entropy

MDPI

Article

Uncertainty Quantification in Inverse Scattering Problems

Carolina Abugattas [1,2], Ana Carpio [2,*] and Elena Cebrián [3]

[1] Facultad de Ingeniería, Universidad Alberto Hurtado, Santiago de Chile 8340575, Chile; cabugattas@uahurtado.cl
[2] Departamento de Análisis Matemático y Matemática Aplicada, Universidad Complutense, 28040 Madrid, Spain
[3] Departamento de Matemáticas y Computación, Universidad de Burgos, 09001 Burgos, Spain; elenac@ubu.es
* Correspondence: ana_carpio@mat.ucm.es; Tel.: +34-91-3944407

Abstract

Inverse scattering problems seek anomalies in a medium given data measured after the interaction with emitted waves. Due to noise, predictions about the nature of these inclusions should be complemented with uncertainty estimates. To this end, we propose a progressive framework for inverse scattering from low- to high-dimensional Bayesian formulations depending on the prior information and the problem complexity. We aim to reduce computational costs by exploiting educated prior information. When we look for a few well-separated inclusions in a known medium with information about their number, we resort to low-dimensional parameterizations in terms of a few random variables representing their shape and material constants. We test this approach detecting anomalies in tissues and deposits in stratified subsoils. In more complex situations where the anomalies may overlap, we propose high-dimensional parameterizations obtained from Karhunen–Loève (KL) or Fourier expansions of the density and velocity fields. We employ these methods to characterize oil and gas reservoirs in a salt dome configuration, where the screening effect of the dome cap prevents the obtention of adequate prior information. We characterize the posterior probability by means of affine invariant ensemble and functional ensemble MCMC samplers depending on dimensionality. This provides information on configurations with the highest a posteriori probability and the uncertainty around them, identifying factors that could reduce the uncertainty. In high-dimensional setups, techniques based on KL developments are more effective and stable. A recurring issue is the choice of the a priori covariance (which strongly affects the results) and the choice of its hyperparameters. Here, we use educated choices. Formulations that include them as additional parameters could be a next step at a higher cost.

Keywords: Bayesian inversion; uncertainty quantification; Markov chain Monte Carlo; Karhunen–Loève expansions; geophysical imaging; medical imaging

Check for updates

Academic Editor: Marcin Kamiński

Received: 7 March 2026
Revised: 4 April 2026
Accepted: 7 April 2026
Published: 17 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Subsurface structures are typically inspected by imaging techniques based on wave emission and recording at specific sites. Depending on the waves employed we have different imaging techniques. We focus here on quantifying uncertainty in imaging approaches that rely on elastic waves, with applications in medical and geophysical imaging.

Geophysical imaging methods include electrical resistivity tomography, ground-penetrating radar, induced polarization seismic tomography, reflection seismology and magnetotellurics. Seismology, for instance, estimates subsurface properties from elastic

energy created by natural and artificial sources [1]. The technique requires a controlled source of elastic waves, such as a seismic vibrator, an explosion, or an air gun in marine data acquisitions. Seismic waves are often modeled as one-dimensional rays, a method referred to as ray theory, which can fit travel time data well through reflection and migration techniques [2,3]. However, the full waveform contains more information than just the travel times, which has led to the development of tomography methods [4]. In geophysical imaging, there are multiple sources of uncertainty, such as the huge size of the regions to explore, multiple reflections and distortions due to the heterogeneity and change in rock type, fluctuations in the terrain and other causes of noise. Therefore, we need to be able to quantify the uncertainty caused by such fluctuations in our predictions.

Medical imaging aims to visualize internal tissue structures by non-invasive techniques [5]. Popular methods include X-ray radiography, nuclear medicine, magnetic resonance imaging, ultrasound, echography, tomography, endoscopy, elastography, thermography, and holography. Elastography is an imaging technique based on the elastic properties of soft tissue [6,7]. Healthy tissue is often softer than cancerous tumors (breast and prostate tumors, for instance, see [7–9]). Similarly, healthy liver tissue is softer than damaged liver tissue. Existing technology [10] can distinguish healthy and unhealthy tissue in some situations. However, the imaging of tissues containing tiny tumors, multiple anomalies, or blurry regions with little contrast is strongly affected by noise and fluctuations. Uncertainty due to different factors needs to be quantified.

Characterizing the structure of reservoirs from full waveform seismic data or the structure of tissues from elastography data are typical examples of inverse problems. Knowing the signal emitted at the sources and the measurements recorded at sets of receivers, we aim to describe the structure and composition of a medium. Inverse problems arise in many other contexts (non-destructive testing of materials and structures, security, etc.) and are severely ill posed [11]. In spite of the vast number of techniques created so far, there is a need for ever-better methods. Specifically, developing uncertainty quantification tools to assess the effect of noise and fluctuations in the predictions is an active field of research nowadays. In this line, Bayesian inversion approaches are a promising tool to handle inverse scattering problems while quantifying uncertainty [12–14]. In the Bayesian framework, noise, unknowns and model parameters are represented as random variables or fields. Our previous knowledge of the solution acts as a regularizing factor through an a priori distribution. The solution to a Bayesian inverse problem is the posterior distribution, that is, the conditional distribution of the unknowns given (i) the prior density of the unknowns, which incorporates prior (expert) knowledge and (ii) the likelihood or conditional probability of the observing the data given the variables according to a certain model of wave propagation. Defining a likelihood requires the choice of a metric. Euclidean distances weighted with a covariance matrix are typical choices; see [15] for a review and for recent developments such as entropy-based metrics. The configurations with the highest probability (maximum a posteriori estimates) are candidate solutions to the original inverse problem, with uncertainty characterized by the posterior distribution. Often, the posterior distributions can be approximated by expressions of the form $e^{-J(\nu)}$, where ν represents the unknown variables or fields. Then, the maximum a posteriori (MAP) estimates are solutions of constrained optimization problems where the equations for wave propagation act as a constraint, which can be solved by Newton and Levenberg–Marquardt type techniques [16,17]. The functional $J(\nu)$ is usually given by the metric employed in the likelihood plus a regularizing term that uses the a priori information. Purely deterministic approaches typically aim to optimize alternative constrained functionals that employ a variety of Tikhonov and total variation regularization terms [17–19]. Deterministic methods provide a solution for each dataset considered, that can even be robust for very small noise

under specific conditions when the material nature is known [17,18]. However, they do not assess how those configurations change with increasing noise levels and what is their range of variability. In the Bayesian context, the MAP estimates provide the most likely anomaly configuration at a moderate computational cost. Resorting to a Laplace approximation, one can quantify uncertainty to a certain extent. This approximation linearizes the posterior distribution about the MAP estimate [20] and replaces the posterior with a Gaussian distribution. Asymmetry and multimodality features are lost [21]. When the MAP point and the mean of the posterior distribution agree, generalized stochastic perturbation techniques [22] may provide a way to obtain information on skewness and other features. However, posterior distributions often encode more information that the MAP estimate and simple approximations. Sampling the posterior distribution by means of Markov chain Monte Carlo (MCMC) methods we can obtain a precise characterization, including the potential presence of several high probability configurations and asymmetry effects in the tails. MCMC sampling is an optimization technique itself: it provides all the minima of the cost $J(\boldsymbol{\nu})$, ranked according to the value of the cost (i.e., their probability), but it also gives information on the uncertainty around these values through sample processing.

A wide variety of strategies to generate Markov chains to sample from target probability distributions are available: Metropolis–Hastings (MH), Hamiltonian Monte Carlo (HMC), preconditioned Crank–Nicholson (pCN), ensemble affine invariant samplers (AIES), ensemble pCN based samplers and functional ensemble based samplers (FES); see [21] for a review and references therein. Compared to MH methods, HMC methods reduce correlations in the samples. The standard HMC algorithm is inherently sequential due to its trajectory-building steps and tends to get trapped around a single mode when sampling multimodal distributions. However, there are recent parallelization attempts [23]. Instead, AIES are designed to be parallelizable and are known to be efficient for poorly scaled and/or multimodal distributions in low to moderate dimensions [24]. As the dimension increases, these methods present vanishing acceptance rates. pCN methods were introduced to overcome this issue, displaying robust acceptance rates as the dimension tends to infinity [14]. Nevertheless, they are slow for poorly scaled or multimodal distributions. Combinations of pCN and AIES such as FES are robust as dimensions tend to infinity, parallelizable, gradient free, and can be designed to handle poorly scaled or multimodal distributions [25]. Since we face poorly scaled and multimodal distributions in increasing dimensions, AIES and FES seem to be well adapted to sample the posterior distributions we will consider. Parallelization is an additional asset since each likelihood evaluation involves solving an initial boundary value problem by hybrid finite difference–finite element techniques, which is computationally quite expensive.

To frame the problems of subsurface characterization in a Bayesian inversion context, we must choose a mathematical representation of the unknown features and determine the a priori information that we can include. This leads to different approaches to the problem, whose computational cost may be prohibitive. In this work, we focus on situations where existing information about the subsurface (provided by other techniques) allows us to adopt low- or moderate-dimensional representations of the unknown subsurface elements. Under these conditions, we can formulate Bayesian versions of the inverse problem with a manageable computational cost. Developing cost reduction strategies, such as reduced-order models or deep learning surrogates [26–28] is an active research field nowadays. However, the applicability these methods to time dependent wave problems with dynamically changing spatiotemporal wave patterns is still a subject of research [29].

The paper is organized as follows. Section 2 formulates the general uncertainty quantification framework for the inverse problem under consideration. We detail the forward wave model in Section 3. Before proposing a Bayesian formulation, we need

to decide how to represent mathematically the details of the subsurface structure that we ignore. In this context, they would be characterized by means of density, velocity and/or elastic moduli fields. In principle, we should determine these fields everywhere in the computational region, which is an infinite-dimensional problem. Finite-dimensional Bayesian formulations that are robust as the dimension grows are suggested in [12], for instance. They represent the unknown fields by their values at all the points of a fine mesh that covers the whole computational region. Here, we propose a progressive framework from low- to high-dimensional Bayesian formulations depending on the prior information, which relies mostly on bases expansions or specific parametrizations on the unknown geometry. We will consider two different situations of increasing complexity. In both, we assume we have some knowledge of the subsurface structure, obtained by other methods. We wish to determine the detailed composition of some regions, where we suspect the presence of inclusions of a different nature. In Section 4 we characterize the inclusions by a small number of parameters that define their shape and material constants. Then, we present uncertainty quantification tests in medical elastography and in geophysics based on AIES (that are known to be effective for poorly scaled problems and multimodal distributions in low dimensions), see Figure 1. Uncertainty tends to be larger in geophysical tests since they typically involve two material constants and stratified configurations, which favors the appearance of several high probability configurations. Section 5 characterizes inclusions in terms of perturbations of the a priori known density and velocity fields. Such perturbations are defined by means of series expansions. In principle this is an infinite-dimensional problem, that we approximate in finite dimensions by means of truncated expansions. We consider two types of parameterizations. The first one is defined by the coefficients of truncated Fourier series expansions, which becomes a finite set of random variables. The second one works with truncated Karhunen–Loève expansions [14,25,30] of the density and velocity random fields. We tests the methods on a salt dome geometry to evaluate the presence of possible oil/gas reservoirs. To sample the resulting posterior probability densities we resort to FES. As previously said, these samplers have a robust acceptance rate as the dimension tends to infinite, are parallelizable and have the potential of handling complex distributions. Finally, Section 6 discuses the results and summarizes our conclusions. A final appendix describes the MCMC ensemble sampling algorithm used, that adapts the original algorithm to work with several fields.

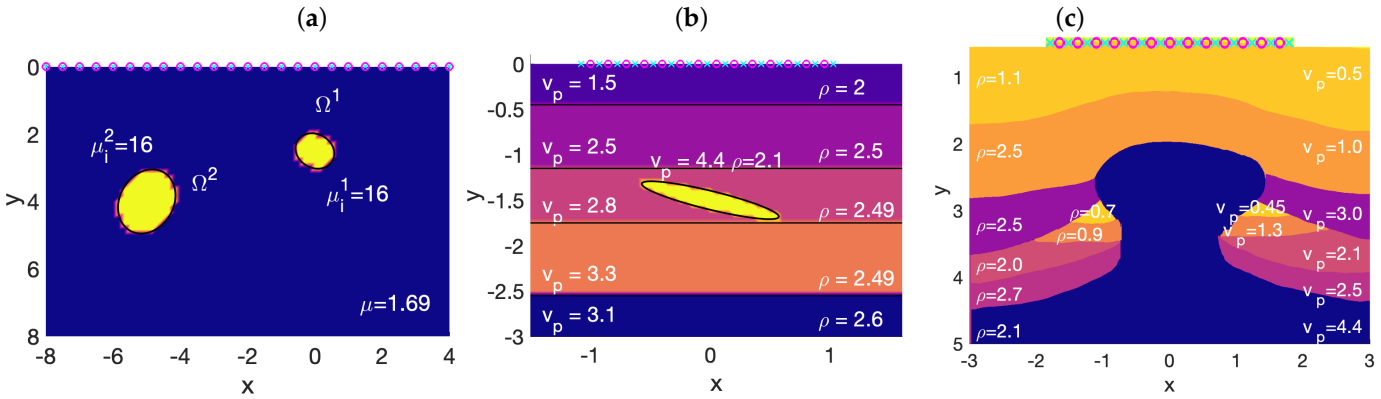

Figure 1. Scheme of typical imaging setups: (**a**) anomalies in tissues, (**b**) inclusions in stratified soil and (**c**) gas and oil reservoirs attached to salt domes. Units for shear moduli μ, densities ρ and velocities v_p are kPa, $\frac{10^3\ \mathrm{kg}}{\mathrm{m}^3} = \frac{\mathrm{g}}{\mathrm{cm}^3}$ and $\frac{\mathrm{km}}{\mathrm{s}}$. Parameter values correspond to healthy and scar liver tissue in (**a**) and to sandstone, limestone, shale and salt in (**b**,**c**), including gas and oil in (**c**). Circles and crosses represent emitters and receivers.

2. Bayesian Inverse Problem

Given a random field ξ, Bayes formula in infinite dimensions [14,31–33] states that the Radon–Nikodym derivative of the posterior probability measure π_{pt} with respect to the prior measure π_{pr} is given in terms of the conditional probability density function

$$\frac{d\pi_{\text{pt}}}{d\pi_{\text{pr}}} = p(\mathbf{d}|\xi)\frac{d\pi_{\text{pr}}}{I}, \quad I = \int \pi(\mathbf{d}|\xi)d\pi_{\text{pr}}, \tag{1}$$

where π_{pr} is the prior distribution, $p(\mathbf{d}|\xi) = L(\xi)$ is the likelihood and $\mathbf{d}$ are the available data. In principle, our inverse problem would consist in characterizing this probability when ξ represents the underground density and velocity fields and $\mathbf{d}$ the data recorded in an imaging setup.

Dimensionality can be reduced in several ways so that the unknown density and velocity fields are approximated by a finite number of parameters. In this framework, such parameters $\boldsymbol{\nu}$ are considered to be a finite set of random variables. According to Bayes's Theorem [33]

$$p_{\text{pt}}(\boldsymbol{\nu}) := p(\boldsymbol{\nu}|\mathbf{d}) = \frac{p(\mathbf{d}|\boldsymbol{\nu})p_{\text{pr}}(\boldsymbol{\nu})}{p(\mathbf{d})}, \tag{2}$$

where $p_{\text{pr}}(\boldsymbol{\nu})$ is the prior probability of the parameters (that encodes available information or known restrictions), $p(\mathbf{d}|\boldsymbol{\nu}) = L(\boldsymbol{\nu})$ is the conditional probability or likelihood (that reflects the fidelity to the measured data of the data that would be observed for $\boldsymbol{\nu}$ as predicted by a model of wave propagation), and $p(\mathbf{d})$ acts as a normalization factor. Different metrics can be selected to compare the measured data and the data predicted by a model in the likelihood definition (see [15] for a review and for new developments such as entropy-based metrics). Here, we mostly resort to Mahalanobis-type metrics, which are essentially Euclidean distances weighted with a covariance matrix. We will consider next two different kinds of finite-dimensional reductions depending on the a priori knowledge on the subsurface. First, we need to discuss the model of wave propagation selected to define the likelihood.

3. Physical Setups and Observation Operator

Typical setups locate a number of sources $\mathbf{x}_k$, $k = 1, \ldots, K$, at the upper surface that launch intense signals into the imaged medium. After interaction with subsurface structures, the waves are recorded at receivers $\mathbf{r}_j$, $j = 1, \ldots, J$, for some time. The recorded signals constitute the data $\mathbf{d}$ and are functions of the wavefield values at the receiver sites. We describe wave dynamics by a simplified model of wave propagation through the subsurface

$$\begin{aligned} &\rho(\mathbf{x})u_{tt} - \operatorname{div}(\mu(\mathbf{x})\nabla u) = f(t)g(\mathbf{x}), \quad x_2 < 0,\, x_1 \in \mathbb{R}, \\ &\tfrac{\partial u}{\partial \mathbf{n}}(x_1, 0, t) = 0, \quad x_1 \in \mathbb{R}, \\ &u(\mathbf{x}, 0) = 0, u_t(\mathbf{x}, 0) = 0, \quad x_2 < 0,\, x_1 \in \mathbb{R}, \end{aligned} \tag{3}$$

where $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, $t > 0$, and $u(\mathbf{x}, t)$ represents the scalar displacements in the (negative) vertical x_2 direction. The term $f(t)g(\mathbf{x})$ describes wave emission at sources. The function f is a Ricker wavelet of frequency f_M related to the spatial resolution expected and the distance between emitters/receivers $f(t) = f_0(1 - 2\pi^2 f_M^2 t^2)e^{-\pi^2 f_M^2 t^2}$. The function g represents the perturbation created by the emitters, for instance, $g(\mathbf{x}) = \frac{1}{(\pi\kappa)^{n/2}} \sum_{k=1}^{K} e^{\frac{-|\mathbf{x}-\mathbf{x}_k|^2}{\kappa}}$, where $\kappa > 0$ is a regularizing parameter and $n = 2$ the spatial dimension. It has zero

normal derivative at the interface. In medical elastography, (3) describes the propagation of shear waves in the tissue and the coefficient μ represents its shear modulus. In geophysical imaging, (3) governs the longitudinal waves that reach deep underground and μ relates to densities and speeds by $\mu = \rho v_{\mathrm{P}}^2$.

We truncate the half space to a finite rectangular domain R and impose non-reflecting boundary conditions at the artificial vertical and horizontal walls to approximate the solution of ((3), see [13,16]). In what follows, we will work with a non-dimensionalized version of the truncated model obtained by the change of variables: $\mathbf{x} = \mathbf{x}'L$, $t = t'T$, $u = u'L$, $R = R'L$, $\Sigma = \Sigma'L$, $\rho = \rho'\rho_0$ and $\mu = \mu'\rho_0 L^2/T^2$. Here, T, L and ρ_0 are characteristic time, length and density scales. Dropping the symbol $'$ for ease of notation, we find

$$\begin{aligned} &\rho u_{tt} - \operatorname{div}(\mu \nabla u) = \tilde{f}(t)\tilde{g}(\mathbf{x}), && \mathbf{x} \in R,\ t > 0, \\ &\tfrac{\partial u}{\partial \mathbf{n}} = 0, && \mathbf{x} \in \Sigma, \\ &\tfrac{\partial u}{\partial \mathbf{n}} = \tfrac{-1}{c} u_t && \mathbf{x} \in \partial R \setminus \overline{\Sigma}, \\ &u(\mathbf{x}, 0) = 0, u_t(\mathbf{x}, 0) = 0, && \mathbf{x} \in R. \end{aligned} \tag{4}$$

Here, $\tilde{f}(t) = \frac{T^3}{\rho_0 L} f(Tt)$ and $\tilde{g}(\mathbf{x}) = g(L\mathbf{x})$. Σ represents the upper boundary. The boundary condition on the artificial boundaries $\partial R \setminus \overline{\Sigma}$, with $c = \frac{\sqrt{\mu}}{\sqrt{\rho}}$, is non-reflecting as long as the velocity and density fields remain piecewise constant nearby and discontinuity curves are orthogonal to the boundary (see [16]). We will make this assumption on the imaged region. Well-posedness and stability results for this model have established in [13,16], where low-cost finite element (FEM) schemes are also discussed.

Given a model of wave propagation such as (4), we define a basic observation operator by recording wave profiles at the receivers $\mathbf{r}_j$, $j = 1, \ldots, J$, for a sequence of times t_m, $m = 1, \ldots, M$. For each realization of the coefficient fields, the observation operator is

$$\mathcal{F} : (\rho, \mu) \to (u_{\rho,\mu}(\mathbf{r}_j, t_m))_{j=1,\ldots,J, m=1,\ldots,M} \tag{5}$$

where $u_{\rho,\mu}$ is the solution of (4). Depending on the nature of the receivers and the expected regularity of the wave field, other choices may be advisable, such as averages in small areas at each receiver location. In practice, we construct numerical solutions of (4) using the finite element method. Therefore, the observation operator (5) will use the numerical solution. For a study on the effect of numerical choices on the observation operator and the numerical solution of inverse Bayesian problems, we refer the reader to [16].

We assume that the measured data are related to the observation operator by

$$\mathbf{d} = \mathcal{F}(\rho_{\text{true}}, \mu_{\text{true}}) + \boldsymbol{\eta},$$

where ρ_{true}, μ_{true} represent the density and the velocity fields for the explored region and $\boldsymbol{\eta}$ is a noise term. To simplify, we assume Gaussian noise with zero mean and correlation $\boldsymbol{\Gamma}_{\text{n}}$, that is, $\boldsymbol{\eta} \in \mathcal{N}(0, \boldsymbol{\Gamma}_{\text{n}})$. In practice, this may not be the case and other probability distributions can be used [34,35].

4. Finite Random Variable Approach

In this section, we assume that we have information on the overall subsurface structure (obtained by other techniques) except for details on the presence of localized inclusions of different materials.

Each inclusion is parametrized in terms of a few parameters $\boldsymbol{\nu}^{(k)}$. Concatenating them, the vector $\boldsymbol{\nu} = (\boldsymbol{\nu}^{(1)}, \ldots, \boldsymbol{\nu}^{(L)})$ represents L possible inclusions. We define $p_{\mathrm{pr}}(\boldsymbol{\nu})$ in terms of a truncated multivariate Gaussian

$$p_{\mathrm{pr}}(\boldsymbol{\nu}) = \begin{cases} \frac{p_0}{(2\pi)^{N/2}} \frac{1}{\sqrt{\det(\boldsymbol{\Gamma}_{\mathrm{pr}})}} \exp(-\frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)^\top \boldsymbol{\Gamma}_{\mathrm{pr}}^{-1}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)), & \boldsymbol{\nu} \in \mathcal{M}, \\ 0, & \boldsymbol{\nu} \notin \mathcal{M}, \end{cases} \tag{6}$$

where N is the dimension of $\boldsymbol{\nu}_0$ and $\mathcal{M}$ incorporates a set of constraints on the coefficients. The symbol $\top$ denotes transposition. The number p_0 is a normalization constant, whose value is not needed in practice. The vector $\boldsymbol{\nu}_0$ represents a priori information on the possible values of the sought parameters, while $\boldsymbol{\Gamma}_{\mathrm{pr}}$ represents the covariance matrix. We define the likelihood as

$$p(\mathbf{d} \mid \boldsymbol{\nu}) = \frac{1}{(2\pi)^{D/2}\sqrt{\det(\boldsymbol{\Gamma}_{\mathrm{n}})}} \exp\left(-\frac{1}{2}(\mathcal{F}(\boldsymbol{\nu}) - \mathbf{d})^\top \boldsymbol{\Gamma}_{\mathrm{n}}^{-1}(\mathcal{F}(\boldsymbol{\nu}) - \mathbf{d})\right), \tag{7}$$

where $\boldsymbol{\Gamma}_{\mathrm{n}}$ is the covariance matrix for the noise affecting the data $\mathbf{d}$. We chose here a $D \times D$ diagonal matrix, with constant diagonal given by σ_{n}^2. $\mathcal{F}(\boldsymbol{\nu})$ represents the observation operator (5) when $\rho = \rho_\nu$ and $\mu = \mu_\nu$, that is, ρ and μ incorporate the presence of the inclusions defined by $\boldsymbol{\nu}$. Neglecting scaling factors, the posterior distribution (2) can be approximated by

$$p_{\mathrm{pt}}(\boldsymbol{\nu}) \approx \exp\left(-\frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)^\top \boldsymbol{\Gamma}_{\mathrm{pr}}^{-1}(\boldsymbol{\nu} - \boldsymbol{\nu}_0) - \frac{1}{2}(\mathcal{F}(\boldsymbol{\nu}) - \mathbf{d})^\top \boldsymbol{\Gamma}_{\mathrm{n}}^{-1}(\mathcal{F}(\boldsymbol{\nu}) - \mathbf{d})\right) \tag{8}$$

for $\boldsymbol{\nu} \in \mathcal{M}$ and zero otherwise. The highest probability configurations minimize the regularized cost

$$J_R(\boldsymbol{\nu}) = \frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)^\top \boldsymbol{\Gamma}_{\mathrm{pr}}^{-1}(\boldsymbol{\nu} - \boldsymbol{\nu}_0) + \frac{1}{2}(\mathcal{F}(\boldsymbol{\nu}) - \mathbf{d})^\top \boldsymbol{\Gamma}_{\mathrm{n}}^{-1}(\mathcal{F}(\boldsymbol{\nu}) - \mathbf{d}) \tag{9}$$

for $\boldsymbol{\nu} \in \mathcal{M}$.

4.1. Homogeneous Background

In homogeneous media, we can obtain educated a priori information on the number of non-overlapping inclusions L, their size and their location, by means of topological energies [13,36]. Iterative methods combining successive calculations of topological fields with deterministic optimization [18] can approximate with precision the geometry of anomalies in cases where the nature of the inclusions is known and noise levels are very small. However, it is quite difficult to predict the nature of the anomalies [36] due to the possible presence of additional local minima. Topological energy methods are robust to noise [37] in the sense that the number of predicted anomalies does not really change as we increase the noise levels. Capturing the right number of anomalies depends on factors such as the overall geometrical arrangement of the anomalies, the employed frequency, the region covered by the receivers and the distance between them, and the use of sweeping strategies (exploiting the finite speed of the waves to calculate topological energies with data from different fractions of receivers to cover different sectors) [13,38]. However, topological methods alone provide no insight on the nature of the inclusions. We can devise a procedure that uses part of the data to produce educated a priori information on the number and geometry of inclusions by topological methods, reserving the remaining amount of data for the Bayesian inversion problem in large noise contexts.

Let us assume that we have characterized the number of inclusions L by topological energy methods and extracted prior information on their geometry encoded in a set of

parameters $\boldsymbol{\nu}_0^{(\ell)}$, $\ell = 1, \ldots, L$. Since we have no information on their nature, we set the a priori values for the material parameters equal to those of the surrounding medium. In the tests presented here, the number of objects predicted by topological energy methods remains the same if we vary the noise intensity: 5%, 15%, 20% .... Up to 70% noise the centers and sizes predicted by topological energy methods to construct the priors remain pretty similar too.

For simplicity, we choose starshaped parameterizations, whose radius is approximated by trigonometric polynomials. The contour of each inclusion Ω^ℓ is given by

$$\mathbf{q}^\ell(\theta) = (\nu^\ell(1), \nu^\ell(2)) + r^\ell(\theta)(\cos(2\pi\theta), \sin(2\pi\theta)), \quad \theta \in [0,1], \tag{10}$$

$$r^\ell(\theta) = \nu^\ell(3) + 2\sum_{q=1}^{Q} \nu^\ell(2q+3)\cos(2\pi q\theta) + 2\sum_{q=1}^{Q} \nu^\ell(2q+2)\sin(2\pi q\theta), \tag{11}$$

and their material properties by $\nu^\ell(2Q+4)$, $\nu^\ell(2Q+5)$, for $\ell = 1, \ldots, L$. The number of parameters $N = L(2Q+5)$. While $(\nu^\ell(1), \nu^\ell(2))$ represents the center, $\nu^\ell(3)$ is the first term of the radius expansion. Other parameterizations are possible (see [39], for instance). In this setup, the waves emitted by the sources are governed by (4) with

$$\rho(\mathbf{x}) = \rho_{\boldsymbol{\nu}}(\mathbf{x}) = \begin{cases} \rho, & \mathbf{x} \in R \setminus \overline{\Omega}_{\boldsymbol{\nu}}, \\ \rho^\ell, & \mathbf{x} \in \Omega^\ell,\ \ell = 1, \ldots, L, \end{cases} \tag{12}$$

$$\mu(\mathbf{x}) = \mu_{\boldsymbol{\nu}}(\mathbf{x}) = \begin{cases} \mu, & \mathbf{x} \in R \setminus \overline{\Omega}_{\boldsymbol{\nu}}, \\ \mu^\ell, & \mathbf{x} \in \Omega^\ell, \ell = 1, \ldots, L, \end{cases} \tag{13}$$

and $\Omega_{\boldsymbol{\nu}} = \cup_{\ell=1}^{L} \Omega^\ell$. We choose a diagonal covariance matrix $\boldsymbol{\Gamma}_{\text{pr}}$ formed by L blocks. The components of each block corresponding to the radius must decay so that we favor positive radii and avoid cusps. We typically enforce power-like decay. This is reminiscent of deterministic inversion studies with star-shaped objects in other contexts, which use weighted metrics for the coefficients of the trigonometric polynomial defining the radius [17].

With these choices, we approximate the posterior distribution by (8). Since the number of parameters forming $\boldsymbol{\nu}$ is usually small, and they are unbalanced, we sample (8) by means of an affine invariant ensemble sampler (AIES) (see [24]). This kind of Markov chain Monte Carlo samplers requires mixing $W > 2N$ chains to properly sample the distribution; see Appendix A for details.

Figure 2 illustrates the behavior of the posterior distribution (8) for a homogeneous tissue containing two anomalies, after 1000 steps of the MCMC algorithm implemented with parameters $a = 2$ and $W = 480$. In tissues, we can assume that $\rho = \rho_0$ everywhere. Thus, the parametrization of each object has only $2Q+4$ components and $N = L(2Q+4)$. The simulation employs data from liver tissue. Anomalies in the liver can represent scar tissue but also malignant tumors. Figure 2a represents the probability that a point belongs to an anomaly. The location, size and shape of the anomalies are well captured. The MAP and mean curves represent the highest probability contour and the contour corresponding to the average of the samples generated, excluding an initial subset discarded as a burn-in period. Figure 2b depicts the values for the shear moduli of the anomalies corresponding to each sample. Most of them lie around the true values (corresponding to scar tissue). However, the distribution has large tails that reach the healthy and malignant regions. These results correspond to data with 10% noise, and to prior information that is quite precise for locations and shapes, but it is blind to the nature of the anomalies. Whereas in healthy tissue $\mu \sim 1.69$ kPa, in scar tissue $\mu_{\text{s}} \sim 16$ kPa, and in malignant tissues is much larger. We consider true inclusions formed by scar tissue, with $\mu_{\text{i}} \sim \mu_{\text{s}}$, but keep μ as the a priori value for them in the posterior probability.

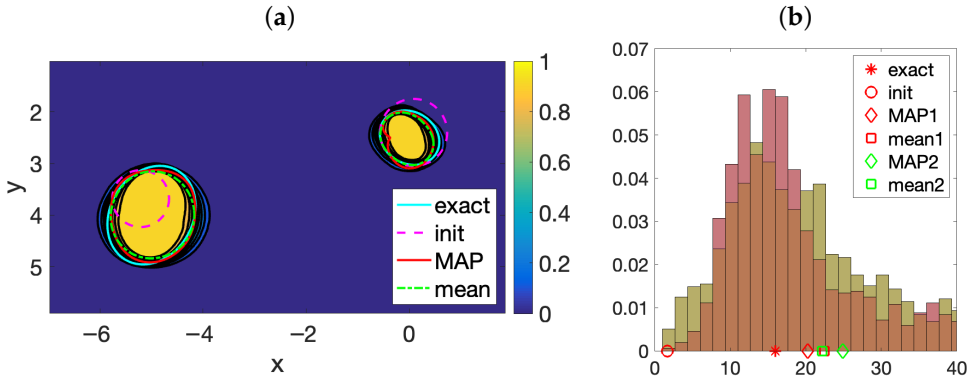

Figure 2. (**a**) Probability of belonging to the anomalies and (**b**) distribution of the shear modulus of the two anomalies (superimposed), constructed by AIES sampling of the posterior probability with liver tissue data. Anomalies are numbered from left (pink histogram, 1), to right (green histogram, 2).

For this test, data have been generated synthetically, solving the forward wave problem (4) with the parameters listed in Table 1 and $\kappa = 2$ in the geometry represented in Figure 1a. We locate transducers at fixed grids of step 0.5 on the upper border ranging from -12 to 12. Notice that the same devices act first as emitters and then as receivers; therefore, we can only record data once the emission is finished. Evaluating the numerical solution on a fixed time grid of step 0.025 in the time interval $[2, 11]$, we obtain the raw data $d^m_{j,\text{true}}$, $j = 1, \ldots, J$, $m = 1, \ldots, M$. To generate the synthetic data, we add noise

$$d^m_j = d^m_{j,\text{true}} + \varepsilon_j, \quad j = 1, \ldots, J, m = 1, \ldots, M, \tag{14}$$

where ε is distributed according to $\mathcal{N}(0, \mathbf{\Gamma}_\text{n})$, with $\sigma_\text{n} = \alpha \max|d^m_{j,\text{true}}|/100$. We set $\alpha = 10$, that is, 10 percent noise. The meshes we use to generate the synthetic data and evaluate the observation operator are different, with the first being finer. We split the data in two halves, corresponding to two interspaced grid with time step 0.05. We use one half to implement topological energy methods [13] to produce educated a priori information on the number of anomalies, their location, and size (dashed magenta circles in Figure 2a). However, we have no a priori information on the nature on the anomalies. We set $\nu^\ell_0(2Q+4) = \mu$, $\ell = 1, 2$, equal to the shear modulus of the healthy tissue. As for the diagonal prior covariance, each block starts with $(\sigma^\ell_1)^2 = (\sigma^\ell_2)^2 = 0.1$ and ends with $(\sigma^\ell_{2Q+4})^2 = 20^2$. Then, $(\sigma^\ell_3)^2 = 0.1$ and $(\sigma^\ell_{2q+2})^2 = (\sigma^\ell_{2q+3})^2 = 0.1/(1+q^2)^s$, $1 \le q \le Q$, s large, as in [40], so that the prior favors regular shapes with $r(t) > 0$. We fix $s = 3$ and $Q = 5$. Notice that $\sigma_{2Q+4} = 20$ is a rather large value. The large deviation allowed in the prior would allow us to infer malignity if necessary.

Table 1. Dimensional parameters used in the tissue simulations.

L	T	ρ_i	ρ	μ_i	μ	c_i	c	f_M	f_0
0.01 m	0.01 s	ρ	$10^3 \frac{\text{kg}}{\text{m}^3}$	16 kPa	1.69 kPa	4 $\frac{\text{m}}{\text{s}}$	1.3 $\frac{\text{m}}{\text{s}}$	50 Hz	$\frac{\rho L}{T^3}$

If we include in the prior distributions more anomalies that needed, some of their shear parameters after sampling may essentially agree with the background medium, suggesting healthy tissue. If we include less anomalies than needed, we may encounter distributions with additional modes, associated to samples that capture the missing anomalies. Multimodality is often a consequence of insufficient prior information. As said before, being able to capture the right number of anomalies L by topological methods depends on the design of the imaging set-up (anomaly/detector arrangement, frequency choice) and the use of sweeping strategies that fraction the data to compute topological energies

covering different regions. Inclusions can be missed if they are screened by other inclusions placed between them and the detectors or their size is below the employed imaging set-up resolution. Methods like the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) [41] are often employed to determine the most likely number of parameters to include in a Bayesian formulation. However, they do not provide a priori values for the centers of the anomalies and their geometrical parameters, so it is not clear how to implement them in this context. Ref. [40] explores a hierarchical Bayesian formulation with model selection through the model evidence to determine the number of anomalies L one should include in the formulation. However, it can only be used to choose between smaller numbers than detected by topological energies.

Topological energy methods would detect overlapping inclusions as one. They are typically employed when small sets of separated anomalies are expected. Since we are assuming here that each inclusion is characterized by constant material parameters we cannot study such anomalies within this framework. Section 5 proposes a different approach applicable to that context.

4.2. Stratified Background

In geophysics, we often work in stratified media. Topological energy tools are hardly useful in this context to obtain educated priors. Instead, we rely on prior information produced by other imaging techniques such as migration analyses [42]. To have a clear idea of the dimensions and orientation of the inclusions, we consider elliptic parameterizations $\boldsymbol{\nu}^\ell$, where $(\nu^\ell(1), \nu^\ell(2))$ represents the center; $\nu^\ell(3)$ and $\nu^\ell(4)$ the x and y semi-axes (respectively); and $\nu^\ell(5)$ the angle of the x semiaxis with the x axis. The last two parameters, $\nu^\ell(6)$ and $\nu^\ell(7)$, represent the density and the velocity, respectively. The total number of parameters is $N = 7L$.

Figure 3 illustrates the behavior of the posterior distribution (8) for stratified media containing one inclusion, after 500 steps of the AIES algorithm, implemented with parameters $a = 2$ and $W = 480$. Figure 3a represents the probability that a point belongs to an anomaly. While the MAP estimate of its boundary is close to the true contour, the mean estimate differs noticeably. There is a broad region about the true inclusion formed by points with a relevant probability of belonging to the inclusion. The posterior probability happens to be multimodal, with families of samples accumulating around different shapes. In addition to the MAP estimate, that is, the curve of highest probability, there are secondary contours with high probabilities. The histograms in panels (b) and (c) visualize the distributions of densities and velocities of the samples. The densities concentrate around the true values and the MAP estimate for the density is very close to the true density. However, the distribution of the velocities is clearly multimodal. The MAP estimate is close to the true velocity, but the mean is just an average of this value and secondary high probability velocities (corresponding to the secondary high probability curves). Notice that this test employs a single frequency. When we combine several frequencies, the secondary modes dilute and the mode representing the true inclusion becomes dominant: uncertainty should diminish.

For this test, data have been generated synthetically, solving the forward wave problem (4) with the parameters listed in Table 2 and $\kappa = 0.04$, in the geometry represented in Figure 1b. We locate the emitters and the receivers at interspaced grids of step 0.04: $K = 51$ sources distributed in $[-1, 1]$ and $J = 52$ detectors ranging from -1.02 to 1.02. Evaluating the numerical solution on a fixed time grid of step 0.1 in the interval $[0, 2.5]$, we obtain the raw data $d^m_{j,\text{true}}$, $j = 1, \ldots, J$, $m = 1, \ldots, M$. To generate the synthetic data, we add noise according to (14) where ε is distributed following $\mathcal{N}(0, \boldsymbol{\Gamma}_\text{n})$, with $\sigma_\text{n} = \alpha/100 \, (\sum_{j=1}^{J} \sum_{m=1}^{M} \frac{|d^m_{j,\text{true}}|^2}{JM})^{\frac{1}{2}}$. We set $\alpha = 10$, that is, 10 percent noise. The meshes

we use to generate synthetic data and to evaluate the observation operator are different, the first one being finer. As before, we choose a diagonal covariance matrix $\mathbf{\Gamma}_{\text{pr}}$ formed by L blocks, L being the number of expected inclusions. The last two components in each block, that is, the standard deviations σ_ρ and σ_{v_p}, are set equal to $\sigma_\rho = (\rho_{\max} - \rho_{\min})/2$ and $\sigma_{v_\text{p}} = (v_{\text{p,max}} - v_{\text{p,min}})/2$, see [16]. For the tests presented here we have set $\mathbf{\Gamma}_{\text{pr}} = \text{diag}\,(1, 1, 0.5, 0.5, 0.1, \sigma_\rho^2, \sigma_{v_\text{p}}^2)$. We assume we have a priori information on the location and size of the anomaly, provided by migration methods, but not on its nature. We select as prior means for ρ and v_p the average of the minimum and maximum values for the layers: $\nu_0^\ell(6) = (\rho_{\max} + \rho_{\min})/2$ and $\nu_0^\ell(7) = (v_{\text{p,max}} + v_{\text{p,min}})/2$. This is a blind choice that does not introduce any specific information on the possible nature of the inclusions.

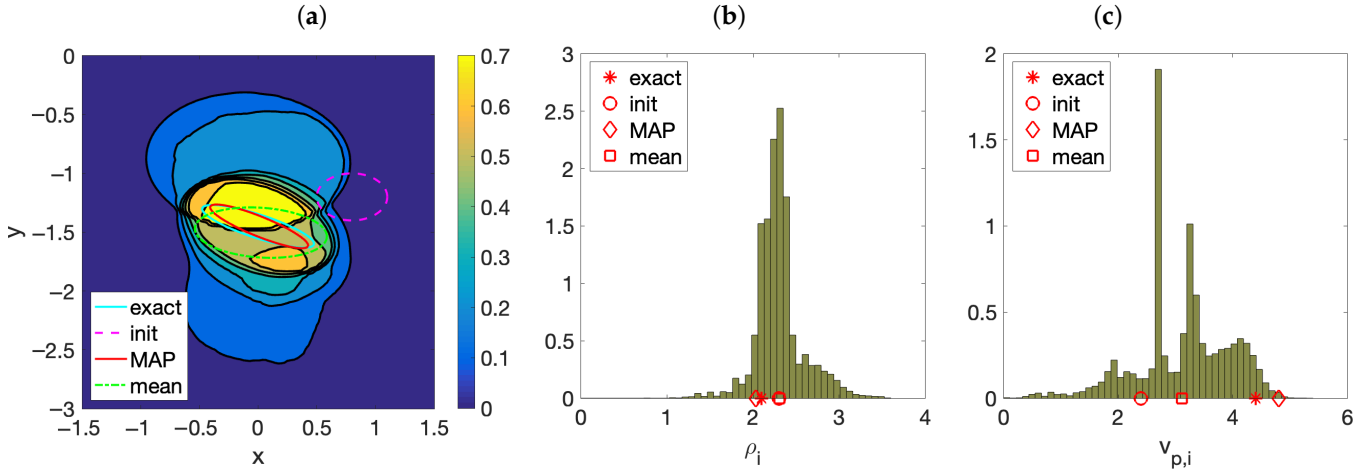

Figure 3. (**a**) Probability of belonging to the anomalies and (**b**,**c**) distribution of the density and the velocity in the inclusion, constructed by AIES sampling of the posterior probability with stratified soil data.

Table 2. Dimensional parameters used in simulations in stratified media.

L	T	ρ_i	ρ_0	$v_{\text{p,i}}$	f_M	f_0
1 km	1 s	$2.1\rho_0$	1000 $\frac{\text{kg}}{\text{m}^3}$	4.4 $\frac{\text{km}}{\text{s}}$	2 Hz	$100\frac{\rho_0 L}{T^3}$

5. Random Field Approach

The previous section focuses on the detection of inclusions in a known medium, given information about their number but not their nature. We were able to estimate their material constants and improve the description of their geometry, with quantified uncertainty. When the number L of inclusions is unknown, it can be considered as another random variable of a discrete nature. The treatment of this type of formulations is complex and has been addressed, for example, in [40]. In this section, we choose to consider the material coefficients as random fields, whose peaks would locate inclusions of diverse nature, without neither imposing their presence, nor introducing information about their shape or location. We assume here that we have reliable information on the background stratified structure, excluding some localized regions which can contain deposits of different materials.

Figure 1c represents a geological structure of particular interest: the salt dome [43]. These domes form when underground salt (less dense than other rocks) moves slowly upwards, deforming and breaking up rocks in the way. These structures often mark the presence of oil-gas traps, when the combination of reservoir and source rocks around them is favorable (a source layer covered by a permeable reservoir rock layer lies under a low permeability cap layer). Salt layers attenuate and distort seismic signals, resulting in poor vertical resolution of standard techniques, as well as poor velocity maps [44]. The deposits

are usually found in carbonate rocks, heterogeneous and with unknown fissure patterns, distribution and connection of pores. The contrast between rock types introduces multiple reflections. Recent work discusses the potential of FWI techniques [45–47] to characterize such reservoirs and underlines the relevance of noise effects in deposit characterization [48] (see also [49] for machine learning methods). We develop here alternative Bayesian approaches, assuming we have information (obtained by other techniques) on the overall subsurface stratified structure, except for the region located under the dome cap, which hides the oil and gas deposits.

We express the velocity and density fields as the sum of known background profiles ρ_0 and $v_{\mathrm{p},0}$ plus correctors ρ_{c} and $v_{\mathrm{p,c}}$:

$$\begin{aligned} \rho(x,y) &= \rho_0(x,y) + \rho_{\mathrm{c}}(x,y), \\ v_{\mathrm{p}}(x,y) &= v_{\mathrm{p},0}(x,y) + v_{\mathrm{p,c}}(x,y), \end{aligned} \tag{15}$$

which we expect to be relevant in the dome area. We propose next two possible formulations.

5.1. Fourier Series

Any square integrable two-dimensional function h admits a Fourier expansion converging to h in L^2, and converging also pointwise where continuous [50,51]. Assuming $\rho_{\mathrm{c}} \in L^2(R)$, we approximate $\rho_{\mathrm{c}}(x,y)$ in $R = [0, L_x] \times [0, L_y]$ by a truncated Fourier expansion

$$\begin{aligned} \rho_{\mathrm{c}}(x,y) \approx \frac{A_{00}}{4} &+ \sum_{q=1}^{N_{\mathrm{x}}} \left[\frac{A_{q0}}{2} \cos\left(\frac{2q\pi x}{L_x}\right) + \frac{B_{q0}}{2} \sin\left(\frac{2q\pi x}{L_x}\right) \right] \\ &+ \sum_{r=1}^{N_{\mathrm{y}}} \left[\frac{A_{0r}}{2} \cos\left(\frac{2r\pi y}{L_y}\right) + \frac{B_{0r}}{2} \sin\left(\frac{2r\pi y}{L_y}\right) \right] + \\ \sum_{q=1}^{N_{\mathrm{x}}} \sum_{r=1}^{N_{\mathrm{y}}} &\left[A_{qr} \cos\left(\frac{2q\pi x}{L_x}\right) \cos\left(\frac{2r\pi y}{L_y}\right) + B_{qr} \sin\left(\frac{2q\pi x}{L_x}\right) \sin\left(\frac{2r\pi y}{L_y}\right) \right] + \\ \sum_{q=1}^{N_{\mathrm{x}}} \sum_{r=1}^{N_{\mathrm{y}}} &\left[C_{qr} \cos\left(\frac{2q\pi x}{L_x}\right) \sin\left(\frac{2r\pi y}{L_y}\right) + D_{qr} \sin\left(\frac{2q\pi x}{L_x}\right) \cos\left(\frac{2r\pi y}{L_y}\right) \right] = \rho_{\boldsymbol{\nu}_\rho}(x,y). \end{aligned} \tag{16}$$

We relabel the coefficients as

$$\begin{aligned} \boldsymbol{\nu}_\rho = (&A_{00}, A_{1,0}, \ldots, A_{N_{\mathrm{x}},0}, B_{1,0}, \ldots, B_{N_{\mathrm{x}},0}, A_{0,1}, \ldots, A_{0,N_{\mathrm{y}}}, B_{0,1}, \ldots, B_{0,N_{\mathrm{y}}}, \\ &A_{1,1}, \ldots, A_{1,N_{\mathrm{y}}}, \ldots, A_{N_{\mathrm{x}},1}, \ldots, A_{N_{\mathrm{x}},N_{\mathrm{y}}}, B_{1,1}, \ldots, B_{1,N_{\mathrm{y}}}, \ldots, B_{N_{\mathrm{x}},1}, \ldots, B_{N_{\mathrm{x}},N_{\mathrm{y}}}, \\ &C_{1,1}, \ldots, C_{1,N_{\mathrm{y}}}, \ldots, C_{N_{\mathrm{x}},1}, \ldots, C_{N_{\mathrm{x}},N_{\mathrm{y}}}, D_{1,1}, \ldots, D_{N_{\mathrm{x}},N_{\mathrm{y}}}, \ldots, D_{N_{\mathrm{x}},1}, \ldots, D_{N_{\mathrm{x}},N_{\mathrm{y}}}), \end{aligned} \tag{17}$$

and perform a similar procedure for $v_{\mathrm{p,c}}$. In this way, we devise a finite-dimensional approximation to the correctors in (15), which becomes

$$\begin{aligned} \rho_{\boldsymbol{\nu}} &= \rho_0 + \rho_{\boldsymbol{\nu}_\rho}, \\ v_{\mathrm{p},\boldsymbol{\nu}} &= v_{\mathrm{p},0} + v_{\mathrm{p},\boldsymbol{\nu}_v}, \end{aligned} \tag{18}$$

in terms of the parameters $\boldsymbol{\nu} = (\boldsymbol{\nu}_\rho, \boldsymbol{\nu}_v)$, where $\boldsymbol{\nu}_\rho$, $\boldsymbol{\nu}_v$ represent the coefficients of the expansions (16) for $\rho_{\boldsymbol{\nu}_\rho}$ and $v_{\mathrm{p},\boldsymbol{\nu}_v}$, respectively. We set the total number of coefficients $N = 2(1 + 2N_{\mathrm{x}} + 2N_{\mathrm{y}} + 4N_{\mathrm{x}}N_{\mathrm{y}})$ and use $\mathcal{F}(\boldsymbol{\nu})$ to denote the observation operator (5) corresponding to the fields (18) with $\mu = \rho v_{\mathrm{p}}^2$.

These coefficients constitute a set of random variables for which we define the posterior probability (2) according to Bayes's Theorem. Assuming we have no information on the presence of traps around the dome, we select a truncated multivariate Gaussian prior distribution p_{pr} with zero mean $\mathcal{N}(0, \boldsymbol{\Gamma}_{\mathrm{pr}})$ and covariance $\boldsymbol{\Gamma}_{\mathrm{pr}}$ when $\rho_{\boldsymbol{\nu}} > 0$ and $v_{\mathrm{p},\boldsymbol{\nu}} > 0$,

and set $p_{\text{pr}}(\boldsymbol{\nu}) = 0$ otherwise. Working with Fourier expansions, the only a priori information we have is that the series expansions should converge. Therefore, the coefficients must tend to zero in a way that their ℓ^2 norm converges. Observing the coefficients of the Fourier series expansions of the type of fields involved in the study, we have not been able to obtain additional information on possible correlations. Therefore, we just chose a diagonal covariance matrix in which the entries corresponding to the modes (q, r) are given by $\frac{\sigma_0^2}{(1+r^2+q^2)^s}$, $s > 1$. We estimate the value of N_x and N_y by approximating piecewise constant functions in the computational region on the computational mesh. With these choices, the likelihood is given by (7) and the posterior distribution to be analyzed is given by (8), with $\boldsymbol{\nu}_\rho$ and $\boldsymbol{\nu}_v$ defined by (16).

We have employed the affine invariant ensemble sampler AIES [24] to characterize (8) for the salt dome problem. These methods do not construct a single chain but W chains that are mixed at each step. The method requires $W > 2N$ to work properly [24]. However, it is known that the chain mixing process becomes inefficient as dimensionality grows. It has been established that AIES methods are expected to be efficient to sample distributions involving small amounts of parameters, typically lower than 100 [24]. In our tests, we set $N_{\text{x}} = N_{\text{y}} = 10$. The resulting number of total parameters N is rather large. The number of chains $W > 2N$ is even larger. Due to this fact, the sampling process becomes inefficient and very costly (weeks using a parallel Matlab server). Worse, there are strong oscillations in the sampled fields due to the diagonal covariance choice that ignores spatial correlations. The resulting maximum a posteriori estimates and mean estimates of the density and velocity fields are not essentially flat with bumps in the reservoir areas as expected. These fields show some oscillations about the known background density and velocity fields, as seen in Figure 4. Still, the sample mean suggests the presence of regions of formed by different materials under the dome (see Figure 4). We have investigated this phenomenon further in one-dimensional setups. The oscillatory behavior observed in the MAP and mean fields increases. This seems to be related to the choice of a diagonal covariance and disappears under other choices. In the next section, we see how a different formulation which includes spatial correlations in the covariances improves the results.

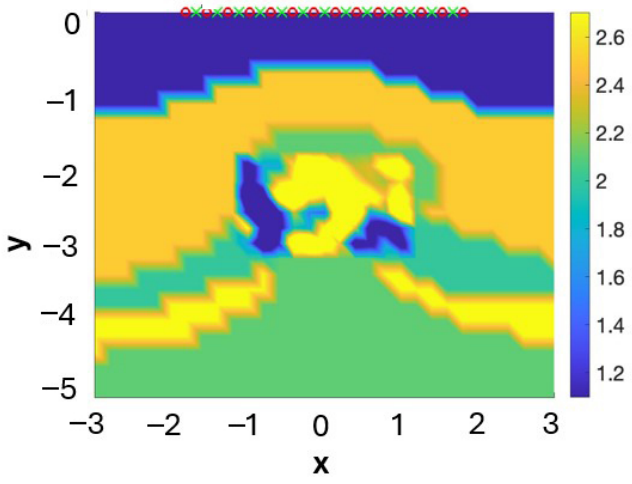

Figure 4. Density field ρ for the geometry in Figure 1c after 1500 steps of the AIES algorithm using a prior information that excludes the presence of oil and gas reservoirs. The image suggests the presence of reservoirs at the sides, under the dome cover.

5.2. Karhunen–Loève Expansions

In this section, we consider the corrections to the pressure and velocity fields as random fields. As previously detailed, in an infinite-dimensional setup [14], the goal of a Bayesian inverse problem is to estimate the posterior distribution

$$\pi_{\mathrm{pt}}(d\xi) \propto L(\xi)\pi_{\mathrm{pr}}(d\xi)$$

where ξ is a square-integrable function on a domain $\Omega \subset \mathbb{R}^n$. Given a likelihood $L(\xi) \propto e^{\phi(\xi)}$ and a prior $\pi_{\mathrm{pr}} \sim \mathcal{N}(0, C)$, Karhunen–Loève expansions decompose random functions ξ drawn from Gaussian distributions $\mathcal{N}(0, C)$ as a linear combination of the eigenfunctions of C [14]

$$\xi = \sum_{\ell=1}^{\infty} \xi_\ell \phi_\ell, \quad \xi_\ell \sim \mathcal{N}(0, \lambda_\ell), \tag{19}$$

associated to its eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_\ell \geq \ldots$. When ξ is a Gaussian random field with zero mean, the coefficients ξ_i are independent Gaussians with mean zero and variances determined by the eigenvalues. The required eigenvalues and eigenvectors are easy to calculate for Matérn type covariance operators $(\tau^2 I - \Delta)^{-\delta}$, where $\tau = 1/l$ represents an inverse length scale and $\delta = \gamma + n/2$ controls the sample regularity [31,52]. If μ_i, ϕ_i are the eigenvalues and eigenfunctions of $\tau^2 - \Delta$ with homogeneous boundary conditions, the eigenvalues and eigenfunctions we seek are $\lambda_i = \mu_i^{-s}$, with $-s = -\delta/2 = -(\gamma + n/2)/2$ and ϕ_i (see [53,54]). The parameter γ controls the degree of irregularity. Gaussian and exponential kernels are particular cases of Matérn kernels. The exponential kernel corresponds to $\gamma = 1/2$ and the Gaussian kernel to $\gamma \to \infty$.

In our framework, we can use zero Dirichlet or Neumann boundary conditions since we expect the correctors to vanish away from the dome. For a rectangle $[0, L_{\mathrm{x}}] \times [0, L_{\mathrm{y}}]$, the eigenvalues of $(\tau^2 I - \Delta)$ with homogeneous Dirichlet boundary conditions are $\lambda_{q,r} = \tau^2 + \pi^2 \left(\frac{q^2}{L_{\mathrm{x}}} + \frac{r^2}{L_{\mathrm{y}}} \right)$ and the normalized eigenfunctions $\phi_{q,r}(x, y) = \frac{4}{L_{\mathrm{x}} L_{\mathrm{y}}} \sin(\frac{\pi q x}{L_{\mathrm{x}}}) \sin(\frac{\pi r y}{L_{\mathrm{y}}})$, $q, r \geq 1$. Thanks to this fact, we can easily construct finite-dimensional approximations of the covariance operator as $C_{\mathbf{\Lambda}} = \mathbf{P}^T \mathbf{\Lambda} \mathbf{P}$ where $\mathbf{\Lambda}$ is a diagonal matrix formed by the selected largest eigenvalues and $\mathbf{P}$ a matrix whose columns are the corresponding discretized basis functions. Since we have some a priori information on the stratified medium, we can select some of the hyperparameters governing the Matérn covariances. The values of σ_ρ^2 and σ_v^2 are set equal to the known values of the variances of the density and velocity fields in the ambient medium. A value for the characteristic length $l = 1$ is suggested by the known dimensions of the dome cap, under which we seek possible reservoirs. To exemplify the method, we will work with Matérn covariances for $\gamma = 1/2$ and dimension $n = 2$. We set $\gamma = 1/2$ because it allows for irregular patterns, and it largely simplifies the calculation of the Matérn covariances due to the availability of explicit formulas for this kernel. More costly Bayesian formulations that include the hyperparameters as additional unknown variables to be sampled are proposed in [55], for instance.

In practice, we will use finite-dimensional approximations constructed as follows. Notice that ρ and v_{p} in (15) must remain positive. We can enforce this constraint in two different ways. In the first, we choose a truncated multivariate normal distribution $p_{\mathrm{pr}} = \mathcal{N}(0, C_{2N})$ as a prior distribution when the resulting ρ and v_{p} are positive and zero otherwise. Here, C_{2N} is formed, combining two finite-dimensional approximations C_N to the covariance operator C, one acting on ν_ρ and the other one on ν_v. We relabel the basis functions as ϕ_ℓ using the relation $\ell(q, r) = (q-1)N_{\mathrm{x}} + r$, $1 \leq q \leq N_{\mathrm{x}}$, $1 \leq r \leq N_{\mathrm{y}}$. N_{x} and N_{y} represent the number of modes selected in each direction. Truncating the

Karhunen–Loève expansions (19) of the random fields ρ_c and $v_{p,c}$ to $N = N_x N_y$ modes and inserting

$$\rho_{\boldsymbol{\nu}_\rho} = \sum_{\ell=1}^{N} \nu_{\rho,\ell}\phi_\ell, \quad v_{p,\boldsymbol{\nu}_v} = \sum_{\ell=1}^{N} \nu_{v,\ell}\phi_\ell, \quad \nu_{\rho,\ell}, \nu_{v,\ell} \sim \mathcal{N}(0, \lambda_\ell) \tag{20}$$

in (15), we have an approximate parametric representation (18) of the density and velocity fields in the truncated model (4) in terms of the random variables $\boldsymbol{\nu} = (\boldsymbol{\nu}_\rho, \boldsymbol{\nu}_v)$. We denote by $\mathcal{F}(\boldsymbol{\nu})$ the observation operator (5) corresponding to the fields (18) with $\boldsymbol{\nu}_\rho$ and $\boldsymbol{\nu}_v$ given by (20).

The second strategy to enforce positivity constraints resorts to lognormal distributions, setting

$$\rho(x,y) = \exp(\log(\rho(x,y))), \quad v_p(x,y) = \exp(\log(v_p(x,y))), \tag{21}$$

where $\log(\rho(x,y))$ and $\log(v_p(x,y))$ follow normal distributions. To determine the distribution parameters, notice that a distribution lognormal(μ, σ^2) has mean $\mu_0 = \exp(\mu + \sigma^2/2)$ and variance $\sigma_0 = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)$. Then, $\mu = \log(\rho_0) - \frac{1}{2}\log((\sigma_0/\rho_0^2) + 1) \sim \log(\rho_0) - \frac{1}{2}\sigma_0/\rho_0^2$ (with $\mu \sim \log(\rho_0)$ if $\sigma_0 << \rho_0^2$) and $\sigma^2 = \log(\sigma_0/\rho_0^2) + 1)$ (see [56]). In this case, we expand

$$\log\rho - \log\rho_0 = \sum_{\ell=1}^{N} \nu_{\rho,\ell}\phi_\ell, \quad \log v_p - \log v_{p,0} = \sum_{\ell=1}^{N} \nu_{v,\ell}\phi_\ell. \tag{22}$$

We denote by $\mathcal{F}(\boldsymbol{\nu})$ the observation operator (5) corresponding to the fields (21) and (22). The prior distribution is $\mathcal{N}(0, C_{2N})$, formed by combining two finite-dimensional approximations C_N to the covariance operator C, one acting on $\boldsymbol{\nu}_\rho$ and the other one on $\boldsymbol{\nu}_v$.

In either case, with these choices, the likelihood is given by (7) and the posterior distribution to be analyzed is given by (8). In view of the spatial correlations, we sample the resulting posterior probability (2) by means of Markov chain Monte Carlo methods, more specifically, a functional ensemble sampler (FES) (see [25]). This sampler follows a Metropolis-within-Gibbs approach that combines AIES for a subset of dominant modes that uses AIES on E_L the low-wavenumber KL components and the preconditioned Crank–Nicolson sampler (pCN) on the high-wavenumber KL components. This method remains stable as we increase dimensionality, that is, the number of total modes we keep. If $E_L = 0$, the method is just pCN. However, for small positive E_L it may converge faster, though it is quite difficult to tune adequate values of E_L. We have adapted this algorithm to our setup (see Appendix A for details). In our tests, we have initialized the chains sampling values $\boldsymbol{\nu}^i$, $i = 1, \ldots W$, $W > 4N + 1$, with probability p_{pr}.

Figure 5 illustrates the results. Notice that inclusions of different materials are identified under the salt dome cap, with spatial variations from top to bottom, in these visualizations of the MAP estimates. These deposits are difficult to identify by standard seismic imaging techniques due to the screening effect of the dome cap. From the theoretical point of view, this approach is more stable that the Fourier approach. Well-posedness and stability as the number of modes tends to infinity can be proved in the framework developed in [14] (see also [55]). However, uncertainty about the details of the deposits is still large. For the MAP estimates, the values of the densities and the velocities in the dome region range in the intervals $[0.4369, 0.8655]$ and $[0.8326, 1.5136]$ (respectively) for oil and $[0.4204, 0.8035]$ and $[0.3329, 0.6595]$ (respectively) for gas. For the mean estimates, the values of the densities and the velocities in the dome region range in the intervals $[0.3861, 0.9026]$ and $[0.3679, 2.4381]$ (respectively) for oil and $[0.5406, 1.0403]$ and $[0.3679, 2.7183]$ (respec-

tively) for gas. Figure 6 represents density-velocity histograms constructed from the sample values in the oil and gas regions. Moreover, the results are strongly dependent on the prior covariances. In practice, the covariance parameters are unknown, and they should be adjusted through a more complex Bayesian formulation that includes them as additional unknown hyperparameters with their own probability distribution [55]. Nevertheless, the question arises of whether Mátern covariances are a good choice for this particular problem [34,35].

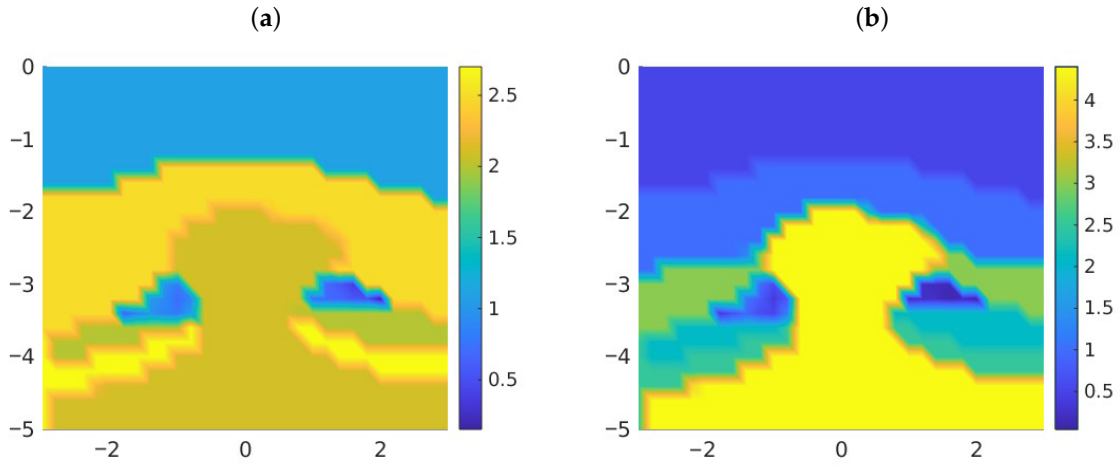

Figure 5. (**a**) Density ρ and (**b**) velocity fields v_p for the salt dome geometry in Figure 1c after 60 steps of the FES algorithm with 400 modes.

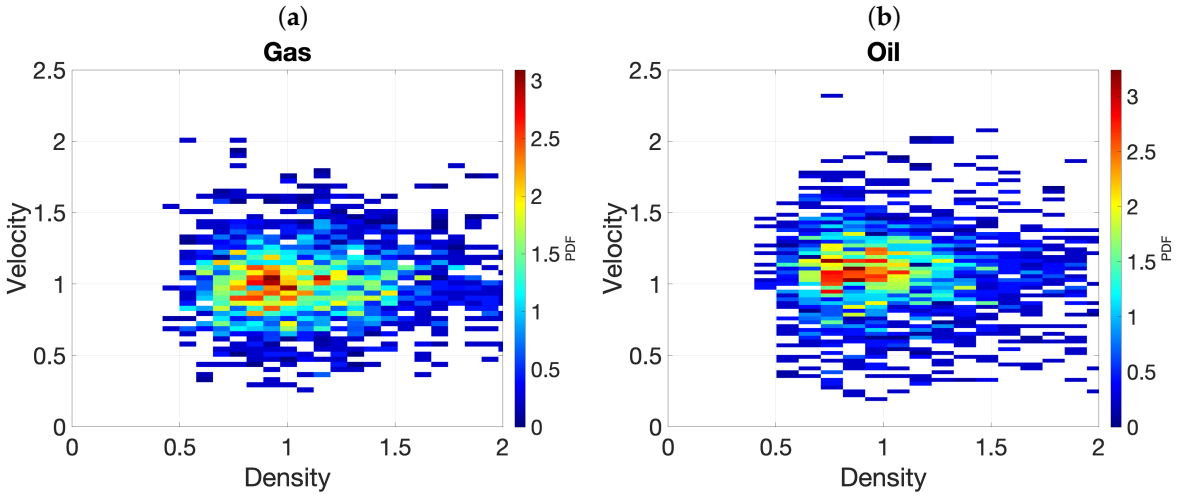

Figure 6. Density–velocity histogram projections constructed from samples in the oil and gas regions.

6. Discussion

We have analyzed different strategies to quantify the uncertainty in inverse scattering problems by means of inverse Bayesian formulations. To this end, we assume we have previous information on the background medium. The goal is to determine the presence of possible anomalies or inclusions in the medium and to characterize them with quantified uncertainty. We focus on imaging applications that employ time-dependent waves, though our techniques are applicable to other imaging setups. In the posterior probability, the likelihood compares the recorded data to the observed values for the proposed anomalies/inclusions as defined by a forward model of wave propagation. Our tests use synthetic data, as detailed at the end of Appendix A.

In homogeneous media, topological energy methods [13] are known to provide information on the number of anomalies and their approximate location and size, but not

on their nature. Working with star-shaped parameterizations, anomalies can be characterized by a low number of random variables. In shear elastography for tissues, Bayesian formulations with topological energy priors for the number and shape of the anomalies, but no prior information on their nature, allow us to extract their shape and material constants with quantified uncertainty from the posterior probability. Tissues are somewhat special because we can approximate their density by a constant everywhere. Anomalies are then characterized by their shear modulus. In stratified media, imaging methods such as migration techniques [42] provide a priori information on the overall layered soil structure. We are interested here in obtaining precise details on the size, orientation, location, and nature of inclusions. Again, assuming no a priori information on the nature of the inclusion, the posterior probability yields information on its nature. However, inclusions now are characterized by two material parameters, density and velocity. This results in an increased uncertainty in the predictions, with several high probability configurations corresponding to different shapes and velocities. Different MCMC tools are adequate depending on the circumstances [21]. In both cases, we have sampled the posterior probability by an affine invariant ensemble MCMC sampler (AIES). These samplers can be parallelized and are efficient for posterior distributions involving few and poorly scaled unknowns [24], even when they are multimodal. In the tests carried out here, the main mode captures the true anomalies/inclusions. These tests use synthetic data generated for incident waves of a single frequency. Combining several frequencies we might be able to reduce uncertainty, suppressing multiple high probability modes. The rationale behind this conjecture is that spurious modes often vary with the frequency. By including data from different frequencies, we may strengthen the main mode (represented in all frequency data) and dilute the variable secondary modes. This phenomenon has already been observed in topological studies [57].

The previous approach is low-dimensional and low in cost. We use prior information on the geometry of inclusions but not on their nature. Often, we ignore that information, and we have to increase dimensionality to capture the details we need. To this end, we represent densities and velocities by random fields and characterize the deposits by analyzing the spatial variation in these fields. We chose to include the available information on the density and velocity of the underlying stratified medium in the direct problem for wave propagation and parameterize the correction, with zero prior information about it. Unlike previous work that parametrizes the fields in terms of their values on a mesh [12], we resort here to global expansions in terms of two different bases. Both approaches lead to Bayesian formulations for a finite number of random variables. In the first one, we represent the corrector fields by truncated Fourier series expansions. The coefficients define a finite collection of random variables for which we formulate the Bayesian inverse problem. In the second one, we represent the corrector fields by random fields, introduce their Karhunen–Loève (KL) expansions and approximate the fields by truncated KL expansions, in terms of a finite number of random variables. This approach can rely on lognormal distributions to ensure the positivity of the resulting densities and velocities. In both cases, we include zero a priori information of the correctors. A delicate point is the choice of the prior covariance since the results depend strongly on this choice. From the theoretical point of view, the KL approach has a number of advantages: it remains stable as the number of modes kept in the truncated expansion grows and the posterior probabilities can be sampled by preconditioned Crank-Nicholson methods (pCN) or functional ensemble (FES) methods in a robust way [25].

We have tested both approaches on a salt dome configuration. We assume we have characterized the main stratified structure by other techniques but lack details on the composition under the dome cap, due to its screening effect. In particular, we ignore

whether oil/gas reservoirs are present. The Fourier strategy barely suggests the presence of deposits under the dome cap. Moreover, the choice of a diagonal covariance allows for unexpected oscillatory behaviors. Instead, the KL strategy relies on matrix covariances that enforce some spatial correlation or definite behaviors on the corrector fields. We have sampled the posterior probability by means of functional ensemble methods, which employ AIES on a small number of modes and pCN on the rest. In this way, we have been able to visualize the presence of reservoirs at the sides of the salt dome. Although the values of the fields obtained are far from the actual values, it is important to mention that despite the initial information on the unknown parameters being zero and the presence of the salt dome attenuating the received signal, it is still possible to identify materials distinct from those of the background fields. Notice that we use zero prior information on the deposits, the results might improve if we add some information or resort to the procedure explained in Section 4. The choice of the covariance matrix is a delicate issue too. To exemplify the procedure, we have used Matérn covariances with ad-hoc hyperparameter choices. That is a restricted class, different covariances are used for other applications [34,35]. We lack a basis to choose a specific type of covariance. On the other hand, the results might improve by considering the covariance hyperparameters as additional unknowns to be sampled [55], which would lead to a more complex Bayesian formulation and a more costly sampling strategy.

Our work shows that Bayesian inversion techniques have a great potential for improving imaging techniques, by allowing us to obtain the most probable solutions, and by quantifying uncertainty in terms of other probable configurations as well as the uncertainty range in the predictions. Starting with simple low-dimensional configurations, we have progressed to study configurations of practical interest in high dimensions. Our study also highlights the importance of having sharp a priori information (which can come from other imaging techniques or previous data studies) and choosing the covariances appropriately.

Author Contributions: Conceptualization, A.C.; methodology, A.C.; software, C.A. and E.C.; validation, A.C., E.C. and C.A.; investigation, A.C., E.C. and C.A.; resources, A.C.; data curation, A.C., E.C. and C.A.; writing—original draft preparation, A.C.; writing—review and editing, A.C., E.C. and C.A.; visualization, E.C. and C.A.; supervision, A.C.; funding acquisition, A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partially supported by the FEDER/Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033) through grant Grant No. PID2024-155528OB-C21 and by a Doctoral Scholarship de la Agencia Nacional de Investigación y Desarrollo, 2022, folio 72220131, Chile (CA).

Data Availability Statement: The paper contains the parameter values necessary to generate synthetic data by solving the forward model and adding noise as explained. In this way, the results are reproducible. The specific realizations employed here could be available upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. AIES and FES Samplers

We next summarize the adapted FES algorithm to sample a posterior probability $p_{\mathrm{pt}}(\boldsymbol{\nu})$, $\boldsymbol{\nu} = (\boldsymbol{\nu}_\rho, \boldsymbol{\nu}_v)$ when $\boldsymbol{\nu}_\rho$ and $\boldsymbol{\nu}_v$ follow distributions $\mathcal{N}(0, \mathbf{C}_\rho)$ and $\mathcal{N}(0, \mathbf{C}_v)$. Set $\mathbf{C}$ equal to a block diagonal matrix with diagonal blocks $\mathbf{C}_\rho$ and $\mathbf{C}_v$. We calculate the eigenvalues and eigenfunctions for the KL expansions and order them in a decreasing order. Then, we store the first E_L eigenvalues and form matrices $\mathbf{J}_\rho$ and $\mathbf{J}_v$ with their ordered column eigenvectors.

Algorithm A1 FES Sampler

1: **Step 1:** Set the size N of $\boldsymbol{\nu}_\rho$ and $\boldsymbol{\nu}_v$ and the number of eigenvalues kept for each $E_L < N$. Compute $\mathbf{P}_\rho = \mathbf{J}_\rho \mathbf{J}_\rho^T$ and $\mathbf{P}_v = \mathbf{J}_v \mathbf{J}_v^T$. Set $\mathbf{P}$ equal to a block diagonal matrix with diagonal blocks $\mathbf{P}_\rho$ and $\mathbf{P}_v$. Set $\mathbf{Q} = \mathbf{I} - \mathbf{P}$.

2: **Step 2:** Select the number of chains W and steps $I_{\max}$. Initialize $\boldsymbol{\nu}_w$ sampling $\mathcal{N}(0, \mathbf{C}_\rho)$ and $\mathcal{N}(0, \mathbf{C}_v)$ for $w = 1, \ldots, W$, $W > 4E_L$, to generate the two blocks of components. Set the sampling parameters $a = 2$ and $\omega = 0.02$.

3: **Step 3:** Calculate the associate densities and velocities $\rho(\boldsymbol{\nu}_w)$ and $v_{\mathrm{p}}(\boldsymbol{\nu}_w)$, by adding the parameterized corrections to their background values according to the selected parameterization, for $w = 1, \ldots, W$.

4: **Step 4:** Calculate the logarithm of the prior distribution for each w: $\mathrm{logprior}_w = -\frac{\boldsymbol{\nu}_w^T \mathbf{C}^{-1} \boldsymbol{\nu}_w}{2}$.

5: **Step 5:** Solve the forward wave problem with fields $\rho(\boldsymbol{\nu}_w)$ and $v_{\mathrm{p}}(\boldsymbol{\nu}_w)$ for $w = 1, \ldots, W$ and evaluate the observation operator $\mathcal{F}(\boldsymbol{\nu}_w)$.

6: **Step 6:** Calculate the logarithm of the likelihood for $w = 1, \ldots, W$: $\mathrm{loglike}_w = -\frac{\|\mathcal{F}(\boldsymbol{\nu}_w) - \mathbf{d}\|_2^2}{2\sigma_{\mathrm{n}}^2}$.

7: **Step 7:** Calculate the logarithm of the posterior probability fpr $w = 1, \ldots, W$

$$\mathrm{logpost}_w = \mathrm{logprior}_w + \mathrm{loglike}_w.$$

8: **Step 8:** Initialize $\boldsymbol{\nu}_{\mathrm{prev}_w}$, $\mathrm{logpostprev}_w$ and $\mathrm{logpriorprev}_w$ as $\boldsymbol{\nu}_w$, $\mathrm{logpost}_w$ and $\mathrm{logprior}_w$.

9: **Step 9:** Apply FES

10: **for** $i = 1$ **to** $I_{\max}$ **do**

11: **Step 9.1:** Apply AIES to the selected first E_L modes in each block

12: **Step 9.1.1:** Generate a permutation `per` of $(1, 2, 3, \ldots, W)$ without fixed sites.

13: **Step 9.1.2:** Generate random numbers and vectors:

14: $\mathrm{vecz} = ((a-1) \times \mathrm{randw} + 1)^2 / a$, and $\mathrm{logvecz} = \log(\mathrm{vecz})$,
where randw is a vector formed by W random numbers $\mathcal{U}(0, 1)$ and operations are performed component wise.

15: **Step 9.1.3:** Perform the cycle

16: **for** $w = 1$ **to** W **do**

17: **Step 9.1.3.1:** Propose the update

$$\boldsymbol{\nu}_{\mathrm{prop},w} = \boldsymbol{\nu}_{\mathrm{prev},w} + (1 - \mathrm{vecz}_w)\, \mathbf{P}\, (\boldsymbol{\nu}_{\mathrm{prev},\mathrm{per}(w)} - \boldsymbol{\nu}_{\mathrm{prev},w})$$

18: **Step 9.1.3.2:** Calculate $\mathrm{logprior}_w = -\frac{\boldsymbol{\nu}_w^T \mathbf{C}^{-1} \boldsymbol{\nu}_w}{2}$.

19: **Step 9.1.3.3:** Generate logr_w as $\log(\mathcal{U}(0, 1))$.

20: **if** $\mathrm{logr}_w < \mathrm{logpriorprop}_w - \mathrm{logpostprev}_w + \mathrm{logvecz}_w \times (2E_L - 1)$ **then**

21: **Step 9.1.3.3.1:** Solve the forward wave problem with fields $\rho(\boldsymbol{\nu}_w)$ and $v_{\mathrm{p}}(\boldsymbol{\nu}_w)$ for $w = 1, \ldots, W$ and evaluate the observation operator $\mathcal{F}(\boldsymbol{\nu}_w)$.

22: **Step 9.1.3.3.2:** Calculate $\mathrm{loglikeprop}_w = -\frac{\|\mathcal{F}(\boldsymbol{\nu}_{\mathrm{prop},w}) - \mathbf{d}\|_2^2}{2\sigma_{\mathrm{n}}^2}$.

Algorithm A1 *Cont.*

Step 9.1.3.3.3: Calculate $\text{logpostprop}_w = \text{logpriorprop}_w + \text{loglikeprop}_w$.

Step 9.1.3.3.4: Calculate the logarithm of the acceptance probability α

$$\log \alpha = \text{logpostprop}_w - \text{logpostprev}_w + \text{logvecz}_w \times (2E_L - 1)$$

Step 9.1.3.3.5: Accept or reject the proposal. Generate logr_w as $\log(\mathcal{U}(0,1))$.

if $\text{logr}_w < \log(\alpha)$ **then**

$\boldsymbol{\nu}_{\text{prev},w} = \boldsymbol{\nu}_{\text{prop},w}$

$\text{logpostprev}_w = \text{logpostprop}_w$

$\text{loglikeprev}_w = \text{loglikeprop}_w$

end if

end if

end for

Step 9.2: Apply pCN to the remaining modes in each block

for $w = 1$ **to** W **do**

Step 9.2.1: Propose the update

$$\boldsymbol{\nu}_{\text{prop},w} = \mathbf{P}\,\boldsymbol{\nu}_{\text{prev},w} + \mathbf{Q}\left(\sqrt{1-\omega^2}\,\boldsymbol{\nu}_{\text{prev},w} + \omega\,\boldsymbol{\xi}_w\right),$$

where $\boldsymbol{\xi}_w \sim \mathcal{N}(0, \mathbf{C})$.

Step 9.2.2: Solve the forward wave problem with fields $\rho(\boldsymbol{\nu}_w)$ and $v_{\text{p}}(\boldsymbol{\nu}_w)$ for $w = 1, \ldots, W$ and evaluate the observation operator $\mathcal{F}(\boldsymbol{\nu}_w)$.

Step 9.2.3: Calculate $\text{loglikeprop}_w = -\frac{\|\mathbf{f}(\boldsymbol{\nu}_{\text{prop},w}) - \mathbf{d}^{\text{obs}}\|_2^2}{2\sigma_{\text{noise}}^2}$.

Step 9.2.4: Calculate the logarithm of the acceptance probability

$$\log \alpha = \text{loglikeprop}_w - \text{loglikeprev}_w.$$

Step 9.2.5: Accept or reject the proposal. Generate logr_w as $\log(\mathcal{U}(0,1))$.

if $\text{logr}_w < \log(\alpha)$ **then**

Calculate $\text{logpriorprop}_w = -\frac{\boldsymbol{\nu}_{\text{prop},w}^T \mathbf{C}^{-1} \boldsymbol{\nu}_{\text{prop},w}}{2}$,

$\boldsymbol{\nu}_{\text{prev},w} = \boldsymbol{\nu}_{\text{prop},w}$,

$\text{loglikeprev}_w = \text{loglikeprop}_w$,

$\text{logpostprev}_w = \text{loglikeprop}_w + \text{logpriorprop}_w$.

end if

end for

end for

As we run the algorithm, we store the moves for the W chains along $I_{\max}$ steps. Discarding B initial samples for each chain as a burn-up stage, the remaining collection of vectors sample the posterior probability. As said earlier, value p_0 is not needed because the employed ratio cancels it. Gelman–Rubin tests [58] are used to test convergence.

To sample with AIES, we set the number of modes equal to the number of parameters for $\boldsymbol{\nu}_\rho$ and $\boldsymbol{\nu}_v$, that is, $E_L = N$, and skip the pCN part. To accelerate the process, we choose educated starting values for the chains $\boldsymbol{\nu}_w$ that correspond to functions with zero or one or two bumps at each side of the lower part of the dome cap. The number of bumps is randomly selected at each side. The location of their center is randomly selected in a box enclosing the lower part of the dome, as well as their radius. Their heights are randomly generated in $[0, \frac{\max(\rho)-\min(\rho)}{2}]$ and $[0, \frac{\max(v_\mathrm{p})-\min(v_\mathrm{p})}{2}]$, considering the known background values. We keep the choices for which the posterior probability is smaller than the probability of the configuration without oil and gas traps. To sample with pCN we set $E_L = 0$ and skip the AIES part. To sample with FES, we choose $E_L < N$.

Notice that the use of logarithms becomes necessary because the exponentials of negative values can be so small as to be interpreted as zero by a computer, rendering the procedure useless. Therefore, the original algorithm is implemented taking the logarithms of the amounts to be compared or processed.

In the tests presented here, we have employed synthetic data $\mathbf{d}$ generated by numerically solving the forward wave model and corrupting with noise. For completeness, we detail the procedure in Algorithm A2.

Algorithm A2 Generation of synthetic data

1: **Step 1:** Create a spatial mesh in the computational region $[ax, bx] \times [ay, by]$ adapted to the true subsurface configuration.

2: **Step 2:** Create a temporal mesh in the time interval $[0, T]$.

3: **Step 3:** Define the locations of sources and receivers, and the recording times (independent of the previous two meshes).

4: **Step 4:** Generate the background reference density and velocity fields, and construct the corresponding FEM matrices for the selected spatial mesh.

5: **Step 5:** Approximate the solution of the forward wave problem by finite difference discretization in time and finite element discretization in space.

6: **Step 6:** Generate the true data at the receivers by evaluating the numerical solution at the required times and locations.

7: **Step 7:** Generate the synthetic noisy data $\mathbf{d}$ by adding noise of magnitude α.

We can use the formula mentioned in Section 4.2 to characterize the noise, though other choices are possible.

References

1. Biondi, B. *3-D Seismic Imaging*; Gardner, G., Ed.; SEG Investigations in Geophysics 14; Society of Exploration Geophysicists (SEG): Tulsa, OK, USA, 2006.
2. Ritchie, W.; Popovici, M.; Fleidner, M.; Saxon, C. Challenges and opportunities in pre-stack depth imaging of legacy seismic data: An overthrust belt case study. In Proceedings of the SEG/Houston 2005 Annual Meeting, Houston, TX, USA, 6–11 November 2005; pp. 400–403.
3. Sheriff, R.E.; Geldart, L.P. *Exploration Seismology*; Cambridge University Press: Cambridge, UK, 1995.
4. Rawlinson, N.; Pozgay, S.; Fishwick, S. Seismic tomography: A window into deep Earth. *Phys. Earth Planet. Inter.* **2010**, *178*, 101–135. [CrossRef]
5. Maier, A.; Steidl, S.; Christlein, V.; Hornegger, J. (Eds.) *Medical Imaging Systems: An Introductory Guide*; Springer: Cham, Switzerland, 2018.
6. Sarvazyan, A.; Hall, T.J.; Urban, M.W.; Fatemi, M.; Aglyamov, S.R.; Garra, B.S. Overview of elastography—An emerging branch of medical imaging. *Curr. Med. Imaging Rev.* **2011**, *7*, 255–282. [CrossRef]

7. Wang, Y.; Yao, B.; Li, H.; Zhang, Y.; Gao, H.; Gao, Y.; Peng, R.; Tang, J. Assessment of tumor stiffness with shear wave elastography in a human prostate cancer xenograft implantation model. *J. Ultrasound. Med.* **2010**, *36*, 955–963. [CrossRef] [PubMed]
8. Do R.H.G. Shear Wave Elastography. In *Blumgart's Surgery of the Liver, Biliary Tract and Pancreas*; Elsevier: Philadelphia, PA, USA, 2017.
9. Hoyt, K.; Castaneda, B.; Zhang, M.; Nigwekar, P.; di Sant'Agnese, P.A.; Joseph, J.V.; Strang, J.; Rubens, D.J.; Parker, K.J. Tissue elasticity properties as biomarkers for prostate cancer. *Cancer Biomark.* **2008**, *4*, 213–225. [CrossRef] [PubMed]
10. Sandrin, L.; Tanter, M.; Catheline, S.; Fink, M. Shear modulus imaging with 2-D transient elastography. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2002**, *49*, 426–435. [CrossRef] [PubMed]
11. Colton, D.; Kress, R. *Inverse Acoustic and Electromagnetic Scattering*; Springer: Berlin, Germany, 1998.
12. Bui-Thanh, T.; Ghattas, O.; Martin, J.; Stadler, G. A computational framework for infinite dimensional Bayesian inverse problems: Part I. The linearized case, with application to global seismic inversion. *SIAM J. Sci. Comput.* **2013**, *35*, A2494–A2523. [CrossRef]
13. Carpio, A.; Cebrián, E.; Gutiérrez, A. Object based Bayesian full-waveform inversion for shear elastography. *Inverse Probl.* **2023**, *39*, 075007. [CrossRef]
14. Stuart, A.M. Inverse problems: A Bayesian perspective. *Acta Numer.* **2010**, *19*, 451. [CrossRef]
15. Yang, L.; Bi, S.; Faes, M.G.R.; Brogg, M.; Beer, M. Bayesian inversion for imprecise probabilistic models using a novel entropy-based uncertainty quantification metric. *Mech. Syst. Signal Process.* **2022**, *162*, 107954. [CrossRef]
16. Abugattas, C.; Carpio, A.; Cebrián, E.; Oleaga, G. Quantifying uncertainty in inverse scattering problems set in layered environments. *Appl. Math. Comput.* **2025**, *500*, 129453. [CrossRef]
17. Hohage, T.; Schormann, C. A Newton-type method for a transmission problem in inverse scattering *Inverse Probl.* **1998**, *1*, 1207–1227. [CrossRef]
18. Carpio, A.; Dimiduk, T.G.; Le Louer, F.; Rapun, M.L. When topological derivatives met regularized Gauss-Newton iterations in holograhic 3D imaging. *J. Comput. Phys.* **2019**, *388*, 224–251. [CrossRef]
19. Esser, E.; Guasch, L.; van Leeuwen, T.; Aravkin, A.Y.; Herrmann, F.J. Total-variation regularization strategies in full-waveform inversion. *SIAM J. Imaging Sci.* **2018**, *11*, 376–406. [CrossRef]
20. Abugattas, C. Métodos Matemáticos y Computacionales en Elastografía. Ph.D. Thesis, UCM, Madrid, Spain, 2025.
21. Carpio, A. Uncertainty Quantification in Scientific Computing: A Short Introduction. In *Mathematical and Computational Modelling Across the Scales*; Diez, P., Giacomini, M., Eds.; EHF-JLL 2023; SEMA SIMAI Springer Series; Springer: Cham, Switzerland, 2025; Volume 39.
22. Kamiński, M. Generalized stochastic perturbation technique in engineering computations. *Math. Comput. Model.* **2010**, *51*, 272–285. [CrossRef]
23. Pozza, F.; Zanella, G. On the fundamental limitations of multi-proposal Markov chain Monte Carlo algorithms. *Biometrika* **2025**, *112*, asaf019. [CrossRef]
24. Goodman, J.; Weare, J. Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci.* **2010**, *5*, 65–80. [CrossRef]
25. Coullon, J.; Webber, R.J. Ensemble sampler for infinite-dimensional inverse problems. *Stat. Comput.* **2021**, *31*, 28. [CrossRef]
26. Fresca, S.; Dede, L.; Manzoni, A. A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized PDEs. *J. Sci. Comput.* **2021**, *87*, 61. [CrossRef]
27. Gunzburger, M.D.; Peterson, J.S.; Shadid, J.N. Reduced-order modeling of time-dependent PDEs with multiple parameters in the boundary data. *Comput. Methods Appl. Mech. Eng.* **2007**, *196*, 1030–1047. [CrossRef]
28. Yang, L.; Zhang, X.; Zhu, F.; Wang, Z.; Zhang, X. Semi-supervised cross-domain fault diagnosis via contrastive pre-training and annotation-efficient alignment strategy. *J. Ind. Inf. Integr.* **2026**, *50*, 101076. [CrossRef]
29. Sanjuan, A. *Caracterización de Anomalías Mediante Aprendizaje Automático*; Trabajo de fin de grado; UCM: Madrid, Spain, 2024.
30. Uribe, F.; Papaioannou, I.; Betz, W.; Straub, S. Bayesian inference of random fields represented with the Karhunen–Loève expansion. *Comput. Methods Appl. Mech. Eng.* **2020**, *358*, 112632. [CrossRef]
31. Dunlop, M.M. Analysis and Computation for Bayesian Inverse Problems. Ph.D. Thesis, University of Warwick, Warwick, UK, 2016.
32. MacKay, D.J.C. *Information Theory, Inference and Learning Algorithms*; Cambridge University Press: Cambridge, UK, 2003.
33. Sullivan, T.J. *Introduction to Uncertainty Quantification*; Texts in Applied Mathematics 63; Springer: Cham, Switzerland, 2015.
34. Díez, P.; Muixí, A.; Zlotnik, S.; García-González, A. Nonlinear dimensionality reduction for parametric problems: A kernel proper orthogonal decomposition. *Int. J. Numer. Methods Eng.* **2021**, *122*, 7306–7327. [CrossRef]
35. Rocas, M.; García-González, A.; Larráyoz, X.; Díez, P. Adaptive surrogates of crashworthiness models for multi-purpose engineering analyses accounting for uncertainty. *Finite Elem. Anal. Des.* **2022**, *203*, 103694. [CrossRef]
36. Carpio, A.; Dimiduk, T.G.; Rapun, M.L.; Selgas, V. Noninvasive imaging of three-dimensional micro and nanostructures by topological methods. *SIAM J. Imaging Sci.* **2016**, *9*, 1324–1354. [CrossRef]
37. Dominguez, N.; Gibiat, V. Non-destructive imaging using the time domain topological energy method. *Ultrasonics* **2010**, *50*, 367–372. [CrossRef]

38. Carpio, A.; Rapún, M.L. Solving inhomogeneous inverse problems by topological derivative methods. *Inverse Probl.* **2008**, *24*, 045014. [CrossRef]
39. Palafox, A.; Capistrán, M.A.; Christen, J.A. Point cloud-based scatterer approximation and affine invariant sampling in the inverse scattering problem. *Math. Methods Appl. Sci.* **2017**, *40*, 3393–403. [CrossRef]
40. Carpio, A.; Iakunin, S.; Stadler, G. Bayesian approach to inverse scattering with topological priors. *Inverse Probl.* **2020**, *36*, 105001. [CrossRef]
41. Liu, Q.; Charleston, M.A.; Richards, S.A.; Holl, B.R. Performance of Akaike Information Criterion and Bayesian Information Criterion in selecting partition models and mixture models. *Syst. Biol.* **2023**, *72*, 92–105. [CrossRef]
42. Zhu, H.; Li, S.; Fomel, S.; Stadler, G.; Ghattas, O. A Bayesian approach to estimate uncertainty for full-waveform inversion using a priori information from depth migration. *Geophysics* **2016**, *81*, R307–R323. [CrossRef]
43. Hudec, M.R.; Jackson, M.P.A. Terra infirma: Understanding salt tectonics. *Earth-Sci. Rev.* **2007**, *82*, 1–28. [CrossRef]
44. Johann, P.; Martini, A.; Maul, A.; Nunes, J.P. OTC 23681 Reservoir Geophysics in Brazilian Pre-Salt Oilfields. In Proceedings of the Offshore Technology Conference, Houston, TX, USA, 30 April–3 May 2012.
45. Chen, G.; Yang, W.; Liu, Y.; Wang, H.; Huang, X. Salt structure elastic full waveform inversion based on the multiscale signed wnvelope. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4508912.
46. Gonçalves, J.F.; Silva, E.C.N. Salt reconstruction in full-waveform inversion using topology optimization techniques. *Geophys. J. Int.* **2023**, *234*, 1484–1504. [CrossRef]
47. Wang, P.; Zhang, Z.; Mei, J.; Lin, F.; Huang, R. Full-waveform inversion for salt: A coming of age. *Lead. Edge* **2019**, *38*, 204–213. [CrossRef]
48. Eikrem, K.S.; Nævdal, G.; Jakobsen, M.; Chen, Y. Bayesian estimation of reservoir properties-effects of uncertainty quantification of 4D seismic data. *Comput. Geosci.* **2016**, *20*, 1211–1229.
49. BinMakhashen, G.M.; Luqman, H.; Ahmed, M. Evaluation of Bayesian classifier for salt dome detection using texture analysis. In Proceedings of the 2nd Smart Cities Symposium (SCS 2019), Bahrain, Bahrain, 24–26 March 2019; pp. 1–6.
50. Arfken, G. *Métodos Matemáticos Para Físicos*; Academic Press: Mexico City, Mexico, 1981.
51. Folland, G.B. *Fourier Analysis and Its Applications*; American Mathematical Society: Providence, RI, USA, 1992.
52. Dunlop, M.M.; Iglesias, M.A.; Stuart, A.M. Hierarchical Bayesian level set inversion. *Stat. Comput.* **2017**, *27*, 1555–1584. [CrossRef]
53. Antil, H.; Pfefferer, J.; Rogovs, S. Fractional Operators with Inhomogeneous Boundary Conditions: Analysis, Control, and Discretization. *Commun. Math. Sci.* **2018**, *16*, 1395–1426. [CrossRef]
54. Antil, H.; Saibaba, A.K. Efficient algorithms for Bayesian inverse problems with Whittle-Matérn priors. *SIAM J. Sci. Comp.* **2024**, *46*, S176–S198. [CrossRef]
55. Afkham, B.M.; Carpio, A. Simultaneous estimation of seabed and its roughness with longitudinal waves. *arXiv* **2026**, arXiv:2602.01099. [CrossRef]
56. Johnson, N.L.; Kotz, S.; Balakrishnan, N. Lognormal Distributions. In *Continuous Univariate Distributions, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 1994; Volume 1.
57. Carpio, A.; Rapún, M.L. Multifrequency topological derivative approach to inverse scattering problems in attenuating media. *Symmetry* **2021**, *13*, 1702. [CrossRef]
58. Gelman, A.; Rubin, D.B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **1992**, *7*, 457–472. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.