Numerical Scheme Impacts on Time Domain Full Waveform Inversion

Pierre Jacquet, Andreas Atle, Hélène Barucq, Henri Calandra, Julien Diaz

Numerical Scheme Impacts on Time Domain Full Waveform Inversion

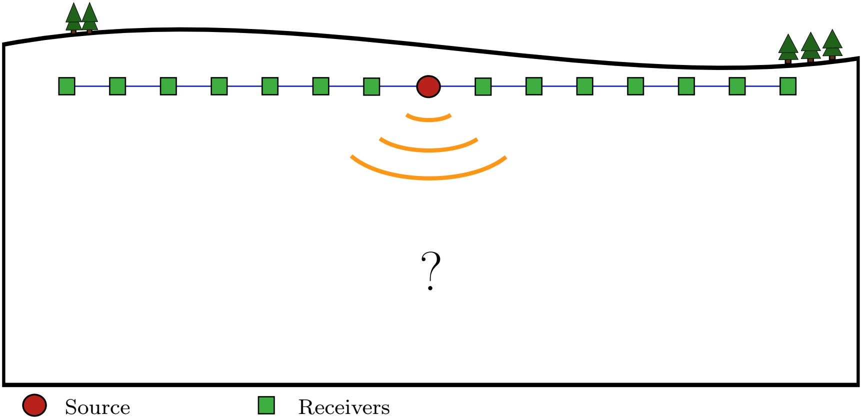
Mathias 2019

Pierre Jacquet
pierre.jacquet@inria.fr

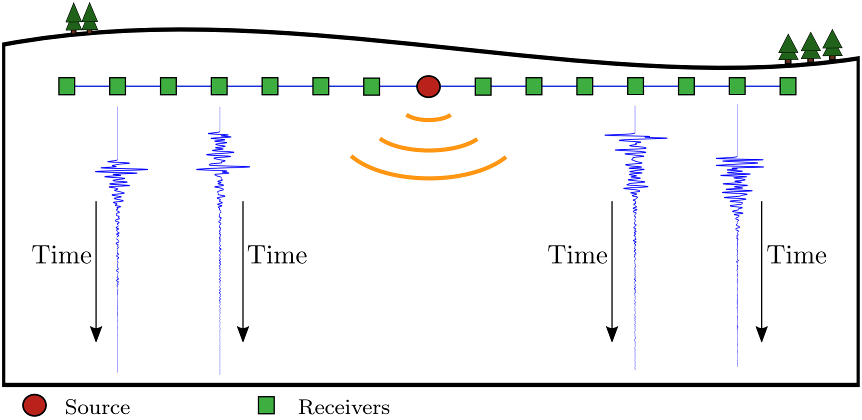
Atle Andreas, Barucq Hélène, Calandra Henri, Diaz Julien
Second year PhD Student
Inria - Magique 3D - DIP

Pau, FRANCE

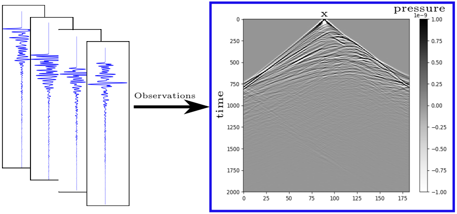
Seismic Acquisition



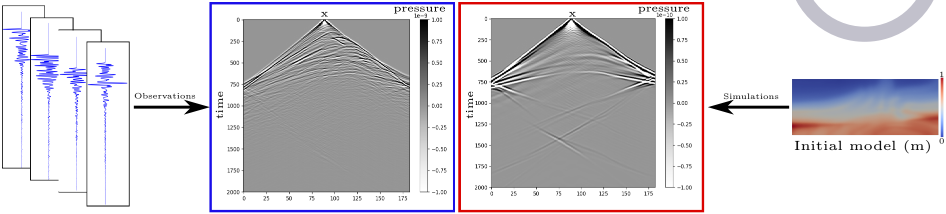
Seismic Acquisition



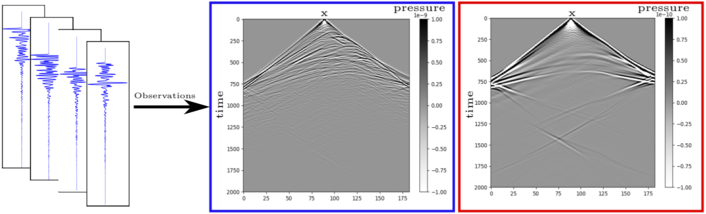
FWI Workflow



FWI Workflow



FWI Workflow



Cost function to minimize :

$$\mathcal{J}(\boldsymbol{m}) = \frac{1}{2}||d_{obs} - \mathcal{F}(\boldsymbol{m})||^2$$

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

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

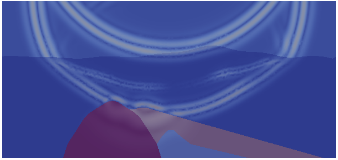
[2] Albert Tarantola
Inversion of seismic reflection data in the acoustic approximation
Geophysics, Vol. 49, 1984

FWI Workflow

Observations (Data)

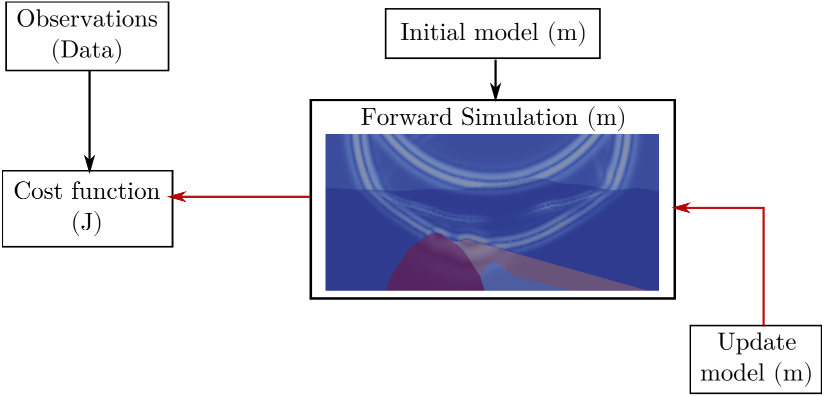
Initial model (m)

Forward Simulation (m)

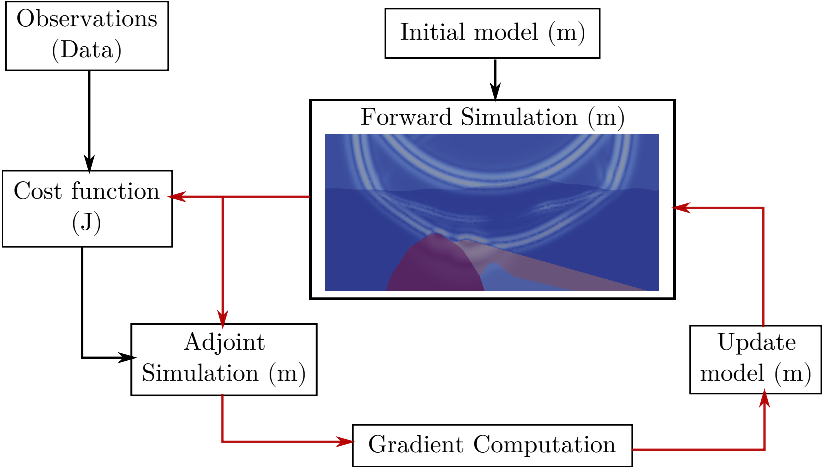


Cost function (J)

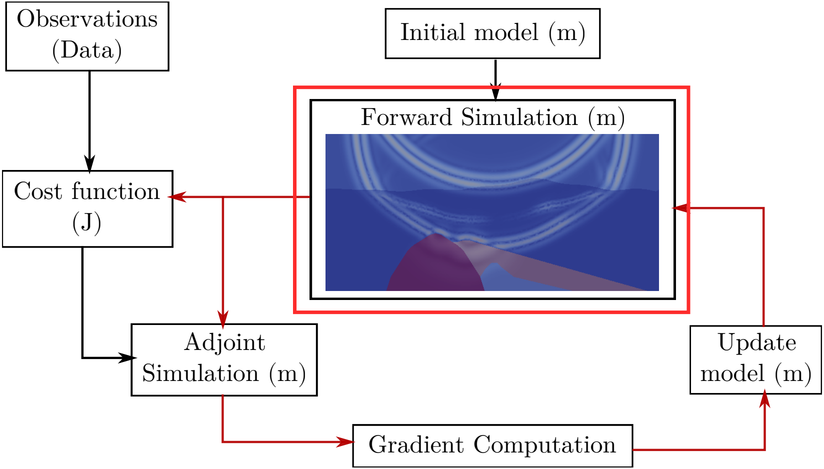
FWI Workflow



FWI Workflow



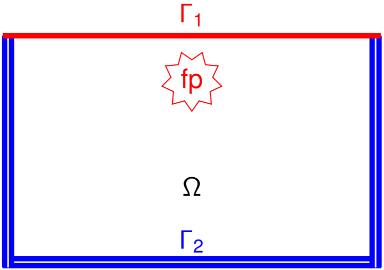
FWI Workflow



Continuous Forward Model

First order acoustic wave equation

$$
\begin{cases}
\dfrac{1}{\rho \boldsymbol{c}^2}\dfrac{\partial \boldsymbol{p}}{\partial t} + \nabla \cdot \boldsymbol{v} = f_p \quad \text{on } \boldsymbol{\Omega} \\
\rho \dfrac{\partial \boldsymbol{v}}{\partial t} + \nabla \boldsymbol{p} = 0 \quad \text{on } \boldsymbol{\Omega} \\
\boldsymbol{p} = 0 \quad \text{on } \boldsymbol{\Gamma_1} \\
\dfrac{\partial \boldsymbol{p}}{\partial t} + \boldsymbol{c} \nabla \boldsymbol{p} \cdot \boldsymbol{n} = 0 \quad \text{on } \boldsymbol{\Gamma_2} \\
\boldsymbol{p}(0) = 0, \quad \boldsymbol{v}(0) = 0
\end{cases}
$$



Domain with Absorbing Boundary Conditions

Discrete Forward Model

Space Discretization :
Discontinuous Galerkin Elements

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

Discrete Forward Model

Space Discretization :
Discontinuous Galerkin Elements

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

Semi-discretized model :

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

with :

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

Discrete Forward Model

Space Discretization :
Discontinuous Galerkin Elements

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

Time schemes :

- Runge Kutta 2/4
- Adams Bashforth 3

Semi-discretized model :

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

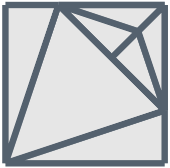
with :

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

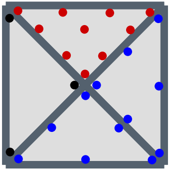
Forward time steps

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



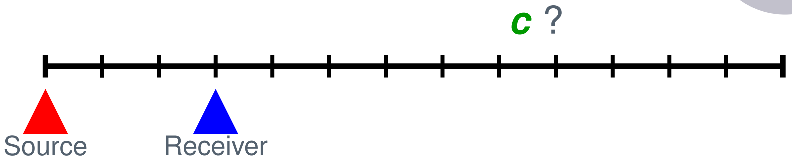
h-adaptivity



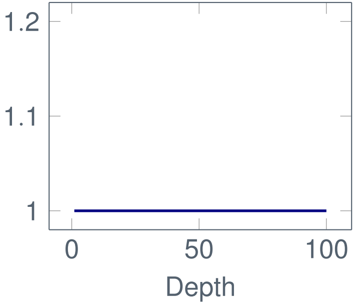
p-adaptivity with P1, P2, P3 elements

Outline

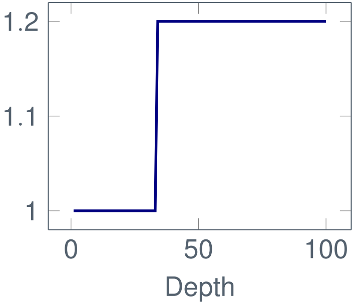
One year ago...



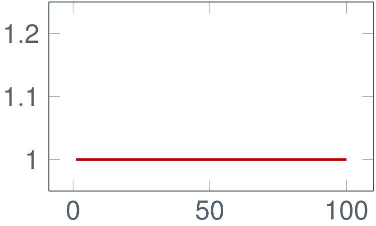
Initial *c* Model



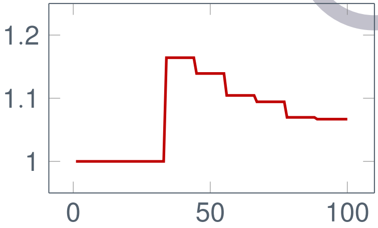
Target *c* Model



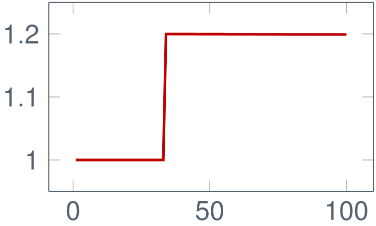
FWI 1D Results



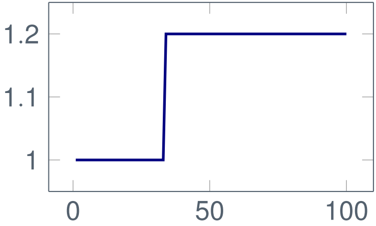
Initial ***c*** model



Itermediate ***c*** model (iter=20)



c model Reconstructed (iter=50)



Target ***c*** model

2D Time Domain Reconstruction

2D FWI :

- Developped in Total environnement (DIP[1])
- Nodal Space Operators (Lagrangian/Jacobian)
- Modal Space Operators (Bernstein-Bézier)
- Runge Kutta 2/4 and Adams Bashforth 3 time-schemes

Gradient expression :

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

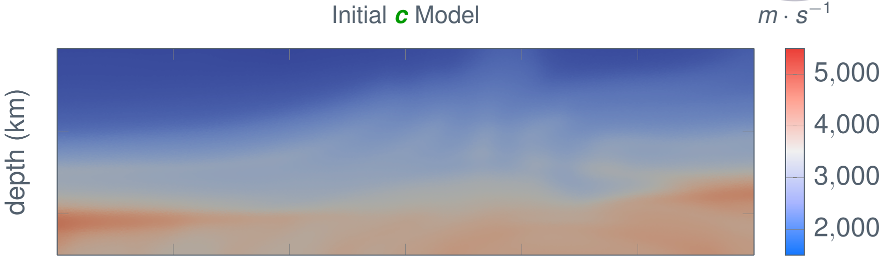
$\boldsymbol{c}$, $\boldsymbol{\rho}$ and $\boldsymbol{\kappa}$ Constant per elements

[1] `http://dip.inria.fr/`

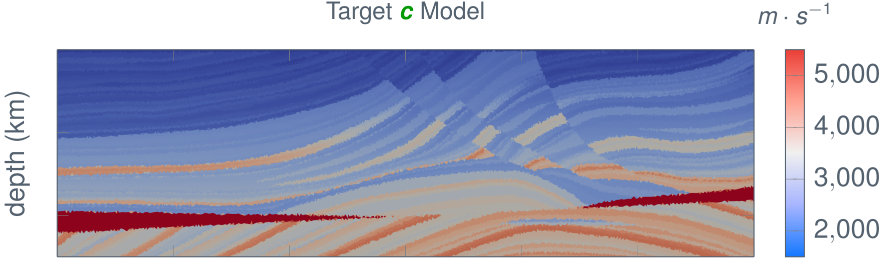
2D Time Domain FWI Reconstructions

Time-schemes comparison

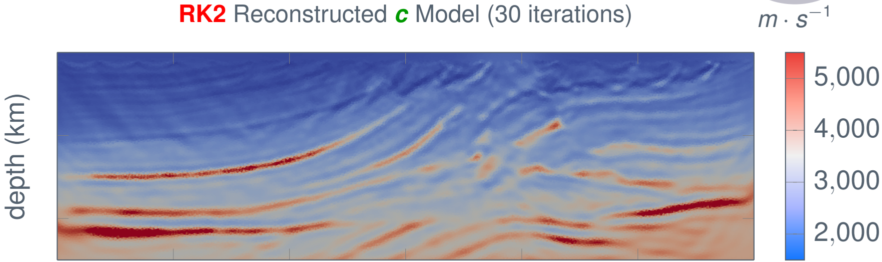
Initial *c* Model



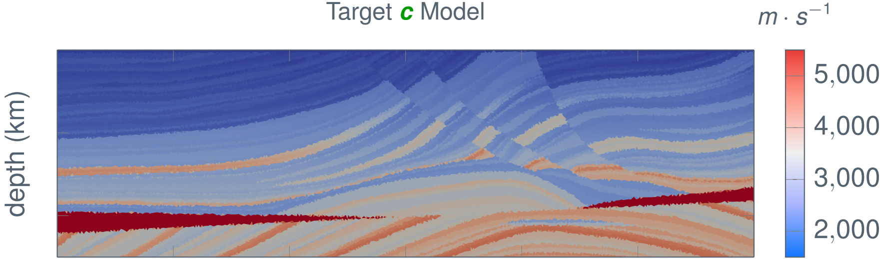
Target *c* Model



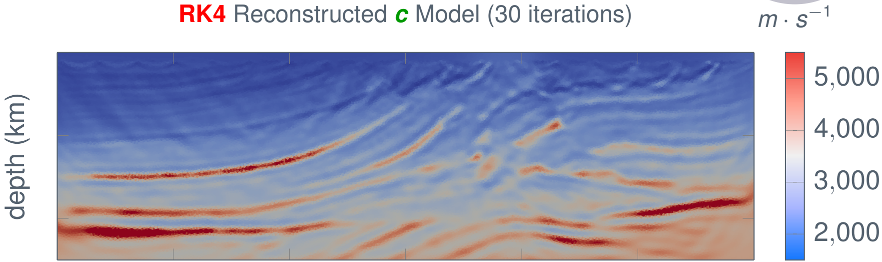
RK2 Reconstructed ***c*** Model (30 iterations)



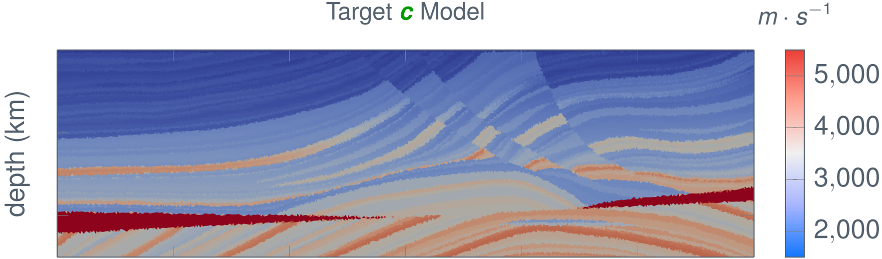
Target ***c*** Model



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



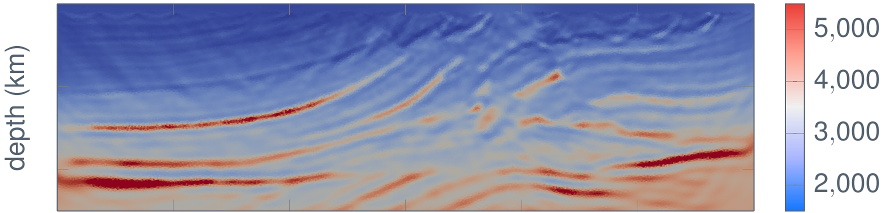
Target ***c*** Model



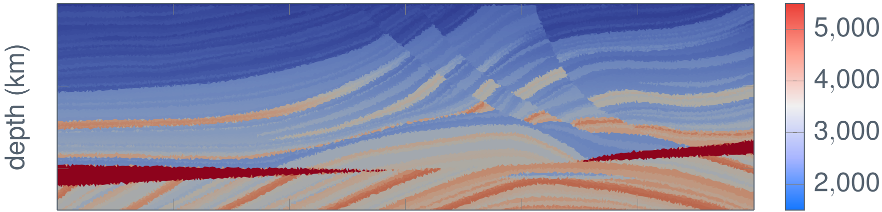
2D Time Domain FWI Reconstructions

Time-schemes comparison

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



Target ***c*** Model

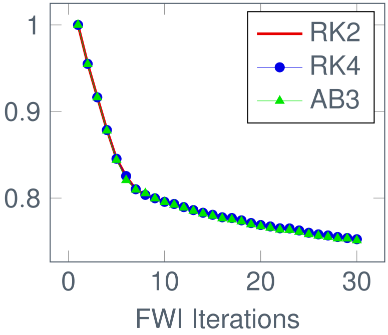


- 47k P1 elements
- Constant ρ model ($\rho = 1$)
- 19 sources / 181 Receivers
- Noise : SNR=10
- 30 iterations
- 120 cores
- Polynomial basis : Nodal

Computational time :

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

Cost function evolution :

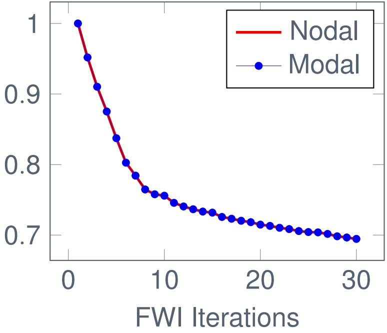


- 47k P1 elements
- Time Scheme : RK2
- Constant ρ model ($\rho = 1$)
- 19 sources / 181 Receivers
- Noise : SNR=10
- 30 iterations
- 120 cores

Computational time :

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

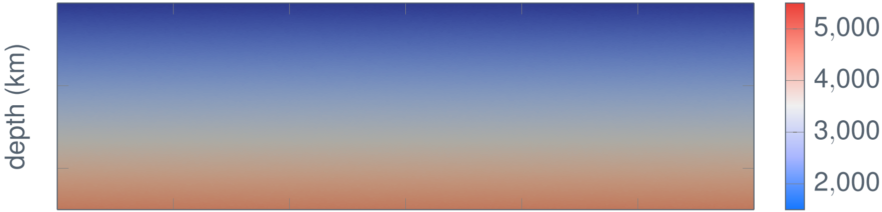
Cost function evolution :



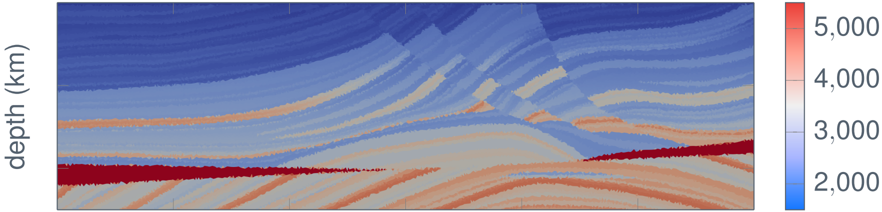
[1] Chan J. and Warburton T.
GPU-Accelerated Bernstein Bézier Discontinuous Galerkin Methods for Wave Problems
SIAM Journal on Scientific Computing 2017

2D Multiscale Reconstructions
Reconstruction with an initial smooth model

Initial $\boldsymbol{c}$ Model $m \cdot s^{-1}$

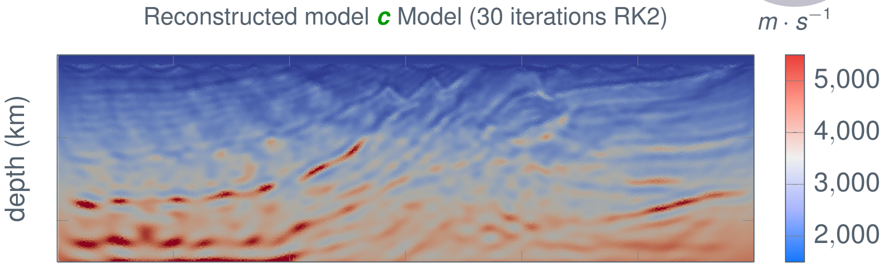


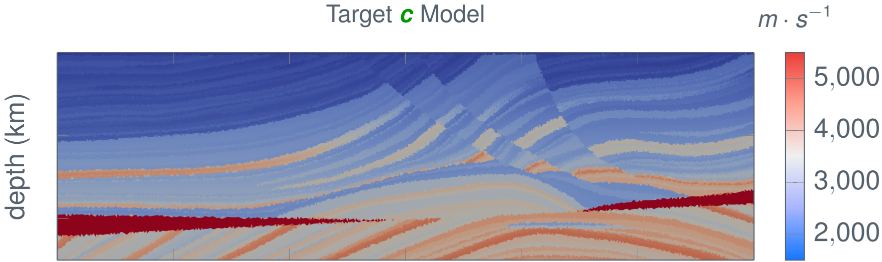
Target $\boldsymbol{c}$ Model $m \cdot s^{-1}$



2D Multiscale Reconstructions

Reconstruction with an initial smooth model





2D Multiscale Reconstructions

Multiscale Principle [1]

Filtered Traces :

Low Frequencies

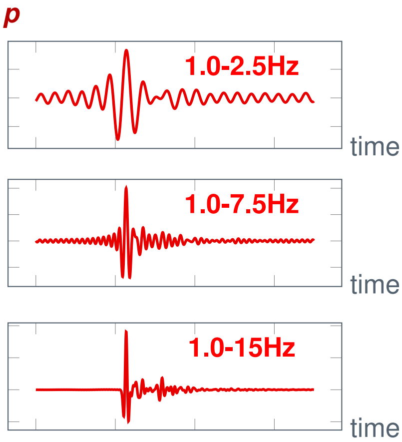
$\Longleftrightarrow$

Reconstruct **coarse** structures

High Frequencies

$\Longleftrightarrow$

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

Multiscale Principle [1]

Heuristic Illustration :

Low Frequencies

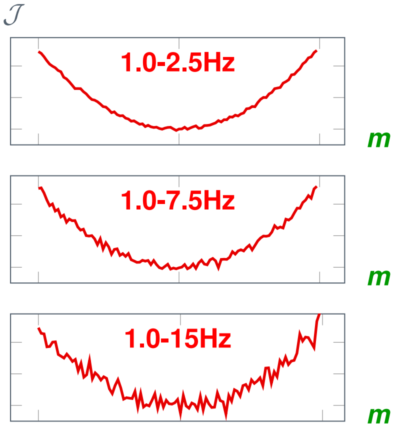
$\Longleftrightarrow$

Reconstruct **coarse** structures

High Frequencies

$\Longleftrightarrow$

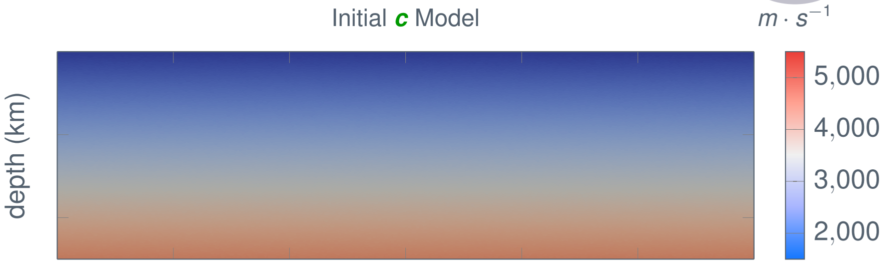
Reconstruct **small** structures

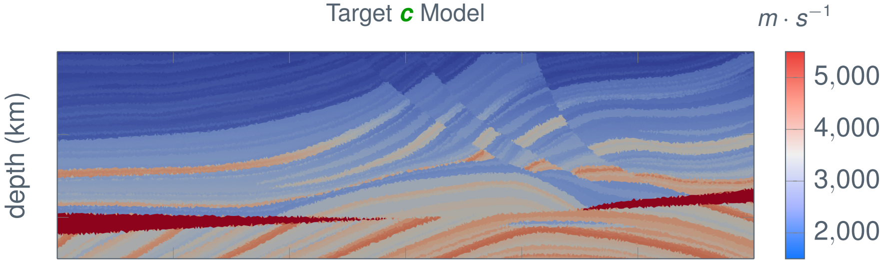


[1] C. Bunks, F. M. Saleck, S. Zaleski, and G. Chavent
Multiscale seismic waveform inversion
GEOPHYSICS, Vol. 60, No. 5, 1995

2D Multiscale Reconstructions

Reconstruction with an initial smooth model

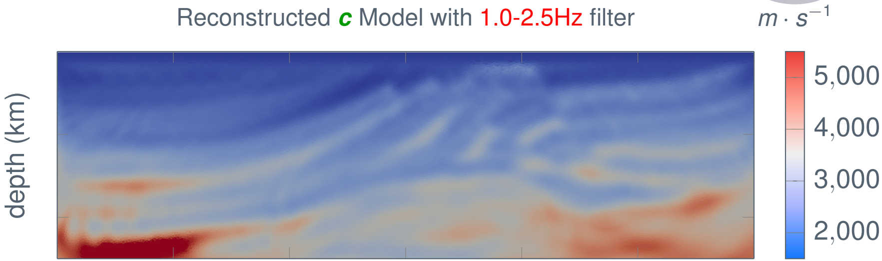




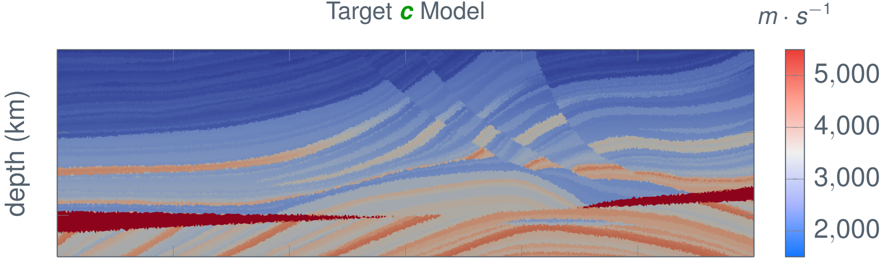
2D Multiscale Reconstructions

Reconstruction with an initial smooth model

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

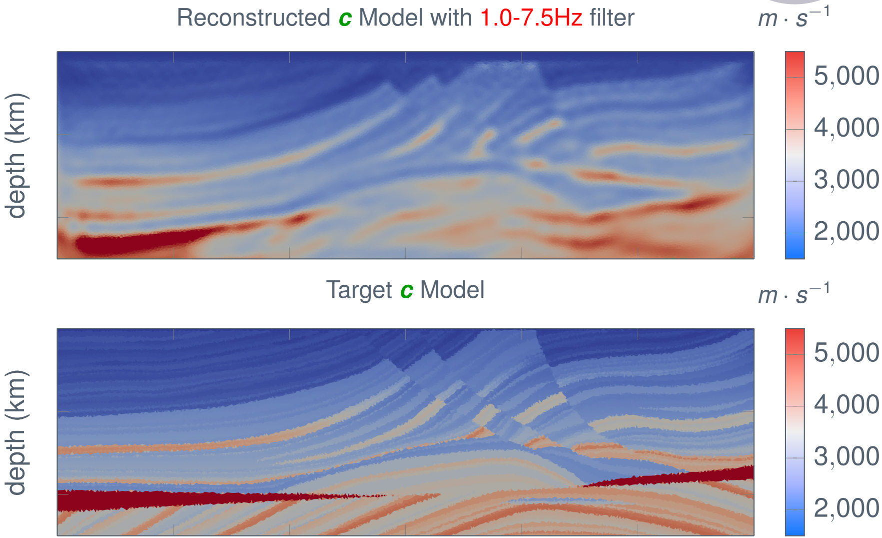


Target ***c*** Model



2D Multiscale Reconstructions

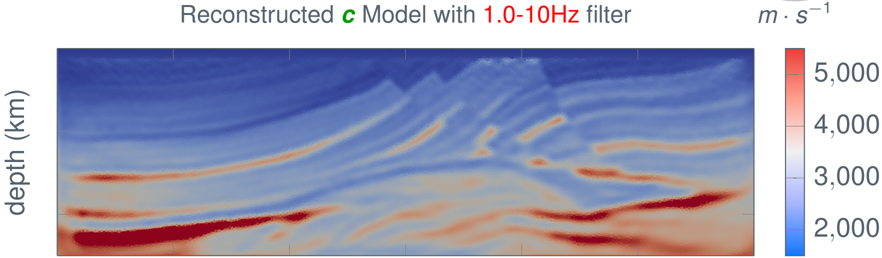
Reconstruction with an initial smooth model



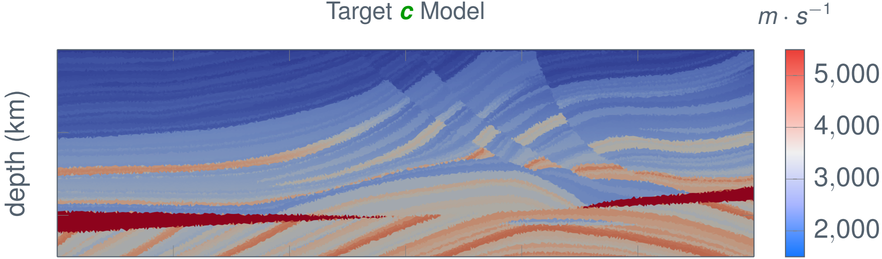
2D Multiscale Reconstructions

Reconstruction with an initial smooth model

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



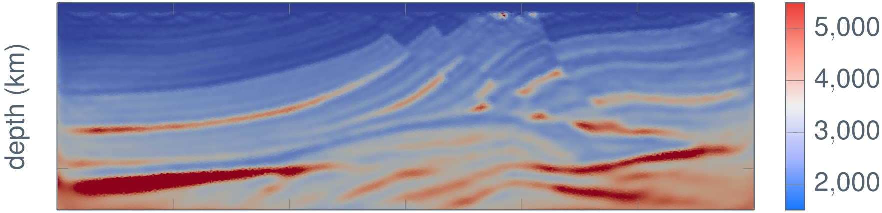
Target ***c*** Model



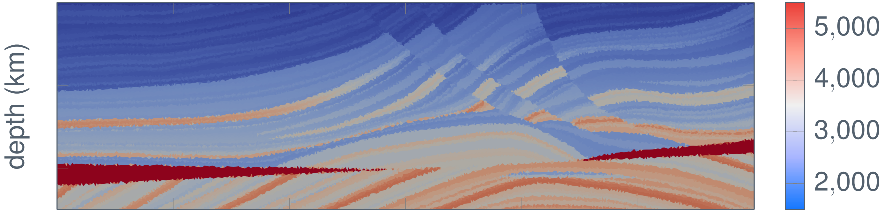
2D Multiscale Reconstructions

Reconstruction with an initial smooth model

Reconstructed ***c*** Model with 1.0-15Hz filter $m \cdot s^{-1}$



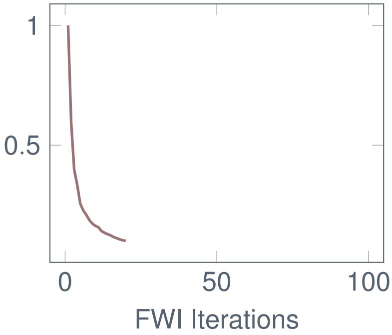
Target ***c*** Model $m \cdot s^{-1}$



2D Multiscale Reconstructions

- 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

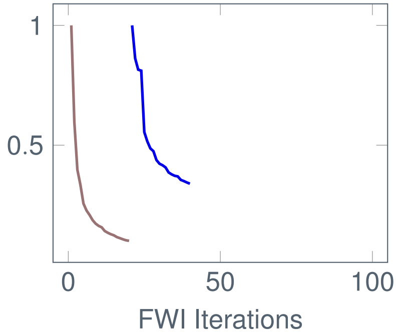
Cost function evolution :



2D Multiscale Reconstructions

- 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

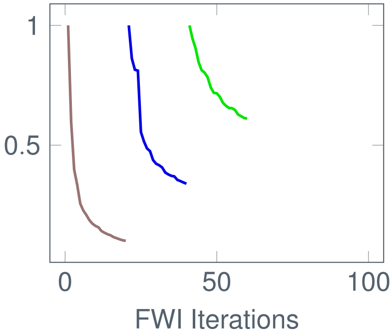
Cost function evolution :



2D Multiscale Reconstructions

- 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

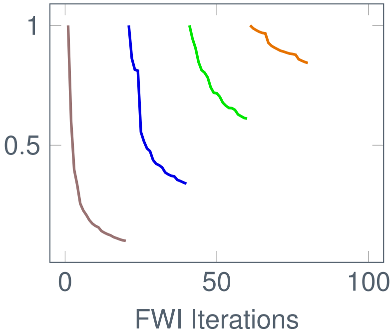
Cost function evolution :



2D Multiscale Reconstructions

- 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

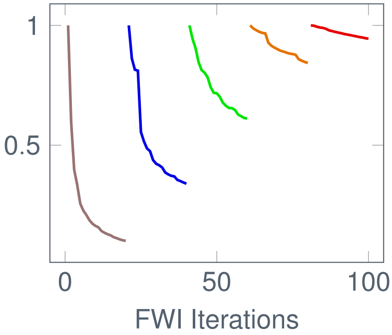
Cost function evolution :



2D Multiscale Reconstructions

- 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 :



Conclusion

Main Results :

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

Conclusion

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)

Conclusion

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.

Annexe

Adjoint State Method

Lagrangian fonctional [1] :

$$\mathcal{L}(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\lambda}}, \boldsymbol{m}) = \frac{1}{2}||d_{obs} - \mathcal{R}(\widehat{\boldsymbol{u}})||^2 + < Forward_{\boldsymbol{m}}(\widehat{\boldsymbol{u}}) - f_p, \widehat{\boldsymbol{\lambda}} >$$

If $\widehat{\boldsymbol{u}} = \boldsymbol{u}$ Solution of the Direct Problem $\Longleftrightarrow$ $(Forward_{\boldsymbol{m}}(\boldsymbol{u}) - f_p = 0)$:

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

[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

Adjoint State Method

Lagrangian fonctional [1] :

$$\mathcal{L}(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\lambda}}, \boldsymbol{m}) = \frac{1}{2}||d_{obs} - \mathcal{R}(\widehat{\boldsymbol{u}})||^2 + < Forward_{\boldsymbol{m}}(\widehat{\boldsymbol{u}}) - f_p, \widehat{\boldsymbol{\lambda}} >$$

If $\widehat{\boldsymbol{u}} = \boldsymbol{u}$ Solution of the Direct Problem $\Longleftrightarrow$ $(Forward_{\boldsymbol{m}}(\boldsymbol{u}) - f_p = 0)$:

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

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

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

[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

Adjoint State Method

Lagrangian fonctional [1] :

$$\mathcal{L}(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\lambda}}, \boldsymbol{m}) = \frac{1}{2}||d_{obs} - \mathcal{R}(\widehat{\boldsymbol{u}})||^2 + < Forward_{\boldsymbol{m}}(\widehat{\boldsymbol{u}}) - f_p, \widehat{\boldsymbol{\lambda}} >$$

If $\widehat{\boldsymbol{u}} = \boldsymbol{u}$ Solution of the Direct Problem $\Longleftrightarrow$ $(Forward_{\boldsymbol{m}}(\boldsymbol{u}) - f_p = 0)$:

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

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

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

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

$$\partial_{\boldsymbol{m}_i} \mathcal{J}(\boldsymbol{m}) = \partial_{\boldsymbol{m}_i} \mathcal{L}(\boldsymbol{u}, \boldsymbol{\lambda}, \boldsymbol{m}) = \partial_{\boldsymbol{m}_i} < Forward_{\boldsymbol{m}}(\boldsymbol{u}), \boldsymbol{\lambda} >$$

[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

Adjoint Formulation

Continuous
Direct Problem

Adjoint Formulation

Continuous Direct Problem $\xrightarrow{\textbf{Adjoint}}$ *Continuous Adjoint Problem*

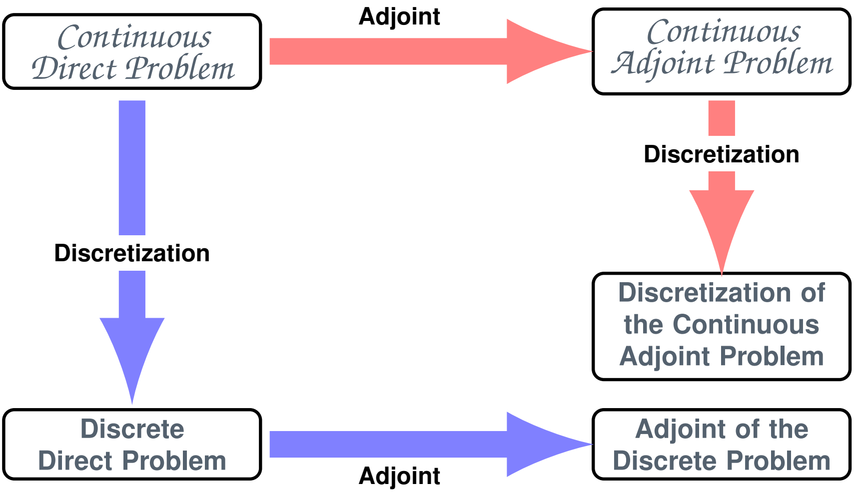
Adjoint Formulation

Continuous Direct Problem $\xrightarrow{\text{Adjoint}}$ *Continuous Adjoint Problem*

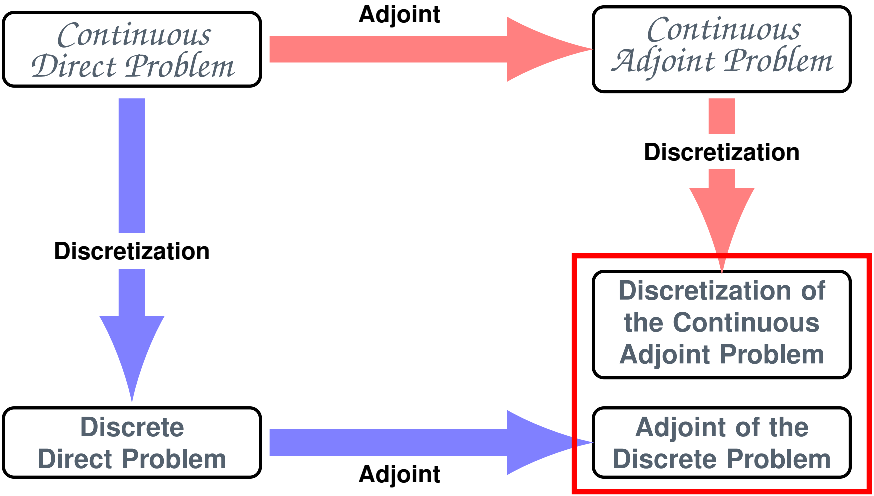
Discretization

Discretization of the Continuous Adjoint Problem

Adjoint Formulation



Adjoint Formulation



AtD : Adjoint then Discretized Strategy

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

$$\begin{cases} \dfrac{1}{\rho \boldsymbol{c}^2}\dfrac{\partial \boldsymbol{p}}{\partial t} + \nabla \cdot \boldsymbol{v} = f_p \quad \text{on } \boldsymbol{\Omega} \\ \rho \dfrac{\partial \boldsymbol{v}}{\partial t} + \nabla \boldsymbol{p} = 0 \quad \text{on } \boldsymbol{\Omega} \\ \boldsymbol{p} = 0 \quad \text{on } \boldsymbol{\Gamma_1} \\ \dfrac{\partial \boldsymbol{p}}{\partial t} + \boldsymbol{c}\nabla \boldsymbol{p} \cdot \boldsymbol{n} = 0 \quad \text{on } \boldsymbol{\Gamma_2} \\ \boldsymbol{p}(0) = 0, \quad \boldsymbol{v}(0) = 0 \end{cases} \qquad \begin{cases} \dfrac{1}{\rho \boldsymbol{c}^2}\dfrac{\partial \boldsymbol{\lambda_1}}{\partial t} + \nabla \cdot \boldsymbol{\lambda_2} = \dfrac{\partial \mathcal{J}}{\partial \boldsymbol{p}} \quad \text{on } \boldsymbol{\Omega} \\ \rho \dfrac{\partial \boldsymbol{\lambda_2}}{\partial t} + \nabla \boldsymbol{\lambda_1} = 0 \quad \text{on } \boldsymbol{\Omega} \\ \boldsymbol{\lambda_1} = 0 \quad \text{on } \boldsymbol{\Gamma_1} \\ \dfrac{\partial \boldsymbol{\lambda_1}}{\partial t} - \boldsymbol{c}\nabla \boldsymbol{\lambda_1} \cdot \boldsymbol{n} = 0 \quad \text{on } \boldsymbol{\Gamma_2} \\ \boldsymbol{\lambda_1}(T) = 0, \quad \boldsymbol{\lambda_2}(T) = 0 \end{cases}$$

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

AtD : Adjoint then Discretized Strategy

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

$$\begin{cases} \dfrac{\partial \bar{\boldsymbol{U}}^n}{\partial t} = A\bar{\boldsymbol{U}}^n + \bar{\boldsymbol{F}}^n \\ \text{With :} \quad \bar{\boldsymbol{U}}^n = \begin{pmatrix} \bar{\boldsymbol{P}}^n \\ \bar{\boldsymbol{V}}^n \end{pmatrix} \end{cases}$$

$$\begin{cases} \dfrac{\partial \bar{\boldsymbol{\Lambda}}^n}{\partial t} = A\bar{\boldsymbol{\Lambda}}^n + R^*(R\bar{\boldsymbol{U}}^n - d_{obs}) \\ \text{With :} \quad \bar{\boldsymbol{\Lambda}}^n = \begin{pmatrix} \bar{\boldsymbol{\Lambda}}_1{}^n \\ \bar{\boldsymbol{\Lambda}}_2{}^n \end{pmatrix} \end{cases}$$

Time-steps going Forward

Time-steps going Backward

All time scheme can be summed-up such as :

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

RK4 time-scheme leads to :

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

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

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

DtA : Discretize then Adjoint Strategy

All time scheme can be summed-up such as :

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

We are looking for a Discrete Adjoint state satisfying :

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

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

$$< \boldsymbol{L}\bar{\boldsymbol{U}}, \bar{\boldsymbol{\Lambda}} > = < \bar{\boldsymbol{U}}, \boldsymbol{L}^{*}\bar{\boldsymbol{\Lambda}} >$$

DtA : Discretize then Adjoint Strategy

All time scheme can be summed-up such as :

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

We are looking for a Discrete Adjoint state satisfying :

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

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

$$< \boldsymbol{L}\bar{\boldsymbol{U}}, \bar{\boldsymbol{\Lambda}} > = < \bar{\boldsymbol{U}}, \boldsymbol{L}^*\bar{\boldsymbol{\Lambda}} >$$

$$< \bar{\boldsymbol{G}}, \bar{\boldsymbol{\Lambda}} > = < \bar{\boldsymbol{U}}, \bar{\boldsymbol{D}} > \quad \text{(Adjoint Test)}$$

Adjoint test succeeds $\Longleftrightarrow$ operator $\boldsymbol{L}^*$ well established

DtA : Discretize then Adjoint Strategy

Example with RK4

RK4 time-scheme leads to :

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

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

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

So :

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

Adjoint Strategies Comparison

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 Strategies Comparison

Adjoint Then Discretize

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

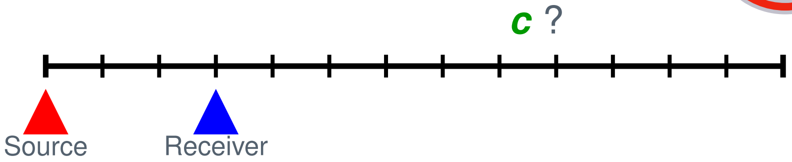
Discretize then Adjoint

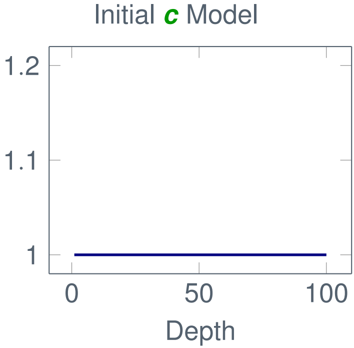
- \+ Numerical approach
- \+ / - Has an Adjoint Test **(in theory)**
- \- Tremendous work to develop the adjoint operators
- Non-consistency of the adjoint state [2]

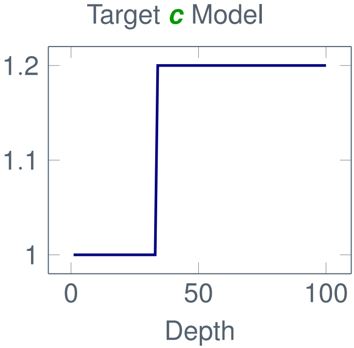
[1] Sirkes, Ziv and Tziperman, Eli
Finite Difference of Adjoint or Adjoint of Finite Difference ?
1997

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

1D Preliminary tests







1D Preliminary tests :

1D FWI :

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

Gradient expression :

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

1D Preliminary tests :

1D FWI :

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

Adjoint test passed with :

- With a canonical space inner-product ($< u, v >_X = \sum_i u_i v_i$)
- With a M-space inner product ($< u, v >_X^M = < Mu, v >_X$)

Gradient expression :

$$\nabla_{\mathbf{c}} \mathcal{J} = -\int_0^T \int_\Omega \frac{2}{\rho \mathbf{c}^3} \frac{\partial \mathbf{p}}{\partial t} \boldsymbol{\lambda}_1 d\Omega dt$$

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

1D Preliminary tests :

1D FWI :

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

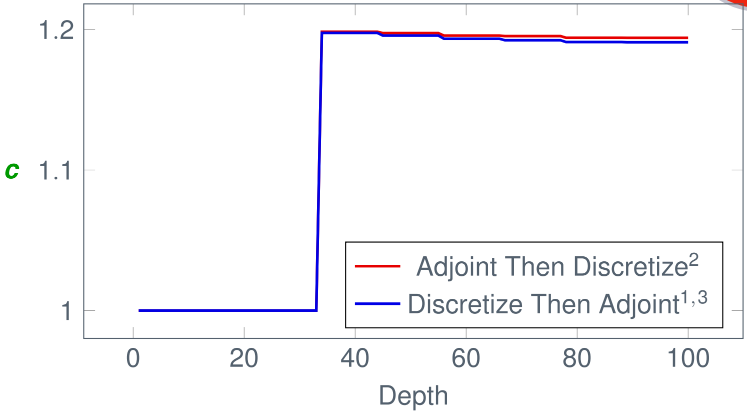
Adjoint test passed with :

- With a canonical space inner-product ($< u, v >_X = \sum_i u_i v_i$)
- With a M-space inner product ($< u, v >_X^M = < Mu, v >_X$)

Gradient expression :

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

```
./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
```



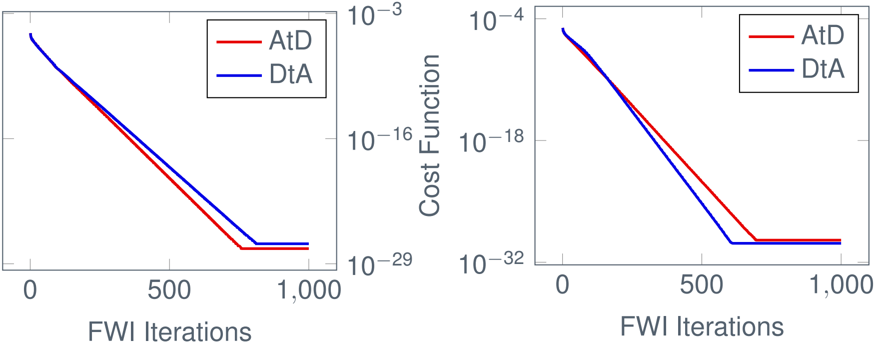
c Model at the 100th FWI iteration

[2] With Bernstein-Bézier elements and AB3 time scheme
[3] With canonical scalar product

1D Velocity Model Reconstructions

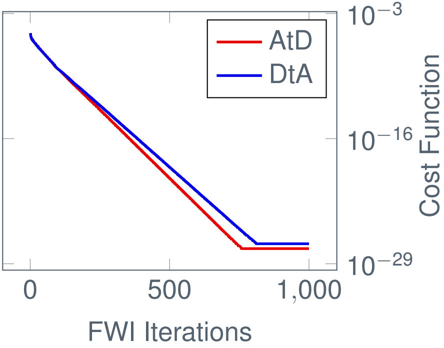
With RK4 :

With AB3 :

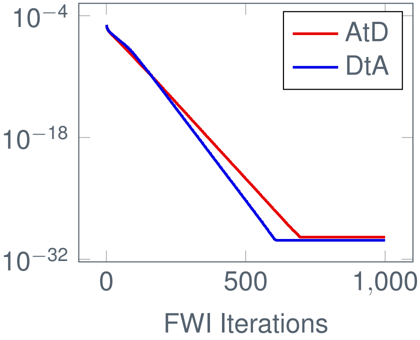


1D Velocity Model Reconstructions

With RK4 :



With AB3 :



- For RK4 scheme : AtD is slighly better than DtA
- For AB3 scheme : DtA is slighly better than AtD
- No predominant behaviour