Bayesian normal modes identification and estimation of elastic coefficients in resonant ultrasound spectroscopy

Simon Bernard, Guillaume Marrelec, Pascal Laugier, Quentin Grimal

Bayesian normal modes identification and estimation of elastic coefficients in resonant ultrasound spectroscopy

Simon Bernard, Guillaume Marrelec, Pascal Laugier, Quentin Grimal

Sorbonne Universités, UPMC Univ Paris 06, INSERM, CNRS, Laboratoire d'Imagerie Biomédicale, Paris, France

E-mail: s.bernard.simon@gmail.com

Abstract.
Resonant ultrasound spectroscopy is an experimental technique for measuring the stiffness of anisotropic solid materials. The free vibration resonant frequencies of a specimen are measured and the stiffness coefficients of the material adjusted to minimize the difference between experimental and predicted frequencies. An issue of this inverse approach is that the measured frequencies are not easily paired with their predicted counterpart, leading to ambiguities in the definition of the objective function. In the past, this issue has been overcome through trial-and-error methods requiring the experimentalist to find the correct pairing, or through involved experimental methods measuring the shapes of the normal vibration modes in addition to their frequencies. The purpose of this work is to show, through a Bayesian formulation, that the inverse problem can be solved automatically and without requiring additions to the usual experimental setup. The pairing of measured and predicted frequencies is considered unknown, and the joint posterior probability distribution of pairing and stiffness is sampled using Markov Chain Monte Carlo. The method is illustrated on two published data sets. The first set includes the exact pairing, allowing validation of the method. The second application deals with attenuative materials, for which many predicted modes cannot be observed, further complicating the inverse problem. In that case, introduction of prior information through Bayesian formulation reduces ambiguities.

1. Introduction

Resonant ultrasound spectroscopy (RUS) is an experimental technique for measuring the stiffness of anisotropic solid materials [1, 2, 3, 4]. In RUS, the free vibration resonant frequencies of a material specimen of simple shape (e.g. cylinder, sphere, parallelepiped) are measured and the stiffness coefficients adjusted until model-predicted frequencies match the experimental results. This inverse approach gives accurate elastic parameters because the resonant frequencies, which are entirely determined by elasticity, specimen dimensions and material mass density, can be measured with high accuracy.

An issue of that inverse approach is that each measured resonant frequency taken individually does not carry information about which of the many predicted normal vibration modes it corresponds to [5, 6, 3, 7]. As the inversion relies on the minimization of the differences between pairs of predicted and measured frequencies, ambiguities in the identification of the frequencies lead to ambiguities in the stiffness estimation. Identification, or pairing, is particularly difficult if the initial guess of the stiffness coefficients is far from the actual values and if some predicted resonant modes are not observed experimentally. It is usually argued that without knowledge of the pairing, RUS cannot lead to accurate estimates of the stiffness coefficients. One purpose of this work is to show, through a Bayesian formulation of the inverse problem, that knowledge of the exact pairing is not necessary to get precise and accurate stiffness estimates.

Several experimental solutions have been proposed to find the frequency pairing, the most popular being that of Ogi *et al.* [7]. They proposed to image the vibration patterns of the resonant modes on a face of the specimen using a laser interferometer. For homogeneous materials and moderate stiffness anisotropy ratios (< 10), these patterns are determined mainly by the specimen shape. Hence, the observed resonant modes can be identified and the stiffness coefficients evaluated from fitting of the resonant frequencies, even in the case of a bad initial guess of their values. This method has been used by several groups [8, 9, 10, 11, 12, 13] but suffers from a number of limitations. (1) It requires specimens with a plane face large enough to be correctly sampled by the laser dot. In consequence, it is not adapted to very small (sub-millimeter) specimens nor to complex shapes. (2) The scanned face must be perfectly plane and highly reflective, which is not always the case, particularly for biological materials such as dentin [14] or bone [15, 16]. (3) The use of a Laser interferometer considerably increases the cost and difficulty of implementation of RUS, preventing its use as a routine method for material characterization in some contexts, e.g. industrial or biomedical. Although difficulties (1) and (2) can be tackled respectively by using optical microscopy and by sputtering an aluminum thin film on the surface of the sample, this further complicates the application of RUS. The growing interest in applications of RUS where the above limitations are problematic motivates the development of an alternative method.

Without measuring the modal shapes, the correct identification can be obtained through trial-and-errors methods [2], or progressive approaches identifying modes one after another [11, 16, 17]. Although each individual frequency does not bring information

on its corresponding vibration mode, the distribution of the frequencies does. Indeed, the correct pairing is expected to lead to the best match between the experiment and the model. These methods however are time-consuming and require the user to manually identify the frequencies.

In this work we propose a Bayesian formulation of the inverse problem that introduces probability distributions for the possible pairings of predicted and measured frequencies and for the stiffness coefficients. Estimating the joint posterior distribution given a set of measured frequencies then solve the combined problem of finding the correct pairing and estimating stiffness. This is done through Markov Chain Monte Carlo (MCMC) sampling of the posterior distribution. It requires no input of the user, except for the specification of *a priori* probability densities. We show on two examples that the method leads to results that are almost as accurate and precise as with the exact known pairing.

2. Method

2.1. Forward problem

The Rayleigh-Ritz method is efficient to predict the resonant frequencies of a solid of simple shape given its stiffness, geometry and mass density. This method has been extensively used in RUS [5, 1, 2, 3, 4]. The general principles of the method are briefly presented here and some details regarding its implementation for rectangular parallelepiped specimens are given, with an emphasis on mathematical properties allowing an efficient computation.

The free-vibration natural frequencies of a solid body are the stationary points of the Lagrangian $\mathcal{L}$

$$\mathcal{L} = \frac{1}{2} \int_V \left[\rho \omega^2 u_i^2 - C_{ijkl} \epsilon_{ij} \epsilon_{kl} \right] dV, \tag{1}$$

where C_{ijkl} are the stiffness coefficients, V and ρ are respectively the volume and mass density, u_i is a component of the displacement field, and ϵ_{ij} is a component of the strain tensor. The Rayleigh-Ritz method expands the displacement field as a finite sum of known basis functions ϕ_λ with yet unknown coefficients $\alpha_{i,\lambda}$, i.e.

$$u_i = \sum_\lambda \alpha_{i,\lambda} \phi_\lambda. \tag{2}$$

in order to solve the stationary equation $\partial \mathcal{L} = 0$ as a generalized eigenvalue problem

$$\omega^2 \mathbf{M} \boldsymbol{\alpha} = \mathbf{K} \boldsymbol{\alpha}. \tag{3}$$

Sufficient details on the manipulations necessary to obtain Eq. (3) from the Lagrangian (1) and the expansion (2) can be found in [1, 2, 3].

In Eq. (3), $\mathbf{M}$ and $\mathbf{K}$ are called respectively the mass and stiffness matrices of the vibration problem. Their elements are

$$M_{i\lambda,k\lambda'} = \rho \delta_{ik} \int_V \phi_\lambda \phi_{\lambda'} dV, \text{ and} \tag{4}$$

$$K_{i\lambda,k\lambda'} = \sum_{j,l=1}^{3} C_{ijkl} \int_V \epsilon_{ij}(\phi_\lambda)\epsilon_{kl}(\phi'_\lambda)dV, \tag{5}$$

where δ_{ij} is the Kronecker symbol. For simple shapes and well-chosen basis functions, Eqs. (4) and (5) reduce to simple analytical expressions and Eq. (3) is numerically solved, giving the resonant pulsations ω_n and the displacement patterns of the resonant modes from the eigenvectors $\boldsymbol{\alpha_n}$. The forward relation between the stiffness coefficients and the vector containing the resonant frequencies $f_n = \omega_n/2\pi$, denoted $\mathbf{f} = \mathbf{g}(C_{ijkl})$, is non-linear and has no analytical expression.

For a rectangular parallelepiped, it is convenient to use the normalized Legendre polynomials of the scaled Cartesian coordinates x_i, i.e.

$$u_i = \sum_{a,b,c=0}^{a+b+c\leq\Lambda} \alpha_{i,abc} P_a\left(\frac{2x_1}{L_1}\right) P_b\left(\frac{2x_2}{L_2}\right) P_c\left(\frac{2x_3}{L_3}\right) \tag{6}$$

where P_a is the normalized Legendre polynomial of order a, and L_i is the length of the parallelepiped in the direction i. Due to the orthogonality of the Legendre polynomials over the interval $[-1, 1]$, $\mathbf{M}$ is the unit matrix and Eq. (3) simplifies to a standard eigenvalue problem, easier to solve numerically. Moreover, many elements of $\mathbf{K}$ are equal to zero and the matrix turns to be block-diagonal, splitting Eq. (3) in several independent smaller problems [5, 2, 3]. For an orthotropic elastic symmetry, involving three orthogonal symmetry planes, the problem splits in eight smaller ones, each corresponding to a different combination of symmetry or antisymmetry in the three directions of space. This split drastically reduces the computation cost and allows labeling of the resonant modes according to their belonging to one of the eight symmetry groups [18]. Additionally, we note that the matrix $\mathbf{K}$ has a linear dependence to the stiffness constants, as it can be seen from Eq. (5). Hence, in an iterative computation of the resonant frequencies for different sets of elastic properties and a given geometry, elements of $\mathbf{K}$ are linearly obtained instead of being computed from scratch.

A RUS measurement usually focuses on a limited frequency band containing a few tenth of resonant frequencies. Hence, it is not necessary to consider all the solutions of the approximation (3) in the inverse problem, but only a number K slightly superior to the number N of frequencies measured in the investigated band ($K > N$ to account for the possibility of non-observed frequencies). For polynomials up to an order $\Lambda = 10$, the accuracy on the $K = 40 - 50$ first resonant frequencies is sufficient [3] (i.e. errors due to the truncation order are much smaller than other sources of error, such as imperfect geometry and measurement uncertainty) and the computing time is of the order of a few hundredths of a second.

2.2. Bayesian formulation of the inverse problem

The Bayesian formulation of the inverse problem includes the estimation of the stiffness coefficients from measured resonant frequencies, as well as the estimation of the correct

pairing of the frequencies. We only briefly review the required concepts of Bayesian inference. Extensive descriptions can be found elsewhere [19, 20].

The purpose of Bayesian inference is to evaluate the posterior probability density function (pdf) $p(\mathbf{m}|\mathbf{f})$ of a d-dimensional vector of parameters $\mathbf{m}$ of a probabilistic model, given some measured data $\mathbf{f}$ and some prior information on the parameters $p(\mathbf{m})$. In our problem, m is a vector of elastic parameters defined from the stiffness coefficients C_{ijkl} (see section 2.3.2 and Appendix A) and f is a vector containing N measured resonant frequencies f_n. Bayes' theorem expresses the posterior pdf as a normalized product of the prior density and the likelihood $p(\mathbf{f}|\mathbf{m})$ (the probability of having obtained the data given the parameters)

$$p(\mathbf{m}|\mathbf{f}) = \frac{p(\mathbf{m})p(\mathbf{f}|\mathbf{m})}{p(\mathbf{f})}. \tag{7}$$

The normalizing term is the marginal likelihood of the data and does not depend on the parameters. The solution of the inference problem can then be obtained from the simpler equation

$$p(\mathbf{m}|\mathbf{f}) \propto p(\mathbf{m})p(\mathbf{f}|\mathbf{m}). \tag{8}$$

The likelihood $p(\mathbf{f}|\mathbf{m})$ describes both the physics of the problem – the assumed relation between the parameters and the measured data $\mathbf{f} = \mathbf{g}(\mathbf{m})$ – and the statistics of measurement and modeling errors. It is usually not possible to separate the sources of errors and their distribution is therefore difficult to specify. A common and convenient practice is to assume a multivariate normal distribution of the errors [19]. We further assume that the errors are uncorrelated and that the variance of each particular frequency is proportional to the squared frequency, with a unique proportionality factor σ^2. This assumption of a constant relative error on the frequencies is popular in the usual formulation of the inverse problem in RUS [3, 16], although the hypothesis of a constant absolute error was also used by some authors [11]. In the present framework, choosing this last hypothesis would lead to a slightly different expression of the likelihood (9), but would not implies changes in the sampling method that is described in the following sections.

As previously stated, we do not know which of the overnumerous predicted frequencies in the investigated band correspond to the measured frequencies. A correspondence between the predicted and measured frequencies is called a *pairing* and is defined as follows: 1) N predicted frequencies among K are selected, and are considered as present in the measurement and 2) each of these selected frequencies is uniquely identified to one of the N measured frequencies. Formally, the possible pairings are represented by a vector $\mathbf{a}$ of K integers. The n-th component of $\mathbf{a}$ is equal to 0 if the corresponding predicted frequency is not paired to a measured frequency (frequency not observed during the experiment) and equal to k if the corresponding frequency is paired to the k-th measured frequency (see Fig. 1). With this notation, the likelihood is

$$p(\mathbf{f}|\mathbf{m}, \sigma^2, \mathbf{a}) \propto (\sigma^2)^{-N/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{n=1}^{N} \left(\frac{g_n(\mathbf{m}, \mathbf{a}) - f_n}{f_n}\right)^2\right], \tag{9}$$
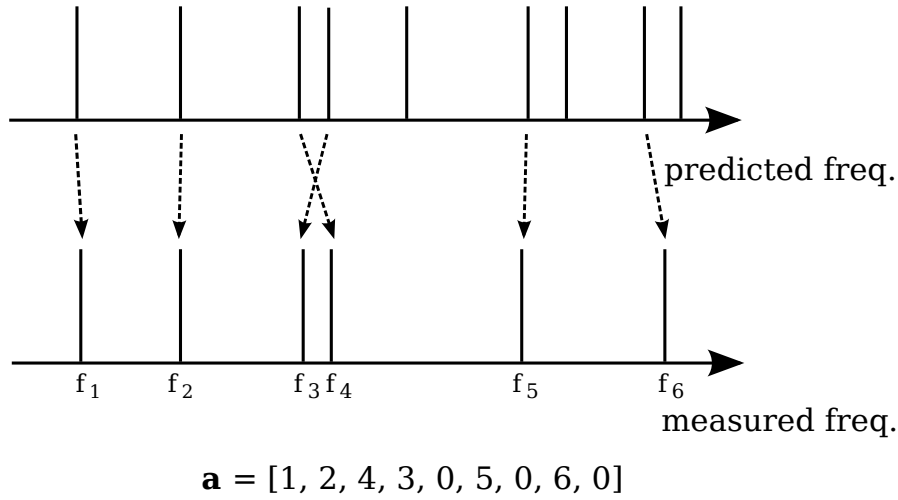
Figure 1. Example of a possible pairing for a simple fictitious case with $N = 6$ measured and $K = 9$ predicted frequencies. The illustrated pairing **a** shows missing of some predicted mode during measurement and different orders of appearance of the vibration modes in the experiment and the model.

where g_n is the predicted frequency paired to the nth measured frequency f_n for the pairing **a**.

If we want to estimate the pairing from the data and if we ignore the magnitude of the errors, the inference problem is threefold. We want to obtain the joint posterior pdf of **m**, **a**, and σ^2,

$$\mathrm{p}(\mathbf{m}, \sigma^2, \mathbf{a}|\mathbf{f}) \propto \mathrm{p}(\mathbf{m})\mathrm{p}(\sigma^2)\mathrm{p}(\mathbf{a})\mathrm{p}(\mathbf{f}|\mathbf{m}, \sigma^2, \mathbf{a}), \tag{10}$$

where we also introduce prior information on the pairing and the error term.

As often in Bayesian inference, the posterior pdf is too complex for being studied analytically. A convenient strategy is then to sample the pdf using Monte Carlo simulations. Then, from a large sample of the pdf, any quantity of interest (e.g. mean and covariances) can be calculated. In the present work we use Gibbs sampling.

Ideally, the prior distributions should reflect our *exact* state of knowledge on the parameters before the data has been collected. However, because that knowledge can be difficult to translate perfectly into probability distributions and because some families of distribution are more mathematically convenient for the sampling of the posterior distribution, the prior distributions are constructed as a *compromise* between convenience and introduced information. We describe the prior distributions $\mathrm{p}(\mathbf{m})$, $\mathrm{p}(\mathbf{a})$, and $\mathrm{p}(\sigma^2)$ in the next sections along the description of the sampling strategy.

2.3. Gibbs sampling

MCMC methods provide a way to sample complicated and multidimensional probability distributions such as the pdf (10). The idea of MCMC methods is to generate dependent samples during a random exploration of the posterior pdf. At a given state, a move

is proposed based only on the current values of the variables (Markov property), and accepted or rejected based on probabilistic rules. Different set of rules exist to generate and accept moves, such as the Metropolis and Metropolis-Hasting algorithms [21, 22, 23], and Gibbs sampling [24, 25, 23]. All ensure that the resulting Markov chain asymptotically samples the target distribution.

In Gibbs sampling only one variable or sub-group of variables is moved at a time. The new state of the updated variable is generated from the conditional distribution of this variable given the data and all other variables at their current values. Our algorithm is initialized at $(\mathbf{m}_0, \sigma_0^2)$ and then samples alternatively from each conditional distribution:

$$\text{step 1: } \mathbf{a}_1 \sim \mathrm{p}(\mathbf{a}|\mathbf{f}, \mathbf{m}_0, \sigma_0^2),$$
$$\sigma_1^2 \sim \mathrm{p}(\sigma^2|\mathbf{f}, \mathbf{m}_0, \mathbf{a}_1),$$
$$\mathbf{m}_1 \sim \mathrm{p}(\mathbf{m}|\mathbf{f}, \sigma_1^2, \mathbf{a}_1), \tag{11}$$
$$\text{step 2: } \mathbf{a}_2 \sim \mathrm{p}(\mathbf{a}|\mathbf{f}, \mathbf{m}_1, \sigma_1^2),$$
$$\dots$$

After an appropriate number of iterations, the $(\mathbf{a}_n, \mathbf{m}_n, \sigma_n^2)$ are samples of the joint pdf (10). One advantage is that there is no need to choose and fine tune a rule to generate new states, on which the efficiency of many MCMC algorithms is dependent [23]. Particularly, we want to sample jointly continuous $(\mathbf{m}, \sigma^2)$ and discrete $(\mathbf{a})$ variables, and it might be difficult to find efficient jumping rules. However, there are convenient ways to sample from the conditional distributions, which make Gibbs sampling a sensible choice. The three following sections expose how we sample from the conditional distributions $\mathrm{p}(\mathbf{a}|\mathbf{m}, \sigma^2, \mathbf{f})$, $\mathrm{p}(\mathbf{m}|\sigma^2, \mathbf{a}, \mathbf{f})$, and $\mathrm{p}(\sigma^2|\mathbf{m}, \mathbf{a}, \mathbf{f})$

2.3.1. Sampling the conditional distribution of $\mathbf{a}$

The conditional distribution of $\mathbf{a}$ is proportional to the product of the likelihood (with known $\mathbf{m}$ and σ^2) and the prior on $\mathbf{a}$:

$$\mathrm{p}(\mathbf{a}|\mathbf{m}, \sigma^2, \mathbf{f}) \propto \mathrm{p}(\mathbf{a})\mathrm{p}(\mathbf{f}|\mathbf{m}, \sigma^2, \mathbf{a}). \tag{12}$$

Due to the complicated dependence of the likelihood on $\mathbf{a}$, there is no simple way to sample from this distribution directly. A solution is to sample it using a MCMC method embedded inside Gibbs sampling (i.e. called at each iteration of the Gibbs sampler). Because $\mathbf{a}$ is a discrete vector, finding the best $\mathbf{a}$ or exploring its distribution is a combinatorial problem which is in some aspects similar to classical combinatorial problems such as the traveling salesman problem [26] (TSP; find the shortest route that visits each city in a given list exactly once and returns to the origin city). An efficient way to solve this kind of problem is simulated annealing [27, 26], briefly explained here.

From an initial state (a route in the TSP, a pairing in our problem), a new state is generated and randomly accepted or rejected based on the ratio of the probabilities of the proposed and current states and on an additional temperature parameter T. For high temperature, almost all states can be reached, even those of low probability, while

for low temperature only the moves leading to states of higher probability are accepted. Temperature is then slowly reduced from high to low values, mimicking the cooling of materials and ensuring convergence to the global optimum of the problem. Here, we are not interested only in the best solution but in the distribution of the solutions. We will therefore stop the temperature decrease at $T = 1$, the value for which the chain samples the target distribution, much before "freezing" of the chain in the most probable state.

From a given state $\mathbf{a}_j$, a new state $\mathbf{a}_{j+1}$ is generated by inverting two randomly selected successive components of $\mathbf{a}_j$. For that purpose a random integer n is uniformly generated in $[1, N]$ and the index i of the component n in $\mathbf{a}$ is determined. Then this component is moved to the right or left with probability $1/2$. For a move to the right we have

$$\begin{cases} \mathbf{a}_{j+1}[i + 1] = \mathbf{a}_j[i] = n, \\ \quad \mathbf{a}_{j+1}[i] = \mathbf{a}_j[i + 1], \\ \quad \mathbf{a}_{j+1}[l] = \mathbf{a}_j[l] \; ; \; \forall \, l \neq (i, \; i + 1). \end{cases} \tag{13}$$

and similarly for a move to the left. If $i = 1$ or $i = K$, the moves to the left and right respectively are not possible; the only possible move is then selected with probability 1.

The chain moves to the new state with probability

$$P = \min \left\{ 1, \frac{\mathrm{p}(\mathbf{a}_{j+1} \to \mathbf{a}_j)\mathrm{p}(\mathbf{a}_{j+1})\mathrm{p}(\mathbf{f}|\mathbf{m}, \sigma^2, \mathbf{a}_{j+1})^{1/T}}{\mathrm{p}(\mathbf{a}_j \to \mathbf{a}_{j+1})\mathrm{p}(\mathbf{a}_j)\mathrm{p}(\mathbf{f}|\mathbf{m}, \sigma^2, \mathbf{a}_j)^{1/T}} \right\}. \tag{14}$$

The ratio of the proposal probabilities $\mathrm{p}(\bullet \to \bullet)$ is equal to 1 except for moves involving a component at index 1 or K, in which cases it can be $1/2$ or 2. In all other cases, only the ratio of the prior probabilities $\mathrm{p}(\mathbf{a}_{j+1})/\mathrm{p}(\mathbf{a}_j)$ and the ratio of the likelihoods need to be calculated. We start the chain at $T = 1000$ and then decrease T using a scheme similar to that of Kirkpatrick et *al.* [27]: after $10 \times N$ accepted moves or $100 \times N$ attempted moves (first reached) T is multiplied by 0.9. This is iterated until $T = 1$, the temperature for which the chain samples the conditional distribution of $\mathbf{a}$. The same stopping criterion is used at $T = 1$ and the last sampled pairing is then selected as the sample from the conditional posterior of $\mathbf{a}$ (12).

Since it is done for fixed elasticity, running the simulated annealing algorithm requires no evaluation of the forward problem, but only computation of the squared difference in Eq. (9), a much less costly operation. Therefore, running a simulated annealing chain long of $\sim 10^4$ samples at each step of the Gibbs sampling only accounts for a small portion of the total computation time.

The formal representation of the pairings considers a fixed number K of predicted frequencies, that is superior to the number of measured frequencies N to account for the possibility of non-observed frequencies. However, the number of frequencies that have been actually missed during the measurement is not always $K - N =$ the total number of zeros in $\mathbf{a}$. The number of missed frequencies is equal to the number of zeros before the last non-null component of $\mathbf{a}$. Indeed, the zeros after this component represents resonant modes whose frequency is higher than the highest measured frequency, so that they are beyond the investigated frequency band. For example, on the simple case

depicted on Fig. 1 there is three zeros in the vector $\mathbf{a}$, but the last one does not count as a missed frequency because it corresponds to a frequency predicted above the last measured frequency. Another approach could have been to consider a fixed frequency band instead of a fixed number of frequencies. However, this would have made the length of the paring vector variable, which would not have been convenient for the implementation.

Our prior distribution on $\mathbf{a}$ depends on the (positive) number k of missed resonant frequencies, i.e. on the number of zeros in $\mathbf{a}$ before its last non-null component. This prior is a truncated Poisson's distribution

$$\begin{cases} \mathrm{p}(k) \propto \dfrac{\lambda^k}{k!} \exp(-\lambda) \text{ if } k \leq k_{max}, \\ \mathrm{p}(k) = 0 \text{ if } k > k_{max}, \end{cases} \tag{15}$$

where $k_{max} = K - N$ and λ is the expected number of missed frequencies. The Poisson's distribution describes the number of occurrences of a random event in a fixed interval of time if this event occurs at random, independently of the time since the last occurrence. In our case, this means that we consider the probability of missing the frequency to be equal for all modes and independent of whether the preceding has been missed or not. For a given k, all the possible pairings are considered equiprobable *a priori*, at the exception of pairings that identify the first predicted frequency as missed, which are not considered, i.e $\mathrm{p}(\mathbf{a}) = 0$ if $\mathbf{a}(1) = 0$. Indeed, in our experience with RUS measurement, the first resonant mode is usually easy to observe, hence we excluded the possibility of having missed it in our treatment of the inverse problem.

The Poisson's distribution (15) can be rather sharply peaked around λ, which may not be adequate if only vague information is known about k. In that case, a gamma prior probability can be attributed to λ instead of a determined value

$$\mathrm{p}(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\beta\lambda). \tag{16}$$

In (16), α and β are positive parameters called the shape and rate parameter, respectively. They determine the mean of the distribution (α/β) and its variance $(\alpha/\beta^2,$ for $\beta > 0)$. Using $\alpha = 1/2$ and $\beta = 0$ provide a non-informative prior distribution [28]. In the case where λ is not fixed, an additional Gibbs step must be done to update λ according its conditional distribution. Given the pairing $\mathbf{a}$ (which sets the value of k), it writes

$$\mathrm{p}(\lambda|\mathbf{a}) \propto \mathrm{p}(\lambda) \frac{\lambda^k/k!}{\sum_{l=0}^{k_{max}} \lambda^l/l!}. \tag{17}$$

2.3.2. Sampling the conditional distribution of $\mathbf{m}$

The conditional distribution of $\mathbf{m}$ is proportional to the product of the likelihood (with known $\mathbf{a}$ and σ^2) and the prior on $\mathbf{m}$

$$\mathrm{p}(\mathbf{m}|\sigma^2, \mathbf{a}, \mathbf{f}) \propto \mathrm{p}(\mathbf{m})\mathrm{p}(\mathbf{f}|\mathbf{m}, \sigma^2, \mathbf{a}). \tag{18}$$

For fixed $\mathbf{a}$, $g(\mathbf{m}, \mathbf{a})$ is still a non-linear function and the conditional probability has no simple exact analytical expression. However, non-linearities are moderate, particularly around the point of maximal conditional probability for $\mathbf{m}$. A sensible strategy for this kind of situations, advocated by Tarantola [19], is to find the point of maximal probability using an iterative gradient-based method and then to approximate the distribution around this point using the partial derivatives of the function. This is indeed what it is usually done in RUS, where the stiffness coefficients are found by minimizing the differences between experimental and predicted frequencies in a least-square sense using the Levenberg-Marquardt algorithm, and the uncertainties are obtained from an expansion of the quadratic cost function around the solution [3].

A method to find the point maximizing the probability in Eq. (18) and to approximate the conditional around this point for a multivariate normal prior on the elastic parameters $\mathbf{m}$ was exposed in a previous work [17] based on Tarantola [19]. It consists of a quasi-Newton iterative algorithm where $\mathbf{m}$ is updated using

$$\mathbf{m}_{n+1} = \mathbf{m}_n + (\mathbf{G}_n^t \mathbf{C_f}^{-1} \mathbf{G}_n + \mathbf{C_m}^{-1})^{-1} \left[\mathbf{G}_n^t \mathbf{C_f}^{-1} (\mathbf{g}(\mathbf{m}) - \mathbf{f}) + \mathbf{C_m}^{-1} (\mathbf{m}_n - \mathbf{m}_{\text{prior}}) \right],$$
(19)

where $\mathbf{G}_n$ is the matrix of partial derivatives of the resonant frequencies with respect to each elastic parameter at step n, $\mathbf{C_f}$ is the (diagonal) covariance matrix of the data and $\mathbf{m}_{\text{prior}}$ and $\mathbf{C_m}$ are respectively the mean and covariance of the prior distribution of $\mathbf{m}$. The conditional (18) is then approximated by a multivariate normal centered on $\mathbf{m}_{\text{final}}$ with covariance

$$\tilde{\mathbf{C}}_{\mathbf{m}} \approx (\mathbf{G}^t \mathbf{C_f}^{-1} \mathbf{G} + \mathbf{C_m}^{-1})^{-1}.$$
(20)

A sample of the approximate conditional can then be generated using the Cholesky decomposition of $\tilde{\mathbf{C}}_{\mathbf{m}}$ [29]. The quasi-Newton method is stopped when the relative change in $\mathbf{m}$ or in the misfit become smaller than a specified value, set to 10^{-2}. Because we generate a sample in a normal approximation of the conditional, a more precise estimation of the center of the distribution (relative to its width) is useless. It usually requires 1 to 8 computations of the forward problem. These few forward problem solving steps at each iteration of the Gibbs sampler account for the main part of the total computation time.

Due to the approximation, the random sample drawn that way is not a sample of the full conditional (18), which causes the stationary distribution of the Gibbs sampler to be different from the target distribution (10). To ensure that we sample the target distribution, the approximation is corrected by introducing a Metropolis-Hasting rejection step for the samples $\mathbf{m}$, considering the approximation as a proposal density [23, 30]. If we denote $\tilde{p}(\mathbf{m})$ the proposal density, the probability of accepting the move from $\mathbf{m}$ to $\mathbf{m}^*$ is

$$P(\mathbf{m}^*, \mathbf{m}) = \min \left\{ 1, \frac{\tilde{p}(\mathbf{m}) p(\mathbf{m}^*) p(\mathbf{f}|\mathbf{m}^*, \sigma^2, \mathbf{a})}{\tilde{p}(\mathbf{m}^*) p(\mathbf{m}) p(\mathbf{f}|\mathbf{m}, \sigma^2, \mathbf{a})} \right\}.$$
(21)

Strictly speaking, we no longer use Gibbs sampling but a Metropolis-Hasting-within-Gibbs algorithm [23]. However, since the normal approximation of the full conditional is

actually very good, the rejection rate is low and we continue to refer to Gibbs sampling for simplicity.

The described approach requires a multivariate normal prior on the d-dimensional vector of the elastic parameters $\mathbf{m}$, where d depends on the elastic symmetry of the material (e.g. for isotropic material $d = 2$). However, a normal prior on the components of the elastic tensor tensor C_{ij} is not consistent with the thermodynamical constraints on the tensor. We therefore did not use the C_{ij} coefficients, but a set of transformed elastic parameters for which a normal prior is more consistent. The new elastic parameters are described in appendix for symmetries ranging from isotropy to orthotropy ($d = 9$).

2.3.3. Sampling the conditional distribution of σ^2

The conditional distribution of σ^2 is proportional to the product of the likelihood (with known $\mathbf{a}$ and $\mathbf{m}$) and the prior $\mathrm{p}(\sigma^2)$:

$$\mathrm{p}(\sigma^2|\mathbf{m}, \mathbf{a}, \mathbf{f}) \propto \mathrm{p}(\sigma^2) \frac{1}{\sigma^N} \exp\left[-\frac{NS^2}{2\sigma^2}\right], \tag{22}$$

with the sum of squares

$$S^2 = \frac{1}{N} \sum_{n=1}^{N} \left(\frac{g_n(\mathbf{m}, \mathbf{a}) - f_n}{f_n}\right)^2. \tag{23}$$

A convenient family of prior distribution $\mathrm{p}(\sigma^2)$ is the scaled-inverse-chi-square distribution, which is conjugate to the likelihood. It means that the conditional Eq. (22) is scaled inverse-chi-square as well [30]. Formally if,

$$\mathrm{p}(\sigma^2) = \mathrm{Inv}\text{-}\chi^2(\sigma^2|\nu, \sigma_{prior}^2), \tag{24}$$

where we use the notation from Gelman et *al.* [30], then the conditional writes

$$\mathrm{p}(\sigma^2|\mathbf{m}, \mathbf{a}, \mathbf{f}) = \mathrm{Inv}\text{-}\chi^2\left(\sigma^2 \middle| \nu + N, \frac{\nu\sigma_{prior}^2 + NS^2}{\nu_0 + N}\right). \tag{25}$$

This choice is particularly convenient since it is easy to sample directly from the Inv-χ^2 distribution, as it is a particular case of the inverse gamma distribution [30], allowing one to build a sampler from any routine sampling the gamma distribution. For $\nu \to 0$ the prior distribution is $\mathrm{p}(\sigma^2) = 1/\sigma^2$ and is nonimformative, as the conditional (25) then depends only on N and S and the parameter σ_{prior} becomes irrelevant. This value ($\nu = 0$) was used in the applications (sections 3 and 4).

In practice, we found useful to truncate the prior on σ^2 to an upper bound. It is useful in the early states of the chain, when no pairings producing good fit have yet been reached. The sampled values of σ^2 then tend to be large, which makes the likelihood flat and the chain to stay in states of poor fit. Truncating the prior on σ^2 breaks this vicious cycle and prevents the chain from staying for a long time in states of low probability before reaching by chance a state of good fit. The upper bound is set larger than the expected misfit in a successful RUS experiment, so that truncation only improves convergence with no effect on the stationary distribution.

2.3.4. Initialization and convergence

To diagnose convergence, L parallel Gibbs chains starting from dispersed points were run. Since the algorithm starts by sampling the distribution of the pairings (11), we have to choose or generate L initial values $\mathbf{m}_0$ and σ_0^2. The initial variance was set to $\sigma_0^2 = 0.01^2$ for each chain. It corresponds to an assumption of relative standard deviation of 1% on the resonant frequencies, which is of the order of the misfit usually obtained in RUS [4]. A simple way to generate initial stiffness coefficients is to sample the prior distribution, but in the case of a broad prior, this starts the chains very far from the high probability zone and leads to slow convergence. To overcome this problem, we generated L initial values $\mathbf{m_0}$ using importance re-sampling [30]. $M \gg L$ samples were generated in the prior distribution, and the distribution of the pairing explored for each sample using the simulated annealing algorithm (section 2.3.1). The probabilities of all the distinct pairings sampled at $T = 1$ were summed, giving an approximation of the marginal probability of each proposed $\mathbf{m_0}$

$$p(\mathbf{m}_0|\mathbf{f}, \sigma_0^2) \propto \sum_{\mathbf{a}} p(\mathbf{a}, \mathbf{m}_0|\mathbf{f}, \sigma_0^2). \tag{26}$$

The marginal probabilities (26) were used as weights to randomly select L values among the M proposed, each having a chance proportional to its weight to be selected.

After letting the chains run for a while, we assessed convergence using the criterion from Gelman et al. [30]. For each scalar parameter, a ratio R of the between-chain to within-chain variance was computed. A ratio close to one indicates that all the chains are sampling the same area of the posterior pdf. Although it is not our purpose to discuss the difficult problem of assessing convergence of MCMC, we note that it is only an *indication* of possible convergence, but in no way a *proof*. We considered that convergence was reached when $R < 1.1$ for all scalar parameters (stiffness and variance).

All the computations in this paper were performed on a desktop computer with 8 computing cores (2 quad-core Intel Xeon E5620 @2.4GHz processors), with a chain running independently on each core. The algorithm was coded using Matlab®. To assess convergence, it was necessary to regularly pause the chains and pool them to compute the ratios R. This was done after each segment of 100 iterations. The $L = 8$ starting points were selected from $M = 1000$ draws from the prior.

3. Application 1 - Data from Ogi et al., 2002

For the first illustration of the proposed method we used the resonant frequencies published by Ogi et al. [7] for a specimen of polycrystalline aluminum alloy (isotropic symmetry). The paper contains all the necessary information about the specimen: dimensions, mass, and the values of the 42 first measured resonant frequencies. Additionally, the exact mode identification obtained by laser interferometry is given, in the form of a group symmetry label identifying uniquely each measured resonant mode [18]. The group symmetry labels are an output of the forward model (section 2.1), and

can be unambiguously attributed to the experimental frequencies in laser-based RUS from a comparison of the predicted and experimental modal shapes. This application is an interesting test for our method, since this dataset was initially presented to demonstrate the ability of laser-based RUS to solve the inverse problem in a case where, according to the authors, it would have been very difficult with the measurement of the frequencies alone. We report here that the method presented above can be used to estimate accurate stiffness coefficients without using the laser data and without requiring an informative prior on the stiffness.

3.1. Prior distributions

Ogi et *al.* [7] emphasize on the ability of laser-based RUS to find the correct pairing of frequencies even in difficult situation by starting with very unrealistic guess of the stiffness constants: $C_{11} = 300$ and $C_{44} = 50$ Gpa ("true" values are close to 109 and 27 Gpa, respectively). We therefore centered our prior distribution on these values by fixing $\mathbf{m}_{prior} = [0 \ ; \ 0]$, with (see appendix)

$$m_1 = \log\left(\frac{C_{11} - \frac{4}{3}C_{44}}{300 - \frac{4}{3}50}\right), \tag{27a}$$

$$m_2 = \log\left(\frac{C_{44}}{50}\right). \tag{27b}$$

We then constructed a weakly informative prior by choosing the covariance matrix to be 0.1 times the 2-by-2 identity matrix $\mathbf{I}_2$, which gives a broad distribution in the stiffness space (Fig. 2).

For the pairing, we considered the $K = 50$ first predicted frequencies, letting room for up to $k_{max} = 8$ potentially non-observed frequencies. The expectation λ was set to 1. It gave equally high prior probability (≈ 0.37) of having missed $k = 0$ and $k = 1$ resonant frequencies and a decreasing probability for larger k, thus reflecting our prior knowledge that the frequencies are easily measured on a very low damping material.

We used the noninformative prior $p(\sigma^2) = 1/\sigma^2$ for the variance by setting $\nu = 0$ in Eq. (25). We set an upper boundary of $\sigma^2 < 0.03^2$, corresponding to a misfit of 3% between predicted and measured frequencies, much larger that what can be expected from a successful RUS experiment.

3.2. Results and discussion

The 8 starting points generated from the procedure described in Section 2.3.4 are located by crosses in Fig. 2. We see that the 8 starting values of C_{44} are close. This is due to the fact that the first predicted mode is usually an almost pure shear mode, depending mainly on C_{44}. Since this mode is always paired with the first measured frequency, it has the effect of almost fixing the shear coefficient C_{44} close to the correct value. The coefficient C_{11} remains largely undetermined at this step.
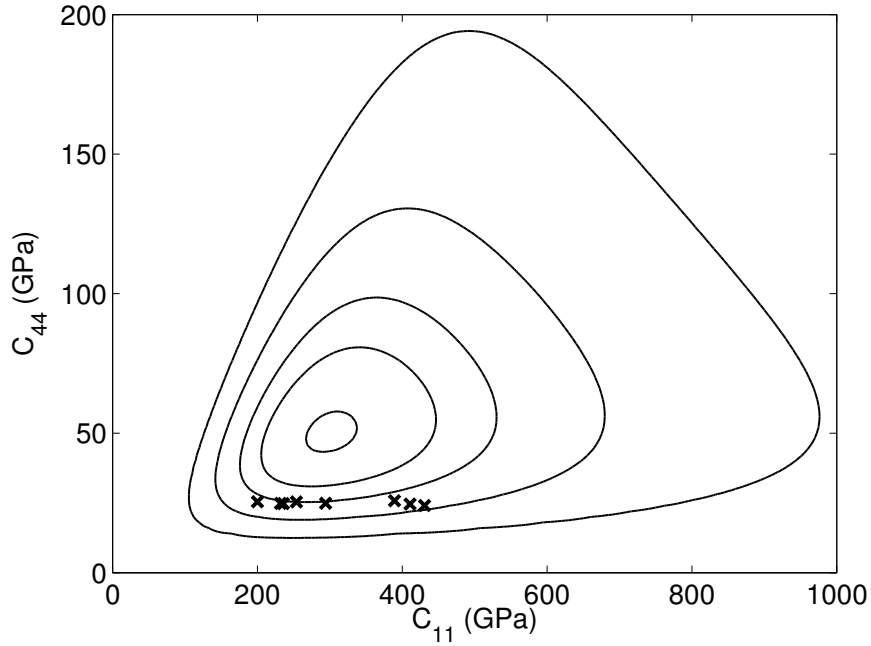
Figure 2. Application 1: prior probability in the space of the stiffness coefficients corresponding to the bivariate normal prior distribution on $\mathbf{m}$, centered on m_1 and m_2 with covariance equal to $0.1 \times \mathbf{I}_2$. The contour lines enclose respectively 10, 68, 95, 99, and 99.99 % of the distribution. The crosses locate the random starting points of the 8 Gibbs chains.
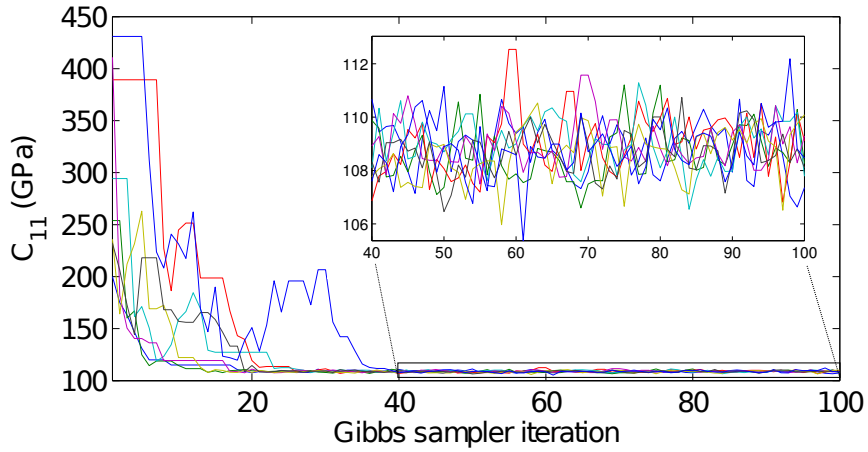
Figure 3. Application 1: values of C_{11} sampled for the 100 first iterations of the 8 Gibbs chain. After 40 iterations, all the chains have converged to the same location.

Convergence of the 8 parallel chains was observed at the first convergence test, after 100 iterations of the Gibbs sampler, but we let them run for 900 more iterations ($\sim$ 4 min of computation time). The convergence was indeed reached after about 40 iterations, as it can be seen on Fig. 3. After discarding the 100 first samples, chains were pooled and the sampled $\mathbf{m}$ transformed back to stiffness coefficients to evaluate the posterior distribution. Any other elastic modulus (e.g. Young's modulus or Poisson's

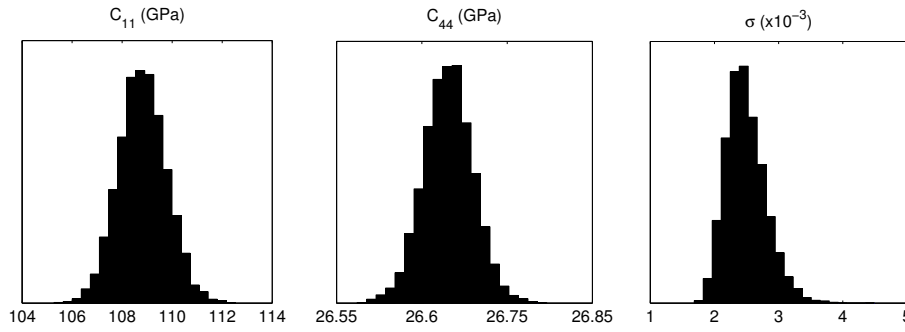**Figure 4. Application 1**: Histograms of the stiffness coefficients and error term σ obtained from Gibbs sampling using the frequency data of the alluminium alloy specimen from Ogi et *al.* [7].

Table 1. Application 1: parameters estimated from the proposed Gibbs sampling method for the data from Ogi et *al.* [7], compared to estimations knowing the exact pairing of the frequencies.

	Gibbs	Exact pair.	Ogi et *al.* [7]
Nb freq.	42	42	> 80
C_{11} (GPa)	108.8 (0.9)	109.0 (0.7)	109.26
C_{44} (GPa)	26.68 (0.03)	26.68 (0.02)	26.72
std (%)	0.24	0.23	0.2

ratio) could also have been calculated.

The histograms for C_{11}, C_{44}, and σ are plotted on Fig. 4. It can be observed that the marginal posterior distributions of the stiffness constants have a normal shape while the marginal for σ is asymmetric. The posterior means and standard deviations for C_{11}, C_{44} were respectively taken as the estimate of the stiffness constant and uncertainties. For σ, the point of maximum marginal posterior probability was estimated from the histogram. The results are given in Table 1, and compared to the estimates obtained assuming the exact pairing. These estimates were obtained from the quasi-Newton method described in section 2.3.2. In that case the estimated standard deviation is the relative root-means-square error between measured and predicted frequencies at the end of the fit. The result obtained with uncertain and exact pairing are almost identical and the uncertainties are only slightly larger in the first case, which shows that the approach proposed here produces results that are almost as accurate and precise as the results obtained using the exact pairing deduced from laser measurements. We also recall the values obtained by Ogi et *al.* using more than 80 frequencies with exact pairing (only the 42 first frequencies were given in the paper). The values are slightly different but agree with the uncertainty intervals.

The 42 published frequencies correspond to the actual 42 first resonant modes of the specimen (no missed frequencies). Our algorithm correctly and unambiguously identified that no frequency was missed, i.e. we obtained p($k = 0$) = 1. Moreover, since the exact pairing is known we can compare the pairings obtained through Gibbs

sampling to the correct pairing. In Table 2, the sampled group symmetry labels for each mode are listed against the exact label obtained from Laser interferometry. We can see that 18 out of 42 modes are uniquely paired and that a total of 32 modes are confidently identified ($> 95\%$ probability) with the correct predicted mode. The other modes are identified with larger uncertainties and with up to 4 possibilities of pairing. However, the most probable mode label is always the correct one, except for two pairs of modes that were inverted (modes $38 - 39$ and $41 - 42$). All cases of uncertain or inexact identifications correspond to resonant frequencies that are very close to each others, compared to the estimated standard deviation of the experimental frequencies (0.24%). It is then not surprising that these frequencies cannot be confidently paired to a unique predicted frequency. But we argue that it is precisely in those cases that the exact pairing matters the less. Indeed, inverting two or more frequencies that are separated by less than the standard deviation can only have a negligible influence on the estimated stiffness.

The correct pairing was obtained by the method (except for some very close modes) using only the frequencies, which shows that the information on the vibration modes is indeed contained in the distribution of the frequencies. Then, getting the modal shapes experimentally only provides redundant information here.

3.3. Additional results

We solved the inverse problem for the same data set but for a more general assumption of orthotropic elastic symmetry. The prior distribution was centered on $C_{11} = 100$, $C_{22} = 150$, $C_{33} = 200$, $C_{12} = 70$, $C_{13} = 60$, $C_{23} = 50$, $C_{44} = 20$, $C_{55} = 30$, $C_{66} = 40$, with a covariance matrix again equal to 0.1 times the identity matrix. After about 1000 iterations the 8 chains were eventually sampling the same peak of the pdf and were run for 20000 additional samples. The sampled values of the orthotropic elastic coefficients, presented in Table 3, are very similar to the isotropic ones and to those obtained with the exact pairing. It shows that the proposed algorithm is able to accurately and automatically estimates 9 orthotropic stiffness constants, without requiring to measure the modal shape of the modes nor to specify an informative prior. Convergence could have been greatly improved with the use of an informative prior, which may actually be available in many practical applications.

We also applied the method to two additional data sets, published in another paper proposing an experimental method for the measurement of the modal shapes of the modes [6]. The data consists in 15 resonant frequencies measured on an isotropic $Ni_{80}P_{20}$ alloy specimen and 31 frequencies measured on a cubic YbAl specimen. In both case the algorithm performed very well and accurately estimated the stiffness constants, starting from a broad prior distribution (results not shown). The first of these two data sets is particularly interesting since three predicted resonant frequencies were not measured. The algorithm unambiguously identified $k = 3$ missing frequencies: $p(k = 3) = 1$.

Finally, we tested the robustness of the method to the expected number of missed

Table 2. Application 1: experimental resonant frequencies and group symmetry label from Laser-based RUS measurement [7] and group symmetry labels proposed during Gibbs sampling with their percentage of occurrence. Labels indicate belongings to one of the 8 distinct symmetry groups and the ordering in each group, following the notation introduced by Mochizuki [18].

#	f_{meas} (kHz)	Mode (Laser) [7]	Modes Gibbs (% of occurrence)			
1	116.716	A_u-1	A_u-1 (100)	-	-	-
2	143.783	A_u-2	A_u-2 (100)	-	-	-
3	158.081	B_{1u}-1	B_{1u}-1 (100)	-	-	-
4	166.500	B_{2u}-1	B_{2u}-1 (100)	-	-	-
5	169.523	B_{1g}-1	B_{1g}-1 (100)	-	-	-
6	177.846	B_{2g}-1	B_{2g}-1 (100)	-	-	-
7	183.875	B_{3u}-1	B_{3u}-1 (99.2)	B_{3g}-1 (0.8)	-	-
8	186.047	B_{3g}-1	B_{3g}-1 (99.2)	B_{3u}-1 (0.8)	-	-
9	190.341	A_g-1	A_g-1 (100)	-	-	-
10	197.386	B_{1u}-2	B_{1u}-2 (100)	-	-	-
11	201.136	A_g-2	A_g-2 (100)	-	-	-
12	207.386	B_{3g}-2	B_{3g}-2 (100)	-	-	-
13	209.836	A_g-3	A_g-3 (100)	-	-	-
14	214.753	B_{2g}-2	B_{2g}-2 (100)	-	-	-
15	223.548	B_{3u}-2	B_{3u}-2 (100)	-	-	-
16	231.266	B_{2u}-2	B_{2u}-2 (100)	-	-	-
17	233.538	B_{3g}-3	B_{3g}-3 (99.1)	B_{1g}-2 (0.9)	-	-
18	234.717	B_{1g}-2	B_{1g}-2 (99.1)	B_{3g}-3 (0.9)	-	-
19	250.980	A_g-4	A_g-4 (55.0)	A_u-3 (44.7)	B_{2g}-3 (0.3)	-
20	251.256	A_u-3	A_u-3 (54.6)	A_g-4 (44.7)	B_{2g}-3 (0.7)	-
21	252.742	B_{2g}-3	B_{2g}-3 (99.0)	A_u-3 (0.7)	A_g-4 (0.3)	-
22	256.122	B_{1u}-3	B_{1u}-3 (98.4)	B_{3u}-3 (1.5)	A_g-5 (0.1)	B_{2u}-3 (0.0)
23	257.595	B_{3u}-3	B_{3u}-3 (72.6)	A_g-5 (19.6)	B_{2u}-3 (6.3)	B_{1u}-3 (1.5)
24	258.118	A_g-5	A_g-5 (51.0)	B_{2u}-3 (25.9)	B_{3u}-3 (23.0)	B_{1u}-3 (0.1)
25	259.035	B_{2u}-3	B_{2u}-3 (67.9)	A_g-5 (29.3)	B_{3u}-3 (2.8)	-
26	268.540	B_{1g}-3	B_{1g}-3 (100)	-	-	-
27	277.113	B_{2u}-4	B_{2u}-4 (99.9)	B_{3u}-4 (0.1)	-	-
28	278.762	B_{3u}-4	B_{3u}-4 (99.9)	B_{2u}-4 (0.1)	-	-
29	282.311	B_{1u}-4	B_{1u}-4 (100)	-	-	-
30	293.686	B_{3u}-5	B_{3u}-5 (100)	-	-	-
31	304.159	B_{2u}-5	B_{2u}-5 (58.9)	B_{1u}-5 (41.1)	-	-
32	304.464	B_{1u}-5	B_{1u}-5 (58.9)	B_{2u}-5 (41.1)	-	-
33	310.109	B_{1u}-6	B_{1u}-6 (100)	-	-	-
34	316.197	B_{1g}-4	B_{1g}-4 (97.7)	B_{2g}-4 (2.3)	-	-
35	317.392	B_{2g}-4	B_{2g}-4 (97.7)	B_{1g}-4 (2.3)	-	-
36	326.462	A_u-4	A_u-4 (100)	B_{3g}-4 (0.0)	-	-
37	329.034	B_{3g}-4	B_{3g}-4 (100)	A_u-4 (0.0)	-	-
38	332.441	A_g-6	B_{2u}-6 (70.9)	A_g-6 (29.1)	-	-
39	333.364	B_{2u}-6	A_g-6 (70.9)	B_{2u}-6 (29.1)	-	-
40	336.650	B_{1g}-5	B_{1g}-5 (69.6)	A_g-7 (18.8)	B_{2g}-5 (11.6)	-
41	337.359	B_{2g}-5	A_g-7 (38.1)	B_{2g}-5 (36.7)	B_{1g}-5 (25.1)	-
42	338.276	A_g-7	B_{2g}-5 (51.7)	A_g-7 (43.1)	B_{1g}-5 (5.2)	-

Table 3. Application 1: stiffness constants (in GPa) estimated from the data from Ogi et *al.* [7] for a more general orthotropic elasticity assumption (see section 3.3).

	Gibbs	Exact pairing	Ogi et *al.* [7]
C_{11}	108.9 ± 0.8	109.3 ± 0.6	109.26
C_{22}	108.7 ± 0.6	109.0 ± 0.5	
C_{33}	109.3 ± 0.6	109.5 ± 0.4	
C_{12}	55.5 ± 0.7	55.9 ± 0.5	55.82
C_{13}	55.5 ± 0.7	55.8 ± 0.5	
C_{23}	56.3 ± 0.6	56.6 ± 0.4	
C_{44}	26.95 ± 0.06	26.96 ± 0.05	26.72
C_{55}	26.52 ± 0.06	26.52 ± 0.05	
C_{66}	26.85 ± 0.05	26.84 ± 0.04	

frequencies λ. We attributed to λ a noninformative prior, in the form of a gamma prior – Eq. (16) – with $\alpha = 1/2$ and $\beta = 0$ [28]. For all the applications described above, the algorithm converged to the same results, although convergence was somewhat longer. It shows that the method is weakly sensitive to λ for these data sets. Hence, setting $\lambda = 1$ or another low value should be adequate for applications to low damping materials.

4. Application 2 - Data from Bernard al., 2013

The second application deals with the recently developed application of RUS to highly attenuating materials [17] and in particular to cortical bone [16]. This application leads to an even more difficult identification problem. Indeed, due to the high amount of vibration damping, the spectrum of the specimen is not composed of sharp resonant peaks, but of broad and overlapping peaks. The frequencies must then be estimated through signal processing [31, 16, 17]. Due to overlapping, it is likely that some frequencies are missed during the measurement and there is no simple way to tell which modes are more likely to be missed [17]. We do not discuss here signal processing and we consider the resonant frequencies as given input data.

We applied the method described above to the 20 resonant frequencies measured by Bernard et *al.* [16] for a specimen of human femoral cortical bone. In that case, the exact pairing of the measured frequencies was not known, but an independent measurement of the stiffness coefficients was available to evaluate the results.

4.1. Prior distributions

The measured specimen was taken in a population of 21 femoral bone specimens whose diagonal elastic coefficients (C_{ii}, $i = 1, ..., 6$) were previously measured using ultrasonic bulk waves velocities [32]. It was then possible to construct a prior distribution using the previous results obtained for the whole population. For the transformed parameters **m** corresponding to the diagonal coefficients (see Appendix A) the center of the prior
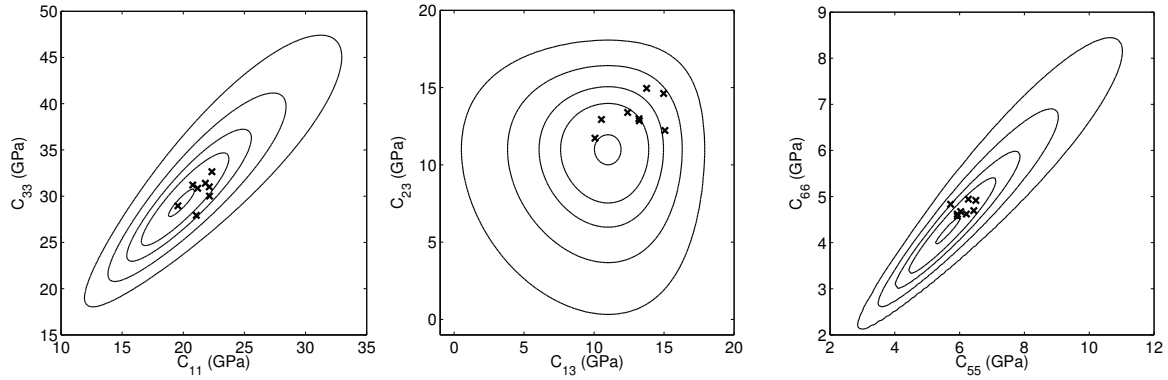
Figure 5. Application 2: prior probability in the space of the stiffness coefficients corresponding to the multivariate normal prior distribution on $\mathbf{m}$. The contour lines embed respectively 10, 68, 95, 99, and 99.99 % of the distribution. The cross locate the randomly generated starting points of the 8 Gibbs chains.

distribution was fixed using the mean values of the measured coefficients, which are $C_{11} = 19.3$, $C_{22} = 19.8$, $C_{33} = 29.2$, $C_{44} = 5.8$, $C_{55} = 5.6$, and $C_{66} = 4.3$ (in GPa). Since the three off-diagonal coefficients C_{12}, C_{13}, and C_{13} were not measured, their prior were centered on $C_{ij} = C_{11} - 2 \times C_{66} = 10.7$ GPa. Variances and covariances of the parameters corresponding to the diagonal coefficients were estimated from the population results. No covariances were assumed between the off-diagonal and the diagonal coefficients, nor between the three off-diagonal coefficients. Their variances were fixed to 0.05. This gives the following prior covariance matrix for $\mathbf{m}$:

$$0.01 \times \begin{pmatrix} C_{11} & C_{22} & C_{33} & \eta_{12} & \eta_{13} & \eta_{23} & C_{44} & C_{55} & C_{66} \\ 1.35 & 1.24 & 1.19 & 0 & 0 & 0 & 1.41 & 1.71 & 1.72 \\ 1.24 & 1.40 & 1.19 & 0 & 0 & 0 & 1.48 & 1.63 & 1.79 \\ 1.19 & 1.19 & 1.27 & 0 & 0 & 0 & 1.42 & 1.66 & 1.64 \\ 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 \\ 1.41 & 1.48 & 1.42 & 0 & 0 & 0 & 1.84 & 1.93 & 2.04 \\ 1.71 & 1.63 & 1.66 & 0 & 0 & 0 & 1.93 & 2.48 & 2.40 \\ 1.72 & 1.79 & 1.64 & 0 & 0 & 0 & 2.04 & 2.40 & 2.58 \end{pmatrix}. \tag{28}$$

Three 2D sections of the prior distribution are plotted on Fig. 5. The assumed *a priori* correlation between C_{11} and C_{33} on one hand, and C_{55} and C_{66} on the other hand, can be observed.

For σ^2, the same nonimformative prior as in application 1 was used. The expected number of missed frequencies λ was not fixed, but an noninformative prior was attributed to it (section 3.3).

4.2. Results and discussion

The 8 randomly generated starting points are shown on Fig. 5. In that case, all coefficients show some scattering. This is due to anisotropy. Indeed, although the first predicted mode is still an almost pure shear mode, it depends on a combination of the three shear coefficients.

Convergence of the Gibbs chains was indicated after about 500 iterations, but the chains were stopped after 5000 samples (about 20 minutes of computing time). The values of the elastic parameters sampled after convergence were transformed back to the coefficients C_{ij}. Histograms of the 9 stiffness coefficients and the standard deviation σ are represented on Fig. 6. Again, the marginal posterior distributions for the elastic coefficient look normal, while the distribution of σ is asymmetric. The mean values of the C_{ij} are given in Table 4. They are compared to the results obtained for a fixed pairing, that was obtained manually [16] and to the independent results obtained from wave velocities measurements. Good agreement is observed between the three results. Particularly, no significant difference is observed between results from RUS with fixed pairing and from Gibbs sampling.

Despite the non-informative prior on λ, the number of missed frequencies was unambiguously identified to $k = 10$ [p($k = 10$) = 0.988], the same number that was obtained from manual pairing [16].

In this application to a highly attenuating material, an informative prior on the stiffness is necessary for the identification of the pairing and then the estimation of the stiffness. Due to the low number of measured frequencies (20) relative to the number of elastic parameters (9), the pairing cannot be entirely determined from the distribution of the frequencies, contrary to the previous application. This is a limitation of the proposed approach, however compensated by the potential availability of prior information in many applications.

The prior distribution of the elastic parameters was constructed from available results for a population of similar samples. In some practical applications, this might be done through review of reported elasticity for similar materials. In some context, like composites materials, a prior distribution might also be obtained from theoretical predictions of elasticity. If no prior is available, it might be necessary to perform preliminary measurements, for instance of ultrasonic bulk wave velocities in some material directions.

5. Conclusion

We presented a method based on Bayesian inference and MCMC sampling to automatically solve the inverse problem in RUS, including the identification of the resonant frequencies, without requiring the measurement of the modal shapes associated to each resonant frequency. The proposed statistical approach required more forward problem solving than the conventional optimization method, but convergence was
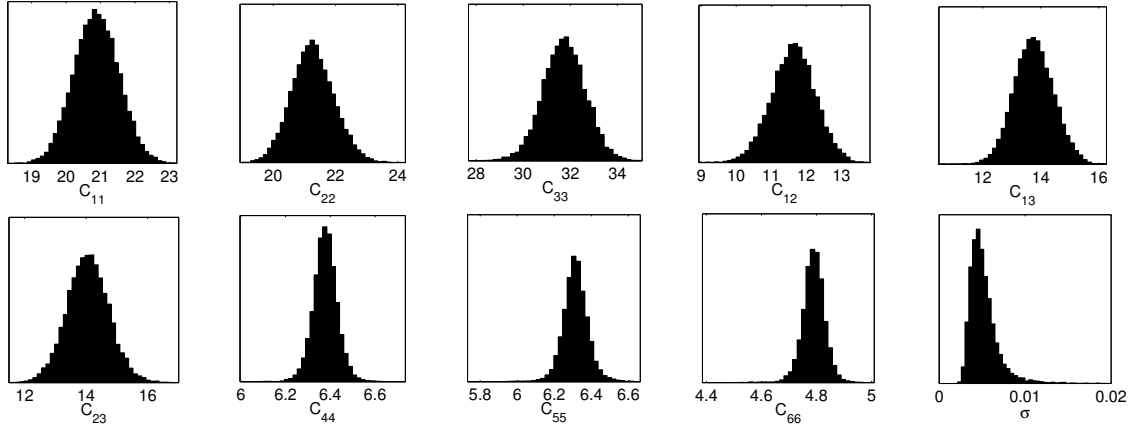
Figure 6. Application 2: histograms of the stiffness coefficients and error term σ obtained from Gibbs sampling using the frequency data of the bone specimen from Bernard *et al.* [16].

Table 4. Application 2: stiffness constants (in GPa) estimated using Gibbs sampling and using a fixed pairing, and results from bulk wave velocity measurements. Data for a human femoral cortical bone specimen [16].

	Gibbs	Fixed pairing [16]	BWV [16]
C_{11}	20.9 ± 0.6	20.3 ± 0.6	21.6
C_{22}	21.3 ± 0.7	20.2 ± 0.6	21.4
C_{33}	31.7 ± 0.9	31.7 ± 0.8	31.3
C_{12}	11.6 ± 0.6	10.7 ± 0.6	-
C_{13}	13.8 ± 0.6	13.4 ± 0.7	-
C_{23}	14.1 ± 0.7	13.4 ± 0.7	-
C_{44}	6.38 ± 0.06	6.38 ± 0.02	6.5
C_{55}	6.31 ± 0.07	6.32 ± 0.03	6.5
C_{66}	4.79 ± 0.04	4.80 ± 0.02	4.8
std (%)	0.43	0.30	-

obtained within a very reasonable time: about a few minutes or a few tens of minutes, depending on the application. In our experience with RUS, this is less than the time required for specimen preparation and spectrum measurement, so that computing time should not be considered as a limitation.

Faster deterministic alternatives to MCMC exist, usually based on the approximation of the posterior distribution in a parametric form. Such methods include Variational Bayes methods, in which the parameters of the approximated posterior are obtained iteratively, and the method we used in section 2.3.2 to sample the elastic parameters by reaching the point of maximum posterior probability and approximating the posterior around this point. Despite the fact that they are faster than MCMC, these methods may not be viable alternatives here because they are less general, prone to trapping into local optima, and might fail to take care of the discrete part of the

problem at hand (i.e. finding the correct pairing).

In the context of RUS applied to weakly attenuative materials (application 1), where many resonant frequencies are easily measured, we showed that the method allows estimating the stiffness constants with high accuracy and precision, without requiring introduction of prior information. For the more difficult application of RUS to attenuating materials (application 2), less resonant frequencies can be measured. The joint estimation of the pairing and stiffness constants can then be ambiguous. However, the Bayesian formalism allows to including the available prior information to resolve the ambiguities.

The method was illustrated for anisotropy up to the orthotropic symmetry (9 stiffness coefficients). In principle, there is no restriction on the number of parameters that can be estimated, and the method could therefore be generalized to lower elastic symmetry and/or to the joint estimation of the Euler's angles in the case of misoriented specimens [33] (i.e. when the principal axis of elastic symmetry are not aligned with the edges of the specimen) or to the measurement of piezoelectric properties of materials [34, 8, 9]. In such applications, the simple sampling method proposed here might be less efficient and the use of methods particularly dedicated to sampling highly complex distributions, such as parallel tempering [35], could be necessary. This is a possible direction for future research.

Acknowledgment

Appendix A. Transformed elastic parameters

Appendix A.1. Isotropic symmetry

The isotropic stiffness tensor is, in the two index notation:

$$
C_{ij} = \begin{pmatrix}
C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\
C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\
C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\
0 & 0 & 0 & C_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & C_{44} & 0 \\
0 & 0 & 0 & 0 & 0 & C_{44}
\end{pmatrix}, \tag{A.1}
$$

with the relation $C_{12} = C_{11} - 2C_{44}$. It involves two parameters. Thermodynamics constraints requiring the tensor to be positive-definite are

$$
C_{11} > \frac{4}{3}C_{44} > 0. \tag{A.2}
$$

The use of C_{11} and C_{44} as random variables in the Bayesian formulation implies two difficulties in the definition of a prior distribution: 1) their positiveness is not consistent with a normal distribution and 2) they must satisfy $C_{11} > (4/3)C_{44}$, and should then be statistically dependent variables. Constraints (A.2) transform to:

$$\left\{ K = C_{11} - \frac{4}{3}C_{44} > 0 \; ; \; G = C_{44} > 0 \right\}, \tag{A.3}$$

where K and G are respectively the bulk and shear modulus, two statistically independent random variables. Their logarithms

$$m_1 = \log(K/K_0), \tag{A.4a}$$

$$m_2 = \log(G/G_0), \tag{A.4b}$$

with arbitrary constants K_0 and G_0 are consistent with a normal distribution on $\mathbf{m} = [m_1; \; m_2]$.

Appendix A.2. Cubic symmetry

The cubic elastic tensor is (A.1), without the relation between C_{12} and the two other parameters. Positive-definiteness is obtained with positiveness of the bulk modulus K and of the two shear modulus G_1 and G_2, defined as [36]:

$$K = (C_{11} + 2C_{12})/3, \tag{A.5a}$$

$$G_1 = (C_{11} - C_{12})/2, \tag{A.5b}$$

$$G_2 = C_{44}. \tag{A.5c}$$

The logarithms m_1, m_2, and m_3 of K, G_1 and G_2 are statistically independent variables consistent with a normal distribution.

Appendix A.3. Transversely Isotropic symmetry

The transversely isotropic stiffness tensor is:

$$C_{ij} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{pmatrix}, \tag{A.6}$$

with the relation $C_{12} = C_{11} - 2C_{66}$, and therefore involves five parameters. A necessary and sufficient set of conditions for positive-definiteness is [37]

$$\left\{ C_{11} > C_{66} > 0; \; C_{44} > 0; \; C_{33}(C_{11} - C_{66}) > C_{13}^2 \right\}. \tag{A.7}$$

Denoting $\kappa = C_{11} - C_{66}$, it transforms to

$$\{\kappa > 0 \; ; \; C_{66} > 0; \; C_{44} > 0; \; C_{33} > 0\} \, , \text{and} \tag{A.8a}$$

$$-1 < x = \frac{C_{13}}{\sqrt{C_{33}\kappa}} < 1. \tag{A.8b}$$

This last variable x can be transformed to a positive variable using $\eta = (1+x)/(1-x)$, and we then have five independent positive variables $\{\kappa; \; C_{33}; \; \eta; \; C_{44}; \; C_{66}\}$ that can be mapped onto variables m_i $(i = 1, ..., 5)$ defined from $-\infty$ to $+\infty$ by taking their logarithm.

Appendix A.4. Orthotropic symmetry

The orthotropic stiffness tensor is:

$$C_{ij} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{pmatrix}, \tag{A.9}$$

with no relation between the coefficients. It involves nine parameters. The conditions for positive-definiteness are that all the principal minors (determinants of the $k \times k$ sub-matrices for $k = 1, ..., 6$) of (A.9) are positives [37]. Although it may exists a set of nine independent real-valued variables that satisfies this conditions, we did not find a simple transformation leading to such a parametrization. Instead, we use the logarithms of the positives parameters $\{C_{11}; \; C_{22}; \; C_{33}; \; \eta_{12}; \; \eta_{13}; \; \eta_{23}; \; C_{44}; \; C_{55}; \; C_{66};\}$ with

$$\eta_{ij} = \frac{1 + C_{ij}/\sqrt{C_{ii}C_{jj}}}{1 - C_{ij}/\sqrt{C_{ii}C_{jj}}}. \tag{A.10}$$

This parametrization ensures that all the principal minors of order 1 and 2 are positive but *does not* guaranty

$$\det \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ C_{12} & C_{22} & C_{23} \\ C_{13} & C_{23} & C_{33} \end{pmatrix} > 0. \tag{A.11}$$

Hence the prior distribution on **m** must be a truncated normal, attributing zero probability for non positive-definite tensors. It was obtained by rejecting the non positive-definite samples during Gibbs sampling .

References

[1] J. D. Maynard, "The use of piezoelectric film and ultrasound resonance to determine the complete elastic tensor in one measurement," *J. Acoust. Soc. Am.*, vol. 91, no. 3, pp. 1754–1762, 1992.

[2] A. Migliori, J. L. Sarrao, W. M. Visscher, T. M. Bell, M. Lei, Z. Fisk, and R. G. Leisure, "Resonant ultrasound spectroscopic techniques for measurement of the elastic moduli of solids," *Physica B*, vol. 183, no. 12, pp. 1 – 24, 1993.

[3] A. Migliori and J. L. Sarrao, *Resonant Ultrasound Spectroscopy.* Wiley, New York, 1997.

[4] A. Migliori and J. D. Maynard, "Implementation of a modern resonant ultrasound spectroscopy system for the measurement of the elastic moduli of small solid specimens," *Rev. Sci. Instrum.*, vol. 76, no. 12, p. 121301, 2005.

[5] I. Ohno, "Free vibration of a rectangular parallelepiped crystal and its application to determination of elastic constants of orthorhombic crystals," *J. Phys. Earth*, vol. 24, pp. 355 – 379, 1976.

[6] A. Stekel, J. L. Sarrao, T. M. Bell, M. Lei, R. G. Leisure, W. M. Visscher, and A. Migliori, "Method for identification of the vibrational modes of a rectangular parallelepiped," *J. Acoust. Soc. Am.*, vol. 92, no. 2, pp. 663–668, 1992.

[7] H. Ogi, K. Sato, T. Asada, and M. Hirao, "Complete mode identification for resonance ultrasound spectroscopy," *J. Acoust. Soc. Am.*, vol. 112, no. 6, pp. 2553–2557, 2002.

[8] H. Ogi, N. Nakamura, K. Sato, M. Hirao, and S. Uda, "Elastic, anelastic, and piezoelectric coefficients of langasite: resonance ultrasound spectroscopy with laser-doppler interferometry," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 50, pp. 553 –560, may 2003.

[9] H. Ledbetter, H. Ogi, and N. Nakamura, "Elastic, anelastic, piezoelectric coefficients of monocrystal lithium niobate," *Mech. Mater.*, vol. 36, no. 10, pp. 941 – 947, 2004.

[10] S. J. Reese, K. Telschow, T. Lillo, and D. H. Hurley, "On the establishment of a method for characterization of material microstructure through laser-based resonant ultrasound spectroscopy," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 55, pp. 770–777, April 2008.

[11] M. Landa, P. Sedlák, H. Seiner, L. Heller, L. Bicanová, P. Šittner, and V. Novák, "Modal resonant ultrasound spectroscopy for ferroelastics," *Appl. Phys., A*, vol. 96, pp. 557–567, 2009.

[12] F. Farzbod and D. Hurley, "Using eigenmodes to perform the inverse problem associated with resonant ultrasound spectroscopy," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 59, pp. –, November 2012.

[13] P. Sedlák, H. Seiner, J. Zídek, M. Janovská, and M. Landa, "Determination of all 21 independent elastic coefficients of generally anisotropic solids by resonant ultrasound spectroscopy: Benchmark examples," *Exp. Mech.*, pp. 1–13, 2014.

[14] J. H. Kinney, J. R. Gladden, G. W. Marshall, S. J. Marshall, J. H. So, and J. D. Maynard, "Resonant ultrasound spectroscopy measurements of the elastic constants of human dentin," *J. Biomech.*, vol. 37, no. 4, pp. 437 – 441, 2004.

[15] T. Lee, R. S. Lakes, and A. Lal, "Investigation of bovine bone by resonant ultrasound spectroscopy and transmission ultrasound," *Biomechan. Model. Mechanobiol.*, vol. 1, pp. 165–175, 2002.

[16] S. Bernard, Q. Grimal, and P. Laugier, "Accurate measurement of cortical bone elasticity tensor with resonant ultrasound spectroscopy," *J. Mech. Behav. Biomed. Mater.*, vol. 18, no. 0, pp. 12 – 19, 2013.

[17] S. Bernard, Q. Grimal, and P. Laugier, "Resonant ultrasound spectroscopy for viscoelastic characterization of anisotropic attenuative solid materials," *J. Acoust. Soc. Am.*, vol. 135, no. 5, pp. 2601–2613, 2014.

[18] E. Mochizuki, "Application of group theory to free oscillations of an anisotropic rectangular parallelepiped," *J. Phys. Earth*, vol. 35, pp. 159–170, 1987.

[19] A. Tarantola, *Inverse Problem Theory and methods for model parameters estimation.* SIAM, 2005.

[20] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms.* Cambridge University Press, 2003.

[21] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.*, vol. 21, pp. 1087–1092, June 1953.

[22] W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications,"

Biometrika, vol. 57, no. 1, pp. 97–109, 1970.

[23] W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice.* London, UK: Chapman & Hall, 1995.

[24] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, pp. 721–741, Nov 1984.

[25] G. Casella and E. I. George, "Explaining the gibbs sampler," *The American Statistician*, vol. 46, no. 3, pp. pp. 167–174, 1992.

[26] V. Cerný, "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm," *Journal of Optimization Theory and Applications*, vol. 45, no. 1, pp. 41–51, 1985.

[27] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, "Optimization by simulated annealing," *science*, vol. 220, no. 4598, pp. 671–680, 1983.

[28] R. Yang and J. Berger, "A catalog of noninformative priors," *ISDS Discussion paper*, vol. 97, no. 42, 1997.

[29] D. P. Kroese, T. Taimre, and Z. I. Botev, *Handbook of Monte Carlo Methods.* Hoboken, New Jersey: Wiley, 2011.

[30] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis.* CRC press, 2013, third edition.

[31] A. V. Lebedev, "Method of linear prediction in the ultrasonic spectroscopy of rock," *Acoust. Phys.*, vol. 48, pp. 339–346, 2002.

[32] M. Granke, Q. Grimal, A. Saïed, P. Nauleau, F. Peyrin, and P. Laugier, "Change in porosity is the major determinant of the variation of cortical bone elasticity at the millimeter scale in aged women," *Bone*, vol. 49, no. 5, pp. 1020 – 1026, 2011.

[33] J. L. Sarrao, S. R. Chen, W. M. Visscher, M. Lei, U. F. Kocks, and A. Migliori, "Determination of the crystallographic orientation of a single crystal using resonant ultrasound spectroscopy," *Rev. Sci. Instrum.*, vol. 65, no. 6, pp. 2139–2140, 1994.

[34] H. Ogi, Y. Kawasaki, M. Hirao, and H. Ledbetter, "Acoustic spectroscopy of lithium niobate: Elastic and piezoelectric coefficients," *J. Appl. Phys.*, vol. 92, no. 5, pp. 2451–2456, 2002.

[35] S. E. Dosso, C. W. Holland, and M. Sambridge, "Parallel tempering for strongly nonlinear geoacoustic inversion," *J. Acoust. Soc. Am.*, vol. 132, no. 5, pp. 3030–3040, 2012.

[36] L. J. Walpole, "The elastic shear moduli of a cubic crystal," *J. Phys. D: Appl. Phys.*, vol. 19, no. 3, p. 457, 1986.

[37] B. A. Auld, *Acoustic fields and waves in solids, 2nd edition*, vol. 1. Malabar, Florida: Krieger, 1990.