Object based Bayesian full-waveform inversion for shear elastography

Ana Carpio (Universidad Complutense de Madrid),
Elena Cebrián (Universidad de Burgos),
Andrea Gutiérrez (Universidad Complutense de Madrid)

May 15, 2023

Abstract. We develop a computational framework to quantify uncertainty in shear elastography imaging of anomalies in tissues. We adopt a Bayesian inference formulation. Given the observed data, a forward model and their uncertainties, we find the posterior probability of parameter fields representing the geometry of the anomalies and their shear moduli. To construct a prior probability, we exploit the topological energies of associated objective functions. We demonstrate the approach on synthetic two dimensional tests with smooth and irregular shapes. Sampling the posterior distribution by Markov Chain Monte Carlo (MCMC) techniques we obtain statistical information on the shear moduli and the geometrical properties of the anomalies. General affine-invariant ensemble MCMC samplers are adequate for shapes characterized by parameter sets of low to moderate dimension. However, MCMC methods are computationally expensive. For simple shapes, we devise a fast optimization scheme to calculate the maximum a posteriori (MAP) estimate representing the most likely parameter values. Then, we approximate the posterior distribution by a Gaussian distribution found by linearization about the MAP point to capture the main mode at a low computational cost.

Keyword. nverse scattering, Full waveform inversion, Topological energy, Bayesian inference, Markov Chain Monte Carlo, PDE constrained optimization, Laplace approximation

1 Introduction

Medical imaging is a part of biological imaging that aims to reveal internal structures hidden in tissues by non invasive techniques [37], such as X-ray radiography, magnetic resonance imaging, tomography, echography, ultrasound, endoscopy, elastography, tactile imaging, thermography, nuclear medicine and holography. From the mathematical point of view, they all pose inverse problems that aim to deduce the properties of living tissues from observed signals. The typical framework is as follows. A set of emitters launch waves which interact with the tissue. The resulting wave field is recorded at a grid of receptors and analyzed to infer the structure of the medium. Different imaging techniques differ in the waves employed (electromagnetic, acoustic, thermal, elastic, etc), the arrangement of emitters and receivers, and the medium properties monitored.

Harmless modalities using light [40], sound [50] or elastic [51] beams are particularly interesting due to the absence of secondary effects. Elastography is a relatively new imaging technique that maps the elastic properties of soft tissue [51, 54]. Cancerous tumors will often be stiffer than the surrounding tissue (prostate and breast tumors, for instance), whereas damaged livers are harder than healthy ones [3, 31, 54]. While existing technology [49] can distinguish healthy from unhealthy tissue in specific situations, the study of tissues containing multiple anomalies, tiny tumors or little contrast regions may benefit from the development of more refined mathematical approaches.

Here we develop an object based Bayesian full-waveform inversion framework for soft tissue shear elastography with topological priors. Instead of tracking spatial variations of the elastic constants or the wave speeds within the tissue (as is often done in many geophysical and medical applications, see [3, 14, 22, 24, 26, 44, 51, 55], for instance and references therein), we take an inverse scattering approach [12] and represent localized anomalies in the tissue, such as tumors or fibromas, by objects with distinctive elastic constants immersed in the background tissue [29]. A first advantage of this approach is that anomalies are characterized by a few unknowns defining their parametrization and their elastic properties, which reduces the computational cost. Moreover, studies in other imaging set-ups [7] suggest that localized inhomogeneities may be more precisely captured by looking for abrupt interfaces defining their boundaries, and for material parameter variations within them, than by tracking the spatial variations of material parameter fields everywhere. A second advantage is that we can define misfit

functionals in terms of object shapes and then use the associated topological energies to construct sharp priors at low cost. Because of their ability to suppress oscillations in configurations with multiple objects, topological energies are used in deterministic inverse scattering frameworks to find first guesses of scatterers in nondestructive materials testing [15, 16] and in biological applications [48].

Classical deterministic full-waveform inversion techniques typically start from a guess of the elastic modulus defined everywhere, see [14, 22] for total variation regularization approaches in geophysics, for instance. If we wish to use the information provided by topological energies about the geometry of anomalies to initialize these methods in our imaging set-up, we need to set initial values for the shear modulus in the regions occupied by them, different from the value for the background tissue. Instead, in our object based Bayesian framework, we can incorporate initial information about the geometry of possible anomalies without choosing a different initial value for the shear modulus of the anomalies, not to bias the predictions with uneducated choices. This can also be done in object based deterministic inversion schemes [11, 30] with Tikhonov type and total variation regularizations, as we will discuss later in Section 6. Compared to them, the Bayesian approach not only provides the most likely configuration for the anomalies and their shear modulus, given a noisy dataset, but also quantifies uncertainty ranges in the predictions of the properties of the anomalies for a noise level.

Depending on the expected complexity, one can choose different representations for the boundaries of the anomalies [25, 28, 33, 43]. Here, we consider two types of star-shaped parameterizations that differ in the way the radius is parameterized. The boundary of star-shaped objects is defined by an angle dependent distance function (the radius) along rays in all space directions [4]. We can reproduce smooth shapes approximating the radius by trigonometric polynomials involving just a few parameters [30]. Rougher boundaries are better described by high dimensional radius functions [18]. Both situations are of interest to study anomalies in tissues. Tumors, for instance, can display smooth or irregular contours depending on their stage and nature. We show that we can infer the structure of anomalies in tissues with quantified uncertainty by Markov Chain Monte Carlo (MCMC) sampling of posterior distributions which use priors constructed by topological energy methods and likelihoods defined in terms of the difference of the recorded elastography data and synthetic observations generated numerically for arbitrary anomalies. Affine invariant ensemble samplers [27, 19] work rea-

sonably well when we can approximate the radius function by combinations of trigonometric polynomials. We can also extract basic information on irregular objects defined by higher dimensional radius functions. However, MCMC methods are computationally expensive. For simple shapes, we have also developed fast methods which first optimize to calculate a maximum a posteriori [5, 9] approximation to the anomaly parameters and then sample a linearized approximation of the posterior distribution. The computational cost is much lower, but details on the structure of the posterior distributions, such as multimodality, may be lost.

The paper is organized as follows. Section 2 describes the mathematical model for shear wave imaging in the tissue. Section 3 formulates the Bayesian inversion framework. Section 4 explains how to construct priors for the number of anomalies and their shapes. Section 5 uses ensemble MCMC samplers to solve the Bayesian inverse problem and quantify uncertainty in the solution for relevant configurations characterized by low dimensional parameter sets. Well defined maximum a posteriori (MAP) approximations are identified. Section 6 presents a low cost approach which combines optimization to calculate the MAP point and linearization of the posterior probability about it to quantify uncertainty. Finally, section 7 adapts affine-invariant ensemble sampling methods to infer the structure of high dimensional irregular shapes. A final Appendix contains details on the numerical schemes employed and parameter choices made. Section 8 presents our conclusions.

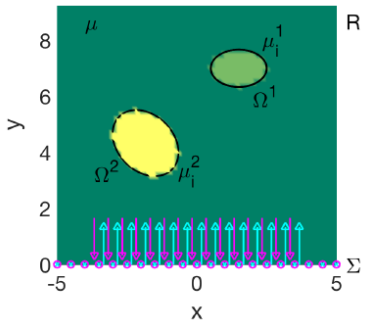

Figure 1: Schematic dimensionless representation of the imaging set-up. The emitters (cyan) generate waves which interact with the medium. The reflected waves are recorded at the receivers (magenta). Emitters and receivers are transducers located at the same position.

2 Physical set-up

Shear elastography tracks variations in the shear modulus μ, which is the property varying more abruptly from healthy to unhealthy tissue, by means of shear waves. Elastic waves in a medium split in shear components (shear S-waves) and compression components (longitudinal P-waves) [35]. Shear waves are adequate for the depths considered in tissues since P-waves travel faster and reach deeper very fast. Moreover, at low frequencies, shear waves are not really affected by attenuation effects in tissues and are governed by standard wave equations, fact already exploited in some medical devices [49]. In principle, we could consider more detailed models to capture the physics of shear wave propagation. However, we keep the model simple to reduce the computational cost in later MCMC studies, which require large amounts of solutions in different configurations.

The imaging set-up is represented in Figure 1. We have in mind applications to inner organs, such as liver, prostate, breast and thyrhoid glands, where source-receiver acquisition devices can be placed closer to the target in specific skin regions. Therefore we consider a partial-coverage acquisition geometry. A medium $R \subset \mathbb{R}^2$ (the tissue) containing a set of anomalies $\Omega = \cup_{\ell=1}^{L} \Omega^\ell$ and locate a set of emitters $\mathbf{x}_j$, $j = 1, \ldots, J$, on a part Σ of the boundary ∂R. We assume that the medium has density ρ and elastic constants μ and λ, while the anomalies have density $\rho_{\rm i}$ and elastic constants $\mu_{\rm i}$ and $\lambda_{\rm i}$. The emitted waves interact with the medium and the resulting wave field is recorded at a grid of receivers $\mathbf{r}_k$, $k = 1, \ldots, K$. Emitter and receivers occupy the same region. They can be interspaced, or, in some set-ups, overlap. Here, we will consider they are transducers located at same positions, playing both roles alternatively. Let us formulate the forward problem that governs the dynamics of the wave field in this framework.

To simplify, we consider that the waves emitted by the sources are governed by the scalar wave equation

$$\begin{aligned} &\rho u_{tt} - \operatorname{div}(\mu \nabla u) = f(t) g(\mathbf{x}), \quad \mathbf{x} \in R,\ t > 0, \\ &u(\mathbf{x}, 0) = 0, u_t(\mathbf{x}, 0) = 0, \qquad \mathbf{x} \in R, \end{aligned} \tag{1}$$

where

$$\rho(\mathbf{x}) = \begin{cases} \rho, & \mathbf{x} \in R \setminus \overline{\Omega}, \\ \rho_{\rm i}^{\ell}, & \mathbf{x} \in \Omega^{\ell},\ \ell = 1, \ldots, L, \end{cases}$$

$$\mu(\mathbf{x}) = \begin{cases} \mu, & \mathbf{x} \in R \setminus \overline{\Omega}, \\ \mu_{\rm i}^{\ell}, & \mathbf{x} \in \Omega^{\ell}, \ell = 1, \ldots, L, \end{cases}$$

with local wave speed $\sqrt{\frac{\mu}{\rho}}$ in the healthy tissue and $\sqrt{\frac{\mu_{\rm i}^{\ell}}{\rho_{\rm i}^{\ell}}}$ inside each anomaly Ω^{ℓ}. In tissues, we have $\rho_{\rm i} \sim \rho$. We assume that the emitters $\mathbf{x}_j$, $j = 1, \ldots, J$, induce source terms of the form $f(t)g_j(\mathbf{x}-\mathbf{x}_j)$, where g_j are smooth functions of narrow support about 0 that we sum to obtain $g(\mathbf{x})$. We represent the function $f(t)$ by a Ricker wavelet $f(t) = f_0(1 - 2\pi^2 f_M^2 t^2)e^{-\pi^2 f_M^2 t^2}$ with peak frequency f_M. The time it takes to move from the initial positive maximum to the negative minimum is $T_D = \frac{\sqrt{6}}{2\pi f_M}$. After that it approaches zero.

A whole organ R can be represented by a domain with zero normal derivative at its physical boundary ∂R

$$\frac{\partial u}{\partial \mathbf{n}} = 0 \quad \text{on} \quad \partial R. \tag{2}$$

Equations (1)-(2) define the forward model, where $\rho \in L^{\infty}(R)$, $\mu \in L^{\infty}(R)$, $\rho \geq \rho_0 > 0$ and $\mu \geq \mu_0 > 0$. Assuming that R and Ω have C^1 boundaries, problem (1)-(2) has a unique solution $u \in C([0,\tau]; H^1(R))$, $u_t \in C([0,\tau]; L^2(R))$, $u_{tt} \in L^2(0,\tau; (H^1(R))')$, for any $\tau > 0$, see [36]. Here, $H^1(R)$ represents the standard Sobolev space and $(H^1(R))'$ its dual space [6]. Since $u_{tt}(\mathbf{x}, 0) = f(0)g(\mathbf{x})/\rho(\mathbf{x}) \in L^2(R)$, we also have $u_t \in C([0,\tau]; H^1(R))$ and $u \in C([0,\tau]; H^2(R \setminus \overline{\Omega}))$, see Appendix. Then, $u(t)$ is defined on Σ and at the receiving sites both in the sense of $L^2(\Sigma)$ traces and pointwise [1, 42].

Assume that $\Omega_{\rm true}$ represents the true anomalies and $\rho_{\rm i,true}$, $\mu_{\rm i,true}$ represent their true material properties. In principle, the values recorded at the receivers constitute the data, that is, $d_{k,\rm true}^m = u(r_k, 0, t_m)$, where u is the solution of (1)-(2) when $\Omega = \Omega_{\rm true}$. In practice, the recorded data d_k^m are corrupted by different sources of noise, that is, $d_k^m = d_{k,\rm true}^m + \text{noise}$. We will assume that the additive noise is distributed as a multivariate Gaussian $\mathcal{N}(0, \mathbf{\Gamma}_{\rm n})$ with mean zero and covariance matrix $\mathbf{\Gamma}_{\rm n}$:

$$d_k^m = d_{k,\rm true}^m + \varepsilon_k^m, \tag{3}$$

for $k = 1, \ldots, K$, $m = 1, \ldots, M$, where ε is distributed according to $\mathcal{N}(0, \mathbf{\Gamma}_{\rm n})$. We consider the noise level for each receiver to be equal and uncorrelated, so that $\mathbf{\Gamma}_{\rm n}$ is the identity matrix of dimension $N = KM$ multiplied by $\sigma_{\rm noise}^2$.

3 Inverse problem

The inverse problem consists in finding the anomalies Ω and their material coefficients $\rho_{\rm i}$ and $\mu_{\rm i}$ such that the solution of the forward problem agrees, in a way to be specified, with the recorded data. For shear elastography in tissues, we take $\rho_{\rm i} \sim \rho$, thus we only have to identify $\mu_{\rm i}$. In a deterministic framework, one typically resorts to optimization formulations: Find objects Ω and parameters $\mu_{\rm i}$ minimizing the cost

$$J(\Omega, \mu_{\rm i}) = \frac{1}{2} \sum_{k=1}^{K} \sum_{m=1}^{M} |u_{\Omega,\mu_{\rm i}}(r_k, 0, t_m) - d_k^m|^2, \tag{4}$$

where $u_{\Omega,\mu_{\rm i}}(r_k, 0, t_m)$ denotes the corresponding solution of the forward problem evaluated at the receivers at the recording times. More refined cost functionals based on optimal transport [21, 39, 17] could be employed. To reduce the occurrence of unphysical minima, the cost (4) is regularized adding additional terms, terms of Tikhonov type, for instance, see [8, 30] and section 6.

To proceed, we need a mathematical representation for the geometry of the anomalies in terms of a set of parameters $\boldsymbol{\nu}$. Star-shaped parameterizations furnish a simple choice to represent the shape of anomalies, though more general representations can be considered too [28, 33, 43]. Star-shaped objects are defined by a center and a radius function that fixes the position of the boundary points along all possible rays emerging from the center. Assume we know the tissue contains L star-shaped anomalies, that is, $\Omega = \cup_{\ell=1}^{L} \Omega^\ell$. Assume $\mu_{\rm i}(x)$ is piecewise constant, equal to $\mu_{\rm i}^\ell$ in Ω^ℓ. Then, different representations of the radius function lead to lower or higher dimensional approaches. We will consider two possibilities.

For smooth star-shaped objects we can approximate the radius function by trigonometric polynomials. Given the data $\mathbf{d} = (d_1^1, \ldots, d_K^1, \ldots, d_1^m, \ldots, d_K^m)$ we wish to predict the $n(L, Q) = L(2Q + 4)$ parameters

$$\boldsymbol{\nu} = (\boldsymbol{\nu}^1, \ldots, \boldsymbol{\nu}^L), \quad \boldsymbol{\nu}^\ell = (c_{\rm x}^\ell, c_{\rm y}^\ell, a_0^\ell, b_1^\ell, a_1^\ell, \ldots, b_Q^\ell, a_Q^\ell, \mu_{\rm i}^\ell), \quad \ell = 1, \ldots, L, \tag{5}$$

representing the centers (c_x^ℓ, c_y^ℓ) and radii $r^\ell(\theta)$ of the anomalies, ordered by blocks, associated to the parameterization

$$\mathbf{q}(\theta)^\ell = (c_x^\ell, c_y^\ell) + r^\ell(\theta)(\cos(2\pi\theta), \sin(2\pi\theta)), \quad \theta \in [0, 1], \tag{6}$$

$$r^\ell(\theta) = a_0^\ell + 2 \sum_{q=1}^{Q} a_q^\ell \cos(2\pi q\theta) + 2 \sum_{q=1}^{Q} b_q^\ell \sin(2\pi q\theta), \tag{7}$$

for $\ell = 1, \ldots, L$. Analogous parameterizations are available in three dimensions replacing Fourier expansions for the radius by expansions in terms of spherical harmonics [8, 25]. The number of modes Q controls the allowed boundary roughness, large values generate more complex shapes [25].

Irregular boundaries are better represented by general radius functions $r(\theta)$, see [2, 4]. In our case, we approximate the boundary by a piecewise linear reconstruction built on a uniform mesh θ_j of $[0, 1]$ with node values $r_j = r(\theta_j)$, $j = 0, \ldots, Z$. The set of anomalies is then represented by the $n(L, Z) = L(3 + Z)$ dimensional parameter set

$$\boldsymbol{\nu} = (\boldsymbol{\nu}^1, \ldots, \boldsymbol{\nu}^L), \quad \boldsymbol{\nu}^\ell = (c_{\mathrm{x}}^\ell, c_{\mathrm{y}}^\ell, r_0^\ell, r_1^\ell, \ldots, r_{Z-1}^\ell, \mu_{\mathrm{i}}^\ell), \quad \ell = 1, \ldots, L, \tag{8}$$

where $r_j^\ell = r^\ell(\theta_j)$, $j = 0, \ldots, Z-1$, and $r_Z = r_0$. The boundary of each object is given by

$$\mathbf{q}(\theta)^\ell = (c_x^\ell, c_y^\ell) + r^\ell(\theta)(\cos(2\pi\theta), \sin(2\pi\theta)), \quad \theta \in [0, 1], \tag{9}$$

$$r^\ell(\theta) = r_j^\ell \frac{\theta - \theta_{j+1}}{\theta_j - \theta_{j+1}} + r_{j+1}^\ell \frac{\theta - \theta_j}{\theta_{j+1} - \theta_j}, \ \theta \in [\theta_j, \theta_{j+1}], \ j = 0, 1, \ldots, Z-1, \tag{10}$$

for $\ell = 1, \ldots, L$. Notice that, while Q is usually small, Z can be very large. Figure 2 compares a star-shaped object defined by (5)-(7), with a star-shaped object defined by a piecewise approximation built from (8)-(10).

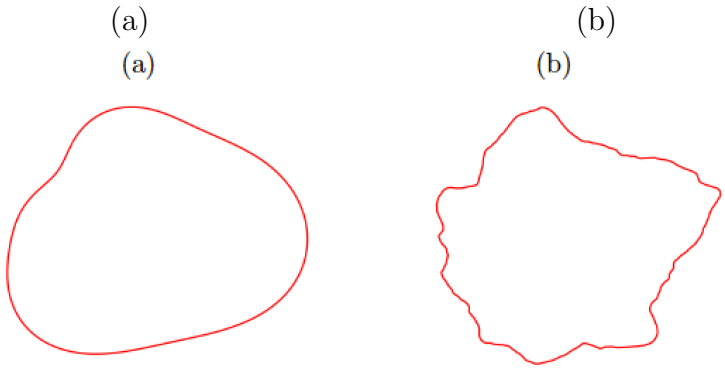

Figure 2: Star-shaped objects with radius defined by (a) a trigonometric polynomial (7) with $Q = 5$ and (b) a piecewise approximation (10) built from a uniform mesh (θ_j, r_j), $j = 0, \ldots, Z$, $Z = 500$, with step $1/Z$.

To quantify uncertainty in the solution of the inverse problem, we resort

to Bayes' formula [32, 52] in finite dimension:

$$p_{\rm pt}(\boldsymbol{\nu}) := p(\boldsymbol{\nu}|\mathbf{d}) = \frac{p(\mathbf{d}|\boldsymbol{\nu})}{p(\mathbf{d})} p_{\rm pr}(\boldsymbol{\nu}). \tag{11}$$

Here, the prior density of the variables $p_{\rm pr}(\cdot)$ incorporates available expert knowledge, while $p(\mathbf{d}|\boldsymbol{\nu})$ represents the conditional probability (or likelihood) of the observations $\mathbf{d}$ given the variables $\boldsymbol{\nu}$. The solution of the Bayesian inverse problem is the posterior density $p_{\rm pt}(\boldsymbol{\nu}|\mathbf{d})$ of the parameters given the data. The density $p(\mathbf{d})$ is a normalization factor that does not depend on the parameters. We choose a likelihood $p(\mathbf{d}|\boldsymbol{\nu})$

$$p(\mathbf{d}|\boldsymbol{\nu}) = \frac{1}{(2\pi)^{N/2}\sqrt{|\boldsymbol{\Gamma}_{\rm n}|}} \exp\Big(-\frac{1}{2} \|\mathbf{f}(\boldsymbol{\nu}) - \mathbf{d}\|^2_{\boldsymbol{\Gamma}_{\rm n}^{-1}} \Big). \tag{12}$$

Here, $\|\mathbf{v}\|^2_{\boldsymbol{\Gamma}_{\rm n}^{-1}} = \overline{\mathbf{v}}^t \boldsymbol{\Gamma}_{\rm n}^{-1} \mathbf{v}$ and $\mathbf{f}(\boldsymbol{\nu})$ represents the measurement operator associated to parameters $\boldsymbol{\nu}$, that is,

$$\mathbf{f}(\boldsymbol{\nu}) = \left(u_{\Omega_{\boldsymbol{\nu}},\mu_{\rm i}}(\mathbf{r}_k, 0, t_m)\right)_{k=1,\dots,K, m=1,\dots,M}, \tag{13}$$

where $u_{\Omega_{\boldsymbol{\nu}},\mu_{\rm i}}$ is the solution of the forward problem and $N = KM$.

We typically choose $p_{\rm pr}(\boldsymbol{\nu})$ as a multivariate Gaussian or a log Gaussian, see Section 4.2 for details. We could implement this approach using prior information obtained by any means, for instance, other imaging systems or other imaging algorithms, see [49]. In the absence of this information, the next section explains how to generate prior knowledge from the data.

4 Topological priors for the anomalies

When we lack prior information on the anomalies, we split the available data and use subsets for different purposes. We insert a fraction of the data in the likelihood (12) and exploit the remainder to obtain prior information. Here, we choose to split the data in two halves: $\mathbf{d}_{\rm odd}$ and $\mathbf{d}_{\rm even}$, which represent the data measured at odd and even times, t_{2m+1} and t_{2m}, respectively. We use $\mathbf{d}_{\rm even}$ to generate prior information on the anomalies from a study of the topological energy of the deterministic cost functional (4). The other fraction, $\mathbf{d}_{\rm odd}$, enters the likelihood, see next section. Here, we show how to extract information on the number, location and size of the anomalies from topological energy methods.

4.1 Calculation of topological energies

Given data $\mathbf{d}_{\text{even}}$, the topological energy [15, 16] for the cost

$$J(R \setminus \overline{\Omega}) = \frac{1}{2} \int_{\Gamma_{obs}} \int_0^{\tau_{\text{end}}} |u_\Omega(\mathbf{x}, s) - \mathbf{d}_{\text{even}}(\mathbf{x}, s)|^2 d\mathbf{x} ds, \tag{14}$$

u_Ω being the solution of (1)-(2) is given by

$$E(\mathbf{x}) = \int_0^{\tau_{\text{end}}} |U(\mathbf{x}, s)|^2 |P(\mathbf{x}, s)|^2 ds, \tag{15}$$

where U and P are adequate forward and adjoint fields. In our set-up, we consider the observation set Γ_{obs} to be a set of receivers. Thus, $\int_{\Gamma_{\text{obs}}}$ in (14) becomes a sum of values at the receivers $\mathbf{r}_k$. Since we record data at discrete time values t_{2m}, we approximate $\int_0^{\tau_{\text{end}}}$ by a sum of values at such times too.

Topological energy fields are related to topological derivatives of shape functionals. Given a shape functional $J(R \setminus \overline{\Omega})$, its topological derivative [7, 29] is a field $D(\mathbf{x})$ defined by the expansion

$$J(R \setminus \overline{B(\mathbf{x}, \varepsilon)}) = J(R) + D(\mathbf{x}) \,\text{meas}(B(\mathbf{x}, \varepsilon)) + o(\text{meas}(B(\mathbf{x}, \varepsilon))), \quad \mathbf{x} \in R,$$

as the radius ε of the ball $B(\mathbf{x}, \varepsilon)$ centered at $\mathbf{x}$ tends to zero. When $D(\mathbf{x}) < 0$, the cost decreases, thus, we expect the true inclusions to be located at regions where the topological derivatives take large negative values. It is usually possible to obtain expressions for $D(\mathbf{x})$ in terms of forward U and adjoint P fields. Topological energies are then defined in terms of their squared absolute values. Their peaks often locate the main regions where the topological derivatives take large negative values. In our case, U is the solution of the forward problem (1)-(2) with inclusion $\Omega = \emptyset$

$$\begin{aligned} &U_{tt} - c^2 \Delta U = f(t) g(\mathbf{x}), && \mathbf{x} \in R, \\ &\tfrac{\partial U}{\partial \mathbf{n}} = 0, && \mathbf{x} \in \partial R, \\ &U(\mathbf{x}, 0) = 0, U_t(\mathbf{x}, 0) = 0, && \mathbf{x} \in R, \end{aligned} \tag{16}$$

for $t \in [0, \tau_{\text{end}}]$ whereas the adjoint field P is a solution of

$$\begin{aligned} &[P_{tt} - c^2 \Delta P](\tau_{\text{end}} - t) = -(U - \mathbf{d}_{\text{even}})(\tau_{\text{end}} - t) \textstyle\sum_{k=1}^K \delta_{\mathbf{r}_k}, && \mathbf{x} \in R, \\ &\tfrac{\partial P}{\partial \mathbf{n}} = 0, && \mathbf{x} \in \partial R, \\ &P(\mathbf{x}, \tau_{\text{end}}) = 0, P_t(\mathbf{x}, \tau_{\text{end}}) = 0, && \mathbf{x} \in R, \end{aligned} \tag{17}$$

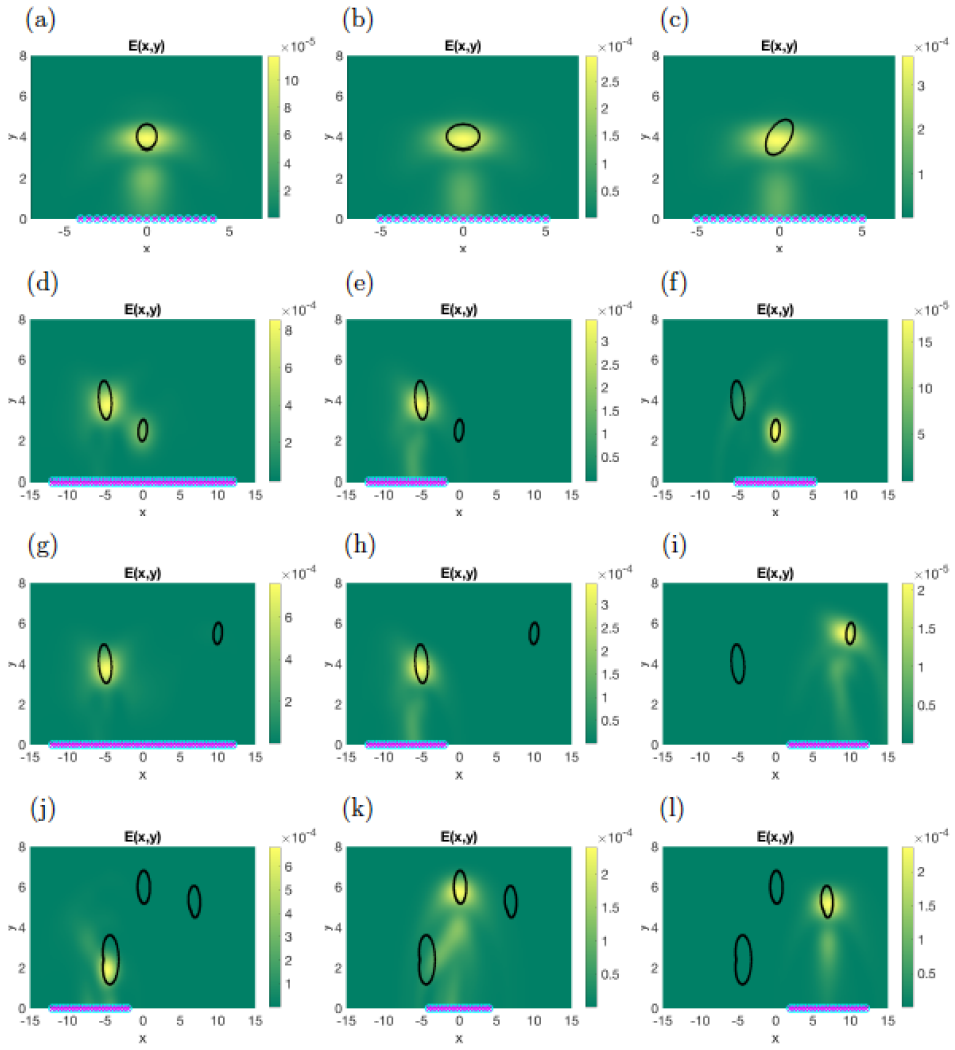

Figure 3: Topological energy fields for: single objects with different sizes and orientations (a) circle, (b) ellipse, (c) rotated ellipse; two objects under different emitter/receiver configurations (d),(g) centered, (e),(h) left sided, (f),(i) right sided; and three star-shaped objects (j)-(l) sweeping the bottom region. Crosses and circles represent emitters and receivers, located at the same position. Black curves represent the true objects. Noise level in the data: 10 %.

for $t \in [\tau_{\text{end}}, 0]$. Here, c is a constant equal to the healthy tissue wave speed everywhere and $\delta_{\mathbf{r}_k}$ represent Dirac masses supported at the receivers. For computational purposes, we replace them by Gaussian regularizations. Notice that problems (16) and (17) can be solved computationally even when c_{i} and μ_{i} are unknown. This is an advantage over alternative methods based on topological derivatives [38], whose calculation usually requires the knowledge of these parameters. The fact that spurious oscillations in the presence of multiple objects are considerably reduced constitutes an additional asset. Topological energies are somewhat related to back-propagation techniques [53] and have been exploited for nondestructive testing of materials and tissues in [15, 16, 48].

The previous description assumes that we record data at the receivers from the time $t = 0$ at which we start to emit. If we start the recording later, at a time τ_{in}, formula (14) integrates from τ_{in} to τ_{end} and the right hand side in (17) is only non zero in $[\tau_{\text{in}}, \tau_{\text{end}}]$. We set the final time $\tau_{\text{end}} \sim 2\frac{H}{c}$, where H is the expected resolution depth.

Figure 3 shows the topological energy fields obtained for several object geometries under different emitter/receiver configurations for the parameter values specified in A.2, after removing dimensions. The data $\mathbf{d}_{\text{even}}$ used to calculate them are synthetic: they are generated by solving numerically the nondimensionalized forward problem (31) in the presence of the true objects, evaluating the solution in the selected space/time data grid and adding 10% noise, as explained in Section 2. To prevent inverse crimes, the fields U in (16) and P in (17) are approximated numerically using rougher meshes: the spatial and time steps for them are twice the steps used when solving numerically to generate the data, and the spatial meshes vary. We have set $\tau_{\text{in}} = 2$ (value at which $f(t)$ almost vanishes for our parameter choice) and $H = 7$ in the calculation of the topological energy.

4.2 Prior construction

Once the topological fields are calculated, we construct a first guess Ω_0 for the anomalies immersed in a background medium R by exploiting the peaks of the topological energy:

$$\Omega_0 = \{\mathbf{x} \in R \mid E(\mathbf{x}) > (1 - C_0) \max_{\mathbf{y} \in R} E(\mathbf{y})\}, \tag{18}$$

where $C_0 \in (0, 1)$ is such that $J(\Omega_0) < J(\emptyset)$. In case several objects are present, we obtain more precise information by sequentially activating frac-

tions of the whole network of emitters/receivers, as shown in Figure 3(d)-(l). We fit circles to the dominant peaks found for each fraction. In this way, we are able to detect all the anomalies. Instead, when we use the information coming from the whole network, we often find the most prominent anomaly only. We use this information to construct priors for the two types of star-shaped parameterizations we consider as follows.

Assuming we locate L peaks, we fit to them circles parametrized by $\boldsymbol{\nu}_0 = (\boldsymbol{\nu}_0^1, \ldots, \boldsymbol{\nu}_0^L)$. When we work with the representation (5)-(7), we set

$$\boldsymbol{\nu}_0^\ell = (c_{\mathrm{x},0}^\ell, c_{\mathrm{y},0}^\ell, a_{0,0}^\ell, b_{1,0}^\ell, a_{1,0}^\ell, \ldots, b_{Q,0}^\ell, a_{Q,0}^\ell, \mu_{\mathrm{i},0}^\ell), \quad \ell = 1, \ldots, L, \tag{19}$$

where $c_{\mathrm{x},0}^\ell, c_{\mathrm{y},0}^\ell$ is the center of mass of each component, $a_{0,0}^\ell$ half the smallest diameter, and $a_{1,0}^\ell = \ldots = a_{Q,0}^\ell = b_{1,0}^\ell = \ldots = b_{Q,0}^\ell = 0$. We also set $\mu_{\mathrm{i},0}^\ell = \mu$, the known background value for the healthy tissue.

Let us denote by $\mathcal{M}$ the set of parameters $\boldsymbol{\nu}$ for which $\mu_{\mathrm{i}}^\ell > 0$ for all ℓ and the curves associated to the parameterization $\boldsymbol{\nu}$ fulfill $r^\ell(\theta) > 0$, for $\theta \in [0, 1]$ and all ℓ, do not intersect, and do not form nested configurations. Notice that $r^\ell(\theta) > 0$ is not a condition on the sign of the curve parameters, but on the sign of the combination (7). Then, we choose $p_{\mathrm{pr}}(\boldsymbol{\nu})$ as a 'modified' multivariate Gaussian with covariance matrix $\boldsymbol{\Gamma}_{\mathrm{pr}}$

$$p_{\mathrm{pr}}(\boldsymbol{\nu}) = \begin{cases} \frac{p_0}{(2\pi)^{n/2}} \frac{1}{\sqrt{|\boldsymbol{\Gamma}_{\mathrm{pr}}|}} \exp(-\frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)^t \boldsymbol{\Gamma}_{\mathrm{pr}}^{-1}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)), & \boldsymbol{\nu} \in \mathcal{M}, \\ 0, & \boldsymbol{\nu} \notin \mathcal{M}, \end{cases} \tag{20}$$

where n is the dimension of $\boldsymbol{\nu}_0$ and p_0 is a normalization constant. In practice, the precise value of this constant is not needed. We choose a diagonal covariance matrix $\boldsymbol{\Gamma}_{\mathrm{pr}}$ formed by L blocks. In our numerical tests, each block starts with $(\sigma_{\mathrm{x}}^\ell)^2 = (\sigma_{\mathrm{y}}^\ell)^2 = 0.1$ and ends with $(\sigma_\mu)^2 = 20^2$. Then $(\sigma_{a_0}^\ell)^2 = 0.1$ and $(\sigma_{a_q}^\ell)^2 = (\sigma_{b_q}^\ell)^2 = 0.1/(1+q^2)^s$, $1 \le q \le Q$, s large, as in [9], so that the prior favors regular shapes with $r(t) > 0$. We fix $s = 3$ and $Q = 5$ for the tests in Sections 5 and 6.

When we work with the representation (8)-(10), we set

$$\boldsymbol{\nu}_0^\ell = (c_{\mathrm{x},0}^\ell, c_{\mathrm{y},0}^\ell, r_{0,0}^\ell, r_{1,0}^\ell, \ldots, r_{Z-1,0}^\ell, \mu_{\mathrm{i},0}^\ell), \quad \ell = 1, \ldots, L, \tag{21}$$

where $c_{\mathrm{x},0}^\ell, c_{\mathrm{y},0}^\ell$ is the center of mass of each component, $r_{0,0}^\ell = r_{1,0}^\ell = \ldots = r_{Z-1,0}^\ell$ is half the smallest diameter and $\mu_{\mathrm{i},0}^\ell = \mu$. Notice that $r_{Z,0} = r_{0,0}$. We choose

$$p_{\mathrm{pr}}(\boldsymbol{\nu}) = p_0 \Pi_{\ell=1}^L p(\boldsymbol{\nu}^\ell), \quad p(\boldsymbol{\nu}^\ell) = p(c_{\mathrm{x}}^\ell) p(c_{\mathrm{y}}^\ell) p(\mu_{\mathrm{i}}^\ell) p(\mathbf{r}^\ell), \tag{22}$$

where $p(c_{\rm x}^\ell)$, $p(c_{\rm y}^\ell)$ are normal distributions with the same means and standard deviations as before, $p(\mu_{\rm i}^\ell)$ a log-normal distribution and $p(\mathbf{r}^\ell)$ is a multivariate log-normal distribution for $r_0^\ell, r_1^\ell, \ldots, r_{Z-1}^\ell$, see [18]. Given a standard normal variable $r = m + \sigma z$ with mean m and standard deviation σ, $e^r = e^{m+\sigma z}$ defines log-normal distribution with parameters m and σ. For $\mu_{\rm i}^\ell$ we use $m = \ln(\mu_{\rm i,0}^\ell)$ and $\sigma = \sigma_\mu^\ell/\mu_{\rm i,0}^\ell$. For $p(\mathbf{r}^\ell)$ we use $\ln(r_{j,0}^\ell)$, $j = 0, \ldots, Z-1$ and Matern covariance matrices, see [18]. We define the Matern covariance between points $(\cos(\theta_i), \sin(\theta_i))$ and $(\cos(\theta_j), \sin(\theta_j))$ separated by a distance d through the expressions [47]

$$C_{\nu,\rho,\sigma}(d) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{d}{\rho}\right)^\nu K_\nu \left(\sqrt{2\nu}\frac{d}{\rho}\right),$$

where Γ is the Gamma function, K_ν the modified Bessel function of the second kind, and ν, ρ, σ are parameters. These parameters govern the length scales and smoothness of curve features, which allows us to consider a wide variety of patterns, see [2] for instance. In our numerical tests we fix $\nu = 3/2$, $\sigma = 0.2$, $\rho = 0.5$. As commented before, the value of the normalization constant p_0 is not needed for MCMC sampling. We will set $L = 1$ in the tests in Section 7.

5 Markov Chain Monte Carlo sampling

We insert the prior distributions obtained in the previous section in the posterior probability $p_{\rm pt}$ given by (11) and (12), with the data $\mathbf{d}_{\rm odd}$ not used to produce the prior information. Then we can sample the unnormalized posterior distribution $q(\boldsymbol{\nu}) = p(\mathbf{d}_{\rm odd}|\boldsymbol{\nu})p_{\rm pr}(\boldsymbol{\nu})$ using Markov Chain Monte Carlo (MCMC) methods. Note that the unknown scaling factor $p(\mathbf{d})$ in (11) is not needed for MCMC sampling. Standard MCMC methods, such as Metropolis-Hastings or Hamiltonian MonteCarlo [41], produce a chain of N-dimensional states $\boldsymbol{\nu}^{(0)} \longrightarrow \boldsymbol{\nu}^{(1)} \ldots \longrightarrow \boldsymbol{\nu}^{(i)} \ldots$ which evolve to be distributed according to the target distribution. One first samples an initial state $\boldsymbol{\nu}^{(0)}$ from the prior distribution, and then moves from one state $\boldsymbol{\nu}^{(i)}$ to the next $\boldsymbol{\nu}^{(i+1)}$ guided by a transition operator. More recent ensemble MCMC samplers [19, 27] draw W initial states (the 'walkers' or 'particles') from the prior distribution and transition to new states while mixing them to construct the chain. This allows us to handle multimodal posteriors [9]

and to parallelize the process for faster exploration of the structure of the posterior distribution.

Different ensemble samplers adapt better to the different parameterizations we consider for the anomalies. Affine-invariant samplers perform well in our set-up. We have considered two. The first one is a stretch move based Affine Invariant Ensemble Sampler (SAIES), see [27]:

- Initialization: Choose W initial states $\boldsymbol{\nu}_w^{(1)} \in \mathbb{R}^d$, $w = 1, \ldots, W$, with probability π (the prior probability $p_{\rm pr}$ in our case) and a value $a > 1$.
- For each step $s = 1, \ldots, S$,
  - For each $w = 1, \ldots, W$
    - Draw $\boldsymbol{\nu}_q^{(s)}$ at random from the set $\{\boldsymbol{\nu}_j^{(s)}\}_{j \neq w}$.
    - Choose a random value z_w from the distribution $g(z) = \frac{1}{\sqrt{z}}$ when $z \in [1/a, a]$, zero otherwise.
    - Set $\boldsymbol{\nu}_{w,\rm prop}^{(s)} = \boldsymbol{\nu}_w^{(s)} + z_w(\boldsymbol{\nu}_w^{(s)} - \boldsymbol{\nu}_q^{(s)})$.
    - Set $\boldsymbol{\nu}_w^{(s+1)} = \boldsymbol{\nu}_{w,\rm prop}^{(s)}$ with probability $\min\left\{1, z_w^{d-1} \frac{p_{\rm pt}(\boldsymbol{\nu}_{w,\rm prop}^{(s)})}{p_{\rm pt}(\boldsymbol{\nu}_w^{(s)})}\right\}$, or else keep $\boldsymbol{\nu}_w^{(s+1)} = \boldsymbol{\nu}_w^{(s)}$.
- Output: The samples $\boldsymbol{\nu}_w^{(s)}$, $w = 1, \ldots, W$, $s = 1, \ldots, S$.

The second one is a general Affine Invariant Ensemble Sampler (AIES), which proceeds as follows, see [19] for instance:

- Initialization: Choose W initial states $\boldsymbol{\nu}_w^{(1)} \in \mathbb{R}^d$, $w = 1, \ldots, W$, with probability π (the prior probability $p_{\rm pr}$ in our case) and a value $\lambda > 0$.
- For each step $s = 1, \ldots, S$,
  - For each $w = 1, \ldots, W$
    - Set $\overline{\boldsymbol{\nu}} = \frac{1}{W-1} \sum_{j \neq w} \boldsymbol{\nu}_j^{(s)}$.
    - Draw z_w with probability $\mathcal{N}(0, 1)$.
    - Set $\boldsymbol{\nu}_{w,\rm prop}^{(s)} = \boldsymbol{\nu}_w^{(s)} + \frac{\lambda}{\sqrt{W-1}} \sum_{j \neq w} z_j(\boldsymbol{\nu}_j^{(s)} - \overline{\boldsymbol{\nu}})$.
    - Set $\boldsymbol{\nu}_w^{(s+1)} = \boldsymbol{\nu}_{w,\rm prop}^{(s)}$ with probability $\min\left\{1, \frac{p_{\rm pt}(\boldsymbol{\nu}_{w,\rm prop}^{(s)})}{p_{\rm pt}(\boldsymbol{\nu}_w^{(s)})}\right\}$, or else keep $\boldsymbol{\nu}_w^{(s+1)} = \boldsymbol{\nu}_w^{(s)}$.

- Output: The samples $\boldsymbol{\nu}_w^{(s)}$, $w = 1, \ldots, W$, $s = 1, \ldots, S$.

Notice that these algorithms involve quotients of posterior probabilities, so that the normalization constants are not needed. We can sample unnormalized distributions.

SAIES is somewhat inspired on fast optimization techniques for multiparametric problems. It is expected to work well for small sets of parameters which are more or less independent. This is the case for star-shaped objects with trigonometric parameterizations. While this sampler evolves quite fast in low dimensions d, it usually requires $W > 2d$ to perform properly. The number of walkers W must be large enough compared to the number of parameters d. As the number of parameters grows, working with too many walkers slows down the mixing and equilibration process. Thus, SAIES is not expected to be effective. In principle, the general AIES does not have that constraint and can be more robust as dimension grows. When the parameterizations use angle meshes for the radius, neighboring points are correlated as reflected by the Matern covariance. We will see that AIES allows us to consider large numbers of such correlated parameters.

Figures 4-8 display results with SAIES for smooth shapes admitting low dimensional parameterizations. AIES provides similar results doubling the number of steps. Section 7 considers high dimensional irregular shapes. There, AIES performs reasonably well with W slightly larger than d. In both cases and for each w, we keep one of each three samples up to a total number of $\tilde{S} = S/3$ to reduce correlations and discard the first $\tilde{S}/5$ as a burn in period. We have set $\sigma_{\text{noise}} = \alpha \max|d^k_{n,\text{true}}|/100$ in $\boldsymbol{\Gamma}_{\text{n}}$ with $\alpha = 10$.

Here, for highly smooth shapes, we consider the parameter set (19) and define $\boldsymbol{\Gamma}_{\text{pr}}$ as in Section 4.2. Then, we insert the prior probability (20) obtained by topological methods and the likelihood (12) in the posterior probability (11) to be sampled. From the samples, we obtain information on the most likely values for $\boldsymbol{\nu}$, that is, the maximum a posteriori (MAP) estimate and the uncertainty about it, depicted in figures 4-8. Figure 4 illustrates the uncertainty in the shape of the anomaly and the value of the parameter representing the dimensionless shear modulus μ_{i} for single shapes: a circle and an ellipse with different orientations. The sample with highest probability defines the MAP point and the mean of the parameters corresponding to all the samples defines a mean estimate. The location and shape of the anomalies is reasonably well captured by both, see also Figure 5 for the uncertainty in geometrical features of interest, such as the location of the center

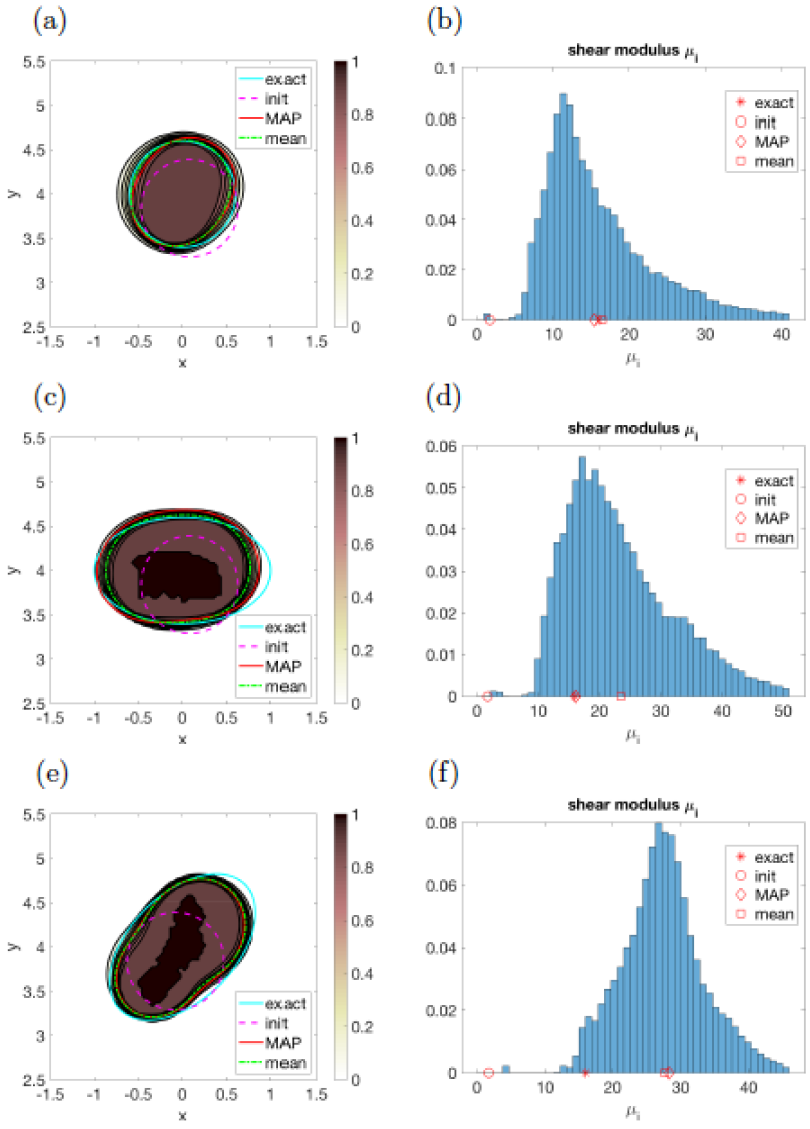

Figure 4: (a), (c), (e) True objects versus MAP estimate and sample mean calculated from MCMC samples for different geometries. The contour levels represent the probability of belonging to the object. (b), (d), (f) Histograms quantifying uncertainty of the MAP estimate and mean values for μ_i. Parameters and samplers: SAIES with $a = 2$, $\tilde{S} = 500$, $W = 480$ and $B = W\tilde{S}/5$.

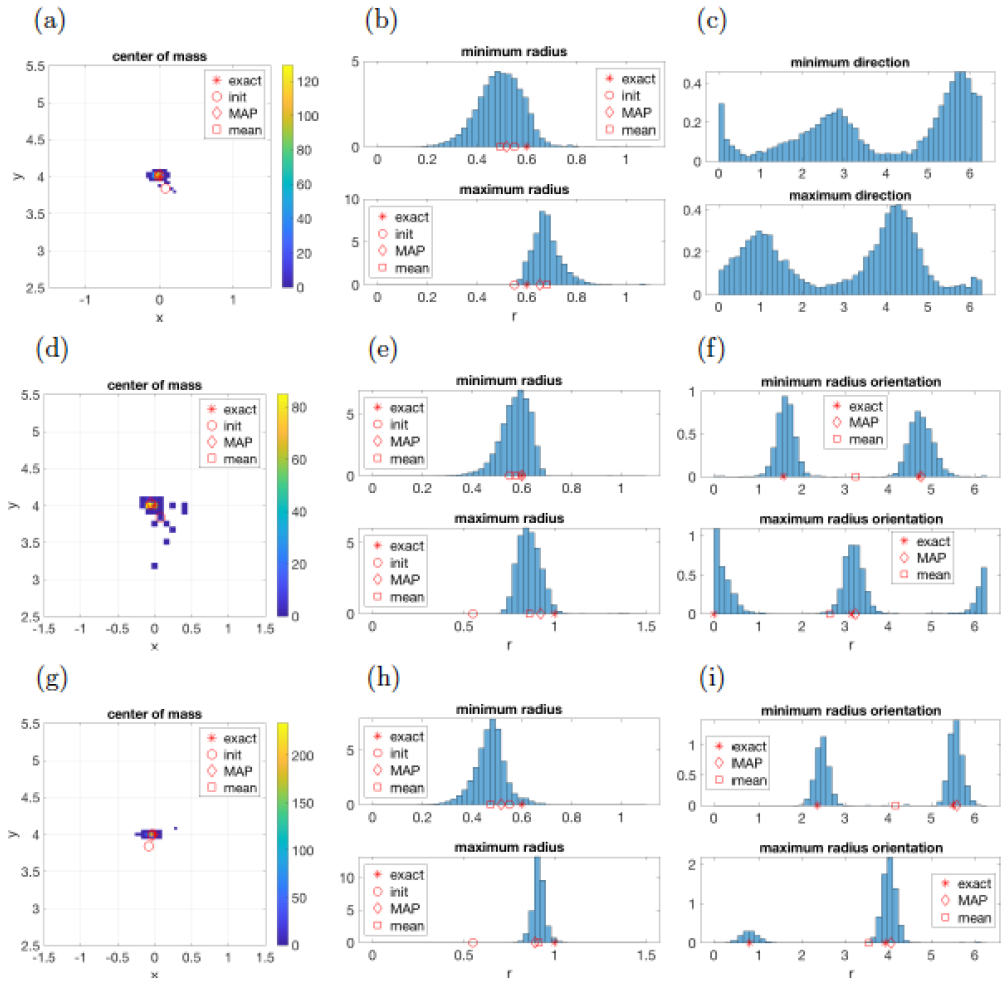

Figure 5: Histograms representing a discrete approximations of the densities for the distribution of the centers of mass (a), (d), (g), shortest and longest radius size (b), (e), (h), and shortest and longest radius orientation (c), (f), (i) in the three test geometries considered in Figure 4. Same sampling parameters.

of mass, the size of the largest and smallest diameters and their orientation. However, the value of the shear modulus displays larger uncertainty, still in the range indicating sickness. The histograms reveal distributions with wide and asymmetric tails. Notice that a change in the orientation of an object can drastically increase uncertainty in the predictions, compare Fig. 4(d)

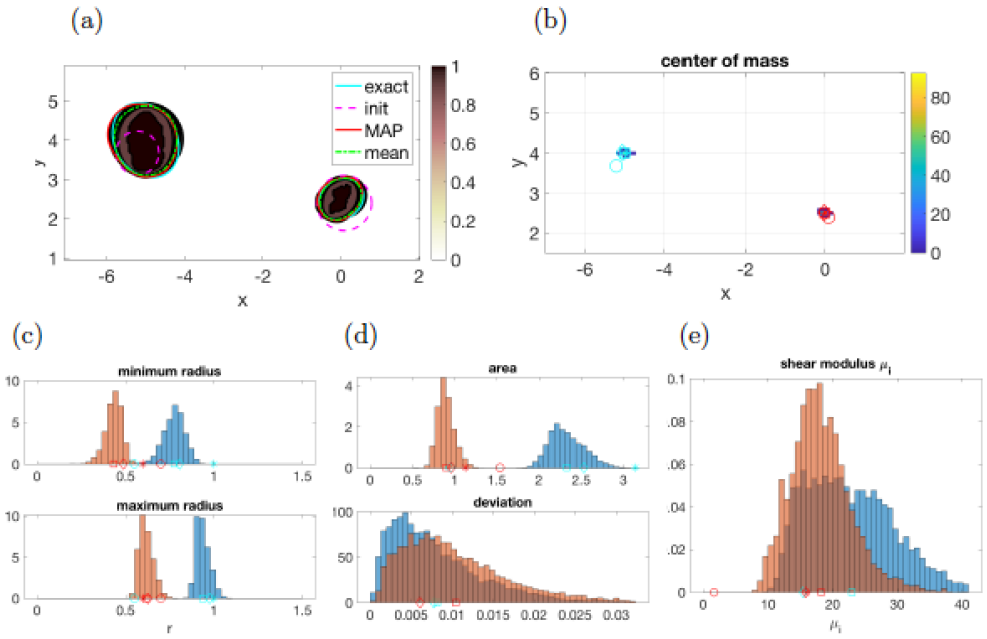

Figure 6: (a) True objects versus MAP estimate and mean calculated from MCMC samples. The contour levels represent the probability of belonging to the object. Superimposed curves represent the exact contours (solid light cyan), the MAP point (solid dark red), the mean (dash-dotted green), and the initial guess (dashed magenta). (b) Histograms representing discrete approximations of the densities for the distribution of the centers of mass. (c)-(d) Histograms for the radius sizes (c), area and deviation from a circular object (d) and shear modulus μ_i (e). Blue histograms and cyan symbols correspond to the large object, orange histograms and red symbols to the small one (asterisk: exact value, circle: initial value, diamond: MAP point, square: mean). Parameters and samplers: SAIES with $a = 2$, $\tilde{S} = 500$, $W = 480$ and $B = W\tilde{S}/5$.

and Fig. 4(f).

Figures 6-8 consider configurations with multiple anomalies. The approximation of the shapes provided by the MAP point and the mean values in figures 6-8 is quite reasonable, regarding both the shapes, their basic geometrical features and the shear modulus, though we observe again wide asymmetric tails.

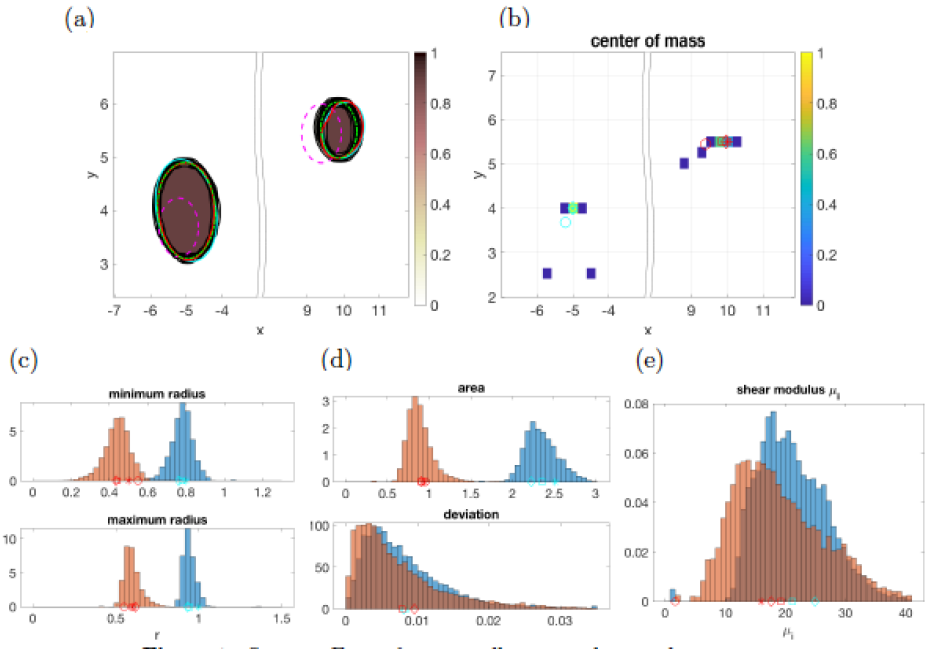

Figure 7: Same as Fig. 6 for two well separated anomalies.

When we include in the prior less anomalies than needed, the distribution may be multimodal: we may spot the missing components. When we include in the prior more anomalies than needed, the spurious ones may essentially vanish because μ_i is basically equal to μ. Notice that our priors contained the correct number of anomalies. The resulting distributions represent a single mode. Also, the means and the MAP points are reasonably close. This suggests that optimization schemes could capture the MAP estimate, allowing for a Laplace approximation of the posterior distribution, which can be sampled at a much lower cost. The computational time drops from a few days to a few minutes.

6 Sampling from a Bayesian linearized formulation

To reduce the computational cost, we analyze the Laplace approximation of the posterior density (11) obtained by linearization at the maximum a

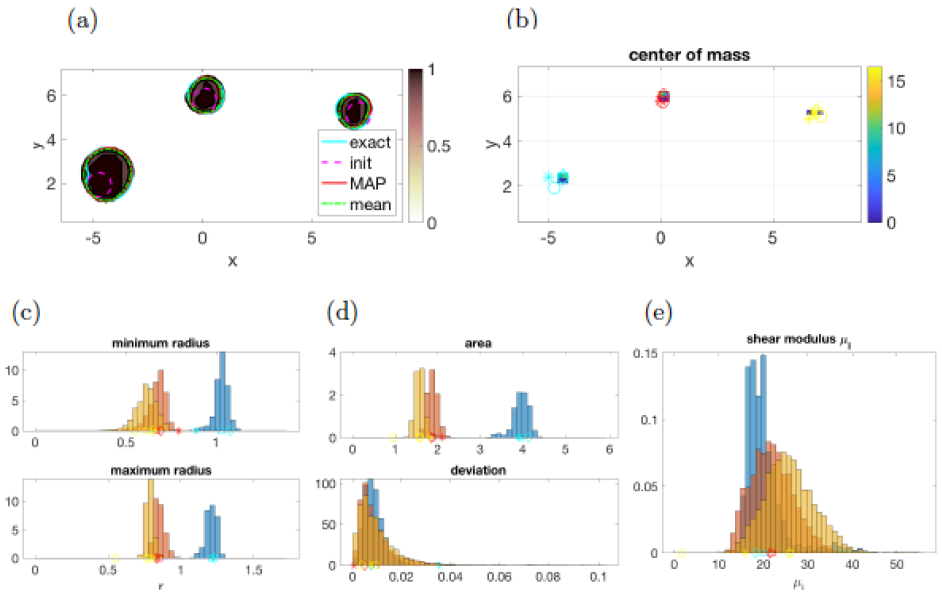

Figure 8: Same as Figure 6 for a configuration with three objects. Orange histograms and red symbols correspond to the middle object, blue histograms and cyan symbols to the left-most one, yellow histograms and symbols to the right-most one.

posteriori (MAP) point $\boldsymbol{\nu}_{\mathrm{MAP}}$. This strategy first computes the vector $\boldsymbol{\nu}_{\mathrm{MAP}}$, which minimizes the negative log likelihood

$$J(\boldsymbol{\nu}) = \frac{1}{2\sigma^2_{\mathrm{noise}}} \sum_{m=1}^{M} \sum_{k=1}^{K} |u_{\boldsymbol{\nu}}(r_k, 0, t_m) - d_k^m|^2 + \frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)^t \boldsymbol{\Gamma}_{\mathrm{pr}}^{-1}(\boldsymbol{\nu} - \boldsymbol{\nu}_0), \quad (23)$$

where $u_{\boldsymbol{\nu}}$ is the solution of the forward problem with object Ω parametrized by $\boldsymbol{\nu}$ given by (5). Then, we approximate the posterior distribution by a multivariate Gaussian $\mathcal{N}(\boldsymbol{\nu}_{\mathrm{MAP}}, \boldsymbol{\Gamma}_{\mathrm{pt}})$ with posterior covariance matrix $\boldsymbol{\Gamma}_{\mathrm{pt}} = \mathbf{H}^{-1}_{\boldsymbol{\nu}_{\mathrm{MAP}}}$, where $\mathbf{H}_{\boldsymbol{\nu}_{\mathrm{MAP}}}$ is an approximation of the Hessian of the measurement operator (13) evaluated at $\boldsymbol{\nu}_{\mathrm{MAP}}$ [5, 55].

6.1 Computing the MAP point

The MAP point is calculated exploiting techniques of deterministic optimization. Taking the prior $\boldsymbol{\nu}_0$ as initial guess of the parametrization, that

is, $\boldsymbol{\nu}^0 = \boldsymbol{\nu}_0$, we can implement the Newton type iteration $\boldsymbol{\nu}^{j+1} = \boldsymbol{\nu}^j + \boldsymbol{\xi}^{j+1}$ where $\boldsymbol{\xi}^{j+1}$ is the solution of

$$\left(\mathbf{H}(\boldsymbol{\nu}^j) + \omega_j \mathrm{diag}(\mathbf{H}(\boldsymbol{\nu}^j))\right) \boldsymbol{\xi}^{j+1} = -\mathbf{g}(\boldsymbol{\nu}^j), \tag{24}$$

see [23], where $\mathbf{H}(\boldsymbol{\nu})$ and $\mathbf{g}(\boldsymbol{\nu})$ represent the Hessian and the gradient of the cost. In practice, to reduce the occurrence of negative radii and the risk of loop formation in the curves, we introduce an additional parameter $\lambda > 0$, replacing (23) by

$$J_\lambda(\boldsymbol{\nu}) = \frac{1}{2\sigma^2_{\mathrm{noise}}} \sum_{m=1}^{M} \sum_{k=1}^{K} |u_{\boldsymbol{\nu}}(r_k, 0, t_m) - d_k^m|^2 + \frac{\lambda}{2} (\boldsymbol{\nu} - \boldsymbol{\nu}_0)^t \boldsymbol{\Gamma}_{\mathrm{pr}}^{-1} (\boldsymbol{\nu} - \boldsymbol{\nu}_0) \tag{25}$$

and (24) by

$$\left(\mathbf{H}^{\mathrm{GN}}_{\lambda_j}(\boldsymbol{\nu}^j) + \omega_j \mathrm{diag}(\mathbf{H}^{\mathrm{GN}}_{\lambda_j}(\boldsymbol{\nu}^j))\right) \boldsymbol{\xi}^{j+1} = -\mathbf{g}_{\lambda_j}(\boldsymbol{\nu}^j). \tag{26}$$

Here, the subscript λ_j indicates that we multiply $\boldsymbol{\Gamma}_{\mathrm{pr}}^{-1}$ by a factor λ_j in the initial iterations to balance the two terms defining the cost J in (23). Notice that we have also replaced the full Hessian by the Gauss-Newton part of the Hessian to reduce the computational cost per iteration. The components of $\mathbf{H}^{\mathrm{GN}}_\lambda(\boldsymbol{\nu})$ and $\mathbf{g}_\lambda(\boldsymbol{\nu})$ are given by:

$$\begin{aligned} (g_\lambda(\boldsymbol{\nu}))_i \quad &= \tfrac{\partial J(\boldsymbol{\nu})}{\partial \nu_i} \quad = \frac{1}{\sigma^2_{\mathrm{noise}}} \sum_{m=1}^{M} \sum_{k=1}^{K} (u_{\boldsymbol{\nu}}(r_k, 0, t_m) - d_k^m) \frac{\partial u_{\boldsymbol{\nu}}}{\partial \nu_i}(r_k, 0, t_m) \\ &\qquad + \lambda [\boldsymbol{\Gamma}_{\mathrm{pr}}^{-1} (\boldsymbol{\nu} - \boldsymbol{\nu}_0)]_i, \end{aligned} \tag{27}$$

$$\begin{aligned} (\mathbf{H}^{\mathrm{GN}}_\lambda(\boldsymbol{\nu}))_{i,\ell} \quad &= \tfrac{\partial^2 J(\boldsymbol{\nu})}{\partial \nu_i \partial \nu_\ell} \quad \sim \frac{1}{\sigma^2_{\mathrm{noise}}} \sum_{m=1}^{M} \sum_{k=1}^{K} \frac{\partial u_{\boldsymbol{\nu}}}{\partial \nu_i}(r_k, 0, t_m) \frac{\partial u_{\boldsymbol{\nu}}}{\partial \nu_\ell}(r_k, 0, t_m) \\ &\qquad + \lambda [\boldsymbol{\Gamma}_{\mathrm{pr}}^{-1}]_{i,\ell}. \end{aligned} \tag{28}$$

The second order derivatives of $u_{\boldsymbol{\nu}}$ are neglected.

To optimize our objective function we implement a double iteration:

- Initially, we set $\omega_0 = 10^{-4}/2$, $\lambda_0 = 0.1\sigma^{-2}_{\mathrm{noise}}$, and $\boldsymbol{\nu}^0 = \boldsymbol{\nu}_0$.
- At each step we calculate $\boldsymbol{\nu}^{j+1} = \boldsymbol{\nu}^j + \boldsymbol{\xi}^{j+1}$, where $\boldsymbol{\xi}^{j+1}$ is the solution of (26). Then

  - We check i) if $r(\theta) > 0$, $r(\theta)$ given by (7), and if $\mu_{\rm i} > 0.5\mu$, ii) if the functional $J_{\lambda_j}(\boldsymbol{\nu})$ decreases replacing $\boldsymbol{\nu}^j$ with $\boldsymbol{\nu}^{j+1}$.
  - If any of these conditions fails, we do not accept $\boldsymbol{\xi}^{j+1}$. We increase ω_j by a factor 2, solve again (26) and check conditions i) and ii) until they are fulfilled.
  - If both conditions are satisfied, we accept $\boldsymbol{\xi}^{j+1}$ and set $\omega_{j+1} = \omega_j/2$ and $\lambda_{j+1} = \max(\lambda_j/5, 1)$.

- After a few steps j_0, $\lambda_{j+1} = 1$ for $j \geq j_0$. When the relative difference between the new value of the cost and the previous one is smaller than a tolerance Tol (here Tol $= 5 \times 10^{-7}$), we freeze all the components except $\mu_{\rm i}^{j+1}$ and iterate with respect to $\mu_{\rm i}$ until variations fall below a threshold 0.02.

To evaluate the derivatives $\frac{\partial u_{\boldsymbol{\nu}}}{\partial \nu_i}(r_k, 0, t_m)$ required for the calculation of (27)-(28) at each step, we use the approximation

$$\frac{\partial u_{\boldsymbol{\nu}}}{\partial \nu_i}(r_k, 0, t_m) \sim \frac{u_{\boldsymbol{\nu}+\eta_i}(r_k, 0, t_m) - u_{\boldsymbol{\nu}}(r_k, 0, t_m)}{\eta_i}$$

with η_i small, $u_{\boldsymbol{\nu}+\eta_i}$ being the solution of the forward problem with ν_i replaced by $\nu_i + \eta_i$. All the forward problems are solved with the same discretization and steps we used in Section 5. In principle, we could calculate these derivatives characterizing Fréchet derivatives in terms of a boundary value problem as we do in [9] for stationary waves governed by Helmholtz equations. However, the resulting boundary value problems involve transmission conditions that are hard to handle in a finite element framework. We would need adapted schemes and codes based on boundary value elements for transmission problems for time dependent waves equations [45], which are still not available.

The values of η_i must be calibrated. Initially, we set for each block $\ell = 1, \ldots, L$ in (5) $\eta_1^\ell = \eta_2^\ell = \eta_3^\ell = \eta_{2Q+4}^\ell = \eta = 0.1$ and $\eta_{3+2i}^\ell = \eta_{2+2i}^\ell = \eta/2$ for $i = 1, \ldots, Q$. As we iterate, we calibrate values for η_i estimating the quotients $\frac{D_{\nu_i} u_{\boldsymbol{\nu}^j}(r_k,0,t_m)}{D^2_{\nu_i} u_{\boldsymbol{\nu}^j}(r_k,0,t_m)}$, where D and D^2 represent approximations of derivatives, and averaging over k and m. In the tests we have performed, the choice

$$\begin{aligned} \eta_1^\ell = \eta_2^\ell = 0.05, \quad \eta_{2Q+4} = 0.15, \\ \eta_{3+2i}^\ell = \eta_{2+2i}^\ell = 0.05, \quad i = 1, \ldots, Q, \end{aligned}$$

gives good results, with η_3 in the range $0.025 - 0.225$.

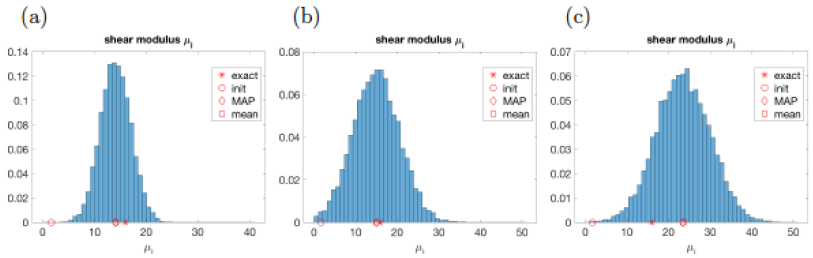

Figure 9: Counterpart of Fig. 4 (b), (d), (f) obtained by calculating the MAP point and linearizing the posterior probability about it. 10000 samples plotted.

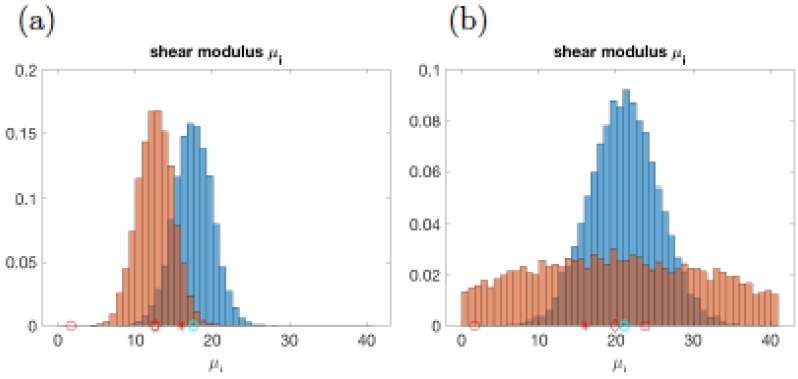

Figure 10: Counterparts of Fig. 6(b) and Fig. 7(b) obtained by linearized Bayesian methods. 10000 samples plotted.

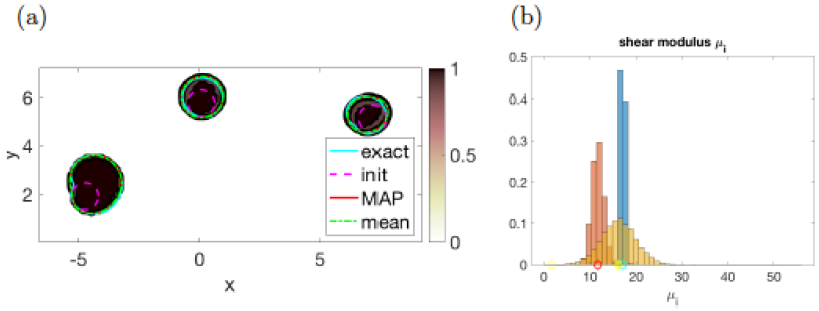

Figure 11: Counterpart of Fig. 8 obtained by linearized Bayesian methods. 10000 samples plotted.

6.2 Sampling

Once we have obtained an approximation to $\boldsymbol{\nu}_{\rm MAP}$, we linearize the posterior distribution about it, approximate by a multivariate Gaussian distribution $\mathcal{N}(\boldsymbol{\nu}_{\rm MAP}, \boldsymbol{\Gamma}_{\rm pt})$ and draw samples from it to quantify uncertainty. We set

$$\boldsymbol{\Gamma}_{\rm pt} = (\mathbf{F}(\boldsymbol{\nu}_{\rm MAP})^t \boldsymbol{\Gamma}_{\rm n}^{-1} \mathbf{F}(\boldsymbol{\nu}_{\rm MAP}) + \boldsymbol{\Gamma}_{\rm pr}^{-1})^{-1} = \mathbf{H}^{\rm GN}(\boldsymbol{\nu}_{\rm MAP})^{-1},$$

where $\mathbf{F}(\boldsymbol{\nu}_{\rm MAP}) = \left(\frac{\partial u_{\boldsymbol{\nu}_{\rm MAP}}}{\partial \nu_i}(p_j)\right)_{j,i}$ and

$$\mathbf{p} = ((r_1, 0, t_1), \ldots, (r_K, 0, t_1), \ldots, (r_1, 0, t_M), \ldots, (r_K, 0, t_M)),$$

that is, $\mathbf{F}(\boldsymbol{\nu}_{\rm MAP})$ is the matrix with ith-column $\frac{\partial u_{\boldsymbol{\nu}}}{\partial \nu_i}(r_k, 0, t_m)$, $k = 1, \ldots, K$, $m = 1, \ldots, M$ evaluated at $\boldsymbol{\nu}_{\rm MAP}$. This can be done by means of the relation

$$\boldsymbol{\nu} = \boldsymbol{\nu}_{\rm MAP} + \boldsymbol{\Gamma}_{\rm pt}^{1/2} \mathbf{n}, \tag{29}$$

$\mathbf{n}$ being a standard normal randomly generated vector (iid).

Figures 9-11 revisit the previous MCMC tests with this procedure. In each case, we optimize to approximate the MAP point and generate a large collection of samples of the posterior distribution by means of (29). The values of the cost for the approximated MAP estimates obtained this way are similar to those for the MAP estimates previously found by MCMC sampling. Comparing the results, we remark that the MAP points and contour curves for the shapes remain similar. Fig 11(a) illustrates this fact in the example with three objects. However, the values of $\mu_{\rm i}$ show larger variability. For single objects, the MAP points and mean values remain alike, while the wide asymmetric tails are lost. The tests with more objects show a similar tendency. Notice that as objects become smaller and distant, information can be lost, as it happens in the red histogram in Fig 10(b) (compare to Fig 7(b)).

The computational cost of this approach is much smaller. The MAP estimate for single objects is obtained in about 10 steps, about 5 minutes in a laptop using MATLAB, while MCMC sampling can take 2-4 days depending on the size of the computational regions.

It is worth comparing these results with predictions obtained by deterministic object based full waveform inversion (FWI) techniques. We have adapted to our framework the object based FWI with total variation regularization methods developed in [11] for Helmholtz constraints and level

set parameterizations, using wave equation constraints and star-shaped parameterizations instead. We can initialize the optimization scheme with our topological energy guess for the shape and constant shear modulus everywhere, as we do here. However, the results we have obtained so far are not good: the shear modulus in the anomaly stays far from the true value and the shape tends to be wavy. Further studies in that direction are needed. With more feneral FWI techniques [14, 22], we cannot use the information on the geometry of the anomalies obtained from topological energy techniques unless we choose an initial value for the shear modulus of the anomalies different from the background value, which changes the starting point. Moreover, the borders of the anomalies tend to be blurred, thus we have not pursued that path.

Notice that the cost functional (23) is the sum of a standard deterministic cost and a regularization term constructed from a scaled parameter norm $\alpha\|\boldsymbol{\nu} - \boldsymbol{\nu}_0\|_X$. It resembles the costs used in deterministic iteratively regularized Gauss Newton methods (IRGN), except that α is fixed, it is not a sequence tending to zero, see [30]. However, the final result is similar using both techniques in the examples studied here, since the regularizing term happens to be small compared to the unregularized cost when we stop. In general, this may not be the case due to the nonvanishing contribution of $\boldsymbol{\nu}_0$ in the Bayesian approach. The Bayesian cost has the potential of inspiring deterministic regularizations for IRGN techniques while providing an interpretation that IRGN techniques lack. As discussed here, we can linearize the posterior probability about the optimum MAP point and sample the resulting Gaussian distribution to extract information on the uncertainty range in the predictions of the properties of the anomalies. Such predictions and uncertainty range may deviate, or not, from the predictions and uncertainty ranges obtained sampling the full posterior distribution by MCMC. This will depend on the particular dataset, the prior knowledge and the data acquisition system under study. In the tests analyzed in this section, we find a reasonable agreement assuming the correct number of anomalies is known. Otherwise, this may not be the case, as it is not in the next Section, where optimization techniques encounter difficulties for a different parameterization and we explore posterior distributions by MCMC techniques.

7 Irregular shapes

Finally, we consider irregular shapes defined by high dimensional parameterizations of the form (8)-(10). We insert the prior distributions (21) obtained by topological methods in the posterior probability given by (11) and (12), with the data $\mathbf{d}_{\text{odd}}$ not used to produce the prior information.

The affine-invariant ensemble sampler AIES described in Section 5 produces the results represented in Figure 12 for an irregular shape. Notice that the use of SAIES would require $W > 2d = 1006$ walkers, which would mix much more slowly. Now, the MAP estimate does not approach the true shape. Nevertheless, the mean profile and the contour plot give an idea of the location and size of the anomaly.

In the previous sections, the probability for negative μ_{i} was set equal to zero. Now, $\mu_{\text{i}} = \exp(\gamma)$, where γ is the random variable that we sample. Notice the peak for μ_{i} near zero. It is due to a family of large samples with small μ_{i}. Figure 12(c) represents the last W samples we obtained. We observe a dominant family of samples which wrap around the true object. A second family is formed by smaller shapes with larger values of μ_{i} placed between the object and the emitters, at the location of a small secondary peak of the topological energy (see Figure 3(c)). The third family corresponds to large samples with small μ_{i} placed behind the true object. The sample distributions we obtain this way are multimodal, though the main mode dominates the rest when averaging to obtain a mean. Notice that uncertainty in the values of μ_{i} with this procedure seems quite large. The prior information we use has low quality in this case. We look for an irregular shape assuming that the prior is a smooth circle and μ_{i} is the value for the healthy tissue. Inconsistency between the prior and the data may lead to multimodality, as pointed out in the previous section.

Working with smooth shapes, we get Figure 13 for the rotated ellipse already studied in Figure 4(c). The samples we generate behave in a similar way as those in Figure 12(c). The MAP point is unlikely to be smooth when we do not enforce the prior knowledge we have on smoothness. However, the information provided by the mean parameters and the statistics of geometrical characteristics and values for shear moduli is still useful, though less precise. Enforcing a smooth parametrization we get better results for smooth shapes, at a lower computational cost, see Figure 4(c). Similarly, the irregular shape studied in Figure 12 could be studied in the smooth framework employed in Section 5 to obtain information about mean values at a

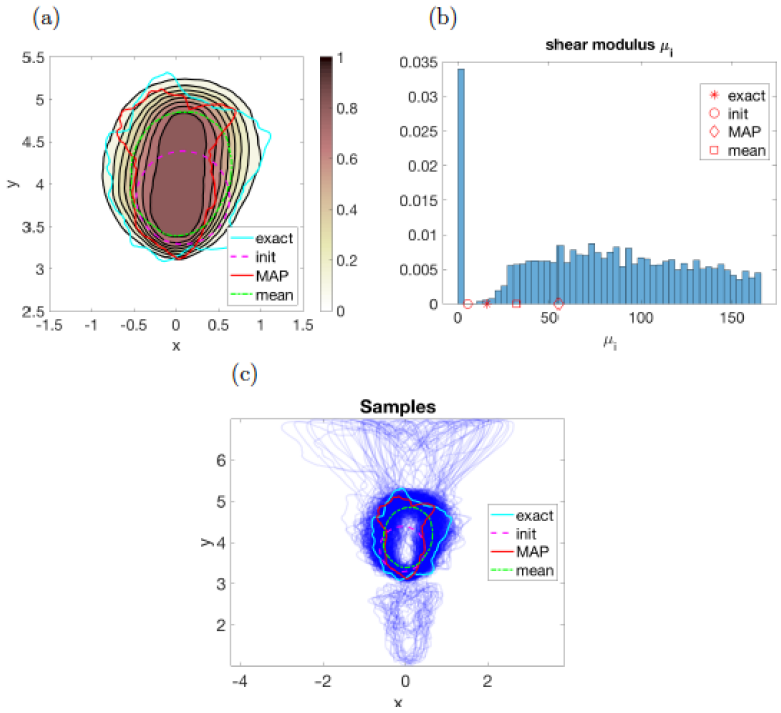

Figure 12: Results for the object in Figure 3(b) for $Z = 500$. MCMC with AIES $W = 600$, $\tilde{S} = 3900$, $B = W\tilde{S}/2$, $\lambda = 0.2$.

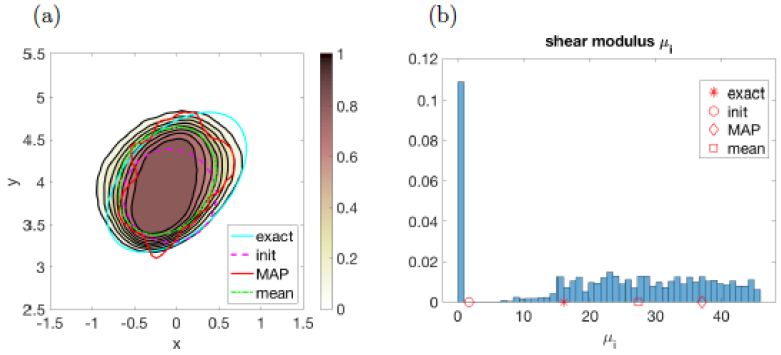

Figure 13: Results for the rotated ellipse, working with high dimensional parametrizations allowing for irregular shapes with $Z = 500$. Same sampling parameters as in Fig. 13 except $\tilde{S} = 2400$.

lower cost.

8 Conclusions

We have developed a Bayesian approach for the detection and characterization of anomalies in tissues which uses topological energies to generate priors. In this framework, anomalies are represented by star-shaped objects whose shear moduli differ from the surrounding tissues. We have considered low dimensional parameterizations for simple smooth shapes and higher dimensional approximations for irregular shapes, which can be used to distinguish encapsulated (smooth) and invasive (irregular) tumors, for instance.

For simple shapes, MCMC methods based on different types of affine invariant ensemble samplers provide a good characterization of the structure of the posterior distribution, which displays asymmetric tails for each mode representing an object. This approach is time consuming, since we must generate a few hundred thousand samples by solving a time dependent wave equation for each of them. We have shown that its is possible to approximate the 'maximum a posteriori' (MAP) estimate of the parameters defining the hardness and geometry of these anomalies. To do so, we minimize a proper cost functional, which can be done in a few iterations by Newton type iterations. Linearizing the parameter-to-observable map about the MAP point, we are able to quantify the uncertainty in nature of the anomalies, their location and shape by generating samples of the Laplace approximation to the posterior distribution at a low computational cost.

We have tested these schemes in 2D shear imaging set-ups, finding reasonable agreement between both sampling techniques for such shapes. While MCMC sampling furnishes a deeper insight in the structure of the posterior, including asymmetry and possible multimodality, the linearization approach provides results quite fast. This is essential for potential technological applications and three dimensional extensions. However, it may miss multimodality and asymmetry details. The calculation of MAP points is related to deterministic iteratively regularized Gauss Newton (IRGN) methods, in the sense that the costs considered involve Tikhonov type regularizations, which vanish in IRGN approaches but keep the prior information in the Bayesian framework. The performance of other deterministic full waveform inversion techniques in this set up, such as those employing total variation regularization, deserves further study.

Irregular shapes lead to higher dimensional problems and optimization approaches to calculate a MAP point encounter difficulties due to fast variations in the boundary. We have seen that affine invariant samplers which are robust as dimension grows still provide some information on basic anomaly properties, though we identify multimodality features due to inconsistency between the data and the prior. Better descriptions of the anomaly shape and shear modulus would probably require improved prior knowledge or a different type of parametrization. We have used here Matern covariances to characterize curves with specified regularity degrees. One could refine this Bayesian approach adding the regularity as an additional hyperparameter and using priors with two components, that is, $p(\boldsymbol{\nu}|m)p(m)$, where $\boldsymbol{\nu}$ are the parameters of the anomaly and m would be the parameters of the Matern covariance governing the regularity of its boundary.

Alternative Bayesian formulations seek variations in the wave speed of the whole tissue, which leads to infinite dimensional problems and very large computational cost. The approach based on seeking shapes described by a moderate number of parameters that we propose here has been tried for simple shapes on time independent imaging problems for which efficient boundary element solvers are available. Lacking similar solvers for time dependent wave problems, we have succeeded in developing fast finite element schemes allowing us to implement our Bayesian formulation in terms of parametrized boundaries, at a low computational cost, which is convenient for practical applications.

The assumption that both healthy and unhealthy tissue have constant shear modulus may not be realistic. We expect that our ideas could be used to deal with field data in different ways. A possibility is to split the data in three sets. The first one, with the assumption of constant shear modulus for the healthy tissue, would be used to locate anomalies by topological energy methods. With the additional assumption of constant shear modulus for the unhealthy tissue, the second one could be used to construct improved priors optimizing a regularized cost, either by MCMC techniques or by more standard optimization methods. The last fraction of data would be used to implement a Bayesian framework with spatially varying coefficients. Hybrid frameworks combining our treatment of boundaries with formulations seeking the wave speed variations in grids of points are a possibility. Another option consists in parameterizing the varying moduli by expanding them in a selected basis.

A Approximate solutions for the forward problem

We recall here the pertinent existence and regularity result for the forward problem, as well as some discretization details and parameter choices.

A.1 Existence and regularity

In the sequel, H^1, H^2 represent the standard Sobolev spaces and $(H^1)'$ is the dual space of H^1 [1, 6]. L^2 stands for the usual space of square-integrable functions.

Theorem 1. *Let R and Ω be C^1 domains, $\Omega \subset R$ [1]. Assume $f \in C^\infty(\mathbb{R}^+) \cup L^\infty(\mathbb{R}^+)$ and $g \in C^\infty(\mathbb{R}^2) \cup L^\infty(\mathbb{R}^2)$. Then, the problem (1)-(2) has a unique solution $u \in C([0,\tau]; H^1(R))$, $u_t \in C([0,\tau]; L^2(R))$, $u_{tt} \in L^2(0,\tau; (H^1(R))')$, for any $\tau > 0$. Furthermore, if $u_{tt}(\mathbf{x}, 0) \in L^2(R)$, we also have $u_t \in C([0,\tau]; H^1(R))$ and $u \in C([0,\tau]; H^2(R \setminus \overline{\Omega}))$.*

Proof. Existence of a solution u with the stated regularity for wave equations with positive and bounded μ and ρ is a particular case of results established in [36, 46]. If $u_{tt}(\mathbf{x}, 0) \in L^2(R)$, u_t solves a problem similar to (1)-(2) with f replaced by f'. Hence, $u_t \in C([0,\tau]; H^1(R))$ and $u_{tt} \in C([0,\tau]; L^2(R))$. Then equation (1) implies that $\Delta u(t) \in L^2(R \setminus \overline{\Omega})$, thus $u(t) \in H^2(R \setminus \overline{\Omega})$ by elliptic regularity theory and u is defined on Σ and the receiving sites both in the sense of $L^2(\Sigma)$ traces and pointwise [6, 10].

A.2 Physical parameters and nondimensionalization

For computational purposes, we nondimensionalize the problem using characteristic times and lengths. Let T and L be two characteristic time and length scales to be chosen. To simplify, one can take $\rho_{\mathrm{i}} = \rho \sim 1000\,\mathrm{kg/m^3}$ in tissues, though $\rho_{\mathrm{i}} > \rho$ in general (slightly). We set $\mathbf{x} = \mathbf{x}'L$, $t = t'T$, $u = u'L$, $\Omega = \Omega'L$, $R = R'L$ and $\Sigma = \Sigma'L$. Making the change of variables

[1] The result remains true with piecewise boundary regularity or when R is a convex Lipschitz domain using Sobolev space theory for them [1, 42].

and dropping the symbol $'$ for ease of notation, we get

$$\begin{aligned} &u_{tt} - \operatorname{div}(\tfrac{\mu T^2}{\rho L^2}\nabla u) = \tfrac{T^2}{\rho L} f(tT)G(\mathbf{x}L) = \tilde{f}(t)\tilde{G}(\mathbf{x}), && \mathbf{x} \in R,\ t > 0, \\ &\tfrac{\partial u}{\partial \mathbf{n}} = 0, && \mathbf{x} \in \partial R, \\ &u(\mathbf{x}, 0) = 0, u_t(\mathbf{x}, 0) = 0, && \mathbf{x} \in R. \end{aligned} \tag{30}$$

Here, $\tilde{f}(t) = f_0 \frac{T^2}{\rho L}(1 - 2\pi^2 f_M^2 T^2 t^2) e^{-\pi^2 f_M^2 T^2 t^2}$. We choose $\tilde{G}$ to have zero normal derivative at the interface, for instance, $\tilde{G}(\mathbf{x}) = \frac{1}{(\pi\kappa)^{n/2}} \sum_{j=1}^{J} \exp(-\frac{|(\mathbf{x}-\mathbf{x}_j)|^2}{\kappa})$, $n = 2$. This function represents the location of the emitters.

Typical experimental conditions [3, 31, 54] suggest the choice $L = 1$ cm $= 10^{-2}$ m and $T = 10^{-2}$ s. For instance, typical anomaly shapes and sizes in a liver framework are ellipsoids of about 0.963×1.15 cm, buried at a depth between 6 and 12 cm. To spot anomalies of size 1 cm, that is, 10^{-2} m, we should need a receiver grid of step about 10^{-3} m distributed or moving over regions of cm length. Typical parameter ranges [3, 31, 54] are $\mu_{\text{i}} = 96 - 241$ kPa (carcinoma), $\mu_{\text{i}} = 55 - 71$ kPa (normal tissue), and $\mu_{\text{i}} = 36 - 41$ kPa (bening hyperplasia) in a prostate gland, for instance. In a liver framework, $\mu_{\text{i}} = 0.4 - 6$ kPa (healthy tissue) and $\mu_{\text{i}} = 15 - 100$ (unhealthy tissue). Breast is less appropriate for these methods because carcinoma may yield $\mu_{\text{i}} = 22 - 560$ kPa, overlapping with fibrous tissue $\mu_{\text{i}} = 96 - 244$ kPa, normal fat $\mu_{\text{i}} = 18 - 24$ kPa, and normal gland $\mu_{\text{i}} = 28 - 66$ kPa, other techniques [34] may be more suitable. Frequencies f_M in shear elastography devices are $4 - 15$ Hz, or 50 Hz, or $100 - 300$ Hz, depending on sizes involved.

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	$4\frac{\text{m}}{\text{s}}$	$1.3\frac{m}{s}$	50 Hz	$\frac{\rho L}{T^2}$

Table 1: Dimensional parameters used in the simulations.

In our numerical tests we work with the parameters listed in Table 1. We set $\mu_{\text{i}} = 16$ kPa and $\mu = 1.69$ kPa, which results in wave speeds $c_{\text{i}} = 4$ m/s inside the anomalies and $c = 1.3$ m/s outside, a low contrast situation. We select f_0 such that $f_0 \frac{T^2}{\rho L} = 1$ and $f_M = 50$ Hz so that $f_M T = 0.5$. Then, the final dimensionless forward problem is

$$\begin{aligned} &u_{tt} - \operatorname{div}(c(\mathbf{x})^2 \nabla u) = \tilde{f}(t)\tilde{G}(\mathbf{x}), && \mathbf{x} \in R, \\ &\tfrac{\partial u}{\partial \mathbf{n}} = 0, && \mathbf{x} \in \partial R, \\ &u(\mathbf{x}, 0) = 0, u_t(\mathbf{x}, 0) = 0, && \mathbf{x} \in R, \end{aligned} \tag{31}$$

for $t > 0$, with

$$c^2(\mathbf{x}) = \frac{\mu(\mathbf{x})T^2}{\rho L^2} = \begin{cases} 1.69, & \mathbf{x} \in R \setminus \overline{\Omega}, \\ 16, & \mathbf{x} \in \Omega, \end{cases} \qquad c(\mathbf{x}) = \begin{cases} 1.3, & \mathbf{x} \in R \setminus \overline{\Omega}, \\ 4, & \mathbf{x} \in \Omega, \end{cases}$$

and

$$\tilde{f}(t)\tilde{G}(\mathbf{x}) = (1-2\pi^2 0.5^2 t^2)\frac{e^{-\pi^2 0.5^2 t^2}}{\pi\kappa}\sum_{j=1}^{J} e^{-\frac{|\mathbf{x}-\mathbf{x}_j|^2}{\kappa}}. \tag{32}$$

We generate synthetic data for our simulations by solving numerically (31)-(32) for different choices of anomalies Ω and adding random noise. We have used finite elements [13, 46] with spatial step $\delta x = 0.08$ and a total explicit spatial discretization with time step $\delta t = 0.00125$, see next section for details. We locate emitter/receivers at fixed grids of step 0.5 (or 0.2) and record the signal at a fixed time grid of step 0.025. The value of κ can be adjusted to the step δx, so that it affects just a few nodes around the emitter. Here, we have set $\kappa = 2$. Alternatively, one could also perform an even extension at the interface $\Sigma = \{(x, y) \,|\, y = 0\}$ to get a problem set in the whole space and resort to boundary elements for wave problems [45] representing the emitters as point sources. However, an adequate framework to implement such boundary value approach is still missing.

A.3 Discretization

To reduce the computational cost we focus on a limited tissue region and truncate the computational region in such a way that R is a rectangular region, as in Figure 1. On the artificial boundaries $\partial R \setminus \Sigma$, we will enforce non reflecting boundary conditions [20]. On Σ, we keep the zero Neumann condition. For the spatial discretization, we use P_1 finite elements on a fixed mesh of step δx in space [13, 46]. If $V = \text{span}\{\phi_1, \ldots, \phi_D\} \subset H^1$ is the resulting finite element space, we approximate u by $u^D = \sum_{i=1}^{D} a_i(t)\phi_i$. Therefore, we must find u^D such that

$$\int_R u^D_{tt}(\mathbf{x}, t)\phi_j(\mathbf{x})d\mathbf{x} + \int_R c(\mathbf{x})^2 \nabla u^D(\mathbf{x}, t)\phi_j(\mathbf{x})d\mathbf{x} - \int_{\partial R\setminus\Sigma} c^2 \frac{\partial u^D}{\partial \mathbf{n}}\phi_j(\mathbf{x})dS_{\mathbf{x}} = \tilde{f}(t)\int R\tilde{G}(\mathbf{x})\phi_j(\mathbf{x})d\mathbf{x},$$

for $j = 1, \ldots, D$. Next, we use the nonreflecting boundary condition $\frac{\partial u^D}{\partial \mathbf{n}} \sim -\frac{1}{c} u_t^D$ on $\partial R \setminus \Sigma$ and total discretization for the time derivatives on a time mesh t_n of step δt [13, 46]

$$u_{tt}^D(\mathbf{x}, t_n) \sim \frac{u^D(\mathbf{x}, t_{n+1}) - 2u^D(\mathbf{x}, t_n) + u^D(\mathbf{x}, t_{n-1})}{\delta t^2},$$
$$u_t^D(\mathbf{x}, t_n) \sim \frac{u^D(\mathbf{x}, t_n) - u^D(\mathbf{x}, t_{n-1})}{\delta t}.$$

To calculate the coefficients $a_i(t_n)$, $i = 1, \ldots D$, we solve the recurrence relations

$$\sum_{i=1}^{D} M_{j,i} a_i(t_{n+1}) = \sum_{i=1}^{D} M_{j,i}(2a_i(t_n) - a_i(t_{n-1})) - \delta t^2 \sum_{i=1}^{D} A_{j,i} a_i(t_n)$$
$$-c\,\delta t \sum_{i=1}^{D} B_{j,i}(a_i(t_n) - a_i(t_{n-1})) + \delta t^2 \tilde{f}(t_n) G_j,$$

for $n \geq 1$, where $M_{j,i} = \int_R \phi_j \phi_i d\mathbf{x}$, $A_{j,i} = \int_R c^2 \nabla \phi_j \nabla \phi_i d\mathbf{x}$, $B_{j,i} = \int_{\partial R \setminus \Sigma} \phi_j \phi_i dS_{\mathbf{x}}$, $G_j = \int_R \tilde{G} \phi_j d\mathbf{x}$. The coefficients $a_i(t_0) = 0$ and $a_i(t_1) = 0$ for $i = 1, \ldots, D$ are determined using the initial conditions. The numerical solutions defined in this way are continuous, so that the costs (4), (23), likelihoods (12), and topological energies (18) are well defined. This is a low order approximation, since the mesh is kept fixed and it is not adapted to the shape of the varying anomalies. In this way, we achieve a reduced computational cost which allows us to solve large amounts of forward problems quite fast for MCMC sampling.

Acknowledgements. This research has been partially supported by the FEDER /Ministerio de Ciencia, Innovación y Universidades - Agencia Estatal de Investigación grants No. MTM2017-84446-C2-1-R and PID2020-112796RB-C21. AC thanks G. Stadler for nice discussions and useful suggestions.

References

[1] Adams R A 1975 *Sobolev Spaces* (Academic Press, New York)

[2] Afkham B M, Dong Y and Hansen P C 2021 Uncertainty quantification of inclusion boundaries in the context of X-ray tomography, arXiv:2107.06607v1.

[3] Shear Wave Elastography, in Blumgart's Surgery of the Liver, Biliary Tract and Pancreas, 2-Volume Set (Sixth Edition), 2017.

[4] Bui-Thanh T and Ghattas O 2014 An analysis of infinite dimensional Bayesian inverse shape acoustic scattering and its numerical approximation *SIAM/ASA J. Uncertain. Quantification* **2** 203-22

[5] Bui-Thanh T, Ghattas O, Martin J and Stadler G 2013 A computational framework for infinite-dimensional Bayesian inverse problems Part I: The linearized case with application to global seismic inversion *SIAM J. Sci. Comput.* **35** A2494-A2523

[6] Brézis H 1987 *Analyse fonctionnelle Théorie et applications* (Paris: Masson)

[7] Carpio A and Rapún ML 2012 Hybrid topological derivative and gradient-based methods for electrical impedance tomography *Inverse Problems* **28** 095010

[8] Carpio A, Dimiduk TG, Le Louër F and Rapún ML 2019 When topological derivatives met regularized Gauss-Newton iterations in holographic 3D imaging *J. Comp. Phys.* **388** 224-251

[9] Carpio A, Iakunin S and Stadler G 2020 Bayesian approach to object detection with topological priors *Inverse Problems* **36** 105001

[10] Cazenave T and Haraux A 1999 *An introduction to semilinear evolution equations* Oxford Lecture Series in Mathematics and Its Applications 13 (Oxford: Clarendon Press)

[11] Chung E T, Chan T F and Tai X C 2005 Electrical impedance tomography using level set representation and total variational regularization, *Journal of Computational Physics* **205** 357-372

[12] Colton D and Kress R 1998 *Inverse Acoustic and Electromagnetic Scattering* (Berlin: Springer)

[13] Dautray R and Lions J L 1984-87 *Analyse mathématique et calcul numérique pour les sciences et les techniques* (Paris: Masson)

[14] Du Z, Liu D, Wu G, Cai J, Yu X and Hu G 2021 A high-order total-variation regularisation method for full-waveform inversion *Journal of Geophysics and Engineering* **18** 241-252

[15] Dominguez N, Gibiat V and Esquerre Y 2005 Time domain topological gradient and time reversal analogy: an inverse method for ultrasonic target detection *Wave Motion* **42** 31-52

[16] Dominguez N and Gibiat V 2010 Non-destructive imaging using the time domain topological energy method *Ultrasonics* **50** 367-372

[17] Dunlop M M and Yang Y 2021 Stability of Gibbs posteriors from the Wasserstein loss for Bayesian full waveform inversion *SIAM/ASA Journal on Uncertainty Quantification* **9** 1499-1526

[18] Dunlop M M 2016 Analysis and computation for Bayesian inverse problems *PhD Thesis* Warwick

[19] Dunlop M M and Stadler G 2022 A gradient-free subspace-adjusting ensemble sampler for infnite-dimensional Bayesian inverse problems, arXiv:2202.11088v1

[20] Engquist B and Majda A 1979 Radiation boundary conditions for acoustic and elastic wave calculations *Commun. Pur. Appl. Math.* **32** 312-358

[21] Engquist B, Froese B D and Yang Y 2016 Optimal transport for seismic full waveform inversion *Comm. Math. Sci.* **14** 2309-2330

[22] Esser E, Guasch L, van Leeuwen T, Aravkin A Y and Herrmann F J 2018 Total-variation regularization strategies in full-waveform inversion *SIAM Journal on Imaging Sciences* **11** 376-406

[23] Fletcher R 1971 Modified Marquardt subroutine for non-linear least squares *Tech. Rep.* 197213

[24] Fichtner A, Bunge H P and Igel H 2006 The adjoint method in seismology - II Applications: traveltimes and sensitivity functionals *Physics of the Earth and Planetary interiors* **157** 105-123

[25] Harbrech H and Hohage T 2007 Fast methods for three-dimensional inverse obstacle scattering problems *J. Integral Equ. Appl.* **19** 237-260

[26] Gebraad L, Boehm C and Fichtner A 2020 Bayesian elastic full-waveform inversion using Hamiltonian Monte Carlo *JGR Solid Earth* **125** e2019JB018428

[27] Goodman J and Weare J 2010 Ensemble samplers with affine invariance, *Commun. Appl. Math. Comput. Sci.* **5** 65-80

[28] Guo Z and De Hoop M V, 2012 Shape optimization in full waveform inversion with sparse blocky model representations *Proceedings of the Project Review, Geo-Mathematical Imaging Group* (Purdue University, West Lafayette IN) **1** 189-208

[29] Guzina B and Chikichev I 2007 From imaging to material identification: a generalized concept of topological sensitiviy *J. Mech. Phys. Solids* **55** 245-279

[30] Hohage T and Schormann C 1998 A Newton-type method for a transmission problem in inverse scattering *Inverse Problems* **1**, 1207-27

[31] Hoyt K, Castaneda B, Zhang M, Nigwekar P, di Sant'Agnese P A, Joseph J V, Strang J, Rubens D J and Parker K J 2008 Tissue elasticity properties as biomarkers for prostate cancer *Cancer Biomark.* **4** 213-225

[32] Kaipio J and Somersalo E 2006 *Statistical and computational inverse problems* Vol 160 (Berlin: Springer)

[33] Käuf P, Fichtner A and Igel H 2010 Object-based probabilistic full-waveform tomography *Master thesis Geophysics* LMU Munchen

[34] Korta Martiartu N, Boehm C, Vinard N, Jovanovic Balic I and Fichtner A 2017 Optimal experimental design to position transducers in ultrasound breast imaging *Medical Imaging* Ultrasonic Imaging and Tomography 10139, 101390M, International Society for Optics and Photonics

[35] Landau L D and Lifshitz L M 1986 *Theory of elasticity* (Oxford: Butterworth-Heinemann 3rd Edition)

[36] Lions J L and Magenes E 1968 *Problémes aux limites non homogénes* (Paris: Dunod)

[37] *Medical Imaging Systems: An Introductory Guide* 2018 Maier A, Steidl S, Christlein V, J Hornegger (Eds.) Springer

[38] Malcolm A and Guzina B 2008 On the topological sensitivity of transient acoustic fields *Wave Motion* **45** 821-834

[39] Métivier L, Brossier R, Mérigot Q, Oudet E and Virieux J 2016 Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion *Geophys. J. Int.* **205** 345-377

[40] Mishra A 2017 Hologram the future of medicine - From Star Wars to clinical imaging *Indian Heart J.* **69** 566-567

[41] Neal R M 2011 MCMC using Hamiltonian dynamics. In: Brooks S, Gelman A, Jones GL, Meng XL, editors. *Handbook of Markov chain Monte Carlo* (London: Chapman & Hall)

[42] Nečas J 1983 Introduction to the theory of nonlinear elliptic equations (Leipzig: Teubner)

[43] Palafox A, Capistrán M A and Christen J A 2017 Point cloud-based scatterer approximation and affine invariant sampling in the inverse scattering problem *Math. Methods Appl. Sci.* **40** 3393-403

[44] Petra N, Martin J, Stadler G and Ghattas O 2014 A computational framework for infinite-dimensional Bayesian inverse problems: Part II. Stochastic Newton MCMC with application to ice sheet flow inverse problems *SIAM J. Sci. Comput.* **36** A1525-55

[45] Hsiao G C, Sánchez-Vizuet T and Sayas F J 2017 Boundary and coupled boundary-finite element methods for transient wave-structure interaction *IMA Journal of Numerical Analysis* **37** 237-265

[46] Raviart P A and Thomas J M 1983 *Introduction a l'analyse numérique des équations aux dérivées partielles* (Paris: Masson)

[47] Rasmussen C E and Williams C K I 2006 *Gaussian Processes for Machine Learning* (MIT Press)

[48] Sahuguet P, Chouippe A and Gibiat V 2010, Biological tissues imaging with Time Domain Topological Energy, *Physics Procedia* **3** 677-683

[49] Sandrin L, Tanter M, Catheline S and Fink M 2002 Shear modulus imaging with 2-D transient elastography, *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* **49** 426-435

[50] Sarvazyan A P, Urban M W and Greenleaf J F 2013 Acoustic waves in medical imaging and diagnostics *Ultrasound Med. Biol.* **39** 1133-46

[51] Sarvazyan A, Hall TJ, Urban MW, Fatemi M, Aglyamov SR and Garra BS 2011 Overview of elastography - an emerging branch of medical imaging *Current Medical Imaging Reviews* **7** 255-282

[52] Tarantola A 2005 Inverse Problem Theory and Methods for Model Parameter Estimation (Philadelphia PA: SIAM)

[53] Tsogka C and Papanicolaou G C 2002 Time reversal through a solid-liquid interface and super-resolution *Inverse Problems* **18** 1639-1657

[54] Wang Y, Yao B, Li H, Zhang Y, Gao H, Gao Y, Peng R and Tang J 2017 Assessment of tumor stiffness with shear wave elastography in a human prostate cancer xenograft implantation model *J Ultrasound Med* **36** 955-963

[55] Zhu H, Li S, Fomel S, Stadler G and Ghattas O 2016 A Bayesian approach to estimate uncertainty for full-waveform inversion using a priori information from depth migration *Geophysics* **81** R307-R323