An application for null-space shuttling for 4D Full Waveform Inversion

Esteban Díaz, Simon Luo, and Madhav Vyas (bp)*

SUMMARY

The effective null-space (all the models that produce an equal or smaller misfit) of the full waveform inversion objective function can be quite large. Once the inversion converges there are many models that yield a similar fit. In the context of 4D FWI, this space is even bigger. The shuttling technology allows us to navigate the null-space to find models with equal or better fit that follow a secondary goal (i.e., minimizing the 4D model difference). In this abstract, we show a way to implement the 4D shuttling algorithm with two different goals and an application to a Gulf of Mexico dataset.

INTRODUCTION

Historically, seismic time-lapse monitoring has been an imaging-based workflow that involves careful and closely co-located acquisition, water velocity corrections, and careful seismic processing. The pre-processing is usually followed by high-resolution imaging (usually reverse time migration) which is followed by post-processing (image registration and cross-equalization). With more robust full waveform inversion (FWI) algorithms together with computational advances, we have been pushing subsurface velocity models to high frequencies (Shen et al., 2018) which in turn has enabled FWI-derived reflectivity (FDR) (Zhang et al., 2020)–a byproduct of the velocity inversion process. This approach can transform how we think about 4D processing projects altogether.

Here, we present a FWI-based workflow that is implemented after the inversion is driven to convergence. The method is derived from Keating and Innanen (2021) and shows a way to navigate the complex 4D inversion null-space and interrogate it in two different ways. We show how the retrieved models can preserve the data fit while driving the models to interesting places. On one hand, we can drive the difference down to unintuitively low NRMS using the model difference as shuttling goal. On the other hand, we can exaggerate the 4D response while also preserving the data-fit for both baseline and monitor surveys. This abstract is organized as follows: we first present the reference 4D FWI workflow (which we run before the shuttling step), we then explain our shuttling workflow, and finally, we present an application to a 4D dataset from Atlantis field, Gulf of Mexico.

THEORY

Here we explain the aspects of our 4D workflows, starting with the benchmark 4D FWI algorithm: let $\mathbf{m}_b$ and $\mathbf{m}_m$ be the baseline and monitor models, and $\mathbf{d}_b$ and $\mathbf{d}_m$ the corresponding data for baseline and monitor surveys, respectively. We can combine both models and data in a joint minimization goal as follows:

$$J(\mathbf{m}_b, \mathbf{m}_m) = ||\mathscr{F}(\mathbf{m}_b) - \mathbf{d}_b||_2^2 + ||\mathscr{F}(\mathbf{m}_m) - \mathbf{d}_m||_2^2 + \varepsilon R(\mathbf{m}_b, \mathbf{m}_m), \quad (1)$$

where $\mathscr{F}$ is a wave equation modeling operator sampling the wavefield at the receiver locations, R is a regularization goal that imposes some constraints on the relationship between baseline and monitor models, and ε it is a scalar that controls the relative weight between the joint regularization goal and the data fitting goals. In practice, we use the L-curve regularization criteria to choose the optimal ε value for our benchmark inversions.

Null space shuttling for 4D

(Keating and Innanen, 2022) expands the shuttling algorithm (Keating and Innanen, 2021) to explore the uncertainty in the 4D FWI problem. Here, we modify the update direction for the shuttling problem and define a new shuttling goal that can be useful for 4D imaging. Nullspace shuttling can be defined as the problem that seeks to minimize a secondary goal F_s under the constraint that the primary optimization goal F_p remains at a lower or equal value through each iteration:

$$\begin{aligned} \min_{\mathbf{m}_i} \quad & F_s(\mathbf{m}_i) \\ \text{s.t.} \quad & F_p(\mathbf{m}_i) \le F_p(\mathbf{m}_{i-1}) \end{aligned} \quad (2)$$

where $\mathbf{m}_{i-1}$ sets the reference function contour $F_p(\mathbf{m}_0)$. We assume that the primary problem F_p has been previously optimized to convergence.

To satisfy both goals, we need to update the model along a direction that minimizes both F_p and F_s:

$$\Delta u = -\hat{g}_p - \hat{g}_s. \quad (3)$$

Where $\hat{g} = g/||g||_2$ is the normalized gradient.

This mixing produces an update direction that can minimize both functions within a given range of suitable steplength α. Once the update direction Δu is constructed, the remaining step is to find the step to properly scale it. Thus, creating the following scalar optimization sub-problem:

$$\begin{aligned} \min_{\alpha_i} \quad & F_s(\mathbf{m}_{i-1} + \alpha_i \Delta u) \\ \text{s.t.} \quad & F_p(\mathbf{m}_{i-1} + \alpha_i \Delta u) \le F_p(\mathbf{m}_{i-1}) \end{aligned} \quad (4)$$

The secondary optimization goal is often less computationally involved compared to the primary goal. Using the chain rule, the gradient of F_s w.r.t. α is:

$$\frac{\partial F_s}{\partial \alpha} = \frac{\partial F_s}{\partial \mathbf{m}} \frac{\partial \mathbf{m}}{\partial \alpha} = \frac{\partial F_s}{\partial \mathbf{m}} \cdot \Delta u. \quad (5)$$

With this gradient, we can go ahead and optimize the secondary goal F_s for alpha to find α_{opt}. Once we found it, we only need to re-scale it such that we meet the constraint

$$F_p(\mathbf{m}_{i-1} + \alpha_i \Delta u) \le F_p(\mathbf{m}_{i-1}) \quad (6)$$

To do so, we only need function evaluations for the primary and more expensive goal F_p within the range $0 < \alpha \leq \alpha_{opt}$.

How to choose the secondary function F_s?
An intuitive function would be to minimize the distance between the model and the reference model:

$$F_s = ||\mathbf{m} - \mathbf{m}_{ref}||_2^2. \quad (7)$$

We can call this strategy "minimum shuttle". As mentioned before, during a 4D shuttling the baseline becomes the reference while we shuttle the monitor, and then we flip it and the monitor becomes the reference while we shuttle the baseline model. One can imagine if both objective functions have similar contours with considerable overlap there is the possibility of ending up with zero distance between baseline and monitor models while maintaining (or improving) the data misfit for both datasets.

Other goals could be entertained as well, for instance, one could find the maximum distance that still maintains acceptable data fit by minimizing the negative of equation 7. Alternatively, we can also mix these two goals, by minimizing away from the area of interest and maximizing the 4D response around the reservoir. This could can be thought of as using a magnifier. A goal that tackles both objectives is:

$$F_s = ||\mathbf{K}(\mathbf{m} - \mathbf{m}_{ref})||_2^2 - ||(\mathbf{I} - \mathbf{K})(\mathbf{m} - \mathbf{m}_{ref})||_2^2. \quad (8)$$

In this "dual shuttle" goal, we use the mask $\mathbf{K}$ to highlight the places where we want to minimize and $\mathbf{I} - \mathbf{K}$ to highlight the places where we want to maximize the distance. Note that this goal will go negative when the second term takes over the inversion.

Solving a toy problem
We present a toy problem that solves equation 2. The primary optimization goal is given by a two-dimensional rotated Gaussian function:

$$F_p(x,y) = 1 - e^{-\frac{(x'-x'_o)^2}{2\sigma_x^2} - \frac{(y'-y'_o)^2}{2\sigma_x^2}} \quad (9)$$

where $x' = (x - x_o)\cos(\theta) - (y - y_o)\sin(\theta)$, $y' = (x - x_o)\sin(\theta) + (y - y_o)\cos(\theta)$, $x'_o = x_o\cos(\theta) - y_o\sin(\theta)$, and $y'_o = x_o\sin(\theta) + y_o\cos(\theta)$. For this problem, we use $x_o, y_o = 0.0, 0.0$ and $\theta = \pi/4$.

We use the "minimum shuttle" as a secondary objective:

$$Fs(x,y) = (x - x_r)^2 + (y - y_r)^2. \quad (10)$$

In this example, the reference point is $x_r, y_r = -1.5, 1.5$.

Figure 1 shows how we minimize the model (moving it to the center of the Gaussian) while also minimizing the distance to the reference point (the dashed line is a reference for the minimum distance path). Note that any point along the segment that connects the reference point, the center of the Gaussian, and it has a lower value of the initial contour will satisfy our constraints. Where we fall in along this segment depends on some of our optimization choices. In practice, in the context of 4D FWI, we alternate between baseline and monitor models to set them as a reference and take one or only a few shuttling steps while immediately changing the reference model. Alternating between the two vintages allows us to quickly converge.

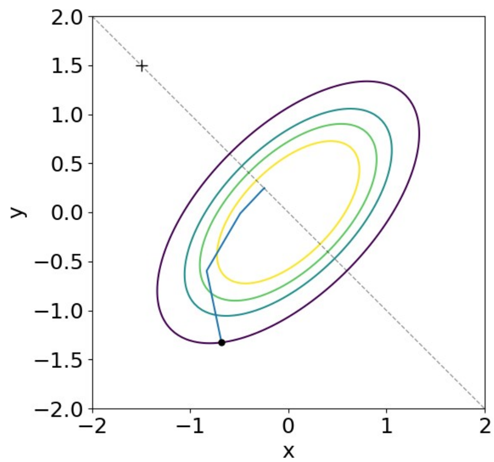
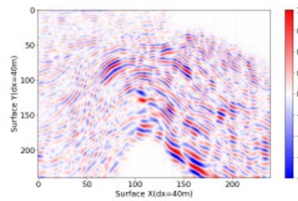


Figure 1: Toy problem setup: the black dot shows a hypothetical final model driven to convergence (and the starting point for the shuttling iterations), the path shows all the models after each shuttling step. The black cross is the reference to which we are minimizing the distance while preserving or improving the convergence. The dashed line connects the reference model with $(x,y) = (0,0)$ which is the lowest point in the function.

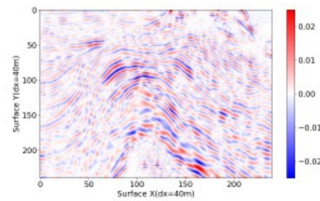
ATLANTIS 4D EXAMPLE

We run a 4D example from the Atlantis field, Gulf of Mexico. Our 4D data consists of two vintages: one from 2005 and one from 2022. The first step is to run the benchmark workflow (joint FWI with regularization constraints) as explained in the previous section. The velocity models were optimized jointly until convergence. For assessing the 4D signal, we take a look at the FDR difference.

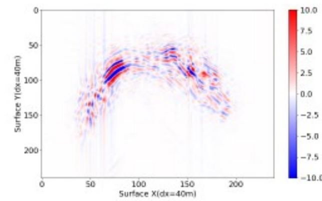
Figure 2a shows a section through the Atlantis model for the FDR difference volume. For the shuttling approach, we perform six shuttling steps (three for the base model and three for the monitor model) using the minimization approach and the dual approach. Figure 2b shows the 4D response on the FDR volumes using the minimization goal. Finally, Figure 2c shows the 4D response on the FDR volumes using the dual approach. An important observation from these sections is how different the 4D FDR amplitude response is. To gain more insights about such amplitude range, let's take a look at the optimization history from the shuttling exercises. Figure 3a shows the shuttling goal optimization for both minimization and dual goals. One can see how our algorithm was able to achieve what we set as a goal. Figure 3b shows the FWI objective functions (F_p) for both base and monitor models, and both shuttling runs (minimization and dual goals). We can observe how the FWI goal was further optimized for both shuttling runs and yet the 4D FDR response yields dramatically different models. This observation gives us an insight into how the topology of both FWI functions (baseline and monitor) perhaps have a great overlap. Hence, there are many combinations of mod-



(a)

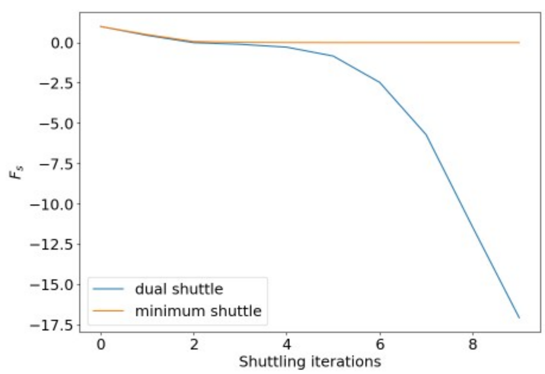


(b)

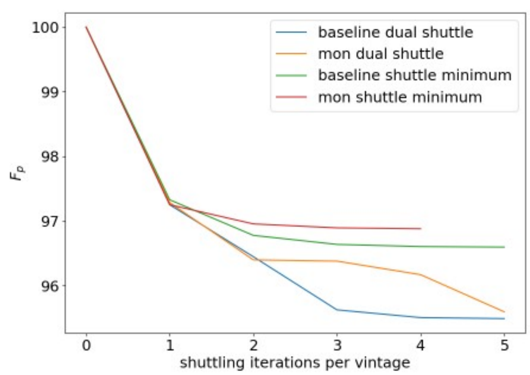


(c)

Figure 2: Zoom through the 4D anomaly on the FDR difference volumes for (a) reference 4D, (b) minimum shuttle distance, and (c) dual shuttle distance. Note the difference in the amplitude scale.



(a)



(b)

Figure 3: Shuttling convergence (a) Secondary goal convergence for both distance minimization and distance maximization goals and (b) Primary goal decrease for shuttling approaches and both vintages (baseline and monitor).

els that are equally or more feasible than the final model from conventional FWI optimization. The shuttling algorithm gives us the possibility to explore this vast space of models given an input goal.

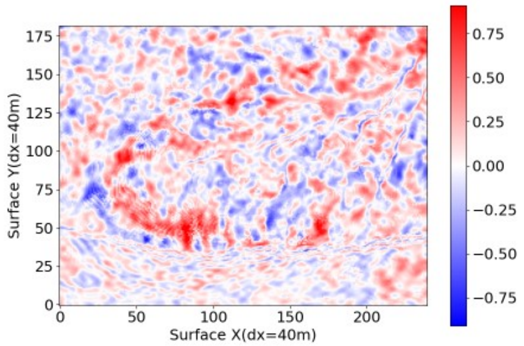
Another way to look at the 4D response is by doing windowed attribute extractions. Figure 4a shows the average amplitude attribute for the benchmark 4D FDR extracted along a 60-sample window around the horizon. Figure 4b shows the corresponding NRMS attribute for the benchmark models. Figures 4c-4d show the equivalent extractions for the minimization shuttling goal, and finally Figures 4e-4f show the same extractions for the dual shuttling approach. One can see that we can use the shuttling framework to drive the NRMS almost to zero in the case of minimization, or also we can use it to maximize or exaggerate the 4D response.

CONCLUSIONS

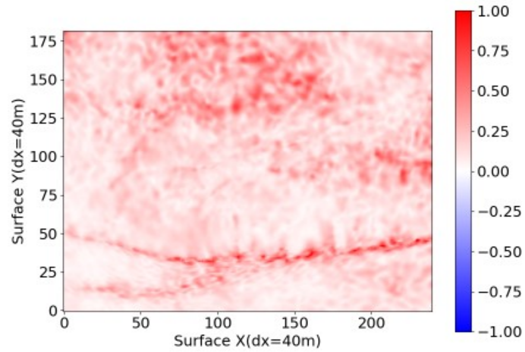
We have presented an alternative algorithm for 4D null-space shuttling that is equivalent in cost to running a few FWI iterations after the conventional optimization step is finished for both baseline and monitor. Through shuttling, we can gain insights into how reliable a 4D signal is and what is the overlap between the baseline and monitor objective functions in the model space. If there is a considerable overlap, it is possible to drive the NRMS attribute to very low levels (a fraction of a percent). At the same time, it is possible to drive the 4D signal to unreasonably high responses. These two scenarios can be achieved while driving the primary goals (the FWI objective functions for baseline and monitor) to lower values (better data fit). It is important to highlight that the method described in this abstract is independent of the chosen FWI flavor, it can give us more information about the function landscape of either acoustic or elastic FWI objective functions.

ACKNOWLEDGEMENTS

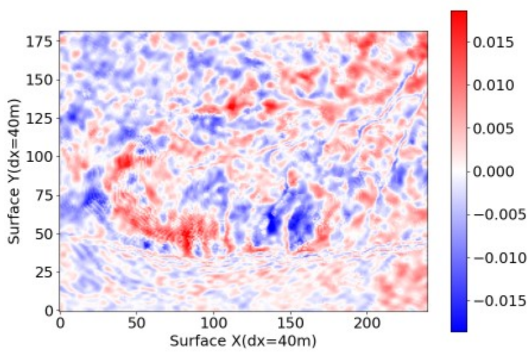
We would like to thank bp and Woodside Energy for the permission to publish this abstract. We also thank Scott Keating (ETH Zürich), Kris Innanen (CREWES, University of Calgary), and Jingfeng Zhang (bp) for discussions regarding the null-space shuttling method.



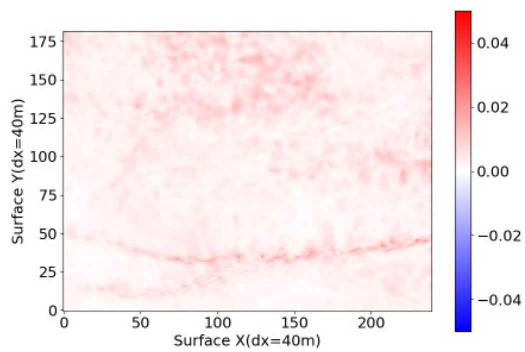
(a)



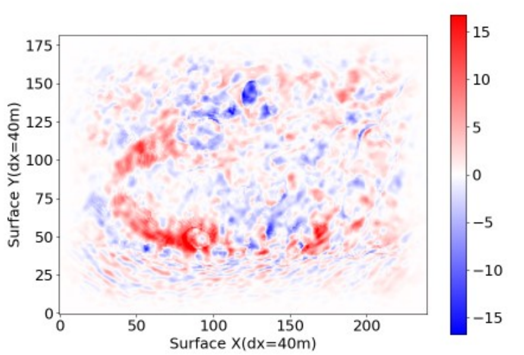
(b)



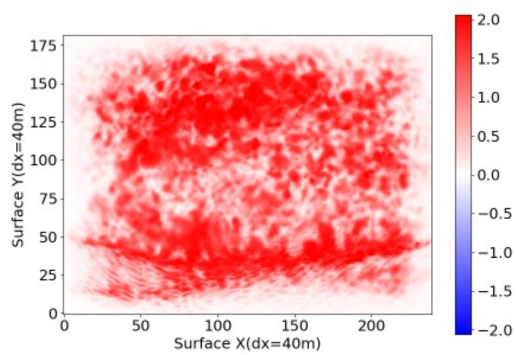
(c)



(d)



(e)



(f)

Figure 4: Attribute extraction along a 60 sample window of the FDR difference. The left column (a), (c), and (e) shows the average difference attribute whereas the right column (b), (d), and (f) show the NRMS attribute for the benchmark, shuttling minimum, and shuttling maximum distance, respectively.