Consistent seismic event location and subsurface parameters inversion through slope tomography: a variable-projection approach

Serge Sambolian, Stéphane Operto, Alessandra Ribodetti, Jean Virieux

Geophysical Journal International

Geophys. J. Int. (2021) **224,** 1956–1979
Advance Access publication 2020 November 18
GJI Seismology

doi: 10.1093/gji/ggaa555

Consistent seismic event location and subsurface parameters inversion through slope tomography: a variable-projection approach

S. Sambolian [1], S. Operto [1], A. Ribodetti[1] and J. Virieux[2]

[1] *Université Côte dAzur, Observatoire de la Côte dAzur, Observatoire de la Côte dAzur, CNRS, IRD, Géoazur,* 06560 *Valbonne, France.*
E-mail: sambolian@geoazur.unice.fr
[2] *Université Grenoble Alpes, ISTerre,* 38400 *Saint-Martin-d'Hères, France.*

Accepted 2020 November 16. Received 2020 October 9; in original form 2020 August 4

SUMMARY

We revisit the hypocentre–velocity problem, which is of interest in different fields as for example microseismics and seismology. We develop a formulation based on kinematic migration of two picked kinematic attributes in the 2-D case, the traveltime and the slope (horizontal component of the slowness vector), from which we are able to retrieve the location and subsequently the origin time correction and the subsurface parameters mainly velocity. We show how, through a variable projection, the optimization problem boils down to a physically consistent and parsimonious form where the location estimation is projected into the subsurface parameter problem. We present in this study a proof of concept validated by a toy test in two dimensions and a synthetic case study on the Marmousi model. The method presented in this study is extendible to three dimensions by incorporating the crossline slope or the backazimuth as a supplementary attribute.

Key words: Inverse theory; Tomography; Earthquake source observations; Seismic tomography.

1 INTRODUCTION

The hypocentre–velocity problem has been a challenging topic of interest in geophysics with its main purpose being the localization of seismic events. The source location problem by itself has been extensively researched for different applications and purposes. On a macro scale, locating the origin of earthquakes is crucial to investigate the geology and dynamics of active margins (Roecker 1982), while at a smaller scale for reservoir monitoring and characterization purposes using arrays (Grechka *et al.* 2010; Deflandre 2016) or borehole recordings (Jones *et al.* 2013). Since the pioneering work of Geiger (1912), the source location is often approximated in a least-squares inversion sense and grid-search methods (Lomax *et al.* 2009). Many variants of traveltime-based localization methods emerged. Some approaches differ in terms of optimization, as an example graphical methods like the master event method (Zhou 1994) and the maximum intersection method (Font *et al.* 2004) versus direct nonlinear location algorithms (Lomax *et al.* 2000), while others differ in the manner of handling the data in terms of acquisition or attributes. The primary physical attribute, picked arrival times, are often supplemented by other attributes like the slowness and the azimuth in order to perceive the wave arrival azimuthal plane and incidence angle. Differential attributes are often extracted from the data through array-based processing techniques (Rost & Thomas 2002) like beamforming (Krüger *et al.* 1993; Verdon *et al.* 2017), double-difference (Waldhauser & Ellsworth 2000) or polarization analysis, which is also utilized in single-station location techniques (Frohlich & Pulliam 1999).

Using directional attributes like the slowness and the backazimuth is crucial to better constrain the location problem. Indeed, knowing the emergence angle and the plane of arrival restrain the grid-search space for optimal location candidates whatever is the technique employed. The latter notion is not only factual for traveltime-based techniques but also finite-frequency waveform-based methods. Most waveform-based techniques are based on time-reversal (McMechan 1982; Fink 1993; Rietbrock & Scherbaum 1994), which consists of propagating backwards in time the recordings of all receivers and eventually refocusing the energy at its point of origination in both space and time. Time-reversal techniques are of interest compared to traveltime-based approaches since they naturally utilize the full-waveform data without the need of picking or labelling arrivals (Gajewski & Tessmer 2005; Larmat *et al.* 2006). Artman *et al.* (2010) proposed an improved imaging condition based on the cross-correlation of *P* and *S* wavefield components, valid beyond the acoustic approximation. In order to control the focusing of the weighted backprojected recording, the location problem was also recast in a seismic migration sense, employing for example interferometry-based (Schuster *et al.* 2004; Li *et al.* 2015) or Kirchhoff-like imaging conditions (Baker *et al.* 2005). We refer the reader to

Li *et al.* (2020) for a comprehensive review on waveform-based source location techniques. In all aforementioned localization approaches, the energy of the backpropagated wavefield would smear around the source location depending on the acquisition spread, the robustness of the imaging condition, the inaccuracy of the subsurface parameters mainly velocity and of course the physics governing the wave equation employed during the modelling. As a remedy, in the same manner as for traveltime-based approaches, array-based processing is introduced like for example Gaussian-beam migration (Rentsch *et al.* 2007), time-domain local stacking (Ishii *et al.* 2007) and slowness-backazimuth weighted migration (Kito *et al.* 2007; Kito & Korenaga 2010).

We now look back at the main problem addressed in this paper: the hypocentre–velocity problem. The success of all localization methods depends on the accuracy of the subsurface-parameter models. The opposite is also true, inverting for the velocity structure using wrong source locations would lead to inaccurate velocity updates. In addition, if the subsurface parameters are wrong, the origin time needs to be constrained too since the time-reversed data do not intersect at their origin time. Even if the main objective of a study is retrieving the location of a seismic event at any scale, the subsurface parameters should be updated in order to account for the inaccuracies of the model. Indeed, the coupling between the source location, the origin time of the event and the subsurface parameters makes the hypocentre–velocity problem challenging (Thurber 1992). We focus for the rest of our discussion solely on the velocity as subsurface parameter since it is the primary parameter of interest. A possible strategy for the hypocentre–velocity problem would be to ignore the event position and velocity structure coupling and proceed with two alternating-direction monoparameter optimizations for velocity and event location (Monteiller *et al.* 2005). Relaxing the original fully coupled problem could be inefficient in terms of optimization and is not even guaranteed to converge (Roecker *et al.* 2006, their appendix A). Jointly inverting for the source parameters and the subsurface parameters is inevitable as shown by Pavlis & Booker (1980) and Spencer & Gubbins (1980). Recently, in accordance with the developments around full-waveform inversion (Tarantola 1984; Virieux & Operto 2009; Fichtner 2010) and the increasing computational capabilities, the hypocentre–velocity problem has been recast as a full-waveform source-focusing problem (Kamei & Lumley 2014; Song *et al.* 2019; Aghamiry *et al.* 2020).

Before introducing our framework, we could imagine a scenario that draws an analogy between seismological arrays and dense seismic acquisitions like towed-streamer acquisition. A scenario where both source and receiver arrays are available at the Earth's surface with a suitable geometry that allows for the extraction of both the initiation and arrival directional attributes of single-scattered phases beyond locally plane wave approximated transmission arrivals. As an example, double beam imaging (DBI; Scherbaum *et al.* 1997) is closely related to seismic tomography methods like controlled directional reception (Riabinkin 1957) and stereotomography (Lambaré 2008). The latter being a reflection tomography method where the slope (horizontal component of the slowness vector) is locally measured at every part of the array on the source and receiver sides in order to constrain the scattering or reflection points during the velocity model building. Of course, DBI and slope tomography have completely different purposes since DBI is used as a relocation imaging technique, while slope tomography is a velocity model building technique where the inverted scattering positions serve as a proxy to attain the sought velocity structures. The similarity lies in the fact that both utilize the local coherency of neighbouring recordings at the surface in order to constrain the scattering positions in depth. Indeed, rays honouring the slope at the source and the receiver respectively would intersect at the scattering location if the velocity model is accurate, while in DBI the energy of the beams is maximized at the intersect around the scattering locations. Finding the intersection of the traveltime isochrone and the ray honouring the slope at the receiver is in fact a kinematic migration, which is utilized for example in the context of migration-based velocity analysis (Chauris *et al.* 2002) and parsimonious slope tomography (Sambolian *et al.* 2019; Fig. 1a). In the latter methods, the scattering location is not an optimization variable (as the subsurface parameters) but a state variable which is projected in the velocity estimation problem through a set of focusing (state) equations, this variable elimination being generally referred to as a variable projection (Golub & Pereyra 2003). This reduction of the model space also shrinks the data space as satisfying the focusing equations amounts to match a subset of observables from the current subsurface model. Put simply, this amounts to make the scattering positions kinematically (or physically) consistent with the available subsurface model.

We show subsequently how the recipe of parsimonious slope tomography (Sambolian *et al.* 2019), a variant of slope tomography that tackles the velocity–position coupling efficiently, could be recast as a hypocentre–velocity method. Indeed, the hypocentre–velocity coupling draws a clear parallel with the ill-famed velocity–position or velocity–depth ambiguity faced in seismic reflection tomography (Stork & Clayton 1986). We refer the reader to Sambolian *et al.* (2019) in order to get further insights on why creating a physical consistency between scattering location and the velocity parameter is advantageous compared to a joint inversion strategy. We note that wave-front tomography (Duveneck 2004; Bauer *et al.* 2017) has been recast into a joint velocity and source location method (Schwarz *et al.* 2016; Diekmann *et al.* 2019). The difference between the latter and our proposed approach goes beyond the comparison between slope tomography and wave-front tomography (Dummong *et al.* 2008) or the framework chosen for solving the forward problem and the gradient calculation. Indeed, since the key difference and main focus of this paper being tackling the velocity–position coupling through a variable projection method.

The method presented in this study is based on the idea that the location problem can be straightforwardly solved for each event in a given model by a one-to-one mapping of two kinematic attributes (traveltime and slope at the station) to the coordinates of the event. For each station, the position found by migration is kinematically consistent with the given model and we seek to collapse all the mapped locations at one position by improving the accuracy of the velocity model. As a result, the velocity model estimation serves only as a proxy to collapse the positions migrated from each station at the true source position. We develop our framework using eikonal solvers as a forward solver (Fomel *et al.* 2009; Tavakoli F. *et al.* 2015) and the adjoint-state method for the gradient computation (Plessix 2006). The presented approach handles tilted transverse isotropy and can be easily extended to three dimensions by incorporating the crossline slope or the backazimuth as an extra attribute. We note that, as any ray-based approach employing picked attributes, our approach is sensitive to the picking process. The arrival time

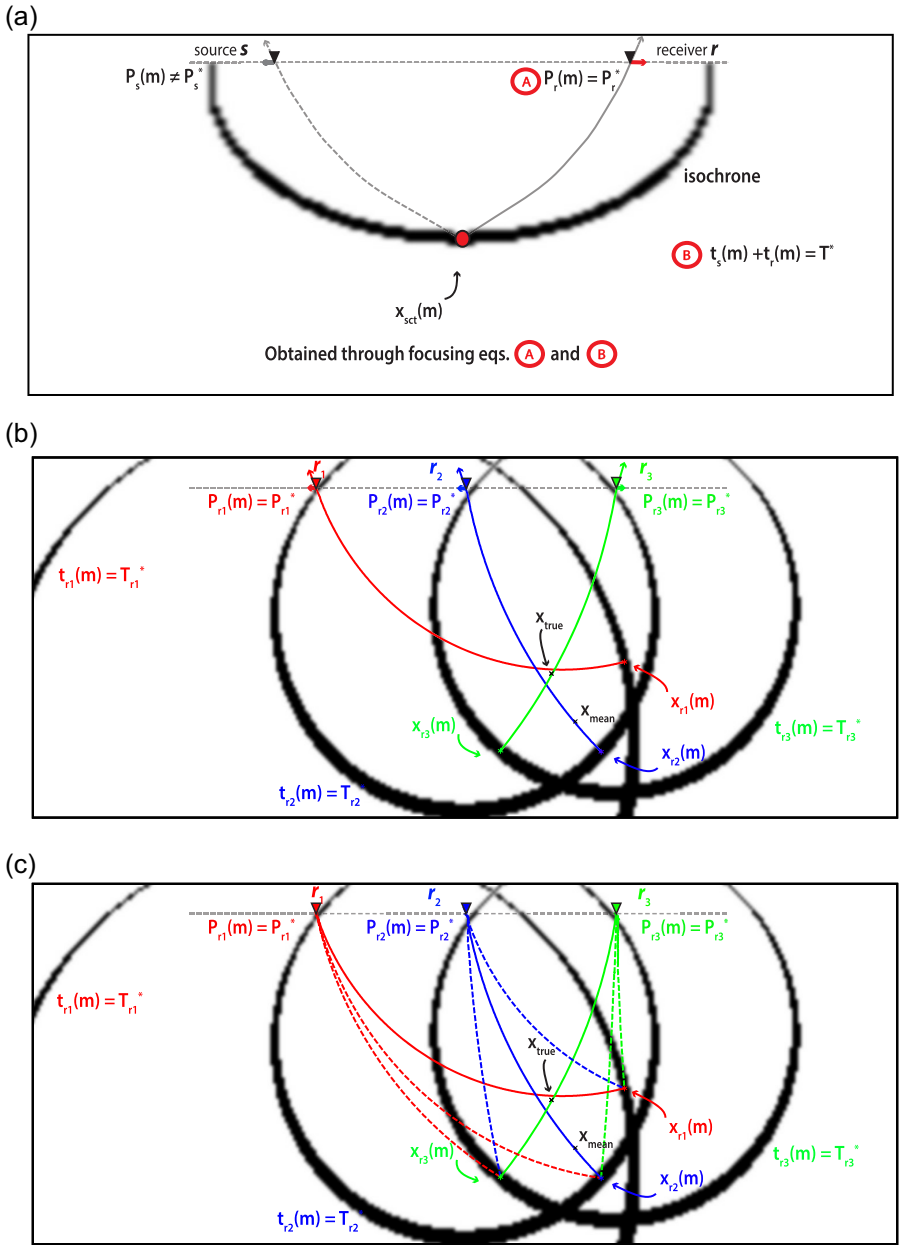

Figure 1. (a) Focusing a locally coherent associated with a reflection in the depth migrated domain through the focusing equations. The two-way traveltime and the slope at the receiver are fitted by construction while the slope at the source is used as objective measure during the inversion (Sambolian *et al.* 2019). (b) Migrating the direct arrival of an event from different receivers by fitting the traveltime and the slope at receivers. Different virtual event locations are obtained due to the inaccuracy of the velocity model. (c) Same as (b) but evaluating the data misfit at every receiver for all virtual events. The solid lines are rays describing the migration of a virtual event, while dashed lines describe rays connecting the virtual event migrated from a receiver r_i to a receiver r_j.

is straightforwardly extracted from the data while keeping in mind the discrepancy between the frequency-dependent nature of the picks and their use under infinite frequency approximation. In practice, the slope estimation is more challenging in seismological contexts, depending on the density of the receiver arrays and the validity of a local plane wave approximation. Nevertheless, the slope is more and more accessible due to the deployment of array groups, sparsity-constrained attributes inversion (Hu *et al.* 2018) and the developments around rotational seismology (Sollberger *et al.* 2018). We validate our proof of concept on a simple toy test, we assess the trade-off between the subsurface parameters and the introduced origin time correction parameter. We finally benchmark our method against the complex Marmousi model.

2 METHOD

In the following, we extend the logic behind the earlier recapped notions of kinematic migration to the hypocentre–velocity problem. First point, trivial since it is the basis of all time-reversal-based techniques but important to raise: knowing the traveltime and the slope at the

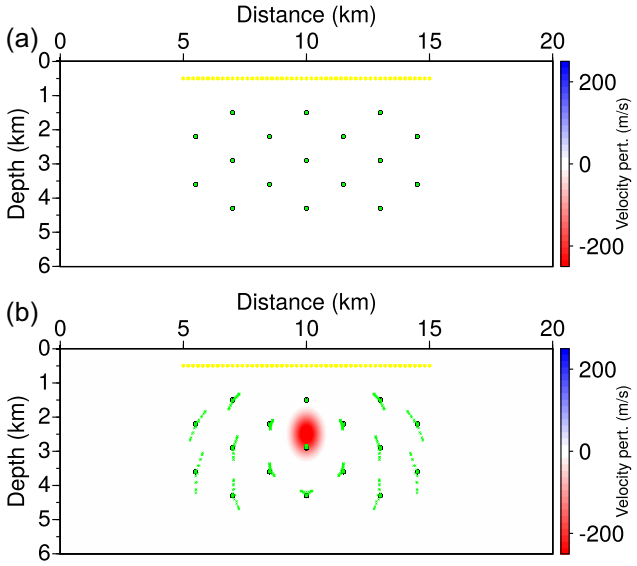

Figure 2. Virtual locations (green cross) of 17 events (black circle) corresponding to all receivers (yellow asterisk). Kinematically migrated picks in the case of the true velocity model (a). Note the focusing of events migrated by all receivers to the true location (a) and the spread caused by the velocity anomaly (b).

receiver for an event (2-D case), we are able to shoot back a ray in a subsurface model using the slope as an initial condition (take-off angle) and the traveltime as its boundary condition (stoppage time; Fig. 1b). If the origin time used in the traveltime estimation is precise and the velocity model accurately represents the subsurface, the ray would stop exactly at the location of the source. Contrarily, in case of an inaccurate velocity model, the ray would reach a different position shifted from the true position depending on the magnitude of the error in the velocity model (Fig. 1b). We emphasize on the notion that the location problem, solved using a receiver array, for a given velocity model is a strictly overdetermined problem. Any attempt at finding a unique solution in a wrong velocity model would lead at best to some best-fit solution, which is of course wrong and shifted from the sought solution. Furthermore, the location found does not honour the attributes migrated in the velocity model, it is hence physically inconsistent. Solving the hypocentre–velocity problem using an approximation or an initialized version of the location aggravates the ill-posedness of the problem, creating a trade-off pattern between both parameters that will impact the whole course of a local optimization scheme especially in the absence of a good approximate Hessian.

Extending the logic of the first point by looking at the recordings of the same event by different stations/receivers, we can conclude that if the subsurface parameters are accurate, the rays shot from all receivers (with their corresponding travel time and slope) would all stop at the same point which is the exact location. The latter notion is illustrated in Fig. 2(a) through a toy test case study, which will be presented later in this paper. We note that, even if the origin time is not accurate, the rays would intersect at the true location of the source without enforcing the stoppage time. On the other hand, in case of an inaccurate velocity model (Fig. 2b), the rays would not end at the same coordinate as they should. In fact, we have at hand as many location solutions (virtual events) as the number of stations/receivers used (Fig. 1b).

The nature of the problem revolving around the focusing of all virtual events makes it easy to constrain through that information. Indeed, it is straightforward to develop a framework where the subsurface parameters are updated to collapse the virtual events to a unique solution. Of course, we remind the reader that the explicit optimization parameters, being the subsurface parameters and the origin time correction, serve merely as a proxy to find the true event location.

We present, in the Appendix, a literal implementation of the aforementioned notion of collapsing all virtual events to a unique location in a constrained time-reversal sense. Literal in the sense that the inversion will seek better parameter estimates by reducing the distance between the virtual events in space, which are solution of the localization problem but serve also as objective measure in the optimization (Fig. 1b). The core of the paper revolves around a more tomographic implementation where the virtual events are used as anchor points to evaluate the data misfit at all other receivers (Fig. 1c) compared to the earlier strategy where the subsurface parameters are only constrained along the transmission paths connecting a receiver to its associated virtual event (Fig. 1b). Indeed, coupling the measurements at the receivers through the different combinations of virtual events and receivers (Fig. 1c) exploits better the redundancy in the data, hence in turn producing an enriched inversion kernel by linking the receiver to virtual events located by other receivers.

2.1 The consistent tomographic framework for tackling the hypocentre–velocity problem

Before proceeding with the development of our formulation, we remind the reader that the traveltime and slope at a receiver are fitted by construction to locate its associated virtual event via the focusing equations but are not fitted when considering the paths connecting this receiver to the virtual events migrated by other receivers (since they are not at the same position due to the inaccuracy of the velocity model).

We seek to optimize the fit of the traveltimes and slopes at all receivers for each virtual event (Fig. 1c). As a result, we define the following nonlinear constrained minimization problem with the aim of retrieving the minimizer $\mathbf{m}$, gathering the subsurface parameters and an origin time correction parameter,

$$\min_{\mathbf{m}} C(\mathbf{m}) = \min_{\mathbf{m}} \frac{1}{2\sigma^2_{T_{e,r,r'}}} \sum_{e=1}^{N_e} \sum_{r=1}^{N_r} \sum_{r'=1, r' \neq r}^{N_r} (T_{e,r,r'}(\boldsymbol{m}) - T^*_{e,r'})^2 + \frac{1}{2\sigma^2_{p_{e,r,r'}}} \sum_{e=1}^{N_e} \sum_{r=1}^{N_r} \sum_{r'=1, r' \neq r}^{N_r} (p_{e,r,r'}(\boldsymbol{m}) - p^*_{e,r'})^2 \quad \text{subject to} \quad \mathbf{F}(\mathbf{u}, \mathbf{m}) = 0, \tag{1}$$

where N_e and N_r denote the number of events and receivers, respectively. The observed data are the traveltime $T^*_{e,r'}$ and the slope $p^*_{e,r'}$ measured at receivers r', namely receivers other than the one denoted by the subscript r. Accordingly, the simulated data, denoted by $T_{e,r,r'}$ and $p_{e,r,r'}$, are the traveltime and the slope at receiver r' evaluated from the virtual event located by kinematic migration of $T^*_{e,r}$ and $p^*_{e,r}$ of receiver r (see the dash lines in Fig. 1c for the corresponding ray paths). Coefficients $\sigma^2_{T_{e,r,r'}}$ and $\sigma^2_{p_{e,r,r'}}$ serve as weighting quantities to make the data space dimensionless. The latter can serve also as the inverse of a diagonal covariance matrix in order to weight the relative contribution of every measurement (Tarantola 1987). The nonlinear forward problem operator $\mathbf{F}$ gathers the forward problem equations related to the data simulation through eikonal-resolved traveltime maps and the focusing equations (Chauris *et al.* 2002). We solve the constrained problem (eq. 1) under a Lagrangian formalism following the adjoint-state method recipe (Haber *et al.* 2000; Plessix 2006). The augmented functional $\mathcal{L}$ in compact form is rewritten as

$$\mathcal{L}(\mathbf{m}, \mathbf{u}, \bar{\mathbf{u}}) = J(\mathbf{u}) - \left\langle \bar{\mathbf{u}} \mid \mathbf{F}(\mathbf{u}, \mathbf{m}) \right\rangle, \tag{2}$$

where $\langle . | . \rangle$ denotes the inner product, $\mathbf{u}$ gathers the state variables, $\bar{\mathbf{u}}$ gathers the adjoint-state variables (or Lagrange multipliers) and $C(\mathbf{m}) = J(\mathbf{u}^*)$, where $\mathbf{u}^*$ stands for a realization of the physical constraints. We proceed with the description of the physical (state) equations gathered by the nonlinear forward problem operator $\mathbf{F}$. We infer the predicted traveltimes and slopes from traveltime maps $t_r(\mathbf{x})$ computed with a finite-difference factored eikonal solver using the receiver positions as injection points (Fomel *et al.* 2009; Tavakoli F. *et al.* 2015). A Dirichlet boundary condition is introduced to zero the traveltime at the receiver positions:

$$H(\mathbf{x}, \nabla t_r(\mathbf{x})) = 0 \quad \text{with} \quad t_r(\mathbf{x}_r) = 0. \tag{3}$$

The operator H stands for the Hamiltonian representation of the eikonal equation in tilted transversely isotropic (TTI) media (Alkhalifah 1998; Waheed *et al.* 2014) given by

$$H(\mathbf{x}, \nabla t(\mathbf{x})) = A(\mathbf{x})((\mathcal{R}\nabla \mathbf{t}(\mathbf{x}))_x)^2 + C(\mathbf{x})((\mathcal{R}\nabla \mathbf{t}(\mathbf{x}))_z)^2 + E(\mathbf{x})((\mathcal{R}\nabla \mathbf{t}(\mathbf{x}))_x (\mathcal{R}\nabla \mathbf{t}(\mathbf{x}))_z)^2 - 1, \tag{4}$$

where $\mathcal{R}$ is a standard rotation matrix and A, C, E are coefficients that embed the model parameters we seek to update depending on the chosen anisotropic parametrization (Alkhalifah & Tsvankin 1995; Plessix & Cao 2011; Gholami *et al.* 2013). We refer the reader to Tavakoli F. *et al.* (2015) and Waheed *et al.* (2015) for a detailed description on the manner of solving eq. (4) in TTI media using the fast sweeping method (Zhao 2005; Luo & Qian 2012) as a global solver and a fixed-point iteration algorithm (Kelley 1995) for handling the quartic term.

In order to extract the traveltime solution at the position $\mathbf{x}_{e,r}$ of an event e from the traveltime map $t_r(\mathbf{x})$ initiated at the receiver r, we introduce a sampling operator $Q_{e,r}$ implemented with a Kaiser-windowed sinc function (Hicks 2002):

$$T_{e,r} = t_r(\mathbf{x}_{e,r}) = Q_{e,r} t_r(\mathbf{x}), \tag{5}$$

while the slope at the receiver r for the event e is obtained in a finite-difference sense

$$p_{e,r} = \frac{\partial T_{e,r}}{\partial x_r} = \frac{\partial t_r(\mathbf{x}_{e,r})}{\partial x_r} \approx \frac{Q_{e,r}(t_{r+1}(\mathbf{x}) - t_{r-1}(\mathbf{x}))}{2\Delta_r}. \tag{6}$$

The computational complexity of the problem scales with $\mathcal{O}(N_r)$ since reciprocity is employed in order to alleviate the computational cost opposed to solving the eikonal equation from each virtual event with a complexity proportional to $\mathcal{O}(N_e \times N_r)$. We note that more precise strategies for the computation of the slopes exist but would involve solving an additional eikonal-based partial differential equation (Qian & Symes 2002; Alkhalifah & Fomel 2010). As mentioned earlier, we solve a kinematic migration through the so-called focusing equations (Chauris *et al.* 2002) as follows:

$$T_{e,r} = T^*_{e,r} \qquad p_{e,r} = p^*_{e,r}, \tag{7}$$

the symbol $*$ denoting the observed data. Enforcing the pair of eq. (7) gives the position of each virtual event $\mathbf{x}_{e,r}$. An origin time correction is mandatory since the estimated origin time contains the bias of the inexact velocity model and in turn contaminates the estimated travel time. In order to correct for the latter error, we introduce the correction term δ_e for every event. We do not search for the origin time but we rather estimate a correction parameter for the error made in the origin time estimation (used as an initial estimate) due to the inaccuracy of the velocity model or some other reason. We note that we could have introduced a different correction for every measurement or in other words every event-receiver pair. The strategy of splitting δ_e into a $\delta_{e,r}$ could be used to absorb some picking and delay errors but would also create an artificial coupling between the locations of the same event since they are not all sensitive in the same manner to the origin time

(depending on location and propagation time). We stick to the first strategy where a correction term is introduced solely for every event and enforced on all picks related to that event. During this study, we suppose that all picked phases recorded at different stations are associated with a unique point-source event. The aforementioned splitting strategy would be beneficial for more complex cases where aftershocks are introduced in the presented framework, especially with the rising interest in unsupervised phase identification and classification algorithms (Bauer *et al.* 2019; Ross *et al.* 2019). The correction parameter δ_e is introduced in the focusing equation related to traveltime as follows:

$$T_{e,r} = t_r(\mathbf{x}_{e,r}) = Q_{e,r}t_r(\mathbf{x}) + \frac{1}{N_r}\delta_e, \tag{8}$$

where δ_e is normalized by N_r to account for the fact that this term is in reality the sum of the same error over all receivers. Enforcing the focusing equations, eq. (7), leads to the position of each virtual event $\mathbf{x}_{e,r}$. As introduced earlier in the section, we evaluate the misfit between the observed data recorded at all receivers and the simulated measurements $T_{e,r,r'}$ and $p_{e,r,r'}$ extracted from the traveltimes maps initiated at the receivers r'. As a result, eqs (8) and (6) lead to

$$T_{e,r,r'} = Q_{e,r}t_{r'}(\mathbf{x}) + \frac{1}{N_r}\delta_e, \tag{9}$$

and

$$p_{e,r,r'} = \frac{\partial t'_r(\mathbf{x}_{e,r})}{\partial x_r r} \approx \frac{Q_{e,r}(t_{r'+1}(\mathbf{x}) - t_{r'-1}(\mathbf{x}))}{2\Delta'_r}. \tag{10}$$

Before proceeding, we note that there is no need to introduce explicitly the equations related to $t_{r'}(\mathbf{x})$ since they are already described by the redundant solutions of $t_r(\mathbf{x})$.

Recasting the minimization problem (eq. 1) under a Lagrangian formalism using the state eqs (3), (7), (9) and (10) and associating the adjoint-state variables $\bar{\mathbf{u}} = (\lambda_r, \alpha_{e,r}, \beta_{e,r}, \xi_{e,r,r'}, \gamma_{e,r,r'})$ to their respective state variables $\mathbf{u} = (t_r(\mathbf{x}), x_{e,r}, z_{e,r}, T_{e,r,r'}, p_{e,r,r'})$ gives

$$\begin{aligned}\mathcal{L}(\mathbf{m}, \mathbf{u}, \bar{\mathbf{u}}) = J(\mathbf{u}) &- \sum_{e=1}^{N_e}\sum_{r=1}^{N_r}\alpha_{e,r}\left(T^*_{e,r} - Q_{e,r}t_r(\mathbf{x}) - \frac{1}{N_r}\delta_e\right) - \sum_{e=1}^{N_e}\sum_{r=1}^{N_r}\beta_{e,r}\left(p^*_{e,r} - \frac{Q_{e,r}(t_{r+1}(\mathbf{x}) - t_{r-1}(\mathbf{x}))}{2\Delta_r}\right)\\
&- \sum_{e=1}^{N_e}\sum_{r=1}^{N_r}\sum_{r'=1,r'\neq r}^{N_r}\gamma_{e,r,r'}\left(T_{e,r,r'} - Q_{e,r}t_{r'}(\mathbf{x}) - \frac{1}{N_r}\delta_e\right) - \sum_{e=1}^{N_e}\sum_{r=1}^{N_r}\sum_{r'=1,r'\neq r}^{N_r}\xi_{e,r,r'}\left(p_{e,r,r'} - \frac{Q_{e,r}(t_{r'+1}(\mathbf{x}) - t_{r'-1}(\mathbf{x}))}{2\Delta_r}\right)\\
&- \frac{1}{2}\sum_{r=1}^{N_r}\Big\langle\lambda_r(\mathbf{x}) \mid H(\mathbf{x}, \nabla t_r(\mathbf{x}))\Big\rangle_\Omega,\end{aligned} \tag{11}$$

where the Lagrangian functional $\mathcal{L}$ depends on the subsurface parameters $\mathbf{m}$ through the eikonal equation solved in the subsurface domain Ω. According to the first-order optimality conditions, namely, the so-called Karush–Kuhn–Tucker (KKT) conditions, a minimizer of an equality constrained optimization problem is reached at the saddle point of the Lagrangian function (Nocedal & Wright 2006) when the three following equations are satisfied:

$$\begin{cases}\partial\mathcal{L}/\partial\mathbf{u} = 0,\\ \partial\mathcal{L}/\partial\bar{\mathbf{u}} = 0,\\ \partial\mathcal{L}/\partial\mathbf{m} = 0.\end{cases} \tag{12}$$

The joint update of the entire system spanned by $\mathbf{u}$, $\bar{\mathbf{u}}$ and $\mathbf{m}$ is avoided due to computational complexity (Akçelik 2002). We thus resort to the reduced-space approach of the adjoint-state method (Haber *et al.* 2000; Plessix 2006) based on a sequence of variable projections.

In other words, the first two KKT conditions of eq. (12) are satisfied by solving the state equations $\partial\mathcal{L}/\partial\bar{\mathbf{u}} = 0$ in the starting model $\mathbf{m}_k$ of iteration k and we then subsequently deduce the Lagrange multipliers by enforcing $\partial\mathcal{L}/\partial\mathbf{u} = 0$ in this manner.

Following the aforementioned recipe, we develop now the adjoint-state equations. We proceed by solving for the first two adjoint-state equations through $\partial\mathcal{L}/\partial_{e,r,r'} = 0$ and $\partial\mathcal{L}/\partial p_{e,r,r'} = 0$ which shows that $\gamma_{e,r,r'}$ and $\xi_{e,r,r'}$ gather the scaled data residuals for every combination of receiver and focused event as follows:

$$\gamma_{e,r,r'} = \frac{1}{\sigma^2_{T_{e,r,r'}}}(T_{e,r,r'} - {T_{e,r'}}^*) = \frac{\Delta T_{e,r,r'}}{\sigma^2_{T_{e,r,r'}}} \tag{13}$$

and

$$\xi_{e,r,r'} = \frac{1}{\sigma^2_{p_{e,r,r'}}}(p_{e,r,r'} - {p_{e,r'}}^*) = \frac{\Delta p_{e,r,r'}}{\sigma^2_{p_{e,r,r'}}}. \tag{14}$$

For each event, $\partial\mathcal{L}/\partial\mathbf{x}_{e,r} = 0$ gives the following 2×2 system of linear equations for $\alpha_{e,r}$, $\beta_{e,r}$:

$$\alpha_{e,r}\frac{\partial Q_{e,r}}{\partial\mathbf{x}_{e,r}}\mathbf{t}_r + \frac{\beta_{e,r}}{2\Delta r}\frac{\partial Q_{e,r}}{\partial\mathbf{x}_{e,r}}(\mathbf{t}_{r+1} - \mathbf{t}_{r-1}) = R_{e,r}, \tag{15}$$

where the right-hand side $R_{e,r} = -\sum_{r'=1,r'\neq r}^{N_r}\gamma_{e,r,r'}\frac{\partial Q_{e,r}}{\partial\mathbf{x}_{e,r}}\mathbf{t}_{r'} - \sum_{r'=1,r'\neq r}^{N_r}\frac{\xi_{e,r,r'}}{2\Delta r'}\frac{\partial Q_{e,r}}{\partial\mathbf{x}_{e,r}}(\mathbf{t}_{r'+1} - \mathbf{t}_{r'-1})$ gathers the terms depending on the already solved $\gamma_{e,r,r'}$ and $\xi_{e,r,r'}$. Solving the latter system through Cramer's rule leads to the closed form expression of $\alpha_{e,r}$ and $\beta_{e,r}$ as follows:

$$\alpha_{e,r} = \frac{\det\left|\frac{\partial(R_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})}\right|}{\det\left|\frac{\partial(T_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})}\right|} = \frac{\partial R_{e,r}}{\partial T_{e,r}}, \tag{16}$$

$$\beta_{e,r} = \frac{\det\left|\frac{\partial(T_{e,r},\ R_{e,r})}{\partial(x_{e,r},\ z_{e,r})}\right|}{\det\left|\frac{\partial(T_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})}\right|} = \frac{\partial R_{e,r}}{\partial p_{e,r}}. \tag{17}$$

Injecting expressions of $\alpha_{e,r}$ and $\beta_{e,r}$ in eq. (15) gives

$$\frac{\partial R_{e,r}}{\partial T_{e,r}}\frac{\partial T_{e,r}}{\partial \mathbf{x}_{e,r}} + \frac{\partial R_{e,r}}{\partial p_{e,r}}\frac{\partial p_{e,r}}{\partial \mathbf{x}_{e,r}} = \frac{\partial R_{e,r}}{\partial \mathbf{x}_{e,r}}, \tag{18}$$

where

$$\frac{\partial R_{e,r}}{\partial T_{e,r}} = \frac{\det\left|\frac{\partial(R_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})}\right|}{\det\left|\frac{\partial(T_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})}\right|} \quad \frac{\partial R_{e,r}}{\partial p_{e,r}} = \frac{\det\left|\frac{\partial(T_{e,r},\ R_{e,r})}{\partial(x_{e,r},\ z_{e,r})}\right|}{\det\left|\frac{\partial(T_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})}\right|}. \tag{19}$$

Knowing that the right-hand sides of the adjoint-state equations contain the partial derivative of $J(R_{e,r})$ with respect to the states $\mathbf{u}$ (Plessix 2006), the chain rule of derivatives (eq. 18) illustrates mathematically how, in the left-hand side of eq. (18), the information carried out by $p_{e,r}$ and $T_{e,r}$ on the positioning of an event via the focusing equations (eq. 7) is passed onto the optimization measure $R_{e,r}$ via the state variables $\mathbf{x}_{e,r}$ (right-hand side of eq. 18). In simple terms, the resultant adjoint-state variables $\alpha_{e,r}$ and $\beta_{e,r}$ describe how $R_{e,r}$ evolves when the coordinates $\mathbf{x}_{e,r}$ of an event are altered by a velocity update.

The information held by $\alpha_{e,r}$ and $\beta_{e,r}$ links the positioning process initiated at the receiver r done through the enforced focusing equations to the data misfit evaluated at other receivers. In other words, the latter terms contain the quantity needed to minimize the residuals held by $\gamma_{e,r,r'}$ and $\xi_{e,r,r'}$ (gathered under the variable $R_{e,r}$) by shifting the position of the event $\mathbf{x}_{e,r}$.

We note that eqs (16) and (17) are only defined for $\det\left|\frac{\partial(T_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})}\right| \neq 0$. The system is always valid since in our case this condition could never be violated. Zeroing the latter term describes physically a tangent ray to its wave front, while in fact that scenario is impossible to occur in weak anisotropy assumptions.

Proceeding with the last derivative $\partial\mathcal{L}/\partial\mathbf{t}_r = 0$ in the same manner as Tavakoli F. *et al.* (2017) in the context of slope tomography, we obtain the adjoint-state equation satisfied by $\lambda_r(\mathbf{x})$. After integrating by parts and enforcing the validity of $\mathcal{L}$ in the subsurface domain Ω, the derivation leads to

$$\begin{aligned}\left(\nabla\cdot(\lambda_r(\boldsymbol{x})\,\mathcal{U}_r)\right)_\Omega = &\sum_{e=1}^{N_e}\sum_{r'=1,r'\neq r}^{N_r}\left(Q^t_{e,r'}\gamma_{e,r',r} - \frac{1}{2\Delta r}\,Q^t_{e,r'+1}\xi_{e,r'+1,r} + \frac{1}{2\Delta r}\,Q^t_{e,r'-1}\xi_{e,r'-1,r}\right)\\ &+\sum_{e=1}^{N_e}\left(Q^t_{e,r}\alpha_{e,r} - \frac{1}{2\Delta r}\,Q^t_{e,r+1}\beta_{e,r+1} + \frac{1}{2\Delta r}\,Q^t_{e,r-1}\beta_{e,r-1}\right).\end{aligned} \tag{20}$$

The adjoint field $\lambda_r(\mathbf{x})$ backprojects the weighted sum of data residuals held by $\gamma_{e,r',r}$ and $\xi_{e,r',r}$ along a ray tube following the group velocity vector $\mathcal{U}_r$ connecting $\mathbf{x}_{e,r'}$ to $\mathbf{x}_r$ (Fig. 3a). In addition, the adjoint field of the receiver r backprojects the weighted data residuals held by $\alpha_{e,r}$ and $\beta_{e,r}$ along a ray tube following the group velocity vector $\mathcal{U}_r$ connecting $\mathbf{x}_{e,r}$ to $\mathbf{x}_r$ (Fig. 3b). The latter information describes how the receiver r controls the migration of its associated virtual events. As the eikonal equation, the adjoint-state eq. (20) is solved with the fast sweeping method (Zhao 2005; Taillandier *et al.* 2009) using a conservative finite-difference scheme as described by Tavakoli F. *et al.* (2019).

We caution the reader that a switch between r and r' occurred in eq. (20). The latter describes the adjoint field $\lambda_r(\mathbf{x})$ of a receiver r which is in turn of course the r' with respect to the other receivers.

From the adjoint-state variables, the gradient of the augmented functional (eq. 11) with respect to the subsurface parameters is straightforwardly obtained by the weighted summation of the adjoint fields λ_r :

$$\nabla_{\mathbf{m_{sp}(x)}}J = -\frac{1}{2}\sum_{r=1}^{N_r}\frac{\partial H(\mathbf{x},\nabla\mathbf{t}_r(\mathbf{x}))}{\partial\mathbf{m_{sp}(x)}}\lambda_r(\mathbf{x}). \tag{21}$$

The adjoint field is weighted by the derivative of the forward operator $H(\mathbf{x},\nabla t(\mathbf{x}))$ with respect to any subsurface model parameter, the gradient of J evolves accordingly. We refer the reader to Tavakoli F. *et al.* (2019, appendix B) for a detailed derivation with respect to every parameter in TTI media. The gradient for the case presented in Fig. 2 is shown in Fig. 4 and illustrates the focusing of the velocity inclusion with some smearing inherited from the footprint of the sensitivity kernels connecting the events to the stations. The only equation left to develop is the gradient of the objective function (eq. 1) with respect to the origin time correction parameter and is written as follows for an event:

$$\nabla_{\mathbf{m}_{\delta_\mathbf{e}}}J = \frac{1}{N_r}\left(\sum_{r=1}^{N_r}\alpha_{e,r} + \sum_{r=1}^{N_r}\sum_{r'=1,r'\neq r}^{N_r}\gamma_{e,r,r'}\right). \tag{22}$$

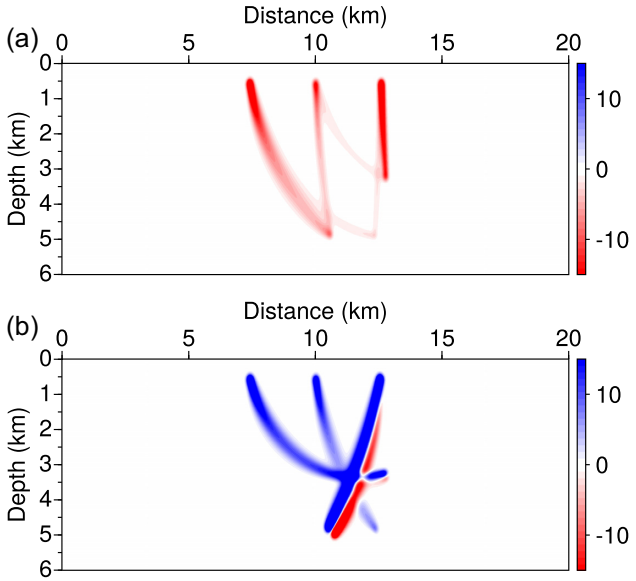

Figure 3. Triple receiver kernel $\lambda_r(\mathbf{x})$ (eq. 20) solved for the case presented in Fig. 1(b). (a) The isolated contributions of $\gamma_{e,r',r}$ and $\xi_{e,r',r}$ backpropagated from the virtual events migrated by receivers r towards receivers r'. (b) The isolated contributions of $\alpha_{e,r}$ and $\beta_{e,r}$ backpropagated from the virtual events migrated by receivers r towards themselves.

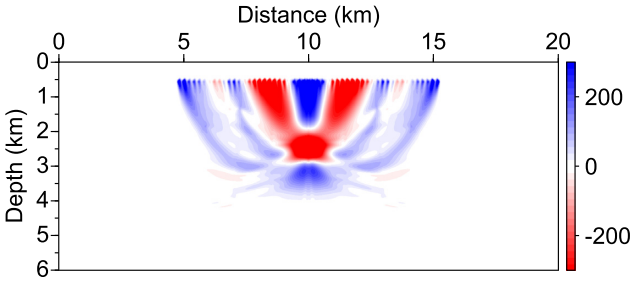

Figure 4. Gradient of $J(\mathbf{m})$ (eq. 11) with respect to velocity for the case presented in Fig. 2.

Once the gradient is computed, we proceed with a Newton-based local optimization scheme:

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha_k \left(\frac{\partial^2 C(\mathbf{m}_k)}{\partial \mathbf{m}^2} \right)^{-1} \left(\frac{\partial C(\mathbf{m}_k)}{\partial \mathbf{m}} \right), \tag{23}$$

where the step length $\alpha_k \in \mathbb{R}_+$ satisfies the Armijo rule and the curvature condition of the Wolfe conditions (Nocedal & Wright 2006). In practice, for all numerical experiments presented in this study, the inexact line search is managed by the SEISCOPE optimization toolbox (Métivier & Brossier 2016) and the inverse Hessian is approximated through a limited-memory Broyden–Fletcher–Goldfarb–Shannon (L-BFGS) algorithm (Byrd *et al.* 1995).

2.2 Validation with a toy numerical example

We validate step by step our method on a toy test of a constant gradient velocity model defined by $v = v_0 + a \times z$, where $v_0 = 1000\,\mathrm{m\,s^{-1}}$ and $a = 1$, with an added Gaussian ball perturbation (centred at $x = 10$ km, $z = 2.5$ km; Fig. 5). The objective of the exercise being the recovery of the Gaussian ball starting with the gradient model as initial guess while finding the correct source locations for 17 events using a dense surface array of 51 stations/receivers. We remind the reader that, since the velocity model is inaccurate, the measurements from different stations/receivers would point at different locations.

2.2.1 The hypocentre–velocity reconstruction

We look first at the case where the exact origin correction times ($\delta t_e = 0.1$ s for all events) are used during the inversion. We note that only for the sake of validation of our formulation we choose this unrealistic experimental setup. In Fig. 5(a), we can see the spread of the virtual locations (green cross) around the exact location (black circle) for the 17 events. We note that the pattern of the spread is solely defined by the inaccuracy of the velocity model since the exact origin corrections times are used for the kinematic migration. Following 73 iterations

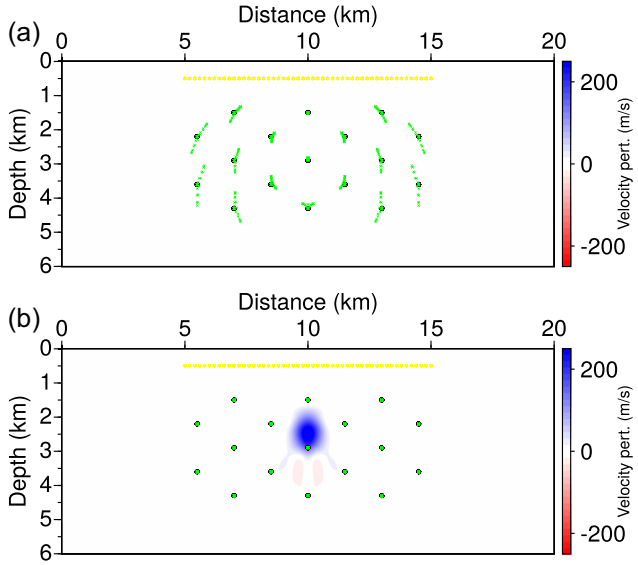

Figure 5. Toy test case: inverting for velocity with exact origin correction time as passive parameters. Virtual locations (green cross) of 17 events (black circle) corresponding to all receivers (yellow asterisk). Focused locations in the case of (a) the initial velocity model and (b) the updated velocity model after 73 iterations.

(Fig. 5b), the points collapse around the true position. We can see that some smearing and inversion artefacts occur below the retrieved ball. The latter is expected using surface acquisition since we have an asymmetric illumination of the medium. We also see that the smearing impacts slightly the focusing for the event directly below the ball. The logs presented in Fig. 6 illustrate the magnitude of the smearing around the recovered perturbation and serve as validation in terms of perturbation amplitude recovery. The results are overall satisfactory and validate our approach for this simple case where we assume that there is no origin time correction needed. We invite the reader to check the animated graphics (i) of the Supporting Information for a look on the evolution of the inverted parameters through the iterations with an insight on the misfit value. An interesting point to raise before proceeding to a more complex test case is the necessity of fitting both traveltime and slope attributes (eq. 1). We remind the reader that both attributes are crucial to perform the kinematic migration (in two dimensions) but fitting traveltimes could be enough especially in dense arrays settings. Undoubtedly, constraining the emergence angle of the trajectory receiver-event by fitting the slopes makes the problem better posed in complex settings or sparse acquisition cases. We present in the Supporting Information (animated graphics ii and iii) the inversion results for the same case as the one presented in Fig. 5 but when fitting slopes and traveltimes, respectively. The results are satisfactory for both cases but more artefacts are built around the retrieved Gaussian ball through fitting traveltimes only. The inversion stopped after 78 iterations when fitting traveltimes and 86 iterations when fitting slopes. The faster convergence in the inversion presented earlier where both attributes are used illustrates the complementary nature of the attributes. Furthermore, we note the fact that the traveltime misfit at the final iteration is worse when they are solely fitted.

2.2.2 The origin time and location–velocity problem

We look closely in this section on the importance of the origin time correction and its induced trade-off with the subsurface parameters via the kinematic migration. The first question would revolve around the impact of ignoring the origin time correction. We examine in the following the ill-posedness of the problem in the case of a wrong estimated origin time. In order to illustrate the answer of the latter question, we repeat the previous test but we introduce an error of 100 ms for all events as if an origin time correction of 100 ms is needed. The error is significant since, depending on the picks, it represents a shift of 3–16 per cent of the observed propagation time (keep in mind that we are using a slow top layer velocity and not very large event-receiver distances so the traveltimes are small). We note that we first assign the same error to all events in order to illustrate the pathology but we will then follow with a different error for each event.

Contrarily to the previous section, we would expect a spread of the locations in the initial model to be defined not only by the inaccuracy of the velocity model but also the error in origin time (Fig. 7a). The inverted model after 118 iterations contains the sought perturbation but is also contaminated by a nearly homogeneous negative velocity perturbation, which balances the overestimation of the origin time (Fig. 7b). The latter trend is produced during the inversion in order to compensate the systematic error introduced on the origin time (which is the same for all events in this case) and enforce the focusing of the virtual locations at one coordinate. This pathology points out perfectly the ill-posed nature of the problem at hand. Indeed, the origin time and the velocity parameter are coupled through the positioning process. We invite the reader to check the animated graphics (iv) of the Supporting Information for a look on the evolution of the inverted parameters through the iterations with an insight on the misfit value.

We repeat the test but we examine now the more realistic case where the origin time error is different for each event. In Fig. 8(a), we note the more complicated patterns drawn by the virtual locations due to the different errors in origin time for each event. Following 47 iterations,

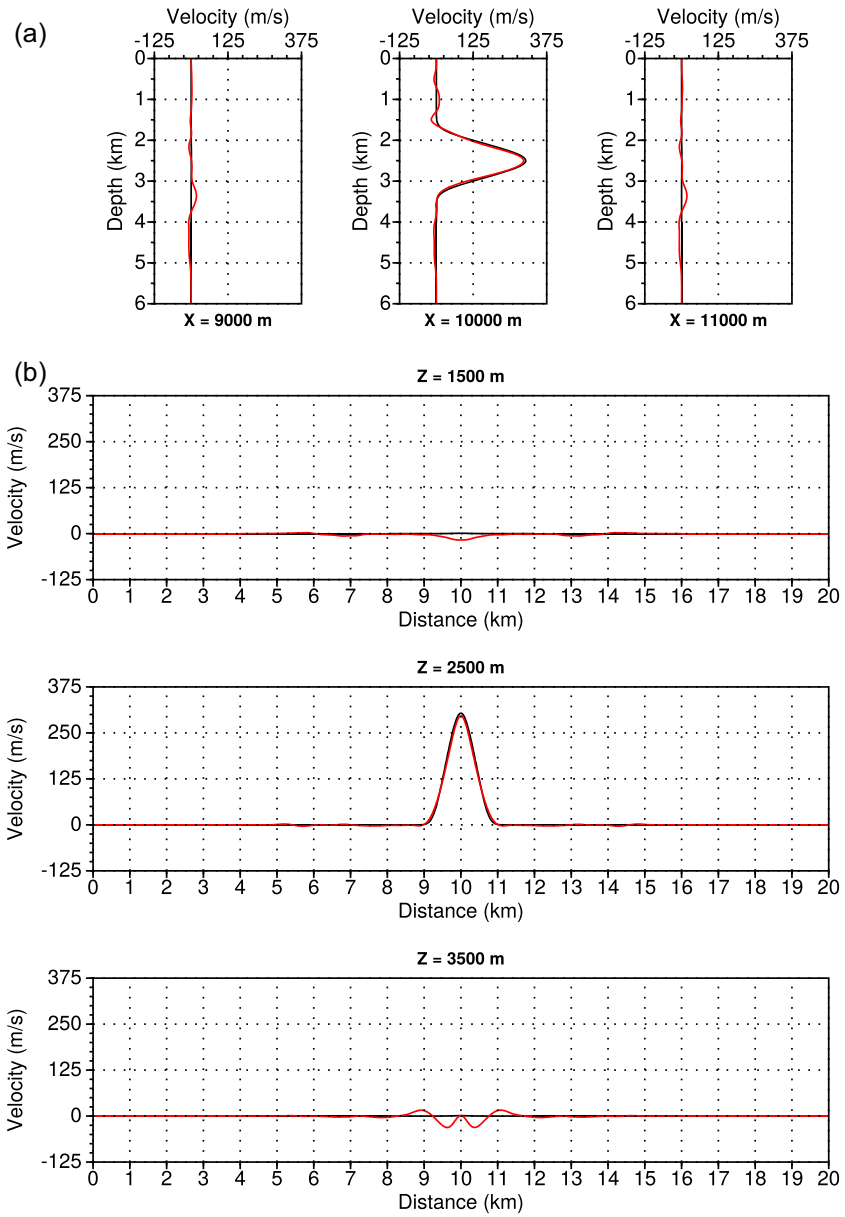

Figure 6. Toy test case: inverting for velocity with exact origin time. (a) Vertical and (b) horizontal comparative velocity perturbation logs. Black and red lines denote the exact and reconstructed perturbations, respectively. The positions of the logs are provided in the figure.

the inversion stopped due to a line search failure. The reconstructed model contains more artefacts than the previous case and in a more heterogeneous manner as one would expect since different origin times are associated with each event. The interesting point compared to the previous result is the fact that the inversion process stopped at a local minimum where the focusing is far from being achieved. The different origin times' errors cannot be compensated by a simple velocity perturbation as the previous test. We invite the reader to check the animated graphics (v) of the Supporting Information for a look on the evolution of the inverted parameters through the iterations with an insight on the misfit value and the values related to the origin time parameter.

In this section, we have highlighted the repercussions of ignoring the origin time parameter and the necessity of accounting for its correction during the inversion. Before proceeding with a multiparameter inversion test, we validate the gradient with respect to the origin time correction parameter. We proceed by doing two inversions in which the velocity is a passive parameter, while we try to invert for the origin time mismatch. In the first test, the true velocity model is used as passive model, while the background velocity gradient model will be used in the second test. We note that, for this validation, we repeat the tests where the same origin time correction is needed for all events. For the first test, we note that the pattern of the spread is solely defined by the origin time mismatch and is systematic for all events since they have the same error in the propagation time (due to the origin time mismatch) and the exact velocity model was used for the kinematic migration (Fig. 9a). Following nine iterations, the correct origin time correction is recovered for all events and the events collapse on the exact

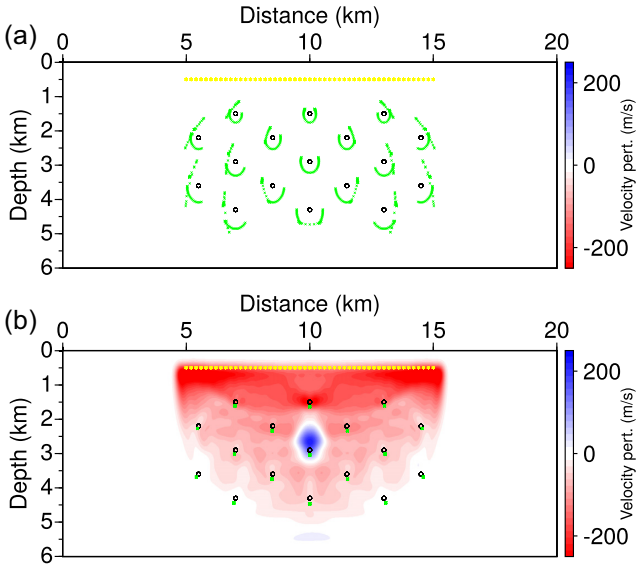

Figure 7. Toy test case: inverting for velocity with wrong origin times (uniform) as passive parameters. Virtual locations (green cross) of 17 events (black circle) corresponding to all receivers (yellow asterisk). Focused locations in the case of (a) the initial velocity model and (b) the updated velocity model after 118 iterations.

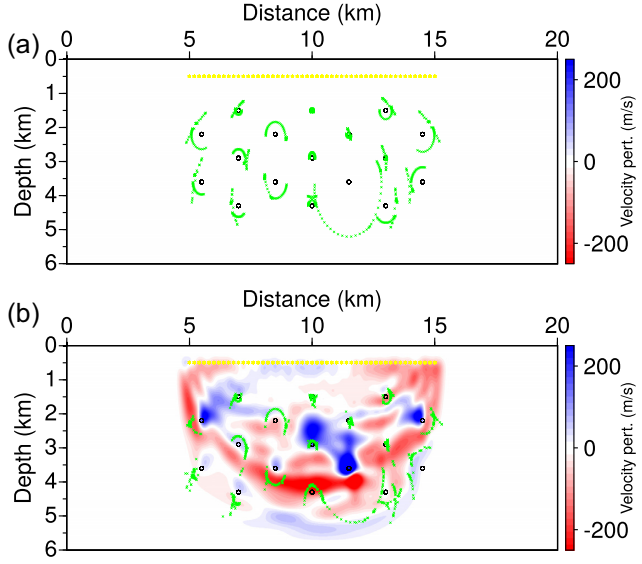

Figure 8. Toy test case: inverting for velocity with wrong origin times (different for every event) as passive parameters. Virtual locations (green cross) of 17 events (black circle) corresponding to all receivers (yellow asterisk). Focused locations in the case of (a) the initial velocity model and (b) the updated velocity model after 47 iterations (bottom).

locations (Fig. 9b). We invite the reader to check the animated graphics (vi) of the Supporting Information for a look on the evolution of the inverted origin time correction parameter through the iterations with an insight on the misfit value.

During the second test, we repeat the same inversion but using the gradient velocity model as a passive quantity during the inversion process. The final result of this inversion will serve as a direct comparison to the initial stage of the case where we inverted for the velocity model while using the exact origin time (Fig. 5a). At the starting point of the inversion (Fig. 10a), the pattern of the spread is as expected more significant compared to Fig. 5(a) since it is generated by both inaccurate origin correction times and an inaccurate velocity model. Following six iterations, the correct origin time correction is recovered for all events and their spreading pattern is identical to the one seen at the initial step of the case where we tried to invert for velocity using the true origin times (Fig. 10b versus Fig. 5a). The fact that we recover the true origin times without retrieving the true positions highlights the necessity of a multiparameter inversion due to the coupling between the origin time and the velocity. We invite the reader to check the animated graphics (vii) of the Supporting Information for a look on the evolution of the inverted origin time correction parameter through the iterations with an insight on the misfit value. We note that, at the reached local minimum, the positions could not be further optimized due to a leakage between the velocity and the origin time parameter. Indeed, we could have imagined a scenario where the final inverted origin times are wrong but the event coordinates collapse around the same position. We

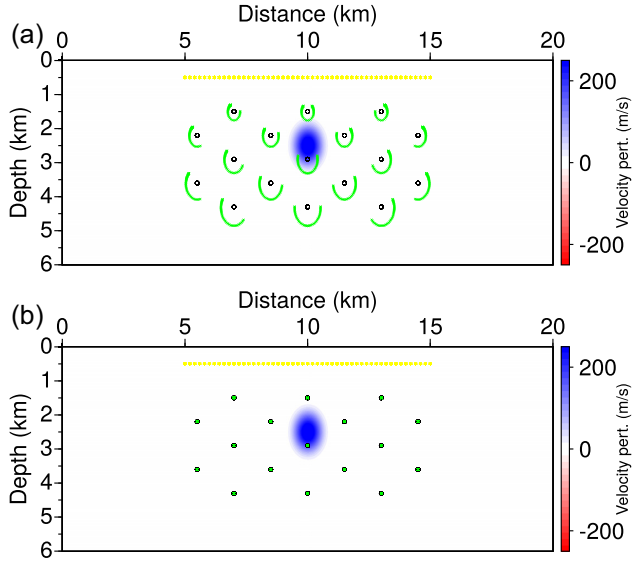

Figure 9. Toy test case: inverting for the origin time correction parameter using the true velocity model as passive parameter. Virtual locations (green cross) of 17 events (black circle) corresponding to all receivers (yellow asterisk). Focused locations using (a) the initial origin time corrections and (b) the final origin time corrections updated after nine iterations.

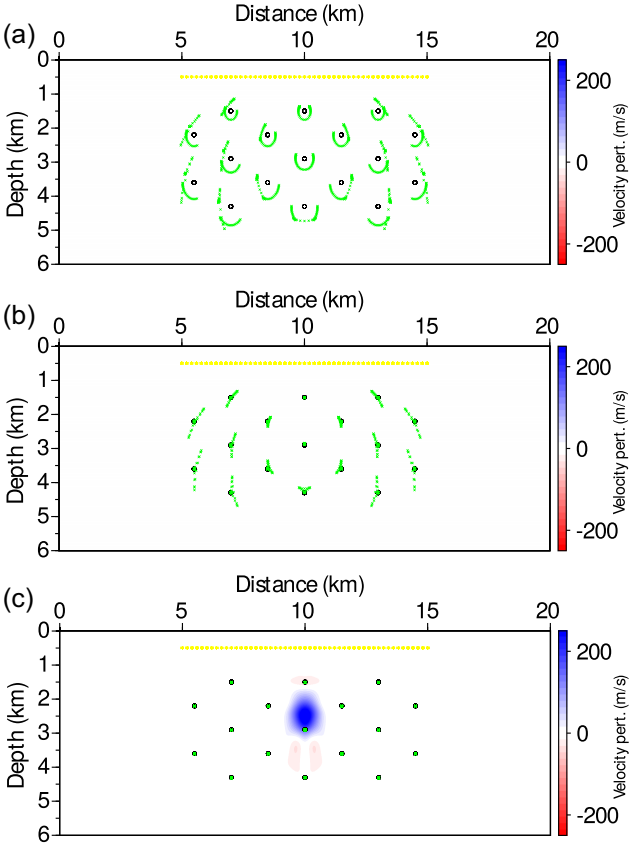

Figure 10. Toy test case: inverting for the origin time correction parameter using the initial velocity model of Fig. 5. Virtual locations (green cross) of 17 events (black circle) corresponding to all receivers (yellow asterisk). Focused locations at (a) the initial stage of the inversion, (b) final stage after 6 iterations where velocity is passive parameter and (c) the joint update result after 83 iterations. The focused locations at the final stage where velocity is passive in the inversion exhibit the same spread as the initial stage of the case where the velocity was inverted using the true origin time (Fig. 5a).

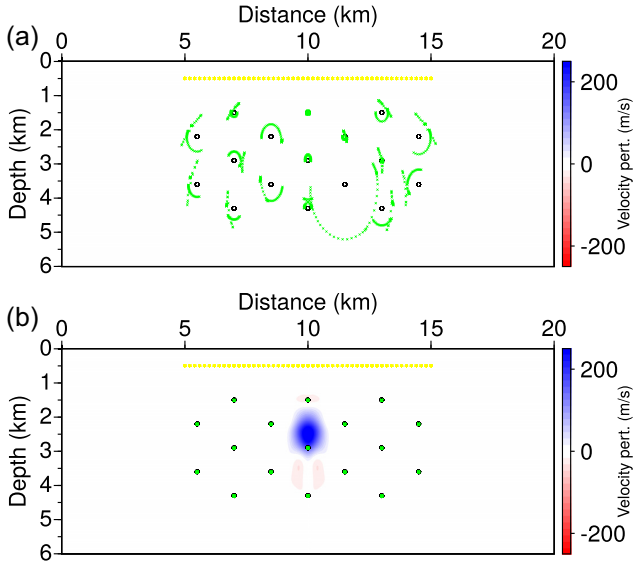

Figure 11. Toy test case: multiparameter inversion (different origin time mismatch for every event). Virtual locations (green cross) of 17 events (black circle) corresponding to all receivers (yellow asterisk). Focused locations in the case of (a) the initial velocity model and (b) the updated velocity model after 76 iterations (bottom).

remind the reader that in a scenario presented previously where we inverted for velocity using the wrong origin time (Fig. 7b), we were able to collapse all positions at one coordinate. The latter was possible since the exercise is a bit simple (same origin time error), the velocity updates compensated for the erroneous origin time but as we have seen for the opposite experiment the result is different.

2.2.3 Multiparameter inversion for wave speeds and origin times

Having presented our formulation and validated the gradient with respect to each parameter, we proceed in this section with the inversion tests in a multiparameter setup. We note that a scaling is applied on the parameters in order to make them dimensionless, hence granting equal contribution in the direction of descent. The scaling factor is kept constant since in the course of the inversion the L-BFGS approximate Hessian is expected to handle the scaling intrinsically. We first look at the simpler version of the previous test where the origin time error is the same for all events (Fig. 10a). In Fig. 10(c), the inversion result after 83 iterations is very satisfactory since it is very similar to the case where we inverted for velocity, while using the exact origin time (Fig. 5b). More smearing is present around the ball but nevertheless the perturbation is well recovered to the same extent while the origin time correction parameter has been correctly estimated leading to a near-perfect focusing of the events. We invite the reader to check the animated graphics (viii) of the Supporting Information for a look on the evolution of the inverted origin time correction parameter through the iterations with an insight on the misfit value. We now proceed with the final inversion where we invert all parameters but for the case where the origin time mismatch for each event is different (Fig. 11a). Following 76 iterations, the result is very similar to the previous case even-though very different origin time errors were introduced (both negative and positive with very different magnitudes; Fig. 11b). We invite the reader to check the animated graphics (ix) of the Supporting Information for a look on the evolution of the inverted origin time correction parameter (and the exact value for each event) through the iterations with an insight on the misfit value.

3 FURTHER NUMERICAL VALIDATION ON THE MARMOUSI BENCHMARK

We benchmark the method on the Marmousi model (Bourgeois *et al.* 1991; Fig. 12a). The latter is a well-known benchmark for exploration scale tomographic methods. The complexity of the structures and the abrupt contrasts are challenging to recover even for tomography techniques (Audebert *et al.* 1997). The following test is evidently more nonlinear than the toy test case presented previously in this study. We remind the reader that the purpose of the approach is the recovery of the events locations while 'absorbing' the errors originating from an erroneous velocity model. In fact, it would be unrealistic to expect a tomographic recovery of the Marmousi model or to invert for the locations starting from crude initial models using few sparse events.

The experimental setup mimics a dense surface acquisition consisting of 227 receivers spaced 40 m apart recording 2 distinct events at ($x = 4560$ m, $z = 1380$ m) and ($x = 5000$ m, $z = 1040$ m). We note that the acquisition design does not represent the sparsity of receivers encountered at all scales and we reiterate on the fact that only two events are used in two dimensions, hence the added challenge of resolving the subsurface parameters structures. We simulate the data set using a smoother version of the original model (Fig. 12b) in order to ensure the validity of the single-arrival assumption while preserving the kinematic properties.

The initial model used during this numerical test is presented in Fig. 13(a). Even though the initial model is a low-frequency representation of the true model, the kinematically migrated virtual events have a significant spreading pattern (Fig. 13a, black circles). The latter is

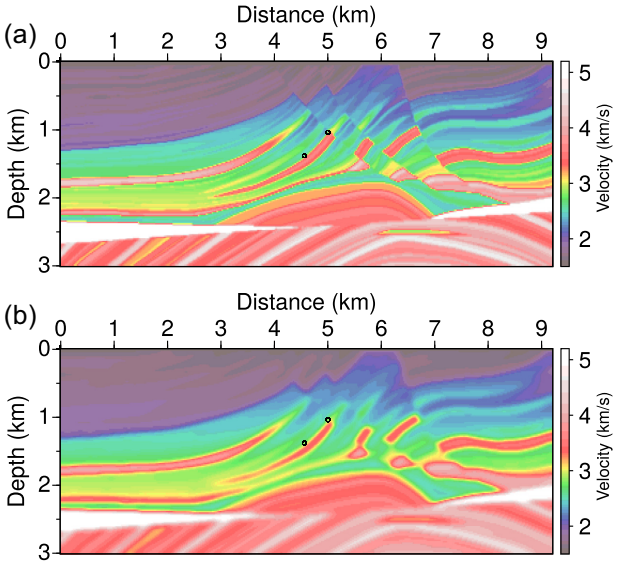

Figure 12. Marmousi case: (a) the true blocky velocity model and (b) the model used for generating the data. Locations of event number one ($x = 5000$ m, $z = 1040$ m) and event number two ($x = 4560$ m, $z = 1380$ m) are denoted by a black circle.

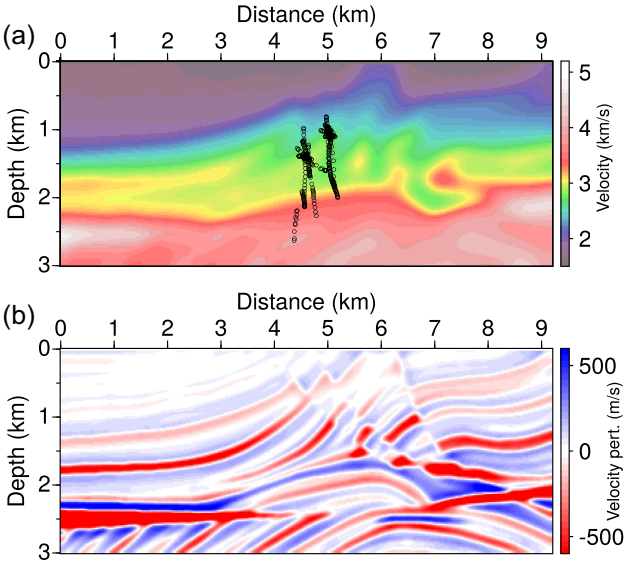

Figure 13. Marmousi case: (a) the initial velocity model and (b) its corresponding velocity perturbations with respect to the true tomographic model. Virtual event locations are denoted by a black circle (using true origin time).

unsurprising since the perturbations missing from the model have a magnitude up to 2000 m s^{-1} (Fig. 13b). The rays in Figs 14(a) and 15(a) reveal the channelling of the rays occurring in the true tomography model along the high velocity layers. On the other hand, the rays shot in the initial model exhibit less channelling due to the smoothing in comparison to the true model (Figs 14b and 15b).

3.1 Can we constrain the subsurface parameters using few events?

3.1.1 All-at-once inversion

As a first test, we invert for the velocity with the objective of finding the event location using the true origin times as passive parameters. The unrealistic experimental setup of this test is chosen in order to assess the difficulty in recovering the main parameters without the bias of the origin time error. The initial velocity model and the virtual events located in this model by kinematic migration are reminded in Fig. 16(a). We first invert the full data set in one go (namely, using the full array of stations). The velocity perturbations and the position of the virtual events after 20 inversion iterations are shown in Fig. 16(b). The result shows that the inversion stopped at a local minimum where a line search failure occurred. In fact, through the iterations, the virtual events tied to the near receivers were collapsing towards the true location especially for

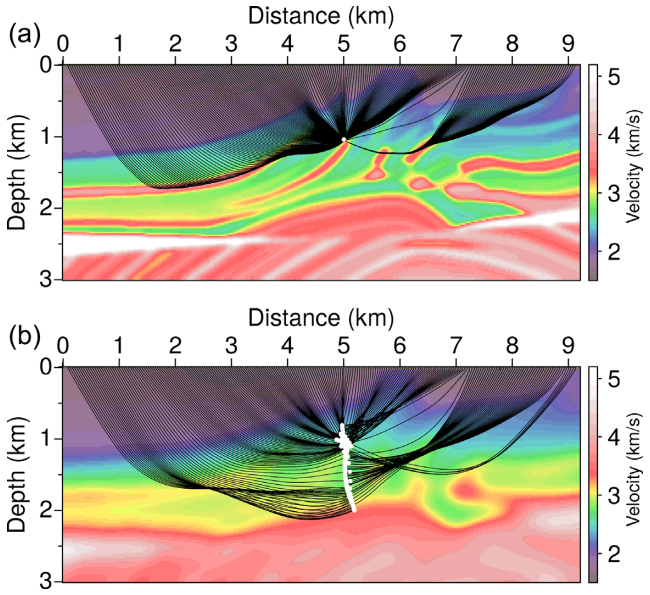

Figure 14. Marmousi case: (a) the true tomographic and (b) the initial models. Virtual event locations are denoted by a solid white circle and black lines are the rays connecting each virtual event related to event number one ($x = 5000$ m, $z = 1040$ m) to its receiver (using true origin time).

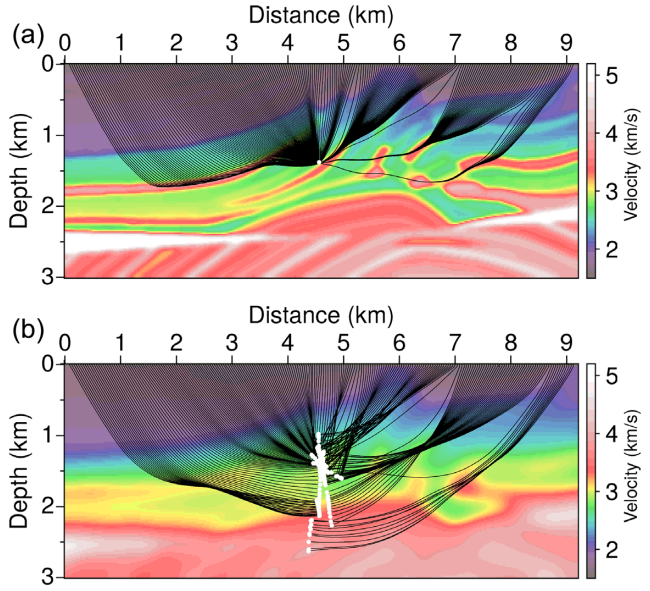

Figure 15. Marmousi case: (a) the true tomographic and (b) the initial models. Virtual events locations are denoted by a solid white circle and black lines are the rays connecting each virtual event related to event number two ($x = 4560$ m, $z = 1380$ m) to its receiver (using true origin time).

the virtual events associated with the rightmost event. On the other hand, for both events, the virtual events tied to the far receivers at 0–1 km and 7–9 km in distance cannot be moved. Looking back at the rays computed in the true velocity model (Figs 14a and 15a), we can see that all the rays connecting the far receivers to the true position of the event are channelled along the high velocity layer (leftmost part) and the dipping high velocity structures (5.5–8 km) at $z = 1.75$ km. We stress that the missing perturbations in these structures are of the order of 700 m s^{-1}, which is significant (Fig. 13b). The latter discrepancy leads to quite different ray trajectories in the true and the starting models (Figs 14 and 15), hence making the tomographic problem highly nonlinear (Hole 1992). Starting from the smooth initial model, the inversion failed to reconstruct such high velocity layer due to insufficient ray-path illumination and remained stuck into a local minimum. In order to further back up the previous statement, we present the traveltime and slope residuals in Figs 17 and 18. The initial stage of the inversion (panels a and e of Figs 17 and 18) generates the highest misfit that is partially reduced at the local minimum at the end of the inversion (panels b and f of Figs 17 and 18). We invite the reader to check the animated graphics (x) of the Supporting Information for a look on the evolution of the inverted parameters through the iterations with a special look on how the virtual events migrated by the central receivers evolved much faster towards a good solution, while the inversion struggled with the virtual events tied to the distant receivers.

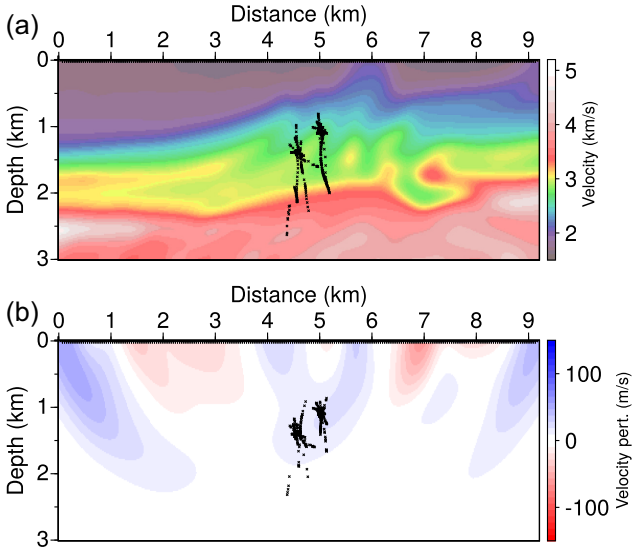

Figure 16. Marmousi case: inverting for velocity with exact origin time. Virtual locations (black cross) of the two events (black circle) migrated from all receiver positions (black asterisk). Focused locations in the case of (a) the initial velocity model and (b) the updated velocity model. In (b), the velocity perturbations are shown after 20 iterations. The inversion remained stuck in a local minimum and failed to collapse the virtual positions at the true positions of the two events.

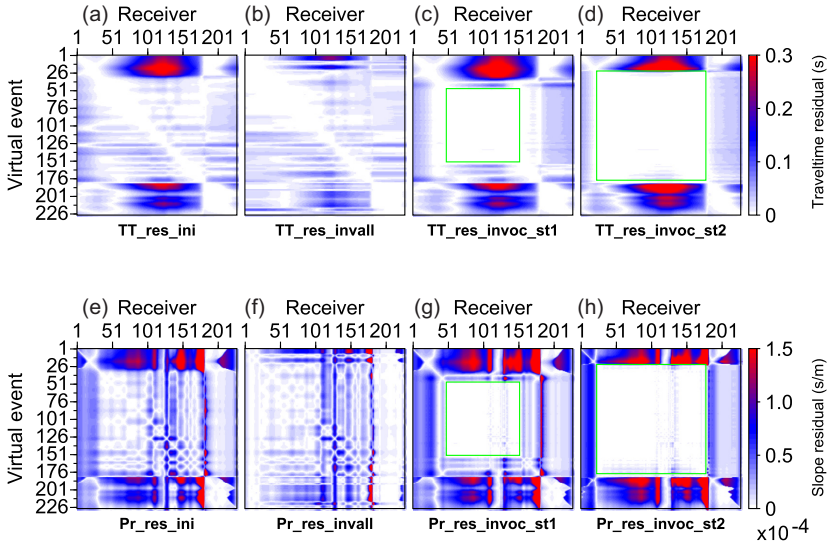

Figure 17. Marmousi case: (a–d) traveltime and (e–h) slope absolute misfit of the event number one at the initial stage (a,e), the final stage of the full acquisition setup (b,f), the first stage of the offset continuation setup (c,g) and the final stage of the offset continuation setup (d,h), respectively. The green squares denote the limited number of receivers used during the offset continuation setup.

3.1.2 Multi-acquisition inversion through offset continuation

Along this line of thought, a possible recipe would be to perform the inversion with a more restrained lateral extension of the acquisition (e.g. considering only stations between 2 and 6 km in distance). The latter strategy could resolve our problem related to the nonlinearity generated by distant stations. However, restraining the angular illumination would also aggravate the velocity–position ambiguity.

Therefore, an improved strategy would be to restrain the acquisition in the earlier steps of the inversion as above-mentioned and then feed the inversion with more picks along the way by involving more distance stations in the inversion. We now present the results of this offset-continuation strategy. During the first stage of the multi-offset inversion, we restrain the acquisition by considering stations located between 2 and 6 km in distance. With this setup, the inversion starts with a more compact spread of the virtual event locations due to the limited acquisition illumination (Fig. 19a). Following 45 iterations, the virtual events spread is fairly minimal and located near the true location but with a slight consistent shift (Fig. 19b). The velocity updates in the well covered zone at the centre are consistent with the dip of the structures, which validates further the obtained result at this stage of the inversion. In addition, the traveltime and slope misfits in the area covered by the restricted station array nearly vanished as illustrated in Panels c and g of Figs 17 and 18, green squares. Following the

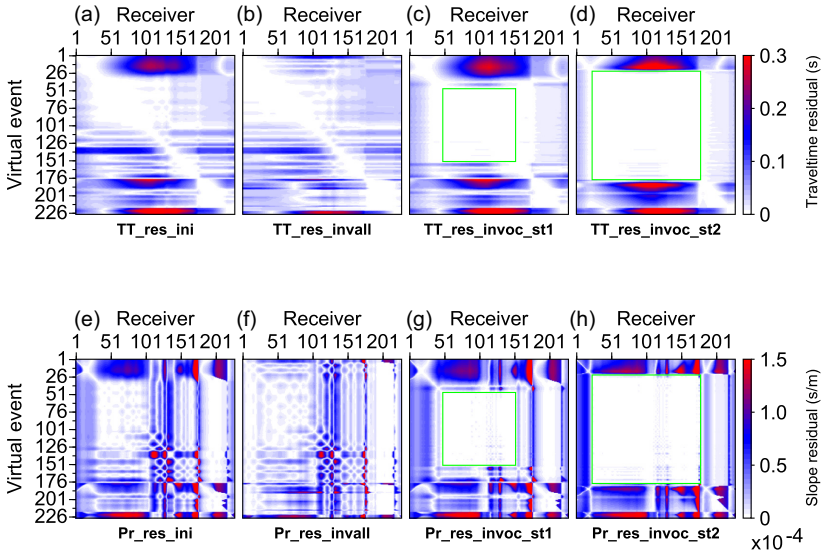

Figure 18. Marmousi case: same as Fig. 17 for event two.

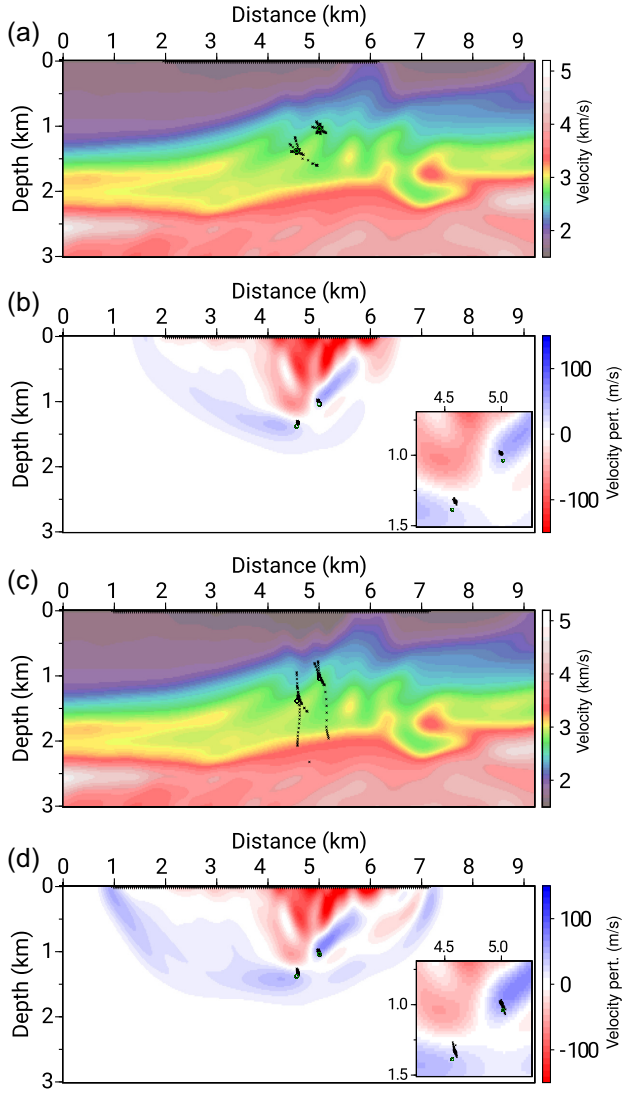

Figure 19. Marmousi case: inverting for velocity with exact origin time. Virtual locations (black cross) of the two events (circle) migrated from all receivers (black asterisk). (a) Focused locations in the case of the initial velocity model, (b) the velocity perturbations added to the initial model after 45 iterations and its focused scatterers, (c) the velocity model at the initial stage of the last extension and its focused scatterers and (d) the velocity perturbations added to the initial model after 71 iterations using extended lateral receiver coverage.

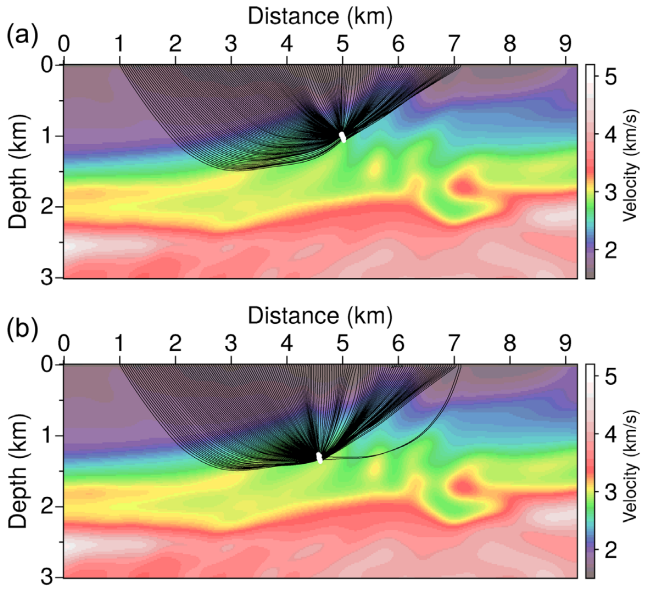

Figure 20. Marmousi case: virtual events and rays computed in the velocity model shown in Fig. 19(c). Virtual event locations denoted by a solid white circle and black lines are the rays connecting each virtual event to their corresponding receivers. Panels (a) and (b) correspond to events 1 and 2, respectively.

offset-continuation strategy, we push further the inversion by injecting more data (1 km on each side; Fig. 19c). We remind the reader that the spreading pattern seen in Fig. 19(c) is coming from the fact that the newly introduced data were not fitted at the earlier stage. The final result shows a slightly bigger spreading pattern due to the larger number of virtual events in the process but their mean position is close to the true position (Fig. 19d). The velocity update reflects the impact of the new injected data since perturbations on both sides of the acquisition have been introduced to the model. The final trajectories of the rays traced between the receivers and their associated events are shown in Fig. 20. The velocity perturbations added at 1.5 km depth between 3 and 4 km (Fig. 19d) sharpened and extended the high velocity contrast, hence permitting the bending of the rays and in turn the focusing of the virtual events at their final location (Fig. 20). Similarly, the updates in the shallow part of the velocity model (4–6 km distances) favoured the channelling of the rays leading to a better kinematically migrated events. As a further quality control, the misfit maps in Figs 17(d)–(h) and 18(d)–(h) reveal that the newly introduced data were properly fitted. We invite the reader to check the animated graphics (xi) of the Supporting Information for a look on the evolution of the inverted parameters through the iterations with a special look on how the inversion reacts to the injection of the extended acquisition in terms of velocity updates and mean position (red cross).

3.2 Multi-parameter inversion for wave speeds and origin times

The final numerical test presented in this report is the joint inversion of the subsurface parameters (namely velocity here) and the origin time correction parameters. We remind the reader again that the main objective is the recovery of the event locations. We introduce an error in the data associated with the origin time corresponding to an overestimation of 0.0712 s and an underestimation of 0.0527 s for events one and two, respectively. The order of magnitude of these mismatches is significant for some receivers since some virtual events have a traveltime as low as 0.3 s. We employ the same multi-acquisition strategy as that presented in the previous test with the sole difference being the need of proper parameter scaling during the different stages of the inversion. At the initial stage of the inversion, the spread of the virtual locations (Fig. 21a) is different compared to the previous case (Fig. 19a) due to the introduced origin time mismatch. The results presented in Fig. 21 show that, at the intermediate and final stages of the inversion, the result is comparable to the mono-parameter inversion case (compare Figs 21b and c & 19b and c). The result is overall satisfying with a slight final mismatch in the origin time correction parameter (the final values were −0.0798 and 0.0498 s). The latter issue could be resolved with better scaling strategies or a more accurate Hessian, which naturally balances the weight of each parameter during the optimization. We invite the reader to check the animated graphics (xii) of the Supporting Information for a look on the evolution of the inverted parameters through the iterations with a special look on how the origin time parameter values evolve during the iterations.

4 DISCUSSION

We have proposed in this study a consistent formulation of the hypocentre–velocity problem under a framework based on eikonal solvers and the adjoint-state method.

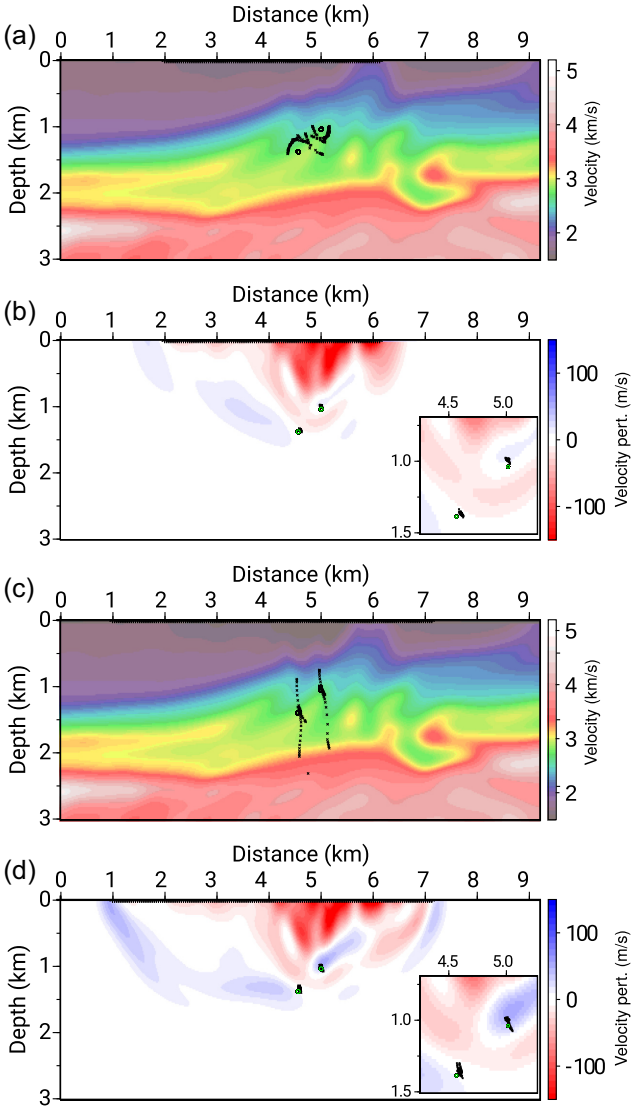

Figure 21. Marmousi case: inverting for velocity and origin time corrections. Virtual locations (black cross) of the two events (circle) migrated from all receivers (black asterisk). (a) Focused locations in the case of the initial velocity model, (b) the velocity perturbations added to the initial model after 49 iterations and its focused scatterers, (c) the velocity model at the initial stage of the last extension and its focused scatterers and (d) the velocity perturbations added to the initial model after 77 iterations using extended lateral receiver coverage.

4.1 A parsimonious variable-projection approach

Our method differs from the vast majority of previous studies in the sense that the unknown positions of the events are not processed as parameters of a least-squares (overdetermined) optimization problem where velocities are either passive quantities or another class of variables. Instead, these positions are eliminated from the optimization problem by tying them to the velocities via the kinematic migration of the observables (traveltime and slope). This implies that the position of the events are computed explicitly at each iteration as part of the forward problem (i.e. as state variables) by solving a pair of focusing (state) equations whose right-hand sides are the observables. In this framework, the only optimization variables to update are the wave speeds leading to a better-posed optimization problem under this parsimonious parametrization. We also tackle the origin time issue through a correction parameter and we discussed its importance in the inversion.

4.2 Time-reversal migration versus tomographic approaches

With this time-reversal migration-based logic, the position of an event is indeed seen differently by each receiver if the velocities are inaccurate. We propose in the Appendix an alternative formulation which relies on the intuitive idea of collapsing the virtual positions migrated from all the receivers to the true event location by updating velocities. In this formulation, the position of each virtual event is successively processed as an observable to be matched by the positions of the other virtual events, the latter being processed as state variables. The drawback of this

formulation is related to its sensitivity kernel that provides a sub-optimal illumination of the subsurface in the sense that it connects each virtual event to the sole receiver from which it was migrated along a transmitted one-way path (similar to the subkernel shown in Fig. 3b).

To overcome the limit of this formulation inspired by time-reversal migration-based methods, we have developed a formulation with a more reflection-tomography oriented logic inspired by Sambolian *et al.* (2019) in the sense that the observables are the traveltimes and slopes at the receivers other than the one used to position the virtual event by kinematic migration. The mismatches between these observables and the attributes simulated from the virtual event position are minimized to update the wave speeds and the origin times with as a direct consequence the collapsing of the smeared positions of the virtual events to their true position. This approach enriches the one-way path kernel of the previous formulation connecting the receiver to the virtual event with multiple paths connecting the virtual event to the other receivers, hence mimicking the two-way paths of reflected waves (Fig. 3a).

4.3 Forthcoming improvements of the method

The Marmousi case revealed however some shortcomings of the method. As any tomographic method, the approach is sensitive to the subsurface illumination, which is itself controlled by the distribution in depth of the events to be located. The optimization problem is extremely ill-posed when the number of events is limited. Moreover, the Marmousi case revealed the non linearity of the relocation/tomograpic problem when sharp contrasts lack in the initial velocity model. These sharp contrasts behave as refractors, which channel the rays connecting the events to far receivers. This prevents uniform ray-path coverage of the subsurface and makes the ray trajectories in the true and initial models quite different, hence violating the ray-stationarity assumption underlying the linearization of nonlinear traveltime tomography (Hole 1992). We have shown that this nonlinearity combined with non-uniform ray coverage can trap the localization problem into a local minimum. To bypass this nonlinearity issue without increasing the number of events, we have proposed a pragmatical solution in our application by using an offset (i.e., the horizontal distance between the event and the station) continuation strategy. The introduction of the restrained lateral extent data set in early stages of the inversion guided the inversion towards reconstructing progressively the missing perturbations in the well covered area. The injection of distant receivers in the late stages of the inversion was guided, through the kinematic migration, by the already recovered solution of the subsurface parameters. By following this offset continuation, we aggravated the velocity–position coupling during the early stage of the inversion (manifested by a consistent shift in position of the virtual events) since we restrained the illumination but we mitigated the non linearity of the tomography, as supported by convergence of the inversion towards far better-focused positions.

The need to mitigate the nonlinearity of the inverse problem generated by inaccurate initial subsurface model together with its ill-posedness generated by uneven illumination may prompt us to implement at the beginning of the inversion a relaxation of the focusing equations governing the kinematic migration. This relaxation will re-extend the search space of the inversion, which was restricted by making the relocation problem consistent with the subsurface parameters through the variable projection. Indeed, the migration of the kinematic attributes for event relocation gives no leeway to the inversion to explore solutions where the virtual events are forced to collapse while the focusing equations are not strictly satisfied. We stress that we don't suggest to go back to strategies where the focusing equations are relaxed without any control over their satisfaction at the sought solution. Instead, strategies that reconcile the relaxation of the focusing equations with more freedom to constrain the distance between virtual events early-on in the inversion while controlling their satisfaction at the convergence point may be implemented with augmented Lagrangian method (or method of multiplier), a versatile method to solve constrained optimization problem by combining a Lagrangian method and a penalty method (for example the spatial spread between the virtual events (eq. A2) could serve as an additional constraint or penalty function). We refer the reader to Nocedal & Wright (2006, their Chapter 17) for a review and Delbos *et al.* (2006) for a tomography oriented example.

Let's conclude by clarifying that the method can be straightforwardly extended to 3-D by incorporating the backazimuth or an additional slope [crossline slope as used by Chalard *et al.* (2000) for stereotomography at exploration scale] as a supplementary attribute. In this framework, the ill-posedness and nonlinearity of the localization problem reviewed above should be significantly mitigated by the richer subsurface illumination provided by areal deployment.

5 CONCLUSION

We revisit the location–velocity problem with a novel angle inspired by slope tomography. We propose our strategy under a framework based on eikonal solvers and the adjoint-state method. The approach was validated on a simple toy test and benchmarked against the Marmousi case. We present a proof of concept but the approach should be tested on a more realistic case since the method is extendible to three dimensions by accounting for the crossline slope or the backazimuth as an extra constraint for the kinematic migration. Future investigations will revolve around coupling the arrival times of *P* and *S* waves under this framework and an extension to reflected arrivals.

ACKNOWLEDGEMENTS

This study was partially funded by the WIND consortium (https://www.geoazur.fr/WIND), sponsored by Chevron, Shell and Total. This study was granted access to the HPC resources of SIGAMM infrastructure (http://crimson.oca.eu), hosted by Observatoire de la Côte d'Azur

and which is supported by the Région Provence-Alpes-Côte d'Azur, and the HPC resources of CINES/IDRIS/TGCC under the allocation A0050410596 made by GENCI. We thank the Editor Prof. Huajian Yao and two anonymous reviewers for their comments.

REFERENCES

Aghamiry, H., Gholami, A. & Operto, S., 2020. Wavefield inversion for microseismic imaging, in *SEG Technical Program Expanded Abstracts 2020,* pp. 2130–2134, Society of Exploration Geophysicists.

Akçelik, V., 2002. Multiscale Newton–Krylov methods for inverse acoustic wave propagation, *PhD thesis*, Carnegie Mellon University, Pittsburgh, Pennsylvania.

Alkhalifah, T., 1998. Acoustic approximations for processing in transversely isotropic media, *Geophysics,* **63,** 623–631.

Alkhalifah, T. & Fomel, S., 2010. An eikonal based formulation for traveltime perturbation with respect to the source location, *Geophysics,* **75**(6), T175–T183.

Alkhalifah, T. & Tsvankin, I., 1995. Velocity analysis for transversely isotropic media, *Geophysics,* **60,** 1550–1566.

Artman, B., Podladtchikov, I. & Witten, B., 2010. Source location using time-reverse imaging, *Geophys. Prospect.,* **58**(5), 861–873.

Audebert, F., Nichols, D.E., Rekdal, T., Biondi, B., Lumley, D.E. & Urdaneta, H., 1997. Imaging complex geologic structure with single-arrival Kirchhoff prestack depth migration, *Geophysics,* **62**(5), 1533–1543.

Baker, T., Granat, R. & Clayton, R.W., 2005. Real-time earthquake location using Kirchhoff reconstruction, *Bull. seism. Soc. Am.,* **95**(2), 699–707.

Bauer, A., Schwarz, B. & Gajewski, D., 2017. Utilizing diffractions in wavefront tomography, *Geophysics,* **82**(2), R65–R73.

Bauer, A., Schwarz, B., Werner, T. & Gajewski, D., 2019. Unsupervised event identification and tagging for diffraction focusing, *Geophys. J. Int.,* **217**(3), 2165–2176.

Bourgeois, A., Bourget, M., Lailly, P., Poulet, M., Ricarte, P. & Versteeg, R., 1991. Marmousi, model and data, in *The Marmousi Experience,* pp. 5–16, European Association of Geoscientists and Engineers.

Byrd, R.H., Lu, P. & Nocedal, J., 1995. A limited memory algorithm for bound constrained optimization, *SIAM J. Sci. Stat. Comput.,* **16,** 1190–1208.

Chalard, E., Podvin, P., Lambaré, G. & Audebert, F., 2000. Principles of 3-D stereotomography, in *Extended Abstracts, 62nd Annual Meeting (29 May–2 June 2000, Glasgow),* P144, European Association of Geoscientists & Engineers, doi:10.3997/2214-4609-pdb.28.P144.

Chauris, H., Noble, M., Lambaré, G. & Podvin, P., 2002. Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I: theoretical aspects, *Geophysics,* **67**(4), 1202–1212.

Deflandre, J.-P., 2016. Induced microseismicity: short overview, state of the art and feedback on source rock production, *Open Pet. Eng. J.,* **9**(1), 55–71.

Delbos, F., Gilbert, J.C., Glowinski, R. & Sinoquet, D., 2006. Constrained optimization in seismic reflection tomography: a Gauss–Newton augmented Lagrangian approach, *Geophys. J. Int.,* **164**(3), 670–684.

Diekmann, L., Schwarz, B., Bauer, A. & Gajewski, D., 2019. Source localization and joint velocity model building using wavefront attributes, *Geophys. J. Int.,* **219**(2), 995–1007.

Dummong, S., Meier, K., Gajewski, D. & Hubscher, C., 2008. Comparison of prestack stereotomography and NIP wave tomography for velocity model building: instances from the Messinian evaporites, *Geophysics,* **73**(5), VE291–VE302.

Duveneck, E., 2004. Velocity model estimation with data-derived wavefront attributes, *Geophysics,* **69**(1), 265–274.

Fichtner, A., 2010. *Full Seismic Waveform Modelling and Inversion,* Springer.

Fink, M., 1993. Time-reversal mirrors, *J. Phys. D: Appl. Phys.,* **26**(9), 1333–1350.

Fomel, S., Luo, S. & Zhao, H.-K., 2009. Fast sweeping method for the factored eikonal equation, *J. Comput. Phys.,* **228,** 6440–6455.

Font, Y., Kao, H., Lallemand, S., Liu, C.-S. & Chiao, J.-Y., 2004. Hypocentral determination offshore eastern Taiwan using the maximum intersection method, *Geophys. J. Int.,* **158,** 655–675.

Frohlich, C. & Pulliam, J., 1999. Single-station location of seismic events: a review and a plea for more research, *Phys. Earth planet. Inter.,* **113**(1–4), 277–291.

Gajewski, D. & Tessmer, E., 2005. Reverse modelling for seismic event characterization, *Geophys. J. Int.,* **163**(1), 276–284.

Geiger, L., 1912. Probability method for the determination of earthquake epicenters from the arrival time only, *St. Louis Univ. Bull.,* **8,** 60–71.

Gholami, Y., Brossier, R., Operto, S., Ribodetti, A. & Virieux, J., 2013. Which parametrization is suitable for acoustic VTI full waveform inversion?—Part 1: sensitivity and trade-off analysis, *Geophysics,* **78**(2), R81–R105.

Golub, G. & Pereyra, V., 2003. Separable nonlinear least squares: the variable projection method and its applications, *Inverse Probl.,* **19**(2), R1, doi:10.1088/0266-5611/19/2/201.

Grechka, V., Mazumdar, P. & Shapiro, S.A., 2010. Predicting permeability and gas production of hydraulically fractured tight sands from microseismic data, *Geophysics,* **75**(1), B1–B10.

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

Hicks, G.J., 2002. Arbitrary source and receiver positioning in finite-difference schemes using Kaiser windowed sinc functions, *Geophysics,* **67,** 156–166.

Hole, J.A., 1992. Nonlinear high-resolution three-dimensional seismic travel time tomography, *J. geophys. Res.,* **97,** 6553–6562.

Hu, J., Zhang, H. & Yu, H., 2018. Accurate determination of *P*-wave back azimuth and slowness parameters by sparsity constrained seismic array analysis, *Geophys. J. Int.,* **181,** 1–18

Ishii, M., Shearer, P.M., Houston, H. & Vidale, J.E., 2007. Teleseismic *P* wave imaging of the 26 December 2004 Sumatra–Andaman and 28 March 2005 Sumatra earthquake ruptures using the Hi-net array, *J. geophys. Res.,* **112,** B11307, doi:10.1029/2006JB004700.

Jones, G., Kendall, J.-M., Bastow, I. & Raymer, D., 2013. Locating microseismic events using borehole data, *Geophys. Prospect.,* **62**(1), 34–49.

Kamei, R. & Lumley, D., 2014. Passive seismic imaging and velocity inversion using full wavefield methods, in *SEG Technical Program Expanded Abstracts 2014,* pp. 2273–2277.

Kelley, C., 1995. *Iterative Methods for Linear and Nonlinear Equations,* SIAM.

Kito, T. & Korenaga, J., 2010. Cross-correlation weighted migration: towards high-resolution mapping of mantle heterogeneities, *Geophys. J. Int.,* **181,** 1109–1127

Kito, T., Rietbrock, A. & Thomas, C., 2007. Slowness-backazimuth weighted migration: a new array approach to a high-resolution image, *Geophys. J. Int.,* **169**(3), 1201–1209.

Krüger, F., Weber, M., Scherbaum, F. & Schlittenhardt, J., 1993. Double beam analysis of the anomalies in the core-mantle boundary region, *Geophys. Res. Lett.,* **20**(14), 1475–1478.

Lambaré, G., 2008. Stereotomography, *Geophysics,* **73**(5), VE25–VE34.

Larmat, C., Montagner, J.-P., Fink, M., Capdeville, Y., Tourin, A. & Clévédé, E., 2006. Time-reversal imaging of seismic sources and application to the great Sumatra earthquake, *Geophys. Res. Lett.,* **33,** L19312, doi:10.1029/2006GL026336.

Li, L., Chen, H. & Wang, X.-M., 2015. Weighted-elastic-wave interferometric imaging of microseismic source location, *Appl. Geophys.,* **12**(2), 221–234.

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

Lomax, A., Virieux, J., Volant, P. & Berge, C., 2000. Probabilistic earthquake location in 3D and layered models: introduction of a Metropolis–Gibbs method and comparison with linear locations, in *Advances in Seismic Event Location,* Vol. 281, pp. 101–134, eds Thurber, C.H. & Rabinowitz, N., Kluwer Academic Publisher.

Lomax, A., Michelini, A. & Curtis, A., 2009, Earthquake location, direct, global-search methods, in *Encyclopedia of Complexity and Systems Science,* pp. 1–33, ed. Meyers, R., Springer.

Luo, S. & Qian, J., 2012. Fast sweeping method for factored anisotropic eikonal equations: multipicative and additive factors, *J. Sci. Comput.,* **52,** 360–382.

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

Métivier, L. & Brossier, R., 2016. The SEISCOPE optimization toolbox: a large-scale nonlinear optimization library based on reverse communication, *Geophysics,* **81**(2), F11–F25.

Monteiller, V., Got, J.-L., Virieux, J. & Okubo, P., 2005. An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to the Kilauea volcano, *J. geophys. Res.,* **110**(B12306), doi:10.1029/2004JB00346.

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

Pavlis, G. & Booker, J., 1980. The mixed discrete-continuous inverse problem: application to the simultaneous determination of earthquake hypocenter and velocity structure, *J. geophys. Res.,* **85,** 4801–4810.

Plessix, R.E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.,* **167**(2), 495–503.

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.

Qian, J. & Symes, W., 2002. An adaptive finite-difference method for traveltimes and amplitudes, *Geophysics,* **67,** 167–176.

Rentsch, S., Buske, S., Lth, S. & Shapiro, S.A., 2007. Fast location of seismicity: a migration-type approach with application to hydraulic-fracturing data, *Geophysics,* **72**(1), S33–S40.

Riabinkin, L.A., 1957. *Fundamentals of Resolving Power of Controlled Directional Reception (CDR) of Seismic Waves,* Vol. 14, pp. 36–60, Society of Exploration Geophysics, Translated and paraphrased from Prikladnaya, 16, 3–36.

Rietbrock, A. & Scherbaum, F., 1994. Acoustic imaging of earthquake sources from the Chalfant Valley, 1986, aftershock series, *Geophys. J. Int.,* **119**(1), 260–268.

Roecker, S., Thurber, C., Roberts, K. & Powell, L., 2006. Refining the image of the San Andreas fault near Parkfield, California using a finite difference travel time computation technique, *Tectonophysics,* **426**(1–2), 189–205.

Roecker, S.W., 1982. Velocity structure of the Pamir-Hindu Kush region: possible evidence of subducted crust, *J. geophys. Res.,* **87**(B2), 945–959.

Ross, Z.E., Yue, Y., Meier, M.-A., Hauksson, E. & Heaton, T.H., 2019. PhaseLink: a deep learning approach to seismic phase association, *J. geophys. Res.,* **124**(1), 856–869.

Rost, S. & Thomas, C., 2002. Array seismology: methods and applications, *Rev. Geophys.,* **40**(3), 1008.

Sambolian, S., Operto, S., Ribodetti, A., Tavakoli, B. & Virieux, J., 2019. Parsimonious slope tomography based on eikonal solvers and the adjoint-state method, *Geophys. J. Int.,* **218**(1), 456–478.

Scherbaum, F., Krüger, F. & Weber, M., 1997. Double beam imaging: mapping lower mantle heterogeneities using combinations of source and receiver arrays, *J. geophys. Res.,* **102**(B1), 507–522.

Schuster, G.T., Yu, J., Sheng, J. & Rickett, J., 2004. Interferometric/daylight seismic imaging, *Geophys. J. Int.,* **157**(2), 838–852.

Schwarz, B., Bauer, A. & Gajewski, D., 2016. Passive seismic source localization via common-reflection-surface attributes, *Stud. Geophys. Geod.,* **60,** 531–546.

Sollberger, D., Greenhalgh, S.A., Schmelzbach, C., Renterghem, C.V. & Robertsson, J.O.A., 2018. 6-C polarization analysis using point measurements of translational and rotational ground-motion: theory and applications, *Geophys. J. Int.,* **213,** 77–97.

Song, C., Alkhalifah, T. & Wu, Z., 2019. Microseismic event estimation and velocity analysis based on a source-focusing function, *Geophysics,* **84**(3), KS85–KS94.

Spencer, C. & Gubbins, D., 1980. Travel time inversion for simultaneous earthquake location and velocity structure determination in laterally varying media, *Geophys. J. R. astr. Soc.,* **63,** 95–116.

Stork, C. & Clayton, R.W., 1986. Analysis of the resolution between ambiguous velocity and reflector position for traveltime tomography, in *SEG Technical Program Expanded Abstracts 1986,* pp. 545–550, Society of Exploration Geophysicists.

Taillandier, C., Noble, M., Chauris, H. & Calandra, H., 2009. First-arrival travel time tomography based on the adjoint state method, *Geophysics,* **74**(6), WCB1–WCB10.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics,* **49**(8), 1259–1266.

Tarantola, A., 1987. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation,* Elsevier.

Tavakoli F., B., Ribodetti, A., Virieux, J. & Operto, S., 2015. An iterative factored eikonal solver for TTI media, in *SEG Technical Program Expanded Abstracts 2015,* Vol. 687, pp. 3576–3581, Society of Exploration Geophysics.

Tavakoli F., B., Operto, S., Ribodetti, A. & Virieux, J., 2017. Slope tomography based on eikonal solvers and the adjoint-state method, *Geophys. J. Int.,* **209**(3), 1629–1647.

Tavakoli F., B., Operto, S., Ribodetti, A. & Virieux, J., 2019. Matrix-free anisotropic slope tomography: theory and application, *Geophysics,* **84**(1), R35–R57.

Thurber, C.H., 1992. Hypocenter-velocity structure coupling in local earthquake tomography, *Phys. Earth planet. Inter.,* **75**(1–3), 55–62.

Verdon, J.P., Kendall, J.-M., Hicks, S.P. & Hill, P., 2017. Using beamforming to maximise the detection capability of small, sparse seismometer arrays deployed to monitor oil field activities, *Geophys. Prospect.,* **65**(6), 1582–1596.

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

Waheed, U.B., Yarman, C.E. & Flagg, G., 2014. An iterative fast sweeping based eikonal solver for tilted orthorhombic media, in *Expanded Abstracts,* pp. 480–485, Society of Exploration Geophysics.

Waheed, U.B., Yarman, C.E. & Flagg, G., 2015. An iterative, fast-sweeping-based eikonal solver for 3D tilted anisotropic media, *Geophysics,* **80,** C49–C58.

Waldhauser, F. & Ellsworth, W., 2000. A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California, *Bull. seism. Soc. Am.,* **90**(6), 1353–1368.

Zhao, H., 2005. A fast sweeping method for eikonal equations, *Math. Comput.,* **74,** 603–627.

Zhou, H.-W., 1994. Rapid three-dimensional hypocentral determination using a master station method, *J. geophys. Res.,* **99**(B8), 15 439–15 455.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

01_CA_VP_TrueT0.gif
01bis_CA_VP_TrueT0_onlyTT.gif
01ter_CA_VP_TrueT0_onlySLP.gif
02_CA_VP_WrongT0.gif
03_CA_VP_WrongDiffT0.gif
04_CA_T0_TrueVp.gif

05_CA_T0_IniVp.gif
06_CA_Multiparam.gif
07_CA_Multiparam_Difft0.gif
08_MARM_ALLOFF.gif
09_MARM_CONTOFF.gif
10_MARM_CONTOFF_T0.gif

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX: FRAMEWORK BASED ON MINIMIZING THE SPREAD OF KINEMATICALLY MIGRATED VIRTUAL EVENTS

In this appendix, we present an alternative formulation of the event location problem inspired by time-reversal migration-based methods. Accordingly, the objective function aims to directly optimize the focusing of the migrated virtual events rather than indirectly fulfilling this task by fitting measurements at the stations. We define the following nonlinear constrained minimization problem with the aim of retrieving the minimizer $\mathbf{m}$, gathering the subsurface parameters and an origin time correction parameter,

$$\min_{\mathbf{m}} C(\mathbf{m}) = \min_{\mathbf{m}} \sum_{e=1}^{N_e} \sum_{r=1}^{N_r} \Phi_{e,r}(\mathbf{m}) \quad \text{subject to} \quad \mathbf{F}(\mathbf{u}, \mathbf{m}) = 0, \tag{A1}$$

where N_e and N_r denote the number of events and receivers, respectively. The function $\Phi_{e,r}(\mathbf{m})$ evaluates the spread of the virtual events migrated from each receiver r for a given event e in function of a given model $\mathbf{m}$. In the context of this example, we use the Euclidian distance

$$\Phi_{e,r} = \frac{1}{N_r - 1} \sum_{r'=1, r' \neq r}^{N_r} \sqrt{(x^*_{e,r} - x_{e,r'})^2 + (z^*_{e,r} - z_{e,r'})^2}, \tag{A2}$$

which is normalized by $N_r - 1$ to keep the physical unit of distance. Note that the position of each virtual event is successively processed as an observable as indicated by the superscript * assigned to the fixed subscript r. In other words, the position of each virtual event is in turn an observable (when the subscript r is assigned to it) and a state variable (when the running subscript r' is assigned to it), the associated state equation being the focusing equations used for kinematic migration. The operator $\mathbf{F}$ gathers the forward problem equations related to the data simulation through eikonal-resolved traveltime maps, the focusing equations (Chauris *et al.* 2002), and the distance employed in Φ. We solve the constrained problem, (eq. A1), under a Lagrangian formalism following the same adjoint-state method employed in Section 2.1 using the state equations (3), (7) and (A2) and associating the adjoint-state variables $\tilde{\mathbf{u}} = (\lambda_r(\mathbf{x}), \alpha_{e,r}, \beta_{e,r}, \gamma_{e,r})$ to their respective state variables $\mathbf{u} = (t_r(\mathbf{x}), x_{e,r}, z_{e,r}, \Phi_{e,r})$ leading to

$$\begin{aligned}\mathcal{L}(\mathbf{m}, \mathbf{u}, \tilde{\mathbf{u}}) = J(\mathbf{u}) &- \sum_{e=1}^{N_e} \sum_{r=1}^{N_r} \alpha_{e,r} \left(T^*_{e,r} - Q_{e,r} t_r(\mathbf{x}) - \frac{1}{N_r} \delta_e \right) - \sum_{e=1}^{N_e} \sum_{r=1}^{N_r} \beta_{e,r} \left(p^*_{e,r} - \frac{Q_{e,r}(t_{r+1}(\mathbf{x}) - t_{r-1}(\mathbf{x}))}{2\Delta_r} \right) \\ &- \sum_{e=1}^{N_e} \sum_{r=1}^{N_r} \gamma_{e,r} \left(\Phi_{e,r} - \frac{1}{N_r - 1} \sum_{r'=1, r' \neq r}^{N_r} \sqrt{(x_{e,r} - x_{e,r'})^2 + (z_{e,r} - z_{e,r'})^2} \right) - \frac{1}{2} \sum_{r=1}^{N_r} \Big\langle \lambda_r(\mathbf{x}) \mid H(\mathbf{x}, \nabla t_r(\mathbf{x})) \Big\rangle_\Omega,\end{aligned} \tag{A3}$$

where the Lagrangian functional $\mathcal{L}$ depends on the subsurface parameters $\mathbf{m}$ through the eikonal equation solved in the subsurface domain Ω.

We develop now the adjoint-state equations. Before proceeding with the first adjoint-state equation $\partial \mathcal{L} / \partial \Phi_{e,r} = 0$, we remind the reader that $\Phi_{e,r}$ is an auxiliary variable; the misfit distance could have been explicitly written in $C(\mathbf{m})$. Having said that the first adjoint-state equation leads to

$$\gamma_{e,r} = 1. \tag{A4}$$

For each event, $\partial \mathcal{L} / \partial \mathbf{x}_{e,r} = 0$. Considering that $\mathbf{x}_{e,r}$ gathers $x_{e,r}$ and $z_{e,r}$ (embedded in $Q_{e,r}$) gives the following 2×2 system of linear equations for $\alpha_{e,r}$ and $\beta_{e,r}$:

$$\alpha_{e,r} \frac{\partial Q_{e,r}}{\partial \mathbf{x}_{e,r}} \mathbf{t}_r + \frac{\beta_{e,r}}{2\Delta r} \frac{\partial Q_{e,r}}{\partial \mathbf{x}_{e,r}} (\mathbf{t}_{r+1} - \mathbf{t}_{r-1}) = -\gamma_{e,r} \frac{\partial \Phi_{e,r}}{\partial \mathbf{x}_{e,r}}. \tag{A5}$$

Solving this system leads to the closed form expression of $\alpha_{e,r}$ and $\beta_{e,r}$ as follows:

$$\alpha_{e,r} = -\frac{\det \left| \frac{\partial(\Phi_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})} \right|}{\det \left| \frac{\partial(T_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})} \right|}, \tag{A6}$$

$$\beta_{e,r} = -\frac{\det \left| \frac{\partial(T_{e,r},\ \Phi_{e,r})}{\partial(x_{e,r},\ z_{e,r})} \right|}{\det \left| \frac{\partial(T_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})} \right|}. \tag{A7}$$

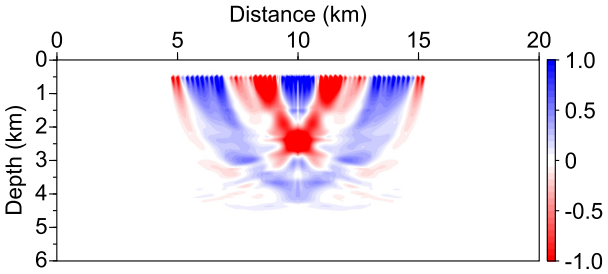

Figure A1. Gradient of $J(\mathbf{m})$ (eq. A3) with respect to velocity for the case presented in Fig. 2.

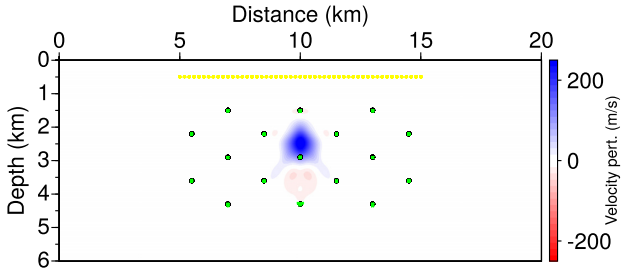

Figure A2. Toy test case: multiparameter inversion (different origin time mismatch for every event) using the time-reversal migration-based formulation. The figure shows the updated velocity model after 100 iterations.

The resultant adjoint-state variables $\alpha_{e,r}$ and $\beta_{e,r}$ describe how $\Phi_{e,r}$ evolves when the coordinates $\mathbf{x}_{e,r}$ of an event are moved by a velocity update (refer to Section 2.1 for a more detailed interpretation of the determinants).

We note that eqs (A6) and (A7) are only defined for $\det \left| \frac{\partial(T_{e,r},\ p_{e,r})}{\partial(x_{e,r},\ z_{e,r})} \right| \neq 0$. The system is always valid since in our case this condition could never be violated as already mentioned in Section 2.1. Proceeding with the last derivative $\partial\mathcal{L}/\partial\mathbf{t}_r = 0$, we obtain the adjoint-state equation satisfied by $\lambda_r(\mathbf{x})$. After integrating by parts and enforcing the validity of $\mathcal{L}$ in the subsurface domain Ω, the derivation leads to

$$\left(\nabla \cdot (\lambda_r(\boldsymbol{x})\,\mathcal{U}_r)\right)_{\Omega} = \sum_{e=1}^{N_e} \left(Q^t_{e,r}\alpha_{e,r} - \frac{1}{2\Delta r}\, Q^t_{e,r+1}\beta_{e,r+1} + \frac{1}{2\Delta r}\, Q^t_{e,r-1}\beta_{e,r-1} \right). \tag{A8}$$

The adjoint field $\lambda_r(\mathbf{x})$ backprojects the weighted sum of residual distances between a virtual event and all its counterparts held by $\Phi_{e,r}$ along a ray tube following the group velocity vector $\mathcal{U}_r$ connecting $\mathbf{x}_{e,r}$ to $\mathbf{x}_r$. From the adjoint-state variables, the gradient of the objective function $J(\mathbf{m})$ (eq. A3) with respect to the subsurface parameters (Fig. A1) is straightforwardly obtained by the weighted summation of the adjoint fields λ_r

$$\nabla_{\mathbf{m_{sp}(x)}} J = -\frac{1}{2}\sum_{r=1}^{N_r} \frac{\partial H(\mathbf{x}, \nabla \mathbf{t}_r(\mathbf{x}))}{\partial \mathbf{m_{sp}(x)}} \lambda_r(\mathbf{x}), \tag{A9}$$

and the gradient of the objective function (eq. A1) with respect to the origin time correction parameter and is written as follows for an event

$$\nabla_{\mathbf{m}_{\delta_e}} J = \frac{1}{N_r}\sum_{r=1}^{N_r} \alpha_{e,r}. \tag{A10}$$

Comparing the right-hand side of the adjoint-state equation, eq. (A8), with that of the tomography-inspired formulation, eq. (20), shows that the latter formulation generates a richer kernel through the cross talk between receivers highlighted by the summation over r' in the right-hand side of the adjoint-state equation solved for receiver r. These different kernels indeed result from the fact that one virtual event migrated by one receiver generates only one observable in the first formulation (Fig. 1b), while this virtual event is processed as an excitation term in the second formulation to match the surface measurements recorded by receivers other than the one used to migrate the virtual event, and hence generate as many observables at stations (Fig. 1c).

The Fig. A1 shows the gradient of the objective function, eq. (A1), corresponding to the experiment of Fig. 2. Comparing this gradient with that computed with the tomography-like formulation (Fig. 4) clearly shows that the enriched kernel of the latter formulation better focused the velocity inclusion.

The result of the multiparameter inversion of the toy test obtained with the formulation presented in this appendix is shown in Fig. A2. For this simple test, the results are similar to those obtained with the tomography-like formulation (Fig. 11), although 100 iterations were necessary to reach the convergence point against 76 iterations for the tomography-like formulation. Moreover, we fail to make the time-reversal migration-based formulation work on the more challenging Marmousi case study where the ill-posedness resulting from the lack of illumination induced by the limited number of events was aggravated by the limited coverage provided by the sensitivity kernels of this formulation.