

# **Full-waveform inversion uncertainty analysis with null-space shuttles**

Scott Keating and Kris Innanen

## **ABSTRACT**

Full-waveform inversion is an effective tool for recovering subsurface information, but quantification of confidence in this information can be very difficult. Uncertainty in a global sense is ever-present when using local optimization, preventing the calculation of an absolute uncertainty. Even when considering local uncertainty, the large dimensionality of the problem means that feasible, intelligible confidence metrics are generally limited to providing a scalar description of confidence for each variable, while a vector is required for completeness. Fortunately, complete characterization of uncertainty is seldom necessary from an applications perspective. More often, the uncertainty in a specific aspect of the inversion is important (for instance, confidence in a recovered anomaly). Here, we investigate the use of null-space shuttles to characterize the maximal change in a chosen metric that can be achieved without altering the inversion objective function. This provides a quantification of the uncertainty in this metric for the inversion result.

## **INTRODUCTION**

Full-waveform inversion (FWI) is a seismic inversion technique which attempts to recover subsurface properties through a data-fitting procedure using the bulk of the measured data (Tarantola, 1984). Substantial successes in P-wave velocity-model building have been achieved using FWI, and multi-parameter implementations, which include more subsurface information, are constantly improving (Virieux and Operto, 2009; Pan et al., 2018). While the results of FWI have consistently improved, the development of tools to assess our confidence in these results has largely been lacking. Recent attempts to improve resolution and uncertainty analyses have improved the state of affairs (e.g Fichtner and Leeuwen, 2015; Thurin et al., 2019), but the problem of uncertainty quantification remains largely unsolved.

Uncertainty quantification in full-waveform inversion (FWI) is challenging. In principle, we seek to determine the confidence with which the properties defined in the inversion result have been determined. In a given inversion problem, the uncertainty identifies the range of possible solutions which could have satisfied the data and prior-fitting requirements of the inversion. Ideally, such an uncertainty estimate would also identify the probabilities of these solutions.

Unfortunately, uncertainty estimation of this kind lies far beyond the reach of modern FWI. The definition above does not preclude points of model space associated with different minima than the that of the inversion result, but these require global assessment of the objective function and this is typically far beyond computational feasibility. Less ambitious approaches attempt to quantify uncertainty only locally, that is, they consider only the uncertainty in in the model-space location of a local minimum. Ideally, this type of uncertainty is measured about the minimum in question, and employs the second derivatives of the problem (the Hessian matrix) to identify the relative uncertainty of different model-

space directions (Tarantola, 2005). Directions with large curvature are less uncertain, as the objective function changes (and, as a minimum is considered, necessarily increases) faster in these directions, while those with smaller curvature are associated with smaller changes in the objective function, and so, greater uncertainty. Two major concerns exist for this approach. Firstly, quantifying the curvature in all possible directions generally greatly surpasses available computational resources, and secondly, the inversion result is generally not a minimum of the objective function. This second point means that directions of large curvature may in reality be those in which the objective decreases fastest: curvature alone cannot be completely informative about uncertainty.

In reality, uncertainty in an arbitrary direction of model space is not usually a useful quantity. The features of an inversion result which impact economic decision-making are generally small in number and highly specific. By focusing the quantification of uncertainty on the subset of model-space directions which have an important impact on decision-making, we can dramatically reduce the computational demands of uncertainty characterization. This allows for more accurate methods to be employed.

In this report, we consider a method for uncertainty quantification with respect to a user-defined scalar function of model-space location. The scalar function defines the extent to which a decision-affecting feature of the inversion result is present. We consider the maximal uncertainty in this feature to be the model-space step which minimizes the function while remaining in the acceptable-solution region of model space. This region should include, at minimum, all points which achieve an equivalent or lower objective function value than the inversion output. Quantifying uncertainty then requires that we explore a subset of the points of model space with the same objective function value as the inversion result. Using the terminology of Deal and Nolet (1996), we refer to the model-space steps which preserve the objective function as ‘null-space shuttles’.

## THEORY

Consider a model output from a greedy FWI inversion,  $m$ . This model corresponds to the lowest value of the objective function which was achieved in the FWI procedure,  $\phi$ , but is not, in general, an exact minimum (these being unfeasible to compute). Suppose that another model,  $m^*$ , corresponding to objective value  $\phi^*$ , is proposed as an alternative to  $m$  as a possible output from the inversion. If  $\phi^*$  is greater than  $\phi$ , then  $m$  will be the preferred inversion output. If  $\phi^*$  is less than  $\phi$ , however,  $m^*$  would generally be preferred to  $m$  by the inversion. The set of models  $\mathbf{m}$ , which contains all possible models  $m^*$  for which  $\phi^* \leq \phi$ , represents the set of models which would be indistinguishable from, or preferred to  $m$  by the inversion procedure. In this sense, there is uncertainty in the inversion between  $m$  and  $\mathbf{m}$ . Quantifying this uncertainty requires that  $\mathbf{m}$  be characterized.

Complete characterization of  $\mathbf{m}$  is computationally infeasible. This is highlighted by the fact that at least one of the models in  $\mathbf{m}$  is an exact minimum of the objective function. As exact minima are generally considered infeasible to compute, the set containing them must also be assumed out of reach. Calculation of elements of  $\mathbf{m}$ , however, is highly feasible. Any sufficiently short step in a model-space direction with a negative projection onto the gradient should result in a model within  $\mathbf{m}$ ; the objective function must go down at

least slightly for some step-length. Very short steps are uninteresting from the standpoint of uncertainty characterization however, as these represent very small changes in the model. What we are most interested in is the maximum uncertainty in a given direction, that is, how large a model change in a given direction could be while still resulting in an objective function value of  $\phi$  or less.

Assuming the FWI result  $m$  lies near a minimum, the objective function near  $m$  can be characterized as a function of model change  $\Delta m$  as

$$\Phi(\Delta m) \approx \phi + g_0 \Delta m + \frac{1}{2} \Delta m^T H_{GN} \Delta m, \quad (1)$$

where  $g_0$  is the gradient at  $m$ , and  $H_{GN}$  is the Gauss-Newton approximation of the Hessian at  $m$ . Because  $H_{GN}$  is positive-definite, the contribution of the third term in equation 1 is always positive. This means that the second term,  $g_0 \Delta m$ , must be negative in order for the objective to decrease or remain stationary. Consequently, the largest step that can be taken in any given direction without increasing the objective function under this approximation will always be a step resulting in the same objective function: any step decreasing the objective could be made longer while still remaining in  $\mathbf{m}$ . We refer to the set of models with an objective function  $\phi$  as  $\mathbf{m}'$ .

Insofar as equation 1 is an accurate representation of the objective function, the model updates,  $\Delta m'$  corresponding to element models of  $\mathbf{m}'$  satisfy

$$g_0 \Delta m' = -\frac{1}{2} \Delta m'^T H_{GN} \Delta m'. \quad (2)$$

For a model-space vector  $\delta m$ , this allows for the calculation of the maximum step parallel to  $\delta m$  expected to lie in  $\mathbf{m}$ :

$$\Delta m' = \alpha \delta \hat{m}, \quad (3)$$

where

$$\delta \hat{m} = \frac{\delta m}{\|\delta m\|_2}, \quad (4)$$

and

$$\alpha = \frac{-2g_0 \delta \hat{m}}{\delta \hat{m}^T H_{GN} \delta \hat{m}}. \quad (5)$$

In this way, given the vector  $\delta m$ , we can approximate the inversion uncertainty in that direction at the computational cost of one Hessian-vector product evaluation:  $H_{GN} \delta \hat{m}$ . A Hessian-vector product can be evaluated at the cost of three wavefield propagations (Métivier et al., 2013), one of which is already performed during the gradient calculation, and can be stored to prevent recalculation. This computation requirement is small relative to the inversion procedure, but if many directions are considered, more Hessian-vector products are required, and the cost can become substantial. Fortunately, the uncertainty of the model in a random direction is not typically of great importance. Instead, the key uncertainties in an inversion result will typically relate to important, specific features of the model, involving relatively few model-space directions.

Suppose there exists a scalar function  $\psi$  which describes a key feature of the model. For instance, the key feature may be an anomalously low  $v_P/v_S$  ratio in a certain region of

the model. The metric  $\psi$  should be defined such that it is large when the feature is present in the model, but small otherwise. We could consider the square of the difference between the  $v_P/v_S$  ratio in the region of interest and the average ratio in the model. The problem of uncertainty quantification with respect to  $\psi$  is that of finding the model in  $\mathbf{m}'$  which minimizes  $\psi$ :

$$\min_{m'} \psi(m') \quad (6)$$

Using the relation defined in equation 3 to connect a given direction  $\delta m$  with the corresponding model in  $\mathbf{m}'$ , we can re-state the optimization problem as

$$\min_{\delta m} \psi(m + \Delta m'(\delta m)) = \psi(m + \alpha \delta \hat{m}). \quad (7)$$

Solving this problem through gradient-based optimization techniques requires the calculation of

$$\frac{\partial \psi}{\partial \delta m} = \frac{\partial \psi}{\partial \Delta m'} \left( \frac{\partial \Delta m'}{\partial \alpha} \frac{\partial \alpha}{\partial \delta m} + \frac{\partial \Delta m'}{\partial \delta \hat{m}} \frac{\partial \delta \hat{m}}{\partial \delta m} \right). \quad (8)$$

This expression can be expanded to

$$\frac{\partial \psi}{\partial \delta m} = \frac{\partial \alpha}{\partial \delta m} * \left( \frac{\partial \psi}{\partial \Delta m'}^T \delta \hat{m} \right) + \alpha \left( \frac{1}{\|\delta m\|_2} \frac{\partial \psi}{\partial \Delta m'} - \frac{\delta m}{\|\delta m\|_2^3} \frac{\partial \psi}{\partial \Delta m'}^T \delta m \right), \quad (9)$$

where

$$\frac{\partial \alpha}{\partial \delta m} = \frac{1}{\|\delta m\|_2} \frac{\partial \alpha}{\partial \delta \hat{m}} - \frac{\delta m}{\|\delta m\|_2^3} \frac{\partial \alpha}{\partial \delta \hat{m}}^T \delta m \quad (10)$$

and

$$\frac{\partial \alpha}{\partial \delta \hat{m}} = -\frac{2g_0}{\delta \hat{m}^T H_{GN} \delta \hat{m}} + 4H_{GN} \delta \hat{m} \frac{g_0 \delta \hat{m}}{\delta \hat{m}^T H_{GN} \delta \hat{m}}. \quad (11)$$

The somewhat cumbersome choice of  $\delta m$  as the optimization variable rather than the unit vector  $\delta \hat{m}$  is made here in order to allow for the use of unconstrained optimization techniques in the solution of equation 7.

An approximate solution to equation 7 can be found through nonlinear optimization. In this report, we used the L-BFGS approach (e.g. Nocedal and Wright, 2006) for the minimization. It is worth noting that the main computational cost of such an approach is the evaluation of the product  $H_{GN} \delta \hat{m}$ , which must be done once per iteration to evaluate  $\psi$  and  $\frac{\partial \psi}{\partial \delta m}$ , and of the product  $H_{GN} d_{BFGS}$  (where  $d_{BFGS}$  is the calculated L-BFGS descent direction) once per iteration for the line search. While the line-search step of the L-BFGS procedure may require many additional evaluations of  $\psi$  and  $\frac{\partial \psi}{\partial \delta m}$ , these will all be evaluated at locations in model space that are linear combinations of  $\delta \hat{m}$  and  $d_{BFGS}$ , so the required Hessian-vector products can be determined from those already calculated. As such, the main computational cost of each L-BFGS iteration is the evaluation of two Hessian-vector products.

Suppose an approximate minimizer of  $\psi$ ,  $\delta m'^*$ , has been determined and the associated  $\Delta m'^*$  has been calculated through equation 3. This represents our best estimate of the model update which maximally alters our interpretation metric without changing the FWI objective function, provided that equation 1 holds exactly (i.e. the objective is exactly linear). In reality, the objective function will not be exactly linear, even close to the

minimum. This has two major implications for our uncertainty estimate. Firstly, it means changing the model by  $\Delta m'$  will not leave the objective function exactly unchanged. To fix this problem, we suggest using a line-search approach to determine the step length in the direction of  $\Delta m^*$  corresponding to no change in the objective. The second implication of non-linearity is that  $H_{GN}$  becomes a function of  $m$ . This means that even an exact solution of the optimization problem in equation 7 may not be the solution to the true problem of interest, equation 6.

In most applications, it is likely that an approximate solution of equation 6 will be a negligible improvement over an approximate solution of equation 7; the FWI objective function should be close to linear when starting from the FWI output. If a solution of equation 6 is preferred, however, it can be treated as an iterative optimization procedure, in which the model updates are calculated through approximate minimization of equation 7. The procedure we consider is outlined in algorithm 1. In this algorithm, the inner loop corresponds to the iterative approximation of the step,  $\Delta m$ , which solves equation 7, while the outer loop attempts to solve equation 6 by updating the model in the direction of  $\Delta m$ , while keeping the FWI objective function constant. If the problem is close to linear, we can limit the procedure to only one outer loop iteration.

```

Input : FWI output:  $m$ 
Output: Minimizer of  $\psi$ :  $m'_{i_{max}+1}$ 
Initialize model estimate
 $m'_1 = m$ 
for  $i = 1, \dots, i_{max}$  do ; // Loop over outer iterations

    Initialize the descent direction  $\delta m_1$ 
    Initialize the L-BFGS inverse Hessian approximation  $Q$ 
    for  $j = 1, \dots, j_{max}$  do ; // Loop over inner iterations

        Set  $H_{v_1} = H_{GN}(m'_i)\delta m_j$  ; // Calculate Hessian - update product
        Set  $g_\psi = \frac{\partial \psi}{\partial \delta m_1}$  ; // Calculate gradient using  $H_{v_1}$ 
        Set  $d = Qg_\psi$  ; // Calculate L-BFGS descent direction
        Set  $H_{v_2} = H_{GN}(m'_i)d$  ; // Calculate Hessian - d product
        Calculate the step length  $\mu$  to minimize  $\psi$  using a line-search
        Use  $H_{v_1}$  and  $H_{v_2}$  for  $\psi$  evaluations in line-search
        Set  $\delta m_{j+1} = \delta m_j + \mu d$  ; // Update the step direction
        Update  $Q$ 
    end
    Set  $\Delta m = \alpha \delta \hat{m}_{j_{max}+1}$ 
    Calculate the step length  $\lambda$  for  $\phi^*(m'_i + \Delta m = \phi)$ 
    Set  $m'_{i+1} = m'_i + \lambda \Delta m$ 
end

```

**Algorithm 1:** Algorithm for  $\psi$  minimization through null-space shuttling.

In the following section, we perform numerical tests to demonstrate this approach.

## NUMERICAL EXAMPLES

For numerical tests of this approach, we consider FWI uncertainty for a synthetic toy model. We use the visco-elastic inversion approach outlined by Keating et al. (2018). The elastic properties of this model are shown in figure 1. This model is based on the Marmousi model, and contains a layer structure which defines similar geometry for each physical property considered. The ratios of these properties are variable throughout the model, however, and differ substantially from the background trends in an anomalous region at about 200 m depth, and 600 m in  $x$ -position. Our goal is to characterize the uncertainty of an inversion based on this model in key model-space directions. The model is defined with a resolution of 10 m in both  $x$  and  $z$  directions. Data up to 20 Hz is assumed to be available in the inversion.

We will characterize uncertainty for multiple acquisition geometries for this problem. For the sake of comparison, we will consider an inversion result which is very similar for each acquisition geometry. To provide this baseline inversion result, we consider an inversion with a relatively complete acquisition geometry - explosive sources and receivers placed at all four edges of the model. On the both the top and bottom edges of the model, 74 evenly spaced explosive sources and 148 evenly spaced multi-component receivers are considered. Similarly, the left and right edges have 29 sources and 58 receivers each. The initial model for  $v_P$ ,  $v_S$ , and  $\rho$  in the inversion was obtained by averaging the true model in the  $x$  direction. The low  $v_P$  and  $\rho$  anomaly was not included in this averaging, and so is not represented in the background model. The initial models for  $Q_P$  and  $Q_S$  were set as constants, given the difficulty in estimating these properties. The background model is shown in figure 2.

In this report we consider frequency-domain inversion. For the baseline inversion, ten frequency bands of five evenly-spaced frequencies were considered, starting with 1-2 Hz, and ending with 1-20 Hz. At each of the first nine frequency bands, one iteration of truncated-Gauss-Newton (TGN) optimization was performed (Métivier et al., 2013), with a maximum of 20 inner loop iterations per FWI iteration. At the last frequency band, which we consider as the available data for uncertainty characterization, 11 iterations of TGN optimization were considered.

The baseline inversion result obtained using this approach is shown in figure 3. This is a relatively strong inversion result, as expected given the comprehensive acquisition geometry and relatively intensive optimization strategy. Used as an inversion output to be considered in our uncertainty characterization, it should represent at least a fair approximation of what an FWI stopping point may be for the other acquisition geometry we consider; at worst it may be an unachievable result for other acquisitions, but not to such an extent that we cannot glean insights from our uncertainty analysis. To ensure an appropriate FWI output is considered, we perform an additional 10 iterations of L-BFGS optimization at the 1-20 Hz frequency band for each specific acquisition geometry considered to give the inversion output for that acquisition. The resulting model changes are small in all cases considered here.

For the examples considered here, we use the null-space shuttling approach outlined

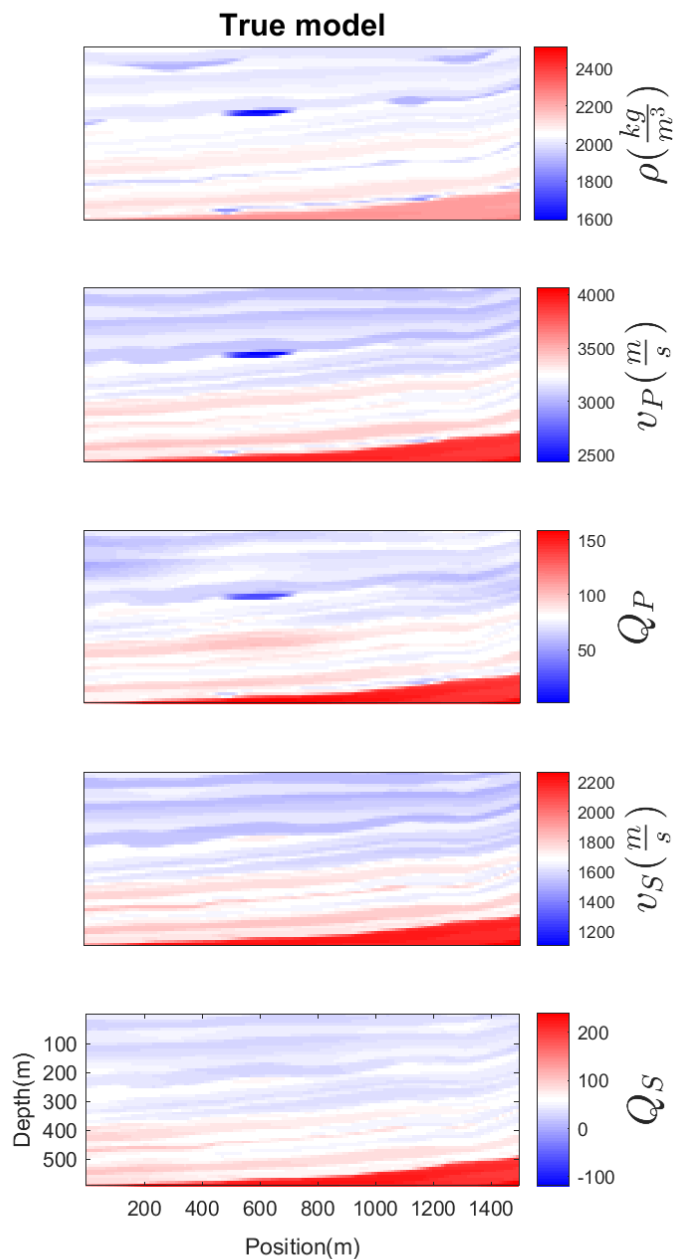


FIG. 1. True model for synthetic tests, defined in terms of  $\rho$ ,  $Q$  values, and  $v_P$ ,  $v_S$  at reference frequency  $\omega_0$ .

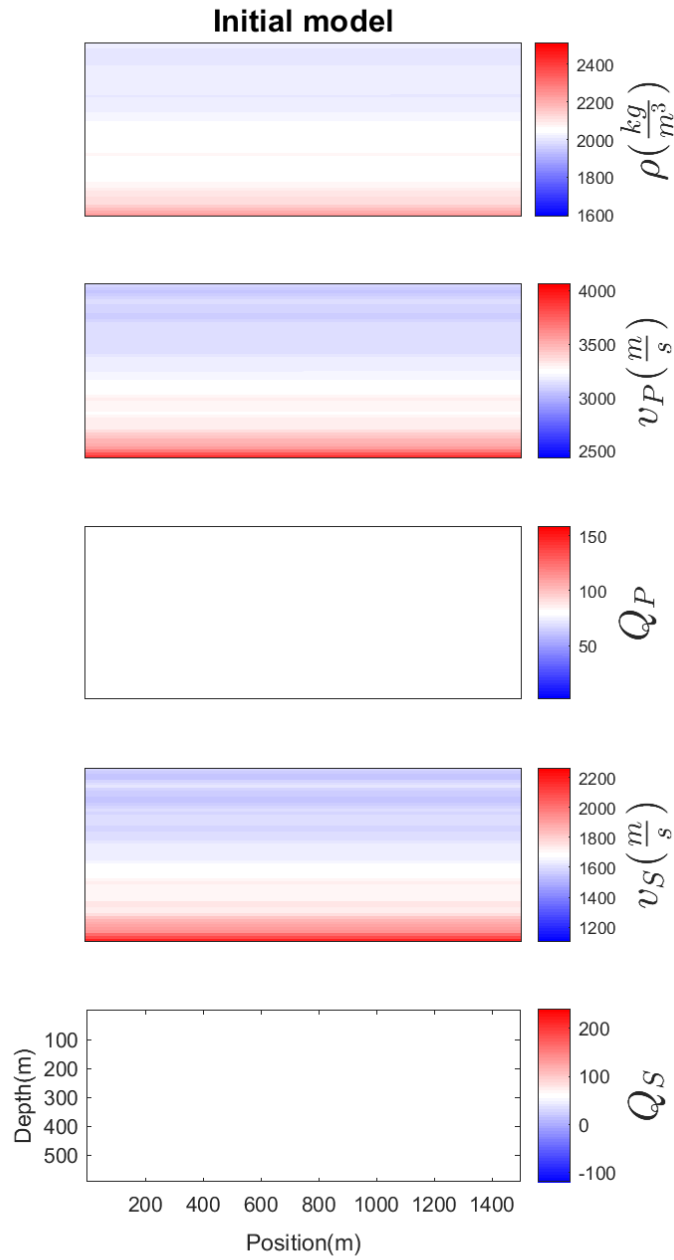


FIG. 2. Initial model for baseline inversion, defined in terms of  $\rho$ ,  $Q$  values, and  $v_P$ ,  $v_S$  at reference frequency  $\omega_0$ .



in algorithm 1, with 20 inner loop iterations per outer iteration. This is a relatively large number of iterations to use, and the computational cost is 40 Hessian-vector product evaluations per outer iteration. As a fraction of the total inversion cost, which included about 400 Hessian-vector product evaluations, however, this is a modest computational expense.

Perhaps the most notable feature in both the true model and the inversion result is the region of anomalously low  $\rho$ ,  $v_P$  and  $Q_P$ . Suppose that the presence or absence of this anomaly is crucial to the interpretation of the inversion result. In this case, we may be relatively uninterested in the uncertainty of the inversion in an arbitrary direction of model space, but very intensely interested in the confidence with which this anomaly has been recovered. With this motivation, we define the interpretation metric  $\psi$  to be large when such a feature is present in the inversion, but negligible otherwise:

$$\psi = \sum_{i=1}^{x_i \in x^*} a (\rho(x_i) - \bar{\rho}(x_i))^2 + b (v_P^{-2}(x_i) - \bar{v}_P^{-2}(x_i))^2 + c (Q_P^{-1}(x_i) - \bar{Q}_P^{-1}(x_i))^2, \quad (12)$$

where  $a$ ,  $b$ , and  $c$  are scaling terms,  $x^*$  are the locations where the anomaly is recovered in the inversion, and  $\bar{\rho}$ ,  $\bar{v}_P^{-2}$ , and  $\bar{Q}_P^{-1}$  are average values of these properties outside of, but near, the anomaly location.

If  $a$ ,  $b$ , and  $c$  are chosen to weight the anomalies in each of  $\rho$ ,  $v_P$  and  $Q_P$  approximately equally, we can learn about the degree of certainty that there is some type of anomaly at that location. We calculate the optimal shuttle for this objective, using three outer-loop iterations, and considering a surface acquisition geometry: 74 evenly spaced explosive sources and 148 evenly spaced multi-component receivers are considered. The calculated shuttle for this choice of penalty term is shown in figure 4. This shuttle represents the step which should maximally reduce our choice of  $\psi$  without changing the objective function value. We can see that this shuttle represents a large change in  $Q_P$  and  $\rho$  at the location of the anomaly, representing large uncertainty in these parameters at this location. By contrast, the change in  $v_P$  is relatively small; the objective function insists on preserving a  $v_P$  low at this location. The model obtained after applying the null-space shuttle in figure 4 is shown in figure 5. With the  $L_2$  objective function we chose to use in this case, the objective function values of these two models differ by less than 1%. As such, the differences between these models represents an uncertainty in the inversion result: our objective function is unable to discriminate between these models.

The available data substantially determine the confidence of the inversion in different model features. To investigate this effect, we attempt to quantify the uncertainty of the same inversion output, in the same metric,  $\psi$ , but now consider data from evenly spaced sources and receivers on each boundary of the inversion, a near-ideal acquisition geometry. The equivalent-objective model obtained by applying the optimal calculated null-space shuttle is shown in figure 6. In this case, the  $\rho$  and  $v_P$  lows at the anomaly location are better preserved; these features are better constrained by the acquisition in this case. The continued absence of the  $Q_P$  low at the anomaly means that there is little confidence in the location of this anomaly, even with the more comprehensive acquisition. The ring of high attenuation around the anomaly region, however, might suggest that the inversion insists on increased attenuation at least nearby the anomaly in this case.

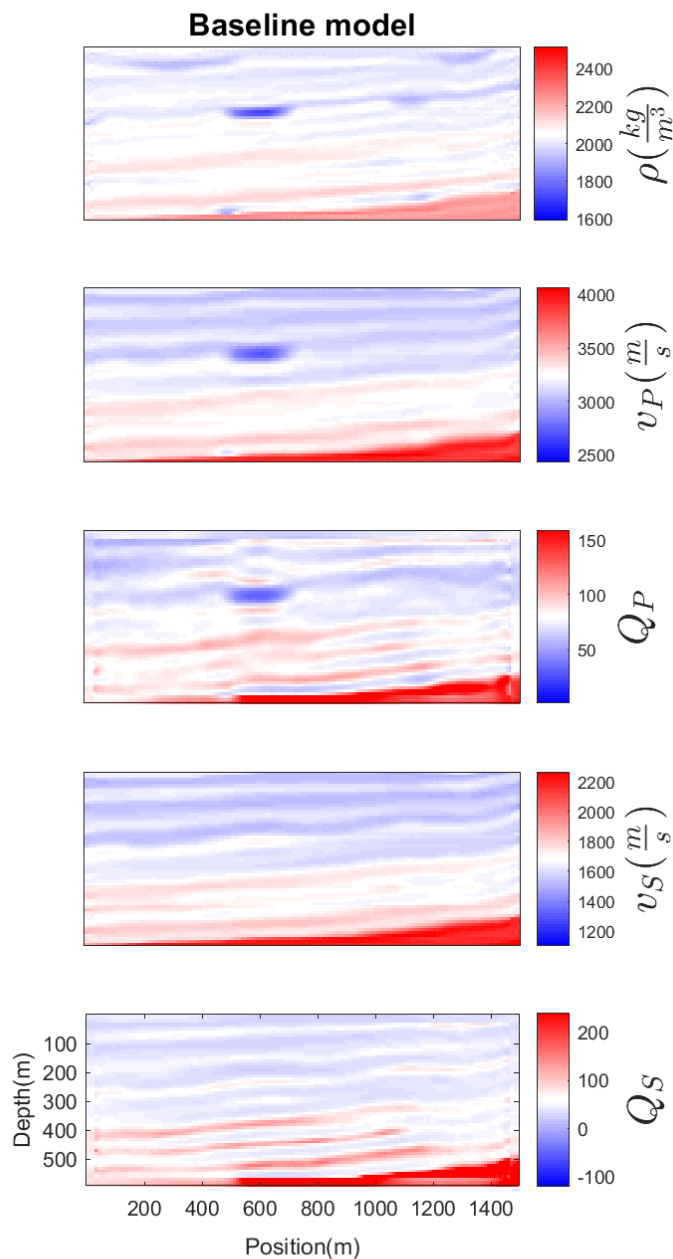


FIG. 3. Baseline inversion result for uncertainty characterization, defined in terms of  $\rho$ ,  $Q$  values, and  $v_P$ ,  $v_S$  at reference frequency  $\omega_0$ .

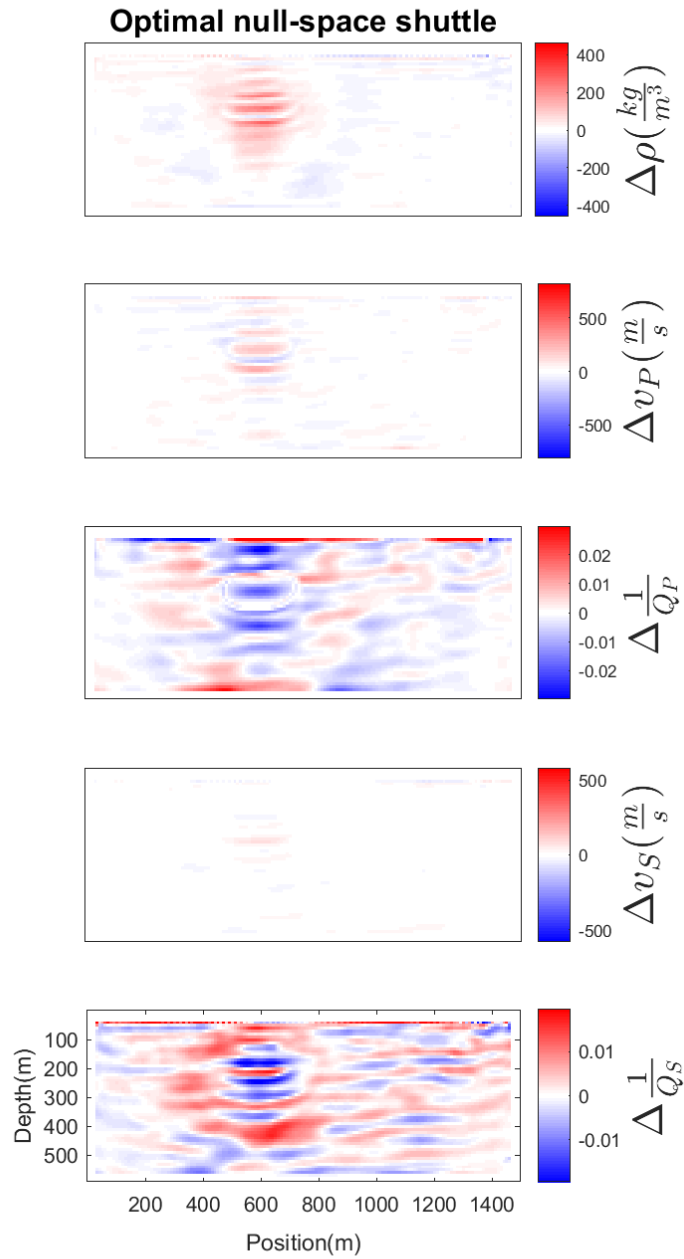


FIG. 4. Optimal null-space shuttle for eliminating anomaly in  $\rho$ ,  $v_P$  and  $Q_P$ , surface-only acquisition.

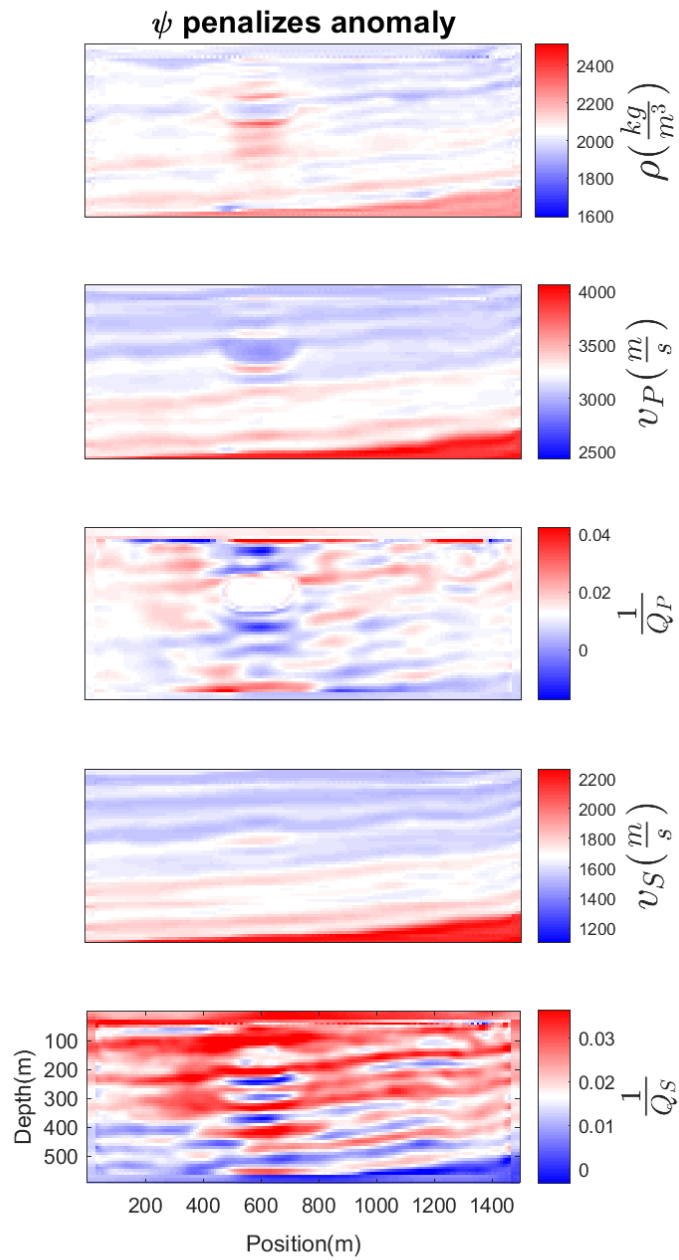


FIG. 5. Equivalent-objective model for eliminating anomaly in  $\rho$ ,  $v_P$  and  $Q_P$ , surface-only acquisition.

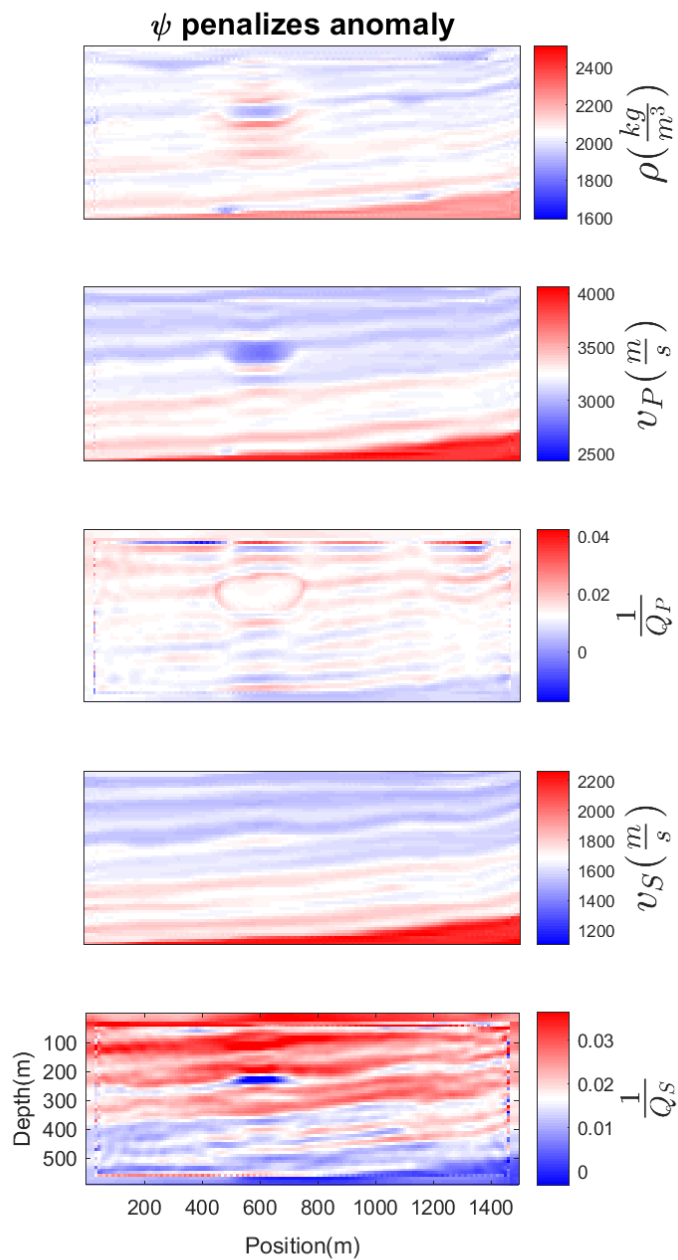


FIG. 6. Equivalent-objective model for eliminating anomaly in  $\rho$ ,  $v_P$  and  $Q_P$ , surround acquisition.

In a third example, we consider a case in which the  $v_P/v_S$  ratio at the anomaly has a substantial impact on the interpretation of the inversion result. To quantify the uncertainty in the  $v_P/v_S$  ratio at the anomaly, we define the inversion metric  $\psi$  as

$$\psi = \sum_{i=1}^{x_i \in x^*} - \left( \frac{v_P(x_i)}{v_S(x_i)} - \tilde{\gamma} \right)^2, \quad (13)$$

where  $\tilde{\gamma}$  is the average  $v_P$  to  $v_S$  ratio in the background model. With this metric, we can investigate the confidence of the inversion in the presence of a substantially different  $v_P/v_S$  ratio at the anomaly. We calculate the null-space shuttle which optimally reduces  $\psi$  using one outer-loop iteration of the approach described above, and apply it to the starting model to achieve the result in figure 7. The objective function of this model is 1.7% less than the initial objective, and so is in the plausible solution region. The  $v_P/v_S$  ratio is increased to an average of about 1.8 in this result, from its value of 1.6 in the inversion output, still substantially lower than its value of about 2.0 outside the anomaly. To this degree, there is uncertainty in the inversion result, but this means that there is high confidence in the presence of  $v_P/v_S$  ratio at least as low as 1.8 here.

## DISCUSSION

One objection to the approach adopted here for uncertainty quantification may be that the FWI objective function is not always a good measure of inversion quality. While two models may share the same objective function value, a typical inversion may not actually treat these model equally. This is because the value of the objective function is not often the main criterion used in determining convergence. Convergence of the inversion problem is typically instead gauged by the rate of objective function descent at each iteration, closely linked to the gradient of the problem. If two models share the same objective function value, but one has a small associated gradient and the other has a large gradient, we typically further iterate the large-gradient model, making the ambiguity between the two a strictly in-inversion phenomenon not corresponding to real inversion output ambiguity. In this case, it may not be the value of the objective, but the amplitude of the gradient which determines the stopping point of the FWI procedure. The approach outlined here could be adapted to such a definition relatively easily; if models with gradients of the same amplitude are those that could be confused in principle, it is easy to alter equations 3 - 5 accordingly. The Hessian still provides the wherewithal to approximate an appropriate step in this case.

## CONCLUSIONS

General uncertainty in FWI is very difficult to quantify, and generally requires large amounts of computation and storage for accurate treatment. The targeted approach to uncertainty quantification investigated here explores the uncertainty in a single scalar function of the model. If appropriately defined, this scalar can allow for uncertainty estimates for the key decision-altering features of an inversion result. The computational cost of these estimates is generally only a fraction of that of the full inversion.

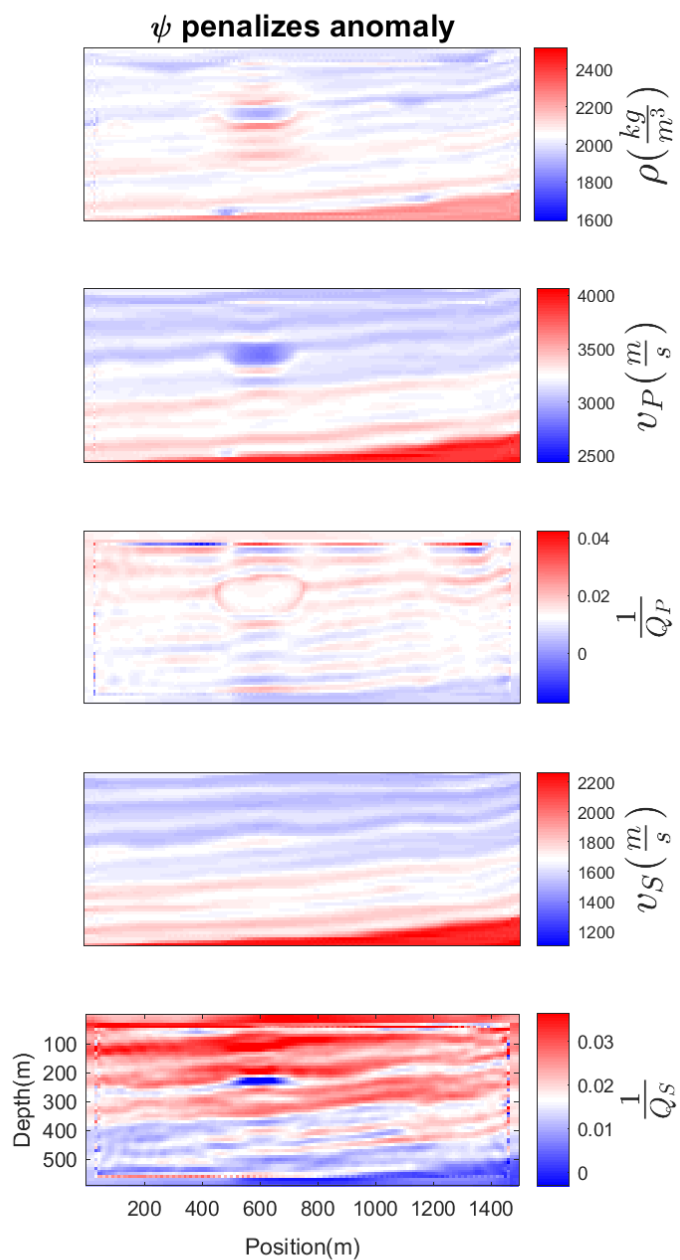


FIG. 7. Equivalent-objective model for eliminating anomaly in  $v_P/v_S$  ratio, surface-only acquisition.

## ACKNOWLEDGEMENTS

The authors thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13. Scott Keating was also supported by the Earl D. and Reba C. Griffin Memorial Scholarship.

## REFERENCES

- Deal, M. M., and Nolet, G., 1996, Nullspace shuttles: *Geophysical Journal International*, **124**, No. 2, 372–380.
- Fichtner, A., and Leeuwen, T. v., 2015, Resolution analysis by random probing: *Journal of Geophysical Research: Solid Earth*, **120**, No. 8, 5549–5573.
- Keating, S., Li, J., and Innanen, K. A., 2018, Viscoelastic fwi: solving for  $q_p$ ,  $q_s$ ,  $v_p$ ,  $v_s$ , and density: *CREWES Annual Report*, **30**.
- Métivier, L., Brossier, R., Virieux, J., and Operto, S., 2013, Full waveform inversion and the truncated newton method: *Siam J. Sci. Comput.*, **35**, B401–B437.
- Nocedal, J., and Wright, P. S., 2006, *Numerical Optimization*: Springer, 2nd edn.
- Pan, W., Geng, Y., and Innanen, K. A., 2018, Interparameter trade-off quantification and reduction in isotropic-elastic full-waveform inversion: synthetic experiments and hussar land data set application: *Geophys. J. Int.*, **213**, 1305–1333.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.
- Tarantola, A., 2005, *Inverse problem theory and methods for model parameter estimation*, vol. 89: siam.
- Thurin, J., Brossier, R., and Métivier, L., 2019, Ensemble-based uncertainty estimation in full waveform inversion: *Geophysical Journal International*, **219**, No. 3, 1613–1635.
- Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, No. 6, WCC1.