Lorentzian-Model-Based Bayesian Analysis for Automated Estimation of Attenuated Resonance Spectrum

Kailiang Xu, *Member, IEEE*, Guillaume Marrelec, Simon Bernard, and Quentin Grimal

***Abstract*—Extracting information from a signal exhibiting damped resonances is a challenging task in many practical cases due to the presence of noise and high attenuation. The interpretation of the signal relies on a model whose order (i.e., the number of resonances) is in general unknown. In this study, the signal is modeled as a sum of Lorentzian lineshapes, and a Bayesian framework is designed to simultaneously remove the baseline distortion, select the number of resonances, and recover the parameters of each lineshape including frequency, damping factor, resonance amplitude, and noise magnitude. The Bayesian problem is solved resorting to a reversible jump Markov chain Monte Carlo (RJ-MCMC) sampling scheme. The algorithm is tested on synthetic signals as well as experimental data from a resonant ultrasound spectroscopy experiment aiming to measure elastic properties. The results show that, compared to the well-known linear prediction singular value decomposition method, the RJ-MCMC method achieves a better performance with the advantages of joint model selection, high accuracy estimation, and uncertainty evaluation. We found that when the signal-to-noise-ratio is larger than 20 dB, the average relative error for frequency extraction is smaller than 0.5%. Such an algorithm enables to estimate the number of resonances and extract tens of resonance parameters from a highly attenuated spectrum, which can significantly facilitate the automated processing of signals exhibiting damped resonances.**

***Index Terms*—Spectrum analysis, Bayesian method, model selection, reversible jump Markov chain Monte Carlo (RJ-MCMC), resonance, deconvolution.**

I. Introduction

RESONANCE is a ubiquitous phenomenon in physics in which a system oscillates with a relatively large amplitude at one or several preferential frequencies. Such a phenomenon occurs widely in nature in the form, e.g., of mechanical and acoustic resonance, electromagnetic waves, nuclear magnetic resonance, electron spin resonance and resonance of quantum wave functions. In mechanical and acoustic spectroscopy methods, resonances have been exploited as a way to infer intrinsic physical properties of a material, such as elasticity, or structural properties, such as defects. A typical way to obtain such properties is to perform a wideband impulse or frequency sweep excitation test on the system. The resulting information takes the form of a resonance spectrum, which usually consists of a series of damped resonant peaks. Each individual resonance is associated with a lineshape closely following a Lorentzian function (the shape of the theoretical frequency response function of an ideal oscillator). Such a Lorentzian function is parameterized by an amplitude, a center frequency, and a quality factor (Q factor driving the width of the peak). Alternatively, in time domain, the resonant signal can be modeled as a sum of exponentially damped sinusoids. From a signal processing point of view, the problem of extracting physical information from a resonant spectrum is twofold, including (i) the detection problem, i.e., to determine the model order or number of modes (sinusoids or Lorentzian curves), and (ii) the estimation problem, i.e., to extract the parameters of each Lorentzian lineshape [1]–[3].

For spectral lineshapes containing weakly damped (high Q factor) resonances, the model order may readily be determined by counting the number of well-resolved sharp peaks. However, for spectral lineshapes containing highly damped (low Q factor) resonances, the Lorentzian lineshapes overlap and, in general, the resonant frequencies do not correspond to the local maxima of the spectrum. For such spectra, the problem of estimating the parameters of the resonances is coupled to the problem of determining the number of resonances. In addition, the spectra are usually corrupted by noise, making the joint problem of model detection and parameter estimation even more challenging.

When the number of resonances is assumed to be known, classical parametric methods can be applied to achieve resonance extraction. Well-known estimators are mainly eigenanalysis-based methods, such as subspace-based MUSIC method, linear prediction singular value decomposition (LPSVD) method, and ESPRIT method [4]. These methods have been reviewed in many articles, e.g., [2], [3]. However, to ensure good performance of the eigenanalysis-based method, the number of

Manuscript received November 30, 2017; revised August 16, 2018 and October 10, 2018; accepted October 24, 2018. Date of publication October 29, 2018; date of current version November 20, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. D. Robert Iskander. This work was supported by the H2020-MSCA-IF-2014 through project FE-RUS-Blast: 656712. *(Corresponding author: Kailiang Xu.)*

K. Xu is with the Department of Electronic Engineering, Fudan University, Shanghai 200433, China, and also with the Laboratoire d'Imagerie Biomédicale, Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, Sorbonne Université, F-75006 Paris, France (e-mail: xukl@fudan.edu.cn).

G. Marrelec and Q. Grimal are with the Laboratoire d'Imagerie Biomédicale, Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, Sorbonne Université, F-75006 Paris, France (e-mail: guillaume.marrelec@inserm.fr; quentin.grimal@upmc.fr).

S. Bernard is with the Laboratoire Ondes et Milieux Complexes, UMR CNRS 6294, Université Le Havre Normandie, 76058 Le Havre, France (e-mail: s.bernard.simon@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2018.2878543

1053-587X © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

dominant eigenvalues, corresponding to the number of resonances, has to be predefined. When the model order is considered as an unknown, two popular model selection rules are the Akaike information criterion (AIC) [5] and the minimum description length (MDL) principle [6]. However, the AIC and MDL rules require a reliable maximum likelihood parameter estimation for each possible model. There is evidence that AIC and MDL tend to provide incorrect estimates of the model order for signals with low signal-to-noise ratio (SNR) and small sample size [7]. With the predefined number of resonances, Astel *et al.*, [24] performed a Bayesian analysis with MCMC implementation to process the magnetic resonance spectroscopy signals of complex biological mixtures.

Another elegant way to simultaneously solve the detection and estimation problems is to jointly infer the posterior probability density functions (pdf) of both the model order and the parameters of interest. The Bayesian framework is particularly relevant when the model order and parameters are difficult to retrieve due to noise and highly-damped resonances. Furthermore, this framework lends itself to the development of an automated method and has the decisive advantage to associate a degree of confidence (probability level) to each quantity of interest. Andrieu and Doucet [8] applied the reversible jump Markov chain Monte Carlo (RJ-MCMC) method [9] for joint Bayesian model selection and posterior distribution estimation. Rubtsov and Griffin [10] applied the RJ-MCMC method to process a blood plasma spectrum measured with nuclear magnetic resonance spectroscopy, highlighting the potential of this method for detecting noisy, low-amplitude, and overlapping resonances peaks. To the best of our knowledge, no method for joint model selection and parameter estimation including both resonant frequencies and damping parameters (Q factors) has been reported. Such a method would be all the more useful that the fundamental Q factor of the resonance has been widely considered for many applications, such as ultrasonic atomic force microscopy [26], piezoelectric force microscopy [27], antennas [25].

The motivation of the study was to develop a method to retrieve all frequencies and attenuation factors from highly damped spectral lineshapes under the condition of unknown number of resonances. In this paper, we present an automated method of general interest for processing a complex-valued resonance spectrum even with an unknown number of resonance peaks and high attenuation. We design an original joint Bayesian model with the unknown model order (number of resonances) and a full set of resonance parameters, i.e., amplitudes, resonant frequencies, Q factor and noise variance. Specifically, we add to the methodology proposed by Andrieu and Doucet [8] a Lorentzian model with the damping parameter for each resonance, a model of the baseline distortion (which often corrupts experimental signals) and a post-processing step (using k-means clustering) to combine several RJ-MCMC outputs, i.e., estimates obtained from multichannel measurements. The method is critically tested using both synthetic and experimental signals containing strongly damped resonances with additive white Gaussian noise. We demonstrate the utility of the method by processing strongly attenuated resonant spectra obtained from a standard resonant ultrasound spectroscopy (RUS) experiment.

II. Method

A. Forward Model of Lorentzian Lineshape Spectrum

The resonant spectrum is often modeled as a finite sum of resonant components [14]. Each component is characterized by a Lorentzian lineshape, which is found in many physical situations involving resonant systems. A Lorentzian function $L(f; A, f_0, \phi, Q)$ characterized by its amplitude A, center frequency f_0, phase angle ϕ, and quality factor Q is defined by

$$L(f; A, f_0, \phi, Q) = \frac{Ae^{i\phi}}{(f_0^2 - f^2) + i\frac{ff_0}{Q}}. \tag{1}$$

Setting

$$\begin{cases} a = & A\cos\phi \\ b = & A\sin\phi, \end{cases} \tag{2}$$

it can be rearranged into a real and an imaginary part as

$$L(f; A, f_0, \phi, Q) = \frac{a(f_0^2 - f^2) + b\frac{ff_0}{Q}}{(f_0^2 - f^2)^2 + \left(\frac{ff_0}{Q}\right)^2} + i\frac{b(f_0^2 - f^2) - a\frac{ff_0}{Q}}{(f_0^2 - f^2)^2 + \left(\frac{ff_0}{Q}\right)^2}. \tag{3}$$

A complex spectrum $y(f)$ can then be modeled as a sum of K Lorentzian functions, a baseline drift, and random noise

$$y(f) = \sum_{k=1}^{K} L(f; A_k, f_{0k}, \phi_k, Q_k) + \sum_{m=1}^{M} g_m W_m(f) + \epsilon(f). \tag{4}$$

In this expression, the drift is assumed to be a linear combination of M complex functions $W_m(f)$, $m = 1, \ldots, M$, where each function $W_m(f)$ is weighted by a real coefficient g_m. $\epsilon(f)$ is a complex noise term, where both the real and imaginary parts are assumed to be zero-mean white Gaussian with identical variance σ^2.

In the following, we use the general convention that any complex quantity $\boldsymbol{q} = (q_k) \in \mathbb{C}^K$ will be coded as a vector in $\mathbb{R}^{2K}$ as the concatenation of its K real values followed by its K imaginary values $\boldsymbol{q} = [\Re(q_1), \ldots, \Re(q_K), \Im(q_1), \ldots, \Im(q_K)]^t$.

Assuming that the spectrum is sampled at N frequencies $f_1, \ldots, f_N$ in $[f_{\min}, f_{\max}]$, we define the data vector $\boldsymbol{y} \in \mathbb{R}^{2N}$, the error vector $\boldsymbol{\epsilon} \in \mathbb{R}^{2N}$, and the drift coefficients $\boldsymbol{g} \in \mathbb{R}^{2M}$. In a similar fashion, we define the vector of projections of the Lorentzian amplitudes $\boldsymbol{c} = (a_1, b_1, \ldots, a_K, b_K)^t \in \mathbb{R}^{2K}$, where each pair (a_k, b_k) is defined as in Eq. (2).

Akin to [8], Eq. (4) can be expressed in vector-matrix form as

$$\boldsymbol{y} = \boldsymbol{\Delta\alpha} + \boldsymbol{\epsilon}, \tag{5}$$

where $\boldsymbol{\alpha} = (\boldsymbol{g}^t, \boldsymbol{c}^t)^t \in \mathbb{R}^{2M+2K}$ is the vector of amplitudes for both the baseline drift and the Lorentzian functions. $\boldsymbol{\Delta} = [\boldsymbol{WX}] \in \mathbb{R}^{2N(2M+2K)}$ contains the definition of the Lorentzian model $\boldsymbol{X} \in \mathbb{R}^{2N\times 2K}$ and the polynomial form of the drift distortion $\boldsymbol{W} \in \mathbb{R}^{2N\times 2M}$. As such, it is a function of the model parameters (see details in Appendix A-A).

B. Bayesian Inference

In the present model, the parameter vector is composed of the number of Lorentzian lineshapes $K \in [0, K_{\max}]$, the parameters (a_k, b_k, f_{0k}, Q_k) corresponding to the parameters for each Lorentzian function k, a parameter g_m for each drift m, the global noise variance σ^2, as well as two other (hyper-) parameters, Λ and δ^2, whose role will be detailed soon. In the following, we group parameters of the same kind belonging to different Lorentzian functions, setting $\boldsymbol{f}_0 = (f_{01}, \ldots, f_{0K})$ and $\boldsymbol{Q} = (Q_1, \ldots, Q_K)$. $\boldsymbol{\alpha}$ is defined as above.

According to Bayes' theorem, given the data $\boldsymbol{y}$, all model parameters $(K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2)$ can be drawn from a posterior probability density function $\mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2|\boldsymbol{y})$, which can be expressed as

$$\begin{aligned}&\mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2|\boldsymbol{y})\\ &\propto \mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2)\\ &\quad\times \mathrm{p}(\boldsymbol{y}|K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2). \end{aligned} \tag{6}$$

In this expression, $\mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2)$ is the prior distribution of the parameters. It represents the degree of belief that we have on the relevant values of the possible parameters before the data are acquired. $\mathrm{p}(\boldsymbol{y}|K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2)$ is the likelihood of the data. It describes how the data are generated according to the model and as a function of the model parameters. The sign "$\propto$" specifies that the left- and right-hand sides of the equation are proportional to each other.

1) Likelihood: According to the model given in Eq. (5), the data are related to the model parameters through the following distribution

$$\begin{aligned}&\mathrm{p}(\boldsymbol{y}|K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2)\\ &= \mathrm{p}(\boldsymbol{y}|K, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2)\\ &\propto (\sigma^2)^{-\frac{2N}{2}} \exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{\Delta}\boldsymbol{\alpha})^t(\boldsymbol{y} - \boldsymbol{\Delta}\boldsymbol{\alpha})\right]. \end{aligned} \tag{7}$$

Note that the dependence on $\boldsymbol{f}_0$ and $\boldsymbol{Q}$ is instantiated through $\boldsymbol{\Delta}$. Eq. (7) shows that $\boldsymbol{y}$ is independent of Λ and δ^2, but only conditionally on all other parameters, namely K, $\boldsymbol{\alpha}$, $\boldsymbol{f}_0$, $\boldsymbol{Q}$, and σ^2. A clearer way to see the relationship between $\boldsymbol{y}$ and (Λ, δ^2) is given by Fig. 1: Λ influences K, and δ^2 influences $\boldsymbol{\alpha}$, and both K and $\boldsymbol{\alpha}$ influence $\boldsymbol{y}$. As a consequence, $\boldsymbol{y}$ and (Λ, δ^2) are not marginally independent. This is why the posterior distributions $\mathrm{p}(\Lambda|\boldsymbol{y})$ and $\mathrm{p}(\delta^2|\boldsymbol{y})$ differ from their respective prior distributions $\mathrm{p}(\Lambda)$ and $\mathrm{p}(\delta^2)$ and provide information regarding the underlying generation process.

2) Priors: We assume that the prior distribution can be expanded as

$$\begin{aligned}&\mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2)\\ &= \mathrm{p}(\Lambda)\,\mathrm{p}(K|\Lambda)\,\mathrm{p}(\boldsymbol{f}_0|K)\,\mathrm{p}(\boldsymbol{Q}|K)\\ &\quad\times \mathrm{p}(\sigma^2)\,\mathrm{p}(\delta^2)\,\mathrm{p}(\boldsymbol{\alpha}|K, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2). \end{aligned} \tag{8}$$

We now have to specify all prior distributions. The prior for the number of components $\mathrm{p}(K|\Lambda)$ is set to a Poisson distribution with rate Λ, truncated to $\{0, \ldots, K_{\max}\}$ (see Appendix A-B). $K_{\max}$ is a heuristic parameter, which should be larger than the maximum number of expected resonances in a given dataset, e.g., twice of the expected number of resonances. The term Λ is related to the expected number of resonances; both quantities would be equal for $K_{\max} = \infty$. $\mathrm{p}(\Lambda)$ is set to a Gamma distribution with parameters $\frac{1}{2} + e_1$ and e_2.

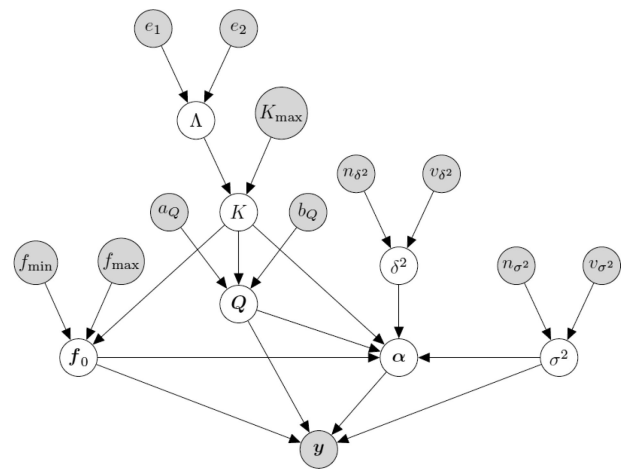


Fig. 1. Bayesian network summarizing the model. Gray nodes represent variables whose values are assumed to be known, while white nodes stand for parameters whose values need to be inferred. Arrows code for dependence. For a given node, the model gives the distribution of that node conditionally to its parents (i.e., nodes with an arrow pointing to that node).

We assume that the resonance frequencies are a priori independent,

$$\mathrm{p}(\boldsymbol{f}_0|K) = \prod_{k=1}^{K} \mathrm{p}(f_{0k}),$$

where the distribution for each resonance frequency f_{0k} is set to a uniform distribution on $[f_{\min}, f_{\max}]$. Similarly, the quality factors are assumed to be a priori independent

$$\mathrm{p}(\boldsymbol{Q}|K) = \prod_{k=1}^{K} \mathrm{p}(Q_k),$$

with a Gamma distributions with parameters a_Q and b_Q for each Q_k (see Appendix A-B).

The distribution for the noise variance $\mathrm{p}(\sigma^2)$ is set to a scaled inverse-chi-square distribution with n_{σ^2} degrees of freedom and squared scale v_{σ^2}. When $n_{\sigma^2} = v_{\sigma^2} = 0$, we obtain Jeffreys' uninformative prior $\mathrm{p}(\sigma^2) \propto 1/\sigma^2$.

Finally, the prior for $\boldsymbol{\alpha}$ is set as a Gaussian distribution with zero mean and covariance $\sigma^2\boldsymbol{\Psi}^{-1}$, where $\boldsymbol{\Psi} = \delta^{-2}\boldsymbol{\Delta}^t\boldsymbol{\Delta}$ (see Appendix A-B). $\mathrm{p}(\delta^2)$ is a scaled inverse-chi-square distribution with n_{δ^2} degrees of freedom and squared scale factor v_{δ^2}.

3) Graphical Model: The relationships between all variables are summarized by a Bayesian network in Fig. 1. The joint distribution of all variables can be read off from this network.

4) Integration of the Nuisance Parameters: The posterior distribution can be obtained by putting the likelihood and the priors distributions together into Bayes' theorem, Eq. (6). The

nuisance parameters $\boldsymbol{\alpha}$ and σ^2 can then be integrated out yielding (see Appendix B-A)

$$\begin{aligned}&\mathrm{p}(K,\Lambda,\delta^2,\boldsymbol{f}_0,\boldsymbol{Q}|\boldsymbol{y})\\&\propto \mathrm{p}(\Lambda)\,\mathrm{p}(K|\Lambda)\,\mathrm{p}(\boldsymbol{f}_0|K)\,\mathrm{p}(\boldsymbol{Q}|K)\,\mathrm{p}(\delta^2)\\&\quad\times(\delta^2+1)^{-\frac{2M+2K}{2}}\left(v_{\sigma^2}+\boldsymbol{y}^t\boldsymbol{\Pi}\boldsymbol{y}\right)^{-\frac{2N+n_{\sigma^2}}{2}}\\&\quad\times\mathbb{I}_{\{0,\ldots,K_{\max}\}}(K)\prod_{k=1}^{K}\mathbb{I}_{[f_{\min},f_{\max}]}(f_{0k}),\end{aligned}\tag{9}$$

for $K\in\{0,\ldots,K_{\max}\}$ and $f_{0k}\in[f_{\min},f_{\max}]$. In this expression, we set $\boldsymbol{\Pi}=\boldsymbol{I}-\boldsymbol{\Delta}\boldsymbol{\Gamma}^{-1}\boldsymbol{\Delta}^t$ and $\boldsymbol{\Gamma}=\boldsymbol{\Delta}^t\boldsymbol{\Delta}+\boldsymbol{\Psi}=(1+\delta^{-2})\boldsymbol{\Delta}^t\boldsymbol{\Delta}$. $\mathbb{I}$ comes from the prior for the frequencies; it is a way to express that all frequencies should be in $[f_{\min},f_{\max}]$, i.e., $\mathbb{I}_{[f_{\min},f_{\max}]}=1$.

C. Reversible-Jump Markov Chain Monte Carlo Method

Introduced in physics in the 1950s [28], Markov chain Monte Carlo (MCMC) methods have been extensively used to generate an approximate sample from a target distribution $\pi(\boldsymbol{x})$, usually a posterior distribution in the Bayesian community [23]. The two most popular MCMC methods are Metropolis-Hastings (MH) method and Gibbs sampler. The MH algorithm consists of two steps, a proposal step, where a new state $\boldsymbol{x}_{\mathrm{prop}}$ is proposed given the current state $\boldsymbol{x}^{(i)}$ according to a proposal probability $q(\boldsymbol{x}_{\mathrm{prop}}|\boldsymbol{x}^{(i)})$, and an acceptance step, where $\boldsymbol{x}_{\mathrm{prop}}$ is accepted, (i.e., we set $\boldsymbol{x}^{(i+1)}=\boldsymbol{x}_{\mathrm{prop}}$), with probability

$$A(\boldsymbol{x}_{\mathrm{prop}},\boldsymbol{x}^{(i)})=\min\left\{1,\frac{\pi(\boldsymbol{x}_{\mathrm{prop}})}{\pi(\boldsymbol{x}^{(i)})}\frac{q(\boldsymbol{x}^{(i)}|\boldsymbol{x}_{\mathrm{prop}})}{q(\boldsymbol{x}_{\mathrm{prop}}|\boldsymbol{x}^{(i)})}\right\};\tag{10}$$

otherwise, the current state is retained, $\boldsymbol{x}^{(i+1)}=\boldsymbol{x}^{(i)}$. The Gibbs sampler is a particular case of MH algorithm, which samples $\boldsymbol{x}=(x_1,\ldots,x_N)$ one component at a time according to the conditional distribution of that component given the remaining components, $\pi(x_j^{(i+1)}|x_1^{(i)},\ldots,x_{j-1}^{(i)},x_{j+1}^{(i)},\ldots,x_N^{(i)})$. In this case, the acceptance rate is equal to 1.

However, neither the MH sampler nor the Gibbs sampler can deal with variables of varying dimension, that is, Bayesian problems of joint model comparison and parameter estimation. In 1995, Green [9] proposed reversible jump MCMC (RJ-MCMC) to specifically deal with that kind of problem. In the specific case of our problem, we build from the RJ-MCMC method developed in [8]. Such an RJ-MCMC sampler is able to jump between subspaces with different dimensions, i.e., differing numbers of Lorentzian lineshapes, which allows us to sample directly from the joint distribution on $\boldsymbol{\Theta}=\bigcup_{k=0}^{K_{\max}}\{k\}\times\boldsymbol{\Theta}_{\boldsymbol{k}}$, where $\boldsymbol{\Theta}_{\boldsymbol{k}}$ stands for the set of parameters corresponding to a signal with k components. Following [8], [9], we applied three kinds of moves:

1) *Birth move:* random birth of a new resonance with frequency and Q factor sampled according to their respective prior distributions.
2) *Death move:* removal of a randomly selected existing resonance.
3) *Update:* update of all unknown parameters. As the model dimension is fixed in the update step, this can be performed by one step of conventional MCMC sampler.

Let $b(K,\Lambda)$, $d(K,\Lambda)$, and $u(K,\Lambda)$ be the probabilities for a birth move, a death move, and an update, respectively. The three probabilities are related by

$$b(K,\Lambda)+d(K,\Lambda)+u(K,\Lambda)=1.\tag{11}$$

When $K=0$, there can be no death, so $d(0,\Lambda)=0$. Similarly, when $K=K_{\max}$, there can be no birth, since the maximum number of components has already been reached, so $b(K_{\max},\Lambda)=0$. Otherwise, we set

$$\begin{aligned}b(K,\Lambda)&=c\,\min\left\{1,\frac{\mathrm{p}(K+1|\Lambda)}{\mathrm{p}(K|\Lambda)}\right\}\\&=c\,\min\left\{1,\frac{\Lambda}{K+1}\right\}\end{aligned}\tag{12a}$$

and

$$\begin{aligned}d(K,\Lambda)&=c\,\min\left\{1,\frac{\mathrm{p}(K-1|\Lambda)}{\mathrm{p}(K|\Lambda)}\right\}\\&=c\,\min\left\{1,\frac{K}{\Lambda}\right\},\end{aligned}\tag{12b}$$

where $\mathrm{p}(K|\Lambda)$ is the prior probability of K (see Section II-B2 and Eq. (13) in Appendix A-B). The choice of the form taken by the probabilities for birth and death moves follows [8], [9]; it ensures $b(K,\Lambda)\mathrm{p}(K|\Lambda)=d(K+1,\Lambda)\mathrm{p}(K+1|\Lambda)$, which would guarantee sure acceptance in the case of an MCMC sampler involving K alone and no data. Furthermore, $b(K,\Lambda)+d(K,\Lambda)\in[c,1]$ for all K and Λ, that is, the probability for a change in the number of components is always larger than c. In the present study, we followed [10] and set $c=0.2$.

For the update step, we followed the hybrid MCMC sampler proposed by Andrieu and Doucet [8], which combines Gibbs and Metropolis-Hasting (MH) steps. More specifically, we sequentially sample $\boldsymbol{f_0}$, $\boldsymbol{Q}$, Λ, and δ^2 from the marginal posterior distribution of each parameter conditioned on the three other parameters.

Details of the numerical sampling strategy can be found in Appendix C. Briefly, RJ-MCMC proceeds as follows:

1) *Initialization:* set $K^{[0]}$ and $\boldsymbol{\Theta_k}^{[0]}$. We use the initial guess of $K=0$, and the order of polynomial drift is empirically set based on the observation. If it is too high, the model will have to deal with unnecessary complexity; too low, the Lorentzian model will (wrongly) have to account for some high-order residual polynomial drift.
2) At iteration i, perform either a birth move, a death move or an update with probability $b(K^{[i-1]},\Lambda^{[i-1]})$, $d(K^{[i-1]},\Lambda^{[i-1]})$, and $u(K^{[i-1]},\Lambda^{[i-1]})$, respectively.
3) Increase i by 1 and go to 2).

III. Simulations

The algorithm was evaluated in terms of resonance detection as well as frequency and Q factor extraction. We considered a synthetic signal consisting of resonant components under

TABLE I
RJ-MCMC RESULTS OF THE SYNTHETIC SPECTRUM

No.	True Values				Estimates											
					$\sigma = 0.01$						$\sigma = 0.05$					
	f_{0k}	Q_k	A_k	ϕ_k	$\hat{f}_{0k}$	f_{0k} %std	f_{0k} %err	$\hat{Q}_k$	Q_k %std	Q_k %err	$\hat{f}_{0k}$	f_{0k} %std	f_{0k} %err	$\hat{Q}_k$	Q_k %std	Q_k %err
1	0.123	22	0.6	0	0.1230	0.07	−0.02	21.68	3.58	−1.47	0.1231	0.63	0.10	15.92	10.00	−27.64
2	0.127	23	0.6	90	0.1273	0.08	0.27	25.39	3.79	10.38	0.1267	0.51	−0.25	19.42	15.89	−15.58
3	0.145	24	0.4	0	0.1450	0.06	0.00	22.70	2.62	−5.42	0.1445	0.22	−0.37	24.03	6.51	0.13
4	0.162	26	0.6	180	0.1621	0.29	0.04	33.68	12.93	29.55	−−	−−	−−	−−	−−	−−
5	0.165	27	1	0	0.1647	0.18	−0.15	25.25	9.78	−6.47	0.1642	0.08	−0.47	43.10	4.43	59.64
6	0.178	24	0.5	180	0.1774	0.13	−0.32	22.22	4.71	−7.42	0.1777	0.52	−0.17	11.48	9.37	−52.18
7	0.185	25	1	0	0.1850	0.06	0.01	23.71	1.81	−5.15	0.1849	0.23	−0.07	25.36	10.60	1.46
8	0.205	25	0.4	0	0.2049	0.06	−0.04	23.13	1.95	−7.47	0.2048	0.30	−0.10	28.63	10.83	14.54
9	0.223	27	0.6	0	0.2235	0.20	0.23	24.90	7.71	−7.79	0.2232	0.58	0.09	24.24	36.10	−10.23
10	0.227	28	0.6	90	0.2263	0.19	−0.33	24.25	6.93	−13.38	0.2258	0.35	−0.53	29.55	67.05	5.54
11	0.245	25	0.4	0	0.2449	0.07	−0.06	24.96	2.31	−0.15	0.2454	0.43	0.15	29.08	31.07	16.31
12	0.262	26	1	144	0.2619	0.26	−0.04	28.69	14.62	10.35	0.2614	0.29	−0.24	17.10	4.49	−34.24
13	0.266	24	1	180	0.2656	0.33	−0.16	18.43	15.04	−23.19	−−	−−	−−	−−	−−	−−
14	0.273	22	1	198	0.2734	0.11	0.13	22.91	3.60	4.12	0.2738	0.29	0.28	24.83	8.46	12.88
15	0.305	21	0.2	90	0.3050	0.18	−0.01	19.91	6.66	−5.17	0.3079	0.69	0.96	14.09	7.30	−32.90
16	0.327	26	1	90	0.3270	0.14	0.01	24.45	5.09	−5.96	0.3278	0.17	0.25	22.28	3.19	−14.32
17	0.333	27	0.7	36	0.3334	0.22	0.12	36.35	19.68	34.62	−−	−−	−−	−−	−−	−−
18	0.345	24	1	180	0.3447	0.07	−0.09	25.08	3.27	4.49	0.3425	0.25	−0.73	20.77	4.78	−13.47
19	0.362	25	1.2	180	0.3629	0.18	0.26	29.25	7.38	17.01	0.3643	0.49	0.65	59.48	42.71	137.91
20	0.367	27	1.6	36	0.3671	0.41	0.04	32.13	37.67	19.00	0.3654	0.46	−0.44	69.19	73.14	156.25
21	0.374	28	1.8	90	0.3743	0.11	0.07	26.51	4.09	−5.32	0.3743	0.21	0.07	23.76	5.33	−15.14
Average							0.12			11.19			0.30			41.36

two different SNRs. Zero-mean Gaussian noises were added to the signal with standard deviations of $\sigma = 0.01$ and $\sigma = 0.05$. As shown in Table I, 21 resonant components were synthesized according to the Lorentzian model with different values of $(f_{0k}, Q_k, A_k, \phi_k), k \in \{1, \ldots, 21\}$. For the damping factors Q_k, an integer was arbitrarily chosen in the range [20, 30], which is considered as high attenuation. For instance, the Q factor of human cortical bone is in a range [20, 30] [12], [13] and that of sandstone in [30, 100] [15], [16]. To assess whether the method could successfully process spectra with overlapping peaks, the amplitudes A_k were set as many close values in a range of $\in [0.2, 1.8]$, such that a quite 'flat' spectrum was obtained (see Figs. 2(g) and 3(g)). The phase parameters ϕ_k were arbitrarily specified in a range $[0°, 180°]$. The polynomial baseline drift was assumed to have an order of 6, i.e., $M = 6$. The total number of iterations was set to 10,000 and the first 3,300 iterations were removed to account for burn-in period. We used the resulting sample to compute estimates for the resonant frequencies $\hat{f}_{0k}$ and damping factors $\hat{Q}_k$, as well as the percentages standard deviations (%std). The percentage root-mean-square errors (%err) were also computed between the estimates ($\hat{f}_{0k}$ and $\hat{Q}_k$) and true values (f_{0k} and Q_k), respectively.

The values used for the various parameters involved in the definition of the prior distributions are summarized in Table II.

Simulated results are shown in Fig. 2 ($\sigma = 0.01$) and Fig. 3 ($\sigma = 0.05$). The marginal posterior distributions $\mathrm{p}(k|\boldsymbol{y})$, $\mathrm{p}(\sigma|\boldsymbol{y})$, and $\mathrm{p}(\Lambda|\boldsymbol{y})$ are shown as histograms in Figs. 2(a)–(c) and 3(a)–(c), respectively. The amplitude spectra of the synthetic signal and of the reconstructed signal are plotted in Figs. 2(g) and 3(g). $\mathrm{p}(f_k|\boldsymbol{y})$ is presented in hologram in yellow color with a magnitude 10-time enlarged. Table I also lists the final estimates of each extracted frequency and Q factor as well as the corresponding true values.

Results from the signal with $\sigma = 0.01$ are depicted in Fig. 2. We used the definition $\mathrm{SNR} = 10\log_{10}(\frac{a_k^2 + b_k^2}{2\sigma^2})$. Accordingly,

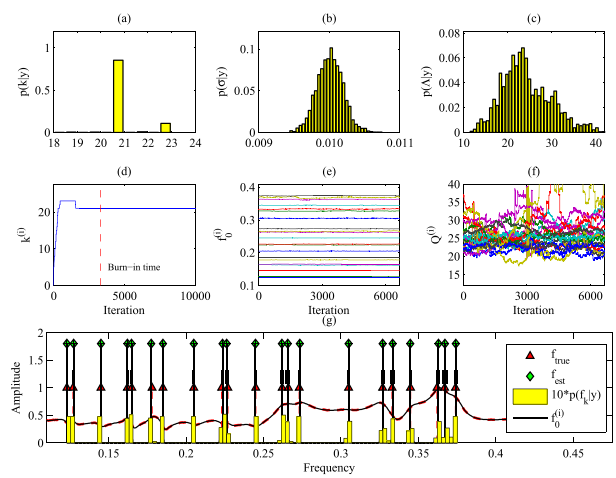


Fig. 2. RJ-MCMC results of simulation with $\sigma = 0.01$. (a)–(c) Posterior distributions of $\mathrm{p}(k|\boldsymbol{y})$, $\mathrm{p}(\sigma|\boldsymbol{y})$, and $\mathrm{p}(\Lambda|\boldsymbol{y})$. (d) Instantaneous estimates of $\mathrm{p}(k|\boldsymbol{y})$. (e)–(f) Instantaneous estimates of frequency and Q factor from the MAP $k = 21$. (g) Comparison between the true frequencies f_{true} ($\triangle$) and the final estimates of the frequencies f_{est} ($\diamond$) with corresponding error bars of standard deviation. The original and reconstructed amplitude spectra are plotted in black solid and red dash lines, respectively.

the SNRs of those Lorentzian lineshapes are in the range between 23.01 dB and 42.10 dB. The estimate of k is 21, in agreement with the true number of sinusoids. The estimate of σ is 0.010, again in accordance with the real value of the noise amplitude. $\mathrm{p}(\Lambda|\boldsymbol{y})$ is shown in Fig. 2(c) with a maximum value at 23.16. The instantaneous estimates of the posterior model $\mathrm{p}(k|\boldsymbol{y})$ is represented in Fig. 2(d). It was found that at the beginning of the burn-in period, the algorithm was able to jump among different model spaces without being trapped at a fixed number of sinusoids. The instantaneous estimates of frequency and Q factor corresponding to the maximum of $\mathrm{p}(k|\boldsymbol{y})$, i.e., $k = 21$, are shown in Fig. 2(e) and (f), respectively.

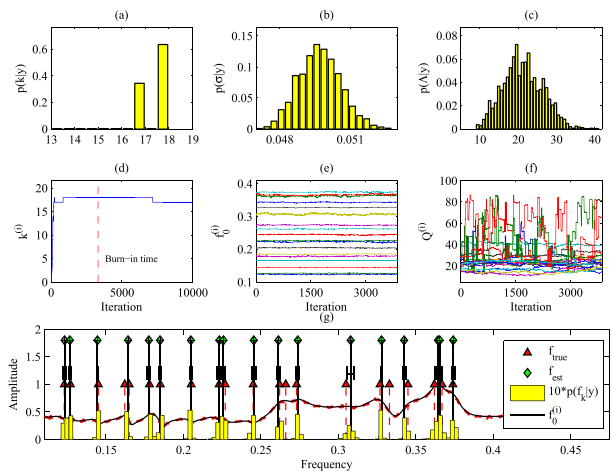


Fig. 3. RJ-MCMC results of the signal with $\sigma = 0.05$. (a)–(c) Posterior distributions of $\mathrm{p}(k|\boldsymbol{y})$, $\mathrm{p}(\sigma|\boldsymbol{y})$, and $\mathrm{p}(\Lambda|\boldsymbol{y})$. (d) Instantaneous estimates of $\mathrm{p}(k|\boldsymbol{y})$, (e)–(f) Instantaneous estimates of frequency and Q factor from the MAP $k = 18$. (g) Comparison between the true frequencies f_{true} ($\triangle$) and the final estimates of the frequencies f_{est} ($\diamond$) with corresponding error bars of standard deviation. The original and reconstructed amplitude spectra are plotted in black solid and red dash lines, respectively.

TABLE II
SUMMARY OF NUMERICAL VALUES USED FOR THE PARAMETERS OF THE PRIOR DISTRIBUTIONS

Parameter	Value
$K_{\max}$	60
e_1	0.1
e_2	0.1
$f_{\min}$	depending on experiment
$f_{\max}$	depending on experiment
a_Q	2
b_Q	24
n_{σ^2}	0
v_{σ^2}	0
n_{δ^2}	2
v_{δ^2}	100

When $\sigma = 0.05$, SNRs of the resonant components are between 9.03 dB and 28.12 dB, which are lower than that of the signals in Fig. 2. In Fig. 3(a), it was found that two possible values of k exist with $\mathrm{p}(k = 17|\boldsymbol{y}) = 0.344$ and $\mathrm{p}(k = 18|\boldsymbol{y}) = 0.636$, respectively. According to $\mathrm{p}(\sigma|\boldsymbol{y})$ plotted in Fig. 3(b), the estimated σ is equal to 0.050. $\mathrm{p}(\Lambda|\boldsymbol{y})$ is shown in Fig. 3(c) with a maximum value at 19.67. The model ambiguity can also been seen from the instantaneous estimates of $\mathrm{p}(k|\boldsymbol{y})$ in Fig. 3(d), where the estimates of $k = 17$ and 18 are both visited many times. In agreement with the maximum of $\mathrm{p}(k|\boldsymbol{y})$, 18 sets of estimates of frequencies and Q factors were chosen, with 3 undetected frequencies; the corresponding estimates are shown in Fig. 3(e) and (f), respectively.

The RJ-MCMC method can furthermore be used to analyze the confidence of each estimate. As an example, Fig. 4 compares the distributions of $\mathrm{p}(f_{01}|\boldsymbol{y})$, $\mathrm{p}(Q_1|\boldsymbol{y})$ and $\mathrm{p}(f_{01}, Q_1|\boldsymbol{y})$. When $\sigma = 0.01$ and 0.05, the pair $(\hat{f}_{01}, \hat{Q}_1)$ is given by (0.1230, 21.68) and (0.1231, 15.92), respectively. When $\sigma = 0.05$, the %std of $(\hat{f}_{01}, \hat{Q}_1)$ are larger than those when $\sigma = 0.01$. As shown in Table I, it was found that in both cases the %std of frequency

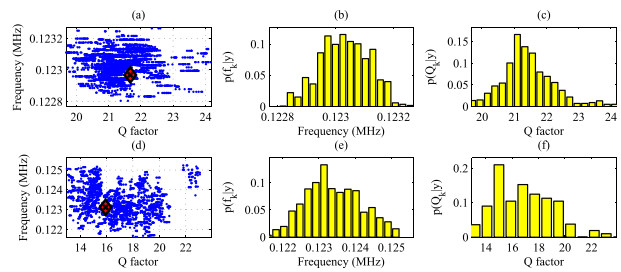


Fig. 4. Comparison of the marginal posterior distributions of f_{01} and Q_1. For $\sigma = 0.01$ (top) and $\sigma = 0.05$ (bottom), we represented the joint posterior distribution $\mathrm{p}(f_{01}, Q_1|\boldsymbol{y})$ (a and d), as well as the marginal posterior distribution $\mathrm{p}(f_{01}|\boldsymbol{y})$ (b and e) and $\mathrm{p}(Q_1|\boldsymbol{y})$ (c and f).

estimates are less than 1%; on the contrary, quite large %std of Q factor estimates were obtained, e.g., when $\sigma = 0.05$, the largest %std of Q factor estimate can be 73.14%. For final estimates of $\hat{\boldsymbol{f}}_0$ and $\hat{\boldsymbol{Q}}$, when $\sigma = 0.01$, the average values of %err are 0.12% and 11.19%, respectively; when $\sigma = 0.05$, those are 0.30% and 41.36%, respectively. In both cases, the %err of the detected frequency estimates are less than 1%.

IV. RUS EXPERIMENTS

In this section, we provide a face-to-face comparison between the frequencies and Q factors extracted using the RJ-MCMC method proposed in this study and previous results obtained using a non-automated method referred to as LPSVD which combines (1) a linear predictive (LP) filter in the time domain associated to an SVD step to select the number of resonances [15], [17], and (2) a non-linear fitting in frequency domain of the complex spectrum to refine the estimation of resonant frequencies and Q values [12]. In practice, the selection of the number of frequencies (model order) after the SVD step requires user interaction.

The experimental data involved [12] were typical acoustic resonance signals, measured by using the resonant ultrasound spectroscopy (RUS) method. RUS is a standard method in material science that uses tens of resonances at ultrasonic bandwidth from a millimeter-scale solid specimen to accurately determine the material elastic properties with high Q factors like metals (usually $Q > 100$) [11]. However, its performance for the characterization of highly attenuated media with low Q factor has long been suffering from the inaccurate frequency extraction of strong overlapping resonant peaks. We here present the RUS signals measured from highly attenuated specimens with Q factors around 20.

A typical RUS measurement consists of placing a rectangular parallelepipedic specimen between an emitting and a receiving ultrasonic transducer and measuring the frequency response of the specimen in a given bandwidth. A RUS spectrum appears to be a combination of several Lorentzian lineshapes, each one corresponding to a mechanical vibrational mode with a certain frequency. Besides, different coupling conditions between the specimen and transducers also lead to different amplitudes and phases of the signal for each mode, which finally results in different resonant spectra. In other words, the lineshapes vary from

one measurement to another after repositioning of the specimen between the transducers. However, in this process, the intrinsic mechanical resonance frequencies are not affected by the way they appear in the lineshapes. Here, we used data from a RUS experiment on a human cortical bone specimen, consisting of an ensemble of six spectra recorded after repositioning the specimen on the RUS setup. In the successive repositionings, the excitation, reception and coupling conditions are slightly modified, resulting in modifications of the relative amplitudes, phases, and noise levels of the spectra. While in principle each spectrum should contain all resonant frequencies within the measured bandwidth, only a number of them, with sufficiently large amplitudes, can actually be extracted. Compiling the information from the six spectra during signal processing is helpful for mode identification and extraction.

Our version of the RJ-MCMC algorithm was applied independently to each RUS spectrum. For each RJ-MCMC computation, parameters were set to values identical to those listed in Table II. In order to pool the results from the six independent RJ-MCMC computations, all frequency outputs were classified using a k-means clustering method. K-means is a clustering method that can automatically partition a dataset into clusters, in which each observation of the dataset belongs to the cluster with the nearest mean [18], [19]. It was here applied as a two-step heuristic post-processing procedure to reorganize the RJ-MCMC estimates from multichannel measurements. Two parameters, cluster size and standard deviation threshold, were introduced. First, chains with large standard deviation values were discarded. The remaining Markov chains were partitioned into N_c groups using k-means method. N_c number was set in accordance with the maximum of $\mathrm{p}(k|\boldsymbol{y})$ obtained from the RJ-MCMC results of all measured spectra. Second, an iterative clustering procedure was performed for those f_{0k} clusters with large numbers of estimates. This procedure stopped when the standard deviation of each new sub-clusters was found to be smaller than the standard deviation tolerance. The sub-clusters with a small number of estimates were further discarded. The final estimates were thus obtained.

As shown in Fig. 5(a)–(f), left, different numbers of frequency estimates were obtained from different spectra. On the right-hand of Fig. 5(a)–(f), a good match was found between the original amplitude spectra and the reconstructed ones. The removed frequencies were selected by the magnitude of %std with a threshold of 1%. After k-means clustering for the frequency estimates obtained from all spectra, the final estimates of $\boldsymbol{f_0}$ are presented in Fig. 5(g). The cluster-size threshold used in k-means step is 10,000.

Table III compares current estimates with results previously published. The estimated frequencies are compared to the theoretical frequencies f_{cal} calculated from the optimized elastic coefficients after solving the inverse problem. In a previous study [12], there were 30 calculated frequencies in the bandwidth 0.10–0.28 MHz, and 20 frequencies could be retrieved using the classical LPSVD method. The RJ-MCMC method enabled to automatically recover 21 frequencies. The average %err of the frequencies in the previous and current studies are 0.25% and 0.28%. The %std of the frequency estimates obtained by

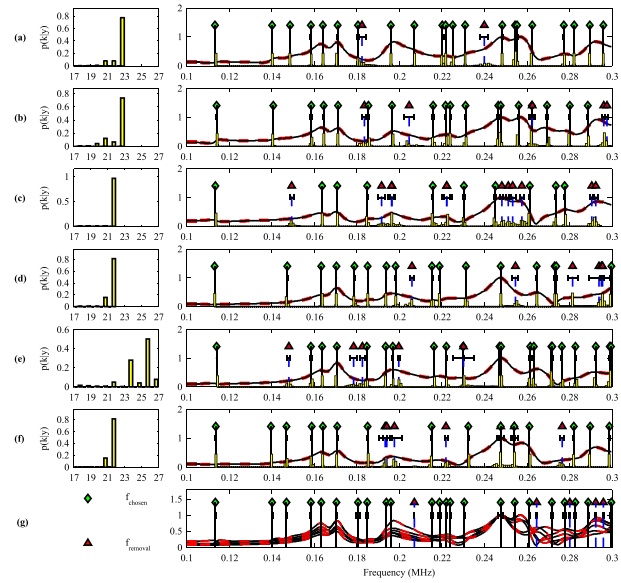


Fig. 5. RJ-MCMC analysis for 6 spectra measured from a human cortical bone specimen [12]. On the left-hand of (a)–(f), $\mathrm{p}(k|\boldsymbol{y})$ obtained from the 6 different spectra are plotted; on the right-hand of (a)–(f), the original and reconstructed amplitude spectra of each signal are depicted in black solid and red dash lines. (g) Final frequency estimates obtained after k-means clustering. The final estimates, i.e., f_{chosen}, are marked in green '$\diamond$'; discarded frequency estimates, i.e., f_{removal}, are marked in red '$\triangle$'. Error bars show the standard deviations for estimates of all frequencies. The $\mathrm{p}(f_k|\boldsymbol{y})$ is presented in hologram in yellow color, but with a magnitude 10-time enlarged.

RJ-MCMC method are less than 1%. Because there is actually no efficient way to know the theoretical Q factors, $\hat{Q}_k$ factor can only be compared to the Q'_k. The average %err of the Q factors between the two studies is 5.72%.

V. Discussion

The accurate estimation of resonant frequencies and attenuation factors in a damped spectrum is a signal processing problem encountered in many applications. Relying on the work of [8], we designed a practical Bayesian framework which allows, in a spectrum, for an automated retrieval of the number of resonances as well as the parameters defining the lineshape of each resonance.

The RJ-MCMC strategy proposed by Andrieu and Doucet [8] can be used to extract the resonant components, but their method cannot deal with complex signal, cannot remove the baseline distortion, and cannot estimate the full set of resonance parameters, such as the Q factors. In response to these limitation, we presented an improved Bayesian framework to automatically and jointly estimate the number of resonances as well as their frequencies and Q factors. Specifically, we further expanded [8]'s approach by considering a more complex model that (i) deals with a complex signal, (ii) takes the attenuation parameter, i.e., the Q factor, into account, (iii) introduces the polynomial terms in the model to fit the baseline distortion, (iv) applies a k-means clustering method to reorganize the RJ-MCMC estimates obtained from a multichannel measurement. The proposed method was used to process highly attenuated and

TABLE III
COMPARISON OF THE RJ-MCMC RESULTS WITH PREVIOUS RESULTS [12]

No.	Previous results				RJ-MCMC results				Deviation	
	f_{calc}	f'_{0k}	Q'_k	f'_{0k} %err	$\hat{f}_{0k}$	$\hat{f}_{0k}$ %std	$\hat{Q}_k$	$\hat{Q}_k$ %std	$\frac{\hat{f}_{0k}-f_{\mathrm{calc}}}{f_{\mathrm{calc}}}$ %err	$\frac{Q'_k-\hat{Q}_k}{Q'_k}$ %err
1	0.11350	0.11367	25.8	−0.15	0.11336	0.28	26.3	17.80	0.28	−1.79
2	0.13997	0.14019	24.2	−0.16	0.14018	0.17	24.4	2.44	0.01	−0.85
3	0.14811	0.14875	21.8	−0.43	0.14686	0.12	23.3	7.57	1.27	−6.97
4	0.15914	0.15901	22.6	0.08	0.15828	0.21	20.8	13.68	0.46	8.19
5	0.16434	0.16375	21.6	0.36	0.16323	0.01	21.8	0.47	0.32	−0.92
6	0.17100	0.17075	28.2	0.14	0.17044	0.03	27.9	0.54	0.18	1.04
7	0.17811	0.17825	24.8	−0.08	−−	−−	−−	−−	−−	−−
8	0.18098	0.18124	25.5	−0.14	0.18053	0.63	27.5	17.45	0.39	−7.66
9	0.18508	0.18535	29.9	−0.14	0.18494	0.24	26.9	26.22	0.22	9.89
10	0.19220	−−	−−	−−	−−	−−	−−	−−	−−	−−
11	0.19402	0.19386	24.9	0.08	0.19383	0.08	25.9	14.00	0.02	−4.14
12	0.19759	0.19719	24.8	0.21	0.19605	0.05	21.7	10.65	0.58	12.32
13	0.20676	0.20631	13.6	0.22	−−	−−	−−	−−	−−	−−
14	0.21432	0.21505	24.8	−0.34	0.21537	0.16	20.9	12.73	−0.15	15.75
15	0.21930	−−	−−	−−	0.21884	0.28	29.8	48.02	−−	−−
16	0.22035	−−	−−	−−	0.22204	0.14	90.0	27.95	−−	−−
17	0.22248	0.223	27.3	−0.23	0.22405	0.07	19.8	8.56	−0.47	27.54
18	0.22949	0.2312	23.8	−0.74	0.23055	0.12	26.2	12.58	0.28	−10.08
19	0.23640	−−	−−	−−	−−	−−	−−	−−	−−	−−
20	0.24348	−−	−−	−−	−−	−−	−−	−−	−−	−−
21	0.24504	−−	−−	−−	−−	−−	−−	−−	−−	−−
22	0.24741	0.24777	24.4	−0.14	0.24758	0.46	25.9	20.89	0.08	−6.33
23	0.24965	−−	−−	−−	−−	−−	−−	−−	−−	−−
24	0.25022	−−	−−	−−	−−	−−	−−	−−	−−	−−
25	0.25308	−−	−−	−−	0.25422	0.11	19.7	21.04	−−	−−
26	0.25771	0.25622	21.2	0.58	−−	−−	−−	−−	−−	−−
27	0.26364	0.26300	24.6	0.24	0.26119	0.09	20.7	58.78	0.69	15.69
28	0.26944	−−	−−	−−	0.27177	0.26	29.3	99.76	−−	−−
29	0.27413	0.27324	21.2	0.33	0.27777	0.19	18.8	57.90	−1.66	11.47
30	0.28001	0.27920	33	0.29	0.28276	0.18	22.8	53.85	−1.27	30.83
Average				0.25					0.28	5.72

Theoretically calculated frequencies (f_{calc}) are given as standard values. We compare the extracted experimental frequency using LPSVD method (f'_{0k}, Q'_k) in our previous study and current RJ-MCMC estimates ($\hat{f}_{0k}, \hat{Q}_k$).

complex-valued resonance spectra. A good estimation of the full set of resonance parameters, i.e., the model order (number of resonances) and the unknown model parameters (amplitudes, resonant frequencies, damping coefficients and noise variance), was achieved.

The algorithm was tested using synthetic spectra and spectra from a resonant ultrasound spectroscopy experiment on a bone specimen. The performance of the algorithm was demonstrated by comparing the frequencies and Q factors estimated with RJ-MCMC to reference: known values in the case of the synthetic signals and estimated values obtained by a trained user with manual processing and theoretical validation after solving the inverse problem in the case of experimental signals. We found that the discrepancy between the reference values and the frequency estimates using the RJ-MCMC method was less than 1%. On the other hand, as shown in Table I, with a high SNR of 20 dB to 40 dB ($\sigma = 0.01$), the Q factors could be estimated with an average relative error of 11.19% and 8.76% of deviation; when the SNR was lower and in a range of 10 dB to 20 dB ($\sigma = 0.05$), the average relative error of Q factor estimates could be 40.36% with 23.42% variation. According to the Lorentzian model in Eqs. (1)–(4), the spectrum shape can be influenced by frequency, damping factor, amplitude, and phase. With only one given spectrum consisting of tens of resonances, the solution of the Lorentzian model parameters is not unique with high uncertainty. The results indicated that the uncertainty of Q factor estimates was larger than those of the frequencies.

The RJ-MCMC method as used in the present work requires to define several parameters (Fig. 1 and Table II), which include $K_{\max}$, (e_1, e_2), (a_Q, b_Q), $(n_{\delta^2}, v_{\delta^2})$, and $(n_{\sigma^2}, v_{\sigma^2})$. $K_{\max}$ can be initially set to a number much larger than that of the expected number of resonances. The prior probability for Λ was set to an uninformative conjugate prior with small values for $(e_1, e_2) = (0.1, 0.1)$. Depending on the resonance systems, (a_Q, b_Q) can be readily adjusted to ensure an efficient sampling of the Q factors in a certain range of interest. In our RUS data with high attenuation, the Q factors of different resonances were always less than 100, so that we set $(a_Q, b_Q) = (2, 24)$. Andrieu and Doucet [8] discussed the specification of $(n_{\delta^2}, v_{\delta^2})$. $n_{\delta^2} = 2$ ensures an infinite variance and there is a weak influence of the choice of v_{δ^2} on the final results. We fixed $(n_{\delta^2}, v_{\delta^2}) = (2, 100)$. The results experimentally indicate that the performance of the algorithm is insensitive to the specification of these parameters.

The RJ-MCMC strategy is very robust and flexible. We implemented three kinds of moves in the study. Some more sophisticated moves, named split and merge, could also be considered, which could be of interest to deal with situations where one spectrum peak contains two or several resonances with quite close frequencies [20]. In our practice, it appeared that the three moves already implemented here were efficient enough. The additional split and merge moves will increase the computation cost, but might not significantly improve the performance of the algorithm.

The algorithm was coded using Matlab. The computation was performed on a laptop computer (Intel Core i7-5500U, 2.4 GHz). Each processing required 300 s on average. We set the iteration and burn-in period to 10,000 and 3,300, respectively. A smaller iteration number could be considered. In addition, parallel computation strategies could be applied.

When processing multiple signals with different numbers of resonances, the posterior distributions $\mathrm{p}(k|\boldsymbol{y})$ and $\mathrm{p}(f_{0k}|\boldsymbol{y})$ obtained from different spectra can be quite different. A clustering method was adopted to pool the results obtained from different measurements. In fact, the Bayesian framework offers great flexibility of the model, which could possibly be revised for simultaneous processing of multiple spectra. Limited by the space, we do not generalize the algorithm to multichannel spectra analysis, which we hope will be clarified in a further study.

The proposed method provides an efficient solution to analyse highly-damped spectra. One application of the method is the processing of RUS signals where elastic properties of a solid are deduced from the resonance frequencies of a sample. Currently, the application of RUS to measure materials with low Q factors (biological hard tissues, reinforced plastic materials, rocks *etc.*) is limited due to the difficulty to extract a sufficient number of resonant frequencies from the RUS spectrum in a reliable manner [12], [13], [16]. The proposed RJ-MCMC method is a mean to process RUS signals automatically (no trial-and- error process with a trained user) and provides a degree of confidence for the estimated signal parameters.

VI. Conclusion

The proposed Lorentzian model based automatic RJ-MCMC Bayesian method provides an efficient way for resonant signal processing. Tens of resonant frequencies can be simultaneously retrieved with damping factors from the posterior distributions of the Markov chains. The algorithm can facilitate automatic processing of highly attenuated resonant signal without a predefined number of resonances and the intervention of a trained operator.

Appendix A
Model Specification

A. Vector-Matrix Definition of the Lorentzian Model

The matrix $\boldsymbol{\Delta} = [\boldsymbol{W}\boldsymbol{X}] \in \mathbb{R}^{2N(2M+2K)}$ consists of a polynomial basis matrix $\mathbf{W}$ and a Lorentzian model basis matrix $\mathbf{X}$. The $2N$-by-$2M$ matrix $\mathbf{W}$ is given by

$$\mathbf{W} = \left[\begin{array}{c|c} \mathbf{R} & 0 \\ \hline 0 & \mathbf{S} \end{array}\right],$$

with the N-by-M matrix

$$\mathbf{R} = \begin{bmatrix} \Re[W_1(f_1)] & \cdots & \Re[W_M(f_1)] \\ \vdots & & \vdots \\ \Re[W_1(f_N)] & \cdots & \Re[W_M(f_N)] \end{bmatrix},$$

and the N-by-M matrix

$$\mathbf{S} = \begin{bmatrix} \Im[W_1(f_1)] & \cdots & \Im[W_M(f_1)] \\ \vdots & & \vdots \\ \Im[W_1(f_N)] & \cdots & \Im[W_M(f_N)] \end{bmatrix}.$$

The $2N$-by-$2K$ matrix $\boldsymbol{X}$ is designed as

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{L}(f_{01}, Q_1) & \cdots & \boldsymbol{L}(f_{0K}, Q_K) \end{bmatrix},$$

where each $2N$-by-2 matrix $\boldsymbol{L}(f_{0k}, Q_k)$ of the form

$$\boldsymbol{L}(f_{0k}, Q_k) = \begin{bmatrix} J_1(f_1; f_{0k}, Q_k) & J_2(f_1; f_{0k}, Q_k) \\ \vdots & \vdots \\ J_1(f_N; f_{0k}, Q_k) & J_2(f_N; f_{0k}, Q_k) \\ -J_2(f_1; f_{0k}, Q_k) & J_1(f_1; f_{0k}, Q_k) \\ \vdots & \vdots \\ -J_2(f_1; f_{0k}, Q_k) & J_1(f_1; f_{0k}, Q_k) \end{bmatrix},$$

with

$$J_1(f; f_{0k}, Q_k) = \frac{f_{0k}^2 - f^2}{(f_{0k}^2 - f^2)^2 + \left(\frac{f f_{0k}}{Q_k}\right)^2}$$

$$J_2(f; f_{0k}, Q_k) = \frac{\frac{f f_{0k}}{Q_k}}{(f_{0k}^2 - f^2)^2 + \left(\frac{f f_{0k}}{Q_k}\right)^2}.$$

B. Prior Distributions

We set the prior distributions for K, Λ, δ^2, $\boldsymbol{f}_0$, and $\boldsymbol{\alpha}$ following [8]. For $\boldsymbol{Q}$, we use a Gamma distribution, which is convenient to provide vague information about a positive parameter.

The prior distribution for the number of components K is set to a Poisson distribution with rate Λ, truncated to $\{0, \ldots, K_{\max}\}$

$$\begin{aligned} \mathrm{p}(K|\Lambda) &\propto \mathrm{Poisson}(K|\Lambda), \qquad K \in \{0, \ldots, K_{\max}\} \\ &\propto \frac{\Lambda^K}{K!} e^{-\Lambda} \\ &= \frac{\frac{\Lambda^K}{K!} e^{-\Lambda}}{\sum_{K=0}^{K_{\max}} \frac{\Lambda^K}{K!} e^{-\Lambda}} \mathbb{I}_{\{0,\ldots,K_{\max}\}}(K), \end{aligned} \tag{13}$$

where $\mathbb{I}(K)$ is equal to 1 for $K \in \{0, \ldots, K_{\max}\}$ and to 0 otherwise. The prior distribution for Λ is set to a Gamma distribution with parameters $\frac{1}{2} + e_1$ and e_2

$$\begin{aligned} \mathrm{p}(\Lambda) &= \mathrm{Gamma}\left(\Lambda \middle| \frac{1}{2} + e_1, e_2\right) \\ &= \frac{e_2^{\frac{1}{2}+e_1}}{\Gamma(\frac{1}{2} + e_1)} \Lambda^{\frac{1}{2}+e_1-1} \exp(-e_2\Lambda), \qquad \Lambda > 0. \end{aligned} \tag{14}$$

The prior distribution for each resonance frequency f_{0k} is set to a uniform distribution on $[f_{\min}, f_{\max}]$

$$\mathrm{p}(f_{0k}) = \frac{1}{f_{\max} - f_{\min}}, \qquad f_{0k} \in [f_{\min}, f_{\max}].$$

The prior distribution for each quality factor Q_k is set to a Gamma distribution with parameters a_Q and b_Q

$$\begin{aligned}\mathrm{p}(Q_k) &= \mathrm{Gamma}\left(Q_k|a_Q, b_Q\right)\\ &= \frac{b_Q^{a_Q}}{\Gamma(a_Q)}\, Q_k^{a_Q-1}\, \exp(-b_Q Q_k), \qquad Q_k > 0. \end{aligned} \tag{15}$$

The prior distribution for $\boldsymbol{\alpha}$ is set as a Gaussian distribution with zero mean and covariance $\sigma^2\boldsymbol{\Psi}^{-1}$

$$\begin{aligned}\mathrm{p}(\boldsymbol{\alpha}|K, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2) &= \mathcal{N}\left(\boldsymbol{\alpha}|\mathbf{0}, \sigma^2\boldsymbol{\Psi}^{-1}\right)\\ &= (2\pi\sigma^2)^{-\frac{2M+2K}{2}}\, |\boldsymbol{\Psi}|^{\frac{1}{2}}\, \exp\left(-\frac{\boldsymbol{\alpha}^t\boldsymbol{\Psi}\boldsymbol{\alpha}}{2\sigma^2}\right),\end{aligned} \tag{16}$$

where $\boldsymbol{\Psi} = \delta^{-2}\boldsymbol{\Delta}^t\boldsymbol{\Delta}$. Such a prior distribution corresponds to a g-prior distribution [21]. $\mathrm{p}(\delta^2)$ is set to a scaled inverse-chi-square distribution with n_{δ^2} degrees of freedom and squared scale factor v_{δ^2}

$$\begin{aligned}\mathrm{p}(\delta^2) &= \text{Scale-inv-}\chi^2\left(\delta^2|n_{\delta^2}, v_{\delta^2}\right)\\ &= \frac{\left(\frac{n_{\delta^2}v_{\delta^2}}{2}\right)^{\frac{v_{\delta^2}}{2}}}{\Gamma\left(\frac{v_{\delta^2}}{2}\right)}\frac{\exp\left(-\frac{n_{\delta^2}v_{\delta^2}}{2\delta^2}\right)}{\left(\delta^2\right)^{\left(\frac{n_{\delta^2}}{2}+1\right)}}.\end{aligned} \tag{17}$$

Finally, the distribution for the noise variance $\mathrm{p}(\sigma^2)$ is set to a scaled inverse-chi-square distribution with n_{σ^2} degrees of freedom and squared scale v_{σ^2}

$$\begin{aligned}\mathrm{p}(\sigma^2) &= \text{Scale-inv-}\chi^2\left(\sigma^2|n_{\sigma^2}, v_{\sigma^2}\right)\\ &= \frac{\left(\frac{n_{\sigma^2}v_{\sigma^2}}{2}\right)^{\frac{v_{\sigma^2}}{2}}}{\Gamma\left(\frac{v_{\sigma^2}}{2}\right)}\frac{\exp\left(-\frac{n_{\sigma^2}v_{\sigma^2}}{2\sigma^2}\right)}{\left(\sigma^2\right)^{\left(\frac{n_{\sigma^2}}{2}+1\right)}}.\end{aligned} \tag{18}$$

APPENDIX B
INFERENCE

A. Integration of the Nuisance Parameters

Substituting the detailed expressions of the various prior distributions in Eq. (8) and incorporating them together with the likelihood of Eq. (7) into Bayes' theorem, Eq. (6), the full posterior distribution reads

$$\begin{aligned}&\mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2|\boldsymbol{y})\\ &\propto \mathrm{p}(\Lambda)\,\mathrm{p}(K|\Lambda)\,\mathrm{p}(\boldsymbol{f}_0|K)\,\mathrm{p}(\boldsymbol{Q}|K)\,\mathrm{p}(\delta^2)\,(2\pi)^{-\frac{2N+2M+2K}{2}}\\ &\quad\times\left(\sigma^2\right)^{-\left(\frac{2N+2M+2K+n_{\sigma^2}}{2}+1\right)}\exp\left(-\frac{v_{\sigma^2}}{2\sigma^2}\right)\\ &\quad\times|\boldsymbol{\Psi}|^{\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^2}\left[(\boldsymbol{\alpha}-\hat{\boldsymbol{\alpha}})^t\boldsymbol{\Gamma}(\boldsymbol{\alpha}-\hat{\boldsymbol{\alpha}}) + \boldsymbol{y}^t\boldsymbol{\Pi}\boldsymbol{y}\right]\right\}\\ &\quad\times\mathbb{I}_{\{0,\ldots,K_{\max}\}}(K)\prod_{k=1}^{K}\mathbb{I}_{[f_{\min},f_{\max}]}(f_{0k})\end{aligned} \tag{19}$$

where $\boldsymbol{\Gamma} = \boldsymbol{\Delta}^t\boldsymbol{\Delta} + \boldsymbol{\Psi} = (1+\delta^{-2})\boldsymbol{\Delta}^t\boldsymbol{\Delta}$, $\hat{\boldsymbol{\alpha}} = \boldsymbol{\Gamma}^{-1}\boldsymbol{\Delta}^t\boldsymbol{y}$, and $\boldsymbol{\Pi} = \boldsymbol{I} - \boldsymbol{\Delta}\boldsymbol{\Gamma}^{-1}\boldsymbol{\Delta}^t$.

While being parameters of the Lorentzian model, $\boldsymbol{\alpha}$ and σ^2 are not of direct interest for resonance spectrum analysis. We therefore integrate out these so-called nuisance parameters. As a function of $\boldsymbol{\alpha}$, the above distribution is proportional to a multivariate Gaussian distribution with mean $\hat{\boldsymbol{\alpha}}$ and covariance matrix $\sigma^2\boldsymbol{\Gamma}^{-1}$. Integration with respect to $\boldsymbol{\alpha}$ therefore implies to multiply by $(2\pi\sigma^2)^{\frac{2M+2K}{2}}\,|\boldsymbol{\Gamma}|^{-\frac{1}{2}}$, yielding

$$\begin{aligned}&\mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2|\boldsymbol{y})\\ &\propto \mathrm{p}(\Lambda)\,\mathrm{p}(K|\Lambda)\,\mathrm{p}(\boldsymbol{f}_0|K)\,\mathrm{p}(\boldsymbol{Q}|K)\,\mathrm{p}(\delta^2)\\ &\quad\times\left(\sigma^2\right)^{-\left(\frac{2N+n_{\sigma^2}}{2}+1\right)}\exp\left(-\frac{v_{\sigma^2}+\boldsymbol{y}^t\boldsymbol{\Pi}\boldsymbol{y}}{2\sigma^2}\right)\\ &\quad\times\left(\delta^2+1\right)^{-\frac{2M+2K}{2}}\mathbb{I}_{\{0,\ldots,K_{\max}\}}(K)\prod_{k=1}^{K}\mathbb{I}_{[f_{\min},f_{\max}]}(f_{0k}).\end{aligned} \tag{20}$$

As a function of σ^2, this previous distribution is proportional to a scaled inverse-chi-square distribution with $2N + n_{\sigma^2}$ degrees of freedom and squared scale $v_{\sigma^2} + \boldsymbol{y}^t\boldsymbol{\Pi}\boldsymbol{y}$. We integrate with respect to σ^2, yielding

$$\begin{aligned}\mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}|\boldsymbol{y}) &\propto \mathrm{p}(\Lambda)\,\mathrm{p}(K|\Lambda)\,\mathrm{p}(\boldsymbol{f}_0|K)\,\mathrm{p}(\boldsymbol{Q}|K)\,\mathrm{p}(\delta^2)\\ &\quad\times\left(\delta^2+1\right)^{-\frac{2M+2K}{2}}\\ &\quad\times\left(v_{\sigma^2}+\boldsymbol{y}^t\boldsymbol{\Pi}\boldsymbol{y}\right)^{-\frac{2N+n_{\sigma^2}}{2}}\\ &\quad\times\mathbb{I}_{\{0,\ldots,K_{\max}\}}(K)\prod_{k=1}^{K}\mathbb{I}_{[f_{\min},f_{\max}]}(f_{0k}).\end{aligned} \tag{21}$$

B. Conditional Posterior Distributions of σ^2 and $\boldsymbol{\alpha}$

Following Eq. (20), the conditional posterior distributions for σ^2 is given by

$$\begin{aligned}\mathrm{p}(\sigma^2|\boldsymbol{y}, K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}) &\propto \mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2|\boldsymbol{y})\\ &\propto \left(\sigma^2\right)^{-\left(\frac{2N+n_{\sigma^2}}{2}+1\right)}\\ &\quad\times\exp\left(-\frac{v_{\sigma^2}+\boldsymbol{y}^t\boldsymbol{\Pi}\boldsymbol{y}}{2\sigma^2}\right),\end{aligned} \tag{22}$$

which means that σ^2 conditionally follows a scaled inverse-chi-square distribution with $2N + n_{\sigma^2}$ degrees of freedom and squared scale $(v_{\sigma^2} + \boldsymbol{y}^t\boldsymbol{\Pi}\boldsymbol{y})/(2N + n_{\sigma^2})$. Then, according to Eq. (19), the conditional posterior distributions for $\boldsymbol{\alpha}$ is given by

$$\begin{aligned}\mathrm{p}(\boldsymbol{\alpha}|\boldsymbol{y}, K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2) &\propto \mathrm{p}(K, \Lambda, \delta^2, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2|\boldsymbol{y})\\ &\propto \exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{\alpha}-\hat{\boldsymbol{\alpha}})^t\boldsymbol{\Gamma}(\boldsymbol{\alpha}-\hat{\boldsymbol{\alpha}})\right],\end{aligned} \tag{23}$$

which implies that $\boldsymbol{\alpha}$ conditionally follows a multivariate Gaussian distribution with mean $\hat{\boldsymbol{\alpha}}$ and covariance matrix $\sigma^2\boldsymbol{\Gamma}^{-1}$.

APPENDIX C
NUMERICAL SAMPLING SCHEME

A. Birth Move

- Propose a new component with a frequency f_{prop} and quality factor Q_{prop} according to the priors $q_f(f_{\text{prop}})$ and $q_Q(Q_{\text{prop}})$, respectively.

$$q_f(f_{\text{prop}}) = \frac{1}{f_{\max} - f_{\min}} \mathbb{I}_{[f_{\min}, f_{\max}]}(f_{\text{prop}})$$
$$q_Q(Q_{\text{prop}}) = \text{Gamma}(Q_{\text{prop}}|a_Q, b_Q).$$

- Insert the new component at location k, and define the new frequency vector as $\widetilde{\boldsymbol{f}}_0 = \boldsymbol{f}_0 \oplus_k f_{\text{prop}}$ and the new vector of quality factors as $\widetilde{\boldsymbol{Q}} = \boldsymbol{Q} \oplus_k Q_{\text{prop}}$. The sign "$\oplus$" specifies the insertion of the proposal component into the vector on the left-hand side.
- The new state $(K+1, \Lambda, \delta^2, \widetilde{\boldsymbol{f}}_0, \widetilde{\boldsymbol{Q}}|\boldsymbol{y})$ is kept with probability $a_{\text{birth}} = \min\{1, r_{\text{birth}}\}$ [22], [23];

$$r_{\text{birth}} = \frac{\text{p}(K+1, \Lambda, \delta^2, \widetilde{\boldsymbol{f}}_0, \widetilde{\boldsymbol{Q}}|\boldsymbol{y})}{\text{p}(K, \Lambda, \delta^2, \widetilde{\boldsymbol{f}}_0, \widetilde{\boldsymbol{Q}}|\boldsymbol{y})} \frac{d(K+1, \Lambda)}{b(K, \Lambda)} \times \frac{\frac{1}{K+1}}{\frac{1}{K+1} q_f(f_{\text{prop}})\, q_Q(Q_{\text{prop}})}. \tag{24}$$

According to Eq. (21), the ratio of posterior distributions can be simplified as

$$\frac{\text{p}(K+1, \Lambda, \delta^2, \widetilde{\boldsymbol{f}}_0, \widetilde{\boldsymbol{Q}}|\boldsymbol{y})}{\text{p}(K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}|\boldsymbol{y})} = \frac{\Lambda}{K+1} \frac{1}{f_{\max} - f_{\min}} \frac{\text{p}(\widetilde{\boldsymbol{Q}}|K+1)}{\text{p}(\boldsymbol{Q}|K)} \times \frac{1}{\delta^2+1} \left(\frac{v_{\sigma^2} + \boldsymbol{y}^{\text{T}} \widetilde{\boldsymbol{\Pi}} \boldsymbol{y}}{v_{\sigma^2} + \boldsymbol{y}^{\text{T}} \boldsymbol{\Pi} \boldsymbol{y}} \right)^{-\frac{2N+n_{\sigma^2}}{2}}. \tag{25}$$

The ratio of death to birth is [8]

$$\frac{d(K+1, \Lambda)}{b(K, \Lambda)} = \frac{\text{p}(K+1|\Lambda)}{\text{p}(K|\Lambda)} = \frac{K+1}{\Lambda}.$$

Finally, the acceptance ratio is

$$r_{\text{birth}} = \frac{1}{\delta^2+1} \left(\frac{v_{\sigma^2} + \boldsymbol{y}^{\text{T}} \widetilde{\boldsymbol{\Pi}} \boldsymbol{y}}{v_{\sigma^2} + \boldsymbol{y}^{\text{T}} \boldsymbol{\Pi} \boldsymbol{y}} \right)^{-\frac{2N+n_{\sigma^2}}{2}}. \tag{26}$$

B. Death Move

- Select one component among the K existing ones with frequency f_{0k} and quality factor Q_k and define $\widetilde{\boldsymbol{f}}_0 = \boldsymbol{f}_{0 \ominus k}$ as the new frequency vector and $\widetilde{\boldsymbol{Q}} = \boldsymbol{Q}_{\ominus k}$ as the new vector of quality factors. The sign "$\ominus$" specifies the removal of the selected component from the vector on the left-hand side.
- The new state $(K-1, \Lambda, \delta^2, \widetilde{\boldsymbol{f}}_0, \widetilde{\boldsymbol{Q}}|\boldsymbol{y})$ is kept with probability $a_{\text{death}} = \min\{1, r_{\text{death}}\}$ [22], [23];

$$r_{\text{death}} = \frac{\text{p}(K-1, \Lambda, \delta^2, \widetilde{\boldsymbol{f}}_0, \widetilde{\boldsymbol{Q}}|\boldsymbol{y})}{\text{p}(K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}|\boldsymbol{y})} \times \frac{b(K-1, \Lambda)}{d(K, \Lambda)} \frac{\frac{1}{K+1}}{\frac{1}{K+1} q_f(f_{\text{prop}}) q_Q(Q_{\text{prop}})}. \tag{27}$$

As for the birth move, the ratio of posterior distributions can be simplified using Eq. (21), yielding

$$\frac{\text{p}(K-1, \Lambda, \delta^2, \widetilde{\boldsymbol{f}}_0, \widetilde{\boldsymbol{Q}}|\boldsymbol{y})}{\text{p}(K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}|\boldsymbol{y})} = \frac{K}{\Lambda} (f_{\max} - f_{\min}) \frac{\text{p}(\widetilde{\boldsymbol{Q}}|K-1)}{\text{p}(\boldsymbol{Q}|K)} \times (\delta^2+1) \left(\frac{v_{\sigma^2} + \boldsymbol{y}^{\text{T}} \widetilde{\boldsymbol{\Pi}} \boldsymbol{y}}{v_{\sigma^2} + \boldsymbol{y}^{\text{T}} \boldsymbol{\Pi} \boldsymbol{y}} \right)^{-\frac{2N+n_{\sigma^2}}{2}}. \tag{28}$$

The ratio of birth to death is

$$\frac{b(K-1, \Lambda)}{d(K, \Lambda)} = \frac{\text{p}(K-1|\Lambda)}{\text{p}(K|\Lambda)} = \frac{\Lambda}{K}.$$

The acceptance ratio is

$$r_{\text{death}} = (\delta^2+1) \left(\frac{v_{\sigma^2} + \boldsymbol{y}^{\text{T}} \widetilde{\boldsymbol{\Pi}} \boldsymbol{y}}{v_{\sigma^2} + \boldsymbol{y}^{\text{T}} \boldsymbol{\Pi} \boldsymbol{y}} \right)^{-\frac{2N+n_{\sigma^2}}{2}}. \tag{29}$$

Note that this is the inverse of the ratio for a birth from $K-1$ to K frequencies.

C. Update

At each stage of the sampling with a fixed number of Lorentzian components, all the parameters are updated as described below.

a) Updating Λ*:* The conditional posterior distribution for Λ reads

$$\text{p}(\Lambda|\boldsymbol{y}, K, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}) \propto \text{p}(\Lambda)\text{p}(K|\Lambda) \propto \frac{\Lambda^{K+\frac{1}{2}+e_1-1}\, e^{-(1+e_2)\Lambda}}{\sum_{K=0}^{K_{\max}} \frac{\Lambda^K}{K!} e^{-\Lambda}} \quad \Lambda > 0. \tag{30}$$

Λ is updated using the MH steps,

- With probability $0 < \lambda < 1$, we perform an MH step with proposal distribution $\text{Gamma}(\Lambda_{\text{prop}}|K + \frac{1}{2} + e_1, 1 + e_2)$.
- With probability $1 - \lambda$, we perform an MH step with proposal distribution $\mathcal{N}(\Lambda_{\text{prop}}|\Lambda, \sigma_\Lambda^2)$, where we selected a small heuristic value of $\sigma_\Lambda^2 = 0.25$.
- Λ is accepted with probability computed according to Eq. (10). The invariant distribution is the conditional posterior distribution of Λ in Eq. (30).

b) Updating δ^2*:* Following the graphic model in Fig. 1, δ^2 can be sampled after sampling σ^2 and $\boldsymbol{\alpha}$ using the Gibbs sampler.

- The conditional posterior distributions of the nuisance parameters $(\sigma^2, \boldsymbol{\alpha})$ are drawn from their full conditional distribution.

$$\begin{aligned}&\mathrm{p}(\sigma^2|\boldsymbol{y}, K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q})\\&= \text{Scale-inv-}\chi^2\left[\sigma^2\middle|2N + n_{\sigma^2}, \frac{\left(v_{\sigma^2} + \boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Pi}\boldsymbol{y}\right)}{2N + n_{\sigma^2}}\right].\end{aligned} \tag{31}$$

and

$$\mathrm{p}(\boldsymbol{\alpha}|\boldsymbol{y}, K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2) = \mathcal{N}(\boldsymbol{\alpha}|\widetilde{\boldsymbol{\alpha}}, \sigma^2\boldsymbol{\Gamma}^{-1}). \tag{32}$$

- We use the Gibbs sampler for δ^2. The conditional posterior distribution is

$$\begin{aligned}&\mathrm{p}(\delta^2|\boldsymbol{y}, K, \Lambda, \boldsymbol{\alpha}, \boldsymbol{f}_0, \boldsymbol{Q}, \sigma^2)\\&\propto p(\delta^2)\,|\boldsymbol{\Psi}|^{\frac{1}{2}}\,\exp\left(-\frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\Psi}\boldsymbol{\alpha}}{2\sigma^2}\right)\\&\propto \left(\delta^2\right)^{-\left(\frac{2M+2K+n_{\delta^2}}{2}+1\right)}\\&\quad\times\exp\left[-\frac{1}{2\delta^2}\left(v_{\delta^2} + \frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta}\boldsymbol{\alpha}}{\sigma^2}\right)\right]\\&= \text{Inv-}\chi^2\left[\delta^2\middle|2M + 2K + n_{\delta^2},\right.\\&\qquad\left.\frac{1}{M + 2K + n_{\delta^2}}\left(v_{\delta^2} + \frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta}\boldsymbol{\alpha}}{\sigma^2}\right)\right].\end{aligned} \tag{33}$$

Details of the parameter specification are given in Discussion V.

c) Updating $\boldsymbol{f}_0$: The conditional posterior probability of each frequency f_{0k} reads

$$\begin{aligned}&\mathrm{p}(f_{0k}|\boldsymbol{y}, K, \Lambda, \delta^2, \boldsymbol{f}_{0,\sim k}, \boldsymbol{Q})\\&\propto \mathrm{p}(f_{0k}|K)\,\left(v_{\sigma^2} + \boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Pi}\boldsymbol{y}\right)^{-\frac{2N+n_{\sigma^2}}{2}}\,\mathbb{I}_{[f_{\min}, f_{\max}]}(f_{0k})\\&\propto \left(v_{\sigma^2} + \boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Pi}\boldsymbol{y}\right)^{-\frac{2N+n_{\sigma^2}}{2}}\,\mathbb{I}_{[f_{\min}, f_{\max}]}(f_{0k}).\end{aligned} \tag{34}$$

Since $\boldsymbol{f}_0$ is a vector, we use a hybrid MCMC sampler that combines Gibbs steps and MH steps. In details, each f_{0k} is randomly selected and sampled using a MH step.

- With probability $0 < \lambda < 1$, we perform an MH step with proposal distribution that is proportional to the squared modulus of the spectrum data.
- With probability $1 - \lambda$, we perform an MH step with proposal distribution as a normal distribution $\mathcal{N}\left(f_{0k\mathrm{prop}}\middle|f_{0k}, \sigma_f^2\right)$, where σ_f^2 should be enough small, and we assigned the frequency sampling step, i.e., $\sigma_f^2 = \frac{f_{\max} - f_{\min}}{N_f}$.
- f_{0k} is accepted with probability computed according to Eq. (10), where the conditional posterior probability in Eq. (34) is the invariant distribution.

d) Updating $\boldsymbol{Q}$: The conditional posterior distribution for each Q_k reads

$$\begin{aligned}&\mathrm{p}(Q_k|\boldsymbol{y}, K, \Lambda, \delta^2, \boldsymbol{f}_0, \boldsymbol{Q}_{\sim k})\\&\propto \mathrm{p}(Q_k|K)\,\left(v_{\sigma^2} + \boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Pi}\boldsymbol{y}\right)^{-\frac{2N+n_{\sigma^2}}{2}}.\end{aligned} \tag{35}$$

Akin to the $\boldsymbol{f}_k$ updating strategy, each Q_k of $\boldsymbol{Q}$ is randomly chosen and sampled using MH steps.

- With probability $0 < \lambda < 1$, we sample Q_k using the prior distribution $\mathrm{Gamma}\left(Q_k\middle|a_Q, b_Q\right)$ as the proposal distribution.
- With probability $1 - \lambda$, the proposal distribution is a normal distribution $\mathcal{N}\left(Q_{k\mathrm{prop}}\middle|Q_k, \sigma_Q^2\right)$, where $\sigma_Q^2 = 0.01$.
- In both MH steps, the acceptance ratio is computed according to Eq. (10), where the conditional posterior distribution in Eq. (35) is the invariant distribution.

Appendix D
Notation

- $\mathbb{I}_E(z)$ is the indicator function of the set E. If $z \in E$, $\mathbb{I}_E(z) = 1$, 0 otherwise.
- Gaussian

$$\mathcal{N}\left(z\middle|\mu, \sigma^2\right) = \frac{1}{\sqrt{2\sigma^2\pi}}\exp-\frac{(z-\mu)^2}{2\sigma^2}.$$

- Gamma

$$\mathrm{Gamma}\left(z\middle|\alpha, \beta\right) = \frac{\beta^\alpha}{\Gamma(\alpha)}\,z^{\alpha-1}\,\exp(-\beta z)$$
$$\alpha, \beta > 0 \text{ and } z > 0.$$

- Scaled inverse-chi-square

$$\text{Scale-inv-}\chi^2\left(z\middle|\nu, \tau^2\right) = \frac{\left(\frac{\tau^2\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)}\frac{\exp\left(-\frac{\nu\tau^2}{2z}\right)}{z^{1+\frac{\nu}{2}}}$$
$$z > 0.$$

- Uniform

$$U_E(z) = \left[\int_E dz\right]^{-1}\mathbb{I}_E(z)$$

Acknowledgment

The authors would like to extend sincere gratitude to Dr. Pascal Laugier in Laboratoire d'Imagerie Biomédicale (LIB) for his constructive comments and excellent advice to improve the quality of this study.

References

[1] G. L. Bretthorst, "Bayesian analysis III. Applications to NMR signal detection, model selection, and parameter estimation," *J. Magn. Reson.*, vol. 88, no. 3, pp. 571–595, 1990.

[2] J. B. Poullet, D. M. Sima, and S. Van Huffel, "MRS signal quantitation: A review of time-and frequency-domain methods," *J. Magn. Reson.*, vol. 195, no. 2, pp. 134–144, 2008.

[3] L. Vanhamme, T. Sundin, P. V. Hecke, and S. V. Huffel, "MR spectroscopy quantitation: A review of time domain methods," *NMR Biomedicine*, vol. 14, no. 4, pp. 233–246, 2001.

[4] P. Stoica and R. L. Mose, *Spectral Analysis of Signals*. Upper Saddle River, NJ, USA: Prentice-Hall, 2005.

[5] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. TAC-19, no. 6, pp. 716–723, Dec. 1974.

[6] A. Barron, J. Rissanen, and B. Yu, "The minimum description length principle in coding and modeling," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2743–60, Oct. 1998.

[7] P. M. Djuric, "A model selection rule for sinusoids in white Gaussian noise," *IEEE Trans. Signal Process.*, vol. 44, no. 7, pp. 1744–1751, Jul. 1996.

[8] C. Andrieu and A. Doucet, "Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2667–2676, Oct. 1999.

[9] P. J. Green, "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.

[10] D. V. Rubtsov and J. L. Griffiin, "Time-domain Bayesian detection and estimation of noisy damped sinusoidal signals applied to NMR spectroscopy," *J. Mag. Reson.*, vol. 188, no. 2, pp. 367–379, 2007.

[11] A. Migliori *et al.*, "Resonant ultrasound spectroscopic techniques for measurement of the elastic moduli of solids," *Phys. B: Condens. Matter*, vol. 183, nos. 1/2, pp. 1–24, 1993.

[12] 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, pp. 12–19, 2013.

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

[14] R. G. Leisure, K. Foster, J. E. Hightower, and D. S. Agosta, "Internal friction studies by resonant ultrasound spectroscopy," *Mater. Sci. Eng.: A*, vol. 370, no. 1, pp. 34–40, 2004.

[15] A. V. Lebedev, "Method of linear prediction in the ultrasonic spectroscopy of rock," *Acoust. Phys.*, vol. 48, no. 3, pp. 391–399, 2002.

[16] M. C. Remillieux, T. J. Ulrich, C. Payan, J. Rivire, C. R. Lake, and P. Y. Le Bas, "Resonant ultrasound spectroscopy for materials with high damping and samples of arbitrary geometry," *J. Geophys. Res.: Solid Earth*, vol. 120, no. 7, pp. 4898–4916, 2015.

[17] R. Kumaresan and D. W. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, no. 6, pp. 833–840, Dec. 1982.

[18] G. A. F. Seber, *Multivariate Observations*. New York, NY, USA: Wiley, 1984.

[19] H. Spath, *Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples*, Translated by J. Goldschmidt. New York, NY, USA: Halsted, 1985.

[20] C. Andrieu, P. M. Djuri, and A. Doucet, "Model selection by MCMC computation," *Signal Process.*, vol. 81, no. 1, pp. 19–37, 2001.

[21] A. Zellner, "Bayesian estimation and prediction using asymmetric loss functions," *J. Amer. Statist. Assoc.*, vol. 81, no. 394, pp. 446–451, 1986.

[22] A. Roodaki, J. Bect, and G. Fleury, "Comments on "Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC"," *IEEE Trans. Signal Process.*, vol. 61, no. 14, pp. 3653–3655, Jul. 2013.

[23] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis* (Texts in Statistical Science), 2nd ed. London, U.K.: Chapman & Hall, 2004, pp. 338–339.

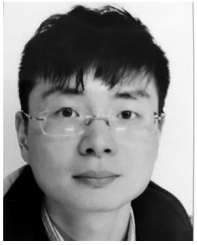
[24] W. Astle, M. De Iorio, S. Richardson, D. Stephens, and T. Ebbels, "A Bayesian model of NMR spectra for the deconvolution and quantification of metabolites in complex biological mixtures," *J. Amer. Statist. Assoc.*, vol. 107, no. 500, pp. 1259–1271, 2012.

[25] A. D. Yaghjian and S. R. Best, "Impedance, bandwidth, and Q of antennas," *IEEE Trans. Antennas Propag.*, vol. 53, no. 4, pp. 1298–1324, Apr. 2005.

[26] K. Yamanaka, Y. Maruyama, and T. Tsuji, "A resonance frequency and Q factor mapping by ultrasonic atomic force microscopy," *Appl. Phys. Lett.*, vol. 78, no. 13, pp. 1939–1941, 2001.

[27] S. Jesse, B. Mirman, and S. V. Kalinin, "Resonance enhancement in piezoresponse force microscopy: Mapping electromechanical activity, contact stiffness, and Q factor," *Appl. Phys. Lett.*, vol. 89, 2006, Art. no. 022906.

[28] N. Metropolis, N. Rosenblutt, A. W. Rosenblutt, M. N. Teller, and A. H. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, pp. 1087–1092, 1953.



Kailiang Xu (M'14) received the Ph.D. degree in biomedical engineering from Fudan University, Shanghai, China, in 2012. He is currently an Associate Professor with the Department of Electronic Engineering, Fudan University, and works with the Medical Ultrasound Laboratory. His research interests include wave simulation, signal processing, inverse problem, and the development of multi-wave imaging techniques, and elasticity characterization for medical ultrasound and non-destructive evaluation.

Guillaume Marrelec received the M.S. degree in engineering and applied mathematics jointly from the École Centrale Paris, Paris, France, and the Universität Stuttgart, Stuttgart, Germany, in 1999, and the Ph.D. degree in medical imaging from the Université Paris XI, Orsay, France, in 2003. He is currently a Senior Research Scientist with the Institut National de la Santé et de la Recherche Médicale, Paris, and works with the Laboratoire d'imagerie Biomédicale. His research interests include brain imaging, functional magnetic resonance imaging, functional connectivity, information theory, and Bayesian analysis.

Simon Bernard received the M.S. and Ph.D. degrees in physical acoustics from the University Pierre et Marie Curie, Paris, France, in 2011 and 2014, respectively. He is currently a Maître de Conférence (Assistant Professor) with the Université Le Havre Normandie, Le Havre, France, and works with the Waves and Complex Media Laboratory. His research interests include ultrasound methods for imaging and material characterization, modeling of wave propagation, and inverse problems.

Quentin Grimal has been with the Department of Engineering, Sorbonne Université, Paris, France, and a member of the Laboratory of Biomedical Imaging since 2004. He is leading a research group, where his main focus is on the measurement of the physical properties of hard biological tissues such as bone with ultrasound, including the development of novel devices for the diagnosis of bone pathologies and the development of laboratory techniques such as resonant ultrasound spectroscopy for the multiscale investigation of bony tissues.