Efficient 3D Frequency-domain Full Waveform Inversion (FWI) with Phase Encoding

H. Ben-Hadj-Ali* (Géoazur/University of Nice Sophia-Antipolis), S. Operto (Géoazur/University of Nice Sophia-Antipolis) & J. Virieux (University Joseph Fourier)

SUMMARY

3D high resolution imaging methods such as full waveform inversion still suffer from their expensive computational cost, for instance because of the huge amount of data characterized by numerous shots and receivers. Performing shot assemblages or super-shots instead of individual shots may reduce the data manipulation necessary for image reconstruction. We may consider a coarse approach where we assemble sources at far distance and a cluster approach where we assemble sources at near distance. We apply simultaneous-shot and phase encoding techniques to 3D frequency-domain waveform inversion. While the imaging resolution is altered by non-physical components associated with interference noise due to shot assembling, we succeed in reducing significantly these artefacts by considering phase encoding technique. We have found that random phase encoding gives optimal results.
Introduction
Because 3D prestack depth imaging methods such as prestack depth migration and full waveform inversion have a high computational cost, the simultaneous-shot technique (Capdeville et al., 2005) can provide an interesting compromise between computational cost and imaging resolution. In fact, computation cost is reduced proportionally to the number of shots summed in each shot assemblage or super-shot. On the other hand, the image resolution is altered by non-physical artefacts associated with the interference among the individual shots of a given super-shot. These artefacts can be reduced by encoding specific phase for each shot in the super-shot. This technique is called phase encoding (PE) (Morton and Ober, 1998; Jing et al., 2000; Romero et al., 2000) and has been originally proposed for prestack migration. Since imaging kernels of prestack migration and frequency-domain FWI are quasi-similar, the technique can be used without any particular modification for FWI approach. We first introduce the basic principle of simultaneous-shot and more sophisticated phase encoding techniques. Then, we assess different strategies for phase encoding with an application of the FWI to a section of the SEG/EAGE overthrust model.

Theory
The inverse problem is an iterative process based on the minimisation of an objective function (least squares problem) (Pratt et al., 1998) defined as below,

$$ C^k (m) = \left| \left| d_o - d_c^k (m) \right| \right|^2 \quad (1) $$

where $m$ is the model, $d_o$ observed data and $d_c^k (m)$ modeled (computed) data at iteration $k$. The model perturbation, based upon the Born approximation and a steepest descent method, is given by

$$ \delta m^k = -\alpha^k \left( \text{diag} H^k + \gamma^k I d \right)^{-1} G^k \quad (2) $$

where $\text{diag} H^k$ is either the diagonal of approximate or pseudo Hessian (Shin et al., 2001), $G^k$ the gradient, $I d$ the identity matrix, $\alpha^k$ the descent value and the prewhitening factor $\gamma^k$ prevents numerical instabilities. The gradient is computed through the adjoint-state method (Plessix, 2006). In the frequency domain, it consists in cross-correlating shot with backpropagated residual wavefields, with a convoluting operator that consists in the radiation matrix $W$:

$$ G^k = -\sum_{\text{shots}} P^t W A^{-1} \left( d_o - d_c^k \right)^* \quad (3) $$

where $P$ is the shot wavefield and $A$ is the impedance matrix. $^t$ denotes transpose and $^*$ conjugate.

Simultaneous-shot technique
Reduce computational cost of seismic imaging techniques is achieved by processing a super-shot as a single shot and, therefore this reduction is proportional to the number of shots inserted into the super-shot. However, this gathering of shots introduces artefacts in the gradient and hessian estimations. These artefacts come from cross-talk effects between different shots of a super-shot. We illustrate our purpose with the gradient $G$ of a super-shot $S$ composed of 2 shots $S_1$ and $S_2$. $R_1$ and $R_2$ are the backpropagated residual wavefields associated with shots $S_1$ and $S_2$ respectively. Note that the same reasoning could apply for the hessian. The gradient can be schematically expressed as

$$ G = S W R^* = [S_1 W R_1^* + S_2 W R_2^*] + [S_1 W R_2^* + S_2 W R_1^*] \quad (4) $$

where ($S=S_1+S_2$; $R=R_1+R_2$) in virtue of the superposition principle.

The first bracketed term corresponds to the standard gradient formed by stacking the contribution of each individual shot while the second term corresponds to cross-talk noise between shots $S_1$ and $S_2$. This extra term alters the imaging result. Minimisation of related artefacts is achieved by the so-called PE technique.

Phase encoding technique
Encoding shots with an arbitrary weight when added into the same super-shot should reduce artefacts...
coming from interferences. Encoding weights $a_1$ and $a_2$ will be expressed with a phase term such that $|a_i| = |\exp(i\phi_i)| = 1$, $i = 1, 2$, and $(S = a_1 S_1 + a_2 S_2; R = a_1 R_1 + a_2 R_2)$. Equation (4) becomes
\[
G = [S_1 W R_1^* + S_2 W R_2^*] + [a_1 a_2^* S_1 W R_2^* + a_1^* a_2 S_2 W R_1^*].
\] (5)

One may notice that these phase terms act only on the cross-talk terms and the PE strategy should minimise the second bracketed term in equation (5) through judicious choice of phases $\phi_i$. Several phase encodings have been proposed such as deterministic PE (Jing et al., 2000) and random PE (Morton and Ober, 1998; Romero et al., 2000). Random PE (RPE) generates random phases in the interval $[0, 2\pi]$. During the summation process, the cross-talk noise tends to stack incoherently and, consequently, is minimised. Deterministic PE (DPE) assumes that shots in the super-shot are very close and expresses the phase $\phi^k$ in function of the previous phases $\phi^i, i = 1, k - 1$ as given by the following equation,

\[
\tan(\phi^k) = -\frac{\sum_{i=1}^{k-1} \cos(\phi^i)}{\sum_{i=1}^{k-1} \sin(\phi^i)}. \tag{6}
\]

An alternative PE (PPE) has been tested as well and consists in taking equidistant phases in the interval $[0, 2\pi]$, i.e. $\phi_i = 2(i-1)\pi/N$ where $N$ equals the number of shots in a super-shot.

**Numerical examples**

Two-dimensional experiments can be designed considering 2.5-D velocity models (laterally invariant in the $y$-direction) and an infinite line shot in the $y$ direction (Ben-Hadj-Ali et al., 2008) in order to simplify our image analysis although we consider 3D numerical simulations. We apply 3D FWI to a dip section extracted from the SEG/EAGE overthrust velocity model (Figure 1(a)), discretized on a $801 \times 187$ grid with a grid spacing $h = 25$ m. For the 3D application, the dip section of the overthrust model is duplicated 3 times in the $y$ direction leading to a $3D 801 \times 3 \times 187$ finite-difference grid. PML absorbing boundary conditions are set on the 4 edges of the 2D model while periodic conditions are implemented in the $y$ direction to mimic an infinite medium. The starting model for inversion is obtained by smoothing the true velocity model with a Gaussian function of horizontal and vertical correlation lengths of 500 meters (Figure 1(b)). For a stable inversion, the true velocity structure is set in the first 100 meters of the starting model. We invert sequentially 7 frequencies ranging from 5 to 20 Hz. For each frequency, we compute 15 iterations. The 2D acquisition geometry consists of a line of 200 shots and receivers equally-spaced on the surface. The final FWI model obtained without shot assembling is shown in Figure 1(c) as reference result. Twenty-five super-shots of 8 shots each are inverted. We build super-shots by summing either 8 near (cluster super-shot) or distant shots (coarse super-shot).

Results obtained with the cluster super-shot approach are outlined in Figure 2. The FWI model obtained with shot assembling but without PE is shown in Figure 2(a). The footprint (cross-talk artefacts) when considering super-shots is clearly visible especially in the shallower part of the model whose imaging is more sensitive to coarse spacing between shots. The FWI models obtained with the three PE strategies, i.e., RPE, DPE and PPE, are shown in Figures 2(b-d) respectively. The three PE techniques succeed to minimise the artefacts in the shallow part of the model. The RPE approach provides the best overall image. This is further confirmed by the L2-norm between the FWI and the true models.

For the coarse super-shot approach, RPE is clearly the best choice of encoding (Table 1). DPE gives the worst result since it violates the basic assumption that requires near shots in a super-shot. Figure 3 illustrates the result obtained when PE technique is not activated in (a) and with RPE in (b).

With appropriate PE strategy, i.e. RPE, cluster super-shot approach gives a slightly better result than the coarse approach, especially in the shallow part of the model.

Super-shots simulations last nearly 7 times less than the simulation without shot gathering. This is consistent with theory since 8 times less forward problems, which is the most intensive part of the CPU time, have been performed.

**Conclusion**

Data reduction using super-shots or shot assemblages is an efficient way for a significant reduction of computational cost of 3D full waveform inversion. In fact, the speed-up is nearly the number of shots in the super-shot. Cross-talk noise associated with interference components could be attenuated using phase...
encoding techniques. Many encodings are relevant and lead to good results. We found that a random phase encoding is the most appropriate one either for the cluster or coarse super-shot techniques.

Acknowledgements
Access to the high performance computing facilities of MESOCENTRE SIGAMM computer center provided the required computer resources and we gratefully acknowledge both this facility and the support of the staff. Finally, this work was carried out within the frame of the SEISCOPE consortium (http://seiscope.unice.fr) sponsored by BP, CGG-Veritas, EXXON-Mobil, SHELL and TOTAL.

References


Table 1: Normalised L2-norm residuals between different FWI and true models.

<table>
<thead>
<tr>
<th></th>
<th>No − PE</th>
<th>DPE</th>
<th>RPE</th>
<th>PPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without shot assembling</td>
<td>3.76E-02</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Cluster approach</td>
<td>5.01E-02</td>
<td>4.66E-02</td>
<td>4.47E-02</td>
<td>4.93E-02</td>
</tr>
<tr>
<td>Coarse approach</td>
<td>5.99E-02</td>
<td>5.86E-02</td>
<td>4.45E-02</td>
<td>5.75E-02</td>
</tr>
</tbody>
</table>

Figure 2: Cluster approach: a) Super-shot no-PE FWI velocity model. b) Super-shot DPE FWI velocity model. c) Super-shot RPE FWI velocity model. d) Super-shot PPE FWI velocity model. e) Comparison between dip-profiles extracted at Inline=5 km (thrust target). f) Comparison between dip-profiles extracted at Inline=13.5 km (channel target). (True model: dotted black line; FWI model: red line; Super-shot no-PE FWI model: green line; Super-shot RPE FWI model: blue line)

Figure 3: Coarse approach: a) Super-shot no-PE FWI velocity model. b) Super-shot RPE FWI velocity model.