SES3D.inv

PART II

FRÉCHET KERNELS AND THE ADJOINT METHOD

1. Setup of the tomographic problem: Why gradients?
2. The adjoint method
3. Practical
4. Special topics (source imaging and time reversal)

Setup of the tomographic problem:

Why gradients?

Problem setup

- Find an Earth model **m** such that a suitably defined misfit X is minimal.
- The number of model parameters and the numerical cost of the forward problem prevent the application of probabilistic methods.
- The minimisation proceeds iteratively:

1. Start from initial Earth model $\mathbf{m}_0$
2. Update according to $\mathrm{m}_{i+1} = \mathrm{m}_i + \underbrace{\gamma_i}_{\text{step length}} \underbrace{\mathrm{h}_i}_{\text{descent direction}}$ with $X(m_{i+1}) < X(m_i)$ 1.1

Problem setup

1. Start from initial Earth model $\mathbf{m}_0$

2. Update according to $\mathrm{m}_{i+1} = \mathrm{m}_i + \gamma_i \, \mathrm{h}_i$ with $X(m_{i+1}) < X(m_i)$ 1.1

step length — descent direction

$$\boxed{\mathrm{h}_\mathrm{i} \propto -\frac{\partial \mathrm{X}}{\partial \mathrm{m}}}$$ 1.2

<u>The family of gradient methods:</u>

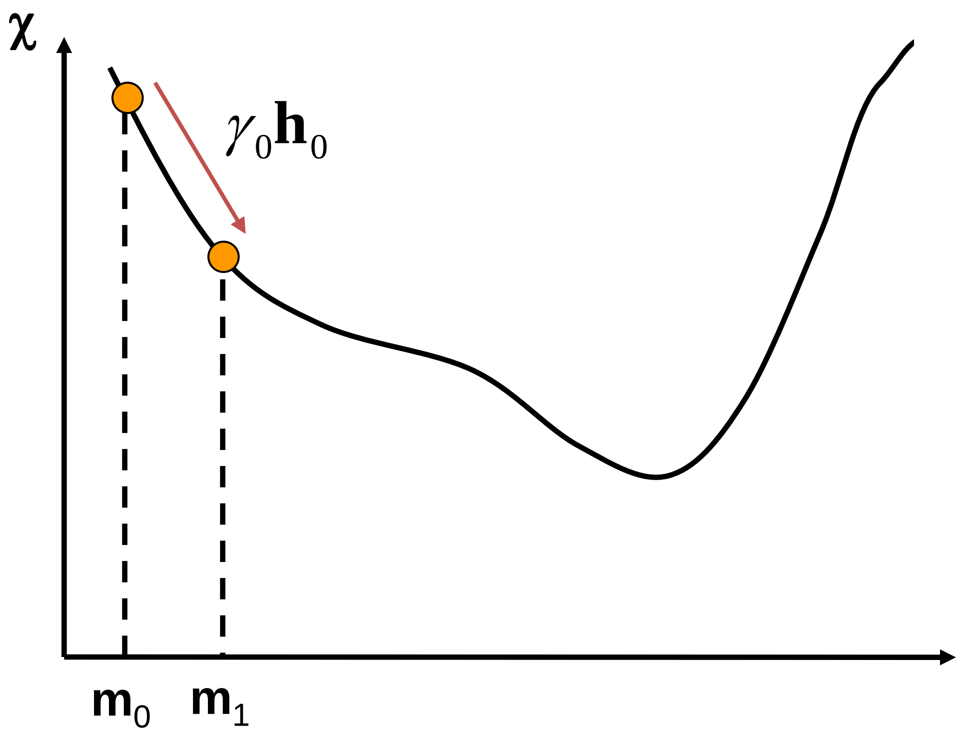
- **method of steepst descent:** $\mathrm{h}_\mathrm{i} = -\partial \mathrm{X}/\partial \mathrm{m}$
- **conjugate-gradient methods**
- **Newton and Newton-like methods**
- **variable-metric methods**
- **…**

Problem setup

1. Start from initial Earth model $\mathbf{m}_0$

2. Update according to $\mathrm{m}_{i+1} = \mathrm{m}_i + \gamma_i \, \mathrm{h}_i$ with $\mathrm{X}(m_{i+1}) < \mathrm{X}(m_i)$ 1.1

step length — descent direction



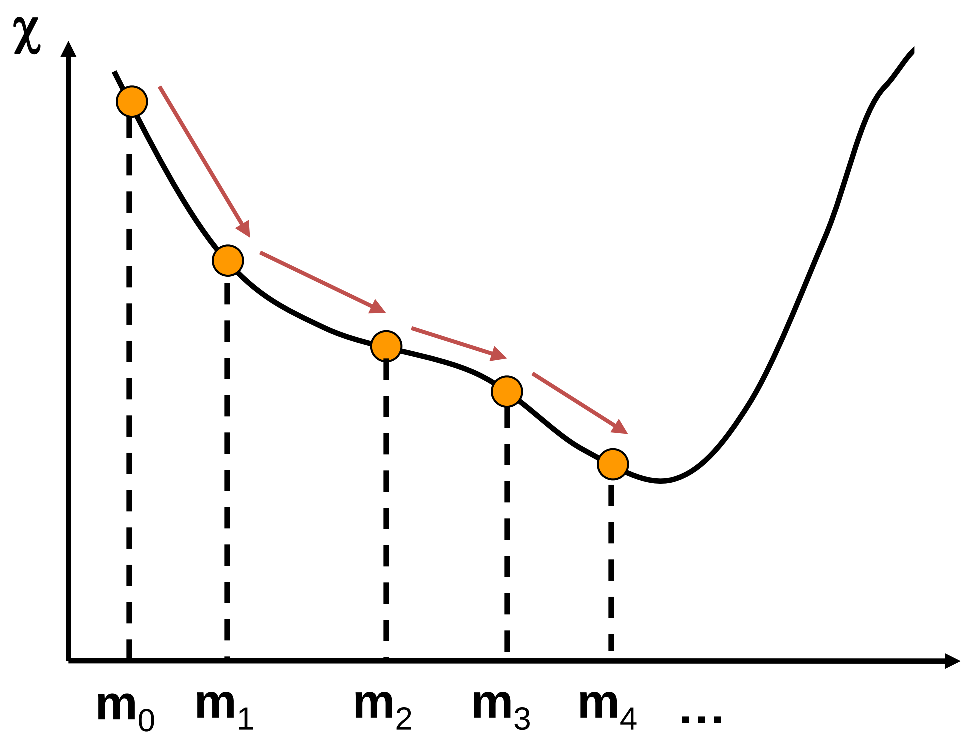
Problem setup

1. Start from initial Earth model $\mathbf{m}_0$

2. Update according to $$\mathrm{m}_{i+1} = \mathrm{m}_i + \gamma_i \, \mathrm{h}_i \quad \text{with} \quad X(m_{i+1}) < X(m_i) \tag{1.1}$$

step length

descent direction



Iteratively approach the minimum misfit by following the local descent directions.

Problem setup

- The full gradient – with all its components - is needed in each iteration.
- The most straightforward approach: approximate the gradient by finite-differences:

$$\boxed{\frac{\partial X(m)}{\partial m_k} \approx \frac{X(\ldots, m_k + h, \ldots) - X(\ldots, m_k, \ldots)}{h}} \qquad 1.3$$

- Example with 500,000 model parameters:

	500,001 forward simulations
x	0.5 h per simulation
x	126 processors
x	50 earthquakes
x	4 simulations per conjugate gradient iteration
x	10 conjugate gradient iterations
	$6.3e^{10}$ cpu hours ≈ 720,000 cpu years

Problem setup

SOLUTIONS:

- **Automatic differentiation** (AD, www.autodiff.org)

  - Differentiation of computer programmes
  - Automatic but inefficient because ignorant about physics

- **Adjoint method**

  - Developed in optimal control theory (J.-L- Lions, 1960s)
  - Express gradients in terms of forward and adjoint fields
  - Can be a bit tedious, but is full of interesting physics.

The adjoint method

1. Prelude: Misfits and Fréchet derivatives
2. The adjoint trick
3. Examples
4. The adjoint wave equation

Prelude I: Fréchet and classical derivatives

- The adjoint method can be derived in many different ways:

    - Born approximation (e.g. Tarantola)
    - Lagrange multipliers (e.g. Liu & Tromp)
    - Data assimilation (e.g. Chen)
    - **Operator approach**

        - elegant and structured
        - very general
        - leads to simple recipes

        - requires a bit more mathematical machinery

Prelude I: Fréchet and classical derivatives

- The Earth model m(x) is a continuously defined function.
- Examples: $\rho(x)$, $v_p(x)$, $v_s(x)$, $Q(x)$, … .
- The Fréchet derivative of the misfit X(m) is the infinitesimal change of X(m) as we pass from Earth model m(x) to m(x)+δm(x):

$$\boxed{\frac{dX}{dm}\delta m := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[X(m+\varepsilon\,\delta m) - X(m)\right]} \tag{2.1}$$

The adjoint method

Prelude I: Fréchet and classical derivatives

- The Earth model m(x) is a continuously defined function.
- Examples: ρ(x), $v_p(x)$, $v_s(x)$, Q(x), … .
- The Fréchet derivative of the misfit X(m) is the infinitesimal change of X(m) as we pass from Earth model m(x) to m(x)+δm(x):

$$\boxed{\frac{dX}{dm}\delta m := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[X(m+\varepsilon\,\delta m) - X(m)\right]} \qquad 2.1$$

- For practical reasons, the continuous Earth model is parameterised in terms of a finite number of basis functions $b_i(x)$:

$$\boxed{m(x) = \sum_{i=1}^{N} m_i\, b_i(x)} \qquad 2.2$$

The adjoint method

Prelude I: Fréchet and classical derivatives

- The misfit X is then a function of the discrete model parameters m_i:

$$X(m) = X[m_1 b_1(x) + \ldots + m_i b_i(x) + \ldots + m_N b_N(x)] \qquad 2.3$$

- The classical partial derivatives with respect to the model parameters m_i can be expressed in terms of the Fréchet derivative:

$$\begin{aligned}\frac{dX}{dm_i} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [&X(m_1 b_1 + \ldots + (m_i + \varepsilon) b_i + \ldots + m_N b_N) \\ &- X(m_1 b_1 + \ldots + \qquad m_i \quad b_i + \ldots + m_N b_N)]\end{aligned} \qquad 2.4$$

$$\frac{dX}{dm_i} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[X(m + \varepsilon b_i) - X(m) \right] = \frac{dX}{dm} b_i \qquad 2.5$$

The adjoint method

Prelude I: Fréchet and classical derivatives

$$\boxed{\frac{dX}{dm_i} = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[X(m+\varepsilon b_i) - X(m)\right] = \frac{dX}{dm} b_i} \qquad 2.5$$

- The classical partial derivative with respect to the discrete model parameter m_i is equal to the Fréchet derivative in the direction of the basis function $b_i(x)$.

- We need to find an efficient way to compute the Fréchet derivative of X for any direction (Earth model perturbation).

The adjoint method

Prelude II: Misfit functionals

- The classical (though not very useful) misfit functional in full waveform inversion is:

$$\mathrm{X} = \frac{1}{2}\int \left[u(m; x^r, t) - u_0(x^r, t)\right]^2 dt \qquad 2.6$$

synthetic waveform | receiver | observed waveform

- This can be rewritten as an integral over both time and space:

$$\mathrm{X} = \frac{1}{2}\iint (u - u_0)^2\, \delta(x - x^r)\, dt\, d^3x \qquad 2.7$$

- The Fréchet derivative of the misfit functional is then:

$$\frac{d\mathrm{X}}{dm}\delta m = \iint \delta u \cdot (u - u_0)\,\delta(x - x^r)\, dt\, d^3x, \qquad \delta u = \frac{du}{dm}\delta m \qquad 2.8$$

The adjoint method

Prelude II: Misfit functionals

- The Fréchet derivative of the misfit functional is then:

$$\frac{dX}{dm}\delta m = \iint \delta u \cdot (u-u_0)\,\delta(x-x^r)\,dt\,d^3x \tag{2.8}$$

- In fact, the Fréchet derivative of **any** misfit functional can be written in this form:

$$\boxed{\frac{dX}{dm}\delta m = \iint \delta u \cdot f^*\,dt\,d^3x} \tag{2.9}$$

- In the special case from above:

$$f^* = (u-u_0)\,\delta(x-x^r) \tag{2.10}$$

- But f* can be anything, depending on the misfit you choose.

The adjoint method

Fréchet derivative of the misfit functional: Problem statement

$$\boxed{\frac{dX}{dm}\delta m = \iint \delta u \cdot f^{*}\, dt\, d^{3}x, \quad with \quad \delta u = \frac{du}{dm}\delta m} \qquad 2.9$$

- The practical difficulty lies in the presence of the Fréchet derivative of the wave field u, that we need to know for **any** model perturbation δm.
- **The adjoint method** is a trick that allows us to eliminiate δu from equation 2.10.
- **The price to pay** is the solution of another differential equation – the **adjoint equation**.

The trick

$$\frac{dX}{dm}\delta m = \iint \delta u \cdot f^{*}\, dt\, d^{3}x, \quad with \quad \delta u = \frac{du}{dm}\delta m \qquad 2.9$$

- To keep the treatment as general as possible, we write the wave equation in a condensed form:
- For the special case of the 1D wave equation:

$$\begin{aligned} \rho \ddot{u} - \partial_x(\mu \partial_x u) &= f \\ L(u,m) &= f \qquad m = (\rho, \mu) \end{aligned} \qquad 2.11$$

- But L could represent any other wave equation (e.g. 3D anisotropic, visco-elastic, …) or even any other PDE.
- Compute the Fréchet derivative of L in the direction δm:

$$\frac{dL(u,m)}{du}\delta u + \frac{dL(u,m)}{dm}\delta m = 0 \qquad 2.12$$

The trick

$$\frac{dL(u,m)}{du}\delta u + \frac{dL(u,m)}{dm}\delta m = 0 \qquad 2.12$$

- We can simplify the first term when L is linear in u (which is the case for the wave equation):

$$\begin{aligned}\frac{dL(u)}{du}\delta u &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[L(u+\varepsilon\,\delta u)-L(u)\right] \\ &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[L(u)+\varepsilon\,L(\delta u)-L(u)\right]=L(\delta u)\end{aligned} \qquad 2.13$$

- Then going back to equation 2.13:

$$L(\delta u)+\frac{dL(u)}{dm}\delta m = 0 \qquad 2.14$$

- … where the explicit dependence on m is omitted for clarity.

The adjoint method

The trick

$$L(\delta u) + \frac{dL(u)}{dm}\delta m = 0 \qquad 2.14$$

- Now we multiply equation 2.14 with an arbitrary (but differentiable) test function u*(x,t), and integrate over time and space.

$$\iint u^* \cdot \left[L(\delta u) + \frac{dL(u)}{dm}\delta m \right] dt\, d^3x = 0 \qquad 2.15$$

- Adding equation 2.15 to the Fréchet derivative of the misfit

$$\frac{dX}{dm}\delta m = \iint \delta u \cdot f^*\, dt\, d^3x \qquad 2.9$$

gives:

$$\frac{dX}{dm}\delta m = \iint \left[\delta u \cdot f^* + u^* \cdot L(\delta u) + u^* \cdot \frac{dL(u)}{dm}\delta m \right] dt\, d^3x \qquad 2.16$$

The adjoint method

The trick

$$\frac{dX}{dm}\delta m = \iint \left[\delta u \cdot f^{*} + u^{*} \cdot L(\delta u) + u^{*} \cdot \frac{dL(u)}{dm}\delta m \right] dt\, d^{3}x \qquad 2.16$$

- We can eliminate δu from equation 2.16 with the help of adjoint of L:

Find u* and L* such that:

$$\iint u^{*} \cdot L(\delta u)\, dt\, d^{3}x = \iint \delta u \cdot L^{*}(u^{*})\, dt\, d^{3}x \qquad 2.17$$

- Finding the adjoints is the actual challenge, but if we manage to do so, we can transform equation 2.16 to:

$$\frac{dX}{dm}\delta m = \iint \delta u \cdot \left[f^{*} + L^{*}(u^{*}) \right] dt\, d^{3}x + \iint \left[u^{*} \cdot \frac{dL(u)}{dm}\delta m \right] dt\, d^{3}x$$

2.19

The trick

$$\boxed{\frac{dX}{dm}\delta m = \iint \delta u \cdot \left[f^* + L^*(u^*)\right] dt\, d^3x + \iint \left[u^* \cdot \frac{dL(u)}{dm}\delta m\right] dt\, d^3x}$$

2.19

- We can eliminate δu when the **adjoint field** u* satisfies the **adjoint equation**:

$$\boxed{L^*(u^*) = -f^*}$$

2.20

- Computing the Fréchet derivative of the misfit X then becomes quite easy:

$$\boxed{\frac{dX}{dm}\delta m = \iint \left[u^* \cdot \frac{dL(u)}{dm}\delta m\right] dt\, d^3x}$$

2.21

- Everything in equation 2.21 is known.

Examples

- The Fréchet derivative of the misfit X:

$$\boxed{\frac{dX}{dm}\delta m = \iint \left[u^* \cdot \frac{dL(u)}{dm}\delta m \right] dt\, d^3x} \qquad 2.21$$

The adjoint method

Examples

- The Fréchet derivative of the misfit X:

$$\boxed{\frac{dX}{dm}\delta m = \iint \left[u^* \cdot \frac{dL(u)}{dm}\delta m \right] dt\, d^3x} \qquad 2.21$$

- 1D wave equation: $\boxed{L(u) = \rho\ddot{u} - \partial_x(\mu\partial_x u)}$ 2.22

- Derivative of the wave operator:

$$\boxed{\frac{dL(u)}{d\rho}\delta\rho = \delta\rho\,\ddot{u}} \qquad 2.23$$

- Fréchet derivative of X:

$$\boxed{\frac{dX}{d\rho}\delta\rho = \iint \delta\rho\,(u^* \cdot \ddot{u})\,dt\,dx = -\iint \delta\rho\,(\dot{u}^* \cdot \dot{u})\,dt\,dx} \qquad 2.24$$

The adjoint method

Examples

- The Fréchet derivative of the misfit X:

$$\frac{dX}{dm}\delta m = \iint \left[u^* \cdot \frac{dL(u)}{dm}\delta m \right] dt\, d^3x \qquad 2.21$$

- 1D wave equation: $\boxed{L(u) = \rho \ddot{u} - \partial_x(\mu \partial_x u)}$ 2.22

- Derivative of the wave operator:

$$\frac{dL(u)}{d\mu}\delta\mu = -\partial_x(\delta\mu\,\partial_x u) \qquad 2.25$$

- Fréchet derivatives of X:

$$\frac{dX}{d\mu}\delta\mu = -\iint \left[u^* \cdot \partial_x(\delta\mu\,\partial_x u) \right] dt\, dx = \iint \delta\mu\,\partial_x u^* \cdot \partial_x u\, dt\, dx \qquad 2.26$$

The adjoint method

Examples

- Fréchet derivatives of X:

$$\boxed{\frac{dX}{d\rho}\delta\rho = -\iint \delta\rho\,(\dot{u}^* \cdot \dot{u})\,dt\,dx} \qquad 2.25$$

$$\boxed{-\iint \delta\rho\,(\dot{u}^* \cdot \dot{u})\,dt\,dx = \int \delta\rho\,K_\rho\,dx} \qquad 2.25'$$

- Fréchet kernel w.r.t. density:

$$\boxed{K_\rho = -\int \dot{u}^* \cdot \dot{u}\,dt} \qquad 2.25''$$

$$\boxed{\frac{dX}{d\mu}\delta\mu = \iint \delta\mu\,\partial_x u^* \cdot \partial_x u\,dt\,dx} \qquad 2.26$$

$$\boxed{\iint \delta\mu\,\partial_x u^* \cdot \partial_x u\,dt\,dx = \int \delta\mu\,K_\mu\,dx} \qquad 2.26'$$

- Fréchet kernel w.r.t. the shear modulus:

$$\boxed{K_\mu = \int \partial_x u^* \cdot \partial_x u\,dt} \qquad 2.26''$$

- Fréchet kernels are the volumetric densities of the Fréchet derivatives.

The adjoint method

Examples

- For the 3D elastic wave equation:

$$
\boxed{
\begin{aligned}
K_\rho^0 &= -\int_T \dot{\mathbf{u}}^\dagger \cdot \dot{\mathbf{u}}\, dt \\
K_\lambda^0 &= \int_T (\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{u}^\dagger)\, dt \\
K_\mu^0 &= \int_T [(\nabla \mathbf{u}^\dagger) : (\nabla \mathbf{u}) + (\nabla \mathbf{u}^\dagger) : (\nabla \mathbf{u})^T]\, dt
\end{aligned}
}
\qquad 2.27
$$

- … and with a different parameterisation:

$$
\boxed{
\begin{aligned}
K_\rho &= K_\rho^0 + (v_\mathrm{P}^2 - 2v_\mathrm{S}^2)\, K_\lambda^0 + v_\mathrm{S}^2\, K_\mu^0 \\
K_{v_\mathrm{S}} &= 2\rho\, v_\mathrm{S}\, K_\mu^0 - 4\rho\, v_\mathrm{S}\, K_\lambda^0 \\
K_{v_\mathrm{P}} &= 2\rho\, v_\mathrm{P}\, K_\lambda^0
\end{aligned}
}
\qquad 2.28
$$

- Fréchet derivatives depend on the parameterisation!!!

The adjoint method

The adjoint operator for the 1D wave equation

- 1D wave equation: $$\boxed{L(u) = \rho \ddot{u} - \partial_x(\mu \partial_x u)}$$ 3.1

- with initial conditions:

$$\boxed{u\,|_{t=0} = \dot{u}\,|_{t=0} = 0}$$ 3.2

- and boundary conditions:

$$\boxed{\partial_x u\,|_{x=0} = \partial_x u\,|_{x=L} = 0}$$ 3.3

- We need to find the **adjoint field** u* and the **adjoint operator** L* such that the following equation holds for **any** δu that satisfies the boundary and initial conditions:

$$\boxed{\int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot L(\delta u)\,dt\,dx = \int_{x=0}^{L}\int_{t=0}^{T} \delta u \cdot L^*(u^*)\,dt\,dx}$$ 3.4

The adjoint operator for the 1D wave equation

- Find u* and L* such that the following equation holds for **any** δu:

$$\int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot L(\delta u)\,dt\,dx = \int_{x=0}^{L}\int_{t=0}^{T} \delta u \cdot L^*(u^*)\,dt\,dx \tag{3.4}$$

- Writing the left-hand side of 3.4 explicitly:

$$\int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot L(\delta u)\,dt\,dx = \int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot \left[\rho\delta\ddot{u} - \partial_x(\mu\partial_x\delta u)\right]dt\,dx \tag{3.5}$$

- The goal is to eliminate δu from the differentiations on the right-hand side:

$$\begin{aligned}\int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot L(\delta u)\,dt\,dx &= \int_{x=0}^{L} \rho u^* \delta\dot{u}\,dx \Big|_{t=0}^{T} - \int_{t=0}^{T} \mu u^* \partial_x \delta u\,dt \Big|_{x=0}^{L} \\ &\quad - \int_{x=0}^{L}\int_{t=0}^{T} \left[\rho\dot{u}^*\delta\dot{u} - \mu\partial_x u^* \partial_x \delta u\right]dt\,dx\end{aligned} \tag{3.6}$$

The adjoint operator for the 1D wave equation

- The goal is to eliminate δu from the differentiations on the right-hand side:

$$\int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot L(\delta u)\,dt\,dx = \int_{x=0}^{L} \rho u^* \delta\dot{u}\,dx\,\Big|_{t=0}^{T} - \int_{t=0}^{T} \mu u^* \partial_x \delta u\,dt\,\Big|_{x=0}^{L} - \int_{x=0}^{L}\int_{t=0}^{T} \left[\rho \dot{u}^* \delta\dot{u} - \mu \partial_x u^* \partial_x \delta u\right] dt\,dx \qquad 3.6$$

- One more integration by parts:

$$\begin{aligned}\int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot L(\delta u)\,dt\,dx &= \int_{x=0}^{L} \rho u^* \delta\dot{u}\,dx\,\Big|_{t=0}^{T} - \int_{t=0}^{T} \mu u^* \partial_x \delta u\,dt\,\Big|_{x=0}^{L} \\ &\quad - \int_{x=0}^{L} \rho \dot{u}^* \delta u\,dx\Big|_{t=0}^{T} + \int_{t=0}^{T} \mu \partial_x u^* \delta u\,dt\Big|_{x=0}^{L} \\ &\quad + \int_{x=0}^{L}\int_{t=0}^{T} \delta u\left[\rho \ddot{u}^* - \partial_x(\mu \partial_x u^*)\right] dt\,dx \qquad 3.7\end{aligned}$$

The adjoint method

The adjoint operator for the 1D wave equation

- The goal is to eliminate δu from the differentiations on the right-hand side:

$$\int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot L(\delta u)\,dt\,dx = \int_{x=0}^{L} \rho u^* \delta\dot{u}\,dx\,\Big|_{t=0}^{T} - \int_{t=0}^{T} \mu u^* \partial_x \delta u\,dt\,\Big|_{x=0}^{L} - \int_{x=0}^{L}\int_{t=0}^{T} \left[\rho \dot{u}^* \delta\dot{u} - \mu \partial_x u^* \partial_x \delta u\right] dt\,dx \tag{3.6}$$

- One more integration by parts (with boundary & initial conditions accounted for):

$$\int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot L(\delta u)\,dt\,dx = \int_{x=0}^{L} \rho u^* \delta\dot{u}\,dx\,\Big|^{t=T} - 0 - \int_{x=0}^{L} \rho \dot{u}^* \delta u\,dx\,\Big|^{t=T} + \int_{t=0}^{T} \mu \partial_x u^* \delta u\,dt\,\Big|_{x=0}^{L} + \int_{x=0}^{L}\int_{t=0}^{T} \delta u \left[\rho \ddot{u}^* - \partial_x(\mu \partial_x u^*)\right] dt\,dx \tag{3.8}$$

The adjoint method

The adjoint operator for the 1D wave equation

- One more integration by parts (with boundary & initial conditions accounted for):

$$\boxed{\begin{aligned}\int_{x=0}^{L}\int_{t=0}^{T} u^* \cdot L(\delta u)\, dt\, dx = &\int_{x=0}^{L} \rho u^*\, \delta\dot{u}\, dx\,\Big|^{t=T} - 0 \\ &- \int_{x=0}^{L} \rho \dot{u}^*\, \delta u\, dx\Big|^{t=T} + \int_{t=0}^{T} \mu \partial_x u^*\, \delta u\, dt\Big|_{x=0}^{L} \\ &+ \int_{x=0}^{L}\int_{t=0}^{T} \delta u \left[\rho \ddot{u}^* - \partial_x(\mu \partial_x u^*)\right] dt\, dx\end{aligned}} \qquad 3.8$$

- We are done, when we manage to make the boundary terms go away. So, we impose conditions upon the adjoint field u* :

$$\boxed{u^*\,\big|_{t=T} = \dot{u}^*\,\big|_{t=T} = 0} \qquad \text{terminal condition} \qquad 3.9$$

$$\boxed{\partial_x u^*\,\big|_{x=0} = \partial_x u^*\,\big|_{x=L} = 0} \qquad \text{boundary condition (same as for u)} \qquad 3.10$$

The adjoint operator for the 1D wave equation

- So, provided that the adjoint field u* satisfies the terminal and boundary conditions

$$\boxed{u^* \,|_{t=T} = \dot{u}^* \,|_{t=T} = 0} \qquad \text{terminal condition} \qquad 3.9$$

$$\boxed{\partial_x u^* \,|_{x=0} = \partial_x u^* \,|_{x=L} = 0} \qquad \text{boundary condition (same as for u)} \qquad 3.10$$

we are left with

$$\boxed{\int\limits_{x=0}^{L} \int\limits_{t=0}^{T} u^* \cdot L(\delta u)\, dt\, dx = \int\limits_{x=0}^{L} \int\limits_{t=0}^{T} \delta u \left[\rho \ddot{u}^* - \partial_x(\mu \partial_x u^*)\right] dt\, dx} \qquad 3.11$$

- What we initially wanted to have is

$$\boxed{\int\limits_{x=0}^{L} \int\limits_{t=0}^{T} u^* \cdot L(\delta u)\, dt\, dx = \int\limits_{x=0}^{L} \int\limits_{t=0}^{T} \delta u \cdot L^*(u^*)\, dt\, dx} \qquad 3.4$$

The adjoint operator for the 1D wave equation

- It follows that the adjoint operator is given by

$$\boxed{L^*(u^*) = \rho \ddot{u}^* - \partial_x(\mu \partial_x u^*)} \qquad 3.12$$

- This is, fortunately, again a wave equation, meaning that it can be solved with pre-existing computer codes.

The adjoint method

Summary

- The Fréchet derivative of the misfit function:

$$\boxed{\frac{dX}{dm} = \iint \delta u \cdot f^{*}\, dt\, d^{3}x}$$

- The Fréchet derivative of the misfit X can be expressed in terms of the regular wave field u and the adjoint field u*:

$$\boxed{\frac{dX}{dm}\delta m = \iint \left[u^{*} \cdot \frac{dL(u)}{dm}\delta m \right] dt\, d^{3}x}$$

- The adjoint field is governed by the adjoint equation:

$$\boxed{L^{*}(u^{*}) = -f^{*}}$$

- The adjoint field satisfies terminal conditions:

$$\boxed{u^{*}\,|_{t=T} = \dot{u}^{*}\,|_{t=T} = 0}$$

The adjoint method

Summary

- The misfit can be expressed as an integral over time and space:

$$\frac{dX}{dm} = \iint \delta u \cdot f^* \, dt \, d^3x$$

$$f^* = (u - u_0)\,\delta(x - x^r)$$

L2 norm

- The Fréchet derivative of the misfit X can be expressed in terms of the regular wave field u and the adjoint field u*:

$$\frac{dX}{dm}\delta m = \iint \left[u^* \cdot \frac{dL(u)}{dm}\delta m \right] dt \, d^3x$$

1D wave equation

$$\frac{dX}{d\rho}\delta\rho = -\iint \delta\rho\,(\dot{u}^* \cdot \dot{u})\, dt\, dx$$

- The adjoint field is governed by the adjoint equation:

$$L^*(u^*) = -f^*$$

$$\rho \ddot{u}^* - \partial_x(\mu \partial_x u^*) = -f^*$$

1D wave equation

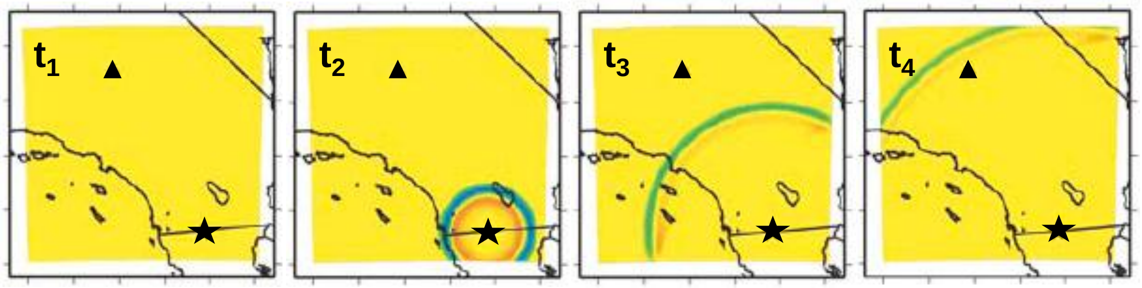
- The adjoint field satisfies terminal conditions:

$$u^*\,|_{t=T} = \dot{u}^*\,|_{t=T} = 0$$

The adjoint method

The recipe

1. Solve the forward problem



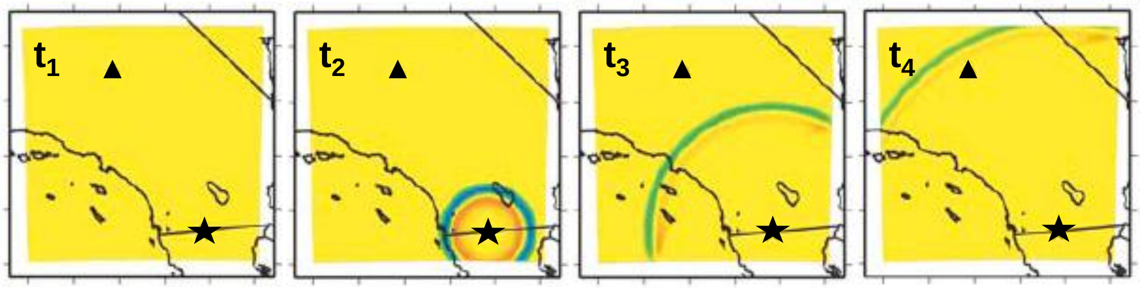
forward field u
synthetic seismograms

Tape et al., 2007

The adjoint method

The recipe

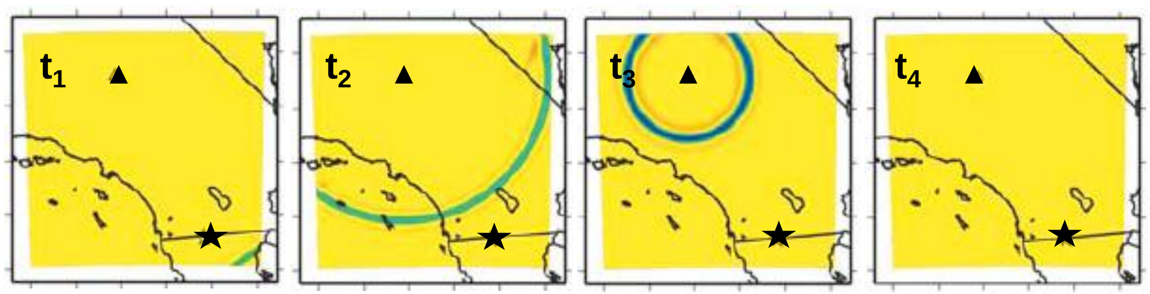
1. Solve the forward problem



➡ forward field u
synthetic seismograms

2. Evaluate the misfit X and compute the adjoint force

3. Solve the adjoint problem



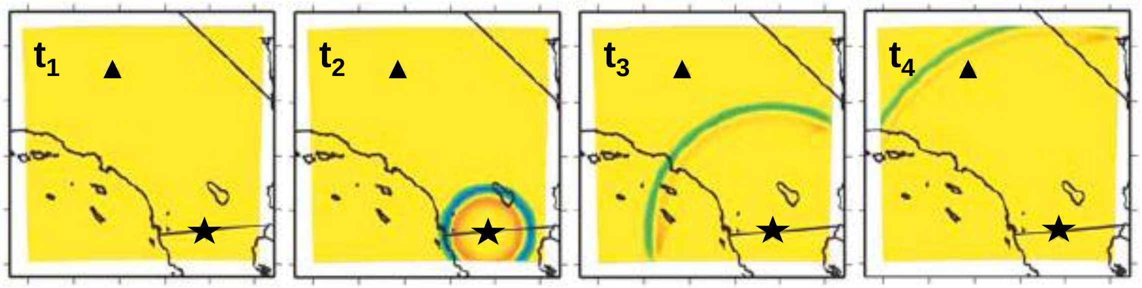
➡ adjoint field u*

Tape et al., 2007

The adjoint method

The recipe

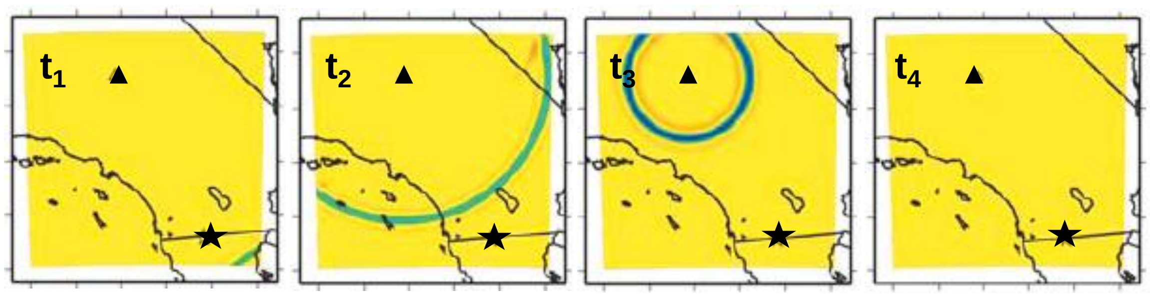
1. Solve the forward problem



➡ forward field u
synthetic seismograms

2. Evaluate the misfit X and compute the adjoint force

3. Solve the adjoint problem



➡ adjoint field u*

4. Compute the Fréchet kernels by integrating u and u^*

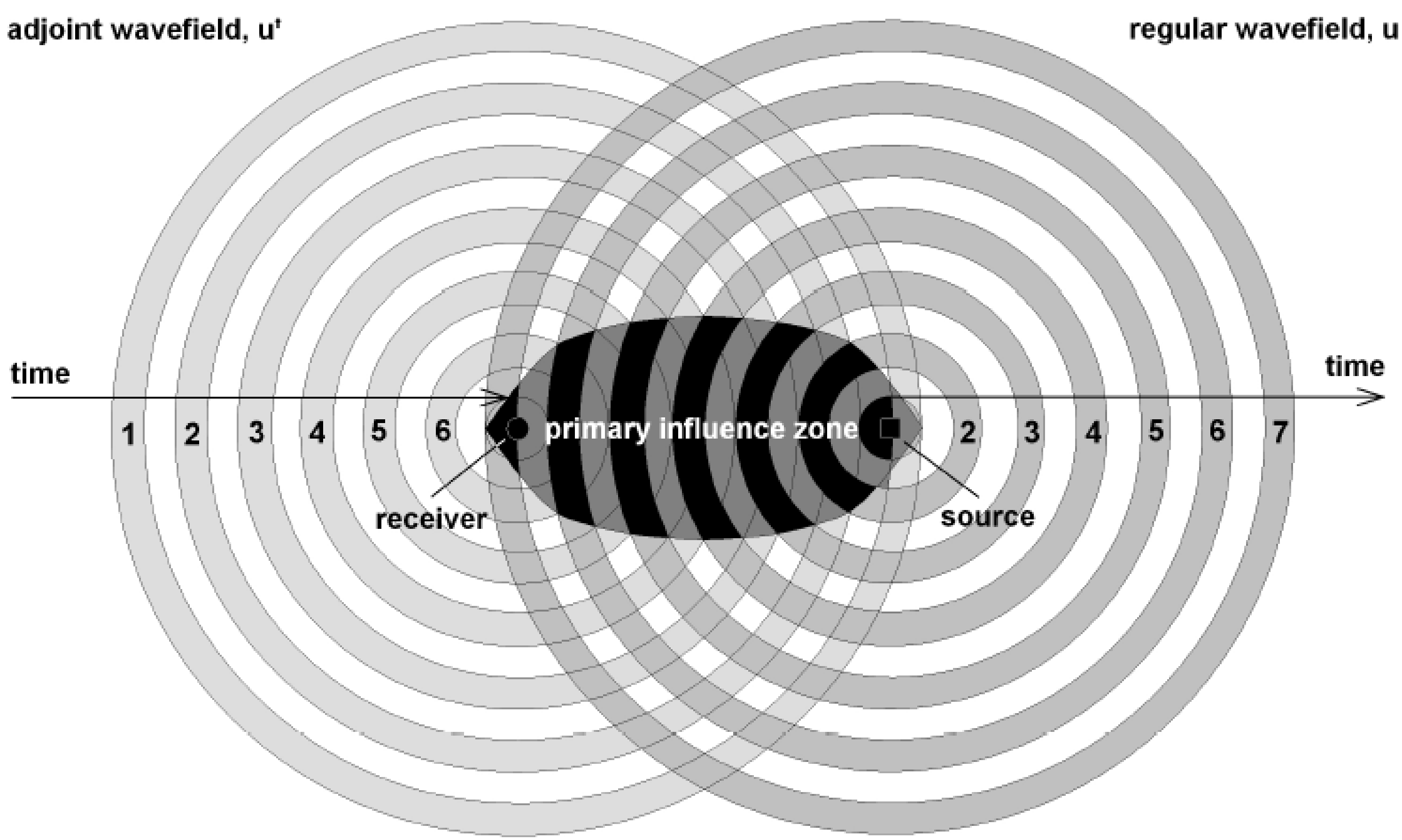
… for instance: $$K_\rho = -\int \dot{u}^* \cdot \dot{u}\, dt$$

Tape et al., 2007

The adjoint method

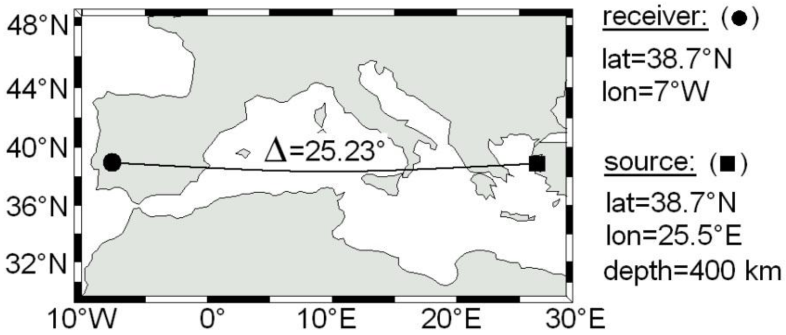
The recipe

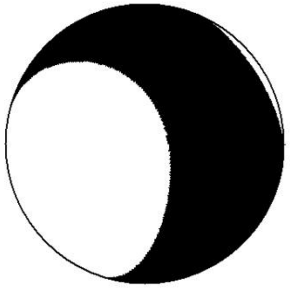
- The interaction of the regular and the adjoint fields generates a primary influence zone.
- First-order scattering from within the primary influence zone affects the measurement.

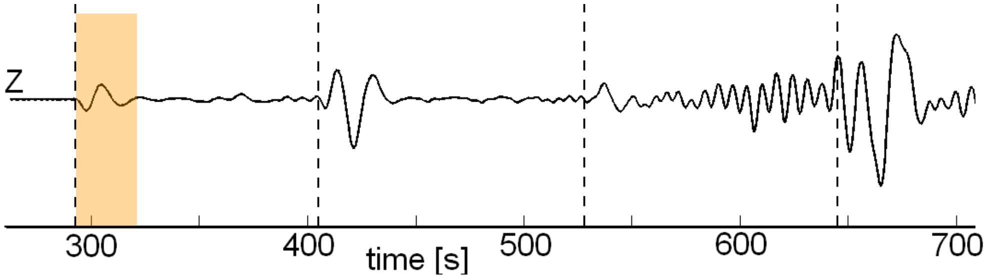


The adjoint method

Fréchet kernel gallery



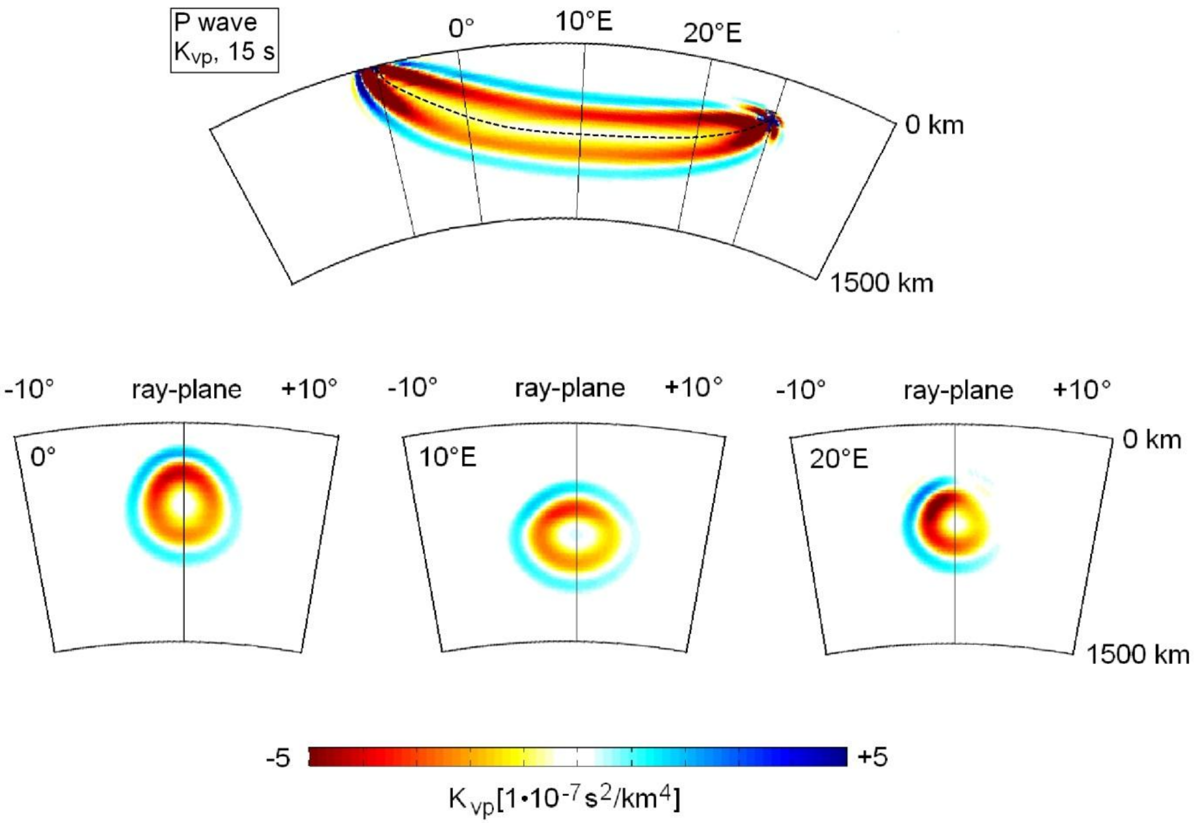




measurement: cross-correlation time shift

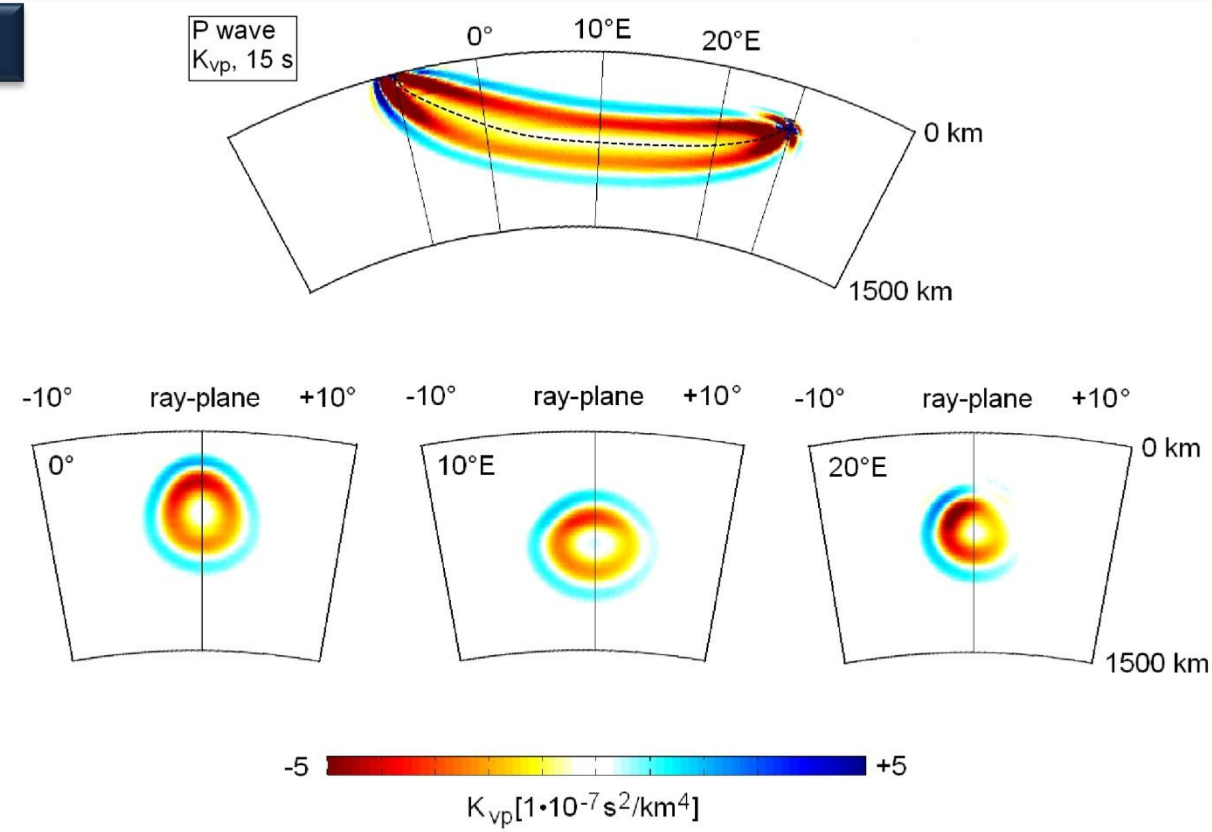
The adjoint method

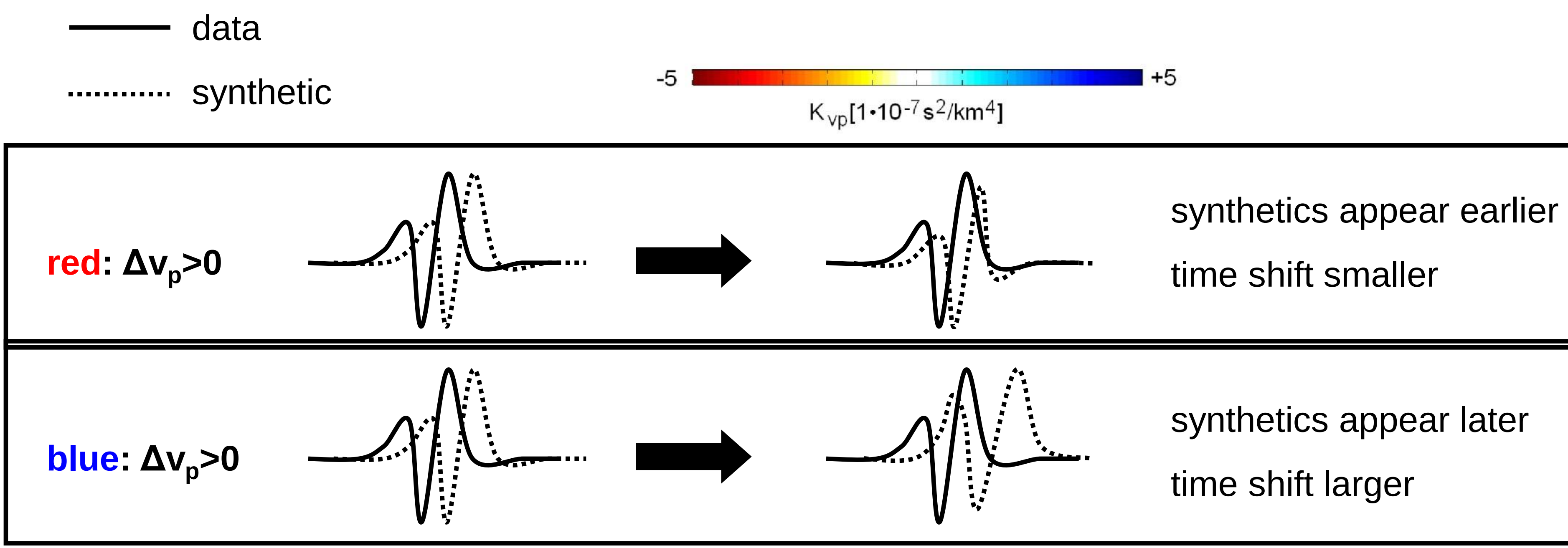
Fréchet kernel gallery



The adjoint method

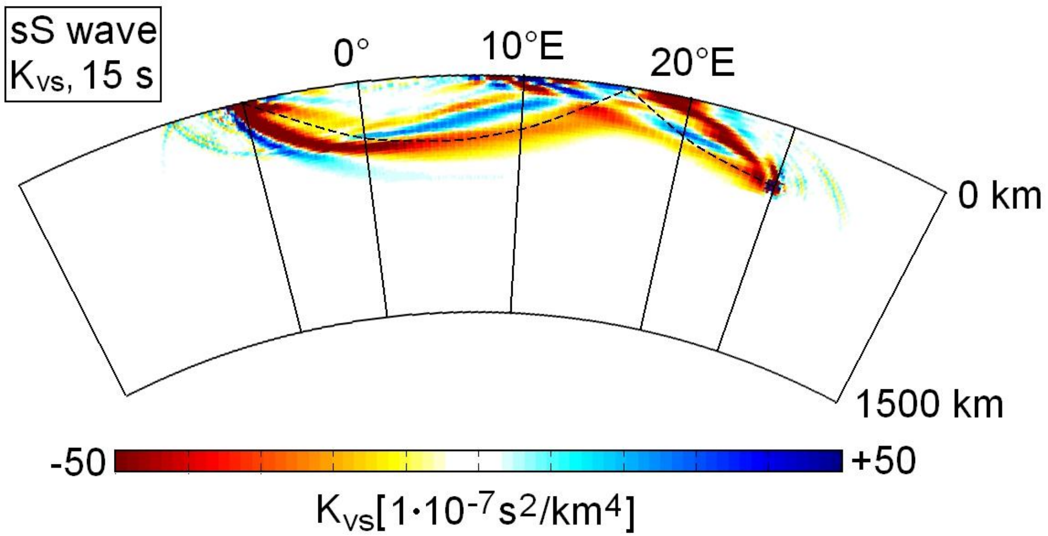
Fréchet kernel gallery

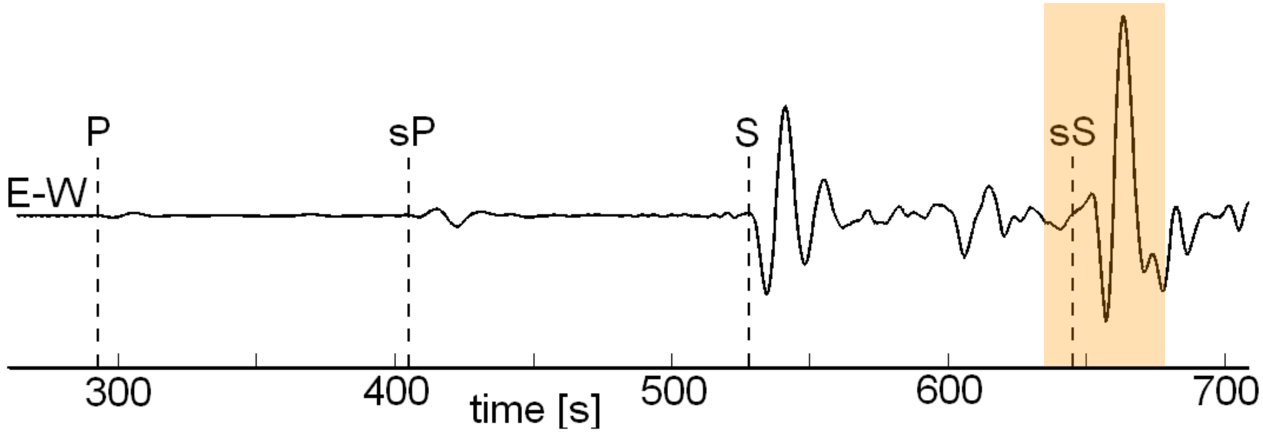




The adjoint method

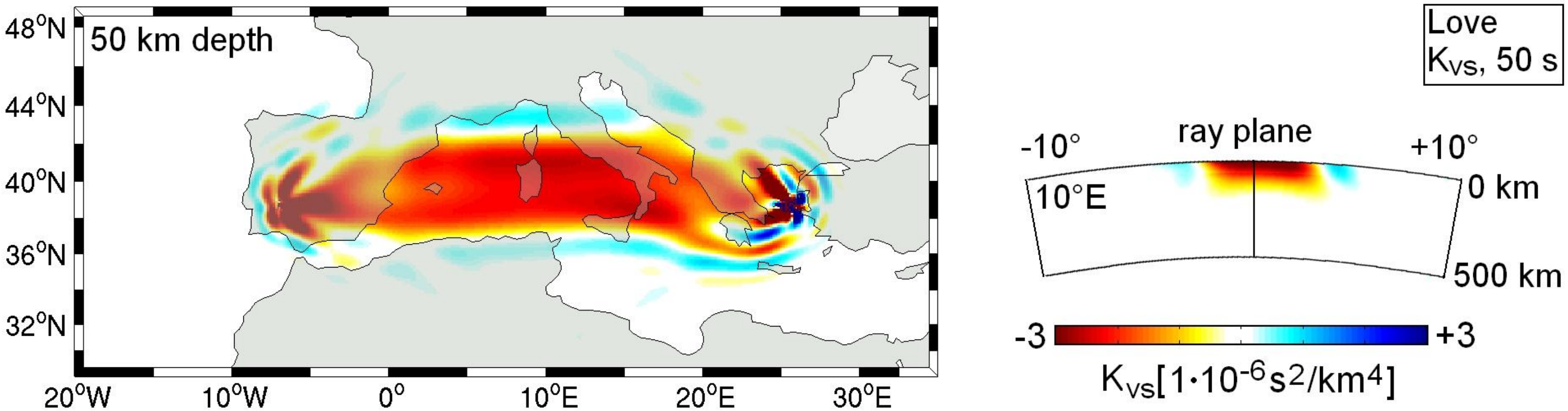
Fréchet kernel gallery

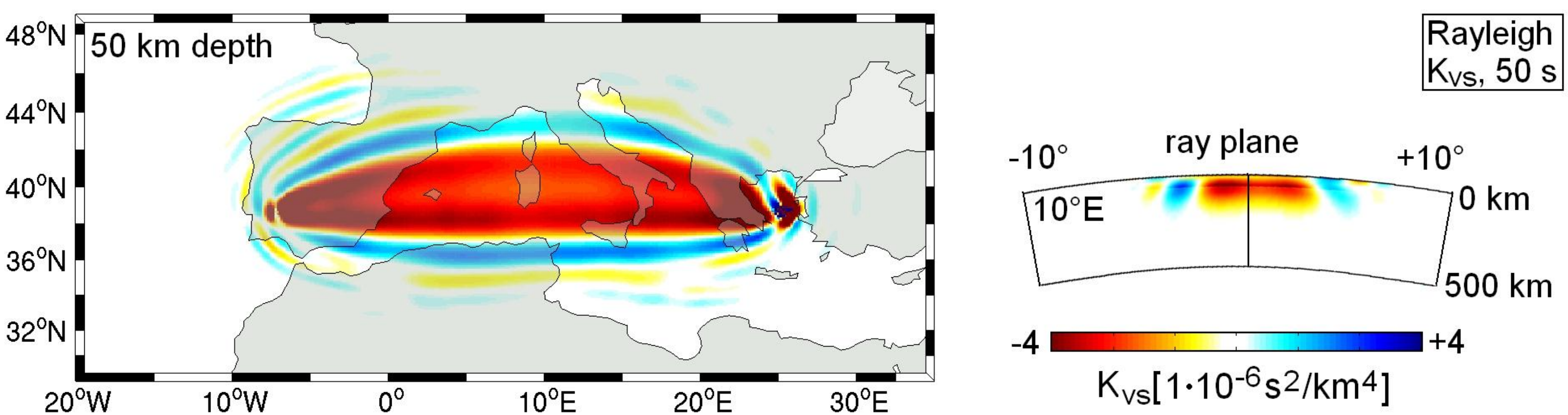




The adjoint method

Fréchet kernel gallery





Practical

1. Storage of the forward field
2. Construction of the adjoint source time function
3. Fréchet kernels

Practical

Storage of the forward field

- To compute Fréchet kernels, the regular velocity and strain fields must be stored during the forward simulation.

Par file

```
OUTPUT DIRECTORY =============================================
../DATA/OUTPUT/1/
OUTPUT FLAGS =================================================
100                                       ! ssamp, snapshot sampling
1                                         ! output_displacement
0                                         ! output_preco
ADJOINT PARAMETERS ===========================================
1                                         ! adjoint flag
10                                        ! samp_ad
../INTERMEDIATE/
```

store velocity and strain field (1=yes, 0=no) every samp_ad (=10) time steps in the directory below (../INTERMEDIATE)

Needed for the computation of sensitivity kernels.

Practical

Storage of the forward field

- To compute Fréchet kernels, the regular velocity and strain fields must be stored during the forward simulation.

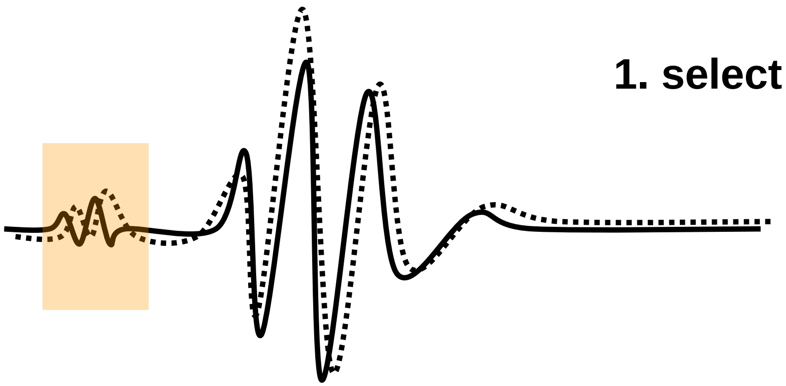
Par file

```
OUTPUT DIRECTORY ==========================================
../DATA/OUTPUT/1/
OUTPUT FLAGS ==============================================
100                                      ! ssamp, snapshot sampling
1                                        ! output_displacement
0                                        ! output_preco
ADJOINT PARAMETERS ========================================
2 }                                      ! adjoint flag
10                                       ! samp_ad
../INTERMEDIATE/
```

adjoint flag = 2: read the stored forward fields and compute Fréchet kernels

Practical

Construction of the adjoint source (for cross-correlation time shifts)



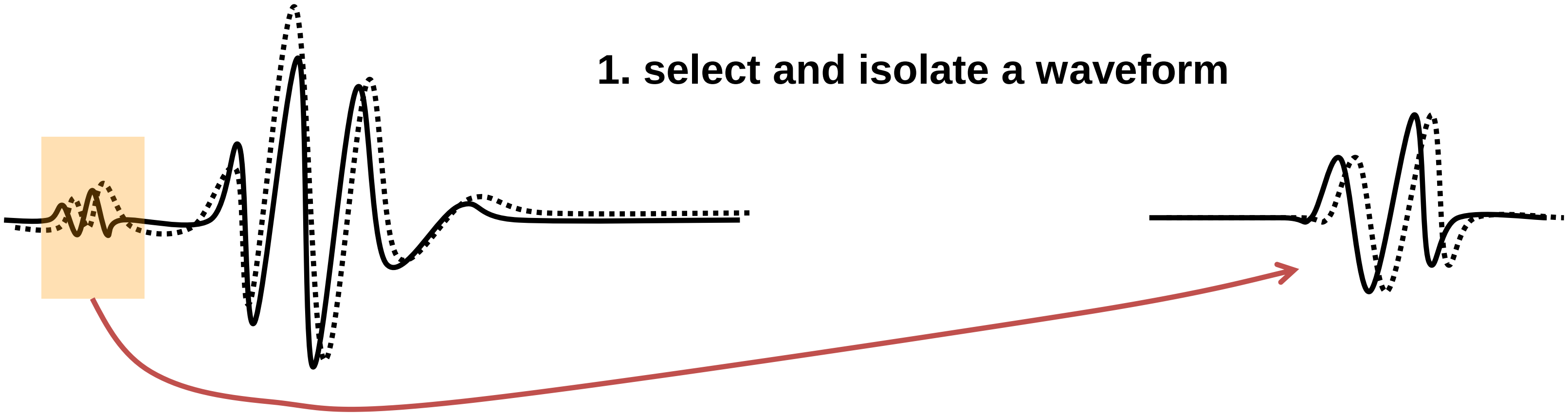
1. select and isolate a waveform

Luo & Schuster, 1991
Used before in surface wave analysis.

Practical

Construction of the adjoint source (for cross-correlation time shifts)

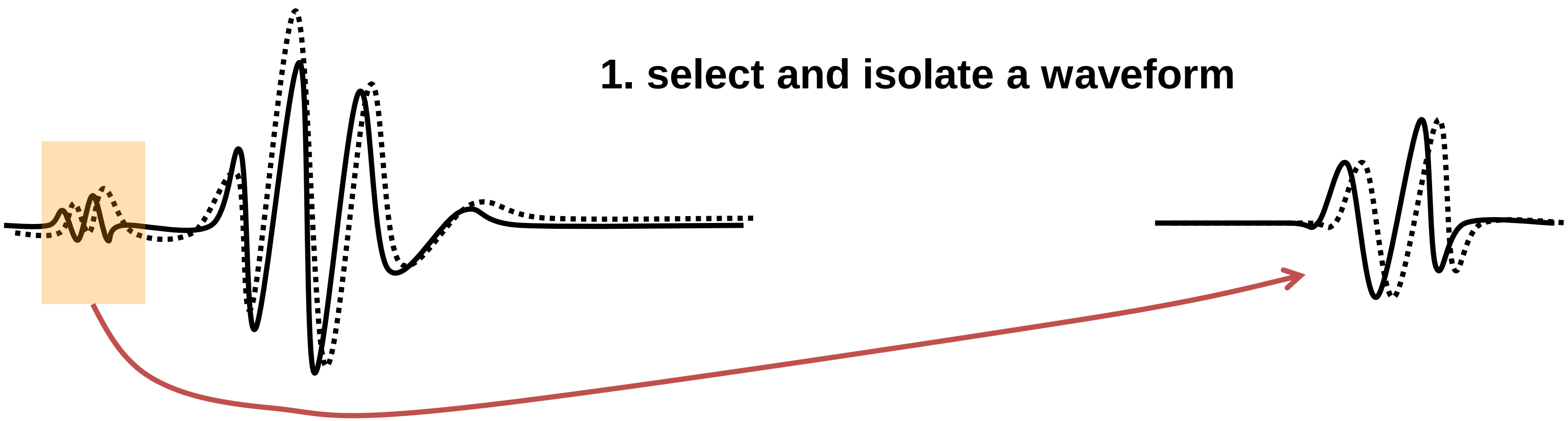
1. select and isolate a waveform



Practical

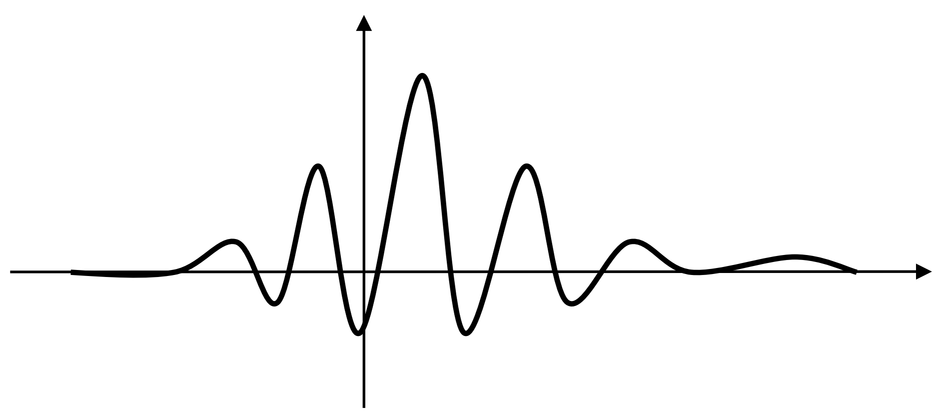
Construction of the adjoint source (for cross-correlation time shifts)

1. select and isolate a waveform



2. compute correlation function

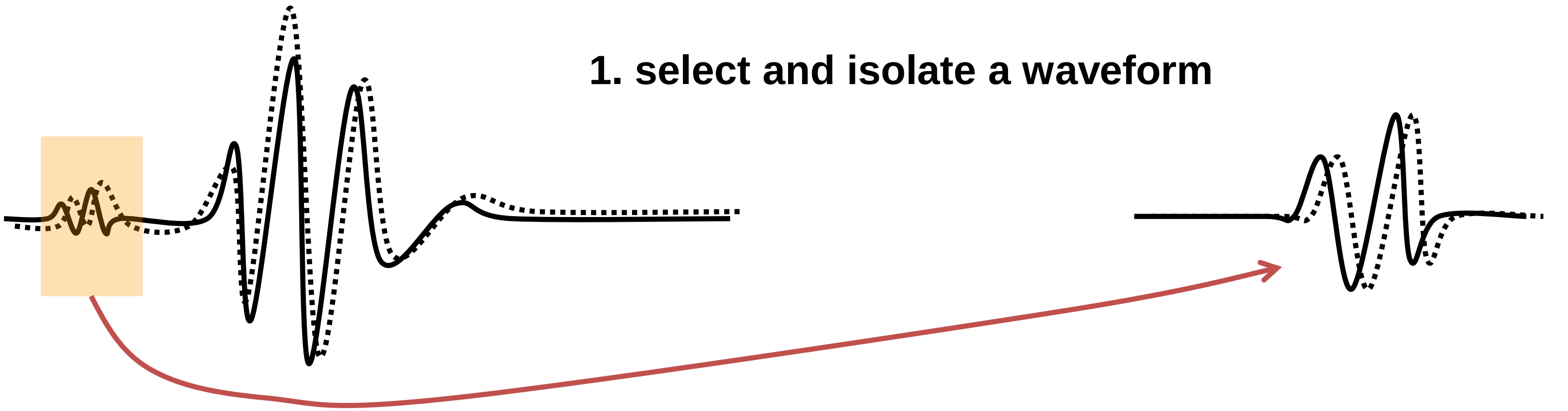
$$C(t) = \int_{\tau} u(t+\boldsymbol{\tau})\,u_0(\boldsymbol{\tau})\,d\boldsymbol{\tau}$$



Practical

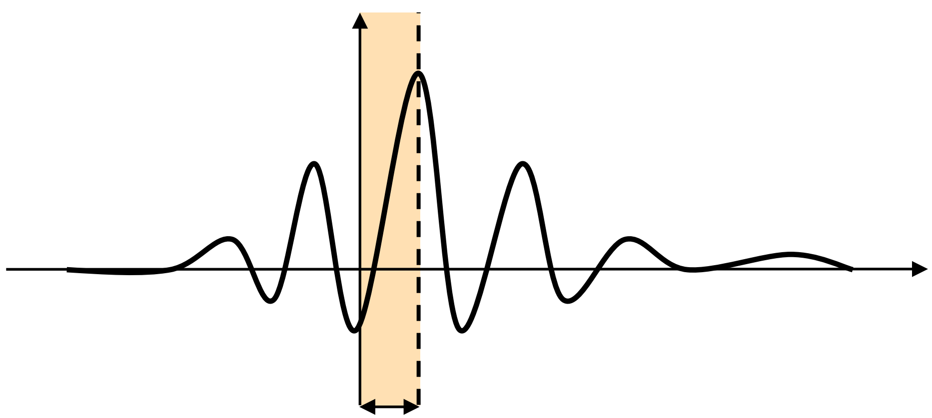
Construction of the adjoint source (for cross-correlation time shifts)

1. select and isolate a waveform



2. compute correlation function

$$\chi = \Delta T^2$$



ΔT = **cross-correlation time shift**

Practical

Construction of the adjoint source (for cross-correlation time shifts)

- The corresponding adjoint source is:

$$f^*(t,x) = \frac{\dot{u}(t,x^r)}{\|\dot{u}(t,x^r)\|^2}\,\delta(x - x^r)$$

amplitude normalisation
→
traveltime kernels are independent of the wave amplitude

acts at the receiver position

$$\boxed{K_\rho^0 = -\int_T \dot{\mathbf{u}}^\dagger \cdot \dot{\mathbf{u}}\, dt}$$

Practical

Construction of the adjoint source (for cross-correlation time shifts)

- The corresponding adjoint source is:

$$f^*(t,x) = \frac{\dot{u}(t,x^r)}{\|\dot{u}(t,x^r)\|^2}\,\delta(x-x^r)$$

amplitude normalisation
→
traveltime kernels are independent of the wave amplitude

acts at the receiver position

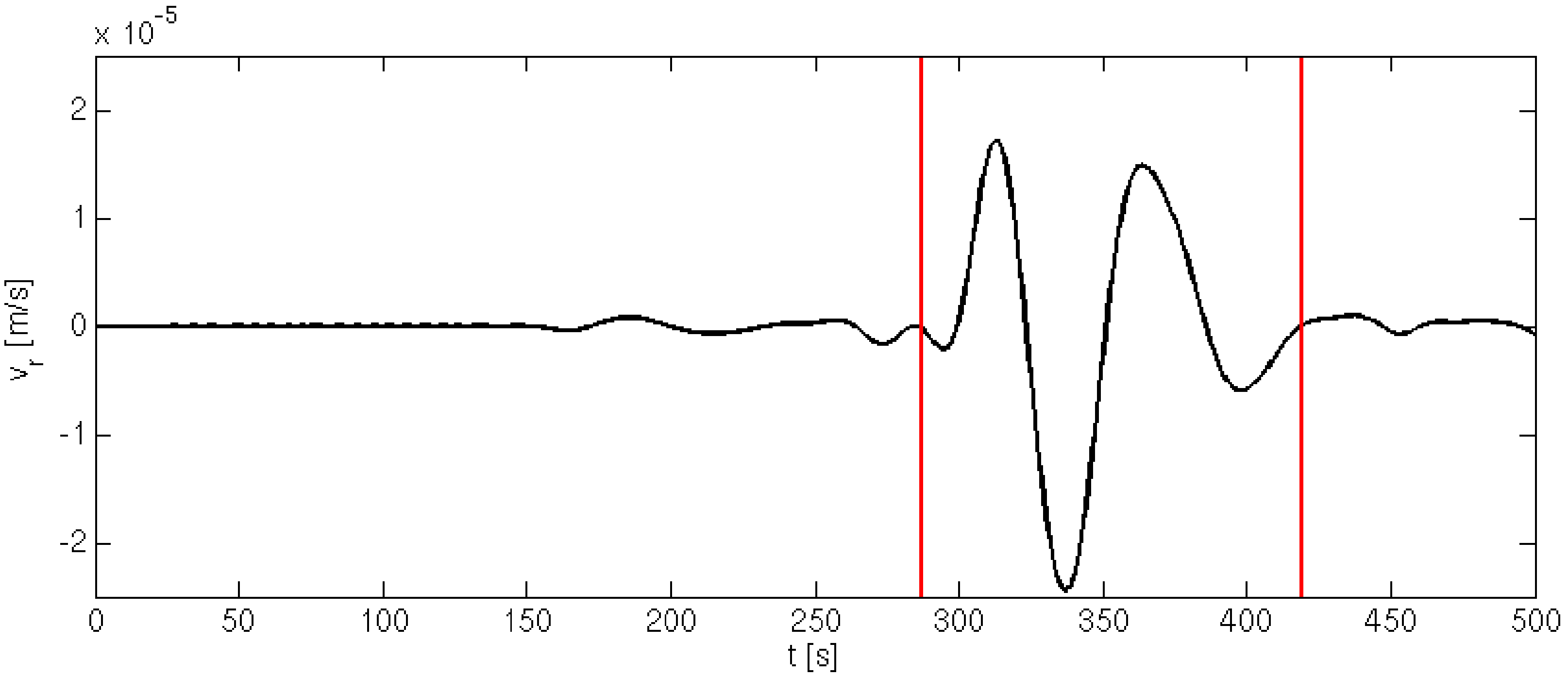
$$K_\rho^0 = -\int_T \dot{\mathbf{u}}^\dagger \cdot \dot{\mathbf{u}}\,dt$$

- The above equation is an approximation! It holds – paradoxically – only when the observed and synthetic waveforms are shifted in time without being otherwise distorted.

Practical

Construction of the adjoint source (for cross-correlation time shifts)

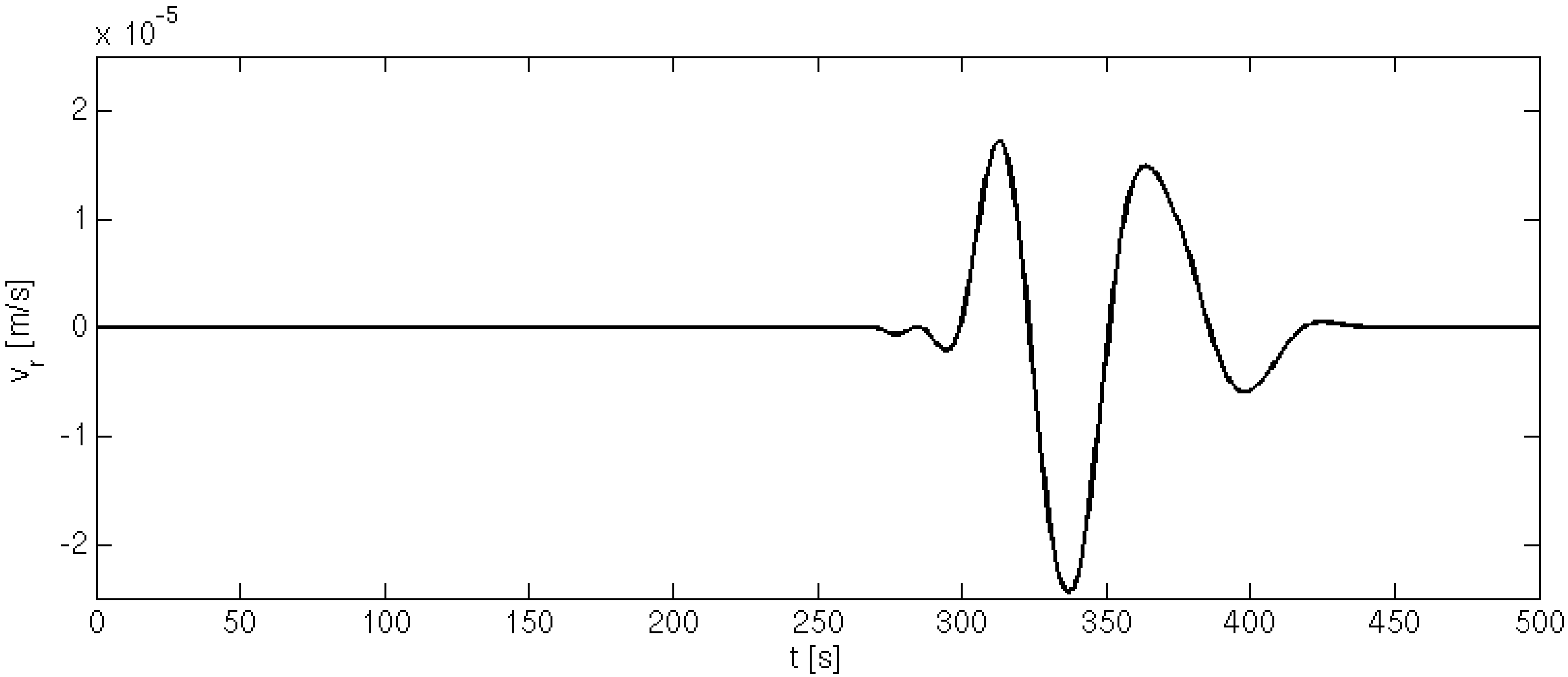
- **Step 1:** read synthetic seismogram and select the waveform of interest



Practical

Construction of the adjoint source (for cross-correlation time shifts)

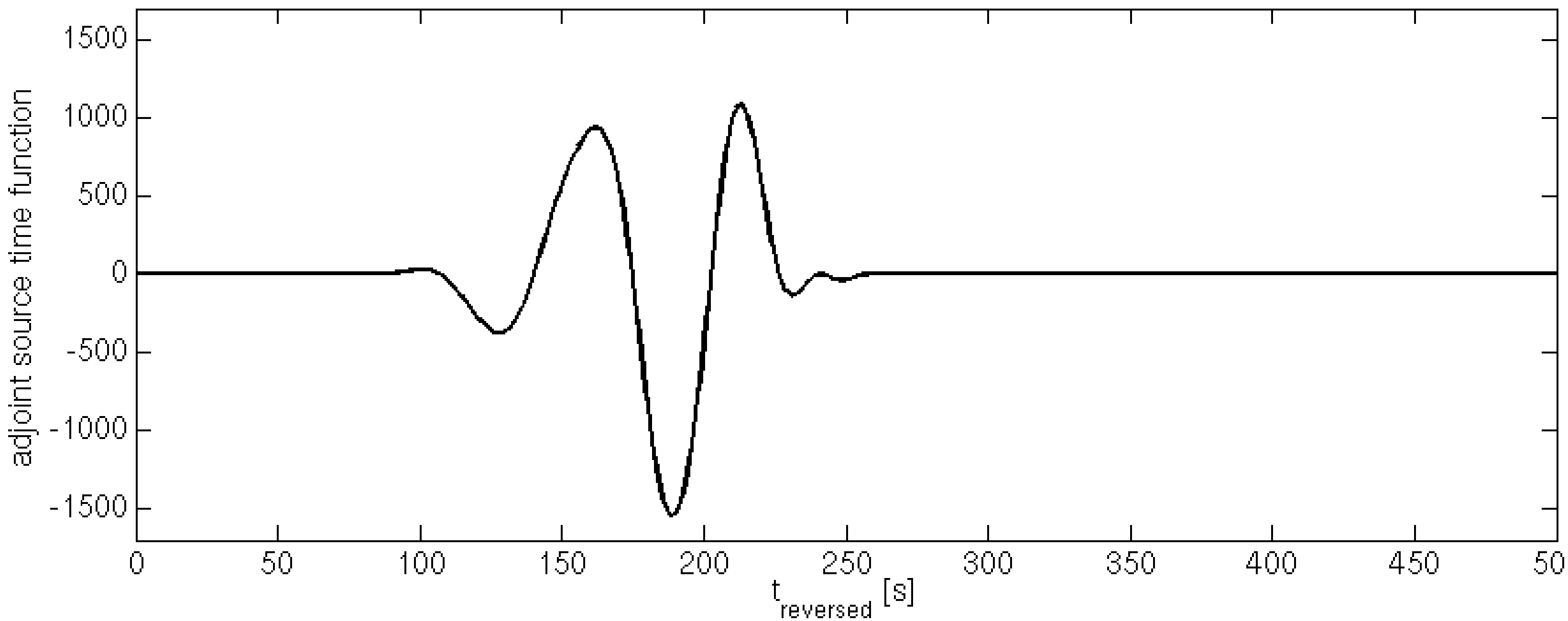
- **Step 2:** apply a window function



Practical

Construction of the adjoint source (for cross-correlation time shifts)

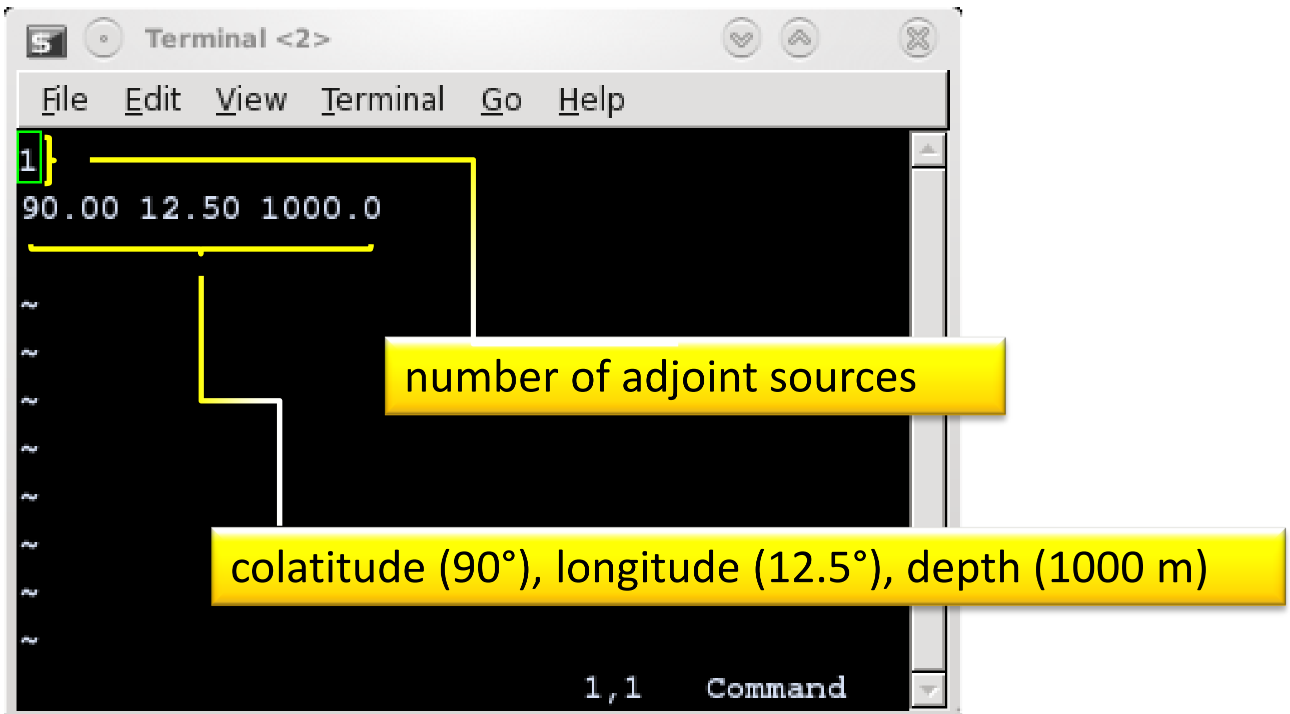
- **Step 3:** scale and reverse in time



Practical

Construction of the adjoint source (for cross-correlation time shifts)

- **Step 4:** write the adjoint source time function and location into the following files:

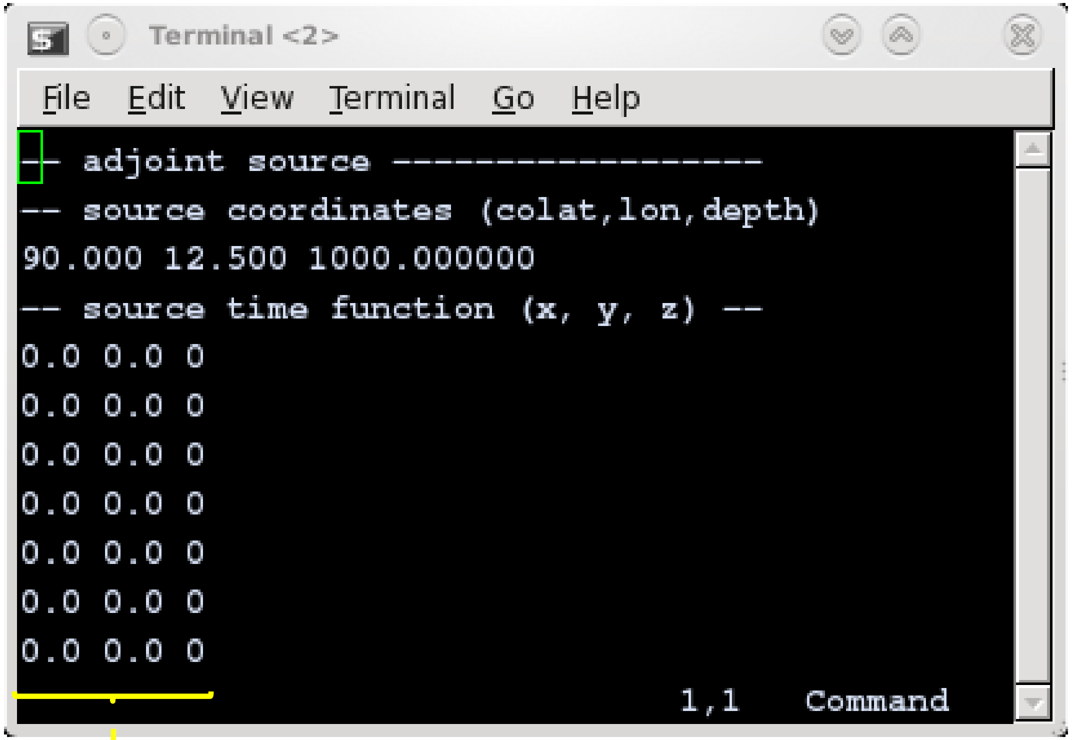


ad_srcfile

Practical

Construction of the adjoint source (for cross-correlation time shifts)

- **Step 4:** write the adjoint source time function and location into the following files:



ad_src_1, ad_src_2, ...

theta-, phi- and r-component of the adjoint source

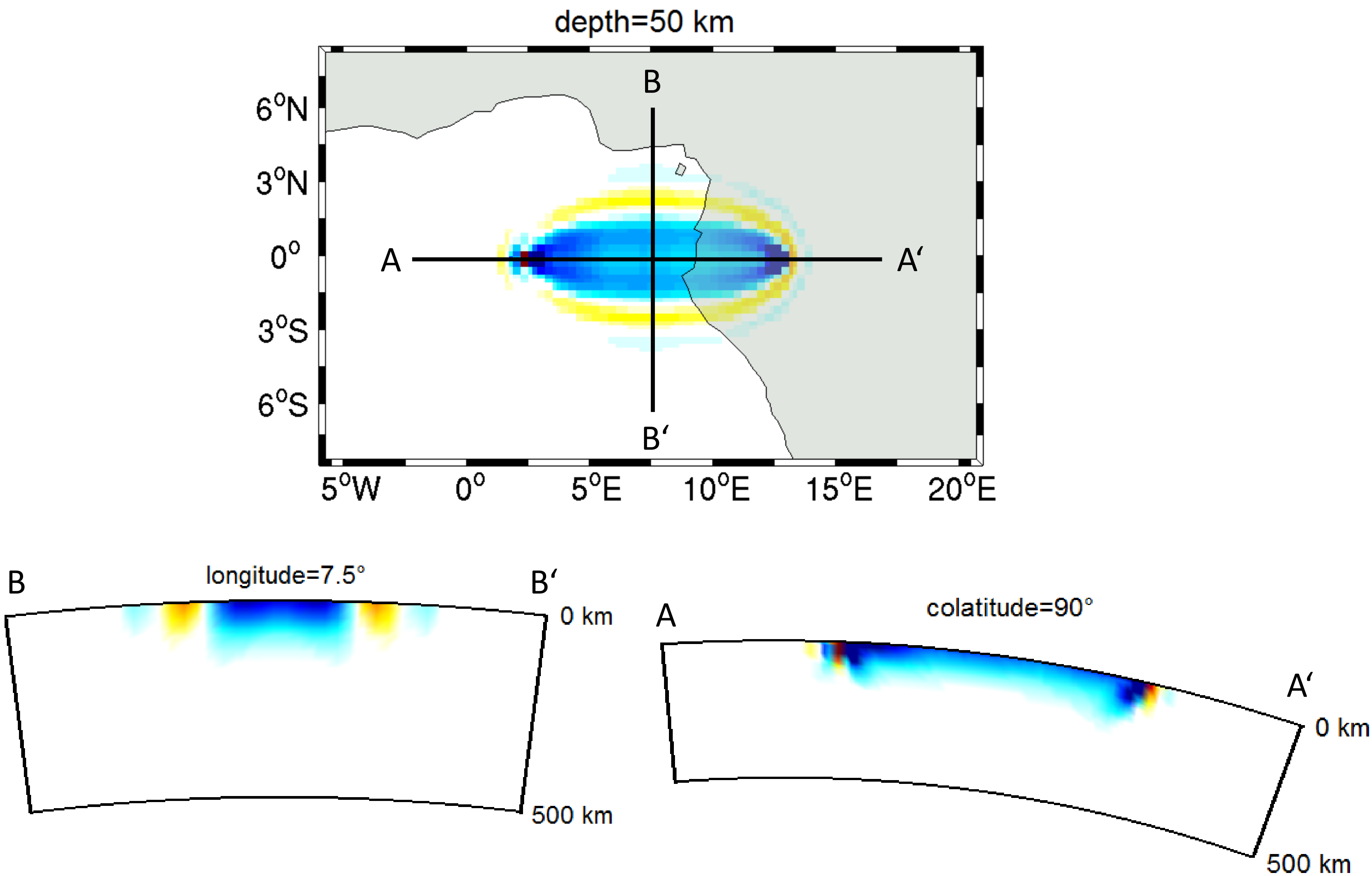
Practical

Construction of the adjoint source (for cross-correlation time shifts)

- **Step 5:** run ses3d (main.exe). The output (Fréchet kernels) is then written to the output directory.

Practical

and finally: **Fréchet kernels**



Special topics

1. The adjoint method for seismic source parameters
2. Time-reversal imaging of seismic sources

The adjoint method for seismic source parameters

- What is the reaction of the seismic wave field to changes in the source parameters?

$$\rho\ddot{u} - \partial_x(\mu\partial_x u) = f \quad \Rightarrow \quad u(x,t) \tag{4.1}$$

$$\rho\ddot{u} - \partial_x(\mu\partial_x u) = f + \delta f \quad \Rightarrow \quad u(x,t) + \delta u(x,t) \tag{4.2}$$

- How does this change in the wave field affect the misfit X ?

$$\frac{dX}{df}\delta f = ??? \tag{4.3}$$

The adjoint method for seismic source parameters

- With the adjoint method we can answer this question in a straightforward way:
- Simply redefine the wave operator …

$$\boxed{\begin{array}{l} \rho\ddot{u} - \partial_x(\mu\partial_x u) = f \\ \qquad L(u,m) = f \qquad m = (\rho,\mu) \end{array}} \qquad 2.11$$

… to:

$$\boxed{\begin{array}{l} \rho\ddot{u} - \partial_x(\mu\partial_x u) - f = 0 \\ \qquad L(u,m) = 0 \qquad m = (\rho,\mu,f) \end{array}} \qquad 4.4$$

- Then repeat the previous derivation.

The adjoint method for seismic source parameters

- The adjoint equations are exactly the same as before:

- The adjoint field is governed by the adjoint equation:

$$\boxed{L^*(u^*) = -f^*}$$

- The adjoint field satisfies terminal conditions:

$$\boxed{u^* |_{t=T} = \dot{u}^* |_{t=T} = 0}$$

- But the Fréchet derivative (kernel) with respect to the source is much simpler:

$$\boxed{\frac{dX}{df}\delta f = -\iint u^*(x,t)\cdot \delta f(x,t)\, dt\, d^3x} \quad \boxed{K_f = -\int u^*(x,t)\, dt} \qquad 4.5$$

- The forward field is not involved in the Fréchet derivative and thus need not be stored.

The adjoint method for seismic source parameters

Example

- Misfit functional:

$$\mathrm{X}(m) = \frac{1}{2}\int [u - u_0]^2 \, dt \qquad 4.6$$

synthetic waveform (u)

observed waveform (u_0)

- Adjoint source:

$$f^* = (u - u_0)\,\delta(x - x^r) \qquad 4.7$$

- Initial source model:

$$f = 0 \quad \Rightarrow \quad u = 0 \quad \Rightarrow \quad f^* = -u_0\,\delta(x - x^r) \qquad 4.8$$

The adjoint method for seismic source parameters

Example

- Adjoint equation (for the 1D case):

$$\boxed{\rho \ddot{u}^* - \partial_x(\mu \partial_x u^*) = u_0\, \delta(x - x^r)} \qquad 4.9$$

- The recorded seismograms are re-injected at the source positions and propagated backwards in time.

- And the Fréchet kernel is just the backward propagated field integrated over time:

$$\boxed{K_f = -\int u^*(x,t)\, dt} \qquad 4.10$$