A fully Bayesian approach based on Bernoulli–Gaussian prior for the identification of sparse vibratory sources from displacement measurements

Charly Faure, Jérôme Antoni, Frédéric Ablitzer, Charles Pézerat

Version of Record: https://www.sciencedirect.com/science/article/pii/S0022460X21007240
Manuscript_9cd33a453e85441e866963d01f98a671

A fully Bayesian approach based on Bernoulli-Gaussian prior for the identification of sparse vibratory sources from displacement measurements

Charly Faure[a], Jérôme Antoni[b], Frédéric Ablitzer[a,*], Charles Pézerat[a]

[a] *Laboratoire d'Acoustique de l'Université du Mans, CNRS – UMR 6613, avenue Olivier Messiaen, 72085 Le Mans cedex 9, France*

[b] *Laboratoire Vibrations Acoustique, Univ Lyon, INSA-Lyon, LVA EA677, F-69621 Villeurbanne, France*

Abstract

This paper introduces a fully Bayesian approach to the Force Analysis Technique (FAT), which aims at identifying sparse vibratory sources from displacement measurements. Being based on the local equation of motion of a structure, the FAT allows for the estimation of vibratory sources without the need of specifying boundary conditions. Nevertheless, since it involves the calculation of higher-order spatial derivatives of the measured displacements, it is highly sensitive to noise perturbations and thus requires careful regularization. Besides, although sparse excitations are commonplace in structural vibrations, standard regularization strategies tend to over-smooth them in the reconstruction process. This paper shows how to reconcile the two goals of regularization and sparsity enforcement in the FAT by setting up a hierarchical Bayesian model rooted on a Bernoulli-Gaussian prior. Inference of all the parameters in the model is achieved with a Gibbs sampler whose convergence is efficiently accelerated with a partial collapsing strategy.

Keywords: vibratory source identification, inverse problem, Bayesian regularization, partially collapsed Gibbs sampler, sparsity, Bernoulli-Gauss.

1. Introduction

In several applications in structural dynamics, vibratory sources happen to be highly localized [1]. The force distribution is thus said to be *sparse* [2]. By sparse it is meant that the force distribution is expected to take zero values almost everywhere except in a few locations. This is the case for example when subsystems are

*Corresponding author
Email address: frederic.ablitzer@univ-lemans.fr (Frédéric Ablitzer)

linked by bolts, rivets or welds. The aim of the present paper is the identification of sparse vibratory sources from measurements of the displacement responses of the structure under test by using the Force Analysis Technique (FAT) [3, 4]. Contrary to most other approaches [5, 6, 7, 8] – which typically involve the inversion of a matrix of impulse responses (time-domain approach), a matrix of transfer functions (frequency domain approach), or a state matrix (state-space approach) – the FAT offers a solution based on the local equation of motion of the structure, whose right-hand side member directly returns the excitation at any spatial position. The definite advantage of the FAT is to not require the (often intricate) specification of boundary conditions. The price to pay is an extreme sensitivity to noise perturbations which are amplified in the calculation of higher-order spatial derivatives (i.e. by means of finite differences of the measured displacements). It should be realized that this sensitivity reflects the ill-posedness of the inverse problem, just as with other inverse methods where noise is typically inflated after application of an inverse operator. This requires the resort to some sort of regularization. In Ref. [4], it has been proposed to solve the FAT within a Bayesian framework, which enjoys intrinsic regularization with an automatic inference mechanisms to tune the regularization parameter, for instance by means of empiricical Bayes or Markov Chain Monte Carlo algorithms. In that work, a Gaussian prior was introduced as the simplest choice; yet this is not the most relevant one with sparse sources. Indeed, the Gaussian prior tends to spread out the reconstructed sources all over the spatial domain, thus missing their sparsity. Consequently, closely spaced point sources could hardly be resolved. It is then more appropriate to replace the Gaussian prior by a "sparse" one, which promotes most of the values of the spatial distribution of the force to be zero or nearly so. Several kinds of such priors exist, such as the Student-t distribution [9], the Laplace distribution [10] or the generalized Gaussian distribution [11, 12]. From the optimization point of view, the latter is equivalent to the minimization of a cost function induced by an ℓ_p norm with $0 < p$ over the source field, with p being the power in the probability density function (pdf). The Laplace and Gaussian distributions are thus special cases of the generalized Gaussian distribution with $p = 1$ and $p = 2$, respectively. Optimization of cost functions with this kind of constraints can be solved using iterative algorithms such as the iterated-reweighted-least-squares (IRLS) [13] or numerical Bayesian approaches [14]. It is also possible to combine the Laplacian and the Gaussian priors to identify both sparse and extended parts of the source field [15]. This specific combination is known as the *Elastic Net*. Also, by mixing several centered Gaussians with different variances, a pdf is obtained with a significant mass at zero and heavy tails, which promotes sparsity [16]. A specific case of the latter model is the Bernoulli-Gaussian, obtained by adding a Dirac distribution at zero (equivalent to a centered Gaussian with nil variance) and a centered Gaussian with free variance. The Bernoulli distribution corresponds to the minimization of

the quasi-norm ℓ_0, which counts the number of non-zero elements. It is thus an especially strong sparsity constraint. The Bernoulli-Gaussian prior has been known for many years [17] and used to model sparse signals in various applications, such as in bio-medical science [18, 19, 20, 21], astrophysics or geophysics [22, 23, 24]. Besides, it naturally fits into a hierarchical probabilistic approach, which easily allows for the construction of other related priors [18, 24], including the Bernoulli-Laplacian [18, 25, 26, 21] or even the Bernoulli-Laplace-Gaussian [27]. In this work, the choice of the Bernoulli-Gaussian prior results from a trade-off between high sparsity, implementation complexity and convergence of the inference algorithm. A Gibbs sampler is proposed for full inference of the resulting Bayesian model. The Gibbs sampler is a relevant choice to explore high-dimensional posterior pdfs, as is the case here. In addition, it is ideally suited to sample hierarchical models. This makes possible the joint inference of the hyperparameters of the problem, which play a crucial role for regularization.

The fully Bayesian paradigm adopted here parallels some recent works which have demonstrated its relevance in inverse problems of structural dynamics, such as system identification [28, 29, 30, 31], model updating [32, 33, 34, 35, 36] and force reconstruction [37, 38, 39, 40]. In Ref. [37], a fully Bayesian methodology based on frequency response functions has been developed to reconstruct an input force from output measurements in the presence of model uncertainties, yet without imposition of sparsity constraints. In Ref. [41], a Bayesian hierarchical model has been solved by relevance vector machine approximation in order to enforce the sparsity of recovered transients in a time formulation based on a matrix of impulse responses. References [38] and [39] have addressed the sparse case in a frequency domain formulation based on the matrix of transfer functions by using a generalized Gaussian "local" prior, which allows the degree of sparsity to change as a function of space; the optimization problem has been first solved by maximum aposteriori estimation regularization tuned with the L-curve ([38]) and next in a fully Bayesian approach where all parameters – including the shape parameter of the generalized Gaussian – are infered together ([39]). A similar approach has been resumed in Ref. [40], with a slightly different Markov Chain Monte Carlo (MCMC) algorithm, by using a state-space formulation of the direct problem. It is noteworthy that the fully Bayesian approach introduced in the present paper differs from the latter works in two aspects: first, it is designed for the FAT, a technique which has its own specificities and advantages as explained in Ref. [4]; second, it uses a Bernoulli-Gaussian prior which is theoretically able to enforce a much higher degree of sparsity than the generalized Gaussian used in the aforementioned references.

The paper is organized as follows. Section 2 first resumes the principle of the FAT. Next, section 2.4

formulates the inverse problem into a fully Bayesian framework with intrinsic regularization and sparsity enforcement by means of the Bernoulli-Gaussian prior. Two versions of a Gibbs sampler are given in order to explore the posterior pdf of the source distribution. The second one includes a "partially collapsed" strategy which turns out decisive in accelerating the convergence of the Markov chains. Finally, sections 4 and 5 validate the proposed methodology on numerical simulations and on experimental data, respectively.

2. Identification of vibratory sources from transverse displacement

This section resumes the principle of the FAT initially introduced in Ref. [3]. The FAT is in essence a local method based on the equation of motion of a structure. It consists in calculating the spatial and temporal derivatives involved in the left-hand side of the partial differential equation and to directly identify them with the source distribution in the right-hand side of the equation. It is trivially converted to the frequency domain after application of the Fourier transform to the equation of motion. In the present work, all the data collected at different spatial positions are taken into account as a whole, in a matrix formulation, which actually makes a slight modification as compared to the original formulation of Ref. [3], where a stencil was used to scan the measurements point by point. Next, a qualitative explanation of the limitations of the FAT with regards to noise perturbations is given, leading to the next section dedicated to regularization.

2.1. Equation of motion

The FAT is based on the equation of motion of a known structure. As an example, the method is presented here on a beam within the Euler-Bernoulli theory. With temporal convention $e^{+j\omega t}$ defined with radial frequency ω, the harmonic transverse displacement $w(x,\omega) \in \mathbb{C}$ of the beam satisfies the partial differiential equation,

$$E(1+j\eta)I\frac{\partial^4 w(x,\omega)}{\partial x^4} - \rho S\omega^2 w(x,\omega) = f(x,\omega), \tag{1}$$

where E is Young's modulus, j the imaginary number ($j^2 = 1$), η the loss factor, I the second moment of inertia, ρ the density, S the cross-section area of the beam and $f(x,\omega) \in \mathbb{C}$ the distribution of the vibratory source or "force" (in N/m) at location x and angular frequency ω. In the above equation, $f(x,\omega)$ stands for the unknown of the problem. The principle of the FAT is to measure the transverse displacement field at a number of points $\{x_i; i = 1, ..., N\}$ (hereafter assumed regularly spaced for simplicity), then to estimate its spatial derivatives, to evaluate the terms in the left-hand side of Eq. (1) and finally to deduce the vibratory

source distribution applied on the structure as it is directly given by the right-hand side of Eq. (1). Since the inverse problem is based on the equation of motion in harmonic regime, different sources operating at different frequencies can be identified, provided that the displacement field at the corresponding frequencies is used to evaluate the left-hand side of Eq. (1).

2.2. Calculation of derivatives

The spatial derivatives needed in the FAT are typically calculated by numerical differentiation, using the centered finite differences at first order,

$$\frac{\partial^4 w(x_i,\omega)}{\partial x^4} \approx \frac{w(x_{i+2},\omega) - 4w(x_{i+1},\omega) + 6w(x_i,\omega) - 4w(x_{i-1},\omega) + w(x_{i-2},\omega)}{\Delta_x^4}, \tag{2}$$

where $\Delta_x = x_{i+1} - x_i$ is the spatial sampling step. Substituting Eq. (2) into Eq. (1) allows the identification of the vibratory source applied at location x_i without having to evaluate the boundary conditions over the domain, i.e.

$$E(1+j\eta)I\frac{w_{i+2} - 4w_{i+1} + 6w_i - 4w_{i-1} + w_{i-2}}{\Delta_x^4} - \rho S\omega^2 w_i = f_i, \tag{3}$$

where w_i and f_i are simplified notations for $w(x_i,\omega)$ and $f(x_i,\omega)$.

The excitation can be classified into two categories: it can result either from pointwise sources or distributed sources. In the first case, an equivalent force amplitude in Newtons is simply obtained by spatially integrating the source distribution around each point. If the source is distributed with constant phase (for example, a transmission through a welding line), spatial integration can again lead to an equivalent force amplitude. However if the phase is seen to change as a function of space (for example, an acoustic excitation), it might be more relevant to interpret the results in terms of a source distribution rather than an equivalent pointwise force – which might integrate destructively. This is indeed an advantage of the FAT over transfer matrix based methods which usually return equivalent forces at a limited number of specific points.

2.3. Matrix formulation

Considering a portion of the structure under test with a constant spatial discretization, Eq. (3) can be expressed in a matrix form,

$$\boldsymbol{D}\boldsymbol{w} = \boldsymbol{f}, \tag{4}$$

where $\boldsymbol{w} \in \mathbb{C}^N$ is the vector of displacements, $\boldsymbol{f} \in \mathbb{C}^{N-4}$ the vector of vibratory sources and $\boldsymbol{D} \in \mathbb{C}^{(N-4)\times N}$ the operator matrix resulting from the discretization of the structural local operator. Due to the fourth order

spatial derivative, $\boldsymbol{D}$ is an $(N-4)$-by-N rectangular matrix, N being the number of measurement points. Since the inverse of this matrix will be required to solve the inverse problem (see section 3), it is proposed to make it square by adding two rows at each boundary. Different strategies are possible. One is to add two extra derivatives at each boundary in order to evaluate local boundary conditions based on first physical principles. The drawback of this strategy is that there is no automatic rule for choosing between imposed displacement, slope, bending moment or shear force as boundary conditions, while only two of them at most are required. More generally, it faces the difficulty of accurately specifying a boundary condition (which is often only partially unknown in complex systems) and, therefore, missing one of the definite advantages of the FAT, which is exactly to avoid this requirement as opposed to other inverse methods. Another drawback that the authors have observed in practice is that the resulting derivatives at boundaries can have much higher magnitudes than the sources evaluated inside the domain, thus making the problem ill-conditioned (see section 3).

Alternatively, another strategy more in the spirit of FAT is to evaluate the source at the boundaries by using fourth order forward and backward finite difference schemes, with stencils $[3; -14; 26; -24; 11; -2]$ and $[-2; 11; -24; 26; -14; 3]$, respectively. Therefore the two vectors $\boldsymbol{w}$ and $\boldsymbol{f}$ end up with the same dimension and matrix $\boldsymbol{D}$ is square. This is the strategy adopted in the present work. It is noteworthy that it can also be implemented within a finite element method to estimate sources at boundaries of a sub-domain of the structure [42, 43], the interpreation of the result then being a combination of internal stresses and external sources located at the boundaries (on the contrary, the use of a finite difference scheme with forward and backward stencil at boundaries only depends on external sources applied on the studied domain). Moreover, although the bias induced by finite difference is higher than with the finite element method, it can be reduced drastically by applying an analytic correction as shown in Ref. [44]. Hence, it seems to be an appropriate model in the present work.

2.4. Additive noise

As each measurement includes noise, the equation of motion is modified as

$$\boldsymbol{y} = \boldsymbol{w} + \boldsymbol{n}, \tag{5}$$

where $\boldsymbol{y} \in \mathbb{C}^N$ is the vector of observed displacement and $\boldsymbol{n} \in \mathbb{C}^N$ the vector of additive noise. White Gaussian noise has been found to be a reasonable assumption as justified in Refs. [33, 4], especially when

one works in the frequency domain. A multiplicative noise part may also be present depending on the experimental setup and the studied physical quantities, but it is often neglected to simplify the approach and mathematical developments. Applying the principle of the FAT to the observable quantity $\boldsymbol{y}$ then yields

$$\boldsymbol{D}\boldsymbol{y} = \boldsymbol{f} + \boldsymbol{D}\boldsymbol{n}. \tag{6}$$

Although the level of noise $\boldsymbol{n}$ is initially much lower than that of displacement $\boldsymbol{w}$, $\boldsymbol{D}$ is a differential operator which, when applied to $\boldsymbol{n}$, may considerably inflate it and thus jeopardize the source identification task. Thus, a regularization step is needed to overcome this issue.

It must be stressed that the idealized noise model considered in the present methodology represents only random errors related to the measurement process. They must be distinguished from modelling errors, which arise if the operator $\mathbf{D}$ used in the inverse problem does not reflect the actual dynamic behavior of the structure. While random errors are likely to make the inverse problem unstable, modelling errors introduce a systematic bias in the results of the inverse problem, even when a regularization is applied. The way modelling errors can lead to such deviations is discussed in Sections 4.3 and 4.4.

3. Regularization within the Bayesian framework

The Bayesian framework for solving the inverse problem is first presented in this section together with its inherent regularization. The Gibbs sampler algorithm is then introduced as an MCMC to solve automatically the inverse problem with a sparse prior on sources. These results are finally compared to those obtained with a Gibbs sampler using the traditional Gaussian prior. From now on, the latter will be considered as a point of reference since it was shown in Ref. [4] to return better results than the classical Tikhonov regularization [45] tuned with the L-curve [46, 47, 8, 48] or GCV [49, 7]. Sections 3.1 to 3.3 take elements already described in [4]. Sections 3.4 to 3.8 contain original points, which are specific to the Bernoulli-Gaussian prior.

3.1. Bayes' theorem

In most inverse problems, some quantities are considered as perfectly known, some as partially known or "uncertain" and some as completely unknown. Working with a probabilistic point of view is then a suitable approach to take into account these different types. The Bayesian framework used in this work allows the inclusion of prior information in the inverse problem.

It is based on the following theorem,

$$[X \mid Y] = \frac{[Y \mid X]\,[X]}{[Y]} \propto [Y \mid X]\,[X]\,, \tag{7}$$

where X represents the unknown and Y the observation (measured data). The posterior density, $[X \mid Y]$ corresponds to the solution of the inverse problem. A regularized point estimate is obtained by finding the maximum of this pdf, which is known as the Maximum *A Posteriori* (MAP). $[Y \mid X]$ is the pdf involved in the direct problem which expresses the information extracted from the experiment and is called the "likelihood". $[X]$ is the prior pdf, delivered from prior knowledge and $[Y]$ is called the "marginalized likelihood" or the "evidence". The evidence acts as a multiplicative constant to ensure that the product of the prior and the likelihood is still a pdf (the integral must be equal to 1). As the shape of the posterior is not impacted by the evidence, the solution maximizing this pdf remains the same. Consequently, the proportional relationship is preferred in Eq. (7) rather than the absolute equality.

3.2. Priors

Considering additive white Gaussian noise, the prior distribution of the noise Fourier coefficient $\boldsymbol{n}$ is expressed as a complex Gaussian

$$[\boldsymbol{n}] \propto \mathcal{N}_c\left(\boldsymbol{n}; \mathbf{0}, \tau_n^{-1}\mathbf{I}\right). \tag{8}$$

The scalar precision (the inverse of the variance) of noise, τ_n, is here preferred to the most usual variance for simplification induced in sampling (see section (3.5)). $\mathbf{I}$ stands for the $N \times N$ identity matrix and $\mathcal{N}_c$ for the N-dimensional multivariate circular complex Gaussian pdf of $\boldsymbol{x}$ (see Ref. [50]), with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, defined as

$$\mathcal{N}_c\left(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) = \frac{1}{\pi^N |\boldsymbol{\Sigma}|} \exp\left(\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^H \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{x}-\boldsymbol{\mu}\right)\right), \tag{9}$$

where $|\boldsymbol{\Sigma}|$ stands for the determinant of the covariance matrix and the exponent H for the Hermitian transposition. The fact that the covariance matrix in Eq. (8) is proportional to the identity matrix reflects that the noise pdf is the same all over the domain and that there is no spatial dependence. Random variables with this type of covariance are then independent and identically distributed. It is analogous to the "rain-on-the-roof" hypothesis of Statistical Energy Analysis.

3.3. Likelihood

From Eqs. (6) and (8), the pdf of the observations $\boldsymbol{y}$ knowing the sources $\boldsymbol{f}$, the structural operator $\boldsymbol{D}$ and the measurement noise $\boldsymbol{n}$ (i.e. the pdf corresponding to the direct problem) can be expressed as

$$[\boldsymbol{y} \mid \boldsymbol{H}, \boldsymbol{f}, \tau_n] \propto \mathcal{N}_c\left(\boldsymbol{y}; \boldsymbol{H}\boldsymbol{f}, \tau_n^{-1}\mathbf{I}\right). \tag{10}$$

It should be noted here that the likelihood is expressed by means of the "transfer matrix" $\boldsymbol{H} = \boldsymbol{D}^{-1}$, the inverse of the dynamic stiffness matrix $\boldsymbol{D}$.

Although the original formulation of the FAT has been chosen as a starting point (see Eq. (4)), the present method requires the use of $\boldsymbol{D}^{-1}$ to solve the inverse problem. This also occurs in other variants of the FAT, such as the one using a finite element operator to identify forces [42].

3.4. Bernoulli-Gaussian prior on source distribution

As announced in the introduction, sparsity on the source distribution is enforced by using a prior in the form of a Gaussian mixture, of which the Bernoulli-Gaussian (hereafter denoted by $\mathcal{BG}$) is a special case with high efficiency. The Bernoulli process follows a discrete pdf allowing only two values, 0 or 1. It can be interpreted as a switch, or a Boolean, for the activation of each node of the source spatial distribution. The relative weight between these two values is driven by the sparsity parameter λ. The Bernoulli pdf of a variable x with sparsity parameter λ is defined by

$$\mathcal{B}ern(x \mid \lambda) = \lambda^x \cdot (1-\lambda)^{1-x} \cdot \mathbb{1}_{\{0,1\}}(x), \tag{11}$$

where the notation $\mathbb{1}_{\{0,1\}}(x)$ specifies that variable x can only be 0 or 1. Hence, for $x = 0$, the associated probability is $(1-\lambda)$ and reciprocally the probability for $x = 1$ is λ. The lower λ is, the more pronounced the sparsity is. It should be noted that the value of λ is not only affected by the number of sources applied to the structure, but also by their spatial extent. As will be shown later, a unique force acting over a finite small area may result in a higher value of λ than several pointwise forces acting at different locations.

When a source node of the grid is activated by the Bernoulli process, another prior pdf is then imposed in order to compel its value. Assuming the latter is totally unknown, with real and imaginary parts either positive or negative, the prior should be symmetrically distributed around zero. Also, to respect a physical behavior, the source energy can not be infinite. So extreme values should be less probable than small ones.

To meet these requirements, it is common to use a complex Gaussian with zero-mean and precision parameter τ. Therefore, the complex Bernoulli-Gaussian pdf of x is expressed as

$$\mathcal{BG}(x \mid \lambda, \tau^{-1}) = (1-\lambda)\underbrace{\mathcal{N}_c(x \mid 0, 0)}_{\equiv\delta(x)} + \lambda\, \mathcal{N}_c(x \mid 0, \tau^{-1}). \tag{12}$$

In the above formulation, the Dirac distribution is seen (and symbolically noted) as a Gaussian with zero-mean and nil covariance as a way to remind the reader that the Bernoulli-Gaussian process is a particular case of the Gaussian mixture model. Thus, with a probability provided by the sparsity parameter λ, a sample is either drawn from the Dirac distribution $\delta(x)$ or from the Gaussian distribution $\mathcal{N}_c(x \mid 0, \tau^{-1})$. This prior being univariate, it only refers to a unique node of the source spatial distribution. To generalize to the whole grid, it is necessary to assume that nodes are independent from each other. Then, the same sparsity and precision parameters are applied to all nodes of the mesh. The multivariate distribution is thus the product of all univariate distributions, i.e.

$$[\boldsymbol{f} \mid \lambda, \tau_f] \quad \sim \quad \prod_{i=1}^{N} [f_i \mid \lambda, \tau_f] \tag{13}$$

with $[f_i \mid \lambda, \tau_f] = \mathcal{BG}(f_i \mid \lambda, \tau_f^{-1})$.

At this stage, it is insightful to interpret the source f_i at each node into the product of an activation variable q_i and a complex amplitude variable r_i,

$$f_i = q_i r_i. \tag{14}$$

In this way, the Bernoulli process applies only to the variable q_i, while the complex Gaussian distribution applies only to the variable r_i. Figure 1 illustrates the pdf of the Bernoulli-Gaussian distribution for different values of the sparsity parameter.

3.5. Conjugate priors

Depending on the state of knowledge on the problem, the hyperparameters τ_n, τ_f and λ needed in the aforementioned pdfs can be either imposed or considered completely unknown. Using a hierarchical Bayesian approach, an unknown hyperparameter can be easily estimated from an appropriate hyperprior, the so-called conjugate prior [51]. More explanations on conjugate priors and how to use them in the present inverse problem can be found in Ref. [4].

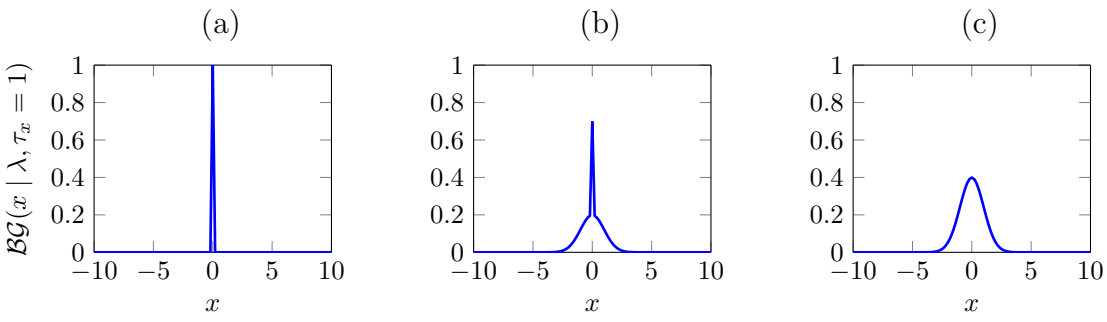

Figure 1: Illustrations of the Bernoulli-Gaussian distribution with zero-mean, precision $\tau_x = 1$ and sparsity parameter (a) $\lambda = 0$, (b) $\lambda = \frac{1}{2}$ et (c) $\lambda = 1$.

3.5.1. Inference of the precision parameters τ_n and τ_f

It turns out that the complex Gaussian distribution $\mathcal{N}_c\left(\boldsymbol{x} \mid \boldsymbol{\mu}, \tau^{-1}\mathbf{I}_N\right)$ (here expressed with the generic variables $\boldsymbol{x}$, $\boldsymbol{\mu}$ and τ) appearing in Eqs. (8) and (12) is endowed with an easy-to-sample conjugate prior for its precision parameter τ, in the form of a Gamma pdf. The gamma density on precision τ with shape parameter k and scale parameter θ is expressed as

$$\mathcal{G}\left(\tau; k, \theta\right) = \frac{\tau^{k-1}\exp\left(-\frac{\tau}{\theta}\right)}{\Gamma(k)\theta^k} \qquad \text{with } k > 0 \text{ and } \theta > 0, \tag{15}$$

where $\Gamma(k) = \int_0^\infty k^{t-1}\exp\left(-t\right)\mathrm{d}t$ is the gamma function.

The conditional pdf of τ given $\boldsymbol{x}$, $\boldsymbol{\mu}$, k and θ is thus given by the update rule

$$\begin{aligned}
\left[\tau \mid \boldsymbol{x}, \boldsymbol{\mu}, k, \theta\right] &\propto \mathcal{G}\left(\tau \mid k, \theta\right)\mathcal{N}_c\left(\boldsymbol{x} \mid \boldsymbol{\mu}, \tau^{-1}\mathbf{I}_N\right) && (16)\\
&= \frac{\tau^{k-1}\exp\left(-\frac{\tau}{\theta}\right)}{\Gamma(k)\theta^k}\frac{\exp\left(-\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^H\tau\mathbf{I}_N\left(\boldsymbol{x}-\boldsymbol{\mu}\right)\right)}{\pi^N|\tau^{-1}\mathbf{I}_N|} && (17)\\
&\propto \tau^{k-1+N}\exp\left(-\tau\left(\frac{1}{\theta}+\|\boldsymbol{x}-\boldsymbol{\mu}\|_2^2\right)\right) && (18)\\
&\propto \mathcal{G}\left(\tau \mid k+N, \left(\frac{1}{\theta}+\|\boldsymbol{x}-\boldsymbol{\mu}\|_2^2\right)^{-1}\right), && (19)
\end{aligned}$$

The two precisions τ_n and τ_f can thus be inferred, as illustrated in Fig. 3.

3.5.2. Inference of the sparsity parameter λ

A popular conjugate prior for a Bernoulli random variable is the Beta distribution, whose support set is the $[0; 1]$ interval, as for the sparsity parameter λ. The Beta distribution of a random variable λ with parameters

(a, b) is given by

$$\mathcal{B}(\lambda \mid a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\ \Gamma(b)} \lambda^{a-1}(1-\lambda)^{b-1} \mathbb{1}_{[0,1]}(\lambda). \tag{20}$$

An illustration of this pdf is displayed in Fig. 2 for different pairs of parameters (a, b). It is symmetric only when $a = b$. The mean is $a/(a+b)$ and the variance is $ab/((a+b)^2(a+b+1))$. So, the higher a and b are, the narrower the pdf is. The mode is $a - 1/(a+b-2)$ (with $a > 1$ and $b > 1$). Then, when a is lower than b, the mode is close to 0, and when a is upper than b, the mode is close to 1. The product of a Beta distribution on λ and a Bernoulli distribution with sparsity parameter λ is another Beta distribution,

$$\begin{aligned} \mathcal{B}(\lambda \mid a, b)\mathcal{B}ern(q \mid \lambda) &= \frac{\Gamma(a+b)}{\Gamma(a)\ \Gamma(b)} \lambda^{a-1}(1-\lambda)^{b-1} \lambda^{q} \cdot (1-\lambda)^{1-q} \cdot \mathbb{1}_{\{0,1\}}(q) &(21)\\ &= \mathcal{B}(\lambda \mid a+q, b+1-q), &(22) \end{aligned}$$

which is proportional to the conditional pdf $[\lambda|q, a, b]$. It is observed that when q is nil, the posterior on λ is shifted towards 0, enhancing also the sparsity.

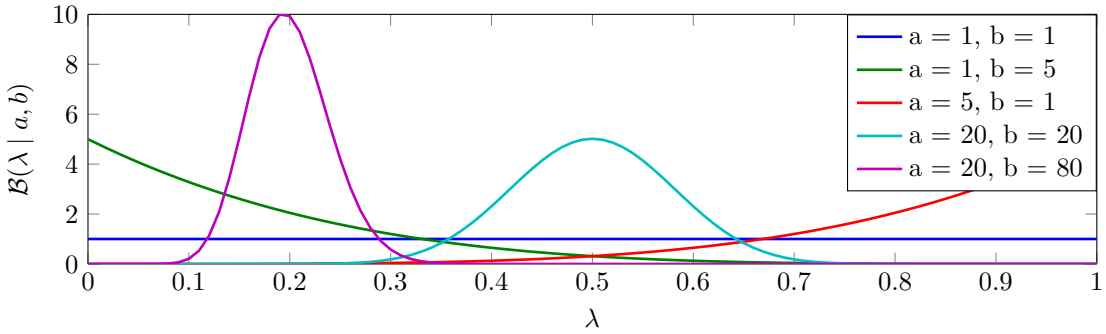

Figure 2: The Beta probability density function for different values of the parameters (a, b).

3.6. Hierarchical graph

Hierarchical graphs are a way to represent relationships between the various variables of a problem, either of random or deterministic natures, into a multi-level tree structure. For a specific variable in the graph, also called a *node*, links to other upper-level variables, or *parent* nodes, act as its priors while links to lower-level variables, or *children* nodes, represent its likelihood. The other parents of children nodes are called *co-parent* nodes. As a result of Bayes theorem, the pdf of a variable θ_j conditionnally to all the others, denoted hereafter as $\infty_{-\theta_j}$, will only depend on its parents, children and co-parents of its children as follows [52],

$$\left[\theta_j \mid \infty_{-\theta_j}\right] \propto \left[\theta_j \mid \text{Parents of } \theta_j\right] \times \prod_{\theta_k \text{ children of } \theta_j} \left[\theta_k \mid \text{Parents of } \theta_k\right]. \tag{23}$$

Based on the probabilities introduced in the previous section, the hierarchical graph of the inverse problem under study is as displayed in Fig. 3.

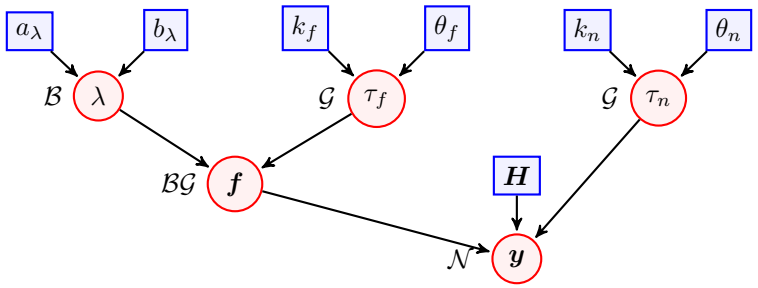

Figure 3: Hierarchical graph with Bernoulli-Gaussian prior on sources. Distributions are denoted by $\mathcal{N}$ for the Gaussian, $\mathcal{G}$ for the Gamma, $\mathcal{BG}$ for the Bernoulli-Gaussian and $\mathcal{B}$ for the Beta. Red circles stand for unknown stochastic variables to infer and blue squares stand for known deterministic variables.

For each random variable in the hierarchical graph, its conditional pdf is obtained from Eq. (23). The four conditional pdfs to sample from are thus (mathematical developments are detailed in appendix A):

$$
\begin{aligned}
[\tau_n \mid k_n, \theta_n, \boldsymbol{y}, \boldsymbol{H}, \boldsymbol{f}] &\propto [\tau_n \mid k_n, \theta_n]\,[\boldsymbol{y} \mid \boldsymbol{H}, \boldsymbol{f}, \tau_n] && (24)\\
&\propto \mathcal{G}(\tau_n \mid k_n, \theta_n)\mathcal{N}_c(\boldsymbol{y} \mid \boldsymbol{H}\boldsymbol{f}, \tau_n^{-1}\mathbf{I}_N) && (25)\\
&\propto \mathcal{G}\left(\tau_n \mid k_n + N, \left(\theta_n^{-1} + \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{f}\|_2^2\right)^{-1}\right), && (26)
\end{aligned}
$$

$$
\begin{aligned}
[\tau_f \mid k_f, \theta_f, \lambda, \boldsymbol{f}] &\propto [\tau_f \mid k_f, \theta_f]\,[\boldsymbol{f} \mid \lambda\tau_f] && (27)\\
&\propto \mathcal{G}(\tau_f \mid k_f, \theta_f)\prod_{i=1}^{N}\left\{\mathcal{BG}(f_i \mid \lambda, \tau_f^{-1})\right\} && (28)\\
&\propto \mathcal{G}(\tau_f \mid k_f + M, \left(\theta_f^{-1} + \|\boldsymbol{f}\|_2^2\right)^{-1}), && (29)
\end{aligned}
$$

$$
\begin{aligned}
[\lambda \mid a_\lambda, b_\lambda, \boldsymbol{f}, \tau_f] &\propto [\lambda \mid a_\lambda, b_\lambda]\,[\boldsymbol{f} \mid \lambda, \tau_f] && (30)\\
&\propto \mathcal{B}(\lambda \mid a_\lambda, b_\lambda)\prod_{i=1}^{N}\left\{\mathcal{BG}(f_i \mid \lambda, \tau_f^{-1})\right\} && (31)\\
&\propto \mathcal{B}(\lambda \mid a_\lambda + M, b_\lambda + N - M), && (32)
\end{aligned}
$$

$$[\boldsymbol{f} \mid \lambda, \tau_f, \boldsymbol{y}, \boldsymbol{H}, \tau_n] \propto [\boldsymbol{f} \mid \lambda, \tau_f]\,[\boldsymbol{y} \mid \boldsymbol{H}, \boldsymbol{f}, \tau_n] \tag{33}$$

$$\propto \prod_{i=1}^{N} \left\{ \mathcal{BG}(f_i \mid \lambda, \tau_f^{-1}) \right\} \mathcal{N}_c(\boldsymbol{y} \mid \boldsymbol{H}\boldsymbol{f}, \tau_n^{-1}\mathbf{I}_N), \tag{34}$$

with M the number of non-zero elements in $\boldsymbol{f}$.

By contrast to the first three probabilities, the posterior of the force distribution $\boldsymbol{f}$ given in the last equation is too difficult to be sampled directly in its multivariate version. Therefore, the posteriors of the elements f_i, $i = 1, ..., N$ will be assessed one by one and in two steps by introducing a local sparsity variable λ_i and an activation variable for the Bernoulli process (q_i from Eq. (14)), similar to a power switch for each index of the spatial mesh.

3.7. Partially collapsed Gibbs sampler

For every node of the spatial mesh, the local sparsity variable (and thereafter, the value that the activation variable q_i is likely to take) is mainly driven by activation values at the other locations. So, f_i strongly depends on its neighbors. However, the mixing of the Gibbs sampler is especially low when the variables involved are strongly correlated because the conditional probabilities are then too narrow. Moreover, this trend is amplified when the dimensions of the target distribution increases (e.g. when the number of mesh nodes increases). Consequently, the Gibbs sampler with Bernoulli-Gaussian prior on the force distribution has extremely slow convergence. An acceleration strategy is proposed in Ref. [23] for the particular case where, instead of identifying a single pointwise source at x_i, the signal is spread out on the nearby nodes x_{i-1} and x_{i+1}. More broadly, Ref. [19] presents two methods to improve the convergence. The first one consists in a block Gibbs sampler to take into account locally the correlations of the target distribution. The second one is based on Refs. [53, 54] and proposes a partial marginalization of the Gibbs sampler, also called *Partially Collapsed Gibbs sampler*. The choice between the two strategies depends on dimensions of the data to work with. Globally, up to a thousand elements, the marginalization strategy is more efficient according to Ref. [19]. This approach is thus adopted in the present study. The concept is illustrated in Fig. 4 with a bi-dimensional distribution on two variables X and Y strongly correlated. While the standard algorithm samples successively from $[X \mid Y]$ and from $[Y \mid X]$, the marginalized one samples successively from $[Y \mid X]$ and from the marginal $[X] = \int [X \mid Y][Y] \mathrm{d}Y$. The marginal distribution can be seen as the projection of the joint distribution $[X, Y]$ on the axis of X. This step makes the Markov chains moving faster towards area of higher probability. The *partial* property of the marginalization is primordial to maintain the correlation. Indeed, marginalizing all conditional pdfs would be equivalent to considering totally independent variables.

This partial marginalization procedure is then applied to the previous Gibbs sampler with a Bernoulli-Gaussian prior on sources. The strategy is developed in appendix (B) and the algorithm 3 of appendix (C) sums up the procedure.

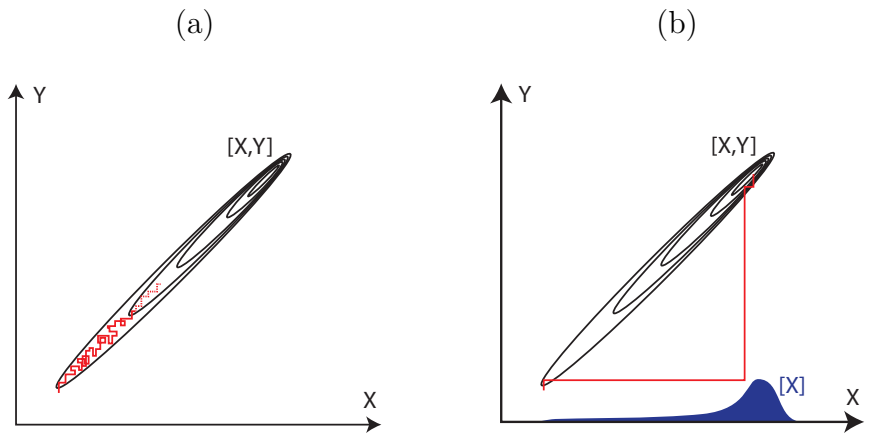

Figure 4: Illustration of the inference of an highly correlated two dimensional probability density function using the traditional Gibbs sampler (a) and the Partially Collapsed Gibbs sampler (b).

3.8. Whole sampling procedure

All the posteriors have now been set up and can be used in the Gibbs sampler for the identification of sparse vibratory sources. As in Ref. [4], the initialization of the algorithm is a key to ensure a good convergence towards the target multivariate probability. The different stochastic variables are thus initialized from results of the standard Gibbs sampler with a Gaussian prior [4]. In particular, the two precisions τ_n and τ_f are initialized with the modes of their Gamma distribution as estimated from the chain histograms. The force field $\boldsymbol{f}$ is initialized from the median of the converged chains. These preliminary results being obtained with a Gaussian prior on sources, the sparsity parameter is implicitly set to 1. The same value is then used to initialize λ in the sparse approach. To monitor the convergence of the chains, one could implement the FREP indicator of Ref. [55] based on the comparison of several parallel Gibbs samplers. If this is deemed too expensive, the chains can be simply stopped after a large number of iterations N_t. Algorithm 1 sums up the whole sampling process.

4. Numerical validations

This section illustrates the performance of the proposed Bernoulli-Gaussian Gibbs sampler through numerical simulations. The displacement field of an Euler beam is synthesized with added white noise to mimic

Algorithm 1 Global (partially collapsed) Gibbs sampler with Bernoulli-Gaussian prior on sources used for the identification of sparse vibration forces.

{Initialization :}
$\tau_n \leftarrow$ mode of Gibbs sampler with Gaussian priors [4]
$\tau_f \leftarrow$ mode of Gibbs sampler with Gaussian priors [4]
$\boldsymbol{f} \leftarrow$ median of Gibbs sampler with Gaussian priors [4]
$\lambda \leftarrow 1$ {Initializations of $\tau_f, \tau_n, \boldsymbol{f}$ are obtained with Gaussian prior, thus $\lambda = 1$ at first step}
$N_t \leftarrow$ {iterations number of the sampler}
$i_t \leftarrow 0$ {iterations count}
{Sampling:}
while $i_t < N_t$ **do**
    $\tau_n \leftarrow$ from Eq. (26)
    $\tau_f \leftarrow$ from Eq. (29)
    $\lambda \leftarrow$ from Eq. (32)
    $\boldsymbol{f} \leftarrow$ from algorithm (2) of section A.4 (without collapsing) or from algorithm 3 of section C (partial collapsing)
    $i_t \leftarrow i_t + 1$
end while

measured data. Several sequences of the Gibbs sampler are run with the aim to investigate the benefit of partial marginalization and the effect of the possible presence of model bias.

4.1. System under study

The simulated structure is an aluminum Euler beam whose structural and geometrical properties are listed in Tab. (1a). A harmonic excitation is applied at multiple points on the structure, with different amplitudes. Table (1b) resumes its properties.

4.2. Synthesis of displacement field and inverse crime

The displacement field of a simply supported beam excited by a pointwise harmonic vibratory source can be synthesized analytically, from the partial differential equations, or numerically, from the Finite Element

Length	Width	Thickness
L [m]	b [m]	h [m]
1	2×10^{-2}	2×10^{-3}
Second moment of inertia	Section	Spatial sampling
I [m^4]	S [m^2]	Δ_x [m]
$\frac{4}{3} \times 10^{-11}$	4×10^{-5}	1×10^{-2}
Young's modulus	Loss factor	Density
E [N/m^2]	η_E	ρ [kg/m^3]
70×10^9	1×10^{-3}	2700

(a)

Location	Amplitude	Frequency
x_0 [m]	A [N]	ν [Hz]
0.18	0.75	330
0.4	1	330
0.42	-1.5	330
0.66	-1	330

(b)

Table 1: (a) Structural and geometrical properties of the simulated beam. (b) Properties of the pointwise sources.

Method or finite differences. The latter method is followed in the present work, by means of the transfer matrix $\boldsymbol{H} \in \mathbb{C}^{101\times101}$ equal to the inverse of the discrete operator matrix $\boldsymbol{D}$ (see Eq. (4)). On the one hand, the analytic approach gives the best accuracy if it correctly describes the system under study, yet this applies only to a few trivial cases. On the other hand, the numerical approach is more flexible as it can address more complex situations (arbitrary geometry, heterogeneity, etc.), yet it may suffer from "numerical errors" such as those induced by discretization of spatial derivatives in Eq (2). In general, care should be taken in simulations to avoid committing *an inverse crime* [56]; this happens when the same operator is used both in the forward problem to synthesize data and in the construction of the inverse operator that is applied

on these same data. This process does not account for possible modeling errors – hereafter referred to as "model bias" – between the physical system and its representation by a model. Therefore, it is important in practice to check the robustness of an inverse approach to the presence of such errors. This is especially critical in the present work since the imposition of sparsity constraints may reduce the space of solutions to a region not included in the image of a biased inverse operator.

In order to the avoid an inverse crime, it is proposed to synthesize the displacement field by using the analytic model of the Euler beam and then to use the numerical operator $\boldsymbol{H}$ to solve the inverse problem. This approach will be compared to the inverse crime which consists in using the numerical operator $\boldsymbol{H}$ in both the direct and the inverse problems.

In the analytic model, the displacement field $w(x,\omega)$ of a simply supported Euler beam is given by [57]

$$w(x,\omega) = \begin{cases} \frac{A}{2k^3E(1+j\eta_E)I}\left(\sin(kx)\frac{\sin(k(L-x_0))}{\sin(kL)} - \sinh(kx)\frac{\sinh(k(L-x_0))}{\sinh(kL)}\right), \text{ for } x \in [0;x_0] \\ \frac{A}{2k^3E(1+j\eta_E)I}\left(\sin(k(L-x))\frac{\sin(kx_0)}{\sin(kL)} - \sinh(k(L-x))\frac{\sinh(kx_0)}{\sinh(kL)}\right), \text{ for } x \in [x_0;L] \end{cases}, \tag{35}$$

with L the beam length and $k = \sqrt[4]{\frac{\rho S\omega^2}{E(1+j\eta_E)I}}$ the flexural wavenumber controlled by the driving angular frequency of the source and structural and geometrical properties of the beam. Parameters $A \in \mathbb{C}$ and x_0 stand for the amplitude and the location of the pointwise source, respectively, whose values are reported in Tab. 1b. The displacement field is thus synthesized on a regular mesh in order to obtain the displacement vector $\boldsymbol{w} \in \mathbb{C}^{101\times 1}$, as defined in Eq. (4).

4.3. Investigation of model bias

Figure 5(a) compares the displacement fields obtained from the analytic and the numerical models. Although trends are similar to a certain extent, the two displacement fields are clearly different; this is perfectly consistent with the fact that the transfer matrix $\boldsymbol{H}$ used in the numerical model has been designed with virtual boundary conditions (completion of four missing rows in $\boldsymbol{D}$ to make it square before inversion) different from the real ones; in addition, approximation errors come with the discretization of spatial derivatives. However, for an inverse method to be robust, it is expected that these differences do not interfere too much in the inversion. This is verified in Fig. 5(b) which displays the source fields recovered by applying the inverse operator based on $\boldsymbol{H}$ on both the analytically and numerically synthesized data of Fig. 5(a) (without noise). It is seen that the two results are almost equal, even though the propagation of model bias in the inverse problem tends to slightly spread the reconstructed force distribution as compared to the ideal

Dirac impulses used in the simulation. This behavior is inherent to the finite difference scheme, which uses neighboring points to estimate the spatial derivatives (note that this effect also occurs if corrected finite difference schemes are considered, as proposed in [44]). The net force obtained after spatial integration in the neighborhood of each peak yet returns the correct magnitude in Newtons. This evidences that a slight loss in sparsity might be expected in practice due the unavoidable presence of modeling errors. A possible inability to distinguish very close sources is also anticipated, as observed near $x = 0.4$ m where the force distribution identified from the displacement field synthesized analytically shows a seemingly single source.

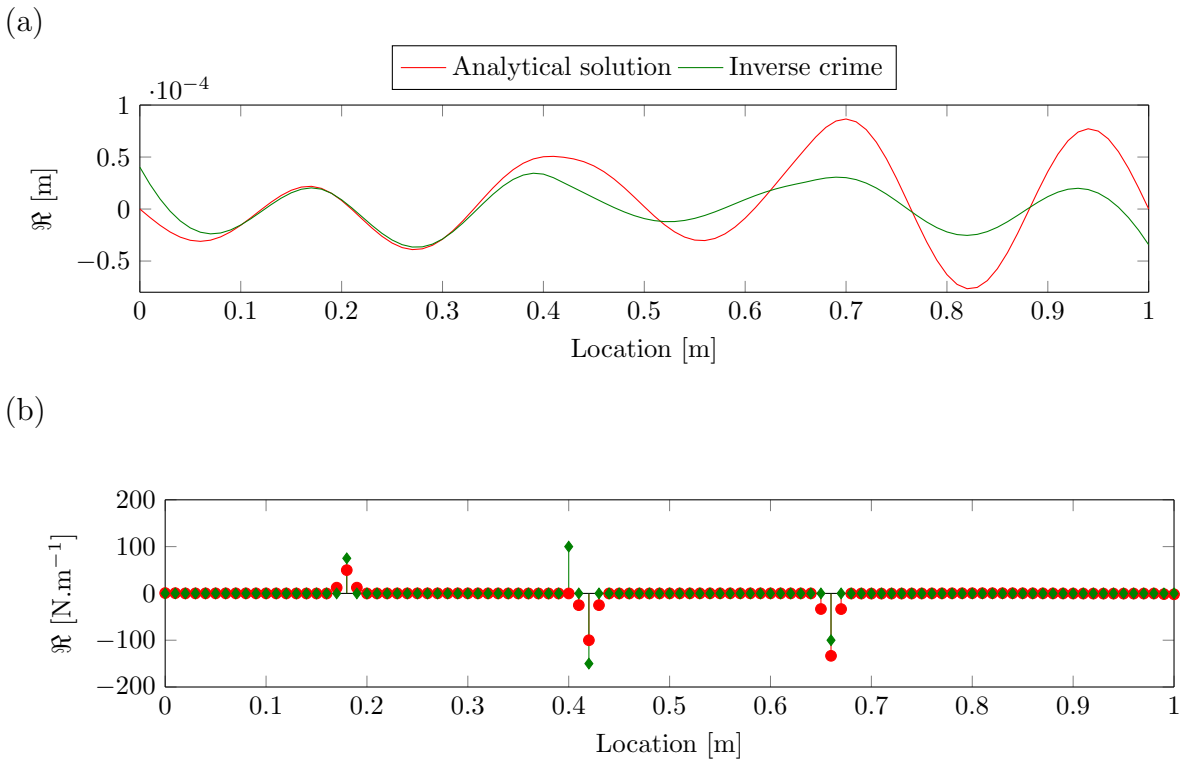

Figure 5: (a) Displacement field synthesized with the analytic model (partial differential equations – red curve) and the numerical model (transfer matrix $\boldsymbol{H}$ – green curve). (b) Reconstruction of the force distribution by applying the inverse numeric operator on the displacement fields synthesized analytically (red dots) and numerically (inverse crime – green diamonds). SNR $= \infty$.

4.4. Additive noise synthesis

Following Eq. (5), additive white noise

$$\boldsymbol{n} = \frac{(\boldsymbol{\epsilon}' + j\boldsymbol{\epsilon}'')}{\sqrt{2}} 10^{-\frac{\mathrm{SNR}}{20}} \tag{36}$$

is next added to the simulated measured displacement, with $\boldsymbol{\epsilon}'$ and $\boldsymbol{\epsilon}''$ two independent random vectors with the same size as $\boldsymbol{w}$ sampled from the standard normal distribution and SNR the signal-to-noise ratio in decibels (dB). The latter is set to 30 dB in the following simulations.

4.5. Effect of partial marginalization

Force identification is then undertaken by following the steps described in Algorithm 1. The Markov chains are initialized with the Gibbs sampler rooted on Gaussian priors of Ref. [4]. Hyperpriors are then tuned to be the least informative possible, so as to allow the likelihoods to bring most of the information. The hyperparameters of the Gamma priors on precisions are as follows: k_n and k_f tend towards 0 while θ_n and θ_f tend towards $+\infty$. The hyperparameters a_λ and b_λ of the Beta prior on the sparsity parameter are both set to 1, so that the probability is uniform on the $[0; 1]$ interval.

First, the traditional Gibbs sampler is performed on the displacement field with inverse crime (i.e. where $\boldsymbol{H}$ is used both in the direct and the inverse problems) and without partial marginalization. Figure 6 displays the Markov chains of the sparsity parameter λ (Fig. 6(a)), the noise precision τ_n (Fig. 6(b)), the source precision τ_f (Fig. 6(c)) and the source field $\boldsymbol{f}$ (Fig. 6(d)). The converged source field is estimated from the median of the second half of the chains. Estimations from the Gaussian (G) and the Bernoulli-Gaussian (BG) cases are presented in Fig. 6(e). It worth noting that the number of iterations is quite high (20000), yet the chains do not evolve so much. Due to the high correlation between variables in the sparse case, the algorithm is stuck in the initialization configuration. As a consequence, the sparsity parameter does not decrease under 0.6 and the Bernoulli-Gaussian prior does not seem to do better than the Gaussian prior.

Second, the Partially Collapsed Gibbs sampler is performed on the same displacement field, still with inverse crime (Fig. 7). A few tens of iterations are sufficient to make the chains converge to a sparse solution, the sparsity parameter being now below 0.1 and the source field being perfectly recovered in localization and in amplitude.

To give an indication of time consumption, the number of iterations has been approximately divided by 100 and the duration by 10 after adopting the partially collapsed strategy. The fast convergence (only 200

samples have been drawn) clearly compensates the higher cost induced by partial marginalization. It is therefore advocated to use the Partially Collapsed Gibbs sampler whenever highly sparse priors are imposed to the source reconstruction.

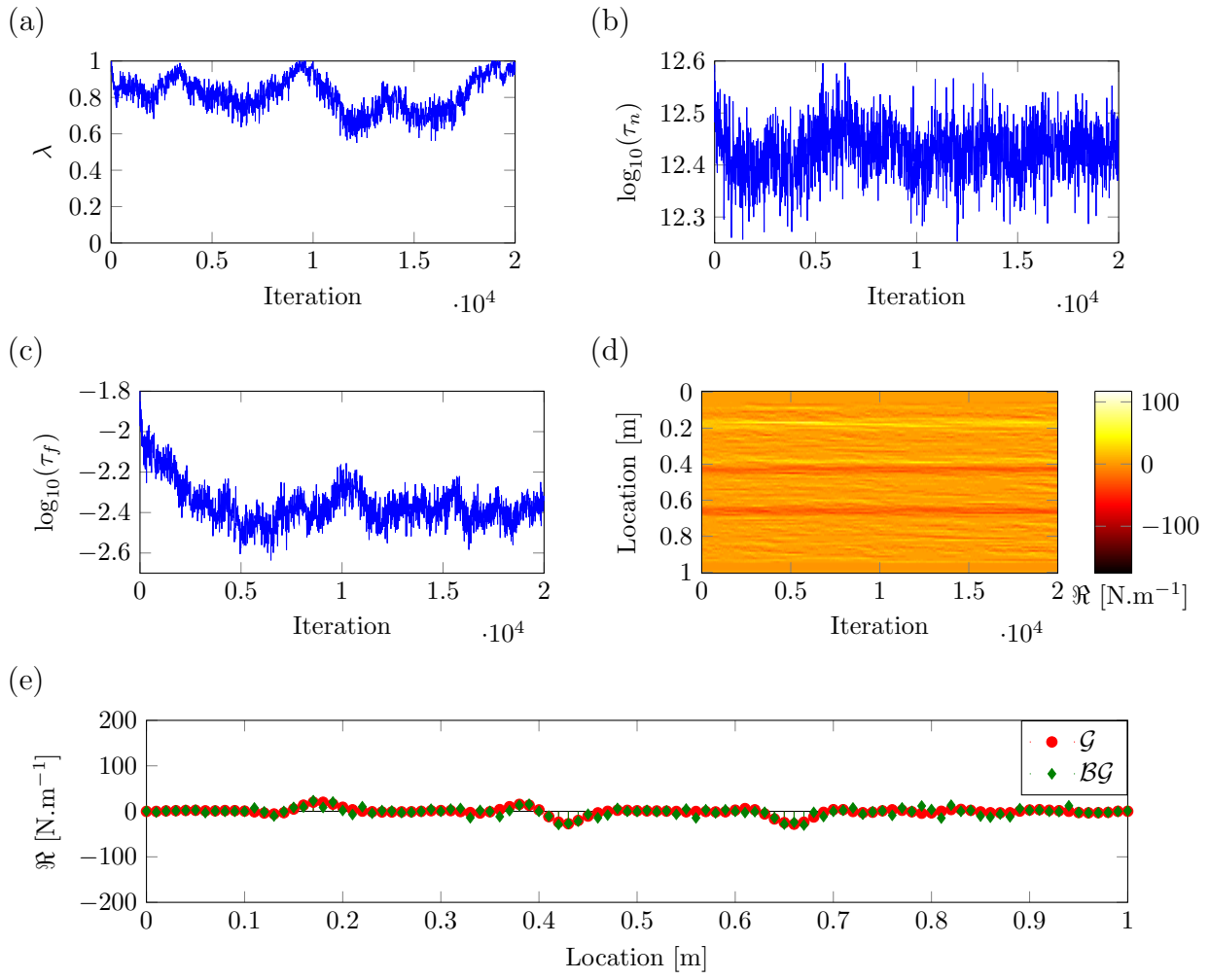

Figure 6: Results of the traditional Gibbs sampler with inverse crime synthesis under Gaussian (red dots) and Bernoulli-Gaussian (green diamonds) priors. Markov chains of sparsity parameter (a) λ, (b) noise precision τ_n, (c) source precision τ_f, (d) source spatial distribution $\boldsymbol{f}$ and (e) median values of the converged Markov chain of the source spatial distribution. SNR = 30 dB.

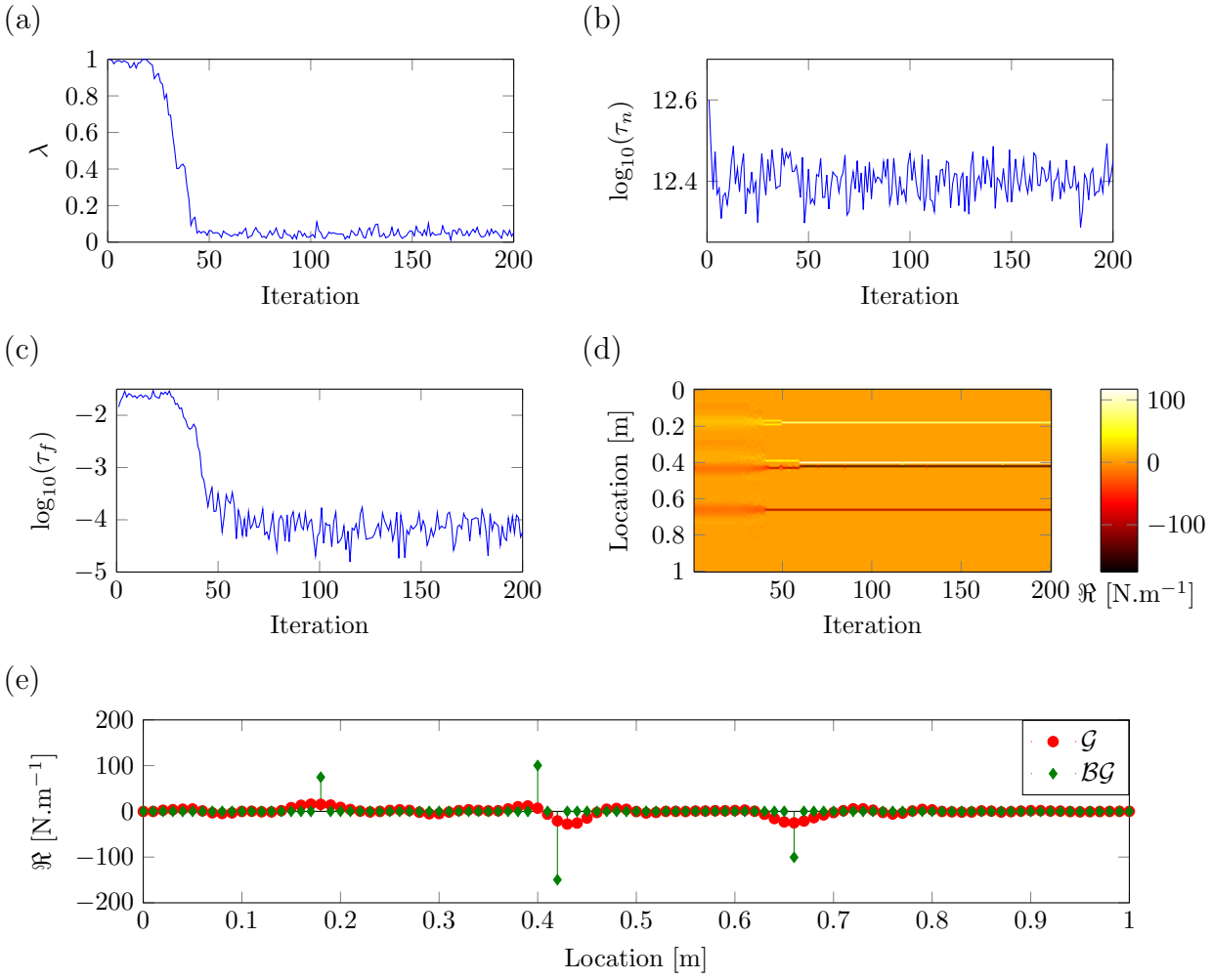

Figure 7: Results of the Partially Collapsed Gibbs sampler with inverse crime synthesis under Gaussian (red dots) and Bernoulli-Gaussian (green diamonds) priors. Markov chains of (a) sparsity parameter λ, (b) noise precision τ_n, (c) source precision τ_f, (d) source spatial distribution $\boldsymbol{f}$ and (e) median values of the converged Markov chain of the source spatial distribution. SNR = 30 dB.

4.6. Influence of model bias

As discussed before, committing an inverse crime is forbidden in the real world. In practice, the model used in the inversion is always biased in many ways (e.g. uncertainties on material and geometrical parameters, nonlinearities, spatial discretization, material fatigue, effect of temperature, etc.). It is therefore compulsory

to investigate the identification of sparse sources with a slightly erroneous model. For the direct problem, the displacement field is now calculated analytically from Eq. (35), while the transfer matrix $\boldsymbol{H}$ obtained from the inverse of the discrete structural operator of Eq. (3) is used in the inverse problem. The results of the Partially Collapsed Gibbs sampler are shown in Fig. 8. In this case, the Markov chains are much more volatile than before, especially the one corresponding to the noise precision τ_n (Fig. 8(b)). These losses of precision are probably linked to the high amplitude peaks seen on the boundaries of the beam on Fig. 8(d), to be interpreted as the search of cohesion forces at boundaries (see [42, 43]). A sparse solution is yet still recovered with accurate localization. It should be noted that the Bernoulli-Gaussian prior tends to counteract the spreading effect of the finite difference scheme. As a consequence, it reveals the existence of the two closely spaced pointwise sources near $x = 0.4$ m, in contrast to the sole application of the inverse operator to the exact displacement field synthesized analytically. It is noteworthy that the point source at $x_0 = 0.4$ m is largely underestimated ($\hat{f}(x_0) \simeq 20$ instead of 100 N/m), which implies an overestimation of the amplitude of the neighboring source at $x_0 = 0.42$ ($\hat{f}(x_0) \simeq -80$ instead of -150 N/m); therefore, the total force in Newton obtained after integration in the sub-domain $[0.4; 0.42]$ remains almost equal to its theoretical value of -0.5 N.

5. Experimental validation

This section reports a validation of the proposed method on experimental data, using the same test structure and measurement procedure as in [4]. It should be noted on the onset that, in this case, the model bias can not be assessed (since the ground truth is unknown) and might be higher than in the previous section.

A freely suspended aluminum beam, with dimensions 72.25 cm $\times$ 2.96 cm $\times$ 2.9 mm, Young's modulus $E = 70 \times 10^9$ N/m^2, mass density $\rho = 2700$ kg/m^3 and structural damping $\eta = 10^{-4}$, is excited by a shaker located at $x_0 = 0.37$ m. An impedance head is placed at the interface to acquire a reference signal. The excitation signal is a linear chirp in the frequency range $[100; 4000]$ Hz and the displacement measurement is performed by a scanning laser vibrometer with a mesh of 105 nodes and a constant spatial discretization of $\Delta_x = 5.7$ mm. Pictures in Fig. 9 show the experimental setup.

The partially collapsed Gibb's sampler is then performed at an arbitrary frequency of 1505 Hz with both Gaussian (see [4]) and Bernoulli-Gaussian priors. Here again, *prior* parameters are set to be as non-informative as possible. Figure 10(a) shows the measured displacement field $\boldsymbol{y}$ while Fig. 10(b) displays the

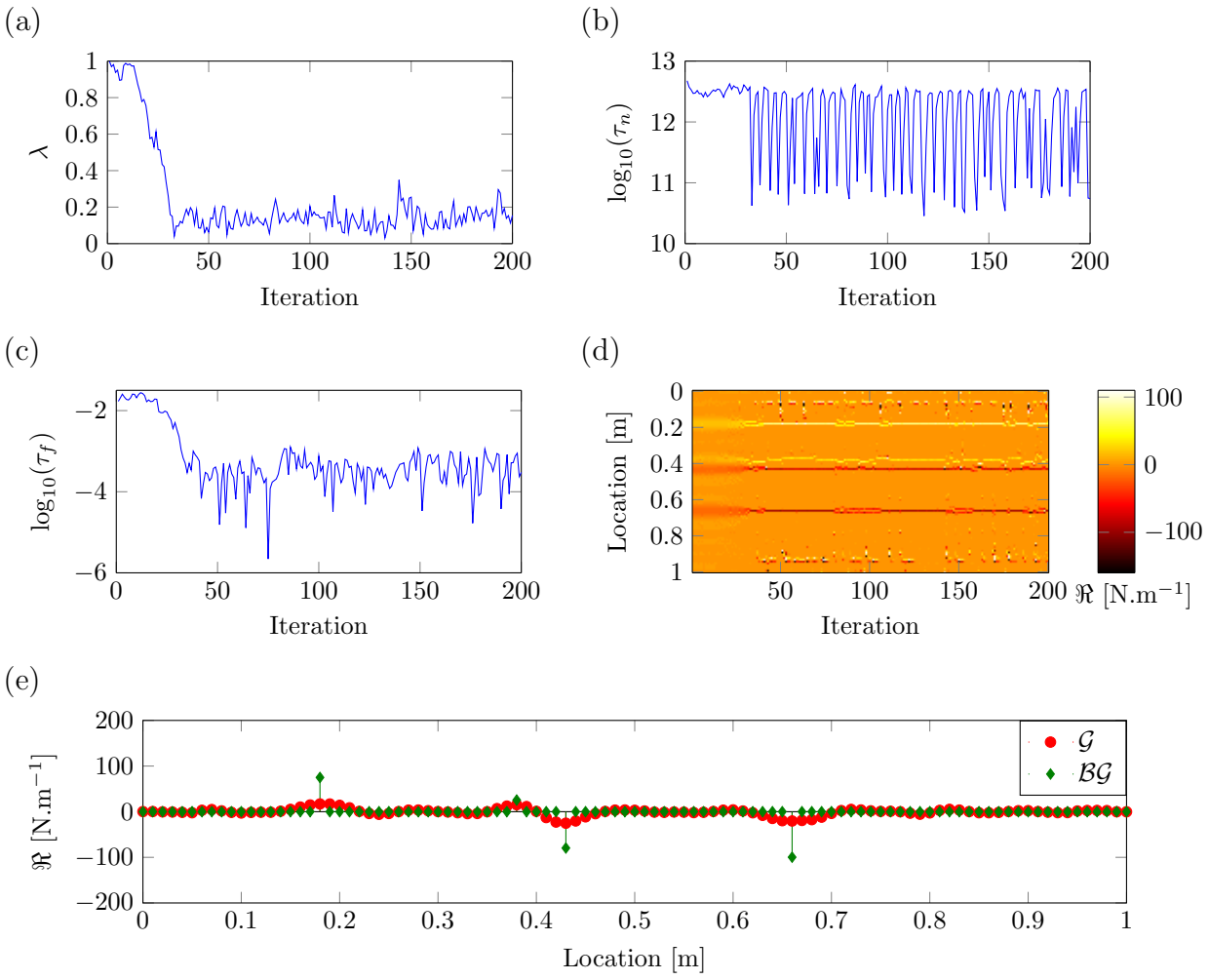

Figure 8: Results of the Partially Collapsed Gibbs sampler with analytic synthesis under Gaussian (red dots) and Bernoulli-Gaussian priors (green diamonds). Markov chains of (a) sparsity parameter λ, (b) noise precision τ_n, (c) source precision τ_f, (d) source spatial distribution $\boldsymbol{f}$ and (e) median values of the converged Markov chain of the source spatial distribution. SNR = 30 dB.

corresponding reconstructed source field from the Gibbs sampler with Gaussian (G) and sparse (BG) priors. The Bernoulli-Gaussian solution shows a sparse solution with non-zero sources only between 0.34 and 0.37 cm, whereas the Gaussian solution is non-zero almost everywhere in the domain. It is noteworthy that a perfectly pointwise source is not expected in this case due to the non-negligible surface of the impedance head that transmits the excitation force from the shaker to the beam. This, plus the fact that the presence

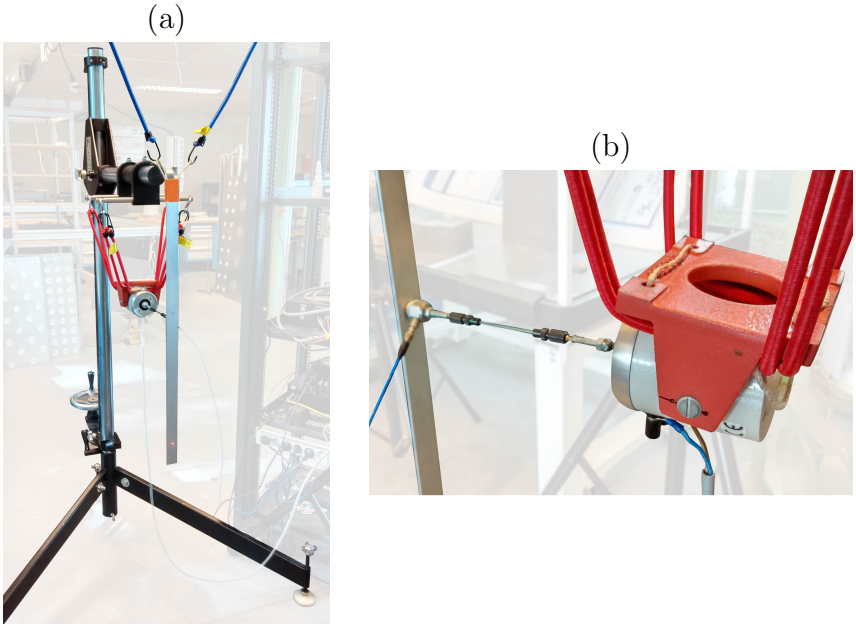

Figure 9: Experimental setup of the free beam excited by a shaker with a linear chirp signal in frequency range $[100; 4000]$ Hz. An impedance head is placed at the interface between the beam and the shaker to acquire a reference signal.

of model bias is likely to spread the reconstructed force distribution as evidenced in the previous section, indicate fairly good results. The presence of two source coefficients with reversed signs at the border of the support set at $x \simeq 0.37$ cm may be due to a slight torque applied by the impedance head resulting from an imperfect excitation apparatus.

The evolution of sparsity parameter λ over the considered frequency range is shown in Figure 11(a). It has a mean value of about 0.3, which is higher than the value obtained from simulations involving several point forces (see Fig 8 (a)). This is an indication that the source applied here does not behave like an ideal point source, for the reasons explained above. Besides, the increase in λ between 2500 and 3000 Hz is associated with the presence of residual forces in the unexcited regions of the beam. It should be noted that the beam model used in the inverse problem does not take into account torsional motion, which exists in practice and can lead to errors in the force identification in some frequency ranges.

Finally, the reconstructed source field is spatially integrated to obtain the force in Newton injected into the structure to allow its comparison with the reference signal of the impedance head. Figure 11(b) compares

the force spectrum obtained with Gaussian and Bernoulli-Gaussian priors to the reference force spectrum provided by the sensor. Overall, the results obtained from the two priors are in agreement, except at low frequencies (below 300 Hz) where the identification with Bernoulli-Gaussian prior leads to wrong values. Regarding the agreement with the reference spectrum, the identification with Gaussian prior is slightly more advantageous. These differences can be explained by the the greater complexity of the sampling algorithm required to perform the identification with Bernoulli-Gaussian prior, which allows a more accurate localization of the source than the Gaussian prior, at the cost of a less accurate estimation of its amplitude.

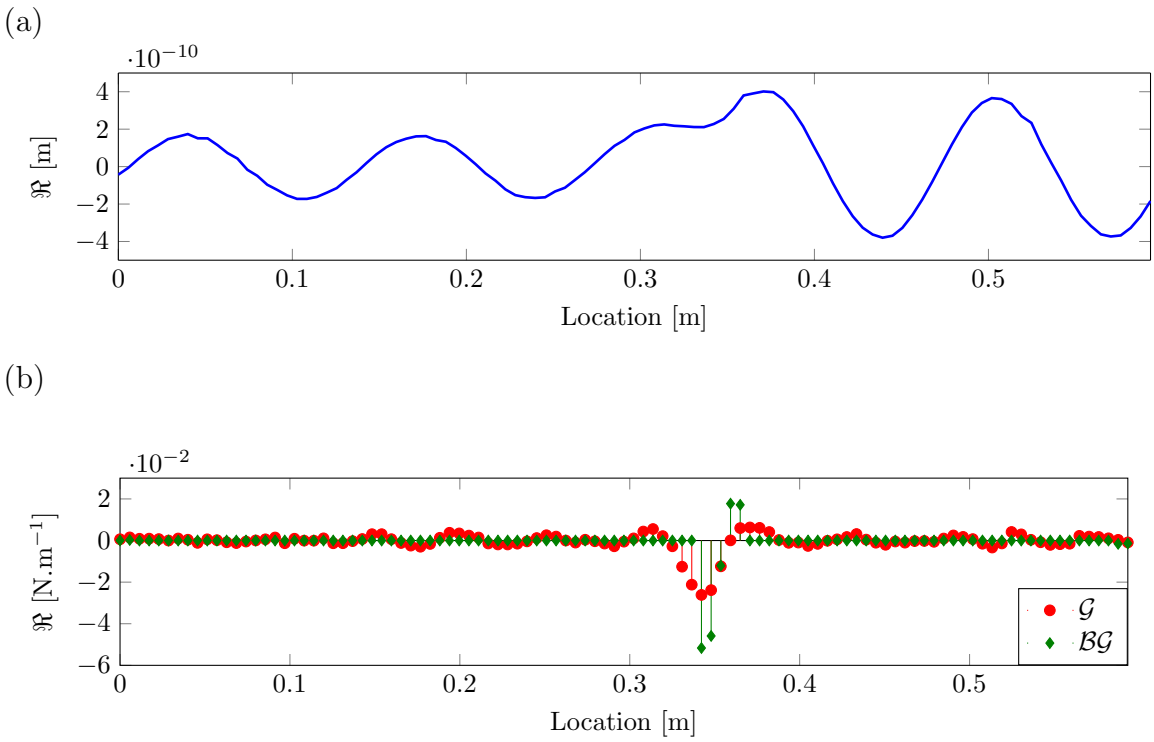

Figure 10: (a) Experimental displacement field at 1505 Hz of a freely suspended beam excited by a pointwise vibratory source. (b) Corresponding experimental source identification with Gaussian priors (G) and Bernoulli-Gaussian priors (BG).

6. Conclusion

This paper has shown how sparsity can be enforced in the recovery of a vibratory source field by following a fully Bayesian approach rooted on an appropriate prior such as the Bernoulli-Gaussian. This has for

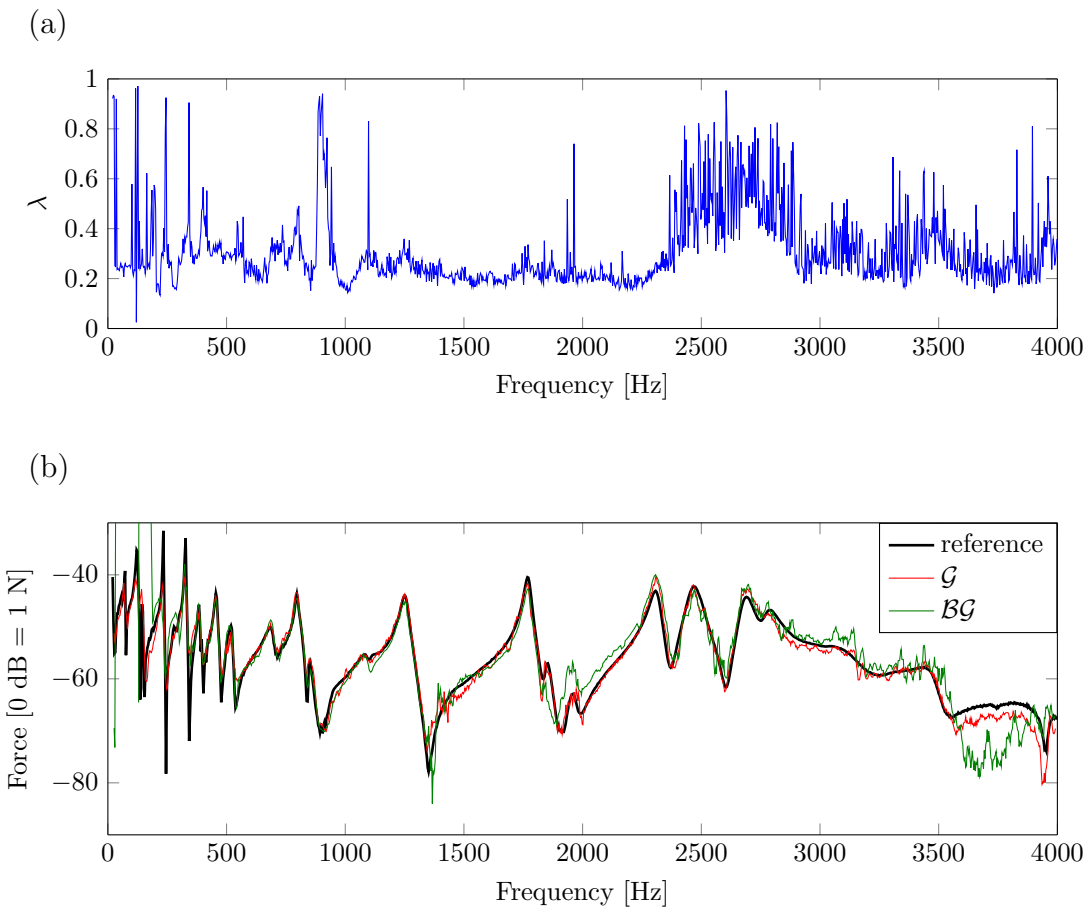

Figure 11: (a) Sparsity parameter λ as a function of frequency. (b) Source level identified with Gaussian priors (G) and Bernoulli-Gaussian priors (BG), compared to the reference signal provided by the force sensor.

consequence to return an almost nil field, except where pointwise sources are actually applied to the vibrating structure, contrary to the usual Gaussian prior which tends to spread non-zero values over the full domain with significant magnitude even away from the excitation points. A much finer spatial resolution is thus achieved, which allows for the separation of closely spaced pointwise sources. Torque excitation can eventually be identified and quantified as it is interpreted by the FAT as two opposite forces placed side by side.

However, the price to pay is that the mathematical developments are more complex with the Bernoulli-Gaussian prior than with the traditional Gaussian prior. The numerical cost is also more expensive despite the fact the Bernoulli-Gaussian prior fits into the conjugate prior strategy. A Partially Collapsed Gibbs

sampler has been proposed to improve the exploration of the posterior probabilities and therefore reduce the numerical cost. Other priors, maybe less sparse but easier to implement, could also be used, such as other Gaussian mixture models. At the same time, a promising prospect concerns the coupling of optimization strategies and Bayesian inference. The numerical cost could thus be significantly reduced while keeping the probabilistic point of view [58, 59].

At last, an important observation made in this work is that the the presence of model bias is likely to deteriorate the identification of sparse sources. Consequently, it would be useful to exploit the assumption of sparsity to jointly update the model and reconstruct the sources [26].

References

References

[1] C. Pezerat, T. Loyau, J. Guyader, Characterisation of vibration sources on a set of plates using the riff technique, Noise Control Engineering Journal 50 (2) (2002) 50–57.

[2] J.-L. Starck, F. Murtagh, J. Fadili, Sparse image and signal processing: Wavelets and related geometric multiscale analysis, Cambridge university press, 2015.

[3] C. Pezerat, J. Guyader, Force analysis technique: reconstruction of force distribution on plates, Acta Acustica united with Acustica 86 (2) (2000) 322–332.

[4] C. Faure, F. Ablitzer, J. Antoni, C. Pézerat, Empirical and fully bayesian approaches for the identification of vibration sources from transverse displacement measurements, Mechanical Systems and Signal Processing 94 (2017) 180–201.

[5] Y. Liu, W. S. Shepard, An improved method for the reconstruction of a distributed force acting on a vibrating structure, Journal of Sound and Vibration 291 (1) (2006) 369 – 387. doi:https://doi.org/10.1016/j.jsv.2005.06.013.
URL http://www.sciencedirect.com/science/article/pii/S0022460X05004050

[6] J. Sanchez, H. Benaroya, Review of force reconstruction techniques, Journal of Sound and Vibration 333 (14) (2014) 2999 – 3018. doi:https://doi.org/10.1016/j.jsv.2014.02.025.
URL http://www.sciencedirect.com/science/article/pii/S0022460X14001527

[7] B. Qiao, X. Zhang, X. Luo, X. Chen, A force identification method using cubic b-spline scaling functions, Journal of Sound and Vibration 337 (2015) 28 – 44. doi:https://doi.org/10.1016/j.jsv.2014.09.038.
URL http://www.sciencedirect.com/science/article/pii/S0022460X14007901

[8] B. Qiao, X. Zhang, C. Wang, H. Zhang, X. Chen, Sparse regularization for force identification using dictionaries, Journal of Sound and Vibration 368 (2016) 71 – 86. doi:https://doi.org/10.1016/j.jsv.2016.01.030.
URL http://www.sciencedirect.com/science/article/pii/S0022460X16000596

[9] C. Févotte, S. J. Godsill, P. J. Wolfe, Bayesian approach for blind separation of underdetermined mixtures of sparse sources, in: International Conference on Independent Component Analysis and Signal Separation, Springer, 2004, pp. 398–405.

[10] T. Park, G. Casella, The bayesian lasso, Journal of the American Statistical Association 103 (482) (2008) 681–686.

[11] L. Chaâri, J.-C. Pesquet, J.-Y. Tourneret, P. Ciuciu, A. Benazza-Benyahia, A hierarchical bayesian model for frame representation, IEEE Transactions on Signal Processing 58 (11) (2010) 5560–5571.

[12] Q. Leclere, A. Pereira, J. Antoni, Une approche bayésienne de la parcimonie pour l'identification de sources acoustiques, Proceedings of CFA (2014).

[13] R. Chartrand, W. Yin, Iteratively reweighted algorithms for compressive sensing, in: 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 2008, pp. 3869–3872.

[14] T. Auranen, A. Nummenmaa, M. S. Hämäläinen, I. P. Jääskeläinen, J. Lampinen, A. Vehtari, M. Sams, Bayesian analysis of the neuromagnetic inverse problem with ℓ_p-norm priors, NeuroImage 26 (3) (2005) 870–884.

[15] Q. Li, N. Lin, et al., The bayesian elastic net, Bayesian Analysis 5 (1) (2010) 151–170.

[16] G. Yu, G. Sapiro, S. Mallat, Solving inverse problems with piecewise linear estimators: From gaussian mixture models to structured sparsity, IEEE Transactions on Image Processing 21 (5) (2012) 2481–2499.

[17] J. Kormylo, J. Mendel, Maximum likelihood detection and estimation of bernoulli-gaussian processes, IEEE transactions on information theory 28 (3) (1982) 482–488.

[18] N. Dobigeon, A. O. Hero, J.-Y. Tourneret, Hierarchical bayesian sparse image reconstruction with application to mrfm, IEEE Transactions on Image Processing 18 (9) (2009) 2059–2070.

[19] D. Ge, J. Idier, E. Le Carpentier, Enhanced sampling schemes for mcmc based blind bernoulli–gaussian deconvolution, Signal Processing 91 (4) (2011) 759–772.

[20] N. Dobigeon, A. Basarab, D. Kouamé, J.-Y. Tourneret, Regularized bayesian compressed sensing in ultrasound imaging, in: Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European, IEEE, 2012, pp. 2600–2604.

[21] F. Costa, H. Batatia, L. Chaari, J.-Y. Tourneret, Sparse eeg source localization using bernoulli laplacian priors, IEEE Transactions on Biomedical Engineering 62 (12) (2015) 2888–2898.

[22] Q. Cheng, R. Chen, T.-H. Li, Simultaneous wavelet estimation and deconvolution of reflection seismic signals, IEEE Transactions on Geoscience and Remote Sensing 34 (2) (1996) 377–384.

[23] S. Bourguignon, H. Carfantan, Bernoulli-gaussian spectral analysis of unevenly spaced astrophysical data, in: IEEE/SP 13th Workshop on Statistical Signal Processing, 2005, IEEE, 2005, pp. 811–816.

[24] V. Mazet, Développement de méthodes de traitement de signaux spectroscopiques: estimation de la ligne de base et du spectre de raies, Ph.D. thesis, Université Henri Poincaré-Nancy I (2005).

[25] L. Chaari, J.-Y. Tourneret, H. Batatia, Sparse bayesian regularization using bernoulli-laplacian priors, in: 21st European Signal Processing Conference (EUSIPCO 2013), IEEE, 2013, pp. 1–5.

[26] L. Chaari, H. Batatia, J.-Y. Tourneret, Sparse bayesian image restoration with linear operator uncertainties with application to eeg signal recovery, in: 2nd Middle East Conference on Biomedical Engineering, IEEE, 2014, pp. 139–142.

[27] L. Chaari, H. Batatia, N. Dobigeon, J.-Y. Tourneret, A hierarchical sparsity-smoothness bayesian model for $\ell_0+\ell_1+\ell_2$ regularization, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2014, pp. 1901–1905.

[28] E. Z. Moore, K. D. Murphy, J. M. Nichols, Crack identification in a freely vibrating plate using bayesian parameter estimation, Mechanical Systems and Signal Processing 25 (6) (2011) 2125 – 2134, interdisciplinary Aspects of Vehicle Dynamics. doi:https://doi.org/10.1016/j.ymssp.2011.01.016.
URL http://www.sciencedirect.com/science/article/pii/S0888327011000264

[29] B. Li, A. D. Kiureghian, Operational modal identification using variational bayes, Mechanical Systems and Signal Processing 88 (2017) 377 – 398. doi:https://doi.org/10.1016/j.ymssp.2016.11.007.
URL http://www.sciencedirect.com/science/article/pii/S0888327016304654

[30] Y. Huang, J. L. Beck, H. Li, Bayesian system identification based on hierarchical sparse bayesian learning and gibbs sampling with application to structural damage assessment, Computer Methods in Applied Mechanics and Engineering 318 (2017) 382 – 411. doi:https://doi.org/10.1016/j.cma.2017.01.030.
URL http://www.sciencedirect.com/science/article/pii/S0045782516305060

[31] T. Baldacchino, K. Worden, J. Rowson, Robust nonlinear system identification: Bayesian mixture of experts using the t-distribution, Mechanical Systems and Signal Processing 85 (2017) 977 – 992. doi:https://doi.org/10.1016/j.ymssp.2016.08.045.
URL http://www.sciencedirect.com/science/article/pii/S0888327016303302

[32] K.-V. Yuen, S.-C. Kuok, Bayesian methods for updating dynamic models, Applied Mechanics Reviews 64 (2011) 010802–18.

[33] E. Zhang, P. Feissel, J. Antoni, A comprehensive bayesian approach for model updating and quantification of modeling errors, Probabilistic Engineering Mechanics 26 (4) (2011) 550 – 560. doi:https://doi.org/10.1016/j.probengmech.2011.07.001.
URL http://www.sciencedirect.com/science/article/pii/S0266892011000439

[34] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou, Hierarchical bayesian model updating for structural identification, Mechanical Systems and Signal Processing 64-65 (2015) 360 – 376. doi:https://doi.org/10.1016/j.ymssp.2015.03.026.
URL http://www.sciencedirect.com/science/article/pii/S0888327015001545

[35] H.-P. Wan, W.-X. Ren, Stochastic model updating utilizing bayesian approach and gaussian process model, Mechanical Systems and Signal Processing 70-71 (2016) 245 – 268. doi:https://doi.org/10.1016/j.ymssp.2015.08.011.
URL http://www.sciencedirect.com/science/article/pii/S0888327015003726

[36] S. H. Cheung, S. Bansal, A new gibbs sampling based algorithm for bayesian model updating with incomplete complex modal data, Mechanical Systems and Signal Processing 92 (2017) 156 – 172. doi:https://doi.org/10.1016/j.ymssp.2017.01.015.
URL http://www.sciencedirect.com/science/article/pii/S0888327017300183

[37] E. Zhang, J. Antoni, P. Feissel, Bayesian force reconstruction with an uncertain model, Journal of Sound and Vibration 331 (4) (2012) 798 – 814. doi:https://doi.org/10.1016/j.jsv.2011.10.021.
URL http://www.sciencedirect.com/science/article/pii/S0022460X1100825X

[38] M. Aucejo, O. D. Smet, Bayesian source identification using local priors, Mechanical Systems and Signal Processing 66-67 (2016) 120 – 136. doi:https://doi.org/10.1016/j.ymssp.2015.05.004.
URL http://www.sciencedirect.com/science/article/pii/S0888327015002253

[39] M. Aucejo, O. D. Smet, On a full bayesian inference for force reconstruction problems, Mechanical Systems and Signal Processing 104 (2018) 36 – 59. doi:https://doi.org/10.1016/j.ymssp.2017.10.023.
URL http://www.sciencedirect.com/science/article/pii/S0888327017305538

[40] Q. Li, Q. Lu, A hierarchical bayesian method for vibration-based time domain force reconstruction problems, Journal of Sound and Vibration 421 (2018) 190 – 204. doi:https://doi.org/10.1016/j.jsv.2018.01.052.
URL http://www.sciencedirect.com/science/article/pii/S0022460X18300749

[41] S. Samagassi, A. Khamlichi, A. Driouach, E. Jacquelin, Reconstruction of multiple impact forces by wavelet relevance vector machine approach, Journal of Sound and Vibration 359 (2015) 56 – 67. doi:https://doi.org/10.1016/j.jsv.2015.08.014.
URL http://www.sciencedirect.com/science/article/pii/S0022460X15006653

[42] C. Renzi, C. Pezerat, J.-L. Guyader, Vibratory source identification by using the finite element model of a subdomain of a flexural beam, Journal of Sound and Vibration 332 (3) (2013) 545–562.

[43] M. Aucejo, Structural source identification using a generalized tikhonov regularization, Journal of Sound and Vibration 333 (22) (2014) 5693–5707.

[44] Q. Leclère, C. Pézerat, Vibration source identification using corrected finite difference schemes, Journal of Sound and Vibration 331 (6) (2012) 1366–1377.

[45] A. N. Tikhonov, V. Y. Arsenin, Solutions of ill-posed problems (1977).

[46] P. C. Hansen, Analysis of discrete ill-posed problems by means of the l-curve, SIAM review 34 (4) (1992) 561–580.

[47] Y. Liu, W. S. Shepard, Dynamic force identification based on enhanced least squares and total least-squares schemes in the frequency domain, Journal of sound and vibration 282 (1) (2005) 37–60.

[48] M. Aucejo, O. De Smet, Bayesian source identification using local priors, Mechanical Systems and Signal Processing 66 (2016) 120–136.

[49] G. H. Golub, M. Heath, G. Wahba, Generalized cross-validation as a method for choosing a good ridge parameter, Technometrics 21 (2) (1979) 215–223.

[50] J. Rice, Mathematical statistics and data analysis, Nelson Education, 2006.

[51] H. Raiffa, Applied statistical decision theory (1974).

[52] W. M. Bolstad, Understanding computational Bayesian statistics, John Wiley & Sons, 2010.

[53] D. A. Van Dyk, T. Park, Partially collapsed gibbs samplers: Theory and methods, Journal of the American Statistical Association 103 (482) (2008) 790–796.

[54] T. Park, D. A. van Dyk, Partially collapsed gibbs samplers: Illustrations and applications, Journal of Computational and Graphical Statistics 18 (2) (2009) 283–305.

[55] A. Gelman, D. B. Rubin, Inference from iterative simulation using multiple sequences, Statistical science (1992) 457–472.

[56] A. Wirgin, The inverse crime, arXiv preprint math-ph/0401050 (2004).

[57] J.-L. Guyader, Vibration in continuous media, John Wiley & Sons, 2013.

[58] S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, Hybrid monte carlo, Physics letters B 195 (2) (1987) 216–222.

[59] M. Pereyra, Proximal markov chain monte carlo algorithms, Statistics and Computing 26 (4) (2016) 745–760.

[60] N. J. Higham, Accuracy and stability of numerical algorithms, Siam, 2002.

[61] F. Champagnat, Y. Goussard, J. Idier, Unsupervised deconvolution of sparse spike trains using stochastic approximation, IEEE Transactions on Signal Processing 44 (12) (1996) 2988–2998.

Appendices

A. Conditional probabilities

Expressions (24) to (34) relating to conditional probabilities are detailed in this section.

A.1. Conditional distribution on τ_n

The noise precision τ_n is the only one stochastic variable which does not depend on the prior of $\boldsymbol{f}$. Its posterior is given by

$$\begin{aligned}
[\tau_n \mid k_n, \theta_n, \boldsymbol{y}, \boldsymbol{H}, \boldsymbol{f}] &\propto \mathcal{G}(\tau_n \mid k_n, \theta_n)\, \mathcal{N}_c(\boldsymbol{y} \mid \boldsymbol{H}\boldsymbol{f}, \tau_n^{-1}\mathbf{I}_N) && (37)\\
&\propto \mathcal{G}\left(\tau_n \mid k_n + N, \left(\theta_n^{-1} + \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{f}\|_2^2\right)^{-1}\right), && (38)
\end{aligned}$$

with N the number of elements in vector $\boldsymbol{y}$.

A.2. Conditional distribution on τ_f

A change of probability on the force distribution modifies the update of the posterior parameters on τ_f. Here, Bayes' theorem yields

$$\begin{aligned}
[\tau_f \mid k_f, \theta_f, \lambda, \boldsymbol{f}] &\propto \mathcal{G}(\tau_f \mid k_f, \theta_f) \prod_{i=1}^{N} \left\{ \mathcal{BG}(f_i \mid \lambda, \tau_f^{-1}) \right\} && (39)\\
&\propto \frac{\tau_f^{k_f - 1} \exp\left(-\frac{\tau_f}{\theta_f}\right)}{\theta_f^{k_f}\ \Gamma(k_f)} \prod_{i=1}^{N} \left\{ (1-\lambda)\delta(f_i) + \lambda\ \frac{\exp\left(-\tau_f \|f_i\|_2^2\right)}{\pi \tau_f^{-1}} \right\} && (40)\\
&\propto \mathcal{G}(\tau_f \mid k'_f, \theta'_f). && (41)
\end{aligned}$$

The Dirac distribution at 0 is independent of τ_f, so the nil elements of the source vector $\boldsymbol{f}$ multiplied by the Gamma distribution do not modify the parameters (k_f, θ_f). They simply act as multiplicative constant and can then be neglected in the proportional relation.

However, when an element of $\boldsymbol{f}$ is non-zero, the Gaussian prior on $\boldsymbol{f}$ (acting as the likelihood here for the inference on τ_f) depends on τ_f and the product with the Gamma distribution updates its parameters. The posterior parameters are then updated using the following rule,

$$\begin{aligned}
k'_f &= k_f + M, && (42)\\
\theta'_f &= \left(\frac{1}{\theta_f} + \sum_{f_i \neq 0} f_i^H f_i \right)^{-1} = \left(\theta_f^{-1} + \|\boldsymbol{f}\|_2^2 \right)^{-1}, && (43)
\end{aligned}$$

with M the quasi-norm ℓ_0 of $\boldsymbol{f}$ (in other words, the number of non-zero elements of the vector). Thus, precision on sources in this context is only affected by non-zero elements of the force distribution. Then, the precision τ_f might be completely different from the one obtained with a classical Gaussian prior on $\boldsymbol{f}$

taking into account the whole force distribution. Therefore, the regularization parameter does not simply correspond to the ratio between noise and source precisions, as is the case with Gaussian priors on both noise and sources.

A.3. Conditional distribution on λ

The posterior on the sparsity parameter is obtained by noting that the Bernoulli process is conditionally dependent on sources only, regardless of the Gaussian process steered by the precision τ_f. By using the activation variables vector $\boldsymbol{q}$,

$$\begin{aligned}
[\lambda \mid a_\lambda, b_\lambda, \boldsymbol{f}, \tau_f] &= [\lambda \mid a_\lambda, b_\lambda, \boldsymbol{q}] \\
&\propto \mathcal{B}(\lambda \mid a_\lambda, b_\lambda) \prod_i \{[q_i \mid \lambda]\} && (44) \\
&\propto \frac{\Gamma(a_\lambda + b_\lambda)}{\Gamma(a_\lambda)\,\Gamma(b_\lambda)} \lambda^{a_\lambda - 1}(1-\lambda)^{b_\lambda - 1} \prod_i \left\{\lambda^{q_i} \cdot (1-\lambda)^{1-q_i} \cdot \mathbb{1}_{\{0,1\}}(q_i)\right\} && (45) \\
&\propto \mathcal{B}(\lambda \mid a'_\lambda, b'_\lambda), && (46)
\end{aligned}$$

with the updated parameters identified by

$$\begin{aligned}
a'_\lambda &= a_\lambda + \textstyle\sum_i q_i &&= a_\lambda + M, && (47) \\
b'_\lambda &= b_\lambda + \textstyle\sum_i (1 - q_i) &&= b_\lambda + N - M. && (48)
\end{aligned}$$

This proves Eq. (32).

The sum of nil elements updating parameter a is equivalent by construction to the quasi-norm ℓ_0 given by M. Logically, the sum of non-zero elements updating parameter b is $N - M$. Thus, when q_i is equal to 0, b is increased by $+1$ and a remains the same. When q_i is equal to 1, a is increased by $+1$ and b remains the same. The update of these parameters is then especially simple.

A.4. Conditional distribution on f

First, parameters of the Gaussian process associated to the amplitude are assessed. Then, from them, parameters of the Bernoulli process on the activation variable are assessed in turn.

Distinguishing each f_i in the same way as in Eq. (13) is equivalent to considering several univariate variables instead of the single vectorial quantity $\boldsymbol{f}$. Hence, in the hierarchical model of Fig. 3, the stochastic node on

$\boldsymbol{f}$ is replaced by N stochastic nodes f_i, all connected to variables λ, τ_f and $\boldsymbol{y}$. Thus for each node f_i, all the other nodes, merged in a set hereafter denoted $\boldsymbol{f}^{(-i)}$, are co-parent nodes and take place in the identification of posterior on f_i, through the likelihood. This posterior is then written as,

$$\left[f_i \mid \lambda, \tau_f, \boldsymbol{y}, \boldsymbol{H}, \boldsymbol{f}^{(-i)}, \tau_n\right] \propto \mathcal{BG}(f_i \mid \lambda, \tau_f^{-1}) \mathcal{N}_c(\boldsymbol{y} \mid \boldsymbol{H}\boldsymbol{f}, \tau_n^{-1}\mathbf{I}_N) \tag{49}$$

$$\propto \left\{(1-\lambda)\delta(f_i) + \lambda \frac{\exp\left(-\tau_f \|f_i\|_2^2\right)}{\pi \tau_f^{-1}}\right\} \frac{\exp\left(-\tau_n \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{f}\|_2^2\right)}{\pi^N \tau_n^{-N}} \tag{50}$$

$$\propto \sum_{l=0,1} \left\{\lambda_l \frac{\exp(-Q_{l,i})}{\pi\ \tau_l^{-1}}\right\}, \tag{51}$$

wherein

$$\lambda_0 = 1-\lambda, \tag{52}$$

$$\lambda_1 = \lambda, \tag{53}$$

$$\tau_0 = +\infty \text{ (symbolic notation for } \delta(x) = \lim_{\tau\to\infty} \mathcal{N}_c(x \mid 0, \tau^{-1})), \tag{54}$$

$$\tau_1 = \tau_f, \tag{55}$$

$$Q_{l,i} = \tau_n \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{f}\|_2^2 + \tau_l \|f_i\|_2^2,\ l = 0, 1. \tag{56}$$

The subscript l refers then to the Dirac distribution when $l = 0$, and to the Gaussian distribution when $l = 1$. The term $\pi^N\ \tau_n^{-N}$ of Eq. (50) is a same constant applied both on the Dirac distribution and the Gaussian distribution, so it is independent of l and can be extracted from this sum. Thanks to the proportional relation, it can be removed. In the posterior referring to f_i, the argument $Q_{l,i}$ has to be explicitly expressed depending on this univariate variable. To this aim, let us introduce

$$\boldsymbol{e}_i = \boldsymbol{y} - (\boldsymbol{H}\boldsymbol{f} - \boldsymbol{h}_i f_i), \tag{57}$$

with $\boldsymbol{h}_i$ the i-th column of matrix $\boldsymbol{H}$, that is to say the global impulse response of the system excited at its i-th spatial node. It can be calculated from the SVD of the operator matrix of the structure $\boldsymbol{D}$ by

$$\boldsymbol{h}_i = \boldsymbol{V}\boldsymbol{S}\boldsymbol{u}_{(:,i)}^H, \tag{58}$$

where $\boldsymbol{u}_{(:,i)}$ corresponds to the i-th row of the left singular vectors matrix $\boldsymbol{U}$ (not to be confused with $\boldsymbol{u}_i$, the i-th column of $\boldsymbol{U}$) and $\boldsymbol{S}$ to the diagonal matrix of singular values. The vector $\boldsymbol{e}_i$ then corresponds to residuals without the contribution of th specific source f_i. Then, developments give

$$\boldsymbol{y} - \boldsymbol{H}\boldsymbol{f} = \boldsymbol{e}_i - \boldsymbol{h}_i f_i, \tag{59}$$

and thus

$$Q_{l,i} = \tau_n \|\boldsymbol{e}_i - \boldsymbol{h}_i f_i\|_2^2 + \tau_l \|f_i\|_2^2,\ l = 0, 1. \tag{60}$$

Gaussian process.

Contrary to the latter cases, the normalizing constant coming from the product between the prior and the likelihood plays in an important role in the sparse model. Indeed, the selection between the Dirac distribution and the Gaussian draw depends on it. To recognize the expression of a Gaussian distribution, $Q_{l,i}$ is expanded in the following form

$$\begin{aligned} Q_{l,i} &= (f_i - \mu_i)^H \tau_p (f_i - \mu_i) + R \quad (61) \\ &= f_i^H \tau_p f_i - f_i^H \tau_p \mu_i - \mu_i^H \tau_p f_i + \mu_i^H \tau_p \mu_i + R, \quad (62) \end{aligned}$$

where μ_i corresponds to the mean of the equivalent Gaussian distribution, τ_p is the new precision and R is the normalizing constant moved into the exponential argument. This latter element is particularly important to identify the normalizing constant. On another note, by expanding expression (60), the exponential argument can also be expressed as

$$\begin{aligned} Q_{l,i} &= \tau_n \boldsymbol{e}_i^H \boldsymbol{e}_i + \tau_n \left(\boldsymbol{h}_i f_i\right)^H \left(\boldsymbol{h}_i f_i\right) - \tau_n \boldsymbol{e}_i^H \left(\boldsymbol{h}_i f_i\right) - \tau_n \left(\boldsymbol{h}_i f_i\right)^H \boldsymbol{e}_i + \tau_l (f_i^H f_i) \quad (63) \\ &= f_i^H f_i \left(\tau_l + \tau_n \|\boldsymbol{h}_i\|_2^2\right) - f_i^H (\boldsymbol{h}_i^H \boldsymbol{e}_i) \tau_n - f_i (\boldsymbol{h}_i^H \boldsymbol{e}_i)^H \tau_n + \|\boldsymbol{e}_i\|_2^2 \tau_n. \quad (64) \end{aligned}$$

From identification with Eq. (62), the equivalent precision τ_p corresponds to

$$\tau_p = \tau_l + \tau_n \|\boldsymbol{h}_i\|_2^2, \quad (65)$$

the equivalent mean μ_i corresponds to

$$\mu_i = \frac{\left(\boldsymbol{h}_i^H \boldsymbol{e}_i\right) \tau_n}{\tau_p} = \frac{\boldsymbol{h}_i^H \boldsymbol{e}_i}{\|\boldsymbol{h}_i\|_2^2 + \frac{\tau_l}{\tau_n}}, \quad (66)$$

and finally, the residuals R equals

$$R = -\mu_i^H \tau_p \mu_i + \|\boldsymbol{e}_i\|_2^2 \tau_n = \left(\frac{-\|\boldsymbol{h}_i^H \boldsymbol{e}_i\|_2^2}{\|\boldsymbol{h}_i\|_2^2 + \frac{\tau_l}{\tau_n}} + \|\boldsymbol{e}_i\|_2^2 \right) \tau_n. \quad (67)$$

Bernoulli process.

At this stage, the mean μ_i and the precision τ_p of the Gaussian posterior have been identified. The relative weight between the Dirac distribution and the Gaussian distribution still has to be evaluate to determine whether the specific node is activated or not. Restarting from Eq. (51) and from identified parameters of

the equivalent Gaussian distribution, the posterior on f_i is written as

$$\left[f_i \mid \lambda, \tau_f, \boldsymbol{y}, \boldsymbol{H}, \boldsymbol{f}^{(-i)}, \tau_n\right] \propto \sum_{l=0,1} \left\{ \frac{\lambda_l \exp\left(+\tau_p \|\mu_i\|_2^2 - \tau_n \|\boldsymbol{e}_i\|_2^2\right)}{\pi\ \tau_l^{-1}} \exp\left(-\tau_p \|f_i - \mu_i\|_2^2\right) \right\}, \tag{68}$$

$$\propto \sum_{l=0,1} \left\{ \frac{\tilde{\lambda}_{l,i}}{\pi\ \tau_p^{-1}} \exp\left(-\tau_p \|f_i - \mu_i\|_2^2\right) \right\}. \tag{69}$$

In Eq. (67), the norm on $\boldsymbol{e}_i$ is independent from l. Indeed, it acts as a multiplicative constant applied both on the Dirac distribution and the Gaussian distribution. Then, it doesn't modify the relative weight and can simply be removed thereafter. The passage to the last expression aims at rewriting the conditional probability on f_i as the general expression of the circular-complex Gaussian distribution with parameters μ_i and τ_p. This step is realized by introducing a new variable $\tilde{\lambda}_{l,i}$ for the "specific sparsity" at node f_i, described by the following relation,

$$\frac{\tilde{\lambda}_{l,i}}{\pi\ \tau_p^{-1}} = \frac{\lambda_l \exp\left(+\tau_p \|\mu_i\|_2^2\right)}{\pi\ \tau_l^{-1}},\ l = 0, 1. \tag{70}$$

By isolating $\tilde{\lambda}_{l,i}$,

$$\tilde{\lambda}_{l,i} \quad = \quad \lambda_l \frac{\tau_l}{\tau_p} \exp\left(+\tau_p \|\mu_i\|_2^2\right),\ l = 0, 1. \tag{71}$$

For the activation of each node, the global sparsity parameter λ (and λ_l as a result, being defined from the first one) acts as a prior for the whole mesh nodes, while the specific sparsity parameter $\tilde{\lambda}_{l,i}$ acts as a posterior for each node. All the expressions are now known (Eqs. (52) to (55), (65), (66) and (67)) and the sparsity parameter of specific node f_i can be calculated. For the Gaussian distribution ($l = 1$), the specific sparsity parameter $\tilde{\lambda}_{1,i}$ is given by

$$\tilde{\lambda}_{1,i} = \lambda \frac{\tau_f}{\tau_n \|\boldsymbol{h}_i\|_2^2 + \tau_f} \exp\left(+ \frac{\tau_n \|\boldsymbol{h}_i^H \boldsymbol{e}_i\|_2^2}{\|\boldsymbol{h}_i\|_2^2 + \frac{\tau_f}{\tau_n}} \right). \tag{72}$$

For the Dirac distribution ($l = 0$), the parameter $\tilde{\lambda}_{0,i}$ is equal to

$$\tilde{\lambda}_{0,i} = (1 - \lambda) \frac{\tau_0}{\tau_n \|\boldsymbol{h}_i\|_2^2 + \tau_0} \exp\left(+ \frac{\tau_n \|\boldsymbol{h}_i^H \boldsymbol{e}_i\|_2^2}{\|\boldsymbol{h}_i\|_2^2 + \frac{\tau_0}{\tau_n}} \right). \tag{73}$$

However, the precision τ_0 tends to infinity, hence

$$\lim_{\tau_0 \to +\infty} \tilde{\lambda}_{0,i} = (1 - \lambda). \tag{74}$$

Since the sum of these two values is not always equal to 1, a last step of normalizing is thus needed,

$$\lambda_{l,i} \quad = \quad \frac{\tilde{\lambda}_{l,i}}{\tilde{\lambda}_{l,i} + \tilde{\lambda}_{1-l,i}} = \left(1 + \frac{\tilde{\lambda}_{1-l,i}}{\tilde{\lambda}_{l,i}}\right)^{-1}. \tag{75}$$

For the Gaussian distribution, the normalized sparsity parameter equals

$$\lambda_{1,i} = \left(1 + \frac{1-\lambda}{\lambda}\frac{\tau_n\|\boldsymbol{h}_i\|_2^2 + \tau_f}{\tau_f}\exp\left(-\frac{\tau_n\|\boldsymbol{h}_i^H\boldsymbol{e}_i\|_2^2}{\|\boldsymbol{h}_i\|_2^2 + \frac{\tau_f}{\tau_n}}\right)\right)^{-1}, \tag{76}$$

and for the Dirac distribution,

$$\lambda_{0,i} = \left(1 + \frac{\lambda}{1-\lambda}\frac{\tau_f}{\tau_n\|\boldsymbol{h}_i\|_2^2 + \tau_f}\exp\left(+\frac{\tau_n\|\boldsymbol{h}_i^H\boldsymbol{e}_i\|_2^2}{\|\boldsymbol{h}_i\|_2^2 + \frac{\tau_f}{\tau_n}}\right)\right)^{-1}. \tag{77}$$

In practice, only one of these two normalized sparsity parameters has to be calculated; since the exponential with a positive argument can be badly handled numerically, it is thus better to calculate the parameter $\lambda_{1,i}$.
710 Therefore,

$$\left[f_i \mid \lambda, \tau_f, \boldsymbol{y}, \boldsymbol{H}, \boldsymbol{f}^{(-i)}, \tau_n\right] = (1-\lambda_{1,i})\delta(f_i) + \lambda_{1,i}\mathcal{N}_c(f_i \mid \mu_i, \tau_p^{-1}) \tag{78}$$

with

$$\mu_i = \frac{\boldsymbol{h}_i^H\boldsymbol{e}_i}{\|\boldsymbol{h}_i\|_2^2 + \frac{\tau_f}{\tau_n}} \tag{79}$$

and

$$\tau_p = \tau_f + \tau_n\|\boldsymbol{h}_i\|_2^2. \tag{80}$$

Sampling in the above distribution is done as follows. A random draw u is made in the the uniform distribution $\mathcal{U}_{[0;1]}$. If $u \geq \lambda_{1,i}$, then the amplitude of f_i is set to 0. Otherwise, it is sampled from the
715 Gaussian distribution with mean μ_i and precision τ_p. This is resumed in algorithm 2.

Algorithm 2 Sampling of force distribution using "standard" Gibbs sampler.

1: **for** $i = 1$ to N **do**
2: Sample $u \sim \mathcal{U}([0;1])$
3: **if** $u < \lambda_{1,i}$ **then** {from Eq. (76)}
4: $f_i \sim \mathcal{N}_c\left(f_i \mid \mu_i, \tau_p^{-1}\right)$ {from Eqs. (79) and (80)}
5: **else**
6: $f_i = 0$
7: **end if**
8: **end for**

B. Posterior on binaries with marginalization over amplitudes

The posterior on binary q_i after marginalization over amplitudes $\boldsymbol{r}$ is obtained by

$$
\begin{aligned}
[q_i \mid \boldsymbol{q}^{(-i)}, \boldsymbol{H}, \boldsymbol{y}, \tau_f, \tau_n, \lambda] \propto & \quad [\boldsymbol{y} \mid \boldsymbol{q}, \boldsymbol{H}, \tau_f, \tau_n][q_i \mid \lambda] && (81)\\
\propto & \quad \mathcal{N}_c\left(\boldsymbol{y} \mid \boldsymbol{0}, \boldsymbol{B}\right) \mathrm{Bern}\left(q_i \mid \lambda\right) && (82)\\
\propto & \quad \tfrac{\exp\left(-\boldsymbol{y}^H \boldsymbol{B}^{-1} \boldsymbol{y}\right)}{\pi^N |\boldsymbol{B}|}\ \lambda^{q_i}(1-\lambda)^{1-q_i} && (83)\\
\propto & \quad \underbrace{\frac{1-\lambda}{\pi^N}}_{C_1} |\boldsymbol{B}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{B}^{-1} \boldsymbol{y}\right) \left(\tfrac{\lambda}{1-\lambda}\right)^{q_i} && (84)\\
\propto & \quad C_1\ \exp\left(\underbrace{-\boldsymbol{y}^H \boldsymbol{B}^{-1} \boldsymbol{y} - \ln(|\boldsymbol{B}|) - q_i \ln\left(\frac{1}{\lambda} - 1\right)}_{-g(q_i)} \right). && (85)
\end{aligned}
$$

In Eq. (82), the Bernoulli distribution results directly from the prior on activation variables while the Gaussian likelihood is obtained from a development similar to the one of Eq. (22) in Ref. [4]. The constant C_1 remains the same whatever the state of the binaries sequence $\boldsymbol{q}$. The covariance matrix associated to this marginalized Gaussian likelihood on sources is defined by

$$
\boldsymbol{B} = \tau_f^{-1} \boldsymbol{H} \lceil \boldsymbol{q} \rfloor \boldsymbol{H}^H + \tau_n^{-1} \mathbf{I}_N, \tag{86}
$$

with $\lceil \boldsymbol{q} \rfloor$ the diagonal matrix composed with elements of vector $\boldsymbol{q}$. In Eq. (85), a change on q_i modifies the probability by two ways, through the term $q_i \ln\left(\frac{1}{\lambda} - 1\right)$, but also through the matrix $\boldsymbol{B}$. The covariance matrix $\boldsymbol{B}_{init}$ associated to an initial sequence of binaries $\boldsymbol{q}_{init}$ can be resumed as

$$
\tilde{\boldsymbol{B}}_{init} = \boldsymbol{B}_{init} \tau_n = \eta^2 \boldsymbol{H} \lceil \boldsymbol{q}_{init} \rfloor \boldsymbol{H}^H + \mathbf{I}_N, \tag{87}
$$

with $\eta^2 = \frac{\tau_n}{\tau_f}$ the signal-to-noise ratio, allowing the isolation of the identity matrix and the simplification of developments in the following. The modification of the i-th binary q_i leads to a perturbation on the reduced matrix $\tilde{\boldsymbol{B}}_{init}$,

$$
\tilde{\boldsymbol{B}}_{mod} = \tilde{\boldsymbol{B}}_{init} + \delta_i \eta^2 \boldsymbol{h}_i \boldsymbol{h}_i^H, \tag{88}
$$

where $\delta_i = \pm 1$ whether the binary q_i is activating $(+1)$ or is deactivating (-1). Thus, when a node is activated, the amplitude of the covariance matrix is increased according to the signal-to-noise ratio η^2, but also according to the system response to this node, $\boldsymbol{h}_i$. Therefore, the binary either remains to its initial

state $q_{i_{init}}$, or it is modified to $q_{i_{mod}}$. In the case of the activation of a binary, $q_{i_{init}} = 0$ and $q_{i_{mod}} = +1$, then $q_{i_{init}} - q_{i_{mod}} = -1$ and $\delta_i = +1$. From a symmetrical reasoning, the deactivation of a binary leads to $q_{i_{init}} = +1$ and $q_{i_{mod}} = 0$, then $q_{i_{init}} - q_i mod = +1$ and $\delta_i = -1$. As a result,

$$q_{i_{init}} - q_{i_{mod}} = -\delta_i. \tag{89}$$

The conditional probabilities of $q_{i_{init}}$ and $q_{i_{mod}}$ are

$$[q_{i_{init}} \mid \boldsymbol{q}^{(-i)}, \boldsymbol{H}, \boldsymbol{y}, \tau_f, \tau_n, \lambda] \propto C_1 \ \exp\left(\underbrace{-\boldsymbol{y}^H \boldsymbol{B}_{init}^{-1}\boldsymbol{y} - \ln(|\boldsymbol{B}_{init}|) - q_{i_{init}} \ln\left(\frac{1}{\lambda} - 1\right)}_{-g(q_{i_{init}})}\right), \tag{90}$$

and

$$[q_{i_{mod}} \mid \boldsymbol{q}^{(-i)}, \boldsymbol{H}, \boldsymbol{y}, \tau_f, \tau_n, \lambda] \propto C_1 \ \exp\left(\underbrace{-\boldsymbol{y}^H \boldsymbol{B}_{i_{mod}}^{-1}\boldsymbol{y} - \ln(|\boldsymbol{B}_{i_{mod}}|) - q_{i_{mod}} \ln\left(\frac{1}{\lambda} - 1\right)}_{-g(q_{i_{mod}})}\right). \tag{91}$$

As for the standard Gibbs sampler, sampling from the Bernoulli distribution requires a normalization of probabilities depending on the different values potentially taken by the variable q_i. Considering only the modified state, the probability on q_i is expressed as

$$\begin{aligned}
[q_{i_{mod}} \mid \boldsymbol{q}^{(-i)}, \boldsymbol{H}, \boldsymbol{y}, \tau_f, \tau_n, \lambda] = & \quad \frac{\exp\big(-g(q_{i_{mod}})\big)}{\exp\big(-g(q_{i_{mod}})\big)+\exp\big(-g(q_{i_{init}})\big)}, & (92)\\
= & \quad \left(1 + \exp\left(-(g(q_{i_{init}}) - g(q_{i_{mod}}))\right)\right)^{-1} & (93)
\end{aligned}$$

so that the sum of the two probabilities is equal to 1. The difference in the exponential argument is then

$$\begin{aligned}
g(q_{i_{init}}) - g(q_{i_{mod}}) = & +\boldsymbol{y}^H \boldsymbol{B}_{init}^{-1}\boldsymbol{y} + \ln(|\boldsymbol{B}_{init}|) + q_{i_{init}} \ln\left(\frac{1}{\lambda} - 1\right)\\
& -\boldsymbol{y}^H \boldsymbol{B}_{mod}^{-1}\boldsymbol{y} - \ln(|\boldsymbol{B}_{mod}|) - q_{i_{mod}} \ln\left(\frac{1}{\lambda} - 1\right).
\end{aligned} \tag{94}$$

To further simplify this expression, the matrix $\tilde{\boldsymbol{B}}_{mod}^{-1}$ can be expressed in terms of $\tilde{\boldsymbol{B}}_{init}^{-1}$ by means of expression (88) and of the Woodbury matrix identity [60],

$$\begin{aligned}
\tilde{\boldsymbol{B}}_{mod}^{-1} &= \tilde{\boldsymbol{B}}_{init}^{-1} - \tilde{\boldsymbol{B}}_{init}^{-1}\eta^2 \boldsymbol{h}_i \underbrace{(\delta_i + \boldsymbol{h}_i^H \tilde{\boldsymbol{B}}_{init}^{-1}\eta^2 \boldsymbol{h}_i)}_{\tilde{\rho}_i}{}^{-1} \boldsymbol{h}_i^H \tilde{\boldsymbol{B}}_{init}^{-1} & (95)\\
&= \tilde{\boldsymbol{B}}_{init}^{-1} - \eta^2 \tilde{\rho}_i^{-1} \tilde{\boldsymbol{B}}_{init}^{-1} \boldsymbol{h}_i \boldsymbol{h}_i^H \tilde{\boldsymbol{B}}_{init}^{-1}. & (96)
\end{aligned}$$

In addition, in comparison with Eq. (88), it can be noted that

$$|\tilde{\boldsymbol{B}}_{mod}| = |\tilde{\boldsymbol{B}}_{init}|\delta_i\tilde{\rho}_i, \tag{97}$$

and then,

$$\ln\left(|\boldsymbol{B}_{mod}|\right) = \ln\left(|\boldsymbol{B}_{init}|\right) + \ln\left(\delta_i\tilde{\rho}_i\right). \tag{98}$$

Injecting expressions (96) and (98) into Eq. (94) and taking into account the division by τ_n done in Eq. (87) on covariance matrices, the difference in the exponential argument is equal to

$$f(q_{i_{init}}) - g(q_{i_{mod}}) = +\tau_n \boldsymbol{y}^H \left(\eta^2\tilde{\rho}_i^{-1}\tilde{\boldsymbol{B}}_{init}^{-1}\boldsymbol{h}_i\boldsymbol{h}_i^H\tilde{\boldsymbol{B}}_{init}^{-1}\right)\boldsymbol{y} - \ln\left(\delta_i\tilde{\rho}_i\right) + \left(q_{i_{init}} - q_{i_{mod}}\right)\ln\left(\frac{1}{\lambda} - 1\right). \tag{99}$$

Finally, the probability associated to the change of state of a binary is

$$\begin{aligned}[q_{i_{mod}} \mid \boldsymbol{q}^{(-i)}, \boldsymbol{H}, \boldsymbol{y}, \tau_f, \tau_n, \lambda] = \Bigg(1 + \exp\Bigg(-\Bigg(&+\tau_n \boldsymbol{y}^H \left(\eta^2\tilde{\rho}_i^{-1}\tilde{\boldsymbol{B}}_{init}^{-1}\boldsymbol{h}_i\boldsymbol{h}_i^H\tilde{\boldsymbol{B}}_{init}^{-1}\right)\boldsymbol{y} \\ &- \ln\left(\delta_i\tilde{\rho}_i\right) - \delta_i\ln\left(\frac{1}{\lambda} - 1\right)\Bigg)\Bigg)\Bigg)^{-1}.\end{aligned} \tag{100}$$

This is easily implemented as follows: for each iteration and for each binary, a sample $u \sim \mathcal{U}([0;1])$ is compared to the value of Eq. (100). If it is lower than $[q_{i_{mod}} \mid \boldsymbol{q}^{(-i)}, \boldsymbol{H}, \boldsymbol{y}, \tau_f, \tau_n, \lambda]$, thus the binary state changes. If the change corresponds to an activation, then the associated amplitude is drawn from the complex Gaussian with mean and precision parameters defined in Eqs. (65) and (66). A strategy for the fast update of the covariance matrix $\tilde{\boldsymbol{B}}_{mod}$ from $\tilde{\boldsymbol{B}}_{init}$ is proposed by [61].

Finally, the force distribution sampling with partial marginalization over amplitudes is detailed in algorithm 3.

C. Source sampling step of the Partially Collapsed Gibbs sampler algorithm

Algorithm 3 Sampling of force distribution using partially collapsed Gibbs sampler.

for $i = 1$ to N **do**
  $\delta_i \leftarrow (-1)^{q_{i_{init}}}$ {add or remove binary depending on its initial state}
  $g(q_{i_{init}}) - g(q_{i_{mod}}) \leftarrow$ Eq. (100)
  Sample $u \sim \mathcal{U}([0;1])$
  if $u < (1 + \exp -(g(q_{i_{init}}) - g(q_{i_{mod}})))^{-1}$ **then**
    $q_{i_{mod}} \leftarrow q_{i_{init}} + \delta_i$ {binary state change accepted}
  end if
  if $q_i = +1$ **then** {sample of corresponding pointwise source amplitude}
    $r_i \sim \mathcal{N}_c\left(r_i \mid \mu_{r_i}, \tau_p^{-1}\right)$ {from Eqs. (65) and (66)}
  else if $q_i = 0$ **then**
    $r_i = 0$
  end if
  $f_i = q_i r_i$
end for