Elastodynamic FWI in 2D with partial stacking

Vladimir Zubov, Gary F. Margrave and Michael P. Lamoureux

ABSTRACT

The present study is devoted to stability and noise-resistance of the partial stack Full Waveform Inversion (FWI) iterative algorithm numerical investigation with the method applied to elastodynamic partial differential equation boundary problem. The economic efficiency of the partial stack was verified on the test model in comparison with full stack FWI. Both non-uniform error in the right hand of the non-linear operator inverse problem and the iterative FWI algorithm convergence behavior were considered as objects of the numerical investigations.

The particular authors' attention was paid to numerical comparison of the density FWI with bulk modulus FWI convergence taking corresponding imaging conditions into consideration. The imaging conditions for both fields FWI are derived explicitly from elastodynamic conservation laws formulated in integral form.

INTRODUCTION

The design of Full Waveform Inversion model usually starts with the mathematical problem formulation as conservation laws in differential form ready for numerical implementation with grid or partials models. Based on the paper on elastodynamic equation numerical solution using a grid finite-differential scheme published by Virieux (1986), we starts with the forward elastodynamic equation boundary problem approximated on the regular staggered grid with central differences providing 2nd order time-spatial accuracy of approximation and the conditional stability to the numerical algorithm.

The corresponding inverse problem formulation in case of the elastodynamic equation gives us a variety in the coefficient field choice for the FWI. As well as in the conventional FWI, which automatically chooses density field as the unknown reflection coefficient vector, we consider density with the corresponding imaging conditions as a target of the standard FWI computational routine. On the other hand, the bulk modulus FWI is of interest in the present study as well. Moreover, the idea to compare the density FWI with bulk modulus FWI leads us in the numerical experiments section of the present report.

In order to organize the computational routine of the FWI iterative model more effectively, we use the well-known partial stack approach which gives us the chance to involve different parts of the non-linear model complexity on different iterations of the FWI with the aim to improve the overall stability of the convergence behaviour of the FWI. This means, that the model needs the proper calibration on preliminary stage of the FWI as prior information on the target solution. One may compare partial stack with simulations shooting in some way doubting the efficiency of the partial stack. The comparison of these two approaches was not made in the present study but would be interesting in the future.

FORWARD ELASTODYNATIC EDUATION BOUDARY PROBLEM

In 2-dimentional spatial elastic domain Ω, let us consider the deformation fields (u_x, u_z) governed by the elastodynamic equation (Virieux 1986) formulated without source of deformation in inner Ω:

$$\begin{cases} \frac{\partial}{\partial t}\rho\frac{\partial u_x}{\partial t} = \frac{\partial}{\partial x}(\lambda+2\mu)\frac{\partial u_x}{\partial x} + \frac{\partial}{\partial x}\lambda\frac{\partial u_z}{\partial z} + \frac{\partial}{\partial z}\mu\frac{\partial u_z}{\partial x} + \frac{\partial}{\partial z}\mu\frac{\partial u_x}{\partial z} \\ \frac{\partial}{\partial t}\rho\frac{\partial u_z}{\partial t} = \frac{\partial}{\partial x}\mu\frac{\partial u_z}{\partial x} + \frac{\partial}{\partial x}\mu\frac{\partial u_x}{\partial z} + \frac{\partial}{\partial z}\lambda\frac{\partial u_x}{\partial x} + \frac{\partial}{\partial z}(\lambda+2\mu)\frac{\partial u_z}{\partial z} \end{cases} \quad (1)$$

where ρ is density, λ and μ are bulk modulus, x represents horizontal spatial direction and z – vertical.

To solve this problem numerically, the same method as described in the Virieux's (1986) paper is used. We transform each 2nd order differential equation into the system of the first order ones adding the source component $f(x, z, t)$ into the deformation tensor explicitly:

$$\begin{cases} \rho\frac{\partial v_x}{\partial t} = \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xz}}{\partial z}, \rho\frac{\partial v_z}{\partial t} = \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{zz}}{\partial z}, \\ \frac{\partial \tau_{xx}}{\partial t} = (\lambda+2\mu)\frac{\partial v_x}{\partial x} + \lambda\frac{\partial v_z}{\partial z} + f(x,z,t), \\ \frac{\partial \tau_{xz}}{\partial t} = \mu\left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right), \\ \frac{\partial \tau_{zz}}{\partial t} = (\lambda+2\mu)\frac{\partial v_z}{\partial z} + \lambda\frac{\partial v_x}{\partial x} + f(x,z,t), \end{cases} \quad (2)$$

where unknowns (v_x, v_z) represent the speed of deformation satisfying the following homogeneous initial and boundary conditions:

$$\begin{cases} v_x(t=0) = v_z(t=0) = \tau_{xx}(t=0) = \\ \quad \tau_{xz}(t=0) = \tau_{zz}(t=0) = 0, \\ \quad v_x(\partial\Omega) = v_z(\partial\Omega) = \\ \quad \tau_{xx}(\partial\Omega) = \tau_{xz}(\partial\Omega) = \\ \quad \tau_{zz}(\partial\Omega) = 0 \end{cases} \quad (3)$$

Everywhere below, we use $f(x, z, t) = \delta(x - x_0) \cdot \delta(z - z_0) \cdot e^{-(\eta(t-t_0))^2} \sin(\omega t)$ as an point source function used in numerical experiments with fixed $\eta = \eta_0$ and $\omega = \omega_0$. The source spatial location (x_0, z_0) is given with the model input data somewhere in the middle of the computational area Ω reasonably far from boundaries to guarantee the time interval [0,T] when boundary conditions support the numerical stability of the model and the uniqueness of the solution but don't provide the reflection of the first arrivals. This time interval [0,T] can be roughly estimated through the mentioned distance of the source point from the boundary, maximum allowed propagation speed and the source appearance delay t_0. All these limitations on the source location and the modelling time interval make us to include the layer of air into the computational domain Ω considering the following spatial area of the boundary problem in numerical experiments (Figure 1) with an approximate ratio of the spatial size 2×1.

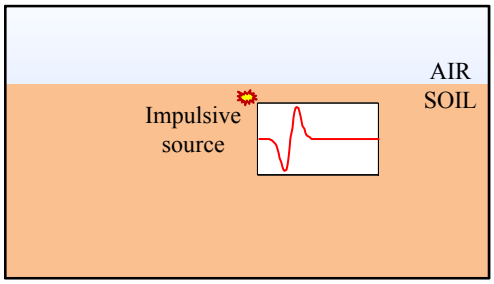


FIG 1. Two-layered spatial area Ω with the impulsive source location and signature.

ADJOINT WAVE REVERSE TIME MIGRATION AND IMAGING CONDITIONS

Integration by parts in the imaging conditions derivation leads us directly to the adjoint problem solution in both acoustic and elastodynamic cases and for either density or bulk modulus components of the FWI. As a result, it is important to be able to solve both direct and adjoint operator in both forward and backward time direction fast.

Let us briefly consider one effective, easy and popular method known as the heat equation way (Hasanov 2009) of imaging conditions derivation without Green functions. The imaging condition we will derive for the bulk modulus λ only but other imaging conditions (Tarantola 1984) are derivable as well.

Let us start with two solutions (u_{1x}, u_{1z}) and (u_{2x}, u_{2z}) of (1) for two different bulk modulus λ_1 and λ_2. Then the difference $\Delta u = u_2 - u_1$ will satisfy

$$\begin{cases} \frac{\partial}{\partial t}\rho\frac{\partial \Delta u_x}{\partial t} = \frac{\partial}{\partial x}(\lambda_1 + 2\mu)\frac{\partial \Delta u_x}{\partial x} + \frac{\partial}{\partial x}\Delta\lambda\frac{\partial u_{2x}}{\partial x} \\ + \frac{\partial}{\partial x}\lambda_1\frac{\partial \Delta u_z}{\partial z} + \frac{\partial}{\partial x}\Delta\lambda\frac{\partial u_{2z}}{\partial z} + \frac{\partial}{\partial z}\mu\frac{\partial \Delta u_z}{\partial x} + \frac{\partial}{\partial z}\mu\frac{\partial \Delta u_x}{\partial z}, \\ \frac{\partial}{\partial t}\rho\frac{\partial \Delta u_z}{\partial t} = \frac{\partial}{\partial x}\mu\frac{\partial \Delta u_z}{\partial x} + \frac{\partial}{\partial x}\mu\frac{\partial \Delta u_x}{\partial z} + \frac{\partial}{\partial z}\lambda_1\frac{\partial \Delta u_x}{\partial x} \\ + \frac{\partial}{\partial z}\Delta\lambda\frac{\partial u_{2x}}{\partial x} + \frac{\partial}{\partial z}(\lambda_1 + 2\mu)\frac{\partial \Delta u_z}{\partial z} + \frac{\partial}{\partial z}\Delta\lambda\frac{\partial u_{2z}}{\partial z}. \end{cases} \quad (4)$$

Now we consider (Φ_x, Φ_z) as a solution of the system adjoint to (1), multiply the first equation in (4) by φ_x, the second equation in (4) by φ_z, integrate over Ω and [0,T] and sum. There is no source function neither in (1) or (4) and the dynamics in the system is defined via non-homogenous boundary conditions of Dirichlet type.

Doing integration by parts, applying boundary conditions to curve integrals and establishing additional easily implementable limitations $\varphi_x(\partial\Omega) = 0$ and $\varphi_z(\partial\Omega) = 0$, we finally get the imaging condition for bulk modulus λ structurally identical to formula (37b) in the Tarantola's paper (1984):

$$\int_\Omega \int_0^T \Delta\lambda \left(\frac{\partial \Phi_x}{\partial x} + \frac{\partial \Phi_z}{\partial z}\right)\left(\frac{\partial u_{2x}}{\partial x} + \frac{\partial u_{2z}}{\partial z}\right) dtdxdz = -\int_{\partial\Omega} \int_0^T \mu \left(\frac{\partial \Phi_x}{\partial \vec{n}} + \frac{\partial \Phi_z}{\partial \vec{n}}\right)(\Delta u_x + \Delta u_z) dtdS \quad (5)$$

The fact that we move in depth the point source in (2) from the boundary air-soil interface makes us expect reasonable difficulties with the explicit imaging conditions implementation in the model (2)-(3) and also makes us require the complete information about the under-surface layer near the source location down to the source depth z_0 which is hardly possible to get in physical experiments.

The standard imaging condition requires the original and adjoint problem's solution integration in time by a space/volume integral. The numerical implementation of this integration would require the time direction turning in either (2) or its adjoint equation. Fortunately, both time direction reversion with the change of variable $t = -t$ and adjoint operator provides the same equation (2) as this equation is invariant to both mentioned operations. Time direction reversion in (1) still requires changes in initial conditions from standard homogenous to the following one

$$u(t = T) = \tilde{u}, u_t(t = T) = \tilde{u}_t \quad (6)$$

The initial conditions for equation adjoint to (1) are almost the same as (6) but homogenous:

$$\Phi(t = T) = 0, \Phi_t(t = T) = 0 \quad (7)$$

The formula (9) in the following section describes the relationships between deformation Φ and the speed of deformation φ which is much easier to compute numerically.

As a result, the imaging condition numerical computing on practice is easily implementable. Moreover, the computational cost of the imaging conditions is 4 times greater than the cost of the forward propagation; only 2 times greater in parallel implementation.

INVERSION ALGORITHM FOR PARTIAL STACK FWI

The partial stack density FWI iterative algorithm for the boundary problem (2)-(3) in its variables used in present study is presented in the block scheme form in Figure 2 and is based on the previous together work of Margrave, Ferguson and Hogan (2011). Bulk modulus FWI will affect imaging condition computing block of the diagram only. The corresponding formula for I_{α_n, f_n} (Figure 2) in λ FWI algorithm is the following

$$I_{\alpha_n, f_n} = \int_0^T \left(\frac{\partial \Phi_x}{\partial x} + \frac{\partial \Phi_z}{\partial z}\right)\left(\frac{\partial u_x}{\partial x} + \frac{\partial u_z}{\partial z}\right) dt \quad (8)$$

where

$$\begin{cases} u_x(t) = \int_o^t v_x(\tau)d\tau\,,\ u_z(t) = \int_o^t v_z(\tau)d\tau\,, \\ \Phi_x(t) = \int_t^T \varphi_x(\tau)d\tau,\ \ \Phi_z(t) = \int_t^T \varphi_z(\tau)d\tau \end{cases} \quad (9)$$

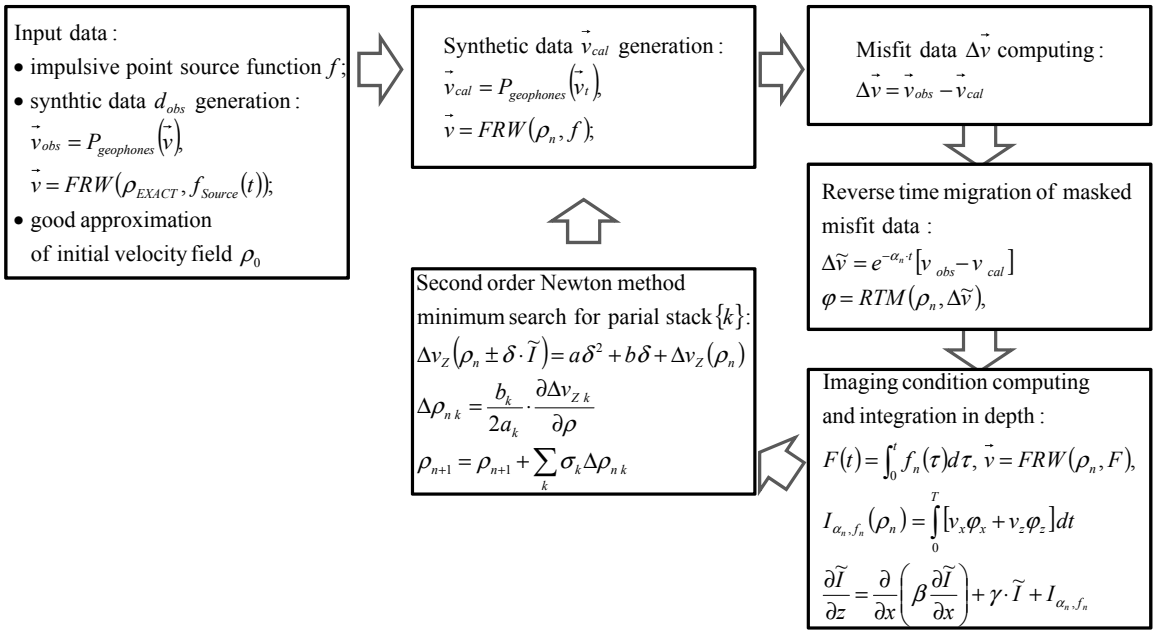


FIG 2. Density FWI iterative algorithm block diagram.

The partial stack approach appears in the iterative strategy of source switching from iteration to iteration of FWI. Depending on the source fraction used on each iteration, the strategy may be either random or conditional. The periodical strategy used in the next section's experiments is presented in the Table 1. It uses just 25% of the whole amount of sources as active on each iteration of FWI.

Table 1. Source switching strategy in partial stack FWI: blue cells – active sources, white cells – inactive sources.

FWI iteration	Sources										
	1	2	3	4	5	6	7	8	9	10	11
1	→				→				→		
2		→				→				→	
3			→				→				→
4				→				→			
5	←				←				←		
6				←				←			
7			←				←				←
8		←				←				←	
9	→				→				→		

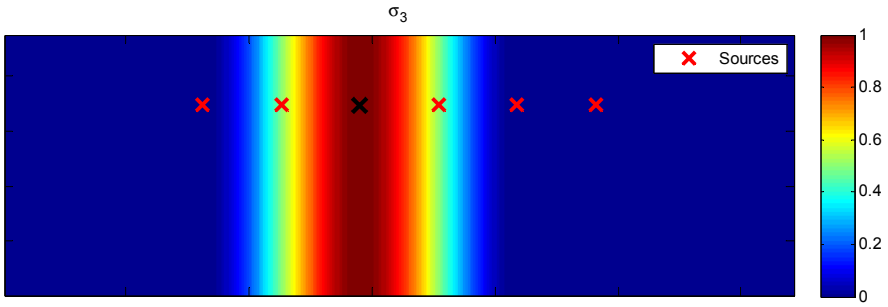


FIG 3. Weight function $\boldsymbol{\sigma_3}$ masking the impact of the black *x* source in either $\sum_i \boldsymbol{\sigma_i}\,\Delta\boldsymbol{\rho_{n\,i}}$ or $\sum_i \boldsymbol{\sigma_i}\,\Delta\lambda_{\boldsymbol{n\,i}}$.

For each particular source the Newton method search is implemented independently with corresponding imaging condition. The synchronization of parallel gradient search appears in density or bulk modulus field correction presented in Figure 2 as the ρ_{n+1} update. In this update, the direct sum weighted with σ corrections of all components of the partial stack is used. The example of the weight function used numerically is presented on the Figure 3 and σ in it is governed with the cosines squared of the horizontal offset from the current source to the next second neighbour source. In this case σ is differentiable once in space which matches the corresponding condition on λ.

Both misfit data $\Delta v = v_{observed} - v_{calculated}$ in the reverse time migration block and the source function $f(x, z, t)$ in the imaging condition calculation block are filtered in frequency domain prior to the use. The filter function is a normal Gaussian with 0-frequency as an expected value and dispersion growing with iterations of FWI. The source filtering example is presented in the Figure 4.

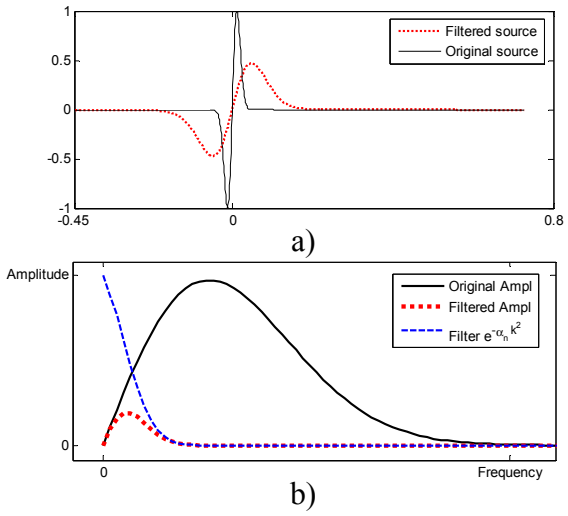


FIG 4. Source function $\boldsymbol{f(t) = e^{-(\eta t)^2} sin(\omega t)}$ filtering result: a) signal in time domain; b) amplitude spectrum and filter in frequency domain.

The objective function Δv of the optimization problem is treated in the energy norm of the positive definite diagonal matrix A where the diagonal is a linearly decreasing function of time such as

$$\|\Delta\vec{v}_x\|_A^2 = \sum_i \frac{T-t_i}{T}\Delta v_{x\,i}^2 \qquad (10)$$

This norm increases the weight of the first arrivals in the same way that the regularization does.

NUMERICAL RESULTS

The partial stack FWI study in this section is implemented using synthetic observation data. We study both the algorithm's stability and its noise resistance. The first step was the model validation and calibration both for full and partial stack density FWI. The number of sources in the full stack was taken as 6 and is consider as small but still enough to operate the Marmousi model for density (Figure 5), simplified horizontally in such a way that the resolution of the whole model is 321×475 with grid point sizes approximately 0.0094×0.0021 in Ω, sized as 4×1 squared units. Each grid point in the finite difference scheme is stretched by about 4 times from the squared one which makes the problem more horizontal than the original Marmousi.

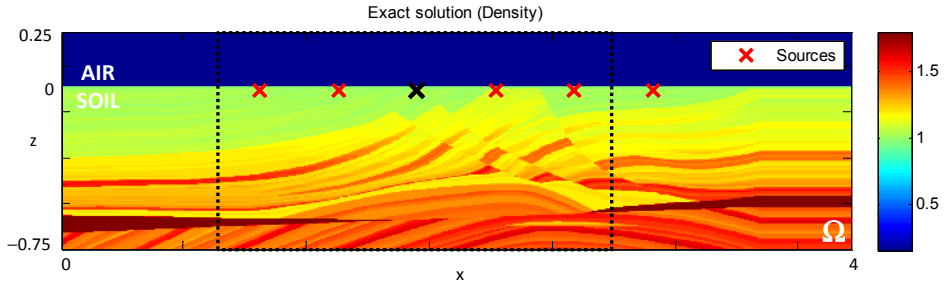


FIG 5. Density field for the synthetic observation's exact solution, with a dotted line representing the spatial window for the demonstrated black x source processing.

The spatial domain consists of 2 layers: air and earth with a horizontal linear interface in between. The propagation speed in air is known so that the FWI is formulated in the earth as usual.

All 6 point sources in the full stack model are put under the surface. The line of geophones available on the surface between air and earth for each of the sources is limited with the corresponding spatial window in which each source is processed. The example of this window is presented in the Figure 4 as well. Both variables λ and μ determining the reflection coefficient fields in (2) are taken as constant 1, while in air we have $\lambda = \mu = 0.01$ and $\rho = 0.16$ to make the gradient in wave propagation speed between air and earth more realistic.

Starting with a good initial guess of the density field obtained through iterative Gauss smoothing with equal weight of until it converges, applying 128 iterations of the FWI we obtain the following approximate density field (Figure 6). In these 128, we changed filtering 32 times starting with low frequencies and continuously involving higher

frequency components. Each filtering is used for 8 iterations of FWI, and for each FWI iteration just 1 iteration of the corresponding Newton gradient search is implemented. As a result, we observe a reasonably good match between high frequency fluctuations in the approximate solution and the exact one (Figures 5 and 6).

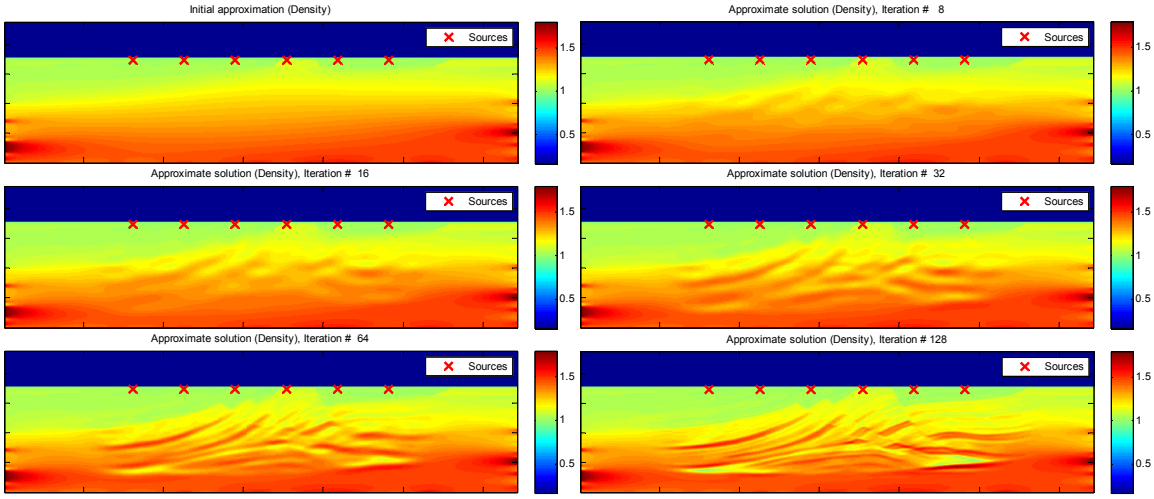
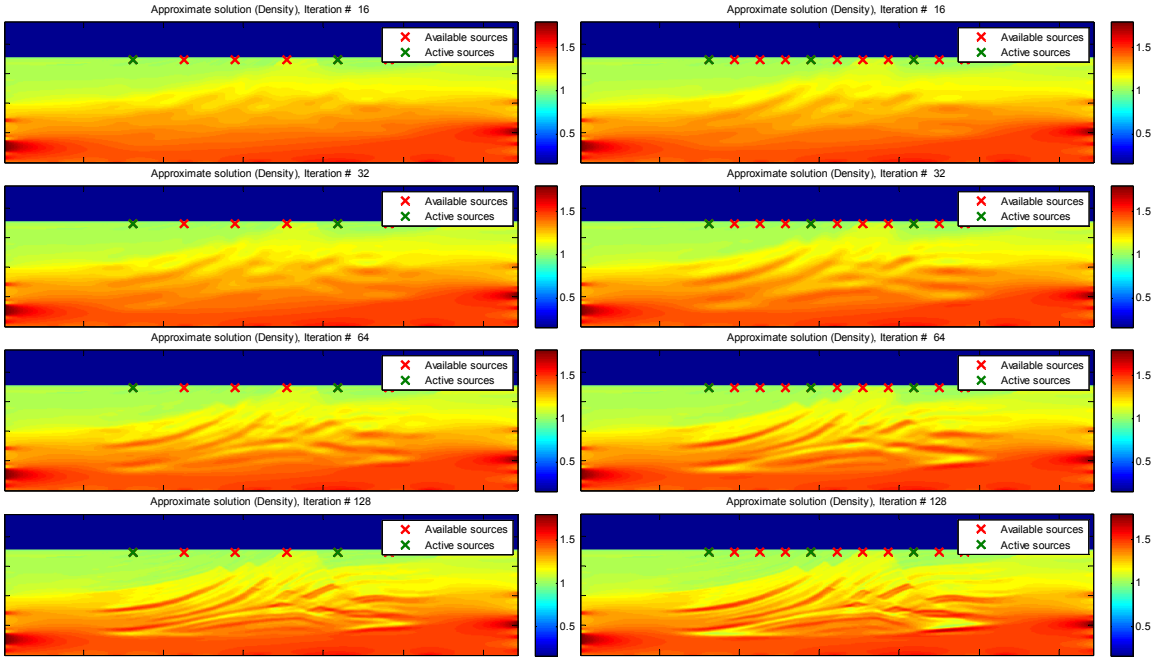


FIG 6. Approximate solution $\boldsymbol{\rho_n}$ convergence for full stack FWI with good initial guess.



a) b)

FIG 7. Approximate solution ρ_n convergence for partial stack FWI with good initial guess: a) – 6 sources in the stack; b) – 11 sources in the stack.

The second step is partial stack implementation in modelling conditions close to the full stack and the comparison of these two approaches. Based on the source switching strategy presented in Table 1, starting with the same initial approximation (Figure 6), we run partial stack FWI with 6 (Figure 7, a) and 11 (Figure 7, b) sources. With 11 sources

the distance between neighbour sources is in 2 times smaller than with 6 sources in the stack and also the coverage of the velocity field with partial stack is better when the same weight function σ (Figure 3) is used in both cases. As a result, near the surface, the convergence of the approximate solution obtained with high frequency filtering is the same for both partial stacks and for full stack. This means that while using high frequency filtering we don't need too many sources in the partial stack. At the same time, with low frequency filtering solving deeper layers FWI, the stack inversion convergence is much slower and requires either many more iterations, use of the full stack, or more precise model calibration which itself requires inside information about the exact solution. In this case, partial stack low frequency FWI may have no strong advantages over either full stack or shots gathering. Neither dynamic partial stack strategy not shots gathering was studied or compared one to another in here.

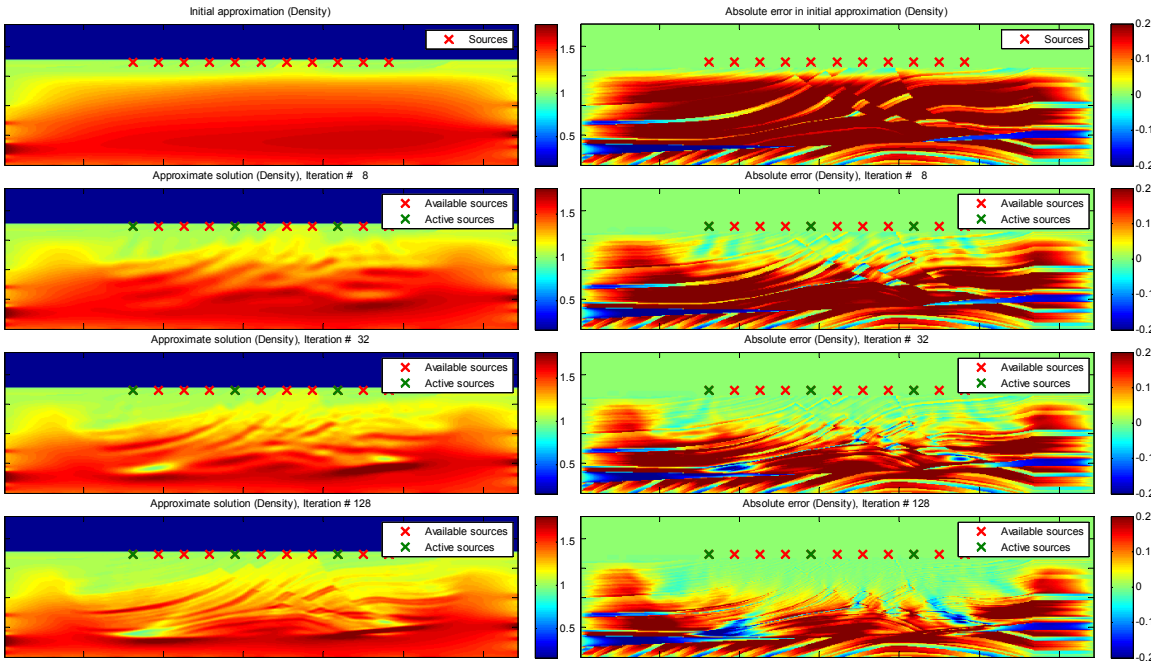


FIG 8. Convergence behaviour with poor initial guess for partial stack FWI with 11 sources: a) – approximate solution ρ_n; b) – absolute error $\rho_n - \rho_{exact}$ distribution.

For the next step, we study the 0-frequency error impact in the initial density guess on the partial stack FWI convergence. To begin, we smoothly increased the initial approximation of the density by 10%-15% in comparison with the previous experiment and ran the same partial stack with 11 sources (Figure 8) with the new initial approximation of the density which we call a "poor guess". The result is a more successful 0-frequency error cancellation on the surface and generally poor convergence in depth for any frequency component of error. This result fully matches the conventional understanding of the initial guess importance and its principal impact on the resulting FWI solution.

The last part of the density field partial stack FWI is devoted to noise impact brief study. We numerically test the hypothesis that the partial stack approach may benefit over the full stack FWI of comparable or even higher computational difficulty in case where the observation data from the geophones includes non-uniform noise.

Let us assume some high frequency noise with unknown location affected one shot and we don't know which shot is affected (Figure 9). Otherwise, knowing the location of the noise and the shot, we can potentially estimate the wavelet of the noise and then cancel it. We also assume that we don't contribute any other error to the existing finite difference approximation and numerical integration.

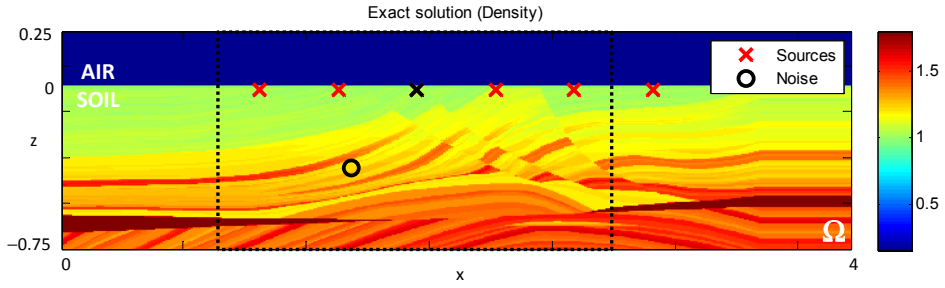
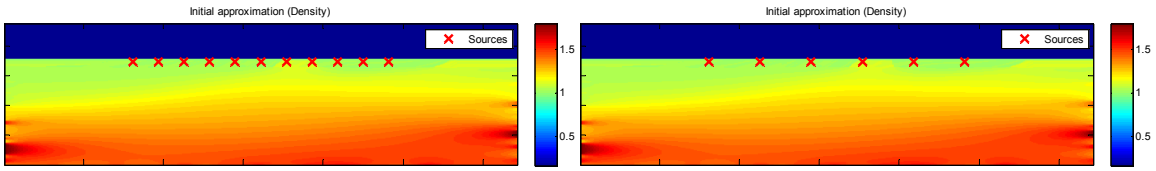


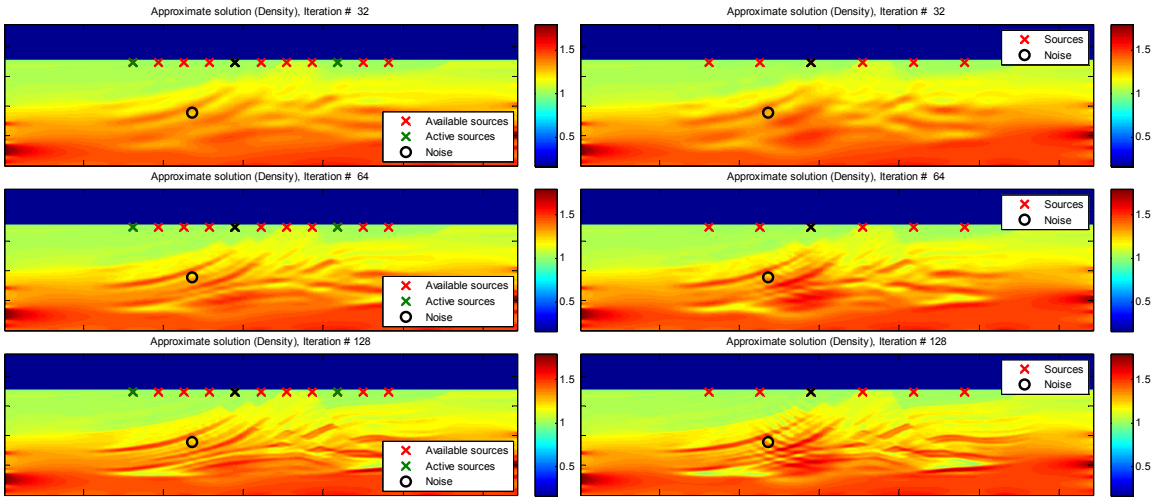
FIG 9. The position of periodical noise affecting black *x* source only.

Now we run the same Marmousi model for both the full and partial stack with one source affected with noise in each case. In the full stack we use 6 sources while in the partial stack we use 25% of 11 sources with switching. Consequently, in the partial stack we average our unknown defective shot with the extra 5 ones without noise and it runs two times faster than the full stack. The level of the noise is two times lower than the level of the misfit data Δv (Figure 2) for the corresponding source affected with it.

The results of the described numerical experiment with the noise impact are presented in the Figure 10. The specified noise has almost no effect on both stacks until the moment when filtering starts increasing the weight of the noise frequency in the misfit data masking. It appears on the iteration 64 for both stacks (Figure 10) simultaneously. Moreover, as we expected, partial stack successfully averaged the defective shot with additional neighbour shots not involved in the full stack in order to keep the comparable overall level of the model computational complexity for both stacks. As a result, the noise significantly affected the final full stack FWI solution while partial stack shows good accuracy.

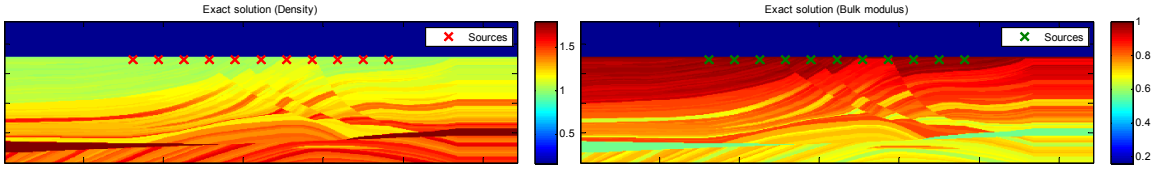
For bulk modulus the same partial stack FWI problem is reformulated. The convergence of the FWI in this case appears to behave similarly to density FWI and the only interest in it is explained with the aim to combine advantages of both in together FWI in the future studies.





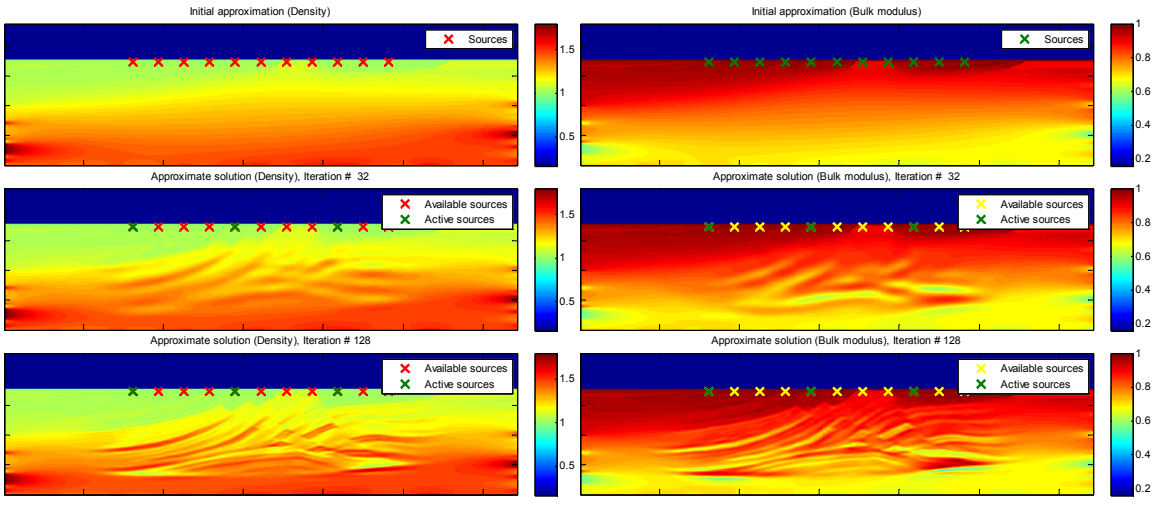
a) b)

FIG 10. Approximate solution ρ_n convergence behaviour with periodical noise affecting black x shot: a) – partial stack FWI for 11 sources; b) – full stack 6 shots FWI.



a) b)

FIG 11. Exact solution for partial stack FWI: a – density field FWI; b – bulk modulus FWI.



a) b)

FIG 12. Approximate solution: a – density partial stack FWI; b – bulk modulus partial stack FWI

A comparison of the results of both approaches is presented in Figures 11-13. The exact solution field in bulk modulus FWI are considered the following (in the layer of

soil only): λ_{exact} (Figure 11 a) is just inversed density in corresponding density FWI (Figure 12 a), $\mu_{exact} \equiv \lambda_{exact}$ and $\rho_{soil} \equiv 1$.

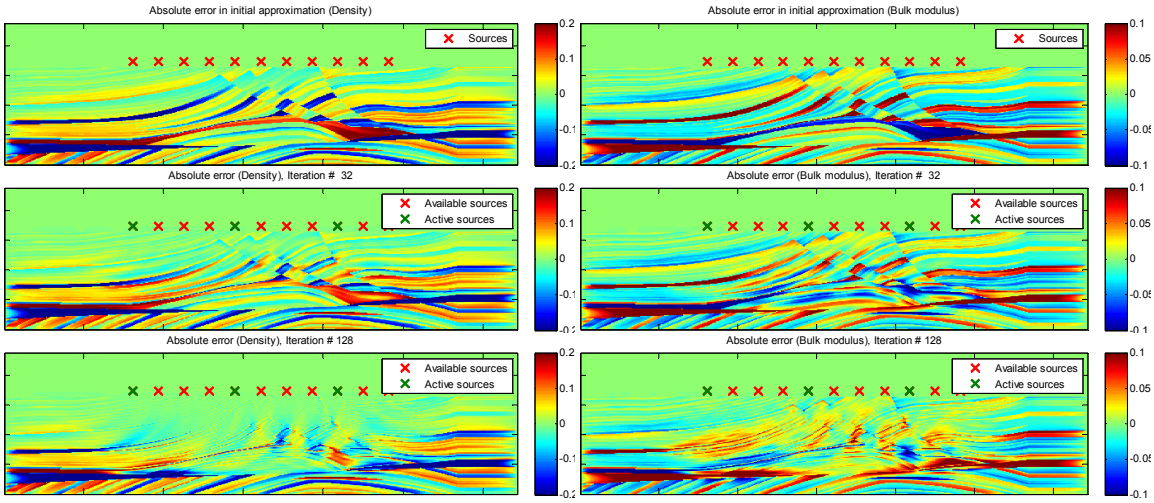


FIG 13. Absolute error cancellation results: a) – density partial stack FWI; b) – bulk modulus partial stack FWI.

In both Figures 12 and 13, at the 32nd iteration, the FWI result for bulk modulus affecting deeper layers are better than density FWI. At the same time, high frequency error cancellation is more successful in density FWI (Figure 13, iteration 128). The way the bulk modulus λ and μ appear in the PDE means that both fields are required to be differentiable. For this reason, it is hard to expect a great success in high frequencies FWI for λ in comparison to ρ. The last observation appears studying absolute error is vertical lines of opposite colors in 128th iteration of density FWI (Figure 13 a). It means that partial stack Newton search chooses opposite minimization directions for neighbour sources based on their imaging conditions. In this case, it could be useful to compare these absolute error fields with shot gathering technique applied to the same experimental conditions in order to compare the robustness of these two algorithms.

CONCLUSION

Several techniques optimizing the FWI algorithm appear to support fast and effective data processing in case of large stacks of shots. In the present work we used partial stack method in order to accelerate the stack convergence for both density and bulk modulus FWI.

The results obtained with the Marmousi density model (with preserved vertical resolution of the field) and partial stack 6 shots FWI (Figure 7 a) supports the hypothesis that the FWI, with either partial stack or shot gathering techniques, will be fast and numerically economical.

Other results obtained are summarized as follows:

- noise resistance behaviour of the partial stack with sufficient "good data" is able to smooth and defective shots via neighbouring shots, makes the partial stack preferable over the full stack;

- the partial stack method performs much better in the upper layers of the earth, with higher frequency filtering and could be used as a dynamic frequency dependent technique;

- Bulk modulus imaging has the advantage with low frequency error in an approximate solution over density FWI and their simultaneous use in one FWI could potentially benefit both inversions;

- partial stack allows independent imaging conditions for each shot which optimally match this particular shot and corresponding misfit function, in contrast to shot gathering and simultaneous shooting.

ACKNOWLEDGMENTS

We thank CREWES and its sponsors, NSERC, PIMS and Department of Mathematics and Statistics for their support of this work. We would also like to thank Joe Wong for his comment and suggestions.

REFERENCES

Hasanov A., Pektas B. and Erdem A. 2011, Comparative analysis of inverse coefficient problems for parabolic equations. Part I: adjoint problem approach, Inverse Problems in Science and Engineering, 19:5, 599-615

Hogan C. and Margrave G., 2006, Measurement of convergence in plane-wave migration: CREWES annual report 38, University of Calgary.

Margrave G., Lamoureux M. and Henley D. 2011, Gabor deconvolution: Estimating reflectivity by nonstationary deconvolution of seismic data: Geophysics 76, No. 3, W15-W30.

Pratt R.G., Shin C. and Hicks G.J. 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysics 133, 341-362.

Tarantola A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, 74, No. 8, 1259-1266.

Shin C., Pyun S. and Bednar J.B. 2007, Comparison of waveform inversion, part 1: conventional wavefields vs logarithmic wavefield: Geophysical Prospecting 55, 449-464

Virieux J. and Operto S. 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics 74, No. 6, WCC1-WCC26.