

***SMAI-JCM***  
SMAI JOURNAL OF  
COMPUTATIONAL MATHEMATICS

Recent advances in numerical  
methods for solving the wave  
equation in the context of seismic  
depth imaging

HENRI CALANDRA, ZOÉ LAMBERT, CHRISTIAN GOUT, ANDREAS ATLE,  
MARIE BONNASSE-GAHOT, JULIEN DIAZ & SIMON ETTOUATI

Volume S5 (2019), p. 47-65.

<[http://smai-jcm.centre-mersenne.org/item?id=SMAI-JCM\\_2019\\_\\_S5\\_\\_47\\_0](http://smai-jcm.centre-mersenne.org/item?id=SMAI-JCM_2019__S5__47_0)>

© Société de Mathématiques Appliquées et Industrielles, 2019  
*Certains* droits réservés.



creative  
commons



Publication membre du

Centre Mersenne pour l'édition scientifique ouverte

<http://www.centre-mersenne.org/>

*l'Inria*



Sousmission sur <https://smai-jcm.math.cnrs.fr/index.php/SMAI-JCM>

## Recent advances in numerical methods for solving the wave equation in the context of seismic depth imaging

HENRI CALANDRA <sup>1</sup>  
ZOÉ LAMBERT <sup>2</sup>  
CHRISTIAN GOUT <sup>3</sup>  
ANDREAS ATLE <sup>4</sup>  
MARIE BONNASSE-GAHOT <sup>5</sup>  
JULIEN DIAZ <sup>6</sup>  
SIMON ETTOUATI <sup>7</sup>

<sup>1</sup> TOTAL SA, 64000 Pau, France

*E-mail address:* henri.calandra@total.com

<sup>2</sup> LMI, Normandie Univ., INSA Rouen, 76000 Rouen, France

*E-mail address:* zoe.lambert@insa-rouen.fr

<sup>3</sup> LMI, Normandie Univ., INSA Rouen, 76000 Rouen, France, and INRIA Bordeaux Sud Ovest - Magique3D, 64000 Pau, France

*E-mail address:* christian.gout@insa-rouen.fr

<sup>4</sup> TOTAL SA, 64000 Pau, France

*E-mail address:* andreas.atle@total.com

<sup>5</sup> TOTAL SA, 64000 Pau, France

*E-mail address:* marie.bonnasse-gahot@total.com

<sup>6</sup> INRIA Bordeaux Sud Ovest - Magique3D, 64000 Pau, France

*E-mail address:* julien.diaz@inria.fr

<sup>7</sup> TOTAL SA, 64000 Pau, France

*E-mail address:* simon.ettouati@total.com.

**Abstract.** In this paper, we present the recent advances in using discontinuous Galerkin method for solving wave equation in the context of seismic depth imaging and full wave inversion. We show some examples and the way forward to some advanced schemes coupling different numerical approximations we believe will provide the necessary tools for building the next seismic depth imaging generation codes for TOTAL Exploration&Production. This contribution is linked to the mini symposium (MS) *Mathematical tools in energy industry* (organized at Arcachon during the 9th International conference Curves and Surfaces).

**Keywords.** Numerical analysis, approximation, energy, HPC, finite elements method, Discontinuous Galerkin method, seismic depth imaging.

### 1. Introduction

Seismic depth imaging is the main tool used to extract information describing the geological structure of the subsurface from the recorded data in Oil and Gas industry. It is an inverse problem which consists in finding the best model minimizing the distance between the observed data and the predicted data. The process to estimate the predicted data is known as the process of forward modeling and is based on the resolution of the wave equation for some sources, initial conditions and boundary conditions. Efficiency for solving the forward modeling is crucial for geophysical imaging as one needs to get solutions of the PDEs for many sources and many iterations as we progressively improve our image and, therefore, the model we consider. Constant progresses in data acquisition and in rocks physics labs, more powerful computers and integrated team including physicists and applied mathematicians have greatly contributed to the development of advanced numerical algorithms integrating more and

more complex physics. For the last 20 years, our industry has been very active in the definition and introduction of different wave equation approximations and corresponding numerical methods for solving forward problem. But the real change came with the implementation of the full wave equation, thanks to the petascale era, giving access to a complete representation of the wavefield. It allowed geo-scientists to re-design imaging algorithm both in time dynamic and time harmonic domain. The most popular numerical scheme used nowadays by the industry is based on finite difference methods (FDM) on regular grids [13], [6]. We refer to [18] for examples of FDM in the geophysics frameworks and to [10] for 3D applications.

Finite Element Methods (FEM) (see [1], [9] or [16]) can use unstructured meshes to discretize accurately complex domains. Two kinds of FEM seem to be adapted to wave propagation simulation: Spectral Elements Methods (SEM) (see [15] for example of SE methods for geophysical applications) and discontinuous Galerkin Methods (DGM) (see [19] for a comparison between DGM and FDM applied to wave equations for seismic applications). SEM use high order functions and need less memory space than classical FEM without hampering the numerical convergence order. Discontinuous Galerkin Methods (DGM) have been increasingly studied for the resolution of differential models of linear wave propagation problems, particularly in the time domain [12] and [7]. DGM are more suitable than classical continuous FEM to deal with hp-adaptivity (interpolation degree  $p$  or mesh step  $h$  can change from element to another) [8], providing a greater flexibility in the mesh construction and the discretization of the different regions of the subsurface. In time domain, the block diagonal structure of the mass matrix leads to explicit representation of the solution and can be coupled with multirate time stepping strategy to relieve the CFL condition. Their main drawback comes from an increase of the number of degrees of freedom leading to an important computational cost (CPU time and memory) as compared to classical FEM. In time domain, following Chan's work, [4], we can take advantage of the Bernstein-Bézier basis, that induce blocked discretization matrices with sparse non-zero blocks. In harmonic domain, we use Hybridizable Discontinuous Galerkin Methods (HDGM) in order to decrease the size of the linear system to be inverted. The principle of HDGM consists in introducing an auxiliary unknown defined only on the faces of the element and in expressing all DG variables as functions of this new unknown. Hence, HDGM inherits all of the advantages of DGM, without the drawback of the increase of degrees of freedom [3].

In this paper, we present recent advances in using DGM for solving wave equation in the context of seismic depth imaging. In the first section, we briefly present the two most commonly used technologies in seismic depth imaging known as reverse time migration and full wave form inversion. In the second section dedicated to the time domain, we successively present the use of Bernstein-Bézier basis functions and the multirate time stepping methods applied to the isotropic elastic wave equation and how we can expect to improve both accuracy and efficiency of the corresponding numerical solver. The last section will introduce the HDGM in the context of solving the seismic full wave form inversion problem in time harmonic domain. It will be shown that the memory footprint of the linear operator can be dramatically reduced and take advantage of the most advanced direct linear solvers. The main objective of this paper is to provide to the reader a quick overview of some research activity on advanced numerical methods for solving the wave equation operator for seismic depth imaging in the context of Oil and Gas exploration. Main results presented here were obtained through an active collaboration with Inria Bordeaux-Sud Ouest (Pau) project-team MAGIQUE3D, Inria Sophia-Antipolis project-team NACHOS and Virginia-Tech.

This work is linked to a mini symposium (MS) organized during the 9th international conference Curves and Surfaces (Arcachon, june-july 2018). The MS was oriented toward energy applications in industry.

## 2. Seismic depth imaging

### 2.1. Introduction

The two main methods used for seismic depth imaging taking advantage of the full implementation of the wave equation are the Reverse Time Migration (RTM) and Full Wave Inversion (FWI). They both take advantage of the redundancy of the huge amount of seismic data recorded during the seismic acquisition. Seismic acquisition, see Figure 2.1, is the starting point to all seismic depth imaging project. It consists in emitting shock waves, the seismic sources, generally placed at the top of the subsurface, and to record on a discrete set of recorders, the seismograms, (fig.2), the echoes: travel time and corresponding amplitudes, of the reflected wave.

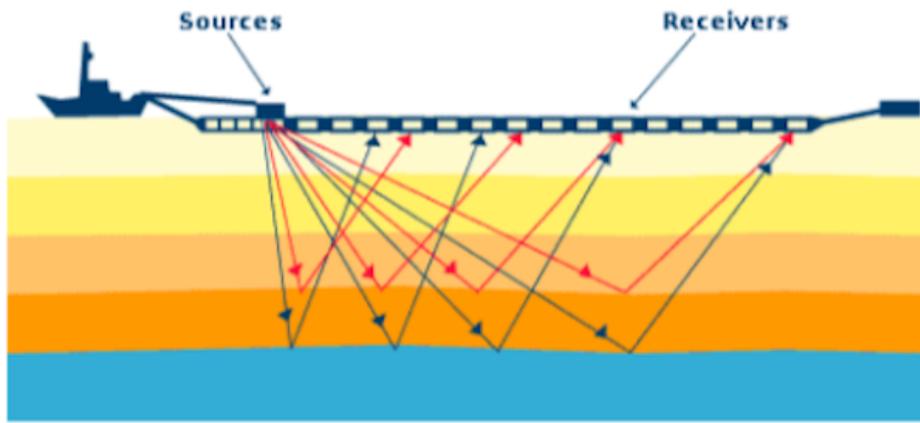


FIGURE 2.1. Principle of seismic acquisition

Seismic depth imaging consists in defining the best model of the subsurface, defined in terms of velocities and densities, explaining the seismic data recorded during the seismic acquisition campaign.

Reverse Time Migration (RTM) and Full Wave Inversion (FWI) are the most widely imaging technology used in the Oil and Gas industry for seismic depth imaging. Both methods are based on the resolution of wave equations.

Reverse time migration is based of the imaging condition introduced by J. Claerbout in 1971 [5] who noted that if there is a reflector in the subsurface, when the downgoing source wavefield hits the reflector, an upcoming reflection is generated traveling up to the receivers. In other words, wherever the source and receiver wavefields are in the same place at the same time, there must be a reflector. Thus, the imaging condition consists for all the shot profiles of the seismic acquisition, in correlating the solution of the forward wavefield with the solution of the backpropagation of the measured data. The imaging condition can be formulated in its simplest form, as

$$\mathcal{I} = \sum_t \sum_s (u_s \cdot u_d^*) \quad (2.1)$$

where  $(s, d)$  are source and receivers for a given shot profile,  $u_s$  is the forward wavefield solution of  $\mathcal{P}(m)u_s = s$  with  $m$  is an appropriate model and  $\mathcal{P}$  is the wave equation operator,  $u_d^*$  is the backpropagated wavefield, solution of the  $\mathcal{P}^*(m)u_d^* = d$ , with  $\mathcal{P}^*$  is the adjoint of the wave equation operator. For more details on the RTM technique, we refer to [5, 2].

Seismic Full Waveform Inversion (FWI) can be seen as an application of an optimal control problem determining the parameters of the PDEs governing wave propagation in a complex heterogeneous

medium. The method relies on the minimization of the cost function  $\mathcal{J}(m)$ , which is the difference between observed and predicted data. On time harmonic domain we have,

$$\mathcal{J}(m) = \frac{1}{2} \sum_{\omega} \sum_s \|\mathcal{R}_s(u_s(m)) - d_s\|^2, \quad (2.2)$$

where  $d$  are the data,  $m$  is the unknown or model,  $s$  represents the sources in the acquisition,  $\omega$  the angular (possibly complex) frequency,  $\mathcal{R}_s$  stands for the restriction operator on to the receivers location for source  $s$ ,  $u$  is solution of  $\mathcal{P}(m)u = g_s$  where  $\mathcal{P}$  is the appropriated wave equation operator and  $\mathcal{R}_u$  is the predicted data.  $\mathcal{J}$  minimization is commonly achieved through the use of Newton method:

$$m^{n+1} = m^n - (\nabla_m \nabla_m \mathcal{J}(m))^{-1} \nabla_m \mathcal{J}(m) = m^n - (H_{mm}(m))^{-1} \nabla_m \mathcal{J}(m), \quad (2.3)$$

where  $H$  refers to the Hessian. However, for the seismic inverse problem, the numerical cost of computing the second order derivative of the functional becomes rapidly overwhelming (due to the size of the discretized domain and the large number of sources in the seismic acquisition).

Alternative techniques have been developed for solving optimization problems, based on the first order derivative only or using Hessian approximation. The simplest approximation is to simply use a gradient descent algorithm, that uses only first order information,

$$m^{i+1} = m^i - \alpha \nabla_m \mathcal{J}(m^i), \quad (2.4)$$

where  $\alpha$  is a scale coefficient selected using a linesearch-type algorithm.

The Gradient computation of the objective function is based on the adjoint state method:

$$\begin{aligned} \nabla_m \mathcal{J}(m, u) &= \Re \left( \frac{d\mathcal{L}}{dm}(m, u, u^*) \right) \\ &= \Re \langle \partial_m (\mathcal{P}(m)u - g, u^*) \rangle \\ &= \Re \langle \partial_m \mathcal{P}(m)u, u^* \rangle, \end{aligned} \quad (2.5)$$

where  $u$  is the predicted wavefield solution of

$$\mathcal{P}(m)u = s, \quad (2.6)$$

$u^*$  is the adjoint state wavefield or backpropagated wavefield, solution of

$$\mathcal{P}^*(m)\gamma = -\mathcal{R}^*(\mathcal{R}(u) - d) \quad (2.7)$$

$\mathcal{P}$  is the wave equation operator, and  $\mathcal{P}^*$  is the adjoint of the wave equation operator.

We refer to [20] for a complete description and analysis of the FWI.

Efficiency for solving the forward and adjoint wave equation operator is crucial for geophysical imaging. Constant progresses in seismic data acquisition give access to more informations such as P-waves and Shear waves, surface waves,... all of them carrying out useful informations about the subsurface. Taking advantage of these huge amount of informations requires both the integration of more physics into the solvers and more accurate numerical methods to tackle the challenging depth imaging problems in seismic depth imaging. Still an important R&D effort on advanced numerical schemes for solving the wave equation operators is pursued in our industry. Figure 2.2 summarizes the importance of wave equation operator in seismic depth imaging.

## 2.2. Discontinuous Galerkin Finite Element for elastic wave propagation

Elastic depth imaging is a real challenge for seismic depth imaging. One of the main challenges is to provide accurate solvers flexible enough to accurately represent the different waves propagations, the high discontinuity material properties, the complex geometries, such as irregular sea-floor or surface acquisition topography. Very efficient on modern computers, Finite Difference Methods, FDM, are the main numerical methods used for seismic depth imaging. However, despite their efficiency, FDM

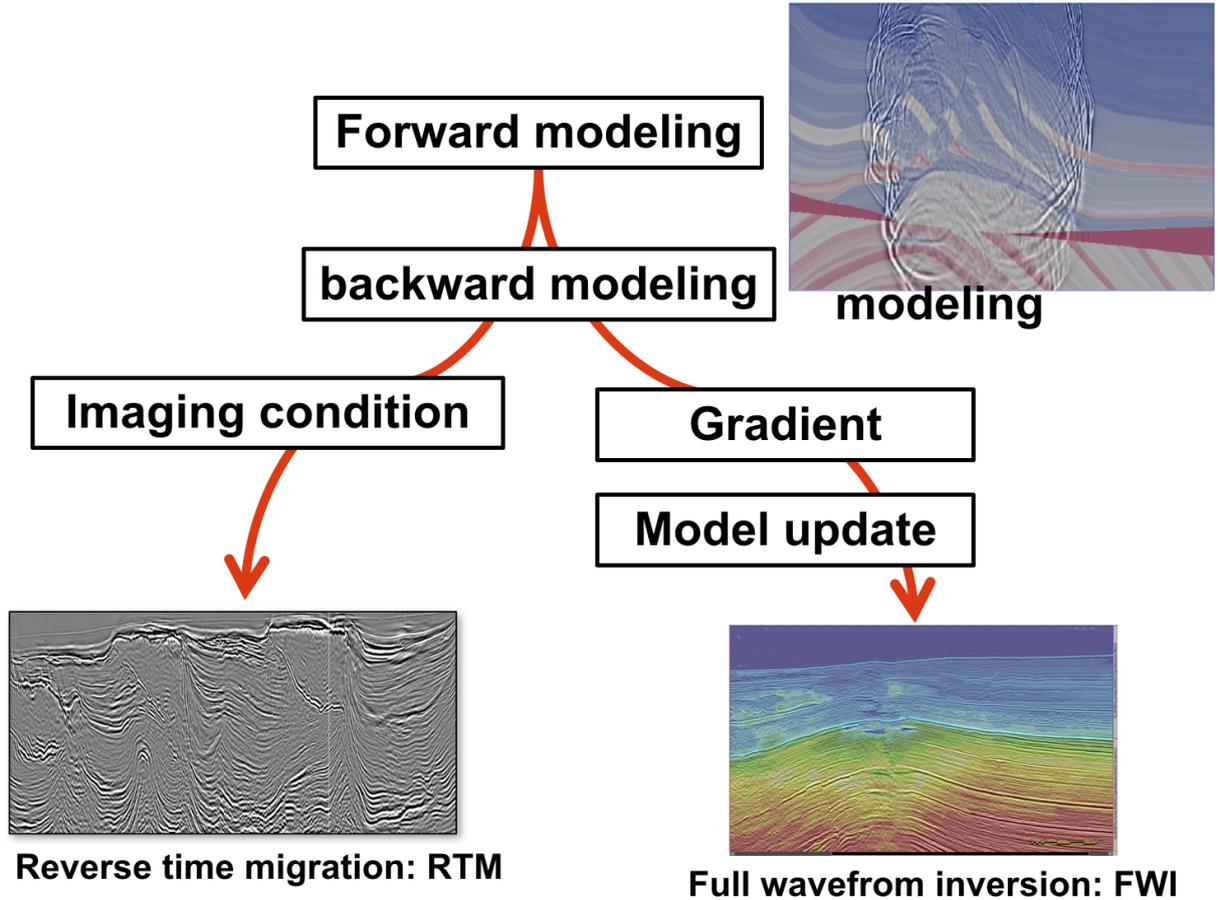


FIGURE 2.2. Wave equation operator and seismic depth imaging

have some limitations for approximating highly heterogeneous media and irregular topographies, for coupling different physics (acoustic-elastic) or are strongly constrained in the case of explicit time marching algorithm by the CFL condition.

By contrast, FEM by construction can use unstructured meshes to discretize accurately complex domains and allow efficient multiscale approach in time and space. However, continuous Galerkin FEM (CGFEM) introduce the need to invert the mass matrix, which is not easily invertible and especially bad for explicit methods.

Discontinuous Galerkin Finite element methods (DGFEM) have been increasingly studied for the resolution of differential models of linear wave propagation problems, particularly in the time domain [12] and [7] and more recently in frequency domain [3]. They are particularly well designed to deal with:

- $hp$ -adaptivity (interpolation degree  $p$  or mesh step  $h$  can change from element to another), providing a greater flexibility in the mesh construction and the discretization of the different regions of the subsurface.
- Explicit representation in time domain, taking advantage of the block diagonal structure of the mass matrix which can be coupled with multirate time stepping strategy to relieve the CFL condition.

- Efficient mapping many supercomputer architectures.

The following Figure 2.3 gives a quick comparison of the different numerical methods and limitations.

	Complex geometries	High-order accuracy and <i>hp</i> -adaptivity	Explicit semi-discrete form
FDM	✗	✓	✓
FVM	✓	✗	✓
FEM	✓	✓	✗
DG-FEM	✓	✓	✓

FIGURE 2.3. Method comparison (From [11])

However, despite their flexibility, DGFEM have still some drawbacks:

- The increasing number of degrees of freedom compared to classical FEM leads to an important increase of the cost in CPU and memory and limits the interest of DGFEM for real applications both in time dynamic and harmonic domain.
- In time dynamic domain, optimal unstructured meshes combined to material properties results in very constrained time discretization sampling, increasing CPU time if not localized.

To address these issues we have explored different strategies:

- For time-domain, we couple the discontinuous Galerkin methods on tetrahedral meshes based on the Bernstein-Bézier basis to multirate time step implementation. Our objective is to take advantage of the optimal complexity algorithms (constant cost per degree of freedom with respect to order of approximation) of Bernstein-Bézier basis function and to reduce time sampling constrain.
- For time harmonic domain, we use Hybridizable Discontinuous Galerkin Finite Element to reduce the number of degrees of freedom and so the size of the resulting linear system to invert.

### 3. Bernstein-Bézier and multirate time stepping for the elastic wave equation

The choice of high-order Lagrange elements requires operations with block dense matrices that create computational bottlenecks, limiting computational efficiency on modern HPC architectures. Figure 3.1 shows an example of the sparsity pattern of derivative operator using nodal basis of polynomial order  $n = 4$ .

Recent work done by J. Chan et al. 2017 [4], addresses this issue by investigating the use of Bernstein-Bézier basis functions, that induces blocked discretization matrices with sparse non-zero blocks. Computational experiments have confirmed the advantage of Bernstein-Bézier DG kernels over both straightforward and block-partitioned nodal DG kernels at high orders of approximation. For a complete description of the history of the Bernstein-Bézier basis functions we refer to Chan et al. [4]

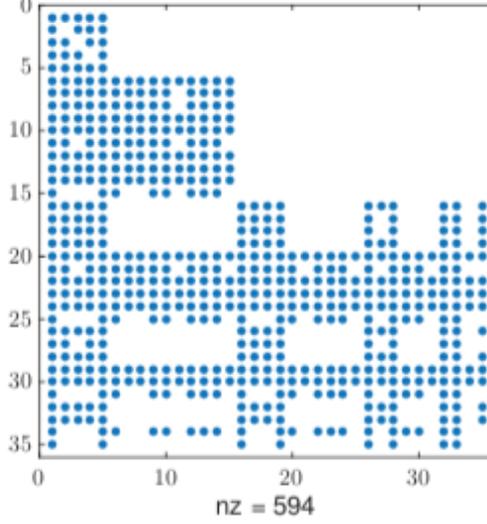


FIGURE 3.1. Nodal derivative matrix sparsity.

The key idea followed by Chan was to use the “strong” DG formulation. In this case, the mass matrix is factored out of the volume terms and removed after multiplying with the inverse mass matrix on both sides. This leaves only the application of a derivative operator, and decreases the number of matrix multiplications required to evaluate the volume term. To evaluate the surface term, the inverse of the mass matrix is fused into the face mass matrices to produce the face lift matrices. The lift can itself be decomposed into two sparse matrices.

The Bernstein-Bézier basis functions are not nodal, but they have sparse matrix operations on each tetrahedron element. The coordinates inside a tetrahedron can be expressed in barycentric coordinates,

$$\bar{x}(\bar{\lambda}) = \sum_{i=1}^4 \lambda_i \bar{v}_i, \quad \sum_{i=1}^4 \lambda_i = 1,$$

where the  $\bar{v}_i$  are the vertices of the tetrahedron and the  $\lambda_i$  are the barycentric coordinates of a point  $\bar{x}$ . The Bernstein-Bézier basis functions on the tetrahedron are the different terms in the expansion of

$$(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)^n, \quad B_{\bar{\alpha}}^n(\bar{\lambda}) = \binom{n}{\bar{\alpha}} \bar{\lambda}^{\bar{\alpha}}, \quad |\bar{\alpha}| = n,$$

where  $n$  is the polynomial order of the basis functions. The property that makes sparse matrix-operations is the degree-elevation operator, that increases the polynomial order by one, but combining up to 4 Bernstein-Bézier functions of order  $n$  to obtain a function of order  $n + 1$  [4].

The degree elevation operator is used to obtain a sparse derivative operator for derivatives w.r.t. barycentric coordinates. These barycentric derivatives are then combined to get the derivatives of the reference element coordinates,

$$\nabla_{\mathbf{x}} = J_K^{-T} \nabla_{\mathbf{r}} = J_K^{-T} \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \nabla_{\boldsymbol{\lambda}},$$

where  $\mathbf{r}$  are the reference coordinates and  $\boldsymbol{\lambda}$  are the barycentric coordinates. The barycentric derivatives are sparse in the Bernstein-Bézier basis as shown in Figure 3.2.

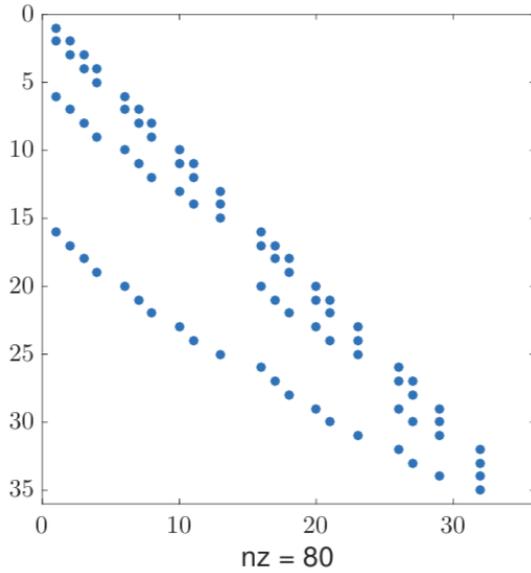


FIGURE 3.2. Sparsity patterns of derivative operators using Bernstein-Bézier for  $N = 4$ .

The LIFT-operator is a projection from the surfaces of the element onto the element as described in [4]. In order to gain efficiency, the Bernstein-Bézier LIFT-operator can be compressed by a factorization,

$$LIFT = E_L \cdot L_0,$$

where  $E_L$  and  $L_0$  are sparse matrices as presented in Figure 3.3 for polynomial order 4.

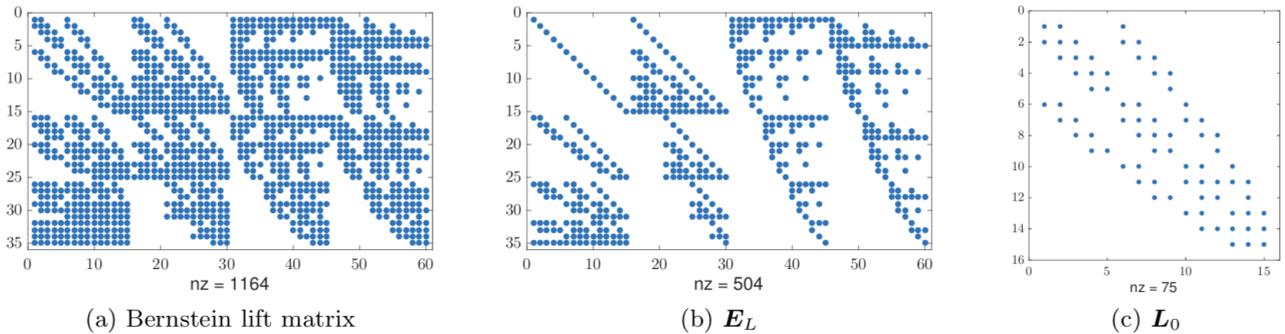


FIGURE 3.3. Sparsity patterns of lift operators  $L$ ,  $E_L$ , and  $L_0$  for a Bernstein-Bézier basis of degree  $N = 4$ .

#### 4. Multirate time stepping

While simulating numerically the wave propagation using DGM, one can expect to have locally refined meshes. This can lead to severe constraints on the stability condition. Figure 4.1 is an example showing the time step distribution for a complex model. In this case smallest time steps represent 0.1% of the mesh when more than 95% of the meshes can accept a time step 16 time longer. Implementing multirate time stepping avoids the requirement of taking small global timesteps and thus enables more efficient computations, see e.g. [14].

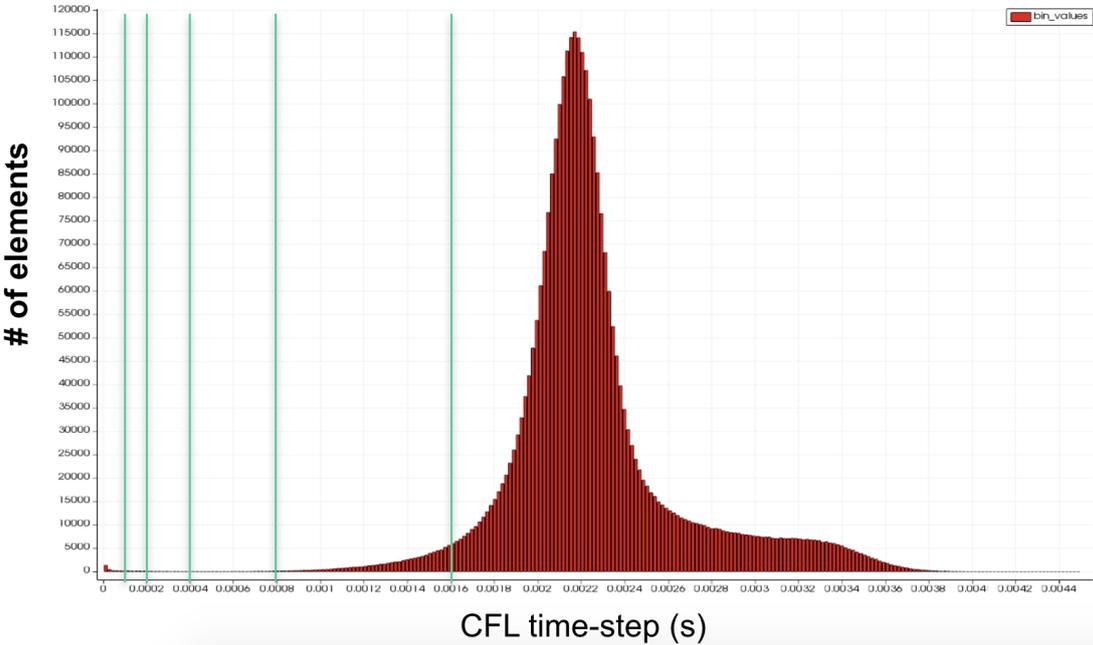
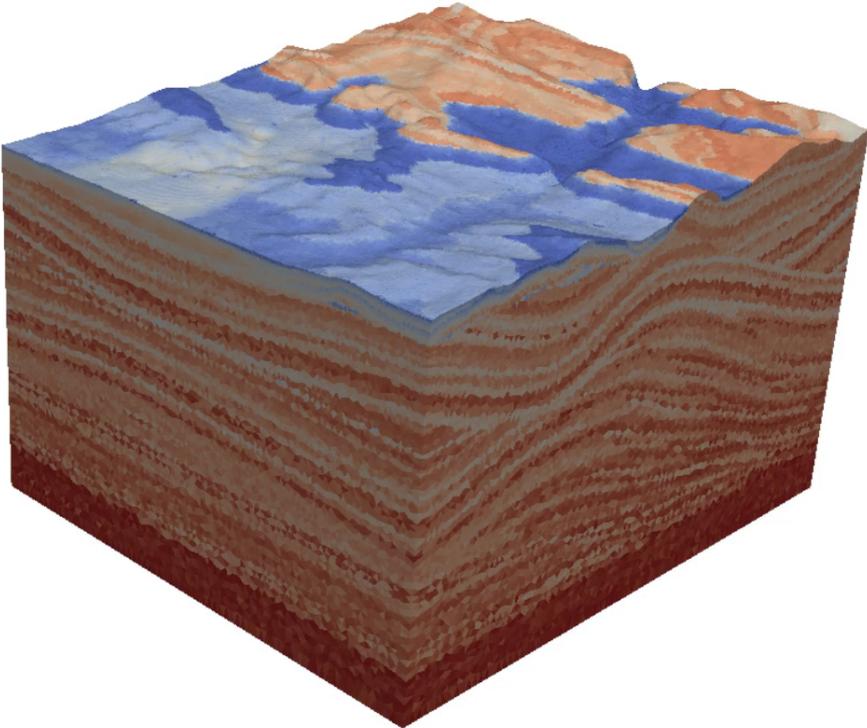


FIGURE 4.1. Time step distribution for a complex geological model.

#### 4.1. General idea

Starting from the following ODE:

$$\frac{d}{dt} \mathbf{U}(t) = \mathbf{F}(t, \mathbf{U}(t)) \quad (4.1)$$

with  $\mathbf{F}(t, \mathbf{U}(t))$  the Right Hand Side (RHS) of the ODE.

Let us recall that it is well known that the Adams-Bashforth (AB) method with local time stepping is very efficient with lower CPU time than many other methods. If we discretize equation (4.1) with a single rate (AB) method then the fully discrete equations are:

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \Delta t \left( a_0 \mathbf{F}(t^n, \mathbf{U}^n) + a_1 \mathbf{F}(t^{n-1}, \mathbf{U}^{n-1}) + a_2 \mathbf{F}(t^{n-2}, \mathbf{U}^{n-2}) \right)$$

with  $\Delta t$  a global timestep and  $(a_0, a_1, a_2)$  the coefficients for the AB time integrator. In the multirate Adams-Bashforth method (MRAB), we use the Adams-Bashforth method with different timesteps for different elements in the domain. We only allow to change the timestep by factors of a half, where each level is determined by the number of factors. We do not allow neighboring elements to differ with more than one level. In order to compute a right-hand-side in Adams-Bashforth, we need the wavefield for the element and its neighbor faces. The not so obvious step is to provide the wavefield for a level where the neighbor element is from a lower level that is not computed for the current timestep.

#### 4.2. Different multirate time integrators

The standard AB method has three parameters  $a_0$ ,  $a_1$  and  $a_2$ , to compute the next timestep

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \Delta t \left( a_0 \mathbf{F}(t^n, \mathbf{U}^n) + a_1 \mathbf{F}(t^{n-1}, \mathbf{U}^{n-1}) + a_2 \mathbf{F}(t^{n-2}, \mathbf{U}^{n-2}) \right). \quad (4.2)$$

The wavefields that we need from the neighbor of lower level can be computed by taking a half step with different parameters  $b_0$ ,  $b_1$  and  $b_2$ ,

$$\mathbf{U}^{n+1/2} = \mathbf{U}^n + \Delta t \left( b_0 \mathbf{F}(t^n, \mathbf{U}^n) + b_1 \mathbf{F}(t^{n-1}, \mathbf{U}^{n-1}) + b_2 \mathbf{F}(t^{n-2}, \mathbf{U}^{n-2}) \right). \quad (4.3)$$

When the lower level has been computed we overwrite the wavefield with the wavefield half the local timestep forward. When we compute the next timestep, we need to take into account that the wavefield has been modified, and the new update that is equivalent to using the “standard” AB becomes

$$\mathbf{U}^{n+1} = \mathbf{U}^{n+1/2} + \Delta t \left( (a_0 - b_0) \mathbf{F}(t^n, \mathbf{U}^n) + (a_1 - b_1) \mathbf{F}(t^{n-1}, \mathbf{U}^{n-1}) + (a_2 - b_2) \mathbf{F}(t^{n-2}, \mathbf{U}^{n-2}) \right). \quad (4.4)$$

Figure 4.2 shows the execution order of those different local time integrators for a two-rate AB3 method.

#### 4.3. DG formulation of the isotropic elastic wave equation

In this section, we extend developments for the isotropic acoustic wave equation operator [4] to the isotropic elastic case. Let us consider the isotropic elastic wave equation

$$\partial_t \bar{\sigma} = C \begin{pmatrix} \mathcal{D}_1 \\ \mathcal{D}_2 \end{pmatrix} \bar{v} \quad (4.5)$$

$$\rho \partial_t \bar{v} = \begin{pmatrix} \mathcal{D}_1 & \mathcal{D}_2 \end{pmatrix} \bar{\sigma}, \quad (4.6)$$

where

$$\bar{\sigma} = \begin{pmatrix} \sigma_{xx} & \sigma_{yy} & \sigma_{zz} & \sigma_{yz} & \sigma_{xz} & \sigma_{xy} \end{pmatrix}^T, \quad \bar{v} = \begin{pmatrix} v_x & v_y & v_z \end{pmatrix}^T,$$

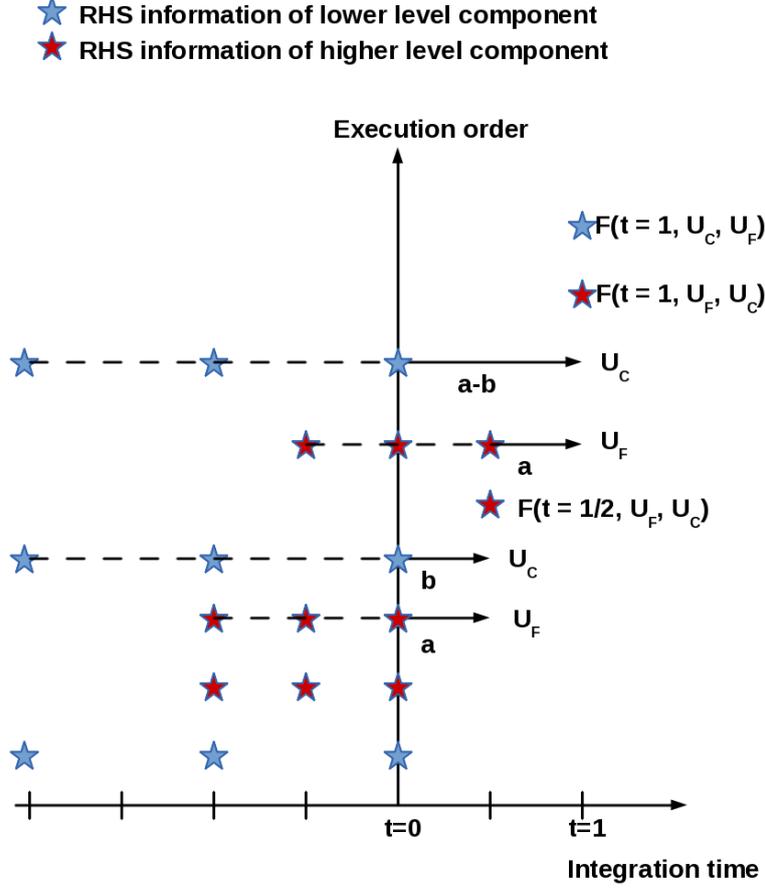


FIGURE 4.2. Two-rate AB3 method diagram.

with  $6 \times 6$  stiffness tensor  $C$  and the derivative matrices

$$\mathcal{D}_1 = \begin{pmatrix} \partial_x & & \\ & \partial_y & \\ & & \partial_z \end{pmatrix}, \quad \mathcal{D}_2 = \begin{pmatrix} & \partial_x & \partial_x \\ \partial_y & & \partial_y \\ \partial_z & \partial_z & \end{pmatrix}.$$

In the DG-formulation, each wavefield is expressed in basis functions  $\sigma_{xx} = \bar{\varphi}^T \bar{\sigma}_{xx}$  on each tetradron element. We can write the coefficients on element  $K$

$$\bar{\sigma}^K = (\bar{\sigma}_{xx} \quad \bar{\sigma}_{yy} \quad \bar{\sigma}_{zz} \quad \bar{\sigma}_{yz} \quad \bar{\sigma}_{xz} \quad \bar{\sigma}_{xy})^T, \quad \bar{v}^K = (\bar{v}_x \quad \bar{v}_y \quad \bar{v}_z)^T.$$

The discretized equations are

$$\partial_t \bar{\sigma} = C \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \bar{v} + LIFT F_{\bar{\sigma}} \quad (4.7)$$

$$\rho \partial_t \bar{v} = (D_1 \quad D_2) \bar{\sigma} + LIFT F_{\bar{v}}, \quad (4.8)$$

with fluxes

$$F_{\bar{\sigma}} = \frac{1}{2} \left( N [\bar{v}] + \frac{1}{\{\rho v_p\}} N N^T [\bar{\sigma}] \right) \quad (4.9)$$

$$F_{\bar{v}} = \frac{1}{2} \left( N^T [\bar{\sigma}] + \{\rho v_p\} \hat{n} \hat{n}^T [\bar{v}] + \{\rho v_s\} (I - \hat{n} \hat{n}^T) [\bar{v}] \right). \quad (4.10)$$

The first term in  $F_{\bar{\sigma}}$  and  $F_{\bar{v}}$  comes from integration by parts, and the remaining terms are penalization. The derivative matrices  $D_1$  and  $D_2$  have the continuous derivatives  $\partial_\xi$  replaced with matrices  $D_\xi$ , and the “normal” matrices  $N_1$  and  $N_2$  have the same shape as the derivative matrices, with e.g.  $D_\xi$  replaced with  $\hat{n}_\xi I$ . We further denote  $N = (N_1, N_2)^T$ . The LIFT-operator projects the wavefields on the faces of the element onto its volume. When the order of two adjacent elements are different, for instance order  $p$  for element  $K$  and order  $q$  for its neighbor, we apply the projection operator  $\Pi_{pq} - M_{pp}^{-1} M_{pq}$  of the wavefield of order  $q$  to the wavefield of order  $p$ , where  $M_{pq}$  is the mass matrix with basis function of order  $p$  and  $q$ . The projection is a non-sparse operator, except for Bernstein-Bezier when  $p > q$ ; then the projection is the sparse degree-elevation operator from order  $q$  to order  $p$ .

At the top of the domain we use free surface conditions that are imposed weakly by changing the jumps in the fluxes. The jumps at the free surface are replaced by

$$[\bar{\sigma}] = -2\sigma^F, \quad [\bar{v}] = 0.$$

At the other boundaries we weakly impose absorbing boundary conditions,

$$[\bar{\sigma}_1] = -\rho v_p N_1 \bar{v}^F - \bar{\sigma}_1^F, \quad [\bar{\sigma}_2] = -\rho v_s N_2 \bar{v}^F - \bar{\sigma}_2^F, \quad [\bar{v}] = 0,$$

where  $\bar{\sigma}_1 = (\bar{\sigma}_{xx}, \bar{\sigma}_{yy}, \bar{\sigma}_{zz})^T$ , and  $\bar{\sigma}_2 = (\bar{\sigma}_{yz}, \bar{\sigma}_{xz}, \bar{\sigma}_{xy})^T$ .

In terms of implementation, note that the fluxes are not precomputed, the values of the solution on the neighbors are obtained on the fly for each face of the element.

## 5. Numerical results

### 5.1. 2D

Multirate time scheme has been validated qualitatively in 2d on a Marmousi test case. We have run the two algorithms on the same configuration and checked that both methods deliver similar results. In Figures 5.1 and 5.2 we depict two snapshots, printed at the same time, of the velocity along the  $z$ -axis for both single rate AB (Fig. 5.1) and three rates AB (Fig. 5.2). The results look the same.

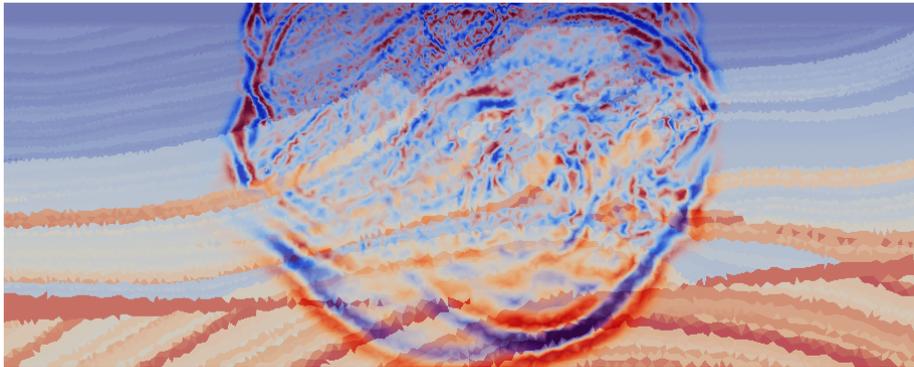


FIGURE 5.1. Snapshot for single-rate AB with DG-elastic-iso propagator.

### 5.2. 3D

MRAB, combined with a variable distribution of polynomial orders has been tested on a 3D isotropic elastic model. Figure 5.3 and 5.4 show the distribution of polynomial orders that is used: mainly P4-elements. Figure 5.5 shows the histogram of the time-step estimation and Figure 5.6 and 5.7 the

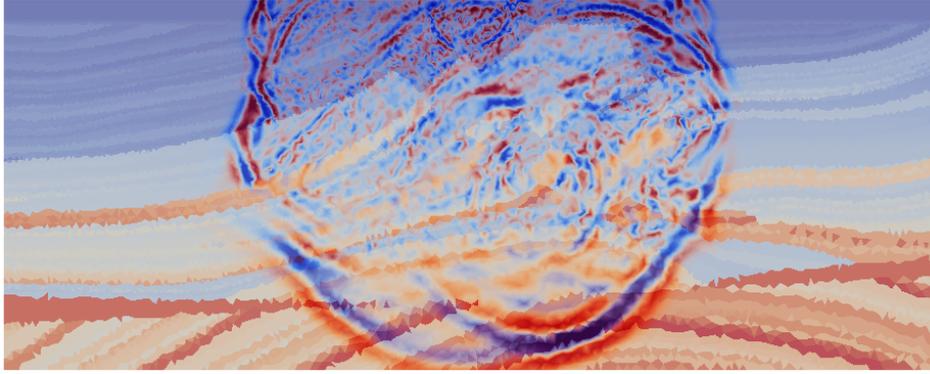


FIGURE 5.2. Snapshot for MRAB (3 levels) with DG-elastic-iso propagator.

corresponding distribution of multirate levels. Let us note that the code has been implemented to satisfy the CFL requirement and not only the mesh size, and therefore our approach is suitable both to fit mesh size and wave speed.

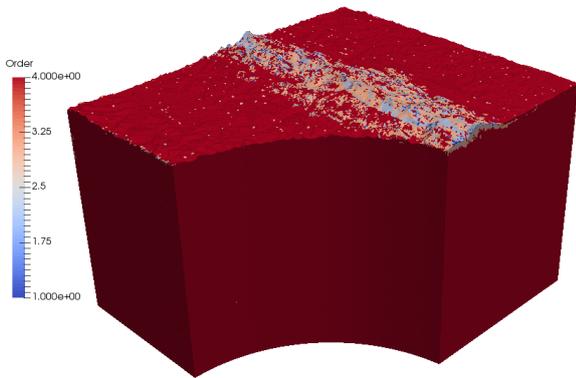


FIGURE 5.3. Distribution of polynomial order for 3D isotropic elastic model test case.

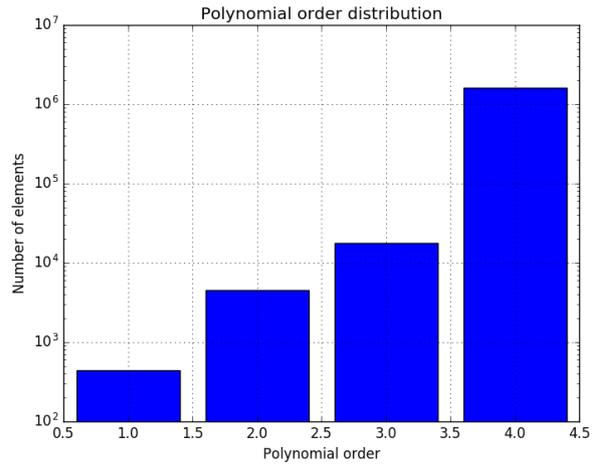


FIGURE 5.4. Bar chart of the polynomial order distribution (log scale).

### 5.3. Isotropic elastic time DG FEM summary

Taking advantage of DG natural hp refinement we can reduced the number of meshes and adapt polynomial order to the size of the mesh for preserving accuracy. Replacing standard Lagrange polynomial basis function by the Bernstein-Bézier basis function combined to a multirate time stepping implementation has improved both accuracy and performance of our isotropic elastic solver as shown in Figure 5.9.

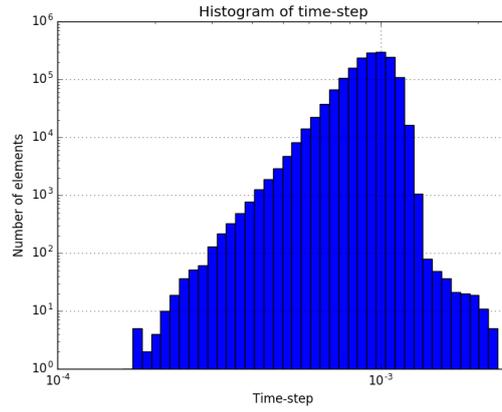


FIGURE 5.5. Histogram of the time-step distribution (log-log scale).

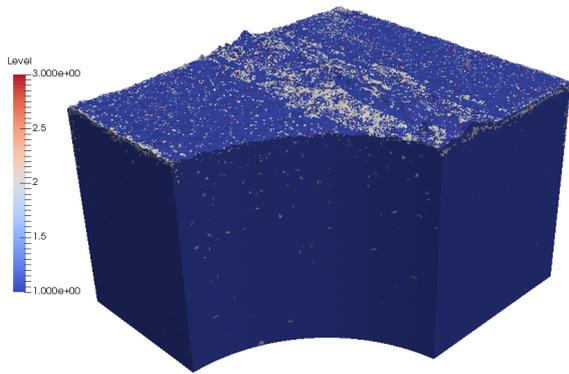


FIGURE 5.6. Distribution of multirate level for 3D isotropic elastic model test case.

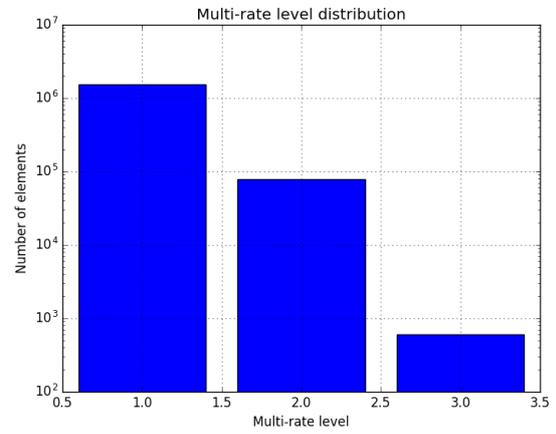


FIGURE 5.7. Bar chart of the multirate level distribution (log scale).

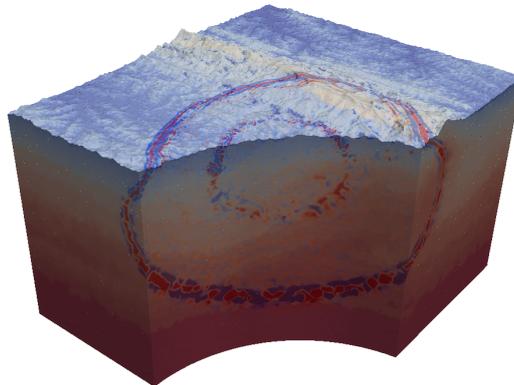


FIGURE 5.8. Snapshot of the wavefield  $\sigma_{zz}$  for MRAB3, with 3 levels.

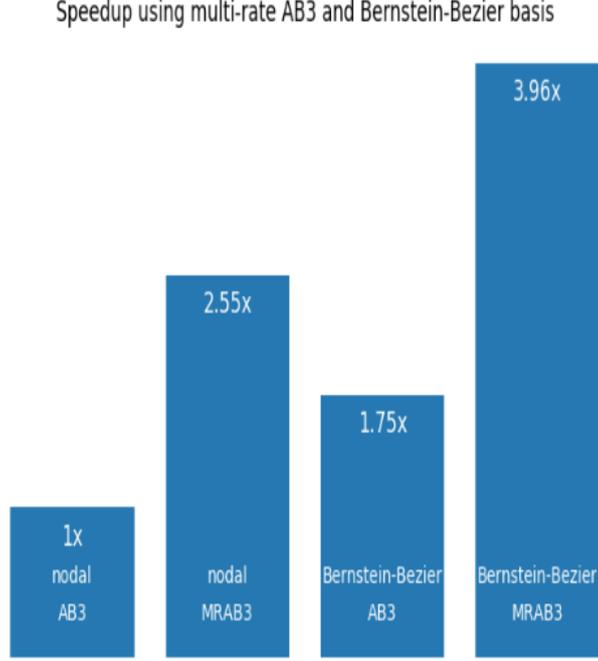


FIGURE 5.9. Performance improvement of the isotropic elastic Time DGFEM implementation using Bernstein-Bézier basis functions and Multirate AB3 (using meshes and parameters specified in Figures 11-15).

## 6. Hybridizable Discontinuous Galerkin method, HDG

When applied to steady-state and time-harmonic problems, the main drawback of classical DG methods is their computational cost (CPU time and memory occupancy) as compared to classical (continuous) finite element (CG) methods because they incur additional degrees of freedom, especially when an arbitrarily high order interpolation of the field components is used. This is due to the fact that the degrees of freedom are local to an element and so, the degrees of freedom placed at element interfaces have to be duplicated. As a consequence, DG methods lead to larger (sparse) linear systems of equations with a higher number of globally coupled degrees of freedom as compared to CG methods on a given mesh. To get around this drawback, we consider here an alternative DG method: the hybridizable DG method (HDG). Instead of solving a linear system involving the degrees of freedom of all volumic cells of the mesh, the principle of HDG consists in introducing a Lagrange multiplier representing the trace of the numerical solution on each face of the mesh. Hence, it reduces the number of unknowns of the global linear systems and the volumic solution is recovered thanks to a local computation on each element (see Fig. 6.1, Fig. 6.2 and Fig. 6.3). For more details and reference, we refer to [3].

### 6.1. Mathematical formulation

We consider the first order formulation of the 3D elastic wave equations in frequency domain. We have, for  $\mathbf{x} = (x, y, z) \in \Omega \subset \mathbb{R}^3$

$$\begin{cases} i\omega\rho(\mathbf{x})\mathbf{v}(\mathbf{x}) = \nabla \cdot \underline{\underline{\sigma}}(\mathbf{x}) + f(\mathbf{x}) & \text{in } \Omega, \\ i\omega\underline{\underline{\sigma}}(\mathbf{x}) = \underline{\underline{C}}(\mathbf{x}) \underline{\underline{\epsilon}}(\mathbf{v}(\mathbf{x})) & \text{in } \Omega, \end{cases} \quad (6.1)$$

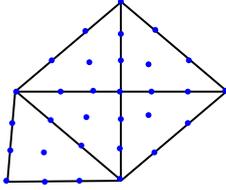


FIGURE 6.1. Distribution of the global degrees of freedom (dof) for the FEM with an interpolation order of 3.

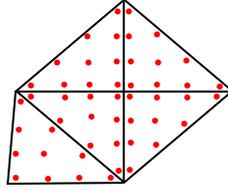


FIGURE 6.2. Distribution of the global dof for the DG method with an interpolation order of 3.

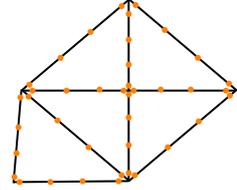


FIGURE 6.3. Distribution of the global dof for the HDG method with an interpolation order of 3.

where  $i$  is the imaginary unit,  $\omega$  the angular frequency;  $\rho(\mathbf{x})$  defines the mass density and  $f(\mathbf{x})$  the source term which is generally associated to volumic forces. The vector  $\mathbf{v}(\mathbf{x}) = (v_x(\mathbf{x}), v_y(\mathbf{x}), v_z(\mathbf{x}))^T$  is the velocity vector and  $\underline{\underline{\epsilon}}$  the strain tensor, where  $\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial v_i}{\partial j} + \frac{\partial v_j}{\partial i} \right)$ ,  $i, j = x, y, z$  and  $\underline{\underline{\sigma}}$  is the stress tensor. In the general case,  $\sigma_{ij} = \sum_{k=\{x,y,z\}} \sum_{l=\{x,y,z\}} C_{ijkl} \epsilon_{kl}$ . The tensor  $\underline{\underline{C}}$  is a fourth order symmetric tensor containing the elastic coefficients. It possesses symmetry properties,  $C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij}$  and it is positive for all symmetric tensors  $\xi : \sum_{i,j,k,l=\{x,y,z\}} C_{ijkl} \xi_{ij} \xi_{kl} \geq \alpha \sum_{i,j=\{x,y,z\}} \xi_{ij}^2$ . For simplicity, we do not describe the boundary conditions.

The main difference between Discontinuous Galerkin Methods and classical Finite Element Method lies in the fact that the basis functions are only assumed piecewise continuous. Assuming that physical parameters are piecewise constant (or polynomial) we approximate  $\mathbf{v}$  and  $\underline{\underline{\sigma}}$  on each element  $K$  of a triangulation (or mesh)  $\mathcal{T}_h$  of  $\Omega$  by polynomial functions.

To obtain the HDG formulation of the elastic waves equations, we write the weak variational formulation of system (6.1) into a linear system

$$\mathbb{A}^K \underline{W}^K + \mathbb{C}^K \underline{\Lambda} = F^K, \quad (6.2)$$

where  $\underline{W}^K$  is a vector containing the unknowns associated to degrees of freedom of  $\mathbf{v}$  and  $\underline{\underline{\sigma}}$  belonging to the element  $K$ , while  $\underline{\Lambda}$  is a vector containing all the unknowns associated to the Lagrange multipliers which represent the numerical trace of the solution. In our case,  $\Lambda$  represents the numerical trace of  $\mathbf{v}$  on each face. As the matrix  $\mathbb{A}^K$  is invertible, we express the initial unknowns of the problem,  $(\mathbf{v}, \underline{\underline{\sigma}})$ , as functions of the Lagrange multiplier  $\Lambda$

$$\underline{W}^K = (\mathbb{A}^K)^{-1} (F^K - \mathbb{C}^K \underline{\Lambda}).$$

To obtain the global linear system, we need to introduce a second equation. Thanks to the physics properties, this last equation is easily deduced by enforcing the continuity of the normal stress over each face of the mesh. It can be written as a linear system

$$\mathbb{B} \underline{W} + \mathbb{L} \underline{\Lambda} = 0, \quad (6.3)$$

Replacing  $\underline{W}$  in this last equation, we obtain the global linear system in  $\Lambda$  that we have to solve

$$\mathbb{A}^{HDG} \underline{\Lambda} = F^{HDG}. \quad (6.4)$$

Once this system is solved, the unknowns  $\mathbf{v}_h$  and  $\underline{\underline{\sigma}}_h$  can be computed locally, element by element, by solving (6.2).

It is worth noting that, except the solution of the global linear system, all the steps of the HDG algorithm are easily parallelizable.

## 6.2. Numerical Results

First, we have compared the performances of the 2D HDG method with those of classical 2D nodal DG methods like the Internal Penalty Discontinuous Galerkin (IPDG) method [17] on the Marmousi test-case. Using a direct solver (Mumps), we have shown that the memory consumption for the solution of the linear system, which is the main bottleneck of harmonic problems, is, for example, divided by 3 using HDG methods (see Fig. 6.4), and that the computational time of the linear solver is divided by 4 for a same interpolation order  $p = 3$  (see Fig. 6.5).

Then, we have considered the 3D case and we have compared the HDG computational performances with the ones of a classical finite elements (FE) method for a same number of degrees of freedom (dof). We have shown that, by increasing the number of dof, the resolution of the HDG linear system requires less memory than the resolution of the FE system. Moreover, we have also been able to reduce the memory consumption by proving that the HDG formulation for the elastic waves equations is a symmetric formulation. The HDG method for seismic imaging has been implemented as a proof of concept and is now being implemented into the FWI algorithm.

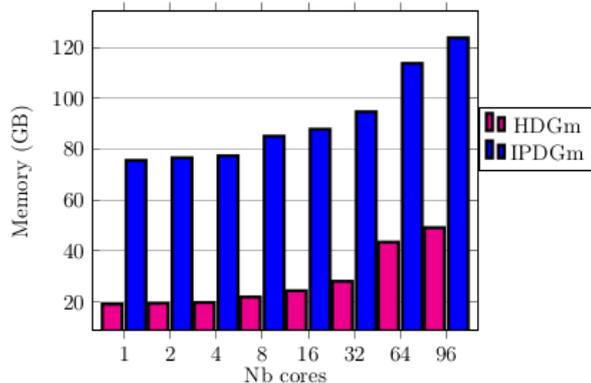


FIGURE 6.4. Memory consumption of classical DG and hybridizable DG.

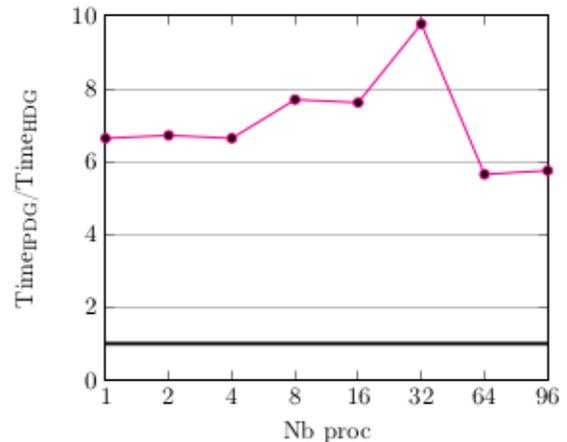


FIGURE 6.5. Speedup of hybridizable DG vs. classical DG.

## 7. Conclusion and perspectives

Efficiency for solving the forward modeling is crucial for seismic depth imaging and still require an important R&D effort. Discontinuous Galerkin (DG) methods have been shown to be well suited for solving problem in highly complex media. However, the increase of the number of degrees of freedom can limit the interest for real application. Taking advantage of recent work, we have shown that the use of appropriate basis function such as Bernstein-Bézier basis coupled with multirate time stepping greatly helps to improve performances. In harmonic domain, Hybridizable Discontinuous Galerkin (HDG) methods reduce the memory footprint helping to take advantage of the most advance direct linear solvers. HP refinement is an important feature we can take advantage of for reducing the CPU and memory cost of the DG. It relies on efficient meshes tools providing optimal meshes based on local material and geometrical properties of the media. The goal is to provide optimal meshes and polynomial

order approximation. This topic is still an important ongoing research topic and has not been addressed in this review. Another important topic we are exploring in order to take advantage of the DG flexibility is the coupling of different meshes, hexahedron- tetrahedron and numerical approximation SEM-DG. By combining the advantage of the different methods we expect to develop very efficient Wave Equation solvers for seismic depth imaging.

## Acknowledgements

C. Gout and Z. Lambert thank Albert Cohen and the organizing committee of the ninth International conference Curves and Surfaces, they also thank Avenia pole, Labex AMIES and MSO net for their support to organize this mini symposium. C. Gout and Z. Lambert thank M2NUM project (M2NUM is co-financed by the European Union with the European regional development fund (ERDF, HN0002137) and by the Normandie Regional Council), they also thank M2SiNum project (Région Normandie and European Union).

## References

- [1] H. Barucq, T. Chaumont Frelet, and C. Gout. Stability analysis of heterogeneous Helmholtz problems and finite element solution based on propagation media approximation. *Math. Comput.*, 86(307):2129–2157, 2017.
- [2] E. Baysal, D. D. Kosloff, and W. C. Sherwood. Reverse time migration. *Geophysics*, 48(11):1514–1524, 1983.
- [3] M. Bonnasse-Gahot. *Simulation de la propagation d’ondes élastiques en domaine fréquentiel par des méthodes Galerkin discontinues*. PhD thesis, Université de Nice - Sophia Antipolis - UFR Science, École Doctorale des Sciences Fondamentales et Appliquées - ED 364, 2015.
- [4] J. Chan and T. Warburton. Gpu-Accelerated Bernstein-Bézier Discontinuous Galerkin Methods for Wave Problems. *SIAM J. Sci. Comput.*, 39(2):628–654, 2017.
- [5] J. F. Claerbout. Toward a unified theory of reflector mapping. *Geophysics*, 36(3):467–481, 1971.
- [6] M. Dablain. The application of high-order differencing to the scalar wave equation. *Geophysics*, 51(1):54–66, 1986.
- [7] M. Dumbser and M. Käser. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - II The three-dimensional isotropic case. *Geophys. J. Int.*, 167:319–336, 2006.
- [8] V. Etienne, E. Chaljub, J. Virieux, and N. Glinsky. An hp-adaptive discontinuous Galerkin finite-element method for 3-D elastic wave modelling. *Geophys. J. Int.*, 183(2):941–962, 2010.
- [9] C. Gout, Z. Lambert, and D. Apprato. *Data approximation : mathematical modelling and numerical simulations*. EDP Sciences; INSA Rouen Normandie, 2019.
- [10] R. W. Graves. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences. *Bull. Seism. Soc. Am.*, 86(4):1091–1106, 1996.
- [11] J. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer, 2008.
- [12] M. Käser and M. Dumbser. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - I. The two-dimensional isotropic case with external source terms. *Geophys. J. Int.*, 166(2):855–877, 2006.
- [13] K. R. Kelly, R. W. Ward, Sven Treitel, and R. M. Alford. Synthetic seismograms: A finite-difference approach. *Geophysics*, 41:2–27, 1976.

- [14] A. Klöckner. *High-Performance High-Order Simulation of Wave and Plasma Phenomena*. PhD thesis, Dipl.-Math. techn., Universität Karlsruhe (TH), Karlsruhe, Germany, 2005.
- [15] D. Komatitsch and J. Tromp. Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophys. J. Int.*, 139(3):806–822, 1999.
- [16] K. Marfurt. Accuracy of finite-difference and finite-elements modeling of the scalar and elastic wave equation. *Geophysics*, 49:533–549, 1984.
- [17] B. Riviere. *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation*. Society for Industrial and Applied Mathematics, 2008.
- [18] J. Virieux. P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, 51:889–901, 1986.
- [19] J. Virieux, V. Etienne, and V. Cruz-Atienza. Modelling Seismic Wave Propagation for Geophysical Imaging. In *Seismic Waves – Research and Analysis*, pages 253–304. IntechOpen, 2012.
- [20] J. Virieux and S. Operto. An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6):1–26, 2009.