# ADMM-based full-waveform inversion for microseismic imaging

Stéphane Operto, H Aghamiry, A Gholami, A Malcolm

Geophysical Journal International

*Geophys. J. Int.* (2022) **228,** 259–274
Advance Access publication 2021 April 16
GJI Seismology

https://doi.org/10.1093/gji/ggab332

# ADMM-based full-waveform inversion for microseismic imaging

H. S. Aghamiry [1], A. Gholami,[2] S. Operto [1] and A. Malcolm [3]

[1] *Université Côte d'Azur, CNRS, Observatoire de la Côte d'Azur, IRD, Géoazur,* 06560, *Valbonne, France. E-mail: aghamiry@geoazur.unice.fr*
[2] *Institute of Geophysics, University of Tehran, Tehran, Iran*
[3] *Department of Earth Sciences, Memorial University of Newfoundland, St John's, NL A1B 3X5, Canada*

Accepted 2021 August 13. Received 2021 July 31; in original form 2021 April 21

## SUMMARY

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 (WRI) method has shown promising results to mitigate the non-linearity of the FWI objective function 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 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 locations. 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.

**Key words:** Waveform inversion; Computational seismolog; and Earthquake source observations..

## 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 (Pratt 1999; Plessix & Cao 2011; Aravkin *et al*. 2012; Rickett 2013; Aghamiry *et al*. 2021b).

This paper focuses 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 location function, both of which strongly influence the FWI results. 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 (Thurber & Engdahl 2000; Han *et al*. 2009). These methods are time-consuming and may not be robust in the presence of noise (Thurber & Engdahl 2000; Han *et al*. 2009). More recently, wavefield-based imaging techniques have received a lot of attention for the relocation of earthquakes and MSEs (McMechan 1982; Gajewski & Tessmer 2005; Michel & Tsvankin 2014; Sjögreen & Petersson 2014; Kaderli

*et al*. 2015; De Ridder & Maddison 2018; Sharan *et al*. 2019; Shekar & Sethi 2019; 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 (McMechan 1982; Gajewski & Tessmer 2005; 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 (Montgomery *et al*. 2010; Li *et al*. 2020).

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; Shekar & Sethi 2019; Sharan *et al*. 2019). Michel & Tsvankin (2014) reviewed how to compute the FWI gradient for source parameters in vertical transverse isotropic (VTI) media. Kaderli *et al*. (2015), Shekar & Sethi (2019), Sharan *et al*. (2019) proposed robust FWI algorithms with sparsifying regularization to refine the location of MSEs and estimate their temporal signatures. 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 and spatial source location functions, while Sharan *et al*. (2019) update both of them jointly. Following this, the source signatures of the identified events are reconstructed separately. Most of the above approaches rely on the acoustic approximation of wave propagation. The extension to elastic media is reviewed in Kaderli *et al*. (2018) where the moment tensor components are estimated together with the spatial distribution of the events and the temporal source signatures. 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 (Sun *et al*. 2016; Song *et al*. 2019). Sun *et al*. (2016) update the source and the velocity model in an alternating manner. The descent direction of the underdetermined 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.

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 non-linear. Part of this non-linearity 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 2013, 2016). 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 in iterations. The Lagrange multipliers are updated with gradient ascent, which controls 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). At this stage, we restrict ourselves to the acoustic approximation of wave propagation. 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, 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 multifrequency 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, which can also represent a single extended source, 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 $\mathbf{b}$ is unknown and can be approximated as a superimposition of $p$ point sources (resembling a blended source, and also approximating an extended source). Accordingly, the source vector $\mathbf{b}(\omega)$ for frequency $\omega$ reads

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

where $\delta(\mathbf{x})$ is the delta function, $\mathbf{x}_j$ is the point source position, $[\mathbf{s}(\omega)]_j$ denotes the source signature, at angular frequency $\omega$, associated to the $j$th point source (at location $\mathbf{x}_j$), $\mathbf{\Phi} \in \mathbb{R}^{n\times p}$ 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 $\mathbf{\Phi}$, the MSE signature vector $\mathbf{s}(\omega)$ and the velocity model (provided that the data set provides sufficient illumination of the model).

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

$$\begin{aligned} &\underset{\substack{\mathbf{m}\in M,\\ \mathbf{u}(\omega_1),\mathbf{b}(\omega_1),\ldots,\\ \mathbf{u}(\omega_q),\mathbf{b}(\omega_q)}}{\text{minimize}} \quad R_m(\mathbf{m}) + \sum_{\omega=\omega_1}^{\omega_q} R_b(\mathbf{b}(\omega)) \\ &\text{subject to} \quad \mathbf{A}(\mathbf{m},\omega)\mathbf{u}(\omega) = \mathbf{b}(\omega), \quad \omega = \omega_1,\ldots,\omega_q \\ &\qquad\qquad \mathbf{P}\mathbf{u}(\omega) = \mathbf{d}(\omega), \qquad\quad \omega = \omega_1,\ldots,\omega_q, \end{aligned} \tag{2}$$

where $\mathbf{m}$ is the squared slowness, $R_m$ and $R_b$ are appropriate regularization functions, $\mathbf{A}(\mathbf{m},\omega) = \nabla^2 + \omega^2 \text{Diag}(\mathbf{m}) \in \mathbb{C}^{n\times n}$ is the Helmholtz operator, $\nabla^2$ is the Laplacian operator, Diag($\bullet$) denotes a diagonal matrix with $\bullet$ on its main diagonal, $\mathbf{u}(\omega) \in \mathbb{C}^{n\times 1}$ 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, $M$ is a convex set defined using our prior knowledge of model parameters. For example, if we know the lower and upper bounds on $\mathbf{m}$ then

$$M = \{\mathbf{m} | \mathbf{m}_{min} \le \mathbf{m} \le \mathbf{m}_{max}\}. \tag{3}$$

The classical implementation of FWI as formulated in (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:

$$\underset{\mathbf{m}\in M,\mathbf{b}(\omega_1),\ldots,\mathbf{b}(\omega_q)}{\text{minimize}} \; R_m(\mathbf{m}) + \sum_{\omega=\omega_1}^{\omega_q} R_b(\mathbf{b}(\omega)) + \frac{\lambda}{2}\sum_{\omega=\omega_1}^{\omega_q} \|\mathbf{G}(\mathbf{m},\omega)\mathbf{b}(\omega) - \mathbf{d}(\omega)\|_2^2, \tag{4}$$

where $\lambda$ is the penalty parameter and $\mathbf{G}(\mathbf{m},\omega) = \mathbf{P}\mathbf{A}(\mathbf{m},\omega)^{-1}$. The problem in eq. (4) is severely underdetermined due to the unknown source and highly non-linear 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, $\mathbf{\Phi}$. Kaderli *et al*. (2015), Shekar & Sethi (2019) and Sharan *et al*. (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, that is 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 IR-WRI (Aghamiry *et al*. 2019b) to solve problem (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 the problem in eq. (2) is given by

$$L(\mathbf{m}, \{\mathbf{u}(\omega)\}, \{\mathbf{b}(\omega)\}, \{\mathbf{v}(\omega)\}, \{\mathbf{w}(\omega)\}) = R_m(\mathbf{m}) + \sum_{\omega=\omega_1}^{\omega_q} R_b(\mathbf{b}(\omega)) + \sum_{\omega=\omega_1}^{\omega_q} \langle \mathbf{v}(\omega), \mathbf{A}(\mathbf{m}, \omega)\mathbf{u}(\omega) - \mathbf{b}(\omega)\rangle + \sum_{\omega=\omega_1}^{\omega_q} \langle \mathbf{w}(\omega), \mathbf{P}\mathbf{u}(\omega) - \mathbf{d}(\omega)\rangle + \frac{\lambda}{2}\sum_{\omega=\omega_1}^{\omega_q} \|\mathbf{A}(\mathbf{m}, \omega)\mathbf{u}(\omega) - \mathbf{b}(\omega)\|_2^2 + \frac{\mu}{2}\sum_{\omega=\omega_1}^{\omega_q} \|\mathbf{P}\mathbf{u}(\omega) - \mathbf{d}(\omega)\|_2^2, \tag{5}$$

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 $\mathbf{m}^0$, $\mathbf{b}^0(\omega) = 0\ \forall\omega$, we compute an initial set of monochromatic wavefields $\mathbf{u}^0(\omega)$ by solving the following overdetermined systems in a least-squares sense:

$$\begin{pmatrix} \sqrt{\lambda}\mathbf{A}(\mathbf{m}^0, \omega) \\ \sqrt{\mu}\mathbf{P} \end{pmatrix} \mathbf{u}^0(\omega) = \begin{pmatrix} 0 \\ \sqrt{\mu}\mathbf{d}(\omega) \end{pmatrix}, \qquad \omega = \omega_1\ ,\ ....,\ \omega_q. \tag{6}$$

Then, beginning with $\mathbf{v}^0(\omega) = \mathbf{0}$ and $\mathbf{w}^0(\omega) = \mathbf{0}$, $\forall\omega$, we solve the multivariate optimization problem, eq. (5), iteratively by using the alternating direction method of multipliers (ADMM) as

$$(\{\mathbf{u}(\omega)^{k+1}\}, \{\mathbf{b}(\omega)^{k+1}\}) = \underset{\{\mathbf{u}(\omega)\},\{\mathbf{b}(\omega)\}}{\arg\min}\ L(\mathbf{m}^k, \{\mathbf{u}(\omega)\}, \{\mathbf{b}(\omega)\}, \{\mathbf{v}(\omega)^k\}, \{\mathbf{w}(\omega)^k\}) \tag{7a}$$

$$\mathbf{m}^{k+1} = \underset{\mathbf{m}\in M}{\arg\min}\ L(\mathbf{m}, \{\mathbf{u}(\omega)^{k+1}\}, \{\mathbf{b}(\omega)^{k+1}\}, \{\mathbf{v}(\omega)^k\}, \{\mathbf{w}(\omega)^k\}) \tag{7b}$$

$$\mathbf{v}(\omega)^{k+1} = \mathbf{v}(\omega)^k + \lambda(\mathbf{A}(\mathbf{m}^{k+1}, \omega)\mathbf{u}(\omega)^{k+1} - \mathbf{b}(\omega)^{k+1}),\ \omega = \omega_1\ ,\ ....,\ \omega_q \tag{7c}$$

$$\mathbf{w}(\omega)^{k+1} = \mathbf{w}(\omega)^k + \mu(\mathbf{P}\mathbf{u}(\omega)^{k+1} - \mathbf{d}(\omega)),\ \omega = \omega_1\ ,\ ....,\ \omega_q \tag{7d}$$

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 the optimization subproblems in eqs (7a) and (7b).

### 2.1 Estimation of the sources and wavefields

Due to the ill-conditioning of the problem of estimating $\{\mathbf{u}(\omega)^{k+1}\}$ and $\{\mathbf{b}(\omega)^{k+1}\}$, we solve the subproblems (7a) in a two step manner. Since the event location matrix $\mathbf{\Phi}$ in eq. (1) is frequency independent, we first solve the optimization problem for a mean source $\bar{\mathbf{b}}^k = \frac{1}{q}\sum_{\omega=\omega_1}^{\omega_q} \mathbf{b}^k(\omega)$ averaged over the frequency to reduce the search space. The optimization problem over $\bar{\mathbf{b}}$ reads

$$\bar{\mathbf{b}}^{k+1} = \underset{\bar{\mathbf{b}}}{\arg\min}\ R_b(\bar{\mathbf{b}}) + \textstyle\sum_{\omega=\omega_1}^{\omega_q} \langle \mathbf{v}(\omega)^k, \mathbf{A}(\mathbf{m}^k, \omega)\mathbf{u}(\omega)^k - \bar{\mathbf{b}}\rangle + \frac{\lambda}{2}\sum_{\omega=\omega_1}^{\omega_q} \|\mathbf{A}(\mathbf{m}^k, \omega)\mathbf{u}(\omega)^k - \bar{\mathbf{b}}\|_2^2. \tag{8}$$

By adding and subtracting the term $\|\mathbf{v}(\omega)^k\|_2^2$ in eq. (8), we have (Aghamiry *et al*. 2019a, Appendix A)

$$\bar{\mathbf{b}}^{k+1} = \underset{\bar{\mathbf{b}}}{\arg\min}\ R_b(\bar{\mathbf{b}}) + \tfrac{\lambda}{2}\textstyle\sum_{\omega=\omega_1}^{\omega_q} \|\mathbf{A}(\mathbf{m}^k, \omega)\mathbf{u}(\omega)^k - \bar{\mathbf{b}} + \frac{1}{\lambda}\mathbf{v}(\omega)^k\|_2^2, \tag{9}$$

where we have ignored the $-\|\mathbf{v}(\omega)^k\|_2^2$-term as it does not impact the optimization result. Eq. (9) is a denoising/proximity problem (Parikh & Boyd 2013) applied to $\bar{\mathbf{b}}$, that is

$$\bar{\mathbf{b}}^{k+1} = \underset{\bar{\mathbf{b}}}{\arg\min}\ R_b(\bar{\mathbf{b}}) + \frac{\lambda q}{2}\|(\frac{1}{q}\sum_{\omega=\omega_1}^{\omega_q} \mathbf{A}(\mathbf{m}^k, \omega)\mathbf{u}(\omega)^k + \frac{1}{\lambda}\mathbf{v}(\omega)^k) - \bar{\mathbf{b}}\|_2^2. \tag{10}$$

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:

$$R_b(x) = \begin{cases} |x| & |x| \le \varepsilon, \\ \frac{x^2+\varepsilon^2}{2\varepsilon} & |x| > \varepsilon, \end{cases} \tag{11}$$

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. (10) admits an explicit

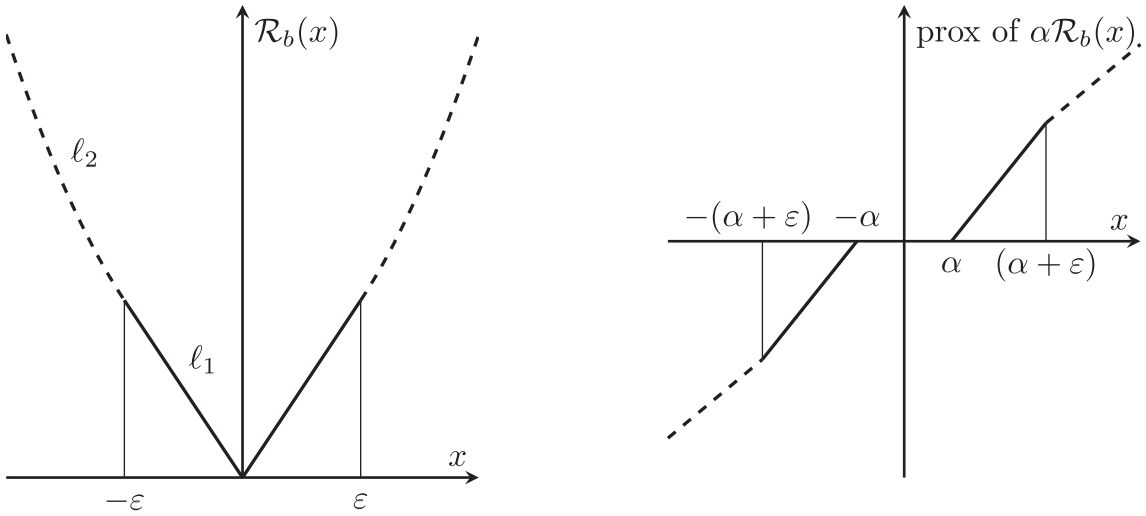


**Figure 1.** Left-hand graph: Berhu regularization function and (right-hand graph) the associated proximity operator. The dashed lines are due to the $\ell_2$ norm, the solid lines are due to the $\ell_1$ norm.

solution for each entry of the source vector

$$\bar{\mathbf{b}}^{k+1} = \text{prox}_{\frac{1}{\lambda q}B}\left(\frac{1}{q}\sum_{\omega=\omega_1}^{\omega_q}\mathbf{A}(\mathbf{m}^k,\omega)\mathbf{u}(\omega)^k + \frac{1}{\lambda}\mathbf{v}(\omega)^k\right), \tag{12}$$

where $\text{prox}_{\alpha B}$ is the proximity operator of the Berhu function defined as

$$\text{prox}_{\alpha B}(x) = \begin{cases} \max(1-\frac{\alpha}{|x|},0)x & |x| \le \alpha+\varepsilon, \\ \frac{\varepsilon}{\alpha+\varepsilon}x & |x| > \alpha+\varepsilon. \end{cases} \tag{13}$$

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:

$$\begin{pmatrix}\sqrt{\lambda}\mathbf{A}(\mathbf{m}^k,\omega)\\ \sqrt{\mu}\mathbf{P}\end{pmatrix}\mathbf{u}(\omega)^k = \begin{pmatrix}\sqrt{\lambda}\bar{\mathbf{b}}^{k+1}\\ \sqrt{\mu}\mathbf{d}(\omega)\end{pmatrix}, \qquad \omega = \omega_1\ ,\ ....,\ \omega_q, \tag{14}$$

and we alternate between solving eq. (14) for $\mathbf{u}(\omega)^k$ given $\bar{\mathbf{b}}^{k+1}$ and updating $\bar{\mathbf{b}}^{k+1}$ via eq. (12) using $\mathbf{u}(\omega)^k$ obtained from eq. (14).

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

$$\begin{pmatrix}\sqrt{\lambda}\mathbf{A}_\omega^k & -\sqrt{\lambda}\mathbf{\Phi}^{k+1}\\ \sqrt{\mu}\mathbf{P} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{u}(\omega)^{k+1}\\ \mathbf{s}(\omega)^{k+1}\end{pmatrix} = \begin{pmatrix}\frac{1}{\sqrt{\lambda}}\mathbf{v}(\omega)^k\\ \sqrt{\mu}\mathbf{d}(\omega) - \frac{1}{\sqrt{\mu}}\mathbf{w}(\omega)^k\end{pmatrix}. \tag{15}$$

### 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 total variation (TV) regularization (Rudin *et al*. 1992) for $\mathbf{m}$. 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

$$R_m(\mathbf{m}) = \sum\sqrt{(\nabla_x\mathbf{m})^2 + (\nabla_z\mathbf{m})^2}, \tag{16}$$

where $\nabla_x$ and $\nabla_z$ are, respectively, first-order difference operators in the horizontal and vertical directions with appropriate boundary conditions. The problem (7b) with non-smooth TV regularization and bound constraints can be written as

$$\mathbf{m}^{k+1} = \underset{\mathbf{m}\in M}{\arg\min}\ R_m(\mathbf{m}) + \lambda(\mathbf{m}^T\mathbf{H}_k\mathbf{m} - \mathbf{g}_k^T\mathbf{m}), \tag{17}$$

where

$$\mathbf{H}_k = \sum_{\omega=\omega_1}^{\omega_q} \left( \frac{\partial \mathbf{A}(\mathbf{m})}{\partial \mathbf{m}} \mathbf{u}(\omega)^{k+1} \right)^T \left( \frac{\partial \mathbf{A}(\mathbf{m})}{\partial \mathbf{m}} \mathbf{u}(\omega)^{k+1} \right), \tag{18}$$

$$\mathbf{g}_k = \sum_{\omega=\omega_1}^{\omega_q} \left( \frac{\partial \mathbf{A}(\mathbf{m})}{\partial \mathbf{m}} \mathbf{u}(\omega)^{k+1} \right)^T (\mathbf{\Phi}^{k+1} \mathbf{s}(\omega)^{k+1} - \frac{1}{\lambda} \mathbf{v}(\omega)^k - \nabla^2 \mathbf{u}(\omega)^{k+1}). \tag{19}$$

The model subproblem in eq. (17) requires a quadratic optimization with bound constrained TV regularization. There are many well documented algorithms to carry out this step (Goldstein & Osher 2009; Maharramov & Levin 2015; Aghamiry *et al*. 2019a, 2020; Gholami & Naeini 2019).

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 set of MSEs is processed with a sufficiently-accurate background subsurface model estimated during a prior inversion. The linear systems for data-assimilated wavefield reconstruction, eqs (6), (14) and (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 data sets 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 after a good background model has been estimated, store them on disk and reuse them to process a new data set efficiently by forward/backward elimination.

**Algorithm 1** ADMM-based FWI for microseismic imaging.

1: Begin with $k = 0$ and an initial model $\mathbf{m}^0$.
2: Set to zero the values of $\mathbf{v}(\omega)^0$ and $\mathbf{w}(\omega)^0$, $\forall\omega$.
3: Calculate $\mathbf{u}(\omega)^0$, $\forall\omega$ (eq. 6).
4: **while** convergence criteria not satisfied **do**
5: **for** $l = 0 : n_l$ **do**
6: Compute $\bar{\mathbf{b}}^{k+1}$ (eq. 12)
7: Update $\mathbf{u}(\omega)^k$, $\omega = \omega_1, \cdots, \omega_q$ (eq. 14) } Source location estimation.
8: **end for**
9: Build $\mathbf{\Phi}^{k+1}$ from $\bar{\mathbf{b}}^{k+1}$ using a peak finder algorithm.
10: Update $\mathbf{s}(\omega)^{k+1}$ and $\mathbf{u}(\omega)^{k+1}$ for $\omega = \omega_1, \cdots, \omega_q$ (eq. 15). } Joint update of wavefields and source signatures.
11: Update $\mathbf{m}^{k+1}$ (eq. 17) } Update velocity model.
12: Update the dual vectors $\mathbf{v}$ and $\mathbf{w}$ (eqs. 7c and 7d) } Iterative refinement.
13: $k = k + 1$ ,
14: **end while**

## 3 NUMERICAL EXAMPLES

We assess the method on a subset of the synthetic Marmousi II model (Martin *et al*. 2006) with size 2.5 km × 1 km in distance and depth with a grid spacing of 5 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. 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. 2(d). The different MSEs, which are used in this section, are shown in Fig. 2(a) 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 1–4 are Ricker wavelets with central frequencies [25, 31, 23, 29] Hz and central times [2.4, 2.56, 2.25, 2.2] s, respectively. Note that we assign a different peak amplitude to each wavelet, so that the sensitivity of the inversion varies from one MSE to the next. The maximum amplitude of the third and fourth MSEs is twice that of the first MSE and four times greater than that of the second MSE. Also, we use a line of receivers spaced 15 m apart at the surface to record data.

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

(i) We set $\varepsilon$, which controls the transition between $\ell_1$ and $\ell_2$, equal to $0.02 \times \max |\frac{1}{q} \sum_{\omega=\omega_1}^{\omega_q} \mathbf{A}(\mathbf{m}^k, \omega)\mathbf{u}(\omega)^k + \frac{1}{\lambda}\mathbf{v}(\omega)^k|$.
(ii) We define $\lambda$ as a percentage of the inverse of the maximum of the diagonal coefficients of $\mathbf{H}_k$, eq. (18).

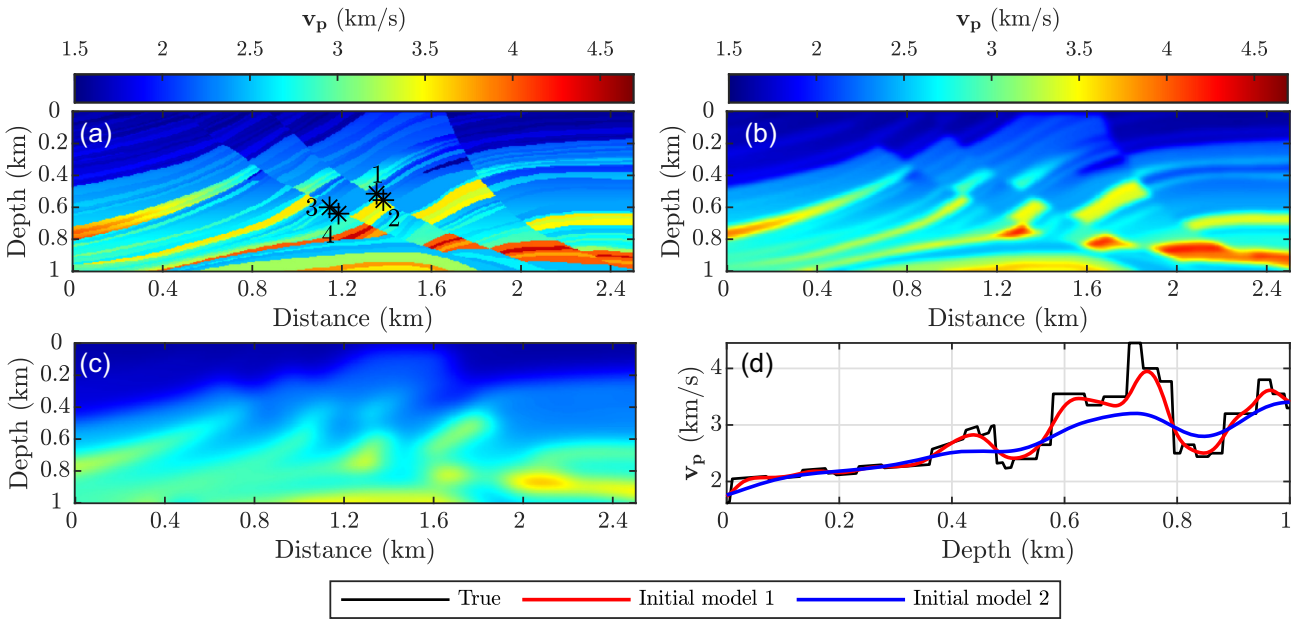


**Figure 2.** True and initial velocity models and the locations of MSEs that are used for the numerical tests. (a) The Marmousi II velocity model, used to generate the data. The location of different MSEs, referred as MSEs 1–4, are shown by black asterisks. (b) Accurate initial model, called initial 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.

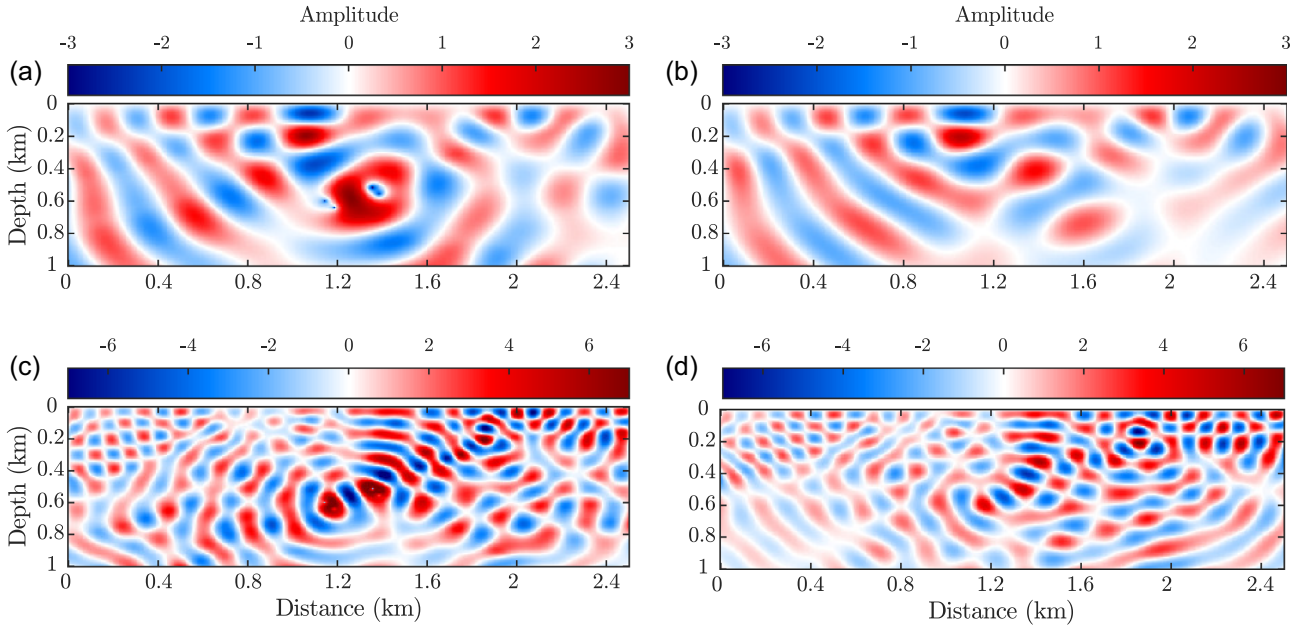


**Figure 3.** Showing the ability of the data-assimilated wavefield to mimic the true wavefields when no information on the source is available. The Marmousi model with a source containing MSEs 1–4 is used. (a) True 7 Hz wavefield. (b) 7 Hz data-assimilated wavefield (eq. 6) with initial model 2. (c–d) Same as (a–b), but for 15 Hz.

(iii) We set $\mu$ such that $\lambda/\mu$ is a small fraction of the highest eigenvalue ($\xi$) of the normal operator $\mathbf{A}(\mathbf{m}^k, \omega)^{-T}\mathbf{P}^T\mathbf{P}\mathbf{A}(\mathbf{m}^k, \omega)^{-1}$ according to the criterion proposed by van Leeuwen & Herrmann (2016). We use $\lambda/\mu = 1\text{e-}5\xi$ and $\lambda/\mu = 1\text{e-}3\xi$ for noiseless and noisy data, respectively.

We first illustrate how the initial data-assimilated wavefield is reconstructed [when no information on the source is available, eq. (6)]. The source contains MSEs 1–4 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 and 15 Hz frequencies, which are reconstructed by back-propagating the data, eq. (6). The ability of the data-assimilated wavefield to mimic the true wavefield can be assessed by comparing the left- and right-hand columns of this figure. This emphasizes how suitable the (transmission) microseismic configuration is to perform accurate wavefield reconstruction

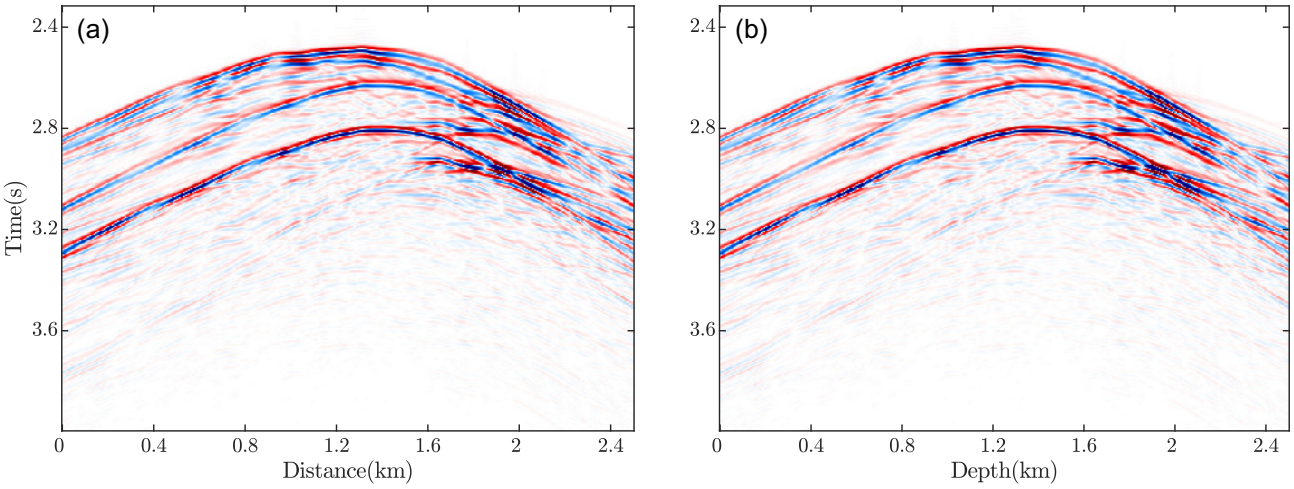


**Figure 4.** Showing the match between the true and data assimilated seismograms when no information on the source is available. The Marmousi model with a source containing MSEs 1–4 is used. (a) True seismograms. (b) Seismograms created using data-assimilated wavefields (eq. 6) with initial model 2.

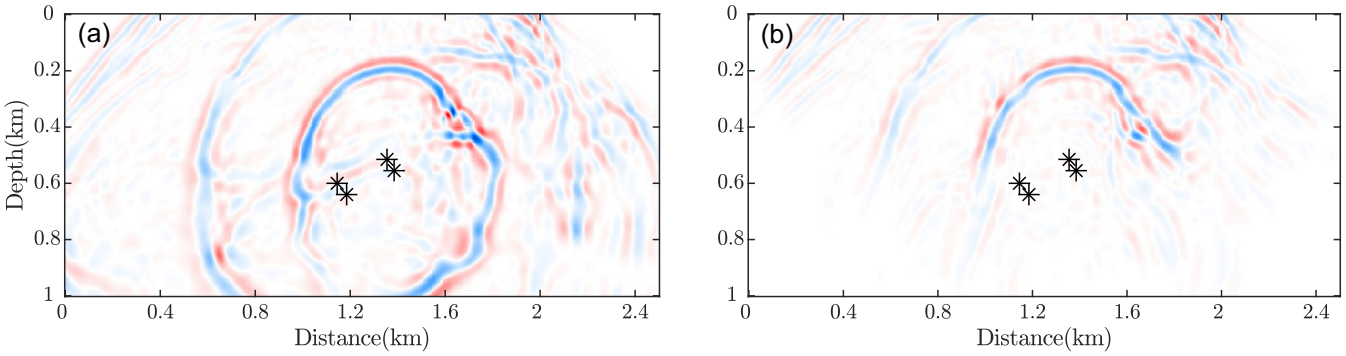


**Figure 5.** Showing the ability of data-assimilated wavefield to capture many features of the true wavefield above the MSEs when no information on the source is available. The Marmousi model with a source containing MSEs 1–4 is used. A time slices at 2.7 s for (a) the true wavefield and (b) the data-assimilated wavefield with initial model 2.

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. 4(a), while those computed in the initial model with data assimilation, when no information about the source is available, are shown in Fig. 4(b). 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. 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. 6(c) in red. 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 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 ($n_l = 9$). We can also consider the relative change of predicted source ($\frac{\|\bar{\mathbf{b}}^{k+1}-\bar{\mathbf{b}}^{k}\|_2}{\|\bar{\mathbf{b}}^{k}\|_2}$) as the stopping criteria, that is when this value becomes small. This second criteria would likely be more appropriate for field data. The estimated $\bar{\mathbf{b}}^1$ at the final iteration of the inner loop is shown in Fig. 6(a). The event location found by the peak finder algorithm is shown in Fig. 6(b) in blue while the true one is shown in red. The results are zoomed in from the full model; the limits of the zoom are shown with the black dashed lines in Fig. 6(a). 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. 6(c) (blue 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 ($\bar{\mathbf{b}}^1$) is shown in Fig. 7(a) and the result of the application of the peak finder algorithm is shown in Fig. 7(b) in blue. 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 blue in Fig. 7(e). We can clearly see that neither the estimated MSE location nor the estimated signature is accurate. To

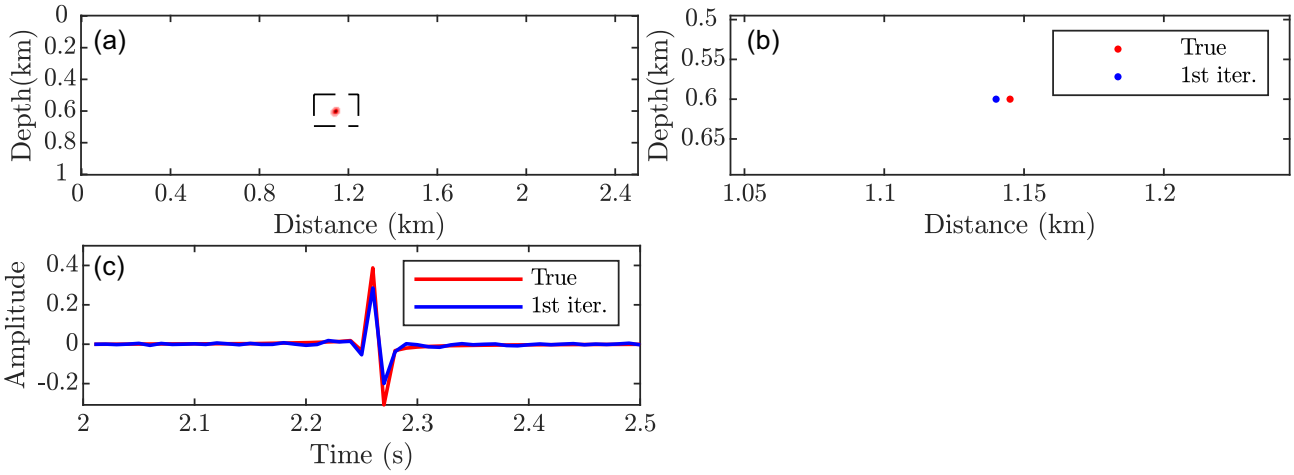


**Figure 6.** Marmousi test with initial model 1 when the source just contains MSE 1. (a) The predicted source at the first outer iteration ($\bar{\mathbf{b}}^1$). (b) The zoomed result of peak finder applied on $\bar{\mathbf{b}}^1$. The boundaries of the zoomed part are shown using black dashed lines in (a). Also, the location of the true source is indicated by a red point. (c) The true source signature (red) and the estimated one (blue).

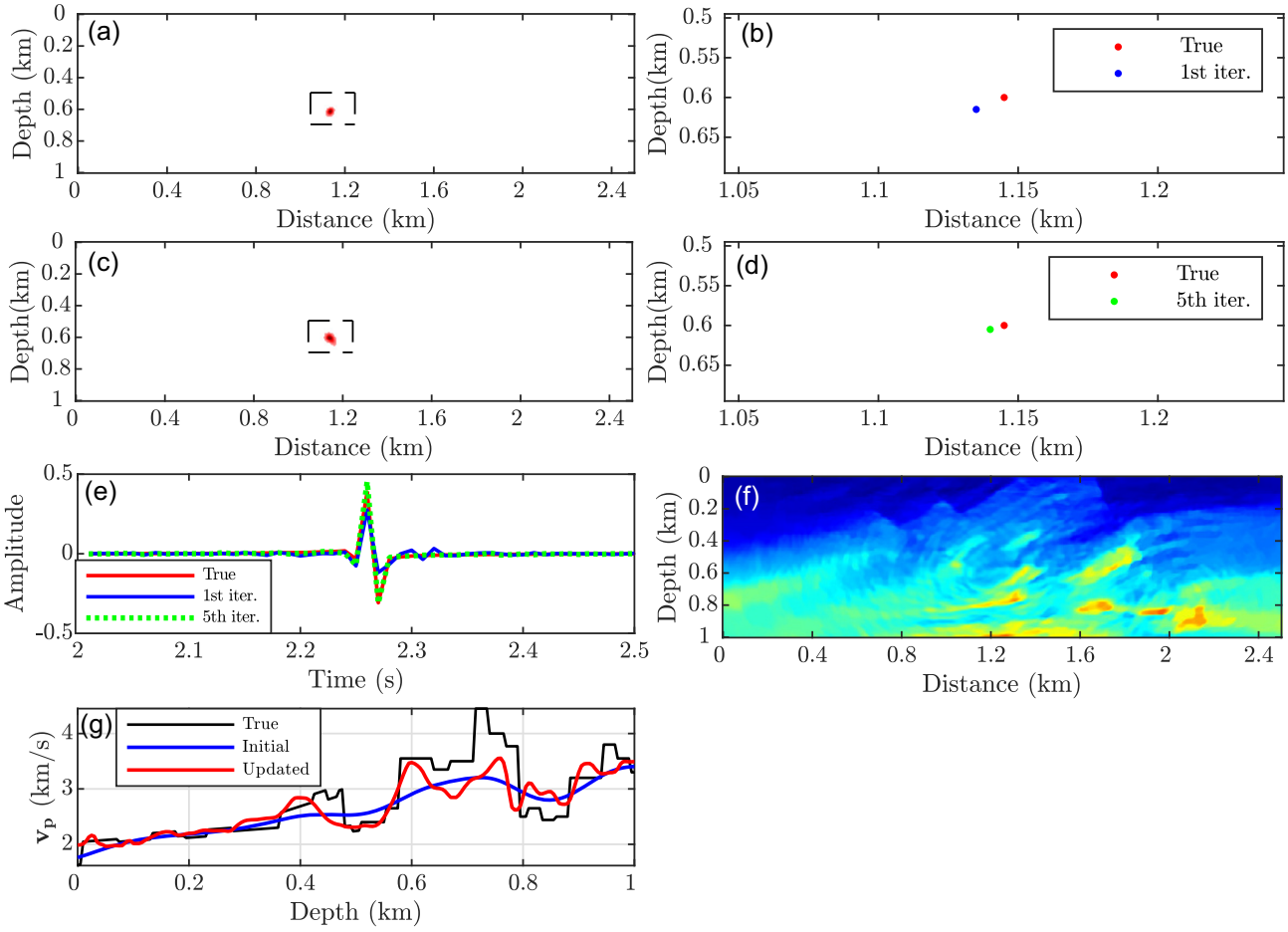


**Figure 7.** Marmousi test with initial model 2 when the source just contains MSE 1. (a) Predicted source at the final inner iteration of outer iteration 1 ($\bar{\mathbf{b}}^1$). (b) The zoomed result of peak finder applied on $\bar{\mathbf{b}}^1$. The boundaries of the zoomed part are shown using black dashed lines in (a). Also, the location of the true source is indicated by a red point. (c–d) Same as (a–b) but for $\bar{\mathbf{b}}^5$. (e) The true source signature (red) and the estimated at outer iteration 1 (blue) and iteration 5 (green). (f) Updated velocity model after five iterations. (g) Direct comparison between true (black), initial model 2 (blue) and updated velocity model (red) at $X = 1250$ m.

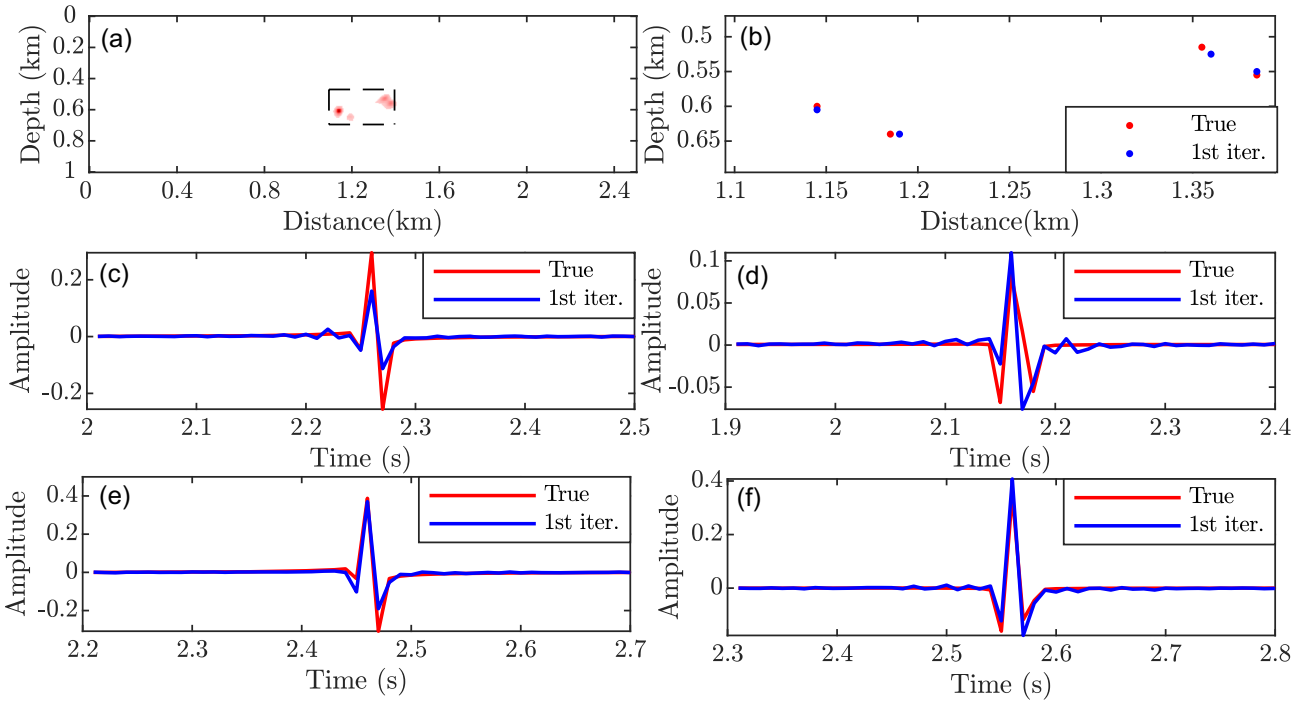


**Figure 8.** The Marmousi test with initial model 1 when the source contains MSEs 1–4. (a–b) $\bar{\mathbf{b}}^1$ and its selected peaks using the peak finder. (c–f) True source signatures (red) and the estimated ones (blue) for MSEs 1–4, respectively.

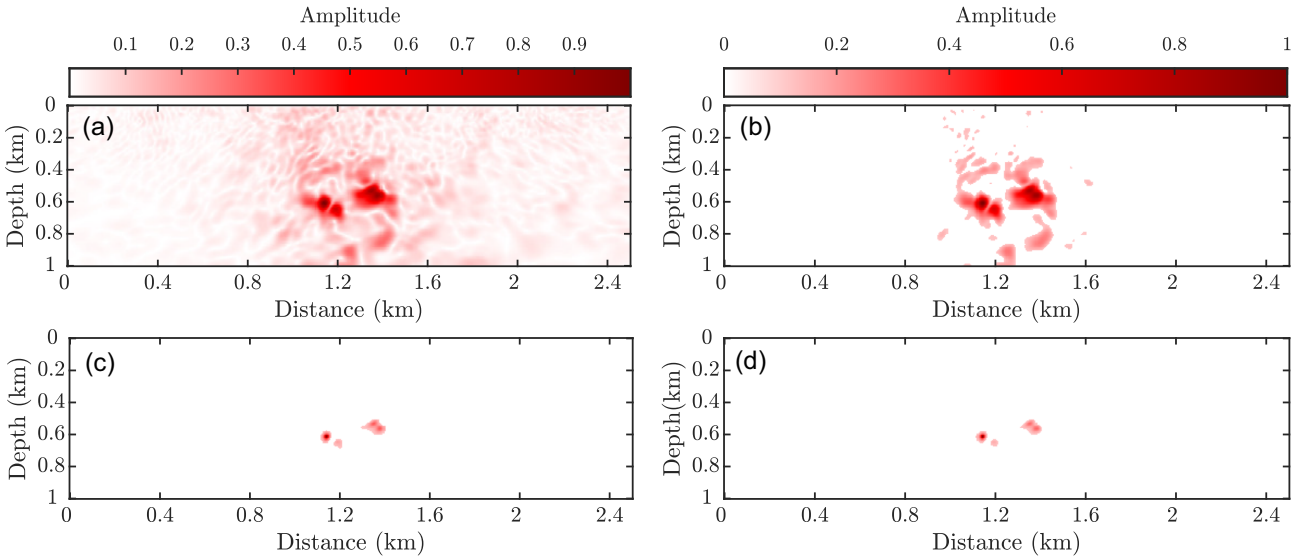


**Figure 9.** Showing the impact of sparsifying regularization on the estimated predicted sources for the Marmousi test with initial model 1 when the source contains MSEs 1–4. (a–b) The predicted source at $k = 0$, $l = 0$ (a) before, (b) after Berhu regularization. (c–d) Same as (a–b) but for $k = 0$,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

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), $\bar{\mathbf{b}}^5$, and the result of the peak finder algorithm are shown in Figs 7(c) and (d), respectively. Finally, the estimated source signature at outer iteration 5 is shown in green in Fig. 7(e). 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. 7(f). 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.

To make the test more representative of the microseismic scenario, we repeat the Marmousi test with the source gathers MSEs 1–4 (Fig. 2a). We start with the initial model 1. The predicted source after 1 outer iteration ($\bar{\mathbf{b}}^1$) and the result of peak finder are shown in Figs 8(a) and (b). As the initial model is kinematically accurate, there are just four spots of energy in Fig. 8. Also, the estimated source signatures for

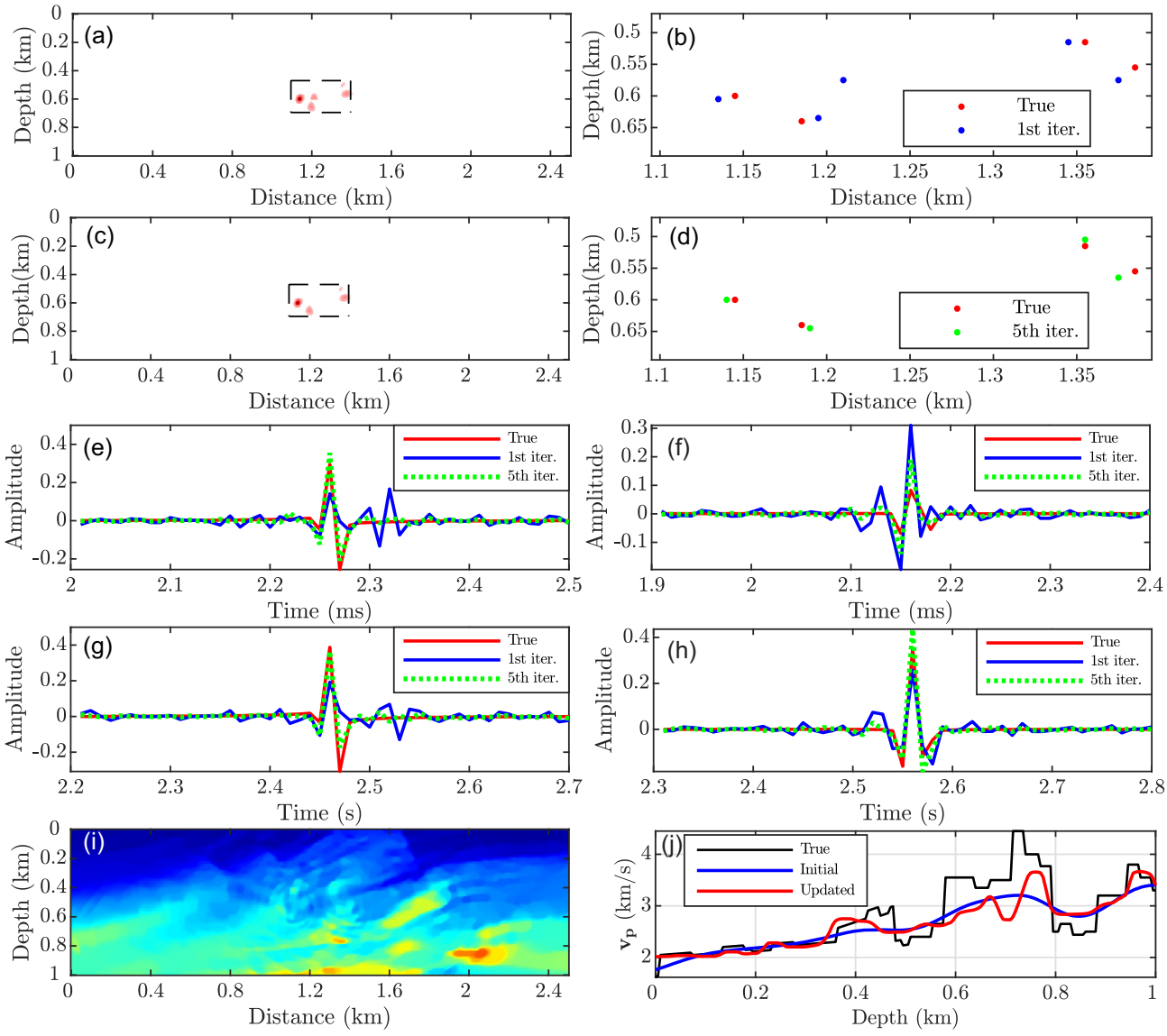


**Figure 10.** The Marmousi test with initial model 2 when the source contains MSEs 1–4. (a–b) $\bar{\mathbf{b}}^1$ and its selected peaks using peak finder. (c–d) $\bar{\mathbf{b}}^5$ and its selected peaks. (e–h) True MSE signatures (red) and the estimated ones for MSEs 1–4, respectively, at outer iteration one (blue) and iteration five (green). (i) The updated velocity model after five iterations. (j) A direct comparison between true (black), initial model 2 (blue) and updated velocity model (red) at $X =$ 1250 m.

MSEs 1–4 are shown in Figs 8(c)–(f). We have accurately recovered both the locations and signatures for all MSEs without trade-off between these variables.

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 sources for this test at inner iteration 1 ($k = 0, \mathrm{l} = 0$) before and after applying Berhu regularization are depicted in Figs 9(a) and (b). We can see the improved focusing provided by the regularization, 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 9(c) and (d). First, we see in Fig. 9(d) that, near the convergence point, the regularization has little effect on the predicted source because it is already sparse (compared to Fig. 9c). Secondly, we show the improved focusing of the MSE location in Fig. 9(d) compared to Fig. 9(b), due to the combined impact of the inner iterations and sparsity-promoting regularization.

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 $\bar{\mathbf{b}}^1$ (at the final inner iteration) and its located peaks are shown in Figs 10(a) and (b) and the same results for outer iteration five are shown in Figs 10(c) and (d). 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. (15). Because the selected MSE at the middle of Fig. 10(b) 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

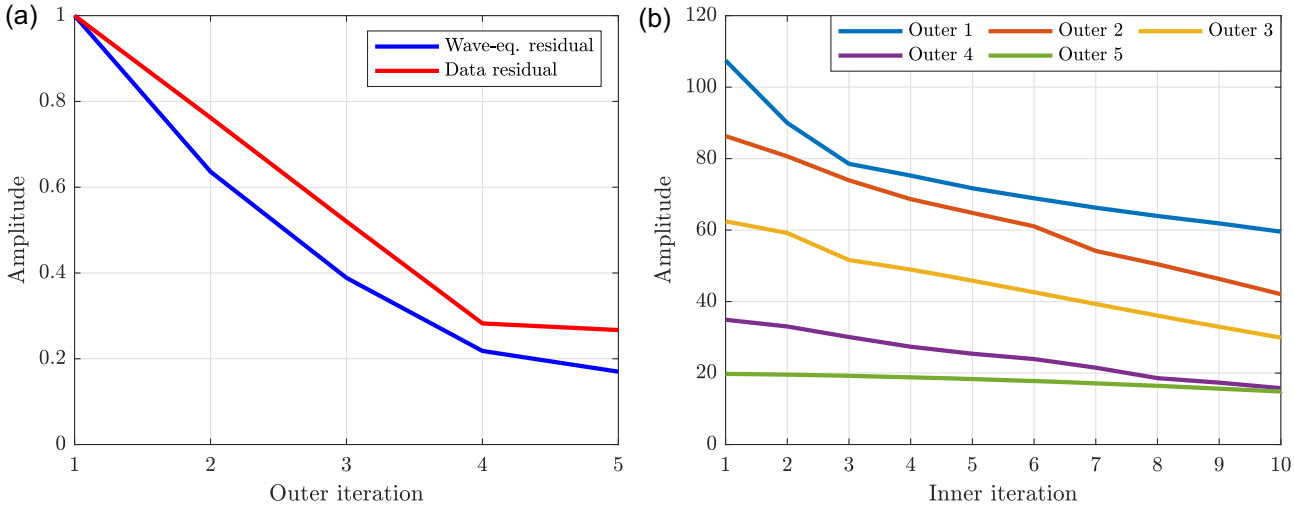


**Figure 11.** Showing the convergence behaviour of the algorithm for the Marmousi test with initial model 2 when the source contains MSEs 1–4. (a) Wave-equation (blue) and and data (red) residuals for different outer iteration where their maximum value is scaled to one. (b) Data residuals of different outer iterations versus inner iterations.

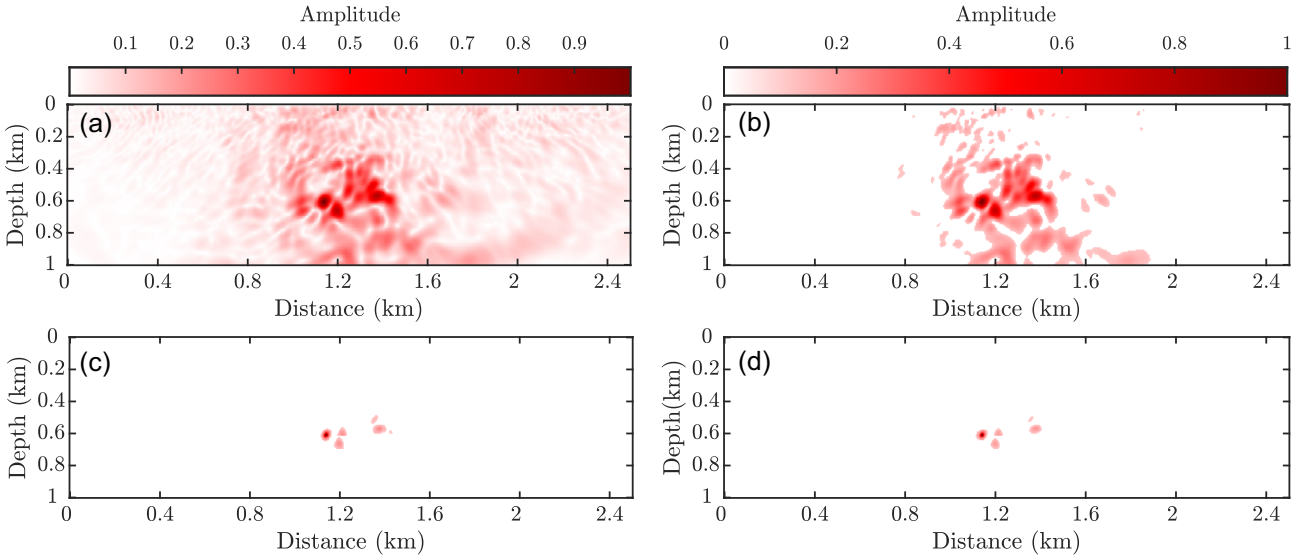


**Figure 12.** Showing the impact of sparsifying regularization on the estimated predicted sources for the Marmousi test with initial model 2 when the source contains MSEs 1–4. (a–b) The predicted source at $k = 0$, $l = 0$ (a) before, (b) after Berhu regularization. (c–d) Same as (a–b) but for $k = 0$, $l = 9$. $k$ and $l$ refer to the outer and inner iteration of algorithm 1, respectively.

true MSEs (Fig. 10d). Also, the estimated MSE signatures at outer iterations 1 and 5 are shown in Figs 10(e)–(h), respectively, for MSEs 1–4 in blue for outer iteration one and in green for outer iteration five. Finally, the updated velocity model after five outer iterations is shown in Fig. 10(i). First, the quality of the estimated signatures of the MSEs is improved by updating the velocity model. Secondly, the updated velocity model captures the trends of the dominant structures, and is improved compared to the initial model 2.

To give some insight into the convergence of the proposed algorithm, the wave-equation and data residuals for different outer iterations are shown in Fig. 11(a). Also, the data residual of each outer iterations versus inner iterations are shown in Fig. 11(b). We see that the data residuals decrease in inner iterations (for example, the blue curve in Fig. 11b), but we should remember that both the source term and velocity model are inaccurate. When the algorithm updates both of these variables, the data residual increases (the red curve in Fig. 11b), and the algorithm again tries to decrease this value. These decrements are not for all the outer iterations. We see that after a few outer iterations, because of insignificant updates in both the velocity model and the source term, we do not see significant decreases in the residual curves.

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 12(a) and (b), respectively, and for inner iteration 10 ($k = 0$, $l = 9$) shown in Figs 12(c) and (d). We see that the energy is less focused compared to the case with an accurate initial model (Figs 9a and b). We also see that the regularization significantly improves the predicted source map, making it ready for MSEs location picking.

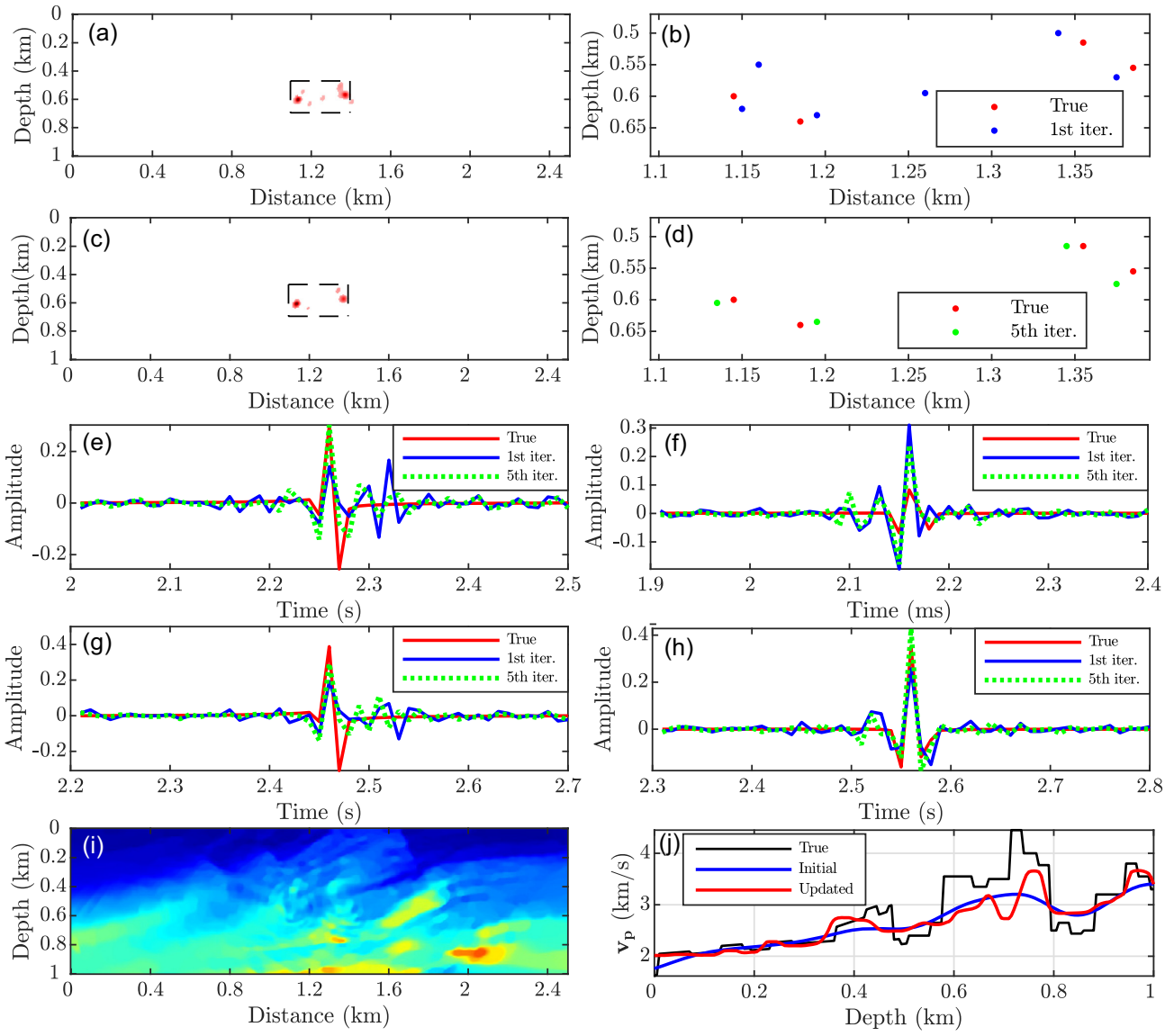


**Figure 13.** The Marmousi test with initial model 2 and noisy data with SNR = 5 dB when the source contains MSEs 1–4. (a–b) $\bar{\mathbf{b}}^1$ and its selected peaks using peak finder. (c–d) $\bar{\mathbf{b}}^5$ and its selected peaks. (e–h) True MSE signatures (red) and the estimated ones for MSEs 1–4, respectively, at outer iteration one (blue) and iteration five (green). (i) The updated velocity model after five iterations. (j) A direct comparison between true (black), initial model 2 (blue) and updated velocity model (red) at $X = 1250$ m.

We now assess 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. 13. 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. 13a), but it is improved at the fifth outer iteration (Fig. 13c). Also, we see that the estimated MSEs are not significantly changed in Fig. 13(c) compared to the noiseless case (Fig. 10c). For the estimated signatures, Figs 13(e)–(h), we see that the quality of the estimated signatures are degraded compared to the noiseless case (Figs 10e–h). Finally, the updated velocity model for this noisy test, Fig. 13(i), is close to the noiseless one in Fig. 10(i). Also, the behaviour of the wave-equation and data residuals as a function of outer iterations as well as data residuals for each outer iteration are shown in Fig. 14. Again, the same behaviour as noiseless data can be seen except for the value of data residuals in Fig. 14(b), which is larger because of the noise in data.

## 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

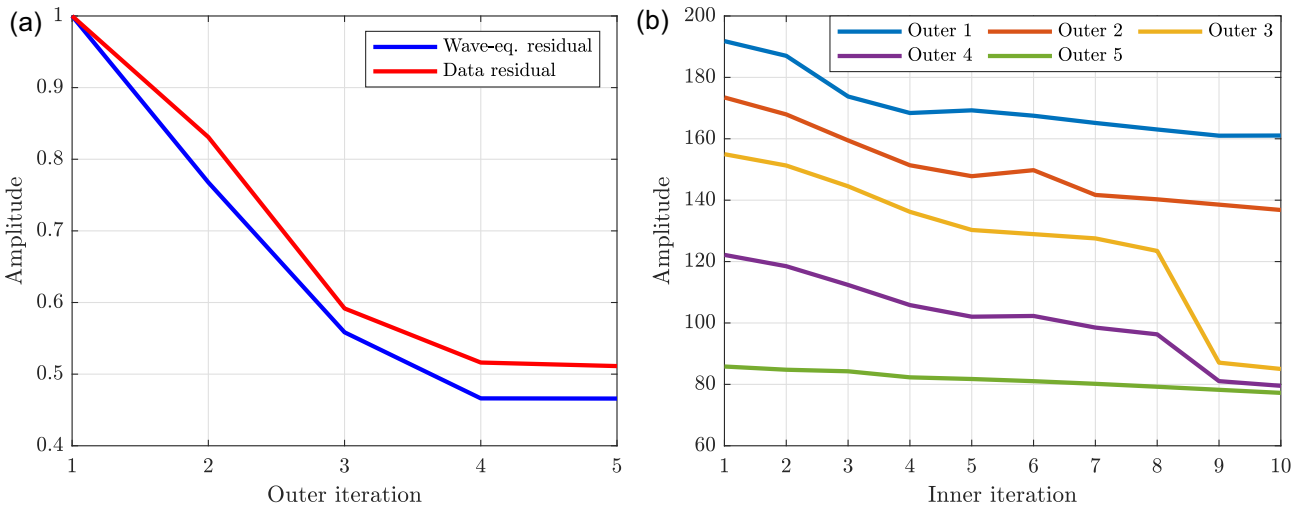


**Figure 14.** Showing the convergence behaviour of the algorithm for the Marmousi test with initial model 2 when the source contains MSEs 1–4 and the data are contaminated with noise. (a) Wave-equation (blue) and and data (red) residuals for different outer iteration where their maximum value is scaled to one. (b) Data residuals of different outer iterations versus inner iterations.

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 signature of seismic events and (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, Fang *et al*. (2018) and Aghamiry *et al*. (2021b) 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. (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 reach a good estimate of the seismic events. This is required to manage the potential ambiguity or trade-off between event location and velocity model. However, IR-WRI can contribute to refine the velocity model in the region through which the recorded waves travel resulting in improved event location. Indeed, for limited subsurface illumination provided by a few seismic events, the velocity-estimation subproblem should be viewed as a proxy to improve the location of the events rather than as a velocity model building tool *per se*. Moreover, since the events are buried in the subsurface, IR-WRI mostly exploits the transmission regime of wave propagation (i.e. forward scattering), hence further limiting the resolution power of the velocity updating to the long to intermediate wavelengths. The reader is referred to Virieux & Operto (2009) for a more comprehensive review of the sensitivity of the FWI resolution to the acquisition design.

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 (14) and (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 pre-conditioner (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 is negligible.

## 5 CONCLUSIONS

We proposed a method based on ADMM-based FWI for characterizing weak seismic events. 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 MSE 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 3-D elastic physics and moment tensor estimation in the prospect of real data applications.

## ACKNOWLEDGEMENTS

We would like to thank editors F. Simons and L. Alexander and also thank two anonymous reviewers for their comments and suggestions. 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. AM acknowledges the NSERC Discovery Grant Program as well as Chevron, InnvoateNL and the NSERC Industrial Research Chair program.

## DATA AVAILABILITY

The data underlying this paper cannot be shared publicly due to the privacy of individuals that participated in the study.

## 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, doi.org/10.1190/1.3250203.

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, *Geophys. J. Int.,* **218**(2), 855–872, doi.org/10.1093/gji/ggz189.

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, doi.org/10.1190/geo2018-0093.1.

Aghamiry, H., Gholami, A. & Operto, S., 2020. Compound regularization of full-waveform inversion for imaging piecewise media, *IEEE Trans. Geosci. Remote Sens.,* **58**(2), 1192–1204.

Aghamiry, H., Gholami, A. & Operto, S., 2021a. Full waveform inversion by proximal newton methods using adaptive regularization, *Geophys. J. Int.,* **224**(1), 169–180, doi: 10.1109/TGRS.2019.2944464.

Aghamiry, H.S., Mamfoumbi-Ozoumet, F.W., Gholami, A. & Operto, S., 2021b. Efficient extended-search space full-waveform inversion with unknown source signatures, *Geophys. J. Int.,* **227**(1), 257–274, doi.org/10.1093/gji/ggab202.

Aravkin, A.Y., van Leeuwen, T., Calandra, H. & Herrmann, F.J., 2012. Source estimation for frequency-domain FWI with robust penalties, in *Proceedings of the 74th EAGE Conference & Exhibition incorporating SPE EUROPEC 2012, Expanded Abstracts,* Copenhagen, Denmark, 4–7 June 2012.

Chen, Z., Cheng, D., Feng, W. & Wu, T., 2013. An optimal 9-point finite difference scheme for the Helmholtz equation with PML, *Int. J. Numer. Anal. Model.,* **10**(2), 389–410.

De Ridder, S. & Maddison, J., 2018. Full wavefield inversion of ambient seismic noise, *Geophys. J. Int.,* **215**(2), 1215–1230, doi.org/10.1093/gji/ggy328.

Dryja, M. & Widlund, O., 1987. *An Additive Variant of the Schwarz Alternating Method for the Case of Many Subregions, Ultracomputer Research Laboratory, Division of Computer Science,* Univ., Courant Inst. of MathematicalSciences.

Fang, Z., Wang, R. & Herrmann, F.J., 2018. Source estimation for wavefield-reconstruction inversion, *Geophysics,* **83**(4), R345–R359, doi.org/10.1190/geo2017-0700.1.

Gajewski, D. & Tessmer, E., 2005. Reverse modelling for seismic event characterization, *Geophys. J. Int.,* **163**(1), 276–284, doi.org/10.1111/j.1365-246X.2005.02732.x.

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 J. Imag. Sci.,* **2**(2), 323–343, doi.org/10.1137/080725891.

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, doi.org/10.1190/geo2017-0557.1.

Lee, K.H. & Kim, H.J., 2003. Source-independent full-waveform inversion of seismic data, *Geophysics,* **68,** 2010–2015, doi.org/10.1190/1.1635054.

Li, L. *et al.*, 2020. Recent advances and challenges of waveform-based seismic location methods at multiple scales, *Rev. Geophys.,* **58**(1), e2019RG000667, doi.org/10.1029/2019RG000667.

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

Martin, G.S., Wiley, R. & Marfurt, K.J., 2006. Marmousi2: an elastic upgrade for Marmousi, *Leading Edge,* **25**(2), 156–166, doi.org/10.1190/1.2172306.

McMechan, G.A., 1982. Determination of source parameters by wavefield extrapolation, *Geophys. J. R. astr. Soc.,* **71,** 613–628, doi.org/10.1111/j.1365-246X.1982.tb02788.x.

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

Montgomery, C.T., Smith, M.B. *et al.*, 2010. Hydraulic fracturing: history of an enduring technology, *J. Petrol. Technol.,* **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),** KS51–KS60.

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

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

Parikh, N. & Boyd, S., 2013. Proximal algorithms, *Found. Trends Opt.,* **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, *Geophys. J. Int.,* **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, *Geophys. J. Int.,* **133,** 341–362, doi.org/10.1046/j.1365-246X.1998.00498.x.

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

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

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

Sharan, S., Wang, R. & Herrmann, F.J., 2019. Fast sparsity-promoting microseismic source estimation, *Geophys. J. Int.,* **216,** 164–181, doi.org/10.1093/gji/ggy415.

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, *J. Scient. Comput.,* **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 *SEG Technical Program Expanded Abstracts,* Society of Exploration Geophysicists, pp. 1405–1410.

Thurber, C. & Engdahl, E., 2000. *Advances in Seismic Event Location,* Springer.

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, *Geophys. J. Int.,* **195**(1), 661–667, doi.org/10.1093/gji/ggt258.

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