Combining global optimization and boundary integral methods to robustly estimate subsurface velocity models

Gregory Ely, Massachusetts Institute of Technology, Alison Malcolm, Memorial University of Newfoundland, David Nicholls, University of Illinois at Chicago

SUMMARY

In this paper, we combine a fast wave equation solver using boundary integral methods with a global optimization method, namely Particle Swarm Optimization (PSO), to estimate an initial velocity model. Unlike finite difference methods that discretize the model space into pixels or voxels, our forward solver achieves significant computational savings by constraining the model space to a layered model with perturbations. The speed and reduced model space of the forward solve allows us to use global optimization methods that typically require numerous evaluations and few unknown variables. Our technique does not require an initial guess of a velocity model and is robust to local minima, unlike gradient descent frequently used in methods for both initial velocity model estimation and full waveform inversion. We apply our inversion algorithm to several synthetic data sets and demonstrate how prior information can be used to greatly improve the inversion.

INTRODUCTION

Most seismic processing techniques rely on an accurate velocity model to obtain meaningful results. Incorrect velocity models can hamper processing and lead to erroneous interpretation of seismic data. In particular, inversion techniques such as Full Waveform Inversion (FWI) are highly sensitive to the initial velocity. Without a good initial velocity estimate, FWI will converge to local minima with artifacts (Virieux and Operto, 2009).

Unlike global seismology, exploration and regional scale velocity models can be poorly constrained. Initial velocity models can be built using a variety of techniques such as travel time tomography, NMO semblance analysis, and even full waveform inversion at very low frequencies (Woodward et al., 2008). Most methods of constructing an initial velocity model rely on expensive travel-time or wave equation solvers and use gradient based approaches that are susceptible to local minima. Furthermore, travel time tomography does not use the full wavefield and requires the picking of arrivals as well as long offsets. Frequently noise levels are too high for full waveform inversion to be performed at frequencies where the problem is sufficiently convex and thus it can be plagued by local minima Pratt (1999); Sirgue (2006).

To combat the problem of local minima in geophysical inverse problems, researchers have applied global optimization techniques that are less susceptible to local minima (Sambridge and Mosegaard, 2002; Sen and Stoffa, 2013). Researchers also make use of different objective functions and regularization techniques (Burstedde and Ghattas, 2009; van Leeuwen and Herrmann, 2013) to make the inversion more convex, mitigating the need for global solvers.

In this paper we integrate a fast Helmholtz solver with the global optimization method Particle Swarm Optimization (PSO) to invert a velocity model without an initial model. We first give an overview of the field expansion method we use for quickly solving the Helmholtz equation and modify it to reduce artifacts and accurately simulate exploration scale data. We then describe two variations of the PSO algorithm for inverting a velocity model from an observed data set. Finally, we present the results of our inversion algorithm on synthetic data and demonstrate how prior information about the velocity model can be used to improve the inversion.

FIELD EXPANSION

In this section we briefly describe the field expansion method to solve the Helmholtz equation for a perturbed layer medium. Similar to some reflectivity codes, such as that by Bouchon (2003), this solver operates in the spatial Fourier domain and as a result describes the response of an infinitely repeating domain subject to a periodic point source, as shown in Figure 1. A detailed description and derivation of the solver can be found in Malcolm and Nicholls (2011).

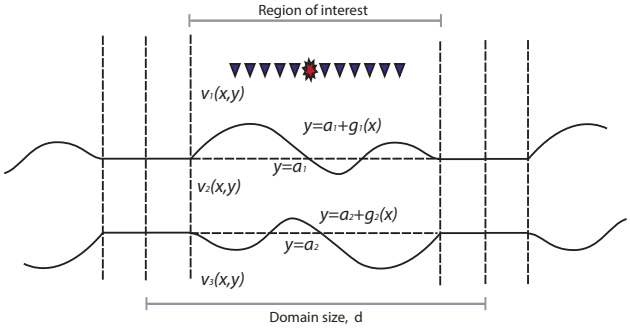

Figure 1: An illustration of the repeating domain structure implicit in the forward solver. The domain is periodic and there is a point source at the same location in each of the repeating domains.

The forward solver describes a velocity model consisting of M layers with velocities c_m. Each layer is divided by an interface at depth a_m; this interface is perturbed from flat by the function $g_m(x)$, which has a zero mean. This is illustrated in Figure 1. The scattered field in each layer $v_m(x,y)$ satisfies the acoustic Helmholtz equation

$$\nabla^2 v_m + k_m v_m = 0 \qquad k_m = \frac{2\pi f}{c_m}, \tag{1}$$

at some frequency f. We begin by describing the case for a set of M flat layers and then expand the forward solver to handle perturbations of the interface shape.

Flat Layers

In the flat layer case we set the perturbation $g_m(x) = 0$ and apply the outgoing wave condition (OWC) to the velocity model. This requires that in the uppermost layer, energy must be propagating upwards away from the model and in the bottom layer, energy must be propagating away from the model downwards. Subject to the OWC and the Helmholtz equation the scattered field, v_m in the spatial frequency domain indexed by modes p has the form,

$$v_m(x,y) = \sum_{p=-\infty}^{\infty} d_{p_m} e^{i(\alpha_p - \beta_{p_m}(y-\bar{a}_m))} + u_{p_m} e^{i(\alpha_p + \beta_{p_m}(y-\bar{a}_m))} \tag{2}$$

where

$$\alpha_p = \alpha + (2\pi/d)p \qquad \beta_{j,p} = \begin{array}{ll} \sqrt{k_m^2 - \alpha_p^2} & \alpha_p^2 < k_m^2 \\ i\sqrt{\alpha_p^2 - k_m^2} & \alpha_p^2 > k_m^2 \end{array} \tag{3}$$

and α is a constant determined by the angle of the incident source, $\bar{a}_m$ is the midpoint of each of the flat layers and d is the domain spacing (Petit, 1980). For the shallowest and deepest layers, the midpoint is equal to the closest interface, $\bar{a}_0 = a_1$ and $\bar{a}_M = a_M$. Note that in Equation 2 the up-going u_{p_m} and down-going d_{p_m} spatial Fourier coefficients completely describe the field as all of the other variables are determined directly from the layered model. To determine these unknown coefficients, we impose the Dirichlet and Neumann boundary

DOI http://dx.doi.org/10.1190/segam2015-5918749.1

conditions

$$\begin{aligned} v_{m-1} - v_m &= \xi_m \quad at \quad y = a_m \\ \partial_{N_m} v_{m-1} - \partial_{N_m} v_m &= \psi_m \quad at \quad y = a_m \end{aligned} \tag{4}$$

at each layer at depth a_m. In Equation 4, ξ and ψ are the incident field and the normal derivative of the incident field respectively. Combining the expression for the scattered field, Equation 2, and the boundary conditions, we can express the equalities given by the Equation 4 as a system of linear equations for each spatial mode p. We write this system as

$$\mathscr{A}_p \vec{z_p} = \vec{r_p}, \tag{5}$$

where $\mathscr{A}_p$ is a penta-diagonal matrix with entries completely determined by the velocity model,

$$\vec{z}_p = (u_0, d_1, u_1, ...d_{M-1}, u_{M-1}, d_M)^T \tag{6}$$

is the unknown vector of up-going and down-going coefficients, and

$$\vec{r}_p = (\xi, \psi..., 0, 0)^T \tag{7}$$

is the source vector. Because the spatial modes are orthogonal, the field coefficients u_p and d_p can be calculated independently from each other mode by mode. Once calculated, the field can be determined by summing across all calculated modes using Equation 2. The entries of $\mathscr{A}_p$ are given in Malcolm and Nicholls (2011) and follow from the insertion of Equation 2 into Equation 4. To insert a point source into this formulation, we specify the incident field for a periodic point source in the uppermost layer at some location (x_0, y_0). This periodic point source infinity repeats in the x direction at an interval of d. For this field, the boundary conditions at a single spatial mode p are

$$\xi_p(x,y) = \frac{1}{2id} \frac{e^{i(\alpha_p(x-x_0)+\beta_p|y-y_0|)}}{\beta_p} \qquad \psi_p = \partial_{N_m} \xi_p. \tag{8}$$

Note that because the matrix $\mathscr{A}_p$ is penta-diagonal with size on the order of the number of layers M, we can calculate the field very rapidly in $O(M)$ time. In addition, its LU decomposition is tridiagonal and extremely efficient to store for multiple right hand solves. Like other Helmholtz solvers, the field calculation for different source locations only requires only the application of the already computed $\mathscr{A}_p^{-1}$ to a new right hand side.

Perturbed Layers

To extend the forward solver to handle a non-trivial interface, we make the approximation that the zero-mean layer perturbation $g_m(x) = \varepsilon f_m(x)$ and take the Taylor series of the field about ε. Taking the Taylor expansion up to some order N, the field is

$$v_m(x,y) = \sum_{p=-\infty}^{\infty} \sum_{n=0}^{N} (d_{(p_m,n)} e^{i(\alpha_p - \beta_{p_m}(y-\bar{a}_m))} + u_{(p_m,n)} e^{i(\alpha_p + \beta_{p_m}(y-\bar{a}_m))}) \varepsilon^n. \tag{9}$$

Like the expression for the flat layer field, $d_{(p_m,n)}$ and $u_{(p_m,n)}$ completely determine the field. In addition, the Dirichlet and Neumann boundary conditions are applied to the perturbed interface, changing the boundary conditions to

$$\begin{aligned} v_{m-1} - v_m &= \xi_m \quad at \quad y = a_m + g_m(x) \\ \partial_{N_m} v_{m-1} - \partial_{N_m} v_m &= \psi_m \quad at \quad y = a_m + g_m(x). \end{aligned} \tag{10}$$

This expansion about ε can be applied to the system of equations in Equation 10 as done in Malcolm and Nicholls (2011). When expanded, a recursion relationship becomes apparent such that field coefficients of order n are determined by the lower order terms $0...(n-1)$ and each higher order coefficient can be determined by solving

$$\mathscr{A}_p \vec{z}_{(p,n)} = \vec{R}_{(p,n)} \tag{11}$$

for a different right hand side $\vec{R}_{(p,n)}$ in Equation 11. For the zeroth order term, $\vec{R}_{(p,0)}$ is equal to the right hand side of the flat layer case given by Equation 5. For higher order terms, the entries are found by equating terms with the same n. The resulting coefficients are given in Malcolm and Nicholls (2011).

We summarize the steps needed to calculate the field for a perturbed velocity model in Algorithm 1.

Algorithm 1 Perturbed Helmholtz Solver

```
Determine number of spatial modes P
Determine order of Taylor series N
for p = 1 to P do
    Construct source vector R_p
    Construct velocity model matrix A_p
    Solve A_p z_(p,0) = r_(p,0) to get [u_(p_m,0) d_(p_m,0)]
    for n = 1 to N do
        from d_(n-1) and u_(n-1) Construct R_(p,n)
        Solve A_p z_(p,n) = R_(p,n) to get [u_(p_m,n) d_(p_m,n)]
    end for
end for
Use [u_(p_mn) d_(p_mn)] to calculate field v_m with Equation 9
```

Practical Considerations

The forward solver described in the previous section works both in the spatial and temporal frequency domain, making it difficult to directly compare to real or synthetic seismic data. If the domain spacing d is small, energy from other domains can cross into each other, generating edge effects. Setting the domain spacing very large relative to region of interest greatly mitigates this effect because the energy of a point source decreases due to geometric spreading. This comes at a small computational cost of increasing the number of spatial frequencies p needed to determine the solution.

In addition to the leakage effects, the source's periodic nature generates a problematic standing wave known as the Rayleigh or Wood's anomalies. The mathematical origin of this problem can be seen in Equation 8. In this series expansion, it is obvious that the expression for $\xi_p(x,y)$ will become very large if one of the low order β_p is very near zero. From the expression for β_p in Equation 3, this generates large amounts of energy at certain combinations of top layer velocity, domain spacing, and temporal frequency as shown in Figure 2 (left). Because the point source is periodic, the forward model is analogous to a diffraction grating in which the slits are replaced by an infinite number of point sources. Applying a source with a particular frequency excites a resonance in the diffraction grating that causes energy to be trapped in the upper layer of the model that does not decay with distance and cannot be mitigated by increasing the domain spacing (Maystre, 2012).

To prevent the energy of far away sources from entering other domains, a dissipation term must be added to the top layer. To implement this dissipation we multiplied the top layer velocity by $1 - 0.025i$. This effectively causes the energy of the wave to die out before it can reach the next domain. Figure 2 (right) shows the scattered field in the time and Fourier domains for a forward solve with a dissipative top layer and a lossless medium. This appears as large amplitude non-causal waves when viewed in the time domain. Without the addition of the dissipation term, the forward solve would not be comparable to real results or synthetic data generated from finite difference or finite element. In addition to complicating comparison to real data, the Wood's anomalies introduces many erroneous local minimia. Figure 3 shows the difference in cost functions as a function of top layer velocity and depth for a simple three-layered medium with the location of the true reflector located at 500 m with a velocity of 1100 m/s at frequency of 3 Hz. With the addition of absorption the minima is well behaved and smooth.

For large perturbations, the Taylor series expansion of the field given by Equation 9 is unstable and it is impossible to simulate the field using the field expansion method. To mitigate this anomaly and allow the

DOI http://dx.doi.org/10.1190/segam2015-5918749.1

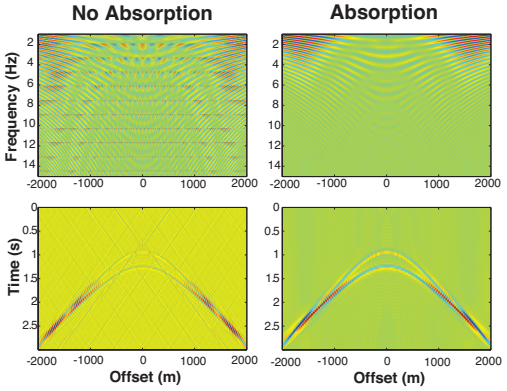

Figure 2: Left: A common shot record in the frequency and time domain for a loseless medium. Right: The same shot record with a small absorption term added to the top layer. We see that the addition of absorption has greatly attenuated the straight line Rayleigh anomalies apparent in the figures on the left.

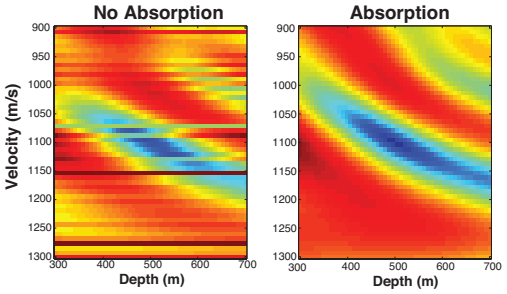

Figure 3: The cost functions as a function of top layer velocity and interface depth. Left shows a loseless medium. Right shows a medium with small absorption term in the top layer. Note the improvement in the behaviors of the cost function when absorption is included.

method to handle large perturbations, a Padé approximant (Baker and Graves-Morris, 1996) is used to enhance both the rate of convergence and permissible perturbation size as detailed in Malcolm and Nicholls (2011). With the addition of the Padé approximant, absorptive top layer, and large spacing between domain size, data generated using the field expansion method is comparable to finite difference simulations, as shown in Figure 4.

OPTIMIZATION METHOD

In this section we describe the global optimization method, PSO, used to solve the inverse problem of estimating a velocity model given a set of measured seismic traces. PSO is modeled on the behavior of flocks in which a collection (swarm) of potential solutions (agents) move through model space and communicate their own successes and failures with the other agents to find a global minima (Eberhart and Kennedy, 1995).

In the PSO algorithm, a collection of agents are randomly initialized across the entire range of model space with a random velocity at which they move through model space. At each iteration, all agents evaluate a cost function based on their current position; in our implementation this cost function is the mean square error between the measured and observed fields. All agents keeps track of the personal best, lowest scoring, location in model space it has visited and its corresponding score. The agents then move through model space according to their velocity. Each agent's velocity is then updated to accelerate the agent

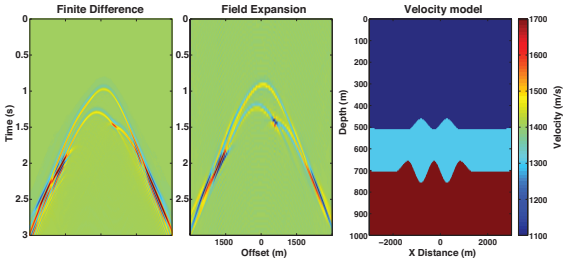

Figure 4: Right: Velocity model with perturbed interface. Middle: Gather generated using the field expansion method. Left: Gather generated using the Seismic Unix 2*nd* order finite difference solver. Good agreement is achieved between the two solvers.

towards its personal best location, the lowest scoring location it has visited in model space, and the neighborhood best location, the lowest scoring location across a subset of agents.

We implemented two forms of PSO: traditional PSO and local best PSO. The two variants of PSO only differ in how the neighborhood best is defined for each agent. In traditional PSO, each agent's velocity is influenced by the best known solution across the entire swarm. In local best PSO, each agent instead only sees the best known solution among its adjacent index neighbors (Bratton and Kennedy, 2007). For example in a swarm of 10 agents, the 5th agent's best neighborhood location is calculated among the personal best of agents with indices [4,5,6] and for the 10th agent the best is determined among indices [9,10,1]. This variation of the algorithm slows the speed of which information is transmitted across the swarm and prevents the algorithm from converging too early to a local minima (Kennedy and Mendes, 2006).

The steps can be summarized through the following pseudo code describing the swarm, consisting of N_s agents each with position vector $\mathbf{x}$ and velocity vector $\mathbf{v}$. For each agent in the swarm,

1. Initiate personal best score to infinity: $p_i^s = \infty$
2. Initiate swarm position from a uniform distribution across all of model space: $\mathbf{x}_i = \mathbf{U}[-1,1]$
3. Initiate swarm velocity from a uniform distribution in some range: $\mathbf{v}_i = \mathbf{U}[-vel_{max}, vel_{max}]$
4. Evaluate forward solve and calculate agent's score: $\mathbf{G}(\mathbf{X}_i)$
5. Record personal best score and location if better than current best: **if** $p_i^s < \mathbf{G}(\mathbf{X}_i)$ **then** $p_i^s < \mathbf{G}(\mathbf{X}_i)$ and $\mathbf{p}_i = \mathbf{x}_i$
6. Determine neighborhood best location, $\mathbf{g}_i$. If local best PSO: $\mathbf{g}_i \leftarrow min(p_{i-1}^s, p_i^s, p_{i+1}^s)$. If normal PSO: $\mathbf{g}_i \leftarrow min(p_i^s ..., p_{N_s}^s)$.
7. Calculate new position: $\mathbf{x}_i = \mathbf{x}_i + \mathbf{v}_i$
8. Calculate stochastic difference vectors, difference between best locations multiplied by uniform distribution: $\mathbf{d}_g = \mathbf{U}[0,1] * (\mathbf{g}_i - \mathbf{x}_i)$, $\mathbf{d}_p = \mathbf{U}[0,1] * (\mathbf{p}_i - \mathbf{x}_i)$
9. Calculate new velocity: $\mathbf{v}_i = \gamma \mathbf{v}_i + a_p \mathbf{d}_p + a_g \mathbf{d}_g$
10. Clamp velocities: If any entries of v_i exceed vel_{max} or $-vel_{max}$ then set the entries to vel_{max} or $-vel_{max}$
11. Go to step 4 until maximum number of iterations is reached.

In the expression for the velocity update above, γ is the inertia term controlling how much the previous velocity is maintained from the prior iteration and a_g and a_p are the acceleration terms that determine how much the personal best and neighborhood best alter each agent's velocity at each iteration. For all of the simulations performed in this paper the following parameters were used: $N_s = 40$, $\gamma = .9$, $a_g = 1.49$, and $a_p = 1.49$. The maximum velocity vel_{max} was clamped to $\pm .05$ of the total search space. Note that in the above algorithm, the evaluation of each agent's cost function is independent from one another and the forward solves can be parallelized across all agents.

To apply PSO to the field expansion framework, the velocity model

DOI http://dx.doi.org/10.1190/segam2015-5918749.1

Variable	Truth	Normal	Local	Prior
Interface 1 (m)	500	585 ± 101	520 ± 57	501 ± 1
Interface 2 (m)	1200	1227 ± 167	1220 ± 155	1197 ± 5
Velocity 1 (m/s)	1500	1433 ± 83	1486 ± 48	1501 ± 1
Velocity 2 (m/s)	2500	2630 ± 288	2539 ± 298	2495 ± 7
Velocity 3 (m/s)	3500	3867 ± 697	3537 ± 385	3494 ± 36

Table 1: Flat model inversion results

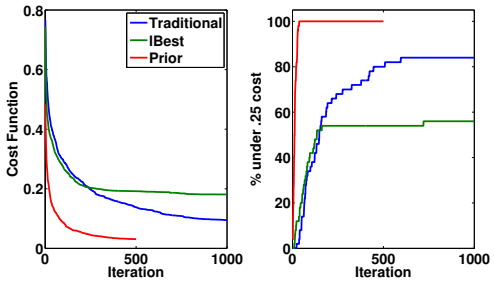

Figure 5: Left: Mean cost function for local and normal PSO as a function of iteration number. Right: Percentage of agents with cost function under .25 as a function of iteration number. Note the improvement seen with the incorporation of prior information.

must be converted into a vector of real numbers. For an M layered flat velocity model, the model can easily be converted into a vector of length $2M-1$ containing M velocities and $M-1$ layer boundaries, one for each of the boundaries. For the perturbed layer case, each of the layers are assigned $N_{perturbed}$ values that described the height of perturbation at $N_{perturbed}$ equally spaced locations across the region of interest of each interface. A cubic interpolation was then applied along each layer to define the surface for all locations. Under this parameterization, a perturbed velocity model could be described by $(M-1)*N_{perturbed}+2M-1$ values. The model is then be normalized to the range of [-1, 1] by dividing each value by the permissible perturbation, velocity, or layer depth limit.

EXPERIMENTS

lBest vs PSO

To compare the convergence characteristics and robustness of local best PSO vs normal PSO, we generated a synthetic dataset using the field expansion method consisting of 3 layers with interfaces at a depth of 500 and 1200 meters and velocities 1500, 2500, 3500 m/s from shallowest to deepest. The domain size or the spacing between periodic point sources was set to 20 kilometers with the synthetic shot fired in the center of the domain with 512 receivers spanning ± 3000 m from the source. Both regular PSO and local best PSO were run for 1000 iterations with a velocity range of [1000 6000] m/s, interface depth range of [100 2000] meters, and perturbation limit of 0 to force the layers to be flat. The inversion was then run 50 times for each algorithm. Figure 5 (left) shows the mean difference in cost function as a function of iteration and Figure 5 (right) shows percentage of simulations whose best score was less than .25 as a function of iteration. The local best PSO is initially slow to converge to a minima, but consistently reaches the minima more frequently. Table 1 shows the mean and standard deviation for the inversions.

Prior

In the previous simulations, we unrealistically assumed that we had no prior information of the velocity model other than constraining the interface and velocity limits to reasonable values. However, in almost all instances we would have some prior knowledge of the velocity obtained from methods such as NMO. For example, one could draw a set of velocity model from the NMO semblance image to generate a collection of prior velocity models. To mimic this, instead of assuming a completely random starting model from the entire velocity and interface limits, we adopt a semi-global approach in this experiment.

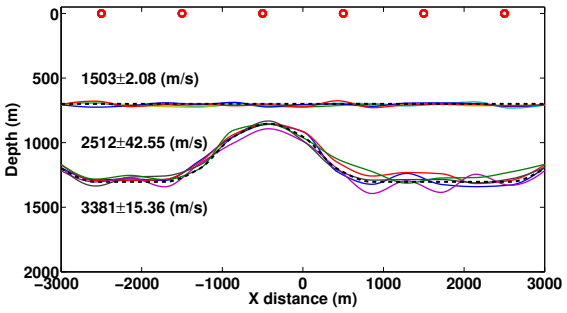

Figure 6: Inversion results for the anticline model. Dashed lines show the true interfaces and 5 solid color lines show estimates from 5 separate PSO runs. The numbers give mean and standard deviations of the 5 inversions for each layer.

We draw an initial model from a uniform distribution around the true model with a spread of [-300, 300] m for the interface limit and [-500, 500] m/s for the velocity distribution. As in the simulations above, the swarm can move through the entire range of velocities and interface limits. Only the initial population was altered to incorporate prior knowledge of the velocity limits. We then ran the inversion using local best PSO 50 times with the same parameters as the two simulations in the previous sub-section, with the maximum number of iterations decreased from 1000 to 500 as the algorithm converged much more quickly.

Changing the initial swarm from a purely random model to a more informed prior model greatly improved the performance of the inversion. As shown in Figure 5, all of the inversions reached a misfit of less than .25 within 100 iterations and the average data misfit as function of iteration is significantly lower than the traditional or local best PSO. The inversions using the prior achieved extremely accurate estimates of the velocity models at 500 iterations, as shown in Table 1.

Anticline with Prior

In the final experiment, we generate a 3 layer velocity model with a velocity of [1500, 2500, 3500] m/s from deepest to shallowest. The top interface is flat while the deeper interface has a anticline shaped perturbation, as shown by the dashed lines in Figure 6. We then generate 6 synthetic shot records at 3 Hz at the locations shown in Figure 6. We next perform local best PSO with an informed prior with a uniform distribution around the true model with a spread of [-400, 400] m for the interface location and [-500, 500] m/s for the velocity distribution. For the experiment, the interface was discretized into 13 interpolation points and we truncate the Taylor series expansion of the field after 8 terms. We ran the inversion 5 times for 500 iterations. The results are summarized in Figure 6. All of the runs of the algorithm were successful in recovering the anticline interface from a limited number of shot records.

CONCLUSION

In this paper we altered the domain size and added a small absorptive term to make the field expansion solver comparable to finite difference methods. Based on this fast solver, we presented an inversion algorithm using PSO to estimate a velocity model without an initial guess. Furthermore, we demonstrated how prior information can be incorporated into the PSO framework to greatly improve accuracy of the recovered velocity model. This opens up the possibility of estimating velocities, with error bounds, without the specification of an initial velocity model.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant No. (DGE-0806676) & (NSF-1115406) and the MIT Earth Resources Laboratory.

DOI http://dx.doi.org/10.1190/segam2015-5918749.1

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

Baker, G. A., and P. R. Graves-Morris, 1996, Padé approximants: Cambridge University Press, 59.

Bouchon, M., 2003, A review of the discrete wavenumber method: Pure and Applied Geophysics, **160**, no. 3, 445–465. http://dx.doi.org/10.1007/PL00012545.

Bratton, D., and J. Kennedy, 2007, Defining a standard for particle swarm optimization: Swarm Intelligence Symposium, IEEE, 120–127.

Burstedde, C., and O. Ghattas, 2009, Algorithmic strategies for full waveform inversion: 1D experiments: Geophysics, **74**, no. 6, WCC37–WCC46. http://dx.doi.org/10.1190/1.3237116.

Eberhart, R. C., and J. Kennedy, 1995, A new optimizer using particle swarm theory: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 39–43. http://dx.doi.org/10.1109/MHS.1995.494215.

Kennedy, J., and R. Mendes, 2006, Neighborhood topologies in fully informed and best-of-neighborhood particle swarms: IEEE Transactions on Systems, Man and Cybernetics. Part C, Applications and Reviews, **36**, no. 4, 515–519. http://dx.doi.org/10.1109/TSMCC.2006.875410.

Malcolm, A., and D. P. Nicholls, 2011, A field expansions method for scattering by periodic multilayered media: The Journal of the Acoustical Society of America, **129**, no. 4, 1783–1793. http://dx.doi.org/10.1121/1.3531931.

Maystre, D., 2012, Theory of Woods Anomalies, in Plasmonics: Springer Berlin Heidelberg, Springer Series in Optical Sciences, No. 167, 39–83.

Petit, R., 1980, Electromagnetic theory of gratings, *in* Electromagnetic Theory of Gratings Series: Topics in Current Physics, R. Petit, ed.: Springer Berlin Heidelberg.

Pratt, R., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901. http://dx.doi.org/10.1190/1.1444597.

Sambridge, M., and K. Mosegaard, 2002, Monte Carlo methods in geophysical inverse problems: Reviews of Geophysics, **40**, no. 3, 1009. http://dx.doi.org/10.1029/2000RG000089.

Sen, M. K., and P. L. Stoffa, 2013, Global Optimization Methods in Geophysical Inversion: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511997570.

Sirgue, L., 2006, The importance of low frequency and large offset in waveform inversion: Presented at the 68th EAGE Conference and Exhibition incorporating SPE EUROPEC 2006. http://dx.doi.org/10.3997/2214-4609.201402146.

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.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration: Geophysics, **74**, no. 6, WCC1–WCC26. http://dx.doi.org/10.1190/1.3238367.

Woodward, M., D. Nichols, O. Zdraveva, P. Whitfield, and T. Johns, 2008, A decade of tomography: Geophysics, **73**, no. 5, VE5–VE11. http://dx.doi.org/10.1190/1.2969907.