Iterative methods for 3D implicit finite-difference migration using the complex Padé approximation

Carlos A. N. Costa, Itamara S. Campos, Jessé C. Costa, Francisco A. Silva Neto, UFPA and INCT-GP, Brazil, Jörg Schleicher and Amélia Novais, Unicamp and INCT-GP, Brazil

Resumo

Conventional implementations of 3D finite-difference (FD) migration use splitting techniques to accelerate performace and save computational cost. However, such techniques are plagued with numerical anisotropy that jeopardizes correct positioning of dipping reflectors in the directions not used for the splitted operators. We implement 3D downward continuation migration without splitting in the space coordinates using a complex Padé approximation and implicit finite differences. In this way, the numerical anisotropy is eliminated at the expense of a computationally more intensive solution of a large banded linear system. We compare the performance of the iterative stabilized biconjugate gradient (BICGSTAB) and the multifrontal massively parallel direct solver (MUMPS). It turns out that the use of the complex Padé approximation is an effective preconditioner for the BICGSTAB, reducing the number of iterations relative to the real Padé expansion of the FD operator. As a consequence, the iterative BICGSTAB method was more efficient than the direct MUMPS method when solving for a single term in the Padé expansion. The properties of these algorithms are evaluated computing the migration impulse response in the SEG/EAGE salt model.

Introduction

Efficient implementations of 3D FD migration use the splitting technique along horizontal coordinates to avoid the solution of large banded linear systems (Claerbout, 1985) at each downward continuation step. Splitting the migration operator along inline and crossline directions reduces 3D migration to a sequence of 2D downward continuation steps making the algorithm highly efficient. As a drawback splitting introduces numerical anisotropy which can damage the image of reflectors (Brown, 1983). Several strategies are used to remedy this problem usually assuming a homogeneous medium and using wavefield interpolation to handle lateral velocity variations (Li, 1991). However, such correction methods add again to the cost and introduce additional errors and artifacts. Exploring another idea to avoid the expensive direct solution of these large systems, Cole (1989) investigated the use of iterative methods to implement FD migration without splitting and evaluated the performace of overrelaxation, Jacobi, and Gauss-Seidel methods Iserles (1996). He reported the poor conditioning of the linear system matrix for low frequencies which degraded the performace of iterative solvers to a prohibitively high computational cost. Along a similar line, Nichols (1991) investigated the performace of the conjugate gradient method (CG) in dependence on the choice of the initial value for preconditioning. He found a degradation of performance of CG for low frequencies but reported that the performance improves when the propagation of energy associated with high angles and evanescent modes is attenuated.

In this respect, it is interesting to note that Amazonas et al. (2007) improved the quality of migrated images by representing the downward continuation operator using the complex Padé approximation (Millinazzo et al., 1997), which is essentially a means of attenuating the propagation of evanescent and high-angle wave modes. Moreover, new iterative methods to solve complex linear system are available that allow to compute the solution of very large linear systems. For example, the stabilized biconjugate gradient (BICGSTAB) method can solve poor conditioned complex linear systems more efficiently than CG (Van der Vorst, 1992). Another development in this area is the multifrontal massively parallel (MUMPS) direct solver which can efficiently solve large complex banded linear systems using Gaussian decomposition (Amestoy et al., 2001, 2006). In this paper, we investigate whether the limitations observed by Nichols (1991) can be overcome when using the complex Padé approximation in the implementation of 3D downward continuation by finite differences without splitting. Our numerical experiments in homogeneous media and in the SEG/EAGE salt model indicate that the complex Padé approximation is an effective preconditioner for 3D FD migration, so that the iterative BICGSTAB approach outperforms the massively parallel MUMPS solver when using a single term of the Padé expansion. If more terms are required to better model high dip events the direct solver is required.

Implicit finite-difference downward continuation

The one-way wave equation for downward continuation in the space-frequency domain is

$$\frac{\partial P(\mathbf{x},\omega)}{\partial z} = -\frac{i\omega}{c(\mathbf{x})}\sqrt{1+\frac{c(\mathbf{x})^2}{\omega^2}D}P(\mathbf{x},\omega), \qquad (1)$$

where $\mathbf{x} = (x,y,z)$ is the position with x and y denoting the coordinates along inline and crossline directions and z denoting depth, ω is the angular frequency, $P(\mathbf{x},\omega)$ is the wavefield spectrum, $D = D_x^2 + D_y^2$, with $D_x \equiv \frac{\partial}{\partial x}$, $D_y \equiv \frac{\partial}{\partial y}$,

and $c(\mathbf{x})$ represents the wave speed.

The square-root operator in equation (1) can expanded using the complex Padé series (Millinazzo et al., 1997; Amazonas et al., 2007):

$$\frac{\partial P}{\partial z} = -\frac{i\omega}{c(\mathbf{r})}\left[C_0 + \sum_{n=1}^{N} \frac{A_n \frac{c^2(\mathbf{r})}{\omega^2} D}{1 + B_n \frac{c^2(\mathbf{r})}{\omega^2} D}\right], \qquad (2)$$

where N indicates the order of the Padé expansion and C_0, A_n e B_n are the complex coefficients:

$$\begin{aligned} C_0 &= e^{i\theta/2}\left[1 + \sum_{n=1}^{N} \frac{a_n\left(e^{-i\theta}-1\right)}{1+b_n\left(e^{-i\theta}-1\right)}\right], \\ A_n &= \frac{a_n e^{-i\theta/2}}{\left[1+b_n\left(e^{-i\theta}-1\right)\right]^2}, \\ B_n &= \frac{b_n e^{-i\theta}}{1+b_n\left(e^{-i\theta}-1\right)}.. \end{aligned} \qquad (3)$$

Here, θ indicates the angle of rotation for the branch cut of the square root in the complex plane, and a_n and b_n are the real Padé expansion coefficients. The constant C_0 is actually an approximation to one that gets better with an increasing number of terms N used in the expansion. Therefore, we can set $C_0 = 1$.

The numerical solution of equation (2) is obtained by solving a sequence of differential equations (Claerbout, 1985). At each continuation step Δz, the solution is computed recursively for each term of the expansion. First one solves:

$$\frac{\partial P(\mathbf{x},\omega)}{\partial z} = -i\frac{\omega}{c(\mathbf{x})} P(\mathbf{x},\omega) \qquad (4)$$

and subsequently for $n = 1,\ldots,N$:

$$\frac{\partial P(\mathbf{x},\omega)}{\partial z} = -i\frac{\omega}{c(\mathbf{x})}\left[\frac{A_n \frac{c^2(\mathbf{x})}{\omega^2} D}{1 + B_n \frac{c^2(\mathbf{x})}{\omega^2} D}\right] P(\mathbf{x},\omega)\ , \qquad (5)$$

where the solution of the previous term is the initial condition for the solution of the next one. This algorithm converges for small Δz.

Using the Crank-Nicolson finite-differences scheme (Iserles, 1996) to approximate the derivative operators, equations 5 are reduced to the linear systems

$$(\mathbf{I} + \alpha\mathbf{D})\,\mathbf{P}^{k+1} = (\mathbf{I} + \alpha^*\mathbf{D})\,\mathbf{P}^k, \qquad (6)$$

where

$$\alpha \equiv B_n + i\frac{\omega\Delta z}{c} A_n\ , \qquad (7)$$

$$\alpha^* \equiv B_n - i\frac{\omega\Delta z}{c} A_n\ , \qquad (8)$$

$$\mathbf{D} \equiv \left(\frac{c^2}{\omega^2}\right)\left[\frac{\mathbf{D}_x}{(\Delta x)^2} + \frac{\mathbf{D}_y}{(\Delta y)^2}\right]\ , \qquad (9)$$

in which $\mathbf{I}$ is the identity matrix, $\mathbf{D}_x$ and $\mathbf{D}_y$ are the finite-differences matrices representing the derivatives along the x and y coordinates, and $\mathbf{P}^k$ and $\mathbf{P}^{k+1}$ represent the wavefield at the grid at successive depth levels. The matrix $\mathbf{D}$ has dimension $N_x \times N_y$. Though sparse, this matrix can have a large band. This increases the computation cost of solving system 6 with direct methods. An efficient and frequently used method to reduce this cost is splitting along inline and crossline directions,

$$\left(\mathbf{I} + \alpha\left(\frac{c}{\omega\Delta y}\right)^2 \mathbf{D}_x^2\right)\left(\mathbf{I} + \alpha\left(\frac{c}{\omega\Delta x}\right)^2 \mathbf{D}_y^2\right)\mathbf{P}^{k+1} = \left(\mathbf{I} + \alpha^*\left(\frac{c}{\omega\Delta y}\right)^2 \mathbf{D}_x^2\right)\left(\mathbf{I} + \alpha^*\left(\frac{c}{\omega\Delta x}\right)^2 \mathbf{D}_y^2\right)\mathbf{P}^k. \qquad (10)$$

The solution of this linear system is a sequence of tridiagonal linear systems each one of dimension N_x and N_y. Unfortunately, this approximation introduces numerical anisotropy. Figure 2 compares the wavefields in a homogeneous medium computed at the same depth without splitting, equation (6), and with splitting along inline and crossline directions. The noncircular format of the result of splitting is the consequence of numerical anisotropy. To reduce numerical anisotropy, Li (1991) proposed a correction in the frequency-wavenumber domain which requires wavefield interpolation in laterally heterogeneous media and Ristow and Rühl (1997) used multiway splitting. However, such correction methods add to the computational cost and introduce additional errors and migration artifacts. For this reason, we are interested in alternative methods to solve system (6) that do not make use of splitting techniques.

Iterative methods

The sparseness of linear system (6) makes it suitable for iterative methods (Iserles, 1996). Cole (1989) investigated the performance of some iterative methods to solve equation (6) using the first order real Padé expansion. He noticed the slower convergence, or even convergence failure, of methods like overrelaxation, Jacobi and Gauss-Seidel when applied for frequencies below $\omega^2 = 8c^2 b_n/(\Delta x)^2$. (Nichols, 1991) used the conjugate gradient method (CG) to solve (6) because of its better theoretical covergence properties. He also noticed the poor performance of this method for low frequencies, but reported some improvement in the convergence speed of CG when evanescents modes are filtered out. However, because CG solves the normal equations, the performance degrades rapidly for ill conditioned matrices.

We investigate whether the limitations of conventional iterative methods as reported in the cited works still hold for more modern techniques. Taking into account the observation of Nichols (1991) that the performance of CG improves when the propagation of energy associated with high angles and evanescent modes is attenuated, we use the BICGSTAB algorithm (Van der Vorst, 1992) to implement downward continuation without splitting using the complex Padé expansion. We chose BICGSTAB because ofits better convergence properties than CG for complex matrices. The performance of iterative methods improves when the matrix is strictly diagonally dominant (Iserles, 1996). For a matrix $\mathbf{A}$ this means

$$|A_{ii}| \geq \sum_{j=1}^{N} |A_{ij}|_{i\neq j} \quad i = 1,\ldots,N\ . \qquad (11)$$

For a homogeneous medium we can verify when the matrix in equation (6) satisfies this condition. In this case its

diagonal entries are

$$A_{ii} = \left[\left(\frac{\omega \Delta x}{c} \right)^2 - 4 \left(B_1 + i \frac{\omega \Delta x}{2c} A_1 \right) \right]$$

and the only five nonzero off diagonal entries in each row are equal to

$$A_{ij} = \left(B_1 + i \frac{\omega \Delta x}{2c} A_1 \right) \quad i \neq j \ .$$

Thus, this matrix is diagonally dominant if

$$\omega \geq \omega_L = \frac{c}{\Delta x} \left[-2\Im\{A_1\} + \sqrt{(2\Im\{A_1\})^2 + 8\Re\{B_1\}} \right] \quad (12)$$

where $\Im\{A_1\}$ and $\Re\{B_1\}$ represent the imaginary and real parts of A_1 and B_1, respectively. This inequality determinines the minimum frequency for which the matrix is strictly diagonally dominant. For frequencies above this limit, iterative methods should perform better. Although this result is only a sufficient condition, it gives some clues on the performance of iterative methods for downward continuation. The limit depends on the ratio between propagation velocity and grid interval. When this ratio increases, the convergence of iterative methods is slower. It also implies that the complex Padé expansion can affect the convergence through $\Im\{A_1\}$. The top left part of Figure 1 shows how the limit frequency $f_L = \omega_L / 2\pi$ depends on the branch cut rotation angle θ, for $\Delta x = 10$ m and $c = 1500$ m/s. This figure indicates that increasing θ should improve the performance of iterative methods.

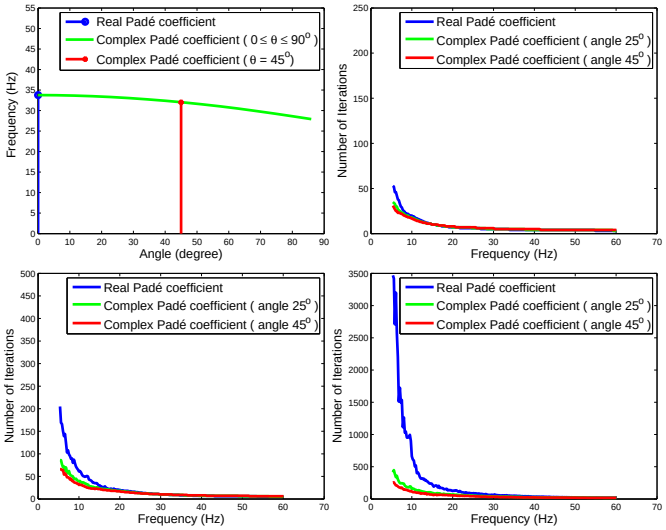

Figura 1: Top left: Limit frequency, f_L, versus rotation angle in degrees. Top right: number of iterations of BICGSTAB for each frequêncy for $c = 1500m/s$ and $h = 20m$. Bottom left: iterations versus frequency, $c = 1500m/s$ and $h = 10m$. Bottom right: iterations versus frequency, $c = 4500m/s$ and $h = 10m$.

To evaluate how the actual behavior of BICGSTAB relates to inequality (12), the remaining graphs in Figure 1 display the number of iterations as a function of frequency for increasing values of the ratio $c/\Delta x$ for real and for complex Padé coefficients with $\theta = 25^\circ$ and $\theta = 45^\circ$. These graphs shows that the complex Padé approximation markedly improves the convergence of BICGSTAB when the ratio $c/\Delta x$ increases. Increasing the branch rotation from 25° to 45°, though reducing the number of iterations, has a less pronounced effect.

Numerical Results

As a first test, we computed the impulse response of the 3D FD migration operator in a homogeneous medium with and without inline and crossline splitting. The propagation velocity is 1500 m/s. We used a single term of the complex Padé expansion with rotation angle $\theta = 45^\circ$ to approximate the migration operator. The impulse response without splitting was computed using BICGSTAB. The grid dimensions are $610 \times 610 \times 210$ along coordinates x, y, and z, respectively. The grid spacing is uniform and equal to 10 m. The volume injection source is located one grid spacing below the surface in the center of the grid. The source signature is a Ricker pulse with peak frequency 25 Hz, and the sample rate is 8 ms.

Figure 2 shows sections of the 3D impulse responses at depth $z = 1050$ m. The red circle in these images indicate the exact position. The numerical anisotropy caused by the splitting is evident, as well as the absence of numerical anisotropy in the impulse response computed by BICGSTAB.

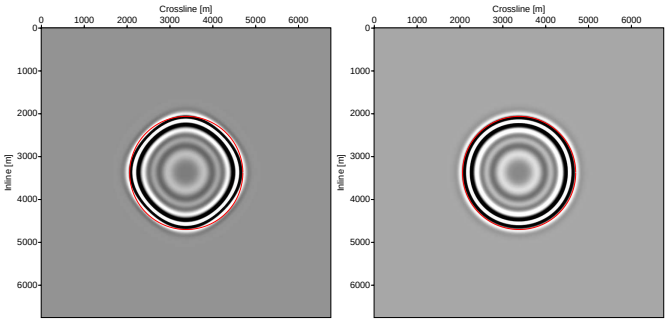

Figura 2: FD impulse response in a homogeneous medium. Left: using splitting along inline and crossline directions. Right: whit-out splitting along coordinate directions.

Next, we use the SEG/EAGE salt model to evaluate the performance of 3D migration without splitting in a complex inhomogeneous medium. The strong lateral variations between the sediments and the salt body and the high velocity of the salt are a challenge for the iterative methods. To filter out the spike velocity contrast used to simulate reflectors in the original model we applied median filter of size $7 \times 7 \times 7$. The grid size, source position and source pulse are the same as for the homogeneous example. We computed FD impulses responses using BICGSTAB with a single term of the complex Padé expansion $\theta = 25^\circ$ and $\theta = 45^\circ$. These impulse responses are compared to FD impulse response computed using the direct solver MUMPS using three terms of the complex Padé response and $\theta = 45^\circ$. Finally, we compare them to the 3D impulse response of reverse time migration.

For a single term of the complex Padé expansion the MUMPS direct solver is two times slower than BICGSTAB. On the other hand, for higher order complex Padé terms, the performance of BICGSTAB degrades rapidly, and increasing the angle θ does not help in this case. This agrees with the observation of Nichols (1991) that iterative

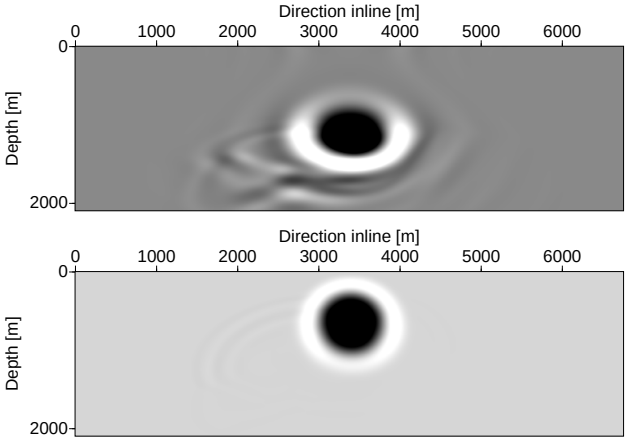

Figura 3: Sections of 3D migration impulse response in the plane $x = 1500$ m. Top: FD migration, $\theta = 25°$. Bottom: RTM

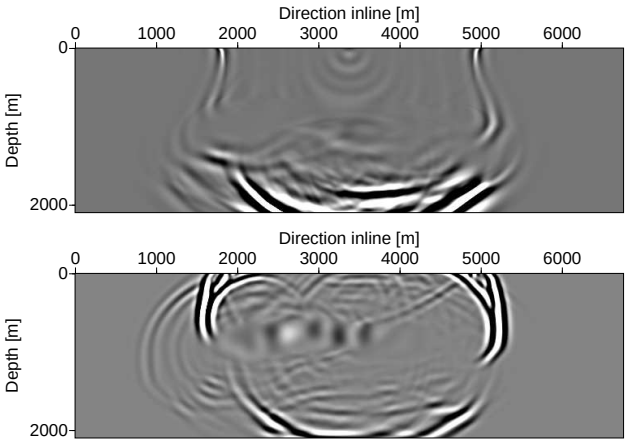

Figura 4: Sections of 3D migration impulse response in the plane $x = 3380$ m. Top: FD migration, $\theta = 25°$. Bottom: RTM

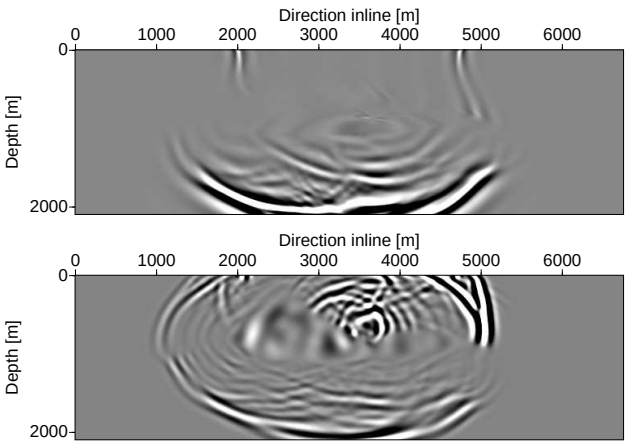

Figura 5: Sections of 3D migration impulse response in the plane $x = 4160$ m. Top: FD migration, $\theta = 25°$. Bottom: RTM

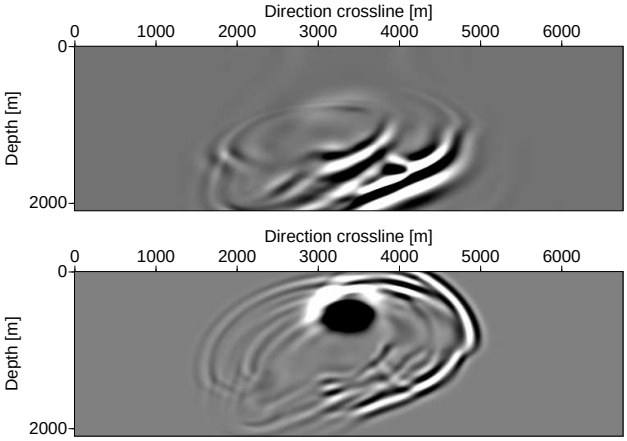

Figura 6: Sections of 3D migration impulse response in the plane $y = 1500$ m. Top: FD migration, $\theta = 25°$. Bottom: RTM.

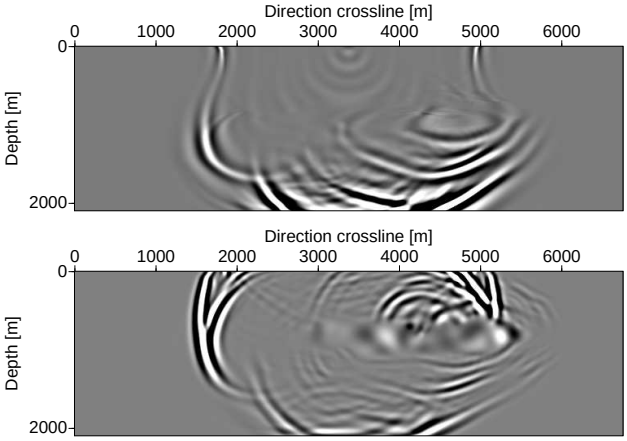

Figura 7: Sections of 3D migration impulse response in the plane $y = 3380$ m. Top: FD migration, $\theta = 25°$. Bottom: RTM.

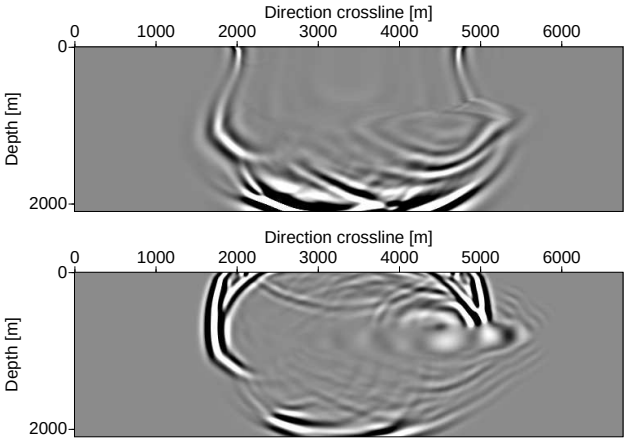

Figura 8: Sections of 3D migration impulse response in the plane $y = 4160$ m. Top: FD migration, $\theta = 25°$. Bottom: RTM.

methods have more difficult modeling the wavefield propagating at higher dip angles. Only the direct solver was able to compute the impulse response with three Padé terms.

Figures 3, 4 and 5 show three vertical sections of the impulse response along the inline direction (x) computed with 3D FD migration using BICGSTAB with a single complex Padé term with $\theta = 25°$ and compare them to corresponding 3D reverse time migration results. Because of its intrinsic restrictions, FD migration cannot properly image events dipping above $45°$. The positioning of FD migrated lower dipping events correctly matches the corresponding RTM events. The relative amplitude of the events in the FD migration results does not reproduce the reverse time migration amplitudes for corresponding

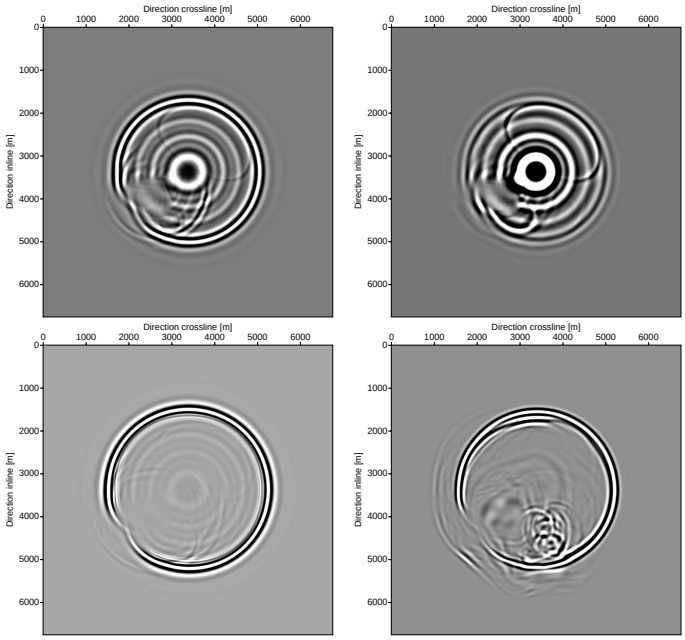

Figura 9: Sections of 3D impulse responses in the plane $z = 550$ m. Top left: FD migration computed using BICGSTAB, $\theta = 25^\circ$. Top right: FD migration using BICGSTAB $\theta = 45^\circ$. Bottom left: FD migration computed using direct solver MUMPS, $\theta = 45^\circ$. Bottom right: RTM.

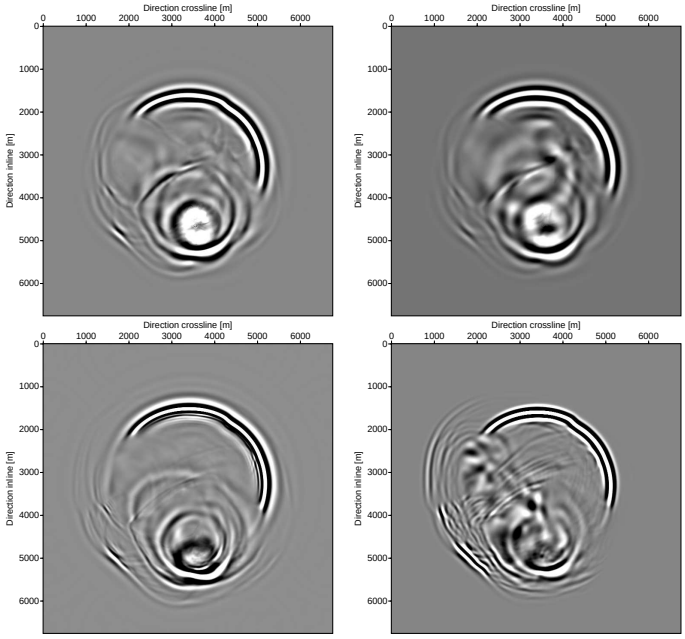

Figura 10: Sections of 3D impulse responses in the plane $z = 1050$ m. Top left: FD migration computed using BICGSTAB, $\theta = 25^\circ$. Top right: FD migration using BICGSTAB $\theta = 45^\circ$. Bottom left: FD migration computed using direct solver MUMPS, $\theta = 45^\circ$. Bottom right: RTM.

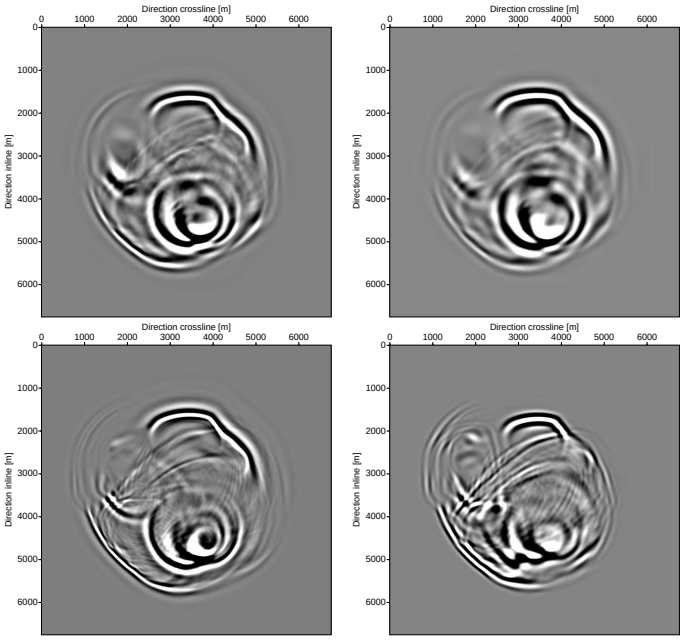

Figura 11: Sections of 3D impulse responses in the plane $z = 1350$ m. Top left: FD migration computed using BICGSTAB, $\theta = 25^\circ$. Top right: FD migration using BICGSTAB $\theta = 45^\circ$. Bottom left: FD migration computed using direct solver MUMPS, $\theta = 45^\circ$. Bottom right: RTM.

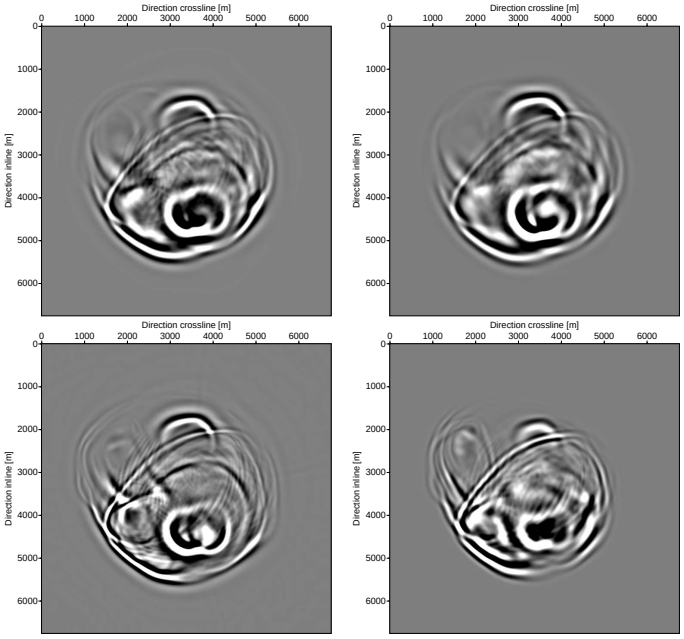

Figura 12: Sections of 3D impulse responses in the plane $z = 1550$ m. Top left: FD migration computed using BICGSTAB, $\theta = 25^\circ$. Top right: FD migration using BICGSTAB $\theta = 45^\circ$. Bottom left: FD migration computed using direct solver MUMPS, $\theta = 45^\circ$. Bottom right: RTM.

events. However, it should be kept in mind that amplitude comparisons are difficult because the presence of upward propagating waves in the RTM image may lead to a different scaling of the figures.

Figures 6, 7 and 8 show the corresponding comparison of vertical sections along the crossline direction (y). Again, note the good match among corresponding events dipping up to around 45° degrees, but the differences for higher propagation angles and in amplitudes.

Figures 9, 10, 11, and 12 show horizontal sections of the 3D impulse response computed by four different methods: FD migration solved using BICGSTAB for one complex Padé term with $\theta = 25^\circ$ (top left) and $\theta = 45^\circ$ (top right), FD migration using the MUMPS direct solver for three

complex Padé terms with $\theta = 45^\circ$ (bottom left) and reverse time migration (bottom right). The comparison to the RTM response is difficult in these horizontal sections, because upward propagating waves cannot be easily recognized as such. The most striking difference between the different methods appears in Figure 9, where the two BICGSTAB results don't position the event correctly and present strong artifacts. These differences in the shallowest section result from the 45° dip limit of the single-term complex Padé expansion. The wide-angle response of three-term complex Padé approximation used in the MUMPS solution is sufficient to correctly recover the position and suppress the artifacts. However, note that this wide-angle response costs six times more computation time than using BICGSTAB with a single term. For the remaining sections the difference among corresponding events diminishes with depth, indicating that the iterative solver is sufficiently accurate for propagation directions below 45°.

Computation time

When comparing non-splitting solvers of 3D FD wave-equation migration, a word on computation time is indispensable. In our implementation, the BICGSTAB solution for one-term complex Padé FD migration took half as much time as the corresponding MUMPS solution and about five times as much as a corresponding solution based on alternating four-way splitting plus Li correction with 10 reference velocities. The three-term MUMPS solution consumes three times more time than the one-term solution, while the three-term BICGSTAB solution did not converge. The computation time of our RTM implementation is not fully comparable, since it uses no disk I/O, which consumes much time in our implementations of the other methods. Under these circumstances, it was of the same order as the one-term MUMPS solution.

Conclusions

The performance of iterative methods for 3D downward continuation without splitting along inline and crossline directions improves markedly using the complex Padé approximation. Numerical experiments using the SEG/EAGE salt model show the feasibility of FD migration without splitting for such a complex velocity model, being less then an order of magnitude slower than conventional splitting approaches. We provided a simple analysis indicating that the complex Padé expansion improves the conditioning of FD downward continuation. For a single term in the Padé expansion, BICGSTAB outperforms the massively parallel direct solver by a factor of two. The performance of iterative methods degrades rapidly when more terms of the Padé are used but does not affect MUMPS in the same way. The computation time of MUMPS increases only linearly with the number of terms.

Efficiency of full solutions without operator splitting still remains an issue. Our results indicate that the combination of the complex Padé approximation with additional preconditioning can further reduce the computational cost, possibly bringing non-splitting solutions within the range of more sophisticated corrections to splitting. Moreover, since low frequency is the main problem affecting the performance of BICGSTAB, multigrid techniques can certainly help to speed up these methods.

Acknowledgments

We acknowledge the financial support from PETROBRAS and the Brazilian Agencies CAPES, CNPq and FINEP, as well as the sponsors of the Wave Inversion Technology (WIT) consortium.

Referências

Amazonas, D., J. C. Costa, J. Schleicher, and R. Pestana, 2007, Wide-angle fd and ffd migration using complex padé approximations: Geophysics, **72**, S215–S220.

Amestoy, P. R., I. S. Duff, J. Koster, and J.-Y. L'Excellent, 2001, A fully asynchronous multifrontal solver using distributed dynamic scheduling: SIAM Journal of Matrix Analysis and Applications, **23**, 15–41.

Amestoy, P. R., A. Guermouche, J.-Y. L'Excellent, and S. Pralet, 2006, Hybrid scheduling for the parallel solution of linear systems: Parallel Computing, **32**, 136–41.

Brown, C. L., 1983, Application of operator separation in reflection seismology: Geophysics, **48**, 288–294.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Sci. Pub.

Cole, S., 1989, Iterative methods for 3-D finite-difference migration and modeling: Technical report, Stanford University.

Iserles, A., 1996, A first course in the numerical analysis of differential equations: Cambridge University Press.

Li, Z., 1991, Compensating finite-difference migration by multiway splitting: Geophysics, **56**, 1650–1660.

Millinazzo, F. A., C. A. Zala, and G. H. Brooke, 1997, Square-root aproximations for parabolic equation algorithms: J. Acoust. Soc. Am., **101**, 760–766.

Nichols, D., 1991, 3-D depth migration by a predictor corrector method: Technical report, Stanford University.

Ristow, D., and T. D. Rühl, 1997, 3D finite-difference erros in 3D migration and modeling: Geophysics, **62**, 554–567.

Van der Vorst, H. A., 1992, Bi-CGSTAB: a fast and smoothly converging variant of BI-CG for nonsymmetric linear systems: SIAM J. Sci. Stat. Comput., **13**, 631–644.