

# Full Waveform Inversion using Proper Orthogonal Decomposition

Julien Besset<sup>1</sup>, H el ene Barucq<sup>1</sup>, Henri Calandra<sup>2</sup>, Stefano Frambati<sup>2</sup>

EPI Makutu, INRIA Bordeaux Sud-Ouest <sup>1</sup>, Universit e de Pau et des Pays de l'Adour, TotalEnergies <sup>2</sup>

October 3rd 2023



## 1 Full Waveform Inversion

- Inverse problem
- Application to seismic : Full Waveform Inversion

## 2 Proper Orthogonal Decomposition

- Introduction
- Proper Orthogonal Decomposition method
- Challenges
- POD application to the Acoustic Wave Equation

## 3 Results

- 2-D forward propagation FEM vs POD
- 3-D forward propagation SEM vs POD with GEOS
- Gradient computation and first FWI results with POD
- Conclusion

## 4 On going work

# FWI process

Full Waveform Inversion (FWI) is a quantitative inversion technique workflow.

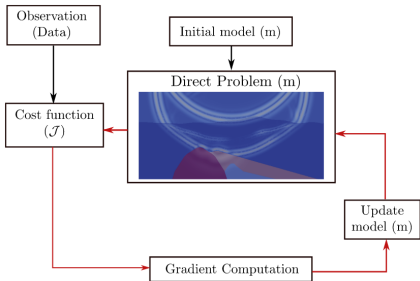


Figure: FWI process (*Jacquet (2021)*).

- Inverse problems aim to retrieve physical parameters of a medium from measured data.
- These parameters intervene in the partial differential equations which characterize the physical phenomenon of interest.
- The principle is general and independent of the physical problem (PDE).

Each step in FWI can be improved to reduce the computational burden. For instance,

- Numerical schemes (IGA formulation : [Frambati \(2021\)](#))
- Mesh and polynomial adaptivity (hp-adaptivity : [Jacquet \(2021\)](#))
- Adjoint state method with high order numerical schemes (frequency domain : [Faucher \(2017\)](#))

Lately, new techniques seem to have aroused the interest such as,

- Machine Learning ([Adler et al. \(2021\)](#)).
- Reduction Order Modeling (ROM) ([Borcea et al. \(2022\)](#)).

## Application to seismic : Full Waveform Inversion

- In a geophysical context, the inverse problem aims at reconstructing the subsurface physical parameters (velocity, density, Lamé parameters...).
- The observed data  $u_{obs}$  are recorded via a seismic acquisition consisting of sources and receivers for a medium. These data are called seismograms.
- The amount of data  $u_{obs}$  is often limited and usually localized in a small area of the domain near the surface.
- The cost function is the difference between the seismograms and simulated data.

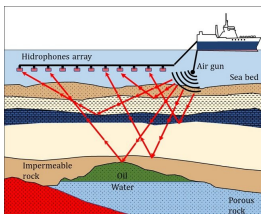


Figure: Data acquisition for seismic marine imaging

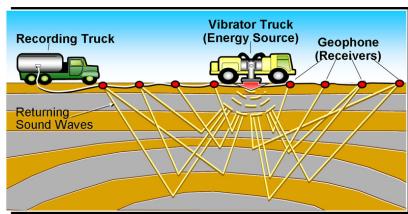


Figure: Data acquisition for seismic land imaging

# State equations

Depending on the physical parameters to be retrieved, different partial differential equations can be considered.

## Acoustic Wave Equation

$$\begin{cases} \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2}(\mathbf{x}, t) - \Delta u(\mathbf{x}, t) = f(\mathbf{x}, t) & \text{in } \Omega \times [0, T] \\ \alpha \frac{\partial u}{\partial t}(\mathbf{x}, t) + \frac{\partial u}{\partial n}(\mathbf{x}, t) = 0 & \text{on } \partial\Omega \times [0, T] \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad \frac{\partial u}{\partial t}(\mathbf{x}, 0) = u_1(\mathbf{x}) & \text{in } \Omega \end{cases} \quad (1)$$

## Elastic Wave Equation

$$\begin{cases} \rho \frac{\partial^2 \mathbf{u}}{\partial t^2}(\mathbf{x}, t) - \nabla \cdot (\underline{\underline{C}}(\underline{\underline{\epsilon}}(\mathbf{u}(\mathbf{x}, t)))) = \mathbf{f}(\mathbf{x}, t) & \text{in } \Omega \times [0, T] \\ \alpha \frac{\partial \mathbf{u}}{\partial t}(\mathbf{x}, t) + (\underline{\underline{C}}(\underline{\underline{\epsilon}}(\mathbf{u}(\mathbf{x}, t)))) \mathbf{n} = 0 & \text{on } \partial\Omega \times [0, T] \\ \mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), \quad \frac{\partial \mathbf{u}}{\partial t}(\mathbf{x}, 0) = \mathbf{u}_1(\mathbf{x}) & \text{in } \Omega \end{cases} \quad (2)$$

# FWI Setup

- Due to the size of the domain (several km) and the wave velocity, we need a huge amount of points to have a good representation of the problem ( $> 10^8$  cells).
- The choice of the discretization is crucial for a good estimation of the gradient.
- The lack of information on the medium makes this problem particularly ill-posed.
- To be able to reduce the space of the solutions, the choice of the starting model is key to avoid local minimum.

- 1 Full Waveform Inversion
  - Inverse problem
  - Application to seismic : Full Waveform Inversion
- 2 Proper Orthogonal Decomposition
  - **Introduction**
  - Proper Orthogonal Decomposition method
  - Challenges
  - POD application to the Acoustic Wave Equation
- 3 Results
  - 2-D forward propagation FEM vs POD
  - 3-D forward propagation SEM vs POD with GEOS
  - Gradient computation and first FWI results with POD
  - Conclusion
- 4 On going work



# ROM Introduction

The Reduction Order Modeling (ROM) is a technique for reducing the computational cost of complex models in numerical simulations by approximating the system of PDE or ODE with a lower order system.

ROM techniques can be classified as follows

- Proper orthogonal decomposition methods.
- Reduced basis methods.
- Balancing methods
- Simplified physics or operational based reduction methods.
- Nonlinear manifold methods.

# POD introduction

- Starting from an arbitrary system and experimental or simulated data, we can determine a new system (reduced model) of PDE or ODE that is representative of the starting system.
- By agreeing to pay the cost of one or several resolutions of our accurate model, we hope, for the same numerical cost, realize a greater amount of simulations with the reduced model.

# POD approximation

## POD approximation

We approach the solution  $u(x, t)$  of an arbitrary system with  $x \in \Omega$  space variable and  $t \in [0, T]$  time variable, in the form of a finite sum of separated variables functions product :

$$u(x, t) \approx \sum_{k=1}^K a_k(t) \phi_k(x) \quad (3)$$

- This representation of  $u$  is not unique.
- We can choose  $\phi_k$  as Legendre, Lagrange polynomial or trigonometric functions. In our case we choose  $\phi_k$  being linked to the solutions  $u$  of the full system.

# POD approximation

Supposing that the base  $(\phi_k)$  is orthonormal we have :

$$\langle u(x, t) | \phi_l(x) \rangle = a_l(t).$$

Solving (3) is equivalent to the following minimization problem:

$$\min \sum_i^{N_t} \left\| u(x, t_i) - \sum_{k=1}^K \langle u(x, t_i) | \phi_k(x) \rangle \phi_k(x) \right\|_2^2 \quad (4)$$

with  $N_t$  the number of discrete solutions in time.

# Snapshot matrix

We suppose that  $u(x, t)$  is known for  $N_x$  values of  $x$  and  $N_t$  values of  $t$ . We gather all the realizations of  $u(x_i, t_j)$  in a matrix called Snapshot:

Snapshot matrix

$$A = \begin{pmatrix} u(x_1, t_1) & u(x_1, t_2) & \cdots & u(x_1, t_{N_t}) \\ u(x_2, t_1) & u(x_2, t_2) & \cdots & u(x_2, t_{N_t}) \\ \vdots & \vdots & & \vdots \\ u(x_{N_x}, t_1) & u(x_{N_x}, t_2) & \cdots & u(x_{N_x}, t_{N_t}) \end{pmatrix} \in \mathbb{R}^{N_x \times N_t}$$

# Link between POD and SVD

Supposing  $u$  can be decomposed in an arbitrary finite element basis :

$$u_n(x, t_i) = \sum_{j=1}^n u_j(t_i) \varphi_j(x)$$

We can then define the following discrete scalar product :

$$\langle u | v \rangle_{\mathcal{M}} = u^T \mathcal{M} v$$

with  $\mathcal{M}$  the finite element mass matrix. We deduce the associated norm :

$$\|u\|_{\lambda} = \langle u | u \rangle_{\mathcal{M}}^{\frac{1}{2}} = \|(\mathcal{M}^{\frac{1}{2}})^T u\|$$

# Link between POD and SVD

We can reformulate the minimization problem (4) as:

$$\min_i \sum_{i=1}^{N_t} \|u_n(x, t_i) - \sum_{k=1}^K \langle u_n(x, t_i) | \phi_k(x) \rangle \phi_k(x)\|_{\mathcal{M}}^2 \quad (5)$$

where we suppose that the POD basis  $(\phi_k)$  can be written in the finite element basis  $(\varphi_j)_{j=1}^n$ .

## POD Minimization Problem

If we collect all the POD coefficient  $\phi_k^j$  in a matrix  $\phi \in \mathbb{R}^{n \times N_t}$ , (5) is equivalent to:

$$\min_{Z \in \mathbb{R}^{n \times K}} \|\tilde{A} - ZZ^T \tilde{A}\|_F^2 \quad (6)$$

with  $\tilde{A} = (\mathcal{M}^{\frac{1}{2}})^T A$  and  $Z = (\mathcal{M}^{\frac{1}{2}})^T \phi$

# Link between POD and SVD

## Expression of $\phi$

Finally, using the Eckart-Young theorem ([Cordier and Bergmann \(2006\)](#)) we find that  $\phi$  is solution to :

$$(\mathcal{M}^{\frac{1}{2}})^T \phi = U_k \in \mathbb{R}^{n \times K} \quad (7)$$

with  $U_k$  the  $k$ -st first column of the matrix  $U$  in the Singular Value Decomposition (SVD)  $\tilde{A} = U\Sigma V$ .

$U$  can be determined as the eigenvectors of  $\tilde{A}\tilde{A}^T$ . However if  $N_x \gg N_t$ , we prefer to evaluate  $V$  as the eigenvectors of  $\tilde{A}^T\tilde{A}$ , then deduce  $U$  using the relation:

$$U = \Sigma^{-1}\tilde{A}V \quad (8)$$



# Challenges

- POD construction requires a first resolution of the high fidelity model.  
⇒ Can't seem to be avoided.
- The POD is constructed for a particular set of physical parameters.  
⇒ We assume that for small perturbation of the physical parameters, the computed basis will still be correct.
- The search for the basis ( $\phi_i$ ) requires a SVD of the Snapshot matrix.  
⇒ Instead of doing a SVD on the full snapshot matrix, we iteratively (as we solve the full problem) compute the POD basis using QR-decomposition with a Gram-Schmidt orthonormalization process.

- 1 Full Waveform Inversion
  - Inverse problem
  - Application to seismic : Full Waveform Inversion
  
- 2 Proper Orthogonal Decomposition
  - Introduction
  - Proper Orthogonal Decomposition method
  - Challenges
  - **POD application to the Acoustic Wave Equation**
  
- 3 Results
  - 2-D forward propagation FEM vs POD
  - 3-D forward propagation SEM vs POD with GEOS
  - Gradient computation and first FWI results with POD
  - Conclusion
  
- 4 On going work

# POD formulation

Using the expression of  $u(\mathbf{x}, t)$  in the POD basis :

$$u(x, t) \approx \sum_{k=1}^K a_k(t) \phi_k(x)$$

We can re-write the previous system of equations as :

## Acoustic Wave Equation POD formulation

$$\left\{ \begin{array}{ll} \sum_{k=1}^K \frac{1}{c^2} \frac{\partial^2 a_k(t)}{\partial t^2} \phi_k(\mathbf{x}) - \Delta(a_k(t) \phi_k(\mathbf{x})) = f(\mathbf{x}, t) & \text{in } \Omega \times [0, T] \\ \sum_{k=1}^K \alpha \frac{\partial a_k(t)}{\partial t} \phi_k(\mathbf{x}) + a_k(t) \frac{\partial \phi_k(\mathbf{x})}{\partial n} = 0 & \text{on } \partial\Omega \times [0, T] \\ a_k(0) = \langle u_0(\mathbf{x}), \phi_k(\mathbf{x}) \rangle, \quad \frac{\partial a_k}{\partial t}(0) = \langle u_1(\mathbf{x}), \phi_k(\mathbf{x}) \rangle & \text{in } \Omega \end{array} \right. \quad (9)$$

# POD discretization

## Space discretization

$$\phi^T \frac{1}{c^2} \mathbf{M} \phi \frac{\partial^2 a(t)}{\partial t^2} + \phi^T \alpha \mathbf{S} \phi \frac{\partial a(t)}{\partial t} + \phi^T \mathbf{K} \phi a(t) = \phi^T f(t) \quad (10)$$

where  $\mathbf{M}$  is the mass matrix,  $\mathbf{K}$  is the stiffness matrix and  $\mathbf{S}$  is the damping matrix.

## Time discretization

$$a^{n+1} = \left( \mathbf{M}_\phi + \alpha \frac{\delta t}{2} \mathbf{S}_\phi \right)^{-1} \left( (2\mathbf{M}_\phi - \delta t^2 \mathbf{K}_\phi) a^n - \left( \mathbf{M}_\phi - \alpha \frac{\delta t}{2} \mathbf{S}_\phi \right) a^{n-1} + \delta t^2 f_\phi^n \right) \quad (11)$$

with  $\mathbf{M}_\phi = \phi^T \frac{1}{c^2} \mathbf{M} \phi$ ,  $\mathbf{S}_\phi = \phi^T \alpha \mathbf{S} \phi$ ,  $\mathbf{K}_\phi = \phi^T \mathbf{K} \phi$  and  $f_\phi = \phi^T f$

- 1 Full Waveform Inversion
  - Inverse problem
  - Application to seismic : Full Waveform Inversion
  
- 2 Proper Orthogonal Decomposition
  - Introduction
  - Proper Orthogonal Decomposition method
  - Challenges
  - POD application to the Acoustic Wave Equation
  
- 3 Results
  - 2-D forward propagation FEM vs POD
  - 3-D forward propagation SEM vs POD with GEOS
  - Gradient computation and first FWI results with POD
  - Conclusion
  
- 4 On going work

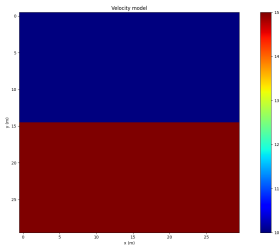


Figure: Two layers model

The model is a two layers model  $10\text{m/s}$  in the upper layer and  $15\text{m/s}$  in the bottom one.

We used a Q4 finite element mesh and the domain is covered with  $30\times 30$  elements.

# Vectors of the POD basis

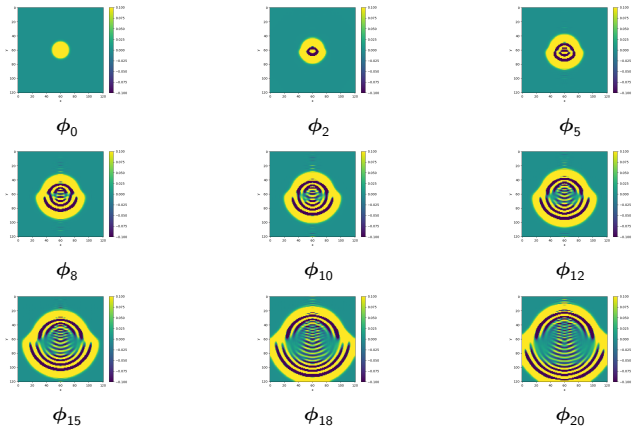


Figure: 9 of the 21 vectors of the POD basis built using QR decomposition

## 2-D forward propagation FEM vs POD

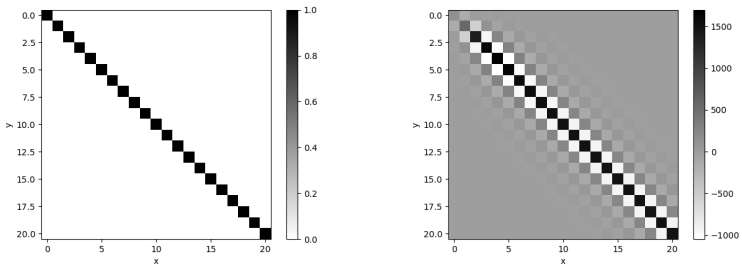
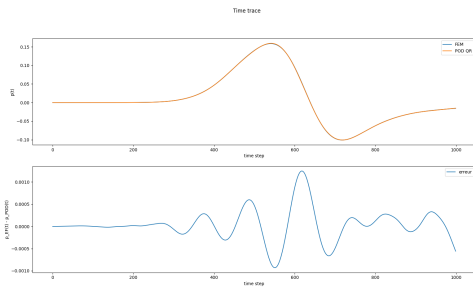


Figure:  $\phi^t \frac{1}{c^2} \mathbf{M} \phi$  and  $\phi^t \mathbf{K} \phi$  using the QR decomposition



## 2-D forward propagation FEM vs POD



**Figure:** Comparison of time trace in one point between the *FEM* solution and the *POD* solution

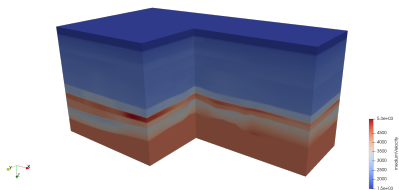


Figure: Equinor model

We have implemented the POD basis approach in the platform GEOS, and tested it on the 3-D Equinor model which has velocities from  $1500\text{m/s}$  up to  $5300\text{m/s}$ . The domain is meshed with Q1 Spectral Element Method (SEM) with  $\simeq 1.500.000$  elements.

# Vectors of the POD basis

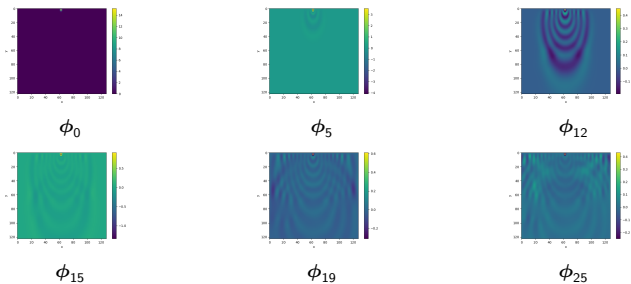


Figure: 6 of the 26 vectors of the POD basis built using QR decomposition

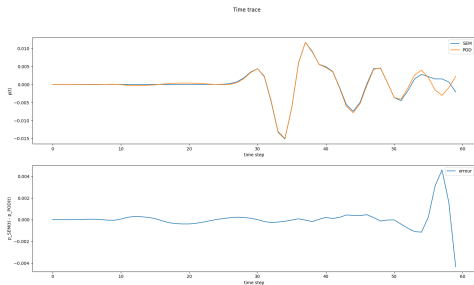
2MPI :

- Simulation = 9.1s
- Simulation + POD basis = 49.3s

12MPI :

- Simulation = 5.0s
- Simulation + POD basis = 13.3s

# Propagation comparison



**Figure:** Comparison of time trace in one point between the *SEM* solution and the *POD* solution

# Time computation

12MPI :

- SEM simulation = 5.0s  $\implies$  for 1.500.000 d.o.f
- POD simulation = 3.5s  $\implies$  regardless the number of d.o.f

Once we have paid the price of the POD basis and POD matrices construction, the more degrees of freedom we have in the domain, the better the POD is in terms of computational time.

# Inversion

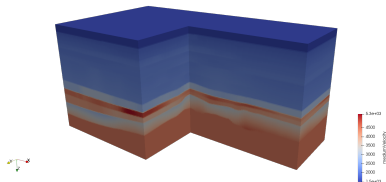


Figure: Equinor model

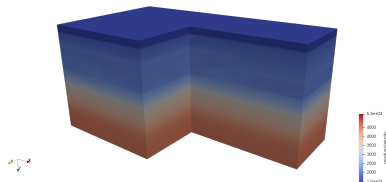
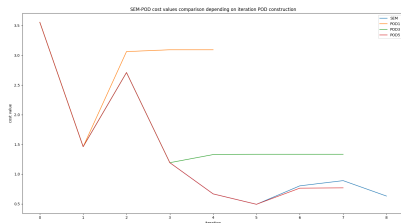


Figure: v0 model for FWI

- $\simeq 1.500.000$  cells
- $2520\text{m} \times 1440\text{m} \times 1520\text{m}$
- 30 sources
- 2560 receivers

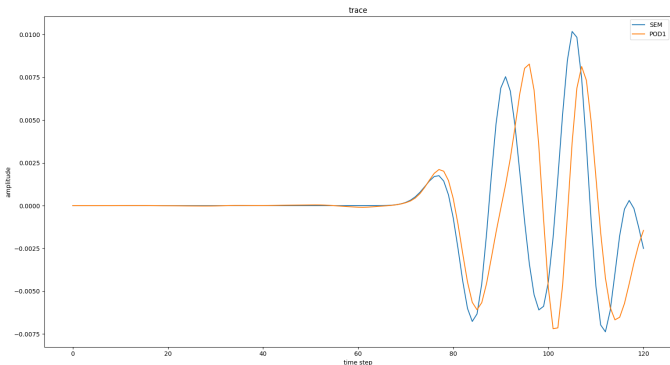
For this study, we decide to build 3 PODs from different SEM model iterations :  $POD_i$  means that the POD (forward and backward) is constructed from model  $m_{i-1}$  during the SEM gradient computation. Once the POD is built, it is used for the next models for gradient and linesearch computations.



**Figure:** Cost functions evolution between SEM and POD constructed at 1st, 3rd or 5th SEM iteration.

## Gradient computation and first FWI results with POD

We compare the seismic trace of the SEM and  $POD_1$  at one receiver location for the model  $m_1$ . We can clearly see a phase shift between the two traces. Because the velocity of the media changed between  $m_0$  and  $m_1$ , the wavelength is also modified, and the vectors of the POD basis, which are constructed from model  $m_0$ , do not have the new wavelength information.



**Figure:** Comparison of the seismic traces at one receiver location for the SEM and  $POD_1$  at model  $m_1$  given by SEM.



## Gradient computation and first FWI results with POD

The phase shift in the simulation has a direct impact on the gradient computation.

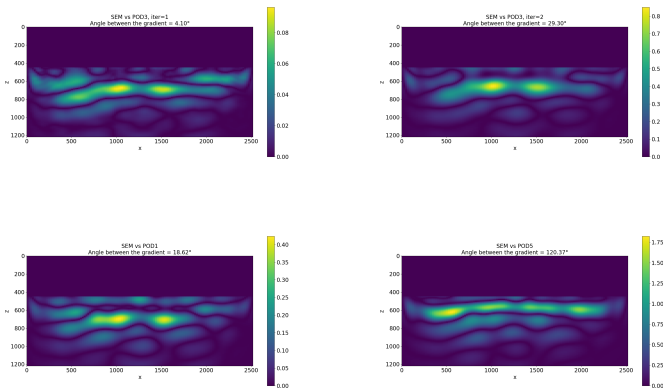


Figure: Relative errors and angle between gradient given by SEM and gradient given by the different PODs for the iterations following their construction.

- The QR approach for computing the POD basis is fast and cheap in terms of memory cost.
- The basis is source dependent.
- The basis is model dependent and not suitable for an FWI process.



# On going work

- Update the POD basis using machine learning techniques.  
⇒ The first tests suggest that we would need a high amount of samples to train the network. The size of the problems seems too big for neural networks techniques.
- Update the POD basis using minimization.
- Update the POD basis using Proper Generalized Decomposition (PGD) (*Oulghelou and Allery (2017)*).
- Try this POD within a Least Square Reverse Time Migration process.  
⇒ The basis is still source dependent.
- Build a POD using mode shapes of velocity model and independent of the source (*Basir et al. (2018)*).

# References

- Adler, A., Araya-Polo, M., and Poggio, T. (2021). Deep Learning for Seismic Inverse Problems: Toward the Acceleration of Geophysical Analysis Workflows. *IEEE Signal Processing Magazine*, 38(2):89–119.
- Basir, H. M., Javaherian, A., Shomali, Z. H., Firouz-Abadi, R. D., and Gholamy, S. A. (2018). Reverse time migration by Krylov subspace reduced order modeling. *Journal of Applied Geophysics*, 151:298–308.
- Borcea, L., Garnier, J., Mamonov, A. V., and Zimmerling, J. (2022). Waveform inversion via reduced order modeling. Number: arXiv:2202.01824 arXiv:2202.01824 [physics].
- Cordier, L. and Bergmann, M. (2006). Réduction de dynamique par décomposition orthogonale aux valeurs propres.
- Faucher, F. (2017). *Contributions à l'imagerie sismique par inversion des formes d'onde pour les équations d'onde harmoniques : Estimation de stabilité, analyse de convergence, expériences numériques avec algorithmes d'optimisation à grande échelle*. Phd, Pau.
- Frambati, S. (2021). *Unstructured Isogeometric Analysis with Applications to Seismic Wave Propagation*. Phd, Pau.
- Jacquet, P. (2021). Time-Domain Full Waveform Inversion using advanced Discontinuous Galerkin method.
- Oulghelou, M. and Allery, C. (2017). Contrôle optimal des écoulements par des modèles d'ordre réduit basés sur la POD et la PGD.

# Cost function

The reconstruction of the parameters is based on the minimization of a misfit function.

$$J(\mathbf{m}) = \frac{1}{2} \sum_t^T \sum_n^N \|(\mathcal{R}\mathbf{u})_{n,t} - (\mathbf{u}_{obs})_{n,t}\|^2 + \mathcal{F}_{n,t} \quad (12)$$

- $\mathbf{m}$  is the model parameters.
- $\mathbf{u}_{obs}$  are the observed data.
- $\mathcal{R}\mathbf{u}$  are the simulated data, with  $\mathcal{R}$  an injective operator.
- $N$  the number of observations.
- $T$  the number of time step.
- $\mathcal{F}$  a regularization term.

# Gradient computation and model update

Using a Taylor development on the expression of  $J$ , and assuming it reaches a minimum for  $\mathbf{m} = \mathbf{m}_0 + \Delta\mathbf{m}$  we find that the model update is given by:

$$\begin{aligned} \left[ \frac{\partial^2 J(\mathbf{m}_0)}{\partial \mathbf{m}^2} \right] \Delta\mathbf{m} &= \frac{\partial J(\mathbf{m}_0)}{\partial \mathbf{m}} \\ \iff H\Delta\mathbf{m} &= \nabla_{\mathbf{m}} J \end{aligned} \tag{13}$$

with  $H$  the Hessian of  $J$ .

$\nabla_{\mathbf{m}} J$  can be evaluated by different method. (ex : Adjoint state method).

# Iterative computation

The problem is then solved using iterative methods such as:

- Steepest descent
- Conjugate gradient
- Nonlinear conjugate gradient
- L-BFGS
- Gauss-Newton methods...
- Stochastic methods...



# Methods comparisons

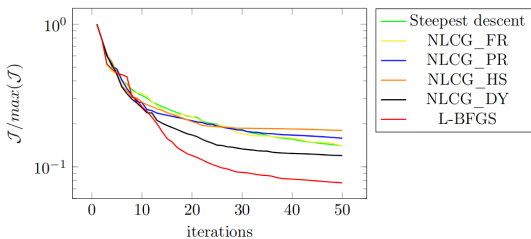


Figure: Cost function evolution for different search direction strategies (?).

# Eckart-Young theorem

The problem (4) is then equivalent to search for  $X$  with the same size as  $A$  but with a lower rank  $k$  such that for a certain norm, the error  $A - X$  is minimal. Taking the Frobenius norm, the solution of this problem is given explicitly by the Eckart-Young theorem :

$$\min_{\text{rank}(X) \leq k} \|A - X\|_F = \|A - A_k\|_F = \left( \sum_{j=k+1}^r \sigma_j^2(A) \right)^{\frac{1}{2}} \quad (14)$$

where

$$\begin{aligned} A_k &= U \begin{pmatrix} \Sigma_k & 0 \\ 0 & 0 \end{pmatrix} V^t \\ &= \sigma_1 u_1 v_1^t + \dots + \sigma_k u_k v_k^t \end{aligned}$$

with  $\sigma_j(A)$  are the eigenvalues of  $A$ .

This theorem establishes a relation between the rank  $k$  of the approximation  $X$  of  $A$  and the eigenvalue  $k$  of  $A$ . In conclusion, if the eigenvalues of  $A$  decrease rapidly enough, we can choose an approximation  $X$  of  $A$  with a very small rank.

# QR introduction

## QR Decomposition

Let  $A$  be a real  $m \times n$  matrix ( $m > n$ ) with  $\text{rank}(A) = n$ . The QR decomposition means that  $A$  can be written as :

$$A = QR \quad (15)$$

where  $Q$  is ( $m \times n$ ) orthogonal ( $Q^t Q = I_n$ ) and  $R$  is ( $n \times n$ ) upper triangular.

A very well known way of computing this decomposition is to use the Gram-Schmidt orthonormalization process.

# Classical Gram-Schmidt

The classical Schmidt algorithm consists of  $n$  steps. For each step, the  $k$ -th column of  $A$  will produce a new column of  $Q$ . Assuming we have already computed the  $k$  first columns of  $Q$ , the idea is to compute the projections of  $A_k$  onto all  $Q_i, i \in \{0, \dots, k-1\}$  :

$$r_{ik} = Q_i^t A_k$$

Then we can compute the vector  $\tilde{Q}_k$  :

$$\tilde{Q}_k = A_k - \sum_{i=0}^{k-1} r_{ik} Q_i \quad (16)$$

which is orthogonal to all previous  $Q_i$ . Finally we normalize this vector :

$$Q_k = \frac{\tilde{Q}_k}{\|\tilde{Q}_k\|}$$

However, it was soon observed that this algorithm is unstable. To be able to use this classical Gram-Schmidt orthonormalization and ensure the orthogonality, we are going to use reorthogonalization.

# Reorthogonalization

The instability in the classical Gram-Schmidt algorithm comes from the numerical approximation that is done by the computation of (16). An indication of this instability is given if :

$$\|\tilde{Q}_k\| \ll \|A_k\|$$

Or we can also use the Rutishauser criterion to decide if reorthogonalization is required, which stands for :

$$\|\tilde{Q}_k\| \leq \frac{1}{10} \|A_k\| \tag{17}$$

If this criterion is verified, then we proceed to a new orthogonalization of  $\tilde{Q}_k$  with respect to all  $Q_i, i \in \{0, \dots, k-1\}$ .

# POD based on QR

Because we want to find a reduce basis that will represent at best our solution, we are going to reject all the  $\tilde{Q}_k$  that verifies the Rutishauser criterion. This leads to the following algorithm :

```
l = 0
for k = 0 : m - 1 do
  Q_l = A_k
  success = 1
  for i = 0 : l - 1 do
    r_ik = Q_i^t A_k
    Q_l = Q_l - r_ik Q_i
    if ||Q_l|| ≤ 1/10 ||A_k|| then
      success = 0
      break
    end if
  end for
  if success then
    Q_l = Q_l / ||Q_l||
    l = l + 1
  end if
end for
```