

# Preconditioning full-waveform inversion with efficient local correlation operators

Paul Wellington<sup>1</sup>, Romain Brossier<sup>2</sup>, and Jean Virieux<sup>3</sup>

## ABSTRACT

Full-waveform inversion (FWI) is an iterative locally linearized data-fitting technique. The FWI method attempts to move from an initial low-wavenumber representation of earth parameters to a broader representation of the medium. An issue with the method is that FWI is an ill-posed problem, oversampled for the numerical forward discretization. The success of the model parameter reconstruction can often be greatly affected by external factors such as the presence of noise in the input field data or other artifacts arising from the imaging condition present in the FWI gradient computation. We have developed a strategy for mitigating against the influence of such external factors by preconditioning the discrete data gradient using an efficient

differential approach instead of the often used integral formulation. Such an application of a smoothing correlation operator allows one to use prior information to locally filter along expected geological dips while being consistent with faults. The application of this preconditioning strategy to real and synthetic 2D data sets illustrates how this incremental additional step makes the FWI workflow less sensitive to noise and spatial aliasing artifacts. Nothing prevents a possible 3D acoustic extension, thanks to the small added computer cost for this local filter application. Three-dimensional elastic FWI may require this inexpensive filtering strategy due to the prohibitive forward-modeling costs that could be partially mitigated by using a coarse shot increment in conjunction with gradient preconditioning.

## INTRODUCTION

Full-waveform inversion (FWI) offers the promise of delivering a broadband representation of subsurface parameters by iterative linearized inversion of the seismic waveform objective function (Lailly, 1983; Tarantola, 1984). Although it is theoretically feasible to invert for the full set of parameters that describe wave propagation in complex media (Toksöz and Johnston, 1981), the FWI method is most frequently used as a high-resolution velocity model building method (Fichtner et al., 2010; Plessix and Perkins, 2010; Sirgue et al., 2010; Tape et al., 2010; Peter et al., 2011; Zhu et al., 2012; Warner et al., 2013; Vigh et al., 2014; Borisov and Singh, 2015; Operto et al., 2015). Theoretically, the global minimum of seismic waveform misfit could be sought using global optimization methods (for an extensive review, see Sambridge and Mosegaard, 2002). Unfortunately, the computational cost of performing seismic waveform modeling many times and at high frequencies does not

allow a thorough exploration of the solution space. The commonly applied approach is a local descent-based optimization with the first derivative of the objective function, the gradient being efficiently computed using the adjoint state method (Plessix, 2006). This first-order information can be complemented with the second-order information from the Hessian operator to improve the result and convergence speed (Pratt et al., 1998; Métivier et al., 2014).

The gradient of the misfit function can be seen as the zero-lag correlation of the forward and adjoint wavefields. The computation of the FWI gradient is therefore very similar to the reverse time migration (RTM) imaging condition. In FWI, instead of correlating the forward-propagated source with the back-propagated receiver wavefields as is the case in RTM, one correlates the incident wavefield with the back-propagated receiver residuals. The current model affects not only the back propagation but also the residual estimation, illustrating the nonlinearity of FWI. The gradient is generally estimated on the same grid as the one used with for the forward modeling.

Manuscript received by the Editor 1 September 2018; revised manuscript received 1 December 2018; published ahead of production 11 January 2019; published online 11 March 2019.

<sup>1</sup>Formerly Univ. Grenoble Alpes, ISTerre, Grenoble, France; presently Chevron Australia, Perth, Australia. E-mail: paul.wellington@chevron.com.

<sup>2</sup>Univ. Grenoble Alpes, ISTerre, Grenoble, France. E-mail: romain.brossier@univ-grenoble-alpes.fr.

<sup>3</sup>Univ. Grenoble Alpes, ISTerre, CNRS, Grenoble, France. E-mail: jean.virieux@univ-grenoble-alpes.fr.

© 2019 Society of Exploration Geophysicists. All rights reserved.

This could result in a too-coarse grid for the seismic imaging condition inducing spatial aliasing artifacts. A second origin of spatial aliasing will come from the use of too few shots/receivers or from a too-large monochromatic frequency increment if FWI is formulated in the frequency domain (Sirgue and Pratt, 2004). It is for these reasons that FWI is intrinsically ill posed. The model parameter reconstruction can be contaminated by artifacts that should be mitigated through the application of a local operator on the gradient controlling the expected parameter resolution and the prior information we may have from other interpretations.

The commonly applied forms of prior constraints can be subdivided into those based on regularization or preconditioning strategies. The regularization strategies apply a model-space penalty term to the waveform misfit objective function to enforce an auxiliary reduction goal, adding a model gradient contribution. Often, this penalty term involves a derivative stencil that will typically look to enforce smoothness (Tikhonov and Arsenin, 1977) or sparsity (Vogel and Oman, 1996; van den Berg and Abubakar, 2001; Vogel, 2002; Loris et al., 2007; Herrmann et al., 2009; Guitton, 2012). It is also feasible to regularize the solution using a geologically based prior model with uncertainties represented in a diagonal inverse covariance matrix (Asnaashari et al., 2012). Interactions between neighboring points of the model space allow us to consider a more general expression of the covariance matrix, but they are more difficult to manage in higher dimensional spaces. A common issue with these approaches is the selection of the correct regularization weight without being required to run the entire FWI process with a range of different values (Hansen, 1998, 2010) and the ambiguity involved in building a meaningful prior model with associated covariance matrix using limited and incomplete information. The strategy we shall consider here could be potentially applied to this regularization option, but we prefer to focus on the following preconditioning strategies related to the data-misfit gradient for this first application.

Preconditioning strategies provide an alternative approach to constrain the inversion approach. Unlike the regularization strategies, the preconditioning strategies can be seen to directly remove unwanted features in the data gradient. Harlan (1995) suggests that, instead of applying a penalty term that discourages complexity through derivative operators, we can perform a change of variable to a model space that builds model simplicity directly. This preconditioning approach has found acceptance in FWI because of its more efficient way of removing high-wavenumber artifacts than regularization procedures. Such preconditioning can be seen as a particular projection to smooth updates. Such projection approaches are nicely reviewed by Peters and Herrmann (2017), who compare regularization and projection methods in the frame of edge-preserving schemes in FWI. Guitton et al. (2012) use local inverse dip filters to ensure that model updates are aligned along local geologic dips. Such dip-based operators are often referred to as steering filters with some key examples being the wave-kill filter (Claerbout, 1992; Clapp, 2001), the plane-wave destruction filter (Fomel, 2002), and filters based on directional Laplacians (Hale, 2007). For these filters, it is efficient to calculate the dip filtering operator, its inverse and adjoint. The inverse of these dip filters are used to shape the updates along a desired direction.

In this paper, we present an alternative strategy based on an efficient differential application of multidimensional correlation filters (Wellington et al., 2017). The key difference between this and the model reparameterization approach is that, instead of shaping updates along the geologic dip with inverse dip filters, we propose to directly

attenuate undesired features from the data gradient using low-pass wavenumber correlation operators. These correlation operators are built with extended correlation lengths in the orientation of the structural dip and very short correlation lengths orthogonal to dip. An added and novel aspect of the operators is that they are efficiently applied using a differential approach rather than an integral approach. We first present an application of this strategy to a subset of the Marmousi model before showing its effectiveness on a 2D field data line acquired over the northwestern continental shelf of Australia with broadband frequency content using legacy Broadseis technology (Soubaras and Whiting, 2011).

## METHODOLOGY

The FWI process seeks to update the model parameters  $\mathbf{m}$  involved in the wave propagation, so that the misfit between the modeled  $\mathbf{d}_{\text{cal}}$  and observed  $\mathbf{d}_{\text{obs}}$  seismic wavefields is minimized using, for example, the least-squares misfit expression:

$$C(\mathbf{m}) = \frac{1}{2} \|\mathbf{d}_{\text{cal}} - \mathbf{d}_{\text{obs}}\|^2. \quad (1)$$

The calculated seismic data are deduced from the model parameters through the formal expression

$$\mathbf{d}_{\text{cal}} = \mathcal{F}(\mathbf{m}), \quad (2)$$

where the operator  $\mathcal{F}$  computes the seismic wavefield and extracts the solution at the receiver positions. Starting from an initial model  $\mathbf{m}_0$ , we update a current model  $\mathbf{m}_n$  at the iteration  $n$  with a perturbation model  $\Delta\mathbf{m}_n$  to define the new model  $\mathbf{m}_{n+1} = \mathbf{m}_n + \Delta\mathbf{m}_n$ . In such a case, we need to look at the shape of the misfit function around the current model  $\mathbf{m}_n$  (for an FWI illustration, see Virieux and Operto, 2009). The model perturbation is given by

$$\Delta\mathbf{m} = - \underbrace{\left[ \frac{\partial^2 C(\mathbf{m}_n)}{\partial \mathbf{m}^2} \right]^{-1}}_{\text{Hessian}=\mathbf{H}_n} \underbrace{\frac{\partial C(\mathbf{m}_n)}{\partial \mathbf{m}}}_{\text{Gradient}=\mathbf{G}_n}. \quad (3)$$

The updated model  $\mathbf{m}_{n+1}$  can then be expressed as

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha_n \mathbf{H}_n^{-1} \mathbf{G}_n, \quad (4)$$

where the so-called “step-length”  $\alpha_n$  is applied to the perturbation vector.

In our approach, we only calculate the gradient using the adjoint state method (Plessix, 2006) and converge to a solution using the quasi-Newton LBFGS method (Nocedal and Wright, 2006; Brossier et al., 2009) that allows us to include approximated second-order information based on the previous model and gradient computations.

### Gradient preconditioning

To limit the influence of undesired features such as previously discussed noise and aliasing issues, we propose to precondition the inversion using a spatial correlation operator applied to the gradient. The following development is performed in two dimensions, but extension to three dimensions would be straightforward. This correlation operator can be thought of as a local, 2D low-wavenumber filter that we use to remove undesired features from the gradient.

The standard FWI gradient

$$\mathbf{G}_n(\mathbf{x}) = \frac{\partial C(\mathbf{m}_n)}{\partial \mathbf{m}} \quad (5)$$

is therefore replaced by the preconditioned gradient

$$\mathbf{G}'_n(\mathbf{x}) = \int_{\mathbf{x}} \mathbf{Corr}_{2D}(\mathbf{x}) \mathbf{G}_n(\mathbf{x}) d\mathbf{x}, \quad (6)$$

where the operator  $\mathbf{Corr}_{2D}(\mathbf{x})$  is a 2D correlation operator with two key features that we would require. First, the application of the operator to the gradient vector should add no significant time to the already costly FWI process. Second, we would like the operator to be locally oriented to prior locally defined geologic dips when available. The decay of the correlation coefficients should be slower when parallel with the local dip where the geology is expected to be more continuous. Correlation coefficients orthogonal to the structure can decay quickly because the geology typically varies much more rapidly along this direction, and we would like to maintain this expected resolution in the inverted results after the application of the correlation operator. Therefore, such a filter is essentially spatially anisotropic. Figure 1 illustrates this operator property for a single point  $(x, z)$ . At this central point, the correlation weight will be at a maximum; however, as we move away to an offset point  $(x', z')$ , the correlation weight decreases more or less rapidly depending on the local geologic dip. The rate of this decay is controlled by the locally oriented correlation lengths  $l_{\hat{x}}$  and  $l_{\hat{z}}$  deduced from the dip and local horizontal/vertical correlation lengths  $l_x$  and  $l_z$ . Moreover, as the orientation of geologic dip is nonstationary, Wellington et al. (2016) show that one can neglect spatial variation of such parameters in the efficient differential formulation of this filtering, which is now shortly described.

To achieve these two objectives, in the case of this study, we use two successive inverse Bessel filter operators developed by Wellington et al. (2017) and Trinh et al. (2017), to approximate a 2D Laplacian correlation function. Such an inverse filter allows to calculate the preconditioned gradient by solving a simple linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad (7)$$

where  $\mathbf{A} = \mathbf{Corr}_{2D}^{-1}$  is the discrete version of the inverse Bessel filter, the equivalent differential expression of the integral equation (equation 6). The solution vector  $\mathbf{x} = \mathbf{G}'_n$  is the filtered gradient, whereas the right side  $\mathbf{b} = \mathbf{G}_n$  is the raw gradient.

As shown in Trinh et al. (2017), the operator  $\mathbf{Corr}_{2D}^{-1}$  is positive definite. The linear-conjugate gradient approach is applied here to this well-posed linear system, which is solved in a few tens of iterations ( $k$ ) until convergence is reached below the stopping criterion:

$$\epsilon = \frac{\|\mathbf{b} - \mathbf{A}\mathbf{x}^k\|}{\|\mathbf{b}\|} < 10^{-4}. \quad (8)$$

The behavior of this preconditioned FWI strategy is highlighted in the following section, in which we consider a synthetic illustration on a subset of the Marmousi model (Martin et al., 2006) and a field data application on a Australian Broadseis data set.

## APPLICATION

We present two 2D FWI applications of our gradient preconditioning strategy. The first uses a subset of the Marmousi model. The importance of the local anisotropic filtering to mitigate against noise is described as well as how dip and fault information can be included into the FWI workflow. The second example we present is a real data case study taken from the Northwest Shelf of Australia.

Time is initially taken to detail this interesting real data example with an FWI workflow that does not use gradient preconditioning. After an initial investigation, our preconditioning strategy is used to mitigate against spatial aliasing effects that arise from coarse shot sampling. Running a coarse shot modeling increment will provide significant promise for decreasing the computational cost involved in FWI especially in 3D and elastic configurations (Trinh et al., 2019).

### Marmousi synthetic model

We apply preconditioned FWI to a subset of the Marmousi model. The subset is selected to illustrate the effect of the filter application in the reconstruction of the tilted fault blocks present in the true velocity model via FWI. We start the inversion from a smoothed initial model constructed by applying a 500 m isotropic Gaussian smoother to the slowness of the true model. The goal of FWI is the reconstruction of intermediate- to high-wavenumber content present in the true  $V_p$  model (stratigraphy in the rifted fault blocks, including two gas-charged sand zones). This level of detail is absent from the initial model that only contains the low-wavenumber velocity trend (Figure 2). The FWI velocity model is discretized using a 10 m grid, and it is composed of 130 cells in the vertical and 390 cells in the horizontal direction. We use frequency-domain FWI with the SEISCOPE open-source tools TOY2DAC and the SEISCOPE Optimization toolbox (Métivier and Brossier, 2016). A fixed-spread acquisition is used with receivers every 10 m and sources every 65 m. Our FWI application simultaneously inverts the nine frequencies between 4 Hz  $\rightarrow$  20 Hz (2 Hz increment), and we investigate the success of FWI for a very noisy (signal to noise = 0.1) and noise-free data sets.

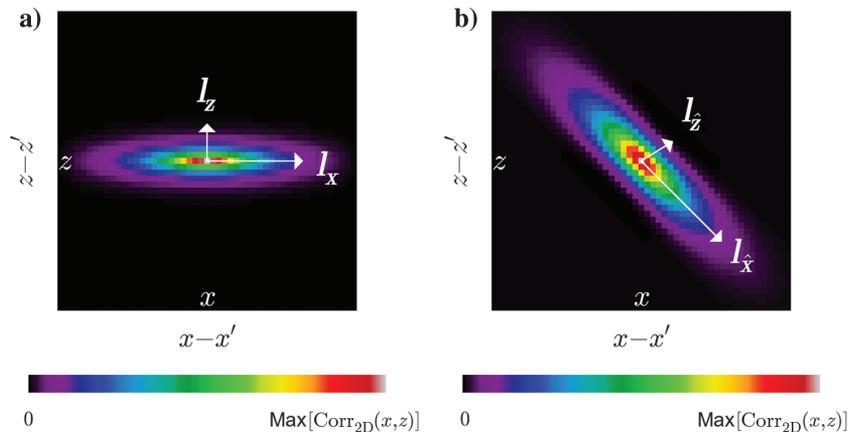


Figure 1. An example of anisotropic correlation functions. (a) The horizontal correlation length  $l_x$  is longer than the vertical  $l_z$  meaning that the correlation weights decay much more rapidly in the vertical direction versus the horizontal direction. (b) The same correlation lengths have been rotated by 45° to the arbitrary orthogonal directions  $l_{\hat{x}}$  and  $l_{\hat{z}}$ . Such a rotation is important if we want to align our anisotropic operator parallel with the stratigraphic dip.

The Marmousi synthetic example is used to explore the effect that incrementally adding prior information has on the model parameter reconstruction. We first look at the effects of preconditioning using an anisotropic correlation filter in which the correlation length in the horizontal direction is extended relative to the vertical direction. The complexities of the Marmousi model also allow us to investigate the effect of orientating this filter with dip and also relaxing it at the structural discontinuities (the faults). The utility of these prior constraints is quite important, especially in the case in which the observed data  $\mathbf{d}_{\text{obs}}$  are contaminated with noise.

Our anisotropic correlation filter is constructed using orthogonal correlation lengths  $l_x$  and  $l_z$  of 75 and 10 m, respectively. There are a couple of considerations to be made when selecting these correlation lengths. First, we are looking to maximize the reconstruction of the geologic contrasts that are in the velocity model. When there is zero dip, the largest variations tend to occur in the vertical direction as we pass through stratigraphic layers. Therefore, we choose our vertical correlation length  $l_z$  to be as small as possible in this orientation. We apply our filter by solving a discretized least-squares linear system and have found that the smallest the vertical correlation length can be in practice to avoid numerical artifacts is half the grid-cell size. The selection of the horizontal correlation length  $l_x$  on the other hand should be much larger than  $l_z$  allowing filtering/summation parallel with the stratigraphic boundaries. Because the filter is in essence a low-pass wavenumber,  $l_x$  should depend on the minimum wavenumber that we should be expecting to be contributing to a coherent update in the model. This will depend on the wavefield energy that is mapped into the gradient imaging condition and will be a function of propagation frequency, local velocity, and acquired scattering angles. There is no reason that these correlation lengths need to be fixed during inversion, and the correlation lengths may also vary spatially. For our example, we have selected constant quantities purely for simplicity.

The geologic dip in this subset of the Marmousi model is not constant or negligible (a maximum dip of  $50^\circ$ ). Therefore, we also investigate the importance of rotating these correlation lengths ( $l_x$  and  $l_z$ ) to the true structural dip ( $l_{\hat{x}}$  and  $l_{\hat{z}}$ ). Horizon information provided by Martin et al. (2006) is gridded to correctly represent this dip field (Figure 3). The longest correlation length  $l_{\hat{x}}$  is locally aligned with the geologic dip. Although not explored in this paper, automatic dip estimation from depth migrated images could be

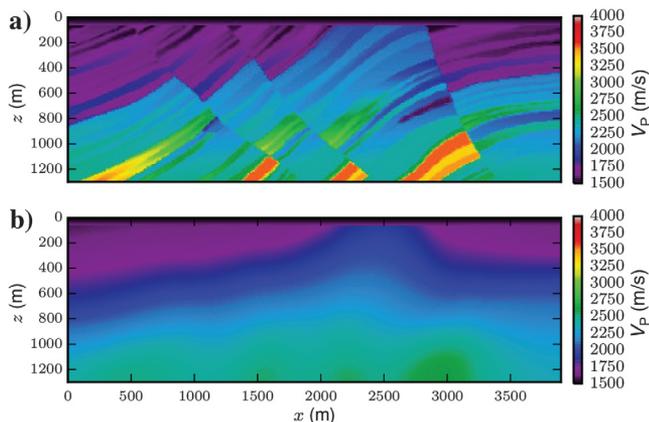


Figure 2. The subset of the Marmousi synthetic model used for FWI investigation. (a) True model and (b) initial model.

implemented in real data cases. Such a strategy is used by Guitton et al. (2012), but it is beyond the scope of this study focused on the efficient correlation filter application.

The final constraint that we test is the relaxing of the correlation filter anisotropy magnitude near the faults. It may be possible to select these fault locations from migrated images or any other equivalent strategy. In our case, we consider them as known inputs. As the Euclidean distance between approximate fault location increases, the large correlation length  $l_{\hat{x}}$  decreases toward the value of the small correlation length  $l_{\hat{z}}$  (Figure 4). The rate of this transition could be very slow if the location of the faults is highly uncertain. The  $l_{\hat{z}}$  correlation length is kept constant at 10 m as in the previous cases. Our discrete operator  $\mathbf{A}$  can handle spatial variations of the correlation parameters as long as they vary smoothly (Trinh et al., 2017).

A schematic behavior of the preconditioning operator for the various correlation configurations is shown in Figure 5. In the case in which just the anisotropic constraint is applied, the raw gradient  $\mathbf{G}_n$  is filtered by the same 2D correlation operator for every point in space to calculate the preconditioned gradient  $\mathbf{G}'_n$ . Because this operator is anisotropic, a contour of constant correlation weight forms an ellipse. When the dip is included, the long axis of the ellipses becomes  $l_{\hat{x}}$  aligned parallel with the local dip field. The final added constraint means that, at the faults, the operator becomes isotropic and mild to preserve potential sharp details at the faults while remaining anisotropic and aggressive elsewhere.

The gradient at the first iteration is shown for the different cases of constrained and unconstrained FWI. Figure 6a shows the unconstrained FWI gradient when no additional noise has been added to the observed data. In this case, we see that the gradient will update the initial model with a lot of structural/stratigraphic features that are present in the true model. Figure 6b shows the gradient that is obtained when Gaussian noise is added to the observed data. One can see that the stratigraphic layers are reconstructed in addition to some weak noise. With no constraint present in the inversion

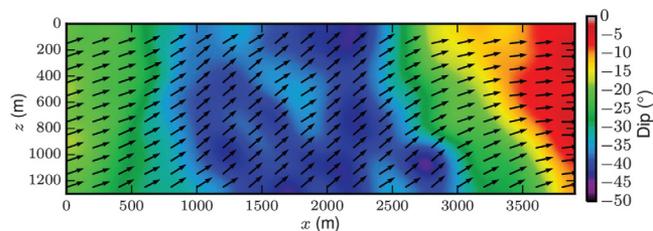


Figure 3. Dip field calculated by gridding the original Marmousi horizons provided in Martin et al. (2006). A value for dip exists for every grid point, and the black arrows show the orientation of the longest correlation length  $l_{\hat{x}}$ .

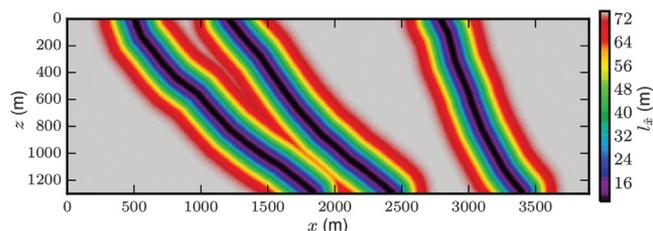


Figure 4. A variable  $l_{\hat{x}}$  model to add details nearby fault planes. The variable  $l_{\hat{x}}$  model increases  $l_{\hat{x}}$  from 10 m at the faults to a maximum of 75 m as the distance from the fault increases.

to mitigate against the undesired noise imprint, it will be added into the reconstructed model result at each iteration. We attempt to combat these undesired artifacts by using our correlation-based preconditioner.

First, we apply the anisotropic correlation filter that does not align it with the local geologic dip (Figure 6c). Although the magnitude of the noise imprint in the preconditioned gradient is slightly minimized compared with the unconstrained gradient, there has been horizontally orientated smearing of the stratigraphic layers and noise content. The resolution of some of the thinner dipping features has been impacted by this smearing, and there is deterioration from the noise-free unconstrained case in Figure 6c. On the right side of the model in which the geology is quite close to horizontal, the reconstruction of subtle stratigraphic layering is good. Several details attenuated in the dipping beds are preserved when the dip is used to correctly align the correlation operator (Figure 6d). An issue with this improved result is that as we move from one fault block the assumption of a large correlation length along the dip is no longer valid: The geology does not vary slowly parallel with the dip due to the discontinuity across the faults. When we decrease the anisotropy of the correlation operator near the faults, we are able to sharpen the terminations on either side (Figure 6e). This result is due to the fact that, at the faults, the correlation filter is mild and it penalizes wave-numbers of all orientations equally. It is important to note that, by relaxing the correlation lengths at the fault, we do allow some additional noise into the inversion. This can be seen when we compare our preconditioned FWI results for the noised (Figure 6e) and noise-free (Figure 6f)  $\mathbf{d}_{\text{obs}}$  cases.

To more clearly understand how the different gradient preconditioning operators affect the reconstructed FWI velocity model, we

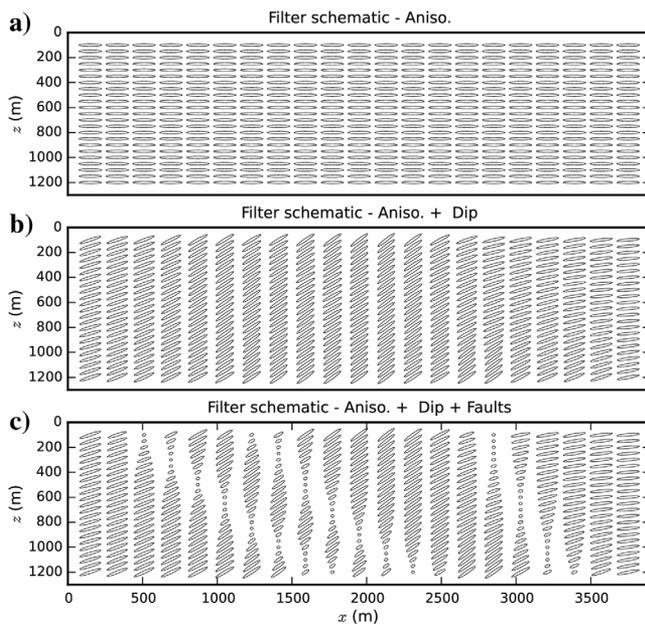


Figure 5. A graphical illustration of the anisotropic correlation operator applied to precondition the FWI result at each iteration. Shown is the results for (a) anisotropic operator ( $l_x = 7.5l_z$ ). (b) Anisotropic operator ( $l_x = 7.5l_z$ ) aligned to the local dip. (c) Anisotropic operator ( $l_x = 7.5l_z$ ) aligned to the local dip and relaxing to a weaker isotropic filter at the fault locations. The ellipses represent a constant correlation weight relative to the center point.

show the result after 30 iterations for all the cases (Figure 7). In the first case when no preconditioning is applied (Figure 7b), a lot of details related to the dipping stratigraphic structure and fault planes are visible in the gradient, but an additional overprinted spatially uncorrelated noise appears. With no constraint to prevent this random noise mapping into the model space, its amplitude will increase as we iterate. In the second case, the filter application with zero-dip values decreases the imprint of the noise and also affects much of the wanted stratigraphic details and faults. By including the true-dip values, we are able to recover more stratigraphic details from dipping events, but only when we include the variable  $l_x$  field are the fault successfully terminations reconstructed. It is important to note

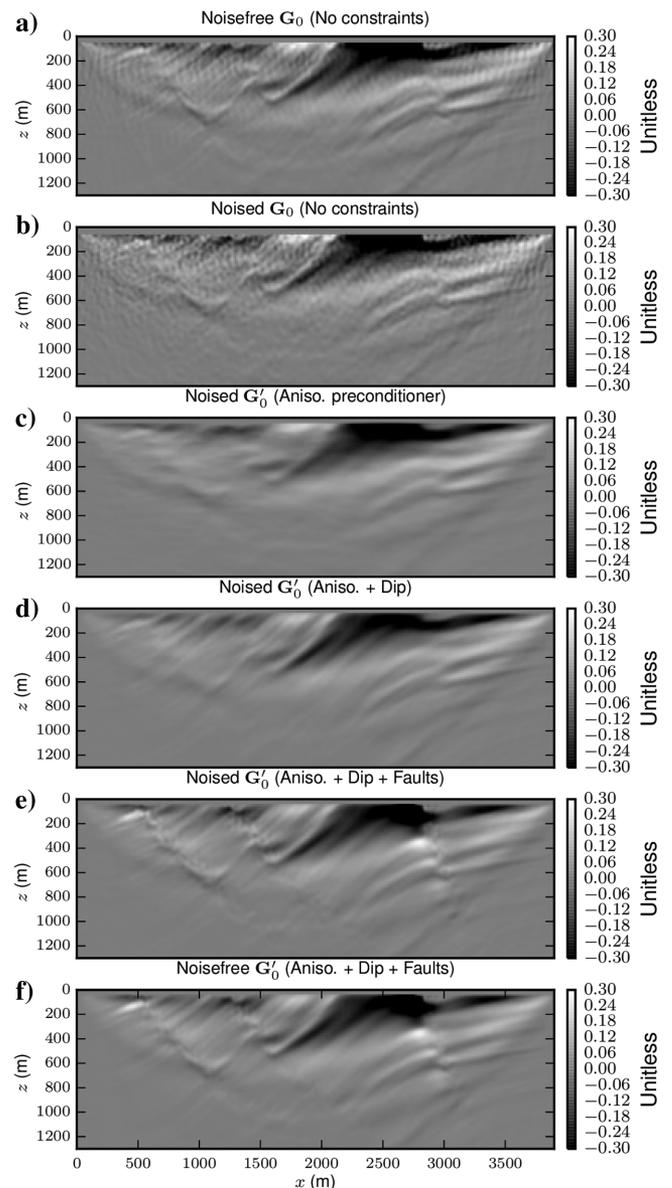


Figure 6. Gradient at the first iteration. (a) Noise-free, no preconditioning. (b) Noised, no preconditioning. (c) Noised, anisotropic preconditioner. (d) Noised, anisotropic preconditioner aligned to dip. (e) Noised, anisotropic preconditioner aligned to dip and honoring faults. (f) Noise-free, anisotropic preconditioner aligned to dip and honoring faults.

that the computational cost of the filtering operation is insignificant when compared with the cost of the FWI gradient computation (<1%). If there is a desire to add additional fine detail, the preconditioner can be further relaxed or alternatively removed for the later iterations.

The convergence behavior of our examples is shown in Figure 8. It is evident that the strong noise limits our ability to significantly decrease the normalized objective function. The noise-free example, though, is able to decrease the objective function by more than 90%. A good sign for the gradient preconditioning strategy is that the convergence rate for all cases does not appear to be severely hampered by the correlation operator.

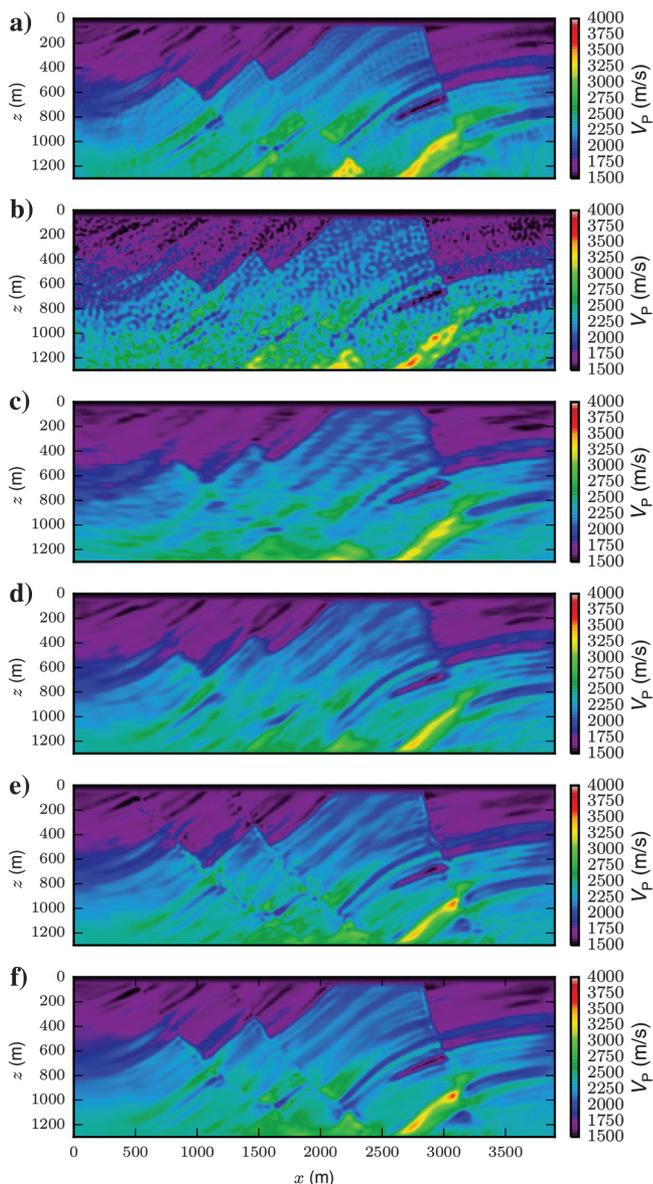


Figure 7. Inversion result after 30 iterations. (a) Noise-free, no preconditioning. (b) Noised, no preconditioning. (c) Noised, anisotropic preconditioner. (d) Noised, anisotropic preconditioner aligned to dip. (e) Noised, anisotropic preconditioner aligned to dip and honoring faults. (f) Noise-free, anisotropic preconditioner aligned to dip and honoring faults.

## Australian Broadseis data set

The potential benefit of our preconditioned FWI is now evaluated on a 2D field example. Our data set comprises a 31.25 km long 2D Broadseis data set acquired with an 8 km streamer. This legacy data set is acquired across the northwestern Australian continental shelf margin. Our investigation involves the imaging improvement that could be obtained by applying FWI in this environment with a focus on the target level at approximately 3 km depth. The anisotropic preconditioning operator is used to explore how effectively we can mitigate undesired artifacts that manifest in the reconstructed velocity parameter from using an increasingly sparse number of shots. Our workflow could save significant computation time for 3D and even elastic FWI without significantly degrading the resulting image. This allows for a significantly decreasing cost of performing FWI.

Previous legacy work on this data set includes a conventional time imaging workflow comprising designature, noise attenuation, stacking velocity analysis, and prestack time migration (preSTM). A full-angle stack of the preSTM (Figure 9) shows the transition of the water depth from the shallow continental shelf (<100 m) on the right of the line to the deeper (>1000 m) water on the left. A series of rifted Triassic fault blocks visible from approximately 3500 ms occur in the Triassic Mungaroo Formation, a known and highly productive hydrocarbon play. Two large vertically elevated Triassic horst blocks are annotated in Figure 9 (the black arrows). In both of these horsts, gas fields were discovered with traps formed due to favorable up-dip juxtaposition of the gas sands in the horst with shales in the adjacent down-thrown grabens. The Triassic Mungaroo Formation exhibits a significant increase in interval velocity relative to the strata above due to a much deeper burial history. The geology up to the seafloor is fairly benign and of constant thickness except for a few key features. A high-velocity feature is present on the right side of the line (the pink arrows). This feature is of importance and will be discussed in detail later because it limits the penetration of diving wave energy. The line also crosses the shelf margin (the red

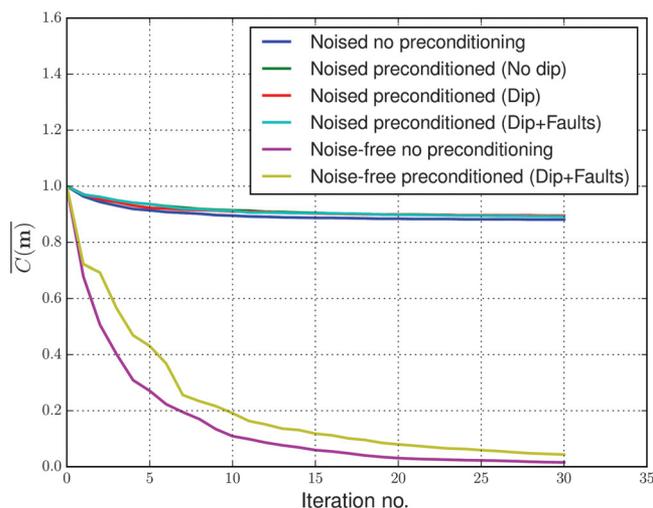


Figure 8. Convergence curve for the inversion results after 30 iterations. A much greater reduction in the normalized objective function  $C(\mathbf{m})$  is obtained for the noise-free cases. The general rate of convergence for the noise-free and noised case is not significantly affected by our gradient preconditioning strategy.

arrows), which presents imaging issues due to the sharp velocity variations over a distance of less than 8 km acquired streamer length. Many of these issues are further complicated due to anticipated “out-of-plane” 3D effects. These will be ignored for this 2D FWI study.

For FWI application, an initial depth velocity model is built up by merging expendable bathythermographic data that describe velocity variation in the water column with a 2 km coarse grid of stacking velocities (for an extensive review, see [Wellington et al., 2015](#)). The resultant interval velocity model is gridded at 25 m and is shown in Figure 9 stretched back to time. This initial velocity model lacks resolution for several features previously discussed. The sharpness and morphology of the water bottom velocity contrast appear to be less than what is seen in the prestack time migration (PSTM) image. Some of the details of the high-velocity zone to the right of the line are also absent. The sharpness of these features within the velocity model should not be expected when using a time-imaging workflow.

For the data input into FWI workflow, minimal preprocessing is performed. Data taken from the navigation merged field tapes have a spectral whitening process and a 3D-to-2D correction ( $T^2$ ) gain applied. FWI is performed in the time domain using a multiscale workflow ([Bunks et al., 1995](#)). We invert over six second-order Butterworth frequency bands. The lowest frequency introduced into the inversion is 3 Hz with the highest one being 12.5 Hz. At the beginning of the inversion, for each frequency band, source estimation is performed using the standard approach ([Pratt, 1999](#)) over the entire shot record. Once the source is inverted, it is kept constant for the entire inversion of the band and velocity and density are inverted simultaneously using an LBFGS algorithm. The inverted velocity result from each band serves as the input to the subsequent frequency band, whereas the density model at each frequency band is initialized

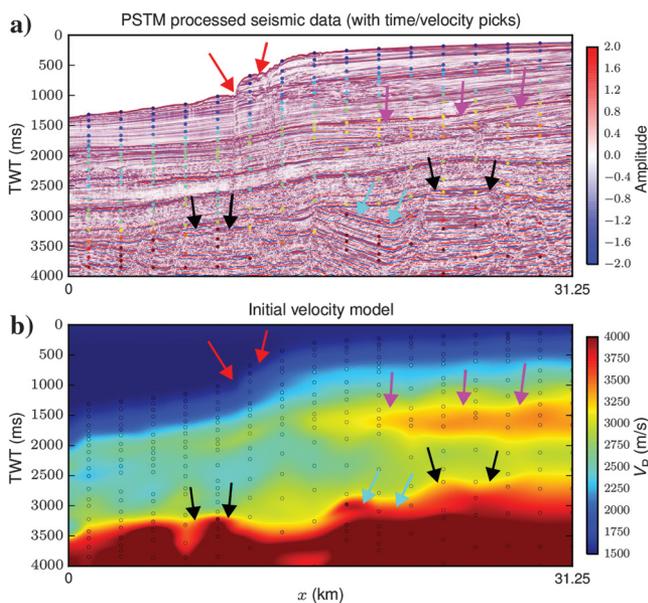


Figure 9. Initial model construction for NWA-006 Broadseis Line. (a) PSTM seismic section (in time) with interpreted time-velocity picks on a 2 km grid. (b) Initial velocity model built from gridding time-velocity picks that are shown in the above figure (the color scale matches the color of the interval velocity picks shown above). Key features are highlighted with the colored arrows: Triassic horst gas fields, black; key Triassic grabens, blue; shelf edge transition, red; and shallow high-velocity zone, pink.

from a constant density model. Using the density in this way allows an effective means for inverting for the density contrast at the seafloor, which we expected to be quite rugose. Several elastic amplitude effects expected to be present in the inversion are also anticipated to be partially absorbed by the density parameter ([Przebindowska et al., 2012](#)).

In this example, FWI has been performed with slightly different configurations regarding the number of modeled shots. The data set has a shot spacing of 17.5 m and a stable inversion result can be obtained using every fifth shot in the computation of the gradient. Figure 10a shows the gradient computed using every fifth shot. This gradient is computed from the first iteration of the first frequency band. When this shot increment is increased (Figure 10b), vertically oriented artifacts are added. As in the Marmousi case presented previously, if no effort is made to attenuate these features, they will manifest in the reconstructed acoustic model  $V_p$ . To mitigate these artifacts, we precondition the FWI gradient with our anisotropic correlation function using correlation lengths of  $l_x = 300$  m and  $l_z = 25$  m. The selection of the vertical correlation length  $l_z$  in this case is set to match the velocity grid cell size (25 m), whereas the horizontal length  $l_x$  was chosen from trial and error. As was discussed in the Marmousi case, it likely makes more sense to vary this horizontal correlation length based on the local velocity field and the highest wavenumber we should expect in our inversion given the propagation frequency. Investigation of the gradient wavenumber spectrum at early iterations would likely provide good clues on how this should be parameterized. For simplicity, these considerations are ignored for the case we show and a constant value of the correlation length is applied. The implementation of this operator is identical to the Marmousi case; however, due to the limited structural dip ( $<5^\circ$ ), there is no need to rotate the correlation filter. This filtering strategy allows for a strong attenuation of the vertically

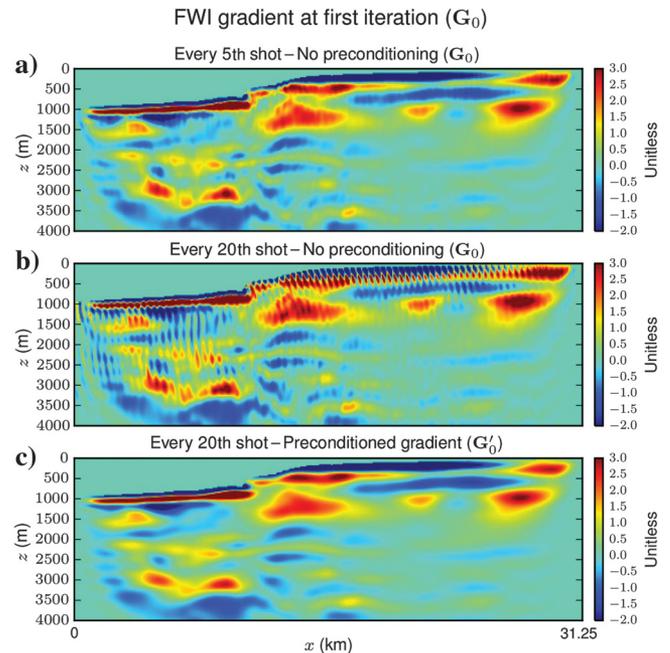


Figure 10. The effect of shot decimation on the gradient at the first iteration of FWI. (a) Every 5th shot, (b) every 20th shot, and (c) every 20th shot with gradient preconditioning ( $l_x = 300$  m and  $l_z = 25$  m).

orientated aliasing features while maintaining the geologic detail that we are looking to reconstruct in our model  $V_p$ .

The inversion results shown (Figure 11) are obtained after approximately 100 iterations, with the inversion of each band terminating when the decrease in the objective function goes below a set threshold or there is a failure in the line-search calculation. The results clearly show the effect of the shot decimation. The inverted velocity model has a lot of undesired artifacts introduced due to the vertically oriented aliasing artifacts mapping into the velocity parameter at each iteration. By applying the correlation filter, we are able to mitigate these artifacts without compromising the velocity model reconstruction. This decreases the cost of the FWI gradient computation by a factor of four. A clearer way of how understanding what details FWI has added to our initial model is shown in Figure 12. This plot illustrates the difference between the initial model  $V_p$  and our inverted results. Unlike the FWI gradient shown at the first iteration (Figure 10), we here see the total changes to the velocity model because we iterate and sweep through higher frequency bands in FWI. The strength of the vertically orientated noise in the shot decimated case is very clear when no preconditioning has been applied. The similarity between the undecimated shot and the decimated shot case with preconditioning is quite striking. This is especially interesting because it suggests that, after approximately 100 iterations, the preconditioning filter has not corrupted the FWI result but has instead acted as a necessary, but relatively mild, constraint that locally attenuates undesired wavenumbers. There are also some interesting features in the velocity model in the every fifth shot case (undecimated). These can be seen near the water bottom in the shallow water on the right and also on the left side of the line. Many of these almost vertically orientated features can be attributed to local differences in illumination and sampling issues that manifest in the gradient imaging condition. The wavenumber spectrum of the total update (Figure 13) high-

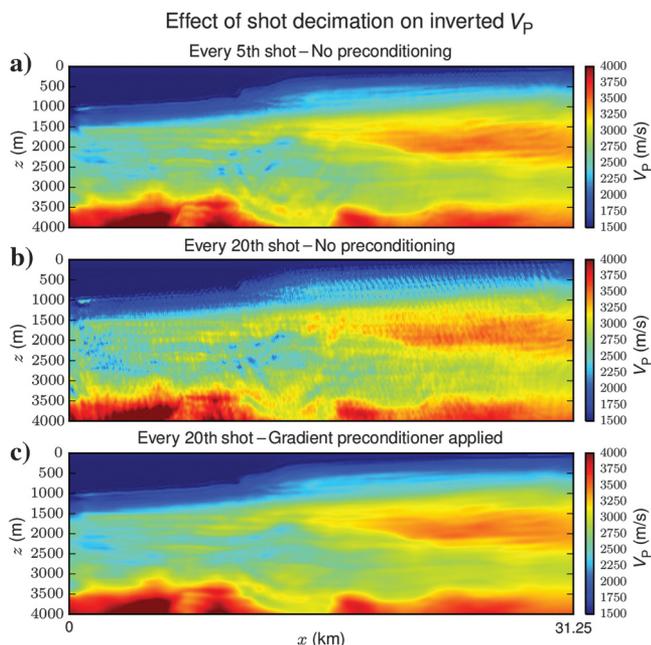


Figure 11. The effect of shot decimation on the final FWI result. (a) Every fifth shot, (b) every 20th shot, and (c) every 20th shot with gradient preconditioning ( $l_x = 300$  m and  $l_z = 25$  m).

lights that the dominant portion in all cases is the vertically orientated wavenumbers close to  $K_x = 0$ . Both cases without preconditioning have strong intermediate to high  $K_x$  features. These are most likely the aliasing effects we see in the imaging condition coming from too coarser shot increment or the local differences in illumination that map into the model parameter reconstruction due to the ill posed nature of the inversion problem. We restricted these wavenumbers via applying a horizontal correlation length and can see improvements in our preconditioned result. As was the case in the Marmousi example, the path and rate of convergence of the objective function (Figure 14) is similar for all the cases.

To better understand the illumination differences, we look at a shot record taken from the location of the Triassic gas field on the left side of the line (Figure 15). The shot record compares the match to real data before versus after FWI. The improvements to the match are seen on the near and far offsets. Two key events on the far offsets are shown in the green (the first arrival) and the red (a deep refracted arrival). Note that, after FWI, there is a good match between the modeled and true shot records. These transmitted events are interpreted to be responsible for most of the significant improvements to the velocity model. Ray-tracing analysis performed in the initial velocity model (Figure 16) highlights the penetration of the first arrivals (red) as well as showing where there will be later transmitted waves received within the acquisition spread on the far left of the line. These refracted arrivals are present due to the lack of a large velocity inversion on the left side of the line. The presence of this refracted arrival coincides with the low-wavenumber uplift we see in our FWI velocity update as well as some of the vertically orientated features that have manifested themselves in the velocity model. The preconditioned result has mitigated against these non-geologic vertical features while preserving the information parallel with the geologic strata.

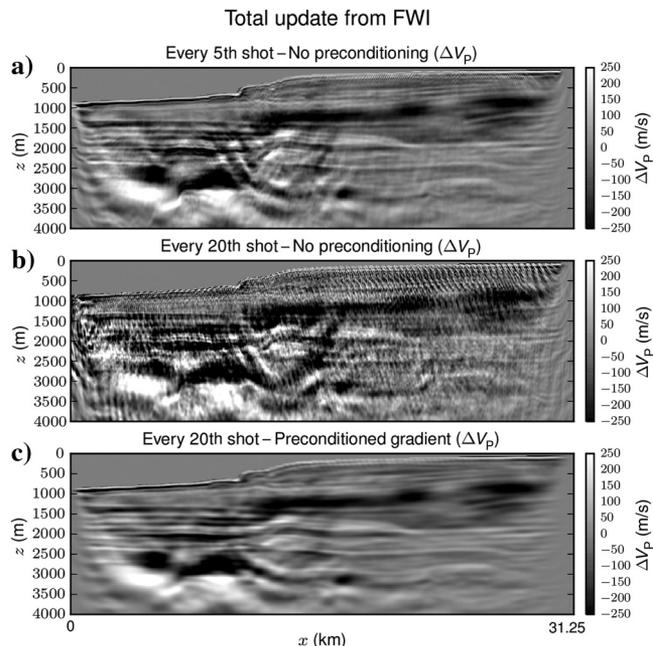


Figure 12. The difference between the FWI result and the initial velocity model for (a) FWI performed using every 5th shot, (b) FWI performed using every 20th shot, and (c) FWI performed using every 20th shot with gradient preconditioning applied.

To assess the improvement that the updated velocity models have provided in the image domain, an RTM has been performed on common-offset sections to allow generation of a stack and common-angle gathers. This was computed over the length of the entire line using the data that were input into the final frequency band of FWI (max frequency of 12.5 Hz). These are computed for the initial velocity model in addition to the various FWI results (Figure 17). One of the key areas of improvement is highlighted by the red arrows, in which the top of the Triassic gas field on the left side of the line is better reconstructed on the stacks and flatter on the common-angle gathers. This is interpreted as a satisfactory FWI result when applied to a 2D line from a rather crude initial model.

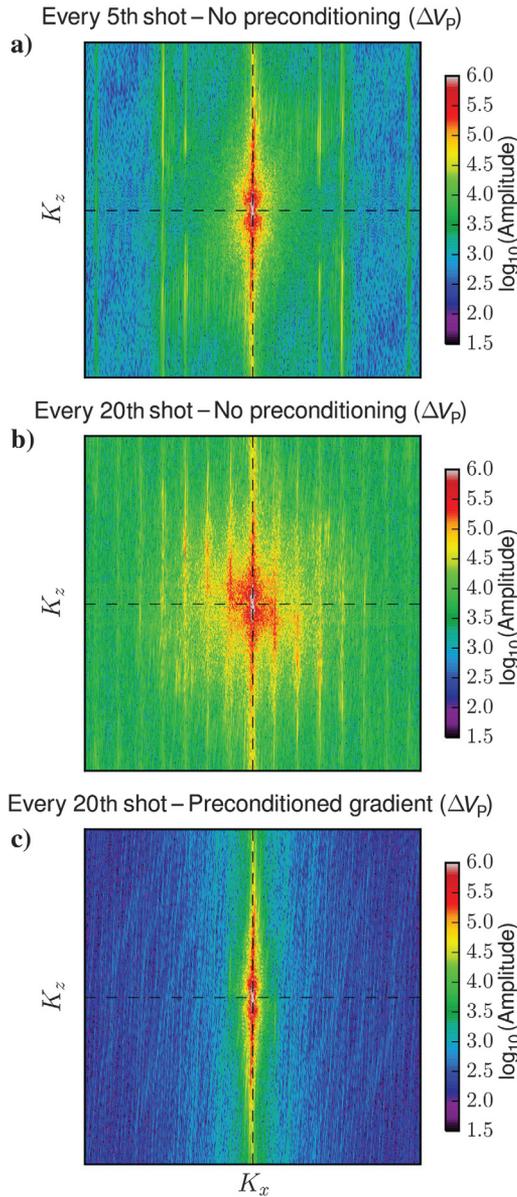


Figure 13. The wavenumber spectrum of the difference between the FWI result and the initial velocity model for (a) FWI performed using every 5th shot, (b) FWI performed using every 20th shot, and (c) FWI performed using every 20th shot with gradient preconditioning applied. The dashed black lines show the location of the zero wavenumber in the  $K_x$  and  $K_z$  orientations.

DISCUSSION

The two discussed examples highlight how undesired artifacts from noise and aliasing can manifest in the reconstructed velocity

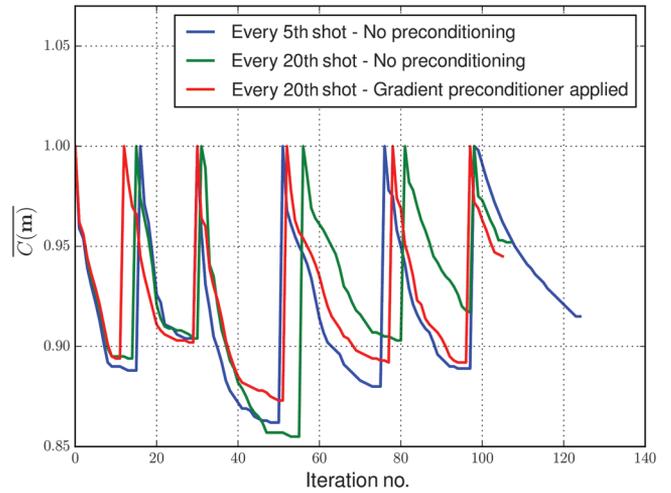


Figure 14. Convergence curve for the inversion results. The normalized objective function  $C(\mathbf{m})$  is shown for each case and returns to 1.0 at the start of each frequency band input into the inversion.

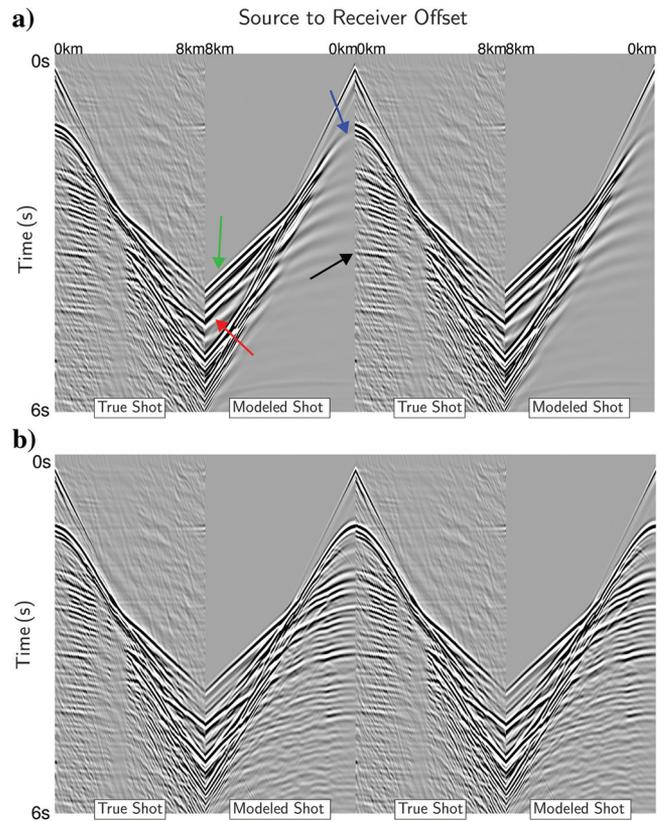


Figure 15. A comparison of a shot at the location of the left gas bearing Triassic horst from (a) before and (b) after the FWI process. The colored arrows highlight several key points of improvement. Blue, water bottom reflection; black, near-offset reflections at the top of the Triassic horst; green, diving wave; and red, postcritical reflection from near top of the Triassic horst.

parameter if prior information is not used to diminish them. Our presented preconditioning strategy is attractive due to several notable advantages that it possesses. The filter behind the preconditioning approach is an efficient directional low-pass filter, which means that, if there were uncertainty over the position of faults, the stratigraphic orientation, or the discontinuities' location, longer correlation lengths could be used to initially add incremental details to the velocity model prior to using shorter and more deterministic correlation length and dip fields. The ability to use variable correlation lengths, and dip fields at the later iterations means that, as the position of faults and fine detail become apparent, the smoothing constraints can be relaxed at these local regions to maximize the reconstruction of high wavenumbers near the discontinuities. At much later iterations, it is an option to remove preconditioning

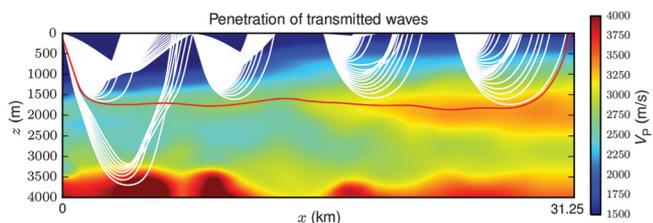


Figure 16. Penetration of transmitted waves modeled using the initial velocity model and acquisition parameters of the recorded shot gathers. The red line represents the maximum penetration of the first arrival calculated using an eikonal solver for all seismic shots. The white lines are the path of the transmitted rays for four equally spaced shots. There is good agreement between the transmitted and eikonal solutions except for on the left side of the line in which we have postcritical reflected energy returning (see Figure 15).

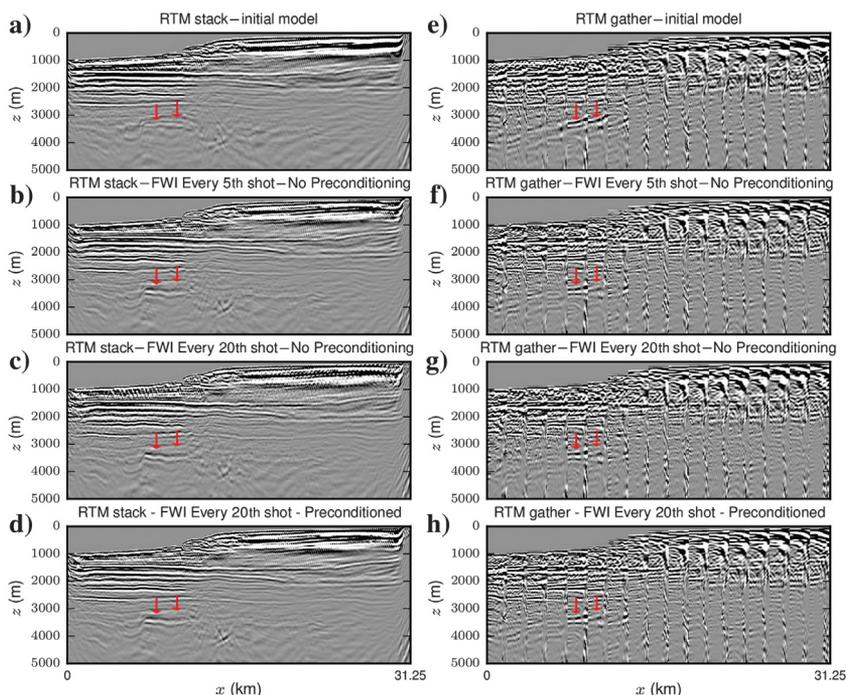


Figure 17. The (a-d) RTM common-angle stack and (e-h) RTM common-angle gathers computed with (a and e) the initial velocity model and (b and f) FWI result using every 5th shot. (c and g) FWI result using every 20th shot. (d and h) FWI result using every 20th shot with gradient preconditioning. The red arrows highlight a Triassic gas fields.

all-together and run the inversion with the unpreconditioned gradient. Although not implemented here, it is also likely that the horizontal correlation length could be automatically parameterized based on the maximum expected wavenumber we should expect given the local velocity and propagation frequency.

Another important result is that in the case of the Marmousi model, when the correct prior information was included, the model parameter reconstruction is very similar to the result without preconditioning. This suggests that our strategy is not overly aggressive and does not drastically alter the optimization path of the model evolution. This is also supported from the convergence curves presented for the synthetic and real data examples.

A second strength of our preconditioning approach stems from a shortcoming of FWI relative to other velocity model building. FWI is computationally intensive due to the cost of the forward modeling of the source and adjoint fields required to calculate the FWI gradient. A significant ongoing effort within the research community is focused on how to decrease this cost. Two key strategies to tackle this problem are (1) increasing the speed of modeling of an individual shot or (2) decreasing the number of shots needed for FWI gradient estimation via source encoding strategies (Herrmann et al., 2009; Krebs et al., 2009). Both of these strategies have potential, but they probably do not provide a complete answer especially in the cases of sparse acquisition, large 3D environments, and in elastic FWI, in which cost and practicalities mean that imaging artifacts can leak into the FWI gradient computation and subsequently the model parameter reconstruction. The 2D field data example we show highlights that we can mitigate against aliased information and imaging artifacts. There is no reason why this method will not be successful in 3D and elastic configurations in which the computational cost is much higher. The computational cost of applying the correlation filter in 3D on regular finite-difference grids typically used for acoustic FWI (Wellington et al., 2017) and also in irregular spectral element method grids for elastic FWI (Trinh et al., 2017) is very small relatively to the cost of the gradient.

## CONCLUSION

An efficient preconditioning strategy for FWI, based on a differential formulation of a local, multidimensional anisotropic correlation filter operator is promoted on 2D synthetic and real examples. A discretization of this differential expression leads us to a linear system to be efficiently solved by a conjugate gradient technique.

By using limited or detailed prior information, we are able to remove unwanted features of the FWI reconstruction coming from data noise, acquisition (sub)-sampling, and inherent over-discretized model representation of FWI protocol. A key benefit of our approach is that the removal of these unwanted artifacts does not appear to involve the need for additional iterations or significantly greater computational effort. Further real data examples, extensions to 3D, and other uses of the multidimensional correlation operators are foreseen in future studies, whereas the extensions to 3D elastic FWI workflow are also in progress.

## ACKNOWLEDGMENTS

This study was partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>), sponsored by AKERBP, CCG, CHEVRON, EQUINOR, EXXON-MOBIL, JGI, PETROBRAS, SCHLUMBERGER, SHELL, SINOPEC, and TOTAL. This study was granted access to the HPC resources of the Froggy platform of the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>), which is supported by the Rhône-Alpes region (grant no. CPER07\_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56), and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the programme Investissements d'Avenir supervised by the Agence Nationale pour la Recherche, and the HPC resources of CINES/IDRIS/TGCC under the allocation 046091 made by GENCI. The authors would like to acknowledge fruitful discussions made and assistance provided by L. Métivier, P.-T. Trinh, and O. Hamitou. I would also like to thank C. Dinneen for casting his careful eyes over the manuscript providing invaluable technical and typographic suggestions. A special mention is reserved for CCG who provided the real data set and helped with some premigration processing in addition to some reflection tomography tests.

## DATA AND MATERIALS AVAILABILITY

Data associated with this research are confidential and cannot be released.

## REFERENCES

- Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2012, Regularized full waveform inversion including prior model information: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, W031.
- Borisov, D., and S. C. Singh, 2015, Three-dimensional elastic full waveform inversion in a marine environment using multicomponent ocean-bottom cables: A synthetic study: *Geophysical Journal International*, **201**, 1215–1234, doi: [10.1093/gji/ggv048](https://doi.org/10.1093/gji/ggv048).
- Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: *Geophysics*, **74**, no. 6, WCC105–WCC118, doi: [10.1190/1.3215771](https://doi.org/10.1190/1.3215771).
- Bunks, C., F. M. Salek, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: *Geophysics*, **60**, 1457–1473, doi: [10.1190/1.1443880](https://doi.org/10.1190/1.1443880).
- Claerbout, J. F., 1992, *Earth sounding analysis*: Blackwell Scientific Publications.
- Clapp, R. G., 2001, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.
- Fichtner, A., B. L. N. Kennett, H. Igel, and H. P. Bunge, 2010, Full waveform tomography for radially anisotropic structure: New insights into present and past states of the Australasian upper mantle: *Earth and Planetary Science Letters*, **290**, 270–280, doi: [10.1016/j.epsl.2009.12.003](https://doi.org/10.1016/j.epsl.2009.12.003).
- Fomel, S., 2002, Applications of plane-wave destruction filters: *Geophysics*, **67**, 1946–1960, doi: [10.1190/1.1527095](https://doi.org/10.1190/1.1527095).
- Guitton, A., 2012, Blocky regularization schemes for full waveform inversion: *Geophysical Prospecting*, **60**, 870–884, doi: [10.1111/j.1365-2478.2012.01025.x](https://doi.org/10.1111/j.1365-2478.2012.01025.x).
- Guitton, A., G. Ayeni, and E. Díaz, 2012, Constrained full-waveform inversion by model reparameterization: *Geophysics*, **77**, no. 2, R117–R127, doi: [10.1190/geo2011-0196.1](https://doi.org/10.1190/geo2011-0196.1).
- Hale, D., 2007, Local dip filtering with directional Laplacians: Technical report, Centre for Wave Phenomena, Colorado School of Mines.
- Hansen, C., 1998, Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion: SIAM, *Mathematical Modeling and Computation*.
- Hansen, P. C., 2010, *Discrete inverse problems: Insight and algorithms*: SIAM.
- Harlan, W., 1995, Regularization by model reparameterization, <http://www.billharlan.com/pub/papers/regularization.pdf>, accessed 1 June 2014.
- Herrmann, F. J., Y. A. Erlangga, and T. T. Y. Lin, 2009, Compressive simultaneous full-waveform simulation: *Geophysics*, **74**, no. 4, A35–A40, doi: [10.1190/1.3115122](https://doi.org/10.1190/1.3115122).
- Krebs, J., J. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M. D. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: *Geophysics*, **74**, no. 6, WCC105–WCC116, doi: [10.1190/1.3230502](https://doi.org/10.1190/1.3230502).
- Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: *Conference on Inverse Scattering, Theory and Application*, SIAM, 206–220.
- Loris, I., G. Nolet, I. Daubechies, and F. A. Dahlen, 2007, Tomographic inversion using l1-norm regularization of wavelet coefficients: *Geophysical Journal International*, **170**, 359–370, doi: [10.1111/j.1365-246X.2007.03409.x](https://doi.org/10.1111/j.1365-246X.2007.03409.x).
- Martin, G. S., R. Wiley, and K. J. Marfurt, 2006, Marmousi2: An elastic upgrade for Marmousi: *The Leading Edge*, **25**, 156–166, doi: [10.1190/1.2172306](https://doi.org/10.1190/1.2172306).
- Métivier, L., F. Breteau, R. Brossier, S. Operto, and J. Virieux, 2014, Full waveform inversion and the truncated Newton method: Quantitative imaging of complex subsurface structures: *Geophysical Prospecting*, **62**, 1353–1375, doi: [10.1111/1365-2478.12136](https://doi.org/10.1111/1365-2478.12136).
- Métivier, L., and R. Brossier, 2016, The SEISCOPE optimization toolbox: A large-scale nonlinear optimization library based on reverse communication: *Geophysics*, **81**, no. 2, F1–F15, doi: [10.1190/geo2015-0031.1](https://doi.org/10.1190/geo2015-0031.1).
- Nocedal, J., and S. J. Wright, 2006, *Numerical optimization*, 2nd ed.: Springer.
- Operto, S., A. Miniussi, R. Brossier, L. Combe, L. Métivier, V. Monteiller, A. Ribodetti, and J. Virieux, 2015, Efficient 3-D frequency-domain monoparameter full-waveform inversion of ocean-bottom cable data: Application to Valhall in the visco-acoustic vertical transverse isotropic approximation: *Geophysical Journal International*, **202**, 1362–1391, doi: [10.1093/gji/ggv226](https://doi.org/10.1093/gji/ggv226).
- Peter, D., D. Komatitsch, Y. Luo, R. Martin, N. Le Goff, E. Casarotti, P. Le Lohier, F. Magnoni, Q. Liu, C. Blitz, T. Nissen-Meyer, P. Basini, and J. Tromp, 2011, Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes: *Geophysical Journal International*, **186**, 721–739, doi: [10.1111/j.1365-246X.2011.05044.x](https://doi.org/10.1111/j.1365-246X.2011.05044.x).
- Peters, B., and F. J. Herrmann, 2017, Constraints versus penalties for edge-preserving full-waveform inversion: *The Leading Edge*, **36**, 94–100, doi: [10.1190/le36010094.1](https://doi.org/10.1190/le36010094.1).
- Plessix, R. E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: *Geophysical Journal International*, **167**, 495–503, doi: [10.1111/j.1365-246X.2006.02978.x](https://doi.org/10.1111/j.1365-246X.2006.02978.x).
- Plessix, R. E., and C. Perkins, 2010, Full waveform inversion of a deep water ocean bottom seismometer dataset: *First Break*, **28**, 71–78, doi: [10.3997/1365-2397.2010013](https://doi.org/10.3997/1365-2397.2010013).
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain. Part 1: Theory and verification in a physical scale model: *Geophysics*, **64**, 888–901, doi: [10.1190/1.1444597](https://doi.org/10.1190/1.1444597).
- Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362, doi: [10.1046/j.1365-246X.1998.00498.x](https://doi.org/10.1046/j.1365-246X.1998.00498.x).
- Przebindowska, A., A. Kurzmann, D. Köhn, and T. Bohlen, 2012, The role of density in acoustic full waveform inversion of marine reflection seismics: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.20148728](https://doi.org/10.3997/2214-4609.20148728).
- Sambridge, M., and K. Mosegaard, 2002, Monte Carlo methods in geophysical inverse problems: *Reviews of Geophysics*, **40**, 1–29, doi: [10.1029/2000RG000089](https://doi.org/10.1029/2000RG000089).
- Sirgue, L., O. I. Barkved, J. Dellinger, J. Etgen, U. Albertin, and J. H. Kommedal, 2010, Full waveform inversion: The next leap forward in imaging at Valhall: *First Break*, **28**, 65–70, doi: [10.3997/1365-2397.2010012](https://doi.org/10.3997/1365-2397.2010012).
- Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: *Geophysics*, **69**, 231–248, doi: [10.1190/1.1649391](https://doi.org/10.1190/1.1649391).
- Soubaras, R., and P. Whiting, 2011, Variable depth streamer — The new broadband acquisition system: 81st Annual International Meeting, SEG, Expanded Abstracts, 4349–4353, doi: [10.1190/1.3628115](https://doi.org/10.1190/1.3628115).
- Tape, C., Q. Liu, A. Maggi, and J. Tromp, 2010, Seismic tomography of the southern California crust based on spectral-element and adjoint methods: *Geophysical Journal International*, **180**, 433–462, doi: [10.1111/j.1365-246X.2009.04429.x](https://doi.org/10.1111/j.1365-246X.2009.04429.x).
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266, doi: [10.1190/1.1441754](https://doi.org/10.1190/1.1441754).
- Tikhonov, A., and V. Arsenin, 1977, *Solution of ill-posed problems*: Winston.
- Toksöz, M. N., and D. H. Johnston, 1981, *Seismic wave attenuation*: SEG 2, *Geophysics Reprint Series*.
- Trinh, P. T., R. Brossier, L. Métivier, L. Tavard, and J. Virieux, 2019, Efficient time-domain 3D elastic and viscoelastic full-waveform inversion using a spectral-element method on a flexible Cartesian-based mesh: *Geophysics*, **84**, no. 1, R61–R83, doi: [10.1190/geo2018-0059.1](https://doi.org/10.1190/geo2018-0059.1).
- Trinh, P. T., R. Brossier, L. Métivier, J. Virieux, and P. Wellington, 2017, Bessel smoothing filter for spectral element mesh: *Geophysical Journal International*, **209**, 1489–1512, doi: [10.1093/gji/ggx103](https://doi.org/10.1093/gji/ggx103).

- van den Berg, P. M., and A. Abubakar, 2001, Contrast source inversion method: State of the art: *Progress in Electromagnetics Research*, **34**, 189–218, doi: [10.2528/PIER01061103](https://doi.org/10.2528/PIER01061103).
- Vigh, D., K. Jiao, D. Watts, and D. Sun, 2014, Elastic full-waveform inversion application using multicomponent measurements of seismic data collection: *Geophysics*, **79**, no. 2, R63–R77, doi: [10.1190/geo2013-0055.1](https://doi.org/10.1190/geo2013-0055.1).
- Virieux, J., and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC1–WCC26, doi: [10.1190/1.3238367](https://doi.org/10.1190/1.3238367).
- Vogel, C., 2002, *Computational methods for inverse problems*: SIAM.
- Vogel, C. R., and M. E. Oman, 1996, Iterative methods for total variation denoising: *SIAM Journal on Scientific Computing*, **17**, 227–238, doi: [10.1137/0917016](https://doi.org/10.1137/0917016).
- Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Stekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: *Geophysics*, **78**, no. 2, R59–R80, doi: [10.1190/geo2012-0338.1](https://doi.org/10.1190/geo2012-0338.1).
- Wellington, P., R. Brossier, S. Garambois, and J. Virieux, 2015, 2D acoustic time domain full waveform inversion – a broadband application in the Carnarvon basin, Australia: 77th EAGE Conference and Exhibition, Expanded Abstracts, doi: [10.3997/2214-4609.201413060](https://doi.org/10.3997/2214-4609.201413060).
- Wellington, P., R. Brossier, O. Hamitou, S. Garambois, and J. Virieux, 2016, Preconditioning FWI with the sparse additive inverse Laplacian correlation function: 86th Annual International Meeting, SEG, Expanded Abstracts, 1462–1466, doi: [10.1190/segam2016-13841790.1](https://doi.org/10.1190/segam2016-13841790.1).
- Wellington, P., R. Brossier, O. Hamitou, P. Trinh, and J. Virieux, 2017, Efficient anisotropic dip filtering via inverse correlation functions: *Geophysics*, **82**, no. 4, A31–A35, doi: [10.1190/geo2016-0552.1](https://doi.org/10.1190/geo2016-0552.1).
- Zhu, H., E. Bozdağ, D. Peter, and J. Tromp, 2012, Structure of the European upper mantle revealed by adjoint tomography: *Nature Geoscience*, **5**, 493–498, doi: [10.1038/ngeo1501](https://doi.org/10.1038/ngeo1501).