



HAL
open science

Multiscale seismic imaging with inverse homogenization

N Hedjazian, Yann Capdeville, T. Bodin

► **To cite this version:**

N Hedjazian, Yann Capdeville, T. Bodin. Multiscale seismic imaging with inverse homogenization. Geophysical Journal International, 2021, 226 (1), pp.676-691. 10.1093/gji/ggab121 . hal-03179337

HAL Id: hal-03179337

<https://hal.science/hal-03179337>

Submitted on 24 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Multiscale seismic imaging with inverse 2 homogenization

3 N. Hedjazian¹, Y. Capdeville², T. Bodin¹

¹*Université de Lyon, UCBL, CNRS, LGL-TPE, 69622 Villeurbanne, France*

²*Université de Nantes, Laboratoire de Planétologie et Géodynamique, CNRS, 44322 Nantes, France*

4 14 October 2020

5 List of changes

6 SUMMARY

7
8 Seismic imaging techniques such as elastic full waveform inversion (FWI) have their
9 spatial resolution limited by the maximum frequency present in the observed waveforms.
10 Scales smaller than a fraction of the minimum wavelength cannot be resolved, and only a
11 smoothed version of the true underlying medium can be recovered. Application of FWI to
12 media containing small and strong heterogeneities therefore remains problematic. This
13 smooth tomographic image is related to the effective elastic properties, which can be
14 exposed with the upscaling or homogenization theory of wave propagation. We study
15 how this theory can be used in the context of FWI. The seismic imaging problem is
16 broken down in a two-stage multiscale approach. In the first step, called homogenized
17 full waveform inversion (HFWI), observed waveforms are inverted for a macro-scale, fully
18 anisotropic effective medium, smooth at the scale of the shortest wavelength present in
19 the wavefield. The solution being an effective medium, it is difficult to directly interpret
20 it. It requires a second step, called downscaling or inverse homogenization, where the
21 macro-scale image is used as data, and the goal is to recover micro-scale parameters. All
22 the information contained in the observed waveform is extracted in the HFWI step. The

23 solution of the downscaling step is highly non-unique as many small scale models may
24 share the same long wavelength effective properties. We therefore rely on the introduction
25 of external *a priori* information, and cast the problem in a Bayesian formulation. The
26 ensemble of potential fine-scale models sharing the same long wavelength equivalent is
27 explored with a Markov chain Monte Carlo algorithm. We illustrate the method with
28 a synthetic cavity detection problem: we search for the position, size and shape of void
29 inclusions in a homogeneous elastic media, where the size of cavities is smaller than
30 the resolving length of the seismic data. We illustrate the advantages of introducing
31 the homogenization theory at both stages. In HFWI, homogenization acts as a natural
32 regularization helping convergence toward meaningful solution models. Working with
33 fully anisotropic effective media prevents the leakage of anisotropy induced by the fine
34 scales into isotropic macro-parameters estimates. In the downscaling step, the forward
35 theory is the homogenization itself. It is computationally cheap, allowing us to consider
36 geological models with more complexity (e.g. including discontinuities) and use stochastic
37 inversion techniques.

38 1 INTRODUCTION

39 Seismic tomography aims at quantifying elastic properties in the Earth’s interior using seis-
40 mic data measured at the surface. It can be formulated as an inverse problem, where un-
41 known parameters defining the Earth model are sought in order to replicate the observed
42 data. Here we consider the case of full waveform inversion, where the seismic data is the
43 seismogram waveform. The primary applications range from the exploration scale to image
44 the subsurface, up to the global scale to study the Earth’s deep interior (Tarantola 1984;
45 Brossier et al. 2009; Tape et al. 2010; Fichtner 2010; Bozdağ et al. 2016; Virieux et al.
46 2017), while applications at the engineering scale or in medical imaging start to develop
47 (e.g. Bernard et al. 2017; Nguyen & Modrak 2018). Replicating the observations requires
48 to compute synthetic seismograms, which constitutes the “forward problem”. For example,
49 waveform modeling can be performed using approximated methods, based on ray theory,
50 normal mode summation, or on the Born approximation (Woodhouse & Dziewonski 1984;

51 Devaney 1984; Li & Romanowicz 1996), or more recently by numerically solving the full
52 wave equation, leading to full waveform inversion (FWI) methods (Fichtner 2010; Virieux
53 et al. 2017). In the standard approach, the misfit between simulated and observed data is
54 minimized using a local optimization strategy based on data derivatives. Global parameter
55 search methods are still out of reach for most real size 3D problems (Käuffl et al. 2013). Once
56 the misfit reaches a minimum, hopefully a global minimum, we obtain a quantitative im-
57 age of the material elastic properties which aims at being interpreted in terms of geological
58 structures.

59 However, most waveform tomography problems represent an ill-posed inverse problem,
60 where the solution is potentially non-unique. Many strategies have been developed to ensure
61 convergence toward a meaningful solution. These include, among others, building a good
62 starting model with alternative imaging methods (such as travel-time tomography), various
63 regularization methods in the parameter space (smoothing, prior information), multiscale
64 imaging by gradually increasing the frequency content of the waveforms (Bunks et al. 1995;
65 Brossier et al. 2009), or constructing more sophisticated misfit functionals (e.g. Bozdağ et al.
66 2011; Métivier et al. 2016). As a result, tomography, and in particular its interpretation,
67 requires a lot of expertise. A recurrent discussion is to distinguish the features in the solution
68 model that are properly resolved and required by the data, from those introduced by the
69 specificity of the implementation, such as an initial model or an *a priori* constraint on the
70 model parameters (promoting smoothness or sparsity, imposing discontinuities or symmetries
71 in the elastic tensor, etc). Also, compromises have often to be made between the fit to the
72 data and the realism of the image, or between model complexity and model constraint.

73 Some of these aspects originate from two fundamental limitations. (i) Due to instrumen-
74 tal limitations or to the high computational cost of waveform modeling, the seismic data
75 are restricted to a finite frequency range. (ii) The imaged media are usually multiscale,
76 but seismic inversions attempt to recover only a finite, limited number of model parame-
77 ters. Accordingly, seismic imaging has difficulties to identify the different scales present in

78 the seismic signal. A common understanding approximates that the spatial resolution of a
79 tomographic image is, at best, the smallest considered wavelength.

80 We take interest in imaging problems involving high contrast, small-scale heterogeneities,
81 especially with discontinuities in elastic properties. This comprises, for example, imaging
82 the subsurface to detect faults, fluids, karsts or tunnels, and also non-destructive testing
83 to detect flaws or rebars in concrete at the engineering scale. To address these problems,
84 some possible methods are based on single-scattering formulations (e.g. Rose 1989), or are
85 based on migration using time reversal and an imaging condition (Almuhaidib & Toksöz
86 2015; Kocur et al. 2016). In particular, the present study deals with a cavity detection
87 problem. This is usually performed with P- or Rayleigh-wave diffraction imaging (Grandjean
88 & Leparoux 2004). These methods usually require complex processing of the data, to separate
89 the different waves that tend to arrive simultaneously, or to handle multiple scattering. Hence
90 FWI methods, in both acoustic and elastic regimes, have been investigated (Bretaudié et al.
91 2013; Tran et al. 2013; Seidl & Rank 2016; Nguyen & Modrak 2018). However, imaging such
92 structures requires a fine-scale description of the medium, which makes the computational
93 cost of modeling the full wavefield prohibitive.

94 An appropriate tool to mitigate these issues is homogenization or upscaling of the elas-
95 tic wave equation. It refers to an ensemble of methods able to compute long wavelength
96 effective equivalents to a fine scale elastic medium. The effective medium does not contain
97 heterogeneities smaller than a given minimum wavelength, but produces the same wavefield
98 as the original medium in a limited frequency band. The historical example is the effective
99 equivalent of a stratified media: an explicit formula was derived by Backus (1962), who
100 showed that a stack of horizontal isotropic layers will be equivalent, at large scales, to a
101 homogeneous hexagonal anisotropic medium with vertical axis of symmetry. The initial idea
102 behind the homogenization theory is to reduce the computational cost of waveform model-
103 ing in a complex medium. By calculating a smooth equivalent medium, it avoids the need
104 of solving the wave equation on a mesh containing all the fine scales and the possible dis-
105 continuities. In the present work, we proceed further and study how homogenization theory

106 can facilitate the inverse problem in seismic imaging. This tool being developed for elastic
 107 media, we restrict ourselves to this case in the following, but an equivalent for the acoustic
 108 case could be considered.

109 In the context of elastic FWI, Backus theory can be used to constrain the solution space
 110 to effective models (Capdeville et al. 2013), or to ensure that fine-scale model updates pre-
 111 serve long-wavelength properties of the model (Afanasiev et al. 2016). Here, we will consider
 112 the more general non-periodic homogenization (Capdeville et al. 2010; Guillot et al. 2010),
 113 a technique able to compute the effective equivalent of any given media, without the need
 114 for spatial periodicity, scale separation or statistical invariance. It enables to properly pose
 115 FWI as a multiscale inverse problem (Fichtner et al. 2013b). This is especially important for
 116 attenuation and anisotropy that are scale-dependent properties. Similar ideas are present in
 117 the field of porous media and subsurface flow, where mostly elliptic equations are consid-
 118 ered. For example, one multiscale inverse problem is to determine the medium properties
 119 (porosity, permeability, conductivity) when large micro-scales fluctuations are present. Ho-
 120 mogenization theory helps at various levels (regularization, uncertainty estimation), and a
 121 natural procedure is to search for an effective model solution of the inverse problem (Nolen
 122 et al. 2012).

123 We propose to solve a seismic imaging problem with the following method (figure 1): we
 124 first perform a waveform inversion where we aim to recover a macro-scale effective (or homog-
 125 enized) elastic medium, based on the homogenized full waveform inversion method (HFWI)
 126 developed by Capdeville & Métivier (2018). Next, we define a second inverse problem where
 127 the obtained macro-scale solution plays the role of the data and the homogenization op-
 128 erator the role of the forward model. That is, we search for micro-scale models having an
 129 effective equivalent close to the macro-scale solution. HFWI and subsequent inverse homog-
 130 enization can therefore be seen as a two-stage resolution method of a multiscale inverse
 131 problem (Frederick & Engquist 2014).

132 Let us first synthesize some elements and results of the HFWI method. For all natural
 133 media considered in seismic imaging, it is possible to define a minimum wavelength λ_{\min} ,

134 associated with the maximum frequency of the waveform data. The spatial wavelength under
 135 which structures cannot be resolved by FWI is directly related to λ_{\min} . We define a constant
 136 wavelength λ_0 , setting the separation between the fine and the large scales. The objective
 137 of HFWI is to recover, from the waveform data, an effective model at the scale of λ_0 .
 138 Knowing the diffraction resolution limit of a FWI is about $\lambda_{\min}/2$ (e.g. Huang & Schuster
 139 2014), a good choice is $\lambda_0 \geq \lambda_{\min}/2$, because smaller scales are poorly resolved by the data.
 140 Capdeville & Métivier (2018) consider numerical examples of a FWI using either isotropic
 141 or anisotropic parameters, adapting the space discretization to maintain identical number
 142 of free parameters. They show that the FWI solution may vary depending on the chosen
 143 parametrization. However, these different solutions have the same effective equivalent at
 144 the scale of $\lambda_0 = \lambda_{\min}$ and are in good agreement at the scale of $\lambda_0 = \lambda_{\min}/2$. Hence,
 145 introducing homogenization in the FWI problem leads to the following observation: while
 146 FWI is parametrization and mesh dependent, HFWI is not. If we assume that the real,
 147 multiscale Earth is solution of a given FWI problem (this may not be true due to the
 148 imperfection of the inversion algorithm), its effective equivalent is also a solution. As the
 149 real Earth is inaccessible, Capdeville & Métivier (2018) suggest to seek for an effective
 150 medium and thereby reduce the non-uniqueness of the FWI solution.

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

161 In this study, we show how the effective anisotropy recovered by the HFWI solution model

162 can be used to resolve small-scale isotropic heterogeneities. The forward operator in the
163 inverse homogenization problem is the non-periodic elastic homogenization (Capdeville et al.
164 2010; Guillot et al. 2010). Thus, no hypothesis such as periodicity or stochastic invariance is
165 required on the fine-scale medium. The misfit is defined by a distance between two effective
166 elastic media: the HFWI solution (taking the role of the observations) and the effective
167 equivalent of the fine-scale model (the predictions). The space of fine-scale models is defined
168 with an object-based approach to reduce the number of inference parameters. Because the
169 solution could be non-unique, we use a Bayesian formalism, where an ensemble solution is
170 sampled with a Markov-Chain Monte Carlo (MCMC) algorithm, thus allowing us to estimate
171 uncertainties.

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

177 We observe the following advantages to split the multiscale inversion into two steps: in
178 the first step, recovering an homogenized elastic medium is in general a better-posed problem
179 than conventional FWI, because the space of acceptable elastic models is contained. We only
180 need to search for smooth solution models, the effective anisotropy contains the information
181 on the small-scale structure seen in the waveform data. This lowers the computational cost of
182 wave propagation if the true medium contains fine-scale heterogeneities or discontinuities. It
183 also mitigates artifacts induced by incorrect parameter restrictions. The level of smoothing is
184 naturally given by the wavelength of homogenization. We introduce as little prior information
185 as possible at this stage, as it would require more advanced mathematical treatments of
186 wave propagation and usually greater computational cost. The second step, the inverse
187 homogenization, can be less well-posed. But it presents the great advantage to having a
188 much cheaper forward model, which makes possible the use of statistical or global search
189 approaches. For example, using a Bayesian framework facilitates the incorporation of *a priori*

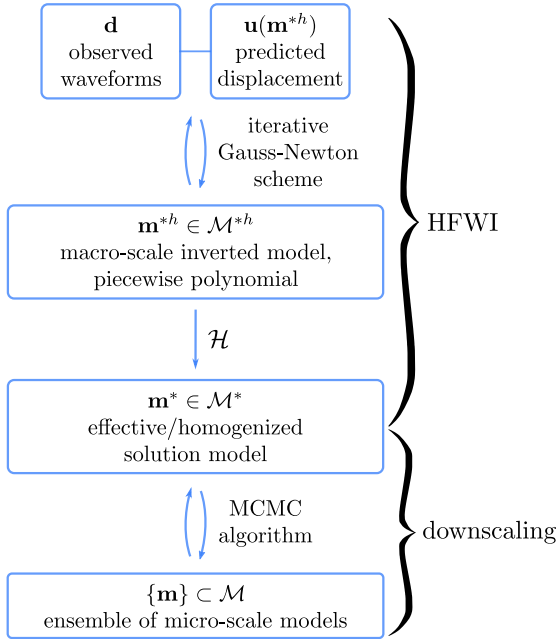


Figure 1. Notations used for the two-stage method proposed here. We invert the waveform data d for a macro-scale effective medium \mathbf{m}^* . \mathbf{m}^{*h} is an intermediate product of HFWI required by the discretization of the problem. Next, we seek for micro-scale models \mathbf{m} having a effective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

190 information. Furthermore, the data is now “localized”: the effective properties at one location
 191 depend only on the fine-scale properties in the neighborhood. This potentially allows to solve
 192 the problem using a large ensemble of tools that are prohibited in waveform inversion (e.g.
 193 Nawaz & Curtis 2016).

194 2 HOMOGENIZATION THEORY

195 This section summarizes the elastic homogenization theory used in this study. More detailed
 196 descriptions can be found in Capdeville et al. (2010); Guillot et al. (2010). We define the en-
 197 semble of elastic media \mathcal{M} . Each $\mathbf{m} \in \mathcal{M}$ is a vector of micro-scale parameters characterized
 198 by its density ρ and its elastic tensor \mathbf{c} for every position in the medium. In the following,
 199 \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, *i.e.* 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 the fine-scale medium is beneficially replaced by a long wavelength effective equivalent.

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 makes it possible to find the effective solution 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 weakly converges toward the effective solution as ε tends toward 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 choice of λ_0 simply determines the minimum length-scale included in the given medium. A smaller

223 λ_0 corresponds to a more detailed structure thus a more accurate solution for the effective
 224 displacement \mathbf{u}^* , whereas a larger λ_0 involves a smoother medium and imprecise solution.
 225 We can introduce:

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

226 which measures the scale separation position in the seismic model.

227 We usually consider that the wavefield “sees” the medium at the scale $\lambda_{\min}/2$, that is
 228 $\varepsilon_0 = 0.5$. Medium heterogeneities of scale $\lambda_0 \ll \lambda_{\min}$ are seen by the wavefield only through
 229 their homogenized/effective properties. Nevertheless, for some strongly scattering media, it
 230 can happen that an ε_0 smaller, or even much smaller, than 0.5 might be required to model
 231 accurately a long coda wave. We refer to Capdeville et al. (2010) for numerical examples.

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

237 Homogenization theory introduces two space variables, macroscopic variations with the
 238 variable \mathbf{x} , and microscopic variations with $\frac{\mathbf{x}}{\varepsilon_0}$, and assumes they can be treated as indepen-
 239 dent variables. A physical quantity depending on macroscopic and microscopic variations
 240 and on time, such as displacement, is written $\mathbf{u}(\mathbf{x}, \frac{\mathbf{x}}{\varepsilon_0}, t)$. The homogenized solution of the
 241 wave equation at the scale of ε_0 is the leading order of the displacement asymptotic expan-
 242 sion and is written \mathbf{u}^* . It can be demonstrated to be independent of the small scale variable
 243 $\frac{\mathbf{x}}{\varepsilon_0}$. The asymptotic expansion to the first order of the true displacement is:

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

244 $\boldsymbol{\chi}$ 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. $[\epsilon(\mathbf{u}^*)]_{ij} = \frac{1}{2}(\partial_i u_j^* + \partial_j u_i^*)$ is the strain tensor of the effective displacement. In practice, (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:

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

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 explicated here (see Capdeville & Marigo (2012) for the treatment of boundary conditions). This equation is analogous to the classical wave equation except that quantities have been replaced by their effective equivalent. 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 equation 4 using a standard wave equation solver. The accuracy of approximating \mathbf{u} by \mathbf{u}^* is given by equation 3: solving the wave equation in the media \mathbf{m} or in \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 media, Backus (1962) showed that the correct effective medium is obtained by filtering non-linear combinations of the

267 fine-scale elastic parameters. In the general case, there exist no analytic solution to the
 268 effective properties. Using homogenization theory, they are obtained following the procedure
 269 described hereinafter.

270 We define effective properties of an elastic model $\mathbf{m} \in \mathcal{M}$ at the scale of ε_0 as $\mathbf{m}^* = (\rho^*$,
 271 $\mathbf{c}^*) \in \mathcal{M}^*$. \mathbf{m}^* is spatially smooth at the scale of ε_0 , or in other words it does not contain
 272 wavelengths smaller than λ_0 . Thus the space of effective model \mathcal{M}^* is a finite dimensional
 273 space. Its dimension is proportional to $(f_{\max}/\varepsilon_0)^d$ with d the dimension (2- or 3-D) of the
 274 problem. In the following, the relation between the fine and large scales is summarized by
 275 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:

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

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:

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

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

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

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

3 HOMOGENIZED FULL WAVEFORM INVERSION

We now present the elements of the homogenized full waveform inversion used in this study (Capdeville & Métivier 2018). We invert seismic waveforms \mathbf{d} for a macro-scale 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 full waveform inversion

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 2D, and 21 parameters in 3D).

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 $\mathcal{H}(\mathcal{M}^{*h}) \subset \mathcal{M}^*$ (figure 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

305 polynomial approximation of degree N_h in each direction. In contrary to \mathbf{m}^* , \mathbf{m}^{*h} may not
 306 be smooth at the scale ε_0 : we may have $\mathbf{m}^{*h} \notin \mathcal{M}^*$.

307 To ensure that the final model \mathbf{m}_f^* is in \mathcal{M}^* , two strategies are possible. A first one is
 308 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
 309 next model $i + 2$ from \mathbf{m}_{i+1}^* . This method acts as a natural regularization of the problem.
 310 A second possibility is to compute each model iteration in the space \mathcal{M}^{*h} and project only
 311 the final model \mathbf{m}_f^{*h} into $\mathbf{m}_f^* \in \mathcal{M}^*$. We rely on the second method in all the inversions
 312 performed in this study, because the regularization at each step was not required to make
 313 the algorithm converge for our examples.

314 **3.2 Waveform inversion strategy**

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

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

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

321 Similar equations E^{*h} and E^* can be written by replacing \mathbf{m} with \mathbf{m}^{*h} or \mathbf{m}^* if the
 322 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
 323 presented hereinafter.

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

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

326 where \mathbf{I} is the identity matrix. The Fréchet derivatives \mathbf{F}^i of the displacement \mathbf{u} with re-

spect 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 micro-scale (or high resolution) information from macro-scale (or low-resolution) variables. In the present 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 equivalent. Therefore, we cast the problem in a Bayesian framework, where we look for an ensemble of plausible solutions. It is described by the following equation:

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

understood in a statistical sense, *i.e.* where ϵ is a random variable representing the errors attributable to the approximate estimation of \mathbf{c}^* in the HFWI. \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. We neglect the theoretical errors of the operator \mathcal{H} , expecting that it should be orders of magnitude smaller than ϵ 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 characterizes the mismatch between the solution of the HFWI \mathbf{m}^* and the effective properties of the true earth \mathbf{m}_t^* , *i.e.* the meaning of \approx in equation 9. Errors in the HFWI solution may have multiple sources, such as: noise on the waveform data, lack of data coverage, damping, convergence toward a local minima, etc. Solving the full waveform inversion 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-fit” 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}_t^* 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 , e.g. 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 equation 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 2D 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 $\hat{\Sigma}_{\text{r}}$, and introduce a scaling parameter h setting $\Sigma = h^2\hat{\Sigma}_{\text{r}}$, 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 misestimation in the amplitude of errors. An

399 inferred value of h close to 1 shows that the estimation $\hat{\Sigma}_r$ given by the user is good. A value
 400 $h > 1$ indicate that some additional parts of the data are not explained by our modeling,
 401 due for example to theoretical errors. In this case, the negative log-likelihood becomes:

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

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

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

405 For a solution space of dimension N , we need to estimate the $N \times N$ parameters. For
 406 now, we rely on a sampling-based estimation of $\hat{\Sigma}^{-1}$ similar to the approach in Hansen et al.
 407 (2014). It consists of first generating a large ensemble of error realizations, from which a
 408 Gaussian model is inferred. We generate M noise realizations on the waveform data \mathbf{d} and,
 409 for each, estimate a HFWI solution. As such, we only consider errors associated to the noise
 410 in waveforms and mapped to the HFWI solution through the tomographic inversion. First,
 411 we calculate the empirical covariance matrix S of the M HFWI solution models. In order to
 412 get a stable inverse of S , a regularized version $\hat{\Sigma}_r$ is computed using the shrinkage formula
 413 (Ledoit & Wolf 2004):

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

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

420 4.4 Prior information

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

425 Prior information is enforced in two different ways: 1) by choosing a specific parametriza-
 426 tion of the model using geometrical objects, and 2) by choosing a relevant prior probability
 427 distribution on these parameters. Translated into notations, the elastic model \mathbf{c} is con-
 428 structed from a vector \mathbf{Z} of n latent variables. This process can be described as a function
 429 $\mathbf{c}(\mathbf{Z})$. Subsequently, a prior distribution is imposed directly on \mathbf{Z} . Hence, we will write the
 430 prior distribution $\pi_{\text{pr}}(\mathbf{Z})$ and keep a similar notation for the other probability distributions
 431 described hereinafter.

432 Because this study deals only with synthetic tests, construction of realistic priors is out
 433 of scope. We will consider a somewhat idealized prior distribution, using an object-based
 434 parametrization made of ellipses. The vector \mathbf{Z} contains parameters describing the position
 435 and shape of each object, as well as the hyper-parameter h , the unknown scaling factor of
 436 the covariance matrix. For the problem at hand, the ellipses are representing cavities. They
 437 are described by their position (x^i, y^i) , the long and short axes (r_1^i, r_2^i) and the angle from
 438 horizontal α^i for each ellipse $1 \leq i \leq n_e$. All parameters are independent. We use a uniform
 439 prior distribution for the object parameters, and a Jeffrey prior for h (*i.e.* proportional to
 440 $1/h$).

441 4.5 Sampling algorithm

442 The posterior distribution is sampled using the reversible-jump Markov chain Monte Carlo
 443 (rj-MCMC) sampler (Green 1995). Rj-MCMC is a transdimensional sampler, where the
 444 number of parameters (here the number of elliptical cavities) is treated as an unknown in
 445 the problem. In our case, it amounts to consider a varying number n_e of objects. Trans-
 446 dimensional inversions are popular in seismic imaging, as the spatial discretization of the

447 solution space is adapted to data quality and coverage (Sambridge et al. 2013). The present
 448 parametrization allows to make some simplifications in the algorithm, a more complete de-
 449 scription of the rj-MCMC algorithm for seismic imaging is outlined in Bodin & Sambridge
 450 (2009).

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

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

458 \mathbf{J} is the Jacobian matrix of the transformation from \mathbf{Z} to \mathbf{Z}' . Its determinant $|\mathbf{J}|$ is
 459 equal to 1 for the transformations considered in this study (see Bodin & Sambridge (2009)
 460 for details). We use a “burn-in” period to ensure the convergence of the Markov chain.
 461 That means, the first samples are discarded until the Markov chain is expected to have
 462 reached its equilibrium. The following samples \mathbf{Z} (and hence \mathbf{c}) constitute samples from the
 463 posterior distribution π , from which any statistical quantity such as mean or variance can
 464 be computed.

465 The proposal distribution comprises the following types of moves:

- 466 (i) perturb a parameter describing an ellipse (position, axis length, angle) using a Gaus-
 467 sian probability density centered in 0;
- 468 (ii) perturb the hyper-parameter h using a similar probability density;
- 469 (iii) birth: add a new ellipse to the vector of parameters \mathbf{Z} , each of the new parameters
 470 being drawn from the prior distribution;
- 471 (iv) death: remove an ellipse, hence its parameters, from \mathbf{Z} .

472 For moves of type (i) and (ii), the proposal distribution is symmetric, $q(\mathbf{Z}'|\mathbf{Z}) = q(\mathbf{Z}|\mathbf{Z}')$,

473 which can be used to simplify expression (17). For moves of type (iii) and (iv), the proposal
 474 equals the prior distribution, they cancel out and the second term of the right-hand side
 475 in equation (17) becomes the ratio of the likelihoods: we do not need to explicit the prior
 476 distribution but only to sample from it (Mosegaard & Tarantola 1995).

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

482 5 SYNTHETIC TESTS FOR THE INVERSE HOMOGENIZATION

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

487 5.1 Setup

488 We consider a 2D square homogeneous plate with 120 m sides (all boundaries are reflective).
 489 The first synthetic model contains 4 ellipsoidal cavities having minor and major diameters
 490 of sizes 2 and 6 m respectively, as seen in figure 2. We compute the effective equivalent of
 491 this model using the homogenization technique described in paragraph 2.2, using a scale
 492 separation parameter $\lambda_0 = 16.6$ m. This value of λ_0 will be kept for the remainder of the
 493 study. We present in figure 2 the resulting shear-wave velocity V_S and a measure of anisotropy
 494 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}
 495 (Fedorov 2013; Browaeys & Chevrot 2004) and $\|\cdot\|_2$ the Euclidean (or Frobenius) matrix
 496 norm. While the original model is discontinuous and isotropic, its effective equivalent is
 497 smooth and anisotropic.

498 We add spatially correlated Gaussian noise to the elastic parameters of the homogenized

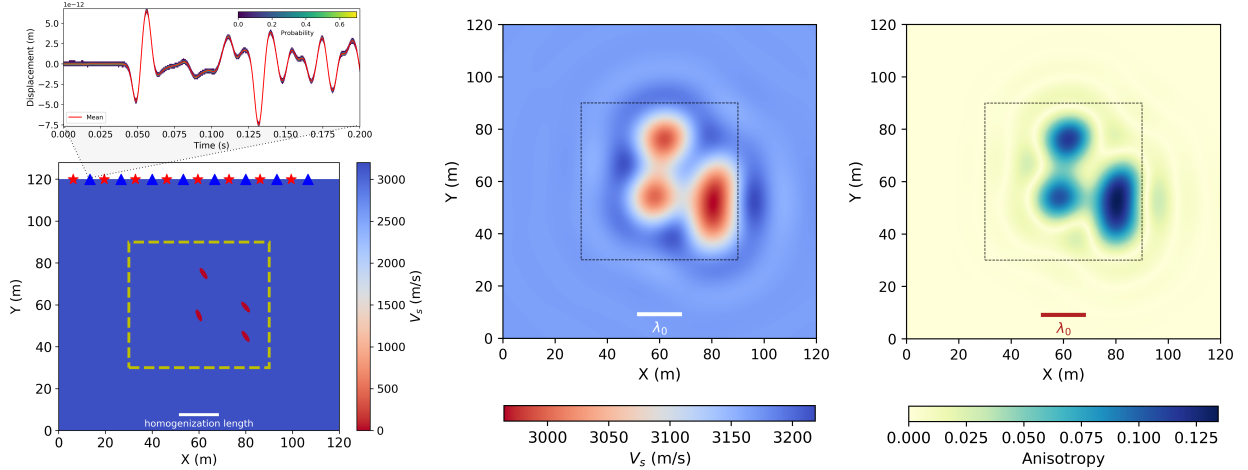


Figure 2. Synthetic model consisting in a homogeneous rectangular domain containing ellipsoidal cavities: (left) the true model and associated synthetic waveform data, (right) shear-wave velocity and anisotropy of its effective equivalent. Red stars show the sources locations, blue triangles the receivers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

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

503 5.2 Results

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

509 For each realization in the Markov chain, we need to build a discretized version of \mathbf{m}
 510 before calculating $\mathcal{H}(\mathbf{m})$ with a finite element method. The elements are triangular and
 511 the interpolation is based on the Fekete points using high order polynomials. The mesh is
 512 generated using the Gmsh tool (Geuzaine & Remacle 2009). Triangular elements are more
 513 versatile than spectral elements meshes, hence allow to generate a new mesh at each iteration

514 of the Markov chain. But this way of proceeding will meet its limits in 3D where automatic
 515 generation of conforming meshes is difficult. Alternative computing methods will be reviewed
 516 in the discussion.

517 The result of the inversion is the posterior distribution of the parameter vector \mathbf{Z} . A visual
 518 representation is obtained by computing the probability of being inside a cavity (figure 3):
 519 it corresponds, at each spatial position of the image, to the number of times this position
 520 is inside an ellipse divided by the number of sampled models. The algorithm is able to
 521 recover the presence of four cavities, as well as their general shape and orientation, which
 522 was not obvious by looking only at the effective model of figure 2. We test the inversion
 523 on an additional setup with 6 cavities of various sizes and orientations. Similarly, the true
 524 model is well recovered. The number of ellipses, n_e , is also an inferred parameter in the rj-
 525 MCMC algorithm. In both proposed tests, its posterior marginal distribution, shown figure
 526 3, indicates a maximum at the correct value of the true model.

527 **5.3 The role of prior information**

528 In the absence of prior information, the inverse homogenization problem is necessarily non-
 529 unique, and does not significantly improve the understanding of the medium. For example,
 530 the observation \mathbf{m}_n^* would be itself a solution (up to the loss of information in the filtering
 531 operation of equation 8). Therefore, we shall emphasize that we are only able to recover
 532 the small scale structure because the chosen parametrization and prior distribution incor-
 533 porate sufficient constraints. To highlight the importance of prior information, we present
 534 an additional test with different choices for the construction of small scale models.

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

540 Figure 5 shows the inverse homogenization results for three different parametrizations,

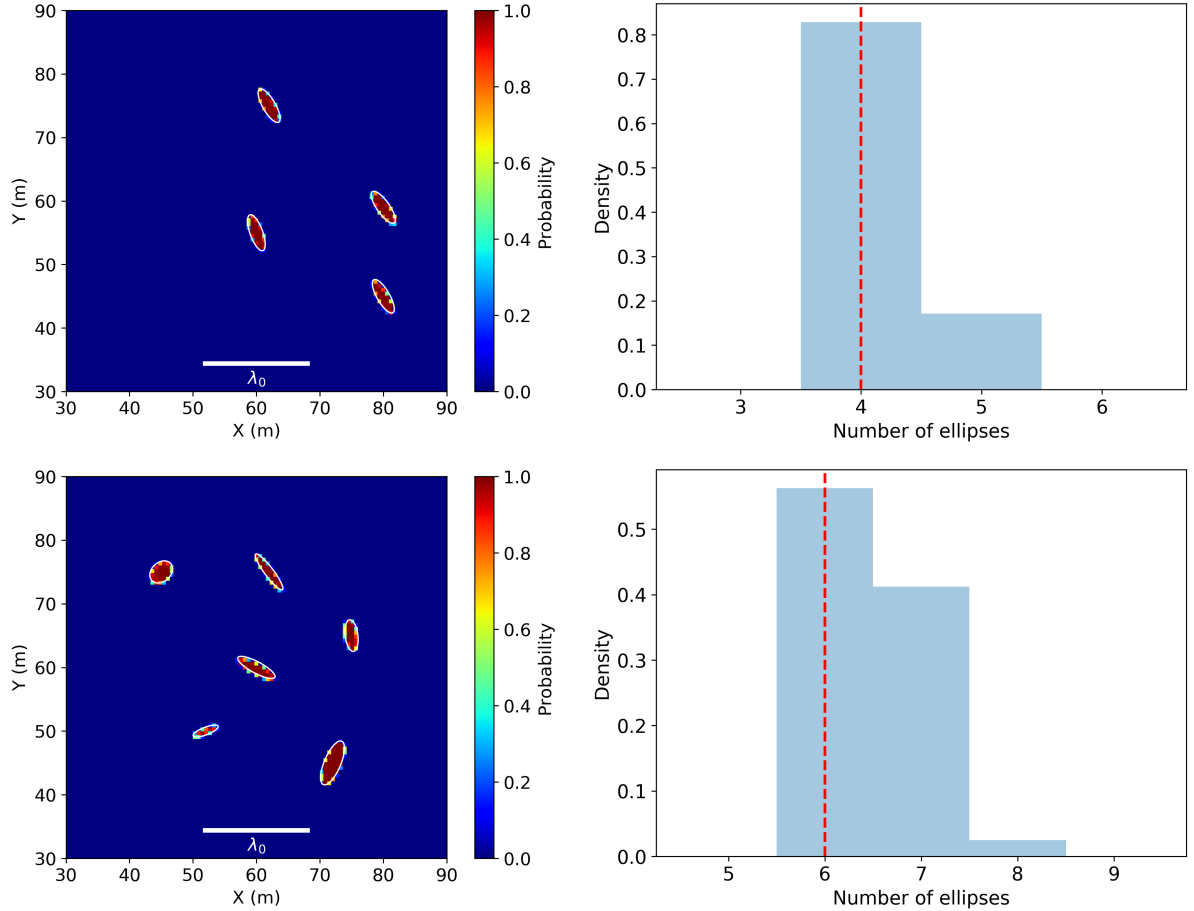


Figure 3. Results of the inverse homogenization. The top left panel shows the probability of recovering void at each spatial point in the model, the top right panel the marginal distribution of the number of ellipses n_e used to parametrize the small-scale model. The true position and shape of the cavities is shown as white ellipses. The bottom panels display similar results for a true small-scale model with 6 cavities of various shape and orientation.

541 from left to right: (a) we consider a stronger prior information, using a fixed-dimensional
 542 MCMC algorithm and choosing the number of ellipses n_e constant and equal to four (the
 543 positions and geometries of the 4 ellipses are still unknown variables to be inverted for); (b)
 544 we use the rj-MCMC algorithm as previously; (c) we parametrize the small scale model with
 545 a regular grid. Each cell has a 1x1 m size and can contain either the elastic properties of
 546 the homogeneous plate or the cavities. We invert for a 30x30 m area, hence comprising 900
 547 parameters, using also a fixed-dimensional MCMC algorithm.

548 As expected, the quality of the inversion is correlated to the incorporated prior informa-
 549 tion. Algorithm (a), which includes the most information and imposes the correct number

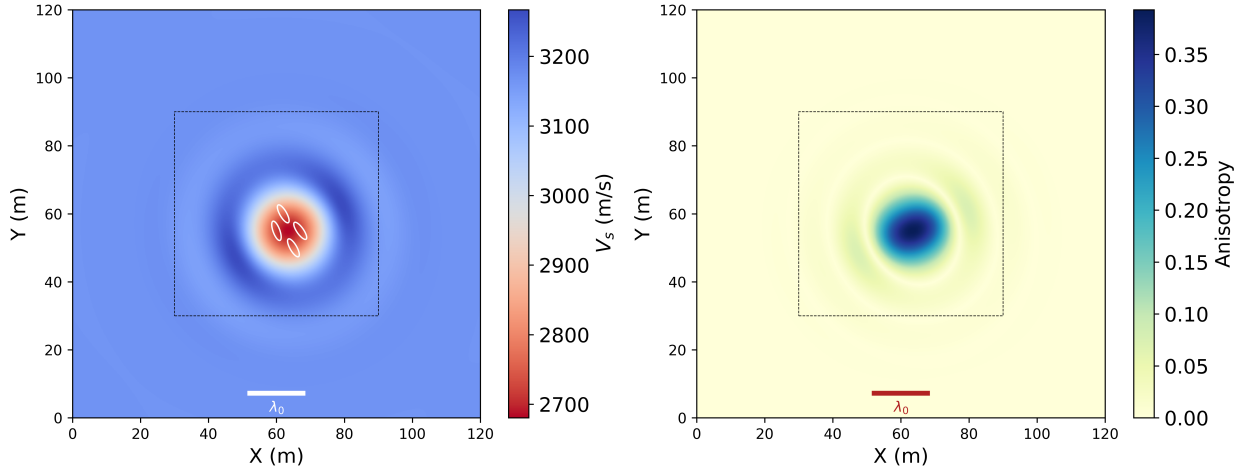


Figure 4. Effective seismic properties of a synthetic model having closer ellipsoidal cavities. The true position and shape of the cavities are represented on the left panel as white ellipses.

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 favored 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 noticeable 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 homogeniza-

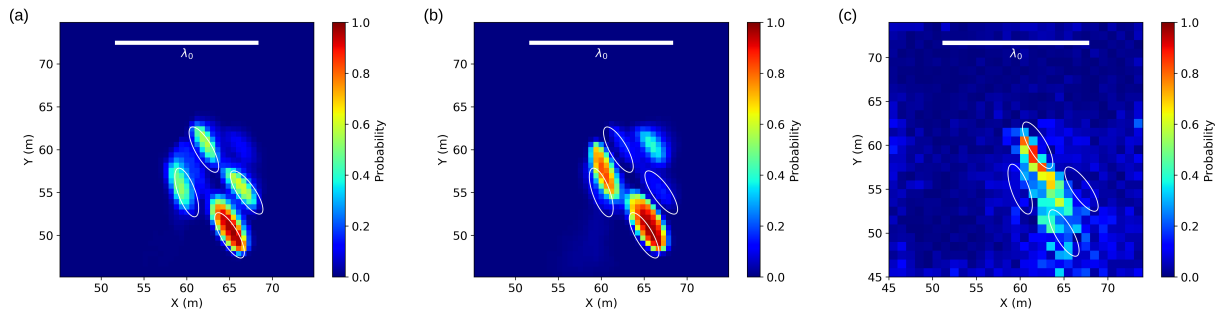


Figure 5. Results of the inverse homogenization for the synthetic model of figure 4, using three different parametrization: (a) a fixed number of ellipses $n_e = 4$, (b) a variable number n_e , (c) a regular grid parametrization.

tion is beneficial for 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 figure 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.

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 modeling 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 2D spectral element solver. Gaussian noise with an amplitude representing 5% of the signal standard deviation is added to the waveform data. $M = 1000$ noise realizations are generated. HFWI

586 is performed independently for each realization in order to produce an ensemble of M to-
 587 mographic solutions, and evaluate uncertainties on the model parameters. An example of
 588 synthetic waveform data distribution is displayed figure 2.

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

593 6.2 Full waveform inversion solutions

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

599 We compare two possible parametrizations for \mathcal{M}^{*h} : 1) a fully anisotropic parametriza-
 600 tion (ρ, \mathbf{c}) where the full elastic tensor is inverted for as prescribed in HFWI; and 2) an
 601 isotropic parametrization (ρ, V_P, V_S) corresponding to a “conventional” FWI (although a
 602 different, spatially denser discretization would have probably been used for a real FWI).
 603 One example of resulting inverted models \mathbf{m}^{*h} and \mathbf{m}^* for both parametrizations are com-
 604 pared in figure 6.

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

612 Once projected into \mathcal{M}^* , results obtained in the case where \mathbf{m}^{*h} is fully anisotropic

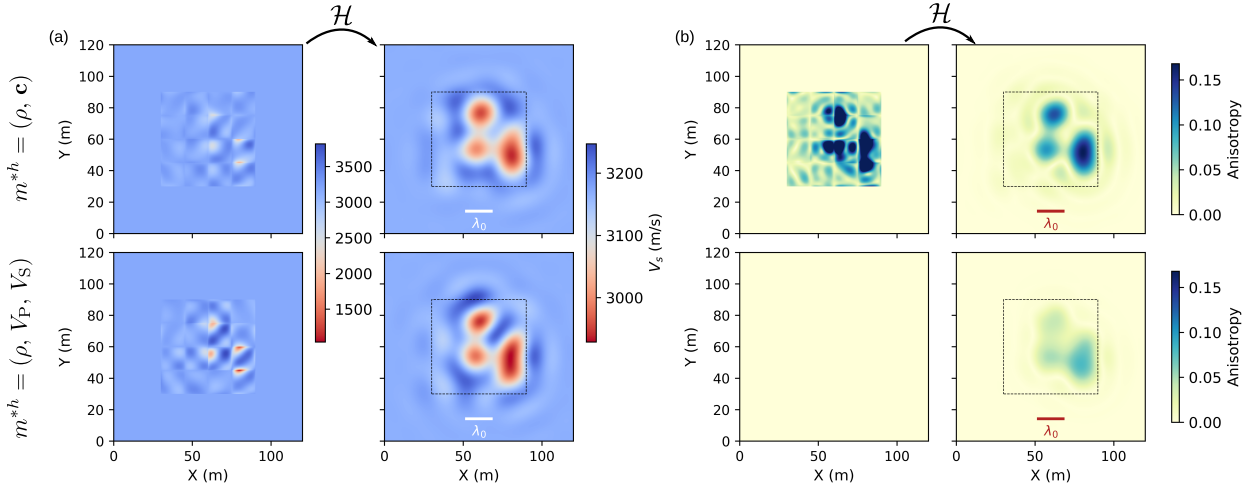


Figure 6. Full waveform inversion results presented in terms of (a) V_S and (b) anisotropy. The top row shows the images obtained with an anisotropic parametrization of \mathcal{M}^{*h} , and the bottom with the isotropic one. The left column of each panel displays the results in the space \mathcal{M}^{*h}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

613 show only circular anomalies, while the strong anisotropy indicates that information about
 614 the shape is preserved. On the contrary, \mathbf{m}^* obtained from an isotropic \mathbf{m}^{*h} depicts weak
 615 anisotropy while the isotropic parameters V_P and V_S already display some structure. The
 616 spatial discretization is too loose to resolve all the small scales and especially the shape
 617 of the anomalies. This example suggests that a fully anisotropic parametrization preserves
 618 more information and is beneficial in the HFWI method. Data misfit reduction is about
 619 76 % versus 61 % for the anisotropic and isotropic parametrizations of \mathbf{m}^{*h} respectively.
 620 This is not surprising since the anisotropic parametrization has a higher complexity.

621 6.3 Uncertainties in HFWI

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

626 We consider $M = 1000$ HFWI resulting models $(\mathbf{m}_1^*, \dots, \mathbf{m}_M^*)$ obtained for each noise
 627 realization on the data and examine their statistical properties. In the inverse homogeniza-

tion 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 $(\mathbf{m}_1^*, \dots, \mathbf{m}_M^*)$ to the one of \mathbf{m}_t^* (figure 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.

We now want to verify the Gaussian hypothesis on ϵ in equation (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 figure 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 pre-
sented paragraph 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

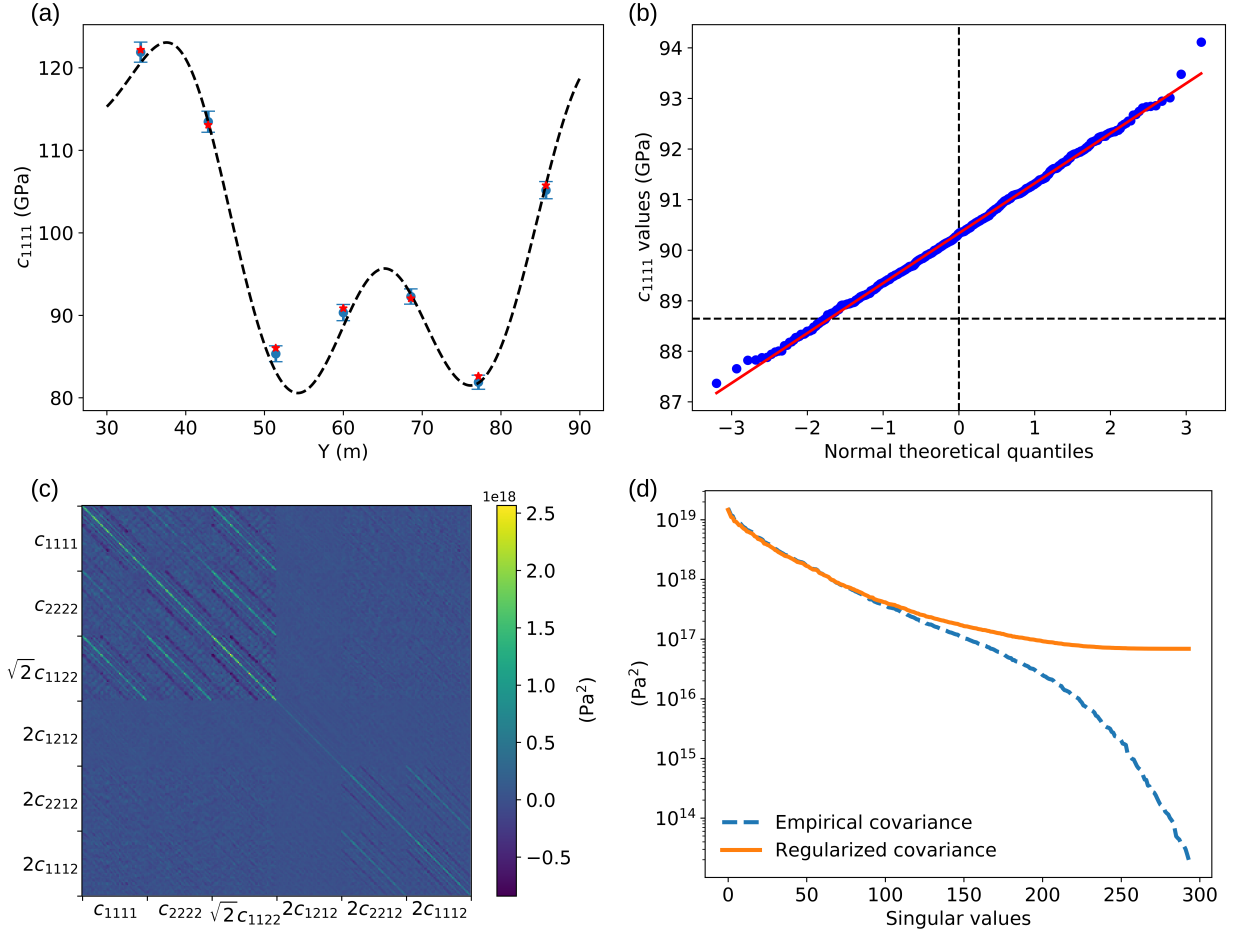


Figure 7. Uncertainty estimation in HFWI: (a) Comparison of the effective equivalent of the true model (dashed black line) presented figure 2 to the distribution of the M images obtained by HFWI, for parameter c_{1111} in a vertical cut at the middle of the model. The blue dot show the mean of the M 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 equation (16) for a parameter $\alpha = 0.06$. Its singular value decomposition (d) exposes the effect of the regularization.

656 matrix, and the effect of the regularization on its eigenvalues is shown figure 7c and figure 7d
 657 respectively. The analysis of its structure indicates that variables are spatially correlated on
 658 a length scale λ_0 . Inter-parameter correlations also exist, for example between parameters
 659 c_{2222} and c_{1122} .

660 The covariance matrix can also provide information about the effective elastic parameters

661 that are best constrained by the waveform data through relative uncertainties. The relative
 662 uncertainty of a parameter is defined as its absolute uncertainty divided by its value. It
 663 is usually considered in seismic imaging that only a restricted part of the elastic tensor is
 664 resolved (for example, only isotropic or transverse isotropic media, depending on the data
 665 type). As expected, the parameters involved in isotropic properties are best resolved ($\sim 1\%$
 666 of relative uncertainty). c_{1112} and c_{2212} are non-zero only in the presence of anisotropy and
 667 have at least 5% of relative uncertainty.

668 6.4 Downscaling results

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

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

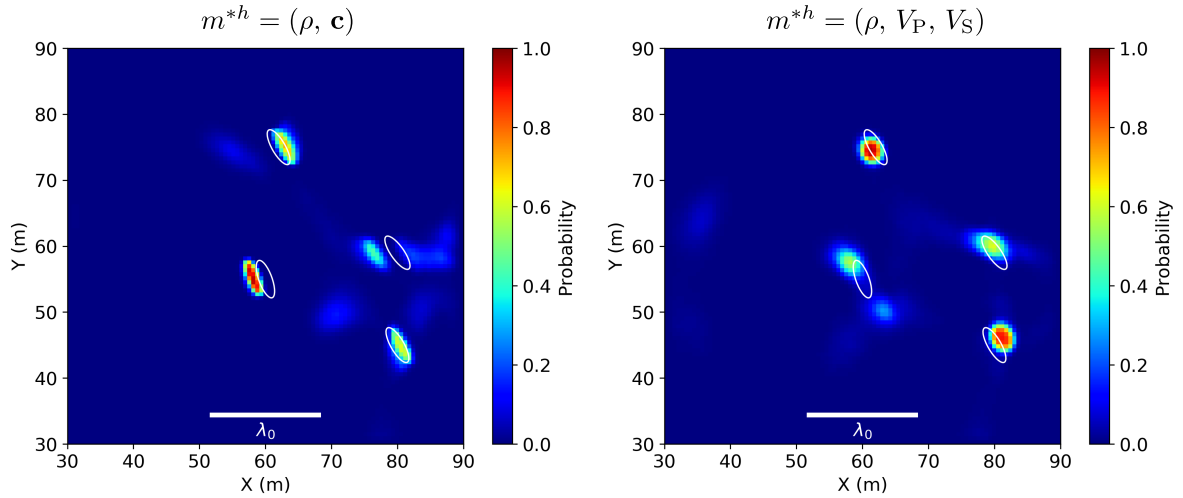


Figure 8. Probability of recovering void at each spatial point in the model. The left panel presents the results when \mathcal{M}^{*h} is parametrized with the full elastic tensor, and the right one for the isotropic parametrization. The true position and shape of the inclusions is shown as white ellipses.

688 To look at the results in more detail, the posterior marginal 1D distributions for the
 689 number of ellipses n_e , the ellipse axes and their orientation are shown in figure 9, top row.
 690 The mode for the distribution of n_e is much higher than the true value of four. The algorithm
 691 uses additional ellipses to match some weak anomalies in \mathbf{m}^* . They are in general small and
 692 circular, hence with an arbitrary orientation. The modes for the distribution of axes and
 693 orientation are close to the true values. The hyper-parameter h is only slightly higher than
 694 1, suggesting that the covariance estimation is reasonable (figure 10).

695 6.5 Isotropic downscaling

696 We repeat the inverse homogenization this time using HFWI solution model obtained from
 697 an isotropic parametrization in \mathcal{M}^{*h} (figure 6, bottom row). In this test, the spatial dis-
 698 cretization of the model in the HFWI is the same as in the anisotropic case (we use the
 699 same mesh), but only 2 elastic parameters are inverted for (V_P and V_S), leading to a re-
 700 duced number of unknowns. Note that, once homogenized, the resulting effective medium
 701 in \mathcal{M}^* is not strictly isotropic. The waveform inversion uses the discontinuous nature of
 702 the discretization mesh in \mathcal{M}^{*h} to preserve a (small) part of the fine-scale information. The
 703 effective medium is weakly anisotropic (figure 6).

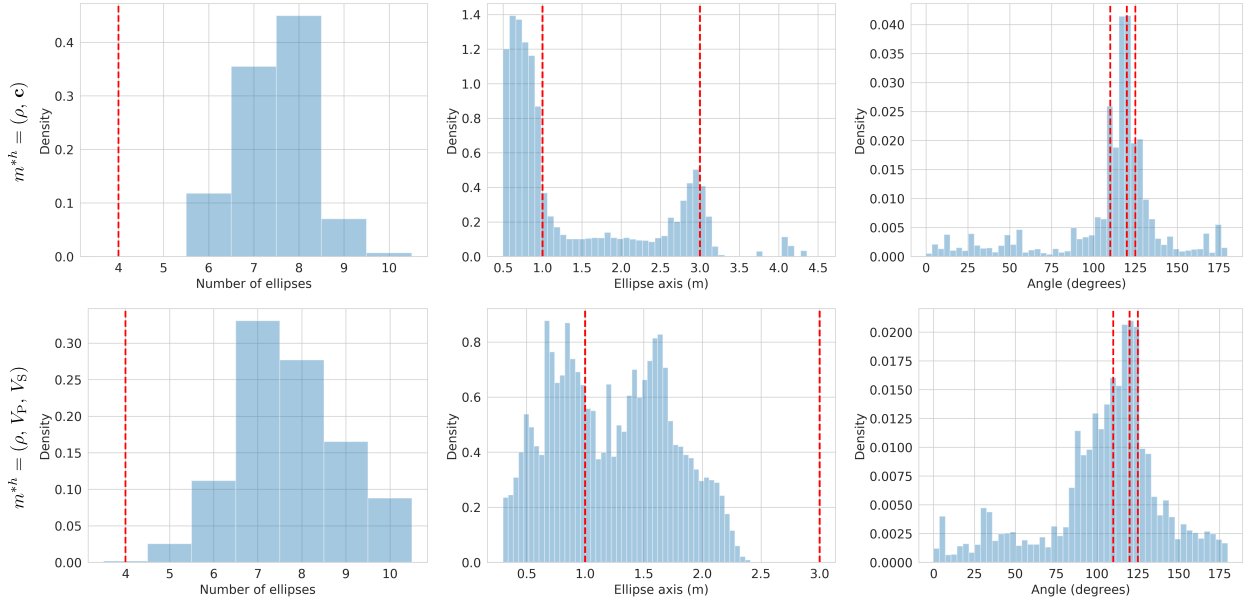


Figure 9. Posterior distribution of the latent variables in Z . The top/bottom panels correspond to anisotropic and isotropic parametrizations \mathcal{M}^{*h} respectively. For each variable type, all the ellipses are combined on the same histogram. The y-axis shows the probability density associated 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 panels display the angle with respect to the horizontal.

704 With a finer spatial discretization, increasing the number of unknowns to the same
 705 amount as in the anisotropic parametrization case, one can recover more details in the
 706 model. Once homogenized, such a finer scale isotropic model can retrieve the correct effective
 707 anisotropy. Some examples are presented in Capdeville & Métivier (2018). The drawbacks are

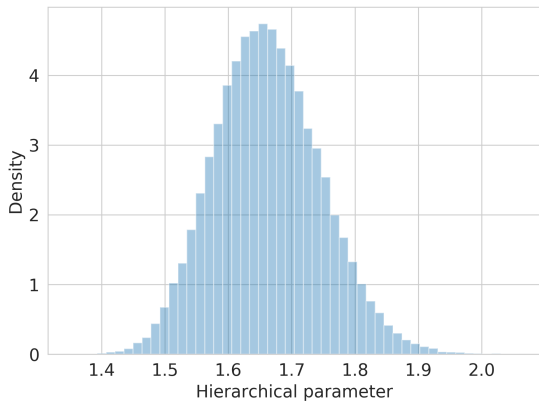


Figure 10. Posterior distribution of the hyperparameter h .

708 an increased computational cost for wave propagation, and a more challenging convergence
709 of the waveform inversion scheme.

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

719 7 DISCUSSION

720 The homogenized full-waveform inversion of Capdeville & Métivier (2018) restricts the imag-
721 ing solution space to effective media, which are not suited for geological interpretation. For
722 this purpose, an additional downscaling step is required. A general summary of the complete
723 procedure is presented in figure 11. Downscaling can be formulated as an inverse problem
724 where the forward operator is the non-periodic homogenization. This study achieves the
725 inverse homogenization of an elastic media and clarifies the possible applications of the
726 method. Because it was restricted to synthetic tests, several additional points need to be
727 discussed.

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

733 The assumption justifying the downscaling step is that the HFWI solution model is the
734 effective equivalent of the true Earth model at scale ε_0 (equation 9). One necessary condition

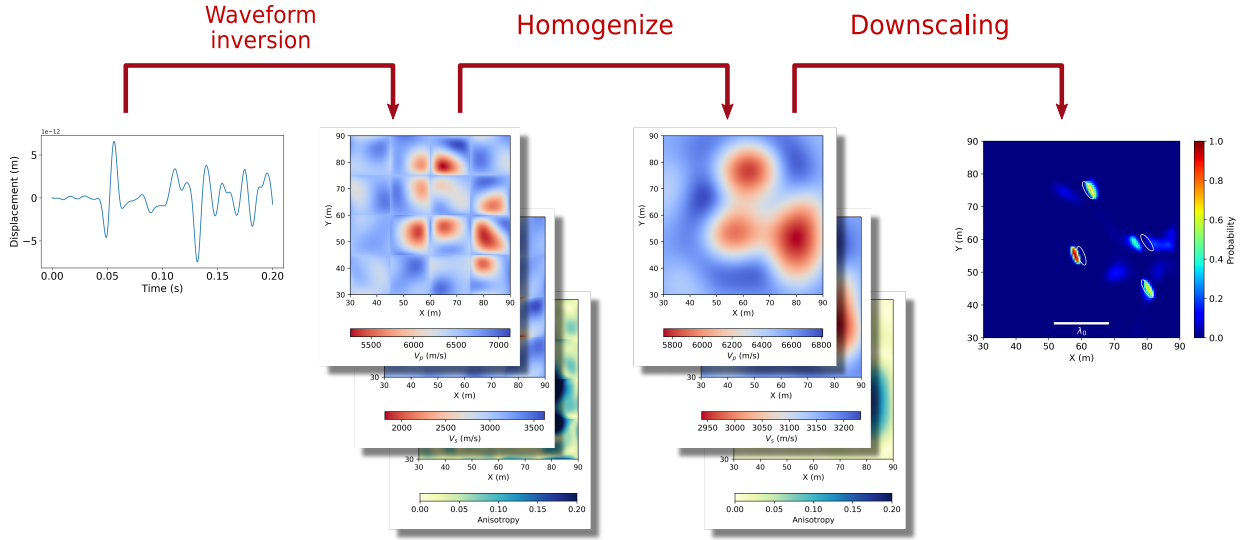


Figure 11. Summary of the seismic imaging method used in the study. The different steps corresponding to figure 1 are illustrated for the reference inversion, here shown in terms of P-wave velocity.

735 to meet this assumption is to find the global minimum of the HFWI. Note that this issue
 736 is not inherent to HFWI but present in all seismic tomography problems. Constraining the
 737 solution to homogenized models is a natural regularization because it reduces the size of the
 738 solution space and yet guarantees the same data fit. It avoids the use of biased smoothing
 739 constraints on the parameter space. More sophisticated strategies used in FWI to guide the
 740 algorithm convergence, such as frequency-continuation, are also possible to implement in
 741 HFWI (Capdeville & Métivier 2018).

742 We formulated the inverse homogenization problem in a Bayesian sense, where the so-
 743 lution is an ensemble of plausible fine-scale models. In this framework, uncertainties on
 744 the HFWI solution models are required. We computed them using a substitute method,
 745 performing multiple HFWI inversions for multiple noise realizations on the synthetic data.
 746 Uncertainty estimation in FWI is a current topical issue (Zhu et al. 2016; Sen & Biswas 2017;
 747 Thurin et al. 2019). One can expect viable solutions in the following years, at least for local
 748 uncertainties around the global minimum. In addition, we observed that the spatial corre-
 749 lation between parameters in effective models is specific and linked to the homogenization
 750 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 full-waveform inversion problem and the statistical model assumed in the inverse homogenization (equation 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-fit” model, according to some *a priori* expert knowledge, is sometimes favored. 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 equivalent 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 2D section of a 3D 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

778 as computing \mathcal{H} only on the perturbed sub-area, is more straightforward and can further
 779 spare computation time.

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

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

798 Structures containing cavities, fluid inclusions or cracks will likely produce significant
 799 effective anisotropy, hence constitute appropriate use cases. They will exist at the engineer-
 800 ing and subsurface scales, where, additionally, strong prior information is usually available.
 801 Although not explored in this study, applications at the continental or global scale are envi-
 802 sioned. In seismology, a long standing issue is the ability to distinguish effective anisotropy
 803 (e.g. induced by small scale heterogeneities or preferentially oriented faults) from intrinsic
 804 anisotropy (induced by anisotropic crystals) (e.g. Fichtner et al. 2013a; Alder et al. 2017). For

805 example, downscaling of global surface wave tomography models could help to discriminate
806 between effective and intrinsic anisotropy.

807 Moreover, our two-step approach is particularly suited to handle discontinuities in elastic
808 properties. FWI methods struggle to build discontinuous models, as most waveform solvers
809 rely on spectral-element methods with rectangular meshes that need to honor the geometry of
810 discontinuities. Inverse homogenization alleviates these limitations, because its forward part
811 solves an elasto-static problem with a standard finite-element method and a triangular mesh.
812 We showed a near-surface / engineering-scale application to image cavities, other possible
813 applications at larger scale are characterization of discontinuous geological structures such as
814 faults in the subsurface or study the depth and shape of major discontinuities in the Earth. A
815 typical example is the issue of crustal corrections in global tomography, required as details
816 within the crust cannot be resolved by long-period seismic waves. Inaccurate corrections
817 lead to artifacts in the less constrained part of the model (e.g. Marone & Romanowicz
818 2007). Instead of using corrections, one could first invert long period waves for a smooth
819 homogeneous earth without crust, and in a second time, together with extra information,
820 apply a downscaling approach to reconstruct crustal structures.

821 **8 CONCLUSION**

822 We have presented a proof of concept of the inverse homogenization of an elastic media,
823 applied to a void detection problem. The seismic imaging problem is solved in two steps.
824 The first one is a waveform inversion, where the solution space is restricted to effective
825 or homogenized media. Second, in the inverse homogenization step, the resulting effective
826 medium is used as data and we look for small-scale models having the corresponding ef-
827 fective equivalent. The main benefits are an easier convergence for the waveform inversion
828 in problems with strong heterogeneities and a better control on the prior information in
829 the inverse homogenization. The theory was presented in general terms and is applicable at
830 different scales. Thus, the benefits of an extension to other seismic imaging problems will
831 be studied in future work. Different implementations might be more appropriate depending

832 on the problem size and on the properties of the heterogeneities (continuity, solid or fluid,
 833 etc). Besides the items for discussion mentioned above, future work on inverse homogeniza-
 834 tion need to focus on building realistic prior models of geological media and adapting the
 835 inversion strategy to these models.

836 ACKNOWLEDGMENTS

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

841 REFERENCES

- 842 Afanasiev, M., Boehm, C., May, D., & Fichtner, A., 2016. Using effective medium theory to better
 843 constrain full waveform inversion, in *78th EAGE Conference and Exhibition 2016*.
- 844 Alder, C., Bodin, T., Ricard, Y., Capdeville, Y., Debayle, E., & Montagner, J., 2017. Quantifying
 845 seismic anisotropy induced by small-scale chemical heterogeneities, *Geophys. J. Int.*, **211**(3),
 846 1585–1600.
- 847 Almuheidib, A. M. & Toksöz, M. N., 2015. Imaging of near-surface heterogeneities by scattered
 848 elastic waves, *Geophysics*, **80**(4), A83–A88.
- 849 Backus, G. E., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. Geophys.*
 850 *Res.*, **67**(11), 4427–4440.
- 851 Bensoussan, A., Lions, J.-L., & Papanicolaou, G., 2011. *Asymptotic analysis for periodic struc-*
 852 *tures*, vol. 374, American Mathematical Soc.
- 853 Bernard, S., Monteiller, V., Komatitsch, D., & Lasaygues, P., 2017. Ultrasonic computed tomog-
 854 raphy based on full-waveform inversion for bone quantitative imaging, *Phys. Med. Biol.*, **62**(17),
 855 7011.
- 856 Bodin, T. & Sambridge, M., 2009. Seismic tomography with the reversible jump algorithm,
 857 *Geophys. J. Int.*, **178**(3), 1411–1436.
- 858 Bodin, T., Capdeville, Y., Romanowicz, B., & Montagner, J.-P., 2015. Interpreting radial
 859 anisotropy in global and regional tomographic models, in *The Earth’s Heterogeneous Mantle:*
 860 *A Geophysical, Geodynamical, and Geochemical Perspective*, pp. 105–144, eds Khan, A. & De-
 861 schamps, F., Springer.

- 862 Bozdağ, E., Trampert, J., & Tromp, J., 2011. Misfit functions for full waveform inversion based
863 on instantaneous phase and envelope measurements, *Geophys. J. Int.*, **185**(2), 845–870.
- 864 Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., & Pug-
865 mire, D., 2016. Global adjoint tomography: first-generation model, *Geophys. J. Int.*, **207**(3),
866 1739–1766.
- 867 Bretaudeau, F., Brossier, R., Leparoux, D., Abraham, O., & Virieux, J., 2013. 2d elastic full-
868 waveform imaging of the near-surface: Application to synthetic and physical modelling data sets,
869 *Near Surf. Geophys.*, **11**(3), 307–316.
- 870 Brossier, R., Operto, S., & Virieux, J., 2009. Seismic imaging of complex onshore structures by
871 2D elastic frequency-domain full-waveform inversion, *Geophysics*, **74**(6), WCC105–WCC118.
- 872 Browaeys, J. T. & Chevrot, S., 2004. Decomposition of the elastic tensor and geophysical appli-
873 cations, *Geophys. J. Int.*, **159**(2), 667–678.
- 874 Bui-Thanh, T., Ghattas, O., Martin, J., & Stadler, G., 2013. A computational framework for
875 infinite-dimensional bayesian inverse problems part i: The linearized case, with application to
876 global seismic inversion, *SIAM J. Sci. Comput.*, **35**(6), A2494–A2523.
- 877 Bunks, C., Saleck, F. M., Zaleski, S., & Chavent, G., 1995. Multiscale seismic waveform inversion,
878 *Geophysics*, **60**(5), 1457–1473.
- 879 Capdeville, Y. & Marigo, J.-J., 2007. Second order homogenization of the elastic wave equation
880 for non-periodic layered media, *Geophys. J. Int.*, **170**(2), 823–838.
- 881 Capdeville, Y. & Marigo, J.-J., 2012. A non-periodic two scale asymptotic method to take account
882 of rough topographies for 2-D elastic wave propagation, *Geophys. J. Int.*, **192**(1), 163–189.
- 883 Capdeville, Y. & Métivier, L., 2018. Elastic full waveform inversion based on the homogenization
884 method: theoretical framework and 2-D numerical illustrations, *Geophys. J. Int.*, **213**(2), 1093–
885 1112.
- 886 Capdeville, Y., Guillot, L., & Marigo, J.-J., 2010. 2-d non-periodic homogenization to upscale
887 elastic media for p–sv waves, *Geophys. J. Int.*, **182**(2), 903–922.
- 888 Capdeville, Y., Stutzmann, E., Wang, N., & Montagner, J.-P., 2013. Residual homogenization for
889 seismic forward and inverse problems in layered media, *Geophys. J. Int.*, **194**(1), 470–487.
- 890 Capdeville, Y., Zhao, M., & Cupillard, P., 2015. Fast Fourier homogenization for elastic wave
891 propagation in complex media, *Wave Motion*, **54**, 170–186.
- 892 Chen, Y., Wiesel, A., Eldar, Y. C., & Hero, A. O., 2010. Shrinkage algorithms for MMSE covari-
893 ance estimation, *IEEE Transactions on Signal Processing*, **58**(10), 5016–5029.
- 894 Devaney, A., 1984. Geophysical diffraction tomography, *IEEE Trans. Geosci. Remote Sens.*,
895 **22**(1), 3–13.
- 896 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, *The 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., 2014. Numerical methods for multiscale inverse problems, *arXiv preprint arXiv:1401.2431*.
- 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 *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*.
- 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.

- 932 Käüfl, P., Fichtner, A., & Igel, H., 2013. Probabilistic full waveform inversion based on tectonic
933 regionalization–development and application to the Australian upper mantle, *Geophys. J. Int.*,
934 **193**(1), 437–451.
- 935 Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P., 1983. Optimization by simulated annealing,
936 *Science*, **220**(4598), 671–680.
- 937 Kocur, G. K., Saenger, E. H., Grosse, C. U., & Vogel, T., 2016. Time reverse modeling of acoustic
938 emissions in a reinforced concrete beam, *Ultrasonics*, **65**, 96–104.
- 939 Komatitsch, D. & Vilotte, J.-P., 1998. The spectral element method: an efficient tool to simulate
940 the seismic response of 2D and 3D geological structures, *Bull. Seism. Soc. Am.*, **88**(2), 368–392.
- 941 Ledoit, O. & Wolf, M., 2004. A well-conditioned estimator for large-dimensional covariance ma-
942 trices, *J. Multiv. Anal.*, **88**(2), 365–411.
- 943 Li, X.-D. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear
944 asymptotic coupling theory, *J. Geophys. Res. Solid Earth*, **101**(B10), 22245–22272.
- 945 Marone, F. & Romanowicz, B., 2007. Non-linear crustal corrections in high-resolution regional
946 waveform seismic tomography, *Geophys. J. Int.*, **170**(1), 460–467.
- 947 Métivier, L., Brossier, R., Méridot, Q., Oudet, E., & Virieux, J., 2016. Measuring the misfit
948 between seismograms using an optimal transport distance: application to full waveform inversion,
949 *Geophys. J. Int.*, **205**(1), 345–377.
- 950 Mosegaard, K. & Tarantola, A., 1995. Monte carlo sampling of solutions to inverse problems, *J.*
951 *Geophys. Res. Solid Earth*, **100**(B7), 12431–12447.
- 952 Nawaz, M. A. & Curtis, A., 2016. Bayesian inversion of seismic attributes for geological facies
953 using a hidden markov model, *Geophys. J. Int.*, pp. 1184–1200.
- 954 Nguyen, L. T. & Modrak, R. T., 2018. Ultrasonic wavefield inversion and migration in com-
955 plex heterogeneous structures: 2D numerical imaging and nondestructive testing experiments,
956 *Ultrasonics*, **82**, 357–370.
- 957 Nolen, J., Pavliotis, G. A., & Stuart, A. M., 2012. Multiscale modelling and inverse problems, in
958 *Numerical analysis of multiscale problems*, pp. 1–34, Springer.
- 959 Pageot, D., Leparoux, D., Le Feuvre, M., Durand, O., Côte, P., & Capdeville, Y., 2017. Improving
960 the seismic small-scale modelling by comparison with numerical methods, *Geophys. J. Int.*,
961 **211**(1), 637–649.
- 962 Plessix, R.-E., 2006. A review of the adjoint-state method for computing the gradient of a
963 functional with geophysical applications, *Geophys. J. Int.*, **167**(2), 495–503.
- 964 Pratt, R. G., 1999. Seismic waveform inversion in the frequency domain; part 1, theory and
965 verification in a physical scale model, *Geophysics*, **64**(3), 888–901.
- 966 Rose, J. H., 1989. Elastic wave inverse scattering in nondestructive evaluation, in *Scattering and*

- 967 *Attenuation of Seismic Waves, Part II*, pp. 715–739, Springer.
- 968 Sambridge, M., Bodin, T., Gallagher, K., & Tkalčić, H., 2013. Transdimensional inference in the
969 geosciences, *Philos. Trans. R. Soc. London A*, **371**(1984), 20110547.
- 970 Sánchez-Palencia, E., 1980. Non-homogeneous media and vibration theory, *Lecture notes in*
971 *physics*, **127**.
- 972 Seidl, R. & Rank, E., 2016. Iterative time reversal based flaw identification, *Computers & Math-*
973 *ematics with Applications*, **72**(4), 879–892.
- 974 Sen, M. K. & Biswas, R., 2017. Transdimensional seismic inversion using the reversible jump
975 hamiltonian monte carlo algorithm, *Geophysics*, **82**(3), R119–R134.
- 976 Tape, C., Liu, Q., Maggi, A., & Tromp, J., 2010. Seismic tomography of the southern California
977 crust based on spectral-element and adjoint methods, *Geophys. J. Int.*, **180**(1), 433–462.
- 978 Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*,
979 **49**(8), 1259–1266.
- 980 Thurin, J., Brossier, R., & Métivier, L., 2019. Ensemble-based uncertainty estimation in full
981 waveform inversion, *Geophys. J. Int.*, **219**(3), 1613–1635.
- 982 Tran, K. T., McVay, M., Faraone, M., & Horhota, D., 2013. Sinkhole detection using 2D full
983 seismic waveform tomography, *Geophysics*, **78**(5), R175–R183.
- 984 Tromp, J., Komatitsch, D., & Liu, Q., 2008. Spectral-element and adjoint methods in seismology,
985 *Communications in Computational Physics*, **3**(1), 1–32.
- 986 Virieux, J., Asnaashari, A., Brossier, R., Métivier, L., Ribodetti, A., & Zhou, W., 2017. An
987 introduction to full waveform inversion, in *Encyclopedia of exploration geophysics*, pp. R1–1,
988 Society of Exploration Geophysicists.
- 989 Woodhouse, J. H. & Dziewonski, A. M., 1984. Mapping the upper mantle: Three-dimensional
990 modeling of Earth structure by inversion of seismic waveforms, *J. Geophys. Res. Solid Earth*,
991 **89**(B7), 5953–5986.
- 992 Zhu, H., Li, S., Fomel, S., Stadler, G., & Ghattas, O., 2016. A Bayesian approach to estimate
993 uncertainty for full-waveform inversion using a priori information from depth migration, *Geo-*
994 *physics*, **81**(5), R307–R323.