An Introduction to Modeling, Imaging and Full Waveform Inversion

Pengliang Yang

ISTerre, Univ. Grenoble Alpes, France
ypl.2100@gmail.com
http://seiscope2.osug.fr

2016 Madagascar school, Zurich

Outline

Outline

1. The wave equation
2. Forward modeling
3. Absorbing boundary condition
4. Seismic imaging/migration
5. Full waveform inversion (FWI)

Particle velocity vector $\mathbf{v} = [v_x, v_y, v_z]^T$; stress vector $\boldsymbol{\sigma} = [\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{yz}, \sigma_{xz}, \sigma_{xy}]^T$

- Newton's law: $\rho \partial_t v_i = \partial_j \sigma_{ij}, \qquad i, j \in \{x, y, z\} = \{1, 2, 3\}$
- Generalized Hooke's law: $\sigma_{ij} = c_{ijkl} \epsilon_{kl}$

Voigt indexing: $(11) \to 1$, $(22) \to 2$, $(33) \to 3$, $(23) = (32) \to 4$, $(13) = (31) \to 5$, $(12) = (21) \to 6$

$$\begin{cases} \rho \partial_t \mathbf{v} = D \boldsymbol{\sigma} + \mathbf{f}_\mathbf{v} \\ \partial_t \boldsymbol{\sigma} = C D^T \mathbf{v} + \mathbf{f}_{\boldsymbol{\sigma}} \end{cases}, C = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ \cdot & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ \cdot & \cdot & c_{33} & c_{34} & c_{35} & c_{36} \\ \cdot & \cdot & \cdot & c_{44} & c_{45} & c_{46} \\ \cdot & SYM & \cdot & \cdot & c_{55} & c_{56} \\ \cdot & \cdot & \cdot & \cdot & \cdot & c_{66} \end{bmatrix}, D^T = \begin{bmatrix} \partial_x & 0 & 0 \\ 0 & \partial_y & 0 \\ 0 & 0 & \partial_z \\ 0 & \partial_z & \partial_y \\ \partial_z & 0 & \partial_x \\ \partial_y & \partial_x & 0 \end{bmatrix} \quad (1)$$

- Triclinic: 21 independent coefficients in C
- Monoclinic: 13 independent coefficients in C
- Orthorombic: 9 independent coefficients in C
- Transverse isotropic: 5 independent coefficients in C

Isotropic elastic medium:

$$C = \begin{bmatrix} \lambda+2\mu & \mu & \mu & 0 & 0 & 0 \\ \mu & \lambda+2\mu & \mu & 0 & 0 & 0 \\ \mu & \mu & \lambda+2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix} \tag{2}$$

Acoustic medium ($\mu = 0$): pressure $p = (\sigma_{xx} + \sigma_{yy} + \sigma_{zz})/3$

$$\begin{cases} \rho \partial_t \mathbf{v} = \nabla p \\ \partial_t p = \kappa \nabla \cdot \mathbf{v}, \kappa = \rho c^2 \end{cases} \tag{3}$$

1st order equation system $\rightarrow$ 2nd order equation

$$\frac{1}{c^2}\partial_{tt} p - \rho \nabla \cdot \left(\frac{1}{\rho} \nabla p \right) = f \tag{4}$$

Constant density:

$$(\frac{1}{c^2}\partial_{tt} - \nabla^2) p = f, \nabla^2 = \partial_{xx} + \partial_{yy} + \partial_{zz} \tag{5}$$

Outline

Modeling=Simulate wave propagation by solving wave equation numerically

Which numerical method?

1. Finite difference method (FDM)
2. Finite element method (FEM) → Spectral element method (SEM)
3. Fourier/spectral method: pseudo spectral method (PSM), lowrank finite difference

Tradeoff between computation and accuracy?

- FDM outperforms FEM and PSM due to excellent efficiency;
- SEM, PSM, lowrank FD arrive at spectral accuracy in space;
- FEM, SEM are more flexible to handle complex topography, generally more expensive than FDM;

In which domain: Time domain or Frequency domain?

1. Finite difference method (FDM) in time domain
   - direct discretization of the original differential equation, both in time and in space
   - explicit time integration: $n \to n+1$, $(\frac{1}{c^2}\partial_{tt} - \partial_{xx})p = f$

$$\frac{p_j^{n+1} - 2p_j^n + p_j^{n-1}}{\Delta t^2 c^2} - \partial_{xx} p^n = f^n \Rightarrow p_j^{n+1} = 2p_j^n - p_j^{n-1} + \frac{\Delta t^2 c^2}{\Delta x^2}(p_{j+1}^n - 2p_j^n + p_{j-1}^n) \quad (6)$$

   - implicit time integration: $n \to n+1$, matrix inversion

$$\frac{p^{n+1} - 2p^n + p^{n-1}}{\Delta t^2 c^2} - \partial_{xx} p^{n+1} = f^n \quad (7)$$

$$\Rightarrow Ap^{n+1} = b, A = \begin{bmatrix} 2+\alpha_2 & -1 & & & \\ -1 & \ddots & \ddots & & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2+\alpha_M \end{bmatrix}, \alpha_j = \frac{\Delta x^2}{c_j^2 \Delta t^2} \quad (8)$$

1. Finite difference method (FDM) in frequency domain: solve Helmholz equation
   - direct discretization of the original differential equation, both in frequency and space
   - implicit scheme with matrix inversion $\tilde{\mathbf{p}} = A^{-1}\tilde{\mathbf{f}}$

$$(\frac{1}{c^2}\partial_{tt} - \partial_{xx})p = f \Rightarrow -\frac{\omega^2}{c^2}\tilde{p} - \partial_{xx}\tilde{p} = f \Rightarrow -\frac{\omega^2}{c_j^2}\tilde{p}_j - \frac{\tilde{p}_{j+1} - 2\tilde{p}_j + \tilde{p}_{j-1}}{\Delta x^2} = \tilde{f} \quad (9)$$

$$\Rightarrow \underbrace{\begin{bmatrix} \ddots & \ddots & \ddots & & \\ \cdots & -\frac{1}{\Delta x^2} & \frac{2}{\Delta x^2} - \frac{\omega^2}{c_j{}^2} & -\frac{1}{\Delta x^2} & \cdots \\ & & \ddots & \ddots & \ddots \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \vdots \\ p^{n-1} \\ p^n \\ p^{n+1} \\ \vdots \end{bmatrix}}_{\tilde{\mathbf{p}}} = \underbrace{\begin{bmatrix} \vdots \\ \tilde{f} \\ \vdots \end{bmatrix}}_{\tilde{\mathbf{f}}} \quad (10)$$

   - very efficient for multiple sources due to only 1 time of matrix inversion: $[\tilde{\mathbf{p}}_1, \tilde{\mathbf{p}}_2, \ldots] = A^{-1}[\tilde{\mathbf{f}}_1, \tilde{\mathbf{f}}_2, \ldots]$
   - easy to incorporate attenuation $\tilde{c}(\omega) = c_0(1 - \frac{isign(\omega)}{2Q}) \rightarrow$ more diagonal dominant for stable matrix inversion
   - filling the matrix in multifrontal parallel solver: MUMPS, umfpack, ...

2. Finite element method (FEM): matrix inversion at each time step
$\Rightarrow$ solving the weak form of the differential equation, test/basis function $\phi_j \in [0, 1]$
Second order wave equation with displacement

$$\int_0^1 \Big(\partial_{tt} u(x,t) - c^2(x) \partial_{xx} u(x,t) - f(x,t) \Big) \phi_j(x) dx = 0 \quad (11)$$

$\Rightarrow$ state variable $u(x,t) = \sum_{i=1}^{N} u_i(t)\phi_i(x)$: polynomial expansion separable over space and time

$$\sum_i \partial_{tt} u_i \underbrace{\int_0^1 \phi_i \phi_j \mathrm{d}x}_{M_{ji}} + c^2 \sum_i u_i \underbrace{\int_0^1 \partial_x \phi_i \partial_x \phi_j \mathrm{d}x}_{K_{ji}} = \int_0^1 f \phi_j \mathrm{d}x \Rightarrow M \partial_{tt} \mathbf{u} + K\mathbf{u} = \mathbf{f}' \quad (12)$$

$$M \frac{\mathbf{u}^{n+1} - 2\mathbf{u}^n + \mathbf{u}^{n-1}}{\Delta t^2} + K\mathbf{u}^n = \mathbf{f}' \Rightarrow \mathbf{u}^{n+1} = 2\mathbf{u}^n - \mathbf{u}^{n-1} + M^{-1}(\mathbf{f}' - K\mathbf{u}^n) \quad (13)$$

- SEM (Komatisch et al., 1998): specific mesh, diagonal mass matrix M^{-1} for efficient wave extrapolation by choosing GLL collocation points

3. Fourier/spectral methods: computing spatial derivative using Fourier transform
   - Pseudo spectral method (PSM):

$$f(x) = \int f(k)e^{ikx}dk, \Rightarrow \partial_x f = \int ikf(k)e^{ikx}dk, \partial_{xx} f = \int -k^2 f(k)e^{ikx}dk \tag{14}$$

$$\frac{1}{c^2}\partial_{tt} p = (\partial_{xx} + \partial_{zz})p \Rightarrow p^{n+1} = 2p^n - p^{n-1} + \Delta t^2 c^2 F^{-1}[-(k_x^2 + k_z^2)Fp^n] \tag{15}$$

   - Lowrank finite difference method (Fomel et al., 2013):

$$p(\mathbf{x}, t+\Delta t) = \int p(\mathbf{k}, t)e^{i\phi(\mathbf{x},\mathbf{k},\Delta t)}\mathrm{d}\mathbf{k}, = \int e^{i\mathbf{k}\cdot\mathbf{x}} W(\mathbf{x}, \mathbf{k}) p(\mathbf{k}, t)\mathrm{d}\mathbf{k}, \tag{16}$$

$$W(\mathbf{x}, \mathbf{k}) = e^{i[\phi(\mathbf{x},\mathbf{k},\Delta t) - \mathbf{k}\cdot\mathbf{x}]}, W(\mathbf{x}, \mathbf{k}) \approx \sum_{m=1}^{M}\sum_{n=1}^{N} W(\mathbf{x}, \mathbf{k}_m) a_{mn} W(\mathbf{x}_n, \mathbf{k}) \tag{17}$$

Approximate lowrank decomposition leads to

$$p(\mathbf{x}, t+\Delta t) \approx \sum_{m=1}^{M} W(\mathbf{x}, \mathbf{k}_m) \sum_{n=1}^{N} a_{mn} \int e^{i\mathbf{k}\cdot\mathbf{x}} W(\mathbf{x}_n, \mathbf{k}) p(\mathbf{k}, t)\mathrm{d}\mathbf{k}, \tag{18}$$

SEISCOPE

PA (Paraxial approximation, radiation) boundary condition (Clayton and Engquist, 1977):
$\rightarrow$ cancel the refection on the boundary using waves propagating in opposite direction according to pseudo differential operator decomposition

$$(\frac{1}{c^2}\partial_{tt} - \partial_{xx} - \partial_{zz})p = -(\partial_x + \frac{1}{c}\partial_t\sqrt{1-S^2})(\partial_x - \frac{1}{c}\partial_t\sqrt{1-S^2})p = 0, S = \frac{\partial_z}{\partial_t/c} \quad (19)$$

- 1st order approximation: $\sqrt{1-S^2} \approx 1$, perfect in 1D $S=0$, accurate within a specific range of angles in multi-dimensions!

$$(\frac{1}{c^2}\partial_{tt}-\partial_{xx})p = (\frac{1}{c}\partial_t-\partial_x)(\frac{1}{c}\partial_t+\partial_x)p = 0 \Rightarrow x \in [0,L] \begin{cases} \frac{1}{c}\partial_t p(0,t) - \partial_x p(0,t) = 0 \\ \frac{1}{c}\partial_t p(L,t) + \partial_x p(L,t) = 0 \end{cases} \quad (20)$$

- 2nd order approximation: $\sqrt{1-S^2} \approx 1 - \frac{1}{2}S^2$

$$\left(\partial_x - \frac{\partial_t}{c} + \frac{c\partial_{zz}}{2\partial_t}\right)p = 0 \Rightarrow \partial_{xt}p - \frac{1}{c}\partial_{tt}p + \frac{c}{2}\partial_{zz}p = 0 (x=0) \quad (21)$$

Sponge layers/Gaussian taper boundary condition (Cerjan et al., 1985):
→ multiply an exponential decaying factor $e^{-d_x \Delta t}$ in boundary layers at each time step

- easy to implement, stable in any anisotropic medium!
- a good choice for the damping profile to achieve PML-like effect:

$$d_x = -\frac{3v_{\max}}{2L}\log(R_c)\left(\frac{x}{L}\right)^2, R_c = 10^{-3} \sim 10^{-6} \quad (22)$$

PML(Perfectly matched layer)(Bérenger, 1994): 1st order wave equation system
→ coordinate transformation in complex-valued anisotropic medium

$$\tilde{x} = x + \frac{1}{i\omega}\int_0^x d_x(s)\mathrm{d}s \Rightarrow \partial_{\tilde{x}} = \frac{1}{1+\frac{d_x}{i\omega}}\partial_x, \quad (23)$$

- intrinsically unstable in anisotropic medium for long duration simulation!
- Frequency domain multiplication → Time domain convolution, implemented by recursion with memory variables

$$\begin{cases}\partial_t \mathbf{v} = \frac{1}{\rho}\nabla p \\ \partial_t p = \kappa \nabla\cdot\mathbf{v}\end{cases} \Rightarrow \begin{cases}\partial_t v_i = \frac{1}{\rho}\partial_i p - d_i v_i, i \in \{x,y,z\} \\ \partial_t p = \kappa(\partial_x v_x + \partial_y v_y + \partial_z v_z) - (d_x + d_y + d_z)p\end{cases} \quad (24)$$

Outline

1. The wave equation
2. Forward modeling
3. Absorbing boundary condition
4. Seismic imaging/migration
5. Full waveform inversion (FWI)

- Deconvolution (U/D) IC (Claerbout, 1971): Recorded signal (upcoming wavefield at receivers) = Earth's reflectivity * downgoing source wavelet (downgoing wavefield).

$$U(\mathbf{x}, t) = R(\mathbf{x}) * D(\mathbf{x}, t) \Rightarrow R(\mathbf{x}) = \frac{U(\mathbf{x}, t)}{D(\mathbf{x}, t)}, \text{unstable} \quad (25)$$

- Zero-lag crosscorrelation IC → illumination compensation

$$\begin{aligned} I(\mathbf{x}) &= \sum_{i=1}^{nt} U(\mathbf{x}, t_i) D(\mathbf{x}, t_i) \Rightarrow \\ I'(\mathbf{x}) &= \frac{\sum_{i=1}^{nt} U(\mathbf{x}, t_i) D(\mathbf{x}, t_i)}{\sum_{i=1}^{nt} D^2(\mathbf{x}, t_i)} = \frac{U_1 D_1 + \cdots + U_i D_{i\max|\cdot|} + \cdots + U_{nt} D_{nt}}{D_1 D_1 + \cdots + D_i D_{i\max|\cdot|} + \cdots + D_{nt} D_{nt}} \end{aligned} \quad (26)$$

requires simultaneous accesing to the source and receiver wavefields

- Excitation amplitude IC (Chang and McMechan, 1986): taking maximum absolute of downgoing wavefield at excitation time t_i

$$I(\mathbf{x}) = \frac{U(\mathbf{x}, t_i)}{\max |D(\mathbf{x}, t_i)|} \quad (27)$$

knowing t_i is enough, no need for storing or reconstructing source wavefield

Extended IC (Sava and Fomel, 2006): redundant dimensions to evaluate energy focusing

- Conventional IC:$I(\mathbf{x}) = R(\mathbf{x}, t = 0), R(\mathbf{x}, t) = U(\mathbf{x}, t) * D(\mathbf{x}, t)$
- Space-shift IC: $\mathbf{h} = 0 \rightarrow$ conventional IC

$$I(\mathbf{x}, \mathbf{h}) = R(\mathbf{x}, \mathbf{h}, t = 0), R(\mathbf{x}, \mathbf{h}, t) = U(\mathbf{x} + \mathbf{h}, t) * D(\mathbf{x} - \mathbf{h}, t) \quad (28)$$

- Time-shift IC: $\tau = 0 \rightarrow$ conventional IC

$$I(\mathbf{x}, \tau) = R(\mathbf{x}, \tau, t = 0), R(\mathbf{x}, \tau, t) = U(\mathbf{x}, t + \tau) * D(\mathbf{x}, t - \tau) \quad (29)$$

- Practical usage in RTM and FWI: horizontal subsurface offset shift, $(x \pm h, y, z)$
- Implementated in frequency domain via Fourier transform

$$R(\mathbf{m}, \mathbf{h}) = \sum_{\omega} U(\mathbf{x} + \mathbf{h}, \omega) D^*(\mathbf{x} - \mathbf{h}, \omega) \quad R(\mathbf{m}, \tau) = \sum_{\omega} U(\mathbf{x}, \omega) D^*(\mathbf{x}, \omega) e^{2i\omega\tau} \quad (30)$$

- Excitation amplitude IC+ extended time/space shift $\Rightarrow$ an easier way to obtain angle gathers with low computation and memory cost? $R(\mathbf{x}, \tau) \Rightarrow R(\mathbf{x}, \theta)$

Outline

Objective=Waveform misfit

$$\min_m C(m) = \frac{1}{2}\|d_{cal}(m) - d_{obs}\| \tag{31}$$

Given an initial model m_0, iteratively update the model

$$m^{k+1} = m^k + \gamma_k \Delta m^k \tag{32}$$

using different optimization methods: $\Delta m^k = -H^{-1}\nabla_m C$

- Steepest descent method (SDM): $\Delta m^k = -\nabla_m C$
- Nonlinear conjugate gradient (NLCG) (Pica et al., 1990): $\Delta m^k = d_k = -\nabla_m C + \beta d_{k-1}$
- Quasi-Newton method, L-BFGS (Brossier, 2011): $H \approx H_a$ approximate Hessian using previously memorized gradients $\nabla C|^k, \nabla C|^{k-1}, \nabla C|^{k-2}, \cdots$
- Newton method, Truncated Newton (Santosa and Symes, 1988; Métivier et al., 2014): compute Hessian vector product using CG method within few iterations
- Gauss-Newton (Shin et al., 2001): $H_a = J^\dagger J$, approximate Hessian using first order scattering information

Frequency/Fourier domain FWI (Pratt et al., 1998): $-(\frac{1}{v^2}\omega^2 + \nabla^2)\tilde{p} = \tilde{f}$

$$C(m) = \frac{1}{2}\sum_s \sum_r \sum_\omega |d_{cal}^{s,r}(m,\omega) - d_{obs}^{s,r}(\omega)|^2, d_{cal}^{s,r}(m,\omega) = R_{s,r}\tilde{p} \quad (33)$$

- Multiscale approach: from low frequency to high frequency
- Extremely efficient for multisource problem: same A^{-1} for any source $\tilde{\mathbf{f}}_i$, $[\tilde{\mathbf{p}}_1, \tilde{\mathbf{p}}_2, \ldots] = A^{-1}[\tilde{\mathbf{f}}_1, \tilde{\mathbf{f}}_2, \ldots]$
- Easy for incorporating seismic attenuation: complex-valued velocity
- Super-big matrix inversion requiring large storage in 3-D
- No way for data windowing and selective inversion
- $\rightarrow$ Laplacian-Fourier domain (Shin and Cha, 2009): Fourier domain with damped traces

Time domain: Convenient for data pre-processing: windowing, filtering, denoising

$$C(m) = \frac{1}{2}\sum_s\sum_r\int_0^T \mathrm{d}t|d_{cal}^{s,r}(m,t) - d_{obs}^{s,r}(t)|^2, d_{cal}^{s,r}(m,t) = R_{s,r}p \tag{34}$$

Lagrange functional with model parameter $m \equiv c$

$$L(m,p,\bar{p}) = C(m) + \langle \bar{p}, (\frac{1}{c^2}\partial_{tt} - \nabla^2)p - f\rangle_{\Omega\times[0,T]} \tag{35}$$

1. forward wave equation:

$$\frac{\partial L}{\partial \bar{p}} = 0 \Rightarrow (\frac{1}{c^2}\partial_{tt} - \nabla^2)p = f \tag{36}$$

2. adjoint state equation: different misfit C only change adjoint source

$$\frac{\partial L}{\partial p} = 0 \Rightarrow (\frac{1}{c^2}\partial_{tt} - \nabla^2)\bar{p} = -\frac{\partial C}{\partial p} = -R^\dagger(Rp - d_{obs}) \tag{37}$$

3. update the model based on the gradient

$$\frac{\partial L}{\partial m} = 0 \Rightarrow \nabla_m C = -\frac{2}{c^3}\int_0^T \mathrm{d}t\bar{p}\partial_{tt}p \tag{38}$$

SEISCOPE

1. Forward equation:

$$\begin{cases} \rho \partial_t \mathbf{v} = D\boldsymbol{\sigma} + \mathbf{f}_\mathbf{v} \\ \partial_t \boldsymbol{\sigma} = CD^T \mathbf{v} - C :: \Gamma \sum_{\ell=1}^{L} y_\ell \boldsymbol{\xi}_\ell + \mathbf{f}_\sigma \\ \partial_t \boldsymbol{\xi}_\ell + \omega_\ell \boldsymbol{\xi}_\ell = \omega_\ell D^T \mathbf{v}, \quad \ell = 1, \cdots, L, \end{cases} \tag{39}$$

2. Adjoint equation:

$$\begin{cases} \rho \partial_t \bar{\mathbf{v}} = D\bar{\boldsymbol{\sigma}} + \sum_{\ell=1}^{L} D\bar{\boldsymbol{\xi}}_\ell + \Delta d_\mathbf{v} \\ \partial_t \bar{\boldsymbol{\sigma}} = CD^T \bar{\mathbf{v}} + C\Delta d_\sigma \\ \partial_t \bar{\boldsymbol{\xi}}_\ell - \omega_\ell \bar{\boldsymbol{\xi}}_\ell = \omega_\ell y_\ell (C :: \Gamma) C^{-1} \bar{\boldsymbol{\sigma}}, \ell = 1, \cdots, L. \end{cases} \tag{40}$$

where

$$\Gamma = \begin{bmatrix} Q_{11}^{-1} & Q_{12}^{-1} & Q_{13}^{-1} & Q_{14}^{-1} & Q_{15}^{-1} & Q_{16}^{-1} \\ Q_{21}^{-1} & Q_{22}^{-1} & Q_{23}^{-1} & Q_{24}^{-1} & Q_{25}^{-1} & Q_{26}^{-1} \\ Q_{31}^{-1} & Q_{32}^{-1} & Q_{33}^{-1} & Q_{34}^{-1} & Q_{35}^{-1} & Q_{36}^{-1} \\ Q_{41}^{-1} & Q_{42}^{-1} & Q_{43}^{-1} & Q_{44}^{-1} & Q_{45}^{-1} & Q_{46}^{-1} \\ Q_{51}^{-1} & Q_{52}^{-1} & Q_{53}^{-1} & Q_{54}^{-1} & Q_{55}^{-1} & Q_{56}^{-1} \\ Q_{61}^{-1} & Q_{62}^{-1} & Q_{63}^{-1} & Q_{64}^{-1} & Q_{65}^{-1} & Q_{66}^{-1} \end{bmatrix} \tag{41}$$

3. Gradient expression (Yang et al., 2016a):

$$\frac{\partial \chi}{\partial \rho} = -\int_0^T \mathrm{d}t \bar{\mathbf{v}}^\dagger \partial_t \mathbf{v},$$

$$\frac{\partial \chi}{\partial C_{IJ}} = \int_0^T \mathrm{d}t \bar{\boldsymbol{\sigma}}^\dagger C^{-1} \frac{\partial C}{\partial C_{IJ}} C^{-1} (\partial_t \boldsymbol{\sigma} - \mathbf{f}_\sigma),$$

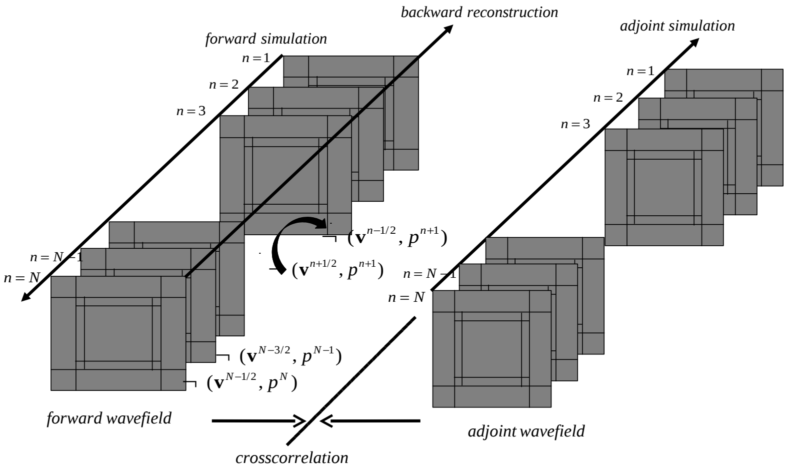
$$\frac{\partial \chi}{\partial Q_{IJ}^{-1}} = \int_0^T \mathrm{d}t \bar{\boldsymbol{\sigma}}^\dagger C^{-1} (C :: \frac{\partial \Gamma}{\partial Q_{IJ}^{-1}}) (\sum_{\ell=1}^{L} y_\ell \boldsymbol{\xi}_\ell), \frac{\partial \chi}{\partial Q_{IJ}} = -Q_{IJ}^{-2} \frac{\partial \chi}{\partial Q_{IJ}^{-1}}. \quad (42)$$

4. Constant-Q approximation by Least-squares method (Blanch et al., 1995; Yang et al., 2016a)

$$\min_{\gamma_\ell} \chi_1^Q, \quad \chi_1^Q = \int_{\omega_{\min}}^{\omega_{\max}} (\sum_{\ell=1}^{L} \gamma_\ell \frac{\omega_\ell \omega}{\omega^2 + \omega_\ell^2} - \gamma^{-1})^2 \mathrm{d}\omega. \quad (43)$$

in which

$$Y_\ell^{IJ}(\mathbf{x}) = y_\ell Q_{IJ}^{-1}(\mathbf{x}) \text{ with } y_\ell = \gamma\, \gamma_\ell. \quad (44)$$


backward reconstruction
forward simulation
adjoint simulation
n = 1
n = 2
n = 3
n = 1
n = 2
n = 3
n = N − 1
n = N
$(\mathbf{v}^{n-1/2}, p^{n+1})$
$(\mathbf{v}^{n+1/2}, p^{n+1})$
n = N − 1
n = N
$(\mathbf{v}^{N-3/2}, p^{N-1})$
$(\mathbf{v}^{N-1/2}, p^{N})$
forward wavefield
adjoint wavefield
crosscorrelation

It is mandatory to access incident wavefield backwards to build FWI gradient (or apply imaging condition in RTM)

1. reading the stored forward wavefield from the disk, even with data compression (Sun and Fu, 2013; Boehm et al., 2015)
   Reading disk memory is extremely slow for I/O intensive problems!
2. inferring from the final snapshots and the saved boundaries via reverse propagation (RP) (Tromp et al., 2005; Dussaud et al., 2008; Clapp, 2008; Yang et al., 2014): Efficient: Computation is much faster compared with disk accessing; may involve large 3D boundary storage, which can be significantly reduced using interpolation (Yang et al., 2016c);
3. remodeling using checkpointing (Griewank, 1992; Griewank and Walther, 2000; Symes, 2007; Anderson et al., 2012) from stored state to another state: Much better than disking reading, more expensive compared with reverse propagation due to repeated modeling!
4. checkpointing-assisted reverse-forward simulation (CARFS) =RP+checkpointing (Yang et al., 2016b) works in both attenuating and non-attenuating medium!

Reverse propagation based on the final snapshot and stored boundaries:

- forward modeling:

$$p^{n+1} = 2p^n - p^{n-1} + \Delta t^2 v^2 \nabla^2 p^n \quad (45)$$

  store boundary at each forward timestepping
- reverse propagation:

$$p^{n-1} = 2p^n - p^{n+1} + \Delta t^2 v^2 \nabla^2 p^n \quad (46)$$

  inject boundary at each backward timestepping

- Goal:
  1. understand the fundamental gadgets in Madagascar
  2. write codes within the framework of Madagascar
  3. reproduce nummerical experiments using SConstruct
- Method: FDM
- Steps:
  1. forward modeling with Clayton-Enquist ABC
  2. wavefield reconstruction by RP+stored boundaries
  3. full waveform inversion
- Where is the exercise?
  1. RSFSRC/book/xjtu/modeling2fwi
  2. https://yangpl.wordpress.com/activities/

- Seiscope consortium for financial support on the travel
- Community effort for developing Madagascar: a great tool
- Sergy Fomel for the invitation
- ETH Zurich for providing the sharing opportunity

Thanks for your attention!

Anderson, J. E., Tan, L., and Wang, D. (2012). Time-reversal checkpointing methods for RTM and FWI. *Geophysics*, 77:S93–S103.

Bérenger, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200.

Blanch, J., Robertson, J. O. A., and Symes, W. W. (1995). Modeling of a constant Q: Methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique. *Geophysics*, 60:176–184.

Boehm, C., Hanzich, M., de la Puente, J., and Fichtner, A. (2015). Wavefield compression for adjoint methods in full-waveform inversion. *submitted to Geophysics*.

Brossier, R. (2011). Two-dimensional frequency-domain visco-elastic full waveform inversion: Parallel algorithms, optimization and performance. *Computers & Geosciences*, 37(4):444 – 455.

Cerjan, C., Kosloff, D., Kosloff, R., and Reshef, M. (1985). A nonreflecting boundary condition for discrete acoustic and elastic wave equations. *Geophysics*, 50(4):2117–2131.

Chang, W.-F. and McMechan, G. A. (1986). Reverse-time migration of offset vertical seismic profiling data using the excitation-time imaging condition. *Geophysics*, 51(1):67–84.

Claerbout, J. (1971). Towards a unified theory of reflector mapping. *Geophysics*, 36:467–481.

Clapp, R. (2008). Reverse time migration: Saving the boundaries. Technical Report SEP-136, Stanford Exploration Project.

Clayton, R. and Engquist, B. (1977). Absorbing boundary conditions for acoustic and elastic wave equations. *Bulletin of the Seismological Society of America*, 67:1529–1540.

Dussaud, E., Symes, W. W., Williamson, P., Lemaistre, L., Singer, P., Denel, B., and Cherrett, A. (2008). Computational strategies for reverse-time migration. In *Society of Exploration Geophysics technical program expanded abstracts 2008*, pages 2267–2271. Society of Exploration Geophysicists.

Fomel, S., Ying, L., and Song, X. (2013). Seismic wave extrapolation using lowrank symbol approximation. *Geophysical Prospecting*, 61(3):526–536.

Griewank, A. (1992). Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation. *Optimization Methods and software*, 1(1):35–54.

Griewank, A. and Walther, A. (2000). Algorithm 799: Revolve: An implementation of checkpointing for the reverse or adjoint mode of computational differentiation. *ACM Trans. Math. Software*, 26:19–45.

Komatisch, D., Tromp, J., and Vilotte, J. P. (1998). The spectral element method for elastic wave equations: application to 2D and 3D seismic problems. In *Expanded abstracts*, volume II, pages 1460–1463.

Métivier, L., Bretaudeau, F., Brossier, R., Operto, S., and Virieux, J. (2014). Full waveform inversion and the truncated Newton method: quantitative imaging of complex subsurface structures. *Geophysical Prospecting*, 62:1353–1375.

Pica, A., Diet, J. P., and Tarantola, A. (1990). Nonlinear inversion of seismic reflection data in laterally invariant medium. *Geophysics*, 55(3):284–292.

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

Santosa, F. and Symes, W. W. (1988). Computation of the hessian for least-squares solutions of inverse problems of reflection seismology. *Inverse problems*, 4(1):211.

Sava, P. and Fomel, S. (2006). Time-shift imaging condition in seismic migration. *Geophysics*, 71(6):S209–S217.

Shin, C. and Cha, Y. H. (2009). Waveform inversion in the Laplace-Fourier domain. *Geophysical Journal International*, 177:1067–1079.

Shin, C., Jang, S., and Min, D. J. (2001). Improved amplitude preservation for prestack depth migration by inverse scattering theory. *Geophysical Prospecting*, 49:592–606.

Sun, W. and Fu, L.-Y. (2013). Two effective approaches to reduce data storage in reverse time migration. *Computers & Geosciences*, 56:69–75.

Symes, W. W. (2007). Reverse time migration with optimal checkpointing. *Geophysics*, 72(5):SM213–SM221.

Tromp, J., Tape, C., and Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160:195–216.

Yang, P., Brossier, R., Métivier, L., and Virieux, J. (2016a). A systematic formulation of 3D multiparameter full waveform inversion in viscoelastic medium. *submitted to Geophysical Journal International*.

Yang, P., Brossier, R., Métivier, L., and Virieux, J. (2016b). Wavefield reconstruction in attenuating media: A checkpointing-assisted reverse-forward simulation method. *submitted to Geophysics*.

Yang, P., Brossier, R., and Virieux, J. (2016c). Wavefield reconstruction from significantly decimated boundaries. *Geophysics, in press*, 80(5):1–13.

Yang, P., Gao, J., and Wang, B. (2014). RTM using effective boundary saving: A staggered grid GPU implementation. *Computers & Geosciences*, 68:64–72.