



HAL
open science

Numerical Scheme Impacts on Time Domain Full Waveform Inversion

Pierre Jacquet, Andreas Atle, H el ene Barucq, Henri Calandra, Julien Diaz

► **To cite this version:**

Pierre Jacquet, Andreas Atle, H el ene Barucq, Henri Calandra, Julien Diaz. Numerical Scheme Impacts on Time Domain Full Waveform Inversion. MATHIAS 2019, Oct 2019, Serris, France. hal-02422842

HAL Id: hal-02422842

<https://hal.science/hal-02422842>

Submitted on 23 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

Numerical Scheme Impacts on Time Domain Full Waveform Inversion

Mathias 2019

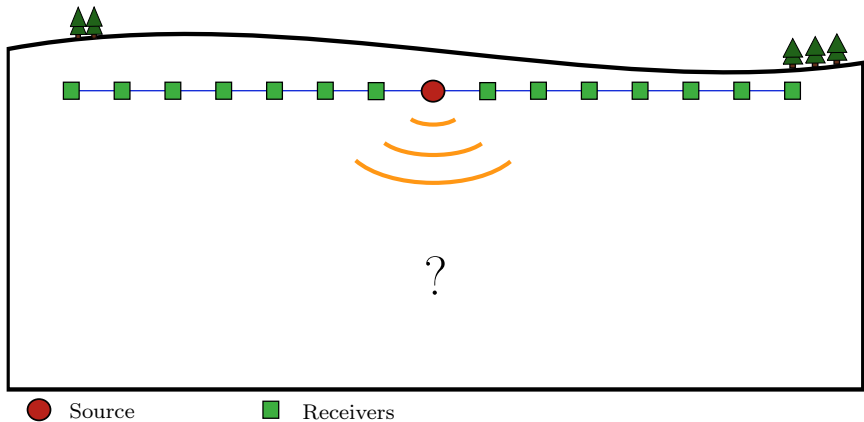
Pierre Jacquet

pierre.jacquet@inria.fr

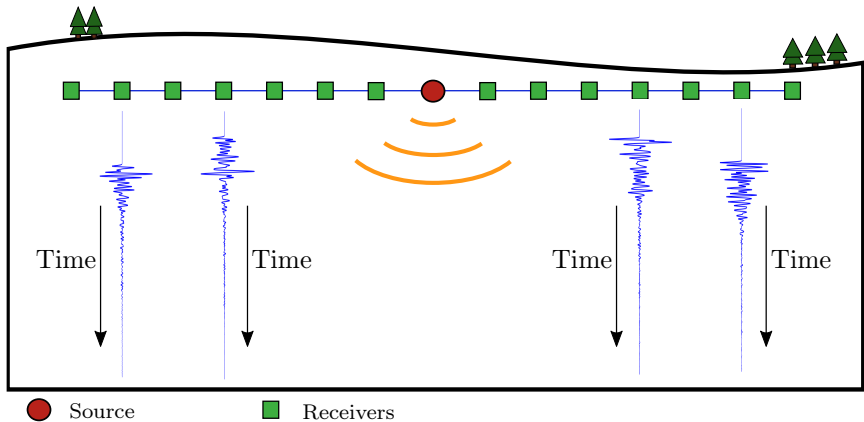
Atle Andreas, Barucq H el ene, Calandra Henri, Diaz Julien
Second year PhD Student
Inria - Magique 3D - DIP

Pau, FRANCE

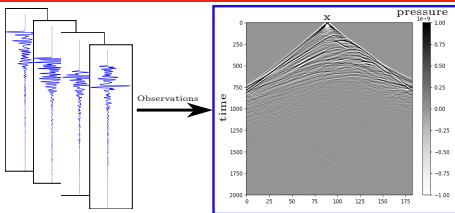




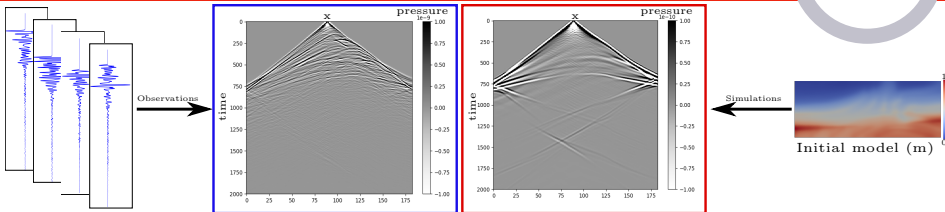
Seismic Acquisition

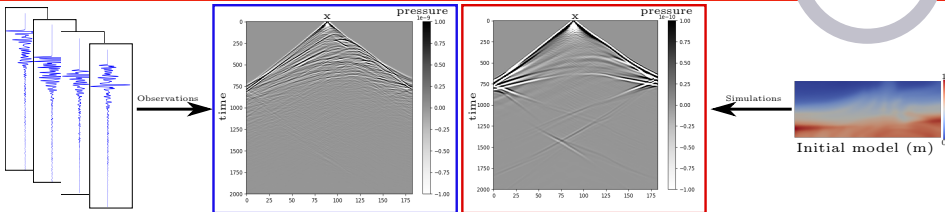


FWI Workflow



FWI Workflow





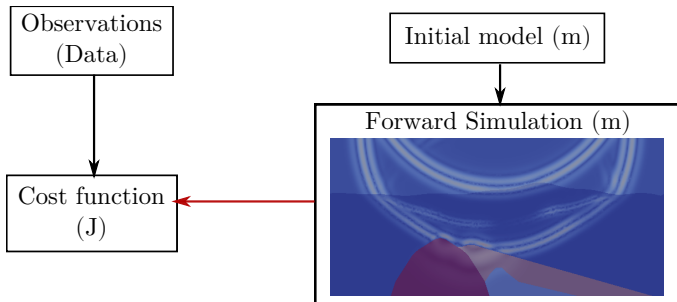
Cost function to minimize :

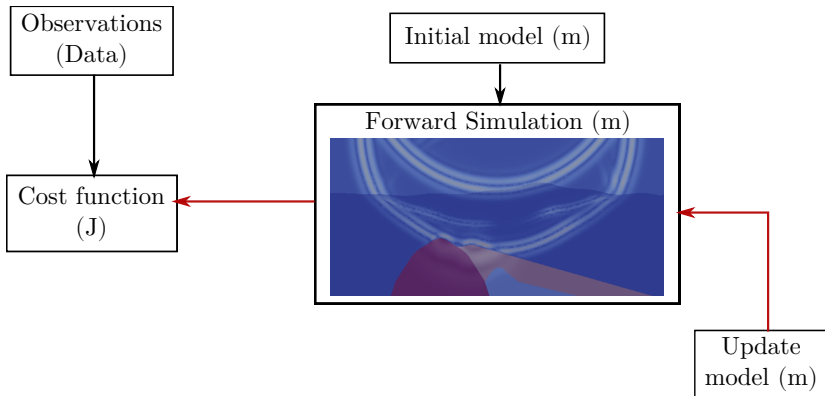
$$\mathcal{J}(\mathbf{m}) = \frac{1}{2} \|\mathbf{d}_{obs} - \mathcal{F}(\mathbf{m})\|^2$$

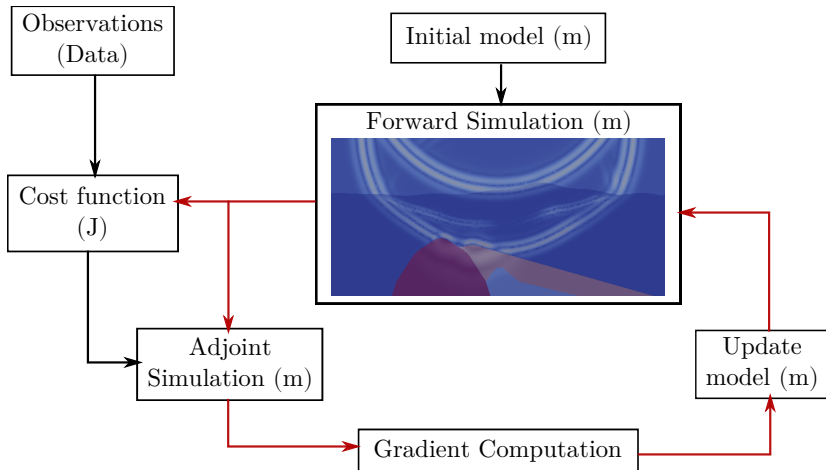
- ▶ $\mathcal{F}(\mathbf{m})$ is the restriction on the receivers of the simulated waves in the medium \mathbf{m} . (With $\mathbf{m} = \mathbf{c}, \rho, \kappa \dots$)
- ▶ FWI iterates until $\mathcal{J}(\mathbf{m}) \rightarrow 0$

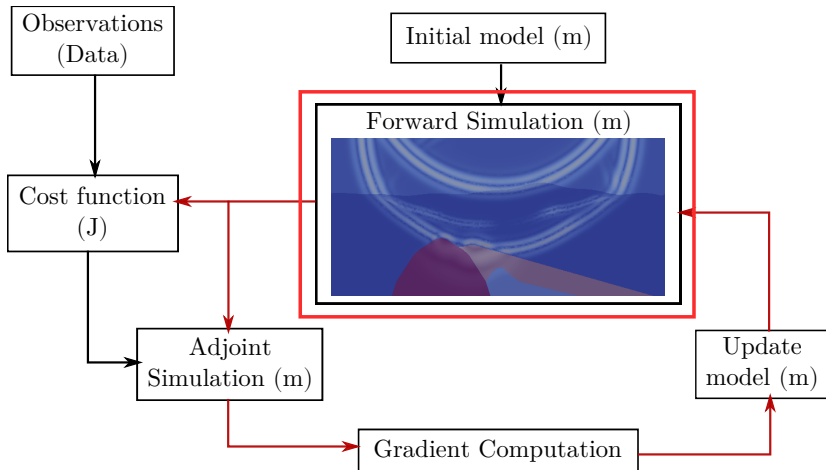
[1] Patrick Lailly
The seismic inverse problem as a sequence of before stack migrations
Conference on Inverse Scattering

[2] Albert Tarantola
Inversion of seismic reflection data in the acoustic approximation
Geophysics, Vol. 49, 1984



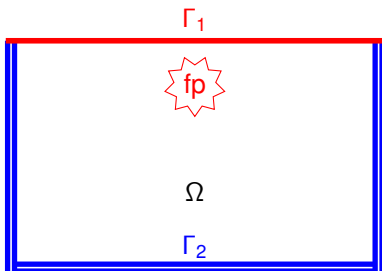






First order acoustic wave equation

$$\left\{ \begin{array}{l} \frac{1}{\rho c^2} \frac{\partial \mathbf{p}}{\partial t} + \nabla \cdot \mathbf{v} = f_p \quad \text{on } \Omega \\ \rho \frac{\partial \mathbf{v}}{\partial t} + \nabla \mathbf{p} = 0 \quad \text{on } \Omega \\ \mathbf{p} = 0 \quad \text{on } \Gamma_1 \\ \frac{\partial \mathbf{p}}{\partial t} + \mathbf{c} \nabla \mathbf{p} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_2 \\ \mathbf{p}(0) = 0, \quad \mathbf{v}(0) = 0 \end{array} \right.$$



Domain with Absorbing Boundary Conditions

Space Discretization : Discontinuous Galerkin Elements

- ▶ Nodal (Lagrangian / Jacobian)
- ▶ Modal (Bernstein-Bézier)

Space Discretization :
Discontinuous Galerkin Elements

- ▶ Nodal (Lagrangian / Jacobian)
- ▶ Modal (Bernstein-Bézier)

Semi-discretized model :

$$\frac{\partial}{\partial t} \bar{\mathbf{U}}(t) = A\bar{\mathbf{U}}(t) + \bar{\mathbf{F}}(t)$$

with :

$$\bar{\mathbf{U}}(t) = \begin{pmatrix} \bar{\mathbf{P}}(t) \\ \bar{\mathbf{V}}(t) \end{pmatrix}$$

Space Discretization :
Discontinuous Galerkin Elements

- ▶ Nodal (Lagrangian / Jacobian)
- ▶ Modal (Bernstein-Bézier)

Time schemes :

- ▶ Runge Kutta 2/4
- ▶ Adams Bashforth 3

Semi-discretized model :

$$\frac{\partial}{\partial t} \bar{\mathbf{U}}(t) = A\bar{\mathbf{U}}(t) + \bar{\mathbf{F}}(t)$$

with :

$$\bar{\mathbf{U}}(t) = \begin{pmatrix} \bar{\mathbf{P}}(t) \\ \bar{\mathbf{V}}(t) \end{pmatrix}$$

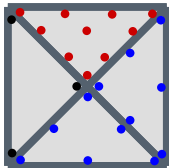


Assets of Discontinuous Galerkin Methods :

- ▶ Unstructured grid (enable to match the topography and medium irregularities)
- ▶ Robust to physical discontinuities
- ▶ hp-adaptivity
- ▶ Massively parallel performance properties



h-adaptivity



p-adaptivity with P1,
P2, P3 elements

Time Domain Full Waveform Inversion

Seismic Acquisition

FWI Workflow

Forward Discretization

Some Results

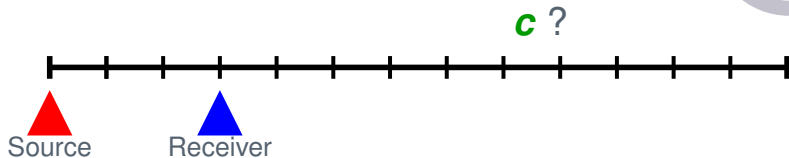
1D Results

2D Time Domain FWI Results

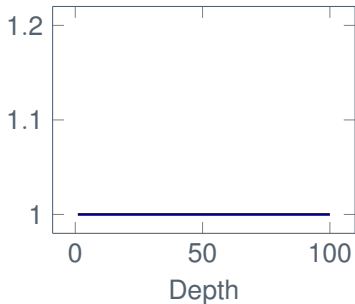
2D Multiscale Reconstruction

1D Preliminary tests

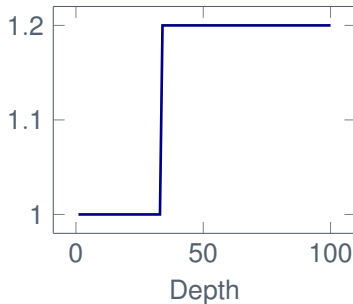
One year ago...



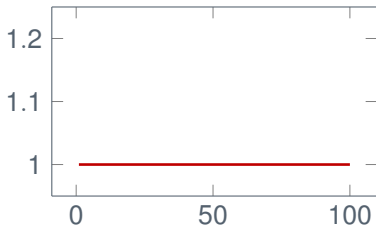
Initial c Model



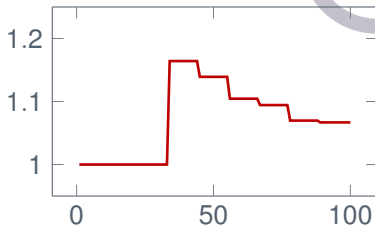
Target c Model



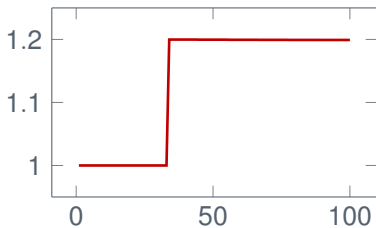
FWI 1D Results



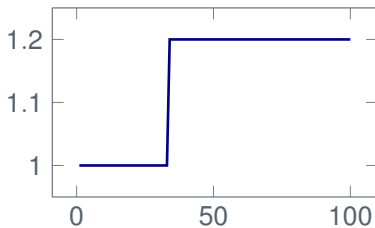
Initial c model



Intermediate c model (iter=20)



c model Reconstructed (iter=50)



Target c model

2D FWI :

- ▶ Developed in Total environnement (DIP¹)
- ▶ Nodal Space Operators (Lagrangian/Jacobian)
- ▶ Modal Space Operators (Bernstein-Bézier)
- ▶ Runge Kutta 2/4 and Adams Bashforth 3 time-schemes

Gradient expression :

$$\nabla_{\frac{1}{\kappa}} \mathcal{J} = \int_0^T \int_{\Omega} \frac{\partial \mathbf{p}}{\partial t} \lambda_1 d\Omega dt \quad \text{with : } \kappa = \rho \mathbf{c}^2$$

\mathbf{c} , ρ and κ Constant per elements

¹<http://dip.inria.fr/>

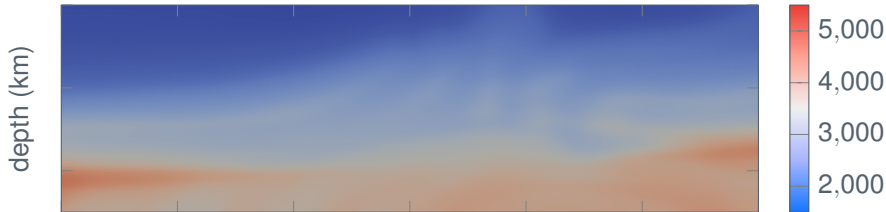
2D Time Domain FWI Reconstructions

Time-schemes comparison



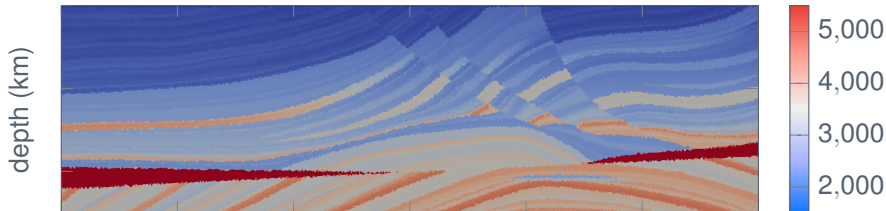
Initial **c** Model

$m \cdot s^{-1}$



Target **c** Model

$m \cdot s^{-1}$



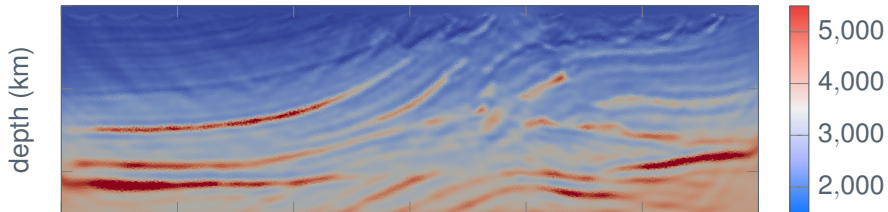
2D Time Domain FWI Reconstructions

Time-schemes comparison



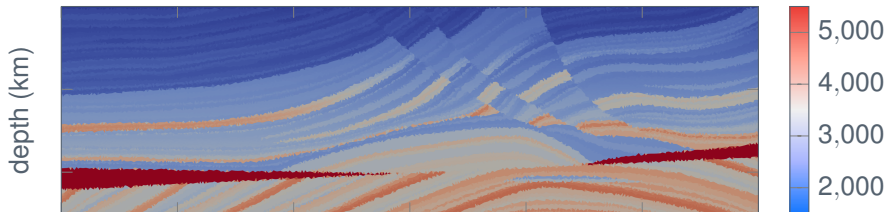
RK2 Reconstructed **c** Model (30 iterations)

$m \cdot s^{-1}$



Target **c** Model

$m \cdot s^{-1}$



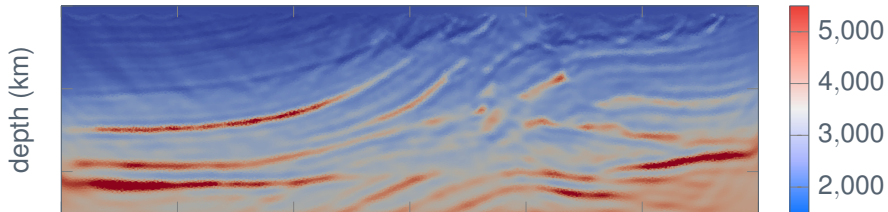
2D Time Domain FWI Reconstructions

Time-schemes comparison



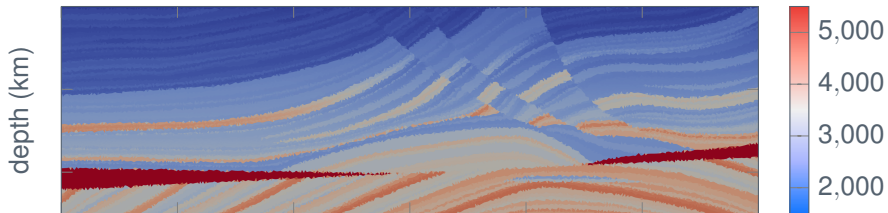
RK4 Reconstructed **c** Model (30 iterations)

$m \cdot s^{-1}$



Target **c** Model

$m \cdot s^{-1}$



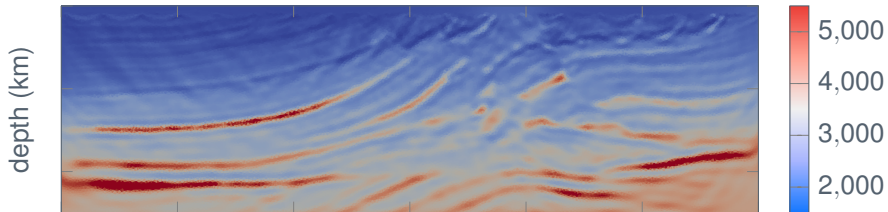
2D Time Domain FWI Reconstructions

Time-schemes comparison



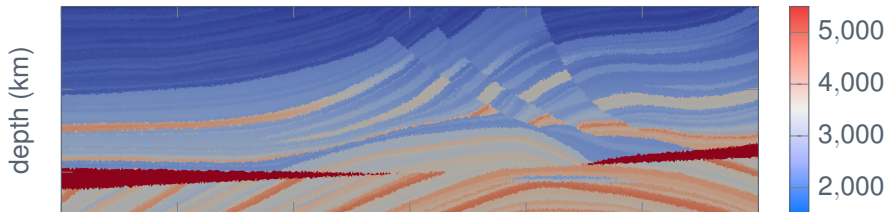
AB3 Reconstructed **c** Model (30 iterations)

$m \cdot s^{-1}$



Target **c** Model

$m \cdot s^{-1}$



2D Time Domain FWI Reconstructions

Time-Schemes Comparison

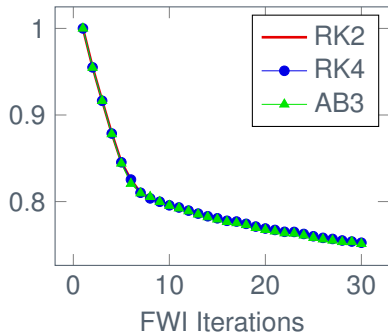


- ▶ 47k P1 elements
- ▶ Constant ρ model ($\rho = 1$)
- ▶ 19 sources / 181 Receivers
- ▶ Noise : SNR=10
- ▶ 30 iterations
- ▶ 120 cores
- ▶ Polynomial basis : Nodal

Computational time :

- ▶ RK2 : **3h15**
- ▶ RK4 : **4h30**
- ▶ AB3 : **5h10**

Cost function evolution :



2D Time Domain FWI Reconstructions

Nodal/Modal Comparison

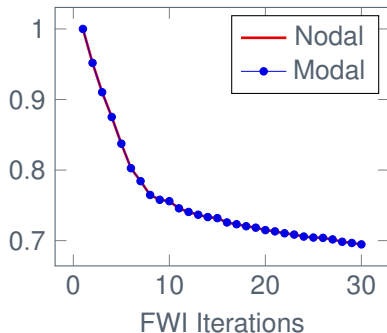


- ▶ 47k P1 elements
- ▶ Time Scheme : RK2
- ▶ Constant ρ model ($\rho = 1$)
- ▶ 19 sources / 181 Receivers
- ▶ Noise : SNR=10
- ▶ 30 iterations
- ▶ 120 cores

Computational time :

- ▶ Nodal : **3h15**
- ▶ Modal: **4h30**^[1]

Cost function evolution :



[1] Chan J. and Warburton T.
GPU-Accelerated Bernstein Bézier Discontinuous Galerkin Methods for Wave Problems
SIAM Journal on Scientific Computing 2017

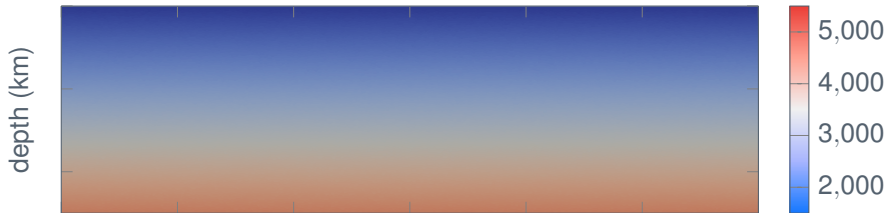
2D Multiscale Reconstructions

Reconstruction with an initial smooth model



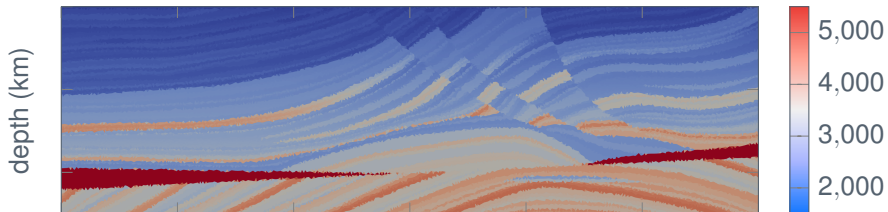
Initial **c** Model

$m \cdot s^{-1}$



Target **c** Model

$m \cdot s^{-1}$



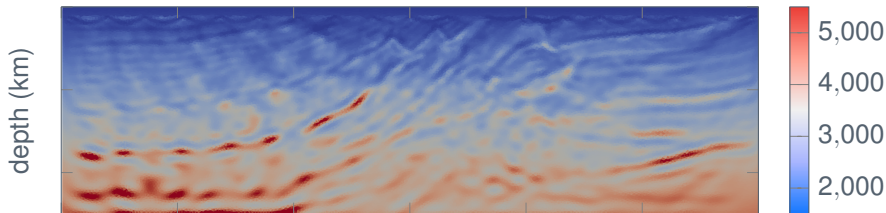
2D Multiscale Reconstructions

Reconstruction with an initial smooth model



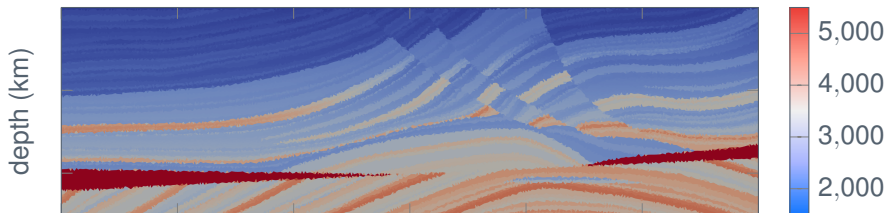
Reconstructed model **c** Model (30 iterations RK2)

$m \cdot s^{-1}$



Target **c** Model

$m \cdot s^{-1}$



2D Multiscale Reconstructions

Multiscale Principle [1]

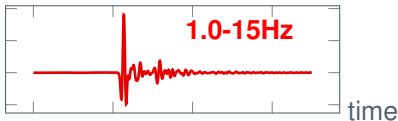
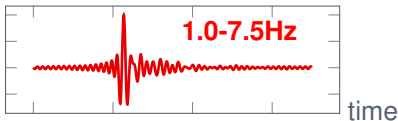
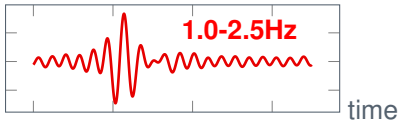


Low Frequencies
↔
Reconstruct **coarse** structures

High Frequencies
↔
Reconstruct **small** structures

Filtered Traces :

p



[1] C. Bunks, F. M. Saleck, S. Zaleski, and G. Chavent
Multiscale seismic waveform inversion
GEOPHYSICS, Vol. 60, No. 5, 1995

2D Multiscale Reconstructions

Multiscale Principle [1]



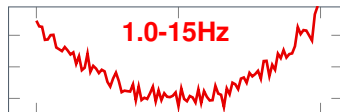
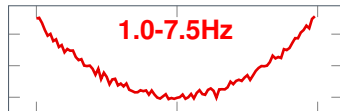
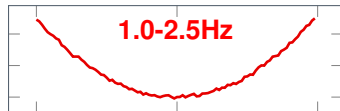
Heuristic Illustration :

\mathcal{J}

Low Frequencies



Reconstruct **coarse** structures



High Frequencies



Reconstruct **small** structures

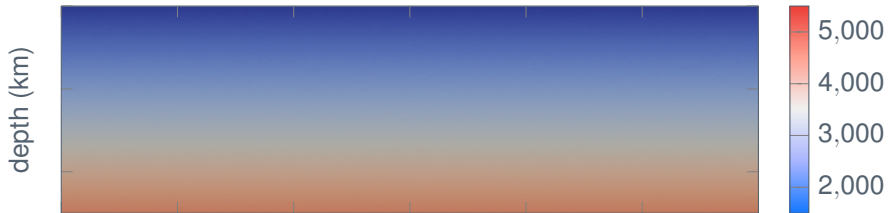
[1] C. Bunks, F. M. Saleck, S. Zaleski, and G. Chavent
Multiscale seismic waveform inversion
GEOPHYSICS, Vol. 60, No. 5, 1995

2D Multiscale Reconstructions

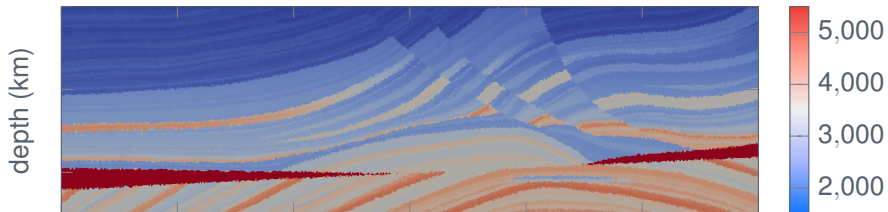
Reconstruction with an initial smooth model



Initial **c** Model



Target **c** Model



2D Multiscale Reconstructions

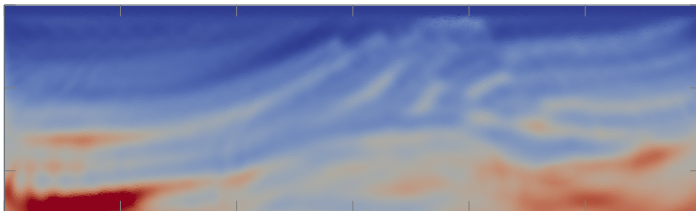
Reconstruction with an initial smooth model



Reconstructed **c** Model with 1.0-2.5Hz filter

$m \cdot s^{-1}$

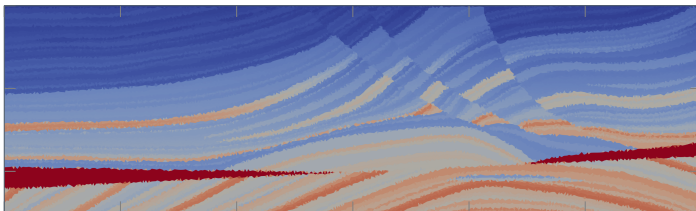
depth (km)



Target **c** Model

$m \cdot s^{-1}$

depth (km)



2D Multiscale Reconstructions

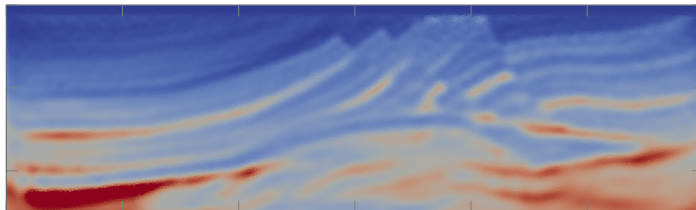
Reconstruction with an initial smooth model



Reconstructed **c** Model with 1.0-7.5Hz filter

$m \cdot s^{-1}$

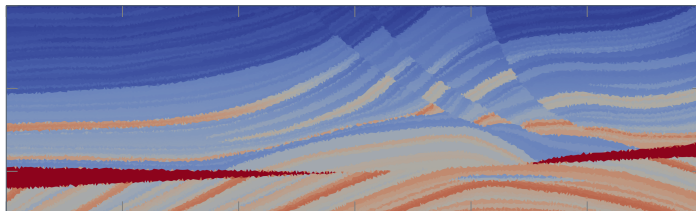
depth (km)



Target **c** Model

$m \cdot s^{-1}$

depth (km)



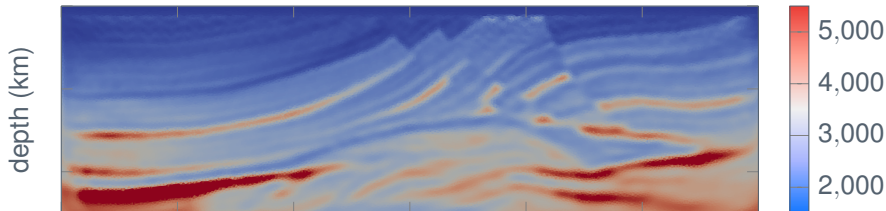
2D Multiscale Reconstructions

Reconstruction with an initial smooth model



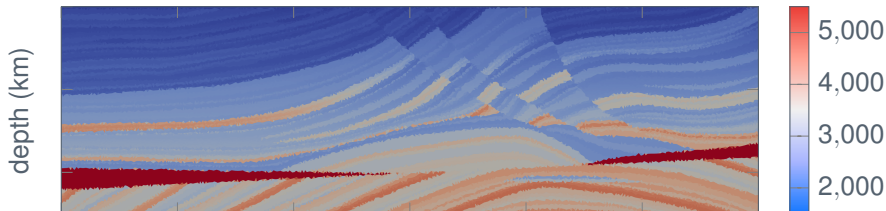
Reconstructed **c** Model with 1.0-10Hz filter

$m \cdot s^{-1}$



Target **c** Model

$m \cdot s^{-1}$



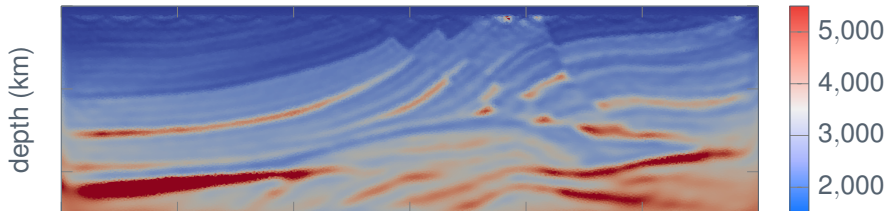
2D Multiscale Reconstructions

Reconstruction with an initial smooth model



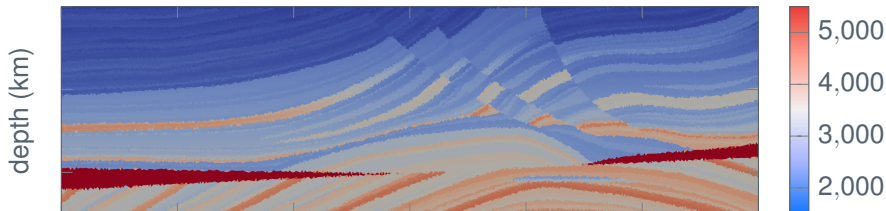
Reconstructed **c** Model with 1.0-15Hz filter

$m \cdot s^{-1}$



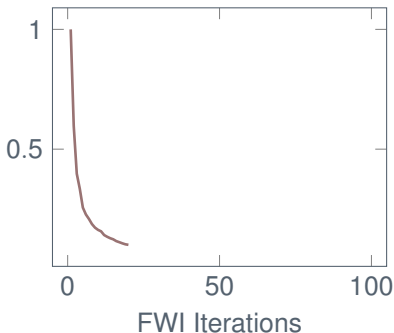
Target **c** Model

$m \cdot s^{-1}$



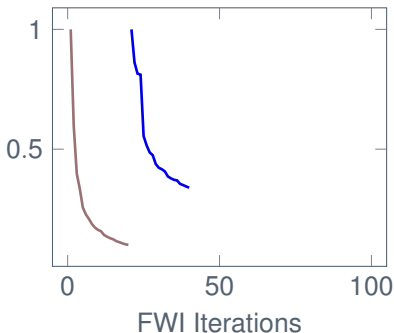
- ▶ 47k P1 elements
- ▶ Time Scheme : RK2
- ▶ Constant ρ model ($\rho = 1$)
- ▶ 19 sources / 181 Receivers
- ▶ Noise : SNR=10
- ▶ 120 cores
- ▶ 20 FWI iterations per filter
- ▶ Computation time : 10h
- ▶ Frequencies : 1-2.5Hz

Cost function evolution :



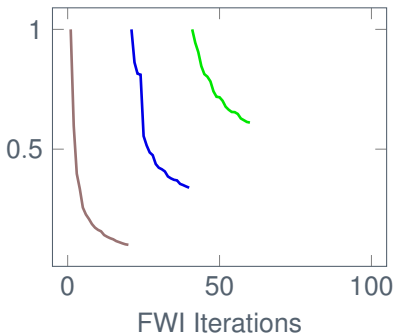
- ▶ 47k P1 elements
- ▶ Time Scheme : RK2
- ▶ Constant ρ model ($\rho = 1$)
- ▶ 19 sources / 181 Receivers
- ▶ Noise : SNR=10
- ▶ 120 cores
- ▶ 20 FWI iterations per filter
- ▶ Computation time : 10h
- ▶ Frequencies : 1-2.5Hz,
1-5.0Hz

Cost function evolution :



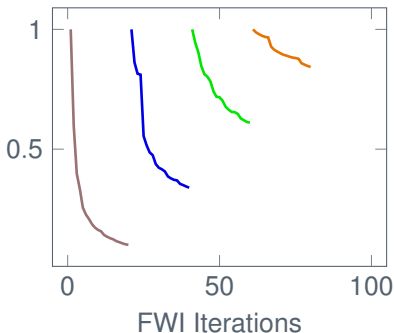
- ▶ 47k P1 elements
- ▶ Time Scheme : RK2
- ▶ Constant ρ model ($\rho = 1$)
- ▶ 19 sources / 181 Receivers
- ▶ Noise : SNR=10
- ▶ 120 cores
- ▶ 20 FWI iterations per filter
- ▶ Computation time : 10h
- ▶ Frequencies : 1-2.5Hz,
1-5.0Hz, 1-7.5Hz

Cost function evolution :



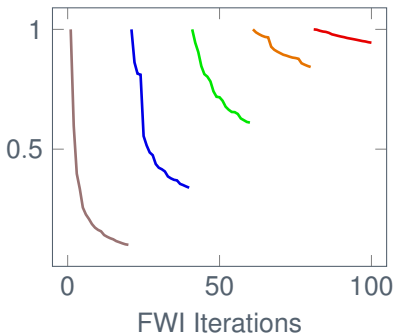
- ▶ 47k P1 elements
- ▶ Time Scheme : RK2
- ▶ Constant ρ model ($\rho = 1$)
- ▶ 19 sources / 181 Receivers
- ▶ Noise : SNR=10
- ▶ 120 cores
- ▶ 20 FWI iterations per filter
- ▶ Computation time : 10h
- ▶ Frequencies : 1-2.5Hz,
1-5.0Hz, 1-7.5Hz, 1-10Hz

Cost function evolution :



- ▶ 47k P1 elements
- ▶ Time Scheme : RK2
- ▶ Constant ρ model ($\rho = 1$)
- ▶ 19 sources / 181 Receivers
- ▶ Noise : SNR=10
- ▶ 120 cores
- ▶ 20 FWI iterations per filter
- ▶ Computation time : 10h
- ▶ Frequencies : 1-2.5Hz,
1-5.0Hz, 1-7.5Hz, 1-10Hz,
1-15Hz

Cost function evolution :



Main Results :

- ▶ 2D Acoustic Reconstruction performed with different discretization
- ▶ Multiscale FWI implemented and working on Marmousi

Main Results :

- ▶ 2D Acoustic Reconstruction performed with different discretization
- ▶ Multiscale FWI implemented and working on Marmousi

Perspectives :

- ▶ Perform reconstruction on other test cases (2D/3D)
- ▶ Develop enhanced optimizers (NLCG, Limited BFGS)
- ▶ Adapt the code to use High order Model
- ▶ Extend the code to elastic and elasto-acoustic propagator
- ▶ Exploit coupled numerical method (SEM/DG) (Aurélien Citrain Thesis)

Main Results :

- ▶ 2D Acoustic Reconstruction performed with different discretization
- ▶ Multiscale FWI implemented and working on Marmousi

Perspectives :

- ▶ Perform reconstruction on other test cases (2D/3D)
- ▶ Develop enhanced optimizers (NLCG, Limited BFGS)
- ▶ Adapt the code to use High order Model
- ▶ Extend the code to elastic and elasto-acoustic propagator
- ▶ Exploit coupled numerical method (SEM/DG) (Aurélien Citrain Thesis)

Thank you.



Lagrangian functional [1] :

$$\mathcal{L}(\hat{\mathbf{u}}, \hat{\boldsymbol{\lambda}}, \mathbf{m}) = \frac{1}{2} \|d_{obs} - \mathcal{R}(\hat{\mathbf{u}})\|^2 + \langle Forward_{\mathbf{m}}(\hat{\mathbf{u}}) - f_p, \hat{\boldsymbol{\lambda}} \rangle$$

If $\hat{\mathbf{u}} = \mathbf{u}$ Solution of the Direct Problem $\iff (Forward_{\mathbf{m}}(\mathbf{u}) - f_p = 0)$:

$$\mathcal{J}(\mathbf{m}) = \mathcal{L}(\mathbf{u}, \hat{\boldsymbol{\lambda}}, \mathbf{m})$$

[1] Plessix R-E

A review of the adjoint-state method for computing the gradient of a functional with geophysical applications
Geophysical Journal International, Volume 167, Issue 2, 2006

Lagrangian functional [1] :

$$\mathcal{L}(\hat{\mathbf{u}}, \hat{\boldsymbol{\lambda}}, \mathbf{m}) = \frac{1}{2} \|d_{obs} - \mathcal{R}(\hat{\mathbf{u}})\|^2 + \langle Forward_{\mathbf{m}}(\hat{\mathbf{u}}) - f_p, \hat{\boldsymbol{\lambda}} \rangle$$

If $\hat{\mathbf{u}} = \mathbf{u}$ Solution of the Direct Problem $\iff (Forward_{\mathbf{m}}(\mathbf{u}) - f_p = 0)$:

$$\mathcal{J}(\mathbf{m}) = \mathcal{L}(\mathbf{u}, \hat{\boldsymbol{\lambda}}, \mathbf{m})$$

Let us choose $\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}$ such as $\frac{\partial \mathcal{L}}{\partial \mathbf{u}} = 0$

$$(\mathcal{R}^* d_{obs} - \mathbf{u}) + Forward_{\mathbf{m}}^*(\boldsymbol{\lambda}) = 0$$

[1] Plessix R-E
A review of the adjoint-state method for computing the gradient of a functional with geophysical applications
Geophysical Journal International, Volume 167, Issue 2, 2006

Lagrangian functional [1] :

$$\mathcal{L}(\hat{\mathbf{u}}, \hat{\boldsymbol{\lambda}}, \mathbf{m}) = \frac{1}{2} \|d_{obs} - \mathcal{R}(\hat{\mathbf{u}})\|^2 + \langle Forward_{\mathbf{m}}(\hat{\mathbf{u}}) - f_p, \hat{\boldsymbol{\lambda}} \rangle$$

If $\hat{\mathbf{u}} = \mathbf{u}$ Solution of the Direct Problem $\iff (Forward_{\mathbf{m}}(\mathbf{u}) - f_p = 0)$:

$$\mathcal{J}(\mathbf{m}) = \mathcal{L}(\mathbf{u}, \hat{\boldsymbol{\lambda}}, \mathbf{m})$$

Let us choose $\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}$ such as $\frac{\partial \mathcal{L}}{\partial \mathbf{u}} = 0$

$$(\mathcal{R}^* d_{obs} - \mathbf{u}) + Forward_{\mathbf{m}}^*(\boldsymbol{\lambda}) = 0$$

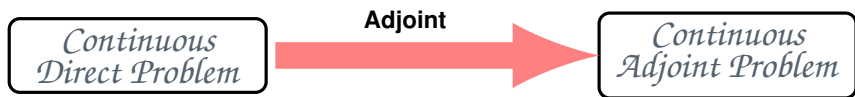
For $Forward_{\mathbf{m}}(\mathbf{u}) - f_p = 0$:

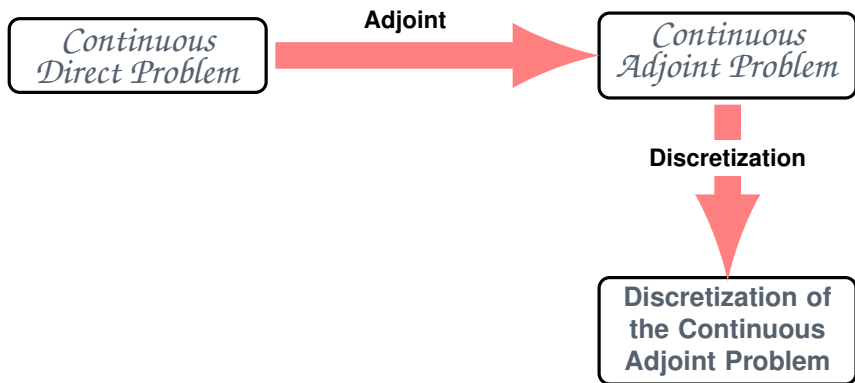
$$\partial_{\mathbf{m}_i} \mathcal{J}(\mathbf{m}) = \partial_{\mathbf{m}_i} \mathcal{L}(\mathbf{u}, \boldsymbol{\lambda}, \mathbf{m}) = \partial_{\mathbf{m}_i} \langle Forward_{\mathbf{m}}(\mathbf{u}), \boldsymbol{\lambda} \rangle$$

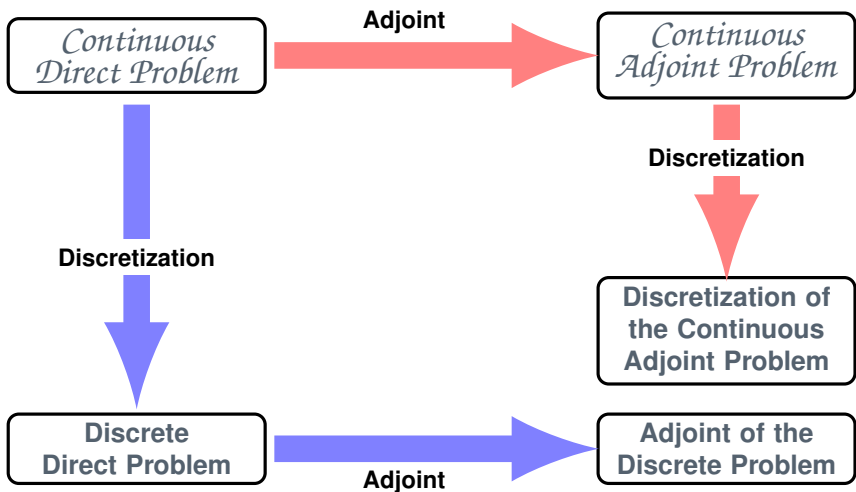
[1] Plessix R-E

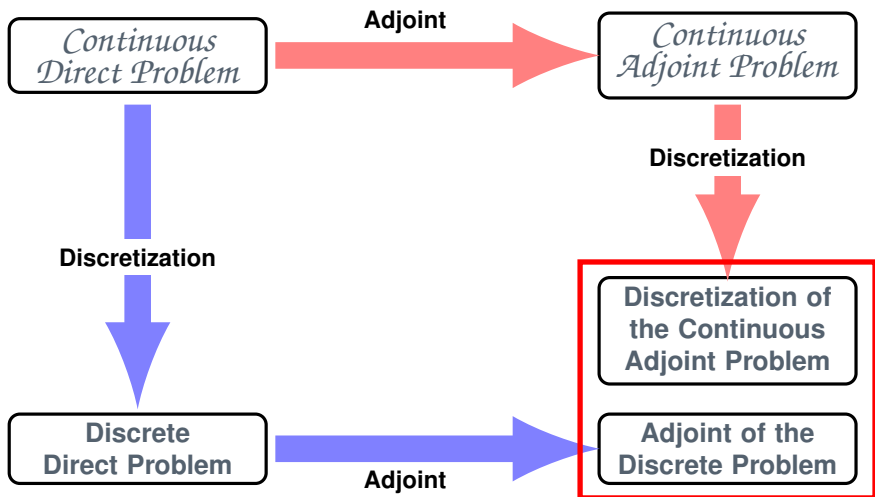
A review of the adjoint-state method for computing the gradient of a functional with geophysical applications
Geophysical Journal International, Volume 167, Issue 2, 2006

*Continuous
Direct Problem*









$$\mathcal{J}(\mathbf{p}) = \frac{1}{2} \|\mathbf{d}_{obs} - R\mathbf{p}\|^2$$

$$\left\{ \begin{array}{l} \frac{1}{\rho c^2} \frac{\partial \mathbf{p}}{\partial t} + \nabla \cdot \mathbf{v} = f_p \quad \text{on } \Omega \\ \rho \frac{\partial \mathbf{v}}{\partial t} + \nabla \mathbf{p} = 0 \quad \text{on } \Omega \\ \mathbf{p} = 0 \quad \text{on } \Gamma_1 \\ \frac{\partial \mathbf{p}}{\partial t} + \mathbf{c} \nabla \mathbf{p} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_2 \\ \mathbf{p}(0) = 0, \quad \mathbf{v}(0) = 0 \end{array} \right.$$

$$t \in [0, T]$$

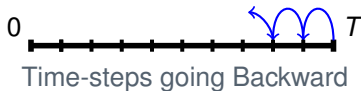
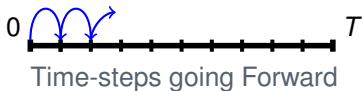
$$\left\{ \begin{array}{l} \frac{1}{\rho c^2} \frac{\partial \lambda_1}{\partial t} + \nabla \cdot \lambda_2 = \frac{\partial \mathcal{J}}{\partial \mathbf{p}} \quad \text{on } \Omega \\ \rho \frac{\partial \lambda_2}{\partial t} + \nabla \lambda_1 = 0 \quad \text{on } \Omega \\ \lambda_1 = 0 \quad \text{on } \Gamma_1 \\ \frac{\partial \lambda_1}{\partial t} - \mathbf{c} \nabla \lambda_1 \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_2 \\ \lambda_1(T) = 0, \quad \lambda_2(T) = 0 \end{array} \right.$$

$$t \in [T, 0]$$

$$\mathcal{J}(\mathbf{p}) = \frac{1}{2} \|d_{obs} - R\mathbf{p}\|^2$$

$$\left\{ \begin{array}{l} \frac{\partial \bar{\mathbf{U}}^n}{\partial t} = A\bar{\mathbf{U}}^n + \bar{\mathbf{F}}^n \\ \text{With : } \bar{\mathbf{U}}^n = \begin{pmatrix} \bar{\mathbf{P}}^n \\ \bar{\mathbf{V}}^n \end{pmatrix} \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{\partial \bar{\boldsymbol{\Lambda}}^n}{\partial t} = A\bar{\boldsymbol{\Lambda}}^n + R^*(R\bar{\mathbf{U}}^n - d_{obs}) \\ \text{With : } \bar{\boldsymbol{\Lambda}}^n = \begin{pmatrix} \bar{\boldsymbol{\Lambda}}_1^n \\ \bar{\boldsymbol{\Lambda}}_2^n \end{pmatrix} \end{array} \right.$$



DtA : Discretize then Adjoint Strategy

Example With RK4



All time scheme can be summed-up such as :

$$L\bar{U} = E\bar{F}$$

RK4 time-scheme leads to :

$$\bar{U}^{n+1} = B\bar{U}^n + C_0\bar{F}^n + C_{\frac{1}{2}}\bar{F}^{n+\frac{1}{2}} + C_1\bar{F}^{n+1}$$

$$L\bar{U} = E\bar{F} = \bar{G}$$

$$\begin{pmatrix} I & & & & & \\ -B & I & & & & \\ & -B & I & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & -B & I \end{pmatrix} \begin{pmatrix} \bar{U}^0 \\ \bar{U}^1 \\ \bar{U}^2 \\ \vdots \\ \bar{U}^n \end{pmatrix} = \begin{pmatrix} \bar{G}^0 \\ \bar{G}^1 \\ \bar{G}^2 \\ \vdots \\ \bar{G}^n \end{pmatrix}$$

All time scheme can be summed-up such as :

$$\mathbf{L}\bar{\mathbf{U}} = \mathbf{E}\bar{\mathbf{F}}$$

We are looking for a Discrete Adjoint state satisfying :

$$\mathbf{L}^*\bar{\boldsymbol{\Lambda}} = -\mathbf{R}^*(d_{obs} - \mathbf{R}\bar{\mathbf{U}})$$

With the adjoint operator \mathbf{L}^* satisfying :

$$\langle \mathbf{L}\bar{\mathbf{U}}, \bar{\boldsymbol{\Lambda}} \rangle = \langle \bar{\mathbf{U}}, \mathbf{L}^*\bar{\boldsymbol{\Lambda}} \rangle$$

All time scheme can be summed-up such as :

$$\mathbf{L}\bar{\mathbf{U}} = \mathbf{E}\bar{\mathbf{F}} = \bar{\mathbf{G}}$$

We are looking for a Discrete Adjoint state satisfying :

$$\mathbf{L}^*\bar{\boldsymbol{\lambda}} = -\mathbf{R}^*(d_{obs} - \mathbf{R}\bar{\mathbf{U}}) = \bar{\mathbf{D}}$$

With the adjoint operator \mathbf{L}^* satisfying :

$$\langle \mathbf{L}\bar{\mathbf{U}}, \bar{\boldsymbol{\lambda}} \rangle = \langle \bar{\mathbf{U}}, \mathbf{L}^*\bar{\boldsymbol{\lambda}} \rangle$$

$$\langle \bar{\mathbf{G}}, \bar{\boldsymbol{\lambda}} \rangle = \langle \bar{\mathbf{U}}, \bar{\mathbf{D}} \rangle \quad (\text{Adjoint Test})$$

Adjoint test succeeds \iff operator \mathbf{L}^* well established

DtA : Discretize then Adjoint Strategy

Example with RK4



RK4 time-scheme leads to :

$$\bar{U}^{n+1} = B\bar{U}^n + C_0\bar{F}^n + C_{\frac{1}{2}}\bar{F}^{n+\frac{1}{2}} + C_1\bar{F}^{n+1}$$

$$L\bar{U} = E\bar{F} = \bar{G}$$
$$\begin{pmatrix} I & & & & \\ -B & I & & & \\ & -B & I & & \\ & & \ddots & \ddots & \\ & & & -B & I \end{pmatrix} \begin{pmatrix} \bar{U}^0 \\ \bar{U}^1 \\ \bar{U}^2 \\ \vdots \\ \bar{U}^n \end{pmatrix} = \begin{pmatrix} \bar{G}^0 \\ \bar{G}^1 \\ \bar{G}^2 \\ \vdots \\ \bar{G}^n \end{pmatrix}$$

So :

$$L^* = \begin{pmatrix} I & -B^* & & & \\ & I & -B^* & & \\ & & \ddots & \ddots & \\ & & & I & -B^* \\ & & & & I \end{pmatrix}$$

Adjoint Then Discretize

- + Physical approach
- + Same discrete operators for Forward and Backward
- - Approximate gradient [1]
- Consistent with the discretization

Discretize then Adjoint

- + Numerical approach
- + Has an Adjoint Test
- Tremendous work to develop the adjoint operators
- Non-consistency of the adjoint state [2]

[1] Sirkes, Ziv and Tziperman, Eli
Finite Difference of Adjoint or Adjoint of Finite Difference ?
1997

[2] Sei Alain and Symes William
A Note on Consistency and Adjointness for Numerical Schemes
1997

Adjoint Then Discretize

- + Physical approach
- + Same discrete operators for Forward and Backward
- - Approximate gradient [1]
- Consistent with the discretization

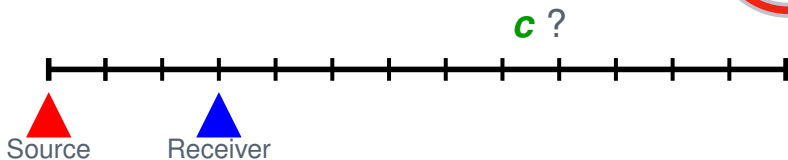
Discretize then Adjoint

- + Numerical approach
- + / - Has an Adjoint Test (**in theory**)
- Tremendous work to develop the adjoint operators
- Non-consistency of the adjoint state [2]

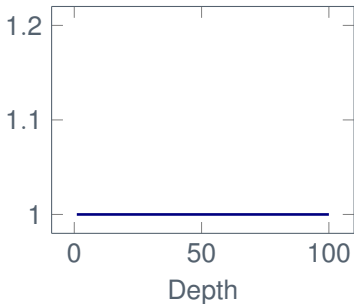
[1] Sirkes, Ziv and Tziperman, Eli
Finite Difference of Adjoint or Adjoint of Finite Difference ?
1997

[2] Sei Alain and Symes William
A Note on Consistency and Adjointness for Numerical Schemes
1997

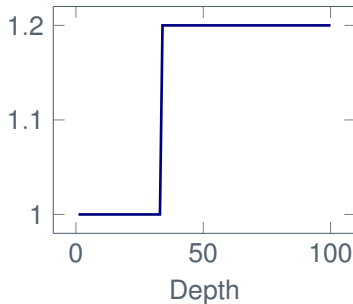
1D Preliminary tests



Initial c Model



Target c Model



1D FWI :

- ▶ Lagrange / B-Bézier Operators
- ▶ RK4 / AB3 time-schemes

Gradient expression :

$$\nabla_{\mathbf{c}} \mathcal{J} = - \int_0^T \int_{\Omega} \frac{2}{\rho \mathbf{c}^3} \frac{\partial \mathbf{p}}{\partial t} \lambda_1 d\Omega dt$$

1D FWI :

- ▶ Lagrange / B-Bézier Operators
- ▶ RK4 / AB3 time-schemes

Adjoint test passed with :

- ▶ With a canonical space inner-product
 $\langle u, v \rangle_X = \sum_i u_i v_i$
- ▶ With a M-space inner product
 $\langle u, v \rangle_X^M = \langle Mu, v \rangle_X$

Gradient expression :

$$\nabla_{\mathbf{c}} \mathcal{J} = - \int_0^T \int_{\Omega} \frac{2}{\rho \mathbf{c}^3} \frac{\partial \mathbf{p}}{\partial t} \lambda_1 d\Omega dt$$

```
./run  
--- Adjoint test ---  
inner product U/D 553123.57586755091  
inner product G/Q 553123.57586756046
```

1D FWI :

- ▶ Lagrange / B-Bézier Operators
- ▶ RK4 / AB3 time-schemes

Adjoint test passed with :

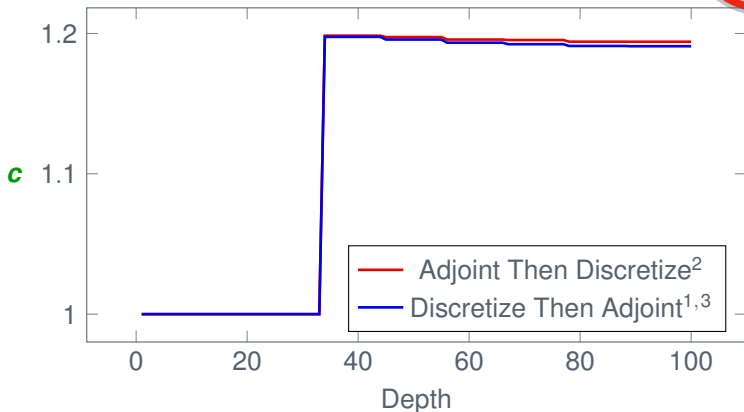
- ▶ With a canonical space inner-product
 $(\langle u, v \rangle_X = \sum_i u_i v_i)$
- ▶ With a M-space inner product
 $(\langle u, v \rangle_X^M = \langle Mu, v \rangle_X)$

Gradient expression :

$$\nabla_{\mathbf{c}} \mathcal{J} = - \int_0^T \int_{\Omega} \frac{2}{\rho \mathbf{c}^3} \frac{\partial \mathbf{p}}{\partial t} \lambda_1 d\Omega dt$$

```
./run
--- Adjoint test ---
inner product U/D 553123.57586755091
inner product G/Q 553123.57586756046
./run
--- Adjoint test ---
inner product U/D -75077.332007383695
inner product G/Q -75077.332007386358
./run
--- Adjoint test ---
inner product U/D 125669.89223600870
inner product G/Q 125669.89223600952
```


1D Velocity Model Reconstructions



c Model at the 100th FWI iteration

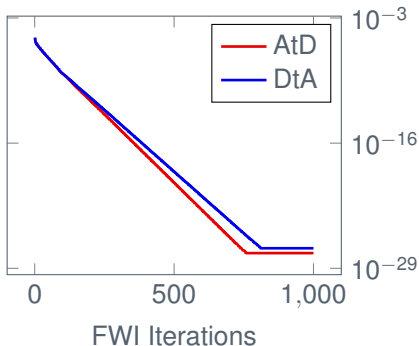
²With Bernstein-Bézier elements and AB3 time scheme

³With canonical scalar product

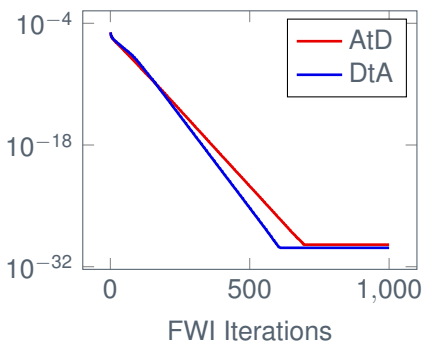
1D Velocity Model Reconstructions



With RK4 :



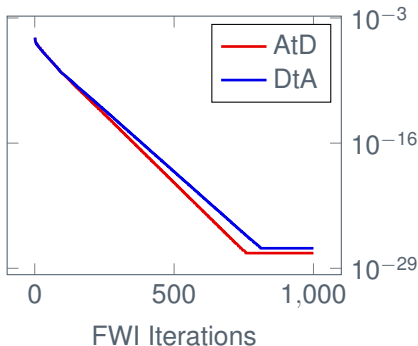
With AB3 :



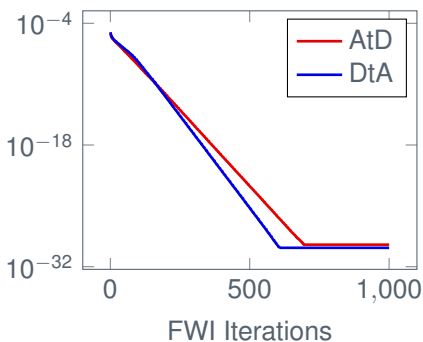
1D Velocity Model Reconstructions



With RK4 :



With AB3 :



- ▶ For RK4 scheme : **AtD** is slightly better than **DtA**
- ▶ For AB3 scheme : **DtA** is slightly better than **AtD**
- ▶ No predominant behaviour