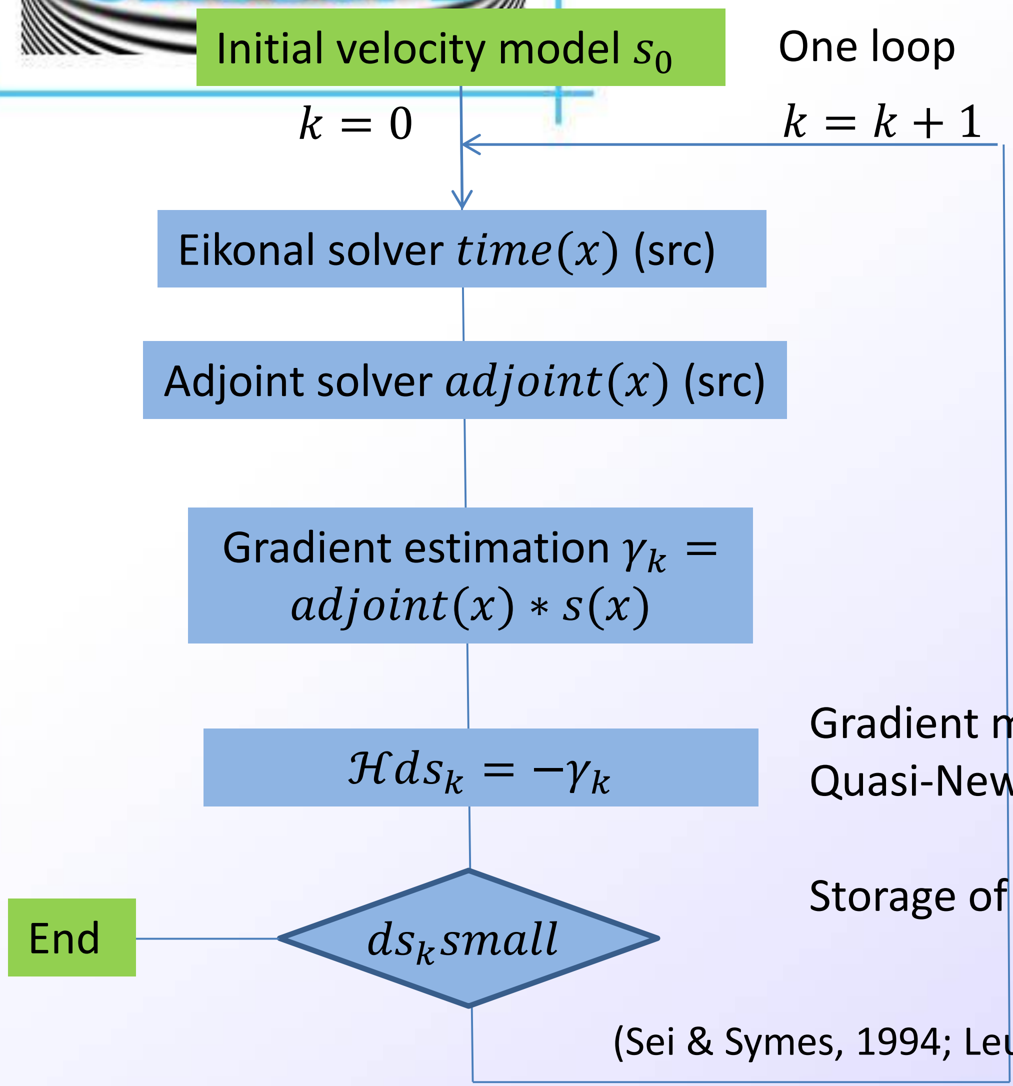
$$\mathbb{C}(s) = \frac{1}{2} dt^t dt$$

- First loop over models : for current model s_k (iteration k)
  - Forward problem
    A. Solve eikonal equation
    B. Compute rays and synthetic travel-times
  - Build sensitivity matrix J_k and delayed times dt_k
- Second loop over linear system: iteratively solve $J_k ds_k = dt_k$, ie $J_k{}^t J_k ds_k = J_k{}^t dt_k$ using conjugate gradient (LSQR, for example)

Update the model $s_{k+1} = s_k + ds_k$

No explicit evaluation of the Hessian $J_k{}^t J_k$:
only products « $J\ ds$ » and « $J^t dt$ » are required in LSQR algorithm
Paige & Sanders (1982)

Linearized Inverse Problem



One-loop procedure

Gradient methods/
Quasi-Newton methods such as l-BFGS method

Storage of previous gradients and models

(Sei & Symes, 1994; Leung & Qian, 2006; Taillandier et al, 2009)

Linearized Inverse Problem

Misfit function

$$\mathbb{C}(s) = \frac{1}{2} dt^t dt$$

One-loop procedure

Complexity: $\mathcal{O}(2 * N_{src} * m * \ln(m))$ for eikonal/adjoint equations

and $\mathcal{O}(m * N_{src})$ for gradient estimation

$(N_{src} \Leftrightarrow N_{rec} \text{ reciprocity})$

Reminder: two-loops procedure

Complexity: $\mathcal{O}(N_{src} * m * \ln(m))$ for eikonal formulation

$\mathcal{O}(k * m * N_{src} * N_{rec})$ for back-ray tracing

$\mathcal{O}(k * m * N_{src} * N_{rec})$ for storage of sensitivity matrix J

Forward and adjoint problems

Forward problem based on Eikonal equation

$$|\nabla T(x,y,z)| = s(x,y,z)$$

$T(x,y,z)$ is the travel time field

Adjoint problem based on Transport equation

$$\nabla.\left(\lambda(x,y,z)\nabla T(x,y,z)\right) = -\sum_r W_r\, \delta t_r$$

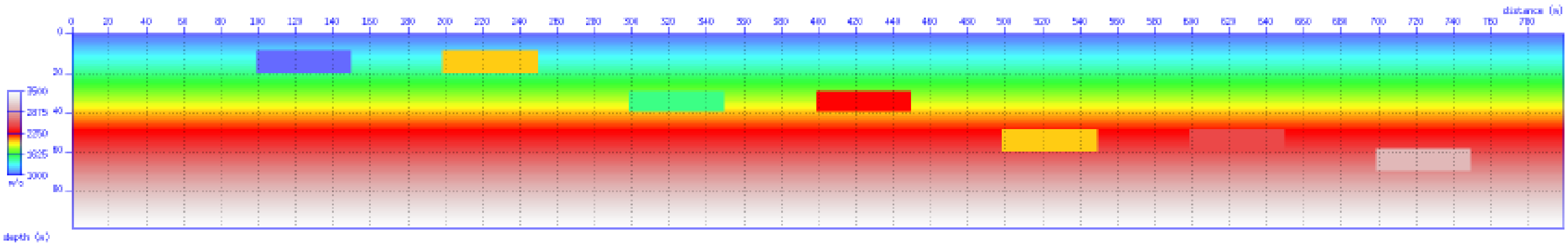
$\lambda(x,y,z)$ is the adjoint field

Non-linear PDE: forward

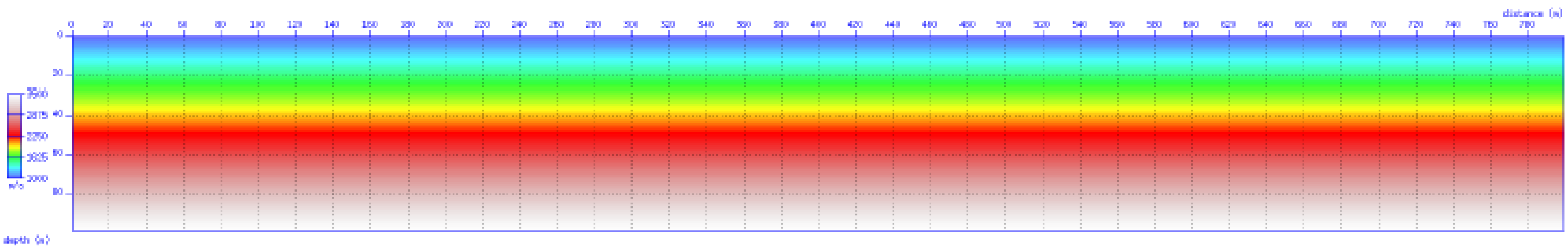
Linear PDE: adjoint

Reflection configuration

Synthetic example:



(a) Exact model V_p

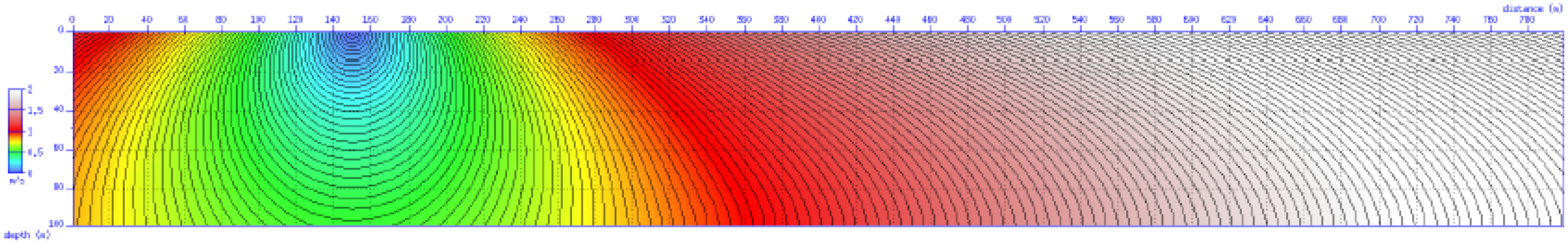


(b) Initial model V_0

Sources and receivers at the free surface

Eikonal solver

Synthetic example:



(c) Traveltime map

Forward problem: the eikonal equation

$$|\nabla \tau| = s$$

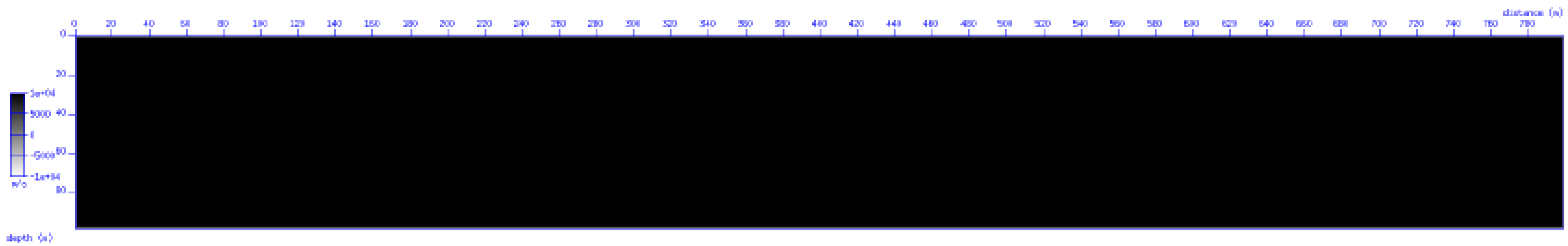
Good accuracy is needed to compute traveltime residuals (and raypaths).

[Podvin and Lecomte, 1991, Zhao, 2005, Li et al., 2008, Fomel et al., 2009, Belayouni, 2012]

(Noble et al, 2014)

Adjoint solver

Synthetic example:



(d) λ initial

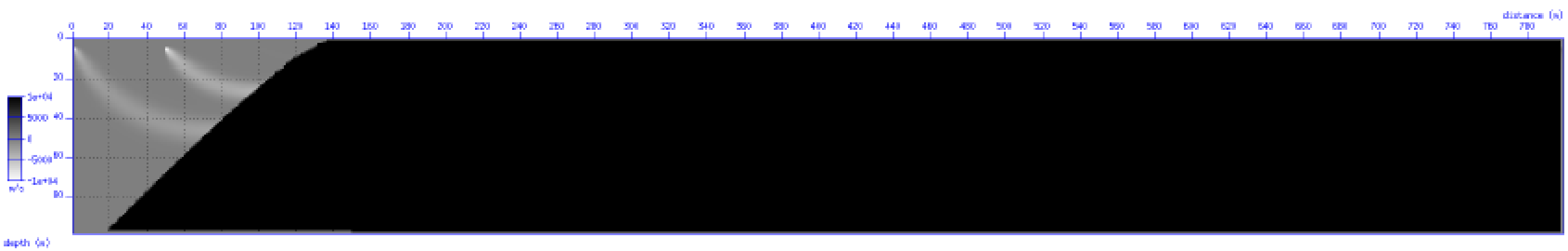
Adjoint problem: a transport equation

$$\nabla(\lambda\nabla\tau) = -\sum_r \mathcal{R}_r^t \delta t_r$$

We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Adjoint solver

Synthetic example:



(e) λ sweep 1 : DOWN-RIGHT

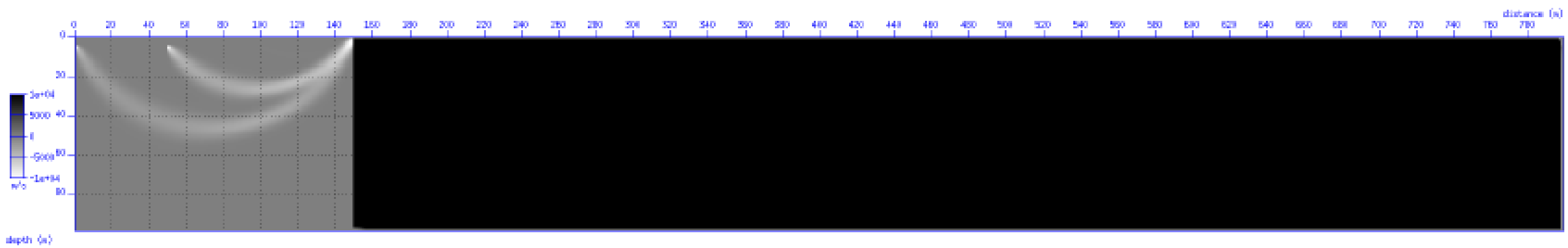
Adjoint problem: a transport equation

$$\nabla(\lambda \nabla \tau) = -\sum_r \mathcal{R}_r^t \, \delta t_r$$

We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Adjoint solver

Synthetic example:



(f) λ sweep 2 : UP-RIGHT

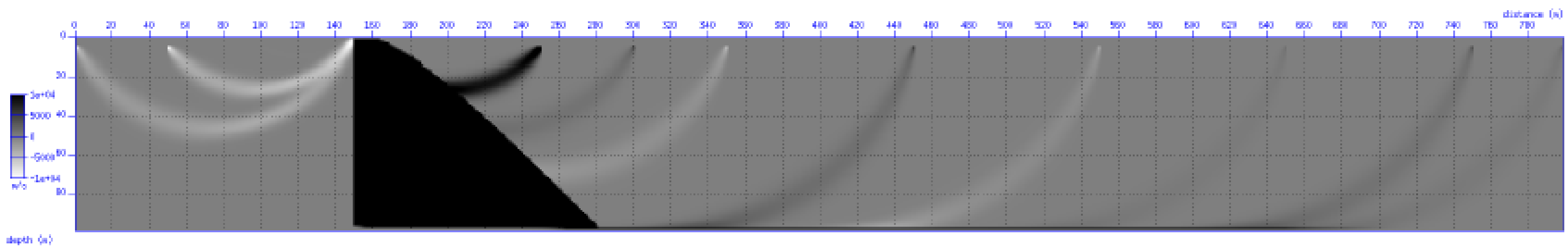
Adjoint problem: a transport equation

$$\nabla(\lambda \nabla \tau) \quad = \quad -\sum_r \mathcal{R}_r^t \, \delta t_r$$

We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Adjoint solver

Synthetic example:



(g) λ sweep 3 : DOWN-LEFT

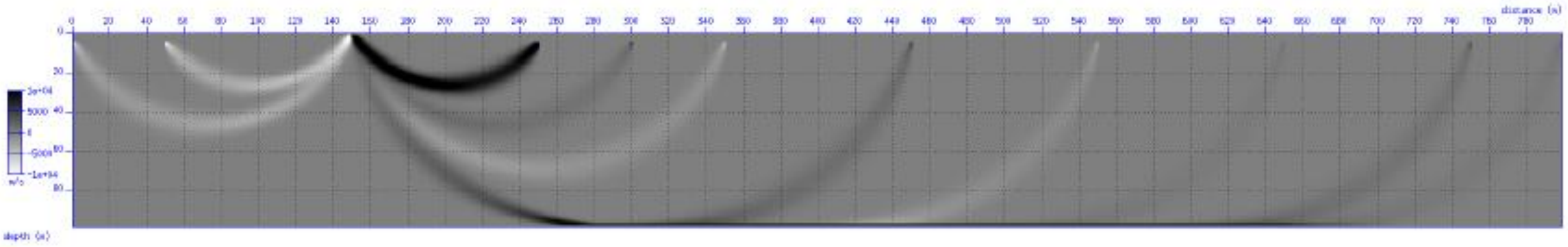
Adjoint problem: a transport equation

$$\nabla(\lambda \nabla \tau) = -\sum_r \mathcal{R}_r^t \, \delta t_r$$

We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Adjoint solver

Synthetic example:



(h) λ sweep 4 : UP-LEFT

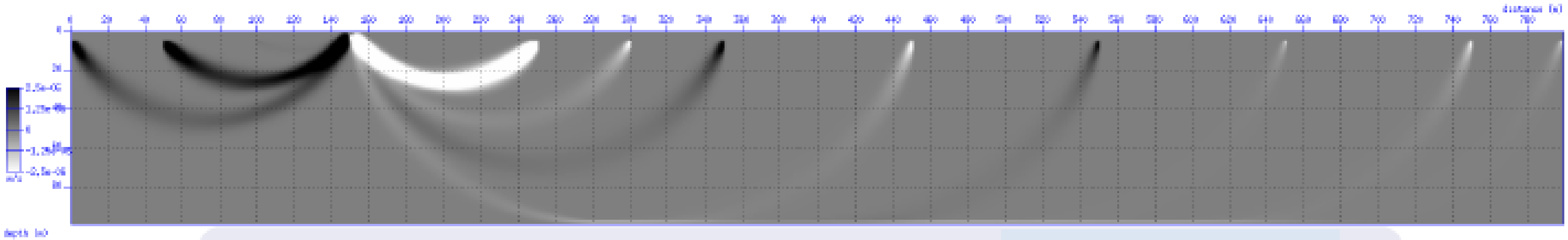
Adjoint problem: a transport equation

$$\nabla(\lambda\nabla\tau) = -\sum_r \mathcal{R}_r^t \delta t_r$$

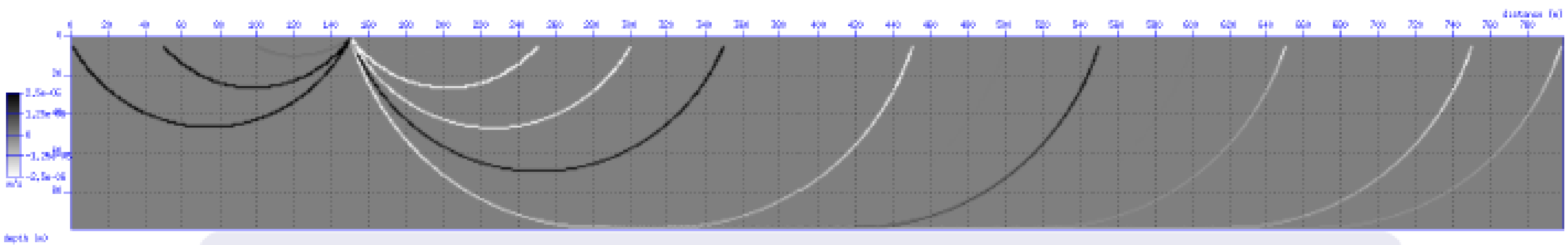
We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Single source gradient

Gradient for a single source (very fine grid ~ to the one for forward modeling)

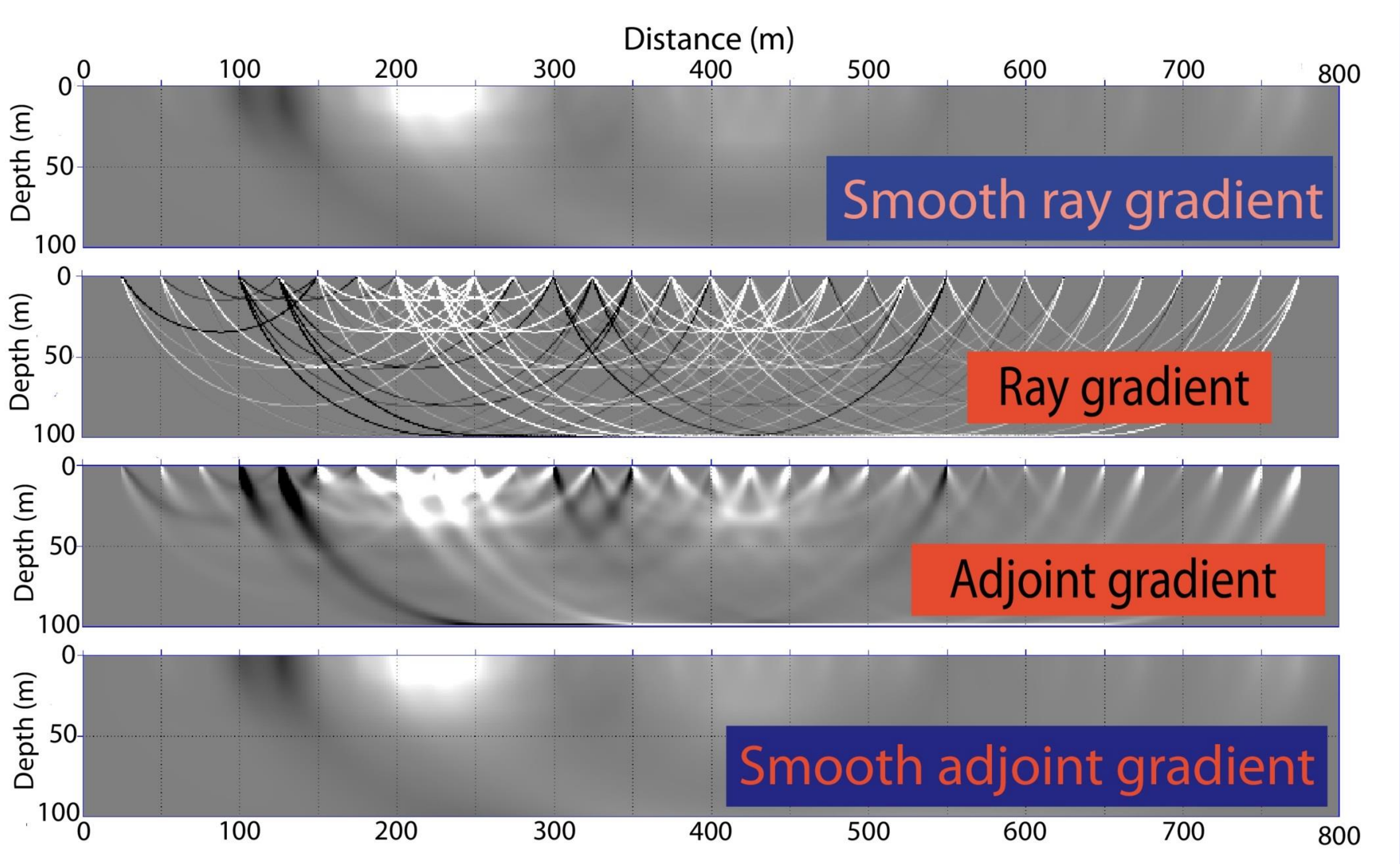


Gradient with adjoint-state: $\gamma = \lambda\, s$



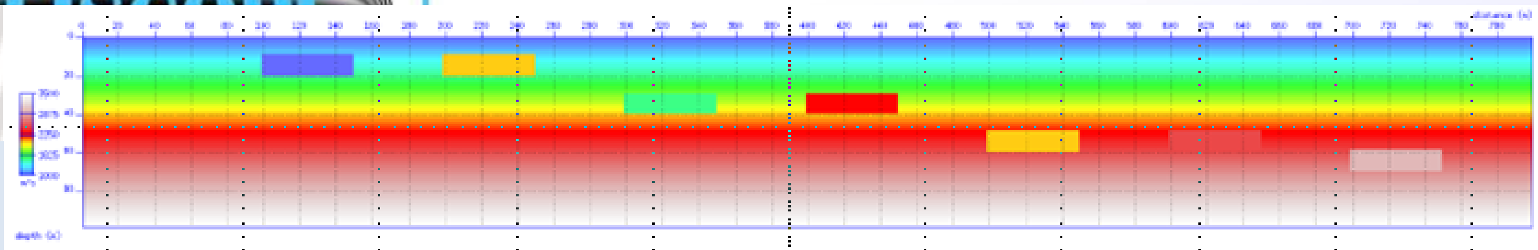
Gradient with residuals projection along back-traced rays: $\gamma = J^t \delta t$

Gradient comparison

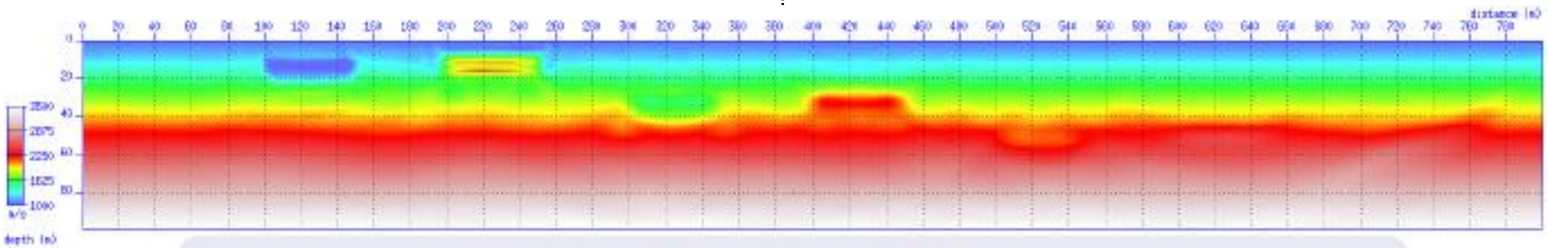


Adjoint versus F-ray inversion

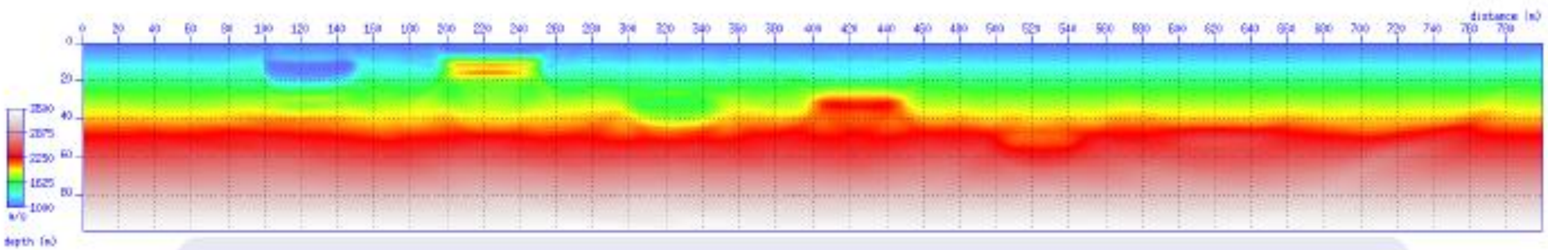
Comparison between adjoint and frozen-ray approach



(a) Exact model V_p



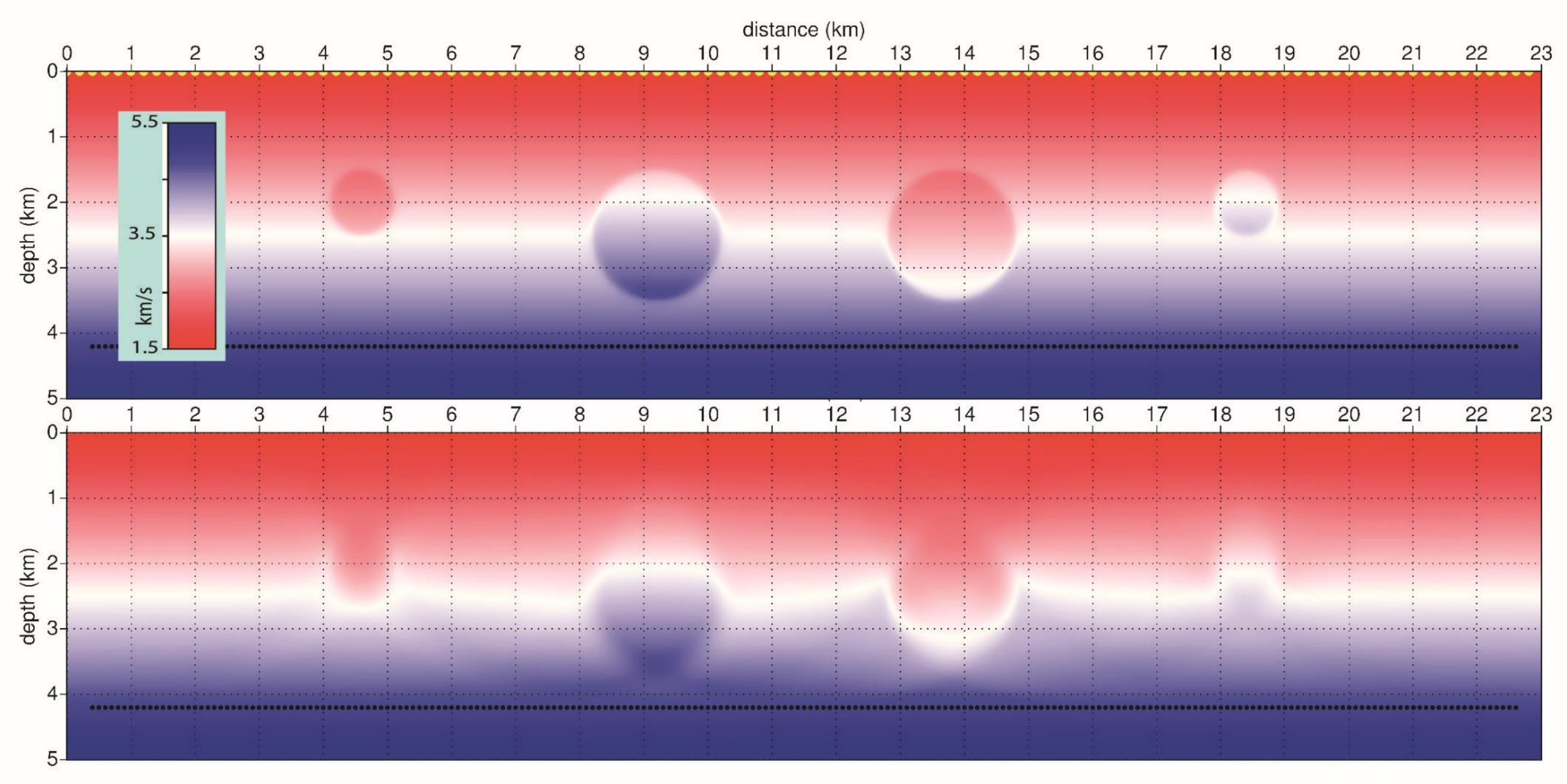
Final result with adjoint-state



Final result with back-traced rays

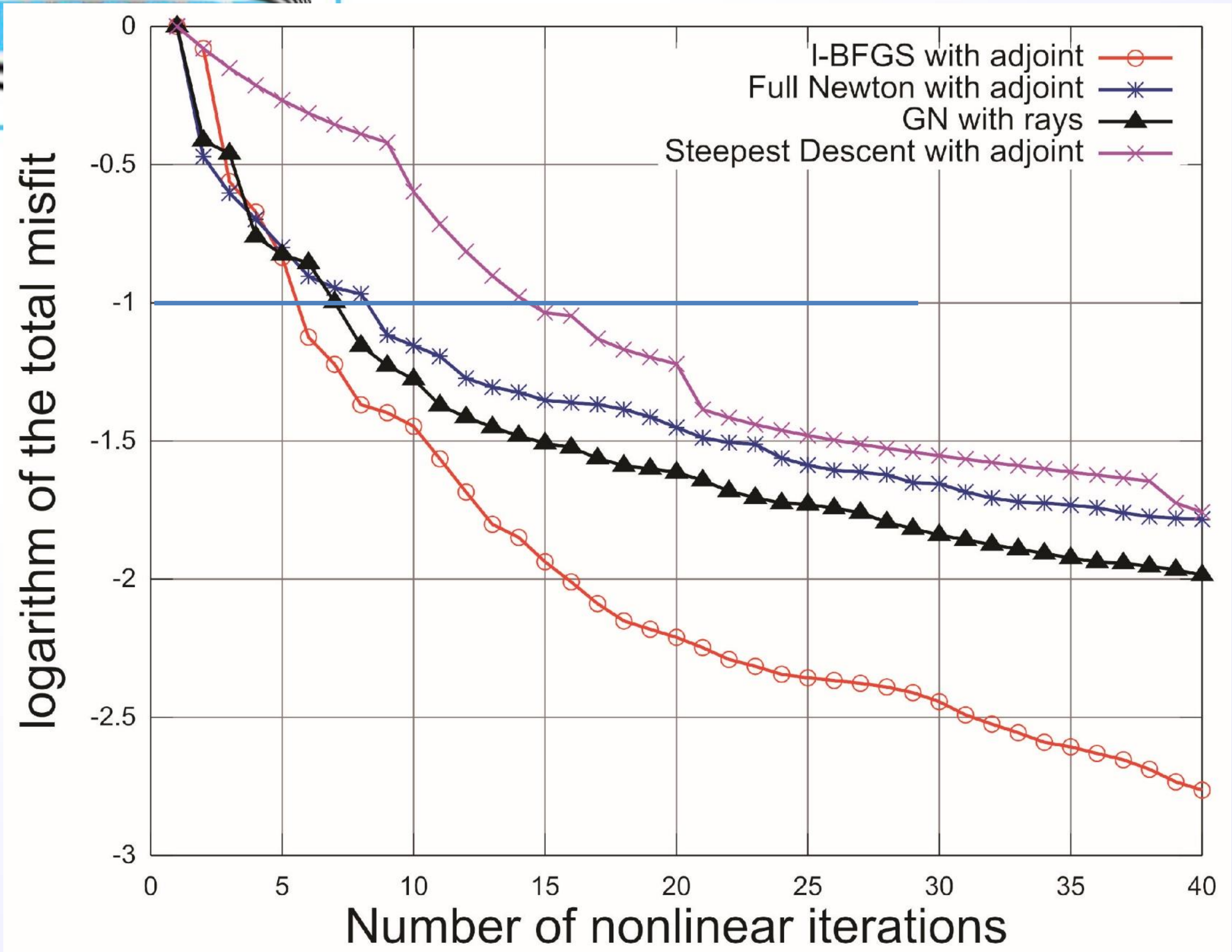
BOTH METHODS GIVE SIMILAR RESULTS FOR MONOPARAMETER INVERSION

Transmission synthetic example



l-GFGS, PCG, PTCN optimization provide almost same results

Transmission synthetic example



Partial Conclusion

- Travel-time tomography could be efficiently performed using adjoint approach and l-BFGS optimization – our best choice -

    A. One-loop procedure
    B. Low complexity increase
    C. Better regularization tuning
    (as the sensitivity matrix derivative is not explicitely evaluated)

- Requirement of accurate Eikonal solvers: on-going work in collaboration with Mines-Paris Tech group (H. Chauris & M. Noble)

- Reduced complexity when considering adjoint formulation: independent of the number of receivers (or sources)

- Joint event-velocity multi-parameter inversion is expected to require the Newton method for proper scaling: on-going work.

Linearized Inverse Problem

Two-loops procedure

Misfit function

$$\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$$

- First loop over models : for current model s_k (iteration k)

  - Forward problem
    A. Solve eikonal equation
    B. Compute rays and synthetic travel-times

  - Build Fréchet derivative matrix J_k and delayed times δt_k

- Second loop over linear system: iteratively solve $J_k \delta s_k = \delta t_k$, ie ${J_k}^t J_k \delta s_k = \delta t_k$ using conjugate gradient (LSQR, for example)

Update the model $s_{k+1} = s_k + \delta s_k$

No explicit evaluation of the Hessian ${J_k}^t J_k$:
only products « $J\ \delta s$ » and « $J^t \delta t$ » are required in LSQR algorithm
Paige & Sanders (1982)

Linearized Inverse Problem

Misfit function

Two-loops procedure

$$\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$$

- First loop over models : for current model s_k (iteration k)
  - Forward problem
    - A. Solve eikonal equation
    - B. Compute rays and synthetic travel-times
  - Build Fréchet derivative matrix J_k and delayed times δt_k
- Second loop over linear system: iteratively solve $J_k \delta s_k = \delta t_k$, ie $J_k{}^t J_k \delta s_k = \delta t_k$ using conjugate gradient (LSQR, for example)

Update the model $s_{k+1} = s_k + \delta s_k$

Complexity: $\mathcal{O}(N_{src} * N_{rec})$ for forward/adjoint formulation and

$\mathcal{O}(N_m * N_{src} * N_{rec} * k)$ for gradient storage (k: sparsity)

Linearized Inverse Problem

Misfit function

One-loop procedure

$$\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$$

- First loop over models : for current model s_k (iteration k)
  - Forward problem
    - A. Solve eikonal equation
    - B. Compute synthetic travel-times and delayed times δt_k
  - Adjoint problem
    - A. Solve adjoint equation
    - B. Build gradient γ_k
- Descent method (one-step gradient or conjugate gradient or l-BFGS) δs_k

Update the model $s_{k+1} = s_k + \alpha_k \delta s_k$
(Sei & Symes, 1994; Leung & Qian, 2006; Taillandier et al, 2009)

No explicit evaluation of the Hessian
l-BFGS provides an evaluation of the H^{-1} from previous stored gradients

Linearized Inverse Problem

Misfit function

$$\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$$

One-loop procedure

- First loop over models : for current model s_k (iteration k)
  - Forward problem
    A. Solve eikonal equation
    B. Compute synthetic travel-times and delayed times δt_k
  - Adjoint problem
    A. Solve adjoint equation
    B. Build gradient γ_k
- Descent method (one-step gradient or conjugate gradient or l-BFGS) δs_k

Update the model $s_{k+1} = s_k + \alpha_k \delta s_k$
(Sei & Symes, 1994; Leung & Qian, 2006; Taillandier et al, 2009)

Complexity: $\mathcal{O}(N_{src})$ for forward/adjoint formulation and $\mathcal{O}(N_m)$ for gradient

Linearized Inverse Problem

Misfit function

$$\mathbb{C}(s) = \frac{1}{2}\delta t^t \delta t$$

Two-loop procedure

- First loop over models : for current model s_k (iteration k)
  - Forward problem
    A. Solve eikonal equation
    B. Compute synthetic travel-times and delayed times δt_k
  - Adjoint problem
    A. Solve adjoint equation
    B. Build gradient γ_k
- Second loop over linear system : $H_k \, \delta s_k = -\gamma_k$
  - Truncated Newton method based on iterative conjugate gradient
  (matrix free approach for product $H_k \, v$, thanks to second-order adjoint formulation)

Update the model $s_{k+1} = s_k + \alpha_k \delta s_k$ (Métivier et al, 2013)

Full Hessian impact $H = J^t J + \frac{\partial J}{\partial s}\delta t$ is evaluated

Partial conclusion

Joint inversion of velocity structure and hypocenter parameters presents strong trade-off.

Estimation of the inverse Hessian operator is quite important for mitigating them.

3D real applications will illustrate the importance of considering Hessian effects.