Using multifrontal hierarchically solver and HPC systems for 3D Helmholtz problem[*]

Sergey Solovyev[1], Dmitry Vishnevsky[1], Hongwei Liu[2]

Institute of Petroleum Geology and Geophysics SB RAS[1]
EXPEC ARC, Geophysics Technology, Saudi Aramco[2]

We present a multi-frontal hierarchically semi-separable solver to perform forward modeling of the 3D Helmholtz acoustic problem. Our frequency-domain solver combines two efficient approaches. First, it uses an optimal 27-point finite-difference scheme to decrease numerical dispersion and reduce required discretization of the model in terms of points per wavelength from 15 to about 4. Second, it uses a supernodal multi-frontal method based on low-rank approximation and hierarchically semi-separable (HSS) structure to improve performance, decrease memory usage and make it practical for realistic size 3D models required by full-waveform inversion. We performed validation and performance testing our new solver using a 3D synthetic model. Performance and OMP scalability of the solver were compared with the Intel MKL PARDISO.

1. Introduction

Frequency-domain modelling and inversion are at the heart of many modern geophysical tools, such as full-waveform inversion of seismic data. Here we focus on developing an efficient algorithm for solving the Helmholtz problem in 3D heterogeneous media. This is a prerequisite for acoustic full-waveform inversion of marine and land seismic data. We present an algorithm based on two features: 1) a 27-point finite-difference approximation scheme to minimize the dispersion errors and 2) low-rank approximation technique and a hierarchically semi-separable (HSS) structure in the multi-frontal direct solver to decrease the computational resources required. Perfectly matched layer (PML) boundary conditions are used in this study. Employing a 27-point scheme as opposed to a simpler 7-point approach allows decreasing the required model discretization from 15 to 4 points per wavelength, while maintaining the dispersion error below 1%. The quality of low-rank compression is improved by using a cross-approximation (CA) approach [6]. We validate this algorithm with numerical tests and confirm accuracy and performance improvements. The high accuracy and memory efficiency of the proposed algorithm is demonstrated using the 3D overthrust benchmark model. The current implementation requires three times less computational resources compared to other sparse direct solvers. As a result, the proposed algorithm significantly decreases the computational resources required for solving realistic 3D acoustic problems. Furthermore, the multi-frontal solver can be efficiently parallelized on both shared (via OMP) and distributed memory systems (via MPI).

2. Optimal 27-point finite-difference scheme

The scalar-wave equation for a homogeneous isotropic medium in the frequency domain can be written as

$$\Delta u + \frac{(2\pi\nu)^2}{V^2}u = \delta(\overline{r} - \overline{r}_s)f \tag{1}$$

where ν is a temporal frequency, V is acoustic velocity, $\overline{r}_s$ denotes source coordinates, and f denotes source function. The standard finite-difference approximation on a parallelepiped grid

uses a 7-point stencil. As result we obtain a second-order approximation that exhibits dispersion errors of the solution in the frequency domain (Figure 1).
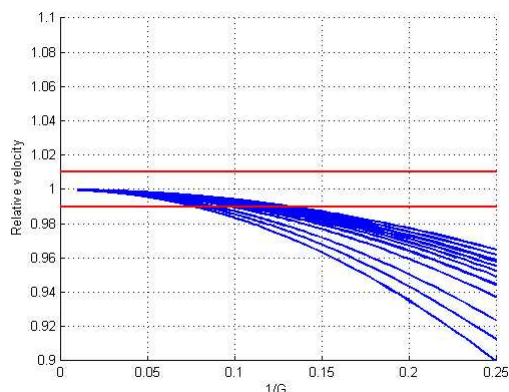

Figure 1. The dispersion curves of the 2-nd order approximation for the 7-point scheme. Numerical phase velocity V_{ph} are normalized with respect to the true velocity V and plotted versus $1/G$, where G is the number of grid points per wavelength.

Dispersion analysis confirms that at least 15 points per wavelength are required to reduce the dispersion to a practically acceptable level of less than 1%. In real-world 3D geophysical problems, this results in a huge computational grid and greatly increased computational time. There are various rotation schemes for decreasing dispersion error both in 2D [4] and in 3D [8] cases. We suggest a different approach that is easier to implement, but provides the same accuracy. To approximate a Laplace operator we use a combination of the three schemes: the first one uses middle face points, the second uses middle edges and the third uses corner points (Figure 3).
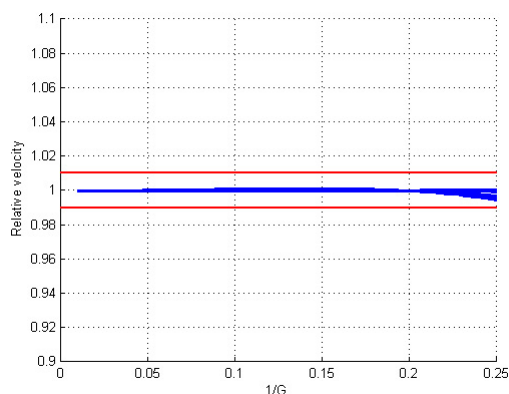

Figure 2. The dispersion curves of the 2-nd order approximation for the optimized 27-point scheme.

To approximate the wavenumber we use four schemes presented in Figure 4. As result we have the universal scheme with seven parameters:

$$\gamma_1 \Delta_1 u + \gamma_2 \Delta_2 u + \gamma_3 \Delta_3 u + \frac{(2\pi\nu)^2}{V^2}[w_1 u_{0,0,0} + \frac{w_2}{6}\sum_{St_2} u_{k_1,k_2.k_3} + \frac{w_3}{12}\sum_{St_3} u_{k_1,k_2.k_3} + \frac{w_4}{8}\sum_{St_4} u_{k_1,k_2.k_3}]$$
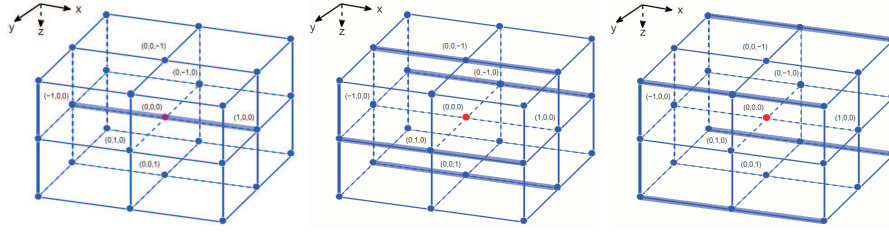
(2)

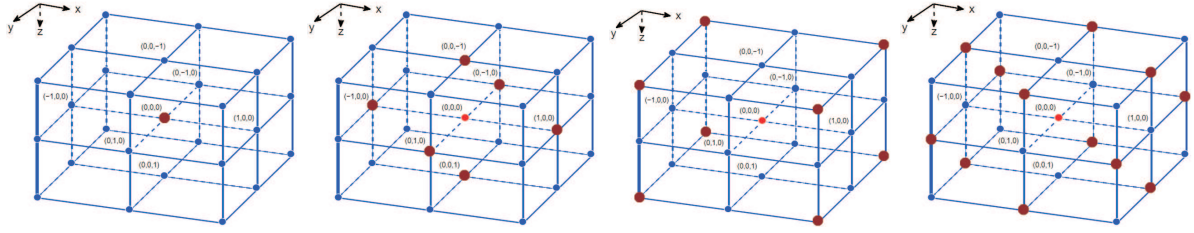**Figure 3.** Three stencils of approximation the second order derivation along X axis.

Figure 4. Four stencils for the approximation of the wavenumber.

where Δ_1, Δ_2 and Δ_3 correspond to the three stencils from Figure 3, St_2, St_3 and St_4 are the set of points corresponding to "bold" points of the stencil from Figure 4.

Additionally, coefficients should satisfy the following constraints where $\sum_{i=1}^{3} \gamma_i = 1$ and $\sum_{i=1}^{4} w_i = 1$. Dispersion analysis gives an explicit formula for phase velocity $V_{ph} = V_{ph}(\overline{\gamma w}; G, \phi, \theta)$, where $\overline{\gamma w} = (\gamma_1, \gamma_2, \gamma_3, w_1, w_2, w_3, w_4)$, G is a number of points per wavelength and (ϕ, θ) is direction of propagation in spherical coordinates. After minimization of the functional (3) (representing average error for all directions), the optimal parameters $\overline{\gamma w}$ are (0.656, 0.294, 0.050; 0.576, 0.187, 0.359, 0.122).

$$J(\overline{\gamma w}) = \iiint\limits_{G=[G_{min},G_{max}];\varphi=[0,\pi/2];\theta=[0,\pi/4]} \left(V_{ph}(\overline{\gamma w})/V - 1 \right)^2 dG d\varphi d\theta \qquad (3)$$

Analysis of numerical dispersion for this scheme shows that four points per wavelength are enough to achieve 0.5% dispersion error in wave propagation velocity.

The proposed scheme, jointly with PML boundary conditions, can lead to a complex sparse symmetric (non-Hermitian) 27-diagonal matrix. Solving such a system of linear algebraic equations (SLAE) is usually done by direct approach, using a nested dissection algorithm and low-rank/HSS technique. For real problems, the SLAE typically has many right hand sides (RHS), corresponding to the number of physical sources. The most compute-intensive stage of the solver, decomposition of the matrix, is performed once and is used for solving SLAE many times for various RHS.

3. Multifrontal low-rank solver

The direct solver is based on decomposition of the matrix A. The traditional way of improving performance of the decomposition is based on row/column reordering. This corresponds to renumbering points of the computational grid. One of the most effective methods is the called nested dissection algorithm [3], which uses a "separator" technique of the binary tree to separate the grid on the unconnected subdomains. As a result, the factorization process of the matrix
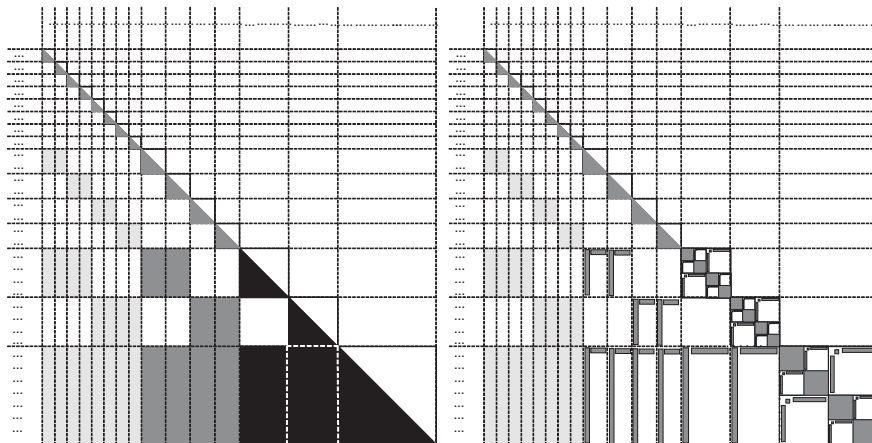
Figure 5. Nested dissection reordering: pattern of L-factor in exact arithmetic (left image) and after low-rank approximation with HSS diagonal blocks (right image).

columns corresponds to the grid nodes from different domains. This process can be performed independently. Such a process is referred to as "multi-frontal." Moreover, after preliminary reordering, the L-factor contains less non-zero elements compared to before reordering. The pattern of such an L-factor is shown in Figure 5 (left). To improve the multi-frontal process we use low-rank approximation based on the fact that large off-diagonal blocks can be efficiently approximated by low-rank matrices [1] while diagonal blocks can be effectively represented in hierarchically semi-separable (HSS) format [9]. The pattern of the L-factor is presented in Figure 5 (right). Matrix operations in the multi-frontal method are performed using low-rank arithmetic, so computational resources can be significantly decreased [5]. To compress dense blocks into low-rank or HSS structures, we use the panel modification of the cross-approximation (CA) approach [6]. The inversion step uses the iterative refinement process to achieve the accuracy, compatible with direct solvers.

4. Numerical experiments

Let us evaluate the performance and precision of our algorithm using a numerical test. Performance is measured on a single node with Intel® Xeon® E5-2690 v2 (Ivy Bridge EP) 3.00GHz processors, RAM 512 GB with 8 threads. A realistic 3D velocity model (3D overthrust benchmark) is used. The velocity varies from 2300 m/s to 6000 m/s. The target domain is with frequencies from 1 Hz to 8 Hz and a source at the corner point. The spatial discretization is 30 m in all directions and the size of the PML boundary varies from 25 grid points for 1 Hz to 10 grid points for 8 Hz. Therefore, the size of the computational grid varies from (351x351x201) to (321x321x171).

First, we compare 7-point and 27-point schemes. We observe that for 10 grid points per wavelength (test for 8 Hz), a standard 7-point scheme delivers unacceptable results (Figure 6, left) with very high numerical dispersion, whereas the 27-point scheme provides good results (Figure 6, right). The wavefield from the 27-point scheme is nearly identical to a numerical solution computed with the Seiscope time-domain solver (reference), thus serving as useful benchmark. Next let us evaluate computer resources required for both the 27-point and 7-point schemes using conventional arithmetic and low rank approximation (Table 1).

5. Parallel computations

To parallelize computations, we are going to use both MPI and OMP parallelization approaches. The idea of the MPI parallelization of the factorization step is based on elimination
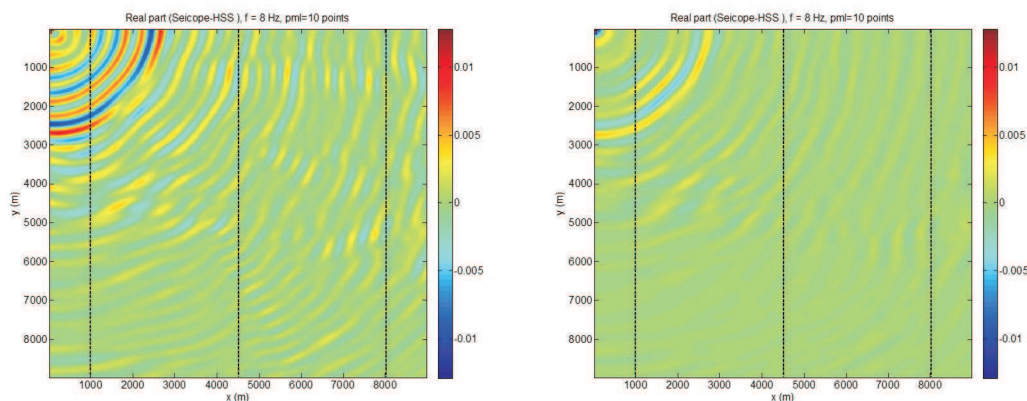
Figure 6. Real part of wavefield error at 8 Hz computed with 7-point scheme (left image) and optimized 27-point one (right image). Observe unacceptable dispersion on the left due to insufficient discretization (10 points per wavelength).

Table 1. Memory usage for storing L-factors (8Hz) and solution times.

	7-point	27-point
Exact arithmetic	848 GB	880 GB
Low-rank approximation	256 GB	272 GB
Factorization	26 339 s	28 741 s
Solution, 5 iterations	258 s	288 s

tree structure as result of 3D Nested Dissection reordering. The factorization of the different elimination tree nodes can be done in parallel on different cluster nodes [7]. Let me note that current MPI version of program is under developing now. The OMP parallelization of our solver is based on OMP functionality of BLAS and LAPACK functions from the Intel MKL. We tested the OMP scalability and compared performance of our solver with direct solver PARDISO from Intel MKL (Table 2). Test case is the result of 7-point finite difference approximation of the Helmholtz equation (12Hz) with PML layer in cube. The number of grid points is $3.4 * 10^6$ (151x151x151) and this problem can be solved by direct solver on our system with 512G RAM as well.

Table 2. Factorization and solution times (inverse LDL^t factors) for 10 right hand sides.

	PARDISO	Low-rank solver
Factorization (1 thread)	13 594 s	4 072 s.
Factorization (8 threads)	2 387 s	1 941 s.
Solution (1 thread)	66 s	28 s
Solution (8 threads)	32 s	14 s

This table shows more than 3 times performance gain of non-parallel factorization of our solver in compare with Intel MKL PARDISO. However the scalability of the PARDISO is better

than our solver, so on 8 threads performance of our solver just a 20% better than PARDISO. Scalability of solution step is the similar both for low-rank solver and PARDISO one: about twice from 1 to 8 threads. Therefore the solution step of our solver is more than twice faster than PARDISO. Let me note that parallelization of the factorization step are going to be improved by using *two-level* approach of combining parallel low-rank compressing of different blocks of L-factors with using OMP parallelization of MKL BLAS/LAPACK functions.

6. Conclusions

We developed a high-performance algorithm for numerical solution of the Helmholtz problem in 3D heterogeneous media. The main advantage of the algorithm is a reduction of number of grid points per wavelength enabled by application of optimal 27-point finite-difference scheme resulting in a significant memory saving. The second advantage is use of an efficient multi-frontal direct solver, based on low-rank approximation technique and hierarchically semi-separable (HSS) structure. The high accuracy of the proposed scheme and memory efficiency is shown using a realistic overthrust benchmark model, which cannot be solved by direct solvers on current systems because of memory limitations. Overall the proposed algorithm allows for a significant reduction of computational resources and enables solving large scale 3D acoustic problems required for FWI. Performance benefits of the multi-frontal solver over the Intel MKL PARDISO was showed. Ways of improving OMP scalability and developing MPI parallelization were proposed.

References

1. *Dewilde P. Gu M. Somasunderam N. Chandrasekaran, S.* On the numerical rank of the off-diagonal blocks of schur complements of discretized elliptic pdes. *SIAM J. Matrix Anal. Appl.*, 31(5):2261–2290, 2010.

2. *Collino F., Tsogka C.* Application of the perfectly matched layer absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media
Geophysics. 2001. **66**. 294–307.

3. *George A.* Nested dissection of a regular finite elementmesh // SIAM Journal on Numerical Analysis. 1973. **10**, N 2. 345–63.

4. *Jo, C., Shin, C., and Suh, J.* An optimal 9-point finite-difference frequency-space 2-D scalar wave extrapolator
Geophysics. 1996. **61**. 294–307.

5. *Xia J.* Robust and efficient multifrontal solver for large discretized pdes // High-Performance Scientific Computing. 2012. 199–217.

6. *Rjasanow S.* Adaptive cross approximation of dense matrices. In *IABEM 2002, International Association for Boundary Element Methods*, UT Austin, TX, USA, May 28-30, 2002.

7. *Solovyev S. A.* Application of the Low-Rank Approximation Technique in the Gauss Elimination Method for Sparse Linear Systems Vychisl. Metody Programm. 15, 441–460 (2014), in Russian.

8. *J. Xia S. Wang, M.V. de Hoop and X.S. Li.* Massively parallel structured multifrontal solver for time-harmonic elastic waves in 3D anisotropic media. In *Project Review, Geo-Mathematical Imaging Group*, volume 1, pages 97–121. Purdue University, West Lafayette IN, 2012.

9. *Hackbusch W.* A sparse matrix arithmetic based on H-matrices. part I: Introduction to H-matrices. *Computing*, 62(2):89 – 108, 1999.