ADMM-BASED FULL-WAVEFORM INVERSION FOR MICROSEISMIC IMAGING

A PREPRINT

Hossein S. Aghamiry
University Cote d’Azur - CNRS - IRD - OCA, Geoazur, Valbonne, France. aghamiry@geoazur.unice.fr

Ali Gholami
Institute of Geophysics, University of Tehran, Tehran, Iran. agholami@ut.ac.ir

Stéphane Operto
University Cote d’Azur - CNRS - IRD - OCA, Geoazur, Valbonne, France. operto@geoazur.unice.fr

Alison Malcolm
Department of Earth Sciences, Memorial University of Newfoundland, St John’s NL A1B 3X5, Canada. amalcolm@mun.ca

May 3, 2021

ABSTRACT

Full waveform inversion (FWI) is beginning to be used to characterize weak seismic events at different scales, an example of which is microseismic event (MSE) characterization. However, FWI with unknown sources is a severely underdetermined optimization problem, and hence requires strong prior information about the sources and/or the velocity model. The frequency-domain wavefield reconstruction inversion method (WRI) has shown promising results to mitigate the nonlinearity of the FWI objective function that is generated by cycle-skipping. WRI relies on the reconstruction of data-assimilated wavefields, which approach the true wavefields near the receivers, a helpful feature when the source is added as an additional optimization variable. We present an adaptation of a recently proposed version of WRI based on the alternating direction method of multipliers (ADMM) that first finds the location of the MSEs and then reconstructs the wavefields and the source signatures jointly. Finally, the subsurface model is updated to focus the MSEs at their true location. The method does not require prior knowledge of the number of MSEs. The inversion is stabilized by sparsifying regularizations separately tailored to the source location and velocity model subproblems. The method is tested on the Marmousi model using one MSE and two clusters of MSEs with two different initial velocity models, an accurate one and a rough one, as well as with added noise. In all cases, the method accurately locates the MSEs and recovers their source signatures.

1 INTRODUCTION

Full waveform inversion (FWI) is a waveform matching procedure that is generally used to reconstruct subsurface models with wavelength resolution (Virieux & Operto, 2009). The FWI optimization problem is commonly solved assuming that the positions of the controlled sources are known, while the temporal source signature is estimated together with the subsurface model by variable projection or in an alternating scheme (Aravkin et al., 2012; Plessix & Cao, 2011; Pratt, 1999; Rickett, 2013). This paper focusses on the case where the source location and time signature are not known. The source can be active (e.g., explosion) or passive (e.g., earthquake, microseismic). Our numerical examples focus on characterizing weak seismic events at different scales, especially for microseismic event (MSE) imaging, but the method can easily be tailored to tectonic earthquakes. When estimating the source, we decompose it into a temporal signature and a spatial
During oil and gas production, CO\(_2\) can seriously affect the estimated model parameters, and so source estimation is an important step in successful FWI. Lee & Kim (2003) show that a small source error can seriously affect the estimated model parameters, and so source estimation is an important step in successful FWI. During oil and gas production, CO\(_2\) injection, and geothermal applications, fluid injections generate fracturing and cause small earthquakes (microseismicity). Locating the MSEs and studying their focal mechanisms are important in monitoring flow mobility and optimizing production. The seismic waves generated by these earthquakes are recorded continuously at the surface or in wells, and their traveltimes or waveforms are inverted to estimate the source time function (source signature) and the location of the MSEs, assuming a known subsurface model. This setting implies that the recorded wavefield has been triggered by a blended source with contributions from all of the MSEs.

Pioneering methods for microseismic location rely on picking the traveltimes of different P and S components to locate the MSEs (Han et al., 2009; Thurber & Engdahl, 2000). These methods are time-consuming and may not be robust in the presence of noise (Han et al., 2009; Thurber & Engdahl, 2000). More recently, wavefield-based imaging techniques have received a lot of attention for the relocation of earthquakes and MSEs (De Ridder & Maddison, 2018; Gajewski & Tressmer, 2005; Kaderli et al., 2015; McMechan, 1982; Michel & Tsvankin, 2014; Sharan et al., 2019; Shekar & Sethi, 2019; Sjögreen & Petersson, 2014; Song et al., 2019). The most basic approaches rely on time-reverse imaging, where the data are propagated backward in time using the adjoint of the wave-equation operator (Gajewski & Tressmer, 2005; McMechan, 1982; Nakata & Beroza, 2016; Sharan et al., 2019). The source location and the origin time are found by tracking where, in both space and time, the maximum focusing occurs. These approaches suffer from a lack of resolution and, as in any migration technique, require a kinematically accurate background velocity model. To mitigate these problems, there has been some interest in applying FWI for imaging such microseismicity (Li et al., 2020; Montgomery et al., 2010). Moving from migration-based techniques to FWI-based techniques opens the door to estimate the source signatures, the source locations and in some cases the subsurface properties, by solving a multivariate optimization problem. This is possible because all of the wave information is involved in the FWI procedure. This multivariate optimization problem attempts to reconstruct wavefields, sources, and model parameters from partial measurements of the wavefield. However, this problem is highly under-determined, even when the subsurface parameters are processed as inactive parameters (i.e. are not updated). To mitigate this, the method requires some priors like the number of events, their approximate locations, the sparsity of the source distribution or an accurate background velocity model to tighten the null space of the inverse problem (Kaderli et al., 2015, 2018; Sharan et al., 2019; Shekar & Sethi, 2019). Michel & Tsvankin (2014) reviewed how to compute the FWI gradient for source parameters in VTI media. Kaderli et al. (2015); Sharan et al. (2019); Shekar & Sethi (2019) proposed robust FWI algorithms with sparsifying regularization to refine the location of MSEs and estimate their temporal signatures. Their approaches first estimate the location of the events by wavefield extrapolation with sparsifying regularization. Following this, the source signatures of the identified events are reconstructed separately. These approaches do not require assumptions about the number or the nature of the events, but require a fairly accurate background velocity model.

A few studies have been presented to jointly locate the events and update the subsurface velocity model (Song et al., 2019; Sun et al., 2016). Sun et al. (2016) update the source and the velocity model in an alternating manner. The descent direction of the under-determined source estimation problem is preconditioned by a weighting of the source term inferred from a cross-correlation time-reversal imaging condition. Song et al. (2019) build the source image using the imaging condition of Nakata & Beroza (2016) and then update the velocity model by penalizing the energy of the source image away from the estimated source location with an annihilator. This approach requires one to process each event separately and hence to separately identify each event in the data. Kaderli et al. (2015) estimate the MSE locations and source signatures in an alternating way when the source is considered as a product of independent temporal source function and spatial source location function, while Sharan et al. (2019) update both of them jointly.

We extend these past works by using a more general form of FWI. In this more general form, FWI can be cast as a constrained optimization problem that aims to estimate the wavefields and the subsurface parameters by fitting the recorded data subject to the constraint that the wave equation is satisfied (Haber et al., 2000). This approach was developed because even when the source location and time signature are known, it is well established that FWI is highly nonlinear. Part of this nonlinearity can be viewed as arising when the full search space encompassed by the wavefields and the subsurface parameters is projected onto the subsurface parameter space. This happens via an elimination of the wavefield variables, by assuming that the wavefields exactly satisfy the wave-equation at each FWI iteration. This variable elimination makes FWI prone to cycle skipping when the initial model is not accurate enough to predict recorded traveltimes with an error smaller than half a period (Virieux & Operto, 2009). To avoid this projection, some approaches implement the wave equation as a soft constraint with a penalty method such that the data can be closely matched with inaccurate subsurface models from the early FWI iterations by not requiring that the wave equation be satisfied exactly (Abubakar et al., 2009; van Leeuwen & Herrmann, 2016, 2013). Then, the subsurface model is updated by solving an overdetermined quadratic optimization problem, which consists of minimizing the source residuals generated by the relaxation of the constraint that the wavefields exactly solve the wave equation. In these extended approaches, the wavefields are reconstructed by solving, in a least-squares sense, an overdetermined linear...
system comprised of the wave equation weighted by the penalty parameter and the observation equation relating the simulated wavefield to the data through a sampling operator. In other words, the wavefields are reconstructed with data assimilation, which makes them approach the true wavefields near the receivers, a helpful feature when the source is added as a new variable. This approach was called Wavefield Reconstruction Inversion (WRI) by van Leeuwen & Herrmann (2013). A variant of WRI, based upon the method of multipliers or augmented Lagrangian method, was proposed by Aghamiry et al. (2019b) to increase the convergence rate and decrease the sensitivity of the algorithm to the choice of the relaxation (penalty) parameter. The augmented Lagrangian method combines a penalty method and a Lagrangian method, where the penalty term is used to implement the initial relaxation of the constraint, and the Lagrangian term automatically tunes the sensitivity of the optimization to the constraint violations. This method is called Iteratively-Refined(IR)-WRI, where the prefix IR refers to the iterative refinement (i.e., defect correction) action of the Lagrange multipliers.

In this study, we propose a new microseismic imaging algorithm for event location and velocity model building based on IR-WRI (Aghamiry et al., 2019b). We extend the IR-WRI method to solve for the signatures and locations of the MSEs as additional variables. Beginning from the initial velocity model, without any assumptions about the MSEs, the first data assimilated wavefield is reconstructed for a band of frequencies starting without a source term (namely, only the data drive the wavefield reconstruction). Then by using the extracted multi-frequency data-assimilated wavefields, we estimate a mean source term averaged over frequencies. During this mean source estimation, we use sparsifying denoising to focus the blended source and hence further facilitate the localization of the MSEs. After a few iterations of this two-step process alternating between wavefield reconstruction and mean source estimation, we apply a peak finder algorithm to the final predicted mean source to extract the location of the MSEs. Then, we jointly update the data assimilated wavefields as well as the source signatures of the picked MSEs keeping the velocity model as a fixed parameter. Finally, we update the velocity model by minimizing the wave-equation errors when the wavefields as well as the locations and signatures of MSEs are kept fixed. The proposed algorithm does not require assumptions about the number or type of MSEs and their locations. However, it does require that the velocity model contains the low wavenumber components of the model.

We first review the different steps of the method. Then, we illustrate the method with the Marmousi synthetic example. Starting from an accurate version of the Marmousi model, we show how the method manages to locate a single source without updating the velocity model. Then, we repeat the same test with a highly-smoothed starting velocity model and show how velocity model updating allows for the accurate location of the event. Then, we complicate the latter test with two small clusters of point sources and the results confirm the potential of IR-WRI for MSE localization. As a last example, we test the method on data with added random noise.

2 METHOD

This paper relies on the frequency-domain formulation of FWI. Accordingly, we review the method with a discrete matrix formalism (Pratt et al., 1998).

In passive experiments, the source \( b \) is unknown and can be approximated as a superimposition of \( p \) point sources (resembling a blended source). Accordingly, the source vector \( b(\omega) \) for frequency \( \omega \) reads

\[
b(\omega) = \sum_{j=1}^{p} [s(\omega)]_j \delta(x - x_j) = \Phi s(\omega),
\]

where \( \delta(x) \) is the delta function, \( x_j \) is the point source position, \( [s(\omega)]_j \in \mathbb{C}^{p \times 1} \) denotes the source signature, at angular frequency \( \omega \), associated to the \( j \)th point source (at location \( x_j \)), \( \Phi \in \mathbb{R}^{n \times p} \) is a tall matrix, the columns of which contain shifted delta functions at the positions of the MSEs \( n \) is the number of discretization points in the model). The goal of microseismic imaging is to find the MSE location matrix \( \Phi \), the MSE signature vector \( s_\omega \) and the velocity model (provided that the data set provides a sufficient illumination of the model).

Frequency-domain FWI with an unknown blended source \( b(\omega) \) can be written as

\[
\text{minimize}_{m \in M, u(\omega), b(\omega), \ldots \atop u(\omega), b(\omega)} \quad \mathcal{R}_m(m) + \sum_{\omega = \omega_1}^{\omega_q} \mathcal{R}_b(b(\omega))
\]

subject to

\[
A(m, \omega)u(\omega) = b(\omega), \quad \omega = \omega_1, \ldots, \omega_q,
\]

\[
P u(\omega) = d(\omega), \quad \omega = \omega_1, \ldots, \omega_q
\]

where \( m \) is the squared slowness, \( \mathcal{R}_m \) and \( \mathcal{R}_b \) are appropriate regularization functions, \( A(m, \omega) = \nabla^2 + \omega^2 \text{Diag}(m) \in \mathbb{C}^{n \times n} \) is the Helmholtz operator, \( \nabla^2 \) is the Laplacian operator, \( \text{Diag}(\cdot) \) denotes a diagonal matrix.
with \( \mathbf{u}(\omega) \in \mathbb{C}^{n \times 1} \) and \( \mathbf{d}(\omega) \in \mathbb{C}^{r \times 1} \) denote the wavefield and the recorded data for frequency \( \omega \), respectively. \( \mathbf{P} \in \mathbb{R}^{r \times n} \) is the observation operator that samples \( \mathbf{u}(\omega) \) at receiver locations, and \( r \) is the number of receivers. Finally, \( \mathbf{m} \in \mathcal{M} \) is a bounding constraint on the model parameters where
\[
\mathcal{M} = \{ \mathbf{m} | \mathbf{m}_{\text{min}} \leq \mathbf{m} \leq \mathbf{m}_{\text{max}} \}. \tag{2.3}
\]

The classical implementation of the FWI as formulated in (2.2) would enforce the wave-equation constraint, \( \mathbf{u}(\omega) = \mathbf{A}(\mathbf{m}, \omega)^{-1}\mathbf{b}(\omega) \), in the observation-equation constraint and process the latter as a penalty term leading to the following optimization problem:
\[
\begin{aligned}
\min_{\mathbf{m} \in \mathcal{M}, \mathbf{b}(\omega_1), \ldots, \mathbf{b}(\omega_q)} & \quad R_m(\mathbf{m}) + \sum_{\omega = \omega_1}^{\omega_q} R_{b}(\mathbf{b}(\omega)) + \frac{\lambda}{2} \sum_{\omega = \omega_1}^{\omega_q} \| \mathbf{G}(\mathbf{m}, \omega)\mathbf{b}(\omega) - \mathbf{d}(\omega) \|_2^2, \tag{2.4}
\end{aligned}
\]
where \( \lambda \) is the penalty parameter and \( \mathbf{G}(\mathbf{m}, \omega) = \mathbf{P}\mathbf{A}(\mathbf{m}, \omega)^{-1} \). The problem in equation (2.4) is severely underdetermined due to the unknown source and highly nonlinear due to the oscillatory nature of the Green’s functions, which makes the waveform inversion prone to cycle skipping. Michel & Tsvankin (2014) solved this problem when they assume a good initial estimate of the number of the point sources, \( p \), and their approximate locations, \( \Phi \). Kaderli et al. (2015); Sharan et al. (2019); Shekar & Sethi (2019) used the sparsity promoting \( \ell_1 \)-norm regularization, \( R_{b}(\mathbf{b}(\omega)) = \| \mathbf{b}(\omega) \|_1 \), to enforce the sparsity of the source term, i.e., predicting the data with a minimum number of point sources. They solved the problem using the time domain formulation of FWI.

In this paper, we extend the iteratively-refined wavefield reconstruction inversion (IR-WRI) (Aghamiry et al., 2019b) to solve problem (2.2). IR-WRI relies on the augmented Lagrangian method, which combines the penalty method with the Lagrangian method (Nocedal & Wright, 2006, Chapter 17). The augmented Lagrangian function associated with problem in Eq. (2.2) is given by
\[
\mathcal{L}(\mathbf{m}, \{ \mathbf{u}(\omega) \}, \{ \mathbf{b}(\omega) \}, \{ \mathbf{v}(\omega) \}, \{ \mathbf{w}(\omega) \}) = R_m(\mathbf{m}) + \sum_{\omega = \omega_1}^{\omega_q} R_{b}(\mathbf{b}(\omega))
\]
\[
+ \sum_{\omega = \omega_1}^{\omega_q} \langle \mathbf{v}(\omega), \mathbf{A}(\mathbf{m}, \omega)\mathbf{u}(\omega) - \mathbf{b}(\omega) \rangle + \sum_{\omega = \omega_1}^{\omega_q} \langle \mathbf{w}(\omega), \mathbf{P}(\mathbf{u}(\omega) - \mathbf{d}(\omega) \rangle
\]
\[
+ \frac{\lambda}{2} \sum_{\omega = \omega_1}^{\omega_q} \| \mathbf{A}(\mathbf{m}, \omega)\mathbf{u}(\omega) - \mathbf{b}(\omega) \|_2^2 + \frac{\mu}{2} \sum_{\omega = \omega_1}^{\omega_q} \| \mathbf{P}(\mathbf{u}(\omega) - \mathbf{d}(\omega) \|_2^2,
\]
where the scalars \( \lambda, \mu > 0 \) are the penalty parameters assigned to the wave equation and the observation equation constraints, respectively, and \( \mathbf{v}(\omega) \) and \( \mathbf{w}(\omega) \) are the Lagrange multipliers. Beginning with an initial model \( \mathbf{m}^0 \), \( \mathbf{b}^0(\omega) = \mathbf{0} \forall \omega \), we compute an initial set of monochromatic wavefields \( \mathbf{u}^0(\omega) \) by solving the following overdetermined systems in a least-squares sense:
\[
\begin{pmatrix} \sqrt{\lambda} \mathbf{A}(\mathbf{m}^0, \omega) \\ \sqrt{\mu} \mathbf{P} \end{pmatrix} \mathbf{u}^0(\omega) = \begin{pmatrix} \mathbf{0} \\ \sqrt{\mu} \mathbf{d}(\omega) \end{pmatrix}, \quad \omega = \omega_1, \ldots, \omega_q. \tag{2.6}
\]
Then, beginning with \( \mathbf{v}^0(\omega) = \mathbf{0} \) and \( \mathbf{w}^0(\omega) = \mathbf{0}, \forall \omega \), we solve the multivariate optimization problem, equation (2.5), iteratively by using the ADMM as
\[
\begin{aligned}
\left\{ \{ \mathbf{u}(\omega)^{k+1}, \{ \mathbf{b}(\omega)^{k+1} \} \right\} & = \arg \min_{\{ \mathbf{u}(\omega), \{ \mathbf{b}(\omega) \}} \mathcal{L}(\mathbf{m}^k, \{ \mathbf{u}(\omega) \}, \{ \mathbf{b}(\omega) \}, \{ \mathbf{v}(\omega)^k \}, \{ \mathbf{w}(\omega)^k \}) \tag{2.7a}\\
\mathbf{m}^{k+1} & = \arg \min_{\mathbf{m} \in \mathcal{M}} \mathcal{L}(\mathbf{m}, \{ \mathbf{u}(\omega)^{k+1} \}, \{ \mathbf{b}(\omega)^{k+1} \}, \{ \mathbf{v}(\omega)^k \}, \{ \mathbf{w}(\omega)^k \}) \tag{2.7b}\\
\mathbf{v}(\omega)^{k+1} & = \mathbf{v}(\omega)^k + \lambda (\mathbf{A}(\mathbf{m}^{k+1}, \omega)\mathbf{u}(\omega)^{k+1} - \mathbf{b}(\omega)^{k+1}), \quad \omega = \omega_1, \ldots, \omega_q \tag{2.7c}\\
\mathbf{w}(\omega)^{k+1} & = \mathbf{w}(\omega)^k + \mu (\mathbf{P}(\mathbf{u}(\omega)^{k+1} - \mathbf{d}(\omega))), \quad \omega = \omega_1, \ldots, \omega_q \tag{2.7d}
\end{aligned}
\]
where \( k \) is the (outer) iteration number. The penalty parameters \( \lambda, \mu > 0 \) are tuned such that a dominant weight \( \mu \) is given to the observation equation at the expense of violating the wave equation during the early iterations to guarantee a data fit that prevents cycle skipping at receivers even at early iterations. The iterative update of the Lagrange multipliers progressively corrects the errors introduced by these penalizations such that both the observation equation and the wave equation are satisfied at the convergence point with acceptable accuracies. In the next two subsections, we review how to solve each of optimization subproblems (2.7a)-(2.7b).
2.1 Estimation of the sources and wavefields

Due to the ill-conditioning of the problem of estimating \( \{ u(\omega) \} \) and \( \{ b(\omega) \} \), we solve the subproblems (2.7a) in a two step manner. Since the event location matrix \( \Phi \) in Eq. (2.1) is frequency independent, we first solve the optimization problem for a mean source \( \bar{b}^k = \frac{1}{q} \sum_{\omega=\omega_1} \omega q b(\omega) \) averaged over the frequency to reduce the search space. The optimization problem over \( \bar{b} \) reads

\[
\bar{b}^{k+1} = \arg \min_b \mathcal{R}_b(\bar{b}) + \frac{\lambda}{2} \sum_{\omega=\omega_1} \omega q \| A(m^k,\omega)u(\omega) - \bar{b} \|_2^2 + \nabla \| v(\omega) \|_2^2.
\]  

(2.8)

By adding and subtracting the term \( \| v(\omega) \|_2^2 \) to problem (2.8), we have (Aghamiry et al., 2019a, Appendix A)

\[
\bar{b}^{k+1} = \arg \min_b \mathcal{R}_b(\bar{b}) + \frac{\lambda}{2} \sum_{\omega=\omega_1} \omega q \| A(m^k,\omega)u(\omega) - \bar{b} \|_2^2 + \frac{1}{\lambda} \| v(\omega) \|_2^2.
\]  

(2.9)

where we have ignored the \( \| v(\omega) \|_2^2 \)-term as it does not impact the optimization result. Equation 2.9 is a denoising/proximity problem (Parikh & Boyd, 2013) applied to \( b \), i.e.

\[
\bar{b}^{k+1} = \arg \min_b \mathcal{R}_b(\bar{b}) + \frac{\lambda}{2} \sum_{\omega=\omega_1} \omega q \| A(m^k,\omega)u(\omega) - \bar{b} \|_2^2 + \frac{1}{\lambda} \| v(\omega) \|_2^2.
\]  

(2.10)

We use the so called Berhu regularizer (Owen, 2007), which is a hybrid function combining the \( \ell_1 \) norm for small values and the \( \ell_2 \) norm for large values:

\[
\mathcal{R}_b(x) = \begin{cases} |x| & |x| \leq \varepsilon, \\ \frac{\varepsilon^2}{x^2 + \varepsilon^2} |x| & |x| > \varepsilon, \end{cases}
\]  

(2.11)

where \( \varepsilon > 0 \) determines where the transition from \( \ell_1 \) to \( \ell_2 \) occurs. This regularizer shrinks small coefficients with the \( \ell_1 \) norm to promote sparsity while damping the large coefficients (Fig. 1). Using the Berhu regularizer, the proximity operator in Eq. (2.10) admits an explicit solution for each entry of the source vector

\[
\bar{b}^{k+1} = \text{prox}_{\frac{\lambda}{2}\mathcal{R}_b} \left( \frac{1}{q} \sum_{\omega=\omega_1} \omega q A(m^k,\omega)u(\omega) + \frac{1}{\lambda} v(\omega) \right),
\]  

(2.12)

where \( \text{prox}_{\mathcal{R}_b} \) is the proximity operator of the Berhu function defined as

\[
\text{prox}_{\mathcal{R}_b}(x) = \begin{cases} \max(1 - \frac{\alpha}{|x|}, 0)x & |x| \leq \alpha + \varepsilon, \\ \frac{|x|}{\alpha + \varepsilon} x & |x| > \alpha + \varepsilon. \end{cases}
\]  

(2.13)

In order to improve the accuracy of the locations, we repeat the process of estimating the locations and updating the wavefields several times. The monochromatic wavefields are updated as follows

\[
\left( \frac{\sqrt{\lambda}}{\sqrt{\mu} P} m^k, \omega \right) u(\omega) = \left( \frac{\sqrt{\lambda}}{\sqrt{\mu} d(\omega)} \right)^{k+1}, \quad \omega = \omega_1, \ldots, \omega_q.
\]  

(2.14)

and we alternate between solving Eq. (2.14) for \( u(\omega) \) given \( \bar{b}^{k+1} \) and updating \( \bar{b}^{k+1} \) via Eq. (2.12) using \( u(\omega) \) obtained from Eq. (2.14).

The final \( \bar{b}^{k+1} \) is then used to construct the location matrix \( \Phi^{k+1} \). To do so, we apply a peak finder algorithm on \( |\bar{b}^{k+1}| \) to determine the location of the peaks. The number of peaks found (i.e., the number of MSE) determines the number of columns of \( \Phi^{k+1} \) and the location of each peak determines the location of the delta function in the corresponding column. Once \( \Phi^{k+1} \) has been determined, the monochromatic wavefields and the time function of each MSE are updated simultaneously:

\[
\left( \frac{\sqrt{\lambda}}{\sqrt{\mu} P} - \frac{\sqrt{\lambda}}{\sqrt{\mu} P} \Phi^{k+1} \right) u(\omega)^{k+1} = \left( \frac{1}{\sqrt{\mu} d(\omega)} \right), \quad s(\omega)^{k+1} = \left( \frac{1}{\sqrt{\mu} w(\omega)} \right).
\]  

(2.15)
2.2 Velocity model update

If we have adequate illumination, we can update the velocity model to improve the source locations; otherwise, this step can be skipped. Due to the blended nature of microseismic data, the velocity update will contain noise from crosstalk. This high-frequency noise can be decreased by applying appropriate regularization. In this paper, we use the first-order isotropic TV regularization (Rudin et al., 1992) for $m$. However, other regularizations such as compound or adaptive regularizations can be used in a similar way (see Aghamiry et al., 2020, 2021a). The isotropic TV regularizer is defined as

$$ R_m(m) = \sum \sqrt{(\nabla_x m)^2 + (\nabla_z m)^2}, $$

(2.16)

where $\nabla_x$ and $\nabla_z$ are respectively first-order difference operators in the horizontal and vertical directions with appropriate boundary conditions. The problem 2.7b with non-smooth TV regularization and bound constraints can be written as

$$ m^{k+1} = \arg \min_{m \in M} R_m(m) + \lambda (m^T H_k m - g_k^T m), $$

(2.17)

where

$$ H_k = \sum_{\omega=\omega_1}^{\omega_q} \left( \frac{\partial A(m)}{\partial m} u(\omega)^{k+1} \right)^T \left( \frac{\partial A(m)}{\partial m} u(\omega)^{k+1} \right), $$

(2.18)

$$ g_k = \sum_{\omega=\omega_1}^{\omega_q} \left( \frac{\partial A(m)}{\partial m} u(\omega)^{k+1} \right)^T (\Phi^{k+1} s(\omega)^{k+1} - \frac{1}{\lambda} v(\omega)^k - \nabla^2 u(\omega)^{k+1}). $$

(2.19)

The model subproblem in Eq. (2.17) requires a quadratic optimization with bound constrained TV regularization. There are many well documented algorithm to carry out this step (Aghamiry et al., 2019a, 2020; Gholami & Naeini, 2019; Goldstein & Osher, 2009; Maharramov & Levin, 2015). The proposed microseismic imaging algorithm is summarized in Algorithm 1. It should be noted that the total algorithm consists of three main steps: (1) estimation of the source location (lines 5-9), (2) estimation of the source signatures and the wavefields (line 10) and (3) update of the model parameters (line 11). Step 3 can be skipped if a new sets of MSEs is processed with a sufficiently-accurate background subsurface model estimated during a prior inversion. The linear systems for data-assimilated wavefield reconstruction, equations 2.6, 2.14, and 2.15, have one right-hand side. Therefore, they may be solved more efficiently with preconditioned iterative solvers instead of direct solvers to tackle large computational domain. However, if multiple datasets are processed over time with the same background subsurface model, it may be beneficial to use direct solvers to compute the Lower-Upper (LU) factors once and for all once a good background model has been estimated, store them on disk and re-use them to process a new dataset efficiently by forward/backward elimination.

3 NUMERICAL EXAMPLES

We assess the method on the selected target of the synthetic Marmousi II model (Martin et al., 2006) with size $2.5 \text{ km} \times 1 \text{ km}$ when the grid spacing is $5 \text{ m}$ (Fig. 2a). We use two different initial models for the tests: an accurate initial velocity model obtained by slightly smoothing the true model, referred to as initial model 1 (Fig. 
**Algorithm 1 ADMM-based FWI for microseismic imaging.**

1. Begin with $k = 0$ and an initial model $m^0$.
2. Set to zero the values of $v(\omega)^0$ and $w(\omega)^0, \forall \omega$.
3. Calculate $u(\omega)^0, \forall \omega$ (Eq. (2.6)).
4. while convergence criteria not satisfied do
   5. for $l = 0 : n_l$ do
      6. Compute $\vec{b}^{k+1}$ (Eq. (2.12))
      7. Update $u(\omega)^k, \omega = \omega_1, \cdots, \omega_q$ (Eq. (2.14))
   8. end for
   9. Build $\Phi^{k+1}$ from $\vec{b}^{k+1}$ using a peak finder algorithm.
   10. Update $s(\omega)^{k+1}$ and $u(\omega)^{k+1}$ for $\omega = \omega_1, \cdots, \omega_q$ (Eq. (2.15)).
   11. Update $m^{k+1}$ (Eq. (2.17))
   12. Update the dual vectors $v$ and $w$ (Eqs. (2.7c) and (2.7d))
   13. $k = k + 1$
5. end while

---

Figure 2: True and initial velocity models and the locations of MSEs that are used for the numerical tests. (a) The Marmousi II velocity model, used to generate the data. The location of different MSEs, referred as MSEs 1 to 4, are shown by black asterisks. (b) Accurate initial model, called initial model 1. (c) Kinematically accurate velocity model, called initial model 2. (d) A direct comparison between true (black), initial model 1 (red), and initial model 2 (blue) at $X = 1250$ m. 2b), and a highly-smoothed, albeit kinematically accurate, initial velocity model referred to as initial model 2 (Fig. 2c). A direct comparison between the true model and initial models 1 and 2 at $X = 1250$ m is shown in Fig. 2d. The different MSEs, which are used in this section, are shown in Fig. 2a using black asterisks and numbers. The MSEs of the right cluster are located at (520,1360) m [MSE 1] and (560,1380) m [MSE 2], and those of the left cluster are located at (605,1150) m [MSE 3], and (645,1175) m [MSE 4]. The source signatures for MSEs 1 to 4 are Ricker wavelets with central frequencies [25, 31, 23, 29] Hz and central times [2.4, 2.56, 2.25, 2.2] s, respectively.

For all of the numerical tests, we compute wavefields with a nine-point stencil finite-difference method implemented with anti-lumped mass, where the stencil coefficients are optimized for each frequency (Chen et al., 2013).
use absorbing boundary conditions along the bottom, right and left sides of the model and a free-surface boundary condition at the surface.

We first illustrate how the initial data-assimilated wavefield is reconstructed (when no information on the source is available, eq. 2.6). The source contains MSEs 1 to 4 and we use model 2 as the starting velocity model (Fig. 2c). Fig. 3 shows the true frequency-domain wavefields and the initial data-assimilated wavefields for the 7 Hz and 15 Hz frequencies which are reconstructed by back-propagating the data, eq. 2.6. The ability of the data-assimilated wavefield to mimic the true wavefield can be assessed by comparing the left and right columns of this figure. This emphasizes how suitable the (transmission) microseismic configuration is to perform accurate wavefield reconstruction between the MSEs and the receivers using data assimilation in the initial model. The time-domain seismograms computed from the MSEs in the true model are shown in Fig. 4a, while those computed in the initial model with data assimilation, when no information about the source is available, are shown in Fig. 4b. These seismograms match almost perfectly because the data assimilated wavefield fits the data at the expense of the accuracy with which they
satisfy the wave equation through the feedback term to the data (Eq. 2.6). Also, a comparison between the time slices at 2.7 s of the true (Fig. 5a) and data-assimilated wavefields (Fig. 5b) shows that the data-assimilated wavefield captures many features of the true wavefield above the MSEs which again emphasizes the ability of data-assimilation to reconstruct an accurate wavefield.

We continue by assessing the performance of Algorithm 1 in estimating the MSE location and signature estimation. We start by a simple setup and complicate it step by step. We perform the first test with a source that contains only MSE 1 using model 1 as the starting model (Fig. 2b). The source signature of this MSE is shown in Fig. 6c in purple. Note that all the wavelets in this section are created with the inverse Fourier transform of a few discrete frequencies, those that are used for inversion in IR-WRI.

We invert frequency components between 5 Hz and 45 Hz with a 2-Hz frequency interval simultaneously. We perform the inner loop for source location estimation (lines 5-8 of Algorithm 1) with 10 inner iterations ($n_l = 9$). The estimated $\bar{b}_1$ at the final iteration of the inner loop is shown in Fig. 6a. The event location found by the peak finder algorithm is shown in Fig. 6b in pink while the true one is shown in purple. The results are zoomed in from the full model; the limits of the zoom are shown with the black dashed lines in Fig. 6a. We observe that the estimated MSE location is close to the true one because of the accuracy of the initial velocity model. Finally, the estimated source signature is shown in Fig. 6c (pink curve), which matches the true one well.

We make this test more complicated by starting with initial velocity model 2. The estimated predicted source at the final inner iteration of outer iteration 1 ($\bar{b}_1$) is shown in Fig. 7a and the result of the application of the peak finder algorithm is shown in Fig. 7b in pink. We see that the error in the estimated location is more than the previous case.
(Fig. 6b) where the initial model was accurate. Also, the estimated source signature is shown in pink in Fig. 7e. We can clearly see that neither the estimated MSE location nor the estimated signature are accurate. To improve the quality of these results, we should update the velocity model. The estimated predicted source after five outer iterations (at the final iteration of inner loop), $\bar{b}^5$, and the result of peak finder algorithm are shown in Figs. 7c-7d, respectively. Finally, the estimated source signature at outer iteration 5 is shown in green in Fig. 7e. It can be seen that the quality of the estimated location and signature improved significantly. The updated velocity model at outer iteration five is shown in Fig. 7f. Although the velocity update is not useful for geological interpretation due to the limited illumination provided by the single event, it improves the structures between the event and the receivers and, as a result, leads to an improved MSE location and signature estimation.

Figure 7: Marmousi test with initial model 2 when the source just contains MSE 1. (a) Predicted source at the final inner iteration of outer iteration 1 ($\bar{b}^1$). (b) The zoomed result of peak finder applied on $\bar{b}^1$. The margins of the zoomed part are shown using black dashed lines in (a). Also, the location of the true source is indicated by a purple point. (c-d) Same as (a-b) but for $\bar{b}^5$. (e) The true source signature (purple) and the estimated at outer iteration 1 (pink) and iteration 5 (green). (f) Updated velocity model after five iterations.

To make the test more representative of the microseismic scenario, we repeat the Marmousi test when the source gathers MSEs 1-4 (Fig. 2a). We start with the initial model 1. The predicted source after 1 outer iteration ($\bar{b}^1$) and the result of peak finder are shown in Figs. 8a-8b, respectively. We see that, because of the good initial model, there are just 4 peaks of energy in Fig. 8. Also, the estimated source signatures for MSEs 1 to 4 are shown in Figs. 8c-8f, respectively, in pink. We have accurately recovered both the locations and signatures for all MSEs.

Even with a good initial model, the estimated predicted source has a lot of unwanted energy, and sparsifying regularization is necessary to find the locations of the MSEs from the predicted source. The predicted source for this test at inner iteration 1 ($k = 0$, $l = 0$) before and after applying Berhu regularization are depicted in Figs. 9a-9b, respectively. We can see the improvement, although it is not enough to have an accurate picking for the location of the MSEs. The same results after 10 inner iterations ($k = 0$, $l = 9$) are shown in Figs. 9c-9d. First, we see that the regularization has little effect on the predicted source in Fig. 9d (compare Fig. 9c) because it is already sparse.
Figure 8: The Marmousi test with initial model 1 when the source contains MSEs 1-4. (a-b) $\hat{b}^1$ and its selected peaks using the peak finder. (c-f) True source signatures (purple) and the estimated ones (pink) for MSEs 1 to 4, respectively.

Second, we see the improvement in Fig. 9d compared to Fig. 9b, where the regularization has significantly improved the MSE location selection.

Figure 9: Showing the impact of sparsifying regularization on the estimated predicted sources for the Marmousi test with initial model 1 when the source contains MSEs 1-4. (a-b) The predicted source at $k = 0, l = 0$ (a) before, (b) after Berhu regularization. (c-d) Same as (a-b) but for $k = 0, l = 9$. $k$ and $l$ refer to the outer and inner iteration of algorithm 1, respectively.

We repeat the Marmousi test with initial velocity model 2 when the source contains MSEs 1-4, and we use five outer iterations to update the velocity model to improve the quality of the estimated MSEs. The estimated $\hat{b}^1$ (at the final inner iteration) and its located peaks are shown in Figs. 10a-10b and the same results for outer iteration five are shown.
in Figs. 10c-10d. At outer iteration 1, the peak finder algorithm finds more peaks (five) than the true number of MSEs because of the inaccurate velocity model (Fig. 10b). Then, the data-assimilated wavefield and source signatures for all these selected MSEs are updated jointly using Eq. 2.15. Because the selected MSE at the middle of Fig. 10b does not contribute to the recorded data, the algorithm finds a small source signature for this fake MSE. As soon as the velocity model improves, we have a better predicted source (Fig. 10c), and the peak finder algorithm selects four points close to the true MSEs (Fig. 10d). Also, the estimated MSE signatures at outer iterations 1 and 5 are shown in Figs. 10e-10h, respectively, for MSEs 1 to 4 in pink for outer iteration one and in green for outer iteration five. Finally, the updated velocity model after five outer iterations is shown in Fig. 10i. First, the quality of the estimated signatures of the MSEs are improved by updating the velocity model. Second, the updated velocity model captures the trends of dominant structures, and is improved compared to the initial model 2.

Like the previous test, the estimated source locations at inner iteration 1 ($k = 0$, $l = 0$) before and after applying Berhu regularization are depicted in Figs. 11a-11b, respectively, and for inner iteration 10 ($k = 0$, $l = 9$) shown in Figs. 11c-11d. We see that the energy is less focused compared to the case with an accurate initial model (Figs. 9a-9b). We also see that the regularization significantly improves the predicted source map, making it ready for MSEs
location picking.

![Image](image.png)

Figure 11: Same as Fig. 9, but for Marmousi test with initial model 2 when the source contains MSEs 1-4.

We assess now the robustness of the method against noise. We repeat the previous test (with initial model 2 and the source that contains MSEs 1-4) when Gaussian distributed random noise with SNR=5 db is added to the data. We use the same configuration as for the noiseless test (Fig. 10). The results are shown in Fig. 12. In comparison with the noiseless case, we see that the predicted source has more peaks than the true number of MSEs at the first outer iteration (Fig. 12a), but it is improved at the fifth outer iteration (Fig. 12c). Also, we see that the estimated MSEs are not significantly changed in Fig. 12c compared to the noiseless case (Fig. 10c). For the estimated signatures, Figs. 12e-12h, we see the quality of the estimated signatures are degraded compare to the noiseless case (Figs. 10e-10h). Finally, the updated velocity model for this noisy test, Fig. 12i, is close to the noiseless one in Fig. 10i.

4 Discussion

In this paper, we have proposed an algorithm based on ADMM-based FWI for characterizing weak seismic events that is valid at any scale. We have focused on microseismic imaging to find the location and signature of MSEs, but the method is general and could be applied to tectonic events or any other situation in which the source location is unknown. The proposed algorithm does not require any information or assumptions about the sources, and it is able to find the location and time signature of seismic events and update the background velocity model if required. The proposed algorithm consists of three steps: 1- finding the number and location of seismic events, 2- jointly updating the data-assimilated wavefield and the event signatures, 3- updating the background velocity model provided that the recorded data provides sufficient illumination of the model.

If the location of the seismic events are known, Aghamiry et al. (2021b); Fang et al. (2018) have shown that it is possible to jointly estimate the data-assimilated wavefield and the event signatures with high accuracy by solving a linear least-squares problem. But, when the location and the number of seismic events are unknown, the first step of the proposed algorithm tries to find this information by applying a peak finder algorithm on the sparsified predicted source map. By using Eq. (2.6) and considering the estimated source as \( \mathbf{A}(\mathbf{m}^{0}, \omega)\mathbf{u}(\omega)^{0} \), the predicted source locations are generated by first propagating the data backward in time from the receiver positions and then the blurring effects induced by the limited bandwidth of the data and the limited spread of the receivers are corrected with the sparsity-promoting regularization. The average of these estimated sources helps to find the number and the location of seismic events. However, the success of this step strongly depends on the sparsifying regularization that is applied on the estimated sources.

At the second step of the algorithm, the seismic event signatures and data-assimilated wavefield are jointly updated from the estimated locations. When the algorithm finds incorrect seismic events in the first step, one contribution of this second step is to mitigate their footprint by assigning low-amplitude signatures to them. This wavefield and
signature refinement, as well as the updating of the background model, help the algorithm to better-estimate the source and seismic event locations in subsequent outer iterations.

The numerical results show that we should have at least a kinematically accurate initial model to have a good estimate of the seismic events. But, if the model is kinematically incorrect, we should have sufficient illumination of the model in the recorded data to update it during the outer iteration. We note that we cannot update the velocity model everywhere, but the region in which we can update the model is the region through which the recorded waves travel. Thus we can update the part of the model that is most important for the estimation of the source.

The proposed algorithm is in the frequency domain when the inversion is limited to a few frequencies. The computational burden of the proposed algorithm is primarily in solving the data-assimilated subproblem with a direct solver, Eqs. 2.14, 2.15. For problems with large computational domains, these subproblems can be solved efficiently with preconditioned iterative solvers. Rezaei et al. (2021) show that conjugate-gradient with additive Schwarz domain-decomposition preconditioner (Dryja & Widlund, 1987) has the best performance for solving data-assimilated subproblem in comparing it to other solvers. The computational burden of all the other steps of the algorithm are negligible.
5 CONCLUSIONS

We proposed a method based on ADMM-based FWI for characterizing weak seismic events at different scales. When the source is added to the unknowns of FWI, in addition to the wavefields and model parameters, it becomes a severely underdetermined problem, and it is challenging to uniquely determine the source without prior information. We present a method to solve this problem that does not require prior information about the sources, although it does require a kinematically correct velocity model. The proposed method consists of three steps: With a finite band of frequencies and appropriate sparsifying regularizations on the source, seismic events are selected with a peak finder algorithm applied on the estimated average (over frequency) source. The time signatures and wavefields are jointly updated for the selected seismic events by solving a linear least-squares problem. The velocity model is updated by applying appropriate regularization on the model. We validate the proposed method with synthetic tests for microseismic event characterizations as a proof of concept. The method works well, even for closely spaced events and has acceptable performance with extremely low signal-to-noise ratio data. The method can be tailored to earthquake relocation or microseismic monitoring. Ongoing work involves extending the method to 3D elastic physics and moment tensor estimation in the prospect of real data applications.

ACKNOWLEDGMENTS

This study was partially funded by the WIND consortium (https://www.geoazur.fr/WIND), sponsored by Chevron, Shell and Total. The authors are grateful to the OPAL infrastructure from Observatoire de la Côte d’Azur (CRIMSON) for providing resources and support. This work was granted access to the HPC resources of IDRIS under the allocation A0050410596 made by GENCI. A. Malcolm acknowledges the NSERC Discovery Grant Program as well as Chevron, InnovateNL and the NSERC Industrial Research Chair program.

References

Abubakar, A., Hu, W., Habashy, T. M., & van den Berg, P. M., 2009. Application of the finite-difference contrast-source inversion algorithm to seismic full-waveform data, Geophysics, 74(6), WCC47–WCC58.
Aghamiry, H., Gholami, A., & Operto, S., 2019a. Implementing bound constraints and total-variation regularization in extended full waveform inversion with the alternating direction method of multiplier: application to large contrast media, Geophysical Journal International, 218(2), 855–872.
Aghamiry, H., Gholami, A., & Operto, S., 2019b. Improving full-waveform inversion by wavefield reconstruction with alternating direction method of multipliers, Geophysics, 84(1), R139–R162.
Aghamiry, H., Gholami, A., & Operto, S., 2020. Compound regularization of full-waveform inversion for imaging piecewise media, IEEE Transactions on Geoscience and Remote Sensing, 58(2), 1192–1204.
Aghamiry, H., Gholami, A., & Operto, S., 2021a. Full Waveform Inversion by Proximal Newton Methods using Adaptive Regularization, Geophysical Journal International, 224(1), 169–180.
Aghamiry, H., Mamfoumbi-Ozoumet, F. W., Gholami, A., & Operto, S., 2021b. Efficient extended-search space full-waveform inversion with unknown source signatures, arXiv preprint arXiv:2102.03916.
Aravkin, A. Y., van Leeuwen, T., Calandra, H., & Herrmann, F. J., 2012. Source estimation for frequency-domain FWI with robust penalties, in Expanded Abstracts, 74th Annual EAGE Meeting.
Chen, Z., Cheng, D., Feng, W., & Wu, T., 2013. An optimal 9-point finite difference scheme for the Helmholtz equation with PML, International Journal of Numerical Analysis & Modeling, 10(2).
De Ridder, S. & Maddison, J., 2018. Full wavefield inversion of ambient seismic noise, Geophysical Journal International, 215(2), 1215–1230.
Dryja, M. & Widlund, O., 1987. An additive variant of the Schwarz alternating method for the case of many subregions, Ultracomputer Research Laboratory, Univ., Courant Inst. of MathematicalSciences, Division of Computer Science.
Fang, Z., Wang, R., & Herrmann, F. J., 2018. Source estimation for wavefield-reconstruction inversion, Geophysics, 83(4), R345–R359.
Gajewski, D. & Tessmer, E., 2005. Reverse modelling for seismic event characterization, Geophysical Journal International, 163(1), 276–284.
Gholami, A. & Naeini, E. Z., 2019. 3D Dix inversion using bound-constrained TV regularization, Geophysics, 84(3), 1–43.
Goldstein, T. & Osher, S., 2009. The split Bregman method for L1-regularized problems, *SIAM Journal on Imaging Sciences*, 2(2), 323–343.

Haber, E., Ascher, U. M., & Oldenburg, D., 2000. On optimization techniques for solving nonlinear inverse problems, *Inverse problems*, 16(5), 1263.

Han, L., Wong, J., & Bancroft, J. C., 2009. Hypocenter location using hodogram analysis of noisy 3c microseismograms, *Consortium for Research in Elastic Wave Exploration Seismology (CREWES)* Research Report 29.

Kaderli, J., McChesney, M. D., & Minkoff, S. E., 2015. Microseismic event estimation in noisy data via full waveform inversion, in *SEG Technical Program Expanded Abstracts 2015*, pp. 1159–1164, Society of Exploration Geophysicists.

Kaderli, J., McChesney, M. D., & Minkoff, S. E., 2018. A self-adjoint velocity-stress full-waveform inversion approach to microseismic source estimation, *Geophysics*, 83(5), R413–R427.

Lee, K. H. & Kim, H. J., 2003. Source-independent full-waveform inversion of seismic data, *Geophysics*, 68, 2010–2015.

Li, L., Tan, J., Schwarz, B., Stanek, F., Poiata, N., Shi, P., Diekmann, L., Eisner, L., & Gajewski, D., 2020. Recent advances and challenges of waveform-based seismic location methods at multiple scales, *Reviews of Geophysics*, 58(1), e2019RG000667.

Maharramov, M. & Levin, S. A., 2015. Total-variation minimization with bound constraints, *arXiv preprint arXiv:1505.05694*.

Martin, G. S., Wiley, R., & Marfurt, K. J., 2006. Marmousi2: An elastic upgrade for Marmousi, *The Leading Edge*, 25(2), 156–166.

McMechan, G. A., 1982. Determination of source parameters by wavefield extrapolation, *Geophys. J. R. astr. Soc.*, 71, 613–628.

Michel, O. J. & Tsvankin, I., 2014. Gradient calculation for waveform inversion of microseismic data in VTI media, *Journal of Seismic Exploration*, 23(3), 201–217.

Montgomery, C. T., Smith, M. B., et al., 2010. Hydraulic fracturing: history of an enduring technology, *Journal of Petroleum Technology*, 62(12), 26–40.

Nakata, N. & Beroza, G. C., 2016. Reverse time migration for microseismic sources using the geometric mean as an imaging condition, *Geophysics*, 81(2), KS1–KS12.

Nocedal, J. & Wright, S. J., 2006. *Numerical Optimization*, Springer, 2nd edn.

Owen, A. B., 2007. A robust hybrid of lasso and ridge regression, *Contemporary Mathematics*, 443(7), 59–72.

Parikh, N. & Boyd, S., 2013. Proximal algorithms, *Foundations and Trends in Optimization*, 1(3), 123–231.

Plessix, R. E. & Cao, Q., 2011. A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium, *Geophysical Journal International*, 185, 539–556.

Pratt, R. G., 1999. Seismic waveform inversion in the frequency domain, part I: theory and verification in a physical scale model, *Geophysics*, 64, 888–901.

Pratt, R. G., Shin, C., & Hicks, G. J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophysical Journal International*, 133, 341–362.

Rezaei, A., Aghamiry, H., Gholami, A., & Operto, S., 2021. Iterative reconstruction of data assimilated wavefields in the extended-source full-waveform inversion, in *submitted to 82th EAGE Conference and Exhibition 2021*, vol. 2021, European Association of Geoscientists & Engineers.

Rickett, J., 2013. The variable projection method for waveform inversion with an unknown source function, *Geophysical Prospecting*, 61(4), 874–881.

Rudin, L., Osher, S., & Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms, *Physica D*, 60, 259–268.

Sharan, S., Wang, R., & Herrmann, F. J., 2019. Fast sparsity-promoting microseismic source estimation, *Geophysical Journal International*, 216, 164–181.

Shekar, B. & Sethi, H. S., 2019. Full-waveform inversion for microseismic events using sparsity constraints, *Geophysics*, 84(2), KS1–KS12.

Sjögreen, B. & Petersson, N. A., 2014. Source estimation by full wave form inversion, *Journal of Scientific Computing*, 59(1), 247–276.
Song, C., Alkhalifah, T., & Wu, Z., 2019. Microseismic event estimation and velocity analysis based on a source-focusing function, *Geophysics*, **84**(3), KS85–KS94.

Sun, J., Xue, Z., Fomel, S., Zhu, T., & Nakata, N., 2016. Full waveform inversion of passive seismic data for sources and velocities, in *86th Annual SEG Meeting*.

Thurber, C. & Engdahl, E., 2000. Advances in seismic event location: Springer Netherlands.

van Leeuwen, T. & Herrmann, F., 2016. A penalty method for PDE-constrained optimization in inverse problems, *Inverse Problems*, **32**(1), 1–26.

van Leeuwen, T. & Herrmann, F. J., 2013. Mitigating local minima in full-waveform inversion by expanding the search space, *Geophysical Journal International*, **195**(1), 661–667.

Virieux, J. & Operto, S., 2009. An overview of full waveform inversion in exploration geophysics, *Geophysics*, **74**(6), WCC1–WCC26.