

Multiscale seismic imaging with inverse homogenization

N. Hedjazian¹, Y. Capdeville² and T. Bodin¹

¹Univ Lyon, UCBL, ENSL, UJM, CNRS, LGL-TPE, F-69622, Villeurbanne, France. E-mail: navid.hedjazian@univ-lyon1.fr

²Laboratoire de Planétologie et Géodynamique, UMR CNRS 6112, Université de Nantes, 44300 Nantes, France

Accepted 2021 March 24. Received 2021 March 10; in original form 2020 October 23

SUMMARY

Seismic imaging techniques such as elastic full waveform inversion (FWI) have their spatial resolution limited by the maximum frequency present in the observed waveforms. Scales smaller than a fraction of the minimum wavelength cannot be resolved, and only a smoothed, effective version of the true underlying medium can be recovered. These finite-frequency effects are revealed by the upscaling or homogenization theory of wave propagation. Homogenization aims at computing larger scale effective properties of a medium containing small-scale heterogeneities. We study how this theory can be used in the context of FWI. The seismic imaging problem is broken down in a two-stage multiscale approach. In the first step, called homogenized FWI (HFWI), observed waveforms are inverted for a smooth, fully anisotropic effective medium, that does not contain scales smaller than the shortest wavelength present in the wavefield. The solution being an effective medium, it is difficult to directly interpret it. It requires a second step, called downscaling or inverse homogenization, where the smooth image is used as data, and the goal is to recover small-scale parameters. All the information contained in the observed waveforms is extracted in the HFWI step. The solution of the downscaling step is highly non-unique as many small-scale models may share the same long wavelength effective properties. We therefore rely on the introduction of external *a priori* information, and cast the problem in a Bayesian formulation. The ensemble of potential fine-scale models sharing the same long wavelength effective properties is explored with a Markov chain Monte Carlo algorithm. We illustrate the method with a synthetic cavity detection problem: we search for the position, size and shape of void inclusions in a homogeneous elastic medium, where the size of cavities is smaller than the resolving length of the seismic data. We illustrate the advantages of introducing the homogenization theory at both stages. In HFWI, homogenization acts as a natural regularization helping convergence towards meaningful solution models. Working with fully anisotropic effective media prevents the leakage of anisotropy induced by the fine scales into isotropic macroparameters estimates. In the downscaling step, the forward theory is the homogenization itself. It is computationally cheap, allowing us to consider geological models with more complexity (e.g. including discontinuities) and use stochastic inversion techniques.

Key words: Inverse theory; Tomography; Waveform inversion; Computational seismology; Seismic anisotropy.

1 INTRODUCTION

Seismic tomography aims at quantifying elastic properties in the Earth's interior using seismic data measured at the surface. It can be formulated as an inverse problem, where unknown parameters defining the earth model are sought in order to replicate the observed data. Here we consider the case of full waveform inversion (FWI), where the seismic data is the seismogram waveform. The primary applications range from the exploration scale to image the subsurface, up to the global scale to study the Earth's deep interior

(Tarantola 1984; Brossier *et al.* 2009; Tape *et al.* 2010; Fichtner 2010; Bozdağ *et al.* 2016; Virieux *et al.* 2017), while applications at the engineering scale or in medical imaging start to develop (e.g. Bernard *et al.* 2017; Nguyen & Modrak 2018). Replicating the observations requires to compute synthetic seismograms, which constitutes the 'forward problem'. For example, waveform modelling can be performed using approximated methods, based on ray theory, normal mode summation or on the Born approximation (Woodhouse & Dziewonski 1984; Devaney 1984; Li & Romanowicz 1996), or more recently by numerically solving the full wave

equation, leading to FWI methods (Fichtner 2010; Virieux *et al.* 2017). In the standard approach, the misfit between simulated and observed data is minimized using a local optimization strategy based on data derivatives. Global parameter search methods are still out of reach for most real size 3-D problems (Käufel *et al.* 2013). Once the misfit reaches a minimum, hopefully a global minimum, we obtain a quantitative image of the material elastic properties to interpret in terms of geological structures.

However, most waveform tomography problems are ill-posed, and their solution is potentially non-unique. Many strategies have been developed to ensure convergence towards a meaningful solution. These include, among others, building a good starting model with alternative imaging methods (such as traveltime tomography), various regularization methods in the parameter space (smoothing, prior information), multiscale imaging by gradually increasing the frequency content of the waveforms (Bunks *et al.* 1995; Brossier *et al.* 2009) or constructing more sophisticated misfit functionals (e.g. Bozdağ *et al.* 2011; Métivier *et al.* 2016). As a result, tomography, and in particular its interpretation, requires a lot of expertise. A recurrent discussion is to distinguish the features in the solution model that are properly resolved and required by the data, from those introduced by the specificity of the implementation, such as an initial model or an *a priori* constraint on the model parameters (promoting smoothness or sparsity, imposing discontinuities or symmetries in the elastic tensor, etc.). Also, compromises have often to be made between the fit to the data and the realism of the image, or between model complexity and model constraint.

Some of these aspects originate from two fundamental limitations. (i) Due to instrumental limitations or to the high computational cost of waveform modelling, the seismic data are restricted to a finite frequency range. (ii) The imaged medium are usually multiscale, but seismic inversions attempt to recover only a finite, limited number of model parameters. Accordingly, seismic imaging has difficulties to identify the different scales present in the medium without *a priori* information. A common understanding approximates that the spatial resolution of a tomographic image is, at best, the smallest considered wavelength.

We take interest in imaging problems involving high contrast, small-scale heterogeneities, especially with discontinuities in elastic properties. This comprises, for example, imaging the subsurface to detect faults, fluids, karsts or tunnels, and also non-destructive testing to detect flaws or rebars in concrete at the engineering scale. To address these problems, some possible methods are based on single-scattering formulations (e.g. Rose 1989), or are based on migration using time reversal and an imaging condition (Almuhaidib & Toksöz 2015; Kocur *et al.* 2016). In particular, this study deals with a cavity detection problem. This is usually performed with *P*- or Rayleigh-wave diffraction imaging (Grandjean & Leparoux 2004). These methods usually require complex processing of the data, to separate the different waves that tend to arrive simultaneously, or to handle multiple scattering. Hence FWI methods, in both acoustic and elastic regimes, have been investigated (Bretaudau *et al.* 2013; Tran *et al.* 2013; Seidl & Rank 2016; Nguyen & Modrak 2018). However, imaging such structures requires a fine-scale description of the medium, which makes the computational cost of modelling the full wavefield prohibitive.

An appropriate tool to mitigate these issues is homogenization or upscaling of the elastic wave equation. It refers to an ensemble of methods able to compute a long-wavelength equivalent medium to a fine scale elastic medium. This effective medium does not contain heterogeneities smaller than a given minimum wavelength,

but produces the same wavefield as the original medium in a limited frequency band. The historical example is the homogenization of stratified media: an explicit formula was derived by Backus (1962), who showed that a stack of horizontal isotropic layers will be equivalent, for long period waves, to a homogeneous hexagonal anisotropic medium with vertical axis of symmetry. The initial idea behind the homogenization theory is to reduce the computational cost of waveform modelling in a complex medium. By calculating a smooth equivalent medium, it avoids the need of solving the wave equation on a mesh containing all the fine scales and the possible discontinuities. In this work, we proceed further and study how homogenization theory can facilitate the inverse problem in seismic imaging. This tool being developed for elastic media, we restrict ourselves to this case in the following, but an equivalent for the acoustic case could be considered.

In the context of elastic FWI, Backus theory can be used to constrain the solution space to effective models (Capdeville *et al.* 2013), or to ensure that fine-scale model updates preserve long-wavelength properties of the model (Afanasyev *et al.* 2016). Here, we will consider the more general non-periodic homogenization (Capdeville *et al.* 2010; Guillot *et al.* 2010), a technique able to compute the effective elastic properties of any given medium, without the need for spatial periodicity, scale separation or statistical invariance. It enables to properly pose FWI as a multiscale inverse problem (Fichtner *et al.* 2013b). This is especially important for attenuation and anisotropy that are scale-dependent properties. Similar ideas are present in the field of porous media and subsurface flow, where mostly elliptic equations are considered. For example, one multiscale inverse problem is to determine the medium properties (porosity, permeability or conductivity) when large microscale fluctuations are present. Homogenization theory helps at various levels (regularization, uncertainty estimation), and a natural procedure is to search for an effective model solution of the inverse problem (Nolen *et al.* 2012).

We propose to solve a seismic imaging problem with the following method (Fig. 1): we first perform a waveform inversion where we aim to recover a macroscale effective (or homogenized) elastic medium, based on the homogenized FWI (HFWI) method developed by Capdeville & Métivier (2018). Hence, HFWI differs from conventional FWI in the definition of the inverted model: it is fully anisotropic with a wavefield-adapted discretization, and it is homogenized after the last iteration (or, if required for helping convergence, after each iteration). Next, we define a second inverse problem where the obtained macroscale solution plays the role of the data and the homogenization operator the role of the forward model. That is, we search for microscale models having an effective medium close to the macroscale solution. HFWI and subsequent inverse homogenization can therefore be seen as a two-stage resolution method of a multiscale inverse problem (Frederick & Engquist 2017).

Let us first synthesize some elements and results of the HFWI method. For all natural media considered in seismic imaging, it is possible to define a minimum wavelength λ_{\min} , associated with the maximum frequency of the waveform data. The spatial wavelength under which structures cannot be resolved by FWI is directly related to λ_{\min} . We define a constant wavelength λ_0 , setting the separation between the fine and the large scales. The objective of HFWI is to recover, from the waveform data, an effective model at scales greater than λ_0 . Knowing the diffraction resolution limit of a FWI is about $\lambda_{\min}/2$ (e.g. Huang & Schuster 2014), a good choice is $\lambda_0 \geq \lambda_{\min}/2$, because smaller scales are poorly resolved by the data. Capdeville & Métivier (2018) consider numerical examples of a

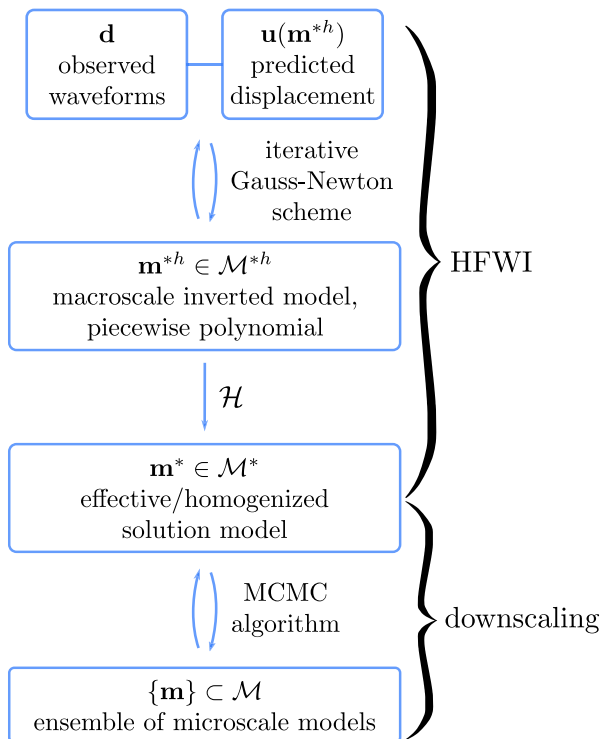


Figure 1. Notations used for the two-stage method proposed here. We invert the waveform data d for a macroscale effective medium \mathbf{m}^* . \mathbf{m}^{*h} is an intermediate product of HFWI required by the discretization of the problem. Next, we seek for microscale models \mathbf{m} having an effective medium close to \mathbf{m}^* . \mathcal{H} designates the homogenization operator.

FWI using either isotropic or anisotropic parameters, adapting the space discretization to maintain identical number of free parameters. They show that the FWI solution may vary depending on the chosen parametrization. However, these different solutions have the same homogenized elastic properties for a scale separation position $\lambda_0 = \lambda_{\min}$ and are in good agreement for a scale separation at $\lambda_0 = \lambda_{\min}/2$. Hence, introducing homogenization in the FWI problem leads to the following observation: while FWI is parametrization and mesh dependent, HFWI is not. If we assume that the real, multiscale Earth is solution of a given FWI problem (this may not be true due to the imperfection of the inversion algorithm), its effective elastic properties are also a solution. As the real Earth is inaccessible, Capdeville & Métivier (2018) suggest to seek for an effective medium and thereby reduce the non-uniqueness of the FWI solution.

By definition, the HFWI solution model is smooth and presents apparent anisotropy. Because of this spatial simplicity, it is appropriate for fast data prediction. It can already prove useful in this form, for example regarding source localization, but it is poorly suited to geological interpretation. For this purpose, we are required to solve a second problem, the downscaling or inverse homogenization, which is the main subject of this study. Some previous examples of inverse homogenization examine the general case of elliptic equations (Hoang & Quek 2019), or, in the FWI context, porosity using seismic attenuation (Dupuy *et al.* 2016). In the 1-D case, Bodin *et al.* (2015) proposed a Bayesian inverse homogenization strategy, where the ensemble of 1-D fine-scale elastic models equivalent to a tomographic profile is explored.

In this study, we show how the effective anisotropy recovered by the HFWI solution model can be used to resolve small-scale

isotropic heterogeneities. The forward operator in the inverse homogenization problem is the non-periodic elastic homogenization (Capdeville *et al.* 2010; Guillot *et al.* 2010). Thus, no hypothesis such as periodicity or stochastic invariance is required on the fine-scale medium. The misfit is defined by a distance between two effective elastic media: the HFWI solution (taking the role of the observations) and the effective medium of the fine-scale model (the predictions). The space of fine-scale models is defined with an object-based approach to reduce the number of inference parameters. Because the solution could be non-unique, we use a Bayesian formalism, where an ensemble solution is sampled with a Markov chain Monte Carlo (MCMC) algorithm, thus allowing us to estimate uncertainties.

We test the method in the case of a 2-D homogeneous medium containing cavities. The objective is to recover the position, size and shape of the cavities. The setup is representative of near-surface or engineering problems, with high wave frequencies and short acquisition times. The rationale of this choice is also the prospect to validate the methodology using reduced-scale experiments in the laboratory (Pratt 1999; Pageot *et al.* 2017).

We observe the following advantages to split the multiscale inversion into two steps: in the first step, recovering an homogenized elastic medium is in general a better-posed problem than conventional FWI, because the space of acceptable elastic models is contained. We only need to search for smooth solution models, the effective anisotropy contains the information on the small-scale structure seen in the waveform data. This lowers the computational cost of wave propagation if the true medium contains fine-scale heterogeneities or discontinuities. It also mitigates artifacts induced by incorrect parameter restrictions. The level of smoothing is naturally given by the wavelength of homogenization. We introduce as little prior information as possible at this stage, as it would require more advanced mathematical treatments of wave propagation and usually greater computational cost. The second step, the inverse homogenization, can be less well-posed. But it presents the great advantage to having a much cheaper forward model, which makes possible the use of statistical or global search approaches. For example, using a Bayesian framework facilitates the incorporation of *a priori* information. Furthermore, the data is now ‘localized’: the effective properties at one location depend only on the fine-scale properties in the neighbourhood. This potentially allows to solve the problem using a large ensemble of tools that are prohibited in waveform inversion (e.g. Nawaz & Curtis 2016).

2 HOMOGENIZATION THEORY

This section summarizes the elastic homogenization theory used in this study. More detailed descriptions can be found in Capdeville *et al.* (2010) and Guillot *et al.* (2010). We define the ensemble of elastic media \mathcal{M} . Each $\mathbf{m} \in \mathcal{M}$ is a vector of microscale parameters characterized by its density ρ and its elastic tensor \mathbf{c} for every position in the medium. In the following, \mathcal{M}^* is the space of effective media, and effective quantities are noted with a^* .

2.1 Homogenized wave equation

Homogenization theory refers to an ensemble of methods that aims at ‘averaging’ the fine scale properties of a heterogeneous medium, that is computing its effective properties at larger scale in a mathematically rigorous and consistent way for a given set of equations.

Such methods have been widely developed for static elastic problems for media with an explicit separation of scales, in particular for periodic media (e.g. Bensoussan *et al.* 2011). In the context of wave propagation, elastic heterogeneities smaller than the minimum wavelength are expected to only have an effective impact on the resulting wavefield. Hence, for solving the wave equation, the fine-scale medium is beneficially replaced by a long wavelength effective medium.

In the case of periodic media, of period λ for example, the effective medium and equations can be obtained with the two-scale homogenization method (e.g. Sánchez-Palencia 1980). It is an asymptotic method based on the following small parameter:

$$\varepsilon = \frac{\lambda}{\lambda_{\min}}, \quad (1)$$

where λ_{\min} is the minimum wavelength of the wavefield. It is possible to find the effective medium and more by solving a series of partial differential equations (called the cell problem), most of the time numerically. It can be mathematically shown that the true solution of the wave equation weakly converge towards the effective solution as ε tends towards zero.

For most geophysical media, the periodic assumption is not valid and the extension of the two-scale homogenization to the non-periodic case is not trivial. Indeed, there exist no natural scale separation λ_0 in the elastic medium. To solve this difficulty, a solution based on the assumption that a minimum wavelength $\lambda_{\min} = V_{S, \min}/f_{\max}$ exists, associated with the maximum frequency contained in the waveform f_{\max} and the slowest velocity in the medium $V_{S, \min}$, has been developed (Capdeville & Marigo 2007; Capdeville *et al.* 2010). In this method, the separation of scale is not performed with respect to the medium but to λ_{\min} . We therefore introduce λ_0 as a user-defined parameter that sets the separation between scales considered as ‘fine’ (or microscopic) and ‘large’ (macroscopic). The value of λ_0 simply determines the minimum length-scale that will be preserved in the homogenized medium, and it does not vary spatially. A smaller λ_0 corresponds to a more detailed structure thus a more accurate solution for the effective displacement \mathbf{u}^* , whereas a larger λ_0 involves a smoother medium and imprecise solution. We can introduce:

$$\varepsilon_0 = \frac{\lambda_0}{\lambda_{\min}}, \quad (2)$$

which measures the scale separation position in the seismic model.

We usually consider that the wavefield ‘sees’ the medium at the scale $\lambda_{\min}/2$, that is $\varepsilon_0 = 0.5$. Medium heterogeneities of scale $\lambda_0 \ll \lambda_{\min}$ are seen by the wavefield only through their homogenized/effective properties. Nevertheless, for some strongly scattering media, it can happen that an ε_0 smaller, or even much smaller, than 0.5 might be required to model accurately a long coda wave. We refer to Capdeville *et al.* (2010) for numerical examples.

In the inversion context, that is recovering the elastic medium properties knowing an observed displacement \mathbf{u} at the surface, the image reconstruction becomes challenging when approaching this $\lambda_{\min}/2$ FWI resolution limit. Thus a value $\varepsilon_0 \geq 0.5$ might be preferable in most cases (for a more comprehensive discussion, see Capdeville & Métivier 2018). The value of $\varepsilon_0 \sim 0.5$ is chosen in this study.

Homogenization theory introduces two space variables, macroscopic variations with the variable \mathbf{x} , and microscopic variations with $\frac{\mathbf{x}}{\varepsilon_0}$, and assumes they can be treated as independent variables. A physical quantity depending on macroscopic and microscopic variations and on time, such as displacement, is written $\mathbf{u}(\mathbf{x}, \frac{\mathbf{x}}{\varepsilon_0}, t)$.

The homogenized solution of the wave equation with a scale separation position ε_0 is the leading order of the displacement asymptotic expansion and is written \mathbf{u}^* . It can be demonstrated to be independent of the small scale variable $\frac{\mathbf{x}}{\varepsilon_0}$. The asymptotic expansion to the first order of the true displacement is:

$$\mathbf{u}(\mathbf{x}, \frac{\mathbf{x}}{\varepsilon_0}, t) = \mathbf{u}^*(\mathbf{x}, t) + \varepsilon_0 \boldsymbol{\chi}(\mathbf{x}, \frac{\mathbf{x}}{\varepsilon_0}) : \boldsymbol{\varepsilon}(\mathbf{u}^*)(\mathbf{x}, t) + O(\varepsilon_0). \quad (3)$$

$\boldsymbol{\chi}$ is a third order tensor called the first-order corrector, it does not depend on time thus needs only to be known at receiver positions. The double dot product notation corresponds to $[\boldsymbol{\chi} : \boldsymbol{\varepsilon}]_i = \chi_{ijk} \varepsilon_{jk}$ with Einstein summation convention. $[\boldsymbol{\varepsilon}(\mathbf{u}^*)]_{jk} = \frac{1}{2}(\partial_j u_k^* + \partial_k u_j^*)$ is the strain tensor of the effective displacement. In practice, eq. (3) is most of the time in $O(\varepsilon_0^2)$.

One can show that the effective displacement \mathbf{u}^* is the solution of the effective wave equation:

$$\begin{aligned} \rho^* \ddot{\mathbf{u}}^* - \nabla \cdot \boldsymbol{\sigma}^* &= \mathbf{f}^* \\ \boldsymbol{\sigma}^* &= \mathbf{c}^* : \boldsymbol{\varepsilon}(\mathbf{u}^*), \end{aligned} \quad (4)$$

ρ^* , \mathbf{c}^* , $\boldsymbol{\sigma}^*$ and \mathbf{f}^* are the effective density, elastic tensor, stress tensor and source term respectively. This equation is subject to effective boundary conditions, which are not explicit here (see Capdeville & Marigo (2012) for the treatment of boundary conditions). This equation is analogous to the classical wave equation except that true quantities have been replaced by effective quantities. In this study, we consider a region containing neither sources nor receivers, thus we also omit the treatment of the effective source \mathbf{f}^* and the first order corrector $\boldsymbol{\chi}$. Hence, the effective elastic medium is uniquely determined by its effective density and elastic tensor. Once ρ^* and \mathbf{c}^* have been determined (see the next section), one can solve the eq. (4) using a standard wave equation solver. The accuracy of approximating \mathbf{u} by \mathbf{u}^* is given by eq. (3): solving the wave equation in the media \mathbf{m} or \mathbf{m}^* produces identical displacements up to the first order in ε_0 .

2.2 The Homogenization operator

To compute the effective medium properties, we introduce a low-pass filter operator noted \mathcal{F}^{λ_0} (see Guillot *et al.* (2010) for its exact analytic definition). Any spatial field filtered by \mathcal{F}^{λ_0} does not contain spatial variations smaller than λ_0 . A ‘naive’ averaging or homogenization would be to simply filter the density ρ and the elastic parameters \mathbf{c} . However, this ‘naive’ solution as well as any other simple option such as filtering the compliance tensor (the inverse of the elastic tensor) or directly the seismic velocities are leading to poorly accurate solutions (Capdeville *et al.* 2010). In the case of a layered medium, Backus (1962) showed that the correct effective medium is obtained by filtering non-linear combinations of the fine-scale elastic parameters. In the general case, there exist no analytic solution to the effective properties. Using homogenization theory, they are obtained following the procedure described hereinafter.

We define effective properties of an elastic medium $\mathbf{m} \in \mathcal{M}$ with a scale separation at ε_0 as $\mathbf{m}^* = (\rho^*, \mathbf{c}^*) \in \mathcal{M}^*$. \mathbf{m}^* is spatially smooth, it does not contain wavelengths smaller than λ_0 . Thus the space of effective media \mathcal{M}^* is a finite dimensional space. Its dimension is proportional to $(f_{\max}/\varepsilon_0)^d$ with d the dimension (2-D or 3-D) of the problem. In the following, the relation between the fine and large scales is summarized by the homogenization operator \mathcal{H} :

$$(\rho^*, \mathbf{c}^*) = \mathcal{H}(\rho, \mathbf{c}). \quad (5)$$

The operator \mathcal{H} is non-linear and implicitly depends on ε_0 and λ_{\min} . In practice, the procedure to compute \mathcal{H} comprises following steps:

(i) Find the initial guess corrector χ_s^{lm} . We solve an elasto-static equation called the cell-problem in the homogenization vocabulary:

$$\nabla \cdot \mathbf{c} : \epsilon(\chi_s^{lm}) = -\nabla \cdot (\mathbf{c} : (\mathbf{e}_i \otimes \mathbf{e}_m)) \quad (6)$$

with periodic boundary conditions, and $\mathbf{e}_i, i \in \{1, \dots, d\}$ the Cartesian unit vector along dimension i .

(ii) Compute the effective density and elastic tensor:

$$\rho^*(\mathbf{x}) = \mathcal{F}^{\lambda_0}(\rho)(\mathbf{x}), \quad (7)$$

$$\mathbf{c}^*(\mathbf{x}) = \mathcal{F}^{\lambda_0}(\mathbf{H}_s) : \mathcal{F}^{\lambda_0}(\mathbf{G}_s)^{-1}(\mathbf{x}), \quad (8)$$

with $(\mathbf{G}_s)_{ijkl} = \frac{1}{2}(\delta_{il}\delta_{jm} + \delta_{jl}\delta_{im}) + (\epsilon(\chi_s^{lm}))_{ij}$ and $\mathbf{H}_s = \mathbf{c} : \mathbf{G}_s$.

First step involves the resolution of an elasto-static problem, which we solve numerically with a the finite element method (Capdeville *et al.* 2010). Second step involves low-pass filtering with operator \mathcal{F}^{λ_0} . This implies that the inverse problem (downscaling, i.e. recovering small-scale properties knowing the large-scale ones) is ill-posed.

3 HOMOGENIZED FULL WAVEFORM INVERSION

We now present the elements of the HFWI used in this study (Capdeville & Métivier 2018). We invert seismic waveforms \mathbf{d} for a macroscale effective medium \mathbf{m}^* . The interest of the HFWI method is that a relation can be established between its solution and the true Earth model \mathbf{m}_t :

$$\mathbf{m}^* \approx \mathcal{H}(\mathbf{m}_t), \quad (9)$$

where \mathcal{H} is the homogenization operator defined above. We will then use this equation to define the inverse homogenization problem.

3.1 Parametrization in the FWI

We consider the inverse problem of FWI. In general practice, the solution space is defined with a restricted number of mechanical parameters (for example, only P -wave velocity or only isotropic parameters) and with a spatial discretization that might be different from the wave equation solver mesh. Here, we chose this solution space as the ensemble of effective elastic models \mathcal{M}^* , the image of \mathcal{M} through the operator \mathcal{H} , for which the parametrization is fully anisotropic (six independent elastic parameters in 2-D, and 21 parameters in 3-D).

In the layered media case, the homogenization problem has an analytic solution, thus the effective medium can be explicitly parametrized with the Backus parameter vector (Capdeville *et al.* 2013). In the general case however, no explicit parametrization of the effective medium space \mathcal{M}^* is known yet. This means, if we suppose having a model $\mathbf{m}_i^* \in \mathcal{M}^*$ at iteration i , the updated model \mathbf{m}_{i+1}^* obtained with an iterative optimization algorithm is not necessarily in \mathcal{M}^* . Hence, in practice, we rely on an approximate finite dimension model space \mathcal{M}^{*h} , wide enough to capture the information in the waveform data, but such that $\mathcal{H}(\mathcal{M}^{*h}) \subset \mathcal{M}^*$ (Fig. 1).

We define a model in \mathcal{M}^{*h} with the notation \mathbf{m}^{*h} . We chose to parametrize such model with a density ρ^{*h} and the full elastic tensor \mathbf{c}^{*h} on a regular mesh, each cell containing a polynomial approximation of degree N_h in each direction. No continuity in

elastic properties is imposed between the elements. In contrary to \mathbf{m}^* , \mathbf{m}^{*h} may not be smooth and we may have $\mathbf{m}^{*h} \notin \mathcal{M}^*$.

To ensure that the final model \mathbf{m}_f^* is in \mathcal{M}^* , two strategies are possible. A first one is to project $\mathbf{m}_{i+1}^{*h} \in \mathcal{M}^{*h}$ into $\mathbf{m}_{i+1}^* \in \mathcal{M}^*$ after each iteration, using \mathcal{H} , and compute the next model $i+2$ from \mathbf{m}_{i+1}^* . This method acts as a natural regularization of the problem. A second possibility is to compute each model iteration in the space \mathcal{M}^{*h} and project only the final model \mathbf{m}_f^{*h} into $\mathbf{m}_f^* \in \mathcal{M}^*$. We rely on the second method in all the inversions performed in this study, because the regularization at each step was not required to make the algorithm converge for our examples.

3.2 Waveform inversion strategy

Waveform observations $\mathbf{d}_s(\mathbf{x}_r, t)$ originating from sources s are measured at receivers positions \mathbf{x}_r . We replicate the waveform data by solving the elastic wave equation and extracting the displacement $\mathbf{u}_s(\mathbf{x}_r, t)$ at the receiver locations. Waveform modelling is performed with a 2-D spectral element solver (Komatitsch & Vilotte 1998).

For a fine-scale model $\mathbf{m} = (\rho(\mathbf{x}), \mathbf{c}(\mathbf{x}))$, the misfit function is defined as a least-squares distance between \mathbf{d} and \mathbf{u} :

$$E(\mathbf{m}) = \sum_{r,s} \int_0^T (\mathbf{d}_s(\mathbf{x}_r, t) - \mathbf{u}_s(\mathbf{x}_r, t; \mathbf{m}))^2 dt. \quad (10)$$

Similar equations E^{*h} and E^* can be written by replacing \mathbf{m} with \mathbf{m}^{*h} or \mathbf{m}^* if the model is in \mathcal{M}^{*h} or is projected into \mathcal{M}^* at each iteration. \mathbf{m}^{*h} will be used in all the tests presented hereinafter.

The misfit function is minimized using a standard damped Gauss–Newton iterative scheme. The updated model is defined as:

$$\mathbf{m}^{i+1} = \mathbf{m}^i + ((\mathbf{F}^i)^T \mathbf{F}^i + \lambda^i \mathbf{I})^{-1} (\mathbf{F}^i)^T (\mathbf{d} - \mathbf{u}(\mathbf{m}^i)), \quad (11)$$

where \mathbf{I} is the identity matrix. The Fréchet derivatives \mathbf{F}^i of the displacement \mathbf{u} with respect to model parameters \mathbf{m}^i are calculated with the adjoint state method (Tarantola 1984; Plessix 2006; Tromp *et al.* 2008). The approximate Hessian $(\mathbf{F}^i)^T \mathbf{F}^i$ is damped by a parameter λ^i which value decreases with iterations. Once the algorithm has converged, the final model \mathbf{m}^{*h} is projected into $\mathbf{m}^* = (\rho^*, \mathbf{c}^*) \in \mathcal{M}^*$ using the homogenization operator.

4 INVERSE HOMOGENIZATION

The term ‘downscaling’ is mostly used in the field of meteorology and defines the procedure of inferring microscale (or high resolution) information from macroscale (or low-resolution) variables. In this study, the upscaling operator is \mathcal{H} , and hence the downscaling is called inverse homogenization. The problem consists in recovering fine-scale information on an Earth model, knowing an estimation of its effective elastic properties \mathbf{m}^* . Typically, \mathbf{m}^* is the result of the HFWI method presented in the previous section. It allegedly contains all the information that could be retrieved from the waveforms. To further learn something about the true Earth \mathbf{m}_t , incorporation of *a priori* information is required. This can include additional data from a non-seismic origin, some known geological features, spatial statistical properties of the rocks, etc. We aim at solving the inverse homogenization using a Bayesian formulation, where the *a priori* information is provided as a probability distribution (e.g. Kaipio & Somersalo 2006). Thus the HFWI and inverse homogenization process is particularly suited to incorporate constraints from geostatistical models.

4.1 Problem setting

Depending on the level of *a priori* constraints, the solution of the downscaling problem might be highly non-unique, as many small-scale models can have the same effective medium. Therefore, we cast the problem in a Bayesian framework, where we look for an ensemble of plausible solutions. It is described by an additive noise model, with the following equation:

$$\mathbf{c}^* = \mathcal{H}(\mathbf{c}) + \epsilon, \quad (12)$$

understood in a statistical sense. \mathbf{c}^* , defining the observed data, is the effective elastic tensor of \mathbf{m}^* obtained by HFWI. \mathbf{c} is the elastic tensor of a microscopic scale model \mathbf{m} . In this equation, the models are described only with their elastic tensor because we did not consider the density in the inverse homogenization. ϵ is a random variable characterizing the mismatch between the solution of the HFWI \mathbf{m}^* and the effective properties of the true earth \mathbf{m}_1^* , i.e. the meaning of \approx in eq. (9). The first source of mismatch is the data uncertainty, the error attributable to the approximate estimation of \mathbf{c}^* in the HFWI. The second source is the theoretical errors of the operator \mathcal{H} . We neglect this source of errors in the following, expecting that in practice, it should be orders of magnitude smaller than the first one if computed correctly.

Properly defining the inverse homogenization with a Bayesian formulation amounts to obtain an accurate idea of the distribution for ϵ . An assumption on its structure is required. We follow the general practice and suppose ϵ to be multivariate-Gaussian distributed with zero mean and covariance matrix Σ . These assumptions are discussed hereinafter and in the synthetic tests presented later on.

The distribution for ϵ represents our knowledge on the all the errors introduced when estimating \mathbf{c}^* by HFWI. Errors in the HFWI solution may have multiple sources, such as: noise on the waveform data, lack of data coverage, damping, convergence towards a local minima, etc. Solving the FWI using a global search is not achievable in real scale problems. They are solved with a local optimization strategy (here the Gauss–Newton method) returning a unique ‘best-fitting’ solution model. Therefore, HFWI shares the same difficulties as any other FWI methods to find a solution close to the global minimum. Still, because HFWI reduces the space of solution models to the effective ones, it at least mitigates the ill-posedness induced by a poor parametrization (Capdeville & Métivier 2018).

If the global minimum is achieved, we can estimate ‘how far’ \mathbf{m}^* is from \mathbf{m}_1^* by computing local uncertainties associated to this solution. This is a widespread subject of research in seismic tomography. Broadly speaking, local uncertainties are usually characterized by a covariance matrix Σ . Methods developed in the FWI context to estimate Σ rely on low-rank matrix approximations (Bui-Thanh *et al.* 2013; Thurin *et al.* 2019). Σ is strongly linked with the Hessian matrix (Fichtner & Trampert 2011). As pointed out by Bui-Thanh *et al.* (2013), the Gauss-Newton part of the Hessian $\mathbf{F}^T \mathbf{F}$ is a good approximation of the full Hessian, especially if the medium is smooth. It could thus be an appropriate choice to estimate Σ . However, in the HFWI formulation used here, we obtain $\mathbf{F}^T \mathbf{F}$ for the \mathcal{M}^{*h} space, while we look for uncertainties in the \mathcal{M}^* space. A proper estimation of Σ , for example by upscaling of the Hessian matrix into \mathcal{M}^* , will be the subject of future work.

4.2 Bayesian inference

The solution of the Bayesian inverse problem is described by the posterior probability density of the model parameters knowing the

observations $\pi(\mathbf{c}|\mathbf{c}^*)$. Using the Bayes formula, we write the posterior as (Kaipio & Somersalo 2006):

$$\pi(\mathbf{c}|\mathbf{c}^*) \propto \pi_{\text{pr}}(\mathbf{c}) \exp(-\phi(\mathbf{c})), \quad (13)$$

where $\pi_{\text{pr}}(\mathbf{c})$ is the prior distribution on \mathbf{c} , and $\phi(\mathbf{c})$ the negative log-likelihood (the Bayesian equivalent of the misfit function).

4.3 Likelihood function

Using the Gaussian additive noise model of eq. (12), the negative log-likelihood is:

$$\phi(\mathbf{c}) = \frac{1}{2}(\mathbf{c}^* - \mathcal{H}(\mathbf{c}))^T \Sigma^{-1}(\mathbf{c}^* - \mathcal{H}(\mathbf{c})), \quad (14)$$

where \mathbf{c} is expressed as a data vector. In the following 2-D examples, this corresponds to the parameters (c_{1111} , c_{1122} , $\sqrt{2}c_{1112}$, c_{2222} , $\sqrt{2}c_{2212}$, $2c_{1212}$) to take into account the symmetries in the elastic tensor.

The choice of Σ^{-1} is crucial in estimating the posterior uncertainties on the model parameters. To ensure that our estimate $\hat{\Sigma}^{-1}$ is reasonable, we only estimate relative uncertainties $\hat{\Sigma}_r$, and introduce a scaling parameter h setting $\Sigma = h^2 \hat{\Sigma}_r$, where h will be treated as an unknown variable in the problem, following the so-called hierarchical Bayes approach (Gelman *et al.* 2013). The role of h is to account for a mis-estimation in the amplitude of errors. An inferred value of h close to 1 shows that the estimation $\hat{\Sigma}_r$ given by the user is good. A value $h > 1$ indicate that some additional parts of the data are not explained by our modelling, due for example to theoretical errors. In this case, the negative log-likelihood becomes:

$$\phi(\mathbf{c}, h) = N \log(h) + \frac{1}{2}(\mathbf{c}^* - \mathcal{H}(\mathbf{c}))^T (h^2 \hat{\Sigma}_r)^{-1}(\mathbf{c}^* - \mathcal{H}(\mathbf{c})), \quad (15)$$

where both \mathbf{c} and h are unknown to be solved for.

In order to compute ϕ , the next step is to estimate the inverse covariance matrix $\hat{\Sigma}_r^{-1}$ of the probability distribution associated to ϵ .

For a solution space of dimension N , we need to estimate the $N \times N$ parameters. For now, we rely on a sampling-based estimation of $\hat{\Sigma}^{-1}$ similar to the approach in Hansen *et al.* (2014). It consists of first generating a large ensemble of error realizations, from which a Gaussian model is inferred. We generate M noise realizations on the waveform data \mathbf{d} and, for each, estimate a HFWI solution. As such, we only consider errors associated to the noise in waveforms and mapped to the HFWI solution through the tomographic inversion. First, we calculate the empirical covariance matrix S of the M HFWI solution models. In order to get a stable inverse of S , a regularized version $\hat{\Sigma}_r$ is computed using the shrinkage formula (Ledoit & Wolf 2004):

$$\hat{\Sigma}_r = (1 - \alpha)S + \alpha \frac{\text{Tr } S}{N} \mathbf{I}, \quad (16)$$

where \mathbf{I} is the identity matrix. It represents an optimal linear combination of S and \mathbf{I} to minimize the mean squared error between $\hat{\Sigma}$ and the true covariance. The formula reduces the ratio between the maximum and minimum eigenvalues while preserving the mean of all eigenvalues, thus a sufficiently large α ensures that $\hat{\Sigma}_r$ is well-conditioned. The parameter α can also be interpreted as a trade-off between bias and variance for the estimator $\hat{\Sigma}_r$. We chose the value of α adapted to Gaussian distributed data (Chen *et al.* 2010).

4.4 Prior information

The prior distribution π_{pr} in eq. (13) encodes the information known about the model \mathbf{m} before the inversion. It is therefore widely dependent on the tackled problem. Basically, any geostatistical model could be used for π_{pr} . The Bayesian formulation permits a large range of applications for the inverse homogenization.

Prior information is enforced in two different ways: (1) by choosing a specific parametrization of the model using geometrical objects and (2) by choosing a relevant prior probability distribution on these parameters. Translated into notations, the elastic model \mathbf{c} is constructed from a vector \mathbf{Z} of n latent variables. This process can be described as a function $\mathbf{c}(\mathbf{Z})$. Subsequently, a prior distribution is imposed directly on \mathbf{Z} . Hence, we will write the prior distribution $\pi_{\text{pr}}(\mathbf{Z})$ and keep a similar notation for the other probability distributions described hereinafter.

Because this study deals only with synthetic tests, construction of realistic priors is out of scope. We will consider a somewhat idealized prior distribution, using an object-based parametrization made of ellipses. The vector \mathbf{Z} contains parameters describing the position and shape of each object, as well as the hyperparameter h , the unknown scaling factor of the covariance matrix. For the problem at hand, the ellipses are representing cavities. They are described by their position (x^i, y^i) , the long and short axes (r_1^i, r_2^i) and the angle from horizontal α^i for each ellipse $1 \leq i \leq n_e$. All parameters are independent. We use a uniform prior distribution for the object parameters, and a Jeffrey prior for h (i.e. proportional to $1/h$).

4.5 Sampling algorithm

The posterior distribution is sampled using the reversible-jump MCMC (rj-MCMC) sampler (Green 1995). Rj-MCMC is a trans-dimensional sampler, where the number of parameters (here the number of elliptical cavities) is treated as an unknown in the problem. In our case, it amounts to consider a varying number n_e of objects. Transdimensional inversions are popular in seismic imaging, as the spatial discretization of the solution space is adapted to data quality and coverage (Sambridge *et al.* 2013). The present parametrization allows to make some simplifications in the algorithm, a more complete description of the rj-MCMC algorithm for seismic imaging is outlined in Bodin & Sambridge (2009).

MCMC algorithms aim at generating samples from a target probability distribution. In a Bayesian inverse problem, we want to sample from the posterior distribution π of eq. (13). At each step of the Markov chain, a new sample \mathbf{Z}' is generated as a small random deviate from the previous parameter vector \mathbf{Z} , according to a proposal distribution noted $q(\mathbf{Z}'|\mathbf{Z})$. From this proposed model \mathbf{Z}' , we construct the elastic tensor \mathbf{c}' , and write its corresponding prior and negative log-likelihood $\pi_{\text{pr}}(\mathbf{Z}')$ and $\phi(\mathbf{Z}')$. The Markov chain converges to the posterior distribution if the acceptance probability of \mathbf{Z}' is:

$$\alpha_{\text{accept}}(\mathbf{Z}'|\mathbf{Z}) = \min \left\{ 1, \frac{\pi_{\text{pr}}(\mathbf{Z}')q(\mathbf{Z}|\mathbf{Z}')\pi(\mathbf{Z}'|\mathbf{c}^*)}{\pi_{\text{pr}}(\mathbf{Z})q(\mathbf{Z}'|\mathbf{Z})\pi(\mathbf{Z}|\mathbf{c}^*)} \times |\mathbf{J}| \right\} \quad (17)$$

\mathbf{J} is the Jacobian matrix of the transformation from \mathbf{Z} to \mathbf{Z}' . Its determinant $|\mathbf{J}|$ is equal to 1 for the transformations considered in this study (see Bodin & Sambridge (2009) for details). We use a ‘burn-in’ period to ensure the convergence of the Markov chain. That means, the first samples are discarded until the Markov chain is expected to have reached its equilibrium. The following samples \mathbf{Z} (and hence \mathbf{c}) constitute samples from the posterior distribution

π , from which any statistical quantity such as mean or variance can be computed.

The proposal distribution comprises the following types of moves:

- (i) perturb a parameter describing an ellipse (position, axis length, angle) using a Gaussian probability density centred in 0;
- (ii) perturb the hyperparameter h using a similar probability density;
- (iii) birth: add a new ellipse to the vector of parameters \mathbf{Z} , each of the new parameters being drawn from the prior distribution;
- (iv) death: remove an ellipse, hence its parameters, from \mathbf{Z} .

For moves of type (i) and (ii), the proposal distribution is symmetric, $q(\mathbf{Z}'|\mathbf{Z}) = q(\mathbf{Z}|\mathbf{Z}')$, which can be used to simplify expression (17). For moves of type (iii) and (iv), the proposal equals the prior distribution, they cancel out and the second term of the right-hand side in eq. (17) becomes the ratio of the likelihoods: we do not need to explicit the prior distribution but only to sample from it (Mosegaard & Tarantola 1995).

This method is appropriate when strong prior information is available, which is the case in this study. For a fast convergence of the Markov chain, numerical experiments suggest that the proportion of accepted moves should be between 10 and 50 per cent. In birth moves, the proposal and the prior are identical. Hence the prior should be close to the likelihood function to achieve a sufficiently high acceptance of proposed samples.

5 SYNTHETIC TESTS FOR THE INVERSE HOMOGENIZATION

This section presents synthetic examples of the inverse homogenization step only. We try to recover small-scale isotropic anomalies, inducing effective anisotropy, from their effective large-scale properties. The goal is to validate the proposed rj-MCMC algorithm and to study the effect of the prior information introduced in the problem.

5.1 Setup

We consider a 2-D square homogeneous plate with 120 m sides (all boundaries are reflective). The first synthetic model contains four ellipsoidal cavities having minor and major diameters of sizes 2 and 6 m, respectively, as seen in Fig. 2. We compute the effective medium of this model using the homogenization technique described in paragraph 2.2, using a scale separation parameter $\lambda_0 = 16.6$ m. This value of λ_0 will be kept for the remainder of the study. We present in Fig. 2 the resulting shear-wave velocity V_S and a measure of anisotropy defined by the ratio $\|\mathbf{c} - \mathbf{c}_{\text{iso}}\|_2 / \|\mathbf{c}_{\text{iso}}\|_2$, where \mathbf{c}_{iso} is the usual isotropic projection of \mathbf{c} (Fedorov 2013; Browaeys & Chevrot 2004) and $\|\cdot\|_2$ the Euclidean (or Frobenius) matrix norm. While the original model is discontinuous and isotropic, its effective medium is smooth and anisotropic.

We add spatially correlated Gaussian noise to the elastic parameters of the homogenized seismic model. We chose an exponential covariance function with characteristic length-scale λ_0 , in order to reproduce the correlations that would be observed in an elastic model obtained from HFWI. There are however no correlations in-between elastic parameters. This noisy elastic field noted \mathbf{m}_n^* is used as data for the inverse homogenization problem.

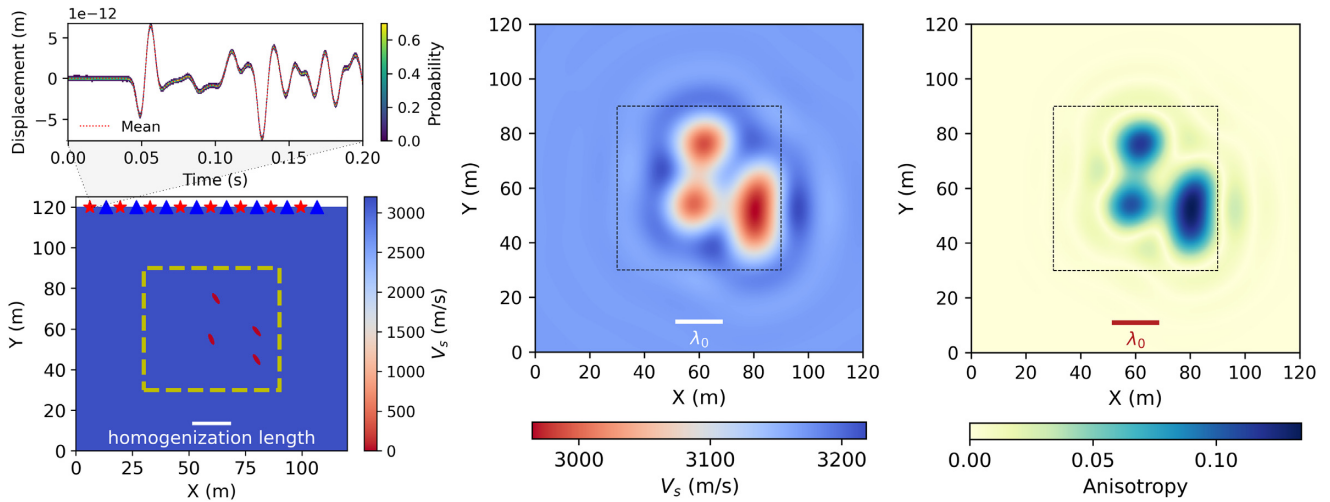


Figure 2. Synthetic model consisting in a homogeneous rectangular domain containing ellipsoidal cavities: (left-hand panel) the true model and associated synthetic waveform data, (right-hand panel) shear wave velocity and anisotropy of its effective medium. Red stars show the sources locations, blue triangles the receivers and a white bar shows λ_0 . The inverted region is inside the dashed-line rectangles.

5.2 Results

We invert \mathbf{m}_n^* for a small-scale model \mathbf{m} using the rj-MCMC algorithm presented in Section 4. Prior information is incorporated in the construction of the model \mathbf{m} and consists of the following: \mathbf{m} is isotropic, its elastic properties are chosen from only two possible materials, either the plate or the void of the cavities, and the cavities are parametrized as (an unknown number of) ellipses.

For each realization in the Markov chain, we need to build a discretized version of \mathbf{m} before calculating $\mathcal{H}(\mathbf{m})$ with a finite element method. The elements are triangular and the interpolation is based on the Fekete points using high order polynomials. The mesh is generated using the Gmsh tool (Geuzaine & Remacle 2009). Triangular elements are more versatile than spectral elements meshes, hence allow to generate a new mesh at each iteration of the Markov chain. But this way of proceeding will meet its limits in 3-D where automatic generation of conforming meshes is difficult. Alternative computing methods will be reviewed in the discussion.

The result of the inversion is the posterior distribution of the parameter vector \mathbf{Z} . A visual representation is obtained by computing the probability of being inside a cavity (Fig. 3): it corresponds, at each spatial position of the image, to the number of times this position is inside an ellipse divided by the number of sampled models. The algorithm is able to recover the presence of four cavities, as well as their general shape and orientation, which was not obvious by looking only at the effective model of Fig. 2. We test the inversion on an additional setup with six cavities of various sizes and orientations. Similarly, the true model is well recovered. The number of ellipses, n_e , is also an inferred parameter in the rj-MCMC algorithm. In both proposed tests, its posterior marginal distribution, shown Fig. 3, indicates a maximum at the correct value of the true model.

5.3 The role of prior information

In the absence of prior information, the inverse homogenization problem is necessarily non-unique, and does not significantly improve the understanding of the medium. For example, the observation \mathbf{m}_n^* would be itself a solution (up to the loss of information in the filtering operation of eq. 8). Therefore, we shall emphasize that

we are only able to recover the small-scale structure because the chosen parametrization and prior distribution incorporate sufficient constraints. To highlight the importance of prior information, we present an additional test with different choices for the construction of small-scale models.

We consider the synthetic model of Fig. 4, where the ellipses are closer to one another, while other parameters are identical to the previous case. Its effective medium displays one large radial anomaly and visually distinguishing the underlying structure is not possible. To construct the noisy data \mathbf{m}_n^* , we increase the level of noise added to the elastic parameters so that the rj-MCMC is not able to recover the small-scale model correctly.

Fig. 5 shows the inverse homogenization results for three different parametrizations, from left to right: (a) we consider a stronger prior information, using a fixed-dimensional MCMC algorithm and choosing the number of ellipses n_e constant and equal to four (the positions and geometries of the four ellipses are still unknown variables to be inverted for); (b) we use the rj-MCMC algorithm as previously and (c) we parametrize the small-scale model with a regular grid. Each cell has a 1×1 m size and can contain either the elastic properties of the homogeneous plate or the cavities. We invert for a 30×30 m area, hence comprising 900 parameters, using also a fixed-dimensional MCMC algorithm.

As expected, the quality of the inversion is correlated to the incorporated prior information. Algorithm (a), which includes the most information and imposes the correct number of parameters, is able to recover the original small-scale model. The rj-MCMC algorithm, however, is inherently parsimonious. We observe a trade-off between models with three or four cavities, as models with less parameters are favoured if they explain the data sufficiently well. Algorithm (c), which considers the least prior information, recovers only one elongated anomaly. Still, all methods are able to recover structures with a preferred orientation, although they are not directly notable on the effective medium. An analysis of the misfit function indicates that the orientation information comes from the anisotropic part of the data, especially from the off-diagonal elements of the elastic tensor. Furthermore, decreasing the level of prior information leads up to an increase of the number of inferred parameters and the convergence of the MCMC algorithm grows more difficult. The inverse homogenization is beneficial for

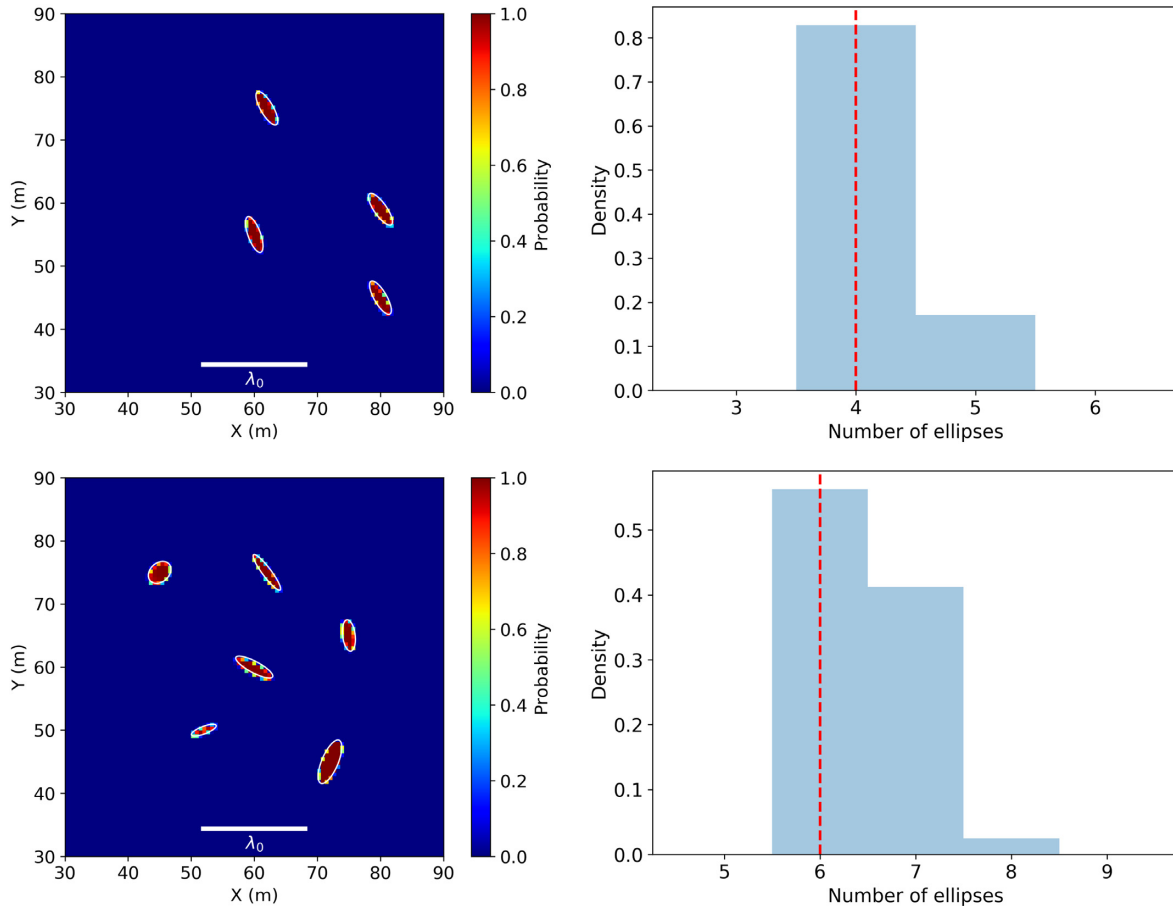


Figure 3. Results of the inverse homogenization. The top left-hand panel shows the probability of recovering void at each spatial point in the model, the top right-hand panel the marginal distribution of the number of ellipses n_e used to parametrize the small-scale model. The true position and shape of the cavities is shown as white ellipses. The bottom panels display similar results for a true small-scale model with six cavities of various shape and orientation.

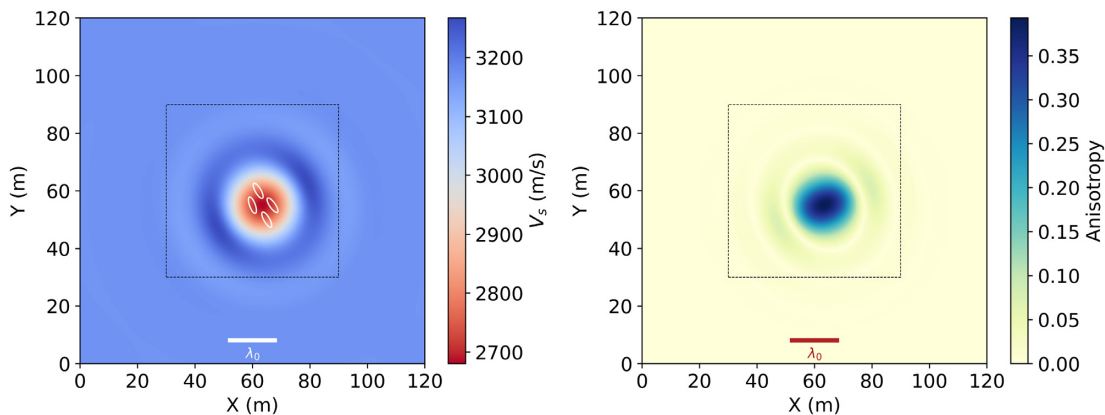


Figure 4. Effective seismic properties of a synthetic model having closer ellipsoidal cavities. The true position and shape of the cavities are represented on the left-hand panel as white ellipses.

problems where significant prior information is available, but, on the contrary, may not be useful if considering non-informative priors.

6 SYNTHETIC TEST FOR COMPLETE TWO-STEPS IMAGING METHOD

We now present synthetic examples of the full proposed procedure (HFWI followed by inverse homogenization). The setup is

identical to the one presented in Fig. 2, with elliptic cavities producing a large effective anisotropy amplitude. This configuration is relevant for subsurface cavity detection, for damaged structure inspection, or for reproducing reduced-scale experiments. Because they combine the information from all types of waves, full waveform methods are attractive for such problems (Bretaudeau *et al.* 2013; Tran *et al.* 2013). We argue here that non-negligible information is already present in the anisotropy at long wavelength (longer than the anomaly size) and can be leveraged with HFWI.

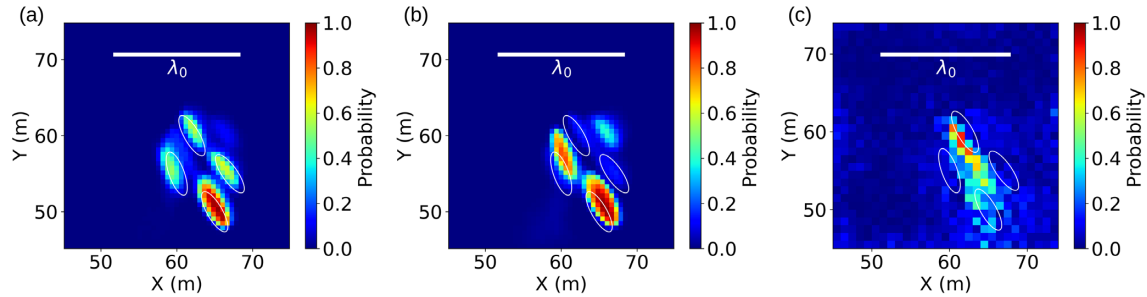


Figure 5. Results of the inverse homogenization for the synthetic model of Fig. 4, using three different parametrization: (a) a fixed number of ellipses $n_e = 4$, (b) a variable number n_e and (c) a regular grid parametrization.

Such problem may also be tackled with a ‘one-step’ FWI method based on shape optimization, where the geometries of the cavities are directly inverted for (e.g. Guo & de Hoop 2013). However, a conventional FWI would involve more expensive wave propagation given the fine scale discretization needed to mesh complex shapes with discontinuities. Derivatives with respect to the parameters defining ellipses position and shape would be difficult to compute, and we would expect convergence issues due to the strong non-linearity of the problem. With the method proposed here, we choose instead a parametrization adapted to waveform modelling and linearized inversion. We leverage information from effective seismic anisotropy, and the inverse homogenization step is solved with a derivative-free algorithm as described above.

6.1 Waveform data

Synthetic data is generated from sources and receivers placed at the top, using the 2-D spectral element solver. Correlated Gaussian noise with an amplitude representing 5 per cent of the signal standard deviation is added to the waveform data. $M = 1000$ noise realizations are generated. HFWI is performed independently for each realization in order to produce an ensemble of M tomographic solutions, and evaluate uncertainties on the model parameters. An example of synthetic waveform data distribution is displayed Fig. 2.

The first step in HFWI is to choose the scale separation between the fine and large scales. The maximum frequency present in the data is $f_{\max} \sim 90$ Hz. The background shear wave velocity is $V_{s, \min} = 3167.0 \text{ m s}^{-1}$, the slowest velocity in the imaged medium is somewhat lower due to the cavities. With $\lambda_0 = 16.6 \text{ m}$, the scale separation position is $\varepsilon_0 \sim 0.5$.

6.2 FWI solutions

The waveform inversion is performed using the approximate model space parametrization \mathcal{M}^{*h} . It is constructed on a 8×8 regular grid, each cell comprising a polynomial of degree 4 in each direction. No continuity is imposed between the elements. This ensures an adequate degree of freedom relative to the minimum wavelength λ_{\min} . The model \mathbf{m}^{*h} obtained after waveform inversion is then homogenized into \mathbf{m}^* with a minimum length-scale λ_0 .

We compare two possible parametrizations for \mathcal{M}^{*h} : (1) a fully anisotropic parametrization (ρ , \mathbf{c}) where the full elastic tensor is inverted for as prescribed in HFWI and (2) an isotropic parametrization (ρ , V_P , V_S) corresponding to a ‘conventional’ FWI (although a different, spatially denser discretization would have probably been used for a real FWI). One example of resulting inverted models \mathbf{m}^{*h} and \mathbf{m}^* for both parametrizations are compared in Fig. 6.

Again, we present the results in terms of shear-wave velocity V_S and anisotropy index. Several features of HFWI emerge. The \mathcal{M}^{*h} space is chosen for efficient wave propagation but does not aim at producing realistic images. As such, for both parametrizations, the four low velocity anomalies are recovered with a poor resolution and display the imprint of the inversion mesh. With the isotropic parametrization, we recover a sharper model. The strong low velocity anomalies help to explain the long coda in the waveform data. Such strong anomalies are not required once anisotropy is allowed.

Once projected into \mathcal{M}^* , results obtained in the case where \mathbf{m}^{*h} is fully anisotropic show only circular anomalies, while the strong anisotropy indicates that information about the shape is preserved. On the contrary, \mathbf{m}^* obtained from an isotropic \mathbf{m}^{*h} depicts weak anisotropy while the isotropic parameters V_P and V_S already display some structure. The spatial discretization is too loose to resolve all the small scales and especially the shape of the anomalies. This example suggests that a fully anisotropic parametrization preserves more information and is beneficial in the HFWI method. Data misfit reduction is about 76 per cent versus 61 per cent for the anisotropic and isotropic parametrizations of \mathbf{m}^{*h} , respectively. This is not surprising since the anisotropic parametrization has a higher complexity.

6.3 Uncertainties in HFWI

Estimating the uncertainty on the HFWI solution model \mathbf{m}^* is crucial for the following Bayesian inversion at the downscaling stage, where \mathbf{m}^* will take the role of the data (Fig. 1). The uncertainty on \mathbf{m}^* is estimated as described in Section 4.1. As in any Bayesian inversion, the result of the downscaling stage directly depends on this estimation.

We consider $M = 1000$ HFWI resulting models (\mathbf{m}_1^* , ..., \mathbf{m}_M^*) obtained for each noise realization on the data and examine their statistical properties. In the inverse homogenization problem, we suppose that the result of HFWI is equal to the homogenized true model \mathbf{m}_t^* . We verify this, at least in expectation, by comparing the mean $\bar{\mathbf{m}}^*$ of the elastic tensors of (\mathbf{m}_1^* , ..., \mathbf{m}_M^*) to the one of \mathbf{m}_t^* (Fig. 7a). We also perform for comparison a reference HFWI using waveform data without noise, and denote its results \mathbf{m}_0^* . All models are similar but not exactly identical and there is in particular a difference between \mathbf{m}_t^* and \mathbf{m}_0^* , which can be imputed to the imperfect illumination with this source–receiver configuration. Also, computing the mean on the elastic parameters is somewhat arbitrary (one could instead, for example, compute the mean on the seismic velocities). This choice may in part explain the difference between $\bar{\mathbf{m}}^*$ and \mathbf{m}_0^* . Ideally, we should construct an ‘homogenized’ mean in the space of effective models.

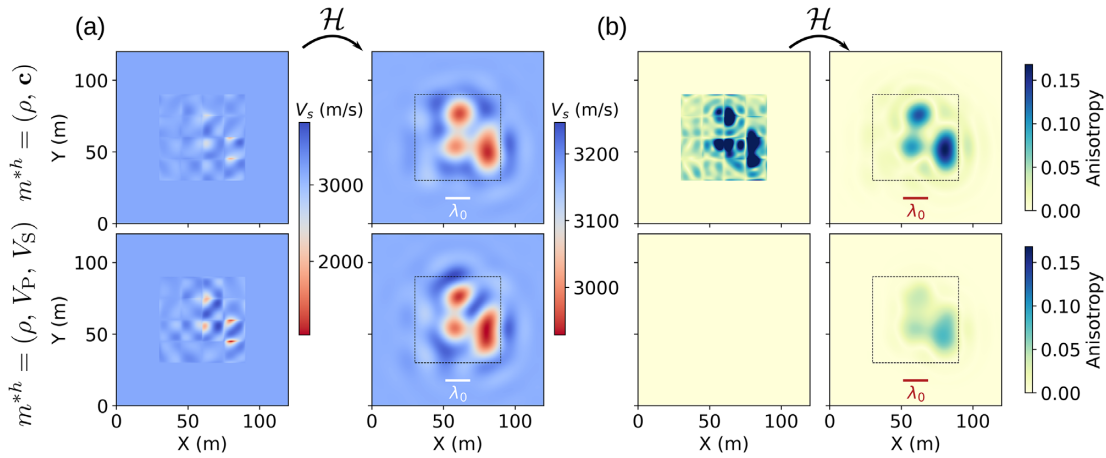


Figure 6. Full waveform inversion results presented in terms of (a) V_S and (b) anisotropy. The top row shows the images obtained with an anisotropic parametrization of \mathcal{M}^{*h} , and the bottom with the isotropic one. The left-hand column of each panel displays the results in the space \mathcal{M}^{*h} , where the inversion mesh is apparent, and the right-hand column their effective medium in space \mathcal{M}^* .

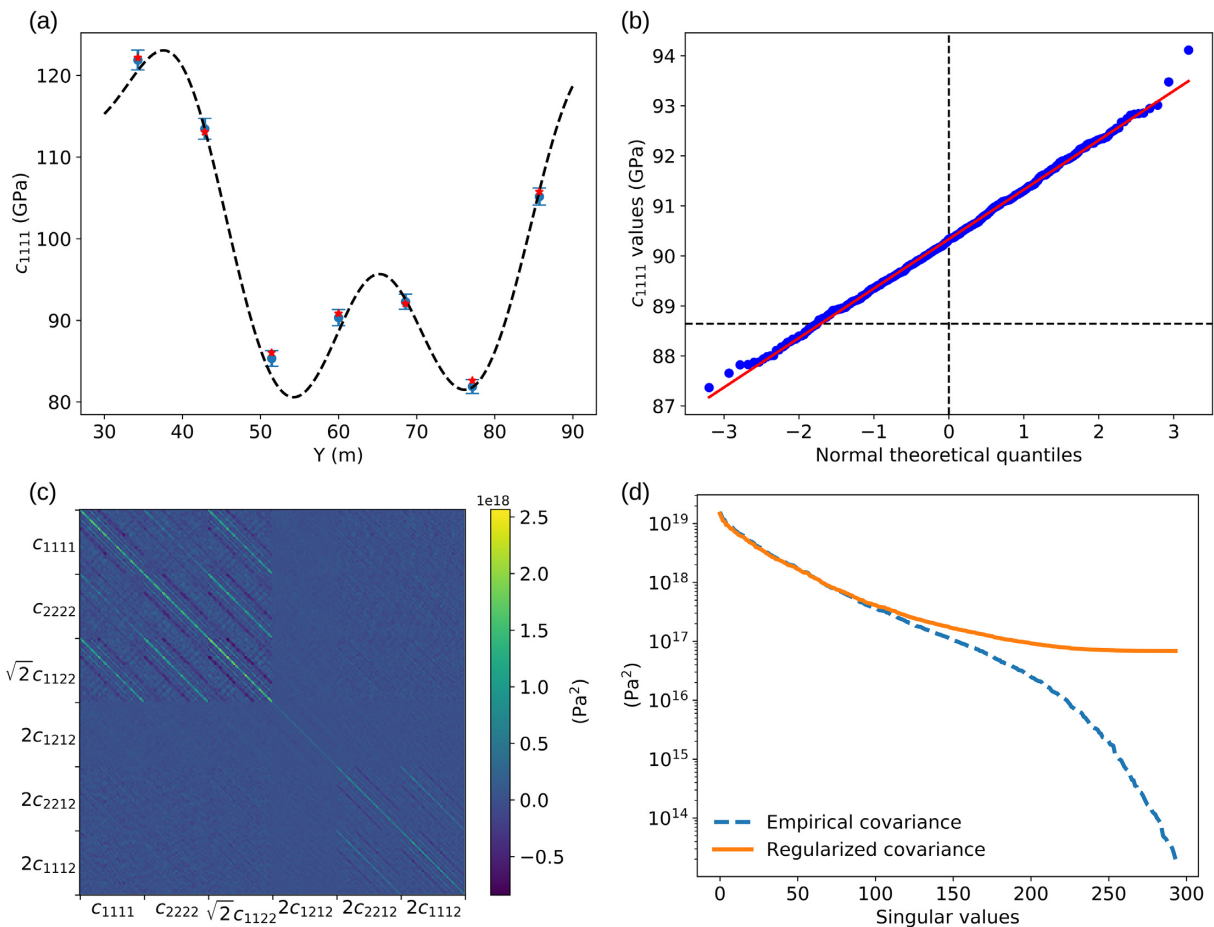


Figure 7. Uncertainty estimation in HFWI: (a) Comparison of the effective medium of the true model presented Fig. 2 (dashed black line) to the distribution of the M images obtained by HFWI, for parameter c_{1111} in a vertical cut at the middle of the model. The blue dot show the mean of the M estimations, and the error bars show the standard deviation. The red star represent a reference HFWI resulting model where no noise was added to the waveforms. (b) Quantile–quantile plot comparing the M samples for a parameter c_{1111} in the middle of the model to a normal distribution. (c) Regularized covariance matrix $\hat{\Sigma}_r$ of the models $(\mathbf{m}_1^*, \dots, \mathbf{m}_M^*)$, obtained with eq. (16) for a parameter $\alpha = 0.06$. Its singular value decomposition (d) exposes the effect of the regularization.

We now want to verify the Gaussian hypothesis on ϵ in eq. (12). We suppose that ϵ is distributed as the ensemble of the HFWI resulting models. The noise on the waveform data is additive and Gaussian. But because the inverse mapping might not be linear,

we need to verify whether the distribution of solution models parameters is also Gaussian. We examine the marginal distribution of individual model parameters from these M realizations. For example, we compare the marginal distribution of c_{1111} for a location

in the middle of the image to a normal distribution on a quantile–quantile plot Fig. 5. c_{1111} is indeed close to normally distributed. Similar conclusions can be obtained for other parameters in \mathcal{M}^* (not shown). All parameters being individually Gaussian distributed is a necessary (but not sufficient) condition for the parameter space to be multivariate Gaussian. Still, these observations suggest that the HFWI result is similar to the homogenized true model, and that the uncertainties are close to normally distributed.

We calculate the covariance matrix of the sample $(\mathbf{m}_1^*, \dots, \mathbf{m}_M^*)$ using the method presented Section 4.1. In practice, because these models are smooth, their properties can be described on a regular grid with 2 points per wavelength λ_0 . This allows to preserve their spatial frequency content, while having a minimal size for the estimated covariance matrix. These chosen grid points determine the solution space dimension N and will be used to compute the likelihood function in the inverse homogenization. The estimated covariance matrix, and the effect of the regularization on its eigenvalues is shown Figs 7(c) and (d), respectively. The analysis of its structure indicates that variables are spatially correlated on a length scale λ_0 . Interparameter correlations also exist, for example between parameters c_{2222} and c_{1122} .

The covariance matrix can also provide information about the effective elastic parameters that are best constrained by the waveform data through relative uncertainties. The relative uncertainty of a parameter is defined as its absolute uncertainty divided by its value. It is usually considered in seismic imaging that only a restricted part of the elastic tensor is resolved (for example, only isotropic or transverse isotropic properties, depending on the data type). As expected, the parameters involved in isotropic properties are best resolved (~ 1 per cent of relative uncertainty). c_{1112} and c_{2212} are non-zero only in the presence of anisotropy and have at least 5 per cent of relative uncertainty.

6.4 Downscaling results

To solve the inverse homogenization problem, we describe the ensemble of fine-scale models using the same object-based approach as previously, described by the vector of parameters Z . This choice may represent an excessively optimistic prior information for a seismology application, but it is appropriate for cavity detection. A good starting model for the rj-MCMC algorithm is required to reduce the ‘burn-in’ period until which the chain reaches its target distribution. For this purpose, we start with an optimization procedure based on simulated annealing (Kirkpatrick *et al.* 1983). Once a starting model is found, we proceed with the rj-MCMC algorithm following standard practice.

We first present the inverse homogenization results for the reference case: a macroscale solution model \mathbf{m}^* with the fully anisotropic parametrization is used as observed data with uncertainties characterized by the covariance matrix of Fig. 7. This model \mathbf{m}^* is shown in Fig. 6 (top row) and corresponds to one element of the M inverted models. Fig. 8 shows the probability of being inside a cavity at each spatial position. The algorithm is able to recover the main structure, identifying four probable cavities with the correct shape. However, the position of the cavities are close but not always centred on the true values. Such deviations did not appear in the synthetic tests of Section 5, and we conclude that the bias originates from a systematic error in HFWI. Possible explanations for this error are thus the effect of the damping in Gauss–Newton iterative scheme, an incorrect estimation of HFWI uncertainties from the M resulting models and an imperfect illumination.

To look at the results in more detail, the posterior marginal 1-D distributions for the number of ellipses n_e , the ellipse axes and their orientation are shown in Fig. 9, top row. The mode for the distribution of n_e is much higher than the true value of four. The algorithm uses additional ellipses to match some weak anomalies in \mathbf{m}^* . They are in general small and circular, hence with an arbitrary orientation. The modes for the distribution of axes and orientation are close to the true values. The hyperparameter h is only slightly higher than 1, suggesting that the covariance estimation is reasonable (Fig. 10).

6.5 Isotropic downscaling

We repeat the inverse homogenization this time using HFWI solution model obtained from an isotropic parametrization in \mathcal{M}^{*h} (Fig. 6, bottom row). In this test, the spatial discretization of the model in the HFWI is the same as in the anisotropic case (we use the same mesh), but only two elastic parameters are inverted for (V_P and V_S), leading to a reduced number of unknowns. Note that, once homogenized, the resulting effective medium in \mathcal{M}^* is not strictly isotropic. The waveform inversion uses the discontinuous nature of the discretization mesh in \mathcal{M}^{*h} to preserve a (small) part of the fine-scale information. The effective medium is weakly anisotropic (Fig. 6).

With a finer spatial discretization, increasing the number of unknowns to the same amount as in the anisotropic parametrization case, one can recover more details in the model. Once homogenized, such a finer scale isotropic model can retrieve the correct effective anisotropy. Some examples are presented in Capdeville & Métivier (2018). The drawbacks are an increased computational cost for wave propagation, and a more challenging convergence of the waveform inversion scheme.

From the M obtained effective elastic models, we compute a new covariance matrix $\hat{\Sigma}_r^{\text{iso}}$, and use it to calculate the likelihood function. One of the M effective models is selected for the inverse homogenization (the one displayed Fig. 6, bottom row). The resulting posterior probability of being inside an ellipse is presented Fig. 8, and 1-D marginals of parameters in Z Fig. 9 (bottom). Without surprise, models with void inclusions close to circular are preferred, as they produce only weak anisotropy. This test confirms that some of the small-scale information has been lost in the isotropic waveform inversion. We conclude that parametrizing \mathcal{M}^{*h} with the full elastic tensor can be beneficial for the inverse homogenization problem.

7 DISCUSSION

The HFWI of Capdeville & Métivier (2018) restricts the imaging solution space to effective media, which are not suited for geological interpretation. For this purpose, an additional downscaling step is required. A general summary of the complete procedure is presented in Fig. 11. Downscaling can be formulated as an inverse problem where the forward operator is the non-periodic homogenization. This study achieves the inverse homogenization of an elastic medium and clarifies the possible applications of the method. Because it was restricted to synthetic tests, several additional points need to be discussed.

The first important point is the choice of the minimum wavelength λ_{\min} . The maximum frequency in the waveform data is deduced from the chosen filter, but the minimum seismic velocity $V_{s, \min}$ is *a priori* unknown before imaging. A strategy to choose λ_{\min} has to be adopted for real data application. For instant, Capdeville & Métivier (2018) suggest to tune λ_{\min} iteratively during the HFWI.

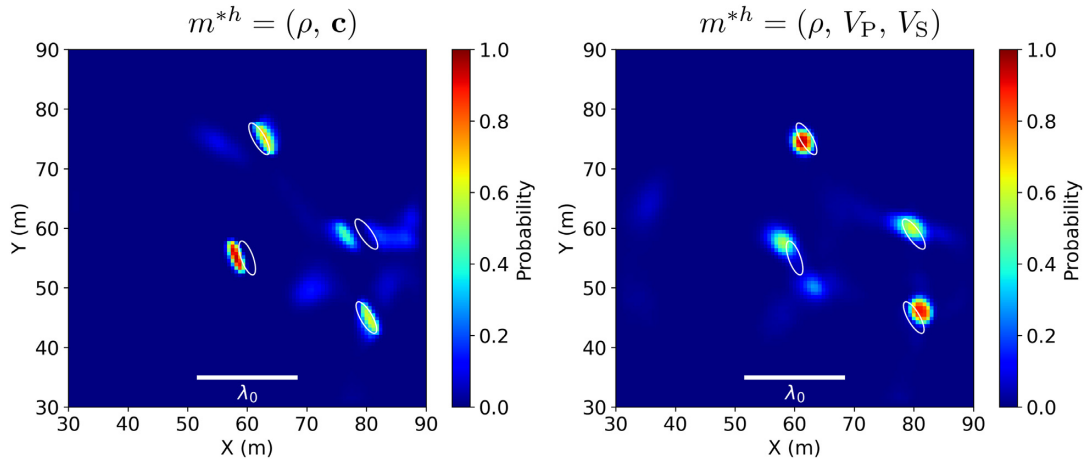


Figure 8. Probability of recovering void at each spatial point in the model. The left-hand panel presents the results when \mathcal{M}^{*h} is parametrized with the full elastic tensor, and the right one for the isotropic parametrization. The true position and shape of the inclusions is shown as white ellipses.

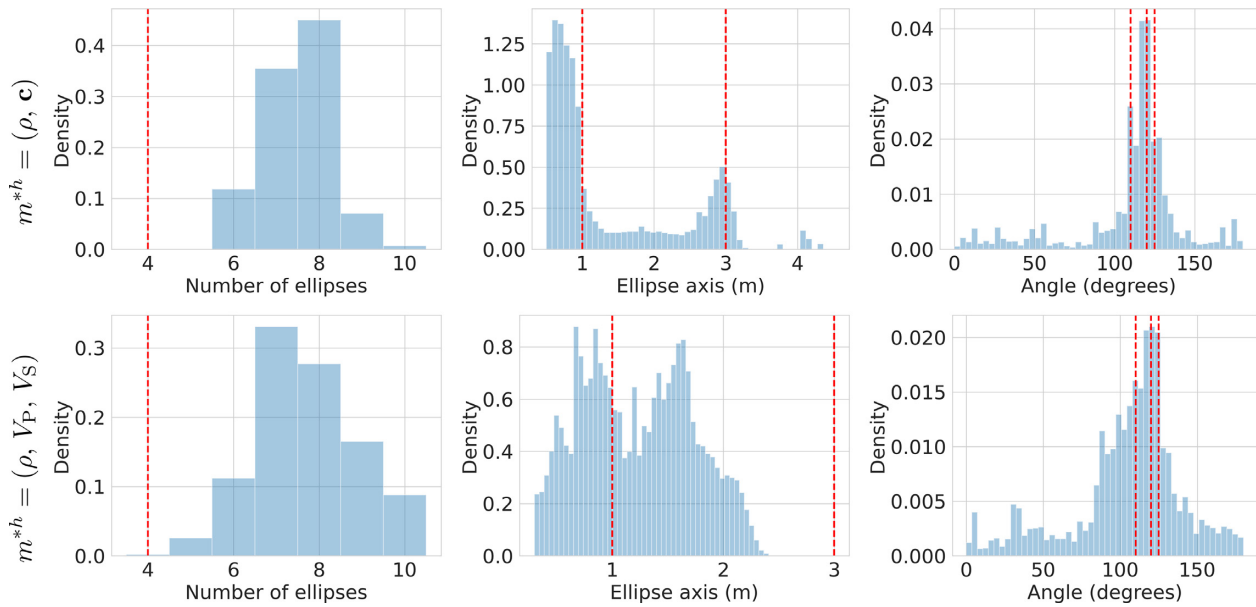


Figure 9. Posterior distribution of the latent variables in Z . The top/bottom panels correspond to anisotropic and isotropic parametrizations \mathcal{M}^{*h} , respectively. For each variable type, all the ellipses are combined on the same histogram. The y -axis shows the probability density associated to each histogram and the vertical dashed lines specifies the true value. On the middle panels, both short and long axis of the ellipses are combined on the same plot. The right-hand panels display the angle with respect to the horizontal.

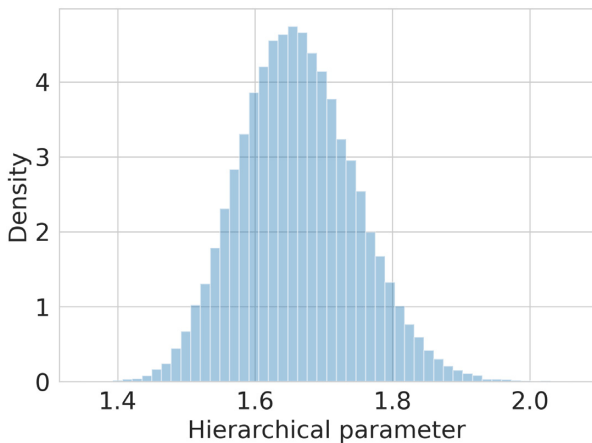


Figure 10. Posterior distribution of the hyperparameter h .

The assumption justifying the downscaling step is that the HFWI solution model is the effective medium of the true Earth model for scale separation position ε_0 (eq. 9). One necessary condition to meet this assumption is to find the global minimum of the HFWI. Note that this issue is not inherent to HFWI but present in all seismic tomography problems. Constraining the solution to homogenized models is a natural regularization because it reduces the size of the solution space and yet guarantees the same data fit. It avoids the use of biased smoothing constraints on the parameter space. More sophisticated strategies used in FWI to guide the algorithm convergence, such as frequency-continuation, are also possible to implement in HFWI (Capdeville & Métivier 2018).

We formulated the inverse homogenization problem in a Bayesian sense, where the solution is an ensemble of plausible fine-scale models. In this framework, uncertainties on the HFWI solution models are required. We computed them using a substitute method, performing multiple HFWI inversions for multiple noise realizations

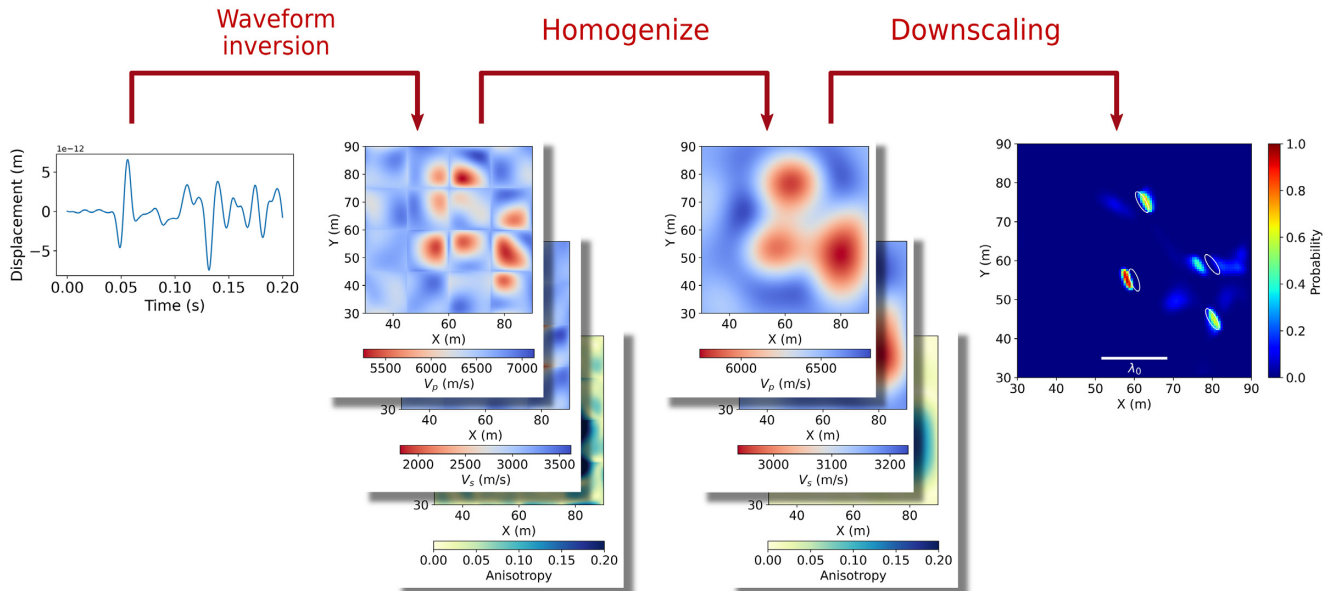


Figure 11. Summary of the seismic imaging method used in the study. The different steps corresponding to Fig. 1 are illustrated for the reference inversion, here shown in terms of P -wave velocity.

on the synthetic data. Uncertainty estimation in FWI is a current topical issue (Zhu *et al.* 2016; Sen & Biswas 2017; Thurin *et al.* 2019). One can expect viable solutions in the following years, at least for local uncertainties around the global minimum. In addition, we observed that the spatial correlation between parameters in effective models is specific and linked to the homogenization filter, thus uncertainty estimation in HFWI might be simpler than in conventional FWI. We are therefore optimistic on the possibility to address this issue. The subsequent stage will be to clarify the theoretical link between uncertainties in the FWI problem and the statistical model assumed in the inverse homogenization (eq. 12).

We used a Bayesian inversion because the inverse homogenization is supposed to be highly non-unique. But posterior probability distributions in high dimension might be difficult to interpret. Using optimization methods to find a single ‘best-fitting’ model, according to some *a priori* expert knowledge, is sometimes favoured. If so, the HFWI uncertainty estimation issue is no longer relevant.

All the presented synthetic tests consider a limited number of inferred parameters, extension to larger problems is now discussed. In the HFWI part, inverting for the full elastic tensor increases the number of model parameters compared to conventional FWI. In return, the spatial discretization is directly defined by λ_0 which could, depending on the application, reduce the solution space size. Additionally, the homogenization concept is beneficial for the forward problem, notably in cases where wave propagation is difficult. For example, if the starting model holds discontinuities or small-scale heterogeneities, the mesh size can shrink drastically. Expensive wave propagation can be spared by considering an effective medium of the starting model. Ultimately, it still must be based on a solver able to manage fully anisotropic media.

The computational cost of the downscaling step is less of a problem. To answer a specific scientific question, one possibility is to select only a sub-area of the image, or a 2-D section of a 3-D image, for subsequent interpretation. Such a ‘target-oriented’ downscaling will have little computational cost compared to the waveform inversion step. The models considered in this study contain discontinuities. Computing \mathcal{H} with a finite elements method, as described in Capdeville *et al.* (2010), is appropriate. Hence, the mesh was

generated again at each new proposed model and the full space of the model was homogenized. For continuous models, an alternative method is to compute \mathcal{H} with an FFT-based solver on a regular grid (Capdeville *et al.* 2015). In this case, designing a proposal distribution with local perturbations, as well as computing \mathcal{H} only on the perturbed sub-area, is more straightforward and can further spare computation time.

For the setup of Fig. 2, the total cost of the forward problem (comprising the mesher and the solver) is less than a second on a computing node with 28 cores. We estimate that MCMC-based methods quantifying the full posterior distribution, where forward modelling is performed thousands to millions of times, will be applicable for 2-D downscaling problems. In 3-D, the downscaling step will be likely restricted to optimization methods.

We now discuss the range of potential applications of the presented 2-steps method. In our tests, we chose a problem comprising small-scale anomalies with maximal contrast and a preferred orientation, which constitutes an ideal case for using the information contained in the effective anisotropy. If we had considered a model with circular cavities smaller than the minimum wavelength, or a model with weaker velocity anomalies, their effective properties would not contain significant anisotropy. When the effective medium seen by the wavefield is close to isotropic, a FWI with isotropic parametrization would be able to reveal the same level of information. This is related to the fact that, in the case of weak velocity contrasts, the homogenization operator is close to a simple low-pass filter. Still, we believe this does not erase all the advantages of HFWI. First, using a fully anisotropic parametrization facilitates the convergence of the waveform inversion. Additionally, many imaged natural media present indeed effective anisotropy and it is generally not possible to know beforehand if a medium is isotropic at a given scale.

Structures containing cavities, fluid inclusions or cracks will likely produce significant effective anisotropy, hence constitute appropriate use cases. They will exist at the engineering and subsurface scales, where, additionally, strong prior information is usually available. Although not explored in this study, applications at the continental or global scale are envisioned. In seismology, a long standing issue is the ability to distinguish effective anisotropy (e.g.

induced by small-scale heterogeneities or preferentially oriented faults) from intrinsic anisotropy (induced by anisotropic crystals, e.g. Fichtner *et al.* 2013a; Alder *et al.* 2017). For example, down-scaling of global surface wave tomography models could help to discriminate between effective and intrinsic anisotropy.

Moreover, our two-step approach is particularly suited to handle discontinuities in elastic properties. FWI methods struggle to build discontinuous models, as most waveform solvers rely on spectral-element methods with rectangular meshes that need to honor the geometry of discontinuities. Inverse homogenization alleviates these limitations, because its forward part solves an elasto-static problem with a standard finite-element method and a triangular mesh. We showed a near-surface / engineering-scale application to image cavities, other possible applications at larger scale are characterization of discontinuous geological structures such as faults in the subsurface or study the depth and shape of major discontinuities in the Earth. A typical example is the issue of crustal corrections in global tomography, required as details within the crust cannot be resolved by long-period seismic waves. Inaccurate corrections lead to artifacts in the less constrained part of the model (e.g. Marone & Romanowicz 2007). Instead of using corrections, one could first invert long period waves for a smooth homogeneous Earth without crust, and in a second time, together with extra information, apply a downscaling approach to reconstruct crustal structures.

8 CONCLUSION

We have presented a proof of concept of the inverse homogenization of an elastic medium, applied to a void detection problem. The seismic imaging problem is solved in two steps. The first one is a waveform inversion, where the solution space is restricted to effective or homogenized media. Secondly, in the inverse homogenization step, the resulting effective medium is used as data and we look for small-scale models having the corresponding effective properties. The main benefits are an easier convergence for the waveform inversion in problems with strong heterogeneities and a better control on the prior information in the inverse homogenization. The theory was presented in general terms and is applicable at different scales. Thus, the benefits of an extension to other seismic imaging problems will be studied in future work. Different implementations might be more appropriate depending on the problem size and on the properties of the heterogeneities (continuity, solid or fluid, etc.). Besides the items for discussion mentioned above, future work on inverse homogenization need to focus on building realistic prior models of geological media and adapting the inversion strategy to these models.

ACKNOWLEDGEMENTS

We are grateful to Michael Afanasiev and an anonymous reviewer for their comments that helped to greatly improve the manuscript. This work was funded by the French National Research Agency, project HIWAI (ANR-16-CE31-0022-01), and the European Research Council (European Union's Horizon 2020 research and innovation program under Grant Agreement No. 716542).

DATA AVAILABILITY

This study uses synthetic data only. All the codes used will be shared on reasonable request to the corresponding author.

REFERENCES

- Afanasiev, M., Boehm, C., May, D. & Fichtner, A., 2016. Using effective medium theory to better constrain full waveform inversion, in *Proceeding of the 78th EAGE Conference and Exhibition 2016*, May 2016, Vol. 2016, pp. 1–3, European Association of Geoscientists & Engineers, <https://doi.org/10.3997/2214-4609.201601614>.
- Alder, C., Bodin, T., Ricard, Y., Capdeville, Y., Debayle, E. & Montagner, J., 2017. Quantifying seismic anisotropy induced by small-scale chemical heterogeneities, *Geophys. J. Int.*, **211**(3), 1585–1600.
- Almuhaidib, A.M. & Toksöz, M.N., 2015. Imaging of near-surface heterogeneities by scattered elastic waves, *Geophysics*, **80**(4), A83–A88.
- Backus, G.E., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. geophys. Res.*, **67**(11), 4427–4440.
- Bensoussan, A., Lions, J.-L. & Papanicolaou, G., 2011. *Asymptotic Analysis for Periodic Structures*, Vol. 374, American Mathematical Soc.
- Bernard, S., Monteiller, V., Komatitsch, D. & Lasaygues, P., 2017. Ultrasonic computed tomography based on full-waveform inversion for bone quantitative imaging, *Phys. Med. Biol.*, **62**(17), 7011.
- Bodin, T. & Sambridge, M., 2009. Seismic tomography with the reversible jump algorithm, *Geophys. J. Int.*, **178**(3), 1411–1436.
- Bodin, T., Capdeville, Y., Romanowicz, B. & Montagner, J.-P., 2015. Interpreting radial anisotropy in global and regional tomographic models, in *The Earth's Heterogeneous Mantle: A Geophysical, Geodynamical, and Geochemical Perspective*, pp. 105–144, eds Khan, A. & Deschamps, F., Springer.
- Bozdağ, E., Trampert, J. & Tromp, J., 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.*, **185**(2), 845–870.
- Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N. & Pugmire, D., 2016. Global adjoint tomography: first-generation model, *Geophys. J. Int.*, **207**(3), 1739–1766.
- Bretaudeau, F., Brossier, R., Leparoux, D., Abraham, O. & Virieux, J., 2013. 2D elastic full-waveform imaging of the near-surface: application to synthetic and physical modelling data sets, *Near Surf. Geophys.*, **11**(3), 307–316.
- Brossier, R., Operto, S. & Virieux, J., 2009. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion, *Geophysics*, **74**(6), WCC105–WCC118.
- Browaeys, J.T. & Chevrot, S., 2004. Decomposition of the elastic tensor and geophysical applications, *Geophys. J. Int.*, **159**(2), 667–678.
- Bui-Thanh, T., Ghattas, O., Martin, J. & Stadler, G., 2013. A computational framework for infinite-dimensional Bayesian inverse problems. Part I: the linearized case, with application to global seismic inversion, *SIAM J. Sci. Comput.*, **35**(6), A2494–A2523.
- Bunks, C., Saleck, F.M., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics*, **60**(5), 1457–1473.
- Capdeville, Y. & Marigo, J.-J., 2007. Second order homogenization of the elastic wave equation for non-periodic layered media, *Geophys. J. Int.*, **170**(2), 823–838.
- Capdeville, Y. & Marigo, J.-J., 2012. A non-periodic two scale asymptotic method to take account of rough topographies for 2-D elastic wave propagation, *Geophys. J. Int.*, **192**(1), 163–189.
- Capdeville, Y. & Métivier, L., 2018. Elastic full waveform inversion based on the homogenization method: theoretical framework and 2-D numerical illustrations, *Geophys. J. Int.*, **213**(2), 1093–1112.
- Capdeville, Y., Guillot, L. & Marigo, J.-J., 2010. 2-D non-periodic homogenization to upscale elastic media for P–SV waves, *Geophys. J. Int.*, **182**(2), 903–922.
- Capdeville, Y., Stutzmann, E., Wang, N. & Montagner, J.-P., 2013. Residual homogenization for seismic forward and inverse problems in layered media, *Geophys. J. Int.*, **194**(1), 470–487.
- Capdeville, Y., Zhao, M. & Cupillard, P., 2015. Fast Fourier homogenization for elastic wave propagation in complex media, *Wave Motion*, **54**, 170–186.
- Chen, Y., Wiesel, A., Eldar, Y.C. & Hero, A.O., 2010. Shrinkage algorithms for MMSE covariance estimation, *IEEE Trans. Signal Process.*, **58**(10), 5016–5029.

- Devaney, A., 1984. Geophysical diffraction tomography, *IEEE Trans. Geosci. Rem. Sens.*, **22**(1), 3–13.
- Dupuy, B., Asnaashari, A., Brossier, R., Garambois, S., Métivier, L., Ribodetti, A. & Virieux, J., 2016. A downscaling strategy from FWI to microscale reservoir properties from high-resolution images, *Leading Edge*, **35**(2), 146–150.
- Fedorov, F.I., 2013. *Theory of Elastic Waves in Crystals*, Springer Science & Business Media.
- Fichtner, A., 2010. *Full Seismic Waveform Modelling and Inversion*, Springer Science & Business Media.
- Fichtner, A. & Trampert, J., 2011. Resolution analysis in full waveform inversion, *Geophys. J. Int.*, **187**(3), 1604–1624.
- Fichtner, A., Kennett, B.L. & Trampert, J., 2013a. Separating intrinsic and apparent anisotropy, *Phys. Earth planet. Inter.*, **219**, 11–20.
- Fichtner, A., Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y. & Villasenor, A., 2013b. Multiscale full waveform inversion, *Geophys. J. Int.*, **194**(1), 534–556.
- Frederick, C. & Engquist, B., 2017. Numerical methods for multiscale inverse problems, *Commun. Math. Sci.*, **15**, 2.
- Gelman, A., Stern, H.S., Carlin, J.B., Dunson, D.B., Vehtari, A. & Rubin, D.B., 2013. *Bayesian Data Analysis*, Chapman and Hall/CRC.
- Geuzaine, C. & Remacle, J.-F., 2009. Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities, *Int. J. Numer. Meth. Engng.*, **79**(11), 1309–1331.
- Grandjean, G. & Leparoux, D., 2004. The potential of seismic methods for detecting cavities and buried objects: experimentation at a test site, *J. appl. Geophys.*, **56**(2), 93–106.
- Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika*, **82**(4), 711–732.
- Guillot, L., Capdeville, Y. & Marigo, J.-J., 2010. 2-D non-periodic homogenization of the elastic wave equation: SH case, *Geophys. J. Int.*, **182**(3), 1438–1454.
- Guo, Z. & de Hoop, M.V., 2013. Shape optimization and level set method in full waveform inversion with 3D body reconstruction, in *Proceedings of the SEG Technical Program Expanded Abstracts 2013*, pp. 1079–1083, Society of Exploration Geophysicists.
- Hansen, T.M., Cordua, K.S., Jacobsen, B.H. & Mosegaard, K., 2014. Accounting for imperfect forward modeling in geophysical inverse problems—exemplified for crosshole tomography, *Geophysics*, **79**(3), H1–H21.
- Hoang, V.H. & Quek, J.H., 2019. Bayesian inverse problems for recovering coefficients of two scale elliptic equations, *Inverse Problems*, **35**(4), 045005.
- Huang, Y. & Schuster, G.T., 2014. Resolution limits for wave equation imaging, *J. appl. Geophys.*, **107**, 137–148.
- Kaipio, J. & Somersalo, E., 2006. *Statistical and Computational Inverse Problems*, Vol. 160, Springer Science & Business Media.
- Käufel, P., Fichtner, A. & Igel, H., 2013. Probabilistic full waveform inversion based on tectonic regionalization—development and application to the Australian upper mantle, *Geophys. J. Int.*, **193**(1), 437–451.
- Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P., 1983. Optimization by simulated annealing, *Science*, **220**(4598), 671–680.
- Kocur, G.K., Saenger, E.H., Grosse, C.U. & Vogel, T., 2016. Time reverse modeling of acoustic emissions in a reinforced concrete beam, *Ultrasonics*, **65**, 96–104.
- Komatitsch, D. & Vilotte, J.-P., 1998. The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. seism. Soc. Am.*, **88**(2), 368–392.
- Ledoit, O. & Wolf, M., 2004. A well-conditioned estimator for large-dimensional covariance matrices, *J. Multiv. Anal.*, **88**(2), 365–411.
- Li, X.-D. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. geophys. Res.*, **101**(B10), 22 245–22 272.
- Marone, F. & Romanowicz, B., 2007. Non-linear crustal corrections in high-resolution regional waveform seismic tomography, *Geophys. J. Int.*, **170**(1), 460–467.
- Métivier, L., Brossier, R., Mérigot, Q., Oudet, E. & Virieux, J., 2016. Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion, *Geophys. J. Int.*, **205**(1), 345–377.
- Mosegaard, K. & Tarantola, A., 1995. Monte carlo sampling of solutions to inverse problems, *J. geophys. Res.*, **100**(B7), 12 431–12 447.
- Nawaz, M.A. & Curtis, A., 2016. Bayesian inversion of seismic attributes for geological facies using a hidden Markov model, *Geophys. J. Int.*, **208**(2), 1184–1200.
- Nguyen, L.T. & Modrak, R.T., 2018. Ultrasonic wavefield inversion and migration in complex heterogeneous structures: 2D numerical imaging and nondestructive testing experiments, *Ultrasonics*, **82**, 357–370.
- Nolen, J., Pavliotis, G.A. & Stuart, A.M., 2012. Multiscale modelling and inverse problems, in *Numerical Analysis of Multiscale Problems*, pp. 1–34, Springer.
- Pageot, D., Leparoux, D., Le Feuvre, M., Durand, O., Côte, P. & Capdeville, Y., 2017. Improving the seismic small-scale modelling by comparison with numerical methods, *Geophys. J. Int.*, **211**(1), 637–649.
- 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.
- Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain; part 1, theory and verification in a physical scale model, *Geophysics*, **64**(3), 888–901.
- Rose, J.H., 1989. Elastic wave inverse scattering in nondestructive evaluation, in *Scattering and Attenuation of Seismic Waves, Part II*, pp. 715–739, Springer.
- Sambridge, M., Bodin, T., Gallagher, K. & Tkalčić, H., 2013. Transdimensional inference in the geosciences, *Phil. Trans. R. Soc. A*, **371**(1984), 20110547.
- Sánchez-Palencia, E., 1980. Non-homogeneous media and vibration theory, *Lecture Notes in Physics*, **127**.
- Seidl, R. & Rank, E., 2016. Iterative time reversal based flaw identification, *Comput. Math. Appl.*, **72**(4), 879–892.
- Sen, M.K. & Biswas, R., 2017. Transdimensional seismic inversion using the reversible jump hamiltonian monte carlo algorithm, *Geophysics*, **82**(3), R119–R134.
- Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.*, **180**(1), 433–462.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**(8), 1259–1266.
- Thurin, J., Brossier, R. & Métivier, L., 2019. Ensemble-based uncertainty estimation in full waveform inversion, *Geophys. J. Int.*, **219**(3), 1613–1635.
- Tran, K.T., McVay, M., Faraone, M. & Horhota, D., 2013. Sinkhole detection using 2D full seismic waveform tomography, *Geophysics*, **78**(5), R175–R183.
- Tromp, J., Komatitsch, D. & Liu, Q., 2008. Spectral-element and adjoint methods in seismology, *Commun. Comput. Phys.*, **3**(1), 1–32.
- Virieux, J., Asnaashari, A., Brossier, R., Métivier, L., Ribodetti, A. & Zhou, W., 2017. An introduction to full waveform inversion, in *Encyclopedia of Exploration Geophysics*, pp. R1–I, Society of Exploration Geophysicists.
- Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle: three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. geophys. Res.*, **89**(B7), 5953–5986.
- Zhu, H., Li, S., Fomel, S., Stadler, G. & Ghattas, O., 2016. A Bayesian approach to estimate uncertainty for full-waveform inversion using a priori information from depth migration, *Geophysics*, **81**(5), R307–R323.