Empirical and fully Bayesian approaches for the identification of vibration sources from transverse displacement measurements

Charly Faure, Frédéric Ablitzer, Jérôme Antoni, Charles Pezerat

▶ **To cite this version:**

Charly Faure, Frédéric Ablitzer, Jérôme Antoni, Charles Pezerat. Empirical and fully Bayesian approaches for the identification of vibration sources from transverse displacement measurements. Mechanical Systems and Signal Processing, 2017, 94, pp.180 - 201. 10.1016/j.ymssp.2017.02.023 . hal-01714495

HAL Id: hal-01714495
https://hal.science/hal-01714495

Submitted on 3 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mechanical Systems and Signal Processing 94 (2017) 180–201

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Empirical and fully Bayesian approaches for the identification of vibration sources from transverse displacement measurements

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

[a] Laboratoire d'Acoustique de l'Université du Maine, UMR CNRS 6613, Avenue Olivier Messiaen, 72085 Le Mans Cedex 9, France
[b] Laboratoire Vibrations Acoustique, Univ Lyon, INSA-Lyon, LVA EA677, F-69621 Villeurbanne, France

ARTICLE INFO

Article history:
Received 12 July 2016
Received in revised form 31 January 2017
Accepted 16 February 2017

Keywords:
Vibration source characterization
Inverse problem
Bayesian regularization
Marginalized maximum *a posteriori*
Gibbs sampler
Credible interval

ABSTRACT

This paper introduces the Bayesian regularization applied to the Force Analysis Technique (FAT), a method for identifying vibration sources from displacement measurements. The FAT is based on the equation of motion of a structure instead of a transfer matrix as it is the case for most of inverse problems. This particularity allows the estimation of vibration sources without the need of boundary conditions. Nevertheless, this method is highly sensitive to noise perturbations and needs a careful regularization. Two Bayesian approaches are thus presented. Firstly, the empirical Bayesian regularization which shows better robustness than L-curve and GCV regularizations while keeping a low numerical cost. Secondly, a fully Bayesian procedure using a Markov Chain Monte Carlo (MCMC) algorithm which provides credible intervals on variables of interest besides the automatically regularized vibration source field. In particular, measurement quality can be evaluated by the noise variance estimation and the uncertainties over the source level are quantified for a wide frequency range, with only a unique measurement scan.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

In vibration engineering, many efforts have been made to improve prediction methods apt to describe vibration propagation so as to provide better estimates of model outputs, such as displacement fields. However, the knowledge of the input sources still remains largely limited. In turn, this leads to unavoidable bias on the predicted outputs. Hence, the identification of vibration sources is an important topic of vibroacoustics.

Originally, vibration sources identification has been investigated in Refs. [1,2] using the concept of structural intensity. This technique generates a field of intensity vectors, allowing the location of vibration sources without quantifying them. Then, different methods have been developed based on modal models or frequency response functions, including Transfer Path Analysis (TPA) described in Ref. [3] or Statistical Energy Analysis described (SEA) described in Ref. [4]. A review of some of these methods can be found in Ref. [5]. As these methods often need global information about the structure, such as boundary conditions, position of sources or number of sources, they can be interpreted as global methods in opposition to local methods which only need local information about the structure. Simultaneously, Ref. [6] and Refs. [7,8] developed a local approach based on the numerical discretization of the equation of motion of the structure to identify the excitation force. The difference between these two methods is the way that they deal with measurement noise. This step represents the regularization inherent to many inverse problems, as will be introduced below. Ref. [6] passes through the wavenumber

* Corresponding author.
E-mail address: charly.faure@univ-lemans.fr (C. Faure).

http://dx.doi.org/10.1016/j.ymssp.2017.02.023
0888-3270/© 2017 Elsevier Ltd. All rights reserved.

domain while Refs. [7,8] uses a convolution with a truncated convolution kernel. From that point of view, the second method, called Force Analysis Technique (FAT, or RIFF in French for Résolution Inverse Filtrée Fenêtrée), is more local and will serve as a starting point for this present work. The FAT has been originally developed for beams and plates and recent works have adapted it to cylindrical shells [9] and to more complex structures by using the Finite Element Method [10].

As mentioned above, the measurement noise causes problems to the FAT and more generally to inverse problems because measurement perturbations are interpreted as the result of multiple unphysical sources. Mathematically, this instability results from the ill-conditioning of the transfer matrix linking inputs to outputs – i.e. the ratio between its maximal and minimal eigenvalues is very large. To ensure that the estimated sources have a physical meaning, one should apply a regularization step. The Truncated Singular Value Decomposition (TSVD) method described in Ref. [11] improves the stability of the inversion by taking into account only eigenvalues higher than a specific value to ensure that the ratio between maximal and minimal eigenvalues is reasonable. As another possibility, the Tikhonov regularization, proposed in Ref. [12] is most widely adopted. The concept is to increase low eigenvalues to get an appropriate ratio. The lowpass filter method presented in Ref. [8] gives similar results with a more physical approach. Nevertheless, this regularization is supervised because the cutoff frequency must be set by the operator. Also, the lowpass filter is not deduced from an optimization but selected for its simplicity. The optimization of the regularization (i.e. the number of eigenvalues to take into account or the threshold of the low eigenvalues) can be realized by Generalized Cross Validation (GCV) or by the L-curve proposed respectively in Refs. [13,14]. The Bayesian framework has also been investigated for the regularization and optimization of inverse problems (see Refs. [15,16]). Applications have been developed mainly in econometrics (see Ref. [17]), biology (see Ref. [18]), social sciences (see Ref. [19]) and image processing (see Ref. [20]). In vibration and acoustics, Refs. [21,22] have used the empirical Bayesian regularization for acoustic holography, Ref. [23] has used it for sparse vibration source identification and Ref. [24] has used a modal model whose parameters are partially unknown to infer the vibration source applied on a beam. In this work, the regularization step is realized within the Bayesian framework because it makes possible to take into account various *a priori* information and to formalize them mathematically, making this approach more practical than Tikhonov method. Moreover, Bayesian approaches usually simplify the inverse problem by using the associated direct problem which is often well described by analytical or numerical models. However, the FAT, being based on the equation of motion, is already written as a direct problem. The application of Bayesian methods is thus more complicated since it will lead to the evaluation of the inverse problem instead of the direct one. Additional steps are then proposed to a better compatibility between the FAT and Bayesian methods, such as specific stencils for numerical approximation of derivatives or a matrix inversion lemma. To go further, a fully Bayesian procedure using Markov Chain Monte Carlo (MCMC) algorithms [25] allows the inference of the whole *a posteriori* density on unknown parameters from numerical sampling, yielding a better exploration of probability density functions (pdf), credible intervals on random variables and an automatic and optimized regularization. The Gibbs sampler [26,27], a special case of MCMC, is particularly suited to problems where the pdf have an analytical expression, such as the Gaussian or the gamma distributions.

In this paper, the vibration source identification method based on measurements of displacement field is first described. The empirical Bayesian regularization leading to a specific optimization is then introduced and compared with other unsupervised optimization procedures, namely the L-curve and GCV. The lowpass filter is not included in this comparison because it is considered as supervised [8]. The robustness improvement with the empirical Bayesian regularization is then highlighted. Then the MCMC part, where the hierarchical model of the problem is discussed, introduces an upper level of *a priori* and the resulting conditional pdf for the Gibbs sampler. The main interest of using the MCMC approach is to obtain credible intervals for each variable besides an automatic and unsupervised regularization. Finally, numerical and experimental validations on a beam excited by a point source are presented to illustrate these methods before a conclusion section.

2. Identification of vibration sources from transverse displacement

This section presents the FAT developed in Ref. [8]. Starting from the equation of motion, spatial and temporal derivatives are simplified to calculate directly the source distribution applied on the structure. After that, a matrix formulation of the approach with specific boundaries is introduced and compared to existing models. All the data are taken into account as a whole with the matrix formulation. This is the main difference with the work presented in Ref. [8], where a stencil is used to scan the measurements. A qualitative explanation of the limitations of the FAT with regards to noise perturbations is then proposed, 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 on a beam within Euler-Bernoulli beam theory. With temporal convention $e^{+j\omega t}$ defined for the frequency $\nu = \frac{\omega}{2\pi}$, the harmonic transverse displacement $w(x,\omega)$ of the beam satisfies Eq. (1),

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

where E is Young's modulus, $j=\sqrt{-1}$ the imaginary number, η the loss factor, I the moment of inertia, ρ the density, S the cross-section area, $w(x,\omega)$ and $f(x,\omega)$ the harmonic transverse displacement and the vibration source distribution in N/m at location x and at angular frequency ω, respectively. The principle of the method is to measure the transverse displacement field, to estimate then its spatial derivatives and, knowing the rest of the terms of the left hand side of equation Eq. (1), to calculate directly the vibration source applied on the structure.

2.2. Estimation of derivates

The spatial derivative is estimated by numerical differentiation, using the centered finite difference method at first order, from

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

where Δ_x is the spatial sampling rate. Substituting Eq. (2) into Eq. (1) allows the identification of the vibration source applied at location x_i without having to evaluate boundary conditions and displacement all over the domain, as proposed in Eq. (3)

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

where $w(x_i,\omega)$ and $f(x_i,\omega)$ are now denoted as w_i and f_i to simplify the notation. The excitation can be classified into two categories. It can result either from pointwise sources or distributed sources. In the case of pointwise sources, the source distribution can simply be spatially integrated around each source to obtain a force amplitude in Newtons. If the source is distributed with all its point sharing the same phase (for example, a transmission through a welding line), the spatial integration can again leads to the force amplitude. It is indeed an advantage of the FAT over transfer function based force identification methods whose identification is given only at a specific point. However, if the source is distributed with different phases (for example, an acoustic excitation) the spatial integration is not pertinent anymore because it would result in a destructive reconstruction.

2.3. Matrix formulation

Considering a portion of the structure with a constant spatial discretization, Eq. (3) can be expressed in a matrix form, as it will be used in the following sections,

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

where $\boldsymbol{w}$ is the vector of displacements, $\boldsymbol{f}$ is the vector of vibration sources and $\boldsymbol{D}$ is the operator matrix resulting from the discretization of the structural local operator. Due to the fourth order spatial derivative, the size of $\boldsymbol{D}$ is $(N-4,N)$, N being the number of measurement points. The inverse of this matrix is required to solve the inverse problem, as it will be shown in Section 3. It is then possible to complete the rectangular matrix $\boldsymbol{D}$ by adding information at each boundary. For example, two other order derivatives at each boundary can be chosen to evaluate local boundary conditions. It is also possible to use fourth order forward and backward finite difference schemes, whose stencils are $[3;-14;26;-24;11;-2]$ and $[-2;11;-24;26;-14;3]$ respectively, in order to evaluate the source at the studied area boundaries. This second option, which will be used in this work, is recommended for two reasons. First, there is no automatic rule for choosing between slope, bending moment or shear force as boundary conditions while only two of them can be evaluated. Secondly, the resulting derivatives at boundaries can be much more important than the evaluated sources inside the domain, yielding an over estimated regularization as it will be introduced in Section 3. The finite element method can also estimate sources at boundaries of a sub-domain of the structure [10,29], but the result is a combination of internal stresses and external sources located at the boundaries. 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 of the finite difference method is higher than the finite element method one, it can be drastically reduced by applying an analytical correction [30]. Hence, it seems to be an appropriate model, at least for academic structures whose equations of motion are known.

2.4. Additive noise

As each measurement includes noise, the equation of observation considering an additive white noise is given by

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

where $\boldsymbol{y}$ is the vector of observed displacement and $\boldsymbol{n}$ is the vector of noise. A multiplicative noise part may also be present depending on experimental setup and studied physical quantities, but is often neglected to simplify the approach and mathematical developments. Applying the method of identification of vibration sources to the accessible quantity $\boldsymbol{y}$ yields

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

Although noise $\boldsymbol{n}$ is much lower than displacement $\boldsymbol{w}$, $\boldsymbol{D}$ is a differential operator which, when applied to $\boldsymbol{n}$, may considerably inflate the level of noise, making the source identification difficult. Thus, a regularization step is needed to overcome noise issues.

3. Regularization within the Bayesian framework

The empirical Bayesian framework for solving the inverse problem is first presented in this section together with its inherent regularization. It is then compared with other approaches such as the Tikhonov regularization [12] tuned with the L-curve [14] and GCV [13]. An empirical optimization of the regularization is then compared with the often used L-curve and GCV.

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 uncertainties. The Bayesian framework used in this work allows the inclusion of the *a priori* information in the identification process. Starting from the conditional probability formula

$$[A, B] = [A|B][B] = [B|A][A] \tag{7}$$

where $[A]$ stands for the pdf of the random variable A, $[A, B]$ for the joint pdf for A *and* B and $[B|A]$ for the conditional pdf of B knowing A, Bayesian probabilities are based on the following theorem

$$[A|B] = \frac{[B|A][A]}{[B]} \propto [B|A][A], \tag{8}$$

to infer the probability of an event from both experiences and *a priori* knowledge. The *a posteriori* density, or posterior, $[A|B]$ 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). $[B|A]$ is the pdf involved in the direct problem which expresses the information extracted from the experience and is called the "likelihood". $[A]$ is the *a priori* pdf, or prior, from *a priori* knowledge and $[B]$ 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. (8) rather than the absolute equality.

3.2. Priors

Considering additive white noise, the *a priori* distribution on $\boldsymbol{n}$ can be expressed as

$$[\boldsymbol{n}] \propto \mathcal{N}_c\left(\boldsymbol{n}; \boldsymbol{0}, \tau_n^{-1}\mathbf{I}\right). \tag{9}$$

The scalar precision (the inverse of the variance) of noise τ_n is preferred to the most commonly used variance for simplification in Section 4. $\mathbf{I}$ stands for the $N \times N$ identity matrix and the N-dimensional multivariate circular complex Gaussian pdf on $\boldsymbol{x}$ (see Ref. [31]) with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is defined as

$$\mathcal{N}_c(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\pi^N |\boldsymbol{\Sigma}|} \exp\left(-(\boldsymbol{x}-\boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})\right), \tag{10}$$

where the exponent H stands for the Hermitian transposition. Although the sources are unknown, their energy is limited, so the *a priori* distribution on $\boldsymbol{f}$ can also be expressed as a zero-mean circular complex Gaussian

$$[\boldsymbol{f}] \propto \mathcal{N}_c\left(\boldsymbol{f}; \boldsymbol{0}, \tau_f^{-1}\mathbf{I}\right), \tag{11}$$

with τ_f the precision of sources. The fact that the two covariance matrices in Eqs. (9) and (11) are proportional to the identity matrix reflects that the pdf is the same all over the domain and that there is no spatial dependence. Random variables with this kind of covariance are then independent and identically distributed. It is equivalent to the 'rain-on-the-roof' hypothesis in Statistical Energy Analysis.

3.3. Likelihood

From Eqs. (6) and (9), the pdf on 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

$$[\boldsymbol{y}|\boldsymbol{D}, \boldsymbol{f}, \tau_n] \propto \mathcal{N}_c\left(\boldsymbol{y}; \boldsymbol{D}^{-1}\boldsymbol{f}, \tau_n^{-1}\mathbf{I}\right). \tag{12}$$

It should be noted that the problem cannot be expressed directly by the dynamic stiffness matrix $\boldsymbol{D}$ but instead by its inverse, that is the transfer functions matrix $\boldsymbol{H}$. As shown later, a matrix inversion lemma will allow avoidingthe explicit inversion of the matrix at the end of the development.

3.4. Posterior

Applying the Bayes theorem of Eq. (8) to the unknown sources $\boldsymbol{f}$ from Eqs. (11) and (12) yields

$$[\boldsymbol{f}|\boldsymbol{D},\boldsymbol{y},\tau_f,\tau_n] \propto \mathcal{N}_c\left(\boldsymbol{f};\boldsymbol{0},\tau_f^{-1}\boldsymbol{I}\right)\mathcal{N}_c\left(\boldsymbol{y};\boldsymbol{D}^{-1}\boldsymbol{f},\tau_n^{-1}\boldsymbol{I}\right) \tag{13}$$

$$\propto \mathcal{N}_c\left(\boldsymbol{f};\boldsymbol{\mu}_{\boldsymbol{f}},\boldsymbol{\Sigma}_{\boldsymbol{f}}\right) \tag{14}$$

with :

$$\boldsymbol{\Sigma}_{\boldsymbol{f}} = \left(\boldsymbol{D}^{-H}\boldsymbol{D}^{-1}\tau_n + \tau_f\boldsymbol{I}\right)^{-1} \tag{15}$$

$$\boldsymbol{\mu}_{\boldsymbol{f}} = \boldsymbol{\Sigma}_{\boldsymbol{f}}\boldsymbol{D}^{-H}\tau_n\boldsymbol{y}. \tag{16}$$

By definition, the highest value of a gaussian pdf is its mean. The MAP corresponding to the regularized solution is then

$$\boldsymbol{f}_{MAP} = \left(\boldsymbol{D}^{-H}\boldsymbol{D}^{-1} + \eta^2\boldsymbol{I}\right)^{-1}\boldsymbol{D}^{-H}\boldsymbol{y} \tag{17}$$

$$= \left(\boldsymbol{D}\boldsymbol{D}^{H}\eta^2 + \boldsymbol{I}\right)^{-1}\boldsymbol{D}\boldsymbol{y}. \tag{18}$$

with $\eta^2 = \frac{\tau_f}{\tau_n}$ the regularization parameter driving the noise filtering. The higher η^2, the more selective the filter. Derived from the application of matrix inversion lemma, the second expression of $\boldsymbol{f}_{MAP}$ no longer depends on $\boldsymbol{D}^{-1}$ (see Appendix A). Although it is mathematically correct, numerical errors may come with practical implementation of Eq. (18), especially when the structural operator matrix is ill-conditioned or has high dimensions. Rewriting these expressions with singular value decomposition (SVD) is an efficient approach to significantly reduce the numerical errors occuring during the inversion. It also make possible to efficiently estimate the regularization parameter η^2 and then the regularized solution, as explained later. However, computing the SVD can be computationnally expensive for large matrices. Let $\boldsymbol{D}$ be decomposed as

$$\boldsymbol{D} = \boldsymbol{U}\lceil s_i \rfloor \boldsymbol{V}^H \tag{19}$$

where $\lceil a_i \rfloor$ stands for a diagonal matrix with diagonal element a_i, $\boldsymbol{U}$ and $\boldsymbol{V}$ for the left and right singular vectors of $\boldsymbol{D}$, respectively, and s_i corresponds to its singular values. Injecting Eq. (19) into Eq. (18) and after few algebraic manipulations, the regularized solution can be expressed as

$$\boldsymbol{f}_{MAP} = \boldsymbol{U}\left\lceil \frac{s_i}{1+\eta^2 s_i^2} \right\rfloor \boldsymbol{V}^H\boldsymbol{y}. \tag{20}$$

3.5. Relationship with other regularization approaches

It should be noticed that, when a Gaussian prior is used in Eq. (11), the Bayesian regularization leads to an expression similar to that of the most commonly used Tikhonov approach. Starting from the gaussian pdf expression in Eq. (10) and minimizing the opposite of the logarithm of the right hand side of Eq. (13) leads to the Tikhonov cost function

$$\hat{\boldsymbol{f}}_{Tikh} = \underset{\boldsymbol{f}}{\operatorname{argmin}} \|\boldsymbol{D}^{-1}\boldsymbol{f} - \boldsymbol{y}\|_2^2 + \eta^2\|\boldsymbol{f}\|_2^2. \tag{21}$$

The regularized solution is then a tradeoff between source energy and residuals or noise energy. According to the previous section, the regularization parameter is the inverse of a signal-to-noise ratio. More specifically, it is not the ratio between the noise and the directly measured displacements (as usually used to assess the measurement quality) but it corresponds to the ratio between the noise and the sources to be infered. The Bayesian approach provides then a more physical point of view on regularization than the Tikhonov one, where the regularization parameter is only a mathematical parameter to adjust. The Tikhonov approach can then be viewed as a particular case of the Bayesian one where Gaussian pdf describe both the likelihood and the prior. It is often easier within the Bayesian framework to design new regularization processing specifically adapted to particularities of inverse problems. For example, the sparsity property can be compelled by an high kurtosis pdf. The use of a Laplace prior yields a cost function with a $\ell 1$-norm while the use of a Cauchy prior, easily taken into account within the Bayesian framework, is more difficult to design into a cost function.

3.6. Optimization of regularization

Several strategies have been developed in order to optimize the regularization parameter η^2 which appears in the cost function of Eq. (21). The two most used strategies are the L-curve [14] and the GCV [13]. The first one simply consists in plotting the source norm and the residual norm in a two dimensional log space for several values of η^2 and searching the maximum of curvature of the resulting curve. In some cases, the L-curve shows multiple corners, that means multiple local maxima of curvature, and a wrong corner can be chosen for the optimization. For a given η^2, the GCV strategy consists in consecutively leaving out a measurement point and predicting it with all the others. The optimization is then performed by adjusting the regularization parameter η^2 which minimizes the residual sum between measurements and predictions. In some cases, the curve associated with this optimization is flat at its minimum and for a wide range of η^2. Therefore, the variability of the regularization parameter can be oversized. Recently, a new approach within the Bayesian framework has been proposed in Ref. [22]. The principle is to apply the Bayesian theorem to express the joint pdf of both precisions only conditionally to measurements and structural operator, without injecting any *a priori* information about precisions

$$\begin{aligned}\left[\tau_n, \tau_f|\boldsymbol{y}\right] &\propto \left[\boldsymbol{y}|\tau_n, \tau_f\right]\left[\tau_n, \tau_f\right] \\ &\propto \mathcal{N}_c\left(\boldsymbol{y}; \boldsymbol{0}, \tau_n^{-1}\mathbf{I} + \tau_f^{-1}\boldsymbol{D}^{-1}\boldsymbol{D}^{-H}\right).\end{aligned} \tag{22}$$

A change of variable is then applied to infer the *a posteriori* pdf of $\left[\eta^2, \tau_f|\boldsymbol{y}\right]$ followed by a marginalization step to only infer $\left[\eta^2|\boldsymbol{y}\right]$. The marginalization step consists in a projection of a multidimensional pdf on a subdimensional domain. The variable marginalized out is then considered as latent or hidden variable and is not expressed in pdf anymore, as defined by

$$\int [A|B][B]\mathrm{d}B = [A]. \tag{23}$$

Then, in the same way as the Tikhonov cost function in Eq. (21) is deduced from the *a posteriori* pdf on sources in Eq. (13), the Bayesian cost function for the regularization optimization is obtained by taking the opposite of the logarithm of $\left[\eta^2|\boldsymbol{y}\right]$. The regularization parameter is thus estimated by minimizing the following cost function

$$J(\eta^2) = \sum_{i=1}^{N} \ln\left(s_i^{-2} + \eta^2\right) + (N-2)\ln\left(\frac{1}{N}\sum_{i=1}^{N}\frac{|\boldsymbol{v}_i^H\boldsymbol{y}|^2}{s_i^{-2}+\eta^2}\right), \tag{24}$$

$$= \sum_{i=1}^{N} \ln\left(1+\eta^2 s_i^2\right) - \ln\left(s_i^2\right) + (N-2)\ln\left(\frac{1}{N}\sum_{i=1}^{N}\frac{s_i^2|\boldsymbol{v}_i^H\boldsymbol{y}|^2}{1+\eta^2 s_i^2}\right), \tag{25}$$

where N is the size of the structural operator matrix $\boldsymbol{D}$, s_i is still its i-th singular value, $\boldsymbol{v}_i$ is the i-th right singular vector – i.e. the i-th column of matrix $\boldsymbol{V}$ in Eq. (19) –and $\boldsymbol{y}$ is the measurement vector. The second formulation given in Eq. (25) illustrates better the relation with the regularization in Eq. (18), from the terms $1+\eta^2 s_i^2$ and $\mathbf{I}+\boldsymbol{D}\boldsymbol{D}^H\eta^2$. It should be noticed that this expression is obtained with the singular value decomposition of $\boldsymbol{D}$, the inverse problem matrix whose inputs are measurements and outputs are sources. For most of the inverse problem applications, the direct problem matrix $\boldsymbol{H}$ mapping the sources to the measurements is known instead of the inverse operator one. On the basis of SVD properties, the direct and inverse matrices are linked by the following relation,

$$\boldsymbol{H} = \boldsymbol{D}^{-1} \iff \boldsymbol{U}_h\boldsymbol{S}_h\boldsymbol{V}_h^H = \boldsymbol{V}_d\boldsymbol{S}_d^{-1}\boldsymbol{U}_d^H. \tag{26}$$

As a result,

$$s_{h_i} = s_{d_i}^{-1}, \quad \boldsymbol{u}_{h_i} = \boldsymbol{v}_{d_i}, \quad \boldsymbol{v}_{h_i} = \boldsymbol{u}_{d_i}. \tag{27}$$

To apply this optimization with the SVD of the direct matrix $\boldsymbol{H}$, $s_i^{-2} = s_{di}^{-2}$ should simply be substituted by $s_i^2 = s_{hi}^2$ and $\boldsymbol{v}_i = \boldsymbol{v}_{di}$ by $\boldsymbol{u}_i = \boldsymbol{u}_{hi}$ in Eq. (24), yielding the same expression as in Ref. [22]. This empirical Bayesian optimization, called the "Marginalized Maximum *A Posteriori*" and denoted hereafter as MMAP, provides the advantage of having a unique global minimum with parabolic shape locally [22]. This approach resolves issues faced by L-curve and GCV, for an equivalent time-efficiency.

4. Markov Chain Monte Carlo

Up to now, only the maximum probabilities are inferred to find optimal values while the rest of pdf contains information about uncertainties. Going further in the Bayesian approach, the whole pdf can be inferred but it would be very difficult and sometimes intractable analytically. The alternative is to evaluate the whole pdf in sufficient detail using suitable sampling algorithms such as MCMC methods [25], which are a class of random samplers whose runs converge to so-called stationary distributions. Histograms can be calculated from these stationary distributions to estimate dispersion of various random variables of the problem, making available credible intervals or propagation of uncertainties besides an automatic regularization. The aim is then to sample into $\left[\boldsymbol{f}, \tau_f, \tau_n|\boldsymbol{D}, \boldsymbol{y}\right]$, the joint density of all the random variables conditionally to all the

deterministic variables of the problem. Putting priors on *a priori* precisions τ_f and τ_n will act as a constraint for the optimization, that is to say a guide for exploring joint density. These so-called *hyperpriors* are chosen in a family of conjugate densities (i.e. such that the product of the prior and the likelihood remains in the same family as the posterior) and the variables are organized according to a hierarchical model to simplify the sampling algorithm.

4.1. Choice of sampling algorithm

Basically, a sample can be drawn from an arbitrary joint pdf on $\boldsymbol{f}$, τ_f and τ_n, called the proposal pdf, centered on an initial point and accepted or rejected by an acceptance rate depending on the target pdf $[\boldsymbol{f}, \tau_f, \tau_n|\boldsymbol{D}, \boldsymbol{y}]$ and the proposal one. If it is accepted, a new sample is drawn from the same arbitrary pdf but centered on the previous sample, making the chain converge. This corresponds to the Metropolis-Hastings algorithm [25]. With an inappropriate proposal pdf or when the dimension of the target pdf increases, the sample rate falls rapidly and the chain stagnates or diverges. To overcome this issue, the component-wise Metropolis-Hastings consists in sampling consecutively from unidimensional proposal pdf over each dimension independently from the others. By contrast, when variables are strongly dependent, unidimensional samplers are less efficient. When the target conditional pdf $[\boldsymbol{f}|\boldsymbol{D}, \boldsymbol{y}, \tau_f, \tau_n]$, $[\tau_f|\boldsymbol{D}, \boldsymbol{y}, \boldsymbol{f}, \tau_n]$ and $[\tau_n|\boldsymbol{D}, \boldsymbol{y}, \boldsymbol{f}, \tau_f]$ are sufficiently simple, the Gibbs sampling algorithm [27] consists in sampling directly from these conditional pdf instead of sampling from the proposal ones. Every sample is then accepted and the chain converges very quickly. The use of hierarchical model simplifies these target conditional pdf, as well as conjugate priors.

4.2. Hierarchical model

Hierarchical models [32] are a way to represent relationships between stochastic and deterministic variables of the problem into a multi-level Bayesian graph. It can also be viewed as a directed acyclic graph within the graph theory. For a specific variable of 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 the Bayes theorem, the pdf of a variable θ_j conditionally to all the others, denoted hereafter as $\infty_{-\theta_j}$, will only depend on its parents, children and co-parents of its children following this expression from Ref. [32]

$$\left[\theta_j|\infty_{-\theta_j}\right] \propto \left[\theta_j|\text{ Parents of } \theta_j\right] \times \prod_{\theta_k \text{ children of } \theta_j} [\theta_k|\text{ Parents of } \theta_k]. \tag{28}$$

This relation is then useful to reduce the number of conditioning variables of the target conditional pdf $[\boldsymbol{f}|\boldsymbol{D}, \boldsymbol{y}, \tau_f, \tau_n]$, $[\tau_f|\boldsymbol{D}, \boldsymbol{y}, \boldsymbol{f}, \tau_n]$ and $[\tau_n|\boldsymbol{D}, \boldsymbol{y}, \boldsymbol{f}, \tau_f]$.

4.3. Conjugate priors

When the likelihood distribution is from the exponential family, it is possible to design a prior, also from the exponential family, so that the posterior is in the same family as the prior. The prior is then a *conjugate prior* for the likelihood [33]. Depending on cases, the prior, and so the posterior, can be described by a common distribution for which a sampling algorithm is available, making the Gibbs sampler even more efficient. It turns out that the Gaussian distribution used in this work has an easy-sampled conjugate prior density for its dispersion parameter. Indeed, the inverse-gamma distribution is a conjugate prior for the variance of a likelihood Gaussian, or the gamma distribution can also be used for the precision. As it is easier to sample from the second one, the precision parameter seems to be a better choice to describe Gaussian pdf, as done in Section 3.2. The gamma density on precision τ with shape parameter k and rate parameter θ is expressed as

$$\mathcal{G}(\tau; k, \theta) = \frac{\tau^{k-1} \exp\left(-\frac{\tau}{\theta}\right)}{\Gamma(k)\theta^k} \text{ with } k > 0 \text{ and } \theta > 0 \tag{29}$$

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

4.4. Choice of hierarchical model

Designing a hierarchical graph is not as trivial as it appears and some tips should be given. Firstly, the noise snapshot $\boldsymbol{n}$ must not appear explicitly. Otherwise, the likelihood $[\boldsymbol{y}|\boldsymbol{f}, \boldsymbol{D}, \boldsymbol{n}]$ would be completely known (equivalent to a dirac distribution). The *a posteriori* of the source field $\boldsymbol{f}$ being the product of a prior and this likelihood, it would also be a dirac distribution. The other pdf would also be impacted in the same way. Each new sample being the same as the initialization value, the Gibbs sampler would be completely frozen. It is much more effective to introduce the noise precision τ_n (or the variance) by $[\boldsymbol{y}|\boldsymbol{f}, \boldsymbol{D}, \tau_n]$, letting the likelihood be uncertain. This time, the product with a prior generates a new pdf whose density is a trade off between prior and likelihood ones, the Gibbs sampler can then converge in this case. Secondly, there may be different ways to describe the physics. For example, the problem of Eq. (6) can be interpreted as a denoising procedure of $\boldsymbol{Dy}$ with a spatially correlated noise $\boldsymbol{Dn}$, or it can also be interpreted as a common inverse problem with a simple additive white

noise by using an alternative expression $\boldsymbol{y} = \boldsymbol{D}^{-1}\boldsymbol{f} + \boldsymbol{n}$. These two descriptions yield different hierarchical graphs, but after calculating each graph conditional pdf needed for the Gibbs sampler it turns out that they are completely equivalent. As different graphs can describe the same physics, it has been decided to use the one corresponding to a common inverse problem because of its generality. The corresponding graph is presented in Fig. 1, where square boxes stand for deterministic quantities while circular boxes correspond to stochastic variables whose *a posteriori* distributions are to be inferred. Arrows represent direct relationships between quantities and $\mathcal{N}$ and $\mathcal{G}$ stand for gaussian and gamma distributions respectively. (k_f, θ_f) are the hyperparameters of the source precision τ_f and (k_n, θ_n) the hyperparameters of the noise precision τ_n. These parameters have to be set depending on *a priori* information but carefully to avoid skewing the inference.

4.5. Conditional probability density functions

Following the graph of Fig. 1 and applying Eq. (28) on stochastic variables, the target conditional pdf are

$$[\boldsymbol{f}|\boldsymbol{D}, \boldsymbol{y}, \tau_f, \tau_n] \propto \mathcal{N}_c\left(\boldsymbol{f}; \boldsymbol{0}, \tau_f^{-1}\mathbf{I}\right)\mathcal{N}_c\left(\boldsymbol{y}; \boldsymbol{D}^{-1}\boldsymbol{f}, \tau_n^{-1}\mathbf{I}\right) \tag{30}$$

$$\propto \mathcal{N}_c\left(\boldsymbol{f}; \boldsymbol{\mu_f}, \boldsymbol{\Sigma_f}\right) \tag{31}$$

with :

$$\boldsymbol{\Sigma_f} = \boldsymbol{D}\boldsymbol{D}^H\left(\boldsymbol{D}\boldsymbol{D}^H\tau_f + \tau_n\mathbf{I}\right)^{-1} \tag{32}$$

$$\boldsymbol{\mu_f} = \left(\boldsymbol{D}\boldsymbol{D}^H\eta^2 + \mathbf{I}\right)^{-1}\boldsymbol{D}\boldsymbol{y} \tag{33}$$

which finally are the same expressions as in Eqs. (15) and (16) but expressed with $\boldsymbol{D}$ instead of $\boldsymbol{D}^{-1}$,

$$[\tau_n|\boldsymbol{D}, \boldsymbol{y}, \boldsymbol{f}, k_n, \theta_n] \propto \mathcal{G}(\tau_n; k_n, \theta_n)\mathcal{N}_c\left(\boldsymbol{y}; \boldsymbol{D}^{-1}\boldsymbol{f}, \tau_n^{-1}\mathbf{I}\right) \tag{34}$$

$$\propto \mathcal{G}\left(\tau_n; k_n + N, \left(\theta_n^{-1} + \left\|\boldsymbol{y} - \boldsymbol{D}^{-1}\boldsymbol{f}\right\|_2^2\right)^{-1}\right) \tag{35}$$

which only depends on the residuals norm, the number of measurement points N and its hyperpriors k_n and θ_n,

$$[\tau_f|\boldsymbol{f}, k_f, \theta_f] \propto \mathcal{G}(\tau_f; k_f, \theta_f)\mathcal{N}_c\left(\boldsymbol{f}; \boldsymbol{0}, \tau_f^{-1}\mathbf{I}\right) \tag{36}$$

$$\propto \mathcal{G}\left(\tau_f; k_f + N, \left(\theta_f^{-1} + \|\boldsymbol{f}\|_2^2\right)^{-1}\right) \tag{37}$$

which similarly only depends on the source norm, its hyperpriors and the number of source points. It can be observed that if the shape parameter k tends towards zero while the rate parameter θ approaches infinity, the conditional pdf on precisions in Eqs. (35) and (37) would no longer depend on hyperpriors. Thus, the hyperpriors would no longer compel the exploration of pdf, the resolution would be totally empirical. It should be noticed that the Gamma distribution is not defined theoretically for $k = 0$ and this shape parameter is often set to $k = 1$. However, this non-zero value is equivalent to injecting *a priori* information, as the expected value for τ is θ and the mode of its distribution is 0. If the prior is not evaluated directly but only the posterior is calculated, it is considered that improper prior can be used if the posterior is well defined. In this way, it is then allowed to set the hyperparameters to $k_f = k_n = \theta_f^{-1} = \theta_n^{-1} = 0$. This configuration is used in Sections 5.3.4 and 6.4 in order to compare the empirical Bayesian optimization from Eq. (25) and the fully Bayesian approach. In brief, putting hyperpriors accelerates convergence of the resolution but it can add a bias if the constraint is too high. Using the SVD of the structural operator and a new parametrization can drastically reduce the numerical cost of each sample as proposed in Appendix A.1. Actually, the most costly step is the SVD of the operator matrix, which is also required for the other optimization pro-

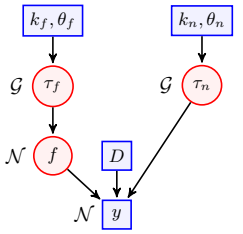

Fig. 1. Hierarchical graph of the vibration source identification problem. Square boxes stand for deterministic quantities while circular boxes correspond to stochastic variables whose *a posteriori* distributions are to be inferred. Arrows represent direct relationship between quantities and $\mathcal{N}$ and $\mathcal{G}$ stand for gaussian and gamma distributions respectively.

cedures presented in this paper. This decomposition has to be computed only once. Thus, the computational complexity is of the same order as other optimization procedures and the additional cost depends on the number of samples requested to reach convergence.

5. Simulations

This section illustrates the previous developments. The displacement field of a simply supported steel beam excited by an harmonic point source is calculated analytically and then corrupted by noise. The Force Analysis Technique is applied to it to demonstrate the high sensitivity of the method to additive noise when not regularized. Pathological L-curve and GCV optimizations are shown in comparison with an adequate empirical Bayesian optimization. Further developments on Gibbs sampling are then presented, including the automatization of sampling and the process to obtain credible intervals. A pseudo code of the algorithm is proposed in Appendix A.2.

5.1. Synthesis of noisy displacements

The harmonic transverse displacement field of a simply supported beam excited by a point source is computed by using an analytical wave decomposition approach [34].

$$w(x,\omega)=\begin{cases}\frac{A}{2k_w^3E(1+j\eta)I}\left(\sin(k_wx)\frac{\sin(k_w(L-x_0))}{\sin(k_wL)}-\sinh(k_wx)\frac{\sinh(k_w(L-x_0))}{\sinh(k_wL)}\right), & \text{for } x\in[0;x_0]\\ \frac{A}{2k_w^3E(1+j\eta)I}\left(\sin(k_w(L-x))\frac{\sin(k_wx_0)}{\sin(k_wL)}-\sinh(k_w(L-x))\frac{\sinh(k_wx_0)}{\sinh(k_wL)}\right), & \text{for } x\in[x_0;L]\end{cases}, \tag{38}$$

with L the beam length, A and x_0 the amplitude and the location of the single point force, $k_w=\left(\frac{\rho S}{E(1+j\eta)I}\omega^2\right)^{\frac{1}{4}}$ the flexural wavenumber controlled by the driving angular frequency of the source and structural and geometrical properties of the beam. Following Eq. (5), an additive white noise $\boldsymbol{n}$ is then added to the simulated displacement as

$$\boldsymbol{n}=\frac{(\boldsymbol{\epsilon}'+j\boldsymbol{\epsilon}'')}{\sqrt{2}}\sigma_w 10^{-\frac{\mathrm{SNR}}{20}} \tag{39}$$

with $\boldsymbol{\epsilon}'$ and $\boldsymbol{\epsilon}''$ two independant random vectors with the same size as w and sampled from standard normal distribution, σ_w the standard deviation of the simulated displacement $\boldsymbol{w}$ and SNR the signal-to-noise ratio in decibels (dB). The parameters used in this simulation are listed in Tables 1 and 2. The resulting noise-free analytical displacement field and the observed field obtained after adding a noise with SNR = 15 dB are presented in Fig. 2. In this simulation, the arbitrary source frequency does not correspond to a resonance because the inverse problem is known to be particularly ill-posed at resonances. Mathematically, is occurs because the displacement at resonance is solution of the equation of motion without external excitation. Physically, this issue can also be explained by the fact that at resonance, the same displacement can be observed from different source location and amplitude, so there is no unique solution. However, considering damping property of the structure, there is no more singularity and the inverse problem can be (roughly) solved.

5.2. Force Analysis Technique

The operator matrix $\boldsymbol{D}$ is then constructed from geometrical and structural parameters of the beam of Table 1 and from the frequency of Table 2. As explained in Section 2.3, forward and backward finite difference schemes are used at boundaries of the studied subdomain. Fig. 3 shows that the location and the amplitude (after a spatial integration) of the point source are revealed with accuracy and precision when this operator matrix is applied to the noisefree displacement $\boldsymbol{w}$ while the point source is totally covered by aberrant forces when applied to the noisy displacement vector $\boldsymbol{y}$. In practice, $\boldsymbol{w}$ is never known, so a regularization step is generally needed to extract information about sources from Fig. 3b.

5.3. Regularization

Regularization by means of the L-curve, GCV, MMAP and MCMC optimizations are compared in this section in order to illustrate the improvement brought by the Bayesian framework. The first two methods are performed by using the Hansen's RegTools package [35]. The following results are obtained from the noisy displacement field of Fig. 2.

5.3.1. L-curve regularization

Fig. 4 shows the L-curve on a 2D plot with residual and source norms as axes, computed for several values of the regularization parameter η^2. The optimal value – i.e. the maximum of curvature – is identified as $\eta^2_{Lcurve}=3.02\times10^{-14}$ but other local curvature maxima are visible, for example near $\eta^2=10^{-10}$, thus illustrating the multimodality of this criterion.

Table 1
Geometrical and structural characteristics of the simulated steel beam.

Length L [m]	Moment of inertia I [m^4]	Section S [m^2]	Spatial discretization Δ_x [m]
1	$\frac{4}{3} \times 10^{-11}$	4×10^{-5}	1×10^{-2}
Young modulus E [N/m^2]	Structural damping η	Mass density ρ [kg/m^3]	
210×10^9	1×10^{-3}	7800	

Table 2
Characteristics of the simulated point source.

Location x_0 [m]	Amplitude A [N]	Frequency ν [Hz]
0.40	1	250

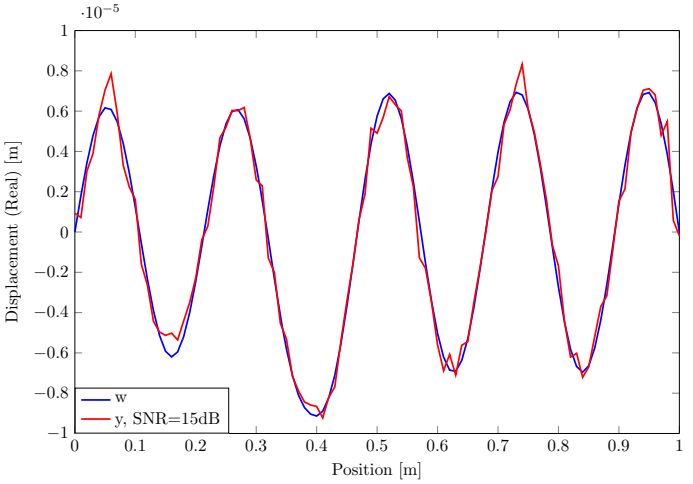

Fig. 2. Real part of the simulated displacement field of the steel beam (w) and of the observed field (y) obtained after adding a noise with SNR = 15 dB.

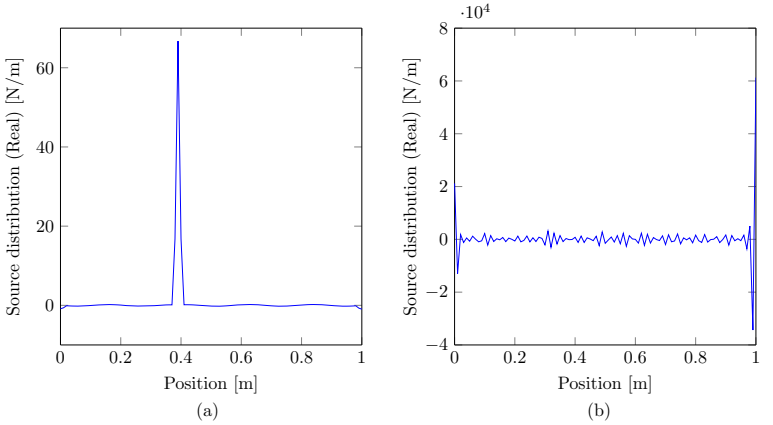

Fig. 3. Real part of the source distribution (N/m) calculated from the FAT, without (a) and with (b) noise added.

5.3.2. GCV regularization

Fig. 5 represents the GCV cost function whose regularization parameter optimum value is found at $\eta^2_{GCV} = 2.61 \times 10^{-15}$. This cost function has only one global minimum but the precision of this criterion is quite large due to the flat shape near the minimum.

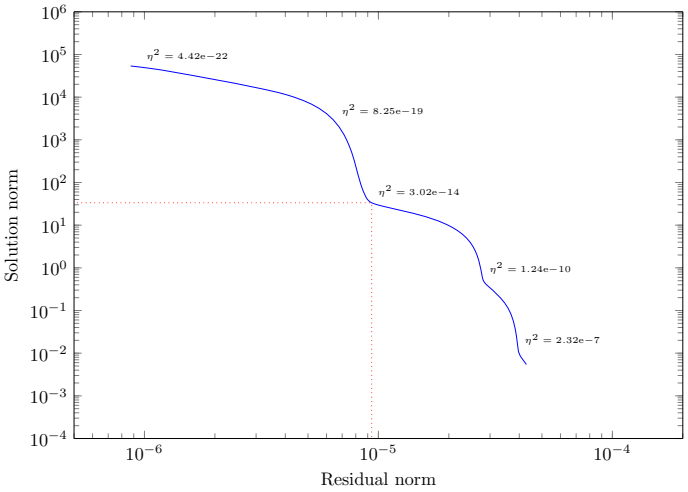

Fig. 4. Cost function of the L-curve optimization, with optimized regularization parameter $\eta^2_{Lcurve} = 3.02 \times 10^{-14}$.

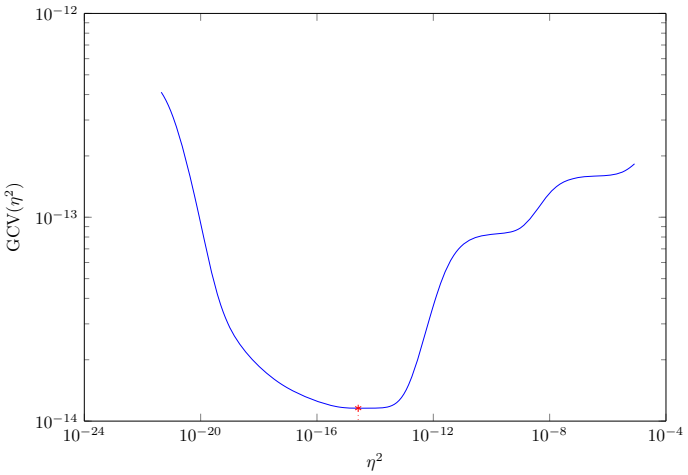

Fig. 5. Cost function of the generalized cross validation (GCV) optimization, with optimized regularization parameter $\eta^2_{GCV} = 2.61 \times 10^{-15}$.

5.3.3. MMAP regularization

The MMAP cost function is computed from Eq. (25) and shown in Fig. 6. There is also a unique global minimum but, contrary to the GCV cost function, the parabolic shape at its minimum allows a better precision. The optimal value $\eta^2_{\text{MMAP}} = 8.11 \times 10^{-15}$ is between the L-curve and the GCV values and can thus be seen as a tradeoff between these two common criteria.

5.3.4. MCMC regularization

The MCMC regularization is perform by the Gibbs sampling and conditional pdf given in Eqs. (30)–(33), (32), (35)–(37). Successive samples are added to Markov chains and allows the estimation of credible intervals of parameters from histograms besides automatic regularization. In order to preserve the empirical approach used by previous criteria, *a priori* parameters of precisions τ_n and τ_f are set to $k_n = k_f = \theta_n^{-1} = \theta_f^{-1} = 0$. In this case, the only maintaining *a priori* information is the shape of the distribution but samples obtained from Eqs. (35) and (37) now depend only on the likelihood.

Initialization of chains. The better the initialization, the faster the convergence. Conversely, a bad initialization slows down the convergence and can even make the chains completely diverge in some cases. We mean by "under/over regularization" the initialization with under/over estimation of the parameter. Fig. 7 shows multiple chains on τ_n, τ_f and η^2 for two different initializations: over initialized τ_n (red) and under initialized τ_n (blue), τ_f being always well initialized. For the under initialization of τ_n, the chains converge quickly to the expected values. In the case of over initialization of τ_n, errors are first propagated to τ_f and η^2 before all chains finally converge in a much higher number of iterations. The reason of this observation is

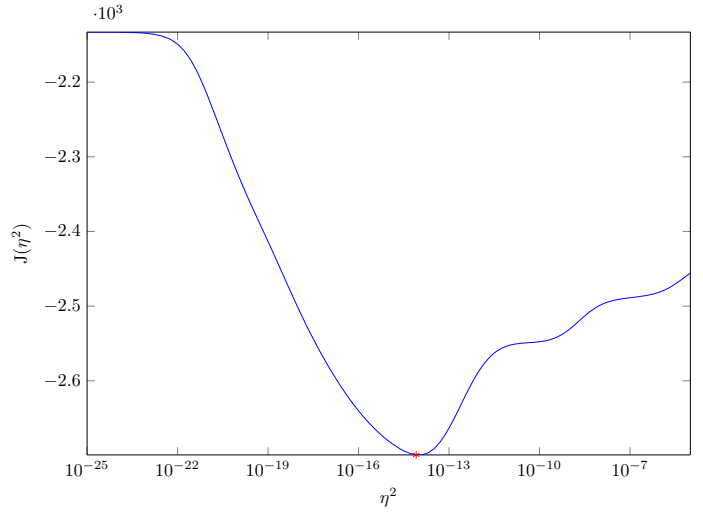

Fig. 6. Cost function of the MMAP optimization, with optimized regularization parameter $\eta^2_{\text{MMAP}} = 8.11 \times 10^{-15}$.

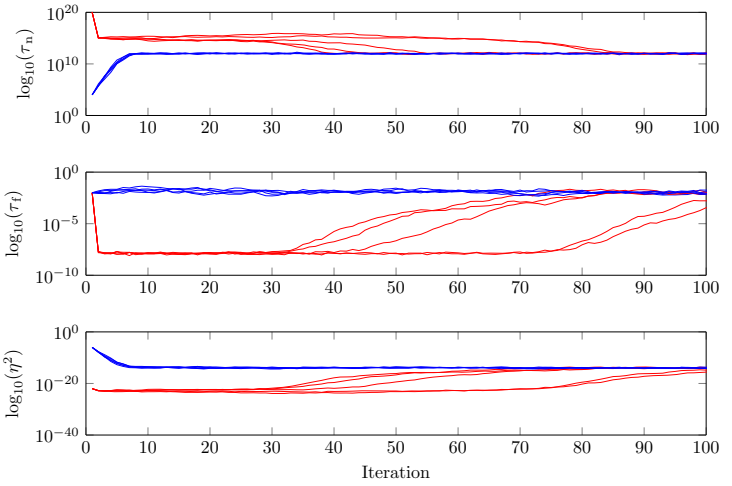

Fig. 7. Markov Chains on τ_n, τ_f and η^2 for two different initializations. Over initialized τ_n (red) and under initialized τ_n (blue), τ_f is always well initialized. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

that the step between two consecutive samples is directly linked to precisions τ_f and τ_n in the Gibbs sampling. If they are too high, chains are treading water. When there is no *a priori*, it is then much better to under initialize than to over initialize. In practice, the L-curve, GCV and MMAP indicators can be evaluated in a first step, then the median value is selected to avoid over or under regularization. The source distribution can then be initialized from Eq. (20) with the median regularization parameter, noted $\boldsymbol{f}_{init}$. The source precision can be directly initialized from $\|\boldsymbol{f}_{init}\|_2^{-2}$. The noise precision is initialized from the comparison between the measurements $\boldsymbol{y}$ and the displacement resulting from the initialized source distribution by $\|\boldsymbol{y} - \boldsymbol{D}^{-1}\boldsymbol{f}_{init}\|_2^{-2}$.

End of chains. To design an algorithm even more unsupervised, an automatic procedure can be added to end sampling when a convergence criterion reaches a target value. Several strategies are proposed in the scientific literature, such as the monitoring of the chain autocorrelation, its mixing [36], a stationary test [37], the comparison of values in the early part of the chain to those in the latter part of the chain [38] or the study of accurary percentiles [39]. Nevertheless, these methods only ensure that the chain has converged, but they do not guarantee that the pdf has been widely explored, so the chain can be stuck in a local maximum of the pdf. The convergence criterion used in this work is then based on multiple sequences of the Gibbs sampling with different chains initializations and monitoring of chains variances [40]. When all of the Gibbs samplers have converged to the same stationary distribution, it can be considered that the global maximum of the pdf has been reached. Within-chain and between-chain variances of second-half sample chains of a specific variable are used to calculate a Potential Scale Reduction Factor (PSRF) which monitors the convergence of this specific variable [40]. When the PSRF tends

to 1, convergence is achieved and all the second-half samples can be concatenated to improve histograms. If convergence is too fast, histograms can be inaccurate so a minimum number of iterations should be imposed. The regularization parameter η^2 being the ratio of noise and source precisions, it can be close to the optimum even if both precisions are wrong (if they have the same multiplicative error). It is then preferable to monitor both precisions chains instead of the regularization parameter ones. Fig. 8 shows three different initializations for τ_n and τ_f. The target value for the PSRF is set to 1.01 to consider the algorithm has converged. The vertical line delineates then the transition from the so-called burn-in time (converging chains) to stationnary distribution (converged chains). Here, the minimum number of iterations is set to 450 to refine parameters histograms even if 200 would be enough. In this work, both precisions are monitored and three chains are used to compute the PSRF indicator. A first noise precision $\tau_{n_{init}}^{(1)}$ is initialized from $\|\boldsymbol{y}-\boldsymbol{D}^{-1}\boldsymbol{f}_{init}\|_2^{-2}$ as explained before. The two other noise precisions are initialized with a gain of ± 15 dB by multiplying $\tau_{n_{init}}^{(1)}$ by $10^{\pm\frac{15}{10}}$. This way, the Gibbs sampler covers a wide measurement quality range. It is then considered that the noise precision perturbation propagates to the other variables, so the source distribution and its precision are initialized with the same value for the three chains, following the procedure detailed in the previous paragraph.

Histograms of both precisions and their ratio η^2 are evaluated from stationary distributions and are presented in Fig. 9 with gamma distribution fits. To test the model sensitivity, these estimated distributions can then be used to perform a more realistic propagation of uncertainty, rather than simply using Gaussian or uniform distributions with *a priori* parameters. Besides, the noise level (i.e. τ_n^{-1}) is a direct indicator of measurement quality and is assessable by Gibbs sampling. As the η^2 histogram is not symmetric, the distribution mode is no longer equivalent to the mean value and to the median indicator. The maximum of the regularization parameter η^2 is then evaluated as the maximum of the associated gamma distribution fit, $\eta^2_{\mathrm{MCMC}} = 1.04 \times 10^{-14}$. The GCV, MMAP and L-curve results are represented by vertical lines on the histogram of η^2 for comparison. The MMAP result is the closest value to the MCMC one and could even be considered as equivalent if the mode of the histogram was as the reference value instead of the maximum of the gamma distribution fit.

Regularized source distribution. Fig. 10 presents the concatenation of corresponding converged chains over the source field. The point source is clearly visible close to $x = 0.4$ m. As this algorithm deals with Gaussians, the MAP estimator can be calculated either by the mean value or by the median. Although these two indicators are asymptotically equivalent, the second one is preferable because it is less sensitive to extreme values when there are few samples.

Credible intervals can also be estimated from chains quantiles, as shown in Fig. 11. The higher is the number of samples, the less disturbed is this source field (with a lack of samples, the empirical median is not totally stabilized). Another possible way is to use the MAP of η^2 obtained by Gibbs sampling and to inject it into Eq. (20) just like for the other indicators. Both fields are quite similar and illustrate the equivalence between the empirical Bayesian cost function of Eq. (25) and the fully Bayesian application with no *a priori* provided by Gibbs sampling. It should be noticed that it would have required several

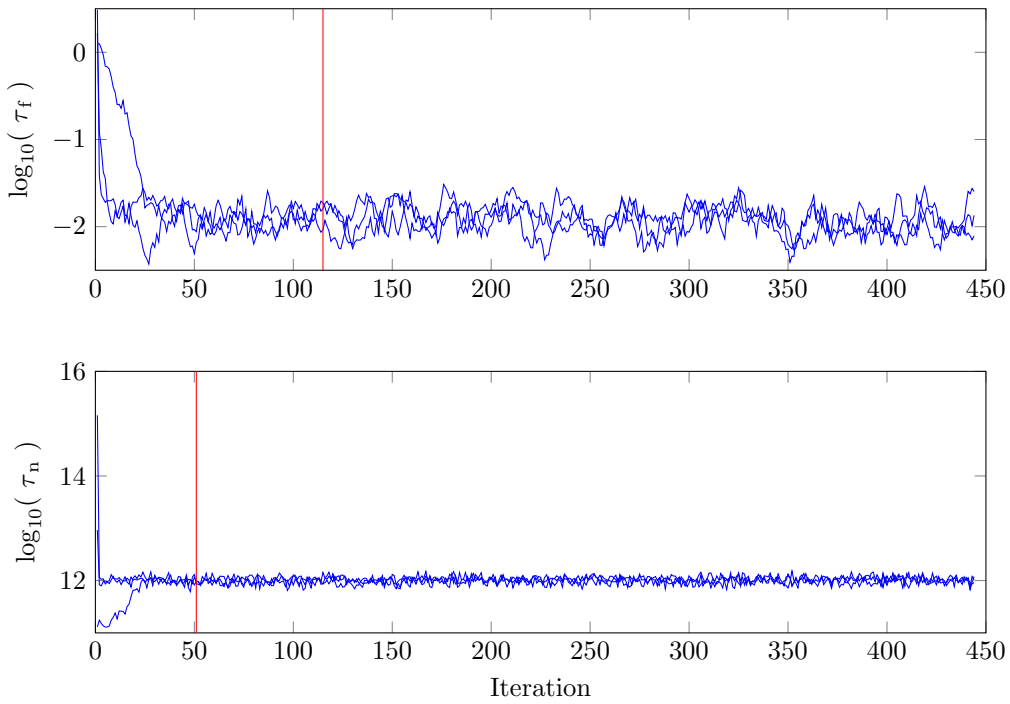

Fig. 8. Multiple Markov chains of noise precision τ_n and source precision τ_f with borderline between burn-in time and stationary distribution.

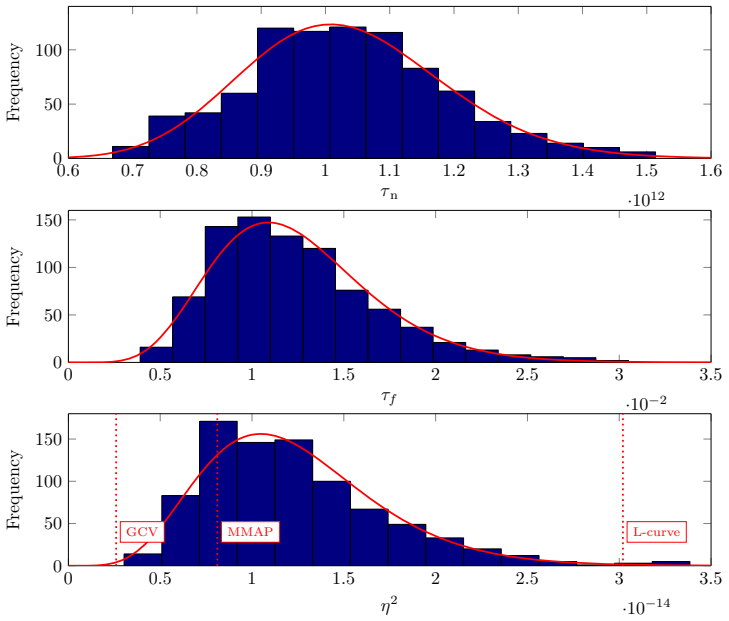

Fig. 9. Histograms of τ_n, τ_f and η^2 with gamma distribution fits.

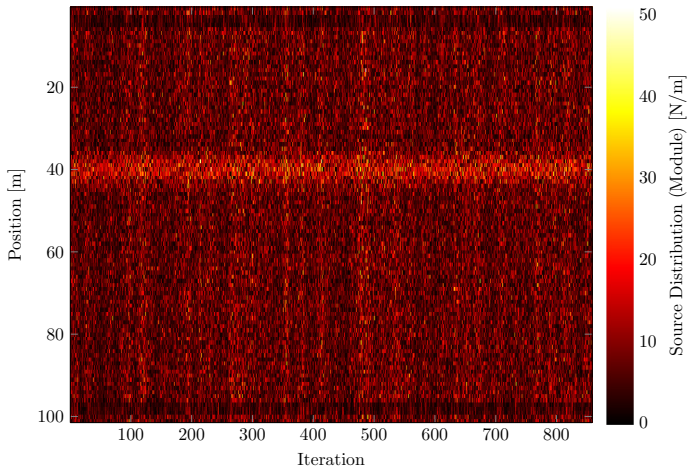

Fig. 10. Concatenation of stationary distribution of multiple Markov chains of source distribution.

measurements to obtain these intervals with traditional methods while a unique scan of the structure is sufficient within the Bayesian framework. These intervals can be used to judge solution relevance. However, the prior defined in Eq. (11) inherently leads to smooth solutions, because such prior does not promote sparsity. Then, the actual sparse source distribution used for this numerical simulation may not lie in the credible intervals everywhere. Hence, credible intervals are prior-dependent (contrary to confidence intervals from a frequentist point of view) and should be interpreted with caution. However, if the source amplitude is the quantity of interest, the spatial integration corrects the smoothing effect of the prior. As it will be shown below, the source amplitude can thus be properly recovered. Nevertheless, these intervals can be used for example to detect automatically the presence of sources. At each space point, if a credible interval of $p\%$ does not include 0, then there is a credibility of $p\%$ that there is actually a source there. This source presence credibility depends on the noise level. Indeed, the intervals increase with the noise, so the credibility to include 0 increases too, so the credibility of a source presence decreases. Depending on a false positive or a false negative tolerated rate, a threshold can be applied to the resulting detection credibility map to assess the presence of sources.

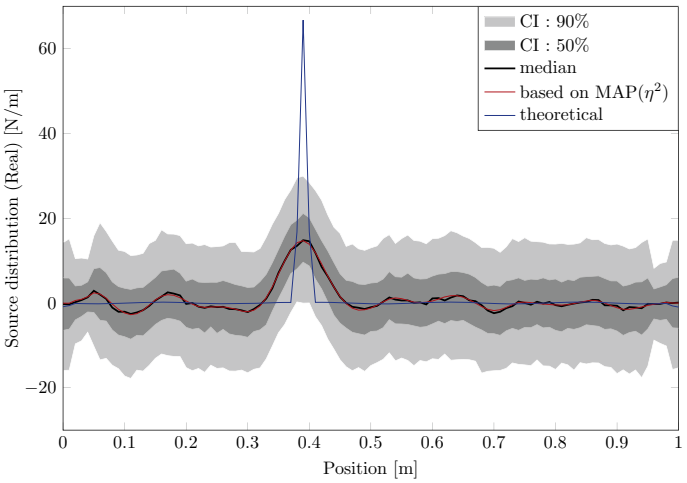

Fig. 11. Real part of the regularized source distribution from Gibbs sampling with credible intervals together with the MAP solution with empirical Bayesian regularization and with the theoretical solution.

5.3.5. MCMC samples illustrated on L-curve and GCV cost function

Fig. 12 illustrates the MCMC samples of a unique Gibbs sampler (i.e. a unique pair of Markov Chains from Fig. 8 in the same representation as the L-curve and the GCV cost function. To compare with the L-curve, the Euclidean norm of each source distribution sample $\boldsymbol{f}^i_{mcmc}$ is computed, where the exponent i corresponds to the i-th iteration of the Gibbs sampler, as well as the reconstruction error norm from $\boldsymbol{y} - \boldsymbol{D}^{-1}\boldsymbol{f}^i_{mcmc}$ instead of using the Hansen's RegTools function. The difference between the L-curve and the MCMC samples could be explained by the fact that the precisions ratio, and so the regularization parameter, is conserved when the same term multiplies both numerator and denominator. Thus, MCMC samples seems to converge towards a higher solution and residual norms than L-curve ones. To compute the GCV cost function from MCMC samples, only the regularization parameter is required instead of using the two precisions. The calculation is then equivalent to the Hansen's RegTools function, MCMC samples and GCV cost function are overlapped. In both cases, the convergence area of MCMC samples is closed but not exactly the same as the optimum values of L-curve and GCV cost function.

5.3.6. Comparison of source fields

Fig. 13 compares the source fields recovered from the four indicators. Even if they are quite close, the L-curve solution tends to over-regularize and the GCV optimization to under-regularize, while the two Bayesian approaches yield almost the same field. As it will be exposed in the next section, these small differences are verified in most cases but can be significantly amplified for some specific frequencies.

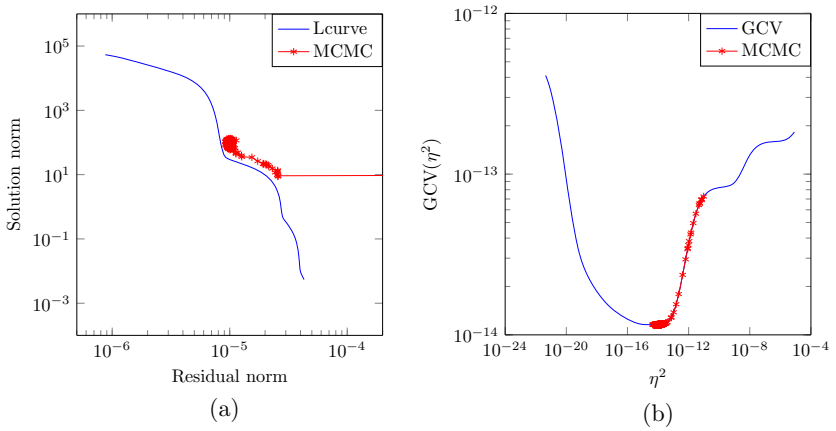

Fig. 12. Comparison between L-curve and MCMC (a), between GCV and MCMC (b).

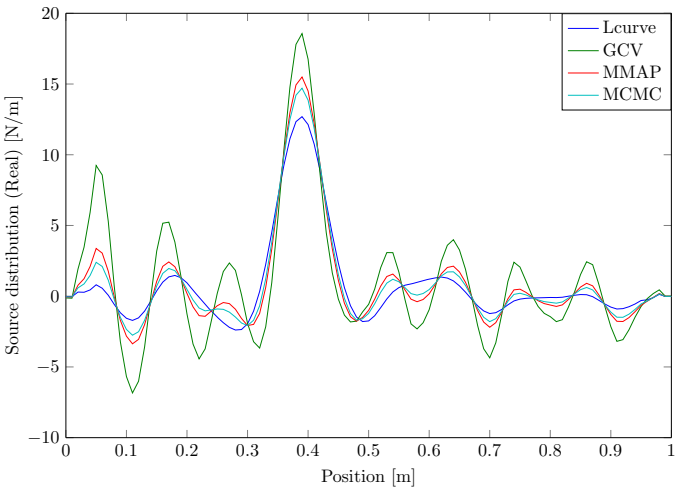

Fig. 13. Real part of the regularized source distribution recovered from L-curve, GCV, MMAP and MCMC.

6. Experimental validations

The applicability of the FAT and Bayesian methods is presented here on experimental data and for a wide frequency range.

6.1. Experimental setup

A freely suspended aluminum beam, whose geometrical and structural characteristics as well as mesh characteristics are listed in Table 3, is excited by a shaker located at $x_0 = 0.37$ m. A force sensor is placed at the interface to acquire a reference signal. The excitation signal is a linear chirp signal in frequency range $[100; 4000]$ Hz with frequency resolution $\Delta_f = 0.625$ Hz, the displacement measurement is performed by a scanning laser vibrometer. Pictures in Fig. 14 show the experimental setup.

6.2. Measurements of noisy displacements

Fig. 15 presents a space-frequency map of the measurements $y(x, \omega)$ with a logarithmic scale on the module, where curved lines correspond to nodal lines. On this simple analytical structure, the force location can easily be evaluated between 0.35 m and 0.40 m but the quantification of injected effort is not as direct.

6.3. Regularization parameters

For each frequency in the range $[100; 4000]$ Hz, the four regularization optimization criteria are calculated, as illustrated in Fig. 16. L-curve and GCV solutions are unstable at several frequencies whereas the two Bayesian solutions are overall stable and almost overlapped. Wrong L-curve optimization tends to force the source field to zero while wrong GCV optimization tends not to filter residuals.

Table 3
Structural and geometrical characteristics of the experimental aluminum beam and characteristics of the mesh.

Young modulus E [N/m^2]	Mass density ρ [kg/m^3]	Structural damping η
70×10^9	2700	1×10^{-4}
Length L [m]	Moment of inertia I [m^4]	Section S [m^2]
72.25×10^{-2}	6.02×10^{-11}	8.58×10^{-5}
Mesh length L_m [m]	Number of nodes N_m	Spatial discretization Δ_x
59.32×10^{-2}	105	5.7×10^{-3}

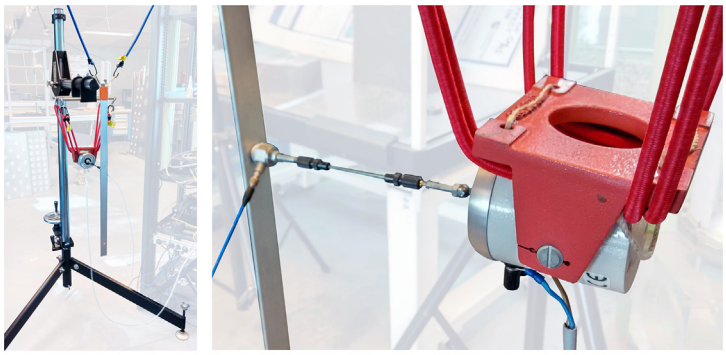

Fig. 14. Experimental setup of the free beam excited by a shaker with a linear chirp signal in frequency range [100; 4000] Hz. A force sensor is placed at the interface between the beam and the shaker to acquire a reference signal.

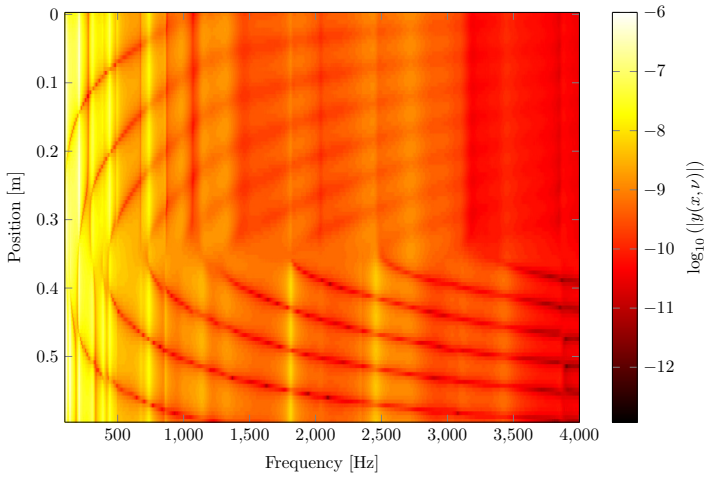

Fig. 15. Map of the measured transverse displacement of a portion of the beam Y as a function of the frequency. Amplitude is in logarithmic scale.

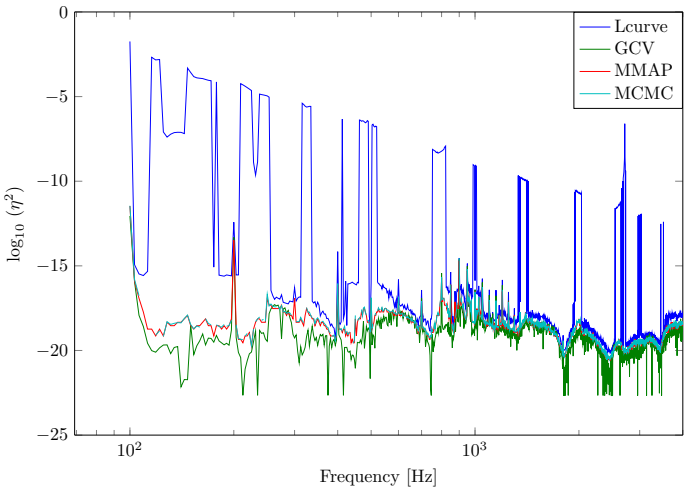

Fig. 16. Regularization parameter optimized by L-curve, GCV, MMAP and MCMC as a function of frequency.

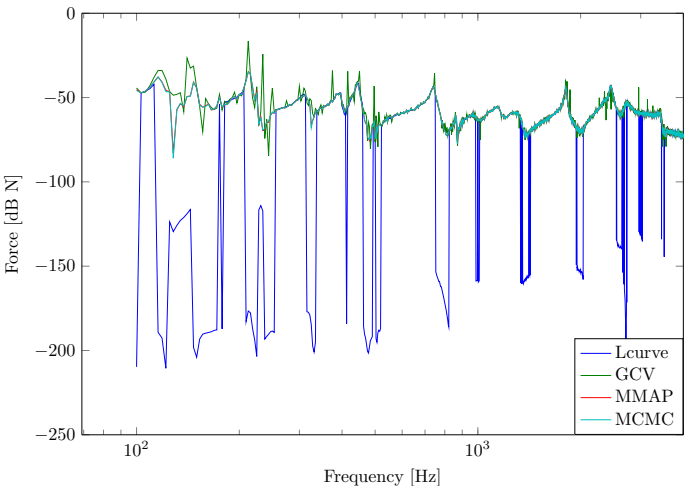

Fig. 17. Source level identified by L-curve, GCV, MMAP and MCMC as a function of frequency.

6.4. Source identification

As first explained in Section 2.2, the identified source field corresponds to a source distribution in N/m and should be spatially integrated around each source to obtain its force amplitude in Newton, following

$$\widehat{A}_\omega = \Delta_x \sum_{i=n_s}^{n_e} \hat{f}(x_i, \omega) \tag{40}$$

with $\widehat{A}_\omega$ the identified force amplitude of a specific source, n_s and n_e the limit nodes for the spatial integration around the specific source and $\hat{f}(x_i, \omega)$ the identified source distribution. A unique point source is considered in this experimental illustration, so the injected force can be obtained by integrating all over the mesh, i.e. with $n_s = 1$ and $n_e = N_m$. Repeating this step for each frequency, Fig. 17 shows the injected force as a function of frequency and for the four criteria. Logically, the Bayesian solutions are almost equivalent and more stable than L-curve and GCV ones.

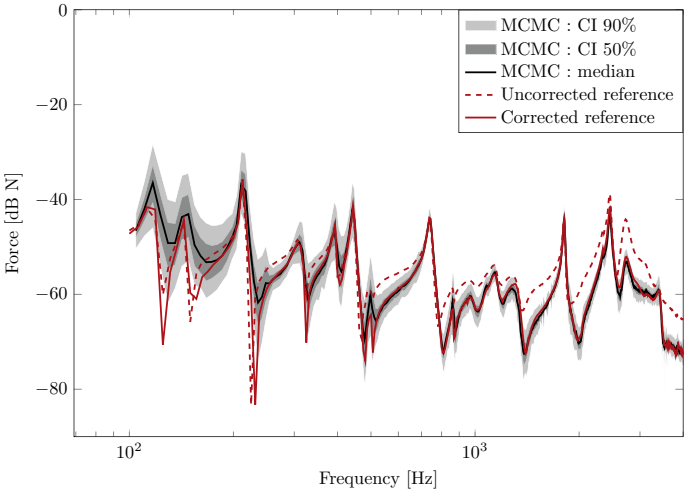

Fig. 18. Source level identified by MCMC with credible intervals as a function of frequency, compared to reference signal before and after the inertial correction.

6.5. Credible intervals and comparison with reference signal

Three main steps are required to obtain the force amplitude in dB from the spatial force distribution: the computation of a sum, a magnitude and a logarithm. Then, it could be hazardous to determine credible intervals on the force amplitude directly by integrating the force distribution with credible intervals from Fig. 11. To propagate the credibility of the force spatial distribution, polynomial chaos could be used. However, the samples must be resumed to a well-known pdf. Also, the pdf of a variable magnitude logarithm is not easily reachable analytically. Sampling strategies could also be applied (for example a Metropolis-Hastings step). Rather than drawing force amplitude samples from a force distribution obtained itself by a sampling strategy, it is more practical to reuse the results of Fig. 10 and to calculate force amplitude samples directly from force distribution samples. This time, each force spatial distribution draw from the stationary distribution of the Markov chain is spatially integrated and quantiles are calculated from all these resulting source level samples. Because the MMAP and MCMC curves are quite similar and because L-curve and GCV ones are unstable, only MCMC results are compared to the reference signal of the force sensor at the interface (see Fig. 18). The raw signal reference is close to the MCMC results only at peaks and far from it near troughs. It must be noticed that the FAT identifies the force actually injected to the structure while the force sensor can be biased by its own inertia. A method to correct the reference signal is proposed in Ref. [41] and applied here. It consists in identifying first the sensor mass from two measurements with two different calibrated masses, then multiplying it for all the frequencies by the measured acceleration of the beam corresponding to the maximum of identified source field, and finally in subtracting this product from the raw signal of reference. After this correction, the source level identified by the inverse method is quite close to the corrected reference signal. At least, the latter is almost always included in the 50% credible interval.

7. Conclusion and perspectives

In the present paper, significant improvement in terms of robustness has been noticed for the vibration source characterization problem by using an empirical Bayesian criterion instead of the L-curve or the GCV regularizations which, for specific frequencies, under-regularizes and over-regularizes respectively. A fully Bayesian procedure has also been set up by adding *a priori* pdf on noise and source precisions, and by identifying the whole *a posteriori* pdf on unknown variables by the Gibbs sampler, a particular case of MCMC algorithms. With an empirical parametrization (i.e. non informative hyperpriors on both precisions), the fully Bayesian procedure yields the same most probable source field as the one obtained by using the empirical Bayesian criterion. This observation reinforces the relevance of the empirical Bayesian criterion which has been criticised for not solving the problem in its entirety. Credible intervals can also be identified from the fully Bayesian procedure, on the spatial source distribution as well as on scalar variables such as the source amplitude (in Newton) or the source and noise precisions. Information about the variability of each variable can be useful for propagation of uncertainties and sensitivity analyzes or for assessing measurement quality. These two Bayesian approaches are completely unsupervised and can be applied easily to a more complex structure requiring a finite element model. The extension of the fully Bayesian approach with sparse prior on sources is a promising outlook for the separation of spatially close sources, which might be used for example for the structure-borne noise characterization of a bolted source at multiple attachment points.

Acknowledgements

This work is funded by the TESSA consortium including Vibratec, CEVAA, Pierburg Pump Technology, PSA Peugeot Citroën, Renault, Renault Trucks, Valeo, Sonorhc, CETIM, Centrale-Lyon, INSA-Lyon and Université du Maine.

Appendix A. Matrix inversion lemma

The matrix inversion lemma (or Woodbury identity) is given by

$$(\boldsymbol{A}+\boldsymbol{BCE})^{-1}\boldsymbol{BC}=\boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{C}^{-1}+\boldsymbol{EA}^{-1}\boldsymbol{B}\right)^{-1}. \tag{41}$$

Let us introduce,

$$\boldsymbol{A}=\eta^2\mathbf{I}, \tag{42}$$

$$\boldsymbol{B}=\mathbf{I}, \tag{43}$$

$$\boldsymbol{C}=\boldsymbol{D}^{-H}\boldsymbol{D}^{-1}=\left(\boldsymbol{D}\boldsymbol{D}^{H}\right)^{-1}, \tag{44}$$

$$\boldsymbol{E}=\mathbf{I}. \tag{45}$$

The expression of the regularized sources given in Eq. (17) can be expressed as

$$\hat{\boldsymbol{f}}_{MAP} = \left(\boldsymbol{D}^{-H}\boldsymbol{D}^{-1} + \eta^2\mathbf{I}\right)^{-1}\boldsymbol{D}^{-H}\boldsymbol{y}, \tag{46}$$

$$= \left(\boldsymbol{D}^{-H}\boldsymbol{D}^{-1} + \eta^2\mathbf{I}\right)^{-1}\boldsymbol{D}^{-H}\underbrace{\boldsymbol{D}^{-1}\boldsymbol{D}}_{\equiv\mathbf{I}}\boldsymbol{y}, \tag{47}$$

$$= (\boldsymbol{A}+\boldsymbol{C})^{-1}\boldsymbol{C}\boldsymbol{D}\boldsymbol{y}, \tag{48}$$

$$= \boldsymbol{A}^{-1}\left(\boldsymbol{C}^{-1} + \boldsymbol{A}^{-1}\right)^{-1}\boldsymbol{D}\boldsymbol{y}, \tag{49}$$

$$= \eta^{-2}\left(\boldsymbol{D}\boldsymbol{D}^{H} + \eta^{-2}\mathbf{I}\right)^{-1}\boldsymbol{D}\boldsymbol{y}, \tag{50}$$

$$= \left(\boldsymbol{D}\boldsymbol{D}^{H}\eta^2 + \mathbf{I}\right)^{-1}\boldsymbol{D}\boldsymbol{y}, \tag{51}$$

with the last expression corresponding to Eq. (18).

A.1. Gibbs sampler from structural operator SVD

From the SVD of the structural operator matrix given in Eq. (19), the covariance matrix for the sampling of $\boldsymbol{f}$ in Eq. (15) can be expressed as

$$\boldsymbol{\Sigma}_{\boldsymbol{f}} = \boldsymbol{U}\left[\frac{s_i^2}{s_i^2\tau_f + \tau_n}\right]\boldsymbol{U}^H, \tag{52}$$

the associated mean vector is given in Eq. (20) so that

$$[\boldsymbol{f}|\boldsymbol{D},\boldsymbol{y},\tau_f,\tau_n] \sim \mathcal{N}_c\left(\boldsymbol{f};\boldsymbol{U}\left[\frac{s_i}{1+\eta^2 s_i^2}\right]\boldsymbol{V}^H\boldsymbol{y},\boldsymbol{U}\left[\frac{s_i^2}{s_i^2\tau_f+\tau_n}\right]\boldsymbol{U}^H\right). \tag{53}$$

Sampling from a multivariate Gaussian distribution usually involves the Cholesky decomposition of the covariance matrice which may be expensive to compute if the dimension of the distribution is large. A new parametrization is proposed to reduce the sampling cost. Let us introduce the notations

$$\tilde{\boldsymbol{f}} = \boldsymbol{U}^H\boldsymbol{f}, \tilde{\boldsymbol{y}} = \boldsymbol{V}^H\boldsymbol{y}. \tag{54}$$

Knowing that a multivariate Gaussian pdf can be expressed as

$$\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu}_x,\boldsymbol{\Sigma}_x) = \boldsymbol{\mu}_{\boldsymbol{x}} + \boldsymbol{\Sigma}_x^{\frac{1}{2}}\mathcal{N}(\boldsymbol{x};\mathbf{0},\mathbf{I}), \tag{55}$$

it is then easier to sample from

$$\left[\tilde{\boldsymbol{f}}|\boldsymbol{D},\boldsymbol{y},\tau_f,\tau_n\right] \sim \left[\frac{s_i}{1+\eta^2 s_i^2}\right]\tilde{\boldsymbol{y}} + \left[\frac{s_i^2}{s_i^2\tau_f+\tau_n}\right]^{\frac{1}{2}}\mathcal{N}_c(\boldsymbol{x};\mathbf{0},\mathbf{I}), \tag{56}$$

equivalent to multiple univariate sampling from standard normal distribution whose numerical cost is marginal. The sample on $\boldsymbol{f}$ is then obtained simply by left-multiplying the sample on $\tilde{\boldsymbol{f}}$ by $\boldsymbol{U}$. It can also be noted that

$$\|\boldsymbol{f}\|_2^2 = \|\tilde{\boldsymbol{f}}\|_2^2, \tag{57}$$

$$\|\boldsymbol{D}^{-1}\boldsymbol{f} - \boldsymbol{y}\|_2^2 = \|\lceil s_i^{-1}\rfloor\tilde{\boldsymbol{f}} - \tilde{\boldsymbol{y}}\|_2^2. \tag{58}$$

From these equivalences and from Eqs. (35) and (37), both precision distributions can be expressed using the new parametrization,

$$[\tau_n|\boldsymbol{D},\boldsymbol{y},\boldsymbol{f},k_n,\theta_n] \propto \mathcal{G}\left(\tau_n;k_n+N,\left(\theta_n^{-1} + \|\lceil s_i^{-1}\rfloor\tilde{\boldsymbol{f}} - \tilde{\boldsymbol{y}}\|_2^2\right)^{-1}\right), \tag{59}$$

$$[\tau_f|\boldsymbol{f},k_f,\theta_f] \propto \mathcal{G}\left(\tau_f;k_f+N,\left(\theta_f^{-1} + \|\tilde{\boldsymbol{f}}\|_2^2\right)^{-1}\right). \tag{60}$$

A.2. Pseudo-code of the Gibbs sampler

Algorithm 1. Gibbs sampler.

1: % *Initialization:*
2: $k_f \leftarrow 0$ % *hyperparameters on* τ_f
3: $\theta_f \leftarrow +\infty$
4: $k_n \leftarrow 0$ % *hyperparameters on* τ_n
5: $\theta_n \leftarrow +\infty$
6: $N \leftarrow 3$ % *number of chains in parallel*
7: $R_n \leftarrow 2$ % *PSRF for* τ_n *(higher than 1) from* [40]
8: $R_f \leftarrow 2$ % *PSRF for* τ_f *(higher than 1)*
9: $R_t \leftarrow 1.01$ % *threshold for the PSRF*
10: $\eta^2 \leftarrow \text{median}\{\eta^2_{Lcurve}, \eta^2_{GCV}, \eta^2_{MMAP}\}$ % *initial regularization parameter*
11: $\tilde{\boldsymbol{y}} \leftarrow \boldsymbol{V}^H \boldsymbol{y}$% *projection of the measurements on the right singular vectors*
12: **for** $k = 1$ to N **do**
13: Initialize $\tilde{\boldsymbol{f}}^{(k)}$ from Eq. (20) and η^2
14: Initialize $\tau_f^{(k)}$ from $\|\tilde{\boldsymbol{f}}^{(k)}\|_2^{-2}$
15: Initialize $\tau_n^{(k)}$ from $\|\lceil s_i^{-1} \rfloor \tilde{\boldsymbol{f}}^{(k)} - \tilde{\boldsymbol{y}}\|_2^{-2} \times 10^{\frac{SNR(k)}{10}}$ %$SNR = [-15; 0; +15]$
16: **end for**
17: % *Sampling:*
18: **while** $R_n > R_t$ and $R_f > R_t$ **do**
19: **for** $k = 1$ to N **do**
20: $\eta^{2^{(k)}} \leftarrow \frac{\tau_f^{(k)}}{\tau_n^{(k)}}$
21: $\boldsymbol{\varepsilon}^{(k)} \leftarrow$ sample from $\mathcal{N}_c(\boldsymbol{\varepsilon}^{(k)}; \mathbf{0}, \mathbf{I})$
22: $\tilde{\boldsymbol{f}}^{(k)} \leftarrow \left\lceil \frac{s_i}{1+\eta^{2(k)} s_i^2} \right\rfloor \tilde{\boldsymbol{y}} + \left\lceil \frac{s_i^2}{s_i^2 \tau_f^{(k)} + \tau_n^{(k)}} \right\rfloor^{\frac{1}{2}} \boldsymbol{\varepsilon}^{(k)}$ % *from* Eq. (56)
23: $\tau_n^{(k)} \leftarrow$ sample from $\mathcal{G}\left(\tau_n; k_n + N, \left(\theta_n^{-1} + \|\lceil s_i^{-1} \rfloor \tilde{\boldsymbol{f}}^{(k)} - \tilde{\boldsymbol{y}}\|_2^2\right)^{-1}\right)$% *from* Eq. (59)
24: $\tau_f^{(k)} \leftarrow$ sample from $\mathcal{G}\left(\tau_f; k_f + N, \left(\theta_f^{-1} + \left\|\tilde{\boldsymbol{f}}^{(k)}\right\|_2^2\right)^{-1}\right)$ % *from* Eq. (60)
25: **end for**
26: $R_n \leftarrow$ compute from [40]
27: $R_f \leftarrow$ compute from [40]
28: **end while**

References

[1] D. Noiseux, Measurement of power flow in uniform beams and plates, J. Acoust. Soc. Am. 47 (1B) (1970) 238–247.
[2] G. Pavić, Measurement of structure borne wave intensity, part i: formulation of the methods, J. Sound Vib. 49 (2) (1976) 221–230.
[3] K.H.B. Janssens, P.P.G. Mas, P.A. Gajdatsy, H. Van Der Auweraer, L.J.P. Gielen, Transfer path analysis, uS Patent 8,731,868, May 20 2014.
[4] R.H. Lyon, R. Lyon, Statistical Energy Analysis of Dynamical Systems: Theory and Applications, MIT Press, Cambridge, 1975.
[5] B. Dobson, E. Rider, A review of the indirect calculation of excitation forces from measured structural response data, Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci. 204 (2) (1990) 69–75.
[6] Y. Zhang, J.A. Mann III, Examples of using structural intensity and the force distribution to study vibrating plates, J. Acoust. Soc. Am. 99 (1) (1996) 354–361.
[7] C. Pezerat, J. Guyader, Two inverse methods for localization of external sources exciting a beam, Acta Acust. 3 (1) (1995) 1–10.
[8] C. Pezerat, J. Guyader, Force analysis technique: reconstruction of force distribution on plates, Acta Acust. United Acust. 86 (2) (2000) 322–332.
[9] M. Djamaa, N. Ouelaa, C. Pezerat, J.-L. Guyader, Reconstruction of a distributed force applied on a thin cylindrical shell by an inverse method and spatial filtering, J. Sound Vib. 301 (3) (2007) 560–575.
[10] C. Renzi, C. Pezerat, J.-L. Guyader, Vibratory source identification by using the finite element model of a subdomain of a flexural beam, J. Sound Vib. 332 (3) (2013) 545–562.
[11] P.C. Hansen, Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank, SIAM J. Sci. Statist. Comput. 11 (3) (1990) 503–518.
[12] A.N. Tikhonov, V.Y. Arsenin, Solutions of Ill-Posed Problems, V.H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York, 1977.
[13] 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.
[14] P.C. Hansen, Analysis of discrete ill-posed problems by means of the l-curve, SIAM Rev. 34 (4) (1992) 561–580.
[15] W.M. Bolstad, Introduction to Bayesian Statistics, John Wiley & Sons, 2004.
[16] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, 2005.
[17] G. Koop, D.J. Poirier, J.L. Tobias, Bayesian Econometric Methods, Cambridge University Press, 2007.

[18] J.P. Huelsenbeck, F. Ronquist, et al, Mrbayes: Bayesian inference of phylogenetic trees, Bioinformatics 17 (8) (2001) 754–755.
[19] S. Jackman, Bayesian Analysis for the Social Sciences, vol. 846, John Wiley & Sons, 2009.
[20] J.M. Bardsley, Mcmc-based image reconstruction with uncertainty quantification, SIAM J. Sci. Comput. 34 (3) (2012) A1316–A1332.
[21] J. Antoni, A bayesian approach to sound source reconstruction: optimal basis, regularization, and focusing, J. Acoust. Soc. Am. 131 (4) (2012) 2873–2890.
[22] A. Pereira, J. Antoni, Q. Leclere, Empirical bayesian regularization of the inverse acoustic problem, Appl. Acoust. 97 (2015) 11–29.
[23] M. Aucejo, O. De Smet, Bayesian source identification using local priors, Mech. Syst. Signal Process. 66 (2016) 120–136.
[24] E. Zhang, J. Antoni, P. Feissel, Bayesian force reconstruction with an uncertain model, J. Sound Vib. 331 (4) (2012) 798–814.
[25] S. Brooks, A. Gelman, G. Jones, X.-L. Meng, Handbook of Markov Chain Monte Carlo, CRC Press, 2011.
[26] S. Geman, D. Geman, Stochastic relaxation, gibbs distributions, and the bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. (6) (1984) 721–741.
[27] G. Casella, E.I. George, Explaining the gibbs sampler, Am. Statist. 46 (3) (1992) 167–174.
[29] M. Aucejo, Structural source identification using a generalized tikhonov regularization, J. Sound Vib. 333 (22) (2014) 5693–5707.
[30] Q. Leclère, C. Pézerat, Vibration source identification using corrected finite difference schemes, J. Sound Vib. 331 (6) (2012) 1366–1377.
[31] J. Rice, Mathematical Statistics and Data Analysis, Nelson Education, 2006.
[32] W.M. Bolstad, Understanding Computational Bayesian Statistics, vol. 644, John Wiley & Sons, 2010.
[33] H. Raiffa, R. Schlaifer, Applied Statistical Decision Theory, Harvard University Press, 1974.
[34] J.-L. Guyader, Vibration in Continuous Media, John Wiley & Sons, 2013.
[35] P.C. Hansen, Regularization tools: a matlab package for analysis and solution of discrete ill-posed problems, Numer. Algorithms 6 (1) (1994) 1–35.
[36] R.E. Kass, B.P. Carlin, A. Gelman, R.M. Neal, Markov chain monte carlo in practice: a roundtable discussion, Am. Statist. 52 (2) (1998) 93–100.
[37] P. Heidelberger, P.D. Welch, Simulation run length control in the presence of an initial transient, Oper. Res. 31 (6) (1983) 1109–1144.
[38] J. Geweke, et al., Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments, vol. 196, Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA, 1991.
[39] A.E. Raftery, S. Lewis, et al, How many iterations in the gibbs sampler, Bayesian Statist. 4 (2) (1992) 763–773.
[40] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences, Statist. Sci. (1992) 457–472.
[41] Q. Leclere, F. Ablitzer, C. Pezerat, Identification of loads of thin structures with the corrected force analysis technique: an alternative to spatial filtering regularization, in: Proceedings of ISMA 2014.