

# Massively parallel hybrid algorithm for simulation of seismic wave propagation in complex media\*

Victor Kostin<sup>1</sup>, Vadim Lisitsa<sup>2,3</sup>, Galina Reshetova<sup>4</sup>, Vladimir Tcheverda<sup>2</sup>, Dmitry Vishnevsky<sup>2</sup>

Intel, Novosibirsk, Russia,<sup>1</sup>

Institute of Petroleum Geology and Geophysics SB RAS, Novosibirsk, Russia,<sup>2</sup>

Novosibirsk State University, Novosibirsk, Russia,<sup>3</sup>

Institute of Computational Mathematics and Mathematical Geophysics SB RAS,  
Novosibirsk, Russia<sup>4</sup>

This paper presents a problem-oriented approach, designed for numerical simulation of seismic wave propagation in models containing geological formations with complex properties such as anisotropy, attenuation, and small-scale heterogeneities. Each of the named property requires special treatment which increases computational complexity of the algorithm in comparison with ideally elastic isotropic media. At the same time these formations are typically relatively small taking up to 25% of the model, thus the computationally expensive approaches can be used only locally to speed-up the simulations. In this paper we discuss both mathematical and numerical aspects of hybrid algorithm paying the main attention to its parallel implementation.

## 1. Introduction

Nowadays numerical simulation of seismic wave propagation in isotropic ideally elastic media is a common tool used as a part of modern algorithms for seismic imaging, full waveform inversion and others. Typically such algorithms are based on the finite differences as they combine high efficiency with suitable accuracy [1], and in particular, standard staggered grid scheme (SSGS) [2], [3] should be mentioned. As it follows from the series of studies, for example [4], this approach is preferred for simulation of wave propagation in isotropic ideally elastic media due to low demand of random access memory (RAM), small number of floating point operations (flops), and low numerical dispersion.

However, there are numbers of geological objects where the wave propagation can not be described (up to suitable accuracy) by the simplest ideally elastic isotropic model. These include anisotropy caused by finely-layered formations [5], fractured rocks [6], formations under initial stresses; seismic attenuation typical for low-consolidated formations or for fluid-saturated rocks. Additional attention should also be paid to multi-scale media and, in particular, fractured reservoirs where the fracture width is much smaller than the wavelength. All these complexities of the model lead to increase of the computational intensity of numerical algorithms. If anisotropy presents in a model use of the SSGS becomes troublesome as the numerical dispersion increases dramatically [7]. As the result two finite-difference scheme are used nowadays to resolve the problem: rotated staggered grid scheme [8] and the Lebedev scheme (LS) [9], [10]. As it is shown in [11] and [12] the LS is preferred due to lower demand on RAM. However, if compared with the SSGS the LS needs four times more RAM and flops to simulate wave propagation. If viscoelasticity is introduced in the model the number of variables and equations doubles in comparison with ideally elastic model [13] which leads to the increase of the computational work. Presence of the small-scale heterogeneities requires use of fine enough grid to match them,

---

\*This research was supported by RFBR grants no. 13-05-00076, 13-05-12051, 14-05-93090, 14-05-00049, 15-35-20022, 15-05-01310. Simulations were performed on "Lomonosov" supercomputer of MSU and on the supercomputer NKS-30T of Siberian Supercomputer Center

about 100 points per wavelength (ppw). Simple estimations (size of the domain  $100^3$  wavelength 12 variables per grid point) show that the amount of RAM, needed for simulation is 100 Tb.

At the same time the formations with the named peculiarities are relatively small taking about 25% of a model. Moreover, anisotropic and viscoelastic formations are typically sub-horizontal layers, while the small scale heterogeneities are concentrated in some volume, which can be embedded in a parallelepiped. Thus it is reasonable to apply the computationally intense algorithms only in the small enough subdomains where the particular properties of the media present. This leads to so-called coupling problems of simultaneous use of different numerical techniques. Moreover, if small scale heterogeneities and fine grid are considered the problem is also known as local mesh refinement. Taking into account the the use of explicit in time finite difference scheme time stepping should be refined as well as spatial one, thus one faces with local time-space mesh refinement problem. Several approaches to resolve this problem for wave equation have been reported, however majority of these techniques have not been implemented in 3D due to several reasons: high complexity of the algorithm, unacceptably high numerical errors, complexity of the parallel implementation. An overview of the modern techniques can be found in [14].

In this paper we present the mathematical, numerical, and computational aspects of the hybrid algorithm for simulation of wave propagation in models containing formations of complex properties, so that the computationally intense approaches are applied only in the subdomains embedding named formations while efficient standard staggered grid scheme is used elsewhere. In particular three situations are considered: coupling of the finite difference approximations of viscoelastic and ideally elastic models; coupling of Lebedev scheme and SSGS to take into account anisotropic properties of the model; local time-space mesh refinement to deal with small scale heterogeneities. Each of the situations is considered in the corresponding section of the paper.

## 2. Seismic attenuation

### 2.1. Mathematical formulation

First of all let us remind the generalized standard linear solid (GSLs) [15] model governing seismic wave propagation in viscoelastic media and in particular  $\tau$ -method [13]:

$$\rho \frac{\partial u}{\partial t} = \nabla \cdot \sigma, \quad \frac{\partial \varepsilon}{\partial t} = (\nabla u + \nabla u^T), \quad \frac{\partial \sigma}{\partial t} = C_1 \varepsilon + \sum_{l=1}^L r^l, \quad \tau_{\sigma,l} \frac{\partial r^l}{\partial t} = -C_2 \varepsilon - r^l, \quad (1)$$

where  $\rho$  is a mass density;  $C_1$  and  $C_2$  are fourth-order tensors, defining the model properties;  $u$  is a velocity vector;  $\sigma$  and  $\varepsilon$  are the stress and strain tensors respectively;  $r^l$  are the tensor of memory variables. Note that the number of memory variables tensors is  $L$ , typically two or three. Proper initial and boundary conditions are assumed.

For the ideally elastic models tensor  $C_2$  equals to zero which means that the solution of the last equation is trivial if zero initial conditions are imposed. Thus the memory variables tensors can be completely excluded from the equations. After that the system turns into that for ideally elastic wave equation. This means that there is no need to allocate random access memory (RAM) for the memory variables in the ideally elastic parts of the model.

Assume now a subdomain  $\Omega \subseteq \mathbf{R}^3$  where full viscoelastic wave equation is stated, while ideally elastic wave equation is valid for the rest of the space. It is easy to prove that the conditions at the interface  $\Gamma = \partial\Omega$  are

$$[\sigma \cdot \vec{n}]|_{\Gamma} = 0, \quad [u]|_{\Gamma} = 0, \quad (2)$$

where  $\vec{n}$  is vector of outward normal and  $[]$  denote jumps of the function over the interface. These conditions are the same as those for elastic wave equation at an interface. Moreover if

standard staggered grid scheme (SSGS) [2] is used these conditions are satisfied automatically [16], [11]. Thus the coupling of the models does not require any special treatment of conditions at the interface and can be implemented by allocation of RAM for memory variables and solving equations for them in viscoelastic part of the model.

## 2.2. Parallel implementation

Parallel implementation of the algorithm is done via domain decomposition technique. However, it has the following features:

- amount of RAM and flops per grid cell is different for elastic and viscoelastic parts of the model thus individual domain decomposition is needed;
- operators used to update solution at the interface are local in all spatial directions, thus one to one correspondence exist between two adjusted subdomains of different type (elastic and viscoelastic);
- performing one time step includes two different types of synchronization points.

The first statement is evident and do not need further explanations. The second point, means that it is natural to change the size of the subdomains only in one direction - normal to the interface, typically this is the vertical direction, due to the structure of geological models; i.e. variation of parameters in vertical direction is much stronger than in horizontal ones. After that no special treatment at the interface is needed, one just have to allocate RAM for memory variables and solve corresponding equations at elementary subdomains where viscoelastic wave equation is solved. The last statement affects the efficiency of the algorithm the most. The first type of synchronization points is at instants just after the velocity components have been updated. This stage requires the same amount of flops per grid cell for both elastic and viscoelastic parts of the algorithm. The second type is after the stresses were computed. This part is strongly different for the two models. This means that regardless to the ratio of the elementary subdomains associated with single core (node) for elastic and viscoelastic parts some of the cores will have latency period. This can be seen in fig. 1.

As the result an optimal domain decomposition was suggested to minimize the overall computational time (core-hours) of the algorithm. The computational time can be estimated by the formula:

$$T(\alpha, \beta) = [\delta \max(1, \beta) + \max(1, \beta\gamma)] \left( \alpha + \frac{1 - \alpha}{\beta} \right),$$

where  $\gamma = 0.33$  and  $\delta = 0.32$  are the ratios of computational time needed to update velocity and stresses respectively for elastic part per grid cell with respect to time for computing stresses for viscoelastic part. These values were measured experimentally. Parameters  $\alpha$  is the relative volume of the viscoelastic part of the model and  $\beta$  is the ratio of the elementary volumes of elastic and viscoelastic parts. It is clear that the optimal domain decomposition is constructed if  $\min_{\beta} T$  is achieved. Figure 2 represents the theoretical estimations of  $T(\alpha, \beta)$  and numerical experiments. One may note that the optimal ratio of the elementary volumes for elastic and viscoelastic parts is equal to 3 for relatively small amount of viscoelasticity. In this case the speed-up of the hybrid algorithm is about 2 with respect to pure viscoelastic simulation.

## 3. Anisotropy

### 3.1. Finite-difference approximation

In this section we consider simulation of seismic wave propagation in ideally elastic anisotropic media. Assume system (1) with zero attenuation term; i.e.  $C_2 \equiv 0$ . If the medium is isotropic

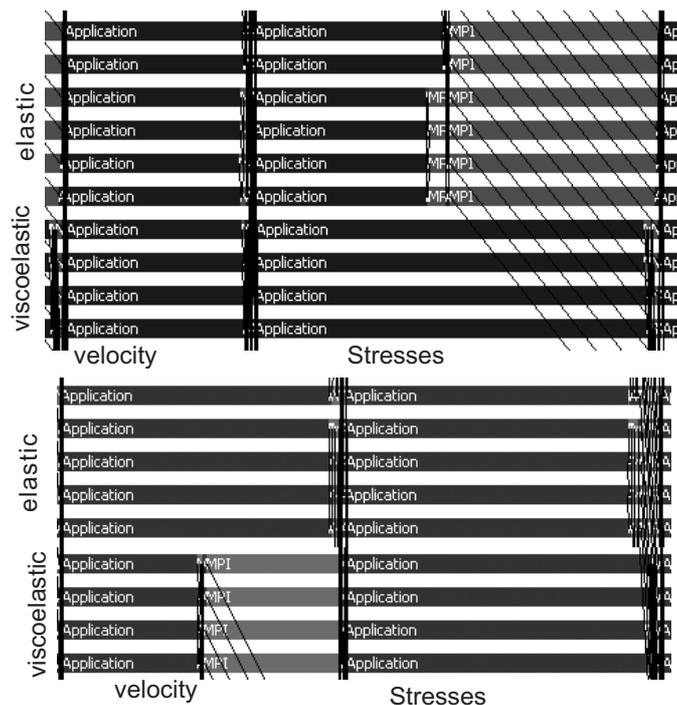


Figure 1: TraceAnalyzer images for equivolumetric (top) and optimal (bottom) domain decomposition. Dark bars correspond to computations, light bars represent waiting time.

the stiffness tensor  $C_1$  has a specific form:

$$C_1 = \begin{pmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{22} & c_{23} & 0 & 0 & 0 \\ c_{13} & c_{23} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{pmatrix}.$$

This means that the stress-strain relation, third equation in (1), decouples into four independent equations which can be solved in different spatial points. As the result one may define different components of the wavefield; i.e. velocity vector and stress tensor components, in different grid points. After that standard central differences can be used to approximate system (1). As the result one needs to store one copy of each variable per grid cell, which is 9 in total. This scheme is known as standard staggered grid scheme (SSGS) [2].

If the medium in anisotropic stiffness tensor  $C_1$  has no special structure except for the symmetry; i.e.  $c_{ij} = c_{ji}$ . As the result no separation of variables can be applied and all components of the stress tensor should be defined in the same grid points. Similarly all components of velocity vector should be stored in each grid point, different from those for stresses. There are two schemes oriented on the anisotropic wave simulation: rotated staggered grid scheme [8] and Lebedev scheme (LS) [9]. Detailed comparison of these approaches is presented in [11], where the authors proved that implementation of the LS is preferred due to lower demand on computational resources. However, if compared with the SSGS, Lebedev scheme requires four times more RAM to store the wavefield variables; i.e. four copies of each component per grid

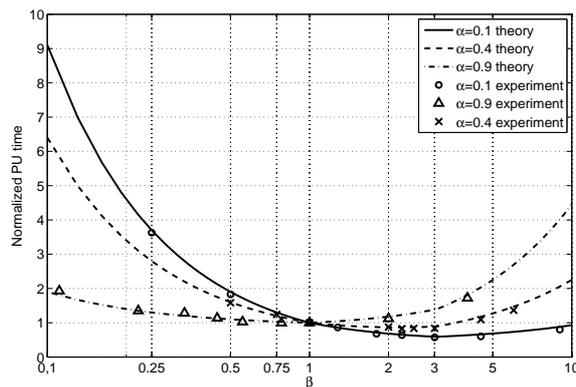


Figure 2: Normalized core-hours with respect to  $\beta$ .

cell, see fig. 3 for the details.

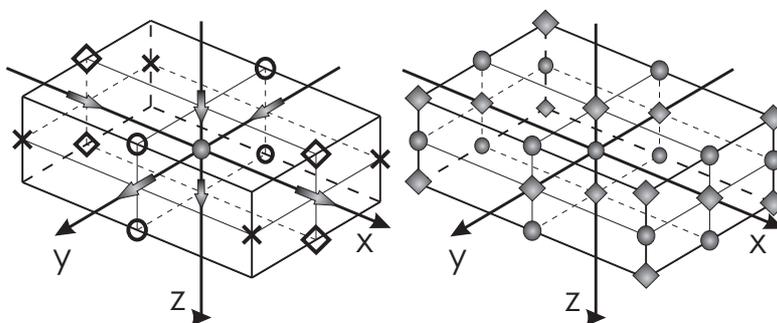


Figure 3: Left - grid cell for the standard staggered grid. Filled circle correspond to point where  $\sigma_{xx}$ ,  $\sigma_{yy}$  and  $\sigma_{zz}$  are defined,  $\sigma_{yz}$  is stored at empty circles,  $\sigma_{xz}$  is at rhombi,  $\sigma_{xy}$  is at crosses. Velocity components are defined at arrows pointing to corresponding directions. Right - grid cell for the Lebedev scheme. All components of the stress tensor are defined at circles, all velocity components are stored at rhombi.

### 3.2. Coupling

Assume a subdomain  $\Omega \subseteq \mathbf{R}^3$  where anisotropic wave equation is stated, while isotropic wave equation is valid for the rest of the space. If differential equations are considered the conjugation conditions at the interface  $\Gamma = \partial\Omega$  are standard and provided by formulae (2). However, if finite-difference approximation is applied number of variables for the LS is four times bigger than that for the SSGS. This gives additional degrees of freedom and leads to the existence of spurious modes for the LS. As it was shown in [17], [18] the LS approximates a system of equations which is four times wider than anisotropic wave equation (1) and possesses 12 plane-wave solutions instead of three. As the result the wavefield, computed by the LS can be decoupled into four parts. Let us denote them by  $u^{++}$ ,  $u^{+-}$ ,  $u^{-+}$ , and  $u^{--}$ , where  $u^{++}$  is the velocity vector corresponding to the true solution, while the others are artifacts. Similar representation take place for stress tensor. Note, that this decomposition is valid at each grid point. Thus the conjugation conditions at the interface  $\Gamma$  are:

- for the true components

$$u^V = u^{++}|_{\Gamma}, \quad \sigma^V \vec{n} = \sigma^{++} \vec{n}|_{\Gamma},$$

- for each artificial mode any three conditions of the set

$$u^{\pm\pm}|_{\Gamma} = 0, \quad \sigma^{\pm\pm}\vec{n}|_{\Gamma} = 0.$$

Here  $\vec{n}$  is the vector of external normal to  $\Omega$  at interface  $\Gamma$ . From the physical point of view these conditions permit the true solution to pass the interface with no reflections but keep all the artifacts inside the domain  $\Omega$  without transformation to any other types of waves.

Finite-difference approximation of the presented conditions leads to the necessity of wavefield interpolation with respect to tangential directions. This interpolation has to be done at each time step. Our suggestion is to use the fast Fourier transform (FFT) based interpolation, as it possesses exponential convergence. Moreover, typical discretization for the finite difference simulation of wave propagation is about 10 points per wavelength, which means that about 30% in 1D (90 % in 2D) of the FFT spectra of the solution corresponding to high frequencies are trivial. This property will be used for the parallel implementation of the algorithm.

So, the formulae to update solution at the interface  $\Gamma$ , where the SSGS and the LS are coupled, are local in normal direction, but nonlocal in tangential directions, requiring use of the whole wavefield in the vicinity of the interface. On the other hand, if FFT based interpolation is used only 10% of the spectrum may be used to reduce the amount of data to be transferred.

### 3.3. Parallel implementation

Let us point out the main features of the algorithm:

- amount of RAM and flops for anisotropic simulation is four times higher than that for solving isotropic problem;
- increase of computational work to update velocity and stresses is equal;
- formulae to update solution at the interface are local in normal direction and nonlocal in tangential ones;
- FFT base interpolation is used, thus the transferred data can be compressed.

Here we need to point out that the anisotropic formations are typically sub-horizontal layers elongated throughout the model. Thus it is reasonable to define the domain  $\Omega$ , where the anisotropic elastic wave equation is solved, as a horizontal layer embedding the formation. As the result one gets two interfaces (for each anisotropic domain), they are  $z = z_+$  and  $z = z_-$ . Within each subdomain an independent 3D domain decomposition is applied to ensure high balancing level. Taking into account the first two peculiarities of the algorithm the optimal ratio of isotropic to anisotropic elementary subdomains is four. To organize the data transfer from isotropic to anisotropic domains master processors are allocated in each groups at both sides of the interfaces. These processors get the wavefield which needed to be interpolated from the corresponding side of the interface. After that the 2D FFT is applied by the master processors and the significative part of the spectra is sent to the master processor from the other side of the interface, where the interpolation and inverse FFT is performed. Finally the interpolated data is sent to the processors which need these data. We understand that this part is a bottle-neck of the algorithm, however this is a reasonable price for nonlocal data transfer. Moreover, use of nonblocking procedures `Isend/Irecv` allow performing communications during updating the solution in the interior of the subdomains, thus the strong scaling of the algorithm is close to 93% for the use of up 4000 cores, see fig. 4.

### 3.4. Numerical experiment

In order to test the algorithm a simulation of wave propagation in anisotropic Gullfaks 2.5D model was performed. Original model was 2D but the third direction was added so that the

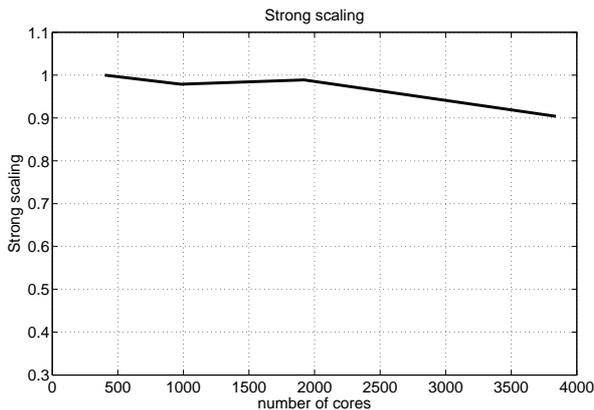


Figure 4: Strong scaling of the hybrid isotropic-anisotropic algorithm.

parameters were constant with respect to this direction. Meanwhile the point source was used, thus the simulation was 3D. The model, P-wave velocity and Thomsen's parameter  $\epsilon$ , is presented in fig. 5 and 5. Note that if  $\epsilon = 0$  the medium is isotropic, thus there is one anisotropic layer in this model. Wave propagation in anisotropic Gullfaks model was simulated by the hybrid algorithm (fig. 6) and that based on the LS after that the difference was considered (fig. 6). One may note that the error caused by the use of the coupled scheme is about 0.2% which is acceptable level for seismic simulations. At the same time the speed-up of the hybrid algorithm was 2.5 times.

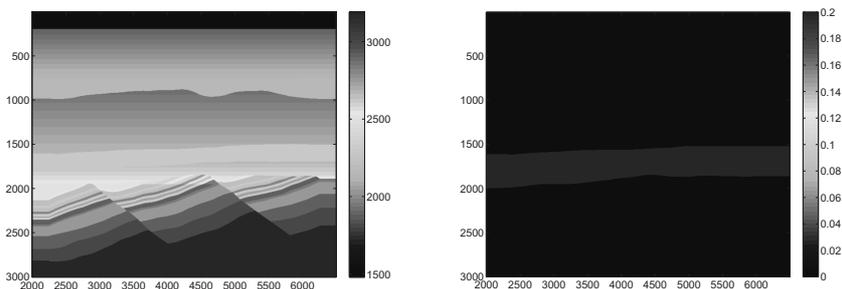


Figure 5: P-wave velocity (left) and Thomsen's parameter  $\epsilon$  (right) for the Gullfaks model.

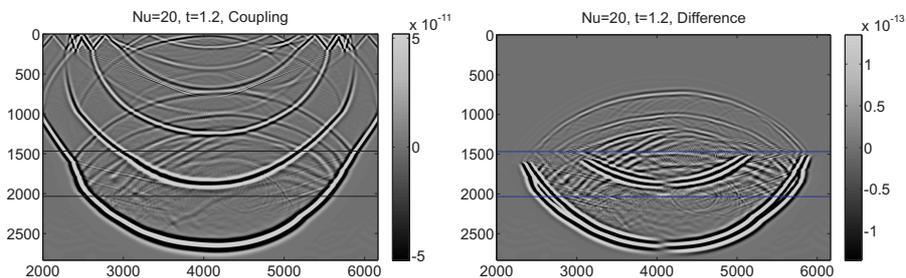


Figure 6: Wavefield computed by the hybrid algorithm (left) and the difference in the wavefields computed by hybrid algorithm and that based on the LS (right).

## 4. Multi-scale simulation

### 4.1. Mathematical formulation

Before proceeding to the mathematical formulation of the algorithm let us point out that it was designed for simulation of seismic wave propagation in a specific class of geophysical models and to study particular physical effects. So, the main object which is under consideration is the cluster of small-scale inhomogeneities such as fractured/cavernous reservoirs. In this situation a detailed description of the of each fracture is not reasonable but their distribution within some relatively small volume is under the interest. So, we assume the reservoir model is provided on some fine enough grid - grid steps are of the size of several centimeters. Typical seismic wave has the wavelength of about several dozens of meters, which assumes the grid steps in the background model of about several meters. Thus a local mesh refinement is needed to perform the full waveform simulation of long wave propagation through a reservoir with fine structure. As the explicit finite differences are used the size time step strongly depends on the spatial discretization thus the time stepping should also be local. As the result the problem of simulation of seismic wave propagation in models containing small-scale structures becomes the mathematical problem of local time-space mesh refinement.

First of all, let us explain how a coarse and a fine grids are coupled to each other. The necessary properties of the finite difference method based on a local grid refinement should be its stability and an acceptable level of artificial reflections. Scattered waves we are interested in have an amplitude of about 1% of the incident wave. Artifacts should be at least 10 times less, that is about 0.1% of the incident wave. If we refine the grid at once in time and space stability of the FDS on this way can be provided via coupling coarse and fine grids on the base of energy conservation, which leads to an unacceptable level (more than 1%) of artificial reflections. We modify the approach so that the grid is refined by turn in time and space on two different surfaces surrounding the target area with microstructure. This allows decoupling temporal and spatial grid refinement and to implement them independently and to provide the desired level of artifacts.

#### 4.1.1. Refinement in time.

Refinement in time with a fixed 1D spatial discretization is clearly seen in fig.7 and does not need any explanations. Its modification for 2D and 3D media is straightforward.

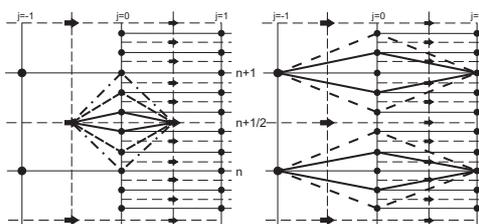


Figure 7: Embedded stencils for local time stem mesh

#### 4.1.2. Refinement in space.

In order to change spatial grids, the Fast Fourier Transform (FFT) based interpolation is used. Let us explain this procedure for a 2D problem. The mutual disposition of a coarse and a fine spatial grids is presented in Fig.8, which corresponds to updating the stresses by velocities (updating stresses by velocities is implemented in the same manner). As can be seen, to update solution at the fine grid at the upmost line it is necessary to know the wavefield at the points marked with black label, which do not exist on the given coarse grid. Using the fact that all of

them are on the same line (on the same plane for 3D statement), we seek the values of missing nodes by FFT based interpolation. Its main advantages are an extremely high performance and exponential accuracy. This accuracy allows us to provide the required low level of artifacts (about 0.001 with respect to the incident wave) generated on the interface of these two grids. For 3D statement we again perform the FFT based interpolation but this time 2D.

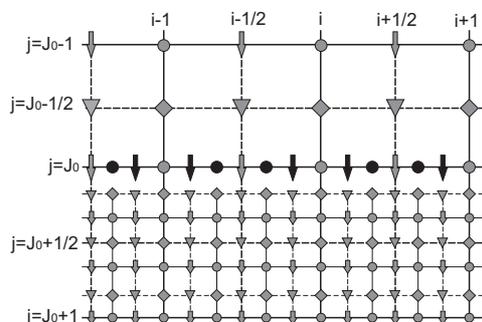


Figure 8: Spatial steps refinement. Black marks correspond to points where the coarse-grid solution is interpolated.

## 4.2. Implementation of parallel computations.

The simultaneous use of a coarse and a fine grids and the need for interaction between them makes it difficult to ensure a uniform load of Processor Units under parallelization of computations based on Domain Decomposition. Besides, the user should be allowed to locate the reservoir anywhere in the background.

This problem is resolved through the implementation of parallel computations on two groups of Processor Units. One of them is fully placed 3D heterogeneous referent environment on a coarse grid, while the fine mesh describing the reservoir is distributed among the PU in the second group (Fig.9). Thus, there is a need for both exchanges between processors within each group and between the groups as well. The data exchange within a group is done via faces of the adjacent Processor Units by non-blocking iSend/iReceive MPI procedures. Interaction between the groups is much more complicated. It is carried out not so much for data sending/receiving only, but for coupling a coarse and a fine grids as well. Let us consider the data exchange from the first group (a coarse grid) of PU to the second (a fine grid) and backwards.

### 4.2.1. From coarse to fine.

The Processor Units are found in the first group which cover the target area, and are grouped along each of the faces being in contact with the fine grid. At each of the faces there is allocated the Master Processor (MP), which gathers the computed current values of stresses/displacements and sends them to the relevant MP on a fine grid (see fig.9). All the subsequent data processing providing the coupling of a coarse and a fine grids by the FFT based interpolation is performed by the relevant Master Processor in the second group (a fine grid). Later this MP sends interpolated data to each processor in its subgroup.

Interpolation performed by the MP of the second group essentially decreases the amount of sent/received data and, hence, the idle time of PU.

### 4.2.2. From fine to coarse.

As in the previous case, primarily there are identified PU from the second group which perform computations on the faces covering the target area. Next, again for each face Master Processor is identified. This MP as its partner from the coarse grid collects data from the

relevant face and performs their preprocessing before sending to the first group of PU (a coarse grid). Now we do not need all data in order to move to the next time, but only those of them which fit the coarse grid. Formally, these data could be thinned out, but our experiments have proved that this way generates significant artifacts due to the loss of smoothness. Therefore for this direction (from fine to coarse) we also use the FFT based interpolation implemented by the relevant MP of the second group (a fine grid). The data obtained are sent to the first group.

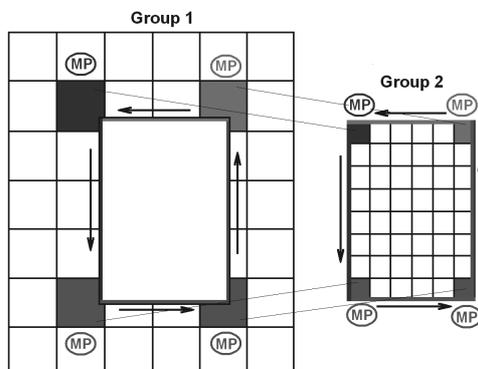


Figure 9: Processor Units for a coarse (left) and a fine (right) grids.

### 4.3. Numerical experiment

Designed algorithm was implemented to study wave propagation in fractured reservoir embedded in complex model of buried channel which was viscoelastic, see fig. 10. Zero off-set seismic data is provided in fig. 11, illustrating presence of scattered waves associated with fractured reservoir.

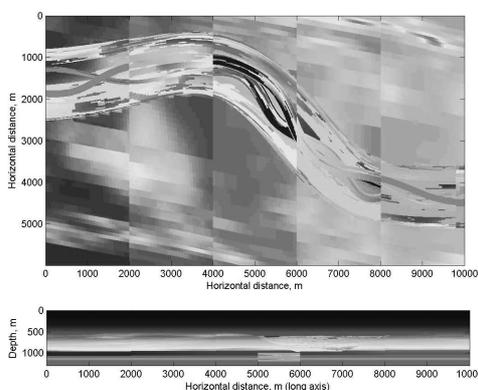


Figure 10: The buried channel model.

## 5. Conclusions

We presented an original algorithm oriented on simulation of seismic wave propagation in complex 3D media. The main idea of the approach is the combination of computationally intense approaches oriented on specific properties of the model with efficient standard staggered grid scheme which is applied in the main part of the model. Three situations were considered and studied in the details.

First we considered the coupling of the elastic and viscoelastic modeling. The main features of the algorithm are the simplicity of mathematical formulation; i.e. the memory variables can be allocated only in viscoelastic part of the model and completely omitted elsewhere. However,

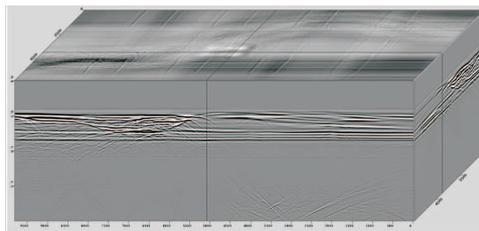


Figure 11: Zero off-set seismogram.

due to presence of two types of synchronization points of the algorithm and different ratio of the computational work of elastic and viscoelastic versions between the synchronization points the balancing of the algorithm is low than 100%. As it was proved if the ratio of elementary volumes assigned to a single PU in elastic to viscoelastic part is equal to 3 the idle time is minimal, thus the overall computational time is also optimal, even though the processor loading is not optimal. The speed-up of the coupled algorithm is up to 2 for the models containing viscoelastic formations which take less than 20% of the computational domain.

The second part of the algorithm is the coupling of the Lebedev scheme with standard staggered grid scheme. The main difficulty of this part is the organization of the data exchange between the PU computed solution in isotropic and anisotropic parts of the model. This communication was organized by means of the FFT interpolation, so that only valuable part of the spectra was the subject of exchange. The computational intensity of the LS is five times higher than that of the SSGS, and this ratio is constant either velocity or stresses are updated, thus the use of individual domain decomposition allows making the hybrid algorithm well-balanced. The speed-up of the hybrid algorithm is about 3 times in comparison with that based on the LS is the anisotropic part of the model is about 25%.

The third part of the hybrid algorithm is local time-space mesh refinement. The main feature of the presented approach is the separation of the refinement of spatial and temporal stepping which ensures stability and low reflectivity of the coupling. The time steps are refined on the base of finite difference approximation of the elastic wave equation. To refine the spatial steps the FFT based interpolation is used. The particular geometry of the grid allows applying 2D interpolation for the 3D modeling. Use of independent domain decomposition leads to a well-balanced algorithm and reduces the requirement of the computational resources up to several orders if compared with fine grid simulations.

## References

1. Virieux, J., Calandra, H., Plessix, R. A review of the spectral, pseudo-spectral, finite-difference and finite-element modelling techniques for geophysical imaging // *Geophysical Prospecting*, 2011, Vol. 59, N. 5, P. 794–813.
2. Virieux, J. P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method // *Geophysics*, 1986, Vol. 51, N. 4, P. 889–901
3. Levander, A. R. Fourth-order finite-difference P-SV seismograms // *Geophysics*, 1988, Vol. 53, N. 11, P. 1425–1436
4. Moczo, P., Kristek, J., Galis, M., Chaljub, E., Etienne, V. 3-D finite-difference, finite-element, discontinuous-Galerkin and spectral-element schemes analysed for their accuracy with respect to P-wave to S-wave speed ratio // *Geophysical Journal International*, 2011, Vol. 187, N. 3, P. 1645–1667
5. Backus, G. E. Long-wave elastic anisotropy produced by horizontal layering // *Journal of Geophysical Research*, 1962, Vol. 67, N. 11, P. 4427–4440

6. Grechka, V., Kachanov, M. Effective elasticity of rocks with closely spaced and intersecting cracks // *Geophysics*, 2006, Vol. 71, N. 3, P. D85–D91
7. Igel, H., Mora, P., Riollet, B. Anisotropic wave propagation through finite-difference grids // *Geophysics*, 1995, Vol. 60, P. 1203–1216
8. Saenger, E. H., Gold, N., Shapiro, S.A. Modeling the propagation of the elastic waves using a modified finite-difference grid // *Wave Motion*, 2000, Vol. 31, P. 77–92
9. Lebedev, V. I. Difference analogies of orthogonal decompositions of basic differential operators and some boundary value problems. I // *Soviet Comput. Math. Math. Physics.*, 1964, Vol. 4, P. 449–465
10. Asvadurov, S., Druskin, V., Moskow, S. Optimal grids for anisotropic problems // *Electronic Transactions on Numerical Analysis*, 2007, Vol. 26, P. 55–81
11. Lisitsa, V., Vishnevskiy, D. Lebedev scheme for the numerical simulation of wave propagation in 3D anisotropic elasticity // *Geophysical Prospecting*, 2010, Vol. 58, N. 4, P. 619–635
12. Bernth, H., Chapman, C. A comparison of the dispersion relations for anisotropic elastodynamic finite-difference grids // *Geophysics*, 2011, Vol. 76, N. 3, P. WA43–WA50
13. Blanch, J., Robertson, A., Symes, W. Modeling of a constant Q: Methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique // *Geophysics* 1995, Vol. 60, N. 1, P. 176–184
14. Lisitsa, V., Reshetova, G., Tcheverda, V. Finite-difference algorithm with local time-space grid refinement for simulation of waves // *Computational Geosciences*, 2012, Vol. 16, N. 1, P. 39–54
15. Christensen, R. M. *Theory of viscoelasticity, an introduction*. Academic press New York and London, 1971
16. Moczo, P., Kristek, J., Vavrycuk, V., Archuleta, R.J., Halada, L. 3D heterogeneous staggered-grid finite-difference modeling of seismic motion with volume harmonic and arithmetic averaging of elastic moduli and densities // *Bulletin of the Seismological Society of America*, 2002, Vol. 92, N. 8, P. 3042–3066
17. Lisitsa V., Vishnevsky, D. On specific features of the lebedev scheme in simulating elastic wave propagation in anisotropic media // *Numerical Analysis and Applications*, 2011, Vol. 4, N. 2, P. 125–135
18. Lisitsa, V., Tcheverda, V., Vishnevsky, D. Numerical simulation of seismic waves in models with anisotropic formations: coupling Virieux and Lebedev finite-difference schemes // *Computational Geosciences*, 2012, Vol. 16, N. 4, P. 1135–1152