An optimal Bayesian regularization for force reconstruction problems

Mathieu Aucejo, Olivier de Smet

An optimal Bayesian regularization for force reconstruction problems

M. Aucejo[a], O. De Smet[a]

[a]*Structural Mechanics and Coupled Systems Laboratory, Conservatoire National des Arts et Métiers, 2 Rue Conté, 75003 Paris, France*

Abstract

In a paper, recently published in *Mechanical Systems and Signal Processing*, we have proposed a full Bayesian inference for reconstructing mechanical sources acting on a linear and time invariant structure. The main interest of this approach is to propose an estimation of all the parameters of the model and quantify the posterior uncertainty associated to each parameter. Since all the necessary information about the problem is available, statistical measures, such as the mean, the median and the mode of the solution, can be easily estimated. In many practical situations, however, one only wants to determine the most probable parameters given the available data. Consequently, it is not relevant to implement a full Bayesian inference to only extract a point estimate. To overcome this potential issue, this paper introduces an optimal Bayesian regularization aiming at computing the Maximum a Posteriori estimate of the Bayesian formulation previously introduced by the authors. In doing so, the most probable parameters are obtained without heavy computations. The validity of the proposed method is assessed numerically and experimentally. In particular, obtained results highlight the

*Corresponding author. E-mail address: mathieu.aucejo@lecnam.net

ability of the proposed regularization strategy in computing solutions with a minimal amount of prior information on the sources to identify.

Keywords: Linear inverse problem, Force reconstruction, Bayesian regularization, Generalized Gaussian priors.

1. Introduction

In structural dynamics, force identification is a special class of inverse problem aiming at reconstructing the excitation sources acting on a mechanical structure from the knowledge of the vibration field measured over its surface. Unfortunately, this problem is known to be ill-posed, meaning that the existence of a unique stable solution is not guaranteed. A classical solution to bypass this difficulty is to apply regularization strategies, which allows constraining the space of solutions by using some prior information on the noise corrupting the data and the sources to identify. The most widespread and, certainly, the most popular technique is the Tikhonov regularization (a.k.a. ℓ_2-regularization) [1, 2, 3, 4, 5, 6, 7]. Although widely used, it is theoretically mainly applicable to the identification of rather smooth excitation fields or excitation signals [8], which is not a desirable effect when a localized source or an impulsive excitation signal have to be identified. To remedy this problem, LASSO regularization (a.k.a. ℓ_1-regularization) has been developed to promote the sparsity of the regularized solutions, while keeping the inverse problem convex. For this particular reason, it has focused many of the research efforts in the recent years [9, 10, 11, 12]. To unify all these approaches, Aucejo introduced the ℓ_q-regularization in the context of force reconstruction, which includes Tikhonov and LASSO regularization

as special cases [13]. However, because the norm parameter q can take any value in $\mathbb{R}^{+*}$, it can lead to more accurate solutions than those obtained from standard regularization strategies [13]. In the procedures described above, the a priori on the sources to identify is global, meaning that poor reconstructions can be obtained, if, for instance, a structure is excited by several sources having different spatial distributions, since the a priori has to reflect a compromise between contradictory distributions. That is why, Aucejo and De Smet extend the ℓ_q-regularization to take advantage of prior local information available on the sources to identify [14]. More specifically, the proposed formulation relies on the definition of several identification regions, in which the norm parameters q can be set independently for each region. Although they have proved their efficiency, all these techniques requires a proper tuning of the parameters of the considered formulation, namely the regularization parameter and, potentially, the norm parameters q. For the regularization parameter, this can be done using external, but expensive, automatic selection procedures [15, 16], while the choice of the norm parameters relies more on the user's skills or experience. Fortunately, because all these regularization strategies derive from the Bayesian statistics, a possible solution is to implement a full Bayesian inference to estimate the value of each parameters of the problem as well as the related posterior uncertainty using either MCMC techniques [17, 18, 19, 20, 21] or variational Bayes-like approaches [22, 23, 24, 25]. Since all the necessary information about the problem is then available, statistical measures, such as the mean, the median and the mode of the solution, can be easily estimated. In many practical situations, however, one only wants to determine the most probable param-

eters given the available data. In this respect, implementing a full Bayesian inference appears to be irrelevant, since it is generally time-consuming and computationally expensive. It results that it seems wiser and more efficient, in this case, to calculate directly the Maximum a Posteriori (MAP) estimate of the considered Bayesian formulation. This idea has given rise to the augmented Tikhonov regularization [26, 27, 28, 29], to the sparse Bayesian blind deconvolution [30] or to the adaptive ℓ_q-regularization [31].

The present paper introduces an optimal Bayesian regularization corresponding to the MAP estimate of the Bayesian formulation proposed by the authors in Ref. [20]. More specifically, the considered regularization strategy allows assessing within a single iterative process the most probable parameters of the problem given the measured vibration field. To avoid any misunderstanding, it is important to note that the proposed Bayesian regularization is optimal in the sense of the underlying Bayesian formulation. In other words, the optimality statement should not be regarded as a statement of superiority over the other existing strategies. To properly highlight the main features of the proposed approach, this article is divided into four parts. For the sake of completeness, the Bayesian formulation behind the proposed regularization strategy is first recalled in section 2 to make the paper self-contained. This section allows clarifying the assumptions made to derive the formulation and allows defining the main parameters of the corresponding probability distributions. Then, section 3 details the optimal Bayesian regularization by thoroughly introducing its mathematical formulation as well as the algorithm used to solve the problem. Finally, the ability of the pro-

posed approach in optimally identifying all the parameters of the problem is illustrated using synthetic and experimental data in sections 4 and 5, where locally sparse sources are reconstructed in the frequency domain. In particular, the proposed validations reveal that the optimal Bayesian regularization is an efficient tool to obtain consistent reconstructions, while avoiding the user to determine an accurate initial solution. In other words, only a rough knowledge of the excitation field to identify is required.

2. Bayesian formulation of the reconstruction problem

Let us consider the practical situation where the vibration field $\mathbf{X}$, measured over the surface of a structure, is caused by an unknown excitation field $\mathbf{F}$. If the structure is linear and time invariant, its dynamic behavior is completely determined by the transfer functions matrix $\mathbf{H}$, relating the vibration field $\mathbf{X}$ to the excitation field $\mathbf{F}$, so that:

$$\mathbf{X} = \mathbf{HF} + \mathbf{N}, \tag{1}$$

where $\mathbf{N}$ is the noise vector related to the measurement errors as well as modeling errors (provided the latter are small enough).

Because the force reconstruction problem consists in estimating the unknown excitation field $\mathbf{F}$ acting on a structure from the knowledge of the measured vibration field $\mathbf{X}$ and the transfer functions matrix $\mathbf{H}$, the noise vector $\mathbf{N}$ is another unknown of the reconstruction problem. To solve this problem efficiently, the Bayesian framework is adopted. Formally, the Bayesian paradigm consists in considering all the parameters of the problem as random

variables. Consequently, the uncertainty about each parameter is modeled by a probability distribution, describing the state of knowledge or the prior on this parameter. From a mathematical standpoint, the standard Bayesian formulation of the reconstruction problem relies on the Bayes' rule:

$$p(\mathbf{F}|\mathbf{X}) \propto p(\mathbf{X}|\mathbf{F})\, p(\mathbf{F}), \tag{2}$$

where:

- $p(\mathbf{F}|\mathbf{X})$ is the posterior probability distribution, representing the probability of observing $\mathbf{F}$ given a vibration field $\mathbf{X}$. In other words, it defines what it is known about the excitation field $\mathbf{F}$ after making vibration measurements;
- $p(\mathbf{X}|\mathbf{F})$ is the likelihood function, representing the probability of measuring $\mathbf{X}$ given an excitation field $\mathbf{F}$. It reflects the uncertainty related to the measurement of the vibration field $\mathbf{X}$;
- $p(\mathbf{F})$ is the prior probability distribution, representing our knowledge of the unknown excitation field $\mathbf{F}$ before measuring the vibration field $\mathbf{X}$.

Generally, the quality of the force reconstruction strongly depends on the choice of the likelihood function and the prior probability distribution. That is why, the probability distributions have to be carefully chosen.

As mentioned previously, the likelihood function reflects the uncertainty related to the measurement of the vibrations field $\mathbf{X}$. By definition, this uncertainty is mainly related to the noise vector $\mathbf{N}$. If the noise is supposed

spatially white and due to multiple independent causes, then the likelihood function can be represented by a complex multivariate normal distribution with zero mean and precision parameter τ_n, namely:

$$p(\mathbf{X}|\mathbf{F},\tau_n) = \left[\frac{\tau_n}{\pi}\right]^N \exp\left[-\tau_n \|\mathbf{X}-\mathbf{H}\mathbf{F}\|_2^2\right]. \tag{3}$$

where $\| \bullet \|_2$ is the ℓ_2-norm and N is the number of measurement points.

Regarding now the prior probability distribution, it reflects the uncertainty related to the unknown excitation field $\mathbf{F}$. If one supposes that the structure is excited in R different regions by uncorrelated excitations of various types (localized or distributed), then local excitation fields $\mathbf{F_r}$ can be considered as independent random vectors. As a result, the prior probability distribution can be written as the product of local prior probability distributions $p(\mathbf{F_r})$, that is:

$$p(\mathbf{F}) = \prod_{r=1}^{R} p(\mathbf{F_r}), \tag{4}$$

where $p(\mathbf{F_r})$ reflects the prior knowledge of the nature of the sources in the region r.

In the present formulation, it is assumed, for practical reasons, that the local excitation vectors $\mathbf{F_r}$ are real random vectors, whose components are supposed to be independent and identically distributed random variables following a generalized Gaussian distribution with zero mean [32]. As a consequence, each local excitation field follows a multivariate generalized Gaussian distribution with zero mean. From the mathematical standpoint,

the local prior probability distributions are thus written:

$$p(\mathbf{F_r}|\tau_{sr}, q_r) = \left[\frac{q_r}{2\,\Gamma(1/q_r)}\right]^{M_r} \tau_{sr}^{\frac{M_r}{q_r}} \exp\left[-\tau_{sr}\|\mathbf{F_r}\|_{q_r}^{q_r}\right], \tag{5}$$

where:

- q_r is the shape parameter of the distribution in the region r. Its value is defined in the interval $]0, +\infty[$;
- $\| \bullet \|_{q_r}$ is the $\ell_{\mathrm{q_r}}$-norm or quasi-norm, if $q_r \geq 1$ and $q_r < 1$ respectively;
- τ_{sr} is the scale parameter of the distribution, which can be viewed as a generalized measure of the precision of the distribution;
- M_r is the number of reconstruction points in the region r;
- $\Gamma(x) = \int_0^{+\infty} t^{x-1}\, e^{-t}\, dt$ is the gamma function.

It should be noted that the choice of a multivariate generalized Gaussian distribution offers a high flexibility for describing prior knowledge of the sources to identify, since it is allows defining either a sparse (or localized) prior for $q_r \leq 1$ or a smooth (or distributed) prior for $q_r = 2$.

From the explanations given above, the Bayesian formulation of the reconstruction problem finally writes:

$$p(\mathbf{F}|\mathbf{X}, \tau_n, \tau_{sr}, q_r) \propto p(\mathbf{X}|\mathbf{F}, \tau_n) \prod_{r=1}^{R} p(\mathbf{F_r}|\tau_{sr}, q_r). \tag{6}$$

At this stage, it is clear that the main shortcoming of the standard Bayesian formulation is the choice of the associated parameters, since the

quality of the identification is conditioned to the knowledge of the shape and precision parameters. If their values are poorly chosen, then the resulting reconstruction won't be representative of the actual target excitation field. To alleviate this limitation, a possible alternative is to infer the shape and precision parameters by also considering them as independent random variables. Such an idea gives rise to the complete Bayesian formulation at the root of the proposed optimal Bayesian regularization [20]. From a mathematical standpoint, this particular Bayesian formulation of the reconstruction problem is given by:

$$p(\mathbf{F}, \tau_n, \tau_{sr}, q_r|\mathbf{X}) \propto p(\mathbf{X}|\mathbf{F}, \tau_n)\, p(\tau_n) \prod_{r=1}^{R} p(\mathbf{F_r}|\tau_{sr}, q_r)\, p(\tau_{sr})\, p(q_r), \quad (7)$$

where $p(q_r)$ is the prior probability distribution of the shape parameters q_r, while $p(\tau_n)$ and $p(\tau_{sr})$ are the prior probability distributions of the precision parameters τ_n and τ_{sr} respectively.

To finalize the definition of the complete Bayesian formulation, it remains to specify the prior probability distributions of the shape and precision parameters. The choice of the priori probability distributions $p(\tau_n)$ and $p(\tau_{sr})$ is first limited to distribution having a strictly positive support, because the precision parameters τ_n and τ_{sr} are real positive numbers. The common choice, made in the literature to use conjugacy properties, is the Gamma distribution [33]. Practically, the Gamma distribution is defined by:

$$\mathcal{G}(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1} \exp(-\beta\, x) \quad \text{with} \quad \alpha > 0,\ \beta > 0, \quad (8)$$

where α and β are respectively the scale parameter and the rate parameter

of the distribution.

The definition of the prior probability distribution of the shape parameters q_r relies on the fact that they are positive and bounded. In absence of more precise knowledge on the parameters, the probability distribution is not only chosen to reflect the available information but also for its mathematical tractability. A probability distribution that meets these requirements is the truncated Gamma distribution defined by:

$$\mathcal{G}_T(x|\alpha,\beta,l_b,u_b) = \frac{\Gamma(\alpha)}{\gamma(\alpha,\beta u_b) - \gamma(\alpha,\beta l_b)}\,\mathcal{G}(x|\alpha,\beta)\,\mathbb{I}_{[l_b,u_b]}(x), \tag{9}$$

where:

- $\mathcal{G}(x|\alpha,\beta)$ is the Gamma distribution defined in Eq. (8);
- $\mathbb{I}_{[l_b,u_b]}(x)$ is the truncation function defined between the lower bound l_b and the upper bound u_b. More precisely, this function simply writes:

$$\mathbb{I}_{[l_b,u_b]}(x) = \begin{cases} 1 & \text{if } x \in [l_b,u_b] \\ 0 & \text{otherwise} \end{cases}; \tag{10}$$

- $\gamma(s,x) = \int_0^x t^{s-1}\exp(-t)\,dt$ is the lower incomplete Gamma function.

From the above considerations, it results that the proposed complete Bayesian formulation is finally expressed as follows:

$$\begin{aligned} p(\mathbf{F},\tau_n,\tau_{sr},q_r|\mathbf{X}) \propto\; & p(\mathbf{X}|\mathbf{F},\tau_n)\,p(\tau_n|\alpha_n,\beta_n) \\ & \times \prod_{r=1}^{R} p(\mathbf{F_r}|\tau_{sr},q_r)\,p(\tau_{sr}|\alpha_{sr},\beta_{sr})\,p(q_r|\alpha_r,\beta_r,l_b,u_b), \end{aligned} \tag{11}$$

where (α_r, β_r) are the hyperparameters related to the shape parameters q_r, while (α_n, β_n) and $(\alpha_{sr}, \beta_{sr})$ are the hyperparameters associated to the precision parameters τ_n and τ_{sr} respectively.

As a final remark, it is worth mentioning that the previous Bayesian formulation allows emphasizing the specificity of the proposed optimal Bayesian regularization with respect to those at the roots of the Augmented Tikhonov regularization and the adaptive ℓ_q-regularization used in the context of force reconstruction. First, the proposed Bayesian regularization generalizes to some extent the Augmented Tihkonov regularization by considering the shape parameters q_r as random variables and not as deterministic parameters. Second, it extends the adaptive ℓ_q-regularization, which is based on the formulation proposed by the authors in Ref. [19], by defining several identification regions and imposing, if necessary, a greater sparsity of local excitation fields by authorizing the shape parameters to be less than 1.

3. Optimal Bayesian regularization

This section aims at introducing thoroughly the proposed optimal Bayesian regularization. To this end, its practical implementation is detailed after having introduced the underlying theoretical formulation. In the present work, the identification regions are supposed to be known a priori. In general, a careful analysis of the mechanical system is sufficient to determine the identification regions if the studied system or the operating conditions are rather simple. In case of more complex systems or operating conditions, the definition of the identification regions could be automatically done using machine

learning strategies such as Relevance Vector Machine [34], but this is outside the scope of this paper.

3.1. Theoretical formulation

The basic idea of the optimal Bayesian regularization is to determine the MAP estimate of the complete Bayesian formulation given by Eq. (11), without performing a full Bayesian inference. Consequently, the solution of the source identification problem is sought such that:

$$\left(\widehat{\mathbf{F}}, \widehat{\tau}_n, \widehat{\tau}_{sr}, \widehat{q}_r\right) = \underset{(\mathbf{F}, \tau_n, \tau_{sr}, q_r)}{\operatorname{argmax}} \; p(\mathbf{F}, \tau_n, \tau_{sr}, q_r | \mathbf{X}). \tag{12}$$

The solution of the previous optimization problem can be classically solved by applying the first-order optimality condition to the dual minimization problem. However, to render the presentation of the optimal Bayesian regularization more didactic, one can notice that an alternative way of solving this optimization problem consists in maximizing the full conditional probability distributions associated to each parameters. Consequently, the MAP estimate of the complete Bayesian formulation can be found by solving the following set of optimization problems:

$$\widehat{q}_r = \underset{q_r}{\operatorname{argmax}} \; p(q_r | \mathbf{X}, \mathbf{F}, \tau_n, \tau_{sr}), \tag{13a}$$

$$\widehat{\tau}_{sr} = \underset{\tau_{sr}}{\operatorname{argmax}} \; p(\tau_{sr} | \mathbf{X}, \mathbf{F}, \tau_n, q_r), \tag{13b}$$

$$\widehat{\tau}_n = \underset{\tau_n}{\operatorname{argmax}} \; p(\tau_n | \mathbf{X}, \mathbf{F}, \tau_{sr}, q_r), \tag{13c}$$

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmax}} \; p(\mathbf{F} | \mathbf{X}, \tau_n, \tau_{sr}, q_r). \tag{13d}$$

Practically, it is generally easier to solve the dual minimization problem, which consists in finding the value of the parameter minimizing the opposite of the logarithm of the corresponding full conditional probability distribution. For the complete Bayesian formulation, the full conditional probability distributions are for:

- the shape parameters q_r:

$$p(q_r|\mathbf{X},\mathbf{F},\tau_n,\tau_{sr}) \propto \frac{\tau_{sr}^{\frac{M_r}{q_r}}}{\Gamma(1/q_r)^{M_r}}\, q_r^{\alpha_r+M_r-1} \exp\left[-\beta_r\, q_r - \tau_{sr}\|\mathbf{F_r}\|_{q_r}^{q_r}\right] \mathbb{I}_{[l_b,u_b]}(q); \tag{14}$$

- the precision parameters τ_{sr}:

$$p(\tau_{sr}|\mathbf{X},\mathbf{F},\tau_n,q_r) \propto \mathcal{G}\left(\tau_{sr}\middle|\, \alpha_{sr}+\frac{M_r}{q_r},\, \beta_{sr}+\|\mathbf{F_r}\|_{q_r}^{q_r}\right); \tag{15}$$

- the precision parameters τ_n:

$$p(\tau_n|\mathbf{X},\mathbf{F},\tau_{sr},q_r) \propto \mathcal{G}\left(\tau_n\middle|\, \alpha_n+N,\, \beta_n+\|\mathbf{X}-\mathbf{HF}\|_2^2\right); \tag{16}$$

- the force vector $\mathbf{F}$:

$$p(\mathbf{F}|\mathbf{X},\tau_n,\tau_{sr},q_r) \propto \exp\left[-\tau_n\|\mathbf{X}-\mathbf{HF}\|_2^2 - \sum_{r=1}^{R}\tau_{sr}\|\mathbf{F_r}\|_{q_r}^{q_r}\right]. \tag{17}$$

Consequently, finding the MAP estimate is equivalent to solve the following set of minimization problems, defining the optimal Bayesian regularization[1]:

$$\widehat{q_r} = \underset{q_r}{\operatorname{argmin}}\, f(q_r|\tau_{sr},\mathbf{F_r}) \;\text{ for } q_r \in [l_b,u_b], \tag{18a}$$

[1]By setting $\alpha_r = \alpha_{sr} = \alpha_n = 1$ and $\beta_r = \beta_{sr} = \beta_n = 0$, the resulting optimal Bayesian regularization is equivalent to that obtained assuming that the prior probability distributions on the shape and precision parameters are uniform.

$$\widehat{\tau}_{sr} = \underset{\tau_{sr}}{\operatorname{argmin}}\ \tau_{sr}\left(\beta_{sr} + \|\mathbf{F_r}\|_{q_r}^{q_r}\right) - \left(\alpha_{sr} + \frac{M_r}{q_r} - 1\right)\log\tau_{sr}, \tag{18b}$$

$$\widehat{\tau}_{n} = \underset{\tau_{n}}{\operatorname{argmin}}\ \tau_{n}\left(\beta_{n} + \|\mathbf{X} - \mathbf{HF}\|_2^2\right) - (\alpha_n + N - 1)\log\tau_n, \tag{18c}$$

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}}\ \|\mathbf{X} - \mathbf{HF}\|_2^2 + \sum_{r=1}^{R}\lambda_r\|\mathbf{F_r}\|_{q_r}^{q_r}. \tag{18d}$$

where $f(q_r|\tau_{sr}, \mathbf{F_r}) = M_r\log\Gamma(1/q_r) - \frac{M_r}{q_r}\log\tau_{sr} - [\alpha_r + M_r - 1]\log q_r + \beta_r\, q_r + \tau_{sr}\|\mathbf{F_r}\|_{q_r}^{q_r}$ and $\lambda_r = \tau_{sr}/\tau_n$.

3.2. Resolution algorithm

The analysis of the set of equations (18) suggests that the optimal Bayesian regularization can only be solved in an iterative manner, since the optimal value of a given parameter explicitly depends on the value of the others. It results that the resolution process is divided into three main steps:

1. Set $k = 0$ and initialize $\widehat{q}_r^{(0)}$, $\widehat{\tau}_{sr}^{(0)}$, $\widehat{\tau}_n^{(0)}$ and $\widehat{\mathbf{F}}^{(0)}$;
2. **while** convergence is not reached
   a. $k = k + 1$;
   b. for each region r, compute $\widehat{q}_r^{(k)}$ from Eq. (18a) given $\widehat{\tau}_{sr}^{(k-1)}$ and $\widehat{\mathbf{F}}_{\mathbf{r}}^{(k-1)}$;
   c. for each region r, compute $\widehat{\tau}_{sr}^{(k)}$ from Eq. (18b) given $\widehat{\mathbf{F}}_{\mathbf{r}}^{(k-1)}$ and $\widehat{q}_r^{(k)}$;
   d. compute $\widehat{\tau}_n^{(k)}$ from Eq. (18c) given $\widehat{\mathbf{F}}^{(k-1)}$;
   e. compute $\widehat{\mathbf{F}}^{(k)}$ from Eq. (18d) given $\widehat{\tau}_n^{(k)}$, $\widehat{\tau}_{sr}^{(k)}$ and $\widehat{q}_r^{(k)}$;

f. monitor the convergence

end while

3. **return** $\widehat{\mathbf{F}}$, $\widehat{\tau}_n$, $\widehat{\tau}_{sr}$ and $\widehat{q}_r$.

The practical implementation of the previous resolution procedure requires a particular attention, since each step has to be carefully and properly defined to converge to the optimal solution. For this particular reason, implementation issues are detailed in the next sections.

3.2.1. Initialization of the resolution algorithm

The initialization step can be one of the keys of the convergence of the iterative procedure when the functional to minimize exhibits local minima. To properly address this issue, it has been decided to initialize the procedure from a starting point having a reasonably high probability. In the present case, the initial force vector $\widehat{\mathbf{F}}^{(0)}$ is obtained from the MAP estimate of the standard Bayesian formulation, which is defined by [see Ref. [14] for more details]:

$$\begin{aligned}\widehat{\mathbf{F}}^{(0)} &= \underset{\mathbf{F}}{\operatorname{argmax}}\, p\left(\mathbf{F}|\mathbf{X}, \widehat{\tau}_n^{(0)}, \widehat{\tau}_{sr}^{(0)}, \widehat{q}_r^{(0)}\right) \\ &= \underset{\mathbf{F}}{\operatorname{argmin}}\, \|\mathbf{X}-\mathbf{H}\mathbf{F}\|_2^2 + \sum_{r=1}^{R} \lambda_r^{(0)} \|\mathbf{F}_\mathbf{r}\|_{\widehat{q}_r^{(0)}}^{\widehat{q}_r^{(0)}},\end{aligned} \tag{19}$$

where $\lambda_r^{(0)} = \widehat{\tau}_{sr}^{(0)}/\widehat{\tau}_n^{(0)}$ is the regularization parameter in the zone r.

To simplify somewhat the calculation, it is supposed that the precision parameters $\widehat{\tau}_{sr}^{(0)}$ are equal to a unique constant value $\widehat{\tau}_s^{(0)}$. Accordingly, Eq. (19) becomes:

$$\widehat{\mathbf{F}}^{(0)} = \underset{\mathbf{F}}{\operatorname{argmin}}\, \|\mathbf{X}-\mathbf{H}\mathbf{F}\|_2^2 + \lambda^{(0)} \sum_{r=1}^{R} \|\mathbf{F}_\mathbf{r}\|_{\widehat{q}_r^{(0)}}^{\widehat{q}_r^{(0)}}, \tag{20}$$

where $\lambda^{(0)} = \widehat{\tau}_s^{(0)}/\widehat{\tau}_n^{(0)}$.

To obtain a relevant initial force vector, it is necessary to determine reasonable values of the initial shape parameters $\widehat{q}_r^{(0)}$ and the precision parameters $\widehat{\tau}_s^{(0)}$ and $\widehat{\tau}_n^{(0)}$. Practically, the values of the shape parameters $\widehat{q}_r^{(0)}$ lie in the interval $]0,\,2]$ for force reconstruction problems. They can be chosen from subjective considerations by recalling that smoothed solutions are promoted for $\widehat{q}_r^{(0)} = 2$, while sparse solutions are promoted for $\widehat{q}_r^{(0)} \leq 1$ [8, 35]. On the contrary, near-optimal values of the precision parameters are difficult to assess a priori, i.e. without any calculation, because they are strongly related to the optimization problem given in Eq. (20) through the regularization parameter $\lambda^{(0)}$, whose optimal value is partly conditioned to $\widehat{q}_r^{(0)}$. In addition, since the shape parameter $\widehat{q}_r^{(0)}$ can take any value in the range]0,2], the solution of the optimization problem has generally no closed-form expression. Incidentally, the optimal value of the regularization parameter $\lambda^{(0)}$ can not be directly estimated from the marginalized MAP [36], the L-curve principle [37] or the Generalized Cross Validation [16] before solving the minimization problem [see Ref. [35] for details]. To this end, the optimization problem is solved iteratively using an adapted Iteratively Reweighted (IR) algorithm, which allows determining in an iterative manner the regularized force vector as well as the optimal regularization parameter associated to the minimization problem.

The proposed IR algorithm is derived by applying the first-order optimality condition to Eq. (20). It results from what precedes that the solution at

iteration j of the iterative process is given by:

$$\widehat{\mathbf{F}}^{(0,j)} = \left(\mathbf{H}^H\mathbf{H} + \lambda^{(0,j)}\,\mathbf{W}^{(j)}\right)^{-1}\mathbf{H}^H\mathbf{X}, \tag{21}$$

where the regularization parameter $\lambda^{(0,j)}$ is updated from an automatic selection procedure such as the L-curve principle [37], while $\mathbf{W}^{(j)}$ is a diagonal global weighting matrix depending explicitly on $\widehat{\mathbf{F}}_{\mathbf{r}}^{(0,j-1)}$ and $\widehat{q}_r^{(0)}$ and defined from local weighting matrices $\mathbf{W_r}$ such that:

$$\mathbf{W}^{(j)} = \mathrm{diag}[\mathbf{W}_{\mathbf{1}}^{(j)}, \ldots, \mathbf{W}_{\mathbf{r}}^{(j)}, \ldots, \mathbf{W}_{\mathbf{R}}^{(j)}]. \tag{22}$$

In the previous equation, each local weighting matrix $\mathbf{W}_{\mathbf{r}}^{(j)}$ is a diagonal matrix, expressed as:

$$\mathbf{W}_{\mathbf{r}}^{(j)} = \mathrm{diag}\left[w_{r,1}^{(j)}, \ldots, w_{r,i}^{(j)}, \ldots, w_{r,M_r}^{(j)}\right], \tag{23}$$

with:

$$w_{r,i}^{(j)} = \frac{\widehat{q}_r^{(0)}}{2}\max\left(\epsilon_r,\, \left|f_{ri}^{(j-1)}\right|\right)^{\widehat{q}_r^{(0)}-2}, \tag{24}$$

where $f_{ri}^{(j-1)}$ is the i$^{\text{th}}$ component of the vector $\widehat{\mathbf{F}}_{\mathbf{r}}^{(j-1)}$ and ϵ_r is a small real positive number acting as a damping parameter. It allows avoiding infinite weights when $\left|f_{ri}^{(j-1)}\right| \to 0$ and $\widehat{q}_r^{(0)} < 2$. Practically, the damping parameter is automatically selected from the cumulative histogram of $\left|\widehat{\mathbf{F}}_{\mathbf{r}}^{(0)}\right|$. More precisely, its value is chosen so that 5% of the values of $\left|\widehat{\mathbf{F}}_{\mathbf{r}}^{(0)}\right|$ are less than or equal to ϵ_r [13, 38].

After convergence of the iterative process, one obtains the optimal force vector $\widehat{\mathbf{F}}^{(0)}$, the global weighting matrix $\mathbf{W}$, as well as the optimal value of the regularization parameter $\lambda^{(0)}$.

At this stage of the initialization procedure, it remains to determine the values of $\tau_n^{(0)}$ and $\tau_s^{(0)}$. For this purpose, we follow the approach proposed by Pereira et al. [36] consisting in finding the the most probable values of $\tau_n^{(0)}$ and $\tau_s^{(0)}$ given the measured vibration field $\mathbf{X}$. The application of this idea to the proposed Bayesian formulation leads to the following values for $\widehat{\tau}_s^{(0)}$ and $\widehat{\tau}_n^{(0)}$ [see Ref. [20] for details]:

$$\widehat{\tau}_s^{(0)} = \frac{N}{\mathbf{X}^H \left(\lambda^{(0)}\mathbf{I} + \mathbf{H}\mathbf{W}^{-1}\mathbf{H}^H\right)^{-1}\mathbf{X}} \quad \text{and} \quad \widehat{\tau}_n^{(0)} = \frac{\tau_s^{(0)}}{\lambda^{(0)}}. \tag{25}$$

As a final remark, it is worth noting that the proposed initialization procedure implies that only the initial shape parameters $\widehat{q}_r^{(0)}$ must be defined by the user, since the other parameters are computer from Eqs. (20) and (25).

3.2.2. Computation of the optimal shape parameters $\widehat{q}_r^{(k)}$

The computation of $\widehat{q}_r^{(k)}$ requires a special attention, since the function to minimize has no analytical solution. However, because the function to minimize, $f(q_r|\tau_{sr}, \mathbf{F_r})$, is known analytically [see Eq. (18)], it is possible to calculate its gradient with respect to q_r. In the present case, the gradient, $g(q_r|\tau_{sr}, \mathbf{F_r})$, is expressed as:

$$\begin{aligned} g(q_r|\tau_{sr}, \mathbf{F_r}) = \beta_r + \tau_{sr} \sum_{i=1}^{M_r} \left(|F_{ri}|^{q_r} \log|F_{ri}|\right) - \frac{\alpha_r + M_r - 1}{q_r} \\ + \frac{M_r}{q_r^2}\left[\log \tau_{sr} - \psi(1/q_r)\right], \end{aligned} \tag{26}$$

where F_{ri} is the i$^{\text{th}}$ component of the force vector $\mathbf{F_r}$ and $\psi(x)$ is the digamma function.

The knowledge of the function and its gradient allows considering several possibilities for finding the optimal values of the shape parameters q_r at iteration k within the interval $[l_b, u_b]$. Indeed, the minimization problem defined by Eq. (18a) can be solved using either a bound-constrained minimization method [39, 40] or a bounded root-finding algorithm applied to the gradient [41] or even a brute-force approach.

It should be noted here that, even if the lower and upper bounds l_b and u_b can theoretically take any positive value, the values of the shape parameters q_r generally lies in the interval $]0, 2]$ [see section 3.2.1]. For this particular reason, we set $l_b = 0.05$ and $u_b = 2$ for all the validations presented in the rest of the paper.

Finally, to complete this section, a comment must be added on the choice of the hyperparameters α_r and β_r. As evoked previously, the choice of the truncated Gamma distribution has been made for mathematical convenience, because other continuous truncated distribution could have theoretically been used. To avoid biasing the optimization process, the shape of the prior distribution needs to be weakly informative [42]. According to Gelman, a prior distribution is said weakly informative if it is proper but is set up so that the information it does provide is intentionally weaker than whatever actual prior knowledge is available [43]. Here, this leads us to set the hyperparameters such that $\alpha_r = 1$ and $\beta_r \to 0$.

3.2.3. Computation of the optimal precision parameters $\widehat{\tau}_n^{(k)}$ and $\widehat{\tau}_{sr}^{(k)}$

The computation of the precision parameters $\widehat{\tau}_n^{(k)}$ and $\widehat{\tau}_{sr}^{(k)}$ is straightforward, insofar as the solutions of minimization problems given by Eqs. (18b) and (18c) can be analytically calculated by applying the first-order optimality condition. In doing so, it readily comes that:

$$\widehat{\tau}_{sr}^{(k)} = \frac{\alpha_{sr} + \dfrac{M_r}{\widehat{q}_r(k)} - 1}{\beta_{sr} + \left\| \widehat{\mathbf{F}}_{\mathbf{r}}^{(k-1)} \right\|_{\widehat{q}_r^{(k)}}^{\widehat{q}_r^{(k)}}} \quad \text{and} \quad \widehat{\tau}_n^{(k)} = \frac{\alpha_n + N - 1}{\beta_n + \left\| \mathbf{X} - \mathbf{H}\widehat{\mathbf{F}}^{(k-1)} \right\|_2^2}. \tag{27}$$

As for the computation of the shape parameters, the choice of the hyperparameters $(\alpha_{sr}, \beta_{sr})$ and (α_n, β_n) is questionable, since the Gamma distribution has been chosen for mathematical convenience and does not reflect any real prior information on the precision parameters, except their positiveness. That is why, the prior distributions on τ_n and τ_{sr} should be as minimally informative as possible. To this end, one sets, as previously, $\alpha_n = \alpha_{sr} = 1$ and $\beta_n = \beta_{sr} \to 0$.

3.2.4. Computation of the optimal force vector $\widehat{\boldsymbol{F}}^{(k)}$

By definition, the force vector at iteration k is solution of the following minimization problem:

$$\widehat{\mathbf{F}}^{(k)} = \underset{\mathbf{F}}{\operatorname{argmin}} \, \|\mathbf{X} - \mathbf{H}\mathbf{F}\|_2^2 + \sum_{r=1}^{R} \lambda_r^{(k)} \|\mathbf{F}_{\mathbf{r}}\|_{\widehat{q}_r^{(k)}}^{\widehat{q}_r^{(k)}}, \tag{28}$$

where $\lambda_r^{(k)} = \widehat{\tau}_{sr}^{(k)} / \widehat{\tau}_n^{(k)}$ is the regularization parameter at iteration k.

Because this minimization problem has the same form as the one used to compute the initial force vector $\widehat{\mathbf{F}}^{(0)}$, the IR algorithm, briefly described in

section 3.2.1, is also implemented to compute the optimal force vector $\widehat{\mathbf{F}}^{(k)}$ at iteration k. In this case, however, all the regularization parameters $\lambda_r^{(k)}$ are considered during the resolution process [20].

3.2.5. Convergence monitoring

As any iterative procedure, the proposed algorithm must be stopped after either a certain stopping criterion is less than some tolerance or a certain number of iterations fixed by the user is reached. In the present paper, the stopping criterion is related to the relative error of the force vector between two successive iterations. Mathematically, the relative error δ is defined such that:

$$\delta\left(\widehat{\mathbf{F}}^{(k-1)}, \widehat{\mathbf{F}}^{(k)}\right) = \frac{\left\|\widehat{\mathbf{F}}^{(k)} - \widehat{\mathbf{F}}^{(k-1)}\right\|_1}{\left\|\widehat{\mathbf{F}}^{(k-1)}\right\|_1}. \tag{29}$$

The choice of the tolerance has a strong influence on the overall performance of the optimal Bayesian regularization, since a too small tolerance can lead to unnecessary calculations, while a too large tolerance can lead to inaccurate solutions. This particular point will be thoroughly studied in the next section.

4. Numerical validation

The present numerical study intends to demonstrate how the proposed optimal Bayesian regularization (OBR) operates with respect to the choice of the initial solution, the tolerance of the iterative process and the measurement noise level. For the sake of completeness, the proposed optimal Bayesian formulation is also applied for different prior structures and comparisons with

results obtained from the corresponding full Bayesian inference are proposed. Finally, the possibility of a parameters reduction is studied.

4.1. Description of the test case

In the present numerical validation, one seeks to identify a point force of unit amplitude (i.e. $F_0 = 1$ N) acting on a thin simply supported steel plate with dimensions 0.6 m $\times$ 0.4 m $\times$ 0.005 m. The coordinates of the point force, measured from the lower left corner of the plate, are $(x_0, y_0) = (0.42$ m, 0.25 m). Practically, this configuration allows studying the influence of the definition of local regularization terms, since the present excitation field exhibits two types of spatial distribution over the structure, namely a smooth distribution of the reaction forces at boundaries and a singular distribution around the location of the point force.

To properly simulate experimental measurements, the exact vibration displacement field $\mathbf{X}_{\mathbf{exact}}$ is first computed from a FE mesh of the plate made up with 187 shell elements, assuming that only bending motions are measurable. Then, in absence of contradictory information, the exact displacement field is supposed to be corrupted by an additive Gaussian white noise with a signal-to-noise ratio (SNR) equal to 35 dB.

Finally, the transfer functions matrix $\mathbf{H}$ is computed from a FE model of the plate with free boundary conditions, assuming that only bending motions are measured. In other words, the computed transfer functions matrix $\mathbf{H}$ is dynamically condensed over the measurable degrees of freedom only [13, 44]. The main interest in using free boundary conditions to model the dynamic

behavior of the plate is to allow the identification of external excitations acting on the structure as well as reaction forces at boundaries [45].

4.2. Application

To numerically validate any force reconstruction strategy, it is first necessary to define the reference force vector $\mathbf{F_{ref}}$ that could serve as a proper benchmark. This reference force vector is computed from the transfer functions matrix $\mathbf{H}$ and the exact displacement field $\mathbf{X_{exact}}$ thanks to the following relation:

$$\mathbf{F_{ref}} = \mathbf{H}^{-1}\mathbf{X_{exact}}. \tag{30}$$

This numerical study is focused on the identification of the excitation field at 350 Hz, i.e. outside the resonance frequencies of the plate. As shown in Fig. 1, the reference force vector corresponds to the description of the test case given in the previous section, since it exhibits smooth reaction forces at boundaries of the plate as well as a unit point force F_0 at $(x_0, y_0) = (0.42$ m, 0.25 m).

Consequently, the analysis of the reference force vector suggests that two identification regions can be defined to apply the OBR: (i) a central region associated to the shape parameter q_1 and containing the point force only and (ii) a region associated to the shape parameter q_2 and corresponding to the boundaries of the plate [see Fig. 2].

Regarding the application of the OBR to the previous synthesized data, it should be noted that the tolerance for stopping the iterative is set to 10^{-4}

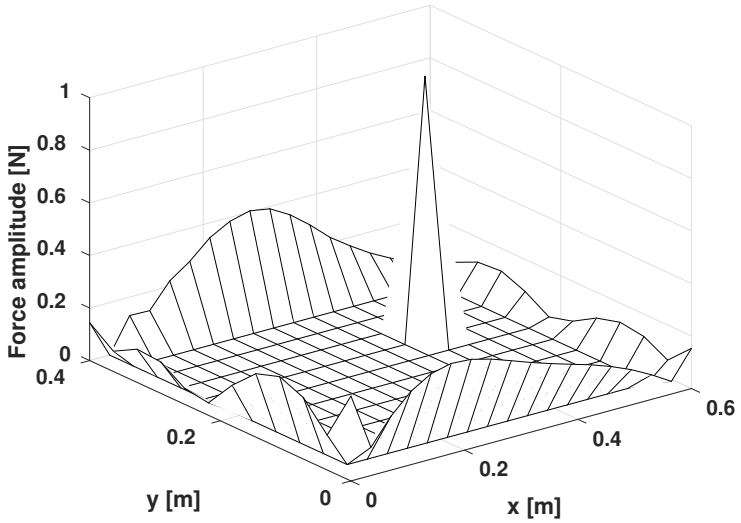


Figure 1: Reference force vector $\mathbf{F_{ref}}$ at 350 Hz

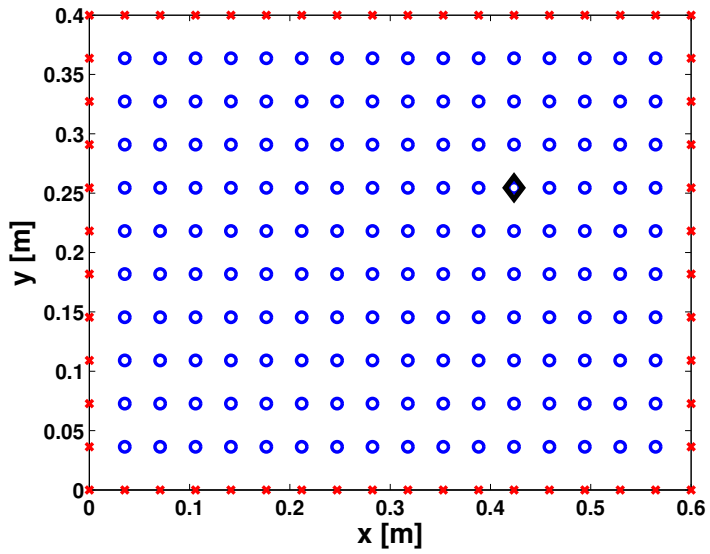


Figure 2: Definition of the identification regions - (∘) region 1 (Point force), (×) region 2 (Reaction forces) and (◊) location of the point force

if no contradictory information is given. Finally, to quantify the accuracy of identified solutions with respect to the considered operating conditions, the relative error (RE) and the peak error (PE) are evaluated. Formally, the relative error δ is defined in Eq. (29) and corresponds, in the present situation, to the relative error between the reference excitation field $\mathbf{F_{ref}}$ and the excitation field $\widehat{\mathbf{F}}$ reconstructed from the OBR. As a result, the relative error must be written $\delta(\mathbf{F_{ref}}, \widehat{\mathbf{F}})$. Practically, the relative error is a global indicator of the reconstruction quality. The peak error is a local indicator that allows assessing the reconstruction quality of the point force amplitude. Mathematically, it is defined such that:

$$\delta_p = \frac{\widehat{F}_p - F_p^{\text{ref}}}{F_p^{\text{ref}}}, \tag{31}$$

where F_p^{ref} is the point force amplitude associated to the reference force vector $\mathbf{F_{ref}}$, while $\widehat{F}_p$ is the point force amplitude associated to the identified solution $\widehat{\mathbf{F}}$.

4.2.1. Influence of the choice of the initial parameters

In section 3.2.1, it has been noted that only the definition of the shape parameters $\widehat{q}_r^{(0)}$ is required to initialize the OBR, since all the other parameters are computed accordingly. In this context, it is interesting to assess the convergence of the OBR with respect to the initial choice of the shape parameters. To this end, it is supposed in the following that $\widehat{q}_1^{(0)} = \widehat{q}_2^{(0)} = \widehat{q}_0$ with $\widehat{q}_0 = \{0.5, 1, 2\}$.

Fig. 3 presents the excitation fields reconstructed from the OBR for the three $\widehat{q}^{(0)}$ previously defined. At first sight, the identified excitation fields

are in a relatively good agreement with the reference one. This qualitative observation is confirmed by the results listed in Table 1. In particular, it should be noted that the values of the identified parameters are very close to each other. It is also worth noting that the point force is properly recovered in all cases. It should however be noted that the overall reconstruction error is about 41%, indicating that the reconstruction of the reaction forces is not as good as expected. This confirms the visual inspection of Fig. 3. Finally, the only noticeable difference appears in the number of iterations performed by the algorithm to reach the convergence.

The main conclusion of this convergence study is that the OBR is almost insensitive to the choice of the initial shape parameter $\widehat{q}_0$.

Table 1: Convergence study of the OBR with respect to the choice of the initial shape parameters – N_{it}: Number of iterations of the algorithm

	Initial shape parameter $\widehat{q}_0$		
	0.5	1	2
q_1	0.51	0.51	0.51
q_2	2	2	2
τ_{s1}	32.02	31.07	31.54
τ_{s2}	19.74	19.71	19.72
τ_n	2.78×10^{16}	2.80×10^{16}	2.79×10^{16}
PE (%)	0.31	0.24	0.27
RE (%)	41.43	41.88	41.66
N_{it}	40	20	10

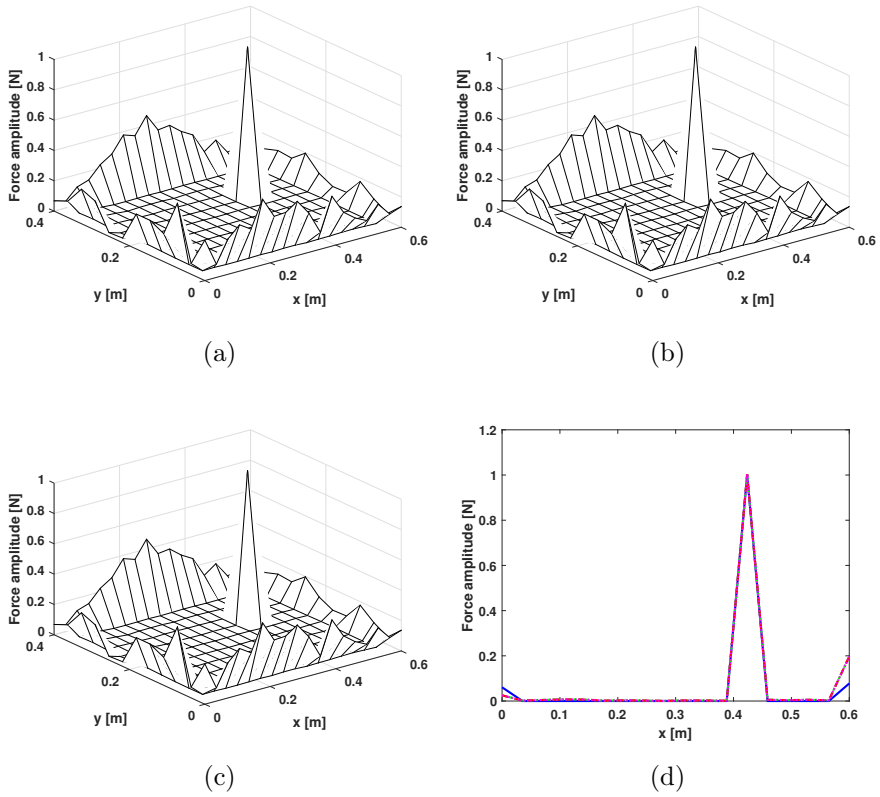


Figure 3: Reconstructed force vector $\widehat{\mathbf{F}}$ at 350 Hz from the OBR with respect to the initial choice of the shape parameter $\widehat{q}_0$. (a) Surface plot – $\widehat{q}_0 = 0.5$, (b) Surface plot – $\widehat{q}_0 = 1$, (c) Surface plot – $\widehat{q}_0 = 2$ and (d) Section view at $y_0 = 0.25$ m – (—) Reference, (– –) OBR for $\widehat{q}_0 = 0.5$, (– · –) OBR for $\widehat{q}_0 = 1$ and (· · ·) OBR for $\widehat{q}_0 = 2$

4.2.2. Influence of the choice of the convergence tolerance

As any iterative procedure, the identification accuracy of the OBR is partly related to the value of tolerance used to stop the iterative process. For this reason, it is important to study the convergence of the OBR with respect to the value of tolerance defined by the user. Table 2 gathers the results obtained with the OBR for tolerance values ranging from 10^{-3} to 10^{-8} and $\widehat{q}_0 = 2$. The analysis of the results show that the OBR converges quickly to the optimal solutions. Consequently, one can choose the value of the tolerance so as to define a compromise between the solution accuracy and the computational efficiency.

Table 2: Convergence study of the OBR with respect to the tolerance fixed to stop the iterative process – N_{it}: Number of iterations of the algorithm

	Tolerance					
	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
q_1	0.51	0.51	0.51	0.51	0.51	0.51
q_2	2	2	2	2	2	2
τ_{s1}	31.45	31.54	31.55	31.55	31.55	31.55
τ_{s2}	19.71	19.72	19.73	19.73	19.73	19.73
τ_n	2.79×10^{16}	2.79×10^{16}	2.79×10^{16}	2.79×10^{16}	2.79×10^{16}	2.79×10^{16}
PE (%)	0.28	0.27	0.28	0.28	0.28	0.28
RE (%)	41.69	41.66	41.66	41.66	41.66	41.66
N_{it}	8	10	13	15	17	20

4.2.3. Influence of the measurement noise level

In the previous sections, it has been supposed that the vibration field is corrupted by an additive Gaussian white noise having an SNR equal to 34 dB. However, it is legitimate to wonder whether the quality of the identified solutions remains the same when the SNR gets lower. To answer this question, the OBR is applied to vibration data having SNR ranging from 35 dB to 10 dB. For this numerical experiment, the initial shape parameter $\widehat{q}_0$ is set to 2.

The results listed in Table 3 indicates that the OBR performs well whatever the SNR values considered, which is an indicator of the robustness of the proposed regularization strategy.

Table 3: Convergence study of the OBR with respect to the measurement noise level corrupting the data – N_{it}: Number of iterations of the algorithm

	SNR					
	10	15	20	25	30	35
q_1	0.36	0.37	0.39	0.39	0.45	0.51
q_2	2	2	2	2	2	2
τ_{s1}	40.24	38.90	37.28	36.94	34.73	32.57
τ_{s2}	24.74	22.56	21.17	20.31	19.84	19.76
τ_n	6.16×10^{13}	2.03×10^{14}	6.78×10^{14}	2.27×10^{15}	7.83×10^{15}	2.77×10^{16}
PE (%)	-0.55	2.81	2.19	1.53	0.59	0.39
RE (%)	28.67	30.25	32.92	35.73	39.90	41.02
N_{it}	18	20	22	10	11	19

4.2.4. Comparison with Bayesian inference results

At this stage of the paper, it is interesting to compare the results obtained in the previous section with those obtained from the full Bayesian inference as described in Ref. [20].

After initially drawing 7000 samples, the values of the parameters corresponding to the mode of posterior distributions of each parameter of the formulation are given in Table 4 for SNR values ranging from 10 dB to 35 dB. Overall, the most probable values of the parameters inferred from the proposed MCMC procedure are in line with the OBR results. However, the analysis of the PE and RE indicators shows that, while the estimation of the point force amplitude is quite satisfying, the reconstruction of the reaction forces differs significantly when the SNR decreases. A possible explanation of these contrasted results is that a normal approximation has been used to draw samples from the full conditional probability distribution $p(\mathbf{F}|\mathbf{X}, \tau_n, \tau_{sr}, q_r)$[2] [see Ref. [20] for details]. Finally, it worth mentioning that the OBR is, as expected, computationally more efficient than the full Bayesian inference, since the solution is obtained in 1 s on average for the OBR, while about 6 min are required on average for the full Bayesian inference.

4.2.5. Influence of the choice of the prior distributions

The proposed OBR being based on a particular prior structure, it is interesting to study the influence of the choice of the prior distributions on the

[2]It should be mentioned that Metropolis-Hastings and HMC updates have also been implemented without success due to convergence issues (low acceptance rate).

Table 4: Most probable parameters inferred from MCMC with respect to the measurement noise level corrupting the data

	SNR					
	10	15	20	25	30	35
q_1	0.62	0.71	0.59	0.60	0.48	0.47
q_2	1.99	1.99	2	2	1.98	1.99
τ_{s1}	29.54	27.10	30.92	30.68	34.27	34.18
τ_{s2}	0.45	1.06	2.07	3.37	5.09	6.24
τ_n	5.81×10^{13}	1.88×10^{14}	5.89×10^{14}	1.89×10^{15}	5.26×10^{15}	1.70×10^{16}
PE (%)	-4.92	-1.57	-4.40	-2.62	-2.94	-2.20
RE (%)	99.90	95.35	71.85	66.89	47.79	41.63

estimation of the optimal solution.

Prior distributions of the shape and precision parameters

As a first modification, let us assume that the shape and precision parameters are uniformly distributed. Formally, this imposes that:

$$p(\tau_n) \propto \mathbb{I}_{\mathbb{R}^+},\ p(\tau_{sr}) \propto \mathbb{I}_{\mathbb{R}^+} \text{ and } p(q_r) \propto \mathbb{I}_{[l_b,u_b]}. \tag{32}$$

From a practical standpoint, the corresponding OBR is obtained by setting $\alpha_n = \alpha_{sr} = \alpha_r = 1$ and $\beta_n = \beta_{sr} = \beta_r = 0$ in Eqs. (18a) and (27). Given this particular formulation, obtained estimated parameters are listed in Table 5 for SNR values ranging from 10 dB to 35 dB. As expected, the results are identical to those presented in Table 3.

Table 5: Convergence study of the OBR with respect to the measurement noise level corrupting the data when the shape and precision parameters are uniformly distributed – N_{it}: Number of iterations of the algorithm

	SNR					
	10	15	20	25	30	35
q_1	0.36	0.37	0.39	0.39	0.45	0.51
q_2	2	2	2	2	2	2
τ_{s1}	40.24	38.90	37.28	36.94	34.73	32.57
τ_{s2}	24.74	22.56	21.17	20.31	19.84	19.76
τ_n	6.16×10^{13}	6.16×10^{14}	6.78×10^{14}	2.27×10^{15}	7.83×10^{15}	2.77×10^{16}
PE (%)	-0.55	2.81	2.19	1.53	0.59	0.39
RE (%)	28.67	30.25	32.92	35.73	39.90	41.02
N_{it}	18	20	22	10	11	19

Another possibility is to define a Jeffrey's prior for the precision parameters τ_n and τ_{sr}, while keeping the prior distribution of the shape parameters defined as a truncated Gamma distribution. It results that the prior distributions of the precision parameters are defined such that:

$$p(\tau_n) \propto \frac{1}{\tau_n} \text{ and } p(\tau_{sr}) \propto \frac{1}{\tau_{sr}}. \tag{33}$$

From a practical standpoint, the corresponding OBR is obtained by setting $\alpha_n = \alpha_{sr} = 0$ and $\beta_n = \beta_{sr} = 0$ in Eq. (27). In doing so, one obtains the results presented in Table 6. Contrary to the previous situation, the results only slightly differ from those listed in Table 3. Consequently, it can be concluded that regulation results are quite insensitive to the choice of the prior distributions of the shape and precision parameters provided that the latter are weakly informative or non-informative.

Prior distribution of the force vector

As for the shape and precision parameters, the choice of the prior distribution of the force vector to identify is not unique. The main advantage of the proposed prior distribution is to be flexible enough to deal with various distributions such as Gaussian and Laplace distributions by properly setting the value of the shape parameters. When the shape parameters are supposed to be deterministic, the corresponding Bayesian formulation is given by:

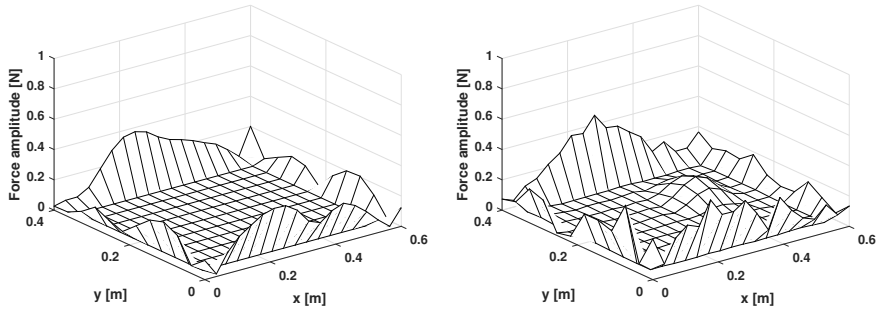
$$p(\mathbf{F}, \tau_n, \tau_{sr}|\mathbf{X}, q_r) \propto p(\mathbf{X}|\mathbf{F}, \tau_n)\, p(\tau_n|\alpha_n, \beta_n) \prod_{r=1}^{R} p(\mathbf{F_r}|\tau_{sr}, q_r)\, p(\tau_{sr}|\alpha_{sr}, \beta_{sr}). \tag{34}$$

A classical choice made in the literature is to suppose that the components of the force vector are normally distributed. In the present case, this implies

Table 6: Convergence study of the OBR with respect to the measurement noise level corrupting the data when the precision parameters follow a Jeffrey's prior – N_{it}: Number of iterations of the algorithm

	SNR					
	10	15	20	25	30	35
q_1	0.36	0.37	0.39	0.39	0.44	0.51
q_2	2	2	2	2	2	2
τ_{s1}	40.44	39.34	37.15	36.16	32.99	31.46
τ_{s2}	24.24	22.11	20.75	19.90	19.43	19.35
τ_n	6.13×10^{13}	2.02×10^{14}	6.75×10^{14}	2.26×10^{15}	7.82×10^{15}	2.77×10^{16}
PE (%)	-0.62	2.80	2.25	1.50	0.67	0.34
RE (%)	28.80	30.39	33.10	36.10	40.40	41.69
N_{it}	15	15	18	12	14	23

that $q_r = 2$ for all the considered regions. This formulation gives rise to an Augmented group Tikhonov regularization. As shown in Fig. 4 and Table 7, this particular formulation is unfortunately unable to properly identify the point force excitation, while the reconstruction of the reaction forces is quite satisfying.



(a) (b)

Figure 4: Reconstructed force vector $\widehat{\mathbf{F}}$ at 350 Hz when the components of the force vector are normally distributed - (a) SNR = 10 dB and (b) SNR = 35 dB

When the components of the force vector are supposed to follow a Laplace distribution (i.e. $q_r = 1$ for each region), the results are not much better as indicated by the analysis of Fig. 5 and Table 8.

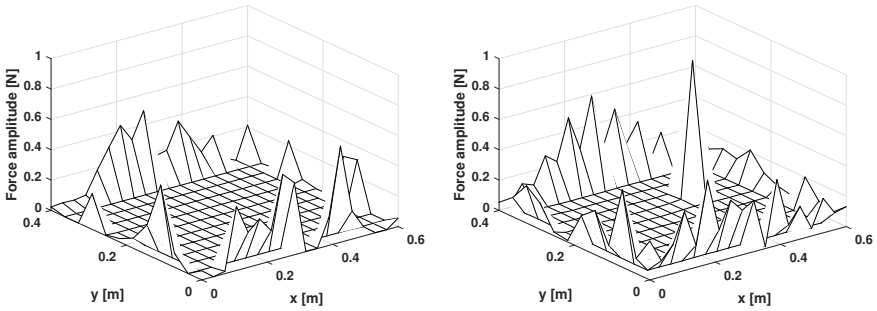
These disappointing results can be explained by the fact that the prior on the force vector is not well adapted to the reconstruction of both localized and distributed sources at the same time.

4.2.6. Parameters reduction

When dealing with parameter optimization, it could be interesting to reduce the number of parameters to constrain the space of admissible solutions

Table 7: Convergence study of the OBR with respect to the measurement noise level corrupting the data when components of the force vector is normally distributed – N_{it}: Number of iterations of the algorithm

	SNR					
	10	15	20	25	30	35
τ_{s1}	6.35×10^{58}	5.45×10^{36}	4.62×10^{35}	3.20×10^{3}	2.32×10^{3}	1.43×10^{3}
τ_{s2}	25.84	22.11	23.48	21.39	20.90	20.04
τ_{n}	4.32×10^{13}	9.05×10^{13}	1.39×10^{14}	1.74×10^{15}	5.74×10^{15}	2.09×10^{16}
PE (%)	-100	-100	-100	-93.76	-91.17	-85.97
RE (%)	36.54	34.11	31.12	53.24	54.94	57.71
N_{it}	9	10	15	13	13	17



(a) (b)

Figure 5: Reconstructed force vector $\widehat{\mathbf{F}}$ at 350 Hz when the components of the force vector follow a Laplace distribution - (a) SNR = 10 dB and (b) SNR = 35 dB

Table 8: Convergence study of the OBR with respect to the measurement noise level corrupting the data when the components of the force vector follow a Laplace distribution – N_{it}: Number of iterations of the algorithm

	SNR					
	10	15	20	25	30	35
τ_{s1}	2.25×10^{10}	5.43×10^{10}	149.32	135.76	127.79	114.05
τ_{s2}	6.27	5.89	5.27	5.02	4.95	4.98
τ_n	4.38×10^{13}	9.15×10^{13}	6.64×10^{14}	2.30×10^{15}	7.81×10^{15}	2.70×10^{16}
PE (%)	-100	-100	-29.17	-18.15	-12.78	-9.37
RE (%)	66.19	58.40	80.65	77.91	68.63	58.82
N_{it}	8	12	100	9	7	6

by eliminating some degrees of freedom. In the present work, two possibilities are studied. The first one consists in defining a unique precision parameter τ_s for all the identification region, which leads to define a single regularization parameter for computing the optimal force vector. From the theoretical standpoint, the corresponding Bayesian formulation is written:

$$\begin{aligned} p(\mathbf{F}, \tau_n, \tau_s, q_r|\mathbf{X}) \propto\, & p(\mathbf{X}|\mathbf{F}, \tau_n)\, p(\tau_n|\alpha_n, \beta_n)\, p(\tau_s|\alpha_s, \beta_s) \\ & \times \prod_{r=1}^{R} p(\mathbf{F_r}|\tau_s, q_r)\, p(q_r|\alpha_r, \beta_r, l_b, u_b). \end{aligned} \tag{35}$$

Regarding the resolution algorithm, this assumption leads to change the calculation of the precision parameter τ_s at step 2.c of the pseudo-code described in section 3.2, whose closed-form expression is given by Eq. (27). Under this

assumption, the precision parameter $\widehat{\tau}_s^{(k)}$ at iteration k is expressed as:

$$\widehat{\tau}_s^{(k)} = \frac{\alpha_s + \sum_{r=1}^{R} \frac{M_r}{\widehat{q}_r^{(k)}}}{\sum_{r=1}^{R} \left\| \widehat{\mathbf{F}}_{\mathbf{r}}^{(k-1)} \right\|_{\widehat{q}^{(k)}}^{\widehat{q}^{(k)}}}. \tag{36}$$

As shown in Table 9, the results obtained for a unique precision parameter τ_s are in a very good agreement with those listed in Table 3. Consequently, it is possible to employ only one precision parameter τ_s for all the considered regions.

Table 9: Convergence study of the OBR with respect to the measurement noise level corrupting the data for a unique precision parameter τ_s – N_{it}: Number of iterations of the algorithm

	SNR					
	10	15	20	25	30	35
q_1	0.34	0.35	0.36	0.34	0.41	0.51
q_2	2	2	2	2	2	2
τ_s	32.62	30.63	28.98	28.52	26.06	24.55
τ_n	6.10×10^{13}	2.02×10^{14}	6.73×10^{14}	2.25×10^{15}	7.88×10^{15}	2.89×10^{16}
PE (%)	1.39	3.39	2.43	1.72	0.94	-0.37
RE (%)	26.73	27.25	29.13	30.53	36.73	42.11
N_{it}	14	14	16	15	12	15

In the light of the previous observations, it is legitimate to wonder whether it is possible to limit the number of the shape parameters. To explore this

possibility, let us defining a unique shape parameter q for all the identification region. From the theoretical standpoint, the corresponding Bayesian formulation is written:

$$p(\mathbf{F}, \tau_n, \tau_{sr}, q|\mathbf{X}) \propto p(\mathbf{X}|\mathbf{F}, \tau_n)\, p(\tau_n|\alpha_n, \beta_n)\, p(q|\alpha, \beta, l_b, u_b) \times \prod_{r=1}^{R} p(\mathbf{F_r}|\tau_s, q)\, p(\tau_{sr}|\alpha_{sr}, \beta_{sr}). \tag{37}$$

Regarding the resolution algorithm, this assumption leads to change the calculation of the shape parameter q at step 2.b of the pseudo-code described in section 3.2. Under this assumption, the shape parameter $\widehat{q}^{(k)}$ at iteration k is solution of the following minimization problem:

$$\widehat{q} = \underset{q}{\operatorname{argmin}}\ f(q|\tau_{sr}, \mathbf{F_r}) \ \text{ for } q \in [l_b, u_b], \tag{38}$$

where $f(q|\tau_{sr}, \mathbf{F_r}) = \beta q - \left[\alpha - 1 + \sum_{r=1}^{R} M_r\right] \log q + \log \Gamma(1/q)(\sum_{r=1}^{R} M_r) - \frac{1}{q}\sum_{r=1}^{R} M_r \log \tau_{sr} + \sum_{r=1}^{R} \tau_{sr} \|\mathbf{F_r}\|_q^q$.

As highlighted in Table 10, the reduction of the number of the shape parameters leads to inconsistent solutions, since the solver converges to the sparsest solution, which is not a desirable effect. However, these results are consistent with the results presented in Tables 7 and 8 indicating that a reduction of the number of shape parameters is detrimental to the solution accuracy.

4.2.7. Recommendations

In the light of the previous results, we propose the following recommendations to properly apply the OBR:

Table 10: Convergence study of the OBR with respect to the measurement noise level corrupting the data for a unique shape parameter q – N_{it}: Number of iterations of the algorithm

	SNR					
	10	15	20	25	30	35
q	0.28	0.28	0.32	0.35	0.44	0.52
τ_{s1}	28.17	30.82	26.90	28.95	30.47	33.84
τ_{s2}	10.97	9.55	8.61	7.47	6.01	5.30
τ_n	5.96×10^{13}	1.97×10^{14}	6.67×10^{14}	2.25×10^{15}	7.87×10^{15}	2.72×10^{16}
PE (%)	4.42	6.45	3.74	2.01	0.83	0.27
RE (%)	122.40	103.05	104.60	86.48	90.32	81.71
N_{it}	55	31	100	31	26	30

1. The OBR being almost insensitive to the choice of the initial shape parameters from which all the other initial parameters are derived, it is recommended to set $\widehat{q}_1^{(0)} = \widehat{q}_2^{(0)} = 2$. This allows computing the initial parameters in an efficient manner;
2. The convergence study, presented in section 4.2.2, has demonstrated that the OBR converges quickly to the optimal solution. Consequently, it is recommended to set the tolerance of the iterative process to 10^{-3}, which leads to a reasonable compromise between the solution accuracy and the computational efficiency;
3. As shown in section 4.2.6, it is possible to limit the number of parameters of the optimization problem. Indeed, a parameter reduction allows constraining more the space of admissible solutions, which can

help for finding an optimal solution. However, our results suggests to reduce the number of the precision parameters τ_{sr} instead of the shape parameters q_r.

5. Experimental validation

This section aims at confirming the main conclusions drawn in the previous section by extending the analysis to a real-world application. Since the proposed method is a natural extension of the work presented in Refs. [14, 20], we have decided to perform this validation using the same experimental set-up.

5.1. Description of the experimental set-up

The structure under test is a steel parallelepiped box, excited on one of its faces by a shaker fed by a white noise signal and equipped with a force sensor [see Fig. 6]. The parameters of this experimental validation are given in Table 11.

Measurements of the vibration field were carried out with a scanning laser vibrometer on a grid of 19×22 points along y and z directions respectively using the force signal as phase reference. In all the subsequent identifications, the measured vibration velocity field is normalized to the force signal delivered by the force sensor. In doing so, the identified point force F_0 should be equal to 1.

Regarding the FE mesh used to model the dynamic behavior of the plate, it has been designed to perfectly match the measurement mesh. Hence, it consists of 378 shell elements, making the model theoretically valid up to

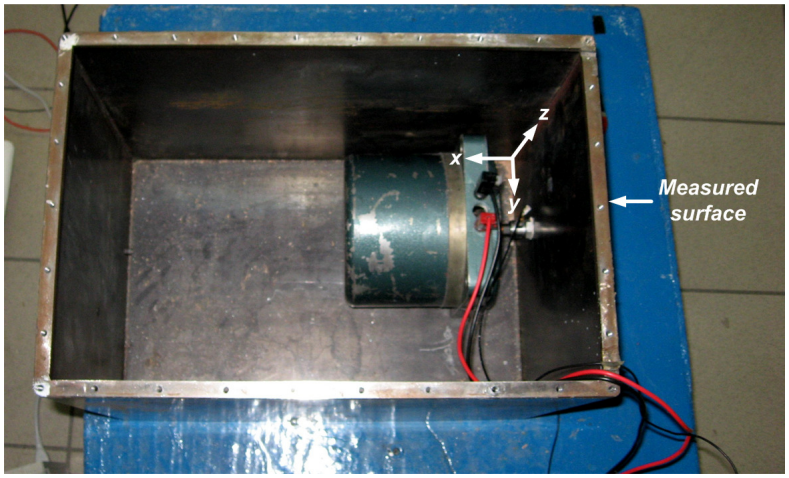


Figure 6: Experimental set-up

Table 11: Experiment parameters

Parameters	**Values**
Length of the parallelepiped	$L_x = 0.45$ m
Width of the parallelepiped	$L_y = 0.3$ m
Height of the parallelepiped	$L_z = 0.35$ m
Wall thickness	$h = 0.005$ m
Young's modulus	$E = 2.1 \times 10^{11}$ Pa
Density	$\rho = 7800$ kg.m^{-3}
Location of the force	$(y_0, z_0) = (0.10 \text{ m}, 0.09 \text{ m})$

5000 Hz. Then, the corresponding FE model with free boundary conditions has been used to compute the transfer functions matrix $\mathbf{H}$, considering the bending motions as the only available data.

5.2. Application

A careful analysis of the experimental set-up suggests the definition of two identification regions. The first region associated to the shape parameter q_1 contains the point force only, while the second region, associated to the shape parameter q_2, corresponds to the boundaries of the plate [see Fig. 7].

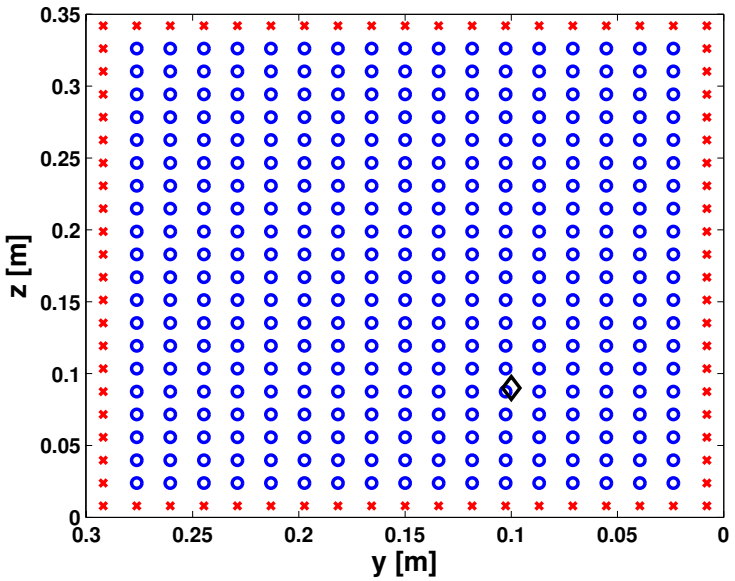


Figure 7: Definition of the identification regions - (○) region 1 (Point force), (×) region 2 (Reaction forces) and (◊) location of the point force

In the experimental application, the OBR is applied at 525 Hz, which is a frequency lying outside the resonance frequency of the plate for a conver-

gence tolerance equal to 10^{-3} and $\widehat{q}^{(0)} = 2$. The aim of this experimental validation is to compare the OBR implemented from the full Bayesian formulations given by Eq. (11) and (35) respectively. In other words, we are going to compare the OBR considering either two precision parameters τ_{sr} (case 1) or a single precision parameter τ_s (case 2). For the sake of completeness, these results are also compared with the MAP solution estimated after initially drawing 7000 samples from the full Bayesian inference introduced in Ref. [20].

Fig. 8 presents the excitation fields reconstructed from the OBR for the two considered parametrizations. This figure shows that the reconstructed excitation fields are qualitatively similar. More precisely, the location and the amplitude of the point force amplitude are recovered. This observation is confirmed by the analysis of the results listed in Table 12. Indeed, the estimates of the shape parameters are consistent with our expectations and the peak error indicates that the point force amplitude is properly estimated. Furthermore, it is interesting to note that, in the present experimental validation, the proposed OBR results are fully consistent with those estimated from the full Bayesian inference.

6. Conclusion

In the present paper, an optimal Bayesian regularization has been introduced for reconstructing mechanical sources acting on a structure. The purpose of the proposed approach is to determine the most probable excitation field as well as the most probable parameters associated to the full

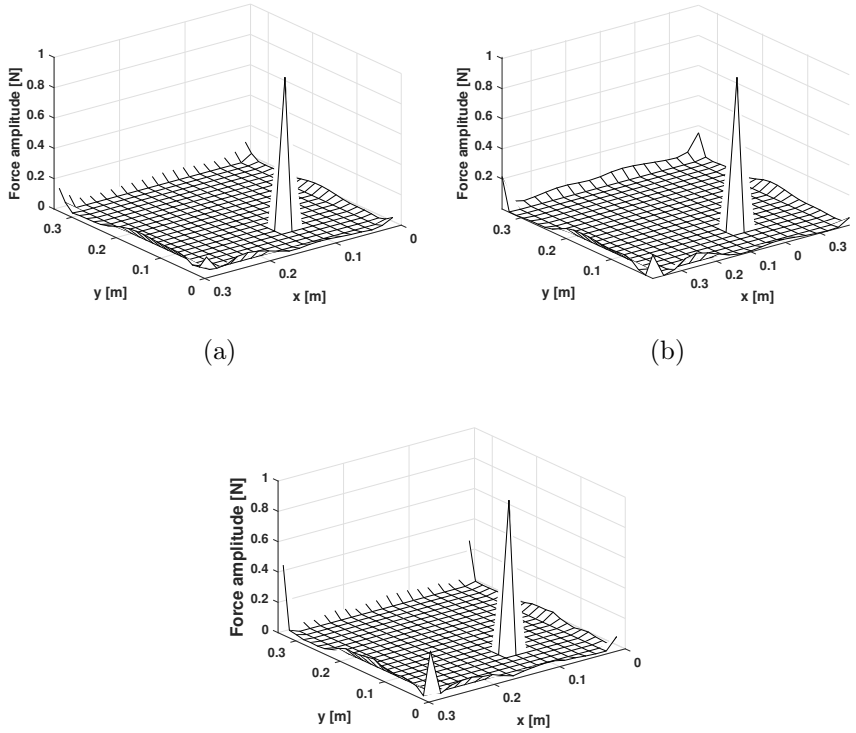


Figure 8: Reconstructed force vector $\widehat{\mathbf{F}}$ at 525 Hz (a) OBR for two precision parameters τ_{sr} (case 1), (b) OBR for a single precision parameter τ_s (case 2) and (c) MAP estimated from MCMC samples

Table 12: Comparison of the OBR when considering two precision parameters τ_{sr} (case 1) or a single precision parameter τ_s (case 2) with the MAP estimated from MCMC samples – N_{it}: Number of iterations of the algorithm

	case 1	**case 2**	**MCMC**
q_1	0.26	0.26	0.40
q_2	2	1.40	1.99
τ_{s1}	64.50	62.79	55.78
τ_{s2}	458.77	-	6.11
τ_n	7.99×10^8	8.25×10^8	9.01×10^8
PE (%)	-0.6	0.8	-0.1
N_{it}	30	70	-

Bayesian formulation previously published by the authors and recalled in this contribution. One of the most interesting features of the proposed strategy lies in its ability in computing optimal solutions from a minimal amount of prior information on the sources to identify. A set of recommendations has been proposed from the numerical and experimental validations we have conducted. It has also been shown that the proposed approach is robust against the initialization, the tolerance set to stop the iterative process and the measurement noise level. Finally, it must be said that the optimal Bayesian regularization is quite general, since it can be directly used to tackle identification problems in frequency or time domains, provided that the reconstruction model is established accordingly.

References

[1] A. N. Tikhonov. Regularization of incorrectly posed problems. *Soviet Mathematics*, 4:1624–1627, 1963.

[2] E. Jacquelin, A. Bennani, and P. Hamelin. Force reconstruction: analysis and regularization of a deconvolution. *Journal of Sound and Vibration*, 265:81–107, 2003.

[3] Q. Leclere, C. Pezerat, B. Laulagnet, and L. Polac. Indirect measurement of main bearing loads in an operating diesel engine. *Journal of Sound and Vibration*, 286 (1-2):341–361, 2005.

[4] Y. Liu and W. Steve Shepard Jr. Reducing the impact of measurement errors when reconstructing dynamic forces. *Journal of Vibration and Acoustics*, 128:586–593, 2006.

[5] A. N. Thite and D. J. Thompson. The quantification of structure-borne transmission paths by inverse methods. Part 2 : Use of regularization techniques. *Journal of Sound and Vibration*, 264 (2):433–451, 2003.

[6] Y.-M Mao, X.-L. Guo, and Y. Zhao. Experimental study of hammer impact identification on a steel cantilever beam. *Experimental Techniques*, 34 (3):82–85, 2010.

[7] E. Turco. Tools for the numerical solution of inverse problems in structural mechanics: review and research perspectives. *European Journal of Environmental and Civil Engineering*, 21 (5):509–554, 2017.

[8] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[9] D. Ginsberg and C. P. Fritzen. New approach for impact detection by finding sparse solution. In *Proceedings of ISMA 2014*, 2014.

[10] B. Qiao, X. Zhang, C. Wang, H. Zhang, and X. Chen. Sparse regularization for force identification using dictionaries. *Journal of Sound and Vibration*, 368:71–86, 2016.

[11] B. Qiao, X. Zhang, J. Gao, and X. Chen. Impact-force sparse reconstruction from highly incomplete and inaccurate measurements. *Journal of Sound and Vibration*, 376:72–94, 2016.

[12] C.-D. Pan, L. Yu, H.-L. Liu, Z.-P., and W.-F. Luo. Moving force identification based on redundant concatenated dictionary and weighted ℓ_1-norm regularization. *Mechanical Systems and Signal Processing*, 98:32–49, 2018.

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

[14] M. Aucejo and O. De Smet. Bayesian source identification using local priors. *Mechanical Systems and Signal Processing*, 66-67:120–136, 2016.

[15] P. C. Hansen. *Discrete Inverse Problems: Insight and Algorithms*. SIAM, 2010.

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

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

[18] C. Faure, F. Ablitzer, C. Pezerat, and J. Antoni. Vibration source characterization using force analysis technique and a Bayesian regularization. In *Proceedings of ICSV23, 23rd International Congress on Sound and Vibration*, Athens, Greece, 2016.

[19] M. Aucejo and O. De Smet. Bayesian formulations for force reconstruction problems. In *Proceedings of the 6th International Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, COMPDYN 2017, in conjunction with the 2nd International Conference on Uncertainty Quantification in Computational Sciences and Engineering, UNCECOMP 2017*, Rhodes Island, Greece, 2017.

[20] M. Aucejo and O. De Smet. On a full bayesian inference for force reconstruction problems. *Mechanical Systems and Signal Processing*, 104:36–59, 2018.

[21] Q. Li and Q. Liu. A hierarchical bayesian method for vibration-based time domain force reconstruction problems. *Journal of Sound and Vibration*, 421:190–204, 2018.

[22] J. T. Ormerod and M. P. Wand. Explaining variational approximations. *The American Statistician*, 64 (2):140–153, 2010.

[23] Y. W. Teh, D. Newman, and M. Welling. A collapsed variational bayesian inference algorithm for latent dirichlet allocation. In *Advances in Neural Information Processing Systems (NIPS) 19*, 2007.

[24] S. D. Babacan, R. Molina, and A. K. Katsaggelos. Variational bayesian blind deconvolution using a total variation prior. *IEEE Transactions on Image Processing*, 18 (1):12–26, 2009.

[25] K. Ishiguro, I. Sato, and N. Ueda. Averaged collapsed variational bayes inference. *Journal of Machine Learning Research*, 18:1–29, 2017.

[26] B. Jin and J. Zou. Augmented tikhonov regularization. *Inverse Problems*, 25:025001 – 25 pp, 2008.

[27] D. Feng, H. Sun, and M. Q. Feng. Simultaneous identification of bridge structural parameters and vehicle loads. *Computers & Structures*, 157:76–88, 2015.

[28] H. Sun and O. Büyüköztürk. Identification of traffic-induced nodal excitations of truss bridges through heterogeneous data fusion. *Smart Materials and Structures*, 24:075032, 2015.

[29] G. Yan, H. Sun, and O. Büyüköztürk. Impact load identification for composite structures using bayesian regularization and unscented kalman filter. *Structural Control and Health Monitoring*, 24 (5):e1910, 2017.

[30] B. Amizic, R. Molina, and A. K. Katsaggelos. Sparse bayesian blind image deconvolution with parameter estimation. *EURASIP Journal on Image and Video Processing*, page 20, 2012.

[31] Q. Li and Q. Lu. Time domain force identification based on adaptive ℓ_q regularization. *Journal of Vibration and Control*, First published March 19, 2018.

[32] S. Nadarajah. A generalized normal distribution. *Journal of Applied Statistics*, 32 (7):685–694, 2005.

[33] C. M. Bishop. *Pattern recognition and machine learning*. Springer, 2006.

[34] M. E. Tipping. The relevance vector machine. In *Advances in Neural Information Processing Systems 12*, pages 652–658, 2000.

[35] M. Aucejo and O. De Smet. A multiplicative regularization for force reconstruction. *Mechanical Systems and Signal Processing*, 85:730–745, 2017.

[36] A. Pereira, J. Antoni, and Q. Leclere. Empirical bayesian regularization of the inverse acoustic problem. *Applied Acoustics*, 97:11–29, 2015.

[37] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. SIAM, 1998.

[38] P. Rodriguez and B. Wohlberg. Efficient minimization method for a generalized total variation functional. *IEEE Transactions on Image Processing*, 18 (2):322–332, 2009.

[39] R. H. Byrd, P. Lu, and J. Nocedal. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, 16 (5):1190–1208, 1995.

[40] S. G. Nash. Newton-type minimization via the lanczos method. *SIAM Journal of Numerical Analysis*, 21:770–778, 1984.

[41] R. P. Brent. *Algorithms for Minimization Without Derivatives*. Prentice-Hall, Englewood Cliffs, N.J., 1973.

[42] P. D. Hoff. *A First Course in Bayesian Statistical Methods*. Springer-Verlag New York, 2009.

[43] A. Gelman. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1 (3):515–533, 2006.

[44] S.R. Ibrahim, A. Fregolent, and A. Sestieri. Structural force identification at unmeasured locations. In *Proceedings of the 14th International Modal Analysis Conference*, Dearborn, USA, 1996.

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