Automatic Wavefield Reconstruction Inversion

Nuno V. da Silva, Imperial College London*

SUMMARY

Wavefield Reconstruction Inversion (WRI) is a method recently introduced by van Leeuwen and Herrmann (2013), that aims to circumvent convergence towards local minima, to which Full Waveform Inversion (FWI), as conventionally implemented, is not immune. The novel aspect of WRI is that it fits the data while relaxing the wave equation constraint, and does not need the computation of an adjoint wavefield. This last aspect, is of particular interest, and represents a substantial saving in terms of computing cost associated to the solution of the inverse problem.

Nonetheless, as WRI is formulated, it requires the estimation of a trade-off parameter. If this is too large, then it implies penalizing for the wave equation and imposing a hard constraint in the inversion. If it is too small, then this leads to a poorly constrained solution as it is essentially penalizing for the data misfit. Here I introduce a new approach, in the context of WRI, which allows the estimation of this parameter automatically based upon a multiplicative cost function.

I present a synthetic example demonstrating that the proposed method allows the convergence towards a correct solution, and that is equivalent to use the additive cost function, however without any assumption or test performed *a priori* in order to estimate a sensible magnitude for this parameter. This is the trade-off is estimated automatically. The method presented here does not add any noticeable computing cost to the inversion scheme.

INTRODUCTION

Full Waveform Inversion (FWI) has been gaining considerable attention from the oil and gas industry. Its popularity is due to the fact that it promises to dramatically change the conventional workflow for identifying potential hydrocarbon prospects.

The concept of FWI, as it is most widely known, has been introduced three decades ago (Tarantola, 1984). If increasing computational power led to some practical results on one hand (Sirgue *et al.*, 2009; Warner *et al.*, 2013), on the other, some theoretical aspects have been hindering a full accomplishment of the technique. These aspects are essentially related to the classical issues that characterize inverse problems, i.e., non-uniqueness of the solution, non-guarantee of existence of a solution and sensitive dependency to the initial conditions.

Because FWI is intrinsically a very large-scale inversion problem, the most favorable techniques are linearized inversion schemes. These, however, are not guaranteed to converge to a global minimum. Instead, local methods, can potentially converge towards local minima, hindering the achievement of an optimal solution. One such expressions of local minima is the existence of cycle-skipping, this is the observed and simulated data are shifted by more than half a cycle.

Avoiding local minima has been one of the most influential topics of research in the scope of FWI. It was early acknowledged that local minima could be avoided by inverting data in a multi-scale fashion, this is, proceeding from lower to higher frequencies (Bunks *et al.*, 1994, Sirgue and Pratt, 2004). This approach, aims to correct the larger wavelengths of model perturbation and gradually increases the resolution of the model.

Even though inverting in a multi-scale fashion has been very effective in tackling the problem of converging towards local minima, it still does not guarantee convergence as there are many factors affecting the inversion. Examples of such factors are for example the quality of the starting model and spectral content in the data. For this reason, other techniques which can be potentially more robust have been developed. For example Ma and Hale (2013) applied dynamic warping, Baek *et al.* (2014) guided the inversion with registration, and Warner and Guasch (2014) implemented a new technique based upon adapting matching filters, also known as Adaptive Waveform Inversion (AWI).

A detailed analysis of the problem of the local minima leads to the conclusion that this is a consequence of the very non-linear relation between the model parameters and the data in the most classical formulation of FWI. This is because FWI is formulated as a data misfit problem only, subject to the solution of a wave equation at each iteration.

van Leeuwen and Herrmann (2013) proposed an alternative approach, where the seismic inverse problem is reformulated in terms of minimizing the data misfit subject to the constraint that the wavefield computed at the entire inversion domain, for a given model, must satisfy the relation between the wave operator, acting on that same wavefield, and the source term. Because this approach enlarges the search space by also penalizing with a wave equation term it is denominated wavefield reconstruction inversion.

Another aspect of the extended space search is that it introduces a parameter linking the data misfit to the constraining term. This parameter balances the contribution of each term to the objective function, nonetheless, it is estimated heuristically, requiring some degree of interaction. The question that naturally raises is if there is an approach for determining this trade-off parameter automatically. Here I present an automatic approach for determining this parameter which does not require any *a priori* information nor assumption on its magnitude, being estimated adaptively throughout the iterations by recasting the additive cost function into a multiplicative cost function (van den Berg *et al.*, 2003).

This paper is structured as follows. First the theory on WRI is reviewed and the minimization problem is recast into a new objective function in the context of the extended space inver-

DOI http://dx.doi.org/10.1190/segam2015-5908659.1

sion. Then a synthetic example is presented demonstrating the feasibility of the proposed approach.

THEORY

I start by reviewing the penalty method as introduced by van Leeuwen and Herrmann (2013). WRI corresponds to minimizing a data functional term constrained by a wave equation:

$$\min_{\mathbf{m},\mathbf{p}} J = \sum_k \frac{1}{2}\|R_k\mathbf{p}_k - \mathbf{d}_k\|_2^2 + \frac{1}{2}\lambda^2\|\mathscr{L}\mathbf{p}_k - \mathbf{s}_k\|_2^2 \quad (1)$$

Where R_k is a restriction operator, $\mathbf{p}_k$ is the acoustic pressure in the modeling domain, $\mathbf{d}_k$ are the measured data, $\mathbf{s}_k$ is the source, $\mathscr{L}$ is a wave operator and λ is the trade-off parameter. The subscript k denotes the number of the source, and the sum is performed over all sources. The wave operator in this case is the Helmholtz equation for acoustic waves, which in its discretized form is given by:

$$\mathscr{L}(\omega,\mathbf{m}) = \nabla_h^2 + \omega^2 \mathbb{1}\mathbf{m} \quad (2)$$

Where ∇_h^2 is a discretized Laplacian operator, ω is the angular frequency, $\mathbb{1}$ is the identity matrix, and $\mathbf{m}$ represents the squared slowness in the grid. The Laplacian operator is discretized with a nine point stencil. The pressure field is a function of frequency, however this is not explicitly represented in the notation.

The minimization of equation 1 leads to the augmented system

$$\begin{pmatrix} \lambda\mathscr{L}(\omega,\mathbf{m}) \\ R_k \end{pmatrix} \mathbf{p} = \begin{pmatrix} \lambda\mathbf{s}_k \\ \mathbf{d}_k \end{pmatrix} \quad (3)$$

which is over-determined and sparse. Its solution is obtained in a least squares sense. As one can immediately identify the trade-off parameter λ is crucial for the solution of the wavefield that satisfies the penalty expression 1. If it is too large the optimization solution will penalize for the wave equation and the iterative scheme behaves as conventional FWI. If it is too small the expression will penalize for the data misfit. Thus determining a suitable trade-off is of paramount importance, and may require several trials when using the additive objective function 1.

An obvious approach in order to determine λ, is implementing a parameter estimation scheme as the one used in Occam's inversion (de Groot−Hedlin and S. Constable, 1990). In such case the optimal parameter requires the solution of the inverse problem for a set of suitable λ's, and the optimal solution is the one that leads to the smallest norm of the data misfit and the smallest norm of the penalty term together. However such an approach is computationally expensive and would be impractical for 3D problems.
Instead I recast the optimization problem in a multiplicative objective function (van den Berg *et al.*, 2003)

$$\min_{\mathbf{m},\mathbf{p}} J = J_d(\mathbf{p})J_e(\mathbf{p},\mathbf{m}) \quad (4)$$

Where $J_d(\mathbf{p})$ is a measure of data misfit and $J_e(\mathbf{p},\mathbf{m})$ is the penalty term (wave equation constraint). The multiplicative objective function 4 is equivalent to the additive objective function 1 when $J_e(\mathbf{p},\mathbf{m}) = 1$ and the trade-off parameter $\lambda = J_d/J_e$. In such case the terms of the cost function are given by

$$J_d = \sum_k J_{d,k} = \sum_k \|R_k\mathbf{p}_k - \mathbf{d}_k\|_2^2 \quad (5)$$

and

$$J_e(\mathbf{p}_k,\mathbf{m}) = J_e(\mathscr{L}\mathbf{p}_k,\mathbf{s}_k) = 1 \quad (6)$$

Thus the trade-off parameter is determined by J_d since $J_e = 1$. This approach eliminates λ as an unknown allowing for an automatic approach for minimizing the cost function of WRI. Because the objective functions are equivalent, this approach also requires the solution of the augmented system 3 and the gradient is computed in a similar fashion as in the case of the additive cost function, yielding

$$\nabla_m J = \Re\sum_k \lambda^2\omega_k^2 \mathbb{1}\mathbf{p}_k^*[\mathscr{L}(\omega,\mathbf{m})\mathbf{p}_k - \mathbf{s}_k] \quad (7)$$

where the star in superscript stands for complex conjugation, and the operator $\Re$ extracts the real part of a complex number. At the minimum the condition $\nabla_m J = 0$ is met.
The Hessian is diagonal making the pre-conditioning of the gradient of the objective function trivial, and is given by

$$\nabla_m^2 J = \Re\sum_k \lambda^2\omega^4\mathbf{p}_k^*\mathbb{1}\mathbf{p}_k \quad (8)$$

Hence the multiplicative cost function leads to the same set or equations for updating the velocity model as in the case of the additive cost function, with the advantage of eliminating the trade-off parameter as an unknown in the minimization problem. The model updates are performed with a steeped-descent where the gradient is scaled by the inverse of the Hessian.

SYNTHETIC EXAMPLE

In this section is presented a synthetic example showing the effectiveness of the proposed approach. The synthetic data is generated using the BP 2004 model, depicted in figure 1 (Billette and Brandsberg−Dahl, 2005). In this example it is considered a constant density background, hence all the phase and amplitude events in the data are related with velocity perturbations only. The velocity model is discretized with a 20 m grid spacing and the data is generated for frequencies in the range 1.5 to 9.0 Hz. The acquisition geometry is a fixed receiver configuration with the sources and receivers placed at 10 m of depth. The data is generated for 500 sources separated by 20 m and 1000 receivers spaced of 20 m. The inversions are carried out in a multi-scale fashion, using the starting model displayed in figure 2 and iterating 16 times at each chosen frequency. The data is generated and inverted for 14 frequencies.

First I exemplify how the wrong choice of the parameter can affect the inversion. The first trial is performed picking a value of $\lambda = 10^4$ and iterating 16 times at 1.5 Hz. Figure 3 shows the resulting inverted model. A second trial is performed for $\lambda = 1$, also at 1.5 Hz and iterating the same number of times. The corresponding result is depicted in figure 4. As one can

DOI http://dx.doi.org/10.1190/segam2015-5908659.1

immediately identify the result of the first trial is very similar to the starting model whereas in the second trial there is some structure being built in the model, suggesting that convergence is occurring. Even though not being showed here, the multiplicative cost function gets a similar result to the the one showed in figure 4. This series of trials suggests that $\lambda = 1$ is a sensible parameter for carrying out the inversion when it is held fixed.

Figure 5 shows the inverted model obtained using all the frequencies and using $\lambda = 1$, which is held fixed throughout the inversion. Figure 6 shows the inverted model obtained using the multiplicative cost function. In this case, λ is estimated automatically changing in each iteration, and its value decreases monotonically with increasing number of iteration, for the same frequency (figure 7).

Comparing figures 5 and 6, one immediately concludes that the solutions are similar. However, in the case of the additive cost function, it was necessary to determine what value is sensible for λ, which works in this example, but may not work with a different model.

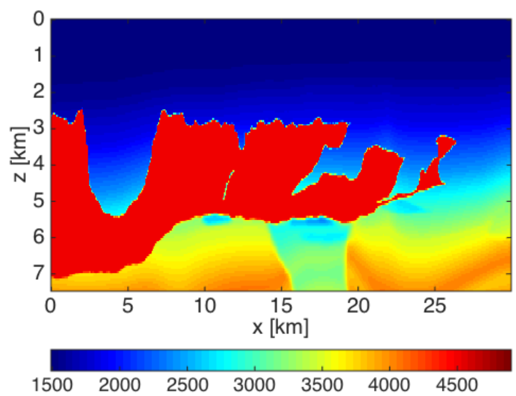

Figure 1: Section of the BP 2004 model.

As illustrated in figure 8 the data residual also decreases faster in the case of using an automatic trade-off parameter. This is because in this case it is determined the best parameter that fits both the model parameters and the wave equation constraint optimally. Thus the entire fit is improved in comparison to use a trade-off held fix. Note that in figure 8 it is only plotted the evolution of data misfit term with increasing number of iteration for the last block of iterations (at 9 Hz). Even though not illustrated here, the data misfit term of the multiplicative cost function decreases faster than the corresponding term in the additive cost function, for all the other frequencies inverted.

Also worth noting, is that in the case of the multiplicative cost function, the value of λ decreases with increasing number of iteration as expected since the multiplicative cost function is equivalent to the additive one, and is expected that all values decrease to zero.

This example demonstrates that the automatic estimation of the trade-off parameter for the WRI solution is an effective approach for determining the solution of the inverse problem in

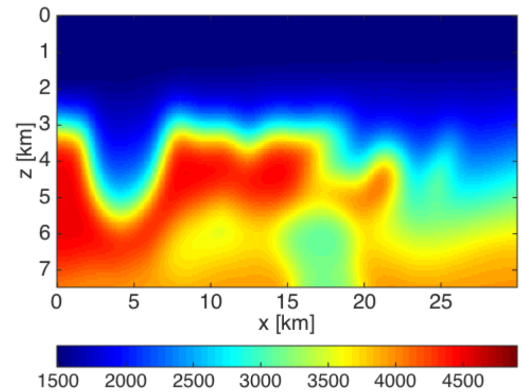

Figure 2: Starting model, obtained after smoothing the model displayed in figure 1.

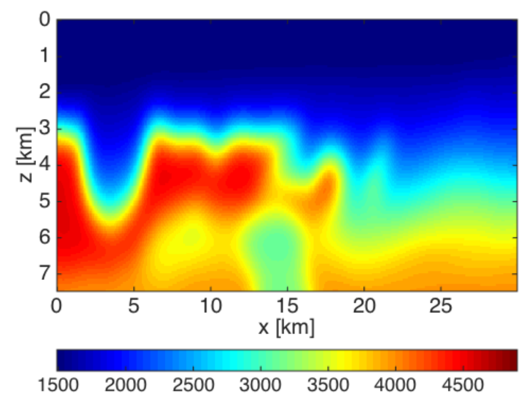

Figure 3: Inverted model after 16 iterations at 1.5 Hz for $\lambda = 10^4$.

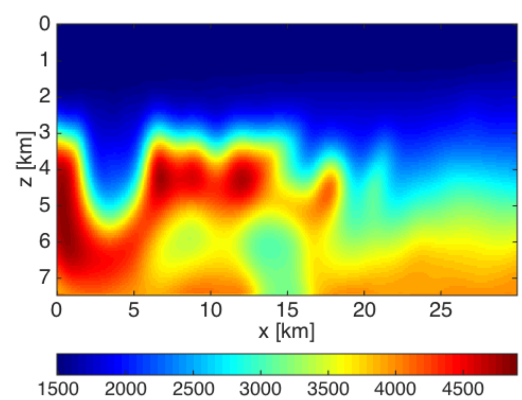

Figure 4: Inverted model after 16 iterations at 1.5 Hz for $\lambda = 1$.

DOI http://dx.doi.org/10.1190/segam2015-5908659.1

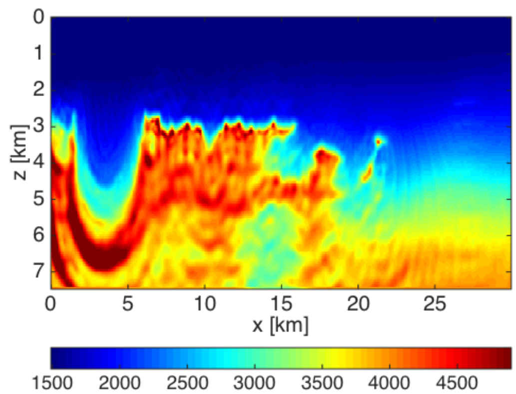

Figure 5: Inverted model using $\lambda = 1$, which is fixed throughout the inversion.

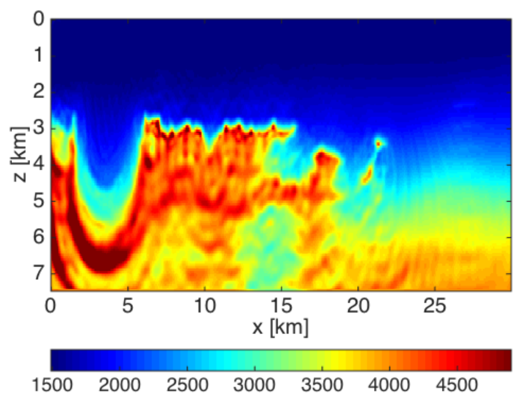

Figure 6: Inverted model using the multiplicative cost function.

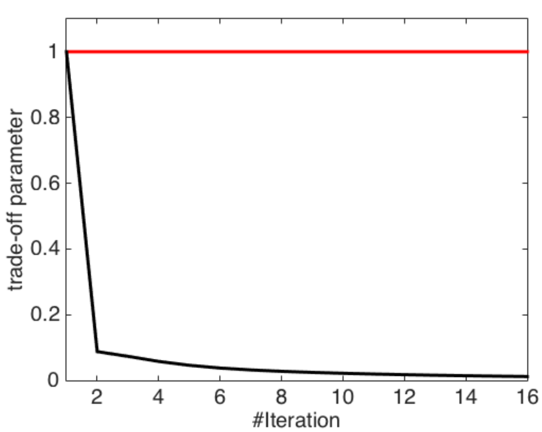

Figure 7: Trade-off parameter with the number of iteration for the multiplicative cost function (black line) and additive cost function (red line).

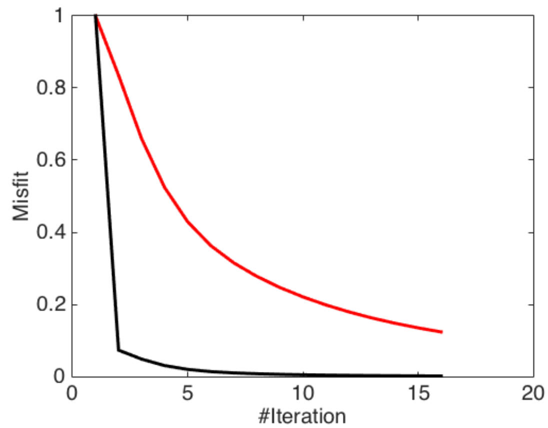

Figure 8: Dependency of the normalized data misfit with the number of iterations for the case of the multiplicative cost function (black line) and additive cost function (red line).

the augmented search space, without the element of subjectivity associated to the use of a purely additive cost function.

CONCLUSIONS

I presented an approach for estimating automatically the trade-off parameter associated to the solution of the optimization problem in WRI. The proposed method uses a multiplicative cost function and is new in this scope. As demonstrated, this method eliminates the trade-off parameter as an unknown in the optimization problem, suppressing any heuristic approach for determining it prior to running the inversion.

The automatic estimation of the trade-off parameter, as presented here, does not have any noticeable impact in terms of computational cost. The synthetic data example demonstrated that the automatic parameter conducted to the smallest data misfit and converged towards a sensible velocity model, demonstrating that this approach is feasible for WRI.
Future research will focus on aspects related to solving the augmented system efficiently, using second order methods, as well as, applications to real data and elastic inversion.

ACKNOWLEDGMENTS

The author is thankful to the sponsors of the FULLWAVE Game Changer research consortium for supporting this research work.

DOI http://dx.doi.org/10.1190/segam2015-5908659.1

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2015 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Baek, H., H. Calandra, and L. Demanet, 2014, Velocity estimation via registration-guided least-squares inversion: Velocity estimation via registration-guided least-squares inversion: Geophysics, **79**, no. 2, R79–R89. http://dx.doi.org/10.1190/geo2013-0146.1.

Billette, F. J., and S. Brandsberg-Dahl, 2005, The 2004 BP velocity benchmark: 67th Annual Internationa Conference and Exhibition, EAGE, Expanded Abstracts, B305.

Bunks, C., F. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophyscis, **60**, 1457–1473. http://dx.doi.org/10.1190/1.1443880.

deGroot-Hedlin, C., S. Constable, and the de Groot-Hedlin and S. Constable, 1990, Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data: Geophysics, **55**, 1613–1624. http://dx.doi.org/10.1190/1.1442813.

Ma, Y., and D. Hale, 2013, Wave-equation reflection traveltime inversion with dynamic warping and full-waveform inversion: Geophysics, **78**, no. 6, R223–R233. http://dx.doi.org/10.1190/geo2013-0004.1.

Sirgue, L., and G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248. http://dx.doi.org/10.1190/1.1649391.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266. http://dx.doi.org/10.1190/1.1441754.

van den Berg, P. M., A. Abubakar, and J. T. Fokkema, 2003, Multiplicative regularization for contrast profile inversion: Radio Science, **38**, no. 2, 8022. http://dx.doi.org/10.1029/2001RS002555.

van Leeuwen, T., and F. J. Herrmann, 2013, Mitigating local minima in full-waveform inversion by expanding the search space: Geophysical Journal International, **195**, no. 1, 661–667. http://dx.doi.org/10.1093/gji/ggt258.

Warner, M., and L. Guasch, 2014, Adaptive waveform inversion — FWI without cycle skipping — Theory: Presented at the 76th Annual International Conference and Exhibition, EAGE. http://dx.doi.org/10.3997/2214-4609.20141092.

Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Štekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: Geophysics, **78**, no. 2, R59–R80. http://dx.doi.org/10.1190/geo2012-0338.1.