Full Waveform Inversion Adjoint Studies MATHIAS 2018

Pierre Jacquet, Hélène Barucq, Julien Diaz, Henri Calandra

Full Waveform Inversion Adjoint Studies

MATHIAS 2018

Pierre Jacquet
`pierre.jacquet@inria.fr`

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

Pau, FRANCE

e2s UPPA energy environment solutions
UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR
Inria inventeurs du monde numérique
TOTAL
Depth Imaging Partnership DIP

Outline

Outline

The continuous model

Continuous Problem:

$$\begin{cases} \frac{1}{\rho_0 {v_p}^2}\frac{\partial p}{\partial t} + \frac{\partial v}{\partial x} = f_p \quad \text{in } \Omega \\ \rho_0 \frac{\partial v}{\partial t} + \frac{\partial p}{\partial x} = f_v \quad \text{in } \Omega \end{cases}$$

$$\begin{cases} p(t=0) = 0 \\ v(t=0) = 0 \\ \frac{\partial p}{\partial t} + v_p \frac{\partial p}{\partial x}.\boldsymbol{n} = 0 \quad \text{on } \Gamma \end{cases}$$

Figure: 1D Domain Model

The discretized model

Discretized Problem:

$$\begin{cases} \dfrac{\partial \bar{\boldsymbol{P}}}{\partial t} = A_{pv}\bar{\boldsymbol{V}} + A_{pp}\bar{\boldsymbol{P}} + \bar{\boldsymbol{F}}_p \\ \dfrac{\partial \bar{\boldsymbol{V}}}{\partial t} = A_{vp}\bar{\boldsymbol{P}} \end{cases}$$

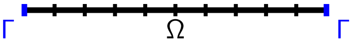


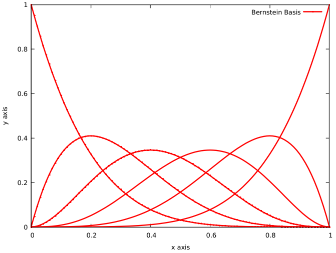
Figure: 1D Discretized Domain

- **Discontinuous Galerkin** space discretization
- Different **time-schemes** (RK4,AB3)
- Two polynomial basis (**Lagrange** and **Bernstein**)
- Constant velocity (v_p) per cells
- Constant density (ρ_0) per cells

Bernstein/DG Properties

Bernstein formulation:

$$B^N_{ijkl} = C^N_{ijkl}\lambda_0^i\lambda_1^j\lambda_2^k\lambda_3^l \quad \text{with:} \quad C^N_{ijkl} = \frac{N!}{i!j!k!l!}$$



P[X^5] Bernstein basis

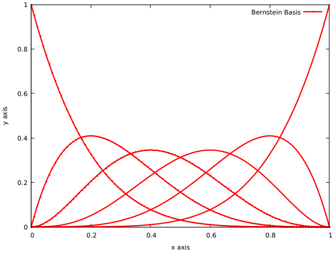
Bernstein/DG Properties

Bernstein formulation:

$$B^N_{ijkl} = C^N_{ijkl}\lambda_0^i\lambda_1^j\lambda_2^k\lambda_3^l \quad \text{with: } C^N_{ijkl} = \frac{N!}{i!j!k!l!}$$

Easy Derivative expression:

$$\frac{\partial B^N_\alpha}{\partial \lambda_p} = N B^{N-1}_{\alpha - e_p} \quad \text{with: } \alpha = (i,j,k,l)$$



P[X^5] Bernstein basis

Bernstein/DG Properties

Bernstein formulation:

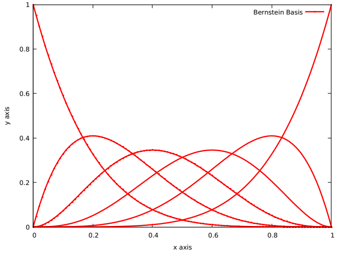
$$B^N_{ijkl} = C^N_{ijkl}\lambda_0^i\lambda_1^j\lambda_2^k\lambda_3^l \quad \text{with: } C^N_{ijkl} = \frac{N!}{i!j!k!l!}$$

Easy Derivative expression:

$$\frac{\partial B^N_\alpha}{\partial \lambda_p} = NB^{N-1}_{\alpha-e_p} \quad \text{with: } \alpha = (i, j, k, l)$$

Sparse Degree Elevation operator:

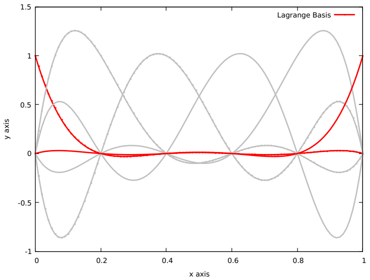
$$B^{N-1}_\alpha = \sum_{p=0}^{d} \frac{\alpha_p + 1}{N} B^N_{\alpha+e_p}$$



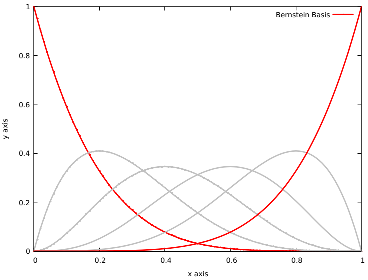
P[X^5] Bernstein basis

Bernstein/DG Properties

Unique boundary condition values :



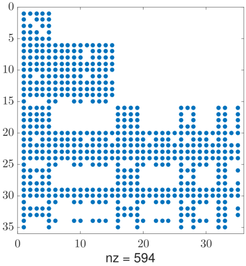
P(X^5) Lagrange basis



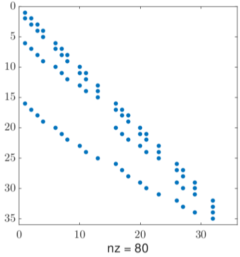
P(X^5) Bernstein basis

$\Rightarrow$ Same Flux Management

Derivative-Operator Analysis



3D Lagrange D matrix



3D Bernstein D matrix

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

1D Results

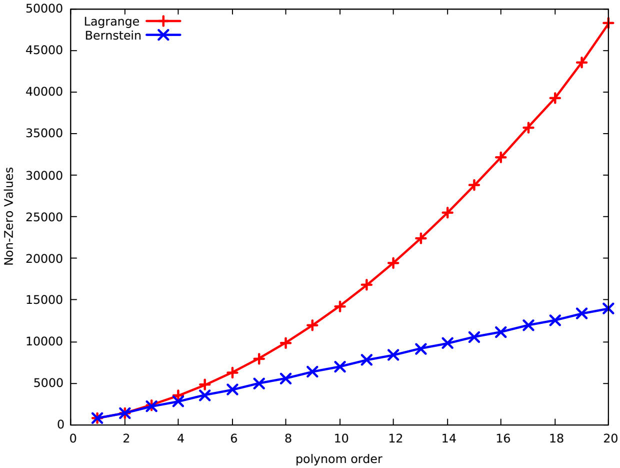
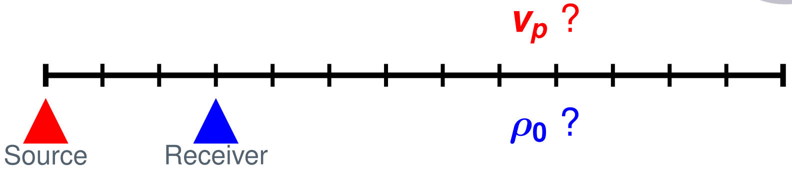


Figure: Operators NZVs as a function of the order

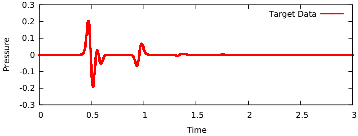
Outline



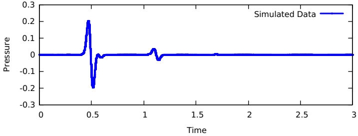
v_p ?
ρ_0 ?
Source
Receiver

v_p ?

ρ_0 ?



Data and Simulation



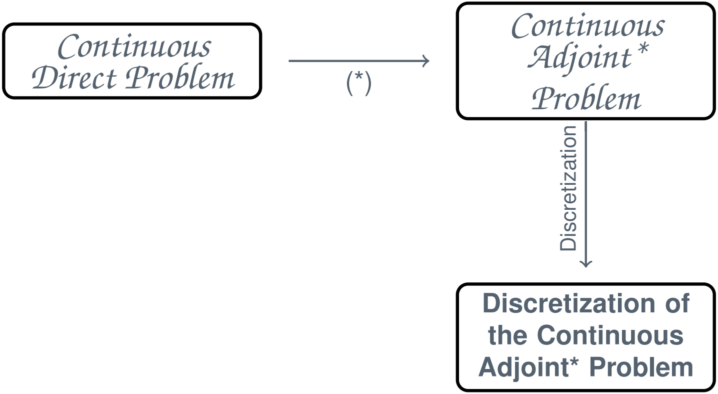
$$J(p) = \frac{1}{2}||Rp - data||^2$$

- Find v_p and ρ_0 to minimize J

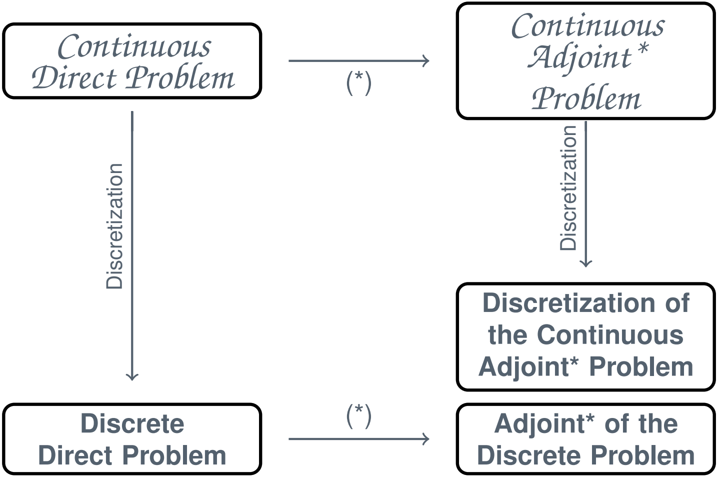
Continuous
Direct Problem

Adjoint Studies

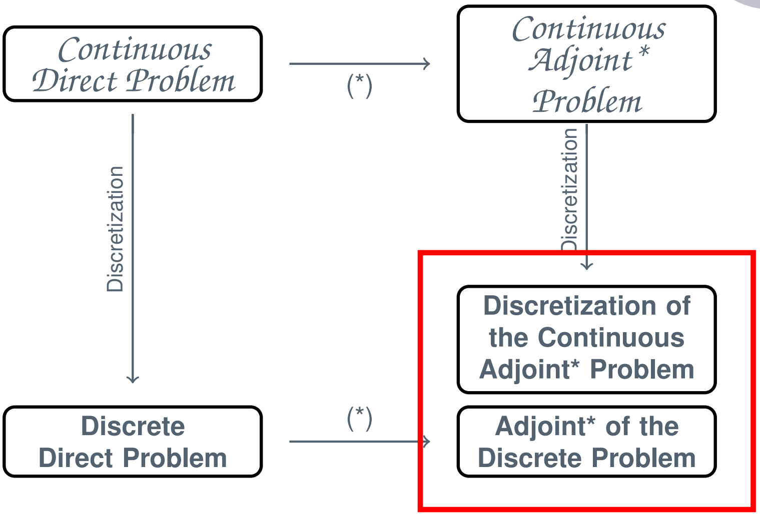
Adjoint Studies



Adjoint Studies



Adjoint Studies



AtD : Adjoint then Discretized Strategy

$$J(p) = \frac{1}{2}||Rp - data||^2$$

$$\begin{cases} \frac{1}{\rho_0 v_p{}^2} \frac{\partial p}{\partial t} + \nabla . v = f_p \\ \rho_0 \frac{\partial v}{\partial t} + \nabla p = 0 \\ p(t=0) = 0 \\ v(t=0) = 0 \\ \frac{\partial p}{\partial t} + v_p \nabla p . \mathbf{n} = 0 \quad \text{on } \Gamma \end{cases} \qquad \begin{cases} \frac{1}{\rho_0 v_p{}^2} \frac{\partial \lambda_1}{\partial t} + \nabla . \lambda_2 = \frac{\partial J}{\partial p} \\ \rho_0 \frac{\partial \lambda_2}{\partial t} + \nabla \lambda_1 = 0 \\ \lambda_1(t=T) = 0 \\ \lambda_2(t=T) = 0 \\ \frac{\partial \lambda_1}{\partial t} + v_p \nabla \lambda_1 . \mathbf{n} = 0 \quad \text{on } \Gamma \end{cases}$$

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

AtD : Adjoint then Discretized Strategy

$$J(\bar{\boldsymbol{P}}) = \frac{1}{2}||R\bar{\boldsymbol{P}} - data||^2$$

$$\begin{cases} \dfrac{\partial \bar{\boldsymbol{P}}}{\partial t}^n = A_{pv}\bar{\boldsymbol{V}}^n + A_{pp}\bar{\boldsymbol{P}}^n + \bar{\boldsymbol{F}}_p^n \\ \dfrac{\partial \bar{\boldsymbol{V}}}{\partial t}^n = A_{vp}\bar{\boldsymbol{P}}^n \end{cases}$$

$$\begin{cases} \dfrac{\partial \bar{\boldsymbol{\Lambda}}_1}{\partial t}^n = +A_{pv}\bar{\boldsymbol{\Lambda}}_2^n + A_{pp}\bar{\boldsymbol{\Lambda}}_1^n + \bar{\boldsymbol{D}}_p^n \\ \dfrac{\partial \bar{\boldsymbol{\Lambda}}_2}{\partial t}^n = A_{vp}\bar{\boldsymbol{\Lambda}}_1^n \end{cases}$$

Time-steps going Forward

Time-steps going Backward

DtA : Discretize then Adjoint Strategy

$$\frac{\partial \bar{\boldsymbol{U}}^n}{\partial t} = A\bar{\boldsymbol{U}}^n + \bar{\boldsymbol{F}}^n \quad \text{With : } \bar{\boldsymbol{U}} = \begin{pmatrix} \bar{\boldsymbol{P}} \\ \bar{\boldsymbol{V}} \end{pmatrix}, A = \begin{pmatrix} A_{pp} & A_p \\ A_v & 0 \end{pmatrix}, \bar{\boldsymbol{F}} = \begin{pmatrix} \bar{\boldsymbol{F}}_p \\ 0 \end{pmatrix}$$

All time scheme can be summed-up such as :

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

We are looking for a Discrete Adjoint state satisfying :

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

DtA : Discretize then Adjoint Strategy

Example with RK4

RK4 time-scheme leads to :

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

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

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

So :

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

Adjoint test

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

Adjoint test

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

$$\begin{cases} L\bar{\boldsymbol{U}} = E\bar{\boldsymbol{F}} = \bar{\boldsymbol{G}} \\ \bar{\boldsymbol{U}}(t=0) = 0 \end{cases}$$

$$\begin{cases} L^*\bar{\boldsymbol{\Lambda}} = -R^*(R\bar{\boldsymbol{U}} - data) = \bar{\boldsymbol{D}} \\ \bar{\boldsymbol{\Lambda}}(t=T) = 0 \end{cases}$$

Time-steps going Forward

Time-steps going Backward

Adjoint test

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

$$\begin{cases} L\bar{\boldsymbol{U}} = E\bar{\boldsymbol{F}} = \bar{\boldsymbol{G}} \\ \bar{\boldsymbol{U}}(t=0) = 0 \end{cases} \qquad \begin{cases} L^*\bar{\boldsymbol{\Lambda}} = -R^*(R\bar{\boldsymbol{U}} - data) = \bar{\boldsymbol{D}} \\ \bar{\boldsymbol{\Lambda}}(t=T) = 0 \end{cases}$$

Time-steps going Forward

Time-steps going Backward

$$< E\bar{\boldsymbol{F}}, \bar{\boldsymbol{\Lambda}} > = < \bar{\boldsymbol{U}}, -R^*(R\bar{\boldsymbol{U}} - data) >$$

Adjoint test

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

$$\begin{cases} L\bar{\boldsymbol{U}} = E\bar{\boldsymbol{F}} = \bar{\boldsymbol{G}} \\ \bar{\boldsymbol{U}}(t=0) = 0 \end{cases} \qquad \begin{cases} L^*\bar{\boldsymbol{\Lambda}} = -R^*(R\bar{\boldsymbol{U}} - data) = \bar{\boldsymbol{D}} \\ \bar{\boldsymbol{\Lambda}}(t=T) = 0 \end{cases}$$

Time-steps going Forward

Time-steps going Backward

$$< E\bar{\boldsymbol{F}}, \bar{\boldsymbol{\Lambda}} > = < \bar{\boldsymbol{U}}, -R^*(R\bar{\boldsymbol{U}} - data) >$$

$$< \bar{\boldsymbol{G}}, \bar{\boldsymbol{\Lambda}} > = < \bar{\boldsymbol{U}}, \bar{\boldsymbol{D}} >$$

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

Outline

Adjoint test

Adjoint test passed for :

- Lagrange Operators
- Bernstein Operators
- Runge Kutta 4 time-scheme
- Adams Bashforth 3 time-scheme
- With a canonical space inner-product ($< u, v >_X = \sum_i u_i v_i$)
- With a M-space inner product ($< u, v >_X^M = < Mu, v >_X$)

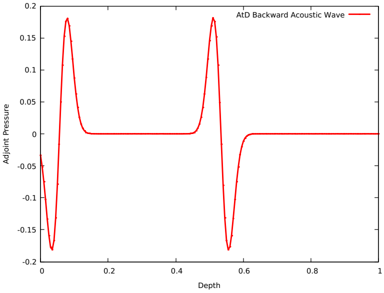
Adjoint test

Adjoint test passed for :

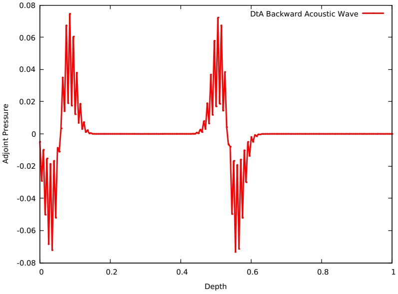
- Lagrange Operators
- Bernstein Operators
- Runge Kutta 4 time-scheme
- Adams Bashforth 3 time-scheme
- With a canonical space inner-product ($< u, v >_X = \sum_i u_i v_i$)
- With a M-space inner product ($< u, v >_X^M = < Mu, v >_X$)

```
./run
--- Adjoint test ----
 inner product UP/DUDP 553123.57586755091
 inner product GPGU/QPQU 553123.57586756046
./run
--- Adjoint test ----
 inner product UP/DUDP -75077.332007383695
 inner product GPGU/QPQU -75077.332007386358
./run
--- Adjoint test ----
 inner product UP/DUDP 125669.89223600870
 inner product GPGU/QPQU 125669.89223600952
./run
--- Adjoint test ----
 inner product UP/DUDP -132852.64215701097
 inner product GPGU/QPQU -132852.64215701059
```

Non consistency of the Adjoint solution



With the AtD strategy

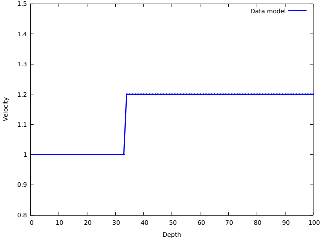


With the DtA strategy using the canonical inner-product (Lagrange+RK4)

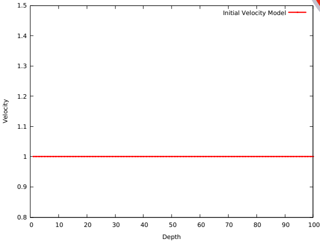
Adjoint test succeeds !

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

FWI Preliminary test (for all strategies)

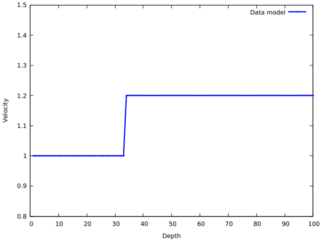


Target model

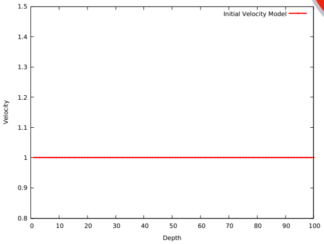


Initial model (iter=0)

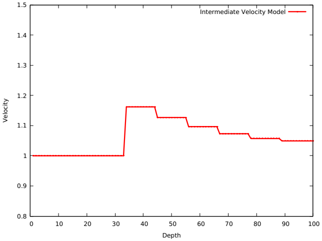
FWI Preliminary test (for all strategies)



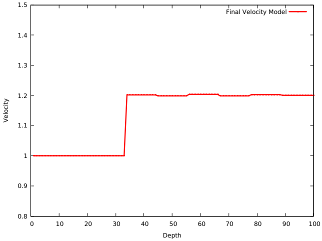
Target model



Initial model (iter=0)



Intermediate model (iter=20)



Final model (iter=50)

Qualitative Cost Function Gradient Study

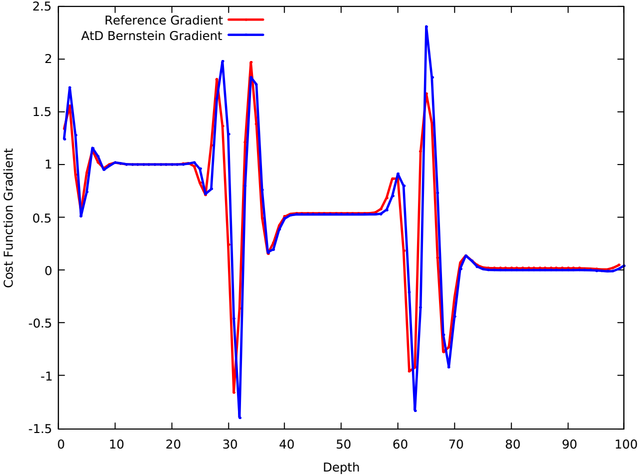


Figure: Comparison between a Reference Gradient and the FWI Gradient with AtD strategy (Bernstein elements and RK4 time scheme)

Qualitative Cost Function Gradient Study

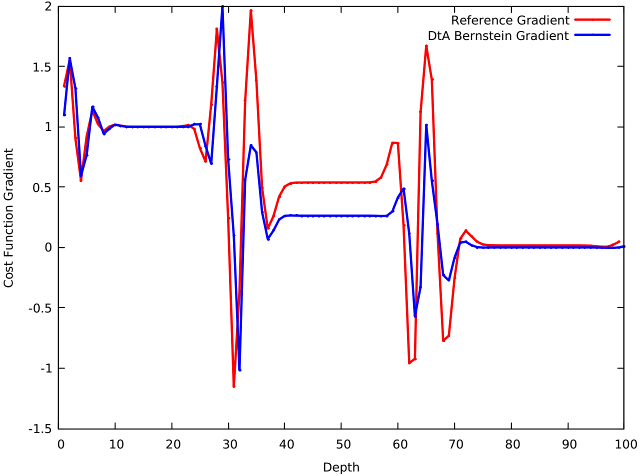


Figure: Comparison between a Reference Gradient and the FWI Gradient with DtA strategy (Bernstein elements and RK4 time scheme)

Qualitative Cost Function Gradient Study

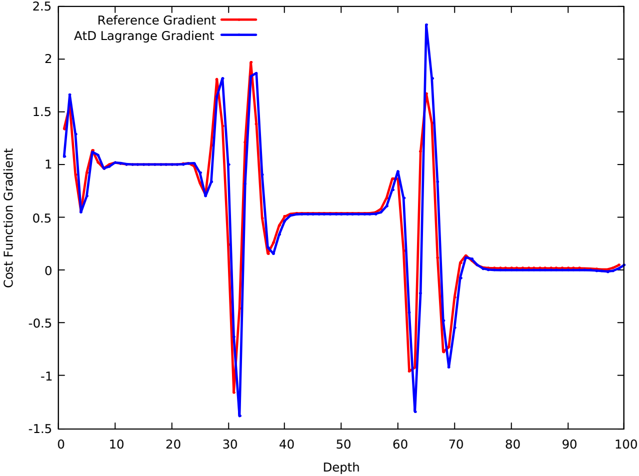


Figure: Comparison between a Reference Gradient and the FWI Gradient with AtD strategy (Lagrange elements and RK4 time scheme)

Qualitative Cost Function Gradient Study

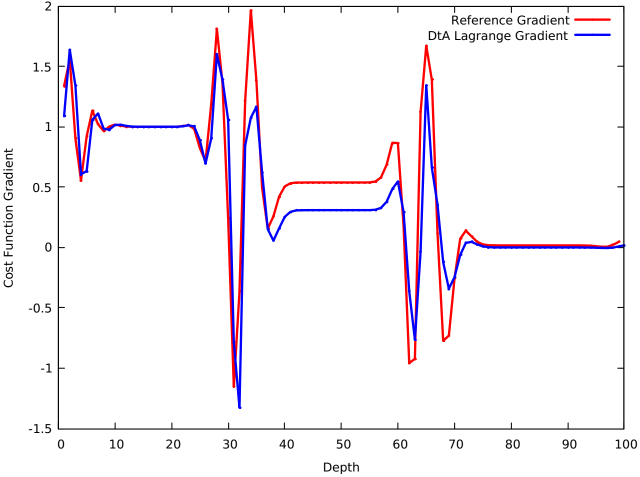


Figure: Comparison between a Reference Gradient and the FWI Gradient with DtA strategy (Lagrange elements and RK4 time scheme)

Conclusion and Perspectives

Conclusion :

- Adjoint then Discretized strategy works
- Discretized then Adjoint strategy has unexpected results (Gradient formulation ? Bug ?)
- The adjoint state is not consistent by using the Discretized and Adjoint strategy (but Adjoint test succeeds)

Conclusion and Perspectives

Conclusion :

- Adjoint then Discretized strategy works
- Discretized then Adjoint strategy has unexpected results (Gradient formulation ? Bug ?)
- The adjoint state is not consistent by using the Discretized and Adjoint strategy (but Adjoint test succeeds)

Perspectives :

- Complementary 1D tests
- 2D FWI + tests
- 3D FWI + tests
- Coupling SEM/DG elements (Aurélien Citrain's thesis)